Trap Assisted Tunneling and Its Effect on Subthreshold Swing of Tunnel FETs

Redwan N. Sajjad, *Member, IEEE*, Winston Chern, *Student Member, IEEE*, Judy L. Hoyt, *Fellow, IEEE*, and Dimitri A. Antoniadis, *Life Fellow, IEEE*

***Abstract*— We provide a detailed study of the oxide–semiconductor interface trap assisted tunneling (TAT) mechanism in tunnel FETs to show how it contributes a major leakage current path before the band-to-band tunneling (BTBT) is initiated. With a modified Shockley–Read–Hall formalism, we show that at room temperature, the phonon assisted TAT current always dominates and obscures the steep turn ON of the BTBT current for common densities of traps. Our results are applicable to top gate, double gate, and gate-all-around structures, where the traps are positioned between the source-channel tunneling regions. Since the TAT has strong dependence on electric field, any effort to increase the BTBT current by enhancing local electric field also increases the leakage current. Unless the BTBT current can be increased separately, calculations show that the trap density D_{it} has to be decreased by 40–100 times compared with the state of the art in order for the steep turn ON (for III–V materials) to be clearly observable at room temperature. We find that the combination of the intrinsic sharpness of the band edges (Urbach tail) and the surface trap density determines the subthreshold swing.**

***Index Terms*— Shockley–Read–Hall (SRH), surface traps, trap assisted tunneling (TAT), tunnel FET (TFET).**

I. Introduction

THE tunnel FET (TFET) [1] is a candidate for low power switching in digital logic circuits for replacing or supplementing standard CMOS technologies because of its potential to reduce power dissipation via reduction of the power supply voltage. In a TFET, over-the-barrier thermionic emission is completely bypassed by triggering a band-to-band tunneling (BTBT) current by the gate voltage, allowing steep "subthermal" change of current and reduced supply voltage. It has been shown that a small reduction in the subthreshold swing (SS) (e.g., to 45–53 mV/decade) in TFET can reduce the dynamic power dissipation by at least 50% [2], [3] with little sacrifice on the switching delay. Such energy saving is calculated for the same OFF current but lower ON current (compared with the CMOS). The energy savings may enable high-frequency operation that currently CMOS cannot provide. Further improvement is possible if higher ON current is achieved, which can be done with III–V semiconductors and heterojunctions [4].

However, the ideal picture of TFET operation is based upon the assumption that the BTBT current is sufficiently higher than any background current that flows before the bands overlap. In an ideal TFET operation, very little current should flow for gate voltages below a threshold voltage [defined as the gate voltage when the conduction band (CB) bottom in the channel and the valence band (VB) extrema in the source first overlap] and a large amount of current should flow above that. Such a notion of steep (or ideal) switching is practically difficult to achieve, since the combined leakage current, e.g., gate or substrate leakage and bulk or interface trap assisted tunneling (TAT), will always be present and can easily obscure steep change of the BTBT current near the threshold voltage. In addition, the steepness of the current change partly depends on the BTBT magnitude, and since it can be weak for multiple reasons, achieving the steep change of current is highly challenging. Despite numerous efforts in this field, experimental demonstrations with steep turn ON are few [5]–[10] and mostly at very low current levels. Most of the demonstrations involved silicon, for which the interface and bulk defect density is by far the best compared with other materials. Except for [9] and [11], most TFET experimental results on III–V semiconductors [12]–[14] do not show subthermal switching. On the contrary, the SS in these experiments shows strong temperature dependence, clearly indicating the existence of a thermal process. The 2-D layered heterostructure-based TFETs have attracted significant attention in recent times with one experiment [15] showing low SS. But, similar structures by other groups [16], [17] have failed to produce such behavior.

In this paper, we show that the oxide–semiconductor interface TAT current, which is known as a leakage current mechanism in conventional p-n junction diodes [18], [19], is also a major parasitic current component in TFETs. TAT is the emission of electrons to a trap state via electron–phonon interaction, followed by tunneling into the CB (Fig. 1). Similarly, a hole emission and tunneling from a trap is possible. This process is strongly temperature-dependent compared with other nonidealities, such as exponential band tails from the heavy source doping [20]. Such an interband transition is also possible when phonon scattering is considered alone.

Manuscript received July 14, 2016; accepted August 22, 2016. Date of publication September 7, 2016; date of current version October 20, 2016. This work was supported in part by the National Science Foundation through the Center for Energy Efficient Electronics Science Center under Award 0939514 and in part by the NCN-NEEDS Program through the Semiconductor Research Corporation under Grant 1227020-EEC. The review of this paper was arranged by Editor A. Schenk.

The authors are with Microsystems Technology Laboratories, Massachusetts Institute of Technology, Cambridge, MA-02139 USA (e-mail: redwansajjad@gmail.com; wchern@mit.edu; jlhoyt@mit.edu; daa@mit.edu).

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.2016.2603468

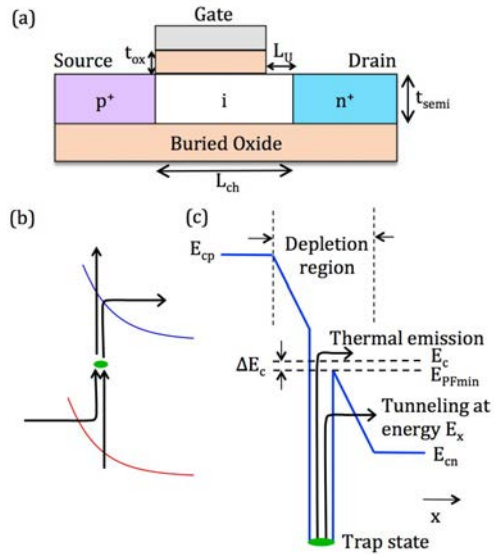

Fig. 1. (a) Schematic of the top gate device considered in this paper. (b) Schematic of the TAT process: an electron can reach the CB from the VB via a combination of phonon absorption and tunneling. Similarly, a hole can be generated. This undesired tunneling is electric-field-dependent in the same way as the ON state BTBT current. The electric field enhanced generation rate is much higher than the classical SRH formalism that does not take the electric field into account. (c) Electron generation part is expanded.

