Conductance Asymmetry of Graphene p-n Junction

Tony Low, Seokmin Hong, Joerg Appenzeller, *Senior Member, IEEE*, Supriyo Datta, *Fellow, IEEE*, and Mark S. Lundstrom, *Fellow, IEEE*

***Abstract*—We use the nonequilibrium Green function method in the ballistic limit to provide a quantitative description of the conductance of graphene p-n junctions—an important building block for graphene electronics devices. In this paper, recent experiments on graphene junctions are explained by a ballistic transport model, but only if the finite junction transition width D_w is accounted for. In particular, the experimentally observed anomalous increase in the resistance asymmetry between n-n and n-p junctions under low source/drain charge density conditions is also quantitatively captured by our model. In light of the requirement for sharp junctions in applications such as electron focusing, we also examine the p-n junction conductance in the regime where D_w is small and find that wave-function mismatch (so-called pseudospin) plays a major role in sharp p-n junctions.**

***Index Terms*—Conductance asymmetry, graphene, nonequilibrium green function, p-n junction.**

I. Introduction

GRAPHENE is a 2-D sheet of carbon atoms arranged in a honeycomb lattice with unique electronic properties due to its linear energy dispersion, with zero bandgap, and a spinor-like two-component wave function [1]–[3]. These characteristics give rise to interesting transport phenomena such as the absence of backscattering [4], anomalies in the quantum Hall regime [2], [5], weak antilocalization [6], [7], so-called Klein tunneling [8], and electron focusing analogous to optical effects that occur in negative refractive index materials [9]. As such, one expects that graphene p-n junctions should differ from traditional semiconductor p-n junctions. The p-n junction is a basic building block for electronic devices. Developing a quantitative understanding of graphene p-n junctions is an important step on the way to realizing novel devices such as graphene lenses [9], [10] and filters [8]. Our goal in this paper is to demonstrate that the nonequilibrium Green function (NEGF) approach [11] quantitatively explains recent experiments on graphene p-n junctions, including the critical role of the junction depletion width and the increased in the odd resistance observed under low source/drain charge density conditions [12].

Electron transmission across a graphene p-n junction occurs by interband tunneling. A theoretical treatment for an abrupt graphene p-n junction predicts that for a symmetric p-n junction (i.e., one in which the hole and electron densities on each side of the junction are equal), the transmission probability is given by $\cos^2\theta$, where θ is the angle between the electron's wave vector and the normal to the junction interface [8]. Realistic p-n junctions will have a transition region of finite width. For a smooth junction transition region of width D_w, the Wentzel Kramers Brillouin (WKB) approximation can be used to show that the transmission probability for a symmetric p-n junction is $e^{-\pi k_f D_w \sin^2(\theta)/2}$ [13], where k_f is the Fermi wave vector. Whether the transition region is abrupt or graded, the transmission is perfect when $\theta = 0$ (i.e., commonly referred to as Klein tunneling), but the transmission decreases as θ and D_w increase. This angular selectivity for electron transport across the p-n junction serves as a filter, allowing states with $|\theta| \leq \sigma_\theta$ (where σ_θ is the spread of the angular distribution) to pass through more effectively. The quantity $2\sigma_\theta$ can be viewed as the bandwidth of this filter and is what gives rise to the larger resistance of a p-n junction as compared to a uniform graphene sheet.

Several research groups have recently fabricated graphene p-n-p devices by using electrostatic gates to create p and n regions [12], [14]–[16]. The typical setup consists of a back gate and top gate, which are used to control the amount of charge density in the source/drain and channel regions, respectively. For example, the back gates can be set to produce n-type source and drain regions, and the top gate can be biased to change the middle (channel) region from n to p type. The asymmetry in the device's source-to-drain resistance as a function of top-gate voltage has been experimentally observed [12]. The amount of this resistance asymmetry is a measure of the intrinsic property of the p-n junction, provided that the mean free path of carriers is larger than the transition length of the junction. One can theoretically compute the junction's transition length accounting for nonlinear screening effects [17]. For recent experiments [12], typical transition lengths for p-n junction are less than 100 nm. Recent experiments indicate that the carrier's mean free path is about 100 nm under low temperature and moderate carrier density conditions of 10^{12} cm^{-2} [18]. In addition, there is experimental evidence of Fabry–Perot interference effects within the channel region in devices with channel lengths less than 100 nm [19], [20], evidently a signature of phase coherent transport. Therefore, a ballistic transport model is sufficient for the study of the experimental p-n junction devices reported in [12].

In this paper, we present a quantitative study of the near-equilibrium IV characteristics of graphene p-n junctions. In particular, we present a systematic study of the impact of the junction transition width D_w on the transport properties of graphene p-n junctions. We use the NEGF approach with a

Manuscript received November 19, 2008; revised January 26, 2009. First published April 28, 2009; current version published May 20, 2009. The review of this paper was arranged by Editor C.-Y. Lu.

The authors are with the Network for Computational Nanotechnology, Birck Nanotechnology Center, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47906 USA (e-mail: tonyaslow@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2009.2017646

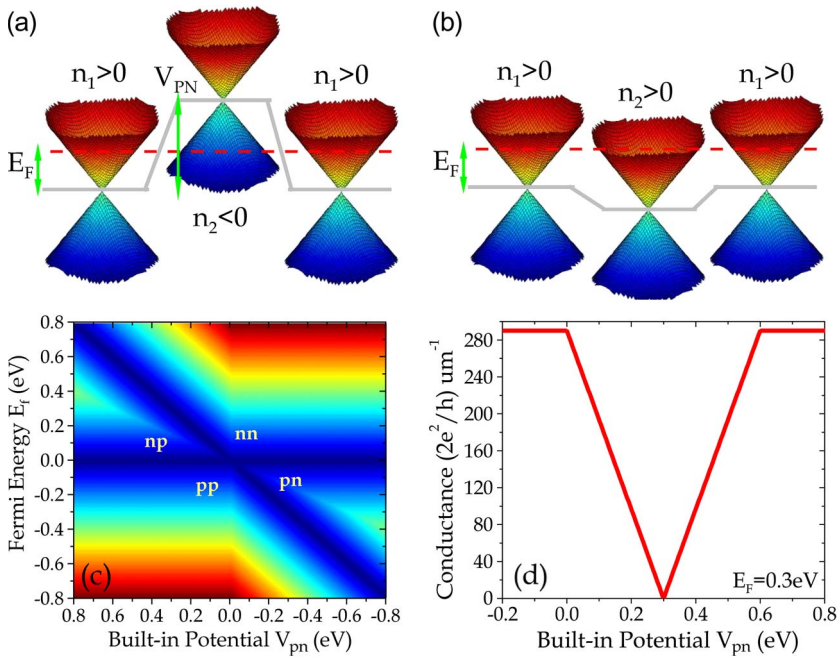

Fig. 1. Energy band diagram of graphene (a) n-p-n and (b) n-n-n devices. (c) Intensity plot of the p-n junction conductance $\sigma_{\rm pn}$ as a function of E_f and $V_{\rm pn}$ using (1), and (d) plot of $\sigma_{\rm pn}$ as a function of $V_{\rm pn}$ at $E_f = 0.3$ eV using (1).

