DISS. ETH NO. 25129

Role of Non-idealities in III-V/Si and All III-V Tunnel Field Effect Transistors

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

Saurabh Sant

M.S. in Electrical Engineering and Information Technology, ETH Zurich, Switzerland
B.Tech. in Electrical Engineering, IIT Bombay, India

born on 11.11.1989
citizen of India

accepted on the recommendation of

Prof. Dr. Andreas Schenk, examiner
Prof. Dr. Mathieu Luisier, co-examiner
Prof. Dr. Alan Seabaugh, co-examiner

2018

Acknowledgments

I would like to heartily express my gratitude towards my supervisor Prof. Andreas Schenk. The long discussions with him taught me the essence of TCAD modeling, its applicability in the electronic device research and its limitations. He developed in me an ability to seek answers to the right questions while performing the analysis and draw proper conclusions. His constant emphasis on validation of the developed theories and TCAD models by comparison with the experimental data was crucial in keeping the analysis realistic. He provided me the freedom to decide the topics at each stage of Ph.D ranging from quantum transport calculations to *ab-initio* simulations for which I am eternally grateful to him.

I am grateful to Prof. Mathieu Luisier for the many discussions on the *ab-initio* analysis and also for placing the valuable computational resources of the Swiss supercomputing center at my disposal. Also, I would like to thank Dr. Petr Khomyakov for instigating my interest in *at-initio* density functional theory simulations and also for helping me with the simulation tool - VASP.

I would like to heartily thank Dr. Heike Riel, Dr. Kirsten Moselund, Dr. Heinz Schmid, Dr. Davide Cutaia, and Clarissa Convertino, all from IBM Research Zurich, for not only sharing the measured transfer characteristics of InAs/Si TFETs but also providing insights of the fabrication process which definitely helped the TCAD modeling efforts. I am especially thankful to Dr. Kirsten Moselund also for providing me an opportunity to work on the simulation of novel plasmonic and photonic lasers as an internship trainee at IBM Research Zurich. It opened a new field of research for me which I am keen to pursue.

I am also grateful to Prof. Lars-Erik Wernersson, Dr. Elvedin Memisevic, and Markus Hellenbrand, all from Lund University, Sweden, for sharing the transfer characteristics of the best TFET till date (in terms of on-current and the subthreshold swing) as well as providing the detailed TEM analysis of the materials at the hetero-junction which enabled us to perform reliable modeling of the devices.

I wish to thank my lab-mates Dr. Andreas Pedersen, Dr. Hamilton Carrillo-Nunez, Dr. Sascha Brücks, Christian Stieger, Paulina Aguirre, Pawel Lenarszyck, Kantawong Vuttivorakulchai, Anne Ziegler, Martin Rau, Hossein Bani-Hashemian, Teute Bunjaku, Cedrik Klinkert, and Dominik Bauer for the discussions on various topics including but not limited to our academic research and for constant encouragement in the pursuit of Ph. D.

In the end, I wish to thank my parents and brother for constantly supporting and encouraging me throughout my studies!

Abstract

Energy scaling of integrated circuits has hit a roadblock as the operating voltage of the MOSFET-based solid state switches has attained a minimum possible value. The thermionic emission mechanism, which governs the switching of the Metal-Oxide Semiconductor Field Effect Transistors (MOSFETs), does not allow to achieve subthreshold swing below 60 mV/decade at room temperature, thus establishing a lower limit on the operating voltage. Tunnel Field Effect Transistors (TFETs) which operate on the principle of field modulation of Band-to-band Tunneling (BTBT) can deliver a subthreshold swing less than 60 mV/dec. Therefore, TFETs are considered as a potential candidate to replace MOSFETs as solid-state switches to achieve further scaling of the supply voltage. However, the swing of a TFET is degraded by non-idealities such as traps, band tails, interface roughness, etc., which are inevitably present. To understand the effect of the non-idealities on the TFETs, physics-based models have been developed in this work and implemented in the Technology Computer Aided Design (TCAD) simulator Synopsys Sentaurus Device.