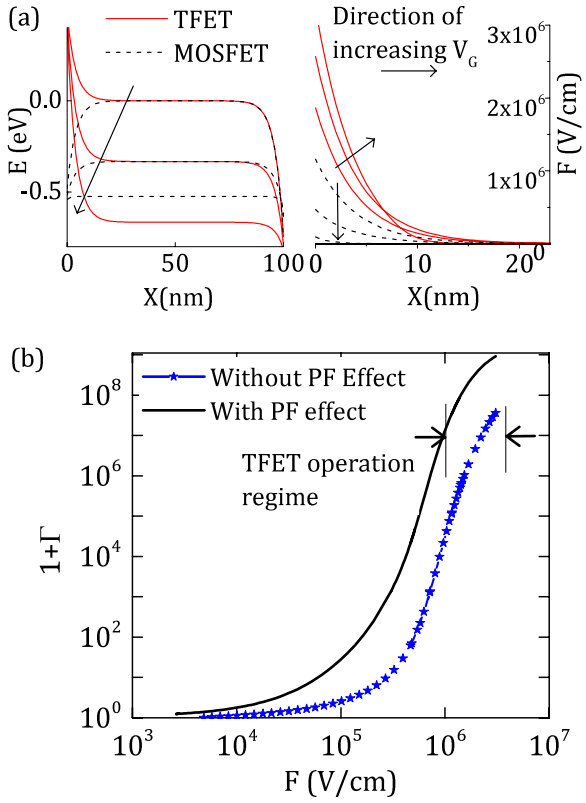

Fig. 2. Relationship of TAT with electric field (F) in TFET. (a) CB profile and the corresponding electric field for silicon TFET at various gate voltages. Solid lines: TFET. Dotted lines: MOSFET configuration. The electric field is increased in TFET as much as possible with gate voltage to increase the BTBT, but in the process, it also increases the undesired trap to CB tunneling. For the MOSFET on the other hand, the electric field is reduced with gate voltage, taking the trap effects out of the picture. (b) Carrier lifetime is decreased as a result of TAT by a factor $1+\Gamma$. Γ is large for the typical electric fields in TFETs and increases the generation rate in the source-channel p-n junction. Here, $1+\Gamma$ versus F is shown at the beginning of the channel ($x = 0$).

Models with phonon scattering (without traps) have shown higher OFF current without sacrificing much on the SS [21]. Although TAT has been identified in the past as a leakage mechanism in TFETs [12], [22]–[26], a detailed quantitative study of its deleterious effects has not been performed. We show that in the presence of traps, electron capture rate prescribed by the Shockley–Read–Hall (SRH) formalism is greatly enhanced due to the high electric field near the source. This is due to the fact that the undesirable electron tunneling from trap to CB depends on the local electric field (Fig. 2), in much the same way as the ON state BTBT current. We show that at room temperature, this TAT current overshadows the steepest part of the BTBT current (Fig. 3) for realistic trap density (midgap $D_{it} = 5 \times 10^{12}/\text{cm}^2\text{-eV}$ for III–V). The steep turn ON of the BTBT current is observable at low temperatures, where the BTBT dominates and the SS becomes less temperature-dependent. In our model, we consider the Poole–Frenkel effect [27]—the lowering of the electron barrier due to the Coulomb interaction of the trap with the lattice. We find that the Poole–Frenkel effect causes a substantial increase in the leakage current by enhancing the trap-channel tunneling. In Section II, we review the electric-field-dependent SRH formalism, the electrostatic model, and the BTBT model used in this paper followed by discussions. The formalism is also applicable to most other device geometries and materials provided that the D_{it} is known and the electrostatic configuration is solved appropriately.

II. Model Description

In this section, we review the electric field enhanced carrier generation rate via phonon and trap states. We incorporate the Poole–Frenkel effect and the tunneling enhanced rates as done in [19] and [27] but apply them using the TFET electrostatics and consider only the surface trap states. The electric field profile from the electrostatic model is used in calculating both the TAT and BTBT current.

A. Trap Assisted Tunneling

The classical SRH formalism [28] describes the generation rate of electron and hole pairs in the presence of traps. An electron in the VB can absorb a phonon to reach a trap state before emitting to the CB by interaction with another phonon. However, in the presence of electric field, the trap-CB (or VB-trap) tunneling rate becomes substantial and greatly increases the electron-hole generation rate [19]. The net generation rate (per unit area) at a given position in the p-n junction space charge region becomes

$$G^n = \int \frac{n_i^2 - np}{\tau_p \frac{n+n_1}{1+\Gamma_p} + \tau_n \frac{p+p_1}{1+\Gamma_n}} D_{it} dE \tag{1}$$

where n_i is the intrinsic carrier concentration, n and p are the electron and hole densities, τ is the minority carrier lifetime, and Γ is a factor that accounts for the tunneling from