nearest neighbor tight-binding description of graphene. This method allows us to accurately simulate the p-n junction conductance for both abrupt and graded junctions under different bias conditions. The value of D_w is not known *a priori* because it depends on charge screening and the gate potential as governed by the Poisson equation. In this paper, we employ the analytical screening model derived in [17]. Our numerical result based on the assumption of ballistic transport is in reasonable quantitative agreement with the experiments reported in [12]. As has been noted previously, we also found that the conductance follows an inverse square root dependence on D_w when D_w is large [13], but we find that strong deviations from this trend occur when D_w is sufficiently small. Understanding this regime of operation has practical importance because devices based on the electron focusing property of graphene p-n junctions, i.e., graphene lenses [9], [10] and filters [8], are expected to operate in this regime.

II. Graphene p-n Junctions and Resistance Asymmetry

Before describing the simulation method used in this paper, we define some terms that will be used in subsequent discussions, present a simple picture of the conductance of a graphene n-p-n or p-n-p junction, and identify the issues that will be addressed in the numerical study. Fig. 1(a) and (b) shows depictions of the energy band diagram for the experiment, which shows the location of the intersection of the conduction and valence bands (the so-called Dirac point) versus position. A back gate controls the location of the Fermi level E_f in the source/drain regions. For $E_f > 0$ (above the Dirac point), there is an increase of electrons with respect to $E_f = 0$, the electron density is greater than zero, $n_1 > 0$, and the material is n-type. If the back gate is biased negatively so that $E_f < 0$, then $n_1 < 0$ and the source/drain regions are p-type. Similarly, the channel (middle) region can be biased by a top gate to be either p-type $(n_2 < 0)$ or n-type $(n_2 > 0)$. The top gate also controls the built-in potential $V_{\rm pn}$ of the n-p/n-n junction. Thus, appropriate top and bottom gate voltages can produce n-p-n, n-n-n, p-n-p, or p-p-p structures. Near-equilibrium conditions are assumed so that the source and drain Fermi energies are the same. In this paper, we shall assume that the channel length is greater than the carrier's phase coherent length, allowing us to ignore wave interference effects within the channel. Therefore, we can reduce the problem by only studying the transport across a single p-n junction.

As a first step to understanding the near-equilibrium conductance of a p-n junction, i.e., $\sigma_{\rm pn}$ (S/um), as a function of $V_{\rm pn}$ and E_f conditions, we adopt a simple density-of-states argument in the Landauer picture. In this simple analysis, $\sigma_{\rm pn}$ can be written in the following form:

$$\sigma_{\rm pn} = \frac{2e^2}{Wh} \min(M_1, M_2) \tag{1}$$

where $M_{1/2}$ is the number of modes in the source/channel, respectively, and W is the device width. Equation (1) mimics the matching of transverse momentum between the source and channel in a ballistic manner, such that the current will be limited by the region with the smaller number of modes. In this simple exercise, we assume a zero temperature treatment, so that $M_{1/2}$ refers to the number of modes at their respective E_f.

Fig. 1(c) shows an intensity plot of $\sigma_{\rm pn}$ computed using (1) versus $V_{\rm pn}$ and E_f. The dark-blue lines are regions of low current. The horizontal line occurs when the source Fermi level is at the Dirac point. Since $n_1 = 0$, so no current flows. The diagonal black line described by $V_{\rm pn} = E_f$ occurs when the channel region is adjusted to place the Dirac point in the

channel at the Fermi energy of the carrier. Since $n_2 = 0$ along this diagonal line, no current flows. Very similar features are observed experimentally [12], [19]. Looking more closely, we can plot $\sigma_{\rm pn}$ versus $V_{\rm pn}$ at a fixed $E_f = 0.3$ eV. As shown in Fig. 1(d), the conductance versus $V_{\rm pn}$ is symmetrical about $V_{\rm pn} = E_f$. In experiments, an asymmetry about $V_{\rm pn} = E_f$ is observed [12], [19]. This asymmetry cannot be captured by simple density-of-states argument because its origin is quantum mechanical in nature, i.e., quantum tunneling and wave-function mismatch. Our goal in this paper is to explain the role that quantum tunneling and wave-function mismatch played in the observed asymmetry and then quantitatively explain the magnitude of the asymmetry observed in experiments.

III. Simulation Methods

In this section, we briefly outline the NEGF formalism for quantum ballistic transport and describe its application to nonequilibrium transport [11], [21] through graphene p-n junctions.[1] The central quantity in the theory is the retarded Green's function

$$G(\epsilon, k_y) = [(\epsilon + i\delta)I - H(k_y) - U - \Sigma_l(\epsilon, k_y) - \Sigma_r(\epsilon, k_y)]^{-1} \tag{2}$$

where δ is an infinitesimal quantity in the channel region but is adjusted to provide nonvanishing density-of-states at the Dirac point for the contact regions [22], [23]. The Hamiltonian is formulated by treating only the nearest neighbor interaction between the p_z orbitals [24], [25]. Assuming that the device width is large and homogeneous along the direction transverse to current flow, we can write the Hamiltonian as

$$H = \begin{bmatrix} \alpha & \beta_1 & & \\ \beta_1^\dagger & \alpha & \beta_2 & \\ & \beta_2^\dagger & \alpha & \ddots \\ & & \ddots & \ddots \end{bmatrix} \tag{3}$$

where α, β_1, and β_2 are all 2×2 matrices given by

$$\alpha = \begin{bmatrix} 0 & t_c \\ t_c & 0 \end{bmatrix} \quad \beta_1 = \begin{bmatrix} 0 & 0 \\ t_y^\dagger & 0 \end{bmatrix} \quad \beta_2 = \begin{bmatrix} 0 & 0 \\ t_y & 0 \end{bmatrix} \tag{4}$$

where t_c is the nearest neighbor orbital coupling energy and $t_y = t_c + t_c e^{ik_y a_0}$. The lattice parameter $a_0 = \sqrt{3} a_{\rm cc}$, where $a_{\rm cc} = 1.44$ Å, is the c–c bond distance. The quantum number k_y is the quantized transverse momentum, to be elaborated upon in the next paragraph. The contacts' self-energies $\Sigma_i(\epsilon, k_y)$ are obtained by solving $\Sigma_i = \tau_i g_i \tau_i^\dagger$, where g_i is the surface Green function associated with the contact. We should mention that for armchair ribbon, an analytical closed-form solution for g_i is possible, since the wave function is known, both in the tight-binding [26] and Dirac formalisms [27]. However, in our numerical treatment, we had employed a nonnegligible δ in the contact regions, which prevents us from using the analytical closed-form solution for g_i. Therefore, g_i is computed numerically with an iterative scheme proposed in [28]. Finally, the current through contact i can then be computed using