Simulations show that channel quantization reduces the on-current of the TFET which is confirmed by comparison with the experimental transfer characteristics of InAs/Si TFETs. Also, interface roughness and band tails are found to degrade the swing. TCAD analysis of experimental InAs/Si and all-III-V TFETs has confirmed that, maximum degradation of the swing results from the traps at the hetero-interface and the oxide interface. Our simulation set-up has achieved good agreement with the experimentally obtained temperature dependent as well as V_{DS} dependent transfer characteristics which confirms reliability of the set-up. Using this set-up, we have shown that, scaling of the

nanowire diameter below 20 nm and alignment of the gate with InAs/Si hetero-interface enables InAs/Si TFETs to deliver sub-60mV/dec swing even in the presence of traps.

Additionally, *ab-initio* modeling is performed to better understand the origin of traps at InAs/Si interface. It reveals that dangling bonds on As atoms at the interface are primarily responsible for the high D_{it}.

Zusammenfassung

Der Tunnel-Feldeffekt-Transistor (TFET) gilt als Alternative zum MOSFET, weil er eine weitere Skalierung der Versorgungsspannung erlauben würde. TFETs, die auf dem Prinzip der Feldmodulation des Interband-Tunnelns (Band-to-Band Tunneling, BTBT) basieren, können eine Steilheit der Sperrkennlinie (Swing) von weniger als 60 mV/dec liefern. Der Anstieg der Sperrkennlinie eines TFETs wird jedoch durch Nichtidealitäten wie Traps, Zustandsdichteausläufer, Grenzflächenrauhigkeit usw. verschlechtert, die zwangsläufig immer vorhanden sind. Um die Auswirkung dieser Nichtidealitäten auf die Leistungsfähigkeit von TFETs zu modellieren, wurden in dieser Arbeit physik-basierte TCAD-Modelle entwickelt und im TCAD-Simulator Sentaurus-Device implementiert. Die Simulationen zeigen, dass die Quantisierung der Zustände im Kanal den ON-Strom des TFETs reduziert, was mittels Vergleich mit den experimentellen Strom-Spannungs-Charakteristiken von InAs/Si-TFETs verifiziert wird. Es zeigt sich, dass Grenzflächenrauhigkeit und Zustandsdichteausläufer den Swing verschlechtern. Die TCAD-Analyse von experimentellen InAs/Si- und All-III-V-TFETs bestätigt, dass die Degradation des Swing hauptsächlich von Traps an der Hetero-Grenzfläche und an den Oxid-Grenzflächen herrührt. Die ab-initio-Modellierung der InAs/Si Grenzfläche führt zu dem Schluss, dass primär die nicht-abgesättigten Bindungen von As-Atomen an der Grenzfläche für deren grosse Zustandsdichte verantwortlich sind. Die Zuverlässigkeit des TCAD-Simulationsaufbau wird durch die gute Übereinstimmung mit den gemessenen temperaturabhängigen IV-Kennlinien bestätigt. Unter Verwendung dieses Aufbaus wird gezeigt, dass selbst im Falle hoher Defektdichten die

Skalierung des Durchmessers eines InAs/Si Quantendraht-TFETs unter 20 nm und die Vermeidung der Überlappung des Gates mit der Position der Hetero-Grenzfläche zu einem Swing kleiner als 60 mV/dec führen würden.

Contents

Chapter 1

Introduction

The past two decades (from the 90's to 2010) witnessed a strong growth of the integrated circuit industry fueled by exponential scaling of the dimensions of the solid-state switch - the MOSFET- which enabled a tremendous increase in the number of transistors per chip, increased operating speed, and the reduction of power consumption. Since the beginning of this decade, the growth of microelectronic industry has been driven by an ever increasing spread of portable electronic gadgets, such as laptops, smart-phones, smart-watches, e-readers, etc. [1]. The personal gadget industry demands low-leakage, power-efficient integrated circuits to reduce frequent charging of the devices. On the other hand, the designers don't wish to compensate on operating speed of these circuits. Additionally, there are other areas such as bio-medical instruments which require integrated circuits operating at low power. Overall, there is a need of low-power, low-leakage solid-state switching devices and circuits.