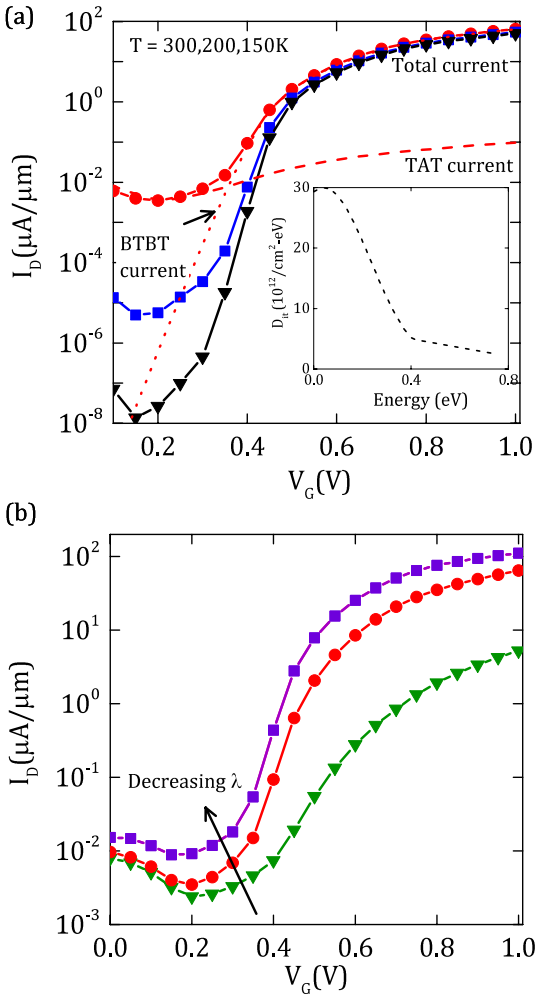

Fig. 3. (a) Total (TAT+BTBT) current in $In_{0.53}Ga_{0.47}As$-based homojunction TFET with the device structure, as shown in Fig. 1 with EOT, t_{ox} and semiconductor thickness t_{semi} 1 and 5 nm, respectively, and a drain bias of $V_{DS} = 0.3$ V. BTBT follows WKB formalism above threshold (when the bands overlap), while below threshold, the BTBT has an exponentially decaying transmission due to the band tails (Urbach tails, in this case at 40 mV/decade at 300 K, 25 mV/decade at 150 K). TAT is temperature-dependent and obscures the steepest part of the BTBT current in the subthreshold regime ($\sim V_G < 0.4$ V) for temperatures above 150 K. The energetic distribution is shown in the inset with mid gap $D_{it} \sim 5 \times 10^{12}/cm^2$-eV. We find that the mid gap traps dominate the net generation rate. (b) Since both TAT and BTBT are electric-field-dependent, the thickness of the oxide and the semiconductor affect the current levels as well as the SS. In this calculation, gradually decrease (from bottom to top) the thicknesses resulting in decreasing scaling lengths λ. t_{ox} is 2, 1, and 0.75 nm and t_{semi} is 10, 5, and 1 nm, respectively. Even for very thin oxide and body thickness, TAT is large enough to overshadow the steep change of BTBT.

trap to CB. With $\Gamma = 0$, (1) reduces to the classical SRH formalism. The interaction between electron and lattice vibration is captured with a phenomenological parameter—capture cross section (σ) in the SRH formalism. The carrier lifetime depends upon σ and the thermal velocity (υ_{th}), $\tau = 1/(\sigma \upsilon_{th})$. Γ is electric-field-dependent and it effectively decreases the minority carrier lifetime. When the electric field is weak, Γ is negligible and (1) reduces to the classical SRH formalism. The terms n_1 and p_1 arise from the principle of detailed balance [28] and are given by $n_1 = n_i \exp^{(E_t - E_i)/k_B T}$ and $p_1 = n_i \exp^{(E_i - E_t)/k_B T}$, where E_i and E_t are the position of the Fermi level for intrinsic semiconductor and the trap state. Fig. 1(b) schematically shows the TAT, which is a two-step process. In the first step of electron emission, the electron is emitted from the VB to the trap state by absorbing a phonon. Afterward, the electron can be partially lifted further and then tunnel into the CB. The amount of the partial lift in the second step can vary from E_{cn}, the position of the CB in the channel, to E_c the position of the CB at the position under consideration. Within the energy range $E_{PF} < E < E_c$, electrons reach the CB without any resistance, since there is no barrier to tunnel, whereas for $E_{cn} < E < E_{PF}$, the transmission probability $T(E)$ through the barrier has to be accounted for. Therefore, the enhancement factor Γ_n is the sum of two components, one for each energy regime

$$\Gamma_n = \Gamma_{PF} + \Gamma_{tunnel} \tag{2}$$

Γ_n is calculated from the net flux (carrier density times the thermal velocity) and the transmission probability [27]

$$\Gamma_n = \frac{1}{k_B T} \int \exp\left(\frac{E_c - E_x}{k_B T}\right) T(E_x) dE_x \tag{3}$$

E_x is the energy to which the electron (or hole) is tunneling to [Fig. 1(b)]. $T(E_x)$ is calculated for a triangular barrier using the Wentzel–Kramers–Brillouin (WKB) approximation

$$T(E_x) = \exp\left(-\frac{4\sqrt{2m^*(E_{PF} - E_x)^3}}{3q\hbar F}\right) \tag{4}$$

where F is the electric field at a particular position in the depletion regime for a given gate voltage. For $E_{PF} < E < E_c$, $T(E_x) = 1$. From (3), it can be shown

$$\begin{aligned} \Gamma_{tunnel} &= \frac{\Delta E_n}{k_B T} \int_0^1 \exp\left[\frac{\Delta E_n}{k_B T} u - K_n u^{3/2}\right] du \\ \Gamma_{PF} &= \frac{1}{4} \exp\left(\frac{E_c - E_{PF}}{k_B T}\right) \\ K_n &= \frac{4}{3} \frac{\sqrt{2m^* \Delta E_n^3}}{q\hbar F} \end{aligned} \tag{5}$$

where $\Delta E_c = E_c - E_{PF}$ is the lowering of the barrier [Fig. 1(b)] due to the Poole–Frenkel effect. ΔE_n is effectively the tunnel barrier height and it also defines the range of energy to which the electron can tunnel to (from the trap). So, ΔE_n is the difference between the top of the barrier and the minimum energy where the electron can tunnel to. Depending upon the position (in the depletion region) under consideration, this can vary from $\Delta E_n = E_{PF} - E_t$ (if $E_t > E_{cn}$) to $E_{PF} - E_{cn}$ (if $E_t < E_{cn}$) [19]. The higher the Poole–Frenkel effect, the higher the ΔE_c and the higher the Γ_n in (5). For typical electric fields, the second term in (5) (which signifies the tunneling contribution) dominates over the first term and increases the exponential term for smaller ΔE_n or larger F. The lowering of the energy barrier ΔE_c is determined by the electric field [29]–[31] $\Delta E_c = q(qF/(\pi\epsilon))^{1/2}$, where ϵ is the electric permittivity. In the same way, Γ_p can be calculated so that all combinations of electron and hole generations (through emission and tunneling), as shown in Fig. 1(b), are included in (1).

Performance degradation in TFET can take place even without the traps due to inelastic phonon scattering [21], [32]. The OFF current is increased in addition to making the transfer I–V ambipolar. But, the phonon limited SS can still be less than 60 mV/decade. Traps on the other hand increase the carrier capture rates to a large extent so that the leakage current dominates over the desired current. TAT affects both the ON–OFF current ratio and the SS.