$$I_i(\epsilon) = \frac{2q}{h} \sum_{k_y} trace\left[\Sigma_i^{\rm in}(\epsilon) A(\epsilon) - \Gamma_i(\epsilon) G^n(\epsilon)\right] \tag{5}$$

where $A = i(G - G^\dagger)$ is the local density-of-states, $\Sigma_i^{\rm in} = f_0(\epsilon)\Gamma_i(\epsilon)$ is the filling function (analogous to the in-scattering function for incoherent case), $f_0(\epsilon)$ is the Fermi function of the contacts, and $\Gamma_i = i(\Sigma_i - \Sigma_i^\dagger)$ is the contact broadening factor. In (5), $G^n(\epsilon)$ is the electron correlation function given by $G(\Sigma_l^{\rm in} + \Sigma_r^{\rm in})G^\dagger$. G and G^n are computed using the recursive Green function algorithm through Dyson's equation [29] and making use of the fact that the Hamiltonian is tridiagonal in nature.

By assuming an armchair ribbon configuration, imposing a box-boundary condition, and solving the Dirac equation, Brey and Fertig showed [27] that the transverse momentum is quantized according to

$$k_y = \left(\frac{2\pi}{3a_0} + \frac{2\pi n}{2W + a_0}\right) \pm \frac{2\pi}{3a_0} \tag{6}$$

for all integer n and W is the width of the device. The last term accounts for the momentum of the Dirac points, $\vec{K}$ and $\vec{K}'$, where the upper/lower sign is used when n is even/odd, respectively.

Treating the problem in terms of transverse modes greatly reduces the computation burden while still providing accurate results when the potential in the transverse direction is uniform [30]–[32]. This approach essentially translates a 2-D real-space transport problem into m decoupled 1-D real-space transport problem, where m is the number of relevant transverse modes. In the limit of large device width W, we have $m \propto W\epsilon$. Fig. 2 compares this mode-space method with 2-D real-space NEGF calculations for a 55-layer metallic graphene ribbon and a 32-layer semiconducting ribbon. "n-layer" refers to the number of layers of carbon atoms along the width direction, where $W = na_0$. In both cases, the two methods give nearly identical

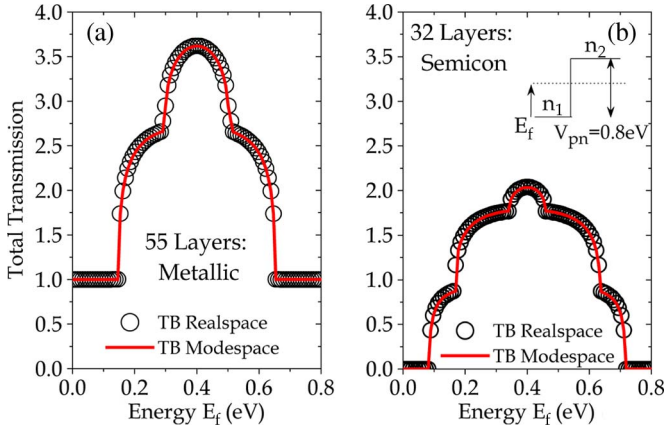

Fig. 2. Total transmission as a function of energy calculated for (a) 55-layer ($W \approx 13.7$ nm) metallic armchair ribbons and (b) 32-layer ($W \approx 8$ nm) semiconducting armchair ribbons transporting across a p-n junction with a built-in potential of $V_{\rm pn} = 0.8$ eV. In both cases, we compared the method outlined in our paper (solid line) with that of real-space tight-binding approach (open symbols).

[1]Note that we are using the NEGF method in the ballistic limit where it is equivalent to the Landauer approach, as explained in [11, ch. 8].

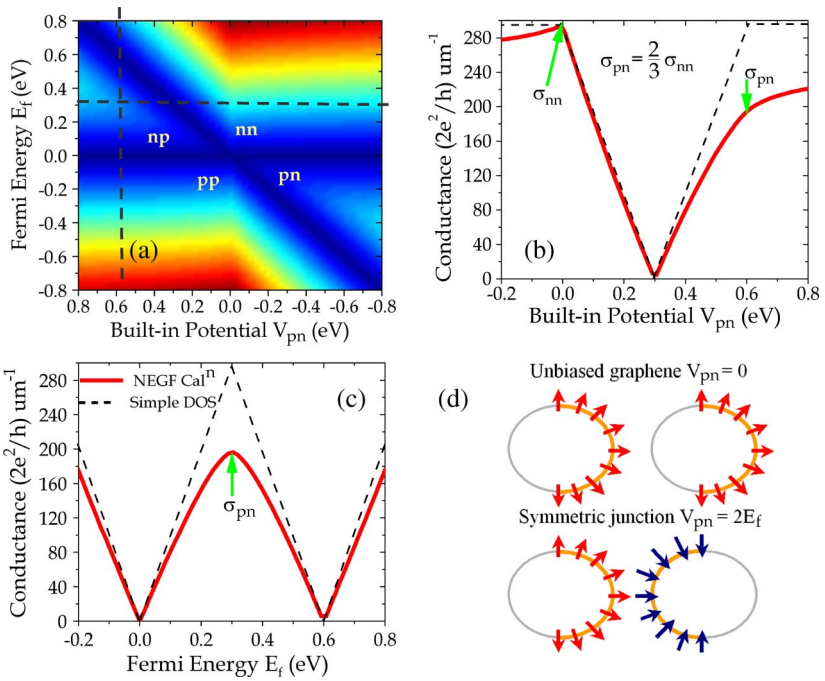

Fig. 3. Simulation results from NEGF calculation of an abrupt graphene p-n junction with device width of 0.5 μm at temperature of 4 K. (a) Intensity plot of conductance as a function of built-in potential $V_{\rm pn}$ and Fermi energy E_f (see inset of Fig. 2(b) for definitions). Blue/red denotes low/high intensity, respectively. [(b) and (c)] Conductance curves as a function of $V_{\rm pn}$ (at $E_f = 0.3$ eV) and E_f (at $V_{\rm pn} = 0.6$ eV), respectively. The dashed lines indicate the estimated conductance from a simple density-of-states argument (see text) for each case. In each case, the conductances of an unbiased graphene ($\sigma_{\rm nn}$) and a symmetric p-n junction ($\sigma_{\rm pn}$) are indicated. (d) Constant energy contour and its pseudospin alignment at each side of the junction for an unbiased graphene and a symmetric p-n junction.

energy-resolved transmission functions for transport across a p-n junction. The mode-space approach has a computational burden that scales linearly with device width, which makes it possible to study typical experimental n-p-n-type structures.