Up to 2010, the necessary power reduction was achieved by scaling the device dimensions and supply voltage, also called Dennard scaling [2]. This scaling of MOSFET dimensions was carried out in such a way that the electric field in the gate oxide remains the same, hence also known as "constant field scaling". However, this scaling concept has hit a road block as the operating voltage of the MOSFET-based solid-state switches has attained a minimum possible value. If one wishes to operate the circuits at a lower supply voltage at the same

frequency (i.e. keeping the on-current constant), the off-state leakage (and static power consumption) would increase exponentially with reduction in supply voltage. Conversely, in order to operate the device under low off-state leakage conditions, frequency needs to be significantly lowered.

This road-block in the "constant field scaling" is a consequence of the principle of operation of a MOSFET. In MOSFETs, electrons undergo thermionic excitation to travel from source to drain. Applying the gate bias gradually reduces the barrier allowing passage for more and more electrons. This emission mechanism does not allow to achieve a swing below 60 mV/decade at room temperature. For a sub-thermal swing, it is necessary to utilize a device that does not operate on thermionic emission. In order to make progress in power scaling, nano-scale solid-state switches based on various other switching mechanisms such as BTBT [3–5], impact ionization [6] and nano-electromechanical switching [7] have been proposed. The TFETs working on the principle of BTBT are considered a potential candidate to replace MOSFETs as solid-state switches.

A schematic diagram of one of the simulated TFETs is presented in Fig. 1.1(a). In a TFET, BTBT is controlled by the gate-induced electric field. At low gate bias, the hole states of the Valence Band (VB) in the source are not energetically overlapped with the electronic states of the Conduction Band (CB) in the channel region. As a result, BTBT at the source-channel interface is inhibited. Note, that although VB states in the source are energetically overlapped with the CB states in the drain, the spatial separation of nearly 80 nm prevents BTBT between the source and the drain. With increasing gate bias, accumulation begins in the channel region. A further increase in the gate voltage pushes the entire band edge down and a tunnel window opens at the source edge of the gate as shown in Fig. 1.1(c). Electrons tunnel from the source to the channel along tunnel paths parallel to the gate as illustrated in Fig. 1.1(b). In this way, energetic overlap of the VB states in the source with the CB states in the channel is initiated by the gate. However, TFETs suffer from an anomaly. If the gate bias is changed to a negative value, VB states in the channel can energetically overlap with the CB states in drain region giving rise to BTBT at the channel-drain interface. This is visible in the transfer characteristics as an increased drain current at $V_{GS} = 0\,V$

(see Fig. 1.1(d)). Device geometry, doping engineering, and selection of the appropriate semiconductor material can reduce ambipolar leakage.

In a TFET, the geometric alignment of the pn-junction with the gate-oxide/semiconductor interface has a profound effect on the BTBT direction and the tunnel rate. For example, in a device with the pn-junction parallel to the gate, the tunneling occurs via vertical tunnel paths (also called line tunnel paths) (Fig. 1.2(a)) [8,9]. On the contrary, when the pn-junction is perpendicular to the gate-oxide/semiconductor interface and the source is aligned with the gate, the BTBT occurs via tunnel paths parallel to the gate (also called point tunneling) (Fig. 1.2(b)). If the heavily doped source in the latter case has a large portion overlapped with the gate (Fig. 1.2(c)), vertically and laterally oriented tunneling paths exist at the same time [10, 11].

