Principles of Metallic Field Effect Transistor (METFET)

Slava V. Rotkin, Karl Hess

Beckman Institute for Advanced Science and Technology,
University of Illinois at Urbana–Champaign,
405 N.Mathews, Urbana, IL 61801, USA;
Email: rotkin@uiuc.edu
Fax: (217) 244–4333

ABSTRACT

Novel type of a field effect transistor (FET) is described. A metallic channel of a metallic nanotube FET is proposed to be switched ON/OFF by applying electric fields of a local gate. Very inhomogeneous electric fields may lower the nanotube symmetry and open a band gap, as shown by tight–binding calculations.

Keywords: metal FET, nanotube, molecular devices, theory, symmetry breaking

1 INTRODUCTION

Field effect transistors in current use are mostly semiconductor devices [1]. While the scaling trend to ever smaller dimensions calls for ever higher doping and channel conductance of these devices [2], the doping level is limited by the nature of a metallic state. A high conductivity in bulk metallic systems inevitably prevents the penetration of the electric field except for extremely short distances, too short to achieve device function.

In this paper we put forward a concept for a novel transistor with a metallic channel. We model a new mechanism to control electronic transport in metallic one dimensional (1D) systems by use of the inhomogeneous electric field induced by a highly localized gate such as an STM tip or a nanotube tip [3]. Also, ultra narrow leads [4], fabricated closely to the nanotube channel, special chemical function groups [5] at the tube sidewalls or inside the tube may be used as a local gate. We notice that use of a dual gate (both local and backgate) may be beneficial because of the uniform backgate controls the Fermi level (charge density) while the local gate controls the conductivity and switches OFF/ON the channel. The weak screening of electric potentials in 1D channels enhances effect of the electric fields. In a subsequent article [6] we discuss that the highly localized gate induces a high electric field in a narrow region and impedes the electron transport.

The key effect which is employed for metallic field effect transistors (METFET's) is the symmetry breaking. The electronic structure of an armchair metallic single wall nanotube (M–SWNT) is robust to perturbations that have a *long range* potential. In contrast, inhomogeneous electric fields have *short range* and may lead to the opening of a band gap in carbon nanotubes due to lowering of the symmetry of the band structure. The possibility to open a band gap by use of inhomogeneous electric fields and the high (ballistic) conductance of the armchair M–SWNTs make them a suitable material to design METFET's.

In this paper we focus on the mechanisms for the band gap opening and discuss the physics of this effect. (For more details on applications of the effect in a METFET we refer a reader to Ref. [6].) We use a tight-binding approximation (TBA) to study the bandstructure of the metallic nanotube. It is well known that for the SWNT which has a moderate curvature the one–band TBA captures most of features of the band structure. Thus we include in the calculation only one orbit per carbon atom and use one–parameter TB Hamiltonian with a hopping integral $\gamma \simeq 2.7$ eV.

We propose two main mechanisms for the local band gap opening in the SWNT: (i) the non-linear Stark effect and (ii) the band-structure modulation by a very non-uniform (multipole) electric field. We notice that other possible mechanisms are thought which may include a combination of an electric field and other perturbation potential, such as molecular/atomic environment (tube wrapping or decoration, for example), deformation or NT–lattice effects. The latter has been already demonstrated to be an effective mechanism for the gap opening in the ropes of armchair SWNTs [7].

The action of the transistor device is based on the gap opening in the following way: the gated region of the channel (with the semiconductor band gap at non zero gate voltage) represents a barrier for charge carriers. The barrier height and length can be modulated by the gate voltage. The transport through the gated region is due to (1) tunnelling and (2) thermionic emission at non zero temperature. We consider examples of typical armchair tubes: [10,10] SWNT of 1.4 nm diameter and [5,5] SWNT of 0.7 nm diameter. The gap is larger for smaller tubes as we will show below.