IV. Results

As discussed in Section II, the experimental setup [12], [14]–[16] uses a back gate and top gate to control the source/drain contact Fermi energy E_f and the potential difference across the p-n junction $V_{\rm pn}$, respectively. The electron densities are n_1 in the source/drain regions and n_2 in the channel; a negative value can be interpreted as a positive hole density. In this paper, we assume that the applied source–drain bias is small, in accordance with experiments [12].

We shall also assume that the channel length is greater than the carrier's phase coherent length, allowing us to ignore any quantum interference effects within the channel, i.e., Fabry–Perot effects. On this premise, the transport process across a p-n junction on one side of the channel would not be influenced by the presence of the p-n junction on the other side of the channel. Hence, we shall focus only on the transport across a single p-n junction. We begin by examining the p-n junction's conductance (σ in units of S/μm) when the junction is abrupt, i.e., $D_w \approx 0$.

A. Conductance for an Abrupt p-n Junction

Fig. 3(a), which shows the computed conductance as a function of E_f and $V_{\rm pn}$ for an abrupt p-n junction, exhibits the features typically observed in experiments [12], [19]. The four distinctive regimes of operations (i.e., n-n, p-p, n-p, and p-n) are partitioned by the $E_f = 0$ and $E_f = V_{\rm pn}$ lines, which correspond to conditions where either/both the contacts and channel are at zero equilibrium charge density as previously discussed in Section II using a simple density-of-states argument. Examining more closely, Fig. 3(b) and (c) shows the conductance as a function of $V_{\rm pn}$ ($E_f = 0.3$ eV) and E_f ($V_{\rm pn} = 0.6$ eV), respectively. The red solid line is NEGF calculation, while the dashed black line is calculated using (1). In Fig. 3(b), one can clearly see the conductance asymmetry with respect to the $V_{\rm pn} = E_f$ point. This conductance asymmetry is an experimentally observed phenomenon [12] which cannot be captured by a simple density-of-states argument.

Conductance asymmetry refers to the difference in conductance between an n-p junction and its n-n counterpart. By "counterpart," we mean that n_2 for n-p junction is $-n_2$ for n-n junction, while n_1 is the same for both junctions. In Fig. 3(b), $\sigma_{\rm nn}$ is the conductance of an unbiased graphene (i.e., $V_{\rm pn} = 0$), while $\sigma_{\rm pn}$ is the conductance of a symmetric n-p junction (i.e., $V_{\rm pn} = 2E_f$). Clearly, $\sigma_{\rm nn} > \sigma_{\rm pn}$, a feature that is observed experimentally and captured by our NEGF calculation. The asymmetry is due to wave-function mismatch at the junction interface.

Electrons in graphene have a two-component wave function, which is often referred to as pseudospin in analogy to the two-component wave function for spins. In the vicinity of the Dirac point, the 2-D Hamiltonian can be written in a form of Dirac equation [3], [4], i.e., $H = \upsilon_f \vec{\sigma} \cdot \vec{P}$, where $\vec{\sigma}$ and $\vec{P}$ are the Pauli spin matrices and momentum operator, respectively. By convention, the definition of pseudospin is such that its direction is parallel to the group velocity, since the group velocity

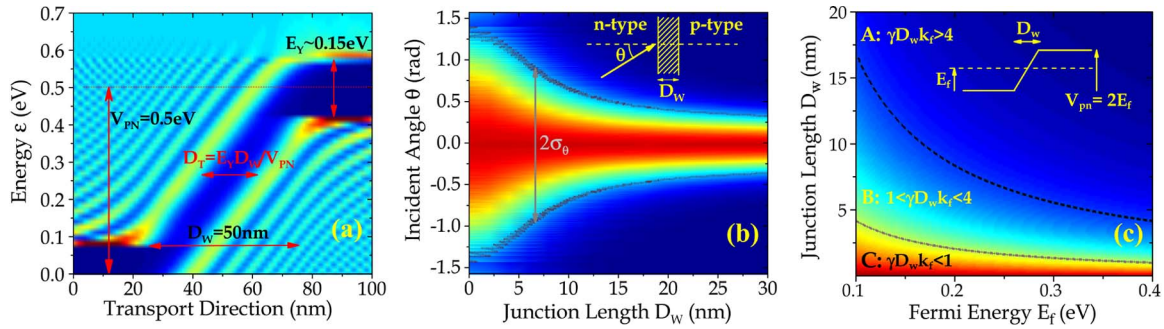

Fig. 4. Analysis of the impact of junction width on conductance modulation of a symmetric p-n junction. (a) Energy-resolved local density-of-states (i.e., $G\Sigma^{\text{in}}G^{\dagger}$) for a $W = 0.5$ μm p-n junction device with a junction length of $D_w = 50$ nm and a built-in potential $V_{\text{pn}} = 0.5$ eV. This is generated for the transverse mode with a transverse energy of $\epsilon_y \approx 0.075$ eV, which yields an apparent "bandgap" of ≈0.15 eV as shown in the plot. $\epsilon = 0$ is set at the Dirac point of the source. (b) Intensity plot of the transmission as a function of incident angle θ and D_w. The device is a symmetric p-n junction with $V_{\text{pn}} = 2E_f$, where $E_f = 0.3$ eV. An illustration of the setup is shown in the inset. (c) Fraction of conductance contribution due to wave-function mismatch, i.e., $\sigma_{\text{pn}}^{\text{wfm}}/\sigma_{\text{pn}}$, for a symmetric p-n junction [see (7)].

operator is defined as $\vec{v}_G = \nabla_P H = v_f \vec{\sigma}$. Fig. 3(d) shows an illustration of the constant energy contour for an unbiased n-n graphene and symmetric p-n junction case. The arrows simultaneously represent the group velocity and pseudospin, which point inwards/outwards for the valence/conduction band, respectively. For the n-n case, the velocity vectors are similarly aligned for each side of the junction. For the p-n case, $\vec{v}_G$ changes sign across the junction. Analogous to spin, the wave function for the n and p regions can be expressed as $|\Psi_n\rangle = (1, e^{i\theta})/\sqrt{2}$ and $|\Psi_p\rangle = (1, e^{i(\theta+\pi)})/\sqrt{2}$, respectively (where $\theta = \tan^{-1}(k_y/k_f)$). The transmission probability across the junction for a particular mode can then be written simply as $|\langle\Psi_n|\Psi_p\rangle|^2$. For $k_y = 0$, the wave function is perfectly matched, i.e., $|\Psi_n\rangle = |\Psi_p\rangle$, so transmission is unity (i.e., Klein tunneling [8]). Through theoretical analysis, we derive $\sigma_{\text{pn}} = 2/3\sigma_{\text{nn}}$, where the factor of 2/3 is due to $\Sigma_{\text{ky}}(1 - k_y^2/k_f^2)$. Similarly, it can also be shown that the conductance when $V_{\text{pn}} \to \infty$ would approach the asymptotic value of $\approx (1/2 + \pi/8)\sigma_{\text{nn}}$. Rigorous NEGF calculations faithfully reproduce these features as shown in Fig. 3(b).

