Modeling the Effect of Interface Roughness on the Performance of Tunnel FETs
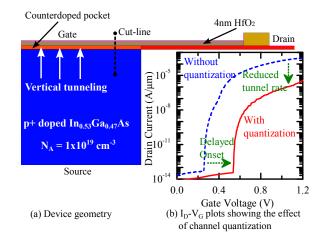
Saurabh Sant and Andreas Schenk

Abstract—The sub-band formation in the triangular-like potential well in the channel of a tunnel field effect transistor (TFET) results in a delayed onset of vertical band-to-band tunneling (BTBT) and in a reduction of the on-current. Furthermore, the roughness of the oxide/semiconductor interface causes density-of-states (DOS) tails, i.e. a smoothing of the otherwise staircase-shaped 2D DOS in the TFET channel. The impact of interface roughness is modeled semi-classically using Ind's perturbation approach. The developed model for the combined effect of channel quantization and interface roughness on TFET performance has been implemented in a commercial device simulator. Simulations of a TFET with predominantly vertical BTBT show that the sharp onset of tunneling, which ideally originates from the 2D-3D DOS matching, is smoothed by the interface roughness. This leads to a degradation of the subthreshold swing of the transistor with increasing amplitude and decreasing auto-correlation length of the roughness.

Index Terms—Tunnel transistors, Interface roughness, Field induced quantum confinement

Fig. 1. (a) InGaAs vertical TFET with counter-doped pocket. The special geometry favors vertical ("line") tunneling and is used to analyze the impact of channel quantization and surface roughness. (b) Simulated transfer characteristics of the vertical TFET in (a) showing the impact of channel quantization.

THE tunnel field effect transistor (TFET) is considered as a potential alternative to the metal oxide semiconductor field effect transistor (MOSFET) in the quest to build low-power logic devices [1]. Various TFET geometries such as bulk, double-gate, and gate-all-around nanowire TFETs have been studied as possible device options. Depending on geometry, doping, and bias voltages, the tunneling paths of band-to-band tunneling (BTBT) are either predominantly lateral or vertical, respectively, with respect to the gate oxide interface, or both directions occur concurrently. TFETs with predominantly vertical tunneling paths, due to special geometry, open the possibility of BTBT between the 2D states in the inversion layer and the 3D states in the "bulk". This mechanism results in steeper onset characteristics compared to the 3D-3D case. Furthermore, the on-current can be simply scaled up by increasing the gate area. However, vertical tunneling is strongly affected by both the quantization of states in the channel [3], [4] and the rough oxide/semiconductor interface which gives rise to density-of-states (DOS) tails in the inversion layer [2]. In this paper, we present a semi-classical model of the combined effect of channel quantization and oxide/semiconductor interface roughness on vertical tunneling.

A vertical ("line-tunneling") TFET was simulated to study the impact of channel quantization and interface roughness. The device consists of p$^+$-doped In$_{0.53}$Ga$_{0.47}$As bulk with a doping concentration of 1×10^{19} cm^{-3}. A counterdoped layer with n$^+$-doping of 4×10^{18} cm^{-3} was introduced in the channel to assist channel formation. The chosen geometry

S. Sant and A. Schenk are with Integrated Systems Laboratory, ETH Zurich, CH-8092 Zurich, Switzerland. Email: sasant@iis.ee.ethz.ch

and doping scheme (shown in Fig. 1) favours vertical tunneling and is, therefore, ideal to study the influence of the above two effects on a TFET.

To account for quantization in the channel, we employed a semi-classical model based on the path-rejection method [3]. In this model, a given tunnel path is selected and the BTBT rate is evaluated along the path, only if it lies energetically above the lowest sub-band in the channel. The sub-band energy is calculated using the triangular well approximation. The applicability of the latter was confirmed by comparing the linearized CB edge with the one obtained from self-consistent simulations at the onset of tunneling. The model was implemented in the commercial device simulator Sentaurus-Device using the "Nonlocal Recombination Physical Model Interface" [5].

