Among the many materials that have been proposed to supplement and, in the long run, possibly succeed Si as a basis for nanoelectronics, carbon nanotubes (CNTs) have attracted the most attention. CNTs are quasi-one-dimensional materials with unique properties ideally suited for electronics. We briefly discuss the electrical and optical properties of CNTs and how they can be employed in electronics and optoelectronics. We focus on single CNT transistors, their fabrication, assembly, doping, electrical characteristics, and integration. We also address the possible use of CNTs in optoelectronic devices such as electroluminescent light emitters and photodetectors.

Phaedon Avouris* and Jia Chen
IBM Research Division, T. J. Watson Research Center, Yorktown Heights, NY 10598, USA
*E-mail: avouris@us.ibm.com

In the last few decades, we have witnessed exponential advances in Si microelectronics. These advances in computing, communications, and automation are affecting just about every aspect of our lives and are responsible for bringing in the 'information age'. To a large extent, these advances have been the result of the continuous scaling, i.e., miniaturization, of electronic devices, particularly of the metal-oxide-semiconductor field-effect transistor (MOSFET), that has led to denser circuitry and faster switching. For example, since 1970 the physical length of the gate and thickness of the gate insulator of MOSFETs have been scaled by factors of about 500 and 120, respectively. However, the scaling cannot continue forever; a number of fundamental scientific as well as technological limitations place lower limits to the size of Si devices. These involve, among others, electron tunneling through short channels and thin insulator films and the associated leakage currents and passive power dissipation, short channel effects, variations in device structure and doping, etc.

The realization of the approaching limits of scaling has inspired a worldwide effort to develop alternative device technologies. Some involve radical departures from the existing technology and are considered for the long term. A shorter term approach discussed here maintains the field-effect transistor (FET) principle, as well as the general current circuit architecture, but replaces the Si channel of the FET by a one-dimensional nanostructure with superior electrical transport properties. In addition to the efforts to develop new electronic devices, direct bandgap one-dimensional nanostructures are attracting attention because of the desire to base both electronic and optoelectronic technologies on the same material. Among the different one-dimensional materials, single-walled CNTs have many highly desirable and distinctive device properties.
We briefly examine the unique electronic and optoelectronic properties of CNTs. We discuss the structure and fabrication of single CNT-FETs, their properties and electrical characteristics, doping and device integration. We also describe simple CNT optoelectronic devices such as electroluminescent light emitters and photodetectors.

**Electronic structure of carbon nanotubes**

Single-walled CNTs are extremely strongly bonded, hollow carbon atomic structures with diameters typically in the range of 1-3 nm. Their electronic structure is usually described in terms of the electronic structure of a folded ‘graphene’ sheet (a layer of graphite)\(^5-7\). The CNT circumference is then expressed by a chirality vector \( \mathbf{C} \) connecting two crystallographically equivalent sites of the two-dimensional graphene sheet (Fig. 1): \( \mathbf{C} = n\mathbf{a}_1 + m\mathbf{a}_2 \), where \( \mathbf{a}_1 \) and \( \mathbf{a}_2 \) are the unit vectors of the hexagonal honeycomb lattice. The structure of any CNT can therefore be described by a pair of integers \((n,m)\) that define its chiral vector.

The interesting electrical properties of CNTs are largely the result of the unusual electronic structure of graphene itself. In going from graphene to a CNT by folding, one has to account for the additional quantization arising from electron confinement around the CNT circumference. This circumferential component of the wave vector \( k_C \) can take only values fulfilling the condition \( k_C \cdot \mathbf{C} = 2n \), where \( n \) is an integer\(^5-7\). As a result, each graphene band splits into a number of one-dimensional subbands labeled by \( n \). These allowed energy states are cuts of the graphene band structure. When these cuts pass through a \( K \) point (Fermi point) of the graphene Brillouin zone, the tube is metallic. Otherwise the tube is semiconducting. It can be shown that an \((n,m)\) CNT is metallic when \( n = m \); it has a small gap when \( n - m = 3i \), where \( i \) is an integer; while CNTs with \( n - m \neq 3i \) are truly semiconducting\(^5-7\).