2 ARMCHAIR NANOTUBES

Use of metallic SWNT's to realize the possibility of 1D METFET's is appealing not only because of the extremely small size of the tube, but also for other rea-

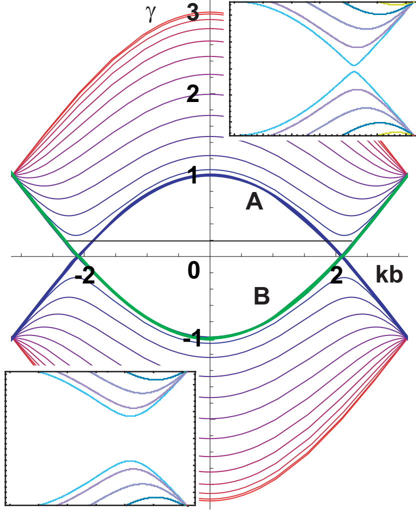

Figure 1: Bandstructure of armchair [10,10] M–SWNT. Upper right Inset shows zoom view of the Fermi point with opening of a gap due to *high–multipole* perturbation as described in the text. Lower left Inset shows how this gap grows linearly with applied potential.

sons, for example high (ballistic) conductance. IBM researchers have enunciated the vision of combining metallic and semiconducting nanotubes (M–SWNT's and S–SWNT's) in circuits [8–11]. The S–SWNT's would serve as active devices and the M–SWNT's as interconnects (also proposed by NASA group for conventional devices [12]) . Here we add a possibility to also use the *metallic tubes* as transistors which would be advantageous for various reasons including their high current capacity and conductance. The METFET will add to the predominantly p-type S–SWNT's an ambipolar device without using extra doping and/or complicated control on a contact work function. The METFET could be combined to circuits in analogy to CMOS.

Armchair M–SWNT's have also the advantage of a low phonon scattering-rate [13] and an equally low rate of scattering by impurities with a long range potential [14,15]. This follows from a special symmetry (selection rules) applicable to the transitions between the closest subbands of the armchair SWNT's.

The wave-function of the electron in a conduction band can be written as a product of an envelope function and an amplitude in a unit cell. Each unit cell consists of two atoms (of different sublattices). Thus, the amplitude has two components, which are often referred as components of a pseudospinor in the nanotube literature. The pseudospin has the opposite sign, ±1, for states of the valence and conduction band. The corresponding TB wave function is given by:

$$\langle\vec{r}|\psi_{m,k,\zeta}\rangle = \frac{e^{ikz}e^{im\alpha}}{\sqrt{2\pi L}}\frac{\langle\vec{r}-\vec{R}_I|\varphi\rangle + \zeta c_{mk}\langle\vec{r}-\vec{R}_{II}|\varphi\rangle}{\sqrt{2}}, \quad (1)$$

here c_{mk} is to be found as an eigenvector of the TB Hamiltonian; the index m stays for an angular momentum of the electron and labels orbital subbands; k is a longitudinal momentum; $\zeta = \pm 1$ is the pseudospin. The components of the pseudospinor vector are atomic-like wave functions, defined in the unit cell, $\varphi_{I/II}$. The electron is considered at the surface of a cylinder. Its coordinates $(\vec{r})$ are z, and α, along the tube axis and along the tube circumference respectively. Here R is the nanotube radius and L is the length of the tube.

The two (closest) crossing subbands are shown in Fig.1 as bold green and blue curves. Other subbands of orbital quantization are shown in color from red to blue (from top to bottom in conduction band) and correspond to $m = 0, 1 \ldots 9$. Two closest subbands, A and B (blue and green) have $m = 10$, but the pseudospin has opposite sign. Hence, the pseudospinor overlap integrals of the states of these subbands are zero [16].

A common approximation of solid state theory is, that the spatial variation of any long range external potential is small at the scale of a unit cell (effective mass theorem). Thus, the matrix element of the potential is a product of the unit cell overlap integral and the Fourier component of the potential: $\langle\psi|U|\psi\rangle = \langle\varphi|\varphi\rangle U_q$. It is evident that this approximation leads to a zero mixing/scattering between two closest subbands of the armchair tube as $\langle\varphi|\varphi\rangle \equiv 0$ and is suggestive for ballistic transport in the armchair SWNT channel [17]. Thus, the ideal armchair SWNT is a great candidate for a 1D METFET as its conductance in the ON state has to be about $2G_o = 2e^2/h$, 4 times of the conductance quantum (for 2 spin and 2 space channels), maximum conductance for a circuit with macroscopic leads [18].