A comparison of the transfer characteristics of the above-described TFET including and excluding channel quantization is shown in Fig. 1(b). The inclusion of channel quantization delays the onset of line tunneling and reduces its strength in the ON-state. The delay in the onset is due to the effective increase of the tunnel gap in the channel region caused by the sub-band formation. In the ON-state, only tunnel paths above the sub-band energy level carry the tunnel current while paths below don't find an empty final state. The tunneling electrons are shielded from the gate field by the inversion layer present between the tunnel paths and the gate oxide which weakens the gate coupling. Also, the sharpness of the band bending reduces as one moves away from the oxide. Therefore, the active tunnel paths are longer than the forbidden ones which further reduces the strength of the tunnel current.
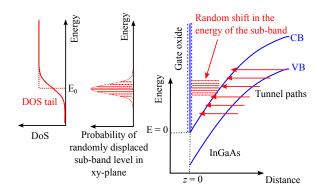
Fig. 2. Interface roughness results in a random shift of the sub-band level which on integration gives rise to DOS tails. A semi-classical model of roughness-induced DOS tails on tunneling is schematically shown. All the tunnel paths are accepted and the BTBT rate along each path is multiplied with the DOS factor.
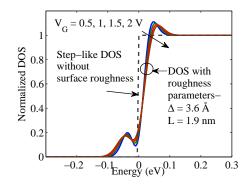

Fig. 3. Normalized DOS in the channel of the device in Fig. 1 calculated by Eq. (1) at different gate bias values. The energies of first subband were set to 0 eV in each calculation. The DOS tails are bias-dependent as a result of the bias dependence of η and $\mathcal{F}$ defined in Eqs. (8a) and (8b).

The oxide interface particularly in III-V-based TFETs is not perfectly smooth, and the roughness is one of the reasons for the broadening of the otherwise discrete sub-band energies. We modeled the effect of interface roughness semi-classically by a spatial variation in the z-position of the infinitely high potential wall. This variation is accompanied by a random shift of the sub-band level which smoothens the otherwise staircase-shaped DOS of the 2DEG in the inversion layer. Band tail states appear below the sub-band energy as shown in Fig. 2(a).

Modeling the DOS in the channel region is necessary to assess the impact of interface roughness on the performance of a TFET. An expression for the DOS of a 2D quantum-mechanical system in the presence of an arbitrary random field with Gaussian correlation was derived by Quang and Tung [7]. According to Ref. [7], the DOS of a 2D system in the presence of a random potential $U(r)$ is given by

$$\varrho(E - E_0) = \frac{m^*}{\pi\hbar^2} \left\{ \frac{1}{2} \left[1 + \mathrm{erf}\left(\frac{E - E_0}{\sqrt{2}\eta} \right) \right] \right.$$
$$\left. - \frac{\hbar^2 \mathcal{F}^2}{24\sqrt{2\pi}m^*\eta^3} \left[1 - \frac{(E - E_0)^2}{\eta^2} \right] \exp\left[-\frac{(E - E_0)^2}{2\eta^2} \right] \right\} , \tag{1}$$

where m^* is the effective mass, E_0 is the energy of the lowest sub-band level, η is the strength of the random potential $U(r)$, and $\mathcal{F}$ is the strength of the field associated with the random potential ($\nabla U(r)$). These quantities are related to the potential

by

$$W(r - r') = \langle U(r) \cdot U(r') \rangle$$
$$\eta^2 = \langle U(r)^2 \rangle = W(r - r')|_{r=r'} \tag{2a}$$
$$\mathcal{F}^2 = \langle (\nabla U(r))^2 \rangle = \nabla_r \nabla_{r'} W(r - r')|_{r=r'} . \tag{2b}$$

