A Multiscale Modeling of Triple-Heterojunction Tunneling FETs

Jun Z. Huang, Pengyu Long, Michael Povolotskyi, Hesameddin Ilatikhameneh, Tarek A. Ameen, Rajib Rahman, Mark J. W. Rodwell, *Fellow, IEEE*, and Gerhard Klimeck, *Fellow, IEEE*

***Abstract*— A high performance triple-heterojunction (3HJ) design has been previously proposed for tunneling FETs (TFETs). Compared with single HJ TFETs, the 3HJ TFETs have both shorter tunneling distance and two transmission resonances that significantly improve the ON-state current (I_{ON}). Coherent quantum transport simulation predicts that $I_{ON} = 460$ μA/μm can be achieved at gate length Lg $= 15$ nm, supply voltage $V_{DD} = 0.3$ V, and OFF-state current $I_{OFF} = 1$ nA/μm. However, strong electron–phonon and electron–electron scattering in the heavily doped leads implies that the 3HJ devices operate far from the ideal coherent limit. In this paper, such scattering effects are assessed by a newly developed multiscale transport model, which combines the ballistic nonequilibrium Green's function method for the channel and the drift-diffusion scattering method for the leads. Simulation results show that the thermalizing scattering in the leads both degrades the 3HJ TFET's subthreshold swing through scattering-induced leakage and reduces the turn-ON current through the access resistance. Assuming bulk scattering rates and carrier mobilities, the I_{ON} is dropped from 460 μA/μm down to 254 μA/μm, which is still much larger than the single HJ TFET case.**

***Index Terms*— Diffusive leads, heterojunction (HJ) tunneling FETs (TFETs), multiscale transport, thermalization scattering, triple-HJ (3HJ) TFETs.**

I. Introduction

TUNNELING FET (TFET) is a steep subthreshold swing (SS) device that can operate at low supply voltage, so it is promising for low-power logic electronics application [1]. The low tunneling probability problem of TFETs can be mitigated by employing staggered- or broken-gap heterojunction (HJ) for the tunnel junction to reduce the tunnel barrier height and tunnel distance [2]–[4]. However, even for broken-gap GaSb/InAs HJ TFETs, the tunnel probability is low due to quantum confinement-induced bandgap overlap [5]–[7]. A number of designs have been proposed to further improve the performance of the GaSb/InAs HJ TFETs [5]–[14]. Among these designs, the triple HJ (3HJ) designs [11]–[14] significantly boost the tunnel probability of the HJ TFETs by adding two additional, properly designed, HJs: one in the source and the other in the channel. For n-type 3HJ TFETs [11], atomistic quantum ballistic transport simulations have shown that extremely high I_{ON} of 800 μA/μm (460 μA/μm) could be obtained at Lg $= 30$ nm (15 nm), $V_{DD} = 0.3$ V, and $I_{OFF} = 1$ nA/μm.

The 3HJ devices' promising ON/OFF current ratio is likely to be degraded by various scattering phenomena, as implied by the large spatial overlap of conduction- and valence-band local density of states (LDOS) at the OFF state [11]–[14]. The effect of electron–phonon scattering on the 3HJ design performance has been examined by the nonequilibrium Green's function (NEGF) approach in the self-consistent Born approximation (SCBA) [14], considering only the diagonal components of the self-energy. It is found that the electron–phonon scattering moderately degrades the transistor's I–V characteristics, but the I_{ON} at a given I_{OFF} and V_{DD} remains much larger than that of a GaSb/InAs HJ TFET, when simulated under equal intensity of phonon scattering [14]. The strong electron–electron scattering, such as the Auger generation [15], present in the heavily doped source (and drain) lead, however, has not been examined yet. Practically, this strong, close to thermalizing, effect of electron–electron scattering is very difficult to model in the NEGF approach with an explicit self-energy due to its nonlocal nature, as explained in more detail in [16]. Therefore, we have developed an efficient empirical scattering model, using the NEMO5 tool [17]. In this model, the central channel is treated as nonequilibrium and ballistic, while the leads are assumed to be in local thermal equilibrium with scattering rate derived from experimental carrier mobility. The local quasi-Fermi levels in the leads are determined by solving the drift-diffusion (DD) equations in them. Note that the single scattering rate takes into account multiple scattering mechanisms present in the heavily doped leads, and the DD solver allows us to include the lead serial resistance naturally.

This multiscale model, originally proposed in [18] for resonant tunneling diode simulation, has been generalized to be able to simulate other structures, such as the ultrathin-body (UTB) superlattice transistors [16]. In this paper, the

Manuscript received January 2, 2017; revised March 15, 2017; accepted March 31, 2017. Date of publication April 13, 2017; date of current version May 19, 2017. This work was supported in part by the nanoHUB.org computational resources operated by the Network for Computational Nanotechnology funded by the U.S. National Science Foundation under Grant EEC-0228390, Grant EEC-1227110, Grant EEC-0634750, Grant OCI-0438246, Grant OCI-0832623, and Grant OCI-0721680, in part by the National Science Foundation under Award 1509394 and Award 1639958, and in part by the Semiconductor Research Corporation under Award 2694.03. The review of this paper was arranged by Editor A. Schenk. *(Corresponding author: Jun Z. Huang.)*

J. Z. Huang, P. Long, M. Povolotskyi, H. Ilatikhameneh, T. Ameen, R. Rahman, and G. Klimeck are with the Network for Computational Nanotechnology and Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47907 USA (e-mail: junhuang1021@gmail.com).

M. J. W. Rodwell is with the Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106-9560 USA.

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.2017.2690669

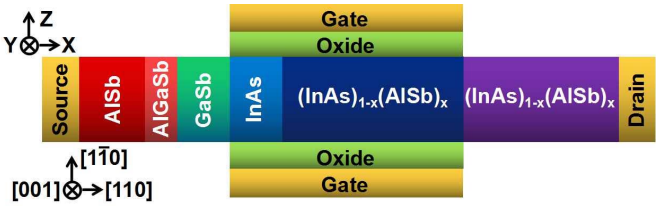

Fig. 1. Device structure and material compositions of the n-type 3HJ TFET. The p-type doped source lead consists of a 5.6-nm AlSb layer with doping density $N_A = 3 \times 10^{19}/\text{cm}^3$, a 1.2-nm $Al_{0.5}Ga_{0.5}Sb$ grade layer, and a 3.3-nm GaSb layer, both with $N_A = 5 \times 10^{19}/\text{cm}^3$. The intrinsic channel consists of a 3-nm InAs layer and a 12-nm $(InAs)_{0.79}(AlSb)_{0.21}$ layer. The n-type doped drain lead consists of a 10-nm $(InAs)_{0.79}(AlSb)_{0.21}$ layer with doping density $N_D = 2 \times 10^{19}/\text{cm}^3$. The body thickness and the oxide thickness are both 1.8 nm with oxide dielectric constant $\epsilon_r = 9$. The confinement (transport) crystal orientation is [1$\bar{1}$0] ([110]).

