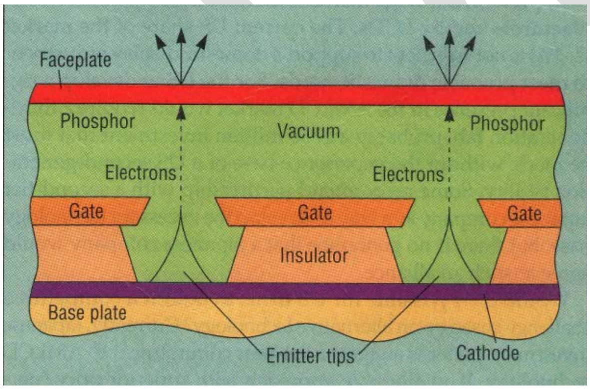
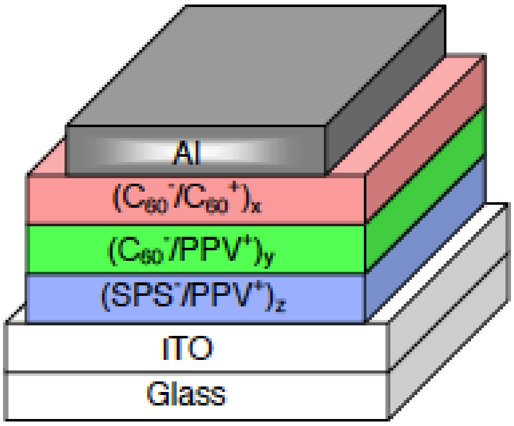
№530785-TEMPUS-1-2012-1-PL-TEMPUS-JPCR

Institute of High Technology, Taras Shevchenko National University of Kyiv

Advanced Nanosystems Design and Fabrication Techniques

Valeriy Skryshevsky, Anatoliy Evtukh, Volodymyr Ilchenko





Textbook

2016

Textbook " **Advanced Nanosystems Design and Fabrication Techniques**" developed to help higher education institutions in Ukraine to introduce new master's educational program "Designing microsystems".

Textbook " **Advanced Nanosystems Design and Fabrication Techniques**" was created with the support of the European Union within the Joint European Project "Curricula Development for New Specialization: Master of Engineering in Microsystems Design" (MastMST), identification number 530785-TEMPUS-1-2012-1-PL-TEMPUS-JPCR.

Project Coordinator prof. Zbigniew Lisik, Lodz University of Technology, Lodz, Poland.

TEMPUS teams:

- LvivPolitechnicalNationalUniversity, Lviv, Ukraine ,
  Coordinator prof. Mykhailo Lobur.
- Taras Shevchenko National University of Kyiv, Ukraine,
  Coordinator prof. ValeriySkryshevsky.
- Kharkiv National University of Radioelectronics, Ukraine ,
  Coordinator prof. Vladimir Hahanov.
- Donetsk National Technical University, Krasnoarmiysk,
  Coordinator prof. VolodymyrSviatny.
- Ilmenau University of Technology, Germany,
  Coordinator prof. Ivo Rangelow.
- Lyon Institute of Applied Sciences, France,
  Coordinator prof. Alexandra Apostoluk.
- University of Pavia, Italy,
  Coordinator prof. Paolo Di Barba.

The Textbook was approved by Editorial Committee (prof. Paolo Di Barba (University of Pavia) - Co-Chair, prof. Alexandra Apostoluk (LyonInstitute of Applied Sciences) – Co-Chair, members: prof. Zbigniew Lisik (Lodz University of Technology), Dr.Jacek Podgorski (Lodz University of Technology), Dr.Janusz Wozny (Lodz University of Technology), Dr.Valentyn Ishchuk (Ilmenau University of Technology), Dr.Maria Evelina Mognaschi (University of Pavia), Dr.Roberto Galdi (University of Pavia) May 6, 2016, Pavia, Italy.

The authors express their deep gratitude to the aforementioned universities for full support of the project.

Content

1. Introduction. History of NEMS and Devices Involving Nanotechnology
2. Nano metal-oxide based gas-sensor
3. Nano and porous silicon for gas, vapor and liquid sensing
4. Silver nanoparticle ink
5. Solar cells of 3-rd generation with nanosilicon. Quantum dots solar cells.
6. Organic, Polymeric and Dye-Sensitized Cells
7. NanoBetavoltaics
8. Environmentally Friendly Energy Sources on nanomaterials
9. Hydrogen storage and generation in nanosilicon
10. Magnetic quantum dots
11. Field emission from nanosilicon and cold cathodes
12. Carbon Nanotubes In 3-D Structured Electrodes
13. Field emission display based on carbon nanotubes
14. Non-volatile memory
15. Nanoparticle Organic Memory Field-Effect Transistor.
16. Flexible electronic circuits
17. Silicon nanophotonics
18. Nanotube and other implant materials
19. Transistors from carbon nanotube and integrated circuits built with nanotube transistors.

Laboratory works.

1. Impedance spectroscopy of nanosemiconductors
2. Manufacturing of porous Silicon based structures by electrochemical etching
3. Photoluminescence (PL) spectra and principles of PL decay time measurement

Chapter 1. Introduction to NEMS and Devices Involving Nanotechnology

The field of Nanotechnology, which aims at exploiting advances in the fabrication and controlled manipulation of nanoscale objects, is attracting worldwide attention. This attention is predicated upon the fact that obtaining early supremacy in this field of miniaturization may well be the key to dominating the world economy of the 21st century and beyond. NanoMEMS exploits the convergence between nanotechnology and microelectromechanical systems (MEMS) brought about by advances in the ability to fabricate nanometer-scale electronic and mechanical device structures. Indeed, the impact of our ability to make and control objects possessing dimensions down to atomic scales, perhaps first considered by the late Richard Feynman in his 1959 talk “There is Plenty of Room at the Bottom” is expected to be astounding [1], [2]. In particular, miniaturization, he insinuated, has the potential to fuel radical paradigm shifts encompassing virtually all areas of science and technology, thus giving rise to an unlimited amount of technical applications. Since high technology fuels the prosperity of the world’s most developed nations, it is easy to see why the stakes are so high.

Progress in the field of miniaturization benefited from the advent of the semiconductor industry in the 1960s, and its race to increase profits through the downscaling of circuit dimensions which, consequently, increased the density and the yield of circuits fabricated on a given wafer area. This density, which derived from progress in photolithographic tools to produce the ever smaller two-dimensional patterns (device layouts) of an integrated circuit (IC), has increased since by more than seven orders of magnitude and has come to be captured by Moore’s law: The number of components per chip doubles every 18 months [2]. The culmination of such miniaturization program, it is widely believed, is the demise of Moore’s law, whose manifestation is already becoming apparent due to an increasing predominance of the quantum mechanical nature of electrons in determining the behaviour of devices with critical dimensions (roughly) below 100 nm.

This line of development is closely related to the field of quantum devices/nanoelectronics, which was prompted by the conception of a number of atomic-level deposition and manipulation techniques, in particular, molecular beam epitaxy (MBE), originally exploited to construct laboratory devices in which the physics of electrons might be probed and explored, following the discovery of electron tunnelling in heavily-doped p-n-junctions [3]. Nanoelectronics did produce interesting physics, for instance, the discovery of Coulomb blockade phenomena in single-electron transistors, which manifested the particle nature of electrons, and resonant tunneling and conductance quantization in resonant tunnelling diodes and quantum point contacts, respectively, which manifested the wave nature of electrons [4-6]. These quantum devices, in conjunction with many others based on exploiting quantum phenomena, generated a lot excitement during the late 1980s and early 1990s, as they promised to be the genesis for a new digital electronics exhibiting the properties of ultra-high speed and ultra-low power consumption [7-8]. While efforts to realize these devices helped develop the skills for fabricating nanoscale devices, and efforts to analyze and model these devices helped to develop and mature the field of mesoscopic quantum transport, the sober reality that cryogenic temperatures would be necessary to enable their operation drastically restricted their commercial importance.

A few practical devices, however, did exert commercial impact, although none as much as that exerted by silicon IC technology, in particular, heterojunction bipolar transistors (HBTs), and high-electron mobility transistors (HEMTs), which exploit the conduction band discontinuities germane to heterostructures, and modulation doping to create 2-D electron confinement and quantization, respectively, and render devices superior to their silicon counterparts for GHz-frequency microwave and low-transistor count digital circuit applications [9-14].

The commercial success of the semiconductor industry, and its downscaling program, motivated emulation efforts in other disciplines, in particular, those of optics, fluidics and mechanics, where it was soon realized that, since ICs were fundamentally two-dimensional entities, techniques had to be developed to shape the third dimension, necessary to create mechanical devices exhibiting motion and produced in a batch planar process [15]. These techniques, which included surface micromachining, bulk micromachining, and wafer bonding, became the source of what are now mature devices, such as accelerometers, used in automobile air bags and pressure sensors, on the one hand, and a number of emerging devices, such as, gyroscopes, flow sensors, micromotors, switches, and resonators, on the other. Coinciding, as they do, with the dimensional features germane to ICs, i.e., microns, these mechanical devices whose behavior was controlled by electrical means, exemplified what has come to be known as the field of microelectromechanical systems (MEMS).

Three events might be construed as conspiring to unite nanoelectronics and MEMS, namely, the invention of a number of scanning probe microscopies, in particular, scanning tunneling microscopy (STM) and atomic force microscopy (AFM), the discovery of carbon nanotubes (CNTs), and the application of MEMS technology to enable superior RF/Microwave systems (RF MEMS) [16-18]. STM and AFM, by enabling our ability to manipulate and measure individual atoms, became crucial agents in the imaging of CNTs and other 3-D nanoscale objects so we could both “see” what is built and utilize manipulation as a construction technique. CNTs, conceptually, two-dimensional graphite sheets rolled-up into cylinders, are quintessential nanoelectromechanical (NEMS) devices, as their close to 1-nm diameter makes them intrinsically quantum mechanical 1-D electronic systems while, at the same time, exhibiting superb mechanical properties. MEMS, on the other hand, due to their internal mechanical structure, display motional behavior that may invade the domain of the Casimir effect, a quantum electrodynamical phenomenon elicited by a local change in the distribution of the modes in the zero-point fluctuations of the vacuum field permeating space [19-21]. This effect which, in its most fundamental manifestation, appears as an attractive force between neutral metallic surfaces, may both pose a limit on the packing density of NEMS devices, as well as on the performance of RF MEMS devices [22].

Quantum Wires to Nanophotonics. There are another interesting application of nanomaterials. It was developed soluble, III-V and II-VI semiconductor nanowires with controlled diameters in the strong-confinement regime of about 3–20 nm by bismuth-nanoparticle-catalyzed Solution-Liquid-Solid (SLS) growth (Figure 1.1) [24,25]. The nanoparticle-catalyzed "Vapor-Liquid-Solid" (VLS) mechanism and its solution-phase variants have emerged as the popular, widely practiced, general methods for the synthesis of semiconductor nanowires[26]. Gold nanoparticles are presently by far the most commonly employed catalysts.

The advantages of nanoparticle-catalyzed nanowire growth include its general applicability to a wide variety of materials, the diameter control afforded, the uniformity of the wires (lack of significant diameter fluctuations) and their oriented, near-single crystallinity. The surface passivation, solubility and length of the wires may also be systematically varied. Small-diameter “quantum” wires are ideal specimens for fundamental studies of two-dimensional (2D) quantum-confinement phenomena and for property comparisons to 3D-confined quantum dots, 1D-confined quantum wells and anisotropically 3D-confined quantum rods. Potential applications of semiconductor nanowires in nanophotonics and lasing, nanoelectronics, solar-energy conversion and chemical detection are under active development. Exciting progress and advances in the semiconductor-nanowire field, enabled by the emergence of nanoparticle-catalyzed growth, are anticipated in the immediate future.

Organic photovoltaics. The ability to create high-efficiency solar cells is a key strategy to meeting growing world energy needs. Nanotechnology is currently enabling the production of high-efficiency organic photovoltaics (OPVs) to help meet this challenge [27]. Organic photovoltaics are nanostructured thin films composed of layers of semiconducting organic materials (polymers or oligomers) [28] that absorb photons from the solar spectrum. These devices will revolutionize solar energy harvesting, because they can be manufactured via solution-based methods, such as ink-jet or screen printing, enabling rapid mass-production and driving down cost. OPVs currently lag behind their “inorganic” counterparts because of low solar energy conversion efficiencies (approximately 4-5%).

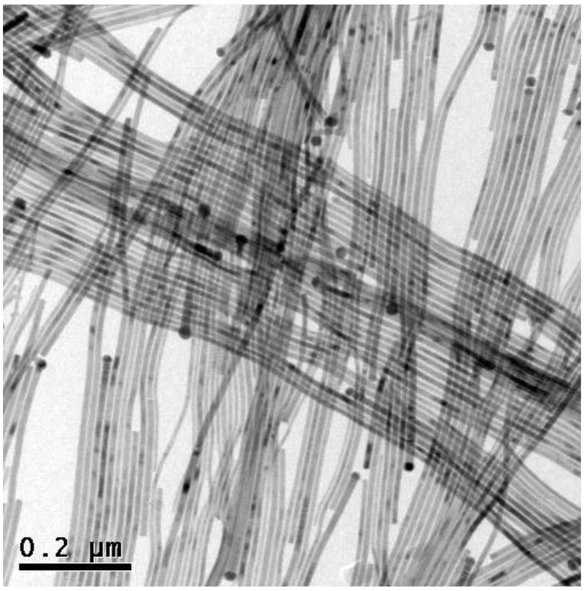


Figure 1.1. CdTe nanowires grown by the SLS mechanism.The bismuth-nanoparticle catalysts are evident at the nanowire tips.

Several research groups are addressing conversion efficiency by employing a combination of nanomaterials and unique nanoscale architectures. These hybrid organic-inorganic photovoltaics consist of light-absorbing polymers in contact with semiconductor nanocrystals, fullerenes or nanostructured metals. The nanomaterials affect electro-optical properties of the conducting polymer, which include assisting in absorption of red and near-IR photons, a significant portion of the solar spectrum.

Examples of OPVs designs employing nanomaterials include [29-32] Polymer-Fullerene Heterojunctions, Organic-Nanocrytsal Solar Cells, Dye-Sensitized Cells, Tandem Cells (Figure 1.2). Significant challenges exist to achieving OPV devices that can be mass-produced. Nanotechnology will assist in meeting the technical challenges of this rapidly evolving field.

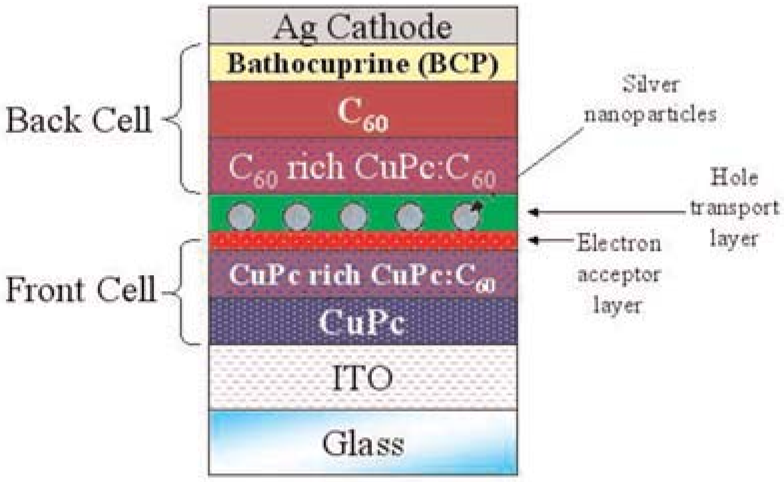


Figure 1.2. Schematic of tandem organic photovoltaic cells [33]

Coatings. Nanopowders and nanoparticle dispersions have seen increasing applications in coatings. Due to their small size, very even coating can be achieved by painting nanoparticle dispersions onto a surface and baking off the residual solvent. Optically Transparent Conductive Coatings**:** Indium tin oxide (ITO) and antimony tin oxide (ATO) are well known, optically transparent, electrically conductive materials. Nanoparticles of these materials can be painted on surfaces such as interactive touch screens to create a conductive, transparent screen without relying on expensive sputtering techniques. In addition, ITO and ATO can be used as an antistatic coating, utilizing their inherent conductivity to dissipate static charge.

Optically Transparent Abrasion-Resistant Coatings: Nanoscale aluminum oxide and titanium oxide are optically transparent and greatly increase the abrasion resistance of traditional coatings. Titanium oxide is of particular interest in many optical applications, since it is highly reflective for most ultra-violet radiation. Zinc oxide and rare-earth oxides are also UV-reflective, but optically transparent and are therefore effective in protecting surfaces from degradation brought about by exposure to UV radiation.

Bio-conjugates. It was developed near-infrared optical biosensors based on single-walled carbon nanotubes, which modulate their fluorescence emission in response to specific biomolecules. The viability of sensor techniques was demostrated by creating a singlewalled carbon nanotube (SWNT) enzyme bio-conjugate that detects glucose concentrations [34].

Carbon nanotubes fluoresce in a region of the near-infrared, where human tissue and biological fluids are particularly transparent to their emission. The sensor could be implanted into tissue, excited with a nearinfrared light source, and provide real-time, continuous sensing of blood glucose level by fluorescence response. Figure 1.3 shows a schematic mechanism of the nanotube sensor. Hydrogen peroxide is produced when glucose reacts with the enzyme, which quickly transforms ferricyanide to modulate near-infrared fluorescence characteristics of the nanotube.

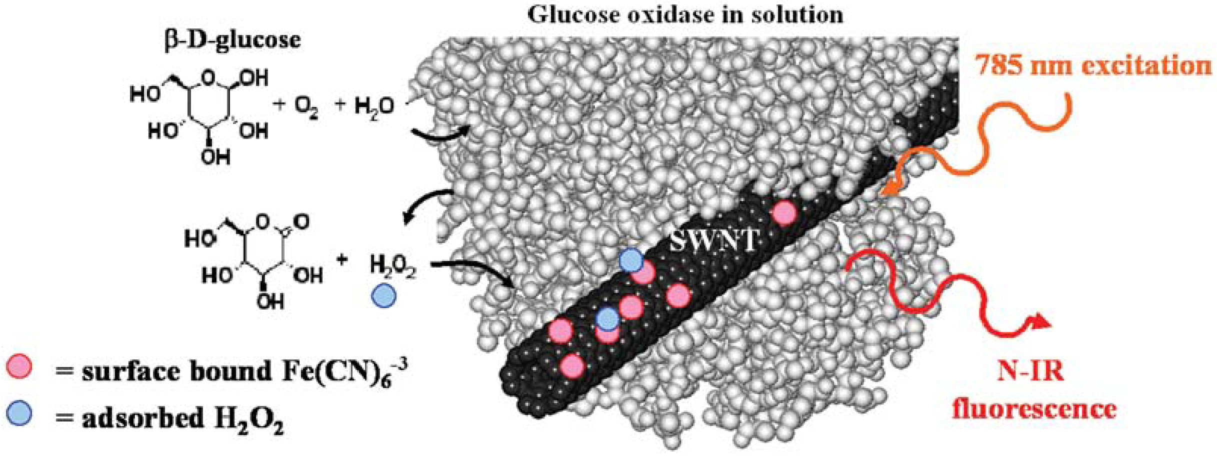


Figure 1.3. Near-IR radiation excites glucose to produce hydrogen peroxide (H2O2). The reaction with surface-bound ferricyanide on the nanotubes modulates the fluorescence characteristics of the nanotube.

The important aspect of this technology is that the technique can be extended to many other chemical systems. New types of non-covalent functionalization are developed, creating opportunities for nanoparticlesensors that operate in strong absorbing media of relevance to medicine or biology.

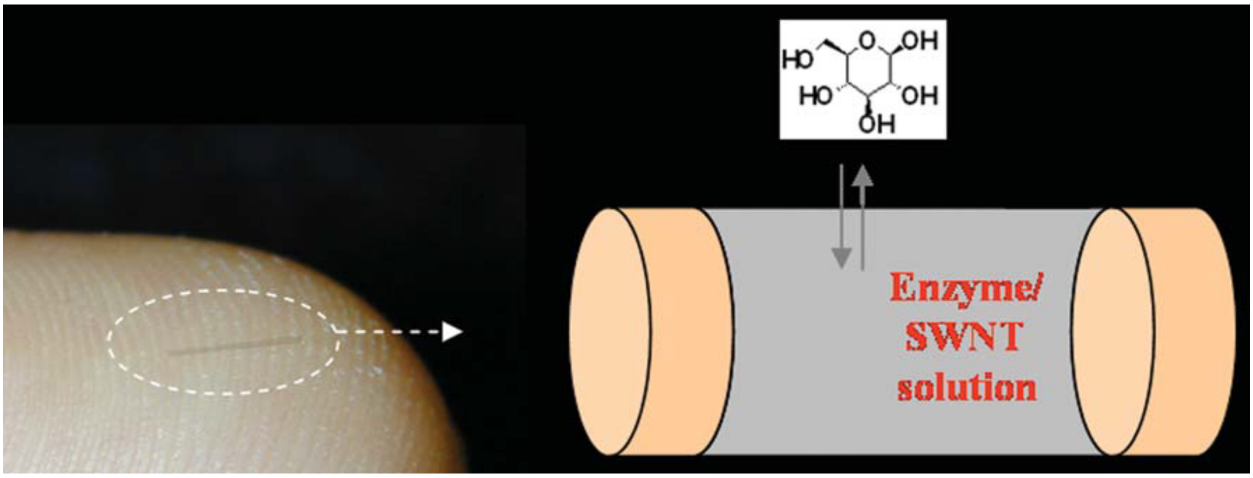


Figure 1.4. A size comparison of a SWNT-based glucose detector on a fingertip.

Over the last two decades, general interest and research in fuel cells has increased, because they have the potential to be more energy efficient than conventional power generation methods [35]. During this time period, researchers have begun using nanomaterials in the catalyst layer of fuel cell electrodes for a variety of reasons, including: increasing the active surface area of the anode and cathode catalyst, increasing the catalytic rate of oxidation or reduction and minimizing the weight of platinum and other precious metals in the fuel cell. The current generated at an electrode is proportional to the active surface of catalyst on the electrode surface, so higher power density fuel cells can be formed from nanomaterials, because nanomaterials have a higher surface area to volume ratio. Researchers have also shown that the electrocatalytic properties of the materials are sensitive to particle size, so increased catalytic activity can be observed for nanoparticles and nanomaterials [36-38]. However, the most important goal has been to decrease the weight of platinum and other precious metals in the catalyst layer of the fuel cell, so that the fuel cell can be costeffective. This has been the main limitation to the widespread use of fuel cells [35]. Researchers have employed carbon nanomaterials as supports for dispersions of platinum nanomaterials. This allows for a decrease in the weight of platinum needed to produce the same surface area of active platinum catalyst. The nanomaterials could be carbon foams containing nanopores [39] different types of nanotubes [40,41] or even singlewalled nanohorns [42]. All of these materials can act as a support and a conductor for platinum nanomaterials, making strides toward cost-effective fuel cell catalysts.

Environmentally Friendly Energy Sources**.** Devising schemes to meet the world's growing energy demands while simultaneously reducing green house emissions and other pollutants, has become one of the major challenges facing materials scientists. Nanomaterials promise to help solve many of the problems associated with new and emerging energy technologies.

Fuel Cells**:** Solid oxide fuel cells (SOFCs) offer the advantage over other fuel cell designs in that they do not require expensive, precious metal catalysts and can operate effectively without extensive purification of fuel sources. The activity of doped rare-earth oxide electrodes such as yttrium stabilized zirconia (YSZ) is directly related to their surface areas. Nanoparticles exhibit the high surface required for developing SOFC technologies.

Cleaner Emissions: Catalytic converters on vehicles around the world have significantly reduced the amount of automotive pollution over the last three decades. These devices require large amounts of expensive metals such as platinum, palladium and rhodium. Doped rare-earth metal oxides offer the promise of increased catalytic activity without the heavy reliance on precious metals. In addition, the increased efficiency of the next-generation catalytic converters will result in cleaner emission of existing internal combustion and diesel engines.

Nanotube-Polymer Composites for Ultra Strong Materials. Exceptional mechanical properties of single-walled carbon nanotubes (SWNT) have prompted intensive studies of SWNT-polymer composites. However, the composites made with nanotubes are still holding a substantial reserve of improvement of mechanical properties. The problem is that pristine SWNTs have very poor solubility in polymers, which leads to phase segregation of composites. Severe structural inhomogeneities result in the premature failure of the hybrid SWNT-polymer materials. The connectivity with the polymer matrix and uniform distribution within the matrix are essential structural requirements for strong

SWNT composites. This problem is being solved by several approaches. First, by using coatings from surfactants and polymers, such as sodium dodecyl sulfate or poly(styrenesulfonate). This enables formation of better dispersions in traditional solvents including water. Polymeric dispersion agents are strongly preferred for the composite preparation because of (a) tighter bonding with the graphene surface, (b) miscibility with polymer matrixes of composites and (c) substantially smaller concentration necessary for the preparation of SWNT dispersions. Among polymers, different poly(vinyl alcohols) work best as a host matrix for SWNTs, providing the composites with high tensile strength and excellent Young's modulus.

References

[1] R. P. Feynman, “There’s plenty of room at the bottom,” *J. Microelectromech. Syst.*, vol. 2, Mar. 1992, pp. 60-66.

[2] G.E. Moore, “Cramming more components onto integrated circuits,” *Electronics*, vol. 38, Number 8, April 19, 1965.

[3] L. Esaki, “New phenomenon in narrow Germanium pn-junctions,” *Phys. Rev.*, vol. 109, 1957, pp. 603-604.

[4] T. A. Fulton and G. J. Dolan, “Observation of single-electron charging effects in small tunnel junctions,” *Phys. Rev. Lett.*, vol. 59, 1987, pp.109–112.

[5] L.L. Chang, L. Esaki, and R. Tsu, “Resonant tunneling in semiconductor doublebarriers,” *Appl. Phys. Lett.*, vol. 24, 1974, p. 593.

[6] B J van Wees *et al.* “Quantised conductance of point contacts in a two dimensional electron gas,” *Phys. Rev. Lett.* , vol. 60, 1988, pp. 848-850.

[7] *Proc. Int. Symp. On Nanostructure Physics and Fabrication*, College Station, TX, March 13-15, M.A. Reed and W.P. Kirk, Editors., Academic Press, San Diego (1989).

[8] *Proc. Int. Symp. On Nanostructure Physics and Fabrication*, Santa Fe, NM, May 20-24, 1991, W.P. Kirk and W.P. Kirk, Editors., Academic Press, San Diego (1992).

[9] H. Kroemer, Theory of a wide-gap emitter for transistors,” Proc. IRE, vol. 45, 1957, pp. 1535-1537.

[10] W.P. Dumke, J.M. Woodall and. V.L. Rideout, “GaAs-GaAlAs heterojunction transistor for high frequency operation,” *Solid State Electron.*, vol. 15, 1972, pp. 1339-1344.

[11] P. M. Asbeck, D.L. Miller, W.C. Petersen, *et al*, “GaAs/GaAlAs heterojunction bipolar transistors with cutoff frequencies above 10 GHz,” *IEEE Electron Device Letts.*, vol. EDL-3, 1982, pp. 366-368.

[12] S.L. Su, et al., “Double heterojunction AlGaAs/GaAs bipolar transistors by MBE with a current gain of 1650,” *IEEE Electron Device Letts.*, vol. EDL-4, 1983, pp. 130-132.

[13] R.J. Malik, J.R. Hayes, F. Capasso et al., “High-gain AlInAs/GaInAs transistors grown by molecular beam epitaxy,” *IEEE Electron Device Letts.*, vol. EDL-4, 1982, pp. 366-368.

[14] T. Miura, S. Hiyamizuk, T. Fujii *et al.*, “A new field effect transistor woth selectively doped GaAs/n-AlGaAs heterostructures,” *Jpn. J. Appl. Phys*, vol. 19, 1980, pp. L255- L277.

[15] A. B. Frazier, A. B., R.O. Warrington, *et al.*, “The Miniaturization Technologies: Past, Present, and Future,” *IEEE Trans. Ind. Electronics*, vol. 42, 1995, p.423.

[16] G. Binnig and H. Rohrer, “Scanning tunneling microscopy,” Helvetica Physica Acta 55, 1982, pp. 726–735.

[17] P. J. F. Harris, *Carbon Nanotubes and Related Structures*. Cambridge Univ. Press,

Cambridge, U.K. (1999).

[18] H.J. De Los Santos, *Introduction to Microelectromechanical (MEM) Microwave Systems*, Norwood, MA: Artech House (1999).

[19] H. B. G. Casimir, "On the attraction between two perfectly conducting plates," *Proc. K. Ned.. Akad. Wet.*, vol. 51, 1948, pp. 793–799.

[20] P.W. Milonni, *The Quantum Vacuum: An Introduction to Quantum Electrodynamics*, Academic Press: San Diego (1994).

[21] C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw-Hill International Editions: New York, N.Y. (1985).

[22] H.J. De Los Santos, "Nanoelectromechanical Quantum Circuits and Systems," *Proc. IEEE*, vol. 91, No. 11, November 2003, pp. 1907-1921.

[23] Héctor J. De Los Santos Principles and Applications of NanoMEMS Physics, Springer, P.O. Box 17, 3300 AA Dordrecht, The Netherlands. 2005

[24] Yu, H. et al. *Nature Mater*. **2003**, *2*, 517.

[25] Yu, H. et al. *J. Am. Chem.Soc*. **2003**, *125*, 16168.

[26] Xia, Y. et al. *Adv. Mater*. **2003**, *15*, 353.

[27] Smalley, R.E. "Nanotechnology, Energy and People" http://www.americanenergyindependence.com/energychallenge.html, accessed Feb 23, 2005.

[28]To review a complete list of semiconducting polymers and oligomers, request your free copy of ChemFiles Vol. 4 No. 6: Organic Semiconductors for Advanced Electronics.

[29] McGhee, M.D.; Coakley, K.M. *Chem. Mater.* **2004**, *16*, 4533.

[30] Shaheen, S.E. et al. *Mater. Res. Bull.* **2005**, *30*, 10.

[31] Hoppe, H.; Sariciftci, N.S. *J. Mater. Res*, **2004**, *19*, 1924.

[32] Nierengarten, J.-F. *New. J. Chem.* **2004**, *28*, 1177.

[33] Xue, J. et al. *Appl. Phys. Lett.* **2005**, *86*, 5757.

[34] Strano, M.S. et al. *Nat. Mater*. **2005**, *4*, 86.

[35] Kordesch, K.; Simader. G. *Fuel Cells and Their Application* (VCH,Weinheim, **1996**).

[36] Brayner, R. et al. *J. Mol. Catal. A: Chem.* **2002**, *182*, 227.

[37] Liu, C. et al. *J. Mol. Catal. A: Chem.* **2000**, *157*, 253.

[38] Drelinkiewicz, A. *J. Catal.* **1999**, *186*, 123.

[39] Joo, S.H. etal. *Nature* **2001**, *412*, 169.

[40] Yuan, F. et al. *J. New Mater. Electrochem. Syst.* **2004**, *7*, 311.

[41] Planeix, J.M. et al. *J. Am. Chem. Soc.* **1994**, *116*, 7935.

[42] Zhu, J. et al. *Nano Lett.* **2003**, *3*, 1033.

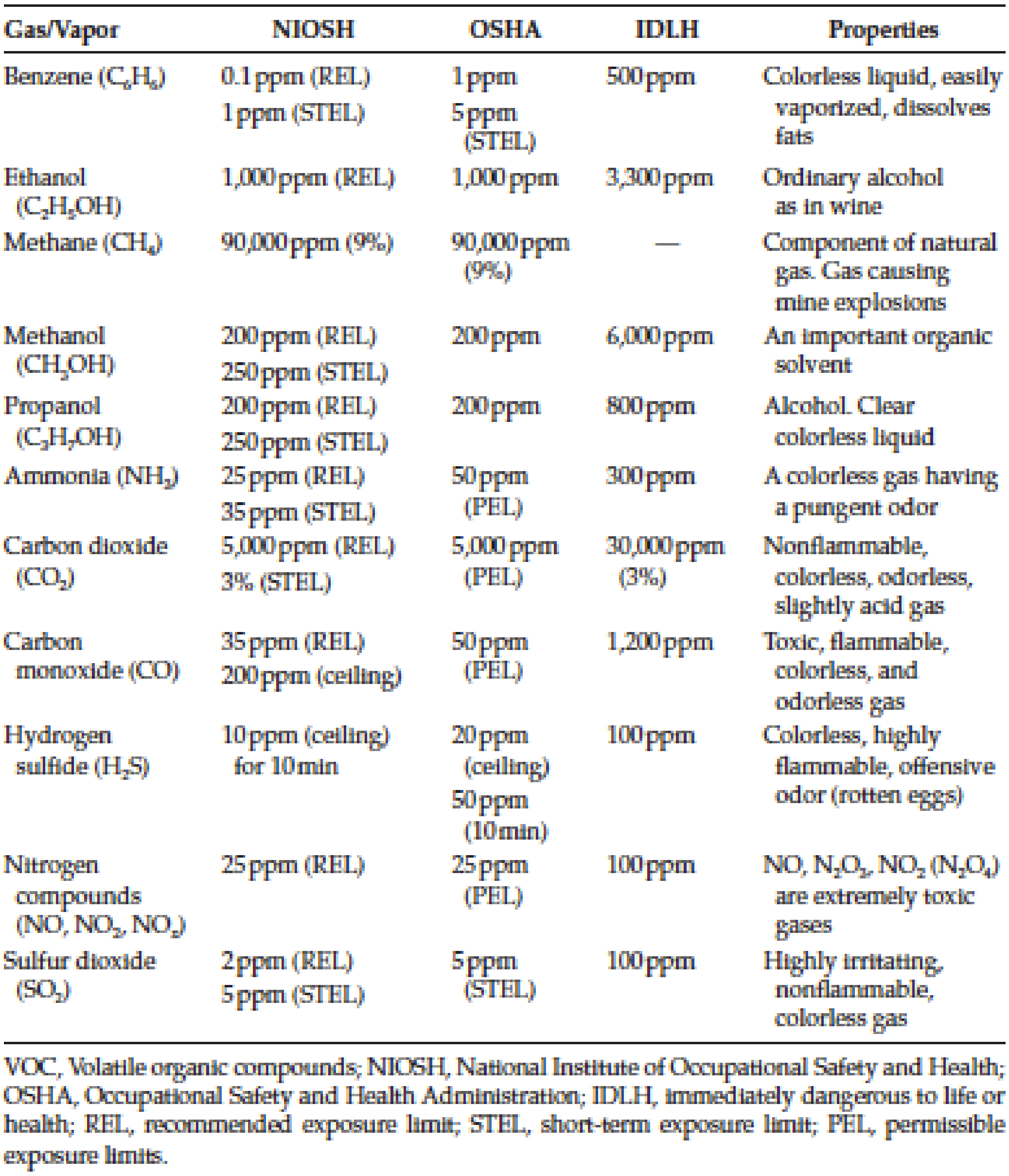
Chapter 2. Nano metal-oxide based gas-sensor

2.1 Introduction

Gas sensors have a great influence in many important areas namely environmental monitoring, domestic safety, public security, automotive applications, air-conditioning in aeroplanes, spacecrafts, etc. Air pollution has been one of Europe's main political concerns since the late 1970s. In 1996, the European Council adopted a Framework Directive on ambient air-quality assessment and management (Directive 96/62/EC of September 27, 1996) based on common methods and criteria in the European Union (EU) Member States. A series of "directives" have been issued to control the levels of identified gas pollutants and also to monitor their concentrations in the ambient air under observation [1]. Table 2.1 briefs the environmental safety standards for some selective volatile organic compound (VOC) vapors and pollutant gases. National Institute of Occupational Safety and Health (NIOSH) and Occupational Safety and Health Administration (OSHA) directives are summarized for benzene (C_6H_6), ethanol (C_2H_5OH), methane (CH_4), methanol (CH_3OH), and propanol (C_3H_7OH) vapors. Similarly, pollutant gases, such as ammonia (NH_3), carbon dioxide (CO_2), carbon monoxide (CO), hydrogen sulfide (H_2S), nitrogen compounds (NO, NO_2), and sulfur dioxide (SO_2), are discussed for their limits of exposure. Recommended exposure limit (REL), short-term exposure limit (STEL), and permissible exposure limit (PEL) parameters are outlined.

Environmental concerns and health and safety regulations have necessitated an increased use of sensors. Wherever sensor technology has been applied, it has proved to be very beneficial in improving energy efficiency and service, product quality, emission reduction, and in general improved the quality of life. There are many polluting species that have drawn attention. Among these gaseous species, the important species are oxides of nitrogen (NO, NO_2); oxides of carbon (CO, CO_2); hydrogen sulfide (H_2S); sulfur dioxide (SO_2); ozone (O_3); ammonia (NH_3); and many VOCs such as methane (CH_4), propane (C_3H_8), liquid petroleum gas (LPG), acetone (CH_3COCH_3), benzene (C_6H_6), and toluene ($C_6H_5CH_3$). In automotive industry, the demand for more fuel efficiency with lesser emission levels has mandated the need for highly efficient sensors in their exhaust systems [3].

Table 2.1



Gas/Vapor	NIOSH	OSHA	IDLH	Properties
Benzene (C_6H_6)	0.1 ppm (REL) 1 ppm (STEL)	1 ppm 5 ppm (STEL)	500 ppm	Colorless liquid, easily vaporized, dissolves fats
Ethanol (C_2H_5OH)	1,000 ppm (REL)	1,000 ppm	3,300 ppm	Ordinary alcohol as in wine
Methane (CH_4)	90,000 ppm (9%)	90,000 ppm (9%)	—	Component of natural gas. Gas causing mine explosions
Methanol (CH_3OH)	200 ppm (REL) 250 ppm (STEL)	200 ppm	6,000 ppm	An important organic solvent
Propanol (C_3H_7OH)	200 ppm (REL) 250 ppm (STEL)	200 ppm	800 ppm	Alcohol. Clear colorless liquid
Ammonia (NH_3)	25 ppm (REL) 35 ppm (STEL)	50 ppm (PEL)	300 ppm	A colorless gas having a pungent odor
Carbon dioxide (CO_2)	5,000 ppm (REL) 3% (STEL)	5,000 ppm (PEL)	30,000 ppm (3%)	Nonflammable, colorless, odorless, slightly acid gas
Carbon monoxide (CO)	35 ppm (REL) 200 ppm (ceiling)	50 ppm (PEL)	1,200 ppm	Toxic, flammable, colorless, and odorless gas
Hydrogen sulfide (H_2S)	10 ppm (ceiling) for 10 min	20 ppm (ceiling) 50 ppm (10 min)	100 ppm	Colorless, highly flammable, offensive odor (rotten eggs)
Nitrogen compounds (NO, NO_2, NO_x)	25 ppm (REL)	25 ppm (PEL)	100 ppm	NO, N_2O_3, NO_2 (N_2O_4) are extremely toxic gases
Sulfur dioxide (SO_2)	2 ppm (REL) 5 ppm (STEL)	5 ppm (STEL)	100 ppm	Highly irritating, nonflammable, colorless gas

VOC, Volatile organic compounds; NIOSH, National Institute of Occupational Safety and Health; OSHA, Occupational Safety and Health Administration; IDLH, immediately dangerous to life or health; REL, recommended exposure limit; STEL, short-term exposure limit; PEL, permissible exposure limits.

Sensing behavior is the most important and well-known property of metaloxide materials. In addition to the sensitivity to light (photon) energy and external pressure, metal oxides demonstrate high sensitivity to their chemical environment. With the capacity to operate in harsh environment, they surpass other chemical sensors in their sensitivity, reliability, and durability. A sensor detects a change in the gas atmosphere due to a change in the electrical properties of sensing elements. It was found that a catalytic reaction and a gas-sensing process on a metal oxide are analogous to each other as both processes involve surface adsorption and chemical reaction with the surrounding gas environment. Metal oxides are being investigated for gas-detection applications (fig.2.1).

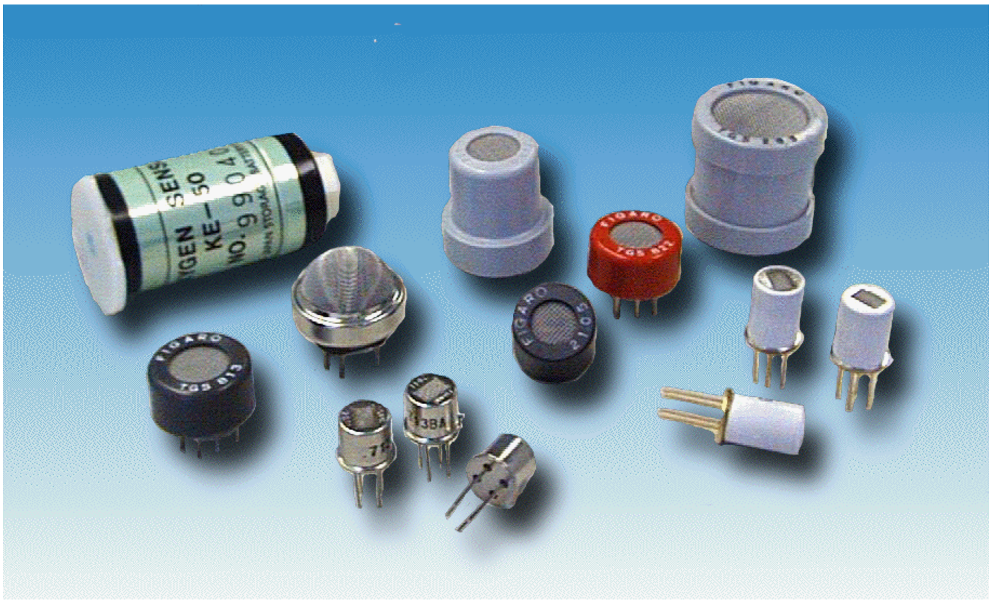


Fig.2.1 Examples of industrial metal-oxide sensors

This is because of their working mechanism being the variation of surface electrical conductance in the presence of a gaseous environment. This induced electrical conductivity is altered by the adsorption of gaseous species from the ambient atmosphere. The metal oxide causes the gas to dissociate into charged ions or complexes on the surface, which results in the transfer of electrons. This property has been exploited and used for the detection of inflammable and toxic gases. The electrical characteristics of these materials critically depend on the chemical composition of the metal-oxide surface and also on the nature and also on the work function behavior. The nature of the surface sites, the electron donor/acceptor properties of the test gas, the adsorption coefficient, the surface reactions, and subsequent desorption of the gases are the key features that determine the performance of semiconductor gas sensors. The interaction between the gas and solid mainly takes place on the surface, and hence the amount of atoms residing in grain boundaries (GBs) and the interface is critical for controlling the properties of these sensors [16]. Generally, these gases induce a resistance increase on n-type semiconductor oxides. This resistance increase can be associated with surface processes and explained in terms of capturing free electrons from the conduction band of the semiconductor by the adsorbed species [17].

It is generally recognized that in metal-oxide samples, the air-conductance behavior mainly represents the influence of surface reactions including electron exchange with the oxide, most notably those involving atmospheric oxygen and water vapor. The sensing mechanism in metal-oxide gas sensors is related to ionosorption of species over their surfaces. The most important ionosorbed species

when operating in ambient air are oxygen and water. The sensing mechanism is mainly governed by the fact that the oxygen vacancies on the oxide surfaces are electrically and chemically active. In this case, two kinds of sensing responses have been observed (Fig.2.1). (1) Upon adsorption of charge-accepting molecules, such as NO_2 and O_2 , at the vacancy sites, electrons are withdrawn and effectively depleted from the conduction band, leading to a reduction of conductivity. (2) On the other hand, in an oxygen-rich environment, gas molecules such as CO and H_2 , could react with the surface-adsorbed oxygen and consequently release the captured electrons back to the channel, resulting in an increase in conductance [18]. Conclusively, if one categorizes such redox-sensing response into reducing and oxidizing, which manifests in an increase and a decrease in the channel conductance, the sensing responses can be represented using two examples:

Reducing response: $CO + O \rightarrow CO_2 + e-$

Oxidizing response: $NO + e^- \rightarrow NO^-$

To avoid long-term changes metal-oxide-based gas sensors, should be operated at temperatures low enough so that appreciable bulk material variation never occurs and is high enough so that gas reactions occur in a time on the order of the desired response time. When a metal oxide is operated in the high-temperature ranges, the charge-transfer process, induced by surface reactions, determines its electrical resistance changes.

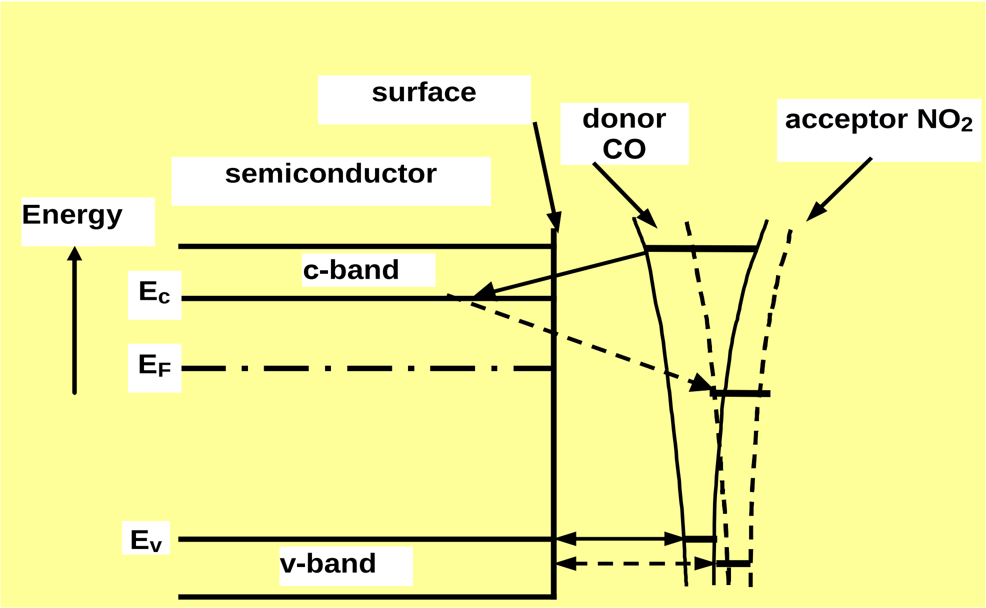


Fig.2.1 The types of adsorbates: Reducting agents (injection of e- to bulk): $H_2 \rightarrow H^+ + H^+ + 2$ e- and Oxiding agents (accept of e- from bulk): $O_2 + 2$ e- → 2O-

The fundamental sensing mechanism of these sensors lies, due to the interaction process, between the surface complexes such as O^- , O^{2-} , H^+ , and OH^- reactive chemical species and the gas molecules to be detected. The effects of the microstructure, namely, ratio of surface area to volume, grain size, and pore size of the oxide particles, as well as film thickness are well recognized. Lack of long-term stability has until today prevented wide-scale application of this type of sensor. The most recent research has been devoted toward nanostructured oxides, since reactions at GBs and complete depletion of carriers in the grains can strongly modify the material transport properties. Unfortunately, the high temperature required for the surface reactions to take place induces a grain growth by coalescence and prevents the achievement of stable materials. This high temperature ultimately decides the operation of these sensors.

Conductometric oxide thin films are the most promising devices among solid-state chemical sensors, due to their small dimensions, low cost, low-power consumption, online operation, and high compatibility with microelectronic processing. The present application field spans from environmental monitoring, automotive applications, air-conditioning in aeroplanes, spacecrafts, and houses, to sensor networks. The progress made on silicon technology for micromachining and microfabrication foreshadows the development of low-cost, small-sized, and low-power consumption devices.

2.2 Surface states in ionic crystals like SnO_2

In an ionic crystal like SnO_2, both anions and cations have poor coordination at the surface. The positively charged Sn ions at the surface have an incomplete shell of negative oxide ions around them. With too few negative ion neighbors, the positively charged ions are more attractive to electrons. So their conduction-band-like orbitals can be at a lower energy than the conduction band edge and can ‘capture’ electrons from the bulk. They also can bond well with a ‘basic’ molecule such as OH^-, which has an electron pair to give to the bond. The surface anions on the other hand do not have their quota of positive ions around them so their anioniclike orbitals can be at an energy level higher than the valence band edge. They can capture holes or give up electrons to the bulk. They can also bond well with an acid molecule like H^+ that has a pair of unoccupied orbitals. Actually at low temperatures it is expected that the ionic solids will normally be covered by chemisorbed water, the OH^- bonded to the surface at cationic and the H+ at anionic sites. At higher temperatures these water molecules can be driven off leaving the actisites open for interaction with bases or acid gaseous species or with electron-donor or electron-acceptor gaseous molecules (fig.2.1).

The electronic energy levels in the bandgap are termed “surface states” and for semiconductors there are both donor and acceptor levels present on the surface. If there no carrier exchange between the surface states and the bulk of the semiconductor the bands remain flat and the energy band model is as shown in Fig. 2.2.

Figure 2.3 illustrates what happens when the electrons move from the region of high E_F, the near surface region of the semiconductor, to a region of lower E_F, the surface states. This separation leads to a double layer voltage that either raises or lowers the energy of the surface states. This movement of bands near the surface is called “band bending”.

The electrical double layer formed may be of three types depending on the movement of electrons in to and out of the double layer. They are:

(1) Accumulation layer-This type of layer forms when electrons are injected in to an “n-type” semiconductor (Fig. 2.4)

(2) Depletion or Exhaustion or space charge laye. This forms when electrons are extracted from the conduction band of the n-type semiconductor (Fig. 2.5)

(3) Inversion layer. This type of layer forms in FET devices when a very strong oxidizing agent such as fluorine is present.

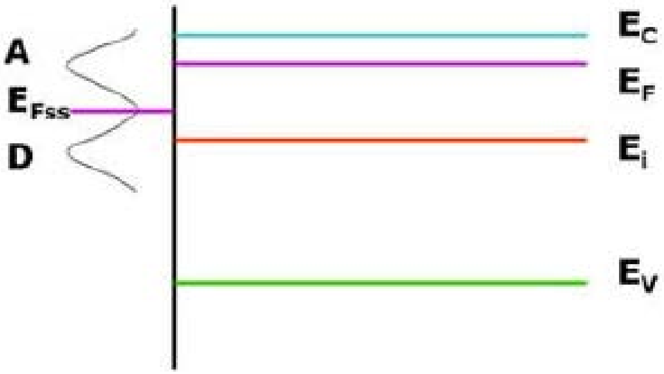


Figure 2.2. ‘Flat Band’ condition-No charge exchange between the surface states and the bulk.

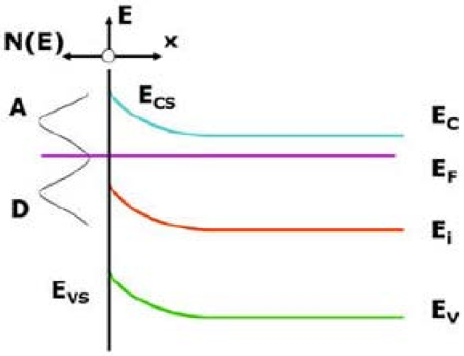


Figure 2.3. ‘Band Bending’-where the electrons from the semiconductor surface have moved to the surface states

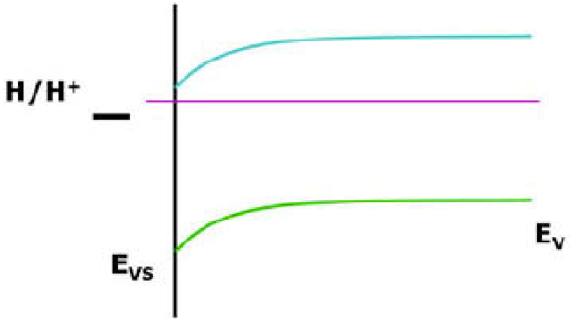


Figure 2.4. Formation of an accumulation layer between the electropositive surface species and the negatively charged semiconductor

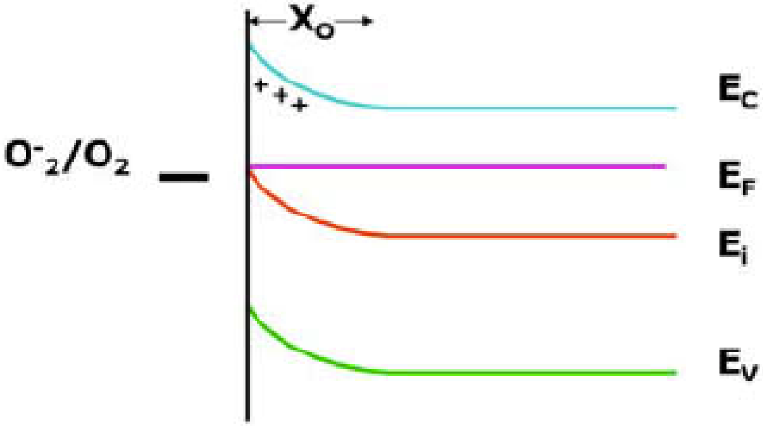


Figure 2.5. Formation of a depletion layer between the negatively charged surface species and the positive donor ions

Of these, the most important type of layer for gas sensing is the depletion layer. Hence it will be discussed in a little more detail. It was said earlier that the depletion layer forms in a ‘n-type’ semiconductor when electrons are extracted from it. Similarly in a p-type semiconductor this type of layer will form when holes are extracted from it. In an n-type semiconductor the double layer forms between the negatively charged surface states and the positively charged donor (immobile) ions in the bulk of the semiconductor. In a p-type semiconductor the double layer forms between the positively charged surface states and the negatively charged acceptor ions in the bulk.

2.3. Mechanism of gas sensing

The surface of a semiconductor bounds the periodicity of the crystal lattice points leading to localized surface states that could be shown by a quantum mechanical analysis. In case the energy levels are lying in the forbidden energy region of the semiconductor material, these surface states can either capture or inject charge carriers into the bulk crystal [20]. The corresponding surface charge induces a counter charge in the surface region built by ionized donors or acceptors present in the crystal.

The charge double layer, which is composed of the negative surface charge and the positive counter charge in the bulk, is equivalent to a band bending of the energy levels in the crystal as can be shown by solving the Poisson equation:

$$\frac{\partial^2\Phi}{\partial x^2} = \frac{q}{\varepsilon\varepsilon_0}N_i \qquad (2.1)$$

with Φ as the electrical potential, N_i as the ion density in the charge region at the semiconductor surface region. The Schottky approximation leads to

$$V_s = \frac{qN_s^2}{2\varepsilon\varepsilon_0 N_i}$$

(2.2)

which describes the relation of the potential difference between the surface and the bulk as a function of the surface charge N_s .

If the semiconductor is exposed to any gaseous atmosphere, additional gas-induced surface states arise due to the adsorption of ambient gas species very close to the surface region. This is shown in Figure 2.6a that shows the mechanism for the adsorption of oxygen species, as an example, which is a common component in most gas-sensing applications. Oxygen adsorbs on the surface by capturing an electron from the conductance band forming an O^{2-} ion. This electron capturing is equivalent to an occupation of a localized adsorbance (O^{2})-induced surface state. This mechanism is possible because the energy level of the surface state is below the Fermi level of the semiconductor without the presence of oxygen. The charge transfer induces a surface charge and a corresponding band bending in the semiconductor. The Fermi level is lowered until so much oxygen has adsorbed so that the surface energy level is equal to the semiconductor Fermi level [20]. This process is called “Fermi level pinning.” Due to the band bending, the surface region is exhausted of free charge carriers thus forming a region of high ohmic resistance, due to charge depletion, for the electrons.

Reducing gases, like CO, get oxidized, to CO_2 , consuming oxygen from the sensing material surface region. A dynamic equilibrium between the rate of desorbing oxygen due to the oxidation of reducing gas components and the rate of adsorbing oxygen of the ambient air is formed and the oxygen concentration on the sensor surface gets reduced. Following Equation 2.2, the energy bending is lowered by a certain amount depending on the concentration of the reducing gases. In most cases, the sensor material has a polycrystalline structure, as is shown in Figure 2.6b. All the polycrystalline grains are attached to each other via the GBs. Individual grains are surrounded with neighboring grains. Contact region of grain, with their high ohmic regions, forms a so-called double Schottky energy barrier for the charge carriers, that is, electrons. If we assume thermal emission as the dominant charge-transport mechanism, in this case, the conductivity G over such barriers depends on the energy barrier and is generally given as $G \sim e^{-qV_s/kT}$.

The electrical response signals of any polycrystalline metal-oxide-based sensors may be formally described by resistance (with $R = G^{-1}$) and capacitance units in RC . They contain high concentration of GBs between nanometer-sized crystalline grains. Their sizes and the corresponding size distribution vary with preparation techniques, deposition conditions, annealing temperature, and time durations. Of particular importance for understanding the electrical response signals of nanocrystalline metal-oxide thin films, upon exposure to reducing or oxidizing gases, are measurements of impedances in comparison with dc conductances. The latter are performed with two- or four-point contact arrangements. This makes it possible to separate specific contributions from contacts, bulk, surface, and GBs, all of which contribute to the sensor response [21]. The details are indicated in Figure 2.7. The variation of these parameters will give rise to a “signal” from the sensing element. The figure shows the schematic representation of elementary steps during detection of molecules. Surface and bulk reactions lead to changes of overall dc or ac conductance of the sensing material, which contain contributions from surface, bulk, boundary, or metal contacts, and finally from GBs. Equivalent circuits with different R–C units, each of which corresponds to a characteristic charge carrier transport and relaxation processes are indicated in it. Usually two ohmic contacts are used in conductance sensors. In Schottky-diode sensors, one contact/oxide interface is a Schottky barrier with nonlinear current–voltage behavior that is sensitively influenced by chemisorbed particles.

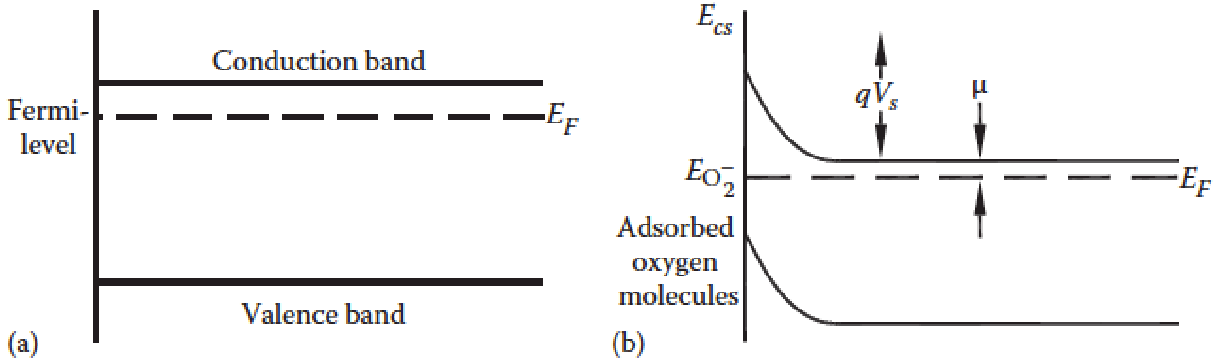


Figure 2.6. (a) Simple energy band diagram showing conduction band, valence band, and Fermi level. (b) Band bending due to adsorption of oxygen species on the sensor surface altering its properties.

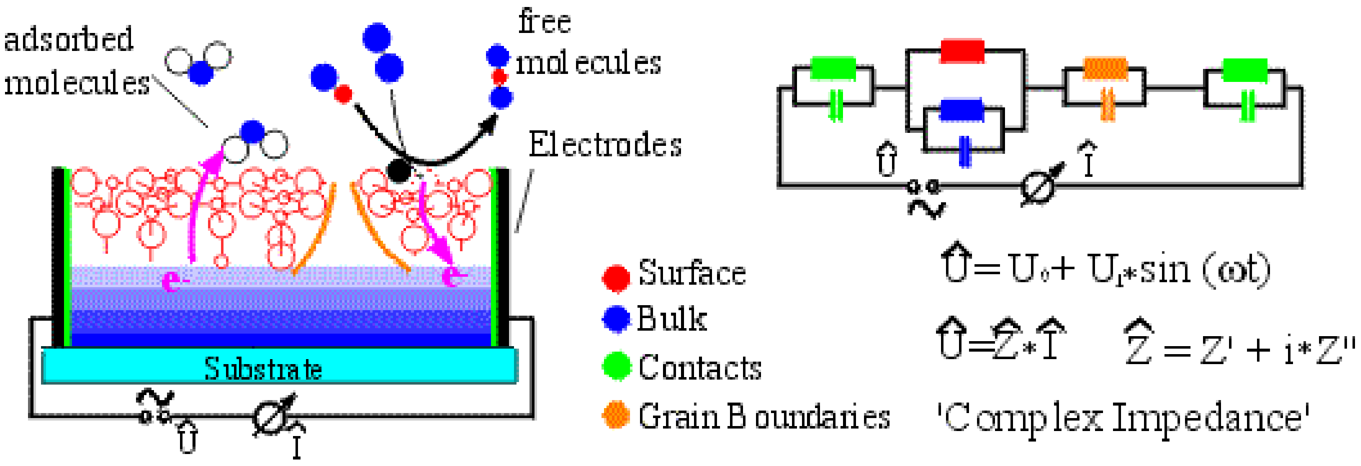


Figure 2.7a. Schematic representation of elementary steps during detection of molecules. Surface and bulk reactions lead to changes of overall dc or ac conductance of the sensing material, which contain contributions from (1) surface, (2) bulk, (3) boundary or contacts, and (4) grain boundaries. Equivalent circuits with different R–C units, each of which corresponds to a characteristic charge-carrier transport and relaxation processes. Usually two ohmic contacts are used in conductance sensors. In Schottky-diode sensors, one contact/oxide interface is a Schottky barrier with nonlinear current–voltage behavior that is sensitively influenced by chemisorbed particles.

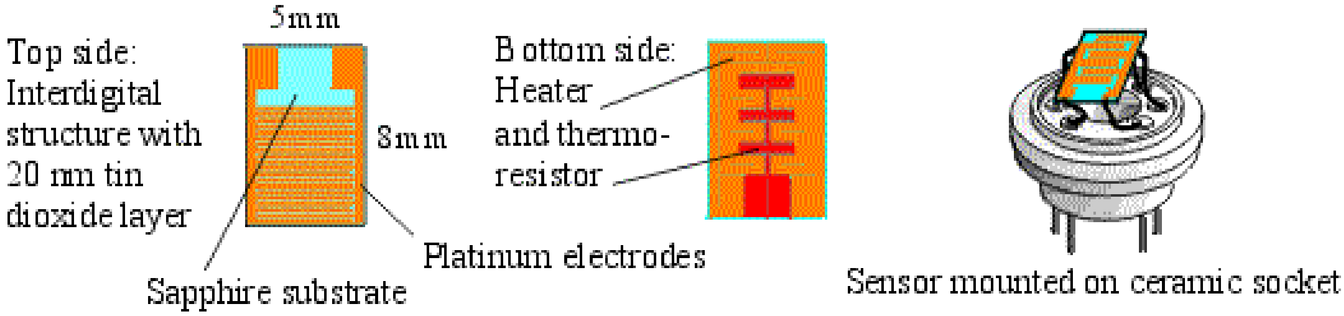


Figure 2.7b The design of sensors

Metal-oxide-based sensors and their corresponding sensor response signals upon interaction with gas molecules may be classified according to their electrical characteristics, that is, the voltage dependence of conductances G and/or capacitances

C . This is schematically shown in Figure 2.8, contact arrangements are also shown here. Figure 2.8a shows resistor-type sensors with ohmic properties, which show linear I–V curves. In this case, their conductance G = dI/dV is independent of the applied voltage and reversibly change as a function of partial pressures of gaseous species. Here, p_1 and $p_=$ are taken as two different partial pressures. Thin film, thick film, pellets, and beads are suitable for such studies. The second type of device is a diode-type sensor, which shows partial-pressure-dependent nonlinear I–V and C–V curves. This typically induces change in cut-in voltage of the diode and dynamic resistances. The third type device is a capacitor-type sensor based on metal-oxide-semiconductor and metal-insulator-semiconductor (M-O-S and M-I-S capacitor) devices, which show voltage-dependent C–V curves. Here, changes in the partial pressure lead to a shift of the flat-band voltage ΔV_{FB}, as determined by the transition between charge accumulation and depletion in the semiconductor close to the surface region. The variations of I–V and ΔVFB are shown in Figure 2.8b and c [21]. Reliable chemical sensors require the control of geometric and electronic structures of the different surfaces and interfaces at the atomic scale. This concerns in particular the metallic surface dopants of oxide surfaces and their electrical contacts.

In solid-state sensors, the gas molecules adsorb onto the sensing material surface bringing a change in the resistance value of it. In the majority of sensors, the initial output of the sensor is linear or very close to it, but as the gas concentration increases, the output signal gradually reduces and drifts toward the nonlinearity region, resulting poor resolution. After certain concentration, the nonlinearity factor disappears and the signal saturates to a fixed value independent of exposed gaseous concentration.

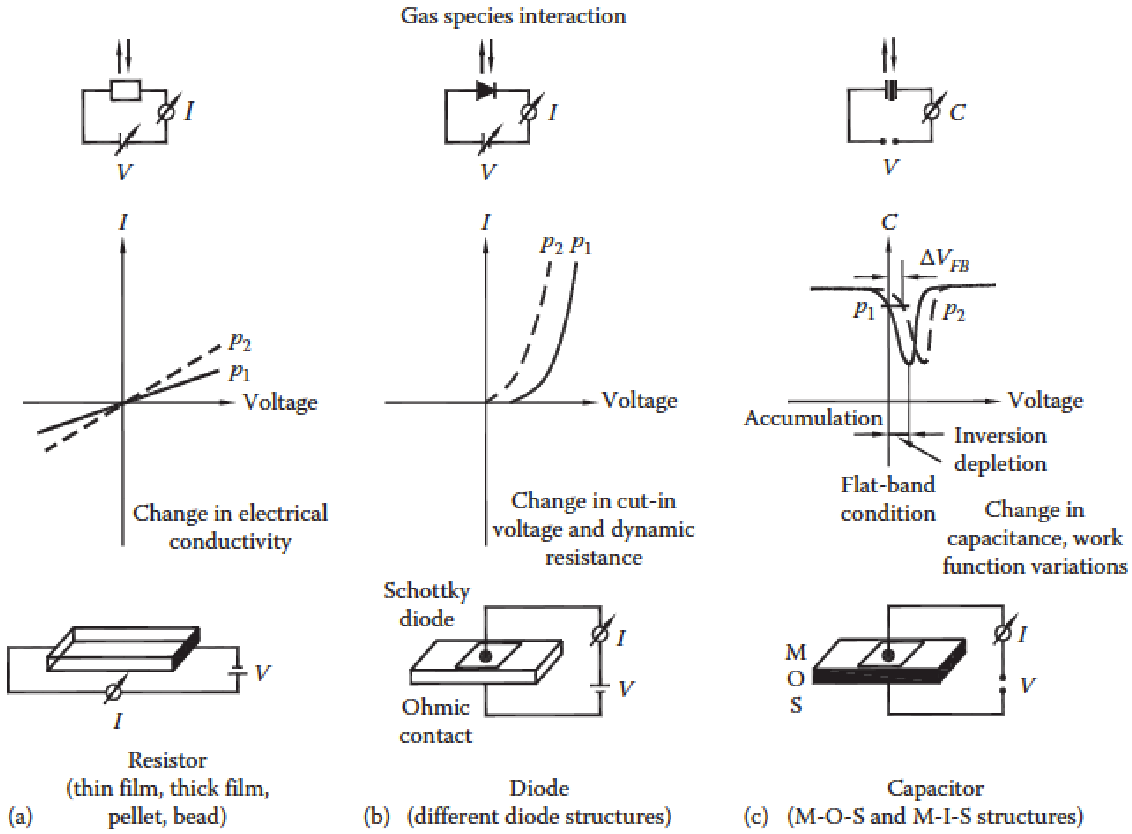


Figure 2.8. Current–voltage (I–V) and capacitance–voltage (C–V) curves of typical metal-oxide-based chemical sensors. (a) Resistor-type sensors with ohmic properties show linear I–V curves, which reversibly change as a function of partial pressure of molecules. (b) Diode-type sensors show partial-pressure-dependent nonlinear I–V curve. (c) Capacitor-type sensors based on MOS devices show voltage-dependent C–V curves

When the gas disappears, the sensor returns to its original condition. In this process, no material consumption takes place and hence provides a long life expectancy in the usage. This is one major advantage of these gas sensors.

The adsorption coverage of the semiconductor surface is given by the Volkenstein isotherm

$$\theta = \frac{\beta p}{1+\beta p}$$

here θ is the adsorption coverage (ratio between number of chemisorbed species to the total chemisorption sites, N_{ad}) β is the coefficient of the Volkenstein isotherm p is the partial pressure of the gas.

The coefficient depends on the position of the Fermi level at the surface. The existence of a strong form of chemisorption implies the presence of a surface charge given by $N^{-} = N_{ad}\theta f$, where f is the occupation probability of the strong form of chemisorption given by the Fermi–Dirac statistics.

2.4 Resistance and Capacitance Variations

The easiest way to transduce the gas sensor is by simply measuring the dc resistance of the sensing element as a function of the surrounding atmosphere (fig.2.10). The design of thin-film resistors normally requires knowledge of the resistance of a certain geometric pattern of the thin-film material between two terminating contacts [23]. The transducer has to provide at least two electrical contacts on the sensing element to measure conductance changes and a heating system in order to maintain the element at the suitable operation temperature. Depending on the material this value is, in general, of the order of a few hundred degrees centigrade.

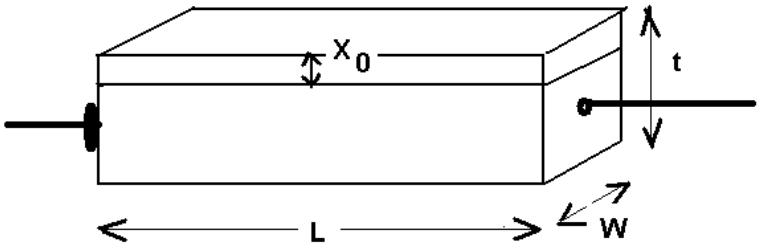


Fig.2.10 The schema of metal-oxide resistive sensor.

The simplest transducer is a bulk insulating substrate with electrical contacts on one side of the substrate and a heater preferably on the backside of the sensing element. The relation between the conductance and the oxygen partial pressure is

$$\sigma = C \exp\left(\frac{-E_a}{kT}\right) p_{O_2}^{1/m}$$

where σ is the conductivity , C is proportionality constant, E_a is the activation energy for the bulk conduction, p_{O2} is the oxygen partial pressure [24]

The constant m , the oxygen sensitivity, depends on the defects involved in the conduction mechanism. When the defects are represented by doubly ionized oxygen vacancies, m = −6, but if the defects are represented by metal vacancies, different values of m are found.

At constant partial pressure of oxygen and working temperature, the experimental dependence between the conductance and the partial pressure of a single gas is given by

$$G = G_0 + \gamma (p_{gas})^m$$

where p_{gas} is the gas partial pressure, γ is a constant that depends on the material
It can be seen that by considering a simplified model of the sensor in which, it was formed by parallel chains of grains between the electrodes, the sensor resistance was mainly determined by the grain boundary resistance R_{gb} (resistance between grains):

$$R_{gb} = R_0 \frac{e^{qV_s/kT}}{N_D} = R_0 e^{(E_C - E_F)_{bulk}/kT} e^{qV_S/kT}$$

which is a function of the free carrier density, N_D , and surface barrier height, V_s . So, resistance variations can be produced mainly by the modification of the density of free electrons (movement of the Fermi level), or by the modification of the barrier height through the change of the charge state of the interfaces of the grains [25]. In the case of solid-state semiconductor gas sensors, the gas-sensitive layer usually consists of a polycrystalline metal-oxide film. The gas-detection principle is based on the variations of the depletion layer at the GBs, in presence of reducing or oxidizing gases, which lead to leading to a change of the conductivity of the sensing material. The majority of these devices play a key role on the behavior of the material used, grain size of the polycrystalline, and the dopants added (present) in them.

In general, the variation of conductivity at gas adsorption is described as

$$\frac{dG}{G} = -\frac{dx_0}{t - x_0},$$

where

$$G = \frac{1}{\rho}\frac{Wt}{L}\left(1 - \frac{x_0}{t}\right)$$

The obtained sensors were characterised electrically by means of gas mixing stations. Typical operation tempertures for the sensors are 200°C-400°C.The humidity can be varied between 0 and 100% RH. In the presence of CO, the sensor resistance decreases . CO concentration down to the ppb range can be detected. The impact of a gas as a sensor can be measured by the sensor signal. For reducing gases the sensor signal is defined as the ratio of the resistance in air (R_O) to the resistance in an ambient containing CO (R_{Gas}).The calibration curve shows the dependence of the sensor signal to CO and CH_4 on their concentrations.

Capacitance Variations

The capacitance variation property for gas-sensing applications has not gained much popularity, in the form of metal-oxide-semiconductor (MOS) device, to study voltage-dependent C–V characteristics. The shift of the flat band voltage ΔV_{FB} , as determined by the transition between charge accumulation and depletion in the semiconductor close to the surface region is still not popular in the case of gas-sensing devices. Probably, the exposed regions of the material in MOS capacitor configuration, available surface area of sensing element limit the gas–surface interaction, which is crucial for the sensors.

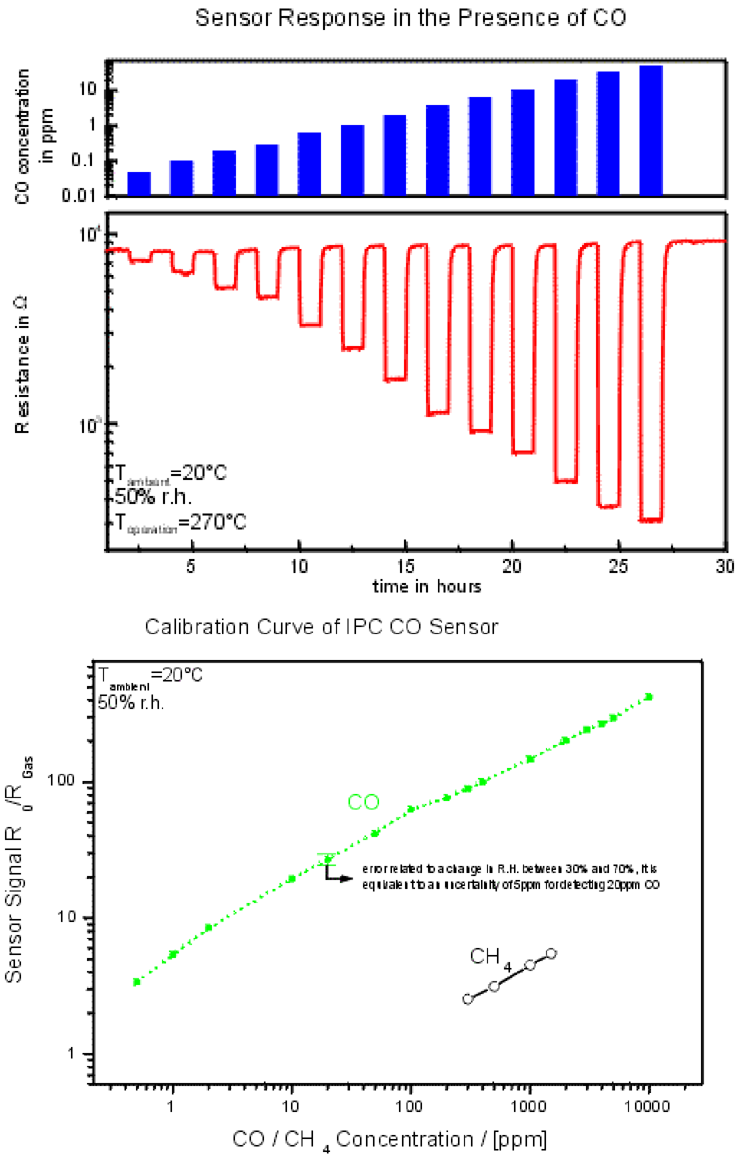


Fig.2.11 Calibration curve of a CO sensor for CO concentrations between 500ppb and 10.000ppm (with CH_4 as interfering gas)

2.6 Sensor Properties

2.6.1 Sensitivity of Gas Sensors

Sensitivity, S, is defined as the ratio of change of resistance in test gas $\Delta R = R_a - R_g$, to the value of resistance in air R_a where R_g is the sensor resistance in the presence of the test gas [16]:

$$S = \frac{\Delta R}{R_a} = \frac{|R_a - R_g|}{R_a}$$

Sensitivity is also defined by the formula $S = (R_0 - R_{gs})/R_{gs}$, where R_0 is the resistance of the sensor before passing gas and R_{gs} is after passing gas and reaching the saturation value [30]. Others have defined this by $S = R_a/R_g$, where R_a and R_g express the resistance of the sensor in air and in detecting gas [31–33].

The sensitivity of the semiconducting oxide gas-sensitive sensor can usually be empirically represented as $S_g = Ap_g^{\beta}$, where p_g is the target gas partial pressure, which is in direct proportion to its concentration, and the sensitivity is characterized by the prefactor A and exponent β [34]. β may have some rational fraction value (usually 1 or 1/2), depending on the charge of the surface species and the stoichiometry of the elementary reactions on the surface.

2.6.2 Concentration-Dependent Sensitivity

At low test-gas concentration values, the sensitivity values increase steeply with an increase in the concentration values; however, at higher concentrations, the increase in the gas-sensitivity values becomes more gradual in nature. With a fixed surface area for each sensing element, a lower gas concentration implies a lower coverage of gas molecules on the available sensor surface, and hence, lower surface reactions occurred. An increase in the gas concentration raises the surface reactions due to a larger surface coverage. Further increase in surface reactions will be gradual when the saturation point on the coverage of molecules at the fixed surface area is reached. Further increase in the concentrations may not yield much sensitivity due to nonavailability of fresh surface area and it may even saturate at a fixed value as no further place is available for surface reactions to take place. This ultimately limits the signal available for analysis.

2.6.3 Temperature-Dependent Sensitivity

The sensitivity also depends on the temperature of the sensing element. Most of the metal-oxide sensors operate at a temperature above the room temperature range extending up to several hundred degrees centigrade and it is necessary to know the correct operating temperatures of these sensing elements. To obtain good sensitivity, one should know and operate these elements accordingly. Depending on the material and the detecting gaseous species, the values differ. In a typical case, the sensitivity increases gradually with temperature and becomes less gradual at higher temperature values.

From linear dependency, it deviates to a maximum value and beyond this point the sensitivity falls rapidly. This type of behavior is being exhibited by almost all the metal-oxide-based sensors. Figure 2.12 shows a sensitivity graph illustrating the response of SnO_2 sensor to 100 ppm methane gas in dry air as a function of sensor operating temperature [35].

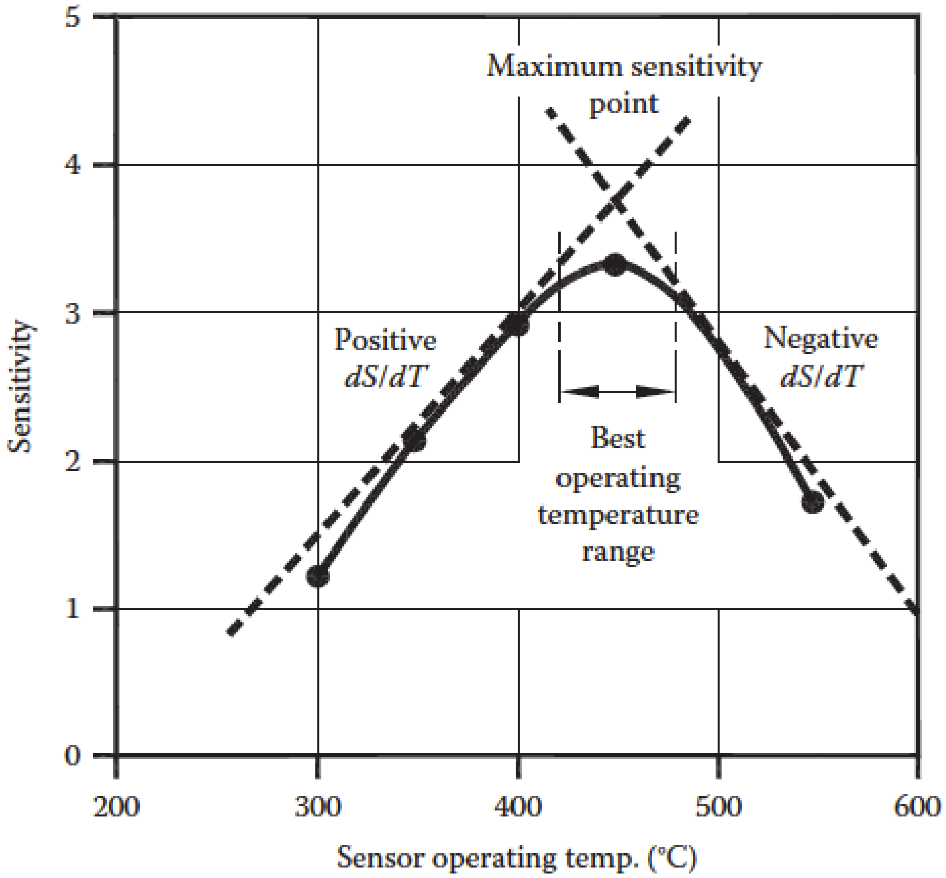


FIGURE 2.12 Response of SnO_2 sensor to 100 ppm of CH_4 gaseous species in dry air as a function of sensor operating temperature. Sensor sensitivity with temperature changes from positive value to negative value. Best operating range is indicated here for sensor operation.

The sensitivity increases continuously from 300oC up to 450oC and this is the best performance point. At higher temperatures, there is a fall in the sensitivity. Sensor sensitivity with temperature changes from positive value to negative value. Further rise in temperature will reduce the sensitivity values. From the experimental results, one can identify the best operating temperature range for the sensing element.

Similar characteristics are seen for the sensors based on In_2O_3 sensor for detecting 2 ppm 2 gas, as shown in Figure 2.13 [35]. Here also, the sensitivity initially rises with temperature and after reaching a maximum value it falls steadily moving toward a saturation value. The sensor sensitivity with temperature changes from positive value to negative value. Best operating range is indicated here for sensor operation. Positive slope indicates the reaction limited case whereas the negative slope is due to binding energy of the gaseous species.

Creation of high-temperature zones at small defined areas is crucial for operating small and micro-gas-sensing devices. Simple silicon surfaces take the temperature range of 50oC–150oC when directly heated. Its good thermal conducting properties limit the temperature rise. In case of silicon-on-insulator structures one can achieve a temperature range between 50oC and 350oC due to the presence of an insulator. Micromachined silicon structures, such as suspended platforms, defined using MEMS bulk and surface micromachining technique, are best suited for high-temperature creation in small localized areas. The air insulation provides the required thermal isolation.

This way, one can create a temperature between 50oC and 500oC. Beyond 500oC, it is better to think of hybrid microstructures with external heating facility and keep the sensor away from the remaining sensor electronics. One has to design by carefully selecting the material, its operating temperature range, and the suitable technique to use these material compositions for sensor applications.

It is reported that by properly selecting the microhotplates, it is easy to achieve a temperature of the order of 350oC where most of the metal oxides show very good response properties. Such fabrication techniques are compatible with standard industrial CMOS technology as well [36]. The hotplate gives excellent temperature homogeneity with fluctuations of less than 2% of the respective operation temperature on the heated area. The membrane power efficiency is reported to be 4.8oC/mW. The system also exhibits excellent thermal isolation between heated membrane area and the circuitry area on the bulk chip.

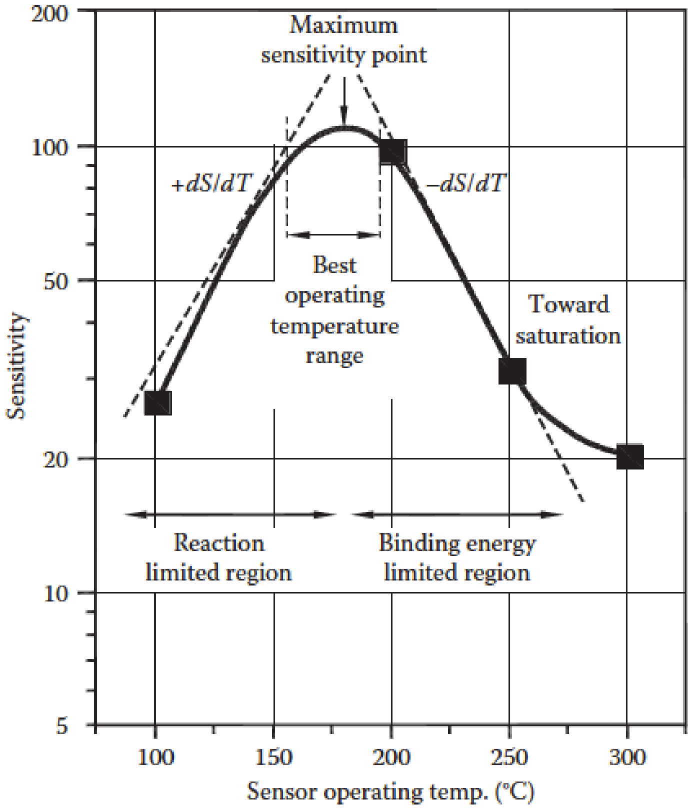


Figure 2.13 Response of In_2O_3 sensor to 2 ppm of NO_2 as a function of sensor operating temperature. Sensor sensitivity with temperature changes from positive value to negative value. Best operating range is indicated here for sensor operation. Positive slope indicates the reaction limited case, whereas the negative slope is due to binding energy of the gaseous species

2.6.4 Response and Recovery Times

For sensor applications, it is also very important for a sensor to react accurately to quick changes in gas concentration values. Optimization of gas sensor to respond quickly to sensing requirements is not free from its problems. Many gas sensors exhibit slower response times. Here, response time (τ_{90}) is typically defined as the time it takes for a sensor to reach a 90% of full-scale reading after being exposed to a full-scale concentration of a given gas. In certain cases, τ_{63} or τ_{80} are also used. In the majority of cases, it is τ_{90} criteria adopted for the evaluation of this time constant. Similarly, the recovery times are also important. These two time constants generally differ and in some cases recovery may pose a problem. In such cases, it is better to increase the temperature of the sensor above 50oC of its operating temperature, a process known as “regeneration of sensor,” before it is reused. This regeneration cycle may be repeated periodically for the better efficiency of the sensing element.

The response time is proportional to particle size, d , of sensing material when the sensor kinetics is controlled by surface reaction and that is proportional to d^2 when the kinetics is controlled by diffusion. Regardless of the rate-limiting step, the response time decreases with decreasing particle size [3]. Therefore, it is expected that response time is much shorter if the sensing material is a nanosized particle.

2.7.Advantages of nanomaterials for chemical sensors

Nanosized materials have been widely used to produce new semiconductor gas sensors. These materials have received considerable interest in recent years because of their unique properties both in physics and chemistry. It is now well established that these materials in this form exhibit some interesting properties. There are also multiple advantages of nanostructured conducting materials as compared to the coarse ones in their applications. The nanostructured materials can effectively increase the surface area for catalytic reaction and electrode reaction and result in an increase in the catalytic activity for oxidation to take place [38]. These materials also greatly reduce the necessary sintering temperature for densification, which is beneficial for many scientific applications. The sensitivity as well as the range of dynamic response of a gas sensor can be improved significantly when nanotechnology methods are used. For nanocrystalline semiconductors, nearly all conduction electrons are trapped in the surface states, which will enhance the sensor response for surface reactions. High porosity and large surface areas are desirable for numerous applications of nanostructures to fulfill the demand of high efficiency and high surface activity.

Nanosized structures show outstanding properties in gas-sensor applications, especially at low operating temperatures, and, high sensitivity, particularly with lower grain sizes. Therefore, ultrafine layers or any feature that resembles lower feature sizes allow various improvements in their performance. The sensitivity of the sensor largely increases with the reduction of the feature size. To explain this property, we select a simple example. Lu et al. [107] showed that the nanosized SnO_2 material is sensitive to 500 ppm concentration of CO at near room temperatures, as shown in Figure 2.14a. In this figure, the effect of nanoparticle size is compared to the sensitivity of CO gas. There is hardly any change in the sensitivity values observed for the particle size that is larger than 50 nm. Slow raising trend is noticed between particle sizes between 50 and 15 nm. When the particle sizes are smaller than 10 nm, the sensitivity shoots up to give a quick increase in the sensitivity values. This type of advantage is expected with almost all the metal-oxide nanostructures for gas-sensing applications (fig.2.14b).

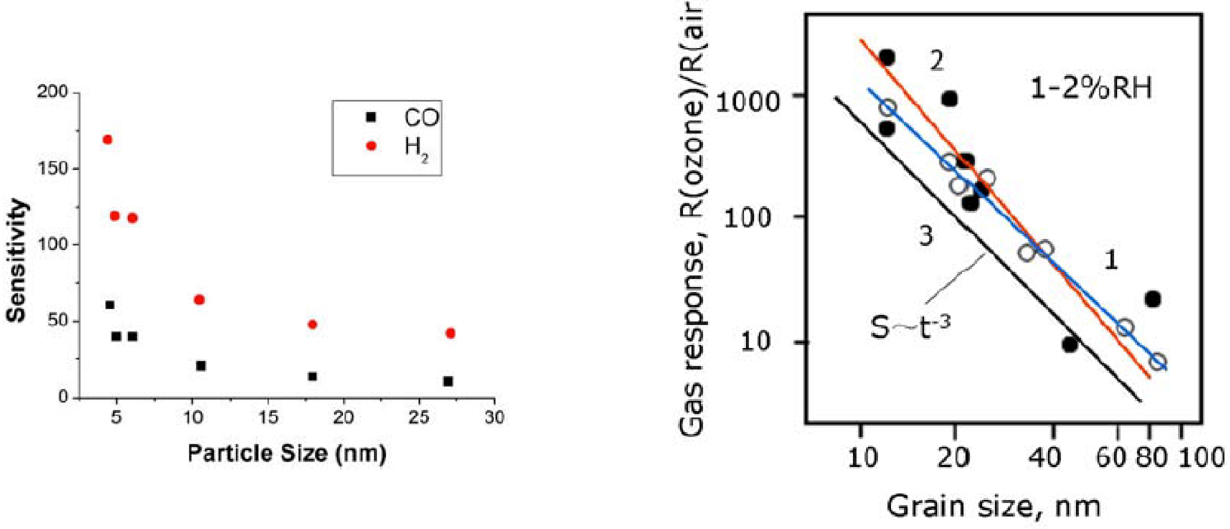


Figure 2.14 Effect of particle size on gas sensitivity of CO and H_2 gas using nanosized SnO_2 material. Influence of In_2O_3 film grain size on the gas response to ozone.

Nanotechnology, which exploits materials of dimensions smaller than 100 nm, is addressing the challenge and offering exciting new possibilities. In addition to a large surface-to-volume ratio and a Debye length comparable to the small size, they demonstrate different but superior sensitivity to surface chemical processes. For several years, many research works and commercial products have sought to exploit the properties of these nanocrystals as a constitutive material in the sensitive layer for gas-sensing applications. Of particular interest for metal-oxide-based conductance sensors are the grain sizes below the Debye length of electrons. Charge-transfer reactions during chemisorption and catalytic reactions at surfaces lead to uniform Fermi energy shifts instead of band-bending effects. Such nanostructured semiconductor oxides offer some advantages: a larger surface-to volume ratio and a grain size of few nanometers, which is comparable to the depth of the space charge layer that surrounds the nanograin [40].

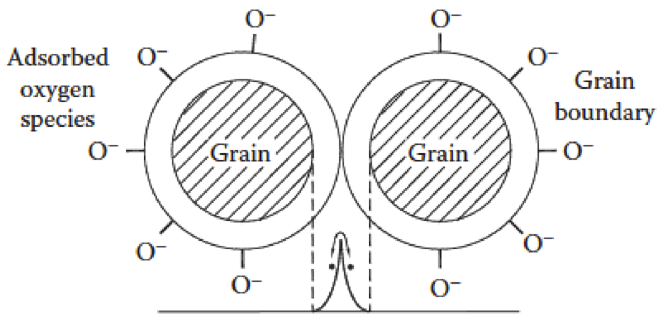
These specific features make nanocrystalline semiconducting oxides very promising for development of high-sensitivity, fast-response gas sensors, in which just surface processes play the key role in the formation of a sensor signal. Every sensor device possesses a sensing body (or resistor), which is a porous assembly of tiny grains of metal oxide. Under exposure to air, oxygen is adsorbed on the grains as anionic species (typically O^-), inducing an electron-depleted layer to increase surface potential and work function [41].

No charge carriers are present in these depleted regions. In cases where the grains are connected to their neighbors through the grain boundaries, a potential barrier for migration of electrons, often called double-Schottky barrier, is formed across each grain boundary, as shown in Figure 2.15a. This barrier plays a dominant role in determining the resistance of the sensing body. When a reducing gas is brought into contact, the oxygen adsorbates are consumed, leading to a mitigation in potential barrier and then also in R .

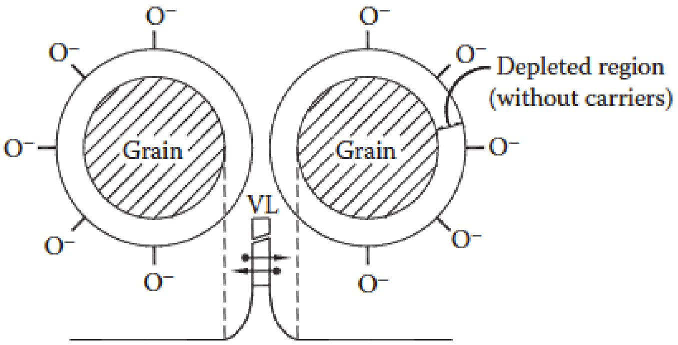
Thus, R decreases with increasing partial pressure (p) of the reducing gas. On the other hand, on contact with an oxidizing gas like NO_2 in air, target gas and oxygen are adsorbed competitively as anionic ions on the same grains. If adsorption of the target gas is stronger than that of oxygen, the surface potential barrier will increase upon exposure to the target gas, resulting in an increase in R . Electrons can also be transported by tunneling through a small gap between oxide grains even when grain boundaries are not formed effectively throughout the sensing body, as shown in

Figure 2.15b. Even in this case, tunneling current is attenuated by the surface potential of oxide grains, resulting in almost the same dependence of R on p . Here, the vacuum level is indicated for reference.

For simplicity, if we assume that the radii of oxide grains (n -type semiconductor) are much larger than Debye length. This allows us to approximate a depletion layer of each grain by a one-dimensional (1-D) model, as shown in Figure 2.16, where x denotes depth from the surface. When a certain surface state is present at an energy level below the Fermi level under the flat band condition, electrons are transferred from the conduction band to the surface state and this transfer continues until the Fermi level becomes constant throughout, from the surface to bulk in equilibrium state [41]. The resulting depletion layer is well described by a depth profile of potential energy of electrons, qV (x), and well characterized by surface potential barrier, qVs , and depletion depth, w, as shown in the figure. This depletion region width is a moving boundary within the grain region.



(a) Carrier cross over on either side



(b) Carrier tunneling on either side

Figure 2.15 Models of electron transfer between grains: (a) double-Schottky barrier model and (b) tunneling model. VL stands for vacuum level

Figure 2.16 shows the moving boundary within the grain region altering depletion region width depending on the surface reactions. Here, q is represented as elemental charge of electron and V(x) and Vs are electric potential at given depth and surface, respectively.

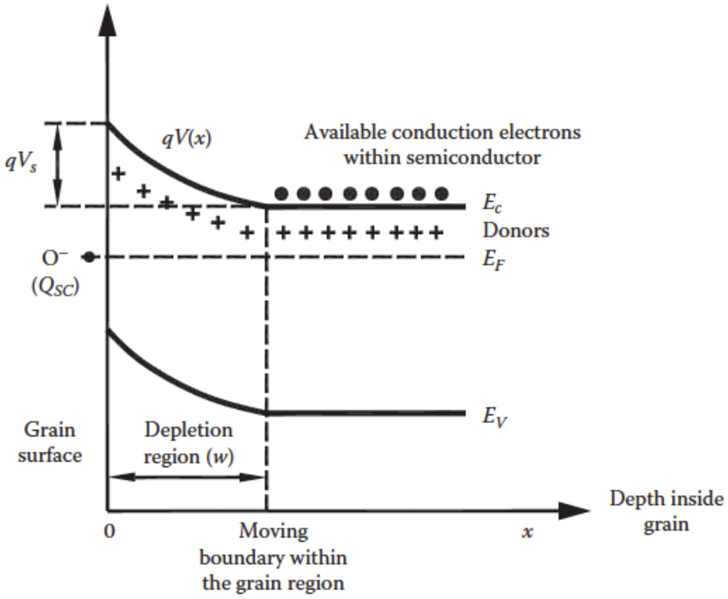


Figure 2.16 Schematic drawing of surface-depletion layer formed by oxygen adsorption for an n-type semiconductor. Diagram shows the moving boundary within the grain region altering depletion region width.

In a low-grain-size metal oxide almost all the carriers are trapped in surface states and only a few thermally activated charge carriers are available for electrical conduction. In this configuration, the transition from “activated” to strongly “nonactivated” carrier density, produced by target gas species, has a great effect on sensor electrical conductance. The challenge became to prepare materials with small crystalline size that were physically stable when operated at high temperature for longer periods. The strong dependence of the grain–grain conductivity from the band bending, as shown in Figure 2.17, is the main reason for the high sensitivity of SnO_2-based gas sensors to reducing gases. The polycrystalline structure could be connected by electrodes in order to measure the gas-depending sensor resistance that results from the network structure and the resistance of the several grains contacts [20,42] as shown in the figure. In the physical model, polycrystalline structure of the sensor surface shows the adsorbed oxygen species and the depletion regions.

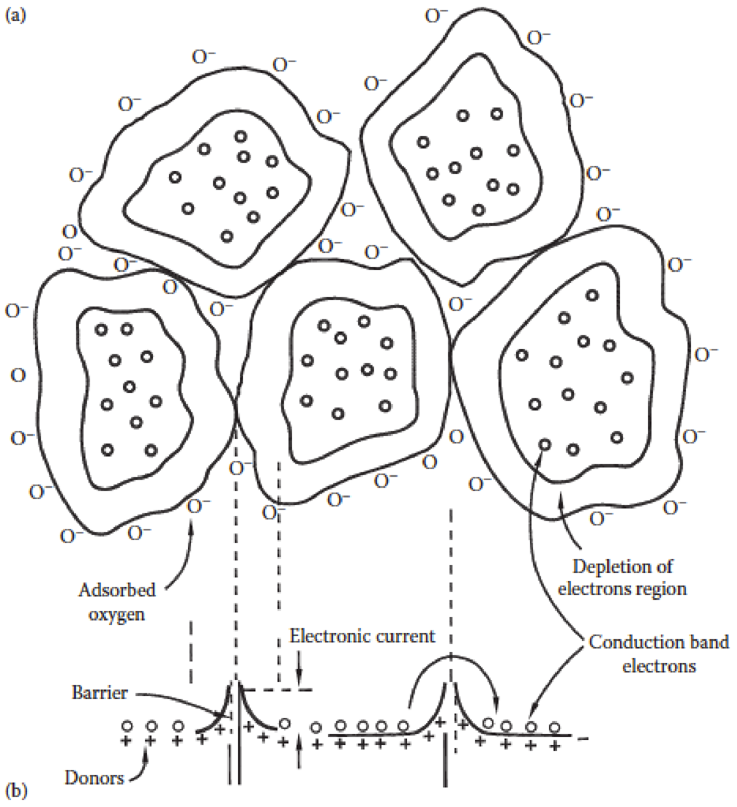


Figure 2.17 Polycrystalline structure of the sensor surface in (a) physical model and (b) the corresponding band model. Diagram shows the adsorbed oxygen species and the depletion regions within the grain regions

In the band model, conduction band electrons and the tunneling of charge carriers across the barriers are shown. Sensor performance depends on percolation path of electrons through intergranular regions, by varying small details in the preparation process. Both thin and thick film electrical properties drift due to the grain coalescence, porosity modification, and grain-boundary alteration [43]. These effects become more critical because the metal-oxide layers must be kept at a relatively high temperature in order to guarantee the reversibility of chemical reactions at surface. Thus, several solutions have been put forward to stabilize the nanostructures. Addition of a foreign element or phase to the material is one such effort and technique generally adopted to maintain the physical structure. The peculiar characteristics and size effects of nanosized structures make them interesting for both fundamental studies and potential nanodevice applications, leading to a third generation of metal-oxide gas sensors. Their nanosized dimensions generate properties that are significantly different from their coarse-grained polycrystalline counterpart. Surface effects appear because of the magnification in the specific surface of nanostructures, leading to an enhancement of the properties related to that, such as catalytic activity or surface adsorption processes. These specific properties are the basic phenomenon underlying the solid-state gas sensors.

In current material science research, the use of nanosized materials for gas sensors is rapidly arousing interest in the scientific community. One reason is the surface-to-bulk ratio and another is that the conduction type of the material is determined by the grain size of the material. When the grain size is small enough (the actual grain size D is less than two times the space charge depth L), the material resistivity is determined by grain control, and the material's conduction type becomes surface conduction type [44]. Hence, the grain-size reduction becomes one of the main factors in enhancing the gas-sensing properties of semiconducting oxides.

The equilibrium phase of nanosized particles may deviate from that of the bulk owing to the contributions of the surface energy and surface stress. Knowledge and understanding of the behavior of small particles and inclusions is of importance for many engineering materials, particularly with the ongoing trends toward materials with a controlled nanostructure.

In this context, the study of phase transitions of small particles and inclusions deserves special attention [45], since it will probably play a key role in the improvement of the understanding and control of the microstructure.

Nanotechnology could provide sensing materials capable of improving the sensitivity, selectivity, stability, and the speed of sensor technology. One recent area of interest involves applying quasi-1-D metal-oxide nanostructures, such as nanowires, nanotubes, nanobelts, etc., as prospective gas-sensing elements [46]. The nanostructures maintain high surface area to bulk volume ratio and the comparability of their radii and their corresponding Debye lengths allow rapid transduction of surface interactions into measurable conductance variations. It also opened a totally new type of electronic device where nanowire-based active device is being proposed.

The best example for one such proposal is an open-gate FET transistor. To explain the effect of grain size on the sensitivity of metal-oxide gas sensors, a semiquantitative model was proposed by Rothschild and Komem [47], which is shown in Figure 2.18. According to this model, the sensing material consists of partially sintered crystallites connected to their neighbors by narrow necks. Those interconnected grains form larger aggregates and are connected to neighbors through the grain boundaries. Three different cases can be distinguished according to the relationship between the grain size D and the width of the depletion layer thickness L that is produced around the surface of the crystallites due to chemisorbed adions.

When $D \gg 2L$ most of the volume of crystallites is unaffected by surface interactions with the gas phase. In this case, the predominant effect of the ambient gas atmosphere on the sensor conductivity is introduced via the GB barriers for intercrystallite charge transport from one grain (agglomerate) to another.

The electrical conductivity σ depends exponentially on the barrier height $q|VB|$: $\sigma \propto \exp(-q|VB|/kT)$, where q is the elementary electron charge, VB is the GB potential, k is Boltzmann constant, and T is temperature. According to depletion approximation $q|VB| \propto (N_t^-)^2$, where N_t^- is the trapped charge density at the surface of the crystallites. N_t^- can be modified by charge-transfer interactions with reactive gases such as O_2, NO_x, and CO, and the conductivities are sensitive to the ambient gas composition.

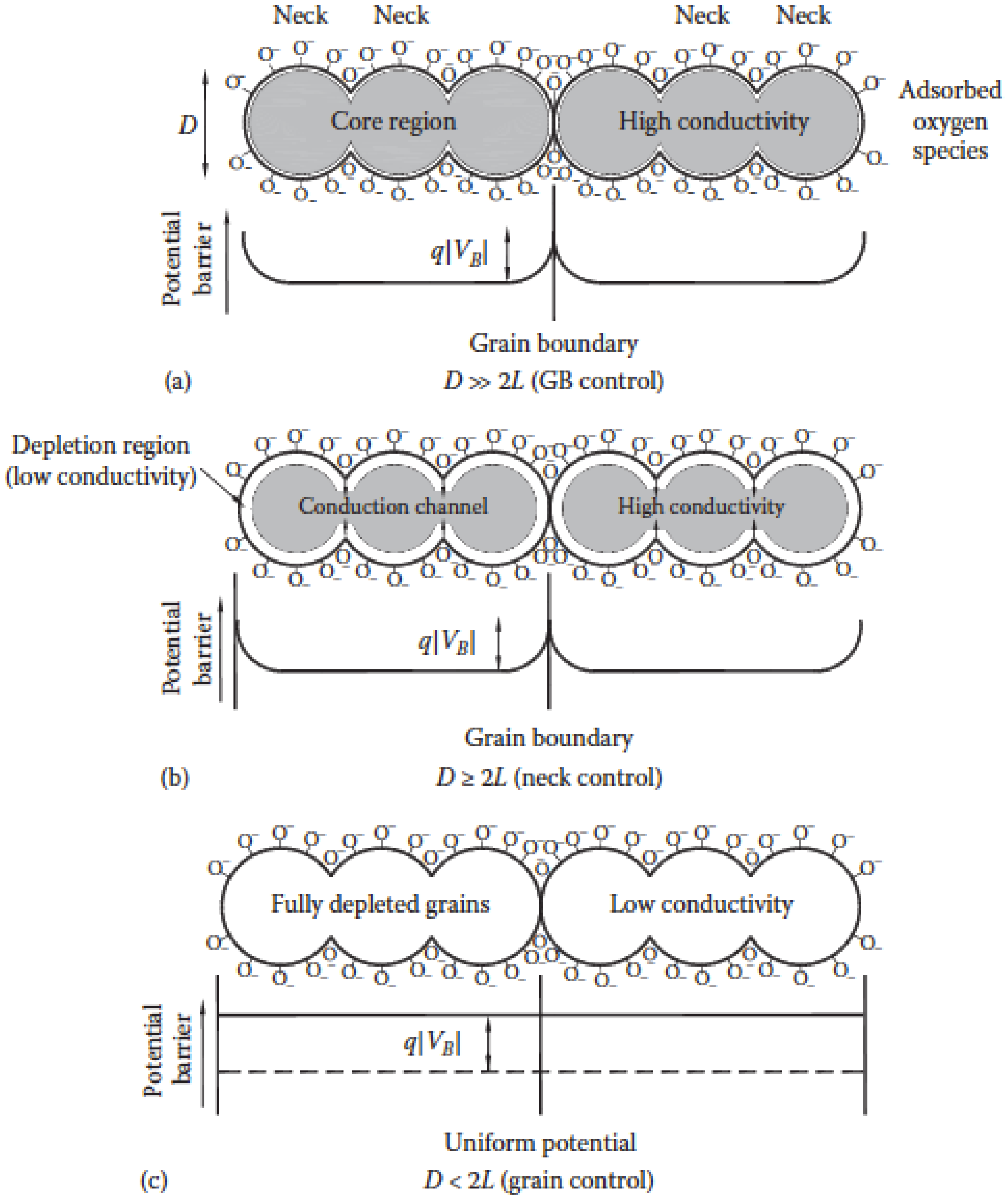


Figure 2.18 Schematic model of the effect of the crystallite size on the sensitivity of metal-oxide gas sensors: (a) D >> 2L (grain-boundary controlled); (b) D ≥ 2L (neck controlled); and (c) D < 2L (grain controlled barriers are independent of the grain size and therefore the sensitivity is independent of D value.

As the grain size decreases, the depletion region extends deeper into the grains and consequently the core region, which is relatively conductive with respect to the depletion region adjacent to the surface, becomes smaller.

When D approaches but still larger than 2L , that is, when D ≥ 2L , the deplet within each aggregate, as depicted in Figure 2.18b. Consequently, the conductivity depends not only on the GB barriers but also on the cross-sectional area of those channels. This area is proportional to $(X-L)^2$, where X is the neck diameter, which is proportional to the grain size D ($X \approx 0.8D$ according to Xu et al. [48]). As a result, the conductivity is a function of the ratio X/L (or X/D). Since $L \propto N_t^-$ and N_t^- is modulated by the surface interactions with the gas phase, the effective cross-sectional area of the current path though the grains is sensitive to the ambient gas composition. The current constriction effect adds up to the effect of the GB barriers, and therefore the gas sensitivity is enhanced with respect to the former case (as shown $D >> 2L$, cf. Figure 5.4a). Furthermore, the sensitivity to gases becomes grain-size dependent and it increases when D decreases. Details of neck-controlled sensitivity are shown in Figure 2.18b.

When $D < 2L$ the depletion region extends throughout the whole grain and the crystallites are almost fully depleted of mobile charge carriers, as shown in Figure 2.18c. As a result, the conductivity decreases steeply since the conduction channels between the grains are vanished. The energy bands are nearly flat throughout the whole structure of the interconnected grains, and since there are no significant barriers for intercrystallite charge transport, the conductivity is essentially controlled by the intracrystallite conductivity (grain controlled). It was found empirically that the highest gas sensitivity is obtained in this case, that is, when $D < 2L$.

Since the electrical conductivity of the films is very crucial and systematic measurement of this important parameter reveals the internal grain property of the films under study. In this connection, a new conductivity sensor structure is defined by Hoel et al. [49] to measure the conductivity change in nanoparticle thin films, as shown in Figure 2.19. Figure 2.19a and b shows two different arrangements for testing the sensor properties of WO_3 nanoparticle films. A “conventional” arrangement is shown in Figure 2.19a, where current passes though the nanoparticle film and the mean resistance change is measured as the layer is exposed to the probing gas. Figure 2.19b illustrates a recently suggested new type of arrangement for nanostructured sensing films [50] for the same purpose. In this case, current is driven through a thin conducting layer, such as gold in the present case, covered by sensor-type nanoparticles. Resistivity fluctuations of the conducting layer are measured as the device is exposed to the test gas. Sensitivity and other parameters are to be studied carefully in this case and this approach is expected to provide better information on nanoparticles.

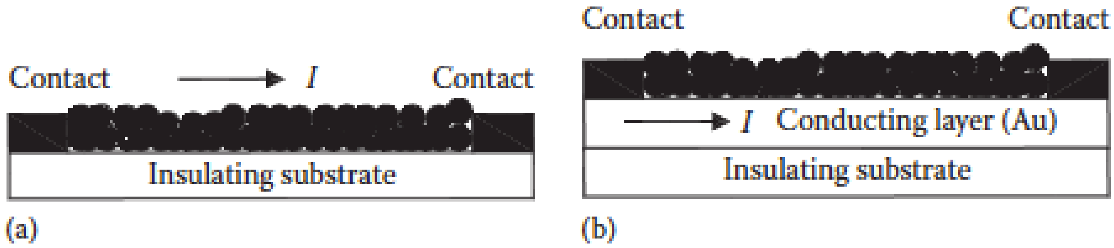


Figure 2.19 Sensor constructions based on nanoparticles. (a) Shows a “conventional” sensor, detecting conductivity change of the nanoparticle film as current I flows through it; (b) illustrates an arrangement to measure conductivity fluctuations of a thin conducting (gold) layer covered by nanoparticle

Although the reduction in the size of the sensing element and/or the transducer in a sensor are important, in order to miniaturize the devices, there are many issues in nanoscience one has to deal with for the nanosensor device development. New effects appear and play an important role that is often related to quantum mechanics and quantum mechanisms. Consequently, important characteristics and quality parameters of the nanosensors can be improved over the case of classically modeled systems by merely reducing the size. For example, sensitivity can increase due to better conduction properties [51]. The limits of detection can be lowered further and very small quantities of samples are possible to analyze. Direct detection is also possible, in this case, without using any labels, and some reagents can be eliminated.

2.8 Nanostructured metal oxide fabrication

This section will try to provide a brief outline of the methods commonly use for the fabrication of nanostructured metal oxides. There are conventional methods for fabrication of nanostructured semiconductor sensors- thick and thin film production methods. Some unconventional metal oxide nanostructures like nanobelts, nanowires, nanodiskettes will also be discussed along with the general approaches for their synthesis.

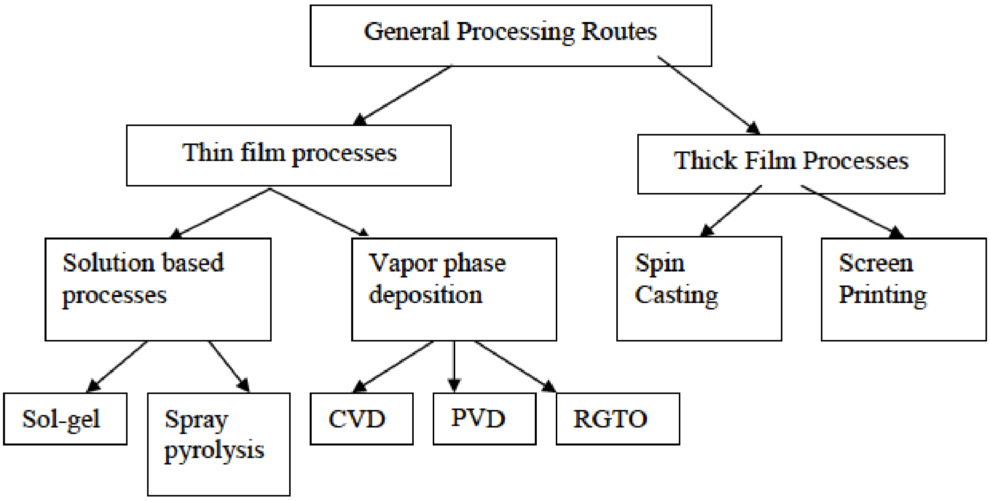


Figure 2.20. Common processing methods for nanostructured materials.

2.8.1 Conventional methods

Thin Film Technologies

I. Vapor Deposition-Nanostructure Synthesis Using PVD and CVD A variety of oxides such as ZnO, Ga_2O_3, In_2O_3, CdO, PbO_2 have been fabricated using a solid vapor process. It involves vaporizing a powder source material at an elevated temperature, and the resultant vapor condenses to formthe desired product under specific operating conditions (temperature, pressure, atmosphere, substrate etc.).

Chemical Vapor Deposition (CVD) involves exposing a substrate of choice to a mixture of volatile precursors that react and/or decompose on the substrate to give the desired product. A wide variety of CVD techniques that are in use are as follows (Park et al., 2003):

(1) Atmospheric pressure CVD (APCVD)

(2) Atomic layer CVD (ALCVD)

(3) Low pressure CVD (LPCVD)

(4) Metal organic CVD (MOCVD)
(5) Microwave plasma assisted CVD (MPCVD)
(6) Plasma enhanced CVD (PECVD)
(7) Metal organic MBE (ultra high vacuum MOCVD)

Physical vapor deposition (PVD) uses physical means rather as opposed to chemical vapor deposition techniques. The various techniques are:

(1) Sputtering
(2) Evaporation
(3) Pulsed LASER deposition

Both PVD and CVD offer enormous amount of control over the film thickness, stoichiometry and microstructure. Hence these films usually have highly controlled properties. But the only disadvantage is that these film deposition techniques are slightly expensive.

Specifically, a vacuum deposition method such as evaporation, plasma assisted chemical vapor deposition, or a sputtering method may be used for forming the thin film metal oxide sensors. In the sputtering method, predominantly neutral atomic ormolecular species are ejected froma target, which may be formed from the material to be deposited, under the bombardment of inert gas positive ions, e.g., argon ions. The high-energy species ejected will travel considerable distances to be deposited on the substrate held in a medium vacuum, e.g. 10^{-4} to 10^{-2} mbar. The positive ions required for bombardment may be generated in a glow discharge where the sputtering target serves as the cathode electrode to the glow discharge system. The negative potential (with respect to ground and the glow discharge) is maintained in the case of insulating target materials by the use of radio frequency power applied to the cathode, which maintains the target surface at a negative potential throughout the process. DC power may be applied when the target is an electrically conducting material. The advantage of such techniques is that control of the target material is greatly enhanced, and the energy of the species ejected is very much higher than with evaporation methods e.g. typically 1 to 10 eV for sputtering as compared with 0.1 to 0.5 eV for evaporation methods.

CVD	PVD	
	Sputtering	Evaporation
• Thermal CVD • Plasma enhanced CVD • Laser induced CVD • Electroless plating • Spray Pyrolysis • Melt dipping • Liquid quenching • Deposition of organic polymers and emulsions	• Reactive sputtering • Cathode sputtering • IBD • Ionized cluster beam • Plasma decomposition	• MBE • Thermal evaporation • Reactive evaporation • Ion plating • Arc evaporation • LASER evaporation

In magnetron sputtering processes, the plasma is concentrated immediately in front of the cathode (target) by means of a magnetic field. The effect of the magnetic field on the gas discharge is dramatic. In that area of discharge where permanent magnets, usually installed behind the cathode, create a sufficiently strong magnetic field vertically to the electric field, secondary electrons resulting from the sputter bombardment process will be deflected by means of the Lorenz force into circular or helical paths. There is an increase in plasma density and a considerable increase in deposition rate.

Bias sputtering (or sputter ion plating) may be employed as a variation of this technique. In this case, the substrate is held at a negative potential relative to the chamber and plasma. The bombardment of the substrate by Argon ions results in highly cleaned surfaces. Sputtering of the target material onto the substrate throughout this process results in a simultaneous deposition/cleaning mechanism.

This has the advantage that the interfacial bonding is considerably improved. In sputter ion-plating systems the substrate is held at a negative potential. In this case, the relative potentials are balanced to promote preferential sputtering of the target material. The target voltage will be typically less than 1 kV, dependant on system design and target material. The substrate may be immersed in its own localized plasma dependent upon its bias potential, which will be lower than that of the target. The exact voltage/power relationship achieved at either target or substrate depends upon many variables and will differ in detail from system to system.

Several heating methods exist, e.g., resistive, inductive, electron beam impingement etc., although the commonly preferred method is an ion beam source where a beam of ions impinge onto the coating material contained in a water cooled crucible. The use of multi-pot crucibles or twin source guns, enable multiple layers and graded stoichiometry layers to be deposited with the aid of electronic monitoring and control equipment.

In ion plating, negative bias applied to the substrate in an inert gas promotes simultaneous cleaning/deposition mechanisms for optimizing adhesion, as described in the sputtering process. Bias levels of -2 kV are typically used but these can be reduced to suit substrates. As operating pressures are higher in the ion plating technique, e.g. 10^{-3} to 10^{-2} mbar, gas scattering results in a more even coating distribution. To protect the filament the electron beam gun in the ion plating technique is differentially pumped to maintain vacuum higher than 10^{-4} mbar.

In the plasma assisted chemical vapor deposition (PACVD) method, the substrate to be coated is immersed in a low pressure (0.1 to 10 Torr) plasma of the appropriate gases/volatile compounds. This pressure may be maintained by balancing the total gas flow-rate against the throughput of the pumping system. The plasma may be electrically activated and sustained by coupling the energy from a power generator through a matching network into the gas medium. Thin films have been successfully deposited from direct current and higher frequency plasmas well into the microwave range. At high frequencies, the energy may be inductively coupled depending on chamber design and electrode configuration. Typically a 13.56 MHz radio-frequency generator would be used having a rating thatwould allowa power density of between about 0.1W/cm^2 . Typical deposition rates for this technique can be favorably compared with those obtained by sputtering. The deposition of the thin film metal oxide may be achieved by immersing a substrate in plasma containing a metal compound, such as molybdenum or tungsten, and oxygen under appropriate processing conditions.

In addition, suitable thin film metal oxides may be formed by a sol-gel deposition method, a plasma ashing method, or a solution coating method.

II. Sol-gel method Sol-gel method (Fig. 2.10) has been used for a long time for the production of nanomaterials (Brinker and Scherrer, 1984). This is a room or slightly elevated temperature process. The process involves the hydrolysis of a metal organic compound such as a metal alkoxide (usually, or can be hexachlorides as well), or inorganic salts such as chlorides (Shieh et al. , 2002) to produce a colloidal sol. The hydrolysis can take place with the help of alcohol, acid or base. The sol is then allowed to age and settle. This is referred to as the gelation step. The inorganic oxide gel is converted to an inorganic oxide glass by a low temperature heat treatment. The metal alkoxides, e.g., alkoxides of molybdenum or tungsten such as molybdenum isopropoxide or tungsten isopropoxide, may

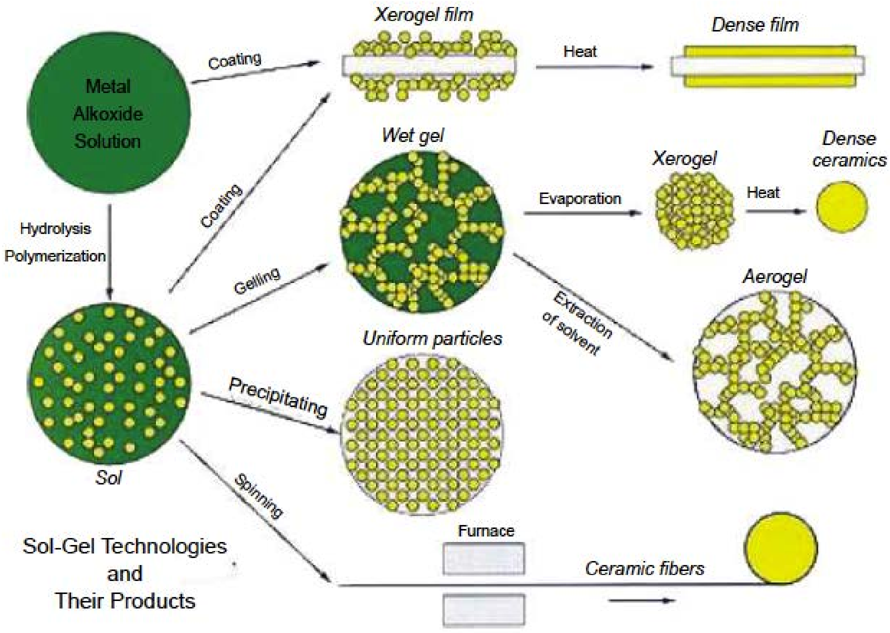


Figure 2.21. Schematic of the sol-gel processing technique used in combination with an alcohol, such as n-butanol, to form themetal trioxide, i.e., molybdenum trioxide or tungsten trioxide, respectively. Because these isopropoxides are reactive to atmosphere, they may be mixed in an enclosed container under a nitrogen atmosphere. The resulting sol is allowed to age and settle, and the sol may then be deposited on a substrate.

The gel coated substrate then undergoes suitable drying and firing stages to convert the coating into an inorganic oxide glass. The precise conditions with respect to temperature and residence time in the various stages of conversion are dependent upon the gel composition and its tolerance to relatively rapid changes in its environment. Porosity and integrity of the coating can be significantly affected by these stages. A suitable conversion process would include drawing the gel-coated substrate through drying ovens in which the temperature is controlled. The versatility of the process lies the in the flexibility available for the form of the end product such as

- The sol can be coated on the substrate by either spin/dip coating to form a ‘xerogel’ film
- The solvent from the sol can be evaporated to precipitate particles of uniform size and then these can be screen printed
- The sol can be allowed to gel completely to obtain either a xerogel or an aerogel.
  - The sol can be spun cast to form ceramic fibers

III. Spray Pyrolysis This process involves the atomization of a liquid precursor through a series of reactors, where the aerosol droplets undergo evaporation, solute condensation within the droplet, drying, thermolysis of the precipitate particle at higher temperature to form a microporous particle which then gets sintered to give a dense particle (Messing et al., 1993). The advantages of using a spray pyrolysis are as follows:

- The process makes use of the wide variety of available solution chemistries compartmentalizing the solution in to unique droplets, thereby retaining very good stoichiometry on the particle surface. This is particularly useful for the synthesis of single and mixed metal oxides
- A variety of particle morphologies can be obtained such as core-shell morphologies, porous particles for catalyst support, fibers, nanocomposites, quantum dots and hollow nanoparticles, to mention a few.

There are a variety of spray pyrolysis processes, and a few of them are aerosol thermolysis, flame spray pyrolysis, aerosol decomposition, spray roasting, and aerosol decomposition.

Thick Film Technologies

Screen printing involves mixing of the high purity oxide powder with organic deflocculant additives like terpineol or propadyol along with inorganic binders like ethyl silicate/ethyl cellulose (Park et al., 2003; Martinelli et al., 1999). The paste should have the correct rheological properties. In addition, adherence to the substrate and correct thermal shrinkage properties are necessary to obtain a good film. The advantages of using screen printing are that it is a very efficient way to produce cheap and robust chemical sensors. Also it offers excellent control over the obtainable thicknesses.

2.9 Use of Catalysts

One of the most popular methods to improve the selectivity and sensitivity of the sensor is by the use of catalysts. There are two possible methods by which the catalystsmay improve the sensitivity by modifying the intergranular contact resistance, namely:

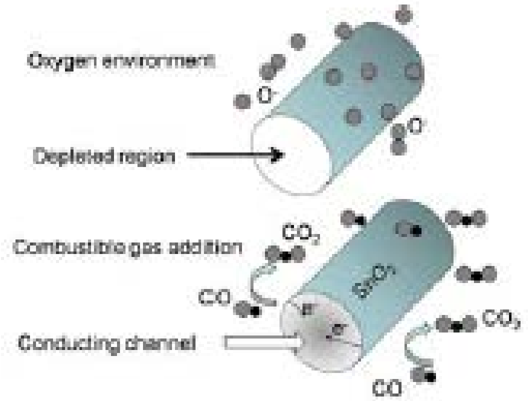


Figure 2.22. Sensing mechanism in a SnO_2 nanowire depicting the complete depletion of charge carriers inside the nanowire

- Chemical Sensitization — The catalyst particles on the surface of the support aid in the dissociation of the reducing/oxidizing gas thereby resulting in the “spillover” of the reaction products onto the surface of the support (The term support in catalytic literature refers to the ionic solid on which the catalyst is deposited. In our case it refers to the semiconductor sensor). The spilled over products react in the usual way (Morrison, 1987).

- Electronic sensitization — The catalyst is involved in electronic exchange with the gas phase and in turn with the support surface, thereby modifying the Fermi energy and the barrier height of the support surface (Morrison, 1987). In both cases a very good dispersion of the catalyst is necessary, in order to obtain improved sensitivity and selectivity.

2.10 1D and 2D Metal Oxides

Now that we have gathered an idea about the sensing mechanism involved in these one-dimensional nanostructures, let us look at some specific metal oxides like ZnO, SnO_2, TiO_2, In_2O_3, WO_3 and MoO_3 as case studies. The case studies presented below are very comprehensive and include all reports of 1D nano metal oxides that have been used for gas sensing applications

Zinc Oxide

Among the binary oxides that are to be discussed in this section, ZnO is probably the only metal oxide that possesses a rich family of one dimensional nanostructures — nanowires, nanobelts, nanorings, nanobows, nanocombs and a variety of other novel and unique architectures (Wang, 2004; Tian et al., 2003). Zinc oxide belongs to the wurtzite family and is the most ionic among the oxides crystallizing in this structure (Heinrich, 1985). It has zinc and oxygen atoms in a tetrahedral coordination with the layers of anions and cations stacked alternatively one on top of the other. Thus anisotropic growth is favored in several directions, including <2110>, <1010>, <0001>. Thus in zinc oxide formation of such one or twodimensional structures with the above planes as the exposure faces is energetically favored. It is a wide band gap (3.37eV) semiconductor with a high binding energy of ~ 60 meV. Due to its non-central symmetry, it exhibits piezoelectricity that makes it viable for using in actuators and transducers (Wang and Song, 2006). ZnO also exhibits near ultraviolet emission (Cross et al., 2005) and transparent conductivity (Banerjee and Chattopadhyay, 2005). With this wide range of functional properties and the ability to be morphologically tailored to unique and novel shapes, ZnO is probably one of a kind.

The nanostructures will be classified based on the route of fabrication; for example ZnO nanostructures that have been grown from the vapor phase either by the vapor-liquid-solid (VLS) growth or vapor-solid growth (VS), those that have been synthesized using the chemical methods such as sol-gel, hydrothermal methods and so forth. The influence of oxide and metallic dopants on the gas sensing properties will also be touched upon. Wherever applicable, the sensing mechanism will be detailed.

a. Growth from the Vapor Phase. This route of fabrication is probably the most commonly employed for the fabrication of one dimensional nanostructures, as shown in Fig. 2.23. Usually growth from the vapor phase occurs via either vapor solid or vapor-liquid solid route. It involves the thermal evaporation of a precursor material, either themetal oxide or themetal (inwhich case evaporation is followed by subsequent oxidation), in the presence of a carrier gas, with or without the influence of pressure. The carrier gas serves the purpose of transporting the species in the vapor phase to the substrate of choice and may or may not aid in oxidation. With control over the supersaturation factor high yields of 1-D nanostructures can be obtained. Probably the first report on the gas sensing properties of ZnO nanowires was by Wan et al. (2004;a,b). In this study the ZnO nanowires were prepared by evaporating Zn pellets at 900◦ C in an argon ambient, with a cross flow of oxygen.

It was found that in this case controlling the oxygen concentration was the key to achieving massive quantities of ZnO. The nanowire sensors were evaluated for their sensitivity towards ethanol up to 1 ppm. The sensitivity at 1 ppm ethanol at an operating temperature of 300◦ C was 1.9 and increased with increase in ethanol concentration with a maximum sensitivity of about 47 at 200 ppm of ethanol. The mechanism of gas sensing was explained by means of the formation of oxygen traps (O^{-2} ion) in ambient due to chemisorption and a resultant increase in surface resistance. On exposure to a reducing gas the electrons are returned to the surface thereby reducing the surface resistance. The high levels of sensitivity are attributed to the dimensions of the nanowire (～25 nm in diameter), since most of the atoms are now surface atoms. Till date this explanation is considered to be the closest approximation to what really happens on the surface of a nanowire and empirical studies are also being carried out simultaneously (Comini et al. , 2004).

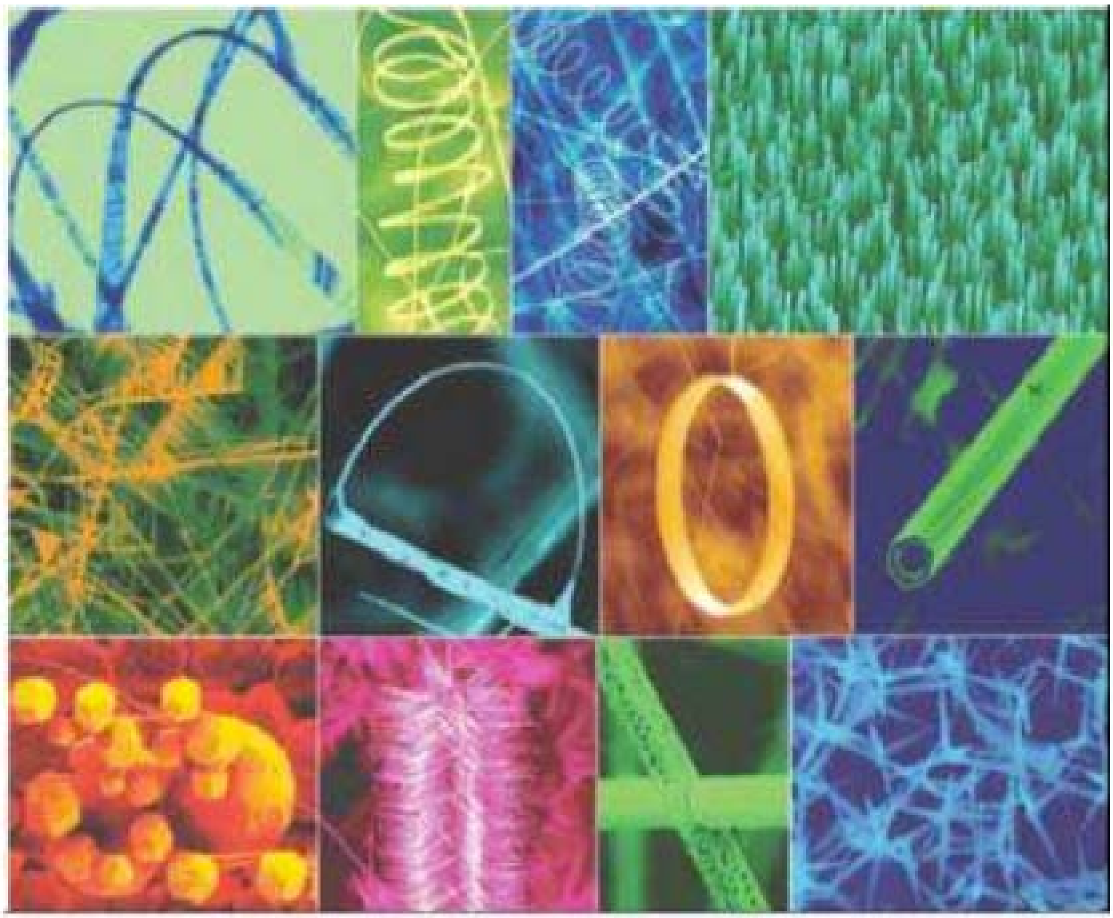


Figure 2.23. ZnO nanostructures synthesized under controlled conditions by thermal evaporation of solid powders

Tin Oxide Nanostructures

Tin dioxide finds use in three major applications (a) as a transparent conductor — has high electrical conductivity and transparency in the visible range of the electromagnetic spectrum (b) as a catalyst in oxidation (c) as a solid state gas sensor. The latter two applications are controlled by the properties of the surface while transparency is more so a bulk property. SnO_2 has a band gap of 3.6 eV and is most commonly found in the rutile crystal structure although a high pressure orthorhombic variant is also known (Suito et al., 1975). Next to zinc oxide, tin oxide is another metal oxide that has been widely synthesized in to nanowires, nanobelts, nanorods, nanoribbons, nanodiskettes and a range of other one dimensional nanostructures (Z.L Wang’s group, 2002–2003). Again as in the section on zinc oxide, attention will be only on those one-dimensional nanostructures that have been explored so far for gas sensing applications and the structures will be discussed based on the synthesis route. Figure 2.24 contains some electron microscopy images of the configurations of SnO_2 one dimensional nanostructures synthesized so far.

The most common method of producing one dimensional nanostructure of tin oxide also is growth from the vapor phase such as thermally evaporating a tin (Chen et al., 2006; Ying et al. 2004;2005) or a tin oxide precursor (Comini et al., 2002;2005) and subsequently condensing from the vapor phase. SnO_2 is known to have very good sensing properties towards reducing gases and in particular ethanol. So almost all these reports are on the ethanol sensing characteristics of ethanol except for a few where the other gases tested have been CO, NO2 (Comini et al., 2002; Ramgir et al., 2005).

Dopants can also be incorporated into the nanostructures by co-evaporating or co-depositing the dopant sources during the synthesis process. CdS, Ru, Sb, CuO, Pd have been doped with SnO_2 nanobelts for improving the specificity towards a particular gas. CdS and Sb doped nanobelt heterostructures have been analyzed for their sensitivity towards ethanol and itwas found that the doped sensors have a better sensitivity towards ethanol than the undoped ones. Ru has been envisaged as a dopant for a room temperature NO2 and LPG nanobelts SnO_2 sensor.

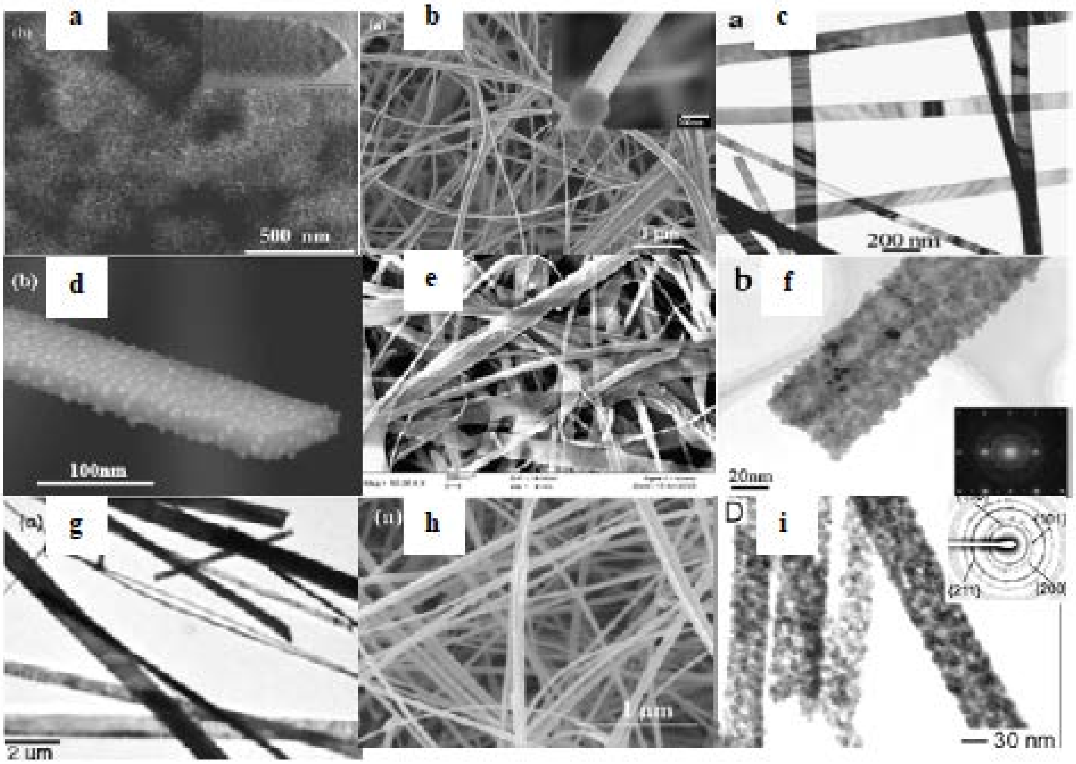


Figure 2.24. Some common morphologies of 1D SnO2nanostructures

CuO on the other hand has been used as a dopant for selective room temperature H_2S sensing, since it is known to form a p-n junction with SnO_2 , and the strong affinity of CuO to H_2S breaks up this p-n junction. Extremely high sensitivities have been achieved with CuO doping. The Pd decorated nanowires were seen to respond to pulses of 2 while the pristine 2 nanowires did not show any response to hydrogen.

Growth of nanostructures using templates has also been explored as a viable method for generating array of nanowires. For instance, porous aluminum oxide (PAO) and cellulose have been tried for templating SnO_2 nanowires (Kolmakov et al. , 2003). The method discussed in the latter involves growing metallic Sn nanowires in the pores of the alumina template, which is subsequently etched away. The Sn nanowires are then topotactically converted in to their oxide by controlled thermal oxidation. The nanowires grown are solid. On the other hand the nanostructures produced as discussed in (Hamaguchi et al. , 2006) are hollow and are the negative replicas of the template. Similarly for the template generated nanostructures the sensing properties are primarily investigated for reducing gases like ethanol, CO, H_2 and ethylene oxide.

Chemical synthetic methods such as hydrothermal route, aqueous synthesishave also been used for synthesis of these nanostructures (Wang et al. , 2003; Chen et al. , 2005). The chemical methods are usually carried out at low temperatures. These 1D structures have ethanol, CO and H_2 .

Indium Oxide

Indium trioxide is a transparent semiconductor that finds use in optoelectronic applications (Ishibashi,1990; Chopra , 1983), and also in flat panel displays due to its high transparency and electrical conductivity (Jarzebski, 1982). Also it interestingly exhibits insulator-superconductor transition at low temperatures (Grantmakher et al. , 1998). Apart from this interesting array of properties In_2O_3 also is n-type semiconductor (owing to its oxygen vacancy doping (Bellingham et al., 1991) that has been used as both nanostructured thin films and one-dimensional nanostructure for resistive gas sensing (Gurlo et al. , 1997;a;b)especially for oxidizing gases like NO_2, O_3. It also serves as a dopant in sensing films for improving the selectivity of sensors (Francioso et al. , 2006).

One dimensional In2 O3 nanostructures, see Fig. 2.25, for gas sensing applications have been primarily produced by laser ablation method using Au as catalyst on a Si/SiO_2 substrate. The In_2 O_3 nanostructures so produced were found to have a cubic crystal structure with [1 1 0] as their primary growth direction (Li et al. , 2003). They have been primarily examined for their gas sensing properties towards NO_2 , NH_3 . Other VOCs tested include butylamine and ethanol for sensing using In_2 O_3 .

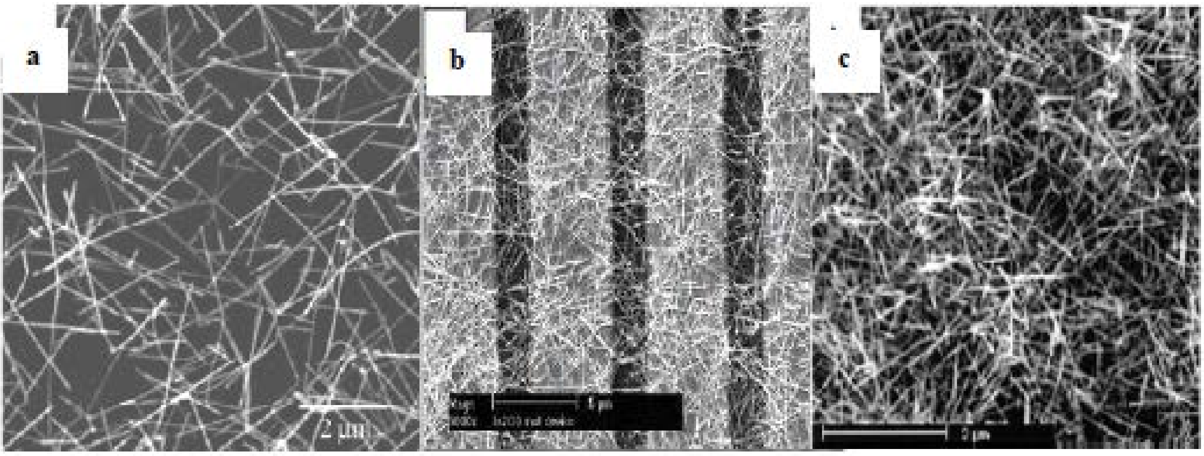


Figure 2.17. One-dimensional In_2O_3 synthesized by the laser ablation method

2.11 Outlook

Nanomaterials have the potential for revolutionizing the sensing field, with the development of one-dimensional nanostructures that provide the maximum surface area to volume ratio. Also the unconventional nanoassemblies will result in the development of miniaturized sensor arrays to be used in electronic noses aiding in the simultaneous detection of multiple analytes. The future lies in the successful integration of these nanoassemblies and arrays with electronics for applications ranging from molecular recognition to biosensing.

References

16. L. Satyanarayana, K.M. Reddy, and S.V. Manorama, Synthesis of nanocrystalline Ni1−xCoxMnxFe2−xO4: A material for liquefied petroleum gas sensing, Sensors and Actuators, B89, 62–67, 2003.

17. A. Gurlo, N. B.rsan, M. Ivanovskaya, U. Weimar, and W. G.pel, In2O3 and MoO3–In2O3 thin film semiconductor sensors: Interaction with NO2 and O3, Sensors and Actuators, B47, 92–99, 1998, and the references therein.

18. J.G. Lu, P. Chang, and Z. Fan, Quasi-one-dimensional metal oxide materials—Synthesis, properties and applications, Materials Science and Engineering, R52, 49–91, 2006, and the references therein.

19. E. Comini, G. Faglia, G. Sberveglieri, Z. Pan, and Z.L. Wang, Stable and highly

sensitive gas sensors based on semiconducting oxide nanobelts, Applied Physics Letters, 81, 1869–1871, 2002.
20. G. Bl.ser, Th. Rühl, C. Diehl, M. Ulrich, and D. Kohl, Nanostructured semiconductor gas sensors to overcome sensitivity limitations due to percolation effects, Physica A: Statistical Mechanics and Its Applications, 266, 218–223, 1999, and the references therein.
21. K.-D. Schierbaum, Engineering of oxide surfaces and metal/oxide interfaces for chemical sensors: Recent trends, Sensors and Actuators, B24–B25, 239–247, 1995, and the references therein.
22. J. Santos, P. Serrini, B. O'Beirn, and L. Manes, A thin film SnO2 gas sensor selective to ultra-low NO2 concentrations in air, Sensors and Actuators, B43, 154–160, 1997.
23. P.M. Hall, Resistance calculations for thin film patterns, Thin Solid Films, 1, 277–295, 1968.
24. G. Sberveglieri, Recent developments in semiconducting thin-film gas sensors, Sensors and Actuators, B23, 103–109, 1995, and the references therein.
25. A. Di.guez, A. Vil., A. Cabot, A. Romano-Rodr.guez, J.R. Morante, J. Kappler, N. B.rsan, U. Weimar, and W. G.pel, Influence on the gas sensor performances of the metal chemical states introduced by impregnation of calcinated SnO2 sol–gel nanocrystals, Sensors and Actuators, B68, 94–99, 2000.
26. F.H. Garzon, R. Mukundan, and E.L. Brosha, Solid-state mixed potential gas sensors: Theory, experiments and challenges, Solid State Ionics, 136–137, 633–638, 2000.
27. P. Keller, H. Ferkel, K. Zweiacker, J. Naser, J.-U. Meyer, and W. Riehemann, The application of nanocrystalline BaTiO3-composite films as CO2-sensing layers, Sensors and Actuators, B57, 39–46, 1999.
28. C. Cantalini, W. Wlodarski, H.T. Sun, M.Z. Atashbar, M. Passacantando, and S. Santucci, NO2 response of In2O3 thin film gas sensors prepared by sol–gel and vacuum thermal evaporation techniques, Sensors and Actuators, B65, 101–104, 2000.
29. G. Sberveglieri, E. Comini, G. Faglia, M.Z. Atashbar, and W. Wlodarski, Titanium dioxide thin films prepared for alcohol microsensor applications, Sensors and Actuators, B66, 139–141, 2000.
30. O.K. Varghese, D. Gong, M. Paulose, K.G. Ong, and C.A. Grimes, Hydrogen sensing using titania nanotubes, Sensors and Actuators, B93, 338–344, 2003.

31. Q. Pan, J. Xu, X. Dong, and J. Zhang, Gas-sensitive properties of nanometer-sized SnO2, Sensors and Actuators, B66, 237–239, 2000.

32. Y.-L. Liu, Y. Xing, H.-F. Yang, Z.-M. Liu, Y. Yang, G.-L. Shen, and R.-Q. Yu, Ethanol gas sensing properties of nano-crystalline cadmium stannate thick films doped with Pt, Analytica Chimica Acta, 527, 21–26, 2004.

33. A. Khanna, R. Kumar, and S.S. Bhatti, CuO-doped SnO2 thin films as hydrogen sulfide gas sensor, Applied Physics Letters, 82, 4388–4390, 2003.

34. Q. Wan, Q.H. Li, Y.J. Chen, T.H. Wang, X.L. He, J.P. Li, and C.L. Lin, Fabrication and ethanol sensing characteristics of ZnO nanowire gas sensors, Applied Physics Letters, 84, 3654–3656, 2004.

35. M. Niederberger, G. Garnweitner, N. Pinna, and G. Neri, Non-aqueous routes to crystalline metal oxide nanoparticles: Formation mechanisms and applications, Progress in Solid State Chemistry, 33, 59–70, 2005, and the references therein.

36. M. Graf, D. Barrettino, M. Zimmermann, A. Hierlemann, H. Baltes, S. Hahn, N. B.rsan, and U. Weimar, CMOS monolithic metal-oxide sensor system comprising a microhotplate and associated circuitry, IEEE Sensors Journal, 4, 9–16, 2004.

37. M.Z. Atashbar, D. Banerji, and S. Singamaneni, Room-temperature hydrogen sensor based on palladium nanowires, IEEE Sensors Journal, 5, 792–797, 2005.

38. W. Zhou, Z. Shao, and W. Jin, Synthesis of nanocrystalline conducting composite oxides based on a non-ion selective combined complexing process for functional applications, Journals of Alloys and Compounds, 426, 368–374, 2006.

39. H. Ogawa, M. Nishikawa, and A. Abe, Hall measurement studies and an electrical conduction model of tin oxide ultrafine particle films, Journal of Applied Physics, 53, 4448–4455, 1982.

40. A.M. Taurino, S. Capone, P. Siciliano, T. Toccoli, A. Boschetti, L. Guerini, and S. Iannotta, Nanostructured TiO2 thin films prepared by supersonic beams and their application in a sensor array for the discrimination of VOC, Sensors and Actuators, B92, 292–302, 2003, and the references therein.

41. N. Yamazoe and K. Shimanoe, Theory of power laws for semiconductor gas sensors, Sensors and Actuators, B128, 566–573, 2008, and the references therein.

42. M.J. Madou and S.R. Morrison, Chemical Sensing with Solid State Devices, Academic Press, London, U.K., 1989.

43. E. Comini, Metal oxide nano-crystals for gas sensing, Analytica Chimica Acta, 568, 28–40, 2006, and the references therein.

44. O.K. Tan, W. Cao, Y. Hu, and W. Zhu, Nano-structured oxide semiconductor materials for gas-sensing applications, Ceramic International, 30, 1127–1133, 2004, and the references therein.

45. B.J. Kooi and J.Th.M.D. Hosson, In-situ TEM analysis of the reduction of nanometre-sized Mn3O4 precipitates in a metal matrix, Acta Materialia, 49, 765–774, 2001.

46. D.C. Meier, S. Semancik, B. Button, E. Strelcov, and A. Kolmakov, Coupling nanowire chemiresistors with MEMS microhotplate gas sensing platforms, Applied Physics Letters, 91, 063118, 2007.

47. A. Rothschild and Y. Komem, The effect of grain size on the sensitivity of nanocrystalline metal-oxide gas sensors, Journal of Applied Physics, 95, 6374–6380, 2004.

48. C. Xu, J. Tamaki, N. Miura, and N. Yamazoe, Grain size effects on gas sensitivity of porous SnO2-based elements, Sensors and Actuators, B3, 147–155, 1991.

49. A. Hoel, L.F. Reyes, P. Heszler, V. Lantto, and C.G. Granqvist, Nanomaterials for environmental applications: Novel WO3-based gas sensors made by advanced gas deposition, Current Applied Physics, 4, 547–553, 2004.

50. A. Hoel, J. Ederth, J. Kopniczky, P. Heszler, L.B. Kish, E. Olsson, and C.G. Granqvist, Conduction invasion noise in nanoparticle WO3/Au thin-film devices for gas sensing application, Smart Materials and Structures, 11, 640–644, 2002.

51. J. Riu, A. Maroto, and F.X. Rius, Nanosensors in environmental analysis, Talanta, 69, 288–301, 2006, and the references therein.

52. S. Iizuka, S. Ooka, A. Nakata, M. Mizuhata, and S. Deki, Development of fabrication process for metal oxide with nano-structure by the liquid-phase infiltration (LPI) method, Electrochimica Acta, 51, 802–808, 2005.

53. T. Maegawa, T. Yamauchi, T. Hara, H. Tsuchiya, and M. Ogawa, Strain effects on electronic bandstructures in nanoscaled silicon: From bulk to nanowire, IEEE Transactions on Electron Devices, 56, 553–559, 2009.

Chapter 3. Nano and porous silicon for gas, vapor and liquid sensing

In the present chapter, we pay attention both on technical and physico-chemical aspects of porous silicon (PS) sensors operation. The transducers and sensors of resistivity, capacitance, Schottky barrier, MIS, FET, EIS, ISFET, LAPS, luminescent and optical sensors with PS are considered. The set of gas sensor`s parameters for detection of humidity, CO, NO_2, different volatile organic compounds, alcohols, H_2S, H_2 and other gases are described.

3.1. Introduction

In present, the monitoring of the environment, recognition of the air, ground and water pollution, food control ensuring of the ecological security of all the countries around the world appears among the most important problems. Detection and observation of the toxic, inflammable or explosive gases are very important due to their harmful influence on human health. Reliable detection of the low level pollution for air quality monitoring is an important motivation to develop the systems of recognition like electronic nose. In the workplace, these include the detection of hazardous materials such as SO_2 and H_2S near petro-chemical plants, volatile organic compounds (VOCs) in coating operations, NO_x in heating, ventilation, and air conditioning equipment and others.

Nowadays, the development of the improved sensors of high sensitivity and selectivity are based on the achievements of modern semiconductor technique and the using of new materials. Application of PS, prepared by electrochemical etching of Si substrate, for sensors are considered in several reviews [1-7] focusing on the analysis of effects that occur in PS at gas adsorption.

In the present chapter, we pay attention both on technical and physico-chemical aspects of PS sensor operation.

3.2. RESISTIVITY (CONDUCTIVITY) SENSORS

Conductance changes in PS layers have been observed upon adsorption of molecules in pores (Fig.3.1), which presumably arise from dielectric-induced changes in the carrier concentration in the Si. Adsorbate-induced conductance changes may also arise from carrier concentration variation due to interaction with dangling bond states on PS surface [7, 8].

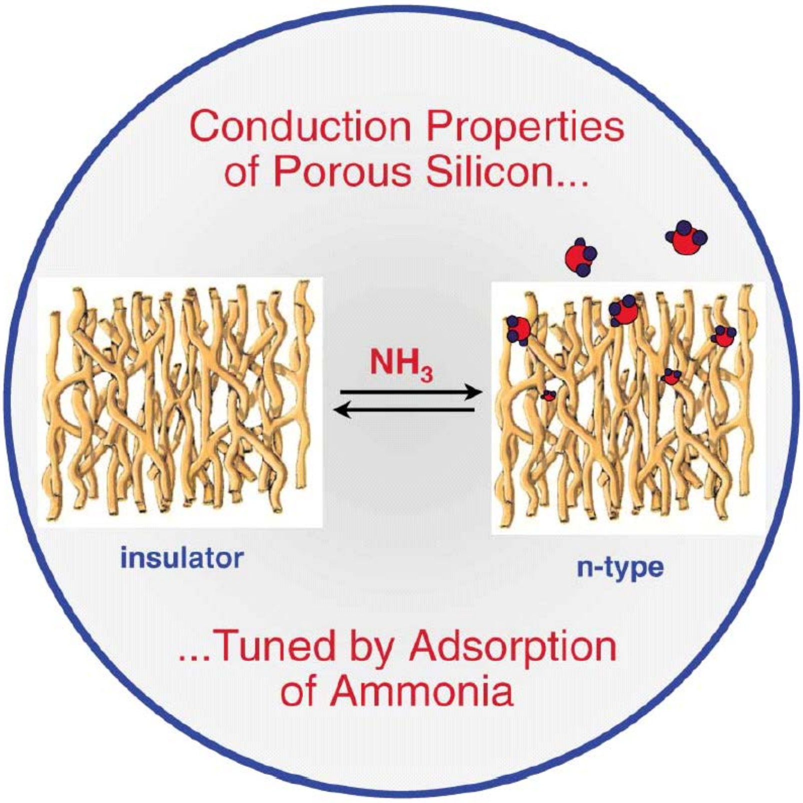


Fig.3.1 Molecules adsorption in pores

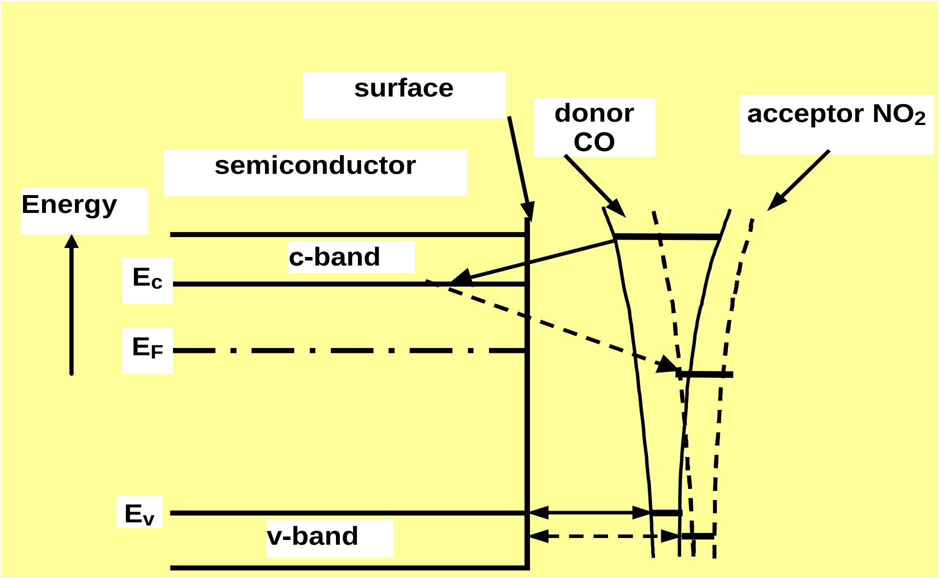


Fig.3.2 Model of donor and acceptor types of molecules.

The change of carrier concentration in bulk of porous semiconductor can arise due to donor or acceptor type of adsobed molecules (Fig.3.2). Typically, reducting agents like adsorbed hydrogen inject the electrons to bulk:

$\mathbf{H_2 \rightarrow H^+ + H^+ +2\,e^-}$

By contrary, the oxiding agents accept the electrons$^-$ from bulk:

$\mathbf{O_2 + 2\,e^- \rightarrow 2O^-}$

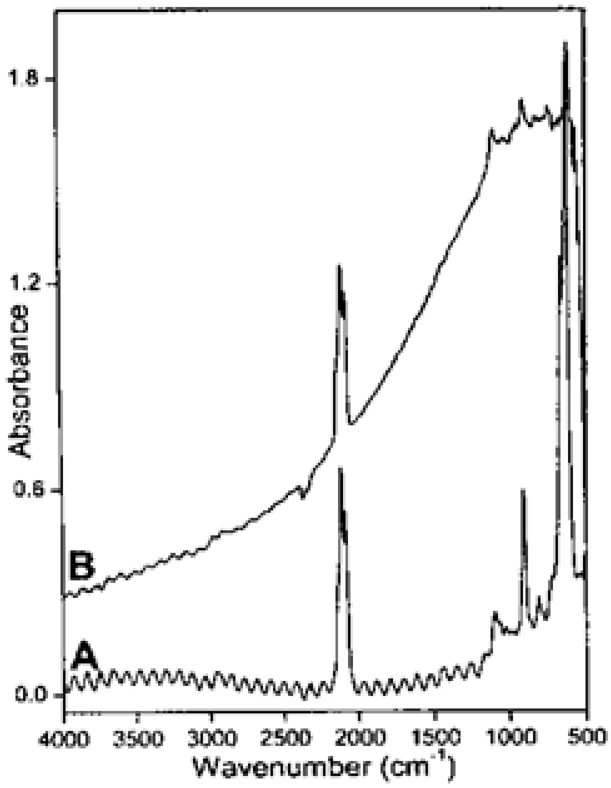


Fig.3.3 IR spectra of PS in vacuum (A) and in the NO_2 atmosphere.

These effects are clearly observed at the measurement of FTIR spectra of PS in different atmospheres. According Drude model, the absorption coeffiecient $\alpha \sim \lambda^2$ **.** Polarization effect is observed at the polar NO molecules adsorption into pores when the induced conductivity increases, i.e. adsorption of NO molecules injects the supplementary holes into high resistivity PS (Fig.3.3).

Besides, the effect of conductivity modulation can be conditioned by change the barrier value between micro (nano) crystalls as is indicted in fig.3.4.

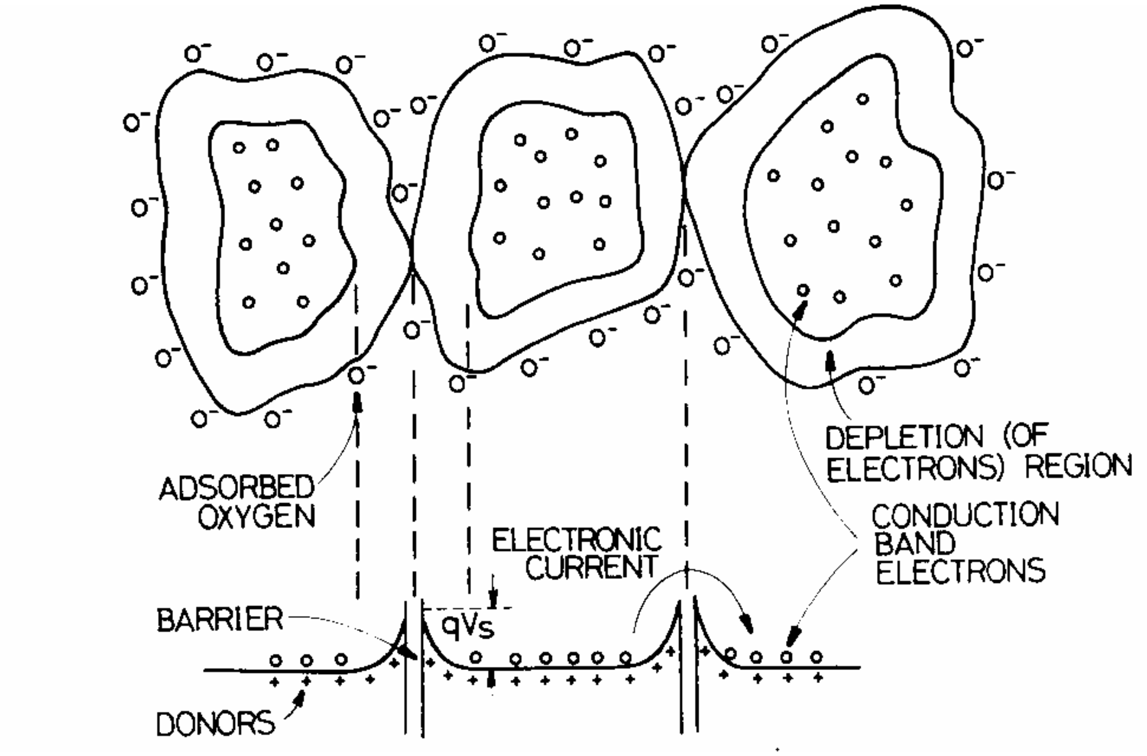


Fig.3.4.Adsorption of proton increases the potential barriers between microcrystals, adsorption of oxygen decreases these barriers.

Usually, *resistivity* (or *conductivity*) sensor is the simplest electrical transducer consisted of PS layer on Si or isolator substrate with, at least, 2 metal terminals (Figure 3.5). To reduce the high PS resistivity, sometimes, the metal terminals are formed as interdigitated electrodes [9, 10]. Supplementary discontinuous layer of catalyst that covers porous surface can be deposited.

Series resistivity sensors were proposed to detect low concentration of nitrogen dioxide (NO_2) [9, 11-18]. NO_2 is a toxic air pollutant emitted by combustion engines with attention level at 106 ppb and alarm level at 212 ppb [17]. The detectable level of NO_2 in air by PS resistivity sensor is well below 100 ppb.

Figure 3.6. shows the dynamic response of the sensor at T_{room} versus the concentration of NO_2 in dry air from 240 till 12 ppb. The sensor shows a very high sensitivity of conductivity: $\Delta G/G$=0.26 at 12 ppb, response and recovery time is about 10 min [17]. The adsorbed NO_2 molecules act as acceptor centers, at its adsorption on PS they inject free holes into PS film increasing the conductivity [12, 20]. The drawback of NO_2 sensors is the strong dependence of sensor current from

humidity (Figure 2, b). An increase of RH leads to a decrease in conductivity due to donor-like character of water molecules adsorbed as PS that lowering of the free holes concentration [20]. At that time the relative response of meso-PS (with and without gold film) to interfering species as CO (up to 1000 ppm), CH_4 (up to 5000 ppm) is negligible and anyhow low for alcohol such as methanol at concentrations up to 800 ppm [12]. The sensitivity of resistivity sensors is determined by the porous microstructure that depended on the composition of the electrochemical solution used for anodization, however, there is no direct correlation between porosity and sensitivity [18].

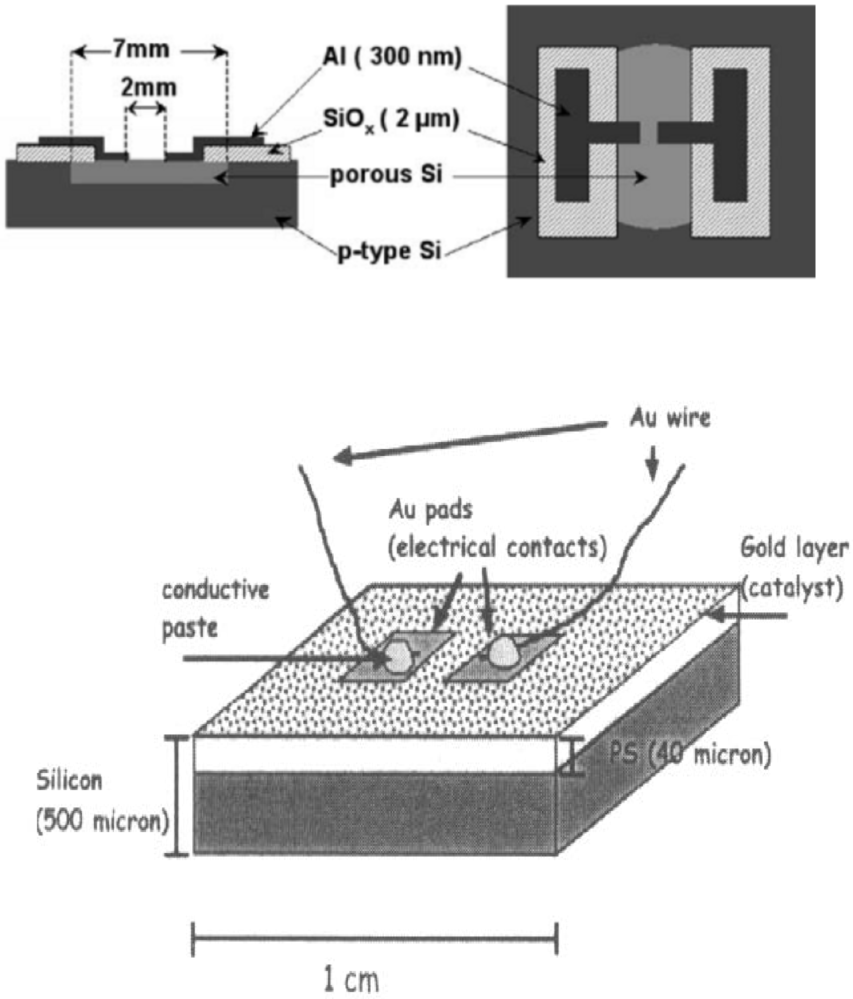
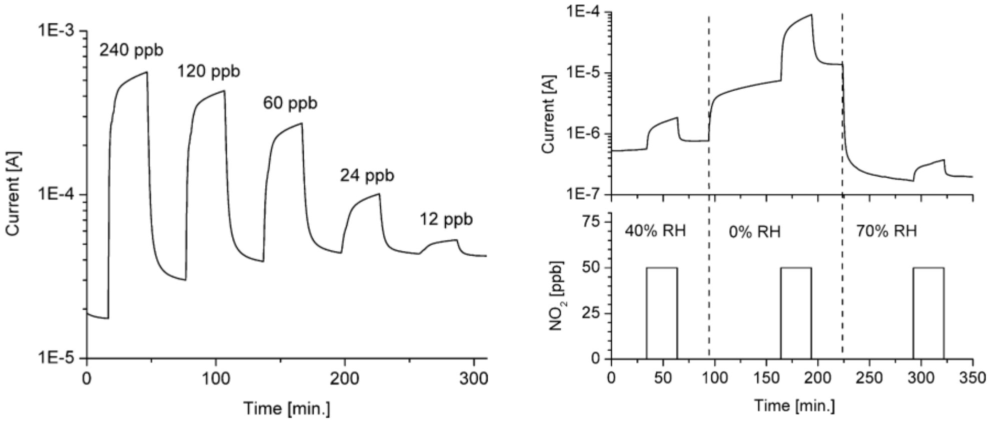


Figure 3.5. Sketch of the PS resistivity sensor.



(a) (b)

Figure 3.6. Dynamic response of the sensor to different concentration of NO_2 in dry air (a) and in the presence of different humidity levels and NO_2 concentrations (b). The graph above represents the sensor response; the graph below represents NO_2 concentrations as a function of time.

Table 3.1. Resistivity (conductivity) PS sensors

N	Si substrate	PS thickness, porosity	Metal contact, coating	Detected gas	Minimal concentration	Sensitivity ΔG/G or ΔR/R	Ref
1	p-Si, 3-13 Ω.cm	Free standing, 15 µm, 45%	interdigital Pt	NO_2 O_3 CO benzene	100 ppb 200 ppb 1000 ppm 20 ppm	13,8	[9]
2	p-Si, 10 Ω.cm	Free standing	planar Au	NO_2	70 ppb	0,39	[15]
3	p^+-Si, 6-15 mΩ.cm	32 µm, 80%	planar Au	NO_2	12 ppb in dry air, 50 ppb in humidity	11,7 1,2	[17], [14]
4	n-Si, 1 Ω.cm	40 µm, 45%	planar, nano Au	NO_2	5 ppm	1,6	[11]
5	p^+-Si, 5-15 mΩ.cm	20-30 µm, 30-75%	Ti-W-Pt	NO_2	3 ppm		[12]
6	p-Si, 12 Ω.cm	8-10 µm	metallophtalo-cyanine	NO_2	100 ppm	29,3%	[13]
7	p^+-Si, 5-15 mΩ.cm		planar Au	NO_2	100 ppb		[18]
8	p-Si, 1-20 Ω.cm		planar Au, nano Au/SnO_2 coating	CO SO_2 NH_3 NO_x	5 ppm 1 ppm 500 ppb 1 ppm		[16]
9	p-Si, 2-20 Ω.cm		interdigital Au	HCl HN_3 NO	100 ppm 100 ppm 100 ppm	0,03 %/ppm 0,1%/ppm 0,1%/ppm	[21]

10	n-Si, 1 Ω.cm	80 µm, 60%	planar Au	Methanol	500 ppm	1,75	[22]
				Ethanol	500 ppm	1,6	
				Chloroform	5100 ppm	0,6	
				Toluene	3000 ppm	2,8	
11	p-polySi		planar Au	ethanol	2000 ppm	19	[8]

The aged PS sensors show improved reversibility compared with fresh samples. The stabilization of the current is also faster in aged rather than in fresh samples. The electrical response dynamics and hysteresis can also be improved by means of a pre-treatment via prolonged exposure to high concentrations of NO_2 [15]. Selectivity and sensitivity of PS resistivity sensors can be enhanced using catalyst materials, for example, by electroless metal deposition to form a gold or tin oxide nanostructured framework interacting with the nanopore-coated microporous surface [16]. The application of PS resistivity sensors to detect several gases is summarized in table 1.

3.3. CAPACITANCE SENSORS

The operating principle of PS *capacitance* sensors is based on the increasing of the dielectric constant of porous materials (and, as results, on the increasing of capacity) after pore filling by solution or vapor of water or organic molecules (Fig.3.1) [23-25].

The sketch of typical PS capacitance sensor is presented in Figure 3.7, a. Silicon wafer (100) with 0.02 Ω.cm resistivity is used as substrate [24, 25]. This device was applied as humidity and ethanol vapor sensor in the range of 0-0.5% alcohol concentration. Alcohol concentration determining in the blood is important in order to prevent traffic accidents (if the alcohol concentration in a driver`s blood exceeds 0.05%, the driver may be punished for violation of drinking and driving regulations). The sensitivity of the capacitance curve measured at 120 Hz indicates an increase of 2.5 % capacitance per 0.1% alcohol solution concentration, while the sensitivity at larger frequency was observed to be less.

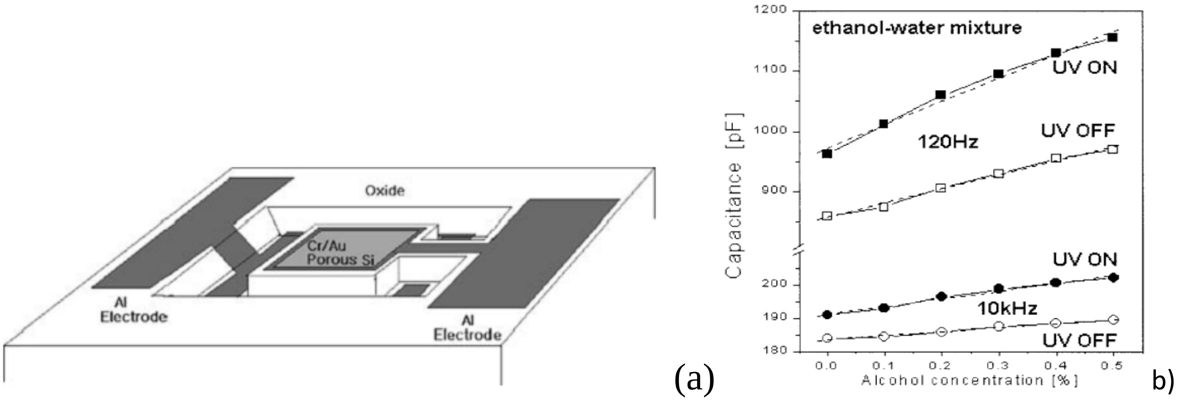


Figure 3.7 The top view (a), dependence of capacitance on alcohol concentration in ethanol-water mixture at 35 C(b).

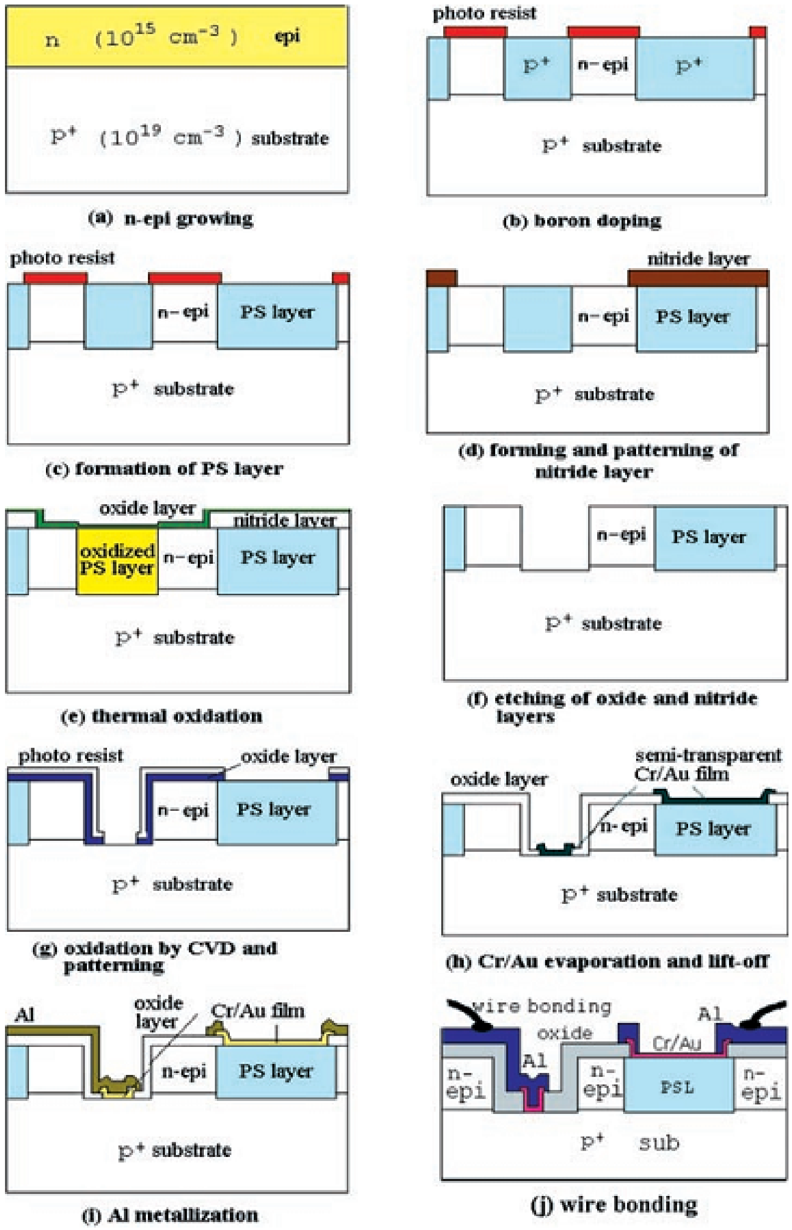


Fig.3.8 Process of fabrication of humidity sensor [24]

Under the UV light illumination (λ=254 nm) of sensor surface the increase in sensitivity and initial capacitance is observed. The sensor was examined in human exhaling gases composed typically of N_2, CO_2, water vapor without alcohol gas. N_2 and CO_2 gas concentrations included in exhaling breath had little effect on capacitance response (because the static dielectrical constant of N_2 and CO_2 gases is close to 1, similar to air in unfilled pores in PS). Process of fabrication of humidity sensor is presented in Fig.3.8.

Capacitance-type humidity sensors in which PS layer is used as a humidity-sensing material were developed also in [26-31] (Table 3.2). The capacitance in the PS layer will strongly depend on the relative humidity RH because the static permittivity of pure H_2O is 80, which is much more than the 12 of the Si. The better humidity sensitivity was obtained for capacitance sensors with hydrothermally-etched PS pillar array [28]. With the RH ranged from 11 to 95%, the variations of the capacitance reach 1500% at a signal frequency of 100 Hz and 800% at 1000 Hz (Figure 3.9). About 15 and 5 s, respectively, are needed for the capacitance to reach 90% of its final/initial values during a RH-increasing process and a RH-decreasing process. The faster response to humidity of PS pillar array sensor might be due to the regular morphology and suitable thickness of the sensing layer.

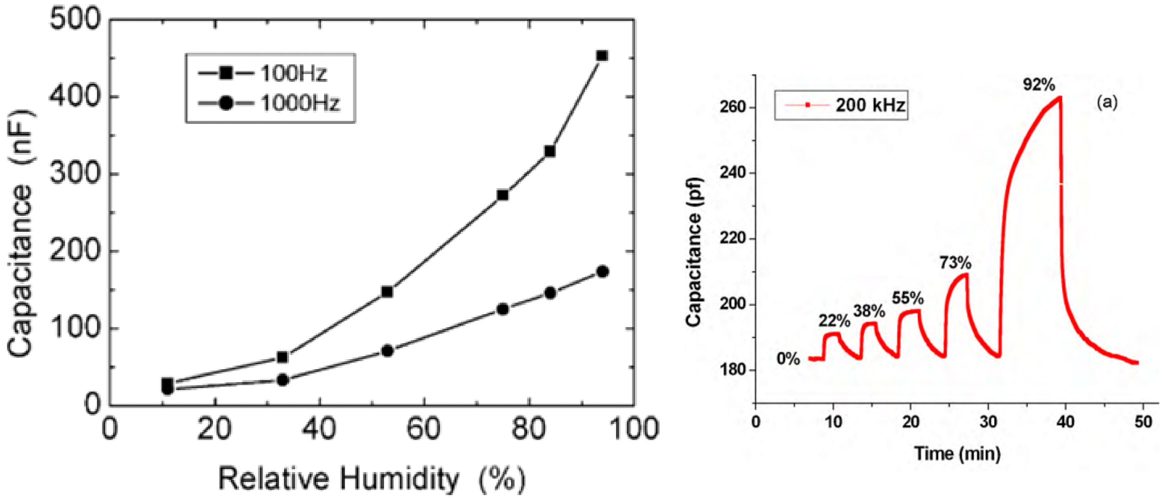


Figure 3.9. Variation of capacitance with RH at two different electrical signal frequencies at T_{room} and dynamic response of the sensor to different RH% levels under 200 kHz

The capacitance sensor with improved humidity sensitivity based on ordered macro-PS with a Ta_2O_5 thin film coating is proposed [31]. The sensor's capacity versus RH shows perfect linearity in two regions respectively over the whole RH range. The sensor shows the fastest response time at 200 kHz: 18–40 s to small RH changes, and 300 s to 100% RH changes, the hysteresis at 200 kHz is 0.3–3.3%.

Other approach that allows to develop high sensitive, stable and reproducible capacitance humidity sensor is thermal oxidation [30] and carbonizations of the PS surface [29]. These treatments change the originally hydrophobic PS surface to hydrophilic, thus improving its humidity sensing properties.

Table 3.2. Capacity PS sensors for liquids and humidity

N	Si substrate	PS thickness	Metal contact, coating	Detected gas	Detection range	Sensitivity ΔC/C	Ref
1	p^+-Si , 20mΩ.cm + n-epi layer	4 μm	Cr/Au	Ethanol/water mixture	0-0,5% ethanol	2,5% per 0,1% with UV light 4% per 0,1%	[25]
2	p^+-Si , 20mΩ.cm + n-epi layer	4 μm	Cr/Au	methanol/ water mixture humidity	0-0,5% methanol 20-100%	1,6 300%	[24]
3	n^+-Si , 6 mΩ.cm		Au	humidity	10-90%		[26]
4	p-Si, 0,15-0,2 Ω.cm		Al	humidity	10-95%	1500%	[28]
5	n-Si	97 μm	Al, Ta_2O_5 coating	humidity	5-100%	300%	[31]
5	p-Si, 10-20 Ω.cm	30 μm, oxidated	interdigital Al	acetone, methanol, ethanol, i-propanol			[30]

3.4 Sensors Based on Schottky Barrier, MIS Structure and Heterojuction

The built- in potentials in the interfaces define the working principle of the sensors with *Schottky barrier, MIS structure or heterojunction* (in contrast to the resistive sensor type in which metal contacts are of Ohmic type). In such structures the electrical characteristics (I-V, C-V, impedance) are changed due to the influence of molecule adsorption on potential barrier or/and the parameters of PS (dielectric function and charge of surface states) [32-36]. The theoretical models of different metal/PS/Si structures were considered elsewhere (fig.3.10)[37-43].

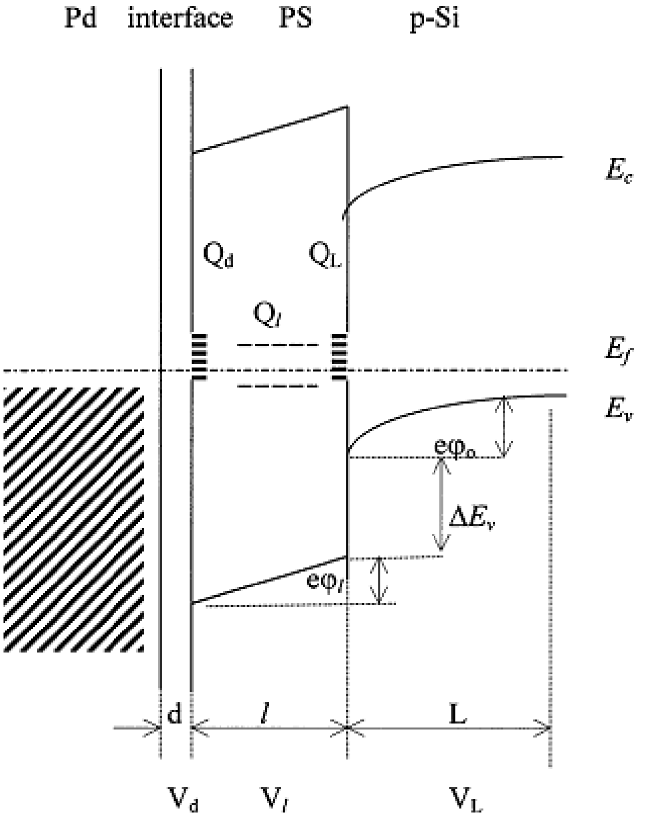


Fig.3.10. Energetic diagram of metal-tunnel thin interface layer- thin PS- p-Si structure.

The spectral dependences of density of surface electronic states for several PS thicknesses obtained from *C-V* curves are presented in Fig.3.11 [13]. The observed curves can be explained in the approach of the quasi continuous distribution of surface electronic states situated in the energetic region from $e\varphi_0+E_f$ to E_c. The thicker PS layer, the bigger density of surface electronic states is observed. surface electronic states which can act as adsorption sites in MIS type sensors are clearly observed in Deep Level Transien Spectroscopy method (DLTS). Fig. 3.11 shows DLTS spectra in different atmospheres. As can be seen, the DLTS curves of nano PS are narrowed and are shifted to the side of high temperature in atmosphere of CO_2, Ar and O_2 comparing with vacuum conditions. The behaviour of DLTS curves of mesoporous Si in different atmosphere is similar to nanoporous Si. It should be mentioned the special case of oxygen atmosphere. At oxygen pressure 0.1 mbar the DLTS spectra is similar as in vacuum condition (10^{-6} mbar). At partial pressure of 1 mbar the positive peak is fully disappeared and only negative two peaks are reiterated.

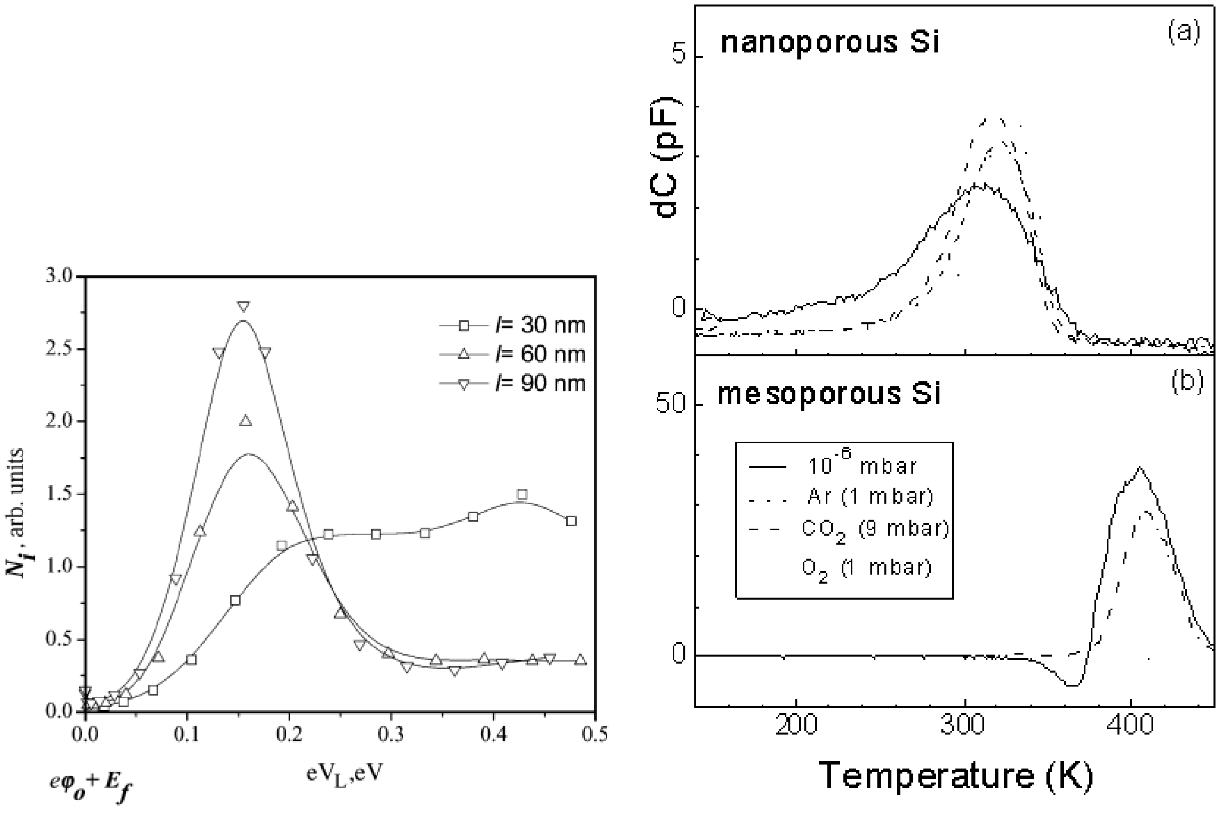


Fig.3.11 Density of surface electronic states versus electron energy for structures with l = 30, 60, 90 nm (a). DLTS spectra of nanoporous and mesoporous Si in vacuum, Ar, CO_2 and O_2 atmospheres (b).

The experimental *I-V* curves exhibit good rectifying behavior for every PS thickness (Fig.3.12). Both the reverse and forward currents decrease for thicker PS layers. The value of potential barrier is $e\varphi_0$=0.46 eV and does not depend on PS thickness. The ideality factor n for forward biases less than 0.3 V increases slowly with PS thickness and ranges from 3.1 to 3.5. Considering that the ideality factor is defined by the re-distribution of voltage drop between the depletion region, interface layer and surface electronic states,

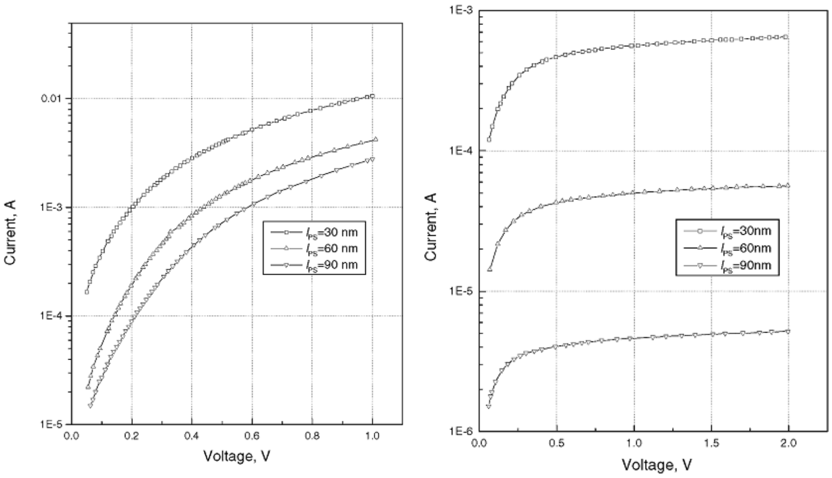


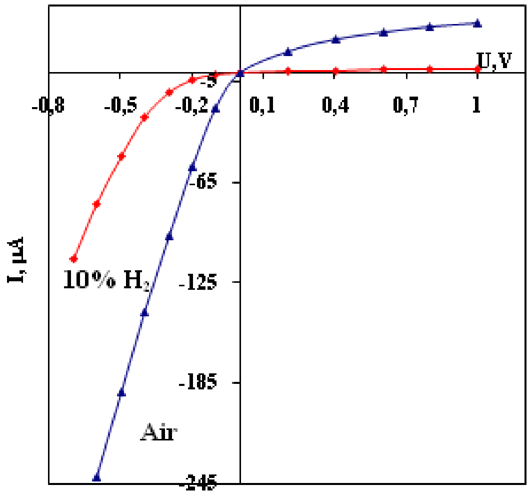
Fig.3.12. Experimental forward and reverse *I-V* characteristics of Pd-PS-p-Si structure at different PS thicknesses.

It is known that the flammability limits of hydrogen in air is 4.0-75 vol%. For this reason it is necessary to develop highly sensitive hydrogen sensors to prevent accidents due to hydrogen - containing gases leakage. Catalyst metal-semiconductor devises utilise the effect of the change in work function of the metal at the interface, because the adsorbed hydrogen atoms in catalyst metal form a dipole layer at the interface [44]. However, operating temperature of such sensors achieves 200-500°C, it corresponds to maximal efficiency of metal catalyst [45] that complicates the measuring instruments.

MIS type structures with PS and Pd as catalyst were applied to detect the H_2 and hydrogen - containing gases [46-51]. Application of MIS type structures with thin Pd and PS layers is a promised way to reduce the operating temperature and improve the hydrogen sensing since: i) the maximal hydrogen dissolubility at room temperature is observed namely for palladium among all catalysts, ii) well-known the island-type catalyst reduces the activation energy to dissociate the hydrogen-contained molecules, iii) the palladium film is semitransparent for penetration of hydrogen atoms that allows to accelerate the sensor response and recovery times. The SEM analysis of considered Pd-PS-Si structure shows the existence of Pd islands of ☐nm dimension.

Fig.3.13 shows the example of influence of hydrogen exposure (10%H_2+synthetic air) on *I-V* characteristics of elaborated structures at room temperature. The hydrogen adsorption decreases both the forward and reverse currents for Pd-PS-p-Si structure with respect to sign of Schottky contact and increased positive charge in interface. Fig.6 shows the dynamic response of Pd-PS-p-Si junction exposed to successively injected H_2 concentrations. It is necessary 200-300 s, approximately, to saturate the current response for given hydrogen concentration.

A good reproducibility of current response $S=\Delta I/I_0$ is observed. Generally, the response time can be defined as the time for which the sensitivity achieves of 90% of its final response: $S(t=t_{res})=90\% \; S(t=\infty)$. The recovery time is defined as time for which the sensitivity recovers 90% of its equilibrium value: $S(t=t_{rec})=10\% \; S(t=\infty)$. The response/ recovery time is 2 min and 5 min, respectively. (Fig.3.14).



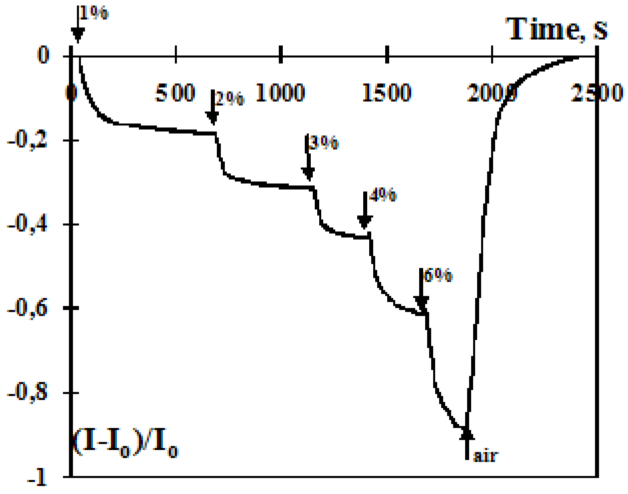


Fig.3.13. The *I-V* characteristics in synthetic air and synthetic air with 10% of H_2 for Pd-(15 nm)PS-p-Si structures.

Fig.3.14. Current responses as function of time for Pd-(15 nm)PS-p-Si diodes under successively injected H_2 concentrations in synthetic air, U=-0.3 V

In general, a gas sensor is said to have good sensitivity when it exhibits both of the following properties:1) it has a high signal- to noise ratio, extremely at low partial pressure of gas to be detected, and, 2) it is sensitive to the gas to be detected over a wide range of partial pressure. The current changes versus hydrogen concentrations shows the sensing in the range 200 ppm- 10% with the linear sensitivity of 0.15 µA/100ppm at hydrogen concentrations from 200 ppm to 6% ppm in air.

The hydrogen sensing mechanism includes the adsorption and dissociation of hydrogen molecules in the Pd catalyst, the transport of hydrogen atoms across the catalytic metal film and the formation of dipoles at the metal-insulator interface.

For Pd-PS-p-Si device the experimental change of *I-V* characteristics at hydrogen adsorption is not fitted only by variation of barrier height. There is the dependence of adsorption effect $(I-I_o)/I_o$ on applied voltage, when $(I-I_o)/I_o$ decreases at the increase of forward voltage biases, and linear dependence of $(I-I_o)/I_o$ on hydrogen concentration. To explain the experimental results we account the surface electronic states insides of pores and/or at PS-Si and PS-SiO_2 interfaces).

The sensor sensitivity evolution as function of its aging is shown on Fig.3.15. There is the small reduction of sensor current with aging in ambient atmosphere. It is conditioned by oxide layer growth that is observed by FTIR as increasing of Si-O-Si absorption modes. By the way, since the slop of curve current versus hydrogen concentration does not change, the sensitivity of sensor remains the same with time. Moreover, by contrary to PS sensors with opened surface the metal covering reduces the temp of surface oxidation.

A typical result of contact potential difference response (CPD) of the Pd/thin PS/n-Si structure in synthetic air (20% O_2+ 80%N_2) at H_2 concentrations from 200 up to 4000 ppm is given in Figure 3.16,a [46]. Sensors are characterised by a considerable sensitivity, time of response and recovery time is in minutes range when operated at T_{room} and reveal the best sensing comparing with Pd/p-Si and Pd/SiO_2/p-Si structures. Experimental values of CPD response are described by the square root dependence of the partial H_2 pressure as predicted by theory [44]. To detect H_2 it is proposed to use Pd/CH_x/PS/Si structures. The role of the CH_x layer is to protect PS against oxidation [51]. Such structures demonstrate a record response and recovery time (2 and 15 s, respectively) at large 0.2 bar hydrogen concentrations.

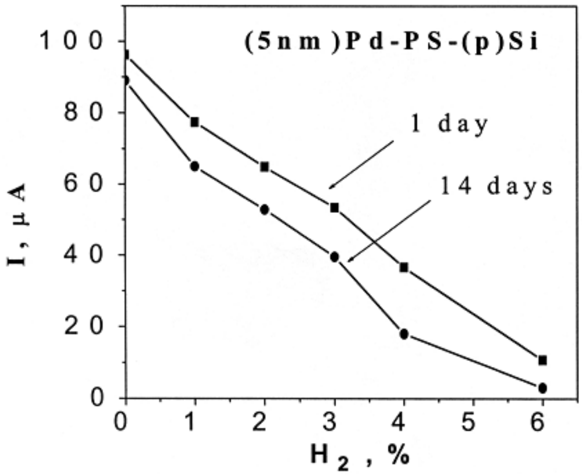
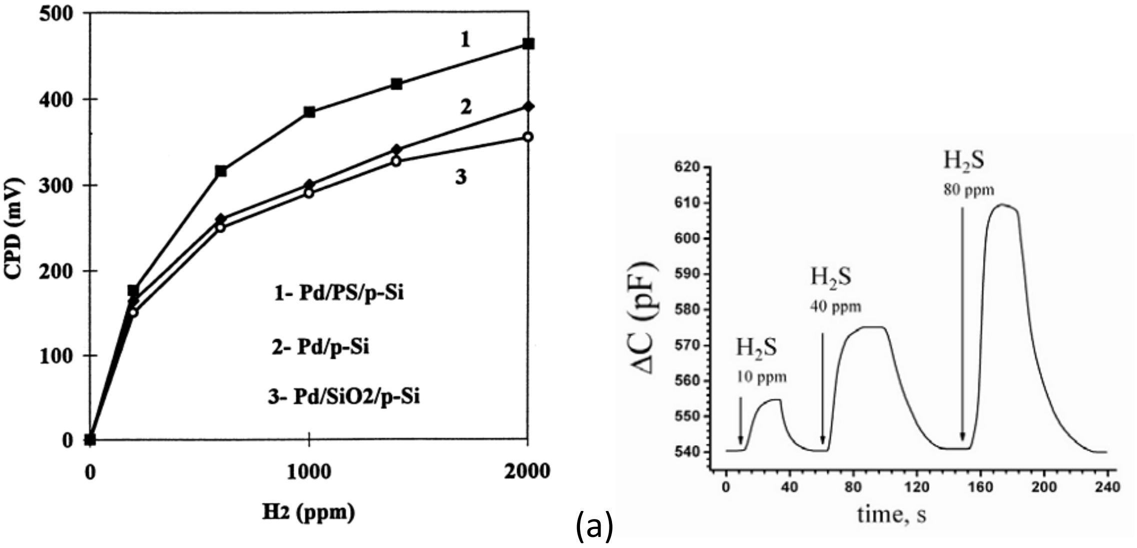


Fig.3.15 Sensor current versus hydrogen concentration 1day and 14 days after structure formation.



(a) (b

Figure 3.16. a) Steady-state CPD response of Pd/thin PS/p-Si, Pd/p-Si and Pd/SiO_2/p-Si structures versus H_2 concentration in synthetic air. Reprinted with permission from [46], b) Kinetics of the signal response of Pd/PS/p-Si under action various concentrations of H_2S.

Schottky-type and MIS structures were applied to analysis of other gases too. The deposition of thin films of catalytic metal which creates island - type covering on the PS surface, or special sensor design that facilitates diffusion of molecules through the pores towards metal layer is applied. The example of Schottky and MIS-type sensors of several gases are presented in Table 3.3

Table 3.3 Schottky and MIS- type sensors with PS interface

N	Potential barrier structure	PS thickness	Measured value	Detected gas	Detection range	Ref
1	Pd/PS/n,p-Si	15-75 nm	I-V	H_2	200ppm-10%	[49]
2	Pd/PS/n-Si		CPD	H_2	200ppm-4%	[46]
3	Pd/PS/p-Si		resistivity	H_2	0-1,5%	[48]
4	Pd/PS/p-Si		I-V	H_2 O_2	500 ppm 500 ppm	[47]
5	Al-Pd/PS/p-Si	1 μm	C-V	H_2 H_2S	10-1000 ppm 5-80 ppm	[50]
6	Pd/CH_x/PS/p-Si	1-10 μm	I-V	H_2	0,2 bar	[51]
7	Pt-Pd/PS/n-Si	5-10 μm	I-V	NO_2	5 ppm	[53]
8	Au-Pd/SnO_2/PS/p-Si	0,6-19 μm	resistivity	NO_2 LPG		[57]
9	Al/CH_x/PS/p-Si	1-10 μm	I-V, C-V	CO_2	500-1500 ppm	[59]
10	Al/CH_x/PS/p-Si	1-15 μm	I-V, C-V	ethylene, ethanol, propane	>115 ppm	[58]
11	Au/PS/n-Si	41 μm	I-V	ethanol, methanol	0,2% 0,2%	[54]
12	Ag/a-Si:H/PS/p-Si	1μm	I-V	water-ethanol	0-50%	[61]
13	Ti/PS/p-Si	100-300 nm	I-V, C-V	humidity	0-50%	[55]
14	Au/PS/p-Si		C-V	humidity	10-100%	[56]

MIS Pd/PS/Si structure is very sensitive to low (5-80 ppm) concentrations of hydrogen sulphide toxic gas [50]. The kinetic dependence of capacitance at gas inlet and outlet is shown in Figure 3.16,b. At low gas concentrations the MIS structures had a high speed of performance and short times of response (<10s).

Heterojunctions based on CH_x/PS structure show a good rectifying behavior, and current sensitivity $\Delta I/I_0$ to ethylene (26.8%), ethane (13.7%) and propane (9.4%) at 575 ppm concentration of these gases, response and recovery times at 115 ppm ethylene are of about 3 and 7 min, respectively [58]. It has been applied also as CO_2 sensors in measured range of 500- 1800 ppm [59]. Amorphous Si/PS heterojunction sensors of organic vapors and molecular oxygen are considered in [60, 61].

3.5. FET AND ISFET SENSORS

The working principle of a gas sensor that is based on the field effect derived from the metal oxide semiconductor field effect transistor (*MOSFET*). In n-channel MOSFET, a p-Si wafer contains two n-type diffusion regions (source and drain), the structure is covered with a thin SiO_2 insulating layer on top of which a metal gate electrode is deposited. When a positive voltage (with respect to the substrate) is applied to the gate electrode, electrons (which are the minority carriers in the substrate) are attracted to the surface of the semiconductor and a conducting channel (inversion layer) is created between the source and the drain, near the SiO_2 interface. The conductivity of this channel can be modulated by adjusting the strength of electrical field between the gate electrode and the Si, perpendicular to the substrate surface. Applying the voltage between the drain and the source results in a drain current between the n-regions (fig.3.17).

In FET gas sensor the space charge layer near SiO_2 interface is induced by the adsorbed gases. Fig.3.17 shows the case of co-adsorption of hydrogen and oxygen. Hydrogen adsorption occurs in 4 steps: adsorption of hydrogen on gate, hydrogen dissociation on Pd gate, penetration of protons through thin Pd film, adsorption as dipole at Pd-SiO_2 interface. As results the supplementary electric field of adsorbed dipoles shifts the C-V and I-V curves (Fig.3.17). The present of oxygen in atmosphere provokes the creation of hydroxyl and water molecules on the surface that prevents the hydrogen adsorption. Application of FET gas sensors for detection of other molecules demands the complex FET structure, for example isle –type gate electrodes.

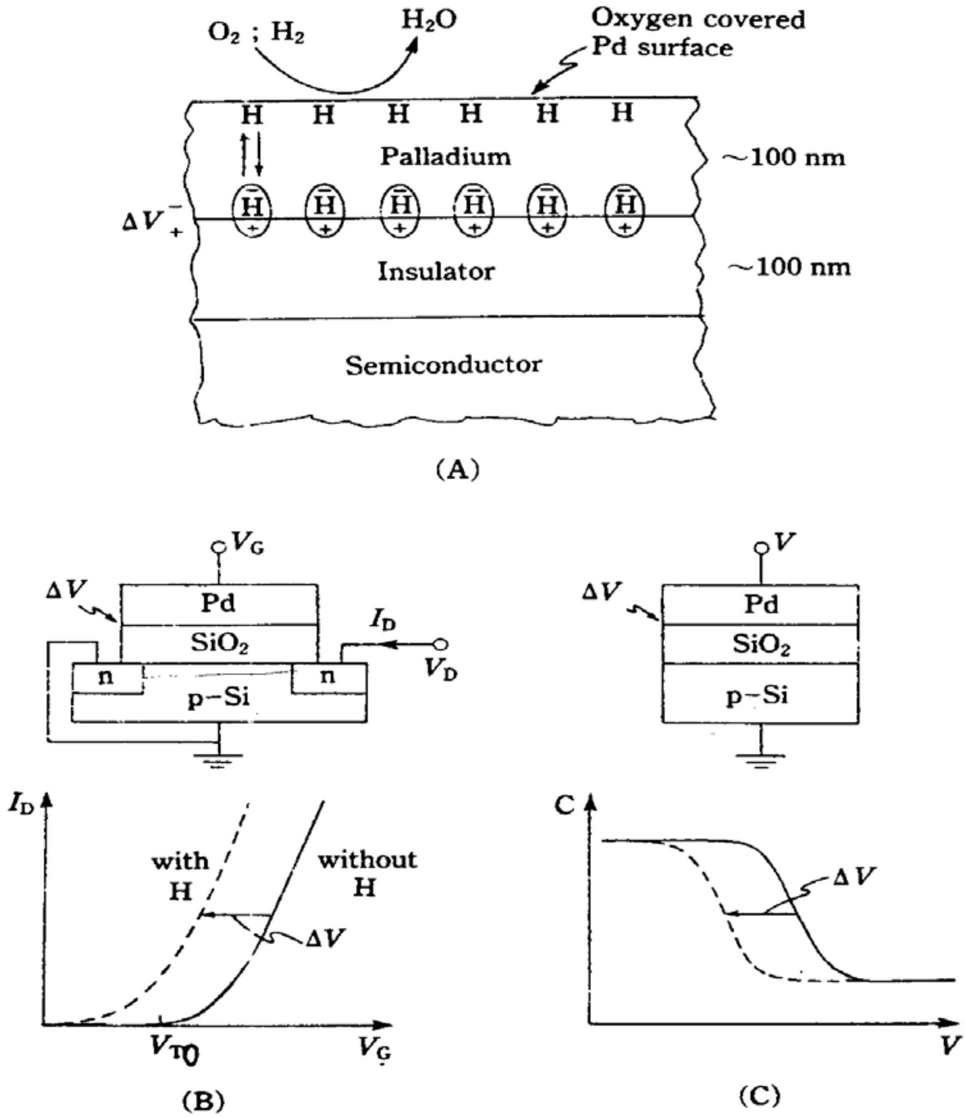


Fig.3.17 Adsorption of hydrogen and oxygen on gate of n-channel MOSFET transistor (a), *I-V* source-drain (b) and C-V characteristics without and with H2 adsorption

The application of PS layer instead of metal gate in FET, seems, very attractive since developed internal porous surface enhances the molecule adsorption and technology of such sensors should be compatible with silicon IC technology. However, the technical realization of declared compatibility of PS fabrication process with the IC technology is one of the bottleneck of PS sensors, since nonstandard processes (like electrochemical etching) are generally used for PS technology. An alternative to this approach is the modification of a standard IC technology with the addition to specific steps for PS fabrication. However, performing the anodization step at PS formation between standard IC processing can result in contamination from PS itself to other wafer parts and PS degradation because of thermal and etching steps. To avoid this problem it is desirable to perform PS fabrication at the end of IC processing [62].

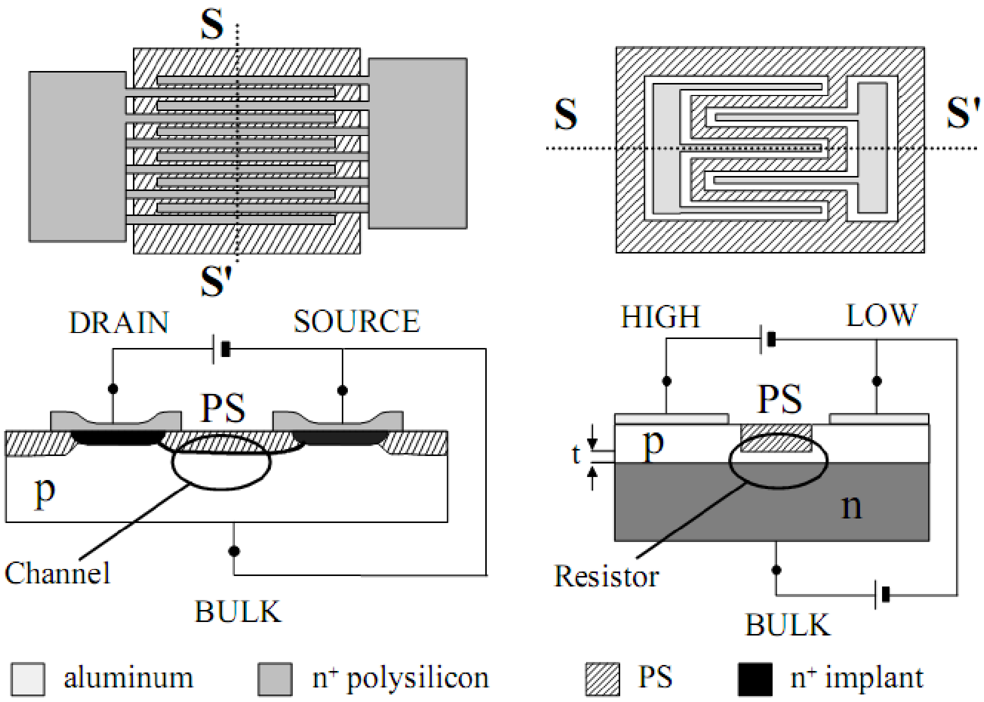


Figure 3.18. Schematic top view (top) and cross-section (bottom) of fabricated gas sensing devices: the APSFET on the left; the resistor on the right.

Adsorption porous Si-based FET (APSFET) gas sensor was proposed [62, 63]. It is a gas sensor based on a standard bipolar +CMOS+DMOS silicon technology in which PS is fabricated at the end of process (Figure 6,a). APSFET works like open gate FET structure with PS layer directly above the conduction channel in crystalline Si. Electrical conduction does not take place in the PS layer, which only plays the role of sensing material. In this way the electrical properties of the PS sensing film are not involved in the measurement, with advantages in terms of sensors sensitivity and reproducibility.

Figure 3.19,a shows a typical *I-V* source-drain characteristics of a device for several isopropanol vapor concentrations.

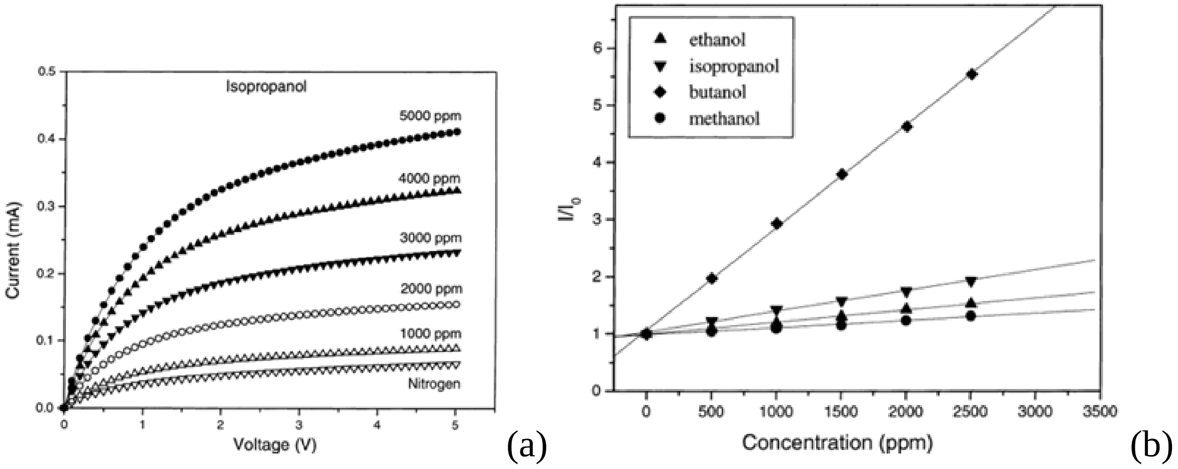


Figure 3.19. I-V characteristics of APSET in presence of several isopropanol concentrations (a) and sensor response to the several alcohols at V= 5V (b).

As typical for FET structures they have linear region for low voltages and a saturation current at high voltages. The remarcable growth of sensitivity is observed at the increase of PS thickness from 0.15 to 0.5 μm. In Figure 3.19,b the effect on the sensor current of alcohols vapor at constant voltage bias (5 V) is shown. For every species the sensor response linearly depends on the concentration of the species in the environment, and the sensitivity of the sensor increases with the number of carbon atoms in the molecule or the molecular weight of the molecule. An opposite effect with respect to alcohol vapors, a reduction of the conduction current is observed at exposition in organic acids (formic, acetic, propanoic and butanoic acid).

NO_2 concentration as low as 100 ppb was detected using the APSFET [63] (Fig.3.20). Nonoxidized sensors show a high sensitivity only for fresh devices reducing with the aging of the sample. Oxidation of the PS layer improves the electrical performance of sensors, in terms of stability, recovery time, and interference with the relative humidity level, keeping the high sensitivity to nitrogen dioxide. For instance, 100 ppb of NO_2 produce a current variation of about one order of magnitude. The drawback of such sensor - RH level affects the percentage variation of the sensor current when exposed to NO_2 and the response time. FET sensor with PS layer under gate layer shows the rapid response (250s) and recovery (240s) time at hydrogen detection (1.2 mbar), however, drain current does not recover completely to the initial value [64] due to hydrogen-induced drift (HID) effect, when hydrogen trapping sites exist in PS [49].

The analog of FET structure with p-n junction gate and adsorbing top layer of PS (that acts as a floating sensing gate) is proposed in [65, 66] (Figure 3.18,b). This device differs in the mechanisms of conduction, if in the APSFET the current flows in the inversion layer of the FET (electrons); in this structure the transport is due to the majority carriers (holes). The adsorption of organic vapors in PS modulates the resistance of a *p*-Si resistor (Fig.3.18,b). The current is measured by biasing one of comb-shaped aluminum electrodes (HIGH) with respect to the other one (LOW in Fig. 3.18,b).

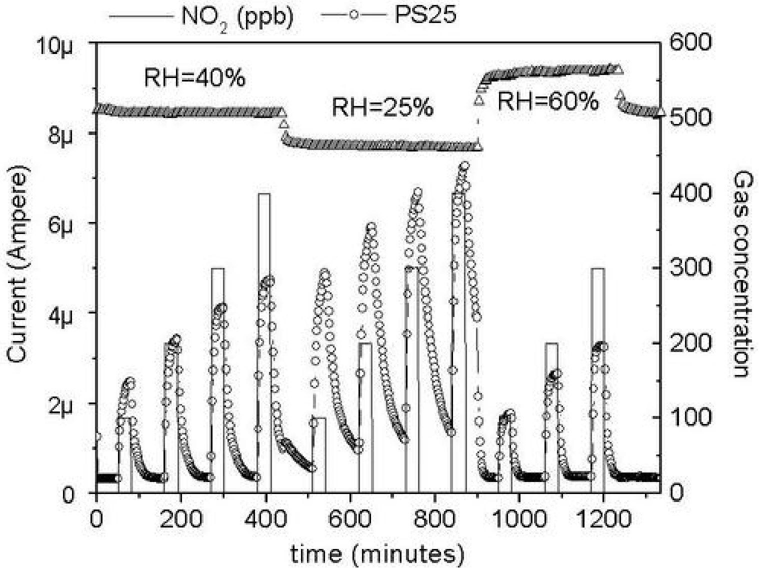


Fig.3.20 Current-time response of an APSFET with oxidized PS in presence of NO_2, for different relative humidity levels [Barillaro 2006].

In order to all the current flows through the p-Si resistor and reduce current leakages toward the bulk n-Si, the bulk is connected to a higher positive voltage than the HIGH terminal. The exposition in isopropanol vapors reduces the resistance of the *p*- resistor (from 150 kΩ in N_2 to 50 kΩ in vapor at V=5V), but the linear behavior is maintained. The adsorption of polar molecules into the PS layer could change the free carriers concentration in the p- resistor or modulate the resistance value through modifying the localized charge on the interfacial states [32-36] that leads to the change of thickness of the underlying space charge layer.

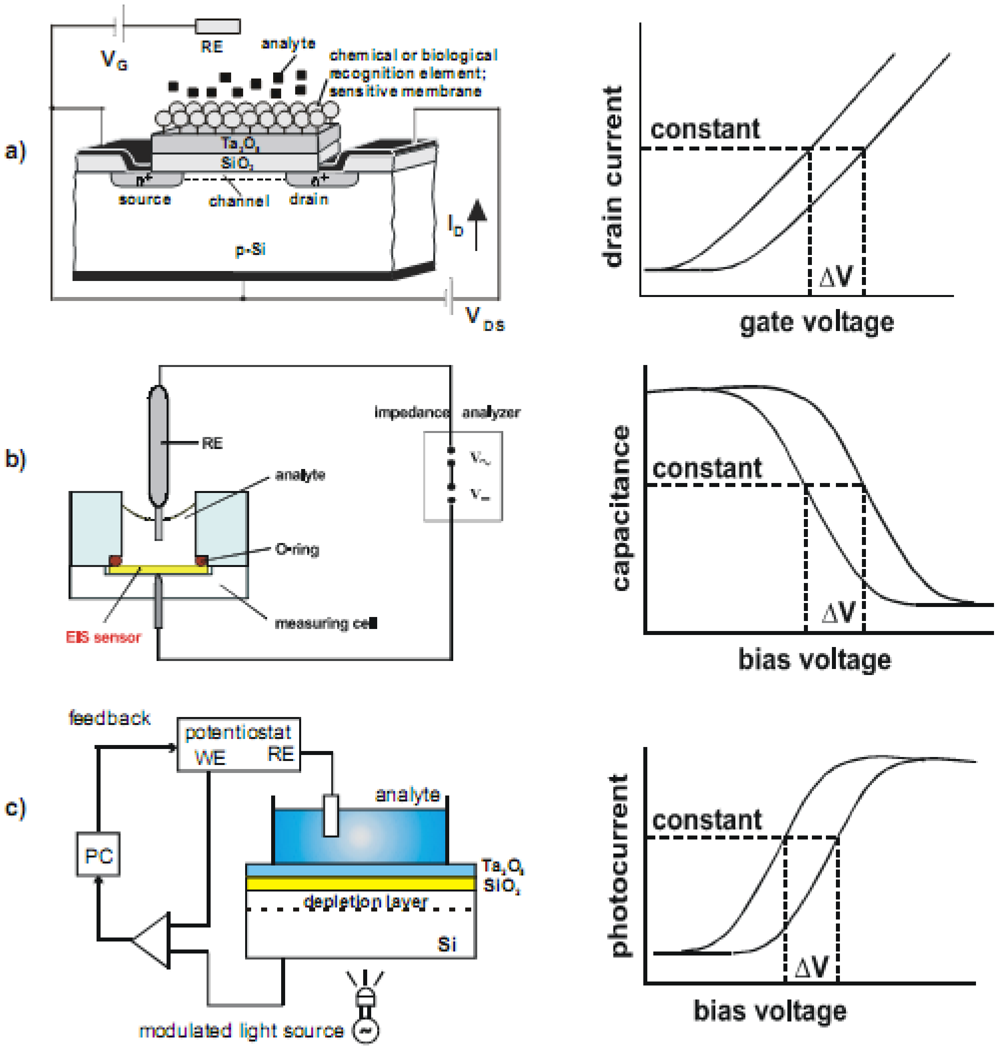


Fig. 3.21**.** Detection principles of ISFET (a), EIS sensor (b) and LAPS (c) with a corresponding change of respective sensor signal (VG: gate voltage, VDS: drain–source voltage, ID: drain current, V~: AC voltage, V=: DC voltage, WE: working electrode, RE: reference electrode)

Other modification of MOSFET structure that can be applied to electrolyte solution measurements is an *ion sensitive field effect transistor (ISFET)* (fig.3.21,a). In this device, the metal gate of MOSFET is replaced by an ion-sensitive membrane, the measured solution, and a reference electrode. The source-drain conductance varies as the function of the *pH value* (pH is defined as a negative decimal logarithm of the hydrogen ion activity in a solution) or ion concentration of the solution in contact with the sensing surface. Input *I-V* characteristics of the ISFET at pH change are shifted along the voltage axis. By measuring this voltage shift, the pH can be determined. ISFET is very sensitive to any kind of potential generation at or near the gate insulator/ electrolyte interface.

Thus, each biological or chemical reaction, leading to chemical or electrical changes (for instance, a pH or ion-concentration change) [67-69] at this interface, can be measured by means of these devices coupled with the respective chemical or biological recognition element. The first realization of ultra high sensitivity ISFET devices in which PS structures are placed on the gate region like ion-sensitive membrane is reported [70].

An *electrolyte-insulator-semiconductor (EIS)* is MIS-type structure in which the metallic gate electrode (like in ISFET) is replaced by a sensor layer, which is in direct contact with the analyte (electrolyte) and a reference electrode (fig.3.21,b). EIS is considered to be an alternative to ISFET structures since for the last the poor adhesion and fast leaching-out of the sensitive materials as well as electrochemical corrosion of the passivation layer sometimes is observed. In contrast, EIS exhibits higher long-term stabilities and possess planar surfaces where no passivation of the electronic circuits is necessary [71,72]. As the ISFET, the EIS sensor is very sensitive for any kind of potential generation at or near the gate insulator/electrolyte interface. The analyte concentration or composition to be detected can be measured in capacitance/ voltage (*C-V*) or constant capacitance modes. The charge carrier distribution at the interface insulator–semiconductor is controlled by an external *dc* voltage applied to reference electrode, a superimposed *ac* voltage is used to measure the space-charge layer capacity. According to Nernstian law a theoretical value of *C-V* curve shift is 59.1 mV per pH decade under standard conditions.

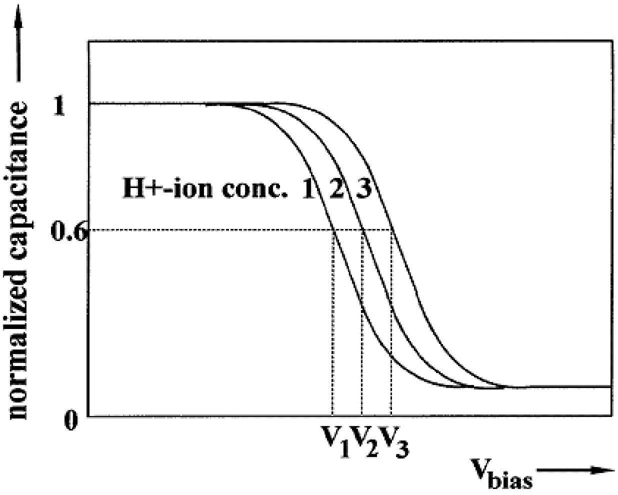


Fig.3.22 C-V curves for EIS structure at increasing of H^+ concentration

For pH sensors due to better pH response, hysteresis and drift characteristics compared with SiO_2, the gate materials such as Al_2O_3, Si_3N_4 and Ta_2O_5 are used instead of SiO_2 [73]. Si_3N_4 is known to be resistant to many chemicals. However, miniaturization of these EIS structures down to the micrometer scale, as for ISFETs, is very difficult since the capacitance values become too small for accurate measurements. To overcome this drawback it was proposed to use PS layers [71]. The scheme of the different capacitive planar and two structured EIS sensors Si_3N_4 (30-70 nm)/SiO_2 (30-50 nm)/Si for the pH determination is given in Figure 3.23. The structured EIS sensor consists of the trenches with the size varied between 2 μm and 8 μm and a depth of about 2 μm or macro-PS (pore diameter and pore depth is 1 and 2 μm, respectively). For all sensors an average pH sensitivity of 54 mV/pH in the concentration range from pH 4 to pH 8 exists. The drift, i.e. the time-dependent shift of the calibration curve varied 4-6 mV/day. However, for the porous EIS sensor, due to the surface enlargement, C_{max} is increased to about 800 nF in comparison to about 40 nF for the structured sensor and to about 25 nF for the planar one.

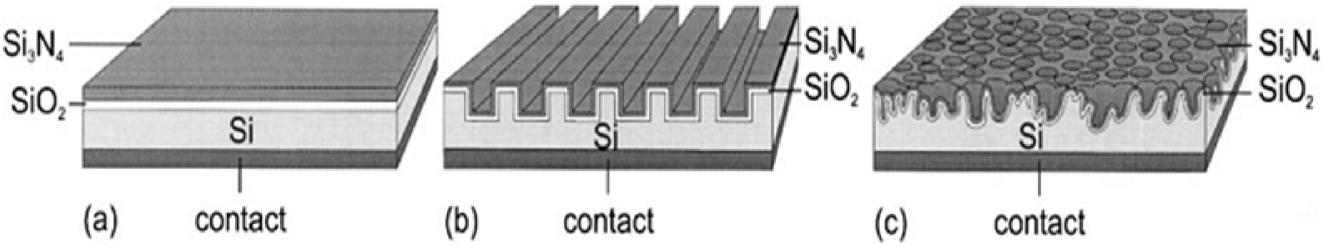


Figure 3.23. The scheme of the semiconductor insulator capacitors: (a) planar, (b) structured and (c) porous.

The *light-addressable potentiometric sensor (LAPS)* is a semiconductor-based chemical sensor with an EIS (fig.3.21,c). The principle of the LAPS is similar to that of the EIS capacitance sensor, in which the capacitance of the EIS system is measured to determine the ion concentration of the solution [72,74]. In the case of LAPS, a *dc* bias voltage is applied to the EIS structure so that a depletion layer appears at the insulator-semiconductor interface. The width, and therefore, the capacitance of the depletion layer vary with the surface potential. This variation of the capacitance is read out in the form of the photocurrent induced by the modulated light. The light-addressability of the LAPS allows an application of the LAPS as a chemical imaging sensor [75]. By using a focused laser beam as the light source, the local values of the surface potential can be measured. A two-dimensional map of the distribution of the pH value or the ion concentration can be obtained by scanning the sensing area with the focused laser beam.

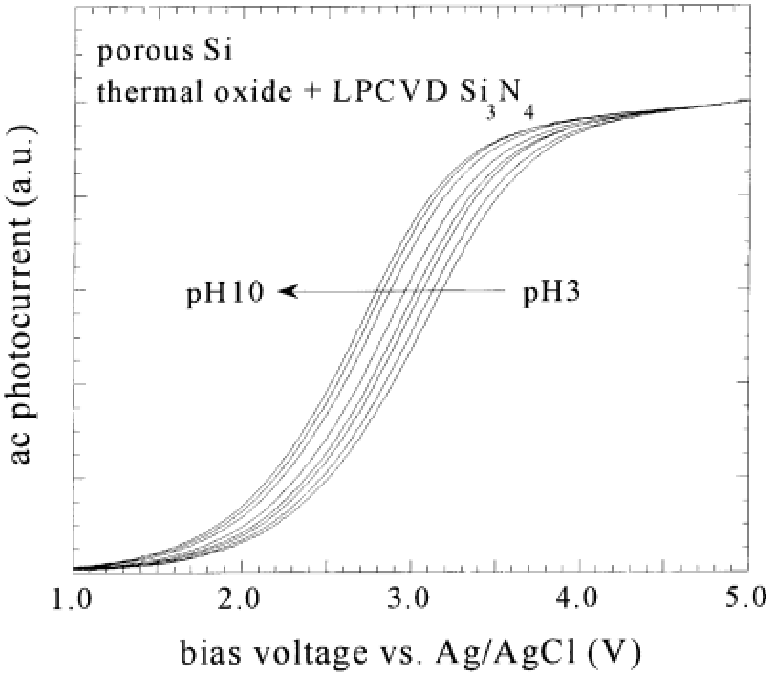


Fig.3.24. LAPS signal at different pH

Table 3.4. FET, ISFET, EIS and LAPS sensors

N	Type of PS structure	Detected gas	Detection range	Ref
1	FET	H_2	1,2 mbar	[64]
2	FET	NO_2 CO_2 ethanol	100 -500 ppb 40-100 ppm 250-500 ppm	[63]
3	FET	isopropanol, ethanol, methanol, acetic acid	<10000 ppm <2500 ppm <1500 ppm	[62]
4	FET	Isopropanol	<15000 ppm	[65, 66]
5	ISFET	pH-meter	pH=2-4	[70]
6	EIS	pH-meter	pH=4-8	[71]
7	EIS	pH-meter	pH=2-9	[68]
8	EIS	pH-meter	pH=5-10	[67]
9	EIS	pH-meter	pH=4-9	[69]
10	LAPS	pH-meter	pH=3-10	[76]

Porous Si LAPS is developed in [76]. LAPS reveals satisfactory high pH sensitivity of 57.3 mV/pH. As an application of the PS LAPS, a penicillin LAPS was fabricated by the immobilization of the penicillinase on the surface. The pH change due to enzymatic reaction was detected for the penicillin concentration between 250 μM and 10 mM. Several field effect sensors are shown in Table 4.

3.5 MEMBRANE SENSORS

PS can be used in gas sensor design as micromachined membranes. Miniaturized hotplates are important parts of pellistor and resistor-type gas sensors [2]. The incorporation of thermally isolated micro- hotplates in the device construction allows a considerable reduction in both the power consumption and thermal transient time. For hotplate fabrication, one can use technologies of Si bulk and surface micromachining [7]. In bulk micromachining technology, the Si wafer is etched from the backside of the Si wafer in defined regions. Surface micromachining technology for free-standing membrane fabrication uses sacrificial layers. An easily etchable sacrificial layer is deposited onto the substrate surface followed by a second layer deposition. This layer will form the free-standing membrane after removing of the sacrificial layer by etching. PS is an ideal approach to combine both technologies and can be fabricated with thicknesses up to several tens of micrometers [77, 78]. PS as a sacrificial layer, due to its thickness, offers the possibility to create a large air gap between the membrane and substrate. Due to very high surface-to-volume ratio the sacrificial PS layer can be perfectly etched away.

The resistive sensor based on free–standing PS membrane displays very high sensitivity to detect NO_2 at T_{room} for concentrations as low as 200 ppb, $\Delta G/G= 111$ [79]. This sensor showed low interference from other gases like ethanol, methanol and ozone and no interference from carbon monoxide. On the contrary the humidity influencing the resistivity is significant.

To overcome some drawbacks of PS resistive sensors (poor selectivity to adsorbed gases, low log-term stubility during ageing) it is proposed to use PS as a physical support for another sensing film [80,81]. To monitor aromatic hydrocarbon molecules the mixed semiconducting oxides Sn–V–O have been used as a sensing layer. These oxides have shown peculiar catalytic properties, promoting selective oxidation of specific hydrocarbon compounds such as alkylaromatics and alcohols. The main feature of the device is represented by a permeated suspended macro- PS membrane few tens of microns thick on top of which a heater resistor and a temperature sensor are integrated (Figure 3.25).

Characteristics of the permeated PS suspended membrane sensor at 10% RH: sensitivity $\Delta G/G$ = 1.0 (at 0.5 ppm C_6H_6), response time (90%) is 80 s, recovery time (70%) is 120 s, measurement cycle duration is 15 min (Table 3.5).

Table 3.5. Membrane sensor parameters

Sensitivity ΔI/I (at 0,5 ppm C_6H_6)	1,0
Response time (90 %), s	80
Recovery time (70 %), s	120
Full time for 1 cycle measurements, min	15
Cross-sensitivity (%) at 1 ppm C_6H_6, 30 ppm CO	30

Analog to EIS working principle membrane sensor is proposed [82] (fig.3.26). The sensor design consists of an macro-PS region connected to a KOH etched fluidic channel. The KOH etched fluidic channels are monolithically integrated with the sensing membrane causing the analyte delivery sites and fluidic channels to be self-aligned. The frontside macro-PS region serves as the sensing membrane for solvent detection.

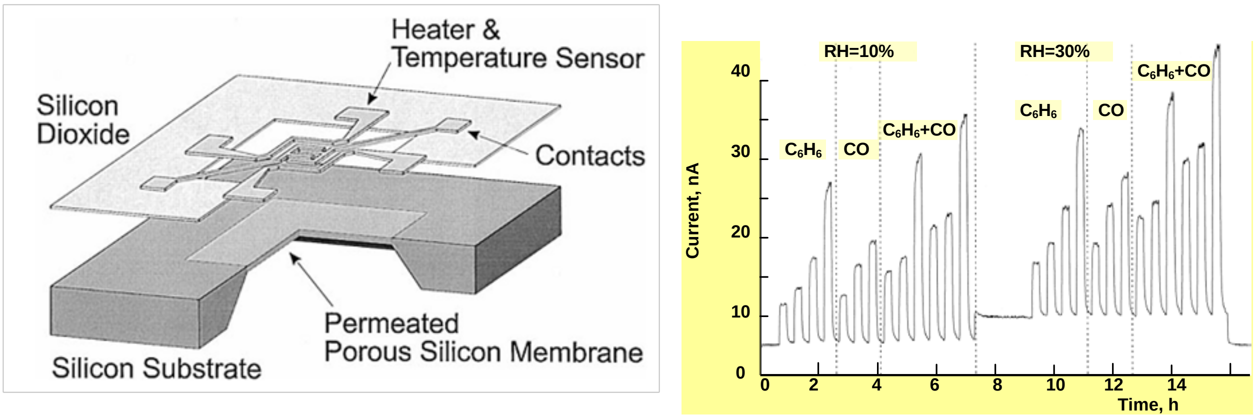


Figure 3.25. Sensor microstructure based on a suspended PS membrane and sensor response for C_6H_6 and CO.

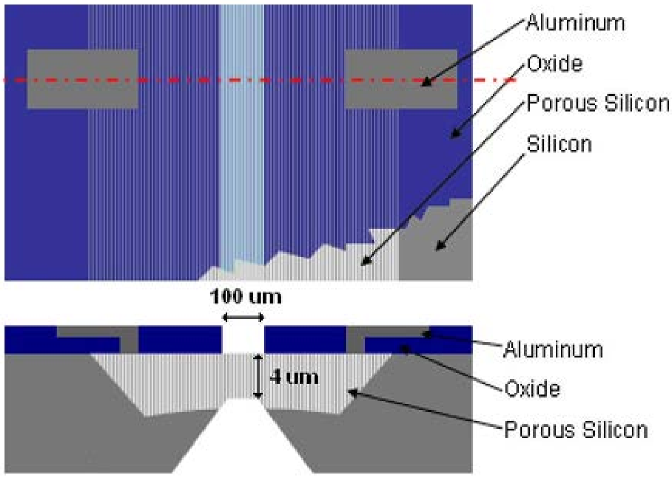


Fig. 3.26. (a) Top view of device layout (not to scale) with bottom right corner oxide “torn” to reveal underlying PS and Si. (b) Cross-sectional view taken at red cut line in (a), which depicts the flow-through membrane, passivating oxide, and aluminum contacts (not to scale) [Clarkson 2007].

3.7 OPTICAL SENSORS

Optical properties of PS have been proposed for gas, vapor and liquid sensing since they appear to be extremely sensitive to the presence of dielectric substances inside of the pores. As the refractive index of the void space increases, the effective refractive index of the PS layer increases, causing the optical spectrum of the layer to shift to longer wavelengths. Thus, by monitoring the reflectance or transmission spectrum, one can detect the binding of molecules inside the pores since the capture of targets inside the void space increases the refractive index. In order to monitor the refractive index of PS samples, a convenient way is to fabricate the samples as suitably thin single layers or stacks of thin layers such as a *Bragg mirror*, *rugate filter*, *Fabry–Perot filter*, *luminescent PS microcavity (PSM)* [3,83-86].

Bragg mirrors are periodic stacks of two quarter-wavelength optical thickness layers of different refractive indices. The periodicity gives rise to a photonic bandgap, in which light propagation is forbidden and incident light is reflected. A microcavity consists of two Bragg mirrors and a layer that breaks the periodicity of the refractive index profile. The reflectance spectrum of a microcavity with a half-wavelength optical thickness layer is characterized by a resonant dip in the stop band (Fig.3.27).

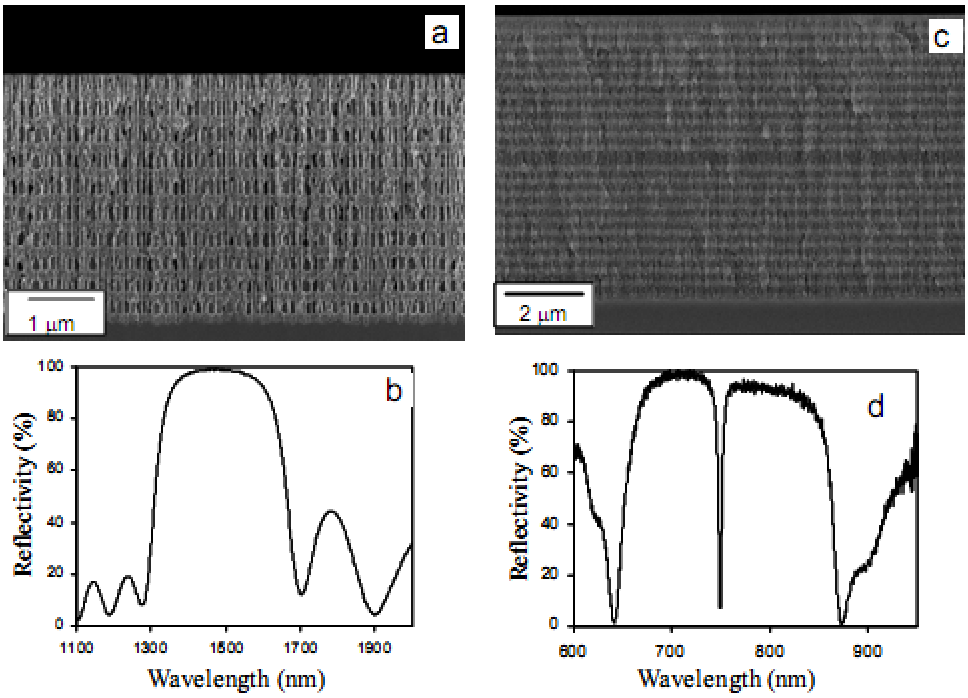


Fig.3.27 Bragg mirror (left) and Fabry–Perot filter (microcavity) (right): design and reflectivity spectra

The spectral position of reflectivity maximum is defined as

$$n_1d_1+n_2d_2=\lambda/2$$

where n_1 and n_2 are refractition indices of first and second layers, d_1 and d_2 are their thickneses, respectively.

Application of BM and PSM for sensors is attractive since PS Bragg reflector can be prepared by anodization process since refractive index depends on porosity, the lst is function of current density (fig.3.28). The pore filling with substance which has refractive iondex n>1 results in shifting of Bragg maxumum that can be easy detected in spectroscopic experiments. (fig.3.29).

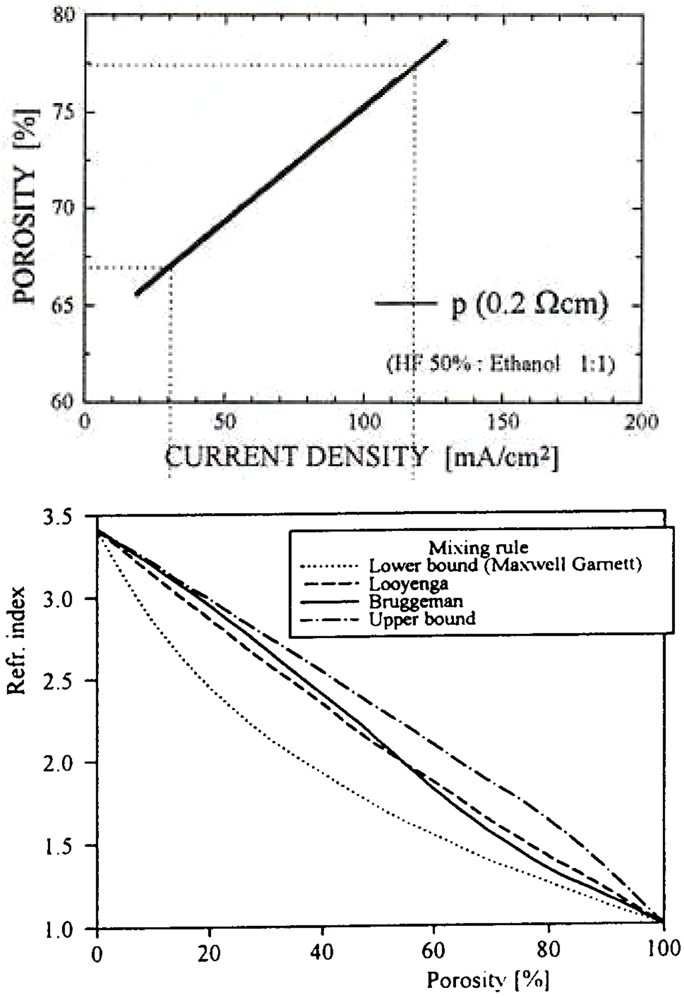


Fig.3.28. Dependence of porosity of PS on anodization current (a) and refractive index versus porosity (b)

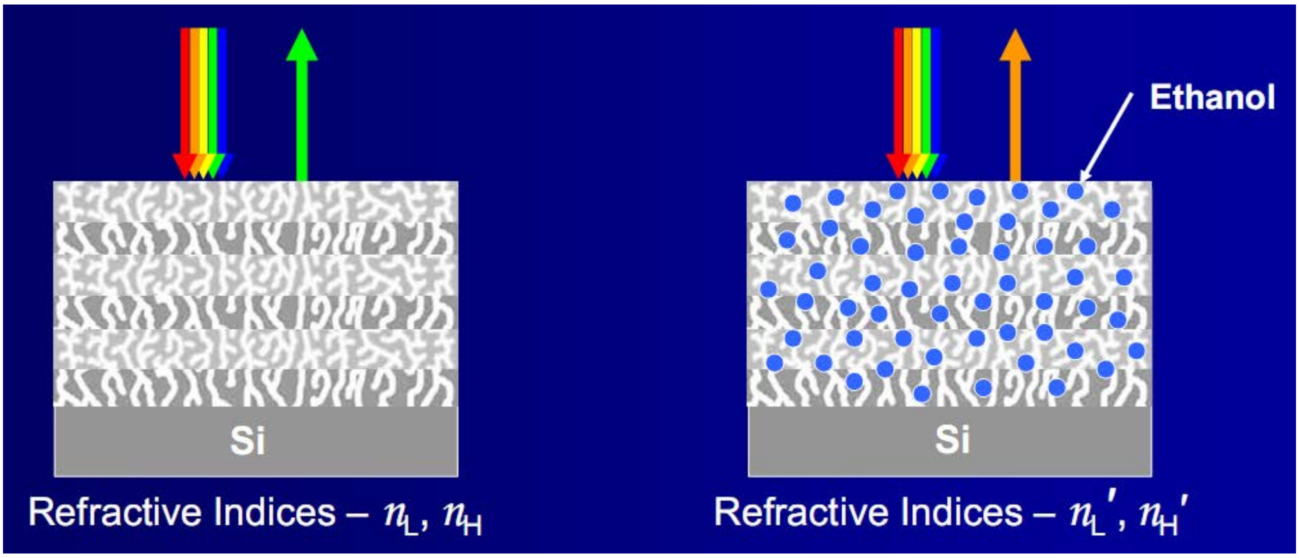


Fig.3.29 The principle of optical sensor based on Brag mirror for liquid detection

The sensitivity of optical sensors is $\Delta\lambda/\Delta n$, where $\Delta\lambda$ is the wavelength shift and Δn is the change of refractive index. For a system able to detect a wavelength shift of 0.1 nm, the minimum detectable refractive index change is 2.10^{-4} [85]. The problem of PS optic sensors is the instability due to the aging of PS itself. One way to prevent this effect is the PS oxidation [84] or PS chemical modification [87].

Figure 3.30 shows the reflectivity at normal incidence for the Bragg mirror (it consists of ten layer pairs, of thickness 0.36 and 0.19 μm having porosities of 71% and 54%, respectively, electrochemically etched in the p^+ doped 0.01 Ω.cm Si wafer, yielding indices of 1.56 and 2.09 for the layers, the reflectivity peak at 957 corresponds to the second Bragg condition peaks). The 70-120 nm shift in the Bragg wavelength of the mirror arises from refractive index changes, induced by capillary condensation of the acetone and chlorobenzene vapors in the meso-PS, in the layers of the mirrors [88].

Fabry–Perot fringes measured from the PS layer in ethanol solution are changed as the concentration of the solution varied (Figure 3.31). Standard deviation of the difference between the fringe patterns obtained in ethanol solutions and deionized water showed an almost linear relationship to the logarithm of ethanol molar concentration in the range between 1.10^{-5} and 1.10^{-14} M [89]. PS vapor sensor based on shift of Fabry–Perot fringes was demonstrated with the ethanol detection limit of 500 ppb and a dynamic range of nearly five orders of magnitude [90].

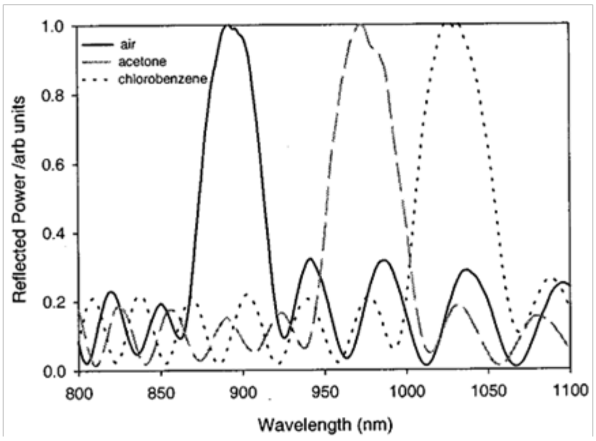
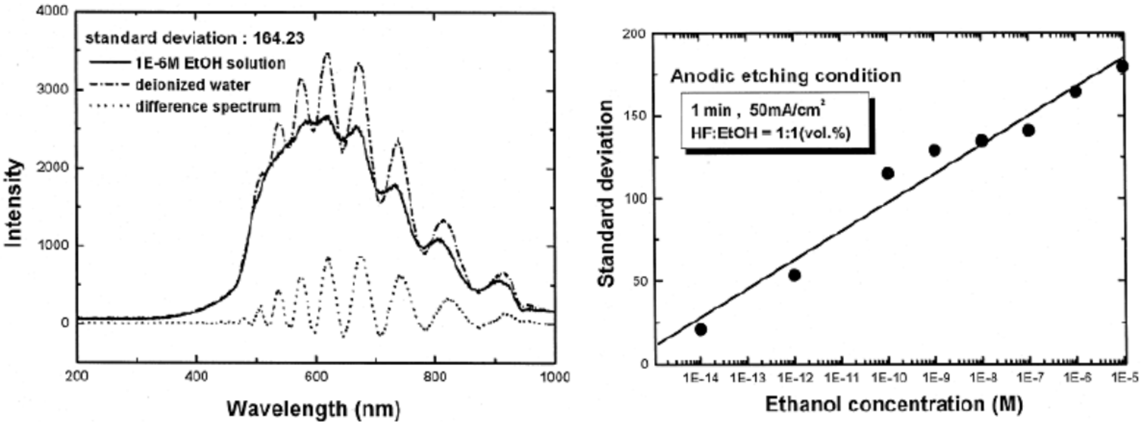


Figure 3.30. Normalized experimental reflectivity spectra showing peak wavelength shift of Bragg mirror after exposure to acetone and chlorobenzene. Reprinted with permission from [88].



(a) (b)

Figure 3.31. a) Fabry–Perot fringe patterns and difference spectra at 1.10^{-6} M ethanol concentrations and b) Standard deviation of difference spectra versus ethanol concentrations. Reprinted with permission from [89].

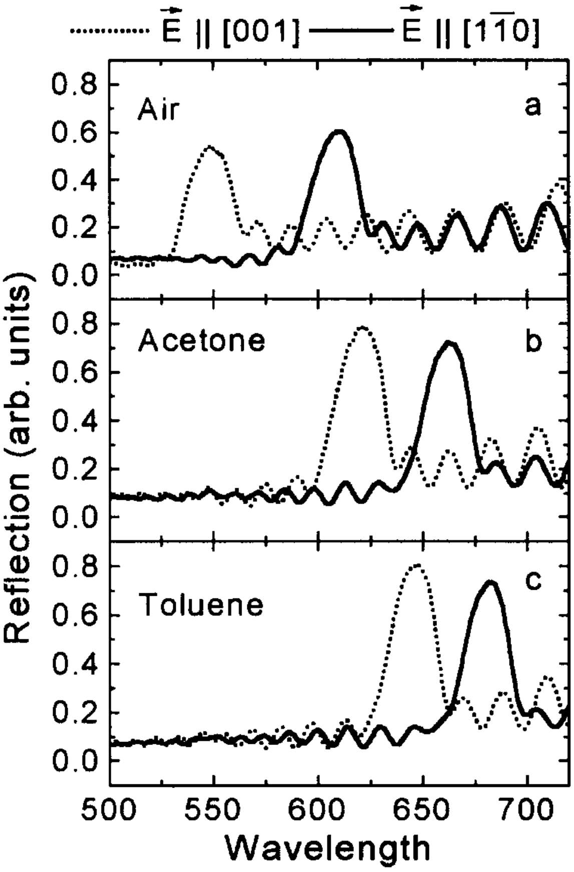


Fig.3.32 Fine tuning of the dichroic behaviour of Bragg reflectors

It should mentioned the dichroic behaviour of Bragg reflectors when reflection depends on light polarisation (fig.3.32).

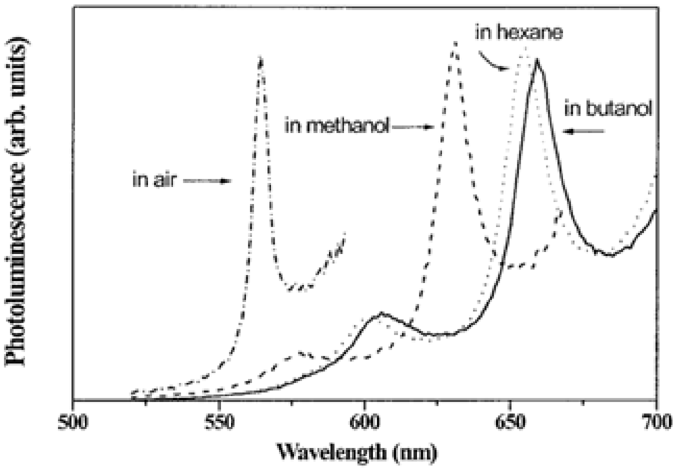
Optical reflectivity spectra of PS Fabry–Perot filters were also applied to detect low concentration of CO_2 (in ppth range) [87], HF gas with detection limit of 30 ppm [86], hydrocarbons [91].

As it was mentioned before, PSM exhibits well resolved Fabry–Perot interference fringes. In the luminescent PSM a narrow peak of photoluminescence (PL) of 5-10 nm width is observed whose spectral position depends on the optical thickness, hence on the refractive index of the cavity (PL peak position $\lambda_c = nd$, where d is the thickness of the central layer), while the PL intensity depends on substrate doping and on the etching and ambient conditions. Marked red shift of the cavity peak, as well as large changes of the integrated PL intensity were observed for PSM in different organic solvents [92].

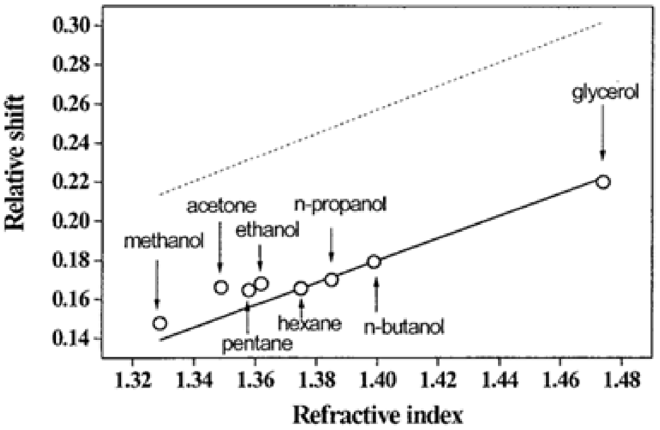
If the refractive index of the environmental gas changes, as, for example, when organic vapor is added to the environment, the PL peak as all Fabry–Perot fringes shift in red region up to 100 of nanometers (Figure 3.33, a). In the case of total filling of the pores the shift can be easily calculated with the Bruggeman effective medium approximation [86, 92]. The results of this calculation are shown together with the experimental data in Figure 3.33,b. Measurements performed in saturated vapors gave the same results obtained for the liquid solvents, confirming that the red shift is due to capillary condensation into pores. As opposed to the peak shift, the calculation of PL intensities in different organic solvents is much more complicated [92]. PL is strongly quenched by high values of dielectric constant.

The effect of ethanol vapor concentration on PL of PSM has been studied [94,93].

During the filling of ethanol, a larger ethanol concentration implies a larger effective refractive index and the position of the peak of the cavity resonance redshifts monotonically. The narrowing of the PL peak in a PSM allows the measuring of small variations of n, i.e. the detection of low gas concentrations.



(a)



(b)

Figure 3.33. a) PL spectra of a PSM initially centered at 570 nm after immersion in different organic solvents. Intensities are not in scale; b) Relative peak shift for total filling versus refractive index for PSM initially centered at 570 nm. Dashed and solid lines: calculated shift without and with inclusion of variation of PS refractive index with wavelength. Reprinted with permission from [92].

Conjugated polymers entrapped in PSM have been studied as optical sensors for low volatility explosives such as trinitrotoluene [94]. The resonance peak of the microcavity reflectance modulates the fluorescence spectra of entrapped polymers in this device. The exposure of the PSM containing entrapped polymer to explosives vapor results in a red shift of the resonance peak, accompanied by the quenching of the fluorescence. The observed redshift in reflectivity is 2–3 nm larger than 1 nm of redshift in fluorescence (fig.3.34).

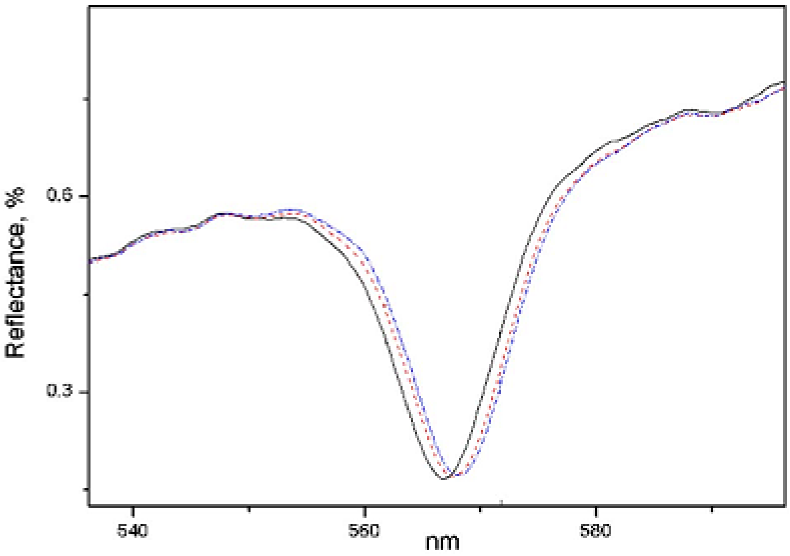


Fig.3.34 Evolution of reflectance of PS microcavity with MEH-PPV under TNT vapor: solid- initial, dash- after 50 sec, dash dot- after 250 sec.

The response time of optic sensors can be very small that is promising to make artificial electronic nose. Both PL quenching and shifts of reflectivity interference fringes of PS Fabry-Perot sample occurs during a couple of seconds (fig.3.35 b). For kinetic measurements of adsorption-desorption the fast-acting CCD detector is usually apllied (fig.3.35 a).

Surfaces of PS chips can be stabilized by an anodic oxidation treatment, leading to long shelf life > 1 month, good baseline stability, and acceptable reversibility [95].

Reflectivity spectra of rugate filter after different surface treatment (alkylated and oxidized) are presented in fig. 3.35. The green solid trace is obtained from the side of the film that contained a mirror etched by using a sinusoidal current varying between 11.5 and 34.6 mA/cm^2. This side of the porous mirror then was hydrosilylated with 1-dodecene. The red dashed trace is obtained from the side of the film that contained a second mirror, etched by using a sinusoidal current varying between 11.5 and 34.6 mA/cm^2 and then removed from the Si substrate. This side of the porous mirror wasthermally oxidized. The total thickness of the porous Si film is 100 μm.

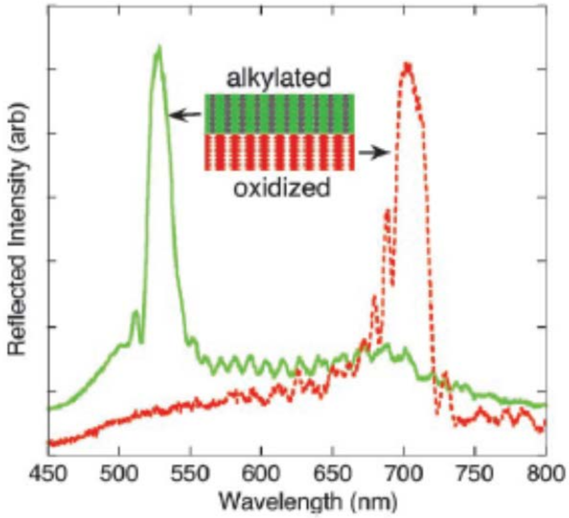


Fig. 3.35**.** Reflectivity spectra of a alkylated and oxidized porous Si rugate film.

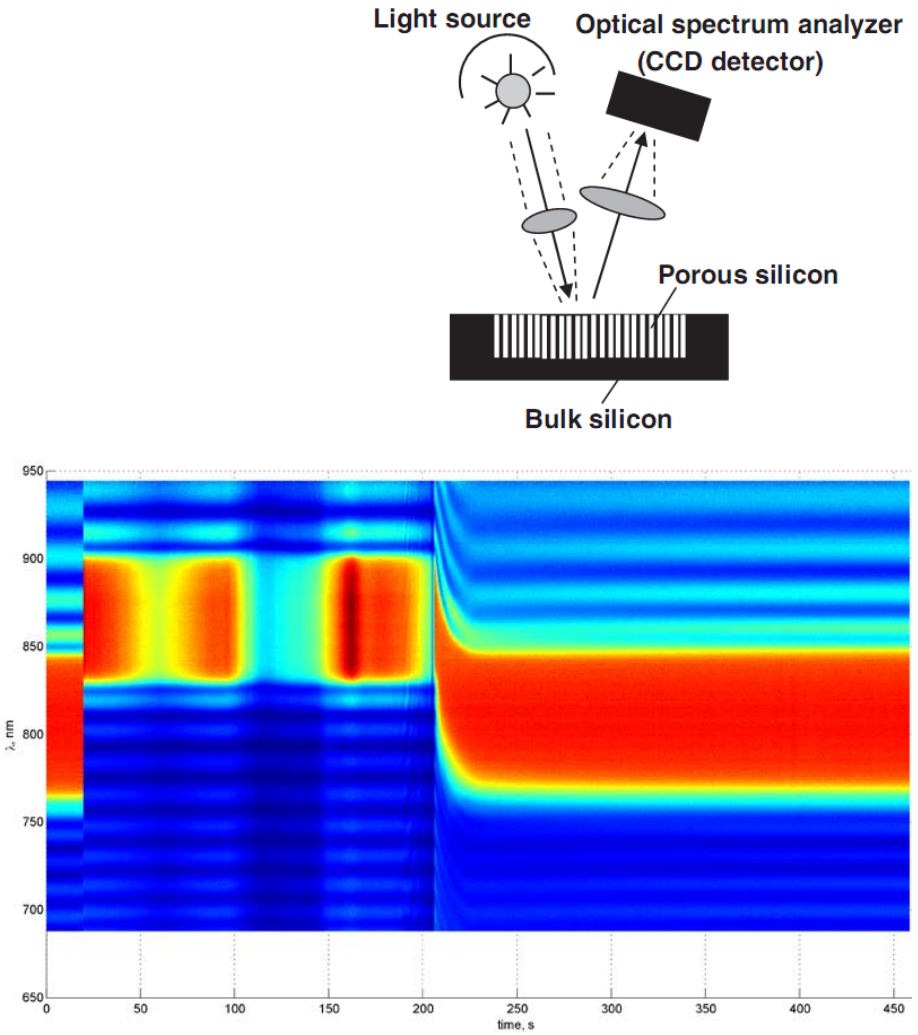


Fig.3.36 The set up for kinetic measurements (a) and relectivity (b) versus time at adsorption-desorption of ethanol in porous silicon Bragg mirror.

Different improved designs of photonic crystals based on PS for sensing application are discussed, among them the methods of PS/polymer micropatterning [96,97]. Micrometer-sized photonic crystals consisting of PS/polymer composite can be prepared by spray coating a fine mist of polymer solution onto PS photonic crystal layer. The uncoated PS is removed by treatment with aqueous base, resulting in composites that retain the spectral interference fringes of the PS photonic crystals but with improved chemical and mechanical stability [96]. The photonic device as planar waveguide in PS with submicrometer lateral patterning was proposed in [98].

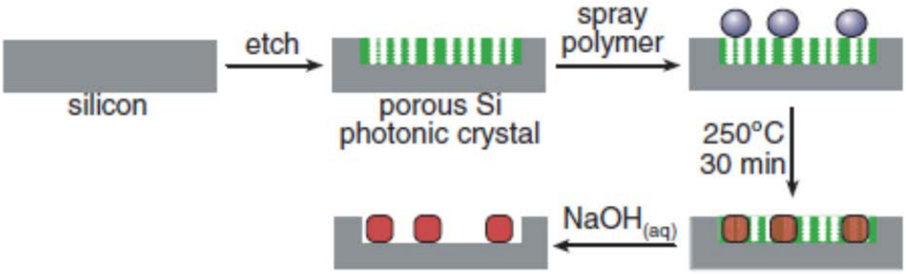


Fig.3.37 Preparation of PS/polymer photonic crystals nanocomposites

The preparation of the polymer nanocomposites follows scheme fig.3.37. This is rugate filter with sinusoidal variation in porosity with depth, which acts as 1D photonic crystal that display sharp resonance in the reflectivity spectum. A fine mist of 16 wt% solution of polystyrene in toluene is then impinged on the surface by spray-coating. The sample is heated in an oven at 250C for 30 min to infuse the polymer into purous matrix.

The method of ellipsometry can also be applied to measure the change in refractive index of film. Ellipsometry is based on measuring of polarization changes occurring upon reflection at onlique incidence of a polarized monochromatic plane wave. The basic quantitaty measured with elliopsometer is the ratio $\rho=\chi_r/\chi_i$, where χ_r and χ_I represents the states of polarization of the reflected and incident beam, respectively. For optically isotropic samples no coupling occurs between the two orthogonal polarizations with electric fields parallel (p-direction) and perpendicular (s-direction) to the plane of incidence as defined in Fig.3.37. The measured quantity is the complex reflectance ratio $\rho = R_p/R_s = tan(\psi) exp(i\Delta)$,where R_p and R_s are the complex reflection coefficients of light parallel and perpendicular to the plane of incidence, respectively. ψ and Δ are called the ellipsometric angles.

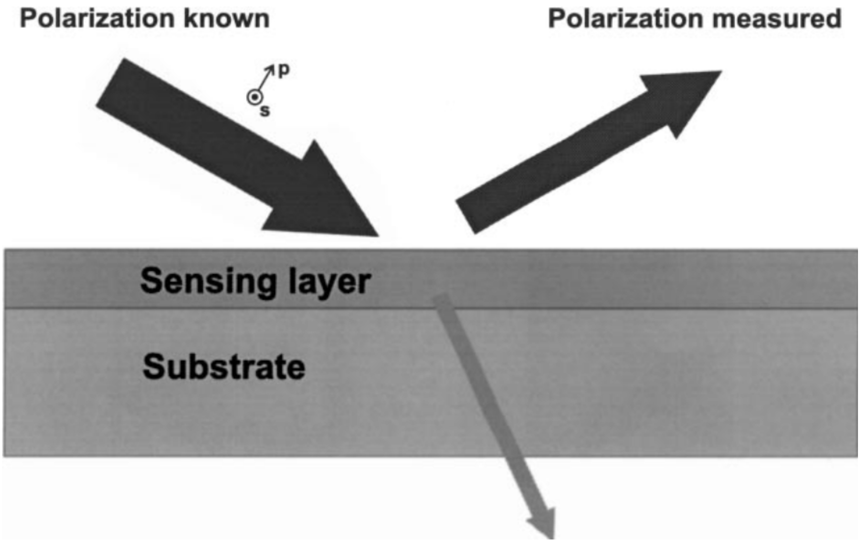


Fig. 3.37 Schematic illustration of ellipsometric sensing

The actual quantities measured are tan(ψ) and cos(Δ) from which the ellipsometric parameters ψ and Δ are obtained. From these instrumental quantities the information on the optical properties in terms the complex refractive index, $n + ik$, or the dielectric function $\varepsilon = (n + ik)^2 = \varepsilon_1 + i\varepsilon_2$ can be obtained. Monitoring over a period of time allows the recording of changes in these parameters caused by adsorption of molecules into pores.

Applicability of ellipsometry for gas sensing was analyzed in [99-102] (fig.3.38, 3.39). For the different types of PS layer, the sensitivity for several gases has been demonstrated including water, ethanol, acetone, toluene. The detection limit threshold for acetone vapors by the layers was 12 ppm. However, as well as other optical methods, ellipsometry of PS is not specific for a certain gas.

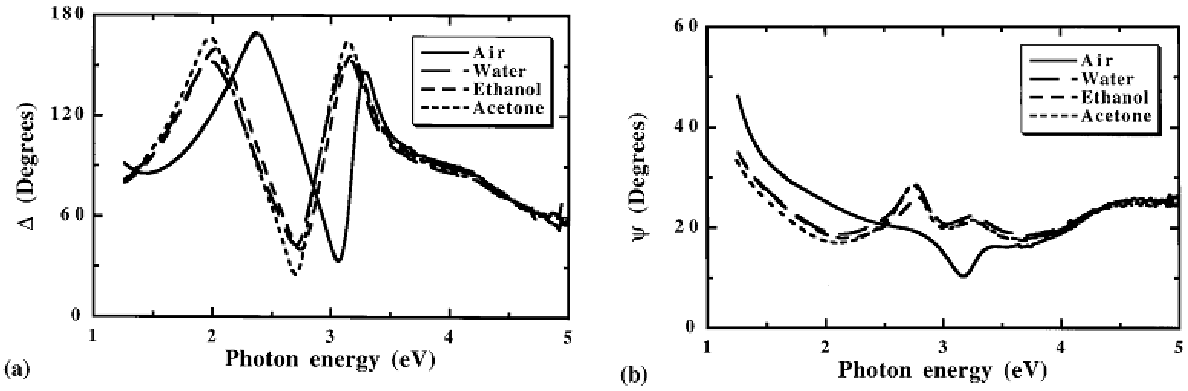


Fig. 3.38. Ellipsometric spectra for a porous silicon layer in air and during exposure to an atmosphere saturated with water, ethanol and acetone, respectively.

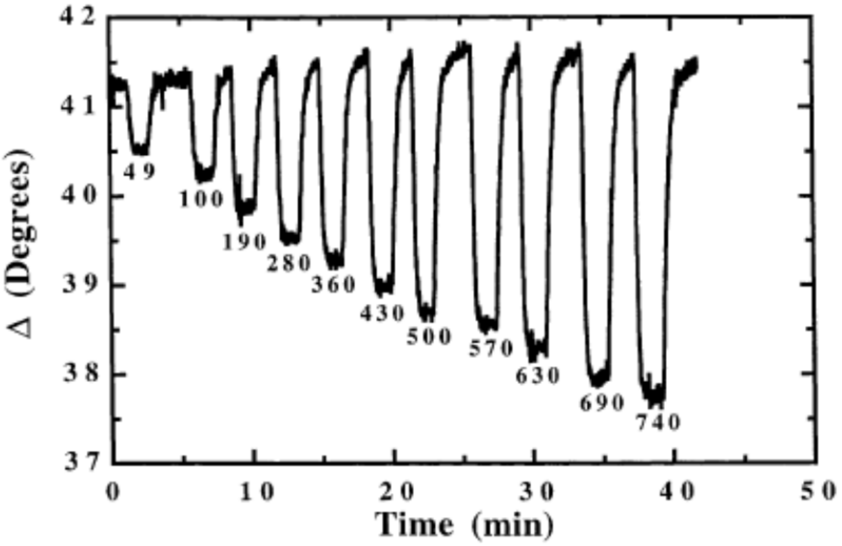


Fig.3.39 Ellipsometric response for 228 nm PS layer exposed to acetone vapor with concentration in range 49-740 ppm.

3.8 Luminescent sensors

Nano and meso-PS with crystallite dimension less than 5 nm shows efficient visible photoluminescence (PL) at UV illumination due to quantum confinement of electrons in semiconductor particles (fig.3.40). It was found that the red-orange PL (S-band) can be strongly quenched at molecule adsorption and that is promising way to make gas sensors [103-115].

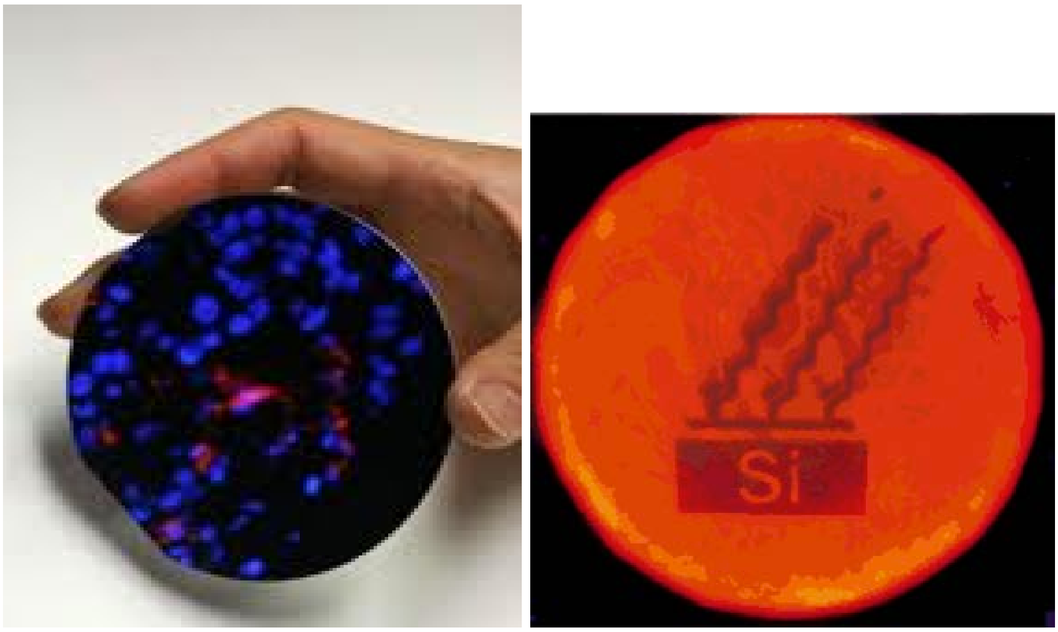


Fig.3.40 Photoluminescence of PD ar room temperature under UV illumination

The phenomenon of PL quenching in PS occurs via different chemical or physical adsorption mechanisms (Fig.3.41). The adsorption of NO_2, amine bases, iodine, bromine, or chlorine vapors results in irreversive PL quenching due to by introduction of surface bonds like Si-I, Si-Br, or Si-Cl that act as efficient non-radiative carrier traps. Si-halogen bonds can be hydrolyzed in air, producing an oxide-terminated PS surface that results in PL recovering [104,108,111].

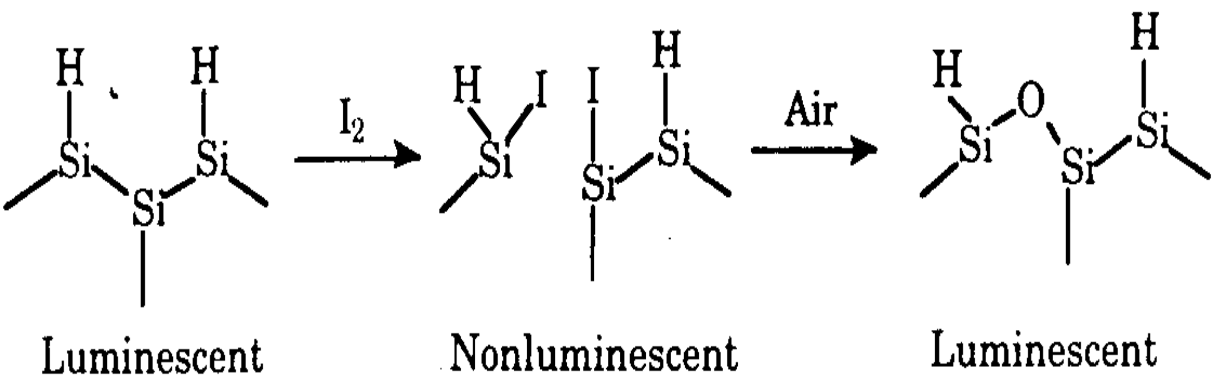


Fig.3.41. Luminescent and non luminescent aditivies for PS.

Reversible adsorption of liquids or vapors of solvents such as alcohol (ethanol, methanol), aliphatic (hexane, pentane), aromatic (pyrene, anthracenes, benzene, toluene), ethers (diethyl ether, tetrahydrofuranor), dichloromethane, causes

PL quenching which is recovered on removal of the solvent. The following PL quenching mechanisms are: the increase of carrier recombination rates at the presence of a dielectric medium outside the silicon nanocrystallites; the arise of non-radiative traps due to strain- induced defects at adsorption, altering the nanoparticle-surface electronic structure, the adsorption of molecules could enhance non-radiative vibronic coupling to surface vibrational modes, some molecules can quench PL of PS by energy or charge - transfer processes.

PL method is efficient to detect low concentration of different gases and vapors with small response time. The molecules NO and NO_2 can be detected by luminescent PS at low ppm and high ppb levels, respectively [111], working SO_2 concentration range that quenches PL is from approximately 400 to 700000 ppm [110], sensitivity of PS photoluminescence to 250 ppm of CO and 5000 ppm of CH_4 was observed [116]. Time response of 3 s and recovery of 30 s was obtained for PL signal at injection of acetone vapor that is much faster comparing with MOS sensors [96]. However, to detect some gases the electrical transducers are preferable. For example, at increasing of methanol and propanol pressure till 200 mbar the PL intensity reduces only by a factor 2, while the conductivity increases by 4 orders of magnitude [105].