In Eqs. (2a), (2b), r and r' are vectors in the xy-plane. To adapt the expression in Eq. (1) for modeling the DOS tails arising from interface roughness, we need to determine the random potential associated with the roughness. It is obtained from Ando's formalism [8] and is given by

$$U(r) := \nu_{\mathrm{SR}}(r) = \int dz\, \zeta_0^*(z)\, V_{\mathrm{SR}}(z, r)\, \zeta_0(z) . \tag{3}$$

Here, $\zeta_0(z)$ is the wave function of the ground state in the triangular well and $V_{\mathrm{SR}}(z, r)$ is the perturbation operator associated with interface roughness as defined in Ref. [8]. If the interface roughness is assumed to have Gaussian correlation, then

$$\langle \Delta(r) \cdot \Delta(r') \rangle = \Delta^2 \exp\left[-\frac{(r - r')^2}{L^2} \right] \tag{4}$$

where $\Delta(r)$ is defined as a random shift in z-position of the potential wall at a location r in the xy-plane. An expression for the roughness potential ($\nu_{\mathrm{SR}}(r)$) was derived in Ref. [8]:

$$\nu_{\mathrm{SR}}(r) = \int dk\, \Delta(k)\Gamma(k) \exp(ik \cdot r) . \tag{5}$$

$\Delta(k)$ is the Fourier transform of the roughness function $\Delta(r)$. The factor $\Gamma(k)$ carries the dimension of an electric field and has the meaning of the Fourier transform of the normal field averaged over z. Note that the averaged normal electric field is r-dependent making $\Gamma(k)$ k-dependent. In the following analysis, k-dependent terms in $\Gamma(k)$ have been neglected. This implies the additional approximation that spatial variations in the average electric field due to interface roughness are negligible. As a consequence, the expression for $\Gamma(k)$ greatly simplifies:

$$\Gamma(k) = \frac{4\pi e^2}{\kappa_s} \left(N_{\mathrm{depl}} + \frac{1}{2} N_{\mathrm{inv}} \right) \approx F_{\mathrm{eff}} . \tag{6}$$

Here, κ_s is the dielectric constant of the semiconductor, N_{depl} is the sheet density of the depletion layer charges, and N_{inv} is the sheet density of the inversion layer charges. The above expression of $\Gamma(k)$ is equal to the "effective electric field" F_{eff} often used in modeling the mobility in MOSFETs. Substituting Eq. (6) in Eq. (5) one obtains for the interface roughness random potential

$$\nu_{\mathrm{SR}}(r) = F_{\mathrm{eff}} \Delta(r) . \tag{7}$$

Inserting this into Eqs. (2a) and (2b) gives the following results for η and $\mathcal{F}$:

$$\eta = F_{\mathrm{eff}} \Delta , \tag{8a}$$
$$\mathcal{F} = F_{\mathrm{eff}} \frac{\Delta}{L} . \tag{8b}$$

Here, Δ is the amplitude and L the auto-correlation length in the Gaussian correlation function describing interface roughness (Eq. (4)). By definition, the effective electric field F_{eff} is

calculated as

$$F_{\text{eff}}^{e}(r) = \frac{\int_0^{z_{\max}} dz\, n(r,z) F_{\perp}(r,z)}{\int_0^{z_{\max}} dz\, n(r,z)} \ , \tag{9}$$

where $n(r,z)$ is the electron density, $z_{\max}$ is the maximum spread of the inversion layer, and $F_{\perp}(r,z)$ is the electric field normal to the oxide/semiconductor interface. A similar expression can be used for the hole channel. The above definition of $F_{\text{eff}}^{e/h}$ has been used in the numerical implementation of the model for DOS tails as described below. It has to be noted that $F_{\text{eff}}^{e/h}$ is bias-dependent. Thus, η and $\mathcal{F}$ defined in Eqs. (8a) and (8b) are also bias-dependent. Their values must be updated at each bias point during the simulation. The normalized DOS factor ($\overline{\varrho} = \frac{\pi \hbar^2}{m^*} \varrho$) calculated using Eq. (1) at different bias values applied to the TFET of Fig. 1 is shown in Fig 3.