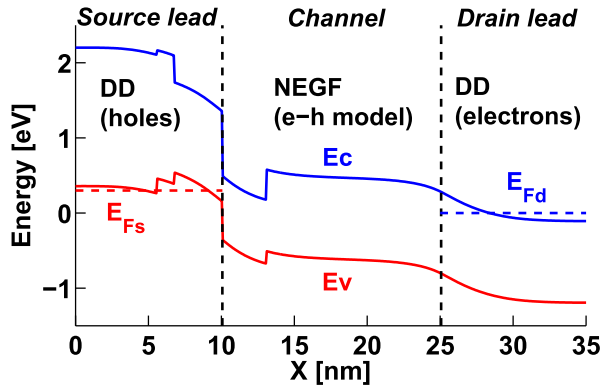

Fig. 2. Region boundaries of the multiscale model. The OFF-state band diagram and the Fermi levels are superimposed. We couple the DD equations for the holes (electrons) in the source (drain) lead and the NEGF equations for both the electrons and the holes in the channel.

model is further extended to consider both electron and hole transport, required for a band-to-band tunneling device simulation. This model is then employed to systematically analyze the 3HJ TFETs by varying the scattering strength, and corresponding mobility, of the leads. The device behaviors under channel length scaling are studied and, finally, the device performances are benchmarked against conventional HJ TFETs.

II. Triple-Heterojunction Design

We consider n-type 3HJ TFETs with double-gate UTB structures [11]. The device geometry and the optimized design parameters are shown in Fig. 1. The added HJ in the channel is InAs/$(InAs)_{0.79}(AlSb)_{0.21}$, while the added HJ in the source is AlSb/GaSb with an $Al_{0.5}Ga_{0.5}Sb$ grading layer. The materials are chosen so that the channel material ($(InAs)_x(AlSb)_{1-x}$ in general) has a higher conduction band edge than in the InAs layer, while the source material ($Al_yGa_{1-y}Sb$ in general) has a lower valence band edge than in the GaSb layer. In addition, these materials have very close lattice constants around 6.1Å [19]; hence, no significant strain is induced. As shown in [11] and [13], the band offsets greatly enhance the electric field at the GaSb/InAs tunnel junction and create two resonant states in the GaSb and InAs quantum wells, improving the tunnel probability at ON state. The larger bandgap and transport effective mass of the $(InAs)_x(AlSb)_{1-x}$ channel, compared with the original InAs channel, also reduce the ambipolar and the source-to-drain tunneling leakage in the subthreshold region. To maximize the electric field at the tunnel junction, the band offsets of the added two HJs should be as large as possible and they should also balance each other, leading to optimal mole fractions $x = 0.79$ and $y = 1$. The GaSb and InAs quantum well widths are then adjusted to have the resonant state levels closely aligned in the Fermi conduction window at the ON state.

III. Multiscale Simulation Approach

The device is divided into three regions, as shown in Fig. 2. The assumption is that, in the doped source and drain regions, the strong scattering drives the free charge carriers into local thermal equilibrium. Therefore, we solve two DD current equations for the thermalized holes in the source and the thermalized electrons in the drain, respectively. Since quantum confinements are critical in the leads, we obtain the hole and electron density of states from Green's functions. The minority carrier current in the leads is neglected, because both the source and the drain leads are heavily doped. In the intrinsic channel region, the scattering rate is much lower, so the quantum-mechanical band-to-band tunneling is the major transport mechanism. Hence, in the channel region, we solve the ballistic NEGF equations involving both the electrons and the holes. The contact self-energies and quasi-Fermi levels for the quantum domain boundaries are taken from the lead surface Green's functions and DD equations, respectively. The central quantum domain provides ballistic current, which is used as a boundary condition for the DD equations in the leads. The DD equations in the leads require LDOS, which is computed from Green's function in a recursive way, using the quantum domain Green's function as a boundary condition.

A. Numerical Model Detail

The device Hamiltonian is modeled using $sp^3d^5s^*$ tight binding (TB) basis with spin–orbital interaction. The parameters are taken from [20]. The alloys are described using the virtual crystal approximation (VCA) with their TB parameters linearly interpolated from their corresponding binaries. In order to utilize the efficient recursive Green's function (RGF) algorithm [21], the whole device is partitioned into a set of layers. The layers are numbered with $l = 1, 2, \ldots, P$ for the layers in the source lead, $l = P+1, P+2, \ldots, Q-1$ for the layers in the channel, and $l = Q, Q+1, \ldots, N$ for the layers in the drain lead.

We solve ballistic NEGF equations in the central channel (between layer P and layer Q), and the current passing through it I_{PQ} is computed by the transmission T_{PQ} weighted by the Fermi function difference

$$I_{PQ} = \frac{q}{h}\int \frac{dk_y}{2\pi}\int \frac{dE}{2\pi} \times T_{PQ}(k_y, E)[f(E - E_{F,P}) - f(E - E_{F,Q})] \quad (1)$$

where q is the electron charge and h is Planck's constant. The integration is performed over the transverse momentum k_y and energy E.

Considering both electrons and holes, the charge density in the central channel per orbital α in layer l can be calculated as [22]

$$\rho_{l,\alpha} = \int \frac{dk_y}{2\pi} \int \frac{dE}{2\pi} \operatorname{sgn}(E - E_{N,l}) \times \left\{ A^L_{l,\alpha}(k_y, E) f[\operatorname{sgn}(E - E_{N,l})(E - E_{F,P})] + A^R_{l,\alpha}(k_y, E) f[\operatorname{sgn}(E - E_{N,l})(E - E_{F,Q})] \right\} \quad (2)$$

where $A^L_{l,\alpha}$ and $A^R_{l,\alpha}$ are the diagonal components of the left- and right-connected spectral functions in the channel, and $\operatorname{sgn}(E)$ is the sign function. $E_{N,l}$ is the layer-dependent threshold (charge neutrality level) defined as the middle of the bandgap

$$E_{N,l} = 0.5(E_{V,l} + E_{C,l}) \quad (3)$$

where $E_{V,l}$ and $E_{C,l}$ are the layer-dependent valence and conduction band edge, respectively.