Group theory proves that the direct matrix element is non–zero if the Fourier transform of the potential $U_{q,\delta m}$ is a full scalar with respect to the rotations of the symmetry group of the nanotube [16]. This is fulfilled for $\delta m = s \cdot n$, where $s = 2, 4, \ldots$ is an positive even integer and n is the integer appearing in the notation [n,n] for armchair SWNT's [19]. Then, the energy dispersion of two new subbands $|A \pm B\rangle$ in vicinity of the Fermi point reads as: $E_\pm = \pm\sqrt{E_A(q)^2 + U_{q,\delta m}^2}$. Here $E_A(q) = -E_B(q) = E_B(-q)$ is the dispersion in A/B subband at zero gate voltage. At the Fermi point $E_A(0) = 0$ and, thus, a gap opens:

$$E_g^{(sn)} = 2|U_{q,sn}| \quad (2)$$

This is a *direct mixing* of the subbands A and B. The band gap is linear in the applied gate potential (see Inset in Figure 1). There is no upper limit for the magnitude of the opened gap in this case (except for a natural condition that the field must not cause an electric breakdown).

The direct mixing of the crossing subbands is tricky as one needs to break the mirror reflection symmetry

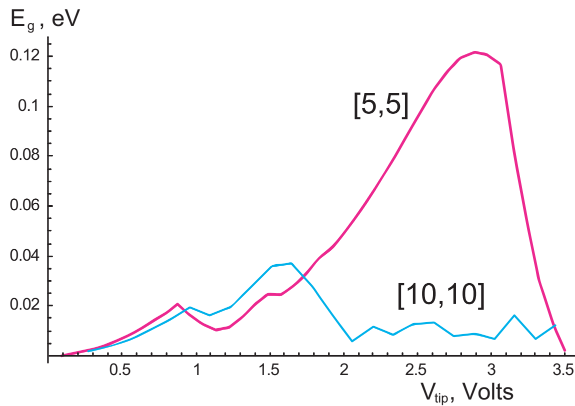

Figure 2: Band gaps in [10,10] and [5,5] SWNTs for *dipole + quadrupole* potential as a function of applied voltage.

along the tube axis *and* keep the full symmetry of the applied potential which requires very high multipole moment of the potential $\delta m = 2n$. Although such potential may be possibly created by applying the gate voltage to a chemically modified/decorated surface of the SWNT (to be discussed elsewhere), this is a challenging problem for future technology.

However, if only first condition is fulfilled (mirror symmetry breaking), the gap opens by *indirect* mixing of the subbands in a higher order of applied potential. In particular, by applying a uniform field (dipole component of the potential) across the armchair SWNT *together* with a higher order (quadrupole) component we obtained the gap which scales as a third power of the gate potential. Perturbation theory for nanotubes (similar to what was used in [20]) predicts a maximum gap which depends only on the size of the tube and scales as R^{-2} or n^{-2} (see below).

To prove the gap opening due to this mechanism, the non–linear Stark mechanism, we consider a number of armchair SWNTs in the fields with two components: the dipole and the quadrupole. A simplest realization of such potential can be given by two tips at the top and the bottom of the SWNT, which are kept at the potentials $+2V_{tip}$ and $-V_{tip}$. (Other potentials were simulated as well and led to similar results.) This potential is also mimic to the potential of a local gate over the SWNT on the insulating substrate with a dielectric permittivity ϵ. The image charge in the substrate is $(\epsilon - 1)/(\epsilon + 1)$ fraction of the tip charge and has the opposite sign. The results of the calculation for [10,10] and [5,5] SWNTs are presented in Fig.2.