Semiconducting CNTs have a diameter-dependent bandgap \( E_g \). A single-particle, tight-binding description of the electronic structure gives \( E_g = \gamma \left( 2a/d_{CNT} \right) \), where \( E_g \) is the bandgap energy, \( \gamma \) is the hopping matrix element, \( a \) is the C-C bond distance, and \( d_{CNT} \) is the diameter of the nanotube. Inclusion of electron-electron interactions raises the size of \( E_g \) significantly, but its \( 1/d_{CNT} \) dependence appears to remain valid. While the electrical and optical bandgaps of semiconducting CNTs were initially considered to be identical, based on the single particle model, we now know that the optical gap is smaller because of the attractive electron-hole (e-h) interaction and that the optical excitations of CNTs involve transitions to exciton states, not interband transitions (Fig. 2)\(^16-20\).

**Transport properties**

In general, scattering of carriers by defects or phonons determines the electrical transport properties of materials and devices. Because of the lack of boundaries in the perfect, hollow cylinder structure of CNTs, there is no boundary scattering, which plagues other nanostructures such as thin body Si MOSFETs, nanowires, or graphene slices. Furthermore, CNTs are quasi-one-dimensional materials in which only forward and backward scattering is allowed (low-angle scattering is suppressed). Yet the fact that they are not truly one-dimensional materials, that the cylinder has a finite size, allows the carriers in them to ‘bypass’ certain defects. Because of the large momentum change involved in backscattering, only scatterers with a strong, short-ranged

Fig. 1 Representation of the CNT atomic structure through the folding of a graphene strip. The chirality vector \( \mathbf{C} \) and the one-dimensional translational vector \( \mathbf{P} \) of a \((5,2)\) CNT are shown as an example.
potential are effective. Thus, elastic scattering mean free paths in CNTs are long, typically of the order of micrometers. At low energies (low applied voltages), the carriers in CNTs can interact with the acoustic phonons of the lattice, but the small electron-acoustic phonon (e-ph) coupling leads to long, temperature-dependent mean free paths of the order of a micrometer. The radial breathing acoustic mode, however, has a stronger e-ph coupling and when resonantly excited does affect transport. In long (many micrometer) CNTs, acoustic phonon scattering leads to diffusive transport, but the mobility can still be extremely high – of the order of 100 000 cm²/Vs.

At high bias, however, when the energy of the carriers reaches those of the zone-boundary (~160 meV) and optical phonons (~180 meV), strong inelastic scattering takes place and the mean free path is reduced to about 10-20 nm. Fig. 2 shows theoretical results for the phonon scattering rate as a function of the carrier energy.

Carbon nanotube field-effect transistors

The first CNT transistors were fabricated back in 1998. In the early efforts, an individual CNT was placed on top of two metal electrodes on a thick (100-200 nm) SiO₂ film and the heavily doped Si substrate itself was used as a back gate. While functional, these devices had low performance, primarily because of high series contact resistances (≥1 MΩ) resulting from the weak van der Waals bonding between the CNT channel and the source/drain metal electrodes. A more intimate metal-CNT contact was achieved by depositing the metal electrodes on top of the CNT and subsequently annealing the contacts. This approach increased the drive currents and transconductance by several orders of magnitude and produced switching I_on/I_off ratios of 10⁶. Fig. 3 shows the output characteristics of such a back-gated CNT-FET.

CNT-FETs fabricated using thick gate dielectrics (e.g. 100 nm SiO₂) and medium-to-high work function metals (e.g. Ti and Pd) are mostly p-type in air. Annealing such a p-type transistor in vacuum can convert the device to ambipolar and even further to n-type. This was one of the first experimental indications that Schottky barriers (SBs) at the metal-CNT interfaces dominate the transistor switching.