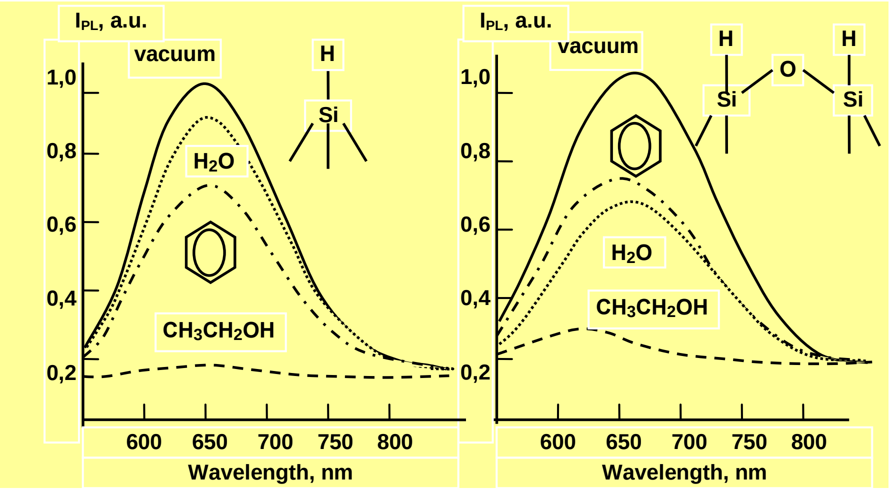


Fig.3.42.PL spectra of as-prepared (a) and oxidized (b) PS in vacuum, H_2O, benzene, ethanol

PL of PS and quenching effect can be tuned by chemical derivatization of the PS surface [118], e.g., at its hydrophobic- hydrophilic transformation. Thus, as-prepared PS has a hydrophobic, hydrogen-terminated surface, and the PL is strongly quenched by ethanol and weakly quenched by water. Mild chemical oxidation (iodine followed by hydrolysis) produces a hydrophilic PS surface; the PL is quenched to a lesser extent by ethanol and to a greater extent by water, relative to the original surface [104]. (fig.3.42).

Derivatization of the PS surface with hydroquinone imparts a greater sensitivity to water vapor and reduces the sensitivity to benzene [119]. Such approach can be useful to improve the selectivity of PS gas sensors. One of the important drawbacks of these sensors is that their working conditions include the UV illumination of PS layer often in corrosive environments. Well-known, the photoexcitation of a high density of electron-hole pairs in a semiconductor material can give rise to possible destructive effects which are usefully classified as photochemical, photothermal, photoelectric and photostructural. All of these four mechanism categories can contribute to significant degradation ‘‘fatigue’’of PL. Thus, the illumination in oxidizing ambient results in photoenhanced oxidation process that substantially decrease PL output [112].

PL based sensors with improved selectivity to solvents can be designed. As it was shown, if PS is immersed in different solvents the PL intensity at maximum of the spectrum is steeply varying function of dielectric constant: PL intensity decreases by more than three orders of magnitude when the dielectric constant of the solvent increases from 2 to 20 (Fig.3.43). The weak variation of the lifetime indicates that the PL quenching in solvents is mostly due to a decrease of the probability of close-pair formation. PL is strongly quenched by high values of dielectric constant, because the Coulomb interaction experienced by the electron-hole pairs is diminished, affecting the radiative recombination rate [113,120].

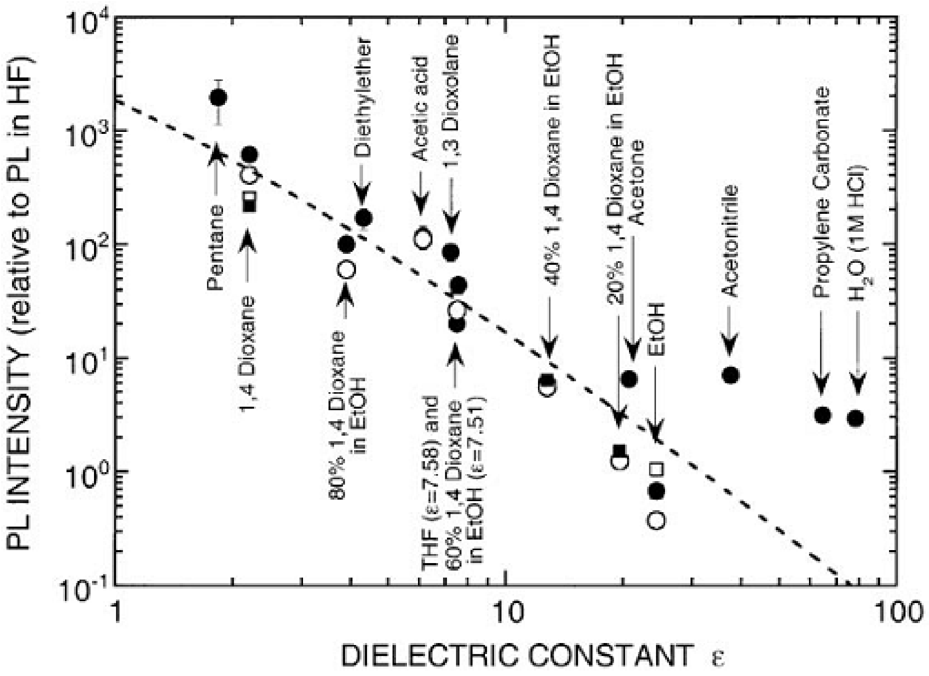


Fig.3.43. PL intensity at maximum emission wavelength as a function of dielectric constant for two emission wavelengths (open circuit – 800 nm, close circuit – 550 nm).

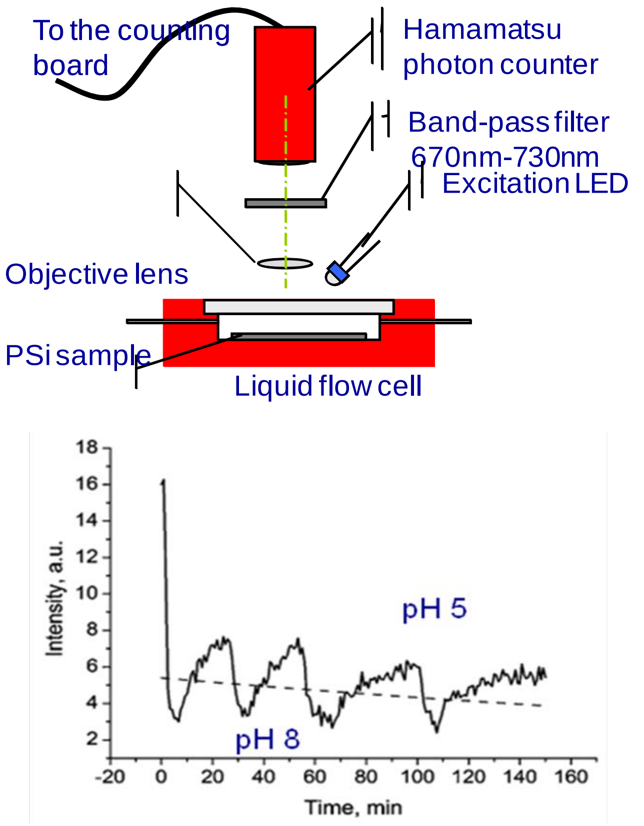


Fig.3.44 a) Set up for measurement of PS photoluminescence kinetics in liquids and b) dependence of photoluminescence intensity on pH solution

PL is as sensing method to detect the pH level of buffer solutions [121] (fig.3.44). Changes of pH value lead to corresponding changes in intensity and lifetime of PS luminescence. PS layer degradation (etching) is observed while the samples are being placed in alkaline solution. The supplementary coating of PS surface by PEDOT (poly-3,4-ethylenedioxythiophene) film prevents against this PL degradation [121]. The covered by PEDOT samples display better sensitivity and stability to pH change. PEDOT seems to be applicable for stabilizing of PS preventing somehow PS layer oxidation and etching and allowing at the same time free access of H^+ ions to the surface. The possible explanation of sensitivity increasing is the swelling of the polymer insides the pores, increasing capillary condensation effects or passivation effects. To improve the stability of PS luminescence another polymers, e.g. polymethylmethacrylate can be applied [101,102]. PS covered by this film exhibits PL quenching and complete recovery in contrast to as-prepared PS in which the quenching is irreversible.

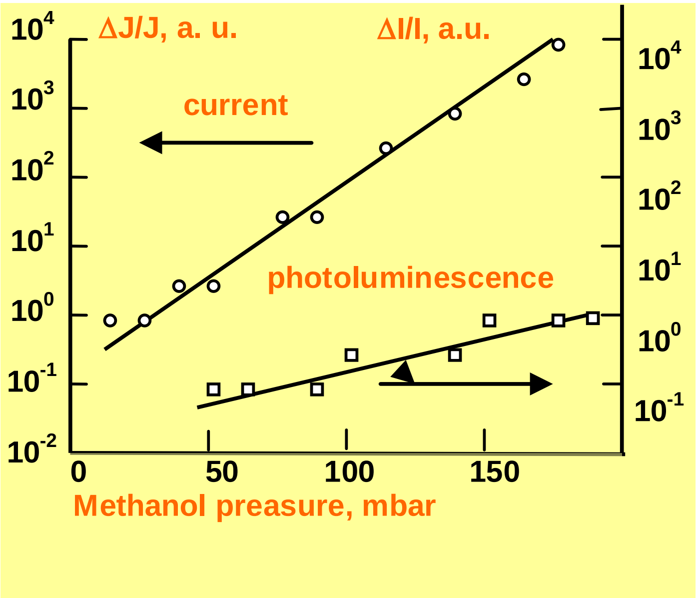


Fig.3.45 Comparison of methanol sensitivity for photoluminescent and electrical transdusers .

The photoluminescent transdusers reveals some advantages (no signal interference, simple signal processing and read out, low noise, etc.). However, their sensitivity is lower comparing with lectrical one (fig.3.45).

CONCLUSION

The advantages of PS application for gas sensor are proved by the several reasons: PS has a huge internal surface area (up to 500-1000 m^2cm^{-2}) and high surface chemical reactivity that can enhance the adsorption efficiency to the gas of interests; it allows to use a number of transduser principles (electrical, optical); simplicity and cheapness of the PS preparation that is compatible with silicon IC technology; there is a possibility to make 3D device and multisensors. It allows to detect very low concentration of CO, NO_2, different volatile organic compounds, alcohols, H_2S, H_2 and organic solutions. The PS gas sensors parameters are among the best gas sensors of different types. However, the PS sensors have some drawbacks. The irreversible PS surface deviation during aging, contact with humidity, UV ilumination leads to the instability of optical and electrical properties of PS and sensors. Thefore, the using of diferent passivation treatments of porous surface is necessary. Selectivity and sensitivity of PS sensors can be enhanced using catalyst materials and using the complex structure as ISFET, Bragg –mirrors, etc.

REFERENCES

[1] Sailor, M. J. (2012). Porous silicon in Practice. Preparation, Characterisation and Application, Wiley-VCH, Weinheim.

[2] Parkhutik, V. (1999). *Solid State Electron.*, 43, pp.1121-1141.

[3] Bisi, O., Ossicini, S., Pavesi, L. (2000), *Surf.Sci.Rep,* 38, pp. 1-126.

[4] Marsh, G. (2002). *Materials Today,* 5, pp. 36-41.

[5] Ozdemir, S. and Gole, J. L. (2007).*Current Opinion in Solid State and Materials Science*, 11, pp.92-100.

[6] Saha, H. (2008)., *Int. J. Smart Sens. Intell. System*, 1, pp. 36-56.

[7] Korotcenkov, G. and Cho, B. K. (2010). *Critic. Rev. Solid State &Mater. Sci.* 35, pp.1-37.

[8] Han, P. G., Wong, H. and Poon, M.C. (2001). *Colloids and Surfaces,* A179, pp.171–175.

[9] Baratto, C., Faglia, G., Comini, E., Sberveglieri, G., Taroni, A., La Ferrara, V., Quercia, L. and Di Francia, G. (2001), *Sensors & Actuators*, B 77, pp. 62–66.

[10] Valera, E., Casals, O., Vetter, M. and Rodriguez, A. (2007). *Spanish Conf. Electron Dev.*, Madrid, pp. 197 – 200.

[11] Baratto, C., Sberveglieri, G., Comini, E., Faglia, G., Benussi, G., La Ferrara, V., Quercia, L., Di Francia, G., Guidi, V., Vincenzi, D. , Boscarino, D. and Rigato V. (2000) *Sensors & Actuators,* B 68, pp. 74–80.

[12] Boarino, L., Baratto, C., Geobaldo, F., Amato, G., Comini, E. , Rossi, A.M., Faglia, G., Le'rondel, G. and Sberveglieri G. (2000), *Mat. Sci. Eng*., B 69–70, pp.210–214.

[13] Chakane, S., Gokarna, A. and Bhoraskar, S.V. (2003). *Sensors & Actuators*, B 92, pp. 1–5.

[14] Pancheri, L., Oton, C.J., Gaburro, Z., Soncini, G. and Pavesi, L. (2004). *Sensors & Actuators*, B97, pp. 45–48.

[15] Massera, E., Nasti, I., Quercia, L., Rea, I. and Di Francia, G. (2004). *Sensors & Actuators*, B102, pp.195-197.

[16] Lewis, S. E., DeBoer, J. R., Gole, J. L. and Hesketh, P. J. (2005). *Sensors & Actuators*, B 110, pp. 54–65.

[17] Pancheri, L., Oton, C.J., Gaburro, Z., Soncini, G. and Pavesi, L. (2003). *Sensors & Actuators*, B89, pp. 237-239.

[18] Gaburro, Z., Bettotti, P., Saiani, M., Pavesi, L., Pancheri, L., Oton, C.J. and Capuj, N. (2004). *Appl. Phys. Lett*., 85, pp. 555-557.

[19] Iraji zad, A., Rahimi, F., Chavoshi, M. and Ahadian, M.M. (2004). *Sensors & Actuators*, B 100, pp. 341–346.

[20] Timoshenko, V.Yu., Dittrich, Th., Lysenko, V., Lisachenko, M.G. and Koch, F. (2001). *Phys.Rev*, B 64, p.085314.

[21] Seals, L., Gole, J. L., Tse, L. A. and. Hesketh, P. J. (2002)., *J. Appl. Phys*., 91, pp.2519-2523.

[22] Baratto, C., Comini, E., Faglia, G. , Sberveglieri, G., Di Francia, G., De Filippo, F., La Ferrara, V., Quercia, L. and Lancellotti, L. (2000), *Sensors & Actuators*, B 65, pp. 257–259.

[23] Anderson, R. C., Muller, R. S. and Tobias, C. W. (1990), *Sensors &Actuators*, A 21–23, pp. 835–839.

[24] Kim, S.-J., Park, J.-Y., Lee, S.-H. and Yi, S.-H. (2000), *J. Phys. D: Appl. Phys*., 33, pp. 1781–1784.

[25] Kim, S.-J., Jeon, B.-H., Choi, K.-S. and Min, N.-K. (2000). *J. Solid State Electrochem*, 4, pp.363-366.

[26] Foucaran, A., Pascal-Delannoy, F., Giani, A., Sackda, A., Combette, P. and Boyer, A. (1997). *Thin Solid Films*, 297, pp. 317–320.

[27] Foucaran, A., Sorli, B., Garcia, M., Pascal-Delannoy, F., Giani, A. and Boyer, A. (2000). *Sensors & Actuators*, A79, pp.189–193.

[28] Xu, Y. Y., Li, X. J., He, J.T., Hu, X. and Wang, H. Y. (2005). *Sensors & Actuators*, B 105, pp.219-222.

[29] Bjorkqvist, M., Paski, J., Salonen, J. and Lehto, V.-P. (2006), *IEEE Sensor. J*., 6, pp. 542 – 547.

[30] RoyChaudhuri, C., Gangopadhyay, S., DevDas, R., Datta, S.K. and Saha, H. (2008). *Int. J. Smart Sens. Intell. System*, 1, pp.638-658.

[31] Wang, Y., Park, S., Yeow, J.T.W., Langner. A., and F. Muller. (2010). *Sensors & Actuators*, B 149, pp. 136–142.

[32] Manilov A. I., Skryshevsky, V. A. (2013). *Mat.Sci.Eng.B*, 178, pp.942-955.

[33] Stievenard, D. and Deresmes, D. (1995). *Appl. Phys. Lett*. 67, pp. 1570-1572.

[34] Matsumoto, T., Mimura, H., Koshida, N. and Masumoto., Y. (1999). *Jpn.J.Appl.Phys*, 38, pp.539-541.

[35] Lee, W.H., Lee, C., Kwon, Y.H., Hong, C.Y. and Cho, H.Y. (2000). *Solid State Commun*, 113, pp.519-522.

[36] Skryshevsky, V.A., Zinchuk, V.M., Benilov, A.I., Milovanov, Yu. S., Tretyak, O.V. (2006). *Semicond. Sci. Technol*., 21, pp. 1605-1608.

[37] Hao, P. H., Hou, X. Y., Zhang, F. L. and Wang, X. (1994). *Appl.Phys.Lett*, 64, pp. 3602-3604.

[38] Deresmes, D., Marissael, V., Stievenard, D. and Ortega, C. (1995). *Thin Solid Films*, 255, pp.258-261.

[39] Dimitrov, D. B. (1995). *Phys.Rev*, B 51, pp. 1562-1566.

[40] Ben-Chorin, M., Moller, F. and Koch, F. (1995), *J. Appl. Phys*., 77, pp.4482-4488.

[41] Martin-Palma, R.J., Perez-Rigueiro, J., Guerrero-Lemus, R., Moreno, J.D. and Martinez-Duart, J.M. (1999). *J. Appl. Phys*., 85, pp.583-586.

[42] Vikulov, V. A., Strikha, V. I., Skryshevsky, V. A., Kilchitskaya, S. S., Souteyrand, E. and. Martin, J.-R. (2000). *J. Phys. D: Appl. Phys.*, 33, pp. 1957-1964.

[43] Tretyak, O. V., Skryshevsky, V. A., Vikulov, V. A., Boyko, Yu. V. and Zinchuk, V. M. (2003). *Thin Solid Films*, 445, pp. 144–150.

[44] Lundstrom, I. and Soderberg, D. (1981/1982). *Sensors & Actuators*, 2, pp.105-138.

[45] Kohl, D. (2001). *J. Phys.D: Appl. Phys.*, 34, pp. R125-R149.

[46] Polishchuk, V., Souteyrand, E., Martin, J.-R., Strikha, V. and Skryshevsky, V. (1998). *Anal.Chem.Acta*, 375, pp. 205-210.

[47] Litovchenko, V.G., Gorbanyuk, T.I., Solntsev, V.S. and Evtukh, A.A. (2004). *Appl. Surf. Sci*., 234, pp. 262-267.

[48] Luongo, K., Sine, A. and Bhansal, S. (2005), *Sensors &Actuators*, B 111–112, pp. 125–129.

[49] Skryshevsky, V. A., Polischuk, V., Manilov, A. I., Gavrilchenko, I. V. and Skryshevsky, R. V. (2008). *Sensor Electron. Microsys. Technol*., 2, pp.21-27.

[50] Solntsev, V.S., Litovchenko, V.G., Gorbanyuk, T.I. and Evtukh, A.A. (2008). *Semicond. Phys,. Quant. Electron. & Optoelectron.*, 11, pp. 381-384.

[51] Lachenani, H., Cheraga, H., Menari, H. and Gabouze, N. (2009). *Mat. Sci. Forum*, 609, pp. 225-230.

[52] Joshi, R. K., Krishnan, S., Yoshimura, M. and Kumar, A. (2009). *Nanoscale Res. Lett*., 4, pp. 1191–1196.

[53] Zhang, W., de Vasconcelos, E. A., Uchida, H., Katsube, T., Nakatsubo, T. and Nishioka, Y. (2000), *Sensors & Actuators*, B 65, pp.154–156.

[54] Alwan, A. M. (2007). *Eng. & Technol*., 25, pp.1023-1027.

[55] Strikha, V., Skryshevsky, V.A, Polishchuk, V., Souteyrand, E. and Martin, J.-R. (2000). *J. Porous Mat.*, 7, pp. 111–114.

[56] O'Halloran, G. M., Kuhl, M., Trimp, P. J. and French P. J. (1997). *Sensors & Actuators,* A 61, pp. 415 -420.

[57] Subramanian, N. S., Sabaapathy, R. V., Vickraman, P., Kumar, G. V., Sriram, R. and Santhi, B. (2007). *Ionics*, 13, pp.323–328.

[58] Belhousse, S., Cheraga, H., Gabouze, N. and Outamzabet, R. (2004). *Sensors & Actuators*, B 100, pp. 250–255.

[59] Chiali, A., Ghellai, N., Gabouze, N. and Sari, N.E.C. (2010). *Adv.Mater. Sci.*, 10, pp.4-10.

[60] De Rosa, R., Di Francia, G., La Ferrara, V., Quercia, L., Roca, F. and Tucci, M. (2000). *Phys.Stat.Sol*, A 182, pp. 489-493.

[61] Tucci, M., La Ferrara, V., Della Noce, M., Massera, E., Quercia, L. (2004). *J. Non-Cryst. Solids*, 338–340, pp. 776–779.

[62] Barillaro, G., Nannini, A. and Pieri, F. (2003). *Sensors & Actuators*, B 93, pp. 263–270.

[63] Barillaro, G., Diligenti, A., Nannini, A., Strambini, L.M., Comini, E. and Sberveglieri, G. (2006). *IEEE Sensors J.*, 6, pp. 19 – 23.

[64] Zouadi, N., Gabouze, N., Bradai, D. and Dahmane, D. (2009). *Mater.Sci.Forum*, 609, pp. 261-264.

[65] Barillaro, G., Diligenti, A., Marola, G. and Strambini, L.M. (2005), *Sensors & Actuators*, B 105, pp.278-282.

[66] Barillaro, G., Diligenti, A., Nannini, A., and Strambini L. M. (2005)., *Phys. Stat. Solidi.*, C2, pp. 3424–3428.

[67] Mizsei, J., Shrair, J.A. and Zolomy, I. (2004). *Appl. Surf. Sci.*, 235, pp.376-388.

[68] Reddy, R. R. K., Basu, I., Bhattacharya, E. and Chadha, A. (2003). *Current Appl. Phys.*, 3, pp. 155–161.

[69] Schoning, M.J., Simonis, A., Ruge, C., Ecken, H, Muller-Veggian, M. and Luth, H. (2002). *Sensors*, 2, pp. 11-22.

[70] Zehfroosh, N., Shahmohammdi, M., Mohajerzadeh, S. (2010). *In Nanotechnology 2010: Electronics, Devices, Fabrication, MEMS, Fluidics and Computational*, pp. 390 – 393.

[71] Schoning, M. J., Malkoc, U., Thust, M., Steffen, A., Kordos, P. and Luth, H. (2000). *Sensors & Actuators*, B 65, pp. 288–290.

[72] Schoning, M. J. (2005). *Sensors*, 5, pp. 126-138.

[73] Chung, D.W.-Y., Yang, C.-H. and Wang, M.-G. (2004). *Proc.SPIE*, 5586, SPIE, Bellingham, WA, pp.18-25.

[74] Yoshinobu, T., Schoning, M.J., Finger, F., Moritz, W. and Iwasaki, H. (2004). *Sensors*, 4, pp.163-169.

[75] Lundstrom, I., Erlandsson, R., Frykman, U., Hedborg, E., Spetz, A., Sundgren, H., Welin, S. and Winquist, F. (1991), *Nature*, 352, pp. 47-50.

[76] Yoshinobu, T., Ecken, H., Poghossian, A., Luth, H., Iwasaki, H. and Schoning, M.J. (2001). *Sensors & Actuators*, B 76, pp.388-392.

[77] Steiner, P. and Lang, W. (1995). *Thin Solid Films*, 255, pp.52-58.

[78] Splinter, A., Bartels, O., Benecke, W. (2001). *Sensors & Actuators*, B76, pp.354-360.

[79] Baratto, C., Faglia, G., Sberveglieri, G., Boarino, L., Rossi, A.M. and Amato, G. (2001), *Thin Solid Films*, 391, pp. 261-264.

[80] Angelucci, R., Poggi, A., Dori, L., Cardinali, G.C., Parisini, A., Tagliani, A., Mariasaldi, M., Cavani, F. (1999). *Sensors & Actuators,* A74, pp.95–99.

[81] Parisini, A., Angelucci, R., Dori, L., Poggi, A., Maccagnani, P., Cardinali, G.C., Amato, G., Lerondel, G. and Midellino, D. (2000). *Micron,* 31, pp. 223–230.

[82] Clarkson, J.P., Fauchet, P.M., Rajalingam, V. and Hirschman, K. D. (2007). *IEEE Sensors J.*, 7, pp. 329-335.

[83] Chan, S., Fauchet, P.M., Li, Y., Rothberg, L.J. and Miller, B.L. (2000). *Phys. Stat Solidi*, A 182, pp. 541-546.

[84] Bettotti, P., Cazzanelli, M., Dal Negro, L., Danese, B., Gaburro, Z., Oton, C. J., Prakash, G. V. and Pavesi, L. (2002). *J. Phys.: Condens. Matter.*, 14, pp. 8253–8281.

[85] Ouyang, H. and P. M. Fauchet, *In Photonic Crystals and Photonic Crystal Fibers for Sensing Applications*, H. H. Du editor (SPIE, Bellingham, WA, 2005), pp 600508 1-15.

[86] I.I. Ivanov, V.A. Skryshevsky, T. Serdiuk, V. Lysenko, *Sensors and Actuators B*, 174, p.521-526, 2012.

[87] Rocchia, M., Garrone, E., Geobaldo, F., Boarino, L. and Sailor, M. J. (2003). *Phys. Stat. Solidi* , A 197, pp. 365–369.

[88] Snow, P.A., Squire, E.K., Russek, P.St.J. and Canham, L.T. (1999). *J.Appl.Phys*., 86 , pp.1781-1784.

[89] Min, H.K., Yang, H.-S. and Cho, S. M. (2000). *Sensors & Actuators,* B 67, pp. 199–202.

[90] Gao, J., Gao, T. and Sailor, M. J. (2000). *Appl.Phys. Lett*., 77, pp.901-903.

[91] De Stefano, L., Moretti, L., Rendina, I. and Rossi, A.M. (2003). *Sensors & Actuators*, A 104, pp.179-182.

[92] Mulloni, V., Gaburro, Z. and Pavesi, L. (2000). *Phys.Stat.Solidi,* A 182, pp. 479-484.

[93] Gaburro, Z., Daldosso, N. and Pavesi, L. (2001). *Appl. Phys. Lett*., 78, pp.3744-3747.

[94] Levitsky, I. A., Euler, W. B., Tokranova, N. and. Rose, A (2007). *Appl. Phys. Lett*., 90, pp.041904.

[95] Letant, S.E., Content, S., Tan, T. T., Zenhaausern, F. and Sailor, M. J. (2000). *Sensors & Actuators,* B 69, pp. 193-198.

[96] Li, Y. Y., Kollengode, V. S. and Sailor, M. (2005). *Adv.Mater*, 17, pp.1249-1251.

[97] Gargas, D. J., Muresan, O., Sirbuly, D. J. and Buratto, S. K. (2006). *Adv. Mater*., 18, pp.3164–3168.

[98] Jamois, C., Li, C., Skryshevskyi, R., Orobtchouk, R., Chevolot, Y., Monnier, V., Souteyrand, E. and Benyattou, T. (2010). *Proc. 7th Int.Conf. Porous semicond. - Science and technol,* Valencia, Spain, pp. 126-127.

[99] Zangooie, S., Bjorklund, R. and Arwin, H. (1997). *Sensors & Actuators*, B 43, pp. 168–174.

[100] Zangooie, S., Jansson, R. and Arwin, H. (1999). *J. Appl. Phys*., 86, pp. 850-858.

[101] Arwin H. (2001). *Sensors & Actuators*, A 92, pp. 43-51.

[102] Wang, G. and Arwin, H. (2002). *Sensors & Actuators*, B 85, pp, 95-103,

[104] Lauerhaus, J. M., Credo, G. M., Heinrich, J. L. and Sailor, M. J. (1992). *J. Am. Chem. Soc*. 114, pp.1911-1912.

[105] Ben-Chorin, M. and Kux, A. (1994). *Appl. Phys. Lett*., 67, pp.481-483.

[106] Fisher, D. L., Harper, J., and Sailor, M. J. (1995). *J. Am. Chem. Soc*., 117, pp. 7846-7847.

[107] Schechter, I., Ben-Chorin, M. and Kux, A. (1995). *Anal. Chem.*, 67, pp.3727-3732.

[108] Chandler-Henderson, R.R., Sweryda-Krawiec, B. and Coffer, J. L. (1995). *J. Phys. Chem*., 99, pp. 8851-8855.

[109] Dittrich, T., Konstantinova, E.A. and Timoshenko, V. Ya. (1995). *Thin Solid Films*, 255, pp. 238-240.

[110] Kelly, M.T., Chun, J.K.M. and A.B. Bocarsly, (1996). *Nature*, 382, pp.214-215.

[111] Harper, J. and Sailor, M.J. (1996). *Anal. Chem*., 68, pp.3713-3717.

[112] Cullis, A. G., Canham, L. T. and Calcott, P. D. G. (1997). *J. Appl. Phys*., 82, pp. 909 -965.

[113] Fellah, S., Ozanam, F., Gabouze, N., and Chazalviel J. N. (2000). *Phys.Stat.Solidi*, A182, pp. 367-372.

[114] Skryshevsky, V. A. (2000). *Appl. Surf. Sci*., 157, pp.145-150.

[115] Salcedo, W. J., Ramirez Fernandez, F. J. and Rubim, J. C. (2004). *Spectrochim. Acta*, A 60, pp.1065-1070.

[116] Di Francia, G., La Ferrara, V., Quercia, L. and Faglia, G. (2000). *J.Por.Mater*, 7, pp.287-290.

[117] Letant, S.E., Content, S., Tan, T.T., Zenhaausern, F. and Sailor, M.J. (2000). *Sensors & Actuators*, B 69, pp. 193-198.

[118] Stewart, M.P. and Buriak, J. (2000)., *Adv.Mat*., 12, pp. 859-869.

[119] Harper, J. and Sailor, M.J. (1997). *J. Am. Chem. Soc*., 119, pp. 6943-6944.

[120] Fellah, S., Wehrspohn, R. B., Gabouze, N., Ozanam, F. and Chazalviel, J. N. (1999)., *J. Lumin*., 80, pp.109-113.

[121] Benilov, A. I. , Gavrilchenko, I. V, Benilova, I. V. , Skryshevsky, V. A. and Cabrera, M. (2007). *Sensors & Actuators*, A137, pp.345-349.

[122] Bakker, J. W., Arwin, H., Wang, G. and Jarrendah K. (2003). *Phys.Stat.Solidi*, A197, pp. 378-381.

[123] Singh, P. K. and Lakshmikumar, S. T. (2002). *Semicond. Sci. Technol*.,17, pp. 1123–1127.

Chapter 4

Silver nanoparticle ink

Printed electronics is a low-cost alternative to standard silicon technology. Several applications are of interest, for example, as an electronic replacement to the simple barcode tags printed on products for identification. Price is the limiting factor and while a bar-code cost around 0.5 cent to print on a box, the electronic version, passive radio frequency identification (RFID) tags from the silicon industry, have now reached a lowest price of 5 cents, which is a too high price for most applications. Printing with ink is a cheaper technology compared with silicon chip production and therefore the concept of printed electronics could serve as the needed low-cost alternative.

Printing by depositing molten metal droplets is problematic, mainly due to the technical difficulties at the high temperature required, leading to problems with adhesion or oxidation of the metal.

Inkjet technology is therefore used in most approaches, mainly for printing electrically conductive pathways but also for more advanced structures like thin film transistors. Commercially available inkjet printers are used together with different types of inks, mainly based on nanoparticles. These nanoparticle inks consist of a solvent (organic or inorganic) together with metallic nanoparticles usually made out of silver, gold, and recently copper.

The nanoparticles are typically in the range 1–100 nm. The smallest sizes, i.e. less than 10 nm, have the advantage of lower melting temperatures. The solvent serves as a stabilizing matrix for the nanoparticles and forms protective shells around them to prevent them from coalescence. The solvent is also designed to satisfy the requirements of the printing nozzle such as drying and viscosity constrains. Printing using nanoparticle inks results in a too low electrical conductivity of the printed structures, especially for low resistance demanding structures such as antennas. To directly melt the nanoparticles in order to increase the conductivity would demand a too high temperature (962uC for silver), for most printed substrates, such as paper, to survive. Therefore a sintering step is used to improve the conductivity, usually the sintering temperature is between 50% to 80% of the melting temperature of the material, but for nanosized powders the sintering temperature can be even lower, i.e. 20% to 30% of the melting temperature.

Usually a furnace is used for sintering, typically at 100–300uC. This heating step does also introduce substrate shrinkage and is time consuming. Control of the percolation threshold is important to prevent too much substrate shrinkage which can result in microcracking. Visible-near IR spectroscopic ellipsometry has been suggested as one method to probe percolation transition during the sintering.

To replace the furnace, laser, microwave radiation, light flashing by utilizing the photothermal effect of nanostructures, and chemical methods have been developed.

Printed electronics will bring to the consumer level great breakthroughs and unique products in the near future, shifting the usual paradigm of electronic devices and circuit boards from hard boxes and rigid sheets into flexible thin layers and bringing disposable electronics, smart tags, and so on. The most promising tool to achieve the target depends upon the availability of nanotechnology-based functional inks. A certain delay in the innovation-transfer process to the market is now being observed. Nevertheless, the most widely diffused product, settled technology, and the highest sales volumes are related to the silver nanoparticle-based ink market, representing the best example of commercial nanotechnology today. This is a compact review on synthesis routes, main properties, and practical applications.

Silver nanoparticle (NP)-based inks represent the most important commercial nanotechnology-derived product and the most widely studied worldwide. To better clarify the motivation of this review, we should therefore focus on the three points highlighted: the raw material (Ag), the morphology it takes (NPs), and the compound through which it is used in practical applications (ink).

Let us start with silver. Why Ag in place of other raw materials? Because it is a noble metal, featuring undisputed advantages in terms of electrical conductivity, resistance to oxidation, and providing interesting plasmonic and antibacterial properties, as we will see further in the text. The topic is far too wide to be synthesized in a single sentence, and there is no single source from which to extract information regarding the different materials that could be used to prepare conductive inks (Au, Cu, brass, nickel, Cr, Fe, Ti, intrinsically conductive polymers, thin conductive oxides, carbon-based materials). It is difficult to imagine a future without the use of Ag, at least for certain critical systems that cannot lose efficiency. The market share will be reduced in favor of other nanoengineered, less expensive materials, but it is not possible to avoid the use of metals to transport electricity without losses.

Then, why NPs? The most important feature is connected with their scale, bringing surface tension and ionic forces to that level of importance that allows a play against gravity, giving stability to a suspension. But many other interesting phenomena occur: collective electron resonances, the so-called plasma waves enabling surface plasmon resonance (SPR), and interactions with the electromagnetic field; a huge enhancement of diffusivity of the surface atoms, enabling “melting” (sintering) at extremely low temperatures, and so on. In a world where nanoengineered materials could have broad application, from the consumer electronics market, building industry, pharmaceutical and cosmetic products, to food and the environment, we can easily imagine that the role of NP synthesis and modification activities will be huge.

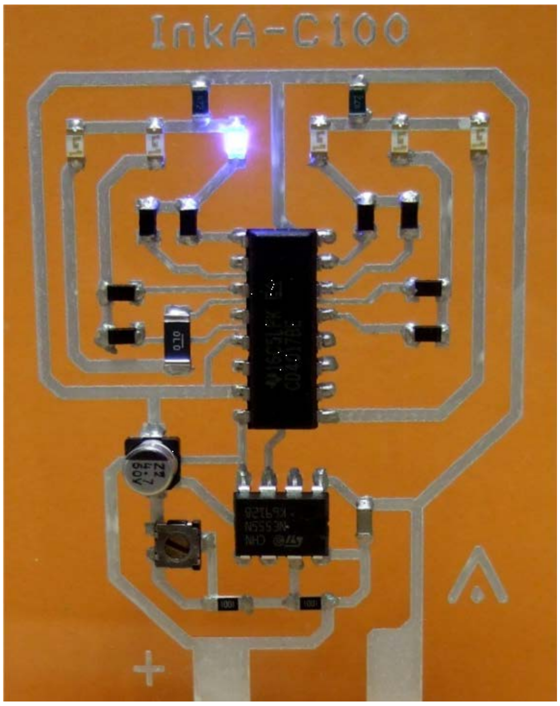


Fig. 1. Silver nanocomposite ink after sintering and resin bonding of discrete electronic components.

Finally, why inks? Inks and writing/printing technology date back to the 23rd century BC (almost 4,500 years ago), presumably being invented in the People's Republic of China by Tien Chu under the empire of Huang Ti. Probably the first nanotechnology ever discovered and applied was that of black ink based on carbon black and bone black, containing fullerenes and a wide variety of aromatic small molecules. Today, printing technology has expanded its horizon toward the realization of electron devices on any substrate, according to two main approaches: 1) analog printing, involving the use of linear/rotary machines that are able to realize multiple copies of the same pattern at a rather high speed (serigraphy, gravure, offset, flexography), involving generally microstructured inks; and 2) digital printing, where raster machines realize at rather slow speed a single copy of a pattern that could be changed simply working at the software level (inkjet printing, 3-D printing), involving nanostructured inks. In Figure 1, an example of a complex circuit realized on an unconventional substrate (borosilicate glass) is shown: thanks to a silver NP-based ink submitted to a sintering treatment and to a bonding phase, it was possible to join the conductive track with discrete traditional components. This is the present cutting-edge level of research. What will be there in the future? We think that once the processes and techniques to produce pure and controlled Ag NPs are settled, more complicated systems involving multimaterial processing could be studied, producing multifunctional, adaptable, and adaptive smart materials. It could be nice to see "futu-retro" technological objects, based on principles as old as 4,500 years but able to realize those functions that are at the basis of our modern “e-society”.

Now that the application domain is clear, to come back to the first question we asked: why is silver so important among nanostructured inks? Silver has optimal electron conductivity and a lower affinity for oxygen if compared to copper, it is 25 times more abundant than gold on Earth’s crust, and hence is less expensive. Silver NPs and nanocomposites (NCs) possess interesting electrical, optical, and chemical properties used in catalysis, surface-enhanced Raman spectroscopy (SERS), nanoelectronics, photonics, and biological and physical sensing. Shape and dimension of Ag NPs are easily controllable, resulting in tunable properties. We will see how silver NP inks are produced and applied in the following sections.

Nanocrystals can be fabricated using two different approaches shown of Figure 2: the first, known as “top-down”, utilizes physical methods to reduce crystal size, while the other, the “bottom-up” approach, is based mainly on solution-phase chemistry and also named wet chemistry. Physical methods usually allow the production of a large quantity of nanocrystals, but is very difficult to control geometry or have a uniform size. In contrast, wet chemical synthesis allows the synthesis of nanocrystals with controlled particle size. Furthermore, as we will see, several nanocrystal shapes can be synthesized by varying the reaction conditions. In the case of inks, the control is really important, because of the dependency of specific properties on the size and shape of the nanocrystal. For these reasons, wet chemical synthesis is generally preferred. In this context, a wide variety of wet-synthesis techniques have been proposed to produce metal nanocrystals and in particular Ag nanocrystals, including chemical reduction, electrochemical and photochemical reduction, sonochemistry, and heat evaporation.

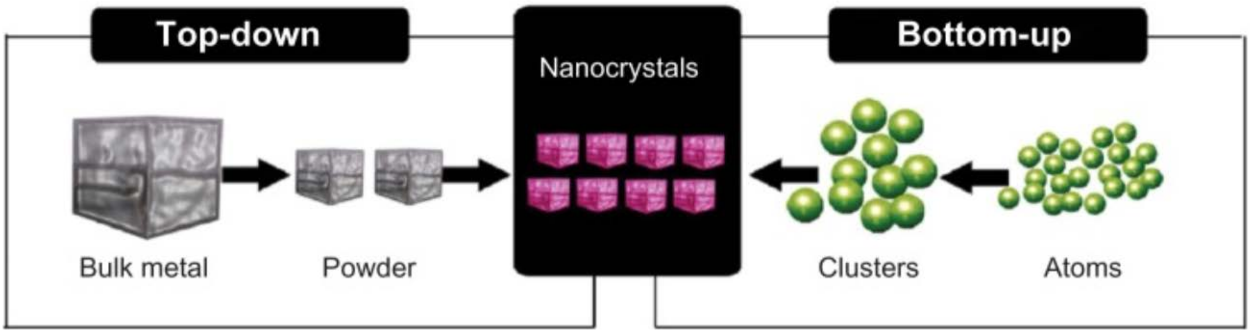


Fig. 2. Top-down and bottom-up approaches to the synthesis of nanocrystals.

The main route involves the bottom-up synthesis, starting from the silver salts and leading to the final nanocrystals. Three distinct stages can be roughly recognized on Figure 3. Nucleation, the clustering of few atoms and/or ions, is the first stage of any crystallization process. In the second step, a seed is formed through atom-by-atom addition to the initial nuclei. In the final step, the seeds grow mainly in size while the shape is largely determined by the structure of the seed.

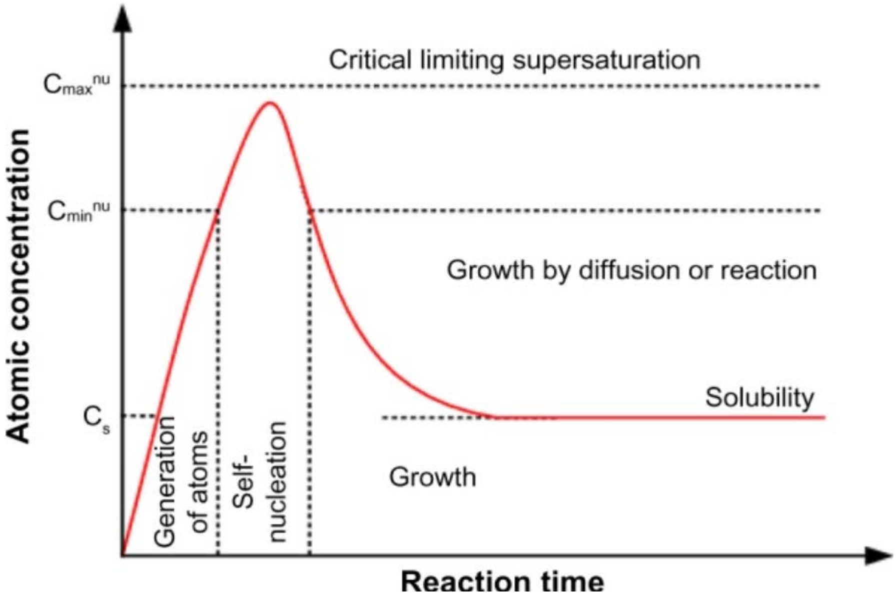


Fig. 3. Plot of atomic concentration against time, illustrating the generation of atoms, nucleation, and subsequent growth.

NP properties, such as catalytic, optical, magnetic, and electronic, have been demonstrated to be size- and shape-sensitive. Nowadays, research efforts are put into not only controlling size and suspension stability but also developing unconventional crystal geometry, eg, synthesizing well-defined anisotropic and/or organized nanostructures. For ink production, usually it is easier to use wet synthesis, because the final NPs are employed in suspension, and thus only bottom-up solution-phase synthesis methods will be discussed.

Bottom-up solution-phase synthesis of metal nanocrystals starts from zerovalent metal compounds or salts dissolved in a solvent. In particular for silver, these precursors are in the +1 oxidation state (Ag^+), and thus, during the reaction, Ag^0 atoms are produced as metal nanocrystal building blocks. Two synthetic pathways are at present under discussion. The first possibility consists in reducing the precursor compound into zerovalent atoms, which then aggregate into the nuclei and grow into nanocrystals. In the second possible reaction pathway, the unreduced metal species associate with nuclei and then are reduced to zerovalent metal species.

Generally, the mechanism depends on reaction conditions: higher precursor concentrations and mild reducing agents shift the reaction from the first to the second pathway. A low reduction rate and high concentration of metal ions prevent the complete reduction into the zerovalent state. A nanosize cluster surface results thus positively charged, and could be stabilized by the capping effect of ionic species, such as halide or carboxylic anions, as well as solvent molecules or polymeric species.

Kinetic control is achieved when the crystal formation is directed by a moderate driving force, thus under conditions far away from the thermodynamic equilibrium. Under kinetically controlled synthesis, the reaction proceeds considerably more slowly than under normal conditions, thus, slowing down the precursor decomposition or reduction.

Silver salts are usually insoluble in any solvent, and thus the most used precursor for Ag nanocrystal production is silver nitrate ($AgNO_3$), which has good solubility in polar solvents. The preferential seed shapes from silver salt reduction are icosahedral and decahedral, thermodynamically favored from the face-centered cubic lattice of metallic silver.

Nanocrystal shape can be controlled by the addition of capping agents, as shown on Figure 4. Surfactants, polymers, biomolecules, small organic molecules, and metal ions or atoms can be used as capping agents. They operate mainly by being adsorbed on a specific crystal plane and thus reduce surface free energy, changing the relative growth rate and inducing growth on the uncapped surfaces. Despite their importance in controlling shape, their mechanism of action is still not completely understood, and thus knowledge on the produced shapes is mainly obtained by trial-and-error attempts.

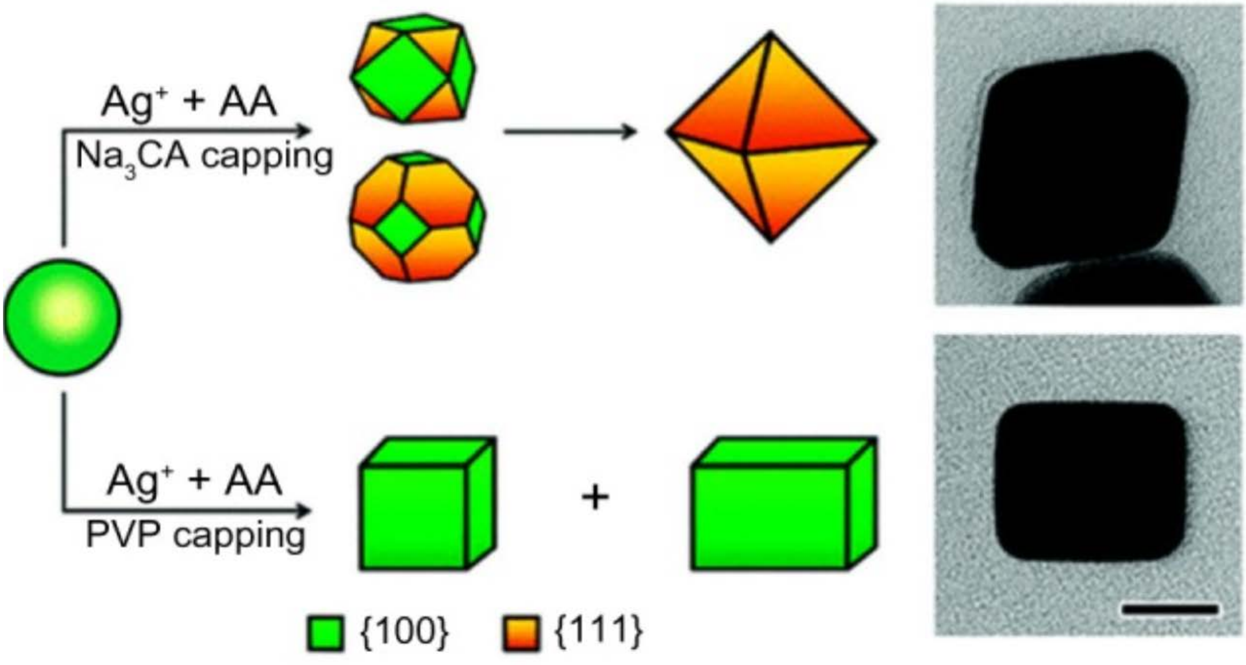


Fig. 4. Role of a capping agent in controlling the evolution of Ag seeds into nanocrystals with different shapes.

Capping agents are used according to two different approaches. In the first approach, the seeds are grown directly into the solution and the capping agents used to orient the addition of the metal atoms on the surface, where the capping agents are either weakly or not bonded. In the second approach, preformed seeds in which capping agents orient the growth are added in the synthesis solution.

The most used capping agents for Ag NPs are polyvinylpyrrolidone (PVP), a polymeric capping agent, and bromine anions. Both the agents tend to be selectively adsorbed onto the {100} facets of Ag nanocrystals, driving the addition of new Ag atoms to other crystal facets. Both capping agents induce the formation of nanocubes, rectangular nanobars, and octagonal nanorods. Bromine ions are much smaller than PVP, and thus the crystals obtained with this capping agent are usually smaller than 25 nm in size. By slowing the reduction rate, thus using mild reducing agents, such as ascorbic acid, it is also possible to obtain Ag nanoplates.

Since the 19th century, silver salts, eg, silver halides, have been used as photosensitive compounds for photography applications. In fact, their light exposure leads to the generation of metallic particles that were used in photography emulsion. Therefore, light radiation is a common way to synthesize silver colloids and NPs.

Photoreduction occurs when photogenerated free electrons react with Ag^+ ions, forming the corresponding Ag^0 metal atom. Ag^+ ions in solution and efficient photogeneration are the main issues to control in order to have an effective production of Ag NPs.

By changing precursors and electron donors, it is possible to control both the dimension and the shape of the NPs synthesized. One of the first syntheses of silver NPs in aqueous and alcohol solution was performed by Hada et al in the 1970s by ultraviolet (UV)-induced photoreduction using the photooxidation of water and alcohols under a deep-UV irradiation. Nowadays, the most common used electron donors are aromatic ketones: they undergo cleavage under UV irradiation, producing radicals that induce reduction of silver. Other molecules reported as photoreducing agents or other parameters have been involved in the reduction process, such as the use of acrylic monomers, sodium citrate to control the pH, reaction performed in nonaqueous media or even applying magnetic fields.

Electrons can be also photogenerated by photoactive semiconductors: under UV irradiation, they are able to promote a free electron that can reduce Ag^+; usually, in this case the redox reaction is balanced by oxidation of the solvent, mostly water. The most investigated materials have been titanium dioxide and zinc oxide.

As reported in the previous section, in order to exploit the potential of metal NPs fully and to provide effective strategies to tune electronic and optical properties of materials, the control of size and morphology of nanostructures are of fundamental and technological interest. Noble metal NPs (Ag, Au, Pt) are extremely interesting because of their unique properties, and among them Ag possess the highest electrical and thermal conductivity, along with other properties, which promote its extensive use in a wide range of applications. Nowadays, Ag NPs are largely used to produce conductive tracks with inkjet printing, thanks to the high conductivity and thermal stability of such materials.

Conductive inks normally are aqueous or organic solvent dispersions of silver NPs that are stabilized by surfactants and polymers that undergo printing, a drying step, and at the end a sintering process that is commonly achieved by heating the printed substrates to a temperature usually higher than 200°C. Alternatively, more unconventional techniques, such as microwave, laser radiation, flash sintering, plasma, and electrical- or chemical-induced sintering, can be pursued.

Sintering at 200°C is much below the melting point of silver (960°C), and it can be attributed to the enhanced surface diffusion of atoms and to surface premelting; therefore, even in the sintering process, the dimension and shape of Ag NPs are one of the first properties to be investigated.

Ag NPs used for inks generally have spherical shape with diameters ranging from 5 to approximately 100 nm with narrow dimensional distribution. Several works are presented in the literature using such NPs. Fuller et al described an inkjet ink based on colloidal silver NPs of spherical shape with a diameter of approximately 5–7 nm dispersed 10 wt% in α-terpineol, which was sintered at 300°C on a hot plate, giving conductive lines of 80 μm and presenting a resistivity of 3 μΩ/cm.

Table 1

Collection of relevant data for inks according to UV-vis measurements, FESEM analysis, and electrical measurements after annealing

Ink name	**NP diameter main mode x ± s (nm)**	**NP diameter main mode 2 × ± s (nm)**	**Main peak x ± w (nm)**	**Secondary peak x ± w (nm)**	**Tertiary peak x ± w (nm)**	**Surface resistance (mΩ/□)**
C10 (46)	10±5	NA	409.7±48.9	NA	NA	30.0
C10 (47)	15±5	100±50	400.1±31.0	365.9±43.5	NA	30.0
C10 (52)	6±2	15±5	420.5±65.7	NA	NA	30.0
C20 (48)	25±15	NA	439.2±77.4	555.3±193.2	365.9±43.5	19.4
C30 (49)	6±2	40±10	469.7±115.0	NA	NA	30.0
C40 (41)	12±2	100±50	441.0±285.1	NA	NA	22.4
C40 (51)	12±2	100±10	470.0±177.6	NA	NA	22.4
C100 (7)	12±2	90±10	NA	NA	NA	30.4

Lee et al proposed a conducting ink composed of silver NPs with diameter around 50 nm dispersed in a water-and-diethylene glycol cosolvent system. Continuous and smooth lines of 130 µm width were printed, and after baking at 260°C for 3 minutes, these lines exhibited a resistivity of 16 µΩ/cm.

Chiolerio et al explored the effects of NP-diameter distribution and composition of Ag NP-based inks for the realization of inkjet-printed microwave circuits. Different NP sizes were measured by numerical analysis of field-emission scanning electron microscopy images, and electrical measurements after annealing gave a surface resistance ranging from 19.4 up to 30 mΩ/□, as specified in Table 1. The best-performing composition was found to be the one containing an added copolymer, and field-emission scanning electron microscopy analysis showed a peculiar NC structure with a percolating network of NP agglomerates, with an extremely low density of metal into the polymeric network.

Different kinds of NP shapes are also presented for the fabrication of conductive inks, such as silver nanowires, which have huge potential applicability in transparent electrodes, but can give rise to problems, such as clogging of the printhead nozzles. In one of the most recent works, authors presented the controlled deposition of networks of silver nanowires (average diameter of 55 nm and an average length of 8.1 µm) in well-defined patterns by inkjet printing from an optimized isopropyl alcohol–diethylene glycol dispersion. The resultant networks, after an evaporation/annealing process at 110°C, presented sheet resistance of 8 Ω/□ and conductivity of 105 S/m, achieved for line widths of 1–10 mm and network thicknesses of 0.5–2 µm deposited from ~10–20 passes. In this case, the thinner networks showed semitransparency.

In 2012, scientists proposed shape-controlled synthesis of Ag NPs by X-ray irradiation for inkjet printing with which various shapes, including spheroidal, prism, rod, and multifaceted NPs, were produced by varying the initial concentration of PVP and $AgNO_3$. It was demonstrated that at an optimized reagent ratio, a mixture of high-aspect-ratio rods (tunable to ~50), and spheroidal particles could be obtained, and such a mixture was proven to have a melting point and dispersive properties suited to inkjet printing of conductive tracks. The resistivity of the printed lines decreased to 77.7 µΩ/cm and 33.1 µΩ/cm after heating to 200°C and 350°C. Nanoplatelets were also proposed for ink applications, allowing the formation of tracks with relatively low resistivity (7.4 µΩ/cm compared to 30 µΩ/cm of a similar track made by NPs), with good stability after external repetitive bending stress as shown on Figure 5. The authors attributed the electrical resistivity and mechanical stability values to the dense microstructure resulting from the NP shape. It was also demonstrated that the pulsed-laser sintering was able to control the shape of the Ag NPs, avoiding the typical “coffee stain” effect and realizing patterned lines with conductivity very close to that of Ag bulk.

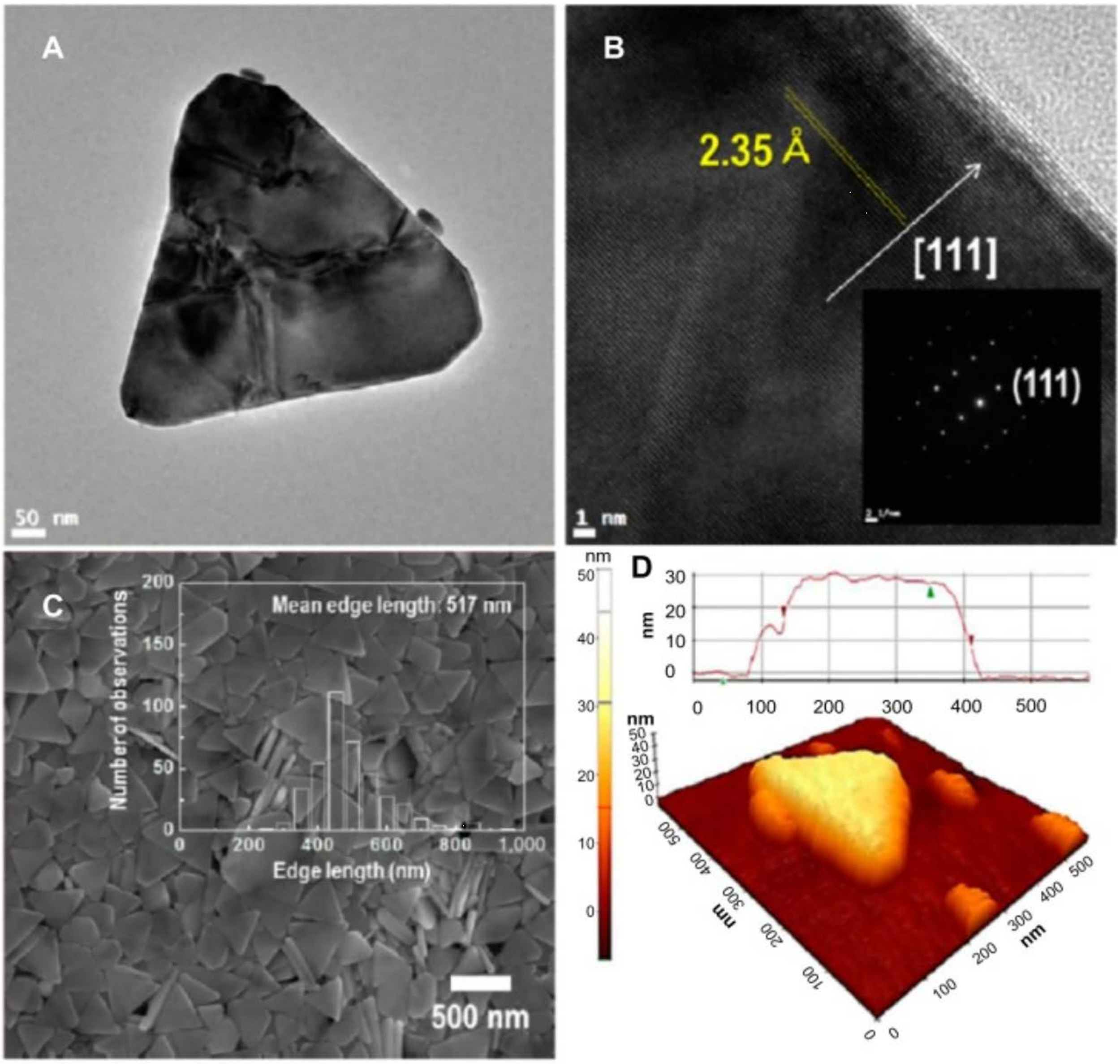


Fig. 5. TEM and AFM of silver nanoplatelets.

An alternative use of Ag inks is related to NP optical properties. It is well known that Ag NPs possess a characteristic plasmon resonance in the visible range that could be controlled by properly tuning their dimensions, number, and relative distance. This property could be exploited for producing optical waveguides. Another property well exploited in the literature is the high transparency that could be obtained from thin films (over 95%) that is necessary in some applications. Also, homogeneous films of Ag NPs show high reflectivity in the visible range; this property was well exploited in order to produce reflective electrodes for solar cells, enhancing considerably solar cell performances.

The following sections deal with the most important applications of Ag NP-based inks. SPR is an effect commonly seen in metals where free electrons collectively oscillate in phase with the incident light, driven by the alternating electric field when irradiated by light of proper wavelength. SPR enables an effective scattering and absorption of light under a resonant condition. For example, this gives to metal colloids, like Ag, their brilliant colors. Concurrently, surface charges are polarized under the excitation of incoming light. In the case of metal NPs, the generated charges cannot propagate as a wave along a flat surface as in bulk metals, but are confined to and concentrated on the NP surface, and thus, this phenomenon is called localized SPR (LSPR).

In these conditions, if organic molecules are adsorbed on the surface of metal NPs, LSPR leads to intense local electric fields within a few nanometers from the particle surface, and thus can be used for the enhancement of the Raman-scattering cross sections of molecules. This would provide an enhanced “fingerprint” spectrum of the molecule, rich in chemical information. This technique is widely known as SERS, and was first demonstrated by Fleischmann and Van Duyne in the 1970s.

It is also known that not only the nanosize dimensions but also the shape of a nanocrystal affects its interaction with electromagnetic waves. Therefore, the intensity and position of LSPR peaks can be fine-tuned by shape control, and a significant Raman-signal enhancement can be achieved by simply selecting nanocrystals with an appropriate shape. The detection of diluted analytes is possible by the signal enhancement of organic molecules. Therefore, the sensitivity of SERS can be greatly enhanced by many orders of magnitude by tailoring the shape of Ag nanocrystals and thus their plasmonic features, ie, LSPR. Particularly, branched silver nanocrystals with tips, such as stars, flowers, and dendrites, have attracted increasing interest for their application in SERS, due to the enhanced plasmonic features.

Embedding highly conductive nanofillers in polymer is a common strategy for producing conductive polymer NCs. One of the common strategies used in order to characterize a noble metal–polymer composite is to evaluate its electrical resistivity. Many works have been reported in literature in this regard.

Silver conductive NCs were synthesized by embedding silver NPs of different shapes in diverse matrices, such as high-density polyethylene, polymethyl methacrylate, polyvinyl alcohol, bisphenol F diglycidyl ether, polyvinylidene difluoride, and polydimethylsiloxane, but also in inks and conductive polymers, such as poly(3,4-ethylenedioxythiophene).

Taking advantage of the electrical conductivity of silver-based NCs that arise upon mechanical stress variation, different pressure and tactile sensors have been produced. These NCs were recently reviewed many authors, partially by Nambiar and Yeow.

Homogeneous dispersion of metallic NPs in polymer matrices remains a critical issue for NC preparation, due to their high surface energy. A common strategy in order to disperse NPs in polymeric matrices is to functionalize the surface of the NPs. An alternate way developed over the last few years envisages the direct dispersion of silver photosensitive precursors in photocurable monomers (often using a cosolvent) followed by UV irradiation, which results in the formation of a polymeric network and in the formation of metal NPs. In literature, several studies have used silver hexafluoroantimonate (in acrylates, epoxies, thiol-ene, and divinyl ether systems), and also in engineered structures, using silver nitrate for synthesizing silver NCs.

Here some innovative strategies for the synthesis of silver NCs are presented in order to illustrate possible future trends in this field. The first strategy concerns the control of the shape of NPs in the solid-bulk phase. Trandafilović et al reported about the synthesis of silver nanoplates in polyampholyte copolymers.

Piezoresistive composite materials have recently found interesting applications in the fields of microsensors, electromechanical devices, circuit breakers, touch-sensitive screens, and tactile sensors for robotics. With respect to commercially available devices, these composite systems can provide cheaper, faster, and more accurate alternatives. By varying the nature and morphology of the type of polymeric matrix and the conductive particles that are used as functional fillers, the properties of these composite materials can be tuned. The percolation effect can be used to explain the conduction mechanism in the case of contact between particles, and the tunneling mechanism where each conductive particle is separated from the others by a thin layer of insulating polymer, which represents the tunneling barrier. In the case of piezoresistive composites, which are based on the tunneling mechanism, a huge change in electrical conductivity is caused, due to an external load-induced deformation. The applied mechanical strain induces a decrease in polymer thickness between the particles, thus reducing the tunneling barrier. In this way, a large reduction in bulk electrical resistance takes place by an increased probability of tunneling.

Silver nanostructures have also been studied and employed as conductive fillers for functional sensing composites. Recently, Hong et al investigated the electrical and thermal conductivities of a silver flake–thermosetting polymer composite. The influence of silver-flake size, distribution, and filler loading on the electrical volume resistivity and thermal conductivity of the composite was studied in detail by the authors.

Concentrated silver as it was shown on Figure 6 NPs are well-recognized materials with potential applications in the field of printing technology. These are used for the preparation of metallic structures on various substrates, because of their high electrical conductivity and resistance to oxidation. Such inks should meet some important requirements: for instance, they should not dry out and clog when in the printhead, they should have good adhesion to the substrate with limited coffee-ring effect and reduced particle aggregation, and they should be characterized by suitable viscosity and surface tension, as they determine drop size, drop-placement accuracy, satellite formation, and wetting of the substrate. These requirements are very well met by Ag NP-based inks.

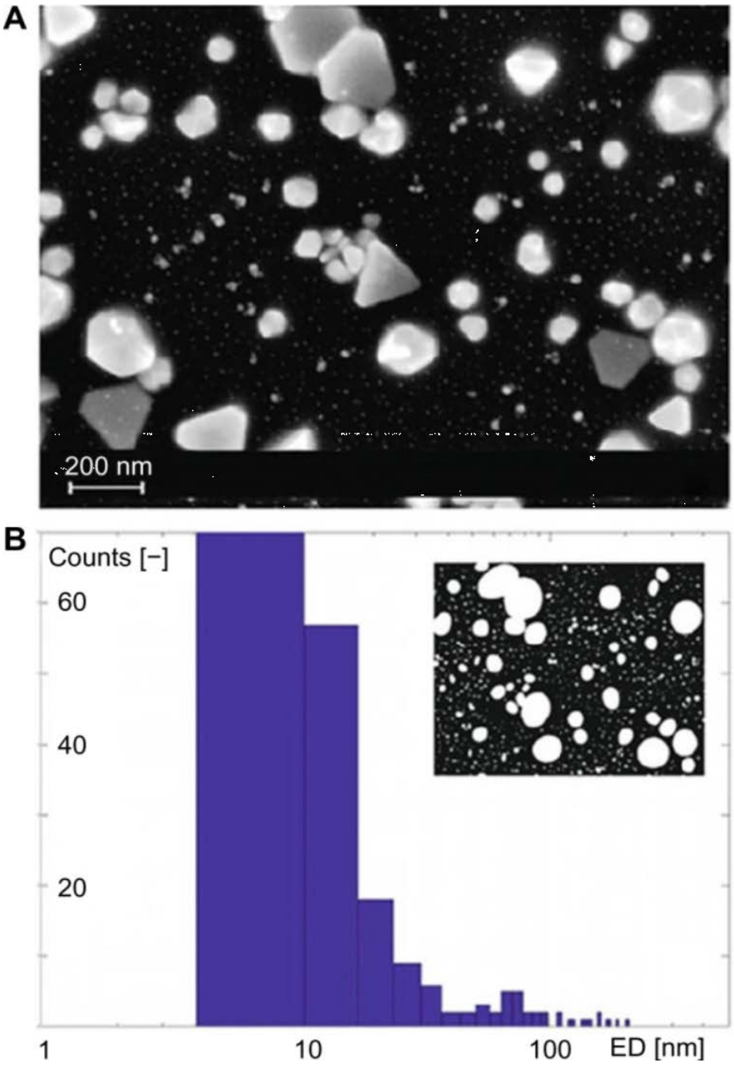


Fig. 6. FESEM image of a water-based Ag nanoink.

In this regard, there are several approaches to formulate Ag-based inks for piezoelectric and thermal inkjet printing that can produce low resistivity and high-resolution conductive traces on different substrates. The primary components for all conductive inks include an appropriate amount of highly conductive metal precursor, such as Ag, Cu, and Au NPs, and a carrier vehicle. The majority of the inks are water-based, and water used in these inks should be very pure so as to limit contaminants. Inks may also contain other additives, such as humectants, binders, surfactants, and bactericides/fungicides. The additives are typically a small percentage with respect to the composition of the ink, and are used to tune ink properties or to add specific properties, thus increasing its performance. Compatibility of the selected ink with a particular inkjet system chosen for deposition is very important, as this influences the interaction among NPs.

In order to avoid precipitation and agglomeration of metal NPs in colloidal inks, dispersants are added to the formulation, which helps to stabilize metal colloids. This helps to increase the loading rate of NPs, thus leading to the synthesis of conductive inks of higher quality. Surfactants and polymers are added to inks in order to interact with the surface of NPs and to form a coating of variable composition and thickness. The resulting modified particle surfaces either attract or repel each other, leading to flocculation or stabilization, respectively. Apart from these, humectants, including alcohols and glycols, are also added to the ink as an additional vehicle or carrier for metal NPs. These control the evaporation of the ink, and help in the reduction of the coffee-ring effect.

Ink transfer to different substrates is facilitated with the help of binder components, which are typically resins that will remain on the substrate or surface along with the NPs. Another important ingredient used in conductive inks is the surfactant, molecules that contain both a hydrophilic and a hydrophobic portion. The main role of a surfactant is to adjust the surface tension of the resultant ink. The addition of a surfactant to a water-based ink will have the result of drastically lowering the surface tension, due to the orientation effects at interfaces caused by the hydrophilic and hydrophobic portions of the surfactant. High surface tension of the ink leads to reduced wettability of the cartridge as well as the substrate, resulting in poor reproduction of the geometry. Growth of bacteria and fungi are common in inks, and this can be avoided by the addition of biocides and fungicides, though with Ag conductive inks it is not necessary, since Ag NPs themselves have antibacterial properties.

One of the most important parameters of an ink is its viscosity. In order to adjust the viscosity to the desired value, a polymeric thickening agent can be used (eg, polyvinyl alcohol). In the case of piezoelectric printheads, the ink viscosity should be in the range of 5–20 cP, while thermal printheads require a viscosity ranging from 1 to 5 cP.

After inkjet printing of a metal NP-based ink, a sintering process has to be performed in order to form a conductive printed pattern. Sintering is the process of welding particles together at temperatures below the corresponding bulk-metal melting point, which involves surface-diffusion phenomena rather than phase change between the solid and the liquid. The conventional approach to sintering metal NPs is heating either with a hot plate or an oven driven by conduction/convection mechanisms (thermal sintering).

In addition to thermal sintering, at present some emerging sintering techniques are being studied and used, such as laser-induced sintering, flash sintering (photonic sintering), microwave oven sintering, and low-pressure Ar plasma sintering (plasma sintering). Sintering can also be obtained by the addition of a positively charged polyelectrolyte, such as polydiallyldimethylammonium chloride, which promotes the coalescence of the NPs due to a decrease in their zeta potential (chemical sintering).

Thermal sintering has been discussed by several authors as a method to optimize the quality of printed silver ink lines, in view of their use as electrodes. A critical drying temperature was found to determine an optimal profile of the printed line, thus also improving the electrical properties of the electrode. In these studies, the authors also considered the effect of other factors on the properties of the printed electrodes, such as drop volume, different substrates, and thicknesses of the printed layers.

The use of inkjet-printed electrodes is important in view of their integration in complex electronic circuits like organic thin-film transistors. In recent years, silver NP-based inks have found a wide range of applications, such as thin-film photovoltaic solar cells, screen printing (which can replace some printed circuitboard interconnections), membrane touch switches, touch screens, automotive sensors, and automatic radio-frequency identification.

The global market demand for high quality and low cost electronic components requires innovative fabrication techniques that are both faster and cheaper compared to traditional production methods. In this regard, inkjet printing with the use of inks based on metal nanoparticles (NPs), metallo-organic decomposition ink (MOD) based on metallo-organic complexes (MC) or salts of various metals is a very attractive low-cost technology for direct metallization.

Inkjet printing, which is widely used in office and home printers, is a non-impact dot-matrix digital printing technology, in which droplets of ink are jetted from a small orifice in a printhead, directly to a specified position on a substrate, as a result of pressure developed after an electronic signal has been sent to the print head to create an image. Today the majority of inkjet printers are based on the drop-on-demand (DOD) methods, as shown on Figure 7: thermal, piezoelectric, electrostatic or acoustic generation of droplets. Most, if not all DOD printers on the market today use either thermal or piezoelectric print heads. In the last two decades, inkjet printing, which has seen commercial success since the late 1970s, has grown to be an important topic in scientific research and technology. The main advantages of inkjet printing compared to other deposition methods are one-step processing, cheap and compact equipment, and applicability to various substrates.

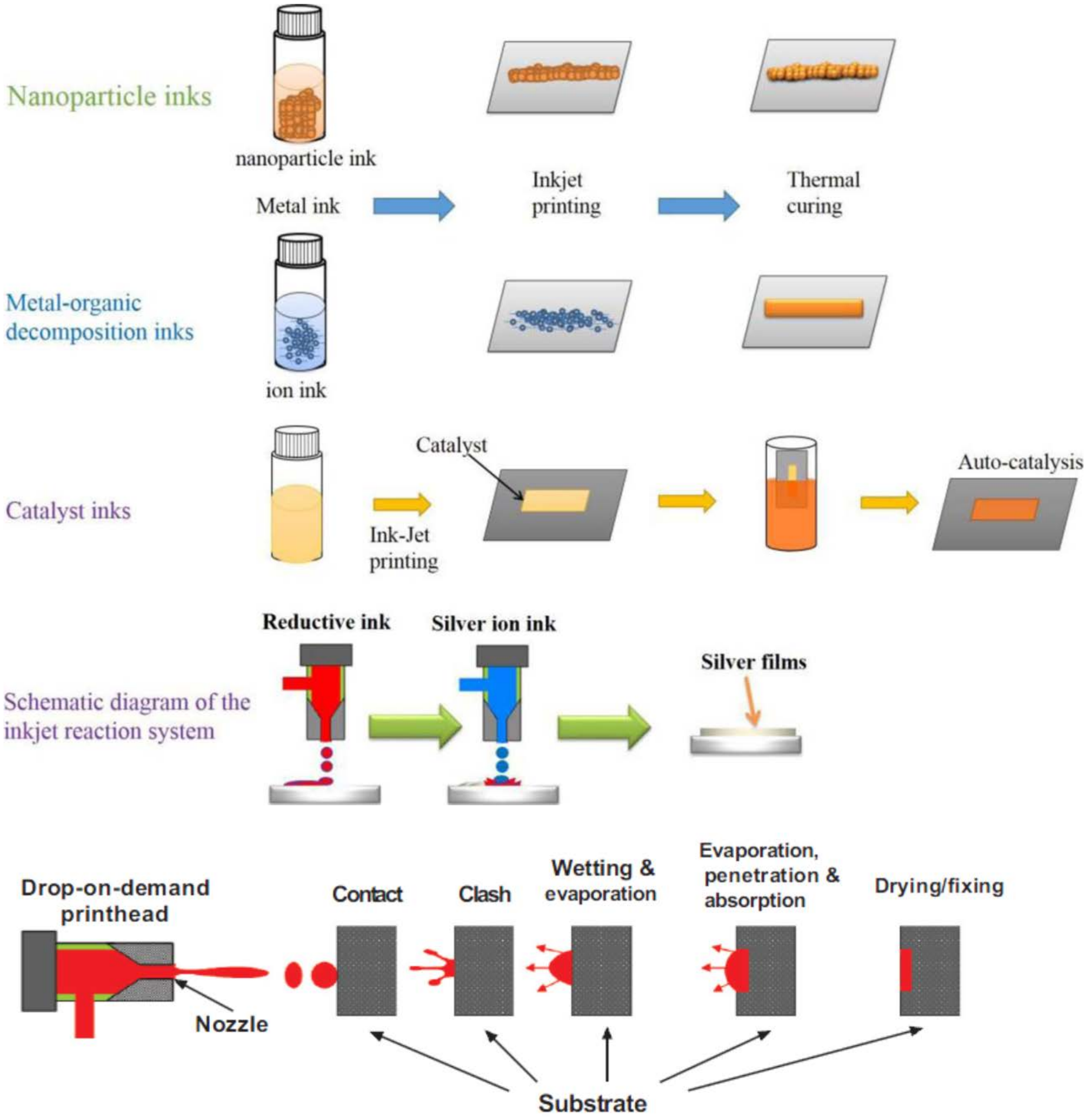


Fig. 7. Schematic presentation of the inkjet printing process.

In addition to conventional graphic applications, inkjet printing has been adapted for patterning various *functional materials* such as organic light emitting diodes (OLED), transistors, integrated circuits, conducting polymer devices, structural polymers and ceramics, biomaterials, and even printed scaffolds for growth of living tissues, as well as for building complicated 3D objects and microelectromechanical systems (MEMS).

Applied to *printed electronics*, inkjet printing enables patterning with high resolution: line and space dimensions can be as small as 10-20 µm, that is at least five times better than the current resolution of screen printing.

In the electronic industry, manufacturing electronic devices such as flexible displays, radio frequency identification (RFID) tags, sensors, OLEDs, PV devices including solar cells (SC), batteries, and printed circuit boards (PCB) by inkjet printing of conductive inks can provide low-cost means of manufacturing large-area electronics on a wide range of substrates (paper, polymers, glass, metals, ceramics etc.) and is attracting tremendous interest.

Nanoparticle inks.— To obtain highly conductive tracks, metals with low resistivity, such as silver, copper, and gold, are commonly used for printed conductive tracks. The most commonly used inks are nanoparticle inks which are metallic nanoparticle suspensions with colloidal stabilizers to prevent particle flocculation or nozzle clogging. At present, due to its high conductivity and low oxidation rate, silver NPs are the most commonly used inks for printed conductive track fabrication. On the other hand, because of low price and high electric conductivity, copper has recently attracted wide attention in printing conductive features. However, because they rapidly oxidize, copper NPs are regularly protected with an anti-oxidation shell of noble metals (Au, Ag, or Pt). The as-printed patterns from NP inks usually have a large resistivity and require a sintering process at elevated temperatures (30–60 minutes at 100–200°C) to remove some electrically non-conductive organic components and/or stabilizers to enhance conductivity. Because of the thermoplastic nature of plastic flexible substrates, the high sintering temperature of NP inks restricts the choice of substrates. To lower the ink sintering temperatures on plastic substrates, researchers have tried to make use of other sintering processes, such as microwave, laser curing, plasma treatments, and chemical sintering. Although these alternative methods have given fruitful and plausible results, however, thermal post-treatment is still extensively used in printing commercial metal inks because of its wide availability and easy access. In addition to high temperature sintering problems, nanoparticle inks with high solid fractions inevitably result in clogging inkjet nozzles. Particles aggregate near nozzle tips in a short time and create a plug at the liquid-air interface. Hence, droplet jetting can be interrupted by these clogs and results in bad printing quality. Proper care should be spent to clean the nozzle tip when the nozzle is idle for a while.

Metallo-organic decomposition (MOD).— MOD inks are highconcentration metal salts dissolved in organic solvents or aqueous solutions. After printed on substrates, the salts decompose into conductive metal under heating conditions. The most commonly used salts are organic silver complex, which can easily be transformed thermally into silver thin films with conductivies approaching that of bulk silver. To reduce the ink sintering temperatures for conductive features, various approaches have been proposed in the literature to formulate silver salt complex solutions at a low sintering temperatures around 90–110°C. These organic salt formulation processes are usually quite complicated and require several synthetic steps. To reduce the preparation procedures, Chen et al. simply mixed silver ammonia solution with diethanolamine to form a clear aqueous solution, which remains transparent for weeks under proper storage. The printed micro-patterns can be easily transformed into highly conductive silver thin films with great adhesion on plastic surfaces after heating at moderate temperatures (20–40 minutes, 50–100°C). Besides thermal treatment, photo-curing processes have also been developed to fabricate metal thin film patterns at room temperature. Valeton et al. printed a UV curable MOD ink on PET substrates and showed rapid silver film formation with a high conductivity of 6.5×10^6 S m−1 at room temperature. Bromberg et al. also inkjetprinted silver nitrate traces and used an efficient plasma process to convert them into tracks with conductivities near bulk silver. Recently, another fast photonic sintering approach has also been developed A solution of copper salts was first printed to form tracks and a high-intensity pulsed light, which can be absorbed by the copper salts, was applied to recover copper. This photonic sintering process provides a fast route for copper or other metal formation, and can be useful for other metals.

Primer inks for electroless plating.— Electroless plating uses a chemical reducing agent to transform metal ions to solid metal. The solid metal film attaches on solid surface by auto-catalysis mechanism occurred on metal surface. The general process of electroless plating involves (i) surface preparation, (ii) surface activation by seeding catalytic metal particles on plated surfaces, and (iii) electroless plating bath to recover metal on the activated surfaces. By controlling the position of activation layers, one can easily determine where the metal crystals grow and create metal thin film patterns. Among all the activation catalysts, palladium (Pd) provides the fastest catalytic effect and thus is frequently used in electroless plating. To fabricate copper conductive tracks, palladium (Pd) colloid inks are first inkjet printed on polymeric substrates following with a subsequent copper electroless plating. This synthetic route creates highly conductive copper micro-patterns on polymeric surfaces for flexible circuits at low temperature with excellent processability. Due to the limited resources and high price of Pd, silver colloids are found to be a good activation agent in electroless copper plating. Although the silver-based materials provide cheaper Pd-free catalysts for copper plating, challenges remain on how to remedy plastic surfaces efficiently for silver-related activating agents. Recent research shows that direct printing of activation agent without surface modification is feasible in preparing copper conductive tracks. A particle-free silver nitrate ink was printed on polyethylene terephthalate (PET) or polyimide (PI) films as an activating agent for copper plating. The printed samples were subjected to electroless plating to create highly conductive copper patterns. Other primer inks, such as polydopamine, can be also used to fabricate copper plated tracks. By replacing the electroless plating solution, this approach can also be applied to form thin film patterns of nickel or other metals.

Metal recovery from redox reactions.— To create metal thin films at room temperature, micro-reaction systems have been developed recently. Inkjet printers were used to eject metal ions and reducing agent separately through two drop generation channels. When these two droplets meet on the printed pattern, metal recovers from the redox reaction to form highly conductive thin films. In 2007, Bidoki et al. printed ascorbic acid and silver nitrate solution sequentially on papers to produce several antenna and capacitor patterns. Kao et al. printed aqueous silver ammonium and formaldehyde solutions through two different nozzles to perform the silver mirror reaction. Silver can quickly recover in the printed area within seconds to form highly conductive thin film features on plastic surfaces. Although these reaction systems are feasible, however, accurate drop landing position and fast drop coalescence or mixing are required in the printing process. Thus, special care should be taken on drop registration process in employing these systems.

In printing conductive tracks on solid substrates, the most frequent problems in printing quality control are bad shape fidelity and non-uniform surface morphology of printed features. These printing problems can lead to severe defects in the printed devices. To resolve these printing problems, numerous remedies have been proposed in the past few years and are summarized in the following sections.

Fluid mechanical instability.— In the applications of printed electronics, the stability during or after pattern creation processes are of crucial importance for printing quality. Different from traditional printing on porous papers, on which inks are mostly absorbed by capillary forces, the printed liquids stay on the flexible plastic surfaces and might move around according to the surface conditions. The motion of liquids during printing or post treatment can lead to significant loss in shape definition and thus greatly affect the functions of printed elements.

Generally speaking, those shape boundary instability problems are especially severe on the hydrophobic substrates with small contact angle hysteresis. Thus, to ensure both wettability and homogeneity of surfaces, special treatments, such as plasma or chemical polishing, are commonly used before the subsequent printing process.

However, even with proper surface modification, the intrinsic fluid mechanical instability always leads to inevitable printing quality deterioration and ink solvent adjustments are usually needed to eliminate the ink-substrate incompatibility.

A large portion of printed electronic devices is composed of straight line patterns. Thus, hydrodynamic stability concerning the printing sequences of a liquid rivulet on solid substrates has been extensively studied. Davis first evaluated the stability of an infinite liquid rivulet, and showed that the printed lines can break into separate drops under various surface conditions. Similar approach was adopted by Kondic et al. to evaluate the stability of liquid rivulets with finite length. The hydrodynamic instability of an inkjet-printed rivulet was further examined experimentally by Schiaffino and Sonin.

The wavelength of bulges between liquid ridges can be estimated theoretically predicted and their theory was experimentally confirmed. Later, Duineveld developed a dynamic model for liquid bulge formation in inkjet printing process with experimental demonstration. The stability of a liquid line, formed by sequential deposition of overlapping droplets, was further examined carefully by Thompson et al. with a rigorous drop-to-drop approach. In 2010, a practical inkjet printing model was developed by Stringer and Derby. Using this model prediction, one can adjust the droplet pitch, printing speed, and droplet sizes to create straight line patterns as it clear from Figure 8.

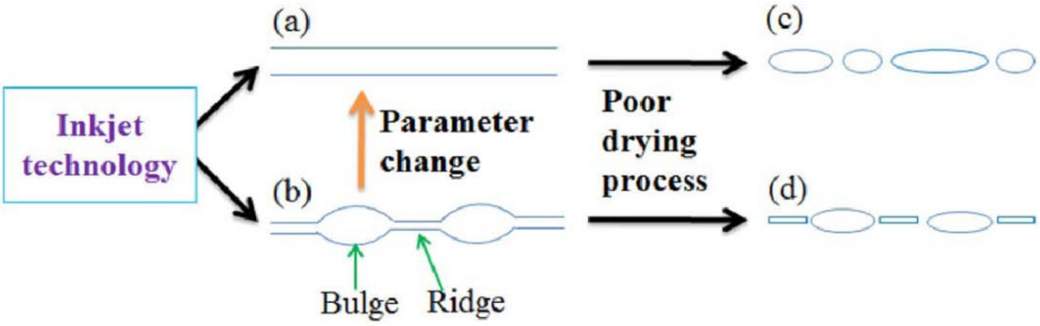


Fig. 8. Instability of inkjet-printed lines can be induced by several factors. (a-b) Printing parameters, such as dot-to-dot spacing and printing speed, can result in geometrical shape variations. With careful parameter tuning, bulges and ridges on printed tracks can be avoided. (c-d)With poor drying conditions, the printed lines can be broken into pieces. Solvent adjustment or substrate heating might be necessary to avoid the shape infidelity of printed features.

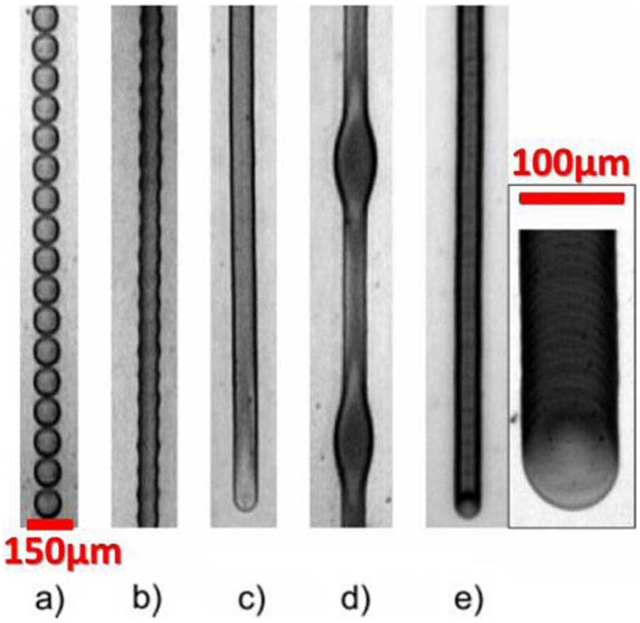


Fig. 9. Examples of principal printed line behaviors: (a) individual drops, (b) scalloped, (c) uniform, (d) bulging, and (e) stacked coins. Drop spacing decreases from left to right.

Although these hydrodynamic stability analyses have greatly reduced the technical difficulties in printing micro-patterns for microelectronic devices, the predictions are limited to simple straight line geometry at constant temperature. It has been shown that with fast evaporation, the printed patterns can be faithful to the original design. Soltman et al. also showed that proper drop registration scheme can help preserve the pattern definition of the printed liquid films (see Figure 9), and the receding contact angles of liquids on the substrate surfaces are of crucial importance to the accuracy of printed features.

With the addition of co-solvent adjustments, one can also adjust the contact angles to preserve the pattern definition. Lin et al. developed a theoretical model to explain the dewetting phenomena during the drying process of a binary solvent system, and successfully predict the contact line pinning conditions in the experiments. The optimized water/PEG solutions can then be used to prepare dye- or particle-based inks, which preserved accurate features after solvent evaporation as it s presented on Figure 10.

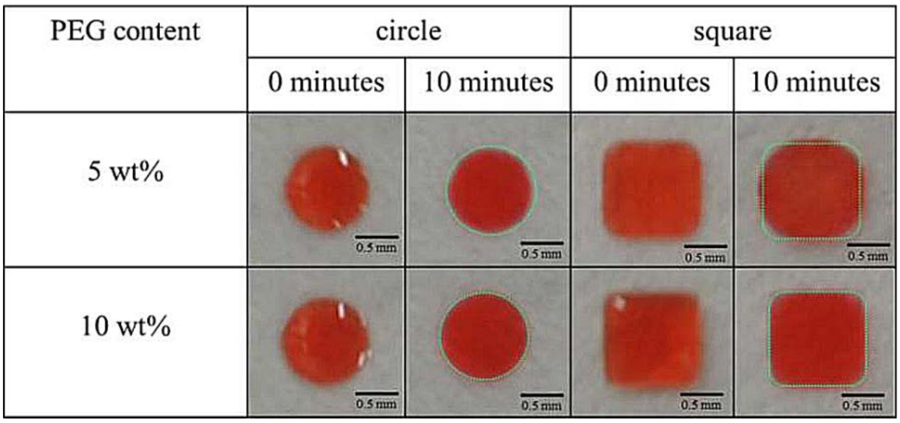


Fig. 10. Optical images of printed circles and squares with various PEG concentrations. The dye-based ink contains 0.2 wt% Congo red. When a 10 wt% aqueous PEG solution was used, the shape definition can be maintained the same as the original design after solvent evaporation. Dotted green lines represent the original shapes for comparison.

Drying and coffee ring effects.— When ink droplets dry on a solid surface, the strong solvent evaporation at contact lines leads to aggregation of solutes or particles at pattern edges. Thus, the solute or particulate matter tend to deposit in ring-like fashion, or known as the coffee-ring effect. The coffee ring effect leads to a strongly non-uniform morphology and performance of patterned devices.

To improve the uniformity of printed thin films, modification of solvent composition has been proposed to provide additional Marangoni flow to cosunteract the evaporating flux at contact lines. To date, significant progress has been achieved to effectively control the coffee-ring effects on printed features.

Basically, the coffee ring formation is largely determined by the time scales of liquid evaporation and particle movement. For a non-volatile solvent, i.e. a long evaporation time scale, one expects no outward flow caused by solvent evaporation and thus negligible coffee ring effects.With this concept, Kim et al. used a high boiling solvent with low surface tension, such as ethylene glycol, in silver nanoparticle inks to reduce the evaporative flux to decrease coffee ring effects on printed conductive tracks. On the other hand,Yunker et al. also showed experimentally that the shape of suspended particles can be used to eliminate the coffee-ring effect as it was shown on Figure 11.

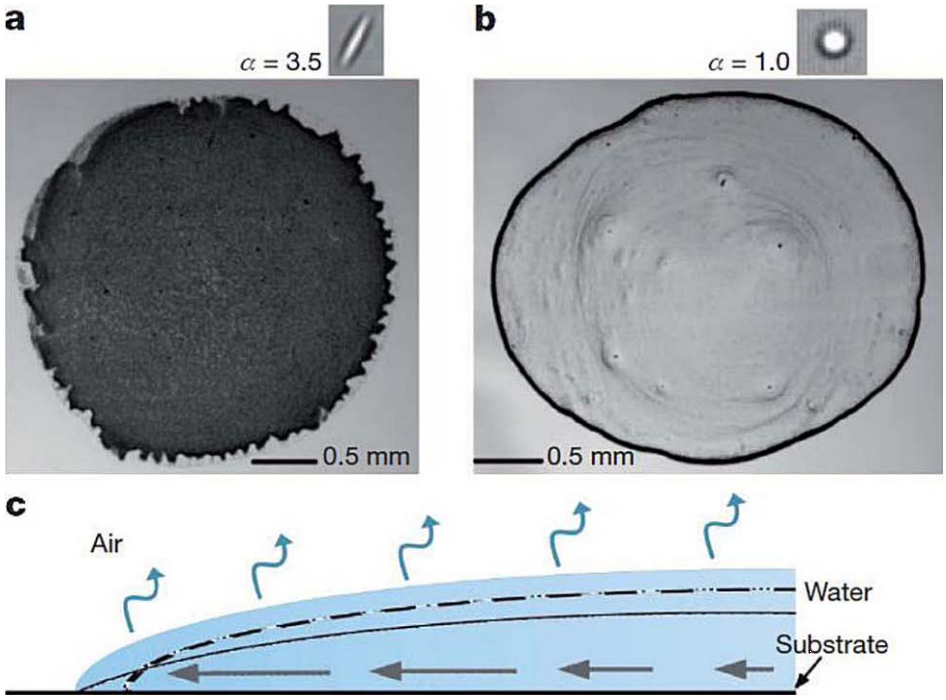


Fig. 11. Deposition of spheres and ellipsoids. a, b, Images of the final distributions of ellipsoids (a) and spheres (b) after evaporation. Insets, particle shape. c, Schematic diagram of the evaporation process, depicting capillary flow induced by pinned edges. Evaporation occurs over the entire drop surface (blue arrows), so if the contact line were free to recede, the drop profile would be preserved during evaporation (dashed line). However, the contact line remains pinned, so the contact angle decreases (solid line). Thus, a capillary flow (black arrows), from the drop’s centre to its edges, is induced to replenish fluid at the contact line.

Moreover, the increase in ink viscosity can lead to slow particle motion and hamper the coffee ring formation. Thus, addition of polymers in the inks can also ease the coffee-ring effects during droplet evaporation. In contrast to diminish the non-uniformity caused by solvent evaporation, enhanced coffee ring effects have also been used to manipulate the microstructures of conductive thin films for greater transparency. Magdassi et al. demonstrated that at room temperature, the replenishing flux during the evaporation of silver nanoparticle suspensions leads to compact solid aggregates near the edges of droplets.

After drying, the silver deposits around coffee rings can yield an electrical conductivity as high as 15% of bulk silver. With this principle, Zhang et al. recently printed silver nanoparticle inks in various patterns, such as lines and meshes (see Figure. 12), with enhanced coffee rings by adjusting substrate wettability.

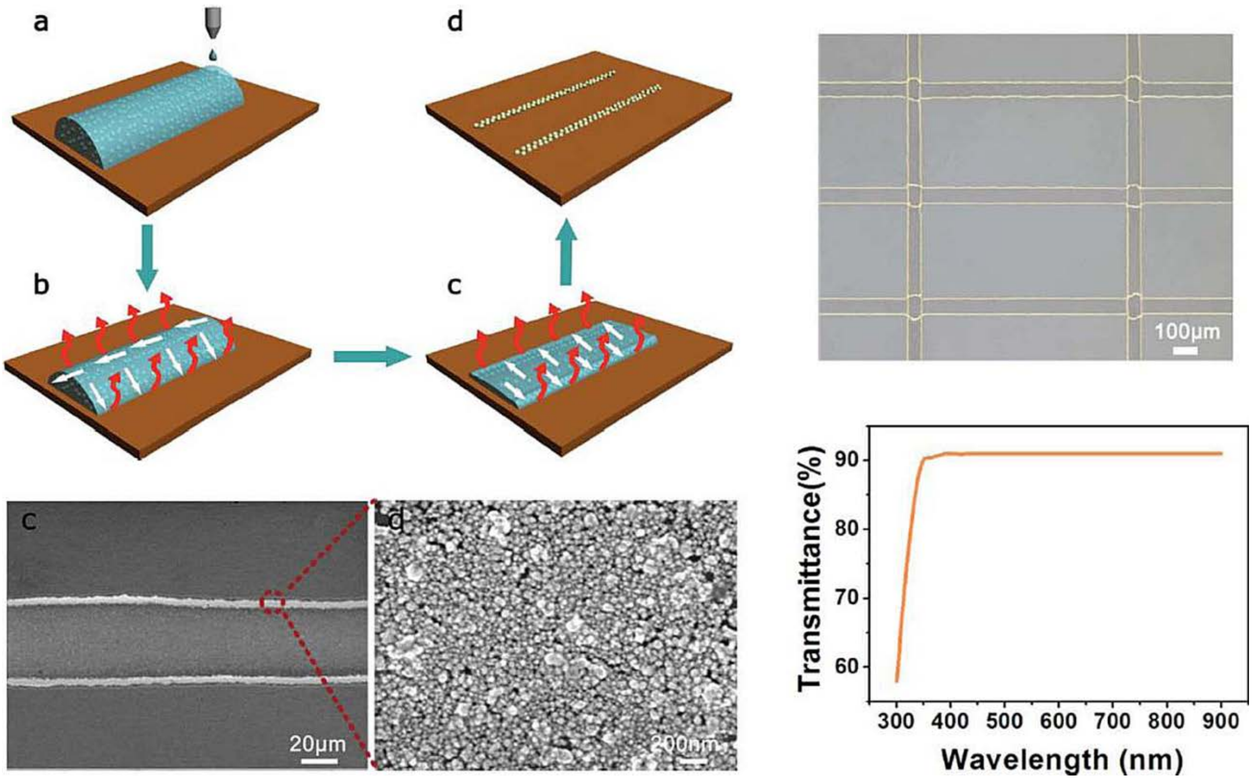


Fig. 12. Schematic illustration of inkjet printing of silver-nanoparticle patterns induced by the coffee-ring effect, SEM image of the printed coffee line and the printed coffee line rim, optical image and transmittance of the glass substrate with the reticular conductive pattern.

The printed rings assembly form a network composed of highly conductive silver rings and create a nearly transparent conductive thin film.

Adhesion and mechanical stability.— The conductive tracks are regularly printed on plastic sheets, such as polyethylene terephthalate (PET) or polyimide (PI) films, to create flexible microelectronic circuits for flexible and light-weight electronic devices. To ensure satisfactory device performances, the printed conductive tracks must adhere to the substrates firmly. Tape tests are regularly performed to test the adhesion: a piece of pressure sensitive tape is applied over the printed area and removed quickly.

As shown in Figure 13a, the printed metal tracks remain totally on the plastic sheets after removing the attached tape, indicating satisfactory adhesion between the tracks and substrates. Besides adhesion, mechanical stability of the printed tracks is also needed to ensure the long term performance of “flexible” electronic devices.

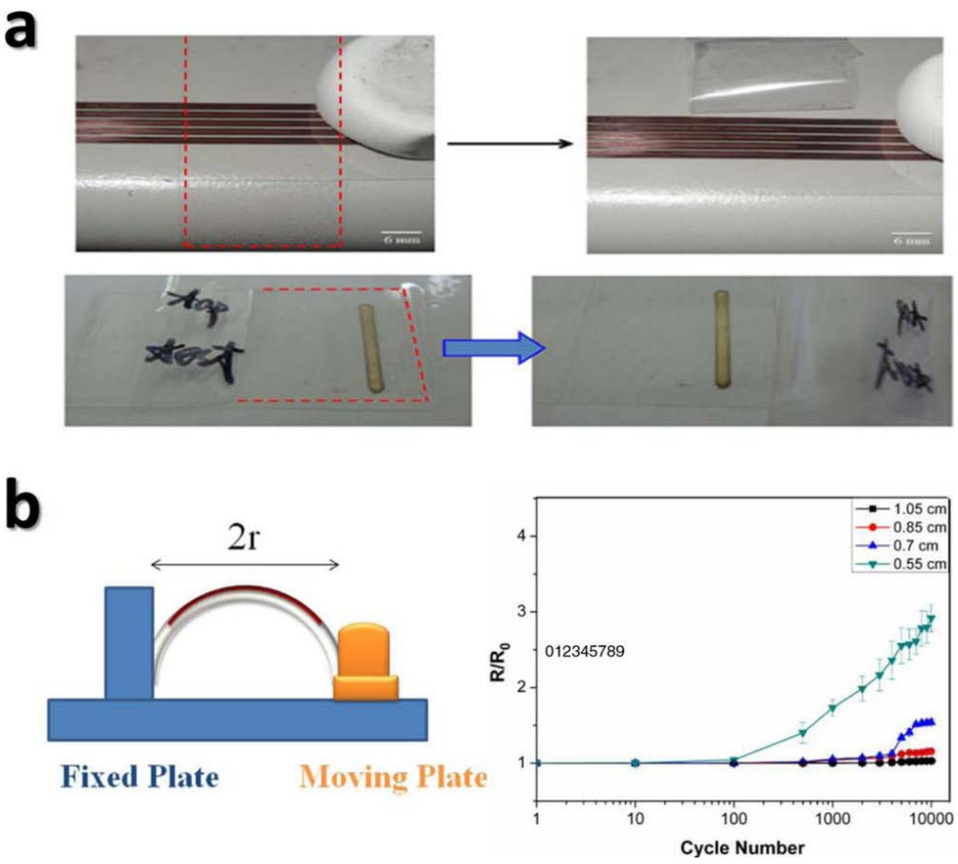


Fig. 13. (a) Typical tape test for metal track adhesion on plastic substrates. Pressure sensitive tapes are firmly attached on the printed area (red dashed line area), and quickly removed. The completeness of the printed feature indicates the good adhesion. (b) Bending tests for mechanical stability of printed metal tracks. The resistance change after repeated bending cycles is recorded. Regularly, smaller radius of curvature leads to larger strain and hence the resistance increases quickly after several bending cycles.

Figure 13b shows a commonly used bending test method. A piece of plastic sheet with printed tracks is placed over a moving stage, which repeatedly moving back and forth to bend the sample. The change in the electronic resistance is recorded with the number of bending cycles to understand the mechanical stability of the sample tracks after being bent repeatedly. The adhesion and mechanical stability can be improved by post-treatments and addition of polymer promoters. At present, most commonly used conductive inks contain plastic binders to help the adhesion and mechanical stability of printed patterns on substrates. However, because binders are regularly electronically non-conductive, the addition can lead to lower conductivity. Thus, more investigation on improving the adhesion of metal tracks on plastic substrates without conductivity loss is still needed.

Smaller feature sizes with faster printing speed.— To meet the demand of fast growing market in printed electronics, a printing system with good printing quality and reliability is needed. To compete with the current semiconductor processes, smaller feature sizes are required for device fabrication. Commercially available inkjet printing systems produce droplets in accordance to the nozzle size with a resolution of 20–30 µm. To obtain higher resolutions, it requires smaller size nozzles, which have manufacturing difficulties as well as worse nozzle clogging problems. On the other hand, to achieve mass production scale, the fast printing processes, which can be up to 100 m/min, require high droplet placement accuracy. However, air motion at fast printing speed also results in larger drifting velocity for smaller droplets and worsens the registration accuracy. Therefore, the quest for high resolutions at faster printing speeds with more precision and accuracy leads to the development of various non-conventional ink-jet printing methods. One technology for getting higher resolutions and better droplet placement accuracy utilizes externally applied electric fields. The socalled electrohydrodynamic inkjet (EHD inkjet) printing technology pulls the liquid inks out of the nozzle rather than pushing inks out. As a result, EHD inkjet is capable of producing superfine ink droplets of size much smaller than the nozzle size with fast ejection speed as it was shown on the scheme on Figure 14.

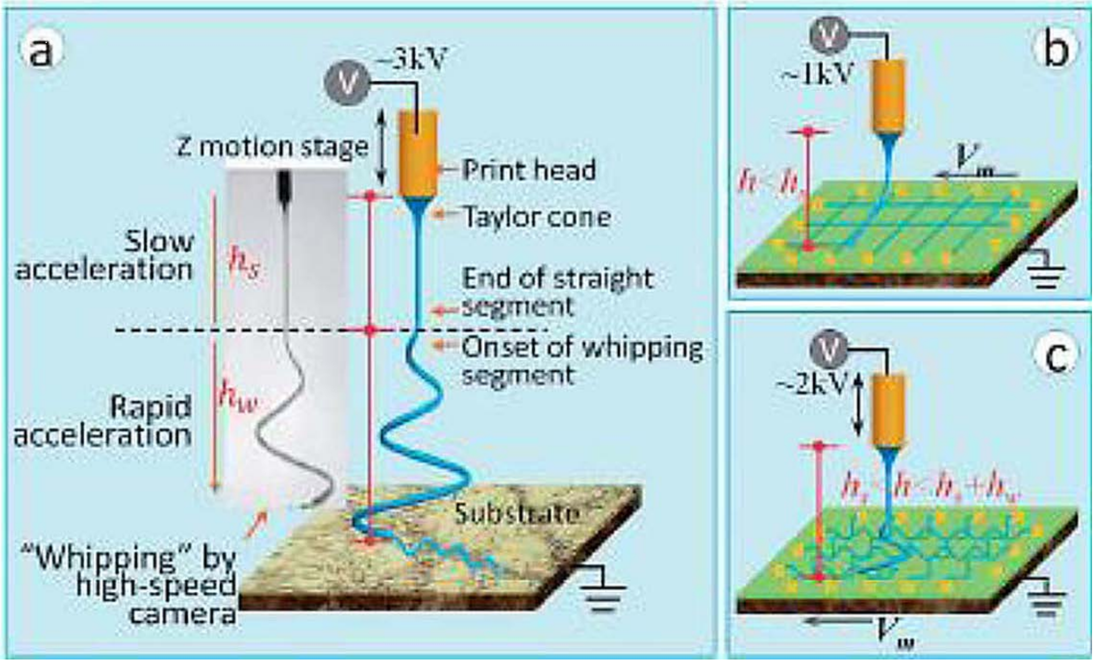
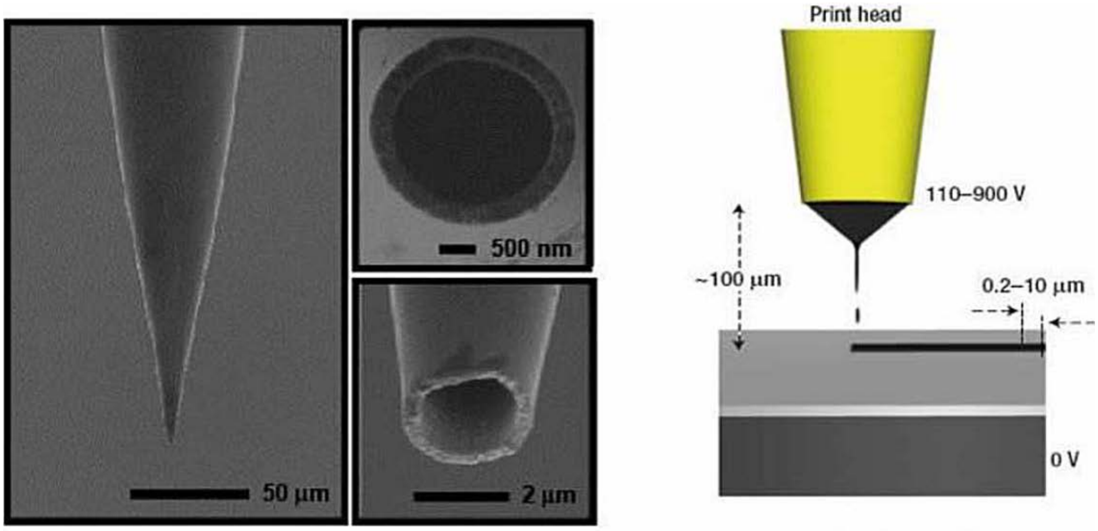


Fig. 14. (a) Standard electrospinning (far-field electrospinning) to fabricate a nonwoven mat. (b) Near-field electrospinning to direct-wire fibers. (c) Medium-field electrospinning to deposit wavy structures.

The droplets size can be ~1/10 in diameter or 1/1000 in volume smaller than conventional inkjet printers. Rogers et al. designed such a system in 2007, and showed the ability to produce conductive features down to sub-micrometer resolution (see Figure 15). A similar but further industrialized system called super inkjet (SIJ) technology was also developed by National Institute of Advanced Industrial Science and Technology (AIST) in Japan. The SIJ system can fabricate nearly sub-mircon features on any substrates.Moreover, it can be also used to create 3D conductive features as well.



Scanning electron micrographs of a representative nozzle; exit apertures with diameters down to 300 nm were used successfully

Schematic illustration of e-jet printing setup

Fig. 15. Nozzle structures and schematic diagrams of a high-resolution ejet printer. a, SEM images of a gold-coated glass microcapillary nozzle (2 µm internal diameter). A thin film of surface-functionalized Au coats the entire outer surface of the nozzle as well as the interior near the tip. The insets on the right show views of this tip region. b, Nozzle and substrate configuration for printing. Ink ejects from the apex of the conical ink meniscus that forms at the tip of the nozzle owing to the action of a voltage applied between the tip and ink, and the underlying substrate. These droplets eject onto a moving substrate to produce printed patterns. In this diagram, the substrate motion is to the right.

Electrodes on conformational surfaces or in 3D printed objects.— Wearable electronic devices have recently attracted widespread attentions due to its extensive usage in the next generation electronics. For this new and emerging application, electronic devices are installed or fabricated on non-planar substrates or even 3D objects, such as contact lens, glasses, and gloves. Thus, new printing technologies for conductive tracks on curved surfaces or 3D structures are required. To date, some attempts have been tried to print conductive materials on curved surfaces using pen writing technology. For inkjet printing, it has been shown that patterns of medicines can be printed on the complicated surfaces of medical stents with special fixtures.

Similar printing tools can be quickly extended to print conductive materials on curved surfaces, but geometry analysis for 3D surface definition and related drop registration schemes might poise challenges for printing tool development.

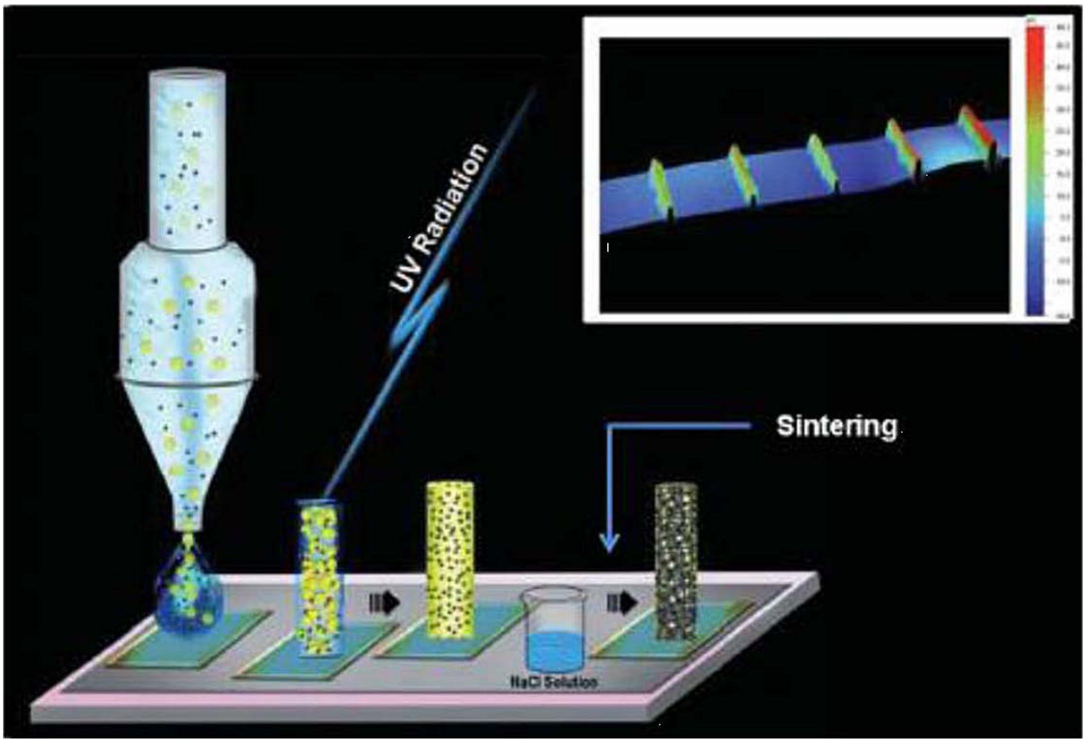


Fig. 16. Printing of a conductive 3D structure with the use of ink composed of an UV-curable emulsion and a dispersion of metal NPs.

Another important perspective is 3D printing of conductive patterns. It was found that geometrical parameters, such as vertex shapes and aspect ratios, can effectively adjust the electron transfer efficiency or surface currents of 3D micro-structures. Thus, it is of great interest for scientists to design new 3D printing tools for conductive materials. This field is at the infancy stage of research, and search for new nanomaterials as well as suitable 3D fabrication tools based on wet deposition is still challenging for scientists. Magdassi et al. developed a new style of 3D conductors by using inkjet printing technology combining with UV curable ink. The synthesized silver nanoparticles were dispersed in water phase and mixed with the oil phase counterpart, which has UV curable characteristic. The printed patterns were exposed to UV light for 1 second and sintered by dipping the substrate into NaCl solution for 10 seconds. The asprinted silver pattern has a conductivity of 1.9×10^{-6} S m^{-1} within the 3D structures as it shown on Figure 16. This new type of ink preparation with an inkjet fabrication method can provide a potential tool for 3D printing electronics.

Inkjet printing with the use of inks based on metal NPs and metallo-organic decomposition inks is considered to be an innovative technique for high quality and low cost metallization. Applied to printed electronics, inkjet printing enables direct high resolution metallic patterning on various substrates, including flexible heat sensitive polymers.

Current high price of commercially available conductive inkjet inks and high sintering temperatures required for obtaining conductive printed patterns impede their use for printed large area plastic electronics. Therefore, R&D efforts should be focused on development of new inks based on low cost metals with high conductivity tailored for low temperature sintering to provide a simple, low cost, energy efficient, environmentally friendly and R2R compatible platform to produce highly conductive structures with high resolution.

Since obtaining conductive printed patterns on plastic substrates, in addition to inkjet formulations, requires also the low-temperature sintering, a detailed understanding of the metal nanoparticles sintering process as well as development of the emerging sintering techniques (photonic, microwave, plasma, chemically triggered room temperature coalescence) are crucial. The efforts here should be focused on optimization of the whole process ink formulation processing-sintering.

Chapter 5 Solar cells of 3-rd generation with nanosilicon. Quantum dots solar cells.

5.1. Introduction

Photovoltaics (solar power) - the direct conversion of sunlight into electricity in semiconductor solar cells (Fig.5.1). Nowadays, photovoltaics (PV) is growing rapidly, a total global installed capacity achieved of 139 GW at the end of 2013 (and 38 GW was installed for the last year) (fig.5.2). PV now covers 3% of the electricity demand and 6% of the peak electricity demand in Europe (EPIA 2014). Monocrystalline, multicrystalline and thin film Si cells and modules actually draw up more than 88% from the total PV production (Aulich et al. 2010). However, the PV electricity still remains more expensive comparing with traditional nuclear or thermal power engineering. That's why now the main efforts of solar power engineering is directed to improve the cell efficiency and to reduce the cell cost, particularly by developing of new structures and materials.

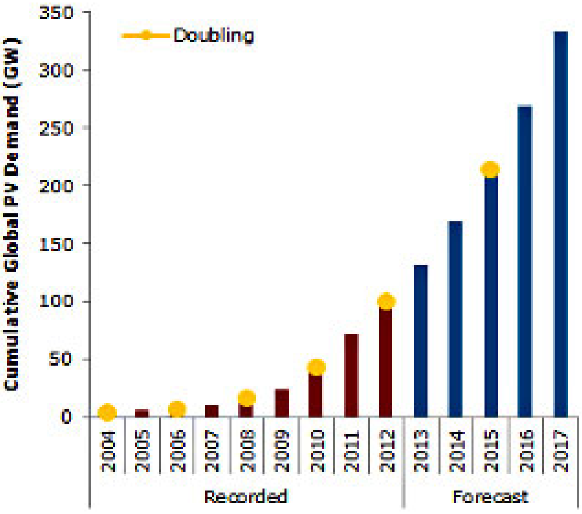


Fig.5.1 Solar electrical station

Fig.5.2 Cumulative global installed PV capacities have doubled every two years on average since 2004

The simple sketch of Si solar cell design with diffusion p-n junction can be presented as following (Fig. 5.3): heavy doped n^+- emitter (N_d~10^{20} cm^{-3}) of 0.3 - 0.5 μm thick is formed on standard textured p-type Si wafer (base) of 150- 250 μm thick with resistivity of 0.5 - 5 Ω·cm. On the surface of n^+- emitter the dielectrical passivating layer and antireflection coating (ARC) are formed as well as metal terminals. Heavy doped p^+ - layer (back surface field, BSF), passivating layer and metal terminals are created on the rear side of a cell. By using of low-cost Si grade materials (upgraded or purified metallurgical Si) the conversion efficiencies of 16% were obtained at cell dimensions of 150 x 150 or 200 x 200 mm^2 (Halm et al. 2010). If as a substrate the electronic quality Si is used then the conversion efficiency can be increased up to 17-19% (Burgers et al. 2008). At that time the laboratory solar cells with small area show 25 % for monocrystalline Si and 20.4% for multicrystalline Si (Green et al. 2011). At that time thermodynamics limit of sunlight conversation is 93% (direct) and 74% for global sun illumination. It is explained by power loss mechanisms.

Optical and recombination losses reduce the cell output from the ideal values. The main from them are: Optical losses, Inability to absorb photons with energy less band gap E_g (1), Thermalisation of photons energies exceeding the bandgap (2), Junction and contact voltage losses (3,4), Recombination losses (5) (Fig.5.4). As results, only small part of solar flux is used by solar cells (Fig.5.5).

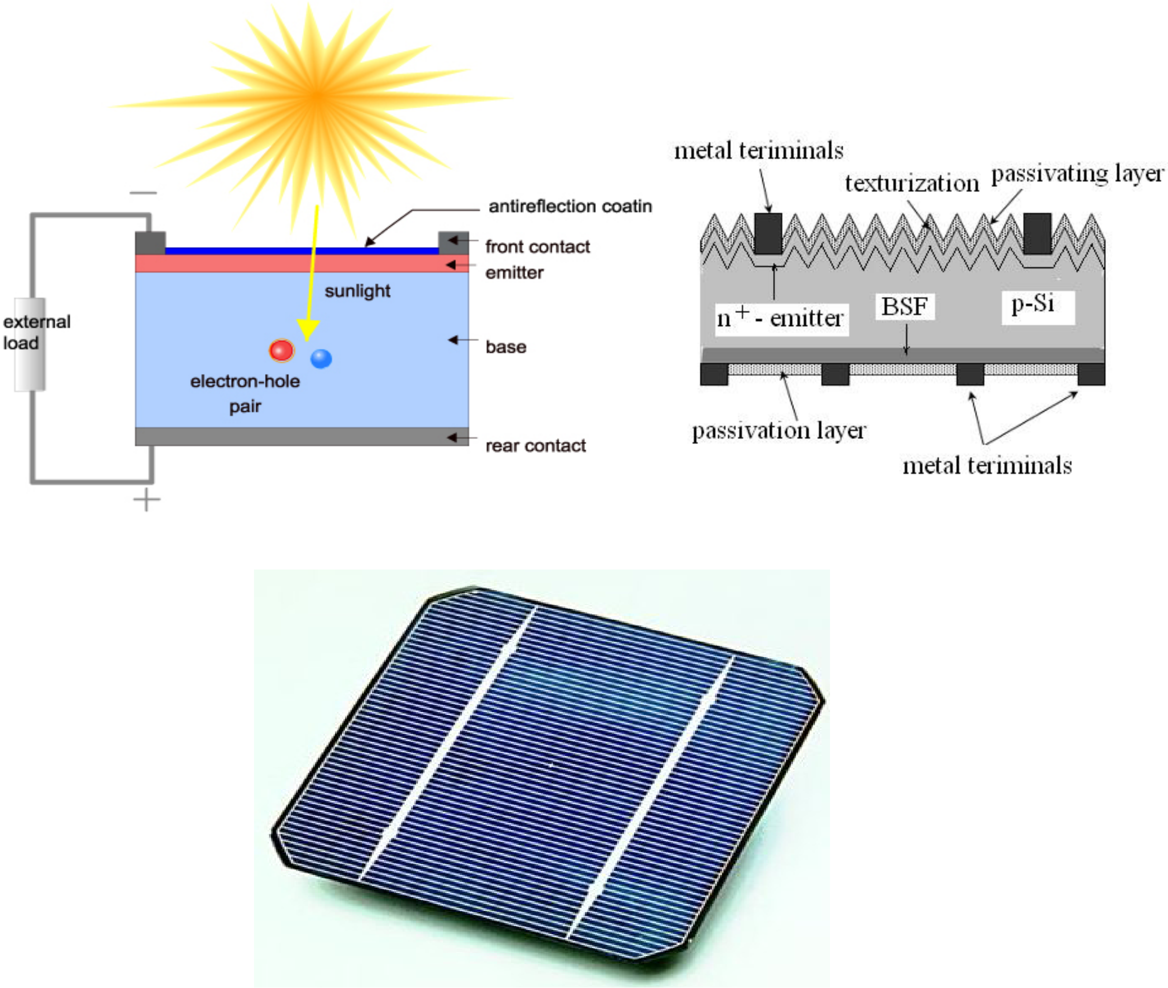


Fig.5.3 The simple sketch of Si solar cell design with diffusion p-n junction

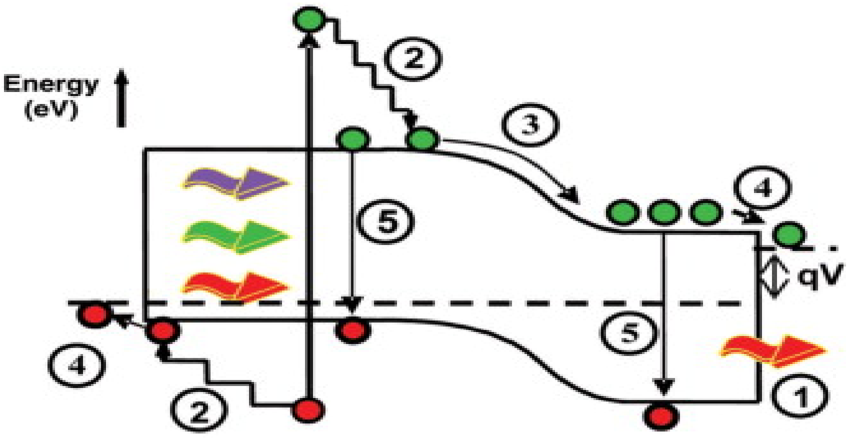


Fig.5.4 Most important power loss mechanisms

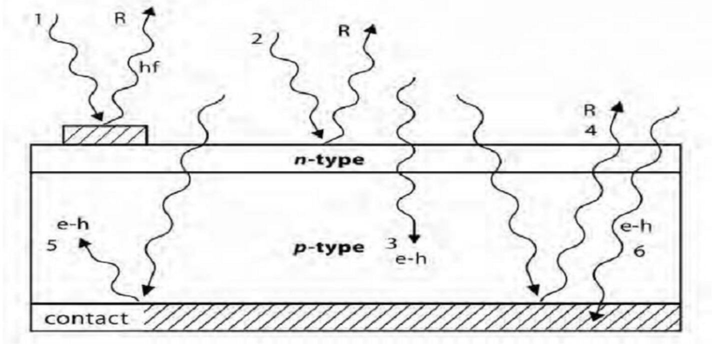


Fig.5.5 a)Optical losses (reflection from front contact (1), reflection from front surface (2), absorption in rear contact (5)) and desired absorption (3), absorption after reflection (5)

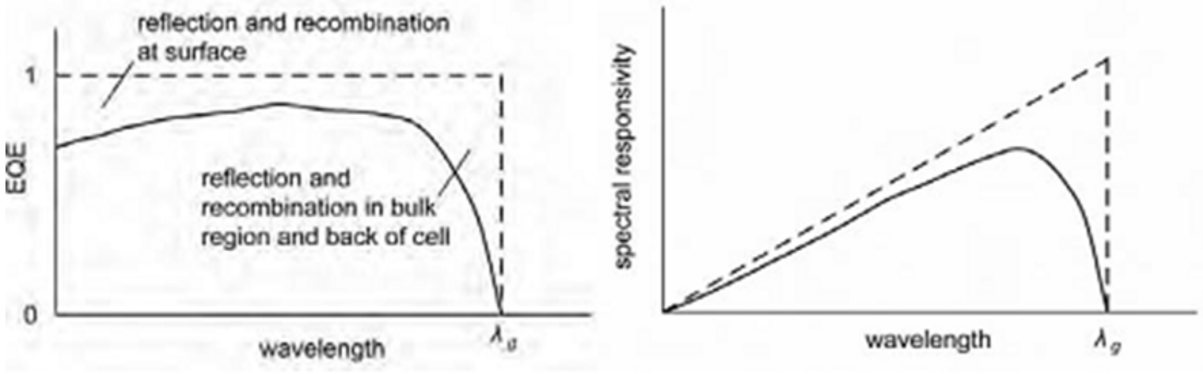


Fig.5.5b) Typical external quantum efficiency and responsivity in actual solar cell, illustrating the impact of optical and recombination losses

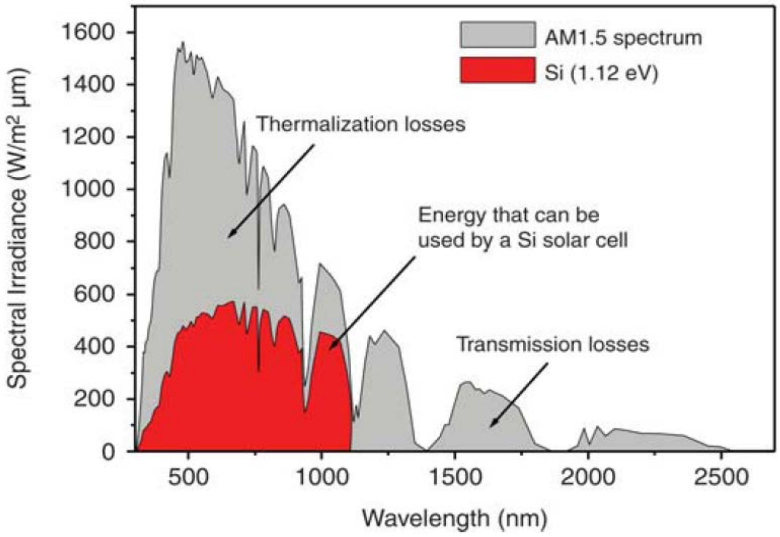


Fig.5.5 Spectrum of solar illumination passed through 1.5 atmosphere and part of solar spectrum that is utilized by silicon solar cell.

Today only silicon solar cell satisfies the requirement of high efficiency, low cost and long life time for commercial applications (fig.5.6). Really, the high efficiency (30-45%) can be obtained in two and 3 junctions cell and concentrated solar flux. However, such technology is expensive since uses the compex A3B5, A2B6, 3 and 4-elements semiconductors. Quantum dots, organic, dye sensitized technology show low efficiency (<12%) or low life time and needs further improvement.

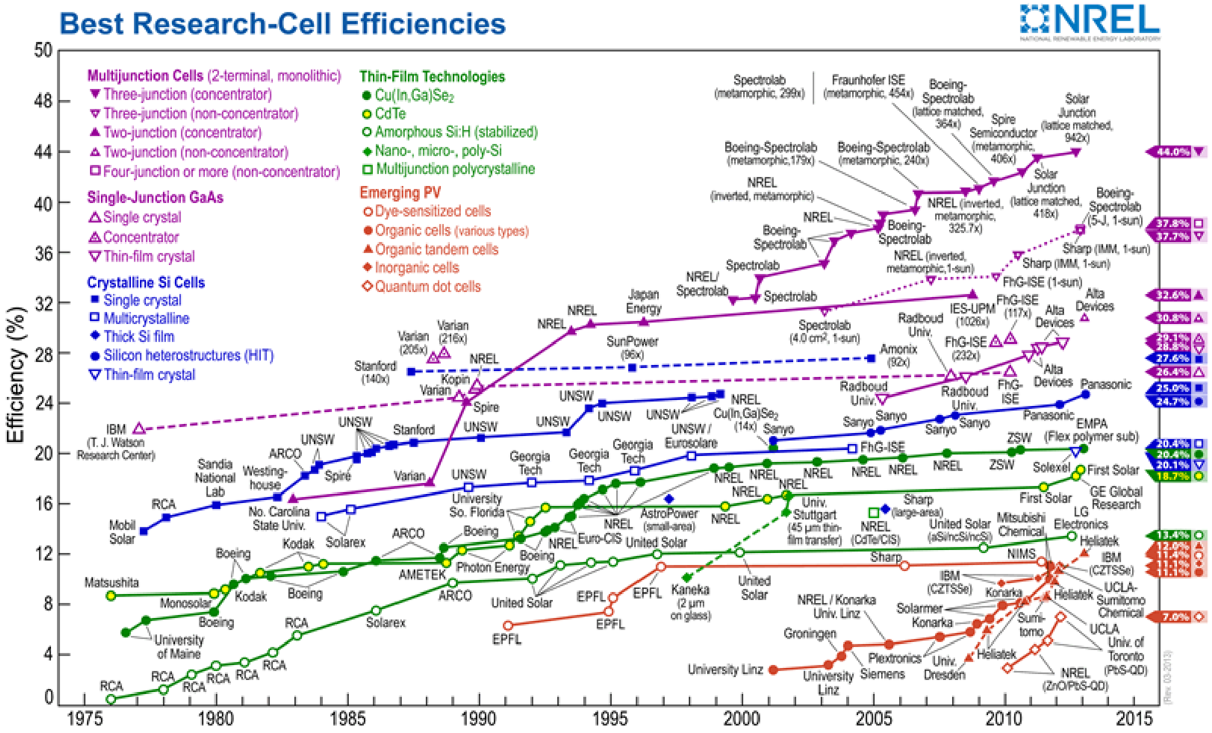


Fig.5.7 The best efficiencies of modern solar cells.

5.2 Why Nano silicon can be used in photovoltaics?

Nanosilicon and porous Si (PSi) formed on crystalline Si wafer by electrochemical or chemical etching exhibits several of optical, electrical and morphological properties that can be useful in the Si solar cells processing (Dzhafarov 2014). The research reports concerning the application of PSi layers for PV began to appear since 1982 (Prasad et al. 1982). The some potential advantages of using of PSi in Si solar cell technology were firstly stated by (Tsuo 1993) and now include the following ideas:

1) Due to the highly textured morphology of PSi, it can be used as efficient antireflection coating and texturing technique to minimize the optical losses in single- and multicrystalline Si solar cells.

2) The band gap of nano PSi can be easily adjusted in the range 1.1- 2.0 eV by varying the porosity. The optimum bandgap for efficient photoconversion[1] thus can be achieved.

3) Wide band gap PSi can be used like a material for the top layer of a heterojunction cell.

4) Strong luminescence of PSi can be used to convert ultraviolet and blue light into longer wavelength emission that has better quantum efficiency in a Si solar cell.

5) PSi layers on Si wafer can act as effective gettering centers for impurity atoms that improves the lifetime of minority charge carriers in Si bulk.

6) PSi can play role of a passivating layer for the front and the back sides of a Si cell.

7) Thin film solar cells can be created using thin film layer transfer with sacrificial porous layers.

8) To enhance long wavelength light absorption in silicon solar cells the multiporous layers like Bragg reflector can be formed on the back side of cells.

9) “All-Si” tandem solar cells can be formed with using of Si quantum wires and quantum dots.

Finaly, very simple technology of PSi fabrication has high potential for large - area applications.

However, there are also many disadvantages associated with using of porous Si in PV. Among them the most important for PV are: *(i)* high resistivity (in the range of 10^7-10^{12} Ω.cm) of nanoporous layers (Ben-Chorin 1997) which eliminates its application to form the junction and *(ii)* chemical composition modification during aging in ambient (Canham 1997)

---

[1] For one junction cell the theoretical limit of efficiency is 31% (or 40,8%) in the case of one solar (or maximal concentrated) flux at optimal semiconductor band gap E_g= 1.3 (or 1.1) eV (Conibeer 2007).

which needs the using of supplementary protective layers. Below we consider how porous Si, nano- Si dots and wires can be used in Si PV.

5.3. Gettering of Si wafers and purification of metallurgical grade Si

It is well known that metallic impurities in solar grade Si create deep electronic levels inside the band gap which act as recombination centers. The presence of large defect densities in Si wafer reduces the lifetimes of minority charge carriers and their diffusion lengths and, as results, strongly limits the efficiency of solar cells. In Si device processing a *gettering* has been widely used to minimize the impact of different contamination (in the first place, the transition metal impurities dissolved in the Si matrix) on device performance. Gettering is a method of reducing impurity concentration in wafer by removing them or localizing them in regions away from the active device regions. The effectiveness of the gettering process depends on the establishment of sinks for absorbing impurities, on the diffusion coefficients of the impurities in bulk semiconductor and in the getter phase, and on the segregation coefficient describing the ability of gettering sites to absorb impurities (Wolf and Tauber 1986).

Nanometer-scaled cavities in PSi provide effective gettering sites [Shieh and Evans 1993; Wong-Leung et al. 1995; Efremov et al. 2000) due to the reaction of the metallic contaminations with Si dangling bonds on the huge developed internal surface of PSi. The probability of tying up contaminants enhances during the post-fabrication annealing step. The gettering of Si wafers consists in the creation of a thin porous layer followed by a thermal annealing in nitrogen or oxygen atmosphere (Tsuo et al. 1996; Khedher et al. 2002; Derbali et al. 2011). The porous layers are removed, along with the gettered impurities, after gettering. Each porous Si gettering process removes up to about 10 μm of wafer thickness. The process can be repeated so that the desired purity level is obtained (Tsuo et al. 1996).

The sufficient diffusion length enhancement was obtained at the combination of porous Si etching and phosphorus diffusion into PSi layer (Zhang et al. 2010; Derbali and Ezzaouia 2012) or Si tetrachloride treatment (Hassen et al.2004). The last technique consists of heating a *p*-Si substrate with a thin porous layer (formed on both sides of wafer by stain-etching in a $HF/HNO_3/H_2O$ solution) under $SiCl_4/N_2$ atmosphere. For example, if for untreated reference sample the diffusion length L_d=86 μm, the sample treated at 950°C in N_2 atmosphere for 60 min shows L_d=150 μm, the sample treated at 950°C in $SiCl_4/N_2$ atmosphere for 90 min displays L_d=210 μm. The L_d increasing of p-Si wafer (from 63 to 134 μm) was obtained when porous Si etching combines with Al layer deposition followed by annealing at 750°C (Evtukh et al. 2001).

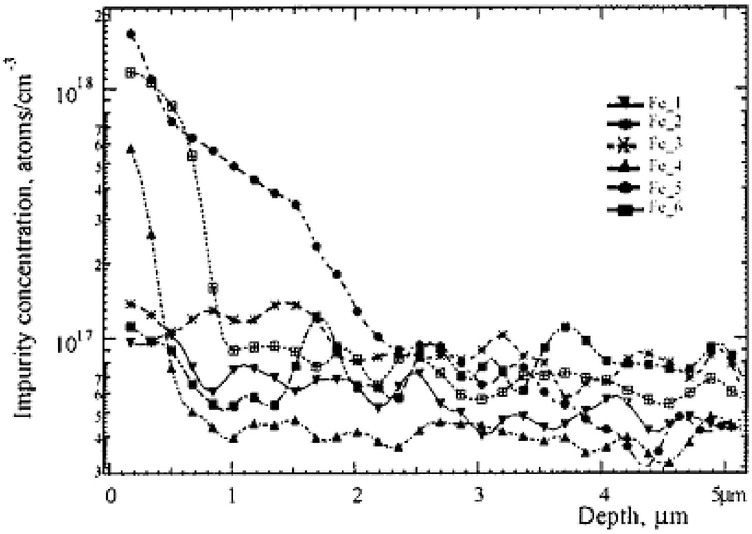


Fig. 5.8 SIMS depth profiles for iron for set of samples at different process conditions (4,5,6 - with porous Si etching, 1,2,3 – without porous Si ething; 2,5 –with light annealing, 3,6- with conventional annealing) [Menna 1998].

Gettering properties of porous Si open the possibility to apply it for the *purification* of metallurgical grade Si. Their importance for PV industry today is obvious. Now, most of commercial Si solar cells are made on wafers sliced from ingots grown from molten Si by Czochralski, casting methods or Si ribbons grown directly from the Si melt. High-purity poly Si rods are normally used as the feed materials for the Si melt. Poly Si is usually obtained by reducing high-purity trichlorosilane, which is converted from metallurgical grade Si. The development of a low-cost supply of Si feedstock still remains of important task since the cost of purifying the trichlorosilane and converting it by CVD to poly Si rods is high (Menna et al. 1998).

Solar grade Si can be obtained by purifying of metallurgical grade Si using pyro- and hydro-metallurgical process, where typical treatments of the melted Si such as casting, grinding, leaching, remelting and controlled solidification are performed (Lian et al.1992). Alternative technologically simple method of purification of metallurgical grade Si by porous Si etching was proposed (Menna et al. 1998). It includes the sequence of porous Si etching in $HF/HNO_3/H_2O$ solution and conventional (or light) annealing at temperature of 1000°C for 30 min in argon of 600 torr. The accumulation of the impurities (B, Al, Cr, Fe, Cu) near the surface in first 1-2 µm is observed, when the metallurgical grade Si wafer undergoes the process sequences PSi etching/annealing. Furthermore, the combined effect of porous Si etching and light-annealing has a marked influence over the diffusion of heavy metal impurities and boron from the bulk to the surface.

5.4 Reduction of optical losses.

There are a number of ways to reduce these losses:

1. Top contact coverage of the cell surface can be minimised (although this results in increased series resistance).
2. Antireflection coatings can be used on the top surface of the cell. A *quarter wavelength* antireflection coating; that is, a transparent coating of thickness *d*1 and refractive index *n*1, such that

$$d_1 = \frac{\lambda_0}{4n_1}$$

will, ideally, cancel the light reflected from the top surface by interference effects from the light reflected at the coating-semiconductor interface, which will be 180° out of phase. This is illustrated in Fig.5.9.

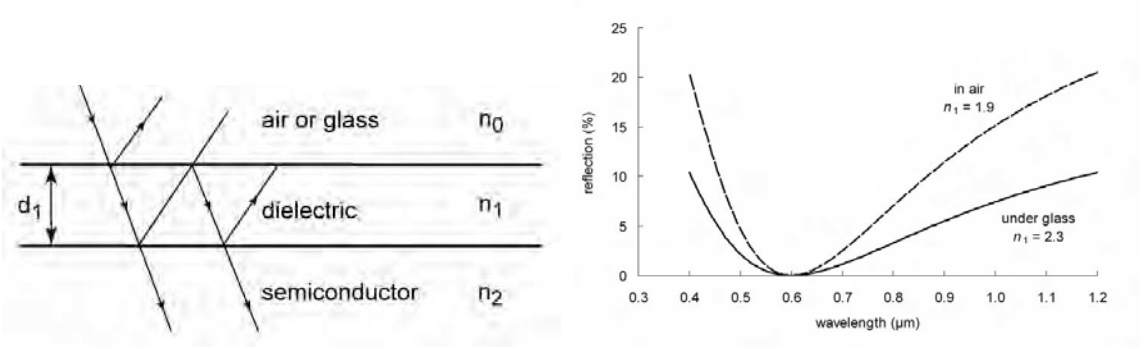


Figure 5.9 a) Use of a quarter wavelength antireflection coating to counter surface reflection. b)Surface reflection from a silicon cell (*n*2 = 3.8) in air (*n*0 = 1) and under glass (*n*0 = 1.5) with an antireflection coating with refractive index and thickness chosen so as to minimise reflection for 0.6 □m wavelength light.

Reflection is further minimised if the refractive index of the antireflection coating is the geometric mean of that of the materials on either side—glass (typically) or air, and the semiconductor—that is, if

$$n_1 = \sqrt{n_0 n_2}$$

Surface reflection can be reduced in this case to zero, as shown in Fig. 5.9.

3. Surface texturing can also be used to minimise reflection. Any ‘roughening’ of the surface reduces reflection by increasing the chances of reflected light bouncing back onto the surface, rather than out to the surrounding air. The surface of crystalline silicon can be textured uniformly by etching along the faces of the crystal planes. The crystalline structure of silicon results in a surface made up of pyramids, if the surface is appropriately aligned with respect to the internal atoms. An electron microscope photograph of a textured silicon surface is shown in Fig. 5.10.

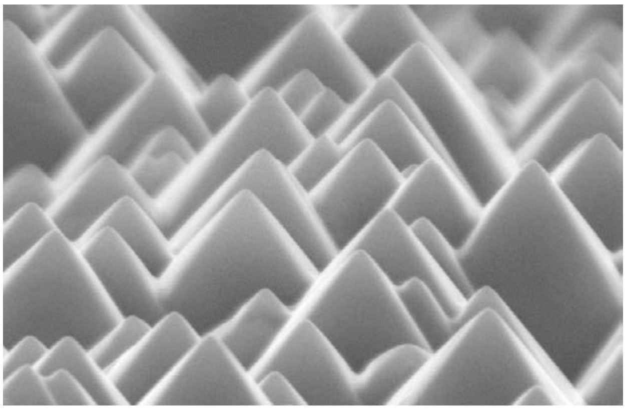


Figure 5.10. Scanning electron microscope image of a textured silicon surface.

An additional benefit of roughened or textured surfaces is that the light is obliquely coupled into Si in accordance with Snell`s law:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

Where θ1 and θ2 are the angles for the light incident on the interface relative to the normal plane of the interface within the mediums with refractive indices n1 and n2 respectively

5.5. MACROPOROUS TEXTURIZATION OF SI WAFERS

5.5.1. Comparison of Alcaline and Macroporous Texturisation

One of the factors which lead to an increase in solar cell efficiency is the reduction of light losses in cell by lowering the surface reflectivity and improving the scheme of light trapping. It can be achieved by surface *texturization* which plays an important role in production both of mono- and multicrystalline Si solar cells (Springer et al. 2000).

Traditionally the standard alcaline texturization (NaOH) has been used in the solar cell industry. Although the usual alcaline etching procedures work well on monocristalline Si (Fig. 5.11a), on multicrystalline wafers (Fig. 5.11b) they are less efficient. Indeed, the etch rate of Si in alcaline solutions is highly anisotropic, i.e. is strongly dependent on the crystallographic grain orientation. Etch rate of Si (111) is considerably lower (100 - 200 times) than that one of Si (100) and Si (110). Alcaline etching of multicrystalline Si wafers having the random grain orientation results in coarse and non-uniform surface morphology (Fig. 5.11b). The alkaline texturization leads to the effective reflectivity (R_E)[2] of about 20% at best for multicrystalline substrates and 12% for (100) monocrystalline ones. It is expected that uniform texturization of multicristalline Si wafer can be obtained by technology of macropore formation.

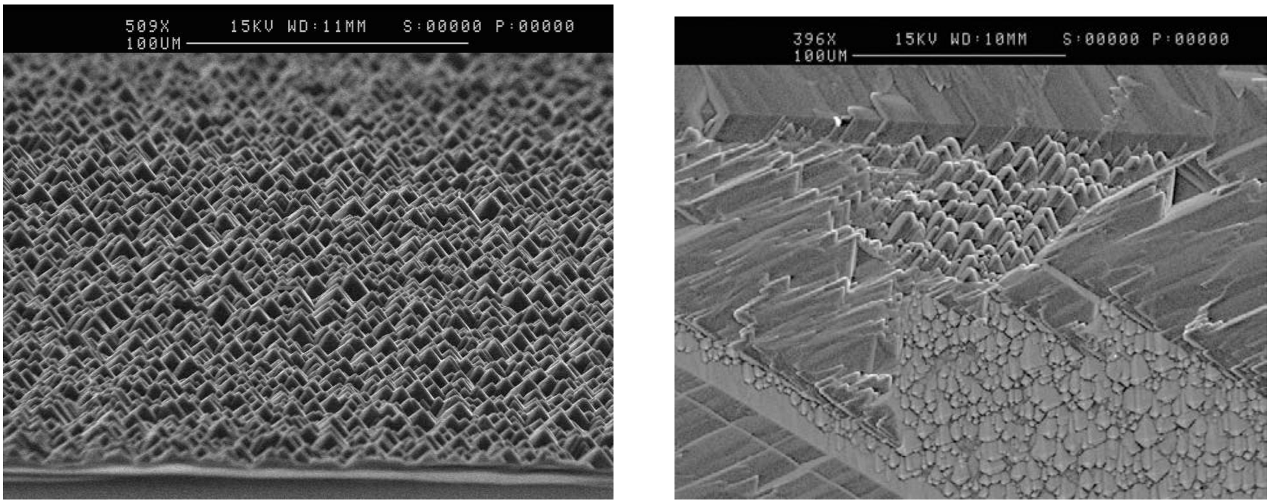


Figure 5.11. Alcaline texturization of (a) mono- (b) and multicrystalline Si wafers. The SEM images were kindly provided by Photowatt Technologies.

Electrochemical oxidation of *p*-type Si in a mixture of hydrofluoric acid with some organic solvents leads to the isotropic formation of macroporous layers (Levy-Clement et al.2003; Bastide et al. 2004), which can be very efficient in lowering the reflectivity of Si wafers. This type of texturization can be performed from a chemically polished surface as well as a saw damaged surface. In the latter case the saw damage removal and the texturization are done simultaneously which is very important for industrial applications. Fig. 5.12 presents a SEM image of macroporous textured multicrystalline Si for (100) and (111) oriented grains. This type of isotropic texturization assures an effective reflectivity of 9-11% on multicrystalline Si wafers. It was found that the main parameter affecting the texturization efficiency is the substrate resistivity. The best results of R_E=9% on *p*-type multicrystalline Si were obtained on low resistivity substrates (0.4 Ωcm) (Bastide et al. 2004). The macroporous texturization is more efficient for monocrystalline Si. R_E <5% on *p*-type monocrystalline Si was obtained on (100) substrate with resisitivity of 0.2 Ωcm (Levy-Clement et al. 2003).

---

[2] The effective reflectivity is the reflectance R(λ) normalized to the solar flux N(λ):
$R_{eff}=\int R(\lambda)N(\lambda)\,d\lambda \,/\, \int N(\lambda)\,d\lambda$

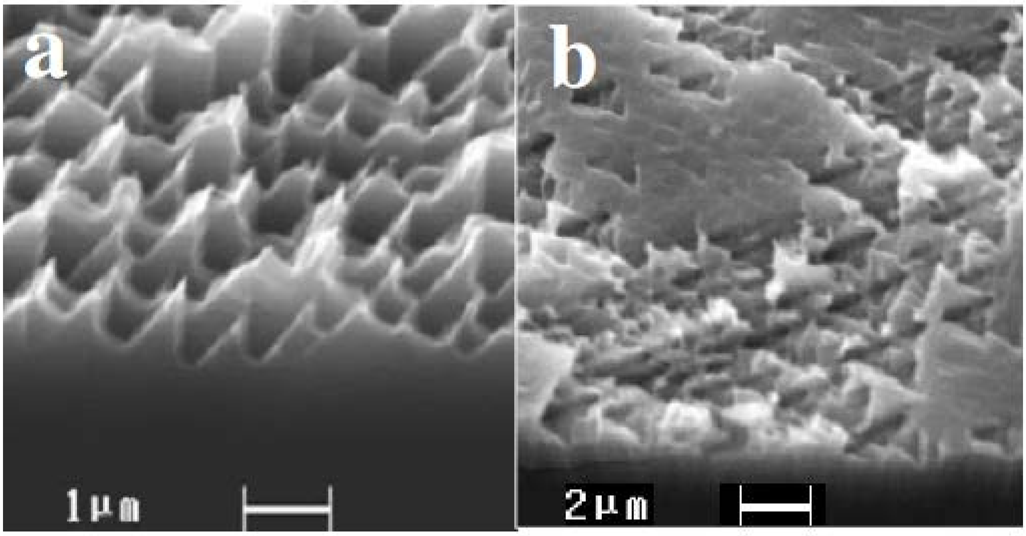


Fig.5.11 SEM images of anodized 0.4 Ω cm multicrystalline Si (POLIX) Si with an optimized morphology for photovoltaic application. 4 M HF/ dimethylsulfoxide, Q = 4.5 C/cm^2. a) (100) oriented grain, b) (111) oriented grain. (Lévy-Clément C. et al., *Phys. Stat. Sol. A.* **197**, 27, 2003)

N-type doped multicrystalline Si presents potential advantages over *p*-type Si, in particular higher diffusion lengths at a given impurity concentration and improved long term stability (Martinuzzi et al. 2005). Micrometer sized macropores have also been isotropically formed on n-type multicrystalline Si wafers under white light illumination, leading to an overall effective reflectivity as low as 8.5% (Grigoras et al. 1998; Bastide et al. 2006; Tena-Zaera et al. 2007).

Despite rather good results in decreasing of reflectivity loses, the electrochemical method possesses several considerable disadvantages which limit its industrial implementation. The necessity to create a good conductive contact on the backside of each sample as well as a development of corresponding equipment to form complete electrode system are some of them.

In turn, the macroporous formation by *chemical etching* does not require the electrical bias applying and corresponding good rear contact formation. The technology of this type of texturization is rather simple and requires only two steps. Description of this method one can find in the Chapter “Alternative methods of silicon porosification” presented earlier. First, the samples are dipped into electrolyte on the base of HF:HNO$_3$ to carry out the etching process and then wafer rising to remove the etching residuals. Contrary to the electrochemical method the macroporous formation occurs on the both sides of the wafer (Yerokhov et al. 2002; Marrero et al. 2008). Figs. 5.12a and 5.12b present the honeycomb surface texturing obtained by this type of isotropical texturization on monocrystalline and multicrystalline Si (Yerokhov et al. 2002). This type of texturing becomes more and more expanded in PV industry. One of the derivatives of the present method is the chemical vapor etching technique (Ben Rabha et al. 2005).

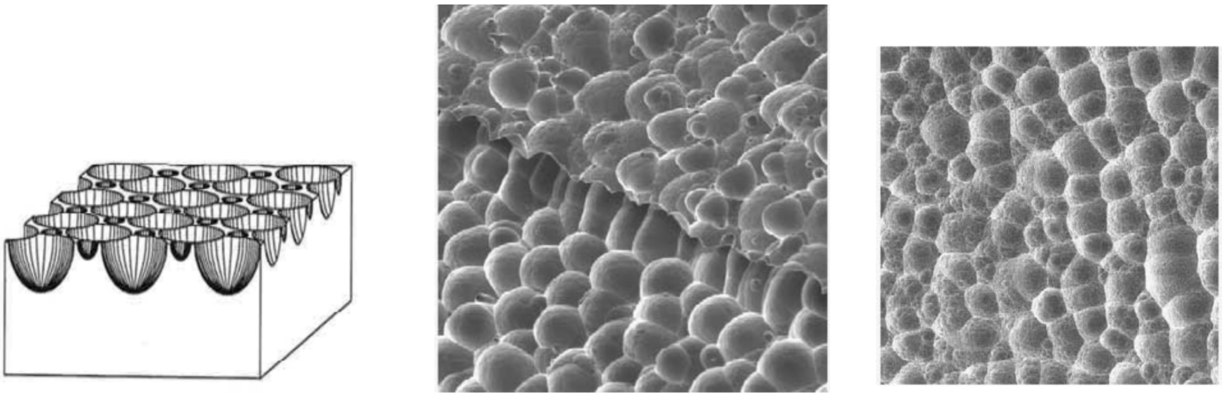


Fig.5.12. Schematic view of honey comb texturing (a) and (b) SEM images of chemical texturing obtained on p-type monocrystalline (a) and multicrystalline silicon (b). (Yerokhov V.Y. et al., *Solar Energy Mater. Solar Cells,* **72**, 291, 2004.)

The comparison of experimental spectral dependence of the global reflectance of polished silicon, random pyramids and chemically and electrochemically prepared macroporous textures, respectively are presented in Fig.5.13.

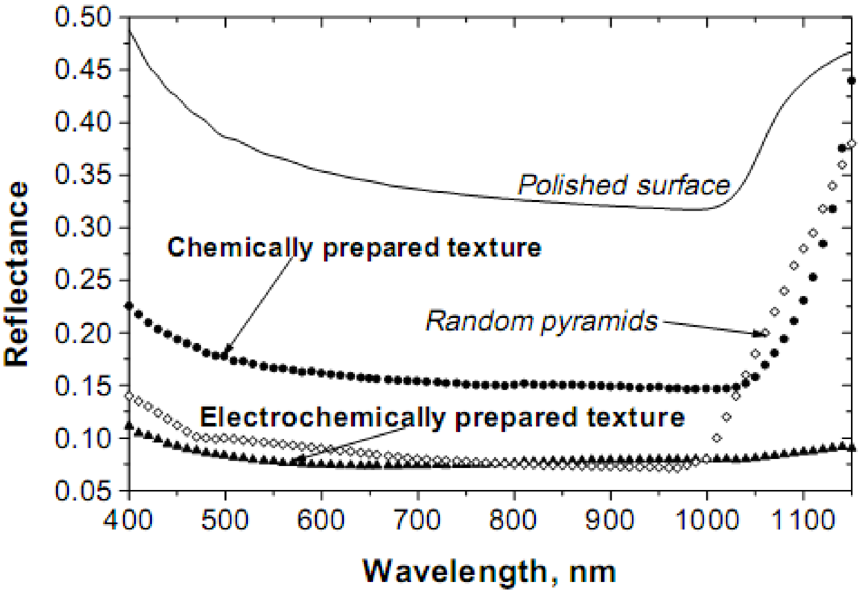


Fig.5.13 Measured spectral dependence of the global reflectance of polished silicon, random pyramids and chemically and electrochemically prepared macroporous textures, respectively (Yerokhov V.Y. et al., *Solar Energy Mater. Solar Cells,* 72, 291, 2004)

Metal assisted etching is another promising approach for surface texturization of Si wafes. Usually metal assisted etching occurred in HF/H_2O_2/H_2O solutions with Ag nanoparticles as catalyst agent (Yae et al. 2005; Chartier et al. 2007; Chaoui et al. 2008; Bastide et al. 2009). This method implied several steps. The first one is electroless deposition of Ag nanoparticles on the surface of Si wafer. Next is the immersion of Ag-loaded Si samples into an HF/H_2O_2/H_2O solution. Ag nanoparticles sink into Si and pore opening with diameter similar to that of the nanoparticles are obtained.

Metal assisted etching forms a meso PSi layer that is subsequently dissolved by NaOH to reveal a macrotextured surface. It should be noted that this texturization technique is anisotropic. Rounded macropores with inversed pyramids at the bottom are observed for (100) oriented Si. The (111) Si surface exhibits craters with a triangular form. The best effective reflectivity result of 5.8% (Chartier et al. 2007) was obtained for (100) oriented monocrystalline samples, compares to a standard NaOH etching of 12%. For multicrystalline substrates the lowest obtained effective reflectivity was 12% (Fig.5.14a) (Bastide et.al. 2009). One of the disadvantages of this texturization technique for PV applications is the necessity to dissolve residual Ag nanoparticles.

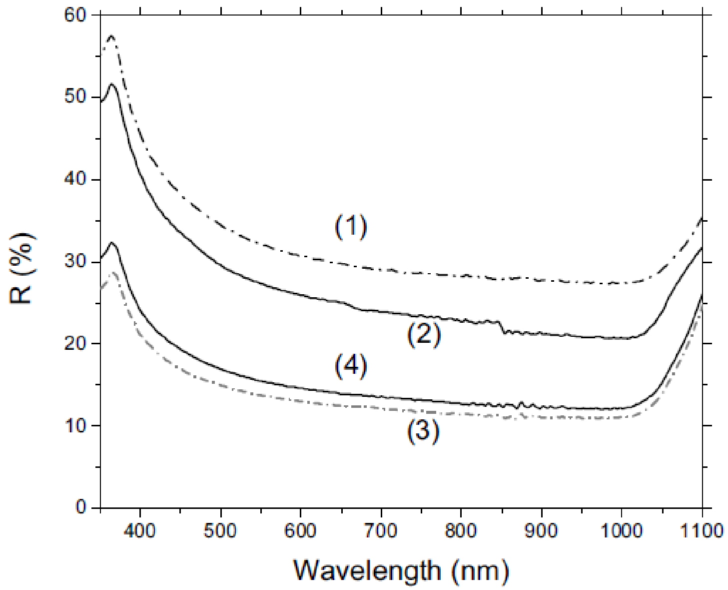


Fig.6 a) Reflectivity spectra of mc-Si: (1) as cut (R_E = 30 %);(2) after NaOH texturisation (25 %); (3) after metal assisted etching (MAE) + PS dissolution for a time t (12 %) and (4) for a time t+2 min (15 %). (Bastide S. et al., *Phys. Stat. Sol. C.* 6,1536, 2009)

5.5.2 Influence of texturization techniques on solar cell performance

Reducing of light loses by lowering the surface reflectivity results in an increase of short circuit current (I_{sc}) which is one of the basic solar cell parameters. It was found that the I_{sc} of electrochemically textured samples can reach 37 mA/cm^2 against 24 mA/cm^2 for the control (no texturing) samples (Gamboa et al. 1998). However, despite the benefit of reduced reflectivity, the internal quantum efficiency (IQE) of the metal assisted etched and electrochemically textured cells are lower over the entire spectral region with a significant difference in a short wavelength region (Bastide et.al. 2009) (Fig. 5.14). The higher recombination loses in the emitter are probably due to the larger surface area of macroporous textured surfaces that results in a larger volume of emitter. This statement is an agreement with the higher fill factor (FF) obtained in the case of macroporous textured cells. Higher FF can be explained by the increase of the contact surface. The lower spectral response of the macroporous textured cell eliminates partially the benefit of their reduced reflectivity. I_{sc} is only 2% higher (vs. NaOH cell) while +6.5% could be expected from the difference in effective reflectance R_E.

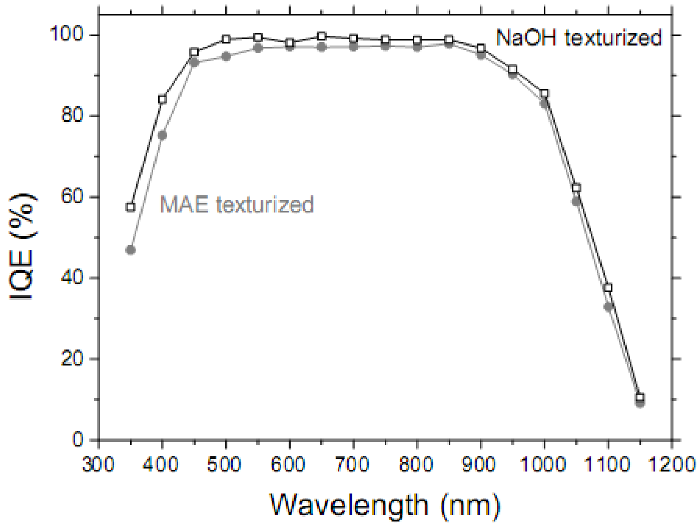


Fig.5.14 Internal quantum efficiency (IQE) as a function of the wavelength in a similar grain of MAE (solid circles) and NaOH texturised solar cells (open square) (Bastide S. et al., *Phys. Stat. Sol. C.* **6**,1536, 2009.)

With respect to solar cell performance, the electrochemical texturization leads to an improvement in solar cell efficiency by 0.8 % (Bastide et al. 2004) when compared to commercial alkaline etched cells. However, this improvement is not due to an increase of photocurrent response as expected but to a higher FF. Additional optimisation of the surface passivation of macroporous textured cells should be performed.

5.6. NANO PSi AS ANTIREFLECTION COATING

The reduction of reflection losses can be realized also by forming of *antireflection coating* (ARC) in form of quarter wavelength dielectric layer on front surface of cell. The principle of a quarter wavelengths ARC is well – known. Light reflected from the dielectric layer - semiconductor interface arrives back at the dielectric-air interface 180° out of phase with that reflected from the dielectric-air interface, cancelling it out to some extent. Refractive index of dielectric layer should obey to:

$$n = \sqrt{n_{air} n_{Si}}, \qquad (10.1)$$

where n_{air}=1, n_{Si}=3.8. Then a single layer ARC on Si has a refractive index of about n=1.9 (Green 1992).

To prevent the penetration of humidity into PV modules and to ensure their long term stability in the standard PV modules the cells are normally embedded in a glass or encapsulated by different materials like ethylene vinyl acetate or resin (Nagel et al. 1999) with a similar refractive index to glass (n=1.5). This increases the optimum value of the index of the ARC to about 2.3. Usualy, one layer coatings are chosen to produce minimum reflection at about 600 nm. The use of multiple layers of ARC materials can improve performance. The design of such coatings is more complex, but it possible to reduce reflection over a broad band.

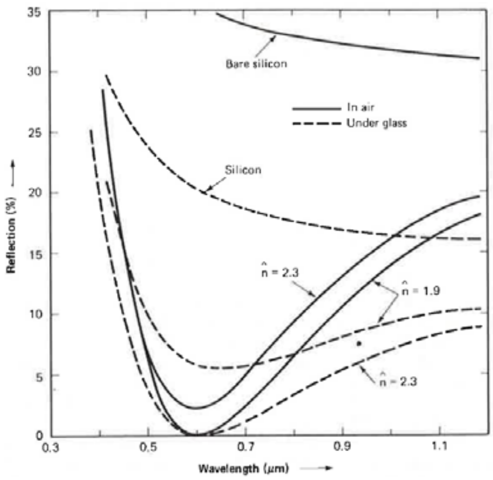


Fig.5.14 Reflection curve of normally incident light reflected from bare Si and Si with antireflecting coating of refractive indices of 1.9 and 2.3.

Since 1997 and up to 2003 an active research work has been paid to find an alternative to TiO_2, ZnS and MgF_2 historically used as ARC for Si cells (Nagel et al. 1999). Indeed, three requirements should be fulfilled when applied in multicrystalline solar cells manufacturing: *(i)* ARC should be low-absorbing in the visible/near IR spectral region, *(ii)* provide surface passivation and *(iii)* induce bulk passivation of the Si wafers.

The use of nanoporous Si as ARC was explored since at least 1997. Thin porous layer is grown on emitter (n^+ or p^+) of preliminary textured mono- or multicrystalline (p/n^+ or n/p^+) Si junctions. The thickness and porosity of porous layer (and respectively the refractive indices) are chosen to produce a minimum reflection at 600 nm. Usually, the thicknesses of thin porous layers do not exceed 100 nm for three reasons:

(i) in order to decrease the influence of the parasitic absorption in porous layer itself,

(ii) taking into account that the thickness of the standard shallow emitter is about 300nm, the electrochemical or stain etched PSi layer should not to damage the emitter;

(iii) in order to reduce the influence of high resistivity of porous Si on series resistance.

There are three main benefits in using of nano PSi as ARC. Firstly, the refractive index of this material covers continuously the range between that of crystalline Si and air. Secondly, it is expected, that chemically treated porous layer can assure a good passivation of emitter. And thirdly, during the porous Si formation the superficial part of emitter is etched back thus removing the dead layer from the highly doped surface[3].

Two main techniques of thin porous layer formation have been employed: electrochemical (Strehlke et al. 1997; Krotkus et al. 1997; Lipinski et al. 2003) and chemical (“stain etching”) (Bilyalov et al. 1997; Schirone et al. 1998; Schnell et al. 2000; Lipinski et al. 2003). The both have relative merits: in particular the main advantage of stain etching is related to the lack of any electrical contact during the process. This makes it particularly suitable for large area processing, which is prerequisite for perspective scale-up in industrial production. On the other hand, porous Si layers obtained by chemical etching exhibit a porosity gradient (Strehlke et al. 1997; Lipinski et al. 2003) which causes the shift of the minimum of reflectivity curve. The electrochemical approach is so preferred, as it gives better reproducibility over stain etching process.

The images of PSi layers formed by electrochemical and chemical methods on the surface of n^+ emitter are presented in Figs. 5.16a and 5.16b, respectively. The effective reflectance

---

[3] In n^+/p Si solar cells, regions of excess phosphorous would lie near the surface of the cell. This can produce a “dead layer” near the surface where light-generated carriers have very little chance of being collected because of the very low minority-carrier lifetimes. This “dead layer” normally should be etched back.

coefficient can be reduced to about 1-3% in the whole visible range (Schirone et al. 1998; Lipinski et al. 2003) (fig.5.17).

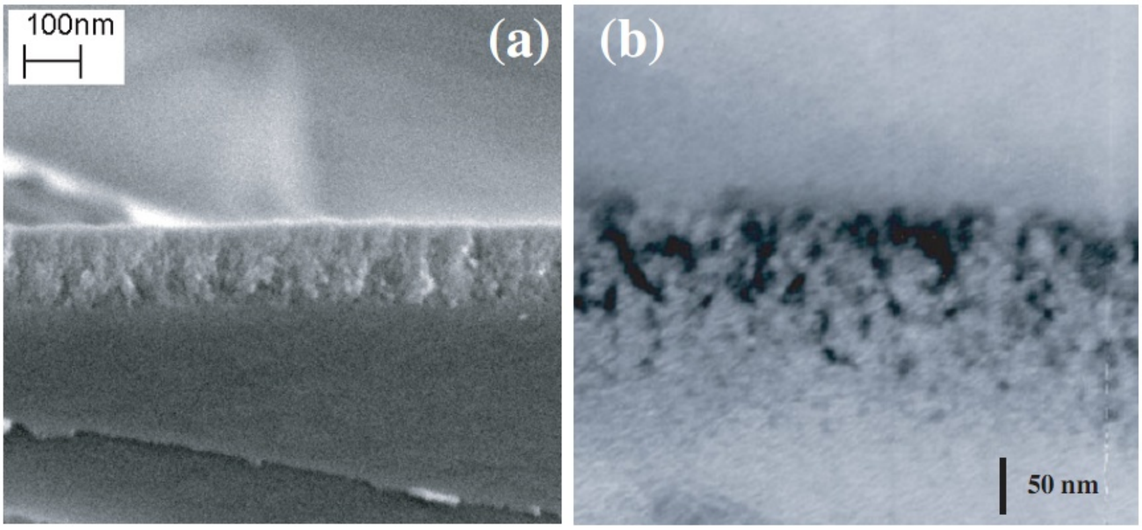


Fig.15.16 Cross section images of a PSi layer formed by a) electrochemical etching (SEM micrograph) and by b) chemical etching (TEM micrograph). (Lipinski M. et al., *Phys. Stat. Sol.A.* **197,** 512, 2003)

A lot of efforts have been made in determining the formation conditions which led to the PSi layers with optimal optical properties resulting in efficiencies comparable to those obtained for commercial multicrystalline solar cells with classical ARC (Strehlke et al. 1997). Single or multilayer structures on the base of PSi (Striemer et al. 2003; Aroutiounian et al. 2004) have been calculated and realized. For example, in diamond-like carbon (DLC) /porous Si double ARC the DLC layer supplementary acts as a protective coating for PSi, increases solar cell stability relating to the effect of proton and UV irradiation, making such solar cells especially promising for space application (Litovchenko and Kluyi 2001; Aroutiounian et al. 2004).

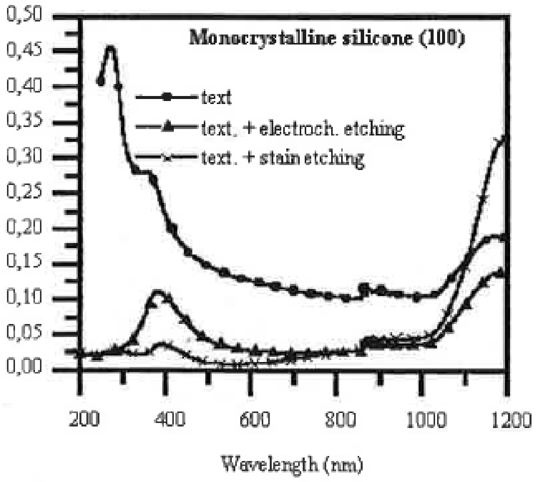


Figure 5.17 Measured total reflectance of PS layer samples in comparison with monocrystalline Si with mirror-polished surface

PSi film which is etched in the shallower portion of the n^+/p junction gives a positive contribution to photocarrier collection efficiency. It reduces optical loses due to the ARC effect and decreases the absorption in the dead layer, allowing the photocurrent increase (Bilyalov et al. 1997; Lipinski et al. 2000).

Besides Fresnel interference thin porous Si layers also display component of diffuse scattering. Indeed, by analyzing the influence of the porous surface on the internal quantum efficiency (IQE) of Si solar cell it was shown that porous layer of 160 nm thick acts as perfect light *diffuser* with an effective entrance angle of 60^0 for the light entering the cell structure after passing through the porous surface layer (Fig. 5.18, insert) (Stalmans et al.1998). This increases the effective length path of incoming photons and IQE in long wavelength region (Fig. 5.1.8). Diffuser effect can be explained by Rayleigh's scattering which occurs in medium if diameters of particles are no larger than about 1/10 of wavelength of incident light and spatial distribution of these particles is fully randomised. Indeed, the dimension of nano PSi obeys to the above mentioned criteria. Theoretical analysis of light scattering in PSi for PV applications were performed (Abouelsaood et al. 2000).

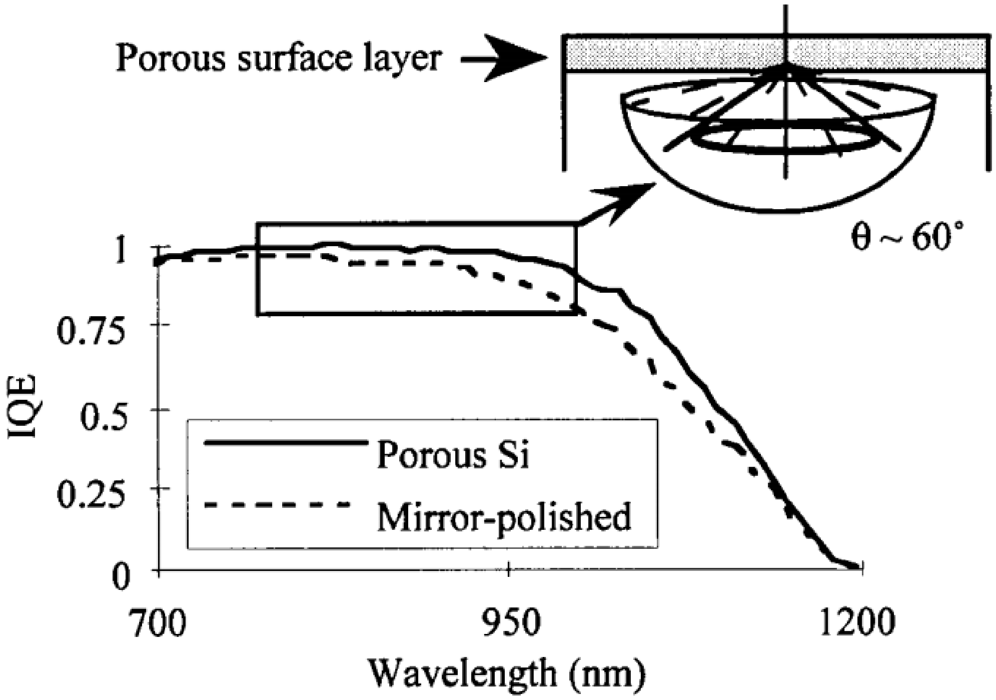


Fig.5.18 Measured IQE data for conventional cell structure with a mirror-polished and porous top surface (Stalmans L. et al., *Prog. Photovolt. Res. Appl.* **6,** 233, 1998)

It was examined also the possible improvement of multicrystalline Si solar cells parameters at the creation in the *n*$^+$-emitter the selective regions in form of PSi pipes (Skryshevsky et al 1999, 2000). It is expected that the light scattering in PSi leads to: *(i)* growth of incoming photon density in parts of the p-n junction between pipes (concentrator effect), *(ii)* increasing of the effective length path of low adsorbed photons of long wavelength part of the sun illumination, *(iii)* possibility of partial absorption of high energy photons in low doped p-base instead top n$^+$ emitter that increases the photogeneration carrier collection in short wavelength region. The difference in crystal orientation of Si grains results in the dispersion of the measured curves of external quantum efficiency (EQE) after PSi forming. However, the general tendency of EQE increasing in the long wavelength region of the sun illumination is observed for the cells with different PSi pipes density. In the short wavelength region the behavior of EQE is more complicated due to the partial light absorption in PSi, light scattering by PSi and change of surface recombination rate in PSi/Si interface. The growth of PSi pipe area increases series and shunt resistances (due to the partial replacing of the top part of the emitter by the high resistive porous layer).

However, technology of PSi ARC displays few disadvantages. During the porous Si fabrication the most heavily doped portion of the n$^+$ region is also etched away, reducing the internal electrical field and eventually the open circuit voltage V_{oc} of the cell. A degradation of the spectral response in the short wavelength (Fig. 5.19) (Bilyalov et al 2000) was rather important for the cells with a PSi ARC compared to cells with classical dielectric ARC. This degradation was caused by the recombination of photogenerated charge carriers at the surface and/or at porous Si/bulk Si interface. A lot of efforts have been undertaken to passivate porous Si ARC .

FF degradation of about 10% is usually observed after formation of PSi. Indeed, if the PSi is made after the contacts, the metallic grids are etched in HF. So, the contacts should be protected by a polymeric film and short etching times (~10 s) should be applied. Another solution is to fabricate the PSi before the contacts. It was shown (Matic et al. 2000) that the high temperature annealing required for contact firing slightly increase the porosity and the thickness of porous layers. As a result no significant changes in reflectance characteristics are observed. Therefore, firing of the contacts through PSi does not destroy its antireflection properties and can be successfully used as an alternative way for preparation of porous Si ARC for Si solar cells.

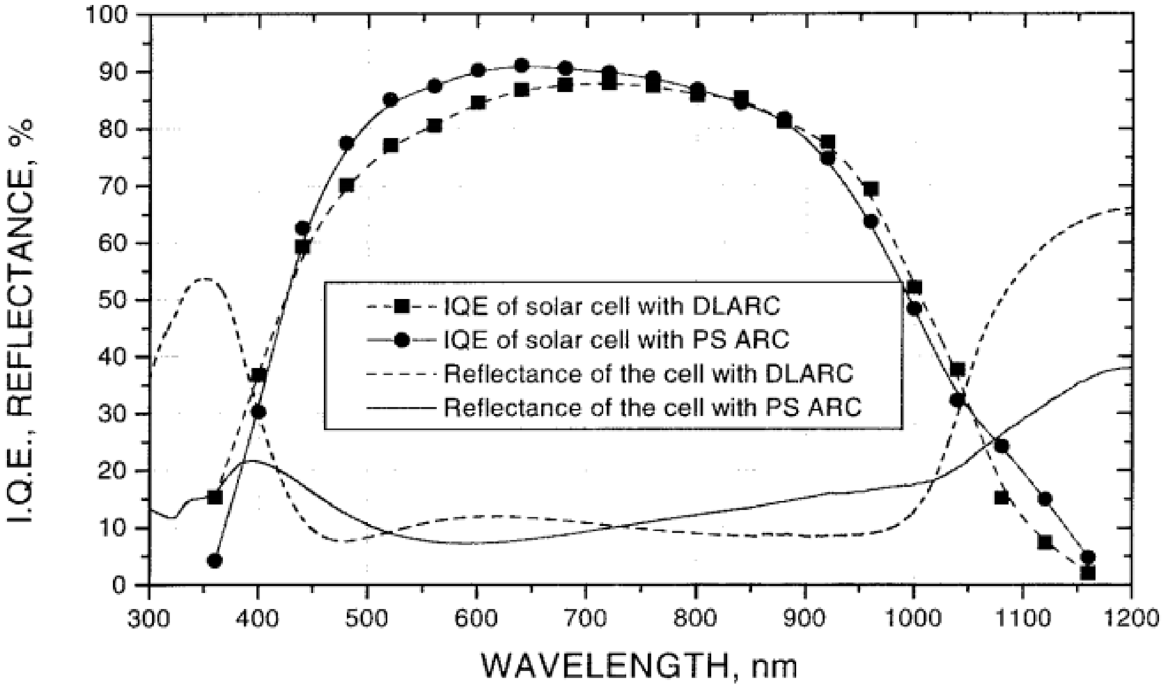


Fig.5.19 IQE and reflectance characteristics of the cells with PSi and double layer TiO_2/MgF_2) ARC. (Bilyalov R.R. et al., *Solar Energy Mater. Solar Cells,* **60**, 391, 2000.)

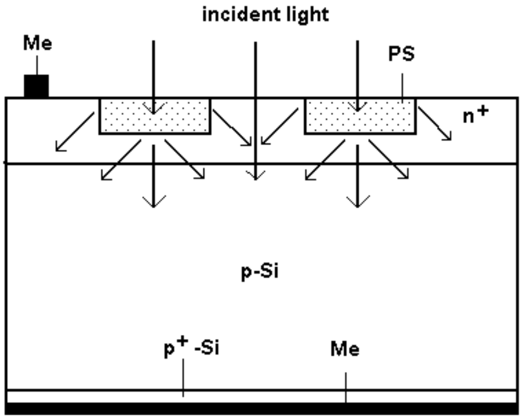
Since 2004 hydrogenated Si nitride (SiN_x:H) has become a dominant material for the ARC in the modern PV technology. Indeed, SiN_x:H layers perfectly fulfill three requirements when applied in solar cells manufacturing: act as non-absorbing ARC, provide surface passivation and induce the bulk passivation of the Si wafer. The letter effect can be accomplished by a thermal activation of the layer (e.g. during a firing of a screen-printed metallisation) by which hydrogen diffuses from the SiN_x:H into the Si wafer where it passivates crystal defects and impurities (Duerinckx and Szlufcik 2002). At the same time the surface passivation is improved by reducing the number of dangling bonds and by creating a positive field effect due to positive charges naturally present in SiN_x:H layers (Aberle 2000). Now there is not a strong activity in PSi implementation as ARC of industrial solar cells, however, the articles devoted to the optimistaion of PSi mono and multilayers ARC we can find in the literature (Ben Rabha et al. 2011; Ramizy et al. 2011; Hsueh et al. 2011, Osorio et al. 2011; Salman et al. 2012; Thogersen et al. 2012; Dzhafarov et al. 2012; Dubey 2013).

Light trapping was shown is improved by inserting porous silicon pipe arrays within Si substrate (Figure 5.20). The main benefits of Si solar cells with porous silicon pipes were expected to be:

- The light scattering in porous silicon leads to the growth of incoming photon density in active parts of the p-n junction. Therefore, the cell efficiency can be enhanced owing to the concentrator effect: I_{sc}~sun flux, and $V_{oc}=nkT/e \ln(I_{sc}/I_d+1)$ enhances as logarithm of sun flux;

-The increasing of the effective length path of incoming long wavelength photons;

- The possibility of partial absorption of high energy photons in low doped p-base instead top n^+ emitter. This increases the photogeneration carrier collection in short wavelength region;



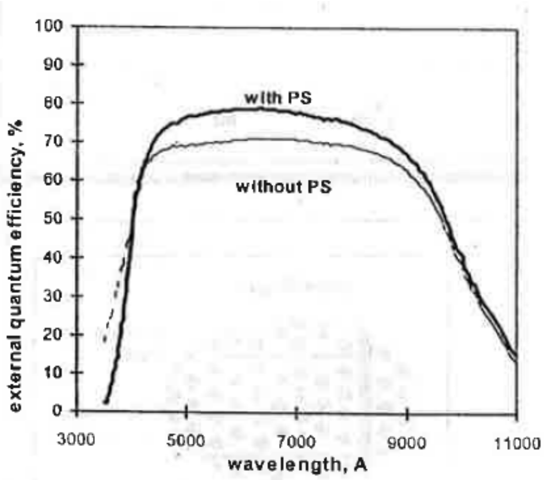


Figure 5.20 Design of improved Si cell with PS pipes. Ray path is shown for two pipes.

Figure 5.21 External quantum efficiency of cell without initial ARC before and after PS pipes forming .

On the other hand, the forming of porous silicon pipes could deteriorate cell parameters due to increasing number of recombination centres, increasing of leakage currents in p-n junction, optical losses of light absorption in porous silicon without photogeneration carrier collection and the increasing of series resistance at the partial replacing of heavy doped n^+ layer by high resistive porous layer. **Fi**gure 5.21 shows the external quantum efficiency of cells without initial ARC before and after porous pipes forming. One can see the growth of the EQE in short wavelength region (effect $1/\lambda^4$ of Rayleigh’s scattering) and in IR spectral region (owing to increasing of light length path). The sharp reducing in blue region can be explained by optical loses due to absorption in PS pipes and recombination losses.

5.7 Influence of porous Si emission on quantum efficiency of solar cell

Since nanoporous silicon displays efficient red-orange photoluminescence at room temperature it will be very attractive offer to apply the effect of conversation of UV-blue part of sun flux absorbed by PS layer into long wavelength light having higher efficiency for silicon SC [Tsuo 1993]. Indeed, the efficient excitation of photoluminescence in nanoporous

Si is occurred at the absorption of photon with energy $h\nu$>2.8 eV (λ<450 nm), that is in the region where Si solar cell photoresponse reduces drastically. Numerically simulation and experimental verification of *re-emission effect* influence on solar cell parameters was fulfilled [Skryshevsky 1996]. As it was shown, effect of re-emission is relatively small and well observed only at high photoluminescence yield and for thick (0.25-2 μm) PS layers which are not compatible with the routine porous silicon ARC.

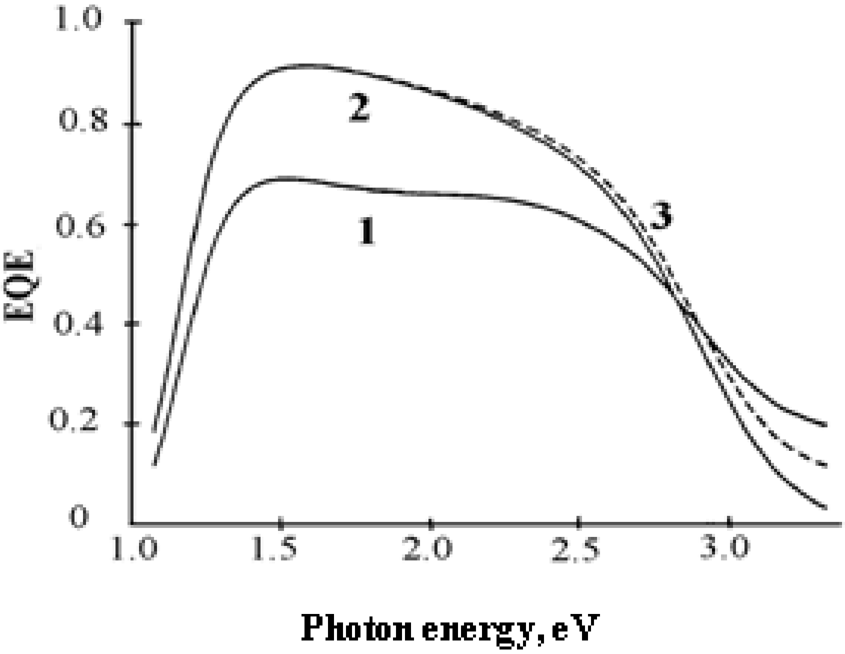


Figure 5.22. External quantum efficiency of Si cells without (1) and with 1 mm PS layer on textured Si surface (2), and at the accounting of re-emission with efficiency 10% (3).

Figure 5.22 shows the simulated external quantum efficiency of standard silicon SC with conversation efficiency of 15% at the presence of 1 μm PS layer taking into account the reduction of reflectivity, optical absorption and photoluminescence yield (10%) in PS layer. Moreover, the application of this effect for silicon sollar cells is restricted by difficulty to obtain effective porous Si photoluminescence on n^+ emitter and photoluminescence quenching at various thermal annealing at cell processing (e.g. at firing of contacts).

5.8. Porous Si selective emitter

The optimal emitter parameters (thickness, doping) of Si solar cells with *homogeneous emitter* is a result of a compromise of the requirements of low dark current and low contact resistance. For passivated surface emitter the surface dopant concentration must be moderated and the emitter must be shallow to keep the dark emitter saturation current density low. However, for nice electrical contact the surface dopant concentration must be higher than 10^{19} - 10^{20}cm^{-3} for n-Si and emitter must be relatively thick (>0.3 μm). Therefore, the homogeneous emitter needs such high surface dopant concentration and emitter thickness that makes the surface passivation almost impossible since recombination increases with doping. For a non screen-printed contact the optimal homogeneous emitter has a sheet resistivity of 60 Ω/□ and a doping level of 2.10^{19}cm^{-3} (Kuthi 2004). But such a highly doped homogeneous emitter decreases the response to short wavelength illumination due to high recombination velocity.

In solar cell devices with *selective emitter* the high doping regions are just under the metallization and sheet resistivity can be increased up to 70-200 Ω/□ between the contacts. Therefore, the recombination losses decrease and a good Ohmic contact and high FF can be obtained along with a low dark current and a good carrier collection yield in short wavelength region due to well passivation between the metal contacts.

.

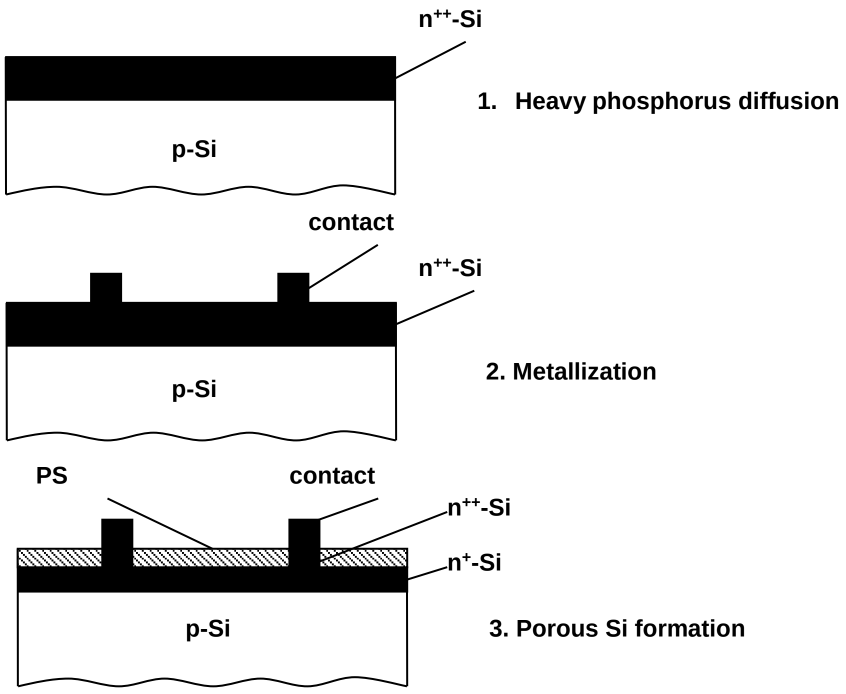


Fig. 5.23. Process of etchback selective emitter formed by porous silicon

The selective emitters are formed in many ways (Bilyalov et al. 2000; Moon et al. 2009). Basically, there are two types: technologies that need scheme aligning (more expensive and more difficult) and self-aligning technologies. The selective emitter based on self- aligning

PSi technique uses that after metallization of heavy doped n^{++}-Si emitter, the emitter is etch back forming porous layer (Fig. 5.23). Since, it was assumed, that PSi will act as ARC and passivating layer, no further layers need (Stalman et al. 1997; Kuthi 2004). This is a strongly simplified processing scheme, aimed to replace the conventional texturing, thermal oxidation and the ARC. In first works, the optimistic expectations of PSi selective emitter were announced for multi and polycrystalline Si materials, because, conventional alkaline texturing only functions well for monocrystalline Si. Efficiencies between 13 and 14 % have been reached on multicrystalline substrates (Strehlke et al. 1997; (Bilyalov et al. 2000) in earlier works

The disadvantages of porous Si selective emitter are the same as for the application of porous Si as ARC. For the case of screen-printed contacts, the underetching of the metallization and of the underlying Si surface takes place due to the presence of HF in the electrochemical etching solution, which leads to the increase of the contact resistivity and to the decrease of FF. For the case of sputtered metal contacts the etching does not strongly degrade the quality of the contacts, however this technique is more expensive (Kuthi 2004).

Now, the effective selective emitter is proposed to form when PSi creation is used as interim technological step at the standard screen printing process. Process sequence uses a heavy diffusion which is masked in the area that will be contacted, the emitter is then etch back in an acidic solution to the desired sheet resistance. Then the porous Si and masking layer are subsequently removed in an alkaline solution. The following process of cell fabrication is the standard: PECVD deposition of SiN_x, screen metallization, cofiring and edge isolation. Cz-Si solar cells of 5`` prepared by this technology shows V_{oc}=640 mV, J_{sc}=37.6 mA/cm^2, fill factor FF=78.4 % and photoconversion efficiency of 18.9%. It was shown that for an etch back emitter as created in this process, the saturation dark current could be reduced to 28 fA/cm^2 for a 68 Ω/□ emitter by etching back from a very low sheet resistance of 17 Ω/□ (Book et al. 2009).

5.9 SURFACE PASSIVATION BY POROUS LAYERS

5.9.1 Mechanism of recombination losses.

The efficiency of a solar cell is also reduced by the recombination of electron-hole pairs before they can be usefully collected (Green, 1986). A number of recombination sites are shown in Fig. 5.24

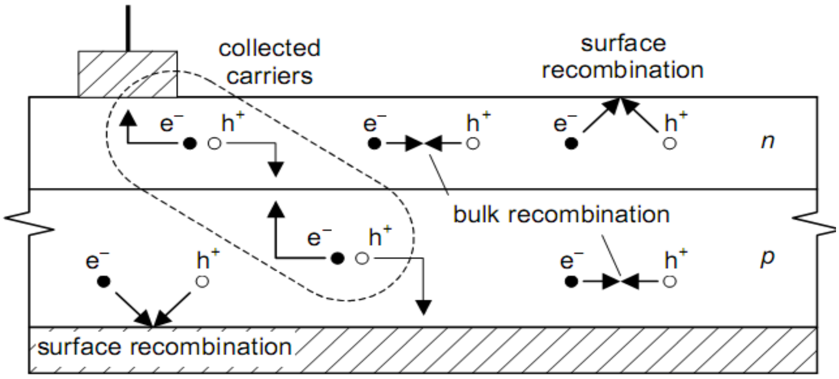


Fig.5.24 Surface and bulk recombinations

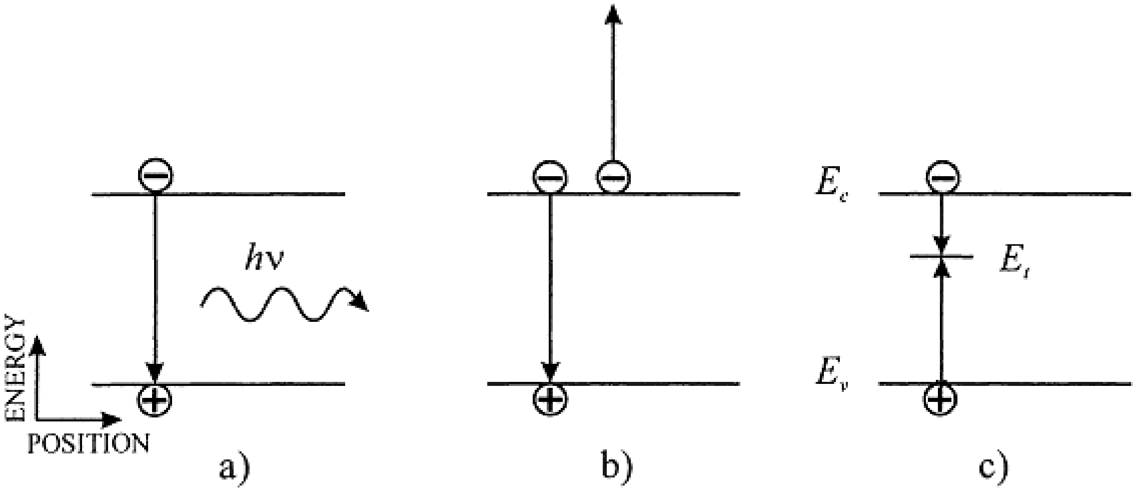


Fig.5.25 Mechanisms of recombination in Si solar cell: a) radiative band-to-band recombination, b) Auger recombination, c) Schockley-Read –Hall recombination via deep traps

Recombination can occur via several mechanisms (fig.5.25):

1. **Radiative recombination**—the reverse of absorption. Electrons in a high energy state return to a lower energy state, with the release of light energy. This form of recombination is used for semiconductor lasers and light emitting diodes, but is not particularly significant for silicon solar cells.

2. **Auger recombination**—the reverse of ‘impact ionisation’ . An electron recombining with a hole gives up the excess energy to another electron, which then relaxes back to its original energy state, releasing phonons. Auger recombination is particularly effective in relatively highlydoped material, becoming the dominant recombination process when impurity levels exceed 10^{17} cm^{-3}.

3. **Recombination through traps**—This can occur when impurities in the semiconductor or interface traps at the surfaces give rise to allowed energy levels in the otherwise forbidden

energy gap. Electrons can thus recombine with holes in a two-stage process, first relaxing to the defect energy level, then to the valence band.

5.9.2 Estimation of surface recombination velocity in PSI / Si interface

Large surface recombination velocity strongly reduces the internal quantum efficiency IQE in the short wavelength region of the cell. For conventional Si solar cell of n^+/p type the front surface recombination is important for the spectral region of 400-600 nm while the back surface recombination is revealed in IR part of the spectrum in the case if the diffusion length of minority charge carriers in the base is bigger than its thickness. The search of *passivating* materials and processes that able to suppress the negative impact of the surface recombination has been one of the important tasks of PV. Application of PSi for the front surface passivation was seemed to be very attractive in the technology in which these layers are used as ARC. The first published results on PSi passivation properties were rather contradictory. In some papers PSi has revealed excellent passivation properties (Zheng et al. 1992) while in the other ones the surface recombination velocity on PSi/Si interface was higher than on SiO_2/Si interface or on free Si surface (Smestad et al. 1992).

The surface passivating capabilities of a porous layer on a top of n^+-emitter can be extracted from the IQE measurement in the short wavelength region. However, the PSi on the top of the cell absorbs a part of the incident light without photogeneration in the cell (the charge carriers are trapped on dangling bonds on developed surface). Hence, the IQE should be corrected according to

$$IQE^* = IQE \exp(\alpha d),$$

where α and *d* are the absorption coefficient and the thickness of PSi, respectively. For example, the absorption from 160 nm porous layer results in decreasing of the short circuit current I_{sc} by more than 2%. The estimated value of surface recombination velocity of Si wafer with porous top surface was found to be $S_p=3.10^5$ cm/s in comparison with that one for unpassivated mirror-polished surface $S_p=3.10^6$ cm/s (Fig. 10.11a) (Stalmans et al. 1997, 1998).

The surface recombination velocity can be obtained and from analysis of dark I-V characteristics. According two diode model the dark I-V characteristics of SC:

$$I = I_{s1}(e^{\frac{qV-IR_S}{n_1kT}} - 1) + I_{s2}(e^{\frac{qV-IR_S}{n_2kT}} - 1) + \frac{V - IR_S}{R_{sh}}$$

The first term corresponds to the recombination current and thus the ideality factor $n_1 = 2$, the second one is the diffusion current with $n_2 = 1$, the third one is the current through the parallel shunt R_{sh}. If the emitter thickness is $d_1 < L_p$ (diffusion length of holes) then the emitter saturation current depends on the surface recombination velocity at the front side of the cell S_p. In this case, according to [Sze 1981] the saturation current density for the diffusion current could be written as a function of S_p (if recombination at rear side of SC in neglected small). The good fitting of experimental I-V characteristics of silicon SC with PS coating was obtained at the S_p =0.8-3.10^5 cm/s [Skryshevsky 2000]. These values are worse comparing with PECVD silicon nitride and silicon dioxide passivation [Nagel 1999]

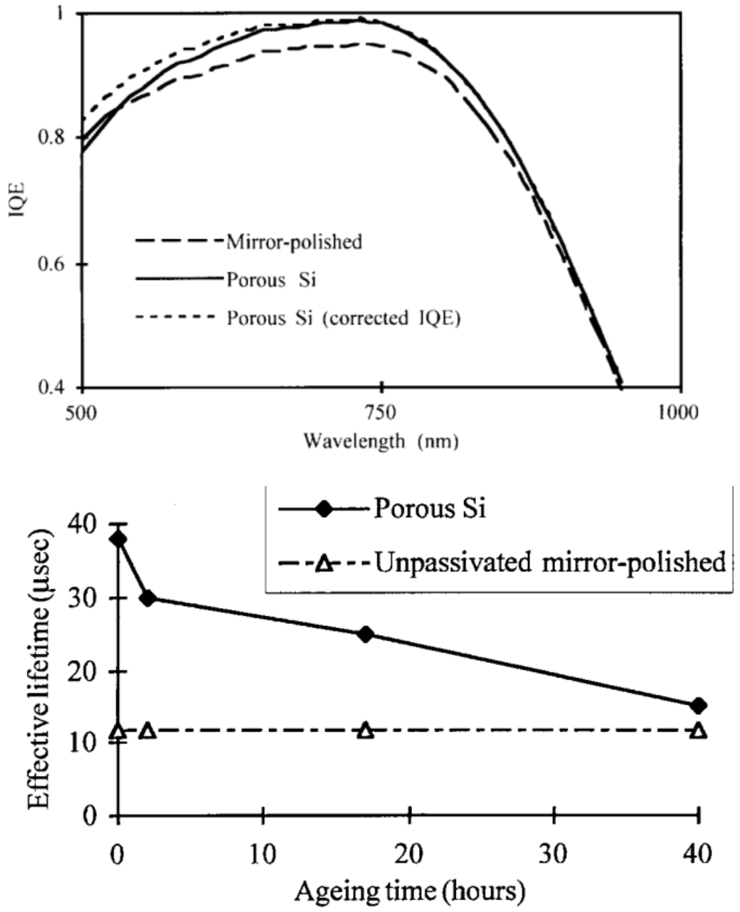


Fig.5.26a) Effect of a porous surface on the IQE in c-Si thin-film cells, b) Degradation in ambient with time of τ_{eff} of p-type Si wafer (unpassivated and passivated with PSi). (Stalmans L. et al., *Prog. Photovolt. Res. Appl.* **6,** 233, 1998)

The other drawback of PSi is its fast degradation that leads to the increase of the surface recombination velocity. The photoconductivity decay measurements showed the degradation of the effective lifetime τ_{eff} during ageing of p-type Si wafer with porous top surface in ambient (Fig. 10.11b) (Stalmans et al. 1998). In the case when the rear side recombination can be neglected (rear side of p-type Si wafer is well passivated) the effective lifetime τ_{eff} is determined by the bulk lifetime τ_b and front recombination velocity

$$\frac{1}{\tau_{eff}} = \frac{1}{\tau_b} + \frac{S_p}{W},$$

where W is the substrate thickness. Since no bulk changes occurred during ageing, the observed decrease of τ_{eff} (from 40 µsec to 10 µsec for 40 h) was attributed to changes at the porous surface (Stalmans et al. 1998). To stabilized surface passivation capabilities of porous Si a several oxidation and hydrogenation treatments have been proposed: rapid thermal oxidation (RTO), electrochemical oxidation, plasma-nitride treatment and electrochemical hydrogenation (Stalmans et al. 1997; Yerokhov et al.1999). Note, the improvement of blue response after the rapid thermal oxidation of porous surface is observed. Furthermore, after the thermal treatment of the cell with PSi top layer, its IQE in the near IR region still remains higher than that one of reference cell without PSi coating (Stalmans et al. 1997). Haddadi et al. (2011) showed the reduction of the surface recombination velocity and an improvement of the diffusion length after the immersion PSi in a lithium bromide (LiBr) aqueous solution. Ben Rabha et al (2013) studied the effect of Al_2O_3/PSi combined treatment on the surface passivation of monocrystalline silicon c-Si and the effective minority carrier lifetime increasing from 2 µs to 7µs is observed.

5.9.3 Light - induced passivation on PSI/Si interface

The *light -induced passivation* of the p-Si surface by a thin ($d \approx 0.1$ µm) porous layer, formed by electrochemical etching is discussed by [Nichiporuk 2006] for interdigitated back contacts (IBC) solar cells. It is known, that IBC cells are very sensitive to front surface recombination velocity [Fossum 1980, Nichiporuk 2005].

The illuminated *I–V* characteristics of the best cell with different front surface coatings are presented in figure.5.27 The best passivation was obtained with UV–CVD silicon nitride passivation, and the worst with chemically formed SiO_2 because of the poor quality of this oxide. The short-circuit current J_{sc} of the solar cell with porous silicon passivation is better than two-times the of an unpassivated silicon surface. Effectively, the maximum J_{sc} gain with optical confinement (porous silicon serves as ARC and light diffuser) for the selected cell configuration is about 40% (calculated with PC1D numerical simulation [Basore 2000). Therefore, the significant gain in J_{sc} after porous silicon elaboration can be explained by the decreasing of the front surface recombination

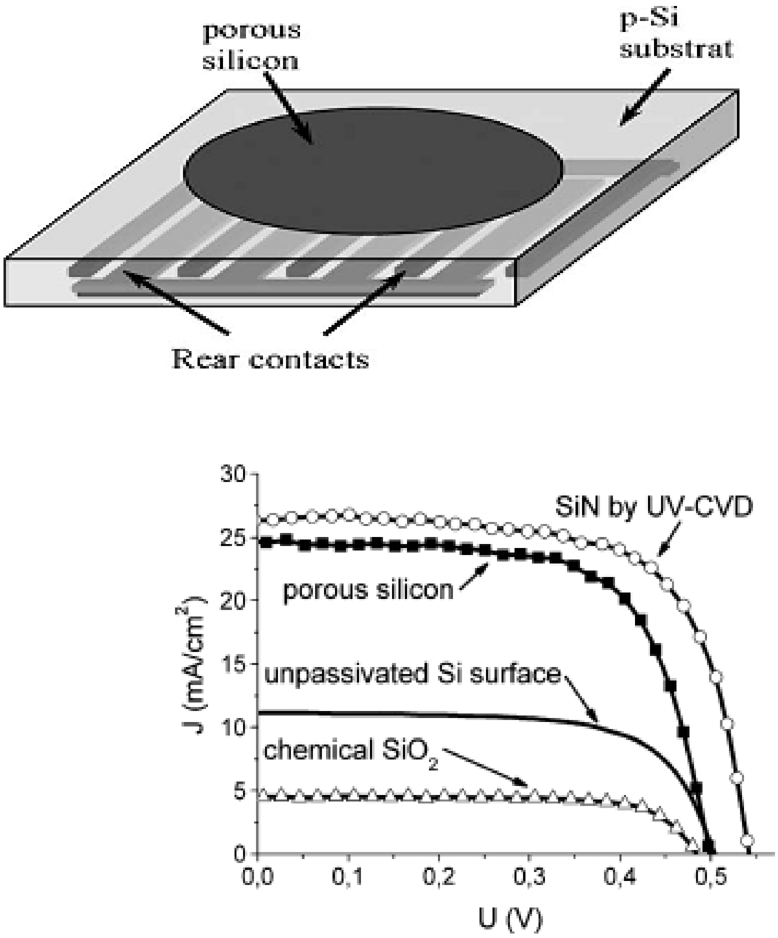


Figure. 5.27. Design of the IBC SC with PS layer formed at the front surface (a) *I–V* characteristics under illumination at AM1.5G spectrum conditions of the same IBC cell with different front surface coatings (b).

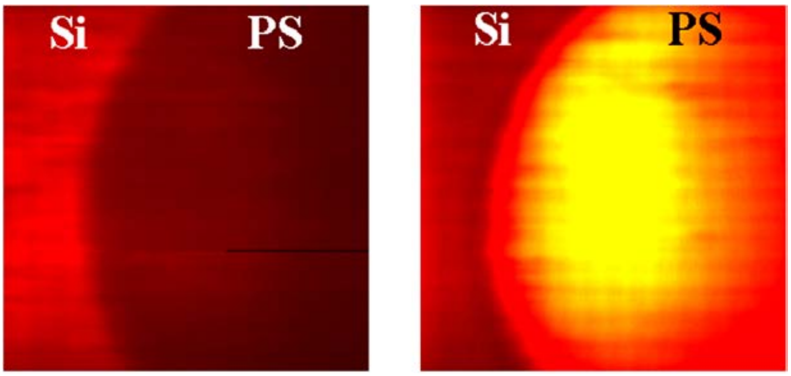


Figure. 5.28 LBIC images of the same IBC solar cell with porous silicon layer at the front surface for two different measurement conditions: a) without additional illumination; b) in presence of permanent blue light.

.

However, the Laser Beam Induced Current (LBIC) measurement of spatial J_{sc} distribution at scanning front surface of the IBC cell with a modulated red laser beam shows that photocurrent is much lower in the porous Si area (right part of the LBIC image in Figure 5.28a) in comparison with an unpassivated silicon surface (left part of the LBIC image in figure 5.28a). When the cell is scanned by red laser beam in presence of a permanent (non-modulated) pumping light with $\lambda < 470$ nm (P ~0.5 mW/cm^2), a strong increase of the photocurrent in the porous Si area is observed (Figure. 5.28b). This J_{sc} growth can be provoked by the decrease of the front surface recombination at the porous Si covered area in presence of pumping light. The light-dependent porous silicon passivation phenomenon can be explained by a significant negative charge accumulation at the porous Si/p-Si interface traps under illumination and formation of an induced hi-low (p^+/p) junction at the front surface of the cell [Mizsei 2004].

5.10 PSi AS OPTICAL REFLECTOR

5.10.1 Application of rear Bragg reflector for Si solar cells

Though Si based solar cells dominate in PV industry, some their drawbacks still are not eliminated. Thus, Si displays a relatively low absorption coefficient in near-infrared spectral region. Si based solar cells must be thick (up to millimeters) to absorb 90% of the incident light at wavelengths from 700 to 1100 nm, respectively, which comprise about one-half of the solar energy available above the band gap of Si (Green and Keevers 1995). However, the cell thickness increasing is not desirable since it leads to growth of bulk recombination losses and Si consumable. Thereby, it is desirable to construct cells with improved scheme of light trapping (e.g. with the elongation of the light path inside cell) without increasing of wafer thickness. As was shown in above the forming of thin PSi diffuser on top of the cell does not solve this problem.

One of the methods to enhance long wavelength light absorption without thickness increasing is the using of a reflector at the backside of the solar cell. Sun light is partially reflected backwards into the cell that results in its additional absorption leading to the supplementary electron/hole generation. The rear surface reflector in n^+/p solar cell should provide both the multi-passing of low adsorbed photons inside a cell and a good passivation. For the reason of strong recombination at metal/Si interface the application of continuous metallic layers evaporated directly on p/p^+ base of Si solar cell is obstacle. Rear side reflector can be formed as multilayers dielectrical films on Si substrate (SiN, SiO_2) with windows for back Ohmic contacts (Fig.5.29).

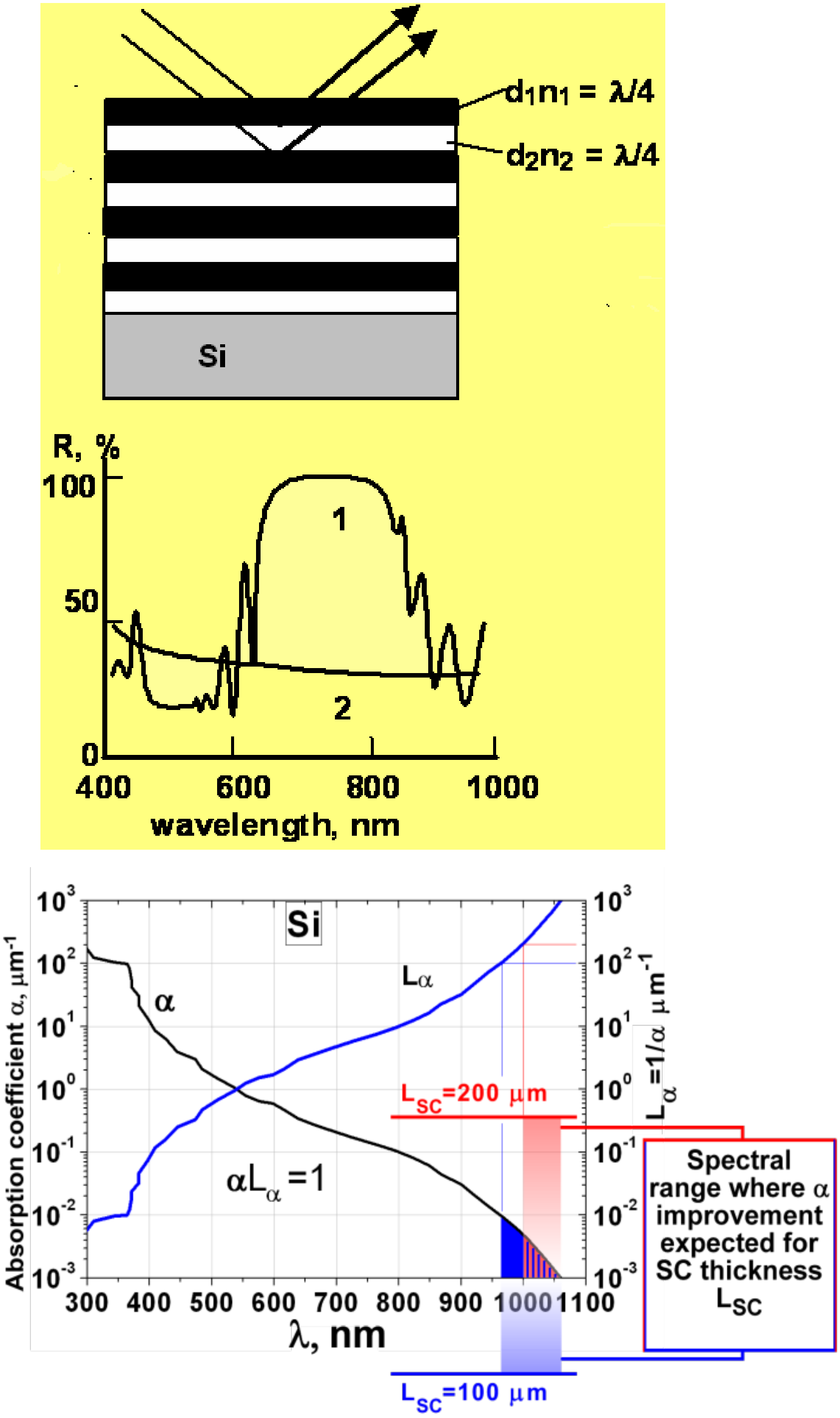


Figure 5.29 Design of simple of Bragg mirror on semiconductor substrate (a) and dependence of reflection versus wavelength (b) and (c) absorption coefficients and depth of light absorption

A simple back side reflector based on PSi is a *Bragg mirror*, which can be easily realized by electrochemical etching. Bragg mirror consist from alternating sequence of quarter wavelength layers of two porosities, correspondingly with two refractive indexes n_H (high) and n_L (low) (Fig.5.29,5.30).

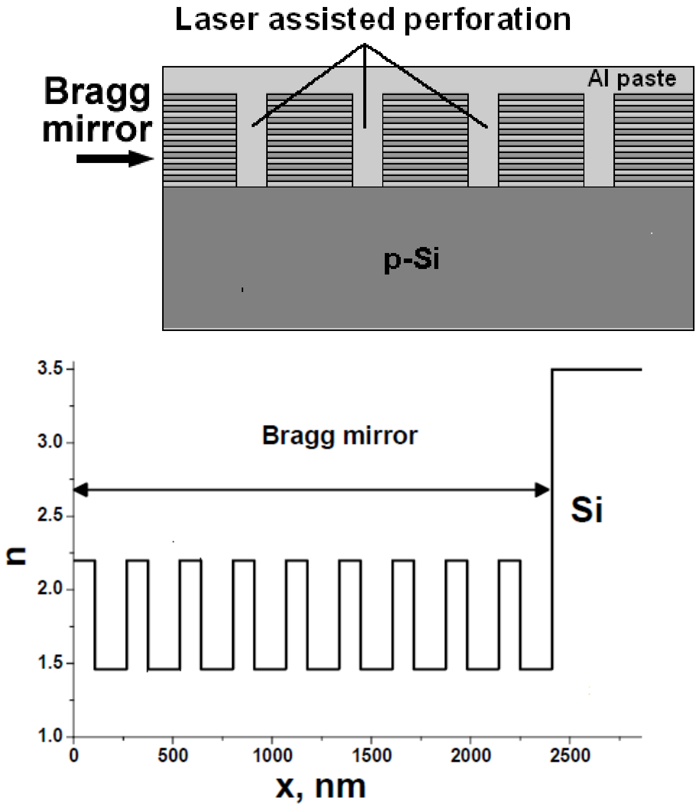


Fig.5.30 a) Scheme of rear side of SC with Bragg mirror and laser assisted perforation in mirror, b) prifile of refractive index at rear side of SC with Bragg mirror

If individual porous layers have the correct refractive index and thickness, constructive interference can lead to a high internal reflection for perpendicular striking photons. The wavelength λ_0 in the maximum reflection is obeyed to equation:

$$m\lambda_0 = 2(d_1 n_H + d_2 n_L)$$

The important characteristics of the Bragg mirror are the value of reflection coefficient R_{max} at λ_0 and FWHM of peak. Fig. 5.31 presents the calculated reflection coefficient in maximum and FWHM versus n_H/n_L ratio for different number of bi-layers. The increasing numbers of layer pairs results in higher refection coefficient R_{max} for lower ratio value n_H/n_L, however, the value of refection coefficient $R_{max}(\lambda_0) \geq 99\%$ can be received for bi-layer number $N_{bi}=4$ when $n_H/n_L > 1.8$. Bi-layer number increasing from $N_{bi}=4$ to $N_{bi}=14$ leads to FWHM increasing. For ratio $n_H/n_L=2$, FWHM increases by 31% when bi-layer number changes from $N_{bi}=4$ to $N_{bi}=6$. The following bi-layer number increasing to $N_{bi}=14$ doesn't lead to considerable FWHM increasing [Ivanov 2009].

Simulation results show that optical light path increasing in long-wavelength spectral region caused by internal multiple bouncing inside of standard cell of 200 µm thick leads to increasing of I_{sc} on 1-2 mA/cm^2. Solar cell conversion efficiency increases up on 1% in absolute when internal reflection coefficients growth from 20% up to 95% (fig.5.32) (Ivanov et al. 2009,2013).

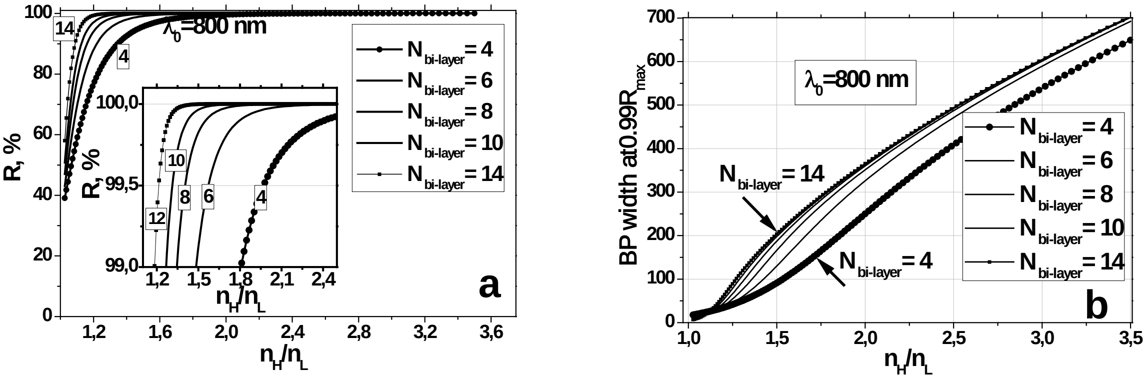


Figure. 5.31Dependence of R_{max}($_0$)□reflection coefficient (a) and FWHM (b) on n_H/n_L ratio for different bi-layer number (digits show quantity of bi-layers).

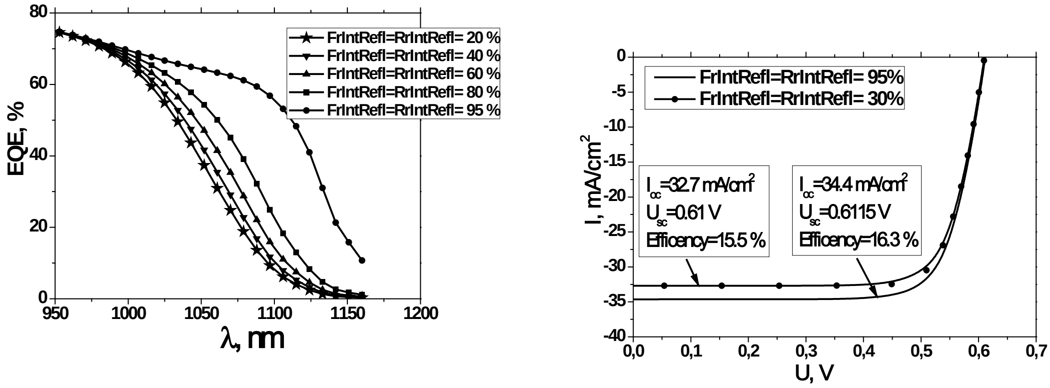


Fig.5.32.a) EQE versus wavelength for different front and rear internal reflection coefficients ($R_{FrIntRefl}$, $R_{RrIntRefl}$) (surface of Si wafer is textured: angle at pyramid base = 54.74°, pyramid height = 5 □m); b) Light I-V dependence for different values of internal reflection coefficients inside of Si wafer.

5.10.2 Porous Si rear reflectors for thin film solar cell

The application of back side reflector has more profit for thin solar cells. Indeed, compared to bulk Si solar cells, in thin Si devices similar levels of V_{OC} and FF can be obtained. The I_{SC}, however, is held back by the optically thin active layer. Light that traverses the epitaxial layer is lost for collection in the highly doped low quality substrate. As a consequence a typical value of I_{sc} in epitaxial thin film solar cells is 26 mA/cm^2. In order to improve the optical pathlength for long wavelength light the intermediate reflector between the epilayer and the substrate can be applied The concept of an intermediate reflector in thin film epitaxial cell technology is illustrated in Fig. 5.33 (Duerinckx et al. 2006; Kuzma-Filipek et al. 2008). The photons reaching epi/substrate interface can now be reflected and pass a second time through the active layer.

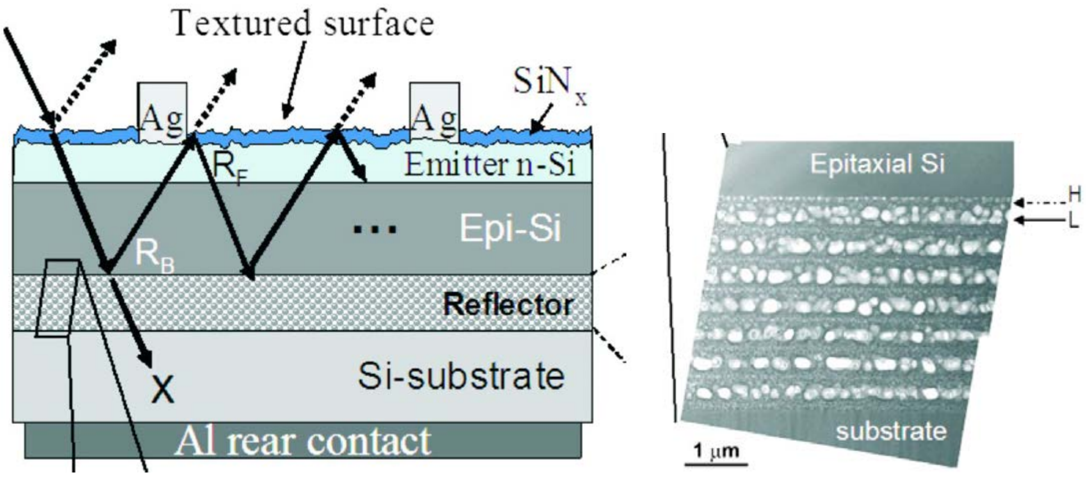


Fig.5.33 a) Schematic cross section of a thin-film epitaxial cell including a Bragg porous mirror, b) TEM image of a reorganized porous Bragg reflector Kuzma-Filipek I. J. Et al., *J. Appl. Phys*. **104,** art. 073529, 2008.

As well as for thick Si solar cells the Bragg reflector can be made by electrochemical growth of a porous stack of alternating high and low porosity layers. Note that the rest solar cell fabrication process remains very similar to that of standard multicrystalline Si process. The main difference is the implementation of the electrochemical etching process for PSi formation after saw damage removal and the deposition of the active layer by CVD. Numerical modelling has demonstrated that PSi multilayer reflector is particularly useful for extremely thin Si cell on textured substrate (Zettner et al. 1998).

The use of PSi single layer and multilayer structures as an optical reflector for thin film epitaxial growth on Si substrate has been reported in numerous research reports (Zettner et al.1998; Kuzma-Filipek et al. 2007, 2008, 2009; Van Hoeymissen et al. 2011). It has also been developed a new method for formation of porous Si with a multilayer structure in one step which makes much easy the up scaling of porous Si into an industrial level (Matic et al. 2000).

There are three ways to use PSi back reflector in thin film devices (Bilyalov et al. 2001). The first one is the high temperature road (>800°C) resulting in quasi-monocrystalline Si and good quality epitaxial layers. During the epitaxial growth on top of the porous stack, the individual layers reorganize into quasi-monocrystalline Si. Using this method efficiency more than 10% have been reached in a 10 µm thick epitaxial cell with PSi reflector on a highly doped multicrystalline Si substrate (Bilyalov et al. 2001). A medium temperature road (700-800°) allows preserving the PSi structure (Jin et al. 2000), but results in pore filling effect during the epitaxy. A low-temperature road (<700°C) would be an attractive solution if deposition on PSi layer could be done with an epitaxial quality and sufficient growth rate.

Reflectance values up to 85% have been obtained by using PSi rear surface reflector. Together with the total internal reflection on the front surface this leads to an additional optical path length enhancement of 7 (Van Hoeymissen et al. 2008). On cell level, the PSi reflector has given considerable increase in I_{sc} (Duerinckx et al. 2006) resulting in efficiencies of 13.5% on low-cost Si substrate.

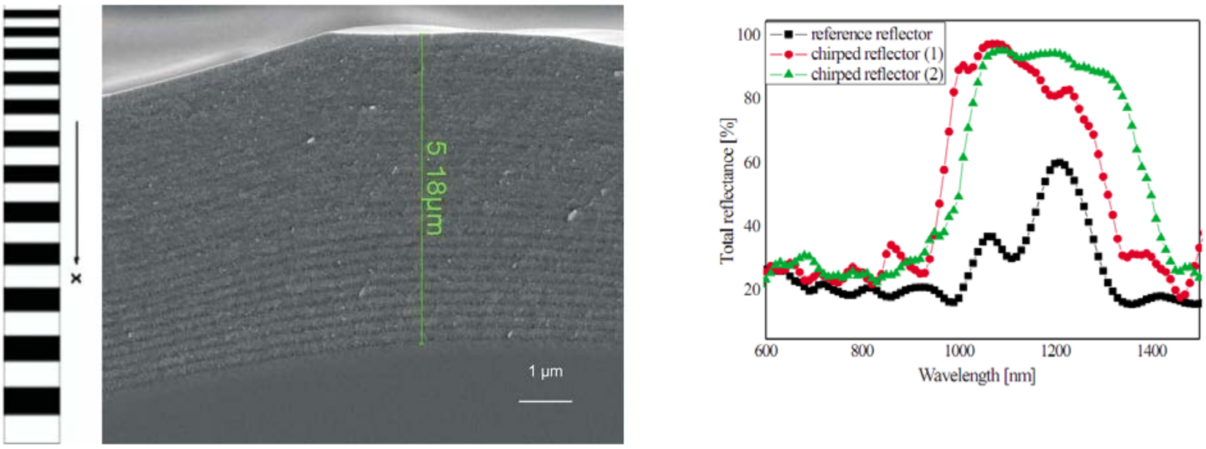


Fig.5.34 a) Linear chirped structure of porous silicon stack as grown. The thickness of porous silicon layers increases with the etching depth of the silicon substrate. b) Experimental reflectance spectrum of various fabricated reflectors as grown porous Si, single mirror Bragg stack of 15 layers and unconventional chirped reflector of 80 layers (substrate is monocrystalline Si). Kuzma-Filipek I. J. et al., *J. Appl. Phys.* **104**, art. 073529, 2008.

Recently a multilayer PSi Bragg reflector was improved by chirping the structure, i.e. varying the periodicity in depth of the alternating PSi sublayers (Van Hoeymissen et al.2008; Kuzma-Filipek et al.2008) (Fig. 5.34a). The experimentally measured reflectance curves of samples with Bragg reflectors, with and without chirping the structure is presented in Fig. 5.34b. The wavelength band with high reflectance was broadened by 50-80%. The cells with chirped reflector containing 60 or 80 sub layers. Efficiencies of large area epitaxial cells on low-cost Si substrates including conventional porous reflectors have reached values of around 13%. With the implementation of alternative designs of porous silicon reflectors such as chirped structures, those efficiencies have increased to almost 14% with the standard screen-printing technology.

5.11 APPLICATION OF POROUS SILICON FOR LAYER TRANSFER

Porous silicon can be applied for layer transfer processes to get a thin monocrystalline silicon film on a foreign substrate for solar cell application. The top surface of the monocrystalline silicon wafer is made porous by electrochemical etching. The porous Si layer has a double porosity structure: low porosity layer at the top and high-porosity one at the bottom. The remaining material in low-porosity layer is of monocrystalline quality and allows the growth of a high-quality epitaxial silicon layer on it after annealing the porous structure at high temperature. The epitaxial layer deposited on porous silicon layer is detachable from the original silicon substrate through a high-porosity layer. The deposited epitaxial layer is transferred onto a foreign substrate. Device fabrication can be obtained before or after the layer transfer onto a foreign substrate and needs different processing sequence and special requirements for the substrate. Thefore, the porous structure serves two purposes. Firstly, it allows the growth of a high quality epitaxial layer and, secondly, it allows separation of the epitaxial layer from the starting substrate and transfer onto a foreign substrate (Brendel 2004; Solanki et al. 2004).

The active layer deposition is generally preceded by the high-temperature treatment of the double porosity structure. The high temperature treatment (>1050°C) of the PSi layer in a H_2 atmosphere causes the reorganization of the porous structure- transforms the 10–50 nm pores to spherical voids ranging from 50 nm to 1 µm depending on the morphology and porosity of the initial PSi layer and top surface of the low-porosity layer closes and becomes smooth. This smooth surface provides a perfect seeding layer for the epitaxial layer deposition. Besides, the high-temperature annealing of a high-porosity layer results in the formation of big voids with remaining thin silicon pillars connecting the initial silicon substrate and the low-porosity layer. This layer with silicon pillars serves as the separation layer (Rinke et al. 1999; Solanki et al. 2004).

There are many processes developed or being developed which use PSi as a sacrificial layer for the layer transfer process. Among these processes are ELTRAN (Yonehara et al. 1994), ψ- process (Brendel 1997), quasi-monocrystalline silicon (QMS) process (Rinke et al. 1999), LAST process (Solanki et al. 2002), Freestanding Monocrystalline thin film Silicon (FMS) process (Solanki et al. 2004) and others. The mentioned layer transfer processes mostly differ in the sequence of the steps that are used to complete the cell fabrication and transfer onto a foreign substrate.

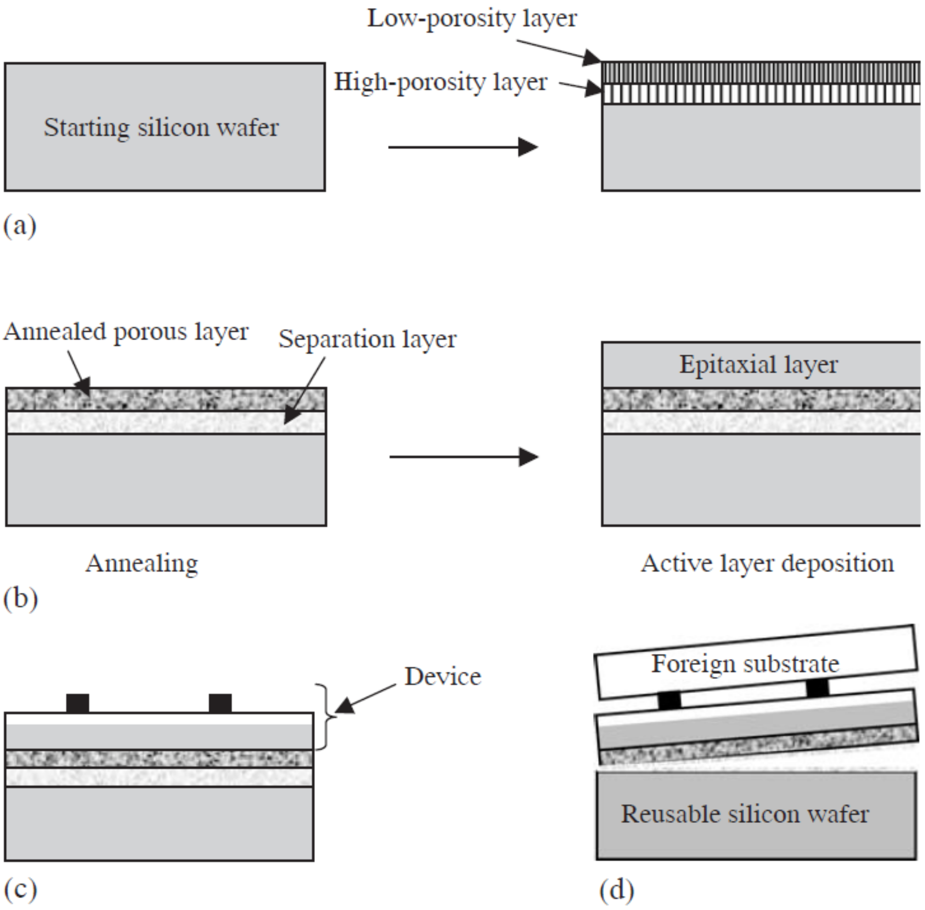


Fig.5.35 Steps of porous layer transfer processes using porous silicon as a sacrificial layer for obtaining thin monocrystalline silicon films on cost-effective substrates: (a) a double porosity structure (high-porosity layer beneath low-porosity layer) formation on starting silicon substrate by anodization, (b) thermal annealing of porous silicon and active layer deposition: annealed low-porosity layer acts as a good seeding layer for epitaxial layer deposition and voids with weak silicon pillars forms in high-porosity layer acts as a separation layer. (c) Device fabrication. (d) Separation of epitaxial layer from the starting silicon substrate and transfer onto foreign substrate by gluing it with an adhesive layer and applying mechanical force.(Solanki C. S. et al., *Solar Energy Mater. Solar Cells*, **83**, 101, 2004)

The QMS process (Rinke et al. 1999; Werner et al. 2003) consists in a double porous layer that restructures after annealing at high temperature under hydrogen. The upper low porosity layer becomes quasi monocrystalline and can be used as the active layer of the solar cell and the high porosity layer permits the detachment. The ψ- process uses the transfer a texturised monocrystalline layer using a texturised silicon wafer for porous layers formating. Efficiencies of 13.3% (Brendel 2004) have been obtained using this process and 15.4% (Feldrapp et al. 2003) have been reached by texturising the wafer after the epitaxy. The main drawbacks of this process are the high consumption of the silicon wafer (12 µm per cycle (Horbelt et al. 2005).

In general, layer transfer processes for silicon solar cells are promising candidates to reduce cell costs because kerf losses are greatly reduced since one substrate wafer can be reused many times. Recently, Petermann et al. (2012a; 2012b) demonstrated the high efficiency potential of this material by reporting a new independently confirmed record efficiency value of 19.1% for this type of solar cell.

5.12. MODERN TENDENCIES IN SI PV

5.12.1 Solar cells with Si nanowires

Advances in semiconductor materials and technology, naturally, have appreciably influence on the modern R&D in PV field. In the last years a new solar cell structures with Si nanowires, quantum dots, quantum wells attract a special attention as promising way to improve the solar cells efficiency and reduce the Si consumable.

The collection efficiency of charge carriers generated in cell outside of p-n junction depends on the diffusion length of minority charge carriers in the quasi-neutral regions which is limited by recombination losses. To minimize charge carrier recombination, cells must be created on material with a large diffusion length of minority charge carriers, which is very expensive for routine Si cell fabrication. A proposed solution to this problem relies on decoupling the long extinction distance of Si and the proximity of generated charge carriers to the p-n junction in cells based on Si *nanowire* (NW) carpets (Sivakov et al. 2009; Hochbaum and Yang 2010). By comparing classical thin film cells based on multicrystalline Si with NW solar cells, the last ones are expected to have higher efficiency due to: *(i)* the possibility to produce the perfect single crystalline structure of wires, *(ii)* the short distances needed for charge separation and *(iii)* the perfect light trapping in NWs arrays.

There exist two methods for production Si NW, a bottom up and a top down approach. The bottom up way relies on the Vapor Liquid Solid (VLS) method which uses metal nanotemplates on Si wafer or on Si thin film (Tian et al. 2007; Tsakalakos et al. 2007; Stelzner et al. 2008; Kelzenberg et al. 2008). Most frequently gold is taken as a template which forms an eutectic with Si at 370^{o}C. The Au nanotemplates are prepared from gold colloids or by deposition of a thin Au film followed by a heating step above the eutectic temperature during which the nanodroplets are formed. The Au nanotemplates act as catalysts to decompose silane at T > 500°C. Si NWs grow with a diameter similar to that of the template droplet. As a result a carpet of perfect single crystalline NWs of 10 to 200 nm in diameter and several micrometers in length (with rate of several 100 nm/min) can be grown on the crystalline substrate (Andra et al. 2008).

In top down approach well aligned Si NW arrays are obtained by electroless metal-assisted chemical etching in HF/$AgNO_3$ solution (Srivastava et al. 2010). Basically, a noble metal is deposited on the surface in the form of nanoparticles which act as catalyst for Si etching in HF solution containing an oxidizing agent. As a consequence, the etching only occurs in the vicinity of the metal nanoparticles and results in the formation of well defined mesopores (20-100 nm in diameter (Peng et al. 2005; Fang et al. 2008). Tapering the NWs by post-KOH dipping achieved separation of each NW from the bunched NW (Fig. 5.36), resulting in a strong enhancement of broadband optical absorption. As electroless etching time increases, the optical crossover feature was observed in the tradeoff between enhanced light trapping (by graded-refractive index during initial tapering) and deteriorated reflectance (Jung et al. 2010).

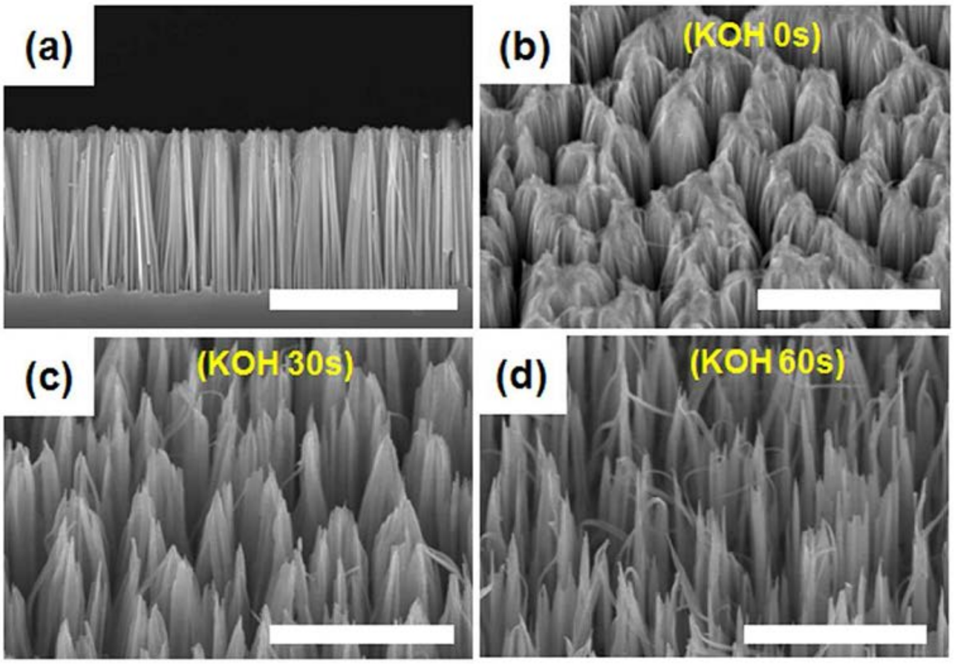


Fig.5.36 Cross-sectional SEM images showing the SiNW array after metal-assisted electroless etching (a). Scale bar is 10 μm. 30°-tilted SEM images also show the morphological change of the SiNW tips according to the post-etching time of KOH: (b) 0 s, (c) 30 s, (d) 60 s .Scale bars in (b)−(d) are 5 μm. Jung J-Y. et al., *Optics Express*, **18**, A286, 2010

Independently of the nanowire preparation method two designs of NW solar cells are now under consideration with p-n junction either radial or axial. In the radial case the p-n junction covers the whole outer cylindrical surface of the NWs (Fig.5.37, 5.38) (Kayers et al. 2005; (Hochbaum and Yang 2010). This was achieved either by gas doping or by CVD deposition of a shell oppositely doped to the wire (Peng et al. 2005; Tian et al. 2007; Fang et al. 2008).

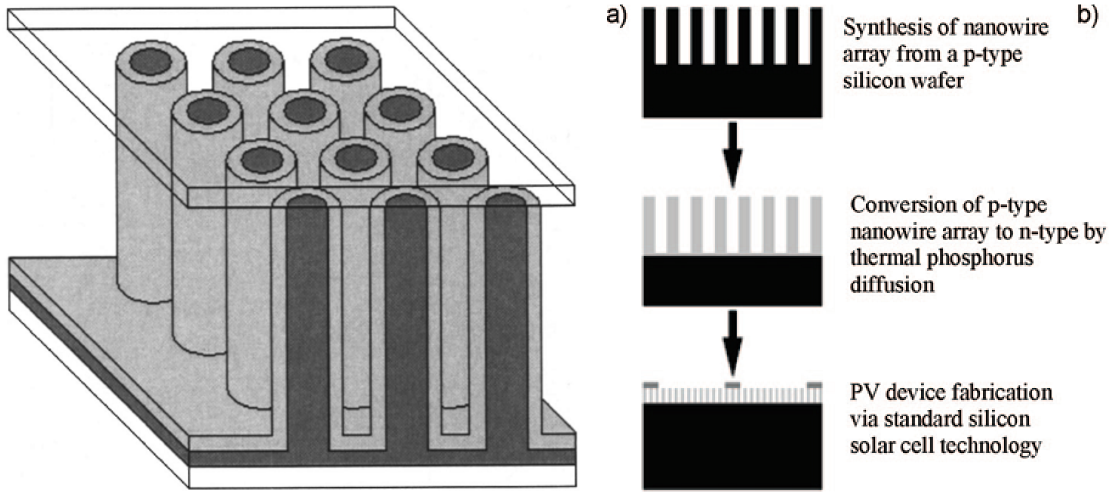


Fig.5.37 a) Schematic cross-section of the radial *p-n* junction nanorod cell. Light is incident on the top surface. The light grey area is *n* type, the dark grey area *p* type. (Kayes B.M. et al., *J. Appl. Phys*. **97**, art. 114302, 2005.)
b) Schematic of a subsurface p-n junction device fabrication process (Hochbaum A.I. and Yang P., *Chem. Rev*.,**110**, 527, 2010.)

In the axial variant, the p-n junction cuts the NW in two cylindrical parts and require minimal processing steps (Andra et al. 2008). However, cells that absorb photons and collect charges along orthogonal directions meet the optimal relation between the absorption values and minority charge carrier diffusion lengths. Fig.5.37b shows the variant of axial p-n junction fabrication process on p-substrate when n-region is formed on nanowire array. Since charge extraction occurs through the nanowires, decreasing I_{SC} is observed due to the larger series resistance of the nanowires and their contacts (Hochbaum and Yang 2010). In order to assure high conversion efficiencies, the NW solar cell parameters should be carefully optimized. The vertical array geometry scatters light efficiently, especially at short wavelengths, and can absorb more light than a comparably thick solid crystalline film. Optical absorption in cell is influenced by NW diameter and their density (NW diameter governs the energy band structure due to the quantum confinement effect). Series resistance depends on NW diameter, length and density. The interface between the nanowires and the encapsulating material impacts photogenerated carriers’ recombination at the nanowires surface.

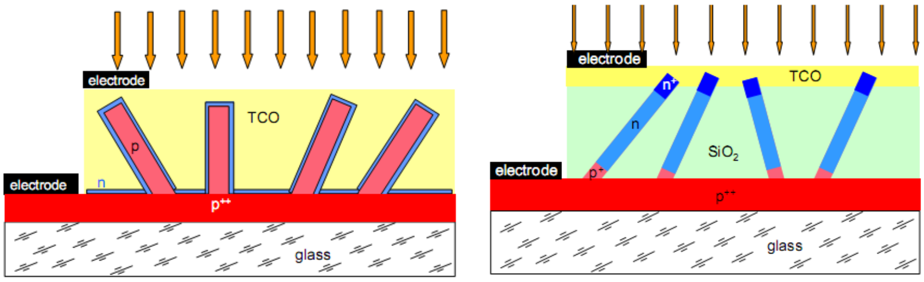


Figure 5.38 Scheme of NW solar cells with radial (a) and axial p-n junction [Hochbaum 2010][Andra 2008].

In Fig. 5.39 the optical absorption (A) of Si NWs on glass substrate and on mc-Si on glass, derived from transmission and reflection data (A=1-T-R), is compared to the absorption of a 375 µm thick Si wafer. A strong broadband optical absorption is observed in the relatively thin Si NW (3-6 µm) due to strong light trapping of the NW. Pay a special attention the strong absorption of light with an energy below the bandgap energy due to light trapping together with absorption by defect states and plasmon coupling of light with the NW and underlying nanocrystalline Au-Si (Stelzner et al. 2008).

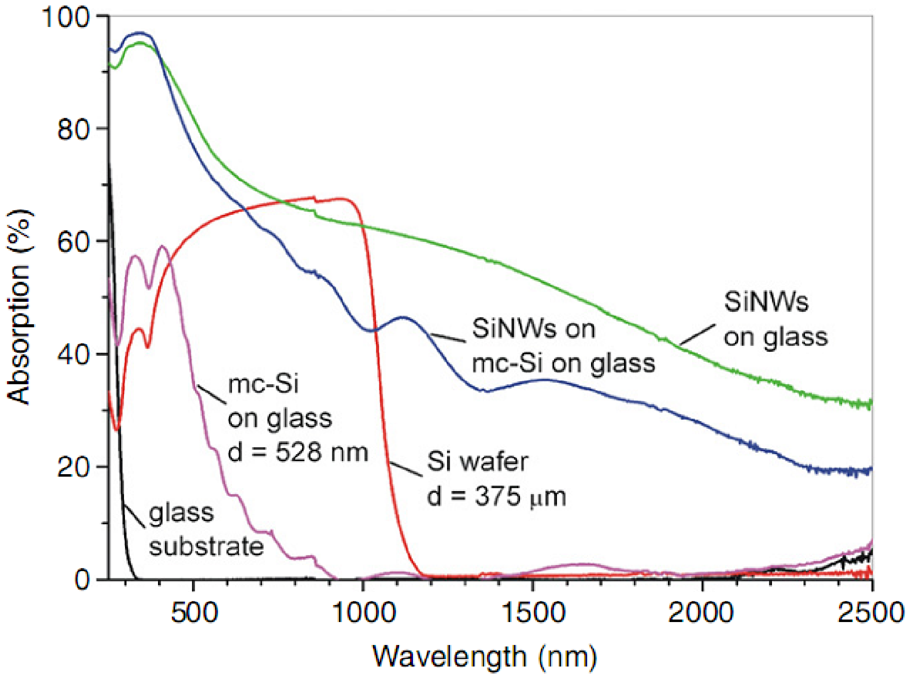


Fig.5.39 . Optical absorption of SiNWs on a glass substrate and on mc-Si on glass compared to the absorption of a 375 μm thick Si wafer. (Stelzner T. et al., *Nanotechnol.*,**19**, 295203, 2008.)

It should be also noted that from economical point of view in VLS technology it is attractive to use low cost and/or flexible substrates like glass or metal foil instead of Si for NW growth (Tsakalakos et al. 2007; Andra et al. 2008). The technique of electroless metal-assisted chemical etching is preferable to create vertically aligned Si NWs of high electronic quality with desirable crystallographic orientation and doping characteristics. These Si NW arrays can be used as an efficient ARC for Si solar cells incorporating this array (Peng et al. 2005). The tapered NW solar cells demonstrated superior photovoltaic characteristics, such as a short circuit current 17.7 mA/cm^2 and a cell conversion efficiency of ~6.6% under 1.5 AM illumination (Jung et al. 2010).

5.12.2 Tandem Si solar cells with Si quantum dots

In standard solar cells, when photon with energy $2E_g>h\nu>E_g$ is absorbed it generates one electron-hole pair. After the photogeneration the conduction band electron and valence band hole quickly lose any energy in excess of the semiconductor band gap (thermal relaxation loses). This loss mechanism alone limits conversion efficiency to 44% (Green 2000) when the one band gap semiconductor is used[4].

Among the proposed concepts, the tandem cells approach is the only one which has already permitted to realize photovoltaic structures with efficiency exceeding the limit for a single band gap device (Green 2003). In tandem cell approach, the cells are stacked on top of one another. By placing the largest band gap cell uppermost, this cell will absorb the highest energy photon, allowing photons of lower energy to pass through to underlying cells, arranged in order of decreasing band gap. The cells can be connected together in series by tunnel junctions and each of them should generate the same current.[5]

Up to now, the tandem cell approach is based on monolithic integration of III-V materials by means of rather expensive technologies of fabrication such as molecular beam epitaxy. Thus, the concept of “all-Si” tandem solar cells appears as an attractive alternative permitting to replace III-V materials by Si and its dielectric compounds.

The concept of “all-Si” tandem solar cell with Si quantum wells or quantum dots sandwiched between layers of a dielectric based on Si compounds such as SiO_2, Si_3N_4, SiC is shown in Fig. 10.21 (Conibeer 2007; 2012). If quantum dots or wells are close enough each other so that neighbors interact, quantum levels broaden out into bands. For quantum dots of 2 nm (or

---

[4] The thermodynamic limit on solar energy conversion to electricity is 93% (Luque and Marti 1997).

[5] The theoretical limiting efficiency of tandem cell depends on the number of subcells in the device. For 1,2,3,4, and ∞ subcells, the theoretical limiting efficiency η are 31.0%, 42.5%, 48.6%, 52.5%, and 68.2%, respectively, for unconcentrated sun light (Conibeer 2007).

quantum wells of 1 nm) the miniband of larger band gap of 1.7 eV is formed which is optimal for tandem cell on top of Si.

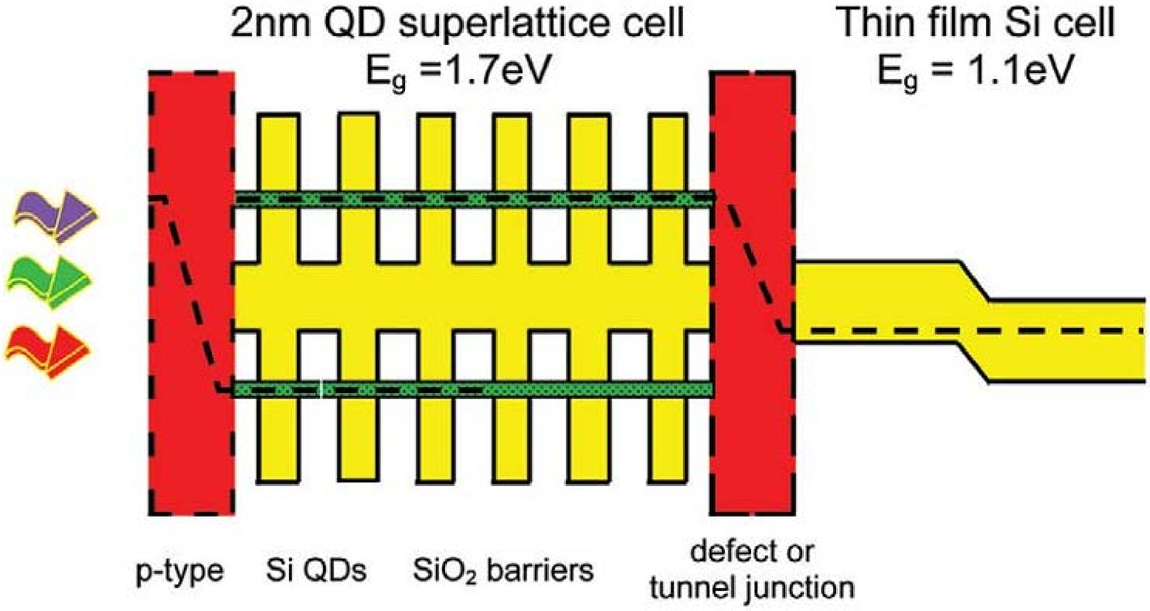


Fig.5.40. “All Si” tandem solar cell. Nanostructured cell consists of Si quantum wells or quantum dots in an amorphous dielectric matrix connected by a defect tunnel junction to Si cell (Conibeer G. , *Materials Today*, **10**, 42, 2007)

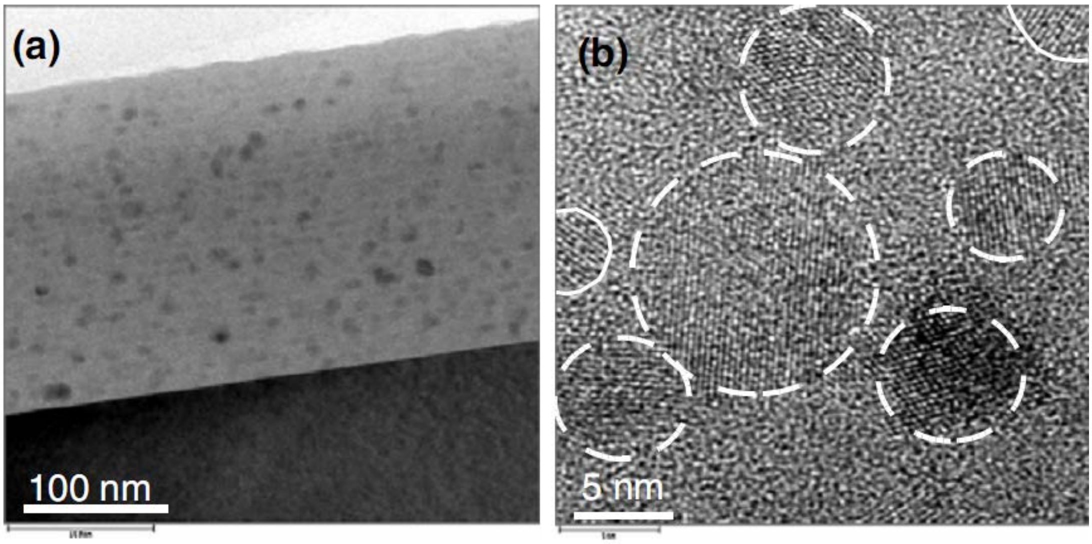


Fig.5.41 TEM images of Si QDs in a SiO2 matrix: (a) low-magnification image and (b) high-resolution image (Cho E.C. et al., *Nanotechnol.*,**19**, 245201 2008)

Optical and structural properties of Si quantum dots in dielectric matrix were studied (Nychyporuk et al. 2008; Delachat et al. 2009). A simple approach to prepare superlattices of Si quantum dots is reported (Zacharias et al. 2002). The Si QDs were formed by alternate deposition of SiO_2 and silicon-rich SiO*x* with magnetron co-sputtering, followed by high-temperature annealing (Fig. 10.22). On heating, surface energy minimization favors the precipitation of Si into approximately spherical quantum dots (Cho et al. 2008). The matrix remains amorphous, thus avoiding some of the problems of lattice mismatch. Other dielectrics (Si nitride, Si carbide) are of interest since their lower band gap compared to SiO_2 should greatly increase current transport in these materials.

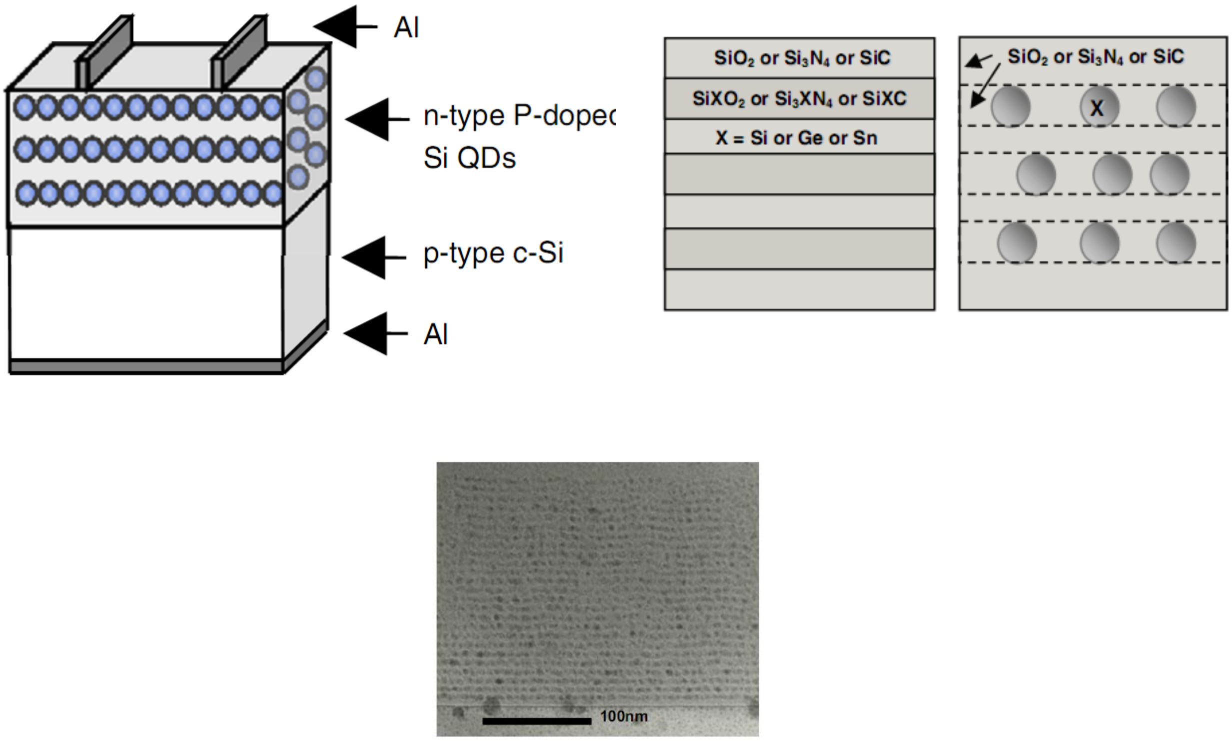


Fig.5.42a) Schematic diagram of an (n-type) Si QD/(p-type) c-Si photovoltaic device And Sketch for growth of dispersed quantum dots (a) and high resolution image showing quantum dots formed in the silicon rich layers (b) [Green 2008].

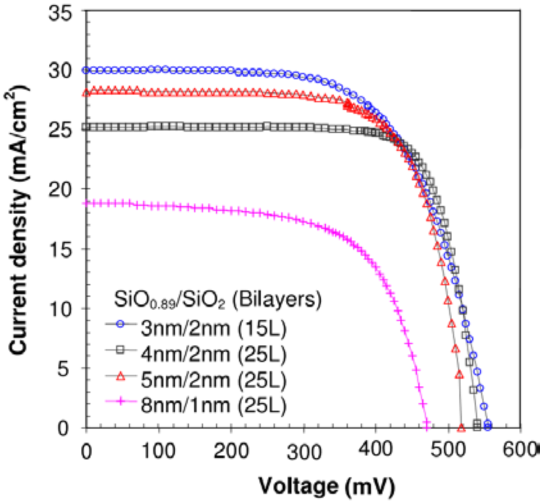


Fig.5.43 One-sun illuminated *I–V* curves of four different (n-type) Si QD/(p-type) c-Si solar cells measured at 298 K.(Cho E.C. et al., *Nanotechnol.*,**19**, 245201 2008.)

Schematic diagram of an (n-type) Si QD/(p-type) c-Si photovoltaic device is presented in Fig.5.42. Solar cells consists of phosphorus-doped Si QDs in a SiO$_2$ matrix deposited on p-type crystalline Si substrates (c-Si). Current tunnelling through the QD layer was observed from the solar cells with a dot spacing of 2 nm or less. To get the required current densities through the devices, the dot spacing in the SiO$_2$ matrix had to be 2 nm or less. The open-circuit voltage was found to increase proportionally with reductions in QD size, which may relate to a bandgap widening effect in Si QDs. One-sun illuminated I–V curves of four different (n-type) Si QD/(p-type) c-Si solar cells measured at 298 K are presented in Fig.5.43. The best cell parameters obtained were an open-circuit voltage V_{oc} of 556 mV, short-circuit current J_{sc} of 29.8 mAcm^{-2}, fill factor FF of 64%, and conversion efficiency of 10.6% from 3 nm Si QDs with a 2 nm SiO$_2$ layer (Cho et al. 2008).

REFERENCES

Aberle A.G. (2000) Surface passivation of crystalline silicon solar cells: a review. *Prog. Photovolt.: Res. Appl.* **8**, 473-487.

Abouelsaood A.A., Ghannam M.Y., Stalmans L., Poortmans J., and Nijs J.F. (2000) Theoretical analysis of light scattering in porous Si for PV applications. In: *Proceedings of 16h European photovoltaic solar energy conference*, 1-5 May, Glasgow, UK, 199-202.

Andra G., Pietsch M., Sivakov V., Stelzner Th., Gawlik A., Christiansen S. and Falk F. (2008) Thin film solar cells based on single crystalline silicon nanowires. In: *Proceeding of 23d European Photovoltaic Solar Energy conference*, 1-5 September, Valencia, Spain, 163-167.

Aroutiounian V., Martirosyan Kh. S. and Soukiassian P. (2004) Low reflectance of diamond-like carbon/porous silicon double layer antireflection coating for silicon solar cells, *J. Phys. D: Appl. Phys*. **37**, L25-L28.

Aulich H. A., Schulze, F–W., and Anspach O. (2010) Large scale crystallisation and wafer production: the way to 2020, In: *Proceeding of 25th European Photovoltaic Solar Energy Conf.*, 6-10 September, Valencia, Spain, pp. 1066-1072.

Bastide S., Le Quang N., Lust S., Korte L., Vard C., and Lévy-Clément C. (2004) Electrochemical macroporous texturization of Polix mc-Si solar cells, In: *Proceeding of 19th European photovoltaic solar energy conference*, 7-11 June, Paris, France, 943-946.

Bastide S., Tena-Zaera R., Lévy-Clément C., Barakel D., Palais O., and Martinuzzi, S. (2006) Photo-electrochemical texturization of N-type crystalline Si. In: *Proceeding of 21st European Photovoltaic Solar Energy conference*, 4-8 September, Dresden, Germany, 911-914.

Bastide S., Le Quang N., Monna R. and Lévy-Clément C. (2009). Chemical etching of Si by Ag nanocatalysts in HF-H_2O_2: application to multicrystalline Si solar cell texturization. *Phys. Stat. Sol. C.* **6**, 1536-1540.

Ben-Chorin M. (1997) Resistivity of porous silicon, In: *Properties of Porous Silicon*, ed. by L. Canham, EMIS datareviews series, N18, INPEC publ., Dera, Malvern, 165-175.

Ben Rabha, M., Boujmil, M. F., Saadoun, M. and Bessaïs, B. (2005). Chemical vapour etching-based porous silicon and grooving: application in silicon solar cell processing. *Phys. Stat. Solidi. C.* **2**, pp. 3505-3509.

Ben Rabha M., Mohamed S.B., Dimassi W., Gaidi M., Ezzaouia H., and Bessais B.(2011) Reduction of absorption loss in multicrystalline silicon via combination of mechanical grooving and porous silicon . *Phys. Stat. Solidi C.* **8**(3), 883–886.

Ben Rabha M., Salem M., El Khakani M.A., Bessais B., Gaadi M. (2013) Monocrystalline silicon

surface passivation by Al_2O_3/porous silicon combined treatment. *Mat. Sci. Eng. B.* **178**, 695–697.

Bilyalov R. R., Lautenschlager H., Schetter C., Schomann F., Schubert U. and Schindler R. (1997) Porous silicon as an antireflection coating for multicrystalline silicon solar cells, In: *Proceeding of 14th European photovoltaic solar energy conference*, 30 June- 4 July, Barcelona, Spain, 788-791.

Bilyalov R.R., Ludemann R., Wettling W., Stalmans L., Poortmans J., Nijs J., Schirone L., Sotgiu G., Strehlke S., Levy-Clement C. (2000) Multicrystalline silicon solar cells with porous silicon emitter. *Solar Energy Mater. Solar Cells.* **60**, 391-420.

Bilyalov R., Stalmans L., Baucarne G., Loo R., Caymax M., Poortmans J. and Nijs J. (2001) Porous silicon as an intermediate layer for thin-film solar cell, *Solar Energy Materials Solar Cells*. **65**, 477-485.

Book F., Dastgheib-Shirazi A., Raabe B., Haverkamp H., Hahn G., and Grabitz, P. (2009) Detailed analysis of high sheet resistance emitters for selectively doped silicon solar cells, In: *Proceeding of 24th European Photovoltaic Solar Energy Conference,* 21-25 September, Hamburg, Germany, 1719-1722.

Brendel R. (1997) A novel process for ultrathin monocrystalline silicon solar cells on glass. In: *Proceeding of Proc. 14th European Photovoltaic Solar Energy Conf.,* 30 June- 4 July, Barcelona, Spain, 1354-1357.

Brendel R. (2004) Thin-film crystalline silicon mini-modules using porous Si for layer transfer, *Solar Energy,* **77,** 969-982.

Burgers A. R., Naber R. C. G., Carr A. J., Barton P. C., Geerlings L. J., Jingfeng X., Gaofei L., Weipeng S., Haijiao A., Zhiyan H., Venema P. R. and Vlooswijk, A. H. G. (2008) 19% efficient n-type Si solar cells made in pivot production, In: *Proceeding of 23d European Photovoltaic Solar Energy Conference,* 1-5 September, Valencia, Spain, 1106-1109.

Canham, L.T. (1997). Storage of porous silicon. In: Canham L. (ed) *Properties of Porous Silicon*. EMIS data reviews series, N18, INPEC publ., Dera, Malvern, 44-50.

Chaoui R., Mahmoudi B. and Si Ahmed Ya. (2008) Porous Si antireflection layer for solar cells using metal-assisted chemical etching. *Phys. Stat. Sol. A.* **205**, 1724-1728.

Chartier C., Bastide S. and Lévy-Clément C., (2007) Metal-assisted wet chemical etching of crystalline Si. In: *Proceeding of 22nd European Photovoltaic Solar Energy Conference,* 3-7 September, Milan, Italy, 1231-1234.

Cho E.C., Park S., Hao X., Song D., Conibeer G., Park S.C., and Green M. A. (2008) Silicon quantum dot/crystalline silicon solar cells. *Nanotechnology.* **19,** 245201 (5pp).

Conibeer G. (2007) Third-genaration photovoltaics, *Materials Today*. **10**, 42-50.

Conibeer G., Perez-Wurfl I., Hao X., Di D. and Lin D. (2012) Si solid-state quantum dot-based materials for tandem solar cells. *Nanoscale Res. Letter.* **7**, art.193.

Delachat F., Carrada M., Ferblantier G., Slaoui A., Keita A.S., Naciri A.En., Kloul M., and Gaston J.P. (2009) Spectrocscopic ellipsometry analysis of silicon-rich silicon nitride layers for photovoltaic applications. In: *Proceeding of 24th European Photovoltaic Solar Energy Conference*, 21-25 September, Hamburg, Germany, 520-523.

Derbali L. Dimassi W., and Ezzaouia H. (2011) Improvement of the minority carrier mobility in low-quality multicrystalline silicon using a porous silicon-based gettering under an O_2 atmosphere. *Energy Procedia.* **10**, 243 – 248.

Derbali L., and Ezzaouia H. (2012) Phosphorus diffusion gettering process of multicrystalline silicon using a sacrificial porous silicon layer. *Nanoscale Res. Letter*. **7**, art.338.

Dubey R. S. (2013) Electrochemical fabrication of porous silicon structures for solar cells. *Nanoscience and Nanoengineering* **1**(1), 36-40.

Duerinckx F., and Szlufcik J. (2002) Defect passivation of industrial multicrystalline solar cells based on PECVD silicon nitride, *Solar Energy Mater. Solar Cells* **72**, 231-246.

Duerinckx F., Kuzma Filipec I., Van Nieuwenhuysen K., Beaucarne G., Poortmans J., Leloup F., Versluys J., and Hanselaer P. (2006) Optical path length enhancement for >13% screenprinted thin film silicon solar cells. In: *Proceeding of 21st European Photovoltaic Solar Energy Conference*, 4-8 September, Dresden, Germany, 726-729.

Dzhafarov T.D., Aslanov S.S., Ragimov S.H., Sadigov M.S., and Aydin Yuksel S. (2012) Effect of nanoporous silicon coating on silicon solar cell performance. *Vacuum* **86**, 1875–1879.

Dzhafarov T. (2014) Porous silicon and solar sells. In: Canham L. (ed) *Handbook of Porous Silicon,* DOI 10.1007/978-3-319-04508-5_95-1, Springer Int.Publ., Switzerland.

Efremov A.A., Klyui N.I., Litovchenko V.G., Popov V.G., Romanyuk A.B., and Romanyuk B.N. (2000) Development of gettering process for the preparation of the solar silicon material. *Opto-Electronics Rev.* **8**, 410-413.

European Photovoltaic Industry Association (EPIA) (2014) Global market outlook for photovoltaics 2014-2018, Brussels, Belgium, www.epia.org.

Evtukh A.A., Litovchenko V.G., Oberemok A.S., Popov V.G., Rasssamakin Yu.V., Romanyuk B.N., and Volkov S.G. (2001) Investigations of impurity gettering in multicrystalline silicon, *Semiconductor Physics, Quantum Electronics & Optoelectronics.* **4**, 278-282.

Fang H., Li X., Song S., Xu Y., and Zhu J. (2008) Fabrication of slantingly-aligned silicon nanowire arrays for solar cell applications, *Nanotechnol.***19,** art. 255703.

Feldrapp, K., Horbelt, R., Auer R. and Brendel R. (2003) Thin-film (25,5 µm) solar cells from layer transfer using porous silicon with 32,7 mA/cm^2 short-circuit current, *Progress in Photovoltaics*, **11**, 105-112.

Gamboa R., Martins M., Serra J.M., Maia Alves J., Vallêra A.M., Ponomarev E. A., and Lévy-Clément C. (1998) First solar cells on electrochemically textured macroporous silicon, In: *Proceeding of 2nd World Conf. and Exhibition on Photovoltaic Solar Energy Conversion*, 6-10 July,Vienna, Austria, 1669-1672.

Green M. A. (1992) *Solar cells : Operating principles, technology and system applications, Prentice –Hall, Englewood Cliffs, Kensington*.

Green M. A. and Keevers, M. J. (1995). Optical properties of intrinsic silicon at 300K. *Prog. Photovoltaics*. **3**, pp. 189-192.

Green, M.A. (2000). Potential for low dimensional structures in photovoltaics. *Mat. Sci. Eng. B*. **74**, 118-124.

Green M. (2003). *Third generation photovoltaics (advanced solar energy conversion). Springer, Berlin.*

Green M. , Emery K., Hishikawa Y., and Warta W. (2011) Solar cell efficiency tables (version 37). *Prog. Photovolt: Res. Appl*. **19**, 84–92.

Grigoras K., Härkönen J., Jasutis V., Kindurys A., Pacebutas V., Sabataityte J., and Simkiene I. (1998) Porous silicon emitter formation from spin-on glasses, In:*Proceeding of 2nd World Conf. and Exhibition on Photovoltaic Solar Energy Conversion*, 6-10 July, Vienna, Austria, 1717-1720.

Haddadi I., Dimassi W., Bousbih R., Hajji M., Ali Kanzari M., and Ezzaouia H. (2011) Improvement of solar cells performances by surface passivation using porous silicon chemically treated with LiBr solution. *Phys. Status Solidi C* **8** (3), 755–758

Halm A., Jourdan J., Nichol S., Ryningen B., Tathgar H., and Kopecek R. (2010) Detailed study on large area 100% SOG Silicon MC solar cells with efficiencies exceeding 16%, In: *Proceeding of 25d European Photovoltaic Solar Energy Conf.,* 6-10 September, Valencia, Spain, 1210-1215.

Hassen M., Hajji M., Ben Jaballah A., Khedher N., Dimassi W., Ezzaouia H., Bessaïs B., Selmi A., and Bouchriha H. (2004) Improvement of minority carrier diffusion length in solar grade silicon: a new method. In: *Proceeding of 19th European Photovoltaic Solar Energy Conf.*, 7-11 June, Paris, France, 766-768.

Hochbaum A.I. and Yang P. (2010) Semiconductor Nanowires for Energy Conversion, *Chem. Rev*.**110,** 527-546.

Horbelt R., Terheiden B., Auer R. and Brendel R. (2005) Manifold Use of Growth Substrate in the Porous Silicon - Layer Transfer Process, *Record 31th IEEE Photovoltaic Specialist Conf.* 3-7 January, Orlando, USA, 1193-1196.

Hsueh T-J.,. Chen H-Y, Tsai T-Y., Dai B-T., Weng W-Y., Hsueh H-T., and Shieha J-M. (2011) A microridge-like structured c-Si solar cell prepared by reactive ion etching. *J. Electrochem.Soc.* **158** (1), H35-H37.

Ivanov I. I., Nychyporuk T., Skryshevsky V. A., and Lemiti M. (2009) Thin silicon solar cells with SiO_x/SiN_x Bragg mirror rear surface reflector. *Semiconductor Physics, Quantum Electronics Optoelectronics*.**12**, 406-411.

Ivanov I.I., Skryshevsky V.A., Nychyporuk T., Lemiti M., Makarov A.V., Klyui N.I., and Tretyak O.V. (2013) Porous silicon Bragg mirrors on single- and multi-crystalline silicon for solar cells. *Renewable Energy*. **55**, 79-84.

Jin S., Bender H., Stalmans L., Bilyalov R., Poortmans J., Loo R., and Caymax M. (2000) Transmission electron microscopy investigation of the crystallographic quality of silicon films grown epitaxially on porous silicon. *J. Crystal Growth.* **212**, 119-127.

Jung J-Y., Guo Z., Jee S-W., Um H-D., Park K-T., and Lee J-H. (2010) A strong antireflective solar cell prepared by tapering silicon nanowires. *Optics Express*. **18**, A286-A292.

Kayes B. M., Atwater H. A., and Lewis N.S. (2005) Comparison of the device physics principles of planar and radial p-n junction nanorod solar cells, *J.Appl.Phys***. 97**, art. 114302.

Kelzenberg M.D., Turner-Evans D.B., Kayes B.M., Filler M.A. Putnam M.C., Lewis N.S., and Atwater H.A. (2008) Photovoltaic measurements in single-nanowire silicon solar cells. *Nano Lett*. **8**, 710- 714.

Khedher N., Hajji M., Bouaïcha M., Boujmil M.F., Ezzaouia H. Bessaïs B., and Bennaceur R. (2002) Improvement of transport parameters in solar grade monocrystalline silicon by application of a sacrificial porous silicon layer. *Solid State Comm*. **123**, 7-10.

Krotkus A., Grigoras K., Pacebutas V., Barsony I., Vazsonyi E., Fried M., Szlufcik J., Nijs J., and Lévy-Clément C. (1997) Efficiency improvement by porous silicon coating of multicrystalline solar cells. *Solar Energy Mater.Solar Cells*. **45**, 267-273.

Kuthi E.B. (2004) Crystalline silicon solar cells with selective emitter and the self-doping contact. *Híradástechnika*. **59**, 21-31.

Kuzma-Filipek, I., Duerinckx, F., Van Nieuwenhuysen, K., Baucarne, G., Poortmans, J. and Mertens, R. (2007). Porous silicon as an internal reflector in thin epitaxial solar cells, *Phys. Stat. Solidi (a),* 204, pp 1340-1345.

Kuzma-Filipek I. J., Duerinckx F., Van E. Kerschaver, Van Nieuwenhuysen K., Beaucarne G., and Poortmans J. (2008) Chirped porous silicon reflectors for thin-film epitaxial silicon solar cells. *J. Appl. Phys*. **104**, art.073529.

Kuzma-Filipek I., Recaman-Payo M., Van Nieuwenhuysen K., Van Hoeymissen J., Baucarne G., Van Kerschaver E., Poortmans J., and Mertens R. (2009) Rear junction epitaxial thin film solar cells with diffused front surface field and porous Si back reflectors. In: *Proceeding of 22nd European Photovoltaic Solar Energy Conf.*, 21-25 September, Hamburg, Gemany, 2584-2588.

Lévy-Clément C., Lust S., Bastide S., Le Quang N., and Sarti D. (2003) Macropore formation on p-type multicrystalline Si and solar cells. *Phys. Stat. Sol. A.* **197**(1), 27-33.

Lian S.S., Kammel R., and Kheiri M.J. (1992) Preliminary study of hydrometallurgical refining of MG-silicon with attrition grinding. *Solar Energy Mater.Solar Cells*. **26**, 269-276.

Lipinski M., Panek P., Bielanska E., Weglowska J., and Czternastek H. (2000) Influence of porous silicon on parameters of silicon solar cells. *Optoelectron. Rev*., 8 (4), 418-420.

Lipinski M., Bastide S., Panek P., and Lévy-Clément C. (2003) Porous Si antireflection coating by electrochemical and chemical etching for Si solar cells manufacturing, *Phys. Stat. Sol.A.* **197**, 512-517.

Litovchenko V. G., and Kluyi N. I. (2001) Solar cells based on DLC film-Si structures for space application, *Solar Energy Mater. Solar Cells*. **68**, 55-70.

Luque A., and Marti A. (1997) Entropy production in photovoltaic conversion. *Phys. Rev. B,* **55**, 6994-6999.

Marrero N., Guerrero-Lemus R., Gonzalez-Diaz B., and Borchert, D. (2008) Effect of porous silicon stain etched on large area alkaline textured crystalline silicon solar cells. *Thin Solid Films,* **517**, 2648-2650.

Martinuzzi S., Palais O., Pasqunelli M., and Farrazza F. (2005) N-type multicrystalline silicon wafers and rear junction solar cells, *European Phys. J. Appl. Phys*. **32,** 187-192.

Matic Z., Bilyalov R.R., and Poortmans J. (2000) Firing through porous silicon antireflection coating for silicon solar cells. *Phys. Stat. Sol. A.***182**, 457-460.

Menna P., Tsuo Y.S., Al-Jassim M.M., Asher S.E., Matson R., and Ciszek T.F. (1998) Purification of metallurgical grade silicon by porous silicon etching, In: *Proceeding of 2nd World Conf. and Exhibition Photovoltaic energy Conversion,* 6-10 July, Vienna, Austria, 1232-1235.