In the Gate-overlapped Source (GOS) TFET category, devices with different geometrical structures such as planar [8,9,12], NW-based [13] or FinFET-like [14] have been proposed. In a TFET, the BTBT rate is enhanced by a small direct band gap of the material which improves the on-state current of the TFET. Therefore, a small-gap material is preferred in the source region. However, the use of a small-gap material in the drain turns on BTBT there which results in a high drain current below the off-voltage (ambipolarity - see Fig. 1.1(d)). Using a wide-gap material in the drain region can suppress ambipolarity by hampering or inhibiting BTBT below the off-voltage. Thus hetero-structure TFETs with a small-gap semiconductor in the source and a wide-gap semiconductor in the drain region can provide a high on-current along with suppressed ambipolar leakage. The ambipolarity could also be suppressed by low drain doping for intermediate band gap materials such as Si. However, for small gap materials, lowering the doping in the drain region cannot suppress ambipolarity as seen in the transfer characteristics of $Si_{0.5}Ge_{0.5}$ homojunction TFETs reported in [14]. As the small gap materials are required for the source region, lowering the drain doping to suppress the ambipolarity might not be a viable alternative. The effect of the band gap on the swing is more involved. Due to the typical field dependence of the BTBT rate, the speed at which the BTBT rate rises with gate voltage becomes higher for a wide band gap. Therefore, the minimum swing worsens in a small-gap material. Additionally, the high intrinsic carrier density of the small-gap material in combination with TAT can result in a high

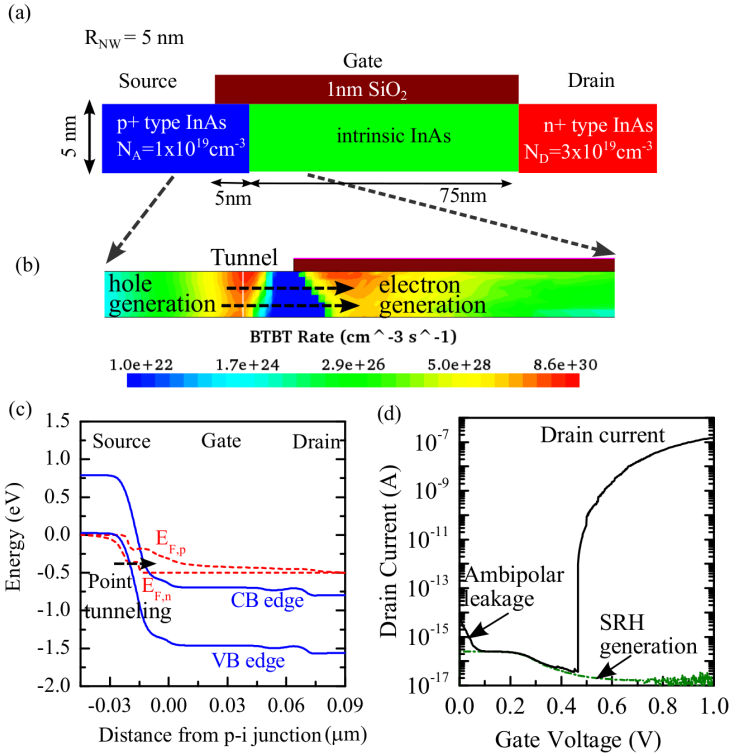


Figure 1.1: (a) Axial cross section of a cylindrical NW TFET. (b) Color-mapped diagram of the BTBT rate in the axial cross section. (c) Band diagram along the axis of the NW showing that the tunnel window is opened at the source edge of the gate. (d) Transfer characteristics of the TFET. Leakage current resulting from SRH generation is also plotted.