The channel quantization model presented above can be extended to include the effect of interface roughness. Accepting a tunnel path with an energy above the lowest sub-band level and rejecting it otherwise is equivalent to multiplying the tunnel rate along each tunnel path by a step function. This step function is equal to the normalized 2D DOS around the lowest sub-band energy (compare Fig. 2). The effect of interface roughness can, therefore, be modeled as follows. Instead of rejecting tunnel paths below the lowest sub-band energy, all tunnel paths are now accepted. The algorithm proceeds in the same manner as the path-rejection algorithm. After calculating the energy E_0 (the bottom of the lowest sub-band) using the triangular well approximation, the effective electric field is computed for each tunnel path using Eq. (9) with the electron density (or the hole density in case of a p-channel TFET) and the electric field in tunnel direction at the discretization point along the tunnel path. This effective electric field is the input for the calculation of the parameters η, $\mathcal{F}$, and the normalized DOS factor $\overline{\varrho} = \frac{\pi \hbar^2}{m^*} \varrho$ for each tunnel path. The tunnel rate for a tunnel path with energy E based on the implemented Kane model in Ref. [5] is multiplied by this normalized DOS factor calculated at the same energy:

$$G_{\text{vc}}(E, x = x_{\text{v}}) = \overline{\varrho}(E - E_0) e F(x_{\text{v}}) \frac{g}{72\hbar} \times$$

$$\frac{1 - \exp\left(-k_m^2 \int_{x_{\text{v}}}^{x_{\text{c}}} \frac{dx}{\kappa}\right)}{\int_{x_{\text{v}}}^{x_{\text{c}}} \frac{dx}{\kappa}} \exp\left(-2 \int_{x_{\text{v}}}^{x_{\text{c}}} \kappa \, dx\right) (f_{\text{c}} - f_{\text{v}}) \ . \tag{10}$$

Here, x is the position along the tunnel path, F is the electric field, x_{v}, x_{c} are, respectively, the start and end points of the tunnel path, κ denotes the imaginary $E(k)$-relation, f_{c} and f_{v} are, respectively, the quasi-Fermi distribution functions of electrons and holes. The modified BTBT rate is then incorporated in the self-consistent simulation as a generation rate of holes and electrons at the beginning and the end of the tunnel path, respectively. It has to be noted that the above semi-classical model is solely phenomenological as an energetic tunnel rate for bulk states is averaged by a 2D DOS.

In order to perform simulations of the vertical TFET in Fig. 1(a), the semi-classical model needs to be implemented
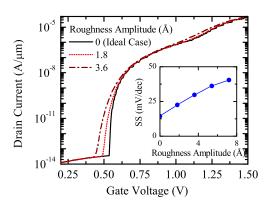

Fig. 4. Transfer characteristics of the vertical TFET for different values of the roughness amplitude Δ. Increasing Δ degrades the SS (inset).

in a 2D device simulator. The model was implemented in Sentaurus-Device [5] via the "Nonlocal Recombination Physical Model Interface". A roughness amplitude Δ of 1.8 Å and a correlation length L of 1.9 nm had been experimentally extracted for a InAs/GaSb heterojunction by Feenstra *et al.* [9] using Transmission Electron Microscopy. Taking these numbers as a guideline, simulations were performed with $\Delta = 1.8 \, \text{Å}, 3.6 \, \text{Å}$ and $L = 1.9 \, \text{nm}$. A comparison of the transfer characteristics for different parameter values is presented in Fig. 4. The DOS tails resulting from interface roughness smoothen the onset of line tunneling. An increasing roughness amplitude degrades the sub-threshold swing (SS). This is a consequence of the direct proportionality of η and $\mathcal{F}$ to the roughness amplitude Δ. Also the simulations show that a decreasing auto-correlation length L is would degrade the SS of the TFET.