3 SWITCHING OFF A QUASI-METALLIC NANOTUBE

While opening the energy gap in the metallic armchair SWNT requires a special symmetry of the gate potential, the quasi–metallic tubes are best candidates for band gap engineering. As we already noticed in Ref. [20], the uniform external electric field applied across a zigzag quasi–metallic nanotube opens the gap. This is just due to the Stark effect and, therefore, the gap is proportional to an interband transition overlap integral $\langle C|V\rangle$ and to the square of the potential, $|V|$ (cf. the third power for non–linear Stark effect in armchair tubes). Though, similarly to the case of the armchair tube, this gap will close eventually with increasing external potential. Thus, there is a maximum gap opening which is defined solely by the tube radius $E_g^{(c)} \sim R^{-2}$.

The closing of the opened gap happens with increasing the gate potential when its value is about the distance between other subbands. Thus, the critical potential is $\sim \hbar v_F/R$, where $v_F = 3\gamma b/2\hbar$ is the Fermi velocity in a M–SWNT, a bond length is $b \simeq 0.14$ nm.

For the armchair SWNT we go in a high order of the perturbation theory and at the critical potential the matrix element of the perturbation cancels with the energy level separation. As a result the maximum gap opening is a universal function for armchair and zigzag nanotubes:

$$E_g^{(c)} \sim \langle C|V\rangle |V^{(c)}| \sim \frac{\hbar v_F b}{2R^2} \propto \frac{\gamma}{n^2}, \qquad (3)$$

and the maximum gap $E_g^{(c)}$ is proportional to the square of the nanotube curvature $(1/R)^2$ (Fig.2) as we also confirmed by a numerical calculation in [20].

Even though the maximum gap due to the Stark effect is small for a quasi–metallic nanotube and thus can be observed only at low temperature, this gap opening might be easier to achieve experimentally, *e.g.*, by use of the inhomogeneous electric field of a tunnelling tip, because of no special symmetry between closest subbands as compared to the case of direct subband mixing by a high–multipole potential. However, the non–armchair SWNT must also have a lower ON conductance due to the same argument. A stronger scattering of carriers happens in this tube, similarly to a S–SWNT [15].

4 FIELD EFFECT MODULATION OF METALLIC CONDUCTANCE

Any opening of the gate induced semiconductor gap along the metallic tube will create a potential barrier for the electrons and, therefore, modulate the conduction. Figure 3 shows typical IV curves (IVC) for a METFET using an armchair metallic nanotube of the radius 0.7 nm ([10,10] SWNT) at T =4K, gate width $W = 15$ nm. The upper (red) IVC corresponds to a zero gate voltage (no gap). The channel is fully open and the ON current is determined by injection from a contact and thus by the quantum conductance $4G_o$ [18]. With increasing gate voltage one observes a substantial decrease of the

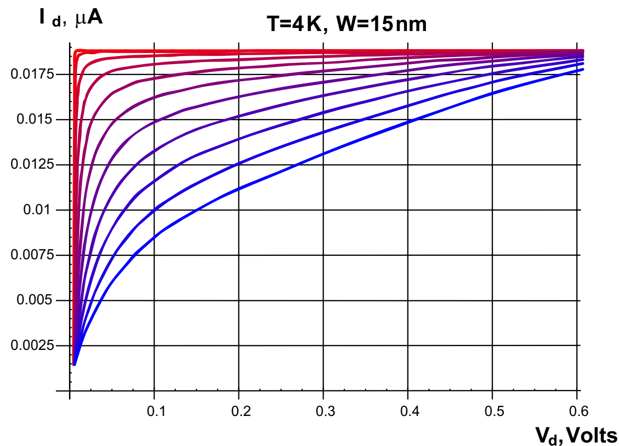

Figure 3: Current of a METFET made of an armchair [10,10] SWNT versus the drain potential at different gate voltages (at different gaps). The gate width is 15 nm. $T = 4$K. Curves from top to bottom correspond to increasing gap from 0 to 0.4 eV (in color: from red to blue).

current due to the opening of the gap and a depletion of electrons in this region. The IVC's correspond to gaps up to 0.4 eV (blue curve).