Mizsei J., Shrair J.A., and Zolomy I. (2004) Investigation of Fermi-level pinning at silicon/porous-silicon interface by vibrating capacitor and surface photovoltage measurements, *Appl. Surf. Sci.* **235**, 376-388.

Moon I., Kim K., Thamilselvan M., Kim Y., Han K., Kyeong D., Kwon T., Ai D.V., Lee J., Ju M.,.Lee K, and Yi J. (2009) Selective emitter using porous silicon for crystalline silicon solar cells. *Solar Energy Mater. Solar Cells*. **93**, 846–850.

Nagel H., Aberle A.G., and Hezel R. (1999) Optimised antireflection coatings for planar silicon solar cells using remote PECVD silicon nitride and porous silicon dioxide. *Prog. Photovolt: Res. Appl,* **7,** 245-260.

Nichiporuk O., Kaminski A., Lemiti M., Fave A., and Skryshevsky V. (2005) Optimisation of interdigitated back contacts solar cells by two-dimensional numerical simulation, *Solar Energy Mater. Solar Cells.* **86**, 517-526.

Nichiporuk O., Kaminski A., Lemiti M., Fave A., Litvinenko S., and Skryshevsky V. (2006) Passivation of the surface of rear contact solar cells by porous silicon, *Thin Solid Films,* **511-512**, 248-251.

Nychyporuk T., Marty O., Rezgui B., Sibai A., Lemiti M., and Bremond G. (2008) Towards the 3rd generation photovoltaic: absorption properties of silicon nanocrystals embedded in silicon nitride matrix. In: *Proceeding of 23d European Photovoltaic Solar Energy Conf.,* 1-5 September, Valencia, Spain, 491-494.

Osorio E., Urteaga R., Acquaroli L.N., Garcı´a-Salgado G., Juare´z H., Koropecki R.R. (2011) Optimization of porous silicon multilayer as antireflection coatings for solar cells. *Solar Energy Mater. Solar Cells*. **95**, 3069–3073.

Peng K.Q., XuY., WuY., YanY., Lee S.-T., and Zhu J. (2005) Aligned single-crystalline Si nanowire arrays for photovoltaic applications. *Small.* **1**, 1062-1067.

Petermann J.H., Ohrdes T., Altermatt P. P., Eidelloth S., and Brendel R. (2012) 19% Efficient thin-film crystalline silicon solar cells from layer transfer using porous silicon: a loss analysis by means of three-dimensional simulations. *IEEE Transactions Electron Devices*. 59 (4), 909-917.

Petermann J. H., Zielke D., Schmidt J., Haase F.,. Garralaga Rojas E, and Brendel R. (2012) 19%-efficient and 43 μm-thick crystalline Si solar cell from layer transfer using porous silicon,. *Progr. Photovol., Res. App.* 20 (1), 1–5.

Prasad A., Balakrishnan S., Jain S.K., and Jain G.C. (1982) Porous silicon oxide anti-reflection coating for solar cells. *J. Electrochem. Soc*. 129, 596-599.

Ramizy A., Hassan Z., Omar K., Al-Douri Y., Mahdi M.A. (2011) New optical features to enhance solar cell performance based on porous silicon surfaces. *Applied Surf. Sci*. **257** 6112–6117.

Rinke T.J., Bergmann R.B., Brugemann R., Werner J.H. (1999) Ultrathin quasi-monocrystalline silicon films for electronic devices. *Solid State Phenomena***. 67–68**, 229-236.

Salman K. A., Omar K., Hassan Z. (2012) Effective conversion efficiency enhancement of solar cell using ZnO/PS antireflection coating layers. *Solar Energy.* **86**, 541–547.

Schirone L., Sotgiu G., Montecchi M., Righini G., and Zanoni R. (1998) Stain etched porous silicon technology for large area solar cells. In: *Proceeding of 2nd World Conf. and Exhibition on Photovoltaic Solar Energy Conversion,* 6-10 July, Vienna, Austria, 276-279.

Schnell M., Lüdemann R., Schaefer S. (2000) Stain etched porous silicon – a simple method for the simultaneous formation of selective emitter and ARC, In: *Proceeding of 16th European Photovoltaic Solar Energy Conf.*, 1-5 May, Glasgow, UK, 1482-1485.

Shieh S.Y., and Evans, J.W. (1993) Some observations of the effect of porous silicon on oxidation induced stacking faults, *J. Electrochem. Soc*. **140**, 1094-1096.

Sivakov V., Andrä G., Gawlik A., Berger A., Plentz J., Falk F., and. Christiansen S. H. (2009) Silicon Nanowire-Based Solar Cells on Glass: Synthesis, Optical Properties, and Cell Parameters. *Nano Letter*. 9(4), 1549-1554

Skryshevsky V.A., Laugier A., Vikulov V.A., Strikha V.I., and Kaminski A. (1996a) Effect of porous silicon layer re-emission on silicon solar cell photocurrent In: *Proceeding of 25th IEEE Photovoltaic Spec.Conf.*, May, Washington,USA, 589-592.

Skryshevsky V.A., Laugier A., Strikha V.I., and Vikulov V.A. (1996b) Evaluation of quantum efficiency of porous silicon photoluminescence. *Mat.Sci.Eng. B.* **40**, 54-57.

Skryshevsky V.A., and Laugier A. (1999) Improved thin film solar cell with Rayleigh's scattering in porous silicon pipes *Thin Solid Films*. **346**, 254-258.

Skryshevsky V. A., Kilchitskaya S. S., Kilchitskaya T. S., Litvinenko S. V., Nichiporuk O., Fave, A. and Laugier A. (2000) Impact of recombination and optical parameters on silicon solar cell with the selective porous silicon antireflection coating, In: *Proceeding of 16th European Photovoltaic Solar Energy Conf.*, 1-5 May, Glasgow, UK, 1634-1637.

Smestad G., Kunst M., and Vial C. (1992) Photovoltaic responce in electrochemically prepared photoluminescent porous silicon, *Solar Energy Mater. Solar Cells*, **26**, 277 - 283.

Solanki C.S., Bilyalov R.R., Poortmans J., and Nijs J. (2002) Transfer of a thin silicon film on to a ceramic substrate. *Thin Solid Films*. **403–404**, 34-38.

Solanki C. S., Bilyalov R. R., Poortmans J., Nijs J., and Mertens R. (2004) Porous silicon layer transfer processes for solar cells. *Solar Energy Mater. Solar Cells.* **83**, 101-113.

Springer J., Poruba A., Fejfar A., and Vanecek M. (2000) Nanotextured thin film silicon solar cells: optical model. In: *Proceeding of 16th European Photovoltaic Solar Energy Conf.*, 1-5 May, Glasgow, UK, 434-437.

Srivastava S. K., Kumar D., Singh P.K., Kar M., Kumar V., and Husain M. (2010) Excellent antireflection properties of vertical silicon nanowire arrays. *Solar Energy Mater.Solar Cells.* **94,** 1506–1511.

Stalmans L., Laureys W., Said K., Caymax M., Poortmans J., Nijs J., Mertens R., Strehlke S., Levy-Clement C., Ventura L., and Slaoui A. (1997) Effect of a porous silicon surface layer on the on the internal quantum efficiency of crystalline silicon solar cells. In: *Proceeding of 14h European Photovoltaic Solar Energy Conf.*, 30 June- 4 July, Barcelona, Spain, 2484-2487.

Stalmans L., Poortmans J., Bender H., Caymax M., Said K., Vazsonyi E., Nijs J. and Mertens R. (1998) Porous silicon in crystalline silicon solar cells: a review and the effect on the internal quantum efficiency. *Prog. Photovolt. Res. Appl.* **6**, 233-246.

Stelzner T., Pietsch M., Andra G., Falk F., Ose E., and Christiansen S. (2008) Silicon nanowire-based solar cells. *Nanotechnol.***19**, art. 295203.

Strehlke S., Sarti D., Krotkus A., Polgar O., Fried M., Roger J. P., and Lévy-Clément C. (1997). Porous Si emitter and high efficiency multicrystalline silicon solar cells, In: *Proceeding of 14th European Photovoltaic Solar Energy Conf.*, 30 June- 4 July, Barcelona, Spain, 2480-2483.

Striemer C., and Fauchet, P. M. (2003) Dynamic etching of silicon for solar cell applications, *Phys. Stat. Sol. A.* **197**, 502-506.

Sze S.M. (1981) *Semiconductor Devices, New York, John Wiley.*

Tena-Zaera R., Bastide S., Lévy-Clément C. (2007) Photoelectrochemical texturization of n-type multicrystalline Si, *Phys. Stat. Sol. A.* **204**, 1260-1265.

Thogersen A., Selj J. H., and Marstein E. S. (2012) Oxidation effects on graded porous silicon anti-reflection coatings. *J. Electrochem. Soc.* **159** (5) D276-D281.

Tian B., Zheng X., Kempa T.J., Fang Y., Yu N., Yu G., Huang J., and Lieber C.M. (2007) Coaxial silicon nanowires as solar cells and nanoelectronic power sources. *Nature*. **449**, 885-889.

Tsakalakos L., Balch J., Fronheiser J., Korevaar B.A., Sulima O., Rand J. (2007) Silicon nanowire solar cells. *Appl.Phys.Lett*. **91**, art. 233117.

Tsuo Y.S., Xiao Y., Heben M.J., Wu X., Pern F.J., and Deb S.K. (1993) Potential applications of porous silicon in photovoltaics, *Record 23th IEEE Photovoltaic Specialists Conf.*, May, Louisville, USA, 287-393.

Tsuo Y.S., Menna P., Pitts J.R., Jantzen K.R., Asher S.E., Al-Jassim M.M., and Ciszek T.F. (1996) Porous silicon gettering. *Record 25th IEEE Photovoltaic Specialist Conf.* May,Washington, USA, 461-464.

Van Hoeymissen J., Kuzma-Filipek I., Van Nieuwenhuysen K., Duerinckx F., Baucarne G. and Poortmans J. (2008). Thin film epitaxial solar cells on low-cost Si substrates: closing the efficiency gap with Si cells using advanced photonic structures and emitters, *Proc.23th European Photovoltaic Solar Energy Conf.*, 1-5 September, Valencia, Spain, pp. 2037-2041.

Van Hoeymissen, J., Depauw V., Kuzma-Filipek I., Van Nieuwenhuysen K., Payo M. R., Qiu Yu., Gordon I., and Poortmans J. (2011) The use of porous silicon layers in thin-film silicon solar cells, *Phys. Status. Solidi A.* **208** (6), 1433–1439.

Werner J.H., Wagner T. A. Gemmer C.E.M., Berge C., Brendle W. and Schubert M.B. (2003). Recent progress on transfer-Si solar cells at IPE-Stuttgart. In: *Proceeding of 3rd World Conf. Photovoltaic Energy Conversion*, 11-18 May, Osaka, Japan, 1272-1275.

Wolf S., and Tauber R.N. (1986) *Silicon processing for the VLSI Era, Vol.**1**, Lattice Press, Sunset Beach, SA.*

Wong-Leung J., Ascheron C.E., Petravic M., Elliman R.G., and Williams J.S. (1995) Gettering of copper to hydrogen-induced cavities in silicon. *Appl. Phys. Lett*. **66**, 1231-1233.

Yae S., Tanaka H., Kobayashi T., Fukumuro N., Matsuda H. (2005) Porous Si formation by HF chemical etching for antireflection of solar cells, *Phys. Stat. Sol. C.* **2**, 3476-3480.

Yerokhov V.Yu., Melnyk I.I. (1999) Porous silicon in solar cell structures: a review of achievements and modern directions of further use. *Renewable Sustainable Energy Rev*. **3**, 291-322.

Yerokhov V.Y. , Hezel R., Lipinski M., Ciach R., Nagel H., Mylyanych A., and Panek P. (2002) Cost-effective methods of texturing for silicon solar cells. *Solar Energy Mater. Solar Cells*. **72** (1-4), 291-298.

Yonehara T., Sakaguchi K., Sato N. (1994) Epitaxial layer transfer by bond and etch back of porous Si. *Appl. Phys. Lett*. **64**, 2108-2110.

Zacharias M., Heitmann J., Scholz R., Kahler U., Schmidt M., Bläsing J. (2002) Size-controlled highly luminescent silicon nanocrystals: a SiO/SiO_2 superlattice approach. *Appl. Phys. Lett.* **80**, 661-663.

Zettner J., Thoenissen M., Hierl Th., Brendel R., and Schulz M. (1998) Novel porous silicon backside light reflector for thin silicon solar cells. *Progress in Photovoltaics: Res.Appl.*, **6**, 423-432.

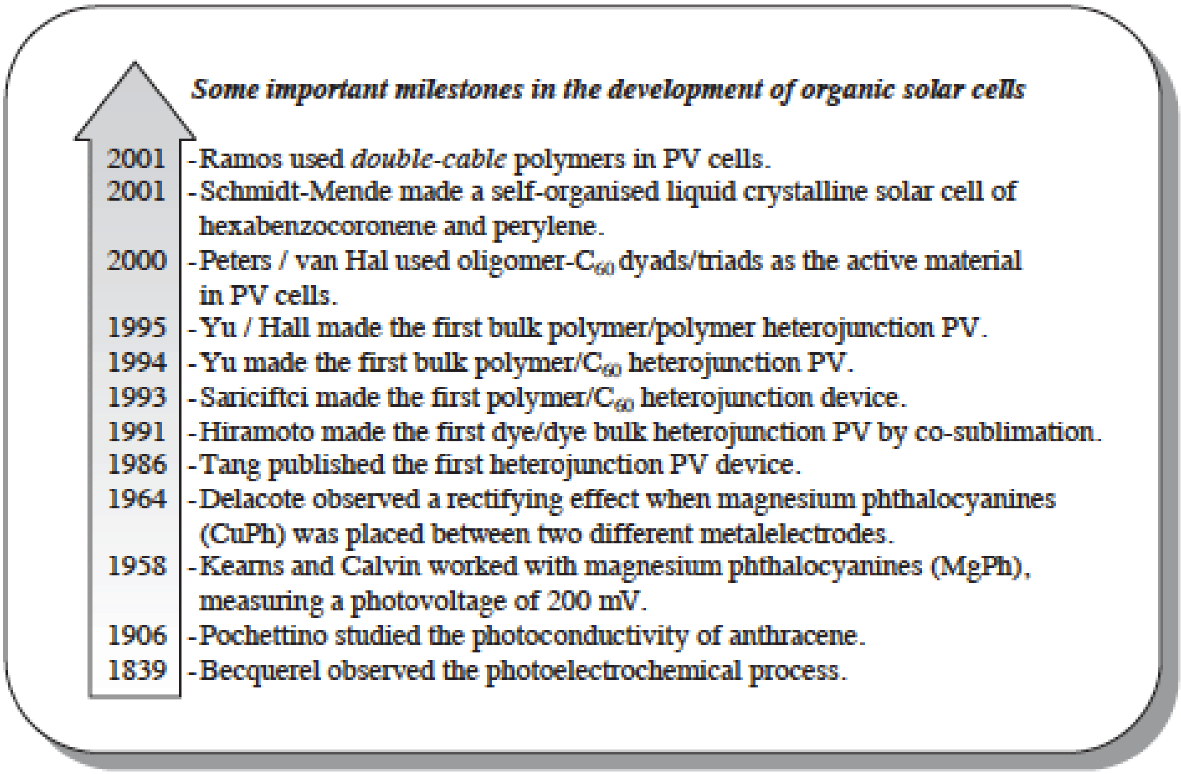
Zhang C., Liu S., Wang Y., Chen Y. (2010) Performance improvements of solar-grade crystalline silicon by continuously variable temperature phosphorous getteringprocess using a porous silicon layer. *Materials Sci. Semicond. Processing*. **13**, 209–213.

Zheng J.P., Jiao K.L., Shen, W.P., Anderson, W.A. and Kwok, H.S. (1992). Highly sensitive photodetector using porous silicon, *Appl. Phys. Lett.*, 61, pp.459-462.

Chapter 6 . Organic, Polymeric and Dye-Sensitized Cells

6.1 Introduction

The observation of photoconductivity in solid anthracene in the beginning of the 19th century marked the start of organic solar cells (SC). The first real investigations of photovoltaic (PV) devices came in the 1950s, where a number of organic dyes, particularly chlorophyll and related compounds, were studied. In the 1980s the first polymers (including poly(sulphur nitride) and polyacetylene) were investigated in PV cells. However simple PV devices based on dyes or polymers yield limited power conversion efficiencies (PCE), typically well below 0.1%. A major breakthrough came in 1986 when Tang discovered that bringing a donor and an acceptor together in one cell could dramatically increase the PCE to 1%.



Some important milestones in the development of organic solar cells

2001 - Ramos used *double-cable* polymers in PV cells.
2001 - Schmidt-Mende made a self-organised liquid crystalline solar cell of hexabenzocoronene and perylene.
2000 - Peters / van Hal used oligomer-C_{60} dyads/triads as the active material in PV cells.
1995 - Yu / Hall made the first bulk polymer/polymer heterojunction PV.
1994 - Yu made the first bulk polymer/C_{60} heterojunction PV.
1993 - Sariciftci made the first polymer/C_{60} heterojunction device.
1991 - Hiramoto made the first dye/dye bulk heterojunction PV by co-sublimation.
1986 - Tang published the first heterojunction PV device.
1964 - Delacote observed a rectifying effect when magnesium phthalocyanines (CuPh) was placed between two different metalelectrodes.
1958 - Kearns and Calvin worked with magnesium phthalocyanines (MgPh), measuring a photovoltage of 200 mV.
1906 - Pochettino studied the photoconductivity of anthracene.
1839 - Becquerel observed the photoelectrochemical process.

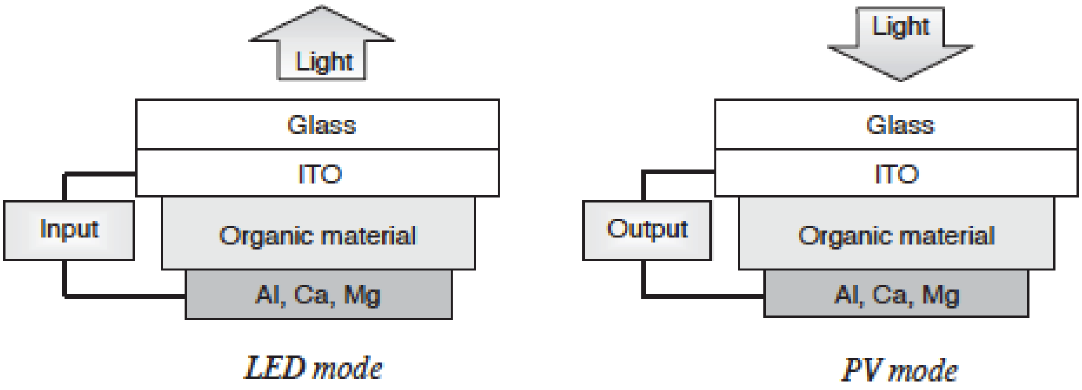
Fig.1 Main milestone of organic SC

This concept of heterojunction has since been widely exploited in a number of donor–acceptor cells, including dye/dye, polymer/dye, polymer/polymer and polymer/fullerene blends. Due to the high electron affinity of fullerene, polymer/fullerene blends have been subject to particular investigation during the past decade. Earlier problems in obtaining efficient charge carrier separation have been overcome and PCE of more than 3% have been reported. Some main milestone of organic SC is shown in Fig.6.1.

6.2 The base of organic solar

Almost all organic solar cells have a planar-layered structure, where the organic light-absorbing layer is sandwiched between two different electrodes. One of the electrodes must be (semi-) transparent, often Indium–tin-oxide (ITO), but a thin metal layer can also be used. The other electrode is very often aluminium (calcium, magnesium, gold and others are also used). Basically, the underlying principle of a light-harvesting organic PV cell (sometimes referred to as photodetecting diodes) is the reverse of the principle in light emitting diodes (LEDs) (see Fig. 6.2) and the development of the two are somewhat related. In LEDs an electron is introduced at the low-workfunction electrode (cathode) with the balanced introduction of a hole at the high-workfunction electrode (anode).

At some point the electron and the hole meets, and upon recombination light is emitted. The reverse happens in a PV device. When light is absorbed an electron is promoted from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO) forming an exciton (see Fig. 6.3). In a PV device this process must be followed by exciton dissociation. The electron must then reach one electrode while the hole must reach the other electrode. In order to achieve charge separation an electrical field is needed, which is provided by the asymmetrical ionisation energy/workfunctions of the electrodes. This asymmetry is the reason why electron-flow is more favoured from the low-workfunction electrode to the high workfunction electrode (forward bias), a phenomenon referred to as rectification. The light harvesting process along with the positioning of energy levels is depicted in Fig. 6.3.



LED mode *PV mode*

Fig. 6.2. A PV device (right) is the reverse of a LED (left). In both cases an organic material is sandwiched between two electrodes. Typical electrode materials are shown in the figure. In PVs electrons are collected at the metal electrode and holes are collected at the ITO electrode. The reverse happens in a LED: electrons are introduced at the metal electrode (cathode), which recombine with holes introduced at the ITO electrode (anode).

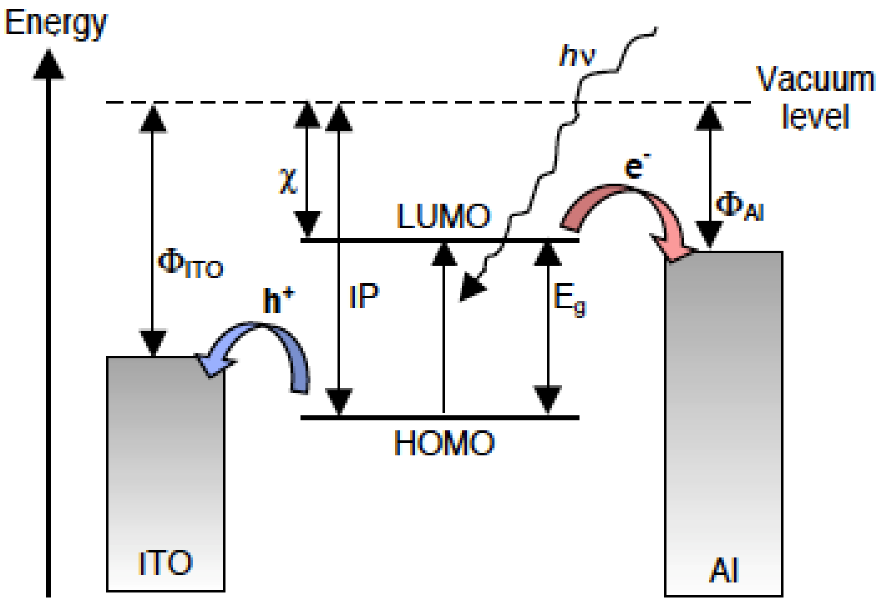


Fig. 6.3. Energy levels and light harvesting. Upon irradiation an electron is promoted to the LUMO leaving a hole behind in the HOMO. Electrons are collected at the Al electrode and holes at the ITO electrode. Φ: workfunction, χ: electron affinity, IP: ionisation potential, Eg: optical bandgap.

In the solid phase, the HOMOs and LUMOs of adjacent molecules may interact and form a conduction band (CB) and a valance band (VB) respectively. The shape of the CB and VB changes when the organic material is put into contact with electrodes (see Fig. 6.4), depending on the conductance of the polymer and on whether the electrodes are connected or not. If the cell is short circuited the Fermi levels of the electrodes align (B and C), and in doing so the CB and VB are pulled skew. In B the polymer material is an insulator. This gives a field profile that changes linearly through the cell. In C a hole-conducting (p-type) semiconductor is used (most polymers are much better hole conductors than electron conductors). If the material is doped or illuminated charge carriers are generated.

Due to the p-conduction properties, the generated holes are allowed to redistribute freely and they will flatten the bands approaching the high-workfunction electrode (a Schottky junction). The distance over which the CB and VB exhibit curvature is called the depletion width. In B the depletion width extends throughout the material. In C the depletion width is less than half the material thickness. Under external bias the relative electrode potentials can be changed, depending on the size and direction (forward or reverse) of the bias (a Schottky junction).

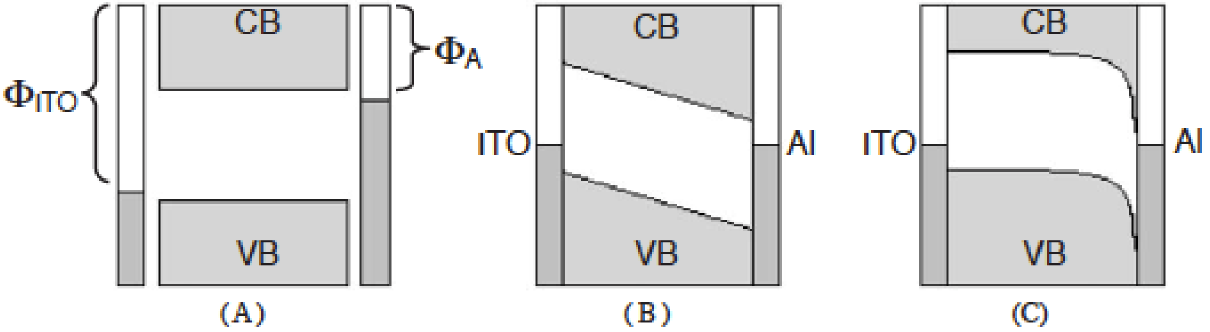


Fig. 6.4. The relative energy levels of the electrodes, CB and VB are shown in three situations, with no external bias. (A) CB and VB are shown along with the low-workfunction electrode (Al) and the high workfunction electrode (ITO) when isolated from each other. (B and C) The cell is assembled and short circuited, causing alignment of the electrode potentials. In (B), an insulating organic material is used. In C a hole-conducting polymer is used forming a Schottky junction at the high workfunction electrode.

The distance over which the CB and VB exhibit curvature is called the depletion width. In B the depletion width extends throughout the material. In C the depletion width is less than half the material thickness. Under external bias the relative electrode potentials can be changed, depending on the size and direction (forward or reverse) of the bias.

6.3. Comparison of inorganic and organic PV

In a crystalline inorganic semiconductor with a 3D crystal lattice the individual LUMOs and HOMOs form a CB and a VB throughout the material. This is fundamentally different from most organic dye semiconductors where the intermolecular forces are too weak to form 3D crystal lattices. Consequently the molecular LUMOs and HOMOs do not interact strongly enough to form a CB and VB. Thus charge transport proceeds by hopping between localised states, rather than transport within a band. This means that charge carrier mobility in organic and polymeric semiconductors are generally low compared to inorganic semiconductors. Also, charge separation is more difficult in organic semiconductors due to the low dielectric constant. In many inorganic semiconductors photon absorption produces a free electron and a hole (sometimes called charge carriers), whereas the excited electron is bound to the hole (at room temperature) in organic semiconductors.

Conjugated polymers lie somewhere between the inorganic semiconductors and organic dyes. In general, excitons are considered to be localised on specific chain segments. However, there are cases where excitons seem to be delocalised. In these cases, the excitons are referred to as polarons. In simple PV devices and diodes based on organic semiconductors the primary exciton dissociation site is at the electrode interface (other sites include defects in the crystal lattice, absorbedoxygen or impurities) . This limits the effective light harvesting thickness of the device, since excitons formed in the middle of the organic layer never reaches the electrode interface if the layer is too thick. Rather they recombine as described above. Typical exciton diffusion distances are on the order of 10 nm.

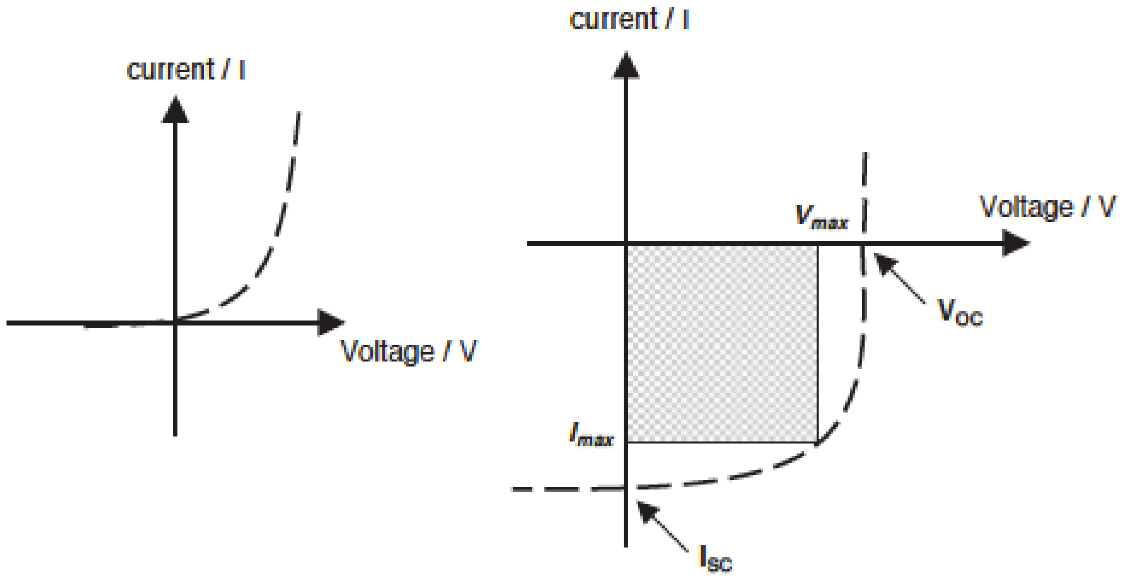


Fig. 6.5. I–V curves of an organic PV cell under dark (left) and illuminated (right) conditions. The open circuit voltage (V_{OC}) andthe short-circuit current (I_{SC}) are shown. The maximum output is given by the square ImaxVmax.

6.4. Characteristics of solar cells

In Fig. 6.5 below some of the characteristics of a typical organic PV device is outlined. To the left, the I– V curve of the cell in the dark is shown. When a cell illuminated, the I– V curve is shiftedd own by the short-circuit current, I_{SC} (Fig. 6.5, right). The open-circuit voltage, V_{OC}, is the maximum voltage difference attainable between the two electrodes, typically around 0.5–1.5V, which is higher than inorganic cells. The cell is placed in an open circuit and illuminated. Electrons and holes separate and flow towards the low- and high–work functions, respectively. At some point the charge build-up will reach a maximum equal to the V_{OC} that is limited by the difference in workfunctions of the two electrodes Rrecently Brabec et al. [18] have shown that in some heterojunction cells V_{OC} is more dependent on acceptor strength than electrode material. The maximum current that can run through the cell is determined by the short-circuit current, I_{SC}. This quantity is determined by connecting the two electrodes, whereby the potential across the cell is set to zero, and then illuminating the cell while the current flow is measured. I_{SC} yields information about the charge separation and transport efficiency in the cell. The magnitude of I_{SC} depends on the illumination strength, but for intensities around 100 mW/cm^2 (AM1.5) I_{SC} is in the 0.20–80 mA/cm^2 range. The square I_{max} V_{max} is the maximum work the cell is able to yield. The fill-factor FF is given by I_{max} V_{max}/ V_{OC} I_{SC}, and is typically around 0.4–0.6. The quantum efficiency (QE) is the number of generated electrons per absorbed photon. For single layered organic PV the QE is typically in the order of 1%, while inorganic PV often reaches a QE of 80–90%. The power conversion efficiency (PCE) is the power output divided by the incident light power. For single layered organic PV this is typically below 0.1%, while crystalline Si cells have up to 24%.

6.5. Chemistry and physics at the electrodes

When a pristine conjugated polymer is brought into contact with electropositive metals typically used as electrodes (i.e. Al, Ca, Mg, Mg/In), the interface is never sharp. Rather, chemistry always occurs to a varying extent. In the case of poly(pphenylene vinylene) (PPV) on aluminium XPS studies reveals the formation of a 30 ^{o}A thick insulating layer. This layer is thought to be formed as Al atoms diffuse into the polymer matrix where Al reacts with the vinyl groups and disrupt the conjugation (see Fig. 6.6). The effect of this layer is to increase the electron injection barrier at the interface. Naturally, as the layer becomes thicker and thicker electron extraction becomes impossible rendering the device useless. Different strategies have been investigated to overcome this problem. In LEDs a thin protective layer between an Al-electrode and the organic layer has been found effective and increasing overall performance. Hung et al. [20] demonstrated that a 5–10 ^{o}A thin LiF or MgO layer improved efficiency of the Al-electrode by lowering the electron-injection barrier. Greczynski et al. [21] has also investigated the protective properties of LiF in a thorough study done on polyfluorene using UPS and XPS. Both studies indicate that LiF does not dissociate and react chemically, but rather serves as a protecting layer between the electrode and the organic material. Also, Al_2O_3 has been shown to have favourable properties as a protective layer reducing drive voltage and increasing device performance in LEDs

[22]. Very recently Brabec et al. [23] have shown that the advantages of using a protective layer on the negative metal electrode carry well over in PV devices. Indeed LiF, has been reported to increase the FF and stabilise the V_{OC} in PV cells.

Transparent ITO is the material most used as the high-workfunction material in PV devices and LEDs. The ITO–polymer interface is not well understood or well controlled. There are large variations of the ITO morphology and workfunction from manufacturer to manufacturer and from batch-to-batch [24]. The uneven surface of ITO on glass results in local areas with high fields under operation that may cause rapid polymer degradation. AFM and UPS studies have shown that acid etching and ozone cleaning can be used to control the surface [25]. Similar to the cathode interface, atoms from the anode can react with the organic material. Thus, in one study it was found that oxygen could diffuse into MEH–PPV forming aromatic aldehydes [26]. Also, indium was found to diffuse into the organic layer where it acts as trapping site for charge carriers [27]. One strategy used to minimize indium and oxygen diffusion is to place an interfacial hole-transporting layer, such as PEDOT-PSS [28] , between ITO and the active material. This layer also serves to smooth out the uneven surface of ITO.

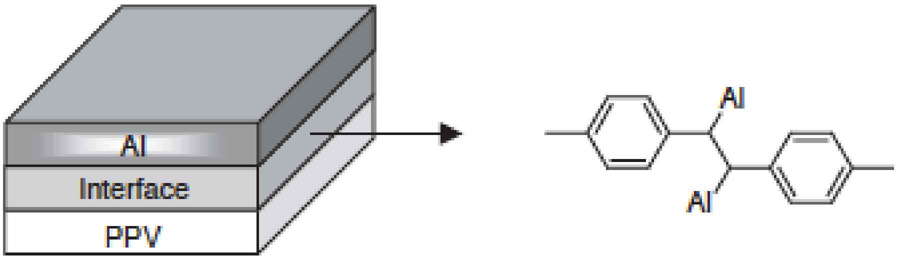


Fig. 6.6. Aluminium andthe pristine polymer will chemically react and form an insulating layer at the interface.

6.6. Dye solar cells

Photoconductivity was first observed on anthracene (see Fig. 6.7) in the beginning of the 20th century [3,4]. From the 1950s anthracene was intensively studied, partly because the crystal structure was accurately determined early [29], and high-purity single crystals was readily available [30]. The first real PV investigations were done on porphyrins and PCS, and this class of compounds has remained among the most investigated dyes. The compounds are easy to prepare, highly coloured, and they form crystalline films by vacuum sublimation with good semiconducting properties. Also, they readily form complexes with a number of metal ions. In 1958 Kearns and Calvin worked with magnesium phthalocyanines (MgPc) between two glass electrodes and measured a photovoltage of 200 mV [31]. Six years later in 1964 Delacote et al. [32] observed a rectifying effect when copper phthalocyanines were placed between two different metals. In 1971 Federov and Benderskii reinvestigated MgPc. They found that the PV properties of MgPc are very dependent on exposure to oxygen [33]. Five Years earlier Kearns et al. [34] hadsuggested that oxygen at the crystal surface of anthracene assisted in exciton dissociation and thus played an important role in the photoconductivity of anthracene.

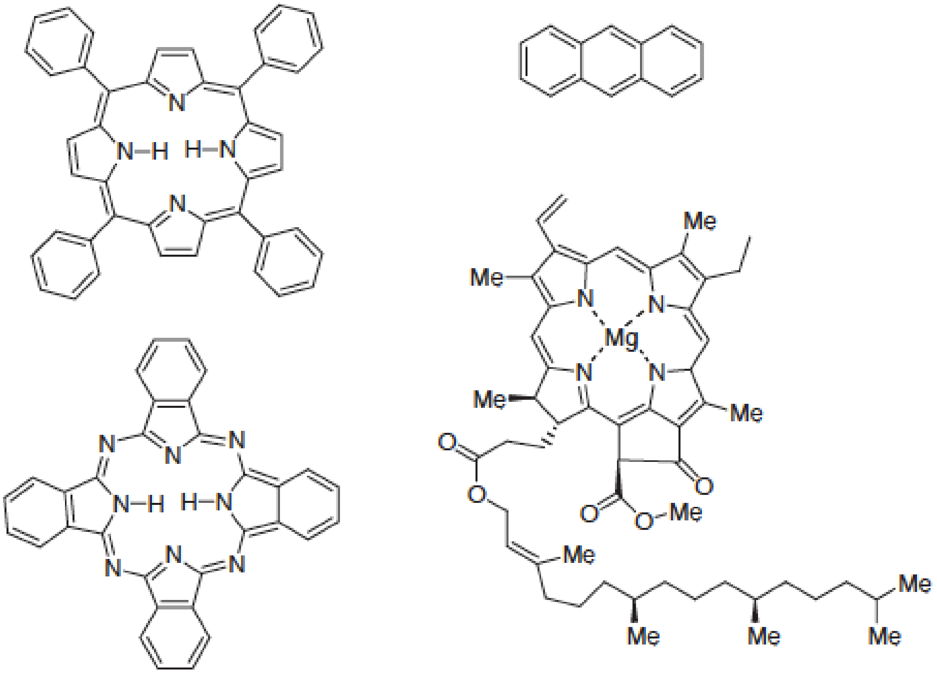


Fig. 6.7 Some of the early investigatedorganic molecules. Top: TPP andanthracene. Bottom: phtalocyanine and Chl- a.

The pronounced effect of oxygen is now known to be related to the electron accepting ability of molecular oxygen. Meilinov et al. [35] worked on chlorophyll–a (Chl–a) in 1970 and measured a photocurrent quantum yield of 10% in Al/Chl–a /Al, which was substantially better than previous reported. In 1975 Tang and Albrecht [35] worked extensively with Chla (extracted from green spinach) in metal1/Chl-a /metal2 cells. By varying the electrode metals they found the optimal cell to be Cr/Chl-a /Hg that had a PCE of 0.01% under monochromatic illumination, orders of magnitude larger than most systems at that time. Earlier work by Putseiko et al. [36] had established that Chl-a was water sensitive, which Ballschmitter et al. [37] found was related to the formation of a Chl–a–H_2O adduct. The adduct formation plays an important role in the formation of ordered microcrystalline domains in Chl-a films. Chlorophyll is still attracting interest, and in 2002 the conducting mechanism of Chl-a in Al/Chl-a /Ag cells was investigated by Mabrouki et al. [38] . The cells had a PCE of 0.1%. A major breakthrough in cell performance came in 1986 when Tang showed that much higher efficiencies are attainable by producing a double-layered cell using two different dyes [39] .

6.7. Conjugated polymer cells

One of the most studied photoconducting polymers is poly(vinyl carbazole) (PVK). The first report came in 1958 by Hoegel et al. [41] who proposedits practical use as an electrophotographic agent. In the 1970s it was discovered that certain conjugatedpolymers, notably poly(sulphur nitride) andpolyacetylene (see Fig. 6.8), could be made highly conducting in the presence of certain dopants [42]. In 1982 Weinberger et al. [43] investigated polyacetylene as the active material in an Al/ polyacetylene/graphite cell. The cell had a low open-circuit voltage of only 0.3V and a low QE of only 0.3%. Later Glenis et al. [44] investigatedd ifferent polythiophenes. Again the systems sufferedlow efficiencies andlow open-circuit voltages in the 0.4V range. The low open-circuit voltages has been ascribed to the formation of polarons (delocalised excitons) that energetically relax in the energy gap, which then becomes smaller than the $\pi - \pi*$ gap. This relaxation results in a large spectral shift when the luminescence spectra are compared to the absorption spectra (Stokes' shift). The result of the relaxations is that it limits the attainable voltage andthe maximum PCE. Different electrode material have been used but not with success.

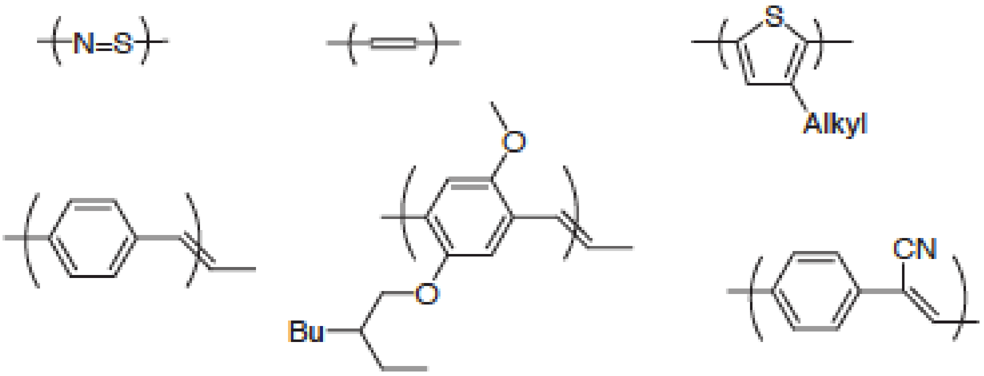


Fig. 6.8. Some conjugatedpolymers investigated in PV cells. Top: poly(sulpher nitride) (SNx,) polyacetylene andpoly(3-alkyl-th iophene). Bottom: poly(p-phenylene vinylene) (PPV), poly(2-methoxy-5-(20-ethylhexyloxy)- 1,4-phenylvinylene) (MEH–PPV), andcyano-PPV (CN-PPV).

Followed by the poly(alkyl-thiophenes) (PATs), PPV and its derivatives is the most investigated conjugated polymer in PV cells. Unlike polyacetylene and polythiophene there is only limited energy relaxation. Karg et al. [45] was the first to investigate PPV in ITO/PPV/Al LEDs and PV devices in 1993. Karg measured VOC of 1V and a PCE of 0.1% under white light illumination. Interestingly enough they had different views on the depletion width in the cells. Marks found that their cells were completely depleted while Antoniadis' cells formed Schottky type barriers at the Al-interface. This divergent behaviour is probably related to the fact that PPV is very sensitive to atmospheric oxygen as an efficient dopant for PPV. Thus in the presence of oxygen, electron abstraction from PPV increases the conductance orders of magnitude, due to the p-type behaviour of PPV. Other investigated PPV derivatives include MEH–PPV [48] andpoly(2,5- diheptyloxy-p phenylenevinylene) (HO–PPV) [49] .

6.8. The concept of a heterojunction

As previously described, excitons do not dissociate readily in most organic semiconductors. The idea behind a heterojunction is to use two materials with different electron affinities and ionisation potentials. This will favour exciton dissociation: the electron will be accepted by the material with the larger electron affinity and the hole by the material with the lower ionisation potential. In the 1950s it was discovered that organic dyes adsorbed on the surface of inorganic semiconductors had an effect on the photoresponse of the semiconductor in the spectral range of the dye. In 1956, for example, Nelson [50] found that the photoconductivity of CdS was increased when sensitised with a cyanine dye in the redspectral range. Nelson argued that the LUMO of the dye was above the CB of the semiconductor. When the dye was excited, the electron in the LUMO of the dye would be transferred to the CB of CdS and thereby enhance the conductivity.

In 1979 Tang [39] filed a patent on his ability to increase PCE to 1% of a bi-layer PV device consisting of copper PC and a perylene derivative. Tang proposed that the observed synergistic effect of bringing two different semiconductors in contact was caused by the field at the heterojunction interface. This local field aids dissociation of excitons diffusing to the interface. Such a situation is depicted in Fig. 6.9 where the donor molecule is excited (upward arrow); the electron is promoted from HOMO to LUMO leaving a hole behind. The electron and hole can recombine (downward arrow), yielding e.g. luminescence, or they can dissociate. If the acceptor LUMO is sufficiently lower than the donor LUMO, the excited electron will relax into the acceptor LUMO and in this way separate from the absorption spectrum, indicating that light has to reach the PPV/C_{60} interface in order to generate a photocurrent. From models, an average exciton diffusion length of 6–8 nm could be derived.

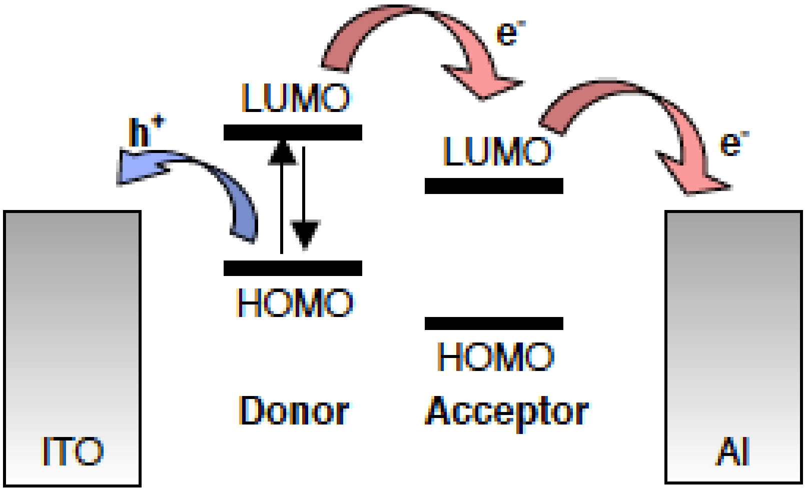


Fig. 6.9. Exciton dissociation at the donor–acceptor interface. The electron goes to the acceptor while the hole stays on the donor.

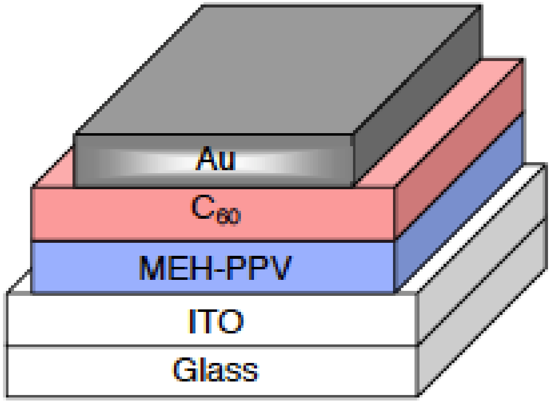


Fig. 6.10. A two-layer heterojunction photovoltaic cell. The electron accepting C_{60}-layer contacts the Au electrode, while the electron donating MEH-PPV layer contacts the ITO electrode

In 1997 Halls [55] also investigated heterojunctions where perylene was the electron acceptor. Bis(phenethylimido)perylene was vacuum sublimed onto PPV films spin coated on ITO covered glass, and Al was used as the counter electrode material. These cells had QE of 6% and PCE of 1% (monochromatic illumination). A high FF value of 0.6 was reported. It became clear that if absorption at the heterojunction interface were increased the photocurrent from the cell wouldincrease as well. The optical field varies through the cell, the incoming light interfere with reflections from the metal surface and the fullerene/polymer interface resulting in a standing wave. In the optimal situation the standing wave will have a node at the heterojunction interface. In 1999 Pettersson et al. [56] calculated the optical field intensity through a heterojunction multilayer cell using complex indices of refraction, determined using spectroscopic ellipsometry. The studied cell consisted of PEOPT (poly(3-(40’ -(100” ,400” ,700” -trioxaoctyl)-phenyl)thiophene)) and C_{60}. Al was used as the electron-collecting electrode, while ITO was the hole-collecting electrode. The ITO electrode was covered with a PEDOT (poly(3,4-ethylenedioxythiophene))–PSS (poly(styrenesulfonate) layer to improve hole transport. By varying the polymer- and C_{60} -thickness the nodes of the standing wave can be moved back and forth through the cell. Pettersson found that the exciton-diffusion range was around 5 nm in PEOPT and around 8 nm in the fullerene domain, thus the node of the optical field becomes very critical. When an 80nm thick fullerene layer was used the optical field had an antinode at the PEOPT/C_{60} interface, but when the fullerene layer was reduced to 35nm the optical field had a node at the interface. Applying this result, Pettersson constructed a cell with a PEOPT thickness of 40nm and a fullerene thickness of 31nm and achieved a peak QE of 21% (using monochromatic 440 nm light), which indicates efficient charge separation.

Very recently in 2003, Durstock et al. [57] use delectrostatic self-assembly to control the layer thickness in PPV/C_{60} heterojunction cells. The thickness of the fullerene layer could be controlled by alternating immersion of the substrate into baths of water-soluble cationic andanionic C_{60} derivatives. In the same way, the PPV layer thickness was controlled by using a PPV precursor as the cation and sulfonated polystyrene (SPS) as the anion. In a subsequent step the films were heated to convert the PPV precursor into PPV. Aluminium and ITO was used as electrode materials. Besides controlling the polymer and fullerene layer, a thin interfacial fullerene/PPV layer was inserted in some of the cells (Fig. 6.11).

For a fixed SPS/PPV layer thickness of z~ 20, varying the fullerene layer thickness gave an optimum cell performance at x~ 50. If x was held at 50, peak performance was around z~ 20. If an interfacial layer was inserted the cell efficiency could be further improved. From this, Durstock concluded that the interfacial layer increases the effective donor–acceptor interface leading to in creased charge separation, however for increasing interfacial layer thickness the distance to the electron- and hole-conduction layer becomes too large to utilise the increased exciton dissociation. This study demonstrates the importance in controlling the interfaces in heterojunction cells.

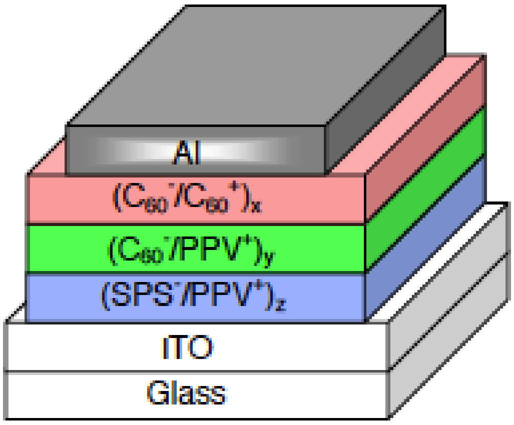


Fig. 10. Electrostatically self-assembled PV cells produced by Durstock. Different layer thicknesses were usedin different cells (x=0–100, y=0–10 and z=0–40).

Finally polymer/polymer donor–acceptor heterojunction should be mentioned. If this type of cell is fabricated by spincoating the challenge is to find a solvent that will not wash away the first layer when the second polymer is spin coated. Since conjugated polymers typically have poor solubility in normal solvents the available donor–acceptor polymer pairs are limited. In 1999 Tada et al. [58] fabricated a heterojunction cell consisting of the acceptor polymer poly(p–pyridylvinylene) (PpyV) andthe donor polymer poly(3-hexylthiophene) (P3HT). PPyV was spincoated from formic acid and P3HT could then be spincoated from chloroform without damaging the first layer. The photocurrent increased some three orders of magnitude when the donor layer was present, indicating efficient charge transfer at the interface between the layers.

6.9. Dispersed (or bulk) heterojunction

It is clear that exciton dissociation is most effective at the interface in heterojunction cells, thus the exciton should be formedwithin the diffusion length of the interface. Since typically diffusion lengths are in the range of 10 nm, this limits the effective light-harvesting layer. However, for most organic semiconductors the film thickness shouldbe more than 100nm in order to absorb most of the light. It follows that thicker film layers increase light absorption but only a small fraction of the excitons will reach the interface and dissociate. This problem can be overcome by blending donor and acceptor, a concept called dispersed (or bulk) heterojunction (see Fig. 6.12).

One of the inherent problems with dispersed heterojunction is that of solid-state miscibility. Large extended conjugated systems are normally not miscible; this goes for dyes and particularly for conjugated polymers. Thus cell fabrication should be carried out so equilibrium is never reached, e.g. by spin coating where solvent evaporation is fast.

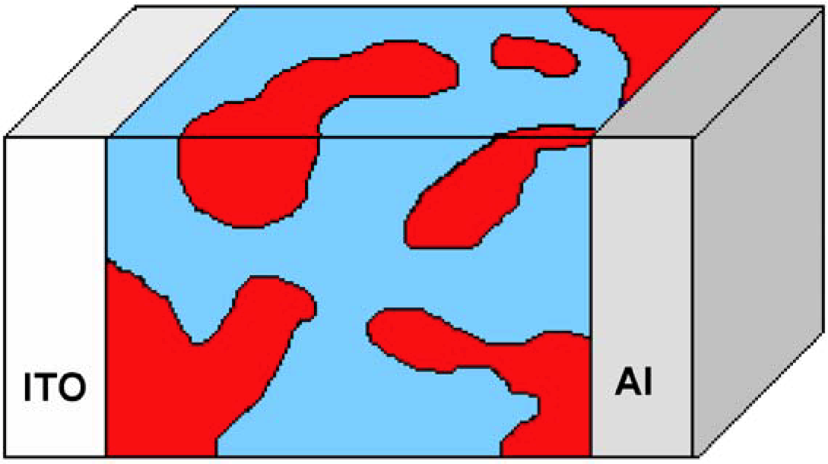


Fig. 6.12. Dispersed heterojunction between a transparent ITO electrode and an Al electrode

Hiramoto [59,60] reported the first dye/dye-dispersed heterojunction in 1991. Working with the same class of p- andn-type organic semiconductors as Tang, metal-free phthalocyanine (H_2 Pc) anda perylene tetracarboxylic derivative, Hiramoto investigated the effect of introducing a third organic layer in the middle of a two-layer heterojunction PV. This middle layer was a mixture of the two dyes made by cosublimation from different thermal sources. Hiramato found that by introducing a third layer the photocurrent was doubled as compared o the two-layered cell.

OPVs go by different names, including polymer solar cells, and represent the other big fields into which alternative solar cell chemistries are developing. In their modern aspect, OPV cells consist of an organic bulk heterojunction (BHJ) sandwiched between a TCO and a semitransparent metal electrode; the BHJ is made of two interpenetrating block copolymers with different electronic properties, namely, with a p-type (or donor) and an n-type (or acceptor) behavior. The interface between the two polymers is where excitonic hole/electron pairs are generated, and dissociation occurs once they reach the metal interface (Fig. 6.13, 6.14).

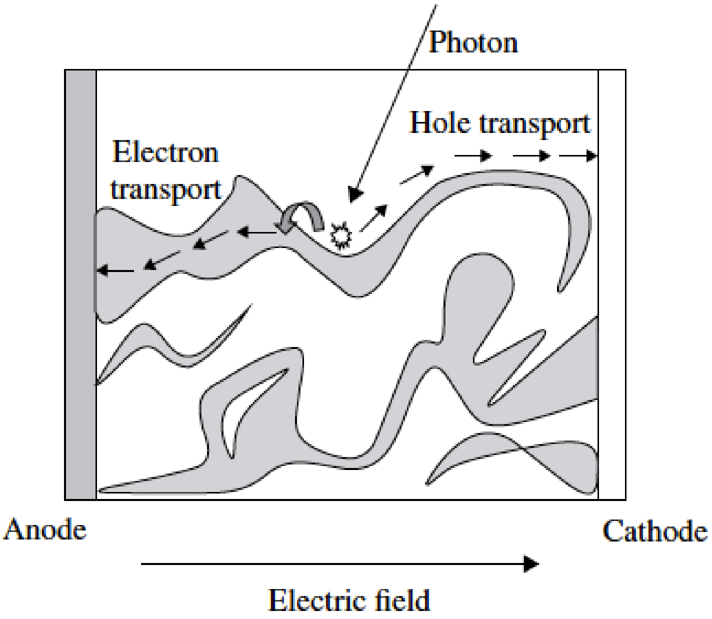


Figure 6.13 Schematic of a bulk heterojunction solar cell.

Although lower PCEs have been attained so far, as compared to DSSCs, a lot of attention has been devoted to this technology, since it will pave the way for easy-to manufacture, cheap, and flexible electronic devices. OPV cells can be fabricated via a wet processing, which does not require the high temperatures generally needed to obtain high-quality silicon, leading to a reduction in energy use as well as energy payback time. Furthermore, it has been forecast that OPV cells in a tandem configuration (i.e., where different materials are stacked on top of each other to absorb different parts of the spectrum) might drive cell efficiencies up to 15%, thus turning polymer solar cells into a clean and sustainable alternative to inorganic PV, even for large-scale energy production [119]. As of April 2012, the world record for commercial OPV cells has already reached 12% [120].

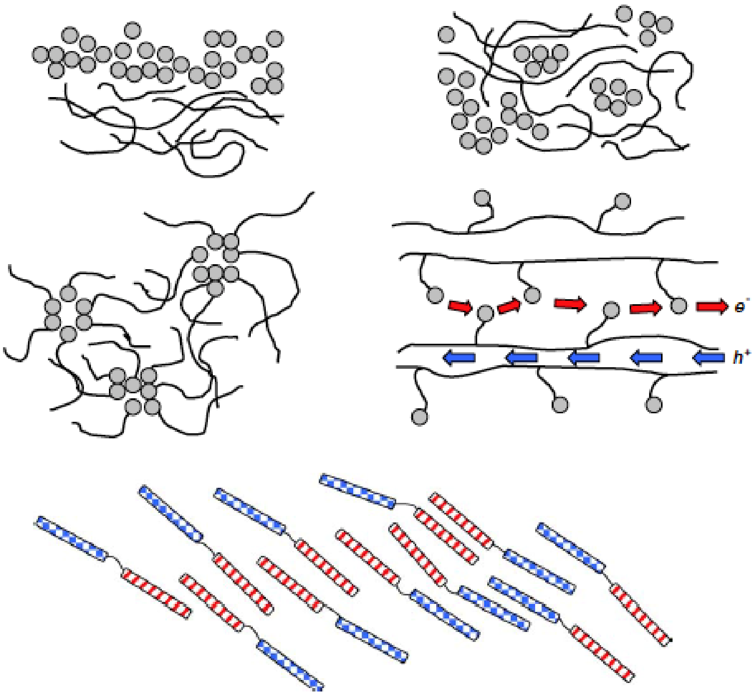


Fig. 6.14. Different morphologies of heterojunction cells. Top, left: Two-layeredstructure of fullerenes and polymer chains. Top, right: dispersed heterojunction. Middle, left: fullerenes with polymer chains attached.Middle, right: self-assembled layered structure of double-cable polymers. Bottom: self-assembled layered structure of diblock copolymers. The layered structure of double-cable polymers and diblock copolymers are expectedto facilitate efficient electron and hole transport.

While some control is introduced by covalently linking the donor and acceptor in polymer/oligomer-C_{60} assemblies the final morphology may suffer from phase separation and clustering of the fullerene (or dye) units (see Fig. 6.14), which potentially limits efficient charge separation due to low donor/acceptor interfacial area. Also, increasedphase separation may disrupt the continuity of the phases and reduce the charge transport properties of the material, due to inefficient hopping between different domains, reducing overall performance. This may be a critical point, since intramolecular charge recombination might occur at a fast rate.

One way to control a bicontinuous phase separation andinsure a large interfacial area between donor and acceptor is to covalently graft fullerene moieties onto the donor-polymer backbone (Fig. 6.14), so-called double-cable polymers (due to their p/ntype conduction properties). These assemblies have been intensively investigated in recent years as promising components in PV devices [81] , but they are also interesting as components in molecular electronics [82] .

6.10 Dye-sensitized solar cells (DSSCs)

DSSCs are a relatively old technology, whose operating principle was first observed in the nineteenth century; they acquired a renewed interest during the early 1990s, thanks to the work of Gretzel and others [69], who were able to boost their power conversion efficiencies up to 10% and turn them into valuable competitors for conventional solar cells. DSSCs consist of three parts:

1). A semiconductor oxide (TiO_2 , ZnO, SnO_2, Nb_2O_5) deposited onto a transparent conducting oxide (typically fluorine tin oxide, FTO, or indium tin oxide, ITO); the oxide particles are coated with a light-absorbing dye.

2). An organic electrolyte containing a redox couple (I^{3-}/I^-).

3. A counter electrode coated with platinum, where the redox shuttle can beregenerated through a catalytic reaction.

A schematic is shown in Figure 6.15. The principle of operation differs from standard inorganic PV in that no hole/electron pairs are generated by the interaction with photons, but rather excitons, an excited state where an electron and a hole are paired together by a binding energy that is higher than thermal agitation [71].

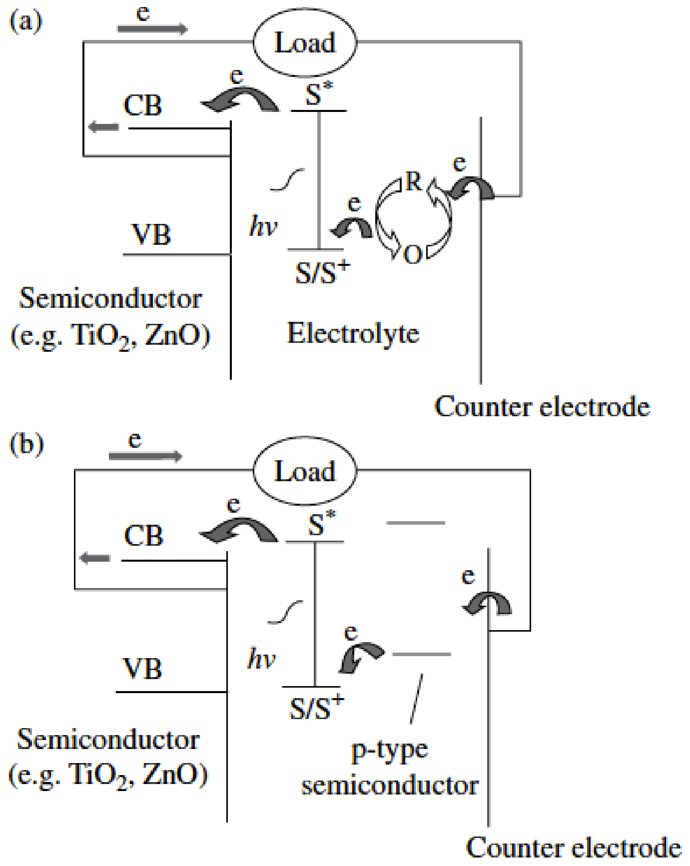


Figure 6.15 General principle of operation of (a) a DSSC with redox couple in the liquid electrolyte. (b) a solid state DSSC with a p-type semiconductor

The reason why this technology had not yielded relatively high efficiencies until Gratzel's contribution is the fact that the oxides used (typically TiO_2 anatase) only absorb a small fraction of the incident light, mostly in the UV spectrum; therefore, a light-sensitive dye coating has to be cast on top of them to ensure sensitivity to a broader spectrum range. A monomolecular layer of cis-RuL2 $(SCN)_2$ dye, a ruthenium complex where L stands for 2,2′ -bipyridyl-4-4′ -dicarboxylate, can however only give a 0.13% incident photon-to-current efficiency (IPCE) at its 530 nm absorption peak when deposited on a flat TiO_2 surface; adding extra layers would further decrease this value by effectively creating a series of filters for light absorption. The use of nanostructured TiO_2 solved this problem by allowing the adsorption of larger amounts of dye and hence overcoming the intrinsic limitations of a flat, unstructured morphology. This resulted in a 600-fold increase in IPCE and paved the way to an organic-based solar harvesting alternative.

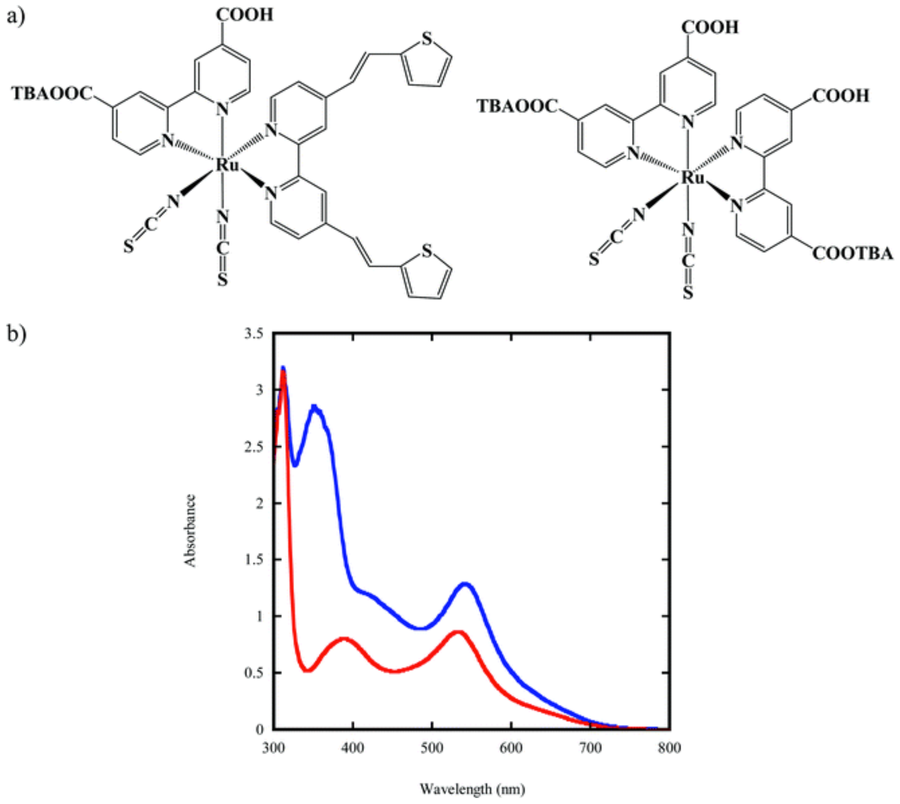


Fig.6.16 Absorption of ruthenium complex

6.11 Graphene and DSSCs

6.11.1 Counter Electrode.

One of the applications of graphene in DSSCs has mainly been devoted toward the replacement of the Pt electrocatalyst at the counter electrode. Platinum is generally added in relatively small amount (<0.1 g m^{-2}), but its high cost and possible side reaction with the iodide/triiodide redox shuttle have driven researchers toward cheaper and more reliable alternatives. Among the first who explored carbonaceous materials are Kay and Gratzel themselves, who added around 20% carbon black to a graphite dispersion and produced a counter electrode for a low-cost solar cell [75]. Later, they also tried grinding together carbon black and TiO_2 to get a 15 μm thick carbon film, which could deliver an outstanding PCE of 9.1% [76]. These promising results led to further efforts from other groups (including ours) who went on to compare the PCEs obtained with graphite, activated carbon, and single- and multiwalled carbon nanotubes (MWCNTs) [77, 78]; the highest efficiency (7.7% under AM 1.5) was obtained by Lee et al., by using defect rich multiwalled carbon nanotubes [79]. More recently, an interesting approach on the production of large-effective-area carbonaceous materials was provided by Lee and coworkers [80]. Large-effective-surface-area polyaromatic hydrocarbons (LPAH) were produced via a hydrogen arc discharge and then assembled with a graphite film to give an all-carbon counter electrode. The average surface roughness of this material was 37.4 nm and contributed to the production of a DSSC with a remarkable PCE of 8.63% and a FF of 80%. Hung et al. [81] obtained graphene oxide (GO) pastes via freeze-drying, which increased the material's porosity, advisable for increased wettability and thus higher number of reduction sites for I^{3-}. This is confirmed by the attained PCE of 6.21%, compared to 5.62 of a blank DSSC. GO was also produced via the Hummer's and the Staudenmaier method [82, 83], but despite obtaining relatively small flakes (down to 5 and 19 nm, respectively), results from DSSC tests were not encouraging (Fig. 6.17).

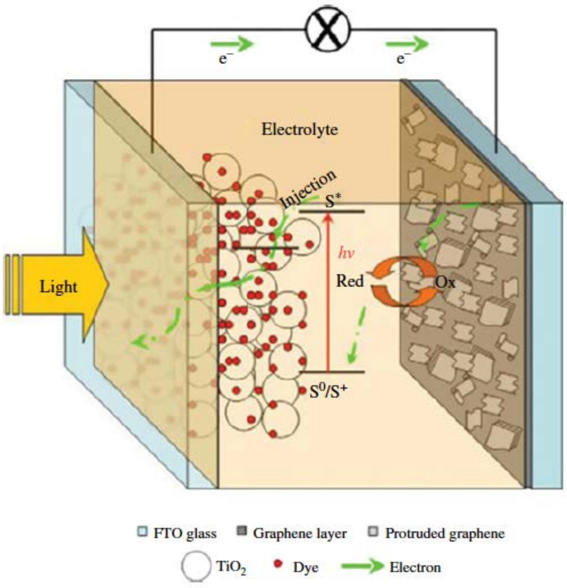


Figure 6.17 Schematic diagram of GN-based DSSC.

6.11.2 Photoanode

As it can be seen, most of the efforts of incorporating grapheme into DSSCs have been devoted to the replacement of expensive platinum and/or transparent conducting oxides (TCO) on the counter electrode. However, graphene's unique versatility has also shown that it can be a good candidate in photoanodes to replace conventional semiconductor oxides like TiO_2 , ZnO, SnO_2 , $Nb_2 O_5$, and so on.

Already in 2008, Hamann et al. had foreseen graphene's advent without directly mentioning it: "an ideal photoanode should be highly transparent, have high surface area and porosity, be easy to fabricate and exhibit fast electron transport [64]." TiO_2 is cheap, naturally abundant, and environmental friendly, but it has a low electron diffusion coefficient and limits the choice of new dyes and redox couples that can be used. On the other hand, it has been proven that when added to TiO2 , graphene can enhance the photocatalytic activity of the latter [103, 104], because of the highly porous anodes that can be fabricated and onto which larger amounts of dye can be chemisorbed. Moreover, graphene is a zero band gap material, as opposed to TiO2 (3.2 eV), and this property can be exploited to prevent charge recombination during the electron transfer process. Nair et al. have created TiO2 –graphene composites via electrospinning, which yielded 150 nm thick fibers [105]. By adding 0.7 wt% of graphene, they were able to increase the short-circuit current of a DSSC device from 13.9 to 16.2 mA cm^{-2} , thereby obtaining a higher PCE than a TiO2 - only control cell (7.6 vs. 6.3%). The synergic contribute of graphene–TiO2 composites with as little as 0.5 wt% graphene content has also been proven via alternative syntheses, such as ball milling [106], heterogeneous coagulation from Nafion [107], spraying [108], and simultaneous reduction hydrolysis of GO with a Ti precursor [109], good for preventing the collapse and restacking of the sheets, even with a graphene content as high as 5 wt%. In all cases, graphene was able to improve the cell performance. The size of TiO2 particles attached to the graphene sheets was tuned by He et al. [110] who could produce nanometer-sized spherical particles as well as TiO2 nanorods and graft them onto graphene via a solvothermal approach. The device containing ultrasmall TiO2 particles (2 nm) showed the best performance with a PCE of 7.25%. Among alternative methods to produce multicomposite graphene photoanodes, the work by Jung et al. shows an interesting approach [111]: graphene-wrapped alumina particles were prepared by the coreduction of CO gas and AlN and later mixed with TiO2 ; this composite, which only contained 1 wt% graphene, showed an 11% increase in performance as compared to a graphene-free device. Ma and coworkers tried to reduce charge recombination at the photoanode by grafting acid-treated MWCNTs onto graphene [112]. The obtained TiO2 -nanostructured hybrid paste, with a 0.07 wt% content of MWCNT and 0.03 wt% graphene, showed a charge transfer resistance of 13 Ω cm2 as compared to a TiO2 -only electrode (25.34 Ω cm2); as expected, an improvement in PCE was attained, not only due to a lower charge recombination but also to a higher degree of dye adsorption.

6.11.3 Prperties of Graphene Electrode

The high electrical conductivity and low optical absorption (2.3% absorption for one-layer graphene) make graphene an excellent transparent electrode. Other advantages of graphene as electrode include good surface smoothness, chemical stability, and flexibility. Peumans et al. [25] have predicted that the sheet resistance (Rs) of graphene will vary with the number of layers as Rs~62.4/N Ω/sq for highly doped graphene, where N is the number of layers. Kim et al. [26] have demonstrated the preparation of large-area graphene films with a sheet resistance of 30 Ω/sq and a transmittance of 90% at 550 nm. This performance is fairly comparable to that of ITO and is sufficient for the transparent electrode application in small-area PSCs. Among various methods for preparing graphene, including micromechanical exfoliation, epitaxial growth, chemical vapor deposition (CVD), and reduction of graphene oxide (GO), the CVD approach can produce high-quality graphene sheets with a large size, low defect content, and high conductivity [27]. The GO-reducing approach suffers from high defect content and low conductivity but has advantages of high-throughput preparation, low cost, and simple film/device fabrication [28].

Although significant progress has been achieved, the development of highperformance PSCs based on graphene electrodes is still hampered by two factors:

(i) the low conductivity reported for many graphene films due to the nonoptimal fabrication technique of graphene electrode and

(ii) (ii) the difficulty of coating HEL. The hydrophobicity of graphene prevents the uniform coating of hydrophilic HEL on the graphene electrode.

Geng et al. [29] had used reduced graphene oxide (rGO) as the transparent electrode to fabricate PSC devices. The rGO was carefully thermal annealed and gave a sheet resistance of 6 Å~ 103 Ω/sq at a transmittance of 78%. Owing to the high sheet resistance, the resulting PSC device showed a PCE of 1.01 Å} 0.05%, which equaled to half of the corresponding value (2.01 Å} 0.1%) for a reference device based on an ITO electrode. Compared to PSCs with an ITO electrode, the rGO-based PSC always shows inferior photovoltaic performance due to the very high sheet resistance of rGO. Highly conductive graphene electrodes can be obtained by synthesis of grapheme via CVD approach, layer-by-layer stacking of graphene sheets, and acid doping, as reported by Kim et al. [26]. Wang et al. [30] used CVD graphene with a sheet resistance of 210 Ω/sq at a transmittance of 72% as the transparent electrode for PSCs. The resulting device with a P3HT:PCBM active layer exhibited a PCE as as 0.21% due to the hydrophobic property of graphene, which prevented the uniform coating of hydrophilic HEL PEDOT:PSS. After the graphene electrode was modified by pyrenebutanoic acid succinimidyl ester to improve the surface wettability for spin coating PEDOT:PSS, the PCE significantly increased up to 1.71% with a V_{OC} of 0.55 V, J_{SC} of 6.05 mA/cm^2 , and FF of 0.51. This work highlighted the importance of interface engineering of graphene electrode for PSC device applications. In a separated but closely related work, Wang et al. [31] made highly conductive grapheme electrode by layer-by-layer transfer method (Fig. 5.3a–c), followed by acid doping.

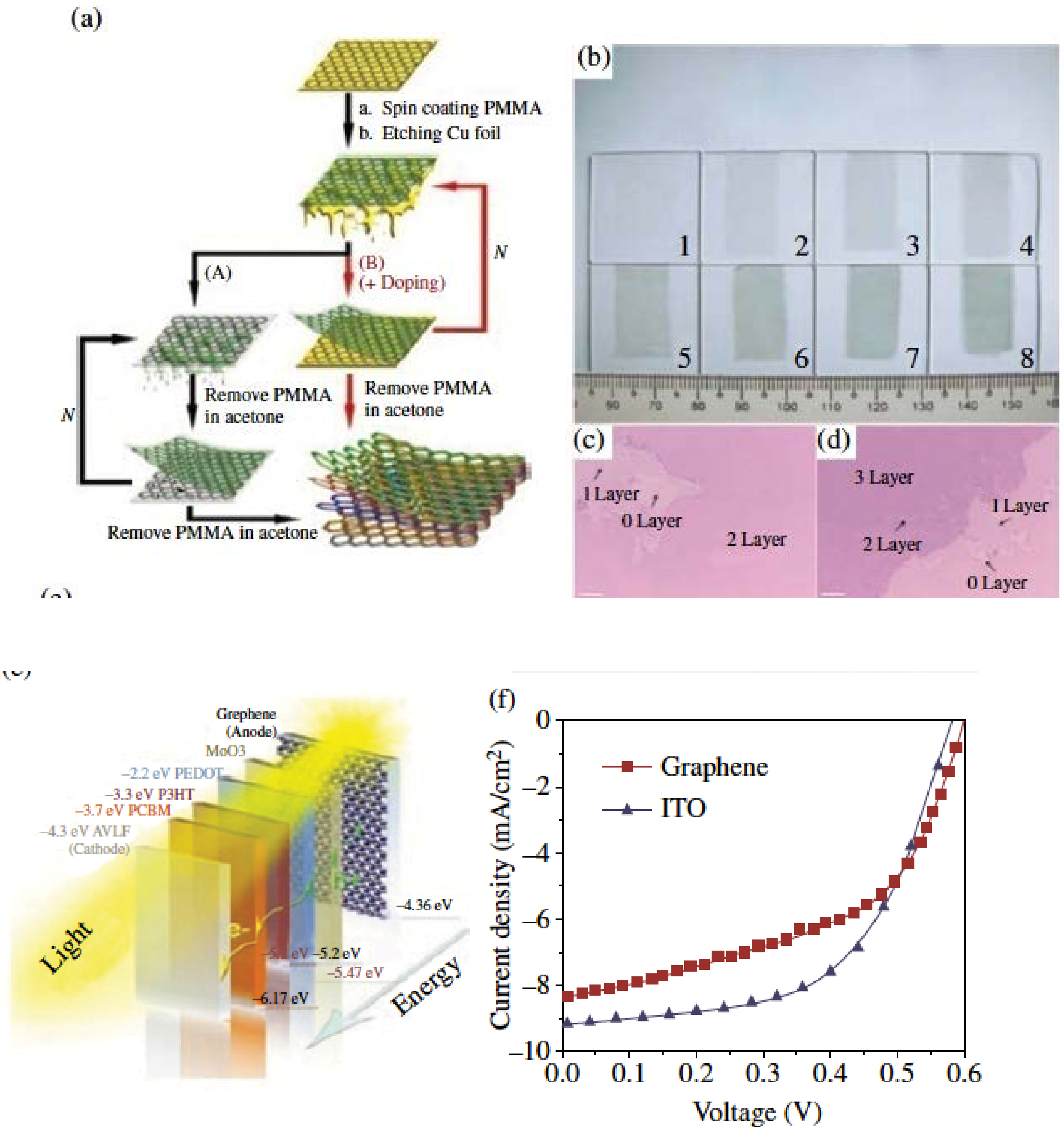


Figure 6.18 (a) Schematic drawing of multilayer graphene films made by normal wet transfer (A) and by LBL assembly (B) (N = 0,1,2,3…). (b) Optical images of multilayer grapheme films (from 1 to 8 layers) on quartz substrates. (c, d) Typical optical microscope images of 2- and 3-layer graphene films on SiO2/Si substrates. (e) Schematic of photovoltaic device structure. (f) Current density–voltage curves of the devices with the anode of ITO or MoO3-coated graphene (device structure: anode/PEDOT:PSS/P3HT:PCBM/LiF/Al)

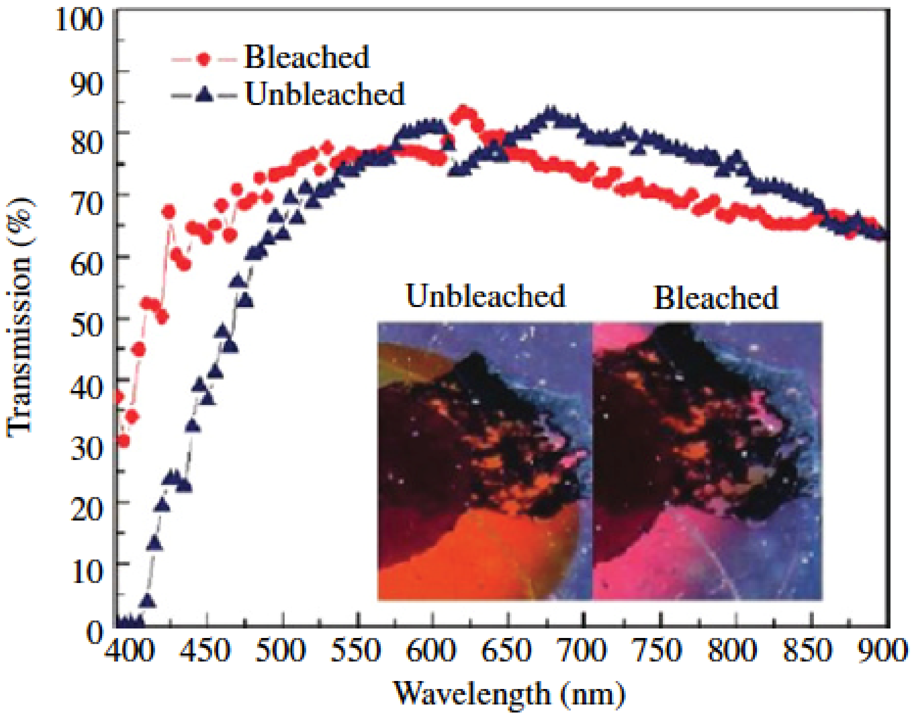


Fig.6.19 Enhanced transparency due to the addition of graphene nanoribbons in the electrolyte.

6.11.5 Electrolyte

A developing area is represented by quasisolid-state electrolytes. Gun’ko et al. incorporated graphene into the ionic liquid 1-methyl-3-propylimidazolium iodide [116]. They found out that a 30 wt% addition of graphene resulted in a quasisolidstate electrolyte that helped deliver a 2.1% PCE without causing an internal short circuit. If 12 wt% graphene and 3 wt% CNTs were incorporated to obtain a hybrid material, this value could be boosted up to 2.5%, with an increase in Jsc from 5.3 to 7.32 mA cm^{-2} but a slight drop in both V_{oc} and FF, down from 0.62 to 0.59 V and from 0.49 to 0.44 respectively. Others incorporated graphene into polyacrylonitrile (PAN) to make a gel polymer electrolyte in much smaller amounts (0.1–1 wt%) [117], and a PCE of 5.41% was obtained, much better than its liquid-state equivalent (3.72%). An interesting finding by Zakhidov et al. [99] (Fig. 6.19) was the bleaching effect of GNR when added to the electrolyte, that is, their ability to increase its transparency. They used GNRs both as a counter electrode and as an electrolyte additive. What was found was that, despite a slight decrease in photocurrent, the overall efficiency was kept constant via a boost in the FF. The unzipping of MWCNTs led to flakes with a GNR morphology, which were then drop-cast onto FTO. The same GNRs were added as a suspension with an estimated concentration of 0.04 mg ml^{-1}, and a 20% boost in PCE was observed with respect to DSSCs made with an unmodified electrolyte.

6.12 Graphene and OPVs

6.12.1 Transparent Conducting Oxide

OPV cells have benefited too from the incorporation of graphene within the cell chemistry. Most of the work has been devoted to its use as a transparent electrode, due to its already mentioned exceptional optical transmittance across a wide range of the visible spectrum. Graphene can provide an economically attractive alternative to ITO while at the same time enabling applications in next-generation IT devices, thanks to its flexibility, which ITO lacks. (Fig. 6.20).

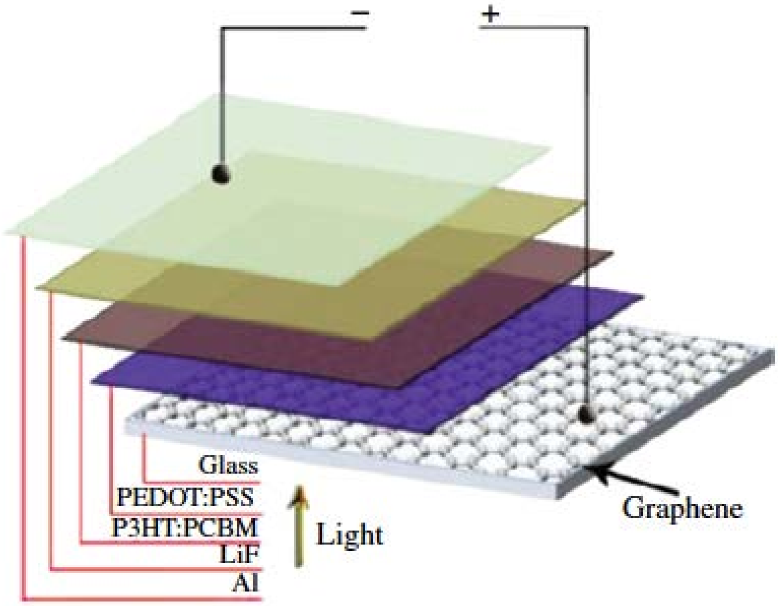


Fig.6.20 Incorporation of a graphene layer within an OPV cell

Among the first reports is the one by Müllen et al. [21], who transposed their previous work on DSSCs onto OPVs, showing that graphene films with an average thickness of 4–30 nm can be obtained from thermal fusion of LPAHs and with transparencies up to 90% at 500 nm that only dip down to 85% when the material is assembled into a functioning device. With a sheet resistance of 1.6 kΩ sq^{-1} and a roughness of 0.4–0.7 nm, the OPV cells fabricated had a 1.53% PCE when exposed to monochromatic light at 510 nm and 0.29% when tested under simulated solar light. Peumans et al. [22] thermally reduced GO and obtained graphene sheets with a thickness less than 20 nm, a transmittance higher than 80% at 550 nm, and a sheet resistance varying between 5 kΩ sq−1 and 1 MΩ sq^{-1} ; transmittance values of 95% and sheet resistances of 100 kΩ sq^{-1} were also obtained with smaller 4–7 nm thick grapheme films. A proof-of-concept OPV cell was assembled and showed a promising PCE of 0.4%, nearly half as much as a reference ITO-based cell. Others have focused on cost reduction by adopting solution-processable graphene [123] or naturally occurring starting products, such as camphor [124]. Significant improvements in performance have rapidly occurred since the early developmental stage, with works like that of Loh [122], where CVD-grown FLG was shown to have similar size and transparency as in previous works (6 nm and 91% at 550 nm, respectively) while lowering the sheet resistance down to 210 Ω sq−1 , thanks to the noncovalent functionalization with pyrenebutanoic acid succinimidyl ester (PBASE). An increase in graphene's work function from 4.2 to 4.7 eV as well as improved contact with PEDOT:PSS led to a PCE of 1.71%. Others have also used fluorine-functionalized graphene [125] or AuCl3− decorated sheets [126].

For interface engineering of the graphene electrodes, Park et al. [32] developed poly(3,4-ethylenedioxythiophene)-block-poly(ethylene glycol) doped with perchlorate (PEDOT:PEG(PC))(Fig. 6.21a). PEDOT:PEG(PC) was deposited on the hydrophobic graphene electrode to facilitate the uniform deposition of PEDOT:PSS (Fig. 6.21a–d). With the interface improvement, the graphene electrode could give a device performance fairly comparable to that of its ITO counterpart in either bilayer organic solar cells or single-layer bulk heterojunction PSCs (Fig. 6.21e).

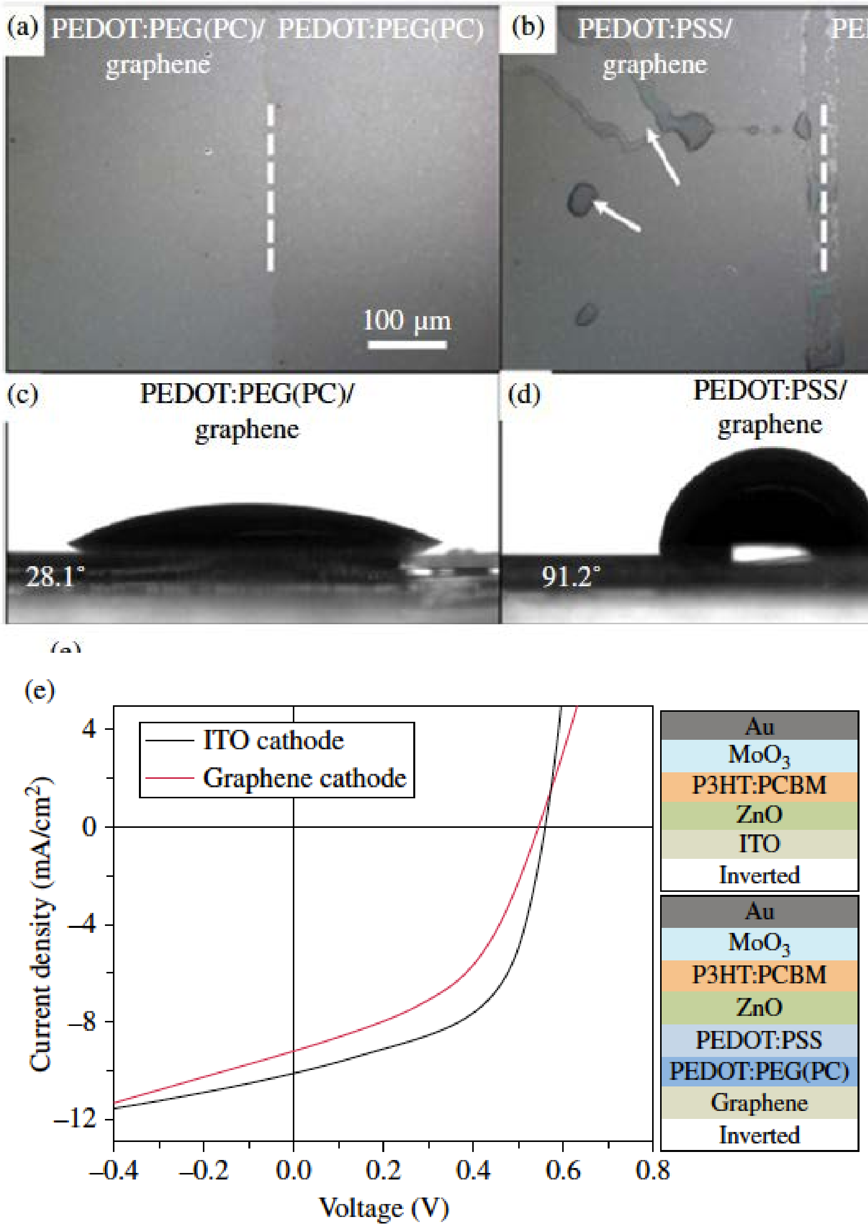


Figure 6.21 Better wettability of PEDOT:PEG(PC) than PEDOT:PSS on the graphene surface.Optical microscopy images of PEDOT:PEG(PC) (a) and PEDOT:PSS (b) on grapheneand bare quartz substrates. The white dotted lines indicate the edge of the graphene, and thearrows denote dewetted PEDOT:PSS. Contact angle images of graphene/PEDOT:PEG(PC)(c) and graphene/PEDOT:PSS (d). (e) Current density–voltage curves for the graphene-based PSC device and the ITO-based PSC device.

Just like composite materials often show properties characteristic of each of the components with a synergistic effect, the combination of CNTs and graphene in transparent electrodes is expected to improve the conductivity. In this context, Tung et al. [33] developed a hybrid nanocomposite film comprised of graphene and CNT (G-CNT) by codissolving GO and CNT in anhydrous hydrazine. The resultant G-CNT film gave a sheet resistance of 240 Ω /sq and a transmittance of 86% (Fig. 6.22a, b). PSC devices with P3HT:PCBM active layer and G-CNT electrode exhibited a V_{OC} of 0.58 V, JSC of 3.47 mA/cm2 , FF of 0.42, and PCE of 0.85% (Fig. 6.22). The low J_{SC} and FF were likely due to the poor contact at the G-CNT/P3HT:PCBM interface.

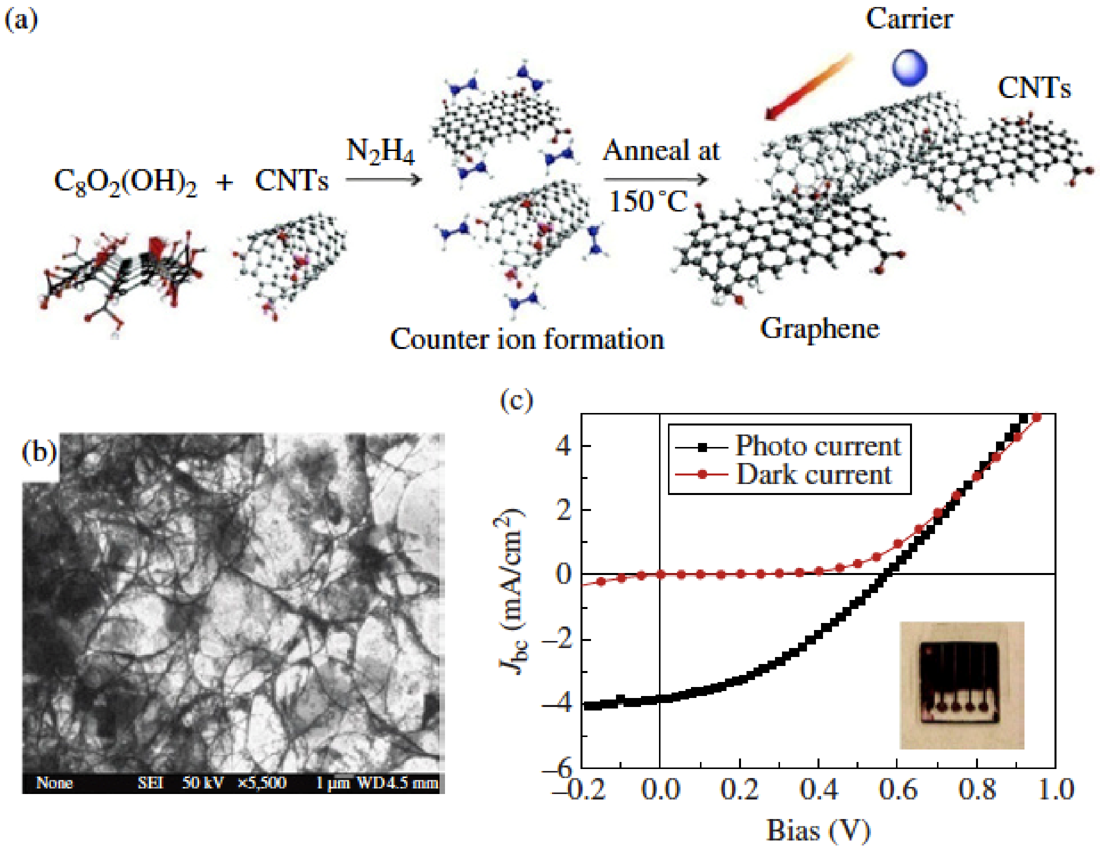


Figure 6.22 (a) Schematic illustration of G-CNT electrode. (b) A representative SEM image of a G-CNT film. (c) Current density–voltage curves of the device with a G-CNT film as the electrode in dark and under AM 1.5G illumination

6.13 Carbon nanomaterials as charge extraction layers

The HEL between the anode and the active layer and the EEL between the cathode and the active layer play critical roles in regulating the overall device performance of PSCs [35]. The function of these charge extraction layers includes (i) to minimize the energy barrier for charge extraction, (ii) to selectively extract one sort of charge carriers and block the opposite charge carriers, (iii) to modify the interface between the electrode and the active layer, and (iv) to act as an optical spacer. An HEL should have a relatively high work function to allow for the built-in electrical field across the active layer and for holes to transport toward the anode.

Similarly, an EEL needs to have a low work function for electrons to efficiently transport to the cathode. Moreover, the hole/electron extraction materials should be solution processable.

The most widely used HEL in PSCs is PEDOT:PSS. However, PEDOT:PSS suffers from its hygroscopicity, which degrades the device efficiency and lifetime. Moreover, PEDOT:PSS itches ITO electrode during longterm device operation due

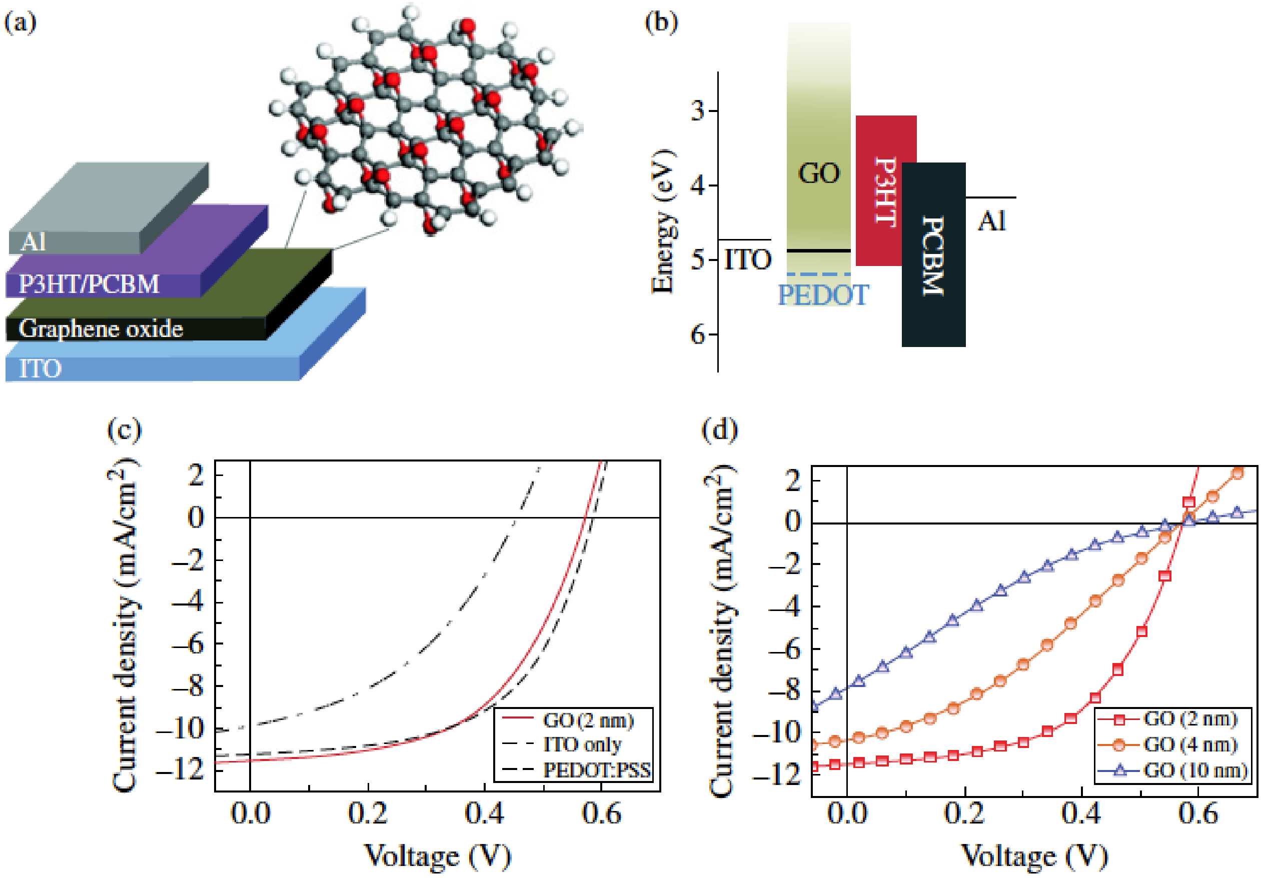


Figure 6.23 (a) Schematic of the PSC device structure with GO as the hole extraction layer (ITO/GO/P3HT:PCBM/Al). (b) Energy level alignment in the PSC devices. (c) Current density–voltage characteristics of PSC devices with no hole extraction layer and with 30 nm PEDOT:PSS layer and 2 nm thick GO film. (d) Current density–voltage characteristics of ITO/GO/P3HT:PCBM/Al devices with different GO layer thickness under simulated AM 1.5

illumination.

to its strong acidity (pH = 1–2). Therefore, certain inorganic oxide semiconductors, such as NiO, MoO_3 , and V_2O_5 , have been recently introduced as HEL for PSCs. The commonly used EEL in PSCs includes, but is not limit to, inorganic salts (e.g., LiF), low work function metals (e.g., Ca), inorganic semiconductors (e.g., TiO_2 and ZnO), and conjugated polymer electrolyte. As we shall see later, GO derivative is the only family that can be used as either HEL or EEL in PSCs.

Li et al. and Gao et al. [36, 37] demonstrated that a thin layer of GO (about 2 nm) could act as an excellent HEL for PSCs to show a fairly comparable device efficiency to that of PEDOT:PSS (Fig. 6.23a–c). GO has a work function of −4.8 eV and band gap of about 3.6 eV and can be spin cast from its aqueous solution into a uniform film. It was found that the FF of the resulting device decreased from 0.53 to 0.19 and the PCE decreased from 3.5% to 0.9% with the GO layer thickness increasing from 2 to 10 nm (Fig. 6.23d). This was due to the insulating property of GO that led to a high series resistance of the resulting device. Murray et al. [38] reported a highly efficient and stable PSC with GO as the HEL and PTB7:PC71 BM as the active layer (Fig. 6.24). The GO‑based device showed a PCE of 7.39%, which was fairly comparable to the corresponding value of 7.46% for a PEDOT:PSS‑based device (Fig. 6.24c). More importantly, the GO‑based device provided a 5 times enhancement in thermal aging lifetime and a 20 times enhancement in humid ambient lifetime compared with the PEDOT:PSS‑based device (Fig. 6.24d, e). The aforementioned results indicated that GO was a promising HEL for efficient and stable PSCs.

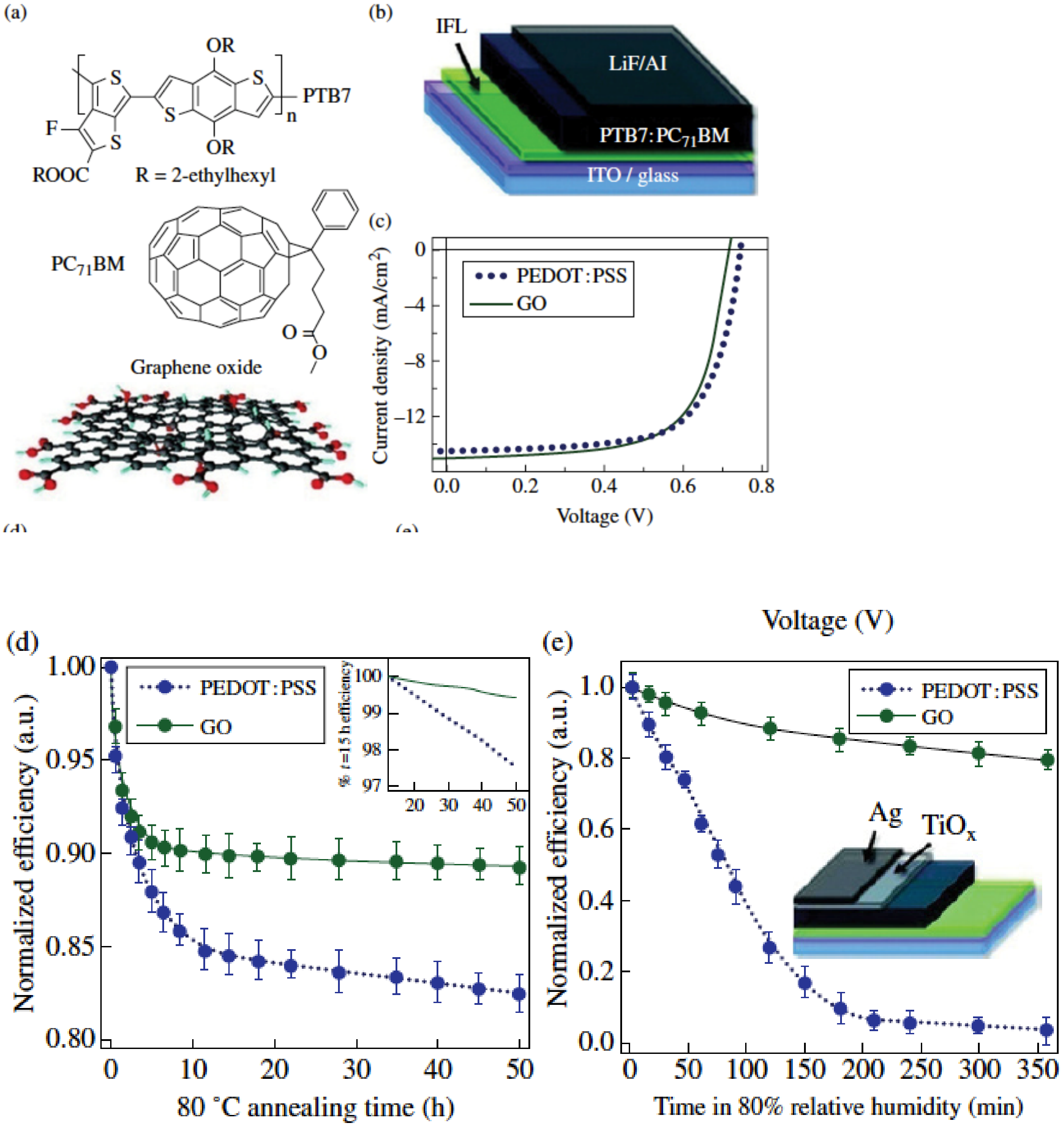
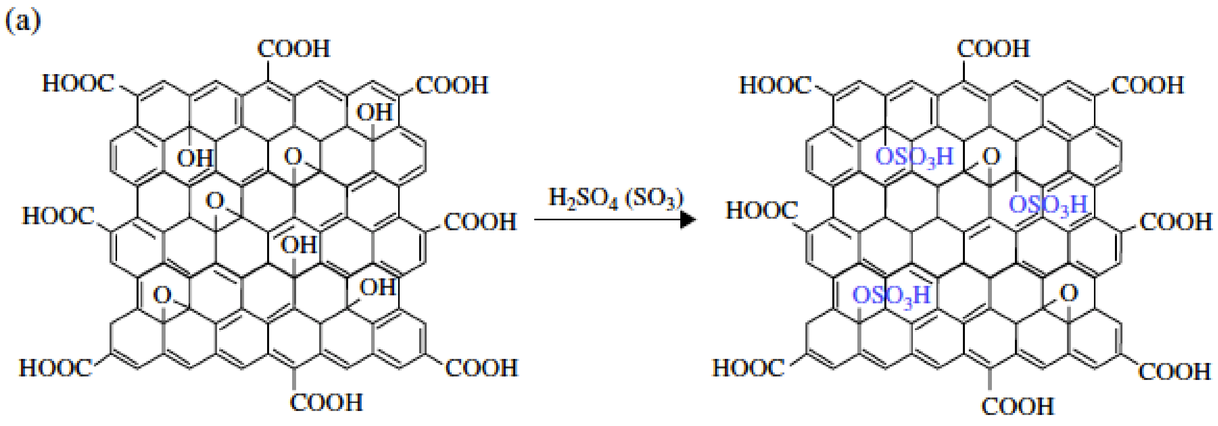


Figure 6.24 (a) Chemical structures of the PTB7 donor, PC71BM acceptor, and GO. (b) Schematic of the PSC device indicating the location of the GO. (c) Representative current density–voltage plots under AM 1.5G solar simulated light for PSCs with PEDOT:PSS and GO as the hole extraction layer. (d) Thermal degradation of encapsulated devices at 80 °C under a N^2 atmosphere. (e) Environmental degradation of unencapsulated devices fabricated with air- stable electrodes at 25ÅãC under 80% relative humidity

Although excellent device performance has been achieved, the performance of PSC device with GO as an HEL is highly sensitive to the GO layer thickness due to its insulating property. To address this problem, Liu et al. [39] treated GO with oleum and developed sulfated graphene oxide (GO–OSO3 H), in which −OSO3 H groups were introduced to the basal plane of reduced GO (Fig. 6.25a). GO–OSO3 H had the advantages of good solubility for solution processing due to the presence of −OSO3 H groups and improved conductivity due to the reduced basal plane. The much improved conductivity of GO–OSO3 H (1.3 S/m vs. 0.004 S/m) led to greatly improved FF (0.71 vs. 0.58) and PCE (4.37% vs. 3.34%) for a PSC device based on GO–OSO3 H HEL with respect to the GO-based counterpart (Fig. 6.25c). On the other hand, Yun et al. [40] reported the use of graphene oxide (pr-GO) reduced by p -toluenesulfonyl hydrazide (p -TosNHNH2) as an HEL. pr-GO could be dispersed with high concentration (0.6 mg/mL) and could give uniform film by spin coating. PSC devices based on the pr-GO HEL exhibited a fairly comparable photovoltaic efficiency and much improved lifetime compared with those of the PEDOT:PSS device. Various studies reported by many other groups have also proven that photovoltaic performance of GO-based PSC devices could be significantly improved by increasing the conductivity of the GO layer. Examples include the blending of highly conductive SWCN in the GO layer to increase the GO layer lateral conductivity [41] and the use of thermal annealing [42] and plasma treatment [43] to reduce GO during device fabrication for increasing its conductivity.



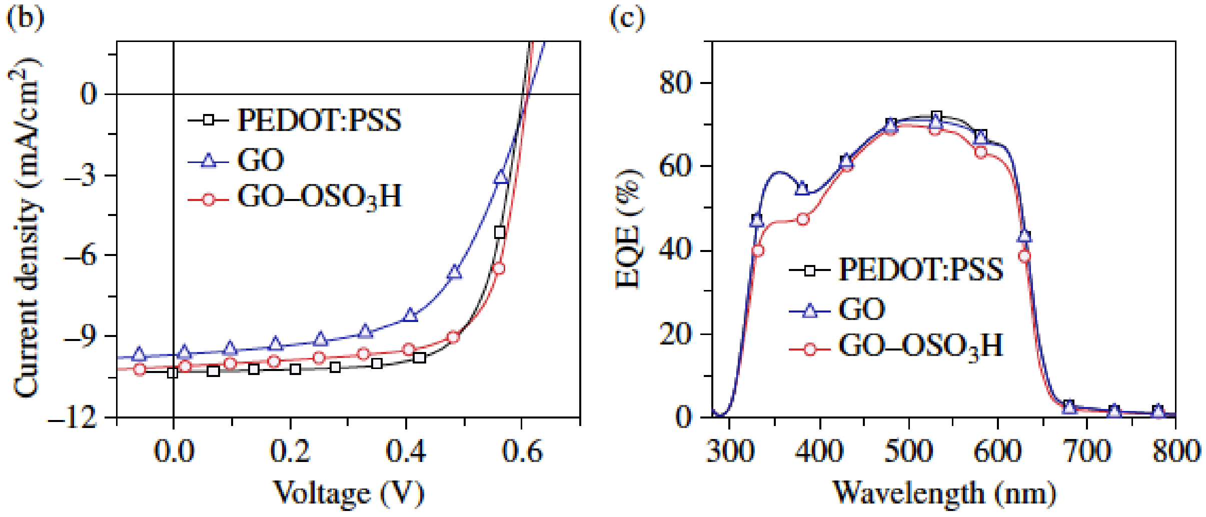


Figure 6.25 (a) Synthetic route to GO–OSO3H. Current density–voltage curves (b) and external quantum efficiency curves (c) of the PSC devices with PEDOT:PSS (25 nm), GO (2 nm), or GO–OSO3H (2 nm) as the hole extraction layer

6.14 QD solar cells containing carbon nanomaterials

Using QDs with tuned band gaps, the single-material-system solar cell is able to contain multiple junctions, having a better use of the sun spectrum even challenging the single-junction limitation. Figure 6.26 illustrates the structure and band diagram of two types of QD-based devices [5].

However, due to the organic ligands and defect states on the surface of QDs, the mobility of photogenerated electrons is very limited, while the recombination of electrons and holes occurs easily on the QD surface. Therefore, the major problems that need to be solved are the separation of photogenerated hole/electron pairs and the transport of carriers from QDs to the electrode. Therefore, C60 , CNTs, and graphene have been applied as transparent conductive electrodes (TCE), electron acceptors, and light absorbers to improve the device performance. Herein, CNTs and graphene as TCE, carbon nanomaterial/QD composites, and graphene QDs (GQDs) will be discussed in the context of QD-based PVs.

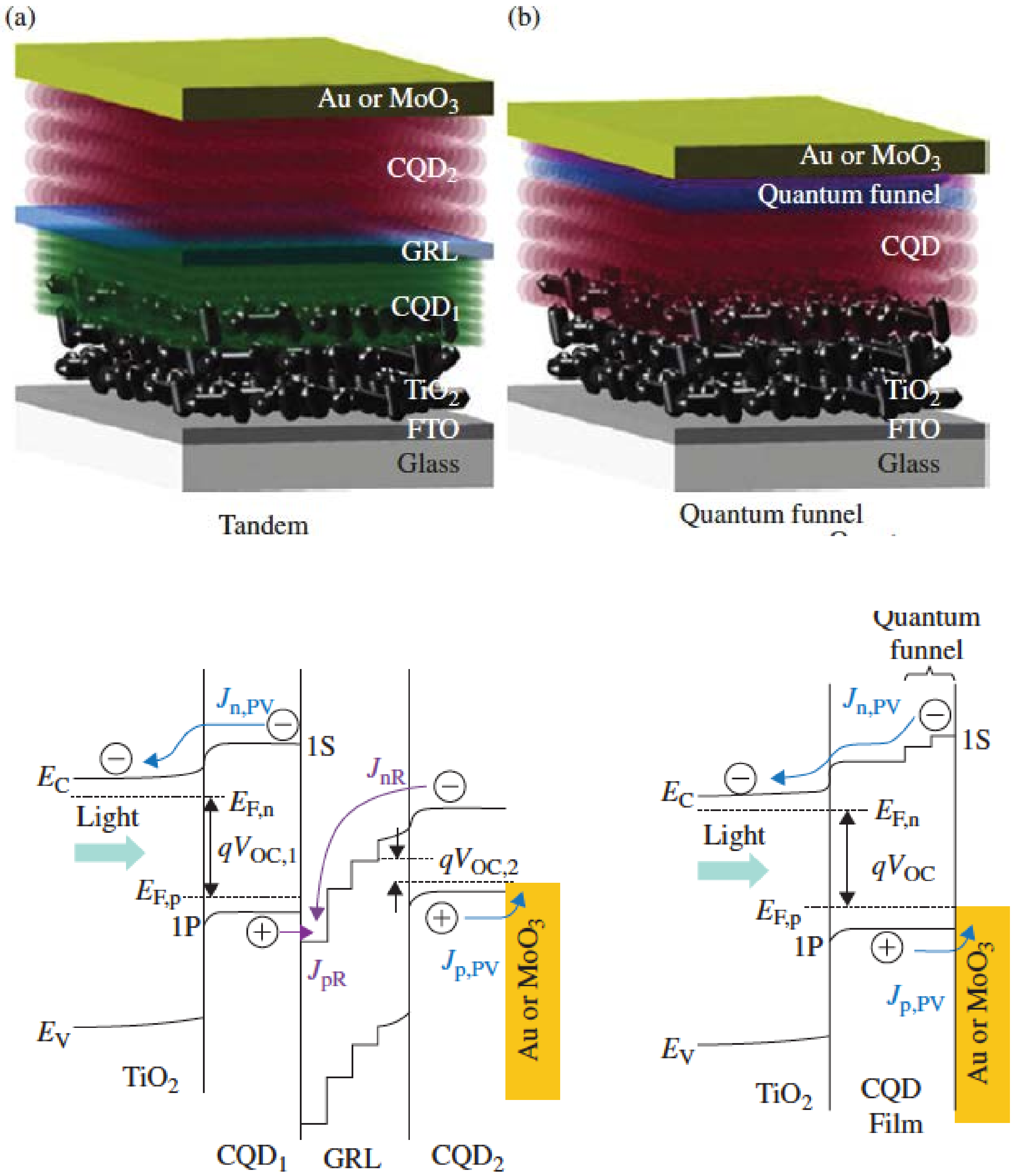


Figure 6.26 Device structure and band diagram of the tandem cell and the quantum funnel solar cells. (a) Device structure and band diagram of the tandem cell. The two junctions contain two layers of QDs of different sizes and a graded recombination layer (GRL) lies in between. (b) Device structure and band diagram of the quantum funnel solar cell. The quantum funnel is formed by the same QDs but in different sizes

C60 and QD Composites

C60 has been widely used in organic PVs as electron acceptor [25] because of its low work function and high electron mobility. Biebersdorf et al. initialized the research of the photosensitization of C60 by nanocrystals [26]. The mixture of trioctylphosphine oxide (TOPO)-/trioctylphosphine (TOP)- capped CdSe QDs and C60 in toluene was dropped cast onto metal contact; then as the toluene evaporated, the needlelike C60 crystals formed. Under illumination, a threefold enhancement of photocurrent was observed from the mixed solid compared with single C60 crystals. However, this preparation method is too rough to build good interaction between the two components. Consequently, Brown et al. proposed a new method to obtain CdSe–n C60 composite clusters by introducing electrophoretic deposition [27]. The formed cluster looks like a CdSe core with C60 shell surrounding and numerous clusters ensemble a thin film. With a more efficient structure, the incident photon to photocurrent generation efficiency (IPCE) of the cluster increased to 4%. The photocurrent generated by the composite is two orders of magnitude greater than pristine CdSe film (Fig. 6.27).

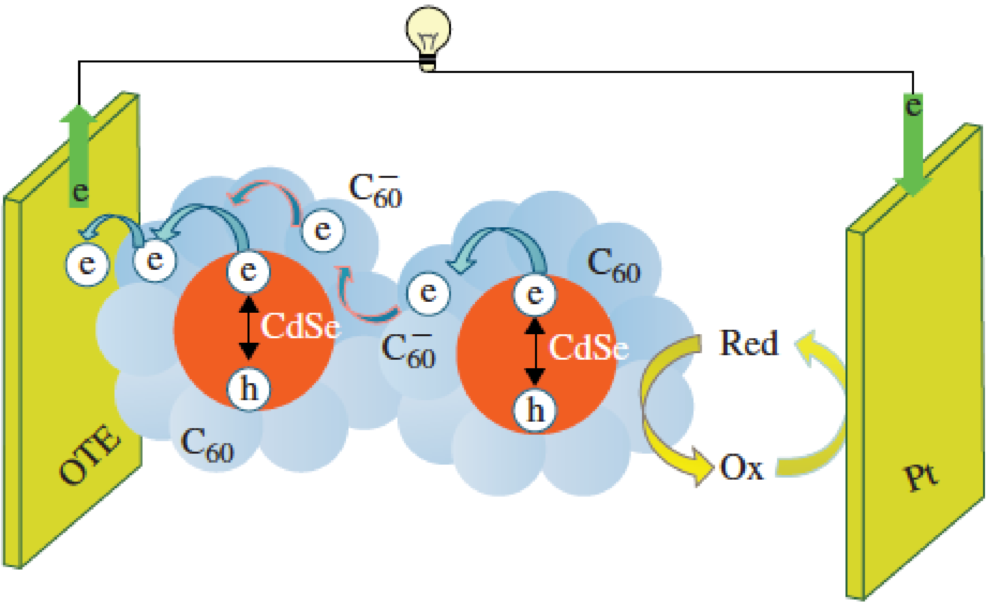


Fig. 6.27 Photocurrent generation by CdSe–nC$_{60}$ clusters

CNTs and QD Composites

One-dimensional nanostructures can not only accept and transport photogenerated electrons from QDs but also provide scaffold for anchoring the light-absorbing QDs. CNTs have been widely studied in combination with various QDs. There are abundant literatures reporting CNT/QD composite synthesis, properties, and its application in solar cells. In general, the preparation methods of CNT/QD composite can be categorized as selective coupling and direct growing or deposit.

The basic idea of the first approach is to attach QDs to functionalized CNTs through the “linker,” which is formed by the functional group on CNT walls and the corresponding ligand on QD surfaces. Different interactions can be used to hold the linkers, such as covalent chemical bonds [31] and noncovalent electrostatic interactions [32]. The drawback of covalent bonding is the property of pristine CNT that can be destroyed during the side-wall functionalization, while using noncovalent surfactants can protect the CNT yet sometimes increase the distance between CNTs and QDs, lowering the charge transfer. The second approach follows a complete distinct concept, based on which the QDs are grown directly from the CNTs without any linker in between. There are also covalent [33] and noncovalent [34] interactions in this kind. Among these synthetic mechanisms, the noncovalent direct growth of QDs onto the CNTs seems the most promising, because it has the least modifications to the original CNTs, and the simple one-step wet chemical method has proven to be effective for a variety of QDs.

After studying the fundamental carrier transport between QDs and CNTs through all aforementioned composites, researchers spare their efforts developing PV devices based on this structure. According to the band gap of QDs, the nanostructured composite can be divided into wide-band gap CNTs/QDs and narrow-band gap CNTs/QDs, enhancing the absorption of ultraviolet (UV) region and near-infrared (NIR) region, respectively. In the first group, CNTs have been combined with CdSe[35], CdS [36], TiO2 [37], CdTe [34], and other nanocrystals. Landi et al. were able to attach CdSe QDs to SWCNTs covalently via the chemical bonds between aminoethanethiol (AET) ligand of QD and carboxylic acid function group of CNTs [35]. In order to achieve a proper energy band alignment (Fig. 6.28), a poly(3-octylthiophene) (P3OT) layer was incorporated, which is a p-type polymer. Although only limited PCE can be gained due to undesirable recombination, a Voc of 0.75 V still indicates a decent configuration of the device. In the case of CdS CNTs/QDs, Li et al. provided another way of fabricating solar cells (Fig. 6.29) [36]. Based on the heterojunction model of SWCNT/Si solar cell, CdS QDs were successfully grafted onto CNTs to further improve light harvesting. With appropriate CdS density on the surface of CNTs, the device gives out a Voc of 0.47 V, Jsc of 6.82 mA/cm2, FF of 43.7%, and PCE of 1.4% under AM 1.5 (100 mW/cm2). One more interesting point is that owing to existence of CdS QDs, the device shows degradation under continuous illumination; however, the mechanism remains unclear. In addition to wide-band gap semiconductor nanocrystals, PbS QDs have also been attached to MWCNTs in the application of NIR PV cells [38]. To form the donor–acceptor structure, P3HT was employed as the hole-conducting layer. The interaction between QDs and CNTs is strong, proven by the complete quenching of PbS, even though no ligand exchange steps applied. Finally, the optimized P3HT:PbS/MWCNT PV device can achieve a PCE of 3.03%.

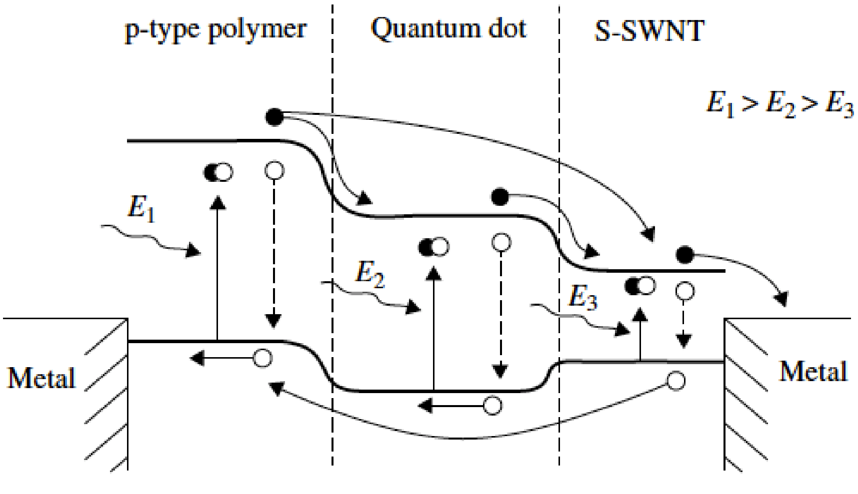
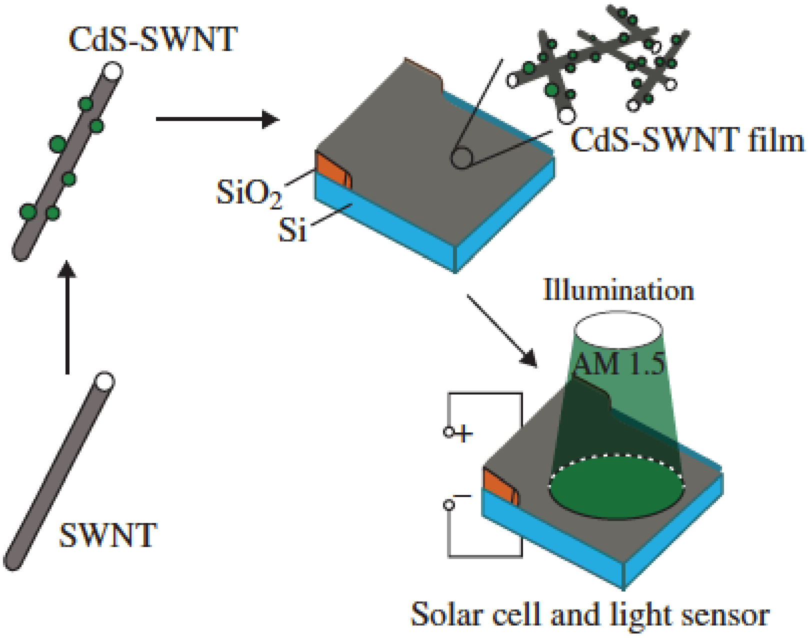


Fig.6.28 Illustration of the equilibrated band diagram for SWCNT/QD polymer solar

cell



6.29 Fabrication process of CdS/SWCNT/Si solar cells.

References

[1] (A) A.E. Becquerel, Compt. Rend. Acad. Sci. 9 (1839) 145.;
(B) A.E. Becquerel, Compt. Rend. Acad. Sci. 9 (1839) 561.
[2] (A) W. Smith, Nature 7 (1873) 303.;
(B) W.G. Adams, R.E. Day, Proc. R. Soc. London 25 (1876) 113.
[3] A. Pochettino, Acad. Lincei Rend. 15 (1906) 355.
[4] M. Volmer, Ann. Physik 40 (1913) 775.
[5] P.M. Borsenberger, D.S. Weiss, Organic Photoreceptors For Imaging Systems, Marcel Dekker,
New York, 1993.
[6] R.H. Bube, Photoconductivity of Solids, Wiley, New York, 1960.
[7] (A) For a review on solar cells basedon dyes see: S. Anthoe, Rom. Rep. Phys. 53 (2002) 427;
(B) For a review on earlier work see: G.A. Chamberlain, Sol. Cells 8 (1983) 47.
[8] D.M. Chapin, C.S. Fuller, G.L. Pearson, J. Appl. Phys. 25 (1954) 676.
[9] For a recent table of inorganic solar cell efficiencies see: M.A. Green, K. Emery, D.L. King, S. Igari,
W. Warta, Prog. Photovolt.: Res. Appl. 11 (2003) 347.
[10] For a comprehensive review on Si andother types of solar cells see: A. Goetzberger, C. Hebling, H.W.
Chock, Mater. Sci. Eng. R 40 (2003) 1.
[11] For a review on hybridcells see: P. Gomez-Romero, Adv. Mater. 13 (2001) 163.
[12] For a review of PEC type cells see: A.K. Jana, J. Photochem. Photobiol. A 132 (2000) 1.

[13] (A) J.H. Burroughes, D.D.C. Bradley, A.R. Brown, R.N. Marks, K. Mackay, R.H. Friend, P.L Burns, A.B. Holmes, Nature 34 (1990) 539.;
(B) N.C. Greenham, S.C. Moratti, D.D.C. Bradley, R.H. Friend, A.B. Holmes, Nature 365 (1993)
628;
(C) For an account of PPVs in LEDs see: R.H. Friend, G.J. Denton, J.J.M. Halls, N.T. Harrison, A.B. Holmes, A. K. ohler, A. Lux, S.C. Moratti, K. Pichler, N. Tessler, K. Towns, Synth. Met. 84
(1997) 463.
[14] For reviews on polaron andexciton formation in PPV see:;
(A) U. Rauscher, H. B.assler, D.D.C. Bradley, M. Hennecke, Phys. Rev. B 42 (1990) 9830;;
(B) E.L. Frankevich, A.A. Lymarev, I. Sokolik, F.E. Karasz, S. Blumstengel, R.H. Baughman, H.H. Horh. old, Phys. Rev. B 46 (1992) 9320.
[15] (A) L.J. Rothberg, M. Yan, F. Papadimitrakopolous, M.E. Galvin, E.W. Kwock, T.M. Miller,
Synth. Met. 80 (1996) 41.;
(B) L.J. Rothberg, M. Yan, S. Son, M.E. Galvin, E.W. Kwock, T.M. Miller, H.E. Katz, R.C. Harddon, F. Papadimitrakopolous, Synth. Met. 78 (1996) 231.
[16] For a thorough treatment see: V.Y. Merritt, in: D.A. Seanor (Eds.), Electrical Properties of Polymer,
Academic Press, New York, 1982, pp. 127 (Chapter 4).
[17] See for example: R.N. Marks, J.J.M. Halls, D.D.C. Bradley, R.H. Friend, A.B. Holmes, J. Phys.:
Condens. Matter 6 (1994) 1379.
[18] C.J. Brabec, A. Cravino, D. Meissner, N.S. Sariciftci, M.T. Rispens, L. Sanchez, J.C. Hummelen,
T. Fromherz, Thin SolidFilms 403 (2002) 368.
[19] For a goodreview on surfaces andinterfaces in polymer-based electronics see: M. Fahlman, W. R. Salaneck, Surf. Sci. 500 (2002) 904.
[20] L.S. Hung, C.W. Tang, M.G. Mason, Appl. Phys. Lett. 70 (1997) 152.
[21] C. Greczynski, M. Fahlman, W.R. Salaneck, J. Chem. Phys. 69 (1996) 599.
[22] H. Tang, J. Shinar, J. Appl. Phys. Lett. 71 (1997) 2560.
[23] C.J. Brabec, S.E. Shaheen, C. Winder, N.S. Sariciftci, P. Denk, Appl. Phys. Lett. 80 (2002) 1288.
[24] (A) C. Xirochaki, G. Kiriakidis, T.F. Pedersen, H. Fritzsche, J. Appl. Phys. 79 (1996) 9349.;
(B) S. Mailis, L. Boutsikaris, N.A. Vainos, C. Xirochaki, G. Vasiliou, N. Garawal, G. Kiriakidis, H. Fritzsche, Appl. Phys. Lett. 69 (1996) 2459.;
(C) C. Xirochaki, K. Moschovis, E. Chatzitheodoridis, G. Kiriakidis, P. Morgen, Appl. Phys. A 67
(1998) 295.;
(D) S. Pissadakis, S. Mailis, L. Reekie, J.S. Wilkinson, R.W. Eason, N.A. Vainos, K. Moschovis, G. Kiriakdikis, Appl. Phys. A 69 (1999) 333.
[25] J.S. Kim, M. Granstrom, R.H. Friend, N. Johansson, W.R. Salaneck, R. Daik, W.J. Feast, F. Cacialli, J. Appl. Phys. 84 (1998) 6859.
[26] J.C. Scott, J.H. Kaufman, P.J. Brock, R. DiPietro, J. Salem, J.A. Gotia, J. Appl. Phys. 79 (1996)
2745.
[27] A.R. Schlatman, D.W. Floet, A. Hilberer, F. Garten, P.J.M. Smulders, T.M. Klapwijk, G. Hadziioannou, Appl. Phys. Lett. 69 (1996) 1764.
[28] (A) S.A. Carter, M. Angelopoulos, S. Karg, P.J. Brock, J.C. Scott, Appl. Phys. Lett. 70

(1997) 2067.;
(B) Y. Cao, G. Yu, C. Zhang, R. Menon, A.J. Heeger, Synth. Met. 87 (1997) 171.;
(C) J.C. Carter, I. Grizzi, S.K. Heeks, D.J. Lacey, S.G. Latham, P.G. May, O.R.d.l. Panos, K. Pichler, C.R. Towns, H.F. Wittman, Appl. Phys. Lett. 71 (1997) 34.
[29] (A) A.M. Mathieson, J.M. Robertson, V.C. Sinclair, Acta. Crystallogr. 3 (1950) 245.;
(B) V.C. Sinclair, J.M. Robertson, A.M. Mathieson, Acta. Crystallogr. 3 (1950) 251.
[30] (A) G.J. Sloan, Mol. Cryst. 1 (1966) 161.;
(B) G.J. Sloan, Mol. Cryst. 2 (1967) 323.;
(C) G.J. Sloan, J.M. Thomas, J.O. Williams, Mol. Cryst. Liq. Cryst. 30 (1975) 167.
[31] D. Kearns, M.J. Calvin, Chem. Phys. 29 (1958) 950.
[32] G.M. Delacote, J.P. Fillard, F. Marco, Solid State Commun. 2 (1964) 373.
[33] (A) M.I.F. Ederov, V.A. Benderskii, Sov. Phys.-Semicond. 4 (1971) 1198.;
(B) M.I. Federov, V.A. Benderskii, Sov. Phys.-Semicond. 4 (1971) 1720.

[34] R.F. Chaiken, D.R. Kearns, J. Chem. Phys. 45 (1966) 451.
[35] C.W. Tang, A.C. Albrecht, J. Chem. Phys. 63 (1975) 953;
C.W. Tang, A.C. Albrecht, J. Chem. Phys. 62 (1975) 2139.
[36] (A) E.K. Putseiko, in: B.S. Neporent (Ed.), Elementary Photoprocesses in Molecules, Consultant Bureau, New York, 1968, p. 281.;
(B) A. Terenin, E.K. Putseiko, I. Akimov, Discuss. Faraday Soc. 27 (1959) 83.
[37] K. Ballschmitter, J.J. Katz, Biochim. Biophys. Acta 256 (1972) 30.
[38] M. Mabrouki, A. Oueriagli, A. Outzourhit, E.L. Ameziane, S. Hotchandani, R.M. LeBlanc, Phys. Stat. Sol. 191 (2002) 345.
[39] (A) C.W. Tang, US Patent 4,164,431, August 14, 1979;;
(B) C.W. Tang, Appl. Phys. Lett. 48 (1986) 183.
[40] (A) For a comprehensive review on polymerbased PVs see: J.J.M. Halls, R.H. Friend, (Eds.), Clean Electricity From Photovoltaics, in: M.D. Archer; R. Hill (Eds.), Imperial College Press, (2001); Chapter 9;
(B) For a goodreview see: C.J. Brabec, N.S. Sariciftci, J.C. Hummelen, Adv. Func. Mater. 11 (2001) 15;
(C) For a recent review see: J. Nelson, Curr. Opin. SolidState M. 6 (2002) 87.
[41] R.F. Cozzens, in: D.A. Seanor, (Ed.), Electrical Properties of Polymer, Academic Press, New York, 1982, p. 93 (Chapter 3).
[42] (A) Handbook of Conducting Polymers, T.A. Skotheim, (Ed.), Vol. 1 and 2, Dekker, New, York, 1986;
(B) Conductive Polymers and Plastics J. M.Margolis, (Ed.), Chapman & Hall, New York, 1989;
(C) For a recent review on semiconducting polymers see: A.J. Heeger J. Phys. Chem. 105 (2001) 8475.
[43] B.R. Weinberger, M. Akhtar, S.C. Gau, Synth. Met. 4 (1982) 187.
[44] S. Glenis, G. Tourillon, F. Garnier, Thin SolidFilms 139 (1986) 221.
[45] S. Karg, W. Riess, V. Dyakonov, M. Schwoerer, Synth. Met. 54 (1993) 427.
[46] R.N. Marks, J.J.M. Halls, D.D.C. Bradley, R.H. Friend, A.B. Holmes, J. Phys.-Condens. Matter 6 (1994) 1379.
[47] H. Antoniadis, B.R. Hsieh, M.A. Abkowitz, S.A. Jenekhe, M. Stolka, Synth. Met. 62 (1994)

265.
[48] G. Yu, C. Zhang, A.J. Heeger, Appl. Phys. Lett. 64 (1994) 1540.
[49] E.L. Frankevich, A. Zakhidov, K. Yoshino, Phys. Rev. B 53 (1996) 4498.
[50] R.C. Nelson, J. Opt. Soc. Am. 46 (1956) 13.
[51] (A) Md.K.H. Bhuiyan, T. Mieno, Thin Solid Films 441 (2003) 187.;
(B) M.T. Rispens, J.C. Hummelen, Dev. Fullerene Sci. 4 (2002) 387;
(C) For review on electroactive fullerenes see: N. Mart!ın, L. S!anchez, B. Illescas, I. P!erez, Chem. Rev.
98 (1998) 2527.
[52] N.S. Sariciftci, L. Smilowitz, A.J. Heeger, F. Wudl, Synth. Met. 59 (1993) 333;
N.S. Sariciftci, D. Braun, C. Zhang, V.I. Srdanov, A.J. Heeger, G. Stucky, F. Wudl, Appl. Phys. Lett.
62 (1993) 585.
[53] S. Morita, A.A. Zakhidov, K. Yoshino, Jpn. J. Appl. Phys. 32 (1993) 873.
[54] (A) J.J.M. Halls, K. Pichler, R.H. Friend, S.C. Moratti, A.B. Holmes, Synth. Met. 77 (1996) 277.;
(B) J.J.M. Halls, K. Pichler, R.H. Friend, S.C. Moratti, A.B. Holmes, Appl. Phys. Lett. 68 (1996) 3120.
[55] J.J.M. Halls, R.H. Friend, Synth. Met. 85 (1997) 1307.
[56] (A) L.A. Pettersson, L.S. Roman, O. Inganas, J. Appl. Phys. 86 (1999) 487.;
(B) L.S. Roman, W. Mammo, L.A. Pettersson, M.R. Andersson, O. Inganas, Adv. Mater. 10 (1998)
774.
[57] M.F. Durstock, R.J. Spry, J.W. Baur, B.E. Taylor, L.Y. Chiang, J. Appl. Phys. 94 (2003) 3253.
[58] (A) K. Tada, M. Onoda, H. Nakayama, K. Yoshino, Synth. Met. 102 (1999) 982.;
(B) K. Tada, M. Onoda, A.A. Zakhidov, K. Yoshino, Jpn. J. Appl. Phys. (Part 2-Lett.) 36 (1997) L306–L309.
[59] M. Hiramoto, H. Fukusumi, M. Yokoyama, Appl. Phys. Lett. 58 (1991) 1062.
[60] M. Hiramoto, H. Fukusumi, M. Yokoyama, Appl. Phys. Lett. 61 (1992) 2580.

[61] L. Schmidt-Mende, A. Fectenk. otter, K. M. ullen, E. Moons, R.H. Friend, J.D. MacKenzie, Science
293 (2001) 1119.
[62] G. Yu, K. Pakbaz, A.J. Heeger, Appl. Phys. Lett. 64 (1994) 3422.
[63] A. K. ohler, H.F. Wittman, R.H. Friend, M.S. Kahn, J. Lewis, Synth. Met. 77 (1996) 147.
[64] J.C. Hummelen, B.W. Knight, F. LePeq, F. Wudl, J. Yao, C.L. Wilkins, J. Org. Chem. 60 (1995) 532.
[65] G. Yu, J. Gao, J.C. Hummelen, F. Wudl, A.J. Heeger, Science 270 (1995) 1789.
[66] S.E. Shaheen, C.J. Brabec, N.S. Sariciftci, F. Padinger, T. Fromherz, J.C. Hummelen, Appl. Phys.
Lett. 78 (2001) 841.
[67] T. Munters, T. Martens, L. Goris, V. Vrindts, J. Manca, L. Lutsen, W. De Ceuninck,
D. Vanderzande, L. De Schepper, J. Gelan, N.S. Sariciftci, C.J. Brabec, Thin Solid Film 403
(2002) 247.
[68] M.T. Rispens, A. Meetsma, R. Rittberger, C.J. Brabec, N.S. Sariciftci, J.C. Hummelen, Chem.
Commun. 17 (2003) 2116.
[69] (A) D. Gebeheyu, F. Padinger, T. Fromherz, J.C. Hummelen, N.S. Sariciftci, Int. J. Photoenergy 1
(1999) 95.;

(B) N. Gamaioni, G. Ridolfi, G. Casalbore-Miceli, G. Possamai, L. Garlaschelli, M. Maggini, Sol.
Energy Mater. Sol. Cells 76 (2003) 107.
[70] P. Schilinsky, C. Waldauf, C.J. Brabec, Appl. Phys. Lett. 81 (2002) 3885.
[71] W. Geens, T. Aernouts, J. Poortmans, G. Hadziioannou, Thin Solid Films 403 (2002) 438.
[72] G. Yu, A.J. Heeger, J. Appl. Phys. 78 (1995) 4510.
[73] J.J.M. Halls, C.A. Walsh, N.C. Greenham, E.A. Marseglia, R.H. Friend, S.C. Moratti, A.B. Holmes,
Nature 376 (1995) 498.
[74] J.J.M. Halls, A.C. Arias, J.D. Mackenzie, W. Wu, M. Inbasekaran, E.P. Woo, R.H. Friend, Adv.
Mater. 12 (2000) 498.
[75] L.C. Chen, O. Ingan.as, L.S. Roman, M. Johansson, M. Andersson, Thin Solid Films 363 (2000) 286.
[76] U. Stalmach, B.d. Boer, C. Videlot, P.F.v. Hutten, G. Hadziioannou, J. Am. Chem. Soc. 112 (2000)
5464.
[77] (A) E. Peeters, P.A.v. Hal, J. Knol, C.J. Brabec, N.S. Sarciftci, J.C. Hummelen, R.A.J. Janssen,
Phys. Chem. B. 104 (2000) 10174;
(B) For a review on fullerene-oligomer dyads see: J.-F. Nierengarten, Sol. Energy Mater Sol Cells,
ibid.
[78] P.A. van Hal, J. Knol, B.M.W. Langeveld-Voss, S.C.J. Meskers, J.C. Hummelen, R.A.J. Janssen, J.
Phys. Chem. A. 104 (2000) 5974.
[79] (A) F.C. Krebs, M. Jørgensen, Macromolecules 35 (2002) 7200.;
(B) F.C. Krebs, H. Spanggaard, N. Rozlosnik, N.B. Larsen, M. Jørgensen, Langmuir 19 (2003) 7873.
[80] F.C. Krebs, Sol. Energy Mater. Sol. Cells 80 (2003) 257.
[81] For review on fullerene in double-cable polymers see: A. Cravino, N.S. Sariciftci, J. Mater. Chem. 12
(2002) 1931.
[82] A. Cravino, N.S. Sariciftci, Nat. Mater. 6 (2003) 360.
[83] T. Benincori, E. Brenna, F. Sannicol ! o, L. Trimarco, G. Zotti, Angew. Chem. 108 (1996) 718.
[84] J.P. Ferraris, M. Hmyene, D.C. Loveday, A. Yassar, Synth. Met. 84 (1997) 231.
[85] (A) A.M. Ramos, M.T. Rispens, J.C. Hummelen, R.A.J. Janssen, Synth. Met. 119 (2001) 171.;
(B) A.M. Ramos, M.T. Rispens, J.K.J. van Duren, J.C. Hummelen, R.A.J. Janssen, J. Am. Chem. Soc. 123 (2001) 6714.
[86] F. Zhang, M. Svensson, M.R. Andersson, M. Maggini, S. Bucella, E. Menna, O. Ingan.as, Adv.
Mater. 13 (2001) 1871.
[87] A. Cravino, G. Zerza, M. Maggini, S. Bucella, M. Svensson, M.R. Andersson, H. Neugebauer, C.J.
Brabec, N.S. Sariciftci, Monatsh. Chem. 134 (2003) 519.
[88] F.C. Krebs, M. Jørgensen, Polym. Bull. 50 (2003) 359.
[89] S.-S. Sun, Sol. Energy Mater. Sol. Cells 79 (2003) 257.
[90] D. Gebeheyu, F. Padinger, T. Fromherz, J.C. Hummelen, N.S. Sariciftci, Int. J. Photoenerg 1 (1999) 95.
[91] S.E. Shaheen, R. Radspinner, N. Peyghambarian, G.E. Jabbour, Appl. Phys. Lett. 79 (2001)

2996.

[92] F.C. Krebs, J. Alstrup, H. Spanggaard, K. Larsen, E. Kold, Sol. Energy Mater. Sol. Cells

[1] G. Yu, J. Gao, J. C. Hummelen, F. Wudl, A. J. Heeger, Polymer Photovoltaic Cells: Enhanced Efficiencies via a Network of Internal Donor-Acceptor Heterojunctions. Science 1995, 270, 1789–1791.
[2] Y. Y. Liang, Z. Xu, J. B. Xia, S. T. Tsai, Y. Wu, G. Li, C. Ray, L. P. Yu, For the Bright Future-Bulk Heterojunction Polymer Solar Cells with Power Conversion Efficiency of 7.4%. Adv. Mater. 2010, 22, E135–E138.
[3] T. Y. Chu, J. P. Lu, S. Beaupre, Y. G. Zhang, J. R. Pouliot, S. Wakim, J. Y. Zhou, M. Leclerc, Z. Li, J. F. Ding, Y. Tao, Bulk Heterojunction Solar Cells Using Thieno[3,4-c] pyrrole-4,6-dione and Dithieno[3,2-b:2’,3’-d]silole Copolymer with a Power Conversion Efficiency of 7.3%. J. Am. Chem. Soc. 2011, 133, 4250–4253.
[4] H. Y. Chen, J. H. Hou, S. Q. Zhang, Y. Y. Liang, G. W. Yang, Y. Yang, L. P. Yu, Y. Wu, G. Li, Polymer Solar Cells with Enhanced Open-Circuit Voltage and Efficiency. Nat. Photonics 2009, 3, 649–653.
186 HIGH-PERFORMANCE POLYMER SOLAR CELLS CONTAINING CARBON
[5] J. B. You, L. T. Dou, K. Yoshimura, T. Kato, K. Ohya, T. Moriarty, K. Emery, C.-C. Chen, J. Gao, G. Li, Y. Yang, A Polymer Tandem Solar Cell with 10.6% Power Conversion Efficiency. Nat. Commun. 2013, 4, 1446.
[6] B. C. Thompson, J. M. J. Fr.chet, Polymer-Fullerene Composite Solar Cells. Angew. Chem. Int. Ed. 2007, 47, 58–77.
[7] Z. C. He, C. M. Zhong, S. J. Su, M. Xu, H. B. Wu, Y. Cao, Enhanced Power-Conversion Efficiency in Polymer Solar Cells Using an Inverted Device Structure. Nat. Photonics 2012, 6, 591–595.
[8] A. K. Geim, K. S. Novoselov, The Rise of Graphene. Nat. Mater. 2007, 6, 183–191.
[9] A. K. Geim, Graphene: Status and Prospects. Science 2009, 324, 1530–1534.
[10] D. R. Dreyer, R. S. Ruoff, C. W. Bielawski, From Conception to Realization: An Historial Account of Graphene and Some Perspectives for Its Future. Angew. Chem. Int. Ed. 2010, 49, 9336–9344.
[11] L. M. Dai, D. W. Chang, J.-B. Baek, W. Lu, Carbon Nanomaterials for Advanced Energy Conversion and Storage. Small 2012, 8, 1130–1166.
[12] J. Liu, Y. H. Xue, M. Zhang, L. M. Dai, Graphene-Based Materials for Energy Applications. MRS Bull. 2012, 37, 1265–1272.
[13] S. Iijima, Helical Microtubules of Graphitic Carbon. Nature 1991, 354, 56–58.
[14] X. J. Wan, G. K. Long, L. Huang, Y. S. Chen, Graphene-a Promising Material for Organic Photovoltaic Cells. Adv. Mater. 2011, 23, 5342–5348.
[15] S. Cataldo, P. Salice, E. Menna, B. Pignataro, Carbon Nanotubes and Organic Solar Cells. Energy Environ. Sci. 2012, 5, 5919–5940.
[16] A. Iwan, A. Chuchmała, Perspectives of Applied Graphene: Polymer Solar Cells. Prog. Polym. Sci. 2012, 37, 1805–1828.
[17] Y. J. He, Y. F. Li, Fullerene Derivative Acceptors for High Performance Polymer Solar Cells. Phys. Chem. Chem. Phys. 2011,13, 1970–1983.
[18] D. S. Hecht, L. B. Hu, G. Irvin, Emerging Transparent Electrodes Based on Thin Films of Carbon Nanotubes, Graphene, and Metallic Nanostructures. Adv. Mater. 2011, 23, 1482–1513.
[19] M. G. Kang, L. J. Guo, Nanoimprinted Semitransparent Metal Electrodes and Their Application in Organic Light-Emitting Diodes. Adv. Mater. 2007, 19, 1391–1396.
[20] Y. J. Xia, K. Sun, J. Y. Ouyang, Solution-Processed Metallic Conducting Polymer Films

as Transparent Electrode of Optoelectronic Devices. Adv. Mater. 2012, 24, 2436–2440.

[21] J. vande Langemaat, T. M. Barnes, G. Rumbles, S. E. Shaheen, T. J. Coutts, C. Weeks, I. Levitsky, J. Peltola, P. Glatkowski, Organic Solar Cells with Carbon Nanotubes Replacing In2O3:Sn as the Transparent Electrode. Appl. Phys. Lett. 2006, 88, 233503.

[22] T. M. Barnes, J. D. Bergeson, R. C. Tenent, B. A. Larsen, G. Teeter, L. M. Jones, J. L. Blackburn, J. vande Langemaat, Carbon Nanotube Network Electrodes Enabling Efficient Organic Solar Cells without a Hole Transport Layer. Appl. Phys. Lett. 2010, 96, 243309.

[23] M. W. Rowell, M. A. Topinka, M. D. McGehee, H. J. Prall, G. Dennler, N. S. Sariciftci, L. Hu, G. Grüner, Organic Solar Cells with Carbon Nanotube Network Electrodes. Appl. Phys. Lett. 2006, 88, 233506.

[24] R. C. Tenent, T. M. Barnes, J. D. Bergson, A. J. Ferguson, B. To, L. M. Gedvilas, M. J. Heben, J. L. Blackburn, Ultrasmooth, Large-Area, High-Uniformity, Conductive Transparent Single-Walled-Carbon-Nanotube Films for Photovoltaics Produced by Ultrasonic Spraying. Adv. Mater. 2009, 21, 3210–3216.

[25] J. B. Wu, M. Agrawal, H. A. Becerril, Z. N. Bao, Z. F. Liu, Y. S. Chen, P. Peumans, Organic Light-Emitting Diodes on Solution-Processed Graphene Transparent Electrodes. ACS Nano 2010, 4, 43–48.

[26] S. Bae, H. Kim, Y. Lee, X. F. Xu, J.-S. Park, Y. Zheng, J. Balakrishnan, Y. Lei, H. R. Kim, Y. I. Song, Y. J. Kim, B. .yilmaz, J. H. Ahn, B. H. Hong, S. Lijima, Roll-to-Roll Production of 30-inch Graphene Films for Transparent Electrodes. Nat. Nanotechnol. 2010, 5, 574–578.

[27] X. S. Li, W. W. Cai, J. H. An, S. Kim, J. Nah, D. X. Yang, R. D. Piner, A. Velamakanni, I. Jung, E. Tutuc, S. K. Banerjee, L. Colombo, R. S. Ruoff, Large-Area Synthesis of High-Quality and Uniform Graphene Films on Copper Foils. Science 2009, 324, 1312–1314.

[28] G. Eda, M. Chhowalla, Chemically Derived Graphene Oxide: Towards Large-Area Thin-Film Electronics and Optoelectronics. Adv. Mater. 2010, 22, 2392–2415.

[29] J. X. Geng, L. J. Liu, S. B. Yang, S. C. Youn, D. W. Kim, J. S. Lee, J. K. Choi, H. T. Jung, A Simple Approach for Preparing Transparent Conductive Graphene Films Using the Controlled Chemical Reduction of Exfoliated Graphene Oxide in an Aqueous Suspension. J. Phys. Chem. C 2010, 114, 14433–14440.

[30] Y. Wang, X. H. Chen, Y. L. Zhong, F. R. Zhu, K. P. Loh, Large Area, Continuous, Few-Layered Graphene as Anodes in Organic Photovoltaic Devices. Appl. Phys. Lett. 2009, 95, 063302.

[31] Y. Wang, S. W. Tong, X. F. Xu, B. .zyilmaz, K. P. Loh, Interface Engineering of Layerby-Layer Stacked Graphene Anodes for High-Performance Organic Solar Cells. Adv. Mater. 2011, 23, 1514–1518.

[32] H. Park, S. Chang, M. Smith, S. Gradečak, J. Kong, Interface Engineering of Graphene for Universal Applications as both Anode and Cathode in Organic Photovoltaics. Sci. Rep. 2013, 3, 1581.

[33] V. C. Tung, L. M. Chen, M. J. Allen, J. K. Wassei, K. Nelson, R. B. Kaner, Y. Yang, Low-Temperature Solution Processing of Graphene-Carbon Nanotube Hybrid Materials for High-Performance Transparent Conductors. Nano Lett. 2009, 9, 1949–1955.

[34] J. H. Huang, J. H. Fang, C. C. Liu, C. W. Chu, Effective Work Function Modulation of Graphene/Carbon Nanotube Composite Films As Transparent Cathodes for Organic Optoelectronics. ACS Nano 2011, 5, 6262–6271.

[35] H.-L. Yip, A. K.-Y. Jen, Recent Advances in Solution-Processed Interfacial Materials for Efficient and Stable Polymer Solar Cells. Energy Environ. Sci. 2012, 5, 5994–6011.

[36] S. S. Li, K. H. Tu, C. C. Lin, C. W. Chen, M. Chhowalla, Solution-Processable Graphene

Oxide as an Efficient Hole Transport Layer in Polymer Solar Cells. ACS Nano 2010, 4, 3169–3174.

[37] Y. Gao, H. L. Yip, S. K. Hau, K. M. Malley, N. C. Cho, H. Z. Chen, A. K.-Y. Jen, Anode Modification of Inverted Polymer Solar Cells using Graphene Oxide. Appl. Phys. Lett. 2010, 97, 203306.

[38] I. P. Murray, S. J. Lou, L. J. Cote,S. Loser, C. J. Kadleck, T. Xu, J. M. Szarko, B. S. Rolczynski, J. E. Johns, J. X. Huang, L. P. Yu, L. X. Chen, T. J. Marks, M. C. Hersam, Graphene Oxide Interlayers for Robust, High-Efficiency Organic Photovoltaics. J. Phys. Chem. Lett. 2011, 2, 3006–3012.

[39] J. Liu, Y. H. Xue, L. M. Dai, Sulfated Graphene Oxide as a Hole-Extraction Layer in High Performance Polymer Solar Cells. J. Phys. Chem. Lett. 2012, 3, 1928–1933.

[40] J. M. Yun, J. S. Yeo, J. Kim, H. G. Jeong, D. Y. Kim, Y. J. Noh, S. S. Kim, B. C. Ku, S. I. Na, Solution-Processable Reduced Graphene Oxide as a Novel Alternative to PEDOT:PSS Hole Transport Layers for Highly Efficient and Stable Polymer Solar Cells. Adv. Mater. 2011, 23, 4923–4928.

[41] J. Kim, V. C. Tung, J. X. Huang, Water Processable Graphene Oxide: Single Walled Carbon Nanotube Composite as Anode Modifier for Polymer Solar Cells. Adv. Energy Mater. 2011, 1, 1052–1057.

[42] X. D. Liu, H. Kim, L. J. Guo, Optimization of Thermally Reduced Graphene Oxide for an Efficient Hole Transport Layer in Polymer Solar Cells. Org. Electron. 2013, 14, 591–598.

[43] D. Yang, L. Y. Zhou, L. C. Chen, B. Zhao, J. Zhang, C. Li, Chemically Modified Graphene Oxides as a Hole Transport Layer in Organic Solar Cells. Chem. Commun. 2012, 48, 8078–8080.

[44] J. Liu, Y. H. Xue, Y. X. Gao, D. S. Yu, M. Durstock, L. M. Dai, Hole and Electron Extraction Layers Based on Graphene Oxide Derivatives for High-Performance Bulk Heterojunction Solar Cells. Adv. Mater. 2012, 24, 2228–2233.

[45] J. X. Geng, T. Y. Zeng, Influence of Single-Walled Carbon Nanotubes Induced Crystallinity Enhancement and Morphology Change on Polymer Photovoltaic Devices. J. Am. Chem. Soc. 2006, 128, 16827–16833.

[46] E. Kymakis, E. Koudoumas, I. Franghiadakis, G. A. J. Amaratunga, Post-Fabrication Annealing Effects in Polymer-Nanotube Photovoltaic Cells. J. Phys. D Appl. Phys. 2006, 39, 1058–1062.

[47] L. Dai, Functionalization of Graphene for Efficient Energy Conversion and Storage. Acc. Chem. Res. 2013, 46, 31–42.

[48] Z. F. Liu, Q. Liu, Y. Huang, Y. F. Ma, S. G. Yin, X. Y. Zhang, W. Sun, Y. S. Chen, Organic Photovoltaic Devices Based on a Novel Acceptor Material: Graphene. Adv. Mater. 2008, 20, 3924–3930.

[49] Y. Li, Y. Hu, Y. Zhao, G. Q. Shi, L. E. Deng, Y. B. Hou, L. T. Qu, An Electrochemical Avenue to Green-Luminescent Graphene Quantum Dots as Potential Electron-Acceptors for Photovoltaics. Adv. Mater. 2011, 23, 776–780.

[50] V. Gupta, N. Chaudhary, R. Srivastava, G. D. Sharma, R. Bhardwaj, S. Chand, Luminscent Graphene Quantum Dots for Organic Photovoltaic Devices. J. Am. Chem. Soc. 2011, 133, 9960–9963.

[51] J. M. Lee, J. S. Park, S. H. Lee, H. Kim, S. Yoo, S. O. Kim, Selective Electron- or Hole-Transport Enhancement in Bulk-Heterojunction Organic Solar Cells with N- or B-Doped Carbon Nanotubes. Adv. Mater. 2011, 23, 629–633.

[52] L. Y. Lu, T. Xu, W. Chen, J. M. Lee, Z. Q. Luo, I. H. Jung, J. I. Park, S. O. Kim, L. P. Yu, The Role of N-Doped Multiwall Carbon Nanotubes in Achieving Highly Efficient Polymer Bulk Heterojunction Solar Cells. Nano Lett. 2013, 13, 2365–2369.

[53] J. M. Lee, B.‐H. Kwon, H. Il Park, H. Kim, M. G. Kim, J. S. Park, E. S. Kim, S. Yoo,
D. Y. Jeon, S. O. Kim, Exciton Dissociation and Charge‐Transport Enhancement in
Organic Solar Cells with Quantum‐Dot/N‐doped CNT Hybrid Nanomaterials. Adv.
Mater. 2013, 25, 2011–2017.
[54] C. Li, Y. H. Chen, S. A. Ntim, S. Mitra, Fullerene‐Multiwalled Carbon Nanotube Complexes
for Bulk Heterojunction Photovoltaic Cells. Appl. Phys. Lett. 2010, 96, 143303.
REFERENCES 189
[55] C. Li, Y. H. Chen, Y. Wang, Z. Iqbal, M. Chhowalla, S. Mitra, A Fullerene‐Single Wall
Carbon Nanotube Complex for Polymer Bulk Heterojunction Photovoltaic Cells.
J. Mater. Chem. 2007, 17, 2406–2411.
[56] B. K. Kuila, K. Park, L. Dai, Soluble P3HT‐Grafted Carbon Nanotubes: Synthesis and
Photovoltaic Application. Macromolecules 2010, 43, 6699–6705.
[57] D. Yu, Y. Yang, M. Durstock, J. B. Baek, L. Dai, Soluble P3HT‐Grafted Graphene
for Efficient Bilayer−Heterojunction Photovoltaic Devices. ACS Nano 2010, 4,
5633–5640.
[58] D. Yu, K. Park, M. Durstock, L. Dai, Fullerene‐Grafted Graphene for Efficient Bulk
Heterojunction Polymer Photovoltaic Devices. J. Phys. Chem. Lett. 2011, 2, 1113–1118.

Wen Lu, Jong-Beom Baek, Liming Dai John. Carbon Nanomaterials for Advanced Energy Systems, 2015,Wiley & Sons, Inc

Holger Spanggaard, Frederik C. Krebs, A brief history of the development of organic and polymeric photovoltaics, Solar Energy Materials & Solar Cells 83 (2004) 125–146

Chapter 7. NanoBetavoltaics

The development of micro-electromechanical systems (MEMS) provoked the search of power sources in range from microwatts to milliwatts which can work for extended period of time. These miniaturized devices (at a scale of micrometers) are often integrated onto silicon chip and typically perform with better precision than similar technologies at the macro-scale. MEMS devices include bio-, chemical and mechanical sensors, optical and RF systems, miniature actuators and gears. MEMS devices are used in satellite systems where they are preferred due to mass/payload requirements, and are also used in various consumer applications including pressure sensors in automobile tires, accelerometers in airbags, and accelerometers in laptop for disabling disk drives [Fang 2006], [Mehregany 2006], [Young 2006], [Wang 2006], [Meier 2008]) (Fig.7.1).

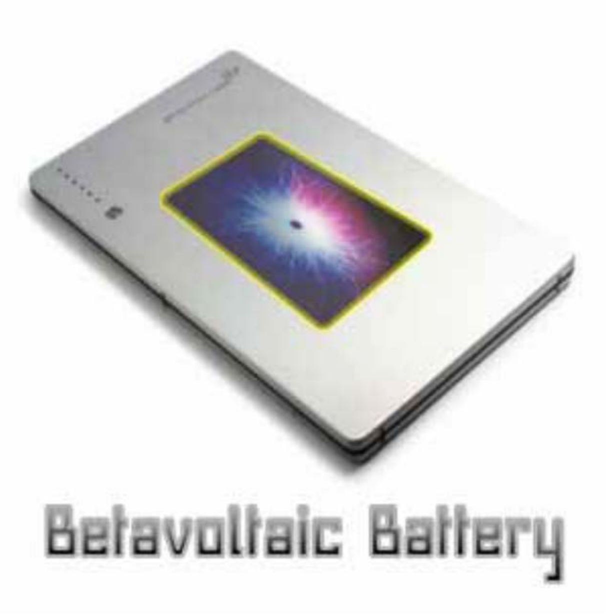


Fig.1 Micro betavoltaic battery

Electricity storage devices for MEMS (other terms- atomic batteries, nuclear batteries, radioisotope generators) are based on the use of the charge particle emissions from radioactive isotope to generate electricity. While a number of conversion schemes can be employed in nuclear batteries (direct charge collection, *betavoltaic*, thermoelectric, thermoionic, thermophotovoltaic, piezoelectric, optoelectric), namely direct voltaic conversion technologies are compatible with the semiconductor manufacturing processes used in MEMS. The direct conversion solid-state voltaic device consists of a p-n junction coupled with charge particle emitting source (e.g. in the form of foil).

Fig. 7.2 explains general operation principle of betavoltaic device which is similar to operation of semiconductor solar cell. The beta particles emitted by radioisotopes lose their kinetic energy by ionizing the semiconductor substrate. Power is generated via the separation of electron-hole pairs by built-in electric field of p-n junction.

Although a large number of radioisotopes emit different particles and electromagnetic radiation, not all are suitable for use in nuclear batteries. Application of radioisotopes with strong gamma emission is restricted due to safety reasons. Alpha emitted isotopes that release energetic He nuclei (typically of 4-6 MeV energy) or beta emitted isotopes (electrons or positrons) with energy more than 200 keV can cause semiconductor materials degradation (e.g. radiation damage threshold for silicon is 150 keV [Gadeken 2007]). Since, today received the priority development the betavoltaic devices.

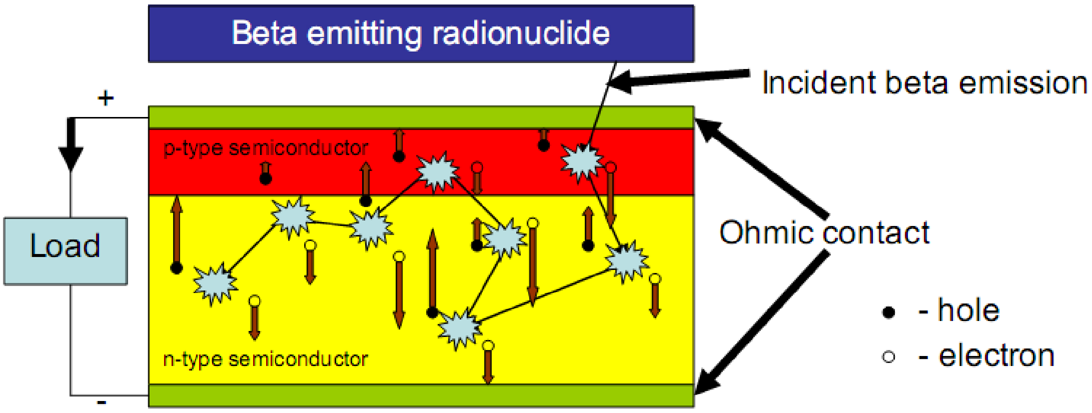


Fig. 7.2. Schema of betavoltaic device [Meier 2008].

Although betavoltaic batteries sound Nuclear they're not, they neither use fission/fusion or chemical processes to produce energy and so (do not produce any radioactive or hazardous waste). Betavoltaics generate power when an electron strikes a particular interface between two layers of material. The process uses beta electron emissions that occur when a neutron decays into a proton which causes a forward bias in the semiconductor. This makes the betavoltaic cell a forward bias diode of sorts, similar in some respects to a photovoltaic (solar) cell. Electrons scatter out of their normal orbits in the semiconductor and into the circuit creating a usable electric current.

The famous isotopes with the long half-lives and low – energy beta emission are the following (Tabl.7.1): ^{63}Ni has a 101 year half-life and 17 keV average beta energy, ^{157}Tb has 99 year half-life and 63 keV average beta energy, ^{151} Sm has 90 year half-life and 76 keV average beta energy, ^{193}Pt has 50 year half-life and 56 keV average beta energy [Ngu 2009]. Half-lives of few tens of years are useful, because then batteries could be designed to whole lifetime of MEMS and would reguire no maintenance. Moreover, electrons emitted from these isotopes has a low energy reducing potential for radiation damage to the n-p junction.

Tritium (symbol T or ^{3}H), a radioisotope of hydrogen, is arguably the ideal candidate for betavoltaic devices, since it decays into helium-3 by beta decay, it has a half-life of 12.3 years and releases electrons with average energy of 5.7 keV and a maximum energy of 18.6 keV [Liu 2006]. It means that tritium decay beta particles pose little radiation damage concern for on-chip energy conversion.

Nuclear batteries using tritium (as well as ^{63}Ni barreries) have power densities an order or two lower than chemical Li-ion batteries and fuel cells (3.4 W/kg for tritium and 0.6 W/kg for ^{63}Ni). However, their energy density 3-5 orders higher compared with chemical batteries and fuel cells (3.10^5 W.h/kg for tritium and 4.10^5 W.h/kg for ^{63}Ni). Therefore, they can supply electricity much longer than chemical power supplies. This is particularly interesting in terms of using nuclear power batteries in laptops during tens years without having to recharge.

TABL.7.1 Beta-Emitting Isotopes

Isotope	$T_{½}$ (year)	**Emission**	E_{ave} (keV)	E_{max} (keV)	P_{sp} (μW/Ci)	**Element**
^{3}H	12.3	β^-	**5.7**	**18.6**	**34**	Hydrogen (tritium)
^{63}Ni	100	β^-	**17.6**	**62**	**100**	nickel
^{14}C	5710	β^-	**59**	**156**	**290**	carbon
^{147}Pm	2.7	β^-, weak γ	62	224	367	promethium
^{45}Ca	0.44	β^-	77	257	456	calcium
^{208}Tl	3.8	β^-, X	243	763	1440	thallium

In spite of these advantages of tritium batteries and fact that tritium is readily and economically accessible, there are some technological problems since tritium exist in the gaseous form and has low power densities of 87 μW/cm^3. Stable incorporation of tritium into semiconductor materials at sufficiently high power density is challenging. Possible approaches to high-density tritium incorporation include metal tritide films, formed by tritiation of evaporated hydride forming metals [Liu 2006], and tritiated amorphous silicon films, deposited using a PECVD technique [Kosteski 2003]. Both methods are relatively complicated.

To improve conversion efficiency of conventional planar (2D) betavoltaic diode the radioisotopes can be encapsulated within multiple layers of semiconductors in a stacked configuration [Meier 2008]. A porous semiconductor materials could also be engineered so that radioisotopes in liquid or gaseous state could occupy a greater surface area in the semiconductor [Guo 2007] [Sun 2005a] [Sun 2005b] [Liu 2009].

The process fabrication of 3D porous silicon p-n diode is described [Sun 2005a] [Sun 2005b] and presented in Fig. 7.3. A macroporous Si of 50 μm thick with pore dimension of 0.8-2 μm is formed by electrochemical etching (current density of 2 mA/cm^2 for 3 h in 4 wt% HF in dimethylformamide) of 20-30 Ω.cm boron doped silicon wafer of <100> orientation. After standard RCA cleaning, p-n junctions are formed from solid P_2O_5 source diffusion at 1000°C in nitrogen for 18 h. The silicon wafer is soaked for 9 min in the same condition to make a shallow n^+ layer of 200 nm thick. Subsequently, the n^+ layer formed on the backside of silicon wafer is removed by SF_6 plasma etching. After the samples are dipped in the buffered oxide etchant, 50 nm and 200 nm aluminium layers are evaporated on the porous side and backside respectively under high vacuum ($<10^7$ Torr). The samples are sintered at 420°C for 15 min to avoid the Schottky barrier at the interface of the aluminium and silicon. Then the wafer is rinsed with acetone, isopropyl and de-ionized water.

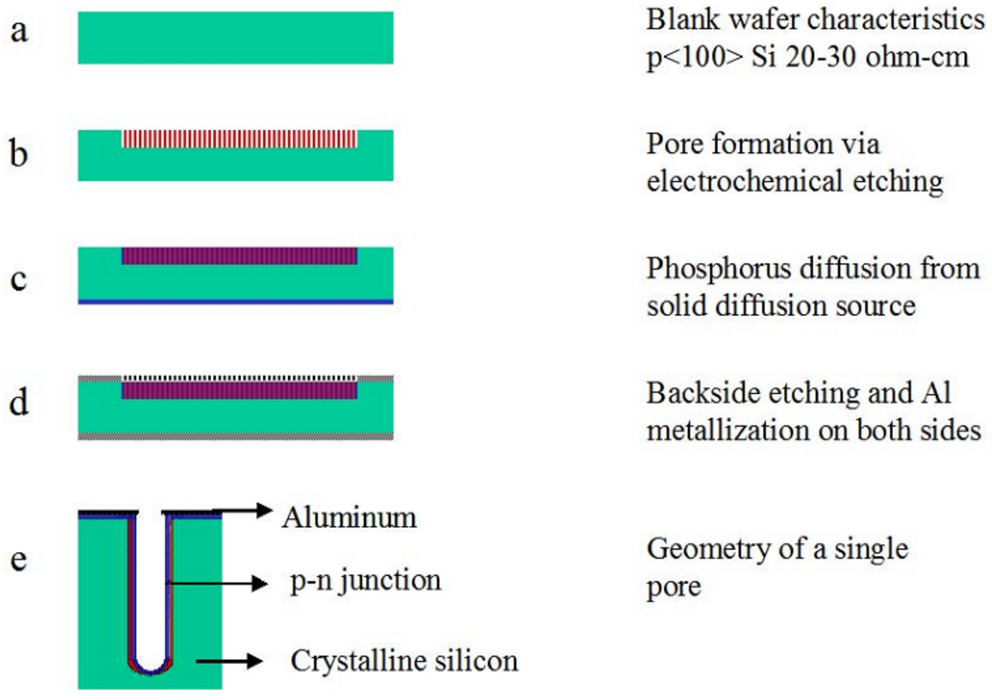


Fig. 7.3. Procedure of fabrication porous silicon betavoltaic device [Sun 2005b]

Pore geometry of the PSi layer was obtained using high-resolution Scanning Electron Microscopy (SEM). SEM viewgraphs are shown in figure 7.4. The average area around each pore and average pore diameter were estimated by measuring the distribution of pore diameters in a 15 x 22.5 μm^2 area. The observed 190 pores gave the average area surrounding each pore as 1.78 μm^2. The average pore diameter, d_{avg}, was 0.837 ± 0.125μm and the average pore depth,

h_{avg}, was 43.1 ± 1.18 μm estimated from the SEM top view and cross-sectional view, respectively. The analysis of the pore geometry together with the area enclosed by the indium seal in each CTF(Chip Test Fixtures) gave the estimated number of pores exposed to tritium as 103 million. The corresponding total internal surface area was 116 cm^2 and the pore volume was 2.43×10^{-3} cm^3. The fraction of the surface containing the pores was 0.31 and the remaining planar surface fraction was 0.69.

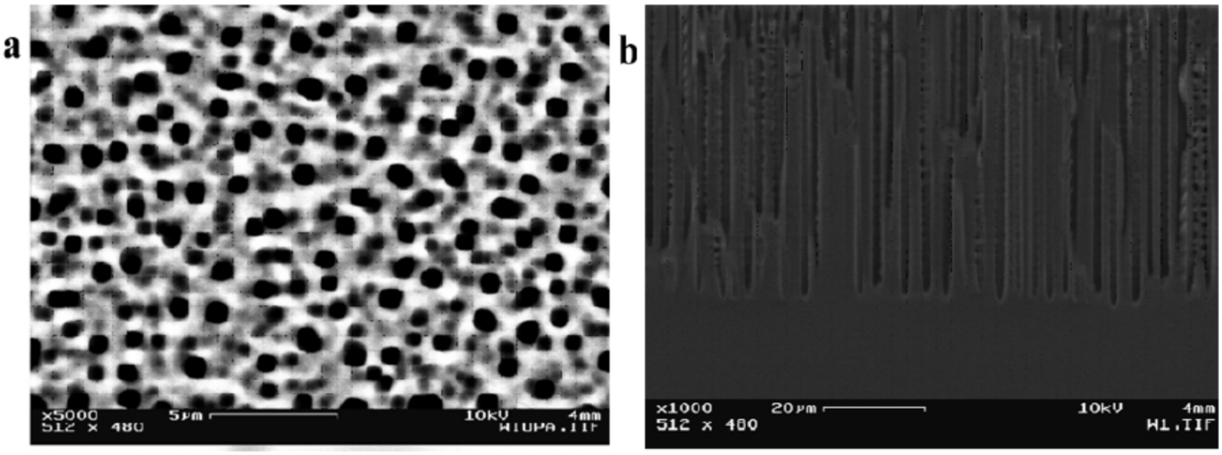


Fig.7.4. SEM images of device geometry.

Data analysis of both betavoltaic and photovoltaic devices is very similar other than the source of energy and generation of excess charge carriers. The energy conversion efficiency, $\eta = P_{ex} FF / P_{\beta}$, where P_{β} is the total available power contained in the reservoir or pore volume. Representative I-V curves for porous and planar diodes are shown in Fig. 7.5 (a) before and after loading with tritium gas. The presence of tritium displaces the I-V curve into 4th quadrant (in the greater extend for 3D diode) indicating that electrical power is being generated. Short circuit current per Curie of tritium is more than an order of magnitude greater for 3D diode compared to the conventional 2D diode geometry (Fig. 7.5 (b,c)).

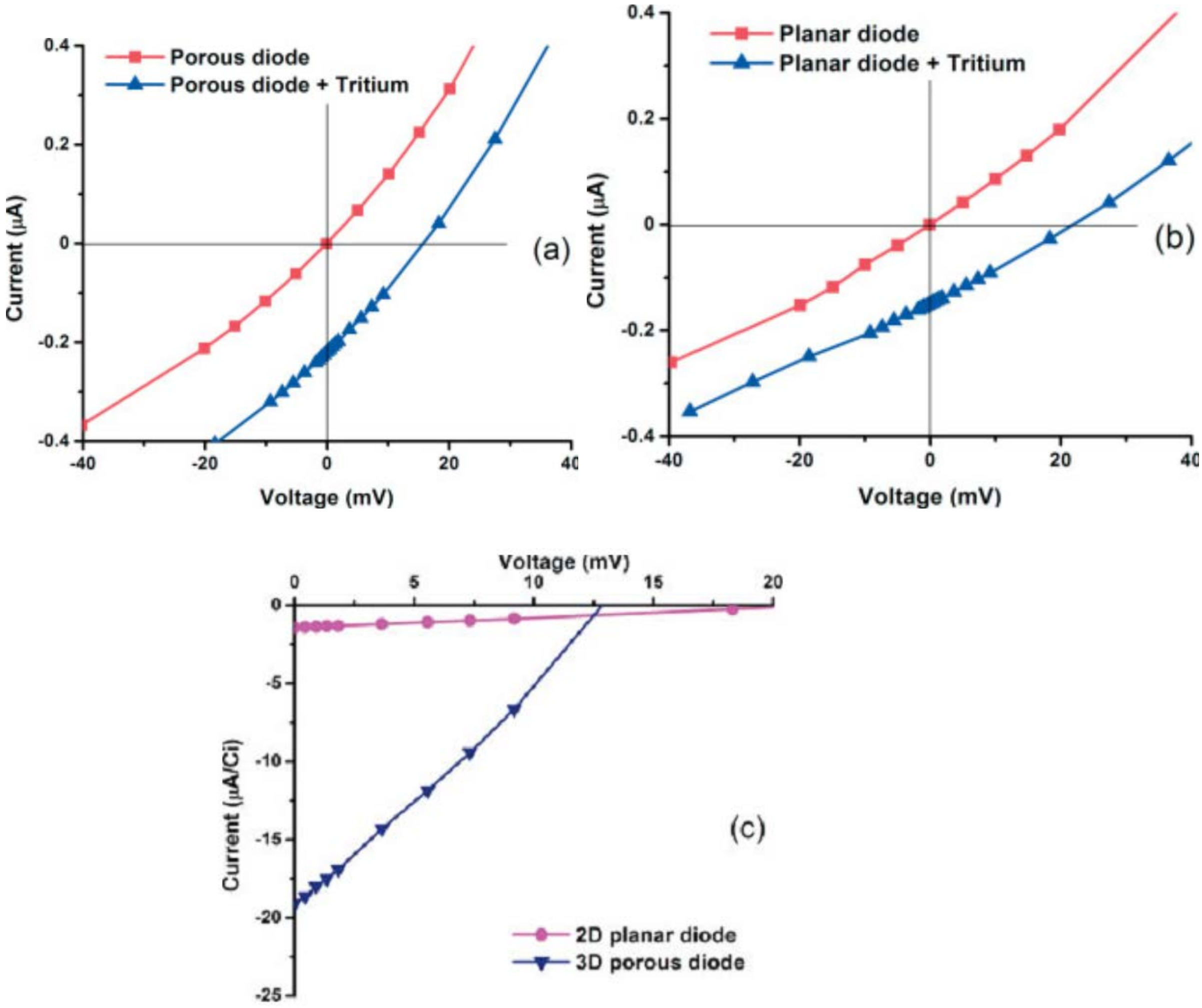


Fig. 7.5. I-V charactersitics of 2D and 3D diodes without and with tritium gas (232 mCi at 1 atm) [Sun 05b] (a,b) and current per unit radiactivity versus voltage (c) [Sun 2005a]

3D Diode in Wafer Test Fixture is presented in Fig.7.6. Extraordinarily Efficient DEC™ Cell (patented) is shown in Fig.7.7. The 3D architecture holds the tritium energy source embedded in a pore walls (Si matrix shown in pink.) 3D pores are the trapping mechanism for the emitted electrons from source (Polymer layer is invisible.) The pores are lined with a conversion layer, allowing Direct Energy capture and Conversion for high efficiency. The emitted electrons bounce around in the pores until absorption occurs. Practically all the tritium decay electrons are captured for conversion.

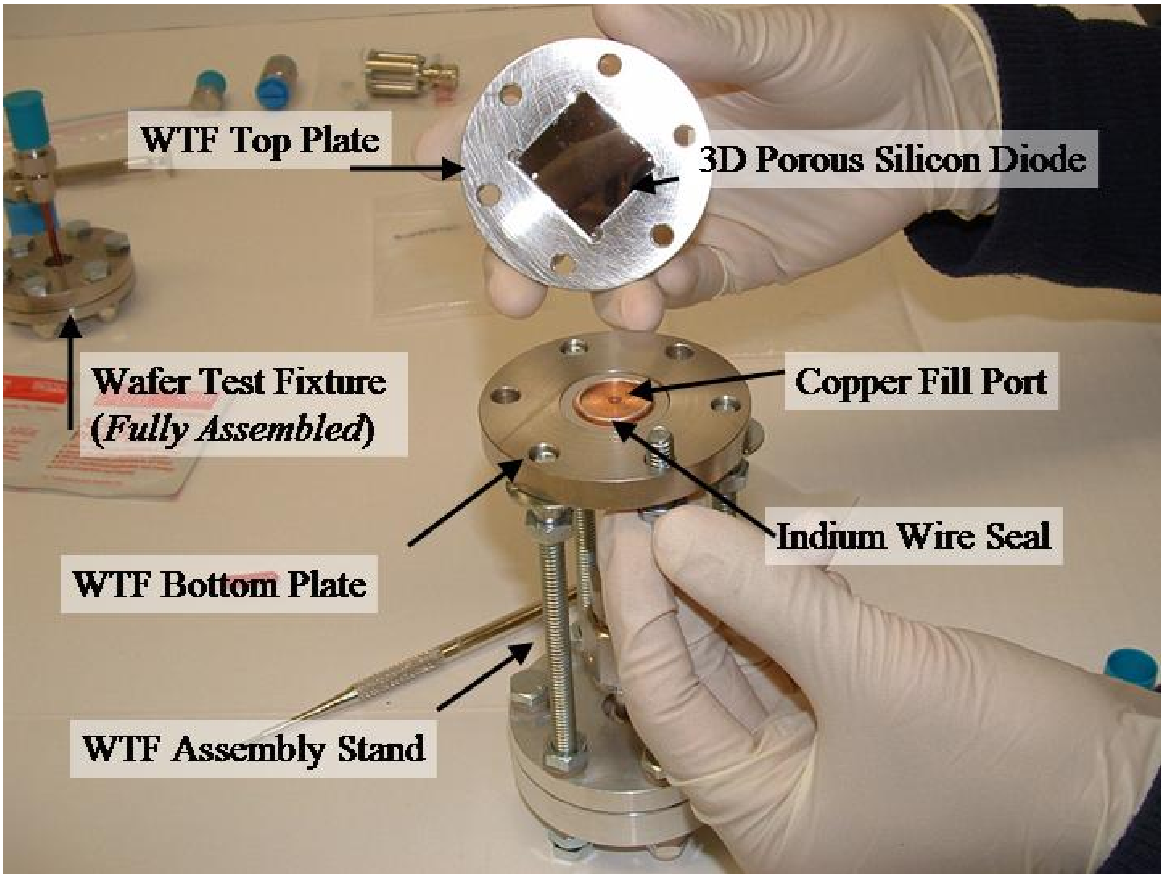


Fig.7.6. 3D Diode in Wafer Test Fixture

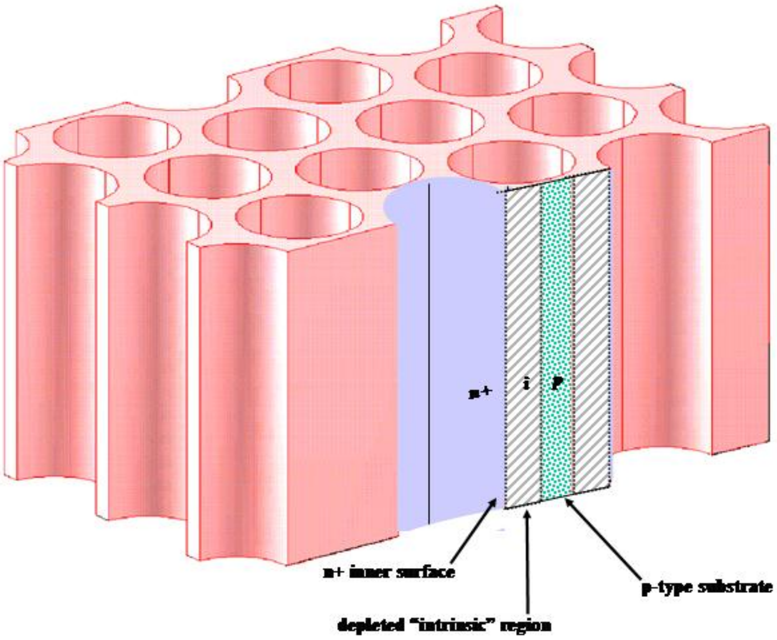


Fig.7.7. The patented DEC™ Cell captures and converts energy in one efficient step

The average observed betavoltaic energy conversion efficiencies achieves 0.023% for 2D diodes and 0.22% for the porous diodes, respectively. These rather low values are mostly attributed to the correspondingly low power available from the small volumes of tritium gas, however, the ratio of the energy conversation efficiency for 3D and 2D planar devices is $\eta_{3D}/\eta_{2D} \approx 10$. This order of magnitude increase in efficiency in the pore channels corresponds to the almost unity probability that tritium decay electrons enter the betavoltaic conversion layers (p-n junctions) in the pore walls. Given that the length to diameter ratio of the pores is ~60:1, whatever the direction of flight, the average path length prior to striking the pore walls is no more than 2 μm. Moreover, the tiny pore diameter eliminates self absorption in the tritium gas. Although it is estimated that~1/3 of the incident betas may be elastically backscattered, this loss is negligible in the pores owing to the close proximity of “adjacent” betavoltaic junctions [Sun 2005a].

It is expected that significant efficiency gains in pre-commercial prototypes of tritium batteries can be achieved via increase energy density using a tritiated polymer instead tritium gas filling the pore space, increase pore channel density, lengthen pore channel dimension, reduce series resistance and minimize parasitic resistances [Gadeken 2007]. The Basis for significant efficiency gains in pre-commercial prototypes are the following:

-Increase energy density using a tritiated polymer (not tritium gas filling the pore space). Estimated improvement factor: ~20,

(Solid energy density >1000 $\times$ *T_2 gas)*

-Increase pore channel density to 25M pores/cm^2 (increase from 16M pores/cm^2)

-Lengthen pore channel dimension to 250-300 μm (up from 50 μm). Estimated improvement factor: ~8, Overall energy density improvement factor: ~160

-Increased energy density improves semiconductor Fill Factor ~70%

-Reduce series resistance and minimize parasitic resistances. Overall efficiency target: 7-9% or greater.

-Target electrical power density: 50 - 125 μW/cm^3 for prototype BetaBatteries.

In order to increase the overall output current of our direct-energy-conversion (DEC) cell we laid out diodes in parallel. The parallel structure is assembled by stacking the diodes in separated 63Ni source pairs. In this arrangement, the copper leads to the p-layers were all connected together with copper tapes to make a single positive leads. The same was done with the leads to the n-layers to make the negative lead of the stacked structure. A schematic of the stacked cell is shown in figure 7.8.

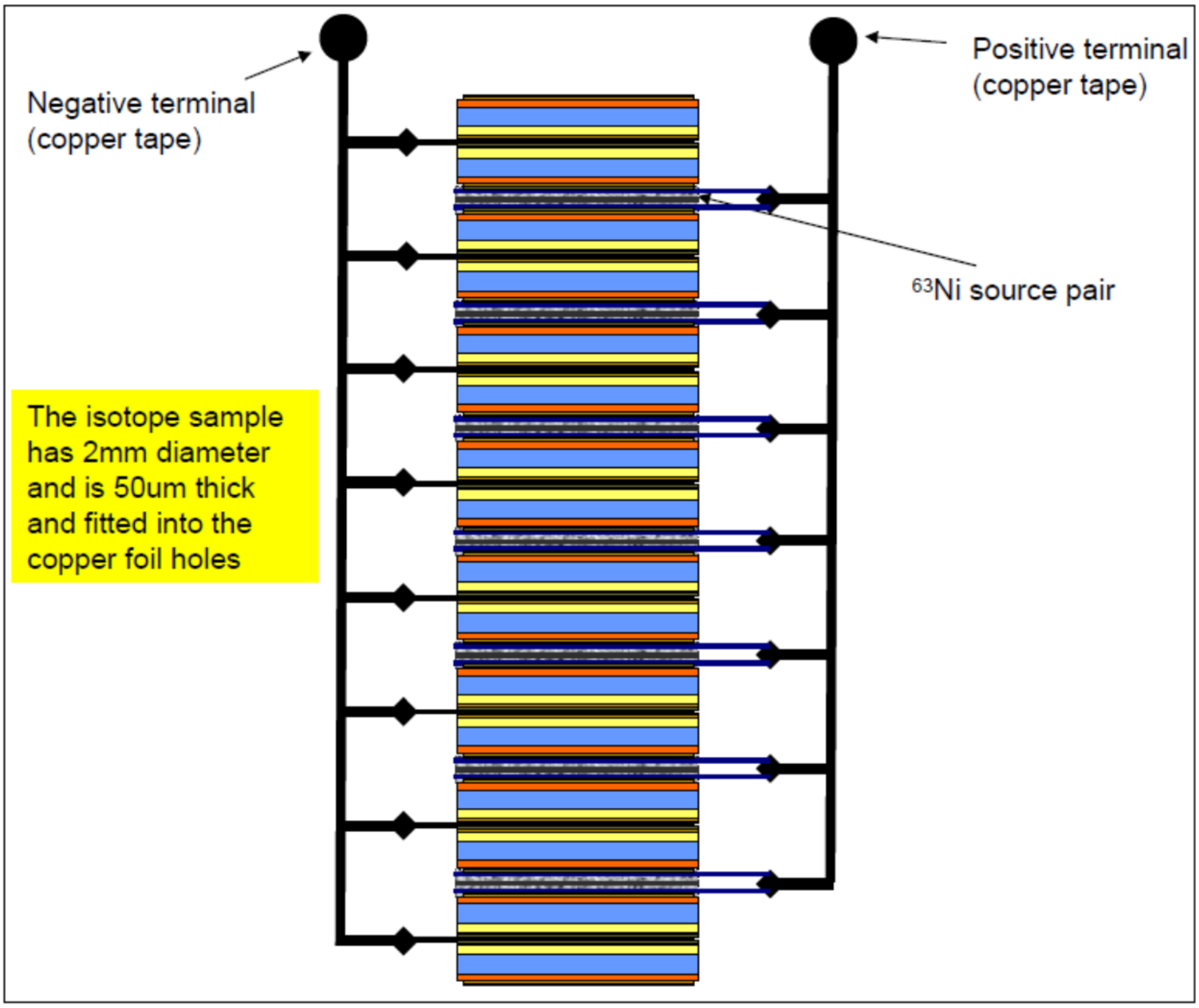


Figure 8. Eight stacked diode pairs with 14 sources (7 pairs).

Last time a new concept of betabateries are developed [Gadeken 2007]. Among them- Hybrid Beta Batteries, which includes tritiated 3D porous silicon diodes and thin film rechargeable lithium battery.

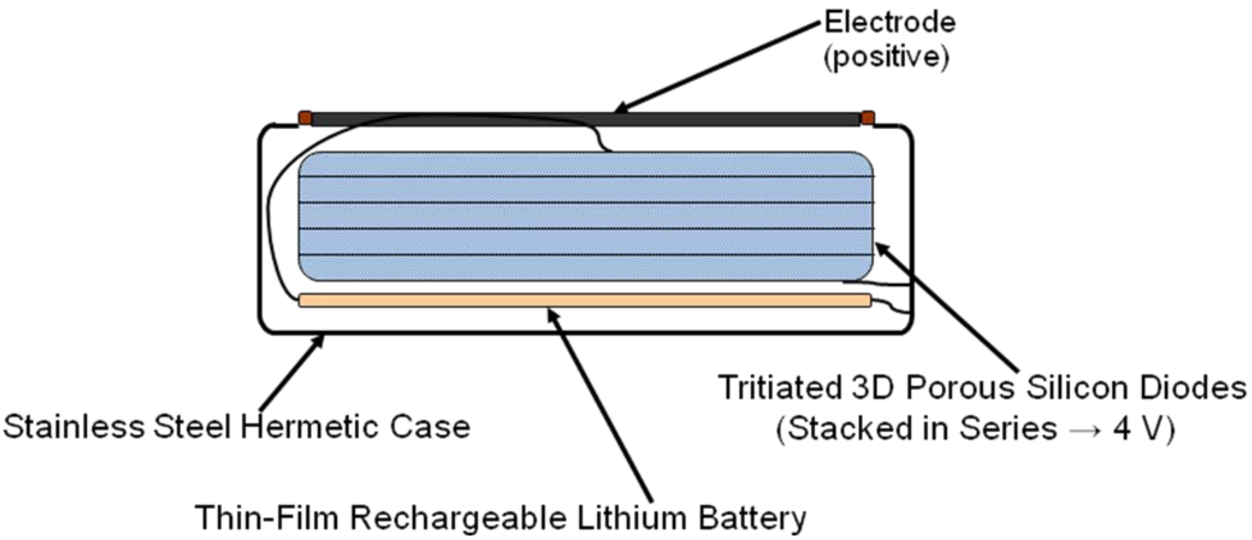


Fig.7.9 Hybrid Beta Batteries

As can be seen from Fig.7.10, BetaBatteries show long life at low power, chemical battery has high power at short life , but Hybrid Betabatteries show middle parameters and possibility to run in mode of pulse power delivery.

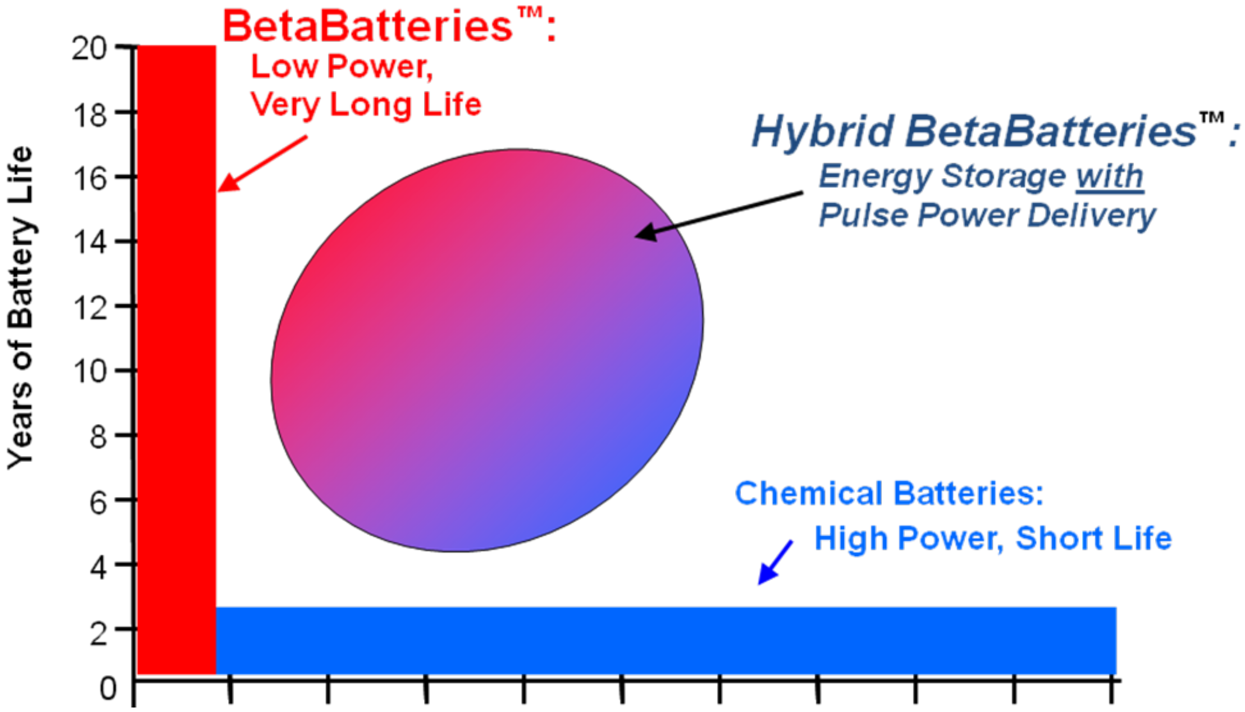


Fig.7.10 The comparison of main parameters of BetaBatteries, Hybrid Betabatteries and Chemical batteriers.

References

Fang, F. Z., Liu, K., Kurfess, T. R. and Lim, G. C. (2006). Tool-based micro machining and applications in MEMS, MEMS/NEMS Handbook - Techniques and Applications, ed. Leondes, Cornelius T. , Springer – Verlag, 3, pp. 63-125.

Mehregany, M., Fu, X.-A., and Chen, L. (2006). Silicon carbide micro/nano systems for harsh environment and demanding applications. *NSTI Nanotechnol. Conf. Trade Show* , 3, pp. 471-474.

Wang, P. (2006), Cell-based biosensors and bio-MEMS for its new applications. *Xiyou Jinshu Cailiao Yu Gongcheng*, 35, pp. 47-50.

Young, D. J., Zorman, C. A., and Mehregany, M. (2006). MEMS/NEMS Devices and Applications, in *The Springer Handbook of Nanotechnology - 2nd Edition*, B. Bhushan, ed., Springer-Verlag, Berlin, pp. 415-442.

Guo, H., Yang, H. and Zhang, Y. (2007). Betavoltaic Microbatteries using Porous Silicon. *Micro Electro Mechanical Systems, MEMS. IEEE 20th Int. Conf.*, Kobe, Japan, pp.867 – 870.

Meier, D.E. (2008). Radionuclide production for radioisotope micro-power source technologies, PhD thesis, University of Missouri-Columbia, 107 p.

Gadeken, L. L. (2007).Tritiated 3D Diode Betavoltaic Microbattery, *IAEA Workshop Advanced Sensors for Safeguards*, Mito, Japan.

Liu, B., Chen, K.P., Kherani,N. P., Kosteski, T., Costea, S., Zukotynski, S. and Antoniazzi, A. B. (2006). Tritiation of amorphous and crystalline silicon using T2 gas. *Appl. Phys. Lett*. 89, p.044104.

Ngu ,Y., and Litz , M. (2009). Study of Beta Radioisotopes Direct Energy Converters , *ARL-TR-4969*, 16 p.

Kosteski, T., Kherani, N. P., Stradins, P., Gaspari, F., Shmayda, W. T., Sidhu, L. S., and Zukotynski S. (2003), Tritiated Amorphous Silicon Betavoltaic Devices , IEE Proc. Circuits, Devices and Syst., special issue on Amorphous and Microcrystalline Semiconductor Devices, 150, pp.274-281.

Liu, B., Chen, K.P., Kherani,N. P., Kosteski, T., Costea, S., Zukotynski, S. (2009) Power-scaling performance of a three-dimensional tritium betavoltaic diode, *Appl. Phys. Lett.,* 95 pp. 233112.

Sun, W. , Kherani, N. P. , Hirschman, K. D., Gadeken, L. L., and Fauchet ,P. M. (2005b).Harvesting betavoltaic and photovoltaic energy with three dimensional porous silicon diodes, Mat. Res. Soc. Symp. Proc. 836, pp.285-290 .

Sun, W. , Kherani, N. P. , Hirschman, K. D., Gadeken, L. L., and Fauchet ,P. M. (2005a). A three-dimensional porous silicon p-nliode for betavoltaics and photovoltaics, Adv.Mat., 17, pp.1230-1233.

Chapter 8 Environmentally Friendly Energy Sources on nanomaterials

8.1 Introduction

Hydrogen is a clean-burning energy carrier and an important chemical feedstock. Benefits of hydrogen sources are the following: it is energy-intensive, environmentally friendly and technologically flexible energy source. However, realization of a possible hydrogen economy presents a trifecta of challenges to the scientific community. Hydrogen must firstly be produced, subsequently stored and then delivered to an end-use application. It is low density gas (0.09 kg / m^3) and for running of 100 km the car based on hydrogen fuel must have the volume of gas ballons ~ 11 m^3 (fig.8.1).

Fig.8.1 Hydrogen vesicle

Traditional storage methods make use of compressed gas cylinders, which require energy intensive processing and have inherent safety risks, notwithstanding the significant weight and volume requirements of such equipment. These factors become increasingly important for mobile applications. Other important considerations include the ease of regeneration of spent material, and overall system cost, operating temperature and stability.

Three main options exist for storing hydrogen (Tzimas *et al*. 2003):

(i) as a highly compressed gas,

(ii) as a cryogenic liquid,

(iii) in a solid matrix.

Most automakers are considering either the high-pressure gaseous hydrogen (passenger vehicles) (fig.2a) or cryogenic liquid (fig.2b), for example, for rockets applications (fig.3)

Fig.8.2 Balloons of highly compressed gas (a) and cryogenic hydrogen (b)

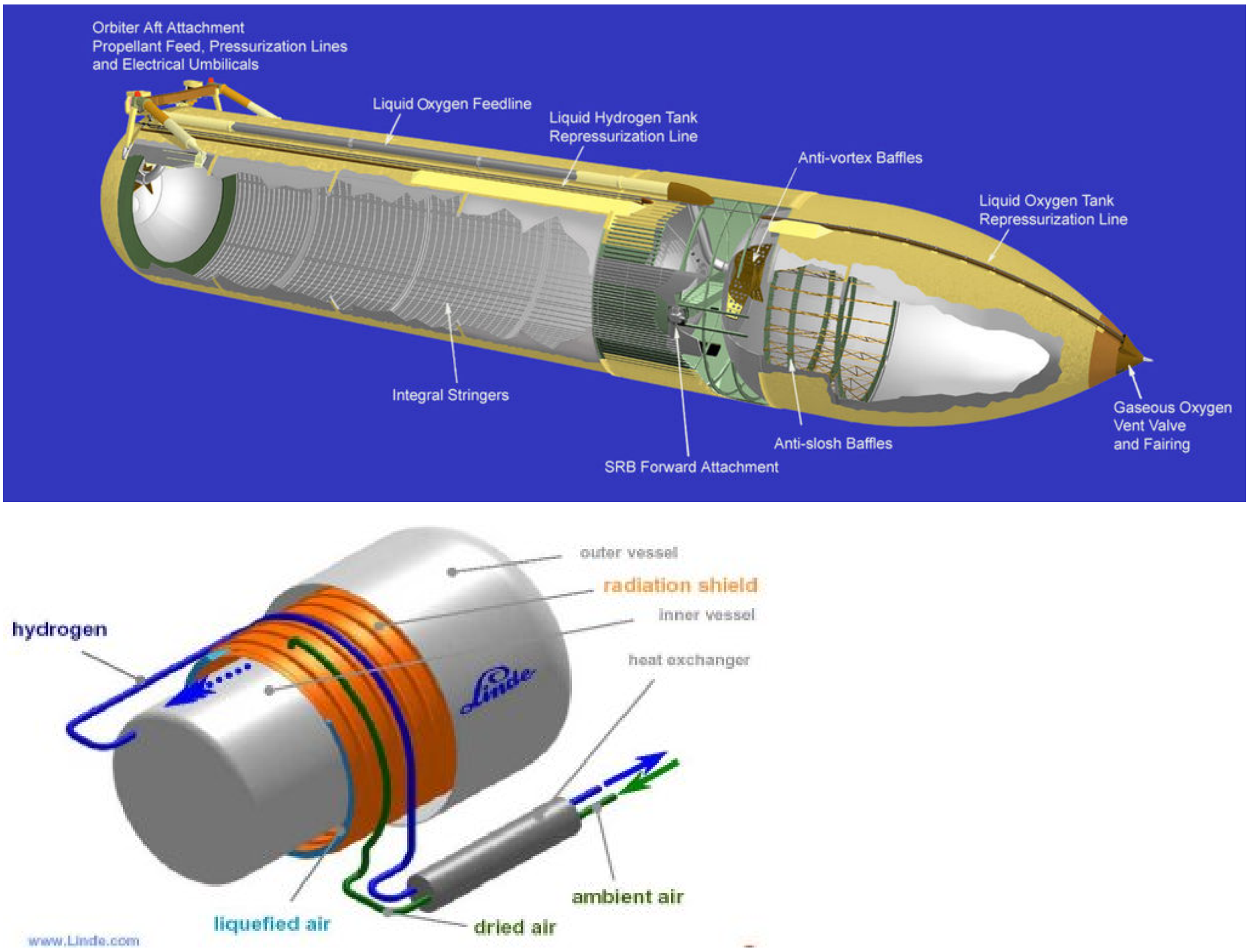


Fig.8.3 Construction of cry tanks for rockets

However, these first two technologies are fraught with public perception issues on safety and cannot, of course, be applied in a large field of portable devices: mobile telephones, portable computers and other gadgets. Other issues of these technologies need to be also addressed, including compression costs and safety, liquidation costs, and dormancy. Advantages of cryogenic hydrogen sources is large hydrogen content ~ 20% by weight of small cryo tanks and 86% by weight in large tanks Aero Space. However, there are significant losses by evaporation (share% per day) and necessity of cryogenics temperatures**.**

The problems of compressed balloon arise from the necessity to increase the thickness of balloon walls at increasing of compressions (see eq.8.1).

$$\frac{S}{D} = \frac{\Delta P}{2\sigma + \Delta P}, \tag{8.1}$$

where S is wall thickness, D is outer diameter, ΔP is overpressure, σ is tensile strength of the material. Due to this reason hydrogen weight decreases at increasing of overpressure (fig.8.4) .

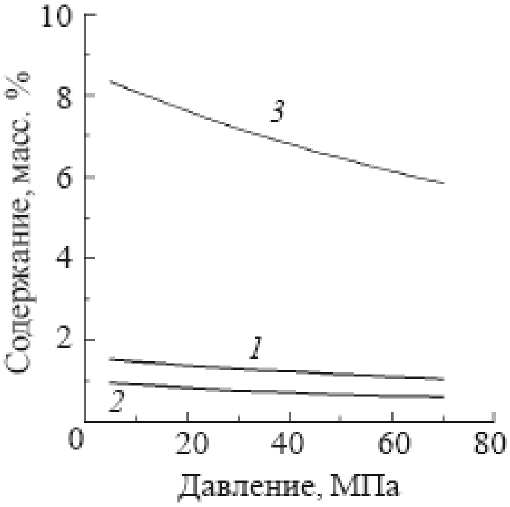


Fig.8.4 Hydrogen weight vs overpressure.

Therefore, it is becoming increasingly accepted that the solid matrix method of hydrogen storage is the only option that has any hope of being efficient for hydrogen storage especially for portable device applications. For mobile applications (fig.8.5), and as such the 2015 DOE (Department of Energy) targets for hydrogen storage are 9 wt% and 81 g L^{-1} for gravimetric and volumetric system capacities respectively (Dalebrook *et al.* 2013), (Krishna *et al.* 2012).

Fig.8.5 Mobile applications of fuel cells with hydrogen batteries

8.2Physisorption of Hydrogen in Porous Matrixes

Nowadays, a wide range of materials are currently being considered as potential reversible hydrogen storage media. Among them- microporous adsorbents, which physisorb molecular hydrogen at low temperatures, the interstitial hydrides, which are reactive metals reversibly absorbing dissociated atomic hydrogen into their bulk as an interstitial, the complex hydrides that bind atomic hydrogen either covalently or ionically and release it via solid state decomposition, and some alternative storage materials (Broom 2011), (Krishna *et al.* 2012).

One of the approaches for reversible hydrogen storage is based on physisorption of hydrogen molecules onto the specific surface of a solid matrix. Hydrogen can be stored in its molecular form by physical adsorption on the surface of a porous solid material, and typical gravimetric capacities for a range of materials are identified in Table 1 (Dalebrook *et al.* 2013).

Table 8.1. Properties of selected physical hydrogen storage systems

[a] Values for pure hydrogen are system gravimetric capacities which measure storage relative to the container weight. [b] Polymers of intrinsic microporosity. [c] Hyper-cross-linked polymers. [d] Covalent organic frameworks. [e] Metal–organic frameworks.

Dalebrook A.F. et al., *Chem. Commun.* **49**, 8735, 2013.

Storage medium	Temperature (C)	Pressure (bar)	Capacity (wt%)
Compressed hydrogen gas	25	200	16[a]
		Up to 700	28[a]
Carbon nanotubes	27	1	0.2
	25	500	2.7
	-196	1	2.8
Graphene oxide	25	50	2.6
PIMs[b]	-196	10	2.7
HCPs[c]	-196	15	3.7
COFs[d]	-196	70	7.2
Zeolites	25	100	1.6
	-196	16	2.07
MOFs[e]	25	50	8
	-196	70	16.4
Clatrate hydrates	-3	120	4

This process is based on Van der Waals interactions of the hydrogen molecules with several atoms at the surface of a solid. Because of the weak interactions, significant physisorption is only observed at low temperatures (<273°K) and the maximum amount of adsorbed hydrogen is proportional to the specific surface area of the adsorbent solid.

8.3 Carbon materials

Activated carbon-based materials, such as nanotubes, fullerenes and graphene, are amorphous carbonaceous materials possessing a high degree of porosity and large specific surface areas (fig.8.5-8.7). The technical and economical potential of activated carbon materials for cryogenic hydrogen storage purposes was first proposed by Schwarz et al [1] In the past two decades, both carbon nanotubes and nanofibers have attracted great interest as materials for hydrogen storage. In 1997, Heben et al. first measured the hydrogen adsorption capacity on soot containing 0.1 to 0.2 wt% single-walled carbon nanotubes [2] and many carbon materials with high hydrogen adsorption capacity were subsequently reported [3-6]. More recent investigations demonstrated that the hydrogen storage capacity of both purified single-walled carbon nanotubes and graphitic nanofibers is limited to 0.2 wt% under ambient conditions [7]. In one study, loadings could be increased to a maximum of 2.7 wt% but only at extremely high hydrogen pressures of 500 bar [8].

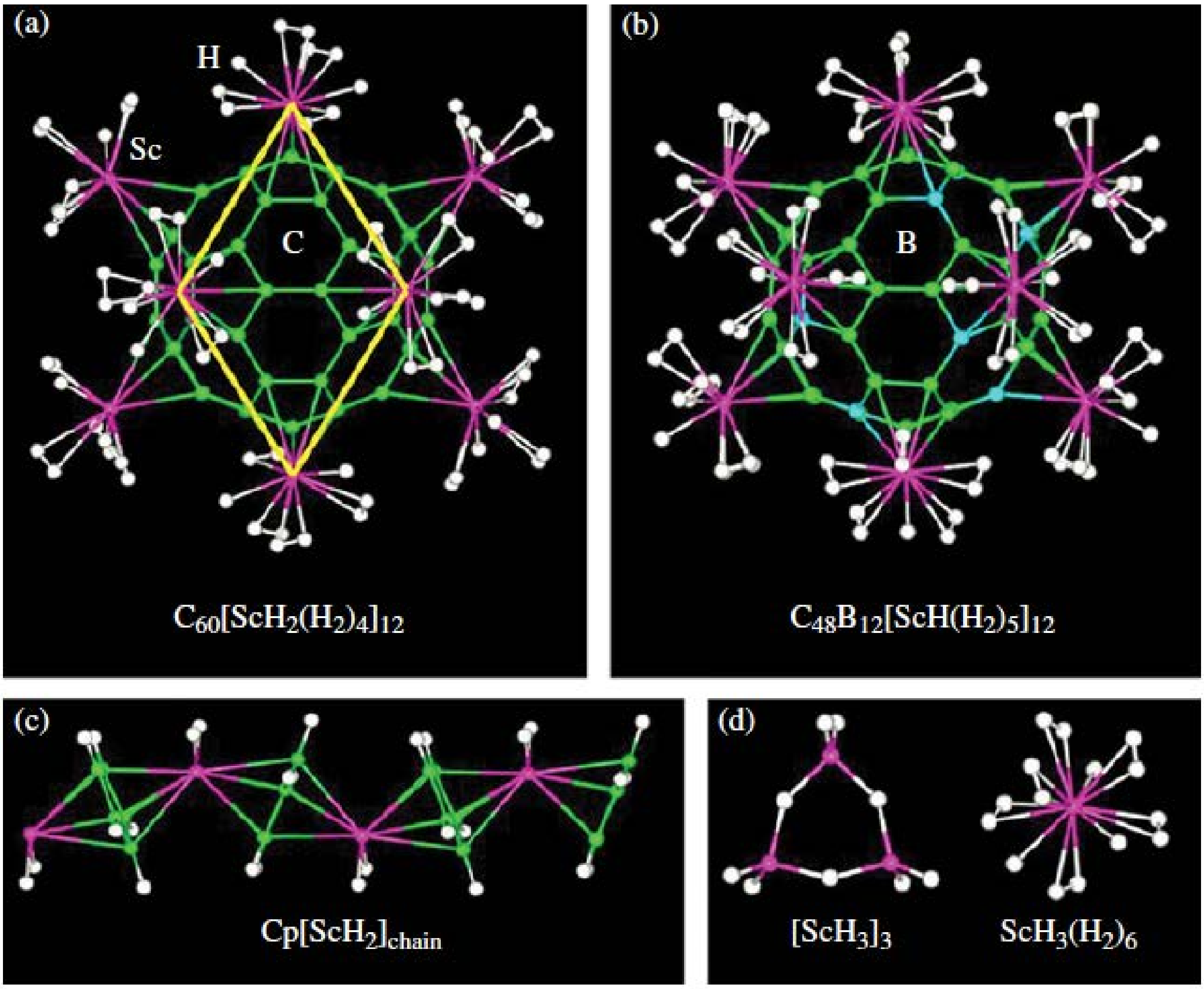


Fig.8.5 (a) $C_{60}[ScH_2(H_2)_4]_{12}$, (b) $C_{48}B_{12}[ScH(H_2)_5]_{12}$, (c) Cp[$ScH_2$]chain, and (d) $[ScH_3]_3$ (left) and $ScH_3(H_2)_6$ (right).

Chahine and Bose reported a storage capacity of 5 wt% at -196 °C and 20 bar pressure for AX-21, an activated carbon with 3000 m^2 g^{-1} specific surface area

[9]. A gravimetric loading of 4.5 wt% was obtained for a comparable activated carbon under similar conditions [10]. Yang and Mokaya recently described the hydrogen storage capacity of a zeolite-like carbon material presenting a high specific surface area of 3200 m^2 g^{-1} and a pore size distribution in the range of 0.5–0.9 nm [11]. The use of an ordered b-zeolite template during the synthesis gave fine control of the pore size distribution, particularly in the micropore range. This solid can store up to 6.9 wt% hydrogen at -196 ^oC and 20 bar, a notable increase relative to activated carbon facilitated by an optimized pore structure.

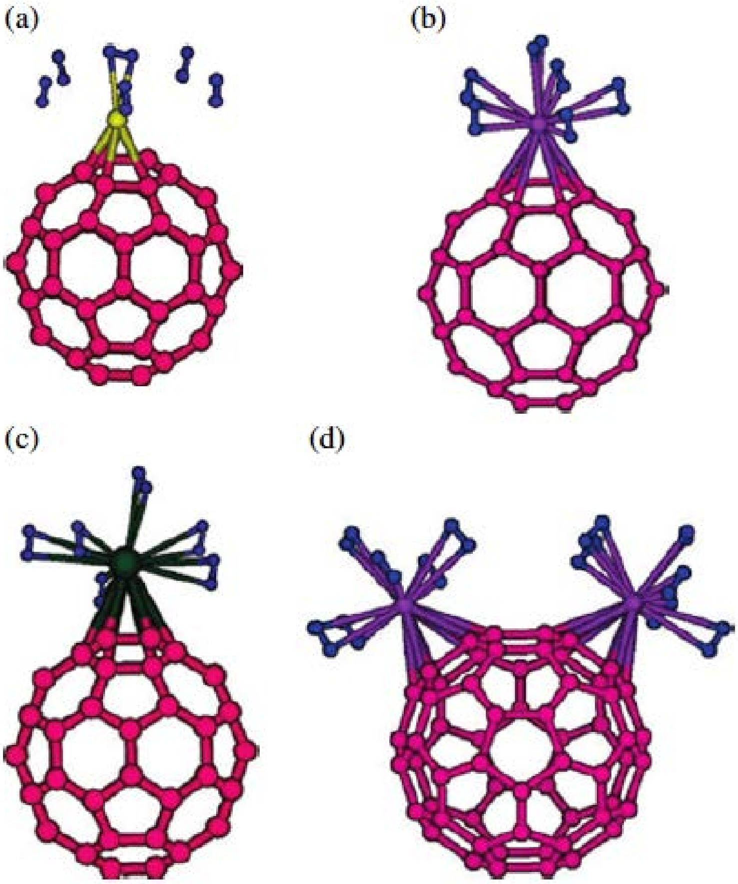


Fig.8.6 Adsorption of hydrogen molecules on alkali metal- doped fullerenes. (a) $C_{60}Li(H_2)_2$, (b) $C_{60}Na(H_2)6$, (c) $C_{60}K(H2)_6$, and (d) $C_{60}Na_2(H_2)_{12}$.

The influence of pore size on the carbon–hydrogen interaction was determined for a series of carbide-derived carbon networks with tunable pore dimensions [12-14]. The results indicate that small pores (<1 nm) lead to an enhanced hydrogen– carbon interaction and an increase in the hydrogen storage capacity at low pressures. Xu et al. later prepared a nanoporous carbon structure with a zeolite-type MOF as both the precursor and template [15]. This material displayed a very high BET surface area of 3405 m^2 g^{-1} and a hydrogen storage capacity of 2.77 wt% at -196 ^oC under atmospheric pressure.

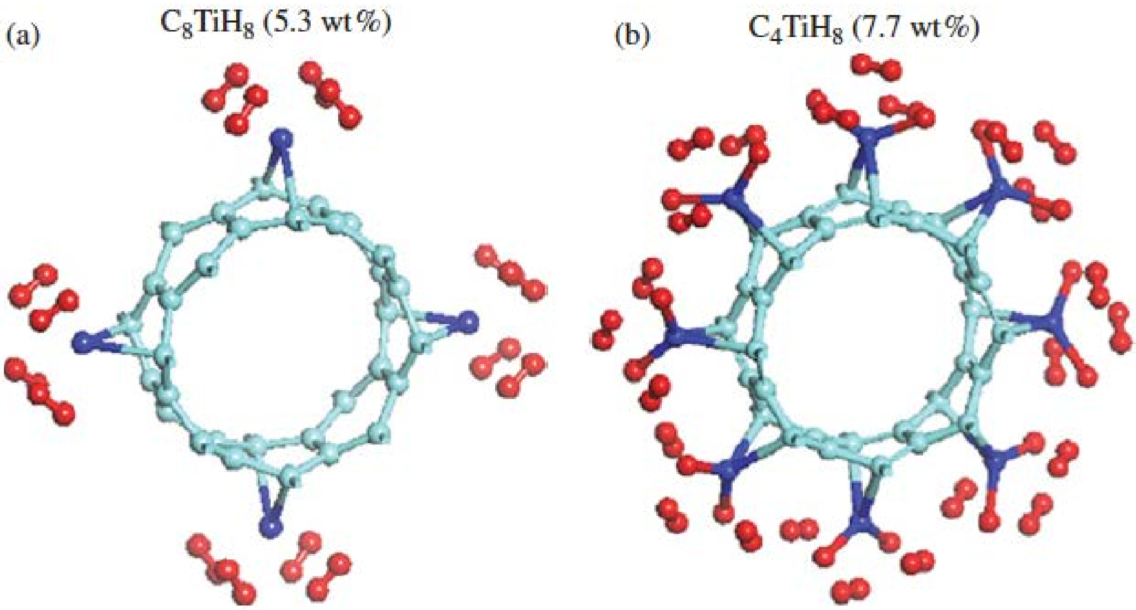


Fig.8.7 Two high-density hydrogen coverages on a Ti-coated (8,0) nanotube. (a) and (b) have different Ti modification ratios.

More recently, graphene has gained consideration as a hydrogen storage material. Tozzini and Pellegrini describe the curved surface of graphene as a tunable substrate, where control of the local curvature induces either adsorption or desorption of hydrogen [16]. Pillared graphene multilayers and the addition of surface metal atom sites were also suggested as promising research directions. One such graphene oxide material was realized by Aminorroaya-Yamini et al. boasting a hydrogen capacity of 2.6 wt% at room temperature [17]. Part of the hydrogen uptake was attributed to the columns of multi-walled carbon nanotubes (MWCNTs) that cooperate synergistically with the graphene oxide surface to increase the adsorption properties. The stored hydrogen in carbon structures as a function of their specific area is shown in Fig.8.8.

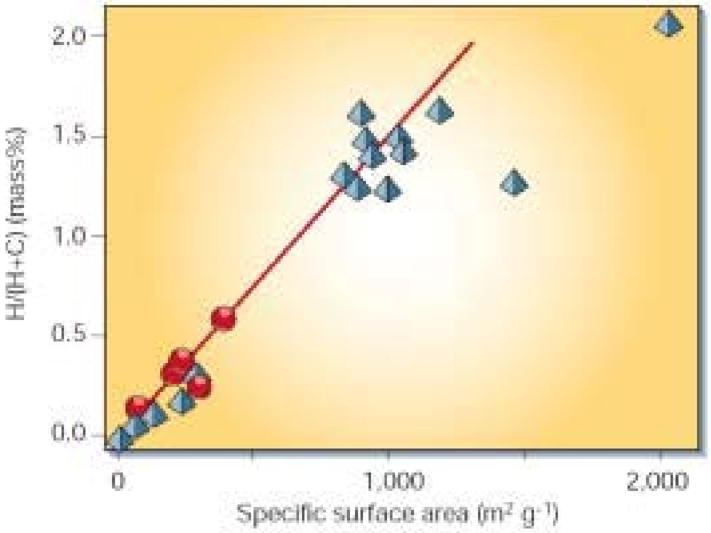


Fig.8.8 Reversibly stored hydrogen in carbon structures as a function of their specific area – circles indicate nanotubes (best-fit line given) while triangles correspond to other nanostructured carbon samples

8.4 Zeolites

Zeolites comprise a regular open pore microstructure of variable framework type and candidates have been trialed for hydrogen storage applications over the last several decades. In 1995 Ernst et al. suggested that the hydrogen storage capacity of a range of zeolites at room temperature was less than 0.1 wt% [18]. Subsequently, Harris et al. probed the influence of different metal cations. A maximum gravimetric storage capacity of 2.19 wt% at 15 bar and -196 ^{o}C was obtained for CaX zeolite [19], with high volumetric hydrogen storage densities of up to 31.0 g L^{-1} and 30.2 g L^{-1} for CaX and KX respectively [29].

Li and Yang studied the hydrogen storage capacities of low silica type X zeolites (LSX, silicon : aluminium = 1) which were fully exchanged by alkali-metal cations (lithium, sodium and potassium) at room temperature [30]. A notable example is Li-LSX, which was shown to reversibly adsorb 1.6 wt% of hydrogen at room temperature under 100 bar. Li et al. investigated the properties of the zeolites Na-LEV, H-OFF, Na-MAZ and Li-ABW. At 16 bar pressure and a temperature of

-196 ^{o}C, maximum hydrogen uptakes of 2.07, 1.75, 1.64 and 1.02 wt% were obtained for each material [31]. It was thought that an optimal channel diameter and large cage volume contributed to the storage capacity. No et al. introduced organic ions, such as pyridine hydrochloride and pyridinium chlorochromate, into the pores of the ion-exchanged zeolite-Y [32]. These modifications increased the hydrogen uptake capacity of the previously ineffi- cient Na-Y zeolite to 0.15–0.34 wt% at 25 oC and 100 bar.

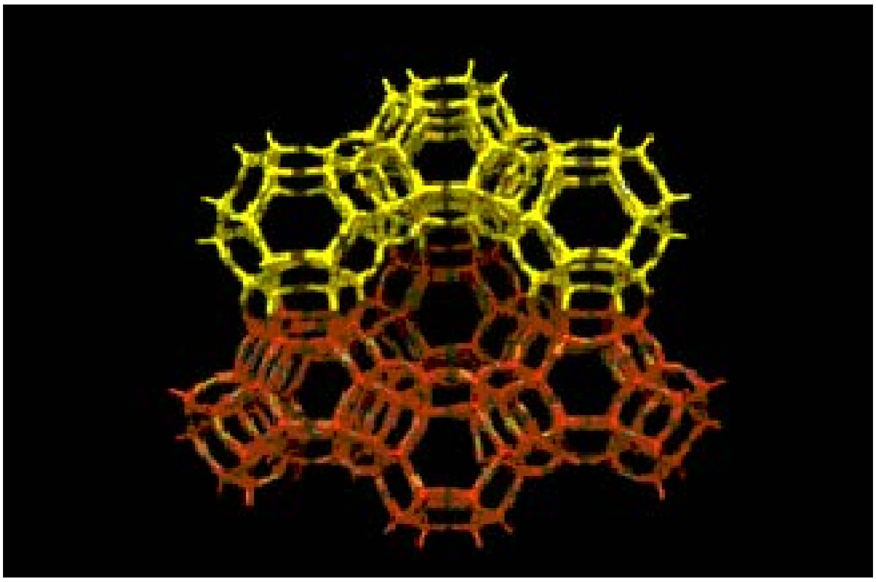


Fig.8.9 Zeolites of type A, X, Y, USY and ZSM-5 are produced on a commercial scale in varying Si to Al ratios. The Si/Al ratio controls total acidity as well as acid site strength. These zeolites are available as powders or formed into beads or monoliths. Zeolite supports are designed for reactions which can take advantage of the shape-selectivity or size exclusion of the zeolite pore. Typical pore sizes are between 4 and 8 Angstroms.

8.5 MOFs

The very high specific surface areas of MOFs (metal–organic frameworks) and IRMOFs (isoreticular metal–organic frame-works) mean these materials are excellent candidates for gas separation and storage (fig.8.10) [33,34]. The first report of hydrogen storage in MOFs was presented in 2003 by Yaghi et al. utilizing MOF-5 [35] and IRMOF-6 and -8 [36]. Capacities for these adsorbents corresponded to 1 wt% at room temperature under 20 bar and 4.5 wt% at -196 ^{o}C under 0.8 bar. It was later discovered that the reported uptake values were inflated by co-adsorption of gas impurities [37].

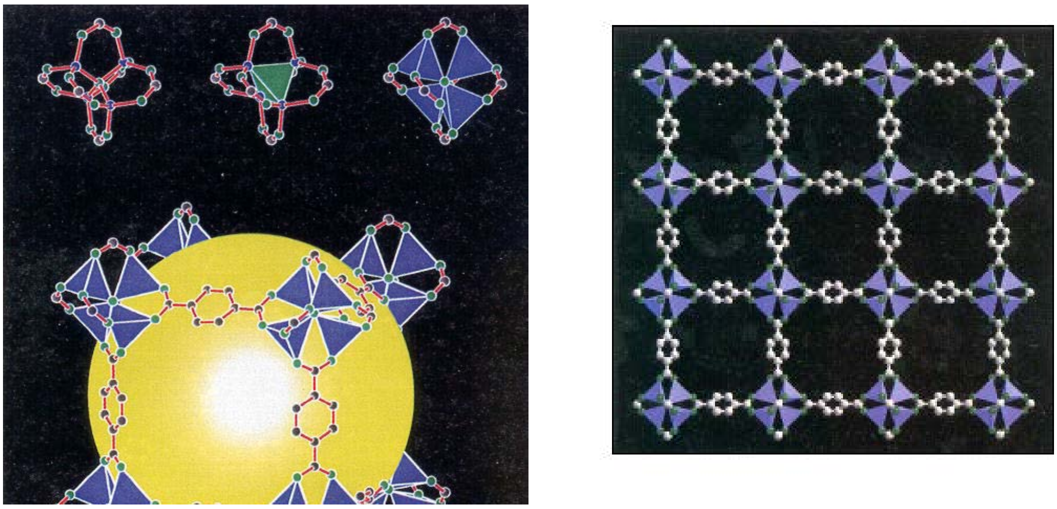


Fig.8.10 Porous Metal-Organic Framework (MOF) Chemical formula $Zn_4O(BDC)_3(DMF)_8(C_6H_5Cl)$, BCD=1,4–benzenedicarboxylate, DMF= dimethylformamide, Cubic crystal structure, 1.294 nm spacing between centers of adjacent clusters

These pioneering efforts have set the stage for considerable growth in this area, giving rise to numerous novel MOF physisorbants for hydrogen storage [33,38]. In particular, MOF-5 has undergone intensive investigation for hydrogen adsorption due to its very high specific surface area and simple structure. In 2006, maximum storage capacities of 4.5–5.2 wt% at approximately 50 bar and -196 ^{o}C were reported for MOF-5 by three different groups [39,40]. An improved sample preparation described by Huang under air- and moisture-free conditions gave a material with a BET specific surface area of 3800 m^2 g^{-1}[41]. This compound demonstrated enhanced capacity, measuring up to 7 wt%^{e} at -196 ^{o}C [42]. The use of hydrogen peroxide in the synthesis of MOF-5 favors the growth of high quality samples with large pore volume and high specific surface area, thereby enhancing hydrogen storage capabilities [43]. A novel mesoporous nanocube of MOF-5 (MNMOF-5) was developed by Bai et al. under dilute solvothermal conditions and displayed improve- ments for the heat of adsorption for hydrogen [44].

Other MOF topologies have been prepared and evaluated for hydrogen storage. Ferey et al. studied the hydrogen uptake in chromium- and aluminium-based MIL-53 MOFs, with a storage capacity of up to 3.8 wt% at -196 °C [45]. Significant uptake values of 6.7 and 7 wt% at -196 °C were later reported for two MOFs possessing a niobium(II) oxide-type framework consist- ing of copper(II)-dimeric paddle-wheel units coordinated by terphenyl and quaterphenyl linkers [46]. An increased storage capacity of 7.5 wt%^{e} at -196 °C was independently confirmed by gravimetric and volumetric measurements for MOF-177 [47], which is comprised of $Zn_4O(CO_2)_6$ units coordinated by the tritopic linker benzene tribenzoate [48]. Subsequently, Long et al. prepared a new type of MOF incorporating four- and six- coordinate manganese subunits with an overall cubic struc- ture [49]. This material demonstrated a total hydrogen uptake of 6.9 wt% at -196 °C and 90 bar. Between 35 and 40 bar, the excess capacity of desolvated polyhedral framework material NOTT-112 is 7.07 wt% [50]. The physisorption properties were attributed to a 3800 m^2 g^{-1} BET surface area and a 1.6 cm^3 g^{-1} pore volume. Kaskel et al. prepared framework DUT-49 with carbazole linker units, displaying a capacity of 8 wt% at 50 bar at room temperature [51]. Aided by computational studies, Hupp et al. designed and synthesized NU-100, with a hydrogen storage capacity of 16.4 wt% at -196 °C and 70 bar [52].

MOFs with promising hydrogen storage abilities described thus far have very high specific surface areas, a feature that correlates with the excellent adsorption properties of these materials [46,47]. A comprehensive overview on significant advances in the development of MOFs, including their use as hydrogen storage materials, can be found in a recent review by Li and Xu [53].

Due to weak interactions between hydrogen molecules and adsorbents, high storage capacities are typically achieved only at low temperatures (-196 °C) when using porous physisorption materials. The advantages of this storage mechanism compared to chemisorption are fast kinetics, complete reversibility and the relatively small amount of heat produced during on-board refueling. In general, the hydrogen storage capacity of physisorption materials correlates with specific surface area. For carbon supports this relation has even been formalized as ‘‘Chahine’s rule’’ (1 wt% increase for each 500 g m^{-2} specific surface area). Based on this phenomenon, increasing the pore volume and surface area is a key target in many efforts to improve hydrogen storage cap- abilities of physisorption media.

Activated carbons were once considered the best disordered materials and zeolites the best crystalline materials for gas storage by physisorption. More recently, new classes of crystalline metal organic frameworks and novel template porous carbons have been developed that possess specific surface areas exceeding those of activated carbons. Current excess hydrogen storage capacities of up to 7 wt% at -196 ^{o}C are an optimum for the best of these materials. New techniques, such as physical mixing, chemical doping and ultrasonication, have the potential to improve hydrogen storage in a variety of materials including carbons, MOFs, zeolites, mesoporous silica and other nano- structures [54,55].

8.6 Solid chemical hydrogen storage media

Contrary to physical storage methods described in the preceding section, chemical storage media contain hydrogen in chemically bonded or complexed forms, or incorporated into small organic molecules. Two categories will be separately addressed where media are comprised of either solid or liquid phase components. The release of molecular hydrogen is initiated when the source material is subjected to thermal or catalytic decomposition. Particular challenges in this area involve ease of reversible charging and discharging, storage capacity and stability of the base material and/or catalytic systems (Dalebrook *et al.* 2013).

8.6.1 Metal hydride and metal hydride alloys

Hydrogen reacts at elevated temperatures with many transition metals and their alloys to form hydrides. Many of these compounds, (MH_n), show large deviations from ideal stochiometry (n=1, 2, 3) and can exist as multiphase systems. The lattice structure is that of a typical metal with hydrogen atoms on the interstitial sites and for this reason they are also called interstitial hydrides. The maximum amount of hydrogen in the hydride phase is given by the number of interstitial sites in the intermetallic compound (Switendick 1978), (Westlake 1983). Metal hydrides can adsorb a large amount of hydrogen at a constant pressure. Some metal hydrides absorb and desorb hydrogen at ambient temperature and close to the atmospheric pressure.

The formation of the metal hydride - a direct interaction of the metal with hydrogen gas or electrochemically (fig. 8.11) by the reaction:

$M(solid) + H_2 (gas) \leftrightarrow MH_x(solid) + Q$

$M(solid) + H_2O + e^- \leftrightarrow MH_x (solid) + OH^- (liquid)$

One of the most interesting features of metallic hydrides is the extremely high volumetric density of hydrogen atoms present in the host lattice. The highest volumetric hydrogen density reported is 150 kg/m^3 in Mg_2FeH_6 and $Al(BH_4)_3$. Both hydrides belong to the complex hydrides family. Metallic hydrides can reach a volumetric hydrogen density of 115 kg/m^3, e.g. $LaNi_5$.

Metal hydrides are very effective as storing large amounts of hydrogen in a safe and compact way. All the reversible hydrides working around ambient temperature and atmospheric pressure consist of transition metals; therefore, the gravimetric hydrogen density is limited to less than 3 mass%. Group 1, 2, and 3 light metals, e.g. Li, Mg, B, and Al, give rise to a large variety of metal-hydrogen complexes. They are especially interesting because of their light weight and the number of hydrogen atoms per metal atom, which is two in many cases. The main difference between the complex and metallic hydrides is the transition to an ionic or covalent compound upon hydrogen absorption.

Metallic hydrides as hydrogen storage materials began intensively studied since 1997, when it was measured the absorption and desorption pressure-concentration isotherms for catalyzed $NaAlH_4$ at temperatures of 180°C and 210°C. The amount of the absorbed/desorbed hydrogen reached 4.2 mass% (Bogdanovich and Schwickardi 1997). The compound with the highest gravimetric hydrogen density at room temperature known today is $LiBH_4$ (18 mass%). This hydride desorbs three of the four hydrogen in the compound upon melting at 280°C and decomposes into LiH and B. The desorption process can be catalyzed by adding SiO_2 and significant thermal desorption has been observed starting at 100°C (Züttel *et al.* 2003).

Complex hydrides constitute relatively new field of hydrogen storage materials. The alanates have been investigated extensively during the last ten years. The borides are especially interesting because of their very high gravimetric and volumetric hydrogen density.

From other side, recent realizations of micro fuel cells (Mex *et al.* 2001), (Meyers and Maynard 2002), (Lee *et al.* 2002), (Yamazaki 2004), (Dzhafarov *et al.* 2008), (Gautier 2014) conceived to be used for energy supplying in portable devices are based on low-cost Si technologies already developed and successfully used mainly in microelectronics field throughout the world. Therefore, finding simple ways for hydrogen storage in such type of microsystems directly compatible with silicon-based microtechnologies is one of the key issues.

Metal hydrides are considered promising candidates for many stationary and mobile hydrogen storage applications. High volumetric densities and low adsorption pressures during hydrogen uptake are two favorable attributes [56]. A new class of nanocrystalline metal hydrides emerged in 1997 [57] and were reviewed for hydrogen storage applications in recent publications [58].

Alloys which are typically used for hydrogen storage are the so-called AB_5 compounds, such as $LaNi_5$, with a potential hydrogen content of 1.4 wt% and an equilibrium pressure of 2 bar at room temperature [59]. Current topics of research involve the design of new alloys and the mechanical effects of ball- milling substrates, including a study presented by Ares et al. on the effects of milling a multi-substituted substrate [60]. However, Fujii et al. later concluded that no improvement in hydrogen sorption properties can be expected from nanocrystalline $LaNi_5$ due to the formation of highly stable hydride species in the grain boundaries [61].

A titanium–iron alloy was recognized as a metal hydride storage material by Reilly and Wiswall in 1974, for which activation is achieved by heating at temperatures between 350 and 354 oC in a vacuum or hydrogen atmosphere [61]. A secondary hydrogenation step is performed after cooling to room temperature [62]. AB_2 Laves phases were similarly considered as prospective hydrogen storage materials, particularly the zirconium-based alloys [64]. These provide relatively good hydrogen storage capacity, kinetics and long cycling lifetimes at low cost, but are sensitive to gas impurities and are prohibitively stable at room temperature [65].

Several groups have reported body-centered cubic (BCC) alloys of various compositions that demonstrate maximum hydrogen capacities up to 4 wt%, with reversible capacities of more than 2 wt% [66]. The lattice parameter and the electron to atom ratio play important roles in the gas uptake. The effect of iron substitution in a titanium–chromium–manganese–vanadium alloy was investigated by Yu et al [67]. Research by Huang et al. undertook the replacement of vanadium by ferrovanadium (FeV) attaining 2.0 wt% of reversible hydrogen storage capacity with a composition of $TiCr_{1.2}(FeV)_{0.6}$ [68].

To date, a number of simple metal hydrides have also been considered for hydrogen storage applications. For example, nanocrystalline magnesium hydride is inexpensive and has a storage capacity of 7.6 wt% [69], although the equilibrium pres- sure of 10 bar at temperatures up to 300 1C (ref. 59) is not attractive for on-board hydrogen storage applications. Nevertheless, magnesium hydride has been extensively studied with the goal of improving the kinetics of hydrogen sorption.

Ball-milling of the substrate creates fresh surface sites during processing and is a simple and cost effective strategy for development. It was shown that ball milling under hydrogen promoted the formation of new hydride species, causing simultaneous hydrogen uptake and mechanical deformation. For example, after 15 h mechanical ball-milling at 300 oC under 10 bar hydrogen, the storage capacity of magnesium hydride increased from 4.86 to 6.01 wt% [70]. A later study by Cuevas et al. investigated the milling of magnesium and titanium in variable ratios under hydrogen, preparing a material with a 3.7 wt% hydrogen content and improved release kinetics promoted by the titanium component [71]. Nanostructuring techniques positively influence reaction kinetics and thus reduce sorption temperatures by lowering diffusion and migration barriers within the substrate [72].

Doping the metal substrate is another research direction for improving reaction kinetics. Catalysts and additives such as palladium [73] and other transition metals, their oxides (niobium(V) oxide)[74] and different metal hydrides were combined with magnesium hydride systems in order to increase reaction rates.

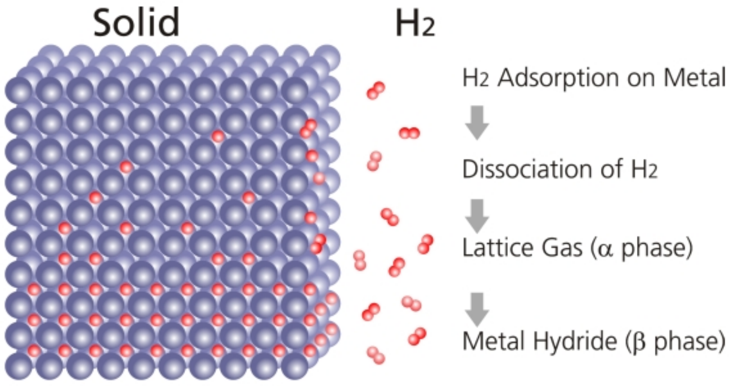


Fig.8.11 Steps of metal hydride formation

Additionally, the thermodynamic parameters of the reaction can be modified to obtain an operational temperature within an acceptable range. In 2009 Chen et al. reported that the use of potassium hydride as an additive in the magnesium amide–lithium hydride system weakens the N–H and Li–N bonds, thus enhancing hydrogen desorption at mild temperatures [75]. Desorption was observed at 80 oC and a reversible hydrogen uptake of 5 wt% was achieved at 107 oC. A technique utilizing chemical vapor deposition was shown to produce nanofibers with a significant increase in specific surface area [76]. Moreover, with the emergence of hydrogen carriers incorporated into thin films, hydrogen storage characteristics of pure and multi-layered magnesium films have been investigated intensively. The size effect of magnesium nanocrystals on hydrogen sorption kinetics was studied by Prieto et al. (Fig. 8.12) [77]. Magnesium nanocrystals of controllable sizes were prepared by reduction of magnesocene with organopotassium. A large rate enhancement of the hydrogen sorption kinetics was achieved using nanocrystals with smaller diameters, probably due to an increase in the defect density present in smaller nanocrystals. This topic is elaborated upon in several reviews [59].

Metal hydrides have been established as potential hydrogen storage materials, but in order to justify their implementation in industrial applications further research is required to develop systems with improved performance. In particular, the develoment of new catalysts and alloys and their respective methods of synthesis (especially for structured nanomaterials) should become a primary focus of the scientific community.

Complex transition metal hydrides are generally not useful as hydrogen release materials due to their high cost, high decomposition temperatures and low storage capacities. Much of the work concerning the preparation and structure determination of these compounds is summarized by Bronger and Auffer- mann [78] and Noreus et al [79]. The widely studied Mg_2FeH_6 has a gravimetric capacity of 5.5 wt% and decomposes at 320 oC under 1 bar of hydrogen with a volumetric capacity of 150 g L^{-1}, more than twice that of liquid hydrogen (71 g L^{-}

^{1}). These properties indicate that thermochemical storage applications for this compound could be promising [80].

At present hydride systems are not suitable for large scale commercial applications. Due to thermodynamic constraints boron hydrides and complex aluminium hydrides are not reversibly hydrogenated under acceptable conditions. The known complex transition metal hydrides suffer from high cost and low gravimetric loading. However, exceptions for specific cases are possible, such as the use of sodium aluminium hydride as a hydrogen carrier for PEM fuel cells. Fundamental research is necessary to understand all the mechanisms involved in the dehydrogenation and rehydrogenation of complex hydrides. The main disadvantages storage metal – hydrides: brittleness starting material, decomposition of some hydrides - an irreversible process, often requires high temperatures, i.e. energy consuming.

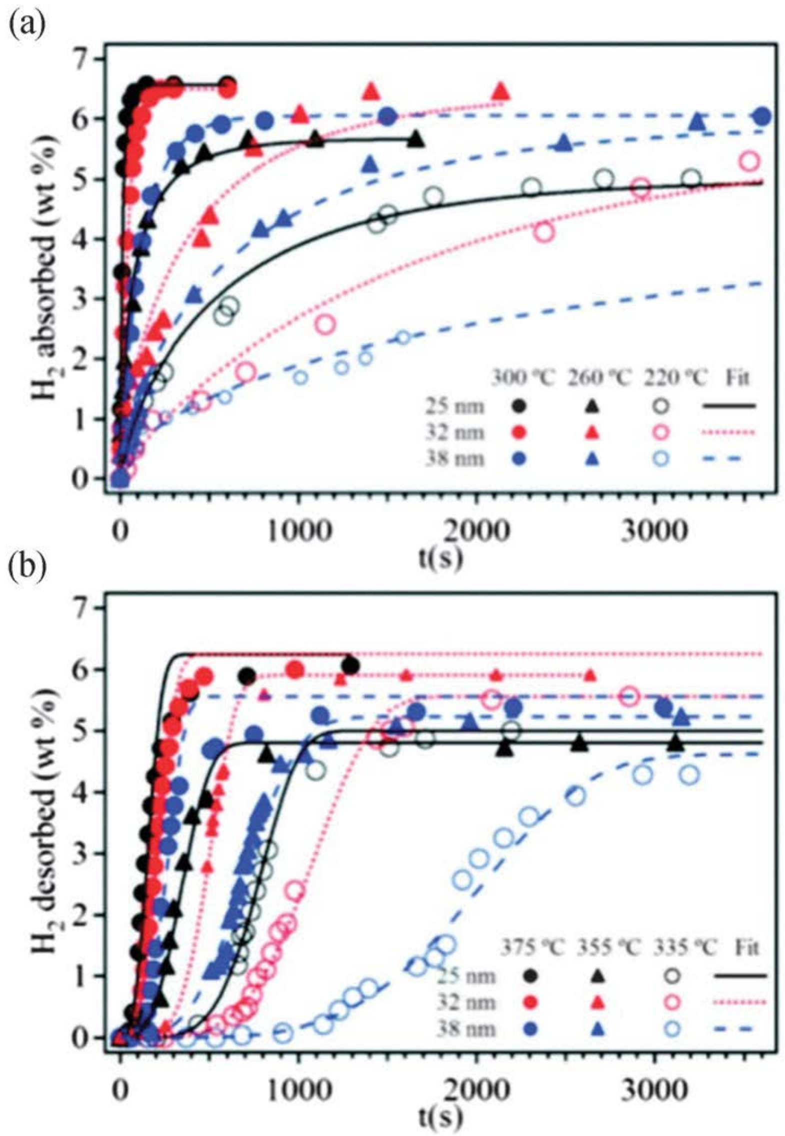


Fig.8.12 Hydrogen absorption (a) and desorption (b) of magnesium nanocrystal samples at different temperatures

Figure 8.13 and 8.14 show the hydrogen storage capacity (considering mass and volume) for metal hydrides, carbon nanotubes, petrol and other hydrocarbons. Gravimetric H2 content is bigger for metal hydrides and other chemical adsorbed systems, however, its need in the high operation temperature. From other side the reversed physical adsorbed systems (carbons, MOFs etc) reveals low hydrogen mass but can operated at low temperature.

So , today the main task is search the optimal materials to meet requirement of low operation temperature and high gravimetric H_2 density.

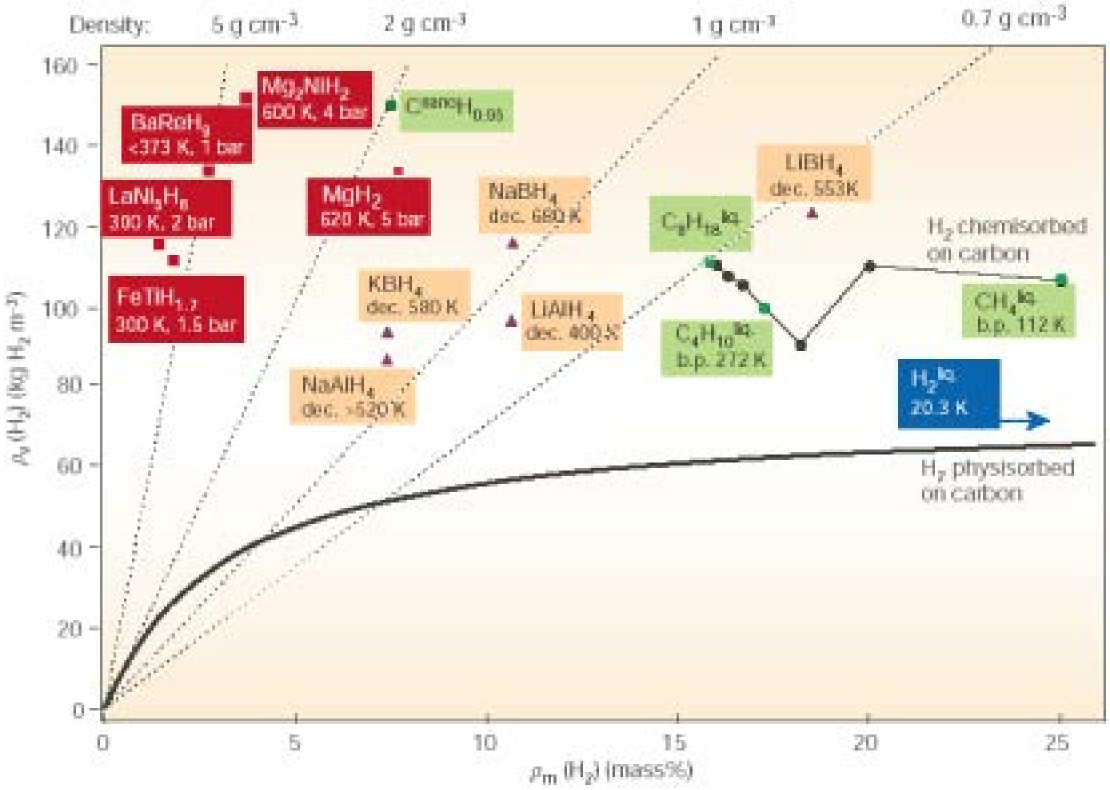


Figure 8.13 – Hydrogen storage capacity (considering mass and volume) for metal hydrides, carbon nanotubes, petrol and other hydrocarbons

Fig. 8.15 illustrates the size of hydrogen storage materials relative to car size for Volume of 4 kg of hydrogen.

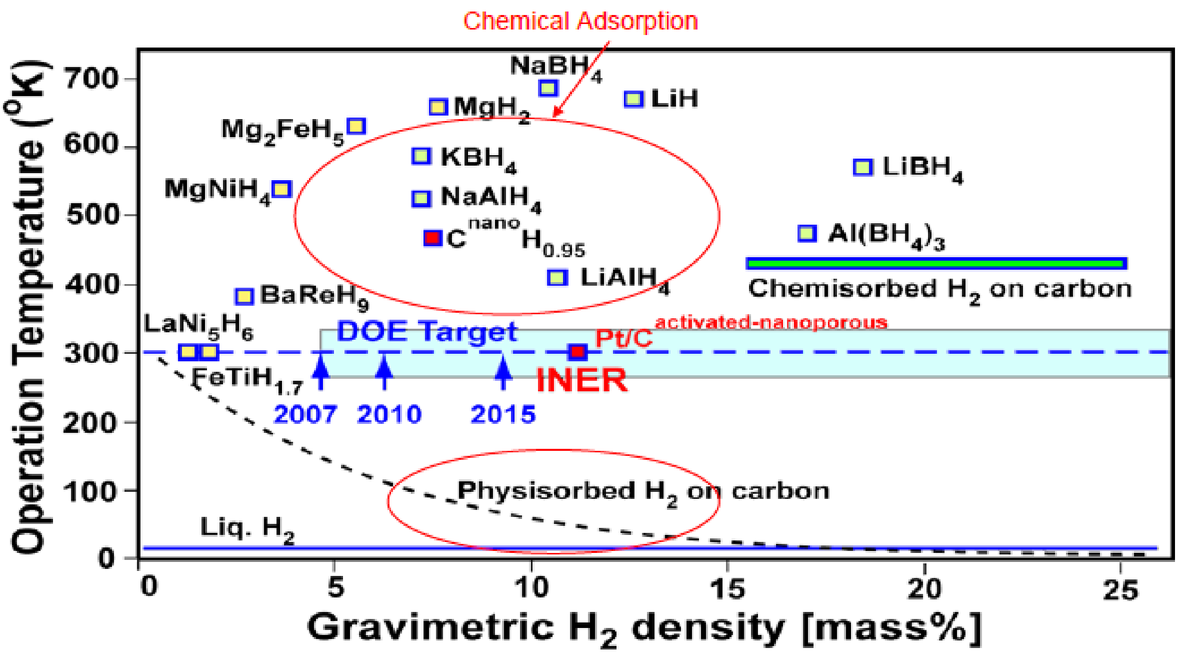


Fig.8.13 Summary of Operation temperature and gravimetric H2 density for different materials. Blue line indicates the DOE target- room temperature and gravimetric H2 density for different years.

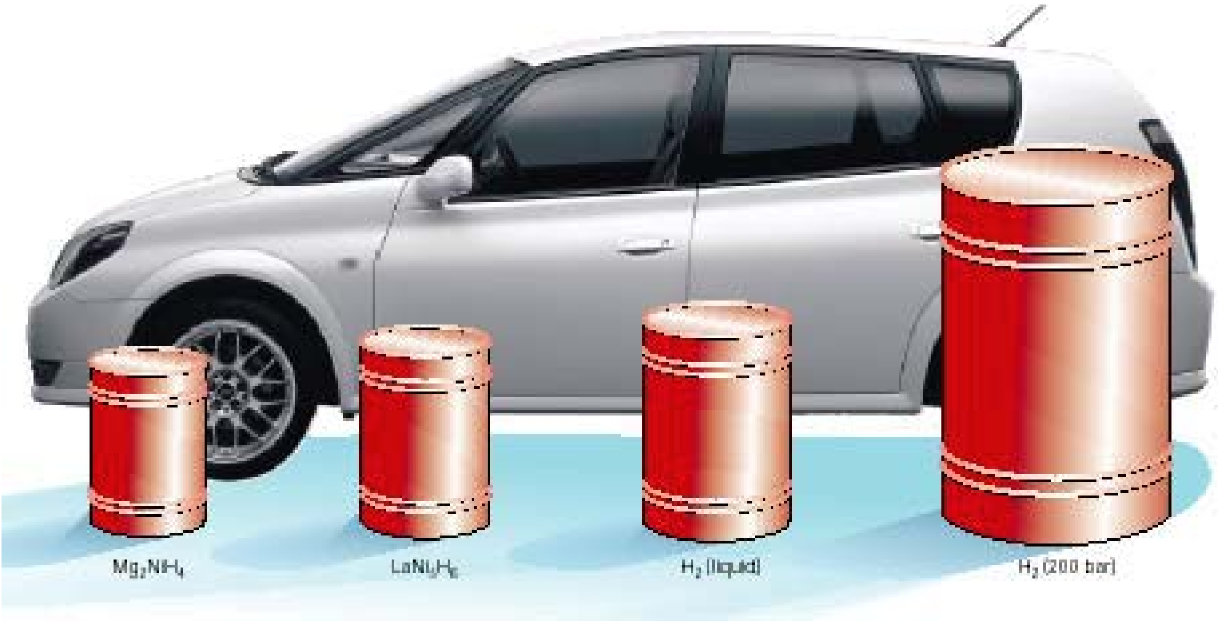


Fig 8.14 Volume of 4 kg of hydrogen (needed to run an electric car with a fuel cell) compacted in different ways as schematically illustrated here; size relative to car size

References

1 K. A. G. Amankwah, J. S. Noh and J. A. Schwarz, Int. J. Hydrogen Energy, 1989, 14, 437–447.
2 A. C. Dillon, K. M. Jones, T. A. Bekkedahl, C. H. Kiang, D. S. Bethune and M. J. Heben, Nature, 1997, 386, 377–379.
3 A. Chambers, C. Park, R. T. K. Baker and N. M. Rodriguez, J. Phys. Chem. B, 1998, 102, 4253–4256.
4 Y. Ye, C. C. Ahn, C. Witham, B. Fultz, J. Liu, A. G. Rinzler, D. Colbert, K. A. Smith and R. E. Smalley, Appl. Phys. Lett., 1999,74, 2307–2309.
5 P. Chen, X. Wu, J. Lin and K. L. Tan, Science, 1999, 285, 91–93.
6 C. Liu, Y. Y. Fan, M. Liu, H. T. Cong, H. M. Cheng and M. S. Dresselhaus, Science, 1999, 286, 1127–1129.
7 M. Rzepka, P. Lamp and M. A. de la Casa-Lillo, J. Phys. Chem. B, 1998, 102, 10894–10898.
8 M. Jorda-Beneyto, F. Suarez-Garcia, D. Lozano-Castello, D. Cazorla-Amoros and A. Linares-Solano, Carbon, 2007, 45, 293–303.
9 R. Chahine and T. K. Bose, Int. J. Hydrogen Energy, 1994, 19,161–164.
10 B. Panella, M. Hirscher and S. Roth, Carbon, 2005, 43, 2209–2214.
11 Z. X. Yang, Y. D. Xia and R. Mokaya, J. Am. Chem. Soc., 2007, 129,1673–1679.
12 M. G. Nijkamp, J. E. M. J. Raaymakers, A. J. van Dillen and K. P. de Jong, Appl. Phys. A: Mater. Sci. Process., 2001, 72, 619–623.
13 G. Yushin, R. Dash, J. Jagiello, J. E. Fischer and Y. Gogotsi, Adv.Funct. Mater., 2006, 16, 2288–2293.
14 H. Jin, Y. S. Lee and I. Hong, Catal. Today, 2007, 120, 399–406.

15 H. L. Jiang, B. Liu, Y. Q. Lan, K. Kuratani, T. Akita, H. Shioyama,F. Q. Zong and Q. Xu, J. Am. Chem. Soc., 2011, 133, 11854–11857.
16 V. Tozzini and V. Pellegrini, Phys. Chem. Chem. Phys., 2013, 15, 80–89
17 S. H. Aboutalebi, S. Aminorroaya-Yamini, I. Nevirkovets, K. Konstantinov and H. K. Liu, Adv. Energy Mater., 2012, 2,1439–1446.
18 J. Weitkamp, M. Fritz and S. Ernst, Int. J. Hydrogen Energy, 1995, 20, 967–970.
19 H. W. Langmi, A. Walton, M. M. Al-Mamouri, S. R. Johnson, D. Book, J. D. Speight, P. P. Edwards, I. Gameson, P. A. Anderson and I. R. Harris, J. Alloys Compd., 2003, 356, 710–715.
20 H. W. Langmi, D. Book, A. Walton, S. R. Johnson, M. M. Al- Mamouri, J. D. Speight, P. P. Edwards, I. R. Harris andP. A. Anderson, J. Alloys Compd., 2005, 404, 637–642.
30 Y. W. Li and R. T. Yang, J. Phys. Chem. B, 2006, 110, 17175–17181.
31 J. X. Dong, X. Y. Wang, H. Xu, Q. Zhao and J. P. Li, Int. J. Hydrogen Energy, 2007, 32, 4998–5004.
32 D. Bae, H. Park, J. S. Kim, J. B. Lee, O. Y. Kwon, K. Y. Kim, M. K. Song and K. T. No, J. Phys. Chem. Solids, 2008, 69, 1152–1154.
33 U. Mueller, M. Schubert, F. Teich, H. Puetter, K. Schierle-Arndt and J. Pastre, J. Mater. Chem., 2006, 16, 626–636.
34 G. Ferey, Chem. Soc. Rev., 2008, 37, 191–214.
35 H. Li, M. Eddaoudi, M. O’Keeffe and O. M. Yaghi, Nature, 1999,402, 276–279.
36 N. L. Rosi, J. Eckert, M. Eddaoudi, D. T. Vodak, J. Kim, M. O’Keeffe and O. M. Yaghi, Science, 2003, 300, 1127–1129.
37 J. L. C. Rowsell, A. R. Millward, K. S. Park and O. M. Yaghi, J. Am. Chem. Soc., 2004, 126, 5666–5667.
38 J. L. C. Rowsell and O. M. Yaghi, Angew. Chem., Int. Ed., 2005, 44,4670–4679.
39 A. Dailly, J. J. Vajo and C. C. Ahn, J. Phys. Chem. B, 2006, 110,1099–1101.
40 B. Panella, M. Hirscher, H. Putter and U. Muller, Adv. Funct. Mater.,2006, 16, 520–524.
41 L. M. Huang, H. T. Wang, J. X. Chen, Z. B. Wang, J. Y. Sun, D. Y. Zhao and Y. S. Yan,

Microporous Mesoporous Mater., 2003, 58, 105–114.
42 S. S. Kaye, A. Dailly, O. M. Yaghi and J. R. Long, J. Am. Chem. Soc., 2007, 129, 14176–14177.
43 S. J. Cheng, S. B. Liu, Q. Zhao and J. P. Li, Energy Convers. Manage., 2009, 50, 1314–1317.
44 Z. F. Xin, J. F. Bai, Y. Pan and M. J. Zaworotko, Chem.–Eur. J., 2010, 16, 13049–13052.
45 G. Ferey, M. Latroche, C. Serre, F. Millange, T. Loiseau and A. Percheron-Guegan, Chem. Commun., 2003, 2976–2977.
46 X. Lin, J. H. Jia, X. B. Zhao, K. M. Thomas, A. J. Blake, G. S. Walker, N. R. Champness, P. Hubberstey and M. Schroder, Angew. Chem.,Int. Ed., 2006, 45, 7358–7364.
47 H. Furukawa, M. A. Miller and O. M. Yaghi, J. Mater. Chem., 2007,17, 3197–3204.
48 H. K. Chae, D. Y. Siberio-Perez, J. Kim, Y. Go, M. Eddaoudi,A. J. Matzger, M. O'Keeffe and O. M. Yaghi, Nature, 2004, 427,523–527.
49 M. Dinca, A. Dailly, Y. Liu, C. M. Brown, D. A. Neumann andJ. R. Long, J. Am. Chem. Soc., 2006, 128, 16876–16883.
50 Y. Yan, X. Lin, S. H. Yang, A. J. Blake, A. Dailly, N. R. Champness,P. Hubberstey and M. Schroder, Chem. Commun., 2009, 1025–1027.
51 U. Stoeck, S. Krause, V. Bon, I. Senkovska and S. Kaskel, Chem.Commun., 2012, 48, 10841–10843.
52 O. K. Farha, A. OÅN. Yazaydın, I. Eryazici, C. D. Malliakas,B. G. Hauser, M. G. Kanatzidis, S. T. Nguyen, R. Q. Snurr andJ. T. Hupp, Nat. Chem., 2010, 2, 944–948
53.S.-L. Li and Q. Xu, Energy Environ. Sci., 2013, 6, 1656–1683.

54 Y. Y. Liu, Z. Ju-Lan, Z. Jian, F. Xu and L. X. Sun, Int. J. Hydrogen Energy, 2007, 32, 4005–4010.
55 L. F. Wang and R. T. Yang, Energy Environ. Sci., 2008, 1, 268–279.

56 J. J. Reilly, Z. Phys. Chem., 1979, 117, 155–184.
57.L. Zaluski, A. Zaluska and J. O. StromOlsen, J. Alloys Compd., 1997,253, 70–79.
58 W. Grochala and P. P. Edwards, Chem. Rev., 2004, 104, 1283–1315.
86 A. W. C. van den Berg and C. O. Arean, Chem. Commun., 2008,668–681.
59 B. Sakintuna, F. Lamari-Darkrim and M. Hirscher, Int. J. Hydrogen Energy, 2007, 32, 1121–1140.
60 J. R. Ares, F. Cuevas and A. Percheron-Guegan, Mater. Sci. Eng., B, 2004, 108, 76–80.
61 H. Fujii, S. Munehiro, K. Fujii and S. Orimo, J. Alloys Compd., 2002,
330, 747–751.
62 J. J. Reilly and R. H. Wiswall, Inorg. Chem., 1974, 13, 218–222.
63 L. Schlapbach, Surface Properties and Activation in Hydrogen in
Intermetallic Compounds II, Springer-Verlag, Berlin, 1992.
64 D. Shaltiel, I. Jacob and D. Davidov, J. Less-Common Met., 1977, 53,117–131.
65 M. Bououdina, J. L. Soubeyroux, P. de Rango and D. Fruchart, Int. J.
Hydrogen Energy, 2000, 25, 1059–1068.

66 Y. G. Chen, Y. G. Yan, H. Liang, X. X. Zhou, C. L. Wu, M. D. Tao andL. J. Pang, J. Alloys Compd., 2008, 454, 427–431.
67 X. B. Yu, Z. X. Yang, S. L. Feng, Z. Wu and N. X. Xu, Int. J. Hydrogen Energy, 2006, 31, 1176–1181.
68 T. Z. Huang, Z. Wu, J. Z. Chen, X. B. Yu, B. J. Xia and N. X. Xu, Mater. Sci. Eng., A, 2004, 385, 17–21.
69 A. Zaluska, L. Zaluski and J. O. Strom-Olsen, J. Alloys Compd., 1999,288, 217–225.
70 S. R. Johnson, P. A. Anderson, P. P. Edwards, I. Gameson, J. W. Prendergast, M. Al-Mamouri, D. Book, I. R. Harris, J. D. Speight and A. Walton, Chem. Commun., 2005, 2823–2825.
71 F. Cuevas, D. Korablov and M. Latroche, Phys. Chem. Chem. Phys., 2012, 14, 1200–1211.
72 V. BeÅLrubeÅL, G. Radtke, M. Dresselhaus and G. Chen, Int. J. Energy Res., 2007, 31,

637–663.
73 H. Fritzsche, M. Saoudi, J. Haagsma, C. Ophus, E. Luber, C. T. Harrower and D. Mitlin, Appl. Phys. Lett., 2008, 92, 121917.
74 G. Barkhordarian, T. Klassen and R. Bormann, Scr. Mater., 2003, 49, 213–217.
75 J. H. Wang, T. Liu, G. T. Wu, W. Li, Y. F. Liu, C. M. Araujo, R. H. Scheicher, A. Blomqvist, R. Ahuja, Z. T. Xiong, P. Yang, M. X. Gao, H. G. Pan and P. Chen, Angew. Chem., Int. Ed., 2009, 48, 5828–5832.
76 I. Saita, T. Toshima, S. Tanda and T. Akiyama, Mater. Trans., JIM, 2006, 47, 931–934.
77 N. S. Norberg, T. S. Arthur, S. J. Fredrick and A. L. Prieto, J. Am. Chem. Soc., 2011, 133, 10679–10681.
78 W. Bronger and G. Auffermann, Chem. Mater., 1998, 10, 2723–2732. 79 M. Olofsson-Martensson, U. Haussermann, J. Tomkinson and D. Noreus, J. Am. Chem. Soc., 2000, 122, 6960–6970.
80 B. Bogdanovic, A. Reiser, K. Schlichte, B. Spliethoff and B. Tesche, J. Alloys Compd., 2002, 345, 77–89.

Chapter 9. Hydrogen storage and generation in nanosilicon

9.1 Introduction

Porous silicon (PS) is one of the modern materials used in the hydrogen power industry. It can serve as a hydrogen source due to the presence of silane groups SiH_X, $x = 1 - 3$, on its surface arising in the course of manufacturing. The advantages of this material are a high content of hydrogen (up to 6 mass percent), the easiness of its release, and the considerable intensity of H_2 desorption. The main drawback of PS is that it is difficult to restore the hydrogen content after one cycle of H_2 desorption, which is caused by the oxidation of silicon. In order to release hydrogen from PS, one uses the following methods: heating, ultraviolet illumination, and reactions with water. The drawback of two former methods lies in the possibility of the oxidation of desorbed H_2 with the formation of water, which requires to carry out the desorption in high vacuum or the inert atmosphere. That is why it is more efficient to obtain hydrogen due to the reaction with water in the presence of a catalyst. Moreover, the hydrogen release is observed both for porous and crystalline phases of silicon .

Modern technologies allow one to create silicon and PS fractions with specified particle sizes, surface mor- phology, and fractality level, which is reflected in the hydrogen content in a material and the intensity of H_2 desorption. In addition, there exists a possibility to increase the content of hydrogen in PS by means of the additional treatment of the material in gaseous H_2. The use of catalysts of hydrogenation reactions allows one to intensify the interaction of H_2 with a material. To such catalysts, one refers metals of the platinum group, palladium being the most efficient of them. This is re- lated to the highest solubility of hydrogen in palladium as compared with the other metals, as well as to the dis- sociation of H_2 molecules into atoms under interaction with a metal surface.

9.2 Estimation of hydrogen contents in PS using FTIR method

It is known that hydrogen is present in the chemical composition of PS, especially in as-prepared samples, as a result of etching technology (Parkhutik 1999), (Föll *et al.* 2002). Thus, silane SiH_x (x = 1, 2, 3) groups are formed on the surface of the obtained PS. In addition, some amount of hydrogen is incorporated into the bulk of PS. Here, it is kept on internal Si states of different types, such as crystalline planes, inner voids and other defects of structure, as well as on impurities (especially oxygen and boron) (Estreicher *et al.* 1999), (Pritchard *et al.* 1999), (Ogata *et al.* 1995). H atoms, incorporated inside the Si lattice, can interact chemically with inner dangling silicon bonds and atoms of impurities. Thus, inner SiH_x, Si-H-B, O_xSiH, SiOH and other complexes are formed. The peculiarities of hydrogen arrangement in the PS lattice are similar to those in c-Si (Cerofolini *et al.* 1995). Moreover, hydrogen in molecular form has the possibility to be captured by the surface or to penetrate into the PS lattice. The structure contains centers that are able to hold H_2 molecules by the forces of dipole–dipole interaction. Some of such centers are formed on the interstitial oxygen atoms. Further, hydrogen molecules can be entrapped and dissociated on defects, particularly on vacancies and self-interstitials (Estreicher *et al.* 1999), (Pritchard *et al.* 1999).

Hydrogen concentration in N_H (mmol/cm^3) in PS can be estimated from absorption spectra by using the following relation used earlier for estimation of hydrogen content in amorphous Si layers (Tolstoy *et al.* 2003):

$$N_H = \frac{1}{\Gamma_S}\int_{h\nu} \frac{a}{h\nu} d(h\nu) = \frac{I_S}{\Gamma_S}, \qquad (9.1)$$

where I_S (cm^{-1}) is the integrated adsorption of the stretching band, α(cm^{-1}) the absorption coefficient, $h\nu$ is the photon energy and Γ_S (cm^2/mmol) is the stretching oscillator strength of the Si–H bonds which was determined from the following relation as it was proposed in (Shanks *et al.* 1980) for amorphous silicon:

$$\Gamma_S = 37.6 \times \frac{I_S}{I_W} \ , \qquad (9.2)$$

where I_W is experimentally measured integrated adsorption of the wagging band (near 640 cm^{-1}).

Otherwise, the hydrogen concentration value can be expressed in atomic, C_H (at.%), or mass C_M (mass%) percents as follows (Tolstoy *et al.* 2003):

$$C_H = \frac{N_H m_{Si}}{1000\rho_{Si}(1-P) + N_H m_{Si}} \ , \qquad (9.3)$$

$$C_M = \frac{1}{1 + (m_{Si}/m_H)\big((1/C_H) - 1\big)} \ . \qquad (9.4)$$

Where m_{Si} and m_H are the atomic masses of silicon and hydrogen, respectively, ρ_{Si} is the monocrystalline Si density (2.33 g/cm^3) and P is the porosity of the PS nanostructure.

Hydrogen in PS is stored in Si-H, Si-H_2 and Si-H_3 bonds and their ratio can be estimated from experimental FTIR spectra. Typical well-known Si-H_x (with x= 1, 2, 3) stretching vibration spectra of meso- PS nanostructures are shown in Fig. 9.1. Three well-defined large bands centered at 2088, 2110, and 2137 cm^{-1}, mainly corresponding respectively to Si-H, Si-H_2 and Si-H_3 stretching bonds, can be clearly seen by (Bisi *et al.* 2000), (Chazalviel and Ozanam 1997). Additionally, a fine structure of the absorption bands consisting of about seven overlapping peaks can be distinguished (especially for a low-porosity meso-PS sample), which is probably due to the specific orientations of Si-H_x bonds in space induced by morphological anisotropy of Si nanocrystallites constituting the porous nanostructures. In particular, oscillation modes corresponding to Si monohydride dimmers can also be present in the range of 2075-2103 cm^{-1}(Timoshenko *et al.* 2003).

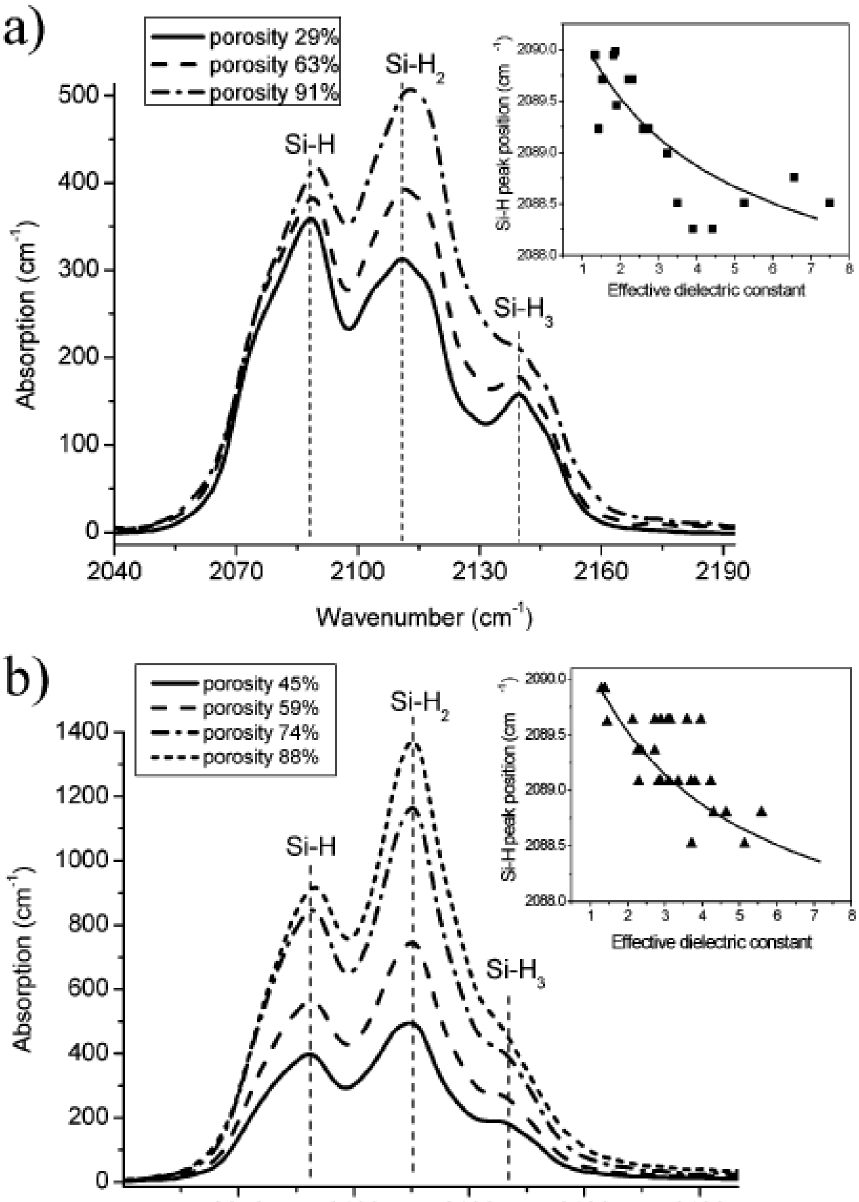


Fig.9.1. Infrared absorption spectra of as-prepared hydrogen terminated meso-PS and nano-PS for three and four porosity values. The insert shows Si–H peaks shift along with effective dielectric constant of the porous silicon. Lysenko V., et.al. *Appl. Surf. Sci.* **230**, 425 – 430

A slight spectral shift of the Si–H and Si–H$_2$ bands toward higher energy edge along with porosity is observed. It can be explained by influence of the porosity dependent effective dielectric constant of meso-PS on the vibrational frequency of SiH$_x$ bonds. Indeed, experimental points corresponding to the SiH peak positions (see insert in Fig. 9.1) can be well fitted (solid curve) by the following expression proposed in (Wieder *et al.* 1979)

$$\omega_{SiH} = \sqrt{\omega_0^2 - A\frac{\varepsilon_{PS}-1}{\varepsilon_{PS}+3}} \tag{9.5}$$

where ω_{SiH} is the energy of Si-H peak, ω_0 and A is fitting parameters equal 2090 cm^{-1} and $1.2x10^4$ cm^{-1}, respectively, ε_{PS} is the porosity dependent effective dielectric constant.