Fig. 2(b) shows the total enhancement Γ (2) in silicon with and without the Poole–Frenkel effect. Γ can be as high as 10^8, which is effectively the enhancement of the SRH rate. Typical TFET electric fields operate around $1-5\times10^6$ V/cm, over which the Γ changes by less than two orders of magnitude.

Finally, the current is calculated from

$$I/W = q\int G^n(x)dx. \tag{6}$$

B. Electrostatic Model

As derived in [33], we use an abridged version of the 2-D Poisson equation for the top gate structure, as shown in Fig. 1(a). For an Silicon On Insulator structure, the electric field at the top and bottom surface of the semiconductor (given by the oxide thickness and gate potentials) can be applied to the 2-D Poisson equation and can be simplified as

$$\frac{d^2\psi}{dx^2}-\frac{\psi-\phi_{gs}}{\lambda^2}=-\frac{\rho}{\epsilon} \tag{7}$$

where ψ is the surface potential and $\phi_{gs}=V_G-V_{FB}$ is the gate potential. Equation (7) captures the 2-D electrostatics quite well for a given characteristic length λ. For the top gated architecture, $\lambda=((\epsilon_{\text{semi}}/\epsilon_{\text{ox}})t_{\text{ox}}t_{\text{semi}})^{1/2}$. The charge density in the channel is mainly populated by the drain injection, since the channel is poorly coupled to the source

$$\begin{aligned}\rho \approx & -qn_0e^{\mathcal{E}_0/k_BT}\log[1+e^{(-\mathcal{E}_0+\psi-V_{DS})/k_BT}]\ldots\\ &+qp_0e^{\mathcal{E}_0/k_BT}\log[1+e^{(-\mathcal{E}_0-\psi)/k_BT}]\\ &+q\int D_{\text{it}}(1-f_t)dE\end{aligned} \tag{8}$$

where n_0 and p_0 are the equilibrium electron concentration in the channel and $\mathcal{E}_0$ is the position of the first subband from the Fermi energy at zero gate bias. f_t is the occupancy for the donor traps derived from the generation and recombination rate equations (shown in Appendix). Equations (7) and (8) are solved iteratively until self-consistency is achieved. For a given ρ, the potential ψ is calculated numerically from (7) using the finite difference method subject to appropriate boundary conditions (for the doped regions). Equation (7) is also valid for double-gate and gate-all-around nanowire structure if the characteristic length λ is changed appropriately [33].

Fig. 2(a) (left) shows the CB profile. On the right, we show the electric field for various gate voltages. For the TFET configuration, the electric field near the source end is greatly enhanced. For an MOSFET configuration on the other hand, the energy barrier (and the CB) is pushed down resulting in a decreased electric field near the source. This opposite trend in the electric field with gate voltage results in a drastically different TAT current in TFET compared with MOSFET, since the TAT is dependent on the local electric field. The TAT for TFETs increases with gate voltage, while for MOSFETs, it diminishes quickly (not shown). Therefore, the role of traps in MOSFETs is mostly limited to decreased gate efficiency, while for TFETs, it affects both the gate efficiency and leakage.

C. BTBT Model

The transmission probability through the tunnel barrier is determined by the WKB approximation [1]. It can be written as

$$J_{\text{wkb}}=aV_{TW}\left(\frac{F}{F_0}\right)^P\exp\left(-\frac{b}{F}\right) \tag{9}$$

where a, F_0, P, and b are material parameters taken from [34] and [35]. V_{TW} is the tunnel window, i.e., the energy difference between the VB in the source and the CB in the channel; it is determined by an Urbach tail below the threshold voltage and it increases linearly with gate voltage above the threshold voltage [35]. $V_{TW}=E_0\log[1+\exp((E_{v,\text{source}}-E_{c,\text{channel}})/E_0)]$. A difference between [35] and our approach is that we find the position of the CB after self-consistency is achieved between carrier density and channel potential, as discussed in Section II-B. So, for any given gate voltage, the position of the CB is $E_{c,\text{channel}}(V_G)=E_{c,\text{channel}}(V_G=0)-\psi$. E_0 is the Urbach parameter and it represents the intrinsic band steepness.

The Urbach tail has been studied in the past in order to understand the sharpness of the optical absorption spectrum in semiconductors. Instead of a steep rise in the absorption coefficient above a threshold photon energy, experimental results typically show an exponential rise following $\alpha=\alpha_0\exp[-((E-E_g)/E_0)]$ [36], [37]. Such nonabrupt absorption has been attributed to the Urbach tail, which originates in heavily doped semiconductors from the smearing of the dopant energy levels. It can also happen in undoped semiconductors due to electron–phonon interaction [38] with a lower Urbach parameter E_0. The temperature variation of E_0 is weak in doped semiconductors compared with an undoped one [39]. Unfortunately, the exact nature of the Urbach tail and its temperature dependence of E_0 is not well understood [40]. In Section III, we will discuss the implication of various cases of Urbach tail and how it affects the TFET performance.

For a given gate voltage V_G, we solve as discussed for the self-consistent channel potential (7) for the top surface $\psi(x)$ and the electric field $F(x)=-(d\psi)/(dx)$. Using the spatial electric field $F(x)$, we calculate the enhancement factors Γ from (3), carrier densities from (8), and x-dependent generation rate $G^n(x)$ from (1).

III. Results and Discussion

We apply the model for the top gate structure, as shown in Fig. 1(a). Effective oxide thickness (EOT), t_{ox} and semiconductor body thickness t_{semi} are 1 and 5 nm, respectively. We use D_{it} profile in [41] for III–V with mid gap D_{it} of $5\times10^{12}/\text{cm}^2$-eV and consider only donor trap states. Although D_{it} is a function of energy in the bandgap, we found that

in most cases the mid gap trap density dominates the trap current. Channel length, $L_{\rm ch}$ is 100 nm. Source and drain contact regions are degenerately doped while the channel is undoped. The capture cross section for electrons and holes is $\sigma_n = 5 \times 10^{-17}$ m^2 and $\sigma_p = 5 \times 10^{-18}$ m^2 [27], [42] and the carrier lifetimes are calculated from there using the thermal velocity, $\upsilon_{\rm th} = ((8k_BT)/(\pi m^*))^{1/2}$. An underlap (10 nm long) at the channel-drain end is used to suppress the electric field in the drain end and, therefore, the ambipolarity. For the transfer curves, we use a drain bias $V_{DS} = 0.3$ V. We ignore channel resistance due to carrier scattering in the channel, since the resistance due to TAT and BTBT is substantially higher.