In summary, the conductance of an abrupt p-n junction can be understood as being controlled by the region in which the number of conducting channels is smallest. The wave-function mismatch reduces the current by a factor of 2/3 in a symmetric n-p junction compared to an unbiased graphene. These features are captured by NEGF simulation, but realistic junctions have a finite transition width D_w which, as we emphasize in the following section, plays an important role.

B. Effect of Junction Width on Symmetric p-n Junction Conductance

For realistic n-p junctions, the transition length across which the charge density changes monotonically from n-type to p-type is finite. The width of the junction transition region has a strong influence on the conductance of the junction. To understand these effects, it is instructive to consider each of the transverse modes individually. According to (6), for a wide graphene sheet, the transverse momentum with respect to the Dirac point is given by $k_y \approx n\pi/W$, where n is an integer. We can view each mode as a ray with an angle of incidence on the junction $\theta = \tan^{-1}[k_y(k_f^2 - k_y^2)^{-0.5}]$. Fig. 4(a) shows the energy-resolved local density-of-states ($G\Sigma^{\text{in}}G^{\dagger}$) for a p-n junction with a transition length of $D_w = 50$ nm and a built-in potential $V_{\text{pn}} = 0.5$ eV. Only one transverse mode with a transverse energy of $\hbar v_f k_y \approx 0.075$ eV is considered. The dark-blue region corresponds to a low density-of-states. Because of the quantization of transverse wave vectors, an apparent bandgap of 0.15 eV is observed. Therefore, electrons at Fermi energy of $E_f < V_{\text{pn}}$ will have to quantum mechanically tunnel through this k_y-dependent bandgap when moving from one side of the junction to the other. This problem is analogous to the classic band-to-band tunneling problem in direct bandgap semiconductors [33].

Based on this physical picture, Cheianov and Falko [13] worked out the WKB tunneling probability for a given transverse mode to be $e^{-\pi k_f D_w \sin^2(\theta)/2}$. This tunneling expression is derived by assuming a symmetric p-n junction. In similar spirit to classic band-to-band tunneling treatment, the electric field across the p-n junction is assumed to be linear, given by $2E_f/D_w$ in this paper. Realistically, the potential energy landscape at the beginning/end of the p-n junction would exhibit a more quadratic profile. However, only the details of the potential energy landscape within the tunneling region contribute to the WKB tunneling probability. In this linear electric field approximation, the tunneling distance is simply given by $D_t = 2\hbar v_f k_y D_w/V_{\text{pn}}$. Fig. 4(b) shows the NEGF-computed electron transmission as a function of incident angle θ and transition length D_w for a symmetric p-n junction with $E_f = 0.3$ eV. As expected, increasing D_w results in a decreased angular bandwidth $(2\sigma_\theta)$ of the allowable transverse modes that can be transmitted across the p-n junction, which subsequently leads to a decreased p-n junction conductance. Based on the WKB tunneling formula, it can be shown that $\sigma_{\text{pn}} \propto \sqrt{k_f/D_w}$. This leads to the question of whether the junction rectification metric can be improved by using a larger D_w. Unfortunately, $\sigma_{\text{nn}}/\sigma_{\text{pn}}$, which depends on D_w in an inverse square root manner, yields only a factor of 10 with $D_w \approx 100$ nm at a typical E_f of 0.3 eV.

Device concepts based on the electron focusing property of p-n junction, i.e., graphene lenses [9], [10] and filters [8], operate best in a symmetric p-n junction. These devices operate in a regime where σ_θ has to be as large as possible so as to

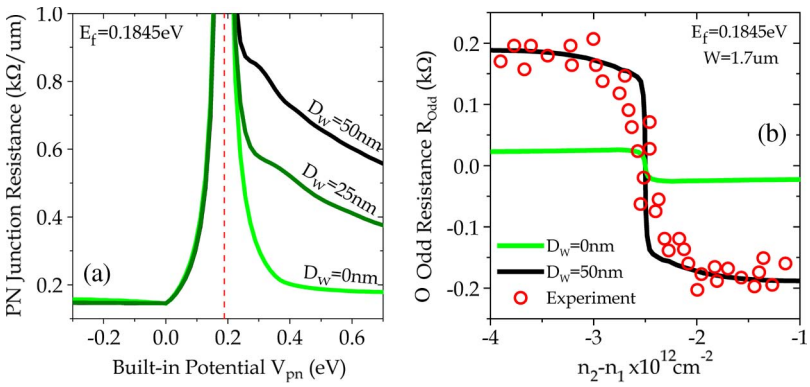

Fig. 5. (a) Theoretical resistance per unit width of a single p-n junction as a function of $V_{\rm pn}$ at $E_f = 0.1845$ eV under different D_w conditions. (b) Comparison of NEGF and experimental odd resistance $R_{\rm odd}$ for different D_w plotting with respect to $n_2 - n_1$. Experimental device has $W = 1.7$ μm with contact Fermi energy also at $E_f = 0.1845$ eV.

reconstruct back a point source image on the other side of the p-n junction. This implies that D_w has to be sufficiently small to enhance tunneling at high incident angle. By accounting for both the wave-function mismatch and tunneling factor in the following manner, $T(\theta) = \cos^2(\theta) e^{-\pi k_f D_w \sin^2(\theta)/2}$, we found that we are able to reproduce the result of NEGF for arbitrary D_w (not shown). By integrating $T(\theta)$ over all transverse modes, we can arrive at a more general result for the conductance of a symmetric p-n junction [in units of $2(e^2/h)$]

$$\sigma_{\rm pn} = \frac{\sqrt{k_f}}{\sqrt{\pi\gamma D_w}}\mathrm{erf}\left(\sqrt{\gamma D_w k_f}\right) + \frac{2}{\pi^2 D_w} e^{-\gamma D_w k_f} - \frac{\sqrt{2}}{\pi^2\sqrt{k_f} D_w^{1.5}}\mathrm{erf}\left(\sqrt{\gamma D_w k_f}\right) \quad (7)$$

where $\gamma = \pi/2$. The first term in (7) is due to the tunneling factor. The last two terms are corrections due to wave-function mismatch. Equation (7) can be written as $\sigma_{\rm pn} = \sigma_{\rm pn}^{\rm tun} + \sigma_{\rm pn}^{\rm wfm}$. Fig. 4(c) shows an intensity plot of $\sigma_{\rm pn}^{\rm wfm}/\sigma_{\rm pn}$ as a function of E_f and D_w. The blue region represents $\sigma_{\rm pn}^{\rm tun} \gg \sigma_{\rm pn}^{\rm wfm}$, while the red regions indicate $\sigma_{\rm pn}^{\rm tun} \ll \sigma_{\rm pn}^{\rm wfm}$. Evidently, conductance modulation is predominantly due to wave-function mismatch only when $\gamma D_w k_f < 1$, which suggests that for the tunneling component not to limit electron focusing applications, $D_w <$ 5 nm is required.