It must be noted that the above models for channel quantization and interface roughness are applicable for vertical BTBT only. Furthermore, interface roughness is among many non-idealities such as interface traps, bulk traps, and disorder-induced band tails which adversely affect the performance of a TFET. In Ref. [10], it was observed that the effect of interface traps is the strongest among them. However, the detrimental effect of traps is expected to fade out as technology improves. In that case, interface roughness may become an important mechanism for the degradation of the SS in TFETs.

In summary, a semi-classical model based on the path rejection method was implemented in a device simulator to analyze the impact of channel quantization on vertical tunneling. Additionally, an analytical model for the DOS of an arbitrary 2D quantum-mechanical system was adapted to capture the effect of DOS tails due to interface roughness. The simulations of a vertical TFET using realistic parameters for roughness amplitude and auto-correlation length have show that the SS degrades due to interface roughness. Eqs. (8a), (8b) as well as the simulations suggest that an increase in the roughness amplitude or a reduction in the auto-correlation length increase the SS of TFETs which are dominated by vertical BTBT.

REFERENCES

[1] A. M. Ionescu, H. Riel, "Tunnel field-effect transistors as energy-efficient electronic switches, Nature vol. 479, pp. 329–337, 2011 [doi: 10.1038/na-

ture10679].

[2] S. Goodnick, R. Gann, D. Ferry, C. Wilmsen, and O. Krivanek, "Surface roughness induced scattering and band tailing, Surface Science vol. 113, pp. 233–238, 1982 [doi: 10.1016/0039-6028(82)90591-X].

[3] W. G. Vandenberghe, B. Soree, W. Magnus, G. Groeseneken, A. Verhulst, and M. V. Fischetti, "Field induced quantum confinement in indirect semiconductors: Quantum mechanical and modified semiclassical model, SISPAD'11, pp. 271–274 (2011) [doi: 10.1109/SISPAD.2011.6035077].

[4] S. Sant, A. Schenk, K. Moselund, and H. Riel, "Impact of Trap-assisted Tunneling and Channel Quantization on InAs/Si Hetero Tunnel FETs", 74th IEEE Device Research Conference, pp. 67–68, 2016 [doi: 10.1109/DRC.2016.7548413].

[5] Synopsys Inc., Sentaurus Device User Guide, Version 2015.06, Mountain View, California, (2015).

[6] E. O. Kane, "Theory of tunneling, J. Appl. Phys., vol. 32, p. 83, 1961 [doi: 10.1063/1.1735965].

[7] D. Quang and N. Tung, "A Semiclassical Approach to the Electron Gas in Two-Dimensional Semiconductor Structures, phys. stat. sol. (b), vol. 207, pp. 111–123, 1998 [doi: 10.1002/(SICI)1521-3951(199805)207:1<111::AID-PSSB111>3.0.CO;2-P].

[8] T. Ando, "Screening effect and quantum transport in Silicon inversion layer in strong magnetic fields, J. Phys. Soc. Japan, vol. 43, pp. 1616–1626, 1977 [doi: 10.1143/JPSJ.43.1616].

[9] R. M. Feenstra, D. A. Collins, D. Z.-Y. Ting, M. W. Wang, and T. C. McGill, "Interface roughness and asymmetry in InAs/GaSb superlattices studied by scanning tunneling microscopy, Phys. Rev. Lett. vol. 72, pp. 2749–2752, 1994 [doi: 10.1103/PhysRevLett.72.2749].

[10] A. Schenk, S. Sant, K. Moselund, and H. Riel, "III-V-Based Hetero Tunnel FETs: A Simulation Study Focussed on Non-ideality Effects", Proc. ULIS-EUROSOI, pp. 9–12, 2016 [doi: 10.1109/ULIS.2016.7440039].