Unlike conventional bulk-switched Si devices, whose source and drain contacts are engineered to be ohmic, switching in CNT-FETs may arise in the bulk or at the contacts, depending on the CNT diameter, nature and geometry of metal electrodes, gate dielectrics, and device geometry. In general, charge transfer at the metal-CNT interface leads to the generation of a SB. Experimental and theoretical work has shown that SBs are much thinner in one dimension than those in three dimensions, and carrier tunneling through the SBs becomes important. Adsorbed species on the electrodes change the surface dipole, i.e. the local work function, thereby modifying the energy level line-up at the metal-CNT contacts. Other types of experiments, such as the observation of large potential drops at the source and drain contacts by scanning gate microscopy, have provided further evidence for the importance of SBs.

In a CNT-FET, Fermi-level pinning at the metal-nanotube interface is weak. To first order, the magnitudes of SBs at metal-CNT interfaces are determined by the band gap of the CNT material (~1/d) and the local work function of the contacting metal. The larger the CNT diameter, the smaller the SB will be. Because of the involvement of tunneling and thermionic emission in the injection of carriers at the contacts, the dependence of the on-current of the transistor on CNT diameter is very strong (exponential), as shown in Fig. 4.

It is this strong dependence that necessitates the precise control of the CNT diameter in electronics. So far, there have been no direct
experimental correlations between CNT-FET function and CNT chirality. However, for CNTs with diameters within the desirable range for electronic applications, i.e. \( d \geq 1.8 \) nm, Fig. 4 suggests that chirality does not play a significant role.

For hole transport, the higher the metal work function, the smaller the SB (see Fig. 4). Thus, a small diameter CNT contacted with low work function electrodes, e.g. Al, has large SBs for \( \rho \)-type transport. On the other hand, a large diameter CNT \( (d > 2 \) nm) contacted by certain high work function electrodes, e.g. Pd, can produce nearly ohmic contacts in \( \rho \)-type transistors\(^{45}\).

In addition to the work function of the metal, other factors, such as the adhesion (wetting) of the CNT by the metal, play important roles in the transport properties of CNT-metal contacts. For example, the work function of Pt is similar to that of Pd, but the on-current for Pt CNT-FETs is significantly larger than that of Pd CNT-FETs\(^{45}\).

The width of the SB depends critically on the electrostatic environment (screening). By reducing the gate oxide thickness and/or increasing its dielectric constant \( k \), the screening length (the distance over which the metal-CNT contact fields penetrate into the CNT channel) decreases, reducing the SB width and increasing its transparency\(^{41}\). Indeed, thin, high-k materials (\( \)HfO\(_2\), ZrO\(_2\)) have been used as gate dielectrics to improve device performance\(^{42,46}\).

Unlike the surface of Si, the bonds on a CNT surface are satisfied and there are no dangling bonds that could form interface traps. A very efficient coupling of the gate to the channel can be accomplished by using an electrolyte solution as the gate\(^{47}\). The high dielectric constant of the electrolyte (~80) and its ultrathin (~0.5 nm) Helmholtz...
layer lead to a very high transconductance, but also to a slow switching rate.

SBs in a CNT-FET impact the characteristics of both the on and off states of the transistor. When the gate voltage $V_{gs}$ is much higher than the threshold voltage $V_{th}$, the CNT-FET is turned on and its drive current increases with increasing CNT diameter: $\log_{10} I_d \propto -\frac{E_g}{d_{CNT}}$ (Fig. 4)\(^{44}\). In addition to the bandgap dependence on $d_{CNT}$, electron and hole effective masses $m^*$, decrease with increasing $d_{CNT}$. Tunneling through the SBs will be more efficient as a result of the lower $m^*$, leading to higher drive currents. This is consistent with reports of higher on-currents for CNT-FETs based on large diameter (2-4 nm) CNTs, usually produced by chemical vapor deposition techniques\(^{45}\), rather than smaller diameter (0.8-1.5 nm) CNTs, produced by high-pressure CO conversion (HiPCO) or laser ablation techniques\(^{36}\).