In summary, the conductance asymmetry of a graphene p-n junction is due to two quantum mechanical processes, i.e., wave-function mismatch and the need to tunnel through an apparent bandgap induced by the quantization of transverse momentum. Increasing D_w results in a decreased angular bandwidth $(2\sigma_\theta)$ of the allowable transverse modes that can be transmitted across the p-n junction. This leads to a decreased p-n junction conductance and would eventually result in a larger magnitude of odd resistance $R_{\rm odd}$.

C. Odd Resistance and Comparison With Experiments

Finally, we shall examine the odd resistance $R_{\rm odd}$ of p-n junction devices and compare our NEGF result with the experimental data reported in [12]. Typically, the resistance asymmetry is characterized by analyzing the resistance of the device as a function of $V_{\rm pn}$ at a given E_f. A quantity known as odd resistance $R_{\rm odd}$ can be obtained by taking the difference between the resistance of the n-p-n and its n-n-n **counterpart** device, i.e., $R_{\rm odd} = 1/2[R_{\rm npn} - R_{\rm nnn}]$, where the channel hole density for n-p-n is equal to the electron density for n-n-n. Fig. 1(a) and (b) shows the energy band of a typical n-p-n device and its n-n-n **counterpart**. $R_{\rm odd}$ for a long-channel device depends only on the odd resistance contribution from individual p-n junctions. Essentially, $R_{\rm odd}$ simply reduces to $R_{\rm odd} = [1/\sigma_{\rm np} - 1/\sigma_{\rm nn}]$ obtained by a simple sum of the resistance of the two adjacent p-n junctions. As a first step, we shall investigate the contributions of wave-function mismatch and quantum mechanical tunneling processes to the magnitude of $R_{\rm odd}$. Fig. 5(a) shows the theoretical p-n junction resistance as a function of $V_{\rm pn}$ at $E_f = 0.1845$ eV under different D_w conditions. The odd resistance contribution due to wave-function mismatch alone (i.e., $D_w = 0$) does not adequately account for the $R_{\rm odd}$ observed in experiments as shown in Fig. 5(b). In fact, it only accounts for 10% of $R_{\rm odd}$. Accounting for finite D_w is essential to match the experimental data.

In our NEGF simulation, D_w is the parameter that we need to determine prior to our NEGF calculations. The determination of D_w is an electrostatics problem which is sensitive to the specific device geometry. In this paper, we used the model presented in [17], which expresses D_w in terms of n_1 and n_2 as follows:

$$D_w \approx 0.196 \times \frac{V_{\rm pn}}{\hbar v_f \alpha^{1/3}}\left(1 - \frac{n_2}{n_1}\right)^{\frac{4}{3}}\left|\frac{T_{\rm ox}}{n_2}\right|^{\frac{2}{3}} \quad (8)$$

where $\alpha = e^2/(\kappa\hbar v_f)$. κ and $T_{\rm ox}$ are the effective dielectric constant and thickness of the oxide between the top gate and graphene device, respectively.[2]

[2]We shall also briefly discuss the validity of the analytical nonlinear screening model of (8). As discussed in [17], the accuracy of the analytical model is dependent on the physical parameter α. The model would yield a possible deviation from numerical results of ≈25% when $\alpha \approx 0.9$ at the charge neutrality point within the transition region. We emphasize that the error is of an oscillatory nature, with the oscillation amplitude deteriorating when it moves away from the charge neutrality point. This would serve to alleviate the average error. On the other limit, $\alpha \approx 0.1$ yields an excellent agreement with the numerical result. In the set of experiments we studied in this paper, $\alpha \approx 0.78$ due to $\kappa \approx 4.5$. Therefore, the average error is intermediate of these two limits. However, when α is large (approaching one), one would need to also account for electron exchange and correlation effects (see, e.g., [35]), making the problem numerically nontractable.

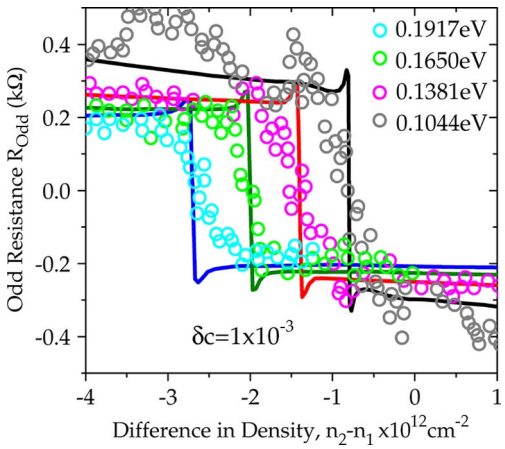

Fig. 6. Comparison of the NEGF-simulated $R_{\rm odd}$ with that of experimental results in [12] at different Fermi energies, i.e., $E_f = 0.1917$, 0.165, 0.1381, and 0.1044 eV. The device width is 1.7 μm, and operating temperature is 4 K. In the simulations, the contacts are assumed to have an energy broadening of 1 meV. The calculations assumed $\alpha = 0.78$ due to a κ of 4.5 [12] and an oxide thickness of $T_{\rm ox} \approx 80$ nm as a fitting parameter.

Fig. 6 shows the computed $R_{\rm odd}$ at different E_f conditions as a function of $n_1 - n_2$ and its comparison with experimental data. The NEGF result achieves quantitative agreement with the experimentally observed odd resistance. In particular, the increase in odd resistance with decreasing E_f, a puzzling feature in the experiments [12], is captured by the simulations. This occurs because of the increasing D_w with decreasing E_f, i.e., smaller n_1 which results in an increase of p-n junction screening length. We note the conductance oscillations at the smallest E_f. These oscillations are likely due to interference effects within the device channel, which are more pronounce at $E_f = 0.1$ eV due to the large D_w which leads to an effectively shorter channel length. The 'spikes' observed in the NEGF simulations when $R_{\rm odd}$ crosses zero are due to the zero density-of-states at Dirac point. Such spikes are not observed in the experiments, probably due to the presence of spatial fluctuations (electron/hole puddles) when E_f approaches the Dirac point [34]. By construction, our NEGF model does not account for these electron/hole puddles.