We solve DD equations for holes in the source lead and for electrons in the drain lead

$$J^h = \mu^h p \nabla E_F \quad (4)$$

$$J^e = \mu^e n \nabla E_F \quad (5)$$

where μ^h and μ^e are the hole and electron mobilities, p and n are the hole and electron densities, and E_F is the major carrier quasi-Fermi level. For simplicity, we assume that the Fermi level is constant over a single layer. Therefore, we have to determine a set of layer-dependent quasi-Fermi levels $E_{F,l}$ ($l = 1, 2, \ldots, P, Q, Q+1, \ldots, N$).

For the hole transport in the source lead, current conservation in the transport direction requires that

$$\mu^h p_l \nabla E_{F,l} - J_{PQ} \approx \mu^h p_l \frac{E_{F,l} - E_{F,l-1}}{\Delta x} - \frac{I_{PQ}}{T_z} = 0 \quad \text{for } l = 1, 2, \ldots, P \quad (6)$$

where the backward difference is used to discretize the differential operator ∇. Δx and T_z are the layer thickness in the x-direction and z-direction, respectively. Similarly, for the electron transport in the drain lead

$$\mu^e n_l \nabla E_{F,l} - J_{PQ} \approx \mu^e n_l \frac{E_{F,l+1} - E_{F,l}}{\Delta x} - \frac{I_{PQ}}{T_z} = 0 \quad \text{for } l = Q, Q+1, \ldots, N \quad (7)$$

where the forward difference is used to discretize the differential operator ∇. p_l and n_l are averaged over layer l.

The hole and electron orbital resolved density in the leads is computed by integrating the diagonals of the lead spectral function $A_{l,\alpha}$ weighted by either hole or electron Fermi-Dirac function with local quasi-Fermi level $E_{F,l}$

$$p_{l,\alpha} = \int \frac{dk_y}{2\pi} \int \frac{dE}{2\pi} \left(\frac{1}{2}[-\operatorname{sgn}(E - E_{N,l}) + 1] \times A_{l,\alpha}(k_y, E) f(E_{F,l} - E) \right) \quad \text{for } l = 1, 2, \ldots, P \quad (8)$$

$$n_{l,\alpha} = \int \frac{dk_y}{2\pi} \int \frac{dE}{2\pi} \left(\frac{1}{2}[\operatorname{sgn}(E - E_{N,l}) + 1] \times A_{l,\alpha}(k_y, E) f(E - E_{F,l}) \right) \quad \text{for } l = Q, Q+1, \ldots, N. \quad (9)$$

To account for a broadening due to scattering in the leads, a small imaginary potential $-i\eta$ is added to the diagonal of the lead Hamiltonian [23]. η is constant below the confined valence band edge and above the confined conduction band edge, while it is exponentially decaying in the bandgap

$$\eta_l(E) = \begin{cases} \eta_h, & E < E_{V,l} \\ \eta_h \exp\left(\dfrac{E_{V,l} - E}{E_0}\right), & E_{V,l} < E < \dfrac{E_{C,l} + E_{V,l}}{2} \\ \eta_e \exp\left(\dfrac{E - E_{C,l}}{E_0}\right), & \dfrac{E_{C,l} + E_{V,l}}{2} < E < E_{C,l} \\ \eta_e, & E_{C,l} < E \end{cases} \quad (10)$$

where η_e and η_h are the electron and hole scattering constants (rates), and decaying constant E_0 is the Urbach parameter. η_h, here, is related to hole momentum relaxation time $\tau_h = \hbar/2\eta_h$ and hole mobility $\mu^h = q\tau_h/m^*_h$, where m^*_h is the hole effective mass; similar expressions relate η_e to electron momentum relaxation time τ_e and electron mobility μ^e.

B. Program Flow

Given an electric potential distribution of the device $V(\mathbf{r})$, the special RGF algorithm in [21] and [23] is employed to compute efficiently the transmission T_{PQ}, $A_{l,\alpha}$ in the leads, as well as $A^L_{l,\alpha}$ and $A^R_{l,\alpha}$ in the channel. Here, the RGF calculations for different transverse momentum k_y and energy E are distributed over different CPU cores. Then, nonlinear equations (6) and (7), together with boundary conditions $E_{F,0} = E_{F,L}$ and $E_{F,N+1} = E_{F,R}$ ($E_{F,L}$ and $E_{F,R}$ are the Fermi levels of the left and right contacts), are solved iteratively using the Newton method to find $E_{F,l}$. Note that the Jacobian matrix is sparse and can be calculated analytically (derivative of the Fermi–Dirac function). With quasi-Fermi levels $E_{F,l}$ determined, the charge density of the whole device is updated using (2), (8), and (9), considering both electrons and holes in the channel [see (2)] and neglecting minority carriers in the leads. The nonlinear Poisson equation is then solved to find a new $V(\mathbf{r})$. The self-consistent simulation flow is summarized in Fig. 3. For the device dimensions shown in Fig. 1 and for one bias point, it usually takes around 1 h to get a converged solution on 96 CPU cores (6 dual 8-core Intel Xeon-E5 CPUs), which is roughly four to five times slower than ballistic simulations with quantum transmitting boundary method [24].

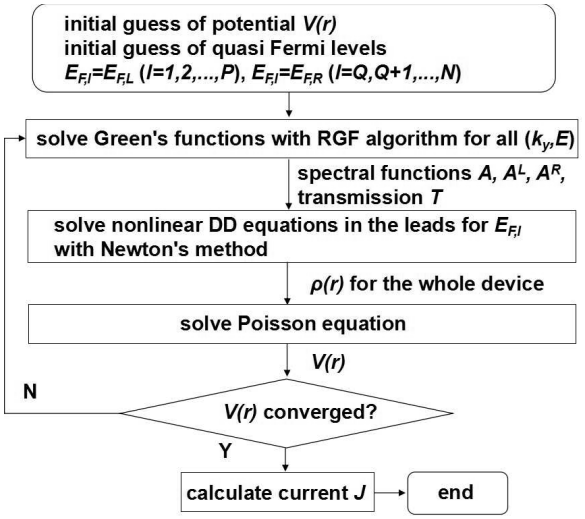

Fig. 3. Self-consistent simulation flow.