Fig. 3(a) shows the transfer plots for $In_{0.53}Ga_{0.47}As$ TFET at various temperatures. For room temperature, the TAT and BTBT current components of the total current are also shown. Well above the threshold voltage ($V_t \sim 0.37$ V), the total current mainly comes from BTBT. The TAT current is just enough so that it intersects with the BTBT current near the threshold voltage, therefore, the total current below V_t is dominated by the TAT. The TAT thus obscures the steepest part of the BTBT (~40 mV/decade in this calculation) and so the minimum SS (~75 mV/decade) is limited by the rate of change of BTBT current just above the threshold voltage. This SS will get much worse for thicker oxide and body thickness. Such transfer behavior with a valley near the minimum current is seen in most experiments on III–V TFETs [13], [43], [44]. At lower temperatures, electron-hole generation rate is reduced leading to lower TAT. For temperatures lower than 200 K, intrinsic SS is observed. The current above the threshold voltage is weakly dependent on temperature while the current below the threshold varies strongly with temperature. In other words, the lowest achievable current at any given temperature is a function of temperature (decreases from ~1 nA at 300 K to ~10 fA at 150 K), similar to what is seen in the experiments [12]–[14], [45], [46]. To demonstrate the effect of the scaling length λ and the local electric field, Fig. 3(b) shows the transfer plots for different oxide and body thicknesses at $T = 300$ K. With $t_{\rm ox} = 0.75$ nm and $t_{\rm semi} = 1$ nm (violet squares), ON current increases substantially due to the increase of the local electric field near the source. SS also improves to ~65 mV/decade, which is still not subthermal. This is due to the fact that the TAT current has also increased, thus limiting the advantage of the higher electric field. We infer that the same effect takes place in heterojunction TFETs, making it difficult to observe subthermal switching for those structures as well.

Fig. 4 shows the model comparison with the experimental data in [12]. Again, both the minimum current and the SS [Fig. 4(b) and (c)] vary strongly with temperature. The traps affect the SS in two ways, through the reduction of the gate efficiency due to charged traps and the TAT. The SS tends to saturate when the TAT is low enough (in this case around 150 K). At this temperature, the overall SS [Urbach tail (E_0) × internal gate efficiency (η_g) × trap limited gate efficiency (η_t)] is 90 mV/decade. From the simulation, $\eta_t(\partial\psi/\partial V_G) \approx 0.55$ for $D_{\rm it} = 10^{12}$/cm^2-eV, and, therefore, $E_0 \times \eta_g \approx 50$ mV/decade for this device. The internal gate efficiency accounts for all factors (such as thick body)

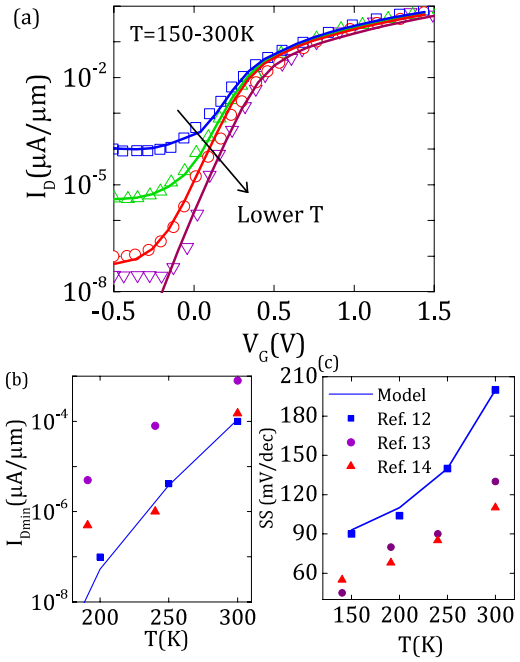

Fig. 4. Comparison of the model with experimental data. (a) Transfer characteristics at different temperatures with $V_{DS} = 0.05$ V of a homogeneous $In_{0.53}Ga_{0.47}As$ TFET as in [12]. As shown in earlier figures, the lower part of the transfer curve is from TAT and the mid gap density that best matched the data is $D_{\rm it} = 1 \times 10^{12}$/cm^2-eV. The gate efficiency for a 4.5-nm EOT is around 0.55 and the intrinsic Urbach tail times the internal gate efficiency turns out to be around 50 mV/decade. (b) Minimum achievable drain current versus temperature shows good agreement with the data and the model. Below 200 K, the minimum current is likely to be limited by other leakage sources (e.g., gate leakage), since the TAT limited current is lower. As reference, the minimum currents found in other experiments ([13], [14]) are also shown. (c) Minimum SS decreases with lower temperature, which is a direct consequence of lower TAT.

other than the charged traps that may be responsible for compromising the SS. Data from [13] and [14] are also shown in Fig. 4 as reference.

Fig. 5 shows the transfer plots for various trap density for two different intrinsic SSs (Urbach tails at 35 and 25 mV/decade) with a motivation to find the trap density required to achieve subthermal switching for multiple decades. We find that a trap density of 1.25×10^{11}/cm^2-eV, which is about 40 times smaller than today's typical mid gap trap density, achieves about two orders of current change at ~40 mV/decade. For the steeper intrinsic swing, we again get two orders of current change at subthermal rate (~28 mV/decade). In this case, the TAT and BTBT intersects at a higher V_G. Since the TAT increases with V_G, a steeper Urbach tail does not necessarily increase the ON–OFF ratio (at subthermal rate). Therefore, the ON–OFF ratio at subthermal rate is determined mainly by the trap density, while the SS is determined by the Urbach tails. We see this again, when the trap density is reduced by 100 times, where we get about three orders of change in current at subthermal rate for both Urbach tails.