V. Conclusion

In this paper, we presented a numerical study of electron transport in graphene p-n junctions. We first presented a very simple minimum density-of-states (or conducting channels) model to account for the overall shape of the conductance versus source carrier density and junction potential. Such a simple model does not capture the resistance asymmetry observed experimentally. We then use NEGF simulation to explore, in detail, the role of wave-function mismatch (also called pseudospin) and quantum mechanical tunneling through the junction transition region. In particular, we examined deviations from the inverse square root dependence of $\sigma_{\rm pn}$ on D_w due to wave-function mismatch at small D_w. Finally, we compared the simulations to a recent experiment and showed that the numerical model is in reasonable agreement with experiments, and explain the increase in odd resistance with decreasing carrier density in the source. The novel features of graphene's electronic structure lead to interesting possibilities for new devices, and this paper shows that NEGF simulation should provide a useful tool to explore and assess device concepts.

Acknowledgment

The authors would like to thank the Nanoelectronic Research Initiative and the Network for Computational Nanotechnology for their support. The authors would also like to thank Prof. J. Guo for useful discussions on NEGF simulation by mode-space methods [32] and also S. Srinivasan for useful discussions and comments on this paper.

References

[1] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, "Electric field effect in atomically thin carbon films," *Science*, vol. 306, no. 5696, pp. 666–669, Oct. 2004.

[2] Y. Zhang, Y. W. Tan, and P. Kim, "Experimental observation of the quantum Hall effect and Berry's phase in graphene," *Nature*, vol. 438, no. 7065, pp. 201–204, Nov. 2005.

[3] G. W. Semeno, "Condensed-matter simulation of a three-dimensional anomaly," *Phys. Rev. Lett.*, vol. 53, no. 26, pp. 2449–2452, Dec. 1984.

[4] T. Ando and T. Nakanishi, "Impurity scattering in carbon nanotubes absence of back scattering," *J. Phys. Soc. Jpn.*, vol. 67, no. 5, pp. 1704–1713, 1998.

[5] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov, "Two-dimensional gas of massless Dirac fermions in graphene," *Nature*, vol. 438, no. 7065, pp. 197–200, Nov. 2005.

[6] H. Suzuura and T. Ando, "Crossover from symplectic to orthogonal class in a two-dimensional honeycomb lattice," *Phys. Rev. Lett.*, vol. 89, no. 26, p. 266 603, Dec. 2002.

[7] X. Wu, X. Li, Z. Song, C. Berger, and W. A. de Heer, "Weak antilocalization in epitaxial graphene: Evidence for chiral electrons," *Phys. Rev. Lett.*, vol. 98, no. 13, p. 136 801, Mar. 2007.

[8] M. I. Katsnelson, K. S. Novoselov, and A. K. Geim, "Chiral tunnelling and the Klein paradox in graphene," *Nat. Phys.*, vol. 2, no. 9, pp. 620–625, Sep. 2006.

[9] V. V. Cheianov and V. I. Fal'ko, "The focusing of electron flow and a veselago lens in graphene p-n junctions," *Science*, vol. 315, no. 5816, pp. 1252–1255, Mar. 2007.

[10] A. V. Shytov, N. Gu, and L. S. Levitov, *Klein Backscattering and Fabri–Perot Interference in Graphene Heterojunctions*, 2007. arxiv 0708.3081v1.

[11] S. Datta, *Electronic Transport in Mesoscopic System*. Cambridge, U.K.: Cambridge Univ. Press, 1995.

[12] B. Huard, J. A. Sulpizio, N. Stander, K. Todd, B. Yang, and D. Goldhaber-Gordon, "Transport measurements across a tunable potential barrier in graphene," *Phys. Rev. Lett.*, vol. 98, no. 23, p. 236 803, Jun. 2007.

[13] V. V. Cheianov and V. I. Fal'ko, "Selective transmission of Dirac electrons and ballistic magnetoresistance of n-p junctions in graphene," *Phys. Rev. B, Condens. Matter*, vol. 74, no. 4, p. 041 403, Jul. 2006.

[14] J. R. Williams, L. C. DiCarlo, and C. M. Marcus, "Quantum Hall effect in a gate-controlled p-n junction of graphene," *Science*, vol. 317, no. 5838, pp. 638–641, Aug. 2007.

[15] B. Ozyilmaz, P. Jarillo-Herrero, D. Efetov, D. A. Abanin, L. S. Levitov, and P. Kim, "Electronic transport and quantum hall effect in bipolar graphene p-n-p junctions," *Phys. Rev. Lett.*, vol. 99, no. 16, p. 166 804, Oct. 2007.

[16] R. V. Gorbachev, A. S. Mayorov, A. K. Savchenko, D. W. Horsell, and F. Guinea, "Conductance of p-n-p graphene structures with air-bridge top gates," *Nano. Lett.*, vol. 8, no. 7, pp. 1995–1999, Jul. 2008.

[17] L. M. Zhang and M. M. Fogler, "Nonlinear screening and ballistic transport in a graphene p-n junction," *Phys. Rev. Lett.*, vol. 100, no. 11, p. 116 804, Mar. 2008.

[18] X. Du, I. Skachko, A. Barker, and E. Y. Andrei, "Approaching ballistic transport in suspended 17 graphene," *Nano. Lett.*, vol. 3, no. 8, pp. 491–495, 2008.

[19] N. Stander, B. Huard, and D. G. Gordon, *Observation of Klein Tunneling in Graphene PN Junctions*, 2008. arXiv:0806.2319.

[20] A. F. Young and P. Kim, *Quantum Interference and Carrier Collimation in Graphene Hetero-Junctions*, 2008. arXiv:0808.0855v1.