When $V_{gs} < V_{th}$, the transistor is in its off state. An important parameter in this regime is the inverse subthreshold slope $S$ that measures the how efficiently the gate controls conduction though the channel. It is defined as $S = \frac{d\log_{10} I_d}{dV_{gs}}^{-1}$. In a transistor with ohmic source and drain contacts (as in a conventional Si MOSFET), $S$ is limited by thermionic emission over the channel and is $-k_BT/q$. Thus, at room temperature, its limiting value is 60 mV/dec. In a transistor with SBs dominating the transport, $S$ is generally higher and depends on the electrostatics of the device, including the thickness and dielectric constant of the gate dielectric, geometry of the electrodes, and trapped charge around the contacts. The early back-gated, bottom-contact CNT-FETs with thick gate oxides had $S$ as high as 1 V/dec. Currently, devices with thin oxides ($<10$ nm) have $S$ in the range of 100-150 mV/dec. Theoretical modeling has also predicted improvements in transistor switching by using thinner and higher dielectric constant gate insulators and smaller contacts\(^{40-42}\).

It is not only the on-current of the transistor that is important, the current in the off state is equally important. It is critical to maintain a low leakage current to keep both the passive power at a minimum and a reasonable $I_{off}/I_{on}$ ratio ($>10^4$ is typically desirable in logic applications). However, the physics that leads to efficient gate switching (e.g. thinner gate insulators) and high on-currents in small bandgap CNTs also tends to increase the off-current.

In a SB-CNT-FET, the injection of carriers is governed by the transparencies of the contacts (dependent on the local electrical fields). Thus, the injection of holes is controlled by the gate field and electrons by the drain field. When one of the SBs is much smaller than the other (e.g. the source contact), transport through the device is dominated by the carriers (holes) injected from that contact. Transport under this condition involves only one type of carrier: it is unipolar. When the gate and drain fields become comparable (the two SBs are comparable), it is then possible to inject both electrons and holes through the two opposite contacts of the transistor. Transport in this regime involves two types of carriers: it is ambipolar. CNT-FETs with thin gate dielectrics and/or small bandgap CNTs under high drain bias tend to become ambipolar. $I-V$ characteristics of such an ambipolar device are shown in Fig. 5\(^{48}\). The drain current is a leakage current that increases exponentially with increasing $V_{ds}$, thus reducing the $I_{off}/I_{on}$ ratio\(^{49}\). This problem can be solved by electrostatically controlling the source and drain contacts separately with multiple gates or introducing charge-transfer doping to obtain different injection rates for electrons and holes\(^{48,50}\).

Most experiments on CNT-FETs have been on back-gated devices because of the simplicity of their fabrication. To build integrated circuits, however, it is vital to be able to gate each CNT-FET independently. Wind et al.\(^{51}\) fabricated the first independently top-gated, thin gate oxide, high-performance CNT-FETs. These CNT-FETs provide the first evidence that CNT-FETs could compete with state-of-the-art Si MOSFETs. Later, Lin et al.\(^{48}\) demonstrated a double-gate CNT-FET structure that allows the independent control of the switching of the SBs and of the CNT channel, and successfully converts ambipolar CNT-FETs to unipolar ones. These double-gate CNT-FETs

---

**Fig. 5** Transfer characteristics of a large diameter CNT-FET (~1.8 nm) at different drain voltages.
produces \( S \) values very close to the thermal limit. Subsequently, by using the same CNT-FETs and filtering the electron energy distribution through band-to-band tunneling, \( S \) values below the thermal limit can be achieved (~40 meV/dec at 300 K). CNT-FETs with very long channels (many micrometers) operating in the diffusive regime are dominated by the bulk barrier and are, therefore, bulk-switched like MOSFETs.

## Doping of nanotubes

For complementary metal-oxide-semiconductor (CMOS)-type applications, one would like to have both p- and n-type CNT-FETs. However, semiconducting CNTs cannot be doped using the traditional approach for bulk semiconductors of ion implantation. Furthermore, especially in small-diameter CNTs, carbon substitution by a dopant such as B or N induces high strain and leads to a defective CNT lattice. For this reason, charge-transfer doping has been employed. Charge-transfer dopants are used primarily for two reasons in CNTs: (a) to convert ambipolar devices to p-type, and (b) to convert p-type devices to n-type by, for example, depositing electron donors on them such as K atoms or amine-containing molecules. Figs. 6a and 6b show examples of both p- and n-type CNT doping. Apart from chemical doping, electrostatic doping through multiple gates has been used successfully to suppress SBs and produce bulk-switched p-i-p and n-i-n structures.