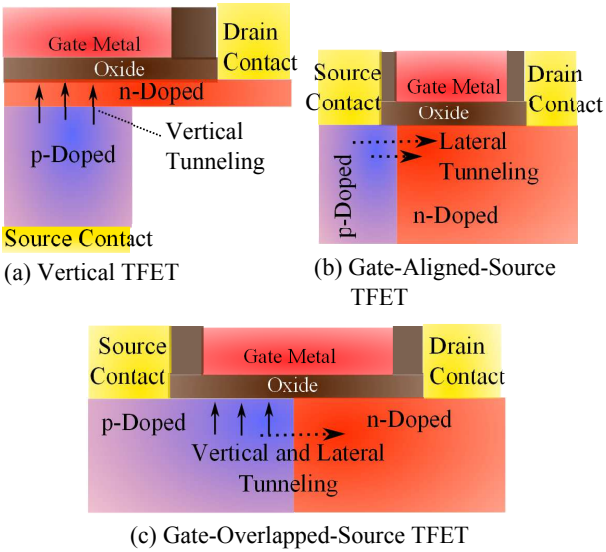


Figure 1.2: (a) In a vertical TFET, a narrow n-doped layer beneath the gate ensures that tunneling is predominantly vertical. (b) In a narrow NW TFET, when the source is nearly aligned with the gate such as in the device of Fig. 1.1(a), lateral tunneling dominates. (c) When the pn-junction is not aligned with the gate, both vertical and lateral tunneling take place.

SRH generation rate which increases the leakage [15] as observed in Fig. 1.1(d).

Although simulations confirm that ideal hetero-TFETs can achieve sub-thermal swing [16], the fabrication of such switches with sufficient on-current and sub-thermal swing over a few decades of drain current is difficult. In Chapter 3, we have shown that field-induced quantum confinement (also called channel quantization) severely degrades the on-current [17, 18]. Other non-idealities such as generation centers (for brevity called traps in this thesis), interface roughness, and band tails are known to increase the swing. Interface and bulk traps are most detrimental in this respect [19]. For instance, the large lattice mismatch between InAs (lattice constant $a_0 = 6.032\,\text{Å}$) and Si ($a_0 = 5.43\,\text{Å}$) makes strain relaxation by defect formation at the interface more favorable than gradual strain relaxation. Additionally, the growth of InAs on Si creates anti-phase domains in InAs [20]. The result is a high density of trap states (D_{it}) which could only be avoided by growing extremely narrow InAs NWs [21]. The lack of a native oxide on InAs causes a large D_{it} at the InAs/oxide interface [22, 23], although InAs channel MOSFETs with low D_{it} have been reported [24]. The TCAD analysis of InAs/Si TFETs (presented in Chapter 3) reveals that, due to a high D_{it} at the InAs/Si hetero-interface, thermionic emission can even become the dominant mechanism in InAs/Si pTFETs thus effectively converting them to MOSFETs. Similarly, as observed in the all-III-V TFETs and described in Chapter 4, the presence of even a single trap at the hetero-interface can degrade the subthreshold swing as well as give rise to device-to-device variability. Additionally, DOS tails in the heavily doped III-V semiconductors have their share of degrading the subthreshold swing. Due to such a strong unfavourable impact of the non-idealities on the device characteristics, it is necessary to take these effects into consideration while simulating the TFETs.

TCAD can help to understand the behavior of hetero-junction TFETs, and can give guidelines to improve their performance by optimization of geometry, doping, composition, gating, and biasing. The best approach is to use an atomistic, 3D quantum transport tool [16]. Unfortunately, the dimensions of real TFETs prohibit the application of such tools. Instead, drift-diffusion transport codes are widely used. In this thesis, all simulations are performed with the commercial device simulator Sentaurus-Device (S-Device) which

is equipped with various local and non-local BTBT models. New, physics-based TCAD models are developed in this thesis to capture the impact of certain physics phenomena (such as size and channel quantization, surface roughness, DOS tails), which are expected to play an important role in a TFET. After careful calibration of the parameters of these models by comparison with experimental data or with results of pseudo-potential calculations, TCAD simulations can be performed to predict trends in the device performance metrics (e. g. on-current, average subthreshold swing, etc.) when various device design parameters such as geometry, doping, or alloy composition can be changed. In this way, physics-based models along with a well-calibrated parameter set can be used to produce a carefully optimized device design.