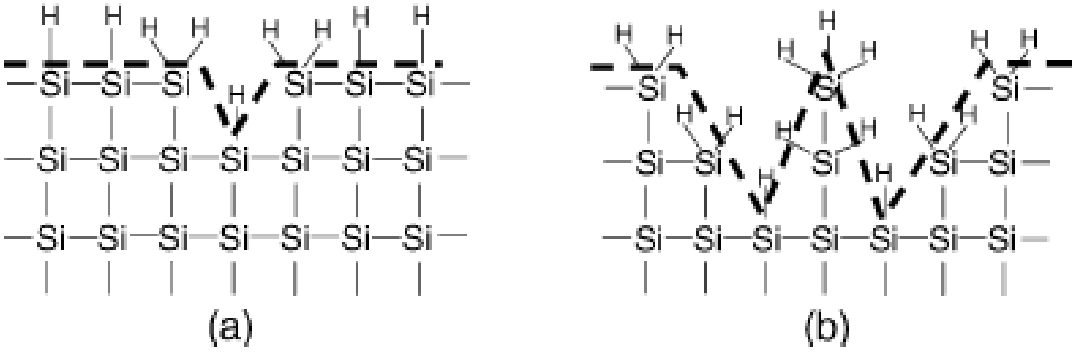


Fig.9.2 Schematic view of porosity dependent fractal-like roughness of Si nanocrystallites surfaces at atomic level: (a) low porosity (<60%); (b) high porosity (>60%). Dashed line points the nanocrystallites surfaces .Lysenko V., et al.. *Appl. Surf. Sci.* **230**, 425 – 430

Another important porosity induced effect can also be observed from the absorption spectra. Indeed, the peak corresponding to Si–H band is higher than that of Si–H_2 one for low porosity values (<63%). Thus, hydrogen is bound with Si atoms by forming mainly monohydride bonds. For high porosity values (>63%) hydrogen is mainly stored under dihydride complexes while for intermediate porosity (~63%) the peak absorption intensities of the mono- and dihydride bonds are the same. Such a remarkable progressive evolution of the Si–H_2 bond intensity is asign of the porosity dependent fractal shape of the Si nanocrystallites surface. It can be explained by increase of fractal-like roughness of the Si nanocrystallites surfaces at atomic level favoring additional appearance of surface Si atoms with two free bonds for preferable formation of Si–H_2 complexes (see schematic representation in Fig. 9.2).

Being relatively smooth at low porosities (<63%), fractal-like surfaces of the large nanocrystallites have a low amount of Si atoms capable to bind two hydrogen atoms (Fig. 9.2a). Porosity enhancement is accompanied by increase of the surface atomic roughness and therefore by number of SiH_2 surface bonds (Fig.9.2) (Lysenko *et al.* 2004).

Fig.9.3 shows the variation of the atomic concentration of hydrogen bound to Si atoms at the PS- specific surface as a function of the porosity of the layer (Lysenko *et al.* 2005). In general, the concentration variation with porosity reflects structural evolution of the PS-specific surface to which hydrogen is bound. Hydrogen concentration values in nano-PS are higher than in meso-PS layers because of the much more important specific surface to which hydrogen atoms may be bound. Indeed, as it was observed from nitrogen adsorption experiments (Herino *et al.* 1987) the specific surface of meso-PS is measured to be about 200 m^2/cm^3, while the nano-PS surface is found to be about 600 m^2/cm^3. Such an important difference of the specific surface values for the meso- and nano-PS morphologies are explained by the difference of the nanocrystallite dimension constituting these two PS morphological types, as can be seen in Fig.9.3b. Indeed, for a given porosity value, the smaller the nanocrystallite dimension is, the higher the corresponding specific surface is.

Fig. 9.4 shows the evolution of Si nanocrystallites dimension, d, estimated from Raman spectroscopy measurements and of specific surface area, S (m^2/g), calculated from the dimension estimations assuming a smooth spherical geometry of the nanocrystallites and using the following relation that can be asily demonstrated:

$$S\,(\mathrm{m}^2/\mathrm{g}) = \frac{6}{d\rho_{\mathrm{Si}}} \qquad (9.6)$$

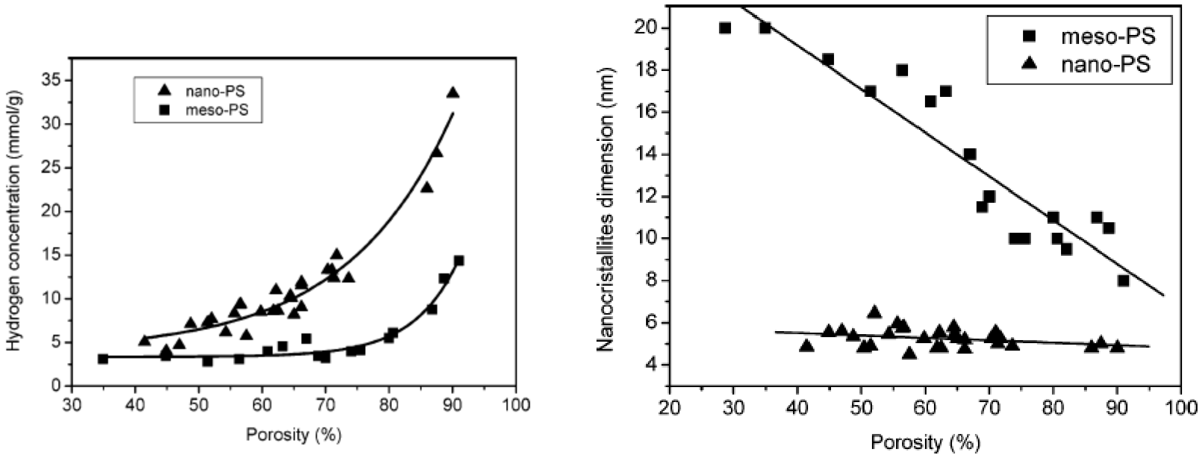


Fig.9.3. a) Hydrogen concentration variation along with porosity in meso- and nano-PS layers. Solid lines added as eye guides represent exponential fit of the experimental data. b) Dimensions of Si nanocrystallites in meso- and nano-PS layers vs porosity. Straight solid lines represent linear fits of the experimental data. Lysenko V. et.al (2005). *J.Phys. Chem. B.* **109**, 19711-19718.

According to this relation, specific surface area is inversely proportional to the nanocrystallites size and its porosity dependence is determined by porosity evolution of the nanocrystallites dimension.

Concerning the meso-PS samples in the 30-70% porosity range, the hydrogen concentration is very low and relatively constant (2.5-5 mmol/g) because the samples are characterized by large (Fig.9.3b) and smooth nanocrystallites. For example, the concentration of atomic hydrogen at the surface of the meso-PS sample of 60% porosity is 3.8 mmol/g that corresponds well to the value of the molecular hydrogen extracted from thermal desorption measurements (~2 mmol/g) (Rivolo *et al.* 2003). Despite the nanocrystallite decrease from 20 to 12 nm, the corresponding specific surface values remain still too small (125-200 m^2/g) in the meso-PS samples with porosities <70%. The corresponding quantity of hydrogen bound at this surface is also too small to be measured with good precision. Therefore, only a slight variation of the hydrogen concentration around a constant level is found in this porosity range. Despite the almost unchangeable nanocrystallite dimensions around 10 nm, a monotonic exponential increase of the concentration (up to 13 mmol/g) in the range of high porosities (>70%) reflects an important enhancement of the PS-specific surface induced by the extremely developed fractal-like surface roughness of the Si nanocrystallites constituting the porous layers (Lysenko *et al.* 2004), (Lysenko *et al.* 2005).

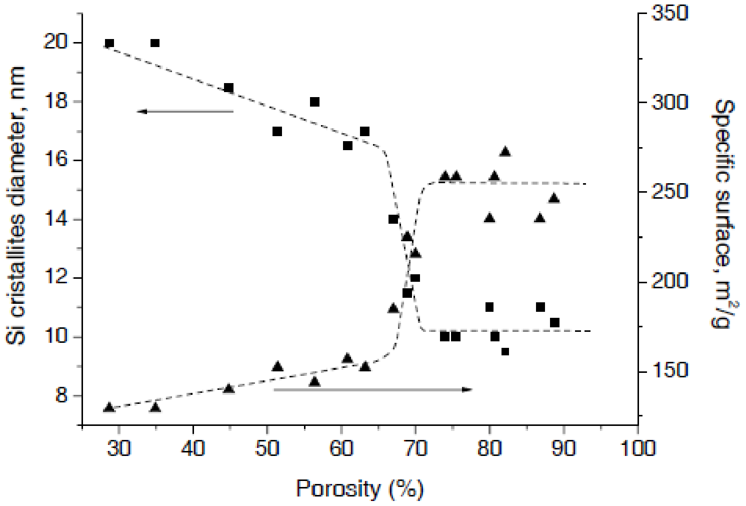


Fig. 9.4. Si nanocrystallites size (&) and specific surface (~) of the meso-porous Si as functions of porosity. Dashed lines are added for eyes guiding Lysenko V., et al.. *Appl. Surf. Sci.* **230**, 425 – 430

The fractal nature of the PS-specific surface roughness is completely confirmed by hydrogen concentration analysis in the nano-PS samples. Indeed, as it can be seen in Fig.9.3, the concentration of hydrogen increases exponentially along with porosity. An excellent exponential fit of the experimental data is represented in Fig.9.3 by a solid line. Quasi- constant values of the Si nanocrystallite dimensions along all porosity ranges means that the porosity increase is ensured only by the decrease of the number of nanocrystallites, and consequently, the observed increase of the total specific surface with porosity is only possible if an important enhancement of the fractal-like surface roughness at the nanometer scale is considered for each nanocrystallite.

Thus, the estimated content of hydrogen in PS is quite high rather for very small nanocrystalls of PS (up to 15-35 mmol/g) and thus it can be considered as quite efficient solid-state hydrogen storage material. The view of Si nanocrystals can be obtained from TEM images. Fig.9.5(a) shows a TEM image of a porosified grain of polycrystalline Si powder and Fig.9.5 (b) shows spherical freestanding Si nanocrystals. As it can be seen, PS is a sponge-like structure containing interconnected Si nanocrystals. The typical sizes of Si nanocrystals in both systems are ~3–10 nm. Immersion of nanosilicon structures in water/alcohol solutions is accompanied by evolution of hydrogen (Goller *et al*. 2011).

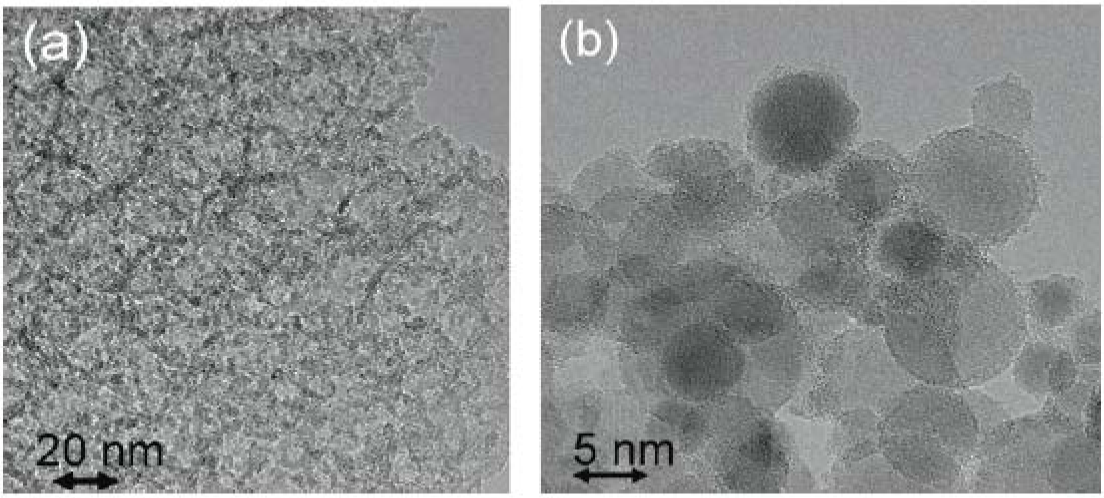


Fig.9.5.TEM images of nanosilicon samples. (a) Chemically porosified polycrystalline Si powder. (b) Freestanding Si nanospheres prepared from the gas phase and subsequently stain etched. The size scale is indicated by arrows (Goller *et al*. 2011)., *Nanotechnology*, **22**, art.305402, 2011.

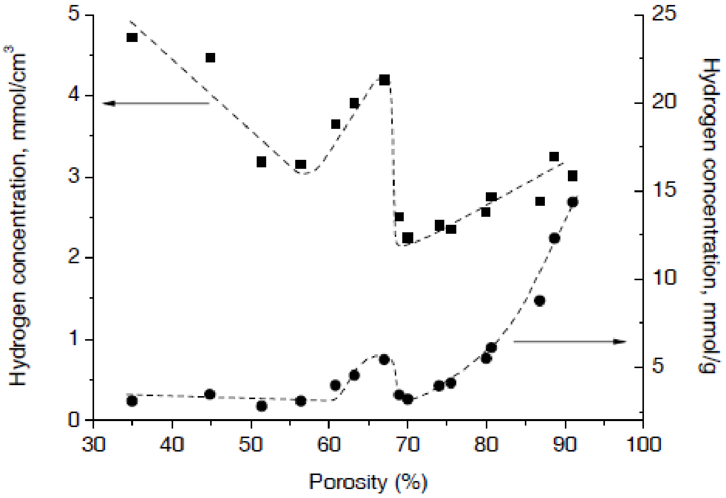


Fig. 9.6. Hydrogen concentration variation along with porosity. Dashed lines are added for eyes guiding. Lysenko V., et al.. *Appl. Surf. Sci*. **230**, 425 – 430

Fig. 9.6 shows the concentration of hydrogen in mmol/cm^3 units stored in meso-porous Si as a function of the layer porosity. As one can note, this dependence is strongly non-monotonous. The hydrogen concentration oscillates in the range 2–5 mmol/cm^3. According to this figure, the concentration value of atomic hydrogen situated at the surface of meso-porous Si of 60% porosity is about 3.5 mmol/cm3 or 3.8 mmol/g taking into account the porous layer density (0.93 g/ cm3). It corresponds to 1.9 mmol/g of molecular hydrogen which is very close to the value extracted by Rivolo et al. from thermal desorption measurements (2 mmol/g). In terms of mmol/g units, being quite constant in the 30–60% porosity range (about 3 mmol/g) the concentration values rise sharply starting from 70% porosity

The theoretical hydrogen storage capacity of different porous Si nanostructures is considered in (Song and Wu 2012), (Bunker and Smith 2011).

9.3 Hydrogen generation from porous silicon

The most widely used methods for H_2 production from PS are:

1. Heating (Gupta *et al*. 1988), (He *et al.* 2013), (Lysenko *et al.* 2005), (Martin *et al.* 2000), (Rivolo *et al.* 2003). This process is accompanied by thermal dissociation of SiH_x bonds and desorption of physically adsorbed atomic and molecular hydrogen.

2. Photodesorption (Tuyen *et al.* 2001), (Collins *et al.* 1992), (Oh *et al.* 2012), (Koropecki *et al.* 2004). Absorption of photons with energy equal to or higher than corresponding binding energies of SiH_x species is accompanied by desorption of hydrogen. Usually, PS is illuminated with ultra-violet laser (Collins *et al.* 1992) or visible light (Oh *et al.* 2012).

3. Chemical reaction with an oxidizer, for example, water (Barabash *et al.* 2006), (Tutov *et al.* 2002), (Bahruji *et al.* 2009), (Tichapondwa *et al.* 2011), (Lysenko *et al.* 2005), (Zhan *et al.* 2011). The reaction between PS and water causes desorption of hydrogen. The presence of bases, even in small amounts, drastically intensifies the reaction.

9.3.1 Thermally Stimulated Desorption of Hydrogen

The desorption of hydrogen from PS is occurred when the latter is heated and it depends on how the hydrogen bound to the porous matrix as well. The convenient method to study the influence of temperature on hydrogen desorption is Temperature-Programmed Desorption (TPD) method. Comparison between TPD curves and mass spectrum of the desorbed species shows that the desorbed phase is almost completely constituted by hydrogen molecules (Rivolo *et al.* 2003). A negligible contribution from desorbed silanes is found to be at least 2 orders of magnitude lower than that observed for hydrogen. Effusion curves for hydrogen desorption from fresh and aged nano-PS samples prepared from p-Si are presented in Fig.9.7 (Lysenko *et al.* 2005).

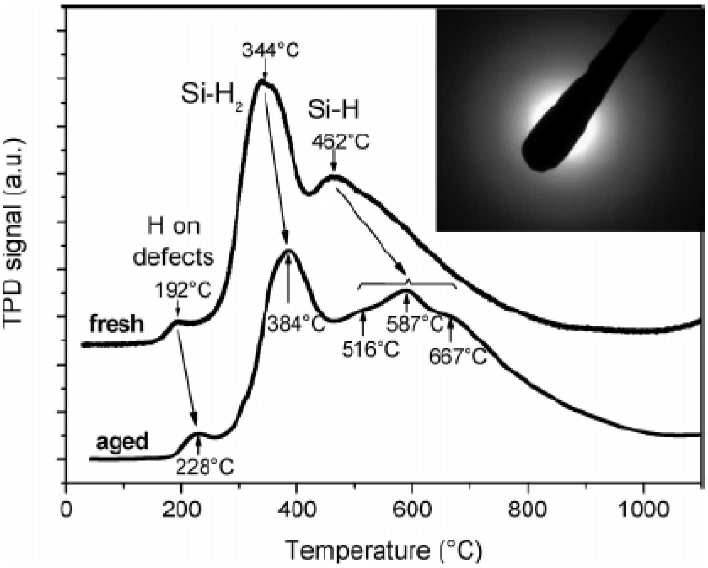


Fig.9.7 Effusion curves for H_2 desorption from fresh and aged nano PS samples. Electron diffraction image given in insert reflects amorphous structure of the PS samples. Lysenko V., (2005). *J.Phys. Chem. B.* **109**, 19711-19718.

Three main regions correspond to the hydrogen desorption from the fresh PS samples: (i) a weak signal at 190 °C, (ii) a more intense one centered at 350 °C, and (iii) another one situated around 460 °C. The two last intensive peaks related to the hydrogen desorption from Si-H_2 and Si-H bonds are also observed on PS formed on *n*-type Si wafers (Martin *et al.* 2000). The first low temperature peak is assumed to correspond to hydrogen desorption from numerous structural defects present in the high nanoporous materials. Indeed, as indicated by the electron diffraction image (see insert in Fig.9.7), the PS used for the thermally stimulated desorption study has amorphous structure, and therefore, atomic and/or even molecular hydrogen can be easily trapped in such completely disordered PS nanostructures (Kitajima *et. al.* 1999). TPD signal of the PS sample after about 50 days of storage mainly preserves its general shape and is shifted by about 40 ºC toward the higher temperatures in comparison with the signal from the fresh samples. This shift corresponds to the storage-induced slight oxidation of the PS samples (Benilov *et al.* 2007) resulting in a slight increase of binding energies of Si-H_x bonds in which the Si atoms are back-bonded to one or more oxygen atoms, as can be deduced from IR spectroscopy (Gupta *et al.* 1991). Indeed, after 50 days of storage, the PS sample has lost about 18% of the initial quantity of hydrogen. The loss of hydrogen is accompanied by an increase of silicon oxide amount due to spontaneous oxidation of PS nanocrystallites in ambient air, as can be concluded from the growth of the IR absorption band corresponding to the Si-O vibrations (Lysenko *et al.* 2005).

The integration of the TPD signal corresponding to the fresh PS sample yielded a maximal value of the total amount of desorbed molecular hydrogen (H_2) being around 33 mmol/g. It corresponds to 66 mmol/g of atomic hydrogen chemically bound at the PS-specific surface. Taking into account that 1 g of Si contains 36 mmol of Si atoms, one can roughly conclude that each Si atom in the PS sample is bound to 2 hydrogen atoms. From the fractal model point of view (Lysenko *et al.* 2004), it means that the fractal dimension of the nanoparticles constituting the PS samples tends toward 2. It means that small (1-3 nm) nanocrystallites constituting the nano-PS layer can be represented as a binary solid constituted by Si and H atoms (Nychyporuk *et al.* 2005).

Despite the fact that the thermodesorption is useful method to define the quantitative of hydrogen in PS and its chemical bonding, from practical point of view this method is rather inconvenient as energy source for portable devices.

9.4 Chemically Stimulated Desorption of Hydrogen

It is well-known that Si-H bonds in the organic compounds react quite easily with water in the presence of bases (Fleming 2001). Evolution of H_2 and formation of Si-O bonds are the results as shown by the following reaction:

(9.6)

The reaction (6) is accompanied by H_2 gas bubbling and can be used for hydrogen desorption from the surface of meso- and nano-PS layers (Lysenko *et al.* 2005), (Goller *et al.* 2011). Application of FTIR spectroscopy allows to control the reaction (6) via analysis of Si-H absorption modes. As it can be seen in Fig.9.8, the intensity of the Si-H_x wagging band near 640 cm^{-1} decreases to 25% after the sample treatment by the solution with 10^{-4}mol/L NH_3 concentration. It should be mentioned that the quantity of NH_3 molecules in such a solution is at least 50 times lower than the quantity of Si-H_x groups in the treated sample. Therefore, the ammonia acts rather as a catalyst than as a reagent in this reaction. In contrast, a treatment with 0.1 mol/L NH_3 solution results in complete desorption of hydrogen from the PS surface, which is illustrated in Fig.10-06 by the disappearance of the Si-H_x wagging band. Treatment of the nano-PS samples of 80% porosity with 10^{-4}mol/L NH_3 solutions results in a 75% decrease of Si-H wagging band. The increase of the treatment efficiency of the nano-PS in comparison with the meso-PS samples is due to the much higher values of the specific surface n the nano-PS structures.

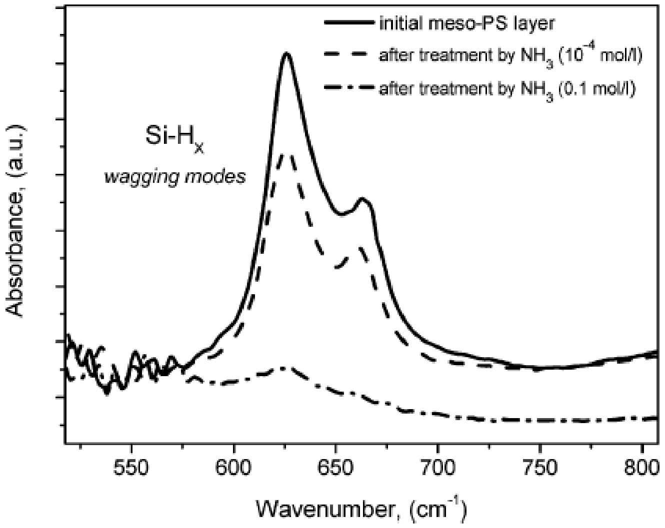
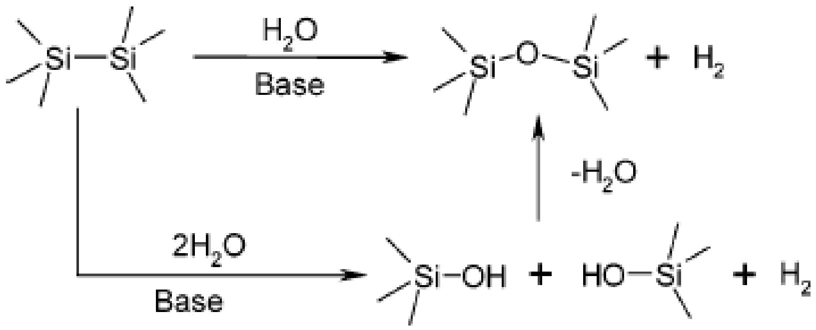


Fig.9.8. Chemically stimulated desorption of hydrogen from meso-PS layers. IR absorption spectra of Si-H_x wagging band corresponding to the hydrogenated meso-PS layers treated by NH_3 solutions during 20 min at room temperature. Lysenko V. (2005). *J.Phys. Chem. B.* **109**, 19711-19718.

It should also be mentioned that the reaction (6) is not the single one taking place during the treatment of the PS samples with bases. Indeed, adsorption bands situated near 2260 cm^{-1} (stretching of Si-H bonds in SiO_ySiH_x complexes) and near 1100 cm^{-1} (stretching of Si-O-Si fragments) present in the IR spectra of the PS samples treated with NH_3 solutions (see Fig.9.9) indicate the interaction of water with numerous highly reactive Si dangling bonds (see reaction scheme 7) appearing at the nanocrystallite surface after the hydrogen release induced by the initial reaction 6:



(9.8)

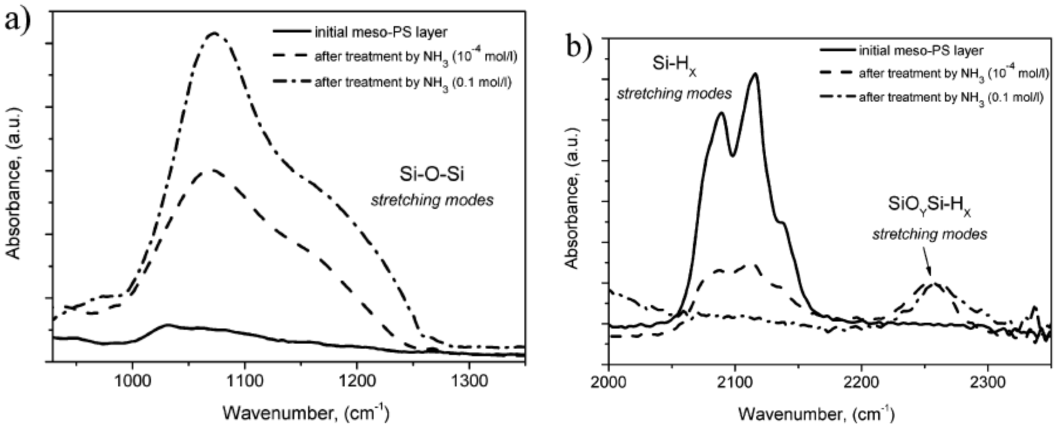


Fig.9.9. IR absorption spectra of: (a) Si-O-Si stretching band and (b) SiO_ySi-H_x band appeared in the PS samples after hydrogen desorption stimulated by NH_3 solutions. Lysenko V., (2005). *J.Phys. Chem. B*. **109**, 19711-19718.

In particular, reaction of the nano-PS samples constituted by the small Si nanoparticles (<5 nm) with 0.1 mol/LNH_3 results in a complete conversion of the PS nanostructure into hydrated dioxide ($SiO_{2x}H_2O$), which was already observed in the IR spectra. (i) Assuming total water induced oxidation of the Si nanocrystallites from the nano-PS samples and (ii) taking into account that the oxidation of each Si atom is accompanied by the formation of two hydrogen molecules according to the following reaction:

$$Si + 2H_2O = SiO_2 + 2H_2, \quad , \qquad (9.9)$$

an additional amount of the molecular hydrogen produced from the oxidation reaction 2 is estimated to reach to a value of 71 mmol/g. Thus, the maximum total amount of the molecular hydrogen issued because of the reactions 7 and 8 can be about 100 mmol per gram of the initial PS nanostructure. Such a theoretical value for the molecular hydrogen that can be obtained from the chemical treatment of the PS layers seems to be quite realistic, especially for the case of nano-PS powder consisting of the smallest

(2-3 nm) amorphous Si nanocrystallites and, therefore, having the highest surface area and smaller nanocrystallite size (Lysenko *et al.* 2005).

An important issue is the ratio between the amount of hydrogen which is generated due to the rupture of silicon-hydrogen bonds in the porous silicon (reaction 7), and hydrogen, which is released by splitting water in presence of silicon nanoparticles (reactions 9 and 8).

Obviously, this ratio will be determined by 1) size of silicon nanoparticles, 2) method of PS production, 3) efficiency of surface coverage by hydrogen, 4) chemical reaction conditions (temperature, catalyst type and its concentration). Impact of these parameters on hydrogen generation for various type of PS and Si microparticles were studied in (Erogbogbo *et al.* 2013), (Goller *et al.* 2011), (Manilov *et al.* 2010a), (Litvinenko *et al.* 2010), (Kale *et al*. 2012).

According to the reaction (8), all produced hydrogen is due to water dissociation, while the reaction of SiH_2 with water results:

$$SiH_2 + 2H_2O = SiO_2 + 3H_2 \quad , \qquad (9.10)$$

e.g. only 2/3 of the obtained hydrogen is due to the spleating of water molecules, and the rest 1/3 is due to rupture of $Si\text{-}H_2$ bonds.

The conceptual illustration of a shrinking core model describing H_2 generation from 10 nm Si nanoparticles prepared by CO_2 laser pyrolysis synthesis of silicon nanoparticles is shown in Fig.9.10. At the beginning of chemical reaction the first 0.5 mmol of H_2 are generated from breaking of Si-H bonds, the following 2.0 mmol are generated from water splitting which is accompanied by size reduction of the silicon nanocrystalline core and the silica cap grows (Erogbogbo *et al.* 2013). The sequence in the reaction of hydrogen evolution from first silicon-hydrogen bonds and then the oxidation process is confirmed by the value of the energy of these bonds (Fig.9.11).

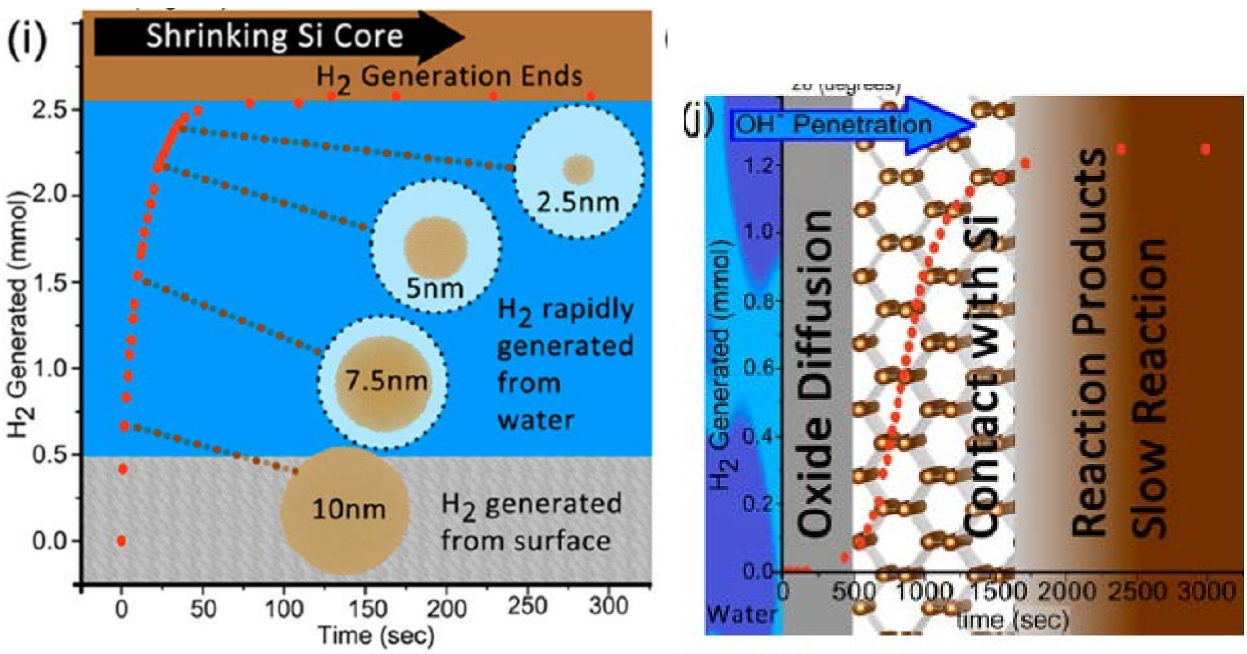


Fig.9.10. a)Shrinking core model describing H_2 generation from 10 nm Si NPs. Brown spheres represent cristalline Si, and the dotted circle indicates shrinkage from the original sphere size. The gray and blue areas represent hydrogen quantity generated from the hydrogenated silicon and water spliting, respectively.(b) A bulk diffusion model for reaction of larger particles for which OH^- diffuses through an oxide layer during the induction period and reaction rates remain limited by diffusion through reaction products. Erogbogbo F., (2013) *NanoLett.* **13**, 451–456

The oxidation processes are claerly can be seen from FTIR absorption spectra of as-prepared PS and after reaction with water (fig.9.12, 9.13).

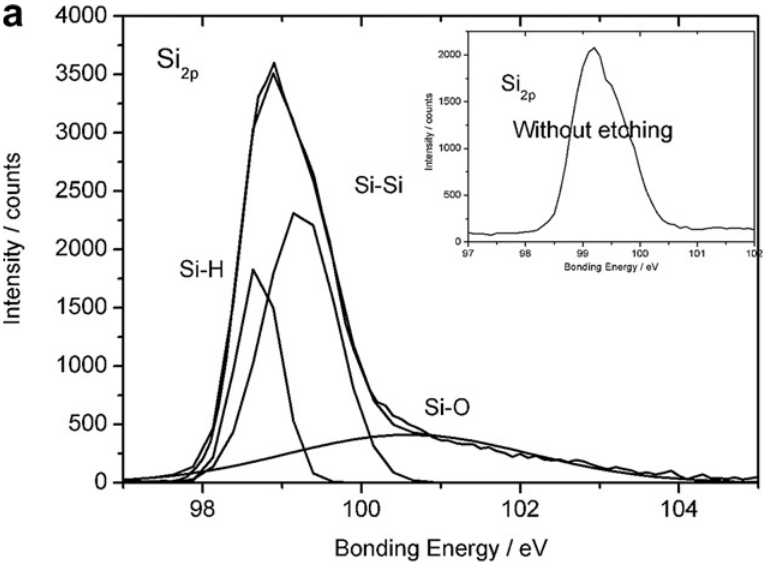


Fig.9.11 Fitted Si2p XPS spectra: (a) freshly prepared porous silicon and bulk silicon and Changyong Zhan *Int. J. Hydrogen Energy*, 36 (7), 4513−4517.

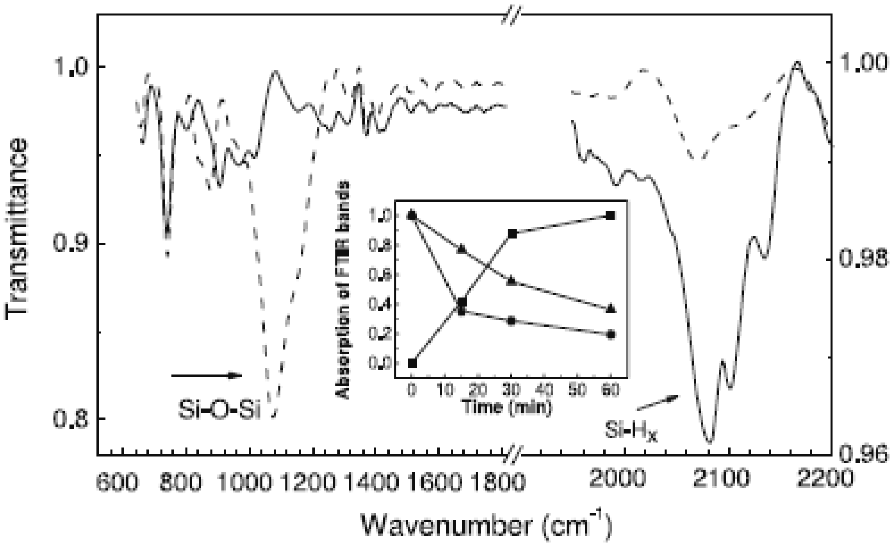


Figure 9.12. FTIR absorption spectra of as-prepared PSi powder (solid) and PSi powder after oxidation in 1:1 water/ethanol solution for 30 min (dashed). Dominant Si surface bonds are indicated. Inset: integrated area below Si–H and Si–O related absorption bands of absorption spectra versus exposure time of PSi powder in the solution. Squares: Si–O–Si bond (1070 cm−1*)*; circles: Si–H stretch bond (2080 cm−1*)*; triangles: Si–H bending mode (622 cm−1*)*.

Goller *et al*. 2011)., *Nanotechnology*, **22**, art.305402, 2011.

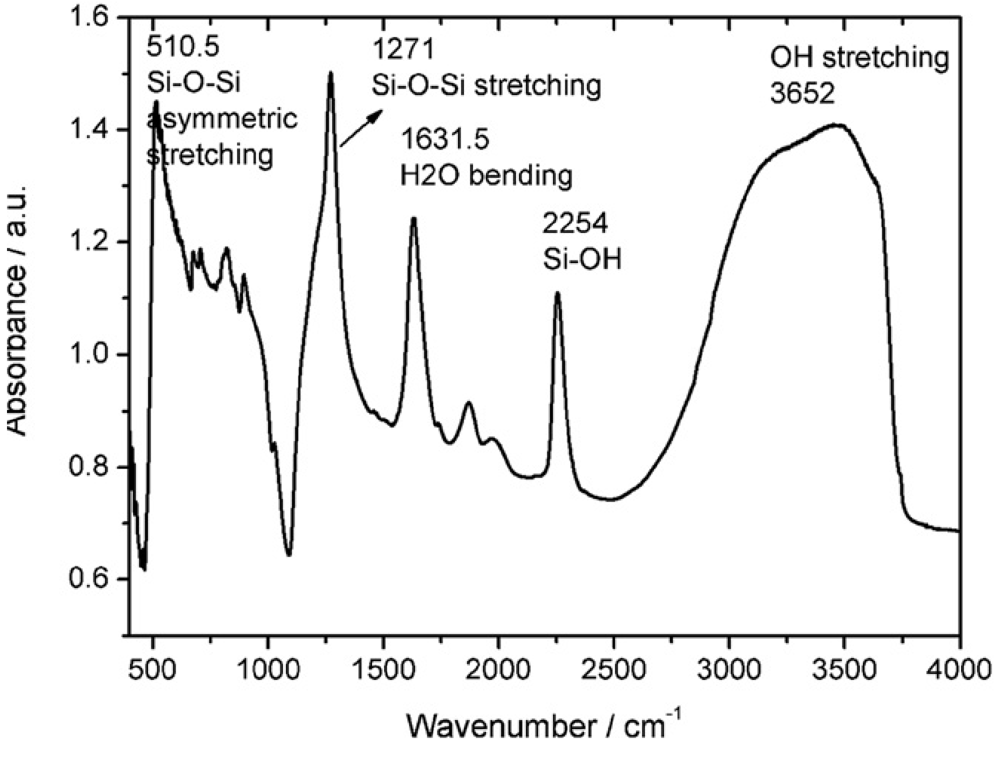


Fig.9.13 FTIR spectrum of the transparent product produced from porous silicon. Changyong Zhan *Int. J. Hydrogen Energy*, 36 (7), 4513−4517.

As shown above, a large quantity of chemically bound hydrogen can be obtained from anodic etching or chemical syntesis methods. However, for high volume production of PS at an industrial level, an electroless technology to form PS should be prefarable. Therefore, electroless stain etching of silicon powder seems to be more attractive than electrochemical etching approach. Moreover, since the total surface area of initial silicon powder is much bigger than the area of bulk silicon substrate for the same mass, the PS production time for crystalline powder etching can be considerably reduced.

Stain etching process of silicon micro-powder was reported in (Litvinenko *et al.* 2010). The metallurgical p-type silicon, 0.001- 0.003 Ohm.cm with size of particles up to several microns was used. Since size distribution in the initial powders is very large, narrower size selection of Si micro-particles constituting the powders was carried out by sedimentation technique. SEM images of the powder fractions obtained during 8 and 120 min of the sedimentation step are shown in Fig. 9.14 (a) and (b), respectively. The used etching solution was H_2O: HF: HNO_3 in proportion 20:4:1 (Polisski *et al.* 2008). The duration of the etching process varied in the range of 10-60 min depending on the temperature and on the HNO_3 concentration. The etching reaction was stopped when the powder became brown. A special attention was paid to prevent dissolution of the created porous shell around the solid Si core. The reaction product in a foam form was collected, rinsed by water, ethanol and dried at room conditions.

According to equation (6) (this process occurs in the presence of a base as a catalyst), 1 g of Si provides L= 1.6 L/g of H_2 at normal conditions. SiH_2 is considered to be the most advantageous solid container of H_2, so one can suggest it as chemical formula of an ideally hydrogenated PS. It also reacts with water and supplies 3 molecules of H_2 (equation 6-8). In this case, 1 g of SiH_2 supplies N= 2.24 L/g H_2. Since PS powder contains both crystalline and hydrogenated phases, both reactions occur, and the output of H_2 should be between 1.6 and 2.24 L per 1 g of the initial Si powder. Therefore, it is possible to calculate the SiH_2/Si ratio from the extracted H_2 volume. The mixture of X grams of Si and Y grams of SiH_2 emits V=LX+NY liters of hydrogen. The relative part of SiH_2 in the best PS powder is Y/M =0.4, and the best PS corresponds to the formula: $SiH_{0.8}$. Thus, it is supposed to correspond to the substance of the shell of the stain etched particles. Such data for porous silicon powder fabricated from different size initial silicon powder is shown in Fig.9.15 (Litvinenko *et al.* 2010).

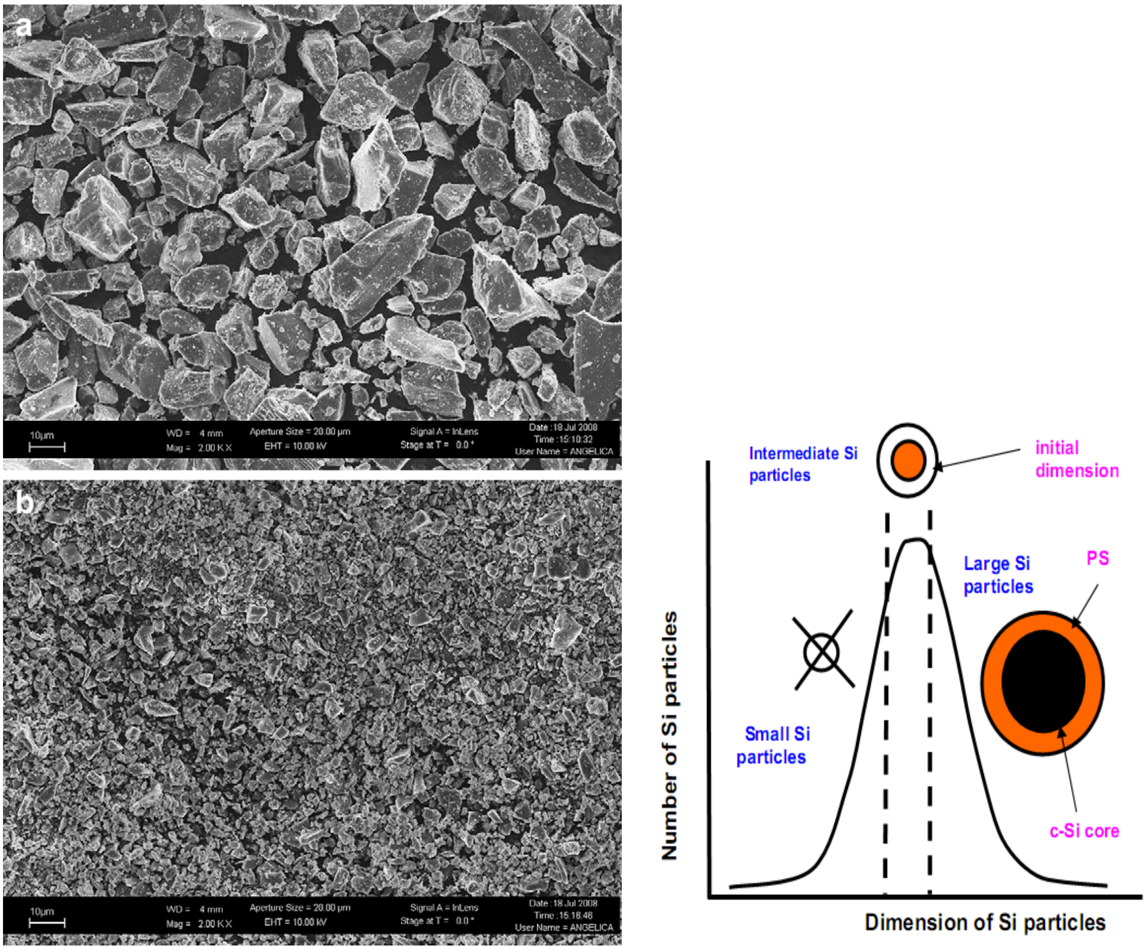


Fig. 9.14 SEM images of the Si powder fractions obtained during 8 (a) and 120 min (b) of the sedimentation step. C) Model of stain etching on Si microgranuls. Litvinenko S. et al., *Int. J. Hydrogen Energy*, 35, 6773, 2010.

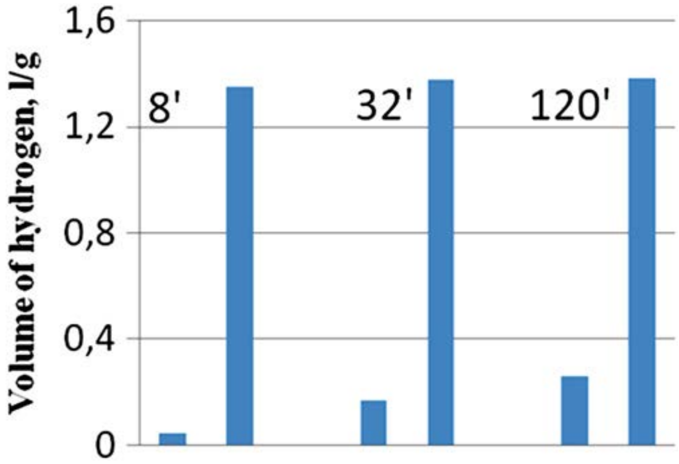


Fig.9.15 Volume of hydrogen produced from stain etched silicon powder for different size fractions of the initial powders. Left (small) peaks correspond to SiH_2, right ones to H_2O for each fraction. Litvinenko S. et al., *Int. J. Hydrogen Energy*, 35, 6773, 2010.

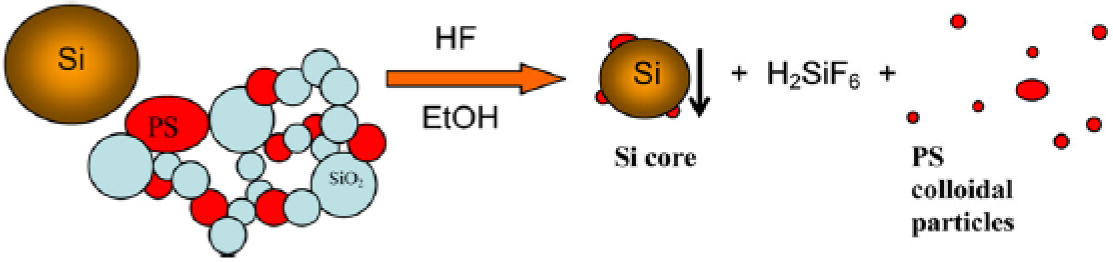


Fig. 9.16 Schema of treatment of stain etched Si in HF:EtOH solutions.

Presence of Si-O bands in the spectrum of stain-etched PS, as well as the loss of SiHx groups under HF treatment can be explained by existence of partially oxidized porous Si shell on the top of the crystalline silicon cores. In this case, the SiO2 dissolves in HF, and PS forms colloidal particles, which are not precipitated during the sedimentation as it is schematically represented in Fig. 9.16. Relative amount of such colloidal phase, however, is not significant, since it originates from the part of the particles’ shells. Small PS particles detached from the crystalline silicon cores are covered predominantly by SiH2 fragments indicating high fractality of the surface (flat Si surface treated in HF solution is covered by SiH groups).

9.4.3 Photocatalytic Hydrogen Evolution

Porous Si is also considered as a perspective candidate for hydrogen-generating materials due to photocatalytic/photoelectrochemical reactions with water (Erogbogbo *et.al*. 2013), (Dai *et al.* 2014), (Mathews *et al.* 2003), (Bahruji *et al.* 2009). Photocatalytic H_2 generation takes advantage of some of the unique abilities of semiconductors; after absorbing photons, electrons (holes) can be generated at the conduction band (valence band) of semiconductors. If the generated electrons and holes have enough energy to overcome the energetic barriers of both water reduction and oxidation and any over potentials, photocatalytic water splitting will take place (Kudo *et al.* 2009). Contrary to usual widegap photocatalist materials like TiO_2, CdS, ZnS-$CuInS_2$ etc., for which UV illumination is necessary to split water, the narrow 1.17 eV band gap of Si is quite appropriate to be photoexcited by the visible part of the solar spectrum, allowing photon capture down to the near-infrared region.

The photocatalytic activities of mesoporous silicon under different illumination conditions were studied by Dai *et al.* (2014). They elaborated a bottom-up synthesis of mesoporous crystalline silicon materials with high surface area and tunable primary particle/pore size via a self-templating pore formation process. The chemical synthesis is based on the use of internal self-forming templates that can be easily removed without any etchants. The advantages of these materials, such as their nanosized crystalline primary particles and high surface areas, enable increased photocatalytic hydrogen evolution rate and extended working life. Such material manifested a certain reactivity in presence of water even under dark conditions (Fig. 9.17). The increased reaction rate of such sample is attributed to its enlarged surface area, which provides extra contact between Si and water. This relationship between surface area and reaction rate is further demonstrated by a chemical reaction of a sample with KOH aqueous solution. An extremely high average H_2 generation rate of 0.095 g H_2 s^{-1} g^{-1} Si (1.33 mol H_2 s^{-1} mol^{-1} Si) was obtained, which is about 30 times higher than the best reported result (0.003 g H_2 s^{-1} g^{-1} Si) (Erogbogbo *et.al*. 2013). In addition, mesoporous crystalline silicon materials demonstrated an efficient photocatalytic reactivity under a visible light illumination (Fig. 9.17).

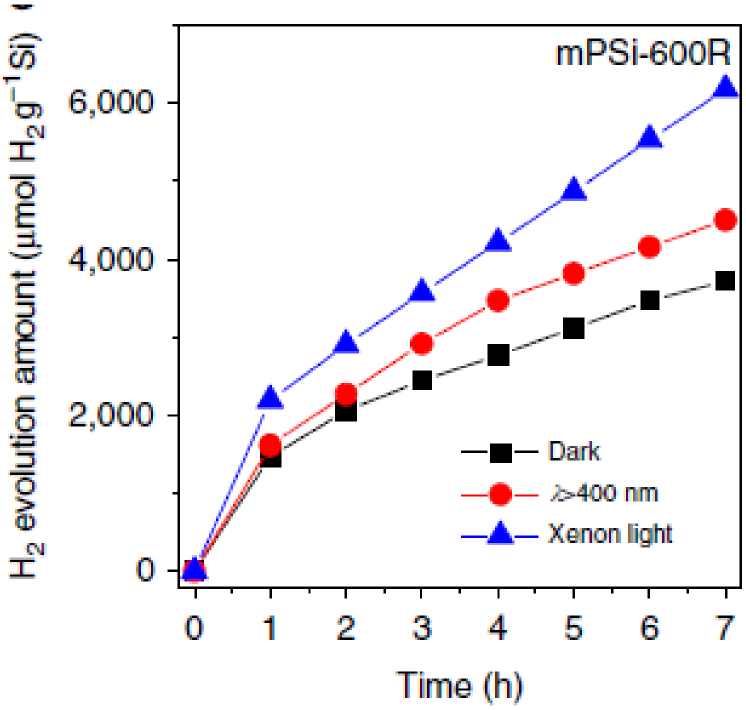


Fig.9.17 Comparison of H_2 evolution activities of mesoporous Si under different illumination conditions. Dai F. et al., *Nature Communications*, 5, art. 3605, 2014

9.5 Rate of hydrogen generation

To determine hydrogen extraction kinetics from the hydrogenated PS, it is possible to use different volumetric or gravimetric measurements (Goller *et al.* 2011) or recording fuel cell electrical output which is proportional to the hydrogen quantity produced during the reactions of PS with water (Manilov *et al.* 2010a). The first measurement technique gives directly the volume or mass of the produced hydrogen, while the last one (see scheme in Fig. 9.18) is much more resolved in time but it has to be calibrated. The calibration could be made, for example, by direct comparison with the first direct methods. A typical curve of hydrogen desorption from the electrochemically formed PS is shown in Fig.9.18 (Litvinenko *et al.* 2010).

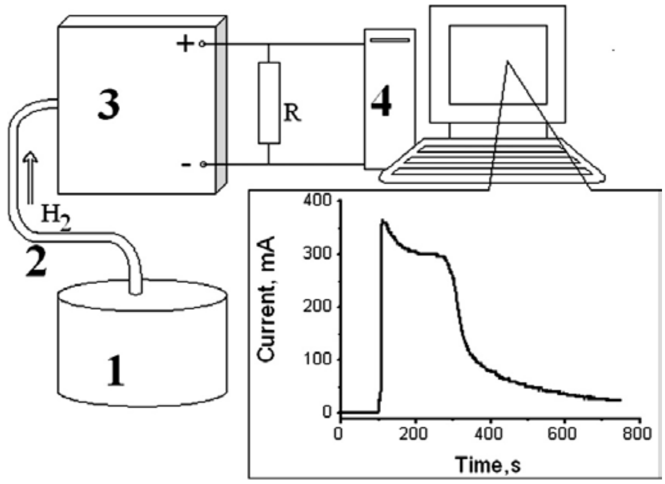


Fig. 9.18 Scheme for measurement of hydrogen desorption with a fuel cell : 1- vessel for chemical reaction, 2 - tube for hydrogen delivery, 3 - fuel cell, 4 - computer with ADC
Litvinenko S. et al., *Int. J. Hydrogen Energy*, 35, 6773, 2010.

In water, freshly prepared Si nanostructures cannot be dispersed because of the hydrophobic nature of their hydrogen-terminated surface while addition of alcohol provides their efficient wetting by water (access of OH^- ions to the nanosilicon surface) and allows dissolution of reaction products. Fig.9.19 shows the rate of hydrogen generation($\Delta W/W$) in wt% with respect to the amount of PS immersed in water/ethanol solution versus exposure time. This experiment has been performed for 2h and the achieved $\Delta W/W$ value was about 9% (squares in fig.9.19). Since ethanol contains about 6% of water, an oxidation process is also possible when PS powder is immersed in hydrous ethanol, however the efficiency of the process is much lower (circles in fig10-13). It was found that about 1:4 mixture of alcohol and water results in the highest reaction rate. Theoretical limit for the $\Delta W/W$ value, assuming that four hydrogen atoms can be generated per single Si atom, is 14%. This value has not been finally achieved because porosified polycrystalline silicon grains, according to direct observation in TEM, have a small bulk core which cannot be completely oxidized (Goller *et al.* 2011).

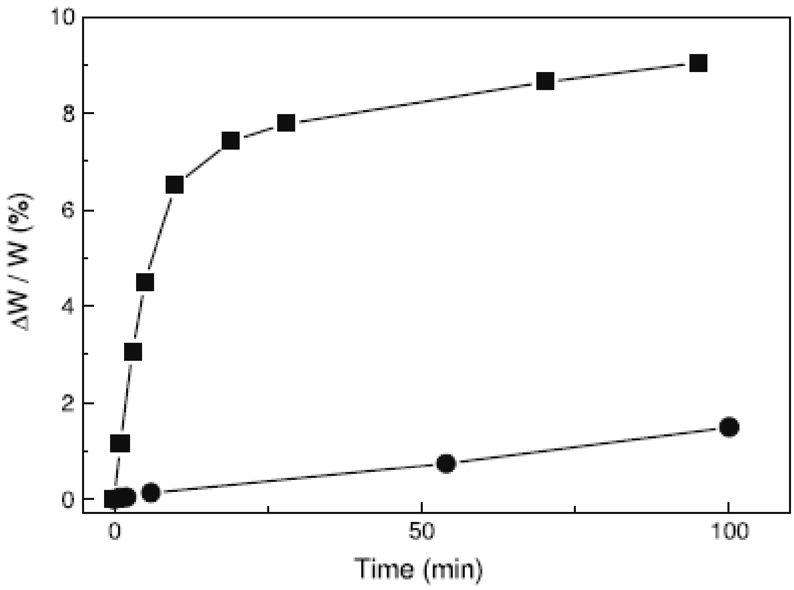


Fig.9.19. Rate of hydrogen generation in wt% versus exposure time of PS to water/ethanol mixture (1:1) (squares) or to hydrous ethanol (circles). Goller B. et al., *Nanotechnology*, **22**, art.305402, 2011

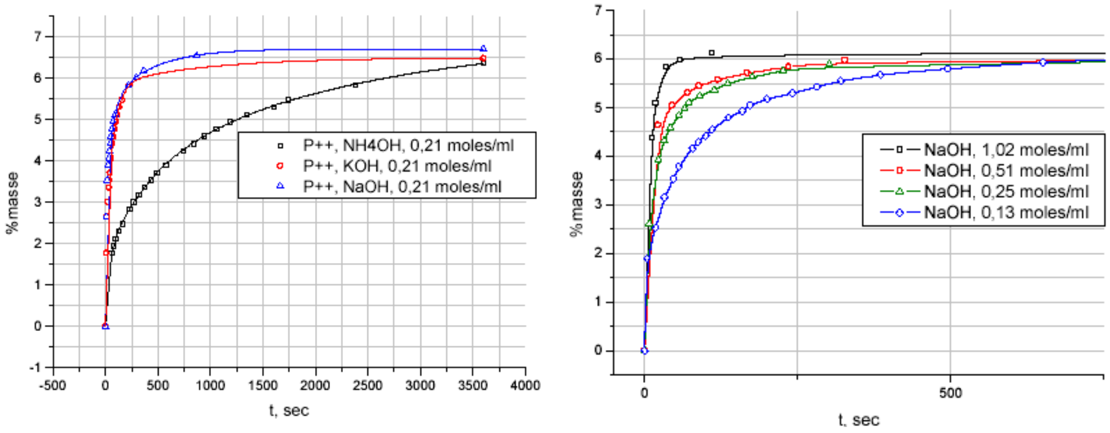


Fig.9.20 a) Influence of base nature on hydrogen desorption kinetics, b) Influence of sodium hydroxide concentration on hydrogen desorption kinetics.

The adding of different bases (acting as catalyzers) drastically increases the generation rate. We tested three different bases: ammonia (NH_4OH), sodium hydroxide (NaOH) and potassium hydroxide (KOH), for identical concentrations of 0.21 mol/l. This experiment was carried out to select a basis ensuring on the most rapid desorption kinetics and the highest overall quantity of the generated hydrogen. The results are shown in Fig.9.20a. One can conclude that the nature of bases has a negligible influence on the amount of the generated hydrogen. In contrast, the rate of hydrogen production depends significantly on the used base. The similar result was obtained from the study of influence of base concentration on the hydrogen generation process (Fig.9.20b).

Comparing results of hydrogen generation from PS obtained by electrochemical etching and stain etching, one can state that H_2 is generated more rapidly with meso-PS fabricated by electrochemical etching (Fig.9.21). Chemically etched micropowders (Fig.9.14) treated with the H_2O:HF: HNO_3 solution are characterized by slower rates of H_2 formation which decreases with increase of the particle size. The lowest intensities of the reaction are recorded in non-etched silicon powders. It is worth noting that a noticeable reaction of the solution containing crystalline Si is initiated at a temperature not lower than 50°C while PS reacts already at room temperature (Manilov *et al.* 2010b).

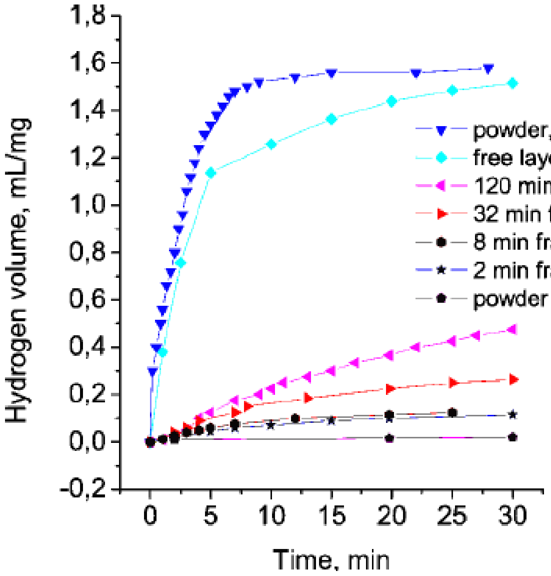


Fig.9.21 Kinetics of hydrogen generation during 30 minutes of the oxidation reaction in dark from various silicon fractions at 50°C (Manilov *et al.* 2010b). Manilov A.I. et al., *Ukr. J. Phys.*, **55**, 928, 2010

The most intense generation of hydrogen from PS is observed during the first minutes of the reaction. Such behavior of the curve is determined by the immediate breakage of Si–H_x surface groups. The developed PS surface allows a considerable intensification of the reactions of Si–Si and Si–H bonds with water. The former effect makes a noticeable contribution to the rate of H_2 generation because silicon is less resistant to the oxidation as compared with Si-H groups, which is explained by the lower energy of the Si–Si chemical bonds (Barabash *et al.* 2006). PS obtained by electrochemical etching contains the largest amount of the porous phase and, therefore, the higher number of Si-H_x groups with $x \geq 2$ (Lysenko *et al.* 2004). The chemically etched Si powders have a less developed surface and, respectively, a smaller number of silane groups and larger dimensions of crystalline Si cores. Moreover, such a worsening of the properties becomes more pronounced with increase in the size of particles. These considerations explain the effect of a decrease of the H_2 generation rate with increase in the size of particles in Si fractions (see Fig.10-15). Nonetched silicon powders do not contain the porous phase and have practically no surface Si–H_x groups. In addition, the surface of non etched silicon is partially oxidized. This results in low intensity of H_2 generation and the absence of the reaction with water at room temperature. The heating of a solution causes a rise of the penetrability of oxide layers and, therefore, stimulates the reaction (Manilov *et al.* 2010b). The increase in the ammonia concentration in the solution results in the intensification of the reaction. The rise of H_2 generation rate by 5–10% was recorded due to the two-fold increase in the ammonia concentration. Moreover, no effect of the ammonia percentage on the total hydrogen yield was observed, which confirms the catalytic role of NH_3.

The hydrogen generation intensifies due to the growth of the solution temperature. This fact is explained by a higher intensity of the interaction between water and silicon, as well as by an increase in the permeability of the silicon oxide layer. In addition, the only phase participating in the reaction at room temperature is the porous silicon phase, whereas the heating of the solution to 50 ^{o}C activates the reaction of water with the crystalline phase. Respectively, the total yield of hydrogen from PS also increases due to the heating, because the c-Si residuals in the structure also get involved into the reaction. In such a way, one can separate the contributions of the crystalline and porous Si phases to the total hydrogen yield by controlling temperature of the solution.

Illumination of PS favors an increase of the hydrogen yield in the reaction, but has some peculiarities. The illumination by a red light-emitting diode had no effect on the kinetics of H_2 release. At the same time, a blue light-emitting diode (470 nm) resulted in a significant effect (Fig. 9.22). In this case, the temperature of the solution was stabilized in order to eliminate its influence on the kinetics. Similar results in the case of the ultraviolet illumination have been obtained by Bahruji et al. (2009). It is worth noting that a relative increase of the H_2 generation at the illumination (470 nm) manifests itself after a certain time of the reaction (1–5 h). At the initial moments, the kinetics of illuminated samples is similar to that of unilluminated ones.

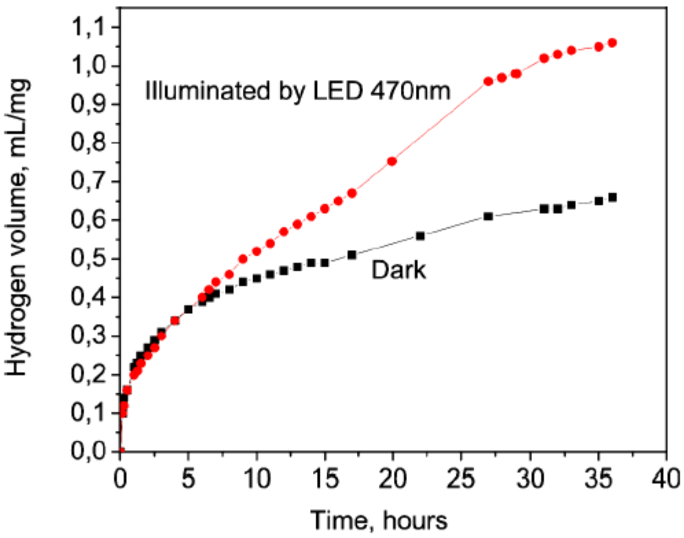


Fig.9.22 Kinetics of hydrogen effusion from PS fraction (32min) in the solution under illumination with a light-emitting diode (470 nm) and in dark (Manilov *et al.* 2010b). Manilov A.I. et al., *Ukr. J. Phys*., **55**, 928, 2010

The maximum and the average rates of hydrogen generation for different silicon powders are shown in Fig.9.23. The maximum rate, time required to generate 1 mmol of H_2, and total produced hydrogen quantity are presented in Table 9.2. The maximum H_2 generation rate with 10 nm Si was 150 times higher than that achieved with 100 nm particles which, in turn, was only 1.3 times than that achieved with 44 μm particles. These differences cannot be accounted by differences in specific surface area. The maximum hydrogen generation from 10 nm silicon is substantially higher than for silicon, stain etched nanoporous silicon, by photoactivated reaction under ultraviolet light and of the the same order of magnitude as nano-Al and nano-Zn (Erogbogbo *et al.* 2013).

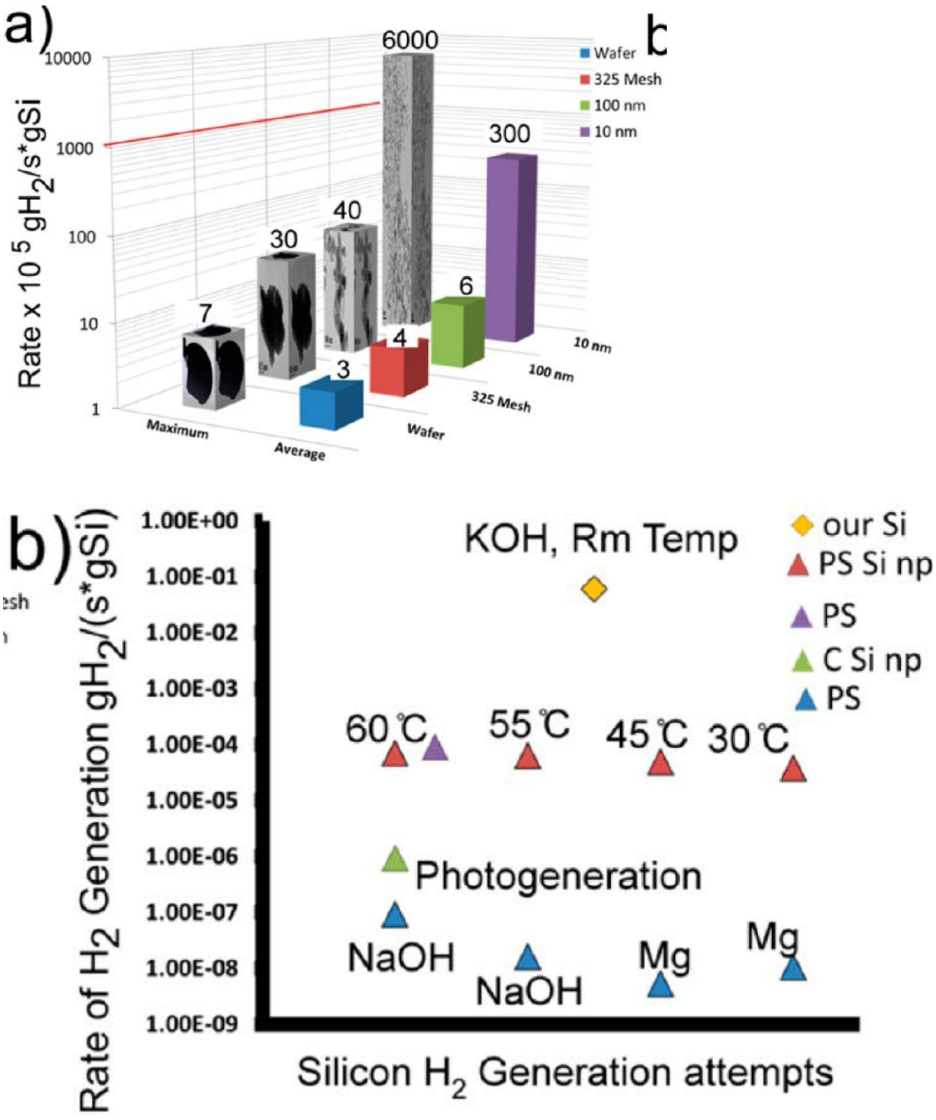


Fig.9.23. . Comparison of hydrogen generation rates for silicon in this and other studies and fuel cell performance using hydrogen generated on demand.(a) Rates for different forms of silicon studied here. Maximum rates are in the left column with images of the samples on them. Average rates are in the right column. The red line indicates the maximum reported rate for hydrogen generated from aluminum. (b) Comparison to literature reports. PS is porous silicon; blue triangles at different horizontal positions represent different chemical conditions. PS Si np is porous silicon with nanoparticles in the pores at the labeled temperatures. C Si is commercial 100 nm silicon nanoparticles. Our Si np refers to silicon nanoparticles around 10 nm synthesized in our laboratories.. Erogbogbo F., et al.(2013) *NanoLett.* **13**, 451–456

9.6 Method of supplementary hydrogenization

The tasks of additional hydrogen accumulation in PS due to absorption in bulk or physical adsorption are still important. Successful experiments of enrichment of crystalline and nano-Si with hydrogen were conducted by way of restraint in molecular H_2 (Manilov, *et al.* 2010a), (Kale *et al.* 2012) or H_2 plasmas (Sriraman *et al.*.2006), (Darwiche *et al.* 2007). Retaining of H atoms and H_2 molecules depends on number of defects, dopants, adsorption centres, which is increasing for more disordered surfaces. The interaction of a hydrogenated PS surface with atomic hydrogen was theoretically and experimentally investigated (Manilov and Skryshevsky 2013), (Glass *et.al*. 1996). The hydrogen accumulation from the gaseous phase is not significant in PS (Manilov, *et al.* 2010a). The main reason is the initial chemically adsorbed monolayer of H atoms existing on the surface of PS (Parkhutik 1999), (Föll *et al.* 2002). Further, it is difficult to incorporate hydrogen inside the nanosilicon particles (Zhen-Yi *et al.* 2012). Therefore, only physical adsorption of hydrogen is possible, due to interaction with Si-H_x species. These processes can be modelled by methods of molecular dynamics in the field of van der Waals forces. The results of calculations show the possibility of physical hydrogen adsorption onto hydrogenated pore walls. However, sufficient desorption of physically adsorbed hydrogen from the PS surface is obtained, which makes the long-time keeping of the adsorbate impossible. Therefore, it can explain the weak additional accumulation of hydrogen from the gaseous phase (Manilov and Skryshevsky 2013).

The maximum experimental hydrogen uptake of 2.25% at temperature of 120°C at 9.76 bar pressure was obtained by Kale *et al.* (2012). In this investigation the Si nanoparticles have been synthesized by sonicating 12 µm thick and 29% porous freestanding PS films for 4 h. Si nanoparticles have been fabricated in the range of 8-20 nm in size. Hydrogen absorption pressure composition isotherm in the pressure range of 1-10 bar and in the temperature range of 29 °C- 150 °C is presented in Fig.9.24a. Room temperature hydrogen storage capacity of Si nanoparticles was obtained about 1.4 wt%. Hydrogen uptake by sample reduces at the temperature of 150^{o}C. This may be attributed to the breaking of some of the Si-H$_x$ bonds and subsequent release of hydrogen. The curves show rising trend with increase in the pressure, meaning hydrogen uptake is directly proportional to the pressure at least in the pressure range of measurements. FTIR transmittance spectrum shown in Fig. 18b exhibits peaks around 640, 900 and 2100 cm^{-1} corresponding to wagging, stretching and bending Si-H bonds produced during SiH$_x$ formation. Increasing sharpness of SiH$_x$ peaks in the FTIR spectra after hydrogenation was attributed to hydrogen uptake by the sample.

Table 9.02. Maximum rate, time required to generate 1 mmol of H_2 and total hydrogen quantitiy mol H_2 per mol Si

Dimention of partic	maximum rate, $gH_2\ s^{-1}\ gSi$	time requir generate 1 m H_2, s	Total hydrogen produced mol H_2 per mo	References
10 nm particles synthese laser pyrolysis	0.06	**5**	2.58	(Erogbogbo *et al.* 2013)
100 nm particles (Sigma Aldrich)	0.0004	811	1.25	(Erogbogbo *et al.* 2013)
<40 µm particles (Sigma Aldrich)	0.0003	3075	1.54	(Erogbogbo *et al.* 2013)
(110) Si wafer	0.00007	12.5 h	1.03	(Erogbogbo *et al.* 2013)
stain etching particles 3- 10 nm	2×10^{-8}			(Goller *et al.* 2011)
Electrochemical etched	2×10^{-8}			(Zhan *et al.* 2011)
Nano PS obtained by sta etching, 60 ^{0}C	1.7×10^{-5}			(Litvinenko *et al.* 2010)
photoactivated reaction UV light	2×10^{-8}			(Bahruji *et al.* 2009)
Photocatalytic hydrogen evolution from self-temp ting synthesis of meso P	0.095		1.33	(Dai *et al.* 2014)
Nano Zn an nano Al	0.01			(Bunker et al. 2011), (Ma and Zachariah 2010

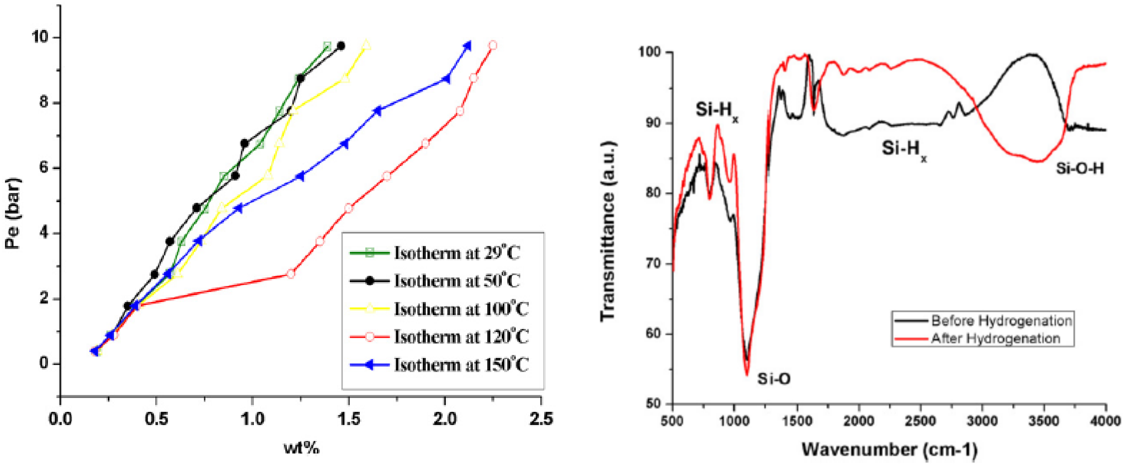
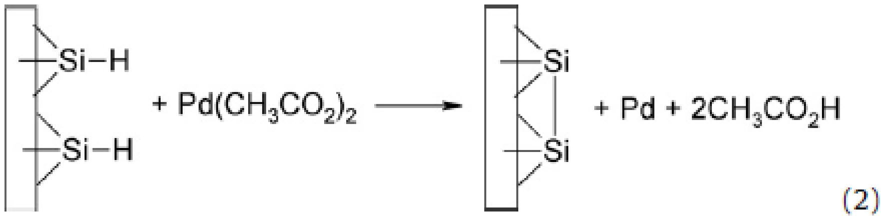


Fig. 9.24a) Hydrogen absorption pressure composition isotherm of Si nanoparticles carried out in Standard Seivert's type apparatus in the pressure range of 1-10 bar and in the temperature range of 29°C-150°C, b) FTIR spectra of Si nanoparticles before hydrogenation and after hydrogenation at P =9.76 bar and T= 150°C Kale P., et al. *Int.J.Hydrogen Energy* **37**, 3741. 2012

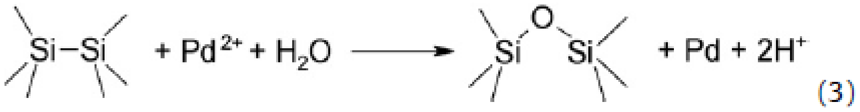
9.7 HYDROGEN STORAGE AND GENERATION IN PS COMPOSITES

The additional hydrogen sorption can be realised by the creation of different PS composites. Thus, creation of a palladium/silicon meso-porous composite is expected to modify the hydrogen generation properties of PS. Firstly, Pd effectively absorbs H_2 molecules dissociating into atoms after interaction with the metal. Therefore, using hydrogen accumulation from a gas phase, an increase of the general hydrogen content in the composite could be achieved (Rahimi and Irajizad 2006), (Rather *et al.* 2007). The composite materials based on porous silicon with palladium particles incorporated inside the pores were prepared by means of Pd electroless deposition by (Manilov *et al.* 2010a), (Manilov *et al.* 2010b) via the following schema:



(9.11)

The first reaction results in decreasing of SiH_x bands intensities in the spectra of PS + Pd samples and can be easily explained by consumption of SiH_x groups in the reaction of Pd^{2+} reduction:



(9.12)

The second reaction leads to appearance of silicon oxide bands and indicates the reduction of Pd^{2+} with the participation of H_2O in such a way that Pd has been included inside the PS nanostructure, but is localized in a near-surface layer (approximately 50μm thick for the PS + 10% Pd sample). Such localization could be explained by the high rate of Pd reduction on the PS surface in comparison with its diffusion inside the PS layer. Furthermore, the Pd particles in the near-surface layer might block the PS pores, preventing further Pd penetration. Apart from the Pd inside the PS nanostructure, it is also located on outer surface of the samples, forming relatively big (20–50 nm) nanoparticles (see SEM image in Fig.9.25).

Adsorption and desorption of molecular hydrogen at 50 kPa and ambient temperature as well as chemical extraction of H_2 during the reaction of initial PS and PS + Pd composites with water and base were studied (Manilov *et al.* 2010a), (Manilov *et al.* 2010b). Reversible uptake of molecular H_2, probably due to the formation of palladium hydride, was detected for PS + Pd composites, otherwise the initial PS was found to be inactive to the H_2 under ambient conditions. The PS + Pd composites react with water much slower and release less H_2 in comparison with initial PS, probably due to partial pore blockage by Pd. However, the pretreatment of the composites in gaseous H_2 increases the volume of H_2 evolving in the reaction with H_2O at ambient temperature (Fig.9.26).

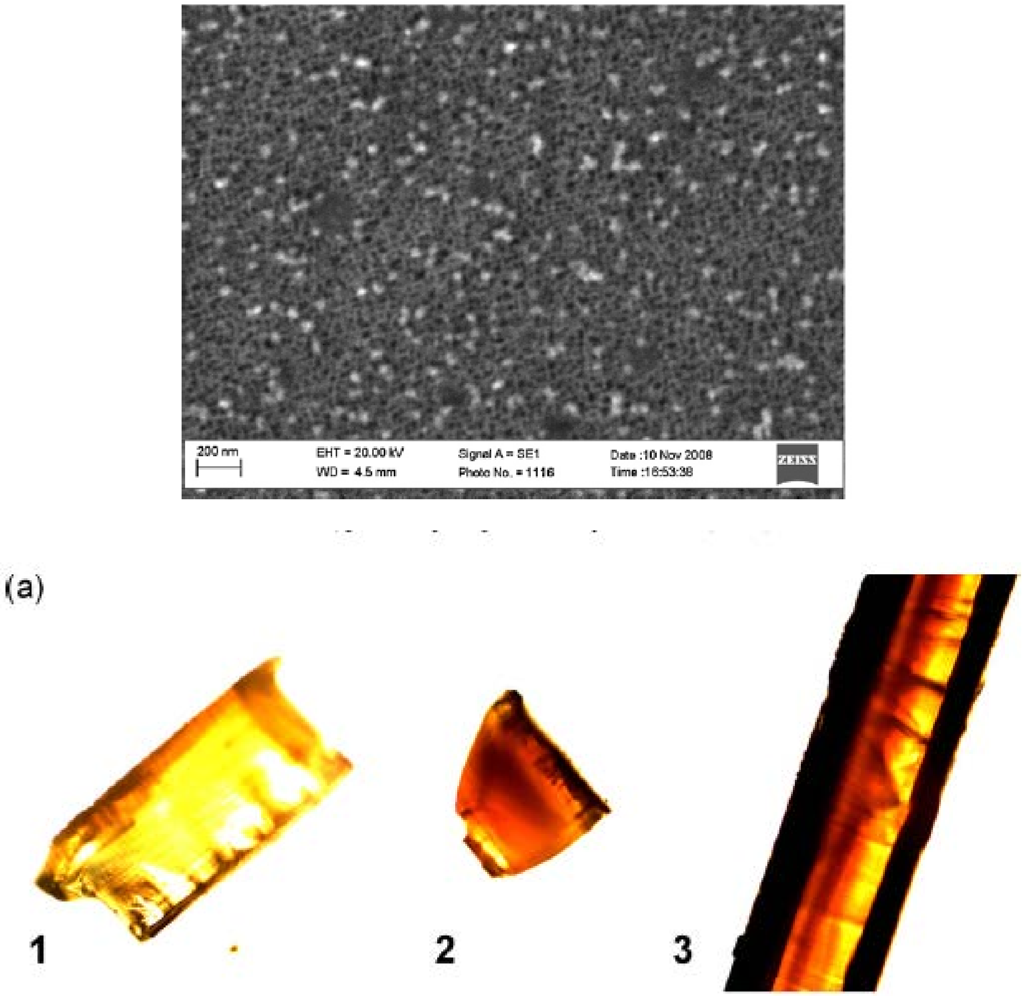


Fig. 9.25 SEM of surface PS + 1% Pd and their optical microphotograph of cleaved facets of the PS, PS + 1% Pd and PS + 10% Pd samples (Manilov *et al.* 2010a). *Journal of Alloys and Compounds* **492**, 466–472, 2010

He *et al.* (2013) proposed to use hydrogen storage system based on hybrid PS-single wall carbon nanotubes (SWNT). The PS was fabricated by anodic etching of silicon wafers in alcoholic solutions of hydrogen fluoride. Direct hydrogen dosing to adsorb hydrogen on the pSi-SWNT hybrid, pure pSi, electrochemically charged pSi and SWNTs, and pure SWNTs, followed by TPD measurements, were used to quantify the hydrogen adsorption process. Before the TPD experiments, a gas mixture containing 10% H_2/Ar was dosed onto the sample at room temperature. The results indicate an increased hydrogen storage capacity at a lower temperature in the pSi-SWNT system relative to that from pure pSi, and pure and charged SWNTs. In addition, hydrogen adsorption in pSi-SWNT is about a factor of 2 to 6 higher than that in pure and charged pSi, and pure and charged SWNTs. In such hybrids SWNT acts as a 'catalyst' that converts inactive Si atoms to active Si atoms towards hydrogen chemisorption; and/or as a bridge for neighbor Si atoms to allow hydrogen spillover to nearby Si atoms. (Fig.9.27).

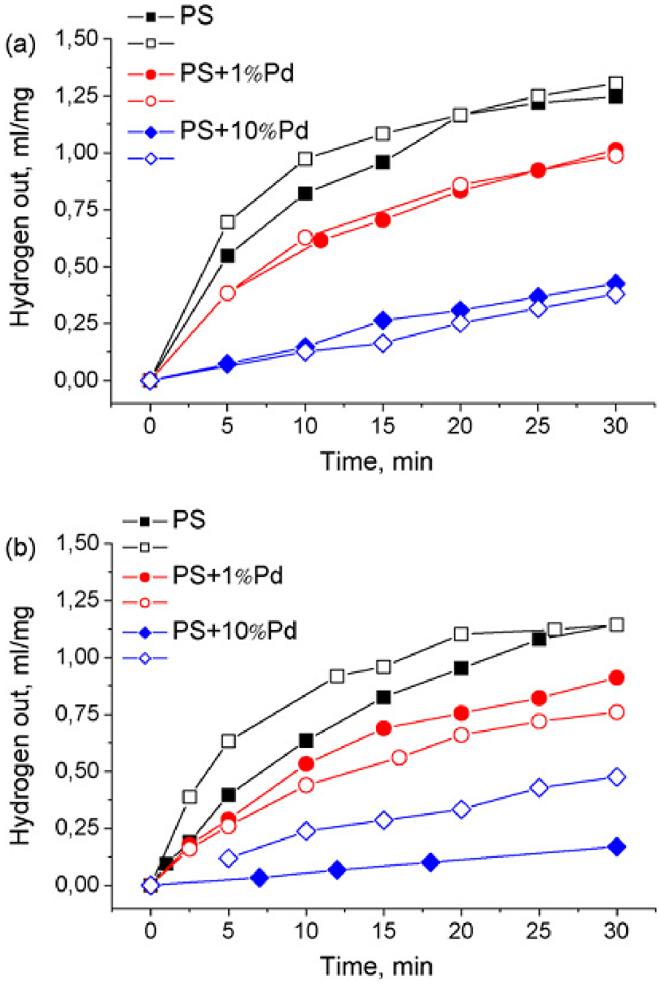
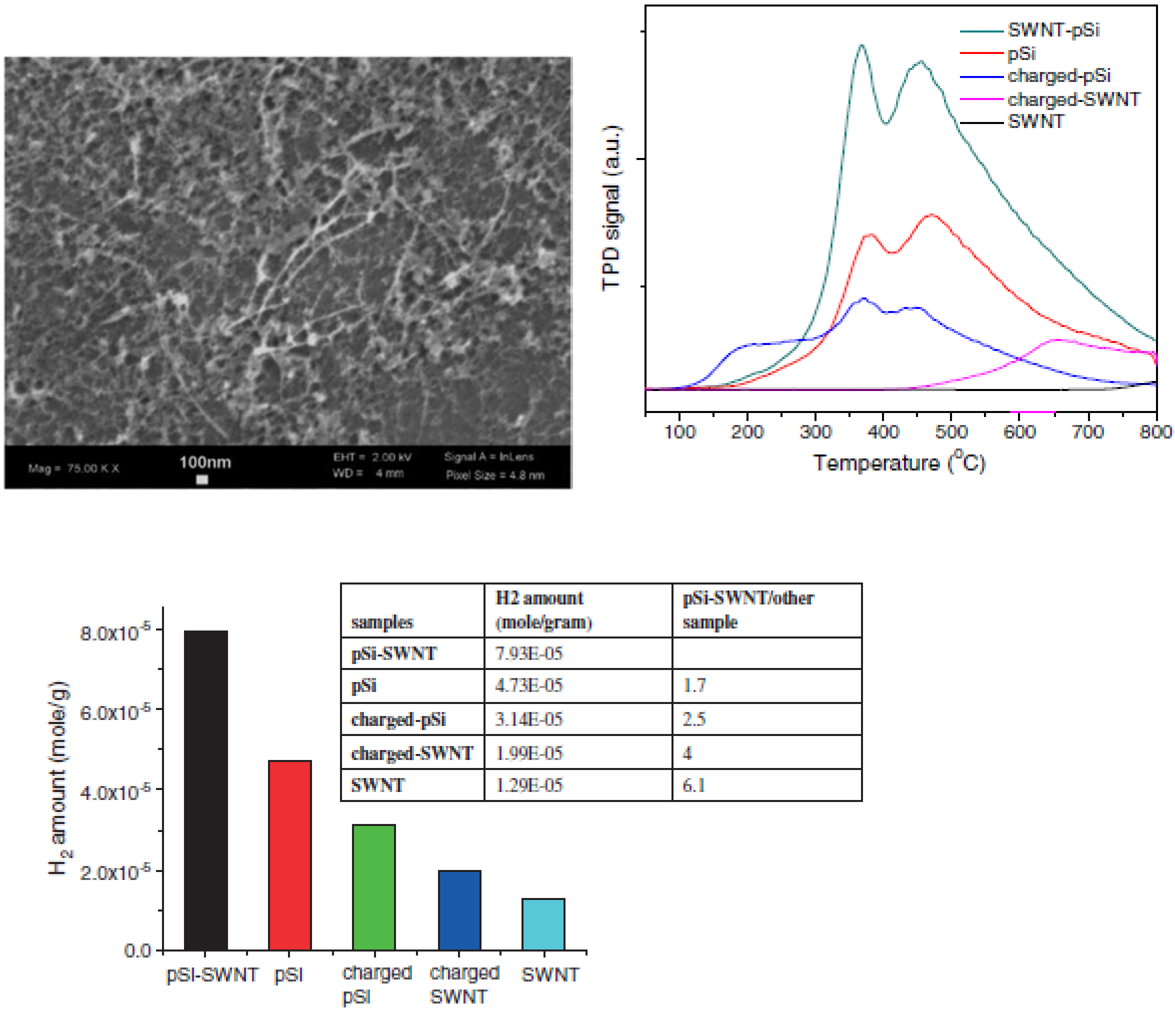


Fig.9.26. Hydrogen extraction kinetic from PS and PS+Pd in H_2O:EtOH:NH_3 (5:2:2) (a) and (5:1:1) (b) solutions. Solid symbols refer to the samples with the hydrogen gas treatment, empty symbols - to the reference samples (Manilov *et al.* 2010a).Manilov A.I., et al*Journal of Alloys and Compounds* **492**, 466–472, 2010.



samples	H2 amount (mole/gram)	pSi-SWNT/other sample
pSi-SWNT	7.93E-05	
pSi	4.73E-05	1.7
charged-pSi	3.14E-05	2.5
charged-SWNT	1.99E-05	4
SWNT	1.29E-05	6.1

Figure 9.27. a)SEM image from an as-prepared pSi sample with electrophoretically assembled SWNTs., b) H2 evolution curves (based on the same weight) as a function of temperature from the temperature-programmed desorption experiments.,c) Hydrogen amount from TPD experiments calibrated with hydrogen dosing experiments. Table : Lists H2 amount and the relative ratios.

He Z., Wang S., Wang X., and Iqbal Z. (2013) Hydrogen storage in hierarchical nanoporous silicon–carbon nanotube architectures. *Int. J. Energy Res*. **37**, 754–760

9.8 ENERGETIC ANALYSIS OF PS NANOSTRUCTURES

Table 9.3 presents a comparative (with other hydrogen storage means) analysis of mass electrical energy that can be potentially extracted from the PS nanostructures through a fuel cell, taking into account atomic hydrogen concentrations (Lysenko *et al.* 2005). Only the maximal concentration values: 13, 34, and 66 mmol.g^{-1} found in meso- and nano-PS structures with enhanced porosities (>90%) are considered. Assuming quasi-identical levels for the fractal-like roughness of the specific surface for the PS nanostructures, the observed difference in hydrogen concentrations is explained by the similar difference in the mean dimension of the Si nanocrystallites constituting the PS samples. The higher the hydrogen concentrationis, the higher the corresponding mass electrical energy is. The PS nanostructures consisting of Si nanocrystallites with a mean dimension of about 2 nm ensure a maximal mass electrical energy value of about 2200 W-h.kg^{-1}, which is quite comparable with energy values ensured by thermolysis or hydrolysis of hydrides. It is important to note that the autonomy of about 100 h could be ensured by the use of 100 g of PS nanostructures as hydrogen reservoirs for a portable device needing 1 W of electrical power.

Table 9.3: Comparative energetic analysis of PS nanostructures for their application as hydrogen reservoirs in portable devices.

materials	Atomic hydroge content, mmol g	Theoretical m energy densit W*h kg^{-1}	Autonomy (h) of a dev consuming 1W and usi 100g of material storin hydrogen *(taking into account 50% efficiency low-temperature fuel c*	References
Meso-PS (90%, 10nm)	13	429	21.4	Lysenko *et al.* 200
Nano-PS (90%, 5nm)	34	1120	56.1	Lysenko *et al.* 200
Nano-PS powder (>95%, 2-	66	2176	108.8	Lysenko *et al.* 200
$MgH_2 \rightarrow Mg+H_2$	76	2505	125.2	MRS Bull. (2002)
$LaNi_5H_6 \rightarrow LaNi_5+3H_2$	14	461	23	MRS Bull. (2002)
$(NaBH_4+2H_2O) \rightarrow NaBO_2+4$	108	3560	178	MRS Bull. (2002)
$(LiBH_4+4H_2O) \rightarrow LiOH+H_3BO_3+4H_2$	85	2802	140.1	MRS Bull. (2002)
$NH_4BH_4 \rightarrow BN+4H_2$	244	8043	402.1	MRS Bull. (2002)
$NH_3BH_3 \rightarrow BN+3H_2$	195	6428	321.4	MRS Bull. (2002)
Li-ion batteries		150	15	Broussely *et.al.*199

Application of Si nanoparticles to hydrogen generation is promising among other methods when we use splitting water as hydrogen source. For example, efficient utilization of hydrogen generated during the reactions of nano-silicon/water in internal combustion engine has been investigated by Mehta *et al.* (2014). Similar to that of other nanoparticles for hydrogen generation, silicon nanoparticles are naturally expected to generate hydrogen more rapidly than bulk silicon. Fig.9.28 provides a schema showing multidisciplinary integrated approach from nanochemistry for generation of silicon nanoparticles (using the SiNP synthesis by laser pyrolisis) to the reaction with water under basic conditions allowing hydrogen generation on demand to the use of generated hydrogen in a fuel cell for portable power. Integration of nanosilicon with appropriate cartridge technologies could provide a “just add water” hydrogen-on-demand technology that would promote adoption of hydrogen fuel cellsin portable power applications. However, scalable and energy efficient processes for nanoparticle production must be implemented to expand the potential use of silicon-based H_2 generation beyond niche applications (Erogbogbo *et al.* 2013).

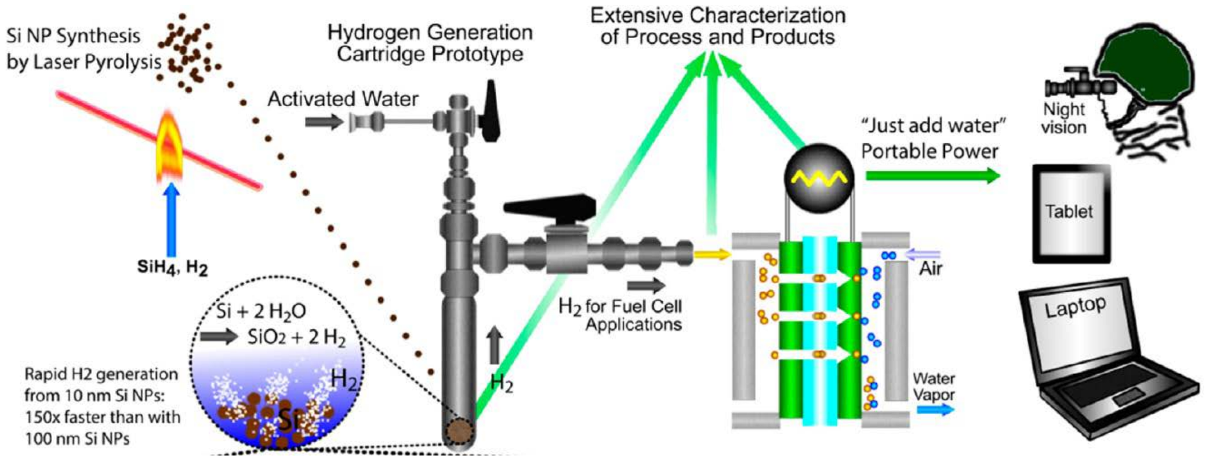


Fig.9.28. Schema showing CO_2 laser pyrolysis synthesis of silicon nanoparticles transferred to a custom stainless steel prototype cartridge used to generate hydrogen for fuel cell applications. Erogbogbo F., (2013) *NanoLett.* **13**, 451–456

Application of Si nanoparticles as catalyst for decomposition of ammonia borane NH_3BH_3 was considered by (Gangal *et al.* 2012). Ammonia borane has high gravimetric hydrogen storage capacity of 19.6 wt% and hydrogen release from it initiates at around 100 ^oC and as such is compatible to meet the present-day requirements of a PEM fuel cell. Variant of laboratory set up for demonstration of porous silicon hydrogen energy sores is presented in fig. 9.29.

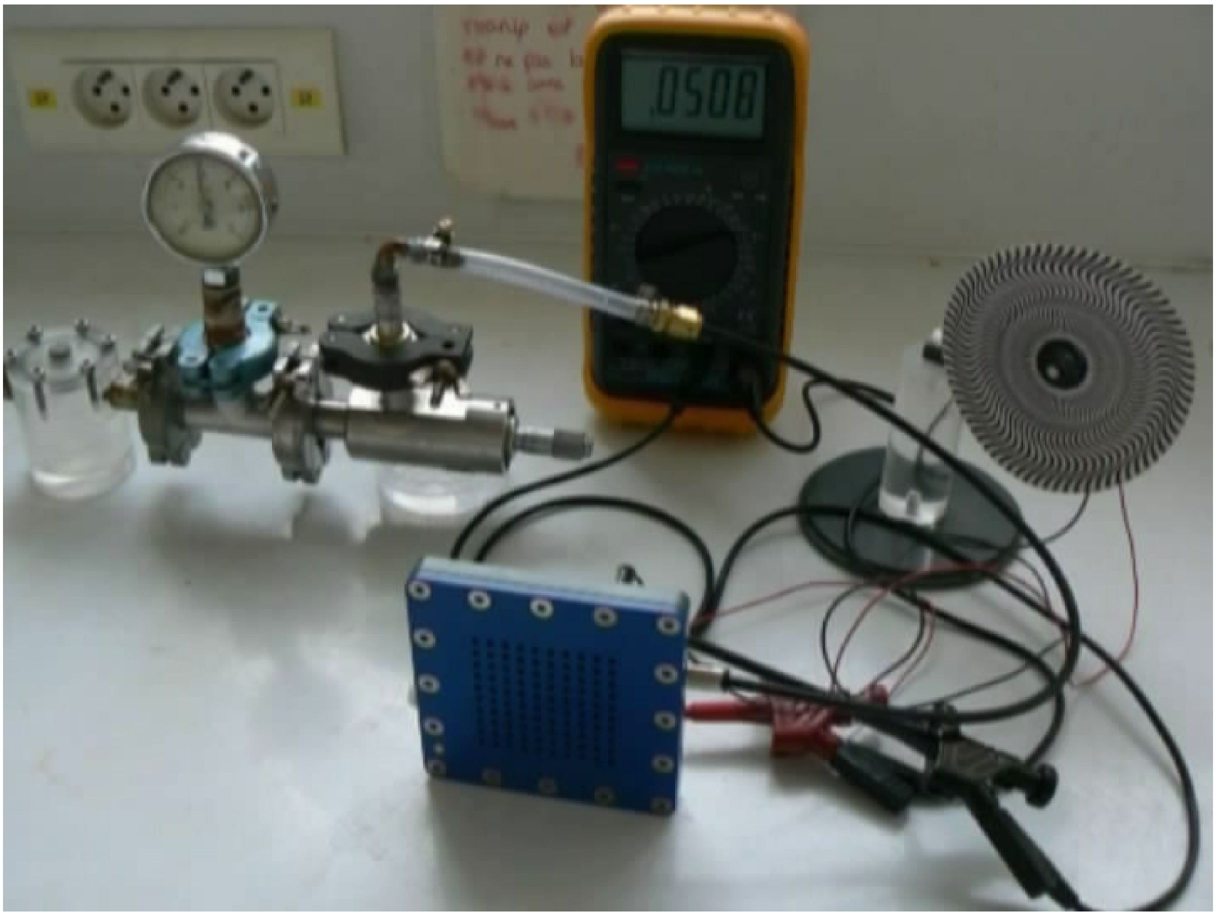


Fig.9.29 Laboratory set up for demonstration of hydrogen tntrgy sourse.

9.9 CONCLUSION

The concentration of hydrogen chemically bound to the porous silicon specific surface is found to be strongly correlated to the dimension and shape of the Si nanocrystallites constituting the PS nanostructures. Maximal values of hydrogen concentration for both meso- and nano-PS samples are mainly due to the fractal-like shape of the nanocrystallite surface. Chemical and thermal ways for hydrogen desorption can be applied. In the last case an additional hydrogen quantity can be generated via water splitting. Thus, the chemically induced production of the molecular hydrogen from the PS nanostructures appears to be much more efficient than that ensured by thermal annealing. However, the chemically stimulated hydrogen desorption is completely irreversible because of the complete transformation of the PS nanostructure into the hydrated dioxide. The PS nanostructures consisting of Si nanocrystallites with a mean dimension of about 2 nm ensure a maximal mass electrical energy value of about 2200 W-h kg^{-1}, which is quite comparable with energy values ensured by thermolysis or hydrolysis of hydrides.

References

Bahruji H., Bowker M., and Davies P.R. (2009) Photoactivated reaction of water with silicon nanoparticles. *Int. J. Hydrogen Energy* **34** (20) 8504–8510.

Barabash R.N., Alekseev S.A., Zaytsev V.N., and Barbier D. (2006) Porous silicon modified with vinylsilanes. *Ukrainian Chem. J.* **72** (10), 78–84.

Benilov A. I. , Gavrilchenko I. V., Benilova I. V. , Skryshevsky V. A. and Cabrera M. (2007) Influence of pH solution on photoluminescence of porous silicon. *Sensors and Actuators A.* **137**, 345-349.

Bisi O., Ossicini S., and Pavesi L. (2000) Porous silicon: a quantum sponge structure for silicon based optoelectronics. *Surf. Sci. Rep.* **38**, 1-126.

Bogdanovich B., and Schwickardi M. (1997) Ti-doped alkali metal aluminium hydrides as potential novel reversible hydrogen storage materials *J. Alloys Compounds.* **253-254**, 1-9.

Broussely M., Perton F., Biensan P., Bodet J.M., Labat J., Lecerf A., Delmas C., Rougier A., and Peres J.P. (1995) Li_xNiO_2, a promising cathode for rechargeable lithium batteries. *Journal of Power Sources.* **54**, 109-114

Broom D. P. (2011) *Hydrogen Storage Materials, Green Energy and Technology*. Springer-Verlag, London.

Bunker C. E. and Smith, M. J. (2011) Nanoparticles for hydrogen generation. *J. Mater. Chem.*, **21** (33), 12173−12180.

Cerofolini G.F., Balboni R., Bisero D., Corni F., Frabboni S., Ottaviani G., Tonini R., Brusa R. S., Zecca A., Ceschini M., Giebel G., and Pavesi L. (1995) Hydrogen precipitation in highly oversaturated single-crystalline silicon. *Phys. Status Solidi* A. **150**, 539–586.

Chazalviel J.-N. and Ozanam F. (1997) Surface modification of porous silicon.In: Canham L. (ed.) *Properties of porous silicon*, INSPEC, The Institution of Electrical Engineers, London, United Kingdom.

Collins R. T., Tischler M. A., and Stathis J. H. (1992) Photoinduced hydrogen loss from porous silicon. *Appl. Phys. Lett.* **61**(14), 1649-1651.

Dai F., Zai J., Yi R., Gordin M.L., Sohn H., Chen S., and Wang D. (2014) Bottom-up synthesis of high surface area mesoporous crystalline silicon and evaluation of its hydrogen evolution performance, *Nature Communications*, **5**, art. 3605, 1-11.

Dalebrook A.F., Gan W., Grasemann M., Moret S. and G.Laurenczy (2013) Hydrogen storage: beyond conventional methods. *Chem. Commun.* **49**, 8735-8751.

Darwiche S., Nikravech M., Morvan D., Amouroux J., and Ballutaud D. (2007) Effects of hydrogen plasma on passivation and generation of defects in multicrystalline silicon. *Sol. Energy Mater. Sol. Cells*. **91**, 195-200.

Dzhafarov T.D, Aydin Y.S., and Oruc L.C. (2008) Porous silicon based gas sensors and miniature hydrogen cells. *Jpn J Appl Phys*.47, 8204-8207.

Erogbogbo F., Lin T., Tucciarone P. M., LaJoie K. M., Lai L., Patki G. D., Prasad P. N., and Swihart M.T. (2013) On-demand hydrogen generation using nanosilicon: splitting water without light, heat, or electricity. *NanoLett.* **13**, 451–456.

Estreicher S.K., Hastings J.L., and Fedders P.A. (1999) Hydrogen-defect interactions in Si. *Mater. Sci. Eng.* B **58,** 31-35.

Fleming, I. (2001) *Science of Synthesis: Houben-Weyl Methods of Molecular Transformations, Vol. 4: Compounds of Group 15 (As, Sb, Bi) and Silicon Compounds.* Georg Thieme Verlag, Stuttgart.

Föll H., Christophersen M., Carstensen J., and Hasse G. (2002) Formation and application of porous silicon. *Mater. Sci. Eng.* B **39**, 93–141.

Gangal A. C., Kale P., Edla R., Manna J., and Sharma P. (2012) Study of kinetics and thermal decomposition of ammonia borane in presence of silicon nanoparticles. *Int. J. Hydrogen Energy* **37**, 6741–6748.

Gautier G.(2014). Porous silicon and micro-fuel cells. In Canham L.(ed) *Handbook of Porous Silicon*, Springer, 957-964.

Glass J.A.,. Wovchko E.A, and Yates J.T. (1996) Reaction of atomic hydrogen with hydrogenated porous silicon – detection of precursor to silane formation. *Surf. Sci.* **348**, 325-334.

Goller B., Kovalev D. and Sreseli O. (2011) Nanosilicon in water as a source of hydrogen: size and pH matter, *Nanotechnology*, **22**, art.305402, 1-4.

Gupta P., Colvin V.L., and George S.M. (1988) Hydrogen desorption kinetics from monohydride and dihydride species on silicon surfaces. *Phys. Rev. B.* **37**(14) 8234–8243.

Gupta P., Dillon A. C., Bracker A. S., and George S. M. (1991) FTIR studies of H_2O and D_2O decomposition on porous silicon surfaces. *Surf. Sci.* **245**(3), 360-372.

He Z., Wang S., Wang X., and Iqbal Z. (2013) Hydrogen storage in hierarchical nanoporous silicon–carbon nanotube architectures. *Int. J. Energy Res.* **37**, 754–760,

Herino R., Bomchil G., Barla K., Bertrand C., and Ginoux J. L. (1987) Porosity and pore size distributions of porous silicon layers. *J.Electrochem. Soc.* **134**(8), 1994-2000.

Kale P., Gangal A. C., Edla R., and Sharma P. (2012). Investigation of hydrogen storage behavior of silicon nanoparticles. *Int.J.Hydrogen Energy* **37** (4), 3741 -3747.

Kitajima M., Ishioka K., Tateishi S., Nakanoya K., Fukata N., Murakami K., Fujimura S., Hishita S., Komatsu M., and Haneda H. (1999) Effects of crystal disorder on the molecular hydrogen formation in silicon. *Mater. Sci. Eng. B.* **58**(1-2), 13-16.

Koropecki R.R., Arce R.D., and Schmidt J.A. (2004) Infrared studies combined with hydrogen effusion experiments on nanostructured porous silicon. *J. Non-Cryst. Solids.* **338-340**, 159-162.

Krishna R., Titus E., Salimian M., Okhay O.,Rajendran S., Rajkumar A., Sousa J. M. G., Ferreira A. L. C., Gil J.C. , and Gracio J. (2012) Hydrogen storage for energy application. In: Liu J. (ed) *Hydrogen storage*. InTech, 243-266.

Kudo A. and Miseki Y. (2009) Heterogeneous photocatalyst materials for water splitting. *Chem. Soc. Rev.* **38**, 253–278.

Langmi H.W., Walton A., Al-Mamouri M.M., Johnson S.R., Book D., Speight J.D., Edwards P.P., Gameson I., Anderson P.A., and Harris I.R. (2003) Hydrogen adsorption in zeolites A, X, Y and RHO. *J.Alloys Compaunds.* **356-357**, 710-715.

Lee S.J., Chang-Chien A., Cha S.W., O'Hayre R., Park Y.I., Saito Y., and Prinz F.B. (2002) Design and fabrication of a micro fuel cell array with "flip-flop" interconnection. *J. Power Sources.***112** (2), 410-418.

Lee S. M., Park K. S., Choi Y.C., Park Y.S., Bok J. M., Bae D. J., Nahm K.S., Choi Y. G., Yu S. C., Kim N.-G., Frauenheim Th., and Lee Y. H. (2000) Hydrogen adsorption and storage in carbon nanotubes. *Synth. Met.* **113** (3), 209-216.

Litvinenko S., Alekseev S.,.Lysenko V, Venturello A., Geobaldo F., Gulina L., Kuznetsov G., Tolstoy V., Skryshevsky V., Garrone E., and Barbier D. (2010) Hydrogen production from nano-porous Si powder formed by stain etching. *Int. J. Hydrogen Energy*. **35**, 6773-6778.

Lysenko V., Vitiello J., Remaki B., Barbier D., and Skryshevsky V. (2004) Nanoscale morphology dependent hydrogen coverage of meso-porous silicon. *Appl. Surf. Sci*. **230**, 425 – 430

Lysenko V., Bidault F., Alekseev S., Turpin C., Geobaldo F., Rivolo P., Garrone E., Zaitsev V., and Barbier D. (2005) Study of porous silicon nanostructures as hydrogen reservoirs. *J.Phys. Chem. B.* **109**, 19711-19718.

Ma X. F., and Zachariah M. R. (2010) Size-resolved kinetics of Zn nanocrystal hydrolysis for hydrogen generation. *Int. J. Hydrogen Energy*, **35** (6), 2268−2277.

Manilov A.I., Alekseev S.A., Skryshevsky V.A., Litvinenko S.V., Kuznetsov G.V., and Lysenko V. (2010a). Influence of palladium particles impregnation on hydrogen behavior inmeso-porous silicon. *Journal of Alloys and Compounds* **492**, 466–472.

Manilov A.I., Litvinenko S.V., Alekseev S.A., Kuznetsov G..V, and Skryshevsky V.A. (2010b) Use of powders and composites based on porous and crystalline silicon in the hydrogen power industry, *Ukr. J. Phys.* **55** (8), 928-935.

Manilov A.I. and Skryshevsky V.A. (2013) Hydrogen in porous silicon — a review. Materials Science and Engineering B. **178**, 942– 955.

Martin P., Fernandez J. F., and Sanchez C. R. (2000) TDS applied to investigate the hydrogen and silane desorption from porous silicon. *Phys. Status Solidi A.* **182**(1), 255-260.

Mathews N. R., Sebastian P. J., Mathew X. and Agarwal V. (2003) Photoelectrochemical characterization of porous Si. *Int. J. Hydrogen Energy* **28** (6), 629–632.

Mehta R. N., Chakraborty M., and Parikh P.A. (2014) Impact of hydrogen generated by splitting water with nano-silicon and nano-aluminum on diesel engine performance, *Int. J. Hydrogen Energy* **39**, 8098- 8105.

Mex L., Ponath N., and Müller J. (2001) Miniaturized fuel cells based on microsystem technologies. *Fuel Cells Bull.***4**, 9-12.

Meyers J. P. and Maynard H. L. (2002) Design considerations for miniaturized PEM fuel cells. *J. Power Sources.***109** (1), 76-88.

MRS Bull. **(**2002) Special issue on Hydrogen Storage. **27,** 675–716.

Nijkamp M.G., Raaymakers J.E.M.J., van Dillen A.J., and de Jong K.P. (2001) Hydrogen storage using physisorption – materials demands. *Appl. Phys. A.* **72**(5), 619-623.

Nützenadel Ch., Züttel A., Emmenegger Ch., Sudan P., and Schlapbach L. (2000) Electrochemical storage of hydrogen in carbon single wall nanotubes. In: Tománek D., and Enbody R. J. (eds.) *Science and Application of Nanotubes*. Kluwer Academic Publishing/Plenum Press, Dordrecht, 205-214.

Nychyporuk T., Lysenko V., and Barbier D. (2005) Fractal nature of porous silicon nanocrystallites. *Phys. Rev. B.* **71**(11), art.115402.

Ogata Y., Niki H., Sakka T., and Iwasaki M. (1995) Hydrogen in porous silicon: vibrational analysis of SiH_x species. *J. Electrochem. Soc.* **142**, 195-201.

Oh I., Kye J., and Hwang S. (2012) Enhanced photoelectrochemical hydrogen production from silicon nanowire array photocathode. *Nano Letters* **12**, 298–302 .

Parkhutik V. (1999) Porous silicon – mechanisms of growth and applications. *Solid State Electron*. **43**, 1121-1141.

Polisski S., Lapkin A., Goller B., and Kovalev D. (2008) Hybrid metal/silicon nanocomposite systems and their catalytic activity. In: Proc. 6th Int. Conf. “Porous Semiconductors Science and Technology”, PSST 2008, Mallorca, Spain, 44.

Pritchard R.E., Ashwin M.J., Newman R.C., and Tucker J.H. (1999) H_2 molecules in crystalline silicon. *Mater. Sci. Eng.* B **58** 1-5.

Rahimi F., and Irajizad, A. (2006) Effective factors on Pd growth on porous silicon by electroless-plating: Response to hydrogen. *Sensors and Actuators B.* **115** (1) 164–169.

Rather S.U., Zacharia R., Hwang S.W., Naik M.U., and Nahm K.S. (2007) Hyperstoichiometric hydrogen storage in monodispersed palladium nanoparticles. *Chem. Phys. Lett.*, **438** (1-3), 78–84.

Rivolo P., Geobaldo F., Rocchia M., Amato G., Rossi A. M., and Garrone E. (2003) Joint FTIR and TPD study of hydrogen desorption from p^+-type porous silicon. *Phys. Status Solidi A*, **197**(1), 217-221.

Rosi N. L., Eckert J., Eddaoudi M., Vodak D. T., Kim J., O'Keeffe M., and Yaghi O. M. (2003) Hydrogen storage in microporous metal-organic frameworks. Science. **300**, 1127-1129.

Shanks H., Fang C.J., Ley L., Cardona M., Demond F.J., and Kalbitzer S. (1980) Infrared spectrum and structure of hydrogenated amorphous silicon. *Phys. Stat. Sol. B.* **100** (1) 43-56.

Song X., and Wu J. (2012) Modeling of hydrogen weight storage capacity in solid porous silicon. *Advanced Materials Research.* **415-417**, 2322-2328.

Sriraman S., Valipa M.S., Aydil E.S., and Maroudas D. (2006) Hydrogen-induced crystallization of amorphous silicon thin films. I. Simulation and analysis of film postgrowth treatment with H_2 plasmas. *J. Appl. Phys.* **100,** art. 053514.

Strobel R., Jörissen L., Schliermann T., Trapp V., Schütz W., Bohmhammel K., Wolf G., and Garche J. (1999) Hydrogen adsorption on carbon materials. *J. Power Sources.* **84**(2), 221-224.

Switendick A.C. (1978) The change in electronic properties on hydrogen alloying and hydride formation. In: Alefeld G., and Völkl J. *Hydrogen in Metals I*. *Topics in Applied Physics*. **28**, Springer, Berlin Heidelberg, 101-129.

Tichapondwa S.M., Focke W.W., Del Fabbro O., Mkhize S., and Muller E. (2011) Suppressing H_2 evolution by silicon powder dispersions, *J. Energetic Mater*. **29**, 326-343.

Timoshenko V. Yu., Osminkina L. A., Efimova A. I., Golovan L. A., Kashkarov P. K., Kovalev D., Kunzner N., Gross E., Diener J., and Koch F. (2003) Anisotropy of optical absorption in birefringent porous silicon. *Phys. Rev. B.* **67** (11), art.113405.

Tolstoy V.P., Chernyshova I.V., and Skryshevsky V.A. (2003) *Handbook of Infrared Spectroscopy of Ultrathin Films.* Wiley, New York.

Tutov E.A., Pavlenko M.N., Protasova I.V., and Kashkarov V.M. (2002) The interaction of porous silicon with water: A chemographic effect. *Tech. Phys. Lett.* **28** (9), 729-731.

Tuyen L.T.T., Tam N.T.T., Quang N.H., Nghia N.X., Khang D.D., and Khoi P.H. (2001) Study on hydrogen reactivity with surface chemical species of nanocrystalline porous silicon. *Mater.Sci. Eng. C.* **15**(1-2), 133–135.

Tzimas E., Filiou C., Peteves S.D. and Veyret J.-B. (2003) *Hydrogen storage:state-of-the-art and future, perspective. European Commission. DG JRC. Institute for Energy*, Petten. The Netherlands.

Weitkamp J., Fritz M., and Ernst S. (1995) Zeolites as media for hydrogen storage. *Int. J. Hydrogen Energy.* **20** (12), 967-1002.

Westlake D. J. (1983) Hydrides of intermetallic compounds: a review of stabilities, stoichiometries and preferred hydrogen sites. *J. Less-Common Metals.* **91**, 1-20.

Wieder H., Cardona M., and Guarnieri C.R. (1979) Vibrational spectrum of hydrogenated amorphous SiC films. *Phys. Stat. Sol. B.* **92**(1), 99-112.

Yamazaki Y. (2004) Application of MEMS technology to micro fuel cells. *Electrochim. Acta* **50** (2-3), 663-666.

Zhan C. Y., Chu P. K., Ren D., Xin Y. C., Huo K. F., Zou Y., and Huang N. K. (2011) Release of hydrogen during transformation from porous silicon to silicon oxide at normal temperature. *Int. J. Hydrogen Energy*, **36** (7), 4513−4517.

Zhen-Yi N.I., Xiao-Dong P.I., and De-Ren Y. (2012) Can hydrogen be incorporated inside silicon nanocrystals? *Chin. Phys. Lett.* **29**(7), 077801.

Züttel A., Wenger P., Rentsch S., Sudan P., Mauron Ph., and Emmenegger Ch. (2003) $LiBH_4$ a new hydrogen storage material. *J. Power Sources*. **118** (1-2), 1 -7.

Chapter 10

Magnetic quantum dots

A variety of experiments are aimed at exploring the classical and quantum mechanical interactions in mesocopic magnetic structures. Researche exploits advances in miniaturization and highly sensitive magnetometry to probe a variety of new physical systems, where two distinct experimental efforts investigate phenomena scaling the classical to the quantum regimes. The magnetic properties of arrays of STM-fabricated ferromagnetic particles are studied as a function of their dimension using a new high sensitivity 2DHG Hall magnetometer.

In contrast to the expectations of classical magnetism, the results reveal surprising magnetic instability of the small structures. At even smaller length scales and lower temperatures, magnetic particles have been proposed as systems where quantum mechanics may produce measurable manifestations on the macroscopic level. Low temperature measurements of the magnetic noise and susceptibility in a variety of artificially-engineered proteins are performed using a combination of superconducting integrated circuits to form a sensitive miniature dc SQUID-based microsusceptometer. The results are compared to theoretical predictions for macroscopic quantum coherence in small antiferromagnetic systems.

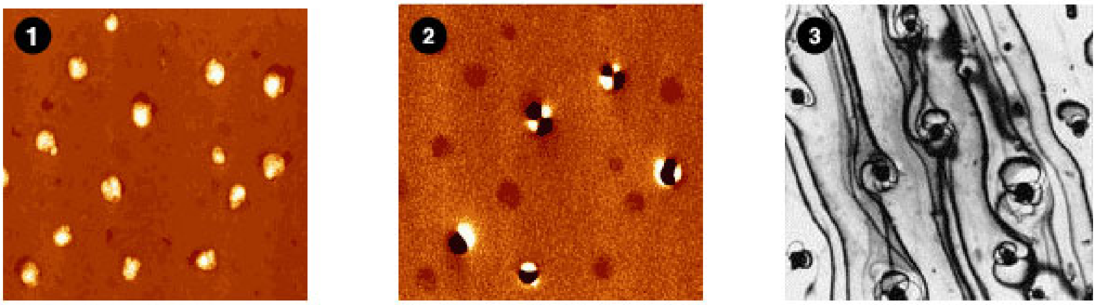


Fig. 1 Submicron ferromagnets are fabricated in GaAs through manganese ion implantation and subsequent rapid thermal annealing. the leftmost image (1) seen below is a room-temperature atomic force microscope image of these precipitates, which are GaMn rich and crystalline. The corresponding magnetic-force microscope image below center (2) demonstrates that many of the precipitates are ferromagnetic.The rightmost image (3) in the figure above is a plan-view transmission electron microscopy (TEM) image of these precipitates.

These measurements reveal that the precipitates form at the sample surface and that average diameters can be varied from 100 nm to 400 nm by changing implantation doses and annealing conditions. SQUID measurements show that after annealing the implanted semiconductor films are ferromagnetic well above 300 K, with coercive fields ranging from 1000-5000 Gauss.

This work was originally started several years ago in the scientific group at IBM in order to produce nanometer-scale ferromagnets for ultra-low temperature studies of magnetic quantum tunneling. The present STM research is an active collaborative effort with Professor Stephan von Molnar at Florida State University and Dr. Andrew Kent. The growth of the magnets is shown in the following picture. A metallorganic precursor, in this case iron pentacarbonyl, is introduced in the UHV chamber. When a tunneling current is produced between the tip of the STM and the substrate, the metallorganic gas is dissociated and iron deposited locally. The iron may be deposited on either the substrate or the tip depending on the sign of the voltage bias. The aspect ratio of the particles may be controlled by maintaining the tunneling current as the tip is withdrawn from the substrate. It is possible to grow magnets on the substrate with aspect ratio almost as high as 10 and on the tip as high as 100.

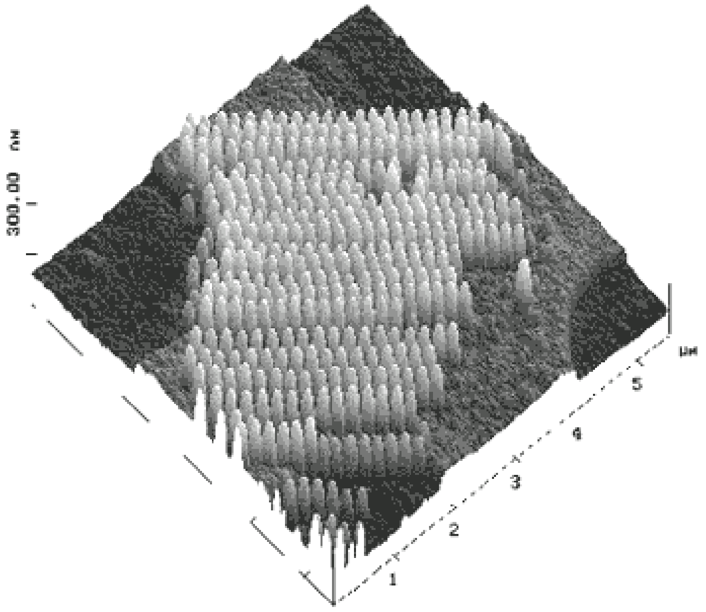


Fig. 2. Atomic Force Microscope (AFM) image of an array of single domain Fe magnets grown by STM deposition on top of a 2DEG Hall magnetometer. The magnets are approximately 40 nm in diameter.

This work was originally started several years ago in our group at IBM in order to produce nanometer-scale ferromagnets for ultra-low temperature studies of magnetic quantum tunneling. The present STM research is an active collaborative effort with Professor Stephan von Molnar at Florida State University and Dr. Andrew Kent. The growth of the magnets is shown in the following picture. A metallorganic precursor, in this case iron pentacarbonyl, is introduced in the UHV chamber. When a tunneling current is produced between the tip of the STM and the substrate, the metallorganic gas is dissociated and iron deposited locally.

The iron may be deposited on either the substrate or the tip depending on the sign of the voltage bias. The aspect ratio of the particles may be controlled by maintaining the tunneling current as the tip is withdrawn from the substrate. It is possible to grow magnets on the substrate with aspect ratio almost as high as 10 and on the tip as high as 100.

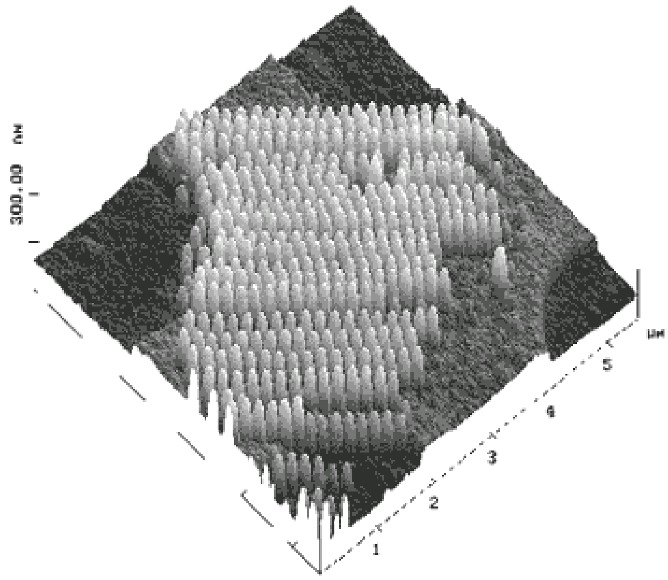


Fig. 3. Atomic Force Microscope (AFM) image of an array of single domain Fe magnets grown by STM deposition on top of a 2DEG Hall magnetometer. The magnets are approximately 40 nm in diameter.

Semiconductor nanoparticles that are smaller than the spatial extension of electrons and holes in the bulk semiconductor show quantum-size effects and are called QDs. These effects typically occur in the 1–10-nm size regime. Because of the very small size of the box available, the (kinetic) energy of the electrons and holes increases with decreasing particle size, causing an increase in the energy difference between the conduction and the valence band (i.e., an increase in the bandgap). This results in strongly size-dependent absorption and emission spectra for QDs, both shifting to higher energies for decreasing particle size. The emission originates from the lowest exciton (electron–hole pair) state and is narrow.

The absorption spectrum is broad owing to the energetic width of the conduction and valence bands, and the QD fluorescence can therefore be excited over a broad spectral range. In this semiconductor heterostructure, lattice mismatch strain between the CdSe quantum dots (QDs) and the ZnCdSe quantum well (QW) produces a three dimensional confinement potential in the QW itself, effectively confining carriers to small regions in the sample termed strain induced quantum well dots, or SIQWD's.

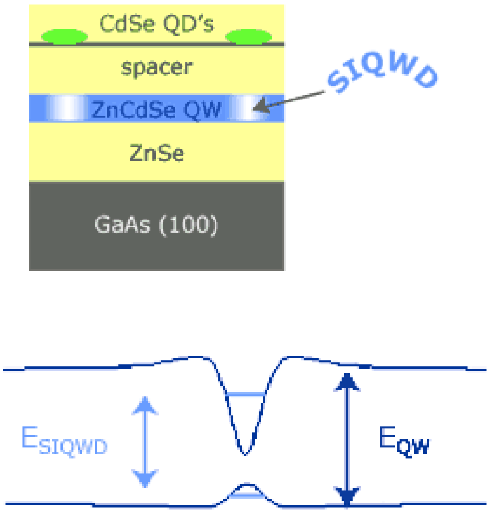


Fig. 4. Schematic picture of the changes in the conduction and valence bands of the ZnCdSe QW due to the strain fields of the CdSe QDs.

Shown below are a photoluminescence (PL) spectra and carrier lifetimes extracted from time-resolved PL measurements. The PL spectra shows that the confinement of carriers is nearly zero-dimensional in both the ordinary CdSe QDs as well as in regions energetically attributed to dots produced in the ZnCdSe QW itself. The intermediate carrier lifetimes in the SIQWD region reflect both shallower confinement relative to CdSe dots, as well as thermally activated detrapping of carriers into the QW host.

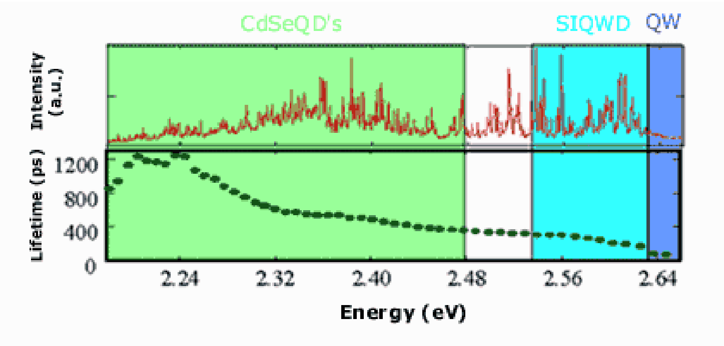


Fig. 5. PL spectra and carrier lifetimes extracted from time-resolved PL measurements.

In the two structures below, carriers confined to quantum dots interact with a local magnetic environment provided by paramagnetic Mn^{2+} ions (spin 5/2). In the left, a layer of MnSe is grown sufficiently close to the CdSe QDs so that there is significant overlap of the carrier’s wavefunction with that of the Mn ions. The right structure is a variant of the SIQWD samples in which the dot-like regions produced by strain are present in a Digital Magnetic Heterostructure (DMH - a quantum well with incorporated Mn ions).

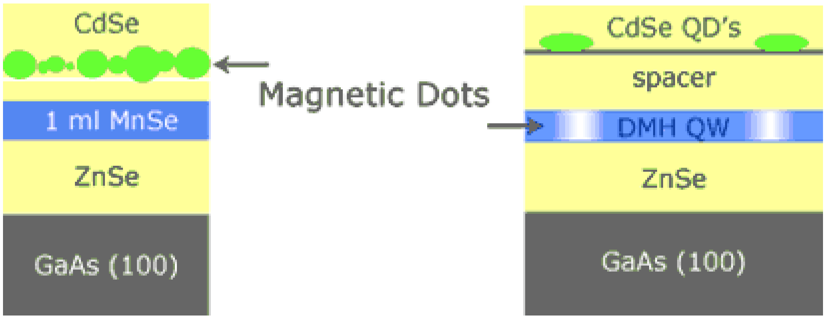


Fig. 6. Scheme of interaction between carriers and paramagnetic Mn^{2+} ions.

The plot below traces the spin splitting of one of the quantum dot photoluminescence peaks as a magnetic field is applied. The single peak at zero field is replaced with two peaks of opposite circular polarizations. The enhanced Zeeman splitting of the magnetic SIQWD sample is attributed to enhanced local fields resulting from contributions of nearby Mn ions.

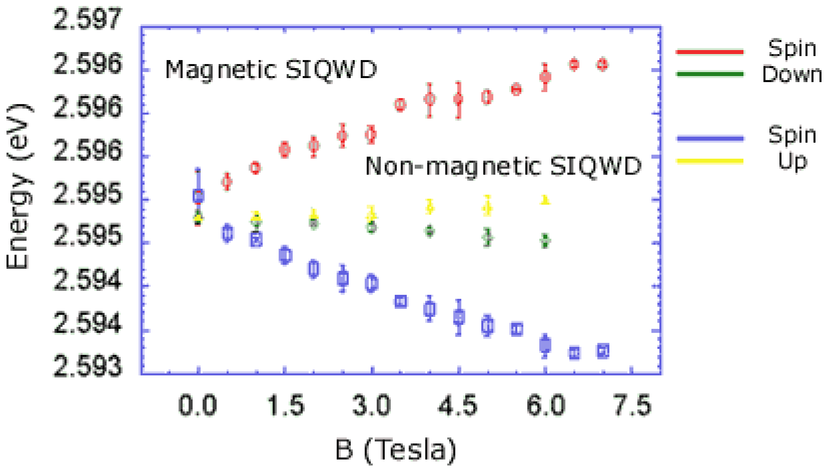


Fig. 7. Spin-splitting phenomenon in magnetic field.

Research into semiconductor quantum dots (QDs) is driven by their promise in exploring carrier behavior in the mesoscopic regime between bulk and molecular systems, and a variety of technological applications which exploit their size-tunable optical properties.

Recent interest in manipulating semiconductor spins for applications ranging from spin-polarized magneto-electronics to quantum computation is based on the ability to control and maintain spin coherence over practical length and time scales. To this end, QDs have been suggested as potential elements for such devices due to control over the structural and electronic environment of localized carriers.

Time-resolved Faraday rotation is used to probe spin dynamics in chemically synthesized CdSe quantum dots (QDs) at temperatures from 6 to 282 Kelvin. The pump pulse used in the experiment excites spins into an ensemble of ~ 10^{10} QDs. The spin precession in a transverse magnetic field indicates that the measured relaxation lifetime of the spin polarization is dominated by inhomogeneous dephasing, ranging from ~3 nanoseconds at zero field to less than 100 picoseconds at 4 Tesla.

Single crystal CdSe QDs ranging from 22 - 80 Angstroms in diameter are synthesized by a solution-phase pyrolytic reaction of organometallic precursors. By controlling the reaction temperature and concentrations of reagents, size distributions from 5-10% can be achieved.

Atomic resolution Transmission Electron Micrographs indicate that the QDs are highly crystalline (with the wurtzite lattice structure), and nearly spherical in shape. Bright spots in the image correspond to spaces between individual atoms in the QD.

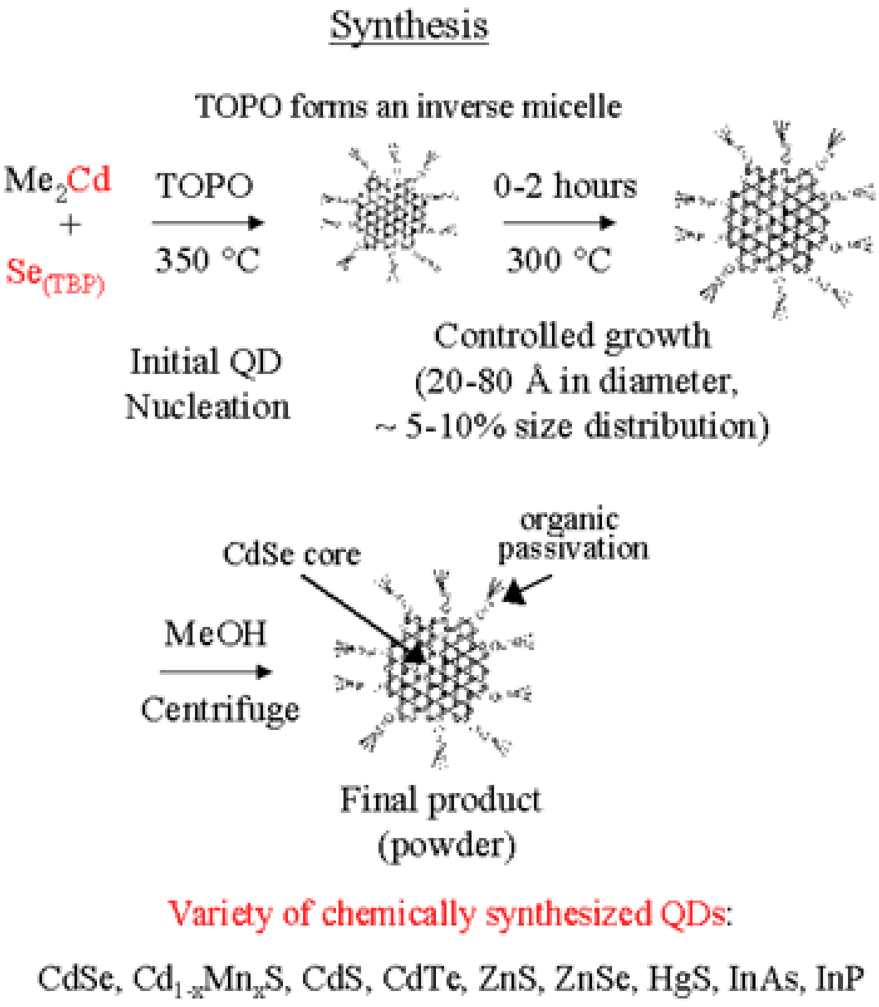


Fig. 8. Synthetic scheme for several types of quantum dots.

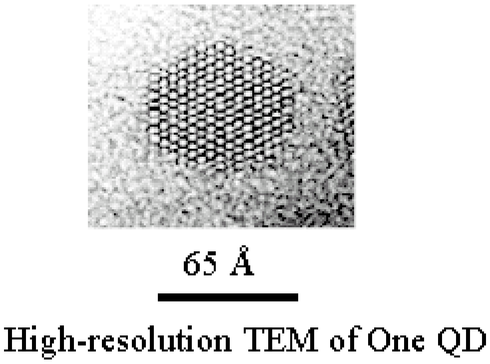


Fig. 9. Image of magnetic quantim dot obtained by TEM.

The figure below compares the spin lifetimes at temperatures of 6K (black) and 282K (red). The spin lifetimes drop from ~ 3 ns at 6K to 0.7 ns at 282K, although little change in the oscillation frequency is observed between these two temperature extremes. The robustness of spin coherence at room temperature is not predicted by conventional theories of spin-scattering in solid state systems.

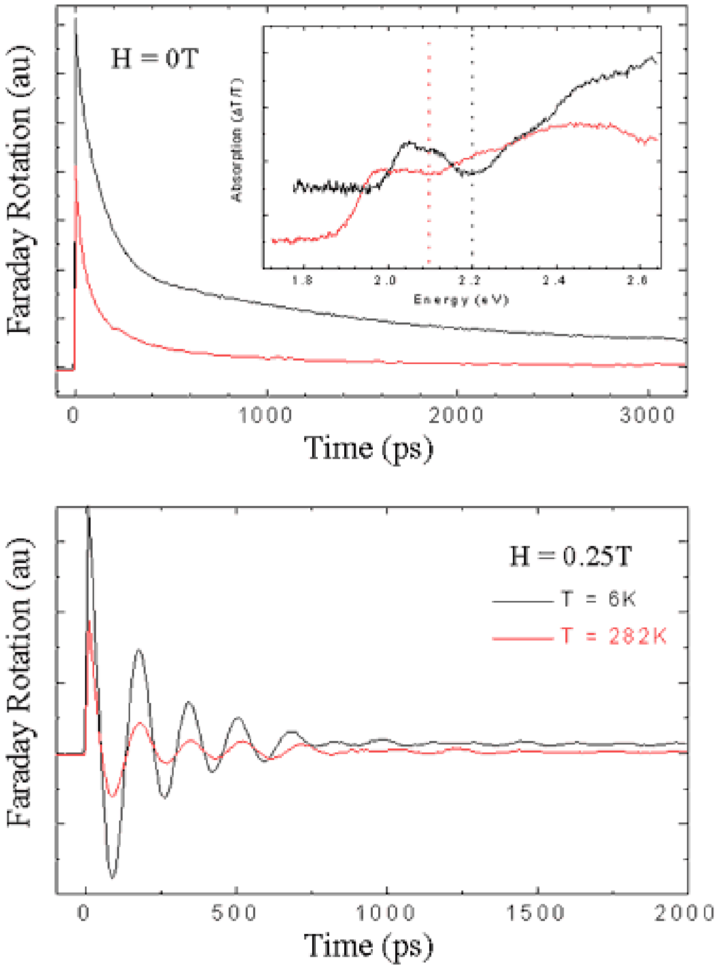


Fig. 10. Spin scattering dependence.

The figure below shows the evolution of the Faraday Rotation as the magnetic field is increased from H = 0-3 T for 80 Å QDs. The oscillation frequency increases with H as expected from classical Larmor precession. Surprisingly, the spin lifetime also shows a strong field dependence. Spin precession becomes unobservable for t > 200ps at H = 3T, but continues for t > 1.7 ns at low fields. This field dependence suggests that the decay of the measured spin coherence is dominated by inhomogeneous dephasing within the ensemble of QDs.

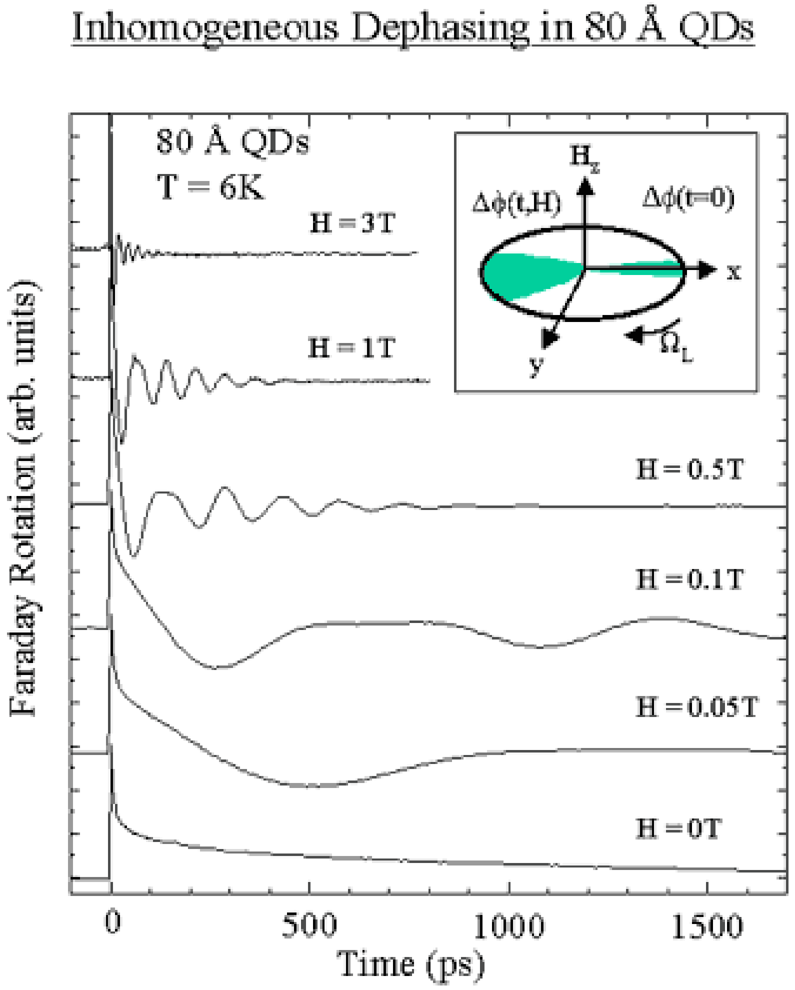

Inhomogeneous Dephasing in 80 Å QDs

Fig. 11. Spin coherence of electrons in quantum dots for different magnetic field.

The inset of the graph illustrates how inhomogeneous dephasing limits the measured spin lifetimes. Spins are initially aligned along the z-axis at t=0. Each spin is characterized by a g-factor, which determines the precession frequency. If these g-factors vary between QDs, then at a given time t, QDs with higher g-factors will have precessed an angle which is slightly greater than QDs with lower g-factors. The spins therefore, "walk" out of phase at a rate which is proportional to the field, thus producing the field-dependent spin lifetimes.

Investigations into chemically synthesized CdSe QDs have revealed spin coherence times exceeding several nanoseconds at room temperature, limited by inhomogeneous broadening. However, the integration of these nanometer- scale building blocks into large-scale functional spintronic devices, while maintaining efficient and tunable spin coupling among QDs, has been extremely challenging.

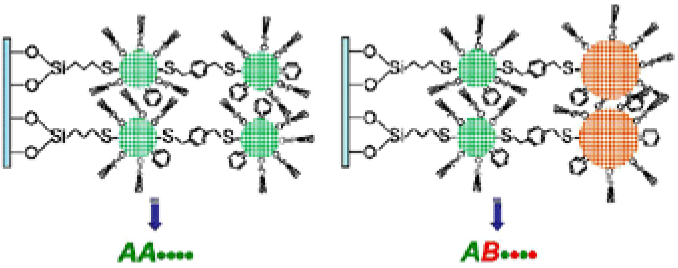


Fig. 12. Building blocks of spintronic devices.

We have implemented a controlled bottom-up layer-by-layer self-assembly scheme to construct artificial solids of semiconductor QDs bridged by organic molecules, as shown in the right figure.

The advantage of such a layer-by-layer assembly scheme is that each layer is independent of the others during the sample preparation. Thus, a variety of multilayer configurations and molecular bridges can be achieved by simply changing the type of QD solutions and/or the dithiol linking molecules used for each step. Specifically, the linking molecule used in our devices is 1,4-benzenedimethanethiol and two different sizes of CdSe QDs are employed (3.4 and 7.0 nm QDs, denoted as **A** and **B** below, respectively) with a large energy separation between their first exciton states (E_{gA} and E_{gB}):

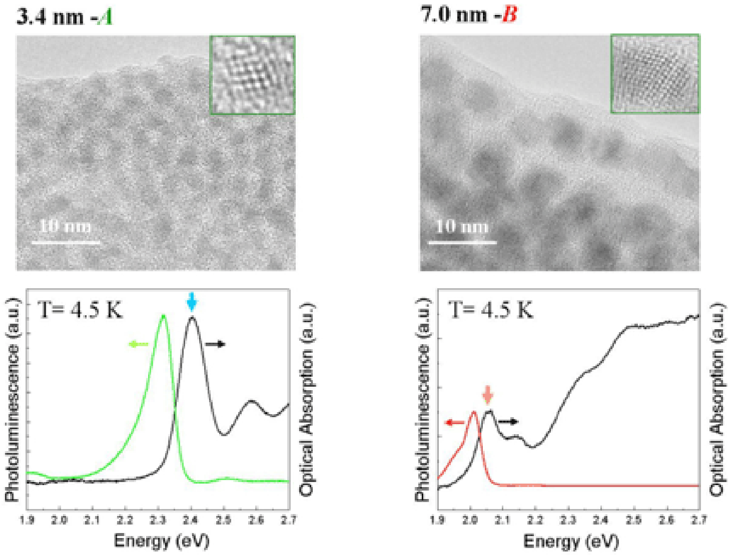


Fig. 13. STM images and photoluminescence spectra for spin-dependent states.

The right figure shows typical TRFR data from two control samples at B=0.8T for a four-layer sample **AAAA** and a two-layer sample **BB**. The TRFR signal from sample **BB** is observed over a broad probe energy range (E_{probe}=1.97-2.33 eV), and yields a constant precession frequency γ_B = 20.72 ± 0.21 GHz. Similarly, TRFR data recorded in sample **AAAA** exhibit comparable spin dynamic features over the detectable range E_{probe}=2.26-2.44 eV (limited at high energies by the tuning range of the probe laser) but with a different characteristic Larmor frequency γ_A = 23.39 ± 0.26 GHz.

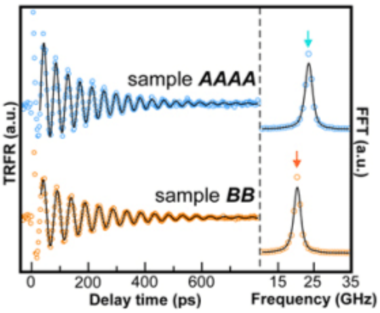


Fig. 14. Spin dynamic features detected in the samples with AAAA and BB layers.

To the left we plot typical TRFR traces of a mixed structure **ABAABA** at three different probe energies with pump energy tuned to the first exciton state of the 7.0nm QDs. Since the pump energy is less than the first exciton state of the smaller QDs, carriers are only expected to be excited in the larger QDs.

Remarkably, however, TRFR signal is observed at the higher probe energies of the smaller QDs. In addition, in contrast to the Figure above, the spin dynamics here exhibit a strong dependence on the probe energy (note the temporal beating pattern highlighted in the middle trace of E_{probe}=2.32eV).

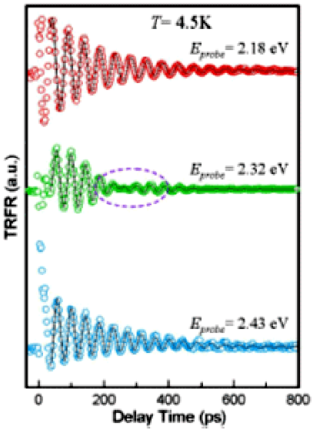


Fig. 15. Typical evolution of the TRFR signal of sample **ABAABA** as a function of probe energy.

The evolution of the TRFR signal of sample **ABAABA** as a function of probe energy can be more clearly seen in the figure below, which shows the normalized FFT of a series of measured TRFR data with probe energies scanned from 1.97 to 2.44 eV while holding the pump energy to the first exciton state of 7.0nm QDs.

Two Larmor frequencies observed in this hybrid sample, 20.82 ± 0.26 and 23.63 ± 0.26 GHz can be attributed to the coherent spin dynamics of carriers in the 7.0nm (B) and 3.4nm (A) QDs, respectively.

Specifically, the presence of the characteristic γ_A of 3.4nm QDs at higher probe energies suggests that spins are coherently transferred from larger to smaller QDs. Further more, probing at intermediate energies (for example, E_{probe}=2.32 eV), spins precessing at both Larmor frequencies can be observed, leading to a temporal beating. Line cuts at three different probe energies are also shown.

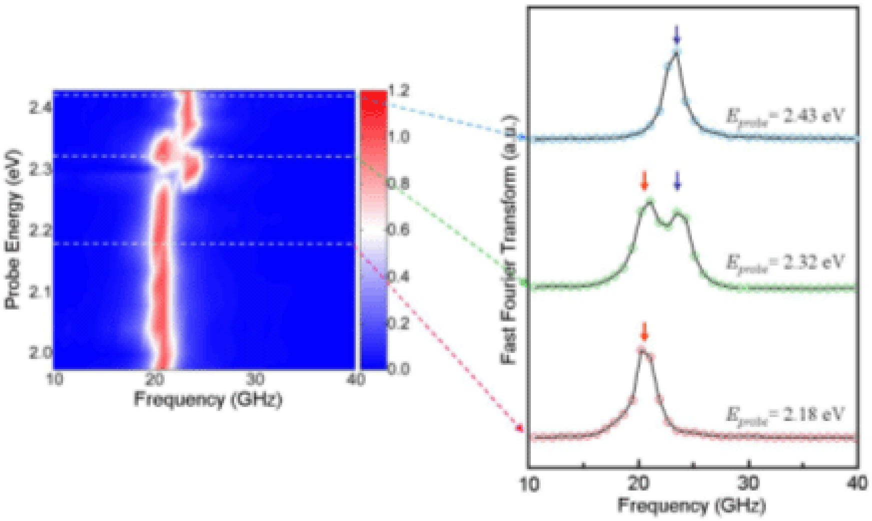


Fig. 16. Degree of spin accumulation measurements.

The left figure further verifies that spins excited in 7.0nm QDs are indeed transferred to 3.4nm QDs. In this experiment, we measure the degree of spin accumulation in 3.4nm QDs as a function of pump energy (E_{pump}) while holding the E_{probe} at 2.43 eV at zero field. The data show a clear onset feature near 2.04 eV, which is coincident with the first exciton state of 7.0nm QDs.

The spin migration from 7.0nm to 3.4nm QDs is therefore confirmed by the sudden growth of TRFR amplitude as pump energy is tuned across the absorption threshold of 7.0nm QDs.

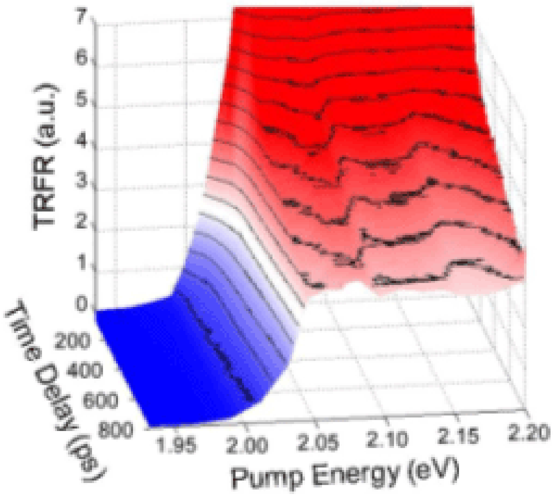


Fig. 17. Comparison of spin dynamics of 3.4nm QDs resulting from either direct spin injection (sample **AAAA**) or spin transfer from 7.0nm QDs (sample **ABAABA**) is shown on the right.

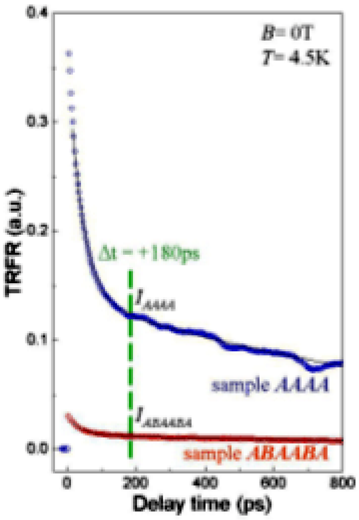


Fig. 18. The observation of time and phase dynamics.

Significantly, there is no observable time and phase delay for spin to transfer from 7.0nm QDs to 3.4 nm QDs within the time resolution of our apparatus (~300fs). We further estimate the efficiency of spin transfer by calculating the spin transfer percentage (STP) as following:

$$STP = \left[\frac{I_{ABAABA}}{I_{AAAA}}\right]_{\Delta t} \times \left[\frac{\gamma_{AAAA}}{\gamma_{ABAABA}}\right] \times \left[\frac{E_{gA}}{E_{gB}}\right] \times 100 ,$$

where I_{AAAA} and I_{ABAABA} are the amplitudes of normalized zero field TRFR data (recorded at the same laser excitation power) of sample **AAAA** and sample **ABAABA** at fixed delay Δt, and γ_{AAAA} and γ_{ABAABA} are the optical absorption of sample **AAAA** and sample **ABAABA** at their excitation energies. Therefore, the STP is ~11.8% at 4.5K.

This observed near-instantaneous spin coherence transfer persists up to room temperature. The figure below shows the evolution of Larmor precession at 200K measured at three different probe energies while spins are injected into 7.0nm QDs in sample **ABAABA**. The spin transfer kinetics exhibit a very similar dependence on E_{probe} as seen at low temperatures. More significantly, the STP slowly increases as the temperature is raised from 4.5K to 25 K, with a sharper increase thereafter. As the temperature is raised above 80K, the STP approximately doubles and then saturates at higher temperatures.

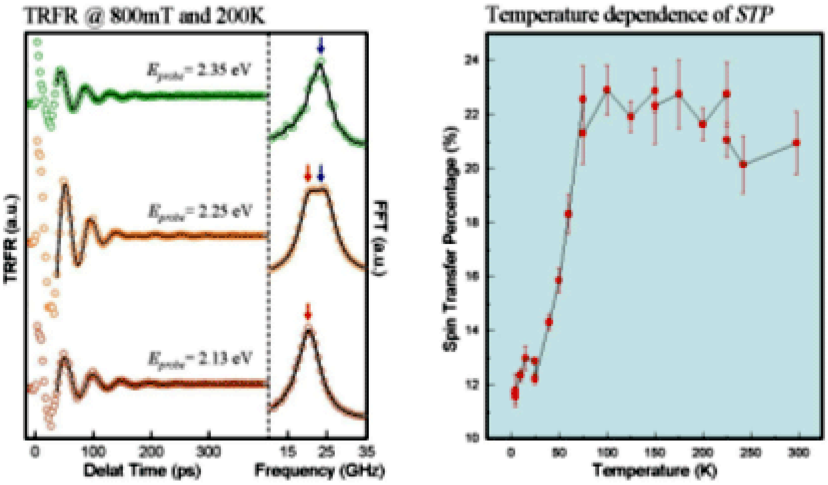


Fig. 19. The spin transfer kinetics.

These spin transfer measurements reveal that small conjugated molecules can serve not only as physical links for hierarchical assembly but also as natural media for shuttling quantum information between QDs.

Chemically synthesized semiconductor quantum dots are promising structures because of their ease of manufacture and highly size-tunable physical properties. To date, colloidal quantum dots have been used in a wide variety of applications including lasers, light-emitting devices, and fluorescent labels. Experiments have also been able to optically excite spin-polarized carriers in such colloidal nanocrystals, leading to measurement of spin dynamics in highly three-dimensionally confined quantum structures. Chemical linking of the quantum dots has led to coherent transport of electron spins between different nanocrystals. With the ability to inject spin information and transfer it between quantum dots, the behavior of these spins in different charge environments becomes important.

Recent experiments have succeeded in controlling the charging of the lowest electron state in colloidal quantum dots in an electrochemical cell. This charging can be detected by a bleaching of absorptive transitions when the lowest quantum state is filled with electrons. Since CdSe quantum dots absorb in the visible range of the spectrum, a film of nanocrystals exhibits a visible color change when bleached in a charged electrochemical cell.

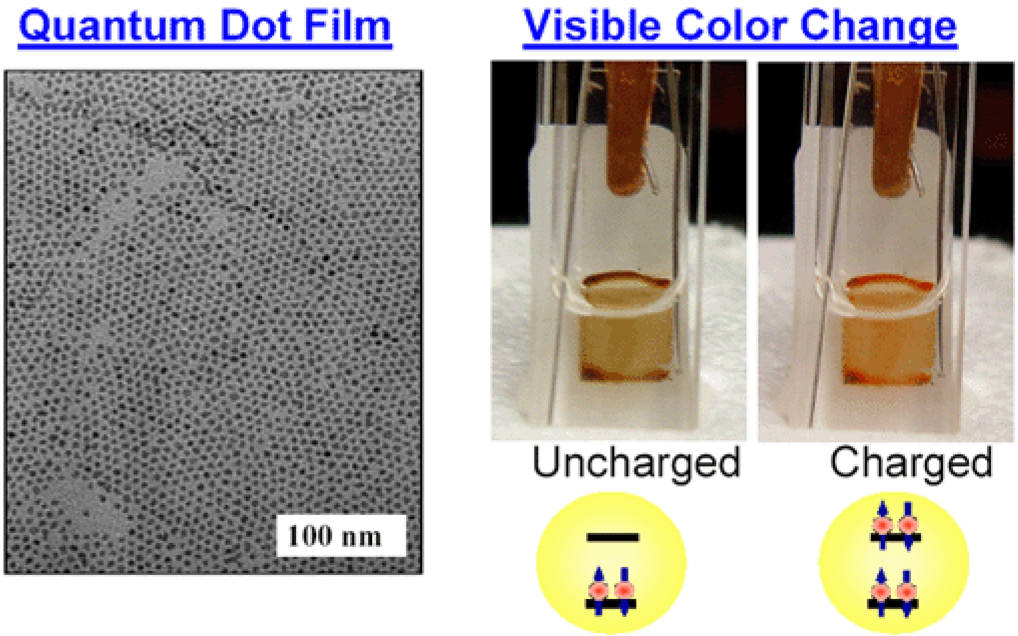


Fig. 20. Spin room temperature dynamics of colloidal CdSe quantum dots.

We use time-resolved Faraday rotation measurements to investigate the spin room temperature dynamics of colloidal CdSe quantum dots in both the charged and uncharged state. The evolution of the spin dynamics reflects the filling of surface states and the voltage-controlled electron occupation of the quantum dot states. Typical time-scans for different electrochemical charging voltages are shown below. Fourier transforms show that the two precession frequencies typically observed in uncharged CdSe quantum dots become a single spin precession frequency upon charging. It has been suggested that the higher frequency component is due to coherent precession of an exciton in the quantum dot. Since extra electrons in the quantum suppress detection of excitons, this frequency component is quenched when the dots are electrochemically charged.

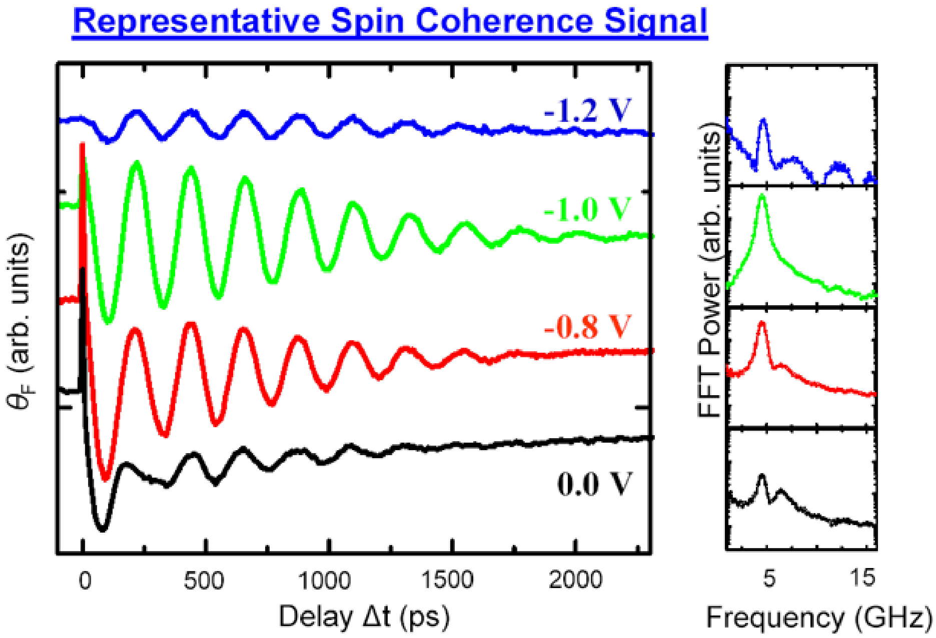


Fig. 21. Spin coherence signal taken at different frequency.

The increase in spin precession amplitude upon charging is likely due to the filling of surface states on the nanocrystal. These localized states have been shown to trap up to 50% of optically injected carriers within a picosecond. When these states are filled by electrochemically injected charges, the charge relaxation pathway is blocked and more charges remain in the conduction band, leading to stronger spin precession signal.

Fits to the Faraday rotation data indicate that the spin lifetime of the optically injected electrons increases, but direct measurement of this effect is precluded by the high degree of inhomogeneous dephasing due to ensemble measurement of quantum dots with random crystallographic orientations.

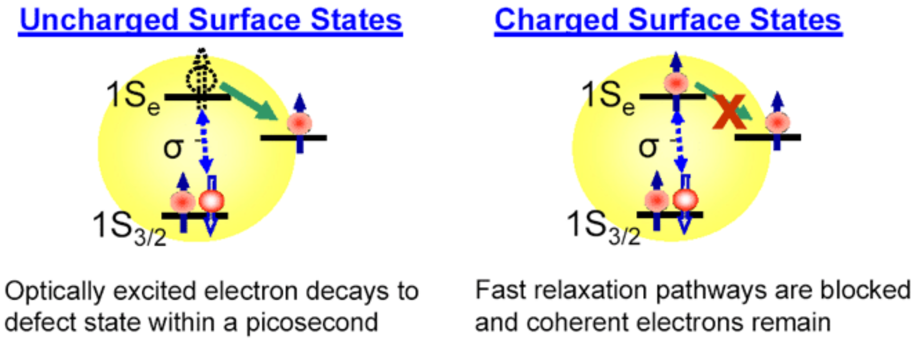


Fig. 22. Scheme of excitation and relaxation of QDs.

Self-assembled quantum dots (QDs) are unique structures in which to study spin coherence in zero-dimensions. They can be gated, doped, or put into more complicated heterostructures to tune their energy spectrum, control carrier occupancy, or modify coupling to the environment or neighboring dots. Such control may help to elucidate the underlying physics as well as enable new device paradigms. The InAs QDs studied here are grown with a partially-covered island technique which allows the ground-state energy of the QDs to be adjusted by changing the dot size and shape. Below (left) is a graph of photoluminescence (PL) intensity vs. detection energy for three samples with emission from the GaAs host at 1.52eV, the InAs wetting layer at 1.45eV, and several QD energy levels. At right is an AFM image showing the random distribution and excellent (<5%) size distribution.

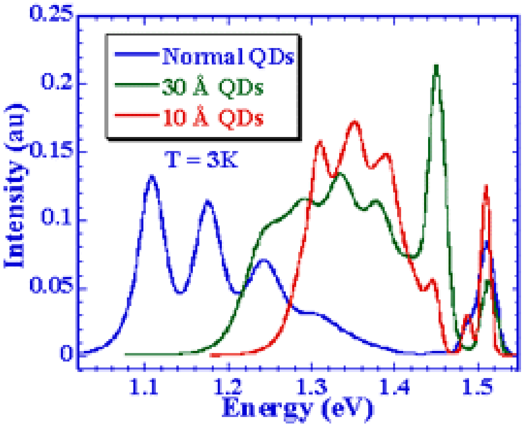


Fig. 23. The effect of degree spin-polarization changing.

The Hanle effect uses the degree of circular polarization r of photons from recombination to measure the component of carrier spin along the direction of observation. As illustrated in the inset below, a magnetic field along z causes the spins (initially along x) to precess in the x-y plane at the Larmor frequency W, which decreases the time-averaged spin along x. The time-averaged spin as a function of magnetic field gives the Lorentzian Hanle curve:

$$\rho(B) \propto \bar{S}_x(B) = \frac{\bar{S}_x(0)}{1+(\Omega T_2^*)^2}$$

where $\Omega=2\Pi g\mu_B B/h$, g is the Lande g-factor, and $_{\mu B}$ is the Bohr magneton. The curve width is inversely proportional to the effective transverse spin lifetime $1/T_2{*}=1/\ \tau_r+1/\tau_S$. The experiment setup is shown below. A circularly polarized laser creates the spin polarized carriers. The resultant PL polarization is analyzed by a variable wave plate (VWP) which alternates between 1/4 and 3/4 – wave retardence and alinear polarizer (LP). Theintensities I_+ and I_- are detected with a photodiode (PD) and used to calculate the polarization as $r = (I_+-I_-)/(I_++I_-)$.

The monochromator enables the spin lifetimes in different QD energy states to be spectrally distinguished.

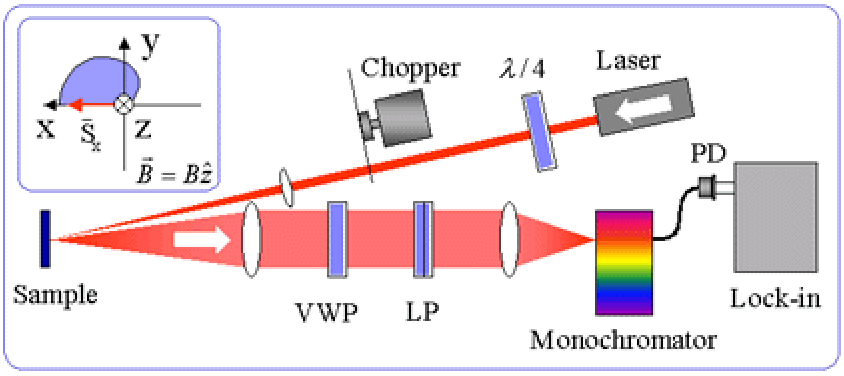


Fig. 24. Scheme of the optical experiments for detecting Hanle effect.

Below (top panel) are Hanle curves from one sample detecting at the ground state for five different temperatures. The curves are not single Lorentzians but can be fit well with bi-Lorentzians. This could result from a bi-exponential spin decay as has been seen in the chemically synthesized QDs (link). The lower panel shows scaled spin lifetimes extracted from the data for all three samples. While the short lifetime component does not change within experimental error, the longer component shows that the largest QDs are the most resilient to changes in temperature.

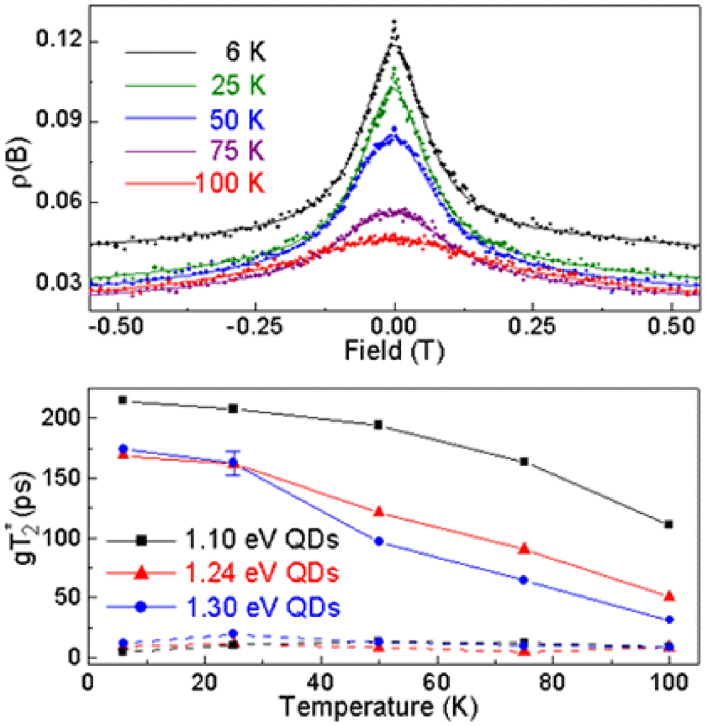


Fig. 25. The results of measurements taken at different magnetic field and temperatures for self-assembled QDs.

The spin lifetime also depends on which QD energy state is probed. Below (middle) is a density plot of PL polarization vs. magnetic field vs. detection energy. By normalizing the zero-field polarization to unity, the change in width of the Hanle curves is more apparent. The PL intensity (right) shows the ground and first excited states. Horizontal line-cuts (top) of the middle panel are Hanle curves at the peaks of the two states. The left panel is the zero-field polarization vs. detection energy.

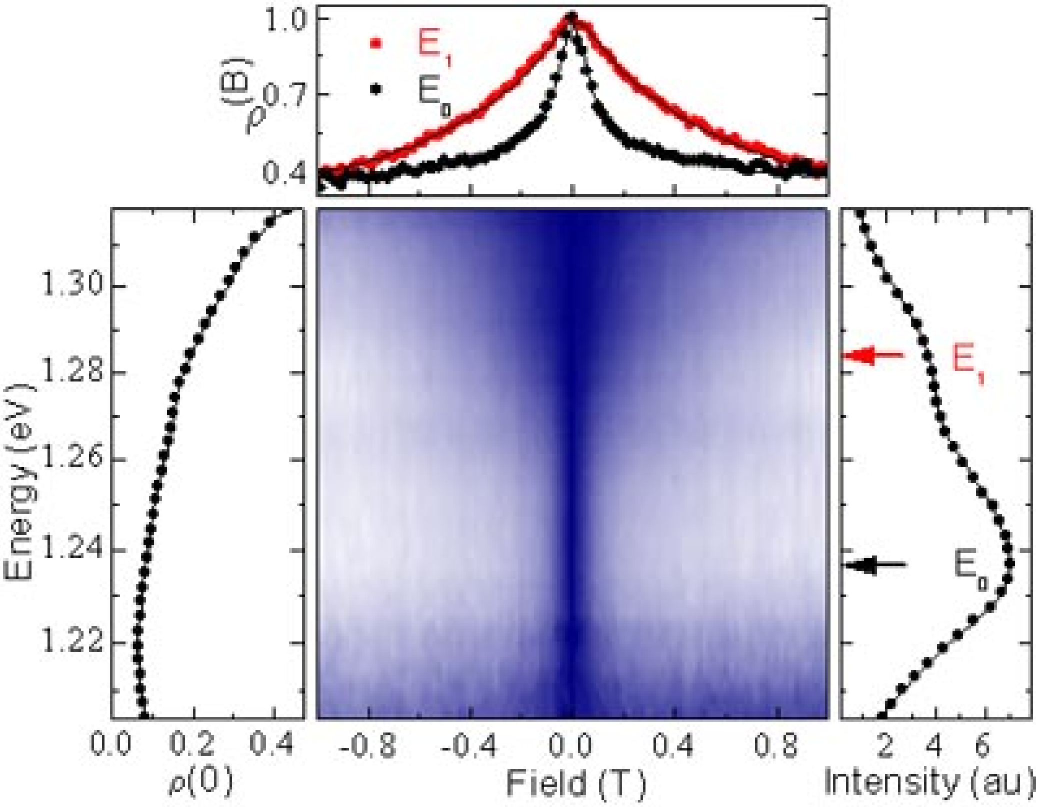


Fig. 26. PL polarization plot.

Compared to exciting above the GaAs energy gap, intensity-induced spin relaxation is dramatically suppressed when exciting at the wetting layer. Below is a series of Hanle curves with increasing excitation intensity exciting above the GaAs band edge. The right inset shows how the middle lifetimes vary with excitation intensity for excitation at 1.58 eV (black; GaAs) and 1.45 eV (red; wetting layer). The lifetime is strongly intensity dependent for the former case but is roughly constant for the latter case.

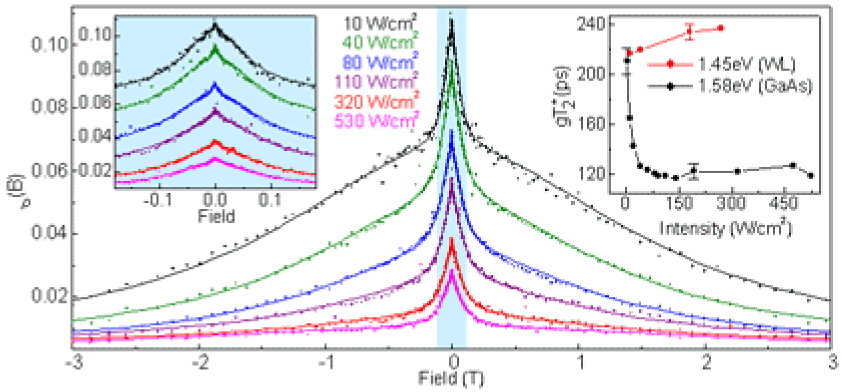


Fig. 27. Three concentric Lorentzian-like peaks appear at three field scales in these Hanle curves.

There are peaks with widths of ~2 T, ~0.1T and ~0.001T. These may be due to different spin relaxation mechanisms each with their own characteristic relaxation time.

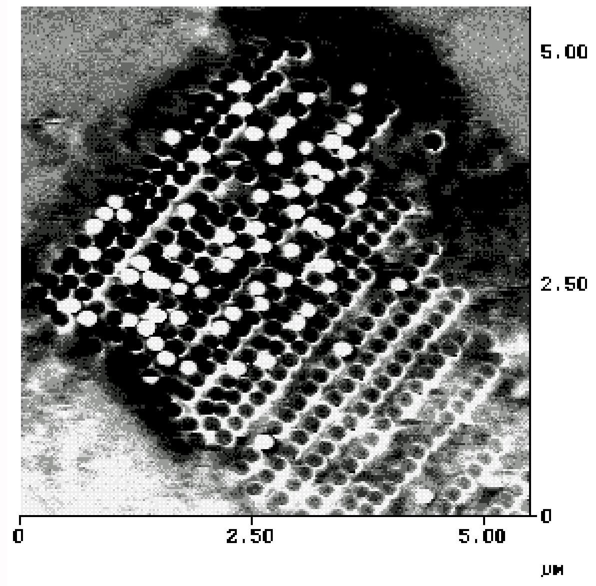


Fig. 28. Magnetic Force Microscope (MFM) image of the array after it was thermally randomized. The magnetic field from each ferromagnet is imaged and seen to be aligned along it's major axis, either pointing "up" (white) or "down" (black).

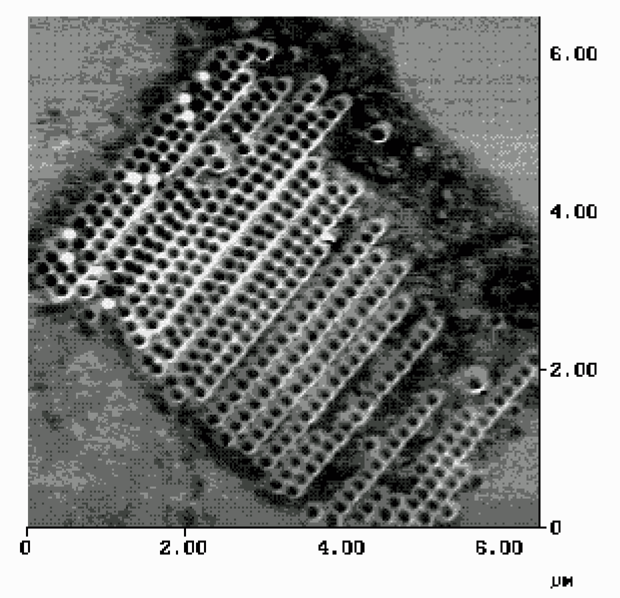


Fig. 29. A magnetic field of 200 Gauss is applied to align (almost) all the magnets.

In mammals, ferritin is used to store iron in the body in the form of an iron oxide bound in a spherical protein (apoferritin). The diagram below shows the assembly of the apoferrtin. First, a dimer (a) is formed burying the largest hydrophobic surface. Then the other hydrophobic surfaces (the shaded ends of the dimers) are buried to form the icosatetramer protein (d). Iron in the form of Fe(II) can enter through the open channels and is then oxidized on catalytic sites on the inside surface of the protein.

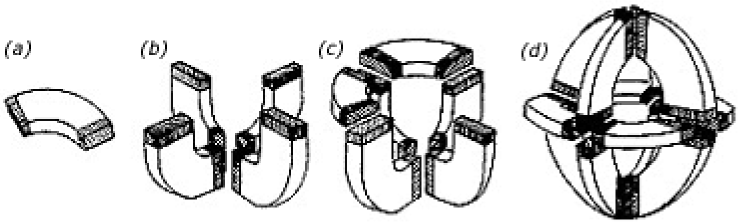


Fig. 30. Artificial ferritins model.

The growth of the crystal is limited by the apoferritin shell, allowing for a maximum of 4500 Fe(III) ions in natural ferritin in the form of ferrihydrite, an antiferromagnetic mineral. In collaboration with Professor Steve Mann at the University of Bath, our work has focussed on the magnetic properties of artificial ferritins where the ferrihydrite core is removed and replaced with either ferrihydrite or other minerals such as magnetite or maghemite, which are ferrimagnetic. Through this technique of biomineralization, the size of the particles may also be controlled from 100 to 3000 ions, allowing for systematic studies of magnetism in nanometer scale particles.

Recent advances in use of Qdots in biology studies support the promise of a quantum leap in the extensive use of Qdots in future biological applications. It is predicted that Qdots will be able to provide unprecedented sensitivity and selectivity over the traditional practices for molecular imaging.

The use of Qdots emitting in the NIR region will provide greater sensitivity and the longer lifetime of their excited states (as compared to organic fluorophores and proteins) will lead to for improved bioimaging. Despite the advantages for Qdots-based bioimaging, several issues related to Qdots need to be addressed before *in vivo* use, especially their toxicity.

at the surface of CdSe Qdots that suppressed their PL intensity and led to the release of Cd^{2+}, which is known to be toxic to human. The only way to partially recover the PL emission was to add excess Cd^{2+}, which is unacceptable for biological application. In addition, bare Qdots were reported to be cytotoxic. Some of the Qdots properties are limiting, such as size of Qdots, which sometimes is larger than the traditional organic marker dyes. As research on nanoparticles with novel properties continues, it should be possible to overcome these drawbacks and to develop multifunctional, multimodal Qdot-based systems for better biological imaging.

Chapter 11. Field emission from nanosilicon and cold cathodes

11.1. Introduction
11.2. Electron Field Emission from Porous Silicon.
11.3. Electron Field Emission from Silicon with Multilayer Coating.
11.4. Peculiarities of Electron Emission from Si Nanoparticles.
    11.4.1. Electron Field Emission from Nanocomposite $SiO_x(Si)$ and $SiO_2(Si)$ Films.
    11.4.2. Electron Field Emission from Si Nanocrystalline Films.
    11.4.3. Laser Produced Silicon Tips with $Si_xO_yN_z(Si)$ Nanocomposite Film.
11.5. Formation of Conducting Channels in SiO_x Coating Film.
11.6. Electron Field Emission from Si Nanowires.
11.7. Metal-Insulator-Metal and Metal-Insulator-Semiconductor Emitters.
11.8. Conclusion.
References to Chapter 11.

Chapter 11. FIELD EMISSION FROM NANOSILICON AND COLD CATHODES

11.1. INRODUCTION

Si tips and field emission arrays (FEAs) are readily fabricated from crystalline silicon using oxidation to create very sharp emitters. The processing steps can be carried out using standard fabrication tools. However, the silicon based emitters have some problems with stability and reliability. As a rule, silicon emitters are contaminated with a layer of molecules that could be adsorbed from the ambient, desorbed from an anode, or segregated from the bulk. The properties of this surface layer can affect the emission properties.

Clean silicon is quite reactive and can be contaminated within hours, even at low pressures. As a rule, Si is coated with grown film of natural oxide SiO_x with uncontrolled properties. To prevent the oxidation of Si in such reactive gases as H_2O, CO, or O_2 it is desirable to cover it by protective film with controlled properties. It can stabilize and, in some cases, increase the electron field emission. In the case of emitters made from silicon or other semiconductors, the electronic properties of the surface can also affect transport through the semiconductor. A number of semiconductor structures can provide nonlinear current-limiting.

Cold electron emitting devices, as a rule, have a complicated structure and they require a high vacuum and a high supply voltage. A large dispersion of angles of emitted electrons results in a poor resolution. The field emission devices based on carbon nanotubes addressed some of these problems. However, lifetime is a serious issue for these devices. Planar type electron emitters based on metal – insulator – metal (MIM) or metal – oxide – semiconductor (MOS) structure are very perspective to overcome the abovementioned problems.

In this Chapter the novel electron sources based on silicon with quantum effects are considered. Silicon electron sources are important first of all due to high developed Si-based technologies and perspectives of integration of solid-state and vacuum nanoelectronics devices. The peculiarities of electron field emission from porous silicon, silicon tips with multilayer coating and laser formed silicon tips with thin dielectric layer are described. In all cases the peaks on emission current-voltage characteristics have been revealed and explained in frame of resonant tunneling theory. Electron field emission from SiO_x(Si) and SiO_2(Si) films containing Si nanocrystals is considered in details. Special attention is paid to Metal-Insulator-Metal and Metal-Insulator-Semiconductor emitters. They are integrated in solid state and have plain design. The peculiarities of electron transport mechanism and the role of nanoparticles are considered in details and perspective of further development of Si-based electron field emitters is shown.

11.2. ELECTRON FIELD EMISSION FROM POROUS SILICON

There is a continuous search of new emission structures that may have certain advantages over existing emitter [1-6]. The porous silicon layers as on the tips and flat surfaces are actively investigated [7-11].

In Ref. [12] the porous silicon (PS) layer was grown both on flat Si surfaces and on Si tips of Si tip array. Arrays of silicon emitter tips were fabricated by wet chemical etching [13-15]. The cathodes were formed on (100) *n*-type Si wafers (N_d=10^{15} cm^{-3}) by patterning with Si_3N_4 as a masking material. The tip sharpening was performed by oxidation of the as-etched tips at 900 °C in wet oxygen followed by etching in HF:H_2O solution. This sharpening technique allowed the production of tips with a curvature radius of 10–20 nm. The height of the silicon tips was 4 μm while the tip density was 2.5×10^5 tips/cm^2. The radii of the tips and their height were estimated from scanning electron microscopy images (Fig. 11.1).

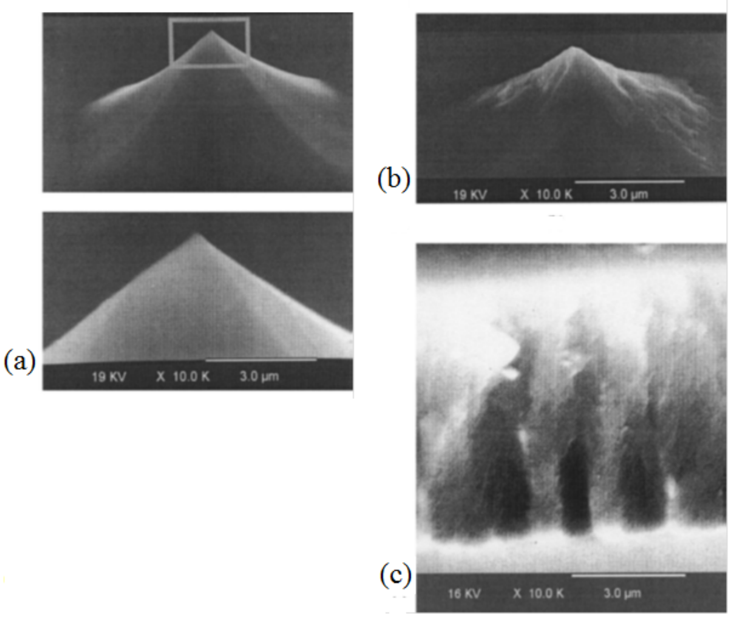


Fig. 11.1. SEM micrographs of silicon tips without (*a*) and with PS layer (*b*), and the formed silicon fibrils (*c*).

Porous silicon layers were formed by electrochemical etching of silicon tips as well as on the flat wafer surfaces [16, 17] (Fig. 11.1). The electrolyte composition was 1:1 of 48% HF and ethanol. The anodization process was held under illumination with the intensity of 30 mW/cm^2 for sufficient hole generation in *n*-type silicon. The anodization current density and the etching time were varied from 5 to 50 mA/cm^2 and from 15 to 60 s, respectively. The thickness of the obtained porous layers and consequently the height of silicon fibrils formed under the investigated conditions increased with the time of anodic etching. The pore width, and, hence, the degree of porosity also increased and the fibril thickness was decreased with the growth of anodization current density. Due to the higher resistivity of porous silicon in comparison with single crystalline Si the PS layer used was thin (<1μm) to ensure that operation was not current limited.

The measurements of the emission current from samples were performed in a vacuum system that could be pumped out to the stable pressure of 10^{-6}–10^{-7} Torr. The emission current was measured at the ungated cathode-anode diode structure. The test diode structures by sandwiching anode and cathode plates were fabricated. The emitter-anode spacing *L* was constant and it was chosen to be in the range 7.5–25 μm. Teflon film spacer was used to keep the two plates separated from each other. The investigated emitting structure was used as a cathode and a heavily doped silicon wafer or a molybdenum wire was used as an anode. The emission current-voltage characteristics were obtained with the current sensitivity of 1 nA over the voltage range up to 1500 V. A 0.56–1 MΩ resistor was placed in series with the cathode to provide short-circuit protection.

The *I–V* characteristics and corresponding Fowler-Nordheim plots of electron field emission from silicon array with a porous silicon layer are shown in Fig. 11.2. The threshold voltage for the silicon tips without PS layer was 530 V. By forming the porous silicon layer the silicon surface becomes rough and therefore an increase of the local electric field and emitting area can be observed in comparison with the case of an untreated silicon tip. An increase of emission area by a factor of 1.5–2 versus anodization time (at fixed etching current density J =25 mA/cm^2) is observed for an etching time of up to 30 s. Thinner fibrils are formed on the silicon surface with an increase of the etching time at a constant current density. These cause an enhancement of local electric field and decrease of the threshold voltage of electron field emission.

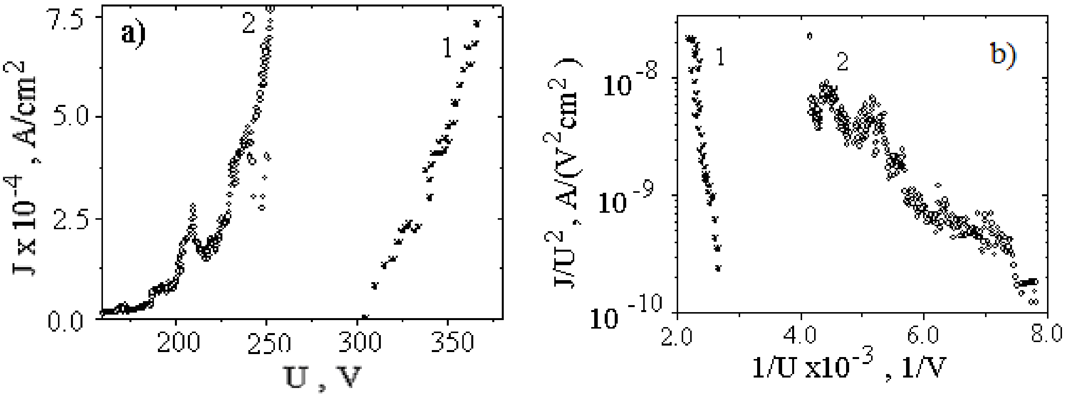


Fig. 11.2. Current-voltage (*a*) and Fowler-Nordheim (*b*) characteristics of electron field emission from Si tip arrays with porous silicon layer. The two regimes of PS formation are: 20 mA/cm^2 (1) and 50 mA/cm^2 (2), both for 15 s.

Depending on porous silicon preparation conditions two types of *I–V* characteristics, with smooth features and presence of peaks, are obtained. The peaks on *I–V* characteristics are caused by the resonance-tunneling effect. As a rule, after electrochemical etching and drying in air the surface of porous silicon is coated by an ultrathin layer of SiO_2 (≥2 nm). Therefore, the one-side double-barrier FE cathode is achieved. Ultrathin diamond-like carbon (DLC) film deposition on porous silicon layers does not suppress the peaks.

Further optimization of PS formation by electrochemical etching included the use of technological conditions proposed in Ref. [18]. In this case the PS is formed at low current densities

(1 mA/cm^2 - 5mA/cm^2) without applying the voltage [19]. The two-layered PS is formed. The lower layer is microstructured and the upper one is nanostructured. After obtaining the desired thickness of PS the samples were washed with acetone and dried at 100°C in air or in an atmosphere of acetone. The process of drying was more critical for samples obtained at high current densities.

The emission *I-V* characteristics of structures c-Si/PS are presented in Fig. 11.3. As can be seen, the increase of current density at electrochemical etching shifts curves in higher fields region (Fig. 11.3, curves 1-2). At lower current densities at electrochemical etching the peaks appear on emission *I-V* characteristics. Their shape does not change with repeated measurements (Fig. 11.3, curves 3*a*-3*c*). The presence of emission peaks can be explained as the result of resonant tunneling of electrons under high electric field. Multiple measurements of electron FE confirm the reproduction of the emission peaks pointing out on quite homogeneous size distribution of Si nanowires coated with the SiO$_x$ shell. In air atmosphere the diameter of Si nanowires is decreased due to partial oxidation and they can be separated on local Si nanocrystals (Fig. 11.4). It restricts the current flow during the electron FE [19]. At removing the partially oxidized layer the threshold voltage at electron FE is shifted to lower values.

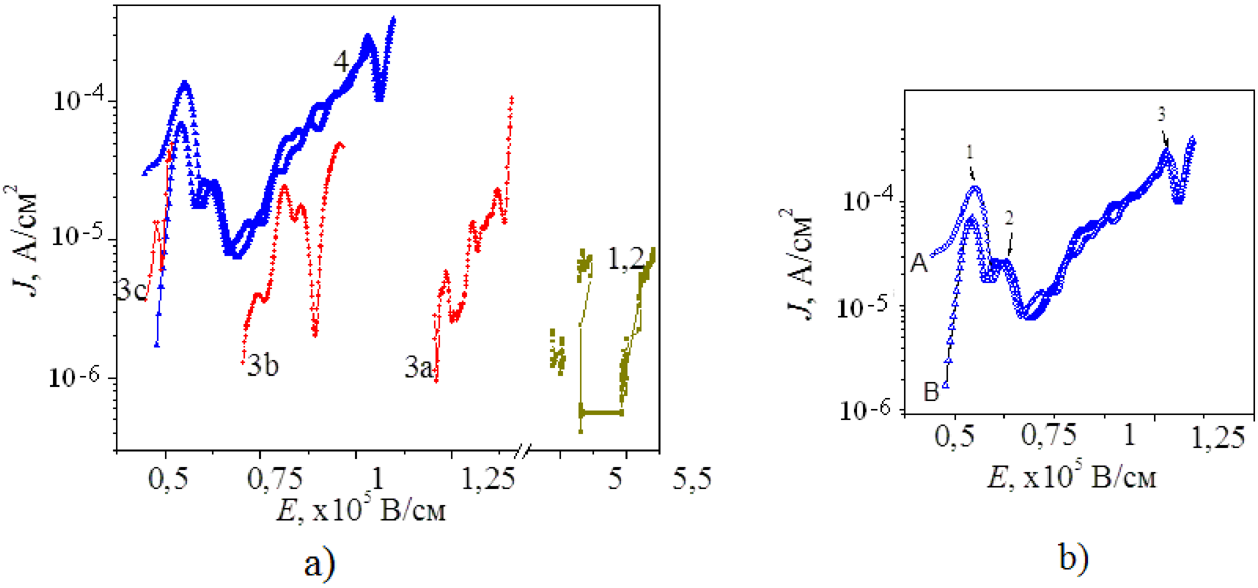


Fig. 11.3. (*a*) *J-E* characteristics of emission current from c-Si/PS with PS layer obtained at different currents densities at electrochemical etching: 1, 2- (4-5) mA/cm^2, 3*a*-3*b*- 2 mA/cm^2, 1- 1 mA/cm^2 and (*b*) behavior of curve 4 at repeated measurements.

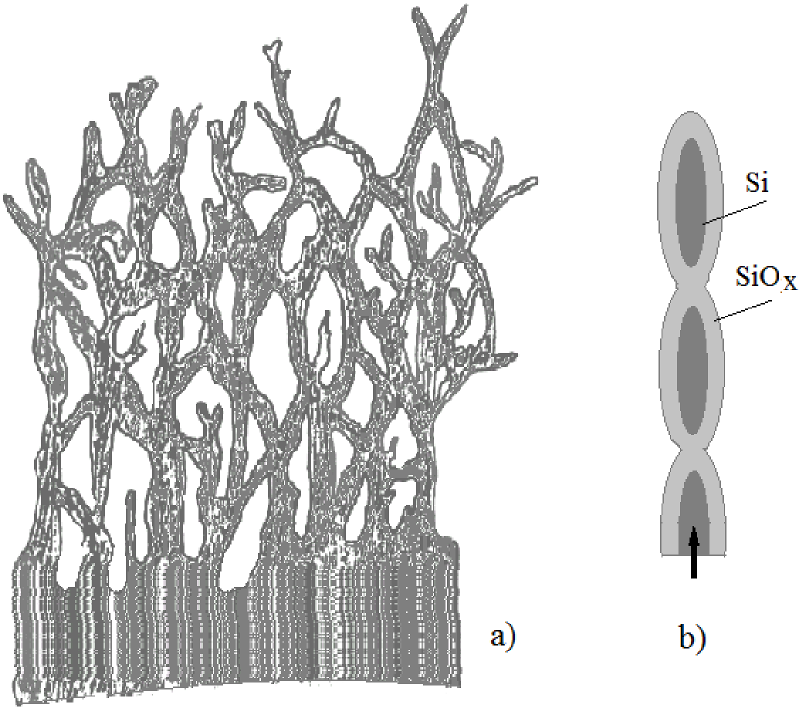


Fig. 11.4. Sponge-like structure of porous silicon (*a*) and view of one nanowire (*b*).

There is an interface between micro- and nanostructured layers of PS [20]. After removing the upper nanocrystalline layer in water solution of KOH the electron FE changed dramatically due to changing of the surface morphology. The efficient of electron field emission can be both increased (for high current densities at electrochemical etching J=4-5 mA/cm^2) and decreased (for low current densities at electrochemical etching J= 1 mA/cm^2).

The comparative detailed analysis of emission *I-V* characteristics (Fig. 11.3) and photoluminescence (PL) spectra [21] of c-Si/PS samples indicates that the samples that produced at the lowest current density have the best emission and low photoluminescence (PL) intensity, but in case of high current densities at electrochemical etching the relationship is opposite.

It is obviously caused by the diameters of the Si nanowires. At lower electrochemical etching current the nanowires are thicker and don't restrict the current flow. In the case of the higher electrochemical etching current the nanowires are thin enough and they can be fully oxidized in some thinnest region with creation of the separated nanocrystals (Fig. 11.4). That restricts the emission current and enhances the PL intensity.

The curve 4 in Fig. 11.3 shows the presence of three peaks at emission *J-F* characteristics of the samples. These peaks are reproduced at repeated measurements (Fig. 11.3(*b*)). The appearance of the emission peaks can be explained as a result of resonant tunneling through the c-Si/Si-NC/SiO$_2$/.../Si-NC/SiO$_2$/.../vacuum structure that is formed during the partial oxidation of PS [15, 22].

To explain the peculiarities of electron FE from c-Si/PS the schematic image of coral-like structure of porous silicon and partially oxidized one nanowire have been used (Fig. 11.4). The energy band diagram of formed resonant tunneling structure is shown in Fig. 11.5.

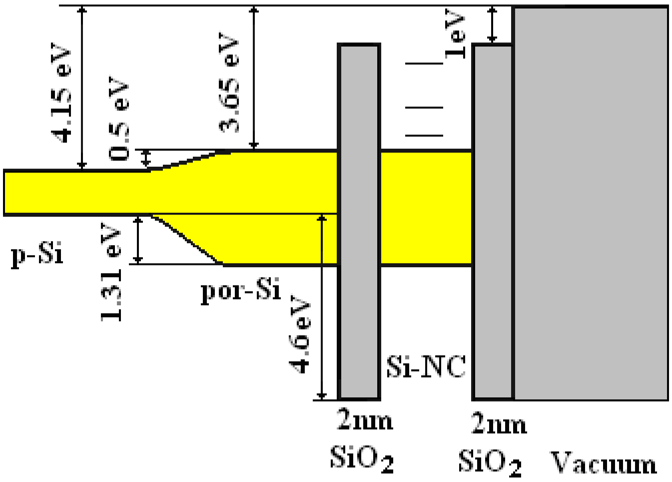


Fig. 11.5. Energy band diagram of the c-Si/PS system.

As a result of size quantization in nanowires there are resonant levels in conduction band of Si NC. The layers of SiO_2 and SiO_2+vacuum are the energy barriers of resonant tunneling structure. In approximation of infinite high barriers it is possible to estimate the diameter of Si-NC (*d*) in nanowires [23]:

$$d^2 = \frac{\left(n_2^2 - n_1^2\right)\hbar^2}{8m_n^*\Delta E_n} \qquad (11.1)$$

where $\hbar$ is the reduced Planck constant, m_n^* is the electron effective mass in Si, ΔE_n is the energy separation between resonance levels, n_1 and n_2 are the numbers of resonance levels. The approximation of infinite high barriers is quite suitable for high barriers of Si/SiO^2 and Si/SiO^2+vacuum. The thickness of the first energy barrier SiO_2 in resonance tunneling structure was supposed to be equal to 2 nm (the thickness of native oxide on Si).

To determine the value of the energy separation ΔE_n from the experimental *J-E* characteristics of the emission current it is necessary to calculate the field enhancement coefficient. For this the effective work function has been determined from slope of curves in F-N coordinate.

$$\Phi_{ef} = \frac{\Phi}{\beta^{*\,2/3}} \qquad (11.2)$$

where Φ is the work function, β^* is the electric field enhancement coefficient. The obtained value of effective work was 5.5×10^{-2} eV. The determination of Φ and β^* for case of PS is complicated task. The value of Φ was estimated based on empiric relationship proposed in Ref. [24]

$$\Phi + E_g \approx 5{,}5\text{eV}. \qquad (11.3)$$

In this case it was suppose that E_g=2 eV (in accordance with PL spectra [21]). The obtained value of work function for PS was 3.5 eV and calculated field enhancement coefficient was 545. The diameters of Si nanowires were in the range 1.5-2.8 nm [19].

After removing the upper nanocrystalline layer of por-Si the electron FE comes from the lower microcrystalline layer. In this case the electric field for emission was above 4 times higher. At repeated measurements the *J-E* characteristics shifted to the region of higher applied fields. The electric field enhancement coefficient calculated from the slope of the curve in the F-N coordinates according to equation (11.2) was significantly smaller $\beta^* = 65$ and no resonance peaks were observed [19].

11.3. ELECTRON FIELD EMISSION FROM SILICON WITH MULTILAYER COATING

The multilayer cathodes ($Si\text{-}SiO_2\text{-}\delta\text{-}Si\text{-}SiO_2$) (MLC) have been formed on silicon tip emitters. The ultrathin SiO_2 and Si layers in MLC structure were deposited using low-pressure chemical-vapor deposition technique, but in some cases the first SiO_2 layer was obtained by thermal oxidation [13,14]. The layers were deposited on flat silicon wafer and their thickness was measured by an ellipsometer. MLC structures with different thickness of individual layers were obtained by changing the deposition time.

During the electron field-emission measurements resonant-tunneling effects have been observed for some MLC structures. The measured current-voltage characteristics and the corresponding Fowler-Nordheim plots are shown in Fig. 11.6 and compared to the calculated electric field strength. The latter was possible by using the field enhancement coefficient obtained from the Fowler-Nordheim plots according to the procedure described previously [1, 25]. Two separate resonant peaks can be seen in experimental curves (Fig. 11.6). The estimated separation between quantum levels is ≈750 meV.

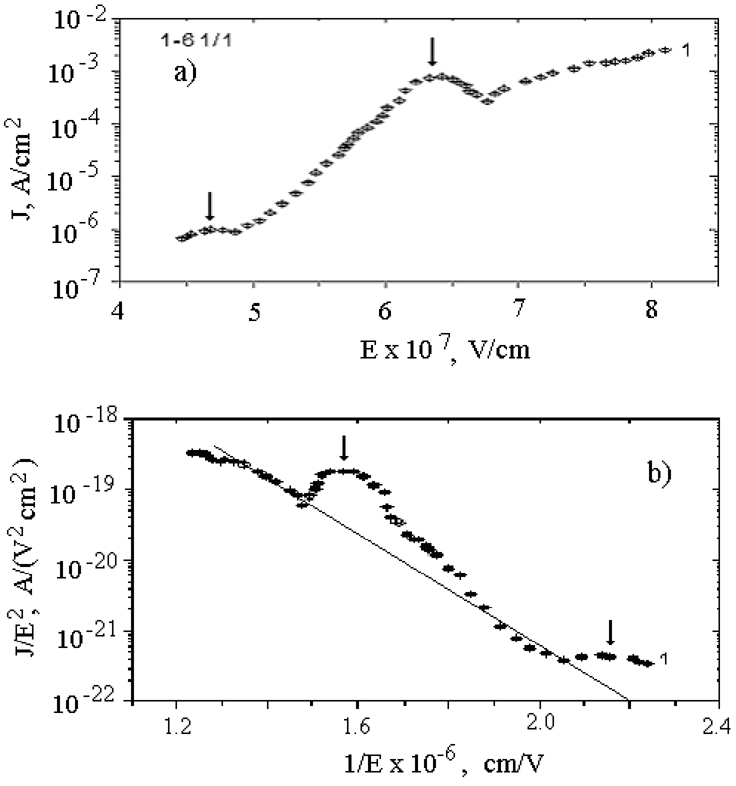


Fig. 11.6. Experimental current density dependence on electric field (*a*) and corresponding Fowler-Nordheim plot (*b*) ($d_1=d_2=d_3=2$ nm, $N_e=N_w=5.6\times10^{17}$ cm^{-3}, T=300 K).

The thickness of δ-doped silicon layer (quantum well) has been determined from the experimental value of the energy separation of two resonant quantum levels according to equation (11.1). The obtained value of d_2 = 1.2 nm. It is lower than given during the structure formation one (d_2=2 nm). The discrepancy between the experimental and calculated values of d_2 is caused by the reproducibility during the deposition, error in determination of d_2 thickness and accuracy of the used approximations.

11.4. PECULIARITIES OF ELECTRON FIELD EMISSION FROM Si NANOPARTICLES

11.4.1. Electron Field Emission from Nanocomposite SiO_x(Si) and SiO_2(Si) Films

In Ref. [26] SiO_x(Si) layers were obtained by thermal evaporation of silicon powder in vacuum ((2 – 3)$\times10^{-5}$ Torr). As it is shown in Refs. [27, 28], the using of thin SiO_2 coating on silicon tips increases the threshold voltage of field emission. The SiO_x film with high excess of Si (low value of *x*) was chosen in order to ease the electron transport from Si substrate to vacuum. The coatings were deposited as on Si tip arrays and flat silicon wafers. During the deposition process the temperature of substrate was 150 °C. Using quartz oscillator technique with an accuracy of the 3% the film thickness (*d*) and deposition rate were estimated. For the samples under investigation the deposition rate was ≈(2–3) nm/s. Measurement of the film thickness was performed by microinterferometer MII-4 and profilometer Dektak 3030. In the electrical conductivity and field emission experiments the oxides thicknesses were in the region of $d\approx10$–100 nm. Some samples were annealed in pure argon ambient under the temperature of 1000 °C for 5–40 min. As it was shown [29], such annealing led to formation of the additional silicon nanoinclusions (nanocrystals) in the oxide film.

The nanorelief of SiO_x(Si) ($x\approx0.3$) film surface and its transformation under subsequent treatments are shown in Fig. 11.7. As it can be seen, the initial sample surface is characterized by high roughness connected with large grain nanoprotrusion asperities up to 20 nm. It is caused apparently by the SiO_x(Si) film deposition technology: the film was produced by the deposition of both silicon clusters and SiO_x. In this respect the atomic force microscopy (AFM) results correlate with earlier obtained optical data [30] that demonstrate the composite nature of the as-grown SiO_x(Si) layer structure. As a result of the thermal annealing, the relief character substantially changes. The film surface becomes more uniform and the film surface structure in this case is characterized by nanoprotrusion asperities in the range of 1–3 nm. Thermal annealing of the sample causes decomposition of the initial material onto SiO_2 and Si phases, the latter being in the form of nanoclusters [31]. This fact is revealed, in particular, in the surface morphology of annealed film.

Subsequent sample partial etching in HF:H_2O solution did not affect on the film surface morphology.

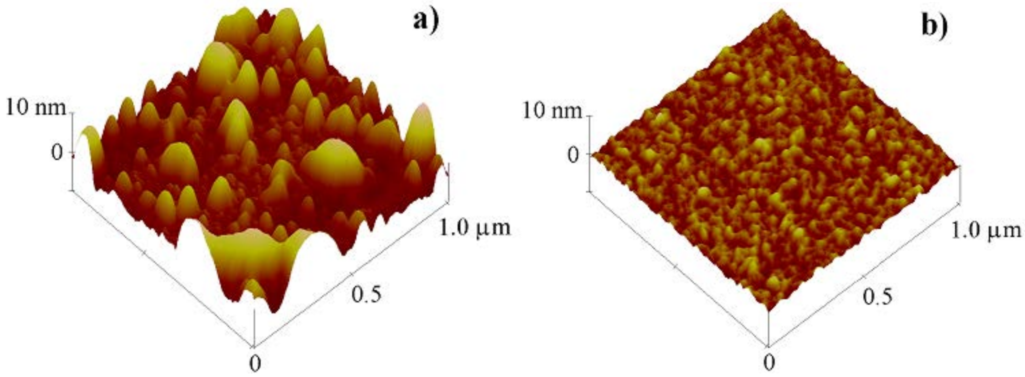


Fig. 11.7. AFM images of nanorelief of SiO_x film (*a*) initial, (*b*) after annealing (T=1000 °C, t=40 min) and following etching in HF solution.

Experimental investigation of electron field emission from silicon tips coated with SiO_x(Si) ($x \approx 0.3$) films [26] has allowed to obtain such main results: (*i*) coating of Si tips with SiO_x(Si) film, as a rule, decreases the electron field emission electron (FE) efficiency; (*ii*) the efficiency of electron FE decreases with the growth of SiO_x(Si) thickness; (*iii*) partial etching of SiO_x(Si) film increases the FE efficiency; (*iv*) high temperature annealing of SiO_x(Si) film and its transformation in SiO_2(Si) film decreases the electron FE efficiency (sometimes the emission is absent at all investigated regions of electric fields); (*v*) partial etching of SiO_2(Si) film in all cases allows to significantly increase the electron FE in comparison with SiO_x(Si) film and clean silicon tips; and (*vi*) long etching of SiO_2(Si) film removes it from Si tips and in such a way decreases the FE.

The significant field emission was observed from nanocomposite SiO_x(Si) and SiO_2(Si) deposited even on flat silicon surface. The results of electron field emission from SiO_x(Si) and SiO_2(Si) films in this case are shown in Fig. 11.8. As it can be seen, in the case of the initial SiO_x film the emission current has been observed at relatively high voltage (570–770 V) and has been 10^{-7}–10^{-6} A. The field emission from annealed samples was not observed in the investigated electric-field region. But after short etching of the annealed film in the HF:H_2O solution the emission is increased in comparison with nonannealed samples. The emission began at ≈375 V and the current reached the level of $\approx 10^{-5}$ A. The nonmonotonous region as a peak of emission current is observed at voltages from 430 to 540 V. For an initial SiO_x film, a peak of emission current appeared at voltage of ≈650 V. The existence of current peaks is explained by quantum-size effects in such structures and the appearance at defined electric fields of the additional transport mechanism - resonance tunneling of electrons.

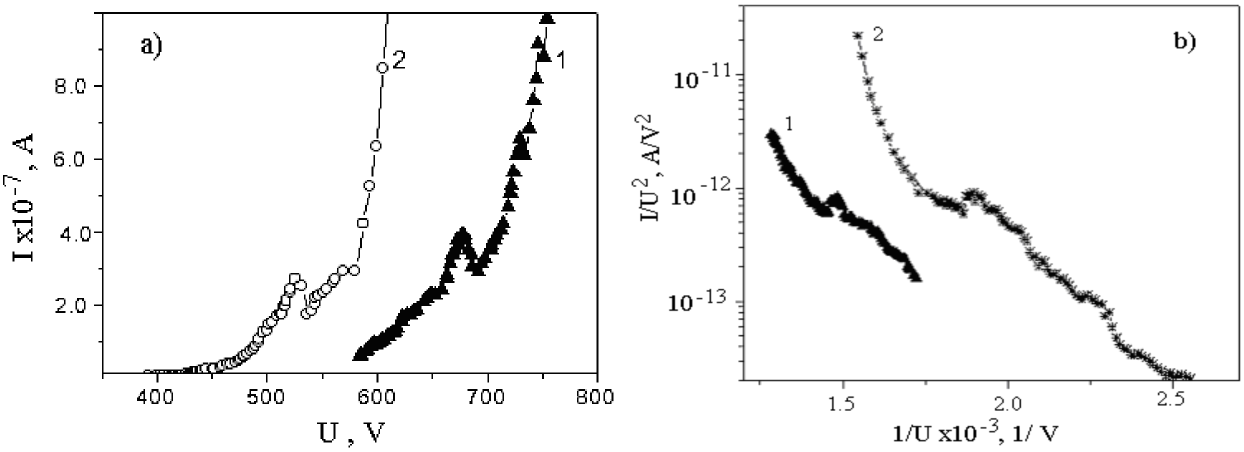


Fig. 11.8. Field emission current from SiO_x(Si) films: 1- initial, 2- annealed and etched in HF:H_2O solution. (*a*) shows the current-voltage characteristics and (*b*) shows the corresponding Fowler-Nordheim plots.

The position of resonant peaks on the voltage axis gives the possibility to estimate the splitting of energy level due to the quantum-size effect by considering the field enhancement coefficient, dielectric permeability, and layer thickness. The estimation of the energy level separation shows that $\Delta E=E_1-E_2=$ 260 meV.

I-V characteristics of emission currents from Si tips coated with SiO_x(Si) ($x\approx0.3$) film of 10 nm thick are shown in Fig. 11.9(*a*). The etching of SiO_x film allowed to increase FE efficiency. Figure 11.9(*b*) shows the *I-V* characteristics of emission from SiO_x(Si) film annealed at T=1000 °C and in such a way transformed in nanocomposite SiO_2(Si) film including Si nanoclusters embedded into SiO_2 matrix. The FE was not observed without etching of as annealed film. But after etching during t=3 s the emission was quite efficient (low electric field, $F<2\times10^5$ V/cm). In this case the two separate peaks were observed on emission *I-V* characteristic (Fig. 11.9(*c*)). Following additional etching (t=10 s) decreases the FE efficiency due to the removing of SiO_2(Si) film (Fig. 11.9(*b*)).

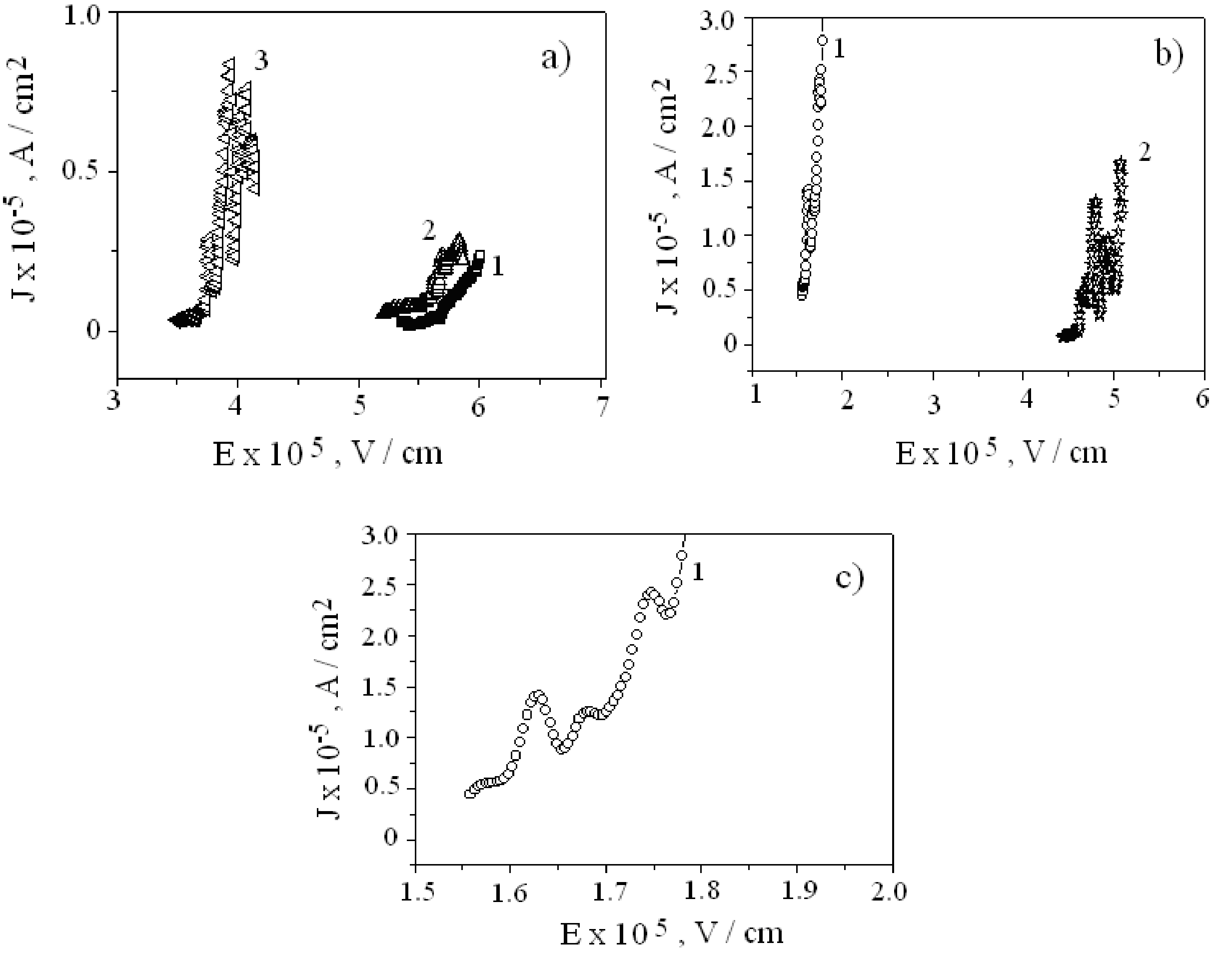


Fig. 11.9. Current-voltage dependences of emission current from Si tips coated with nanocomposite SiO_x(Si) film: (*a*) (1) initial Si+SiO_x(Si) (*d*=10 nm) system, (2) after following etching *t*=3 s, and (3) after additional following etching *t*=10 s; (*b*) (1) after high temperature annealing and etching *t*=3 s, (2) after additional following etching *t*=10 s, and (*c*) curve 1 taken from (*b*) in other scale.

Experimental results show that in some cases there are peaks on *I-V* characteristics of emission current. The observed peculiarities of the electron FE from SiO_x(Si) and SiO_2(Si) films on semiconductor tips can be explained, taking into account (*i*) phase composition of the initial and annealed oxide films, (*ii*) amplification of the electric field on the surface silicon clusters, and (*iii*) current carrying mechanisms during field emission.

At least a part of relief nanoprotrusions on the film surface observed by AFM are connected with silicon inclusions in SiO_x matrix. These Si clusters due to the local increasing of the electric field promote the electron field emission and current flow through the film. To leave the film, electrons overcome the combined Si–SiO_x+vacuum barrier at external surface. This barrier is caused by oxide layer coated at the surface of silicon clusters (Fig. 11.10(*a*)). The breakdown of this layer determines initial sudden change of field emission current. Further electron emission passes through Si-vacuum barrier.

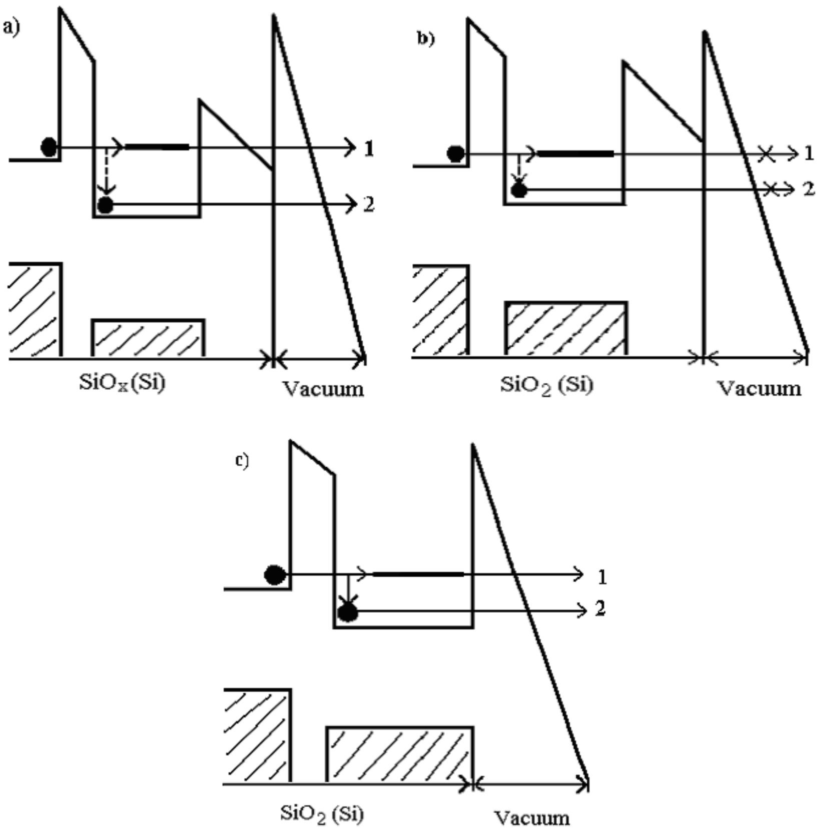


Fig. 11.10. Schematic energy diagrams for field emission mechanisms from Si–SiO_x(Si) system: (*a*) initial, (*b*) after annealing, and (*c*) after annealing and processing in HF:H_2O solution ((1) resonance tunneling process; (2) Fowler-Nordheim tunneling process).

Energy band diagram for annealed structure (Fig. 11.10(*b*)) is similar to initial one. However, in this case the values of the energy barriers at Si–SiO_2 interface are sufficiently higher than at Si–SiO_x interface. Significant barrier height at Si–SiO_2 interface (3.2 eV) is the reason for the absence of field emission at applied electric fields for annealed SiO_x(Si) films since external SiO_2+vacuum barrier has small probability of the electron tunneling. During partial etching in HF:H_2O solution the external SiO_2 layer is removed from Si inclusions on the surface. In this case, to be emitted, electron should be tunneling through only vacuum barrier (Fig. 11.10(*c*)).

The above mentioned decrease of the threshold electrical field for field emission from annealed and subsequently partially etched SiO_x(Si) films in comparison with initial ones (Fig. 11.9) can be connected with the appearance of two factors: (*i*) increase of the emission center (silicon clusters) concentration that leads to decrease of the distance between them; and (*ii*) decrease of their sizes that causes increase of the local electric field in the field emission center region. As it can be seen in Fig. 11.7, the sizes of Si inclusions for annealed film are smaller, and their density is higher.

The current peak observed on *I-V* characteristics can be explained by the appearance of the electron resonance tunneling transport. The possibility of resonant tunneling mechanism for multilayer Si–SiO_2–δ-Si–SiO_2 structure was theoretically predicted and experimentally proven in work [14]. Due to the similarity of the energy band diagrams of multilayer Si–SiO_2–δ-Si–SiO_2 system and SiO_2(Si) films it is possible to make qualitative suggestion about the resonance tunneling transport at electron field emission from SiO_2(Si) films. In conduction band of surface Si cluster the discrete resonant energy levels (minibands) may arise due to quantum-size effect that provides the additional resonant tunneling of electrons under certain electric field. The following experiments with varying of the Si cluster size and their density are needed for further support of the resonance tunneling mechanism realization. The optimization of cluster size and nanocomposite film thickness will allow to obtain the positions of current peaks on *I-V* characteristics in agreement with resonance tunneling theory.

11.4.2. Electron Field Emission from Si Nanocrystalline Films

SiO_2(Si) films with silicon nanocluster embedded in SiO_2 dielectric matrix is a good candidate for field emission cathodes coating [30, 32]. But in order to obtain the efficient emission it is necessary to etch the upper layer of SiO_2. At electron FE from SiO_2(Si) films the effect of resonant tunneling has been observed [30]. It is expected that direct deposition of film of the nanocrystal materials (nc-Si) on a flat surface can be promising for efficient field emitters [33].

In Ref. [34] the results of studies of the electron field emission from flat nc-Si films formed by laser induced decomposition of silane (SiH_4) [35] have been presented. The silane from gas pipeline flows to the center of the reactor where the CO_2 laser beam has crossed it. The reactor was connected to the molecular-beam analyzer, which included the time of flight spectrometer (TOFMS). Reaction products extracted from a conical nozzle into a vacuum and then passed through a sieve in deposition chamber, which preceded the camera TOFMS. Coated with carbon micro rids (sieves) were fixed towards the flow at a distance of 30 cm from the exit nozzle. They served to capture the particles from the cluster beam with or without selection by size using the rotated discontinuous wheel. In typical experiments the nc-Si films were deposited at such technological conditions: the silane flow rate was 40 cm^3/s, the flow rate of helium was 1100 cm^3/s, the total pressure was of 350 mbar and power of CO_2 laser pulse was 50 mJ.

In this experiment the Si nanoparticles were selected by size. Their diameter was in the range of 3.7-4.2 nm. Si nanoparticles were deposited on silicon substrates. The size of Si nanoparticles is very important to enhance the electric field at electron FE. Morphology of the cathode surface was studied with AFM. AFM image of Si nanoparticles on the surface of the wafer shown in Fig. 11.11. As it can be seen, there are a lot of individual nanoparticles, which are grouped in larger clusters. Current-voltage characteristics of emission current and dependence in the Fowler-Nordheim coordinates are shown in Fig. 11.12(*a*,*b*), respectively. Figure 11.12(c) shows the characteristic in the different scale. The efficiency of field emission is quite high for the flat structure (without special formation of tip). There are some nonmonotonous regions of *I-V* characteristics. In the Fowler-Nordheim coordinates these nonmonotonous regions appear as emission current peaks (Fig. 11.12(*b*)). Repeated measurements have revealed that the low voltage peak 1 in Fig. 11.12(*b*) is stable (reproducable), but peaks 2 and 3 have disappeared. To explain the results we consider a layer structure of Si nanoparticles. There wasn’t special heat treatment for oxidation of Si particles. But the existence of oxygen as impurities in gases and in the reactor cannot be excluded. In Ref. [36] the particle sizes were estimated using the electron microscope with high resolution (HREM). There was revealed the existence of oxide shell around the Si particles. Thickness of the oxide shell decreased linearly from 2.9 to 0.8 nm with decreasing particle size from 33 nm to 6 nm. In our case, the particle size was 3.7-4.2 nm and thickness of oxide shell was <0.8 nm.

At the field emission the electrons from the silicon substrate are injecting into Si nanoparticles layer, pass through it by direct tunneling through the oxide shells between Si nanoparticles and are emitted into vacuum. Due to very thin oxide shells the tunneling through the oxide doesn’t limit the emission current.

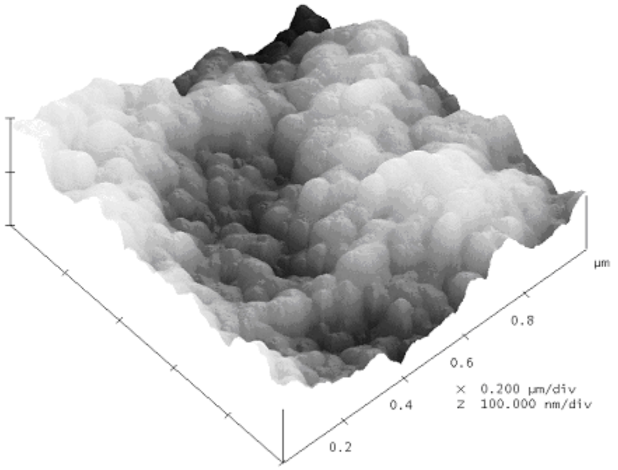
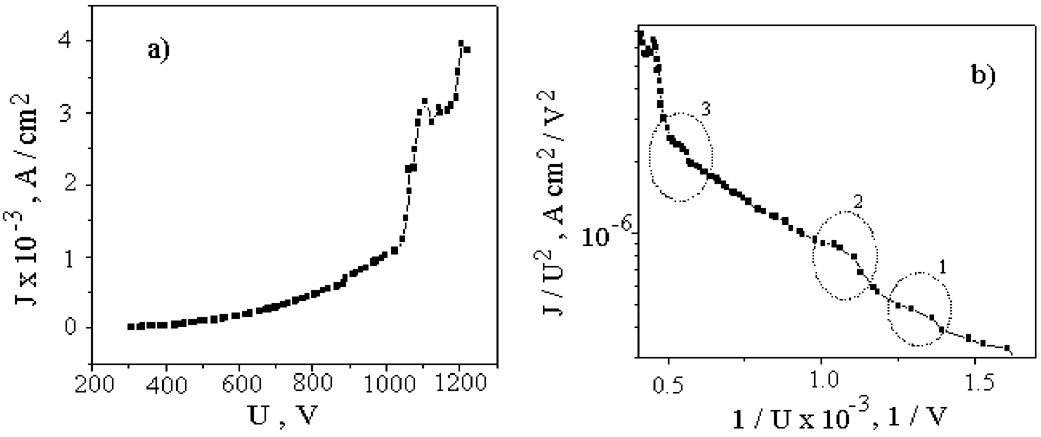


Fig. 11.11. AFM image of Si nancrystalline film on the silicon wafer surface.



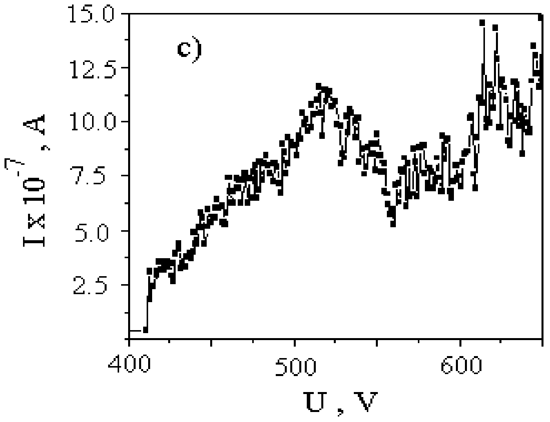


Fig. 11.12. *I-V* characteristics (*a*, *c*) and the corresponding Fowler-Nordheim plot (*b*) of emission current from Si nanoparticles ((*c*) is different scale).

High efficiency of electron FE can be explained by the enhancement of the electric field on Si nanoparticles. In Ref. [37] for calculation of the enhancement coefficient the model of "floating sphere" has been used. Distribution of electric field in this case is

$$E(\theta) = (\frac{h}{r})E_0 + 3E_0 \cos\theta \qquad (11.4)$$

where r is the radius of the sphere, h is the altitude of the sphere above the ground (substrate), E_0 is the plane-parallel field of space, and θ is the polar angle. In our case the h is the distance from the substrate surface to the particles, i.e. the thickness of the layer of nanoparticles. Accordingly, the electric field near the top of the sphere equals $E_{max} = (h / r + 3) E_0$.

Non-monotonic regions of *I-V* characteristics of emission current are caused by resonant tunneling effect. Because each Si nanoparticle is surrounded by the layer of silicon oxide and it is very small, it can be represented as the quantum well (dot) surrounded by energy barriers. Since Si nanoparticles have small size they show the quantum size effect. As a result, there are resonant levels in the conduction band of Si nanoparticles. At certain values of the electric field the additional mechanism of electron transport, namely, resonance tunneling is realized. Due to this the resonance peaks are observed in in current-voltage characteristics of emission current.

11.4.3. Laser Produced Silicon Tips with $Si_xO_yN_z$(Si) Nanocomposite film

The silicon tip array was formed by a series of single laser pulses which led to the formation of single conical tips [38, 39]. The arrays were relatively uniform. In this process, the silicon substrate (*n*-Si) was heated locally above its melting point by a pulse from a YAG:Nd^{3+} laser. The diameter of the focused laser spot was 20–30 μm. The region exposed to the laser beam was heated to a high temperature and melted leading to strong material evaporation. When the temperature decreased to the crystallization point the conical surface hardened and the arrays of silicon cones could be prepared. Their distance to each other was about 50 μm, the height was in the range 0–100 μm and the curvature radius at the top was ≈1 μm. The laser ablation of *n*-Si took place in air. Due to the pressure of the air (10^5 Pa), most of the silicon particles created by laser heating that had escaped from Si were deposited back on the target and on the conical surface. As the pulse energy was high (0.2 J), there was a broad distribution of particle sizes ranging from several nm to several μm. The lighter particles were backscattered further from the plume axis (as far as several mm) than the heavier ones. As a result, the conical surface had many protuberances on it and was covered with the nanocomposite film consisting of the nanocrystalline silicon grains in a $Si_xO_yN_z$ matrix. (Fig. 11.13).

Fig. 11.13. SEM image of top of the silicon tip formed by laser.

To analyze the content of the surface layer of the silicon tip, Auger profiling was used (Fig. 11.14) [38]. Oxygen and nitrogen were observed down to 30 nm from the surface. The real thickness of the nanocomposite film on a silicon surface was smaller and a value of ≈30 nm was obtained due to profiling of a nonflat surface during the Auger measurement.

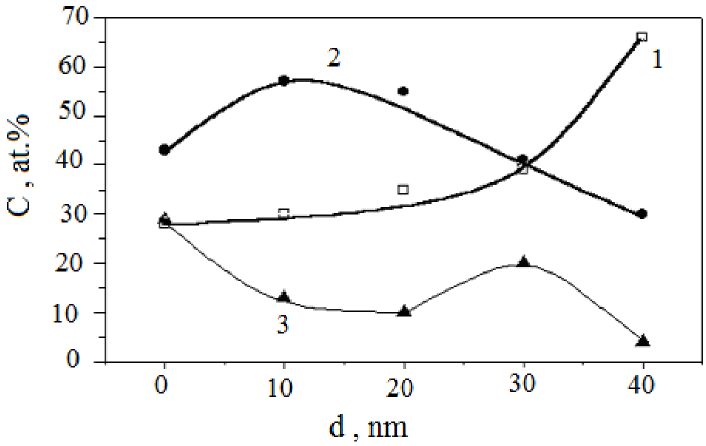


Fig. 11.14. Auger profiling spectra of silicon tip array coated with nanocomposite film (1- silicon, 2- oxygen, 3- nitrogen).

Photoluminescence spectra were observed from laser-produced silicon tips (Fig. 6.15) [38] and their origin lies in the presence of the nanocomposite film $Si_xO_yN_z$(Si) on the surface of the silicon tips. Due to small-size nanocrystals in the $Si_xO_yN_z$ matrix, the quantum-size effect causes direct optical transitions. After removing the nanocomposite film $Si_xO_yN_z$(Si) in a solution of HF in water, no peaks were observed in the photoluminescence spectra.

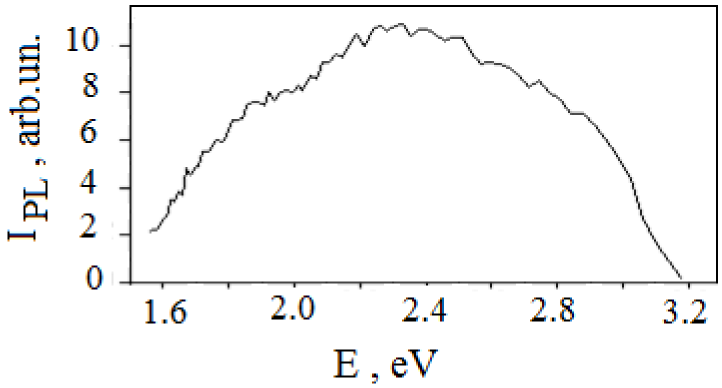


Fig. 11.15. Photoluminescent spectrum from Si tips coated by nanocomposite $SiO_xN_y(Si)$ film.

The typical current–voltage characteristic of electron field emission from laser produced silicon tip array corresponds to Fowler–Nordheim law. For the explanation of the features of current–voltage characteristics and their comparison for different tips with nanocomposite film the turn-on (threshold) voltage (V_{th}), emitting areas (α), and local field enhancement coefficient (β^*) have been determined from the Fowler–Nordheim equation according to procedure described in [1, 25, 40]. For the laser produced silicon tip array relatively high emission parameters, namely effective emission areas around $\alpha = 1.75\times10^{-8}$ cm^{-2} and local field enhancement coefficient $\beta^*=360$ have been obtained.

Resonant-tunneling phenomena have been seen on some laser produced samples that have a nanocomposite $Si_xO_yN_z(Si)$ film on their surface. As a rule, one or two resonant peaks have been observed. The current-voltage curves show asymmetry around the maximum in the current, which is typical for electron transport through resonant-tunneling structures. These structures consist of Si quantum-size nanocrystals in the oxynitride matrix on a silicon cone surface. Due to the quantum-size effect, there are some energy levels in the quantum well region that cause an increased tunneling probability under certain values of the electric field.

11.5. FORMATION OF CONDUCTING CHANNELS IN SiO_x COATING FILM

A simple and low-cost method to form high roughness structures on Si substrates is the chemical or electrochemical etching that produces porous Si (por-Si) [41-45]. The EFEs from emitters based on the por-Si structures have been actively studied and discussed [19, 46-48]. However, por-Si is a complicated material with a wide bandgap due to the quantum-size effect. It contains an ultrathin oxide layer on the surface as well as different kinds of adsorbed atoms such as hydrogen, oxygen, and fluorine. All these factors influence the work function of electrons at field emission and cause some modifications of emission current. Deposition of diamond-like carbon (DLC) and related films onto the surface of por-Si is known as a good method to decrease the instability of emission current [8, 49-52]. SiO_x(Si) films were used to enhance and stabilize the EFE from semiconductors such as Si [26]. The SiO_x(Si) films were annealed to form nanocomposite SiO_2(Si) films containing Si nanocrystals embedded in the dielectric SiO_2 matrix [26]. Nanocomposite SiO_x(Si) films with different Si compositions can be compared with the por-Si structures regarding their application as EFE emitters. In contrast to Si wires in the por-Si structures, Si nanocrystals are embedded in the nonstoichiometric SiO_x(Si) matrix and covered with protection coating (SiO_2). The protection role of this coating is the same as in the case of DLC films [52, 53]. The size of Si nanocrystals embedded in dielectric matrix and their concentration after annealing depend on the Si concentration in the films before annealing. The initial films with different Si concentrations were grown by thermal evaporation of Si powder [26, 30, 53] or plasma-enhanced chemical vapor deposition (PECVD) from gas mixture [54-56]. The Si nanocrystals embedded in SiO_2 matrix are separated by dielectric layers. Therefore, to obtain conducting Si wires through the SiO_x(Si) films in Ref. [57] the method previously proposed in Ref. [58] has been applied. Electrical conditioning at high current densities through the films leads to a connection of separate Si nanocrystals and in such a way to formation of conducting Si wires.

For investigations described in Ref. [57], different kinds of substrates were used. Glass and flat Si wafers were used for the film thickness and refractive index measurements. Si tip arrays and flat Si substrates were used for current-voltage and EFE current-voltage measurements.

SiO_x(Si) films with $x\approx1.2$ were deposited by PE CVD method from the gas mixture of $N_2O:SiH_4=0.19:1$ at the substrate temperature of 75 °C. The SiO_x(Si) films with $x\approx0.3$ were deposited by thermal evaporation of Si powder under the residual gas pressure and substrate temperature of $(2.7\text{-}4.0)\times10^{-3}$ Pa and 150 °C, respectively. The values of a stoichiometry index x for SiO_x films were determined by two methods described in Ref. [59]. In the first method, the compositions of SiO_x films were determined from the experimentally measured positions of the edges of the bands of optical absorption spectra. The second method was based on the determination of the positions of the bands of the infrared absorbance spectra of SiO_x films in the range of 1000–1100 cm^{-1} and referring their positions to the film composition. The thicknesses of deposited films on flat Si substrates measured by the Dektak 3030 profilometer were 10–100 nm. All samples were annealed during 5-40 min at 1000 °C under argon flow to induce phase decomposition of nonstoichiometric SiO_x(Si) films and formation of Si nanocrystals in SiO_2 matrix

[26, 29, 31,]. In order to etch the coating the ammonium fluoride [NH_4]F solution was applied. It etched selectively SiO_2 and opened the Si nanocrystals on the film surface [60]. The etch rate of the SiO_2 film was 150 nm/min. The thicknesses of 7.5 and 25 nm were etched off in 3 and 10 s, respectively. Following chemical etching, all samples were rinsed in double de-ionized water.

The procedure used to form the conducting channels in DLC films [58] was applied in Ref. [57] for SiO_x(Si) films. A Si tip array was used as the formation electrode. The mean contact surface area of the Si tip was 3×10^{-9} cm^2. The conducting channels were formed between the flat Si substrates coated with SiO_x(Si) films and the formation electrode. The bias voltage was increased until the current through the film was $I\approx1\times10^{-6}$ A. To create single conducting channels, a gold wire of 80 μm in diameter was used. The tip of the gold wire was etched to obtain a contact surface area of 3×10^{-9} cm^2.

In the described below experiment a 60 nm SiO_x(Si) film with a stoichiometric index of $x\approx1.2$ was deposited onto a flat Si substrate by PE CVD and annealed afterward. The Si inclusion density in the dielectric matrix was small and it was used for conducting channel formation inside the films according to Ref. 58. The current-voltage characteristics were measured using the Si tip array as the formation electrode and they are shown in Fig. 11.16. The current-voltage characteristics after electrical conditioning (curves 1*a*–4*a* in Fig. 11.16) compared to the nontreated curves (1–4) have much smaller threshold voltage and up to four to five times higher current in the low voltage region. The current transport characteristics through the films tested after electrical conditioning did not change during subsequent measurements. The current transport mechanism during the channel formation was determined with the assistance of a gold electrode. In Fig. 11.17, the current-voltage characteristics before (curve 1) and after the electrical conditioning (curve 2) of annealed SiO_2(Si) film are shown. The mechanism of current transport changed after the electrical conditioning (curve 2). At increased applied voltages the electrons tunnel into the conductive band of dielectric and flow according to the Poole–Frenkel (P-F) law (curve 2).

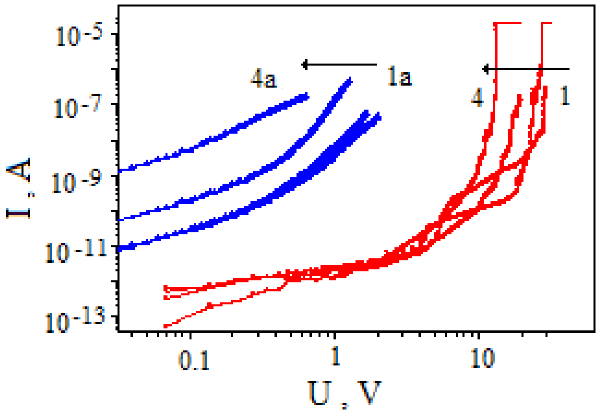


Fig. 11.16. Current-voltage characteristics of SiO_x(Si) ($x\approx1.2$) films before (curves 1–4) and after electrical conditioning (curves 1*a*–4*a*).

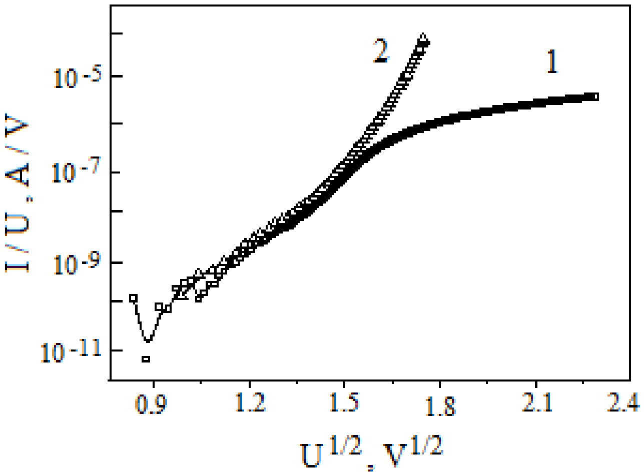


Fig. 11.17. Current-voltage characteristics of SiO_x(Si) ($x \approx 1.2$) film measured by using the single gold electrode before (curve 1) and after electrical conditioning (curve 2) in the P-F coordinates.

Some explanations of the film conditioning mechanism were presented in Refs. 61-63. A possible reason for the start of channel formation is the energy release due to the electron transport from the conduction band to the traps in the bandgap or from the conduction band to the Fermi level of the formation electrode. When the electrons flow through the dielectric-metal interface (at the anode contact) they will release energy that can be estimated from the expression $E=Q \times \varphi$ [61-63] (here Q is the charge flow and φ is the contact potential difference at the dielectric-metal interface). For example, the released energy is E_1=0.015 J /cm^2 and E_2=0.02 J /cm^2 at the Al–SiO_2 ($q\varphi$=3.3 eV) and the Pt–SiO_2 ($q\varphi$=4 eV) interfaces, respectively. The estimations show that the energy released by electron trapping or falling down to the Fermi level of metal electrode is two to three times higher than the bond breaking energy in SiO_2 (33 kJ/cm^3). The higher is the work function of the formation electrode, the higher is the released energy density. As shown in Ref. 64, the electrical conditioning procedure depends on the anode material. The electrical conditioning and channel formation capability are very high for the anodes with the work functions ≥4 eV.

The amorphous SiO_2 films have many Si–H and Si–OH bonds that are the main source of film disordering. During the formation of the films the bonds, mentioned above, act as centers for creation of conducting channels [65]. Break of Si–OH and Si–H bonds produces uncoupled Si atoms during the current transport. These uncoupled Si atoms add energy levels into the bandgap of the energy spectrum of SiO_2. In the course of the current transport through the films, electrons are captured in the traps that lie in the middle of the bandgap and heat up. When the applied voltage is increased, electrons tunnel to the conduction band and then flow according to the P-F law. They are frequently trapped at deep levels in the energy bandgap or fall down to the Fermi level of the anode with some energy release. The released energy is higher than that needed for the bond breaking $SiO_2 \rightarrow SiO+O$ (550.6 kJ/mol). Therefore, the number of Si atoms with uncoupled bonds increases and results in the increase in allowed states in the bandgap of dielectric. Increased concentration of free Si atoms is the main reason for the uncoupled Si bonds to connect to each other. It can result in the formation of conducting Si channels and change in the current transport behavior. As shown in

Refs. 61-63, the heat loss was small, and the temperature effect could be neglected during the formation procedure.

To characterize the structure of SiO_x ($x≈1.2$, 60 nm) films the x-ray diffraction has been used [57]. X-ray spectrum of film after annealing, but before conditioning confirms the amorphous structure of SiO_x film with small quantity of Si inclusions [66]. But after conditioning the x-ray diffraction shows the existence of crystalline peaks that confirm the increase in the number of Si inclusions after conditioning treatment. The morphology of the sample surface measured by atomic force microscopy (AFM) did not change after the treatment. It can be explained by the fact that no critical breakdown (with melting and cracking of material) occurred in the films under study.

The emission properties of SiO_x(Si) films with various values of stoichiometric index x before and after thermal treatment, and after electrical conditioning were compared. The stoichiometric index influences both the concentration and the size of Si inclusions.

Coating of emitter surface with SiO_x(Si) ($x≈0.3$) films improves the efficiency of EFE. Maximum improvement was achieved when the thickness of coating was in the range of 40–50 nm. The chemical etching and annealing influence were stronger for thicker SiO_x(Si) films. At optimized chemical etching treatment very small Si inclusions remained on the top of films, which resulted in higher field enhancement factors. For thicker SiO_x(Si) films (ranging from 50 to 100 nm), the EFE process began only after short-time etching. The EFE from 100 nm thick film disappeared after annealing. It can be explained by the fact that thick dielectric SiO_2 upper layer is almost nontransparent for field emission. The emission curves of the initial silicon tips, after SiO_x(Si) film deposition and following thermal annealing, and after the chemical etching were fitted with linear functions and plotted in Fig. 11.18. As was estimated the work function values decreased from 4.15 eV to 3.16 eV [57].

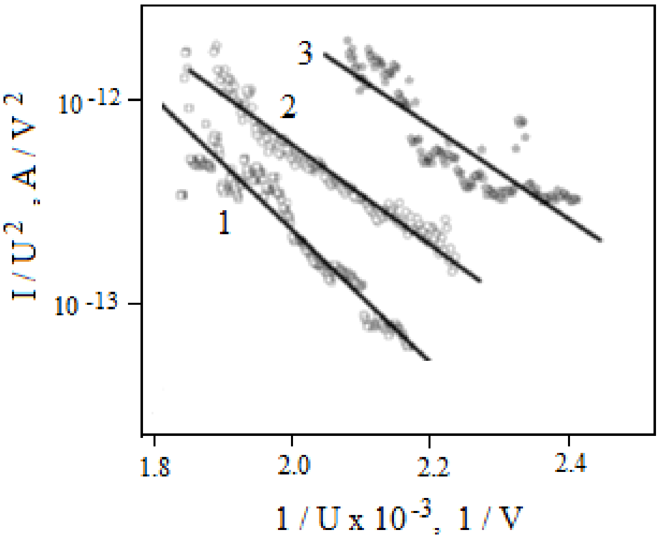


Fig. 11.18. F-N characteristics of the Si tip array coated with the SiO_x(Si) ($x≈0.3$) film: (1) initial array of Si tips; (2) Si tips coated with 100 nm thick annealed SiO_2(Si) film; (3) 10 s etched (the thickness is 75 nm).

The F-N characteristics of SiO_x(Si) ($x \approx 1.2$, 60 nm thick) film deposited onto flat Si substrate by the PECVD method before and after different treatments are shown in Fig. 11.19. Curves 1 and 2 have been taken from Fig. 11.18. As it can be seen from Fig. 11.19, the curve 3 related to the initial SiO_x(Si) ($x \approx 1.2$) has another slope in comparison with the curves 4 and 5 related to the films after electrical conditioning (two different samples with the same condition). The initial film before thermal annealing (curve 3) contained very small concentration of Si inclusions which could be neglected (x was very high). The slope decrease depends only on the work function. The effective emission area and the field transformation factor can be estimated [1, 25, 40, 67-69]. The field transformation factors are practically the same for curves 3 and 4. The slope change of curve 4 can be explained by the work function increase.

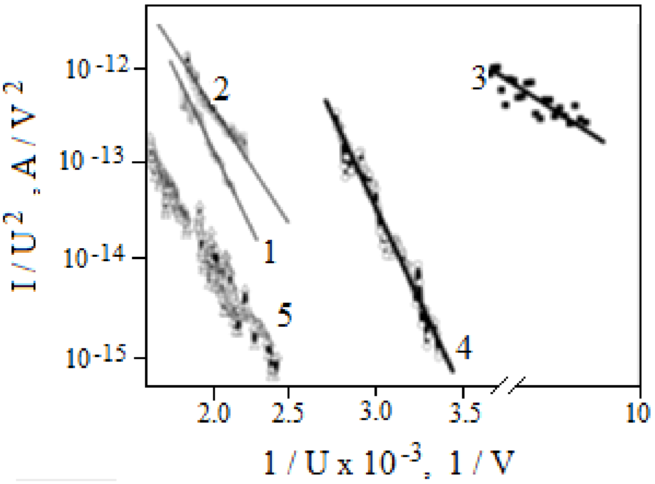


Fig. 11.19. F-N characteristics of the SiO_x(Si) deposited onto Si tip array and flat Si substrate: (1) initial Si tips; (2) SiO_x(Si) ($x \approx 0.3$) film, after annealing (tips); (3) SiO_x(Si) ($x \approx 1.2$), before treatment (flat); (4) and (5) series of experiments SiO_x(Si) ($x \approx 1.2$), after annealing and electrical conditioning (different samples with the same condition, flat).

After annealing and further electrical conditioning of SiO_x(Si) ($x \approx 1.2$) it transforms in SiO_2(Si) with embedded Si inclusions. Therefore, the work function of the created Si conducting channels should be identical to the initial Si tips (curve 1, work function is 4.15 eV). As can be seen from Fig. 11.19 the slope of curve 1 is almost identical to the slope of curve 4. This indicates that the EFE process is identical in both cases, which is only possible when the emission spot areas and the corresponding work functions for both emitting structures are equal. The initial tip and the created conducting channel in the SiO_2(Si) film have different ratios for tip height divided by the tip radius. The estimation of the created conducting channel size can be performed according to [26, 37, 58]

$$\beta^* \approx \frac{h}{r}, \qquad (11.5)$$

where β^* is the local electric field enhancement coefficient, r is the tip radius of conducting channel, and h is the tip height (film thickness), respectively.

To calculate the conducting channel radius, the effective emission area should be estimated [1, 25, 40, 67-69]. This expression can be determined as the sum of all conducting channels with arithmetic mean of channel radius r in assumption that only the top of each tip provides with the field emission as follows:

$$\alpha = \sum_{i=0}^{N} \pi r^2, \qquad (11.6)$$

where $N=n\times S$ is a quantity of tips, n is the tip density of the formation array (1.44×10^8 tips/cm^2), and S is the hole area of emitter-anode spacer (7.854×10^{-3} cm^2).

The effective emission area related to curve 2 (see Fig. 11.19) is found to be 3.35×10^{-8} cm^2. From Eqs. (11.5) and (11.6) one finds that the mean values of conducting channel radius and the electric field enhancement coefficient β^* are 0.97 nm and 62, respectively. The diameter of Si inclusion is equal to ≈2 nm. The obtained result is in good agreement with the sizes of Si inclusions measured by AFM (1-3) nm and the calculated ones (2.3 nm) given in Ref. 26.

11.6. ELECTRON FIELD EMISSION FROM Si NANOWIRES

One-dimensional (1D) structures such as nanotubes, nanobelts, nanorods, and nanowires are ideal candidates for achieving high field emission current density at a low electric field because of their high local electric field enhancement by the high aspect ratio [70-74]. Among various 1D nanostructured materials, carbon nanotubes have attracted extensive efforts [70, 71, 73, 74]. However, field emission from other materials, including Si, ZnO, SiC, etc. is also interesting and being explored [75-77]. The semiconducting properties of Si nanowires (Si-NWs) opens distinct possibilities compared to metallic emitters and carbon nanotubes. Si nanowires exhibit a unique sp^3-bonded crystal structure and a low work function (3.6 eV) [75]. Meanwhile Si has been the backbone of the microelectronics industry for decades and it would be desirable to have Si field emitters to be integrated onto Si substrates along with the driving circuitry. Si nanowires can be synthesized by several methods, and field emission from various Si nanostructures has been reported [75, 78-82]. For Si nanowires grown on Si substrates in Ref. [83] the current density of 1 mA/cm^2 at such low electric field as 3.4 V/μm has been obtained which is comparable with that of carbon nanotubes [73]. The measurements of EFE from Si-NWs arrays show the linear F-N behavior as for metallic emitters. In addition to strong current saturation in EFE due to the bandgap, their properties could be strongly influenced by surface states because of their large surface-to-volume ratio.

The Si nanowires are usually synthesized from silane by a chemical vapor deposition method using Au as catalyst. In Ref. [83] it has been done at a temperature of 480 °C. At the beginning the native SiO_2 layer has been removed from Si substrate in a 5% hydrofluoric acid water solution for 5 min, followed by cleaning in alcohol by ultrasonication, then Au as catalyst was deposited onto the Si wafers by immersing the Si wafers into alcohol solution of hydrogen gold tetrachloride ($HAuCl_4 \cdot 3H_2O$, Aldrich). Then the Au-coated Si wafers were loaded into the growth tube furnace for the growth of Si nanowires. After the Si wafers were loaded, the furnace was pumped down to about 1 mTorr, followed by heating up to 480 °C within 20 min with flowing H_2 gas and Ar gas. As soon as the temperature reached 480 °C, a helium-diluted silane gas (10%) was introduced into the furnace to build up a pressure to start Si nanowire growth. During growth, the pressure was maintained in the range of 590–610 Torr. The typical growth lasts 30 min.

The scanning electron microscope (SEM) and the high-resolution transmission electron microscope (HRTEM) images of Si nanowires (NW) are shown in Fig. 11.20. As it can be seen, the nanowires are very long (at least 100 μm), the diameter is of about 100 nm and it is fairly uniform. The TEM images clearly show that the nanowire has a core and shell structure. The core is about 45 nm in diameter and the shell is about 30 nm thick. The amorphous shell is mainly Si (97%) with a very small amount of oxygen (3%) evidenced by the strong Si peak and weak oxygen peak in the energy dispersive spectroscopy (EDS) spectrum [83].

The *I*-*V* characteristics of emission current density and corresponding F-N plot are shown in Fig. 11.21. A turn-on electric field of 5.5 V/μm was obtained at an emission current density of 0.01 mA/cm^2 for the as-grown Si nanowires, and the highest saturated current density obtained was only 0.03 mA/cm^2 due to probably weak mechanical and electrical contact with the Si wafer. To improve the contact, the same sample was annealed in vacuum at 550 °C for 24 h and measured again. After annealing the turn-on voltage was drastically reduced to 2.0 V/μm and an emission current density of 1 mA/cm^2 was obtained at 3.4 V/μm.

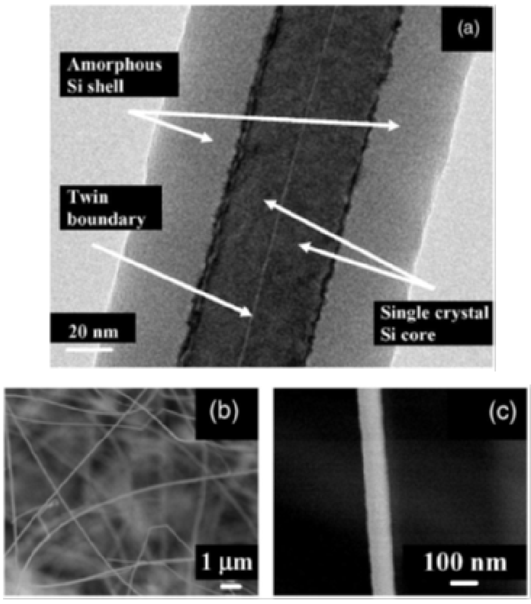


Fig. 11.20. TEM (*a*) and SEM (*b*, *c*) micrographs of Si nanowires grown by low-pressure CVD.

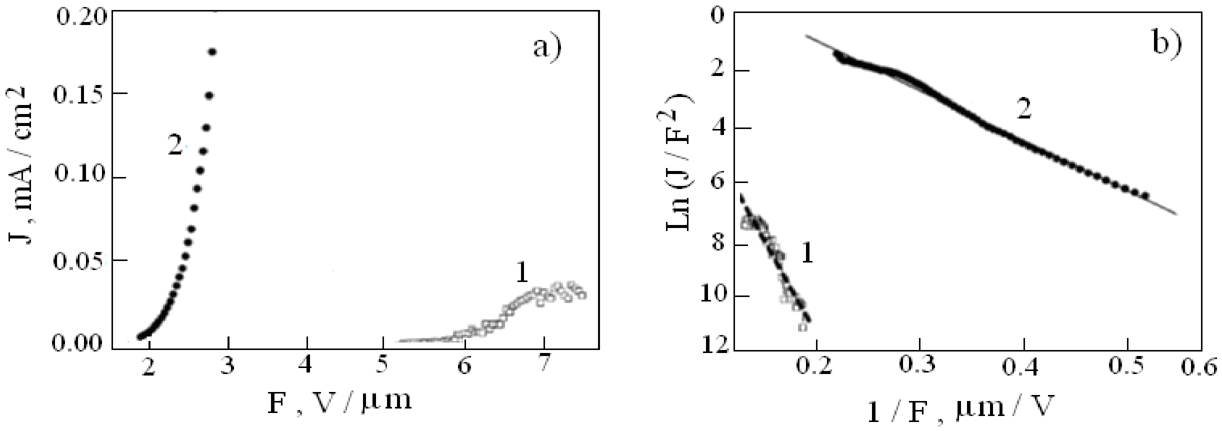


Fig. 11.21. Field emission current density dependencies on electric field (*a*) and Fowler-Nordheim plots (*b*): 1- as-grown Si NW, 2- annealed during 24 h.

To assess the size effect of the Si-NWs on field emission, samples with different nominal diameters have fabricated by high temperature laser ablation method [84] with varying the synthesis parameters. After that the Si-NWs containing paste has been prepared. The turn-on field of the Si-NW emitters with a nominal diameter ~10 nm, ~20 nm, and ~30 nm are 4.5, 13, and 23 V/μm, respectively [75]. These results clearly show that the diameter of Si-NW emitters plays an important role in the field emission properties. The turn-on field and threshold field decrease with decreasing diameter of the Si-NWs. This phenomenon is readily understandable in terms of the field enhancing factor of the Si-NW emitters. As the field enhancing factor is expected to increase with decreasing radius of curvature of the emitting tip, which is in turn related to the diameter of the Si-NWs, it is natural that smaller Si-NWs have larger field enhancing factor and thus smaller turn-on field.

In an attempt to remove the oxide layer and thus enhance the field emission of the Si-NWs, the film made with Si-NWs of ~20 nm in diameter was treated with hydrogen (H_2) plasma using electron cyclotron resonance (ECR) chemical vapor deposition [75]. The ECR was operated at a power of 550 W with pure H_2 of 0.01 Torr. The plasma treatment was made for 1 h. After the H_2 plasma treatment it was found from SEM observations that the Si-NWs became thinner. These observations of EFE indicate that H_2 plasma treatment improved the uniformity of the Si-NW emitters. The uniformity improvement was possibly due to the reduction of the barrier for field emission as a result of the thinner oxide shell of Si-NWs.

The study of the EFE from individual Si-NW allows understanding the surface effects and optimizing the EFE characteristics. In Ref. [85] the individual high crystalline Si-NWs with controlled surface passivation were investigated. The Si-NWs were batch-grown by vapor-liquid-solid using Au catalyst with no intentional doping [86]. Individual Si-NWs were mounted on standard tungsten tips. The measurements were performed in ultra-high vacuum 2×10^{-10} Torr. Quasi-ideal saturation was obtained accompanied by a strong sensitivity to temperature (current increasing up two orders of magnitude) (Fig. 11.22). These strong saturation effects must be associated with the high quality and good passivation of the Si-NWs and can be explained by an unintentional *p*-type doping inherent in the Si-NWs growth process. The curves obtained are predicted by the theory of EFE from semiconductor [87]. These are similar to those in *p-n* junctions in reverse bias where the field penetration induces an internal *p-n* junction in the Si-NWs (Fig. 11.23).

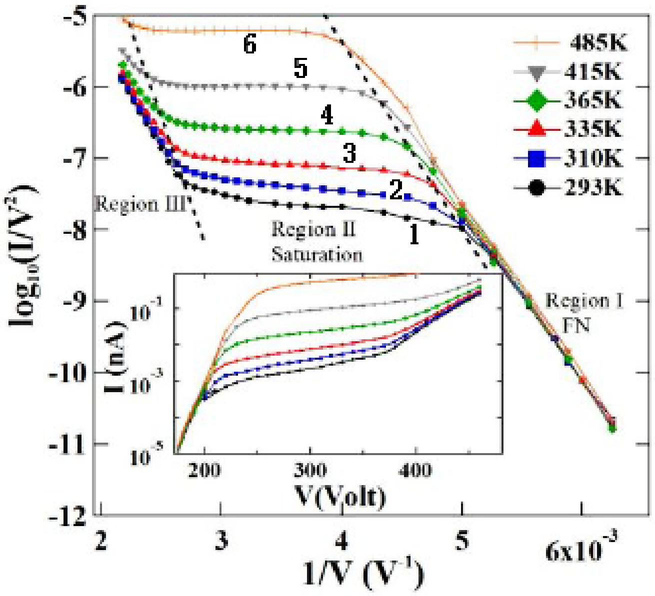


Fig. 11.22. Typical Fowler-Nordheim plots for individual Si-NWs with increasing temperature: region I - a standard F-N emission (I<1 pA), region II - saturation of current with high sensitivity to temperature, and region III - rapid increase of current (1- *T*=293 K, 2- *T*=310 K, 3- *T*=335 K, 4- *T*=365 K, 5- *T*=415 K, 6- *T*=485 K).

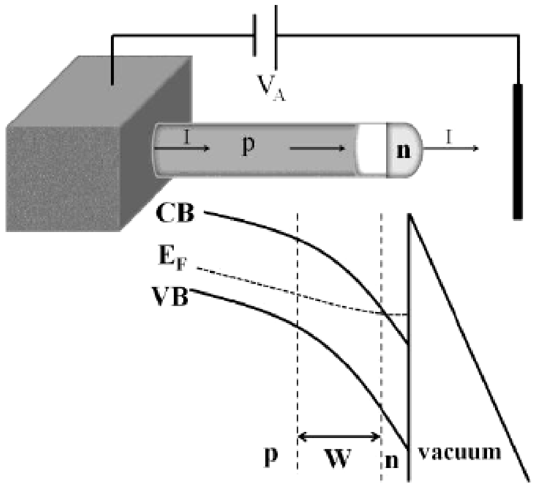


Fig. 11.23. Schematics of the emission geometry and the energy band diagram of the internal *p-n* junction formed in *p*-type nanowire during field emission: a depletion region (*W*) build up behind a pocket of electrons (*n*) formed in the degenerated conduction band.

The role of the surface states was demonstrated by cyclic heating and hydrogen passivation treatments done in situ (Fig. 11.24). This treatment allowed determining that the current saturation was linked to the conducting properties of the Si/SiO$_2$ interface of the Si-NWs. The increasing of the surface states led to suppression of saturation and then to a linear F-N curve (Fig. 11.24).

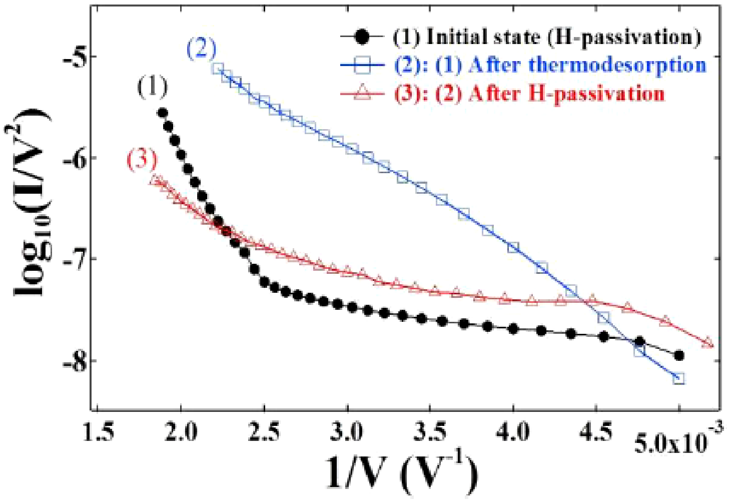


Fig. 11.24. F-N plots with heat and hydrogen passivation: the thermodesorption induced a linear F-N plot due to an increase in the surface states (1 - initial state (H-passivation), 2 - after thermodesorption, and 3 - after H-passivation) [85].

Various types of Si-NW show excellent emission parameters. One-dimensional boron-doped Si nanoparticle chains synthesized in bulk quantity using laser ablating SiO powder mixed with B$_2$O$_3$ powder have the outer diameters of the nanoparticles in the chains around 15 nm. High-resolution transmission electron microscopy showed that the nanoparticles had perfect lattices with 11 nm crystalline core and 2 nm amorphous oxide outer layer while the distance of the inter particles was 4 nm. Field-emission measurement showed that the turn-on field of the Si nanoparticle chains, which was defined as the electric field leading to a current density of 0.01 mA/cm^2, was 6 V/µm. That was much lower than that of undoped Si nanowires (9 V/µm) [78].

Well-aligned quantum silicon nanowires arrays have been synthesized by chemical vapor deposition template method without catalyst. The superior field emission behavior (turn-on field for electron emission 14 V/µm) is believed to originate from the oriented growth and the sharp tips of Si-NWs [82].

The taper-like Si-NWs exhibit a turn-on field of (6.3-7.3) V/µm and a threshold field, defined as the electric field leading to a current density of 10 mA/cm^2, of (9-10) V/µm. The excellent field emission characteristics are attributed to the taper-like geometry of the crystalline Si nanowires [79].

To improve emission parameter the cesiated silicon nanowires grown by the vapor-liquid-solid (VLS) technique have been also prepared and investigated. The average threshold field of cesiated Si NWs was found to be ~7.76±0.55 V/µm and showed a significant improvement over that of as-grown NWs (average threshold field ~11.58 V/µm) [80].

A multistep template replication route was employed to fabricate highly ordered silicon nanotube (Si-NT) arrays, in which annular nanochannel membranes were produced first, and then silicon was deposited into the annular nanochannels by pyrolytic decomposition of silane. Field emission characterization showed that the turn-on field and threshold field for the Si-NT arrays weree about 5.1 V/μm and 7.3 V/μm, respectively [81].

As it was shown, the considered results on the EFE from Si-NWs opened up numerous perspectives for cathodes application, thermal and optical modulation.

11.7. METAL-INSULATOR-METAL EMITTERS

Cold electron emitting devices have been studied intensively for their use in flat panel type displays [88], X-ray tubes [89], vacuum microwave devices [90], electron sources in ultrafast electron microscopy [91], electron beam lithography [92] etc. Spindt-type electron emitters [49, 93] based on field emission from Si and metal cones have been studied. High brightness and the ability to operate at high speed were demonstrated. However, these devices have a complicated structure and they require high vacuum and high supply voltage. Large dispersion of angles of emitted electrons results in a poor resolution. The field emission devices based on carbon nanotubes have addressed some of these problems [94, 95]. However, lifetime is a serious issue for these devices. Planar type electron emitters based on metal – insulator – metal (MIM) or metal – oxide – semiconductor (MOS) structure are very perspective to overcome the above mentioned problems.

Metal-insulator-metal cathodes are attractive for application in vacuum micro- and nanoelectronics because they have some important advantages among others. Significant advantages of MIM emitters are their flat thin film structure. The cathode is less susceptible to surface contamination due to the fact that the emitting material is buried inside, and electrons tunnel through interfacial Schottky barriers instead of surface barriers [96]. Many efforts have been applied to develop this type cathode to practical application [97-102]. The schematic image and energy band diagram of such type of cathode are presented in Fig. 11.25. Typically, the device consists of a thin insulating film (e.g., Al_2O_3) sandwiched between two metal electrodes (e.g., Al and Au). The insulating layer is so thin (several nanometers) that an electron can tunnel through it when an electric field is applied across the layer. As the energy diagram in Fig. 11.25(*b*) shows, the tunneling electrons are injected from the negative electrode (Al, the emitter) through the insulator into the positive electrode (Au, the gate) as hot electrons and are detected as a diode current I_d. A portion of the injected electrons that have kinetic energies larger than the work function of the Au surface can go through Au and emit to vacuum, which is collected as an emission current I_e. However, a majority of the tunneling electrons lose their kinetic energies while they pass through the structure because of scattering events in both the insulator and the gate metal, resulting in a very inefficient emission process with low current transfer ratios ($I_e/I_d < 10^{-3}$). Reducing the thickness of both the insulator and the gate electrode would increase I_e/I_d, but for the insulator, it needs to have a certain thickness to withstand the applied voltage that must be higher than the work function of the gate electrode. The use of low work function materials for the gate electrode would also help to enhance I_e/I_d. In addition, very high current transfer ratio (0.7%) has been reported from the metal-oxide-semiconductor (MOS) structure, a variant of the MIM cathode [103].

It is important to note that stable and reproducible electron emission from MIM and MOS cathodes requires the fabrication of ultrathin (a few nanometers thick) insulating and metal or semiconductor layers, the structures of which need to be sufficiently controlled on an atomic scale. Roughness at the interfaces or the presence of defects and structural inhomogeneity in these thin layers will cause significant fluctuations in emission current and emission nonuniformity. The emission current density from MIM cathodes is also generally too low (~50 μA/cm^2) for practical applications. Although higher currents can be obtained by increasing the electric field applied across the insulating layer, the insulator would degrade quickly under such high fields, and electroforming of the MIM cathodes can occur [104]. This would further result in unstable emission and poor emission uniformity. Further researches allowed to enhance the emission current density by 2 orders of magnitude and obtain the peak current density of 5.8 mA/cm^2 from 30×30 MIM cathode array operating at the gate voltage of 10 V [105].

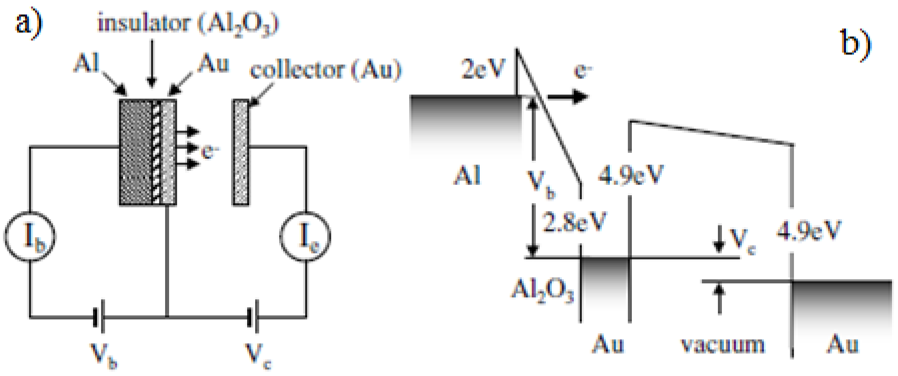


Fig. 11.25. Schematic depicting the MIM emission mechanism: a) MIM structure, b) energy band diagram of MIM structure under applied voltage [102].

The perspective application of MOS cathodes has been proposed in Ref. 106. A MOS tunneling cathode is a promising candidate as a fine electron source, because it has a potential for high current density and pressure insensitivity of emission current and is a flat cathode in contrast to field emission cathode. The emission characteristics of MOS cathodes were demonstrated in Ref. [103,107,108]. In a MOS cathode, when a voltage exceeding the work function of the gate electrode is applied, electrons travel through the conduction band of the oxide and the gate electrode after tunneling through the potential barrier in the oxide, and some of them are emitted into vacuum. Since electrons are easily scattered by phonons and many types of electron traps during traveling in the oxide, the mean-free path of electrons in the conduction band of oxide is very short (0.4-0.7 nm) [108]. Thus, the phase coherence of electron waves is lost in the oxide for a usual MOS cathode. However, if an oxide layer is sufficiently thin and the interface between SiO_2 and gate electrode is abrupt to keep the coherence, resonant tunneling due to the interference of incident and reflected electron waves at the interface, so-called resonant Fowler–Nordheim (F–N) tunneling, occurs in the triangular potential and resonantly tunneling electrons are emitted into vacuum. In a MOS diode with an ultrathin oxide layer the resonant F–N tunneling is reported for the electrons injected into the SiO_2 layer from the gate metal because the interface of Si–SiO_2 is abrupt in an atomic scale [109]. On the contrary, in Ref [106] they have achieved the resonant F-N tunneling due to reflection at the interface between SiO_2 and the gate electrode by preparing an extremely abrupt interface with a polycrystalline Si (poly-Si) gate electrode. Using this MOS structure the author confirmed electron emission based on resonant tunneling.

Schematic drawing of MOS cathode and its fabrication process are shown in Fig. 11.26(*a*) and (*b*), respectively [106]. An *n*-type Si wafer with the carrier concentration of 1×10^{15} cm^{-3} is oxidized to form thick SiO_2 by wet oxidation. After patterning the gate area, the Si surface is oxidized again in dry O_2 as the gate oxide in the MOS cathode. Although this oxide thickness must be thin enough to satisfy the phase coherence of electron waves, the thickness of 6.3 nm has been employed from consideration of breakdown voltage of SiO_2 because the gate voltage larger than the work function of the gate electrode is necessary to extract electrons into vacuum. The phosphorus-doped poly-Si as a gate electrode instead of Al to form a sufficiently abrupt interface in an atomic scale has been used. The phosphorus-doped poly-Si has been prepared by the following processes; amorphous Si is first deposited on the gate oxide by low pressure chemical vapor deposition using Si_2H_6 and PH_3 at a substrate temperature of 550 °C and annealed at 700 °C in dry N_2. The thickness of the poly-Si is chosen to be 20 nm by considering the electron mean-free path of poly-Si of about 4 nm [108] and gate resistance to apply a uniform field on the gate. Finally, an Al electrode is deposited on both sides to reduce the gate resistance and to form an Ohmic contact on the substrate by usual vacuum evaporation. The diode and emission currents have been measured at a pressure of 1.0×10^{-7} Torr. A single cathode with a gate size of 0.25 mm^2 has been used for the experiments (Fig. 11.26).

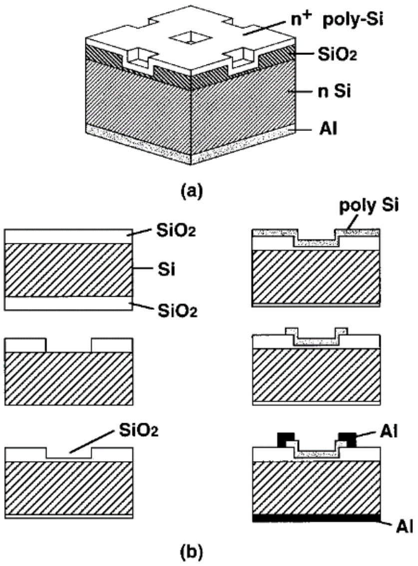


Fig. 11.26. Schematic drawing (*a*) and the fabrication process (*b*) of the MOS cathode [106].

Figure 11.27 shows the band diagram of the MOS cathode, and Fig. 11.28 shows the diode and emission currents (*a*) and the transfer ratio, i.e., the ratio of the emission current to the total current (*b*) of the MOS cathode as a function of the gate voltage. The emission current increases abruptly at 4 V, which corresponds to the work function of the poly-Si gate as shown in Fig. 11.27. This indicates that among the electrons tunneling through the potential barrier only electrons with the energy higher than the work function of the poly-Si gate contribute to the emission current and the rest result in the diode current. The transfer ratio has a peak at a gate voltage of about 4.5 V. This means that a considerable number of electrons travel ballistically through the conduction band of the oxide layer below 4.5 V. On the other hand, electrons lose their energy considerably by scattering in the oxide layer above the gate voltage of 4.5 V, because the traveling distance in the conduction band of SiO_2 becomes longer.