In conclusion, we put forward a novel type of electronic switching device based on carbon nanotubes: a one–dimensional metallic field effect transistor (1D METFET). We have described band gap opening mechanisms in the armchair and zigzag metallic SWNT's. Our calculations demonstrate, at least principle, the possibility to open a band gap by application of inhomogeneous electric fields that may be created by special local gates to break the SWNT symmetry. Two mechanisms of the symmetry breaking and subsequent gap opening in an armchair SWNT are outlined: a subband mixing by a high–multipole potential and a non–linear Stark effect. A quasi–metallic zigzag SWNT can be also used for the METFET. In that case the proposed gap opening mechanism is the Stark effect (in a low order of a perturbation theory). Assuming virtually ballistic transport, the IV curves for the METFET's are obtained.

ACKNOWLEDGMENTS

Authors acknowledge Ms. Yan Li for help in numerical computations. This work was supported by NSF Grant No. 9809520, by the Office of Naval Research grant NO0014-98-1-0604 and the Army Research Office grant DAAG55-09-1-0306. S.V.R. acknowledges partial support of DoE Grant No. DE-FG02-01ER45932, and NSF grant No. ECS–0210495.

REFERENCES

[1] Advanced Theory Of Semiconductor Devices// Karl Hess. New York: IEEE Press, 2000.

[2] International Technology Roadmap for Semiconductors (Semiconductor Industry Association, San Jose, CA, 2001); http://public.itrs.net/.

[3] http://www.itg.uiuc.edu/exhibits/iotw/2003-09-09/

[4] J.-F. Lin, J.P. Bird, L. Rotkina, P.A. Bennett, Appl. Phys. Lett. **82** (5), 802, 2003.

[5] S.V. Rotkin, I.Zharov, Int. Journal of Nanoscience **1** (3/4), 347–356, 2002.

[6] S.V.Rotkin, K.Hess, Appl.Phys.Lett. (submitted).

[7] P. Delaney, H.J.Choi, J.Ihm, S.G.Louie, M.L.Cohen, Nature, **391**, 466, 1998

[8] J. Appenzeller, Device Research Conference, Utah, June 26, 2003

[9] J. Appenzeller, R. Martel, V. Derycke, M. Radosavljevic, S. Wind, D. Neumayer, and Ph. Avouris, Microelectronic Eng., **64**, 391, 2002.

[10] Ph.G. Collins, Ph. Avouris, Sci.Am., 12, 62, 2002. Ph. Avouris, Chemical Physics, **281**, 429, 2002.

[11] S. Heinze, J. Tersoff, R. Martel, V. Derycke, J. Appenzeller, Ph. Avouris, Phys.Rev.Lett. **89**, 106801, 2002.

[12] J. Li, Q. Ye, A. Cassell, H. T. Ng, R. Stevens, J. Han, and M. Meyyappan, Appl.Phys.Lett. 82, 2491 (2003).

[13] H. Suzuura and T. Ando, Phys. Rev. **B 65**, 235412, 2002.

[14] T. Ando and T. Nakanishi, J. Phys. Soc. Jpn. **67**, 1704, 1998. T. Ando, T. Nakanishi and R. Saito, J. Phys. Soc. Japan **67**, 2857, 1998.

[15] A.G.Petrov, S.V.Rotkin, submitted.

[16] T. Vukovic, I. Milosevic, M. Damnjanovic, Phys.Rev. B 65(04), 5418, 2002.

[17] Although, there is a possibility of scattering to the other valley of the same subband, this scattering amplitude is small due to a large momentum transfer $q \simeq 2\pi/\sqrt{3}b$ (see thin line between left and right Fermi points in Fig.1), where $b \simeq 0.14$ nm is a carbon–carbon bond length.

[18] For a circuit with macroscopic leads to the M–SWNT channel the total conductance will be about $4G_o = 2e^2/h$. This gives a minimum resistance, a contact resistance, of the SWNT device $\sim 6.5k\Omega$. The lower resistance can be expected in the case of entirely nanotube circuit [8]. The quantum contact resistance will not limit anymore the ON current in this case. That device can fully exploit all advantages of the METFET.

[19] Y.Li, S.V.Rotkin, U.Ravaioli, and K.Hess, unpublished.

[20] Y.Li, S.V.Rotkin and U.Ravaioli, Nano Letters, vol. 3(2), 183–187, 2003.