[21] H. Haug and A. P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors*, ser. Springer Series in Solid State Sciences. New York: Springer-Verlag, 1996, p. 123.
[22] H. Schomerus, "Effective contact model for transport through weakly-doped graphene," *Phys. Rev. B, Condens. Matter*, vol. 76, no. 4, p. 045 433, Jul. 2007.
[23] R. G. Mojarad and S. Datta, *Effect of Contact Induced States on Minimum Conductivity in Graphene*, 2007. arxiv:0710.2727.
[24] P. R. Wallace, "The band theory of graphite," *Phys. Rev.*, vol. 71, no. 9, pp. 622–634, May 1947.
[25] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, *Physical Properties of Carbon Nanotubes*. London, U.K.: Imperial College Press, 1998.
[26] H. Zheng, Z. F. Wang, T. Luo, Q. W. Shi, and J. Chen, "Analytical study of electronic structure in armchair graphene nanoribbons," *Phys. Rev. B, Condens. Matter*, vol. 75, no. 16, p. 165 414, Apr. 2007.
[27] L. Brey and H. A. Fertig, "Electronic states of graphene nanoribbons studied with the Dirac equation," *Phys. Rev. B, Condens. Matter*, vol. 73, no. 23, p. 235 411, Jun. 2006.
[28] M. P. L. Sancho and J. M. L. Sancho, "Quick iterative scheme for the calculation of transfer matrices: Application to Mo (100)," *J. Phys. F, Met. Phys.*, vol. 14, no. 5, pp. 1205–1215, May 1984.
[29] R. Lake, G. Klimeck, R. C. Brown, and D. Jovanovic, "Single and multiband modeling of quantum electron transport through layered semiconductor devices," *J. Appl. Phys.*, vol. 81, no. 12, pp. 7845–7869, Jun. 1997.
[30] R. Venugopal, Z. Ren, S. Datta, M. Lundstrom, and D. Jovanovic, "Simulating quantum transport in nanoscale transistors: Real versus mode-space approaches," *J. Appl. Phys.*, vol. 92, no. 7, pp. 3730–3739, Oct. 2002.
[31] J. Guo, S. Datta, M. Lundstrom, and M. P. Anantram, *Towards Multi-Scale Modeling of Carbon Nanotube Transistors*, 2003. arxiv:0312551.
[32] P. Zhao and J. Guo, "Modeling edge effects in Graphene Nanoribbon Field-effect Transistors with real and mode space methods," *J. Appl. Phys.*, vol. 105, p. 034 503, 2009.
[33] E. O. Kane, "Theory of tunneling," *J. Appl. Phys.*, vol. 32, no. 1, pp. 83–91, Jan. 1961.
[34] E. H. Hwang, S. Adam, and S. D. Sarma, "Carrier transport in two-dimensional graphene 18 layers," *Phys. Rev. Lett.*, vol. 98, no. 18, p. 186 806, May 2007.
[35] S. Adam, E. H. Hwang, E. Rossi, and S. D. Sarma, *Theory of Charge Impurity Scattering in Two Dimensional Graphene*, 2008. arxiv:0812.1795v1.

Tony Low received the Ph.D. degree in electrical engineering from the National University of Singapore, Singapore, in 2008, on the numerical modeling of electronic band structure and transport in thin-film semiconductor devices.

Since 2007, he has been a Postdoctoral Scholar with the Network for Computational Nanotechnology, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN. His current research interest includes modeling of electronic transport in graphene electronics, spin transport, and band-structure effects in spintronic devices.

Dr. Low was the recipient of the IEEE Electron Devices Society and the Singapore Millennium Fellowships.

Seokmin Hong received the B.S. degree in electrical engineering from Seoul National University, Seoul, Korea, in 2003, and the M.S. degree in electrical engineering and computer science from the University of Florida, Gainesville, in 2007. He is currently working toward the Ph.D. degree in the Network for Computational Nanotechnology, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN.

His research interests include the physics, modeling, and simulation of nanodevices.

Joerg Appenzeller (SM'04) received the M.S. and Ph.D. degrees in physics from the Technical University of Aachen, Aachen, Germany, in 1991 and 1995, respectively. His Ph.D. dissertation investigated quantum transport phenomena in low-dimensional systems based on III/V heterostructures.

He worked for one year as a Research Scientist with the Research Center, Juelich, Germany, before he became an Assistant Professor with the Technical University of Aachen in 1996. During his professorship, he explored mesoscopic electron transport in different materials, including carbon nanotubes and superconductor–semiconductor-hybride devices. From 1998 to 1999, he was a Visiting Scientist with Massachusetts Institute of Technology, Cambridge, exploring the ultimate scaling limits of silicon MOSFET devices. From 2001 to 2007, he was a Research Staff Member with the IBM T. J. Watson Research Center, Yorktown, NY, where he was mainly involved in the investigation of the potential of carbon nanotubes and silicon nanowires for future nanoelectronics. Since 2007, he has been a Professor of electrical and computer engineering and a Scientific Director of Nanoelectronics in the Birck Nanotechnology Center, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN. His current research interests include novel devices based on low-dimensional nanomaterials as nanowires, nanotubes, and graphene.

Supriyo Datta (F'96) received the B.Tech. degree from the Indian Institute of Technology, Kharagpur, India, in 1975, and the Ph.D. degree from the University of Illinois at Urbana–Champaign, in 1979.

In 1981, he joined Purdue University, West Lafayette, IN, where he has been the Thomas Duncan Distinguished Professor with the School of Electrical and Computer Engineering since 1999. Since 1985, he has focused on current flow in nanoscale electronic devices and is well known for his contributions to spin electronics and molecular electronics. His most important contribution, however, is the approach his group has pioneered for the description of quantum transport far from equilibrium, combining the nonequilibrium Green function formalism of many-body physics with the Landauer formalism from mesoscopic physics as described in his books *Electronic Transport in Mesoscopic Systems* (Cambridge, 1995) and *Quantum Transport: Atom to Transistor* (Cambridge, 2005).

Dr. Datta is a fellow of the American Physical Society. He started his career in the field of ultrasonics and was selected by the Ultrasonics Group as its Outstanding Young Engineer to receive an IEEE Centennial Key to the Future Award and by the ASEE to receive the Terman Award for his book on *Surface Acoustic Wave Devices*. He has received a number of awards including the IEEE Technical Field Awards both for research and for graduate teaching.

Mark S. Lundstrom (F'94) received the B.E.E. and M.S.E.E. degrees from the University of Minnesota, Minneapolis, in 1973 and 1974, respectively, and the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, IN, in 1980.

From 1974 to 1977, he was with Hewlett-Packard Corporation, Loveland, CO, where he worked on integrated circuit process development and manufacturing support. In 1980, he joined the School of Electrical Engineering, Purdue University, where he is currently the Don and Carol Scifres Distinguished Professor of electrical and computer engineering and the Founding Director of the Network for Computational Nanotechnology. From 1989 to 1993, he served as the Director of Purdue University's Optoelectronics Research Center, and from 1991 to 1994, he served as an Assistant Dean of Engineering. His current research interests center on the physics of small electronic devices, particularly nanoscale transistors, and on carrier transport in semiconductor devices.

Prof. Lundstrom currently serves as an IEEE Electron Devices Society Distinguished Lecturer and is an elected member of the IEEE Electron Devices Society Administrative Council. Over the course of his career, he has received a number of awards for his contributions to research and education. In 2009, he was elected to the U.S. National Academy of Engineering.