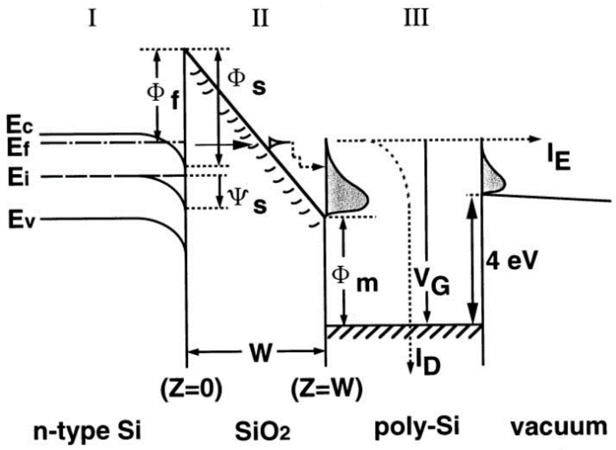


Fig. 11.27. Band diagram of the MOS cathode [106].

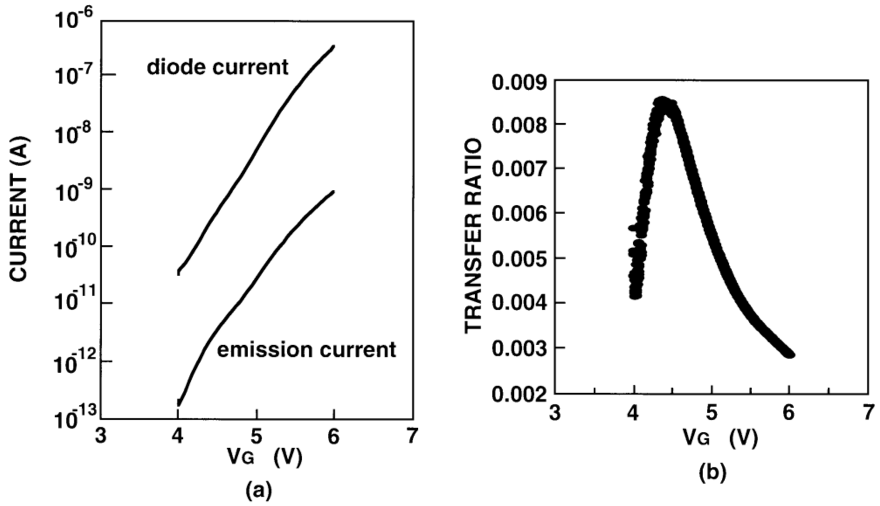


Fig. 11.28. Diode and emission currents (*a*) and the ratio of the emission current to the total current (*b)* of the MOS cathode as a function of the gate voltage [106].

Figure 11.29(*a*) shows the F–N plot of the emission current. The F–N fitting using the Wentzel–Kramers–Brillouin (WKB) approximation is also shown by a straight solid line. The emission current is slightly perturbed from the straight line. To enhance this deviation the ratio of the emission current to the fitting current using the WKB calculation (I_e/I_{WKB}) has been shown in Fig. 11.29(*b*). The oscillatory feature in the emission current becomes clearly visible.

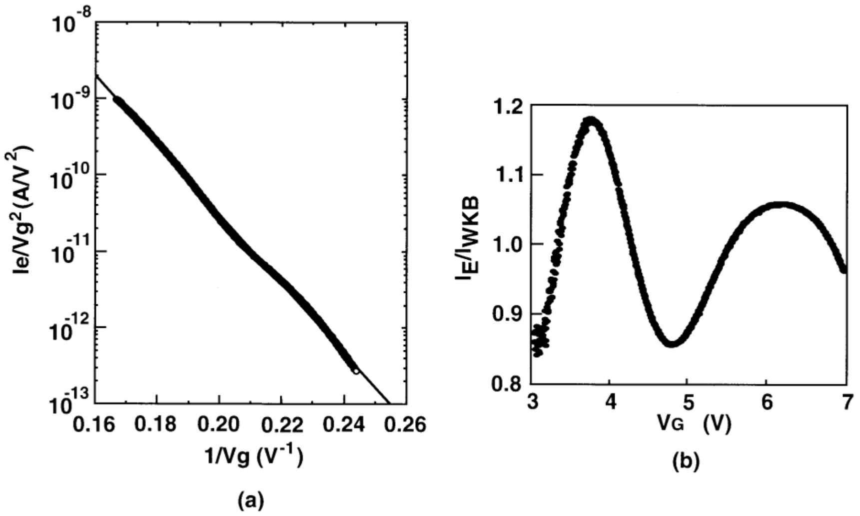


Fig. 11.29. (*a*) F–N plot of the emission current. The F–N fitting using the WKB approximation is also shown by a straight solid line. (*b*) Ratio of the emission current to the fitting current using WKB approximation as a function of the gate voltage [106].

In order to confirm that the oscillation is caused by the interference of electron waves reflected at the SiO_2–poly-Si interface the tunneling probability by solving the Schrodinger equation describing electron tunneling through a triangular potential have been calculated and compared it with that of the WKB approximation [106]. The comparison shows that although the magnitude of tunneling probabilities is almost the same, the exact solution exhibits oscillatory behavior which does not appear in the WKB approximation. This oscillation is due to the interference of the incident and reflected electron waves at the interface between the SiO_2 and the poly-Si gate. As ratio of currents I_e/I_{WKB} and ratio of tunneling probabilities T_e/T_{WKB} for exact solution and solution at WKB approximation show the peaks on applied field dependences. The peak positions between experimental and theoretical results agree reasonably, though the peak position at the high field region shifts slightly to high field in the experiment [106]. This disagreement is probably due to the voltage drop by the relatively large tunneling current in the gate electrode. These results indicate that the peaks in the emission current arise from the resonant effect of electrons tunneling. The smaller amplitude of the oscillation in the experiment compared with that of the calculation indicates that the scattering of electrons is still significant in the conduction band of the SiO_2.

The perspective direction of MOS emitter research and development is connected with using of the low-dimensional structures in or instead of the oxide layer.

A planar type electron emitter with narrow beam dispersion formed by porous Si (PS) has been proposed in Refs. [110, 111]. It has been assumed from detailed experimental analyses of the emission characteristics that the observed cold emission is based on the hot-electron tunneling mechanism [112-115]. Under the biased condition, electrons injected into the PS layer from the substrate are drifted toward the outer surface and reach the top electrode as hot electrons, owing to high electric field in the PS layer. Thus, hot electrons are easily emitted into vacuum through tunneling. The emission characteristics of the PS cold cathode strongly depend on the high-field conduction mode in the PS layer. If the carrier transport in PS is appropriately controlled, therefore, further improvement in the emission efficiency and stability should be obtained. To change and control the conductivity of PS the modification of the PS structure during its preparation by electrochemical etching has been performed [110]. Nanocrystalline PS layers has been formed by photoanodizing heavily doped (0.01–0.03 Ω) n^+-type (100) Si wafers in an ethanol solution of 50% HF (HF:ethanols = 1:1). Three kinds of PS layers (3-10 μm thick) has been formed under the following different conditions of the anodization current as shown in Fig. 11.30: (*i*) normal: conventional galvanostatic anodization, (*ii*) multilayered: periodic modulation of anodization current, and (*iii*) graded-multilayer: periodic modulation during increase of anodization current. In the case of (*ii*) and (*iii*), the structures of PS layers have been controlled by modulating the anodization current such that the multilayered and graded-band structures with high- and low-porosity layers were formed periodically, as indicated in Fig. 11.30 (*b*) and (*c*). In some cases, the PS layers have been treated by rapid thermal oxidation (RTO). Finally, thin Au films (10 nm thick) have been evaporated onto the PS layers and used as top electrodes. The investigations have been also conducted for non-doped polycrystalline Si (poly-Si) films deposited onto n^+-type Si wafers to confirm applicability for large-area devices [111]. The porous poly-Si layer is denoted as PPS.

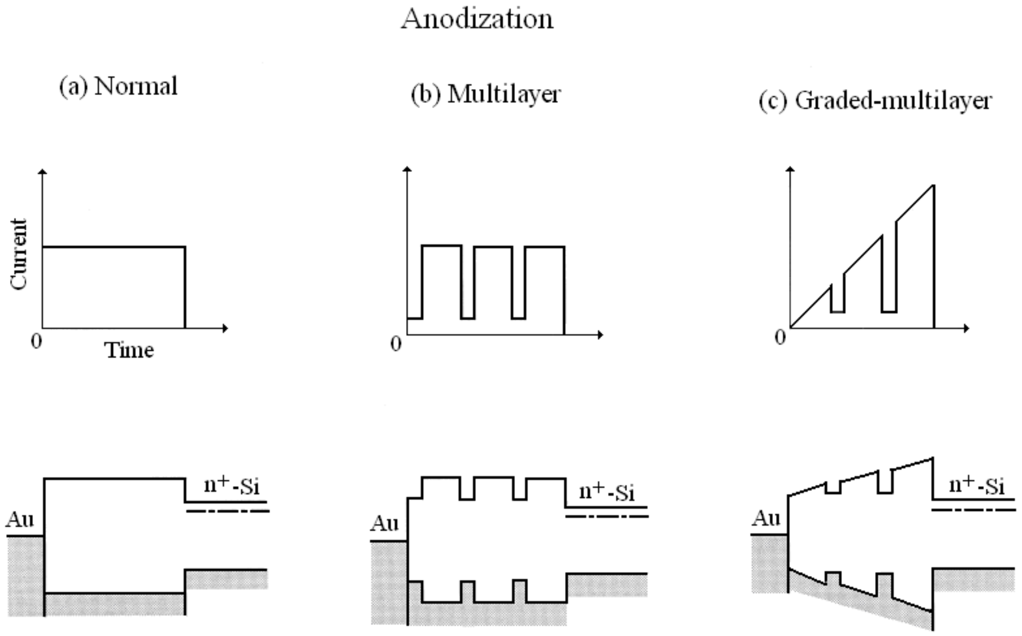


Fig. 11.30. Anodization current modulation and corresponding PS diode band structure: (*a*) normal structure, (*b*) multilayer, (*c*) graded-multilayer [110].

The introduction of RTO into the normal-structured PS diode is very useful for enhancing the electron emission. The emission efficiency, defined as the ratio I_e/I_d, is improved by employing the multilayer PS structure and RTO treatment. In this case, emission current comparable to the normal structure case is obtained, while the diode current is significantly suppressed. As a result, the emission efficiency is enhanced up to 12%.

The improvement in the emission efficiency for the multilayered device is presumably caused by a uniform electric field established in the PS layer, since the low-porosity compact layer which has a high electrical conductivity acts as an equipotential plane. It is self-regulated reformation of the electric field distribution in the PS layer, especially in the outer surface of PS close to the top contact. Consequently, hot electrons, which contribute to emission, are efficiently generated there. The introduced low-porosity layers also act as heat sinks owing relatively high thermal conductivity. This contributes to keeping the diode current constant without thermal effects and to stabilizing the electron emission.

Further efficient and stable cold emission was obtained from the PS diode with a graded-multilayer structure [116] and from the RTO-treated PPS diode with a multilayer structure [117]. The electron emission of the graded-multilayer PS diode is quite uniform. There was the strong correlation between the emission stability and output energy distribution. The energy peak and maximum energy E_{max} are shifted toward higher energy fairly in accordance with an increase in V. These PS results strongly indicate that there is very little serial scattering loss during the drift in the PS layer, and that electrons are emitted quasiballistically. This situation is quite different from that of conventional PS devices, in which the emission fluctuates with spike noises and the corresponding energy distribution curves exhibit broad Maxwellian behavior whose peak energy and width are almost independent of V. It appears that PS in PS diodes with normal or multilayer structure, injected electrons are thermalized after serial scattering losses in PS due to possible potential fluctuations. Thus, some fluctuation is induced in both diode and emission currents.

More efficient quasiballistic emission has been observed in PPS diodes with a multilayer structure. The emission efficiency at $V > 15$ V reaches 1%, and emission current density at $V = 30$ V becomes higher than 200 μA/cm^2. Both the diode and emission currents are quite stable without any signs of spike-like fluctuations. The experimental results of the output electron energy distribution show behavior characteristic of ballistic emission.

The observation of the emission image on the fluorescent screen, the area of the emission shows that the electrons are emitted with a small angular dispersion. Another advantageous feature of this ballistic emitter is the insensitivity of I to ambient pressure. In fact, the emission current of the PPS device, for instance, has showed no changes with Ar gas pressure up to about 1 Pa [110].

For a graded multilayer structure, high electric field exists throughout the PS layer, because the high-porosity wide-gap PS near the substrate owns higher resistivity than that at the outer surface region. In addition, the periodic low-porosity compact layers reform the electric field in the same way as the mentioned above multilayered PS devices. The PS layer consists of a large-number of confined Si nanocrystallites with the same band dispersion as c-Si [118]. The size of Si nanocrystallites in PS is considerably small compared with a drift length of electrons in c-Si. Under a biased condition, major potential drop is produced at the insulating electronic boundaries such as surface skin oxides. In the case of the device with a well-controlled structure such as graded-multilayer, electrons injected into PS can travel for a long distance quasiballistically by multiple tunneling as schematically shown in Fig. 11.31. The effective drift length near the outer surface of PS where a high-field of higher than 10^5 V/cm exists possibly reaches about 1 μm. Thus, emitted electrons can easily obtain the energy gain of 10 eV. The increased drift length in PS under a high electric field was also suggested in Ref. [119] based on time-of-flight measurements. As a result of significantly reduced scattering and energy dissipation, the diode current is also fully stabilized. The observation of quasiballistic emission provides very important insights for the high-field transport and device physics of PS as a nanocrystalline confined system.

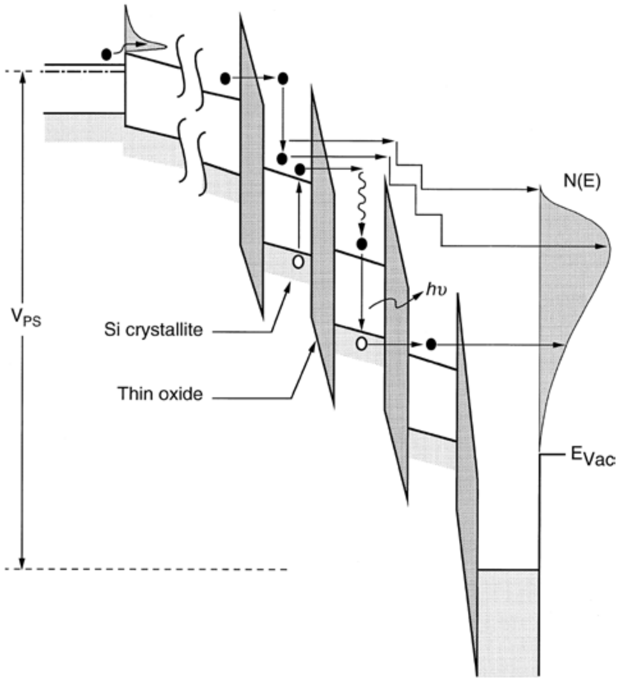


Fig. 11.31. Band diagram showing generation and emission processes of quasiballistic electrons in PS diodes under biased condition [110].

Considered emission device based on PS and PPS does not require the fine control of the position and the size of Si nanostructures in the PS layer because the transport of a large number of electrons is involved. One of the most important mechanisms in these devices is the tunneling process of electrons through a SiO_2 film.

Thermal SiO_2 around nc-Si dots is expected to be more stable than SiO_2 formed by an electrochemical etching process for the PS formation. Thus, an electron emitter using nc-Si dots is expected to deliver higher performance. The attention was focused on nc-Si dots with small spherical shape for a hot electron emission device which behaves like thin film electron emitter devices [120, 121]. The device structure consisted of dots buried in oxide and sandwiched between thin electrode films [122]. It was expected that the enhancement of the electric field around the dots due to the small radius of curvature of the dots should enable a higher yield of hot electrons at the fixed applied voltage. In this device, the problem of size and position distribution of nc-Si dots would not significantly affect the device performance because of macroscopic averaging. An n^+-Si

(0.01 Ω×cm) wafer was used as the substrate and the electron source. The nc-Si dots were deposited onto the substrate at room temperature by plasma decomposition of SiH_4 [122]. This method enabled to control of the growth time of nc-Si dots. A high-resolution transmission electron microscopy (TEM) image of nc-Si dot revealed that the crystalline dot was covered with native oxide. The total thickness of the nc-Si layer varied from 0.1 to 1.5 μm. In the next step the samples were oxidized at 700 °C for 1 h and at 1000 °C for 5 min. The low temperature oxidation was performed for covering the dots with SiO_2, and the high temperature oxidation was for forming a thin oxide layer near the top surface. The cross sectional scanning electron microscopy (SEM) images of the oxidized sample are shown in Fig. 11.32. Near the interface between the Si substrate and the nc-Si layer, many voids were observed between spherically shaped nc-Si dots because nc-Si dots grown in the gas phase had been deposited randomly onto the Si substrate. There was no charge up during the SEM observation because only very thin SiO_2 was formed around each dot. On the other hand, near the top surface, the porosity was small because the molecular volume of SiO_2 was larger than the atomic volume of Si, and the observed image was dark because of charge up effect. This image showed that a thick SiO_2 layer was formed near the top surface. Some samples were planarized by a reflow annealing process. The deep nc-Si layers were not oxidized completely, and nc-Si dots, which were not completely oxidized, remained after the annealing process. For the complete planarization for the thicker nc-Si layer, two optional techniques were added. One was an impurity doping into SiO_2. The other method used for the complete planarization was repeated oxidation and annealing processes.

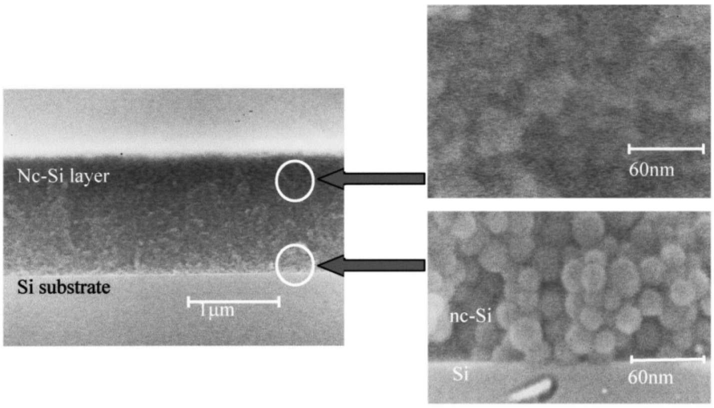


Fig. 11.32. Cross-sectional SEM images of the oxidized samples with nc-Si layer [122].

Finally, an Al ohmic electrode on the backside of Si and a 10-nm-thick Au film on the front surface were formed by electron-beam evaporation. The device structure and the measurement system are shown in Fig. 11.33. The current and emission characteristics of the sample with the nc-Si layer thickness of 0.6 μm are shown in Fig. 11.34. The emission efficiency in this case was 0.8%. It was observed that the samples with a rough surface had larger average currents. In the sample with smoother surface, the diode current and the electron emission increased with the bias voltage without saturation. The devices planarized by the phosphorus diffusion, which includes the oxidation at 850 °C and following phosphorus diffusion at 1150 °C with P_2O_5 for a few minutes, has showed an emission current of 0.25 μA/cm^2 and the high efficiency of 10%. The investigation of the effect of the size of nc-Si dots to the emission characteristic has showed that the smaller nc-Si dots give higher efficiency (for $d_{nc\text{-}Si}$= 6 nm η =0.99% and for $d_{nc\text{-}Si}$= 8 nm η =0.54%) [122].

The mechanism for electron emission includes some steps. First, the electrons are injected from the silicon wafer into nc-Si dots without scattering due to the small size of the dots. The electric field is applied mainly within SiO_2 regions covering the dots. Thus the electrons from nc-Si are accelerated into SiO_2 by the high electric field, allowing ballistic transport through subsequent nc-Si layers. These processes allow electrons to reach the Au electrode with high energy. Electrons with energies larger than the work function of Au can ballistically transport through the Au electrode and reach the collector.

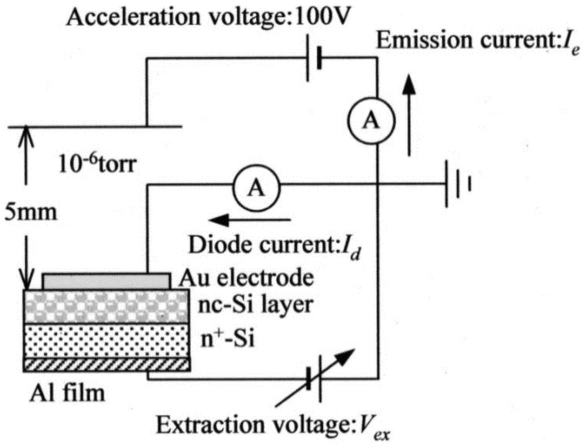


Fig. 11.33. Diagram for the measurement of the electron emitter [122].

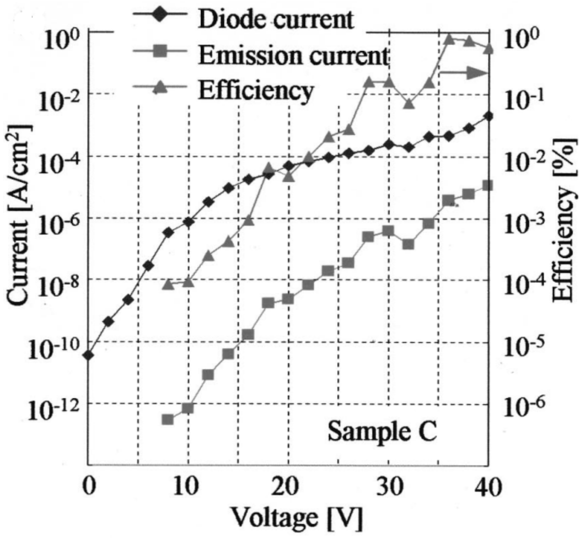


Fig. 11.34. Diode and emission current characteristics as a function of the applied voltage [122].

For more detailed investigation, the energy band profile and the electric field in the nc-Si layer were calculated from the Poisson equation under the estimated structure as shown in Fig. 11.35(*a*). The dielectric constant of Si and SiO_2 were assumed to be 12 and 3.9, respectively, and the quantum confinement effect in nc-Si dots was ignored. Figure 11.35(*b*) shows the energy band profiles of the nc-Si layer with and without a single nc-Si dot. The thickness of the tunnel barrier without dot, t_{ox}, is divided into two thinner SiO_2 layers, t_1 and t_2. Thus, electrons go through the thinner SiO_2 layer labeled as ‘‘A’’ by direct tunneling. Then, accelerated electrons travel through the nc-Si dot with a low scattering rate, followed by a FN tunnel transport through the SiO_2 layer ‘‘B’’ with thinner and lower tunnel barrier than the SiO_2 layer without the nc-Si dot. This means that the electron transport through the SiO_2 layer is enhanced due to the presence of Si dots. Moreover, the electric field around the nc-Si dot is enhanced due to the spherical shape. The nc-Si dots enhance the electron transport density. The calculation of tunnel rates with using the Wentzel–Kramers–Brillouin (WKB) approximation for the 50-nm-thick SiO_2 layer with and without one nc-Si dot showed that the existence of nc-Si increased the tunneling rate by more than 4 of magnitude compared with the cases of without a dot.

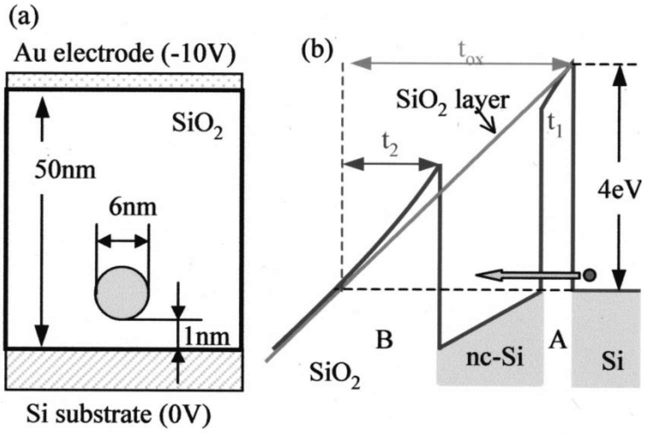


Fig. 11.35. (*a*) Structure of the measured sample, which is estimated from TEM and SEM images, for calculations of the electric field and the potential profile. (*b*) The band diagram around the nc-Si dot. The straight line marked by the label ‘‘SiO_2 layer’’ indicates the band diagram of the SiO_2 layer without the nc-Si dot. ‘‘$t_{1,2}$’’ and ’’t_{ox}’’ indicate the thickness of tunnel barriers for electrons in the conduction band of the Si substrate with and without the nc-Si dot, respectively. The regions labeled as ‘‘A’’ and ‘‘B’’ are the SiO_2 layers between the Si substrate and the nc-Si dot and between the Au film and the nc-Si dot, respectively [122].

Effect of the top electrode. Electrons accelerated in nc-Si dots and the SiO_2 layers are extracted into the vacuum through the top Au electrode. In the Au film, electrons lose their energy due to scattering, and electrons with energy less than the work function of 5eV are absorbed into the Au electrode without emission. The electron mean free path in a metal film, which is determined by electron-electron interaction and electron-phonon interaction, is shorter than that of Si [123, 124]. Although the mean free path depends on the electron energy, the mean free path of hot electrons with energy of higher than 5 eV is estimated to be 5–7 nm [125]. Thus, a thinner Au electrode is desirable. However, a very thin film, less than 10 nm, formed by an evaporation method can hardly cover the surface completely [126, 127]. Especially, the poor cohesion between Au film and SiO_2

makes the formation of very thin Au film difficult. The effect of the Au film with the thickness of more than 10 nm was investigated [122]. As was revealed the number of emitted electrons is proportional to $exp(-t/l_m)$, where t is the Au film thickness, l_m the electron mean free path. The mean free path estimated from the slope of the emission efficiency – Au film thickness characteristic is 5.8 nm. When the slope is extrapolated, the interception at zero thickness is not 100% but 10%. This means that 90% of all electrons reaching the Au film are either reflected quantum mechanically at Au surface or having an energy of less than the work function of the Au film. While electrons go through the conduction band of SiO_2, electrons lose their energy. Under this consideration, polycrystalline (poly-Si) or amorphous (a-Si) Si may be good choices of the electrode films because the work function of 4 eV is smaller and the mean free path is longer than those of metal films. Moreover, a thinner film formed by CVD can be obtained easily, and cohesion is also improved. In Ref. [107] the comparison of the Al top electrode with the a-Si top electrode in the electron emitter has been composed of a few-nm-thick SiO_2 sandwiched between the top Al or a thin a-Si electrode and a Si substrate has been performed. They indicated that the a-Si top electrode enhanced the electron efficiency by more than one order of magnitude. So, replacing the Au top electrode with an *a*-Si film may enhance the efficiency of the electron emitter using nc-Si dots.

The MOS emitters with nc-Si dots have been used in Refs. [128, 129] at investigation of the photoresponse of such cathodes. It is important for formation of pre-modulated electron beam (a train of periodic electron bunches).

Nc-Si MOS cathode arrays having emission area of 50 μm circles on a 500 μm-diameter active area were fabricated on heavily doped *p*-type silicon with the resistivity of 0.3-0.5 Ω×cm, as shown in Fig. 11.36 [129, 130]. Nanocrystalline silicon was deposited by pulse laser ablation (PLA) and surfaces of nc-Si particles were oxidized by an oxygen radical beam exposure during deposition. The Pt gate electrode of 3 nm thicknesses and nanocrystalline silicon layer of about 300 nm were formed. Thick metal for contact pad was deposited on a Pt gate around the active area.

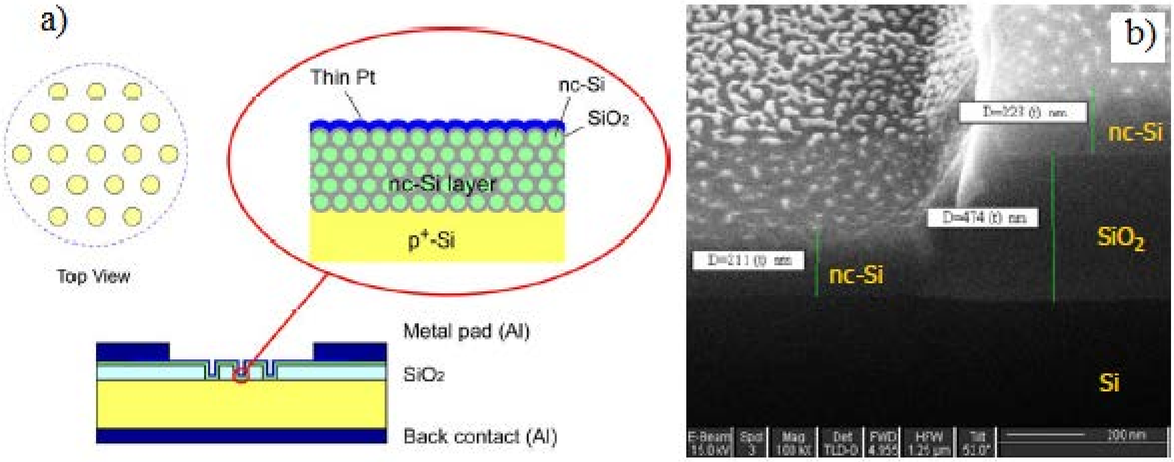


Fig. 11.36. A *p*-type MOS cathode based on nanocrystalline silicon: a) schematic view; b) scanning electron microscopy image [129].

The schematic of experimental setup for optically modulation is shown in Fig. 11.37. The anode was biased at 200 V. The array was irradiated with a He-Ne laser having an optical power of 1 mW and a beam spot of 1 mm in diameter with oblique incidence of 40 degree from front side. The laser pulses were produced by an optical chopper at frequency of 10 Hz.

The optical response of emission current with irradiating He-Ne laser pulses is shown in Fig. 11.38. The gate voltage is set at 14 V. The pulsed emission current is observed corresponding to the pulse frequency and duty, indicating that a modulated electron beam is generated directly from the cathode by He-Ne laser pulse excitation. In this experiment, minimum delay of 10 μs was needed due to the measurement system. Therefore, the disappointing slow responses in the rising and falling of emission are not due to a diffusion process of electrons photogenerated outside the depletion region, but due to the experiment system [128].

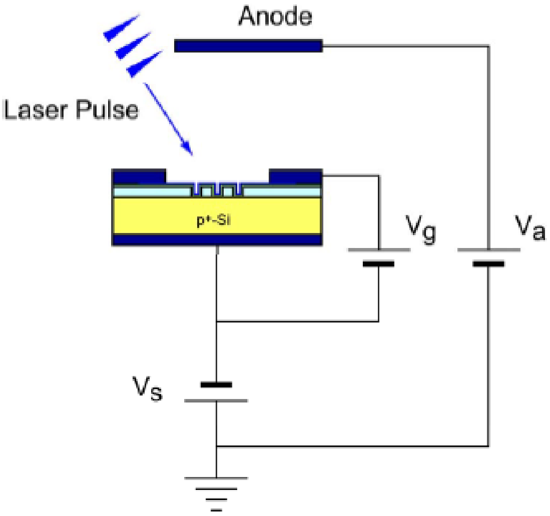


Fig. 11.37. Schematic of experimental set up for optical modulation of nc-Si MOS cathode [128].

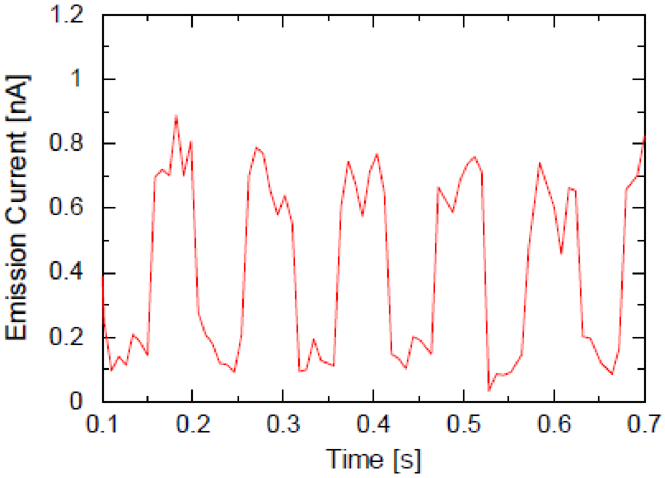


Fig. 11.38. Photoresponse of the emission current under He-Ne laser pulse irradiation at the gate voltage of 14 V [128].

In other experiment [129] the nc-Si MOS cathode was irradiated with 405 nm pulse laser with the repetition frequency of 10 kHz with oblique incidence of 40 degree from front side. The original rise and fall times of the laser pulse were less than 1 ns. The photoresponse produced by nc-Si MOS cathode and reference photodetector is shown in Fig. 11.39.

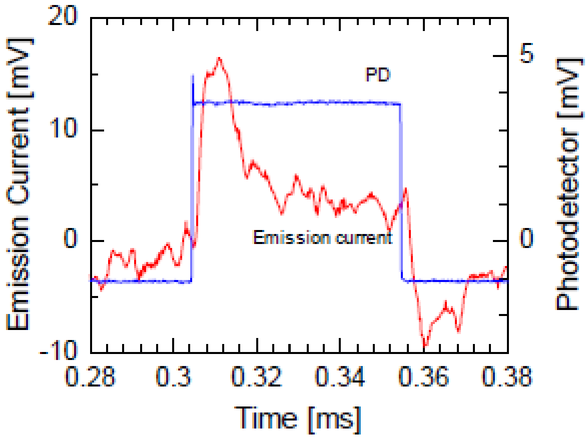


Fig. 11.39. Photoresponse of the emission current and photodetector under illumination of a 405 nm pulse laser [129].

The modulated emission current is found to be synchronized with the laser pulse. Both the rise and fall times of the photoresponse in the cathode device were about 4 μs. This indicates that the measured rise and fall times are determined by those of the amplifier used in this experiment and the real rise and fall of the photoresponse in the cathode device are at least less than 4 μs.

11.8. CONCLUSION

The considered novel Si based cathodes potentially offer the important advantages of high emission parameters, ease in fabrication and low-cost manufacturing. They have unique and interesting field emission characteristics. In many experiments the uncontrolled native oxide on the silicon surface is replaced by coatings with determined properties. Electron field emission from SiO_x(Si) and SiO_2(Si) films containing Si nanocrystals was considered in details. New emission transport mechanisms were revealed in silicon cathodes coated with multilayer films and films with Si nanocrystals. The potential advantages of Metal-Insulator-Metal and Metal-Insulator-Semiconductor emitters were shown. They are integrated in solid state and have plain design. The device applications based on novel Si-based cathodes are being actively researched and explored.

References to Chapter 11.

1. I. Brodie, C.A. Spindt, Adv. in Electronics and Electron Phys. **83**, 1 (1992).

2. K. Derbyshire, Solid State Technol. **11**, 55 (1994).

3. J.E. Jaskie, MRS bulletin. **21**, 59 (1996).

4. E.P. Sheshin. *Surface Structure and Electron Field Emission Properties of Carbon Materials* (MFTI, Moskow, 2001).

5. J.-L. Kwo, M. Yokoyama, J. Vac. Sci. Technol, B. **19**, 1023 (2001).

6. V. Filip, D. Nicolaescu, F. Okuyama, C.N. Flavitac, J. Itoh, J. Vac. Sci. Technol. B. **17**, 520 (1999).

7. Yu.A. Skryshevskii, V.A. Skryshevskii, J. Appl. Phys. **89**, 2711 (2001).

8. A.A. Evtukh, V.G. Litovchenko, R.I. Marchenko, N.I. Klyui, V.G. Popov, V.A. Semenovich, J. de Phys. IV **6**, C5-119 (1996).

9. P.R. Wilshaw, E.C. Boswell, J. Vac. Sci. Technol. B **12**, 662 (1994).

10. M. Takai, M. Jamashita, H. Wille, S. Yura, S. Haribata, M. Ototake, Appl. Phys. Lett. **66**, 422 (1995).

11. I. Kleps, D. Nicolaescu, C. Lungu, G. Musa, C. Bostan, F. Caccavale, Appl. Surf. Sci. **111**, 228 (1997).

12. V. Litovchenko, A. Evtukh, Yu. Kryuchenko, N. Goncharuk, O. Yilmazoglu, K. Mutamba, H.L. Hartnagel, and D. Pavlidis, J. Appl. Phys. **96**, 867 (2004).

13. A. A. Evtukh, V. G. Litovchenko, R. I. Marchenko, and S. Yu. Kydzinovski, J. Vac. Sci. Technol. B **15**, 439 (1997).

14. V. G. Litovchenko, A. A. Evtukh, Yu. M. Litvin, N. M. Goncharuk, and V. E. Chaika, J. Vac. Sci. Technol. B **17**, 655 (1999).

15. V. G. Litovchenko and A. A. Evtukh, Phys. Low-Dimens. Semicond. Struct. **3/4**, 227 (1999).

16. A.A. Efremov, A.A. Evtukh, D.V. Fedin, V.G. Litovchenko, Yu.M. Litvin, Yu.V. Rassamakin, A.V. Sarikov, Phys. Low-Dimens. Semicond. Struct. **1/2**, 65 (2001).

17. A.A. Evtukh, V.G. Litovchenko, Yu.M. Litvin, D.V. Fedin, Yu.V. Rassamakin, A.V. Sarikov, A.G. Chakhovskoi, T.E. Felter, Mater. Res. Soc. Symp. Proc. 685E, D1541 (2001).

18. D.N. Goryachev, L.V. Byelyakov, O.M. Sreseli, Semiconductors **37**, 477 (2003).

19. A.A. Evtukh, V.G. Litovchenko, M.O. Semenenko, N.I. Klyui, E.B. Kaganovich, E.G. Manoilov, Int. J. Nanotechnology, **3**, 89 (2006).

20. L.V. Byelyakov, T.L. Makarova, V.I. Sakharov, I.T. Serenkov, O.M. Sreseli, Semiconductors **32**, 1003 (1998).

21. A.A. Evtukh, E.B. Kaganovich, E.G. Manoilov, M.O. Semenenko, Semiconductors, **40**, 175 (2006).

22. A. Evtukh, H. Hartnagel, V. Litovchenko, O. Yilmazoglu, Materials Science and Engineering, A **353**, 27 (2003).

23. A.F. Kravchenko, V.N. Ovsyuk, *Electron Processes in Solid-State Low-Dimensional Structures* (Novosibirsk, 2000) (in Russian).

24. V.G. Litovchenko, Phys. Rev. B. **65**, 1531 (2002).

25. D. V. Branston and D. Stephani, IEEE Trans. Electron Devices **ED-38**, 2329 (1991).

26. A. A. Evtukh, V. G. Litovchenko, and M. O. Semenenko, J. Vac. Sci. Technol. B **24**, 945 (2006).

27. Yang, K. K. Chin, and R. B. Marcus, IEEE Trans. Electron Devices **ED-38**, 2373 (1991).

28. A. A. Evtukh, V. G. Litovchenko, Yu. M. Litvin, S. Yu. Kudzinivsky, and D. V. Fedin, Phys. Low-Dimens. Semicond. Struct. **1/2**, 21 (2001).

29. V. Ya. Bratus', V. A. Yukhimchuk, L. I. Berezhinsky, M. Ya. Valakh, I. P. Vorona, I. Z. Indutnyi, T. T. Petrenko, P. E. Shepeliavyi, and I. B. Yanchuk, Semiconductors, **35**, 821 (2001).

30. A.A. Evtukh, I.Z. Indutnyy, I.P. Lisovskyy, Yu.M. Litvin, V.G. Litovchenko, P.M. Lytvyn, D.O. Mazunov, Yu.V. Rassamakin, P.E. Shepeliyavyi, Semicond. Phys., Quantum and Optoelectron. **6**, 32 (2003).

31. D. J. DiMaria, D. W. Dong, C. Falcony, T. N. Theis, J. R. Kirtley, J. C. Tsang, D. R. Young, F. L. Pesavento, and S. D. Brorson, J. Appl. Phys. **54**, 5801 (1983).

32. V. Ichizli, H. Mimura, H.L. Hartnagel, Proc. 12th Int. Vac. Microelectron. Conf. (Darmstadt, Germany, July 6-9, 1999), 336, 1999.

33. T. Ifuku, M. Otobe, A. Itoh, S. Oda, Jpn. J. Appl. Phys. **36**, 4031 (1997).

34. A. Evtukh, V. Litovchenko, M. Semenenko, O. Steblova, V. Yukhimchuk, O. Yilmazoglu, H. Hartnagel, and H. Mimura, Proc. 25th Int. Vac. Microelectron. Conf. (Jeju, Korea, July 6-13, 2012), 100, 2012.

35. M. Ehbrecht, F. Huisken, Phys. Rev. B. **59**, 2975 (1999).

36. H. Hofmeister, F. Huisken, B. Kohn, Eur. Phys. J.: D. **9**, 137 (1999).

37. T. Utsumi, *IEEE Trans. Electron. Dev.* 38, 2276 (1991).

38. A.A. Evtukh, E.B. Kaganovich, V.G. Litovchenko, Yu.M. Litvin, D.V. Fedin, E.G. Manoilov, S.V. Svechnikov, Mater. Sci. Eng., C **19**, 401 (2002).

39. A.A. Evtukh, E.B. Kaganovich, V.G. Litovchenko, Y.M. Litvin, D.V. Fedin, E.G. Manoilov, S.V. Svechnikov, Semicond. Phys., Quantum Electron. Opt. **3**, 474 (2000).

40. V.G. Litovchenko, A.A. Evtukh, R.I. Marchenko, N.I. Klyui, V.A. Semenovich, J. Micromech. Microeng. **7**, 1 (1997).

41. A. Uhlir, Bell Syst. Tech. J. **35**, 333 (1956).

42. D. R. Turner, J. Electrochem. Soc. **105**, 402 (1958).

43. B. Hamilton, Semicond. Sci. Technol. **10**, 1187 (1995).

44. A. G. Cullis, L. T. Canham, and P. D. J. Calcott, J. Appl. Phys. **82**, 909 (1997).

45. O. Bisi, S. Ossicini, and L. Pavesi, Surf. Sci. Rep. **38**, 1 (2000).

46. E. C. Boswell, M. Huang, G. D. W. Smith, and P. R. Wilshaw, J. Vac. Sci. Technol. B **14**, 1895 (1996).

47. J. R. Jessing, D. L. Parker, and M. H. Weichold, J. Vac. Sci. Technol. B **14**, 1899 (1996).

48. M. Tejashree, S. V. Bhave, and S. V. Bhoraskar, J. Vac. Sci. Technol. B **16**, 2073 (1998).

49. C. A. Spindt, *J. Appl. Phys.* 39, 3504 (1968).

50. A. A. Evtukh, V. G. Litovchenko, R. I. Marchenko, N. I. Klyui, V. A. Semenovich, and C. Nelep, J. Vac. Sci. Technol. B **14**, 2130 (1996).

51. J. Robertson, Mater. Sci. Eng. **37**, 129 (2002).

52. A. A. Evtukh, H. Hartnagel, V. G. Litovchenko, M. O. Semenenko, and O. Yilmazoglu, Semicond. Sci. Technol. **19**, 923 (2004).

53. A. Szekeres, T. Nikolova, A. Paneva, A. Cziraki, Gy. Kovacs, I. Lisovskyy, D. Mazunov, I. Indutnyy, and P. Shepeliavyi, J. Optoelectron. Adv. Mater. **7**, 1383 (2005).

54. A. A. Evtukh, V. Ya. Rassamakin, and Yu. V. Rassamakin, Mater. Sci. **10**, 40 (2001).

55. A. Evtukh, O. Bratus’, T. Gorbanyuk, and V. Ievtukh, Phys. Stat. Sol. (c) **5**, 3663 (2008).

56. O. L. Bratus’, A. A. Evtukh, V. A. Ievtukh, and V. G. Litovchenko, J. Non-Cryst. Solids **354**, 4278 (2008).

57. M. O. Semenenko, A. A. Evtukh, O. Yilmazoglu, H. Hartnagel, and D. Pavlidis, J. Appl. Phys. **107**, 013702 (2010).

58. A. Evtukh, V. Litovchenko, M. Semenenko, O. Yilmazoglu, K. Mutamba, H. L. Hartnagel, and D. Pavlidis, Semicond. Sci. Technol. **21**, 1326 (2006).

59. V. A. Dan’ko, I. Z. Indutnyy, V. S. Lysenko, I. Yu. Maidanchul, V. I. Min’ko, A. N. Nazarov, A. S. Tkachenko, and P. E. Shepeliavyi, Semiconductors, **39**, 1197 (2005).

60. M. S. Dunaevskii, J. J. Grob, A. G. Zabrodskii, R. Laiho, and A. N. Titkov, Semiconductors, **38**, 1254 (2004).

61. D. R. Wolters and J. J. Van der Schoot, Philips J. Res. **40**, 115 (1985).

62. D. R. Wolters and J. J. Van der Schoot, Philips J. Res. **40**, 137 (1985).

63. D. R. Wolters and J. J. Van der Schoot, Philips J. Res. **40**, 164 (1985).

64. G. Dearnaley, A. M. Stoneham, and D. V. Morgan, Rep. Prog. Phys. **33**, 1129 (1970).

65. G. Revesz, Phys. Status Solidi A **57**, 235 (1980).

66. H. Bouridah, F. Mansour, R. Mahamdi, N. Bounar, and P. Temple-Boyer, Phys. Status Solidi A **204**, 2347 (2007).

67. C. A. Spindt, I. Brodie, L. Humphrey, and E. R. Westerberg, J. Appl. Phys. **47**, 5248 (1976).

68. V. V. Zhirnov, C. Lizzul-Rinne, G. J. Wojak, R. C. Sanwald, and J. J. Hren, J. Vac. Sci. Technol. B **19**, 87 (2001).

69. A. Modinos, *Field, Thermionic, and Secondary Electron Emission Spectroscopy,* (Plenum, New York, 1984, p. 320).

70. W. A. de Heer, A. Chatelain, and D. Ugarte, Science, **269**, 1179 (1995).

71. Y. B. Li, Y. Bando, D. Golberg, and K. Kurashima, Appl. Phys. Lett. **81**, 5048 (2002).

72. Y. Tu, Z. P. Huang, D. Z. Wang, J. G. Wen, and Z. F. Ren, Appl. Phys. Lett. **80**, 4018 (2002).

73. G. Z. Yue, Q. Qiu, B. Gao, Y. Cheng, J. Zhang, H. Shimoda, S. Chang, J. P. Lu, and O. Zhou, Appl. Phys. Lett. **81**, 355 (2002).

74. C. S. Hsieh, G. Wang, D. S. Tsai, R. S. Chen, and Y. S. Huang, Nanotechnology, **16**, 1885 (2005).

75. F. C. K. Au, K. W. Wong, Y. H. Tang, Y. F. Zhang, I. Bello, and S. T. Lee, Appl. Phys. Lett. **75,** 1700 (1999).

76. D. Banerjee, S. H. Jo, and Z. F. Ren, Adv. Mater. (Weinheim, Ger.) **16**, 2028 (2004).

77. J. Zhou, L. Gong, S. Z. Deng, J. Chen, J. C. She, N. S. Xu, R. Yang, and Z. Wang, Appl. Phys. Lett. **87**, 223108 (2005).

78. Y. H. Tang, X. H. Sun, F. C. K. Au, L. S. Liao, H. Y. Peng, C. S. Lee, S. T. Lee, and T. K. Sham, Appl. Phys. Lett. **79**, 1673 (2001).

79. Y. L. Chueh, L. J. Choua, S. L. Cheng, J. H. He, W. W. Wu, and L. J. Chen, Appl. Phys. Lett. **86**, 133112 (2005).

80. N. N. Kulkarni, J. Bae, C. K. Shih, S. K. Stanley, S. S. Coffee, and J. G. Ekerdt, Appl. Phys. Lett. **87**, 213115 (2005).

81. C. Mu, Y. X. Yu, W. Liao, X. S. Zhao, D. S. Xu, X. H. Chen, and D. P. Yu, Appl. Phys. Lett. **87**, 113104 (2005).

82. M. Lu, M. K. Li, L. B. Kong, X. Y. Guo, and H. L. Li, Composites, Part B **35**, 179 (2004).

83. B. Zeng, G. Xiong, S. Chen, S. H. Jo, W. Z. Wang, D. Z. Wang, and Z. F. Ren, Appl. Phys. Lett. **88**, 213108 (2006).

84. Y. F. Zhang, Y. H. Tang, N. Wang, D. P. Yu, C. S. Lee, I. Bello, and S. T. Lee, Appl. Phys. Lett. **72**, 1835 (1998).

85. M. Choueib, R. Martel, S.C. Cojocaru, A. Ayari, P. Vincent, and S.T. Purcell, in Proc. of IVNC 2012, p.154-155.

86. E. Lefeuvre, K.H. Kim, Z.B. He, J.-L. Maurice, M. Châtelet, D. Pribat, C.S. Cojocaru, Thin Solid Films, **519**, 4603 (2011).

87. L. M. Baskin, O. I. Lvov and G. N. Fursey, Phys. Stat. Sol. (b), **47**, 49 (1971).

88. M. Nagao, IEEJ Trans. Electric. & Electron. Engin. **1**, 171 (2006).

89. A. Haga, S. Senda, Y. Sakai, Y. Mizuta, S. Kita, and F. Okuyama, Appl. Phys. Lett. **84**, 2208 (2004).

90. Y. Neo, Y. Suzuki, K. Sagae, H. Shimawaki, and H. Mimura, J. Vac. Sci. Technol. B **23**, 840 (2005).

91. W. E. King, G. H. Campbell, A. Frank, B. Reed, J. F. Schmerge, B. J. Siwick, B. C. Stuart, and P. M. Weber, J. Appl. Phys, **97**, 111101 (2005).

92. Y. Tanaka, H. Miyashita, E. Tomono, M. Esashi, and T. Ono, Tech. Digest of 24 th IVNC, 2011, p.76-77.

93. I. Brodie and C. A. Spindt, Appl. Surf. Sci. **2**, 149 (1979).

94. S. C. Lim, H. J. Jeong, Y. S. Park, D. S. Bae, Y. C. Choi, Y. M. Shin, W. S. Kim, K. H. An, and Y. H. Lee, J. Vac. Sci. Technol. A **19**, 1786 (2001).

95.Y. C. Choi, Y. W. Jin, H. Y. Kim, S. J. Lee, M. J. Yun, J. H. Kang, Y. S. Choi, N. S. Park, N. S. Lee, and J. M. Kim, Appl. Phys. Lett. **78**, 1547 (2001).

96. W. Zhu, P.K. Baumann, C.A. Bower, *Novel Cold Cathode Materials*, in *Vacuum Microelectronics*, Ed. W. Zhu, (John Wiley & Sons, New York, 2001).

97. C. A. Mead, J. Appl. Phys. **32**, 646 (1961).

98. Y. Kumagai, K. Kawarada, and Y. Shibata, Jpn. J. Appl. Phys. **6**, 290 (1967).

99. J. G. Simmons, J. Appl. Phys. 34, 1793 (1963).

100. K. Ohta, J. Nishida, and T. Hayashi, Jpn. J. Appl. Phys. **7**, 784 (1968).

101. T. Kusunoki, M. Suzuki, S. Sasaki, T. Yaguchi, and T. Aida, Jpn. J. Appl. Phys. **32**, L1695 (1993).

102. H. Adachi, J. Vac. Sci. Technol. B **14**, 2093 (1996).

103. K. Yokoo, H. Tanaka, S. Sato, J. Murota, and S. Ono, J. Vac. Sci. Technol. B **11**, 429 (1993).

104. T. W. Hickmott, Thin Solid Films, **9**, 431 (1972).

105. T. Kusunoki and M. Suzuki, IEEE Trans. Electron Devices, **47**, 1667 (2000).

106. H. Mimura, Y. Abe, J. Ikeda, K. Tahara, Y. Neo, H. Shimawaki, and K. Yokoo, J. Vac. Sci. Technol. B **16**, 803 (1998).

107. K. Yokoo, S. Sato, G. Koshita, I. Amano, J. Murota, and S. Ono, J. Vac. Sci. Technol. B 12, 801 (1994).

108. K. Yokoo, G. Koshita, S. Hanzawa, Y. Abe, and Y. Neo, J. Vac. Sci. Technol. B **14**, 2096 (1996).

109. J. Maserjian and N. Zamani, J. Appl. Phys. **53**, 559 (1982).

110. N. Koshida, X. Sheng, T. Komoda, Appl. Surf. Sci. **146**, 371 (1999).

111. T. Komoda, X. Sheng, and N. Koshida, J. Vac. Sci. Technol. B **17**, 1076 (1999).

112. N. Koshida, T. Ozaki, X. Sheng, H. Koyama, Jpn. J. Appl. Phys., Part 2, **34**, L705 (1995).

113. X. Sheng, T. Ozaki, H. Koyama, N. Koshida, T. Yoshikawa, M. Yamaguchi, K. Ogasawara, Thin Solid Films, **297**, 314 (1997).

114. X. Sheng, H. Koyama, N. Koshida, S. Iwasaki, N. Negishi, T. Chuman, T. Yoshikawa, K. Ogasawara, J. Vac. Sci. Technol. B **15**,1661 (1997).

115. X. Sheng, H. Koyama, N. Koshida, J. Vac. Sci. Technol. B **16**, 793 (1998).

116. X. Sheng, N. Koshida, Mater. Res. Soc. Symp. Proc. **509**, 193, (1998).

117. T. Komoda, X. Sheng, N. Koshida, Mater. Res. Soc. Symp. Proc. **509**, 187 (1998).

118. Y. Suda, K. Obata, N. Koshida, Phys. Rev. Lett. **80**, 3559 (1998).

119. O. Klima, P. Hlinomaz, A. Hospodkova, J. Oswald, J. Kocka, J. Noncryst. Sol. **164/166**, 961 (1993).

120. C. A. Meed, J. Appl. Phys. **32**, 646 (1961).

121. J. Cohen, Appl. Phys. Lett. **1**, 61 (1962).

122. K. Nishiguchi, X. Zhao, and S. Oda, J. Appl. Phys. **92**, 2748 (2002).

123. D. J. Bartelink, J. L. Moll, and N. I. Meyer, Phys. Rev. **130**, 972 (1963).

124. S. M. Sze, J. L. Moll, and T. Sugano, Solid-State Electron. **7**, 509 (1964).

125. M. Heiblum, Solid-State Electron. **24**, 343 (1981).

126. H. L. Caswell and Y. Budo, J. Appl. Phys. **35**, 238 (1964).

127. J. L. Robins and J. N. Rhodin, Surf. Sci. **1**, 346 (1964).

128. H. Shimawaki, Y. Neo, H. Mimura, K. Murakami, F. Wakaya, and M. Takai, Tech. Digest of the 24st IVNC, 2011, pp. 220-221.

129. H. Shimawaki, Y. Neo, H. Mimura, K. Murakami, F. Wakaya, and M. Takai, Tech. Digest of the 25st IVNC, 2012, pp. 324-325.

130. H. Shimawaki, Y. Neo, H. Mimura, K. Murakami, F. Wakaya, and M. Takai, Tech. Digest of the 21st IVNC, 2008, pp. 128-129.

Chapter 12. Carbon Nanotubes In 3-D Structured Electrodes

12.1. Introduction

12.2. Synthesis approaches

- 12.2.1. 3D CNT Structure
- 12.2.2. 3D Graphene Structures
- 12.2.3. CVD Grown3D Graphene
- 12.2.4. 3D CNT-Graphene Hybrids

12.3. Functionalization

12.4. Applications

- 12.4.1. Fuel Cells
- 12.4.2. Biofuel Cells
- 12.4.3. Supercapacitors
- 12.4.4. Lithium Ion Batteries
- 12.4.5. Integrated Self-powering Systems

12.5. Conclusion and Perspectives

References to Chapter 12.

Chapter 12. CARBON NANOTUBES IN 3-D STRUCTURED ELECTRODES

12.1. INRODUCTION

In the past decades, carbon nanotubes (CNTs) and graphene have changed the landscape of many fields in science and technology, including energy storage and conversion [1 - 6]. Although differing in structure, synthetic routes, and physicochemical properties, the two carbon allotropes share common merits for energy applications, such as good electrical conductivity, high-specific surface area, excellent mechanical strength, and flexibility. In frequent cases, integration of 1D CNT and 2D graphene materials exerts synergistic effects. Despite the great success achieved, further extending their potential, to a great extent, depends on the possibility to assemble these nanocarbon materials into macroscopic 3D architectures while preserving the intrinsic properties of individual building components.

The electrochemical electrodes used in energy devices or for other purposes are conventionally planar. Electrodes based on 3D porous nanocarbon structures can greatly improve the capacity and efficiency because of largely expanded working volume, multiplexed conduction network, 3D interfacing, or intercalation with other system components (e.g., electrolytes, reactants). However, it is non-trivial to assemble CNTs and graphene sheets into 3D architectures due to their small size as well as stacking or aggregation caused by strong hydrophobic and pi–pi interactions. In this chapter, we will consider the latest development on synthesis of macroporous CNT/graphene structures and their applications in energy devices.

12.2. **SYNTHESIS APPROACHES**

The 3D porous CNT and graphene structures can be readily fabricated from CNT/graphene solvent dispersions via self-assembly or substrate-based deposition, which are cost-effective and easily scalable. On the other hand, chemical vapor deposition (CVD) is able to produce 3D porous CNT/grapheme monoliths, with high quality and well-defined architectures. Some other approaches have also been recently developed.

12.2.1. 3D CNT Structure

Direct deposition of dispersed CNTs onto a 3D substrate is simple and effective to construct porous architectures. For example, sponge with continuous 3D surface and good mechanical flexibility can serve as a supporting substrate. CNT-sponge composites fabricated *via* dip-coating method are highly porous (98%, pore size 200–500mm) with large specific surface area of ~10^4 m^2 m^{-3} [7]. The thin CNT coating layer provides the composite a conductance of ~1 S cm^{-1}. Without use of any template, self-assembled CNT foam has been made by a low temperature chemical fusion method (Fig. 12.1) [8]. Meso- /macroporosity of the CNT foam can be tuned by the amount of ammonia carbonate, which acts as the pore former. The as-synthesized CNT foam can withstand a stress pressure as high as 1.39 MPa. Catalytic growth of CNT forests via CVD method is advantageous in well controlling the length, quality, and density of CNTs [9, 10]. Fe catalysts on the surface of Ni foam for the growth of 3D randomly entangled CNTs has been sputtered in [11]. 3D sponge-like N-doped CNT architecture, employing the mixture of ferrocene, thiophene, and pyridine as CVD precursors has been synthesized in work [12]. The N and S are proposed to synergistically promote the formation of “elbow” and “welded” junctions between CNTs. In addition, the diameter of N-doped multi-walled CNTs can be easily controlled by modulating thiophene concentration, which in turn influences the mechanical and electrical properties of N-doped MWCNT sponge.

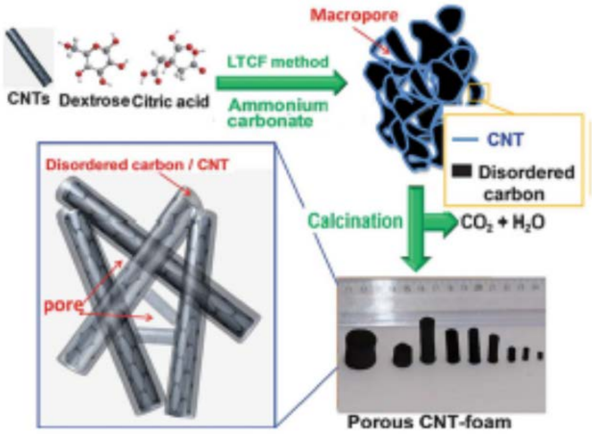


Fig. 12.1. |Illustration of a low temperature chemical fusion method for the synthesis of porous CNT-foam composites [8].

12.2.2. 3D Graphene Structures

Graphene hydrogels and aerogels. The hydrophilic edges and hydrophobic basal planes of graphene oxide (GO) render it to be amenable to solution-based processes [13]. Direct gelation of GO dispersion into porous bulk structures can be achieved by freeze-drying and hydrothermal methods [14 – 16]. Recently, a modified freeze casting process has been devised [17]. They found that the microstructures and properties of the resulting porous graphene aerogel can be widely adjusted by the freezing temperature. Liquid nitrogen freezing confers the as-synthesized aerogel an isotopic and highly elastic structure. The authors of work [1] covalently cross-linked GO sheets into a monolith by conjugating the epoxy groups on GO with amine groups of a gelating polymer *via* ring opening reaction [17]. The obtained GO hydrogel is highly elastic and can be easily converted into reduced GO (rGO) hydrogel *via* thermal treatment. Without using any linkage additives or chemical reducing agents, substrate-assisted spontaneous reduction and assembly of GO sheets into 3D network on various conductive substrates has been demonstrated [18] (Fig. 12.2A). The 3D graphene structures have also been obtained by repetitive dip-coating GO sheets into sponge, followed by GO reduction and removal of sponge by heating (Fig. 12.2B) [19]. Interestingly, a chemical unzipping approach to convert CNT sponge into graphene nanoribbon aerogel as has been reported [20] (Fig. 12.2C). The resulting aerogel inherits the 3D network and high porosity of CNT sponge, and gains enhanced surface area and chemical functionality. It has also been shown that further hydrazine reduction helps the recovery of mechanical properties of graphene nanoribbon aerogel impaired by unzipping process. Graphene aerogel with hierarchical porous structure has also been fabricated using a so-called "sugar-blowing" approach by carbonization of a mixture of glucose and ammonia chloride (Fig. 12.3A) [21]. Such ultralight (3.0 mg cm^{-3}) structure possesses a good mechanical strength and is highly compressible.

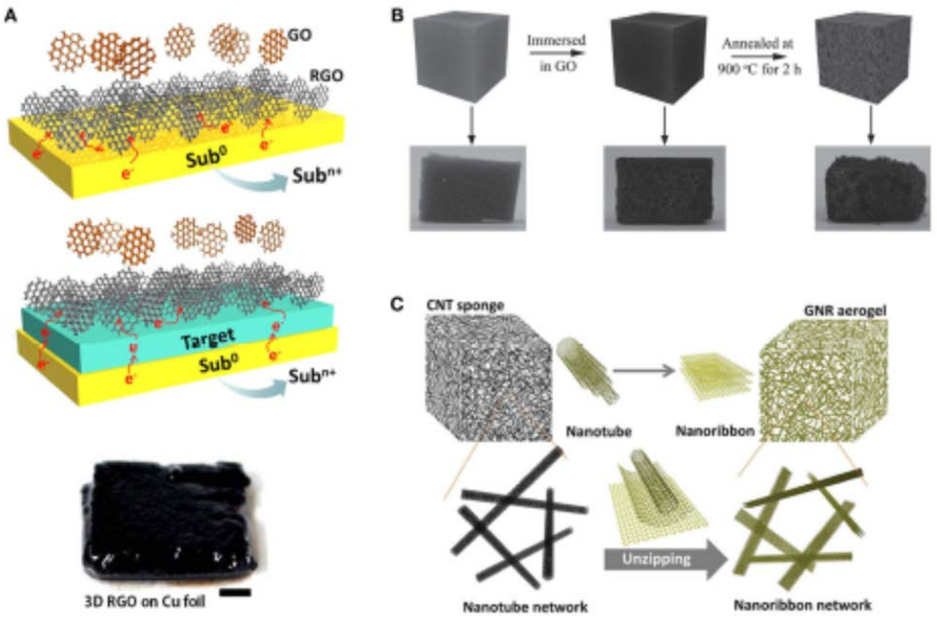


Fig. 12.2. (A) Scheme of the substrate-assisted reduction and assembly of GO on active metal substrate and arbitrary conductive target supported by active metal [18]. (B) Schematic representation and photoimages of the procedure for preparing sponge template grapheme aerogel [19]. (C) Illustration of the process in which a CNT sponge is directly converted into a GNR aerogel by unzipping method [201.

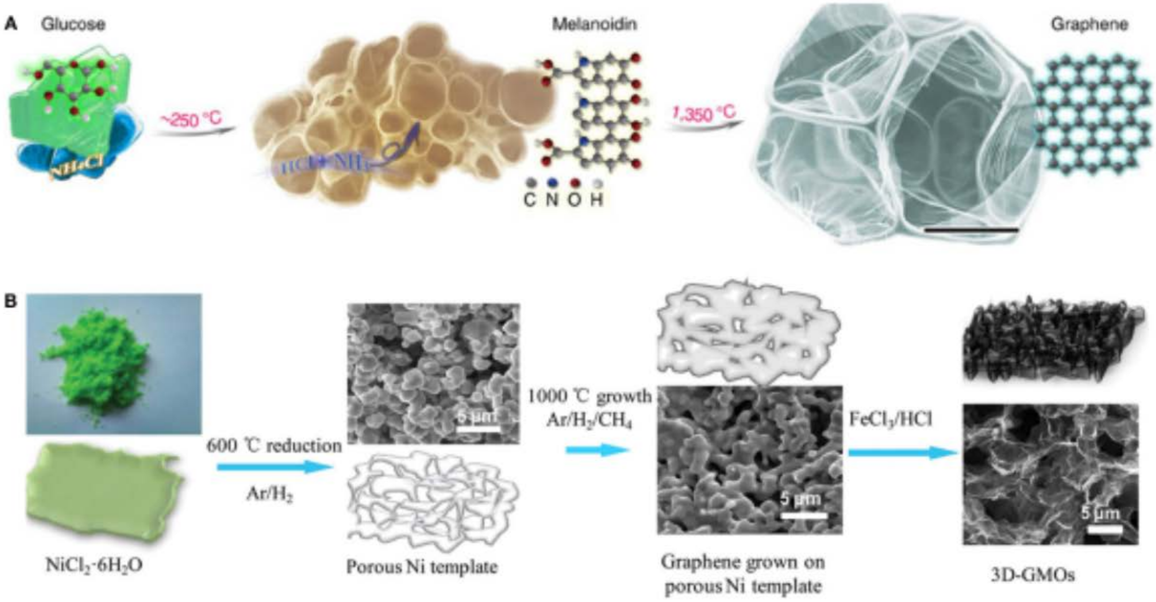


Fig. 12.3. (A) SEM image and schematic growth process of porous graphene structure prepared by sugar-blowing method [21]. (B) Schematic illustrations displaying the preparation process of 3D graphene macroscopic objects [22].

12.2.3. CVD Grown3D Graphene

As first demonstrated in work [23], 3D defect-free graphene monolith can be obtained by CVD using Ni foam as the growth substrate. Such graphene foam exhibits macro - porous structure (99.7% porosity), high-specific surface area (~850 m^2 g^{-1}), and high conductivity (~10 S cm^{-1}). The Ni foam can be subsequently removed by acid, leaving light weighted and free-standing graphene foam [24]. Recently, microporous 3D graphene foams via a rapid CVD process using nickel chloride hexahydrate as the catalyst precursor has been fabricated [22] (Fig. 12.3B). Prolonging the growth time enhances the thickness and strength of graphene layers. Using plasma enhanced CVD process, dense vertically oriented graphene sheets were grown on Ni foam, offering a large surface area [25].

12.2.4. 3D CNT-Graphene Hybrids

CNT and graphene possess a plethora of different characteristics in structural and physicochemical properties. To enjoy the synergistic effects of these two graphitic allotropes, 3D CNT/graphene hybrid architectures have been prepared by hydrothermal treatment of CNT and graphene solution mixture [26]. In this composite, the intercalating CNTs not only prevent stacking of graphene sheets but also facilitate electron and ion transport. Two-step CVD processes have been devised to grow graphene foam covered with CNT matrix (Fig. 12.4) [27]. CNT nano-meshes on top of monolithic graphene skeleton further enhance surface area, conductivity, and charge-transfer conductance. CNT-graphene hybrid can also simply obtained by dip-coating CNT dispersion on pre-synthesized graphene foam, taking advantage of the strong hydrophobic and pi-pi interactions between CNTs and graphene [28]. 3D CNT-graphene sandwich structure from catalyst precursor-coated GO sheets has been prepared [29].

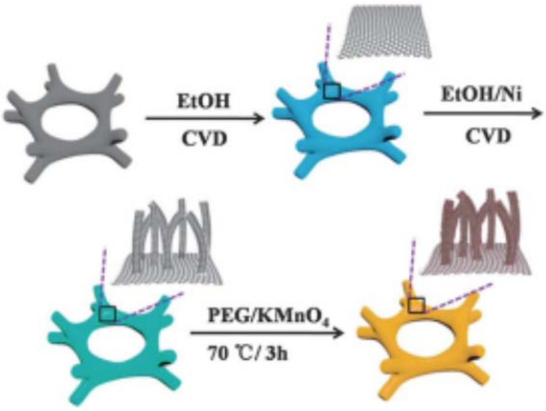


Fig. 12.4. Schematic illustrations of the fabricated 3D MnO$_2$–CNT–graphene–Ni hybrids [27].

12.3. FUNCTIONALIZATION

The 3D nanocarbon structures can be functionalized with biomolecules, polymers, or other functional nanomaterials (e.g., precious metal nanoparticles and nanostructured metal oxides) to attain improved or new functionalities [30, 31]. Functionalization of enzymes via physisorption or covalent bonding renders the 3D nanocarbon bio-catalytic capabilities. Decoration of metal or metal oxide nanoparticles renders high electrochemical or catalytic properties. Integration of conducting polymers enhances the conductivity, mechanical strength, as well as electrochemical activities. The 3D architectures not only provide a large supporting/loading surface area but also alleviate the commonly encountered problem of nanomaterial aggregation. Heteroatom doping is another effective route to modify the properties of 3D nanocarbon structures [32].

12.4. APPLICATIONS

12.4.1. Fuel Cells

The 3D CNT/graphene structures have been employed as the electrochemical electrodes for fuel cells that convert chemical energy into electricity [33]. In a direct methanol fuel cell, Pt-decorated graphene/g-C3N4 composite has been used as the anode for methanol oxidation (Fig. 12.5A) [34]. It provides higher forward-scan peak current, better anti-poisoning capability, and superior long-term stability than Pt/graphene, Pt/carbon nitride, and commercial Pt/C electrodes. The 3D porous graphene structure is advantageous for rapid transport of ions, fuel molecules, and electrons, as well as homogeneous and dense loading of Pt nanoparticles. A novel iron nitride/nitrogen doped-graphene aerogel hybrid synthesized by a two-step hydrothermal process was used as the fuel-cell cathode for oxygen reduction reaction (ORR) **(**Fig. 12.5B**)** [35]. It demonstrated more positive onset reduction potential, higher current density, and lower charge-transfer resistance than electrodes based on iron nitride or N-doped graphene aerogel. With comparable catalytic activity, the hybrid outperforms the commercial Pt/C catalyst in methanol crossover resistance and stability. Heteroatom doping can confer 3D graphene/CNT architectures with remarkable electrocatalytic ability for ORR. The 3D N-doped CNT-arrays have been reported to exhibit excellent ORR activity with the ideal four electron pathway in both acid and alkaline conditions [36, 37]. Using CVD method, B/N co-doped graphene foam as efficient metal-free ORR catalyst has been (2013) synthesized [38]. It offers a higher reduction peak current than the commercial Pt-C/glassy carbon electrode.

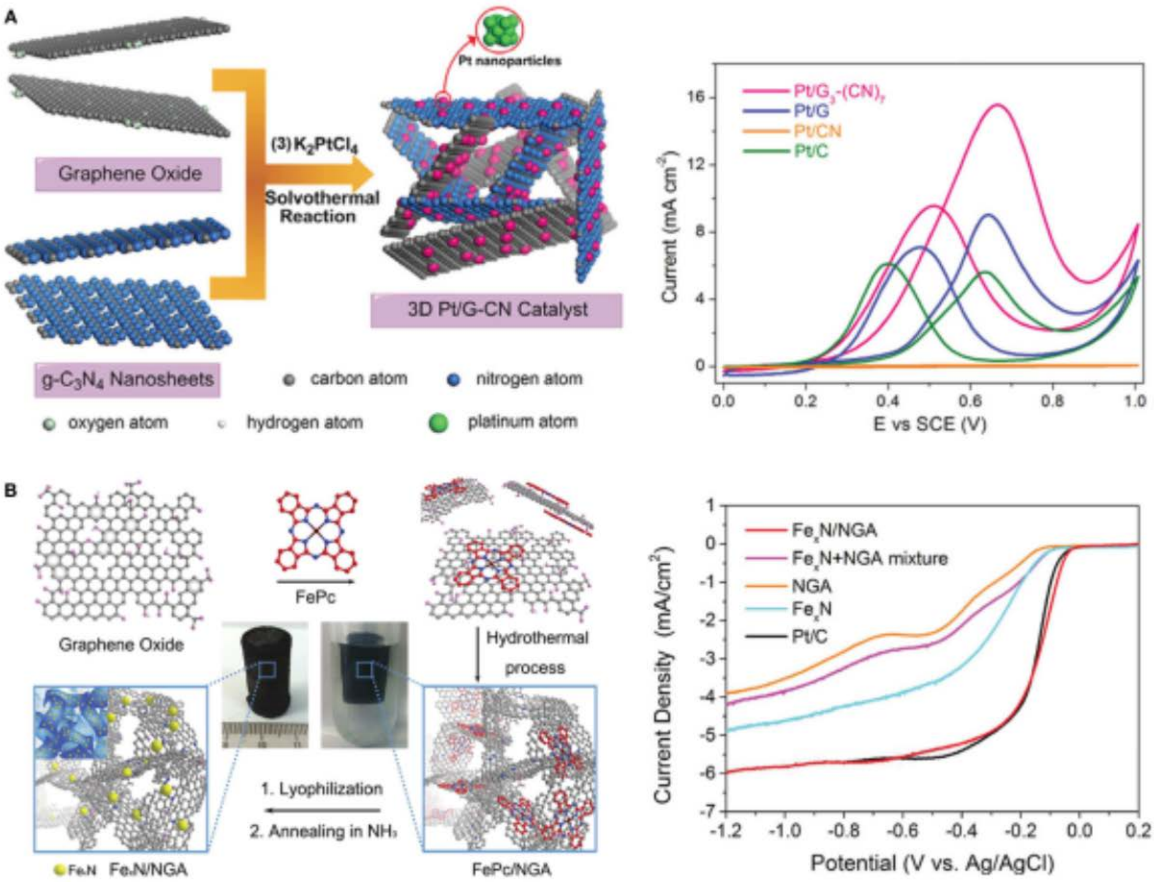


Fig. 12.5. Illustration of the synthesis of the 3DPt/graphene–carbon nitride catalyst and its electrochemical performance in direct methanol fuel cell [34]. (B) Illustration of the preparation procedure of Fe_xN/N-doped graphene aerogel and its oxygen reduction reaction performance [35].

12.4.2. Biofuel Cells

Biofuel cells share the same design with traditional fuel cell while replacing anode catalyst with biological components (enzymes or microorganisms), which are able to harvest biochemical energy into electricity. One major problem of enzymatic biofuel cell is the poor electron transfer from the buried catalytic centers of the enzyme molecules. To tackle this issue, enzyme decorated 3D hybrid of graphene foam and single-walled CNTs as novel biofuel cell electrodes has been employed [28]. The as- fabricated device attains the theoretical limit of open circuit voltage (~1.2 V) and a high power density of 2.27 mW cm^{-2}, which benefits from the large surface area and high conductivity of the 3D electrodes, and the nearly perfect electrical coupling ensured by the intimate contact between enzyme molecules and nanotopographic electrode surface. In comparison to the conventional planar electrodes, 3D carbon electrodes can greatly improve the output of microbial fuel cells by promoting colonization and loading of electrogenic bacteria and 3D electrical coupling. 3D CNT-coated macroporous sponge as the bioanode and achieved a maximum current density of 10.63 mA cm^{-3} in glucose media has been employed [7], which was 48% higher than that of CNT-textile electrode. A novel 3D macroporous anode based on PANI decorated CVD- grown 3D graphene foam, which reached a powder density of ~190 mWm^{-2} has been demonstrated [39]. It significantly outperforms planar carbon electrodes, e.g., carbon cloth or felt electrode with or without PANI coating.

12.4.3. Supercapacitors

Supercapacitors, also called as ultracapacitors, are electrochemical energy storage devices that combine the high energy-storagecapability of conventional batteries with the high power-deliverycapability of conventional capacitors [40 - 44]. Being able to achieve higher power and longer cycle life than conventional dielectric capacitors and batteries, supercapacitors have been developed to providepowerpulses for a large variety ofapplications, ranging from consumer electronics through hybrid electric vehicles (HEVs) to industrial electric utilities [45]. Therefore, supercapacitors play an important role in achieving better fuel economy, decreasing harmful emissions, and reducing the reliance on petroleumsources. The world market for supercapacitors has been growing steadily and rapidly [41]. To improve the performance of state-of-the-art supercapacitors to meet the stringent requirements for the applications mentioned above, and many other advanced applications not discussed (e.g. portable, transparent and wearable electronics), new electrode materials with superior properties over those of current activated carbon electrodes are needed and new device structures (e.g. all-solid state supercapacitors [46, 47], optically transparent [48, 49], mechanically flexible and stretchable [50–52], and even fiber shaped [53–55] supercapacitors) are highly desirable.

Nanotechnology has opened up new frontiers by offering unique enabling technologies and new materials for energy storage. In particular, graphitic carbon nanomaterials (e.g. carbon nanotubes, graphene sheets) have been playing a more and more important role in the development of high-performance supercapacitors [44, 45]. The aim of this article is to summarize recent progress in the development of supercapacitors (especially electrical double layer capacitors) based on carbon nanomaterials and to provide various rational concepts for materials engineering to improve device performance.

A typical supercapacitor consists of three essential components, namely the electrodes, the electrolyte, and the separator. The overall performance of supercapacitors is determined by the physical properties of both the electrode and the electrolyte materials. Nevertheless, the electrode is one of the most important components for charge storage/delivery, and plays a crucial role in determining the energy and power densities of a supercapacitor. The electrochemical performance of a supercapaictor can be characterized by cyclic voltammetry and galvanostatic charge–discharge measurements [41–44]. The capacitance (C) is determined from the constant current discharge curves according to Eq. 12.1

$$C = \frac{I}{(dV/dt)} \tag{12.1}$$

where I is the discharge current and dV/dt is calculated from the slope of the discharge curve. Then, the specific capacitance (CSP) for one electrode in a supercapacitor can be calculated using the following equation:

$$C_{SP}(Fg^{-1}) = \frac{4C}{m} \qquad (12.2)$$

where C is the measured capacitance for the two-electrode cell and m is the total mass of the active materials in both electrodes. The mass can also be replaced by volume or area of the electrodes depending on the nature of the targeted applications. The stored energy (E) and the power density (P) in a supercapacitor can then be calculated from Eqs 12.3 and12.4, respectively:

$$E = \frac{CV^2}{2} \qquad (12.3)$$

$$P = \frac{V^2}{(4R_S)} \qquad (12.4)$$

where C (F g^{-1}) is the total capacitance of the cell, V is the cell voltage, and R_S is the equivalent series resistance.

The principle of energy storage in a supercapacitor can be either (i) electrostatic charge accumulation at the electrode/electrolyte interface (electrical double layer capacitance, EDLC), as schematically shown in Fig. 12.6, or (ii) charge transfer, via reversible (Faradaic) redox reaction(s), to redox materials (e.g. conductive polymers, metal oxide nanoparticles) on the surface of electrode (pseudo-capacitance). In practical supercapacitors, the two storage mechanisms often work simultaneously [56]. Different charge transfer processes involved in the EDLC and pseudo-capacitance [44, 45, 56]. In EDLC, the energy is stored through ion adsorption (a purely electrostatic process) at the electrode-electrolyte interface with no charge transfer across the electrodes, suggesting a non-faradic process. By contrast, pseudo-capacitance arises from reversible redox reaction(s) between the electrolyte and active species on the surface of electrodes. Although pseudo-capacitance higher than EDLC capacitance can be achieved, supercapacitors based on pseudo-capacitance often suffer from the poor electrical conductivity of the electroactive species, and hence demonstrate low power density and cycling stability. Therefore, the combination of both EDLC and pseudo-capacitance presents an effective means to improve the overall capacitance of a supercapacitor.

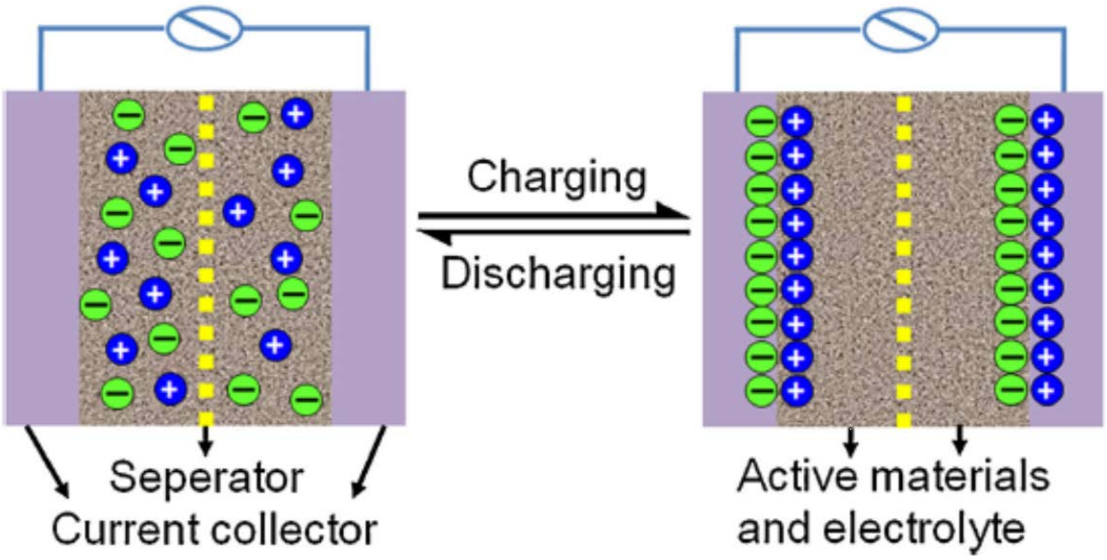


Fig.12.6. Schematic illustration of the charging/discharging process in a supercapacitor.

Since both EDLC and pseudocapacitance are surface phenomena, high-surface-area mesoporous carbon and activated carbons (specific surface area: 1000–2000 m^2 g^{-1}) have been widely used as electrode materials in both academic and commercial supercapacitors [56–59]. Taking a specific surface area of 1000 m^2 g^{-1} for carbon as an example, its ideal attainable capacitance could be 200–500 F g^{-1}. However, the practically obtained values are of only a few tens of F g^{-1}. Activated carbons have a wide pore size distribution, consisting of micropores (<2 nm), mesopores (2–50 nm), and macropores (>50 nm) [57, 58], with most of the surface area of activated carbons being on the scale of micropores [60]. Pores of this size are often poorly or non-accessible for electrolyte ions (especially for organic electrolytes), and thus are incapable of supporting an electrical double layer. By contrast, mesopores contribute the most to the capacitance in an electrical double layer capacitor [61–63]. However, recent experimental and theoretical studies have demonstrated that charge storage in pores 0.5–2 nm in size (smaller than the size of solvated electrolyte ions) increased with decreasing pore size due to the closer approach of the ion center to the electrode surface in the smaller pores [58, 64–66]. Pores less than 0.5 nm wide are too small for double layer formation [66]. Currently available activated carbon materials have a high surface area but unfortunately a low mesoporosity, and hence a limited capacitance due to a low electrolyte accessibility [60]. This translates to the limited energy density of the resultant supercapacitors (Eqn 3). The low electrolyte accessibility of activated carbons, coupled with their poor electrical conductivity, produces a high internal resistance and hence a low power density for the capacitors (Eq. 12.4) [60]. Consequently, a limited energy density (4–5 Wh kg^{-1}) and a limited power density (1–2 kW kg^{-1}) have been obtained for currently available supercapacitors based on the activated carbon electrodes [60]. Clearly, therefore, new materials are needed to overcome the drawbacks of activated carbon electrode materials to improve the performances for supercapacitors.

Owing to their large surface area, high mesoporosity and electrolyte accessibility, and good electrical properties, carbon nanomaterials, especially graphene and carbon nanotubes (CNTs), are very promising candidates to replace activated carbons as the electrode materials in high-performance supercapacitors [67, 68]. Graphene, a single-atom-thick layer of sp^2 carbon atoms densely packed into a two-dimensional (2D) honeycomb lattice, can be viewed as the basic building block for carbon materials of all other dimensionalities, such as 0D fullerene, 1D nanotubes, and 3D graphite (Fig. 12.7) [69]. For instance, a carbon nanotube (CNT) may be conceptually viewed as a graphene sheet that is rolled into a nanoscale tube form as a single-walled carbon nanotube (SWNT), or as a multiwalled carbon nanotube (MWNT) with additional graphene coaxial tubes around the SWNT core [69].

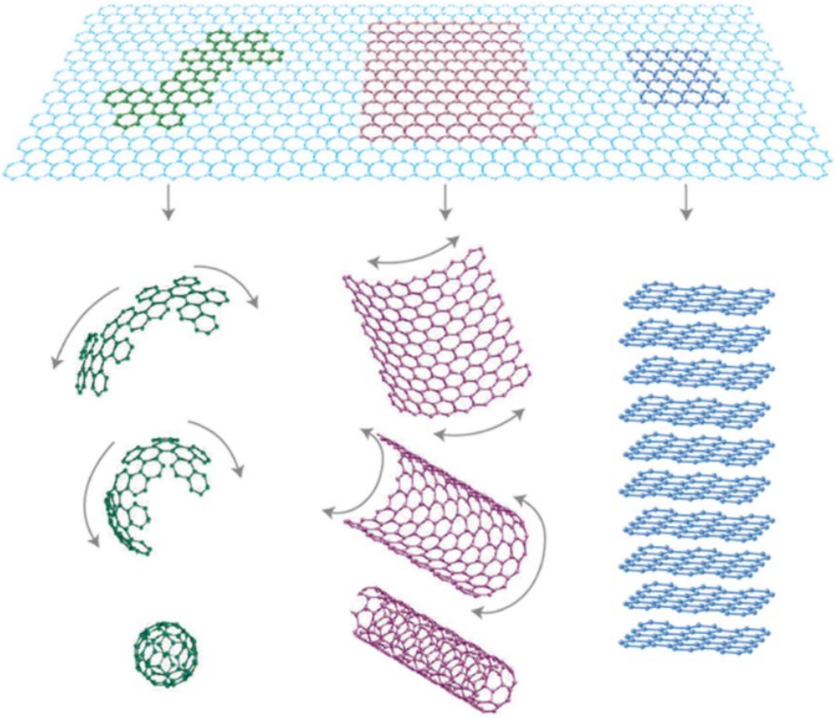


Fig. 12.7. The mother of all graphitic forms. Graphene is a 2D building block for carbon materials of all other dimensionalities. It can be wrapped up into 0D buckyballs, rolled into 1D nanotubes or stacked into 3D graphite [67].

CNTs with a high aspect ratio, large specific surface area (SWNT >1600 $m^2 g^{-1}$, MWNT >430 $m^2 g^{-1}$) [70, 71] as well as good mechanical and electrical (~5000 S cm^{-1}) properties have been widely used as the active electrodes in supercapacitors [68, 72–77]. Having many similarities to CNTs in structure and properties, including its high aspect ratio (the ratio of lateral size to thickness), large surface (2630 $m^2 g^{-1}$) [78–80], excellent carrier mobility (15,000 $cm^2 V^{-1} s^{-1}$ for both electrons and holes) [81], and good mechanical properties [82], graphene is an attractive candidate for many potential applications where CNTs have been exploited. Superior to CNTs, one-atom-thick graphene sheets with a 2D planar geometry could offer additional advantages as more effective electrode materials in supercapacitors [83]. Furthermore, the availability of many rational strategies for materials engineering and rich chemistries for controlled functionalization of both CNTs and graphene should facilitate materials engineering for the development of various high-performance supercapacitors based on carbon nanomaterials, as described below.

High-performance supercapacitors by controlling the orientation and tip structure of carbon nanotubes. As mentioned above, CNTs with a high aspect ratio, large specific surface area [70, 71] and high electrical conductivity have been widely used as the active electrode in supercapacitors [68, 72–77]. For instance, a supercapacitor of a specific capacitance of 102 F g^{-1} with a power density >8 $kWkg^{-1}$ and an energy density of ~1Wh kg^{-1} has been fabricated from electrodes based on freestanding mats of entangled MWNTs in H_2SO_4 electrolyte [71]. More recently, a higher specific capacitance of 180 F g^{-1} has been achieved when a random SWNT network (Fig. 12.8) was used as the electrode in a KOH electrolyte [72]. In this particular case, the maximum power density and energy density approached 20 kW kg^{-1} and 10 Wh kg^{-1}, respectively. This value of power density is higher than those attainable by activated carbon-based commercial capacitors, but the energy density still needs to be improved. The randomly entangled or bundled structure could have significantly reduced the CNT specific area with a concomitant decrease in the electrochemical energy-storage performance [75, 84, 85].

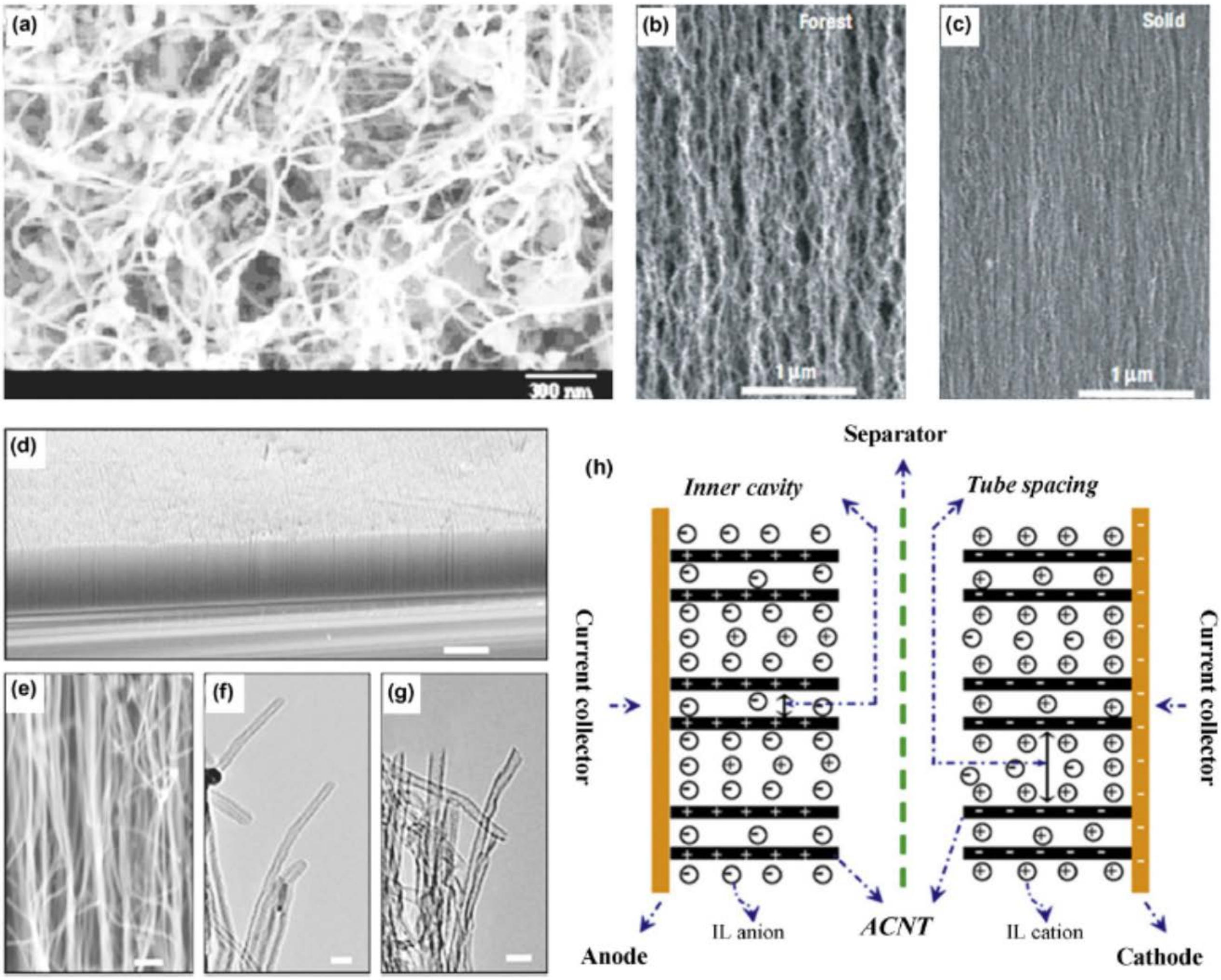


Fig. 12.8. Typical structures of random, VA-CNTs and the schematic illumination of VA-CNT based supercapacitors. (a) SEM image of the as -grown randomly entangled SWNTs [72]. (b,c) SEM images of the as-grown forest (b) after having been solidified with a droplet of liquid (c) [75]. (d) SEM image of a plasma etched VA-CNT electrode (scale bar: 100 mm). (e) Higher magnification view of the electrode (scale bar: 100 nm). TEM images of the CNTs before (f) and after (g) plasma etching (scale bar: 20 nm) (h) Schematic of an electrochemical capacitor based on plasma-etched VA-CNT electrodes and ionic liquid electrolyte [84].

Compared to their randomly entangled counterparts, vertically aligned CNT arrays (VA-CNTs) have been demonstrated to be advantageous for supercapacitor applications. Unlike the irregular pore structures of random CNTs (Fig. 12.8), the vertically-aligned structure and the well-defined tube-spacing in a VA-CNT array (Fig. 12.8b,c) can provide a more electrolyte-accessible surface. The aligned structures should also provide improved charge storage/delivery properties as each of the constituent aligned tubes can be connected directly onto a common electrode to allow them to effectively participate in the charging/discharging process (Fig. 12.8d). Moreover, the top end-caps of the VA-CNTs can be properly opened (e.g. by plasma etching [84, 86]) (Fig. 12.8e–g) to allow the electrolyte access to the otherwise inaccessible inner cavity of the VA-CNTs for charge storage. Indeed, recent research has demonstrated the improved rate capability of VA-CNTs over randomly entangled CNTs [75, 84–89], along with enhanced energy density (148 Wh kg^{-1}) and power density (315 kW kg^{-1}) [90–95]. Specifically, a high capacitance has been obtained in 1 M H_2SO_4 for a VA-CNT array electrode (365 F g^{-1}) prepared by template-assisted CVD [96] and in ionic liquid electrolytes [97] for an VA-CNT electrode (440 F g^{-1}) prepared by a template-free CVD approach (Fig. 12.8e–h) [84, 89]. To further optimize the performance of VA-CNT supercapacitors, conductive polymer (e.g. polyaniline [98] and polypyrrole [99]) and metal oxide (e.g. MnO_2 [100], TiO_2 [101] and RuO_2 [102]) can be deposited onto the surface of CNTs to introduce pseudocapacitance.

High-performance supercapacitors by controlling the edge-structure and π-π stacking of graphene sheets. Supercapacitors based on CNT electrodes do not exhibit the expected performance (e.g. specific capacitance below 500 F g^{-1}) due to the high contact resistance between the CNT electrode and current collector and its inefficient interaction with the electrolyte. Besides, it is difficult, if not impossible, to scale up the production of VA-CNTs for the commercialization of supercapacitors [103]. The recent availability of solution-processable graphene oxide (GO) by exfoliation of graphitepowder via solution oxidation [104] and edgefunctionalized graphene (EFG) sheets via ball milling [105, 106] has allowed for the large-scale production of graphene materials at a relatively low cost for various device applications through solution processing [107]. Due to its large surface area, high carrier transport mobility and excellent thermal/mechanical stability, graphene has recently been studied as an alternative carbon based electrode in supercapacitors [108]. Theoretically, the double-layer capacitance value of a graphene electrode can reach up to 550 F g^{-1}, the highest value of intrinsic capacitance among all carbon-based electrodes [109, 110]. Using chemically reduced graphene oxide electrodes, a supercapacitor with specific capacitances of 135 and 99 F g^{-1} in aqueous and organic electrolytes, respectively, has been fabricated [79], as was a supercapacitor with ultrahigh specific energy density (85.6 Wh kg^{-1} at room temperature and 136 Wh kg^{-1} at 80°C) [110]. Recently, supercapacitors with high specific capacitance (276 F g^{-1}), power density (20 Wcm^{-3}, 20 times higher than that of the activated carbon counterpart), energy density (1.36 mWh cm^{-3}, 2 times higher than that of the activated carbon counterpart), and excellent stability during bending from 08 to 1808 have also been developed by using a standard Light Scribe DVD optical drive to directly reduce graphene oxide films to graphene electrodes (Fig. 12.9a–c) [111]. The performance can be further enhanced by using a conductive polymer [102, 103] or metal oxide [114, 115] coating to introduce pseudocapacitance.

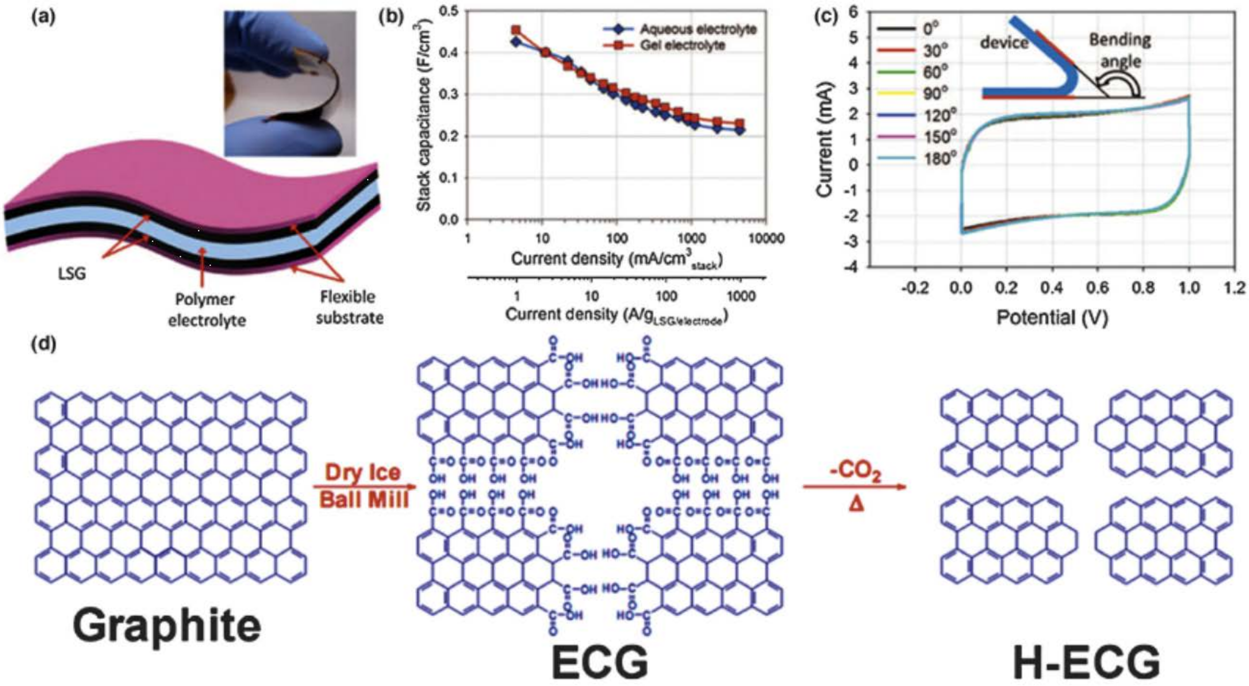


Fig. 12.9. Structure and electrochemical properties of an all-solid-state supercapacitor based on laser-scribed graphene (LSG) and the synthetic route to edge functionalized graphene. (a) Schematic diagram of the supercapacitor based on LSG. Inset is a photograph showing the flexibility of the device. (b) A comparison between performances of LSG-EC using gelled versus aqueous electrolytes. (c) Flexibility of the LSG-based supercapacitor: current–voltage measurements were collected at a scan rate of 1000 mV s^{-1} [111]. (d) Syntheses and proposed structures of edge-carboxylated graphene (ECG) and heat-treated (decarboxylated) ECG (H-ECG) [105].

Recent work on the EFG with various edge groups (e.g. –H, – COOH and –SO_3H) by effective and eco-friendly ball milling of graphite (Fig. 12.9d) provides an effective means for the development of functionalized graphene materials with tailor-made chemical structures and electronic properties attractive for multifunctional applications [105], including large-area transparent and conducting electrodes for electronics and metal-free catalysts for oxygen reduction reaction in fuel cells [106]. Although EFGs have hardly been exploited for energy-storage application, the use of EFGs, having abundant active sites at the edge and perfect conjugation (conductivity) on its basal plane, as the electrode materials could lead to high-performance supercapacitors with a high rate capability. Through controllable self-assembling, EFGs should also facilitate the formation of hierarchically structured electrodes in supercapacitors for specific applications.

High-performance supercapacitors based on 3D pillared graphene-carbon nanotube networks. Although graphene sheets with a large surface area are ideal electrode materials for energy storage, a large portion of the surface area associated with individual graphene sheets in the grapheme electrode is inaccessible due to restacking via the strong π-π interaction. Therefore, it is crucial to physically separate 2D grapheme sheets to preserve their high surface area, for example by template-assisted construction of 3D graphene foams (Fig. 12.10a,b) with a porous structure and large surface area [116–119]. In addition, a solution layer-by-layer self-assembling approach to prepare multilayered hybrid carbon films of alternating poly(ethyleneimine)- modified graphene sheets and acid-oxidized MWNT layers was developed to construct supercapacitors with a relatively high specific capacitance of 120 F g^{-1} (higher than those of vertically- aligned [75] and random CNT electrodes [72]) and rate capability since the well-defined interlayers of CNT networks allowed for not only a fast ion diffusion but also efficient electron transport [120]. However, it is difficult to control the porosity and pore distribution within the multilayered CNT/graphene hybrid film as each of the constituent CNT and graphene layers is randomly assembled, and hence it is infeasible to further improve the device performance. To address this issue, 3D pillared VA-CNT/grapheme architectures with alternating VA-CNT and graphene layers, whose porous structure and pore distribution can be easily controlled by tuning the length and packing density of VA-CNTs, have been produced [121–123] and investigated as ideal electrode materials for advanced supercapacitors [124–128]. By intercalated growth of VA-CNTs into thermally-expanded highly ordered pyrolytic graphite (HOPG), a 3D pillared VA-CNT/graphene architecture was created (Fig. 12.10c,d) [129], which showed a specific capacitance of about 110 F g^{-1} in an electrical double layer supercapacitor. The resulting 3D pillared structure hybridized with nickel hydroxide coating showed a high specific capacitance 1065 F g^{-1} (Fig. 12.10e) with a remarkable rate capability and excellent long-term electrochemical stability – only 4% capacity loss after 20,000 charge–discharge cycles (Fig. 12.10f) [129].

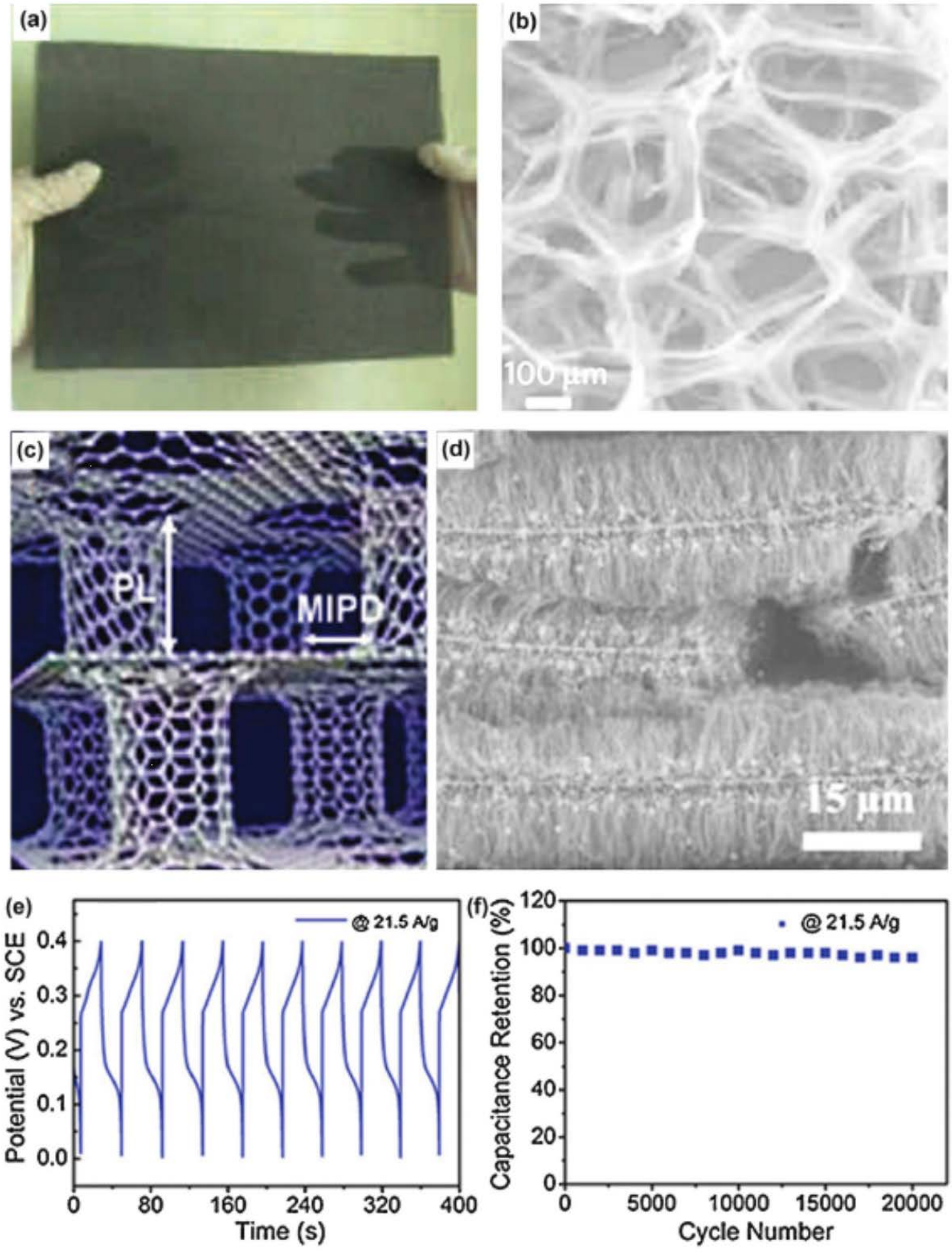


Fig. 12.10. Structure of graphene foam, 3D pillared VA-CNT/graphene architecture and the electrochemical properties of its composite with $Ni(OH)_2$. (a) Photograph of a 17 cm × 22 cm free-standing graphene foam. (b) SEM image of a graphene foam [116]. (c) Schematic diagram of a 3D pillared VA-CNT/graphene nanostructure. (d) Typical SEM images of the 3D pillared VA-CNT/grapheme architectures. (e) Galvanostatic charge and discharge curves for the $Ni(OH)_2$-coated VA-CNT/graphene electrode at a current density of 21.5 A/g. (f) Dependency of

the specific capacitance versus cycle number for the $Ni(OH)_2$-coated VA-CNT/graphene electrode at a galvanostatic charge and discharge current density of 21.5 A g^{-1} [129].

The value of 1065 F g^{-1} is about 10 times that of the high surface-area activated carbons (<100 F g^{-1}) [130] and within the range of 953–1335 F g^{-1} for graphene-supported single-crystalline nickel hydroxide hexagonal nanoplates [131]. In the 3D hierarchical structure, VA-CNTs can act as not only mechanical supports for the graphene layers but also good conductive paths for electrons and ions, and hence the high capacitance and excellent rate capability.

High-performance supercapacitors with novel structures. As discussed above, significant progress has been achieved with conventional supercapacitors based on liquid electrolytes. However, they cannot satisfy the requirements for certain specific applications, including portable, transparent and wearable electronics. In this regard, recent work on the development of lightweight, flexible, stretchable (Fig. 12.11a–c) [50–52] and/or transparent (Fig. 12.11d,e) [8, 9] supercapacitors with novel structures (e.g. all-solid [46, 47], fiber-shaped [53–55]; Fig. 12.11f–h) has attracted a great deal of attention. Highly-stretchable supercapacitors based on buckled SWNT macrofilm electrodes have also been developed by coating a thin SWNT film onto a pre-strained elastomeric substrate (polydimethylsiloxane, PDMS) and followed by relaxation of the prestrained substrate (Fig. 12.11a) [51]. Even when these supercapacitors were stretched up to 30% stains, their electrochemical properties and stabilities remained almost unchanged with respect to the pristine state (Fig. 12.11b,c). By extension, stretchable electrodes based on graphene or aligned CNT films should be potential candidates for stretchable supercapacitors with a high efficiency [132, 133], which can be used as the energy source for stretchable electronics. On the other hand, supercapacitors with high optical transparency (Fig. 12.11d,e) are useful for personal electronics and energy windows as power sources; they are also crucial components for many other power-integrated devices (e.g. self-powered light-emitting diodes, vide infra). Recently, supercapacitors in a fiber form (Fig. 12.11f–h) have been demonstrated to show many unique advantages, including being lightweight, wearable and flexible, for use as power sources by integrating them into clothing, bags and other textiles [53–55]. Specific capacitance as high as 38 mF cm^{-2} has been achieved for a fiber-shaped supercapacitor based on the CNT/ PANI composite fiber electrodes [55].

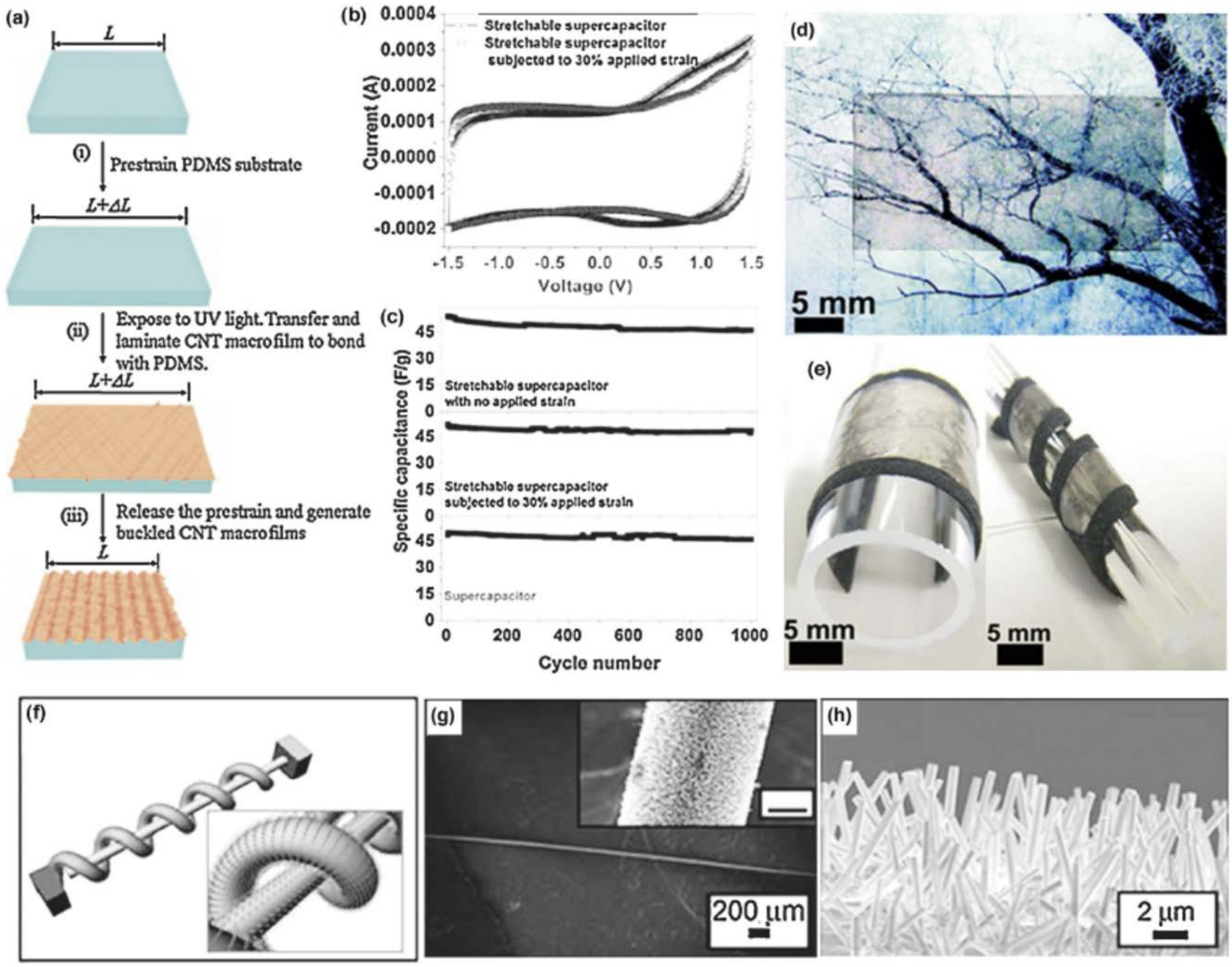


Fig. 12.11. Supercapacitors with novel structures and their electrochemical performances: (a) Schematical illustration of the fabrication steps of a buckled SWNT macrofilm on an elastomeric PDMS substrate for stretchable supercapacitor. (b) Cyclic voltammograms of the stretchable supercapacitors measured at scan rates of 100 mV s^{-1}. (c) Long charge–discharge cycling at a constant current density of 1 A g^{-1} demonstrates the stability of the stretchable supercapacitor under 0 and 30% applied tensile strain [50]. (d,e) Optical pictures demonstrating transparent (d) and flexible (e) natures of the supercapacitors [48]. (f) Schematically illustration of a fiber-based supercapacitor. (g) Low-resolution SEM image of a Kevlar fiber covered with ZnO nanowire arrays. (Inset) A close-up view (scale bar: 10 mm). (h) Higher-magnification SEM image of the plastic wire, showing arrays of NWs [53].

The 3D CNT/graphene architectures, with low level aggregation and agglomeration of individual CNTs of graphene sheets, are able to provide large electric double-layer capacitance (EDLC) owing to large surface area. In addition, they permit high loading of electrochemically active nanomaterials to gain high electrochemical pseudocapacitance. The 3D carbon nanocup-CNT structures have been synthesized on anodized aluminum oxide for EDLC based supercapacitors [134]. The CNT modification gives carbon nanocup an extremely high-specific surface area (1340

m^2 g^{-1}) and enhances its capacitance to 45 F g^{-1} in 1M $LiPF_6$. Open-porous rGO foam prepared from autoclave leavening GO film is also able to provide a large EDLC (110 F g^{-1}) which is one order higher than that of rGO films (17 Fg^{-1}) [135]. Seamless and sp^2-carbon connection of CVD grown graphene are advantageous to exploit the synergistic effects between CNT and graphene. The 3D graphene-CNT hybrids grown on nickel foam prepared by two-step CVD approach have been reported to exhibit a five-time-higher specific capacitance than CNT directly grown on inconel without graphene as the supporting substrate [136].

Vertically aligned CNT forest with uniformly embedded nickel oxide nanoparticles has been prepared [137], which delivered a 5.7 times higher specific capacitance (1.26 F cm^{-3}) than pure CNT forest samples because of combination of both EDLC and pseduocapacitance (Fig. 12.12A). A general strategy to prepare functional porous graphene hybrid architectures (Fig. 12.12B) has been reported [138]**.** As an example, rGO/PANI nanofiber electrode offers a high-specific capacitance of 475 F g^{-1}, much larger than that of pure rGO architecture. CVD-grown graphene foam with well-defined macroporous structure can serve as an excellent platform to prepare various functional supercapacitor electrodes [139, 140]. For example, deposition of nanostructured Co_3O_4 on CVD-graphene foam offers a high-specific capacitance of 1100 F g^{-1} [24]. 3D MnO_2-CNT-graphene-nickel hybrid foam. Such porous and highly conductive electrode delivers a specific capacitance of -251 F g^{-1} has been prepared [27]. Heteroatom doping and integration of CNTs can be employed to enhance the capacitive behavior. N-doped 3D CNT-graphene aerogel electrode for supercapacitor application has been prepared in [26]. It gives a maximum specific capacitance of 180 F g^{-1} at a current density of 0.5 A g^{-1}. Addition of CNTs effectively prevents graphene aggregation and improves the conductivity.

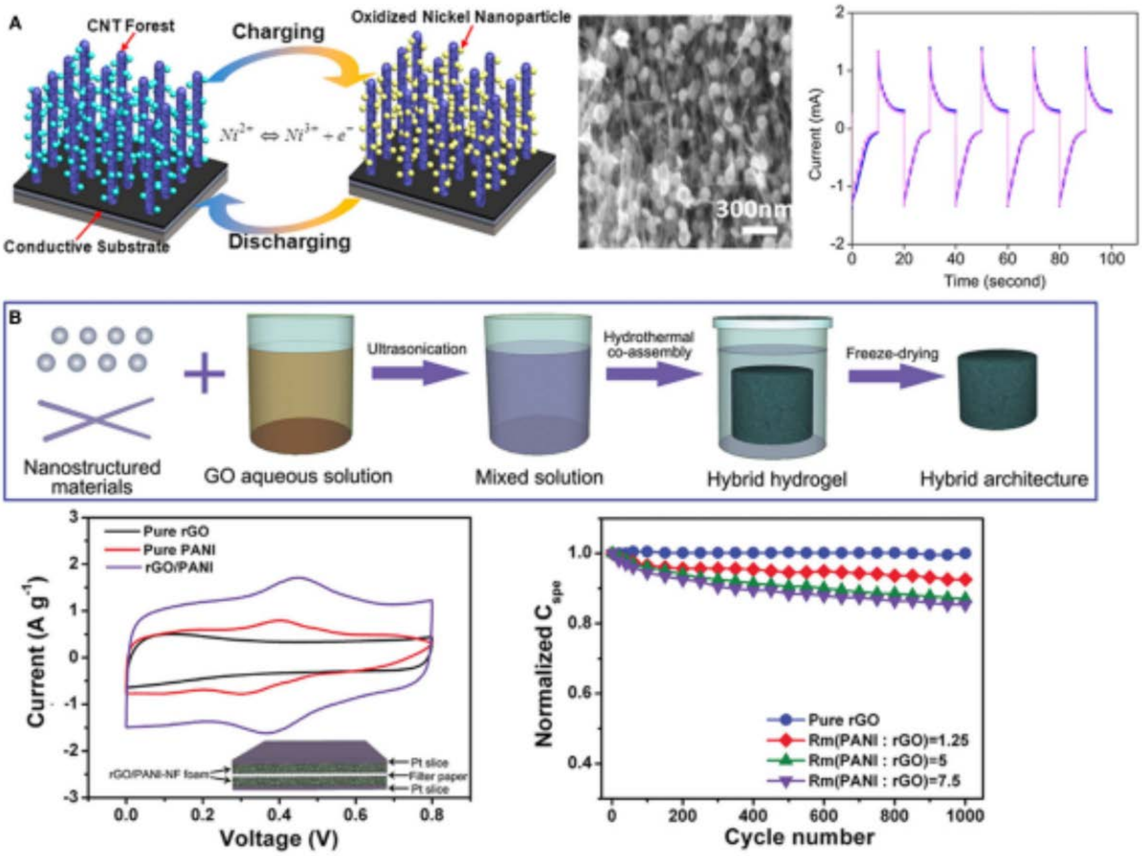


Fig. 12.12. Schematic of the pseudocapacitor using CNT forests functionalized with oxidized nickel nanoparticles as the electrode and its cyclic voltammetry curves in 0.1MKOH at scan rates of 100mVs^{-1} [137]. (B) Schematic of preparing porous grapheme hybrid structures by a hydrothermal co-assembly process and the capacitive behavior of resultant 3D graphene/polyaniline example [138].

12.4.4. Lithium Ion Batteries

The 3D porous graphene/CNT architectures are also attractive for lithium ion batter (LIB), which can improve Li ion diffusion and tolerance to mechanical stress induced by ion deintercalation process. CVD-grown 3D graphene for LIB cathode has been first applied [141]. Accounting the total mass, such 3D graphene loaded with lithium iron phosphate exhibited a 23 and 170% higher specific capacity than that of Al- and Ni-foam loaded with lithium iron phosphate, respectively. 3D graphene aerogel/Fe_2O_3 composite by a one-pot hydrothermal process has been prepared in [142]. The synergistic interaction between uniformly dispersed Fe_2O_3 nanoparticles and robust 3D graphene aerogel endows such electrode with a high reversible capacity of 995 mAh g^{-1}, which is almost four times to that of physically mixed Fe_2O_3-graphene electrode. Integration of CNTs to a similar electrode further increases the reversible capacity to 1118 mAh g^{-1} [143]. TiO_2 nanocrystal/graphene aerogel as LIB anode, using glucose as dispersion and linking agent has been prepared [144]. A reversible capacity of 956.2 mAh g^{-1} was achieved, which is superior to the previously reported TiO_2/carbon electrodes [145]. Even at a high current density of 5000 mAh g^{-1}, the hybrid foam electrode still gives a reversible capacity of 99 mAh g^{-1}, four times higher than that of pure TiO_2 based electrode. A bottom-up method to construct 3D LIB electrode from GO sheets and exfoliated MoS_2 nanosheets has been developed [146]. The resultant MoS_2-graphene electrodes exhibit a high reversible capacity of ~1200 mAh g^{-1}, good stability, and excellent rate performance. Nanostructured MoS_2 on CVD-grown graphene foam has been coated by dip-coating and annealing method and demonstrated a high reversible capacity of 877 mAh g^{-1} at a current density of 100 mAh g^{-1} [147].

Flexible Li-ion batteries. Flexible Li-ion batteries (hereafter referred to as LIBs) are used as the energy storage devices for flexible electronics [148 - 150]. To date, several leading research groups have demonstrated nanomaterial based LIB with the flexible polymer substrates. The flexible nanocomposite thin-film battery composed of CNT-cellulose paper electrode and showed a reversible capacity of 110 $mAhg^{-1}$ over several tens of cycles has been fabricated [151]. The chemical vapor deposition (CVD) grown CNT-carbon layer hybrid anodes have been prepared for a paper LIB and demonstrated its high reversible capacity of 572 $mAhg^{-1}$ after 100 cycles [152]. The single-walled carbon nanotubes (SWCNTs) spray has been deposited onto polyethylene terephthalate (PET) films and obtained high specific capacitance ranging from ~90 to 120 F/g by using various aqueous electrolytes [153]. However, these previous results have drawbacks such as low loading amount of electrode materials and complicated coating process. Nevertheless, higher specific and volumetric capacities of the anode have been the recent strong motivation for emerging flexible electrodes comprising no metallic current collectors or even polymer substrates (e.g. SWCNT papers) [154].

Another major research direction for LIBs has been the development of high-efficiency LIBs composed of nanomaterial-based electrodes for the next-generation electric vehicle as one of the most attractive applications. The strong candidates of the nanomaterials can be listed as carbon nanotubes (CNTs) [155], carbon nanofibers (CNFs) [156, 157], graphene [158], silicon nanowires [159], transition metal oxide nanoparticles [160] and so on. In particular, LIB cells assembled by those nanostructured anodes relative to conventional graphite anodes have demonstrated more excellent electrochemical properties in terms of safety, high energy density, C-rate capability and long cycle lifetime. In our previous study, we used directly grown multiwall carbon nanotubes (hereafter referred to as MWCNTs) on metal films with multitude advantages over conventional mesocarbon microbeads (MCMB) or hard carbon- based electrodes since they have their unique properties including high surface area to accommodate more Li^+ ion intake, good electric conductivity for the faster electron charge transfer and good structural resilience for Li^+ ion insertion and extraction into and from the host structure [161, 162]. However, it is difficult to obtain high volume or mass loading amount of nanostructured active materials for their real application by conventional methods of nanomaterial coating or direct growth on the simple 2-dimensional (2D) metal substrates [163]. To address this issue, 3D electrode structures have been implemented in supercapacitors and LIB applications by many leading research teams. 3D network electrode structure has been introduced for effective electron and ion pathways and assembled a prototype lithium-ion pouch battery, which is constituted by a bicontinous lithiated MnO_2 cathode and a graphite anode. They demonstrated relatively high specific capacity value at 60°C charge and discharge (ca. 110 $mAhg^{-1}$) [164]. Very recently, the structure of CNT-graphene seamlessly grown on 3D-Ni foam by a thermal CVD method and its excellent performance of the electrode used in field emitter devices and supercapacitors (double layer capacitors) have been demonstrated [165]. In the structure, the use of 3D-Ni foam enhanced the surface-area utilization efficiency of the metal current collectors, and graphene served as a linking agent between CNTs and metal interface to improve bonding strength and electrical properties between the two different phases. Additionally, our team proposed 3D anode architecture of MWCNTs directly grown on Cu mesh substrate, and the areal mass density of MWCNTs in the 3D structure (2.7 mg/cm^2) was four times higher than that of MWCNTs directly grown on 2D Cu foil, resulting in 160% enhancement of specific capacity in the cycling performance [166].

The architecture of large-scale patternable 3D MWCNTs-graphene hybrids fabricated on transparent flexible films, where graphene serves as an underlayer film to promote the electrical conductivity for 3D network structure of MWCNTs has been we proposed in [148]. The 3D patterned structure of MWCNTs was directly grown on 3D-Cu mesh by a CVD method and transferred onto graphene-PET film by a simple lamination and Cu etching process [18]. As-prepared 3D MWCNTs-graphene-PET structure was inserted in a coin-cell type LIB to evaluate its electrochemical performance for a high-efficiency flexible LIB anode. The 3D MWCNTs-graphene-PET anode does not include any binder which may have harmful effect on LIB operation. In addition, bending of 3D MWCNTs-graphene on PET substrates leads to no significant film cracking and peel-off, which is different from the previous report [154]. Moreover, a commonly established roll-to-roll lamination method has potential for high-throughput production of the 3D MWCNTs-graphene-PET anode structure.

Preparation for the anode structure A Cu mesh (200 Mesh Copper, TWP) with average thickness of 50 lm and hole size of 65 lm was prepared. 3D MWCNTs grown on Cu mesh were used as anodes for LIB. The 3D MWCNTs-Cu anode materials were prepared by a two-step process including catalysts deposition of Ti (10 nm)–Ni (10 nm) by using a magnetron sputtering system followed by direct growth of MWCNTs by thermal chemical vapor deposition (CVD). In the first step, Ti underlayer played an important role in preventing Ni thin layer from being diffused into the Cu mesh underneath, maintaining activity of catalytic Ni thin layer. In addition, formation of titanium carbide (TiC) thin layer at around 750 °C could improve electrical conductivity and be an effective bonding bridge between MWCNTs and Cu mesh. The presence of TiC thin layer in the interface between MWCNTs and Cu was previously confirmed by high resolution transmission electron microscopy (HRTEM) [167, 168]. In the second step, the Ti–Ni deposited Cu mesh was rapidly heated to 750 °C in an Ar gas environment. Carbon precursor ethylene (C_2H_4) gas carried by hydrogen (H_2) gas (1:2 volume ratio of C_2H_4 to H_2) was fed into the chamber for growth time 50 min. Once the growth time was reached, mixture gases were turned off and the chamber was cooled down to the room temperature while flowing inert Ar gas.

A simple hot lamination process to transfer MWCNTs grown on 3D Cu mesh onto graphene laminated polyethylene terephthalate (PET) flexible substrate has been introduced. The growth procedures of graphene on 2D Cu foil and its transferring methods onto PET films were mentioned in our previous report [169]. Figure 12.13 shows the processing flow for the transferring of 3D MWCNTs on the graphene-PET films. The as-grown MWCNTs-Cu mesh and the graphene-PET were held in a ‘‘face-to-face’’ orientation. The assembled samples of 3D MWCNTs-graphene-PETwere hot-laminated at around 70 °C. After the lamination process, MWCNTs were strongly adhered to the graphene on PET films. To completely remove the Cu mesh from the sandwich-like hybrid structure, the concentrated $FeCl_3$ etching solution was used. After four hour etching at room temperature (~25 °C), the Cu mesh was completely etched away, leaving 3D MWCNT mesh patterns on graphene-PET films. Such a patterned hybrid structure was thoroughly washed with de-ionized water and dried in the oven at 80 °C for 15 min. As presented in the digital image of Fig. 12.13(e), as-prepared structure was flexible enough to form various geometries without any damage of the MWCNT patterns. The primary reason behind this is the important role of pressure-sensitive adhesives coated on the PET films to make it possible to strongly bond between MWCNTs-graphene and PET.

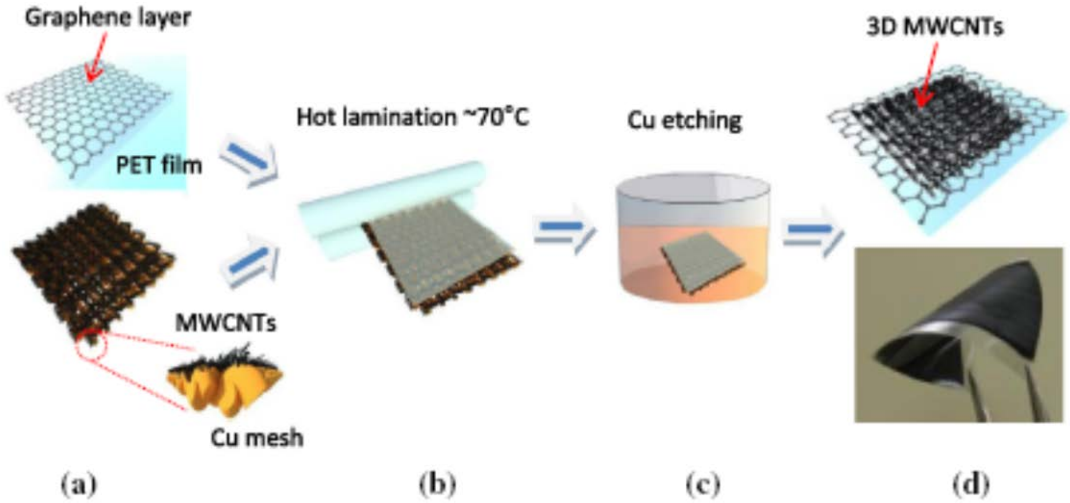


Fig. 12.13. Process flow of transferring technique of 3D MWCNTs-Cu mesh onto the transparent flexible graphene-PET substrate. (a) Hot laminated graphene-PET film and 3D MWCNTs on Cu mesh directly grown by CVD, (b) Hot lamination of 3D MWCNTs onto graphene-PET, (c) Etching process of Cu mesh by using FeCl3 oxidation acid, (d) 3D pattern of MWCNTs transferred on the graphene-PET film. The digital image shows high flexibility of the 3D MWCNTs-graphene-PET films.

Structural characterization of the anode structure. Surface morphology and structural properties of the prepared 3D MWCNTs-graphene-PET structures were observed using a field emission scanning electron microscope (FESEM) (JEOL, JSM-7000F; HITACHI, S-4800), an energy dispersive spectroscope (EDS) (Thermo Electron Corporation, NORAN System SIX), a Raman spectrometer (Ar^+ laser with λ = 514 nm, 33mW power), micro Raman system (Uni-G2D, Uninanotech; low noise 532 nm Diode-pumped solid-state (DPSS) laser (3.7 mW)) and a field emission transmission electron microscope (FETEM) (FEI, Tecnai G2 F20). Electron conductivity of the structure was measured using a four-probe station (MS Tech, Korea) connected to a digital multimeter (Keithley, 2400). Measurements of sheet resistance were performed after flexing the samples at different bending angles, ranging from 0° to 180°.

Electrochemical characterization of the anode structure. Prior to using the as-prepared structure as an anode for LIB, the anode was cut into a disk with 14 mm diameter and dried at 50 °C under high vacuum for 12 h to ensure complete water removal. For the electrochemical performance tests, a coin cell (CR 2032) assembly was carried out in an argon filled glove box under extremely low level of humidity (<1 ppm). The 3D MWCNTs-graphene-PET structure was used as working electrode, whereas Li metal foil as both counter and reference electrodes. PET acted as a flexible substrate for the 3D MWCNTs-graphene, however due to its insulating nature, a crescent shaped Cu current collector was used to make electrical contact between the active 3D MWCNT materials and bottom of coin cell. 1 M solution of $LiPF_6$ salt in 1:1 (v/v) mixture solvent of ethylene carbonate (EC) and dimethylene carbonate (DMC) was used as electrolyte. Typical polypropylene (PP) material (Separator 2400) was used as separator. The cell was finally assembled by using a CR 2032 coin-cell crimping tool (Honsen Corp.). The charge and discharge tests of the assembled cells were performed in a multichannel battery testing unit (MACCOR) at room temperature in galvanostatic mode (constant current density).

Structural properties of 3D MWCNTs-graphene-PET anodes. Fig. 12.14(a) shows that the regularly patterned 3D MWCNT structure is successfully transferred and uniformly distributed over the graphene-PET film. From the two spots of 3D MWCNTs on graphene-PET film, EDS spectra represent higher than 99 atom% of carbon and less than 0.2 atom% of copper. Zoomed-in SEM plane view of the transferred 3D MWCNTs on PET structure is shown in Fig. 12.14(b). When the MWCNTs were transferred over the PET film, the 3D MWCNT architecture was retained and turned the structure upside down. The SEM image of the darker region (not the brighter box regions) represents the bottom side of the 3D MWCNT architecture directly grown on the original Cu mesh substrate. Once pressed by the laminator, the MWCNTs in the upper region were filled into the free spaces (i.e. the brighter box regions in the SEM image corresponding to the holes in the original Cu mesh), presenting their randomly oriented morphology (see Fig. 12.14(b)). As the areal mass density of MWCNTs increases, the more MWCNTs are filled, creating high density pillars and the transparency of the film accordingly decreases.

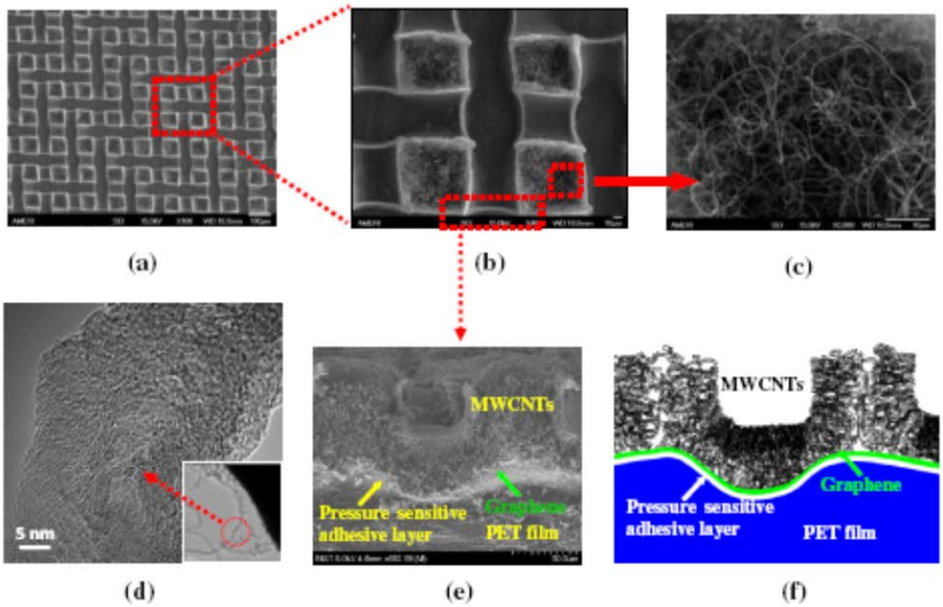


Fig. 12.14. Structural analysis of as-fabricated 3D MWCNTs-graphene-PET film. (a) The SEM plane view of 3D MWCNTsgraphene-PET film, (b) Zoomed-in SEM image showing detailed morphology of the structure, (c) The magnified SEM image of the randomly oriented MWCNT structures in the patterned structure, (d) HRTEM image of the MWCNTs grown on 3D Cu mesh showing a herringbone-like internal structure, (e) The cross-sectional SEM image of 3D MWCNTs-graphene-PET film, (f) Schematic diagram (not to scale) of the cross-sectional image shown in Fig. 12.14(e).

The areal density of the tested MWCNT based flexible anodes showed from 0.65 to 1.75 mg/cm^2 [148]. Compared with that of MWCNTs-polycarbonate (PC) film fabricated by a small scale and costly laser-jet patterning technique and a rolled quartz cylinder, structure of patterned 3D MWCNTs-graphene-PET film is unique, considering the following two features [170]. Firstly, graphene layers were incorporated into the heterogeneous two materials MWCNTs and PET and served as electric conducting agent. Secondly, 3D structure of MWCNTs-graphene-PET film was created by simple lamination process of MWCNTs grown on large scale 3D Cu mesh [170]. The Cu mesh is commercially available, easily etched away by acid solution, low cost, scalable and suited to a roll-to-roll process. Furthermore, MWCNTs and graphene layer could be physically bonded through high pressure-sensitive adhesive not chemical reaction which are different from the past reports (e.g. layer by layer assembly [171, 172], a self-assembly [173] and intercalated growth of vertically aligned carbon nanotubes (VACNTs) into thermally expanded highly ordered pyrolytic graphite (HOPG) by the pyrolysis of iron phthalocyanine (FePc) [174]). The promising advantages of PET as a substrate are its flexibility, good adherence with active materials, and ease of transferring of thin layers like graphene without structural degrading for its application to the flexible electronic devices. Some researchers previously used PET as a substrate to fabricate a flexible electrode for supercapacitors [153, 170]. As presented in Fig. 12.14(c), the high magnification SEM image shows the randomly oriented MWCNT structure grown by thermal CVD and their diameters are in the range of 200 nm to 300 nm which are relatively higher than those in the previous results [167, 168]. Nevertheless, damage of the PET films by electron beam irradiation limited the observation of the transferred graphene on PET. In addition, we selected an MWCNT with smaller diameter around 30 nm among the randomly oriented MWCNTs (see the inset TEM image of Fig. 12.14(d)) for HRTEM analysis of the internal structure of MWCNTs. The structures of 3D MWCNTs are correlated to the diameter of MWCNTs mainly governed by the size of the Ni catalytic islands. When the diameter is around 50 nm, the internal structures of 3D MWCNTs are composed of lower ordered graphene layer stacking with central tube structure (e.g. herringbone-like structure), as shown in Fig. 12.14(d). On the contrary, most 3D MWCNTs of our flexible anode system comprised lower ordered graphene layer stacking (e.g. turbostratic structure). As illustrated in Fig. 12.14(e), the 3D MWCNTs are strongly bonded to the PET film with their average height around 40 lm and such adhesion is due to the intrinsic adhesion layer coated on the commercial PET film. Although polymer binders or adhesives have been used to enhance bonding strength between nanomaterial based film and polymer flexible substrates, these additives can also decrease in the electric conductivity of the film and deteriorate the cyclability of LIB [154]. Thus, our binder-free transfer method can result in not only more simple fabrication process of LIB anode but also more reliable LIB performance of the fabricated anode. The bonding strength of graphene and CNTs is negligibly weak so that the adhesion layer can physically attach the both non-attractive materials. During the thermal lamination process, the adhesion layer acted as a glue between MWCNTs and graphene on PET film. To illustrate the good adhesion of graphene and CNTs, Fig. 12.14(d) shows the adhesion layer conforming the overall contact area of the transferred MWCNTs. Also, Fig. 12.14(e) is an interpretable schematic diagram representing such pseudosinusoidal conformation curve with the white line. HRTEM image reveals that graphene layer stacking inside a MWCNT of ~300 nm diameter is randomly oriented (turbostratic structure), forming internally filled layer structure instead of being aligned parallel to each other (co-centric tube structure).

Structural integrity without defilement and retention of the electrical property of active materials are essential for their flexible LIB performance. In spite of its advantages such as flexibility and lightweight, PET film cannot be effectively used as a current collector for the 3D MWCNTs-graphene- PET structure due to the insulating nature of PET. However, 3D MWCNT networks in this study showed high electric conductivity (low sheet resistance ~95 Ω/sq), which was comparable to ~40–50 Ω/sq obtained from the highly SWCNT-spray coated on PET substrates [153]. Therefore, 3D MWCNT networks can be used as both current collectors and electrochemical active materials. Furthermore, true flexible nature of the structure can be proved from sustainable properties even after severe bending. To examine the requirement, the sheet resistance of the 3D MWCNTs-graphene-PET structure was measured as a function of bending angle. As a result, even after severe flexing around 180°, the normalized sheet resistance with respect to the structure without bending (R_{deg}/R_o) was around 1 (see Fig. 12.15). Such excellent ability to preserve structural integrity and electrical property is attributed to the strong adhesion of the MWCNTs to the graphene-PET films and the macroscopic, pliable, porous morphology of the 3D MWCNT structure, releasing the bending strain. In many cases, MWCNTs and graphene are characterized by two main features which are the G mode, coming from the first-order scattering of the E_{2g} phonon of sp^2 C atoms (usually observed at ~1580 cm^{-1}) and the D mode, a breathing mode of k-point photons of A1g symmetry (~1350 cm^{-1}) [156]. Raman results also shows the typical G and D peaks presented in Fig. 12.16(c), indicating that graphene and MWCNTs were successfully grown on Cu and transferred over PET. In addition, the MWCNTs contained defective structure according to the intensity ratio of the D and G mode peaks (I_D/I_G ~1) which was due to the relatively lower growth temperature (750 °C). Structural characterization is consistent with the herringbone and turbostratic internal structures of the 3D MWCNTs as illustrated in Fig. 12.14(d). The similar structure was also reported in our past publication [166].

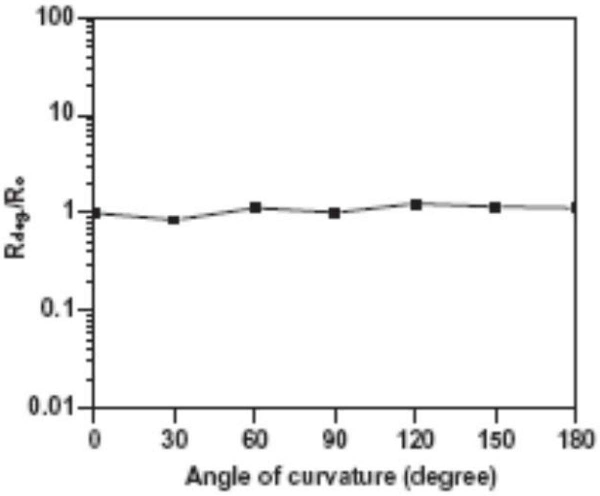
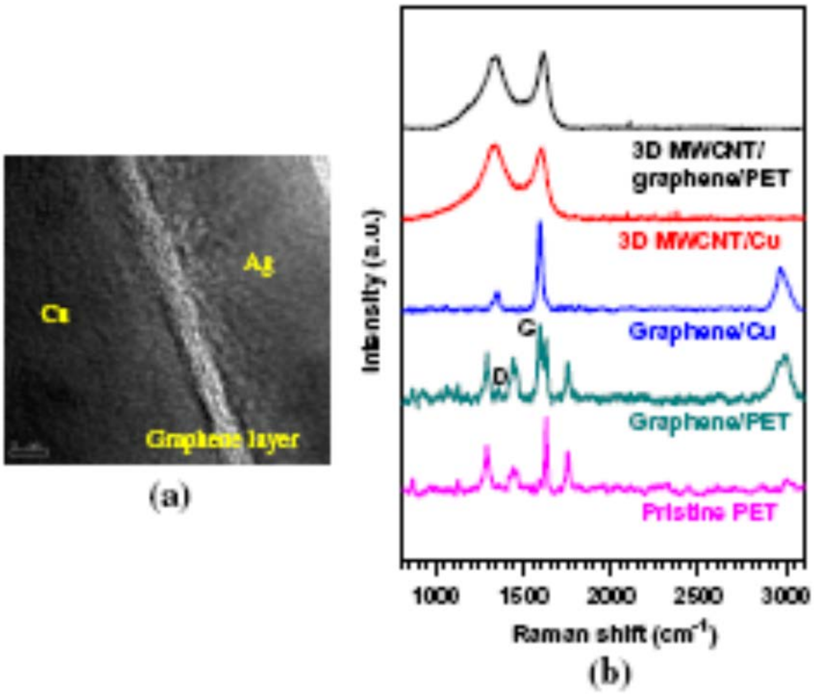


Fig. 12.15. Variation of resistance (R_{deg}/R_o) as a function of the bending angle of curvature. (R_{deg} is sheet resistance with angle of curvature, and R_o is sheet resistance with zero degree angle of curvature.)



(a)

(b)

Fig. 12.16. (a) HRTEM image of graphene on Cu substrate showing a few layer of graphene is grown on Cu, and (b) Raman spectra of 3D MWCNTs and graphene on two different substrates Cu foil and PET film and pristine PET film.

Li-ion battery performance of 3D MWCNTsgraphene-PET anodes. To characterize the electrochemical performance of the 3D MWCNTs-graphene-PET anode, a coin cell was fabricated by using Li metal foil as reference and counter electrode (see Fig. 12.17). Fig. 12.17(a) shows the initial charge and discharge cycle, measured at a current rate of 63 mAg^{-1} (0.17C). The primary lithiation and delithiation capacities are 279 $mAhg^{-1}$ ($LiC_{8.0}$) and 249 $mAhg^{-1}$ ($LiC_{9.0}$), respectively. The values of the discharge and charge specific capacities are lower than those obtained from 3D MWCNTs on Cu mesh anode systems [166], but they are higher than the reversible capacity (110 $mAhg^{-1}$) from the recently demonstrated flexible LIB comprising CNT-cellulose paper electrode [151]. In addition, the low sheet resistance (95 Ω/sq) of the anode materials plays an important role in activating the electrochemical redox reactions by improving electron transfer property between current collector and MWCNTs.

Nevertheless, the primary reason for the relatively higher irreversibility is related to the electrolyte decomposition and the solid electrolyte interface (SEI) formation on the surface of MWCNTs occurred around at 0.8 V, resulting in a plateau region (see Fig. 12.17(a)). This phenomenon is typically observed in the carbon nanostructure-based anodes as in the previous reports [162, 175]. Graphene is known as a high potential anode material for LIB based on the excellent LIB performance in past research [176]. However, the graphene layer lying between 3D MWCNTs and PET film was too thin and light to be considered a practical, electrochemical active material for lithium ion insertion and extraction. In the whole charge and discharge cycles, Coulombic efficiency was higher than 99%, indicating the high reversibility of MWCNT anode structures. As the main reason for the excellent result, electrolyte could effectively penetrate into and spread over the 3D MWCNTs' structures so that lithium ions could be facilely intercalated and decalated into and from the 3D MWCNTs. The relatively higher irreversible capacity at the first charge–discharge cycle (see Fig. 12.17(a)) is considered a critical limitation of carbonaceous anode materials and the capacity loss may be mainly associated with solid electrolyte interphase (SEI) or a permanent alloy formation in defect sites residing in the MWCNTs and their entangled structures [177]. The larger surface area of MWCNTs of 268m^{2}g^{-1} may cause more electrolyte decomposition, which result in larger SEI formation [178]. However, the high Coulombic efficiency was observed after the first cycle, even in the high rate of 1.7C, which is attributed to the porous structure of the 3D MWCNTs anode. In this scenario, lithium transfer reactions may directly happen with the outer walls of MWCNTs. It is also noted that the interface of 3D MWCNTs and the grapheme demonstrated excellent structural integrity during the C-rate cyclability test. Also, the pressure-sensitive adhesives could remain during the charge and discharge cycling. If the interface or the pressure-sensitive adhesive layer were damaged by electrochemical reaction of electrolyte, the structural integrity of MWCNTs on graphene/PET might be deteriorated, leading to lower electrochemical performance. However, it was found that the structure of graphene/3D MWCNTs showed good structural integrity after long cycling test, which manifests that the adhesion of the graphene and 3D MWCNTs were remained stable during the cell performance.

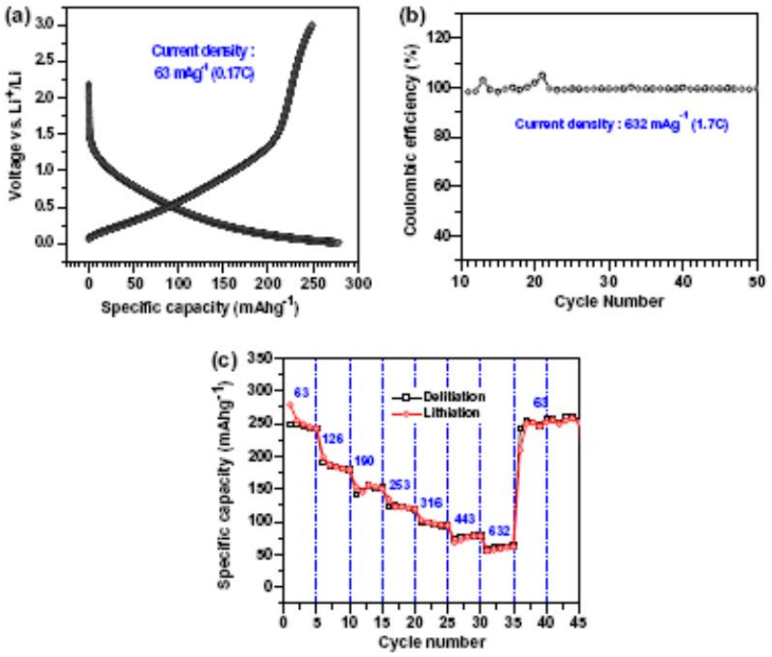


Fig. 12.17. Electrochemical performance of the 3D MWCNTs-graphene-PET as a flexible LIB anode structure. (a) Initial galvanostatic charge and discharge profiles of 3D MWCNTs-graphene-PET film anode structure, (b) Coulombic efficiency of the anode at 1.7C, (c) C-rate capability of the anode at seven different current densities. Numerical values represent current densities, at which charge and discharge cycles were conducted.

Another essential performance of the LIB anodes is the rate capability. The values of specific capacities were measured at different current densities, as illustrated in Fig. 12.17(c). A typical staircase type of decreasing behavior is shown from 254 $mAhg^{-1}$ to 59 $mAhg^{-1}$ with the increase of the rate from 0.17C to 1.7C, respectively. Such trend can stem from slower rate of charge transfer reaction between lithium ions and electrons through the anode structure at the higher C-rate condition. However, the specific capacity retains to earlier ~254 $mAhg^{-1}$ (at 0.17C) at the 37th cycle even after undergoing the high rate till 1.7C, which indicates the persistent and enormous capability of withstanding different C-rate of the flexible LIB anode. The reversible stable specific capacity is comparable to the value of the free-standing MWCNT film (300 $mAhg^{-1}$ in the 40th cycle at 0.07C) [179]. The ability of the flexible anode to withstand against bending angle till ~180°, a relatively simple transfer process without distinct structural variation, and the good LIB performance reveal the effectiveness of the flexible 3D MWCNTs-graphene-PET anode.

12.4.5. Integrated Self-powering Systems

Self-powering systems with integrated energy conversion (e.g. solar cells) and storage (e.g. supercapacitors) devices have recently attracted more and more attention [180-182]. Using an SWNT network as the integration platform (Fig. 12.18a,b), for instance, a printable all-solid integrated device consisting of a polymer solar cell (PSC) and a supercapacitor has been developed [181]. The supercapacitor charged when the integrated device was under an illumination of 100 mW cm^{-1}, whereas discharging took place once the light source was switched off and the device was connected to a resistor as a load (Fig. 12.18a,b). Fig. 12.18c,d shows the voltage and current profiles versus time of the integrated system during the charging and discharging process, respectively, with a capacitance of 28 F g^{-1} [181]. Integrated devices containing an energy conversion unit and a supercapacitor in a single wire have also been devised for wearable electronics [183, 184]. Specifically, Fig. 12.18e–g shows an integrated self-powering energy wire, consisting of a wire-shaped dye-sensitized solar cell (DSSC) and a supercapacitor both based on an aligned CNT fiber and modified titanium wire as the two electrodes, for simultaneous energy conversion and storage [184]. The supercapacitor in this integrated device can be rapidly charged to a voltage (Fig. 12.18h) close to the open-circuit voltage of the DSSC upon light irradiation. The entire energy conversion and storage efficiency up to 1.5% was obtained by multiplying the energy conversion efficiency of the solar cell part and the energy storage efficiency of the storage part. We anticipate considerable technical challenges ahead but also believe many more developments will be made over the next few years to advance the fiber-shaped integrated self-powering systems for wearable electronic applications.

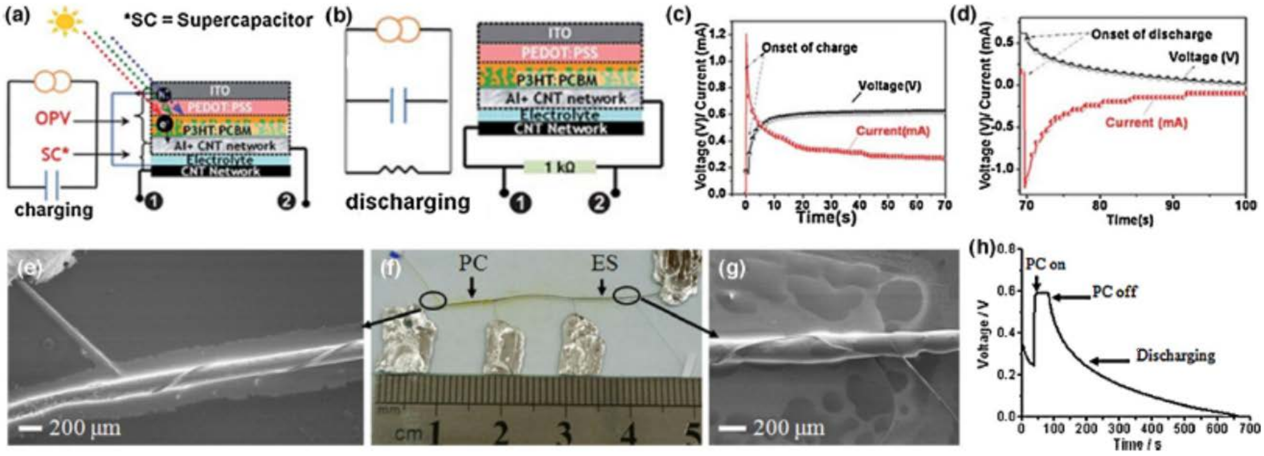


Fig. 12.18. Structures and properties of integrated self-powering devices in planar and fiber forms. (a,b) Schematic and circuit illustrations of a PSC during the charging (a) and discharging (b) process. (c,d) The voltage and current profiles versus time for the charging (c) and discharging (d) process of a PSC [181]. (e) Photograph of a typical integrated wire-shaped device consisting of photoelectric conversion (PC) and energy storage (ES). (f) SEM image of the end PC part. (g) SEM image of the end ES part. (h) A typical photocharging–discharging curve of the integrated energy wire. The discharging current is 0.1 mA [184].

12.5. CONCLUSION AND PERSPECTIVES

With global energy consumption and CO_2 emissions increasing exponentially, it is crucial to develop clean and renewable energy systems and advanced energy storage devices. Nanotechnology has opened up new frontiers in materials science and engineering to meet this challenge by creating new materials and technologies for efficient energy conversion and storage. Of particular interest, carbon nanomaterials (e.g. CNTs and graphene) have been shown to be attractive for advanced energy storage applications.

The electrochemical electrodes based on 3D nanocarbon structures, as compared with conventional planar electrodes, can greatly enhance the capacity, efficiency, or stability for the energy storage and conversion devices.

Recent research and development clearly indicate that high-performance supercapacitors can be prepared by using electrodes based on vertically-aligned CNTs with opened tips, graphene sheets with tunable through-thickness π–π stacking interactions and/or edge functionalities, and 3D pillared graphene-carbon nanotube networks. Their performance can be further enhanced by coating with conductive polymers and/or metal oxide to introduce pseudocapacitance. However, it is still challenging to further improve the performance of supercapacitors based on carbon nanomaterials. On one hand, the aggregation of CNTs and/or graphene materials tends to result in a loss of surface area, and hence inferior device performance. Carbon nanomaterials with various 3D architectures (such as CNT arrays, graphene foams, and 3D pillared VA-CNT/ graphene networks), along with the 2D graphene sheet transferred onto a crumpled structure [190, 191], have been developed to prevent the aggregation. On the other hand, the relatively high cost compared to commercial mesoporous and/or activated carbon is another challenge for carbon nanomaterials to be scaled up for practical application in supercapacitors. Therefore, it is highly desirable to develop carbon nanomaterials with high charge capacity at a low cost (e.g. by ball milling).

For potential applications in portable and wearable optoelectronics, conventional supercapacitors are too heavy and bulky. To address these challenges, a few optically transparent, mechanically stretchable, and/or wearable wire-shaped supercapacitors have been developed based on limited electrode materials. The wireshaped supercapacitors and the integrated self-powering systems play important roles in the development of flexible wearable optoelectronics. Aligned CNTs and/or graphene films with high transparency, stretchability, and charge mobility are promising electrodes for transparent, flexible, and/or stretchable supercapacitors, although it is still in the initial research stage. Continued research efforts in this embryonic field could give birth to a flourishing area of supercapacitor technologies.

A novel 3D MWCNTs-graphene-PET have been demonstrated anode structure for flexible LIB. The 3D MWCNTs were directly grown on Cu substrates by using a CVD method and transferred onto graphene on flexible PET film using a thermal lamination method. According to the Raman spectra, graphene and MWCNT exhibited similar G and D peak intensities before and after their transfer over a PET substrate, which indicated no degradation of the crystallographic order during the transfer process. The low sheet resistance (~95 Ω/sq) of 3D MWCNTs-graphene-PET even after bending the structure reflected the flexibility of the anode. Also, the anode structure showed high reversible specific capacity of ~250 $mAhg^{-1}$ and good cycling stability. Considering all those features, our proposed 3D MWCNTs-graphene-PET flexible structure is expected to be a promising anode for next-generation flexible and light weight Li-ion batteries.

Although this Chapter focuses on energy devices, electrochemical electrodes based on 3D architecture of CNTs or graphene or their hybrid are also instrumental to other applications. For example, hybrid foam of graphene and CNT for selective removal of oil and organic solvent from water surface has been developed [185]. 3D graphene/metal oxide composite for effective capacitive deionization from water has been used [186]. Fluorescent graphene quantum dots from 3D graphene foam have been successfully synthesized [187]. Furthermore, electrochemical sensors based on 3D nanocarbon electrodes offer higher sensitivity or larger detection range as compared to traditional planar electrodes [188, 189]. The 3D nanocarbon substrates can also serve as supporting scaffolds to incorporate other 2D materials such as transition metal dichalcogenide nanosheets. Although with challenges, developing new strategies to controllably fabricate functional 3D electrodes while preserving the intrinsic merits of CNTs or graphene will continue to provide tremendous thrusts to energy as well as many other applications.

References to Chapter 12.

1. X. Wang, G. Sun and P. Chen. Three-dimensional porous architectures of carbon nanotubes and grapheme sheets for energy applications. Fronties in Energy Research, 2014, **2**, 1-8.

2. Sun, Y.Q., Wu, Q.O., and Shi, G.Q. (2011b). Graphene based new energy materials. *Energy Environ. Sci.* 4, 1113–1132.

3. Liu, Y.X., Dong, X.C., and Chen, P. (2012). Biological and chemical sensors based on grapheme materials. *Chem. Soc. Rev.* 41, 2283–2307.

4. Weiss, N.O., Zhou, H.L., Liao, L., Liu, Y., Jiang, S., Huang, Y., et. al. (2012).Graphene: an emerging electronic material. *Adv. Mater.* 24, 5782–5825.

5. Jariwala, D., Sangwan, V. K., Lauhon, L. J., Marks, T. J., and Hersam, M. C. (2013). Carbon nanomaterials for electronics, optoelectronics, photovoltaics, and sensing. *Chem. Soc. Rev.* 42,2824-2860.

6. Zhang, Q., Huang, J. Q., Qian, W. Z., Zhang, Y. Y., and Wei, F. (2013). The road for nanomaterials industry: a review of carbon nanotube production, post- treatment, and bulk applications for composites and energy storage. *Small* 9, 1237-1265.

7. Xie, X., Ye, M., Hu, L. B., Liu, N., Mcdonough, J. R., Chen, W., et al. (2012). Carbon nanotube-coated macroporous sponge for microbial fuel cell electrodes. *Energy Environ. Sci.* 5, 5265-5270.

8. Liu, Y. F., Ba, H., Nguyen, D. L., Ersen, O., Romero, T., Zafeiratos, S., et al. (2013b). Synthesis of porous carbon nanotubes foam composites with a high accessible surface area and tunable porosity. /. *Mater. Chem. A* 1, 9508-9516.

9. Zheng, L. X., Sun, G. Z., and Zhan, Z. Y. (2010). Tuning array morphology for high-strength carbon-nanotube fibers. *Small* 6,132-137.

10. Zhan, Z. Y., Zhang, Y. N., Sun, G. Z., Zheng, L. X., and Liao, K. (2011). The effects of catalyst treatment on fast growth of millimeter-long multi-walled carbon nanotube arrays. *Appl. Surf. Sci.* 257, 7704-7708.

11. Zhao, D. D., Yang, Z., Zhang, L. Y., Feng, X. L., and Zhang, Y. F. (2011). Electrodeposited manganese oxide on nickel foam-supported carbon nanotubes for electrode of supercapacitors. *Electrochem. Solid State Lett.* 14, A93-A96.

12. Shan, C. S., Zhao, W. J., Lu, X. L., O'brien, D. J., Li, Y. P., Cao, Z. Y., et al. (2013). Three-dimensional nitrogen-doped multiwall carbon nanotube sponges with tunable properties. *Nano Lett.* 13,5514-5520.

13. Sun, G..Z., Pan, Y.Z., Zhan, Z.Y., Zheng, L.X., Lu, J.Y., Pang, J.H.L., et. al. (2011a). Reliable and large curvature actuation from gradient-structured grapheme oxide. *J. Phys. Chem. C* 115, 23741–23744.

14. Jiang, X., Ma, Y. W., Li, J. J., Fan, Q. L., and Huang, W. (2010). Self-assembly of reduced graphene oxide into three-dimensional architecture by divalent ion linkage. /. *Phys. Chem. C* 114,22462-22465.

15. Xu, Y. X., Sheng, K. X., Li, C., and Shi, G. Q. (2010). Self-assembled graphene hydrogel via a one-step hydrothermal process. *ACS Nano* 4, 4324-4330.

16. Sun, G.Z., Zheng, L.X., Zhan, Z.Y., Zhou, J.Y., Liu, X.B., and Li, L. (2014). Actuation triggered exfoliation of grapheme oxide at low temperature for electrochemical capacitor applications. *Carbon N.Y.* 68, 748–754.

17. Xie, X., Zhou, Y. L., Bi, H. C., Yin, K. B., Wan, S., and Sun, L. T. (2013). Large- range control of the microstructures and properties of three-dimensional porous graphene. *Sci. Rep.* 3,2117.

18. Hu, C. G., Zhai, X. Q., Liu, L. L., Zhao, Y., Jiang, L., and Qu, L. T. (2013). Spontaneous reduction and assembly of graphene oxide into three-dimensional graphene network on arbitrary conductive substrates. *Sci. Rep.* 3, 2065.

19. Yang, Z. Y., Jin, L. J., Lu, G. Q., Xiao, Q. Q., Zhang, Y. X., Jing, L., etal. (2014). Sponge-templated preparation of high surface area graphene with ultrahigh capacitive deionization performance. *Adv. Funct. Mater.* 24, 3917-3925.

20. Peng, Q.Y., Li, Y.B., He, X.D., Gui, X.C., Shang, Y.Y., Wang, C.H., et.al.(2014). Graphene nanoribbon aerogels unzipped from carbon nanotube sponges. *Adv. Mater.* 26, 3241–3247.

21. Wang, X. B., Zhang, Y. J., Zhi, C. Y., Wang, X., Tang, D. M., Xu, Y. B., et al. (2013). Three-dimensional strutted graphene grown by substrate-free sugar blowing for high-power-density supercapacitors. *Nat. Commun.* 4, 2905.

22. Li, W., Gao, S., Wu, L., Qiu, S. Q., Guo, Y. F., Geng, X. M., et al. (2013). High-density three-dimension graphene macroscopic objects for high-capacity removal of heavy metal ions. *Sci. Rep.* 3, 2125.

23. Chen, Z. P., Ren, W. C., Gao, L. B., Liu, B. L., Pei, S. F., and Cheng, H. M. (2011). Three- dimensional flexible and conductive interconnected graphene networks grown by chemical vapour deposition. *Nat. Mater.* 10,424-428.

24. Dong, X. C., Xu, H., Wang, X. W., Huang, Y. X., Chan-Park, M. B., Zhang, H., etal. (2012f). 3D Graphene-cobalt oxide electrode for high-performance supercapacitor and enzymeless glucose detection. *ACS Nano* 6, 3206-3213.

25. Ren, G. F., Pan, X., Bayne, S., and Fan, Z. Y. (2014). Kilohertz ultrafast electrochemical supercapacitors based on perpendicularly-oriented graphene grown inside of nickel foam. *Carbon N. Y.* 71,94-101.

26. You, B., Wang, L. L., Yao, L., and Yang, J. (2013). Three dimensional N-doped graphene-CNT networks for supercapacitor. *Chem. Commun.* 49, 5016-5018.

27. Zhu, G. Y., He, Z., Chen, J., Zhao, J., Feng, X. M., Ma, Y. W., etal. (2014). Highly conductive three-dimensional MnC>2-carbon nanotube-graphene-Ni hybrid foam as a binder-free supercapacitor electrode. *Nanoscale* 6, 1079-1085.

28. Prasad, K. P., Chen, Y., and Chen, P. (2014). Three-dimensional graphene-carbon nanotube hybrid for high-performance enzymatic biofuel cells. *Acs Appl. Mater. Interfaces* 6, 3387-3393.

29. Ma, Y. W., Sun, L. Y., Huang, W., Zhang, L. R., Zhao, J., Fan, Q. L., et al. (2011). Three-dimensional nitrogen-doped carbon nanotubes/graphene structure used as a metal-free electrocatalyst for the oxygen reduction reaction. /. *Phys. Chem. C* 115, 24592-24597.

30. Georgakilas, V., Otyepka, M., Bourlinos, A. B., Chandra, V., Kim, N., Kemp, K. C., etal. (2012). Functionalization of graphene: covalent and non-covalent approaches, derivatives and applications. *Chem. Rev.* 112, 6156-6214.

31. Chang, H. X., and Wu, H. K. (2013). Graphene-based nanocomposites: preparation, functionalization, and energy and environmental applications. *Energy Environ. Sci.* 6,3483-3507. 27.

32. Wang, X. W., Sun, G. Z., Routh, P., Kim, D. H., Huang, W., and Chen, P. (2014). Heteroatom-doped graphene materials: syntheses, properties and applications. *Chem. Soc. Rev.* doi:10.1039/C4CS00141A.

33. Maiyalagan, T., Dong, X. C., Chen, P., and Wang, X. (2012). Electrodeposited Pt on three-dimensional interconnected graphene as a free-standing electrode for fuel cell application. /. *Mater. Chem.* 22, 5286-5290.

34. Huang, H. J., Yang, S. B., Vajtai, R., Wang, X., and Ajayan, P. M. (2014). Pt-decorated 3D architectures built from graphene and graphitic carbon nitride nanosheets as efficient methanol oxidation catalysts. *Adv. Mater,* doi: 10.1002/adma.201401877 %

35. Yin, H., Zhang, C. Z., Liu, F., and Hou, Y. L. (2014). Hybrid of iron nitride and nitrogen-doped graphene aerogel as synergistic catalyst for oxygen reduction reaction. *Adv. Funct. Mater.* 24, 2930-2937.

36. Gong, K. P., Du, F., Xia, Z. H., Durstock, M., and Dai, L. M. (2009). Nitrogen-doped carbon nanotube arrays with high electrocatalytic activity for oxygen reduction. Science 323, 760-764.

37. Xiong, W., Du, F., Liu, Y., Perez, A., Supp, M., Ramakrishnan, T. S., etal. (2010). 3-D Carbon nanotube structures used as high performance catalyst for oxygen reduction reaction. *J. Am. Chem. Soc.* 132, 15839-15841.

38. Xue, Y. H., Yu, D. S., Dai, L. M., Wang, R. G., Li, D. Q., Roy, A., etal. (2013). Three-dimensional B,N-doped graphene foam as a metal-free catalyst for oxygen reduction reaction. *Phys. Chem. Chem. Phys.* 15, 12220-12226.

39. Yong, Y. C., Dong, X. C., Chan-Park, M. B., Song, H., and Chen, P. (2012). Macroporous and monolithic anode based on polyaniline hybridized three-dimensional graphene for high-performance microbial fuel cells. *ACS Nano* 6, 2394-2400.

40. T. Chen, L. Dai. Carbon nanomaterials for high-performance supercapacitors. Materials Today, 16, 272-280 (2013).

41. M. Winter, R.J. Brodd, Chem. Rev. 104 (2004) 4245.

42. P. Simon, Y. Gogotsi, Nat. Mater. 7 (2008) 845.

43. M.D. Stoller, R.S. Ruoff, Energy Environ. Sci. 3 (2010) 1294.

44. L.L. Zhang, X. Zhao, Chem. Soc. Rev. 38 (2009) 2520.

45. W. Lu, L. Dai, in: J.M. Marulanda (Ed.), Carbon Nanotubes, In-Tech, 2010, pp. 563–589.

46. C. Meng, et al. Nano Lett. 10 (2010) 4025.

47. L. Yuan, et al. ACS Nano 6 (2012) 656.

48. H.Y. Jung, et al. Sci. Rep. 2 (2012) 773.

49. Z. Niu, et al. Small 9 (2013) 518.

50. C. Yu, et al. Adv. Mater. 21 (2009) 4793.

51. X. Li, et al. Nano Lett. 12 (2012) 6366.

52. Z. Niu, et al. Adv. Mater. 25 (7) (2013) 1058.

53. J. Bae, et al. Angew. Chem. Int. Ed. 50 (2011) 1683.

54. Z. Cai, et al. J. Mater. Chem. A 1 (2013) 258.

55. K. Wang, et al. Adv. Mater. 25 (2013) 1494.

56. L.L. Zhang, et al. J. Mater. Chem. 20 (2010) 5983.

57. K. Kierzek, et al. Electrochim. Acta 49 (2004) 515.

58. E. Raymundo-Pinero, et al. Carbon 44 (2006) 2498.

59. E. Raymundo-Pinero, et al. Adv. Mater. 18 (2006) 1877.

60. E. Frackowiak, F. Be´guin, Carbon 39 (2001) 937.

61. B. Kastening, S. Spinzig, J. Electroanal. Chem. 214 (1986) 295.

62. S.T. Mayer, et al. J. Electrochem. Soc. 140 (1993) 446.

63. I. Tanahashi, et al. J. Electrochem. Soc. 137 (1990) 3052.

64. J. Chmiola, et al. Science 313 (2006) 1760.

65. J. Huang, et al. Angew. Chem. Int. Ed. 47 (2008) 520.

66. T.E. Rufford, et al. J. Phys. Chem. C 113 (2009) 1933.

67. A.K. Geim, K.S. Novoselov, Nat. Mater. 6 (2007) 183.

68. L. Dai, et al. Small 8 (2012) 1130.

69. S. Iijima, Nature 354 (1991) 56.

70. M. Cinke, et al. Chem. Phys. Lett. 365 (2002) 69.

71. C.M. Niu, et al. Appl. Phys. Lett. 70 (1997) 1480.

72. K.H. An, et al. Adv. Funct. Mater. 11 (2001) 387.

73. E. Frackowiak, et al. J. Power Sources 97–98 (2001) 822.

74. R.H. Baughman, et al. Science 297 (2002) 787.

75. D.N. Futaba, et al. Nat. Mater. 5 (2006) 987.

76. M. Kaempgen, et al. Nano Lett. 9 (2009) 1872.

77. A. Izadi-Najafabadi, et al. ACS Nano 5 (2011) 811.

78. C. Liu, et al. Nano Lett. 10 (2010) 4863.

79. M.D. Stoller, et al. Nano Lett. 8 (2008) 3498.

80. Y. Zhu, et al. Science 332 (2011) 1537.

81. K. Novoselov, et al. Nature 438 (2005) 197.

82. C. Lee, et al. Science 321 (2008) 385.

83. J.S. Wu, et al. Chem. Rev. 107 (2007) 718.
84. W. Lu, et al. J. Power Sources 189 (2009) 1270.
85. C.S. Du, et al. Nanotechnology 16 (2005) 350.
86. S.M. Huang, L.M. Dai, J. Phys. Chem. B 106 (2002) 3543.
87. H. Zhang, et al. Nano Lett. 8 (2008) 2664.
88. L.J. Gao, et al. Solid State Commun. 146 (2008) 380.
89. W. Lu, et al. ECS Trans. 6 (2008) 257.
90. H. Chen, et al. Mater. Sci. Eng. R: Rep. 70 (2010) 63.
91. A. Thess, et al. Science 273 (1996) 483.
92. B. Zheng, et al. Nano Lett. 2 (2002) 895.
93. L. Dai, et al. ChemPhysChem 4 (2003) 1150.
94. Y. Yan, et al. Small 3 (2007) 24.
95. L. Qu, L. Dai, J. Mater. Chem. 17 (2007) 3401.
96. Q.L. Chen, et al. Electrochim. Acta 49 (2004) 4157.
97. H. Zhang, et al. J. Electrochem. Soc. 155 (2008) K19.
98. F. Huang, et al. ChemSusChem 5 (2012) 888.
99. Y. Hu, et al. Electrochim. Acta 66 (2012) 279.
100. Y. Hou, et al. Nano Lett. 10 (2010) 2727.
101. Q. Wang, et al. Adv. Funct. Mater. 16 (2006) 2141.
102. Y.T. Kim, et al. J. Mater. Chem. 15 (2005) 4914.
103. R. Van Noorden, Nature 469 (2011) 14.
104. S. Park, R.S. Ruoff, Nat. Nanotechnol. 4 (2009) 217.
105. I.-Y. Jeon, et al. Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 5588.
106. I.-Y. Jeon, et al. J. Am. Chem. Soc. 135 (2012) 1386.
107. L. Dai, Acc. Chem. Res. 46 (2013) 31.
108. J. Liu, et al. MRS Bull. 37 (2012) 1265.
109. J.L. Xia, et al. Nat. Nanotechnol. 4 (2009) 505.
110. C.G. Liu, et al. Nano Lett. 10 (2010) 4863.
111. M.F. El-Kady, et al. Science 335 (2012) 1326.
112. J. Xu, et al. ACS Nano 4 (2010) 5019.
113. Q. Wu, et al. ACS Nano 4 (2010) 1963.
114. X. Dong, et al. Carbon 50 (2012) 4865.
115. S.H. Lee, et al. Adv. Funct. Mater. 21 (2011) 1338.
116. Z.P. Chen, et al. Nat. Mater. 10 (2011) 424.
117. Y.H. Xue, et al. Angew. Chem. Int. Ed. 51 (2012) 12124.
118. X.H. Cao, et al. Small 7 (2011) 3163.
119. Y. Zhao, et al. Adv. Mater. 25 (2013) 591.
120. D. Yu, L. Dai, J. Phys. Chem. Lett. 1 (2010) 467.
121. G.K. Dimitrakakis, et al. Nano Lett. 8 (2008) 3166.

122. Y. Mao, J. Zhong, New J. Phys. 11 (2009) 093002.

123. V. Varshney, et al. ACS Nano 4 (2010) 1153.

124. Z. Fan, et al. Adv. Mater. 22 (2010) 3723.

125. R.K. Paul, et al. Small 6 (2010) 2309.

126. G.-P. Dai, et al. RSC Adv. 2 (2012) 8965.

127. V. Sridhar, et al. ACS Nano 6 (2012) 10562.

128. J. Lin, et al. Nano Lett. 13 (2013) 72.

129. F. Du, et al. Chem. Mater. 23 (2011) 4810.

130. O. Barbieri, et al. Carbon 43 (2005) 1303.

131. H.L. Wang, et al. J. Am. Chem. Soc. 132 (2010) 7472.

132. Y. Zhu, et al. Adv. Mater. 22 (2010) 3906.

133. K.S. Kim, et al. Nature 457 (2009) 706.

134. Hahm, M.G., Reddy, A.L.M., Cole, D.P., Rivera, M., Vento, J.A., Nam, J., et. al. (2012). Carbon nanotube-nanocup hybrid structures for high power supercapacitor applications. *Nano Lett.* 12, 5616–5621.

135. Niu, Z. Q., Chen, J., Hng, H. H., Ma, J., and Chen, X. D. (2012). A leavening strategy to prepare reduced graphene oxide foams. *Adv. Mater.* 24, 4144-4150.

136. Yan, Z., Ma, L. L., Zhu, Y., Lahiri, I., Hahm, M. G., Liu, Z., etal. (2013). Three-dimensional metal-graphene-nanotube multifunctional hybrid materials. ACS *Nano* 7,58-64.

137. Jiang, Y. Q., Wang, P. B., Zang, X. N., Yang, Y., Kozinda, A., and Lin, L. W. (2013). Uniformly embedded metal oxide nanoparticles in vertically aligned carbon nan- otube forests as pseudocapacitor electrodes for enhanced energy storage. *Nano Lett.* 13, 3524-3530.

138. Niu, Z.Q., Liu, L.L., Zhang, L., Shao, Q., Zhou, W.Y., Chen, X.D., et. al. (2014).A universal strategy to prepare functional porous grapheme hybrid architectures. *Adv. Mater.* 26, 3681–3687.

139. Dong, X. C., Cao Y.F., Wang, J., Chan-Park, M. B., Wang, I.H., Huang W.,et al. (2012a). Hybrid structure of zinc oxide nanorods and three dimensional graphene foam for supercapacitor and electrochemical sensor applications. *RSC Adv.* 2,4364-4369.

140. Dong, X. C., Wang, J.X., Wang, J., Chan-Park, M. B., Li, X.G., Wang, L. H., et.al. (2012d). Supercapacitor electrode based on three-dimensional graphene-polyaniline hybrid. *Mater. Chem. Phys.* 134, 576-580.

141. Ji, H.X., Zhang, L.L., Pettes, M.T., Li, H.F., Chen, S.S., Shi, L., et. al.(2012). Ultrathin graphite form: a three-dimensional conductive network for battery electrode. *Nano Lett.* 12, 2446–2451.

142. Xiao, L., Wu, D. Q., Han, S., Huang, Y. S., Li, S., He, M. Z., etal. (2013). Self-assembled Fe_2O_3/graphene aerogel with high lithium storage performance. *ACS Appl. Mater. Interfaces* 5,3764-3769.

143. Liu, J. L„ Jiang, J. B., Qian, D., Tan, G. R., Peng, S. J., Yuan, H. M., et al. (2013a). Facile assembly of a 3D rGO/MWCNTs/Fe_2O_3 ternary composite as the anode material for high-performance lithium ion batteries. *RSC Adv.* 3, 15457-15466.

144. Qiu, B. C.,Xing,M. Y.,andZhang, J. L. (2014). Mesoporous TiC>2 nanocrystals grown in situ on graphene aerogels for high photocatalysis and lithium-ion batteries. */. Am. Chem. Soc.* 136, 5852-5855.

145. Yang, S., Cai, Y., Cheng, Y. W., Varanasi, C. V., and Liu, J. (2012). Monolithic co-aerogels of carbon/titanium dioxide as three dimensional nanostructured electrodes for energy storage. *J. Power Sources* 218, 140-147.

146. Gong, Y. J., Yang, S.B., Zhan, L., Ma,L.L., Vajtai,R., and Ajayan, P.M.(2014). A bottom-up approach to build 3D architectures from nanosheets for superior lithium storage. *Adv. Funct. Mater.* 24, 125–130.

147. Cao, X. H., Shi, Y. M., Shi, W. H., Rui, X. H., Yan, Q. Y., Kong, J., et al. (2013). Preparation of MoS_2-coated three-dimensional graphene networks for high- performance anode material in lithium-ion batteries. *Small* 9, 3433-3438.

148. C. Kang, R. Baskaran, J. Hwang, B.-C. Ku, W. Choi. Large scale patternable 3-dimensional carbon nanotube –graphene structure for flexible Li-ion battery. Carbon, 68, 493-500 (2014).

149. Hu L, Cui Y. Energy and environmental nanotechnology in conductive paper and textiles. Energy Environ Sci 2012;5:6423–35.

150. Hu L,Wu H, Mantia FL, Yang Y, Cui Y. Thin, flexible secondary Li-ion paper batteries. ACS Nano 2010;4(10):5843–8.

151. Pushparaj VL, Shaijumon MM, Kumar A, Murugesan S, Ci L, Vajtai R, et al. Flexible energy storage devices based on nanocomposite paper. Proc Natl Acad Sci 2007;104(34):13574–7.

152. Chen J, Minett AI, Liu Y, Lynam C, Sherrell P, Wang C, et al. Direct growth of flexible carbon nanotube electrodes. Adv Mater 2008;20(3):566–70.

153. Kaempgen M, Chan CK, Ma J, Cui Y, Gruner G. Printable thin film supercapacitors using single-walled carbon nanotubes. Nano Lett 2009;9(5):1872–6.

154. Hu L, Choi JW, Yang Y, Jeong S, Mantia FL, Cui LF, et al. Highly conductive paper for energy-storage devices. Proc Natl Acad Sci 2009;106(51):21490–4.

155. Wang XX, Wang JN, Chang H, Zhang YF. Preparation of short carbon nanotubes and application as an electrode material in Li-ion batteries. Adv Funct Mater 2007;17(17):3613–8.

156. Fan ZJ, Yan J, Wei T, Ning GQ, Zhi LJ, Liu JC, et al. Nanographene-constructed carbon nanofibers grown on graphene sheets by chemical vapor deposition: highperformance anode materials for Lithium Ion batteries. ACS Nano 2011;5(4):2787–94.

157. Zou G, Zhang D, Dong C, Li H, Xiong K, Fei L, et al. Carbon nanofibers: synthesis, characterization, and electrochemical properties. Carbon 2006;44(5):828–32.

158. Guo P, Song HH, Chen XH. Electrochemical performance of graphene nanosheets as anode material for lithium-ion batteries. Electrochem Commun 2009;11(6):1320–4.

159. Chan CK, Ruffo R, Hong S, Huggins RA, Cui Y. Structural and electrochemical study of the reaction of lithium with silicon nanowires. J Power Sources 2007;189(1):34–9.

160. Poizot P, Laruelle S, Grugeon S, Dupont L, Tarascon JM. Nanosized transition-metal oxides as negative-electrode materials for lithium-ion batteries. Nature 2000;407:496–9.

161. Landi BJ, Ganter MJ, Cress CD, DiLeo RA, Raffaelle RP. Carbon nanotubes for Lithium Ion batteries. Energy Environ Sci 2009;2:638–54.

162. Winter M, Moeller KC, Besenhard JO. Carbonaceous and graphitic anodes. In: Nazri GA, Pistoia G, editors. Lithium batteries: science and technology. New York: Springer; 2004. p. 144–94.

163. Gogotsi Y, Simon P. True performance metrics in electrochemical energy storage. Science 2011;34(6058):917–8.

164. Zhang H, Yu X, Braun PV. Three-dimensional bicontinuous ultrafast-charge and-discharge bulk battery electrodes. Nat Nanotechnol 2011;6:277–81.

165. Yan Z, Ma L, Zhu Y, Lahiri I, Hahm MG, Liu Z, et al. Threedimensional Metal–Graphene–Nanotube multifunctional hybrid materials. ACS Nano 2013;7(1):58–64.

166. Kang C, Lahiri I, Baskaran R, Kim WG, Sun YK, Choi WB. 3- dimensional carbon nanotube for Li-ion battery anode. J Power Sources 2012;219:364–70.

167. Lahiri I, Oh SW, Hwang JY, Cho SJ, Sun YK, Banerjee R, et al. High capacity and excellent stability of Lithium Ion battery anode using interface-controlled binder-free multiwall carbon nanotubes grown on copper. ACS Nano 2010;4(6):3440–6.

168. Lahiri I, Seelaboyina R, Hwang JY, Banerjee R, Choi WB. Enhanced field emission from multi-walled carbon nanotubes grown on pure copper substrate. Carbon 2010;48(5):1531–8.

169. Verma VP, Das S, Lahiri I, Choi WB. Large-area graphene on polymer film for flexible and transparent anode in field emission device. Appl Phys Lett 2010;96:203108.

170. Cole M, Hiralal P, Ying K, Li C, Zhang Y, Teo K et al. Drytransfer of aligned multiwalled carbon nanotubes for flexible transparent thin films. J Nanomater 2012; Article ID 272960, 8 pages.

171. Yu D, Dai L. Self-assembled graphene/carbon nanotube hybrid films for supercapacitors. J Phys Chem Lett 2010;1(2):467–70.

172. Hong TK, Lee DW, Choi HJ, Shin HS, Kim BS. Transparent, flexible conducting hybrid multilayer thin films of multiwalled carbon nanotubes with graphene nanosheets. ACS Nano 2010;4(7):3861–8.

173. Kim YK, Min DH. Durable large-area thin films of graphene/ carbon nanotube double layers as a transparent electrode. Langmuir 2009;25(19):11302–6.

174. Huang S, Dai L, Mau AWH. Patterned growth and contact transfer of well-aligned carbon nanotube films. J Phys Chem B 1999;103(21):4223–7.

175. Casas CDL, LiW. A review of application of carbon nanotubes for lithium ion battery anode material. J Power Sources 2012;208(15):74–85.

176. Liang M, Zhi L. Graphene-based electrode materials for rechargeable lithium batteries. J Mater Chem 2009;19: 5871–8.

177. Guoxiu W, Xiaoping S, Jane Y, David W, Jung-ho A. Hydrothermal synthesis of carbon nanotube/cobalt oxide core-shell one-dimensional nanocomposite and application as an anode material for lithium-ion batteries. Electrochem Commun 2009;11:546–9.

178. Hiraoka T, Izadi-Najafabadi A, Yamada T, Futaba DN, Yasuda S, Tanaike O, Hatori H, Yumura M, Iijima S, Hata K. Light supercapacitor electrodes from a surface-only solid by opened carbon nanotubes with 2200 m2g_1 surface area. Adv Funct Mater 2010;20:422–8.

179. Ng SH, Wang J, Guo ZP, Chen J, Wang GX, Liu HK. Single wall carbon nanotube paper as anode for lithium-ion battery. Electrochim Acta 2005;51(1):23–8.

180. Z. Yang, et al. J. Mater. Chem. A 1 (2013) 954.

181. G. Wee, et al. Energy Environ. Sci. 4 (2) (2011) 413.

182. W. Guo, et al. Nano Lett. 12 (2012) 2520.

183. J. Bae, et al. Adv. Mater. 23 (2011) 3446.

184. T. Chen, et al. Angew. Chem. Int. Ed. 51 (2012) 11977.

185. Dong, X. C., Chen, J., Ma, Y. W., Wang, J., Chan-Park, M. B., Liu, X. M., et al. (2012b). Superhydrophobic and superoleophilic hybrid foam of graphene and carbon nanotube for selective removal of oils or organic solvents from the surface of water. *Chem. Commun.* 48,10660-10662.

186. Yin, H. J., Zhao, S. L., Wan, J. W., Tang, H. J., Chang, L., He, L. C., etal. (2013). Three-dimensional graphene/metal oxide nanoparticle hybrids for high- performance capacitive deionization of saline water. *Adv. Mater.* 25,6270-6276.

187. Ananthanarayanan, A., Wang, X. W., Routh, P., Sana, B., Lim, S., Kim, D. H., et al. (2014). Facile synthesis of graphene quantum dots from 3D graphene and their application for Fe^{3+} sensing. *Adv. Funct. Mater.* 24, 3021-3026.

188. Dong, X. C., Ma, Y. W., Zhu, G. Y., Huang, Y. X., Wang, J., Chan-Park, M. B., et al. (2012c). Synthesis of graphene-carbon nanotube hybrid foam and its use as a novel three-dimensional electrode for electrochemical sensing. *J. Mater. Chem.* 22, 17044-17048.

189. Xi, F. N., Zhao, D. J., Wang, X. W., and Chen, P. (2013). Non-enzymatic detection of hydrogen peroxide using a functionalized three-dimensional graphene electrode. *Electrochem. commun.* 26, 81-84.

190. J. Luo, et al. ACS Nano 5 (2011) 8943.

191. J. Luo, et al. ACS Nano 7 (2013) 1464.

Chapter 13. Field emission display based on carbon nanotubes

13.1. Introduction
13.2. Field Emission Displays
    13.2.1. Principle of Operation
    13.2.2. Emitter Technologies
    13.2.3. Uniformity of Emission Enhancement
    13.2.4. Emitter Array Fabrication
    13.2.5. Phosphors and Phosphor Plate Fabrication
    13.2.6. Beam Focusing
    13.2.7. Encapsulation
    13.2.8. Driver Electronics
    13.2.9. Aging/Surface Chemistry
13.3. Other Display Technologies
    13.3.1. Hybrid FED
    13.3.2. Picture Element Tubes
    13.3.3. Direct View Infrared FED
    13.3.4. AMLCD and other technologies
13.4. FED Based on Carbon nanotubes.
13.4. Summary
References to Chapter 13.

13.1. INTRODUCTION

A field emission display (FED) is, in principle, a simple device. It consists of a baseplate containing an array of addressable gated field emitters whose electrons are projected towards a phosphor plate that is positioned in close proximity (0.2–2 mm) to the array. In contrast to the cathode ray tube (CRT) where three hot filament electron guns address all of the red, green, and blue phosphor spots on the screen, in the FED, each color pixel or a small group of color pixels is addressed by a field emitter source. The virtues of the FED over other flat panel devices are higher brightness, better viewing angles, lower power consumption, and a larger range of operating temperatures. Table 13.1 compares some of the characteristics of FEDs with those of active matrix liquid crystal displays (AMLCDs), electroluminescent, and plasma displays. The table was adapted from Ref. [1] with an added column for organic light emitting diode (OLED) from Ref. [2]. The table is intended as a guide and will change as these display technologies evolve.

Table 13.1. Comparison of FED with Other Flat Panel Displays

Feature	Thin Film Transistor LCD	Electro-luminescent Display	FED	Plasma Display OLED Panels	OLED Display
Brightness (cd/m^2)	200	100	150 (low-V) >600 (high-V)	300	300
Viewing angle (degrees)	±40	±80	±80	±80	±80
Emission efficacy (lm/W)	3–4	0.5–2	1.5–3 (low-V) 10–15 (high-V)	1.0	10–15
Response time (ms)	30–60	<1	0.01–0.03	1–10	<0.001
Contrast ratio (intrinsic)	>100:1	50:1	300:1	100:1	100:1
Number of colors (millions)	16	16	16	16	16
Number of pixels	1024 × 768	640 × 480	800 × 600	852 × 480	640 × 480
Resolution (mm in pitch)	0.31	0.31	0.27	1.08	0.012
Power consumption (W)	3 (25.4)[a]	6 (25.4)	2 (25.4)	200 (106.7)	6 (15.2)
Maximum screen size in diagonal (cm)	55.9 (22)[b]	25.4 (10)	35.6 (14)	106.7 (42)	15.2 (6)
Panel thickness (mm)	8	10	10	75–100	3
Operating temperature range (°C)	0–50	−5 – +85	−5 – +85	−20 – +55	−25 – +65

[a] Values in parentheses in this row indicate panel size in centimeters.
[b] Values in parentheses in this row indicate panel size in inches.

The FED was first proposed by Crost, Shoulders, and Zinn and is described in US Patent 3,500,102 issued in 1970. The first working monochrome prototype was demonstrated at Japan Display in 1986 by a group from LETI, France [3]. The first color display (6 in.) was demonstrated by Pixel/LETI in 1993 [4]. Initially, it was conjectured that, by using standard CRT phosphor and encapsulation technologies, FEDs could be manufactured at lower cost than LCDs because of simpler processing and proven CRT technology. In reality, FED manufacturers encounter tremendous difficulties, mostly caused by the transition from the low surface-to-volume ratio CRT to a very high surface-to-volume ratio device. Issues of spacer breakdown, vacuum flash over, gettering, placement of getters, device outgassing, uniformity of emission, emitter array yield, emitter positioning, and others—all of which are of minor concern in CRT processing—are becoming major problems that have to be resolved in FEDs.

The FED has been the driving engine for the continuous development of vacuum microelectronics (VME), similar to microprocessor and memory chips in microelectronics and pressure sensors and accelerometers in microelectromechanical systems (MEMS). This chapter will describe in detail the basic FED devices that are intended for desktop applications. This will be followed by profiles of other display applications, such as electronic billboards, wall television, and infrared-to-visible light converters.

The display market is one of the fastest growing segments of the semiconductor industry. As computers and other intelligent electronics shrink to pocket size, replace traditional dialand-switch control panels with graphical user interfaces, and insinuate themselves into other products, they carry displays with them. Field emission displays appear to have potential as such an alternate technology to liquid crystal display (LCD), although much further development will be needed (Table 13.2).

Table 13.2. Field emission displays versus AMLCDs

	FED (target)	**Typical TFT-LCD**
Thickness	6-10 mm	23 mm
Weight	<0.20 kg	0.33 kg
Contrast ratio*	>100:1	60:1 to 100:1
Viewing angle	>80°	±45°H.±30°V
Maximum brightness	>200 Cd/m^2	60 Cd/m^2
Power consumption at 60 Cd/m^2	<1 W	4 W
Operating temperature	-50 to +80°C	0 to 50°C

*Assumes dim room light. Very bright and very dark environments require different design tradeoffs.

The table compares a current subnotebook-type AM LCD display with a potential FED display for the same application. Both displays would have different characteristics if optimized for aircraft or other high-performance applications.

Not only is the FPD market enormous by itself — every cathode ray tube (CRT) on every factory floor or desktop and in every hospital room would take up less space and consume less power if replaced with a flat panel — but displays are an enabling technology for many other applications. For instance, laptop computers, now the fastest growing segment of the personal computer market, were made possible by FPDs. In the not too distant future, personal digital assistants, virtual reality-driven robotics for hazardous material handling, automotive global positioning systems, and many other civil and military applications will depend on portable, efficient, high-performance displays to communicate with users. Eventually, integration of auxiliary circuits with the display may enable even more novel applications.

Unfortunately, those displays may not be available. Sharp, the world's leading producer with 55% of the market in AMLCDs, has stated that it is not interested in low-volume, customized displays, such as those required by military applications.

Despite massive investment in their development, AMLCDs do have significant limitations. Fabricated from amorphous Si on glass, they are inherently slow. Manufacturing complexity is high and yield is an ongoing problem. They require a power-hungry fluorescent backlight, then throw 90 to 97% of the light away. Worse, an AMLCD image can be difficult to see in bright sunlight or at extreme viewing angles, a major concern in aircraft and automotive applications. Research now underway promises to address some of these limitations.

FEDs have emerged as one of the most attractive alternatives to AMLCDs. At least in theory, FEDs offer reduced manufacturing complexity and cost, increased brightness, improved viewing angle, and reduced power consumption (Table 13.2). It remains to be seen which approach is more manufacturable in practice.

Unlike AMLCDs, FEDs produce their own light using colored phosphors. Therefore, they don't need complicated, powerconsuming backlights and filters: almost all the light generated by an FED is visible to the user. Further power savings result because, like any emissive display, no power is consumed by pixels in the off state. Since text only uses about 5% of the pixels in a display, and even video averages only 20% "on" time, compaisons between power consumption of a fully on FED display and an AMLCD backlight can be misleading.

Phosphors are inherently faster and more persistent than liquid crystal, and field emission devices provide a threshold voltage for emission. Therefore, video-rate displays are possible using simple matrix addressing schemes similar to those employed in passive matrix LCD displays. The AMLCD requirement for large arrays of thin-film transistors (TFTs) - a major source of yield problems - is eliminated. On the other hand, FED designs based on arrays of microtips replace TFT arrays with micron-scale machining over large areas, which presents other yield problems.

13.2. FIELD EMISSION DISPLAYS

13.2.1. Principle of Operation

In principle, an FED is similar to a conventional CRT. Electrons are liberated from a cathode and impinge on phosphors on a transparent faceplate to produce the image. This cathodoluminescent process is one of the most efficient means known for generating light. Unfortunately, with a CRT's single electron gun and its distance from the screen, high vacuum is necessary to ensure a sufficiently long mean free path for the electrons. This, in turn, requires the use of strong (and heavy) faceplate glass. So, the volume and weight of the tube increase as the cube of the screen diagonal.

In an FED, each pixel has its own electron source, typically an array of emitting microtips (Fig. 13.1). The voltage difference between the cathode and gate extracts electrons from the cathode (Fig. 13.2). As the Fowler-Nordheim equation indicates, the emission current - and thus the display brightness - is strongly dependent on the work function of the emitting material. Cleanliness and uniformity of the emitter surfaceare are essential. All materials used in fabrication of thedisplay must be extremely clean. Sulfur deposition from conventional phosphors on the emitters can cause tip degradation, so sulfur-free phosphor materials, such as yttrium oxysilicate (Y_2SiO_5) and gadolinium oxysulfide (Gd_2O_2S), have been considered.

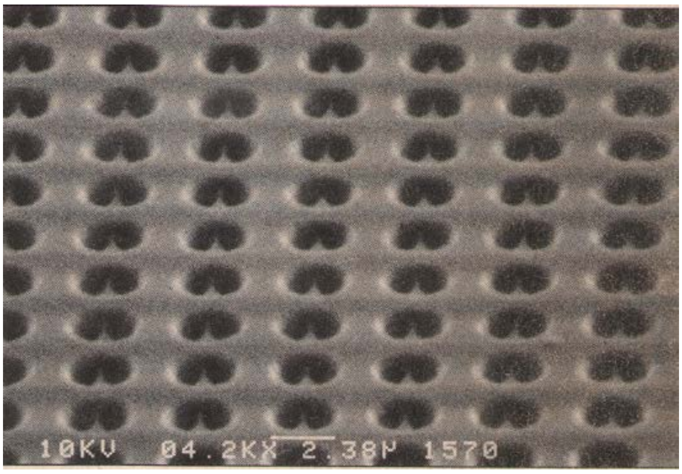


Figure 13.1. An array of field emitters. Depending on the design, as many as 1600 emitters may make up a single visible pixel.

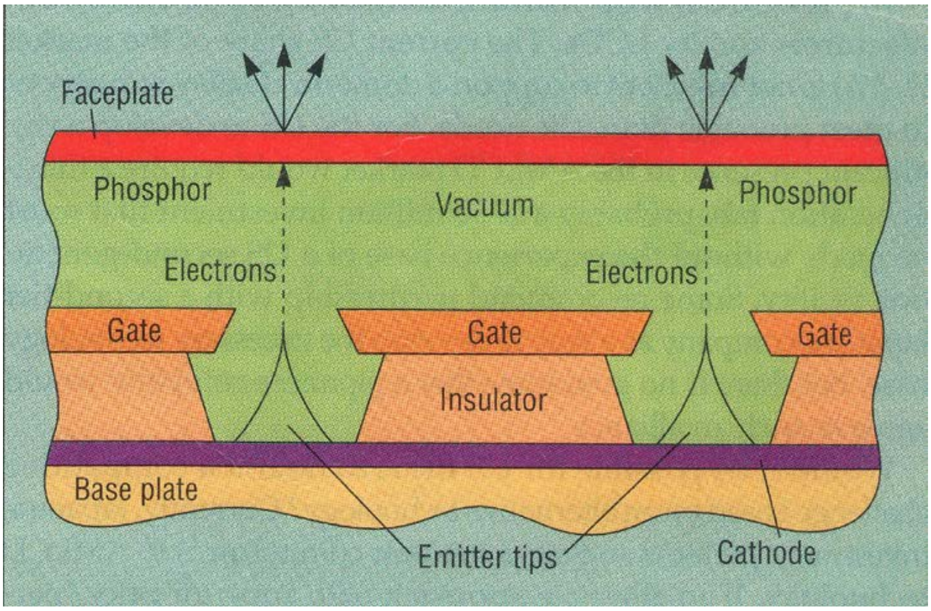


Figure 13.2. Cross-sectional view of field emission cathodes. Voltage applied to the gate controls emission from the tips.

Most FEDs must be evacuated to low pressures (about 10^{-7} torr) to provide a long mean free path for emitted electrons and to prevent contamination and deterioration of the tips. Large display screens will bend at such low pressures unless built with heavy CRT glass or held apart with spacers. The design need for display thinness thus leads to important tradeoffs. With minimal display thickness, spherical spacers will be small enough not to be noticed by the eye. However, a small standoff distance between gate and screen requires a very low anode voltage, or the anode voltage would overpower the gate and simply turn the entire display on. As the emission of microtips is a 60° cone, pulling the gate and tips back from the screen to permit the use of higher anode voltages (and the resulting greater brightness) will result in a very large spread of the electron beams – and much lower picture resolution.

Picture resolution is improved by use of a focus grid to collimate the electrons drawn off the microtips. This permits the use of larger standoff distances and thus, efficient high-voltage phosphors running at high brightness, yet permits accurate focusing on the subpixel level. This large spacing, however, vastly compounds the spacer problem.

Triode Structure. FEDs can be operated as either diode or triode devices with either low-voltage (<2000 V) or high-voltage phosphors (2000–8000 V).

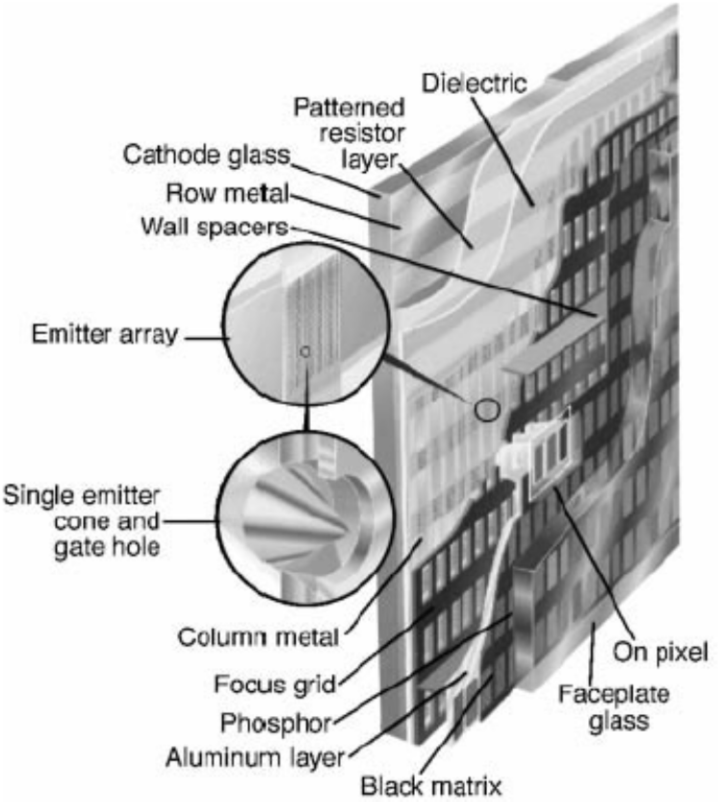


Figure 13.3. Perspective view of a high-voltage phosphor FED architecture, including cut-away enlargements of a subpixel emitter array and a single gated tip [5].

Diode devices are usually fabricated using planar carbon-type emitters, and triode devices use Spindt-type gated emitters fabricated in molybdenum or silicon. Other triode approaches such as surface conduction electron emitters and gated edge emitters have been proposed and used. At present, the most mature device is the Spindt-type triode device using highvoltage aluminized phosphors. Figure 13.3 shows a perspective cutaway drawing of such a device [5]. The glass baseplate consists of row metal lines upon which a patterned resistor layer is fabricated. This layer is necessary for obtaining uniform emission per subpixel array by providing a voltage drop to the emitters that is proportional to the emission current. Thus, the gate voltage is reduced by the largest amount for the dominant emitters. On top of the resistor film, the gate dielectric is deposited, followed by the deposition of the gate material. Gate column lines are then formed, followed by the processing of gate openings and tip formation. A gated tip is shown in the enlargement on the left side of the drawing. For obtaining uniform emission, it is desirable to fabricate as many emitters per subpixel as possible (ranging from hundreds to several thousands).

The faceplate that contains the red, green, and blue phosphor regions is separated from the baseplate by spacer rails. These are needed to avoid collapse of the device by the atmospheric pressure once it is evacuated. By aluminizing the phosphors, higher brightness can be achieved, since light that would reflect back into the display from the glass/air interface is reflected towards the viewer by the aluminum layer.

A black matrix surrounding the rectangular phosphor pixels enhances the contrast of the display. By applying a positive voltage to the gate columns with respect to the emitter rows, selective pixels can be energized. The electrons generated at the tips traverse through the vacuum space, penetrate through the thin aluminum layer, and cause the phosphor pixels to light up via the electroluminescent process. If the distance between the phosphor plate and the emitters is short, the electrons will not spill over to adjacent pixels. This situation is called proximity focusing. For larger distances that are generally needed for high-voltage phosphor operation to avoid flashover of the spacers, the electron beam has to be focused to avoid illumination of adjacent pixels. In Fig. 13.3, this is accomplished by surrounding each subpixel emitter area by a conductive thick film “waffle” structure that is negatively biased with respect to the gate. Not shown in this cutaway drawing are the ancillary items such as the device seal, pumping ports, getters, electrical feedthroughs, and placement of the driver chips.

To reduce the number of emitter arrays of a device such as the one shown in Fig. 13.3 by a factor of 3 and therefore, increase the processing yield, one array can be shared with the corresponding red, green, and blue subpixels. To accomplish this, the phosphor lines must be switched, which limits the device to low-voltage phosphor operation. The first color FED prototype from Pixel/LETI utilized this scheme [6]. Figure 13.4 shows a perspective view of the switched phosphor triode device [7].

Diode Structure. The devices shown in both Figs. 13.3 and 13.4 are of the triode type, in which the phosphor voltage is constant. A diode device is shown in perspective view in Fig. 13.5 [8]. It consists of a glass baseplate onto which thin film emitter lines and metal spacers are fabricated. The emitters are fabricated from diamond-like carbon (DLC). Because of the fragile nature of these films and the possibility of contamination from subsequent processing steps, the emitters are deposited at the end of the processing sequence using a shadow-mask technique. The top plate contains groups of red, green, and blue nonmetalized phosphor stripes. In operation, the phosphor lines have to be switched to enable selection of a given subpixel. This limits the device to low-voltage phosphor technology, since achieving line-to-line isolation at high voltages (>2 kV) is difficult. In addition, the device would also require high-voltage addressing circuitry. In order to obtain low-voltage operation, the distance between the two plates is only about 50–100 μm. This is dictated by the turn-on field of carbon films which range from 5 to 10 V/μm. The advantage of the diode FED is that a much simpler device structure is used; the disadvantage is the need for higher voltage driver circuitry.

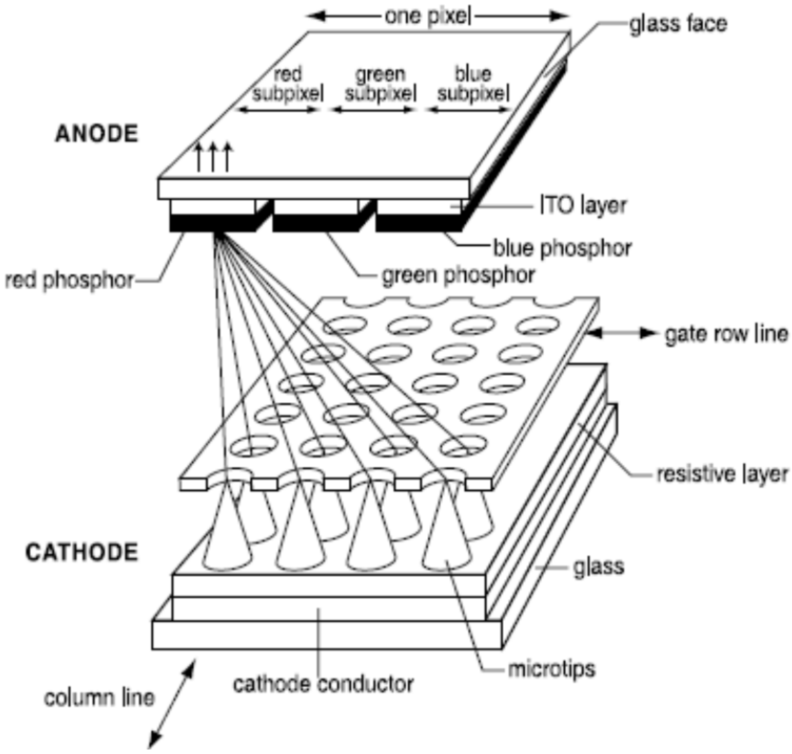


Figure 13.4. Perspective view of an FED in which one emitter array is shared by a red, green, and blue subpixel [7]. Selection of the subpixel is accomplished by biasing the appropriate phosphor

line. The drawing is not to scale. The distance between the emitters is at least 100 times larger than the tip height, meaning that the electron paths are not as distorted as shown.

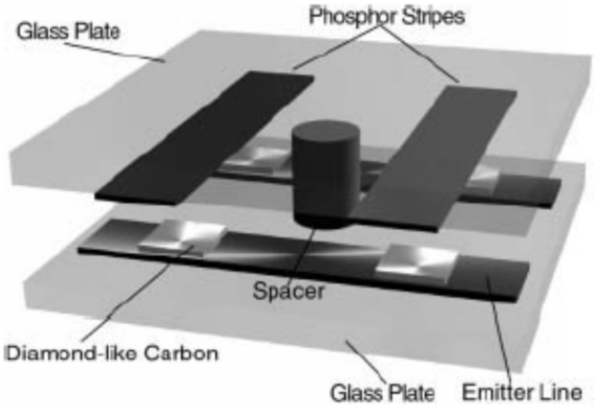


Figure 13.5. Perspective view of a diode device using DLC emitters.

13.2.2. Emitter Technologies

Generalizations about possible FED performance are difficult; the some most promising approaches all have unique advantages and disadvantages.

Metal microtip emitter (Spindt Emitters). The most prominent and widely used emitter technology for FEDs is the Spindt emitter using e-beam evaporated molybdenum cones ranging in heights from about one to several micrometers, and gate diameters ranging from 0.15 to 1.5 µm. Other materials, such as tungsten, tantalum, hafnium, metal silicides, carbides and nitrides, have been proposed, along with thin layers (nanometer thick) of insulators, semiconductors, metals, and others on the tips.

The first field emission cathodes, developed by C. A. Spindt and his coworkers at SRI International, used molybdenum microtip emitters, and Mo microtips are still a core FED technology. In a Spindt emitter array, a silicon wafer is oxidized to produce a thick SiO_2 layer (Fig. 13.6). A metallic gate layer is deposited on top of the oxide, the patterned to form gate holes. Etching the SiO_2 underneath the holes undercuts the gate and creates a well. Next, a sacrificial layer, usually Ni, is deposited at grazing incidence to prevent deposition of nickel into the emitter well. Mo is then deposited at normal incidence so that a cone with a sharp point grows inside the cavity until the hole closes above it. Finally, the sacrificial layer is removed, leaving an emitter cone.

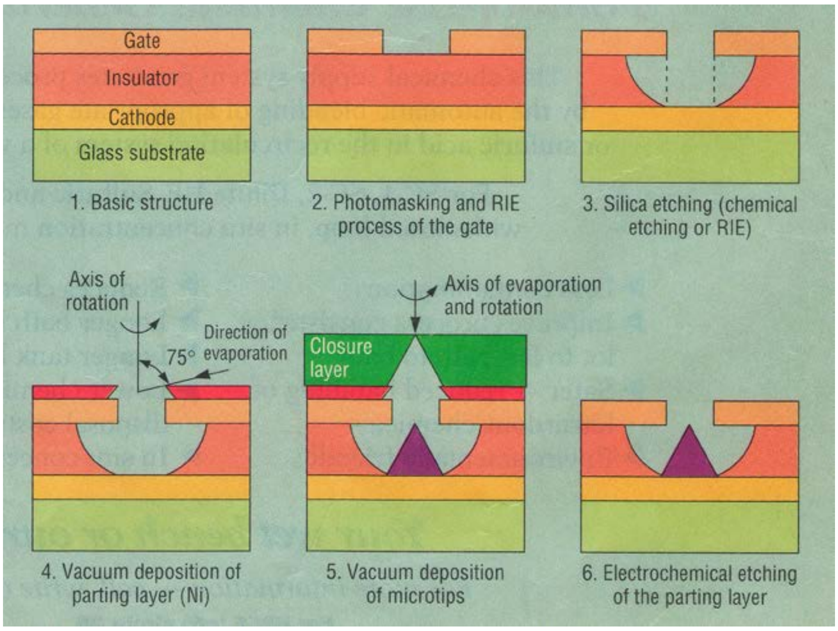


Figure 13.6. Microtip fabrication process. Anode phosphor deposition is related to CRT technology. Final assembly is done in vacuum.

Fabrication of a Mo microtip array requires micron-scale lithography over the entire area of the display. Several hundred emitter cones are fabricated for each pixel. Dozens of these can fail with no discernible loss of display brightness. This built-in redundancy may improve yield and manufacturability. If the emitter yield is not uniform over the total display area, however, nonuniform brightness will result.

Metal microtips have some of the most visible supporters among FED developers. Since individual tips may have varying I-V characteristics, the lateral resistors provide a load to limit emission current and ensure uniform brightness. They also serve as fuse links, protecting the display from the catastrophic effects of a shorted emitter by burning out and electrically isolating the defective emitter.

LETI has also developed a switched anode design, in which all the microtips for a particular pixel are cycled through the three color phosphors for that emitter during a screen refresh cycle. Thus, the average current per emitter is one-third of what would be necessary if all three phosphors were refreshed at once, and emitter lifetime is improved.

Emission results have been reported in the literature, notably in proceedings of the International Vacuum Microelectronics Conferences (IVMCs) that have been held annually since 1988 (Williamsburg, VA.) Even though some improvements in emission behavior over the molybdenum devices have been reported, none of the companies pursuing product development are using any of these materials besides molybdenum at this point. One reason is historical, since at the time when most companies launched their product development efforts, the only device for which sufficient data were available was the Spindt device. Another reason is that the molybdenum cone deviceworks well in comparison to other refractory metals and silicon. For instance, initial conditioning is not that critical as compared to silicon, which is attributed to the more conductive molybdenum oxide as compared to the native silicon oxide.

Silicon Emitters. Next in popularity is the gated silicon emitter. Silicon emitters are the other major FED approach. By using the entire arsenal of VLSI processing, an impressive body of work has been developed, ranging from pyramid to wedges, to cone, and to tip-on-post emitters. The advantage of silicon technology is that new device concepts can be proven in a relatively short period of time and in most academic institutions. By combining field emitters with novel current-limiting circuits on a silicon substrate, low-noise emitter arrays have been demonstrated.

Monocrystalline silicon, however, limits the size of FEDs to applications such as helmet displays. For larger products, in which all the emitters are fabricated on one substrate, amorphous or polycrystalline silicon cone devices are used.

In these devices, thermal oxidation of silicon is followed by patterning of the oxide and selective etching to form silicon tips (Fig. 13.7). Next, further oxidation or etching protects the silicon, sharpens the point (since Si is consumed in oxidation), and provides a sacrificial layer. An insulating layer and gate metal follow, and then the sacrificial oxide is removed [9, 10]. This oxide sharpening process uses the crystal structure of silicon to produce [111]-aligned atomically sharp tips. Of course, a polycrystalline silicon substrate is required to make use of this effect. The design illustrated in Fig. 13.7 still requires large-area lithography, but the IC industry has developed extensive silicon fabrication technology. Furthermore, self-aligned emitter designs that eliminate the lithography requirement appear to be possible [3].

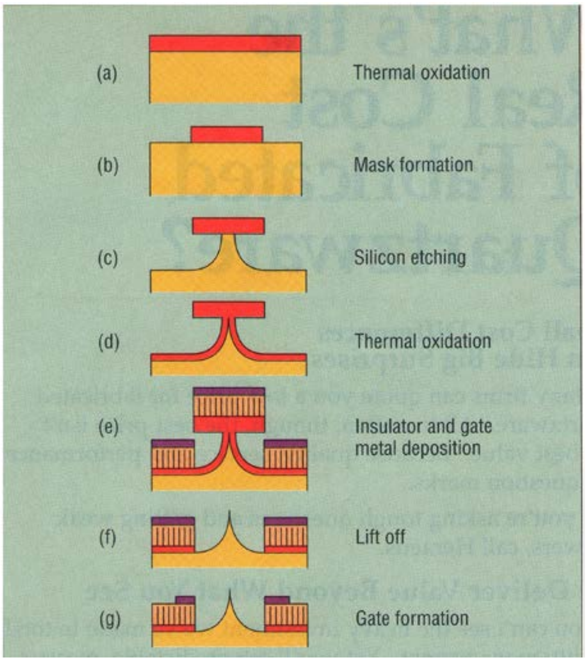


Fig. 13.7. Processing sequence to form a self-aligned silicon emitter array [9].

Diamond emitters. Diamond cathodes are the newest and perhaps the most intriguing field emission technology. SI Diamond, a Houston-based company using technology developed at the University of Texas [4], has deposited nanocrystalline diamond/graphite composite films, dubbed "Amorphic Diamond," by laser ablation of graphite. Diamond is one of the most chemlcally inert materials known, which means that diamond FEDs require only 10^{-4}-torr vacuum (required to achieve an adequate electron mean free path). Spacer design, substrate cleaning, and device evacuation and sealing are all greatly simplified. Emission currents as high as 100 mA/mm^2 have been obtained with emitters fabricated from the SI Diamond material [11].

The mechanism of emission from diamond films is not well understood. Calculation from the Fowler-Nordheim equation gives an effective work function of only 0.2 to 0.3 eV, compared to about 4.5 eV for silicon. AD emitters do exhibit greatly enhanced emission current and thus, a brighter display, without the micromachining inherent in microtip approaches.

Negative electron affinity is not observed in natural diamond. Moreover, diamond tends to form sharp facets, and the same emission current could be obtained by reducing the emitting area rather than by invoking a novel emission mechanism. Nonetheless, the emitting properties of diamond are - for whatever reason –such that simple diamond cathode stripes and phosphor-coated indium tin oxide anode stripes can be used to make an emitter array.

Such an array is extremely simple to fabricate, requiring only printed-wirmg-board-type lithography. However, diamond deposition technology is still in its infancy, and obtaining uniform emission over the entire area of the display is a challenge. The extremely high switching voltages involved in this approach (350 V) also present switching and phosphor adhesion problems.

Carbon Nanotube Emitters.

Carbon nanotubes are excellent field emitters because of their small tip size, high aspect ratio, high chemical and mechanical stability, as well as their ability to produce very high current density at low operating voltages. Carbon nanotubes, grown on the cathode of arc discharge process, were first observed by Iijima in 1991 [1]. Nowadays, CNTs can be synthesized by arc discharge [10,11], laser ablation [12], and chemical vapor deposition process [13,14], etc. Arc discharge is suitable for its throughput. The needle-like tubes comprised coaxial tubes of graphite sheets and ranged in size from 4 to 30 nm in diameter and up to 1μm in length. The discovery of carbon nanotubes has been attracting considerable attention because of their own unique physical properties and their potential for a variety of applications [2-4]. Due to their high aspect ratios and small tip radii of curvature, the nanotubes (CNTs) exhibit excellent field emission characteristics. High field emission current density of 10 mA/cm2 [5] and low turn-on electric field of 0.8 V/pm have been demonstrated [6,7]. Carbon nanotubes are excellent sources of electron beams that can be used for many applications, including displays, material analysis and various sensor technologies. Historically, the focus of electron emission from carbon nanotubes has been on large area display applications. While televisions and desktop computer monitors have traditionally used cathode-ray tubes to aim electron beams from electrically heated wires onto phosphor-coated screens, FEDs replace the CRT's single source of electrons with a large array of fine electron emitting tips. Ideally suited for large flat screens like televisions, they maintain high resolution but in a much thinner enclosure.

Thin Film Edge Emitters. Thin film edge emitter is, in principle, much simpler to fabricate, since the cone deposition process (either by etching or collimated deposition through a diminishing hole diameter) is eliminated. Some of the basic device structures are revealed in US patents 5,382,185 by Gray and Hsu and 5,874,808 by Busta and Pogemiller, and some of the initial devices were discussed in Ref. [12]. Figure 13.8 shows the perspective view of a single gated lateral device with combshaped thin film emitters [13]. Figure 13.9 shows the SEM micrograph of a 0.2 μm tungsten device with a *p*/*a* ratio (defined in Fig. 13.8) of 3.3 and an emitter-to-gate distance of 0.5 μm. Figure 13.10 shows the *I*–*V* curves of two 150-comb devices with different *p/a* ratios.

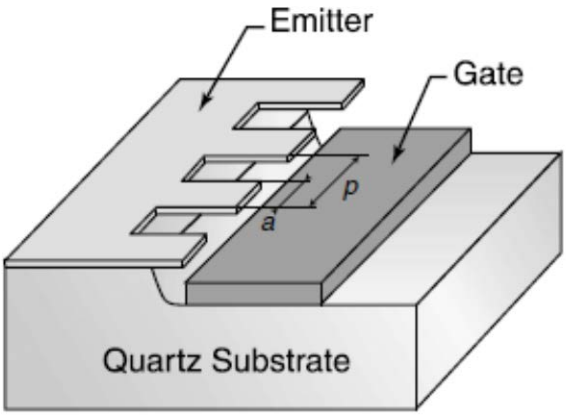


Figure 13.8. Perspective drawing of a lateral, self-aligned comb-shaped field emitter device [13].

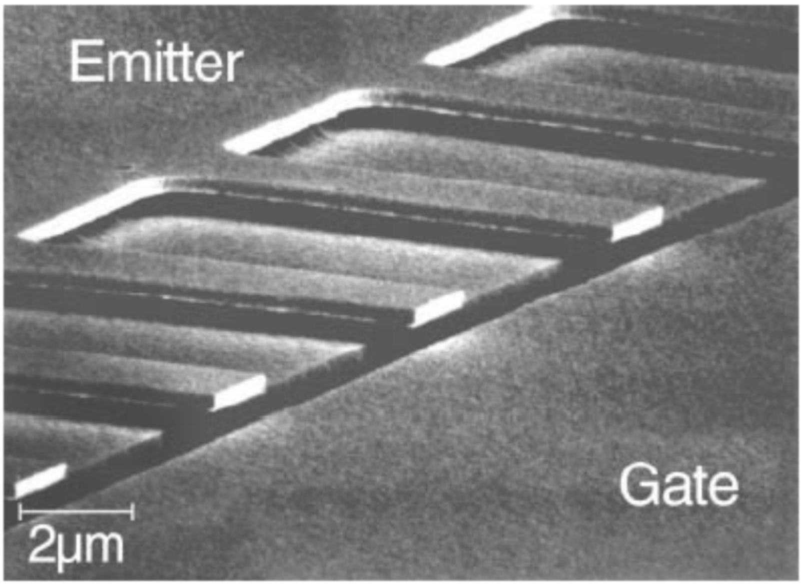


Figure 13.9. SEM micrograph of a lateral edge emitter with a pitch-to-edge width ratio (*p*/*a*) of 3.3 [13].

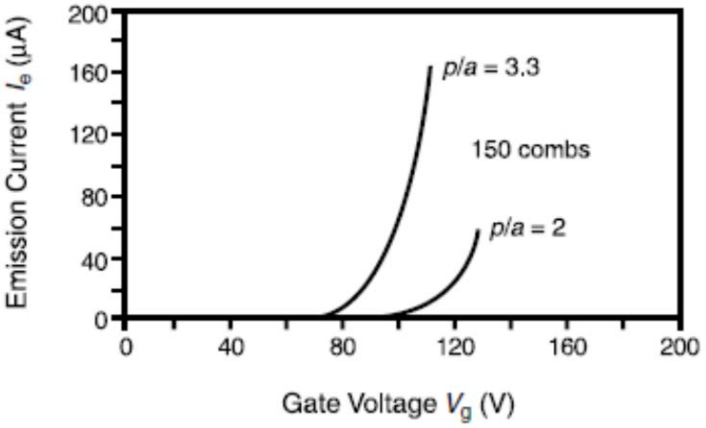


Figure 13.10. *I*–*V* characteristics of two 0.2-μm thick tungsten edge emitters with different *p*/*a* ratios [13].

By folding the combs and gates into cylindrical structures and rotating them by 90°, gated edge emitters can be fabricated. The cross section of a stepped oxide device is shown in Fig. 13.11, and the SEM micrograph of a TiW–Au emitter is shown in Fig. 13.12 [14]. By stepping the oxide near the gate plateau, the turn-on voltage can be reduced without increasing the gate-to-emitter capacitance significantly. The device in Fig. 13.12 was processed by an etch process, taking advantage of the pile-up of photoresist near the gate posts and the thinning on top of the plateau. For large-area implementation, a chemical mechanical planarization (CMP) process is preferred. Sputtered SiC emitters fabricated by CMP were described in Ref. [15]. Other edge emitters, such as a dispenser-type, vertical multilayer emitter using a low work function lithium film sandwiched between two noble metals, were described in Ref. [16]. Double-gated edge emitters, which are formed by the inclusion of another insulating and conductive layer, are shown schematically in Fig. 13.13. Such devices can avoid the excessive gate currents recorded in devices with single or asymetrically positioned gates, thus reducing the power consumption of FEDs.

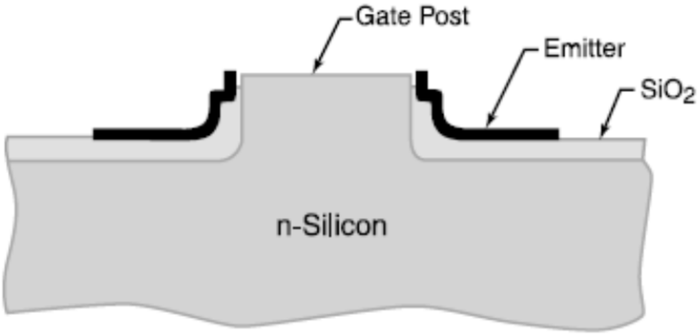


Figure 13.11. Cross section of a single gated edge emitter using a stepped oxide approach to reduce the gate-to-emitter distance near the emission area without significantly increasing device capacitance [14].

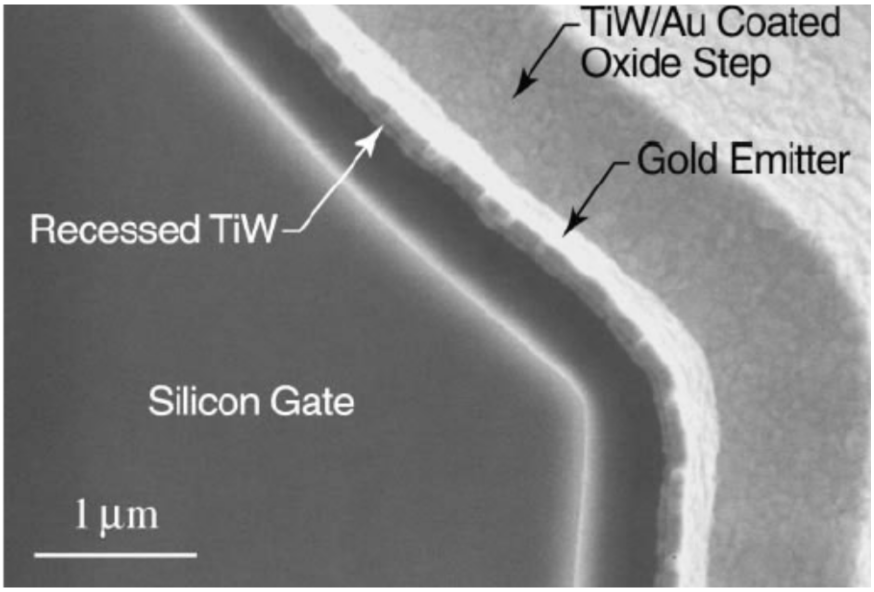
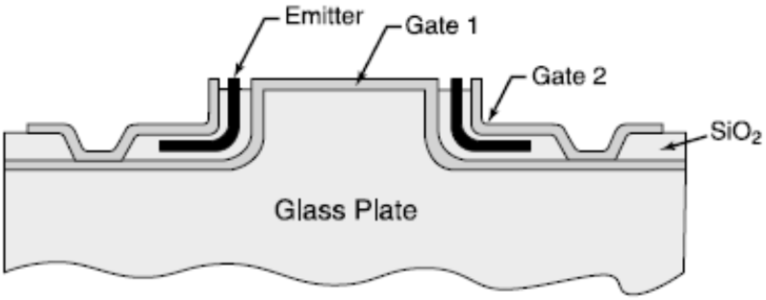


Figure 13.12. SEM micrograph of a stepped oxide TiW–Au vertical edge emitter [147.



Fsgure 13.13. Cross section of a vertical edge emitter with symmetrically placed gate.

Emission from these structures originates from sharp features at the thin film edges and is not very controlled at present. Although this development appears promising, no FED prototypes currently use thin film edge emitters. Considerable work must be done to raise the level of performance of these edge emitters to that of cone devices.

Surface Conduction Electron Emitters. Avery intriguing technology based on surface conduction electron emitters has entered the field in recent years [13,18]. The basic phenomenon was described in 1965, but had not been identified as a possible candidate for display applications until recently. The device is based on a fine particle film of PdO (5–10 nm) that is deposited between two Pt electrodes. A forming process generates a nanofissure parallel to the edges of the two electrodes that acts as an electron emission source. Electrons emitting from the fissure are scattered towards the positive side of the electrode and some eventually reach the anode. The efficiency, i.e., the ratio of anode current to electrode current, is below 1%, but still high enough for a low-power display. It is estimated that the power consumption for a 40-in. display is only 53 W.

For FED applications, the electrodes are screen-printed in air, and the PdO film is deposited via an ink jet printer. This eliminates thin film processing completely and brings this technology into the arena of plasma display manufacturing. When keeping the electrode voltage below 20 V, the electrons focus sufficiently so that no extra focusing electrodes are needed even for an emitter-to-anode spacing of 3–5 mm. A 10-in. diagonal prototype with 240 × 240 × 3 pixels and 8.6-mm panel spacing has been demonstrated using stripe-patterned P-22 (R/G/B) phosphors and a black matrix [17]. Little information is available at present about the reliability of this display. Since there is only one fissure formed per subpixel, there is no redundancy, as is the case for multiple cone emitters.

Other Emitters. Other emitters gaining popularity in recent years, such as diamond, DLC, and carbon nanotubes. The main objective in developing alternative cold electron sources for display applications is low-cost manufacturability over large areas. One such alternative might exist in printable inks consisting of semiconductive and conducting particles dispersed in an insulator-forming precursor. Such inks need to be vacuum compatible at both the fabrication stage and throughout the life of the display. Initial results yielding an emission site density of about 1×10^5 cm^{-2} at an extraction field of about 25 V/μm were described in Ref. [19]. Field emission from polymers has also been recently demonstrated [20]. By embedding transistors fabricated from the same semiconducting polymer, polyalkylthiophenes, inexpensive FEDs are being proposed (Bill Eccleston, University of Liverpool, UK).

13.2.3. Uniformity of Emission Enhancement

When refractive and semiconductive cone emitters are fabricated on conductive substrates such as thin film metal emitter lines, they typically show poor emission uniformity within an array. This is caused by the extreme sensitivity of the Fowler–Nordheim (F–N) tunneling process to small variations in tip radii, cone angles and heights, gate diameters, position of the gate holes, and the work function of the tips. A key contribution in the development of FEDs was made by introducing resistive layers between the tips and the emitter lines. Vertical resistors were proposed in Ref. [21] and resulted in significant improvements in emission uniformity. Additional improvements, especially in cases where some of the tips cause shorts between the emitter and gate metals, were obtained by the introduction of a lateral resistor mesh [22]. For silicon tips and tips fabricated on silicon substrates, active resistors in the form of on-substrate integrated field effect transistors (FETs) have been reported with excellent results.

Figure 13.14 shows the cross section of two Spindt emitters fabricated on a resistive layer. To demonstrate the effect of the resistive layer, we assume that the second tip has a field enhancement factor β that is 20% higher than the first tip. This could be caused, for instance, by a sharper tip. The current can be calculated using the F–N equation

$$I = \left(\frac{1.54\times10^{-6}\alpha E^2}{\Phi}\right)\exp\left(\frac{-6.87\times10^7\Phi^{3/2}v(y)}{E}\right) \qquad (13.1)$$

where $y = 3.79 \times 10^{-4} E^{1/2}/\Phi$, $v(y) = 0.95 - y^2$,$E = \beta V_g$, $\alpha = 4 \times 10^{-15}$ cm^2, $\beta = 8 \times 10^5$ cm^{-1}, and $\Phi = 4.5$ eV.

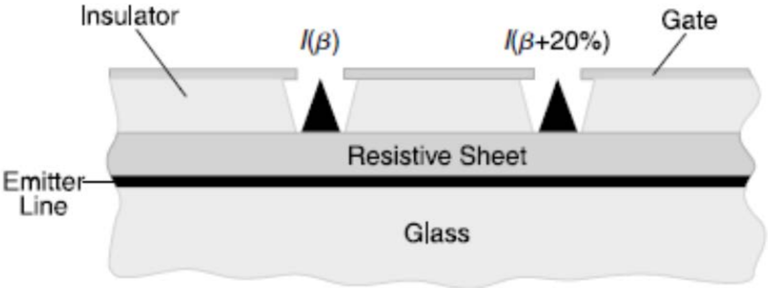


Figure 13.14. Cross section of two tips fabricated on a resistive sheet.

This is shown in Fig. 13.15 as the line designated with (R = 0, β). Assuming that the resistance between the base of the tip and the emitter line is 2MΩ, the effect of the voltage drop IR is calculated by substituting Vg with $V_g - IR$ and solving the F–N equation numerically. The result is shown as the line designated with (R =2MΩ, β). It can be seen that, at a gate voltage of 145 V, the emission current is reduced from 9 μA without the resistive layer to 5 μA with the resistive layer. Now the field enhancement factor β of the second tip is 20% higher, i.e., 9.6 × 10^5 cm^{-1}, the calculated I–V curves with and without the resistive layer are shown by the lines designated with (R = 2MΩ, β + 20%) and (R = 0, β + 20%), respectively. At 145 V gate bias, the emission current of the dominant second tip reduced from 32 μA to 11 μA by including the resistance.

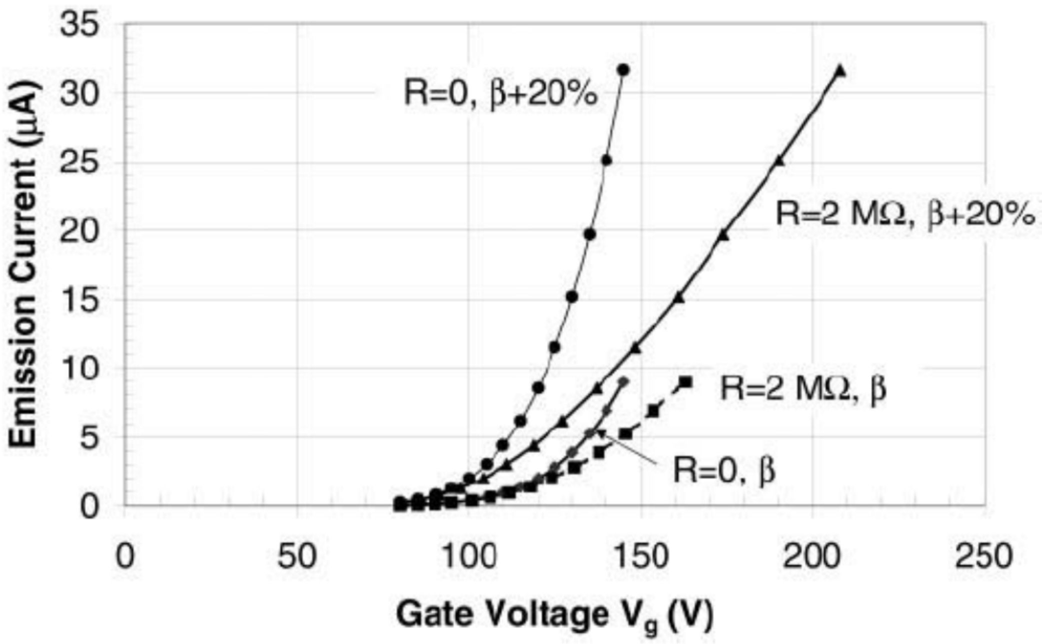


Figure 13.15. Calculated emission currents of two single-tip devices with and without a vertical resistor of 2MΩ, assuming that one of the devices has a 20% higher field enhancement factor compared to the other one.

Assuming that the base radii of these two emitters and the film thickness of the resistive layer are all 1 μm, the resistivity of the film can be calculated to be 6.3 × 10^2 Ωcm, and the sheet resistance is 6.3×10^6 Ω/□. This is a relatively high resistivity, and very little information is available concerning actual materials that are used by the FED manufacturers. Sputtered amorphous silicon, Cr cermets, and SiC are some of the candidates [23]. Care has to be taken that the voltage drops across the resistive layer do not reach magnitudes that can cause film breakdown.

To avoid the large voltage drop, and particularly to reduce the tip shorting problem, the lateral resistor mesh was introduced. In addition to increasing the emission uniformity, the lateral resistor mesh allows faster tip conditioning and results in a more rugged array [24]. However, the lateral film consumes more real estate at the expense of tip density. Figure 13.16 shows the top view and cross section of an FEA with a lateral resistor mesh onto which a 4 × 4 microtip array is positioned [25]. The bias to the emitters is supplied by a thin film metal grid, on top of which the resistive film is deposited. The gate lines are positioned perpendicular to the emitter grid. The performance of this 4 × 4 array was calculated in Ref. [26] by assuming an emission current of 1 µA from each emitter at a gate voltage of 100 V and a sheet resistance of 10^7 Ω/□ for the film.

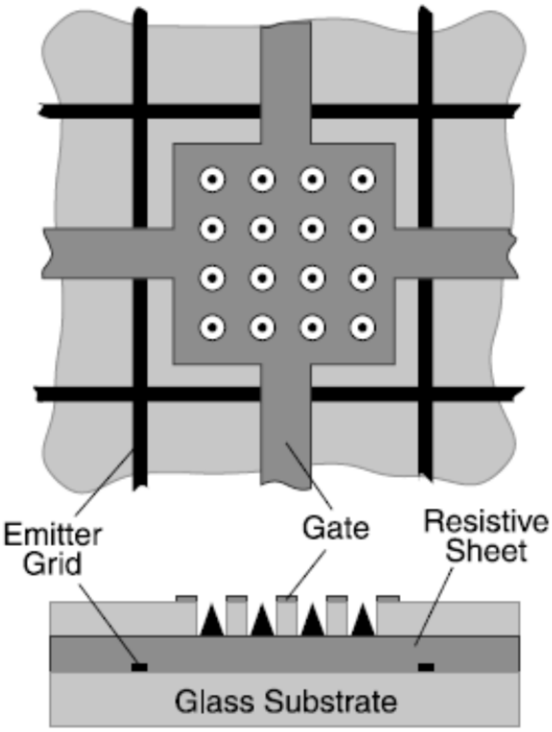


Figure 13.16. Design of a 4 × 4 emitter array with a lateral resistor film [21].

Figure 13.17 shows the voltage drops and currents for each of the emitters. The inner 2 × 2 matrix shows identical results which, for reasons of symmetry, is easily understandable. Some nonuniformity exists for the outer 12 emitters. The resistive layer causes a reduction of the total current from 16 µA to 4.53 µA.

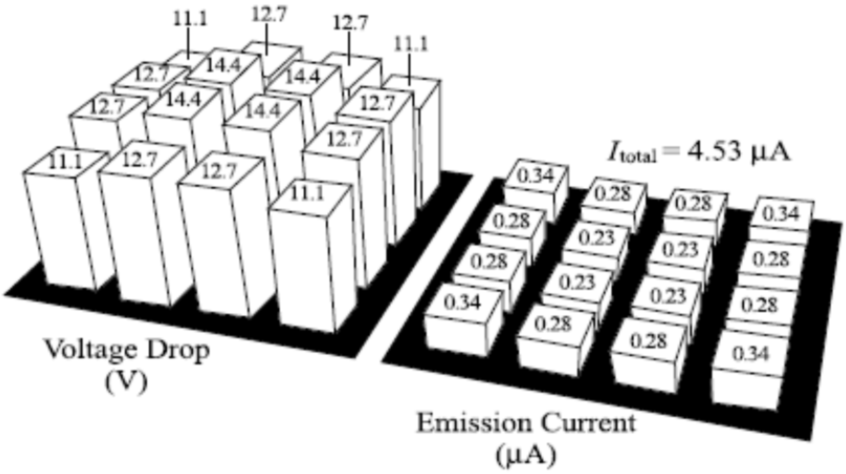


Figure 13.17. Voltage drop and current distribution of a 4 × 4 emitter array embedded in a lateral rasistor mesh with a sheet resistance of 10^7 Ω/□ and in which each emitter emits 1 µA at 100 V [26].

The effect of one tip shorting to the gate is shown in Fig. 13.18. Shorting can be caused during processing or due to the tip exploding during operation, leaving a conductive film on the sidewall of the insulator between the gate and the emitter line [27]. It is assumed that the upper left emitter in the 4 × 4 array is shorted to the gate. The voltage drop is then 100 V, drawing a non-emissive current of 9.8 µA. The total emitted current from the other 15 tips is now 1.75 µA as compared to the 4.53 µA for the ideal case of identical emitters with the resistive layer. If the short is assumed at a different position, slightly different results are obtained.

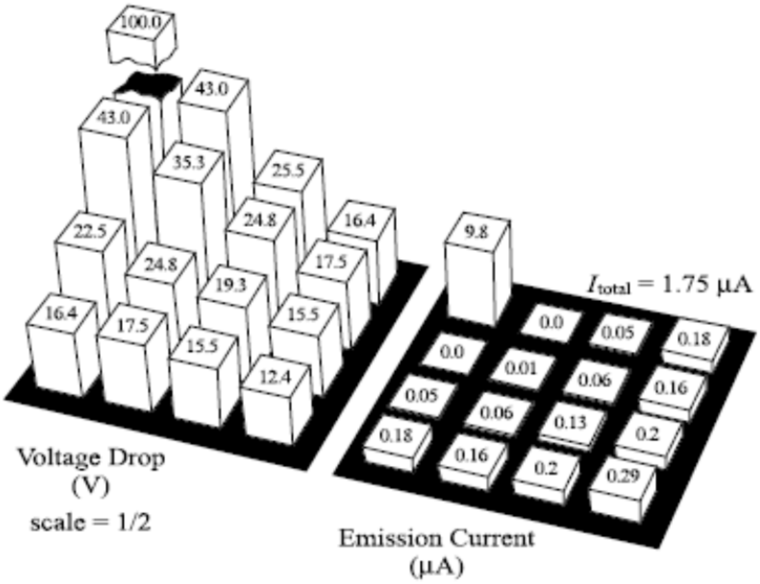


Figure 13.18. Voltage drop and current distribution of a 4 × 4 emitter array embedded in a lateral resistor mesh with a sheet resistance of 10^7Ω/□ in which the upper right emitter forms a short to the gate. Each emitter is capable of delivering 1 µA at 100 V [26].

A combination of vertical and lateral resistors was described in Ref. [28], and the cross section of such a configuration is shown in Fig. 13.19. It consists of a mesh structure and a conductive plate underneath the emitters. A comparison study was performed with a 21 × 21 array of mesh and mesh with plate structures, with the same distance between the cathode grid and the nearest emitter in the array and the same 0.5-µm thick amorphous Si resistive film. With the mesh structure, a current density of about 0.1 A/cm^2 was reached and 1.5 A/cm^2 when the conductive plate structure was included. The authors attributed this result to higher gate voltages per tip and improved uniformity of the voltage drops within the array. Postmortem inspection of failed arrays showed significant damage at the periphery of the array for the mesh structure and a more uniform distribution of failed devices for the mesh with conductive plate array.

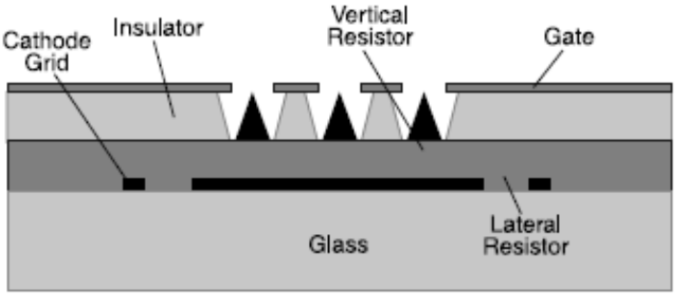


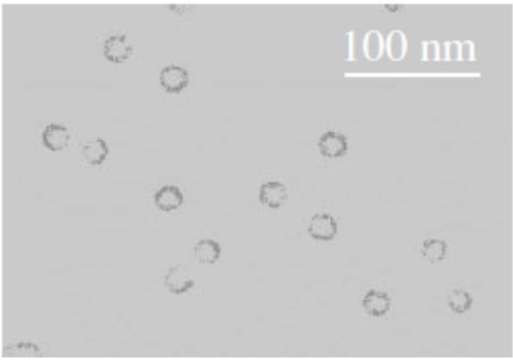
Figure 13.19. Cross section of an emitter array combining the vertical and lateral resistor concepts [28].

13.2.4. Emitter Array Fabrication

The objective of array fabrication is to produce low-cost, high-yield bottom plates that can be easily mated with the phosphor plate and also survive the sealing and evacuation process without imposing any detrimental effects to the emitters. In addition to fabricating uniform arrays containing millions of microtips, it is necessary to properly design the periphery of the bottom plate so that sealing can be performed without damaging the interconnect lines, and the thin film getter material can be incorporated without interfering with the performance of the arrays.

The array formation uses the classical Spindt process. For optimized tip formation, the incoming molybdenum beam should be collimated, which means a large source-to-substrate distance; one-to-two story high evaporators are on the market to satisfy this demand. In addition to this approach, other processing strategies such as random ion beam damage, laser interferometric lithography, and chemical mechanical polishing have been developed.

Ion Tracking Lithography. Ion tracking is used by one of the FED houses and was described in Ref. [5]. After the emitter lines, resistive material, gate dielectric, and gate rows are photolithographically defined, a polymer is applied and exposed to a low current density ion beam that is scanned across the emitter plate. Every ion induces a damage track in the resist. These tracks are then exposed to an etchant with a large (100:1) etch selectivity between the damaged and undamaged polymer. This causes the removal of polymer at the damaged sites, resulting in gate holes with diameters of about 0.15 µm. Other etchants are then used to remove the gate electrode material and to form the cavities in the interelectrode dielectric layer. After that, tip processing proceeds in a modified Spindt-type fashion. Figure 13.20 shows several gated tips in a small region of a subpixel that have been formed by this method. Tip densities exceeding 10^8 tips/cm^2 have been achieved for substrate sizes exceeding 14-in. diagonal.



Fsgure 13.20. Top view SEM micrograph of several gated tips fabricated by the ion beam tracking method [5].

Laser Interferometric Lithography. Large area arrays of dots or holes with submicron diameters and submicron-to-micron size spacings can be fabricated using laser interferometry [29], and working FED prototypes using this technique have been reported [30]. The technique, in some of its implementations, offers essentially unlimited depth of field, large-area (100 cm×100 cm) and submicrometer patterning with the use of existing lasers, and relatively inexpensive optics.Atwo-beam exposure system can form a dense array of patterns as shown in Fig. 13.21(a) with the distance between the dots or holes given by

$$d = \lambda / 2\sin(\theta) \qquad (13.2)$$

where λ is the wavelength of the laser, and θ is the angle of incidence.

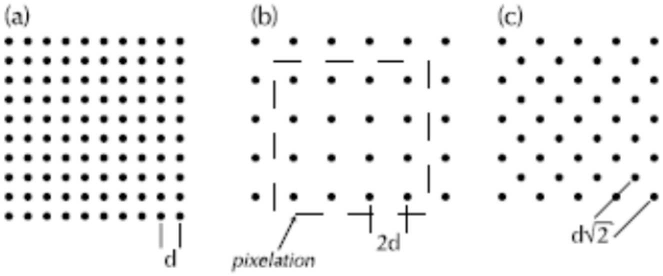


Figure 13.21. Hole pattern formation for (a) two-beam double exposure, (b) three-beam double exposure and five-beam single exposure, and (c) four-beam single exposure [31].

For obtaining sparser arrays, three-, four- and five-beam exposure strategies, each with its own subtlety, were described in Ref. [31], and the corresponding patterns are shown in Figs. 13.21(b) and (c). Sparser arrays are desirable from a reliability point of view, since failure propagation from tip-to-tip is reduced as the tip distance is increased. The trade-off of the three- and five-beam methods compared to the twoand four-beam systems is that the advantage of near infinite depths of focus is lost, necessitating the use of extremely flat substrates. Since arrays of holes are formed over the entire substrate in both the ion tracking and laser interferometric lithographies, an additional processing step is needed to block formation of tips in unwanted regions. This is indicated in Fig. 13.21 as pixelation.

Chemical Mechanical Polishing. Another attractive, large-area tip forming process is that of chemical mechanical polishing/planarization (CMP). Figure 13.22 shows the basic processing steps in forming gated tips [32]. After tip etching, the gate dielectric is deposited, followed by the gate material (Fig. 13.22(a)). At this point, the CMP process is performed, which consists of polishing off the protruding gate metal and insulator portions over the tips in a slurry that is chemically altered to facilitate enhanced removal (Fig. 13.22(b)). Once the caps are polished off, the pressure of the polishing pad increases, because the substrate is planarized. The point of increased pressure serves as an endpoint indicator of the process, and the polishing is subsequently stopped. The gate lines are then photolithographically defined, and the insulator near the tips is removed by an etching process (Fig. 13.22(c)). Commercial CMP machines capable of polishing 100 cm × 100 cm plates are readily available.

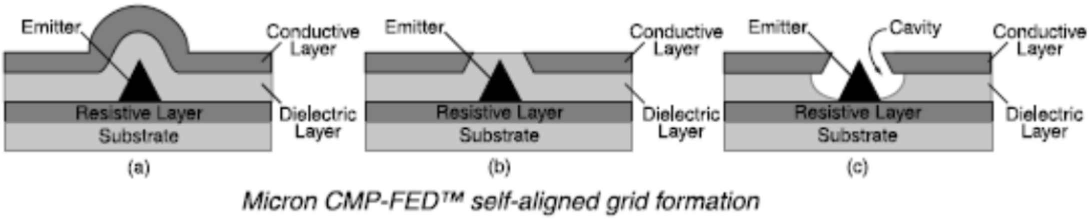


Figure 13.22. Cross sections of a gated device at different stages of the CMP process [32].

13.2.5. Phosphors and Phosphor Plate Fabrication

The color quality and brightness of FEDs should be at least as good or better than LCDs for desktop monitors or as good as the CRT for video image applications. When most of the efforts concentrated on the development of the emitter array during the early days of FED development, it was thought that standard CRT phosphors, which were developed to operate at 20–30 kV for optimum performance, could be employed (perhaps with minor modifications). It was soon learned that the maximum phosphor voltage for FEDs had to be limited below10 kV dictated by the short distance between the two plates. Much lower voltages are desired to eliminate flashover problems that can occur at those high fields.

Definition of Terms. Phosphors are characterized by their chromaticity, efficacy, and luminance. For FED applications, it is also important to understand the degradation mechanisms under e-beam bombardment and interactions of the phosphors with the emitters. The color of a given phosphor is defined by its *chromaticity coordinates*, x and y, and is displayed within the *chromaticity diagram* or the so called CIE (Commission Internationale de l'Eclairage) diagram [33] as shown in Fig. 13.23 (including the chromaticity coordinates for the P22 red, green, and blue phosphors that are currently used in CRTs, see Table 13.3).

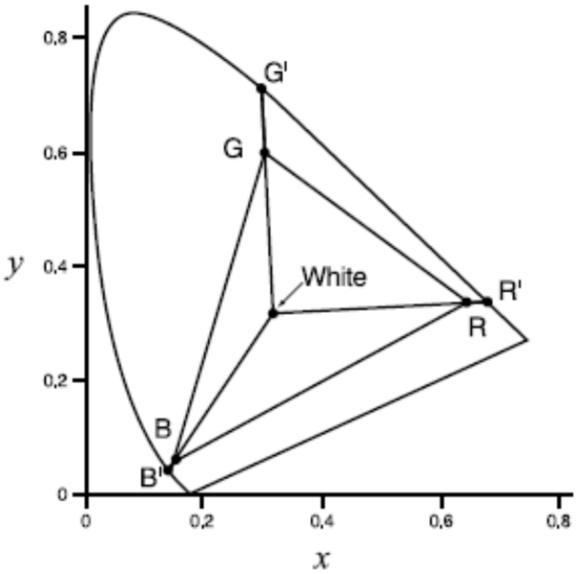


Figure 13.23. The CIE chromaticity diagram including x and y coordinates for P22 red, green, and blue CRT phosphors.

Table 13.3. Chromaticity coordinates for P22 CRT phosphors [34].

Phosphor	Color	*x*	*y*	Efficacy (lm/W at 25 kV)
ZnS:Ag:Cl	Blue	0.150	0.0060	10
ZnS:Cu:Au:Al	Green	0.305	0.605	65
Y2O2S:Eu	Red	0.650	0.340	20

The *color gamut* is the range of colors that can be obtained with a group of phosphors and is defined by the triangle that connects the three primary colors. The triangle (RGB) should be as large as possible, and the R, G, B coordinates should lie close to the envelope (R'G'B') of the diagram. By extending lines from the white point through the R, G, B coordinates to the envelope, the degree of *color saturation* can be obtained by determining the ratio of distances. For instance, the green in Fig. 13.23 is the least saturated color amongst the three phosphors, since the ratio of the distances, W − G/W – G', is the smallest. A good phosphor should be color saturated.

The chromaticity can shift as a function of excitation current density. This is shown in Fig. 13.24 for three ZnS phosphors [35], where increased current density follows the arrows. The gamut of colors decreases because of the relatively large shift of the red and green phosphors. Prior to 1964, ZnCdS:Ag was used for P22 red, but has since been replaced by Y_2O_2S:Eu due to the fact that ZnCdS:Ag exhibits a large current density shift, is hard to screen, and has a low efficacy.

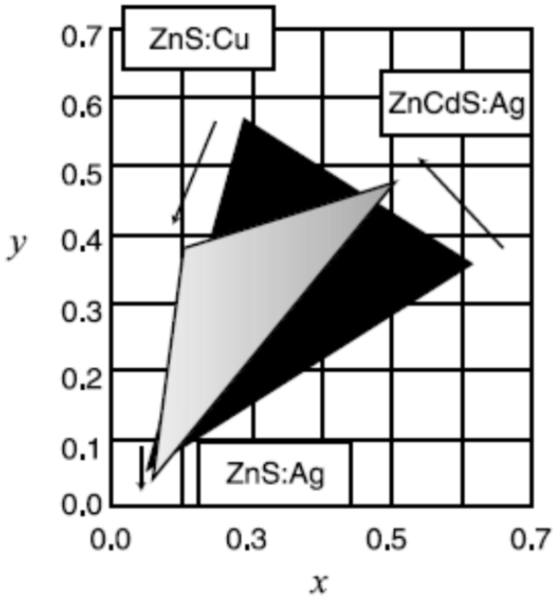


Figure 13.24. Chromaticities of red, green, and blue zinc sulfide phosphors for different current excitation densities at 5 kV anode voltage [35].

There are other CRT phosphors listed in Table 13.4 that are used for monitor and projection TV applications. The degree of current saturation in these phosphors is less than that of the P22 phosphors, and their Coulomb maintenance is better. However, these improvements are counterbalanced by their inferior chromaticities.

Table 13.4. Other Popular Phosphors for CRT Applications [34].

	x	*y*

Phosphor	Color	Low Current	High Current	Low Current	High Current	Efficacy (lm/W at 25 kV)
P3l – ZnS:Cu:Cl	Green	0.226	0.193	0.528	0.420	60
P43 – Gd2O2S:Tb	Green		0.333		0.556	35–40
P53 – Y3Al5O12:Tb	Green		0.368		0.539	30
P56 – Y2O3:Eu	Red		0.640		0.335	18

The *efficacy* of a phosphor under electron beam excitation is defined as [36]

$$\varepsilon = \pi LA/P \qquad (\mathrm{lm/W}) \tag{13.3}$$

where L is the luminance in cd/m^2, A is the area of the electron beam spot in m^2, and P is the power of the incident electron beam in watts obtainable by multiplying the anode voltage by the current. Note that the dimension on the left side of the equation is in lm/W and on the right side in cd/W. *Luminance* is a measure of the total energy output of the light source emitted in the visible region of the spectrum. The subjective sensation produced in the eye by this energy is known as brightness. Several units are used for luminance: 1 nit = 1 cd/m^2 = 0.292 fL (foot-Lambert). An objective measure of phosphor efficacy is defined by the integrated photon energy divided by the input power and is measured in W/W.

The *intrinsic luminous* efficacy is the efficacy of a powder sample of the phosphor. The *screen luminous efficacy* is the efficacy of a thin layer of phosphor powder deposited onto a substrate. Screen efficacies are typically lower than intrinsic efficacies because of the presence of binders in thin film phosphors that can absorb fractions of the excitation electrons. The binders may also chemically react with the phosphor. Screen efficacies can be measured in back reflection mode or transmission mode [36]. For FEDs to compete with LCDs on a power consumption level, the phosphors must have screen efficacies of 11, 22, and 3 lm/W for the red, green, and blue components, respectively [37]. To produce a white pixel of a given luminance, about 60% has to be generated by the green, 30% by the red, and 10% by the blue phosphor [34].

Phosphors developed for theCRT industry are most efficient in the 20–30 kV range. In general, for a given current, luminance increases with anode voltage, and for a given anode voltage, it increases with current. An example is shown in Fig. 13.25 for picture element tubes developed at Sarnoff using P56(Y_2O_3:Eu), P53($Y_3Al_2(AlO_4)_3$:Tb), and P55(ZnS:Ag) phosphors for R, G, B, and gated Si emitters for the excitation source. The phosphor voltage was 10 kV.

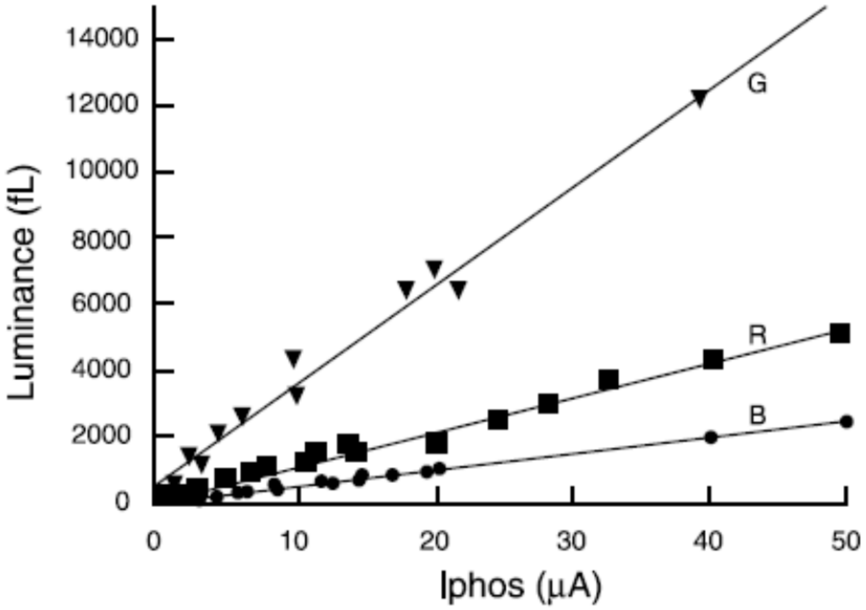


Figure 13.25. Luminance for P56 (red), P53 (green), and P55 (blue) phosphors as a function of phosphor current at a phosphor voltage of 10 kV. Spindt-type arrays were used as the excitation source.

At lower anode voltages, higher current densities are required in order to obtain the same brightness. Unfortunately, phosphors start to saturate at higher current densities due to the generation of higher defect densities that are caused by the decreasing penetration depth, and thus smaller excited phosphor volume. In addition to current saturation, phosphors decrease in efficacy under prolonged electron bombardment. This is known as *Coulomb aging.* Since low-voltage phosphors are excited at much higher current densities, the time at which the initial luminance is reduced to onehalf becomes much shorter. Coulomb aging is attributed to the formation of color centers (point defects that act as traps for electron–hole pairs) and to surface damage. Electron-stimulated chemical reactions between the phosphor and constituents of the residual atmosphere in vacuum (H_2, CO, CO_2, H_2O) are another potential Coulomb aging mechanism. Typical values for Coulomb aging when the luminance decreases to half its initial value are 100–200 C/cm^2 [38].

The *dwell time* is the time that the electron beam spends to excite the phosphor. For CRTs, it is several tens of nanoseconds. After the beam scans to the next phosphor dot, the luminance decays. Typical decay times for the P22 phosphors are 800 μs for red, 100 μs for green, and 60 μs for blue. As the beam scans the screen, it can select the same phosphor dot 1⁄30 s later. During this time interval, the luminance of the dot has sufficiently decayed and is ready for new excitement. FEDs are generally addressed in a one-line-at-a-time mode. This means that, with a repeat time of 1⁄30 s and a display of 480 lines, the dwell time is 0.033⁄480 or 69 μs. Exciting a pixel during this time period at increased current densities and reduced phosphor voltages causes severe current saturation with P22 phosphors. It has been shown that the current saturation can be reduced with phosphors in which the decay time (defined as the time to reach 10% of the initial luminance) is lower than the dwell time [39, 40]. This means that the phosphor can be excited, relaxed by luminescing, and re-activated several times during the dwell time.

Low and High-Voltage Phosphors. Phosphors can be divided as lowvoltage (<2 kV) and high-voltage phosphors (>2 kV). Most FED companies currently use high-voltage phosphors. To improve contrast, a black matrix is used between the color stripes (triade) or the rectangular subpixels (quad). A thin aluminum (Al) layer, 50–100 nm, is also used to increase luminance, since the light that would be radiated towards the emitter plate is reflected by the Al screen towards the viewer. The layer also minimizes outgassing of the phosphor. Depending on the thickness, anode voltages of 800–2000 V are needed for the electron beam to penetrate the Al layer.

Since the low-voltage phosphors are usually not metallized, resistive effects can severely limit their performance [41]. Mixing of conductive particles such as In_2O_3 with the phosphor powder is common [42]. ZnO:Zn is a low-voltage phosphor with the highest efficacy (10 lm/W at 100 V). Unfortunately, its chromaticity does not meet color TV standards. One reason for its high efficacy is that it has the highest conductivity. Since the electron penetration depth is reduced at low voltages, it is important to synthesize phosphors with long diffusion length and also investigate surface coatings using wide bandgap dielectrics such asWO_3 and V_2O_5 for increased efficacy and stability [43].

The Phosphor Technology Center of Excellence (PTCOE) at the Georgia Institute of Technology, Atlanta, Georgia (C.J. Summers) is maintaining a database on phosphors, and Sandia National Labs, Albuquerque, New Mexico (L.E. Shea) is establishing characterization protocols for the proper evaluation of phosphors. Tables 13.5 and 13.6 summarize some of the phosphors under consideration for low voltage (Table 13.5) and high voltage (Table 13.6) FEDs. The numbers in the parenthesis are efficacies in lm/W at 1000 V for Table 13.5 and at 10000 V for Table 13.6. Most phosphors used at this point are fabricated from powders and have particle sizes of several micrometers. Thin film phosphors deposited by sputtering and e-beam evaporation are also considered. They have better heat sinking properties and lower outgassing rates. However, they are less efficient because of lower optical scattering and increased light piping.

Table 13.5. Low-Voltage Phosphors for FED Applications [34].

In Use	Alternate Phosphors	Experimental
Y_2SiO_5:Ce – Blue (.5)	$Y_3(Al, Ga)_5O_{12}$:Tb – Green (YAGG) (7)	$SrGa_2S_4$:Ce – Blue
Gd_2O_2S:Tb – Green (14)	$ZnGa_2O_4$:Mn – Green (3)	$CaMgSi_2O6$:Eu –Blue
Y_2O_3:Eu – Red (4.5)	ZnO:Zn – Bluish green (8)	$Sr_3MgSi_2O_8$:Eu – Blue (.5)
		nGa_2O_4 – Blue (.3)
		$SrGa_2S_4$:Eu – Green

Table 13.6. High-Voltage Phosphors for FED Applications [34].

In Use	Alternate Phosphors	Experimental
ZnS:Ag:Cl – Blue (10)	Gd_2O_2S:Tb – Green (35)	$SrGa_2S_4$:Ce – Blue (4)
ZnS:Cu:Au:Al – Green (65)	Y_2O_3:Eu – Red (9)	$SrGa_2S_4$:Eu – Green (45)
Y_2O_2S:Eu – Red (20)		

Deposition Methods. Deposition of phosphors from powders can be performed by the slurry process, screen printing, electrophoresis, dusting, and several other methods. As shown in Fig. 13.26 for the entire steps in the slurry process [44], the starting glass substrate (7059 Corning or equivalent) is first coated with indium tin oxide (ITO) (transmission 94–95%) with a sheet resistance of 20–25 Ω/□. After cleaning the substrate, the black matrix pattern is formed. This can be accomplished by using silicate-based oxide powders (0.1–0.5-µm size particles) through suitable screens by screen printing. The substrate is then baked, and loose particles are blown off. The phosphors, in the form of a slurry containing sodium or ammonium dichromate, are coated onto the substrate with polyvinyl alcohol (PVA) by spin coating, dried, and exposed to light through a shadowmask. After removal of the mask, the film is developed by washing away unwanted portions of the photoresist/ phosphor pattern and dried. This sequence has to be repeated for the other two colors.

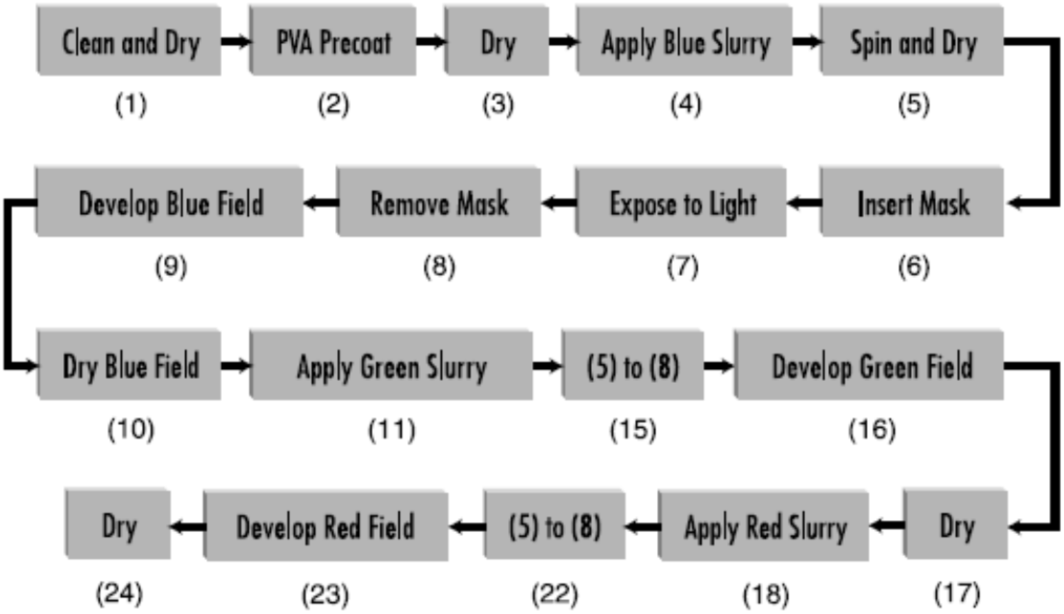


Figure 13.26. Steps in forming a phosphor plate by the slurry process [44].

After the R, G, B phosphor patterns are defined, a thin film of lacquer (acrylic polymer) is deposited, followed by aluminum deposition. The lacquer acts as a planarizer to obtain highly reflective Al coverage and is subsequently evaporated (with the PVA) in air at 440°C by escaping through pinholes in the Al. For high-voltage phosphors, the initial ITO layer is not necessary, since the Al layer is conductive. For low-voltage phosphor plates, the ITO layer is essential, since no Al is deposited. The cross section of the finished phosphor plate is shown in Fig. 13.27. An application of this method for fine-line (25 μm) phosphor deposition was given in Ref. [45].

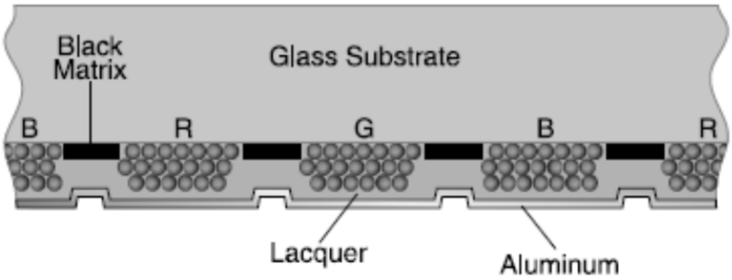


Figure 13.27. Cross section of a high-voltage FED phosphor plate. At least three layers of phosphor powder are needed per subpixel.

For the screen printing process shown in Fig. 13.28, the phosphor slurry is forced through the opening in a silk or metal screen with a rubber squeegee and deposited onto the glass substrate. The screen is then lifted off, and the plate is dried. This process needs to be performed with all three colors. Similar to the slurry process, both the ITO layer and black matrix are formed prior to phosphor screening, and both the lacquer and Al are deposited later.

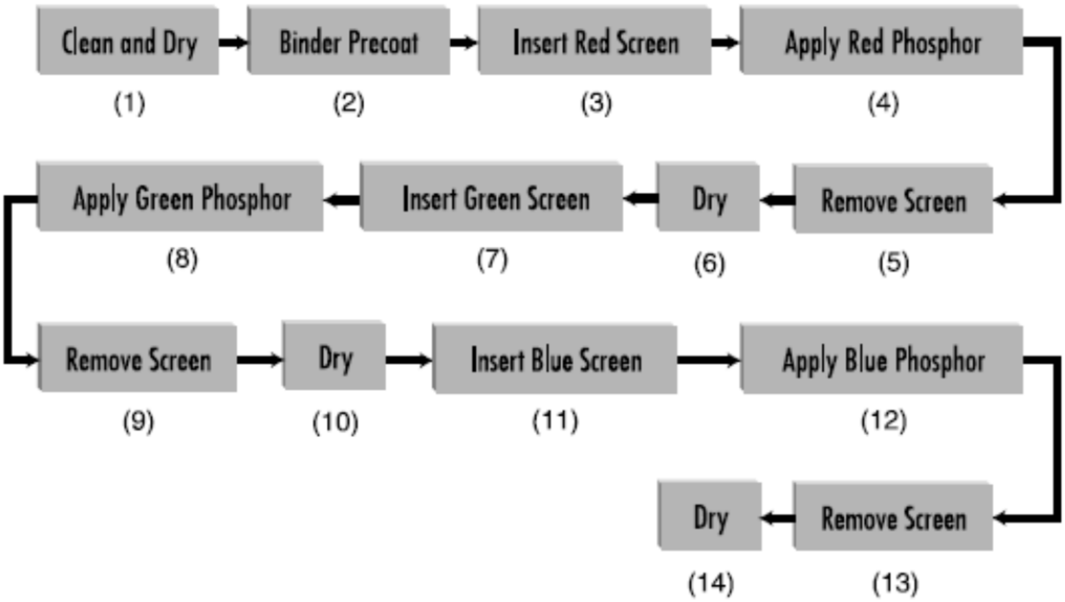


Figure 13.28**.** Steps in forming a phosphor plate by screen printing [44].

Electrophoresis is a transport process in which conducting or nonconducting particles in a suspension move under the influence of an applied field and deposit on substrate surfaces. A phosphor deposition process based on electrophoresis was described in Ref. [46] and in additional references given therein.

Phosphor Summary. It is clearly desirable to use low voltage phosphors in FEDs for reduced voltage operation and simplified device designs. However, efficient low-voltage phosphors are still under development. To be efficient, they should possess high electrical and thermal conductivities. Since the screens will not be aluminized, these phosphors can become potential emitter contaminants, and it is, therefore, important to minimize the outgassing from these phosphor surfaces. The degradation by total charge injection also needs to be minimized, and no saturation or emission wavelength shifts at high current densities should take place. The emission spectra should be narrow and lie near the boundaries of the CIE diagram.

High voltage phosphors, on the other hand, do exist and are currently being used at voltages ranging from 5–7 kV. They are efficient and are usually aluminized. The thin aluminum layer provides electrical conductivity and additional viewer intensity. It also reduces the contamination of the emitters from phosphor and binder byproducts. Because of the lower current densities required, the degradation due to charge loading is minimized. Disadvantages include the need for increased panel spacings and high voltage spacer technology. The larger distance between the emitters and phosphor screen takes the display out of the realm of proximity focusing, thus requiring a more complex scheme that includes the addition of focusing electrodes.

13.2.6. Beam Focusing

When the subpixel size is less than about one-tenth of the anode-to-cathode spacing, the electron beam is generally required to be focused. For an 80 μm subpixel size, the maximum distance for proximity focusing is about 0.8 mm. This is far too short compared to the 1.5 mm spacing that is typically required for high-voltage operation. Figure 13.29 shows several focusing or beam confinement schemes, and theoretical analysis of beam focusing was described in Ref. [47].

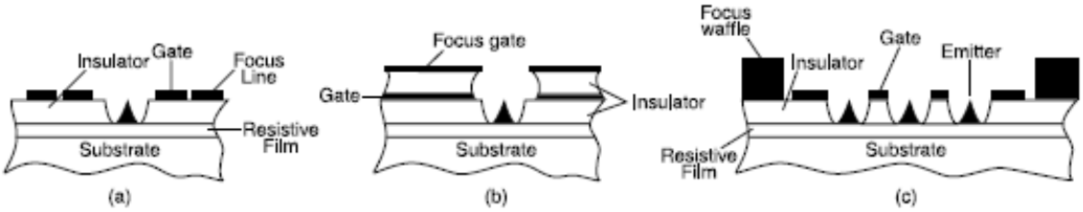


Figure 13.29. Methods for generating focused electron beams: (a) in-plane, thin film focusing of each emitter [47]; (b) double-gate, thin film focusing of each emitter [47]; and (c) thick film focusing of a group of emitters [34, 48].

Applying a focusing electrode to each tip as shown in Fig. 13.29(a), which would be the most effective means, is not very practical for FED applications, since it complicates manufacturing considerably. For the configuration shown in Fig. 13.29(b), experimental results indicated that, at a gate voltage of 80 V and anode voltage of 1000 V, focusing can be achieved at voltages between 10 and 20 V [49]. Below 10 V, the anode current drops drastically. At present, the method of choice in the FED industry is that of Fig. 13.29(c), which is more of a beam confinement method rather than true individual tip focusing [30, 48]. Modeling results have shown that the best focusing is obtained with the double gate structure, followed by the in-plane structure, and then by the waffle structure [50]. A microphotograph taken with a digital camera of light emitted from a display using a 40-µm thick film focusing “waffle” is shown in Fig. 13.30 [48]. The column pixel pitch is 94 µm.