## Integrated nanotube electronic circuits

Following the optimization of individual transistors, the obvious next step is to attempt to integrate them into logic circuits. The first CNT logic gate, a NOT gate, was demonstrated by Derycke et al. In this case, a single CNT was patterned into a p-type and an n-type (through K-doping) FET pair to form a voltage inverter. This work was followed by demonstrations of a number of different logic gates, usually built with FETs involving separate CNTs.

More recently, a more complex and sophisticated circuit, a five-stage ring-oscillator (RO), was built on a single, long CNT. In this RO circuit, five pairs of p- and n-type CNT-FETs in CMOS configurations are wired along the length of an ambipolar CNT. Pd metal is used for all source and drain electrodes. Since the threshold voltages for the p- and n-branches of the ambipolar CNT are quite different, two different metals are used to construct the top gates: Pd for the p-type and Al for the n-type transistors. The work function difference of the two metals effectively shifts the FET thresholds and allows the formation of the CMOS structure. This novel approach eliminates the need for doping altogether. As gate insulator, a high-\( k \) material, AlO\(_x\), was used so as to improve the coupling of the CNT and the gate. In addition to the five-stages, two more inverters are added, one at the beginning of the chain to determine the ideal operating conditions and one at the end to avoid interference with the measurement setup (spectrum analyzer). The complete circuit is shown in Fig. 7.

With this first circuit, an oscillation frequency of up to ~70 MHz can be achieved (at \( V_{DD} = 1 \) V), which corresponds to a delay of 1.4 ns per stage. Although this frequency is much higher (~10\(^5\) times) than those achieved by oscillators based on different CNTs, it is still relatively low. However, this frequency does not represent an inherent limitation of the CNTs, which is expected to be in the terahertz range, but is determined by the parasitics of the circuit that can be eliminated by improvements in fabrication.

## Self-assembly of nanotube devices

In addition to their superior electrical properties, CNTs may offer the possibility of simpler, less expensive device fabrication. As a molecular entity, CNTs should be able to self-assemble into desirable structures. For the most part, two different strategies have been pursued: (a) prepatterning of the substrate with chemical species to which CNTs...
tend to adhere or be repulsed by; and (b) functionalizing the CNTs themselves with groups that selectively adsorb to a particular structure on the substrate. Assembly of CNTs on insulators, particularly high-

$k$ materials such as the basic oxides HfO$_2$ and Al$_2$O$_3$, is of particular interest in electronics. Acidic groups, such as phosphonic (-PO$_2$OH) or hydroxamic (-NHOH) groups, attach rather strongly to these oxides. Thus, selective adsorption can be achieved by patterning such an oxide surface, for example by imprint lithography, with molecules of the type $L-R-PO_2$OH, where $L$ is a group with affinity for CNTs, e.g. a -NH$_2$ group, or a nonadhering group, e.g. -CH$_3$.

A typical example of CNT assembly through functionalization is provided by DNA wrapping. Functionalizing the CNTs themselves could provide even better control of their placement. However, the functionalization must be reversible so that it does not degrade the excellent intrinsic electrical properties of the CNT in the final device. Covalent functionalization tends to destroy these properties by converting the sp$^2$ hybridization of the C atoms to sp$^3$. Recently, an approach was proposed that bypasses this problem. The CNTs are reacted with derivatized diazonium salts of the type $X-N_2^+-R-L$, to give covalently bonded derivatives of the type CNT-$R-L$, which as discussed above, adhere to the appropriate oxide structures through the group $L$. In this case, however, the CNT-$R$ bond can be cleaved cleanly after deposition by thermal annealing, and the electrical properties of the pristine CNT can be recovered.