IV. Simulation Results

We use $\eta_h = 6$ meV for the holes at the GaSb source, which translates to $\tau_h = 55$ fs and $\mu^h = 241$ cm^2/(V · s), assuming bulk heavy hole effective mass $m^*_{hh} = 0.4m_0$. This mobility value is comparable to the experimental hole mobility of bulk GaSb at 5×10^{19} cm^{-3} doping density, which is around 250 cm^2/(V · s) [25]. We use $\eta_e = 9$ meV for the electrons at the $(InAs)_{0.79}(AlSb)_{0.21}$ drain, which translates to $\tau_e = 36.6$ fs, and $\mu^e = 1497$ cm^2/(V · s) using bulk electron effective mass $m^*_e = 0.043m_0$, obtained from TB calculation using VCA. This mobility value is comparable to the experimental electron mobility of bulk $In_{0.53}Ga_{0.47}As$ at 2×10^{19}cm^{-3} doping density, which is 1422 cm^2/(V · s) [26]. We use the experimental mobility value of $In_{0.53}Ga_{0.47}As$ as a reference, because there is no experimental mobility data for $(InAs)_{0.79}(AlSb)_{0.21}$ to the best of our knowledge, and the electron effective mass of $(InAs)_{0.79}(AlSb)_{0.21}$ is very close to that of $In_{0.53}Ga_{0.47}As$, which is $0.041m_0$. Quantum confinement may change the carrier mobility and scattering rate that needs to be assessed in future studies. We set the Urbach parameter E_0 to be the thermal energy at room temperature, i.e., 26 meV, which has been justified and used in [27] for studying the band tail effects in TFETs.

The transfer characteristics for different scattering rates, and, hence, mobilities, are simulated and shown in Fig. 4(a). The extracted SS is shown in Fig. 4(b). They show that both the SS and the turn-ON current change in the presence of scattering. With fixed $I_{OFF} = 1$ nA/μm and at $V_{DD} = 0.3$ V, the I_{ON} is 460 μA/μm for $\eta = 0$ meV. It reduces by 23% to 352 μA/μm for $\eta_{h(e)} = 2(3)$ meV and by 45% to 254 μA/μm for $\eta_{h(e)} = 6(9)$ meV. In the subthreshold region, as explained in Fig. 5(a)–(d), the scattering in the leads induces an additional peak in the current spectra, which becomes bigger, as the scattering rate increases. At the deep OFF state ($V_{GS} = -0.1$ V), as shown in Fig. 5(a) and (b), the direct source-to-drain tunneling dominates, and, therefore, the total current does not increase significantly. At the OFF state ($V_{GS} = 0$ V), as shown in Fig. 5(c) and (d), the

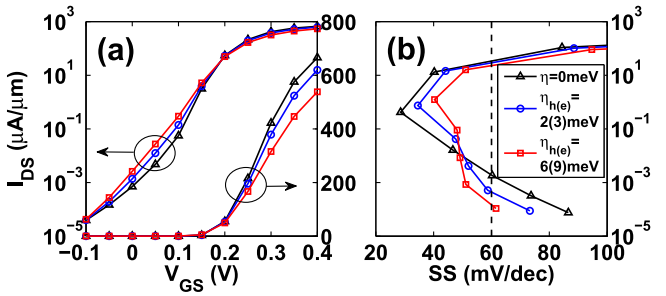

Fig. 4. (a) I_{DS}–V_{GS} characteristics (at $V_{DS} = 0.3$ V) in both logarithmic scale and linear scale. Different η values are compared. η are associated with μ. $\eta = 0$ corresponds to $\mu = \infty$ for both leads. $\eta_h = 2$ meV and $\eta_e = 3$ meV correspond to $\mu^h = 724$ cm^2/(V · s) and $\mu^e = 4492$ cm^2/(V · s) for the source and drain leads, respectively. $\eta_h = 6$ meV and $\eta_e = 9$ meV correspond to $\mu^h = 241$ cm^2/(V · s) and $\mu^e = 1497$ cm^2/(V · s) for the source and drain leads, respectively. (b) I_{DS}-SS extracted from the curves in (a).

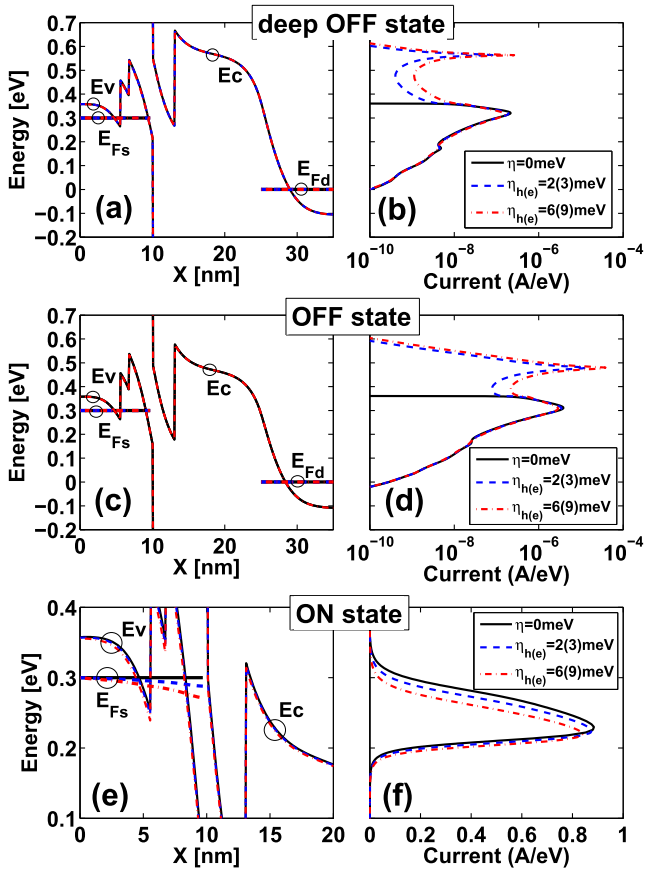

Fig. 5. Band diagrams ($V_{DS} = 0.3$ V) and current spectra ($k_y = 0$) at (a) and (b) $V_{GS} = -0.1$ V, (c) and (d) $V_{GS} = 0$ V, and (e) and (f) $V_{GS} = 0.3$ V. Different values of η and corresponding μ are compared. In (a), (c), and (e), the quasi-Fermi levels in the leads are also superimposed.