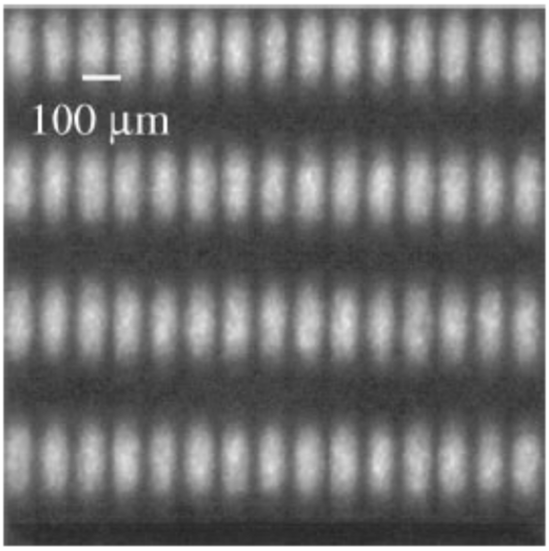


Figure 13.30. Microphotograph of light originating from a display with a pixel pitch of 94 µm using a thick film “waffle” around each subpixel for focusing. The waffle is biased at a few volts (dc) negative with respect to the gate voltage [48].

13.2.7. Encapsulation

In addition to the fabrication of emitter arrays on the baseplate and phosphor subpixels on the faceplate, spacers need to be inserted or fabricated on one of the plates, the plates have to be brought together and aligned, and the periphery of the plates must be sealed to form a vacuum envelope. Getters, which are placed at the periphery of the device (near or inside the pumping stem), are activated as the final step to provide chemical pumping throughout the life of the display. Figure 13.31 shows some details of an FED near the periphery [51]. The row and column address leads of the baseplate have to be extended beyond the FEA, and the lower lying emitter leads must have the gate insulator material removed to form pads that can be connected to the driver electronics via solder or pressure connectors. Provisions also need to be made for the anode pad(s), the pump port(s), and the glass frit seal.

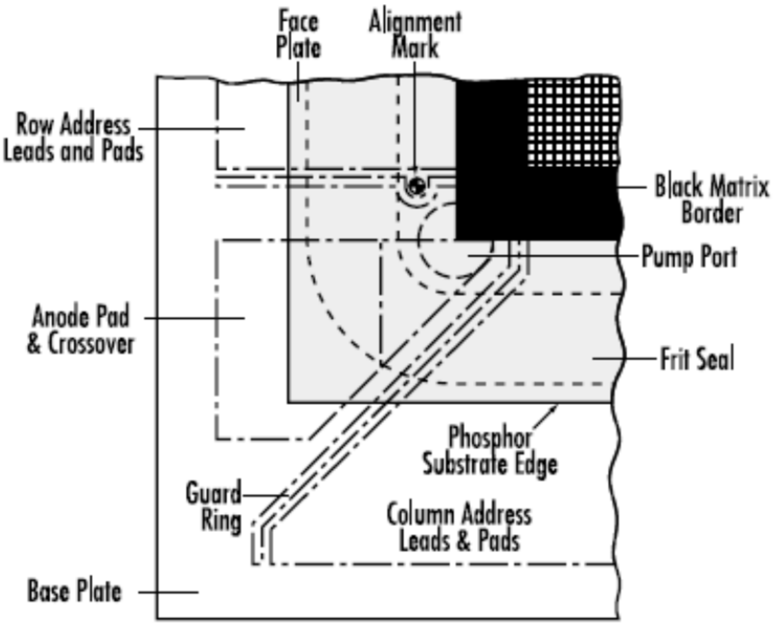


Figure 13.31. Segment of the lower left edge of an FED showing conductor pads, pumping port, and glass seal. Not shown are the getters [51].

Spacers. Spacers are the “orphans” of the FED technology. It is interesting to observe that during the first 10 years of the IVMC, no papers on spacers were presented. Only recently breakdown results on small gaps in vacuum without spacers [52] and with spacers [53-55] have been seen in publications. Nevertheless, none of the companies involved in FED product development are sharing performance results of their spacer technologies.

An ideal spacer should possess the following properties: (i) good mechanical strength to support the baseplate and faceplate under the influence of the atmospheric pressure, (ii) low cost and easy to manufacture and to position, (iii) good dielectric strength (>5V/µm), (iv) low coefficient for secondary electron emission, (v) high vacuum flashover resistance, (vi) slightly conductive to avoid charging, (vii) transparent to the viewer, and (viii) matched thermal expansion coefficient with glass. Spacers can take the forms of spheres, rails, or posts. Materials that have been investigated include glass, ceramic, polyimide, and silicon nitride coated polyimide. Resistive coatings that have been investigated to reduce charging include metal oxides (Fe_2O_3, Cu_2O, Cr_2O_3), DLC films, amorphous silicon, and SiC [56]. Spacer fabrication can be an integral part of the faceplate or baseplate manufacturing process. Alternatively individually processed spacers can be attached to either one of these plates. The maximum dimensional requirement for a spacer is that it fits within the black matrix. It should be somewhat smaller to allow for alignment tolerances in placing the spacer (if it has not been lithographically defined) and in aligning of the two plates. For instance, for a 17-in. diagonal super extended graphics array (SXGA) (1280 × 1024) FED, the black matrix is 60 µm between the rows of the display and smaller between the columns. By placing the 2.4-mm tall glass fiber posts only in the rows of the black matrix, 40-µm diameter spacers were chosen [57]. This corresponds to an aspect ratio of 2400/40 or 60. Spacer geometries and placement need to be determined very carefully for each display application. The spacer has to be electrically invisible, meaning that no electrons should strike it. Nevertheless, stray electrons and secondary electrons from the faceplate can still end up striking the spacer, which causes charging and distortion of the image around the spacer. This problem must be eliminated by the proper choice of the spacer materials and technology.

The breakdown voltage of thin film electrodes deposited on glassplates and measured in a relatively poor vacuum of 1 × 10^{-6} torr was reported to be larger than 18 kV for a 750-µm gap using Ni/Cr electrodes [52], corresponding to a breakdown field of 24 V/µm. By inserting teflon disks and wall spacers, breakdown voltages ranged from 13 to 16 kV and 15 to 20 kV, respectively, for the same 750-µm gap [53]. By replacing the teflon spacers with thin-wall machinable glass ceramic (MACOR), Coors ceramic, Coors YTZP Zirconia, and AMZIROX Zirconia, breakdown fields ranging from 6.4 to 24.5, 5.9 to 8.8, 8.0 to 16.7, and 4.0 to 10.4 V/µm, respectively, were obtained [53]. These values were obtained when the spacers were inserted between upper and lower thin film electrodes. It is anticipated that, if these experiments are performed using FED faceplate and baseplate, the breakdown fields will be further reduced, because the spacers are now subjected to bombardment from primary electrons from the emitters and secondary electrons from the phosphors.

Several breakdown mechanisms have been proposed to explain spacer flashover. Figure 13.32 shows two of the proposed spacer breakdown mechanisms: (a) electron injection, hopping, secondary electron emission, and impact ionization, and (b) electron injection and the formation of a surface conducting channel. Most of these explanations relate to macrosize spacers with diameters much larger than the spacers used in FEDs. Breakdown generally occurs in two stages: prebreakdown electron injection, followed by vacuum flashover. The role of the triple junction, which is defined as the intersection of conductor, insulator, and vacuum, plays a critical role in spacer breakdown. Triple junctions can be formed at both the cathode and anode. In Fig. 13.3, the cathode triple junction (CTJ) is formed at the intersection of the conductive thick film focusing grid, the wall spacer, and the vacuum. The anode triple junction (ATJ) is formed at the intersection of the Al-coated phosphor plate, the spacer, and the vacuum. It is conjectured that high fields established at the conductor–insulator interface can cause electron injection from the conductor into the insulator. For macrospacers, this injection is also aided by microparticles and rough surface features near the spacer–conductor periphery. Early theories in the 1980s [56, 58] proposed that the electrons injected into the insulator “hopped” along the surface of the insulator, generating more electrons by secondary electron emission at the electron–insulator impact points, thereby causing an electron avalanche and positive charging of the spacer. In parallel with the charging process, electron-stimulated desorption of neutral or positively ionized atoms and molecules from the insulator surface can also take place. The positive ions drift towards the cathode (in the case of Fig. 13.3, the focus grid) and, with the positively charged insulator from impact ionization, give rise to an enhanced field at the triple junction. As a result, through a combination of gas desorption, regenerative secondary electron emission and impact ionization, the necessary conditions can be established for striking an avalanche-based plasma discharge across the surface of the spacer. Other models, such as hot-electron processes within the surface layers of the insulator that could give rise to ionization sufficient to enable a solid-state avalanche discharge [59] and explosive relaxation of trapped charges [60], have also been proposed.

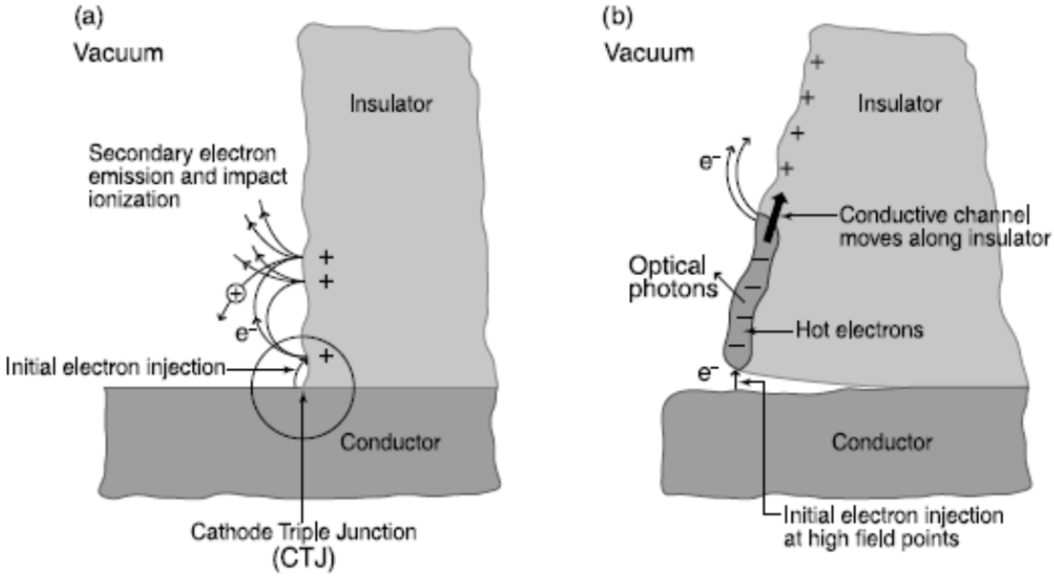


Figure 13.32. Proposed spacer breakdown mechanisms: (a) electron injection, hopping, secondary electron emission, and impact ionization [56] and (b) electron injection and the formation of a surface conducting channel. An essential prerequisite for growth of the conductive channel is the emission of electrons from the surface of the insulator [58].

For high-voltage macrospacers, it has been shown that suppression of spacer breakdown can be achieved by specially shaping the insulator near the triple junction [61, 62] or reducing secondary emission by coating the spacer with materials exhibiting low secondary electron yield, ideally with the yield $\delta_e \leq 1$. These coatings, which should also be slightly conductive, reduce the build-up of the positive charge. Truncated cone-shaped spacers with one end larger than the other end (the larger end should be positioned on the cathode electrode) have also been shown to increase the breakdown fields by six-fold (from 5 to 30 V/µm) when the truncated cone angle is greater than 45° [62]. This approach might be suitable for low resolution FEDs, because real estate considerations preclude such an approach in high-resolution displays. So far, the most prominent method is to coat the spacers with a low secondary electron yield material with its conductivity just enough to inhibit charge build-up.

Recent studies of 2-mm tall, conically shaped alumina macrospacers have yielded additional information about spacer breakdown mechanism [58]. It is shown that the pre-breakdown conduction process is “nucleated” at a highly localized site on the CTJ with simultaneous emission of optical photons. This process is believed to be associated with the injection of electrons from a particulate-based emission site located under the bounding edge of the insulator. These sites are switched on by the enhanced fields in the range of 5–10 V/µm. Because of the high field within the insulator, the injected electrons will be ballistically “heated” and, as a result of internal ionization processes, give rise to a point source of light at the CTJ. Depending on the microscopic configuration of the emission site, the switch-on process may result in an energetically favorable regime for the subsequent growth of a conducting channel along the surface of the insulator. Some of the field-emitted electrons may bombard the insulator surface in front of the ATJ and give rise to an anode conducting channel. Typically, it can take several seconds for the channels to form. The final breakdown or flashover of the gap is a very rapid process and generally occurs at some random times after the growth of a conduction channel.

It is yet to be reported how these findings relate to microspacers and what breakdown mechanisms are the most relevant. In addition to spacer breakdown, vacuum breakdown can also occur at sharp points on the baseplate and from dislodged particles (such as phosphors, aluminum) derived from the faceplate. It is critical for high voltage robustness that all components and assembly be handled in a clean room environment to minimize particles and contamination issues.

There are some promising developments in the search of innovative spacer technologies. The MEMS approach might offer some interesting solutions. Spacers with features that inhibit charge initiation and propagation could be embedded into the cathode or anode manufacturing process and then released and actuated as a final step. Actuation could be achieved via built-in stresses in the spacer or other means. One such approach, in which a Ti–Ni alloy filmwas used for activation,was described in Ref. [63].

Alignment/Sealing. After fabrication or placements of the spacers, the two plates must be aligned and sealed. There are two basic methods of sealing, namely at atmospheric pressure or in vacuum. The atmospheric method is similar in many ways to the standard CRT sealing procedure. However, due to the sensitivity of the emitters towards oxidation and poisoning, special precautions should be taken. If air is the sealing ambient, the tips must be coated to prevent oxidation or specially treated after sealing. Sealing in nonoxidizing atmospheres such as nitrogen and argon has also been used [64]. Depending on the glass used for the manufacture of the baseplate and faceplate (borosilicate, float, sodalime, etc.), it is important to work with the glass manufacturers to select the appropriate frits for sealing and follow the recommended temperature cycles. Typical frit firing temperatures range from 350 to 450◦C. Shrinkage of the glass during sealing also needs to be considered. Additionally, one must consider the CTE (coefficient of thermal expansion) of the spacers.

For a 17-in. SXGA display, an alignment tolerance of about 5 μm is required [57], and the appropriate alignment/sealing fixturing systems have to be designed. Generally, fixturing is less expensive when aligned and sealed at atmospheric pressure as compared to alignment/sealing in vacuum. The basic atmospheric sealing process involves dispensing of the glass frit on either the faceplate or the baseplate, followed by pre-firing. The two plates are then aligned and heated together. As the frit softens, the plates are squeezed together, and the frits are extruded. The plates are then cooled, and a vacuum seal is established. The package is then evacuated through an exhaust tube. Evacuation is usually performed at elevated temperatures to desorb moisture and outgas the envelope prior to pinching off the exhaust tube. Pinch-off is done by heating the tube until it softens and then pulling on it until a seal is formed. A chemical pump (getter) is usually placed around the exhaust tube and activated by rf heating. In most cases, evaporable getters such as barium (as they are used in CRTs) are employed. Care has to be taken that none of the barium is deposited onto the emitter tips.

In the vacuum alignment/sealing process, the two plates are aligned and placed into a high vacuum station. A gap exists between the two plates for the efficient removal of air. The temperature in the vacuum chamber is raised to allow moisture to be desorbed and outgassing to take place. The temperature is then raised to the point where the glass frit begins to soften, and the two plates are forced together, causing the glass frit to extrude. The glass frits used for vacuum sealing must be specifically processed to minimize outgassing. At this point, the envelope is sealed, and the temperature is slowly lowered to prevent thermal stressing of the glass plates. A thermally activated getter (e.g., by the temperature of the sealing process) can be used in the vacuum process. This vacuum seal process is better suited for FEDs, since the baseplate is not contaminated by air. In principle, the two plates can be outgassed prior to alignment and joining, and the faceplate can be exposed to an electron beam to condition and pre-age the phosphor prior to sealing. Disadvantages include a more expensive system and potentially reduced throughput.

By replacing the air sealing process with sealing in nonoxidizing atmospheres, some of the advantages of the vacuum sealing process can be retained in the less expensive atmospheric pressure sealing method. In Ref. [64], Spindt-type Mo tips were mounted on glass substrates that contain pre-baked frits. The substrates were then placed into the sealing furnaces filled with N_2, Ar, and air. Severe degradation of the *I*–*V* characteristics was observed in the air-sealed devices. Reduced degradation was seen with the nitrogen-sealed devices, and the best results were obtained when Ar was used. For example, for a constant current, the gate voltages shifted upwards by 250 V in air, 30 V in N_2, and 15 V in Ar. The surface morphologies of Mo tips as measured by atomic force microscopy (AFM) also roughened significantly, by an average roughness of 39 °A (air), 36 °A (N_2), and only 9.7 °A (Ar), respectively.

In Ref. [65], the authors showed that Mo Spindt emitters that were treated at 470 °C for 20 min in highly purified N_2 and Ar gases displayed no degradation in their *I–V* characteristics, similar to those vacuum-sealed (10−6 torr) devices. The authors concluded that if the partial pressure of oxygen in Ar or N_2 was kept below 7×10^{-6} torr, no detrimental effects upon the *I–V* characteristics would be observed. In Ref. [66], the authors observed a 20-V upward shift in the *I–V* characteristics of Mo tips when they were sealed in an Ar ambient, while no shift was observed when using Ar with 1.5% H_2. They claimed that the small amount of hydrogen keeps the Mo tips from oxidizing. Further, the Ar/H_2 treatment did not result in any detrimental changes in the luminance and chromaticity of the devices.

Vacuum Requirements. Although CRTs can operate at pressure of about 10^{-6} torr, Spindt-type emitters fabricated from refractory metals and silicon should be operated below 5×10^{-7} torr. Above these pressure, noise increases and life expectancy decreases due to ion bombardment effects. Some of the planar emitters such as BN [67] and carbon-type emitters [68] were reported to be capable of operating at 10^{-4} torr. However, no good lifetime data exist at these pressures.

The basic panel structure consists of the faceplate and baseplate joined together with a glass frit as shown in Fig. 13.3 as well as in Fig. 13.33 in a cross-sectional form. To increase pumping efficiency, a panel structure with an auxiliary tank as shown in Fig. 13.34 was investigated [69]. Several pumping holes are fabricated in the baseplate, and the device is evacuated through an exhaust tube that is positioned at the back of the auxiliary tank. The advantages of this structure are that placement of the getters is more convenient, and that sufficient getter volume can be used. The disadvantages are a relatively thicker display panel and the need for additional spacers as indicated in the figure.

The evacuation stems are a nuisance during the final packaging of the displays and can cause outgassing problems during heating prior to pinch-off. A stemless process that employs an anodic bonding technique was described in Ref. [70]. As shown in Fig. 13.35, the anodic bonding takes place at the sputtered silicon–glass interface at an elevated temperature of 280 °C under dc bias. By applying a negative bias to the glass seal-off plate via the ITO layer, positive metal ions (Na, Al, K, Mg, . . .) migrate to the ITO layer, leaving a space-charge layer near the silicon–glass interface that supports most of the applied bias. The resulting large electrostatic force pulls the plates together, resulting in Si O Si bonds. The advantages of this method are that no frit is needed, and that the sealing temperature is only 280 °C. The disadvantage is that the bonded surfaces have to be extremely smooth, with peak-to-valley roughness not exceeding 200 °A.

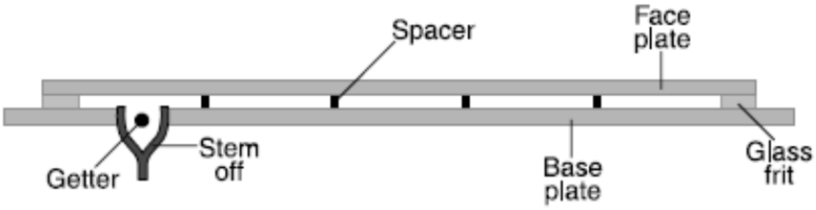


Figure 13.33. A conventional FED panel consisting of two glass plates [69].

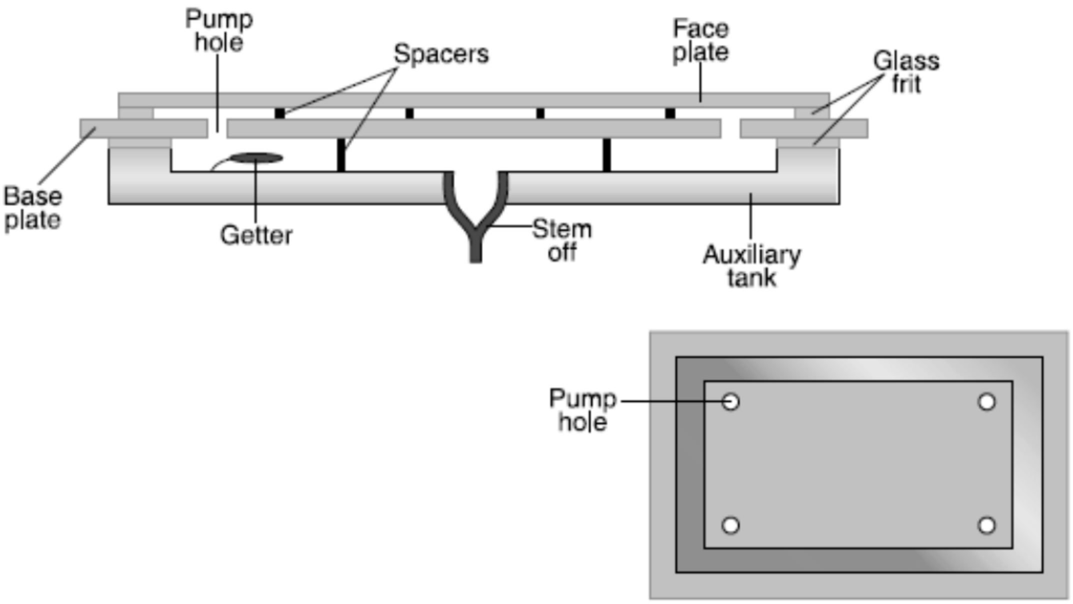


Figure 13.34. Cross section and top view showing four pump holes of a panel structure with an auxiliary tank [69].

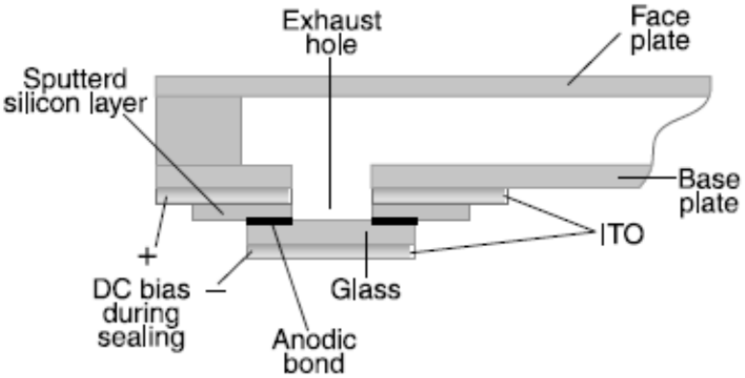


Figure 13.35. Cross section of a stemless panel in which the tipped-off exhaust tube is replaced by an anodically bonded glass plate [70].

Some of the pumping results from the conventional structure shown in Fig. 13.33 and the structure shown in Fig. 13.34 are presented in Fig. 13.36. The 5.7-in. Diagonal faceplate and baseplate are fabricated from sodalime glass with a panel separation of 200 µm. The diameter of the pump hole is 5 mm, and that of the exhaust opening is 5.8 mm. The results show that after 250 min of pumping, the pressure in the conventional panel has still not reached its lowest value. The pumping speed is similar when one pump hole in the baseplate is used. The condition improves quite considerably only when four pump holes in the baseplate are used.

It is interesting to observe that the equilibrium pressure is rather high at about 6×10^{-5} torr. This is caused by the low conductance of the 200-µm spaced panels and by the outgassing of the two plates. By assuming a panel that is closed at one end and pumped at the other with infinite pumping speed, the pressure at a distance x along the panel for $t > 3600$ s can be calculated as [71]:

$$P(x) = \frac{64}{\pi^3}\left(\frac{L}{a}\right)^2 Q\left(\frac{mk_BT}{2\pi}\right)\sum_{n=0}^{\infty}\left[\frac{(-1)^n}{(2n+1)^3}\right]\cos\left(\frac{\pi(2n+1)x}{2L}\right) \qquad (13.4)$$

where L is the length of the panel, a is the panel spacing, Q the outgassing rate in number of molecules/m^2 s, m the gas mass (g/mole), k_B the Boltzmann constant, T the absolute temperature, and x the distance along the panel. Results for a 23-cm long panel with 200, 400, and 800 µm spacings are shown in Fig. 13.37. It is assumed that the sticking coefficient is zero, i.e., no entrainment of gas to the surface takes place once it is released by outgassing. In the calculations, *Q* is assumed to be 6.4 × 10^{15} molecules/m^2 s for the outgassing rate of stainless steel, and *m* is 29 g/mole for air. These data show that spacing and outgassing rates are the most important parameters in obtaining low pressure inside the panel. Unfortunately, they also show that if flashover events occur away from the getter pump, the getter might not be able to reduce these local pressure bursts fast enough to avoid localized damage.

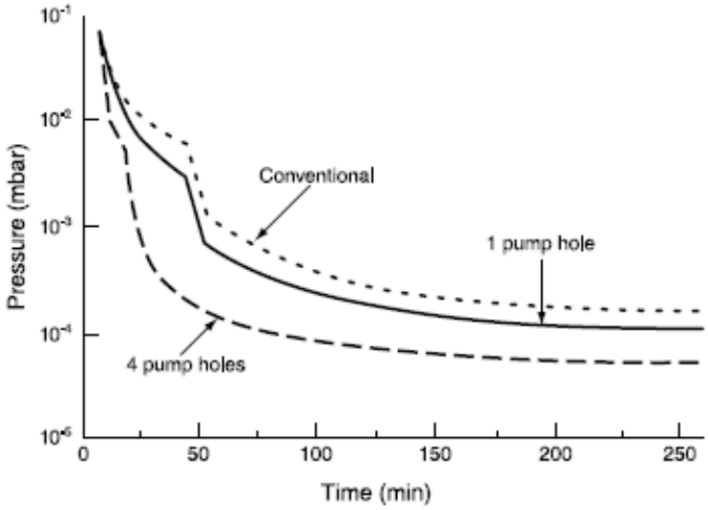


Figure 13.36. Pressure vs. pump time for a conventional 5.7$^{//}$ diagonal panel and a panel with an auxiliary tank in which the baseplate has one and four pump holes of 5-mm in diameter [69]. Note that the pressure unit of mbar is used. As a refresher: 1 Pa = 1 Newton/m^2; 1 torr = 133.32 Pa; 1 mbar = 100 Pa; 1 atmosphere = 1.013 × 105 Pa; and 1 torr = 1 mmHg.

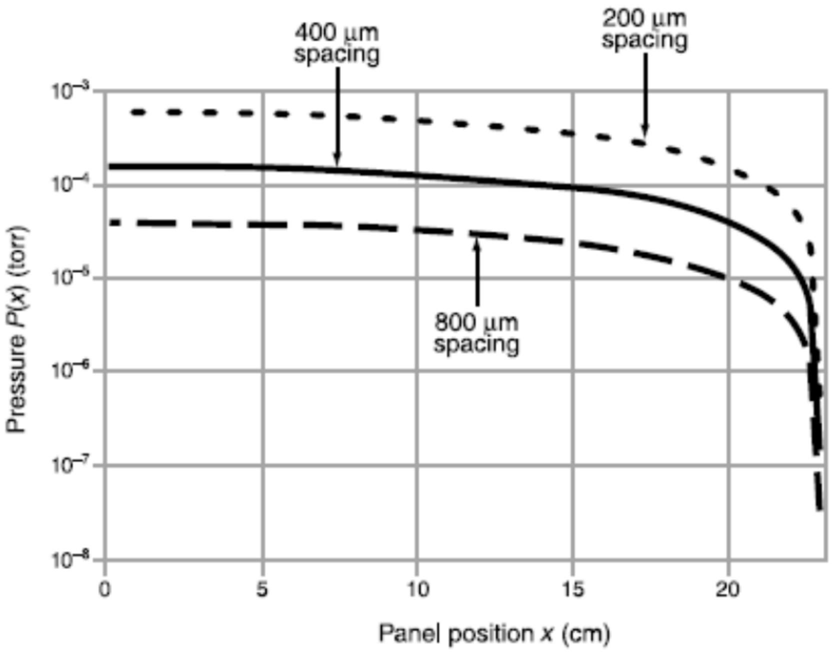


Figure 13.37. Pressure vs. distance along a 23-cm long panel with panel spacings of 200, 400, and 800 µm. The panel is closed on the left side (x = 0) and pumped on the right side (x = 23 cm) with infinite pumping speed [71].

Chemical Pumping or Gettering. Once a panel is sealed, changes in pressure over its life can be caused by leaks, permeation, and outgassing. Assuming that materials and processes are chosen to eliminate the first two causes, outgassing remains. Outgassing is caused by thermal fluctuations in the materials used in panel production, electron bombardment of spacers and phosphors, and electron-induced desorption of species at the tips. To combat the pressure increase, getters are installed inside the panel and activated. A good review concerning getters was provided in Ref. [72].

Gettering is routinely used in CRTs, X-ray tubes, particle accelerators, and many other vacuum devices. The getter material chemically interacts with active gas species, forming very stable chemical compounds both at the surface and in the bulk. At room temperature, getters can chemically absorb hydrogen, oxygen, water, nitrogen, and carbon monoxide and dioxide in an irreversible manner. The sorption capacity of getters at room temperature is related to the available effective surface area, while at elevated temperatures (above a threshold), it is close to the theoretical stoichiometric capacity for the chemical compounds that are formed such as oxides and nitrides. Sorption of hydrocarbons (e.g. methane) only occur at high temperatures. Noble gases such as Ar and He cannot be gettered.

Getters are available in two forms: evaporable and nonevaporable. The evaporable getters are commonly based on a mixture of $Ba–Al_4$ alloy and Ni powders compressed together into ring- or wire-shaped containers. These containers are then heated, by rf coils, to 800–900°C to promote an exothermic reaction between the two components. These getters are mainly used inCRTs, since large areas are available for the deposition of a porous gettering film. Although they are also used in FED prototypes, their effectiveness is limited, because large surface areas for getter deposition are not available.

Over the last fewyears, getter manufacturers have been increasinglyworking on the development of non-evaporable getters based on Zr–Al and Zr–V–Fe alloys. Of great interest for FED applications are high porosity (60–65%), screen printable getters based on Zr–V–Fe and Ti. Sorption capacities of about 1 ($cm^3 \cdot torr/cm^2$) have been measured at room temperature for carbon monoxide. The capacities are 10 times higher for O_2 and H_2O, and 100 times higher for H_2. These getters are typically screen-printed with 100, 200, and 400-μm thick lines and are activated at 500 °C for about 10 min. Figure 13.38 shows the sorption characteristics of a screen printable HPTF (high porosity thin film) NEG (non evaporable getter).

To design an appropriate gettering system, it is important to measure the electroninduced outgassing rate of different gaseous species in the displays. The outgassing rate typically decreases as a function of time t in the form of t^{-1} for most gases and $t^{-1/2}$ for hydrogen. By knowing the outgassing rate, the total gas load and the corresponding increase in pressure over the life of the displays can be calculated. The amount of getter material can, in turn, be estimated from the total gas load. For example, a 20 cm × 20 cm display area has an estimated gas load of 1.5 (cm^3 mbar) after 10000 h of operation [38,72]. This corresponds to a pressure build-up of 10^{-2} mbar, and only an area of several square centimeters of HPTF NEG would be needed to absorb this gas load.

The real devices are more complicated. Getters are typically placed at the periphery of the panel, near or in the evacuation stem. This limits their effectiveness in terms of absorbing pockets of outgassing deep inside the panel. Distributed getters that consist of a network of getter material uniformly distributed across the entire panel are, therefore, considered more effective. However, for high resolution panels, this is difficult to achieve and would further increase the complexity of the display structure. One possibility that is worth investigating is that the focusing grid shown in Fig. 13.3 can be fabricated from the gettering material.

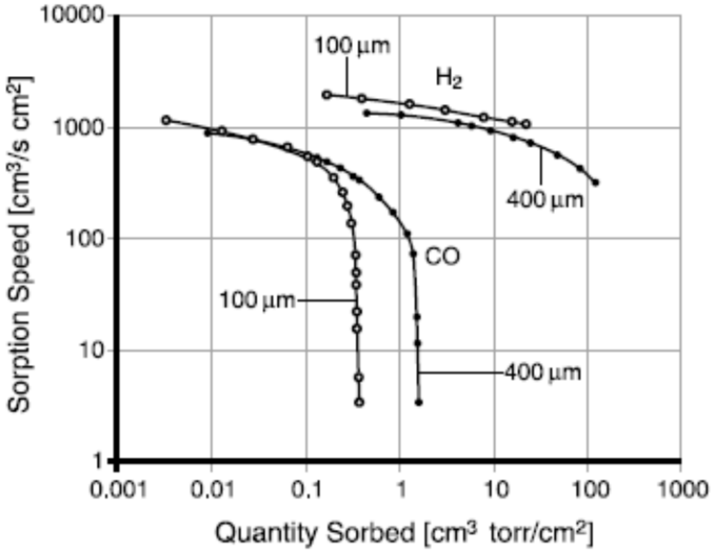


Figure 13.38. Sorption speed vs. quantity absorbed for CO and H_2 of a screen printable getter with thicknesses of 100 and 400 µm [72].

13.2.8. Driver Electronics

Figure 13.39 shows the basic dimensions of a 10-in. diagonal video graphics array (VGA) display that consists of 480 row grid lines and 640 × 3 column emitter lines [73]. The row lines are 300-µm wide with a 30-µm spacing, and the column lines are 92 µm wide with a 18-µm spacing. Thus, the size of a white pixel is 330 µm × 330 µm. Each 92 µm × 300 µm color pixel is partitioned into 4 × 13 = 52 subpixels of 23 µm×23 µm dimensions, and each subpixel contains four tips. The total number of tips per color pixel is thus 208 (only three tips are shown in Fig. 13.39).

Both emitter and gate drivers must be able to supply enough charging current I

$$I = \frac{cdV}{dt} \qquad (13.5)$$

to support the required scan rates [74]. Charging and discharging the matrix conductors dissipates power P in the resistance of the circuit, which is given by

$$P = cV^2 f \qquad (13.6)$$

where C is the capacitance, V the switching voltage, and f the frequency. Although the transient power is dissipated in the resistance of the circuit, the magnitude of the lost power is independent of the value of the resistance. Nevertheless, the resistance should be kept low in order to avoid excessive voltage drops along the emitter lines when most of the pixels are excited. In addition, the line resistance, capacitance, and switching voltage should all be kept low in order to minimize the driver circuit power requirements.

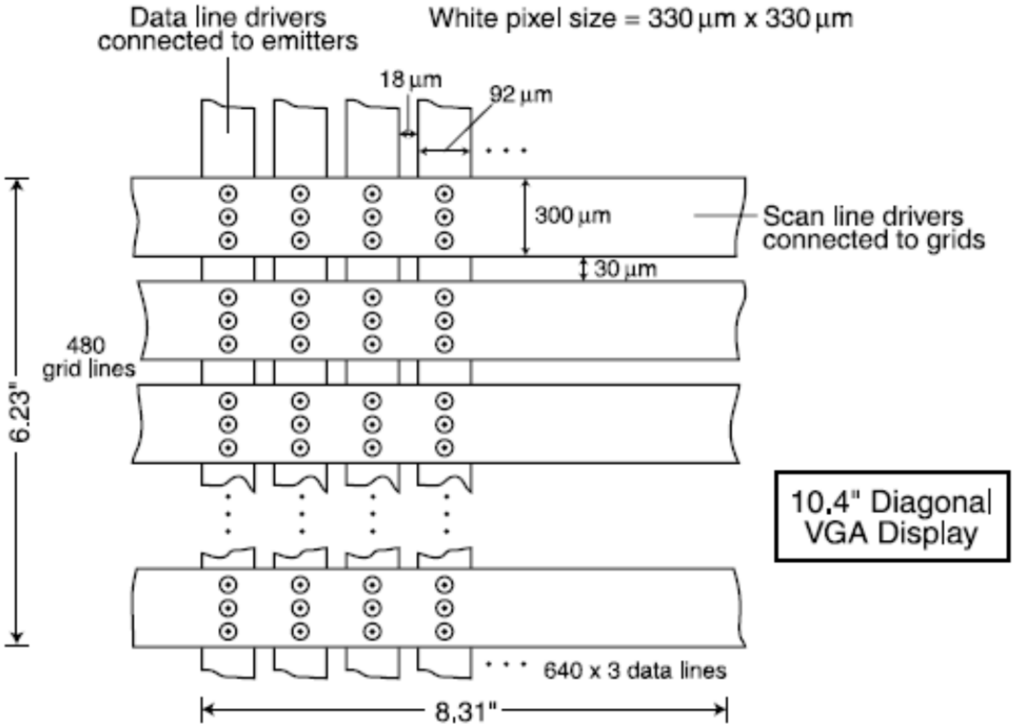


Figure 13.39. The row and column arrangement of a 10″ VGA FED prototype display [73].

The line resistance R of a 6.23-in. (15.82 cm) long and 92-µm wide line in Fig. 13.39 can be estimated by

$$R = l\rho / A \qquad (13.7)$$

where l is the line length, ρ the resistivity (~20 × 10^{-6} Ωcm), and A the cross section of the line. Assuming a 1-µm thick line, R is estimated as

$$R = \frac{15.82 \times 20 \times 10^{-6}}{92 \times 10^{-4} \times 1 \times 10^{-4}} = 344\Omega$$

The gate-to-emitter capacitance is given by

$$C = \varepsilon_0 \varepsilon A / d \qquad (13.8)$$

where $\varepsilon_0 = 8.86 \times 10^{-14}$ F/cm, $\varepsilon = 3.9$ for SiO_2, A is the area of the gate line to emitter line overlap (300 µm × 92 µm), and d is the thickness of the dielectric.
Assuming d = 1 _m, one obtains

$$C = \frac{8.86 \times 10^{-14} \times 3.9 \times 300 \times 10^{-4} \times 92 \times 10^{-4}}{1 \times 10^{-4}} = 9.54 \times 10^{-13} \quad \text{or 0.95 pF/color pixel overlap}$$

There are 480 overlaps along one emitter line, which means that the line capacitance is 480×0.95 pF=450 pF. In addition to the parallel plate capacitance, the tip-to-gate stray capacitance should be included in the calculations. The value of this capacitance depends strongly on both the gate diameter and tip geometry and is usually below 0.5 fF per tip [75]. Thus for 208 tips, the additional capacitance is less than 0.1 pF/color pixel, which would increase the total capacitance by less than 10%.

The required pixel current can be estimated from the equation for luminance, which is given by [27]

$$L = \frac{1}{\pi A} V I D_c \varepsilon \qquad (13.9)$$

where L is the luminance in cd/m^2, A is the faceplate area in m^2, V the anode voltage in volts, I the emission current in amps, D_c the duty cycle, and ε the luminous efficacy of the phosphor in lm/W. The current can then be calculated as

$$I = L\pi A / V D_c \varepsilon \qquad (13.10)$$

We assume that the display in Fig. 13.39 yields an overall luminance of 200 cd/m^2 at an anode voltage of 4000 V and a screen luminous efficacy of 10 lm/W. With a duty cycle of 1/480 (for sequential line addressing) and an area of 6.23 in. × 8.31 in. = 51.77 in.^{2} (0.0334 m^2), a total current of 0.252 A can be derived. The individual pixel current can be obtained by dividing this value by the number of pixels, which yields an average current per pixel of $0.252/480 \times 640 \times 3 \sim= 0.3\ \mu A$. Operating this display at 400 V increases the current to about 30 μA due to a 10-fold decrease in both voltage and luminous efficacy. This estimated range of 0.3–30 μA agrees well with the quoted values of 1–10 μA in Ref. [74] for the currents needed to excite a white pixel.

With the knowledge of current/color pixel, line capacitance, line resistance, and $I - V_g$ characteristics of a subpixel, the driver circuitry can then be designed. Typically, the row connections are made to the gates and the column connections to the emitter lines. The rows are scanned sequentially from top to bottom. During each row select time, the column connections are used to imprint intensity information to the pixels of the selected row.

There are three basic methods to modulate the pixel intensity with the column drivers [74]: pulse height (amplitude) modulation (PHM), pulse width modulation (PWM), and hybrid approach (PHM/PWM).Within each group, either the voltage or current drive methods can be selected. By choosing thePHMvoltage drive method, the column driver is a voltage source with sufficient current to support the requirements for slew rate. The drive capability based on this PHM voltage method is good. The circuit is relatively simple, but must maintain high accuracy at low emission levels. At low currents, the effect of ballast resistors is reduced due to insufficient voltage drops across the resistors. As a result, pixel emission tends to become nonuniform. Transient power is low, though, because the voltage excursions from one row select time to the next are minimal.

In the PHM current drive scheme, a current proportional to the pixel intensity is applied to the column conductor. The column voltage floats to a value corresponding to the applied current. Circuit implementation is easier in this current mode than the voltage drive, because the applied current is linear with brightness and requires less accuracy at lowbrightness levels. Transient power is also minimal, since there are little changes in column voltage from row to row. However, the current approach suffers from two significant disadvantages. First, since the current is limited to the emission level (which is much smaller than the charging currents), the voltage changes on the column conductors are too slow to support typical scan rates. Second, emitter defects that result in pixel leakages will distort the pixel brightness. This does not occur in the voltage drive method.

In the PWM, the pixel intensity is modulated in time. The pixels are operated at a constant current, but the pixel on-time is varied as a function of the pixel intensity. This can be accomplished with either voltage or current drive, or a combination of both. In the PWM voltage drive, during the row select time, a voltage pulse to the column turns the pixel on for a fraction of the time. Dark pixels are not turned on at all, while white pixels remain on for the entire row select time. Gray scale is achieved by varying the on-time between these two extremes. Since the driver is digital, the output voltage accuracy is much relaxed as compared to the PHM approaches. The transient power is large, since the same voltage transitions occur whether the pixel is at a low or high brightness level. Depending upon the number of gray scales, the slew rate can be quite high, resulting in 100 ns pulse width. This leads to high transient currents. The image uniformity is very good, since the pixel is always at a current level where significant voltage drops can develop at the ballast resistors.

In the PWM current drive, a constant current is forced to the column conductors with the on-time proportional to the pixel intensity, and again, the column voltage floats to a value corresponding to the applied current. This can potentially produce images with very high quality, except that limiting the currents also restricts the possible slew rate and, therefore, the number of possible gray levels. The pixel leakage will result in brightness distortion in the pixel. Transient power is much reduced from the PWM voltage method, but so is the ability to support the required bandwidth and gray scales. The PWM can also be operated in a combined voltage and current drive mode. This offers advantages of fast slew rates of the voltage drive and uniformity of the current drive. When the pixel is turned on, a voltage source is used to precharge the line to a voltage near that needed for the required pixel current. The voltage source is then switched to a high impedance state, and a current source is switched on for emission control. The main disadvantage of PWM is the requirement for very short pulse widths for low brightness pixels. Short pulses require fast rise and fall times, and result in very high charge currents.

The hybrid PHM and PWM method has distinct advantages over the PWM approaches, because it does not require a very high charging current at low pixel brightness. The PHM/PWM approach can also be implemented in either the voltage or current drive mode. Typically, the video data are split into two nibbles, and multiple current/voltage sources are used for the implementation.

In addition to the technical considerations, the overall driver electronics cost, driver size, IC technology, and chip processing yield have to be taken into account in order to select an appropriate display drive method. Based on the current cathode technology, the PWM voltage drive appears to be the better solution, since cathode nonuniformity and leakage prevents the use of either the PHM approach or the current drive. As the technology advances, other drive approaches will become possible. In the long run, if the problem of pixel leakage can be resolved, the current drive technique will offer the best solution for display uniformity.

Two custom chip implementations of data line drivers were described in Refs. [73] and [76]. In Ref. [73], a PHM voltage scheme was considered for driving the display shown in Fig. 13.39. The functional block diagram of the data line driver is presented in Fig. 13.40. The features of this data driver include 120 driver outputs, the support of interlaced and noninterlaced video signals, the support of interleaved and noninterleaved panel connections, the support of three types of color filter arrangements: vertical stripe, diagonal mosaic (top-left to bottom-right), and bicolor triangular patterns, and three phase clock inputs with frequencies up to 20 MHz. The major functional blocks are bidirectional shift register, level shifter, analog multiplexer, and sample-and-hold circuits.

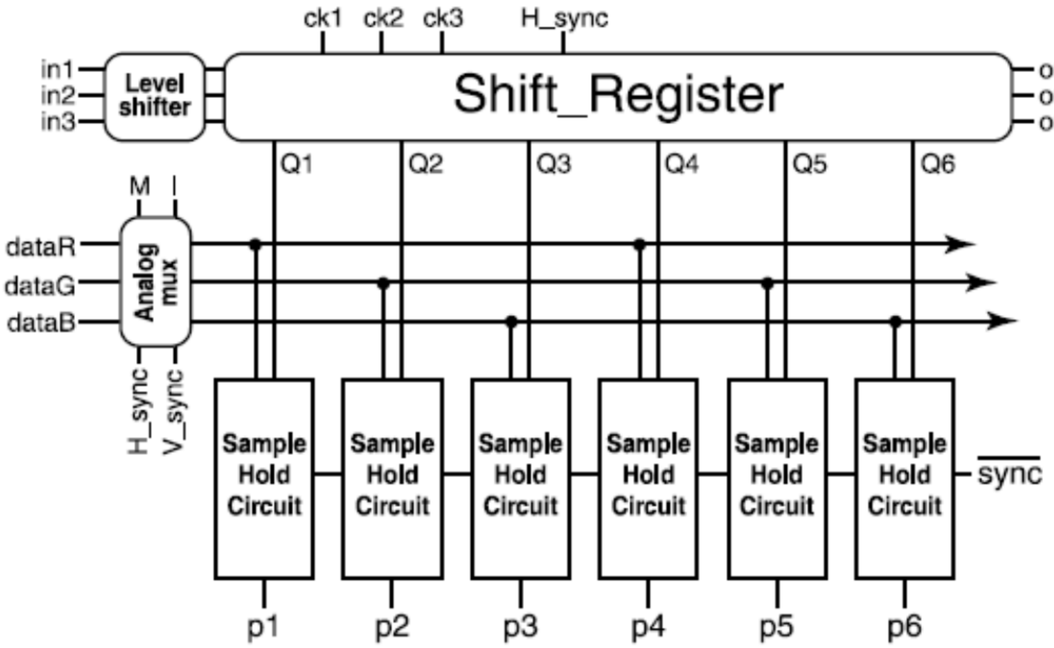


Figure 13.40. Functional block diagram of a pulse height modulation (PHM) data line driver [73].

The operation of the data line is repetitive per scan line and is synchronized by the horizontal synchronization signal (H sync). At the beginning of each scanning line, input pins In1, In2, and In3 are held high for one clock cycle. Thus, the outputs of the first three shift registers are high. These outputs are used to put the first three sample-and-hold circuits in the sampling mode. After one clock cycle, the outputs of the first three shift registers become low, and the outputs of the next three shift registers are high. Thus, the first three sample-and-hold circuits are in the hold mode, and the next three are in the sampling mode. As the clock goes on, the three logic l's in the shift register are shifted to the right by three steps, and each sample-and-hold circuit samples the video input in sequence. With 120 stages, the data line driver can handle 40 × 3 data lines. A panel of 640 × 3 data lines requires 16 driver chips. The driver is designed to support direct cascade connects between chips. The output pins Out1, Out2, and Out3 are connected to In1, In2, and In3 pins, respectively. At the end of a scan line, all 640 × 3 sample-and-hold circuits are holding the appropriate video signals. At the beginning of the next scan line, these values are output through the output buffer during the entire scan line period, while the other set of sample-and-hold circuits are sampling the current scan line video signals.

By using three phase clocks that are offset by 120°, the maximum operating frequency is reduced by a factor of three. Every input signal (except the R, G, B video signals) is connected to a level shifter. These shifters change the 0 to −5 V input signal swing to −19 to −5 V internal swing, which is needed to obtain the appropriate brightness on the selected rows. A MOS (metal-oxide-semiconductor) switch and a MOS capacitor are used to perform the sample-and-hold function. Two sets of sample-and-hold were used, thus allowing the unity gain buffer the entire scan line period to drive the data. To reduce power, a class AB operational amplifier was used as the output buffer. These buffers are able to supply large currents when the outputs are changing, and draw little or no current when the outputs reach their steady state. The dimensions of the chip are 7.62 mm × 17.50 mm.

In Ref. [76], the authors described a PHM current method and claimed that the resistive layer was not needed, and nonuniform pixel I–V characteristics could be allowed (probably assuming that there was no gate leakage). Figure 13.41 shows the schematic diagram of this driver scheme. The chip consists of both high- and lowvoltage devices. The high-voltage devices protect the low-voltage *n*-channel current sources when capacitively coupled high voltages develop at the cathode lines during the scan time. T1, T2, T3, T4 are low-voltage n-channel devices used for current sources. Their channel widths increase from W, $2W$, $4W$ to $8W$, which accommodate 16 grey scales for this experimental chip. When V_{gate} goes from ground to high (>60 V) during the scan time, the cathode lines will reach relatively high voltages due to the gate-to-cathode capacitive coupling. The induced high voltage is protected by T5 (high-voltage *n*-channel LDMOS (laterally diffused metal-oxide-semiconductor) with a breakdown voltage of 120 V). When the scan pulse (Vgate) goes from ground to high, node (B) is grounded by circuit block (A) for preventing floating the source of T5, and $V_{control}$ is low. After $V_{control}$ is high (T5 turns on), the digitized four-bit video data are applied to the gates of current sources (T1 through T4), and the proper current level is supplied to the cathode line for a given gray scale.

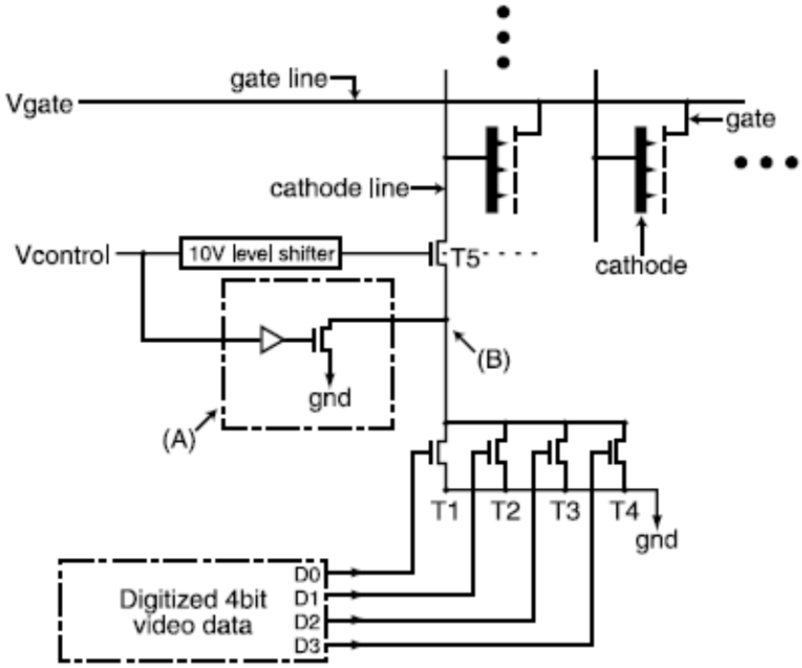


Figure 13.41. Schematic diagram of a constant-current data driver scheme [76].

The methods described above are primarily for passive matrix addressing. Active matrix addressing can be achieved by including thin film transistors (TFT) per subpixel in a similar scheme to the AMLCD display [77].

13.2.9. Aging/Surface Chemistry

Coulomb Load. In contrast to thermionic emitters where chemical impurities are burned off, field emitters are extremely sensitive to changes in surface chemistry. Submonolayer coverage of certain species can change the work function of field emitters significantly. In addition, sputtering events can change the surface morphologies of emitters. Not only do the emitter arrays have to maintain the performance during the life of the display, which should last at least 10 000 h, but the phosphors also need to operate satisfactorily and should not reach their Coulomb limit of 100–200 C/cm^2 during the projected operation period. It is interesting to observe that a phosphor dot in a CRT TV set is only excited by the electron beam for a total of 1.6 min during 10 000 h of operation [78]. If we assume a TV signal of 525 lines per frame and a typical 4:3 aspect ratio, there are about 700 addressable points in each horizontal line. Thus, the number of phosphor dots is 525 × 700 = 367,500. The dwell time is the frame time divided by the number of dots. With a frame time of 1/30 s, the dwell time is 90.7 ns. The number of times a phosphor dot gets excited (hit) in 10 000 h is 10 000 × (3600 s/h)/frame time (sec), and the total time that the dot is irradiated is the number of hits (=1.08 × 10^9) times the dwell time. This turns out to be 100 s or 1.6 min. The Coulomb loading that this phosphor dot experiences is $Q = I\ t/A = 1 \times 10^{-3} \times 100/10^{-2} = 10$ C/cm^2. Here we assume an average current of 1 mA/dot and an area of about 1 mm^2 per phosphor dot.

For a passive addressing scheme as discussed in relation to Fig. 13.39, the dwell time is (1/30)/480 (number of lines) or 69 μs. Since the number of hits one phosphor pixel sees during 10 000 h of operation is the same as in the CRT, i.e., life time/frame time, the total time one pixel is energized is $1.08 \times 10^9 \times 69 \times 10^{-6} = 74.52 \times 10^3$ s or 20.7 h. The average current per pixel for the same brightness can be estimated by equating the energy densities/pixel for the two cases. For the CRT, assuming l mA/phosphor dot, 1 mm^2 area, 30 kV phosphor voltage and a dwell time of 90 ns, the energy density is $I\ Vt_{Dwell}/A = 1 \times 10^{-3} \times 30000 \times 90 \times 10^{-9}/0.01 = 2.7 \times 10^{-4}$ Ws/cm^2. For the FED shown in Fig. 13.35, the area is 92 μm × 300 μm, and the dwell time is 69 μs. Assuming an anode voltage of 4 kV, the current is then calculated to be 0.27 μA/color pixel. This agrees well with the current of 0.3 μA obtained from Eq. (13.10). This then corresponds to a Coulomb load of

$$Q = \frac{It}{A} = \frac{0.27 \times 10^{-6} \times 20.7(3600s/h)}{92 \times 10^{-4} \times 300 \times 10^{-4}} = 72.9C/cm^{2}$$

The increase in the Coulomb load is only proportional to the ratio of the two operating voltages of 30 kV/4 kV. One can immediately see that the situation becomes critical when low-voltage FEDs are involved. Going from 4 kV operation to 400V, for instance, will increase the Coulomb loading to 729 C/cm^2, assuming that the luminous efficiency remains the same. The efficiency, however, also decreases by almost 1 order of magnitude as compared to the high-voltage phosphors. Expanding this argument further, one can conclude that an ideal application for FEAs lies in high brightness video billboards fabricated from individual picture element modules. These elements are usually operated at 10 kV, which brings the Coulomb loading to 30 C/cm^2. Thus, a factor of six in brightness can be achieved by increasing the pixel current without reaching the Coulomb loading limit. Since the average TV set operates at about 200 cd/m^2, a billboard of 1200 cd/m^2 can be built by operating it at 10 kV. By increasing the operating voltage to 20–30 kV, true sunlight readable billboards can be built with luminance exceeding 3000 cd/m^2 and operating lifes of longer than 10 000 h. Table 13.7 summarizes these comparative results.

Table 13.7. Summary of Coulomb Loads from Different Displays.

Display	Display Operating Time (h)	Operating Time of Emitters (h)	Operating Time of Phosphor (h)	Anode (Phosphor) Voltage V_{ph} (V)	Luminance L (cd/m^2)	Coulomb Load Q (C/cm^2)
CRT	10 000	10 000	0.027	30 000	200	10
FED (high voltage)	10 000	20[a]	20	4000	200	70
FED (low voltage)	10 000	20[a]	20	400	200	>700
Video billboard[b]	10 000	20	20	30 000	3000	150

[a] The 20 hours correspond to the emitter on-time. In most cases the column drivers have a continuous dc offset that could stress the gate dielectric and thus reduce product life.

[b] We assume a 480 Ч 640 Ч 3 pixel billboard that is a-line-at-a-time addressed.

Aging of Mo Emitters. In a typical packaged and sealed FED, the main gases present are oxygen, water, carbon dioxide, carbon monoxide, methane, and hydrogen. Gas evolution can take place by outgassing during sealing, electronstimulated desorption from the phosphor and other components, and desorption of process residues. Figure 13.42 shows some of the results obtained from a controlled experiment in an UHV (ultra-high vacuum) analytical chamber using a platinum-coated silicon wafer as the anode. This avoided any potential outgassing contribution from a phosphor anode. In the figure, the percentage changes in emission current are plotted against gas exposure ranging from 1 to almost 10 000 Langmuir (1 Langmuir = 1 × 10^{-6} torr s).

It can be seen that both methane and hydrogen increase the emission current, helium has no effect up to 1000 Langmuir, and carbon dioxide, oxygen, and water decrease the performance. The authors conjectured that in the case ofwater, the degradation mechanism is the interaction of oxygen with molybdenum, forming surface molybdenum oxides that have a higherwork function than clean molybdenum. Similar situations exist for oxygen and carbon dioxide where the tips become oxidized, thus resulting in higher local work function and decreased emission. Methane undergoes dissociation and ionization by the field-emitted electrons, in which the energetic hydrogen species react with Mo to form volatile Mo hydrides, which are pumped away by the vacuum system, leaving clean tip surfaces; the carbon free radicals interact with Mo to form stable Mo carbides. Hydrogen undergoes similar dissociation and ionization upon impact with electrons and forms volatile Mo hydrides. It was found that the effects of gas exposure are similar in both dc and pulsed mode operation of the device. Thus, dc mode testing can be used as an effective acceleration method in establishing the device lifetimes under various vacuum conditions.

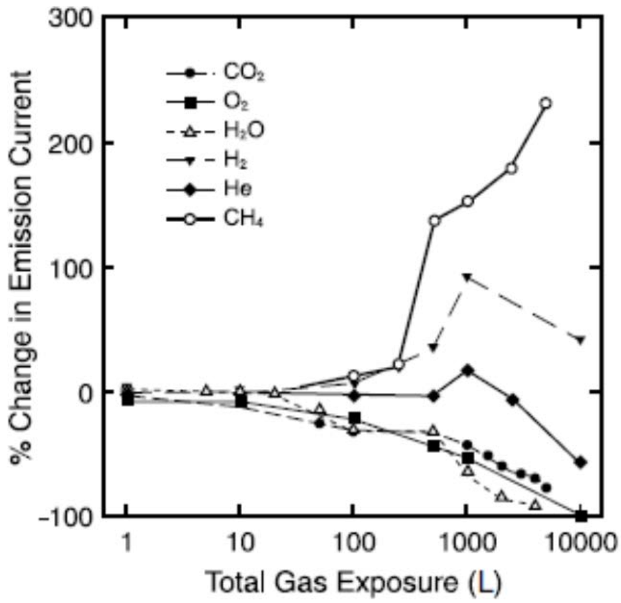


Figure 13.42. Effects of gas exposure on the emission characteristics of a 2.8 × 10^5 Mo Spindt emitter array [79].

Although the above measurements showsome trends in tip behavior upon exposure to certain gases, different results might be obtained in an enclosed FED vacuum envelope, where volatile species cannot be pumped away as readily as in a high vacuum test system. By using micro-Auger and X-ray photoelectron spectroscopy (XPS) measurements, tips that were part of a 320 × 240 matrix addressable array were investigated [80]. The panel was divided into two sections. One section served as a control area, and in the other the tips were operated at 50 μs pulses with an interval of 13 ms for 1800 h when the emission current had decreased by 50%. The anode phosphor used was ZnO:Zn. The vacuum was then broken, and analytical measurements were performed. Table 13.8 summarizes the micro-Auger results, in which the values quoted are the averages of nine tips measured, both in the test and control areas.

Table 13.8. Carbon and Oxygen Concentrations on Mo Tip Surfaces in Test and Control Areas as Measured by Micro-Auger Spectroscopy

Impurity	Tips in Test Area (Average of 9 Tips)	Tips in Control Area (Average of 9 Tips)
Carbon	10%	28%
Oxygen	47%	31%

From the Auger data on tips and XPS measurements performed over larger areas, it was concluded that the tips oxidized during operation in the vacuum envelope, and that the reduction in current was caused by the increased work function of the Mo oxide. Zinc from the phosphor anode was also found on the emitters with its XPS intensity three times higher in the test area as compared to the control area. No explanation was given for the reduction in carbon in the test area and how that might contribute to the change in emission behavior.

Aging of Carbon-Based Emitters. High temperature CVD deposited carbon thin film emitters are under intense investigation for display applications [81, 82]. Experiments similar to the ones described here showed severe degradation in emission currents when the carbon cathodes were operated in both water and oxygen at pressures of about 10^{-6} torr, and showed almost no changes when operated in hydrogen [83]. It is believed that both the oxygen and water molecules were ionized by the electrons, and they could subsequently modify and etch the surface of the carbon films. The fact that hydrogen has little effect on cathode lifetime might indicate that the surface termination of the carbon films by hydrogen was crucial. Figure 13.43 shows the emission current as a function of time for two vacuum frit sealed picture element tubes (PETs). It can be seen that the current decreases monotonically with time. Similar results of decreasing emission current as a function of time were obtained from the operation of PETs using carbon nanotubes as cathodes [84].

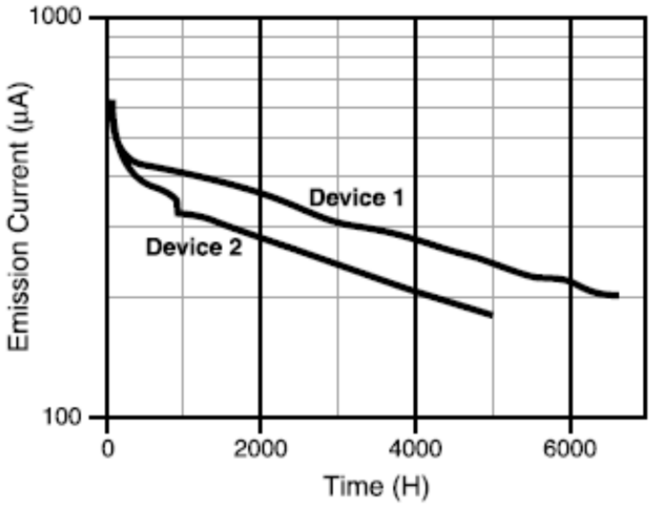


Figure 13.43. Cathode emission as a function of time for two frit sealed PETs [83]. The measurements were performed at an anode voltage of 7 kV using aluminized phosphor.

When graphite emitters were operated with a copper anode (instead of phosphors used in the experiments described above), no significant current decreasewas observed as shown in Fig. 13.44 [85]. Initially, the tube was operated under continuous pumping for about 500 h, and a slight decrease in current was observed. After pinch-off, the current increased monotonically until it showed signs of saturation at about 4000 h (the discontinuity at 1500 h was due to power supply failure during the measurement).

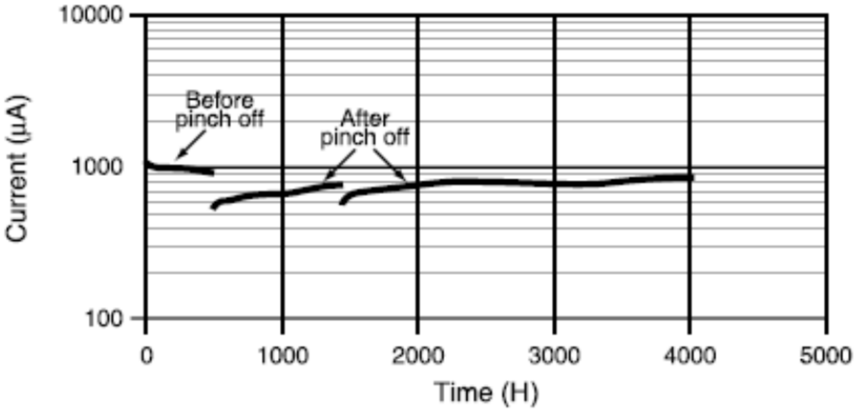


Figure 13.44. Emission aging of a nanocrystalline graphite emitter as a function of time. The tube was operated in a dc diode mode using a copper anode of 2 mm in diameter [85].

From the different aging behaviors of these carbon-based emitters and the Mo emitters, it becomes apparent that the interplay between emitters, phosphors, and sealing conditions must be both understood and controlled in order to manufacture a reliable and reproducible display. Some of the peculiarities of the *I–V* characteristics of carbon emitters when operated in a diode mode with different phosphors have been reported in Ref. [86]. The authors observed different *I–V* characteristics with ZnS:Cu,Al and $Y_3Al_5O_{12}$:Tb,Ge phosphors, as compared to an ITO-covered glass anode and an yttrium phosphor. Similar results were obtained in PETs using graphite emitters operated in the triode mode [85]. The triode PETs were formed by placing a metal mesh about 100 _m above the emitter surface. The phosphor anode was placed about 60 mm away from the grid. The area of the openings in the grid was 25% of the total area. Sets of R, G, B aluminized P22 phosphor tubes were fabricated using CRT frit sealing techniques. It was found that the electron transmission coefficient, which is defined as

$$T = \frac{I_{phosphor}}{I_{grid} + I_{phosphor}} \times 100 \qquad (13.11)$$

was 25% before the sealing for all three tubes, consistent with the grid area consideration. After frit sealing, the green tubes had emission coefficients ranging from 25 to 50%, the red tubes from 5 to 24%, and the blue tubes from 0.6 to 5%. This suggested that during the sealing process, changes on the emitter surfaces occurred, possible related to the outgassing from the phosphors. These changes apparently affected either the work function of the emitters, the number of active emission sites, or emission areas. Figure 13.45 shows the voltage dependencies of the grid and phosphor currents of green, red, and blue tubes that are shown in the upper left corner in Fig. 13.48 later in Section 13.3.2. For the green tube, the phosphor current was about half of the grid current, resulting in a transmission coefficient of 33%. For the red tube, the grid current was larger than that of the green tube, but the phosphor current was reduced by an order of magnitude at V_{grid} = 800 V. For the blue tube, the transmission coefficient is only about 6%. The offset in operating voltage could be caused by a larger grid-to-emitter distance or by surface changes during the sealing.

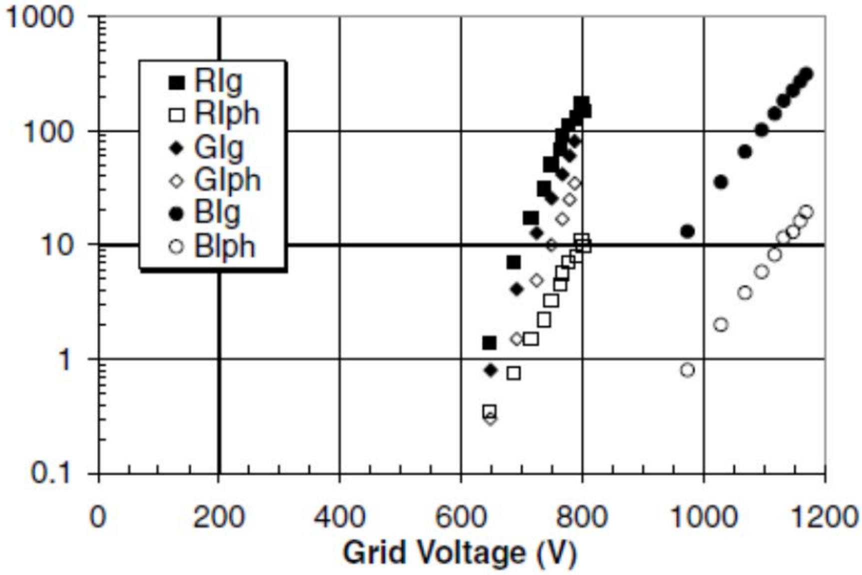


Figure 13.45. *I–V* characteristics of green, red, and blue PETs using carbon emitters. These tubes are shown in the upper left corner of Fig. 13.48. (RI_g = grid current of the red tube, RI_{ph} = phosphor current of the red tube, GI_g = grid current of the green tube, . . . etc.).

13.3. OTHER DISPLAY TECHNOLOGIES

So far, the main emphasis has been placed on FEDs intended for laptop applications. There are a couple of companies that have demonstrated 13- and 15-in. Prototypes and are probably investigating the possibility of scaling the technology to 40 in. For true, very thin, TV-on-the-wall applications. This is likely a very costly endeavor, since up-scaling to 40-in. requires cost-effective solutions to several technological hurdles. A potentially interesting solution was presented by SI Diamond Technology (SIDT) in its hybrid FED concept. Other applications for field emitters include high brightness and large area video billboards and a direct view IR image converter, in which IR-activated gates modulate the field at the tips and the optical readout is performed by a phosphor plate. There are other potential products such as X-ray fluorescence spectrometer, e-beam lithography, etc., but these are beyond the scope of this chapter.

13.3.1. Hybrid FED

In the hybrid FED (HyFED) structure, the three electron guns of a CRT are replaced by a matrix of $N \times M$ field emitting electron guns. Each gun rasters a small section of the screen; thus, the distance between the guns and the phosphor plate can be reduced. No spacers are needed in HyFEDs, and standard CRT faceplates can be used in manufacturing. But the device is still relatively bulky and heavy as compared to 40-in. plasma displays with 1-mm thick glass plates and the proposed TV-on-thewall FEDs using spacers. Implementation of the HyFED technology is not simple, since several metal grids are needed to provide focusing and deflection capabilities. Also, compensation schemes have to be developed to obtain seamless images near the boundaries of the individual electron guns. Figure 13.46 illustrates the reduction in display thickness from L (conventional CRT) to approximately L/N for the HyFED, where N is the larger number in the (N,M) matrix. Figure 13.47 shows the cross section of one of the $N \times M$ gun elements. Note that, in addition to the field emission extraction grid (B), there are eight more grids involved with a total thickness of the grid assembly at 3 mm (excluding the thickness of the emitter panel and the phosphor panel). This HyFED concept has been demonstrated in a 4-in. diagonal display with a 0.45 mm pixel pitch [87].

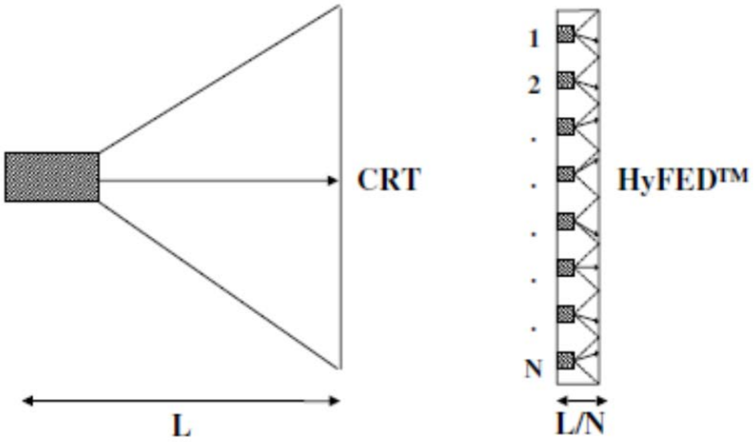


Figure 13.46. Cross sections of conventional CRT (left) and HyFED (right) devices.

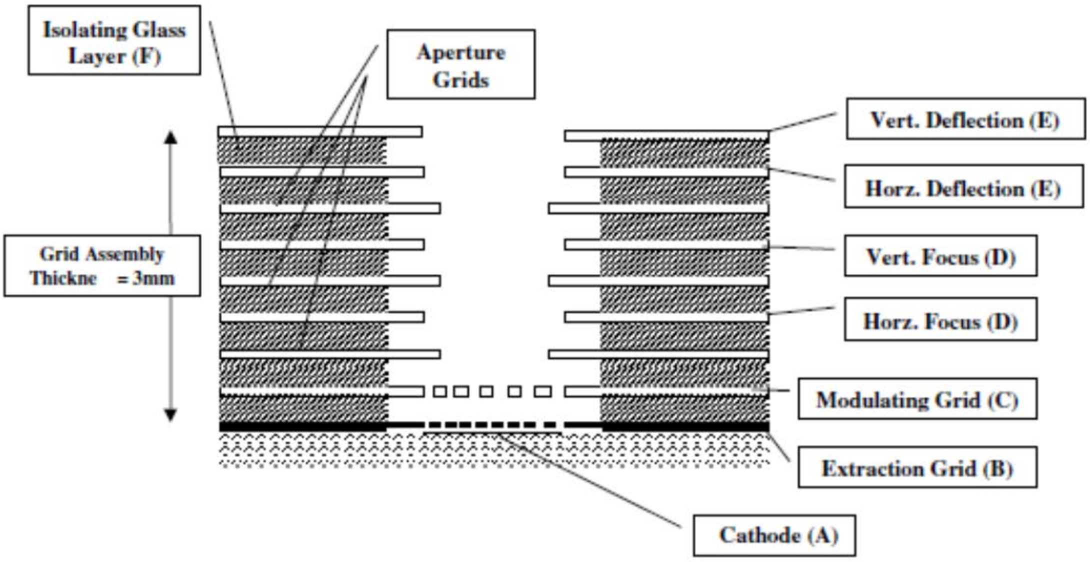


Figure 13.47. Cross section of one of the individual electron guns in a *N*×*M* HyFED gun array.

13.3.2. Picture Element Tubes

Picture element tubes (PETs) with hot filament emitters are used in giant screen displays [88]. They are available in either individual tube form with each tube having one color or in the modular form with each module having at least four white pixels. An exaggerated image of a giant display with the Statue of Liberty as foreground is shown in Fig. 13.48, and single color PETs and a four-white pixel module are shown in the upper and lower left corners, respectively. Implementing a large array of hot wires into one envelope is not an easy manufacturing task. Replacing these hot wires with the cold emitter technology is thus a logical choice in improving manufacturability and reducing power consumption. A typical hot cathode in a PET is operated at 45 mA and 0.45 V, which corresponds to 0.02 W. The same tube at a 10 kV phosphor voltage and an anode current of 100 μA dissipates a total of 1 W. Thus, the filament power is 2% of the maximum brightness (about 20000 fL) power of the tube. If the hot cathode is replaced with field emitters, this would translate into a 24.5 kW reduction in standby power for a 480 × 640 × 4 white pixel billboard (60 × 80 feet).

Figure 13.48. A conceptual giant video billboard composed of individual, high luminance, cold cathode PETs. Prototypes of these tubes, in single color elements and in a four white pixel module, are shown on the left (developed at Sarnoff under partial funding by Applied Photonics Technology, San Jose, CA.)

Figure 13.49 shows the cross section of a PET module that is operated with hot cathode filaments [89]. These modules are usually 3 in. × 3 in. in area and about 1 in. in depth. Sixteen of these modules are usually packaged into individually serviceable 1-ft × 1-ft units. Billboards of any size can then be assembled using these components. The metal anode prevents spillover of electrons to adjacent pixels. Highvoltage contact to the phosphors is provided by a spring-loaded contact placed inside the evacuation tube. The center of the plug is used for the outside high-voltage connection. There is plenty of space available for effective gettering, and the peripheral glass envelope of the module functions as the spacer. The front shield grid is a defocusing lens to spread the electrons as uniformly as possible across a given color subpixel. The processing yield is high, since the number of emitters per module is low.

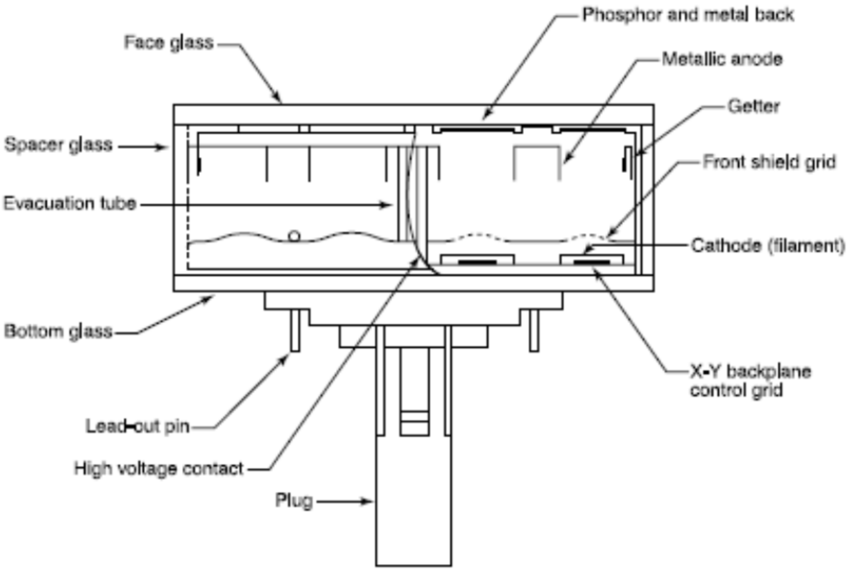


Figure 13.49. Cross section of a PET module that is operated with hot filament emitters [89].

Replacing the hot filaments with cold cathodes can be accomplished by several means. Using area emitters such as thin film carbon or carbon nanotubes, diode or triode structures can be formed. Triode structures are preferred, since addressing voltages are lower. At present, metal grids are placed about 10–100 μm above the emitters. Gated metal or Si emitter arrays can also be used, but only as point sources, since area emitters would be too expensive. All the advantages of built-in ballasting and flashover protection of silicon chips as described in Chapter 5 can be used in this approach. Silicon technology in the single chip form is only economical for low resolution displays, since pick-and-place operations and the cost per chip prohibit manufacturing of billboards with a large number of silicon chips.

Some of the implementation techniques are shown in Fig. 13.50 with corresponding prototype devices shown in Figs. 13.48 and 13.51. Method (a) was jointly developed by SIDT and ISE Electronics and was implemented in the device shown on the left side in Fig. 13.51. It shows a 64 white pixel module energized by carbon emitters and operated at 10 kV. For comparison, a 16-pixel conventional module with the same cross section shown in Fig. 13.49 was placed on the right side in Fig. 13.51. The reduction in module thicknesswas achieved by the use of proximity focusing. Method (c) was implemented in the device shown in the lower left corner in Fig. 13.48. It is a four white pixel (G, G, R, B), 3-in. × 3-in. module energized by individual gated silicon tip arrays, one array per color pixel. The size of the chip is about 1 mm × 1 mm, containing about 6000 tips. Method (d) was used in fabricating the individual picture element tubes shown in the upper left corner in Fig. 13.48. Spreading of the electrons exiting the grid takes place due to the lensing action of the extraction grid. The distance between the grid and the phosphor in these tubes is 2.5 in.

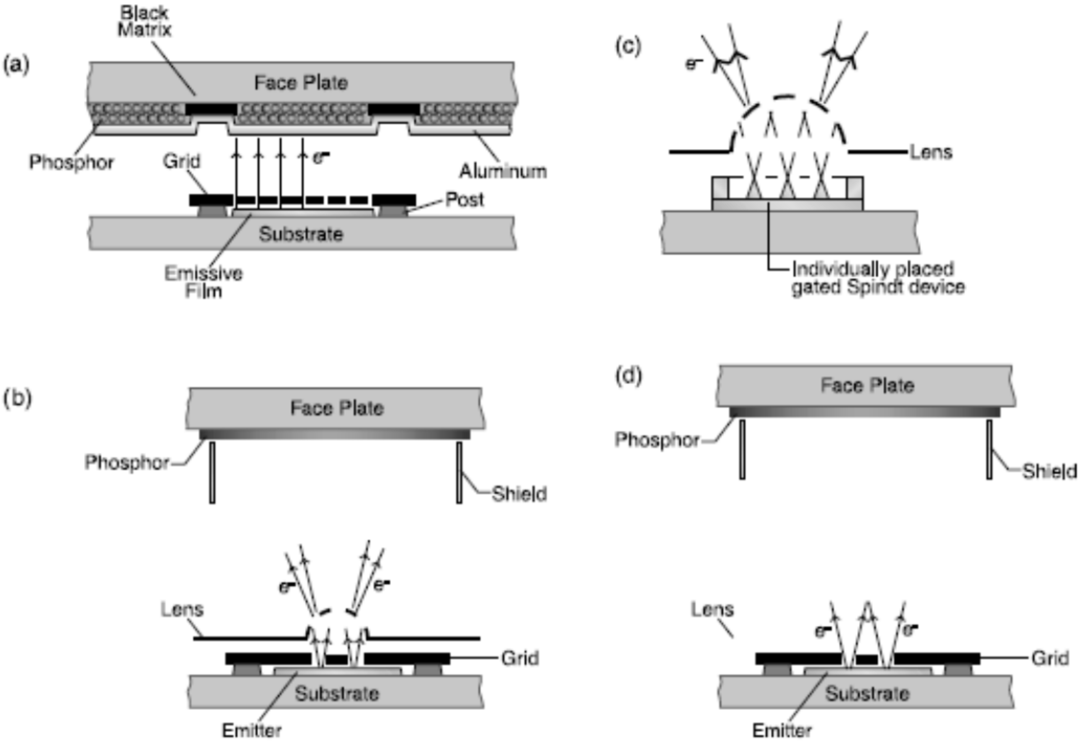


Figure 13.50. Emitter implementation schemes for replacing thermionic emitters with field emitters: (a) large area planar emitter triode operated in proximity mode; (b) “point” source planar emitter triode with defocusing lens; (c) “point” source gated Spindt emitter with defocusing lens; and (d) planar emitter “point” source without defocusing lens (some defocusing is obtained by the grid) [85].

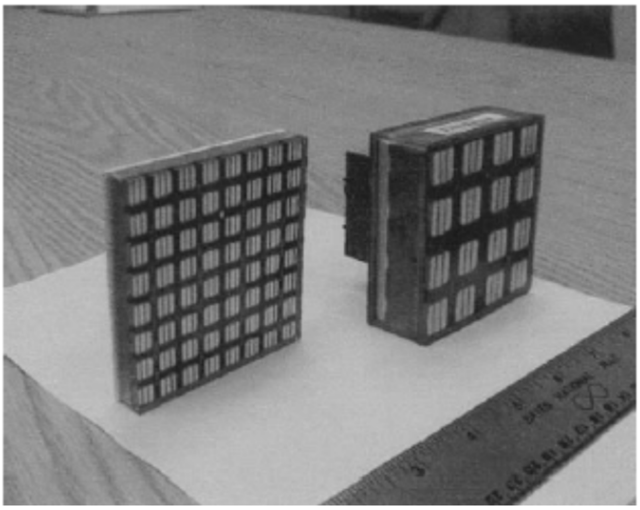


Figure 13.51. PET modules. Left, a 64 white pixel array using method (a) depicted in Fig. 13.46 (joint development of SIDT and ISE Electronics). Right, a conventional hot filament cathode module with its cross section shown in Fig. 13.49 [89].

13.3.3. Direct View Infrared FED

An interesting application that combines MEMS with FEAs for infrared (IR) imaging has been proposed [90]. The basic operation is illustrated in Fig. 13.52. The gate plate of an FEA is released with respect to the tips and is supported by two double-layer (SiC/Al) arms that end in thermal isolation elements anchored to the substrate (anchors are not shown). The infrared radiation that enters the silicon chip from the back is absorbed in the gate plate, causing the double-layer arms to move with respect to the tips. Thus, for a given gate voltage, the field at the tips is modulated, which causes the emission current to vary as a function of incoming IR power. Readout is accomplished by a monochrome FED phosphor. Sensitivity calculations predict that a 1K change in temperature should move a gate plate that is suspended by a 50-μm long cantilever arm by 0.2 μm, which causes the emission current to be modulated by a factor of 2–10, depending on the tip geometry.

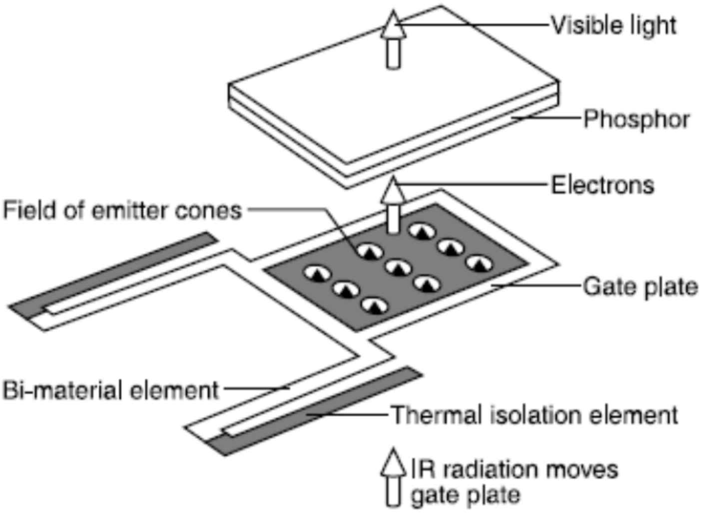


Figure 13.52. Schematic of a direct view IR FED [90].

However, when a gate voltage is applied, an electrostatic pull-in between the gate and the emitter substrate occurs. For a 0.85 μm gate-to-substrate separation, the pullin voltage is only 2–5

V for a 50-μm long cantilever for the chosen double-layer arm materials and thicknesses. One way to overcome this problem is to include a gate shield. By applying the gate voltage also to the shield, a field-free region (except near the tips) is created, allowing emission to occur without pulling in the gate. The shield can be processed in the region enveloping the tips, which avoids potential IR loss through the shield.

Figure 13.53 shows a micrograph of a single pixel element of a 16 × 16 array device, with each single pixel element containing 30 field emitters. The IR absorber plate is shield free and electrically isolated from the Ti–W gate. An additional contact is provided for electrically moving the structure by a small amount from its equilibrium position for compensation purposes. The cross section of the prototype package for such an IR-to-visible display is shown in Fig. 13.54. The intended application of this device is for night vision operation in an IR wavelength regime of 8–12 μm.

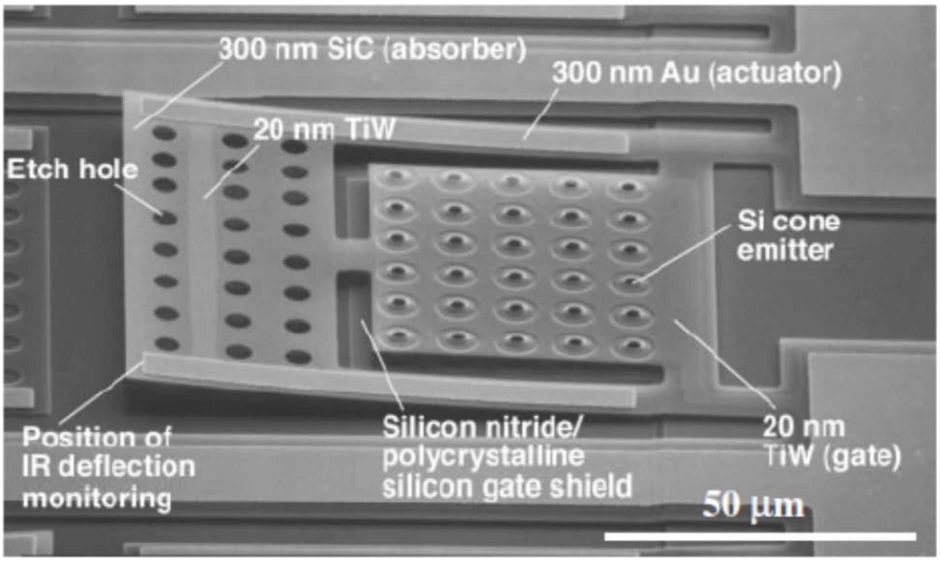


Figure 13.53. SEM micrograph of a 30 emitter IR sensitive pixel. The fully released cantilevered gate moves toward the tips when IR is absorbed from the back of the chip. IR-induced motion of about 1 μm has been demonstrated with this structure.

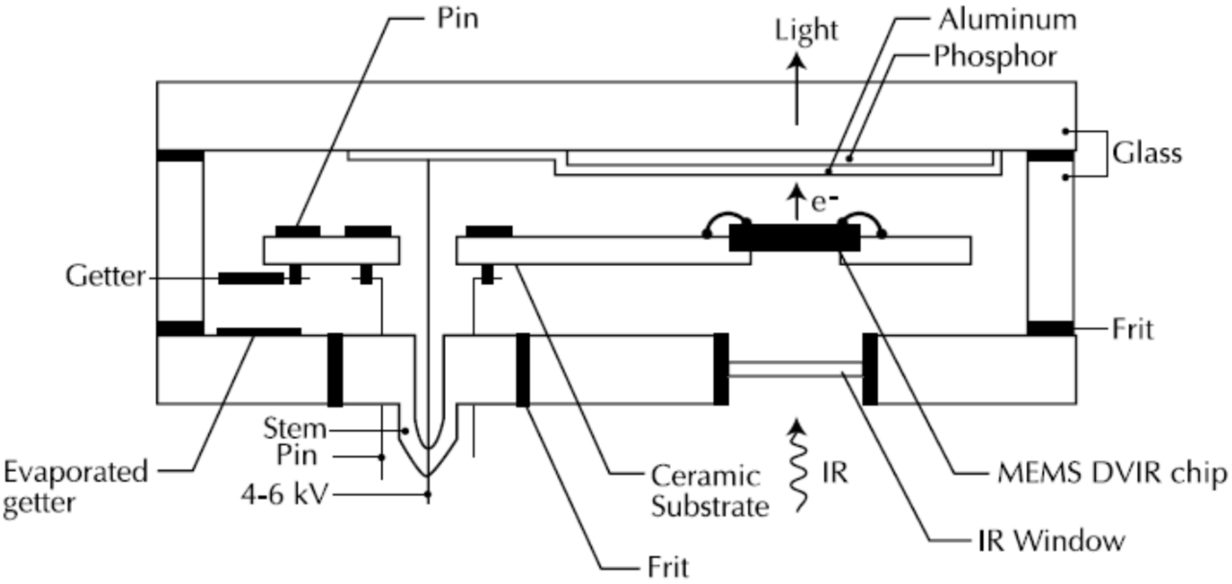


Figure 13.54. Cross section of a prototype direct view IR FED [90].

13.3.4. AMLCD and other technologies

Active matrix liquid crystal displays (AM LCDs) are the acknowledged leader in the portable flat panel display market They are the standard against which present and future alternatives will be compared (Fig. 13.55).

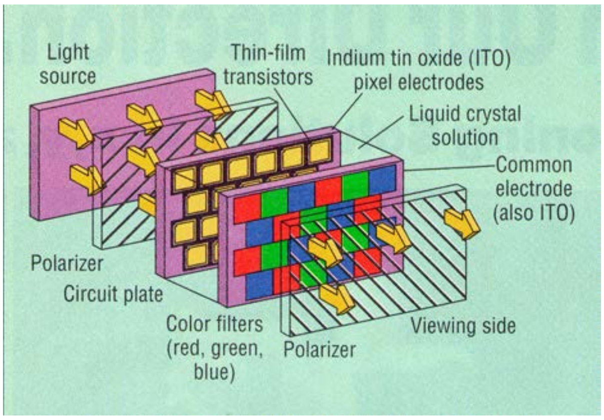


Figure 13.55. Cutaway view of a color AMLCD shows the components that go into the display.

Nematic liquid crystal cells, the foundation of these displays, are made of twisted rod-like molecules forming a dielectric between two electrodes. In the normal, unperturbed state, light passes freely through the translucent cell. When a voltage is applied across the gap, the molecules twist, blocking polarized light and turning the pixel black. Gray-scale values are obtained by varying the voltage across the cell and thus the degree of twist in the liquid crystal. Colors are obtained with discrete filters, so three red, green, and blue (or sometimes four, RGBG) subpixels make up a single pixel able to produce the color white.

In a passive matrix display, external circuitry pulses row and column electrodes. A particular pixel switches only when both its row and column electrodes are charged. The pixels of a passive matrix display can be addressed only by intersection of row and column drives, not individually. This means each picture element in a passive liquid crystal display (LCD) is driven sequentially rather than a line at a time and - just like a cathode ray tube - must stay "on" for the entire picture frame time. This long decay time requires the use of slow-response (and low-contrast) liquid crystal material. Slow response creates undesirable visual artifacts and slows the screen refresh rate. Video-rate displays - displays fast enough to keep up with rapidly moving objects, such as laptop mouse cursors or enemy aircraft - are impossible with passive LCD addressing.

In an AMLCD, each subpixel (or color dot in a color display) forms a capacitor connected to the drain of a thin-film transistor (TFT). When the TFT is switched on, it applies a persistent voltage across a single capacitor. Fast-response liquid crystal can be used. AMLCDs offer brilliant color, high contrast, and relatively fast screen refresh rates.

However, AMLCDs are not perfect. Video-rate displays are still difficult. All LCDs require an external light source, either ambient light or (usually) a backlight mounted behind the device plane. Unlike an emissive display, where only lit pixels consume power, the backlight must be on at full power no matter what is being displayed on the screen. In an active matrix display, about 95% of this light, and the power used to generate it, is lost to polarizers, color filters, circuitry, and the liquid crystal itself. Furthermore, since the liquid crystal manipulates *polarized* light, viewing angle is limited. In laptop computers, presently the most visible segment of the display market, these limitations are irritating. In automotive, aircraft, and head-mounted applications, they can affect the success or failure of a product.

Even more important, especially in the extremely competitive consumer electronics market, AMLCDs are expensive to manufacture. Unlike conventional 1C manufacturing where hundreds of devices can be built on one wafer, only six to eight laptop-size displays will fit on a glass plate. Each display has hundreds of thousands of TFTs (at least three dots per pixel, with 640 x 480 pixels in a standard VGA display), and it does not take many defective TFTs to ruin the entire panel. Worse, this demanding specification must be met with amorphous silicon (a-Si) deposited on glass: a much less uniform surface than that of a typical single-crystal silicon wafer. Amorphous Si cannot switch fast enough for video-rate driver circuitry, so separate driver ICs must be attached to each of the 500 or so back plane data lines - a tedious process that accounts for as much as 30% of the cost of the display. Some analysts estimate that AMLCD yields are in the 50% range.

It has been found a solution to the a-Si problem. The transistor arrays are built on a thin film of single-crystal silicon on top of a sacrificial layer. After the display and driver circuitry is fabricated using standard IC processing, the entire circuit is lifted off, transferred to a glass plate, and packaged with polarizers, filters, etc. Since the crystalline silicon transistors are smaller than their a-Si counterparts, they allow a greater percentage of backlighting to pass through the LCD panel for a brighter display. This technology is best suited to small displays, such as head-mounted units, and to projection systems, and the company has chosen to focus on these areas.

Meanwhile, other companies are investigating use of polysilicon for both driver circuitry and TFT arrays. Use of polysilicon would simplify display manufacture, since the a-Si and therefore require expensive quartz substrates. However, it is possible to introduce a laser annealing system for fast recrvstallization on ordinary glass substrates.

Ferroelectric liquid crystal displays (FLCDs) are also being considered by some researchers. Unlike nematic liquid crystal, which relaxes to its untwisted state as soon as the switching voltage is removed, ferroelectric crystal remains stable in the opaque state until deliberately switched off. Thus, simple, passive, matrix-like addressing would be sufficient for an FLCD display.

Liquid crystal displays aren't the only technology, either. In addition to field emission displays, electroluminescent (EL) displays are practical for some applications. An EL panel (Fig. 13.56) consists of a thin, luminescent phosphor layer sandwiched between transparent dielectric layers and a matrix of row and column electrodes. EL displays are emissive - they produce their own light - so the complications of backlighting, polarizing, and filtering are avoided, as are the yield problems associated with large transistor arrays. According to company president Jim Hurd, EL display achieves 80% yield with a zero pixel defect specification. EL displays are well-suited for large public information boards, medical equipment, and other displays that need to be visible under a variety of conditions and viewing angles (Fig. 13.57). Their brightness and manufacturing simplicity are an advantage in head-mounted displays as well. However, the full-color products, which require a much more complex manufacturing process and an efficient blue phosphor, present a major challenge.

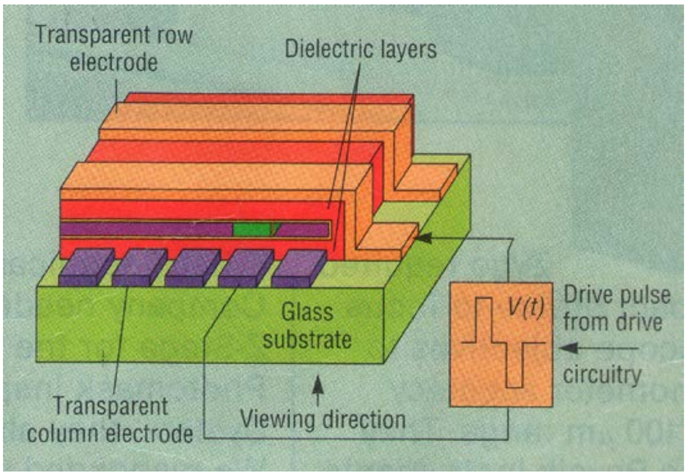


Figure 13.56. Cross-sectional view of an electroluminescent display. The display consists of an electroluminescent glass panel and a mounted circuit board with drive and control electronics.

Figure 13.57. A large electroluminescent display in the passenger waiting area, Helsinki Airport, Finland.

In a plasma display, matrix address electrodes excite neon or other gases to produce ultraviolet radiation, which excites phosphors to produce colored light. Plasma displays also consume much more power than AMLCDs and are probably unsuitable for portable display applications. Their brightness may earn them a niche in the large stationary display market, however. Plasma is the most likely technology for wallhung TVs in the near future.

13.4. FED BASED ON CARBON NANOTUBES.

Recently, Ise Electronic Corp. has demonstrated a CRT-lighting-element, which used MWNTs as the cold emitter source, with a lifetime of more than 5000 hours [8]. Further more, Samsung shows a matrix-addressable diode display using CNTs-emitters [9]. CNTs show potential in the application for vacuum fluorescent display (VFD) and field emission display (FED). In this work, we use DC arc discharge process to synthesize CNTs. The electron emission characteristics of CNTs films, which were patterned by printing process, will be investigated and a prototype of CNTs-FED will be demonstrated in this paper.

The anode used in DC arc discharge process was a pure graphite rod with 15 cm in length, and the cathode was a graphite disc (18 mm in diameter) mounted into water-cool copper. The chamber was maintained in helium atmosphere (500mbar). Arc constant DC current (70-80 A) was applied to the anode and cathode separated with a gap about 2-3 mam, whereas the voltage across the gap was 20-24 V.
The morphology and structure of the CNT-bulks were examined using scanning electron microscopy (SEM). Field emission characteristics of the CNT-emitters were measured by a diode technique using an electrometer. Anode plate, a glass coated with ITO layer and P15 phosphor, was separated from the cathode by using 90 um spacers. The turn-on field (E0) were designated as the field at which the Fowel-Nordheim plot, In (I/V 2)-(1/V) deviates from the straight line. To fabricate CNT emitters, silver paste was first printed on glass as bottom electrode. The tube was crushed using ball milling to reduce its length to 1-2 um. The slurry containing CNTs was then printed upon silver paste. The fully sealed CNTs-FED was fabricated using VFD-like vacuum sealing process. All the interfaces between the anode plate/spacer 94 In *Display Technologies Ill*, I-Wei Wu, Heiju Uchiike, Editors,
Proceedings of SPIE Vol. 4079 (2000) **e** 0277-786X/00/$15.00
bar/CNTs-emitter plate were sealed with glass frits, followed by thermal annealing to enable the reaction of the glass with
glass frits.

SEM micrographs of the bulk, shown in Fig. la and **lb,** indicate that consists of multi-wall nanotubes embedded in fibrous bundles. Each bundle is approximately 50 um in diameter. The nanotubes are about several tens of nm in diameter. The carbonaceous particles attached on the CNTs can easily be removed by post treatment at 600Z in air. The yields of nanotubes embedded in fibrous bundles are higher than 70% after thermal treatment. To reduce the length of nanotubes, ball milling was introduced to crush it for 2 hours. To study the characteristics of CNTs emitters, the silver was first screen-printed in dot pattern on the glass as cathode. Each dot has 1 x **I** mm2 in area. The slurry containing CNTs and binders was then screen-printed upon silver paste, followed by oven curing. Figure lc shows a cross-sectional SEM micrograph of CNTs-emitter. Most of CNTs are aligned perpendicularly to substrate after surface treatment process on the surface. The density of CNTs-emitters was markedly larger than the typical density of microtips in conventional Spindt-type FEDs. Further, the P15 phosphor was screen printed on another ITO glass plate as anode. The emission property of CNTs-emitter was measured using diode structure, as shown in Fig. 2. The CNTs-emitters turns on at a field as low as 1.45 V/urn and the current rise to 0.5 mA under electric field of 5.5 V/um. The current significantly increased to 3.4 mA under 5.5 V/urn applied field after surface treatment. The resultant fabrication process is shown in Fig. 3. The gap spacing between both plates, separated by the spacer bar, is 90 um. All the interfaces between the anode/plate/sapcer bar/CNTs-emitter plate was sealed with glass frits, followed by thermal annealing to enable the reaction of the glass with glass frits. Exhaust process was performed with a conventional vacuum pump, which is almost the same as VFD vacuum sealing process.

To study the image and measure the characteristics of sealed FED device, the anode is applied with high voltage (300- 350 V) and the cathode electrodes are switch-driven. Fig. 4 shows a sequence of the images of a counting clock. The images indicated that each dot could be controlled well using an MOSFET as a current switch. At the on-state, the voltage between gate and source was 5 V, i.e., turned on the channel. The voltage drop between the drain and source are neglected which compared to anode voltage. At the off-state, the voltage between gate and source was 0 V, thus there is no induced current in the channel due to floating.

Moreover, to demonstrate the driving of CNTs-FED, a seven-segment numerical indicator was designed using diode structure. Each segment has 1 mm wide and 5 mm long in size. The system of indicator consists of two components, control and driving circuits, as shown in Fig. 5. The control circuit, using FPGA, operated under duty cycle of **1/10-9/10.** The high voltage driving circuit was n-channel open drain interface IC. Fig. 6 shows the emission image of seven-segment indicator.

The brightness is as high as 500 nits when applied with 300 volts between anode and cathode. The voltages of 5 V and 12 V are applied to FPGA and driving IC, respectively. During 1000 mm testing, the emission current decay is about 14% and saturated after 50 hours under 280 V applied anode voltage. It is still testing.

4. **CONCLUSION**

In summary, the seven-segment numerical indicator shows the excellent electron emission characteristics of CNTs-FED using diode structure. The brightness is as high as 500 nits when applied with 300 volts between anode and cathode, using open drain driver IC. Most of CNTs are aligned perpendicularly to substrate and its emission current density significant increase after surface treatment process on the surface. During 1000 min testing, the emission current decay is about 14% and saturated after 50 hours under 280 V applied anode voltage. It indicates that CNTs-FED is very promising for application of flat panel display. A 25-inch color proof-of-concept CNT-based TV has been demonstrated in September 2005.

13.5. SUMMARY

In this chapter, we have attempted to profile the excitement that is associated with the development of FED technologies. The excitement lies in the interdisciplinary approach of combining large-area FEAs, high-voltage phosphors or the still developing low-voltage phosphor technology, and large-area, small-gap vacuum technology for the production of high resolution, medium-size displays. Material science, surface chemistry, and physics all become intimately involved in the development of key display components such as stable emitters, efficient phosphors, and reliable flashover-resistant spacers.

The main portion of the chapter relates to FEDs using cone-shaped FEAs that are brought in close vicinity to a phosphor plate. The rest of the chapter focuses on encapsulation and aging issues and methods of addressing the panel. There is currently great motivation in exploring non-FEA-based emitter technologies, particularly carbon-based emitters, with the aim of lowering the production cost and enhancing the device reliability. For example, researchers from Samsung recently presented display prototypes based on two different gated versions (undergate triode and mesh triode) of carbon nanotube emitters as low cost alternatives to Spindt-type emitters [91]. The same group also incorporated electron multiplying microchannel plates (MCPs) into the display [92,93]. By generating secondary electrons inside of SiO_2 covered pores, the brightness of the display increased by a factor of three to four for the same emission current from the tips. By properly designing the MCPs, focusing electrodes can be eliminated for high voltage phosphor displays, because the MCPs can act as a focusing element.

FEDs are still a nascent technology that requires a fair amount of development for the devices to reach the marketplace. The most exciting prototypes today are the 13-in. 600 × 800 × 3 Candescent display operating at 5000 V and the 15-in. PixTech display operating at 6000 V. These prototypes are currently being refined to meet the stringent requirements for laptop applications and groomed for mass production so that they can compete effectively with other mature display technologies, most notably LCDs. For video billboard displays of sizes 2×4 ft^2 and larger, the specifications are not as stringent as lap top FEDs. Here, the conventional hot filament approach can be adequately replaced by the use of more efficient field emitters, because spacers are no longer needed, very efficient high-voltage phosphors can be used, and enough space exists within the vacuum envelope to provide for sufficient gettering. Other applications such as TV-on-the-wall using a matrix of field emitting electron guns and an IR imaging display are also being contemplated.

Reference to Chapter 13

1. S. lijima, "Helical microtubules of graphite carbon", Nature 354,pp56-58, 1991.

2. Saito,Susumu, Science ,V278,p77(1997).

3. P. G. Collins, A. Zettle, and R. E. Smalley, "Nanotube Nanodevice", Science, 278,pOO(1997).

4. Walt A. De Heer, A. Chatelain, D. Ugarte , "A carbon nanotube field emission electron source", Science, 270, pp 1179-1180, 1995.

5. J.M. Bonard, J.P. Salvetat, T. Stockli, "Field emission from single-wall carbon nanotube films", Appl. Phys. Lett. 73, pp9l8-920, 1998.

6. Q. H. Wang , T. D Corrigan , J. Y. Dai and R. P. H. Chang ,"Field emission from nanotube bundle emitters at low fields", Appl. Phys. Lett., 70, p3308(1997).

7. P. G. Collins, and A. Zettl ,,"Unique characteristics of cold cathode carbon-nanotube-matrix field emitters", Phys. Rev.

95

B55, pp9391-9399, 1997.

8. S. Uemura , T. Nagasako , J. Yotani ,T. Shimojo, SID 98 DIGESTp1052(1998).

9. W. B. Choi, D. S. Chung, S. H. Park, J. M. Kim, "Fully sealed high-brighmess carbon-nanotube field emission display", Appl. Phys. Lett.75, pp3l29-3131, 1999.

10. J. M. Lauerhaas, J. Y. Dai, A. A. Setlur, and R. P. H. Chang, J.Mater. Res., 12 No.6, p15 3 6(1 9 9 7).

11. P. M. Ajayan ,Ph. Redlich, and M. Ruhle, J.Mater. Res., 12, No. 1, p24 4 (19 9 7).

12. A. G. Rinzler, J. Liu, and R. E. Smalley, "Large scale purification of single wall carbon nanotubes", Appl. Phys. A, 67, pp2 9 -3 7 1998.

13. J. M. Mao, L. F. Sun, L. X. Qian, Z. W. Pan, B. H. Chang, W. Y. Zhou, G. Wang and S. S. Xie, "Growth of highly orinted carbon nanotubes by plasma-enhanced hot filament chemical vapor deposition", Appl. Phys. Lett. 72, p3297-3299, 1998.

14. G. Che, B. B. Lakshmi, and E. R. Fisher, Chem. Mater, 10, p2 6 0 (1998).

1. B. R. Chalamala, B. E. Gnade, and Y. Wei, *IEEE Spectrum*, **35**, 42 (1998).

2. R. Dawson, Y. Fukuda, S. Miyaguchi, S. Ishizuka, T. Wakimoto, J. Funaki, H. Kubota, T. Watanabe, H. Ochi, T. Sakamto, M. Tsuchida, I. Ohshita, H. Nakada, and T. Tohma, Organic LED full-color passive-matrix display, in *SID Digest*, 1999, p. 430.

3. H.F. Gray, G.J. Campisi, US Patent 4,964,946 (1990).

4. A. Wagal et al., US Patents 4,987,007 and 5,098,737.

5. T. S. Fahlen, Performance advantages and fabrication of small gate openings in Candescent’s thin CRT, in *Proceedings IVMC*, 1999, p. 56.

6. R. Meyer, in *Workshop Digest*, *Eurodisplay*, 1993, p. 189.

7. K. Werner, *IEEE Spectrum* **34**, 40 (1997).

8. N. Kumar, H. K. Schmidt, M. H. Clark, A. Ross, B. Lin, L. Fredin, B. Baker, D. Patterson, W. Brookover, C. Xie, C. Hilbert, R. L. Fink, C. N. Potter, A. Krishnan, and D. Eichman, Development of nano-crystalline diamond-based field-emission displays, in *SID Digest*, 1994, p. 43.

9. K. Betsui, *Technical Dig. 4th Int. Vacuum Microelectronics Conf.*, Nagahama, Japan, p. 26 (1991).

10. R.B. Marcus et al., *Appl. Phys. Lett.*, vol. 56, p. 236 (1990).

11. C. Xie et al., "Electron field emission from amorphic diamond thin films", *Proc. Intl. Vacuum Microelectronics Conf.,* Rhode Island (1993).

12. H. H. Busta, *J. Micromech. Microeng*. **2**, 43 (1992).

13. J. Itoh, K. Tsuburaya, S. Kanemaru, T.Watanabe, and S. Itoh, *Jpn. J. Appl. Phys.* **32**, 1221 (1993).

14. H. Busta, G. Gammie, S. Skala, J. Pogemiller, R. Nowicki, J. Hubacek, D. Devine, R. Rao, andW. Urbanek, Volcano-shaped field emitters for large area displays, in *Technical Digest IEDM*, 1995, p. 405.

15. H. H. Busta, *J. Micromech. Microeng.* **7**, 37 (1997). Also US Patent 5,930,590, Fabrication of volcano-shaped field emitters by chemical-mechanical polishing, (1999).

16. D. S. Y. Hsu and H. F. Gray, A low-voltage, low-capacitance, vertical multi-layer thinfilm-edge dispenser field emitter array electron source, in *Proceedings IVMC*, 1998, p. 82.

17. E. Yamaguchi, K. Sakai, I. Nomura, T. Ono, M. Yamanobe, N. Abe, T. Hara, K. Hatanaka, Y. Osada, H. Yamamoto, and T. Nakagiri, A 10-in. surface-conduction electron-emitter display, in *Technical Digest SID*, 1997, p. 52.

18. M. Okuda, S. Matsutani, A. Asai, A. Yamano, K. Hatanaka, T. Hara, and T. Nakagiri, Electron trajectory analysis of surface conduction electron emitter displays (SEDs), in *SID Digest*, 1998, p. 185.

19. A. P. Burden, H. E. Bishop, M. Brierley, C. Hood, P. G. A. Jones, W. Lee, R. J. Riggs, V. L. Shaw, and R. A. Tuck, Field emitting inks for consumer-priced broad-area flat-panel displays, in *Proc. IVMC*, 1999, p. 62.

20. I. Musa, D. A. I. Munindrasdasa, G. A. J. Amaratunga, and W. Eccleston, *Nature*, **395**, 362 (1998).

21. M. Borel, J. F. Borenat, R. Meyer, and P. Rambbaud, Electron source with micropoint emissive cathodes and display means by cathodoluminescence excited by field emission using said source, US Patent 4,940,916 (1990).

22. R. Meyer, Electron source with microtip emissive cathodes, US Patent 5,194,780 (1993).

23. R. S. Nowicki, H. H. Busta, J. E. Pogemiller, A. R. Forouhi, I. Bloomer, W. M. Clift, J. L. Yio, and T. Felter, Studies of rf-sputtered CrSiOx and SiC as series resistor films in FEDs, in *SID Digest*, 1996, p. 456.

24. J. D. Levine, R. Meyer, R. Baptist, T. E. Felter, and A. A. Talin, *J. Vac. Sci. Technol. B* **13**, 474 (1995).

25. C. Constancias and R. Baptist, *J. Vac. Sci. Technol. B* **16**, 841 (1998).

26. R. Baptist, F. Bachelet, and C. Constancias, *J. Vac. Sci. Technol. B* **15**, 385 (1997).

27. J. D. Levine, *J. Vac. Sci. Technol. B* **14**, 2008 (1996).

28. S. Itoh, T. Niiyama, M. Taniguchi, and T.Watanabe, A new structure of field emitter arrays, in *Proc. IVMC*, 1995, p. 99. Also US Patent 5,786,659, Field emission type electron source, (1998).

29. J. P. Spallas, R. D. Boyd, J. A. Britten, A. Fernandez, A. M. Hawryluk, M. D. Perry, and D. R. Kania, *J. Vac. Sci. Technol. B* **14**, 2005 (1996).

30. M. Anandan, G. Jones, and A. P. Gosh, Field emission display for commercial avionics applications, in *Workshop Digest of Technical Papers, Asia Display*, 1998, p. 131.

31. X. Chen, S. H. Zaidi, and S. R. J. Brueck, *J. Vac. Sci. Technol. B.* **14**, 3339 (1996).

32. T. T. Doan, J. B. Rolfson, T. A. Lowrey, and D. A. Cathey, Method to form self-aligned gate structures around cold cathode emitter tips using chemical mechanical polishing technology, US Patent 5,229,331 (1993).

33. L. E. Tannas, Jr. (Ed.), *Flat-Panel Displays and CRTs*, Van Nostrand Reinhold Co.: Ch. 3, New York, 1985.

34. N. Yocom, Sarnoff Corporation, private communications.

35. H. Bechtel and H. Nikol, Applicability of zinc sulfide phosphors in low voltage cathode ray displays, in *Workshop Digest, Asia Display*, 1998, p. 283.

36. L. E. Shea, The Electrochemical Society *Interface* **7**, 24 (Summer 1998).

37. S. S. Chadha, in *3rd Int. Conf. on the Science and Tech. of Display Phosphors*, Huntington Beach, CA, Nov. 1997.

38. P. H. Holloway, J. Sebastian, T. Trottier, H. Swart, and R. O. Peterson, *Solid State Technol*. **38**, 47 (1995).

39. S. Jacobsen, *J. SID* **4**, 331 (1996).

40. C. Stoffers, S. Yang, S. M. Jacobsen, and C. J. Summers, *J. SID* **4**, 337 (1996).

41. S. S. Kim, S. H. Cho, J. S. Yoo, S. H. Jo, and J. D. Lee, The effect of the resistivity of ZnGa2O4:Mn phosphor screens on the emission characteristics of field emitter array, in *Proc. IVMC*, 1997, p. 676.

42. S. A. Bukesov, N. V. Nikishin, A. O. Dmitrienko, S. L. Shmakov, and J. M. Kim, Electrophysical characteristics and low-energy cathodoluminescence of VFD and FED screens, in *Proc. IVMC*, 1997, p. 663.

43. C. J. Summers and B. K.Wagner, The role of surface modifications and coating in improving phosphor performance, in *Workshop Digest, Asia Display*, 1998, p. 261.

44. R. Ravi, Plasmaco, Private communications.

45. J. E. Jang, J. H. Gwak, Y. W. Jin, N. S. Lee, J. E. Jung, and J. M. Kim, Polyvinyl alcoholslurry screening of phosphors for full-color SVGAFED applications, in *Proc. IVMC*, 1999, p. 402.

46. M. J. Shane, J. B. Talbot, B. G. Kinney, E. Sluzky, and K. R. Hess, *J. Electrochem. Soc.* **165**, 334 (1994).

47. W. D. Kesling and C. E. Hunt, *IEEE Trans. Electronic Devices* **42**, 193 (1995).

48. T. S. Fahlen, Focusing properties of thin CRTTM display, in *Technical Digest SID*, 1999, p. 830.

49. J. Itoh, Y. Tohma, K. Morikawa, S. Kanemaru, and K. Shimizu, *J. Vac. Sci. Technol. B* **13**, 1968 (1995).

50. Y. C. Lan, J. T. Lai, S. H. Chen, W. C. Wang, C. H. Tsai, K. L. Tsai, and C. Y. Shen, Simulation of focusing field emission devices, in *Proc. IVMC*, 1999, p. 72.

51. Notes from Technology Forum, Coloray Display Corporation, November 1994.

52. X. Ma, P. G. Muzykov, and T. S. Sudarshan, *J. Vac. Sci. Technol. B* **17**, 769 (1999).

53. X. Ma and T. S. Sudarshan, *J. Vac. Sci. Technol. B* **17**, 1 (1999).

54. X. Ma and T. S. Sudarshan, Flashover performance of thin-wall spacers in field emission displays, in *IS&T/SPIE Conference on Flat Panel Display Technology*, Vol. 3636 0277-786/99, 1999, p. 143.

55. W. Taylor, N. S. Xu, R. V. Latham, B. J. Goddard, W. Kalbreier, J. Tan, H. J. Hopman, J. Verhoeven, and D. J. Chivers, The effects of surface treatments on the high voltage vacuum performance of alumina based insulators, Poster presented at *IVMC*, Asheville, NC, 1998.

56. R. V. Latham, Flashover across solid insulator in a vacuum environment [Appendix], in *High Voltage Vacuum Insulation*: *The Physical Basis*, Academic Press: New York, 1981, p. 229.

57. J. Browning, C.Watkins, D. Dynka, D. Zimlich, and J. Alwan, Field emission displays for desktop monitors, in *Workshop Digest of the 18th Int. Display Research Conference, Asia Display*, 1998, p. 149.

58. R. V. Latham, J. A. Mir, B. Goddard, J. Tan, andW. Taylor, The electro-optical signature of pre-breakdown electron injection from “electron pin-holes” at the cathode triple junction of a bridged vacuum gap, in *Proceedings of the 3rd International Conference on Space Charge in Solid Insulators*, Tours, France, June 1988.

59. N. C. Jaitly and T. S. Sudarshan, *IEEE Trans. on Elec. Insul.* **23**, 261 (1988).

60. G. Blaize and C. LeGressus, Charge trapping-detrapping processes and related breakdown phenomena, in *High Voltage Vacuum Insulation: Basic Concepts and Technological Practice* (R. V. Latham, Ed.), Academic Press: London, UK, 1995, Ch. 9, p. 329.

61. M. J. Kofoid, *Trans. AIEE (USA), Pt. III*, **79**, 991 (1960).

62. R. A. Anderson, Time-resolved measurements of surface flashover of conical insulators, Annual Report CEIDP, p. 475, 1978.

63. A. D. Johnson, H. H. Busta, R. S. Nowicki, Fabrication system, method and apparatus for microelectromechanical devices, US Patent 5,903,099 (1999).

64. H. Kim, B. K. Ju, K. B. Lee, M. S. Kang, J. Jang, and M. H. Oh, Influences of ambient gases upon emission characteristics of Mo-FEAs during frit sealing process, in *Proc.IVMC*, 1998, p. 67.

65. S. J. Jung, K. J.Woo, K. S. Kim, G. J. Moon, M. S. Kim, N. Y. Lee, and S. Ahn, The effect of ambient gas in sealing process on field emission characteristics, in *Proc. IVMC*, 1999, p. 40.

66. J. M. Kim, N. S. Lee, J. E. Jung, H.W. Lee, J. Kim, Y. J. Park, S. N. Cha, J.W. Kim, J. H. Choi, N. S. Park, J. E. Jang, S. S. Hong, Y.W. Jin, I. T. Han, and S. Y. Jung, Development of high voltage operated 5.2__ full color field emission displays, in *Proc. IVMC*, 1999, p. 54.

67. H. H. Busta and R. W. Pryor, *J. Vac. Sci. Technol. B* **16**, 1207 (1998).

68. H. H. Busta and R. J. Espinosa, The need for standardization of field emitter testing, in *Workshop Digest, Asia Display*, 1998, p. 113.

69. Y. R. Cho, J. Y. Oh, H. S. Kim, J. D. Mum, and H. S. Jeong, A new panel structure for field emission displays, in *Proc. IVMC*, 1997, p. 271.

70. W. B. Choi, B. K. Ju, Y. H. Lee, S. J. Jeong, N. Y. Lee, M. Y. Sung, and M. H. Oh, *J. Electrochem. Soc.*, **146**, 400 (1999).

71. B. Gnade and J. Levine, SID Seminar Lecture Note, Vol. 1, M-8/3, 1995. Eq. 4 was derived by Prof. Wiley P. Kirk of Texas AMU.

72. S. Tominetti, Chemical pumping in flat CRTs, in *Workshop Digest, Asia Display*, 1998, p. 101.

73. C. C. Wang, J. C. Wu, and C. M. Huang, Data line driver design for a 10__ 480 × 640 × 3 color FED, in *Proc. IVMC*, 1996, p. 557.

74. R. T. Smith, Drive approaches for FED displays, in *Proc. Int. Display Research Conference*, 1997, p. F-35.

75. H. H. Busta, J. E. Pogemiller, W. Chan, and G. Warren, *J. Vac. Sci. Technol. B* **11**, 445 (1993).

76. Y. S. Na, O. K. Kwon, C. H. Hyun, and J.W. Choi, A new driver circuit for field emission display, in *Workshop Digest, Asia Display*, 1998, p. 137

77. Y. D. Lee, Seoul National University, Private communication.

78. A. Silzars, Society for Information Display, Private communication.

79. B. R. Chalamala, R. M. Wallace, and B. E. Gnade, Residual gas effects on the emission characteristics of active Mo field-emission cathode arrays, in *SID Digest*, 1998, p. 107.

80. Y.Wei, B. R. Chalamala, B. G. Smith, and C.W. Penn, Surface chemical changes on field emitter arrays due to device aging, in *Proc. IVMC*, 1998, p. 46.

81. Z. Li Tolt, R. L. Fink, and Z. Yaniv, Carbon thin film cathode for large area displays, in *Proceedings 5th Int. Display Workshops*, Kobe, Japan, 1998, p. 651.

82. V. M. Polushkin, S. N. Polyakov, A. T. Rakhimov, N. V. Suetin, M. A. Timofeyev, and V. A. Tugarev, *Diamond Relat. Mater*. **3**, 531 (1994).

83. Z. Li Tolt, L. H. Thuesen, R. L. Fink, and Z. Yaniv, Field emission carbon thin film and its life time and stability, in *Proc. IVMC*, 1999, p. 160.

84. J.Yotani, S. Uemura, T. Nagasako,Y. Saito, and M.Yumura, Proceedings, High-luminance triode panel using carbon nanotube field emitters, in *Proc. 6th Int. Display Workshop*, Sendai, Japan, 1999, p. 971.

85. H. Busta, R. Espinosa, A. Rakhimov, N. V. Suetin, M. A. Timofeyev, M. E. Kordesch, P. Bressler, M. Schramme, J. R. Fields, and A. Silzars, *Solid State Electron*. **45**, 1039 (2001).

86. B. V. Seleznev, A. V. Kandidov, A. T. Rakhimov and N. P. Sostchin, Peculiarities of emission current flow in diode-structured FEDs, in *Workshop Digest, Eurodisplay*, 1996, p. 207.

87. R. C. Robinder, Z. Yaniv, R. L. Fink, Z. Li Tolt, and L. Thueseu, Application of diamond thin film cathodes to the hybrid field emission display/CRT (HyFED), in *Proceedings 5th International Display Workshops*, Kobe, Japan, 1998, p. 655.

88. M. Morikawa, Y. Seko, H. Kamogawa, S. Uemura, and T. Shimojo, Study to improve the flood-beam CRT for giant screen display, in *Workshop Digest*, *Japan Display*, 1992, p. 385.

89. N. Hasegawa, *Display Devices*, Dempa Publications, 1998, p. 33.

90. H. Busta, R. Amantea, D. Furst, C. Ling, L. White, and J. M. Chen, MEMS direct view infrared vision system (DVIR) with FED readout, in *Proc. IVMC*, 2000, p. 232.

91. N. S. Lee, J. H. Kang,W. B. Choi, Y. S. Choi, Y. J. Park, H. Y. Kim, Y. J. Lee, D. S. Chung, J. E. Jung, C. J. Lee, J. H. Kim, J. H. Yon, S. H. Jo, C. G. Lee, and J. M. Kim, Triode structure field emission displays using carbon nanotube emitters, in *Proc. IVMC*, 2000, p. 193.

92. W. Yi, S. Jin, T. Jeong, J. Lee, S. G. Yu, and Jongmin Kim, New electron multiplying channel for high efficiency field emission display, in *Proc. IVMC*, 2000, p. 9.

93. T. Jeong, J. Lee, S. G. Yu, S. Jin, J. Heo,W. Yi, and Jongmin Kim, The secondary electron emission characteristics for sol-gel based SiO_2 thin films, in *Proc. IVMC,* 2000, p. 20

Chapter 14

Non-volatile memory

Nonvolatile semiconductor memory (NVSM) is one of the most important discoveries in the field of electron devices since the invention of the transistor. Invented by D. Kahng and S.M. Sze in 1967, NVSM, today, is the largest electronic industry in the world. This single invention has subsequently given rise to a large family of memory devices including flash memory and electrically erasable programmable read-only memory (EEPROM).

1. All modern electronic devices, like mobile phones, notebook computers, digital cameras, MP3 music players, personal digital assistants, digital televisions, automotive engine control units, portable medical diagnostic systems, USB flash personal discs, and global positioning systems, are all based on NVSM. In the past 20 years, NVSM's minimum feature length has been scaled down from 0.8 µm to sub-20 nm—a 40-fold reduction. Many alternative NVSM technologies are being developed to reduce the size (sub-10 nm) and improve device performance.

Nonvolatile memory technologies in Si-based electronics date back to the 1990s. Ferroelectric field-effect transistor (FeFET) was one of the most promising devices replacing the conventional Flash memory facing physical scaling limitations at those times.

A variant of charge storage memory referred to as Flash memory is widely used in consumer electronic products such as cell phones and music players while NAND Flash-based solid-state disks (SSDs) are increasingly displacing hard disk drives as the primary storage device in laptops, desktops, and even data centers.

The integration limit of Flash memories is approaching, and many new types of memory to replace conventional Flash memories have been proposed. Emerging memory technologies promise new memories to store more data at less cost than the expensive-to-build silicon chips used by popular consumer gadgets including digital cameras, cell phones and portable music players. They are being investigated and lead to the future as potential alternatives to existing memories in future computing systems.

Emerging nonvolatile memory technologies such as magnetic random-access memory (MRAM), spin-transfer torque random-access memory (STT-RAM), ferroelectric random-access memory (FeRAM), phase-change memory (PCM), and resistive random-access memory (RRAM) combine the speed of static random-access memory (SRAM), the density of dynamic random-access memory (DRAM), and the nonvolatility of Flash memory and so become very attractive as another possibility for future memory hierarchies.

Many other new classes of emerging memory technologies such as transparent and plastic, three-dimensional (3-D), and quantum dot memory technologies have also gained tremendous popularity in recent years. Subsequently, not an exaggeration to say that computer memory could soon earn the ultimate commercial validation for commercial scale-up and production the cheap plastic knockoff.

Data storage devices can be classified based on many functional criteria. Of them, silicon-based semiconductor memories are categorized into two: volatile and nonvolatile. In volatile memories, the information eventually fades while power supply is turned off unless the devices used to store data will be periodically refreshed. Both types of memories can be further classified based on the memory technology that they use and based on data volatility as shown in the classification flow chart depicted in Figure 1. Volatile memories consist mostly of DRAM, which can be further classified into SDRAM and mobile RAM which only retain information when current is constantly supplied to the device. Another small but very important memory device is SRAM.

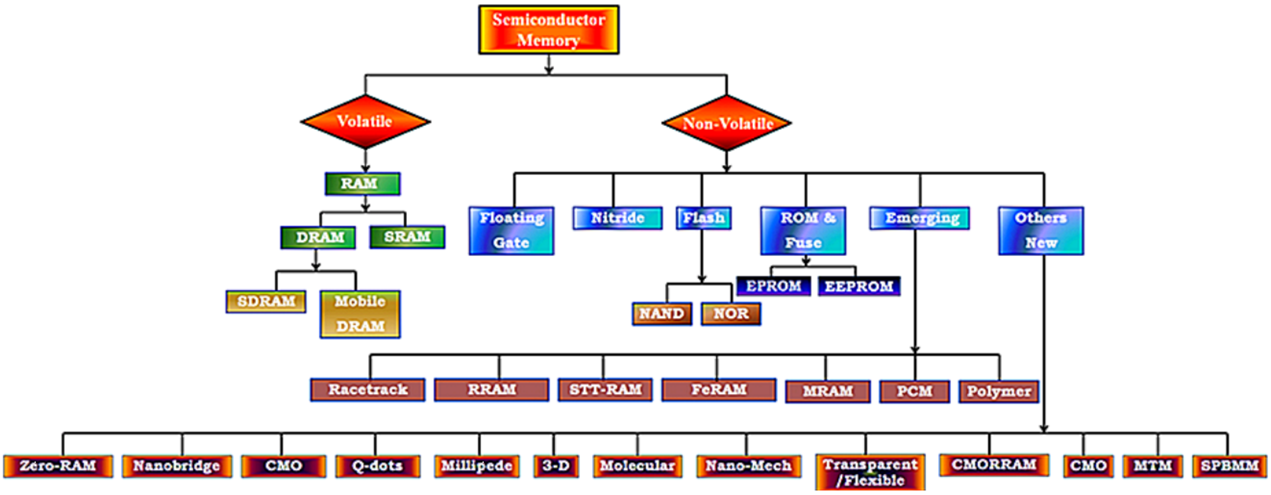


Fig. 1. Flow chart for the semiconductor memory classification according to their functional criteria.

The market for DRAM devices far exceeds the market for SRAM devices, although a small amount of SRAM devices is used in almost all logic and memory chips. However, DRAM uses only one transistor and one capacitor per bit, allowing it to reach much higher densities and, with more bits on a memory chip, be much cheaper per bit. SRAM is not worthwhile for desktop system memory, where DRAM dominates, but is used for its cache memories. SRAM is commonplace in small embedded systems, which might only need tens of kilobytes or less. Forthcoming volatile memory technologies that hope to replace or compete with SRAM and DRAM include Z-RAM, TTRAM, A-RAM, and ETA RAM. In the industry, new universal and stable memory technologies will appear as real contenders to displace either or both NAND Flash and DRAM.

Flash memory is presently the most suitable choice for nonvolatile applications for the following reasons: Semiconductor nonvolatile memories consist mostly of the so-called ‘Flash’ devices and retain their information even when the power is turned off. Other nonvolatile semiconductor memories include mask read-only memory (MROM), antifuse-based one-time programmable (OTP) memory, and electrically erasable read-only memory (EEPROM). Flash is further divided into two categories: NOR, characterized by a direct write and a large cell size, and NAND, characterized by a page write and small cell size. Nonvolatile memory is a computer memory that can retain the stored information even when not powered. Nonvolatile semiconductor memories are generally classified according to their functional properties with respect to the programming and erasing operations, as shown in the flow chart described in Figure 1.

These are floating gate, nitride, ROM and fuse, Flash, emerging, and other new next-generation memory technologies. Today, these nonvolatile memories are highly reliable and can be programmed using a simple microcomputer and virtually in every modern electronic equipment, which are expected to replace existing memories.

Among them, emerging nonvolatile memories are now very captivating. The next-generation memory market will cover up these emerging memory technologies. There are mainly five types of nonvolatile memory technology: Flash memory, ferroelectric random-access memory (FeRAM), magnetic random-access memory (MRAM), phase-change memory (PCM), and RRAM. Nonvolatile memory, specifically ‘Flash’ memory, which is characterized by a large-block (or ‘sector’) erasing mechanism, has been the fastest growing segment of the semiconductor business for the last 10 years. Some of these newer emerging technologies include MRAM, FeRAM, PCM, spin-transfer torque random-access memory (STT-RAM), RRAM and memristor. MRAM is a nonvolatile memory. Unlike DRAM, the data is not stored in an electric charge flow, but by magnetic storage elements. The storage elements are formed by two ferromagnetic plates, each of which can hold a magnetic field, separated by a thin insulating layer. One of the two plates is a permanent magnet set to a particular polarity; the other's field can be changed to match that of an external field to store memory.

STT-RAM is an MRAM (nonvolatile) but with better scalability over traditional MRAM. The STT is an effect in which the orientation of a magnetic layer in a magnetic tunnel junction or spin valve can be modified using a spinpolarized current. Spin-transfer torque technology has the potential to make MRAM devices combining low current requirements and reduced cost possible; however, the amount of current needed to reorient the magnetization is at present too high for most commercial applications. PCM is a nonvolatile random-access memory, which is also called ovonic unified memory (OUM), based on reversible phase conversion between the amorphous and the crystalline state of a chalcogenide glass, which is accomplished by heating and cooling of the glass.

It utilizes the unique behavior of chalcogenide (a material that has been used to manufacture CDs), whereby the heat produced by the passage of an electric current switches this material between two states. The different states have different electrical resistance which can be used to store data. The ideal memory device or the so-called unified memory would satisfy simultaneously three requirements: high speed, high density, and nonvolatility (retention). At the present time, such memory has not been developed. The floating gate nonvolatile semiconductor memory (NVSM) has high density and retention, but its program/erase speed is low. DRAM has high speed (approximately 10 ns) and high density, but it is volatile. On the other hand, SRAM has very high speed (approximately 5 ns) but limited from very low density and volatility. It is expected that PCM will have better scalability than other emerging technologies. RRAM is a nonvolatile memory that is similar to PCM. The technology concept is that a dielectric, which is normally insulating, can be made to conduct through a filament or conduction path formed after application of a sufficiently high voltage.

Arguably, this is a memristor technology and should be considered as potentially a strong candidate to challenge NAND Flash. Currently, FRAM, MRAM, and PCM are in commercial production but still, relative to DRAM and NAND Flash, remain limited to niche applications. There is a view that MRAM, STT-RAM, and RRAM are the most promising emerging technologies, but they are still many years away from competing for industry adoption. Any new technology must be able to deliver most, if not all, of the following attributes in order to drive industry adoption on a mass scale: scalability of the technology, speed of the device, and power consumption to be better than existing memories. The NVSM is in inspiring search of novel nonvolatile memories, which will successfully lead to the realization and commercialization of the unified memory. In progress, another new class of nonvolatile memory technologies will offer a large increase in flexibility compared to disks, particularly in their ability to perform fast, random accesses.

Unlike Flash memory, these new technologies will support in-place updates, avoiding the extra overhead of a translation layer. Further, these new nonvolatile memory devices based on deoxyribonucleic acid (DNA) biopolymer and organic and polymer materials are one of the key devices for the next-generation memory technology with low cost. Nonvolatile memory based on metallic NPs embedded in a polymer host has been suggested as one of these new cross-point memory structures. In this system, trap levels situated within the bandgap of the polymer are introduced by the NPs.

Memory devices play a massive role in all emerging technologies; as such, efforts to fabricate new organic memories to be utilized in flexible electronics are essential. Flexibility is particularly important for future electronic applications such as affordable and wearable electronics. Much research has been done to apply the flexible electronics technology to practical device areas such as solar cells, thin-film transistors, photodiodes, light-emitting diodes, and displays. Research on flexible memory was also initiated for these future electronic applications. In particular, organic-based flexible memories have merits such as a simple, low-temperature, and low-cost manufacturing process. Several fabrication results of organic resistive memory devices on flexible substrates have been reported.

In addition, with growing demand for high-density digital information storage, NAND Flash memory density has been increased dramatically for the past couple of decades. On the other hand, device dimension scaling to increase memory density is expected to be more and more difficult in a bit-cost scalable manner due to various physical and electrical limitations. As a solution to the problems, NAND Flash memories having stacked layers are under developing extensions. In 3-D memories, cost can be reduced by building multiple stacked cells in vertical direction without device size scaling. As a breakthrough for the scaling limitations, various 3-D stacked memory architectures are under development and expecting the huge market of 3-D memories in the near future. With lots of expectation, future-generation memories have potential to replace most of the existing memory technologies.

Currently, there have been increasing demands on reducing the feature size in microelectronic products and more interest in the development of Flash memory devices to meet the growing worldwide demand. A conventional FG memory device must have a tunnel oxide layer thickness of 8 nm to prevent charge loss and to make 10 years' data retention certain. This necessity will limit scalability for Flash memory devices. Thus, in order to meet technology scaling in the field of memory and data storage devices, mainstream transistor-based Flash technologies will be developed gradually to incorporate material and structural innovations. Dielectric scaling in nonvolatile memories has been reached near to the point where new approaches will be required to meet the scaling requirements while simultaneously meeting the reliability and performance requirements for future products. High-dielectric-constant materials are being explored as possible candidates to replace both the traditional SiO2 and oxide/nitride/oxide (ONO) films used in Flash memory cells. Flash cell scaling has been demonstrated to be really possible and to be able to follow Moore's law down to the 90-nm technology generations. The technology development and the consolidated know-how are expected to sustain the scaling trend down to the 50-nm technology node and below as forecasted by the International Technology Roadmap for Semiconductors (ITRS) in Figure 2, which indicates that the silicon MOSFET was already in the nanoscale.

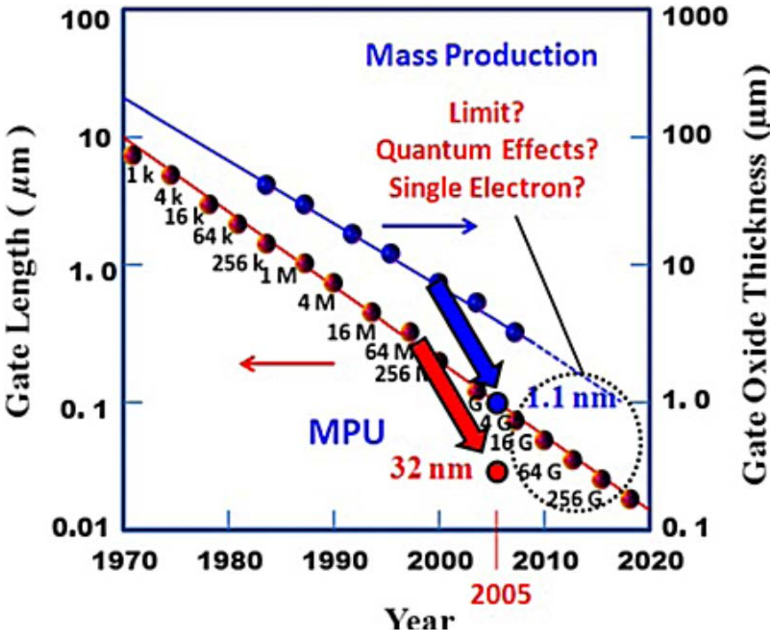


Fig. 2. The trend of MOSFET scaling from ITRS. Reproduced from ITRS Corp.

The minimum feature size of an individual CMOSFET has shrunk to 15 nm with an equivalent gate oxide thickness (EOT) of 0.8 nm in 2001. However, semiconductor Flash memory scaling is far behind CMOS logic device scaling. For example, the EOT of the gate stack in semiconductor Flash memory is still more than 10 nm. Moreover, semiconductor Flash memory still requires operation voltages of more than 10 V, which is still far from the operation voltage of CMOS logic devices. It is important to scale the EOT of the gate stack to achieve a small memory cell size and also prolong battery life. Another limitation of FG technology is that tunnel oxide scaling is limited by stress-induced leakage current (SILC) related to charge transfer problem as indicated in Figure 3.

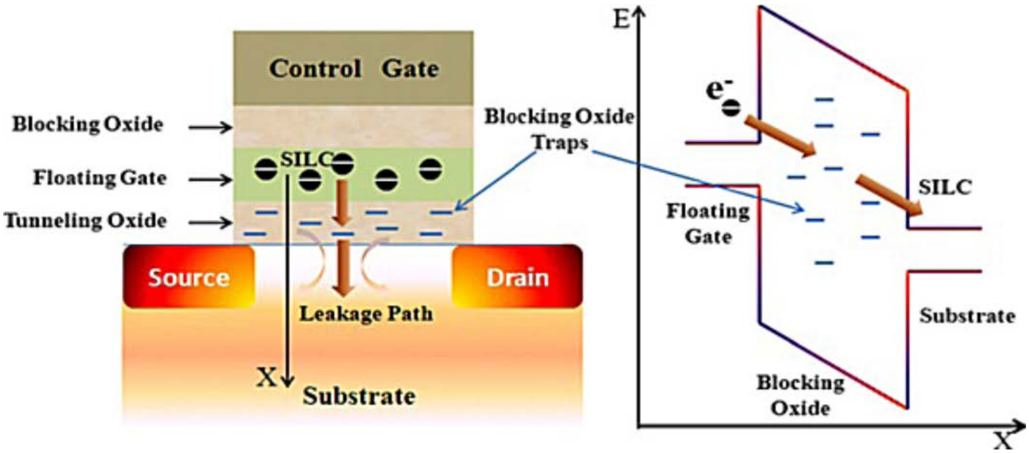


Fig. 3. Schematic plots of a Flash memory cell and the degradation of its tunnel oxide. The degradation leads to the formation of percolation paths responsible for the FG charge loss, hence the loss of the stored information.

The presence of traps in the energy barrier yields is a basic building block of Flash memory, which is based on FG thin-film storage (TFS) memories that have been developed with the addition of an erase gate configuration. The conventional FG memory (Figure 4a) consists of a MOSFET configuration that is modified to include polysilicon as a charge storage layer surrounded by an insulated inner gate (floating gate) and an external gate (control gate). This what makes Flash memory nonvolatile and all floating gate memories to have the same generic cell structure. Charge is transferred to or from the floating gate through a thin (8 to 10 nm) oxide.

Because the floating gate is electrically isolated by the oxide layer, any electrons placed on it are trapped there. Flash memory works by adding (charging) or removing (discharging) electrons to and from a floating gate. A bit's 0 or 1 state depends upon whether or not the floating gate is charged or discharged. When electrons are present on the floating gate, current cannot flow through the transistor and the bit state is ‘0’. This is the normal state for a floating gate. When electrons are removed from the floating gate, current is allowed to flow and the bit state is ‘1’. The FG memory has achieved high density, good program/erase speed, good reliability, and low operating voltage and promotes endurance for Flash memory application.

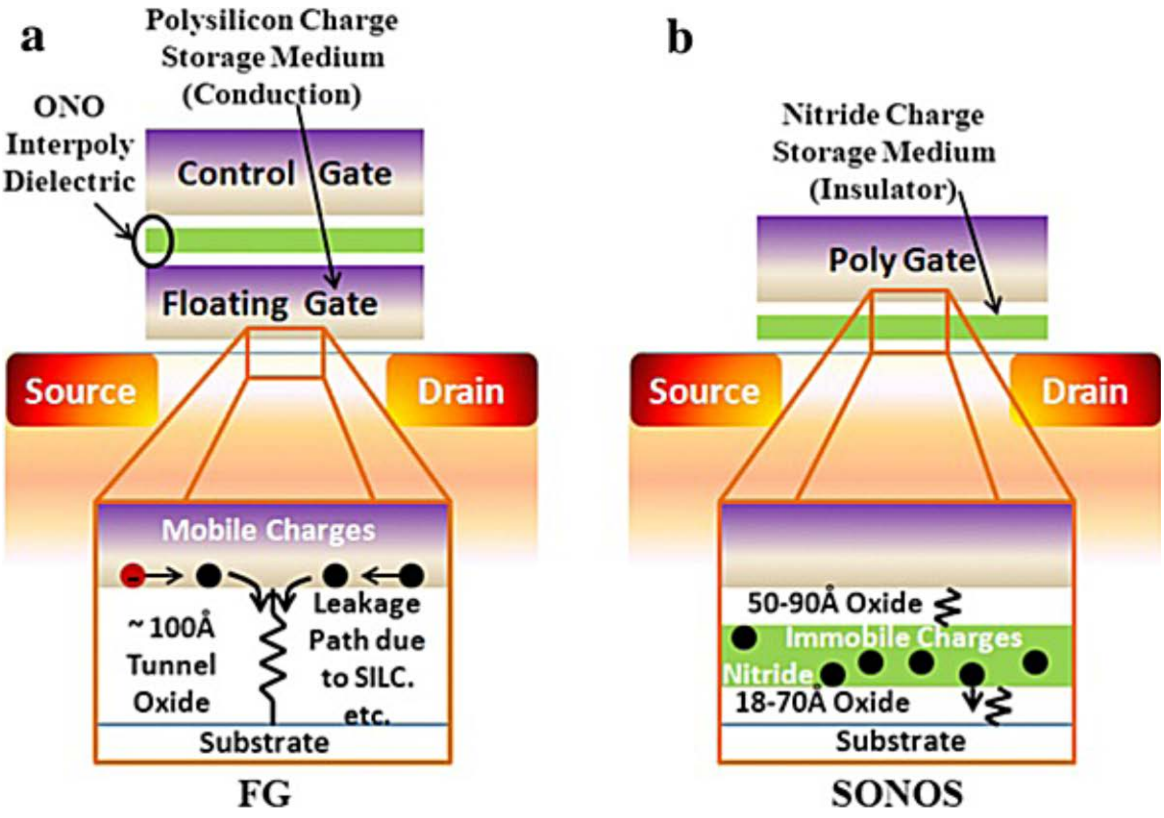


Fig. 4. Schematics of the conventional FG memory and SONOS. Schematics of (a) floating gate and thin-film storage-based embedded nonvolatile memory bit cells, depending on the charge stored inside the gate dielectric of a MOSFET, and (b) the nitride traps (SONOS), embedded into the gate oxide of a MOSFET.

In order to solve the scaling issue of the FG memory, the SONOS memory has been proposed as a Flash technology since the 1980s. The acronym SONOS is derived from the structure of the device as shown in Figure 4b. The SONOS device is basically a MOSFET, where the gate has been replaced by an ONO dielectric. The SONOS memory has a better charge retention than the FG memory when the FG bit cell's tunneling oxide layer is below 10 nm. Moreover, the SONOS memory exhibits many advantages, e.g., easy to fabricate, high program/erase (P/E) speed, low programming voltage and power consumption, and better potential for scalability below the 70-nm node, according to the ITRS. The charge, holes or electrons, are injected into the nitride layer using direct tunneling through the tunnel oxide layer.

The nitride layer is electrically isolated from the surrounding transistor, although charges stored on the nitride directly affect the conductivity of the underlying transistor channel. Since the SONOS memory possesses spatially isolated deep-level traps, a single defect in the tunneling oxide will not cause discharge of the memory cell. The thickness of the top oxide is important to prevent the Fowler-Nordheim tunneling of electrons from the gate during erase. When the polysilicon control gate is biased positively, electrons from the transistor source and drain regions tunnel through the oxide layer and get trapped in the silicon nitride. This results in an energy barrier between the drain and the source, raising the threshold voltage Vth (the gate-source voltage necessary for current to flow through the transistor).

Moreover, the nitride layer is electrically isolated from the surrounding transistor, although charges stored on the nitride directly affect the conductivity of the underlying transistor channel. The oxide/nitride sandwich typically consists of a 2-nm-thick oxide lower layer, a 5-nm-thick silicon nitride middle layer, and a 5- to 10-nm-thick oxide upper layer. However, SONOS-type Flash memories have several drawbacks such as shallow trap energy level, erase saturation, and vertical stored charge migration. The programming speed and operating voltage problems can be solved by reducing the tunnel oxide thickness. At low tunnel oxide thickness, the issues that impact SONOS-type memories include erase saturation and vertical charge migration, which seriously degrade the retention capability of the memory. Thus, many concerns still remain for the SONOS type of memories, which will be discussed in the next section.

Scaling demands very thin gate insulators in order to keep short channel effects and control the shrinkage of the device size and maximize the performance. When the tunneling oxide thickness is below 10 nm, the storage charge in the FG is easy to leak due to a defect in the tunneling oxide formed by repeated write/erase cycles or direct tunneling current. The tunneling gate oxide thickness in a conventional Flash memory cannot be scaled down to sub-7 nm because of charge retention. The SONOS Flash memory can relieve the problem but still has a relatively thick gate dielectric thickness of about 7 nm. Therefore, conventional SONOS Flash memory also has a scaling-down problem. Many studies have shown that the charge retention characteristics in scaled SONOS nonvolatile memory devices with a low gate oxide thickness and at high temperature are problematic with shallow-level traps. For the conventional SONOS memory, erase saturation and vertical stored charge migration are the two major drawbacks; the most challenging tasks are how to maintain an acceptable charge capability of the discrete storage nodes and how to fabricate nanocrystals with constant size, high density, and uniform distributions. When the trap energy level is shallow, erase saturation and vertical migration occur and the electron charge decay rate increases due to low tunnel oxide thickness, issues that impact SONOS-type memories as shown in Figure 5.

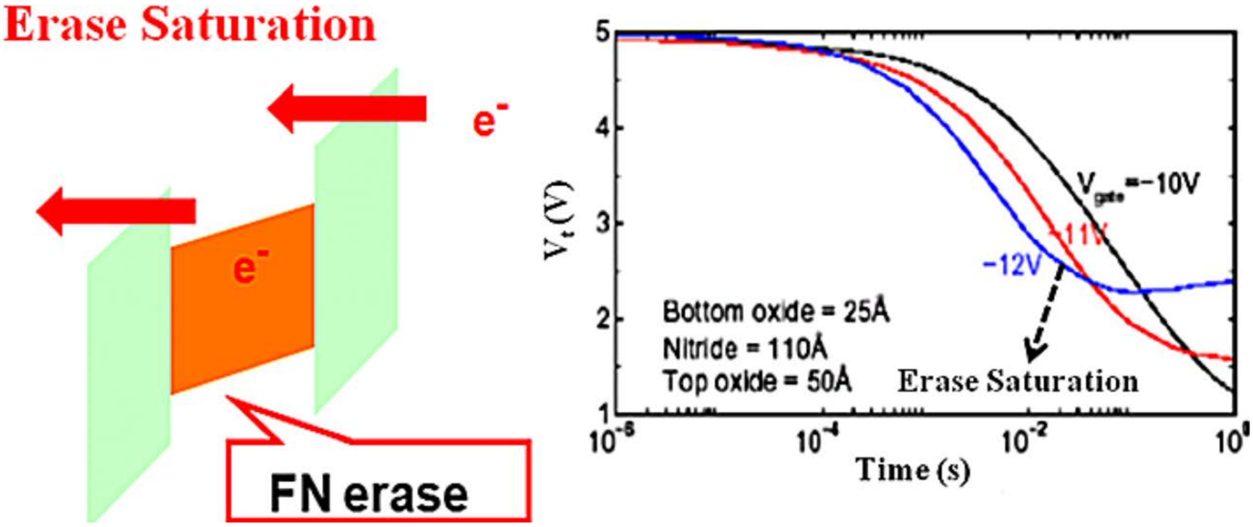


Fig. 5. Fowler-Nordheim (FN) tunneling of electrons from the gate during erase and erase saturation in SONOS nonvolatile memory.This indicates the reduced memory window as the erase voltage is increased.

This erase saturation makes SONOS erase less as the erase voltage or the tunnel oxide thickness is increased. Since the SONOS memory uses silicon nitride as a charge trapping layer, the electrons in the Si sub-conduction band will tunnel through the tunneling oxide and a portion of the nitride, and this consequently degrades the program speed.

Besides this, the conduction band offset of nitride is only 1.05 eV and back-tunneling of the trapped electron may also occur. Although applying a very high electric field may accelerate the de-trapping rate, the gate electron injection current exceeds the de-trapping but resulting in practically an increase in charge and no erasing. Using an ultra-thin (<2 nm) tunnel oxide offers an efficient charge direct tunneling erase and opens a memory window. However, the direct tunneling cannot be turned off at a low electric field, leading to poor retention and read disturb.

Thus, the SONOS memory cannot be used for NAND Flash without further innovation of new memory technologies. The main reason for the growth of emerging NVM technologies is that scaling has now become a serious issue for the memory industry. Not only are many of these new technologies inherently more scalable, but also they seem well suited to the next generation of mobile computing and communications that will demand high-capacity memories capable of storing and rapidly accessing video and a large database without overburdening battery power sources.

Many alternate device structures are proposed to hopefully circumvent these scaling challenges and to improve the device performance. In an effort to continue Moore's law and overcome the ultimate limitations of MOS-based memory devices, other storage concepts have been proposed in search of the ‘unified memory’.

The ideal memory device or the so-called ‘unified memory’ would satisfy simultaneously three requirements: high speed, high density, and nonvolatility. At the present time, such an ideal memory has not been developed. FGNVSM has high density and nonvolatility, but its P/E speed is low. DRAM has high speed (approximately 10 ns) and relatively high density, but it is volatile. SRAM has very high speed (approximately 5 ns), but it suffers from very low density and volatility. Many nonvolatile memory devices have been proposed on the basis of changing charge storage materials and new device concepts for the ‘unified memory’. These structures will be considered in the next sections. In light of such issues, emerging memory solutions seem to be a key technology.

Recent studies have revealed that there is a close correlation among existing and emerging memory technologies in view of scalability. The scaling trend of memory transition leads to smaller and smaller memory devices, which have been routinely observed. To further support this assertion, another set of current progress in memory technology is described to the increasing importance of memory to users' experience and the importance of memory to system performance.

There are many emerging memory technologies which are trying to replace existing memory technologies in the market. These new memory devices such as RRAM, PCM, and STT-RAM have read/write/retention/endurance characteristics different from those of conventional SRAM, DRAM, and Flash. But the ideal characteristics of new emerging memory technologies have to be meeting the performance of SRAM and the density of NAND Flash in terms of stability, scalability, and switching speed. Thus, going beyond the traditional bistable memory, the possibilities of multilevel, high-performance memory devices suitable for market must be explored.

Currently, there are several technologies that show some promise; some of these new emerging technologies are MRAM, FeRAM, PCM, STT-RAM, nano-random-access memory (NRAM), racetrack memory, RRAM and memristor, molecular memory, and many others. Each of these memory technologies will be briefly outlined and discussed in the following sections. In view of the commercial production, currently, MRAM, FeRAM, and PCM are in commercial production but still remain limited to niche applications relative to DRAM and NAND Flash. There is a prospect that among the emerging memory technologies, MRAM, STT-RAM, and RRAM are the most promising ones, but they are still many years away from competing for industry adoption. It is necessary for any new technology to be able to deliver most for industry adoption. For industry adoption on a mass scale, some parameters must be matched with existing memory technologies. In consideration of new technology for industry application, the scalability of the technology, speed of the device, power consumption to be better than existing memories, endurance, densities, better than existing technologies and most importantly the cost; if the emerging technology can only run one or two of these attributes, then, at most desirable, it is likely to be resigned to niche applications.

MRAM or magnetic RAM is a nonvolatile RAM technology under development since the 1990s. RRAM methods of storing data bits use magnetic charges instead of the electrical charges used by DRAM and SRAM technologies. MRAM, first developed by IBM in the 1970s [85], is expected to replace DRAM as the memory standard in electronics. MRAM is basically based on memory cells having two magnetic storage elements, one with a fixed magnetic polarity and another with a switchable polarity. These magnetic elements are positioned on top of each other but separated by a thin insulating tunnel barrier as shown in the cell structure in Figure 6.

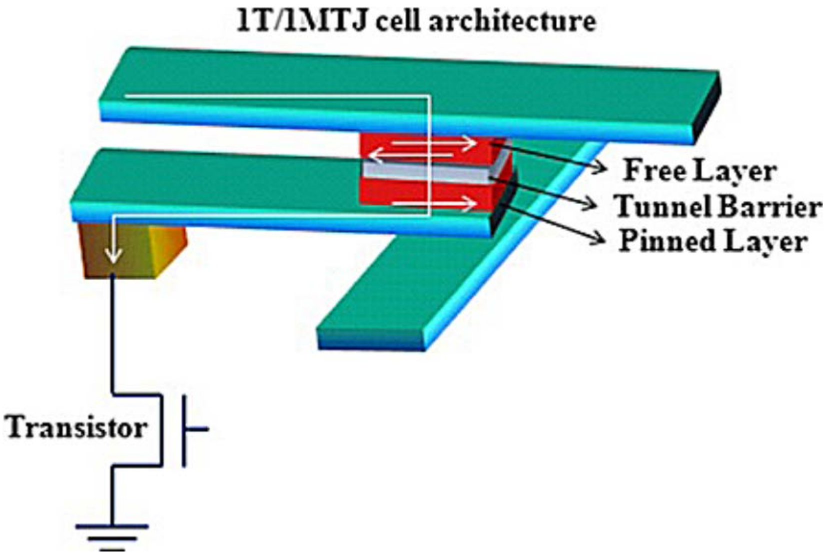


Fig. 6. Basic MRAM cell structure.

Moreover, scientists define a metal as magnetoresistive if it shows a slight change in electrical resistance when placed in a magnetic field. By combining the high speed of static RAM and the high density of DRAM, proponents say that MRAM could be used to significantly improve electronic products by storing greater amounts of data, enabling it to be accessed faster while consuming less battery power than existing electronic memories.

Technically, it works with the state of the cell, which is sensed by measuring the electrical resistance while passing a current through the cell. Because of the magnetic tunnel effect, if both magnetic moments are parallel to each other, then the electrons will be able to tunnel and the cell is in the low resistance ‘ON’ state. However, if the magnetic moments are antiparallel, the cell resistance will be high. The memory characteristics of MRAM of writing and erasing are fulfilled by passing a current through the write line to induce a magnetic field across the cell. MRAM has been slowly getting off the ground but has now entered the market and will become increasingly available for mass production in the couple of years and beyond.

Currently, it has reached some level of commercial success in niche applications. Various companies such as Samsung, IBM, Hitachi and Toshiba, and TSMC are actively developing variant technologies of MRAM chips. In view of power consumption and speed, MRAM competes favorably than other existing memories such as DRAM and Flash, with an access time of a few nanoseconds. Although it has some limitation during the ‘write’ operation, the smaller cell size could be limited by the spread of the magnetic field into neighboring cells and need an amendment to compete completely as a universal memory. The price of MRAM is also another issue and considered a limiting factor, with prices far in excess of all the currently established memories at approximately £2 to £3 ($3 to $5) per megabyte. According to this price level, MRAM is in excess of 1,000 times the price of Flash memory and over 10,000 times the price of hard disk drives. It is expected that of the next-generation memory technologies, MRAM, in the future, will have the biggest market, followed by FeRAM, PCRAM, and memristors.

STT-MRAM is a magnetic memory technology that exerts the base platform established by an existing memory called MRAM to enable a scalable nonvolatile memory solution for advanced process nodes. It is a new kind of magnetic RAM with the following features: fast read and write times, small cell sizes, potentially even smaller, and compatibility with existing DRAM and SRAM. As we have discussed in the previous section, MRAM stores data according to the magnetization direction of each bit and the nanoscopic magnetic fields set the bits in conventional MRAM. On the other hand, STT-MRAM uses spin-polarized currents, enabling smaller and less energy-consuming bits. The basic cell structure of STTRAM is depicted in Figure 7.

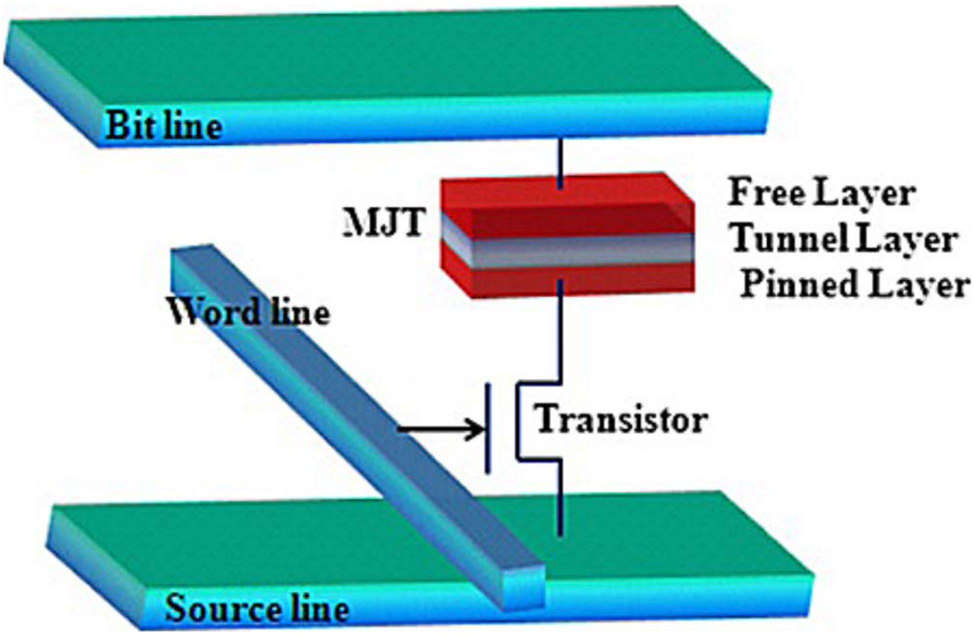


Fig. 7. Basic STT-RAM cell structure.

In addition, STT-RAM writing is a technology in which an electric current is polarized by aligning the spin direction of the electrons flowing through a magnetic tunnel junction (MTJ) element. Data writing is performed by using the spinpolarized current to change the magnetic orientation of the information storage layer in the MTJ element. The resultant resistance difference of the MTJ element is used for information readout. STT-RAM is a more appropriate technology for future MRAM produced using ultra-fine processes and can be efficiently embedded in subsequent generations of such semiconductor devices as FPGAs, microprocessors, microcontrollers, and SoC. A special bonus for embedded designers is the fact that the internal voltage STT-RAM requires is only 1.2 V.

The difference between STT-MRAM and a conventional MRAM is only in the writing operation mechanism; the read system is the same. The memory cell of STT-MRAM is composed of a transistor, an MTJ, a word line (WL), a bit line (BL), and a source line (SL). Currently, STT-RAM is being developed in companies including Everspin, Grandis, Hynix, IBM, Samsung, TDK, and Toshiba. However, for STT-RAM to be adopted as a universal mainstream semiconductor memory, some key challenges should be resolved: the simultaneous achievement of low switching current and high thermal stability. It must be dense (approximately 10 F^2), fast (below 10 ns of read and write speeds), and operating at low power.

FeRAM is a nonvolatile RAM that combines the fast read and write access of DRAM cells, consisting of a capacitor and transistor structure as shown in Figure 8.

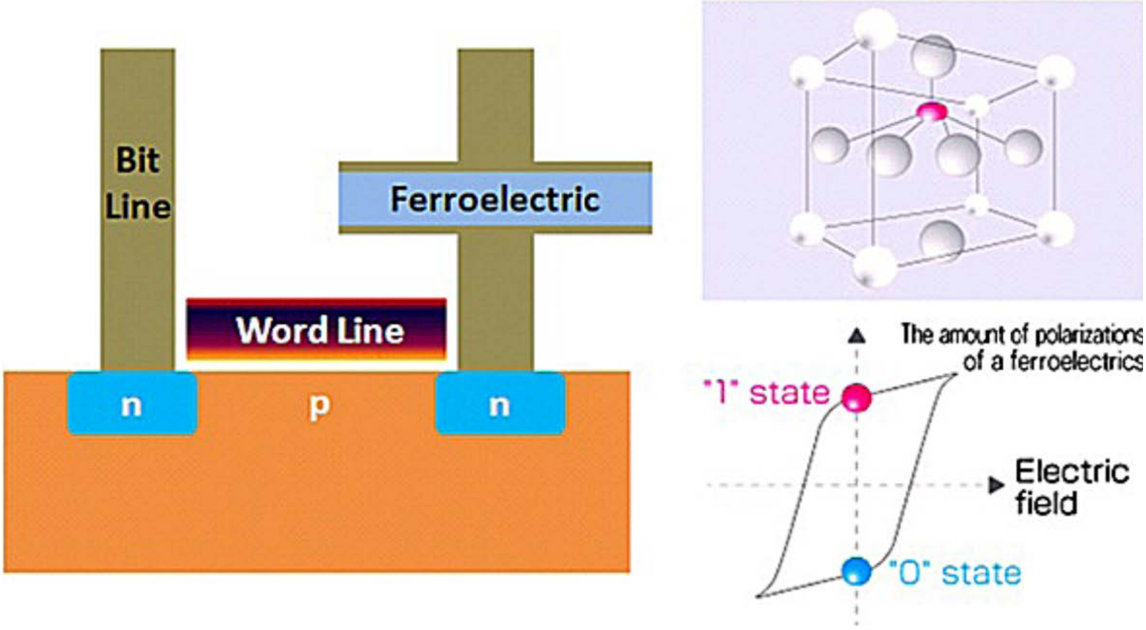


Fig. 8. Basic structure of a FeRAM cell. The crystal structure of a ferroelectric and an electric polarization-electric field hysteresis curve are also shown.

The cell is then accessed via the transistor, which enables the ferroelectric state of the capacitor dielectric to be sensed. In spite of its name, FeRAM does not contain iron. The polarization properties of a ferroelectric substance are used as a memory device. Today's FeRAM uses lead zirconate titanate (PZT); other materials are being considered.

The main developer of FeRAM is Ramtron International. FeRAM is the most common kind of personal computer memory with the ability to retain data when power is turned off as do other nonvolatile memory devices such as ROM and Flash memory. In a DRAM cell, the data periodically need refreshing due to the discharging of the capacitor, whereas FeRAM maintains the data without any external power supply. It achieves this by using a ferroelectric material in the place of a conventional dielectric material between the plates of the capacitor. When an electric field is applied across dielectric or ferroelectric materials, it will polarize, and while that field is removed, it will depolarize. But the ferroelectric material exhibits hysteresis in a plot of polarization versus electric field, and it will retain its polarization. One disadvantage of FeRAM is that has a destructive read cycle. The read method involves writing a bit to each cell; if the state of the cell changes, then a small current pulse is detected by indicating that the cell was in the OFF state. However, it is a fast memory that can endure a high number of cycles (e.g., 10^{14}), meaning that the requirement for a write cycle for every read cycle will not result in short product lives with a very low power requirement. It is expected to have many applications in small consumer devices such as personal digital assistants (PDAs), handheld phones, power meters, and smart cards, and in security systems. FeRAM is faster than Flash memory.

It is also expected to replace EEPROM and SRAM for some applications and to become a key component in future wireless products. Even after FeRAM has achieved a level of commercial success, with the first devices released in 1993, current FeRAM chips offer performance that is either comparable to or exceeding current Flash memories, but still slower than DRAM.

PCRAM, also known as PCM, perfect RAM (PRAM), OUM, and chalcogenide RAM (CRAM), is a type of nonvolatile RAM based on a class of material called chalcogenide glasses that can exist in two different phase states (e.g., crystalline and amorphous). The basic PCRAM cell structure is depicted in Figure 9.

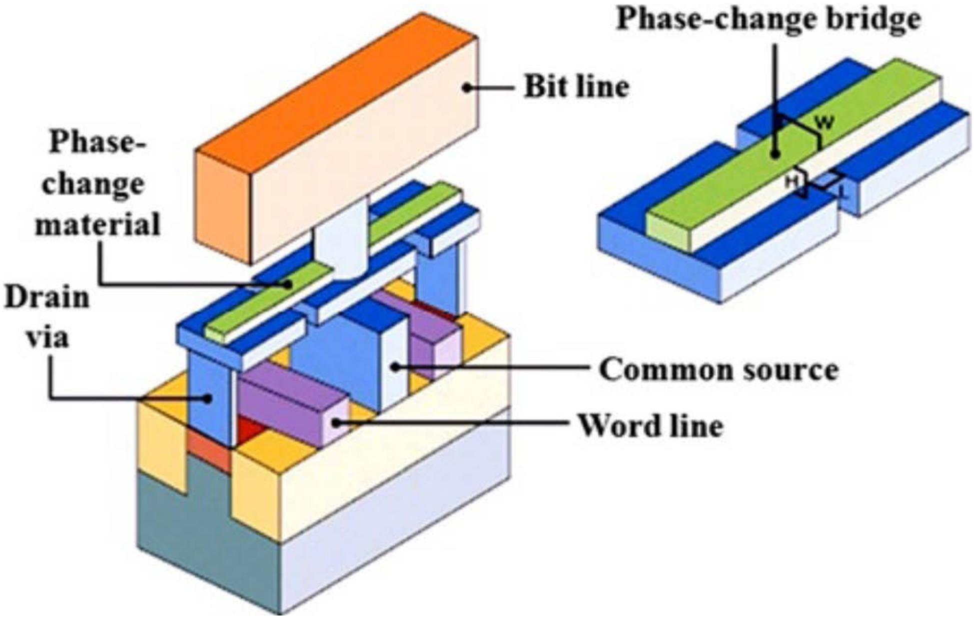


Fig. 9. Basic PCRAM cell structure. Reproduced from IBM-Macronix-Qimonda.

Most phase-change materials contain at least one element from group 6 of the periodic table, and the choice of available materials can be further widened by doping these materials. In particular, the most promising are the GeSbTe alloys which follow a pseudobinary composition (between GeTe and Sb_2Te_3), referred to as GST. These materials are in fact commonly used as the data layer in rewritable compact disks and digital versatile disks (CD-RW and DVD-RW) where the change in optical properties is exploited to store data. The structure of the material can change rapidly back and forth between amorphous and crystalline on a microscopic scale. The material has low electrical resistance in the crystalline or ordered phase and high electrical resistance in the amorphous or disordered phase. This allows electrical currents to be switched ON and OFF, representing digital high and low states.

This process has been demonstrated to be on the order of a few tens of nanoseconds, which potentially makes it compatible with Flash for the read operation, but several orders of magnitude faster for the write cycle.

This makes it possible for PCM to function many times faster than conventional Flash memory while using less power. In addition, PCM technology has the potential to provide inexpensive, high-speed, high-density, high-volume nonvolatile storage on an unprecedented scale. The physical structure is three-dimensional, maximizing the number of transistors that can exist in a chip of fixed size. PCM is sometimes called perfect RAM because data can be overwritten without having to erase it first. Possible problems facing PCRAM concern the high current density needed to erase the memory; however, as cell sizes decrease, the current needed will also decrease. PCM chips are expected to last several times as long as currently available Flash memory chips and may prove cheaper for mass production. Working prototypes of PCM chips have been tested by IBM, Infineon, Samsung, Macronix, and others. Also, the production of PCM has been announced recently by both collaborations between Intel and STMicroelectronics as well as with Samsung .

Before going to other emerging memories, we herein provide a comparison among MRAM, FeRAM, and PCM. Relatively mature, new-material memories such as MRAM, STT-RAM, FeRAM, and PCM can offer a variety of features that have potential to be the candidates for next-generation nonvolatile memory devices.

Brand-new concepts such as RRAM, molecular, organic/polymer, and other nanowire-based memory technologies have also been proposed. These are discussed in detail in the following section.

RAM is a disruptive technology that can revolutionize the performance of products in many areas, from consumer electronics and personal computers to automotive, medical, military, and space. Among all the current memory technologies, RRAM is attracting much attention since it is compatible with the conventional semiconductor processes.

Memristor-based RRAM is one of the most promising emerging memory technologies and has the potential of being a universal memory technology. It offers the potential for a cheap, simple memory that could compete across the whole spectrum of digital memories, from low-cost, low-performance applications up to universal memories capable of replacing all current market-leading technologies, such as hard disk drives, random-access memories, and Flash memories. RRAM is a simple, two-terminal metal-insulator-metal (MIM) bistable device as shown in the basic configuration in Figure 10.

It can exist in two distinct conductivity states, with each state being induced by applying different voltages across the device terminals. RRAM uses materials that can be switched between two or more distinct resistance states. Many companies are investing metal oxide nanolayers switched by voltage pulses.

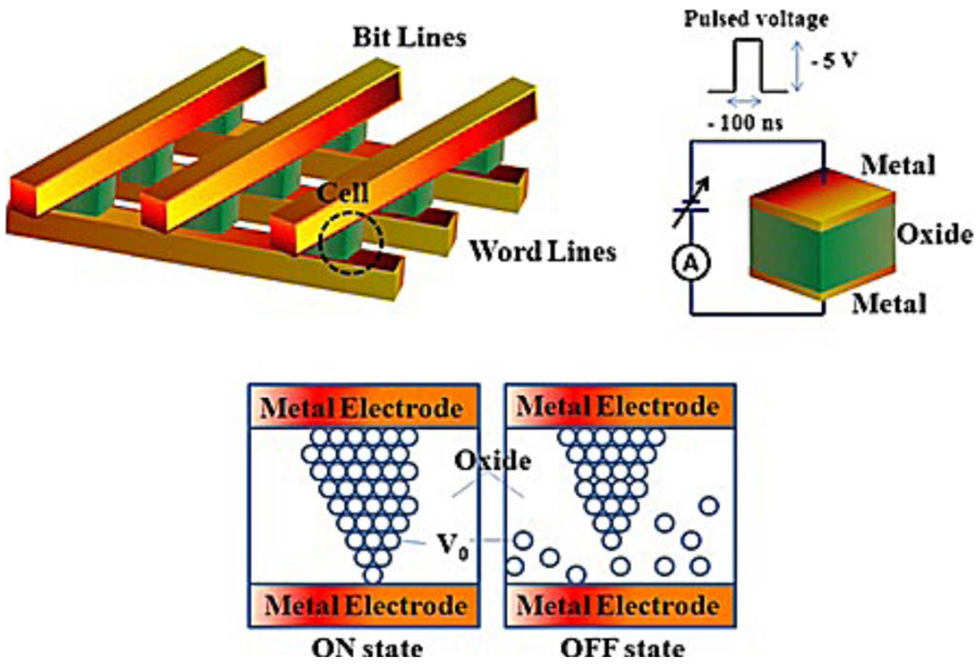


Fig. 10. Basic RRAM cell structure. A schematic diagram of the mechanism of the resistive switching in a metal/oxide/metal-structured memory cell is also shown.

Researchers generally think that the pulses' electric fields produce conducting filaments through the insulating oxide. HP Labs plans to release prototype chips this year based on ‘memristors’ in which migrating oxygen atoms change resistance.

As a result of experimental research has been also defined that among all the technology candidates, RRAM is considered to be the most promising as it operates faster than PCRAM and it has a simpler and smaller cell structure than magnetic memories (e.g., MRAM or STT-RAM). In contrast to a conventional MOS-accessed memory cell, a memristor-based RRAM has the potential of forming a cross-point structure without using access devices, achieving an ultra-high density. This device is based on the bistable resistance state found for almost any oxide material, including NiO, ZrO2, HfO2, SrZrO3, and BaTiO3. Currently, Samsung and IBM are actively investigating RRAM. Kamiya et al. have revealed by a theoretical mechanism that RRAM shows filamentary-type resistive switching, where the oxygen vacancy is considered to form conductive filaments in the resistive material as shown in Figure 10.

The formation and disruption of these filaments are thus the mechanisms responsible for the ON-OFF switching in RRAM devices. The key issue is therefore, to reveal electronic roles in the formation and disruption of the vacancy filaments. RRAM can be switched between the low resistance state (LRS) and the high resistance state (HRS) of the resistive material by applying voltages to the electrodes.

Researchers have explained that during the SET process, the current level increases from HRS to LRS as the voltage increases from 0 V to the critical point which is called the set voltage (V set), while the current level abruptly decreases from LRS to HRS at the reset voltage (V reset) under the RESET process. The SET and RESET processes are repeatedly carried out by sweeping the gate voltage with the binary states LRS and HRS. Wang and Tseng and Lin et al. have indicated that the interface plays an important role in enhancing the performances of RRAM. Recently, Goux et al. have explained that using a stacked RRAM structure has been shown to be one of the most promising methods to improve the memory characteristics. Although being a most promising memory element, critical issues for the future development of RRAM devices are reliable, such as data retention and memory endurance. A data retention time of over 10 years can be extrapolated from retention characteristics measured at high temperatures and a memory endurance of over 10^6 cycles. Therefore, a statistical study of reliability, availability, and maintainability is essential for the future development of RRAM.

Throughout the last few years, polymers have found growing interest as a result of the rise of a new class of nonvolatile memories. In a polymer memory, a layer consists of molecules and/or nanoparticles in an organic polymer matrix is sandwiched between an array of top and bottom electrodes as illustrated in Figure 11.

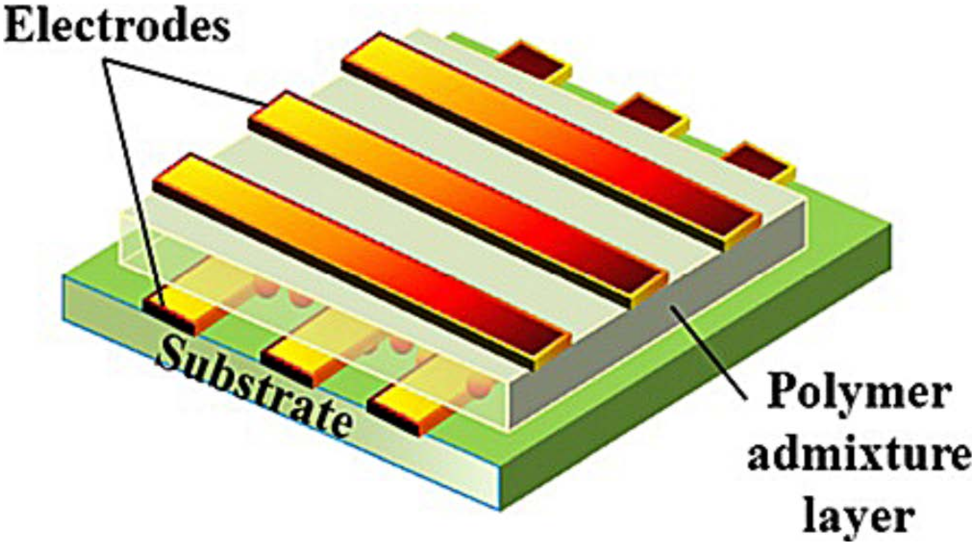


Fig. 11. Structure of a polymer memory device.

Moreover, polymer memory has the advantage of a simple fabrication process and good controllability of materials. Polymer memory could be called digital memory with the latest technology. It is not possible for a siliconbased memory to be established in less space, but it is possible for polymer memory. Ling et al. explained that polymer materials have simplicity in structure, free read and write capability, better scalability, 3-D stacking ability, low-cost potential, and huge capacity of data storage. They revealed that a polymer memory stores information in a manner that is entirely different from that of silicon-based memory devices. Rather than encoding ‘0’ and ‘1’ from the number of charges stored in a cell, a polymer memory stores data on the basis of high and low conductivity while responding to an applied voltage. Among the large number of emerging memory technologies, polymer memory is the leading technology. It is mainly because of its expansion capability in 3-D space since most polymers are organic materials consisting of long chains of single molecules.

Prior to polymer memory fabrication, deposition of an organic layer is usually done by the sol-gel spin coating technique. All the other necessary constituent materials are dissolved in a solvent which is then spin-coated over a substrate. When the solvent is evaporated, a thin film of material with 10- to 100-nm thickness is successfully deposited at bottom electrodes. Top electrodes are deposited as the final step. The conductivity of the organic layer is then changed by applying a voltage across the memory cell, allowing bits of data to be stored in the polymer memory cell. When the polymer memory cell becomes electrically conductive, the electrons are introduced and removed. Even the polymer is considered as a ‘smart’ material to the extent that functionality is built into the material itself of switchability and charge store. This will open up tremendous opportunities in the electronics world, where tailor-made memory materials represent an unknown territory. The nonvolatileness and other features are inbuilt at the molecular level and offers very high advantages in terms of cost. But turning polymer memory into a commercial product would not be easy. Memory technologies compete not only on storage capacity but on speed, energy consumption, and reliability. ‘The difficulty is in meeting all the requirements of current silicon memory chips,’ says Thomas, the Director of Physical Sciences at IBM's Watson Research Center in Yorktown Heights, NY. They are likely to be limited to niche applications.

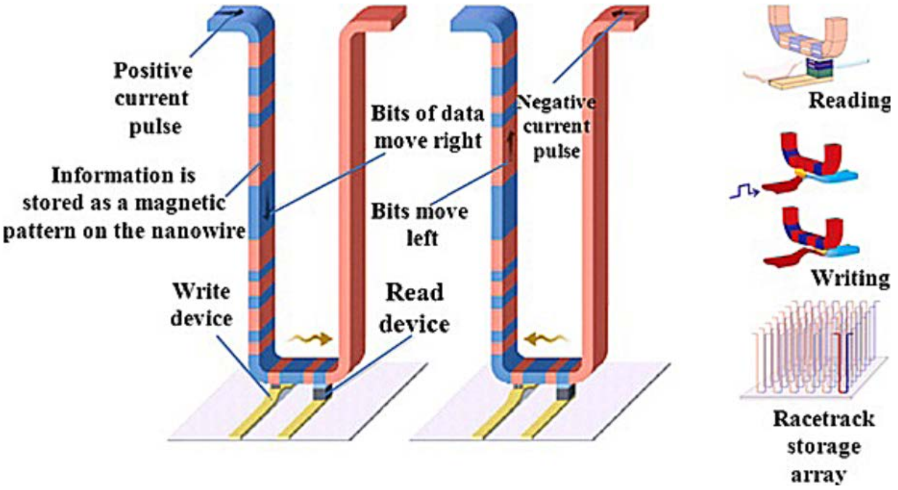


Fig. 12. Racetrack memory diagram showing an array of U-shaped magnetic nanowires. The nanowires are arranged vertically like trees in a forest and a pair of tiny devices that read and write the data. Adopted from IBM.

In a racetrack memory, information is stored on a Ushaped nanowire as a pattern of magnetic regions with different polarities. The U-shaped magnetic nanowire is an array of keys, which are arranged vertically like trees in a forest as shown in Figure 12.

Achieving capacities comparable to vertical RM or hard drives would requirestacks of these arrays. The nanowires have regions with different magnetic polarities, and the boundaries between the regions represent 1 or 0 s, depending on the polarities of the regions on either side. The magnetic information itself is then pushed along the wire, past the write and read heads by applying voltage pulses to the wire ends.

The magnetic pattern to speed along the nanowire, while applying a spin-polarized current, causes the data to be moved in either direction, depending on the direction of the current. A separate nanowire perpendicular to the U-shaped 'racetrack' writes data by changing the polarity of the magnetic regions. A second device at the base of the track reads the data.

Data can be written and read in less than a nanosecond. A racetrack memory using hundreds of millions of nanowires would have the potential to store vast amounts of data. In this way, the memory requires no mechanical moving of parts and it has a greater reliability and higher performance than HDDs, with theoretical nanosecond operating speeds. For a device configuration where data storage wires are fabricated in rows on the substrate, conventional manufacturing techniques are adequate. However, for the maximum possible memory density, the storage wires are proposed to be configured rising from the substrate in a 'U' shape, giving rise to a 3-D forest of nanowires.

While this layout does allow high data storage densities, it also has the disadvantage of complex fabrication methods, with so far, only 3-bit operation of the devices demonstrated. As the access time of the data is also dependent on the position of the data on the wire, these would also be performance losses if long wires are used to increase the storage density further. The speed of operation of the devices has also been an issue during development, with much slower movement of the magnetic domains than originally predicted. This has been attributed to crystal imperfections in the permalloy wire, which inhibit the movement of the magnetic domains. By eliminating these imperfections, a data movement speed of 110 m/s has been demonstrated.

Researchers are already working hard on several emerging technologies, as discussed in previous sections, to pursue storage-class memories with a more traditional design than that of the racetrack memory, which places the bits in horizontal arrays.

A molecular memory is a nonvolatile data storage memory technology that uses molecular species as the data storage element, rather than, e.g., circuits, magnetics, inorganic materials, or physical shapes. In a molecular memory, a monolayer of molecules is sandwiched between a cross-point array of top and bottom electrodes as shown in Figure 13.

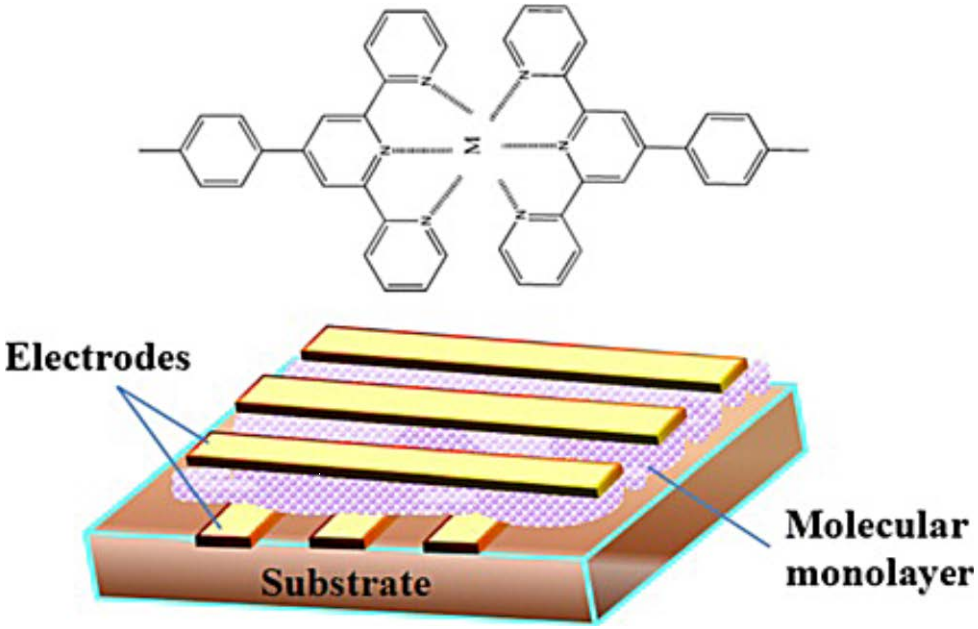


Fig. 13. Cell structure of a molecular memory device.

The molecules are packed in a highly ordered way, with one end of the molecule electrically connected to the bottom electrode and the other end of the molecule connected to the top electrode, and this molecular component is described as a molecular switch. Langmuir-Blodgett (LB) deposition is ideally suited for depositing the molecular layer for the fabrication of molecular memory devices. Then, regarding the molecular memory operation, by applying a voltage between the electrodes, the conductivity of the molecules is altered, enabling data to be stored in a nonvolatile way.

This process can then be reversed, and the data can be erased by applying a voltage to the opposite polarity of the memory cell. The increasing demand for nonvolatile electronic memories will grow rapidly in order to keep pace with the requirements for subsystems involved in flight demonstration projects and deep space operations. At the same time, mass, volume, and power must be minimized for mission affordability concerning these requirements; molecular memory could be a very promising candidate to fill this need. Recently, Plafke has revealed clearly via an article that like most experimental technology that sounds so amazing that we want it right now, the molecular memory cell does not provide enough power for a commercial device. This is currently only able to produce a 20% jump in conductivity.

However, the area of molecular switching memory is promising, having eliminated the need for near-absolute zero temperatures and removed some of the constraints of the shape and number of layers of the molecule sheets which intend to convey that two of the biggest barriers are taken away.

Thus, molecular memory requires strong attention to work over such issues and needs immediate amendment to see the possibility of a universal memory in the future. In the last two decades, an increasing interest is observed for electronics-related devices and the search for a universal memory data storage device that combines rapid read and write speeds, high storage density, and nonvolatility is driving the investigation of new materials in the nanostructured form.

As an alternative to the current Flash memory technology, a novel transistor architecture using molecular-scale nanowire memory cells holds the promise of unprecedently compact data storage. The molecular nanowire array (MNW) memory is fundamentally different from other semiconductor memories; information storage is achieved through the channel of a nanowire transistor that is functionalized with redox-active molecules rather than through manipulation of small amounts of charge. It is relatively slow and lacks the random access capability, wherein data that can be randomly read and written at every byte are being actively pursued.

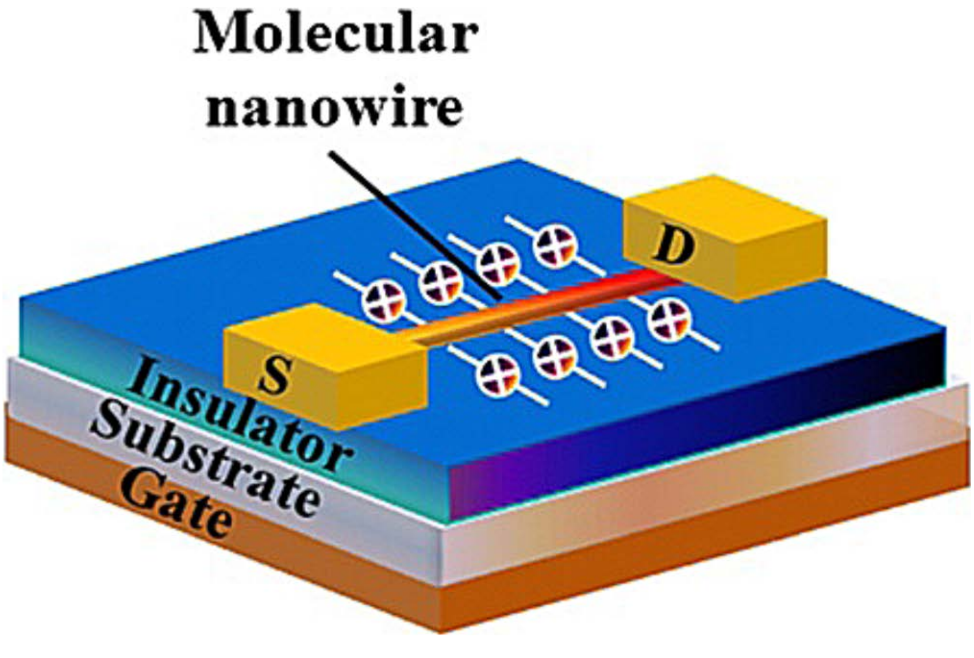


Fig. 14. A MNW memory cell structure.

Figure 14 shows the schematic design of a MNW memory cell. Lieber, and Agarwal and Lieber have revealed that the nanowire-based memory technology is a powerful approach to assemble electronic/photonic devices at ultra-small scales owing to their sub-lithographic size, defect-free single-crystalline structure, and unique geometry.

Nanowires synthesized by chemical or physical processes are nearly perfect single-crystal structures with a small geometry and perfect surface. The channel of a nanowire transistor is functionalized with redox-active molecules. During programming, control of the voltage acting on the substrate is possible to change the oxidation and reduction states of the active molecules. Finally, by measurement of the conductance of the nanowire with the gate bias fixed at 0 V or a small voltage and from the hysteresis, the two states can be defined as a high-conductance ON state and a lowconductance OFF state.

The MNW memory has advantages of low power dissipation, ultra-high density, simple fabrication process, 3-D structure, and multilevel storage, and it functions at the nanoscale with a few electrons but limited by low retention time parameter. Moreover, the deposition of metals onto a monolayer of molecular wires can lead to low device yield, and this problem remains a major challenge. However, mentioning the term emerging class memory, it could be expected that the MNW memory represents an important step towards the creation of molecular computers that are much smaller and could be more powerful than today's silicon-based computers.

Semiconductor memory is essential for information processing as a key part of silicon technology; semiconductor memory has been continuously scaled to achieve a higher density and better performance in accordance with Moore's law. Flash memory may reach fundamental scaling limits, however, because a thick tunneling oxide is required to prevent charge leakage and achieve 10 years' retention. As Flash memory approaches its scaling limit, several alternative strategies have been proposed to extend or replace the current Flash memory technology.

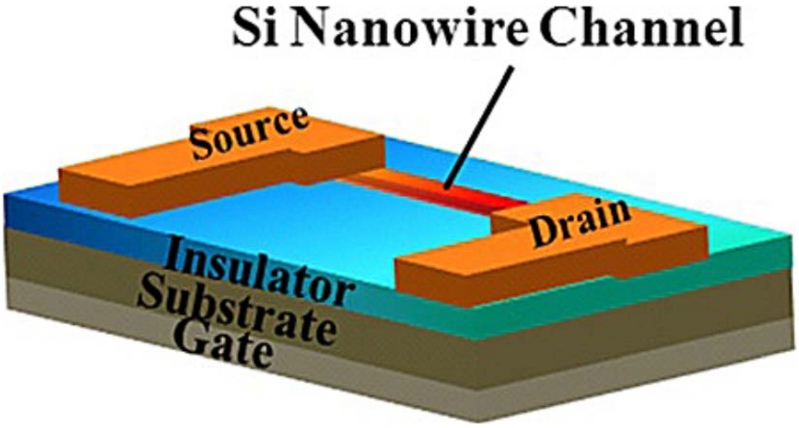


Fig. 15. A bottom-gate FET-based nonvolatile SNW memory device.

These approaches are revolutionary, but major challenges must be overcome to achieve small memory size and aggressive technology design architecture. In addition to the engineering of trapping layers, the device performance can also be improved by using innovative nonplanar channel geometries.

Among the various nanostructure materials, semiconductor nanowire memory (SNW) has induced great scientific interest as possible building blocks for future nanoelectronic circuitry. In a SNW memory device, nanowires are integrated with SONOS technology. The basic schematic design of SNW is depicted in Figure 15. The SNW memory shows high mobility, less power dissipation, and high performance.

Moreover, being 3-D-stacked, the SNW memory enhances cell density and data capacity without relying on advances in process technology. The nanowire-based memory device can store data electrically and is nonvolatile, meaning it retains data when the power is turned off, like the silicon-based Flash memory found in smart phones and memory cards, with minimal increase in chip size. In addition, the SNW device exhibits reliable write/read/erase operations with a large memory window and high on-to-off current ratio, which are highly advantageous for applications in nonvolatile memory. The SNW memory cannot hold data as long as the existing Flash, but it is slower and has fewer rewrite cycles and it could potentially be made smaller and packed together more densely. And its main advantage is that it can be made using simple processes at room temperature, which means that it can be deposited even on top of flexible plastic substrates. The SNW could, for instance, be built into a flexible display and could be packed into smaller spaces inside cell phones, MP3 players, plastic RFID tags, and credit cards.

NRAM is a carbon nanotube (CNT)-based memory, which works on a nanomechanical principle, rather than a change in material properties. NRAM uses carbon nanotubes for the bit cells, and the 0 or 1 is determined by the tube's physical state: up with high resistance, or down and grounded. NRAM is expected to be faster and denser than DRAM and also very scalable, able to handle 5-nm bit cells whenever CMOS fabrication advances to that level. It is also very stable in its 0 or 1 state. Produced by Nantero, these memories consist of the structure shown in Figure 16a with an array of bottom electrodes covered by a thin insulating spacer layer. CNTs are then deposited on the spacer layer, leaving them freestanding above the bottom electrodes.

Unwanted CNTs are removed from the areas around the electrode, with top contacts and interconnects deposited on top of the patterned CNT layer. During the time that the CNTs are freestanding, there is no conduction path between the bottom and top electrodes and hence the memory cell is in the OFF state. However, if a large enough voltage is applied over the cell, the nanotubes are attracted to the bottom electrode where they are held in place by van der Waals forces. Due to the conductive nature of the CNTs, the electrodes are now connected and the cell reads the low conductivity ON state as shown in Figure 16b.

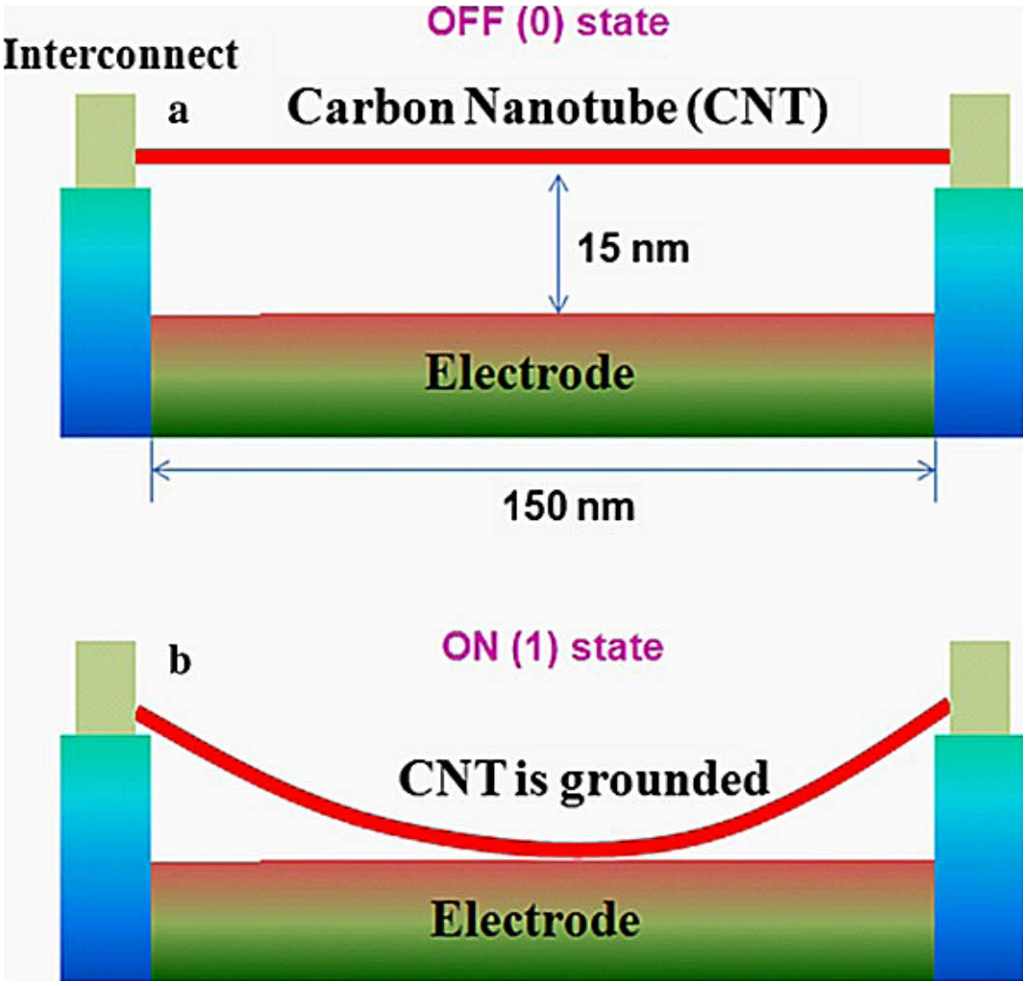


Fig. 16. NRAM structure with (a) OFF state and (b) ON state.

The OFF state can be returned by repelling the nanotubes with the opposite electrode polarity. Nonvolatility is achieved due to the strength of the van der Waals forces overcoming the mechanical strain of the bent nanotubes, hence holding the cell in the ON state. NRAM offers the possibility of a simple cell architecture, which could operate at much higher speeds than the conventional Flash and with low power use. Cui et al. reported CNT memory devices exhibiting an extraordinarily high charge storage stability of more than 12 days at room temperature. However, as NRAM is based on CNTs, it suffers from fabrication problems that are inherent in carbon nanotube-based devices. The issues include the cost and fabrication complexity of producing the CNTs, ensuring uniform dispersions of nanotubes, and difficulties in removing nanotubes from the unwanted positions on the substrate.

In 2002, IBM developed a punch card system known as Millipede, which is a nonvolatile computer memory stored in a thin polymer sheet with nanoscopic holes to provide a simple way to store binary data. It can store hundreds of gigabytes of data per square centimeter. However, the polymer reverts to its pre-punched form over time, losing data in the process.

Millipede storage technology is being pursued as a potential replacement for magnetic recording in hard drives, at the same time reducing the form factor to that of Flash media. The prototype's capacity would enable the storage of 25 DVDs or 25 million pages of text on a postage stamp-sized surface and could enable 10 GB of storage capacity on a cell phone. Millipede uses thousands of tiny sharp points (hence the name) to punch holes in athin plastic film. Each of the 10-nm holes represents a single bit. The pattern of indentations is a digitized version of the data. The layout of the millipede cantilever/tip in contact with the data storage medium is shown in Figure 17.

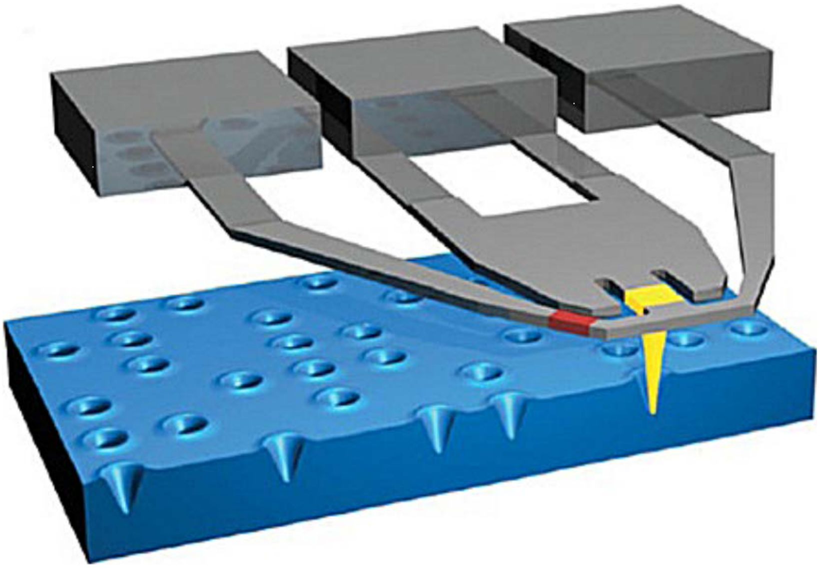


Fig. 17. Schematic layout of the millipede cantilever/tip in contact with the data storage medium.

According to IBM, Millipede can be thought of as a nanotechnology version of the punch card data processing technology developed in the late nineteenth century. However, there are significant differences: Millipede is rewritable, and it may eventually enable storage of over 1.5 GB of data in a space no larger than a single hole in the punch card.

Storage devices based on IBM's technology can be made with existing manufacturing techniques, so they will not be expensive to make. According to P. Vettiger, head of the Millipede project, there is not a single step in fabrication that needs to be invented. Vettiger predicts that a nanostorage device based on IBM's technology could be available as early as 2005. Now, researchers at IBM's Zurich Research Laboratory in Switzerland have clocked the rate of data loss. They have calculated that at 85°C, a temperature often used to assess data retention, it would lose just 10% to 20% of information over a decade, comparable to Flash memory. The use of DNA is well known as a good model for metal NP synthesis due to its affinity to the metal ions.

In recent years, DNA has also been shown to be a promising optical material with the material processing fully compatible with conventional polymer for thin-film optoelectronic applications. Researchers from National Tsing Hua University in Taiwan and the Karlsruhe Institute of Technology in Germany have created a DNA-based memory device, that is, write-once-readmany-times (WORM), that uses ultraviolet (UV) light to encode information.

The device consists of a single biopolymer layer sandwiched between electrodes, in which electrical bistability is activated by in situ formation of silver nanoparticles embedded in a biopolymer upon light irradiation, as it is shown on Figure 18.

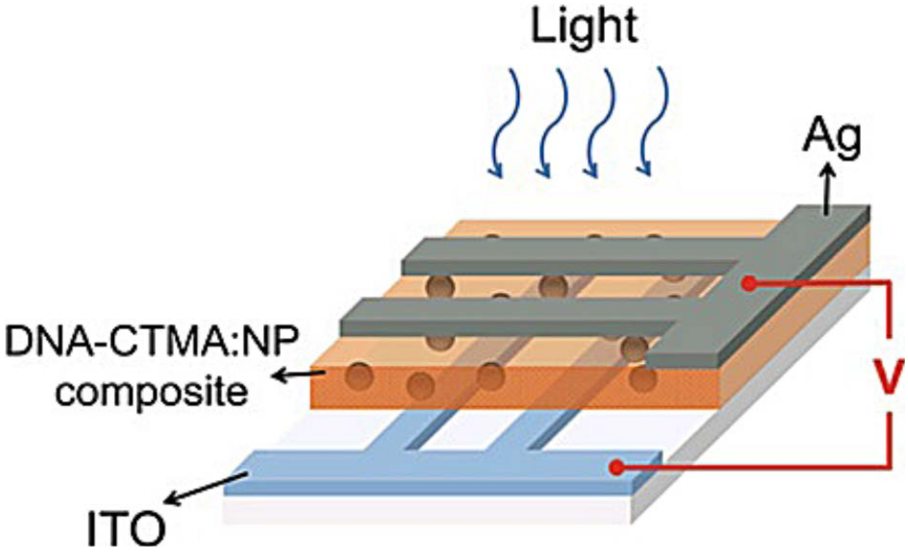


Fig. 18. Schematic design of a memory device consisting of a thin DNA biopolymer film sandwiched between electrodes. The memory switching effect is activated upon light irradiation.

The device functionally works when shining UV light on the system, which enables a light-triggered synthesis process that causes the silver atoms to cluster into nanosized particles and readies the system for data encoding. For some particular instance, the team has found that using DNA may be less expensive to process into storage devices than using traditional, inorganic materials like silicon, the researchers say. They said that when no voltage or low voltage is applied through the electrodes to the UV-irradiated DNA, only a low current is able to pass through the composite; this corresponds to the ‘OFF’ state of the device. But the UV irradiation makes the composite unable to hold a charge under a high electric field, so when the applied voltage exceeds a certain threshold, an increased amount of charge is able to pass through. This higher state of conductivity corresponds to the ‘ON’ state of the device. The team found that this change from low conductivity (‘OFF’) to high conductivity (‘ON’) was irreversible: once the system had been turned on, it stayed on, no matter what voltage the team applied to the system.

Once information is written, the device appears to retain that information indefinitely. The researchers hope that the technique will be useful in the design of optical storage devices and suggest that it may have plasmonic applications as well.

Consequently, WORM memories based on DNA a biopolymer nanocomposite have emerged as an excellent candidate for next-generation information storage media because of their potential application in flexible memory devices. This work combines new advances in DNA nanotechnology with a conventional polymer fabrication platform to realize a new emerging class of DNA-based memory.

Memory made from tiny islands of semiconductors - known as quantum dots - could fill a gap left by today's computer memory, allowing storage that is fast as well as long lasting. Researchers have shown that they can write information into quantum dot memory in just nanoseconds.

Memory is divided into two forms: DRRAM and Flash. Computers use DRAM, for short-term memory, but data does not persist for long and must be refreshed over 100 times per second to maintain its memory. On the other hand, Flash memory, like that used in memory cards, can store data for years without refreshing but writes information about 1,000 times slower than DRAM.

New research shows that memory based on quantum dots can provide the best of both: long-term storage with write speeds nearly as fast as DRAM. A tightly packed array of tiny islands, each around 15 nm across, could store 1 terabyte (1,000 GB) of data per square inch, the researchers say. Dieter Bimberg and colleagues at the Technical University of Berlin, Germany, with collaborators at Istanbul University, Turkey, demonstrated that it is possible to write information to the quantum dots in just 6 ns.

The key advantages of quantum dot (QD) NVMs are the high read/write speed, small size, low operating voltage, and, most importantly, multibit storage per device. However, these features have not been realized due to variations in dot size and lack of uniform insulator cladding layers on the dots.

Incorporating QDs into the floating gate results in a reduction in charge leakage and power dissipation with enhanced programming speed. Researchers in India and Germany have now unveiled the memory characteristics of silicon and silicon-germanium QDs embedded in epitaxial rare-earth oxide gadolinium oxide (Gd2O3) grown on Si (111) substrates as shown in the DQM structure in Figure 19.

Multilayer Si as well as single-layer $Si_{1-x}Ge_x$(where x = 0.6) QDs show excellent memory characteristics, making them attractive for next-generation Flashfloating-gate memory devices.

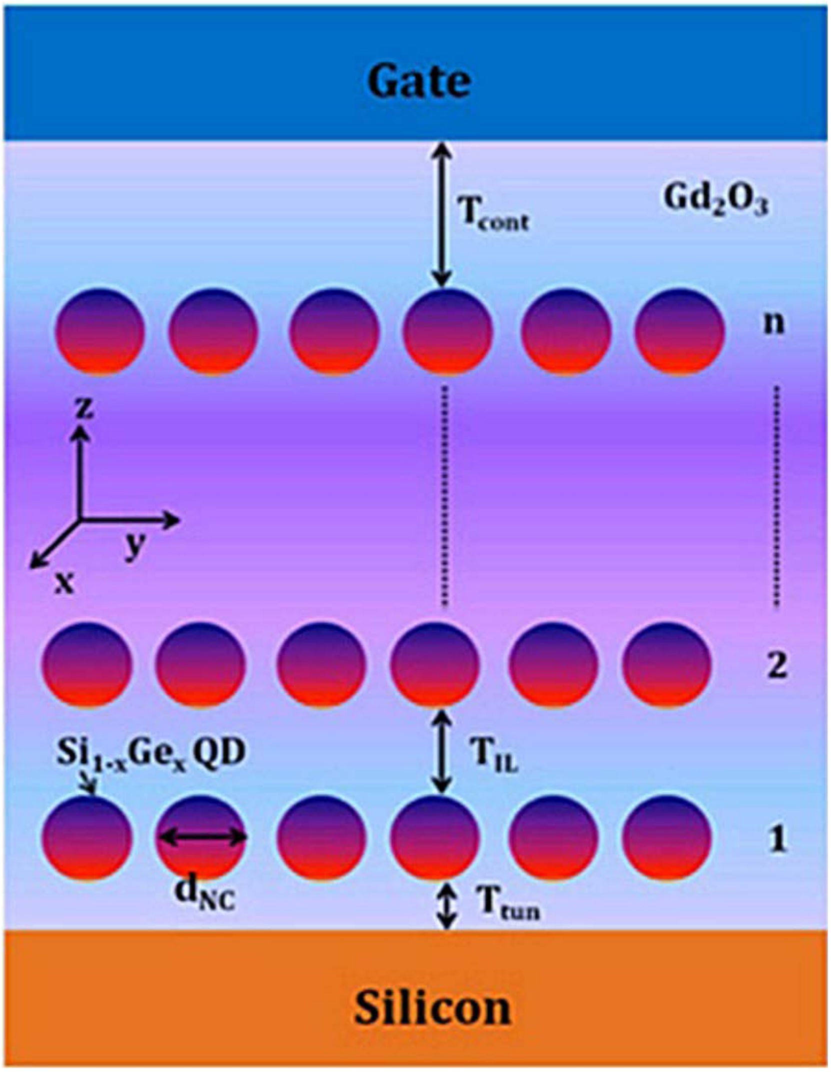


Fig. 19. Structure of quantum dot memory.

Memory producers are also trying to develop alternative technologies that may be scalable beyond 20-nm lithography. For true scalability beyond 20-nm technology nodes, it is necessary to design a cross-point memory array which does not require diodes for access elements. The cross-point memory architecture could be designed such that it can be easily fabricated in multiple layers to form a stacked 3-D memory. The 3-D technology has brought to high volume an NVM where arrays of memory cells are stacked above control logic circuitry in the third dimension, and stacking 3-D memory directly over CMOS allows for high array efficiency and very small die size. The 3-D technology uses no new materials, processes, or fabrication equipment, which control logic circuitry composed of typical CMOS. The memory construction uses typical back-end processing tools, and each memory layer is a repeat of the layers below it. The basic design of the 3-D cell consists of a vertical diode in series with a memory element as shown in Figure 20.

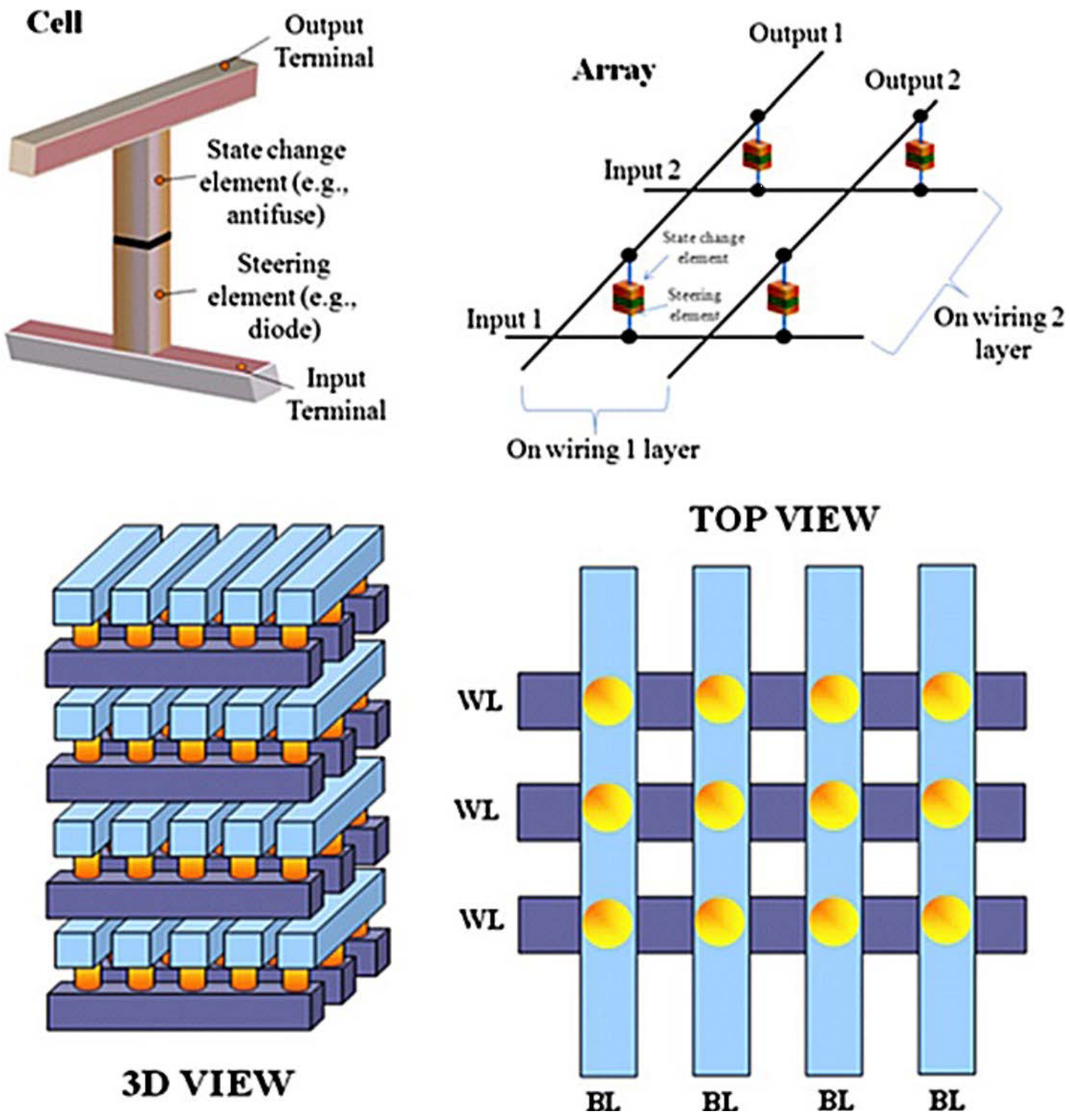


Fig. 20. The basic design of a 3-D cell that consists of a vertical diode in series. (top) Side view, (bottom right) top view, and (bottom left) 3-D view.

Building integrated circuits vertically allows for a reduced chip footprint when compared to a traditional 2-D design, by an approximate factor of the number of layers used. This offers significant advantages in terms of reduced interconnect delay when routing to blocks that otherwise would have been placed laterally. The process for the 2-D cross-point array can be built into a multilayer 3-D architecture.

Traditionally, a 3-D integrated circuit (3-D-IC) has used more than one active device layer. While resistance-change memory cells are not active devices, they function as rectifying devices in design.

Further characterization of the resistance-change material is also necessary in order to guarantee that the 3-D cross-point memory will be practical for data storage.

Also, the scalability of metal-oxide resistance-change materials beyond 20-nm technology nodes still needs to be studied. Moreover, the programming operation is expected to be competitive with both NAND and NOR Flash in terms of speed because of the relatively low voltage requirements of resistance-change materials. If the peripheral circuitry for accommodating the write operation can be made sufficiently compact, then the 3-D cross-point memory will indeed be a viable replacement for NAND and NOR Flash in future process generations.

Transparent and flexible electronics (TFE) is, today, one of the most advanced topics for a wide range of device applications, where the key component is transparent conducting oxides (TCOs), which are unique materials that oxides of different origin play an important role, not only as a passive component but also as an active component. TFE is an emerging technology that employs materials (including oxides, nitrides, and carbides) and a device for the realization of invisible circuits for implementing next-generation transparent conducting oxides in an invisible memory generation. In general, the TF-RRAM device is based on a capacitor-like structure (e.g., ITO/transparent resistive material/ITO/transparent and flexible substrate), which provides transmittance in the visible region. For such new class of memory technology, data retention is expected to be about 10 years. The basic structural design of the new memory chips is configured, namely with two terminals per bit of information on a transparent and flexible substrate rather than the standard three terminals per bit on a rigid and opaque substrate (Figure 21).

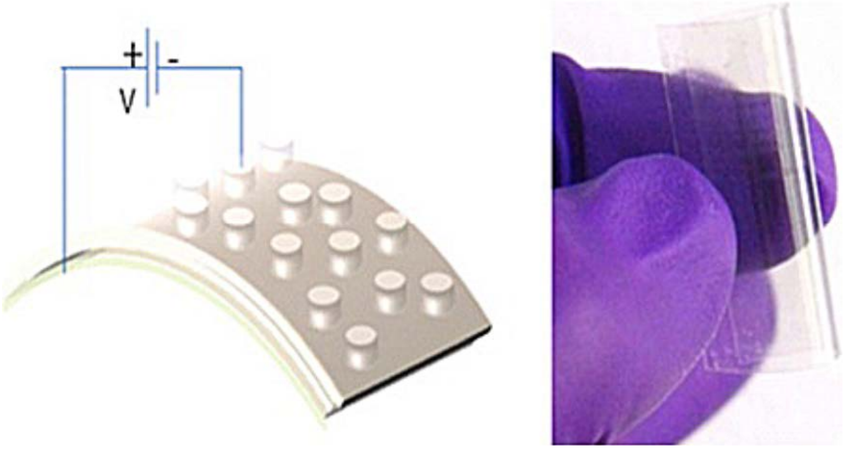


Fig. 21. A schematic design of FT-RRAM and a flexible, transparent memory chip image created by researchers at Rice University.

They are much better suited for the next revolution in electronic 3-D memory than Flash memory. These new memory chips that are transparent, are flexible enough to be folded like a sheet of paper, shrug off 1,000°F temperatures twice as hot as the max in a kitchen oven, and survive other hostile conditions could usher in the development of next-generation Flash-competitive memory for tomorrow's keychain drives, cell phones, and computers, a scientist reported today.

Speaking at the 243rd National Meeting and Exposition of the American Chemical Society, the world's largest scientific society, he said that devices with these chips could retain data despite an accidental trip through the drier or even a voyage to Mars. And with a unique 3-D internal architecture, the new chips could pack extra gigabytes of data while taking up less space. Despite the recent progress in TF-RRAM, it needs lots of work to satisfy the dual requirements of resistance to repeated bending stress and transparent properties. Thus, it is supposed that an achievement of such TF-RRAM device will be the next step towards the realization of transparent and flexible electronic systems. We hope that FT-RRAM devices will mark a milestone in the current progress of such unique and invisible electronic systems in the near future.

One-transistor one-resistor (1T1R)-RRAM is also one class of emerging memory technology with impressive characteristics. It meets the demands for next-generation memory systems. It is expected that 1T1R-RRAM could be able to meet the demand of high-speed (e.g., performance) memory technology. The 1T1R structure is chosen because the transistor isolates current to cells, which are selected from cells which do not. The basic cell structure of 1T1R is depicted in Figure 22.

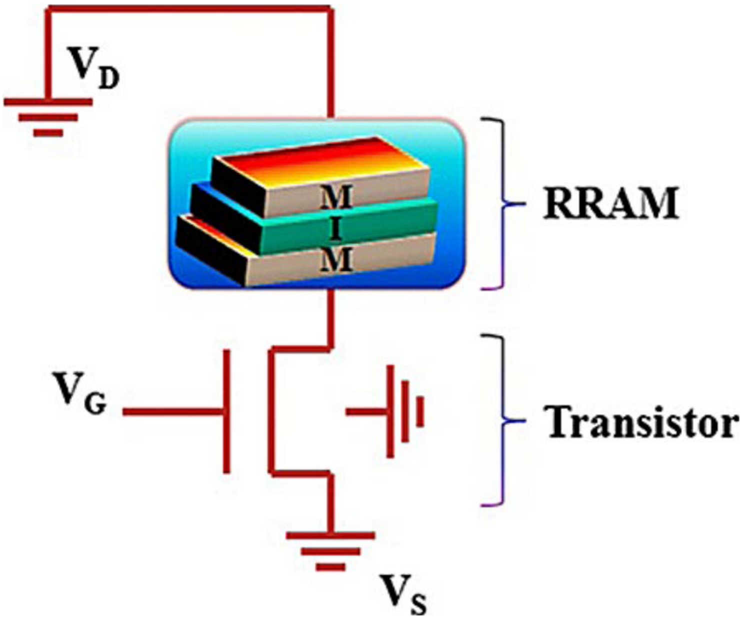


Fig. 22. The basic cell structure of 1T1R-RRAM.

1T1R-RRAM consists of an access transistor and a resistor as a storage element. Zangeneh and Joshi have also mentioned that the 1T1R cell structure is similar to that of a DRAM cell in that the data is stored as the resistance of the resistor and the transistor serves as an access switch for reading and writing data. In reference to this, they revealed the 1T1R cell as the basic building block of a NVRRAM array as it avoids sneak path problem to ensure reliable operation.

Moreover, the 1T1R structure is more compact and may enable vertically stacking memory layers, ideally suited for mass storage devices. But, in the absence of any transistor, the isolation must be provided by a ‘selector’ device, such as a diode, in series with the memory element, or by the memory element itself. Such kinds of isolation capabilities have been inferior to the use of transistors, limiting the ability to operate very large RRAM arrays in 1T1R architecture. 1T1R memory polarity can be either binary or unary. Bipolar effects cause polarity to reverse reset operation to set operation. Unipolar switching leaves polarity unaffected but uses different voltages. Other potential emerging classes of memory technologies, we are describing in short, are molecular tunnel memory (MTM), polymeric ferroelectric RAM (PFRAM), spinpolarized beam magnetic memory (SPBMM), light memory, and complex metal-oxide RRAM (CMORRAM). We can say that these are sister memory technologies of molecular memory, ferroelectric/polymer memory, magnetic memory, and metal-oxide RRAM, respectively.

Although these new technologies will almost certainly result in more complex memory hierarchies than their family memories, they are likely to allow the construction of memory chips that are nonvolatile, have low energy, and have density and development close to or better than those of DRAM chips, with improved performance and allowing memory systems to continue to scale up.

This chapter reviewed the historical development to the recent advancement on memory architecture and scaling trend of several conventional types of Flash within the MOS family and projected their future trends. With great progress being made in the emerging memory technologies, current trends and limitations were discussed before leading to some insight into the next generation of memory products.

For the past three and a half decades in existence, the family of semiconductor memories has expanded greatly and achieved higher densities, higher speeds, lower power, more functionality, and lower costs. In the past 40 to 50 years, NVSM has grown from the FG concept to FAMOS, SAMOS, Flash memory, multilevel cells, RRAM, 3-D structures, and TF-RRAM.

Since 1990, NVSM is an inspired technology, which has ushered in the digital age, enabled the development of all modern electronic systems, and brought unprecedented benefit to humankind. At the same time, some of the limitations within each type of memory are also becoming more realized. As the device dimension is reduced to the deca-nanometer regime, NVSM faces many serious scaling challenges such as the interface of neighboring cells, reduction of stored charges, and random telegraph noise.

As such, we hope and are confident that there are several emerging technologies aiming to go beyond those limitations and potentially replace all or most of the existing semiconductor memory technologies to become a USM.

Despite these limitations, the field of conventional semiconductor memories would continue to flourish and memory device scientists will find the way to meet these challenges and may even develop a ‘unified memory’ with low cost, high performance, and high reliability for future electronic systems.

Progress towards a viable new resistive memory technology relies on fully understanding the mechanisms responsible for switching and charge transport, the failure mechanisms, and the factors associated with materials reliability. Moreover, the development of current redoxbased resistive switching will help to improve our old technologies, and further research will produce more impressive results that will benefit industries and society to improve the quality of life for billions of people around the world.

Chapter 15

Nanoparticle Organic Memory Field-Effect Transistor

Organic transistors have emerged as a revolutionary semiconductor technology because of their customizable surface chemistry, low temperature organic deposition and their much simpler fabrication processes compared to their inorganic counterparts. Conventional Silicon technology involves high-temperature and high-vacuum deposition processes and intense photolithographic patterning methods. However because of the relatively low mobility of the organic semiconductor layers, OTFTs cannot outperform single-crystalline inorganic semiconductors, such as Si and Ge, which have charge carrier mobility (μ) about three orders of magnitude higher. So, organic semiconductors are not suitable for use in applications requiring very high switching speeds. But they can be competitive as organic semiconductor devices like OFETs and OTFTs which can be fabricated easily by low-temperature deposition and solution processing methods1. In addition, the mechanical flexibility of organic materials makes them naturally compatible with plastic, paper substrates making lightweight products.

Nowadays, transparent and flexible electronics are one of the technologies with the widest range for innovative products, so they are the focus of several research groups and enterprises. One reason for this interest is the chance to integrate identification tags, smart cards or flexible displays using almost the same integration processes. Moreover, the cost of these devices is kept low due to the innovative use of newly created or adapted processes for large-area and flexible substrates. Thin-film transistors (TFT) are an essential element of this technology. They are responsible for the driving current in memory cells, digital circuits or for light-emitting diodes (LEDs). Another characteristic of this technology is the opportunity to introduce new materials, which improve the electrical performance, simplify the integration process or even add new mechanical properties to the final product.

Organic and inorganic semiconductors have been used to integrate TFTs for more than 20 years. They exhibit better characteristics in comparison to amorphous silicon-based transistors due to the achieved performance and low production cost. Nevertheless, when pursuing a cost-efficient production with a high reliability and restrained mechanical characteristics, there are new requirements. Most of them are related to the fact that flexible substrates can only be handled at low temperatures, due to their flexibility, *i.e.,* all deposited materials suffer from mechanical stress during their fabrication or use. For instance, the use of non-malleable metal connections or dielectrics may induce a failure after bending the substrate. For this sake, all components (materials and integration processes) have to be selected carefully to fulfill these requirements.

Thin-film transistors basically consist of three elements: the active semiconductor, the electrical connections (source, drain and gate) and the dielectric layer. The dielectric layer separates the gate electrode from the active semiconducting film, whereas the source and drain contacts are directly connected to the semiconductor, determining the transistor's dimensions. In general, the electrical connections consist of thermally evaporated or sputtered metal layers. The choice of the drain/source contact material is crucial, as it determines whether electron or hole injection is preferred. This depends on the position of its Fermi level with respect to the frontier molecular orbitals (highest occupied molecular orbital, HOMO, and lowest unoccupied molecular orbital, LUMO) and the valence/conduction band in the case of organic and inorganic semiconductors, respectively. Considering the metal-semiconductor interface, a Schottky barrier will be formed (Figure 1), which has poor blocking properties in comparison to a pn-junction. As a voltage is applied between drain and source contacts, an off-current occurs due to the tunneling of charge carriers across the formed potential barrier (Figure 1b). The potential barrier width can be reduced by applying a voltage to the gate electrode and, as a result, the drain current increases (Figure 1c).

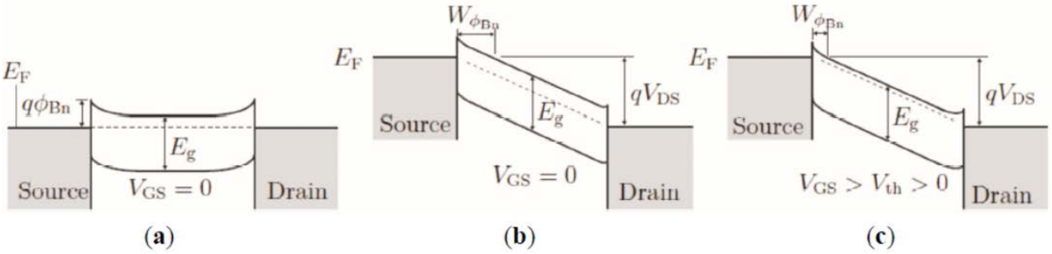


Fig. 1. Band diagram of the Schottky contacts to an n-type semiconductor with metal contacts: (**a**) without applied voltages; (**b**) with drain voltage and (**c**) with drain and gate voltages.

In general, TFT operation is expressed using the same models as for the metal-insulator-semiconductor field-effect transistor (MISFET). Although it may not provide quantitative agreement with the device, it represents the transistor's basic operation. As discussed above, the three elements, *i.e.*, semiconductor, dielectric and metal contacts, are the main components of a TFT. The arrangement of these elements has a strong influence on both the device performance and the integration process itself. The devices are divided into two main groups, the staggered and the coplanar setup. The position of the drain/source electrodes and the gate dielectric in relation to the channel area defines the transistor configuration. In coplanar structures (Figure 2b,d), the drain/source electrodes and the dielectric are located at the same side of the channel, whereas in staggered structures (Figure 2a,c) the drain/source electrodes and the dielectric are placed at opposite sides. The position of the gate electrode is also used to classify the setups as either bottom-gate (Figure 2a,b) or top-gate devices (Figure 2c,d). Usually, bottom-gate devices are referred as an inverted structure.

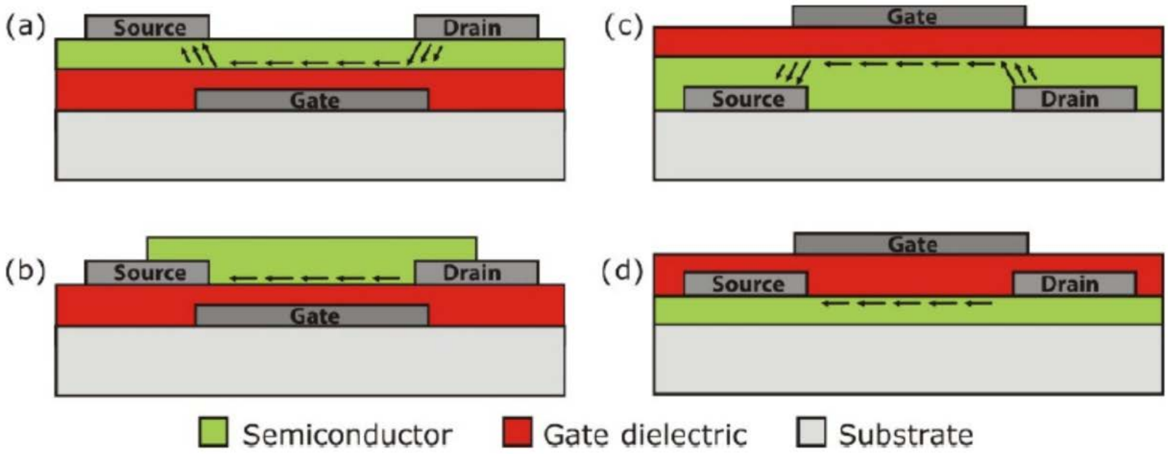


Fig. 2. General thin-film transistor configurations: (a) staggered bottom-gate; (b) coplanar bottom-gate; (c) staggered top-gate and (d) coplanar top-gate.

The contact resistances between the metal (drain and source) electrodes and the semiconductor reduce the maximum current that will flow through the device in the on-state. Nevertheless, as a Schottky barrier is formed at those contacts, the contact resistance during the transistor off-state should be high enough to prevent a high leakage current. For staggered structures, a major contribution to the contact resistance is related to access resistance. This mainly occurs because the charge carriers have to travel through the semiconducting layer thickness before reaching the gate dielectric/semiconductor interface. In the case of coplanar setups, the drain and source electrodes are already in contact with the formed accumulation layer, which results in a low access resistance. However, results show that if drain/source electrodes are deposited on the active semiconductor, the access resistance is reduced due to valleys in the semiconducting film. An additional reduction of the contact resistance is observed due to the increased contact area when the drain/source metal is deposited on the semiconductor layer.

Moreover, the integration process has to be adapted for the different setups. The inverted coplanar layout permits the semiconducting layer to be deposited in the last step. For this reason, the semiconducting film will not suffer from chemical impacts and other integration process steps.

For other layouts, the semiconductor has to withstand different processes, as for example lithography, etching or annealing steps. Deposition methods for metals, dielectrics, semiconductors and all materials used in the integration process should be also entirely suitable for flexible electronics. However, there are variations, either in the quality of the material, temperature or production costs. For the deposition of the semiconductor, commonly vacuum techniques produce a denser film with higher charge carrier mobilities.

During the last few decades, pentacene seems to be the most promising candidate for organic semiconductor applications. Typically, pentacene layers are deposited by evaporation under high- or ultra-high-vacuum conditions, whereby the substrate temperature can be vary between room temperature and 80 °C to improve the self-ordering of the semiconductor molecules. Here, charge carrier mobilities up to 3 cm2/Vs could be achieved. Furthermore, deposition under low vacuum conditions assisted by an inert gas stream is possible.

Nowadays, new materials promising good environmental stability and higher mobilities have been synthesized. Various research groups focus on materials such as dinaphtho(2,3-b:2′,3′-f)thieno (3,2-b)thiophene (Cn-DNTT) and 2,7-dioctyl(1)benzothieno(3,2-b)(1)benzothiophene (C8-BTBT), which can be deposit by thermal evaporation or from solutions, respectively. Currently, the highest achieved field-effect mobility is about 43 cm2/Vs for C8-BTBT deposited by a special off-centered spin-coating technique.

For inorganic semiconducting films, the highest mobilities and denser films are commonly achieved using sputtering techniques. With this technique it is possible to deposit binary compounds such as ZnO, In_2O_3, and semiconductor compositions with different ratio between elements, such as, IGO, ZTO, GIZO. In 2003, Hoffman *et al.* have demonstrated a fully functional transparent ZnO TFT. Even using high annealing temperatures of around 600 °C–800 °C, the device mobility was found to be 0.35–0.45 cm^2/Vs. Using RF magnetron sputtering, Fortunato *et al.* reported a high mobility (70 cm^2/Vs) ZnO based TFT. The transistor performance can be further improved by using a GIZO target or co-sputtering techniques. Additionally, adjusting the pressure in the chamber and the power during the process, the defect concentration as well as the film density can be controlled. When oxygen is added, a reactive process occurs and the amount of oxygen in the film can be adjusted. Atomic layer deposition (ALD) is another method to achieve high-density layers as demonstrated by. Nevertheless, these films are deposited under vacuum conditions, increasing the production cost. Conjointly, when large area substrates are used, the equipment size will limit the template size.

For this reason, methods using solutions have greater potential to fulfill the requirements imposed for large-area and flexible substrates. They can be divided into two main groups, *i.e.*, molecular precursors and nanoparticle suspensions. For molecular precursors, the temperature in which the semiconductor film will be synthesized depends on the type of precursor. Commonly nitrate precursors require lower annealing temperatures than acetate or chloride precursors, and greater performance is achieved.

In the case of nanoparticle suspensions, the semiconductor is already synthesized in the form of nanoparticles being dispersed in a solution, either water or other solvents, like ethanol, isopropanol and buthyl acetate. That means it is possible to fabricate high-quality nanoparticles in mass production, using high temperature and vacuum processes. It maintains an attractive cost base and has almost no influence on the transistor integration process. After the nanoparticle deposition, an annealing process removes the solvent.

One of the main advantages of using a solution based process is the variety in coating methods. Most of the literature refers to spin-coating methods; however, methods like inkjet printing, spray coating, doctor blade and Meyer rods have attracted the interest of research groups, due to the opportunity to integrate low-cost devices on large area and flexible substrates either using organic or inorganic semiconductors.

To investigate and integrate thin-film transistors at submicron scales, a suitable process is required. In general, nanostructures are established using expensive equipment, like high-resolution lithography or electron beam lithography. Nonetheless, these methods are not entirely suitable for large area and flexible substrates. The side-wall deposition and etch-back (SWEB) technique is a cheap method to fabricate nanostructures applying only conventional lithography. The SWEB technique can be used to integrate nanoscaled field-effect transistors in bulk silicon.

Figure 3 depicts the schematic integration process of this method on silicon. Since silicon is stable at high temperatures, a process temperatures of around 800 °C are used for the nanostructure formation.

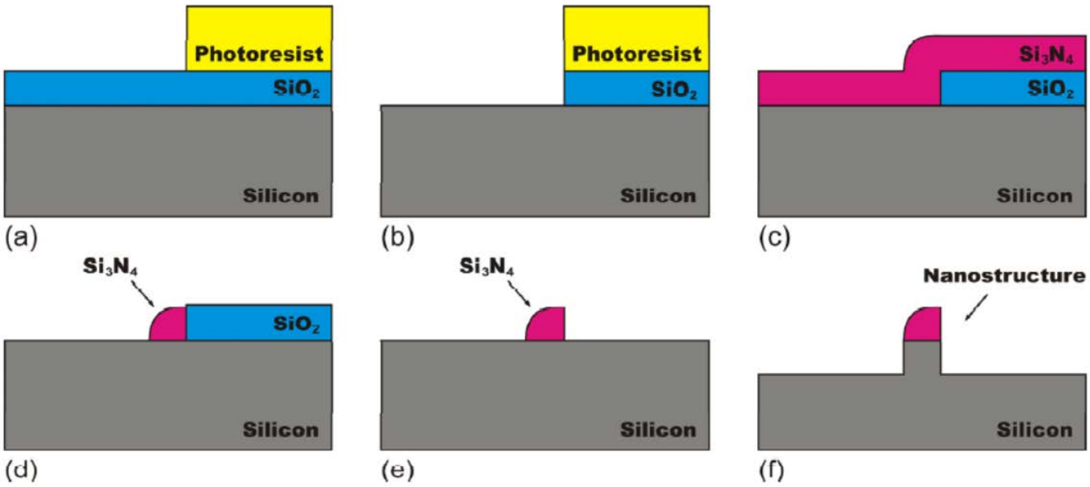


Fig. 3. Schematic process of the side wall deposition and etch-back technique. (a) The photoresist is structured by conventional photolithography and (b) transferred to the underneath SiO_2 layer. (c) Subsequent to the photoresist removal, Si_3N_4 is deposited conformably and (d) etched back by a directional dry etching process. (e) The SiO_2 layer is removed by selective wet etching and a Si_3N_4 spacer stays on the surface. (f) Applying the spacers as a mask, the structure can be transferred to the underneath layer.

The main advantage of the SWEB technique is that the transistor size is not defined by the lithography itself, but by the deposition of a sacrificial layer in combination with an etching process. After the thermal oxidation of the wafer or the deposition of SiO_2 by plasma-enhanced chemical vapor deposition (PECVD), the position of the nanostructure is defined by a standard photolithography step and a subsequent anisotropic etching (Figure 3a–c). As a result, a rectangular edge, perpendicular to the wafer surface, is formed in the oxide film. After the photoresist removal, silicon nitride (Si_3N_4) is deposited conformably by low pressure chemical vapor deposition (LPCVD). Due to the very high uniformity of this deposition (better than +/−1%), the thickness of the deposited film, which defines the transistor channel length afterwards, can be controlled carefully. Subsequently, the Si_3N_4 is etched back by a directional dry etching process. After removing of SiO_2 by selective wet etching, a Si3N4 spacer stays on the surface, as shown in Figure 3d,e. As the structure size is defined by the thickness of the deposited silicon nitride film, it is possible to achieve nanostructures using standard lithography. Applying the spacers as a mask, the structure can be transferred to the underneath layer (Figure 3f). An example of nanostructures fabricated using the SWEB technique is shown in Figure 4a. Nanoscaled MOSFETs can be integrated by structuring polysilicon. Figure 4b depicts the output characteristics of a 70 nm channel length transistor using μm resolution photolithography. Additional information and discussion about this technique can be found in.

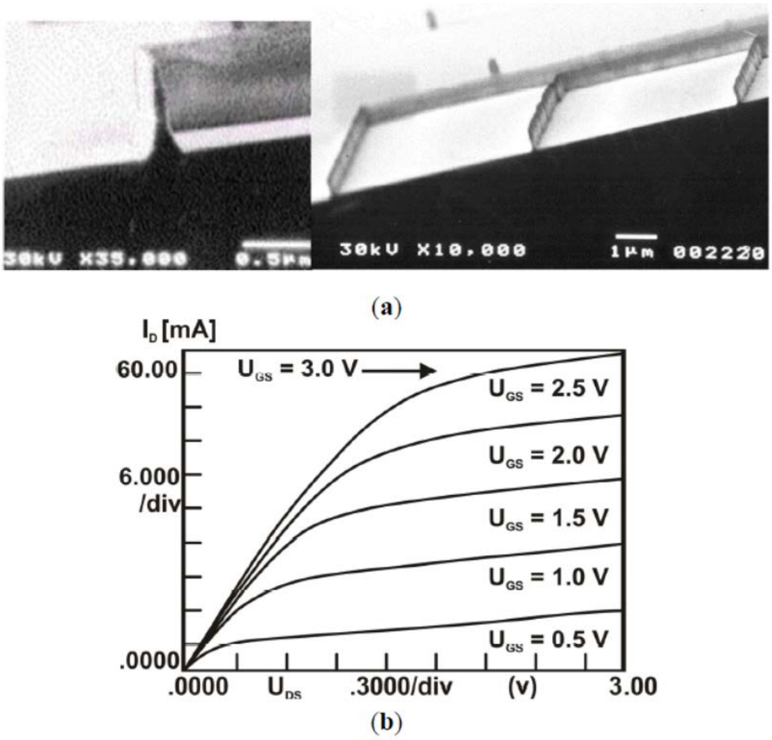


Fig. 4. (a) Example of structures fabricated using the SWEB technique on silicon bulk and polycrystalline Si on SiO_2; (b) electrical characteristic of a MOSFET integrated by SWEB with L = 70 nm and W = 300 μm.

In the MOSFET technology, the creation of a nanoline acting as the gate electrode is sufficient as the source and drain contact areas are defined subsequent to the ion implantation. In contrast, to integrate nanoscaled TFTs on glass and flexible substrates an additional step is required to define the channel length of the transistor by a nanogap. Figure 5 shows the adaptation of the SWEB technique to avail the process to flexible substrates. First, the position of the structure is defined by conventional photolithography, where the sacrificial layer is the photoresist itself (Figure 5a). A PECVD-SiO_2 layer is deposited at low temperature and etched anisotropically (Figure 5b,c). After the photoresist removal, the nanostructure (nanoline) stays back (Figure 5d). In order to achieve a nanogap from a nanoline, a metal is evaporated anisotropically (Figure 5e). By using a selective etching process, the PECVD-SiO_2 is removed creating a nanogap (transistor channel) as shown in Figure 5f.

Depending on the integrated setup (staggered/coplanar or inverted/noninverted), the SWEB is done directly on the substrate or after the integration of the gate electrode and dielectric. Nevertheless, when the technique is performed over the gate dielectric, the anisotropic etching of the PECVD-SiO_2 can damage the dielectric layer, if the etching time is not defined carefully. Furthermore, all processes should be selected to fulfill the substrate requirements; for instance, the process temperature is limited to 150 °C.

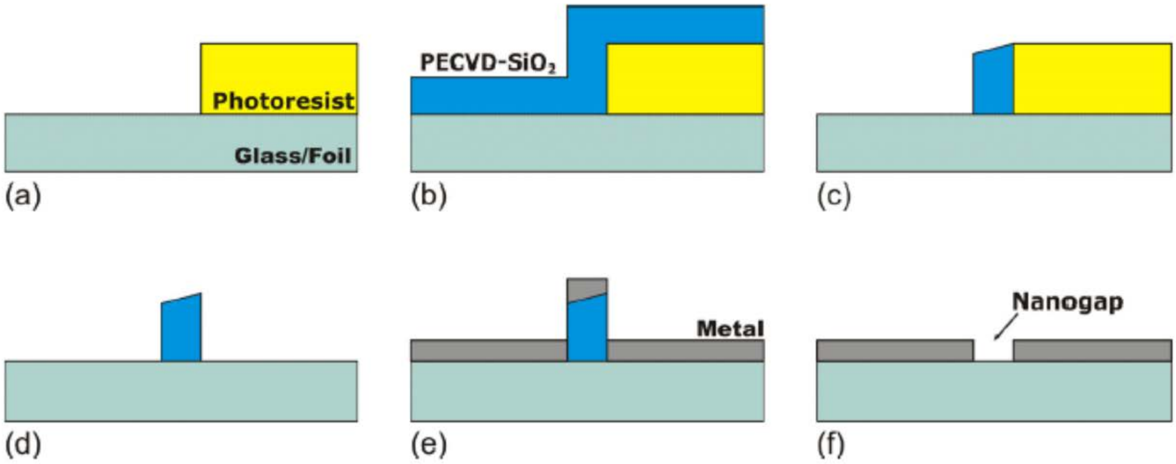
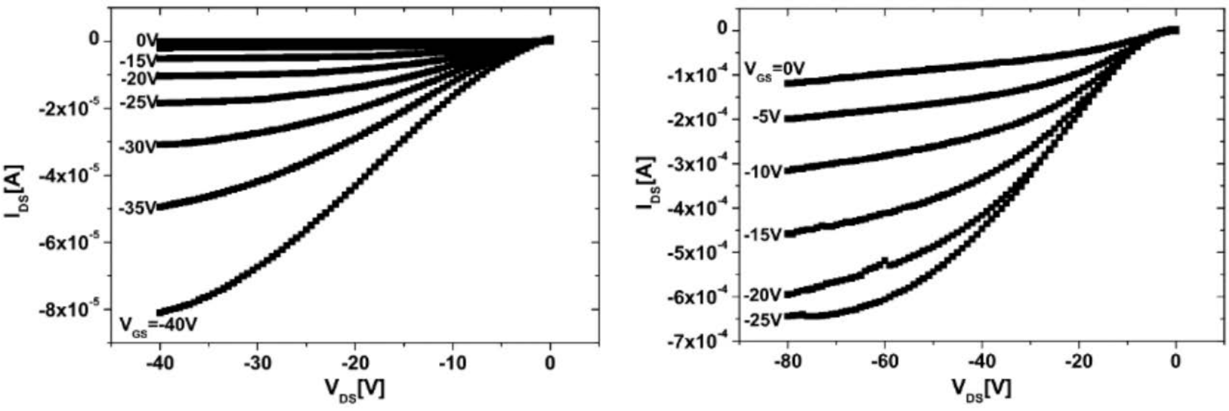


Fig. 5. Process flow for integrating nanoscaled gaps on glass or foils. (a) The photoresist is structured by conventional photolithography. (b) A SiO2 layer is deposited at low temperature and (c) etched anisotropically. (d) Subsequent to the photoresist removal, a nanoline stays back. (e) A metal is evaporated anisotroically and (f) the SiO_2 is removed by a selective etching process.

The properties of the gate dielectric are essential for the performance of OFET-devices. Roughness and polarity of the insulating layer surface are crucial for the morphology and distribution of the electronic states of the grown organic semiconductor. The first OFET-devices were developed using silicon dioxide, a standard dielectric material in MOSFET technology. The growth process of SiO_2 can be well-controlled, providing a low defect density and a smooth surface to produce well-oriented organic semiconducting films. Besides the high electrical stability, thermally grown SiO_2 on a heavily doped Si wafer acts as template to evaluate new synthesized organic semiconducting materials. As a first investigation, 30 nm pentacene were evaporated under high-vacuum conditions as an active semiconducting layer, in which the drain and source contacts consist of gold. The transistor exhibits a threshold voltage of about 1.5 V and an on-off ratio of 10^3 at a drain-source voltage of VDS = −40 V.

Notwithstanding of the high gate voltage, no electrical breakdown could be observed. In order to improve the growth process of the pentacene film, an oxygen plasma surface treatment of the dielectric material was performed before the evaporation. Figure 6 shows the electrical characteristics of OFET-devices using thermally grown SiO2 as insulating layer (a) without and (b) with prior treatment by an oxygen plasma. Evidently, the surface treatment leads to a threshold voltage shift to 17.2 V regarding the polarization of the dielectric due to the generation of free bonds at its surface. Furthermore, the electrical strength diminished, resulting in an increasing leakage current at off-state. Nevertheless, the surface treatment leads to higher drain currents at lower gate voltages, regarding the increased pentacene crystallites at the dielectric surface. AFM studies found a crystallite diameter of 250 nm at untreated and 1 μm at treated SiO_2 surfaces, respectively.



(a) (b)

Fig. 6. Output characteristics of an inverted coplanar pentacene OTFT using SiO_2 as gate dielectric (L = 1 μm, W = 1000 μm). (a) Without oxygen plasma treatment (t_{ox} = 150 nm); (b) with oxygen plasma treatment in a parallel-plate reactor (t_{ox} = 100 nm).

As the thermal growth of SiO2 cannot occur on other substrate materials with individual gate contacts, other deposition techniques for the gate dielectric are required for the realization of digital circuits.

A low-temperature-oxide (LTO) can be deposit by CVD techniques. Triethylsilane forms silicon dioxide by the separation of ethyl groups from the residual silane under low pressure and continuous oxygen flow. Due to the process temperatures of 550 °C, titanium can be used as a gate material. To avoid the oxidation of the gate metal during the LTO-deposition, an additional step to nitrify the titanium can be performed. Nevertheless, large leakage currents and a low on-off ratio that are related to carbon atoms from ethyl groups prevent the building of digital circuits. Even after an additional thermal treatment, just a slightly better breakdown voltage can be observed.

By using PECVD-deposition techniques, a further decrease of the process temperature can be achieved. At typical process temperatures between 130°C and 300°C, silicon dioxide is formed by a chemical reaction of silane and oxygen. Due to the low temperatures, the nitridation of the titanium gate electrode is not necessary, but a high defect density and low electrical stability restrain the exclusive use of PECVD-oxide for OTFT. Nevertheless, PECVD-oxide can be used as an oxidation barrier for the metal gate electrode, replacing the time-consuming step of nitrifying titanium. Admittedly, a reduction of the leakage current can be observed compared to the above mentioned LTO-process, but a rather low on-off ratio exhibits the impossibility of building digital circuits.

Flexible electronic devices using polymeric substrates such as polyethylene terephthalate (PET) or polypropylene (PP) are sensitive to thermal treatments. High process temperatures, electrical instability and the low dielectric constant of SiO_2 ($k = 3.9$) create a demand for other deposition techniques and materials. Reactive sputtering techniques provide process temperatures down to room temperature, but it should be mentioned that sputtering of the gate dielectric leads to a high surface roughness.

Experiments using tantalum pentoxide Ta_2O_5 ($k = 23$) as dielectric material show a rather low mobility of $8.2 \times 10{-}5$ cm^2/Vs of the OFET-device regarding to an amorphous growth of the pentacene film.

Furthermore, it is shown that the stoichiometric ratio of oxygen and tantalum is important to control. An excessive amount of tantalum leads to a metallic character in the insulating layer.

Inorganic dielectric materials have in common that they lack mechanical elasticity, hence they are partly usable in flexible electronic devices. However, polymeric dielectrics are featured with high mechanical flexibility and rather low processing temperatures. A commercially available coating varnish, Bectron, which is based on modified alkyd chemistry, was used for initial investigations. Its low curing temperatures of 80°C and deposition by spin-coating show possible application for flexible electronics.

Limited stability against solvents that are part of optical lithography and insufficient control of the deposited layer thickness led to efforts being discontinued.

Organic-inorganic nanocomposites are intertwining the advantages of inorganic and polymeric dielectrics. While the polymeric matrix leads to the flexible attributes of the nanocomposite material, the permittivity can be adjusted by the involved inorganic component. Therefore, a high-*k*-resist based on hydrolyzed and partially condensed ethyl silicates was used. The high-*k*-resist contains different amounts of zirconium dioxide or titanium dioxide resulting in dielectric constants between 9 and 12.

The deposition of the resist can be done by spin-coating. The curing process consists of a soft-bake step at temperatures of 80°C, followed by an UV-curing step for cross-linking of the polymeric matrix material. The resulting layers exhibit a thickness in the range of 200 nm and 400 nm and a sufficient stability against the common solvents used in optical lithography.

Figure 7 shows the electrical characteristics of an OFET-device using a 200 nm thick high-*k*-resist layer as dielectric material. A 150 nm thick aluminum layer serves as the gate electrode. As active semiconducting material, 30 nm of pentacene was thermally evaporated under high-vacuum conditions. The transistor exhibited a threshold voltage of about 7.4 V, an on-off ratio greater than 102 and a saturation regime mobility of 0.022 cm^2/Vs.

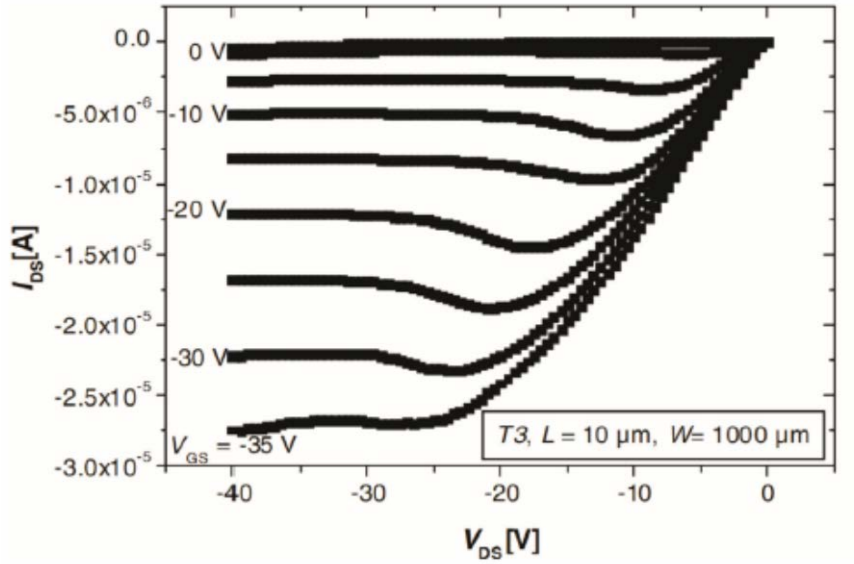


Fig. 7. Output characteristics of an inverted coplanar pentacene OTFT using a high-*k* dielectric (L = 10 µm, W = 1000 µm and tins = 200 nm).

In general, organic materials are influenced by environmental conditions such as ambient humidity and oxygen. As organic semiconductors act as active layers in OTFT, the device’s performance is directly affected. Water vapor and oxygen lead to modifications in the charge carrier field-effect mobility and in the threshold voltage. To investigate the impact of ambient humidity and oxygen, long-term experiments were executed. Therefore, an OFET-device using 110 nm of thermally grown SiO_2, as insulating layer, on a heavily doped Si substrate was prepared. A 58 nm thick pentacene layer was evaporated under high-vacuum conditions. The long-term experiments were performed over a period of 9 months at intervals of 3 months. The device was stored under ambient dark atmosphere conditions in a shielded metal box. The measurements were performed under ambient conditions.

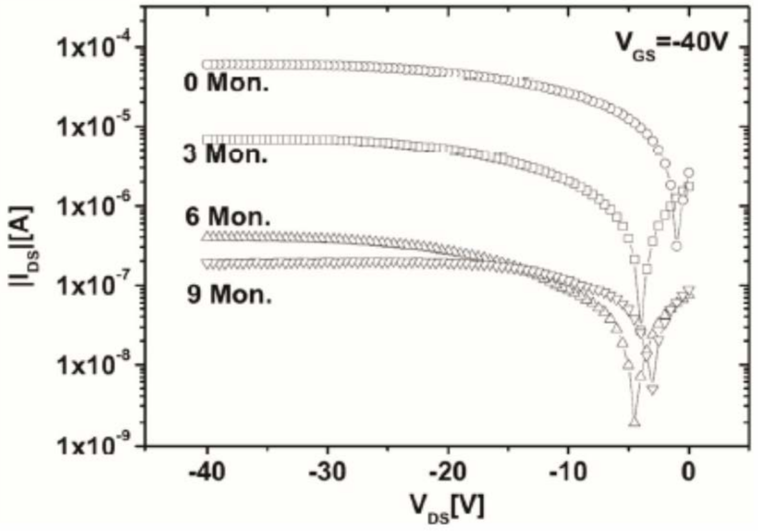


Fig. 8. VDS—IDS characteristics of an inverted coplanar pentacene OTFT at VGS = −40 V.

Figure 8 shows the electrical characteristics of the device during the long-term experiment. Over the 9 months interval, the drain current decreases from −61 µA to −187 nA, the charge carrier mobility in the linear regime diminishes from 2 × 10−3 cm^2/Vs to 1.2 × 10−5cm2/Vs, and the threshold voltage shifts from 4.8 V to −8 V. Table 1 shows an overview of the results. Further experiments specify the dominating agent in the device’s degradation. Therefore, a chip carrying a proceeded OTFT template was bonded and connected to the measurement setup, followed by placement in a vacuum chamber. A 30 nm thick layer of pentacene was evaporated under high-vacuum conditions. Subsequent to the measurement under high-vacuum conditions, the chamber was flooded with pure technical oxygen before repeating the measurement. The relevant transfer characteristics are shown in Figure 9.

Table 1. Behavior of the electrical parameter of an inverted coplanar pentacene OFET-device

Time	I_{DS} (µA)	V_{th} (V)	µ (cm^2/Vs)
0 months	−60.9	4.8	2.0×10^{-3}
3 months	−6.8	2.3	2.4×10^{-4}
6 months	−0.457	−3.4	2.2×10^{-5}
9 months	−0.187	−8	1.2×10^{-5}

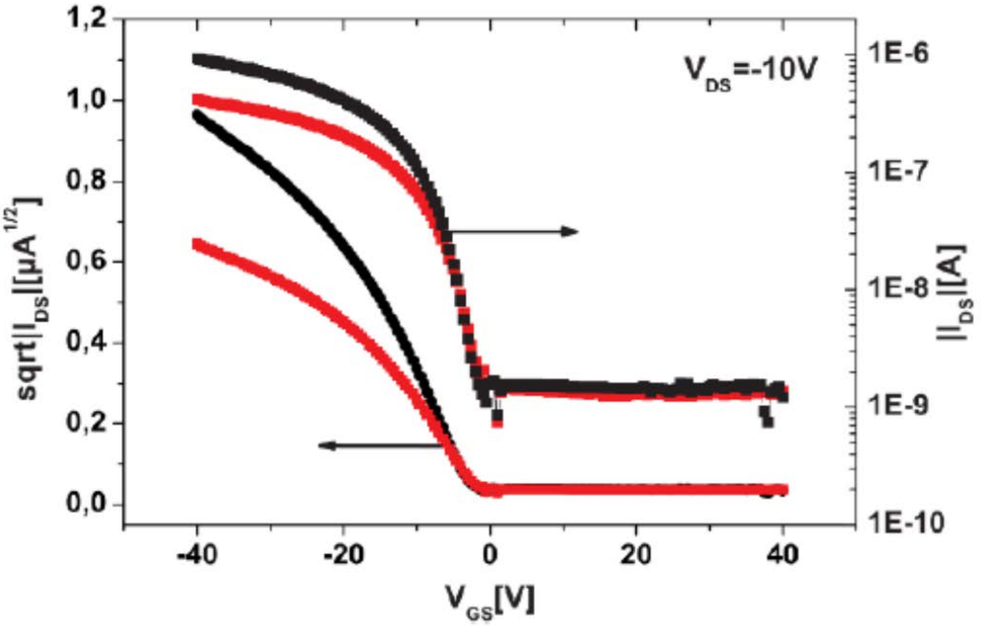


Fig. 9. Transfer characteristics of an inverted coplanar pentacene OTFT using SiO_2 as gate dielectric (L = 1 µm, W = 1000 µm and tox = 110 nm). Measurement under vacuum conditions (black curves) and after the impact of oxygen (red curves).

Subsequent to the flooding of the chamber, the drain current in the saturation regime decreases from −15.9 µA to −6.4 µA, whereas the threshold voltage shifts from −2.6 V to −0.7 V. This threshold voltage shift is related to the influence of oxygen on the molecular structure of pentacene. Oxygen is connected to the middle benzene ring of the pentacene molecule, creating acceptor states for electrons. As a result, negative charges at grain boundaries and at the dielectric surface were introduced.

It should be mentioned that the threshold voltage shifted in the opposite direction as observed in the long term degradation experiments. From the literature it is well-known that water molecules can diffuse into the organic semiconducting material and dissociate under the impact of an electric field. The formed H+ and OH− ions create defects in the pentacene crystal, acting as traps for positive charge carriers.

Therefore, the current decreases, the interface between the gate dielectric and the semiconductor becomes polarized and a threshold voltage shift in the negative direction can be observed. It should be mentioned that the sign of the permanent bias stress has an impact on the direction of the threshold voltage shift. Nevertheless, regarding both experiments, the influence of water seems to dominate the degradation of pentacene OTFT devices. As a result, an efficient encapsulation of pentacene OTFT devices is needed to ensure long term stability. As oxygen also diffuses through polymeric substrates, a flexible encapsulation of the completed device is required.

It is now well recognized that electronic circuits based on von Neuman paradigm are unable to catch the complex, real-world environment, behaviors as do a biological neural systems (e.g. human brain). One of the reason is the so-called von Neuman bottleneck due to the physical separation of computing units and memories. In the brain, memory and computation are mixed together and allow the processing of information both in time and in space via the time dependent properties of interconnected neurons. This challenge for the development of a new generation of computers has induced a lot of efforts in neuroscience computation activities and the framework for the spatiotemporal processing of information seems to be theoretically achievable.

One key element that is still a limitation concerns the integration of neurons and synaptic connections in order to realize a brain-like computer. Even if silicon CMOS chips have been designed and fabricated to emulate the brain behaviors, this approach is limited to small system because it takes several (at least 7) silicon transistors to build an electronic synapse. As the human brain contain more synapses than neurons (by a factor of ~ 10^7), it is mandatory to develop a nano-scale, low power, synapse-like device if we want to scale neuromorphic circuits towards the human brain level. This feature has recently prompted the research for nano-scale synaptic devices. Proposals for such programmable memory devices include optically-gated CNTFET, organic/hybrid Si nanowire transistor memristor.

Here, we show that mixing nanoparticles (NPs) and molecules to implement computation and memory in a single synapse-like device is a powerful approach towards such objectives. NPs and molecules are nano-size objects suitable for nano-device fabrication, they can be manipulated and assembled by low-cost, bottom-up, techniques (e.g. self-assembly), they are prone to work on flexible, plastic, substrates (see the tremendous efforts on plastic, printable, organic electronics).

The nanoparticle organic memory field-effect transistor (NOMFET) demonstrated here can be programmed to work as a facilitating or depressing synapse; it exhibits short-term plasticity (STP) for dynamical processing of spikes. This behavior is obtained by virtue of the combination of two properties of the NOMFET: the transconductance gain of the transistor and the memory effect due to charges stored in the NPs. The NPs are used as nanoscale capacitance to store the electrical charges and they are embedded into an organic semiconductor (pentacene). Thus, as detailed in section 4, the transconductance of the transistor can be dynamically tuned by the amount of charges in the NPs. We previously demonstrated that this type of device works as a memory but with a “leaky” behavior. The retention times are in the range of few seconds to few 10^3 s. This behavior is used here to implement the synaptic weight wij with a possible dynamic working in this time range, a mandatory condition to obtain the training/learning capabilities of a spiking neural network. A transistor is basically a multiplier, it is used to realize the basic function of the synapse described as $S_j=w_{ij}S_i$, where S_i and S_j are the pre- and post-synaptic signals (here the source/drain current and voltage of the NOMFET). The synaptic weight wij is a timedependent parameter whose value depends on the activity of the pre- and post-synapse neurons. We demonstrate that we can tailor the dynamic behavior of the NOMFET in the frequency/time domain (0.01-10 Hz) by adjusting the size of both the NPs (5-20 nm in diameter) and the size of the NOMFET (channel length L from 200 nm to 20 μm). We also demonstrate that models developed to explain and simulate the plasticity of biological synapses can be successfully adapted to the NOMFET behavior. These results open the way to the rate coding utilization of the NOMFET in neuromorphic computing circuits.

The NOMFET (Figure 11) consists of a bottom-gate, bottom source-drain contacts organic transistor configuration. We immobilized the gold NPs (20, 10 and 5 nm in diameter) into the source-drain channel using surface chemistry (self-assembled monolayers or SAM). They were subsequently covered by a thin film (35 nm thick) of pentacene (see experimental section). This device gathers the behavior of a transistor and a memory. We selected several sizes for the NPs and channel lengths of the NOMFET (200 nm to 20 μm) to study their impact on the synaptic behavior of the NOMFET. We also studied the dependence of this behavior as a function of the NP density. Figure 1 shows scanning electron microscope (SEM) images of NP arrays with NP densities ranging from ~10^{10} cm-2 to a quasi-continuous 2D film.

The density of the NP network is controlled by the density of NPs in solution and the duration of the reaction (see experimental section). The optimum NP density to observe the synaptic behavior reported here is between 10^{11} and 10^{12} cm-2 as shown in Figs. 1-b to 1-d for the 5 nm, 10 nm and 20 nm NPs, respectively. A too low density, <10^{11} NP/cm^2, (fig. 1-a) leads to a too weak memory and synaptic behaviors of the NOMFET. At high density, >10^{12} NP/cm^2, (above the percolation threshold, such as in fig. 1-e), the device does not exhibit a transistor behavior due to metallic shorts between source and drain, and screening of the gate voltage by the metallic film of NPs. In the following, we focus on devices with a NP density of around 10^{11} cm^{-2}. In all cases (figs 1-b to 1-d), we obtained a rather uniform distribution of NPs (no NP aggregation) with a density of ~ 10^{11} NP/cm^2.

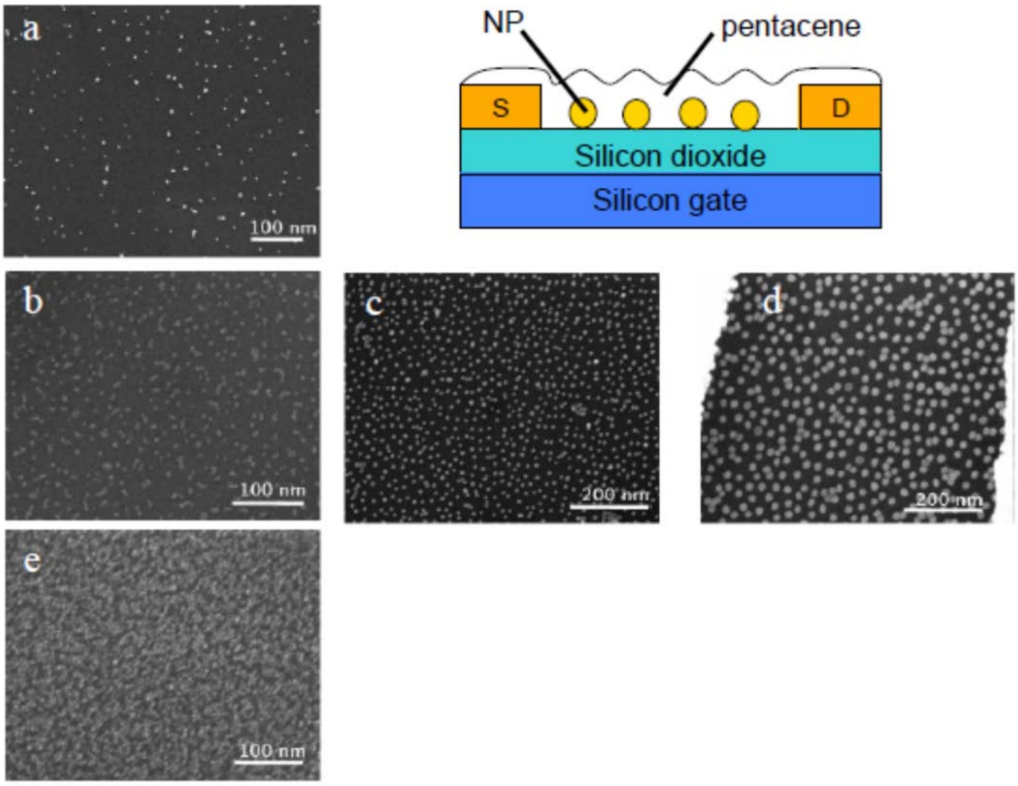


Fig. 11. Scanning electron microscope image of the NP arrays before the pentacene deposition. a) 5 nm NP with a low density of 4.4×10^{10} NP/cm^2, b-d) 5, 10 and 20 nm NP, respectively, with a medium density of 3.7×10^{11}, 1.8×10^{11} and 0.9×10^{11} NP/cm^2, respectively, e) 5 nm NP with a high density $>10^{12}$ NP/cm^2 showing a quasi 2D film with a strong percolation.

Even if we have optimized the process (see experimental section) to got a reproducibility (i.e. a density of about 10^{11} NP/cm^2 without aggregation) of nearly 100% in the case of the 10 and 20 nm NPs (tested on three different fabrication runs), we have observed in few cases the formation of some aggregates (see supporting information, fig. S3). A reproducibility of only 50% is obtained for the 5 nm NPs, mainly due the lack of reproducibility to form the thiol-terminated SAM on the gate oxide. Noteworthy, we always observed the synaptic behavior of the NOMFET, even when the arrays of NPs seem less under controlled. This feature means that the synaptic behavior of the NOMFET is very robust and defect-tolerant, and this feature can be viewed as an important advantage for the envisioned applications in neuronal computing at the nanoscale.

We checked by TM-AFM the morphology of the pentacene layer deposited over the NP networks. Figure 12 shows the pentacene morphology for the reference sample (no NPs) and for the 10 nm NP NOMFET. The film without NPs shows the usual polycrystalline structure of pentacene film with large grain size and terraces, each terrace corresponding to a monolayer of pentacene molecules. The pentacene film is more disordered in the presence of the NPs with a smaller grain size, and the terraces are more difficult to see.

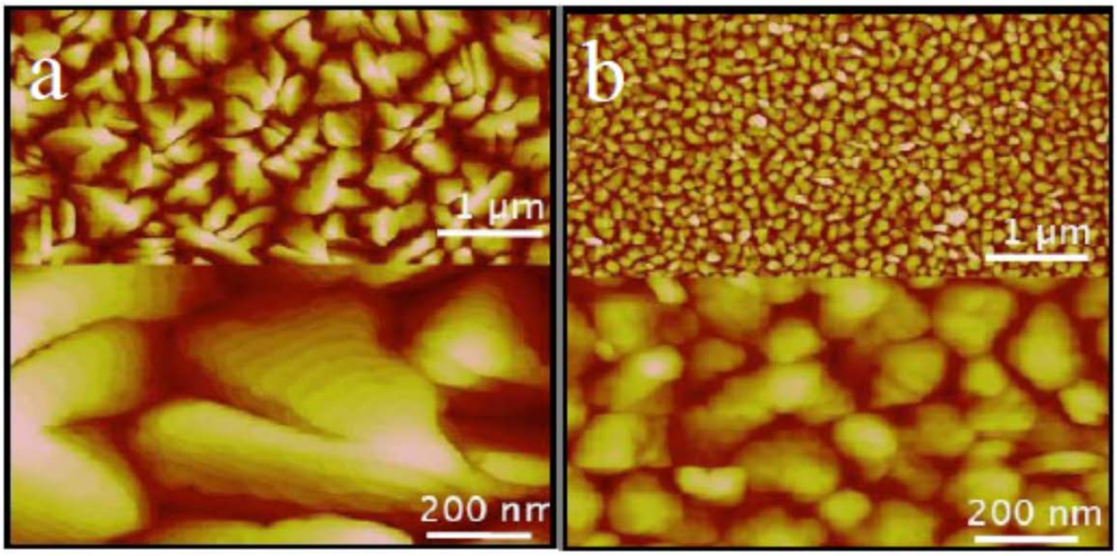


Fig. 12. TM-AFM images of the 35 nm thick pentacene films at two magnifications. a)without the NPs and b) with an array of 5 nm NPs at a density of around 10^{11} NP/cm^2 as in figure 11-b.

This result may be due to both the presence of NPs that hinders the surface diffusion of pentacene molecules during the deposition, and the presence of the organic SAM used to anchor the NPs on the surface. Smaller grain size has been often observed for OFET with a gate dielectric functionalized by a SAM.

This more disordered pentacene film explains the decrease in the output drain current ID and hole mobility µh measured for the NOMFET (~ 10^{-3} cm^{2}V-1s-1) compared to the reference device (~ 0.1 cm^{2}V-1s-1) – see Table 1.

Table 1

Main transistor parameters of the NOMFET with different NP size and of the reference device (no NP). All measurements for a L=5 µm NOMFET. Average values and error bars correspond to the mean and standard deviation of a normal distribution fitted on the experimental data. The threshold voltage shift is the difference between the threshold voltages after and before charging the NP by a negative pulse voltage applied on the gate (-50V during 30s). Note that the small value measured for the reference device is known and due to traps in the pentacene or at the SiO2/pentacene interface (see Ref. 18 and reference therein).

	Reference	**5 nm NP**	**10 nm NP**	**20 nm NP**
Average Id (A)	$1.8\text{x}10^{-4}$ (± $8.5\text{x}10^{-5}$)	$2.3\text{x}10^{-6}$ (± $1.6\text{x}10^{-5}$)	$4.1\text{x}10^{-7}$ (± $7.5\text{x}10^{-7}$)	$3.3\text{x}10^{-7}$ (± $2.8\text{x}10^{-7}$)
Average mobility ($\text{cm}^2\text{V}^{-1}\text{s}^{-1}$)	0.13 (± 0.05)	$2\text{x}10^{-3}$ (± $2\text{x}10^{-3}$)	$7.6\text{x}10^{-4}$ (± $7.9\text{x}10^{-4}$)	$2.0\text{x}10^{-3}$ (± $2.9\text{x}10^{-3}$)
Vt shift (V)	1.1 (± 0.4)	7.5 (± 1.4)	9.4 (± 0.9)	12.8 (± 1.8)

These values are averaged over 6-8 devices, and we notice that the presence of NPs induces a larger dispersion of the FET parameters. The error bars given in table 1 are the standard deviation of a normal distribution fitted on the data. However, we do not observe a specific trend as function of the NP size (see Table 1).

Electronic non-volatile memory technology is actually based on flash memories. A flash memory cell is fabricated using silicon technology and its principle of operation is closely related to the operation of a metal oxide field effect transistor (MOSFET). In a flash cell, a continuous layer of poly-silicon film is buried within the insulator of the MOSFET, separated by a thin insulator, called tunnelling oxide, from the channel area of the transistor and by a thicker insulator, named control oxide, from the gate.

The device operates through the application of voltage pulses to the gate, allowing electrons from the silicon channel to cross the tunnelling oxide barrier and charge the floating gate. Thus, the electrostatic potential of the floating gate screens the electrons of the channel, and the current between source and drain is substantially reduced. To remove the electrons from the floating gate, an opposite polarity voltage pulse is applied, which brings them back to the silicon channel. In this case, we observe that the source–drain current increases again. These two states represent the write and erase states of the device. The reading operation is performed at a gate voltage that will not disturb the writing and erasing states of the device.

Research in scaling flash memory technology is underway using the approach described above or using different approaches based on local storage of the electrons. One important advantage of local storage is the possibility of further scaling the thickness of the tunnelling oxide, as any defect present in that oxide would allow only local loss of stored information, which is not the case for a continuous storage medium, where any defect between the floating gate and the channel of the transistor will result in complete loss of the charge. A nitride layer can replace poly-silicon for local storage. Nanoparticles have been introduced at first time for charge storage by Tiwani et al. at 1996 as another method of local storage of the charge as shown on Figure 13.