**Optoelectronic devices**

Electron and hole carriers in semiconductors can recombine by a variety of different mechanisms. In most cases, the energy will be released as heat (phonons), but a fraction of the recombination events may involve the emission of a photon. This process is termed electroluminescence and is widely used to produce solid-state light sources such as light-emitting diodes (LEDs).

In order to fabricate LEDs or any other electroluminescent device, one must recombine significant populations of electrons and holes. Conventionally, this is achieved at an interface between a hole-doped and an electron-doped material (e.g. a p-n junction). In ambipolar CNT-FETs at an appropriate bias, however, electrons and holes can be simultaneously injected at the opposite ends of the CNT channel. This allows electroluminescence to occur. While the emission mechanism is the same as that in p-n junctions, ambipolar CNT-FETs do not require chemical doping.

CNT electroluminescence exhibits a variety of interesting properties. The emitted light is strongly polarized along the tube axis. The radiation also has a characteristic energy that depends on the diameter and chirality of the excited CNT, just as the optical bandgap does, and the length of the electroluminescent region is $l_{rec} \leq 1 \mu m$. In short devices ($L < l_{rec}$, where $L$ is the channel length), the light emission encompasses the entire CNT. In long devices ($L >> l_{rec}$), on the other hand, the emission will be localized wherever the concentrations of electrons and holes overlap most strongly. This overlap region can be physically moved using a gate electrode, since the relative contributions of electrons and holes to the total current is strongly gate dependent. Therefore, a CNT-LED is a movable light source. The gate bias $V_{g}$ can smoothly and continuously position the site of emission. In Fig. 8, we demonstrate the translation of the emission spot between two electrodes by applying different gate voltages.
In addition to this translatable emission, localized electroluminescence is also observed from particular spots on a CNT under unipolar transport conditions. In this case, the current is carried by only one type of carrier (electrons or holes). Since both types of carriers are necessary to produce light, these sites must actively generate e-h pairs. This process occurs near defects, trapped charges in the insulator, or any other inhomogeneities that produce voltage drops along the CNT and generate large, local electric fields. The resulting ‘hot’ carriers produce e-h pairs through a most efficient intra-nanotube impact excitation process. The efficiency of this process can be traced in the quasi-one-dimensional confinement and the weak screening of the Coulombic interactions. The monitoring of localized electroluminescence provides a new tool for detecting defects in CNT devices. Artificial structures can also be fabricated that create the conditions locally, i.e. sudden change in the potential, that generate e-h pairs and light emission. Such a device is shown in Fig. 9b. Unlike the ambipolar device emission, the light intensity of the unipolar devices depends exponentially on the current, a fact that supports the impact mechanism.

Photoconductivity is the reverse of electroluminescence, with optical radiation producing electron and hole carriers. An example of a photoconductivity measurement is shown in Fig. 10. The resonant excitation of a CNT generates an electric current, which can be used as a nanosized photodetector or as a spectroscopic tool. Alternatively, in the open-circuit configuration, the device generates a photovoltage. Thus, a CNT-FET device can be used as a transistor.
light emitter, or light detector. Choosing among these different modes of operation only requires the bias conditions to be changed.

**Conclusions**

It is clear that the unique properties of CNTs make them excellent candidates for nanoelectronics and photonics, and the devices already demonstrated prove this point. However, as is true with any new technology, there are numerous technical and other types of problems that need to be resolved before a competitive CNT-based technology can be developed. It is very unlikely that highly developed Si technology will be replaced by CNTs or, for that matter, by any other type of technology anytime soon. The speed at which CNT technology evolves will depend strongly on breakthroughs, such as the development of novel integration processes. CNTs also provide the opportunity for a wealth of new types of applications, e.g. in bioelectronics, and most importantly for new, low-cost methods of fabrication. Guided self-assembly techniques could be used to fabricate circuits at a low level of integration, but with very high performance characteristics, for example in telecommunications, without the multibillion dollar costs of current facilities.

**REFERENCES**

10. Avouris, Ph., MRS Bull. (2004) 29, 403
78. Perebeinos, V., and Avouris, Ph., unpublished results