scattering leakage dominates, and, therefore, the total current increases significantly. At the ON state ($V_{GS} = 0.3$ V), as explained in Fig. 5(e) and (f), the source serial resistance induces notable potential and quasi-Fermi level drops in the source lead, which become more pronounced as the mobility decreases. The potential drop reduces the transmission and the quasi-Fermi level drop reduces the Fermi conduction window, both decreasing the total current passing through the central channel. Note that attempt to reduce the source serial

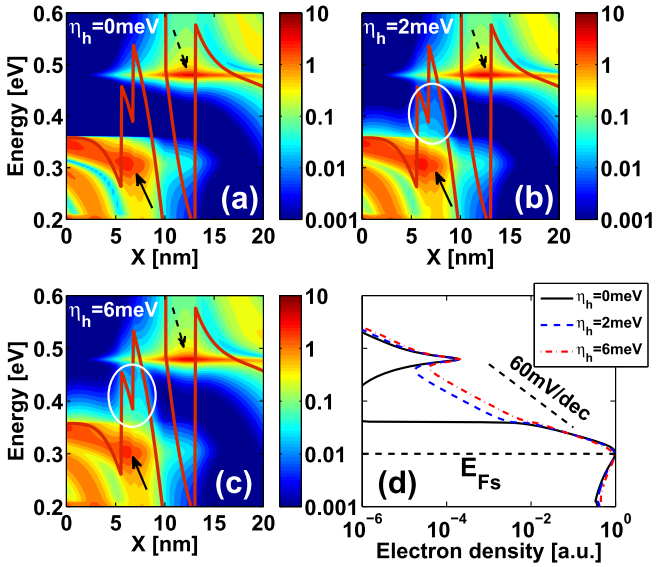

Fig. 6. Logarithmic scale LDOS with the superimposed band diagram at $k_y = 0$, $V_{DS} = 0.3$ V, and $V_{GS} = 0$ V, for (a) $\eta_h = 0$, (b) $\eta_h = 2$ meV, and (c) $\eta_h = 6$ meV. The source (channel) resonant state is marked with solid (dashed) arrows. The scattering-induced LDOS in the source notch is marked with white circles. (d) Electron spectral density at $(x, y, z) = (6.9, 0.46, 1.1)$ nm.

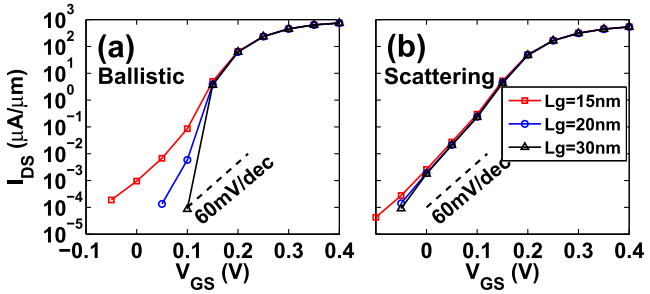

Fig. 7. I_{DS}–V_{GS} (at $V_{DS} = 0.3$ V) for a different channel length Lg = 15 nm, Lg = 20 nm, and Lg = 30 nm. (a) Ballistic simulations and (b) scattering simulations with $\eta_h = 6$ meV ($\eta_e = 9$ meV) and $\mu^h = 241$ cm^2/(V · s) [$\mu^e = 1497$ cm^2/(V · s)] for the source (drain) lead.

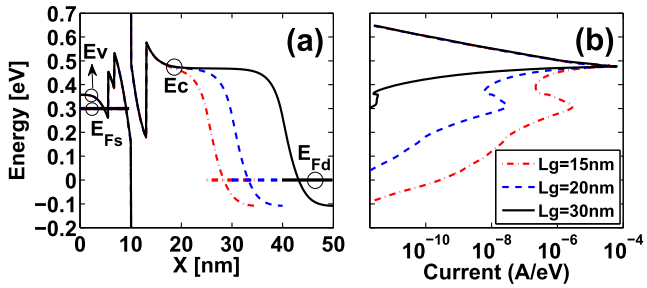

Fig. 8. (a) Band diagram and (b) current spectra at $k_y = 0$, $V_{DS} = 0.3$ V, and $V_{GS} = 0$ for channel length Lg = 15 nm, Lg = 20 nm, and Lg = 30 nm. For all curves, $\eta_h = 6$ meV ($\eta_e = 9$ meV) and $\mu^h = 241$ cm^2/(V·s) [$\mu^e = 1497$ cm^2/(V · s)] are used for the source (drain) lead.

resistance by increasing the doping density would result in large source Fermi degeneracy that in turn degrades the SS. The potential and quasi-Fermi level drop in the drain lead are less pronounced (not shown here) due to the smaller drain serial resistance as a result of the higher electron mobility. To quantify the most relevant mechanism among these two, i.e., the scattering-induced leakage and the serial resistance, for the I_{ON} reduction, we simulated $\mu = \infty$ [keeping $\eta_{h(e)} = 6(9)$ meV] and $\eta = 0$ [keeping $\mu^{h(e)} = 241(1497)$ cm^2/(V·s)] separately, and found that the I_{ON} (at $V_{DD} = 0.3$ V and $I_{OFF} = 1$ nA/μm) drops by 29% due to the scattering-induced leakage and by 24% due to the serial resistance. Therefore, the two mechanisms have similar effects on the device I_{ON}/I_{OFF} ratio.