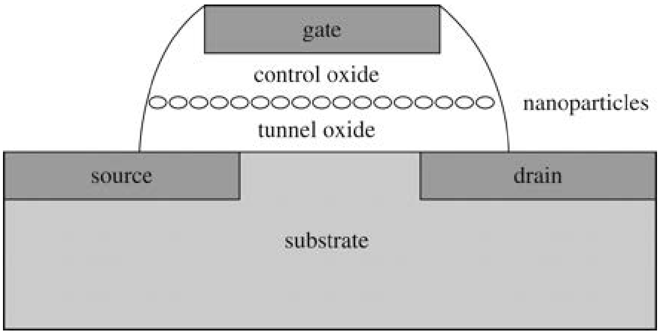


Fig. 13. Schematic of a nanoparticle memory device.

To achieve that goal nanoparticles should be electrically isolated from each other. Silicon and germanium have initially been widely used as nanoparticle materials because of their compatibility with silicon technology. The formation of a two-dimensional layer of nanoparticles is a critical step in the realization of a nanoparticle memory device. For that purpose, two major techniques have been widely investigated: ion-beam synthesis of nanoparticles and chemical vapour deposition. Ion-beam synthesis is performed by implanting silicon at very low implantation energy within a thin silicon oxide layer, followed by annealing at a high temperature. Both approaches have resulted in the demonstration of high-performance devices.

Metallic nanoparticles have been less well explored as charge storage elements in nanocrystal memory devices. In principle, metal nanoparticles should present advantages over silicon nanocrystal memories. The higher electron affinity of several metals compared with silicon allows the potential well of the storage nodes to be engineered to create an asymmetric barrier between the silicon channel and the storage nodes as it clear form Figure 14.

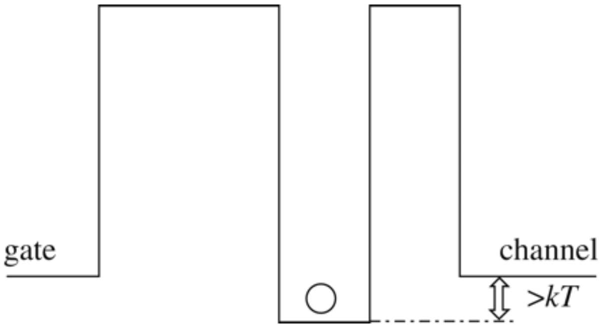


Fig. 14. Energy barriers in the particular case of a metallic nanoparticle with higher work function than silicon (channel area).

This form of barrier makes the write operation easier by creating a higher energy barrier to movement in the other direction by electrons, which favours retention characteristics. During the erase operation, the nanocrystal Fermi level is brought above the silicon conduction band edge, enabling a high erase current to pass through the tunnelling oxide. On the other hand, the high density of states around the Fermi level in a metallic nanoparticle reduces the influence of traps at the nanoparticle/oxide interface, which seems not to be the case for silicon nanocrystals. Quantum confinement determines the conditions for charge retention in the silicon nanocrystals since strongly confined nanocrystal electron states lie at higher energies than the conduction band edge in the silicon substrate and should ease out tunnelling. If, however, experimental data show good retention of silicon nanocrystal memories, this rather suggests that the injected charges fall into interface traps of nanocrystals with the oxide matrix.

Research on metallic nanoparticles for floating gate memories is more limited and more recent than corresponding research efforts on semiconductor nanoparticles. Investigations are based on the deposition of thin metallic films on top of a tunnelling thermal oxide using evaporation. Deposition is followed by rapid thermal annealing (RTA) at 800˚C, resulting in the formation of metallic nanoparticles whose size and density distribution depends on RTA conditions and the thickness of the deposited thin film. The resulting floating gate memory devices prepared using this technique show clear memory windows without any contamination of the silicon substrate, which was monitored by carrier lifetime measurements. This technique will be further explored with future generations of metallic gate MOS devices since it allows more flexibility in nanoparticle energy barrier engineering. A more versatile technique for room temperature deposition of metallic nanoparticles under high vacuum has also been presented recently.

Realizations of memories that are based on the nanoparticle concept discussed above but combining hybrid structures (silicon with organic materials) or purely organic materials have appeared more recently in the literature and will be covered in the next sections of this review. Research on these devices is of an exploratory nature and the purpose is not to replace silicon-based memories but rather to find new areas of application. One important issue for these devices is the possibility of fabricating them at low temperatures, which opens up possibilities for their integration in three-dimensional architecture.

Besides the vacuum technique mentioned above for the low temperature deposition of nanoparticles, in recent years we have also investigated methods using chemical self-assembly. Along these lines, a silicon FET was used in a similar way to flash memory device architecture in order to probe the charge storage properties of gold nanoparticles. The device structure is shown in Figure 15. and the silicon device fabrication process has been described in detail Kolliopoulou at 2003. The memory stack is made up of a 5 nm thermal SiO_2 bottom, numbered (3), a gold nanoparticle layer in the middle (2) and a 54 nm organic insulator on top (1).

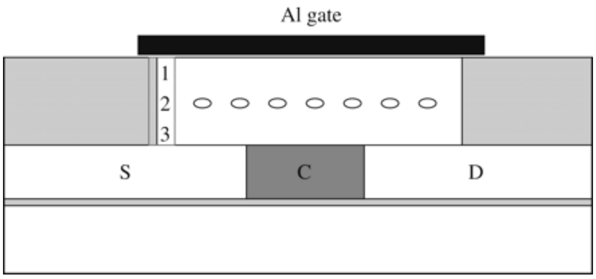


Figure 15. Schematic of a device realized on a silicon on insulator wafer. S and D are the source and drain of the device and C is the channel area. The memory stack is made of a 5 nm SiO_2 bottom (3), a gold nanoparticle layer in the middle (2) and an organic insulator on top (1).

The gold nanoparticles were deposited on the SiO_2 substrates at room temperature by two methods. The first method was the Langmuir–Blodgett (LB) technique using a Molecular Photonics LB700 trough. The subphase was purified water obtained from a reverse osmosis/deionization/UV sterilization system; the film depositions were undertaken at a subphase pH of 5.8±0.2 and a temperature of $20\pm2^{\circ}C$. These nanoparticles were of a nominal diameter of 10 nm and were passivated with tri-*n*-octylphosphine oxide/octadecylamine, which makes the nanoparticles soluble in various organic liquids but mainly insoluble in water; the Q-Au is thus suitable for LB deposition. More processing details can be found in Paul et al at 2003. where a hysteresis in MOS capacitors has been also observed using the above film within the insulator stack.

The second method used for gold nanoparticle layer deposition on the SiO_2 substrate was chemical processing by self-assembly. Using this process, the SiO_2 surface was first functionalized with an amine by placing the wafers into a 10 per cent silane solution (1 ml 3-aminopropyltriethoxysilane in 9 ml toluene) for 1 h in a nitrogen environment. They were then washed in toluene and sonicated in a fresh toluene solution and dichloromethane solution for 2 min each. This process was repeated twice. The chemical reaction that takes place at the oxide surface covers the SiO_2 layer with an amine compound, leaving a $–NH_2$ functionality exposed. The functionalized substrate was dried with nitrogen and held under running ultra-pure water for 1–2 min to encourage charging of the amino groups prior to exposure to the nanoparticles. This surface was then dipped into a solution of carboxylic acid (–COOH) derivatized gold nanoparticles. The gold nanoparticles were passivated with organic ligands. Provided that the pH was adjusted correctly, the acid and amine were mutually attracted. The nanoparticles were therefore positioned at a distance from the SiO_2 surface equal to the length of the amine plus the acid ‘units’, probably 5–6 nm.

Cadmium arachidate (CdA) film was used to cap the layer with gold nanoparticles. Twenty CdA layers, corresponding to a total thickness of 54 nm, were deposited by the LB technique using a Molecular Photonics LB700 trough. CdA films were obtained by spreading arachidic acid (Sigma, 99% purity) on a water subphase containing 2.5×10^{-4} M cadmium chloride (BDH, Analar grade). The deposition pressure for these fatty acid salt films was $22\,mN\,m^{-1}$.

Tests and measurements of the memory characteristics of these devices consist of the successive application of positive or negative voltage pulses on the gate of a previously unstressed device, keeping source and drain electrodes grounded. The voltage pulse height progressively increases while the pre-selected pulse duration is kept constant. The injected (rejected) charges into (out of) the gold nanoparticles cause a shift of the transistor threshold voltage to higher or lower values compared with the unstressed device. The high V_{th} state is usually called the *write* state and the low V_{th} state is called the *erase* state.

The final device exhibits a clear memory window under different gate bias pulses of 1 s duration, as shown in Figures 16 and 17.

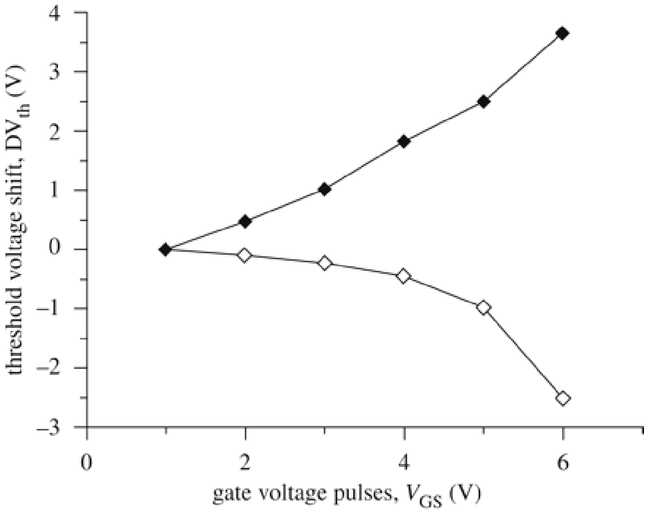


Fig. 16. The effect of the programming voltage on the memory window for a pulse of 1 s using chemical self-assembly as the gold deposition technique. V_{DS}=100 mV, W/L=10 µm/1.5 µm, pulse duration = 1 s. Open diamond, *erase*; filled diamond, *write*.

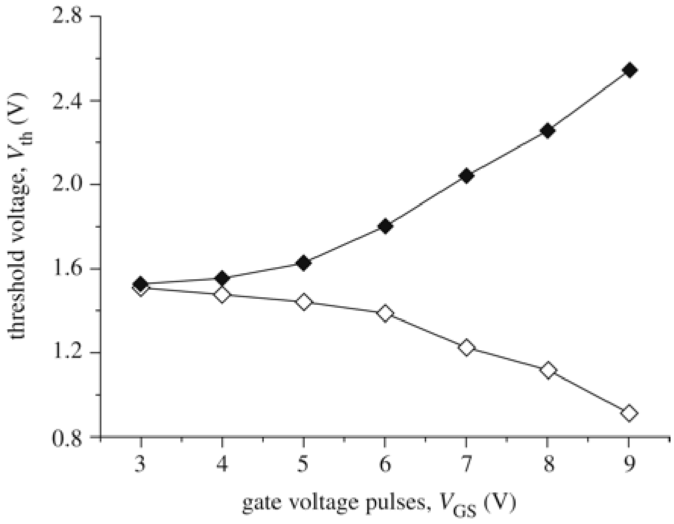


Fig. 17. *Write*/*erase* memory window after application of gate voltage pulses with 1 s duration using LB deposition of gold. V_{DS}=100 mV, W/L=10 µm/1.5 µm, pulse duration = 1 s. Open diamond, *erase*; filled diamond, *write*.

For this pulse duration, the charge exchange takes place between the nanocrystals and the metal gate electrode, leading to counterclockwise hysteresis. The hysteresis direction indicates that electrons are extracted from the nanoparticles for positive gate voltages and injected into them from the gate electrode for negative gate voltages.

Although the thermal SiO_2 layer used is relatively thin (5 nm), for self-assembly deposited nanoparticles the distance between the surface of the silicon and the gold particles is effectively increased to over 10 nm because of the presence of the amine capping layer plus the acid units associated with the gold nanoparticles and for LB-deposited nanoparticles is increased to 7 nm, with 2 nm corresponding to the length of the organic ligand. In the latter case, the application of pulses with a shorter time duration, 300 ms, reveals the effect of nanoparticle charge exchange with the channel (see Figure 18).

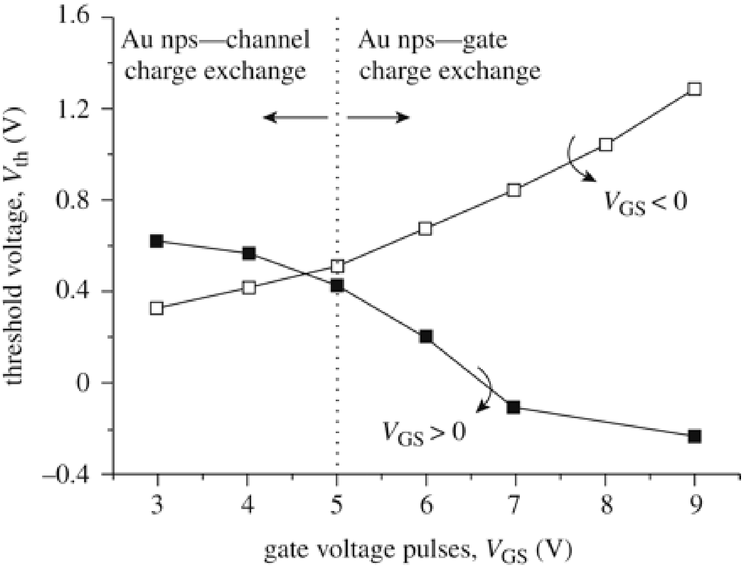
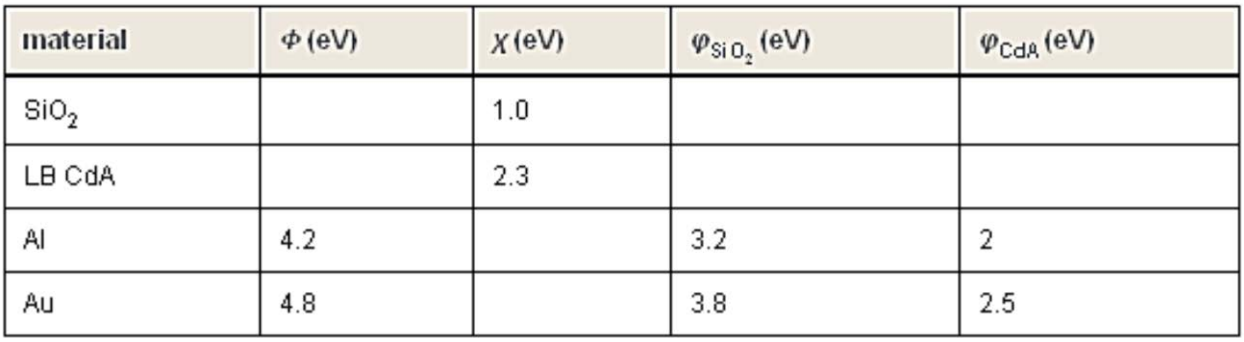


Figure 18. The device memory window for a pulse duration of 300 ms (gold nanoparticles were deposited by LB). $I_{DS}(V_h)$=50 nA.

This charge exchange takes place when the gate voltage does not exceed a value of 5 V. At this voltage the leakage from the organic insulator that separates the nanoparticles from the gate is low but time dependent. After 300 ms, the current density increases, compensating for charges from the channel area. At pulse voltages higher than 5 V, the current density is high enough even for durations of 300 ms or less, so the nanoparticles start to interact with the gate from the very beginning of the pulse application. This enhanced conduction masks the effect of channel charge exchange. The organic insulator breaks down at voltage values higher than those shown in Figure 16 and 17, and therefore no plateau can be reached within the measurement range. We have investigated the insulating properties of CdA in more detail in a previous work Kolliopoulou. We observed that the current flowing through 20 monolayers of CdA is higher than the current through a 5 nm SiO_2 layer. According to works of Nabok at 2002 the conduction mechanism of LB insulating films is due to (i) direct tunnelling through each LB bilayer and (ii) thermally activated hopping within the plane of carboxylic head groups. Table 2 summarizes the electron affinities (χ) and the work functions (Φ) of these materials.

Table 2
Electronic properties of metals and insulators used.



material	Φ (eV)	χ (eV)	φ_{SiO_2} (eV)	φ_{CdA} (eV)
SiO_2		1.0		
LB CdA		2.3		
Al	4.2		3.2	2
Au	4.8		3.8	2.5

Comparison of the electron barrier heights φ leads to the conclusion that transport through the CdA is much easier than through the SiO_2 insulating layer for the same equivalent thickness.

The replacement of the organic material used above as the control insulator by a less leaky one is a challenge for that type of device. Promising new results show the way ahead for high-quality organic insulating materials. Inorganic insulators deposited at low temperature with reduced fixed charge could also be investigated as control oxides. Recently poly(methyl methacrylate) as well as pentacene have been evaluated as organic insulators quite successfully Mabrook et al at 2008.

The planar device fabrication process presented above has also been extended to a three-dimensional architecture that exhibits similar non-volatile memory behaviour with planar devices. This scenario involves the combination of new processes such as wafer bonding at low temperatures, anisotropic etching of silicon for the realization of source–drain areas together with chemical self-assembly and LB deposition of CdA as the organic insulator. The device channel is made of SiGe and the complete process does not exceed the 400°C deposition temperature limit of low temperature oxide (LTO) that plays the role of tunnelling oxide. The complete structure is shown in Figure 19. It is also interesting to remark that in order to realize the three-dimensional memory devices it was also necessary to successfully overcome the problem of LB deposition of a CdA organic insulator over non-planar structures.

More recently, efforts to demonstrate flash-type memory devices that also replace the channel with an organic semiconductor have seen the light of publicity. Below, this research is briefly reviewed, and it is expected that in the coming years other interesting new results will appear in this direction. It appears that there is plenty of room for the investigation of new organic materials not only for use as the semiconductive channel but also for the tunnelling and control insulators.

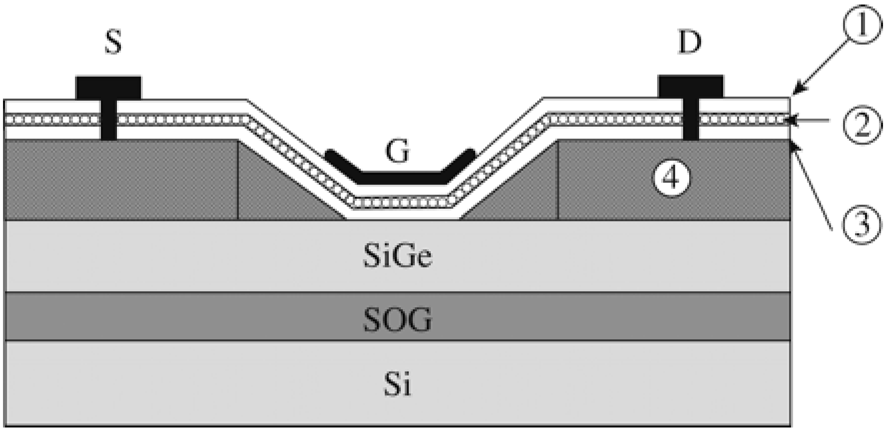


Fig. 19. The schematic of a three-dimensional memory device. Layer SOG stands for spin on glass, a material that is used to bond the two wafers (substrate and active), and LTO is used as the tunnelling insulator. (1), LB insulator; (2), gold nanoparticles; (3), LTO; (4), n^+ Si.

It is important also to report that memory organic FET devices with ferroelectric insulators have been fabricated by Schroeder et al at 2004, but this is not the focus of the current chapter.

The first attempt to demonstrate a nanoparticle-based memory organic FET device was performed byLiu at 2006. Their device, which uses poly(3-hexylthiophene) as the channel material, also uses an organic material (poly(4-vinylphenol)) for the tunnelling barrier but the control oxide is still thermal SiO_2. Another innovation of this fabrication process is the ink-jet printing of metal contacts for the source and drain. The device and its fabrication process sequence are shown in Figure 20.

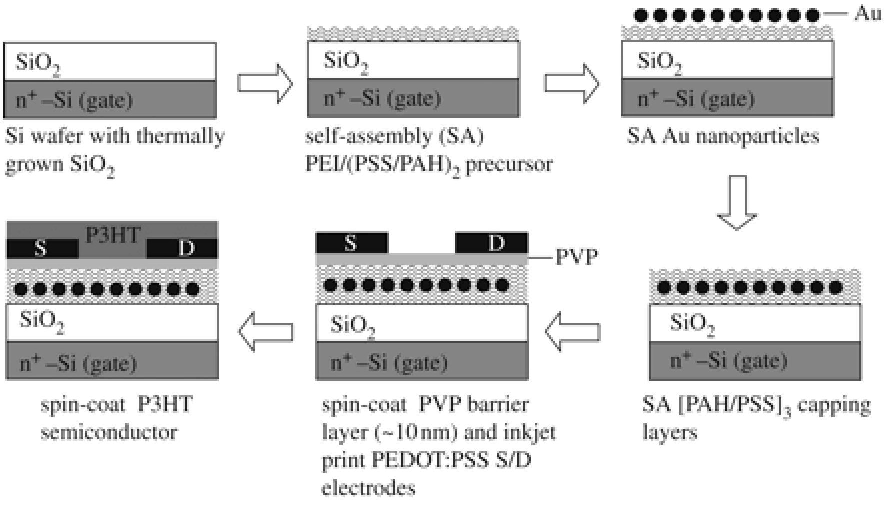


Figure 20. Processing sequence and final three-terminal nanoparticle organic memory (PEI, polyethyleneimine; PSS, polystyrenesulphonate; PAH, poly(allylamine hydrochloride); PVP, poly(4-vinylphenol); PEDOT, poly(3,4-ethylenedioxythiophene); P3HT, poly(3-hexylthiophene); S, source; D, drain).

The effect of the gold nanoparticle layer on the memory performance of the device was clearly demonstrated, although trapping was also partially attributed to the organic insulators used and their interfaces. The on/off ratio of the current is 1500, which degrades after 100 cycles in ambient.

Memory organic FETs using pentacene deposited in a vacuum as the active semiconductive channel over a layer of gold nanoparticles previously self-assembled on a thick SiO_2 layer have also been demonstrated by Novemder et al at 2008. These devices operate at very high voltages owing to the large thickness of the insulator. The effect of nanoparticle contribution to the high current on/off ratio observed after applying opposite polarity voltage pulses to the gate (silicon substrate) was clear. A similar approach has previously been presented by Leoned et al at 2007 using two-terminal devices with a smaller SiO_2 oxide thickness (4.5 nm) that allows gate control at much smaller voltages. The device structure is shown in Figure 21.

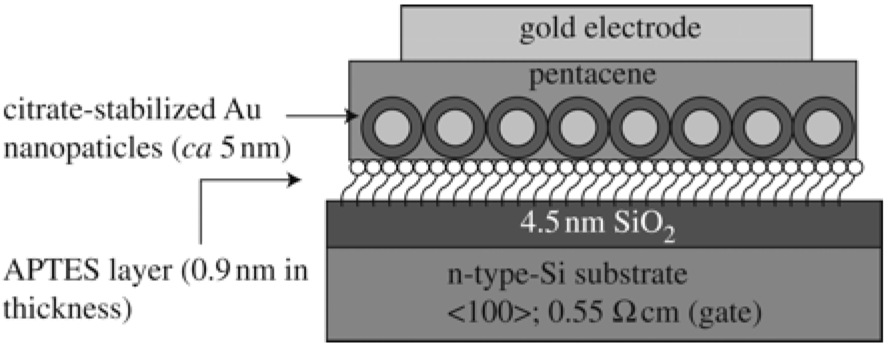


Figure 21. Two-terminal organic device (APTES, aminopropyltriethoxysilane).

The charging of nanoparticles takes place from holes injected from the gold electrode through the influence of the gate (silicon substrate). In both these research works, it appears that charge is trapped in the nanoparticle, although there is no insulator that separates the nanoparticles from the channel (except from the ligand molecules). It is expected that the retention of such devices will be limited. More recently, two-terminal devices using gold nanoparticles synthesized within a block copolymer (polystyrene-block-poly(4-vinylpyridine)) over a pentacene layer demonstrated charge storage following charge exchange between the nanoparticles and pentacene. In that case, the block copolymer acts as the insulating material and retention is good.

For the sake of clarity and comparison with the NOMFET, let us start with a brief description of how a biological synapse works. The most important feature of a synapse is its ability to transmit in a given way an action potential (AP) from one pre-synapse neuron N1, to a post-synapse neuron N2. When a sequence of APs is sent by N1 to N2, the synaptic behavior determines the way the information is treated. The synapse transforms a spike arriving from the presynaptic neuron into a chemical discharge of neurotransmitters (NT) detected by the post-synaptic neuron and transformed into a new spike.

The synapse possesses a finite amount, *U*, of resources: the chemical neurotransmitters. Each spike activates a fraction *aU* (a<1) of these resources and the amplitude *I* of the transmitted spike is a function of this fraction. The fraction of neurotransmitters spent to transmit the information is then recovered with a characteristic time τrec (typically in the range of a second). The response of a synapse to a train of pulses depends on the time interval between successive pulses that determines the amount of available NTs.

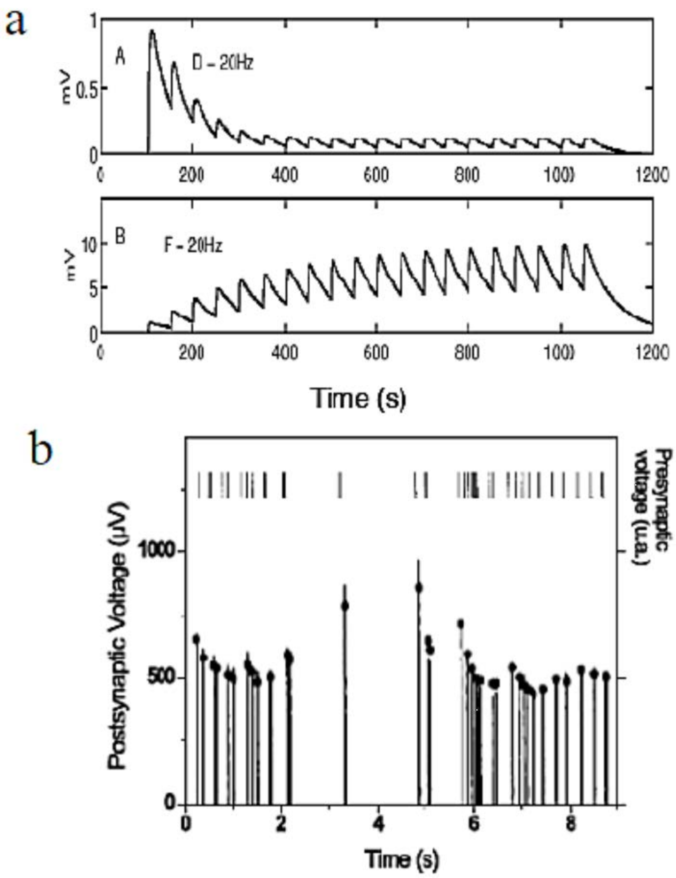


Fig. 22. a) simulation of post-synaptic potential for a depressing (top figure) and a facilitating (down figure) synapse submitted to the same pre-synaptic pulse train at 20 Hz, b) comparison between the frequency-dependent post-synaptic potential response of a depressing synapse (lines) and the iterative model of Varela et al. (dots), adapted as a function of frequency of the pre-synaptic input signal.

Depending on the nature of the synapse, the response to a constant frequency train of pulses can be either depressing or facilitating (see Figure 22-a). Moreover, the biological synapse can process dynamical information when the frequency of the train of pulses is changed (Figure 22-b).

In the case of a purely depressive synapse, a depressing behavior is obtained at a "high" frequency train of pulses (decrease of the NTs available due to the low recovery between each spike separated by a time interval < τrec), while the response of the synapse increases (facilitating behavior) at a lower frequency (the NTs get enough time to recover completely). This property has been extensively studied in biological synapse and is referred to as Short Term Plasticity – STP.

This simple behavior gives to the synapse the main property that is necessary for dynamical processing of information. Varela et al. have developed a simple iterative model to simulate the STP of biological synapses. Based on the work of Magleby et al., they describe the amplitude I of a given spike in the post-synapse neuron by

$$I = \tilde{I} F_1 ... F_n D_1 ... D_m \qquad (1)$$

where the F and D terms are attributed to different facilitating or depressing mechanisms with specific time constants. Thus, the output of the simplest depressing synapse follows $I = \tilde{I} D$, where only one depressive term is considered (this expression is also used by Abbott et al. $\tilde{I}$ is the intrinsic spike intensity delivered by the synapse (i.e. the intensity of the first spike after a long period of rest of the synapse) and D is a dynamical factor representing the depression ($0 < D < 1$).

The response of a synapse to a train of pulses at variable frequency is calculated by an iterative model: at each pulse arriving to the input, the output is calculated by combining the depression factor D and the amount of NTs recovered (with a time constant τrec) between two successive pulses. This quantitative model has shown a good agreement with the biological synapse behavior. Here we present in Figure 22-b the comparison between the biological synapse response and the iterative model of Varela et al. with one facilitating and two depressing contribution. We now demonstrate how the NOMFET can well mimic these typical synapse behaviors. We first express the NOMFET output current taking into account the effect of the charged NPs.

The pentacene being a p-type semiconductor, the NOMFET is active for negative gate voltage only. Such voltage has two effects. (i) As for usual transistors, it creates holes in the pentacene thin film at the interface with the insulator. (ii) It charges positively the gold particles via the organic material. Adapting the percolation theory of Vissenberg and Matters for thin film organic transistors, the drain/source current of the NOMFET is expressed as (see supporting information)

$$G = A_0 e^{-\beta \varepsilon_F} e^{-\beta \Delta} \square \qquad \square(2)$$

where G is the channel conductance, $\beta = \frac{1}{k_B T}$, T - the temperature, A_0 - is a temperature dependent parameter, ε_F - is the Fermi energy fixed by the gate voltage and Δ - is the shift of the Fermi energy induced by the charged NPs. The last term is caused by the repulsive electrostatic interaction between the holes trapped in the NPs and the ones in the pentacene.

The effect of the positively charged NPs is to reduce the current. To do an analogy with the biological synapses, the holes play the role of the neurotransmitters. They modify the output signal, *IDS*, in a way that depends on the numbers of trapped charges via the Δ term: the more charges are stored in the NPs the more the current is reduced. Note that with electrons trapped in the NPs, the result is the opposite: the charges then increase the Fermi energy and the current.

These behaviors correspond to our experimental observations reported elsewhere. Charging NPs with holes by applying a negative voltage pulse on the gate induces a negative threshold voltage shift as measured on the drain current – gate voltage (ID-VG) curves of the NOMFET, and (at fixed VD and VG), a decrease in the current. A positive gate voltage pulse shifts the NOMFET characteristics backward. From these threshold voltage shifts (Table 1) and for the optimized NOMFET with a NP density of around 10^{11} cm^{-2} (such as those shown in Fig. 1-b to d), we estimated the number of charge per NP (for a gate voltage pulse of -50 V during 30 s, measured on a 5 µm channel length NOMFET at VD = - 30 V). They are typically ~2, ~6 and ~15 holes/NP for the 5nm, 10nm and 20 nm NP NOMFET, respectively.

Using these properties, we can mimic the different behaviors of the biological synapses by initially charging the NPs with holes (negative gate voltage) or discharging the NPs (positive gate voltage) depending whether we want a depressing or a facilitating behavior, respectively, before measuring the response of the NOMFET to a train of spikes. The input signal is the drain-source voltage, the output is the drain/source current (the transistor is used in a pass-transistor configuration, source and drain are reversible). The gate has two functions.

In a first phase (see a chronogram of the signals in the supporting information, Figure 2), we used it to program the NOMFET by applying a gate voltage pulse *VP* while source and drain are grounded, and during the working phase, we applied a dc voltage *VG* to maintain the transistor in its “on” regime while an input spiking voltage (between *VD1* and *VD2*) is applied on the drain. Figures 23-a and b demonstrate that the facilitating and depressing behaviors are obtained on the same NOMFET, for rigorously the same application of pulses at the input, by initially programming the device with the gate signal *VP*. For the depressing case (Figure 23-a), we applied $VP > 0$ before to run the device, and $VP < 0$ for the facilitating case (Figure 23-b). In a biological synapse, the facilitating behavior means that an incoming signal with a given frequency and duty cycle induces a post-synaptic signal having an increasing trend, whereas in the case of a depressing synapse, the post-synaptic signal tends to decrease as shown in Figure 22-a.

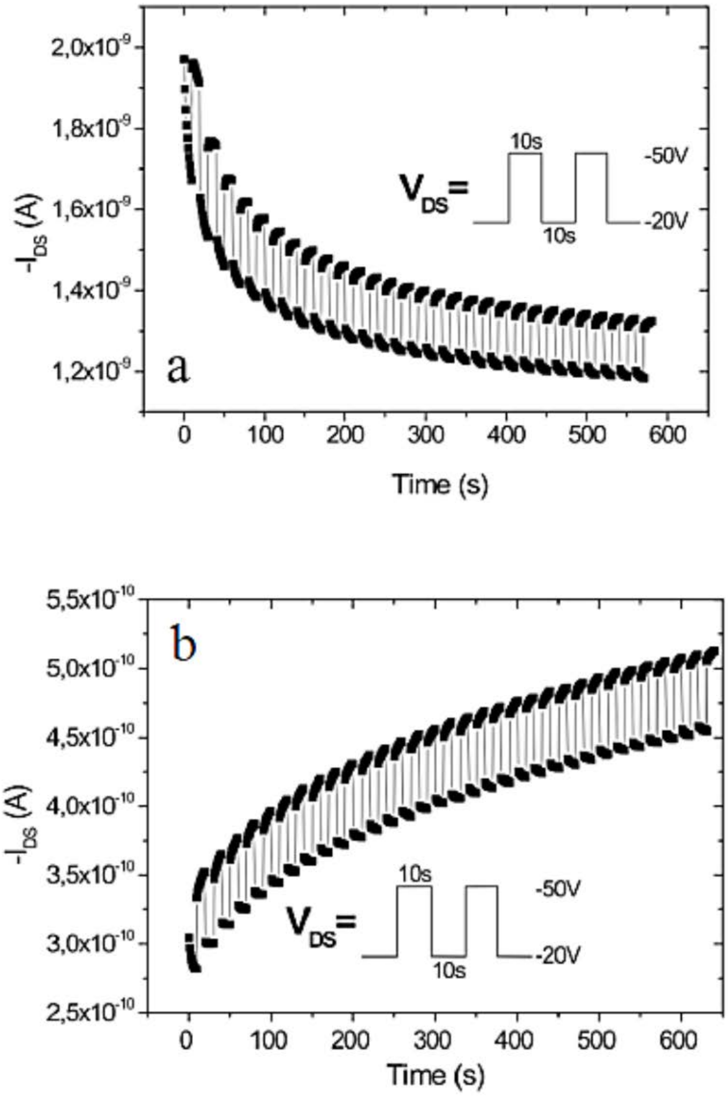


Fig. 23. a) response (drain current) of the NOMFET (5 nm NP, L/W = 12 µm/113 µm) to a constant frequency (0.05 Hz), train of pulses (*VD1 = -20 V, VD2 = - 50V*) after programming the device by a *VP* = 50V (40s) gate pulse to discharge the NPs. The decreasing trend of the output current (in absolute value) mimics a depressing biological synapse. b) response (drain current) of the NOMFET (5 nm NP, L/W = 12 µm/113 µm) to a constant frequency (0.05 Hz) train of pulses (*VD1 = -20 V, VD2 = - 50V*) after programming the device by a *VP* = -50V (40s) gate pulse to charge the NPs. The increasing trend of the output current (in absolute value) mimics a facilitating biological synapse.

We demonstrated in Figure 23 exactly the same behavior for the NOMFET, where the programming negative/positive gate voltage pulse induces the facilitating/depressing behavior, respectively.

Interestingly, the same behavior (Figure 24) is also obtained for NOMFET with a less controlled deposition of NPs (i.e. NPs forming aggregates, Fig. 22 in supporting information). This result means that the synaptic behavior of the NOMFET is robust against process variations.

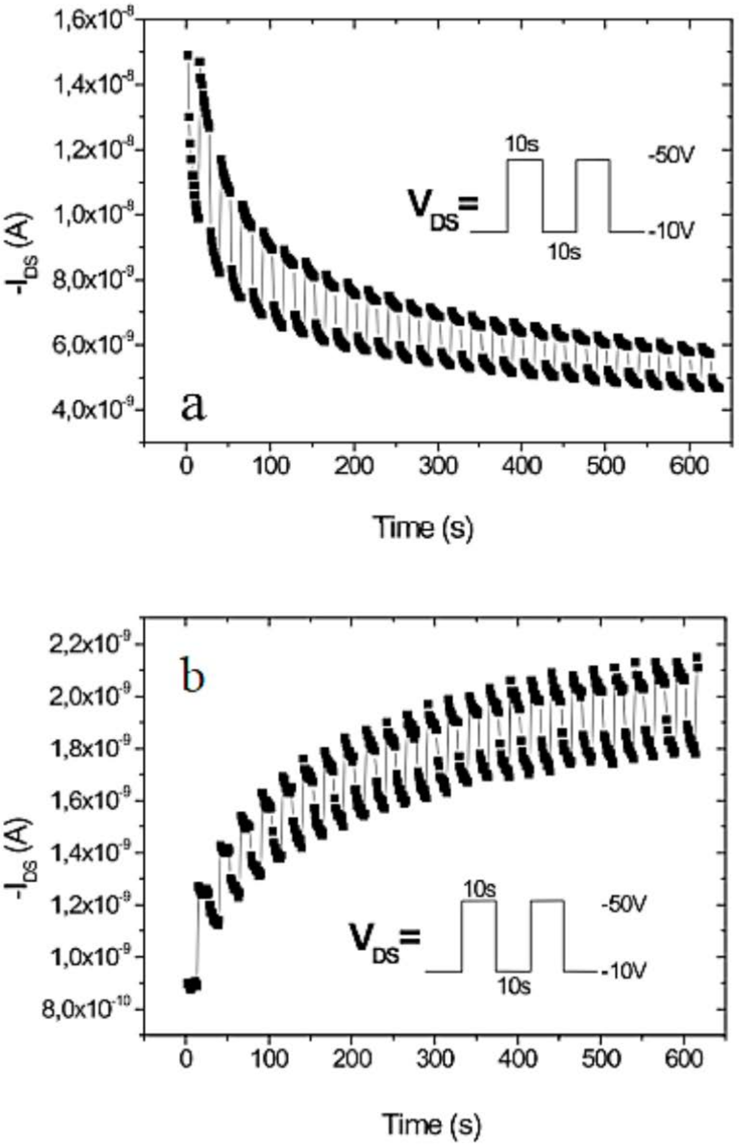


Fig. 24. a) response (drain current) of the NOMFET (20 nm NP, L/W = 12 µm/113 µm) to a constant frequency (0.05 Hz), train of pulses (*VD1 = -10 V, VD2 = - 50V*) after programming the device by a *VP* = 50V (20s) gate pulse to discharge the NPs. The decreasing trend of the output current (in absolute value) mimics a depressing biological synapse. b) response (drain current) of the NOMFET (20 nm NP, L/W = 12 µm/113 µm) to a constant frequency (0.05 Hz) train of pulses (*VD1 = -10 V, VD2 = - 50V*) after programming the device by a *VP* = -50V (20s) gate pulse to charge the NPs. The increasing trend of the output current (in absolute value) mimics a facilitating biological synapse.

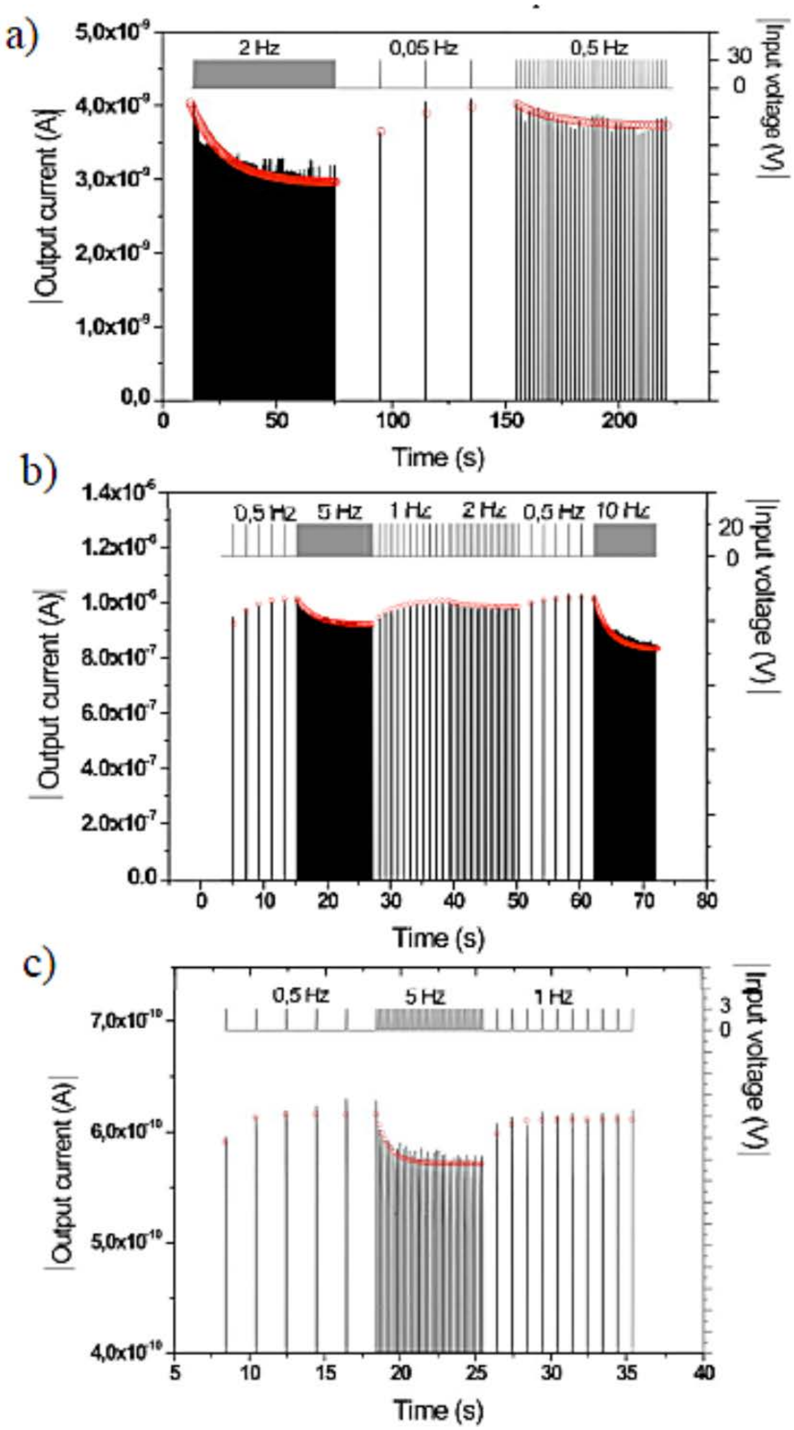


Fig. 25. a) response (drain current) of NOMFET with L/W ratio of 12 µm/113 µm and NPs size of 5 nm to sequences of spikes at different frequencies (pulse voltage *VP* = - 30V). b) response of NOMFET with L/W ratio of 2µm/1000µm and NPs size of 5 nm (pulse voltage *VP* = - 20V). **c**) response of NOMFET with L/W ratio of 200nm/1000µm and NPs size of 5 nm (pulse voltage *VP* = - 3V). In both cases, the red circles correspond to the iterative model calculation (see text), the black lines are the output current measurements. Increasing the frequency evidences a depressing behavior and a facilitating one is observed by decreasing the frequency.

More importantly, we can reproduce the STP behavior without initial programming. We used the NOMFET as a "pseudo two-terminal device". The gate receives the same input voltage (a train of pulses at frequency *1/T, T* is the period, and amplitude *VP*) as the drain/source electrode (supporting information, Fig. 22). We measured the response of the NOMFET to sequences of pulses with different periods, *T* (Figure 25). During such experiments, the NPs are alternatively charged during the pulse time and discharged between pulses. The value of the current at a certain time (i.e. after a certain number of spikes) depends on the past history of the device that determines the amount of charges presents in the NPs.

To illustrate this point, let us consider the NOMFET (L = 12 µm, 5 nm NPs) at the beginning of a particular sequence with period T (Figure 25-a), where the NPs contain some charges.

If T $<<\tau_d$ (τ_d is the NP discharge time constant of about 20 s here, see below), more and more holes are trapped in NPs and the NOMFET presents a depressing behavior as observed for the 0.5 and 2 Hz spike sequences. As expected, the depressing behavior is more pronounced when increasing the frequency. This result is also in good agreement with the behavior of a spiking biological synapse (see supporting information).

Then, for a larger period T (0.05 Hz signal in Fig. 25-a), the NPs have enough time to be discharged between pulses and the sequence presents a facilitating behavior. This feature exactly reproduces the behavior of a biological synapse (Fig. 22-b). The holes trapped in the NPs play the role of the neurotransmitters and the output signal, *IDS,* is a decreasing function of the number of holes stored in the NPs - Figure 25.

At each spike, a certain amount of holes are trapped in the NPs. Between pulses the system relaxes: the holes escape with a characteristic time τ_d. We did not observe such a synaptic behavior for the reference pentacene OFET (no NPs, see supporting information, Fig. 25). Based on eq. (2) for the output current, we adapted the simplest iterative model of Varela et al. to simulate the dynamical behavior of the NOMFET, and we developed an analogous modelisation . A pulse induces positive charges in the NPs: the NOMFET channel conductance is then reduces by a multiplicative factor K < 1. Between each pulses, the NPs tends to discharge with a characteristic time τ_d. The general equation describing the iterative model is (see supporting information):

$$I_{n+1} = I_n K e^{-(T-P)/\tau_d} + \tilde{I}\left(1 - e^{-(T-P)/\tau_d}\right)K \qquad (3)$$

where *In+1* and *In* are the current of the NOMFET at the end of the n+1th and nth voltage pulses sent to the NOMFET, $\tilde{I}$ - is the intrinsic drain current of the NOMFET i.e. the current that would have the device if the charges in the NPs would be kept in the equilibrium configuration (i.e. after a long period of rest, with no charge, or with a constant residual charge induced by the static dc bias of the device), *T* is the period between two pulses and *P* is the width of the pulses.

We fitted the iterative model to simulate the NOMFET behavior (red circle on Figure 25). The same fitted $\tilde{I}$, K parameters and τ_d (here 4.1x10-9 A, 0.9 and 20 s) are used in the three successive sequences shown in Fig. 6-a proving a good agreement between the model and the experiments. For instance, let us consider the system at the beginning of a particular sequence with period *T* (T =0 .5 s or F = 2 Hz in Fig. 6-a), where the NPs contain some charges. Since $T << \tau_d$, the first term of the iterative function, $I_n K e^{-(T-P)/\tau_d}$, is the more important one and the sequence of spikes present a depressing behavior ($I_{n+1} < I_n$). At the opposite, for a larger period *T* (T = 20 s, F = 0.05 Hz), the second term $\tilde{I}\left(1-e^{-(T-P)/\tau_d}\right)K$ becomes the larger one and the sequence presents a facilitating behavior ($I_{n+1} > I_n$). Again, the depressing behavior for T = 2 s (F = 0.5 Hz) is well reproduced by the model. We applied this model to different NOMFET with various W/L ratio (L = distance between electrodes and W = width of the electrodes) down to L = 200 nm. For instance, the fits shown in Fig. 25-b for the L = 2µm NOMEFT correspond to τ_d = *3 s* (*I*□ = 10^{-6} A, K = 0.99) and for a L = 200nm NOMFET (Fig.25-c) to τ_d = *0.9 s* ($\tilde{I}$ = $6.3\text{x}10^{-10}$ A, K = 0.98). We report in Fig. 26 the evolution of τ_d with the channel length and size of the NPs. For comparison, we also plot (for some devices) the time constant extracted from basic charge/discharge measurements (see supporting information, Fig. 26).

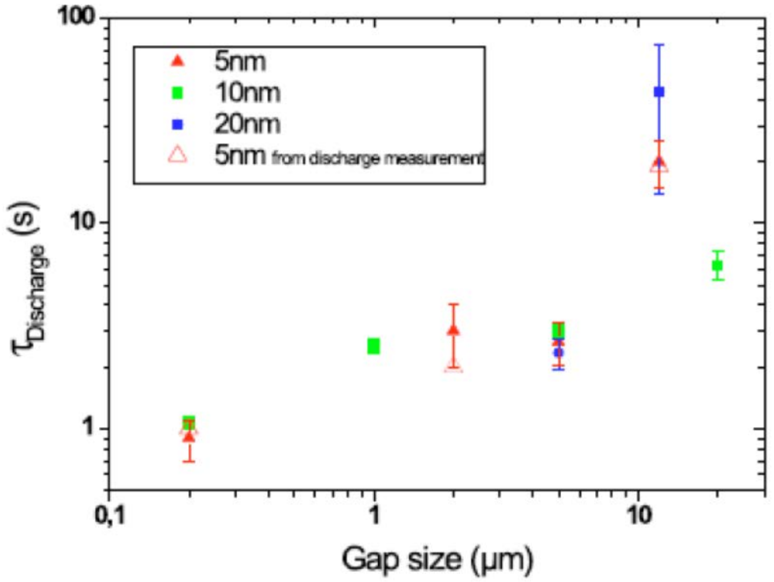


Fig. 26. Evolution of the NP discharge time constant as function of the NOMFET channel length and for the different sizes of NPs. Full symbols are from data extracted by fitting the model on the experiments (as in Fig. 6), open symbols are from discharge experiments.

We note that: (i) the characteristic time from the fitted model and the direct discharge measurement are of the same order of magnitude, (ii) the characteristic time, thus the working frequency range of the NOMFET, can be adjusted by changing the channel length L of the NOMFET, (iii) τ_d is weakly dependent of the NPs size. The RC charge/discharge time constant is roughly governed by the channel resistance of the NOMFET, which scales as L, and the self-capacitance of a NP $C_{self} = 2\pi\varepsilon D$, which scale as the NP diameter D (ε - the dielectric constant), thus scaling down L and D should, in principle, decrease τ_d. This feature is clearly observed for L. At the macroscopic scale of the NOMFET, we have to take into account the total capacitance, which is NC_{self} with N the number of NPs in the channel. This number, while controlled to be around 10^{11} NP/cm2 (see Figure 1), varies from device to device, and we believe that these dispersions can hinder the intrinsic role of the NP size. Moreover, the NPs are capped with surfactant molecules (see experimental section), which can act as a tunnel barrier between the NPs and the pentacene.

It is likely that this tunnel barrier play a role in the discharge phenomena (a thicker tunnel barrier will increase discharge time constant). We did not control the structural quality of this tunnel barrier, which can represent an additional source of dispersion. More detailed experiments, for instance by systematically varying the length of the alkanethiol capping the NPs, will be necessary to increase the control of the charging/discharging phenomena in the NOMFET. A more accurate description of the relationship between these time constants, the size of the NPs and geometry of the NOMFET would require a more sophisticated modelization using a 2D network of distributed RC time constants (taken into account statistical dispersion) rather than a single macroscopic one.

Finally, we note that the iterative model used here to fit the experimental data can be implemented in usual device simulator (SPICE-like) allowing a reliable conception and simulation of hybrid NOMFET/CMOS neuronal circuits. These device/circuit simulations can easily take into account the experimental dispersion of the NOMFET performances to test the robustness of these neuronal circuits against the actual device variation.

Recently, a simple associative memory has been built using a purely CMOS-based emulator of memristor (acting as the synapse) and neurone. In these papers, the memristor-synapse is emulated using a combination of CMOS analog-to-digital converter and microcontroller. It is likely that the NOMFET can be used as the synapse in such associative memory architecture. Such hybrid NOMFET/CMOS neuromorphic computing circuits and architectures are currently under investigation in our laboratory. At this stage, an interesting question can be raised, whether the NOMFET is more related to a memristor device or more related to a memcapacitor as defined by Di Ventra et al.

While a definitive answer would probably require more experiments and simulations of the NOMFET, we believe that the NOMFET is more related to a charge-controlled memristor (or charge-controlled memconductance), since the channel conductance of the NOMFET is history-dependent of the amount of charges stored in the NPs.

So, this experiment it has been demonstrated a hybrid nanoparticle-organic device, a NOMFET, that makes use of the charge storage capability of the nanoparticles and the amplification factor of the organic transistor to mimic the short-term plasticity of a biological synapse. The NOMFET can be programmed to exhibit both a facilitating or depressing behavior. By adjusting the size of the device down to 200 nm and the diameter of the nanoparticle down to 5 nm, we can range the working frequency between 0.01 Hz and 10 Hz. We simulated the synapse behavior of the NOMFET adapting a model developed for the biological synapse. For the NOMFET, by approximating the relaxation function of the NP discharge by a simple exponential, we get the simplest model of this type: our NOMFET behaves within this approximation as the simplest depressive synapse with only one *D*-term.

These results open the way to rate coding utilization of the NOMFET in neuromorphic computing circuits. We can also envision the NOMFET as a building block of neuroelectronics for interfacing neurones or neuronal logic devices made from patterned neuronal cultures with solid-state devices and circuits.

Chapter 16. Flexible Electronic Circuits

16.1. Introduction
16.2. Flexible-circuit technology
16.3. Flexible-Circuit Materials
16.4. Flex circuit Market and Applications
16.5. Manufacturing and Design of Flexible Circuits
16.6 Flexible-Circuit Design
16.7. Flexible-Circuit Technology and Process Trends
16.8. Conclusions
References to Chapter 16.

16.1. INTRODUCTION

Flexible electronics, also known as *flex circuits*, is a technology for assembling electronic circuits by mounting electronic devices on flexible plastic substrates, such as polyimide, PEEK or transparent conductive polyester [1, 2] film. Additionally, flex circuits can be screen printed silver circuits on polyester. Flexible electronic assemblies may be manufactured using identical components used for rigid printed circuit boards, allowing the board to conform to a desired shape, or to flex during its use.

Flexible printed circuits are found in everything from automobiles, VCR's, camcorders, portable phones and SLR cameras to sophisticated military and avionics systems. High-profile applications of flexible circuits are many. One example is the application of flexible-circuit technology in a rigid flex wire harness used on Sojourner, the robot that roamed the surface of Mars collecting data during the summer of 1997.

On a somewhat less romantic level, something as common as a notebook computer or fliplid mobile phone would not be possible without flexible printed circuit technology which allows components to be electronically connected in a dynamic, three dimensional way.

Flexible-circuit technology has a well-established history that goes back nearly one hundred years. Early patent activity highlights the fact that concepts for flexible-circuit materials and designs, which have only come into commercial use within the last few decades, were speculated upon by inventors such as Thomas Edison, Frank Sprague and others in the early twentieth century. The heart and soul of Flexible Printed Circuits (FPCs) are the flexible films and thin layers of conductive circuit traces. These typically constitute the base flexible-circuit laminate, which can be utilised to interconnect electronic devices – such as the LCD screen and keyboard of a laptop – as a reliable wiring replacement, or can have electronic components directly attached to it via soldering or conductive adhesive, to form a finished, pliable circuit board.

Any assessment of the technology of flexible circuits quickly identifies a whole range of benefits that complement and surpass the capabilities of rigid printed circuit boards (PCBs). For many, the technology of flexible circuits and their wide applications may be new, and the view of flexible circuits may be restricted to that of simple point-to-point connections, as a replacement for traditional electrical wire for example. This is currently far from the case and the promise of flexible circuitry is highly significant. With new applications and new materials continually being designed and developed, the technology promises to revolutionise many aspects of electronic circuit design.

16.2. FLEXIBLE-CIRCUIT TECHNOLOGY

The advance of electronic systems into our everyday lives is evidence of a major digital technology revolution. The success stories of the personal computer and the mobile phone serve to demonstrate that consumer and business demand for innovative products are significant. Increasingly electrical and electronic systems are entering our lives in many unanticipated ways. They can be found in our homes in the form of cordless phones and digital TVs, in our cars in the form of hands-free communications and telematics, and in business in the form of notebook computers and mobile personal data assistants (PDAs).

Importantly, and also covertly, within the above applications flexible printed circuits have also been entering our lives. Traditionally employed in the role of wire replacement, removing the need for complex wire harnesses, and replacing costly and increasingly complicated wired assemblies, flexible circuits offer a much simpler and often significantly more cost-effective interconnection method.

However, alongside increasingly innovative applications flexible-circuit technology is branching out significantly from this initial role and it is poised to be a technology that will provide enormous design freedoms for electronic engineers and product designers over the coming years. As the demands of modern electronic systems call for increasing functionality, greater circuit density, higher connectivity, better environmental performance, and all at lower cost, flexible circuitry is poised to deliver on the promise of twenty-first century electronics.

A Definition for Flexible Circuits. Confusion still exists regarding what constitutes a flexible circuit. When asked to envisage a flexible circuit, the image in most people’s mind will be of a bendy printed circuit, typically consisting of a flexible film with a pattern of copper conductors on it. Whilst the image is not far from the truth, in order to better understand flexible circuits it is important at the outset to establish a working definition. The IPC (formerly the Institute for Interconnecting and Packaging Electronic Circuits), through its role of setting standards and guidelines for the electronics industry, has established such a definition: *Flexible Printed Circuit - a patterned arrangement of printed circuitry and components that utilizes flexible base material with or without flexible cover lay* [3] (Fig. 16.1).

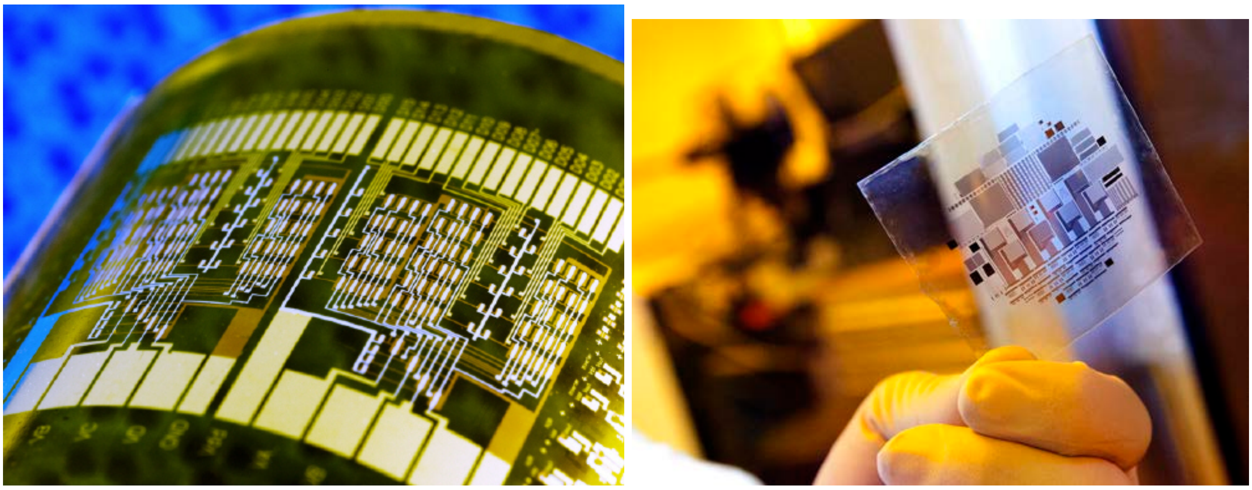


Fig. 16.1. Examples of flexible circuits.

The above definition, although strictly accurate, does little justice to the complexity of the technology but does serve to convey some of the potential given the available variations in base materials, conductor materials, and protective finishes.

Flexible-Circuit Constituents. From the above definition, there are a number of basic material elements that constitute a flexible circuit: a dielectric substrate film (base material), electrical conductors (circuit traces), a protective finish (cover lay or cover coat), and, not least, adhesives to bond the various materials together. Together the above materials form a basic flexible-circuit laminate suitable for use as a simple wiring assembly, or capable after further processing of forming a compliant final circuit assembly.

Within a typical flexible-circuit construction the dielectric film forms the base layer, with adhesives used to bond the conductors to the dielectric and, in multilayer flexible circuits, to bond the individual layers together. Adhesives can also be used in a protective capacity to cover the final circuit to prevent the ingress of moisture and dirt, when they are termed 'cover lays' (also 'cover layers') or 'cover coats'.

Materials Diversity Overview. Many individual materials exist that time and extensive prototyping have proven suitable for application in flexible circuits (Fig. 16.2). There are numerous substrate materials (termed dielectrics) available as very thin films of 12–120 microns in thickness that have been prototyped as base materials upon which to build flexible circuits. However, the two most common dielectric substrate materials are polyester and polyimide. Both are widely available from a number of global sources and both have unique advantages that make them suitable as base materials.

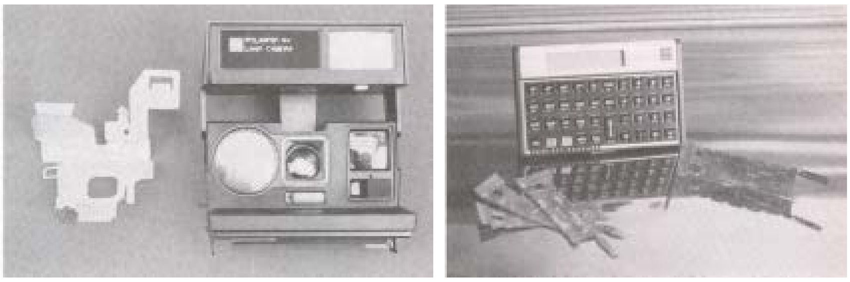


Fig. 16.2. Applications of flexible circuits in consumer goods. FPCs provide interconnection solutions for a diverse range of end-user applications, including (a) the flash and control circuit in an instamatic camera and (b) modern calculators, which consist entirely of low-cost flexible circuitry.

At costs of pennies per square meter, polyester materials are used to provide millions of exceptionally low-cost flexible circuits that find their way into calculators, cameras, touch panels, keypads and automotive dashboards. Polyesters are also highly flexible and are the material of choice for dynamic flexing applications. One example is the connection between a notebook PC keyboard and its screen, an application where many thousands of flexing operations are required.

Polyimide is the material of choice for more demanding flexible-circuit applications. Unlike polyesters, polyimides have excellent high-temperature characteristics and low thermal expansion, which has led to their use being effectively standard practice within the demanding aerospace and defense sector, where complex multilayer circuits are required. They can withstand service temperatures approaching 700 °C.

In the area of conductors, fine metallic foils are used, with copper being the material of choice. Whilst there are more than a half-dozen variations of copper officially recognised by the flexible-circuit industry, each with its own particular characteristics, there are two main classification for copper conductor material. These are electrodeposited copper and rolled-annealed copper.

There are important distinctions between the two classifications of material. Being easier to deposit onto the base substrate, by spraying or sputtering, electrodeposited (ED) copper foils offer the industry low-cost circuitry, whilst rolled-annealed (RA) copper foils, processed between rollers and bonded onto the base laminate, offer high resistance to continuous flexing required of circuits in dynamic applications. New developments are blurring the lines between the various categories of copper, allowing designers continually increasing freedom.

Aside from copper, just about any conductive metal that can be supplied as a foil, sprayed, sputtered or electrodeposited, such as gold, aluminium, nickel or silver, can be used as a conductor.

Also many adhesive systems exist. Common practice is to utilise an adhesive system that offers maximum compatibility with the chosen base material. Hence, polyimide and polyester adhesives are common, as are ‘universal’ adhesives such as acrylics, epoxies, and phenolics, which have migrated from the world of rigid circuit boards.

In conjunction with the basic building blocks many other materials find applications in flexible circuits. It is not uncommon that appropriate stiffening materials – aluminium, steel and moulder polymers – are integrated into the circuitry to provide unique solutions to electrical interconnect problems.

Increasingly within areas such as the automotive and aerospace sectors, where thermal performance is highly important, flexible-circuit laminates are populated with surface mounted electronic components, to form a complete circuit assembly, to overcome the thermal expansion mismatch problem that has plagued the use of rigid circuit boards and surface mounted components.

Research and industry knowledge has demonstrated that during exposure to large changes in temperature, the rigid nature of conventional circuit boards, typically made from stiff FR4 composite, tend to expand at a slower rate than the surface mounted electronic components [4]. Such a mismatch in thermal expansion levels leads to excess stress being generated in the components and their joints, in the worst cases leading to either joint or component cracking, and circuit failure.

On the other hand, a flexible substrate, populated with components, bonded to a rigid base makes for a circuit assembly which though unsuited to use in a dynamic, flexing, capacity exploits the inherently pliable nature of flexible laminates to provide a thermally compliant circuit, with high reliability, at low cost. The area of flexible circuits as compliant, flexible ‘packaging’ for electronic components and assemblies is currently one of the most intense areas of flexible-circuit research.

As stated, many variations on the basic theme exist. Flexible circuits with or without cover lays with or without adhesives, with or without substrates, rigidised and stiffened, are possible. The approach, like the technology, is highly flexible.

Flexible-Circuit Construction. Despite the variability of flexible-circuit materials there is a topology to flexible-circuit construction that follows a number of generic variations. Many of the flexible circuits found in the vast proportion of interconnection and flexible packaging applications follow six basic designs.

*Single-Sided Flexible Circuits***.** Single-sided flexible circuits are the most common types of flexible circuit available. They consist of a single conductor layer on a flexible dielectric film with access to circuit-termination features accessible from one side only. They can be manufactured with or without cover lays and protective coatings, and their relatively simple design makes them highly cost effective. The conductors used can be conventional metal foil, or, for low-cost, polymer thick-film (PTF) ink can be used. This is simply printable conductive ink, loaded with carbon or silver particles, which is directly applied to the flexible substrate in the circuit pattern required by a variety of printing and stenciling techniques.

Single-sided circuits can offer the lowest cost and relative ease of production. Because of their thin and lightweight construction such circuits are best suited to dynamic-flexing or wiring-replacement applications such as computer printers and disk drives. Nearly all of the world’s calculators consist of PTF flexible circuits on polyester film, a combination that offers an exceptionally low circuit cost.

Double Access Flexible Circuits. Double access flexible circuits have been developed to meet the demand for low-cost circuitry that can handle an increase in component real-estate demand. Such circuits allow designers to place components on both sides of the flexible dielectric film.

To access both sides of the film requires careful design to allow exposed areas of the conductor to be available for underside component attachment. This often means punching through-holes in the dielectric film prior to its lamination with the conductor.

Other methods involve after-lamination machining – laser or chemical milling for example – to provide rear access to the conductive layer. Because of the process steps required to produce double access circuitry it is not widely used.

Double-Sided Flex Circuits. The double-sided flexible circuits are also very popular. With the demand to place more components on a circuit and increasing circuit density and power-handling capabilities comes the need for greater conductor numbers. This can be met by incorporating more than a single conductive layer on the same base film.

Double-sided circuits can be constructed by various means such as separate conductors on both sides of the base film and printed conductors separated by printed insulating cover lays. With double-sided circuits an issue is ensuring reliable connectivity paths between components mounted on the top and the bottom of the board. Various techniques have been developed to provide connectivity through such multilayered laminates. Early examples include conductive metal staples, pins and rivets. The most popular flexiblecircuit through-board interconnectivity technique is the plated through hole (PTH), which is also the most popular approach in the rigid-circuit world, from which it has successfully transferred.

In PTH, holes are drilled or laser cut in or through the conductors and base filmlaminate. These holes are then primed and plated with conductive materials to producea reliable interconnect feature.

Sculptured Flex. Sculptured flex is a derivative of flexible-circuit architectures in which a specialised, patented technique is used to yield a conductor layer of varying thickness. The conductive layer is selectively etched back in places to provide thin layers where flexibility is required, and thicker layers for joining and circuit interconnection. It is typical that the thicker layer forms leads that protrude from the circuit, to provide plug in connectors or greater lands for improved solder joint formation. Typically such leads also provide the circuit with improved mechanical strength and rigidity.

Multilayer Flex Circuits. Flexible circuits that have three or more layers of conductors are referred to as multilayer flex. These circuits are complex to construct and have high costs, but they meet designers’, manufacturers’ and consumers’ demands for even greater circuit density. A multilayer circuit consists of bonded conductive layers that are interconnected by means of plated through-holes. Unlike their rigid multilayer counterparts, the individual circuit layers in flexible multilayer circuits may or may not be continuously laminated together, depending upon the flexing and dynamic characteristics required.

Flexible multilayer circuits are popular within the defence and aerospace sectors where they provide dynamic high-density circuits. Their drawback is that with current substrate and conductor materials they are often restricted to a maximum of twenty-five layers. Even with flexible circuit there is a degree of mismatch between the coefficient of thermal expansion of various materials used in their construction, particularly the adhesives. This means that over multiple layers laminate stress can cause throughhole interconnects to barrel and stretch, restricting their reliability.

Rigid-Flex Circuits. Rigid-flex circuits are hybrid constructions consisting of rigid and flexible substrates laminated together. Predominantly, the rigid circuits are used to house the components, whilst the flexible circuitry provides the necessary interconnects between them.

Like double-sided and multilayer circuits they make use of PTH interconnects where required. These types of boards have found particular favour in the defence sector where the combination of reliability, strength and flexibility has not been lost on equipment designers. They are used in a wide variety of commercial microelectronics applications such as laptop computers and notebooks and extensively in the construction of hearing aids.

There are a number of variations of rigid-flex available (Fig. 16.3.) Amongst them is ‘rigidised’ flex which is in effect a flexible circuit which has a stiffening material attached, to support the weight of mounted components and to provide the circuit with some rigidity to aid assembly. Suitable stiffer materials depend upon the application at hand but plastic, composite and metal backing materials are commonly used.

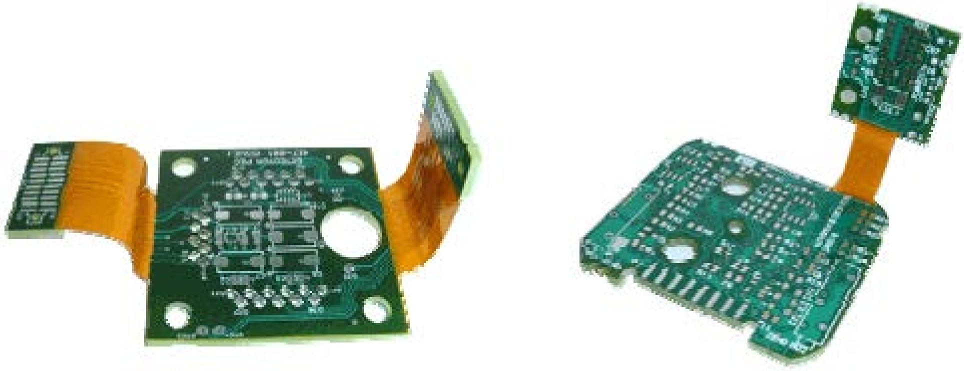


Fig. 16.3. Rigid-flex circuitry. Rigid-flex circuits are a combination of rigid circuitry and flexible interconnects. The interface between the flexible and rigid elements requires careful design, particularly if it is to be subjected to repeated flexing. If this is the case, compliant materials are often applied to the join to reduce the direct flexing of the interconnects as they reach the rigid board.

Beyond the generic variations of flexible-circuit constructions there are a number of alternatives. One such major variation is moulded circuits. These are typically threedimensional moulded plastic components with mechanical capabilities, into which electrical circuitry is incorporated. For some types of moulded circuits the electrical functionality is provided via a flexible circuit that is introduced into the mould at the time of manufacture. Other variations utilise complex moulding and selective plating techniques to form suitable conductor patterns in and on the component.

16.3. FLEXIBLE-CIRCUIT MATERIALS

As discussed initially, flexible circuits typically represent a composite (laminate) of materials, chosen to work together to deliver a desired overall combination of electrical and mechanical performance. However, these criteria are not alone in enabling the determination of an appropriate combination of circuit materials. Other typical factors that play heavily on the selection of suitable materials are:

- Application environment
- Volumes
- Reliability requirements
- Dynamic or static flexing required
- Duration of flexing or dynamic operation (cycles)
- Additional electrical requirements of the circuit – e.g. impedance
- Connections to components and other circuitry
- Method used for component assembly
- Costs

Material Configuration. Within a typical flexible circuit, four distinct classes of materials are used:

- Base material (dielectric film/flexible substrate)
- Conductors (foil or conductive coating)
- Adhesive (optional)
- Cover lay (film or coating)

Other materials utilised include numerous surface finishes and anti-tarnishing coatings, and integral stiffeners or backing substrates, all designed to give additional properties and performance capabilities to the circuit assembly or enhance the ease of manufacture. A simple typical circuit make-up is detailed in Fig. 16.4 below.

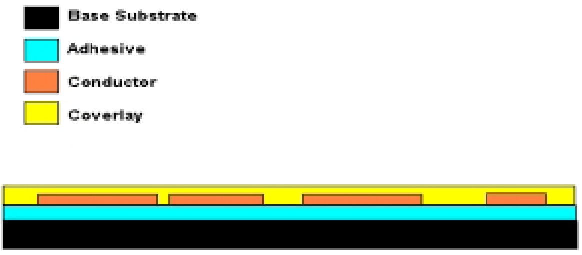


Fig. 16.4. Flexible-circuit make-up.

Base Materials. A suitable base material has to perform a variety of important functions. It must electrically insulate the conductive circuit tracks from one another and it must be compatible with any adhesives used for conductor or cover-lay bonding. Under normal circumstances the base material will also provide the circuit with much of its mechanical characteristics, such as its flexing strength and durability. In the case of adhesiveless laminates, which will be discussed later, the base substrate provides all of the circuit's strength.

Typically, the major criteria and properties required of a suitable substrate are:

- High dimensional stability
- Good thermal resistance
- Tear resistance
- Good electrical properties
- Flexibility
- Low moisture absorption
- Chemical resistance
- Low cost
- Consistency from batch to batch
- Wide availability

A singular benefit of flexible-circuit substrates is that unlike their counterparts within the rigid-circuit world they are not restricted to processing in sheet form. Many manufacturing processes for flexible circuits take advantage of the nature of the materials used in their construction to manufacture circuits in a continuous roll-to-roll fashion. Base materials such as polyester are supplied in roll form and processed as a single web, a metre or so wide. The processed circuits (minus components) can then be rolled up for further processing and component attachment, or die-cut from the web to the final circuit shape in the case of simple wiring interconnects.

Many of the substrate materials used in flexible laminates are themselves cast or produced from web-based processes, making their supply in roll form straightforward and highly economic. A significant proportion of manufacturers have enabled their manufacturing stages – plating, cleaning and rinsing amongst others – to use roll-to-roll techniques.

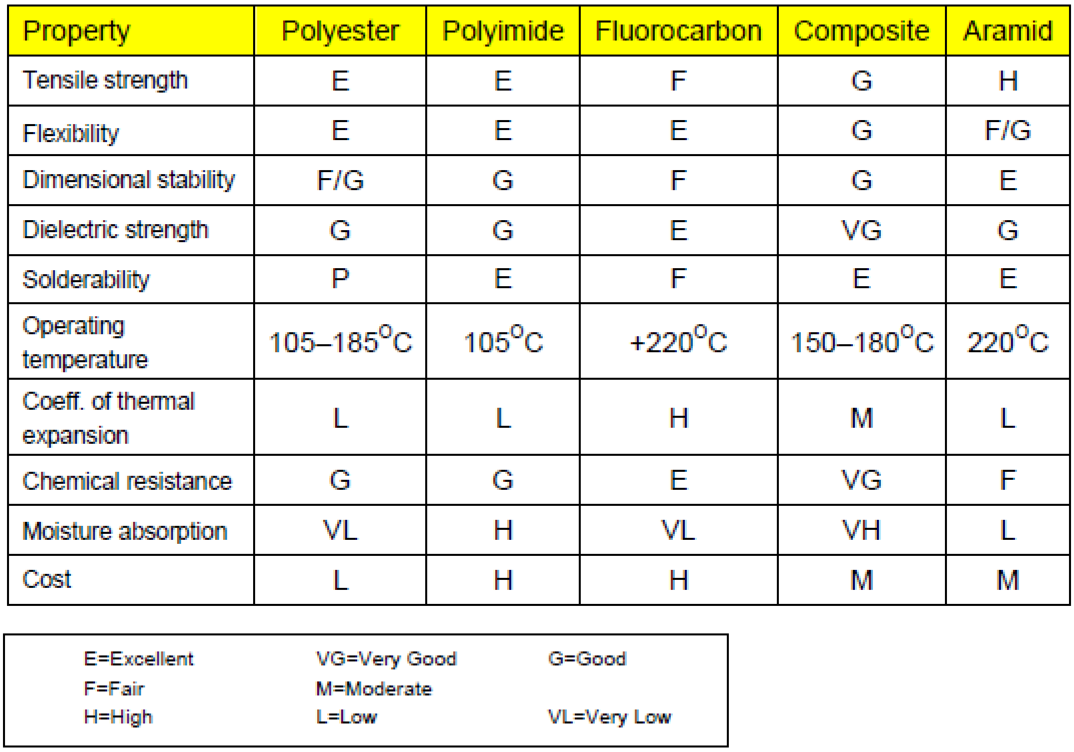
Roll-to-roll processing poses distinct technical challenges such as ensuring accurate layer registration and correct web tension. However, there is no doubt that compared to the conventional press lamination process, which uses rigid sheets, it offers higher production rates of thousands of circuits per hours, improved reliability and improved material handling.

There are many materials that have been used as substrate materials with varying degrees of success over recent decades. Materials that have been widely used for various FPC applications include:

- Polyimide
- Polyester
- Fluorocarbon films
- Aramid papers
- Composites

These materials present a range of differing properties (see Table 16.1), which are called upon by circuit designers where their blend of electrical and mechanical performance capabilities and costs best suit the application at hand. As previously discussed, the two materials that receive by far the most attention are polyimide and polyester films.

Table 16.1. Typical properties of dielectric materials.



Property	Polyester	Polyimide	Fluorocarbon	Composite	Aramid
Tensile strength	E	E	F	G	H
Flexibility	E	E	E	G	F/G
Dimensional stability	F/G	G	F	G	E
Dielectric strength	G	G	E	VG	G
Solderability	P	E	F	E	E
Operating temperature	105–185°C	105°C	+220°C	150–180°C	220°C
Coeff. of thermal expansion	L	L	H	M	L
Chemical resistance	G	G	E	VG	F
Moisture absorption	VL	H	VL	VH	L
Cost	L	H	H	M	M

E=Excellent	VG=Very Good	G=Good
F=Fair	M=Moderate	
H=High	L=Low	VL=Very Low

Polyimide. Polyimide films are manufactured from a condensation polymer of an aromatic dianhydride and an aromatic diamine. The film is often cast in various thicknesses of 7.5–125 microns.

Polyimides are non-flammable thermoset polymers and as such do not exhibit a softening or melting point. However, unlike many thermosetting plastics, polyimide films are highly flexible, exhibit good flexing and electrical properties across a wide range of temperatures, have good high temperature resistance, and resist soldering conditions.

Polyimide films are produced in several varieties by companies such as DuPont (USA), Ube Chemical (Japan) and Kaneka (USA) under trade names such as Kapton, Apical and Upilex. The standard use of polyimide substrates is significantly more prevalent in Japan than in other regions of the flexible-circuit manufacturing world.

Polyimide is the flexible-circuit dielectric of choice for about 80–85% of applications, as reflected in the materials relative FPC market volume sales. Whilst they are relatively expensive, up to nine times the cost of polyesters, their tough nature and resistance to thermal or chemical damage make them an obvious choice where circuitry is required with a high degree of reliability and immunity to environmental influences.

A downside to polyimides is the fact that they readily absorb moisture, typically up to 3% by weight. This requires the material to be thoroughly dried prior to processing, and requires monitoring of the material as it passes through various production stages to ensure that the chances of further moisture uptake are minimised. It is also conventionally the case that polyimide flex is produced via a panel process, with the material being bonded to a conductive substrate via a number of methods. This means that polyimide flex is provided in sheet form, negating the processing advantages achievable by roll-to-roll processing. However, developments are underway to develop roll-processable polyimides.

Polyester. Polyesters, also known as polyethylene terephthalate (PET), are similar in most ways to polyimides but because they have a melting point of 250ºC and a glass transition temperature of 80ºC, they generally have lower heat resistance that precludes their use in assemblies that require the use of soldering. However, many OEM manufactures have developed in-house processes that enable them to solder to various grades of polyester with high degrees of success.

Due to their lower raw-material cost and ease of roll-to-roll manufacture, polyesters are found in a high proportion of high-volume, low-cost, environmentally undemanding applications such as calculators and VCRs. In these applications they are often used as simple wire replacements or as rigidised assemblies with surface mounted components. Despite its low temperature resistance many companies have explored techniques for mounting components onto polyester. A commercially proven approach is through the use of conductive adhesives.

Polyester films are the preferred dielectric material for the remaining 20% of the market that polyimide has not captured. Polyester films are generally available in thicknesses of 25–125 microns. They are characterised by their excellent flexibility, good electrical properties, and high chemical and moisture resistance. Given the recent advances in overcoming the poor temperature resistance of polyesters, the material looks set to approach even closer to becoming the ideal flexible-circuit substrate. Developments are well under way with various polyester formulations, such as polyethylene naphthalate (PEN), which offer the increasingly higher service temperatures suitable for soldering. Overall, polyester is easily modified at low cost. It can be readily drilled, punched, embossed, thermally formed, coated and dyed

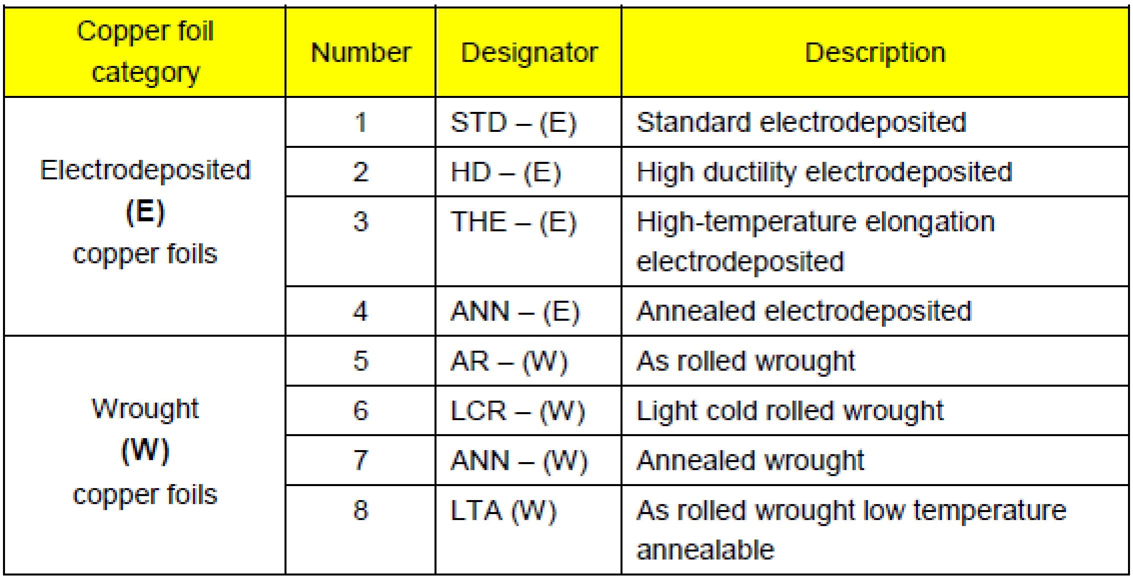
Conductor Materials. Material considerations for FPC conductors are similar to those of rigid circuit boards. The conductor material must survive processing and provide adequate electrical and mechanical performance in the application environment. The list of conductor candidate materials includes elemental metal foils, such as copper and aluminium, and metal mixtures including stainless steel, beryllium-copper, phosphor-bronze, coppernickel and nickel-chromium resistance alloys. Both silver and carbon polymer thick-film (PTF) inks are also used.

Copper is the material of choice for flexible-circuit conductors. In practice, of the wide variety of possible conductor materials, only a selected few have found use within volume applications. As well as providing the electrical connectivity and electrical performance features of flexible circuits, conductor properties greatly influence the fatigue life, stability, and mechanical performance of FPC assemblies. In many static applications bending is limited to installation and general servicing. In dynamic applications, the assembly is flexed or folded repeatedly during normal use. As a general rule, for dynamic applications, conductors should be of the minimum acceptable thickness and their material of construction must be carefully chosen, along with their grain orientation and deposition technique, to match the performance levels required.

Copper. The relatively low cost of copper, its high workability, good plating and good electrical characteristics make it an excellent material for flexible-circuit conductors. It is also the case that there are several different kinds of copper available, which can be matched by the circuit designer to specific applications.

Copper foils suitable for use in flex circuits typically fall into two categories – electrodeposited (ED) or wrought (W). The IPC standard IPC-MF-150 (‘Metal Foil for Printed Wiring Applications’)3 details these categories and defines four types of copper within each, giving flex designers eight types of copper to select from. Table 16.2 is a guide to the common grades of copper available.

Table 16.2. Summary of IPC-MF-150 copper foil categorization [5].



Copper foil category	Number	Designator	Description
Electrodeposited **(E)** copper foils	1	STD – (E)	Standard electrodeposited
	2	HD – (E)	High ductility electrodeposited
	3	THE – (E)	High-temperature elongation electrodeposited
	4	ANN – (E)	Annealed electrodeposited
Wrought **(W)** copper foils	5	AR – (W)	As rolled wrought
	6	LCR – (W)	Light cold rolled wrought
	7	ANN – (W)	Annealed wrought
	8	LTA (W)	As rolled wrought low temperature annealable

Electrodeposited copper (E) is recommended for use in applications where the need for dynamic flexing is minimal. The grain structure of E copper consists of vertical grain boundaries that extend through the material’s deposited thickness. These grain boundaries allow cracks to propagate through the material very quickly causing E copper’s electrical and mechanical performance to fall dramatically.

Wrought (W) copper, sometimes referred to as ‘rolled and annealed’ or ‘RA’ copper, is the material of choice for flexing applications. Unlike the vertical grain structure of E copper, W copper is produced by heating and mechanically rolling ingots of pure copper to the desired thickness through rollers. Whilst this places restrictions on the ultimate thickness of the foil that can be economically rolled, typically 18 microns, the process produces a grain structure that resembles overlapping plates. This plate-like structure has a significantly longer crack propagation path and gives W copper a higher tensile strength and a much higher resistance to repeated bending.

Other differences between the two main forms of copper are that typical E coppers have a lower conductivity and more pinholes and inclusions than W coppers. However, E copper foils can be readily heat-treated to form improved grain structures, and they generally have a more uniform response to surface treatments which can significantly improve their adhesion to base laminates.

Whilst the copper options defined in IPC specifications provide scope for the design of a wide variety of copper circuits, a number of alternative copper platings are unaccounted for. These include vapour-deposited, sputtered and electroplated copper, each of which has its own relative merits and pitfalls.

Electroplated copper, using either electroless or electrolytic plating techniques, allows copper to be directly plated onto the base laminate material. It is distinguished from electrodeposited copper by its ‘as plated’ properties, which are significantly different from the properties of E copper. The plating process allows greater control over the materials grain structure, which, given the addition of suitable additives, can produce a layer of material that has an amorphous or equiaxed grain structure resulting in properties that are superior to W copper.

Sputter copper films are similar to electroplated foils but allow the application of a very thin copper layer (typically less than 1 micron) that can be used as a seed layer for an additive electroplating process. Here copper is added to the laminate in specified areas to give the level of circuit thickness required. The very thin layers produced by sputtering have been proven to be beneficial in the fabrication of very fine-lined circuits, which find applications in specialist areas such as low-temperature (cryogenic) circuits.

Very fine-lined circuits are also of benefit in highly dynamic circuit applications such as computer disk heads, where hundreds of thousands of dynamic flexing operations are required.

Other Metal Foils. As stated, where the occasion presents itself, circuit manufacturers have utilized metallic foils other than copper. For example, aluminium foil has found use in circuits where low costs and weight reduction are driving factors, or where electric shielding is required. Unfortunately, aluminium cannot be processed with conventional soldering equipment and so, to keep costs down, is often used in circuits that employ conductive adhesives to form connections.

Metal mixtures, such as phosphor bronze and beryllium copper foils, have been employed where a combination of reasonable conductivity, good mechanical strength and spring-like mechanical properties has been required. Whilst typical beryllium copper offers only a quarter of the conductivity of copper, such material also provides improved corrosion resistance in electrical contacts. Table 16.3 shows the fundamental properties associated with some of the candidate conductor materials.

Table 16.3. Conductor material properties [6].

Property	Aluminium	Copper	Gold	Iron	Nickel	Silver
Resistivity ohm-cm x l0^{6}	2.8	1.7	2.4	10.0	6.8	1.7
Density oz/ft^{2} @ l mil	0.22	0.74	1.6	0.64	0.74	0.87
Harness Brinell	15	42	28	80	110	95
Thermal conductivity Cal/ sec/cm^{3}/°C	0.48	0.92	0.70	0.16	0.14	0.97
Coeff. of thermal expan. (TCE) ppm/°F x l0^{-5}	1.3	0.93	0.79	0.51	0.76	1.05

Polymer Thick-Film Processes. A further method for generating conductors on FPC base films is by screen printing or stencilling conductive inks onto polymer films to directly create circuit traces (Fig. 16.5). The method is commonly referred to as the polymer thick-film (PTF) method. PTF has been used for decades as the world's most cost-effective and successful fully additive, waste-free circuitry and assembly technology. Computer keyboards, hand calculators and telephones are key examples of high-volume goods that benefit from this economical method of circuit manufacture.

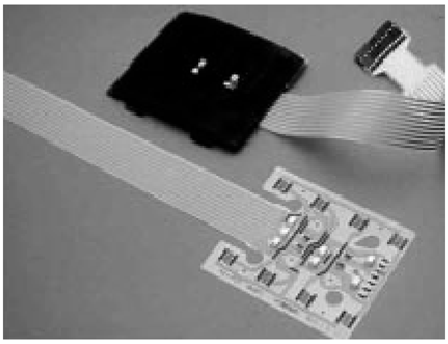


Fig. 16.5. PTH keyboard.

The most common substrate for PTF is thick polyester film of 3–5 microns thickness, but almost any non-conductive flexible film or rigid non-conductor can be used. Films are often heat-stabilised by the film producer or the flexible-circuit manufacturer so that very little shrinkage will occur at ink- and adhesive-processing temperatures of up to 150°C [7].

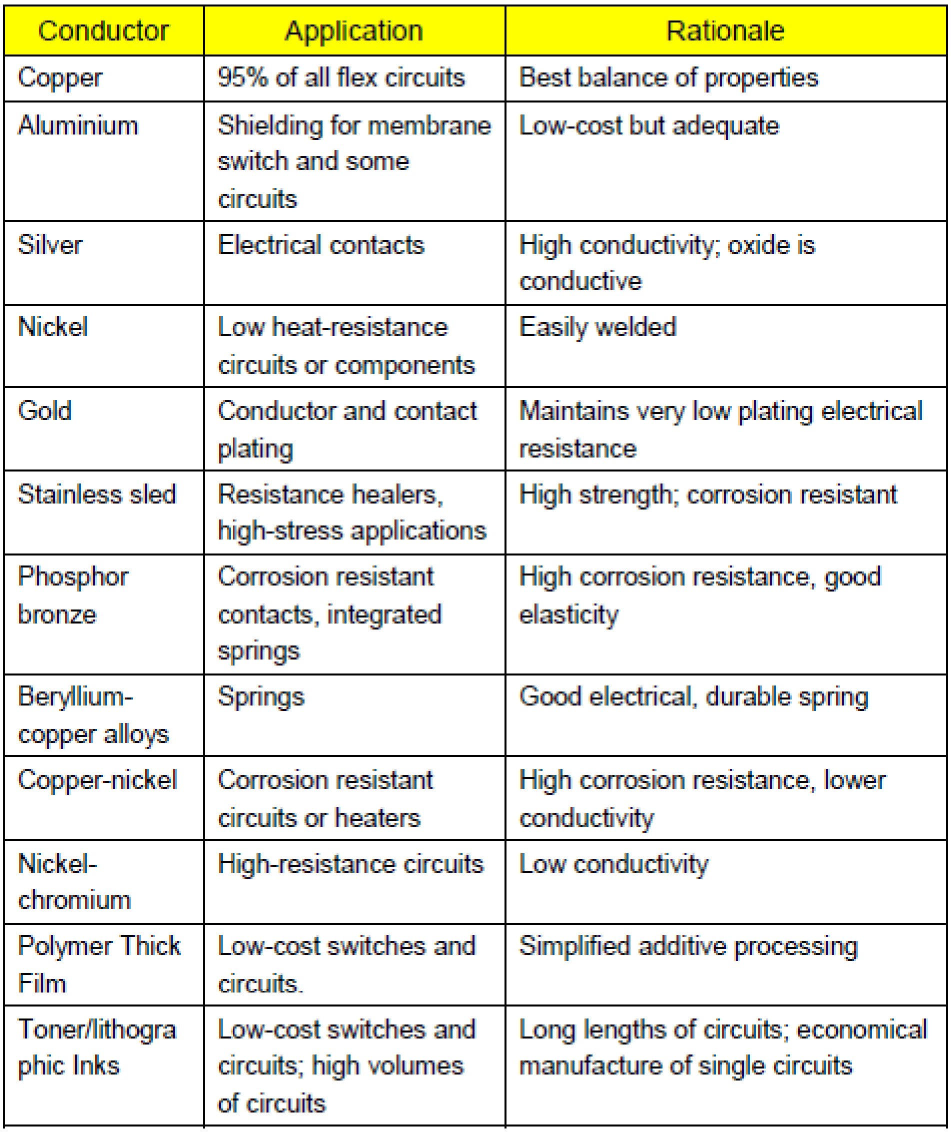
PTF inks consist of a mixture of a polymer binder, such as polyester, epoxy or an acrylic, and a finely granulated conductive material such as silver or resistive carbon. Alternatively a blend of both silver and carbon may be used. The ink is applied directly to the base laminate without the use of a resist, often in a single-step operation. This capability makes PTF a very low-cost process and it is considered more environmentally friendly because it negates the need for laminate stripping, cleaning and treatment, as would be required with other ways of applying conductors.

Aside from the environmental benefits, PTF inks have been demonstrated to have a flex life that is similar to that of copper foil of equal thickness. The most commonly used metal for filling PTF inks is silver. Whilst silver metal has one of the highest conductivities of all metals, the conductivity of silver-filled polymer inks is relatively low, six to sixty times more resistive than copper metal (7–50 milliohms per square at 25 microns thickness versus approximately 0.9 milliohms per square for 25 microns copper). PTF inks are also difficult to solder to using conventional means. Connections with PTF circuitry is usually made via pressure contact using a conductive adhesive, as is the case with keyboards and touch pads.

The relatively low-cost and versatility of the PTF approach, combined with an upsurge in the use of conductive adhesives for circuit interconnection, has made PTF a much researched technology.

Emerging Direct Apply Technologies. The benefits of PTF technology have spurred much research in the development of direct conductor application techniques. Emerging technologies that are considered good candidates for commercial applications include a number of printing techniques transferred from other industry sectors (Table 16.4).

Table 16.4. Typical Conductor Applications



Conductor	Application	Rationale
Copper	95% of all flex circuits	Best balance of properties
Aluminium	Shielding for membrane switch and some circuits	Low-cost but adequate
Silver	Electrical contacts	High conductivity; oxide is conductive
Nickel	Low heat-resistance circuits or components	Easily welded
Gold	Conductor and contact plating	Maintains very low plating electrical resistance
Stainless sled	Resistance healers, high-stress applications	High strength; corrosion resistant
Phosphor bronze	Corrosion resistant contacts, integrated springs	High corrosion resistance, good elasticity
Beryllium-copper alloys	Springs	Good electrical, durable spring
Copper-nickel	Corrosion resistant circuits or heaters	High corrosion resistance, lower conductivity
Nickel-chromium	High-resistance circuits	Low conductivity
Polymer Thick Film	Low-cost switches and circuits.	Simplified additive processing
Toner/lithographic Inks	Low-cost switches and circuits; high volumes of circuits	Long lengths of circuits; economical manufacture of single circuits

The company Extended Length Flex Technologies have patented a technique for producing long lengths of flexible circuits through the use of catalytic toners which can be electrostatically deposited onto a substrate material by means of a modified laser printer. The toner pattern produced can then be electroplated over with copper to produce an additive circuit-production method.

Other direct application methods include lithographic printing and using conductive or resistive inks. The process employs standard offset lithographic printing – used to produce books and magazines – to deposit electrically conductive traces onto a wide range of flexible media. Lithography ink deposition lends itself to high-volume production; a typical lithographic machine can perform between six thousand and ten thousand impressions per hour, and line thickness can be in the order of 3 microns, which compares favourably with the 50-micron line thicknesses achieved by screen printing of low-cost circuits [8].

Adhesives. Adhesives play an important role in flexible circuits. They are used to provide a secure join between the substrate and the chosen conductor material, to join circuits together where multi-layer or rigid-flex constructions are required, and to provide a protective cover lay over exposed conductors once they are formed.

There are a number of methods for applying adhesives but generally they are coated onto the dielectric substrate and then laminated to the conductor foil. Depending upon the nature of the base this can be done via a heated press for sheet processed materials such as polyamides, or through heated rollers for roll-to-roll materials such as polyester. Some form of post curing at elevated temperature is usually required after roller lamination due to the relatively short contact time between roller and laminate.

Adhesives must be carefully chosen to offer compatibility with both substrate and conductor materials. They must be able to provide adequate mechanical strength, have good chemical resistance, and be able to withstand the conditions used in FPC manufacture without delamination. Typical adhesives systems used for flexible-circuit manufacture include:

- Polyester
- Polyimide
- Acrylics
- Epoxies
- Fluoropolymers
- Phenolics

It is also important that adhesives act as part of the dielectric packaging of the signal, power, and ground circuit traces. They determine a fundamental part of the circuit’s electrical behaviour. Adhesives are typically available in a range of thicknesses from 0.5 mil to 5 mils in 0.1 mil increments.

It is often the case that the chosen dielectric material determines the type of adhesive used. For example, polyester adhesives are typically used with polyester laminates, and new formulations of high-temperature polyimide adhesives are increasingly used for polyimide substrates.

Importantly, developments are ongoing within the industry regarding the commercialisation of adhesiveless laminates, in which the substrate material and conductor layer are intimately joined without adhesives. Such laminates allow for improved environmental performance because it is often the adhesive layer that is the limiting factor in high-temperature applications. Developments in this area will be covered in more detail later in this report.

Polyimide Adhesives. Polyimide adhesives are thermoplastic materials, their high-temperature performance capabilities, withstanding temperatures as high as 500OC, has made their use and development attractive within the defence, aerospace and satellite sectors. Projects are underway to develop commercial systems at the likes of NASA, Rogers Corporation (USA), NEC (Japan) and others. Aside from providing matched performance with polyimide substrates used for demanding flex-circuit applications, the adhesive also has a very low coefficient of thermal expansion, which makes it a good choice for use in demanding multilayer circuits.

The downside to the material is its high cost and the limited sources of supply. Proven processing experience with high commercial-scale volumes of flexible circuits using polyimide adhesives is currently limited.

Polyester Adhesive. Polyester adhesives are low-temperature thermoplastics. They are relatively low-cost materials and can be processed using low temperatures. Their drawback is that they exhibit poor high-temperature performance and have relatively low bond strength.

However, modified polyester adhesives are available that have better high-temperature properties and high flexibility and can withstand many soldering operations. Polyester adhesives are widely used in applications where polyester is the dielectric substrate material, and where the application itself does not present extremes of temperature or forces that will significantly stress the circuit.

Acrylic Adhesive. Acrylic adhesives are thermosetting materials and have a higher resistance to soldering conditions than polyesters and modified epoxies. The materials are relatively low cost and their ease of processing and low flow characteristics during coating have made them a popular choice for flexible-circuit manufacture.

On the downside, the material does have a higher coefficient of thermal expansion and higher moisture absorption than polyesters and epoxies, which means that it can swell during processing. Flexible laminates with acrylics are generally available in sheet form, which can significantly increase the cost of FPCs made with acrylics. However, acrylic resin systems are being developed and marketed that offer better performance for roll-to-roll processing.

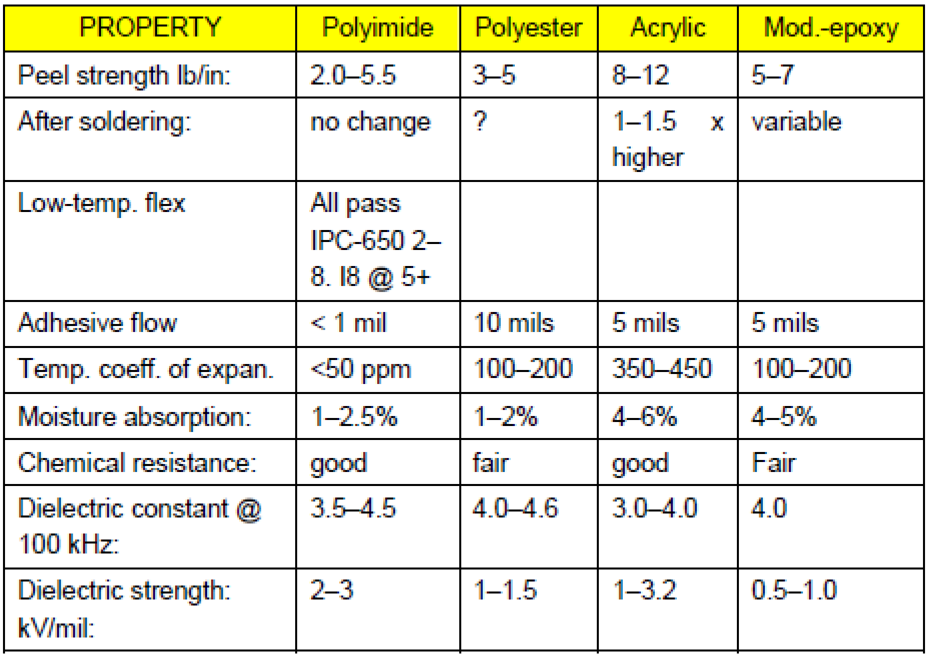
Epoxies and Modified Epoxies. Epoxies and modified epoxies are the most widely used adhesives for rigid circuit boards. They possess excellent resistance to high temperatures and modified grades offer excellent bond strength and material compatibility.

They are generally less flexible than other adhesive systems but they can be modified by the addition of other polymers, like polyesters, to increase their flexibility.

Other Adhesives. In developing flexible circuits to meet their applications, designers have used many varieties of adhesive materials. A number of these adhesive systems have enjoyed commercial success in other areas of electronics production and have been transferred over to flex by designers familiar with their capabilities and performance (Table 16.5).

Examples of other adhesives include phenolics, which have been used in TAB (tape automated bonding). The material also demonstrates very low flow characteristics which reduces the risk of the material migrating to cover neighbouring connection lands and circuit pads. They are thermosetting materials and have physical properties similar to epoxies. Fluoropolymer-based adhesives are also used and offer a good range of electrical and mechanical capabilities over diverse environmental conditions. Their main drawback is their relatively high cost.

Table 16.5. Properties of flexible-circuit adhesives.



PROPERTY	Polyimide	Polyester	Acrylic	Mod.-epoxy
Peel strength lb/in:	2.0–5.5	3–5	8–12	5–7
After soldering:	no change	?	1–1.5 x higher	variable
Low-temp. flex	All pass IPC-650 2–8. I8 @ 5+			
Adhesive flow	< 1 mil	10 mils	5 mils	5 mils
Temp. coeff. of expan.	<50 ppm	100–200	350–450	100–200
Moisture absorption:	1–2.5%	1–2%	4–6%	4–5%
Chemical resistance:	good	fair	good	Fair
Dielectric constant @ 100 kHz:	3.5–4.5	4.0–4.6	3.0–4.0	4.0
Dielectric strength: kV/mil:	2–3	1–1.5	1–3.2	0.5–1.0

Adhesiveless Laminates. As discussed previously, adhesiveless laminates represent a new emerging class of base material for flexible circuits. Today many flexible circuits consist of a base film and foil bonded together with an adhesive. Industry experience, research and product histories have shown that it is often the adhesive used in flexible circuits that is the principal limiting factor in many of the circuit's capabilities. The performance of the adhesive chosen hinders high-temperature performance in particular but also chemical resistance and multilayer build capability. For example, the relatively poor thermal expansion characteristics of adhesives causes excessive expansion problems in multilayer circuits, which causes stress in conductors and delamination of fine circuit features if unchecked. Adhesiveless materials avoid all of these pitfalls by doing away with the adhesive layer.

Adhesiveless laminates are manufactured in a number of different ways. A thin coating of seed conductor material, typically copper of less than 1 micron, may be placed directly onto the base laminate via techniques such as sputtering and electroless plating. These thin surface layers can then be selectively electrodeposited up to the required circuit thickness, and finally the whole circuit is given a flash etch back to remove the web of copper connecting the circuit traces to produce the final circuit. Overall, additive circuit manufacturing processes produce less waste and can generate extremely thin circuit tracks suitable for high-density circuits. Alternatively, the base laminate material, such as polyamide, may be cast in resin form onto a carrier foil, such as copper, and when processed through heated rollers and allowed to cure, forming a continuous adhesiveless laminate suitable for further processing.

Driving the market uptake of adhesiveless laminates is demand for complex highspeed circuits that are capable of operating at higher temperatures and frequencies. These must make use of multilayer construction techniques. In this regard an adhesiveless copper-polyimide, for example, will be more suited to the application, and has the added advantage of being thinner.

Traditional foils adhesively bonded to base films are proving too thick for high-density circuitry. Also the extra copper thickness associated with adhesively bonded laminates requires extra processing time to etch back and is less environmentally friendly. Adhesiveless copper-polyamide laminates are now offered with thicknesses ranging from 0.3 microns to about 35 microns, with the thinner material proving much more suitable for fine line circuitry.

Since coming onto the scene in the mid 1980's, adhesiveless manufacturing techniques have improved, as has appreciation of their worth. They have now captured a significant proportion, some 10%, of the flexible-circuit market and research into their development and future potential is ongoing.

Protective Coatings. Protective films or coatings may be selectively applied to the surface of FPC to protect it from moisture, contamination and abrasion, and to reduce conductor stress during bending. The protective layer is placed over the circuit once the conductor pattern has been established.

Cover Lays. A cover lay (also known as a 'cover layer') is usually a combination of a flexible film and a suitable pressure-sensitive or thermosetting adhesive. The most commonly used materials are polyester film coated with polyester adhesive, polyimide film with acrylic adhesive, and polyimide film with epoxy adhesive. As stated, in circuit design the usual practice is to match the cover-lay film to the material of the base substrate.

The purpose of a cover lay is threefold: to provide circuit and conductor protection; to allow access to circuit pad and contact areas for further processing such as soldering and conductive adhesive bonding of components; and to enhance circuit flexibility and reliability.

To enable access to required conductor features beneath the cover lay, such as pads and contact points, registration holes are drilled, punched, or laser machined into the film. The cover lay is then registered over the conductor pattern and laminated using heat and/or pressure according to the adhesive's requirements.

To reduce conductor damage from frequent bending, the thickness of the cover lay should be the same as the thickness of the dielectric layer. This arrangement places the conductor traces near the neutral axis of the finished circuit assembly, in effect in the centre of the layered construct, which significantly reduces conductor stress during flexing.

An increasingly popular alternative to pre-punched and drilled adhesive films is the photoimageable cover lay. A layer of light-sensitive material, either in film or liquid form, is placed over the top surface of the conductor trace layer. The layer is exposed to light through a photographic negative that acts as a mask, selectively exposing areas of the film to the light. The light-sensitive coating cures in the exposed areas and subsequent processing strips uncured material to leave a patterned covering which provides access to contact pads and soldering lands.

Cover Coatings. 'Cover coating' is a broad term denoting a growing range of thin coatings applied to flexible circuits instead of cover lays. Such coatings are usually applied in liquid form via techniques such as screen printing. This allows access features to be generated at the stage when the coating is applied. The coatings are then rapidly cured thermally or by exposure to UV radiation.

Cover coats are best suited to applications where no or minimal flexing is required as, unlike thicker cover lays, they do not protect the conductor layer to the same degree for flexing-induced forces. Typical cover-coat materials are acrylated epoxy and acrylated polyurethane, both of which are applied as liquid polymers, are solvent free, and are rapidly cured by exposure to UV.

16.4. FLEX CIRCUIT MARKET AND APPLICATIONS

The diversity of flexible-circuit design is indicative of the numerous applications to which the technology has been applied. It is clear that flexible-circuit technology is suited to a wide range of circuit applications where rigid board technology is currently used. As circuit designers and engineers become more familiar with the capabilities of FPCs it is anticipated that the technology will begin to make serious headway in the heartland of rigid-circuit applications.

Market. According to the report *"World Market for Printed Wiring Boards and Substrate Materials*," by the IPC Technology Market Research Council (TMRC), world production of printed wiring boards (PWB) and flexible circuits reached a record high of US$42.7 billion in 2000. Japan ranked number one in 2000, producing the most rigid PWBs in the global market with 27%, followed by the United States, 25%, Taiwan, 11%, China/Hong Kong, 9%, and South Korea, 5%.

The same report also comments that global flexible-circuit production has grown significantly, reaching $3.9 billion in 2000. The IPC states that the figures were generated by input from industry experts, manufacturers and organisations around the world. Accordingly, the flex-circuit world market in 2000 was also led by Japan, accounting for 36% of production, followed by the US producing 28%, Taiwan, 7%, Thailand, 6%, and Germany, 4%.

In general terms the largest single usage of FPCs, more than a third of all sales volume, is in computers, a category which includes peripheral equipment, such as printers and scanners. There is also a large consumption of FPCs in ink-jet cartridges and similar consumables (Fig. 16.6).

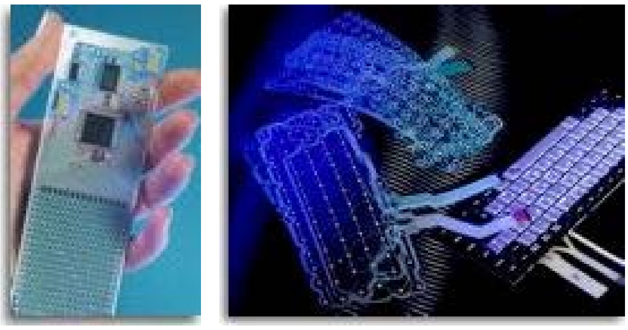


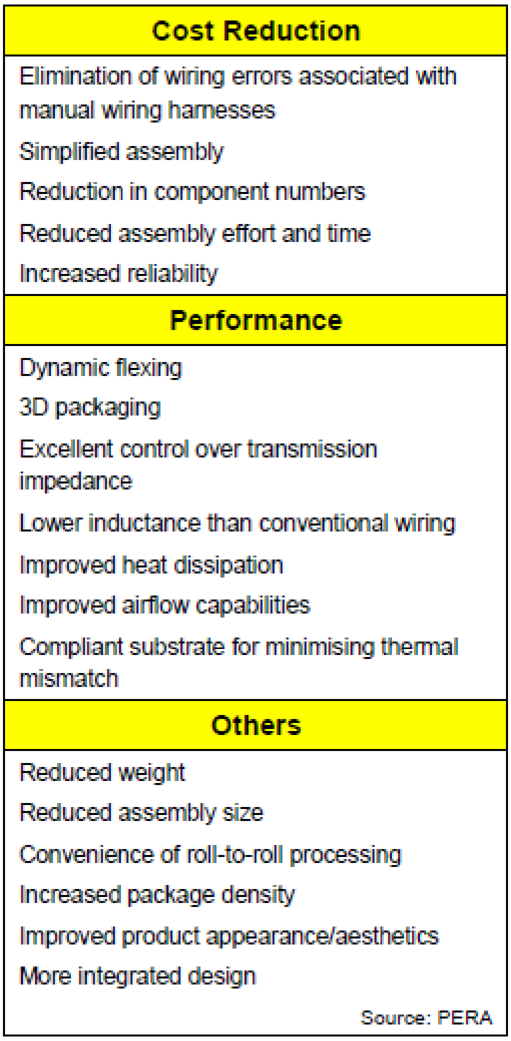
Fig. 16.6. Calculator touch pad and keyboard.

The second biggest application sector is the automotive, where FPCs can be found in many locations from the relative calm of the dashboard to the more demanding underbonnet, where engine management and control and ABS control systems are to be found. Automotive applications account for a fifth of the market.

Telecommunications, including mobile phones, pagers and all manner of communications infrastructure equipment, accounts for a further fifth of the market. The remaining areas of the market are typically low-volume applications for defence or aerospace, principally high-tech, in the form of very expensive but indispensable flexible and hybrid circuitry.

The Benefits of Flex. By far the biggest drivers for flexible-circuit technology are the range of benefits and capabilities that the technology offers. Some of the advantages of flexible printed circuits are highlighted in Table 16.6.

Table 16.6. The benefits of FPCs.



Cost Reduction
Elimination of wiring errors associated with manual wiring harnesses
Simplified assembly
Reduction in component numbers
Reduced assembly effort and time
Increased reliability
Performance
Dynamic flexing
3D packaging
Excellent control over transmission impedance
Lower inductance than conventional wiring
Improved heat dissipation
Improved airflow capabilities
Compliant substrate for minimising thermal mismatch
Others
Reduced weight
Reduced assembly size
Convenience of roll-to-roll processing
Increased package density
Improved product appearance/aesthetics
More integrated design

Source: PERA

In the face of increasing industry challenges and consumer demands such as miniaturisation, lightweight products, lower cost, greater product design freedom, high reliability and more environmentally extreme applications, flexible circuits are proving their worth.

There are many advantages to using FPCs. They utilise the thinnest dielectric substrates available for electronic interconnection, down to 0.002”, and are known for their ability to reduce package size as well as package weight. FPCs can reduce the weight of an electronic package significantly – by up to 75%. This weight reduction makes flexible circuits extremely popular in the aerospace industry. Another advantage of FPCs is assembly costs. Reduction of assembly costs is achieved by reducing the number of assembly operations and having the capability to test the circuit prior to committing it to assembly. A properly designed flexible circuit is an excellent means of reducing the number of levels of interconnection required in an electronic package. Flexible printed circuits can eliminate wiring errors associated with hand-built wire harnesses, since it is not possible to route flexible printed circuits to points other than those designated in the artwork.

Developments within the field of electronic components are also positive developments for flexible circuits. The growth in surface mount technology (SMT) and the development of conductive adhesives, used to attach such components to circuit boards at relatively low temperatures, has favored the use of flexible substrates. Such components are highly sensitive to thermal mismatch between substrate, mounting and component materials. It is recognised that flexible circuits offer a highly compliant material that is able to counteract the effects of thermal stress with more success than rigid laminates, making their use in environmentally taxing conditions highly appealing.

Applications. There are currently two basic use categories of flexible-circuit applications, flex-toinstall and dynamic operation. As the names imply, flex-to-install applications, which represent the majority of use, are those that require the circuit to be formable at the time of assembly, to fit into a product with maximum ease (Fig. 16.7).

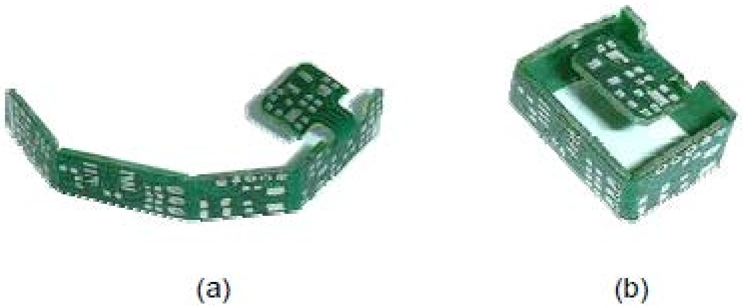

(a) (b)

Fig. 16.7. Assembly of flexible circuit. A single-sided flexible circuit for generating a 3D interconnection to minimise space requirements (a) as produced and (b) after assembly.

Dynamic applications are those that make full use of the circuit's dynamic capabilities, often resulting in applications that require many thousands of flexing operations throughout the circuit's lifetime. Examples of such applications include typewriters and printers, where the requirement is for a connection between a moving element and its control system (Fig. 16.8).

Fig. 16.8. Printer flexible circuits. The major application within computer printers – the interconnect between the printing head and its control system.

Within the consumer arena, the major application currently demanding the greatest dynamic flexing applications is in the area of computer disk drives, where the readwrite head tracks along the surface of the hard disk. This highly demanding application requires billions of flexing operations of the products life and so requires a highly reliable interconnect between the head and its control module.

Automotive. Polyester substrate and copper-conductor-based flexible-circuit constructions have been utilised for high-volume instrument-cluster wiring and interconnections for over twenty-five years (Fig. 16.9, Fig. 16.10). Indeed, the automotive instrument-cluster market was one of the original, flex-to-install, volume drivers for the flexible-circuit market. Such circuits are used to deliver both power and connectivity to instruments mounted on vehicle dashboards, including lighting, power, sensor connection and touch buttons.

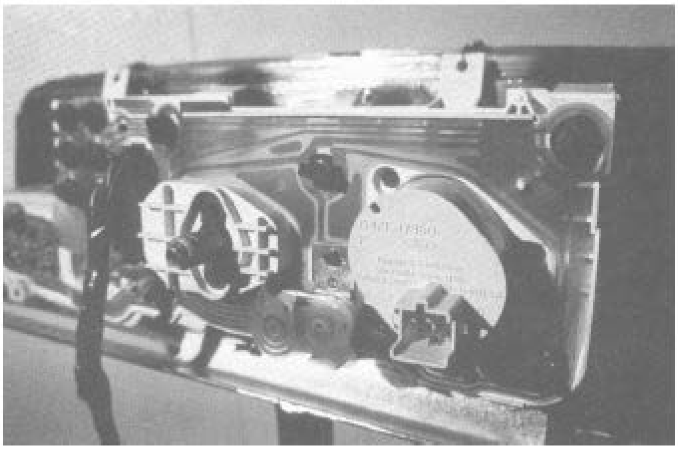


Fig. 16.9. Instrument cluster.

Fig. 16.10. Automotive dash flexible circuit.

In the early years automotive cluster circuits were single-sided and of low complexity. They could be accurately formed by die stamping. However, the modern trend within vehicles of incorporating an increasing array of sophisticated sensors and the general proliferation of automotive electronics has turned dashboards into electronic nerve centres. Current designs of dashboard-based instrument-cluster flexible circuits can typically consist of a number of interconnected layers, with greater circuit area to handle increased power requirements.

The major growth in in-vehicle electronics is a significant area of opportunity for FPC technology. It is anticipated that in 2002, 30% of a car's price will be accounted for by its electronic equipment, notably antilock brakes, air bags, electric windows and power steering. The growth in driver safety and maintenance features such as ABS and engine management and control requires the application of flexible interconnects and circuitry in increasingly new and more demanding environments. Under-bonnet applications are on the increase, where circuits are exposed to temperature extremes, chemical contaminants and dirt and an environment sensitive to, and generating a lot of, electrical interference.

The original volume growth of automotive circuits was met by low-cost polyester circuits of relatively low technology. An increasing call for circuits that are required to form more safety-critical functions in an environment where they are exposed to extremes of temperature, aggressive chemicals and high levels of dirt has led to the adoption of more robust circuit constructions and base materials.

The scope of the potential for flexible circuits within the automotive arena can not be underplayed, and the technology is seen as key to reducing the complexity of modern vehicle assembly, reducing assembly errors and meeting increasing demands for low weight, low cost and reliable connectivity. These demands are set to escalate as the vehicles are forecast to contain an increasing array of electronic automotive systems.

The automotive sector makes extensive use of wiring in harness for a variety of on vehicle control applications. The length of cable used varies significantly from motor vehicle model to model, with 500 metres in the Fiat Punto, typically some 1,600 metres in an S-Class Mercedes and up to 2 or 3 kilometres in a larger BMW. For some vehicles the figure for copper-cored cable is increased to five kilometres by multiplying the length of multi-core cable by the number of cores enclosed within a wiring sheath [9].

Typical automotive applications of wire and cable are shown in Table 16.7. Current wiring-harness technology is finding itself increasingly strained as it struggles to cope with the increase in on-board electronics. Even the task of routing numerous wires around a modern vehicle is becoming a challenge. Also, as vehicle manufacturers find themselves increasingly driven by regulation and legislation regarding recycling vehicles and better vehicle fuel economy, greater attention will turn to FPCs, which use less raw material, are easier to recycle, reduce weight and offer a reliable low-cost interconnection method.

Table 16.7. Automotive applications of wire and cable.

Engine	
Special cables	ABS
Winding wires	Starter/alternator
Coils and transformers	Starter/alternator harness
Optical bus	Fibre-optic control
Power source	Electric vehicles
Microsystem	Injection
Interior / Exterior	
Special cables and sensors	Air bags
Winding wires	Electric windows, seats
Printed circuits	Dashboards
Harnesses	Air conditioning, lighting
Antenna	Radio/mobile phone
Lighting fixtures	

Telecommunications. Telephone handsets have utilised flex-to-install FPC technology for many years, the principal advantage being that FPCs are well suited to the cured design of handsets and offer a low-cost, reliable circuit technology.

Flexible circuits can be found in a whole range of telecommunications equipment from low-cost fax machines, where they deliver low-cost interconnectivity, to automated switchboards, where they afford significant space saving and, through the use of multilayered circuits, a high degree of connectivity.

Significantly, Flexible boards and laminates have greatly benefited from the explosion in cellular technology. Sales of polyester and polyimide flex circuitry in the telecommunications sector amounted to $56 million and $471 million, respectively, in 1998.

Within mobile phones, flexible circuits have been employed to solve many of the technical problems that have enabled the technology to keep pace with changing customer and designer demands. Initial applications of flexible circuits in mobile phones were in touch-pad type applications. At the time, button keypads were being replaced with backlit pliable dome buttons, and polyester flex circuits combining the attractions of low cost and transparent substrate were employed to provide electrical connectivity.

Flexible circuitry can be found in the rechargeable battery packs that power many phones, where it proved an effective way to regulate discharge in nickel metal hydride and lithium ion batteries. In this regard, not only did flex circuits simplify battery-pack assembly, but they also reduced cost and gave designers an easy way to locate fuses, typically on the flex circuitry itself.

As mobile phone screens and LCDs have become more complex, showing an ever greater amount of information, the increasing demands for more inputs and outputs has led to the application of flex circuits to link screens to their driver chips and power supply. Again, designers have saved space, improved reliability and lowered assembly complexity by locating components directly onto the flex substrate. Materials of choice revolve around polyimide due to it strength and ability to accept soldered components.

A further widespread use of flex circuits is for internal and external antennas for handsets. Low-cost polyester flex has been found to be an innovative technology for their production in the face of the tough challenge of improving transmission range while reducing exposure of the user to stray radiation.

Various grades of low-cost polyester laminates are used for antenna construction, and their flexible and lightweight construction make them ideal for flex-to-fit placement within the phone housing. Flexible-circuit tracks and circuit traces have the ability to effectively channel radiation patterns. They can maximise signal effectiveness while keeping radiation away from the user. In external applications the circuits are often coated with a protective rubber or plastic material, to produce a desired antenna shape.

Overall, flexible circuits have evolved to meet many of the major challenges faced by the functionality and miniaturisation requirements experienced by the mobile phone industry. Whilst the use of flex has become advantageous for some aspects of phone design, it is considered essential in delivering the functionality and durability required of both phone, PDA and laptop designs that incorporate hinged flip lids. Previously, only RA copper circuits were considered capable of meeting the demands for many thousands of flexing operations over the lifetimes of such products. However, the developments in adhesiveless laminates with the same level of flexing and durability have been eroding RAs stronghold in the market.

Aviation. Large quantities of cables are used in aircraft. Typically Airbus models have 70–100 kilometres of cable while helicopters require over 12 kilometres of cable [10]. The importance of the sector, in wiring terms, is underlined by the recent moves of automotive harness manufacturers into this lucrative market.

The FPC applications within the aviation sector are almost exclusively based on polyimide film with copper conductors. The aviation sector and in particular the defence sector were pioneers in the development of flex-rigid boards, which offer the advantages of both rigid and flexible circuits.

Within the aerospace sector flexible-circuit technology has come into its own, driven by the aerospace market’s stringent requirements for both reliability and extreme environmental performance.

However, given the Asian financial crisis and the events of September 11th there has been a significant downturn in the sector, which is experiencing restructuring on a global scale. The market is not expected to recover the levels it reached prior to September 11th for some significant time.

16.5. MANUFACTURING AND DESIGN OF FLEXIBLE CIRCUITS

The design and manufacture of flexible circuits demands production techniques and a design approach that are significantly different from those used for their rigid cousins.

The development of a successful flexible circuit requires an approach that combines mechanical and electrical disciplines to produce the end product. It is this combination of disciplines that more often than not leads to ambiguity in the circuit's design, which in turn comes back to haunt developers as expensive manufacturing problems later on in the production cycle.

In this regard it is essential to take the best path and consider all of the relevant options during circuit design to avoid future problems. As with any design and development approach, consideration of manufacturing requirements will be critical if tolerances are to be held, standard materials and processes used, and costs kept low.

Manufacturing Flexible Circuits. The manufacture of flexible circuits is a highly variable process, with the inherent design flexibility offered by FPCs resulting in many combinations of manufacturing stages for the basic types of circuit constructions.

Some of the keys processing steps for flexible circuits have a high degree of similarity with the manufacture of their rigid counterparts. One of the most significant differences between the two is the amenability of flexible circuits to roll-to-roll processing. As discussed, this approach bestows a significant number of processing advantages on flexible circuits in terms of their high-volume manufacture and low circuit cost.

However, as discussed roll-to-roll circuit manufacturing raises distinct production problems, specifically the required accuracy of image registration and difficulties with holding circuit tolerances given the material's ability to expand during processing and its susceptibility to changes in web tension.

Given the number of basic flexible-circuit constructions it would be impractical to discuss all of the various forms of manufacture. For more detailed discussion of these requirements a number of excellent publications exist [9-12].

The discussion here has been restricted to the major processes associated with flexible-circuitry manufacture and seeks to highlight some of the essential differences between the production of rigid and flexible circuits. It is anticipated that significant opportunities are evident in the area of flexible-circuit manufacture and design that may stand to generate significant business opportunities if identified and exploited.

The Manufacturing Route. It is not surprising to realise that some of the most widely used manufacturing processes in flexible-circuit production have been imported and adapted from the most popular techniques used to manufacture rigid boards. Circuit designers, familiar with the constraints and advantages of rigid-board manufacture, have adapted techniques to the needs of flexible circuits. This factor, along with the widespread use of copper as the conductor material of choice, has meant that established techniques such as electroplating and PTH have successfully migrated to the flexible world.

In considering the process stages for flexible-circuit manufacturing the requirements of a single-sided circuit will be exemplified. The production of double-access and multilayer circuits requires the use of increasingly sophisticated manufacturing techniques and involves additional process stages. Where these involve technologies or process steps that are well established and considered of interest, they will be touched upon in the following discussion.

The broad process stages of single-sided circuit manufacture are shown in Fig. 16.11:

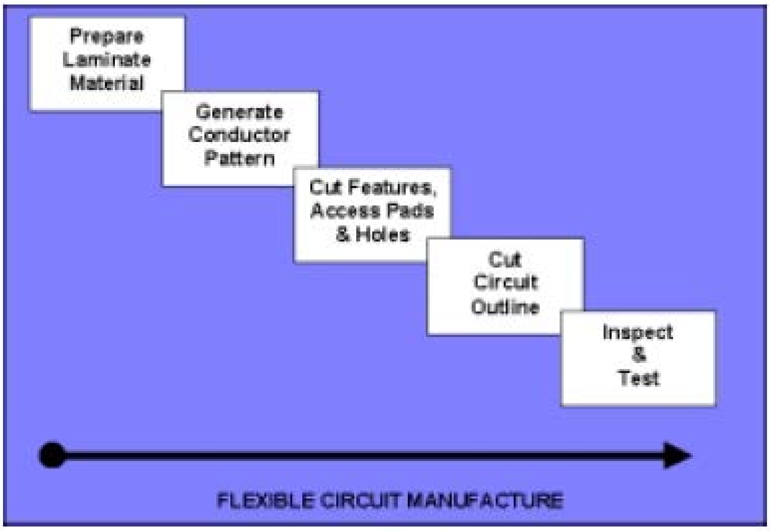


Fig.16.11. Basic manufacturing stages

Clean the Laminate. The typical starting point for circuit manufacture is a copper-clad laminate. This may consist of a base material with an adhesive-bonded copper layer or it may be an adhesiveless laminate as previously outlined. There are a variety of reasons for cleaning the laminate prior to processing. Copper- and foil-clad laminates will most likely be provided with an anti-tarnish treatment on their surface, placed there by the laminate or foil supplier to prevent oxidation and tarnishing. This coating is detrimental to further processing and so is removed by dipping in an acid bath or using an acid etching spray. The exposed foil can then be micro-etched, typically with a treatment of sodium persulphate, to promote resist adhesion. Finally, after a suitable rinse, the board is usually coated with an in-process oxidation agent to prevent tarnishing during further operations.

Image Resist. A pattern is generated to reflect required circuit traces. The two most widely used techniques are screen printing and photoimaging. Screen printing is an established technique that can deliver deposits on the laminate surface of 4–50 microns in thickness. Control of the screen printing process can be an art form as there are many variables such as deposit (liquid resist) consistency, screen tension, application squeegee profile and screen geometry. In screen printing the desired resist pattern is printed directly onto the laminate.

Photoimaging is typically undertaken using a dry photoresist film that is applied to the entire laminate surface. A photomask, consisting of artwork depicting the desired circuit traces, is placed in close contact with the film resist. The assembly is exposed to UV light which hardens the resist where it is unprotected by the mask. The uncured resist is then chemically removed to reveal a patterned laminate.

Etch the Exposed Conductor. This stage aims to remove the unprotected conductor material to leave a selective conductor pattern on the laminate surface. The resist previously applied is impervious to a selected etchant. This etchant removes the copper in a controlled process, as in the initial cleaning stage. The laminate may be exposed to the etchant by dipping, in an etch bath or by being sprayed with the etchant solution. It is accepted that a spray etchant process allows for the generation of finer line tolerances with straighter edges on circuit traces; dip etching undercuts the circuit traces.

Resist Removal. The resist, having performed its function in protecting the required circuit traces is now removed via a separate chemical or mechanical processing stage.

Cover-Lay or Cover-Coat Placement. It is undesirable to have a circuit where the conductor pattern is vulnerable to scratching and electrical shorting via accidental contact with the conductor layer. In this regard a protective dielectric film or coating is placed on top of the exposed conductive tracks to prevent this, stopping environmental ingress and maintaining electrical integrity.

Produce Holes and Outline. One of the penultimate manufacturing stages is to produce the required circuit outline and any through-laminate holes. A number of manufacturing techniques can be employed at this stage. One of the most popular is die cutting, which consist of a male and female set of tools in a punch press. Modern die cutting machinery offers highly accurate outlining of both circuit outlines and holes, and various tooling arrangements can be developed to deliver precise cutting for both sheet and web-fed materials.

A less costly alternative is the use of a steel rule die, which use a series of punches and blades held in a low-cost wooden or plastic base. This technique offers less accuracy than die cutting and significantly lower tool life. However, it does offer a rapid tooling approach for lower-volume production and a relatively low cost. The cost of steel rule dies is typically between one-tenth and one-third that of die cutting tools.

Test and Verification. The final stage in successful flexible-circuit manufacture is test and verification of circuit performance and quality against the set specification. Table 16.8 outlines the key process stages for the three main manufacturing routes. Two generic process routes exist for generating an accurately patterned circuit on a suitable laminate material – additive and subtractive. Subtractive processes, as outlined in the manufacturing stages above, are the most popular route, growing from their use within the rigid-board electronics industry.

Table 16.8. Flexible-circuit manufacture.

Additive Printing Processing	**Subtractive Etching Processing**	**Additive Plating Processing**
Clean the laminate	Clean the laminate	Clean the laminate
Print circuit pattern	Apply the image resist	Apply the image resist
Cure	Etch conductors	Plate the conductors
Cover coat/cover lay	Strip resist	Strip resist
Cut holes	Cover coat/cover lay	Etchback/flash etch
Die-cut outline	Cut holes	Cover coat/cover lay
Inspection and test	Die-cut outline	Cut holes
	Inspection and test	Die-cut outline
		Inspection and test

Additive processes are under development and are used by a number of manufacturers. In additive plating, the starting material is typically an adhesive less laminate possessing a very thin layer of exposed conductor material over its entire surface. A negative resist image is applied to the laminate, one that leaves the desired conductor paths exposed. This exposed area is then subjected to an additional plating operation that builds up conductive material in the unprotected areas. The resist is removed and the assembly is given a final flash etch to remove the original thin conductor layer, thus electrically isolating the built-up conductive tracks.

The main advantage of the additive process is the general agreement that it results in circuits that have improved performance over subtractive circuit manufacturing. This performance differential is due mainly to the advantage that additive plating removes the need for an adhesive layer between the conductor and the substrate, which typically limits performance at higher temperatures. Additive plating also allows the reproduction of finer circuit tracks because of the way the track are formed.

Whilst a number of manufacturers are unfamiliar with the additive plating process it is anticipated that the popularity of the technique will grow alongside the demand for greater circuit density and the use of adhesiveless laminates.

16.6 FLEXIBLE-CIRCUIT DESIGN

With their higher line density, improved impedance control, lighter weight and greater reliability, flexible circuits can be designed for a range of applications beyond the scope of traditional cabling and connectors. Often, flex circuitry is the only design option available that can provide a required combination of interconnect density, and electrical, mechanical and environment performance.

The discussion below lays out some of the basic rules for beginning the process of specifying flexible circuits. Further advice can also be obtained by communication with a chosen flexible-circuit manufacturer.

Basic Design Rules. Many years of heartache, woes and lost profits have served to harden the flexiblecircuit industry to the invariable teething problems that accompany the introduction of new technologies. Basic rules of thumb have evolved that help to minimise design costs and improve manufacture.

- Avoid using non-standard base dielectric material thickness and adhesiveless laminate, both of which would entail special orders. Dielectric material and adhesiveless laminate come in standard sizes. Consult the chosen FPC vendor at the earliest stage of the design process and take the time to study the design guides that major flexible-circuit manufactures have compiled to showcase their products against their competitors’.
- Careful consideration should be given to manufacturing tolerances. It is expensive to request levels that approach the limits of the tooling used and it will potentially result in a significant number of rejects. For example, steel rule dies, used to produce circuit outlines and trim laminates, can hold 0.010" but specified allowances should be 0.015". Go tight on tolerances only where it is absolutely necessary to the functioning of the final circuit.
- Leave sufficient clearance between the edge circuit trace and the circuit outline. This error has been responsible for many circuit failures, not leaving enough clearance in the case of circuits for dynamic application circuit may prevent adequate cover-lay adhesion and promote delamination of the circuit material at the outline. Minimum distance should be in the region of 0.010". Similarly, the edge of any stiffener materials should be pulled back from the final circuit outline. It is advantageous to select a stiffening material that is easily punched. Moulded stiffeners are expensive to fabricate but are commonly found in computer hard disk drives. Tooling holes should be placed within the area to ensure easy registration of the stiffener to the circuit.

Implementation Recommendations. There are few general implementation recommendations that can be of great assistance to those seeking to implement flexible circuits for the first time. For newcomers to the technology the first major step is the realisation that flexible circuits can help.

Information is usually the second step, seeking a source of flexible-circuit knowledge that can verify any initial finding and help advance the process from potential use to design reality. There is no doubt that among the best sources of such knowledge are flexible-circuit manufacturers themselves.

There are very few generic rules when it comes to flexible-circuit design, which is understandable given the variety of the material, process and assembly options available. There are good-practice guidelines that will help ensure that the design process starts off on the best possible footing.

Costing Flexible Circuits. It is difficult for any justification of a switch to a new technology, such as flexible circuits, not to consider principally the cost implications and assess the potential savings that may result from change. From an initial perspective flexible circuits offer cost-saving opportunities over traditional rigid circuit board assemblies and can typically reduce costs by 20% to 70% through the elimination of wiring errors, meeting space restrictions, reducing inspections and simplifying assembly.

A point to be borne in mind is that many of the inherent advantages of flexible circuits, such as simplified assembly and reduced errors, are system-level advantages. Therefore it is important to note that often, for a given application, the cost of a single flexible circuit will be more expensive than a rigid counterpart. However, the system level benefit of moving to flexible circuits in terms of design freedom, flexibility, dynamic performance, reliability and reduced errors, represent the real opportunity to derive value and generate savings.

Simple break-even calculations can be applied to gauge the relative advantage of moving to flexible-circuit technology. To complete such calculations requires an assessment of the recurring and non-recurring costs associated with an application – no simple feat. By way of example, consider a typical scenario in which a wiring harness is replaced by a flexible-circuit alternative. For such a situation the factors for consideration in determining cost are as laid out in Table 16.9.

Table 16.9. Costing flexible-circuit applications [11].

Break Even Costing Of Flexible Circuits	
Flex Circuit NRE	Wire Harness NRE
Circuits layout Artwork Documentation Mock-ups Hard & soft tooling	Wiring diagrams Assembly drawings Harness assembly boards Potting moulds
Flex Circuit RC	Wire Harness RC
Unit price Inspect and test circuits Assemble and / or install Inventory	Wire inventory Cut, strip and mark wires Lay out wires on boards Solder wires to connector Test and rework errors Pot wires into connector

The required formula is:

Break-Even Point Equals

$$\frac{\text{NRE (Flex)} - \text{NRE (alternative)}}{\text{RC (alternative)} - \text{RC (flex)}}$$

Key: NRE = non-recurring engineering costs; RC = recurring costs

16.7. FLEXIBLE-CIRCUIT TECHNOLOGY AND PROCESS TRENDS

The flexible-circuit industry is highly dynamic in terms of new applications. There is a certain degree of maturity in some of the processes and materials that are currently the mainstay of the market. Many emerging applications are pushing existing materials and processes to their limits – sophisticated electronic assemblies, optoelectronic modules, liquid crystal displays, microelectromechanical systems (MEMS), disk drives and highfrequency circuits amongst others.

Those within the flexible-circuit sector have responded to market demands by developing new interconnection methods and fostering the development of new materials and new processes designed to meet the challenges ahead.

Currently there is much activity within the flexible-circuit arena that holds promise for improved flexible circuitry for new applications. Whilst developments are numerous, it is clear that certain developments will play a significant role in meeting market demands in the short to medium term by realizing new materials and new processes.

Major developments worth noting relate to new substrate materials, new cover coats, the development of micro-via technology and new manufacturing techniques. What is clear is that with each development, flexible circuits will continue to encroach upon the application areas currently the reserve of their rigid counterparts and that the traditional barrier that has divided rigid and flexible applications will crumble.

Comment is offered below on some of the most promising developments within the flexible-circuit sector, key areas of development identified are:

- Adhesiveless laminates
- Direct circuit application methods
- Cover lays
- Conductive plastics
- Chip-scale packaging opportunities

Adhesiveless Laminates. One of the most significant developments within the flexible-circuit market over the past ten years has been the growth and advance of adhesiveless laminates. These materials are an example of synergy between advances in new materials and new processes that have delivered functionality above and beyond initial expectations.

With the demand for greater interconnect density, the market penetration and development of adhesiveless laminates that offer higher levels of interconnectivity through finer conductor lines look set to continue apace. Currently the number of adhesiveless laminate suppliers is small but growing. The challenge ahead for flexiblecircuit manufacturers is to make the switch to the technology.

Such a move holds many industry challenges, not least to the expertise of designers in becoming more familiar with the materials, but also in production processes. These represent a radical departure for many production facilitates which are biased towards sheet processes while the current mainstays of adhesiveless materials – polyesters and polyimides – are best manufactured by roll-to-roll processing.

It is predicted by some that the unavoidable advantages of adhesiveless material technology will undoubtedly lead to its use in both PWB and FPW constructions and an elimination of the distinction between these two industries. It is anticipated when this change occurs that there will be an order-of-magnitude increase in the utilisation of flexible circuits because of the enormously large and well established PCB market.

Market analysts have concluded that equipment for producing adhesiveless laminates in roll-to-roll format is of such efficiency that there is enough existing capacity to service any increased market demand for the first few years of high growth.

Direct Apply Technologies. New techniques for circuit manufacture are continually evolving. Among the more promising production routes are the direct printing of conductor patterns. The hunt is on for processes that marry the distinct advantages of technologies such as PTF with high-density interconnection capability. Direct PTF-type production of conductors offers cost advantages over the expensive plating and etching processes widely utilized today.

An intermediate direct-print technology of promise is an adapted laser printing technique, which directly deposits material onto the film substrate in the pattern required. The deposited material can then be plated over to deliver a final circuit.

There are a number of companies currently experimenting with techniques that enable flexible circuits to be made without the use of conductive or resistive metal foils through the use of inherently conductive plastics (ICPs) [13].

Examples of conductive plastics tailored to deliver specific conductivity are anticipated to revolutionise areas like display technology. It is not unexpected that such technology will find applications in the production of all-plastic flexible circuits that may offer advantageous performance characteristics, most likely by precisely matching material characteristics of conductor and substrate.

The concept of ICP is not new. PTF uses polymers that mimic the conductivity of metals (particularly copper and steel) by compounding into the resin conductive metal particles or resistive carbon powder. Whilst this achieves a measure of conductivity, there are always compromises in terms of manufacture, performance or total part economics.

Research into conductive plastics and polymers is longstanding. In the mid-1970s, the first polymer capable of conducting electricity, polyacetylene, was discovered. By the mid-1980s, the number of patents issued in the field of electrically conductive polymers increased significantly, and the first application, the polymer electrode, was registered. By the mid-1990s, commercialisation of ICP was still in its infancy and production of these sometimes exotic materials has only been scaled up from grams to pounds. Whilst overall global production and consumption totals are still negligible hundreds of papers and many patents on inherently conductive polymers are published each year. There are a great many scientists and corporations who are optimistic about ICP's commercial potential.

New Substrate Materials. It is hard to believe that the current commercial dominance of polyimide and polyester materials could be radically altered by new material developments in the foreseeable future as it is very hard for any new material to penetrate the market significantly in the short to medium terms.

However, there are always new materials being offered to circuit manufacturers that offer incremental improvements in performance over existing material. Examples of polymers that have shown major improvements in dielectric or physical properties include Avatrel (BFGoodrich); PBO and PIBO (Dow), and liquid crystal polymers (LCPs) such as Vectra and Superex.

An area of intense research is that of LCP for flexible-circuit applications. In December 2001, 3M introduced their LCP dielectric film for flexible circuits that provides several enhanced properties required for demanding flex circuit applications. Principally cited were lower moisture absorption, 0.1% versus polyimide's 1–3%; an optimised dielectric constant; and higher tensile strength and tensile modulus.

Cover Coats: The Non-Cover-Lay Process. As discussed, cover coats are thin protective films that can be applied to circuits in preference to more sturdy cover lays. Their principal advantage is that their methods of application eliminate the need for some of the hole cutting and film/circuit alignment required with conventional cover-lay materials.

Cover lays, whilst proven for a wide variety of applications, generate dimensional uncertainties, which test their application with increasingly finer circuitry and highdensity interconnect component assemblies. The cover coating is evolving and is set to take over wherever finely detailed and accurately registered openings are needed.

Photoimaging technology, which is also used to define the circuit traces accurately, offers higher levels of placement accuracy. Screen printing technology can also be used to deposit cover coatings with high accuracy. Overall cost is reduced because the materials are applied only where needed, the labour and material consumption involved in press lamination is eliminated, and the inherent cost of the coatings is less than the cost of a dielectric film with adhesive coating.

Like adhesiveless laminates, cover coatings remove the need for the adhesive used to bond protective cover lays. So, like adhesiveless laminates, they offer the potential for improved thermal performance and greater dimensional stability – circuit parameters that are hindered by the presence of an adhesive. With the uptake of adhesives laminates and their penetration into the mainstream flexible-circuit market, cover coat technology looks set to follow.

The Next Generation of Applications – Chip Scale Packaging. An expanding area for flexible-circuit technology is that of chip-scale packaging. A number of new circuit constructions are emerging that further demonstrate the ability of flexible technology to increase circuit density in unusual ways [14].

The most compelling application for flexible circuits may well be as an IC packaging medium, termed chip-scale packaging. Flexible substrates have clearly demonstrated that when married with SMT component technology they provide assemblies that can survive many more thermal cycles before failure than assemblies using rigid boards.

There are three primary ways of employing flex circuits to package integrated circuits in an area array format (see Fig. 16.12):

- Attaching the semiconductor IC face-up on a flex circuit and wire bonding the leads to the leads on the flex circuit;
- Flipping the chips onto the flex substrate and joining them with solder followed by an underfill to protect the solder joints or by use of a conductive adhesive;
- Applying a flex circuit or flex film to the surface of the semiconductor IC facedown and interconnecting it to the die using an appropriate joining technique such as wire bonding or ribbon-lead connection.

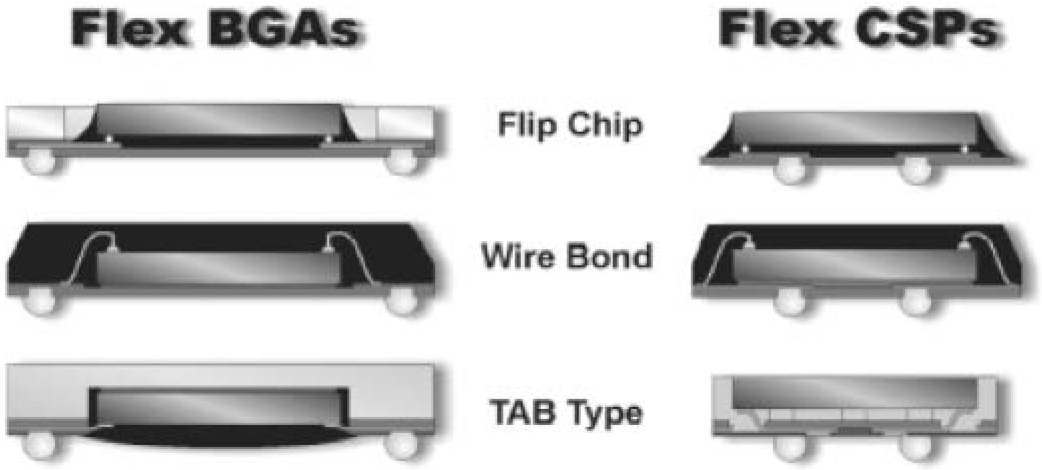


Fig. 16.12. Flex in IC packaging. A popular substrate choice among package IC designers, flex circuits are being used in BGAs and CSPs assembled by three common technologies: flip chip, wire bond and TAB (or TAB-like in the case of the μBGA® CSP at lower right).

The idea of utilising flexible laminates as a compliant interface for semiconductor packaging is not new. They were originally in use in the 1960s and 1970s. Developments in component technology in the interim, specifically SMT, have revived interest. Flexible circuits as semiconductor packaging are currently among the most active research areas of flexible-circuit application (Fig. 16.13).

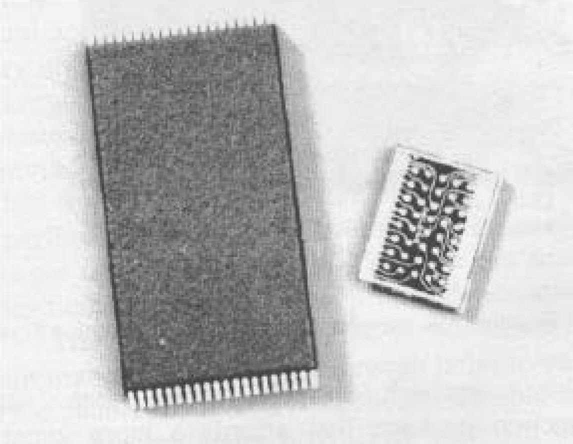


Fig. 16.13. Packaging size reduction. Micro BGA chip-scale package (on the right) uses a miniature flexible circuit for direct interconnection to the silicon chip. It is shown next to the TSOP package it replaced [11].

16.8. CONCLUSIONS

Flexible circuits represent a strongly growing sector of the electrical interconnectivity market. Thanks to the high level of demand for electronic products and the increasing penetration of electronic and electrical systems in the home, car and work environment, the need for interconnect and circuitry solutions that can meet a variety of demanding requirements will begin to favour flexible circuitry.

Flexible circuits offer a number of clear system-level benefits over rigid board technology. Factors such as weight reduction, mechanical performance and thermal compliance are becoming ever more important as electronic systems find themselves performing in ever more diverse applications and environments. Particularly key is the forecast growth in mobile electronics such as phones, PDAs, notebooks and portable DVD, which are pushing rigid circuitry beyond its current limits.

There is a clear shift away from the traditional wire-harness replacement of the past to more robust applications of the technology. This is being accompanied by a steady shift to high-quality commercial products and away from government/military production.

Materials of construction continue to be primarily polyimide with adhesively bonded rolled copper foil, but the application of adhesiveless systems is on the increase. With improved production of adhesiveless laminates and their availability from a wider range of vendors should come a greater penetration of their use in the flexible-circuit market.

As discussed, adhesiveless systems offer a number of clear performance benefits over adhesively bonded laminates, which make them better suited to a range of demanding applications.

New materials and new processes promise to make flexible circuits an exciting area of future promise for those seeking innovative solutions to a wide range of interconnection problems. Whilst polyimide and polyester material will continue to dominate as the base substrates of choice, new material developments are underway that, as they evolve to greater commercialisation, look set to provide significantly more circuitry options. Specifically developments with LCPs and PEN look set to challenge conventional material wisdom. The growth of direct apply technologies such as toner printing and lithographic circuits also look set to offer improved manufacturing processes which may in turn open up new markets for this exciting technology.

References to Chapter 16.

1. P. Macleod. A Review of Flexible Circuit Technology and its Applications. Pera Knowledge, 2002. ISBN 1-84402-023-1.

2. D. Shavit: The developments of LEDs and SMD Electronics on transparent conductive Polyester film, Vacuum International, 1/2007, S. 35.

3. IPC (1996) *IPC-T-50: Terms and Definitions for Interconnecting and Packaging Electronic Circuits*, Revision F (June 1996), IPC, Northbrook, IL

4. Kühl, Reiner W. (1999) Mechanical stress and deformation of SMT components during temperature cycling and PCB bending, *Soldering & Surface Mount Technology* Vol. 11, No. 2, 35–41

5. IPC (1991) *IPC-MF-150: Metal Foil for Printing Applications*, Revision F (October 1991), IPC, Northbrook, IL

6. Gilleo, Ken (1992) *Handbook of Flexible Circuits*, Chapman & Hall, London (ISBN 0-442-00168-1)

7. Gilleo, Ken, Bob Boyes, Steve Corbett, Gary Larson and Dave Price (1999) High volume, low cost flip chip assembly on polyester flex, *Circuit World*, 25 February, 11– 17

8. Evans, P. S. A., P. M. Harrey, B. J. Ramsey and D.J. Harrison (2001) Lithographic film circuits – A review, *Circuit World*, Vol. 27, No. 3, 31–35

9. Cousins, Keith (2000) *Polymers for Wire and Cable*, RAPRA (ISBN 1-85957-183-2)

10. Fjelstad, J. (1997) Inspection and Test of Flexible Circuits, *Circuit World*, Vol. 23, No. 3, 20–27

11. Fjelstad, J. (1998) *Flexible Circuit Technology*, Silicon Valley Publishers Group, Sunnyvale, California (ISBN 0-9667075-0-8)

12. Fjelstad, Joseph (1999) Flexible Circuitry – technology background and important fundamental issues, *Circuit World*, Vol. 25, No. 2, 6–10

13. Strohmeyer, Robert (2000) Flexible circuits - Roll out the next generation of hardware, *Smart Business for the New Economy*, 1 July.

14. Fjelstad, Joseph, Konstantine Karavakis and Belgacem Haba (1999) Manufacture of high density interconnection substrates by co-lamination of inner layers and programmed interconnection joining layers, *Circuit World*, Vol. 25, No. 3, 9–12

Chapter 17. SILICON NANOPHOTONICS (Evtukh A.)

17.1. Introduction
17.2. Photonic Crystals From Macroporous Silicon
    17.2.1. Various Shapes of Macroporous Silicon
    17.2.2. 2D Photonic Crystals
    17.2.3. Bulk Photonic Crystals
    17.2.4. Waveguides
    17.2.5. 3D Crystals
17.3. Outlook
17.4. Conclusion
References to Chapter 17.

Chapter 17. SILICON NANOPHOTONICS

17.1. INTRODUCTION

As is well known, there is the following relation between the frequency ν, the velocity c, and the wavelength λ_0, of the radiation field in free space:

$$c = \lambda_0 \nu \tag{17.1}$$

When we define the wave number k by,

$$k = \frac{2\pi}{\lambda_0}, \tag{17.2}$$

we obtain the relation between the angular frequency ω and k:

$$\omega = ck \tag{17.3}$$

This equation is called the dispersion relation of the radiation field. If one thinks of the radiation field in a uniform material with refractive index η, one can obtain its dispersion relation by replacing c by $v = c/\eta$ and λ_0 by $\lambda = \lambda_0/\eta$ in (17.2) and (17.3). The density of states of the radiation field in the volume V of free space, $D(\omega)$, is proportional to ω^2 (Fig. 17.1):

$$D(\omega) = \frac{\omega^2 V}{\pi^2 c^3} \tag{17.4}$$

The density of states in the uniform material is obtained by replacing c by v in this equation. The optical properties of atoms and molecules strongly depend on $D(\omega)$. As an example, let us consider the spontaneous emission of a photon from an electronic excited state of an atom or a molecule. Quantum mechanics tells us that the rate of the spontaneous emission is proportional to $\omega D(\omega)$. Since the spontaneous emission is an origin of the energy dissipation and the fluctuation of the radiation field, it suppresses the occurrence of laser oscillations. This suppression is marked in the high frequency region since $D(\omega)$ is proportional to ω^2. This is one of the reasons why the laser oscillation is difficult to realize at high frequencies.

Now, if we can design and modify $D(\omega)$, we can substantially change the optical properties of atoms and molecules [1, 2]. This is a key idea of contemporary optical physics, and it is possible. One method is to use optical microcavities while another is to use *photonic crystals*.

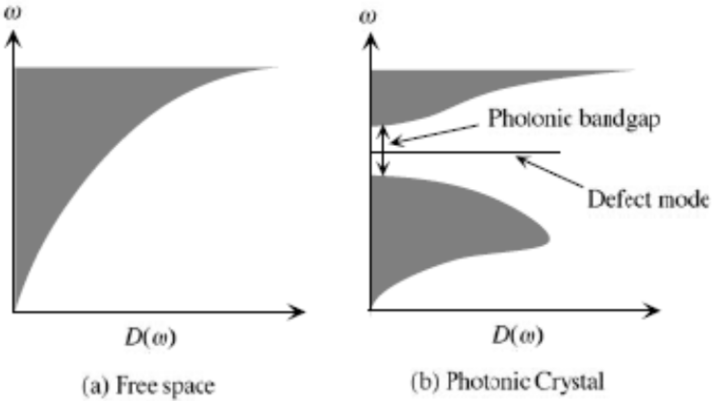


Fig. 17.1. Schematic illustration of the density of states of the radiation field (a) in free space and (b) in a photonic crystal. In the illustration for the photonic crystal, a photonic bandgap and a localized defect mode with a delta-function like density is included.

Photonic crystals are regular arrays of materials with different refractive indices. Figure 17.2 shows the simplest case in which two materials denoted by A and B are stacked alternately. The spatial period of the stack is called the *lattice constant*, since it corresponds to the lattice constant of ordinary crystals composed of a regular array of atoms. Actually, many basic ideas are common to both crystals and they will be utilized to build fundamental theories of the photonic crystals.

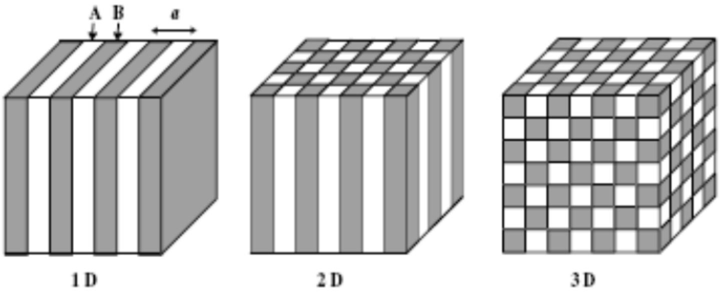


Fig. 17.2. Schematic illustration of one-dimensional (1D), two-dimensional (2D), and three-dimensional (3D) photonic crystals. *a* is the lattice constant crystals and how the optical properties of atoms and molecules embedded in them are altered.

However, one big difference between them is the scale of the lattice constant. In the case of ordinary crystals, the lattice constant is on the order of angstroms. On the other hand, it is on the order of the wavelength of the relevant electromagnetic waves for the photonic crystals. For example, it is about 1 μm or less for visible light, and is about 1 cm for microwaves.

Photonic crystals are classified mainly into three categories, that is, one-dimensional (1D), two-dimensional (2D), and three-dimensional (3D) crystals according to the dimensionality of the stack (see Fig. 17.2). The photonic crystals that work in the microwave and far-infrared regions are relatively easy to fabricate. Those that work in the visible region, especially 3D ones are difficult to fabricate because of their small lattice constants. However, various technologies have been developed and applied to their fabrication in the last ten years, and many good crystals with a lattice constant less than 1 μm are now available. For example, Fig. 17.3 shows an SEM (scanning electron micrograph) image of an fcc (face-centered-cubic) lattice composed of silica spheres [3]. The diameter of the spheres is 0.3 μm. On the other hand, Fig. 17.4 is an SEM image of a 3D structure made of Si [4].

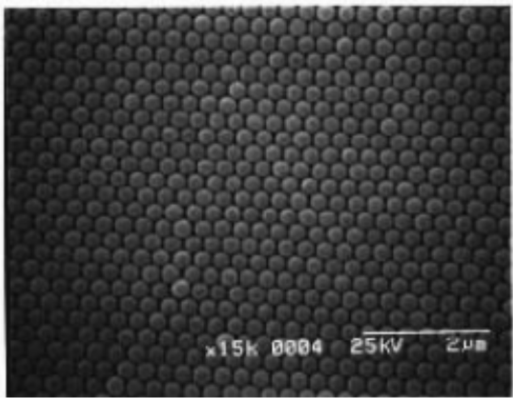


Fig. 17.3. SEM image of a 3D photonic crystal composed of an fcc array of SiO_2 spheres with a diameter of 300 nm [3].

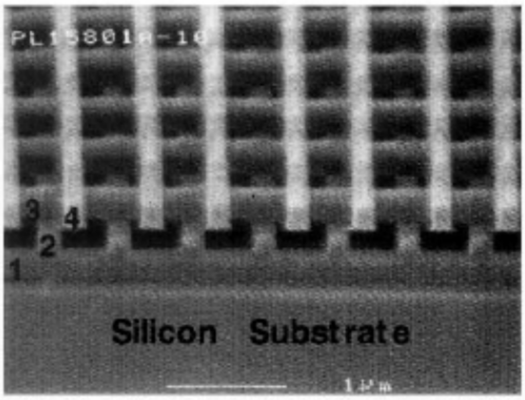


Fig. 17.4. SEM image of a 3D photonic crystal made of silicon [4].

If we design a 3D photonic crystal appropriately, there appears a frequency range where no electromagnetic eigenmode exists. Frequency ranges of this kind are called *photonic bandgaps*, since they correspond to bandgaps of electronic eigenstates in ordinary crystals. Moreover, if we introduce a disorder into the regular dielectric structure of the photonic crystal, we may obtain midgap modes whose eigenfunctions are strongly localized around the disorder. These modes are called *localized defect modes*. The density of states for a 3D photonic crystal with a photonic bandgap and a localized defect mode is schematically illustrated in Fig. 17.1(b). If the emission frequency of an atom or a molecule embedded in the photonic crystal lies just in the photonic bandgap, the spontaneous emission of a photon from its electronic excited state is completely forbidden, since there no exists photon in the gap. On the other hand, if the emission frequency coincides with the eigenfrequncy of the localized mode and the atom is located near the defect, the spontaneous emission is accelerated.

Another class of photonic crystals known as *photonic crystal slabs* [5–14] is illustrated in Fig. 17.5. Photonic crystals of this type are usually fabricated on a substrate made of a semiconductor or an insulator. They have been investigated energetically in recent years, because many sophisticated technologies such as electron beam lithography and thin-layer formation developed in the field of electronics and opto-electronics can be applied to their fabrication. Monolayers made of polymer micro-spheres [15–17] may also be regarded as photonic crystal slabs.

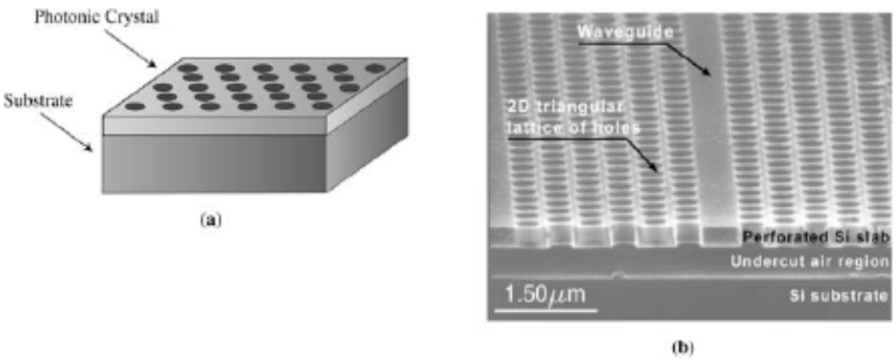


Fig. 17.5. (a) Schematic illustration of a photonic crystal slab and (b) the scanning electron micrograph of an actual specimen fabricated on a Si substrate [5].

17.2. PHOTONIC CRYSTALS FROM MACROPOROUS SILICON

Photons traveling in a periodic dielectric medium show interesting new behavior. The photons behave similar to electrons moving through the periodic potential in a crystal and their movement may be described by a photonic band structure. These periodic dielectric media are called Photonic Crystals (PCs) and have attracted much interest in the last decade [18 - 20].

Because of fabrication deficiencies, these predictions were first demonstrated by experiments in the microwave regime [21]. However, extending the application of PCs to the near infrared or visible spectrum demands very high precision on the fabrication of the structures. Since the periodicity determines the wavelength of the photonic band gaps, these structures must be fabricated with a precision much better than the wavelength, i.e., on a sub-□ m level. The translational symmetry is desired in all 3 directions of space, but already for the 2D case interesting new devices have been proposed [18 - 20].

Electrochemically-etched macroporous silicon exhibits pores with a diameter in the micro- to submicrometer range and a depth of up to the thickness of the silicon wafer. Because of the high uniformity, the low absorption coefficient and the high refractive index contrast, macroporous silicon can be used to fabricate two-dimensional photonic crystals with a complete band gap in the near or mid infrared spectral range [22 - 24].

17.2.1. Various Shapes of Macroporous Silicon

Macroporous silicon with its well defined pore shape is often used as a template material. Amples with straight pores and different aspect ratios were prepared and applied as a template to form polymers filled inside the pores or as a microreactor array for oxidation experiments.

For the membrane fabrication method a treatment in KOH is required to open the pore bottoms of the etched samples from the back side. However, the macroporous layer can be already released from the bulk silicon during the etching process. Therefore, the back side illumination intensity is strongly increased by a factor of 2.5 over a length of 2 μm at the end of the etching process (Fig. 17.6a). Now, neighboring pores are almost touching each other and the whole macroporous layer can be easily separated by a mechanical force from the substrate or by an oxidation step followed by an HF-dip (Fig. 17.6b). Beside the released freestanding macroporous silicon membrane the remaining bulk silicon is of interest, too. As can be seen in Fig. 17.6c the surface forms sharp spikes arranged in a regular lattice given by the initial lattice geometry. Since silicon is a very hard material, this micro 'fakir pillow' could be used as a micrometer sized structuring tool, e.g. as a stamp for imprint lithography.

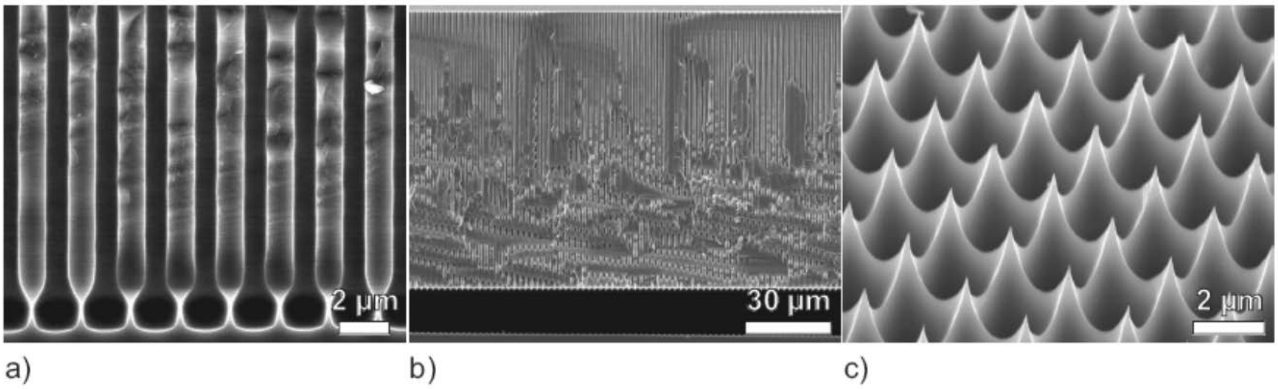


Fig. 17.6. SEM micrographs of macroporous silicon with 2 μm interpore distance in a square pore arrangement. a) Cross sectional view of straight pores with increased pore diameter at the pore bottom. b) Released layer of straight pores. The gap between the porous layer and the remaining bulk silicon at the bottom edge of the picture indicates the successful release of the porous membrane. c) Bird's eye view of the remaining bulk silicon surface after the macroporous layer has been released. The sharp silicon spikes can be used for instance as an imprint stamp.

To a certain extent tubes and cylinders are the inversion of pores. The easiest way for tube preparation can be realized by partial oxidation of a straight pore sample and a subsequent KOH etching process. Depending on the macroporous silicon treatment prior to the oxidation, microtubes with closed pore bottom (like a test tube) or open pore bottom (like a straw) can be realized. A back side view of closed SiO_2 tubes partly released from the silicon is shown in Fig. 17.7a.

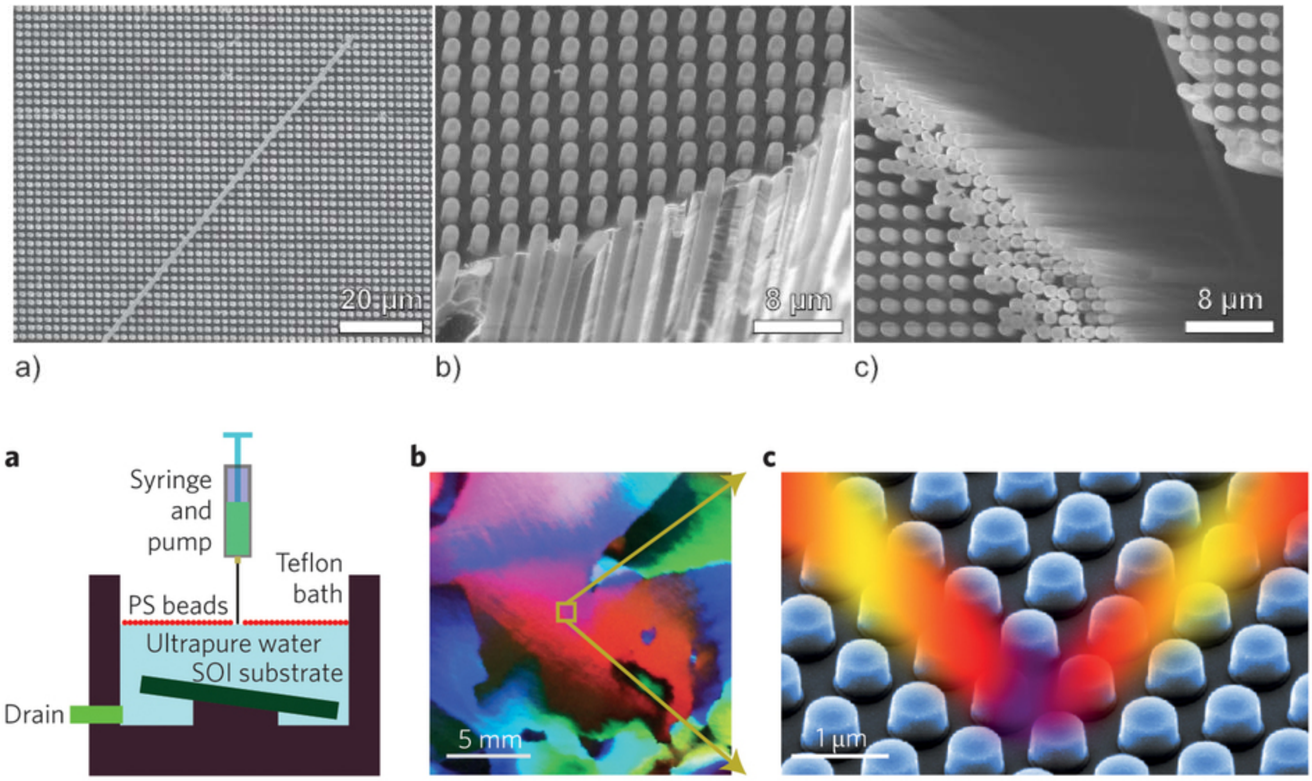


Fig. 17.7. SEM micrographs of macroporous silicon with 2 µm interpore distance in a square pore arrangement. a) An array of oxidized macroporous silicon tubes (bright spots) after partly removing the silicon from the back side with KOH. On top of this array a silicon dioxide tube is shown. b) Bird's eye view onto a bottom edge of this sample. The thickness of the tube's pore walls is about 100 nm and depends on the oxidation parameters time, temperature, and atmosphere. c) Fully released silicon dioxide microtubes after several hours of etching in KOH. The length of these tubes is 200 µm. Illustration of light beams movement.

In order to get a high etch contrast between silicon and silicon dioxide low concentrated KOH (2wt%) at room temperature was used. The etch ratio between the silicon <100> direction and silicon dioxide exceeds a factor of 10000 : 1. This ensures the complete release of the hollow silicon dioxide tubes from the silicon, even for tubes with high aspect ratios and thin pore walls (Figs. 17.7b and c). The length of the tubes is given by the length of the etched pores. The inner and outer diameter as well as the wall thickness are dependent on each other and determined by the initial pore diameter and the thickness of the grown silicon dioxide layer. A macroporous silicon pore diameter of 1 µm and a grown silicon dioxide layer with 100 nm thickness would result in an outer tube diameter of 1.1 µm.

17.2.2. 2D Photonic Crystals

Reparation of macroporous silicon. A detailed description of macroporous silicon formation can be found in [25, 26]. Here, we just give a short summary. First, an n-type silicon wafer with <100> orientation is prepatterned by standard photo-lithography. Subsequent alkaline etching produces inverted pyramids acting as initial pores. Under anodic bias and backside illumination, the wafer is then etched in hydrofluoric acid. The electronic holes generated by the illumination near the back surface diffuse through the whole wafer and promote the dissolution of silicon mainly at the pore tips. As a result, pores grow straight along the <100> direction with very high aspect ratio. The arrangement of these pores can be controlled by the lithographic mask, the diameter by the illumination intensity. By controlling these parameters, variations of pore diameters both between neighboring pores and with depth can be made negligible.

Most of the samples we report about were etched on 0.5 Ω×cm n-type FZ silicon substrates using a hexagonal pore arrangement. The pores have a center-to-center distance of 1.5 µm and a depth of 100 µm. The pore diameter after electrochemical etching is 0.9 µm. By subsequent oxidation/etching steps the pore diameter is increased up to 1.36 µm.

For optical investigations or applications of such structures the porous silicon has to be processed further. The two-dimensional PC has translational symmetry perpendicular to the pore axes. Therefore, analyzing the band structure requires the photons also to travel perpendicular to the pore axis. Investigations of the properties of defects demand access to the end of a waveguide and require removal of part of the porous silicon with a precision of about one pore lattice constant or about 1 µm. A special microstructuring technique was developed to meet these specifications [27]. First, the pore walls are passivated by a thermal oxide and a CVD nitride. Afterwards, an aluminum layer is sputtered onto the porous silicon and structured by conventional photo-lithography. Although the structure sizes of this second mask are in the 10 µm range, the precision of these structures and the alignment relative to the defects is better than 1 µm. In the opened window of the aluminum mask, the passivating oxide and nitride are removed by chemical etching. In a subsequent isotropic plasma etching process the porous silicon in the areas without passivation is etched leading to the desired bar structure. The quality of this process is demonstrated in Fig. 17.8. A bar of porous silicon is shown with about 22 layers of pores or about 33 µm width. The microstructuring technique leads to sharp edges. The whole transition for the 100 µm deep pores is within one pore layer! Overall, we get bars of porous silicon on the silicon substrate which are 100 µm high, 2–200 µm wide and several mm long. The bars are aligned with the designed defect structures in the porous silicon.

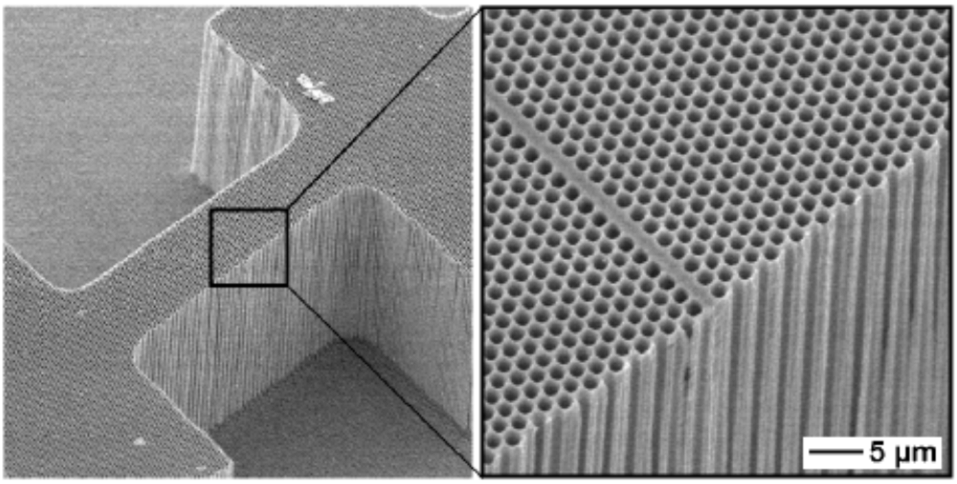


Fig. 17.8. Laterally-structured porous silicon sample with a defect line. The H-like structure facilitates the positioning of a fiber for coupling light in and out. Pore distance is 1.5 µm, height of the porous silicon is 100 µm.

17.2.3. Bulk Photonic Crystals

The processed macroporous Si samples described above are now well suited for investigations of the optical properties for light traveling perpendicular to the pores. The transmission measurements have bttn done on such bars for different pore diameters, polarizations, and directions. In Fig. 17.9 spectra are shown for the Γ - M direction, H-polarization, and two quite different diameters of a sample having 22 layers of pores and no defects. Depending on the pore diameter the center of the stop band due to the band gap shifts from 1650 cm^{-1} (6 µm wavelength) for a pore diameter of 1.06 µm up to 2800 cm^{-1} (3.6 µm wavelength) for 1.36 µm. Repeating this measurements for other directions and E-polarization, we find a complete gap centered at 3.2 µm for 1.36 µm pore diameter and no overlap of all the different gaps for the 1.06 µm sample in good agreement with theoretical predictions for such structures.

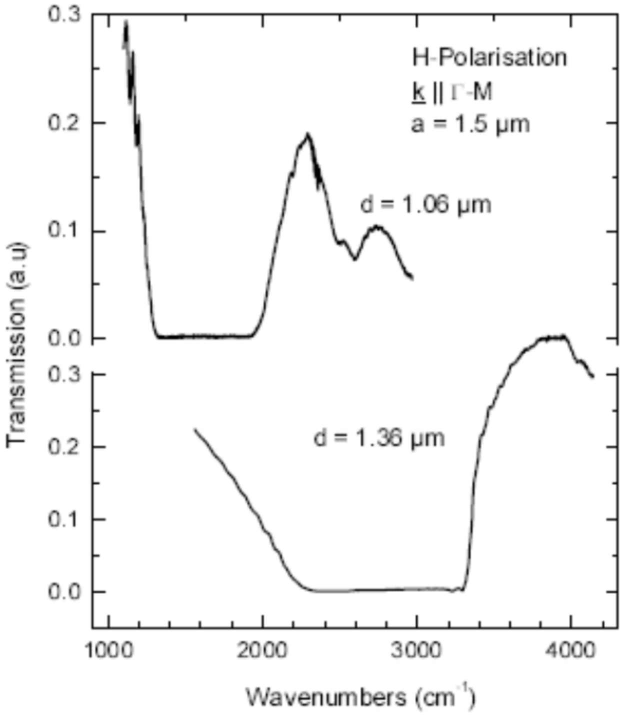


Fig. 17.9. Transmission spectra of two samples with different pore diameter *d* along the Γ - M direction, i.e., along the pore rows. Lattice constant *a* is 1.5 µm, the width of the porous silicon bar is 33 µm. Measurement is in H-polarisation, i.e., the magnetic field is parallel to the pore axes.

In Fig. 17.10 the resulting bandgaps of such measurements for a whole set of samples with varying pore diameters (symbols) are presented. For comparison the theoretical predictions from a plane wave expansion method [28] are shown as solid lines. For the lower filling factors the agreement is quite perfect. Only for the very high porosities there are some slight deviations due to the difficulty in preparing and handling these very fragile samples. Small differences between the real dimensions of the samples and the values used in the calculations are, therefore, responsible for the observed discrepancies [29].

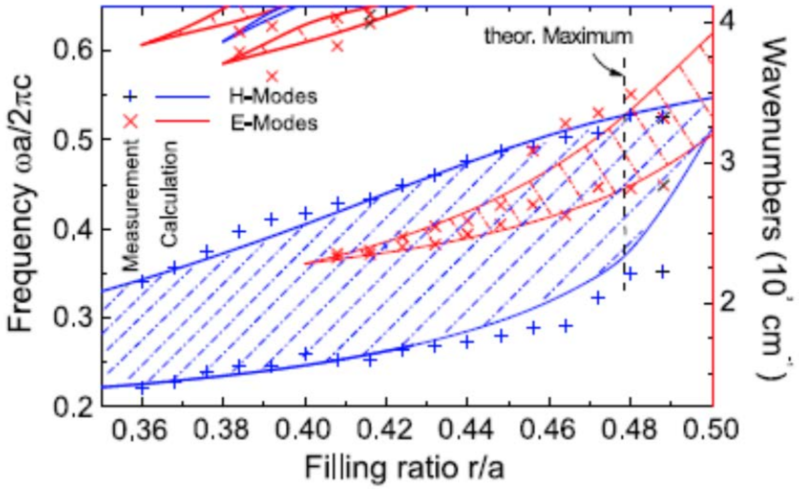


Fig. 17.10. Dependence of the observed bandgaps on polarisation and filling factor (gap map). Theoretical (solid lines) and experimantal values (symbols) show very good agreement.

These measurements were carried out using a Bruker IFS66 FTIR spectrometer with an MCT detector element of 0.5 mm diameter. In the band gap region, transmission was reduced by more than 2 orders of magnitude. However, this value reflected the limitation of the FTIR setup and was mainly determined by the nonlinearity of the MCT detector. Recently, measurements by a tunable laser on very thin samples showed a damping of about 10 dB per pore row [30] for frequencies laying in the band gap.

17.2.4. Waveguides

If some etch pits by using a suitable mask for the photo-lithography were omited, the electronic holes generated at the back side by illumination are consumed by Fig. 17.11. Different defect structures realized in macroporous silicon with 1.5 μm interpore distance. the neighboring pores without great influence on their position. They get somewhat larger because they collect more carriers. To minimize this effect, pores are grown with a small diameter and afterwards widened by pure chemical etching or by several oxidation/etching steps. The result of this procedure is a quite perfect structure with some missing pores at predefined positions, as shown in Fig. 17.12 for a single missing pore. These missing pores disturb the translational symmetry and for proper conditions lead to localized states in the forbidden spectral region. Arranging such defects in a line creates defect modes with transmission bands inside the bandgap. As propagation is forbidden in the surrounding medium, waveguides with very sharp bends should be possible according to theory. Different defect structures: linear and bent waveguide [31], Y-branch and microresonator have been analyzed in [29] (Fig. 17.11). More detail discussion ofbthe the impact of the linear defect onto the optical properties is given in [31].

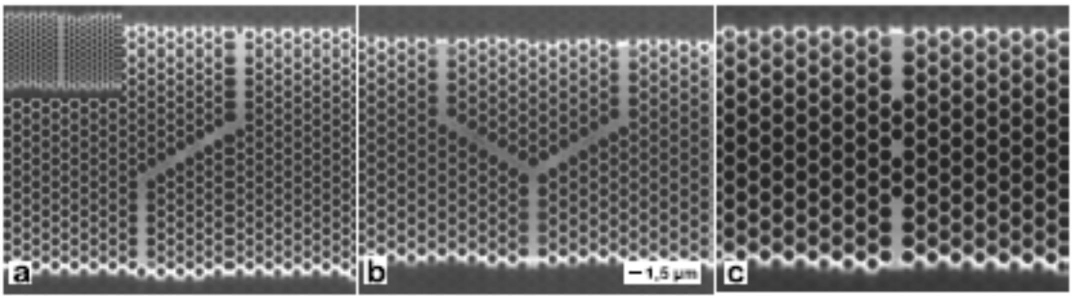


Fig. 17.11. Different defect structures realized in macroporous silicon with 1.5 μm interpore distance.

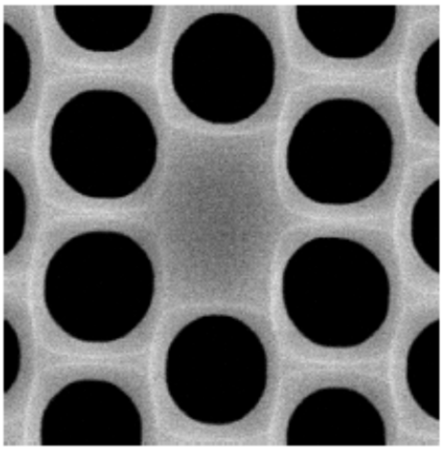


Fig. 17.12. SEM top view of the region around a missing etch pit after electrochemical pore growth and subsequent pore widening by oxidation/etching steps. The distance between the pores is 1.5 µm, pore diameters are 1.15 µm.

Band structure calculations show additional states all over the bandgap localized at the defect line. As it wase experimentally coupling in from a plane wave, only to the even symmetry states were coupled. The theoretical transmission spectrum is depicted in the upper part of Fig. 17.13. For the lower frequencies single mode transmission is expected. At the in- and the out-coupling facets the waveguide modes are partially reflected leading to Fabry-Perot resonances. At higher frequencies a small gap is expected where no even states are available. Above the gap there are again states available, but the density of these states is so high, that the resonances can no longer be resolved. The measured spectrum which is shown in the lower part of Fig. 17.13 shows a remarkable good agreement with these predictions. For the lower frequencies the expected single mode resonances are observed. There is a well pronounced stop band and a broad transmission band at higher frequencies. Again, the observed small discrepancies in finesse and stop band frequency can be well explained by the strong influence of small deviations in the real dimensions of structures.

The observed high finesse results from the low coupling efficiency between the strongly localized waveguide modes and the external plane waves. In a real device the waveguide facets could be avoided by connecting other optical ele ments directly. These linear defects act then as quite ideal waveguides where the light is very well confined within a few pore rows.

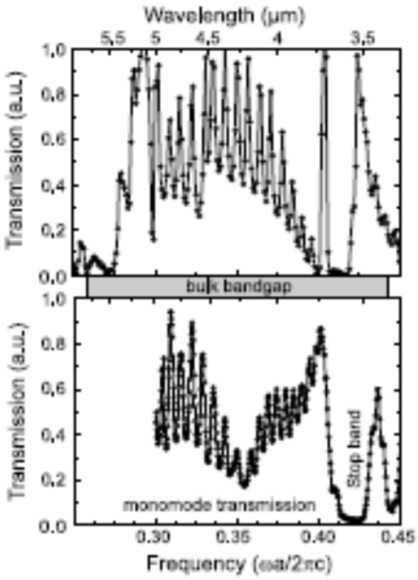


Fig. 17.13. Experimental transmission spectrum (bottom) of a linear waveguide structure compared to theoretical predictions (top) [31]. The measurement was carried out for similar structure as in Fig. 17.8 with a thickness of 18a with a = 1.5 µm and for H-polarization. The bulk bandgap is marked by a gray bar.

17.2.5.. 3D Crystals

According to the established growth model of Lehmann [24, 26], the current density at the pore tips is always equal to the critical current density j_{ps}. The porosity p is therefore determined by the ratio of the total current density to the critical current density, or for regular arrangements, where all pores have the same area A_{pore} the porosity is $p=A_{pore}/A_{cell} = j/j_{ps}$.

The total current is controlled by the illumination intensity. Therefore, variations of the pore diameter with depth can be achieved by controlling the light source. This concept has proven its validity in the past for compensating the variation of other growth parameters like temperature or the dilution of HF with increasing pore depth. For the application of macroporous silicon as 2D photonic crystals, this compensation is well established to get homogeneous pore diameters for pores several 100 µm deep [23, 32].

Samples were prepared from 5 Ω×cm substrates with a photolithographically defined hexagonal pattern of rtsu 4.2 µm pitch. A circular area with a diameter of 20 mm was exposed to the 6 wt-% HF at 10wC. The samples reported here were etched using a sawtooth-like current density.

The SEM cross-section image in Fig. 17.14 demonstrates the quality of this etching process. The resulting samples have the expected strong asymmetrically varying pore shapes as defined by the current profile. The sample is laterally homogeneously etched over the whole exposed area without noticeable defects. With increasing pore depth, the HF concentration at the pore tips and therefore also the critical current density j_{ps} and the etching speed v are reduced. If the sawtoothlike current density is applied on a linear time scale, this leads to a strong variation of about 30% in the length of a period from top to bottom for a 100 µm deep porous film. Using the reduction of growth speed from the homogeneous model [26] improves this effect quite remarkably. Up to 25 periods could be etched without noticeable deviation of the linear fit leading to a total thickness of over 200 µm [33].

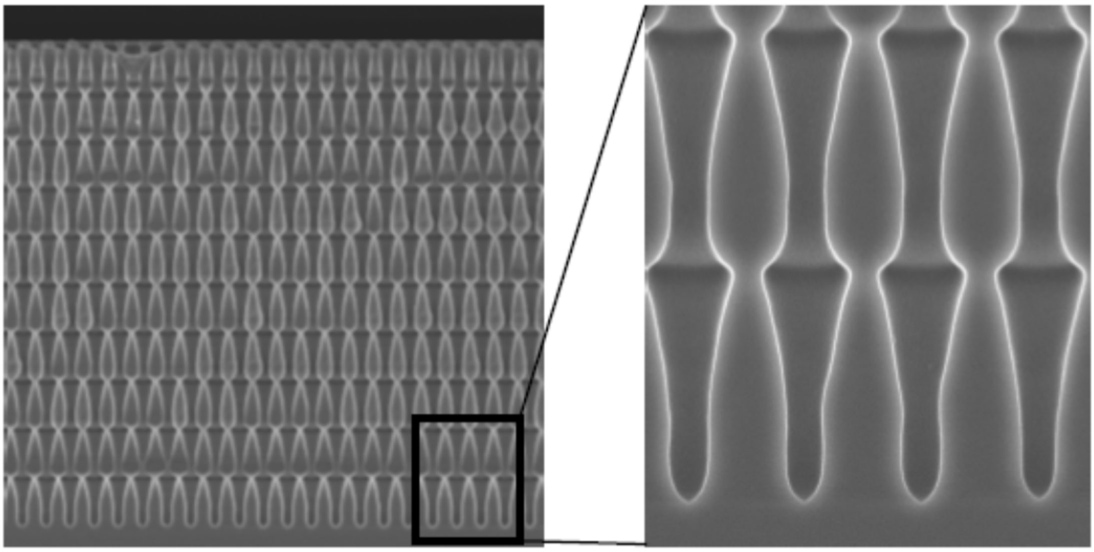


Fig. 17.14. SEM cross-section image for a sample etched with 10 periods of modulated light.

In Fig. 17.15 the transmission spectrum in growth direction of the pores for a sample with a period of l_z=7.2 µm is shown. Two strong stop bands can be observed near 320 cm^{-1} and near 610 cm^{-1}. As the lateral period is significantly smaller than the period in the growth direction, a first approximation of the optical behavior can be obtained by using an effective medium model. From the pore diameter, as measured from the SEM cross section images, the depth-dependence of the porosity and then the effective refractive index using the Bruggeman formalism have been determined [34]. The calculated transmission spectrum for a multilayer model [35] using 10 slabs for each period and 8 periods in total is shown as a dotted line in Fig. 17.15. The positions of the stop bands are very well reproduced by this simple approximation. The Fabry-Perot interferences from the reflections at the front and back surfaces, which are very strong in the calculated spectrum, were not resolved in the experiment. The high transmission around 550 cm^{-1} and below 180 cm^{-1} are artifacts due to the low background intensity of the spectrometer.

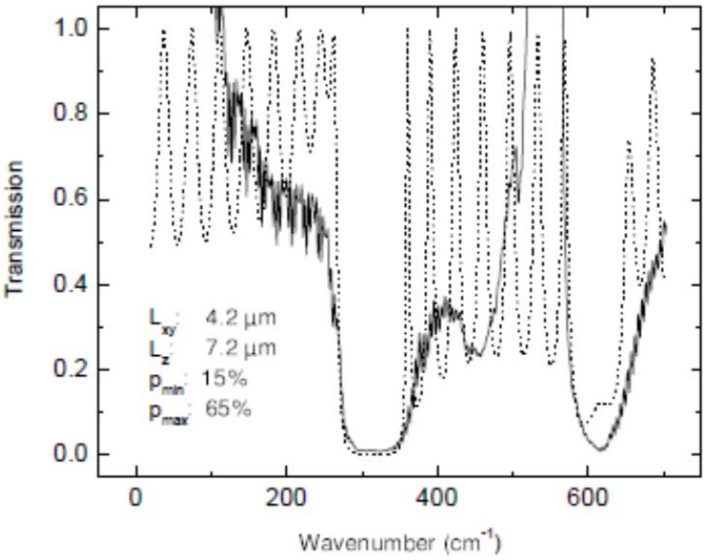


Fig. 17.15. Transmission spectrum for light travelling parallel to the pores axes (straight line) together with the calculated spectrum from a 1D Bruggeman approximation (dotted line). The lattice constant l_{xy} in the xy - plane is 4.2 µm, in z-direction l_z = 7.2 µm. Porosity varies from p_{min} = 0.15 to p_{max} = 0.65.

These samples very impressively demonstrate the ability to generate real 3D PCs by modulating the backside illumination. Although the structures produced so far, do not exhibit a complete bandgap, they have strongly nonlinear dispersion relations in all directions, a behavior which will be very useful in nonlinear optic experiments to fulfill phase matching conditions. Especially for applications like mixing experiments of beams with different wavelength, the freedom to design the third z -direction independently from the periodicity in the xy -plane will provide very high flexibility.

17.3. Outlook

To obtain bandgaps around 1.3 µm a downscaling of the above described triangular pore lattice is necessary. Therefore the pitch was adjusted to a= 0.5 µm. The pores fabricated had a radius r = 0.18 µm and a depth of 100 µm. During the process they were widened to r = 0.215 µm since a complete bandgap opens up at ratios $r/a \geq 0.4$. To investigate the 2D-bandgaps in the near IR reflection measurements were performed. For this purpose the samples were cleaved along the Γ-K direction to gain access to the side walls of the pores and to get a cleancut interface. For the reflection measurements an IR microscope equipped with a polarizer and a mirror objective with an opening angle of 30 degrees was used similar to Ref. [19]. The spectra were detected by a FTIR-spectrometer containing a tungsten lamp, CaF_2 –beam splitter and a MCT-detector. The reflections for E- and H-polarisation were measured along the Γ-M direction using a gold mirror with a nominal reflectivity of 98% as reference. The resulting sample spectra are shown in Fig. 17.15. Due to the relatively wide opening angle the measured signal comprises contributions ranging from normal incidence reflection up to 30 off axis and off-plane incidence reflections. The first order bandgap investigated here weakly depends on the direction of incidence [36]. Therefore bandstructure calculations performed along the wavevector path Γ-M-K-Γ can still be employed as a good approximation for the interpretation of the reflection spectra. They were computed for $r/a \geq 0.425$ using 967 plane waves. The shaded spectral ranges represent theoretically expected regions of high reflectivity. They mainly coincide with the bandgaps along the Γ-M direction. Figure 17.16 reveals good agreement between theoretical predicted ranges of total reflection and experimentally determined high reflectivity regions. Finally, from the calculated bandstructure, a complete bandgap for E- and H-polarisation can be derived for the spectral range from 1.22 µm to 1.3 µm, thus incorporating the wavelengths of the 2nd telecommunication window.

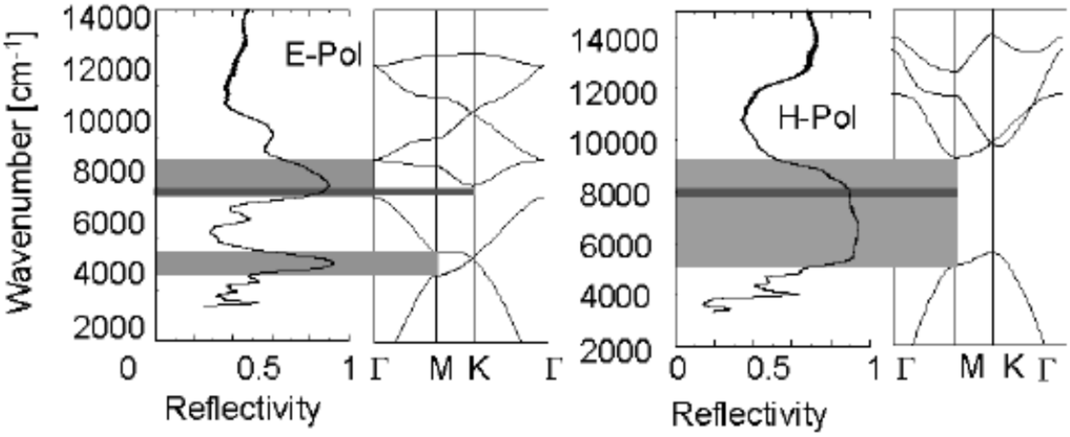


Fig. 17.16. Measured reflectivity for S -M direction and comparison with bandstructure; (right) E-Polarization, (left) H-Polarization. Grey shaded ranges correspond to regions of high reflectivity. Dark shaded range shows the complete bandgap around 8000 cm^{-1} (=1.25 µm).

One possible next step is towards an active device, in the simplest case an optical switch, but also an optical transistor can be envisaged. A possible realization would be the introduction of a material into the pores which changes its refractive index as a function of an applied field (electrical, magnetical, mechanical, etc.). Recently, as a proof of principle the feasibility of such an optical switch by introducing a liquid crystal into the pores and switching it thermally has been shown for the first time [37].

17.4. CONCLUSION

Silicon based photonics is under intensive research and development. Two-dimensional photonic crystals based on macroporous silicon with a lattice constant of 1.5 µm exhibiting a complete gap from 3 to 4 µm wavelength have been prepared. Transmission measurements show good agreement with theoretical predictions for a wide range of pore diameters. Different defects like waveguides with or without bends have been fabricated. For the linear waveguide the good agreement between theoretical calculations and measured transmission spectra is demonstrated. Periodicity in the third direction is achieved by modulated light intensity leading to real 3D photonic crystals. The possibility to grow macropores with a lattice constant of 0.5 µm has been shown. The precision of the achieved samples and the presented optical data prove the suitability of electrochemically etched macroporous silicon as a candidate for new optical devices based on PCs in the near and mid infrared spectral range.

References to Chapter 17.

1. K. Sakoda, Optical Properties of Photonic Crystals, Springer-Verlag Berlin Heidelberg, 2005.

2. E. M. Purcell: Phys. Rev. **69**, 681 (1946)

3. K. Yoshino, Y. Shimoda, Y. Kawagishi, K. Nakayama, M. Ozaki: Appl. Phys. Lett. **75**, 932 (1999).

4. J. G. Fleming, Shawn-Yu Lin: Opt. Lett. **24**,49 (1999).

5. M. Lonˇcar, D. Nedeljkovirc, T. Doll, J. Vuˇckovirc, A. Scherer, T. P. Pearsall: Appl. Phys. Lett. **77** 1937 (2000).

6. P. L. Gourley, J. R. Wendt, G. A. Vawter, T. M. Brennan, B. E. Hammons: Appl. Phys. Lett. **64**, 687 (1994).

7. M. Kanskar, P. Paddon, V. Pacradouni, R. Morin, A. Busch, J. F. Young, S. R. Johnson, J. Mackenzie, T. Tiedje: Appl. Phys. Lett. **70**, 1438 (1997).

8. S. Fan, P. R. Villeneuve, J. D. Joannopoulos, E. F. Schubert: Phys. Rev. Lett. **78**, 3294 (1997).

9. D. Labilloy, H. Benisty, C. Weisbuch, T. F. Krauss, R. M. De La Rue, V. Bardinal, R. Houdrre, U. Oesterle, D. Cassagne, C. Jouanin: Phys. Rev. Lett. **79**, 4147 (1997).

10. B. D'Urso, O. Painter, J. O'Brien, T. Tombrello, A. Yariv, A. Scherer: J. Opt. Soc. Am. B **15**, 1155 (1998).

11. D. Labilloy, H. Benisty, C.Weisbuch, C. J. M. Smith, T. F. Krauss, R. Houdrre, U. Oesterle: Phys. Rev. B **59**, 1649 (1999).

12. O. Painter, R. K. Lee, A. Scherer, A. Yariv, J. D. O'Brien, P. D. Dapkus, I. Kim, Science **284**, 1819 (1999).

13. S. G. Johnson, S. Fan, P. R. Villeneuve, J. D. Joannopoulos: Phys. Rev. B **60**, 5751 (1999).

14. S. Kuchinsky, D. C. Allan, N. F. Borrelli, J. C. Cotteverte: Opt. Commun. **175**, 147 (2000).

15. M. Haraguchi, T. Nakai, A. Shinya, T. Okamoto, M. Fukui, T. Koda, R. Shimada, K. Ohtaka, K. Takeda: Jpn. J. Appl. Phys. **39**, 1747 (2000).

16. T. Fujimura, T. Itoh, A. Imada, R. Shimada, T. Koda, N. Chiba, H. Muramatsu, H. Miyazaki, K. Ohtaka: J. Lumin. **87–89**, 954 (2000).

17. H. T. Miyazaki, H. Miyazaki, K. Ohtaka, T. Sato: J. Appl. Phys. **87**, 7152 (2000).

18. *Photonic Bandgaps and Localization*, edited by C. M. Soukoulis, NATO ASI Series, Ser. B, Vol. **308**, (Plenum Press, New York 1993).

19. J. D. Joannopoulos, R. D. Meade, and J. N. Winn, *Photonic Crystals*, (Princeton University Press, New Jersey 1995).

20. *Photonic Band gap materials*, edited by C. M. Soukoulis, NATO ASI Series, Ser. E, Vol **315**, (Kluwer Academic Publishers, London 1996).

21. E. Yablonovitch, T. J. Gmitter, and K. M. Leung, *Phys. Rev. Lett.* **67**, 2295 (1991).

22. U. Gruning, V. Lehmann, S. Ottow, and K. Busch, *Appl. Phys. Lett.* **68**, 747 (1996).

23. A. Birner, U. Gr¨uning, S. Ottow, A. Schneider, F.M¨uller, V. Lehmann, H. Foll, and U. Gosele, *Phys. Stat. Sol.* (a) **165**, 111 (1998).

24. V. Lehmann and U. Gr¨uning, *Thin Solid Films* **297**, 13 (1997).

25. V. Lehmann and H. F¨oll, *J. Electrochem. Soc.* **137**, 653 (1990).

26. V. Lehmann, *J. Electrochem. Soc.* **140**, 2836 (1993).

27. S. Ottow, V. Lehmann, and H. F¨oll, *J. Electrochem. Soc.* **143**, 385 (1996).

28. K. Sakoda, *Phys. Rev. B* **51**, 4672 (1995).

29. A. Birner, A.-P. Li, F. M¨uller, U. G¨osele, P. Kramper, V. Sandoghdar, J. Mlynek, K. Busch, V. Lehmann, to be published in *Mater. Sci. Semicon. Proces.* (2000).

30. S. W. Leonard, H. M. van Driel, K. Busch, S. John, A. Birner, A.-P. Li, F. Muller, U. Gosele,and V. Lehmann, *Appl. Phys. Lett.* **75**, 3063 (1999).

31. S.W. Leonard, H. M. van Driel, A. Birner, U. Gosele, P. R. Villeneuve, submitted to *Opt. Lett.*

32. F. Muller, A. Birner, U. G¨osele, V. Lehmann, S. Ottow,and H. Foll, *J. Porous Mater.* **7**, 201 (2000).

33. F. Muller, A. Birner, J. Schilling, U. G¨osele, Ch. Kettner and P. H¨anggi, *Phys. Stat. Sol.* (a).

34. D. A. G. Bruggeman, *Ann. Physik.*, Leipzig, **24**, 636 (1935).

35. F. Abel`es, *Ann. de Physique* **5**, 596 (1950).

36. S. Rowson, A. Chenokov, C. Cuisin, J.-M. Lourtioz, *IEEE Proc.-Optoelectron.* **145**, 403 (1998).

37. S. W. Leonard, J.P. Mondia, H.M. van Driel, O. Toader, S. John, K. Busch, A. Birner, U. Gosele and V. Lehmann, *Phys. Rev. B* **61**, R2389 (2000)

Chapter 18. Notubes and other implant materials.(A. Evtukh)

25.1. Introduction
25.2. Carbon Nanotubes
    25.2.1. Molecular structure
    25.2.2. Synthesis of CNT
    25.2.3. Electrical Conduction in Nanotubes
    25.2.4. CNTs in Medicine and Biology
25.3. Nanotubes for Medical Implants
25.4. Nanotubes for implant surfaces: The biomimickingeffect for osseointegration
25.5. Nanotubes for Implant surfaces: The antimicrobialand anti-inflammatory properties
25.6. Nanotubes for implant surfaces: The biomimickingeffect for osseointegration
25.7. Hybrid Micro/Nano-Topography of a TiO_2Nanotube-Coated Implant
25.8. Potential Hazards of Using Nanotubes
25.8. Conclusion and Future Prospects
References to Chapter 25.

25.1. Introduction

With ever-growing interest and rapid development, the field of biomedical engineering has quickly assumed great importanceas one of the pillars of modern science. Especially significant has been the stress given on research and developmentin the field of orthopedics with implants being a subject of much interest. Among much new advancement in the field ofdevelopment and design of orthopedic implants, one of the most important ones has been the inculcation of nanotechnologyfor more viable and reliable implant design. The application of nanotechnology in implants has seen the extensive use ofnanotubes i.e. nanoporous structures with tubular shape. There are some of the unique features,which make nanotubes an interesting addition to the already diversifying field of implant design.Nanotubes have been usedin orthopedic implants in three principal aspects[1]: 1) The use of drug-loaded nanotubes on the implant surface as novel reservoirsfor systematic and controlled drug release in order to act as prophylactic or to prevent certain specific post-implantationinfections. 2)The use of certain nanotubular structures with high mechanical properties as reinforcing agent in conjunctionwith bone-simulating bio-active materialsto create durable prosthetics. 3) The use of metal oxide or alloy nanotubes on implantsurface promoting greater osseointegration and cell viability due to the alteration from a slightly hydrophobic surfaceto an extremely hydrophilic surface. Also, owing to prevalent concerns regarding the cytotoxicity and malignancy inducingcharacteristics of certain nanotubes, the safety aspect forthe use of nanotubes in the orthopedic implants is also importent.

Implants have been known to be used for the aid of normalhuman body functioning since the middle ages[2]. The installedmedical devices ranged from nails to screw to plateto entire limbs made out of wood or metals. Common applicationsof implants are for orthopedic reconstruction likerepairing fractures[3], providing joint arthrodesis[3], fixingnon-unions[3], complete or partial joint arthroplasty reconstructionof the spinal cord[3]. These devices, among ahost of different functions, primarily aim to enable properalignment of bones and facilitate usual functioning in thecourse of physiologic loading by giving the structure mechanicalstability. Thus the implants provide relief from pain andrender natural use of the replaced/reinforced body part. Thisis considered to be critical for the process of recovery andresumption of normal functioning of the body parts[4].Thebiological aspects of bone repair are assisted to an extent bydecreasing unwanted shear stress due to greater stability fosteredin the vicinity of the bone fractures. Also, devices thatminimize the extent of micromotion at the junctions betweenbone and implants in apatite-free joint replacements, andundesired motion between opposed bone surfaces in spinalreconstruction aid in bone formation and remodeling[5,6].The ultimate clinical outcome is squarely dependent on boththe mechanical and biological aspects of bone healing, whichare also very closely interlinked. Orthopedic implants havehistorically been designed as mechanical installation withconsideration of the biological repercussions of the implantbeing sidelined as byproducts of stable fixation, internally orexternally, of the fixture to the ambient bones or soft tissues.Especially in case of fracture fixation it has been largely acceptedthat bones, if provided with adequate support and stabilizationare prone to self-heal. This approach is indicativeof a rather dangerous notion however as only in the US peryear there are around 100,000 cases of nonunion and in excessof 500,000 fractures with unusually slow union[7]. Osseointegrationof the implant with the surrounding bone is notalways possible in cementless joint replacements ultimatelyculminating in implant migration and possible loosening[8].

Physicochemical compatibility and mechanical stability arespecifically important in the case of bone implants. Orthopedicimplants use distinct materials because of specific advantagesthose materials have for particular applications. A materialwhich possesses all the desirable properties at once has provento be rather elusive. Ceramics like aluminum oxide or zirconiumoxide and metallic alloys such as nickel-titanium are usuallyselected due to their high stiffness and strength for hardtissue replacement. But a potential pitfall of this application ispossible bone atrophy resulting from substantial difference betweenstiffness of the implant and that of the host tissue whichleads to the phenomenon of stress-shielding[9]. Additionally,these materials are bio inert in nature since active stimulate ofbone-formation process is not accomplished by them. Moreover,introduction of implants into the body brings with it therisk of microbial infection more so in case of fixation of jointrevisionsurgeries and open-fractured bones implant infectionsare caused by bacterial attachment to the surface of theimplant and consequent biofilm production in the vicinity ofthe installed implant[10]. Orthopedic devices need to foster integrationbetween host tissues and the implant surface as wellas oppose colony formation by microbes. The conundrum liesin the fact that the biomaterial surfaces which allow greater degreesof implant cell integration, proliferation and growth alsoprovide congenial environment for microorganisms whichhave similar attachment mechanisms as host cells. Moreover,the implant layers which are specifically intended to resist bacterialinfestation and creation of biofilm are less likely to haveeffective osseoconductive and osseointegrative properties.These various problems plaguing the effective installation anduse of orthopedic implants have been recently addressed effectivelyand extensively through the emerging field of Nanotechnology.Nanotubes specifically have been found to beof immense usefulness in this context.

25.2. Carbon Nanotubes

Carbon is arguably the most fascinating element in the periodic table. It is the base for DNA and all life on Earth. Carbon can exist in several different forms. The most common form of carbon is graphite, which consists of stacked sheets of carbon with a hexagonal structure. Under high pressure diamond is formed, which is a metastable form of carbon.

The discovery of fullerenes[11], carbon nanotubes and graphene initiated a new era in carbon chemistry. In addition tothe two well-known allotropes, graphite and diamond, alarge number of carbon structures were synthesized[12-18].Thefullerene molecule (primarily the most abundant fullerene,C_{60}), the polymerized, and doped fullerenes will be first considered.

A related quasi-one-dimensional form of carbon, carbon nanotubes, have been known for several decades[19] and the single walled nanotubes since 1993[20,21]. These can be formed from graphene sheets which are rolled up to form tubes, and their ends are half spherical in the same way as the fullerenes(Fig. 25.1.). The electronic and mechanical properties of metallic single walled nanotubes have many similarities with graphene.

It was well known that graphite consists of hexagonal carbon sheets that are stacked on top of each other, but it was believed that a single sheet could not be produced in isolated form such that electrical measurements could be performed. It, therefore, came as a surprise to the physics community when in October 2004, Konstantin Novoselov, Andre Geim and their collaborators[22]showed that such a single layer could be isolated and transferred to another substrate and that electrical characterization could be done on a few such layers[23]. The single layer of carbon is what we call graphene.

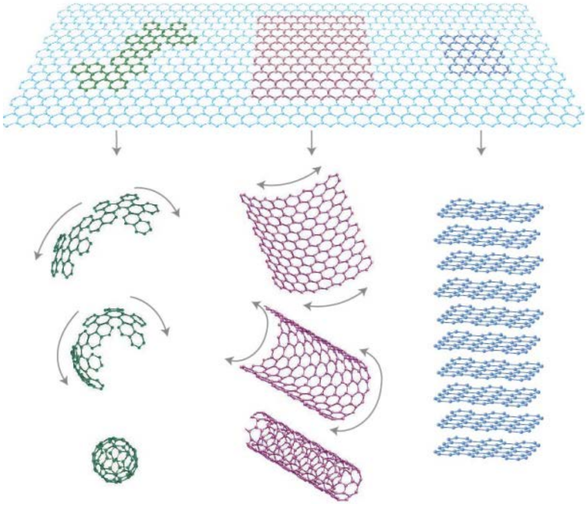


Fig.25.1. C_{60} fullerene molecules, carbon nanotubes, and graphite can all be thought of as being formed from graphene sheets, *i.e.* single layers of carbon atoms arranged in a honeycomb lattice[24].

25.2.1. Molecular structure

If a long strip ofgraphene is cut out and bent into a cylinder, so that loosecarbon bonds match up, we get a nanotube. The matchingcarbon atoms can be labeled by a chiral vector $\mathbf{C_h}$. Thisvector is represented by two basic vectors ($\mathbf{a_1}$ and $\mathbf{a_2}$) andby two integers (n_1 and n_2); $\mathbf{C_h}=n_1\mathbf{a_1}+n_2\mathbf{a_2}$(Fig. 25.2).Accordingly, the structure of every nanotube is unambiguouslycharacterized by the $\mathbf{C_h}$ vector; in a shorthandnotation $\mathbf{C_h} = (n_1, n_2)$, so that $n_1 \geq n_2$. Based on thepatterns visible on the nanotube they may be called‘‘zigzag’’ (for $n_2=0$), ‘‘armchair’’ (for $n_1=n_2$), or‘‘chiral’’ (for any other, see, for example, Fig. 25.2). Thediameter of the nanotube is given by $d = (a/\pi)[3(n_1^2 + n_1 n_2 + n_2^2)]^{1/2}$, where a=0.144 nm is the carbon–carbon distance. There is no upper limit on the size of thenanotube, although at large diameter the structurebecomes very soft (except for multiwalled nanotubes(MWNT)). For small diameter the carbon–carbon bonds are more strained, and the thinnest stablenanotube corresponds to $\mathbf{C}_h$=(5,5) (Fig. 25.3).

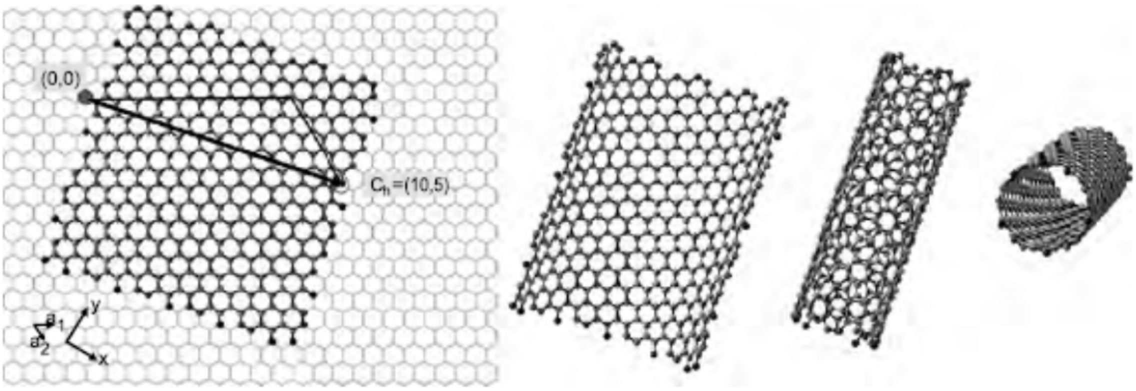


Fig. 25.2.Schematic illustration of deriving a nanotube structure from graphene. The unit vectors $\mathbf{a}_1$ and $\mathbf{a}_2$, and the chiral vector $\mathbf{C}_h$ areindicated in the first panel. Bending the graphene strip results in the nanotube shown in the last panel.

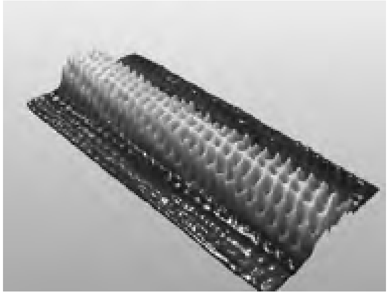


Fig. 25.3.Scanning tunneling microscope image of a nanotube.The peaks correspond to individual carbon atoms; the chiralityof the structure is quite visible [25].

The nanotubes derived from a single graphene strip arecalled single-walled nanotubes (SWNT)[15, 16]. If a set ofsubsequently smaller diameter nanotubes are stacked likeconcentric cylinders, a MWNT is obtained(Fig. 25.4). The distancebetween the carbon layers in an MWNT is about the sameas the distance between the graphene layers in graphite.MWNTs are produced more commonly, and the very firstreport of nanotubes by Iijima was about MWNTs[14].

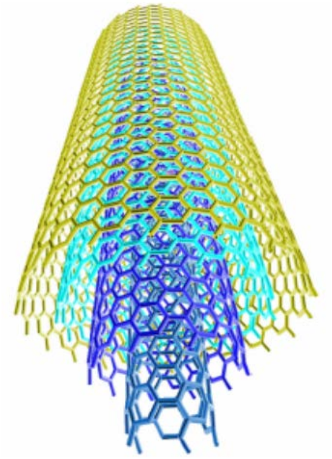


Fig. 25.4.Multiwalled nanotubes(MWNT).

Nanotubes generally have a length to diameter ratio of about 1000 so they can be considered as nearly one-dimensional structures. More detailed, a SWNT consists of two separate regions with different physical and chemical properties. The first is the sidewall of the tube and the second is the end cap of the tube. The end cap structure is similar to or derived from a smaller fullerene, such as C_{60}.

C-atoms placed in hexagons and pentagons form the end cap structures. It can be easily derived from Euler's theorem that twelve pentagons are needed in order to obtain a closed cage structure which consists of only pentagons and hexagons. The combination of a pentagon and five surrounding hexagons results in the desired curvature of the surface to enclose a volume. A second rule is the isolated pentagon rule that states that the distance between pentagons on the fullerene shell is maximised in order to obtain a minimal local curvature and surface stress, resulting in a more stable structure. The smallest stable structure that can be made this way is C_{60} the one just larger is C_{70} and so on. Another property is that all fullerenes are composed of an evennumber of C-atoms because adding one hexagon to an existing structure means adding two C-atoms.

The other structure of which a SWNT is composed is a cylinder. It is generated when a graphene sheet of a certain size that is wrapped in a certain direction. As the result is cylinder symmetric we can only roll in a discreet set of directions in order to form a closed cylinder (Fig. 25.2). Two atoms in the graphene sheet are chosen, one of which servers the role as origin. The sheet is rolled until the two atoms coincide. The vector pointing from the first atom towards the other is called the chiral vector and its length is equal to the circumference of the nanotube (Fig. 25.5).The direction of the nanotube axis is perpendicular to the chiral vector. SWNTs with different chiral vectors have dissimilar properties such as optical activity, mechanical strength and electrical conductivity.

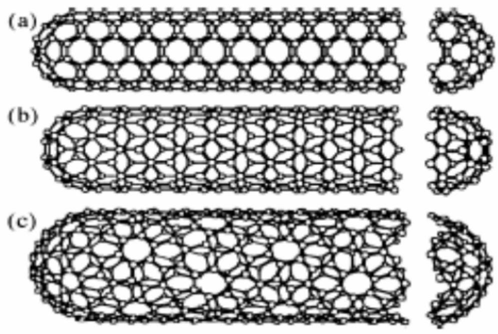


Fig. 25.5.Some SWNTs with different chiralities. The difference in structure is easily shown at the open end of the tubes: a) armchair structure, b) zigzag structure, c) chiral structure.

A nanotube rope[16] is a parallel arrangement ofSWNTs in a crystalline order, as shown in Fig. 25.6. Therope is held together by weak van der Waals forces (justlike graphite), but it is extremely strong along the long thedirection. SWNT strands as long as a few centimetershave been produced[27].

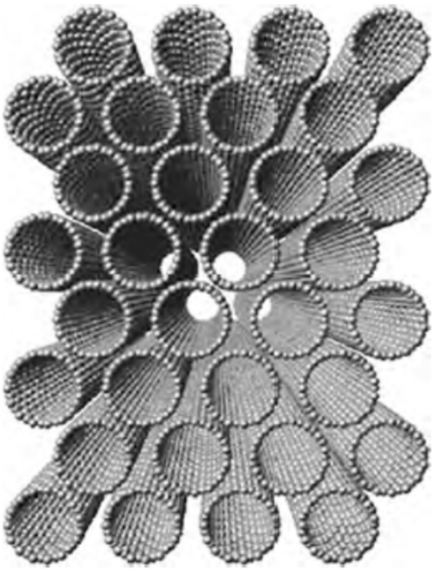


Fig. 25.6.Cross-sectional view of a nanotube rope.

After ideal structures without flaws, we discuss the possible desirable or undesirable defects. Deformations, such as bends and nanotube junctions, are introduced by replacing a hexagon with a heptagon or pentagon. Deformations can be inward or outward and, among others, electrical properties are seriously changed by these deformations. Another class of defects is caused by impurities that are built in during or after the nanotube growth process; Compounds that can be incorporated into the structure are for example catalyst particles.

Multi Walled Nanotubes (MWNT) can be considered as a collection of concentric SWNTs with different diameters. The length and diameter of these structures differ a lot from those of SWNTs and, of course, their properties are also very different(Fig. 25.7). In Fig. 25.7carbon cones are also shown. It can be considered as a gradual transition from a large diameter to a smaller one without defects in the wall of the cone but with fewer pentagons in the end cap.

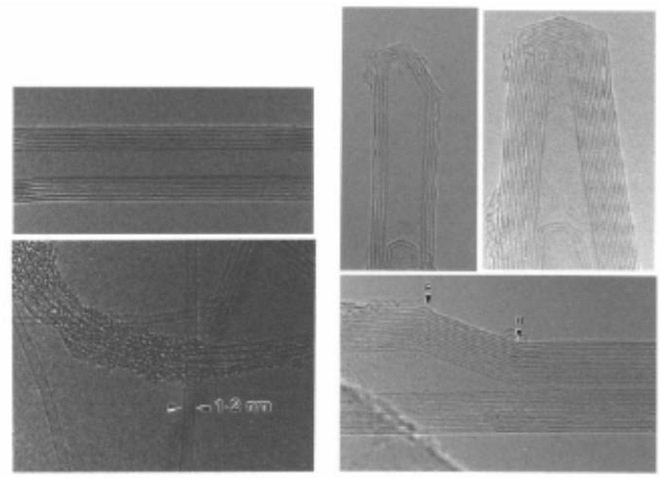


Fig. 25.7.Different structures of MWNTs. Top-left: cross-section of a MWNT the different walls are obvious, they are separated by 0.34nm. Rotation around the symmetry axis gives us the MWNT. Top-right: Symmetrical or non-symmetrical cone shaped end caps of MWNTs. Bottom-left: A SWNT with a diameter of 1.2nm and a bundle of SWNTs covered with amorphous carbon. Bottom-right: A MWNT with defects. In point P a pentagon defect and in point H a heptagon defect[28].

Introduction of defects can also result in various new structures such as Y-branches T-branches or SWNT junctions(Fig. 25.8).Under certain circumstances, these defects can be introduced in a ‘controlled’ way. These defects result in special structures that will have other, but even more interesting, properties than their original forms.

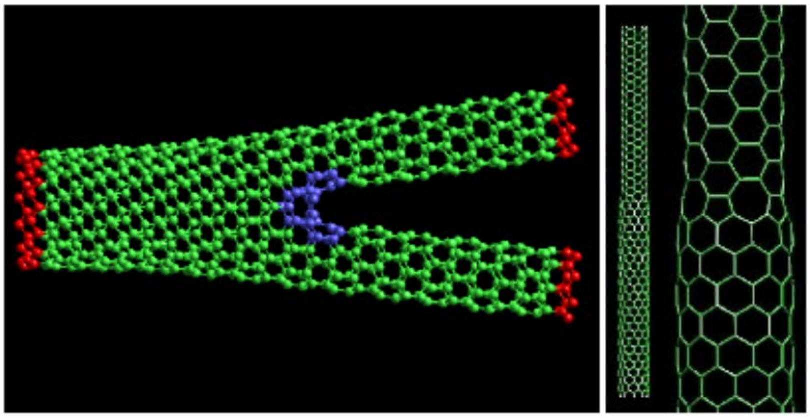


Fig. 25.8.Left: A Y-branch, the defects are marked in blue. Right: A transition from a metallic to a semiconducting SWNT. The change is made by insertion of pentagons and heptagons.

A final type of interesting structures is the so-called peapods, carbon nanotubes with C_{60} molecules enclosed in the nanotube.

25.2.2. Synthesis of CNT

Carbon nanotubes are generally produced by three main techniques, arc discharge, laser ablation andchemical vapour deposition. Scientists are researching more economic ways to produce thesestructures. In arc discharge, a vapour is created by an arc discharge between two carbon electrodeswith or without catalyst. Nanotubes self-assemble from the resulting carbon vapour. In the laserablation technique, a high-power laser beam impinges on a volume of carbon –containing feedstockgas (methane or carbon monoxide). At the moment, laser ablation produces a small amount of cleannanotubes, whereas arc discharge methods generally produce large quantities of impure material. Ingeneral, chemical vapour deposition (CVD) results in MWNTs or poor quality SWNTs. The SWNTsproduced with CVD have a large diameter range, which can be poorly controlled. But on the otherhand, this method is very easy to scale up, what favours commercial production.

Growth mechanism.The way in which nanotubes are formed is not exactly known. The growth mechanism is still a subjectof controversy, and more than one mechanism might be operative during the formation of CNTs. Oneof the mechanisms consists out of three steps. First a precursor to the formation of nanotubes andfullerenes, C_2, is formed on the surface of the metal catalyst particle. From this metastable carbideparticle, a rodlike carbon is formed rapidly. Secondly there is a slow graphitisation of its wall. Thismechanism is based on in-situ TEM observations(Fig. 25.9)[29].

The exact atmospheric conditions depend on the technique used later on these will be explained foreach technique as they are specific for a technique. The actual growth of the nanotube seems to be thesame for all techniques mentioned.

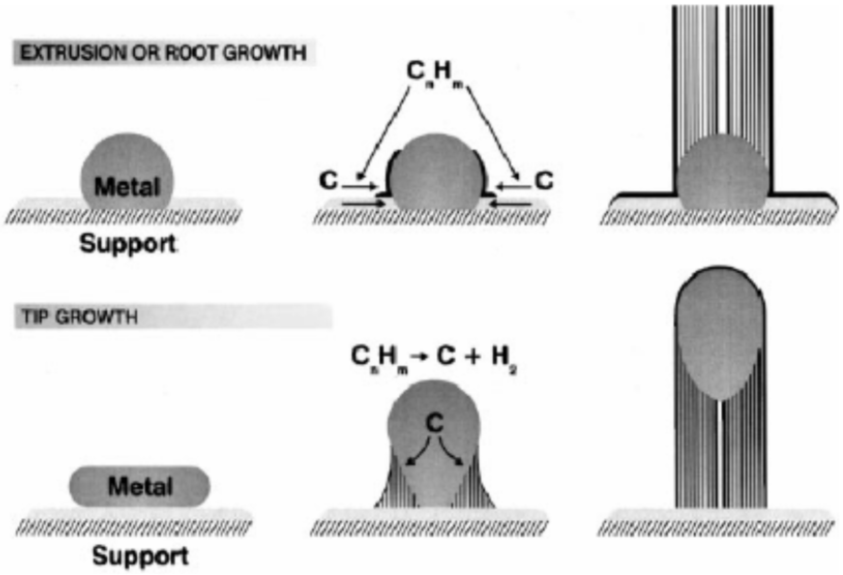


Fig. 25.9.Visualisation of a possible carbon nanotube growth mechanism.

There are several theories on the exact growth mechanism for nanotubes. One theory[30] postulates thatmetal catalyst particles are floating or are supported on graphite or another substrate. It presumes thatthe catalyst particles are spherical or pear-shaped, in which case the deposition will take place on onlyone half of the surface (this is the lower curvature side for the pear shaped particles). The carbondiffuses along the concentration gradient and precipitates on the opposite half, around and below thebisecting diameter. However, it does not precipitate from the apex of the hemisphere, which accountsfor the hollow core that is characteristic of these filaments. For supported metals, filaments can formeither by ‘extrusion (also known as base growth)’ in which the nanotube grows upwards from themetal particles that remain attached to the substrate, or the particles detach and move at the head of thegrowing nanotube, labelled ‘tip-growth’. Depending on the size of the catalyst particles, SWNT orMWNT are grown. In arc discharge, if no catalyst is present in the graphite, MWNT will be grown onthe C_2-particles that are formed in the plasma.

Arc discharge.The carbon arc discharge method, initially used for producing C_{60} fullerenes, is the most common andperhaps easiest way to produce carbon nanotubes as it is rather simple to undertake. However, it is atechnique that produces a mixture of components and requires separating nanotubes from the soot andthe catalytic metals present in the crude product.

This method creates nanotubes through arc-vaporisation of two carbon rods placed end to end,separated by approximately 1mm, in an enclosure that is usually filled with inert gas (helium, argon) atlow pressure (between 50 and 700 mbar). Recent investigations have shown that it is also possible tocreate nanotubes with the arc method in liquid nitrogen[31]. A direct current of 50 to 100 A driven byapproximately 20 V creates a high temperature discharge between the two electrodes. The dischargevaporises one of the carbon rods and forms a small rod shaped deposit on the other rod. Producingnanotubes in high yield depends on the uniformity of the plasma arc and the temperature of the depositform on the carbon electrode[32].

Insight in the growth mechanism is increasing and measurements have shown that different diameterdistributions have been found depending on the mixture of helium and argon. These mixtures havedifferent diffusions coefficients and thermal conductivities. These properties affect the speed withwhich the carbon and catalyst molecules diffuse and cool, affecting nanotube diameter in the arcprocess. This implies that single-layer tubules nucleate and grow on metal particles in different sizesdepending on the quenching rate in the plasma and it suggests that temperature and carbon and metalcatalyst densities affect the diameter distribution of nanotubes[32].

Depending on the exact technique, it is possible to selectively grow SWNTs or MWNTs, which isshown in Fig. 25.10. Two distinct methods of synthesis can be performed with the arc dischargeapparatus.

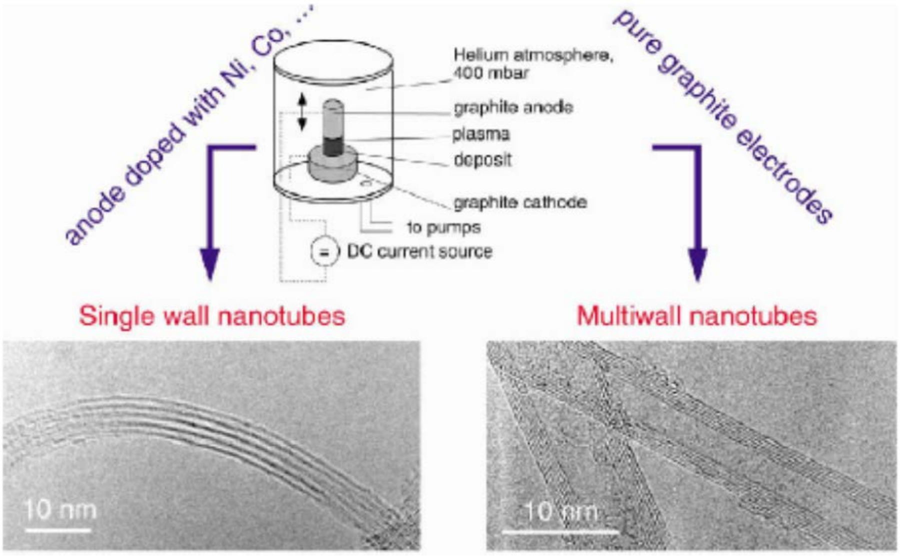


Fig. 25.10.Experimental set-up of an arc discharge apparatus.

*Synthesis of SWNT.*If SWNTs are preferable, the anode has to be doped with metal catalyst, such as Fe, Co, Ni, Y or Mo.A lot of elements and mixtures of elements have been tested by various authors[33] and it is noted thatthe results vary a lot, even though they use the same elements. This is not surprising as experimentalconditions differ. The quantity and quality of the nanotubes obtained depend on various parameterssuch as the metal concentration, inert gas pressure, kind of gas, the current and system geometry.Usually the diameter is in the range of 1.2 to 1.4 nm. A couple of ways to improve the process of arcdischarge are stated below.

The most common problems with SWNT synthesis are that the product contains a lot of metal catalyst,SWNTs have defects and purification is hard to perform. On the other hand, an advantage is that thediameter can slightly be controlled by changing thermal transfer and diffusion, and hencecondensation of atomic carbon and metals between the plasma and the vicinity of the cathode cancontrol nanotube diameter in the arc process. This was shown in an experiment in which differentmixtures of inert gases were used[34].It appeared that argon, with a lower thermal conductivity and diffusion coefficient, gave SWNTs witha smaller diameter of approximately 1.2 nm. A linear fit of the average nanotube diameter showed a0.2 nm diameter decrease per 10 % increase in argon helium ratio, when nickel/yttrium was used(C/Ni/Y was 94.8:4.2:1) as catalyst.

A second way of control is plasma control by changing the anode to cathode distance (ACD). TheACD is adjusted in order to obtain strong visible vortices around the cathode. This enhances anodevaporisation, which improves nanotubes formation. Combined with controlling the argon-heliummixture, one can simultaneously control the macroscopic and microscopic parameters of the nanotubesformed.With a nickel and yttrium catalyst (C/Ni/Y is 94.8:4.2:1) the optimum nanotube yield was found at apressure of 660 mbar for pure helium and 100 mbar for pure argon. The nanotube diameter rangesfrom 1.27 to 1.37 nanometre.

Knowing that chemical vapour deposition (CVD) could give SWNTs with a diameter of 0.6–1.2 nm,researchers tried the same catalyst as used in CVD on arc discharge. Not all of the catalysts usedappeared to result in nanotubes for both methods. But there seemed to be a correlation of diameter ofSWNTs synthesised by CVD and arc dischargeAs a result, the diameter can be controllably lowered to a range of 0.6-1.2 nm with arc-discharge.Using a mixture of Co and Mo in high concentrations as catalyst resulted in this result. Thesediameters are considerably smaller than 1.2-1.4 nm[33], which is the usual size gained from arcdischarge.

There is also progress in developing methods to improve the oxidation resistance of the SWNTs,which is a consequence of the defects present in nanotubes. A strong oxidation resistance is needed forthe nanotubes if they have to be used for applications such as field emission displays. Recent researchhas indicated that a modified arc-discharge method using a bowl-like cathode (see Fig. 25.11),decreases the defects and gives cleaner nanotubes, and thus improves the oxidation resistance[35].The Raman spectrum of the newly synthesised nanotubes shows that the nanotubes formed are cleanerand less defective compared with those synthesised by conventional methods. The anode rodcontained Ni and Y catalyst (C /Ni/Y is 94.8:4.2:1). No information is given about the diameter size.

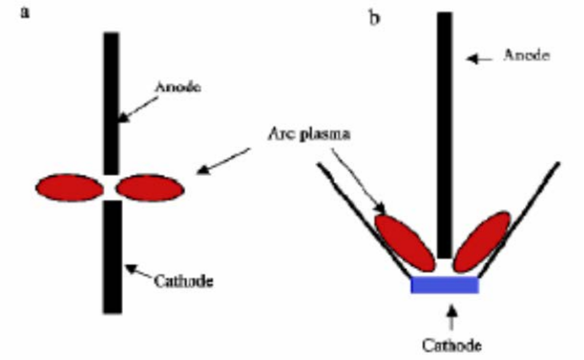


Fig. 25.11. Schematic drawings of the electrode set-ups for (a) the conventional and (b) the new arc dischargeelectrodes.

The researchers discovered that it was possible to form MWNTs in open air [36].A welding arc torch was operated in open air and the process was shielded with an argon gas flow.The anode and cathode were made of graphite containing Ni and Y (Ni/Y is 4.2:1 at. %).This method was modified for preparing SWNTs[37]. A plate target made of graphite containing metalcatalyst Ni and

Y (Ni/Y is 3.6:0.8 at. per cent), was fixed at the sidewall of a water–cooled, steel basedelectrode. The torch arc aimed at the edge of the target and the soot was deposited on the substratebehind the target (see Fig. 25.12). The arc was operated at a direct current of 100 A and shieldingargon gas flowed through the torch, enhancing the arc jet formation beyond the target.In the soot, carbon nanohorns (CNHs) and bundles of SWNT with an average diameter of 1.32 nmwere found. However, the yield was much lower than for the conventional low-pressure arc dischargemethod. There are two reasons for this fact. At first, because of the open air, the lighter soot willescape into the atmosphere. Secondly, the carbon vapour might be oxidised and emitted as carbondioxide gas. In order to improve the yield in this method, contrivances for collecting soot anddevelopment of an appropriate target are required.This method promises to be convenient and inexpensive once the conditions for higher yield areoptimised. With a Ni/Y catalyst (Ni/Y is 3.6:0.8), SWNT bundles and CNHs are formed. In this casethe SWNTs have a diameter of approximately 1.32 nm [37].

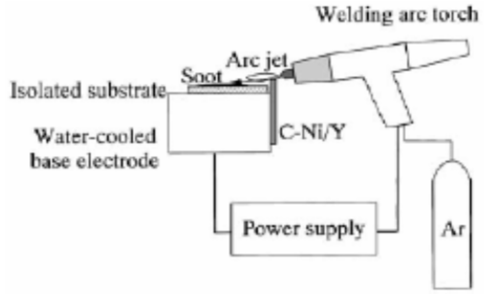


Fig. 25.12.Experimental set-up of the torch arc method in open air.

*Synthesis of MWNT.*If both electrodes are graphite, the main product will be MWNTs. But next to MWNTs a lot of sideproducts are formed such as fullerenes, amorphous carbon, and some graphite sheets. Purifying theMWNTs, means loss of structure and disorders the walls. However scientist are developing ways togain pure MWNTs in a large-scale process without purification.Typical sizes for MWNTs are an inner diameter of 1-3 nm and an outer diameter of approximately 10nm. Because no catalyst is involved in this process, there is no need for a heavy acidic purificationstep. This means, the MWNT, can be synthesised with a low amount of defects.

A first, possibly economical route to highly crystalline MWNTs is the arc-discharge method in liquidnitrogen[31], with this route mass production is also possible(Fig. 25.13). For this option low pressures andexpensive inert gasses are not needed.The content of the MWNTs can be as high as 70 % of the reaction product. Analysis with Augerspectroscopyrevealed that no nitrogen was incorporated in the MWNTs. There is a strong possibilitythat SWNTs can be produced with the same apparatus under different conditions.

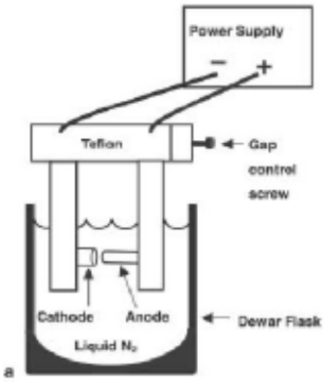


Fig. 25.13.Schematic drawings of the arc discharge apparatus in liquid nitrogen.

Synthesis of MWNTs in a magnetic field[38] gives defect-free and high purity MWNTs that can beapplied as nanosized electric wires for device fabrication. In this case, the arc discharge synthesis wascontrolled by a magnetic field around the arc plasma.Extremely pure graphite rods (purity > 99.999 %) were used as electrodes. Highly pure MWNTs(purity > 95 %) were obtained without further purification, which disorders walls of MWNTs.

A second possibly economical route to mass production of MWNTs is synthesis by plasma rotatingarc discharge technique(Fig. 25.14)[39]. The centrifugal force caused by the rotation generates turbulence andaccelerates the carbon vapour perpendicular to the anode. In addition, the rotation distributes the microdischarges uniformly and generates stable plasma. Consequently, it increases the plasma volume andraises the plasma temperature.At a rotation speed of 5000 rpm a yield of 60 % was found at a formation temperature of 1025 °Cwithout the use of a catalyst. The yield increases up to 90% after purification if the rotation speed isincreased and the temperature is enlarged to 1150 °C. The diameter size was not mentioned in thispublication.

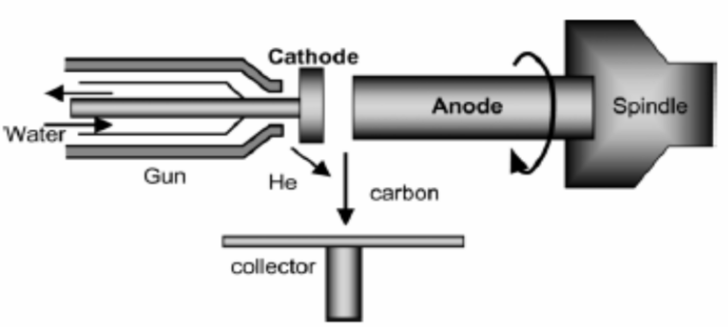


Fig. 25.14.Schematic diagram of plasma rotating electrode system.

Laser ablation.In 1995, Smalley's group[40] at Rice University reported the synthesis of carbon nanotubes by laservaporisation. The laser vaporisation apparatus used by Smalley's group is shown in Fig. 25.15. Apulsed[41, 42], or continuous[43, 44] laser is used to vaporise a graphite target in an oven at 1200 °C. Themain difference between continuous and pulsed laser, is that the pulsed laser demands a much higherlight intensity (100 kW/cm^2 compared with 12 kW/cm^2). The oven is filled with helium or argon gasin order to keep the pressure at 500 Torr. A very hot vapour plume forms, then expands and coolsrapidly. As the vaporised species cool, small carbon molecules and atoms quickly condense to formlarger clusters, possibly including fullerenes. The catalysts also begin to condense, but more slowly atfirst, and attach to carbon clusters and prevent their closing into cage structures[45]. Catalysts may evenopen cage structures when they attach to them. From these initial clusters, tubular molecules grow intosingle-wall carbon nanotubes until the catalyst particles become too large, or until conditions havecooled sufficiently that carbon no longer can diffuse through or over the surface of the catalystparticles. It is also possible that the particles become that much coated with a carbon layer that theycannot absorb more and the nanotube stops growing. The SWNTs formed in this case are bundledtogether by van der Waals forces[45].

There are some striking, but not exact similarities, in the comparison of the spectral emission ofexcited species in laser ablation of a composite graphite target with that of laser-irradiated C_{60} vapour.This suggests that fullerenes are also produced by laser ablation of catalyst-filled graphite, as is thecase when no catalysts are included in the target. However, subsequent laser pulses excite fullerenes toemit C_2 that adsorbs on catalyst particles and feeds SWNT growth. However, there is insufficientevidence to conclude this with certainty.

Laser ablation is almost similar to arc discharge, since the optimum background gas and catalyst mixis the same as in the arc discharge process. This might be due to very similar reaction conditionsneeded, and the reactions probably occur with the same mechanism.

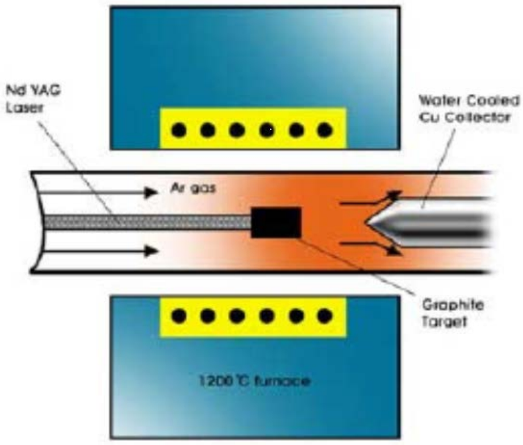


Fig. 25.15.Schematic drawings of a laser ablation apparatus.

*SWNT versus MWNT.*The condensates obtained by laser ablation are contaminated with carbon nanotubes and carbonnanoparticles. In the case of pure graphite electrodes, MWNTs would be synthesised, but uniformSWNTs could be synthesised if a mixture of graphite with Co, Ni, Fe or Y was used instead of puregraphite. SWNTs synthesised this way exist as 'ropes' (Fig. 25.16)[43, 45]. Laser vaporisation resultsin a higher yield for SWNT synthesis and the nanotubes have better properties and a narrower sizedistribution than SWNTs produced by arc-discharge.Nanotubes produced by laser ablation are purer (up to about 90 % purity) than those produced in thearc discharge process. The Ni/Y mixture catalyst (Ni/Y is 4.2/1) gave the best yield.The size of the SWNTs ranges from 1-2 nm, for example the Ni/Co catalyst with a pulsed laser at1470 °C gives SWNTs with a diameter of 1.3-1.4 nm. In case of a continuous laser at 1200 °C andNi/Y catalyst (Ni/Y is 2:0.5 at. %), SWNTs with an average diameter of 1.4 nm were formed with 20-30 % yield, see Fig. 25.16[43].

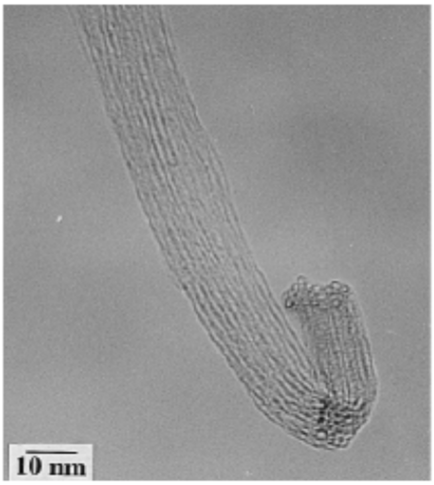


Fig. 25.16. TEM images of a bundle of SWNTs catalysed by Ni/Y (2:0.5 at. %) mixture, produced with a continuous laser.

Chemical vapour deposition.Chemical vapour deposition (CVD) synthesis is achieved by putting a carbon source in the gas phaseand using an energy source, such as a plasma or a resistively heated coil, to transfer energy to agaseous carbon molecule. Commonly used gaseous carbon sources include methane, carbon monoxideand acetylene. The energy source is used to “crack” the molecule into reactive atomic carbon. Then,the carbon diffuses towards the substrate, which is heated and coated with a catalyst (usually a firstrow transition metal such as Ni, Fe or Co) where it will bind. Carbon nanotubes will be formed if theproper parameters are maintained. Excellent alignment[46], as well as positional control on nanometer scale[47], can be achieved by using CVD. Control over the diameter, as well as the growth rate of thenanotubes can also be maintained. The appropriate metal catalyst can preferentially grow single ratherthan multi-walled nanotubes[30].

CVD carbon nanotube synthesis is essentially a two-step process consisting of a catalyst preparationstep followed by the actual synthesis of the nanotube. The catalyst is generally prepared by sputteringa transition metal onto a substrate and then using either chemical etching or thermal annealing toinduce catalyst particle nucleation. Thermal annealing results in cluster formation on the substrate,from which the nanotubes will grow. Ammonia may be used as the etchant[42 - 45]. The temperaturesfor the synthesis of nanotubes by CVD are generally within the 650–900 oC range[46 - 49]. Typicalyields for CVD are approximately 30%.

These are the basic principles of the CVD process. In the last decennia, different techniques for thecarbon nanotubes synthesis with CVD have been developed, such as plasma enhanced CVD, thermalchemical CVD, alcohol catalytic CVD, vapour phase growth, aero gel-supported CVD and laserassistedCVD.

*Plasma enhanced chemical vapour deposition.*The plasma enhanced CVD method generates a glow discharge in a chamber or a reaction furnace by ahigh frequency voltage applied to both electrodes. Figure25.17shows a schematic diagram of a typicalplasma CVD apparatus with a parallel plate electrode structure.

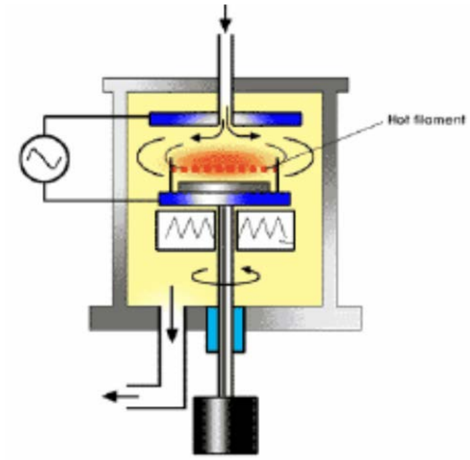


Fig. 25.17.Schematic diagram of plasma CVD apparatus.

A substrate is placed on the grounded electrode. In order to form a uniform film, the reaction gas issupplied from the opposite plate. Catalytic metal, such as Fe, Ni and Co are used on for example a Si,SiO$_2$, or glass substrate using thermal CVD or sputtering. After nanoscopic fine metal particles areformed, carbon nanotubes will be grown on the metal particles on the substrate by glow dischargegenerated from high frequency power. A carbon containing reaction gas, such as C$_2$H$_2$, CH$_4$, C$_2$H$_4$,C$_2$H$_6$, CO is supplied to the chamber during the discharge[50].

The catalyst has a strong effect on the nanotube diameter, growth rate, wall thickness, morphology andmicrostructure. Ni seems to be the most suitable pure-metal catalyst for the growth of aligned multiwalledcarbon nanotubes (MWNTs)[51]. The diameter of the MWNTs is approximately 15 nm. Thehighest yield of carbon nanotubes achieved was about 50% and was obtained at relatively lowtemperatures (below 330°C)[50].

*Thermal chemical vapour deposition.*In this method Fe, Ni, Co or an alloy of the three catalytic metals is initially deposited on a substrate.After the substrate is etched in a diluted HF solution with distilled water, the specimen is placed in aquartz boat. The boat is positioned in a CVD reaction furnace, and nanometre-sized catalytic metalparticles are formed after an additional etching of the catalytic metal film using NH$_3$ gas at atemperature of 750 to 1050°C. As carbon nanotubes are grown on these fine catalytic metal particlesin CVD synthesis, forming these fine catalytic metal particles is the most important process. Figure 25.18 shows a schematic diagram of thermal CVD apparatus in the synthesis of carbon nanotubes.

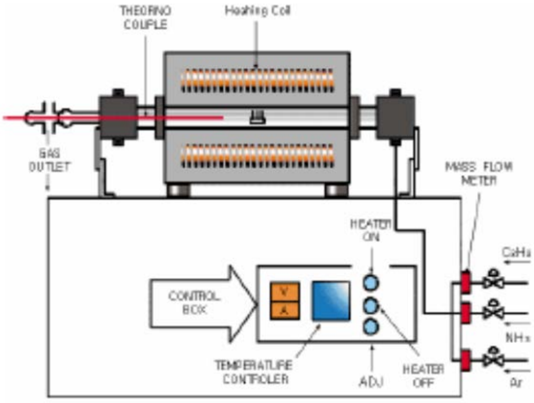


Fig. 25.18.Schematic diagram of thermal CVD apparatus.

When growing carbon nanotubes on a Fe catalytic film by thermal CVD, the diameter range of thecarbon nanotubes depends on the thickness of the catalytic film. By using a thickness of 13 nm, thediameter distribution lies between 30 and 40 nm. When a thickness of 27 nm is used, the diameterrange is between 100 and 200 nm. The carbon nanotubes formed are multiwalled[52].

Flame synthesis.This method is based on the synthesis of SWNTs in a controlled flame environment, that produces thetemperature, forms the carbon atoms from the inexpensive hydrocarbon fuels and forms small aerosolmetal catalyst islands[53 - 55]. SWNTs are grown on these metal islands in the same manner as in laserablation and arc discharge.

These metal catalyst islands can be made in three ways. The metal catalyst (cobalt) can either becoated on a mesh[53], on which metal islands resembling droplets were formed by physical vapourdeposition. These small islands become aerosol after exposure to a flame. The second way[55], is tocreate aerosol small metal particles by burning a filter paper that is rinsed with a metal-ion (e.g. ironnitrate) solution. The third way is the thermal evaporating technique in which metal powder (e.g. Feor Ni) is inserted in a trough and heated[54].

In a controlled way a fuel gas is partially burned to gain the right temperature of ~800 °C and thecarbon atoms for SWNT production. On the small metal particles the SWNTs are than formed(Fig. 25.19). Asoptimisation parameters the fuel gas composition, catalyst, catalyst carrier surface and temperature canbe controlled[53].

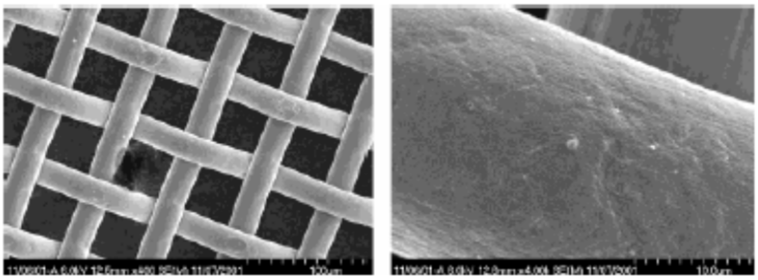


Fig. 25.19.Meshes on which the metal catalyst is coated, used in flame synthesis.

In Table 25.1, a short summary of the three most common CNT obtaining techniques used nowadays is given.

Table 25.1.A summary of the major CNT production methods and their efficiency.

Method	**Arc discharge method**	**Chemical vapour deposition**	**Laser ablation (vaporization)**
Who	Ebbesen and Ajayan, NEC, Japan 1992 15	Endo, Shinshu University, Nagano, Japan 53	Smalley, Rice, 199514

How	Connect two graphite rods to a power supply, place them a few millimetres apart, and throw the switch. At 100 amps, carbon vaporises and forms a hot plasma.	Place substrate in oven, heat to 600 °C, and slowly add a carbon-bearing gas such as methane. As gas decomposes it frees up carbon atoms, which recombine in the form of NTs	Blast graphite with intense laser pulses; use the laser pulses rather than electricity to generate carbon gas from which the NTs form; try various conditions until hit on one that produces prodigious amounts of SWNTs
Typical yield	30 to 90%	20 to 100 %	Up to 70%
SWNT	Short tubes with diameters of 0.6 - 1.4 nm	Long tubes with diameters ranging from 0.6-4 nm	Long bundles of tubes (5-20 microns), with individual diameter from 1-2 nm.
MWNT	Short tubes with inner diameter of 1-3 nm and outer diameter of approximately 10 nm	Long tubes with diameter ranging from 10-240 nm	Not very much interest in this technique, as it is too expensive, but MWNT synthesis is possible.
Pro	Can easily produce SWNT, MWNTs. SWNTs have few structural defects; MWNTs without catalyst, not too expensive, open air synthesis possible	Easiest to scale up to industrial production; long length, simple process, SWNT diameter controllable, quite pure	Primarily SWNTs, with good diameter control and few defects. The reaction product is quite pure
Con	Tubes tend to be short with random sizes and directions; often needs a lot of purification	NTs are usually MWNTs and often riddled with defects	Costly technique, because it requires expensive lasers and high power requirement, but is improving

25.2.3. Electrical Conduction in Nanotubes

Depending on their chiral vector, carbon nanotubes with a small diameter are either semi-conducting or metallic. The differences in conducting properties are caused by the molecular structure that results in a different band structure and thus a different band gap. The differences in conductivity can easily be derived from the graphene sheet properties [56]. It was shown that a (n,m) nanotube is metallic as accounts that: $n=m$ or $(n-m) = 3i$, where i is an integer and n and m are defining the nanotube. The resistance to conduction is determined by quantum mechanical aspects and was proved to be independent of the nanotube length[57, 58].

In order to understand the electrical properties of ananotube, we first have to look at a single graphene sheet.Graphene is on the borderline of being metallic: the bandstructure calculation results in a ‘‘zero-gap semiconductor’’behavior[59].Fig. 25.20 shows an illustration of theelectron energy as a function of the two-dimensional wavevector. The Brillouin zone of graphene is a hexagon, andthe allowed electronic states (wave vectors) correspond tothe points inside the hexagon; the electronic bands reachthe Fermi level at the six corners of the hexagon.Therefore the Fermi ‘‘surface’’ consists of six points, asindicated in the two panels in Fig. 25.21. (In graphite theinfluence of the neighboring layers causes the bands tocross the Fermi level, and there will be finite Fermisurface, yielding a semimetallic behavior.)

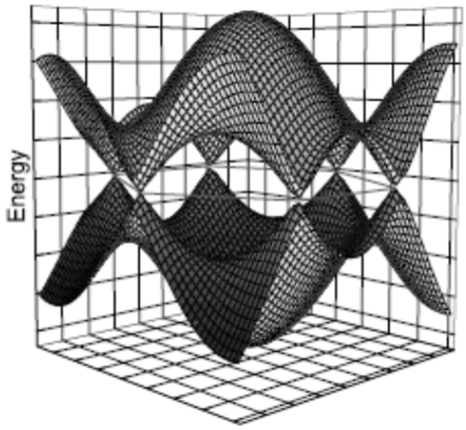


Fig. 25.20.Schematic representation of the energy bands ingraphene. The possible electron states are constrained in ahexagon in the base plane. The lower band is full, the upper bandis empty. The energy of the electrons is zero (equal to the Fermienergy) at the six corners of the hexagon.

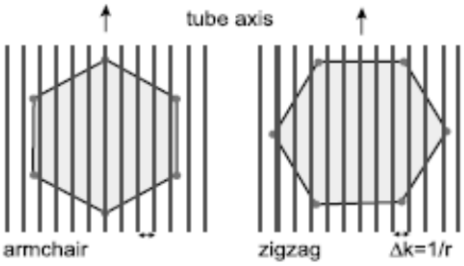


Fig. 25.21. The six corners of the hexagon correspond to electronicstates available for conduction. The lines represent the allowedelectronic states in a zigzag and an armchair nanotube.

By bending the graphene sheet into a nanotube themost drastic change will be the constraint introduced dueto the periodic boundary condition along the circumferenceof the cylinder. Accordingly, in the directionperpendicular to the tube axis the allowed wave vectorswill be constrained to lines. The larger the diameter of thenanotube the more dense are the lines. Naturally, for avery large diameter nanotube the allowed wave vectorsare just about the same as for the graphite, and theproperties are also similar. However, for smaller diametersthe constraint is really important. As shown in the twopanels of Fig. 25.21, the relative orientation of the linescorresponding to the allowed states is sensitive to theactual structure of the nanotube[18, 60]. For armchair nanotubesone of the lines always intersects the corner of thehexagon, therefore these nanotubes are metallic; for thezigzag structure it is possible that the lines miss thecorners, and in those cases the nanotube is semiconducting.By carrying out this analysis in more detail, we obtainthat a nanotube is metallic if the integers in the chiralvector describing the structure $\mathbf{C_h}=(n_1,n_2)$ satisfy$n_1-n_2=3j$, where j is an arbitrary integer. More detailedcalculations, taking into account the curvature of thegraphene sheet, lead to the conclusion that for $n_1=n_2$(armchair) the nanotubes are metallic, for $n_1-n_2=3j$ thereis a tiny gap at the Fermi surface, and for all other casesthe gap is large.

Much progress has been made in the experimental testof these predictions[61]. One prefers to work with aSWNT, because for a MWNT it is unlikely that all of thecomponent nanotubes are of the same type (metal orsemiconductor). The first conductance measurementswere done on a nanotube rope, a collection of similarSWNTs. Several single SWNTs of various chirality havebeen tested later. The measurements largely support theconclusions drawn above. For the complete interpretationof the experimental results one needs to step beyond thesimple theory of bulk metals: the electrical leads to thenanotube must be treated as tunneling contacts. Theresonance between higher quantum levels and theCoulomb repulsion of electrons on the nanotube playimportant roles[62, 63].

Electronic property measurements of single SWNTs.Studying the electronic properties of SWNTs, scientists have been able to calculate models for energybands and Density of States (DOS) in single walled carbon nanotubes. If these models have to beexperimentally confirmed, a way has to be found in which atomic structures and electron densities canbe visualised. Scanning Tunnelling Microscopy (STM) can do this job as it can display atomicstructures as well as measure the DOS.

As is shown in Fig. 25.22and Fig. 25.23A, chirality can clearly be determined from STMmeasurements. Combining this fact with the ability to measure electronic properties allows studyingthe effect of nanotube chirality on electronic properties. The spectroscopic image (Fig. 25.23B)shows the density of states as measured by a STM. As these two images show, this technique can alsobe used to determine properties of intra-molecular SWNT junctions. In addition, it can also be used formeasuring influence of symmetry, defects, and doping, electronic contacts and so on.

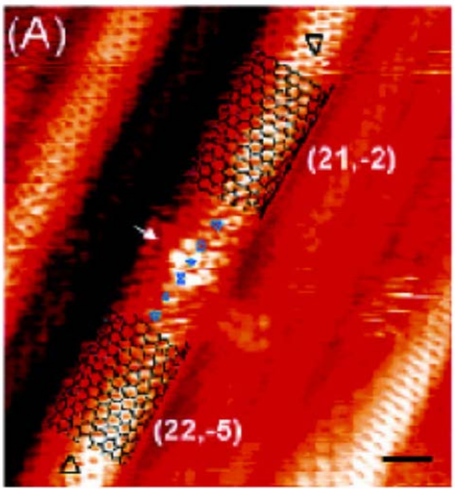


Fig. 25.22.STM image of a SWNT. The chirality can be determined easily.

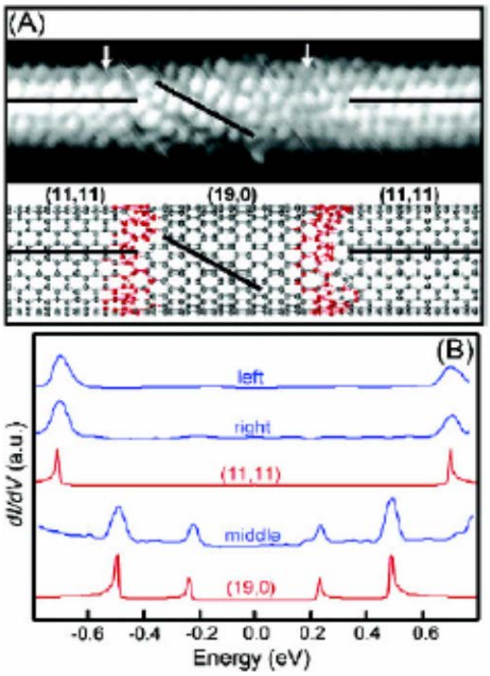


Fig. 25.23. A) STM atomic resolved image of Intra Molecular Junction of SWNT. B) STM image measurement representing the Density of States corresponding to lines in A[64].

25.2.4. CNTs in Medicine and Bioligy

Carbon nanomaterialsare being used to develop the next generation of biomaterialsfor applications in therapeutics and regenerative medicine.Carbon nanomaterials, mainly in the form of nanotubesand graphene, have become the focus of intensive researchbecause of their unique physical and chemical propertiessuch as their hollow structure, their high surface area-tovolumeratio, electrical conductance, thermal conductivity,mechanical stiffness, and the possibilities of functionalizingthem to change their intrinsic properties. Functionalizationcan increase their solubility and biocompatibility underphysiological conditions. The nanomaterials can be furtherconjugated with specific biomolecules such as polymers,peptides, proteins, nucleic acids, and other therapeuticagents, which can target specific types of cells, tissue, andorgans[65].

Carbon nanomaterials demonstrate several significantfeatures that have promise for use in nervous systemrepair. Carbon nanomaterials have the type of nanosurfacefeatures that have been demonstrated to encourage nervoustissue regeneration, including the physical shape (a lineargeometry), the nanoscale surface topology, and the highaspect ratio of nanomolecules or larger structures madefrom carbon nanotubes, like carbon nanotube thread. Theseclosely resemble the microenvironment that nerve fibersmigrate along during embryonic development and regeneration.Carbon nanomaterials also offer high mechanicalstrength to support process outgrowth and flexibility toavoid further damage of soft surrounding tissues duringmovement. The applications of carbon nanomaterials in neurobiology are also very important, in particularly, theiruse in scaffolds to repair damaged nervous tissues, andtheir use as biocompatible electrodes for recording from orstimulating nervous tissues.

Carbon nanotube research has reached a maturity stage thatoffers an adequate understanding of the structural and physicochemicalproperties and allows the exploration of their usein various applications, including the biological application.Biotechnological applications of CNT are being anticipated ina variety of fields ranging from microfluidics to bioinformatics.Ever since methodologies to improve the aqueous solubility ofCNT (also in biological fluids) were developed, transformationof CNT as viable components of therapeutic agents appropriatefor clinical use has been feasible. Table 25.2describes the differentcategories of biology-related applications for which CNThave already been explored[66].

Table 25.2.Carbon nanotubes in biomedical applications[66].

	CNT	Functionalization	In vitro model	In vivo model	Stude aim	Ref.
Biomedical sensors	SWNT	Unmodified	n/a	n/a	Role as catalyst in biochemical reactions (PCR)	[67]
	SWNT MWNT	Oxidized and then coupled covalently with DNA	n/a	n/a	CNT for amplifeing enzyme-based bioaffinity electrical sensing of proteins and DNA	[68]
	MWNT	GOD enzyme immobilized on the electrode of gold/multi-walled carbon nanotubes	n/a	n/a	Glucose biosensors	[69]
	SWNT	Oxidized (COOH)	n/a	n/a	Bioelectrodes	[70]
	SWNT	Ferrocenoyl-functionalized	n/a	n/a	Glucose biosensors and phosphate anions	[71] [72]
	MWNT	Immobilized on the surface of a glassy carbonelectrode by mixing with horse-radish peroxidazse (HRP)	n/a	n/a	On-line glucose and lactate biosensors	[73]
	MWNT	Oxidized and covalently bounded to DNA	n/a	n/a	Miniature DNA biosensors	[74]
	MWNT	Carbon nanotube-containing composite with	n/a	n/a	Development of stable	[75]

		encapsulated enzyme			biosensors	
	SWNT	Polyethylene oxide-funtionalized CNT	n/a	n/a	Highly specific electronic biomolecule detectors	[76]
Tissue engineering	MWNT	Oxidized (COOH) and chemically bounded to PABS and ethylene-ediamine	Hippocampal neurons	n/a	Neural prosthesis	[77]
	MWNT	Unmodified or coated with the bioactive molecule 4-hydroxynon-enal	Embryonic rat-brain neirons	n/a	Substrate for neuronal growth	[78]
	MWNT	Oxidized (COOH)	L292	n/a	3D network substrate for cellular growth	[79]
	Carbon nanofibers/ nanotubes	CNT reinforced polycarbonate urethane composites	Rat astrocyte cells, human osteoblast, PC-12 and 3T3	n/a	Neural or orthopedic prosthetic devices	[80]
	MWNT	Nanocomposite consisting of blends of polylactic acid and CNT	Osteoblasts	n/a	Substrate to expose cells to electrical stimulation	[81]
Ion channel blockers	SWNT	Oxidized (COOH)	CHO	n/a	Block K’ channels	[82]
Translocation into cells	SWNT	FITC or peptide from G protein covalently bounded to CNT	3T6, 3T3	n/a	CNT as delivery systems and their cellular uptake	[83]
	SWNT	FITC or biotin-fluorescent streptavidin covalently linked to CNT	HL60	n/a	Cellular internalization of CNT conjugates	[84]
Cytotoxicity	SWNT	Unrefined	HaCaT	n/a	Risk evaluation of the dermal exposure to SWNT	[85]
	SWNT	Unrefined	n/a	n/a	Risk evaluation of the airborne and dermal exposure to SWNT	[86]
	SWNT	FITC or peptide from G protein covalently bounded to CNT	3T6, 3T3	n/a	Cellular viability	[83]
	SWNT	FITC and biotin-fluorescent streptavidin covalently linked to CNT	HL60	n/a	Cellular viability	[84]
	SWNT	Unrefined	n/a	Crl:CD(SD)IGS BR mice	Evaluation of the acute lung toxicity of intratracheally instilled SWCN	[87]
	SWNT	Unrefined	n/a	B6C3F mice	Evaluation of the pulmonary toxicity after intratracheally instillation of SWCN	[88]
	SWNT	NH_3^+	HeLa	n/a	Cellular viability	[89]

Antigen delivery	SWNT	FMDV peptide covalently linked to CNT	n/a	n/a	Study of the antigenicity of the FMDV peptide-CNT conjugate	[90]
	SWNT	Mono and bis functionalization with FMDV peptide	n/a	BALB/c mice	Study of the immune response of CNT conjugates	[91]
Gene delivery	SWNT MWNT	NH_3^+covalently linked and plasmid DNA by electrostatic interactions	HeLa CHO	n/a	CNT as delivery system, cellular distribution of CNT conjugates and gene expression	[89]

Biosensors.The detection of molecules (DNA, proteins, enzymes, infectiousagents) present in diminutive concentrations in biologicalenvironments is difficult if not impossible. Ultra-sensitive assaysare required to achieve very low detection limits. Carbonnanotubes have been extensively explored in this area due totheir physical and electrochemical properties. Most bioassaysdepend on hybridization or antigen-antibody interactions [69].Also, CNT are used to immobilize DNA, proteins, enzymes oranions on their surface, in this way acting as bioelectrodes [68-76] responsible for transduction and amplification of events oras catalysts in biochemical reactions [67]. Nanoscale systemsbased on carbon nanotubes have already shown great improvementsin bioelectronic devices, protein analysis and medicaldiagnostics [67-76].

Tissue engineering.Neural and bone implants commonly fail due to glial scartissueformation (central nervous system) or fibrous encapsulation(bone) at the tissue-implant interface. This leads toshort-term functionality of the implant and can result in poorclinical efficacy [78]. The rough surface of carbon nanotubesat the nanometer scale has been shown to permit sufficientnatural bonding between the implant and the surrounding tissuewithout wound-healing [78]. Moreover, carbon nanotubematrices have offered a range of possibilities for the design ofnew composites for neural and orthopedic implants combiningnumerous features. Besides the electric properties of CNT thatallow electrical stimulation of cellular growth [77, 81] andtheir strong mechanical properties needed to ensure implantrobustness [80], CNT are biocompatible and non biodegradablematerial, permitting their use as long-term implants [77]. Veryrecent studies *in vitro* [77-81] have explored the possibilitiesfor the use of CNT in tissue regeneration after damage of thespinal cord, brain or bone tissues.

Ion channel blockers.Ion channels play an essential role in cellular communication(cell-cell and cell-environment). Disturbances in theionic balance generate information for cells, and the cellularresponse will interfere with the body homeostasis. Controllingthe functionality of ion channels can be very useful in severalsituations when for example an ionic imbalance results in adisease. Several toxins and therapeutic drugs have ion channelsas their targets. One study [82] has shown the susceptibility ofK- channels responding in a dose-dependent manner to CNTof certain diameters. Further exploration in this area can leadto CNT per se as therapeutic agents.

Cytotoxicity studies.The increasing interest in CNT for electronics to biomedicalapplications, inevitably will lead to the mass production ofthese materials in all their conformation: single or multi-walled,functionalized or not. A few studies have lately appeared onthe toxicity of carbon nanotubes. Concerning handling ofunrefined CNT material by CNT production workers, thesestudies have focused on dermal [85] and pulmonary [87-88]exposure. Current studies on the utilization of *f*-CNT as carriersfor drugs, DNA and proteins and their internalization into cells,have produced new information related with the cytotoxicityof carbon nanotubes [83, 84, 89] on the cellular level.

25.3. Nanotubes for Medical Implants

Nanotubes are cylindrical hollow structures, which have diametersin the range of 1-800 nm and very high length todiameter ratios to the extent of 132,000,000:1[92]. Carbonnanotubes (CNTs) (Fig. 25.24) are also can be considered forbiomedical implants because of their mechanical, thermal, orelectrical properties[92]. Recently, specific interests have beengenerated over the fabrication of nanotubes made of metal oxides.The most commonly used nanotube materials are titania(Fig. 25.25), alumina, silicon, boron nitride, manganese dioxide,tungsten disulphide, zinc oxide, molybdenum disulphide andtin sulphide[93].

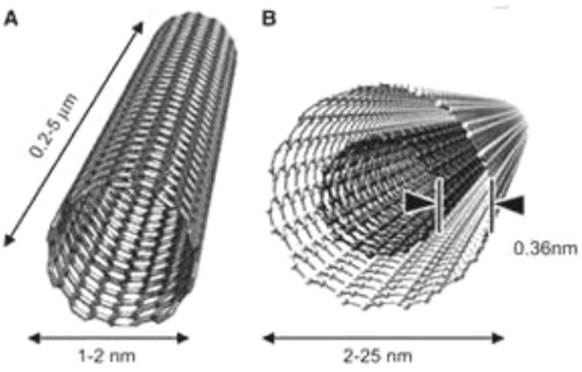
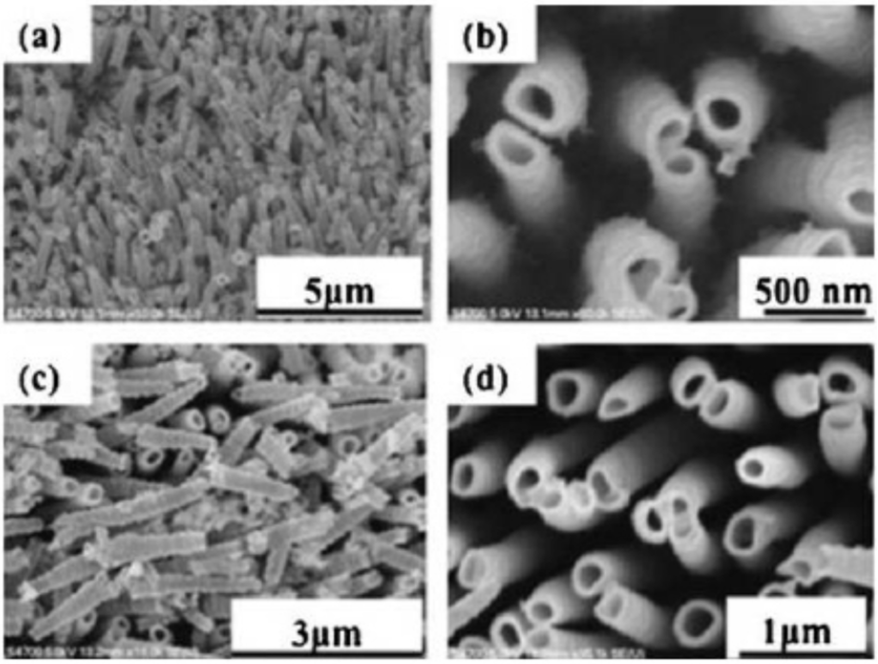


Fig. 25.24.Conceptual diagram of single-walled carbon nanotube (SWCNT)(A) and multiwalled carbon nanotube (MWCNT) (B) delivery systems showingtypical dimensions of length, width, and separation distance between graphenelayers in MWCNTs[94].



The SEM images of (a) as-synthesized TiO2 nanotubes arrays and (b) close up of nanotubes shown in (a)

Fig. 25.25. Titanium dioxide nanotubes prepared by Pt electrode (a)–(f). In general,titania nanotubes have lengths between 7 and 10µm andinner diametersranging from 60 to 130 nm. (a) and (b) are amorphous, (c) and (d) are crystalline,(e) is as-synthesized and (f) is annealedTNT[95].

The reasons for use of nanotubular structures in orthopedicimplants can be classified as follows:

(a) Improved levels of osseointegration: The structure of thebones consist of elements at three different dimensional scales:(a) the macrosized cortical as well as cancellous bones; (b) themicroscale structures including haversian systems, osteons,and lamellae; (c) the nanoscale features like non-collagenousorganic proteins, fibrillar collagen and embedded mineral crystals.A hybrid of micro-scale pits in the implant surface itselfand nano-ordered layers of tubes are used to mimic the cellularenvironment thus favoring rapid bone accrual process (Fig. 25.26).

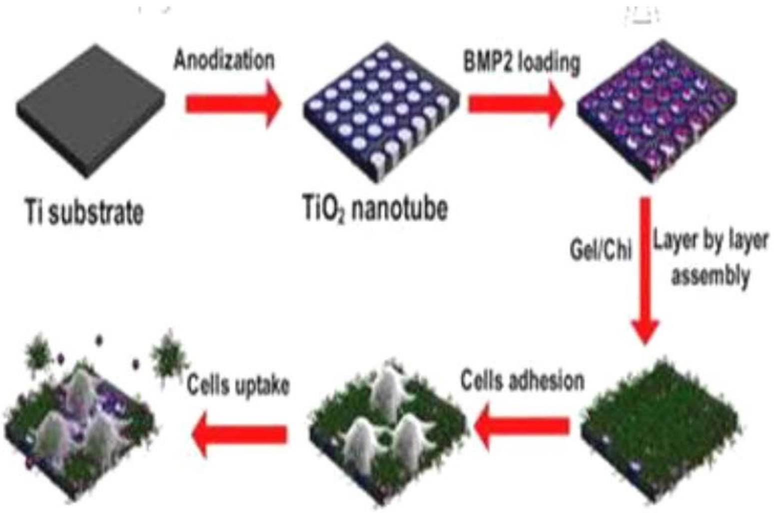


Fig. 25.26.Graphical illustration showing the process of TiO$_2$nanotube layerformation on Ti substrate and its effect on cell uptake and adhesion[96].

(b) Anti-inflammation/anti-microbial functionality: Nanotubesby means of their particular ordered structure have beenfound to possess antimicrobial properties. Moreover, loadingnanotubes with anti-inflammatory drugs or antimicrobialagents can further reduce the risks of post-operative infection.This could be done through anti-inflammatory agentslike dexamethasone, drugs such as Indomethacin, Gentamicinand Vancomycinor antibacterial chemicals like silver nanoparticlesor zinc oxide nanoparticles (Fig. 25.27). Functionalizednanotubes can target specific cells and then release the loadeddrugs in response to triggers like change in pH or temperature.