However, the fitting of various model parameters of the BTBT model to the measured transfer/output characteristics may result in a flawed analysis if very general models are used. Due to the possibility that multiple parameter sets can yield the same good fit (perhaps because the model is too general), TCAD results based on fitted parameters cannot always be generalized, and the predictability remains limited. Therefore, minimum fitting parameters are used in order to make the TCAD study in this work more reliable. The nonlocal BTBT model depends on band gap, band alignments, and effective masses. All these parameters are set to their experimental values taking into account the position dependence of material composition, uniaxial stress, and crystal structure. This information is extracted from transmission electron microscopy (TEM) and other physical characterizations. Also, measured interface trap densities are used wherever available. Using minimum fitting parameters in this way would make the TCAD analysis more general and more reliable.

Chapter 2

Modeling Non-idealities in TCAD

Non-idealities in a TFET such as traps at the oxide/semiconductors interface or at the semiconductor hetero-interface, band tails in the semiconductor, field induced quantum confinement in the channel, and roughness of the oxide/semiconductor interface exhibit a strong influence on the TFET characteristics. To quantify the impact of the above non-idealities on the TFETs, it is necessary to develop physically accurate models which can account for the effect of non-idealities within the drift-diffusion formalism. Physics-based models which describe the non-idealities in the semiclassical approach are presented in this chapter.

2.1 Modeling Band-to-band Tunneling

Various local and nonlocal models are available to model BTBT in TCAD. Local models such as Kane's model or Schenk's model calculate the BTBT rate using the local electric field. A general expression for these models is given by,

$$G_{\mathrm{btb}} = A \cdot F^{p} \cdot \exp(-\frac{B}{F^{q}}) \tag{2.1}$$

where A, B are constants which are calibrated to I-V plots of Esaki tunnel diodes. The exponents are $(p, q) = (1, 2)$ in the Kane model [25] and $(p, q) = (1.5, 3)$ in the Schenk model [26]. In a general model, these exponents are taken as fitting parameters. The above models cannot take into account the effect of the non-uniform electric field on the tunnel rate. Also, the models do not check whether tunnel path between the CB and the VB exists which results in overestimation of the tunnel current particularly in the subthreshold region. As shown in Fig. 2.1, a tunnel path starting from the CB edge at $x = x_1$ ends at the VB edge. Hence, electrons tunnel from VB to the CB resulting in generation of holes at x_1. However, a tunnel path starting at $x = x_2$ on the VB edge does not reach the CB edge, thus, must be ignored. A calculation of the electron generation at x_1 using Eq. (2.1) would yield correct results since a uniform electric field is applied on the semiconductor (note that the expression in Eq. (2.1) is derived for a semiconductor under constant electric field). However, employing the same local expressions to calculate the generation rate at x_2 would over-estimate the tunnel current, as the local model will also include the tunnel rate along the path starting from x_2 in calculating the generation rate. This limitation of the above model has been corrected in the formalism developed in [27] for calculating the electric current in Esaki diodes. However, the treatment outlined therein is difficult to implement in the drift-diffusion formalism. Instead, a Dynamic Non-local Path (DNLP) BTBT model, which dynamically searches whether the tunnel path ends at the CB or not, is used to model BTBT in this work.

2.1.1 Dynamic non-local path BTBT model

In this model, an iteration is performed over all the vertices in the active region. At each vertex, the model dynamically determines the tunnel path that starts from the VB edge, continues along the direction of the electric field at the vertex, and ends at the CB edge. Thus, the energy of the tunnel path (E_{tun}) is equal to the VB edge at the vertex. The CB edge, the VB edge, and the effective masses are sampled at discrete points along the tunnel path. The tunnel path ends as soon as it reaches the CB edge i.e. when the CB energy at the head of the path is less than E_{tun}. The generation rate is calculated

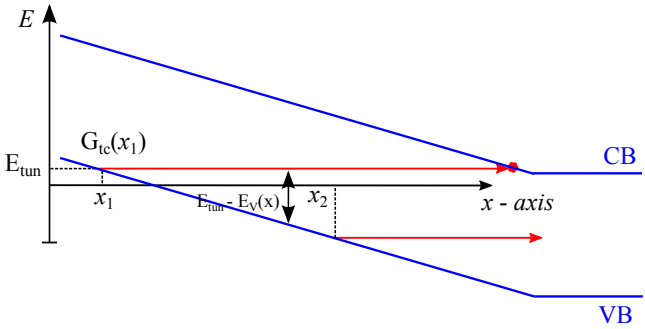


Figure 2.1: Tunnel path which starts from the VB edge at x_1 and ends at the CB edge. Another path starting from x_2 does not end at the CB edge, thus, must be ignored.