To explain the additional spectral current peak in the subthreshold region when scattering is present, we compare the LDOS at the OFF state for different scattering rates. For $\eta_h = 0$, essentially ballistic, as shown in Fig. 6(a), the channel resonant state cannot connect to the source contact due to the lack of states in the source bandgap. Therefore, the only leakage path is the direct source-to-drain tunneling leakage, peaked at the source resonant state energy level. While for $\eta_h = 2$ meV and $\eta_h = 6$ meV, as shown in Fig. 6(b) and (c), the source valance band states, in particular, the source resonant state, are broadened, introducing states into the source bandgap and into the source potential notch. As a result, the channel resonant state is connected to the source, giving rise to the additional leakage path. It is worth mentioning that inelastic scattering mechanisms, such as the optical phonon scattering and the electron–electron scattering, are captured here, because in our model $-i\eta_h$ is also added to the bandgap (through the exponentially decaying function) and to the source potential notch where the ballistic LDOS is zero. The electron spectral densities for the three cases, sampled in the source quantum well, are compared in Fig. 6(d). It is shown that the slope of the spectral density increases as the scattering rate increases, and with bulk scattering rate $\eta_h = 6$ meV, the slope is still below 60 mV/decade thermal limit. This explains why the SS of the scattering I–V curves, shown in Fig. 4(b), is still less than 60 mV/decade (for 10^{-3} μA/μm $< I_{DS} <$ 10 μA/μm). For 1-D broken-gap TFETs featuring a similar potential notch in the source [28], in the presence of electron–phonon scattering, sub-60 mV/dec SS was also obtained, which was ascribed to the suppression of DOS in the notch and nonequilibrium carrier distribution below the thermionic limit in the notch. In our simulation, the reduced DOS in the source notch solely accounts for the sub-60 mV/decade SS.

We found that the channel length scaling behaviors are very different when lead scattering is included, as compared in Fig. 7. For the ballistic case, the SS improves as the channel length (Lg) increases, which is easily understood because longer Lg suppresses the source-to-drain tunneling leakage. For the scattering case, when Lg is increased from 15 to 20 nm, the SS is slightly improved, but when Lg is further increased to 30 nm, the SS is not improved further. The explanations are shown in Fig. 8 for the scattering case, where the band diagrams and current spectra at different Lg values are compared. As can be seen, there are two peaks in the spectra, the lower one corresponds to direct source-to-drain tunneling through the source resonant state and the upper one corresponds to the scattering-induced leakage through the channel resonant state. As Lg increases, the direct tunneling leakage is gradually

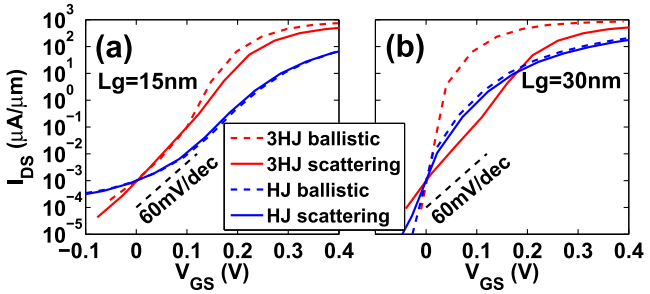

Fig. 9. Comparison of I_{DS}–V_{GS} (at V_{DS} = 0.3V) between the 3HJ TFET and the HJ TFET for channel length (a) Lg = 15 nm and (b) Lg = 30 nm. For all scattering curves, η_h = 6 meV (η_e = 9 meV) and μ^h = 241cm^2/(V · s) [μ^e = 1497cm^2/(V · s)] are used for the source (drain) lead. The threshold voltages are all adjusted for the same I_{OFF} = 1 nA/μm.

suppressed but the scattering-induced leakage remains unchanged. At Lg = 15 nm, the direct tunneling leakage is significant and, therefore, increasing Lg helps reduce the total leakage. At Lg = 20 nm, the scattering-induced leakage already dominates and, consequently, increasing Lg further does not help reduce the total leakage further.

To benchmark the ultimate device performances, we compare I_{ON} of the 3HJ TFETs with the HJ TFETs at V_{DD} = 0.3 V and I_{OFF} = 1 nA/μm, for a short channel length case (Lg = 15 nm) and a long channel length case (Lg = 30 nm), as shown in Fig. 9. At Lg = 15 nm, the scattering reduces I_{ON} of the 3HJ TFET from 460 to 254 μA/μm, while it does not change I_{ON} of the HJ TFET that much, which is 13 μA/μm. At Lg = 30 nm, the scattering reduces I_{ON} of the 3HJ TFET from 767 down to 280 μA/μm and reduces I_{ON} of the HJ TFET from 87 down to 75 μA/μm. Inclusion of lead scattering is, therefore, crucial for predicting the performances of the 3HJ devices. The relatively small reduction of I_{ON} of the HJ TFETs is due to the band tail in the source. Note that [27] predicted a stronger band tail effect using a much larger scattering rate. Overall, the 3HJ TFETs still possess much larger I_{ON} than the HJ TFETs under equal scattering conditions. We note that the 3HJ TFETs are more sensitive than the HJ TFETs to the gate misalignment: simulations (for Lg = 15 nm) show that although the 3HJ TFET can tolerate a small amount (1 nm) of gate underlap with the source, a 1-nm gate overlap with the source leads to a 47% decrease in I_{ON} and a 140% increase in I_{OFF} for the 3HJ TFET, compared with a 18% decrease in I_{ON} and a 99% increase in I_{OFF} for the HJ TFET. Therefore, processes with gate self-alignment are likely necessary and will be subject of future experimental studies. Furthermore, the 3HJ TFET performance is also sensitive to the variations of the quantum well lengths, since the quantum well length critically determines the resonant energy level. Simulations show that, for Lg = 15 nm, +1 (−1) monolayer variation of the GaSb layer length leads to a 5.1% increase (11% decrease) in I_{ON} and a 9.5% increase (14% decrease) in I_{OFF}; +1 (−1) monolayer variation of the InAs layer length leads to a 5.4% increase (19% decrease) in I_{ON} and a 140% increase (54% decrease) in I_{OFF}. Note that one monolayer thickness in the [110] orientation is about 0.216 nm. Therefore, the device performance is more sensitive to the channel well length variation and it needs to be controlled with ±1 monolayer accuracy. Such precision in growth is achievable using atomic layer epitaxy [29].