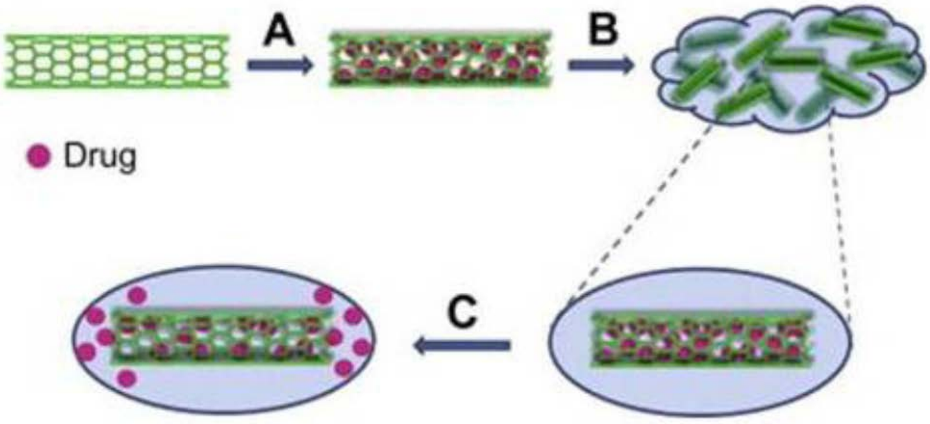


Fig. 25.27. Schematic of the drug loading and release process of CNT nanoreservoirs.A) Drug solution is filled into the interior of acid treated CNTsthrough sonication; B) Pyrrole is added to the suspension containing CNTsand Dex and electropolymerization is carried out; C) Drug is released fromCNT nanoreservoirs to surroundings through diffusion or electric stimulation[97].

(c) Shielding/Scaffolding effect: Use of certain nanotubularstructures (like Carbon Nanotubes/Boron NitrideNanotubes) with high stiffness, shear resistance and compressivestrength as reinforcing agent in conjunction withbone-simulating bio-active material like Hydroxyapatiteor Polypropylene fumarate to create effective prostheticsallowing for rapid osseointegration (Fig. 25.28).

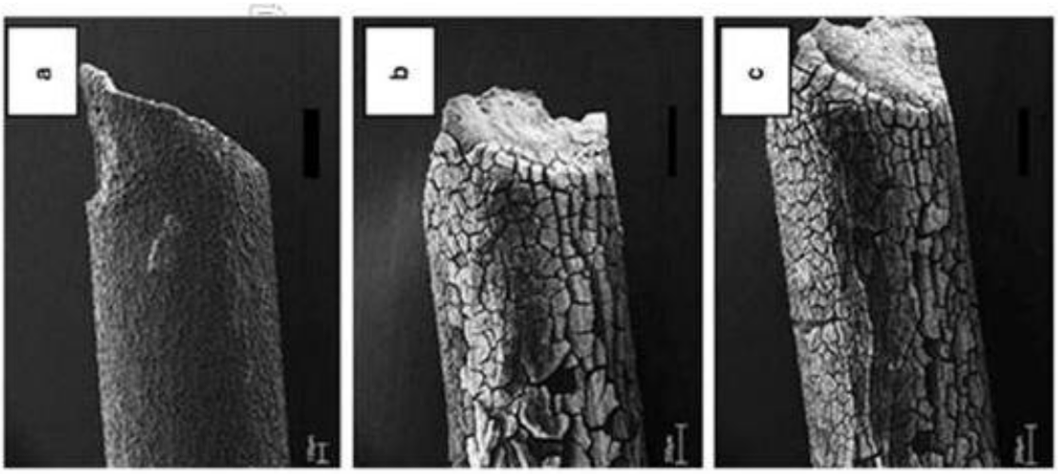


Fig. 25.28. SEM images of HA–CNTs coating on Ti alloy wire with differentdeposition times (a) 1 min, (b) 2 min, and (c) 3 min[98].

25.4. Nanotubes for implant surfaces: The biomimickingeffect for osseointegration

All complex organisms are composed of an architecturalhierarchy ranging from the nano-level through micro-levelto the macroscale. Complex macro scale actions are effected bysynergistic combination of processes at smaller dimensionalscales. It starts with extracellular proteins like collagen whichbuild the tissue structure and have dimensions of about 1.5 nmin diameter and 200 nm in length, a 5 to 10 nm thick encapsulatingcell membrane, proteins for cell adhesion that effect thisthrough binding sites of length 5nm situated at the end of their20nm long body and biochemical compounds like DNA whichthrough changes in conformity can produce structural andfunctional effects[99]. Cellular activities such as cell attachment,locomotion, growth, gene expression and fate of stemcells are mostly controlled through nanosized features. Hence,development of biomaterials incorporating nano-dimensionedfacets has received considerable attention with a view to providefavorable environment for normal cell behavior (Fig. 25.29).

In case of nano-patterned surfaces for implants,the chemical stability and structural properties of the layer isuseful for providing greater reactive surface area and highernumber of sites for preferential adsorption of proteins. Suchbiologically inspired nanotube layers are better suited for attachmentof osteoblast cells due to the presence of nanoscaleholes allowing greater penetration of cell filopodia. The elevatedsurface roughness and increased surface affinity for calciumalso causes preferential deposition of apatitic minerals thusrendering the surfaces the valuable property of bioactivity.

One of the most important parameters governingthe successful proliferation of cells on any surface is the cell-cell/Extracellular matrix (ECM)-cell interaction[100]. Nearly300 different proteins are responsible for forming the core ofthe ECM. The proteins are polymeric units formed by crosslinking of insoluble monomeric singular units which have afiber like appearance. This ECM constitutes the characteristicenvironment, which determines cell behavior and proliferation.Changes in the elasticity of the environment have a determiningeffect on the cell behavior. As the number of sitesof adsorption increases there is a subsequent increase in thearea over which the force is distributed thereby reducing thelocalized stress, thus simulating an increase in elasticity. Apartfrom the interaction between the cell and the ECM cell signalingdetermines the functioning of the tissues. These intercellularinteractions are a consequence of the action of nanoscalemolecules, which also regulate the production, propertiesand composition of the ECM and translocation of biologicalmatter. Disruption in the natural order in the cell and unwarrantedremodeling and production of the ECM can result fromthe disturbance of intracellular interaction mechanisms. ECMcontains hydrocarbon in the form of collagen and inorganiccarbonate apatite which acts as the mineral constituent of thebone. The nanotubes, owing to their particular dimensionalrange, are considered to be analogous to the fibrillar proteinconstituents in the ECM, especially those in the collagen.

During installation of the implant, blood comesin contact with the implant initially. Blood owing to its highconcentration of plasma proteins adsorbs to the surface ofthe implant and the extent of adhesion is controlled by thehydrophilicity of the surface[99].This interaction betweenthe implant and the blood is responsible for plasma proteinaccrual and is a determining factor in the osseoconductionstage of osseointegration. Nanotubular surfaces can bemanufactured to have superhydrophilic surfaces, which inthis regard, would significantly improve the interaction characteristics,besides allowing for easier absorption of nutrientsover the course of the recovery period. Thus, wettabilityof the nanotubular surface is another factor, which makes itsuitable for heightened cell adhesion allowing for easier osseointegrationfollowing implant placement in the body.

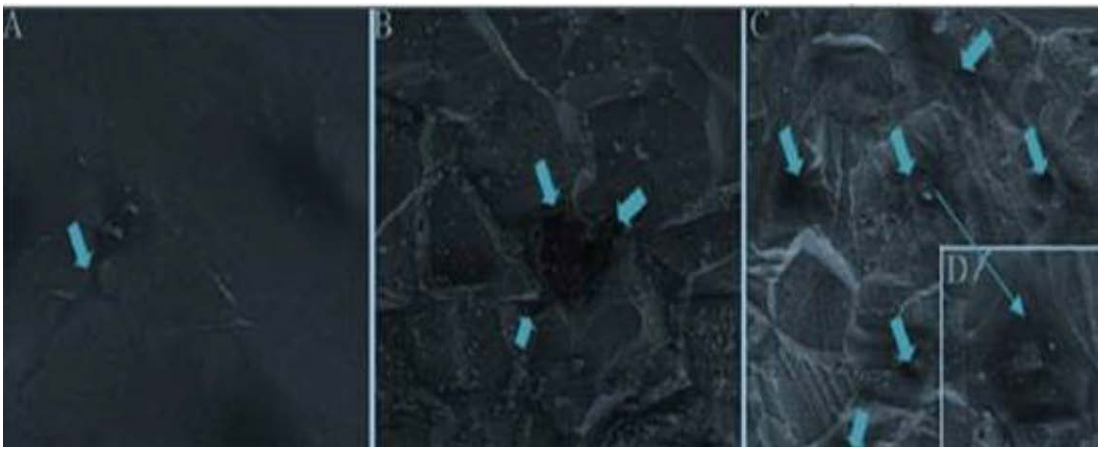


Fig. 25.29. The SEM images after culturing and adhering osteoblasts on threedifferent Ti substrate surfaces for 2 h. (A) Smooth interface. (B) Micro-treatedinterface. (C) Hierarchical micro/nano interface. (D) Local amplification ofsingle osteoblast’s adhesion on micro/nano interface[101].

25.5. Nanotubes for Implant surfaces: The antimicrobialand anti-inflammatory properties

Orthopedic implant installations are executed with utmost attentionbeing paid to sterilizing the wound, the implant and theimplantation site to minimize the risk of infection. This, however,does not always prove to be enough. According to medicalreports, up to 30% of all transcutaneous fracture fixations and13% of bone supplementation procedures are followed by infectionin the recipient’s body causing unbearable pain and necessitatingboth painful and extremely expensive surgery[102].

The chances of infection are determined by thesurgical site and the procedure. The pathogen colonizationof hardware is enhanced by the host response to implantation.The recipient’s body secretes copious quantities ofserum proteins which accumulate over the implant materialand are responsible for cell proliferation and repair ofdamaged tissues. However, these serum proteins are alsoused by microbes for adsorption and virulence[102]. Impartingimplants with drug-eluting properties to ensure localizeddelivery of drugs is a favorable remedy to this problem.

The nanotube approach towards addressing the issueof post-surgery infection can be broadly divided into twocategories- functionalized carbon nanotubes loaded with antimicrobialagents and metal oxide/alloy nanotubes loadedwith nanoparticles/antibacterial drugs. Carbon nanotubeshave been widely studied for their antibacterial drug loadingand releasing property and constitute one of the most efficientways of achieving this goal (Fig. 25.30). Pristine CNTaggregates when in contact with cells cause damage to thecell membrane resulting in cell lysis[103]. However, CNTshave their antibacterial capacity regulated by their physicochemicalproperties. Thus, functionalization with metallicnanoparticles or peptides has been suggested to improve theantimicrobial activity of CNTs[104]. However, the antimicrobialefficiency is diminished due to the low chemical stabilityof the groups used for functionalizing. This could be addressedthrough covalent functionalization with cationic chemicalgroups. It is common to deploy functional groups which havecationic charge and through strong oxidizing potential producehigh oxidative stress in the bacterial membrane. Carbonnanotubes are prime candidates for such antimicrobial drugrelease agencies owing to their elevated levels of cellular uptake,increased surface area and the ability to be easily conjugatedwith different drugs showing superior efficacy, highspecificity and reduced side effects. Common mechanismfor CNT functionalization involves sonication with lysinewith microwave irradiation. Also, it is common to functionalizeCNTs with carboxylic and cephalexin linkers. Other approachesfor functionalizing CNTs involve coating the surfaceof the nanotubes with shallow antibacterial films based onlayer by layer assembly of biological polyelectrolytes such asanionic poly (L-glutamic acid) and cationic poly (L-lysine).

In functionalized state the CNTs are capable of alteringthe glycolysis pathway in the bacterial cell, which isresponsible for producing energy to sustain stress[105]. Apartfrom this, the miniscule nanotube diameter cause partitioningand partial penetration of nanotubes into the cell wall.Oxidative stress is the other significant factor, which contributesto the antimicrobial effect of the nanotubes[105]. Thedrug release characteristics from nanotubes are functionsof the nanotube dimensions and the drug loading concentrationand can be accurately controlled to get release overspecified periods of time and in particular quantities. Antibacterialagents like azithromycin, gemifloxacin, and hydroquinonecan be effectively dispersed using these nanotubes.Also, nanoparticles of Ag, ZnO, CdS and Ag_2S are effectiveantibacterial agents owing to their strong oxidative nature.

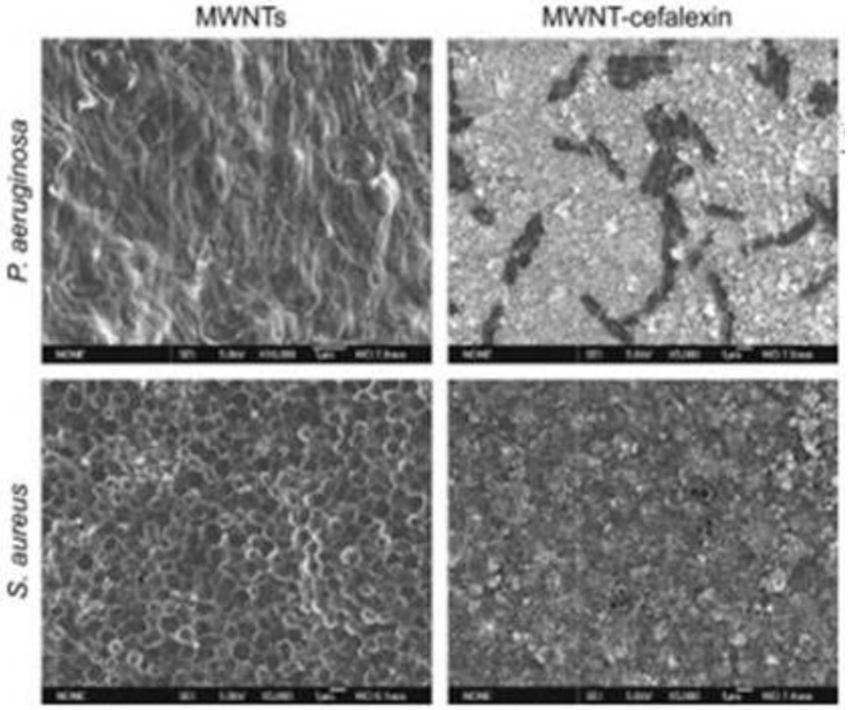


Fig. 25.30. Scanning electron microscopyimages of P. aeruginosa and S. Aureuson the MWNTs and MWNT-cephalexin upon 3-hour exposure[106].

In case of metal oxide or metal alloy nanotubes the anti-infectioneffect comes from four possible sources: the nanostructuringof the nanotubes itself, the functionalizing of the nanotubeswith anti-inflammatory agents, nanotubes decoratedwith antimicrobial nanoparticles and nanotubes loaded withantibacterial drugs. Nanotubes of metal oxides and alloys, especiallythose made of Titanium, Titanium-Al-V alloy and Zirconiahave been widely studied for their inherent antimicrobialproperties. Report[101] shows a substantial decrease in bacterialadhesion and proliferation on the nanotubular surfaces ascompared to polished or micro-rough implant surfaces. Thisis the result of three principal factors: (1) the nanostructuringof the surfaces cause protrusion of nanotubes into the cell wallof the bacteria thus causing cell death by triggering of a stressresponse.(2) The presence of fluorine on the surface of thetitania nanotubes as a remnant from the fluorine containingelectrolyte has antibacterial effect due to the strong oxidizingnature of fluorine that can cause hydrolysis of enzymatic pathwaysthat are responsible for protein adsorption (Fig. 25.31).This can result in muting of genes, which are responsible forcell proliferation. (3)The ability to absorb UV light and subsequentformation of highly reactive radicals such as those ofhydroxyl groups or peroxide groups influence the antibacterialnature of the titania. These allow TiO_2 to possess bactericidalabilities whilst also enabling them to act as cleaning agents forswift degradation and removal of harmful bacterial excreta.

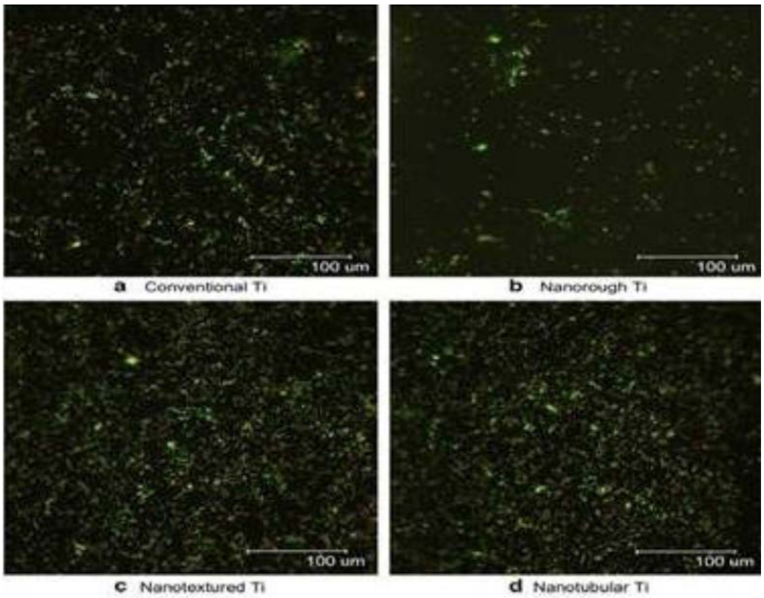
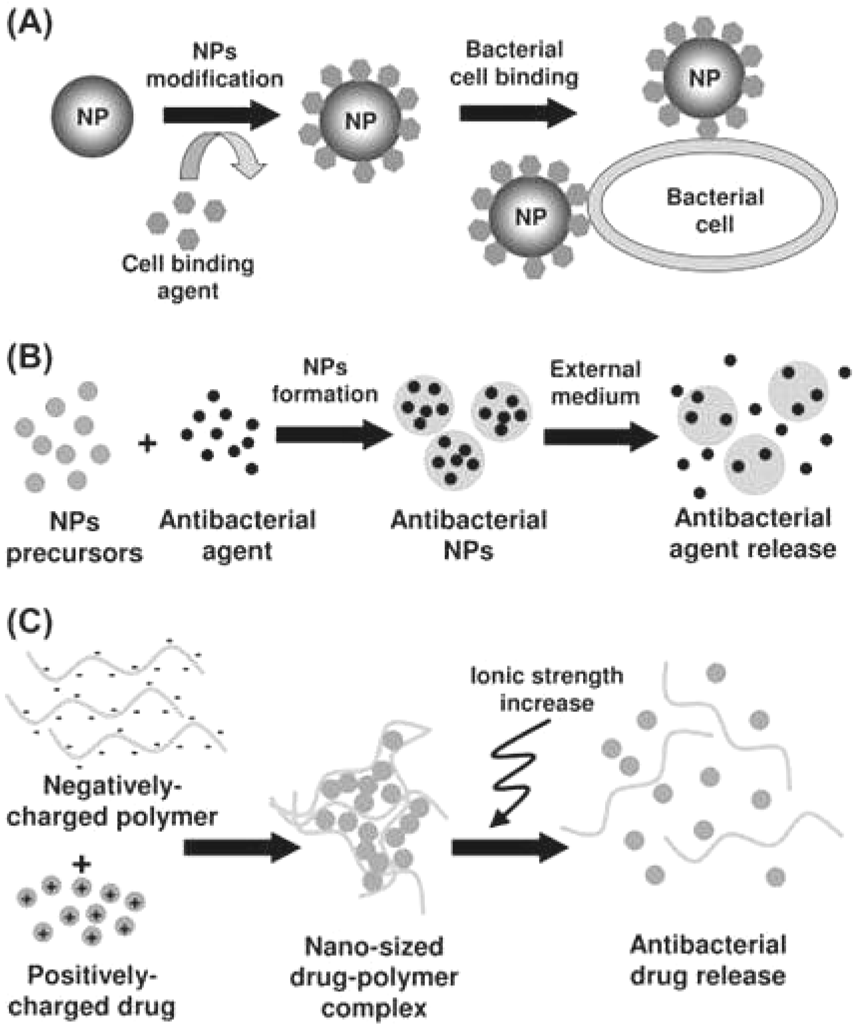


Fig. 25.31. Fluorescent micrographs of decreased S. aureus colonies on (b)nanorough Ti compared to all other substrates and increased bacteria colonieson the (c) nanotextured and (d) nanotubular Ti compared to (a) conventionalTi after 1 hour [107].

Inflammation and its aftermath in the period following implantinstallation surgery constitute serious challenges forthe wellbeing of the patient and effectiveness of the implant.Chronic inflammatory response to debris from the implant orosteogenic cell stress is also another major cause of concernin this context[108]. Deficiency of cellular anti-oxidant capacityis responsible for many pathological inflammatory conditions,which results in high release of reactive oxygen species (ROS).ROS are known to cause osteoclast damage through loweringof the bone mineral density[109]. Most metallic or alloy implantsurfaces despite their biocompatibility cause increase inintracellular ROS levels which can potentially lead to chronicinflammation and reduced bone regeneration[110, 111]. Thus osseoconductionand regeneration in the implant vicinity canbe positively influenced through the use of pharmacologicalanti-oxidizing agents for surface modification of the implants.

Antimicrobial agents such as selenium/silver nanoparticles,ZnO nanoglobules, quantum dots of metal salts,organic molecules like Chitosan are being extensively studiedfor their potential role in limiting the occurrence of postoperativeinfection[112]. Among these silver nanoparticles,ZnO, and chitosan are specifically important. Silver in molecularstate are inert but upon hydrolysis in the body streamrelease Ag+ ions. Silver ions catalyze the oxidation of the hydrogenatoms present in the thiol groups of enzymes thus releasingwater while causing respiratory arrest and death of thecell due to the formation of disulfide bonds between two thiolgroups (Fig. 25.32) [113].Also, the hydrogen bonding amongtwo anti-parallel strands is disrupted when Ag^+ enters thecell and intercalates between the purine and pyrimidine basepairs thereby denaturing the DNA molecule. Inclusion of thesilver ion in the cell of the bacteria causes high amount of oxidativestress, which can result in the extermination of the cell.



Nano-sized drug formulations with antibacterial activity: (A) NPs binding on bacterial cell, (B) drug release from organic NPs matrix and (C) drug release from nano-sized complex.

Fig. 25.32. A schematic drawing showing the various mechanisms of antibacterialactivities exerted by silver nanoparticles[113].

ZnO is a metal oxides possessing substantial photocatalyticand photo-oxidizing capability and hence the capacityto undergo strong chemical interaction with biologicalspecimen[114]. ZnO works as an antibacterial agent by meansof producing highly volatile functional groups called reactiveoxygen species. The bacterial cell wall is attacked by these ionsand under the influence of electrostatic agitation between thezinc nanoglobules and the surface of the cell increased levels ofstress are produced in the membrane ultimately leading to ruptureand death of the cell. The external lipid bilayer in the microbesis ruptured by physical bombardment of the nanoparticlesas well, causing the cytoplasmic material to be drained out.

Chitosans are complicated but useful non-aromaticsemi-crystalline polysaccharides formed through incompleteethanoylation of a ubiquitous natural polymer calledchitin. The unique chemical properties of chitosan are renderedpossible because of the occurrence of the protonableamino group in vicinity of the D-glucosamine residues. Thechitosan, which has a net positive charge (Fig. 25.33), is capableof electrostatic interaction with the negatively chargedradicals and ions present on the microbial cell layers and dueto this interaction the cell permeability is changed[115]. Thiscauses unnatural exchange of material between the cytoplasmof the cell and the surrounding environment, especially ofimportant genetic materials, ultimately leading to cell lysis.Also, the binding of cell DNA with the chitosan by virtue ofthe protonated amino groups causes the bacterial RNA synthesisto be negatively affected. The bactericidal nature ofchitosan is through an amalgamation of these two methods.

Different mechanisms can be used for doping the nanotubeswith chitosan to render the antimicrobial property. Themost common method is that of electrospinning where a dilutedsolution of chitosan is dropped onto the substrate and the coatingof chitosan is obtained on the surface by rotating the substrateat 500~100 rpm for a few seconds[112]. Upon drying the chitosanlayer is firmly ensconced on top of the nanotube surface.

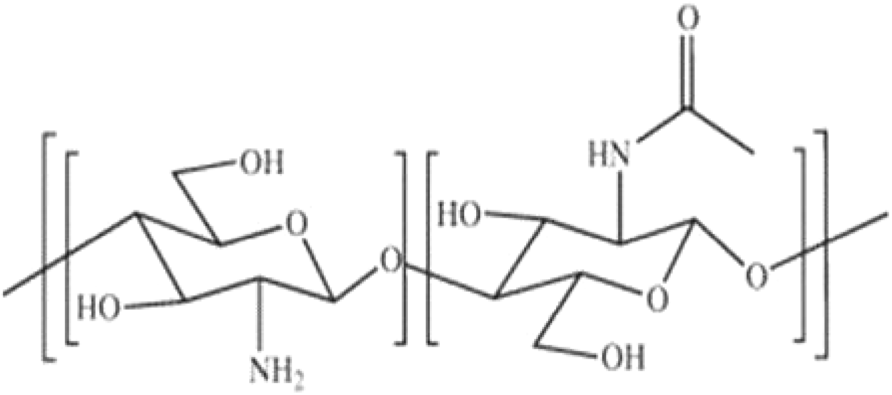


Fig. 25.33. Chemical structure of chitosan showing the protonable groups thatenables its unique chemical reactivity[112].

Another important way of using nanotubes for antimicrobialapplications is through their use as nano-scale reservoirs forcontrolled and calculated drug release. Both carbon nanotubesand other metal-based nanotubes are exceptionally potent optionsfor this purpose due to their increased aspect ratio, functionalizablesurface, high pace of cellular integration and surfacechemical reactivity. In case of carbon nanotubes a functionalizingprocedure is carried out, usually with peptide base compounds,followed by drug loading through lyophilization[116].In case of composite or metal nanotubes, drugs can be loadedsimply by lyophilization or nano–pipetting[117]. The exact relationbetween the rates of release of these drugs is yet to bediscovered but its dependence on the aspect ratio of the nanotube,the concentration of drug loading, and the time of releaseare known[118]. Also, since the chances of infection are highestimmediately after the implant surgery, it is favorable that thesenanotubes tend to release the drug in copious quantities initiallyand then gradually the release is reduced in amount overtime. Commonly used drugs for such applications are Cefuroxime,Gentamicin, Curcumin, Indomethacin and Vancomycin.

25.6. Nanotubes for implant surfaces: The biomimickingeffect for osseointegration

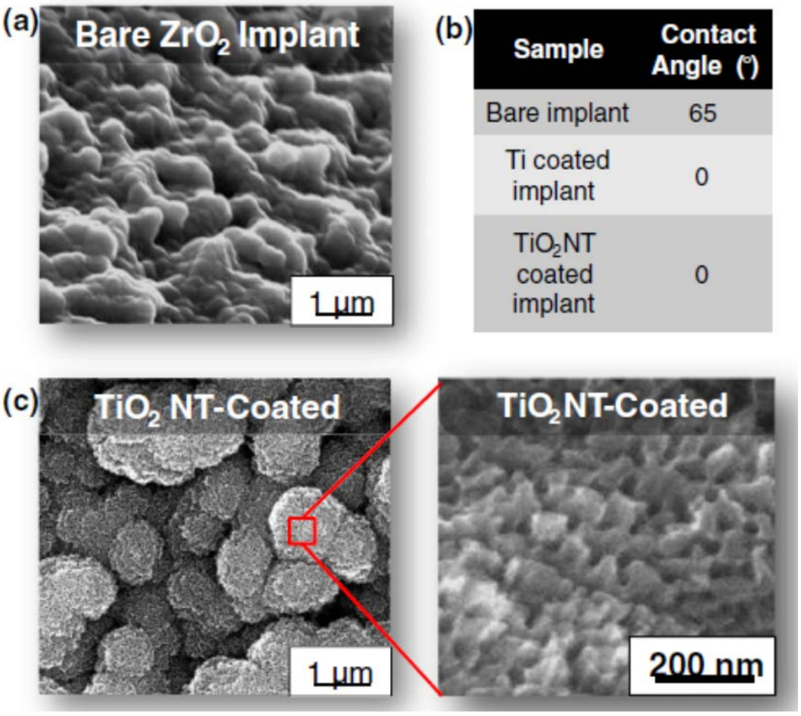
Orthopedic implant materials are usually of two types theimplant surface made of metals, alloys or hard ceramics andthe orthopedic scaffolds for tissue regeneration, which are importantto ensure the implant is integrated into the body. Thislatter is often made of hydroxyapatite because of its chemicalcomposition, which closely mimics that of naturally occurringapatite in the bones[98]. The biocompatibility and the highlevels of osseointegration provided by HA make it an idealsurface to promote growth, proliferation and integration ofbones. This however is offset by its punitive fracture resistance,minimal shear strength and insufficient wear resistance[98].Thus, a second phase material is used as a reinforcing layer toHA. Ceramics and composite layers have been tested but thesefall short on accounts of biocompatibility. However, the incorporationof a nanotube layer as a reinforcing sheath has beenfound to alleviate both the problems. Both carbon nanotubesand titania nanotubes have been extensively studied for thisparticular purpose and have both shown a lot of promise[119].

25.7. Hybrid Micro/Nano-Topography of a TiO_2Nanotube-Coated Implant

It has been well-established that the presence of nanotopographyaffects basic cell behavior in almost all types of mammalian cells[120]. Of particular interest are titanium oxide (TiO_2) nanotubes, sincetitanium is a well-known biocompatible orthopedic material. TiO_2nanotubes have been found not only to significantly accelerate osteoblastcell growth, but also improve bone-forming functionality anddirect stem cell fate. These results are especially promising forwell-established titanium bone implants currently on the market,since aseptic loosening is still a relevant problem in orthopedic implants.The nanotube surface structure can essentially be added toany shape of titanium implant by a simple anodization procedure. Additionally,TiO_2 nanotubes can be grown from a thin film of titaniumdeposited onto another surface. This expands the possible applicationsto not only current titanium implants, but other types of implants aswell.

*TiO_2 nanotube-coated implant preparation[121].*A zirconia ceramic femoral knee component was obtained fromKinamed, Inc. (GEM™ Total Knee System, Cat # 20-120-1001) andwas cleaved into six pieces in a manner which maximized the samplesurface area of relatively flat backside implant surface. In order to obtainregular surfaces for the experiments, the ceramic prosthesis wasslowly cut using a diamond saw into rectangular pieces measuring~1 cm × 2 cm. The backside of the samples was then ground to createa flat and even surface to deposit metallic Ti followed by anodizationto create TiO_2 nanotubes. Two of the pieces were set aside to be usedas controls in this study, while the remaining four pieces were takenfor further processing. Before film deposition, ZrO_2 implant materialswere cleaned successively in acetone and isopropyl alcohol withultrasonication. Ti thin film was vacuum-deposited using DentonDiscovery 18 sputter system. Base pressurewas 1×10^{-6} Torr and substratebeing rotated was heated to 400 °C during sputtering. Plasmapower was 200 W and Ar pressure was 3.0 mTorr. Deposition ratewas 0.25 nm/s and 1 μm thick Ti film was deposited. Electrical contactfor the anodization stepwas provided by copper tape and a lacquer protectivepaint. The TiO_2 nanotube surfaces were created using atwo-electrode-setup anodization process. An organic-based electrochemicalsolution was used which was composed of 0.25 wt.% NH_4Fin 2 vol.% deionized water in ethylene glycol. A platinum electrode(99.9% pure, Alfa-Aesar, USA) served as the cathode. The sampleswere anodized at 20 V for 15 min, followed by awashing stepwith ethanol,and an acetone soak to remove the lacquer paint. The sampleswere then dried at 80 °C overnight and heat-treated at 500 °C for 2 hin order to crystallize the as-fabricated amorphous structured TiO_2nanotubes into an anatase structure. The samples used for all experimentswere sterilized by autoclaving prior to use. The bare zirconiaimplant samples were used as a control after being chemically cleanedby acetone and isopropanol, dried and autoclaved.

Figure 25.34(a) shows SEM micrographs of the as-received zirconia implantsurface, and the TiO_2 nanotube coated implant surface at lowand high magnifications (Fig. 25.34(c)). The micrograph in Fig. 25.34(c)revealsthe nanotopography present on the TiO_2 nanotube surfacewhich is lacking on the bare implant surface in Fig. 25.34(a). Theself-assembled nanotube surface was obtained by anodizing thetitanium sputter-coated zirconia femoral implant at 20 V of anorganic-based electrolytic solution. The nanotube pore size is~50 nm, with a length of ~200 nm.



Sample	Contact Angle (°)
Bare implant	65
Ti coated implant	0
TiO_2NT coated implant	0

Fig. 25.34. Physical characterization of the as-received zirconia implant and TiO_2 nanotube coated implant surfaces. (a) SEM micrographs of bare implant surface (tilted 45° angle).(b) Table with surface contact angle measurements of water droplets on the bare implant, titanium coated implant, and TiO_2 nanotube coated implant surfaces. (c) SEM micrographsof the TiO_2 nanotube coated implant surface at low magnification (left), and high magnification (right).

While the top-view of the TiO_2 nanotube surface appears slightlymore porous in nature, tube morphology was verified to be presentunderneath a porous top layer, as seen in Fig. 25.34(c).This configurationof TiO_2 nanotubes with a porous surface morphology was clearlyexplained by Wang et al. to be the result of an imbalance in two competitivereactions which occur during TiO_2 nanotube formation [122].Further optimization of our anodization conditions could result in atube-like surface morphology. However, for the purposes of thisstudy we analyzed the porous TiO_2 nanotube surface morphology incomparison with the bare implant material.

The SEM images show that a micron-scale surface roughness ispresent on the as-received implant; the nanotube coated implantpossesses the same microtopography, in addition to the nanotopographycreated by the porous nanotube surface. From the SEM microstructuressuch as shown in Fig. 25.34, the surface roughness of the TiO_2nanotube sample in macro/micro scale appears to be on the order ofa few to several micrometers, while the local roughness in nanoscaleis roughly 50–100 nm. The surface contact angle measurements ofwater droplets on each surface revealed that the titanium surfacecoating altered the as-received implant surface to become extremelyhydrophilic in nature, while the as-received implant was only slightlyhydrophilic. The TiO_2 nanotube coated implant was also superhydrophilic,with a contact angle of 0° (Fig. 25.34(b)).

*Osteoblast cell test.*For this study, MC3T3-E1 mouse osteoblast (CRL-2593, subclone 4,ATCC, USA) was used. Each 1 mL of cells was mixed with 10 mL ofalpha minimum essential medium (α-DMEM; Invitrogen, USA) inthe presence of 10 vol.% bovine serum (FBS; Invitrogen, USA) and1 vol.% penicillin–streptomycin (PS; Invitrogen, USA). The cell suspensionwas plated in a cell culture dish and incubated at 37 °C in a5 vol.% CO_2 environment. When the concentration of the MC3T3-E1osteoblastic cells reached confluency, they were seeded onto the experimentalsubstrate of interest (TiO_2 nanotubes or bare implant),which was placed on a 6-well polystyrene plate and stored in a CO_2incubator for 24 and 48 h to observe the cell morphology and adhesion.The concentration of the cells seeded onto the substrate was1.5×10^5 cells per well.

Fluorescein diacetate (FDA; Sigma, USA) staining was conductedto visualize cell viability and to quantify cell spreading. At 24 h afterplating, the cells on the substrates were washed with phosphatebufferedsaline (1× PBS) solution (Invitrogen, USA) and incubatedfor approximately 30 s with FDA stock (5 mg dissolved in 1 mL ofacetone) dissolved in PBS (10 µL/10 mL), and washed once more.The samples were then inverted into new wells, visualized andphotographed using a fluorescence microscope with a green filter(DM IRB, Leica Co., USA). Six fields were randomly chosen fromeach sample. The digital images were stored in a 672 × 512 pixelsfile and imported to a TIF format. Stored images were imported toImageJ image processing program for digital analysis. The numberof adherent cells were counted in each image. Additionally, the totalcell spreading area was quantified.

To investigate cell behavior in response to the as-received and modifiedimplant surfaces, osteoblast cells were seeded on the comparativesurfaces at a cell density of 1.5 × 105 perwell. As indicated in Fig. 25.35, celladhesion and spreading after 24 h of culturewere found to be increasedon the TiO_2 nanotube surface in comparison to the bare implant surface.The plots of the quantification of the fluorescent images (Fig. 25.36) clearlyconfirm these trends. Both the number of adhered cells (Fig. 25.36(a)) andcell spreading area (Fig. 25.36(b))were significantly increased on the nanotubesurface. These results are in agreement with our previouslyreported investigations of osteoblast behavior on TiO_2 nanotube surfaceswhen compared to flat Ti foil [123, 124]. It is likely that the increasein cell adhesion and spreading is a direct result of the superhydrophilicnature of the nanotube-coated surface. Many researchers have demonstratedthe positive effects of superhydrophilic biomaterials on proteinand cell adhesion properties [125, 126].

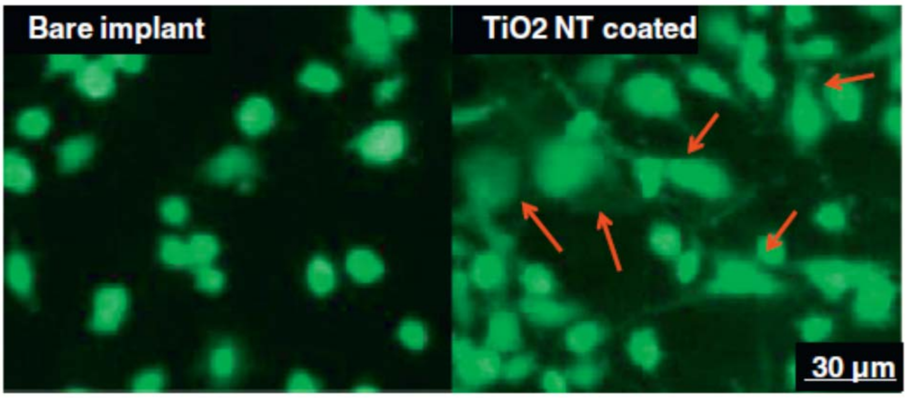


Fig. 25.35. FDA viability of osteoblast cells after 24 h of incubation on the bare zirconia implant and TiO_2 nanotube coating on the zirconia implant. More spreading is evident on thenanotube surface which indicates greater cell adhesion. Red arrows indicate significant cell spreading.

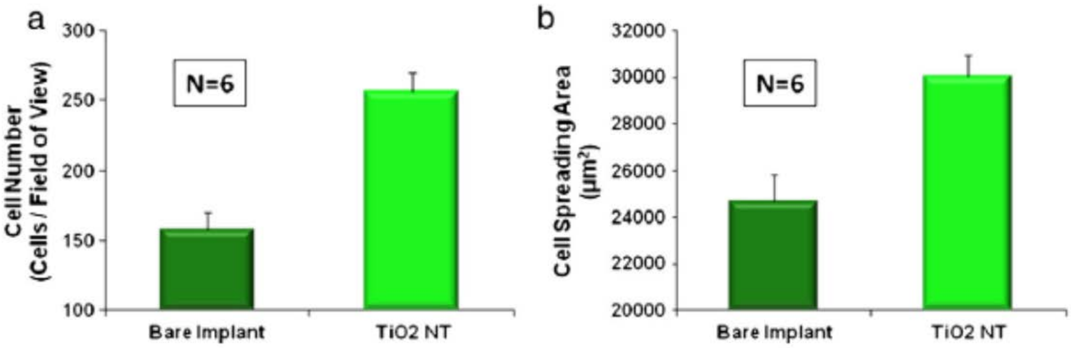


Fig. 25.36. Cell number (a) and spreading area (b) after 24 h of incubation. The bar graphs show the average ± standard error bars. The p-value after performing a *t*-test confirmed astatistical significance ($p < 0.005$).

*Adhesion strength test.*For adhesion strength of TiO_2 nanotube coating on the zirconiaimplant materials, a simple adhesion set-up was utilized. Commercialthermosetting epoxy was attached on the surface of the TiO_2nanotubes and a hook-shaped metal wire end was tightly embeddedinside the epoxy. The adhered area of epoxy was 0.2 cm^2. After theepoxy was fully cured, uniaxial force was applied normal to the surface.Adhesion strength was indirectly determined whether fractureoccurred along the epoxy/TiO_2 nanotube interface or TiO_2 nanotube/implant interface. In result, adhesion strength was at least greaterthan 460 lb/in^2 in which epoxy was detached from TiO_2 nanotubesurface with a sharp facture interface.

The higher degree of cell adhesion and spreading was further confirmedby SEM, as shown in Fig. 25.37. On the bare zirconia implant surface(Fig. 25.37(a) and (c)), the cells appear to be unable to attach tothe surface. Although there are many filopodial extensions present,rarely are they in contact with the surface, as indicated by the red arrows.In contrast, the cells on the TiO_2 nanotube coated implant surface(Fig. 25.37(b) and (d)) are clearly integrating into the surface, andno floating filopodia are evident. The higher magnification images ofthe cell edge (Fig. 25.37(c) and (d)) emphasize the unmistakable cell–surface interaction evident on the TiO_2 nanotube coated implant.

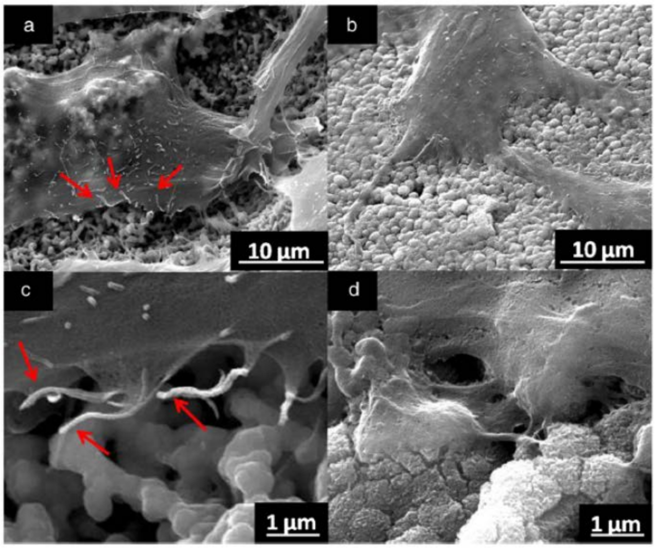


Fig. 25.37. SEM micrographs of osteoblast cells after 24 h of incubation on the bare zirconia implant (a, c), and TiO_2 nanotube coating on the zirconia implant (b, d). (c) and (d) arehigher magnification of the cell edges on the respective surfaces. Red arrows indicate floating filopodia.

Itcan be speculated that the increase in cell–surface interaction canbe contributed to two key factors. Firstly, as mentioned previously,the alteration from a slightly hydrophobic surface to an extremelyhydrophilic surface is likely to enhance cell adhesion and spreadingcapabilities. Secondly, the presence of nanotopography in additionto the bare implant microtopography may have an impact on howthe cells behave. A hierarchical hybrid micro/nano-textured titaniumsurface has been recently considered to create an improved surfacestructure for osseointegration [127]. While the nanotopography canbe assumed to induce an increase in bone functionality, the microtopographycontributes to the mechanical interlocking ability of thesurface.

25.8. Potential Hazards of Using Nanotubes

Despite the obvious benefits of using nanotubular structuresin orthopedic implants there have been persistent doubts andconcerns raised by engineers, doctors and the general public.While many of these are invalidated, there are issues of medicalsignificance that require addressing before large-scale incorporationof nanotubes in implants can be implemented.

Severe and chronic disorders like granuloma, fibrosisof the lung and generation and accumulation of fluid inthe lungs can be caused by long term deposition of nanoparticleswithin the human body[128]. This is of importance sincenanotubular debris studies are limited and it can be logicallyassumed that over an extended period of time it is possiblefor them to undergo degradation and ultimately disintegrateinto nanoparticulate matter. Factors such as physical dimensions,chemical activity and method of preparation are integralto the cytotoxicity considerations for nanotubes. Most of thetoxicity studies conducted on nanotubes concentrate on CNTs,titania Nanotubes and boron nitride nanotubes.

Nanotubes of both single walled and multiwallednature have been studied for their cytotoxic effect (Fig. 25.38). Multiwalledcarbon nanotubes have been reported to form biggerclusters in biological tissues without transportation whileSWCNTs tend to form tiny particles, which are ingested byphagocytes and then transferred to lymph nodes. Nanotubescan easily navigate through capillaries and adhere to bloodvessels. SWCNTs in particular can potentially block the ionictransportation pathways for potassium due to its low diameterand tube-like structure. Cationic surfaces have the capabilityof effecting organism hemolysis and platelet aggregationand most NTs owing to their cationic charge are suspected tohave similar effects, leading to accelerated vascular thrombosis[129]. Oxidative stress can result from activation of oxidativeenzymatic channels by the nanotubes. As a result ROSgeneration is greatly enhanced. Long-chained unsaturatedfatty acids within the cell lining undergo peroxidation whenthe ROS concentration is too high which causes a change inthe permeability of the membrane of the mitochondria aswell as the cell membrane causing rupture and release of cytoplasmiccontent. The dead or damaged cells give out toxinsand debris which accumulate on the surface of blood vessels,RBC, organs like lungs, brain, heart and kidney and the GItract. This is a potential cause of concern due to its carcinogenicnature. Also, nanotubes are known to be responsiblefor retardation of DNA reparation mechanisms owing toDNA strand breaks by suppressing proteins that are responsiblefor DNA repairs. In case of CNTs there have also beenconcerns over possible fibrosis being caused by individualcarbon nanofibers, which remain intermingled with the nanotubes.In case of metal oxide nanotubes there are concernsregarding the release of heavy metal ions from the reactivespecies that remain on the substrate even after repeated cleaning.These ions have been known to potential carcinogens.Apart from the concern over the nanotubesthemselves, there has been widespread fear in the medicalcommunity regarding the health hazards posed by the dopingagents, which are proposed for use as antimicrobial surfacetreatment agents. Silver has been widely studied as a surfacebactericidal agent, however, silver is known to be slightly toxicto humans in elemental form, the most dramatic effect beingobserved in patients with a condition called Argyria where theskin turns blue and patches appear on the conjunctiva[130].However, silver in ionic state is a strong oxidizing agent andmuch like the mechanism in which it causes bacterial cell lysisit can cause irreparable damage to normal human cells[131].The concentration level beyond which silver is toxic to normalhuman cells is still under debate and control of the silver bactericiderelease under this control value is of utmost importance.

The biggest challenge in assessing these risks whenit comes to testing nanotubes for implants is that cytotoxicityassays carried out on implant material surfaces involveonly cells of a certain type incubated in isolation in simulatedbody conditions on the implant surface[132]. While thiscan provide an idea about the level of cytocompatibility betweenthe nanotubes and the cells being studied, it is a verycrude approximation of how the implant containing the nanotubesbehaves in the far more complicated and dynamicenvironment inside the host’s body. In vivo studies are farmore conclusive for such assessment and are currently beingused to study long-term effects of nanotubes. But eventhen, the in vivo studies are limited in number due to ethicalconcerns, infrastructural and economic constraints. Also,human in vivo testing does not seem to be a possibility anywherein the near future due to strict regulatory policies.

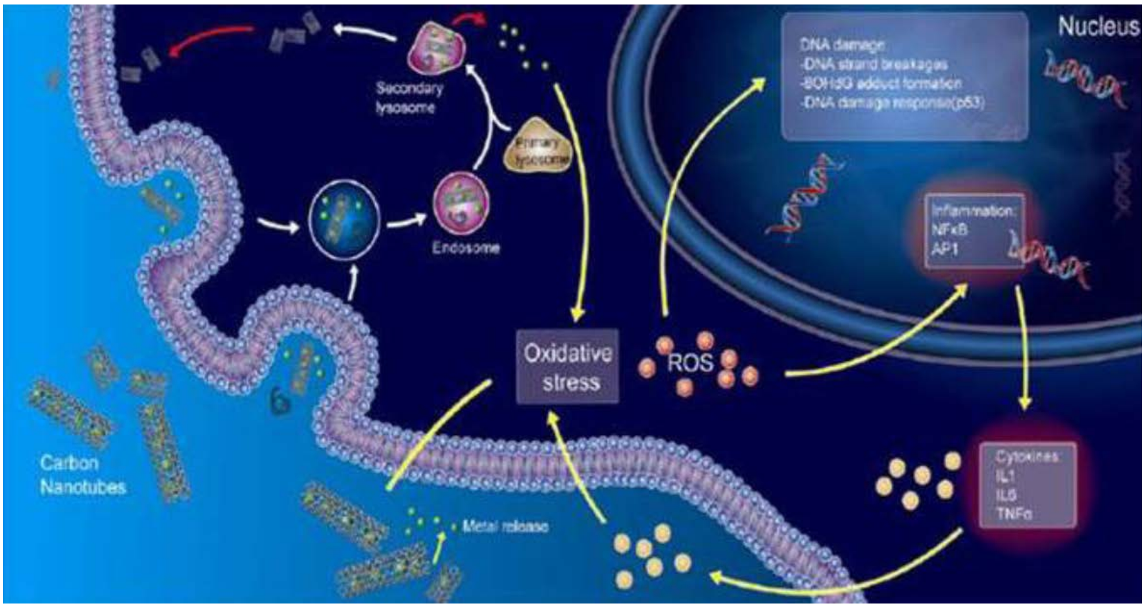


Fig. 25.38. Schematic showing different cytotoxicity mechanisms for CNTs[129].

25.8. Conclusion and Future Prospects

Nanotubes have revolutionized our approach to orthopedicimplants in many ways. They drastically improve the osseointegrationwithin the body, the mechanical strength of the implant,and adding infection resistance or drug delivery capabilitiesto the implants.Nanotube arrays allow for mimickingof the natural micro-nano hierarchical structure in the bodythus stimulating greater levels of bio-compatibility. Essentially,this is a step towards the ultimate goal of self-regenerativemedicine. Through targeted delivery of drugs infections can beresisted by orthopedic implants, thus eliminating the cause forfailure of more than 40% of all medically installed implants.

Despite the positives there are still concerns in themedical community and among the greater masses aboutthe safety aspect of nanotubes. Some of these concerns arelegitimate and basically stem from the fact that any foreignelement is resisted by the body and when of a nano-scale dimensioncan act as a carcinogen by initiating genetic mutation.There are, however, many fears surrounding the use ofnanotubes, which are not based on any scientific evidence andmust be dispelled. In summary, nanotubes can definitely behailed as one of the most promising horizon on the front oforthopedic implants and in the years to come it is expected toyield solutions to many problems that have plagued the field.

Everyday new ideas are being generated for more directed andadvanced uses of the unique properties of nanotubes for orthopedicimplant. Many of these are aimed at improving thecontrol over the behavior of nanotubes with respect to drugrelease. Nickel-titanium alloys, commonly known as Nitinolare known to exhibit strong shape memory effects[133]. Shapememory alloy (SMA) based implants made with nickel-titanium(Nitinol/NiTi) are under intense investigation as a meansof providing the implant real time adaptability to the healingprocess. Dramatic advances in bone mending can be effectedthrough use of SMA as implant biomaterials which allow alterationof the stiffness of the implants at a particular time.Currently studies[134] have been conducted where Titania nanotubeshave been formed on the surface of NiTi implants renderingit biocompatibility and also utilizing the shape memoryeffect of the ‘smart material’. The primary concern in this contextremains regarding the possible cytotoxicity and carcinogenicproperties of Ni^{2+} ions that are potentially releasedfrom the implant over an extended period of time[135, 136].DNA/RNA nanotubes are being investigated for potentialuse as drug delivery vehicles[137]. The inherent biocompatibilityof DNA/RNA strands makes them an effective meansof delivering drugs into the body. Also, of particular interestare protein nanotubes, which can be grown on Alumina templatesand then functionalized with aptamers. These nanotubescan be lyophilized with drugs of choice and the aptamersbased on their selective binding with particular hormonalproteins can provide a basis for targeted drug release[138].

Investigations are being carried out to increase controlof release rate of drugs/anti-inflammatory agents fromthe nanotubes by means of external stimuli such as ultrasound,UV light and laser radiation[139, 140]. UV radiationin particular has been extensively studied for metal oxidenanotubes, which have photocatalytic properties wherebythe radiation intensity can be used to regulate the numberof electron/ion species available on the surface of the implant.This would, in turn, determine the selective chemicalreactivity and adhesion property of the implant coating[139].Since in-vivo testing of implants remains one of the most challengingaspects of nano-orthopedics, alternative ways of testingsuch as organ-on-chip are being looked into. Organs-on-chipare multichannel three dimensional microfluidic chips for culturingcells, which are made of respective organ cells and havetiny hydraulic or pneumatically actuated control mechanismsthat simulate cellular motion in dynamic body conditions[141].

ReferencestoChapter25.

1. D. Ganguly, R. Shahbazian-Yassar, T. Shokuffar. Recent advanced in nanotubes for orthopedic implants. J. Nanotechnology and Smart Materials, 1, 201 (2014).

2. F. Witte (2010) The history of biodegradable magnesium implants: A review.Acta Biomaterialia 5: 1680-1692.

3. Kaplan EG, Kaplan GS, Kaplan DM, Kaplan RK (1984) History of implants.Clinics in Podiatry 1: 3-10.

4. Carter DR, Beaupre GS, Giori NJ, Helms JA (1998) Mechanobiology ofskeletal regeneration. Clinical Orthopaedics and Related Research 355: 41-S55.

5. X. Liu, G.L. Niebur (2008) Bone ingrowth into a porous coated implantpredicted by a mechano-regulatory tissue differentiation algorithm. Biomechanicsand Modeling in Mechanobiology 7: 335–344.

6. Kienapfel H, Sprey C, Wilke A, Griss P (1999) Implant fixation by boneingrowth. Journal of Arthroplasty 14: 355-368.

7. Bishop JA, Palanca AA, Bellino MJ, Lowenberg DW (2012) Assessment ofcompromised fracture healing. Journal of the American Academy of OrthopaedicSurgeons 20: 273-282.

8. Aro HT, Alm JJ, Moritz N, Makinen TJ, Lankinen P (2012) Low BMD affectsinitial stability and delays stem osseointegration in cementless total hiparthroplasty in women: a 2-year RSA stuady of 39 patients. ActaOrthopaedica2: 107-114.

9. Luthringer BJ, Ali FF, Akaichi H, Feyerabend F, Ebel T, et al. (2013) Production,characterisation, and cytocompatibility of porous titanium-basedparticulate scaffolds. Journal of Material Science. Materials in Medicine 24:2337-2358.

10. Del Pozo JL and R Patel (2009) Infection associated with prosthetic joints.The New England Journal of Medicine. 361: 787–794.

11. H. W. Kroto, J. R. Heath, S. C. Obrien, R. F. Curl, and R. E. Smalley, Nature **318**, 162 (1985).

12. W.Kratschmer, L.D.Lamb, K.Foristopoulos,D.R.Huffman, Solid C_{60}: A new form of carbon.Nature 1990, **347**, 354.

13. S.Iijima, Helical microtubules of graphitic carbon,Nature, **354**, 56 (1991).

14. S.Iijima, T.Ichihashi, Single-shell carbon nanotubesof 1-nm diameter. Nature 1993, **363**, 603.

15. D.S.Bethune, C.H.Kiang, M.S.DeVries, G.Gorman, R.Savoy, R.Beyers, Cobalt-catalysedgrowth of carbon nanotubes with single-atomiclayerwalls. Nature 1993, **363**, 605.

16. M.S.Dresselhaus, G.Dresselhaus, P.C.Eklund, Science of Fullerenes and Carbon Nanotubes;Academic Press: San Diego, 1996.

17. Carbon Nanotubes: Synthesis, Structure, Properties and Applications. M.S.Dresselhaus, G.Dresselhaus, Ph. Avouris, Eds.; Topics in Applied Physics, Springer: Berlin, 2001; vol. 80.

18. N.Hamada, S.Sawada, A.Oshiyama, New onedimensionalconductors: Graphite microtubules.Phys. Rev. Lett. 1992, **68**, 1579.

19. A. Oberlin, M. Endo, and T. Koyama, Journal of Crystal Growth **32**, 335 (1976).

20. S. Iijima and T. Ichihashi, Nature **363**, 603 (1993).

21. D. S. Bethune, C. H. Kiang, M. S. Devries, G. Gorman, R. Savoy, J. Vazquez, and R. Beyers, Nature **363**, 605 (1993).

22. K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, Science **306**, 666 (2004).

23. K. S. Novoselov, D. Jiang, F. Schedin, T. J. Booth, V. V. Khotkevich, S. V. Morozov, and A. K. Geim, Proceedings of the National Academy of Sciences of the United States of America **102**, 10451 (2005).

24. A. K. Geim and K. S. Novoselov, Nature Materials **6**, 183 (2007).

25. J.W.G.Wildoer, L.C.Venema, A.G.Rinzler,R.E.Smalley, C.Dekker, Electronic structure ofatomically resolved carbon nanotubes. Nature 1998,391, 59.

26. A.Thess, R.Lee, P.Nikolaev, H.Dai, P.Petit,J.Robert, C.Xu, Y.H.Lee, S.G.Kim, D.T.Colbert, G.Scuseria, D.Toma´nek, J.E.Fischer,R.E.Smalley, Crystalline ropes of metallic carbonnanotubes. Science 1996, 273, 483.

27. H.W.Zhu, C.L.Xu, D.H.Wu, B.Q.Wei, R.Vajtai, P.M.Ajayan, Direct synthesis of long single-walledcarbon nanotube strands. Science 2002, 296,884.

28. P. M. Ajayan, andT. W. Ebbesen, Rep.Prog.Phys.,**60**, (1025-1065, 2003.

29. Yasuda, Ayumu, Kawase, Noboru, andMizutani, Wataru, JournalofPhysicalChemistryB, 106, (51), 2002.

30. S. B. Sinnot, R. Andrews, D. Qian, A. M. Rao, Z. Mao, E. C. Dickey, and F. Derbyshire, Chem.Phys.Lett., **315**, (25-30, 1999.

31. S. H. Jung, M. R. Kim, S. H. Jeong, S. U. Kim, O. J. Lee, K. H. Lee, J. H. Suh, and C. K. Park, Applied Physics A-Materials Science & Processing, **76**, (2)**,** 285-286, 2003.

32. T. W. Ebbesen, and P. M. Ajayan, Nature, **358**, (220-222, 1992.

33. C. Journet, and P. Bernier, Applied Physics A-Materials Science & Processing, **67**, (1)**,** 1-9, 1998.

34. S. Farhat, M. L. La Chapelle, A. Loiseau, C. D. Scott, S. Lefrant, C. Journet, and P. Bernier, Journal of Chemical Physics, **115**, (14)**,** 6752-6759, 2001.

35. H. J. Huang, J. Marie, H. Kajiura, and M. Ata, Nano Letters, **2**, (10)**,** 1117-1119, 2002.

36. Takikawa, Hirofumi, Tao, Yoshitaka, Hibi, Yoshihiko, Miyano, Ryuichi, Sakakibara, Tateki, Ando, Yoshinori, Ito, Shigeo, Hirahara, Kaori, and Iijima, Sumio, AIP Conference Proceedings, 590, (Nanonetwork Materials), 2001.

37. H. Takikawa, M. Ikeda, K. Hirahara, Y. Hibi, Y. Tao, P. A. Ruiz, T. Sakakibara, S. Itoh, and S. Iijima, Physica B: Condensed Matter (Amsterdam, Netherlands), 323, (1-4), 2002

38. Anazawa, Kazunori, Shimotani, Kei, Manabe, Chikara, Watanabe, Hiroyuki, and Shimizu, Masaaki, Applied Physics Letters, 81, (4), 2002.

39. Lee, Seung Jong, Baik, Hong Koo, Yoo, Jae eun, and Han, Jong hoon, Diamond and Related Materials, 11, (3-6), 2002.

40. T. Guo, P. Nikolaev, A. Thess, D. T. Colbert, and R. E. Smalley, Chemical Physics Letters, 243, (1,2), 1995.

41. M. Yudasaka, R. Yamada, N. Sensui, T. Wilkins, T. Ichihashi, and S. Iijima, Journal of Physical Chemistry B, 103, (30), 1999.

42. P. C. Eklund, B. K. Pradhan, U. J. Kim, Q. Xiong, J. E. Fischer, A. D. Friedman, B. C. Holloway, K. Jordan, and M. W. Smith, Nano Letters, 2, (6), 2002.

43. W. K. Maser, E. Munoz, A. M. Benito, M. T. Martinez, G. F. de la Fuente, Y. Maniette, E. Anglaret, and J. L. Sauvajol, Chemical Physics Letters, 292, (4,5,6), 1998.

44. A. P. Bolshakov, S. A. Uglov, A. V. Saveliev, V. I. Konov, A. A. Gorbunov, W. Pompe, and A. Graff, Diamond and Related Materials, 11, (3-6), 2002.

45. C. D. Scott, S. Arepalli, P. Nikolaev, and R. E. Smalley, Applied Physics A: Materials Science &Processing, 72, (5), 2001.

46. Z. F. Ren, Z. P. Huang, J. W. Xu, J. H. Wang, P. Bush, M. P. Siegel, and P. N. Provencio, Science (Washington, D.C.), 282, (5391), 1998.

47. Z. F. Ren, Z. P. Huang, D. Z. Wang, J. G. Wen, J. W. Xu, J. H. Wang, L. E. Calvet, J. Chen, J. F. Klemic, and M. A. Reed, Applied Physics Letters, 75, (8), 1999.

48. Yudasaka, Masako, Kikuchi, Rie, Matsui, Takeo, Ohki, Yoshimasa, Yoshimura, Susumu, and Ota, Etsuro, Applied Physics Letters, 67, (17), 1995.

49. Yudasaka, Masako, Kikuchi, Rie, Ohki, Yoshimasa, Ota, Etsuro, and Yoshimura, Susumu, Applied Physics Letters, 70, (14), 1997.

50. M. Chen, C. M. Chen, and C. F. Chen, Journal of Materials Science, **37**, (17)**,** 3561-3567, 2002.

51. Z. P. Huang, D. Z. Wang, J. G. Wen, M. Sennett, H. Gibson, and Z. F. Ren, Applied Physics AMaterials Science & Processing, **74**, (3)**,** 387-391, 2002.

52. J. B. Park, G. S. Choi, Y. S. Cho, S. Y. Hong, D. Kim, S. Y. Choi, J. H. Lee, and K. I. Cho, Journal of Crystal Growth, **244**, (2)**,** 211-217, 2002.

53. R. L. V. Wal, L. Randall, L. J. Hall, J. Lee, and G. M. Berger, Journal of Physical Chemistry B, 106, (51), 2002.

54. R. L. V. Wal, L. Randall and T.M. Ticich, Journal of Physical Chemistry B, 105, (42), 2001.

55. R. L. V. Wal, G. M. Berger, and L. J. Hall, Journal of Physical Chemistry B, **106**, (14**),** 3564-3567, 2002.

56. P. Avouris, Chemical Physics, **281**, (2-3**),** 429-445, 2002.

57. S.J. Tans, M.H. Devoret, H. Dal, A. Thess, R.E. Smalley, L.J. Geerligs, and C. Dekker, Nature (London), 386, (6624), 1997.

58. P. M. Ajayan and O. Z. Zhou, Carbon Nanotubes, **80**, (391-425, 2001.

59. S.Reich, J.Maultzsch, C.Thomsen, P.Ordejon, Tight-binding description of graphene. Phys. Rev.,B 2002, 66, 035412.

60. R.Saito, G.Dresselhaus, M.S.Dresselhaus, Electronic structure of chiral graphene tubules.Appl. Phys. Lett. 1992, 60, 2204.

61. S.J.Tans, M.H.Devoret, H.Dai, A.Thess,R.E.Smalley, L.J.Geerligs, C.Dekker, Individualsingle-wall carbon nanotubes as quantum wires.Nature 1997, 386, 474.

62. M.Bockrath, D.H.Cobden, P.L.McEuen, N.G.Chopra, A.Zettl, A.Thess, R.E.Smalley, Singleelectrontransport in ropes of carbon nanotubes.Science 1997, 275, 1922.

63. M.R.Buitelaar, W.Belzig, T.Nussbaumer, B.Babic, C.Bruder, C.Scho¨nenberger, Multiple andreevreflection in a carbon nanotube quantum dot. Phys.Rev. Lett. 2003, 91, 057005.

64. M. Ouyang, J. L. Huang, and C. M. Lieber, Accounts of Chemical Research, **35**, (12**),** 1018-1025, 2002.

65. V. Giridharan, Y. H. Yun, P. Hajdu, L. Conforti, B. Collins, Y. Jang, and J. Sankar, Microfluidic platforms for evolution of nanobiomaterials: a review. J. Nanomaterials, Art. ID 789841 (2012).

66. K. Kostarelos, L. Lacerda, C.D. Partidos, M. Prato, A. Bianco, Carbon nanotube-mediated delivery of peptides and genes to cells:translating nanobiotechnology to therapeutics. J. Drug Del. Sci. Tech., 15 (1) 41-47 2005.

67. Cui D.X., Tian F.R., Kong Y., Titushikin I., and Gao H.J.- Effects of single-walled carbon nanotubes on the polymerasechain reaction. - Nanotechnology, **15**, 154-157, 2004.

68. Wang J., Liu G., and Jan M.R. - Ultrasensitive electricalbiosensing of proteins and DNA: carbon-nanotube derivedamplification of the recognition and transduction events. - J.Am. Chem. Soc., **126**, 3010-3011, 2004.

69. Wang S.G., Zhang Q., Wang R., and Yoon S.F. - A novelmulti-walled carbon nanotube-based biosensor for glucosedetection. - Biochem. Biophys. Res. Commun., **311**, 572-576, 2003.

70. Gooding J.J., Wibowo R., Liu J., Yang W., Losic D., Orbons S., Mearns F.J., Shapter J.G., Hibbert D.B. - Protein electrochemistry using aligned carbon nanotube arrays. - J. Am. Chem. Soc., **125**, 9006-9007, 2003.

71. Callegari A., Cosnier S., Marcaccio M., Paolucci D., Paolucci F., Georgakilas V., Tagmatarchis N., Vázquez E., Prato M. - Functionalised single wall carbon nanotubes/polypyrrole composites for the preparation of amperometric glucose biosensors. - J. Mater. Chem., **14**, 807-810, 2004.

72. Callegari A., Marcaccio M., Paolucci D., Paolucci F., Tagmatarchis N., Tasis D., E. Vázquez E., Prato M. - Anion recognition by functionalized single wall carbon nanotubes. - Chem. Commun., 2576-2577, 2003.

73. Yamamoto K., Shi G., Zhou T., Xu F., Xu J., Kato T., JinJ.Y., Jin L. - Study of carbon nanotubes-HRP modified electrodeand its application for novel on-line biosensors. - Analyst, **128**,249-254, 2003.

74. Guo M., Chen J., Liu D., Nie L., Yao S. – Electrochemicalcharacteristics of the immobilization of calf thymus DNA moleculeson multi-walled carbon nanotubes. - Bioelectrochem.,**62**, 29-35, 2004.

75. Gavalas V.G., Law S.A., Ball J.C., Andrews R., BachasL.G. - Carbon nanotube aqueous sol-gel composites: enzymefriendlyplatforms for the development of stable biosensors.- Anal. Biochem., **329**, 247-252, 2004.

76. Chen R.J., Bangsaruntip S., Drouvalakis K.A., Kam N.W., Shim M., Li Y., Kim W., Utz P.J., Dai H. – Noncovalent functionalization of carbon nanotubes for highly specific electronic biosensors. - Proc. Natl. Acad. Sci. USA, **100**, 4984-4989,2003.

77. Hu H., Ni Y.C., Montana V., Haddon R.C., Parpura V.- Chemically functionalized carbon nanotubes as substratesfor neuronal growth. - Nano Lett., **4**, 507-511, 2004.

78. Mattson M.P., Haddon R.C., Rao A.M. - Molecular functionalizationof carbon nanotubes and use as substrates forneuronal growth. - J. Mol. Neurosci., **14**, 175-182, 2000.

79. Correa-Duarte M.A., Wagner N., Rojas-ChapanaJ., Morsczeck C., Thie M., Giersig M. - Fabrication andbiocompatibility of carbon nanotube-based 3D networks asscaffolds for cell seeding and growth. - Nano Lett., 2004.

80. Webster T.J., Waid M.C., Mckenzie J.L., Price R.L., Ejiofor J.U. - Nano-biotechnology: carbon nanofibres as improved neural and orthopaedic implants. - Nanotechnology, **15**, 48-54, 2004.

81. Supronowicz P.R., Ajayan P.M., Ullmann K.R., Arulanandam B.P., Metzger D.W., Bizios R. - Novel currentconducting composite substrates for exposing osteoblasts to alternating current stimulation. - J. Biomed. Mater. Res., **59**,499-506, 2002.

82. Park K.H., Chhowalla M., Iqbal Z., Sesti F. - Single-walledcarbon nanotubes are a new class of ion channel blockers. - J.Biol. Chem., **278**, 50212-50216, 2003.

83. Pantarotto D., Briand J.P., Prato M., Bianco A. -Translocation of bioactive peptides across cell membranes bycarbon nanotubes. - Chem. Commun., 16-17, 2004.

84. Shi Kam N.W., Jessop T.C., Wender P.A., Dai H. – Nanotube molecular transporters: internalization of carbon nanotubeprotein conjugates into mammalian cells. - J. Am. Chem. Soc., **126**, 6850-6851, 2004.

85. Shvedova A.A., Castranova V., Kisin E.R., Schweglerberry D., Murray A.R., Gandelsman V.Z., Maynard A., Baron P. - Exposure to carbon nanotube material: Assessment of nanotube cytotoxicity using human keratinocyte cells. - J. Toxicol Environ. Health - Part A, 66, 1909-1926, 2003.

86. Maynard A.D., Baron P.A., Foley M., Shvedova A.A., Kisin E.R., Castranova V. - Exposure to carbon nanotube material: Aerosol release during the handling of unrefined single-walled carbon nanotube material. - J. Toxicol Environ. Health - Part A, **67**, 87-107, 2004.

87. Warheit D.B., Laurence B.R., Reed K.L., Roach D.H., Reynolds G.A.M., Webb T.R. - Comparative pulmonary toxicity assessment of single-wall carbon nanotubes in rats. - Toxicol. Sci., **77**, 117-125, 2004.

88. Lam C.W., James J.T., Mccluskey R., Hunter R.L. -Pulmonary toxicity of single-wall carbon nanotubes in mice 7and 90 days after intratracheal instillation. - Toxicol. Sci., **77**, 126-134, 2004.

89. Pantarotto D., Singh R., Mccarthy D., Erhardt M., Briand J.P., Prato M., Kostarelos K., Bianco A. - Functionalized Carbon Nanotubes for Plasmid DNA Gene Delivery. - Angew. Chem. Int. Ed. Engl., **43**, 5242-5246, 2004.

90. Pantarotto D., Partidos C.D., Graff R., Hoebeke J.,Briand J.P., Prato M., Bianco A. - Synthesis, structuralcharacterization, and immunological properties of carbon nanotubesfunctionalized with peptides. - J. Am. Chem. Soc., **125**, 6160-6164, 2003.

91. Pantarotto D., Partidos C.D., Hoebeke J., Brown F., Kramer E., Briand J.P., Muller S., Prato M., Bianco A. - Immunization with peptide-functionalized carbon nanotubes enhances virus-specific neutralizing antibody responses. - Chem. Biol., **10**, 961-966, 2003.

92. Wang X, Li Q, Xie J, Jin Z, Wang J, et al. (2009) Fabrication of Ultralongand Electrically Uniform Single-Walled Carbon Nanotubes on Clean SubstratesNano Letters 9: 3137–3141.

93. S. Wu, X. Liu, Kelvin W.K. Yeung, H. Guo, et al. (2013)Surface nano-architectures and their effects on the mechanical properties andcorrosion behavior of Ti-based orthopedic implants. Surface and CoatingsTechnology 233: 13-26.

94. Raymond M. Reily (2007) Carbon Nanotubes: Potential Benefits and Risksof Nanotechnology in Nuclear Medicine. The Journal of Nuclear Medicine 48:1039-1042.

95. Shokuhfar T, Gao Q, Ashtana A, Walzack K, Heiden P, et al. (2010) Structuralinstabilities in TiO_2 nanotubes. Journal of Applied Physics 108: 104310.

96. Y. Hu, K. Cai, Z. Luo, D. Xu, D. Xie, et al. (2012)TiO_2 nanotubes as drug nanoreservoirs for the regulation of mobility and differentiationof mesenchymal stem cells Acta Biomaterialia 8: 439-448

97. X. Luo, C. Matranga, S. Tan, N. Alba, X.T. Cui (2011) Carbon nanotube nanoreservior for controlled release of antiinflammatorydexamethasone. Biomaterials 32: 6316-6323.

98. D. Lahiri,S. Ghosh, A. Agarwal (2012) Carbon nanotube reinforcedhydroxyapatite composite for orthopedic application: A review. MaterialsScience and Engineering 32: 1727–1758.

99. P. Newman, A. Minett, R. Ellis-Behnke, H. Zreiqat(2013) Carbon nanotubes: Their potential and pitfalls for bone tissue regenerationand engineering. Nanomedicine: Nanotechnology, Biology and Medicine9: 1139-1158.

100. B. M. Gumbiner (1996) Cell adhesion: the molecular basis of tissuearchitecture and morphogenesis Cell 9: 345-57.

101. F. Wanga, L. Shic, W.-X. Hea, D. Hand, Y. Yan, et al. (2013)Bioinspired micro/nano fabrication on dental implant–bone interface. AppliedSurface Science 265: 480– 488.

102. Noreen J. Hickok, Irving M. Shapiro (2012) Immobilized antibiotics toprevent orthopaedic implant infections. Advanced Drug Delivery Reviews 64:1165-1176.

103. Aslan S, Loebick CZ, Kang S, Elimelech M, Pfefferle LD (2010) Antimicrobialbiomaterials based on carbon nanotubes dispersed in poly(lactic-coglycolicacid). Nanoscale 2: 1789-1794.

104. Zhou J, Qi X (2011) Multi-walled carbon nanotubes/epilson-polylysinenanocomposite with enhanced antibacterial activity. Letters in Applied Microbiology52: 76-83.

105. Y.-F. Young, H.-J. Lee, Y.-S. Shen, S.-H. Tseng, C.-Y. Lee,et al (2012) Toxicity mechanism of carbon nanotubes on Escherichia coli. MaterialsChemistry and Physics 134: 279-286.

106. X. Qi, P. Gunawan, R. Xu, M. W. Chang (2012)Cefalexin-immobilized multi-walled carbon nanotubes show strong antimicrobialand anti-adhesion properties. Chemical Engineering Science 84: 552-556.

107. S. D. Puckett, E. Taylor, T. Raimondo, T. J Webster(2010) The relationship between the nanostructure of titanium surfaces andbacterial attachment. Biomaterials 31: 706–713.

108. T. Fiedler, A. Salamon, S. Adam, N. Herzmann, J.Taubenheim, K. Peters (2013) Impact of bacteria and bacterial componentson osteogenic and adipogenic differentiation of adipose-derived mesenchymalstem cells. Experimental Cell Research 319: 2883-2892.

109. S. Basu, K. Michaëlsson, H. Olofsson, S. Johansson, H.Melhus (2001) Association between oxidative stress and bone mineral density.Biochemical and Biophysical Research Communications 288: 275-279.

110. Lee Y.H., Lee N.H., Bhattarai G, Oh Y.T., Yu M.K., et al. (2010) Enhancementof osteoblast biocompatibility on titanium surface with terrein treatment.Cell Biochemistry and Function 288: 678-685.

111. Eckhardt A, Gerstmayr N, Hiller KA, Bolay C, Waha C, et al. (2009)TEGDMA-induced oxidative DNA damage and activation of ATM and MAPkinases. Biomaterials 30: 2006–2014.

112. F. Croisier, C. Jérôme (2013) Chitosan-based biomaterialsfor tissue engineering. European Polymer Journal 49: 780-792.

113. M. Moritz, M. Geszke-Moritz (2013) The newest achievementsin synthesis, immobilization and practical applications of antibacterialnanoparticles. Chemical Engineering Journal, 228: 596-613.

114. S. T. Hojati, H. Alaghemand, F. Hamze, F.A. Babaki, R. Rajab-Nia, et al. (2013) Antibacterial, physicaland mechanical properties of flowable resin composites containing zinc oxidenanoparticles. Dental Materials 29: 495-505.

115. Sutha S., Karunakaran G., Rajendran V. (2013) Enhancement of antimicrobialand long-term biostability of the zinc-incorporated hydroxyapatitecoated 316L stainless steel implant for biomedical application. Ceramics International39: 5205-5212.

116. A. Amiri, H. Zare Zardini, M. Shanbedi, M. Maghrebi,M. Baniadam, et al. (2012) Efficient method for functionalization of carbonnanotubes by lysine and improved antimicrobial activity and water-dispersion.Materials Letters 72: 153-156.

117. Maho A., Linden S., Arnould C., Detriche S., Delhalle J., Mekhalif Z. (2012)Tantalum oxide/carbon nanotubes composite coatings on titanium, and theirfunctionalization with organophosphonic molecular films: A high qualityscaffold for hydroxyapatite growth. Journal of Colloid and Interface Science371: 150-158.

118. S. Boncel, P. Zając, K. Koziol KK (2013) Liberation ofdrugs from multi-wall carbon nanotube carriers. Journal of Controlled Release169: 126-140.

119. Kodama A., Bauer S., Komatsu A., Asoh H., Ono S., et al. (2009) Bioactivationof titanium surfaces using coatings of TiO_2 nanotubes rapidly preloadedwith synthetic hydroxyapatite. Acta Biomaterialia 5: 2322-2330.

120. C.J. Bettinger, R. Langer, J.T. Borenstein, Angew. Chem. Int. Ed. Engl. 48 (30)(2009) 5406–5415.

121. C.J. Frandsen, K. Noh, K.S. Brammer, G. Johnston, S. Jin. Hybridmicro/nano-topographyofaTiO_2nanotube-coatedcommercialzirconiafemoralkneeimplantpromotesbonecelladhesioninvitro, MaterialsScienceandEngineeringC 33, 2752-2756 (2013).

122. D. Wang, et al., Chem. Mater. 21 (7) (2009) 9.

123. S. Oh, et al., J. Biomed. Mater. Res. A 78A (1) (2006) 97–103.

124. K.S. Brammer, et al., Acta Biomater. 5 (8) (2009) 3215–3223.

125. T. Miyauchi, et al., Biomaterials 31 (14) (2010) 3827–3839.

126. T. Ishizaki, N. Saito, O. Takai, Langmuir 26 (11) (2010) 8147–8154.

127. L. Zhao, et al., Biomaterials 31 (19) (2010) 5072–5082.

128. Song Y, Li X, Du X (2009) Exposure to nanoparticles is related to pleuraleffusion, pulmonary fibrosis and granuloma. European Respiratory Journal34: 559-567.

129. Liu Y, Zhoa Y, Sun B, Chen C (2012) Understanding the Toxicity of CarbonNanotubes. Accounts of Chemical Research 46: 702-713.

130. N. Hadrup, H. R. Lam (2014) Oral toxicity of silver ions, silvernanoparticles and colloidal silver – A review. Regulatory Toxicology andPharmacology 68: 1-7.

131. P. Chairuangkitti, S. Lawanprasert, S. Roytrakul, S.Aueviriyavit, D. Phummiratch, et al. (2013) Silver nanoparticlesinduce toxicity in A549 cells via ROS-dependent and ROS-independentpathways. Toxicology in Vitro 27: 330-338.

132. Wadhwa S., Rea C., O'Hare P., Mathur A., Roy S.S., et al. (2011) Comparativein vitro cytotoxicity study of carbon nanotubes and titania nanostructureson human lung epithelial cells. Journal of Hazardous Materials 191: 56-61.

133. Huang W.M., Song C.L., Fu Y.Q., Wang C.C., Zhao Y., et al. (2013) Shapingtissue with shape memory materials. Advanced Drug Delivery Reviews 65:515-535.

134. R. Hang, X. Huang, L. Tian, Z. He, B. Tang(2012) Preparation, characterization, corrosion behavior and bioactivity ofNi_2O_3-doped TiO_2 nanotubes on NiTi alloy. Electrochimica Act 70: 382-393.

135. Y.-J. Park, Y.-H. Song, J.-H. An, H.-J. Song, K. JAnusavice (2013) Cytocompatibility of pure metals and experimental binarytitanium alloys for implant materials. Journal of Dentistry 41: 1251-1258.

136. S. Gu, B. Yan, L. Liu, J. Ren (2013) Carbon nanotube–polyurethane shape memory nanocomposites with low trigger temperature.European Polymer Journal 49: 3867-3877.

137. W. Cheung, F. Pontoriero, O. Taratula, A. Chen M.,H. He (2010) DNA and carbon nanotubes as medicine. Advanced DrugDelivery Reviews 62: 633-649.

138. Hou S., Wang J., Martin C.R. (2005) Template-Synthesized Protein Nanotubes.Nano Letters 5: 231-234.

139. M. S. Aw, D. Losic (2013) Ultrasound enhanced release oftherapeutics from drug-releasing implants based on titania nanotube arrays.International Journal of Pharmaceutics 443: 154-162.

140. Z. Markovic M., L. Harhaji-Trajkovic M., B. Todorovic-Markovic M., D. Kepić P, K. Arsikin M., (2011) In vitro comparisonof the photothermal anticancer activity of graphene nanoparticles and carbonnanotubes. Biomaterials 32: 1121-1129.

141. Š. Selimović, M. R. Dokmeci, A. Khademhosseini (2013) Organson-a-chip for drug discovery. Current Opinion in Pharmacology 13: 829-833.

Chapter 19

Transistors from carbon nanotube and integrated circuits built with nanotube transistors

Carbon nanotube (CNT) is considered as a cylinder formed by rolling a piece of graphene. It has been studied for many aspects in recent years. Researchers have demonstrated that CNTs possess unique properties such as high current density exceeding 109 A/cm2, excellent thermal conductivity of about 6600W/m·K, ballistic transport, high mechanical flexibility with extremely high Young's modulus, about 1.2 TPa, and high photo transparency which can be higher than 90%.

Because of these remarkable properties, CNTs, especially the single-walled carbon nanotubes (SWCNTs), have been expected to work as wiring and interconnect material, as well as alternative channelmaterial forfieldeffect transistors.They are the candidates for very large-scale integration circuits. Carbonbased nanoelectronics have been considered as the most promising emerging research device (ERD) technologies targeting for commercial demonstration in the 5–10-year horizon. Thin film transistors (TFTs) are widely used for flat panel display, flexible electronics, and sensor applications. The most common TFTs are using amorphous silicon (α-Si) or polysilicon as transistor channel. Amorphous silicon TFTs can satisfy the requirements of large area, lowto-middle displaying speed, good uniformity, and fair stability.

Poly-silicon TFTs own an advantage of high mobility. However, either of these two types of TFTs has its critical limitation so that neither of them can be widely applied in the more advanced displays. Amorphous silicon is sensitive to light. Also, the carrier mobility of the α-Si device is less than 2 cm2/V·s, which cannot satisfy the requirement of the high-speed display with a frame rate of 120Hz or higher. Though poly-silicon TFT's mobility is large enough, it lacks flexibility and transparency, which is fatal for flexible devices.Metal oxide TFT is one of the innovations to meet the requirements of mobility and transparency simultaneously. However, present metal oxide TFT is instable because it is sensitive to light, temperature, and water vapor. Also, it is instable and subjects to the negative bias illumination stress (NBIS) which can cause threshold voltage to shift to the negative voltage direction. Single-walled carbon nanotubes possess high mobility, high transparency, and good flexibility simultaneously.

These attractive properties satisfy the requirements of thin film transistors, making CNTs the most promising candidates as the high-performance TFT channel material. This research field has attracted much attention and intensive efforts since 2004. The number of published papers on carbon nanotube thin film transistors has increased year by year. At this chapter we present the status of the fabrication technologies and the applications of single-walled carbon nanotube thin film transistors. Their electrical properties, mechanical properties, and flexibility are studied.

Carbon is an outstanding element. The isolated carbon atom has a ($1s^2 2s^2 2p^2$) configuration with filled 1s and 2s states and two electrons in the 2p state. Through sp^2 hybridization of the atomic orbitals, two neighboring atoms can bond strongly via both s- and p-states, with the latter involving the p_z orbitals that are oriented perpendicular to the intersection line between the atoms. Due to the high p-bonding energy, graphite is energetically slightly more stable than sp3-bonded diamond-like structures under ambient pressure and temperature. Although its neighbors to the right and left in the periodic table (boron and nitrogen) can also form strong p bonds, only carbon occupies the favored position of having exactly four electrons and requiring exactly four bonds to attain a closed shell. Indeed, the ability of carbon to form a wide variety of anisotropic and stable 2D structures is in contrast to clusters of nearly all other elements, which are essentially always 3D. The fundamental 2D carbon structure is graphene, a honeycomb atomic layer (Fig. 1A) representing the basic structural constituent of graphite.

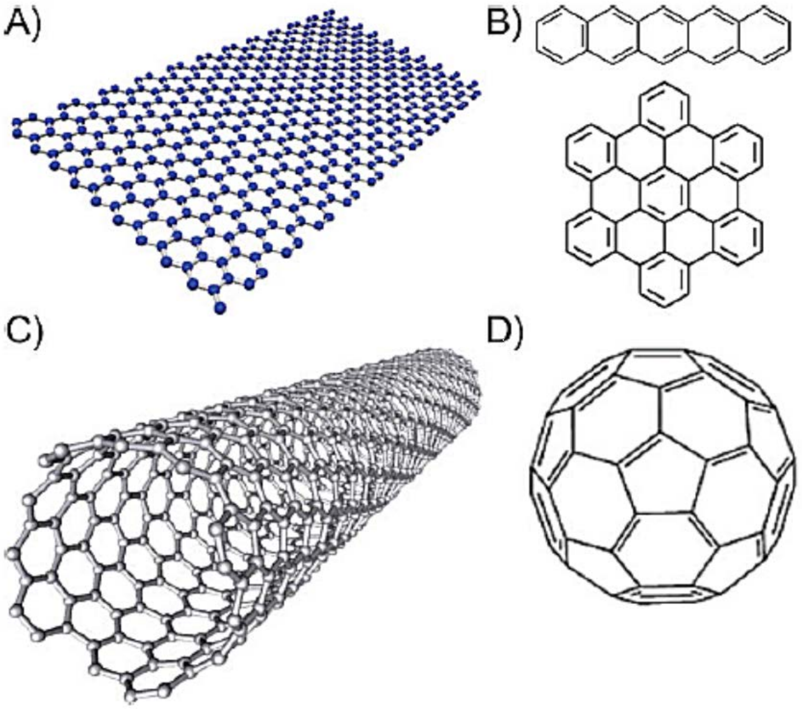


Fig.1. Chemical structures of A) a graphene sheet with its 2D delocalized electron system, B) pentacene (top) and hexabenzocoronene (bottom) as two representative p-conjugated molecules that can be conceived as small graphene sheets with hydrogen-saturated edges, C) a chiral single-wall carbon nanotube with a quasi-1D electron system, and D) the C60 fullerene representing a 0D electron system.

Graphene can be conceptually viewed as an indefinitely extended, 2D aromatic macromolecule. Capping off the edges of very small graphene sheets by hydrogen atoms yields stable p-conjugated molecules such as pentacene, hexabenzocoronene, or picene (Fig. 1B). Further to this, a rich family of structures can be formed by distorting the grapheme sheet in the third dimension.

Graphene, being a single atom layer, can be bent without appreciably changing the in-plane bond lengths, rendering the energy cost for the distortion relatively small. In order to eliminate dangling bonds, graphene must be wrapped around itself to form a closed structure without edges. As one possibility, the roll-up of graphene results in carbon nanotubes (Fig. 1C), whose ends may be closed by appropriate caps. Another option involves bending the graphene sheet in two dimensions, which yields fullerenes (such as C_{60}, in Fig. 1D).

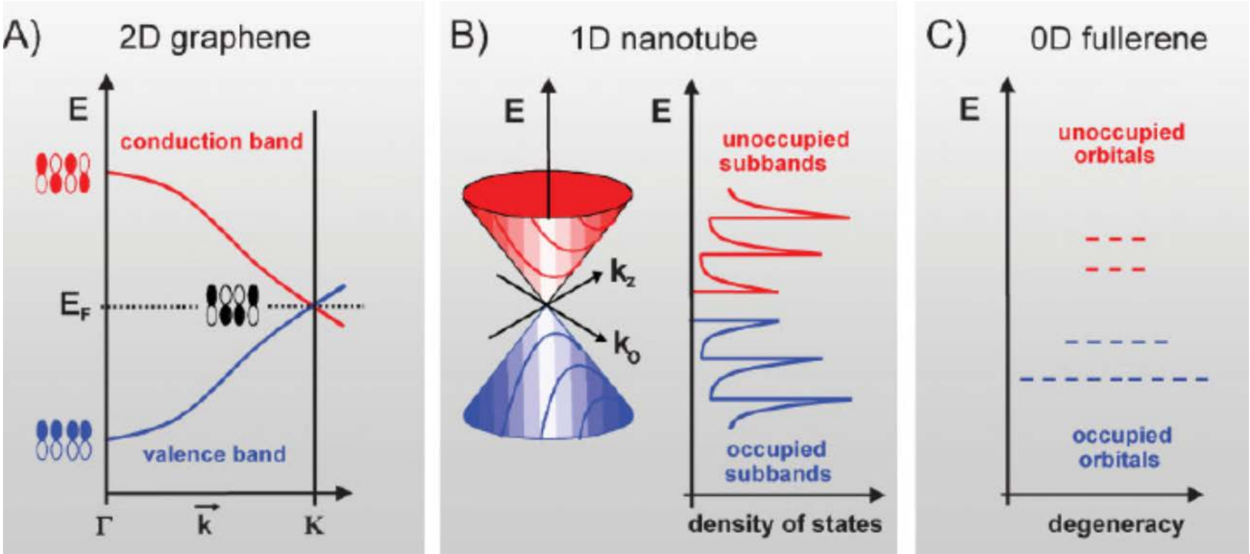


Fig. 2. Schematic representation of A) a 1D section of the 2D band structure of graphene, B) the formation of 1D sub-bands in a carbon nanotube by restricting the wave vectors in circumferential direction and the resulting electronic DOS displaying van Hove singularities at the sub-band minima, and C) the set of discrete energy levels arising from the 3D confinement in C_{60}.

Similar to the nanotube caps, this folding into a closed cage requires the presence of pentagons in addition to hexagons. The electronic-band structure of graphene combines semiconducting and metallic characteristics, as it can be conceived both as a metal with vanishing Fermi surface and as a semiconductor with a vanishing band gap. This peculiar property stems from graphene's unit cell, comprised of two carbon atoms, which renders the p- and π-bands indistinguishable at the Fermi energy, such that they remain equal in energy and the two bands touch each other at the K-point in the Brillouin zone (Fig. 2A). Moreover, the close carbon–carbon distance in graphene leads to a pronounced interatomic overlap that spreads these bands out over a wide energy range, resulting in a rather large Fermi velocity of the electrons ($v_F \sim 10^6$ m s^{-1}). As a consequence of the linear dispersion up to ~1 eV around the Fermi level, the electrons and holes in graphene behave as if they have no mass, and their electronic density of states (DOS) depends linearly on energy. The 1D band structure of a carbon nanotube can be derived by restricting the 2D energy dispersion of graphene to those wave vectors that remain allowed by quantization along the tube circumference (Fig. 2B). If the corresponding slices pass through graphene's K-point, the nanotube is metallic, and otherwise semiconducting. In the simple tight-binding model, the band gap Eg of a semiconducting tube is given by

$$E_g = \frac{4\hbar \upsilon_F}{3d_{CNT}} = \gamma \left(\frac{2R_{C-C}}{d_{CNT}} \right),$$

where γ is the hopping integral (~3 eV), R_{C-C} the C–C bond length, and d_{CNT} is the tube diameter. The electronic DOS of a nanotube features van Hove singularities at each band minimum, associated with $1/\sqrt{(E-E_0)}$ - like tails reflecting the free-electron character along the tube axis. Finally, in the p-conjugated molecules, the quantum confinement in all three dimensions gives rise to a set of discrete energy levels, as exemplified for C_{60} in Fig. 2C.

While all the carbon nanostructures in Fig. 1 have made their way into prototype or even close-to-application electronic devices, the inherent charge-transport characteristics and the factors that limit the charge-carrier mobility differ significantly among them. These differences directly mirror the electronic dimensionality and achievable dimensions of the respective p-systems. Graphene sheets can be easily prepared with micrometer sizes sufficient to connect electrodes accessible through standard lithography, thus opening the possibility to transport charges along graphene's extended p-system. By application of a control (gate) voltage, the charge carriers in graphene can be continuously tuned between electrons and holes in concentrations as high as 10^{13} cm^{-2}.

However, the zero band gap of grapheme prevents effective current switching, which is essential for field-effect transistor (FET) operation. It is well documented that the carrier mobility in graphene is strongly influenced by the environment. Graphene sheets resting on a Si/SiO_2 substrate routinely reach mobilities of the order of 10 000 cm^2 V^{-1} s^{-1}. In such devices, carrier scattering is likely to arise from charged surface impurities, which are typically present on Si/SiO2 substrates with a density of ~10^{12} cm^{-2}.

Scattering by trapped charges is expected to be especially effective close to the Dirac point, where the low carrier densities provide only weak screening. As an alternative scattering source, the formation of static ripples within the graphene sheet or the presence of graphene impurity states with an energy close to the charge neutrality (Dirac) point have been proposed.

Very recently, immense electron mobilities approaching 200 000 cm2 V^{-1} s^{-1} have been detected at liquid-helium temperature in freely suspended graphene sheets (Fig. 3), indicative of near-ballistic transport over a micrometer length scale. The suspended sheets, moreover, display a temperature-dependent minimum conductivity that approaches the theoretically predicted value of $4e^2 \cdot \pi h^{-1}$ at low temperatures (Fig. 3B).

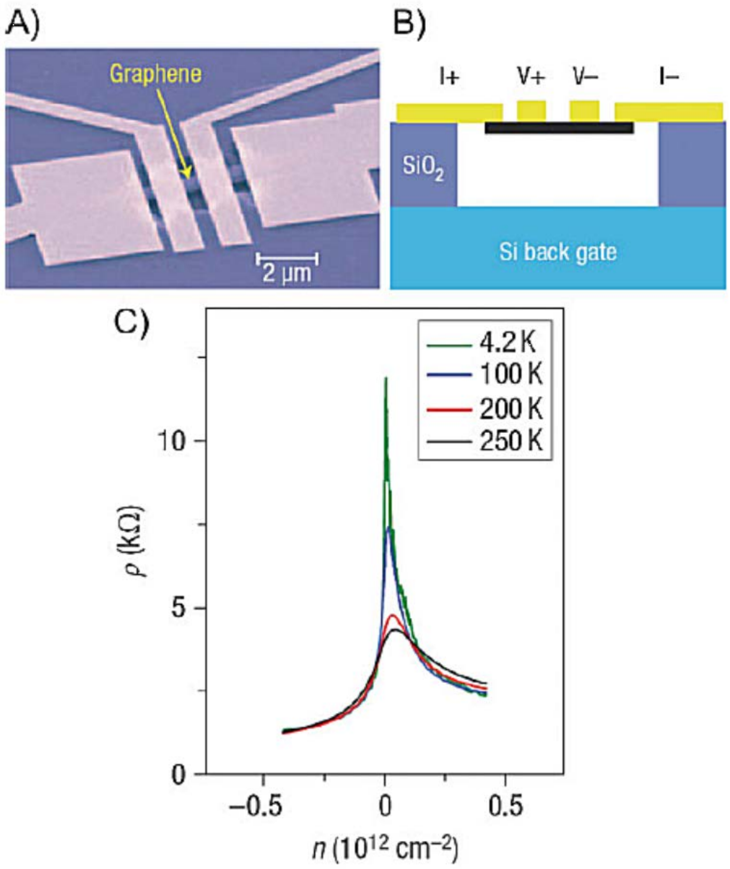


Fig. 3. A) Electron microscopy image of a suspended graphene sheet in four-probe contact configuration implemented on a Si/SiO$_2$ substrate; B) schematic side view of the device. C) Plot of the resistivity of a suspended graphene sheet (0.5mm length and 1.4mm length), measured at four different temperatures, as a function of carrier density.

Notably, current-induced thermal annealing is a means to further remove surface contaminants. With rising temperature, the carrier mobility is notably reduced, which has been attributed to acoustic phonon scattering or extrinsic scattering by surface phonons at the underlying SiO$_2$ layer, with the latter mechanism setting in at temperatures above 200 K. In contrast to extended graphene sheets, the much smaller size of p-conjugated molecules renders it almost impossible to utilize the extended wave functions within the molecular plane for charge transport, unless considerable effort is spent, such as in specifically designed experiments on single molecules positioned inside nanogaps.

Therefore, technologically useful devices have to rely upon p-stacked molecule ensembles in the form of thin films, wherein the charge-transport direction is preferably perpendicular to the p-systems.

Due to the fact that the molecular interactions in organic films are governed by relatively weak van der Waals forces and large intermolecular distances, the carrier-transport efficiency is expected to be substantially smaller than in covalently bound solids. In addition, the p-stack arrangement renders the charge transport highly sensitive with respect to structural defects, so that even moderate distortions, such as a single misplaced molecule, can significantly impede the charge transport through the stack. Indeed, the carrier field-effect mobilities of early organic transistors were quite small, on theorder of 10^{-5}–0.01 $cm^2 V^{-1} s^{-1}$.

Because of the small bandwidth of π-stacked organic semiconductor films, charge transport is typically described by carrier hopping. For organic semiconductors with mobilities of less than about 0.01 $cm^2 V^{-1} s^{-1}$, the variable-range hopping (VRH) model has been developed. However, in organic semiconductors with optimized thin-film morphology, mobilities up to ~1 $cm^2 V^{-1} s^{-1}$ are routinely achieved. To describe the electronic transport in organic films with such large mobilities (above about 0.1 $cm^2 V^{-1} s^{-1}$), the multiple trapping and release (MTR) model has provenvaluable, although the existence of a true band-like transport in organic semiconductors is still under debate.

Similar to graphene, single-walled carbon nanotubes (SWCNTs) can be easily obtained with sufficient size to connect electrode pairs defined by standard lithography. Chemical vapor deposition (CVD) methods enable the routine growth of tens of micrometer long nanotubes from controlled surface sites by catalyst patterning on a substrate. As a major difference to p-stacks of conjugated molecules, a CNT has a stronger capability to tolerate defects, because their impact is averaged out to some extent along the tube circumference, rendering them quite robust against elastic scattering. High-quality SWCNTs exhibit at low temperatures ballistic transport over a micrometer length scale, as manifested by Fabry–Pe´rot-like interferences in the gate-dependence of the conductance.

At low applied bias, the charge transport is determined by inelastic scattering involving acoustic phonons, which imparts a 1/T dependence of the carrier mobility on temperature in semiconducting nanotubes.

Due to the weak electron-acoustic phonon coupling, semiconducting SWCNTs show extraordinary field-effect mobilities of up to several 10 000 $cm^2 V^{-1} s^{-1}$.

FETs are by far the most common electronic devices. Currently, about 10^{19} FETs are manufactured per year, most of them (~99%) on single-crystal silicon wafers as the building blocks of integrated circuits for microprocessors, solid-state memories, or mobile phones.

The silicon serves as the substrate and as the semiconductor in which a channel of charge carriers (either negative electrons or positive holes) is accumulated by a transverse electric field from a conducting gate electrode. The gate electrode is usually a metal (or heavily doped polycrystalline silicon deposited by CVD) and is separated from the semiconductor by an insulating layer, the gate dielectric. Since, the carrier channel is located in close vicinity of the semiconductor/dielectric interface, the mobility of carriers in the channel is usually smaller than in the semiconductor bulk, due to interface scattering. In the case of silicon, a gate dielectric with excellent bulk and interface quality is readily obtained by oxidizing the silicon surface at high temperature in dry oxygen. The thickness of the resulting SiO_2 layer can be scaled down to a few nanometers. This gate-dielectric scaling is necessary as the channel length of the transistor (the distance between the source contact, at which carriers are injected into the channel, and the drain contact, at which carriers are extracted) is further and further reduced. While the first microprocessors in the early 1970s had FETs with a channel length of 10 mm, 64-bit multicore processors currently in production utilize FETs with a channel length of 35 nm. In order to assure that the carrier density in the channel is controlled by the transverse gate field rather than by the lateral drain-source field, the thickness of the gate dielectric must also be aggressively reduced.

As a result of the continuous lateral and vertical scaling, state-of-the-art silicon FETs are nanoelectronic devices, but theminiaturization comes at a price. As the gate dielectric thickness is reduced below 2 nm, charge leakage due to quantummechanical tunneling increases to more than 1 A cm^{-2}, causing severe problems related to off-state current, power consumption, and heat dissipation. The leakage-current problem can be alleviated by replacing the thermally grown SiO_2 with a deposited metal oxide with larger permittivity (such as HfO_2), so that the same gate capacitance corresponds to a thicker dielectric (and hence smaller gate leakage). However, a deposited dielectric renders it more difficult to control the interface quality and scattering, so that the carrier mobilities are further reduced, to about 200 $cm^2\ V^{-1}\ s^{-1}$ for electrons and 40 $cm^2\ V^{-1}\ s^{-1}$ for holes, and future scaling is likely to reduce the mobilities even further.

Therefore, alternative semiconductors that are compatible with silicon technology platforms but provide higher carrier mobilities than silicon are of great interest, especially if they fit naturally into a nanoscale FET architecture. An important milestone in this direction was the fabrication of the first carbon nanotube-based FET (Fig. 4) in 1998.

Since graphene and carbon nanotubes are characterized by very large mobilities and near-ballistic transport, these materials hold great promise for next-generation nanoelectronics. While graphene and individual carbon nanotubes are of interest as an extension of silicon-based nanoelectronics, thin films of p-conjugated hydrocarbons offer the possibility of implementing electronic functionality in the form of FETs and integrated circuits on substrates other than silicon.

For example, because all the functional materials required to build organic transistors can be deposited and patterned at or near room temperature, these devices can be manufactured on flexible, conformable, and even stretchable plastic substrates.

This opens a range of new applications, such as roll-up displays, spherical focal-plane imaging arrays, and artificial skin (providing robots maneuvering in unstructured environments with a human-like sense of touch and temperature). These applications require that FETs and circuits be distributed on unconventional substrates over large areas, but they typically do not require high-frequency operation, so that the carrier mobilities of organic semiconductor films ($\sim 1\ \mathrm{cm}^2\ \mathrm{V}^{-1}\ \mathrm{s}^{-1}$) are usually sufficient. It is noteworthy that in addition to nanoscale FETs (based on graphene or carbon nanotubes) and flexible FETs (based on thin films of conjugated hydrocarbons), carbon-based transistors are also useful as chemical or biological sensors, because the environmental sensitivity of carbon-based molecules can be efficiently tailored by synthetic chemistry. In the following sections, the focus will be laid on important recent advances in the development of FETs comprising the three different types of carbon nanostructures.

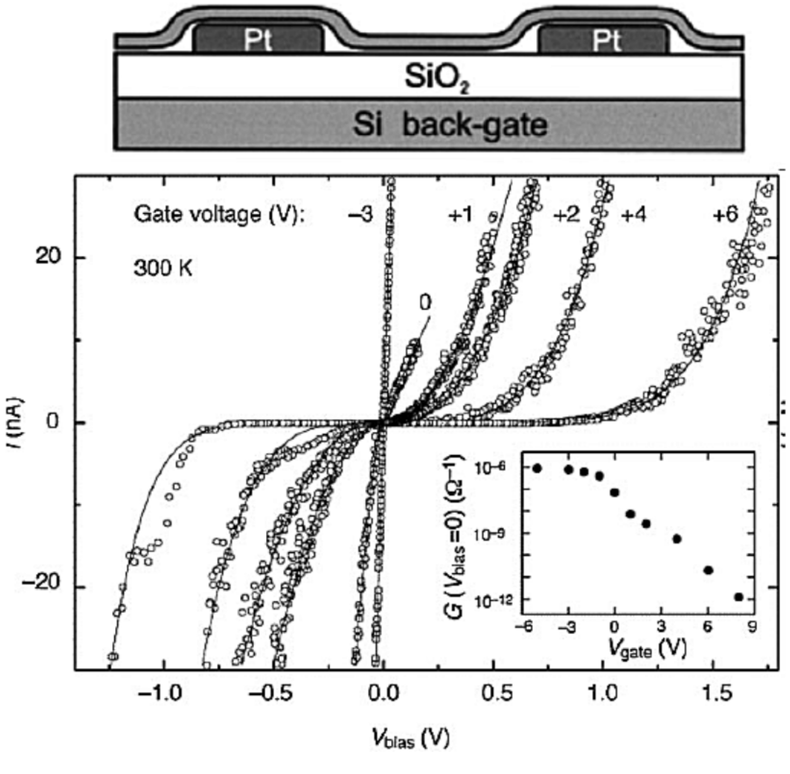


Fig. 4. Schematic depiction of a FET comprising a single semiconducting SWCNT as conducting channel (top) and the gate-dependent current (I)-voltage (V) curves of the SWCNT-FET at room temperature (bottom). The inset shows the transfer characteristics of the device. Reproduced with permission from.

Although the operation mechanism of SWCNT-FETs is not yet fully understood, there is general consensus that they typically function as Schottky-barrier transistors rather than conventional bulk transistors. In the former type of transistor, the applied gate potential modulates not only the carrier density, but also the transmission through the Schottky barriers at the contacts, that is, both contact and channel resistance are affected by the gate. The Schottky barrier for one carrier type can be minimized through a proper choice of contact metal, as has been first demonstrated by palladium contacts enabling nearly barrier-free access to the valence band of semiconducting tubes. More recently, close-to-perfect contacts to the tube's conduction band have been accomplished using scandium, rendering both p- and n-type SWCNT-FETs available.

The influence of the Schottky barriers can be reduced by selectively doping the contact regions, as has been realized on the basis of strong chemical charge-transfer between the tube and adsorbed molecules. In close correspondence to graphene devices, a convenient method to estimate the Schottky-barrier height in a semiconducting nanotube is SPCM performed under gate modulation. It relies upon the fact that zero photocurrent is generated in the transistor off-state corresponding to the flat-band situation, whereas maximum photocurrent is observed at the switching threshold, that is, when the conduction-band edge aligns with the Fermi level of the contacts, for instance.

In addition, SPCM has been used to determine the depletion layer width in the different device-operation regimes. In order to optimize the gate switching (minimize the sub-threshold swing S), the capacitive coupling of the gate electrode has to be enhanced. In the optimum case, the classical electrostatic capacitance Cg would become larger than the quantum capacitance Cq of the tube ($Cq=10^{-16}$ F mm^{-1}), and thus dominate the switching action. Strong gate coupling has been implemented through a top gate separated from the nanotube by a layer a few nanometers thick of a high-permittivity (high-k) dielectric, such as TiO_2 (k=80), or HfO_2 (k=25). The corresponding SWCNT-FETs display a sub-threshold swing close to the theoretically expected value of S=60mV per decade at room temperature, along with a large-scale transconductance of up to 5000 S m^{-1}.

An alternative promising approach takes advantage of the excellent insulating capability of high-quality organic SAMs in combination with a thin, oxygen-plasma-grown oxide layer. SWCNT-FETs with a hybrid SiO2/silane-based SAM gate insulator feature excellent device characteristics, including a low operation voltage of ~1 V and a sub-threshold swing of 60mV per decade. Another advantageous characteristic of these transistors is that the hysteresis is much smaller compared with devices in which the SWCNT is in direct contact with a hydrophilic SiO2 surface. Moreover, recent studies have shown that FETs with organic SAM-based gate dielectric do not only exhibit excellent device performance, but also high operation stability.

This is exemplified in Figure 5 for an FET comprising an individual semiconducting nanotube with a lithographically patterned aluminum gate and a thin oxide/SAM gate insulator.

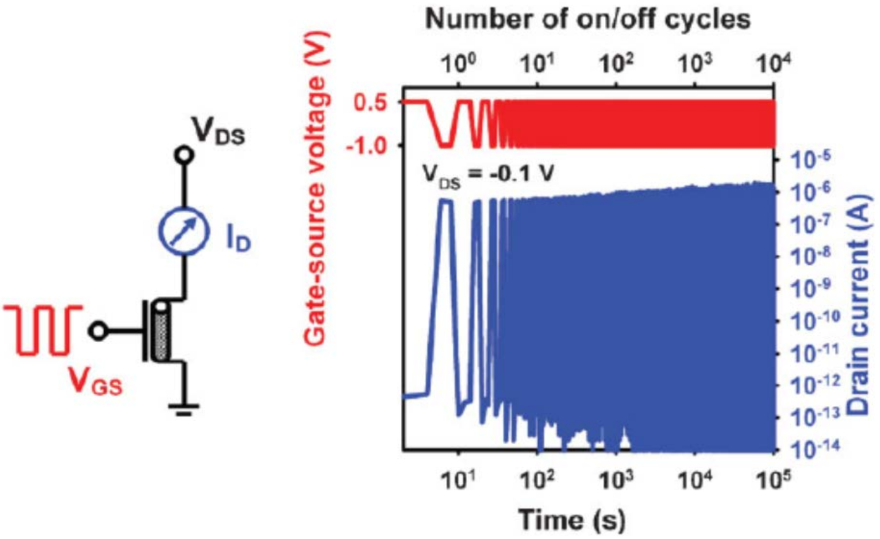


Fig. 5. Gate cycling between the on and off states of an FET with an individual SWCNT on an SAM of an organic phosphonic acid attached as gate insulators onto the underlying aluminum electrode used as back gate. All measurements were performed under ambient conditions using a source-drain bias of -0.1 V.

Such devices endure more than 104 on/off cycles without detectable loss of performance. In addition, they are stable for more than 300 days when stored in ambient air.

A long-standing problem in the fabrication of SWCNT-FETs is the lack of synthetic methods that yield exclusively semiconducting nanotubes, which has stimulated numerous attempts to either separate semiconducting tubes from the as-prepared material or to selectively eliminate the metallic tubes. The separation approach has mainly relied upon noncovalent chemical functionalization by various types of polymers capable of selectively wrapping semiconducting SWCNTs, most prominently DNA and polyfluorenes.

Moreover, self-sorting semiconducting-SWCNTs networks have been obtained by spin-coating nanotubes from solution onto appropriately surface-functionalized Si/SiO$_2$ substrates. The principle of this method is based upon the selective binding of semiconducting tubes by the terminal amino groups of the silane layer on the silica. Efficient chemical methods to eliminate the metallic tubes in nanotube ensembles include the coupling of benzene diazonium salts and methane plasma etching. In the latter process, the metallic nanotubes within an SWCNT film are preferentially modified, while semiconducting tubes with diameters larger than 1.4nm remain largely unaffected.

The modified metallic tubes can then be permanently broken through thermal annealing. A reasonable explanation for the higher chemical reactivity of metallic tubes invokes their more-abundant delocalized electronic states favoring the formation of a transition state with electrophilic species.

Preferential destruction of metallic tubes has also been achieved through selective photo-oxidation using laser irradiation of appropriate wavelength. More recently, it has been demonstrated that also long-arc Xe-lamp irradiation under ambient enables faster destruction of metallic over semiconducting tubes.

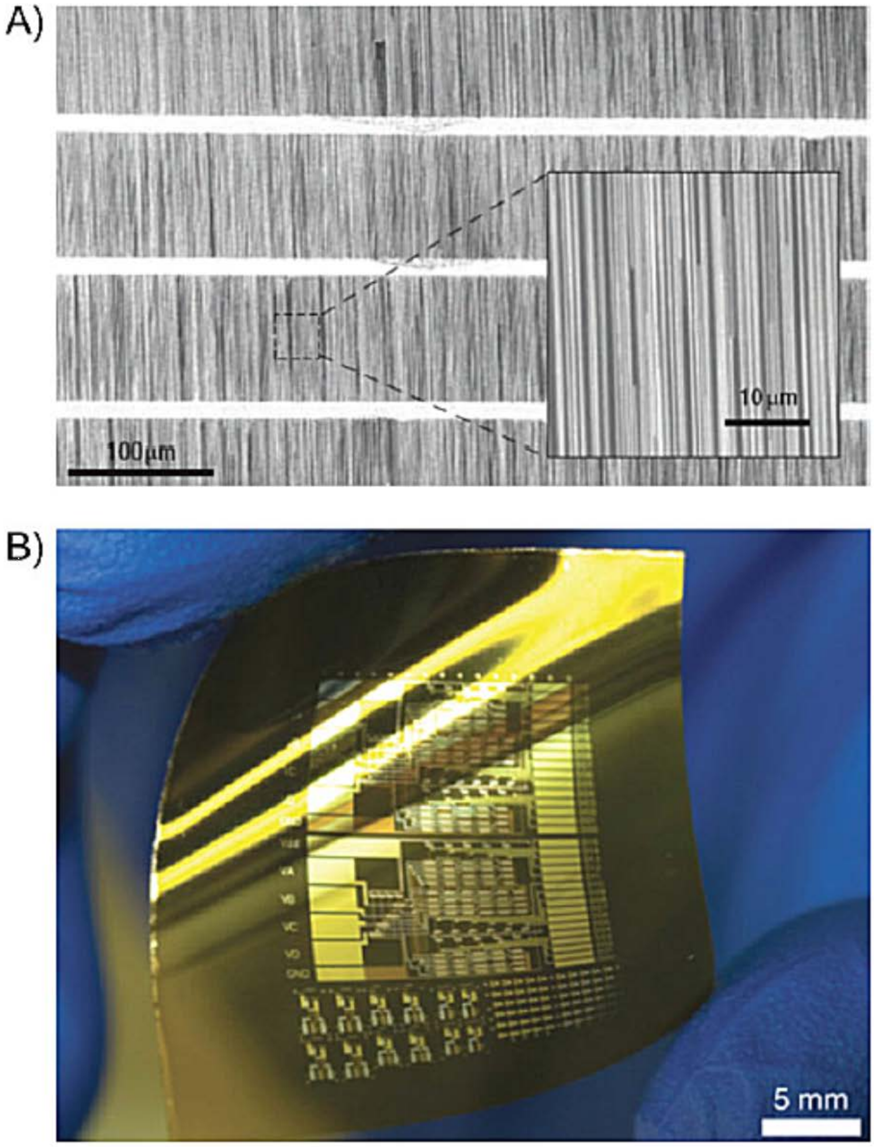


Fig. 6. A) A pattern of laterally aligned SWCNTs grown by CVD on a quartz substrate. The densely packed arrays have been used as FET channels after electrical breaking of the contained metallic tubes. B) Optical microscopy image of an integrated circuited consisting of several dozens of SWCNT-network transistors on a thin polyimide sheet.

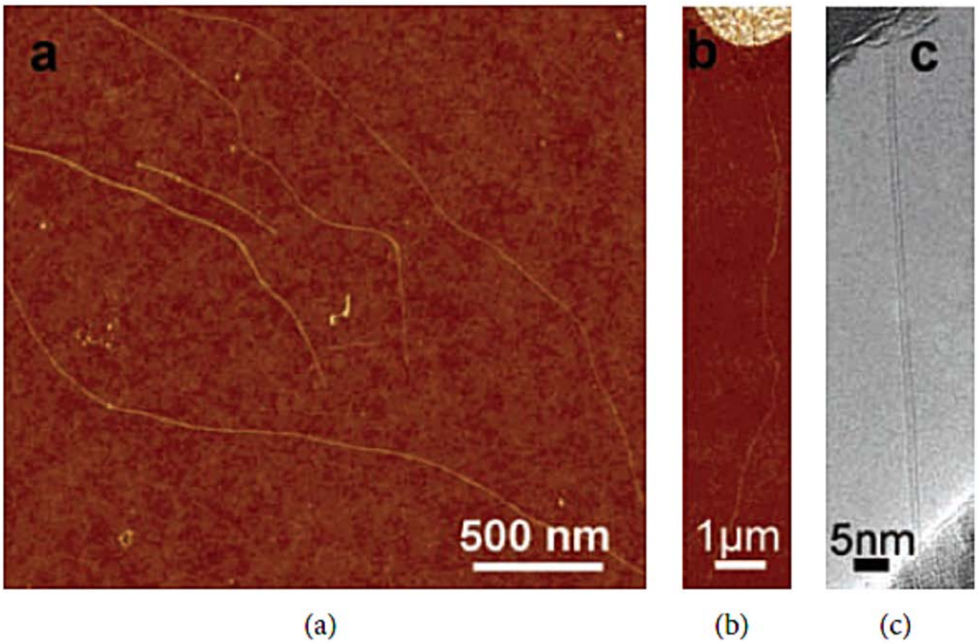
Another promising strategy involves substitution-doping of nanotubes. As a first step in this direction, the synthesis of SWCNTs co-doped with boron and nitrogen has recently been reported. In agreement with theory, which predicts a band gap in the order of 0.5 eV for a 10% content of each boron and nitrogen, BCN-SWCNTs have been found to be purely semiconducting and well-suited as FET channels. FETs made of enriched semiconducting SWCNT ensembles easily reach large on/off ratios exceeding 105, which is sufficient for many applications.

Significant progress has also been achieved in the development of FETs incorporating highly ordered SWCNTarrays produced via oriented CVD growth on quartz substrates, as illustrated in Figure 6A.

After transferring the arrays onto a polymer substrate and selective electrical breakdown of the metallic tubes, the devices display very good performance, including carrier mobilities approaching 1000 cm2 V-1 s-1, scaled transconductances as high as 3000 Sm-1, and current outputs of up to 1 A. By combining p- and n-type FETs, with the latter obtained by chemical doping using a polymer rich in amino groups, logic gates such as CMOS inverters have also been realized.

An extension of this work is the fabrication of integrated circuits comprising up to 100 random-network SWCNT-FETs on flexible plastic substrates (Fig. 6B). Remarkably, even without enrichment of semiconducting tubes, the transistors display excellent performance, as reflected in sub-threshold swings as low as 140mVper decade, mobilities of up to 80 cm2 V-1 s-1, and operation voltages below 5 V. A key factor to achieve this has been to reduce the probability of metallic pathways through use of sufficiently narrow network stripes. Recent SPCM studies have provided electric-potential landscapes of SWCNT networks under different gating conditions, which could aid further optimization of the corresponding FET devices.

Traditional processes to synthesize carbon nanotubes include arc discharge, laser ablation, and chemical vapor deposition (CVD). Among these methods, CVD is the most widely adopted due to its advantages of low cost, controllable synthesis, and high throughput. By controlling the parameters of catalyst, carbon source, gas pressure, and reaction temperature, researchers can control the location, the orientation, and the diameter of CNTs for the specific applications. CNTs were first discovered in arcevaporation of fullerene in 1991. Laser ablationwas used by Guoet al. togrowSWCNT in 1995. Arc discharge was first used to synthesize carbon nanotubes in large scale by Ebbesen and Ajayan in 1992. The laser and arc methods require very high temperature (1200^{o}C or above) to fabricate CNTs, while the location and alignment of the CNTs are difficult to control. Carbon nanotubes for thin film transistors are usually synthesized by thermal CVD or plasma-enhanced chemical vapor deposition (PECVD). Floating catalyst chemical vapor deposition (FCCVD) is also an alternative for CNT growth. Thermal CVD is a traditional approach for CNT growth. Many groups have utilized the method in the experiments. In thermal CVD, CNTs grew on Fe or its salt catalyst at temperature of 900°C or higher. Huczko tried a low-temperature process ranging from 500 to 750°C. The catalyst thickness is a key factor to determine the density and the diameter of the CNTs. Figure 1 shows a picture of the CNTs grown from different catalyst thickness. Thinner catalyst produces CNTs with smaller diameter and higher density. PECVD has been used to synthesize CNTs since 1992. Li et al. obtained high ratio of semiconducting CNTs using this method. Carbon nanotubes were grown on the SiO2/Si wafer at 600°C in Ar, with monodispersed ferritin particles and Fe thin film as catalyst. The CNT diameters are from 0.8 nm to 1.5 nm with a mean of ~1.2 nm. One of the advantages for PECVD is that the products are mainly semiconducting nanotubes,with the ratio as high as 90%.The semiconducting CNT ratio is very critical for fabricating thin film transistors. Figure 7 shows that the as grown CNTs are high-quality single-walled nanotubes.



(a) (b) (c)

Fig. 7: AFM and TEM views of SWCNTs grown by PECVD. (a) AFM image of nanotubes grown from low-density ferritin catalyst. (b)AFM image of a tube grown from a catalyst particle. (c) TEM image of an SWCNT with diameter of 1.2 nm.

Liu et al. proposed to use FCCVD to fabricate CNTs. They have demonstrated that it is possible to use FCCVDto growhigh qualityCNTs even at a lowtemperature. This method is cheap and it suits for mass production of CNTs with different diameters. Though the products in Liu's experiment are multiwalled carbon nanotubes, Moisala et al. demonstrated SWCNTs grown from ferrocene and carbon monoxide (CO) in the temperature ranging 891–928°C using FCCVD.

Metallic nanotubes and semiconducting nanotubes coexist in the synthesized carbon nanotubes all the time. This is a fatal obstacle impeding the development of SWCNT-based electronics because the metallic CNTs lack gate control and degrade the ON/OFF ratio of devices. Separating metallic and semiconducting nanotubes is a very critical technology, which also have attracted extensive research efforts. Different innovative approaches have been proposed to solve this problem, including electrical breakdown, density gradient ultracentrifugation, gel-based separation, dielectrophoresis, and DNA sequence separation. Each of them has its advantages. For example, utilizing electrical breakdown to separate nanotubes does not need any extra-process step during the device fabrication. Using density gradient ultracentrifugation to separate CNTs can achieve 98% purity of semiconducting carbon nanotubes. Here, we review several commonly used CNT purification methods.

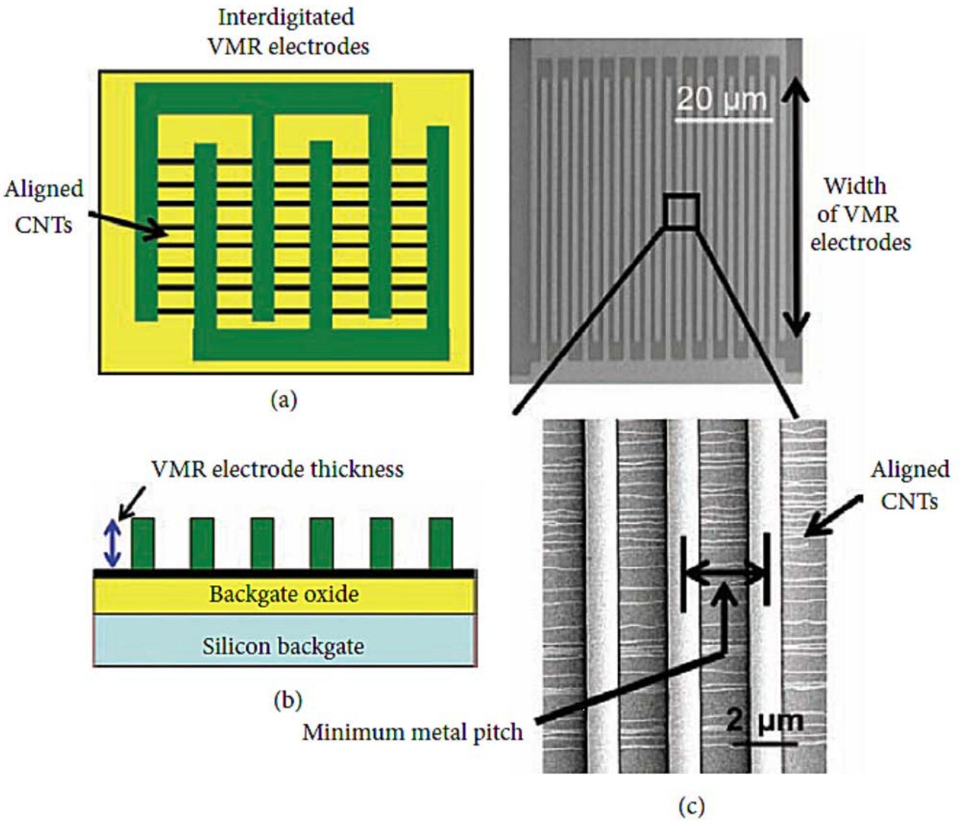
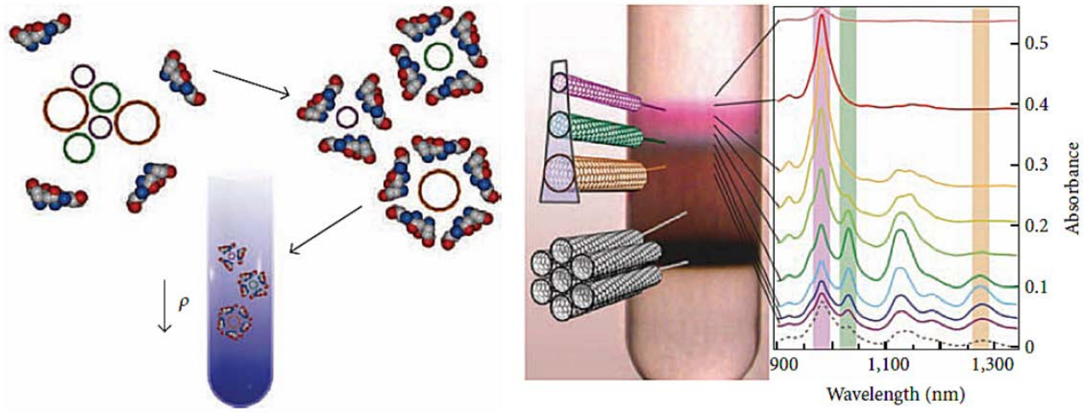


Fig. 8. Viewof VLSI compatibleMetallicCNTRemoval (VMR) structure. (a) Top view. (b) Cross-sectional view. (c) SEMimage (top view). The high voltage is applied to the interdigitated electrodes. The silicon backgate with an appropriate voltage turns off the semiconducting CNTs. Metallic CNTs between digitation will breakdown.

VLSI-Compatible Metallic CNT Removal. Patil et al. proposed a method to separate metallic tubes from semiconducting ones, which is called VLSI compatible Metallic CNT Removal (VMR).

It is a process of using electrical field to break metallic nanotubes in a CNT integrated circuit. In their approach, an oxide layer of 100 nm worked as backgate oxide. Nanotubes were deposited on the oxide and then the interdigital VMR electrodes were fabricated on them to complete the structure, as shown in the Figure 8. Finally, a breakdown voltage was applied between the two electrodes and the backgate turns off the semiconducting CNTs. As a result, the metallic tubes were electrically removed at high electrical current.



(a) (b)

Fig. 9. (a) Schematic of SWCNTs surrounded by surfactant and their sorting. (b) Picture of SWCNTs after separation and the optical absorbance spectra. SWCNTs were sorted evidently.

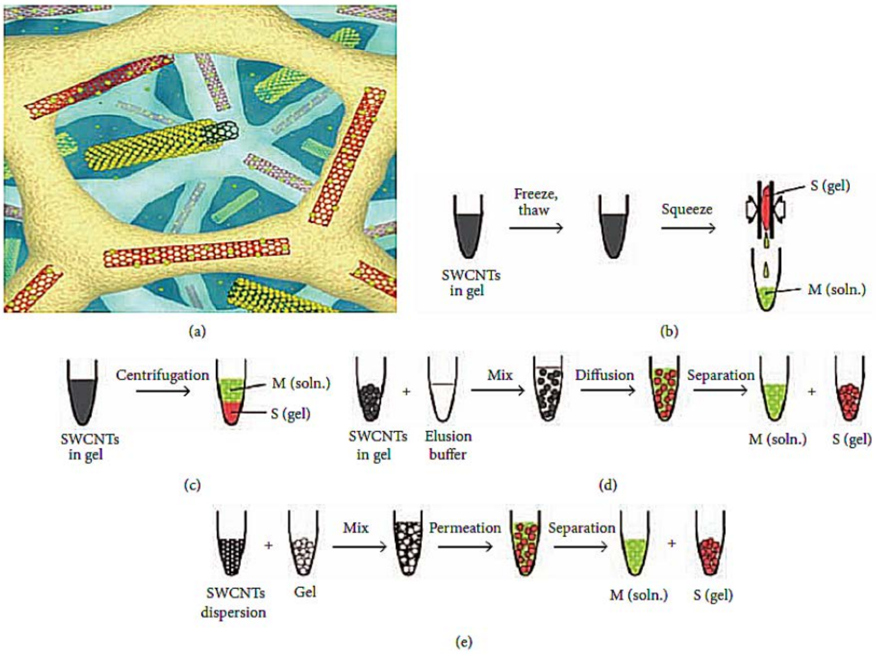


Fig. 10: Schematic of gel separation. (a)Model ofMS separation using agarose gel. Red, semiconducting SWCNTs; beige, agarose gel matrix; green, metallic SWCNTs; yellow, SDS. (b) Freeze and squeeze. (c) Centrifugation. (d) Diffusion. (e) Permeation.

Arnold et al. adopted an approach of density gradient ultracentrifugation, which is widely used in biochemistry, to separate two kinds of SWCNTs. The method is based on the differences in the buoyant densities of different SWCNTs to separate metallic and semiconducting nanotubes. In this method, carbon nanotubes with different diameters, bandgaps, and electronic types (metallic or semiconducting) were surrounded by surfactants. Nanotubes were spatially separated under the effect of centripetal force, as shown in Figure 9.

Tanaka et al. discovered that metallic semiconducting (MS) separation occurred during electrophoresis of an SWCNT-containing agarose gel. SWCNTswere dispersed in sodiumdodecyl sulfate (SDS) solution, and thenagarosewas added to formgel.The reasonwhyusing gel method can sort nanotubes is that the gel adsorbs semiconducting SWCNTs while the metallic ones are left in the solution. Figure 10 shows the steps of the gel separation.

Besides the main previous approaches, still other methods were also reported to sort nanotubes. For example, Liyanage et al. have reported a chirality-based sorting technique that relied on regioregular poly (3-dodecylthiophene) to aid the dispersion of semiconducting single-walled carbon nanotubes.

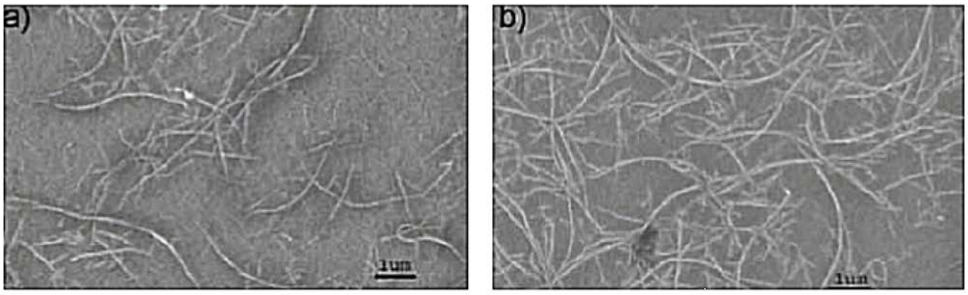
After purification of CNTs, the semiconducting ones can be used to fabricate field effect transistors or thin film transistors. For thin film transistors, CNTs are used in form of CNT film, which is composed of random networks. There are various methods to fabricate carbon nanotube random networks, including filtration, dip coating, transfer printing, ink-jet printing, spray coating or spray deposition, and direct dry transfer.

These methods are mainly by means of transferring CNTs from one substrate to a target substrate. Most target substrates are polymers and therefore cannot sustain high temperature. Each method has its specialty. Using filtration process can form uniform CNT films, and the films allow to be transferred to the other substrates. However, the process is relatively complicated. Dip coating, spray coating, and transfer printing are all low-temperature processes. Okimoto et al. have fabricated devices by ink-jet printing with a temperature as low as 80°C.

Spray coating and dip coating are both simple fabrication processes.However, spray coating cannot be used for large area or mass production. Films fabricatedby spray coating lackuniformity. Dip coating lacks controllability. Let’s consider several main purification methods in detail.

Hu et al. used a vacuum filtration method to produce uniform films of single-walled carbon nanotubes. They dissolved the sonicated CNTs in the chloroform, obtaining a carbon nanotube suspension with a concentration of 0.2mg/L.

Then, the diluted CNT suspension was vacuum filtered over a porous alumina filtration membrane. After the filtration, nanotubes were trapped on the filtration membrane, forming an interconnected network, as shown in Figure 11.



(a) (b)

Fig. 11: SEM images of the CNTs on alumina membrane. (a) and (b) show different density of the network.

Dip coating is a conventional method to get thin film. Figure 12 is the process of dip coating and device fabrication by Xiong et al. They used modified parylene-C as adhesive layer on the silicon substrate and then defined microchannel by optical lithography. After that, the chip was dipped into the aqueous SWCNT solution which was terminated with carboxylic acid groups and pulled up slowly.

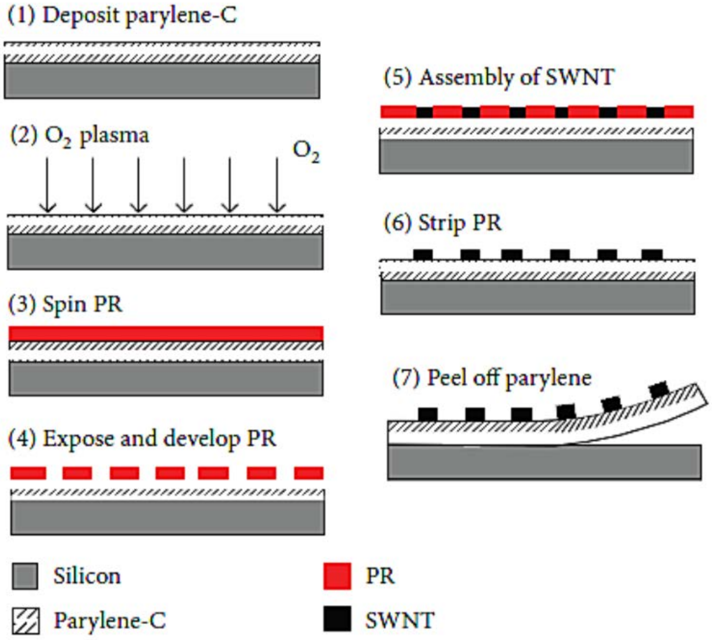


Fig. 12: Schematic drawing of the direct patterning of SWCNTs onto a flexible substrate by Xiong et al. The parylene-C film was disposed by O2 plasma for 30 s to change hydrophobic surface into hydrophilic and defined microchannel by optical lithography. Then dipped the chip into aqueous SWCNT solution and pulled it up slowly with a speed of 0.1 mm/min. Photoresist was removed by acetone after completing dip coating.The flexible parylene-C SWCNT film could be peeled off from the substrate.

The SWCNTs adhering to the surface of the parylene-C formed stripe structure, as shown in Figure 12. It has been demonstrated that SWCNT thin films formed by dip coating show good electrical and optical properties. Ink-jet printing was used to fabricate organic light-emitting diode (OLED) by Hebner et al. in 1998. After that, this technology has been widely used in electronic device fabrication. Ink-jet printing fabricates devices by a printer with objective material in it as ink. SWCNT printing uses SWCNT solution as ink and prints nanotube onto the substrate to form thin film channel. Other parts of the device can be fabricated by the standard fabrication processes like deposition, optical lithography, etch, liftoff, and so forth, or, as an alternative, they can also be fabricated directly by printing. Ink-jet printing results in little material waste and it is friendly to the environment.

Therefore, it has the potential to reduce the ecological impact and energy consumption during manufacturing and also to reduce the costs in production. Besides, it may be used in industrial scale production.

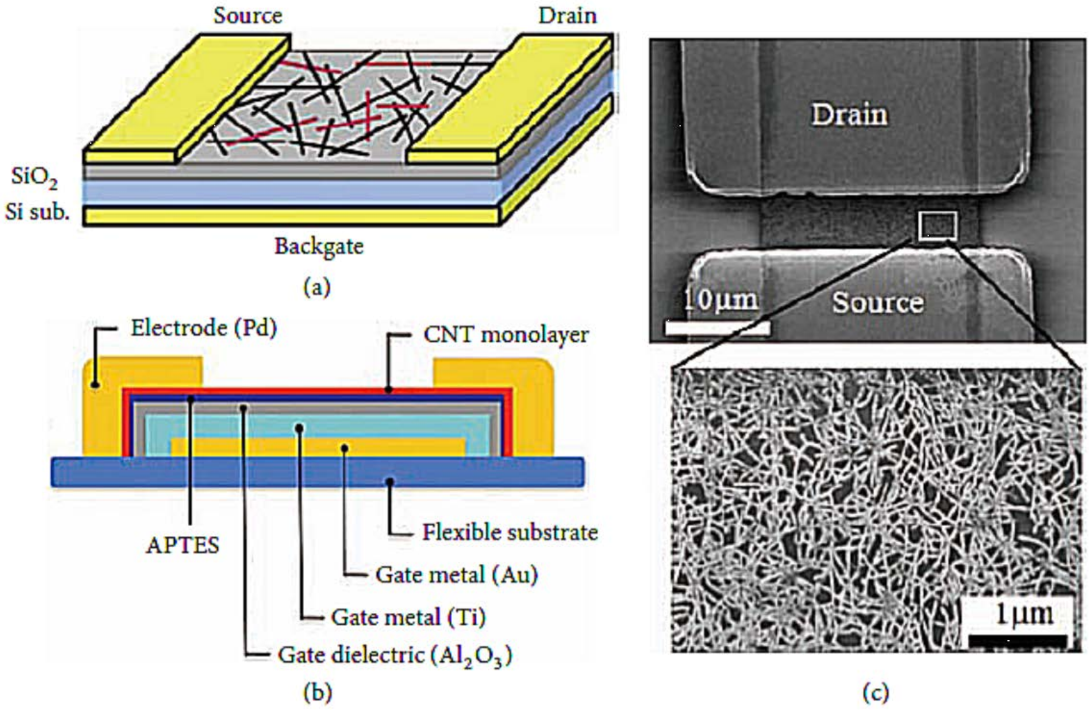


Fig. 13. Typical structure of SWCNT-TFT. (a) Schematic diagram 1. This type of TFT is fabricated on silicon wafer.The silicon substrate works as the backgate and SiO2 works as the gate dielectric. (b) Schematic diagram 2. The other type of TFT is fabricated on flexible substrate like polyamide. (c) SEM image of a TFT fabricated on silicon substrate with SWCNT network as its channel.

Transfer printing has been widely used to transfer CNTs grown by CVD at a high temperature on quartz or silicon substrate. In transfer printing process, the CNTs are adhered to some sacrificial layer, then transferred to an objective substrate, and the sacrificial layer is removed afterwards. Transfer printing technology is quite mature. Using transfer printing can form random CNT network or aligned CNTs.

And another advantage of transfer printing is that it can be applied to nonplanar substrates. Besides the previousmethods,Ma et al. and Zhu andWei have developed technologies to synthesize CNT films in situ using floating catalyst CVD growth.

It is a direct and efficient method, but it requires a high temperature of at least 600°C. Therefore, it is not compatible to the standard flexible thin film transistor fabrication process, which can only sustain a low temperature because the substrates are usually glass or polymer.

There is another form of SWCNT film besides the aforementioned. Liu et al. proposed a novel amorphous indium zinc oxide/ carbon nanotube (IZO/CNT) hybrid film to be used for thin film transistors.

The major difference between the aforementioned SWCNT network film and the hybrid film is that the former utilizes semiconducting CNTs to fabricate devices while the later uses metallic CNTs to provide a superior transporting channel. Experimental results show that IZO/CNT hybrid film possesses better electrical and mechanical performance.

Figure 13 shows the typical structure of a single-walled carbon nanotube thin film transistor. The substrate is silicon wafer or flexible polymer. The channel is a network consisting of random CNTs.

In Figure 8(a), the device has silicon substrate as the backgate and SiO2 as the gate dielectric. In Figure 8(b), the device uses polymer as substrate. The commonly used polymers are Polycarbonate (PC), Polyimide (PI), Polynaphthalene two formic acid glycol ester (PEN), and so forth.

For the polymer substrate device, at the beginning of the fabrication, the flexible substrate was affixed on a silicon wafer to keep the interface flat. Then, titanium (Ti) and Gold (Au) were defined on the substrate as backgate electrodes. A layer of aluminumoxide (Al2O3)with thickness ranging from20 to 40 nm, sometimes with a thin layer of SiO2, was formed over the gate electrodes using atomic layer deposition, serving as the gate dielectrics.

To enhance the adhesion between SWCNTs and the dielectric layer, the wafer was immersed in the solution of 3-aminopropyltriethoxysilane (APTES) in deionized (DI) water for 30 min. After that, SWCNTs were transferred to the interface. The last step was forming source and drain electrodes by deposition of Ti/Palladium(Pd).

We have introduced the ink-jet printing technique to produce CNT films previously. Another innovative process has been proposed to fabricate the whole device with inkjet printing technology. As shown in Figure 14, the device, including source and drain electrodes, is completely fabricated by ink-jet printing.

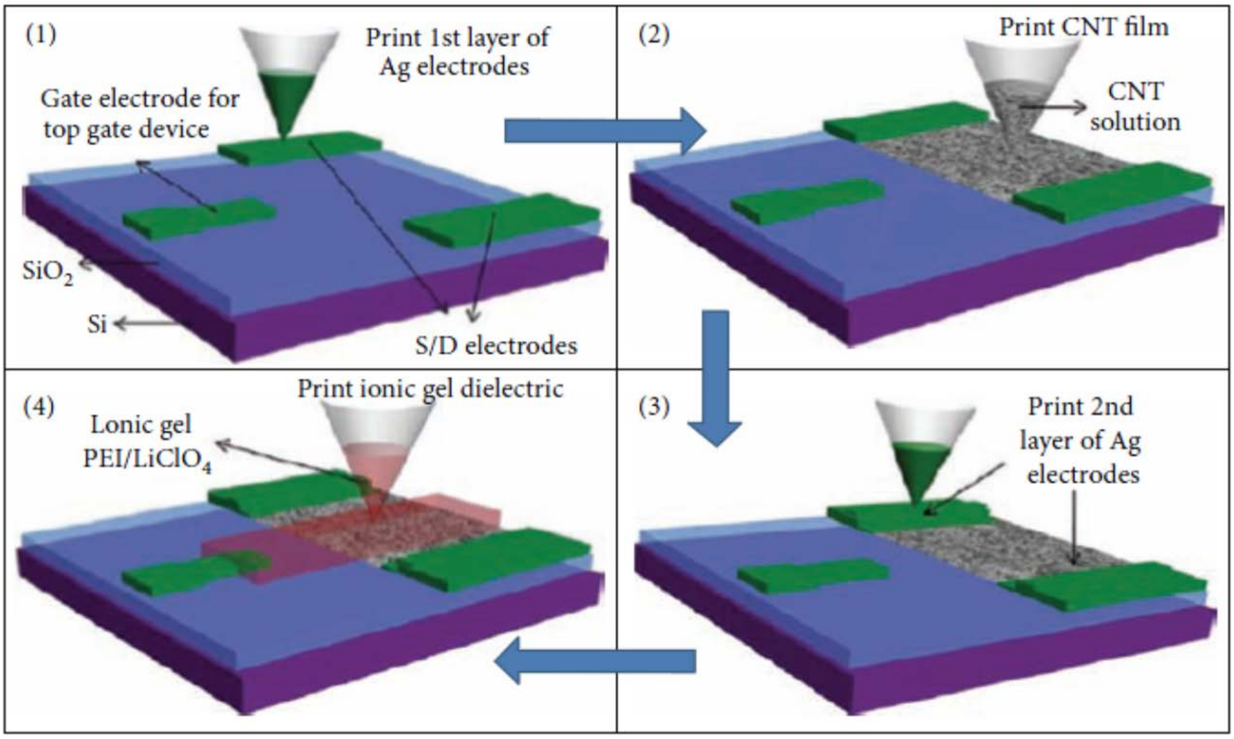


Fig. 14. Schematic for fabrication process of a SWCNT-TFT by ink-jet printing. The whole device was fabricated by printing. First, a layer of silver ink was printed on the substrate to form source and train electrode.The silver layer was sintered at a temperature of 180°C to obtain low resistance. Then SWCNT solution was used as ink to print the channel. After that, another layer of silver ink was printed to improve the contact between SWCNT and silver electrode. Finally, the gate dielectric and gate electrode were fabricated.

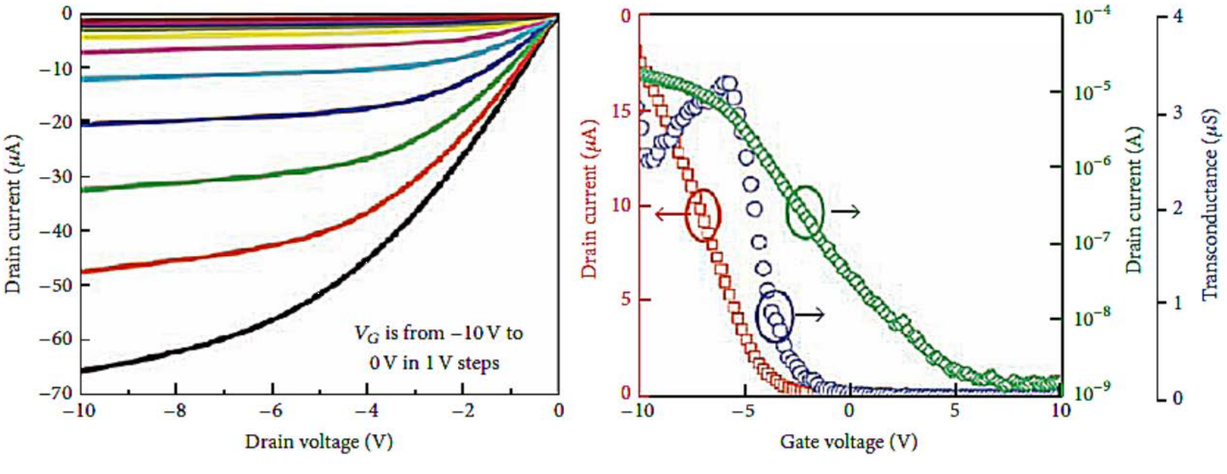


Fig. 15. Drain current and transconductance versus voltage characteristics of SWCNT-TFTs.

The electrical properties of a SWCNT-TFT sample are shown in Figure 15. From the drain current-drain voltage curve, the transistor shows typical p-type characteristics.

There is a peak in the transconductance curve, and it suggests a near linear subthreshold slope.When drain voltage is small, the device shows a linear relationship between drain current and voltage. With the increase of drain voltage, it shows semiconducting characteristics and acts as a p-type transistor. Drain current is influenced by channel length. With the increase of channel length, the on current decreases, as well as the off current. As shown in Figure 16(a). The current density changes as well with the change of channel length. The current on/off ratio of the SWCNT-TFT device is closely related to the CNT purity. The existence of metallic CNTs provides a percolation path. The off current increases drastically and degrades the on/off ratio. The on/off ratio also depends on the channel length, shown in Figure 16(b) by Sun et al. The on/off ratio increases with channel length increase in most experiments, due to the presence of a small amount of metallic nanotubes in the channel.The highest on/off ratio reported is 6•10^6. It is large enough to operate devices such as OLED.

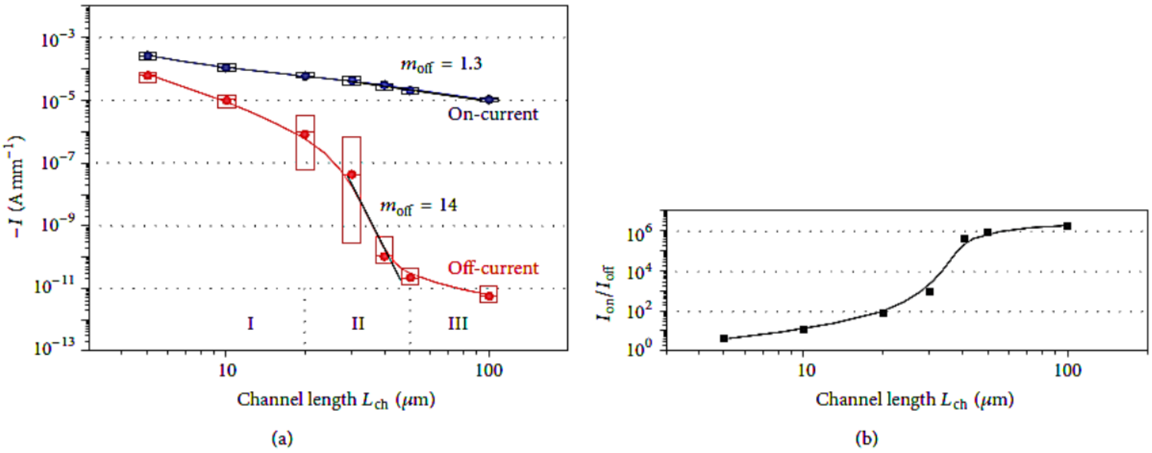


Fig. 16. Plot of current and on /off ratio of SWCNT-TFT. (a) On-current and Off-current versus channel length; (b) *I*on/*I*off ratio versus channel length.

The properties of SWCNT-TFTs reported by different research groups are benchmarked in Table 1.

Table 1

Benchmarking for properties of SWCNT-TFTs reported by different groups.

No.	L_{ch} (μm)	I_{on}/W (μA/μm)	μ (cm^2/V·s)	I_{on}/I_{off}	g_m/W (μS/μm)
1	10	15	50	3×10^3	4
2	4	10	52	$>1 \times 10^4$	0.033
3	100	0.01	35	6×10^6	—
4	50	0.356	67	482	0.046
5	20	—	140	1×10^7	—
6	0.12	125	—	—	40
7	1	0.8	3500	$>1 \times 10^3$	~0.096
8	50	0.6	1300	3×10^4	—

Note: sample No. 5 is metallic SWCNT/IZO-TFT sample, while others are semiconducting SWCNT-TFTs; sample No. 6, 7, and 8 are aligned CNT channels, while others are random CNT network films.

From the comparison, properties of the SWCNT-TFTs by different processes vary a lot from each other. From Table 1, Ion/W of sample No. 6 can reach 125 μA/μm. The on/off ratio can be higher than $6\bullet 10^6$. Mobility (μ) for aligned carbon nanotube thin film transistors (CNT-TFTs) can be as high as 3500 cm^2/V·s. The transconductance (g_m /W) can reach 40 μS/μm. The comparison in this table is relative because their Ion, g_m, and mobility were measured in different working voltage and different conditions. Different channel lengths also affect their measured electrical parameters. In the random CNT network devices, the mobility can be calculated through the equation $\mu = g_m\ L_{\text{ch}}/C_{\text{ox}}V_{\text{DS}}W_{\text{ch}}$ here C_{ox} is the channel capacitance of the random CNT networks. It canbe calculate dusing the method of reference. From Table 1, the highest mobility can be 67 cm^2/V·s. Chandra et al. show that the mobility decreases as the channel length increases. The explanation for the result is that the device performance is dominated by the electron scattering in the channel region rather than at the metal nanotube contacts. For transistors of similar channel length, using longer nanotubes would lead to less nanotube-nanotube junctions and consequently higher mobility. The mobility is also related to the density of CNTs in the network.The mobility increases with increasing the density of CNTs.

For aligned SWCNT-TFTs, the mobility is extremely higher than that of the random CNT network TFTs. It is because the aligned CNTs drastically decrease the tube-to tube junctions. The carriers can transport by ballistic in some degree. The existence of metallic CNTs increases the mobility too.

There is also a relationship between channel length Lch and nanotube length L_{CNT}. When L_{CNT} is comparable to L_{ch}, or when density of the metallic nanotubes exceeds the percolation threshold $\rho_{th} = 4.24^2 / \pi L_{CNT}^2$, the metallic nanotubes can connect the source and drain directly, resulting in a drastic increase in the devicemobility, as well as both the on and the off current.

Thus, there is a trade-off betweenmobility and the on/off ratio. To describe the device performance, g_m is another important parameter. g_m /W is the normalized transconductance, which was reported as high as 4 μS/μm. From Cao's recent work, the value can reach 40 μS/μm. g_m/W is inversely proportional to channel length. It is similar to the device mobility. From the benchmarking in Table 1, the CNT/IZO hybrid thin film transistors has much higher device mobility (140 cm2/V·s) than that of CNT network TFTs, because the active CNTs in its channel are metallic nanotubes, which can provide a high-speed path.The on/off ratio of CNT/IZO hybrid thin film transistors can reach 1.3•107. In CNT/IZO hybrid thin film transistors, both on/off ratio and device mobility depend on CNT weight concentration. In the range of 0 wt% to 1 wt%, the on/off ratio increases as CNT weight concentration increases, as well as the mobility.However, when the CNT weight concentration is larger than 2 wt%, the TFT cannot be switched off.

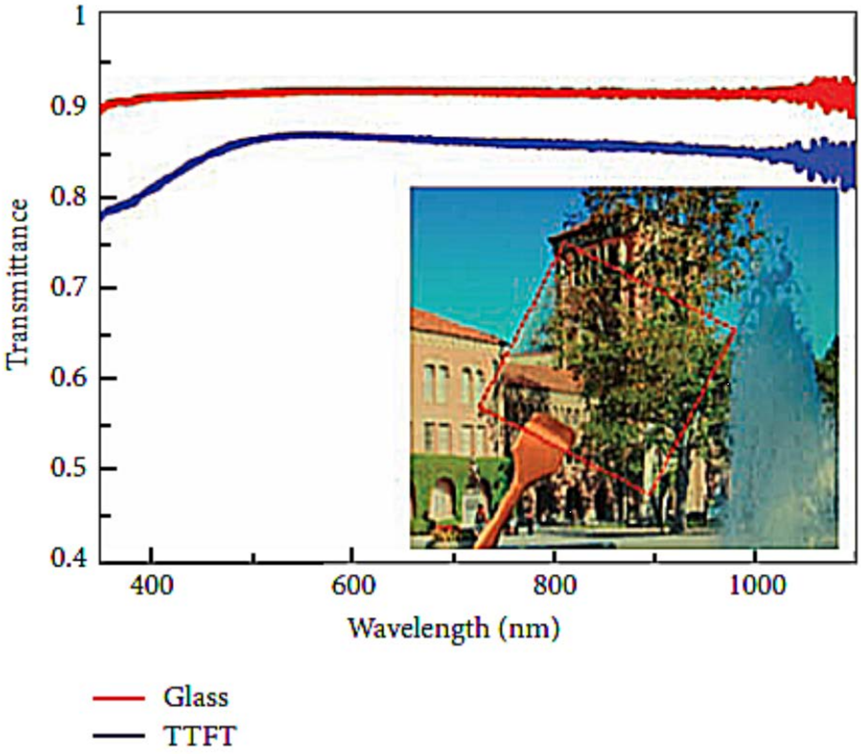
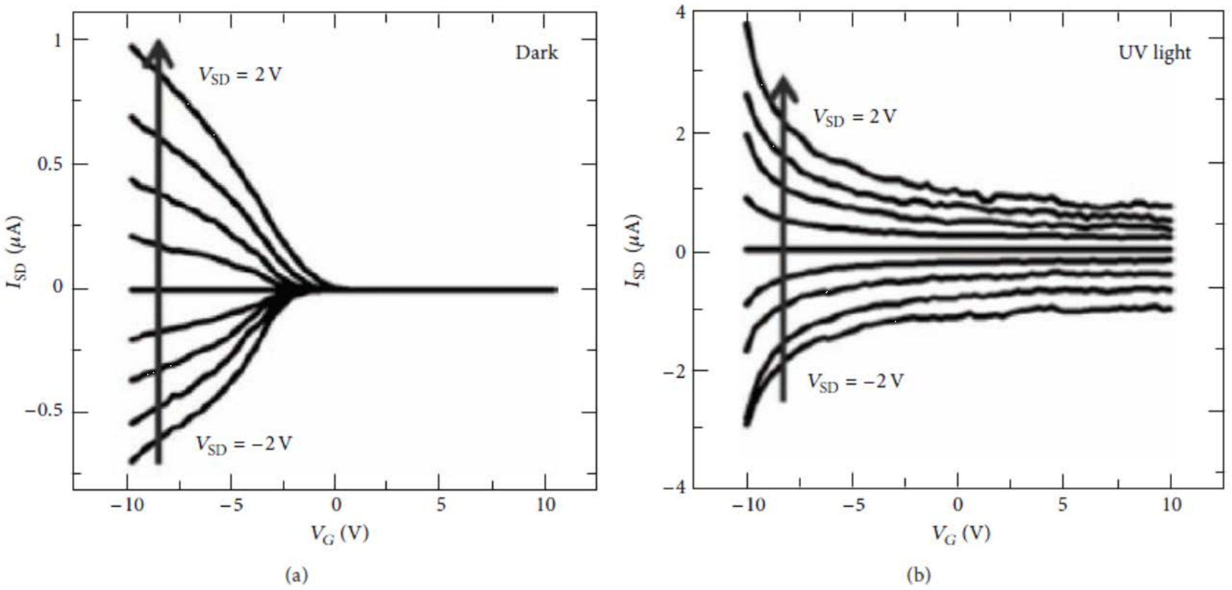


Fig. 17. Transmittance curve of carbon nanotube thin film transistors.The red curve is the transmittance of the glass substrate; the blue curve is the transistors. Inset: the image of the devices on a 2 in. square glass substrate.

SWCNT-TFTs have a good performance on optical transmittance. Wu et al. have obtained 50nm thick CNT thin films by filtration process, with transmittance greater than 70% over visible range and 90% in the near infrared at 2 μm wavelength. Zhang et al. have fabricated carbon nanotube thin film transistor with high transparency of 82% in the visible range. Figure 17 is the transmittance curve of the thin film transistor and the inset is image of the devices on a 2 in. square glass substrate.

Besides high transparency, the SWCNT-TFTs have a high photosensitivity to the exposure of UV/visible light. In Park et al.'s research, they measured the photo response over visible and UV light. Experimental results are shown in Figure 18. When the transistor was exposed to UV light, the drain current increased obviously due to the photo-induced charge carriers.

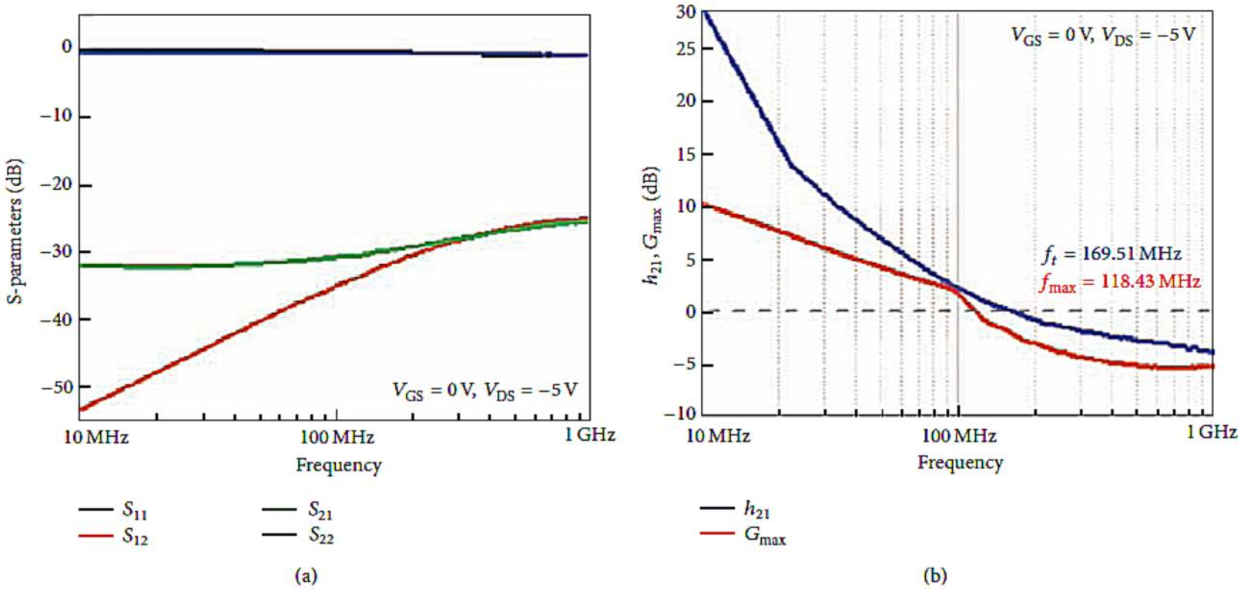
Carbon nanotubes have a good property of mechanical flexibility with extremely high Young's modulus of about 1.2 TPa. Many researchers have done bending test to demonstrate the high mechanical flexibility of carbon nanotube thin film transistors. The bending test shows that the thin film transistors perform well even after bending to a small radius.



(a) (b)

Fig. 18. Current-voltage curves of a transistor with drain voltage from −2 to 2V. (a) without UV light irradiation; (b) with UV light irradiation.

In Chandra et al.'s experiment, the on current remains nearly unchanged and shows reliability. Chae et al. showed that their SWCNT/graphene thin film transistor remained in high performance after stretching and releasing itmore than 1,000 times. For CNT/IZO hybrid thin film transistors, the dynamic loading test shows that after the transistors were bended to a minimum radius of 700 μm, the devices remained in high performance and stability. The repeated bending test shows that after 300 cycles, the normalized resistance varied by less than 8%.



(a) (b)

Fig. 19. Carbon nanotube TFT's RF characteristic. (a)Measured S-parameters from 10MHz to 1GHz.The transistor is biased at VGS = 0V and VDS = −5V for optimal performance. (b) Intrinsic current gain h_{21}, and maximum available gain $G_{\max}$ derived from the S-parameters by Wang et al.

Single-walled carbon nanotubes are promising for a diverse range of applications including high-strength hybrid nanocomposites, growth platforms for neuronal circuits, highly sensitive chemical/biological sensors, electrode material for solar cells, and channel of active elements in electronic devices.

SWCNT-TFTs have been a potential candidate for the future electrical and photonic applications, including integrated circuits, control circuits for liquid crystal display (LCD) and OLED displays, and flexible and transparent display. Because SWCNT-TFTs are sensitive to UV/visible light, they can be used as photo devices as well.

In the work of Wang et al. in the University of California, Berkeley, the SWCNT-TFT showed a cutoff frequency of 170MHz, as shown in Figure 19, which can meet the requirements of certain wireless communication. Zhang et al. measured the RF transmission properties of SWCNT in field effect transistors using S parameter and found the CNT transmission showed no degradation until 12GHz.

The result indicates that there is further room for the cutoff frequency of SWCNT-TFT to improve. Other researchers have used SWCNT-TFTs to fabricate integrated circuits successfully, such as digital logic gates, including inverter, NAND, and NOR gates. Figure 20 shows the inverter circuit and its transfer characteristics.

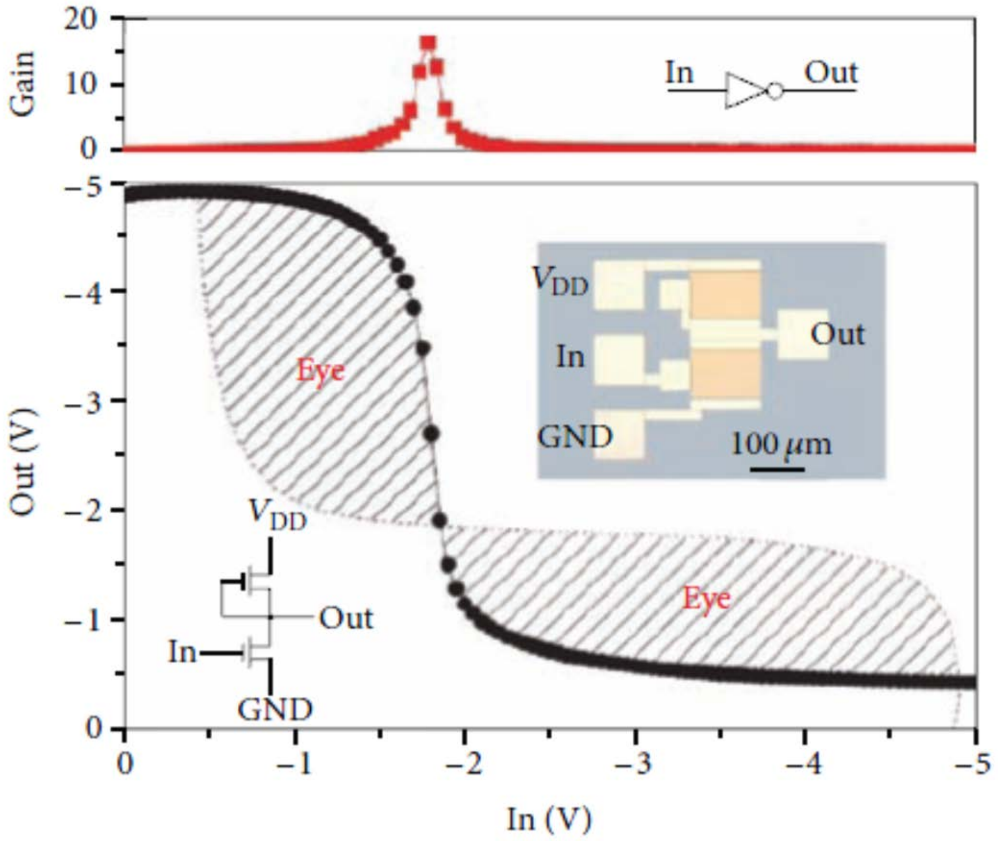


Fig. 20. Inverter realized by SWCNT-TFTs and its transfer curve.

Though lots of studies have been conducted for single-walled carbon nanotube thin film transistors, some issues remain to be solved before their practical applications in the near future. They are summarized as follows.

The resistance of the contacts between tube and tube is high and varies with different contact situations. From Hu et al., the intertube contact resistance is as high as 100MΩ, which is 4 orders of magnitude larger than the resistance of the tubes themselves, 10 kΩ. So, it is important to reduce the tube-to-tube contacts. Using better aligned carbon nanotubes can solve the problem so as to reduce the contact resistance. Some researchers have achieved higher device mobility with aligned nanotube arrays.

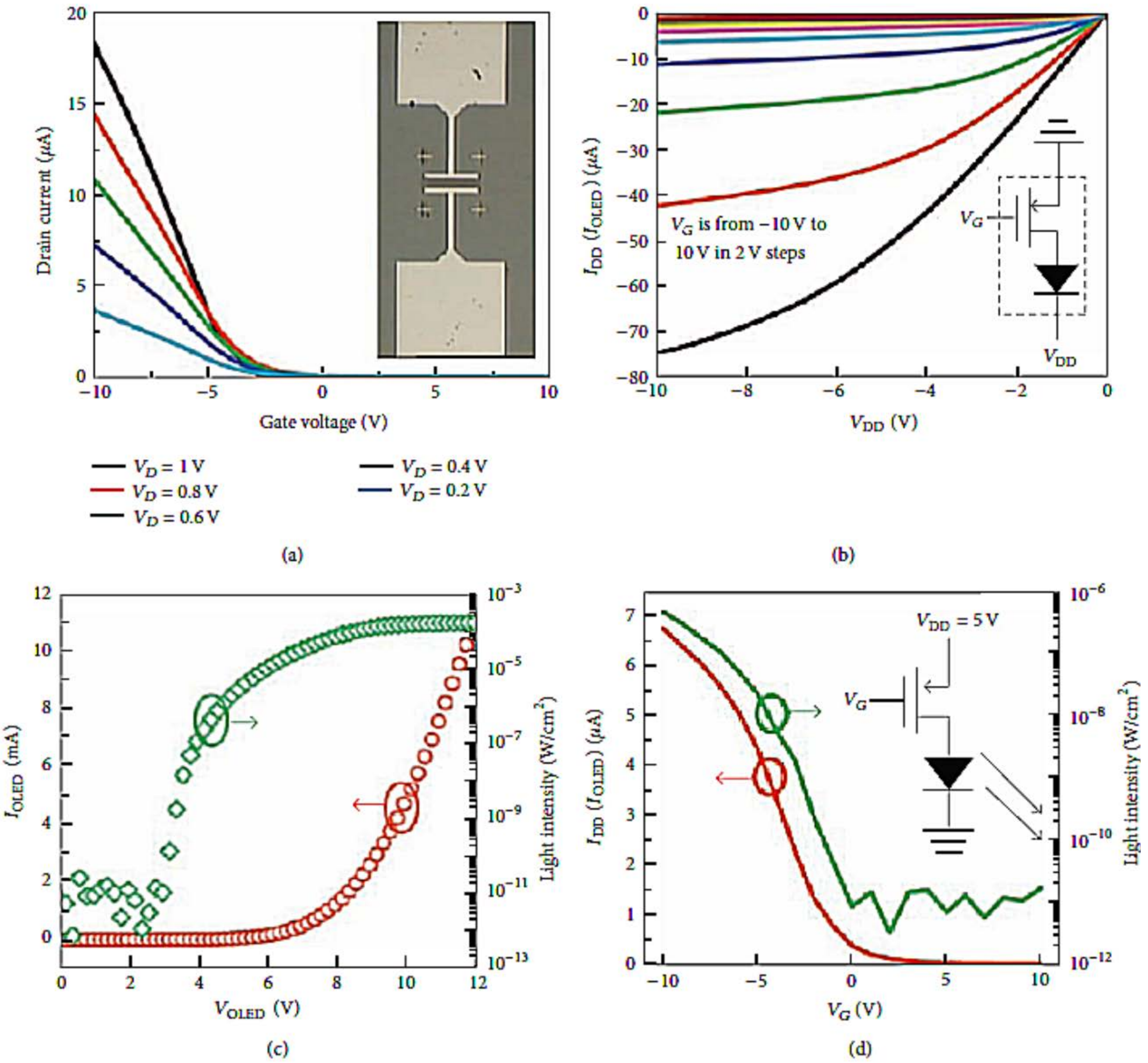
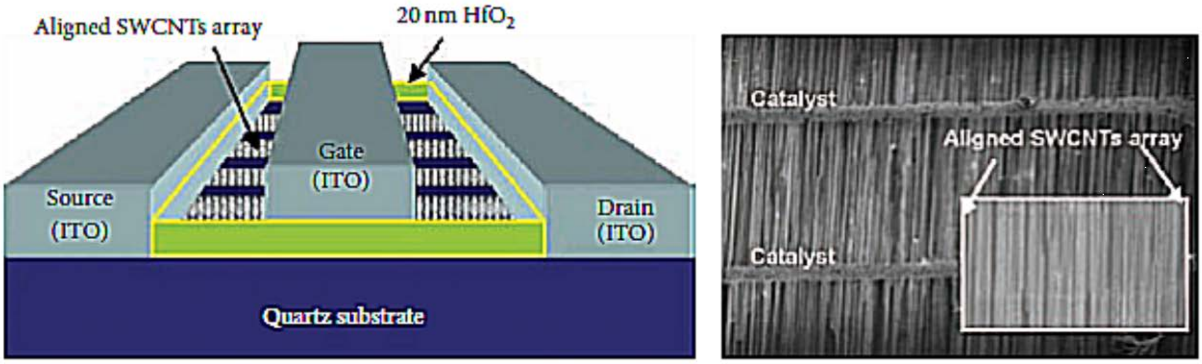


Fig. 21. An OLED control circuit using SWCNT-TFTs (a) Transfer characteristics under different drain voltages for the device, Inset: optical microscope image of the device. (b) Characteristics of the OLED control circuit. (c) The OLED current and light intensity versus the OLED voltage. (d)The OLED current and light intensity versus V_G.

Thin film transistors have been widely used in LCD and OLED displays. Nowadays, the popular TFT channel materials used in mass production are α-Si and poly-silicon. α- Si TFTs have been used in many display products. However, this kind of TFTs cannot meet the requirements of highprecision and high-frame rate for advanced LCD and OLED display circuits in the future. The three-dimensional (3D) displays appearing on the market require frame rates of 240Hz. Even higher frame rates, 480Hz or above, are further required to improve the picture quality. Tomeet these requirements, SWCNT-TFT is a promising candidate with higher carrier mobility and on/off ratio. Recently, a SWCNT-TFT driving OLED circuit has been proposed. In this design, the SWCNT-TFT is connected to an OLED, and VDD (<0V) is applied to the cathode of the OLED, as shown in Figure 21. SWCNT-TFT's on/off ratio can meet the requirement of OLED switching. The output light intensity modulation is over 10^4.



(a) (b)

Fig. 22. View of well-aligned SWCNT-TFT. (a) Schematic of the device. Aligned SWCNTs were synthesized on the quartz substrate at 900◦ C. Source, drain, and gate electrodes were deposited ITO at room temperature with HfO2 gate dielectric. (b) SEM image of aligned SWCNT arrays.

Figure 22 is one of the device schematics. CNTs were in situ grown on the quartz substrate at 900°C. From the picture, CNTs were aligned very well. Although the density is low (about 0.5 tubes/μm), devices fabricated by this method performed well, with effective mobility of $\sim$3500 cm^2/V·s and on/off current ratio higher than 10^3. However, this *in situ* growth process was operated at a temperature of 900°C, which cannot fit for flexible TFT fabrication.

A preferable design has been proposed by Ishikawa et al.. They used a transfer process to transfer aligned CNTs synthesized on quartz to a flexible substrate. They used Au as a sacrifice layer to transfer CNTs. The final CNTs on the flexible substrate reached a density of 2-3 tubes/μm. The mobility is 1300 cm2/V·s and on/off ratio is 3•104. The CNT array density and alignment can be further improved. Ding et al. have obtained CNT density of higher than 50 tubes/μm, and Cao et al. have obtained more than 500tubes/μm by using the Langmuir-Schaefer method in their recent work.

The contacts between nanotubes and metal electrodes are not perfect ohmic contact. The widely used metals to contact CNTs include Ti, Pd, Au, and Pt.

Their work functions are close to those of CNTs. Xu et al. have demonstrated that the contact between Pt and CNT films has the lowest resistance. However, the barrier still exists. Chai et al. inserted graphitic interfacial layer between the metal and the CNT so as to reduce the contact resistance. A new material/composite whose work function matches well with a nanotube is expected to avoid the Schottky barrier and to reduce the resistance.

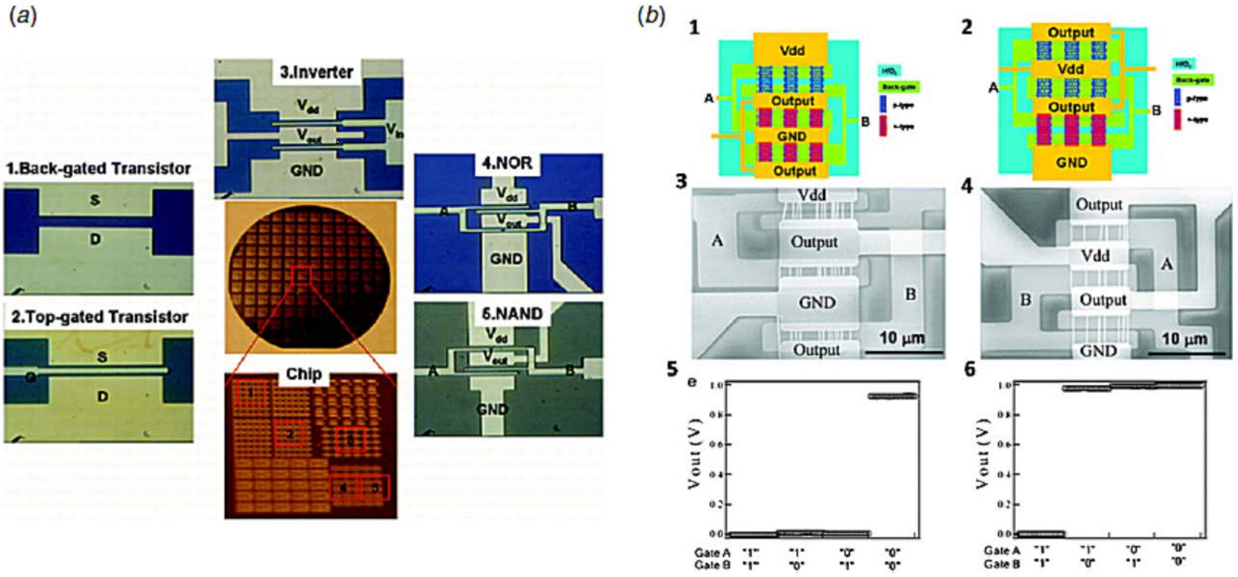


Fig. 23. (*a*) Optical images of CNT transistors and circuits built on a 4 inch Si/SiO2 wafer: 1, back-gated transistor; 2, top-gated transistor; 3, CMOS inverter; 4, NOR logic gate; 5, NAND logic gate. (*b*) Defect-tolerant CMOS NOR and NAND with individual back-gated transistors. (1), (2) Schematic diagrams of CNT-based CMOS NOR and NAND, respectively. (3), (4) SEM images of CNT-based CMOS NOR and NAND, respectively. (5), (6) Output characteristics of CNT-based NOR and NAND.

On the basis of massive aligned nanotube arrays synthesized with CVD method on sapphire, our group reported a high-yield and registration-free nanotube-oninsulator approach to fabricate nanotube transistors, similar to the silicon-on-insulator process adopted by the semiconductor industry. Our group demonstrated truly integrated nanotube circuits and wafer-scale fabrication, overcoming the challenges of small sample size, micrometer-scale channel length, and a lack of controlled doping (Fig. 23). The processing started with the wafer-scale synthesis and transfer of aligned nanotubes arrays up to 4 inches in size. Aligned nanotube array transistors with top-gate were used. Based on the sub-micrometer-scale device platform, controlled electrical-breakdown was used to remove metallic and highleakage semiconducting nanotubes to improve the on/off ratio.

Meanwhile, potassium and electrostatic doping were applied to convert p-type nanotube transistors to n-type. In this way, a truly integrated CMOS logic inverter based on nanotube array transistors was successfully realized. In addition, defecttolerant circuit design was proposed and employed for NAND and NOR gates, as an essential feature for integrated nanotube circuits. Shulaker et al in Stanford also developed a largescale integration (VLSI)-compatible metallic CNT removal technique, so-called VLSI-compatible metallic CNT removal (VMR) technology.

This technology combines design and processing to create CNT transistors and circuits immune to metallic and mispositioned nanotubes. In this work, combinational and sequential CNTFET logic circuits such as half-adder sum generators and D-latches have been realized as the fundamental building blocks of VLSI digital systems. Moreover, this CNTFET-based digital electronics shows great potential in low-power application due to the large improvement in energy-delay product. They further presented a complete sub-system of a sensor interface circuit, which is entirely implemented using CNTFETs.

The CNTFETbased sub-system was demonstrated to interface successfully with a sensor to control a handshaking robot with correct operation. Recently, the first CNT computer has been demonstrated as an important milestone in the practical application of CNT digital electronics. A scanning electron microscopy (SEM) image of the fabricated CNT computer is presented in Figure 24(*a*).

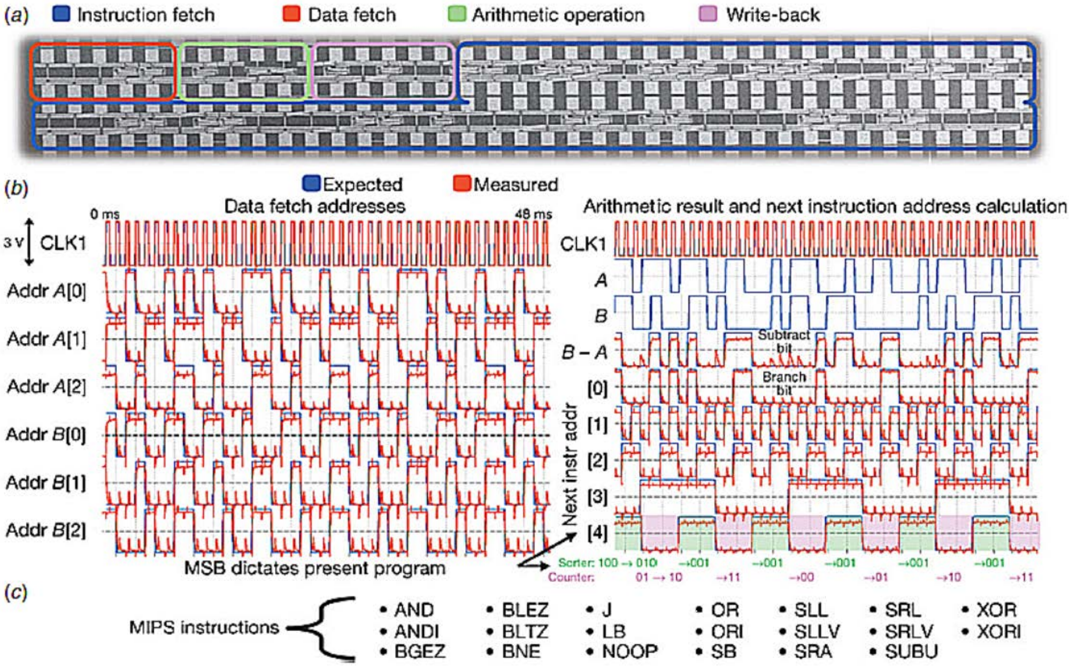


Fig. 24. CNT computer (*a*) SEM of the CNT computer. (*b*) As-measured and expected output waveforms for the CNT computer. In this program, the CNT computer is switching between performing counting and sorting. (*c*) A list of MIPS instructions tested on the CNT computer.

It is demonstrated that the CNT computer runs an operating system and achieves multitasking of counting and integer-sorting simultaneously. Even though the operation frequency of this CNTFET-based is only 1 kHz due to the capacitive loading introduced by the measurement setup and academic experimental limitations, the experimental demonstration in this work is a considerable advancement in the development of complex and highly energy-efficient CNTbased electronic system.

With all the achievement discussed above, CNTFETbased digital electronics, as an exciting complement to existing semiconducting technologies, is a good candidate that can outperform silicon.

In order to realize CNTbased CMOS integrated circuits, both n-type and p-type CNTFETs are required. Recently, through applying metal contact engineering to the existing nanotube platform, both n-type and p-type CNTFETs have been demonstrated.

Based on metals with a work function larger than the valence band edge of CNTs, such as palladium (Pd), people have realized perfect p-type CNTFETs with a barrier-free contact for hole carrier transport. In order to achieve nanotube transistors with n-type behavior, metals with small work function need to be used. This allows the Fermi level of the electrodes to align with the conduction band edge of CNTs, providing an Ohmic contact for electron carrier transport. Chemical and electrostatic doping methods have been applied to achieve n-type CNT transistors by providing excess positive charges in the vicinity of contacts to facilitate electron tunneling. Recently, different kinds of technologies have also been developed to form n-type Ohmic contact in CNTFETs, ideally with no parasitic resistance.

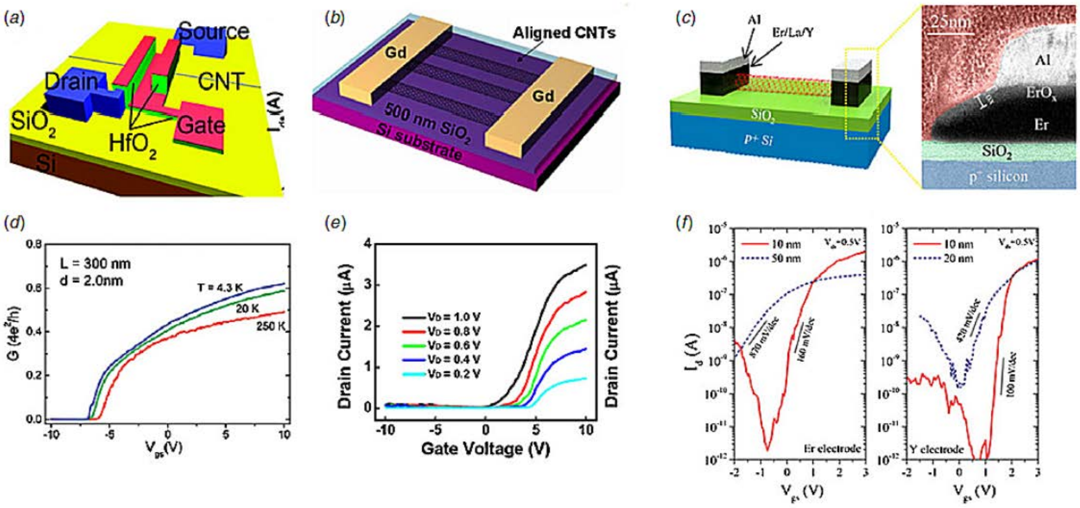


Fig. 25. (*a*)–(*c*) Schematic of N-type transistor from reference, respectively. (*d*)–(*f*) Representative transfer characteristics of N-type transistor corresponding to transistor structure (*a*)–(*c*).

Ding et al applied Y to the contact of CNTFETs and discussed that Y makes a perfect Ohmic contact with the conduction band of theCNT. As illustrated in Figure 25(*a*), a self-aligned top-gate device has been fabricated with Y metal contact. This device showed high performance with a room temperature conductance approaching the theoretical quantum conductance limit of CNT-based devices (4*e2/h*).

Figure 25(*d*) presents the transfer characteristics of an as-fabricated top-gate CNTFET with gate length of ~0.8 μm and a gate dielectric HfO2 with thickness of 15 nm. Based on the analysis, the n-type CNTFET reported in this work shows a subthreshold swing (SS) of 73 mV/decade at room temperature, approaching the theoretical limit (~60 mV/decade).

Meanwhile, our group also introduced the small work function metal gadolinium (Gd) for n-type contact and demonstrated air-stable n-type aligned nanotube transistors. Gd with awork function of ~3.1 eVwas used as themetal

contact in source and drain electrodes, as shown in Figure 25*b*). The transfer characteristics of the transistor after electrical breakdown are presented in figure 25(*e*) with different drain bias voltages. The linear output characteristics reported in our work also indicate that n-type ohmic contacts can be formed between the Gd electrodes and CNTs.

Recently, Shahrjerdi et al performed a thorough experimental study of n-type contacts for CNTFETs based on several kinds of low work function metal contacts, including erbium (Er, Φ_M = 3.0 eV), lanthanum (La, Φ_M = 3.5 eV), and yttrium (Y, Φ_M = 3.1 eV).

Figure 25(*c*) shows a representative transmission electron microscopy (TEM) image of Er contacting CNT transistor, where the Er contact was capped in situ with aluminum (Al). The corresponding n-type transfer characteristics of the resulting devices with Er and Y contacts are plotted in Figure 25(*f*), showing that the performance of n-type CNTFETs degraded with contact layer thickness increased, due to a higher degree of themetal contact oxidation.

It has been discussed that high oxidation rates and sensitivity to deposition conditions of low work function metal contacts results in lower yield and large variation in performance of n-type CNTFETs. To avoid this problem, proper passivation of Er electrodes with a monolayer of hydrophobic polymerwas applied to improve device durability in this work. To further advance CNT-based CMOS integrated circuits, a lot of efforts have been made in improving the scalability and stability of n-type CNTFETs, and those results suggest practical and promising approaches for nanotube-based CMOS integrated circuit applications.

In order to become a viable alternative to silicon technology, nanotubebased electronics require scalable assembling of high-density aligned CNT arrays. Dense nanotube arrays are essential to optimize CNT-based electronics through maximizing device packing density and providing sufficient drive current. Previously, our group reported the combined use of lowpressure CVD and stacked multiple transfer to achieve aligned nanotubes with high density up to 55 tubes μm^{-1}.

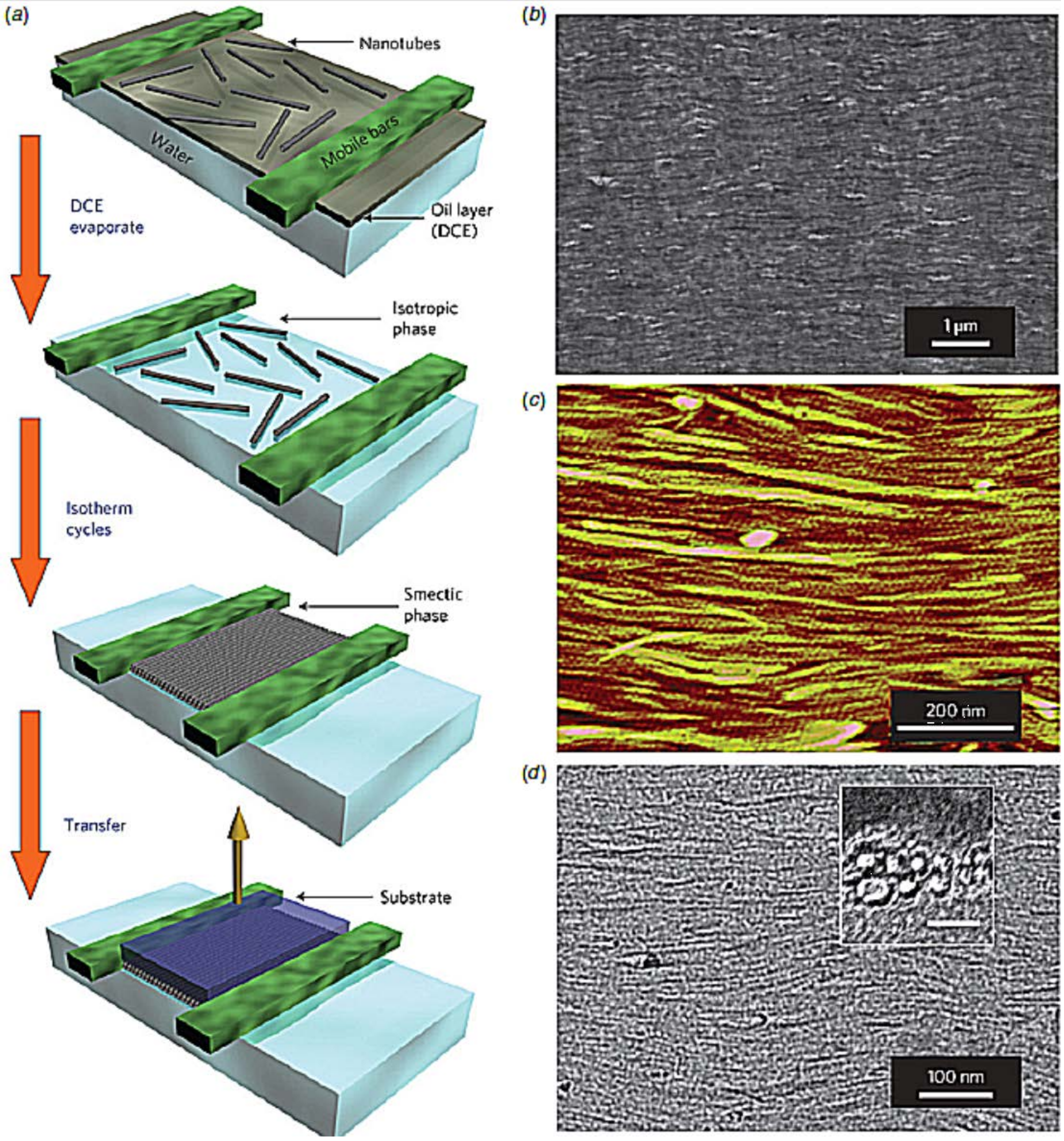


Fig. 26. Langmuir–Schaefer assembly of full-coverage aligned semiconducting CNT arrays. (*a*) Schematic of the Langmuir–Schaefer assembly process flow. (*b*)–(*d*) SEM(*b*), AFM (*c*), TEM (*d*) images of aligned nanotube arrays on substrates.

Recently, Cao *et al* presented the Langmuir–Schaefer method to assemble aligned arrays of semiconducting CNT (Figure 26). Starting with pre-processed nanotube solution with a semiconducting purity of 99%, they reported that nanotube arrays assembled using this method can fully cover a surface with a density of more than 500 tubes μm−1.

Due to the high density, semiconducting purity, and quality of alignment, the CNTFETs fabricated using this approach present improved electrical properties with a drive current density of more than 120 mA mm–1, transconductance greater than 40 mS mm^{-1}, and on/off ratios of ~103. These CNT-assembling results show the great potential for using CNT arrays in scalable high-performance beyond-silicon electronics through CMOS-compatible circuitand system-level implementation. Meanwhile, the scale assembling of CNT also motivates other emerging applications, including thin-film electronics, transparent electronics, and stretchable electronics.

The unique characteristics of SWCNT such as high mobility, small dimension, low capacitance, and large transconductance generate great interest in CNT-based analogue electronics, which is of great importance. Analogue electronics only requires high transconductance but not high on/off ratio for RF transistors. Recently, RF transistors and circuits that incorporated densely aligned arrays of SWCNTs enabled comprehensive experimental and theoretical evaluation of their intrinsic properties, toward practical application of SWCNTs.

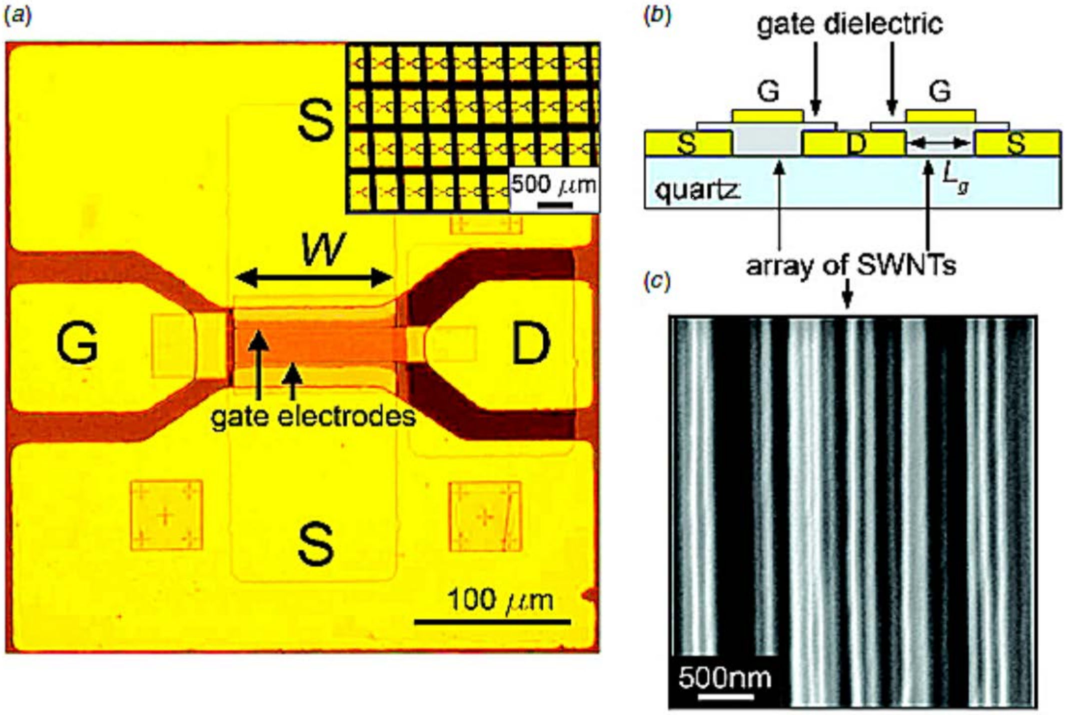


Fig. 27. (*a*) Optical images of nanotube array RF transistors on a quartz substrate. (*b*) Schematic of aligned nanotube RF transistor. (*c*) SEM of aligned nanotube arrays.

The typical layouts used in CNT RF transistor are presented in Figure 27(*a*) consisted of a double channel configuration in which two gate electrodes and two source electrodes surround a common drain electrode. This layout design is fully compatible with conventional small signal models of RF response.

Previously, device structure in Figure 27(*b*) was commonly used, where horizontally aligned arrays of SWCNTs occupy the channel region of RF transistor and provide an electrically continuous and independent pathway for charge transport. On the basis of this device design, Kocabas *et al* reported submicrometer channel length RF transistors that involve perfectly aligned nanotube array with densities of 2 or 5 SWCNTs μm−1. The transistors reported in this work show unity current gain (f t) and unity power gain frequencies (f max) as high as ~5 and ~9 GHz, respectively. The small signal models of the devices provide the essential intrinsic parameters: intrinsic f t of ~30 GHz for a gate length of 700 nm. The results provide clear insights into the challenges and opportunities of CNT arrays for applications in RF electronics.

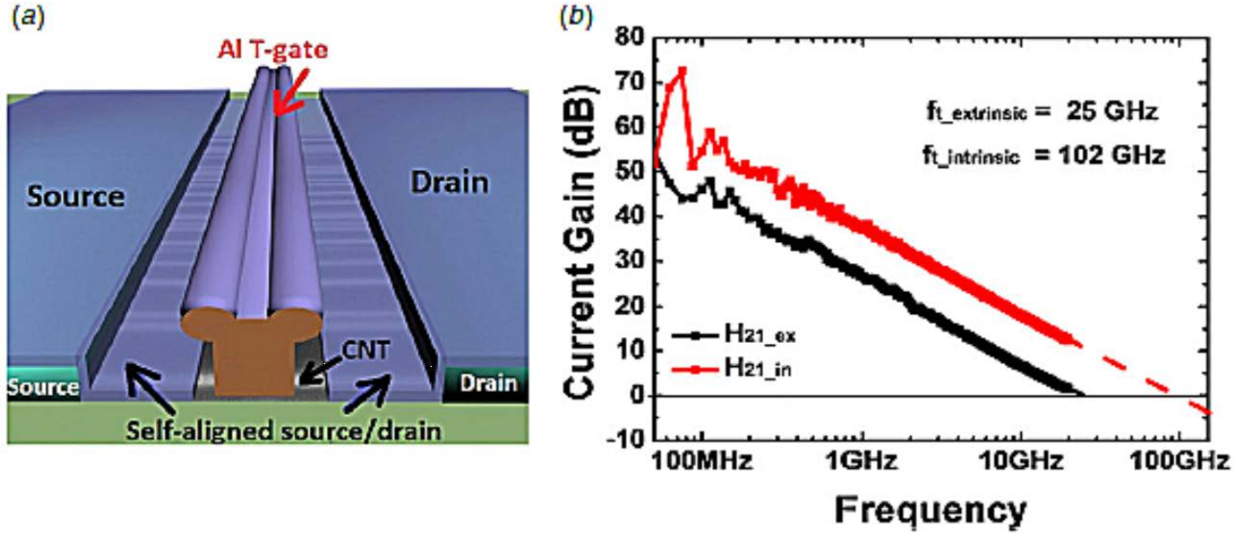


Fig. 28. (*a*) Schematic of self-aligned T-gate aligned nanotube RF transistor. (*b*) Extrinsic and intrinsic frequency response of self-aligned T-gate aligned nanotube RF transistors.

Recently, a self-aligned T-shaped gate fabrication approach developed by our group has been introduced for high-performance SWCNT RF transistors, representing an important step toward RF applications (Figure 28(*a*)).

With this self-aligned design, the parasitic effects of fringe gate capacitance, access resistance, and gate charging resistance can be significantly reduced. Furthermore, the channel length can be scaled down to 100 nm and the Al2O3 gate dielectric was reduced to 2–3 nm, contributing to the quasi-ballistic and quasi-quantum capacitance operation for nanotube transistors.

An extrinsic f_t up to 25 GHz before any de-embedding procedure was achieved for RF transistors consisted of aligned nanotube arrays synthesized by CVD method (density ~5 tubes μm−1), as the highest extrinsic cut-off frequency reported to date for nanotube RF transistors (Figure 28(*b*)).

Meanwhile, Steiner et al reported a planar device platform with embedded electrodes combined with dense aligned high-purity semiconducting CNT assembled through dielectrophoresis (Figures 29(*a*), (*b*)). In this work, as measured, extrinsic *f* t and *f* max, respectively, of 7 and 15 GHz were obtained for a RF transistor of 100 nm channel length (Figure 29(*c*)). After de-embedding, intrinsic *f* t and *f* max of 153 and 30 GHz were observed (Figure 29(*d*)). These high-performance nanotube RF transistors pave the path toward practical RF circuits operated in gigahertz regime.

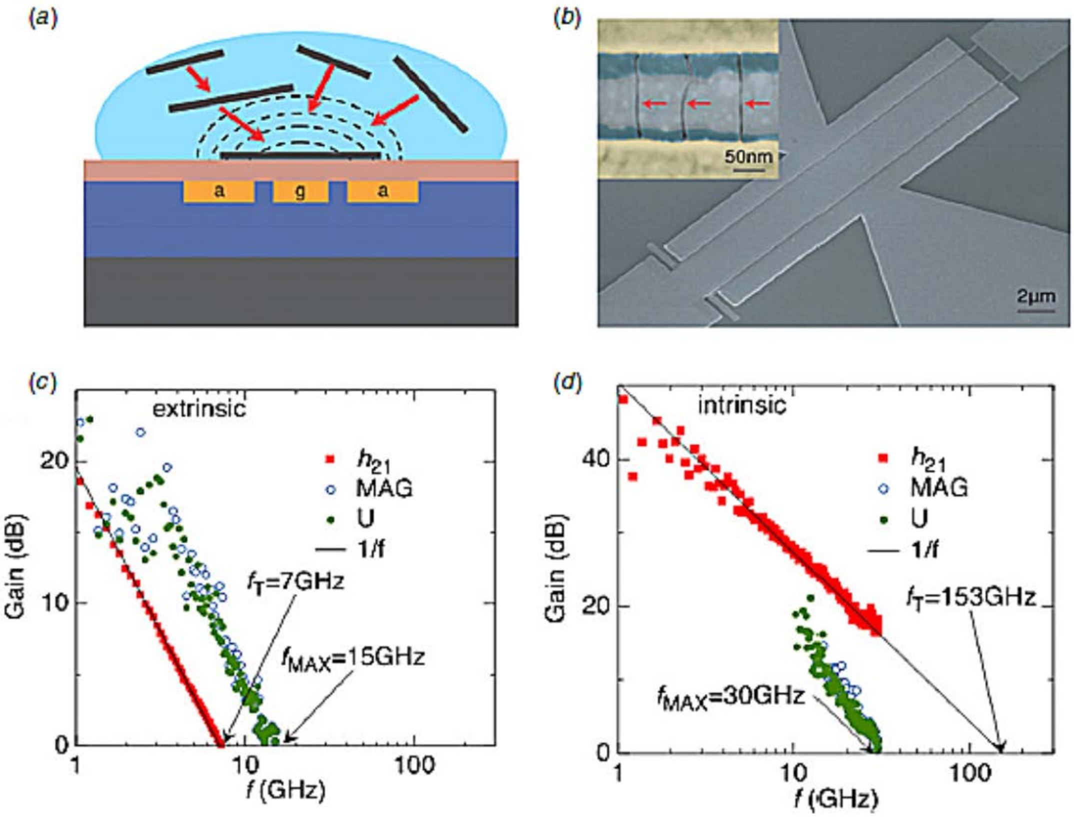


Fig. 29. (*a*) Schematic of a planar CNT RF transistor with embedded gate structure. (*b*) SEM image of the CNT RF transistor. (*c*) The extrinsic current gain and extrinsic maximum available power gain. (*d*) The intrinsic current gain and intrinsic maximum available power gain.

In addition to the study of CVD-based CNT random network electronics, there has also been extensive research in the area of solution-based CNT electronics. The as-grown CNT contain approximately 1⁄3 metallic tubes and 2⁄3 semiconducting nanotubes. This presents a negative effect in CNT-based transistors because the metallic nanotubes in the CNT filmmay contribute to direct conducting paths between the drain and source electrodes in a FET, and this can result in low *I*on/off for the transistors. Due to the advancement in enrichment of semiconducting CNT, CNT solution with higher percentage of semiconducting tubes can be sorted with techniques such as density-gradient ultracentrifugation.

Many methods of separating metallic/semiconducting or even single chirality of nanotubes have been achieved. Particularly, DNA-wrapped nanotubes can be separated into nearly single-chirality nanotubes by ionic exchange chromatography. Density-gradient ultracentrifugation can realize the separation of nanotubes dispersed by widelyused cheap surfactants.

Furthermore, the simple and scalable gel chromatography can also be applied to achieve the separation of the surfactant-dispersed nanotubes without ultracentrifugation. Selectively dissolving nanotubes by designed polymers has also achieved great success, but the polymers are not readily available. Recently, the aqueous two phase separation, which could tune the hydrophobicity and hydrophilicity of the two phases and cause the partitioning of the tubes, is a promising, scalable, and fast separation method.

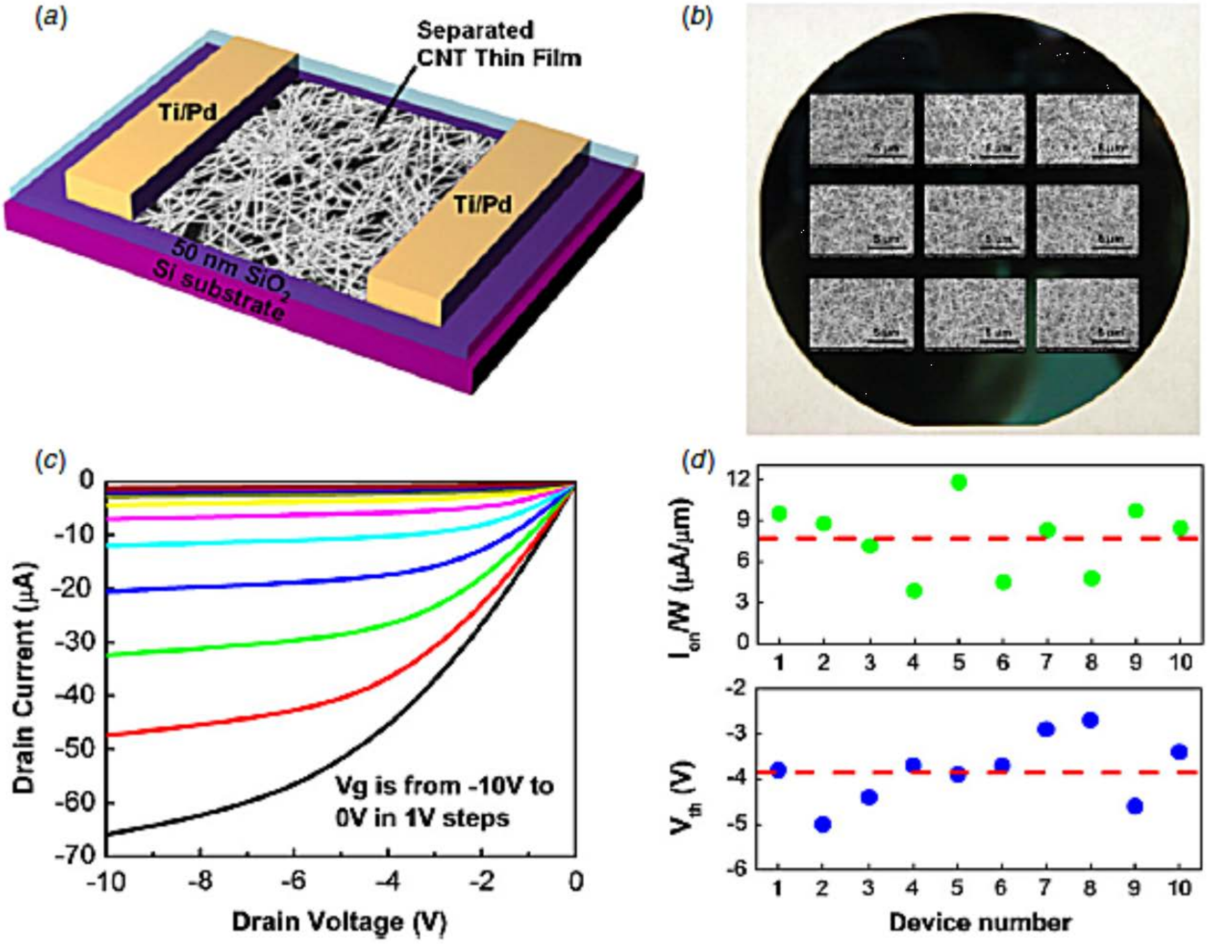


Fig. 30. (*a*) Schematic diagram of a back-gated transistor based on separated semiconducting enriched CNT random network. (*b*) SEM images of CNT random network captured at different region of a wafer after CNT deposition. (*c*) Output characteristic of a CNT network-based transistor. (*d*) Statistical study of current density (*I*on/W) and threshold voltage (*V*th) measured ten CNT-based devices.

Many research groups including our own have demonstrated high performanceCNTTFTs based on the sorted semiconducting CNT solution. The CNT TFTs have been employed in the applications of drivers for AMOLEDs, pressure sensors and integrated digital circuits. CNT random network thin film exhibits excellent transparency, which makes it a desirable channel material for the driver circuitry in transparent display applications. In comparison with amorphous silicon FETs and organic transistors, CNT TFTs have been demonstrated to exhibit superior performance in terms of mobility.

CNT TFTs with mobility greater than 100 $cm^2\ V^{-1}\ s^{-1}$ have been reported, which exceed the mobility of amorphous and organic FETs by at least two orders of magnitude. And comparing CNT TFT with polycrystalline silicon FETs, they can be processed at room temperature, which is essential for the fabrication of flexible electronics on polymeric substrates. In addition to the aforementioned merits of CNT TFTs, CNT solution can be printed with various techniques such as inkjet printing, screen printing or gravure printing etc. Printing for electronics is an inexpensive and scalable technique that can be expended to fabricate circuits in a large area of flexible substrates at low cost. This advantage is unparalleled to the existing silicon process.

We are among the first fewresearch groups to demonstrate CNT TFTs comprised of semiconducting enriched CNT random network. Figure 30(*a*) illustrates a common back-gated TFT based on a random network of separated CNT solution. In the study, CNT solution consisted of 95% and 98% enriched semiconducting tubes were used in the channel of the FETs. The CNT solution was uniformly deposited onto the device substrate by first functionalizing with aminopropyltriethoxy to terminate the surface of the substrate with a layer of amino-groups, which enhances the attraction between the CNTs and the surface.

As can be observed in Figure 30(*b*), the CNT thin film can be deposited invariantly over the entire surface of a 3 inch Si/SiO2 wafer, as shown by the SEM images captured at different regions of the wafer after CNT deposition. This proves that the deposition process is scalable, and can be used in industrial-scaled fabrication.

The geometry of the channel in the transistors can be defined by standard photolithography, followed by O_2 plasma etching of the CNT film in the region outside the channel. This technique eliminates the issue of assembly of the nanotubes.

The transistors fabricated with the 98% enriched CNT solution exhibited ideal p-type behavior with mobility as high as 52 $cm^2\ V^{-1}\ s^{-1}$, while maintaining a current on/off ratio of 104. Although the mobility of the devices is not as high as that exhibited by silicon transistors, it is still an invaluable thin film material for applications such as drivers for display, or flexible electronics. The output characteristic of the CNT TFTs is exhibited in figure 30(*c*).

As can be observed in the plot, the output curves can be fully saturated. The uniformity of the devices is illustrated in Figure 30(*d*). The normalized on current (*I*on/W) and threshold voltage (V_{th}) of 10 CNT TFTs were delineated in the figure. The results provided evidence for the applicability to implement the separated CNT thin films in large-scale fabrication processes.

In order to provide evidence for the practicality of employing CNT TFTs in actual applications, our research group has conducted a series of researches in CNT TFTs for displays and digital circuits. Due to the excellent transparency of CNT thin films, they were used as the channel material for the FETs in the drivers of arrays of AMOLEDs fabricated on glass substrate, which is an essential demonstration for the practicality of adopting the CNT TFTs in transparent display applications.

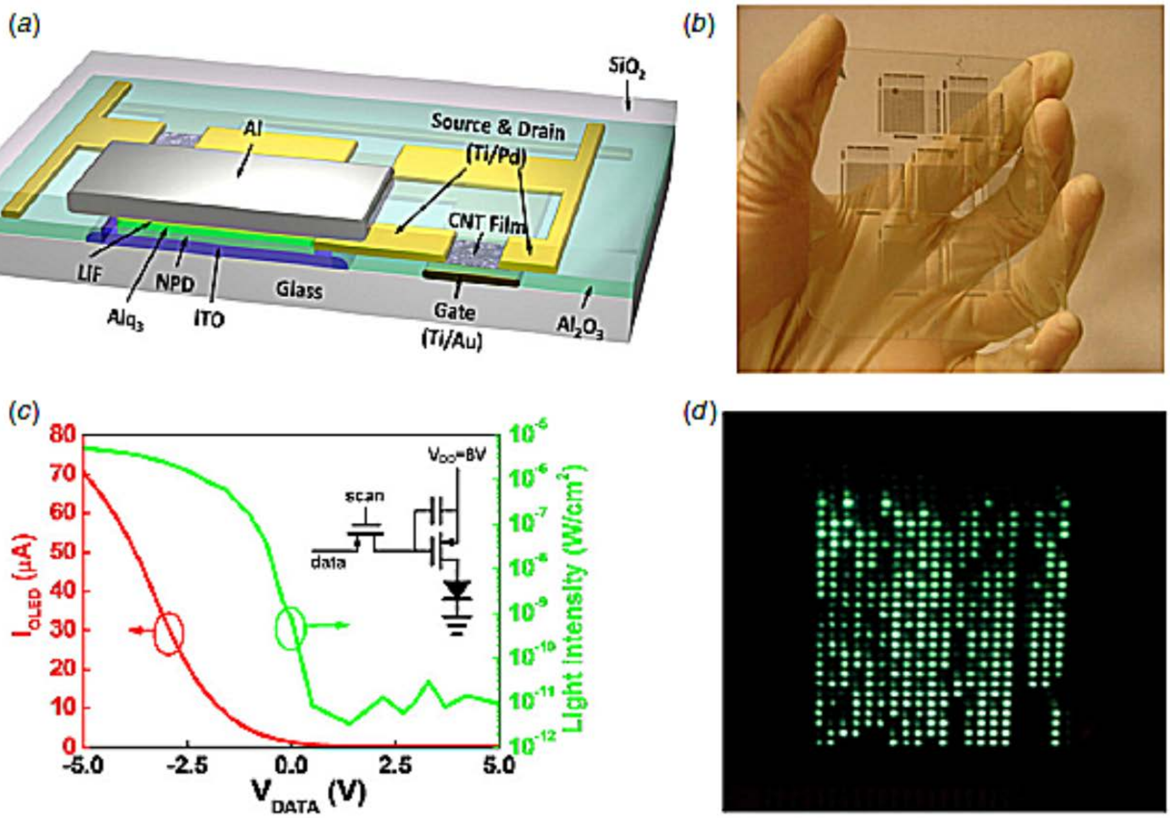


Fig. 31. (*a*) Schematic diagram of one pixel of OLED controlled by CNT network-based driver circuitry. (*b*) Arrays of AMOLEDs controlled CNT circuits fabricated on a transparent glass substrate. (*c*) Current (IOLED) and light intensity behavior of a pixel of the CNT controlled OLED unit. (*d*) Arrays AMOLED illuminating green light after being turned on by CNT network-based driver circuit.

As it was described in Figure 31(*a*), the device structure of one unit of the transparent AMOLED array was illustrated conceptually in the schematic diagram. This is the first demonstration of monolithically integrated AMOLED arrays with 500 pixels driven by 1000 CNT TFTs.

Figure 31(*b*) shows that all of the display elements in conjunction with the CNT TFT-based driver circuitry were fabricated on a glass substrate. The CNT TFTs in the drivers exhibited desirable electrical properties with mobility of around 30 cm^2 V^{-1} s^{-1} and Ion/off of ~104.

The mobility and Ion/off were extracted based on the transfer characteristic presented in Figure 31(*c*). Figure 31(*d*) illustrates a photographic image of 500 AMOLEDs pixels that were turned on by the CNT TFTs in the drivers, and the yield of the pixels was 70%. Most of the failure in the AMOLEDs was caused by the failure of the OLEDs themselves, not the CNT TFTs. The results clearly demonstrate the uniformity and reliability of the CNT TFTs.

One of the aforementioned advantages of CNT thin films is their flexibility, and that makes them an invaluable candidate in flexible electronics. Our group has previously demonstrated extremely flexible circuits based on CNT thin films exhibiting desirable electrical performance. CNT thin films have also been employed in control circuits for pressure sensing elements in electronic skin.

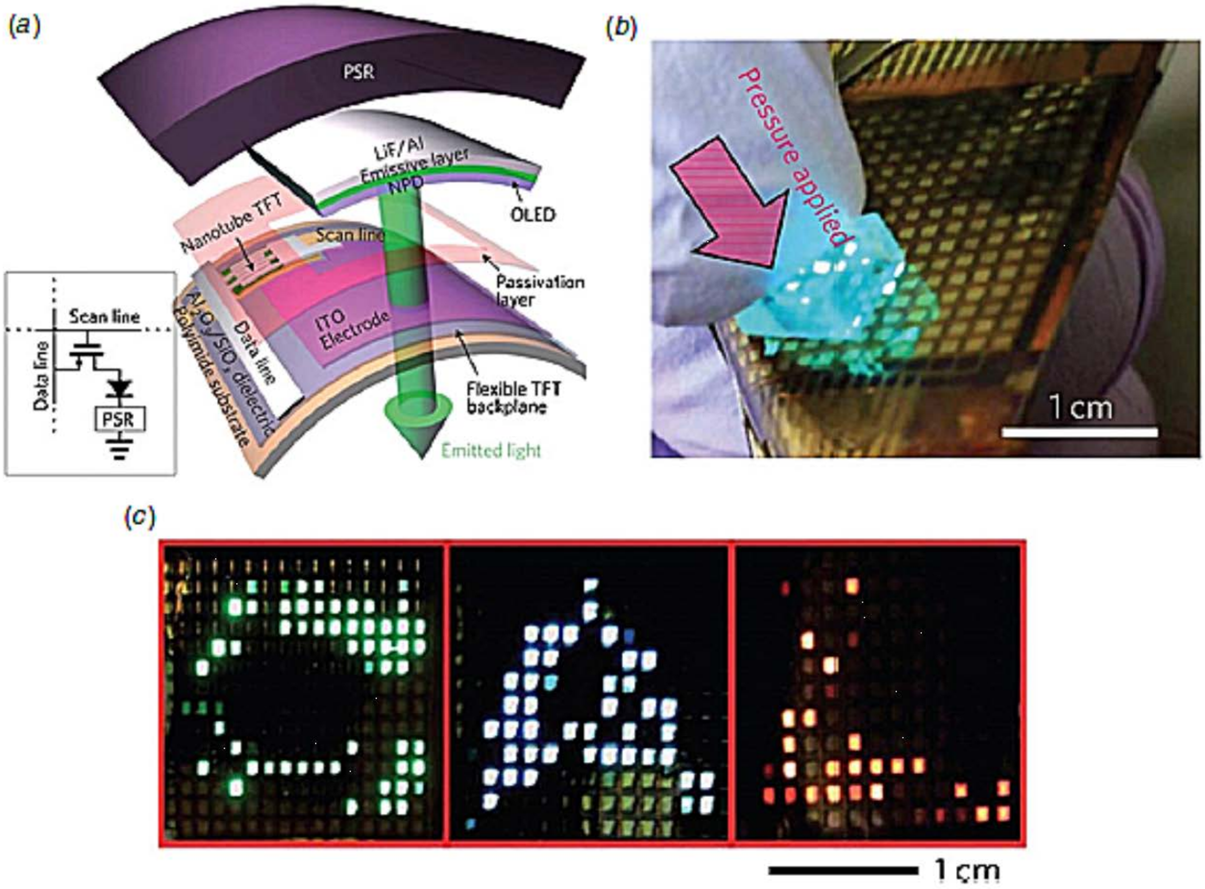


Fig. 32. (*a*) Schematic diagram of a pixel of PSR-based sensor integrated with an OLED in conjunction with a CNT thin film-based driver. (*b*) Demonstration of lighting arrays of OLEDs under the region where pressure is applied on the e-skin. (*c*) Illustration of letter C, A and L illuminated by the e-skin as pressure is applied to PDMS stamps formed in the letter shapes.

Figure 32(*a*) illustrates the schematic diagram of an element in user-interactive electronic skin arrays. The user-interactive electronic skin is a combination of CNT TFT-based control circuits, AMOLEDs and pressure sensing elements.

When pressure is applied to a certain region on a continuous piece of pressure sensitive rubber (PSR), arrays of CNT TFTs located under the region will turn on the OLEDs connected to the corresponding CNT TFTs under the same region. With this design mechanism, an area on the electronic skin is lit up whenever pressure is exerted to that area, and this mechanism is illustrated in Figure 32(*b*).

This is further demonstrated in Figure 32(*c*) with pixels of different colors of OLEDs connected to the CNTbased control circuits and the PSR. In the figure, it shows that polydimethylsiloxane (PDMS) slabs of different letters were placed onto the electronic skin, and a high proportion of the OLED pixels underneath the PDMS slabs were lit when pressure was applied to the configuration.

Among the different carbon nanostructures, carbon nanotubes are probably closest to real technological implementation into high-performance FET devices. This advance stems from the fact that (semiconducting) carbon nanotubes are the sole carbon nanostructure that combines several crucial features, namely the presence of a band gap, a big-enough size to enable electrical contacting with reasonable effort, as well as a low defect density and high stability.

While still considerable effort is spent to drive SWCNT-based FETs to the ultimate performance limit, there is also an emerging trend of integrating the various nanostructures into novel, all-carbon device architectures. Theoretical studies have predicted that carbon-based conductive materials are advantageous over metal contacts for electrically contacting carbon nanostructures.

This is due to the possibility of direct bond formation between the carbon material and the nanostructure, with a character very similar to the internal bonding in the latter, thus ensuring a well-matched bonding network and a good continuity of electronic structure.

As another advantage, the relatively small work-function difference among carbon nanostructures helps avoiding changes in the electrically addressed material due to contact doping. Moreover, in the particular case of carbon nanotubes, the tube diameter of 1–2nm renders them close-to-ideal contacts for molecules.

Experimentally, first steps have already been undertaken toward integrating grapheme and p-conjugated organics, as reflected by a very recent study using patterned graphene as bottom source and drain electrodes for efficient hole injection into a pentacene conduction channel. Further to this, metallic SWCNTs have been utilized as electrical leads to contact pentacene nanocrystals.

Due to the favorable gate electrostatics of the sharp 1D electrode geometry, FETs obtained in this manner display orders of magnitude conductance modulation for channels of only several molecules in length.

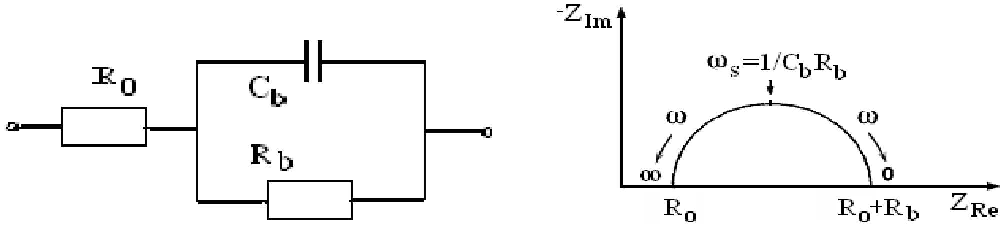
Another noteworthy development in this context is the use of carbon nanotubes as crystallization templates for the patterned growth of organic single crystals directly into device structures. Besides integration approaches, also novel device concepts can be expected to become increasingly important in the future.

Laboratory work 3-1.

Impedance spectroscopy of nanosemiconductors

Impedance spectroscopy is one of the most informative methods of conduction study of the systems with developed porous structure. Analysis of impedance hodograph within the framework of equivalent circuit method allows to separate the contribution to the electrical conductivity of the grain volume, surface and intercrystalline boundaries section, to define the mechanisms of the ions movement, their mobility and the diffusion coefficient [1-2].

Linear circuit of resistors and capacitances. Typically, in the high-frequency spectrum region the experimental results of the ultradisperse system study are consistent with the calculations of the impedance of independent on frequency equivalent linear resistors and capacitor circuit. The dielectric layer resistance is shunted by their capacity, the value of the active component of impedance Z_{Re} is determined by volume clusters ohmic resistance R_v, connected in series with $R_B C_B$ chain of intergranular interface (Fig. 1).



a b

Fig. 1. The equivalent linear circuit of resistors and capacitances of dispersed system (a) and the impedance dependence (b).

The elements of resistance R_o, R_b, and capacitance C_b are active and reactive components of general frequency dependent complex impedance Z(ω):

$$Z(\omega) = Z_{\text{Re}} - jZ_{\text{Im}} \tag{1}$$

where $Z_{\text{Re}} = R_o + \frac{R_o}{1+\omega^2 C_b^2 R_b^2}$; $Z_{\text{Im}} = \frac{\omega C_b R_b^2}{1+\omega^2 C_b^2 R_b^2}$, Z_{Re}, Z_{Im} – respectively, the real and imaginary components of the total impedance.

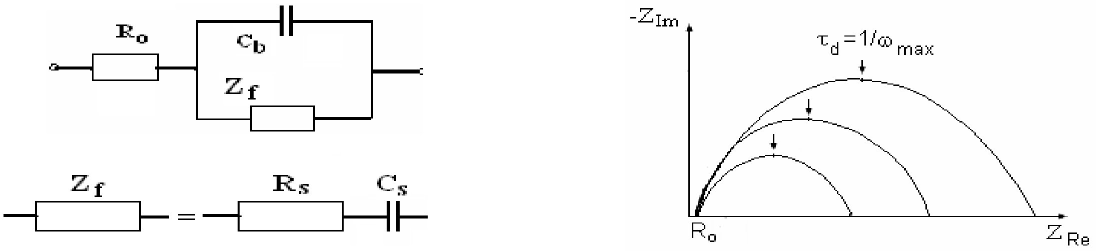
Equation (1) in the complex plane coordinates Z_{Re},$-Z_{Im}$ represents the circle equation with the center coordinates $Z_{Re}=(R_0+R_b)/2$and $Z = 0$ *and* $Z_{Im}=0$ and give the possibility of simple extrapolation of the frequency spectrum (Fig. 1 b). The maximum of this dependence is observed at the frequency $\omega_s=1/R_BC_B$. When $\omega=0$ and $\omega=\infty$ imaginary component of the impedance $Z_{Im}=0$ and the current flow is determined only by active components. When $\omega=\infty$ the ohmic resistance R_0 is determined, and when $\omega=0$ – values $R_B + R_0$.

Nonlinear circuit of resistors and capacitances. In the low-frequency spectrum region should take into account the dispersion of the impedance amplitude and phase angle shift due to the presence of ion-proton current component. The impedance hodographs have the form of semicircles whose centers can move relative to the imaginary axis, depending on the relaxation frequency.

In the equivalent circuit (Fig. 2.) a similar feature of the impedance hodograph is considered by replacement of capacity C_B by constant phase element (CPE) [1]. The use of the constant phase element is useful in the case of non-ideal capacitor when the distribution of impurity and defect centers is nonuniform by oxide thickness or when the composition and surface of the electrode is inhomogeneous. The frequency dependence of impedance is described by the expression:

$$Z_{\mathrm{Re}} = R_0 + \frac{R_b}{1+(i\omega\tau_d)^s} \qquad (2)$$

where $\tau_b = (R_bC_b)^{1/s}$ - in first approximation corresponds to the average value of effective relaxation time of the conductive cluster and s - describes the displacement of the center of half circle hodograph.



a **b**

Fig. 2. The equivalent circuit of disperse system (a) and impedance

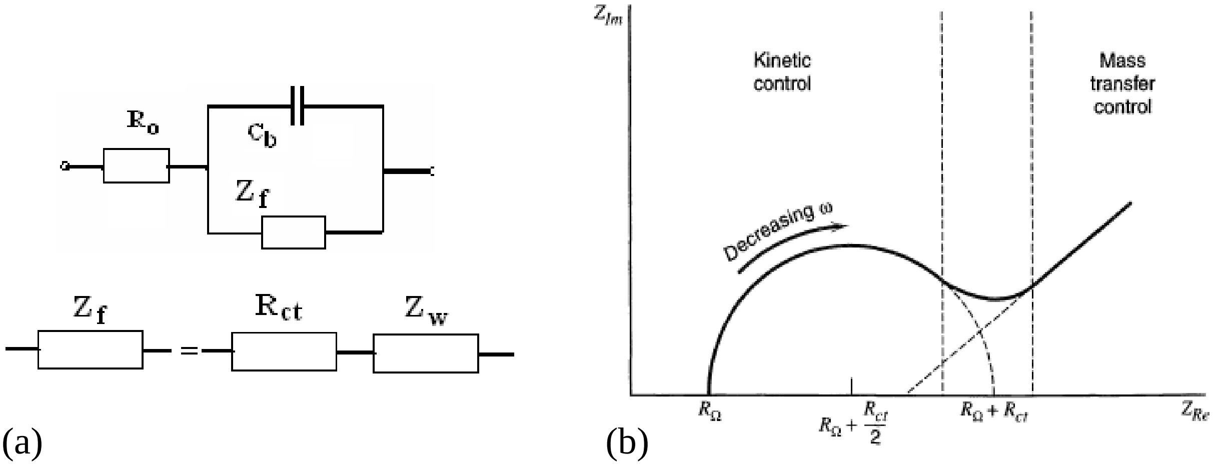
dependence (b) by CPE approximation.

Nonlinear circuit of resistors and capacitances. Concentration Warburg impedance. It is used in the case of quick reaction or adsorption when limiting factor is the slow nonstationary diffusion of electroactive material (Fig. 3). The diagram shows both kinetic area (half circle) and diffusion (direct line). The reaction resistance includes the charge transfer resistance R_{ct} and Warburg impedance $Z_v = Rw + Cw$.

A characteristic feature of Warburg diffusion impedance is linear dependence of active and capacitance resistance of $\omega^{-½}$:

$$Z_{\mathrm{Im}} = \frac{1}{\omega C} = A\omega^{-\frac{1}{2}} \qquad\qquad Z_{\mathrm{Re}} = R_o + A\omega^{-\frac{1}{2}} \tag{3}$$

that allows to determine the concentration constant by the angle of obliquity: $A = \frac{\gamma}{z^2 F^2 C_i \sqrt{2D_i}}$ and, accordingly, the ions diffusion coefficient D_i and their concentration C_i in volume.



(a) (b)

Fig. 3. The equivalent circuit of dysperse system (a) and impedance dependence (b) taking into account the concentration Warburg impedance.

2. Measurements.

Measurements on alternating signal (impedance and phase) were carried out using immittance meter E7-20 in automate mode (Fig. 4). The features management and registration were insured by computer via serial port using the application programme package. The device calibration was carried out on an empty sample holder. Voltage range was set from -4 to +4 V, the amplitude of the signal - 0.1 V, that was selected to prevent nonlinear effects in samples [2]. Internal averaging in the device was set to 100, when external during the reading to 5.

The reliability of the voltage dependence was also provided by prior tracking of impedance kinetics that allowed to consider this component during the resulting characteristics analysis.

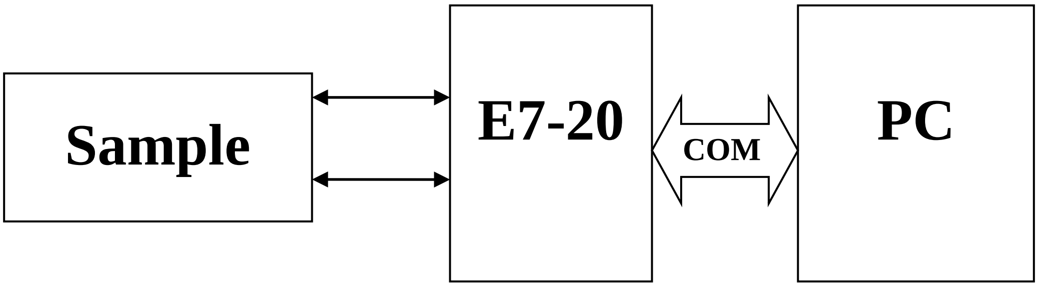


Fig. 4. The complex impedance measuring setup.

The sample holder was made on the basis of hold-down contacts.

In the case where the samples had no vacuum deposited contacts, connection to device was carried by clamping of the free layer between metal contacts using clothespeg with a platinum coating. The wires were directly outputed from the contact and were connected to the holder of immittance meter. In this case, the sample derived parameters were further transferred to specific values that were made to take into account the changes in area and thickness of the structure during the experiments. Additionally, the input parameters of the sample holder were excluded from the results.

3. Safety

1. Working alone in the laboratory is strictly forbidden. Students may work in laboratories only during regularly scheduled times and then only when supervised by an authorized Teaching Assistant or a member of the faculty.
2. Students may only perform authorized experiments and then only at the assigned time
3. Open flames (burners, matches, etc.) are not permitted.

4. All laboratory workers must know the location and proper use of all laboratory safety equipment, including eyewash, safety shower, fire extinguisher, and telephone. You will be expected to locate all safety equipment during the first week's lab.

5. You should know how to exit the lab via two different exits, in the event of an emergency.

6. All accidents, including contact with chemicals, cuts, burns, or inhalation of fumes must be reported to an instructor immediately. Any treatment beyond emergency first aid will be referred to the student infirmary.

7. All laboratory workers must know the location and proper use of all laboratory safety equipment, including eyewash, safety shower, fire extinguisher, and telephone. You will be expected to locate all safety equipment during the first week's lab.

8. It is your responsibility to read and abide by the "Laboratory Safety" section of the lab manual and to keep it with you in the laboratory. Any other safety handouts or special precautions mentioned during lab lecture must be scrupulously observed.

9. Laboratory equipment and work area must be cleaned after finishing work each week.

10. Failure to observe laboratory safety rules and procedures may result in injury to you or to fellow students. For a first violation, you may be asked to leave the laboratory for the day, while for repeated violations, you may be dropped from the course with a grade of E, at the discretion of the instructor.

4. Tasks.

1) To measure impedance dependence of tested structure Ti-porous TiO_2-Si.

2) Calculate the equivalent schema parameters.

3) To make conclusion about mechanisms of charge transport in porous TiO_2.

5. Reference information

1. Bard A.J. Electrochemical methods. Fundamentals and applications. / Bard A.J., Faulkner L.R. - N.-Y.: Willey, -2001. – 833 p.

2. Boukamp B.A. Electrochemical impedance spectroscopy in solid state ionics: recent advances / Boukamp B.A. // Solid State Ionics. -2004. –169. – P. 65–73.

Lab work 3-2

Manufacturing of porous Silicon based structures by electrochemical etching

1. Introduction

1.1. Porous Silicon as material

Porous Silicon is a promising material because of its nano- and microstructure. Sponge-like form introduces unique properties to porous Silicon, like large internal surface area (300-1000 m^2 per cm^3), modulated optical properties and quantum effects. One of fields of porous Silicon based structures usage is a sensor application. Majority of semiconductor sensors are based on application of electrical and electronic transducers. Nanoscale semiconductors give opportunity to produce optical and luminescence transducers. Large internal surface of porous Silicon can be used for adsorption of analyzed substance or placing specified receptors in pores. Change of optical properties like refraction coefficient after analyte filling of pores can be detected and measured.

Porous Silicon is the material made of bulk mono- or polycrystalline silicon by electrochemical etching in alcohol-fluoride acid (HF) solutions. This process forms connected cavities with size larger or equal 1 nm. Cavity size, porosity (fraction of dissolved silicon to the whole sample volume) depends on several conditions like doping rate of silicon, proportion of HF in initial solution, current density. Initial conditions are used for producing porous Silicon with predefined properties. Refraction coefficient depends on porosity of the prepared material. Porous Silicon based multilayer structures like Bragg reflectors (BR), Fabry–Pérot interferometers are produced varying etching current density. Cavities are much smaller than visible wavelength, so formulas of efficient medium can be applied to compute spectral dependences of the reflection coefficient of porous Silicon based single and multilayered structures.

1.2. Production of porous Silicon by electrochemical etching

There are several known methods porous Silicon manufacturing like anodization, mechanical means, photo-induced, vapor etching, stain etching, etc. Most simple and tunable methods is anodization or electrochemical etching. The porosification of the surface of a silicon wafer is generally referred to as “anodization” and occurs when the wafer is anodically biased in a fluoride-based electrolyte solution. The most commonly used electrolyte component is hydrofluoric acid (HF), with ammonium fluoride being less common. The degree of anodization is defined by the layer formation rate and porosity and, together with pore morphology, depends on wafer type and resistivity, the applied current density and time, and the electrolyte composition (HF concentration, with or without additives). Secondary parameters include electrolyte temperature and pH. The surface of a silicon wafer, as received from the manufacturer, will always be covered with a native oxide film. The oxide layer will be removed when the wafer is immersed in HF, although the cleanliness of the underlying surface can influence the anodization process and certain applications may therefore require wafer pre-cleaning prior to anodization.

A heavily doped wafer ($p^{+,++}$, $n^{+,++}$) can be readily anodized in a variety of HF-based electrolytes to form mesoporous silicon. A lightly doped wafer (p^-, n^-, and majority carrier concentration $<10^{18}$ cm^3) generally requires an ohmic contact to be made on the backside, either by ion implantation or by depositing a thin metal layer such as Al (for p-type wafers) and Au–Sb (for n-type wafers); mechanical abrasion of the backside of the wafer, before metal deposition, can often improve the ohmic contact.

Lightly doped p-type wafers can be anodized to form either microporous, mesoporous, or macroporous layers. Lightly doped n-type wafers have the additional requirement for light assistance, which involves illuminating the back or front side of the wafer to be anodized with high intensity white light (while minimizing radiative heating of the wafer/electrolyte). The light generates electron–hole pairs in the exposed area, the holes being “injected” into the silicon valence band, thereby facilitating the electrochemical dissolution process. With frontside illumination, a pseudo bilayer structure is formed, comprised of a very thin high porosity microporous layer at the surface above a much thicker (and lower porosity) macroporous layer, the relative thicknesses being dependent on the intensity (and wavelength range) of light and relative absorption coefficient of the wafer. Without illumination, or with backside illumination, a higher voltage is required and a low porosity macroporous layer results.


Fig. 1. Examples of Teflon etching cells

Etching process can be performed in chambers which are passive to Fluoride acid solutions made of materials like Polyethylene, Polypropylene, Polytetrafluoroethylene, etc. Usually these chambers look like in fig. 1. Chamber forms vessel where one of surfaces formed by Silicon wafer.

Anode which located in electrolyte should be acid-tolerant as well. Materials like Carbon, Platinum, and Gold can be used (Fig. 2).

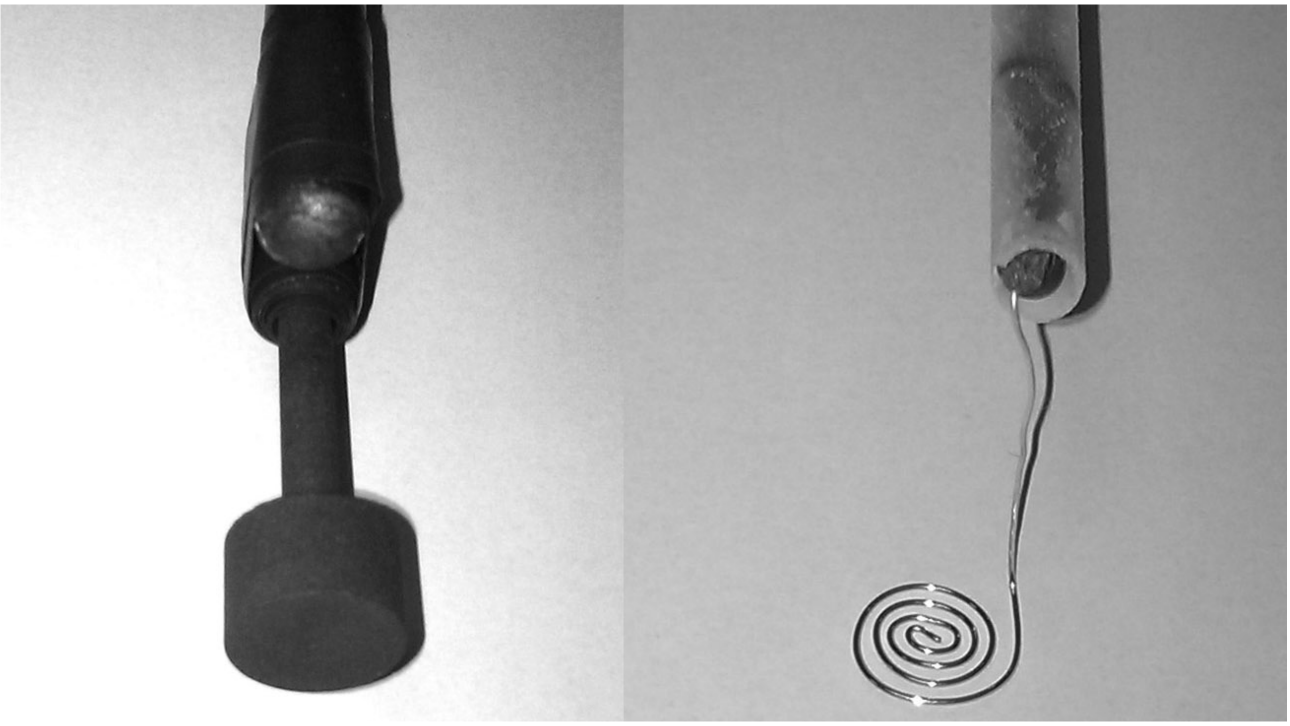


Fig. 2. Carbon and platinum electrodes

1.3. Multilayer porous Silicon based structures

Wide range of characteristics that affect properties of porous layer gives opportunity to produce structures with unique optical properties. For example change of etching current will affect porosity of silicon which will provoke change in refraction coefficient. It can be changed in both discrete and continues manner. Distributed Bragg reflectors (DBRs) are one dimensional (1D) photonic crystals consisting of a stack of discrete layers of two different refractive indices. The DBR allows rejection of a wide range of wavelengths of light in a specific region called photonic bandgap (PBG) of the structure. Design of DBR requires that each layer should have optical thickness (i.e. n*L , where n and L are the refractive index and thickness of a layer, respectively) of one quarter of design wave- length and the constituent layers must be phase-matched. Therefore, simple and accurate method is required to determine optical constants of PSi layers. Structure and optical spectrum are shown on fig. 3.

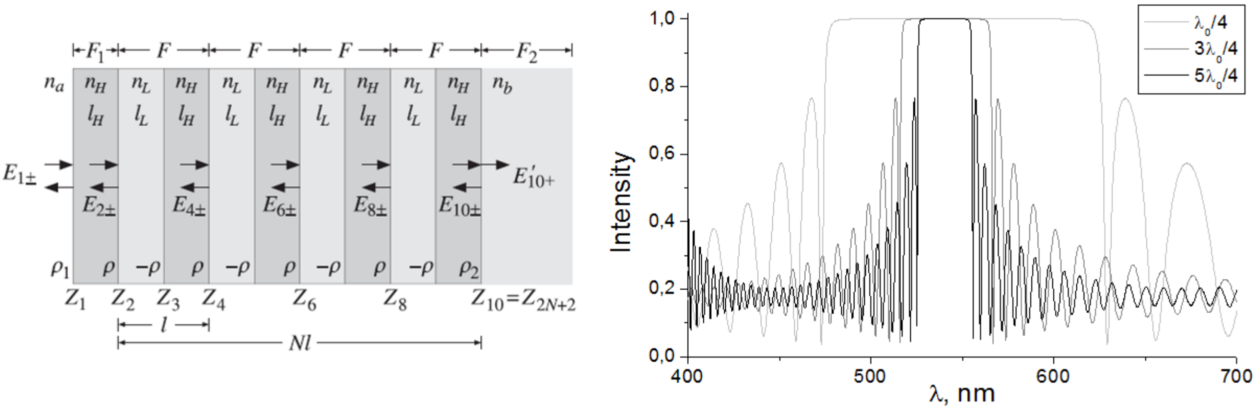


Fig. 3. Structure of Bragg mirror(left) and typical optical spectrum.

2. Safety

2.1. General code of conduct in chemical laboratory

1.Working alone in the laboratory is strictly forbidden. Students may work in instructional laboratories only during regularly scheduled times and then only when supervised by an authorized Teaching Assistant or a member of the faculty.

2. Approved safety goggles must be worn at all times when in a laboratory. You are responsible for bringing your own pair of approved goggles to lab each week.

3. Students may only perform authorized experiments and then only at the assigned time.

4. Open flames (burners, matches, etc.) are not permitted.

5. Use every precaution to keep all chemicals off your skin and clothing, out of your nose, mouth and eyes, and away from flames. It is strictly forbidden to eat or drink anything (including water) in the laboratory.

6. Long hair and billowy clothing must be confined when in the laboratory. Shoes are mandatory; sandals or open-toed shoes are not allowed, even if socks are worn. Shorts and skirts must cover the knees.

7. All laboratory workers must know the location and proper use of all laboratory safety equipment, including eyewash, safety shower, fire extinguisher, and telephone. You will be expected to locate all safety equipment during the first week's lab.

8. You should know how to exit the lab via two different exits, in the event of an emergency.

9. All accidents, including contact with chemicals, cuts, burns, or inhalation of fumes must be reported to an instructor immediately. Any treatment beyond emergency first aid will be referred to the student infirmary.

10. It is your responsibility to read and abide by the "Laboratory Safety" section of the lab manual and to keep it with you in the laboratory. Any other safety handouts or special precautions mentioned during lab lecture must be scrupulously observed.

11. Laboratory equipment and work area must be cleaned after finishing work each week.

12. Failure to observe laboratory safety rules and procedures may result in injury to you or to fellow students. For a first violation, you may be asked to leave the laboratory for the day, while for repeated violations, you may be dropped from the course with a grade of E, at the discretion of the instructor.

2.2. Actions with Fluoride acid and work in fume hood

Fluoride acid is weak but very corrosive liquid. It dissolute Silicon oxide and because of that can be stored and used only in non-glass vessels. Because of small molecule size acid is very volatile and have good penetration properties. In the lab plastic bottles with Fluoride acid are stored in fume hood in plastic cuvette.

Before work with acid hood compressor should be switched on. The person who works with acid should wear protection glasses and special grooves. All actions in fume hood are performed with pulled down protective glass.

If there is a suspicion that acid get on a skin first of all place of supposed acid contact should be immediately washed with ammonia hydroxide. Place of acid spoil should be neutralized with Calcium carbonate.

3. Process preparations

3.1. Etching parameters calculations and current source programming

Several parameters should be chosen or calculated for desirable pore size, porosity and porous layer thickness. For pore size usually choose Silicon type and doping level. Highly doped P-type Silicon will be used in current work which corresponds to mesoporous Silicon. Porosity-etching current calibration curve should be chosen according to desirable porosity depending on chosen Silicon type. Lower etching current is better. Each calibration curve has corresponding electrolyte parameters which should be used during solution preparation. We will use calibration curve for P++ Silicon and alcohol solution of Fluoride acid with ratio 1:1. Etching current density is chosen according to desired porosity on previously chosen calibration curve. There are different etching speed for different etching solutions and Silicon types. Another calibration curves are used to determine porous Silicon formation speed which corresponds to chosen etching current density. According to porous Silicon formation speed etching time can be easily calculated. While calibration curves indicates only current density current value should be calculated according to requested area of porous Silicon layer. Calibrating curves can be found in annex 1.

3.2. Etching procedure

Piece of Silicon with corresponding size is mounting inside etching cell. Fix force should enough to avoid flow. This is checked by pouring some amount of ethanol and checking its level for 3-5 minutes. If there no flow cell can be used for etching otherwise re-fix required. Cell should be installed as shown on fig. 4. Cathode positioned above Silicon surface inside cell. Mixer can be located above cathode if required. Cell's wire should be connected to "-" of current source and cathode – to "+".

Fig. 4. Cell installation.

After fume hood launch cell can be filled with required Fluoride acid solution. Work with acid solution can be performed only in grooves. Pour some solution from a bottle to plastic cup and carefully fill cell with solution from the cup. Solution should cover cathode and partially cover mixer. When preparations are completed etching process can be launched.

When etching time outs mixer and cathode should be taken out from the cell and rinsed with ethanol. Solution should be taken away with pipette to waste bottle. Then inner surfaces of cell being rinsed with ethanol twice and Silicon sample can be taken away from the cell and then put into ethanol for final rinsing.

Used chemical vessels should be washed in tap water and clean vessels rinsed with distilled water. Clean porous Silicon should be dried on air at room temperature for 10-15 minutes.

4. Task

Manufacture 2 porous Silicon layers on monocrystaline Silicon with parameters:

1) n = 3; d = 300nm
2) n = 2; d = 300nm

5. Reference information

1. Handbook of Porous Silicon by Leigh Canham 2014

2. F.A. Harraz, et al., A capacitive chemical sensor based on porous silicon for detection of polar and non-polar organic solvents, Appl. Surf. Sci. (2014)

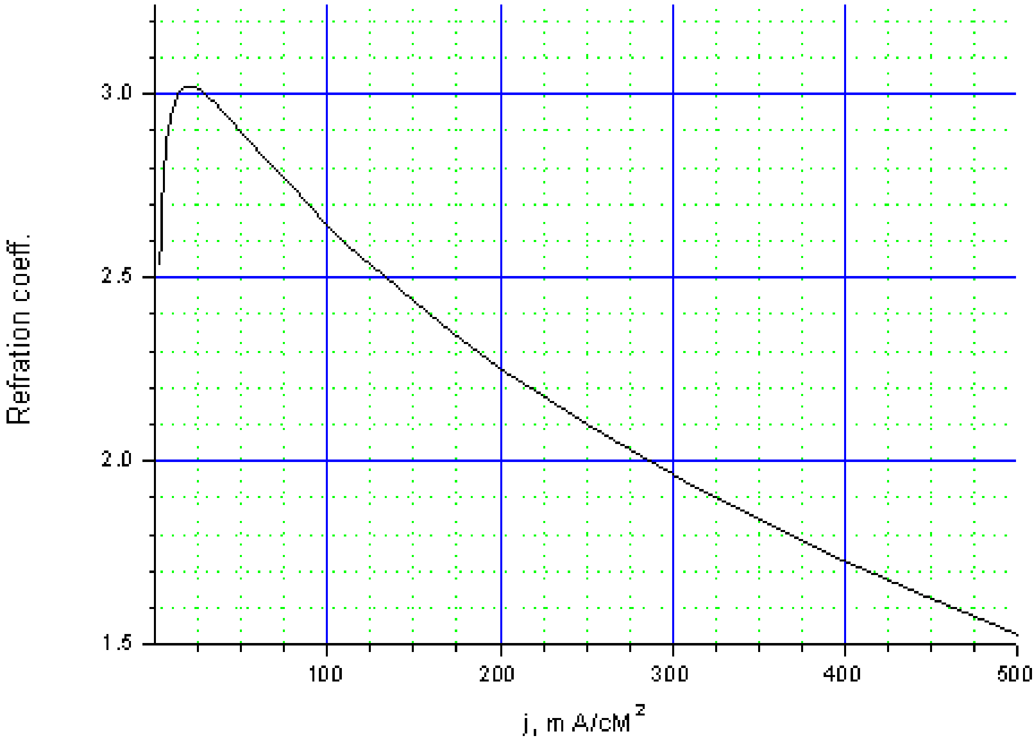
3. I.I. Ivanov, V.A. Skryshevsky, T. Serdiuk, V. Lysenko Kinetics of adsorption–desorption processes of alcohol molecules in porous silicon Bragg mirror // Sensors and Actuators B 174 (2012) 521– 526

4. Nalin H. Maniya, Sanjaykumar R. Patel, Z.V.P. Murthy, Simulation and fabrication study of porous silicon photonic crystal, Optik - International Journal for Light and Electron Optics, Volume 125, Issue 2, January 2014, Pages 828-831

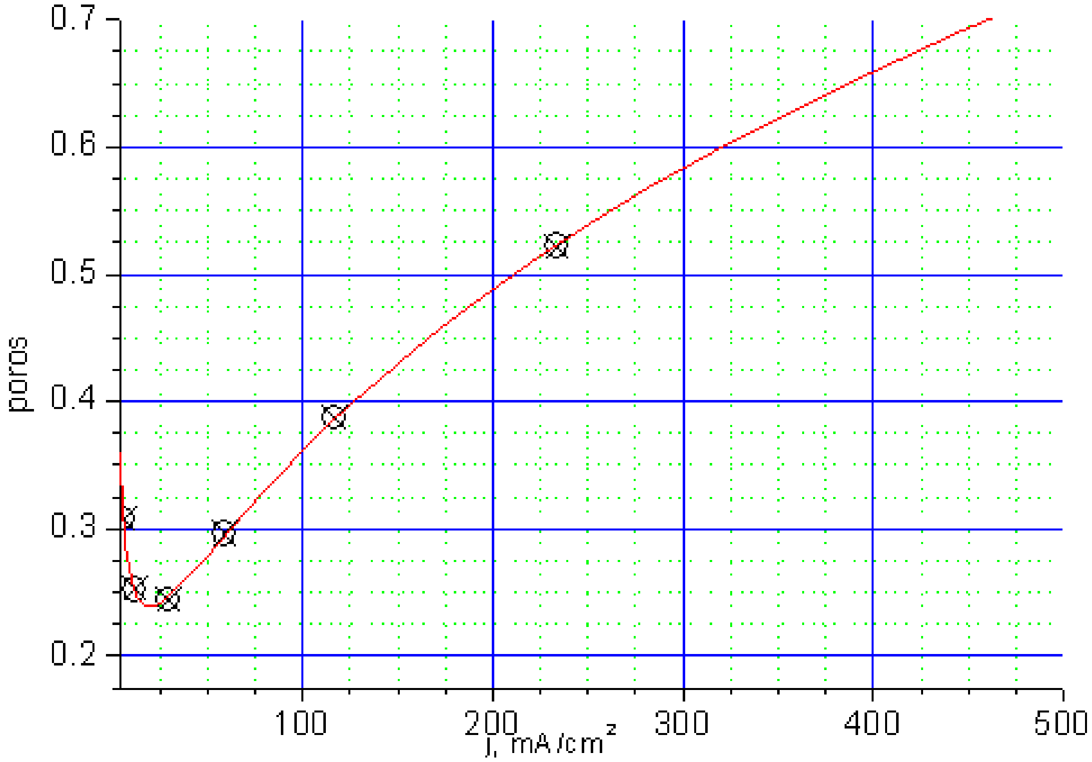
5. https://www.dartmouth.edu/~chemlab/info/safety/rules.html

Annex 1. Calibrating curves

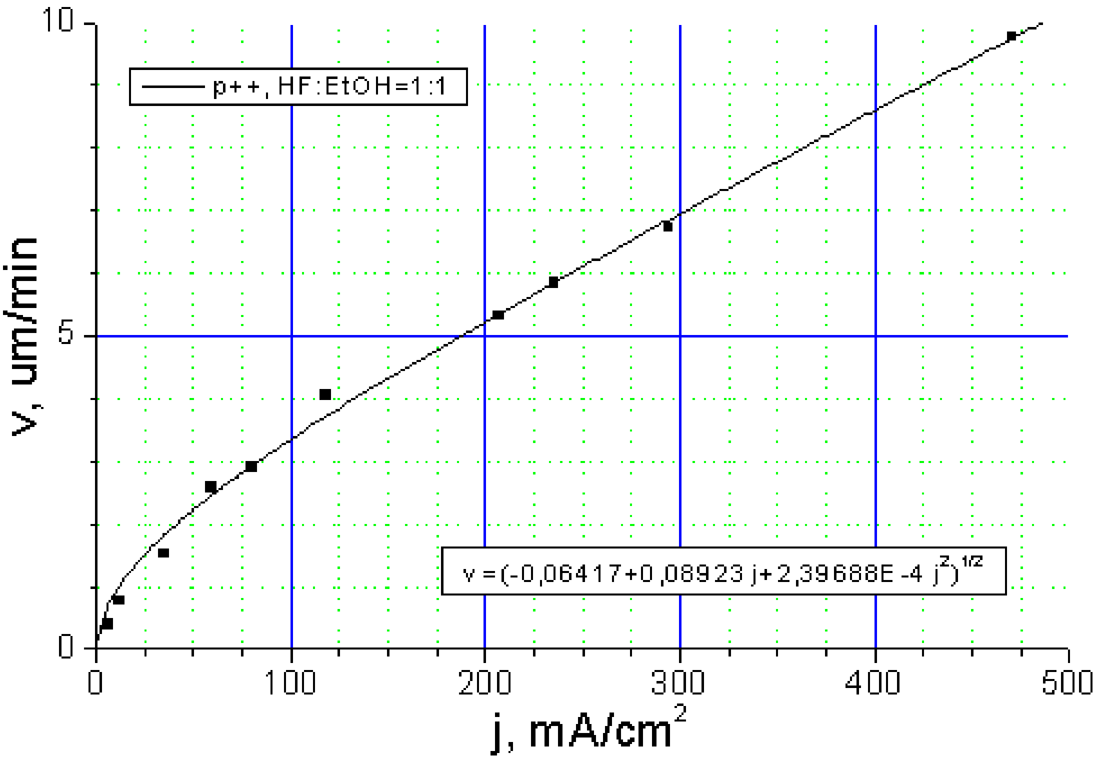
Calibrating curves for Ethanol solutions of 40% Fluoride acid in proportion 1:1



Refraction coefficient vs current density



Porosity vs current density



Etching velocity vs current dencity

Lab work 3-3

Photoluminescence (PL) spectra and principles of PL decay time measurement

1. Introduction

Last 15 years photoluminescence (PL) in porous silicon (PS) has been devoted much attention. Recently, much effort has been directed to clarify the mechanism of visible radiative recombination yield that can achieve 5 and more percents at room temperature, and seems to give a new positive pulse for sustainable development of silicon optoelectronics. The first interpretation of this phenomenon was based on quantum confinement in Si nanoparticles that results in a bandgap up shift and arising of a strong visible PL at room temperature. Another important interest to visible PL in PS is related with development of chemical and biochemical sensors that use luminescent transducer. Chemical sensors based on PS display some advantages comparing with other transducers. Thus, it is observed the reversible quenching of PL intensity in presence of different gases, the quenching growths with partial pressure and dipole moment of adsorbed molecules [1]. PL intensity of PS drops by three orders of magnitude when PS is immersed in organic solvent of dielectric constant varying from 2 to 20 [2]. However, varying PS samples preparation, conditions of PL measurement the contradictory effects (increasing of PL intensity in gas atmosphere, independent of PL intensity versus dielectric constant, etc) are observed too [1, 3, 4]. The detail knowledge about kinetic of radiative recombination can give an important new insight in the interpretation of PL lifetime and assists to elaborate high sensitive and reversible chemical sensors.

2. Principles of photoluminescence PL decay time measurement

PL spectra were studied at a room temperature by excitation the samples with a pulse of nitrogen laser λ = 337.1 nm, FWHM = 8 ns, power in pulse is 20 kW. The light spot was focused on the samples surface with diameter of 1 mm approximately. The proper band of emission was selected by a monochromator (MS2004, SOLAR T II) and registered by a combination of photomultiplier (HAMAMATSU C6270) and analog-to-digital board (up to 1 GHz sample rate). The signal was processed and analyzed by PC.

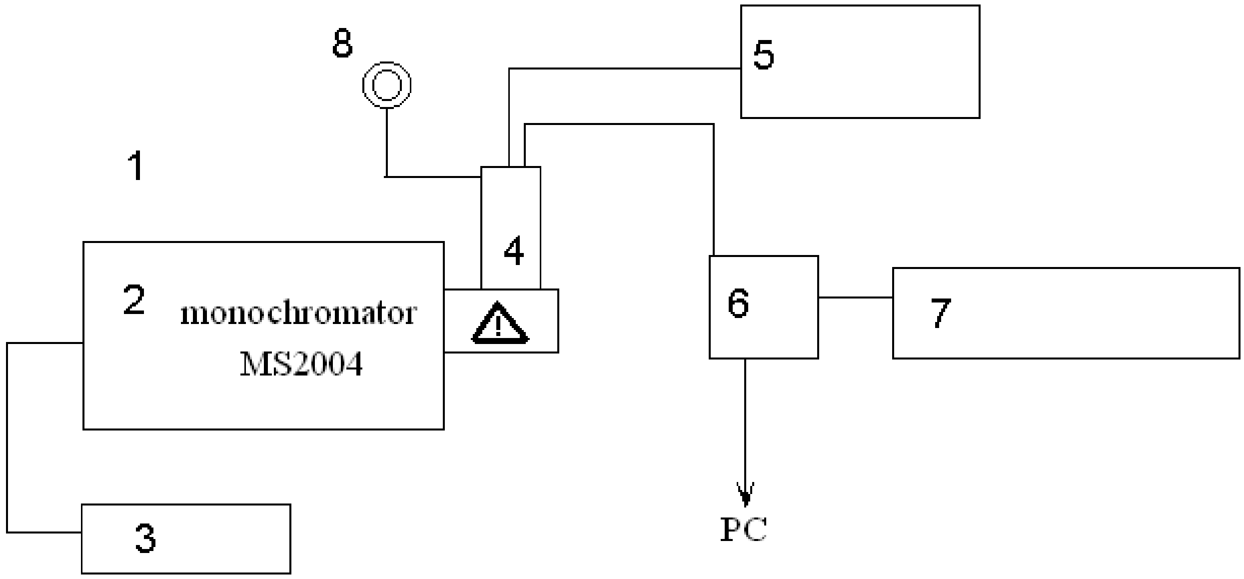


Fig.1. The general scheme of photoluminescence PL decay time measurement, (1)- monochromator MS2004, (2) - monochromator aperture, (3)- the power supply of monochromator, (4)- photomultiplier, (5)- the power supply of photomultiplier, (6) – amplifier, (7) - the power supply of amplifier.

Monochromator MS2004 is designed to operate in the spectral range from 260 nm to 3000 nm. The operating range overlaps by using 4-diffraction gratings, of which only one can be working at any given time. Control of monochromator is carried out via a serial RS-232 port (COM port PC) using the libraries provided by the developer.

The apparatus as detector used photomultiplier with built-in source of high voltage (1500 V) (fig.1). Photomultiplier is attached to the output window of monochromator and carefully isolated from extraneous light. Power of Photomultiplier carried a constant voltage of 15 V from source (5). Regulation of high voltage anode photomultiplier (and hence the output level) is using the potentiometer (8) face value of 50 ohms. Output from the anode photomultiplier is fed to the amplifier.

Since output resistance of photomultiplier is several tens of MOhm, direct connection of the photomultiplier anode to measuring circuits is impossible without matching amplifier. In this scheme applied dc amplifier with an input impedance of about 10 MOhm. Schematic diagram of the amplifier is presented in fig.2.

Amplifier is on the base of the operational amplifier AD812 (bandwidth 145 MHz, slew rate front 1600 V / us). The first stage is a follower amplifier with high input impedance and signal level limiter (diodes D1, D2). The second stage - inverting amplifier with a gain equal to 10 feeds power from 2-polar voltage source + 8 V. The output amplifier is connected directly to analog-to-digital board using coaxial cable.

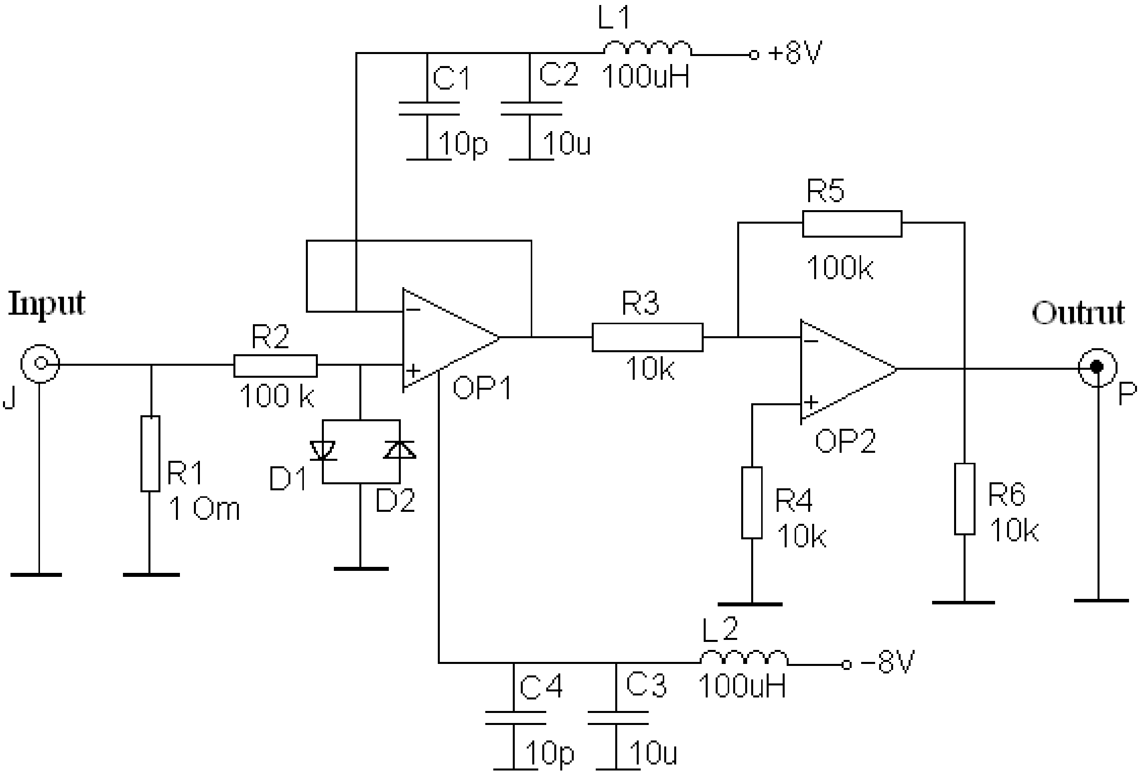


Fig.2. The scheme of amplifier

The time correlated single photon counting technique (TCSPC) is applied for time-resolved PL spectra and decay time measurements [4]. The method uses an idea that luminescence intensity at a given time is proportional to the probability of a photon occurring at this time. Thus in TC-SPC technique we wait for a single photon coming from the luminescent sample and measure a time at witch this photon was captured. Continuing such measurements we’ll get a set of time intervals passed between excitation pulse and photon capturing. Having enough of these time intervals one can build a photon appearing probability distribution and in such way obtain a luminescence transient signal. TCSPC displays an excellent sensibility and signal acquisition precision but takes too much time to gain a necessary statistics (from tens minutes till hours) for signal reconstructing.

In our work we use a slightly different method to obtain PS luminescence transient signal: multiple photon counting or multichannel scaler (MS). MS operates within fixed time windows (from several ns to μs) at witch several photon pulses can come from photon counter tube. This method is less time-precise as TCSPC but saves a lot of time and computer memory while data acquisition. A sequence of repetitive after excitation signals can be accumulated over many signal periods. This allows obtaining of transient waveform of desired precision but with a fixed time step corresponding to the data fetching time window. It uses a fixed time window of 200 ns to fetch the data from the counter and a memory array to keep 1000 counts. This allows monitoring up to 200 µs of signal decay (200 ns time window steps per 1000 steps). Hamamatsu photon counter is assembled with band-pass filter and objective lens.

2. Task

1) Switch on equipment:
    - a) laser power supply (10 min)
    - b) the power supply of monochromator
    - c) the power supply of photomultiplier
    - d) the power supply of amplifier
2) To measure of photoluminescence spectra
3) To measure photoluminescence lifetime

3. Reference information

1) Handbook of Porous Silicon by Leigh Canham 2014

2) S. Fellah, F. Ozanam, N. Gabouze, J. — N. Chazalviel // Phys.Stat.Sol. (a) — 2000. — 182. — P. 367.

3) H. A. Lopez, X. L. Chen, S. A. Jenekne, P. M. Fauchet // J.Lum. — 1999. — 80. — P. 115.

4) A. I. Benilov, I.V. Gavrilchenko, I. V. Benilova, V. A. Skryshevsky, M. Cabrera Sensors and Actuators A 137 (2007) P. 345–349

Annex 1. Monochromator MS2004

Structure of monochromator MS 2004 and principle of its operation. Optical scheme MS 2004

MS2004 (Fig.3) is built on an asymmetric scheme of Czerny-Turner, which uses different focal distances for the collimator and chamber of mirrors in order to minimize the effect of the coma.

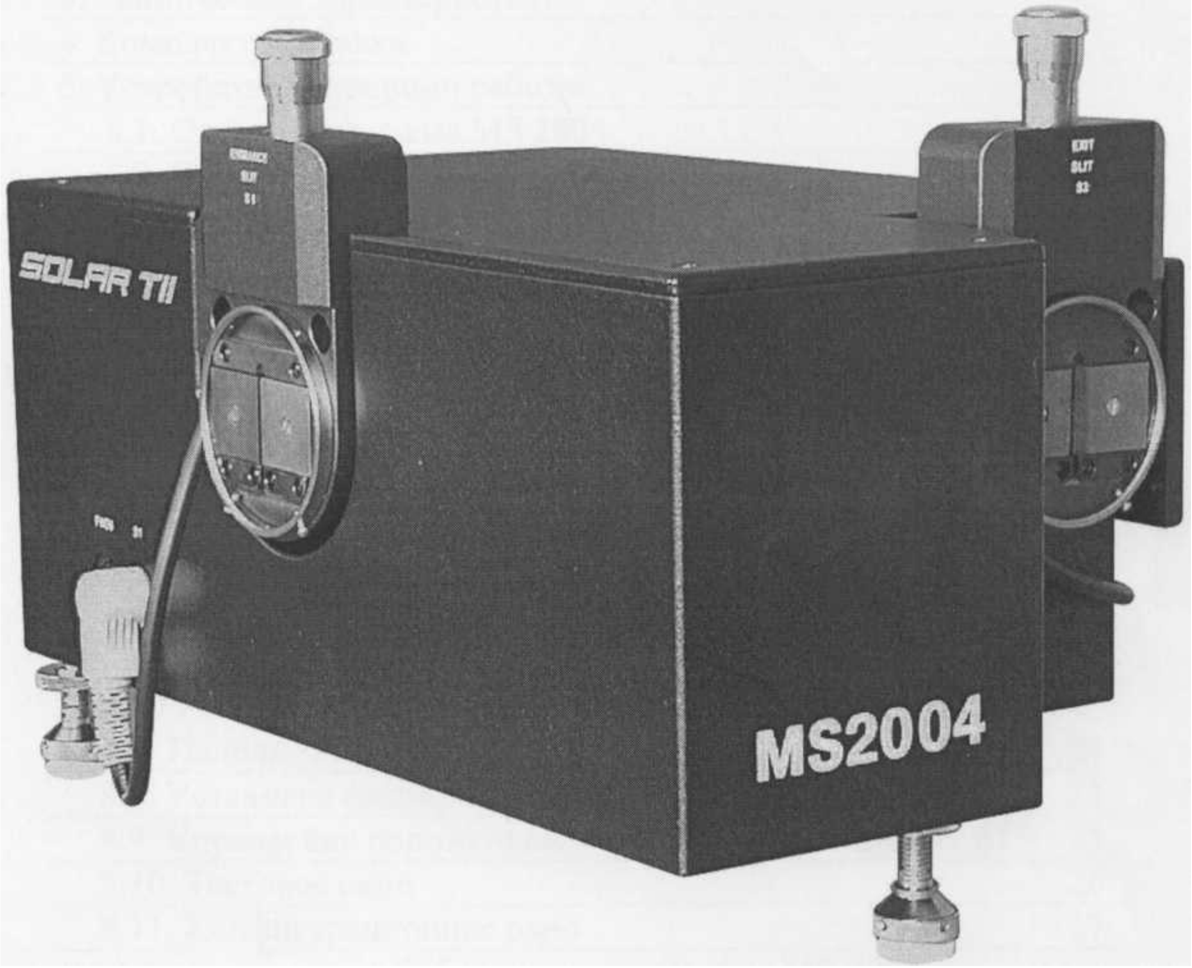


Fig.3 Monochromator MS 2004

Optical scheme MS2004is shown in Fig. 4. After passing through the input slit (1) and the rotary mirror (2), the radiation is directed by collimating spherical mirror (3) on the diffraction grating (4). The diffraction grating (4) converts the parallel beam from each point of the input slit (1) into a fan of monochromatic parallel beams. The chamber spherical mirror (5) forms on the focal plane (6) monochromatic image of the input slit (1). These all images form the spectra with specific range of wavelengths.Light gate, consisted of a shutter (9), blocks the light passing through the input slit (1).

The diffraction grating 4 is mounted on the four-turret. The diffraction gratings are change by turning of turret on 90 ° around the axis PP’. By turning of the grating (4) around the axis 00 ', it is possible to modify the spectral interval of wavelengths in the focal plane (6). The axis 00' passes through the center of grating (4) and coincides with the direction of the strokes.

The folding mirror (7) is used to select the detector port. When folding mirror is in position 7', the side port (output 2) is selected as the output port, when the folding mirror is in position 7, the axial port (output 1) is used as an output port.

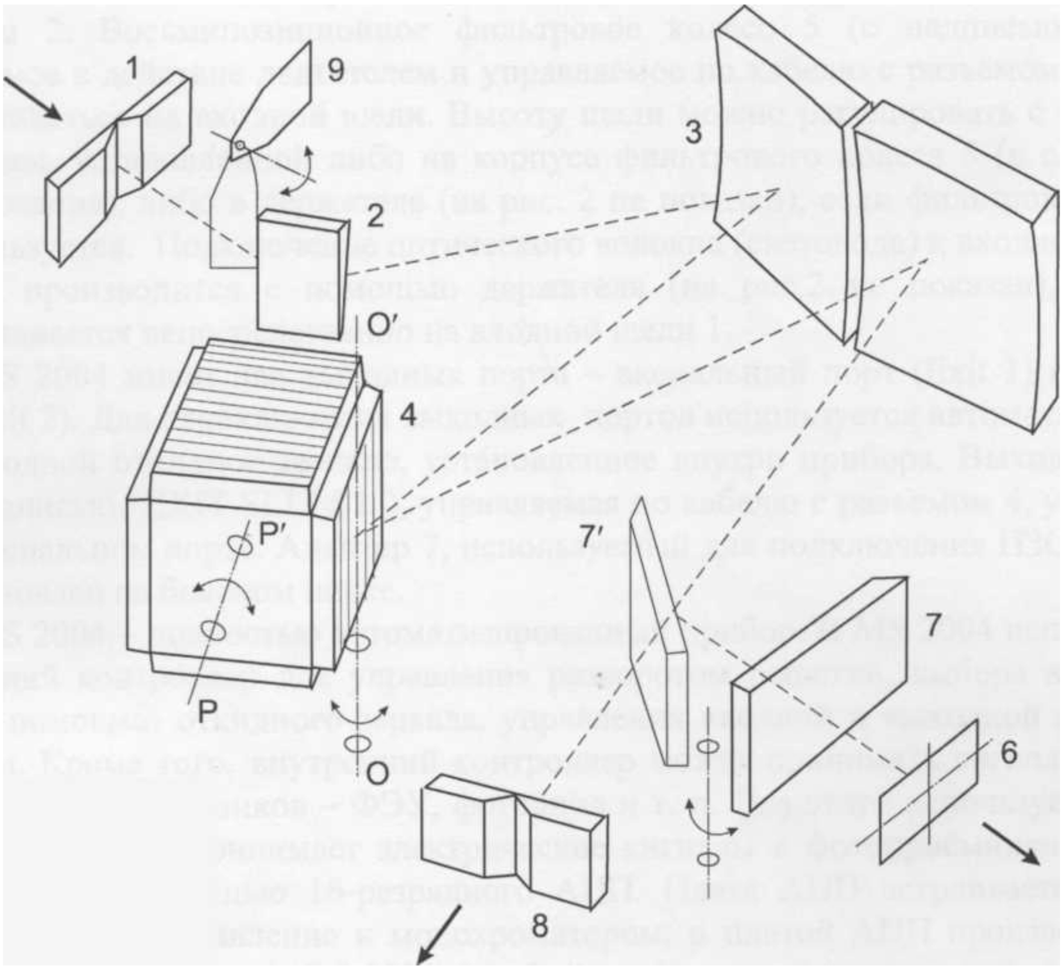


Fig.4. Optical scheme of MS2004

1 - input slit, 2 - rotating mirror, 3 - collimator mirror, 4 - diffraction grating, 5 - chamber mirror, 6 - the focal plane, 7 - folding mirror, 8 - output slot, 9 - gate.