We applied the same model to silicon to see how the transfer characteristic changes at reduced trap density and different material properties. In Fig. 6, we see that both BTBT and TAT decrease substantially due to heavier effective mass and higher

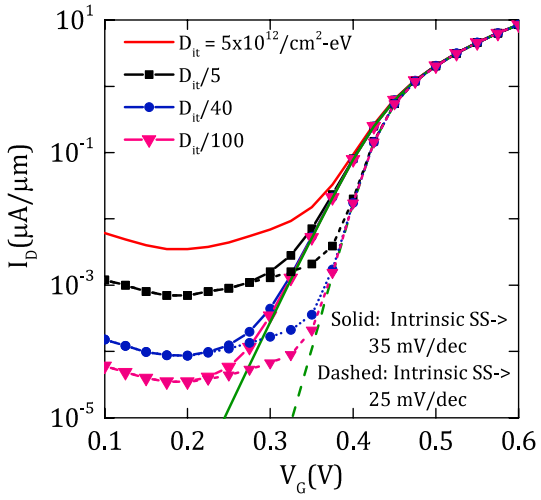

Fig. 5. Impact of D_{it} magnitude on the transfer characteristics for $t_{ox} = 1$ nm and $t_{semi} = 5$ nm for two different Urbach tail parameter E_0. Total current at different mid gap D_{it} levels. At roughly 10^{11}/cm^2-eV (typical D_{it}/40), the TAT current is low enough for the steep BTBT current to be manifested (for two orders of magnitude at ~40 and ~28 mV/decade).

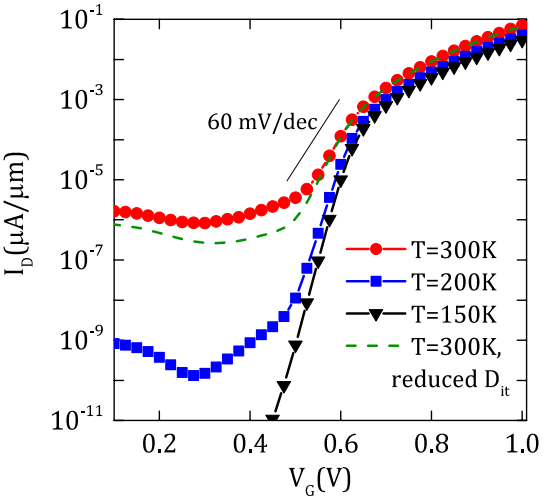

Fig. 6. TAT+BTBT current at different temperatures for homojunction silicon TFET $t_{ox} = 1$ nm and $t_{semi} = 5$ nm. Since silicon has much lower D_{it} (mid gap density assumed here is 5×10^{10}/cm^2-eV), TAT to BTBT transition takes place at a higher temperature (~200 K) compared with III–V. For a slightly lower mid gap D_{it} of 1×10^{10}/cm^2-eV (dashed lines), we get two orders of current change at 50 mV/decade at room temperature.

bandgap with a mid-gap trap density of 5×10^{10}/cm^2-eV, which is typical in today's silicon technology. Similar to III–V, the steepest part of the BTBT is not seen due to the TAT. However, at 1×10^{10}/cm^2-eV, we found (dashed line) two orders of current change at 50 mV/decade (in pA range). Such D_{it} is much easier to achieve in silicon than the requirements mentioned earlier for III–V. This also explains, why most experiments reporting subthermal switching at very low currents involved silicon, where it is likely that such trap density may be achieved.

IV. Conclusion

We provide an analysis of the parasitic TAT current in TFETs. We show that in most cases, the subthreshold current in TFETs is dominated by TAT, regardless of channel material. The takeover from TAT to BTBT depends on the temperature, electrostatic characteristic length, material parameters (e.g., effective mass), and the rate of change of the exponential band tails (Urbach tails). We show that the engineering efforts to increase the ON current are also likely to increase the subthreshold current, since both BTBT and TAT are driven by the same mechanism (tunneling through a barrier). The TAT current is much more deleterious than just the electron–phonon scattering without traps. We find that to get a reasonable ON–OFF ratio with steeper than 60-mV/decade SS at room temperature, trap density has to be reduced by 40–100 times (in the 1×10^{11}/cm^2-eV range) compared with the state of the art for III–V semiconductors, for reasonable structural device parameters. The continuum approach used in this paper is valid for large scale TFET devices that we typically see in most experimental setups. Quantized device structures, such as nanowires and FinFETs, where a single trap can make the overall trap density quite large, will require a different approach than the SRH.

Appendix
Trap Occupancy

The modified electron and hole generation–recombination rates in presence of TAT can be written as

$$G^e = \frac{n_1}{\tau_n}(1+\Gamma_n)\,f_t; \quad R^e = \frac{n}{\tau_n}(1+\Gamma_n)(1-f_t)$$
$$G^h = \frac{p_1}{\tau_p}(1+\Gamma_p)(1-f_t); \quad R^h = \frac{p}{\tau_p}(1+\Gamma_p)\,f_t$$

where f_t is the probability that the trap state at energy E_t is occupied. The net electron generation rate $(G^e - R^e)$ must equal the net hole generation rate $(G^h - R^h)$ under steady-state condition. This leads to

$$f_t = \frac{\sigma_n(1+\Gamma_n)n + \sigma_p(1+\Gamma_p)p_1}{\sigma_n(1+\Gamma_n)(n+n_1) + \sigma_p(1+\Gamma_p)(p+p_1)}. \quad (10)$$

Using this in the net electron generation rate $(G^e - R^e)$ gives us the expression for the net generation rate G^n, as shown in (1). When donor traps are considered alone, Γ_n becomes much larger than Γ_p due to the fact that the Poole–Frenkel effect lowers tunnel barrier height for electrons only. In such a case, we have

$$f_t \approx \frac{n}{n+n_1} \quad (11)$$

which essentially reduces to the Fermi–Dirac distribution, also used in the MOSFET literature [28]. If both types of traps (donors and acceptors) are present, (7), (8), and (10) need to be solved self-consistently.

Acknowledgment

The authors would like to thank E. Yablonovitch (UC Berkeley), P. Xiao (UC Berkeley), S. Agarwal (Sandia), U. Radhakrishna (MIT), and A. Seabaugh (University of Notre Dame) for useful discussions.

REFERENCES

[1] A. C. Seabaugh and Q. Zhang, "Low-voltage tunnel transistors for beyond CMOS logic," *Proc. IEEE*, vol. 98, no. 12, pp. 2095–2110, Dec. 2010.