V. Discussion

The implementation of this model in this paper has two major simplifications. First, the mobility is constant in the source and drain leads. However, the electron density distribution, especially in the source lead, is nonuniform due to the quantum well structure. This implies that the scattering rate and carrier mobility should be spatially varying. Moreover, the scattering rate is constant below E_V and above E_C. Given the complicated band diagram and LDOS, the scattering rate should also be energy-dependent and momentum-dependent. Therefore, a more accurate approach would need a more sophisticated density-dependent mobility model, as well as LDOS-dependent scattering rate. Second, the channel is assumed to be ballistic in this model. In practice, phonon-assisted tunneling would also occur in the channel to form a third leakage path. To model such leakage, the ballistic NEGF approach in the channel needs to be upgraded to include electron–phonon scattering in the usual SCBA way. We note, however, that the scattering in the leads is much stronger than that in the channel due to the high carrier concentrations in the leads. This has been observed even when only the electron–phonon scattering was included [14]. Consequently, the presented model captures the dominant scattering leakage path.

VI. Conclusion

We have discussed that the effect of scattering in the leads is critical for 3HJ TFETs, and, at the same time, is very challenging to model in the NEGF approach. A multiscale model, which solves quantum DD equations in the thermalizing scattering leads and NEGF equations in the ballistic channel, is developed to assess the realistic device performances. This model captures the important scattering leakage current in the subthreshold region and serial resistance above threshold. Simulation results show, that, although the scattering degrades the SS and reduces the turn-ON current, with aligned I_{OFF}, the 3HJ TFETs still operate with <60 mV/decade SS and have much larger I_{ON} than the HJ TFETs for both 15- and 30-nm channel lengths. The 3HJ TFETs are thus very promising in future fast low-power computing applications. The model developed in this paper can be used to study the effects of carrier thermalization and serial resistance in a variety of high-current TFETs, especially those featuring a potential notch in the source [5], [30], [31].

References

[1] A. M. Ionescu and H. Riel, "Tunnel field-effect transistors as energy-efficient electronic switches," *Nature*, vol. 479, no. 7373, pp. 329–337, Nov. 2011.

[2] O. M. Nayfeh, C. N. Chleirigh, J. Hennessy, L. Gomez, J. L. Hoyt, and D. A. Antoniadis, "Design of tunneling field-effect transistors using strained-silicon/strained-germanium type-II staggered heterojunctions," *IEEE Electron Device Lett.*, vol. 29, no. 9, pp. 1074–1077, Sep. 2008.

[3] D. K. Mohata *et al.*, "Demonstration of MOSFET-like on-current performance in arsenide/antimonide tunnel FETs with staggered heterojunctions for 300mV logic applications," in *IEDM Tech. Dig.*, Dec. 2011, pp. 33.5.1–33.5.4.
[4] A. W. Dey *et al.*, "High-current GaSb/InAs(Sb) nanowire tunnel field-effect transistors," *IEEE Electron Device Lett.*, vol. 34, no. 2, pp. 211–213, Feb. 2013.
[5] M. G. Pala and S. Brocard, "Exploiting hetero-junctions to improve the performance of III–V nanowire tunnel-FETs," *IEEE J. Electron Devices Soc.*, vol. 3, no. 3, pp. 115–121, May 2015.
[6] P. Long, J. Z. Huang, M. Povolotskyi, G. Klimeck, and M. J. W. Rodwell, "High-current tunneling FETs with (1$\bar{1}$0) orientation and a channel heterojunction," *IEEE Electron Device Lett.*, vol. 37, no. 3, pp. 345–348, Mar. 2016.
[7] J. Z. Huang, P. Long, M. Povolotskyi, G. Klimeck, and M. J. W. Rodwell, "Scalable GaSb/InAs tunnel FETs with nonuniform body thickness," *IEEE Trans. Electron Devices*, vol. 64, no. 1, pp. 96–101, Jan. 2017.
[8] G. Zhou *et al.*, "Novel gate-recessed vertical InAs/GaSb TFETs with record high I_{ON} of 180 μA/μm at V_{DS} = 0.5 V," in *IEDM Tech. Dig.*, Dec. 2012, pp. 32.6.1–32.6.4.
[9] U. E. Avci and I. A. Young, "Heterojunction TFET scaling and resonant-TFET for steep subthreshold slope at sub-9nm gate-length," in *IEDM Tech. Dig.*, Dec. 2013, pp. 4.3.1–4.3.4.
[10] H. Carrillo-Nuñez, M. Luisier, and A. Schenk, "InAs-GaSb/Si heterojunction tunnel MOSFETs: An alternative to TFETs as energy-efficient switches?" in *IEDM Tech. Dig.*, Dec. 2015, pp. 34.6.1–34.6.4.
[11] P. Long *et al.*, "Extremely high simulated ballistic currents in triple-heterojunction tunnel transistors," in *Proc. 74th Annu. Device Res. Conf. (DRC)*, Jun. 2016, pp. 1–2.
[12] P. Long, J. Z. Huang, M. Povolotskyi, D. Verreck, G. Klimeck, and M. J. W. Rodwell, "High-current InP-based triple heterojunction tunnel transistors," in *Proc. 28th Int. Conf. Indium Phosphide Rel. Mater. (IPRM)*, Jun. 2016, pp. 1–2.
[13] J. Z. Huang, P. Long, M. Povolotskyi, G. Klimeck, and M. J. W. Rodwell, "P-type tunnel FETs with triple heterojunctions," *IEEE J. Electron Devices Soc.*, vol. 4, no. 6, pp. 410–415, Nov. 2016.
[14] P. Long *et al.*, "A tunnel FET design for high-current, 120 mV operation," in *IEDM Tech. Dig.*, Dec. 2016, pp. 30.2.1–30.2.4.
[15] J. T. Teherani, S. Agarwal, W. Chern, P. M. Solomon, E. Yablonovitch, and D. A. Antoniadis, "Auger generation as an intrinsic limit to tunneling field-effect transistor performance," *J. Appl. Phys.*, vol. 120, no. 8, pp. 084507-1–084507-15, Aug. 2016.
[16] P. Long, J. Z. Huang, Z. Jiang, G. Klimeck, M. J. W. Rodwell, and M. Povolotskyi, "Performance degradation of superlattice MOSFETs due to scattering in the contacts," *J. Appl. Phys.*, vol. 120, no. 22, pp. 224501-1–224501-8, Dec. 2016.
[17] S. Steiger, M. Povolotskyi, H.-H. Park, T. Kubis, and G. Klimeck, "NEMO5: A parallel multiscale nanoelectronics modeling tool," *IEEE Trans. Nanotechnol.*, vol. 10, no. 6, pp. 1464–1474, Nov. 2011.
[18] G. Klimeck, "Quantum and semi-classical transport in NEMO 1-D," *J. Comput. Electron.*, vol. 2, no. 2, pp. 177–182, Dec. 2003.
[19] H. Kroemer, "The 6.1 Å family (InAs, GaSb, AlSb) and its heterostructures: A selective review," *Phys. E, Low-Dimensional Syst. Nanostruct.*, vol. 20, no. 3, pp. 196–203, 2004.
[20] Y. Tan, M. Povolotskyi, T. Kubis, T. B. Boykin, and G. Klimeck, "Transferable tight-binding model for strained group IV and III-V materials and heterostructures," *Phys. Rev. B, Condens. Matter*, vol. 94, no. 4, pp. 045311-1–045311-17, Jul. 2016.
[21] R. Lake, G. Klimeck, R. C. Bowen, 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.
[22] J. Guo, S. Datta, M. Lundstrom, and M. P. Anantam, "Toward multiscale modeling of carbon nanotube transistors," *Int. J. Multisc. Comput. Eng.*, vol. 2, no. 2, pp. 257–276, 2004.
[23] G. Klimeck, R. Lake, R. C. Bowen, W. R. Frensley, and T. Moise, "Quantum device simulation with a generalized tunneling formula," *Appl. Phys. Lett.*, vol. 67, no. 17, pp. 2539–2541, Oct. 1995.
[24] M. Luisier, A. Schenk, W. Fichtner, and G. Klimeck, "Atomistic simulation of nanowires in the $sp^3d^5s^*$ tight-binding formalism: From boundary conditions to strain calculations," *Phys. Rev. B, Condens. Matter*, vol. 74, no. 20, p. 205323. Nov. 2006.
[25] D. Martín and C. Algora, "Temperature-dependent GaSb material parameters for reliable thermophotovoltaic cell modelling," *Semicond. Sci. Technol.*, vol. 19, no. 8, pp. 1040–1052, Jul. 2004.
[26] M. Sotoodeh, A. H. Khalid, and A. A. Rezazadeh, "Empirical low-field mobility model for III–V compounds applicable in device simulation codes," *J. Appl. Phys.*, vol. 87, no. 6, pp. 2890–2900, Mar. 2000.
[27] M. A. Khayer and R. K. Lake, "Effects of band-tails on the subthreshold characteristics of nanowire band-to-band tunneling transistors," *J. Appl. Phys.*, vol. 110, no. 7, pp. 074508-1–074508-6, Oct. 2011.
[28] S. O. Koswatta, S. J. Koester, and W. Haensch, "On the possibility of obtaining MOSFET-like performance and sub-60-mV/dec swing in 1-D broken-gap tunnel transistors," *IEEE Trans. Electron Devices*, vol. 57, no. 12, pp. 3222–3230, Dec. 2010.
[29] S. M. Bedair, "Selective-area and sidewall growth by atomic layer epitaxy," *Semicond. Sci. Technol.*, vol. 8, no. 6, pp. 1052–1062, 1993.
[30] D. Verreck *et al.*, "Improved source design for p-type tunnel field-effect transistors: Towards truly complementary logic," *Appl. Phys. Lett.*, vol. 105, no. 5, pp. 243506-1–243506-4, Dec. 2014.
[31] Z. Jiang *et al.*, "Quantum transport in AlGaSb/InAs TFETs with gate field in-line with tunneling direction," *IEEE Trans. Electron Devices*, vol. 62, no. 8, pp. 2445–2449, Aug. 2015.