[2] U. E. Avci *et al.*, "Energy efficiency comparison of nanowire heterojunction TFET and Si MOSFET at Lg = 13 nm, including P-TFET and variation considerations," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, Dec. 2013, pp. 33.4.1–33.4.4.

[3] I. A. Young, U. E. Avci, and D. H. Morris, "Tunneling field effect transistors: Device and circuit considerations for energy efficient logic opportunities," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, Dec. 2015, pp. 22.1.1–22.1.4.

[4] J. Knoch and J. Appenzeller, "Modeling of high-performance p-type III–V heterojunction tunnel FETs," *IEEE Electron Device Lett.*, vol. 31, no. 4, pp. 305–307, Apr. 2010.

[5] J. Appenzeller, Y. M. Lin, J. Knoch, and P. Avouris, "Band-to-band tunneling in carbon nanotube field-effect transistors," *Phys. Rev. Lett.*, vol. 93, no. 19, p. 196805, Nov. 2004.

[6] W. Y. Choi, B.-G. Park, J. D. Lee, and T.-J. K. Liu, "Tunneling field-effect transistors (TFETs) with subthreshold swing (SS) less than 60 mV/dec," *IEEE Electron Device Lett.*, vol. 28, no. 8, pp. 743–745, Aug. 2007.

[7] T. Krishnamohan, D. Kim, S. Raghunathan, and K. Saraswat, "Double-gate strained-Ge heterostructure tunneling FET (TFET) with record high drive currents and ≪60 mV/dec subthreshold slope," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, Dec. 2008, pp. 1–3.

[8] K. Jeon *et al.*, "Si tunnel transistors with a novel silicided source and 46 mV/dec swing," in *Proc. Symp. VLSI Technol. (VLSIT)*, 2010, pp. 121–122.

[9] G. Dewey *et al.*, "Fabrication, characterization, and physics of III–V heterojunction tunneling field effect transistors (H-TFET) for steep subthreshold swing," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, Dec. 2011, pp. 33.6.1–33.6.4.

[10] R. Gandhi, Z. Chen, N. Singh, K. Banerjee, and S. Lee, "Vertical Si-nanowire *n*-type tunneling FETs with low subthreshold swing (≤50 mV/decade) at room temperature," *IEEE Electron Device Lett.*, vol. 32, no. 4, pp. 437–439, Apr. 2011.

[11] B. Ganjipour, J. Wallentin, M. T. Borgström, L. Samuelson, and C. Thelander, "Tunnel field-effect transistors based on InP-GaAs heterostructure nanowires," *ACS Nano*, vol. 6, no. 4, pp. 3109–3113, 2012.

[12] S. Mookerjea, D. Mohata, T. Mayer, V. Narayanan, and S. Datta, "Temperature-dependent *I–V* characteristics of a vertical $In_{0.53}Ga_{0.47}As$ tunnel FET," *IEEE Electron Device Lett.*, vol. 31, no. 6, pp. 564–566, Jun. 2010.

[13] T. Yu, U. Radhakrishna, J. L. Hoyt, and D. A. Antoniadis, "Quantifying the impact of gate efficiency on switching steepness of quantum-well tunnel-FETs: Experiments, modeling, and design guidelines," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, Dec. 2015, pp. 22.4.1–22.4.4.

[14] X. Zhao, A. Vardi, and J. A. del Alamo, "InGaAs/InAs heterojunction vertical nanowire tunnel FETs fabricated by a top-down approach," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, Dec. 2014, pp. 25.5.1–25.5.4.

[15] D. Sarkar *et al.*, "A subthermionic tunnel field-effect transistor with an atomically thin channel," *Nature*, vol. 526, no. 7571, pp. 91–95, Oct. 2015.

[16] T. Roy *et al.*, "Dual-gated MoS_2/WSe_2 van der Waals tunnel diodes and transistors," *ACS Nano*, vol. 9, no. 2, pp. 2071–2079, 2015.

[17] A. Nourbakhsh, A. Zubair, M. S. Dresselhaus, and T. Palacios, "Transport properties of a MoS_2/WSe_2 heterojunction transistor and its potential for application," *Nano Lett.*, vol. 16, no. 2, pp. 1359–1366, 2016.

[18] A. Schenk, "A model for the field and temperature dependence of Shockley–Read–Hall lifetimes in silicon," *Solid-State Electron.*, vol. 35, no. 11, pp. 1585–1596, 1992.

[19] G. A. M. Hurkx, D. B. M. Klaassen, and M. P. G. Knuvers, "A new recombination model for device simulation including tunneling," *IEEE Trans. Electron Devices*, vol. 39, no. 2, pp. 331–338, Feb. 1992.

[20] 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, p. 074508, Oct. 2011.

[21] S. O. Koswatta, M. S. Lundstrom, and D. E. Nikonov, "Influence of phonon scattering on the performance of *p-i-n* band-to-band tunneling transistors," *Appl. Phys. Lett.*, vol. 92, no. 4, p. 043125, 2008.

[22] A. L. Vallett, S. Minassian, P. Kaszuba, S. Datta, J. M. Redwing, and T. S. Mayer, "Fabrication and characterization of axially doped silicon nanowire tunnel field-effect transistors," *Nano Lett.*, vol. 10, no. 12, pp. 4813–4818, 2010.

[23] M. G. Pala and D. Esseni, "Interface traps in InAs nanowire tunnel-FETs and MOSFETs—Part I: Model description and single trap analysis in tunnel-FETs," *IEEE Trans. Electron Devices*, vol. 60, no. 9, pp. 2795–2801, Sep. 2013.

[24] Y. Qiu, R. Wang, Q. Huang, and R. Huang, "A comparative study on the impacts of interface traps on tunneling FET and MOSFET," *IEEE Trans. Electron Devices*, vol. 61, no. 5, pp. 1284–1291, May 2014.

[25] U. E. Avci *et al.*, "Study of TFET non-ideality effects for determination of geometry and defect density requirements for sub-60 mV/dec Ge TFET," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, Dec. 2015, pp. 34.5.1–34.5.4.

[26] S. Agarwal and E. Yablonovitch, "The low voltage TFET demands higher perfection than previously required in electronics," in *Proc. 73rd Annu. IEEE Device Res. Conf. (DRC)*, Jun. 2015, pp. 247–248.