Jun Z. Huang received the Ph.D. degree from The University of Hong Kong, Hong Kong, in 2013. His Ph.D. thesis was on fast numerical algorithms for quantum electronic device simulations.

In 2013, he joined the Network for Computational Nanotechnology, Purdue University, West Lafayette, IN, USA, as a Post-Doctoral Researcher, where he is currently involved in developing an atomistic nanoelectronics modeling package NEMO5.

Pengyu Long received the B.S. degree in electronic science and technology from the Huazhong University of Science and Technology, Wuhan, China, in 2012, and the M.S. degree in electrical engineering from Purdue University, West Lafayette, IN, USA, in 2015, where he pursuing the Ph.D. degree in electric engineering with the Network For Computational Nanotechnology.

Michael Povolotskyi received the Ph.D. degree in electrical engineering from the University of Rome Tor Vergata, Rome, Italy, in 2004.

In 2008, he joined Network for Computational Electronics with Purdue University, West Lafayette, IN, USA, where he is currently involved in the design and development of an atomistic nanoelectronics simulation package NEMO5, theoretical study of semiconductor devices and development of simulation methodologies.

Hesameddin Ilatikhameneh received the M.S. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2007. He is currently pursuing the Ph.D. degree with Purdue University, West Lafayette, IN, USA.

He is currently a member of the NEMO5 Development Team, where he is involved in the development of quantum transport, strain, mode-space, and phonon solvers for this tool.

Rajib Rahman received the bachelor's degree in physics from the Gettysburg College, Gettysburg, PA, USA, in 2002, and the master's and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, USA, in 2005 and 2009, respectively.

He is currently a Research Assistant Professor with Purdue University, where he is involved in atomistic modeling of electronic structure and transport properties of nanoscale devices.

Mark J. W. Rodwell (M'89–SM'99–F'03) received the Ph.D. degree from Stanford University, Stanford, CA, USA, in 1988.

He currently holds the Doluca Family Endowed Chair in electrical and computer engineering with the University of California at Santa Barbara (UCSB), Santa Barbara, CA, USA. He also directs the UCSB node of the NSF Nanofabrication Infrastructure Network. His Research Group develops nm MOSFFETs, THz transistors, and mm-wave/THz ICs.

Tarek A. Ameen received the B.Sc. degree in electronics and electrical communications engineering and the M.Sc. degree in engineering physics from Cairo University, Giza, Egypt, in 2009 and 2013, respectively. He is currently pursuing the Ph.D. with Purdue University, West Lafayette, IN, USA, under the supervision of Prof. Klimeck.

He is currently a member of the NEMO5 Development Team.

Gerhard Klimeck (S'91–M'95–SM'04–F'13) received the Ph.D. degree from Purdue University, West Lafayette, IN, USA, in 1994.

He was with Texas Instruments and NASA/JPL for ten years. He is currently the Reilly Director of the Center for Predictive Materials and Devices, and the Network for Computational Nanotechnology, Purdue University, and also a Professor of Electrical and Computer Engineering.