[27] J. Furlan, "Tunnelling generation–recombination currents in a-Si junctions," *Prog. Quantum Electron.*, vol. 25, no. 2, pp. 55–96, 2001.

[28] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*. Hoboken, NJ, USA: Wiley, 2006.

[29] J. C. S. Woo, J. D. Plummer, and J. M. C. Stork, "Non-ideal base current in bipolar transistors at low temperatures," *IEEE Trans. Electron Devices*, vol. 34, no. 1, pp. 130–138, Jan. 1987.

[30] L. Pelaz, J. L. Orantes, J. Vincente, L. A. Bailon, and J. Barbolla, "The Poole–Frenkel effect in 6H-SiC diode characteristics," *IEEE Trans. Electron Devices*, vol. 41, no. 4, pp. 587–591, Apr. 1994.

[31] Q.-A. Huang, M. Qin, B. Zhang, J. K. O. Sin, and M. C. Poon, "A field-enhanced generation model for field emission from p-type silicon," *IEEE Electron Device Lett.*, vol. 18, no. 12, pp. 616–618, Dec. 1997.

[32] Y. Yoon and S. Salahuddin, "Dissipative transport in rough edge graphene nanoribbon tunnel transistors," *Appl. Phys. Lett.*, vol. 101, no. 26, p. 263501, 2012.

[33] R.-H. Yan, A. Ourmazd, and K. F. Lee, "Scaling the Si MOSFET: From bulk to SOI to bulk," *IEEE Trans. Electron Devices*, vol. 39, no. 7, pp. 1704–1710, Jul. 1992.

[34] K.-H. Kao, A. S. Verhulst, W. G. Vandenberghe, B. Sorée, G. Groeseneken, and K. De Meyer, "Direct and indirect band-to-band tunneling in germanium-based TFETs," *IEEE Trans. Electron Devices*, vol. 59, no. 2, pp. 292–301, Feb. 2012.

[35] H. Lu, D. Esseni, and A. Seabaugh, "Universal analytic model for tunnel FET circuit simulation," *Solid-State Electron.*, vol. 108, pp. 110–117, Jun. 2015.

[36] J. L. Pankove, "Absorption edge of impure gallium arsenide," *Phys. Rev.*, vol. 140, no. 6A, p. A2059, 1965.

[37] F. Urbach, "The long-wavelength edge of photographic sensitivity and of the electronic absorption of solids," *Phys. Rev.*, vol. 92, no. 5, p. 1324, Dec. 1953.

[38] A. V. Subashiev, O. Semyonov, Z. Chen, and S. Luryi, "Urbach tail studies by luminescence filtering in moderately doped bulk InP," *Appl. Phys. Lett.*, vol. 97, no. 18, p. 181914, 2010.

[39] S. R. Johnson and T. Tiedje, "Temperature dependence of the Urbach edge in GaAs," *J. Appl. Phys.*, vol. 78, no. 9, pp. 5609–5613, 1995.

[40] C. W. Greeff and H. R. Glyde, "Anomalous Urbach tail in GaAs," *Phys. Rev. B*, vol. 51, no. 3, p. 1778, 1995.

[41] G. Brammertz, H.-C. Lin, M. Caymax, M. Meuris, M. Heyns, and M. Passlack, "On the interface state density at $In_{0.53}Ga_{0.47}As$/oxide interfaces," *Appl. Phys. Lett.*, vol. 95, no. 20, p. 202109, 2009.

[42] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*. New York, NY, USA: Springer, 2012.

[43] D. K. Mohata *et al.*, "Barrier-engineered arsenide–antimonide heterojunction tunnel FETs with enhanced drive current," *IEEE Electron Device Lett.*, vol. 33, no. 11, pp. 1568–1570, Nov. 2012.

[44] R. Pandey *et al.*, "Demonstration of p-type $In_{0.7}Ga_{0.3}As/GaAs_{0.35}Sb_{0.65}$ and n-type $GaAs_{0.4}Sb_{0.6}/In_{0.65}Ga_{0.35}As$ complimentary heterojunction vertical tunnel FETs for ultra-low power logic," in *Proc. Symp. VLSI Technol. (VLSI Technol.)*, 2015, pp. T206–T207.

[45] S. Mookerjea *et al.*, "Experimental demonstration of 100nm channel length $In_{0.53}Ga_{0.47}As$-based vertical inter-band tunnel field effect transistors (TFETs) for ultra low-power logic and SRAM applications," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, Dec. 2009, pp. 1–3.

[46] M. Noguchi *et al.*, "High I_{on}/I_{off} and low subthreshold slope planar-type InGaAs tunnel FETs with Zn-diffused source junctions," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, Dec. 2013, pp. 28.1.1–28.1.4.

Redwan N. Sajjad (M'07) received the B.Sc. and M.Sc. degrees from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 2006 and 2008, respectively, and the Ph.D. degree from the University of Virginia, Charlottesville, VA, USA, in 2014, all in electrical engineering.

He is currently a Post-Doctoral Research Associate with Microsystems Technology Laboratories, Massachusetts Institute of Technology, Cambridge, MA, USA. His current research interests include quantum transport and computational study of emerging nanoelectronic devices.

Winston Chern (S'12) received the B.S. degree in materials science and engineering from the University of Illinois at Urbana-Champaign, Champaign, IL, USA, in 2010, and the M.S. degree in electrical engineering and computer science (EECS) from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2012, where he is currently pursuing the Ph.D. degree in EECS.

His current research interests include Si/Ge and III–V semiconductor devices and their applications.

Judy L. Hoyt (F'08) received the Ph.D. degree in applied physics from Stanford University, Stanford, CA, USA, in 1987.

In 2000, she was a Faculty Member with the Department of Electrical Engineering and Computer Science, MIT Cambridge, MA, USA. Her current research interests include the areas of silicon-germanium heterostructure devices and technology, epitaxial growth, solar cells, Ge-on-Si photodetectors, and CMOS front-end processing.

Dimitri A. Antoniadis (LF'14) received the B.S. degree in physics from the National University of Athens, Athens, Greece, in 1970, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1972 and 1976, respectively.

He joined the MIT Faculty in 1978 and is currently the Ray and Maria Stata Professor of Electrical Engineering. His current research interests include technology and modeling of nanoscale electronic devices in Si, Ge, and III–V materials.