along the tunnel path using the following equation obtained by Kane's formalism [25, 28]:

$$G_{\mathrm{btb}} = \frac{g}{72\hbar} \frac{1 - \exp\left(-k_m^2 \int_0^L \frac{dx}{\kappa(x)}\right)}{\int_0^L \frac{dx}{\kappa(x)}} \exp\left(-2 \int_0^L \kappa(x) dx\right) (f_{\mathrm{c}} - f_{\mathrm{v}}). \tag{2.2}$$

Here, x is the position along the tunnel path, g is the degeneracy of the joint DOS (given by $g = 2g_{\mathrm{C}}g_{\mathrm{V}}$, where g_{C} and g_{V} denote, respectively, degeneracies of the CB and the VB, excluding spin degeneracy factor), F is the electric field, L is the length of the tunnel path, $\kappa(x)$ denotes the imaginary wave vector, f_{c} and f_{v} are, respectively, the quasi-Fermi distribution functions of electrons and holes. $\kappa(x)$ is calculated using Kane's two band dispersion model throughout this work, unless specified otherwise. An intuitive explanation of the above model can be found in the appendix of Ref. [29].

2.1.2 Comparison with OMEN data

To confirm that the above TCAD model reliably calculates the BTBT rate, a narrow InAs/GaSb nanowire (diameter of 3 nm) GOS TFET is

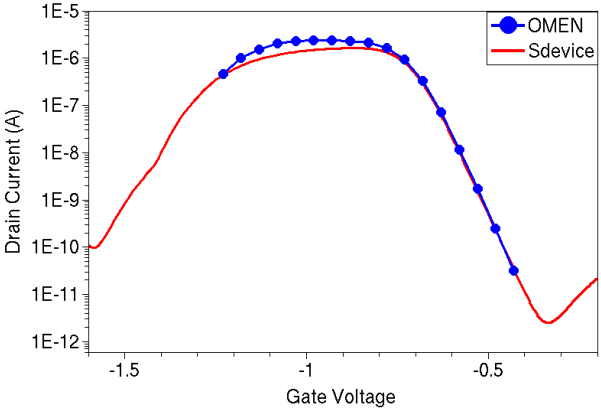


Figure 2.2: Comparison of the transfer characteristics of an InAs/GaSb nanowire ($d = 3\,\mathrm{nm}$) TFET simulated with S-device and the ones obtained by the full-band $\mathrm{sp^3d^*}$ tight-binding quantum transport simulator OMEN.

simulated with S-Device using the above model. The p-TFET consists of a heavily n-doped InAs source ($N_\mathrm{D} = 4 \times 10^{19}\,\mathrm{cm^{-3}}$), and intrinsic GaSb as channel and drain. The gate is 80 nm long and half of the gate is overlapped with InAs source making it a GOS geometry. Since the device is radially symmetric, an axial cross-section of the device is simulated using cylindrical coordinates. Band gaps and the effective mass values of InAs and GaSb nanowires with $d = 3\,\mathrm{nm}$ are different from bulk values due to geometrical confinement. They have been obtained from $\mathrm{sp^3d^*}$ tight-binding calculations and plugged into the DNLP model. Band gaps of InAs and GaSb nanowires with $d = 3\,\mathrm{nm}$ are 0.767 eV and 1.028 eV, respectively. The electron effective mass in the InAs nanowire ($d = 3\,\mathrm{nm}$) is 0.058 $\mathrm{m_0}$ while the hole effective mass in the GaSb nanowire ($d = 3\,\mathrm{nm}$) is 0.22 $\mathrm{m_0}$. Since electrons tunnel from the VB of GaSb to the CB of InAs, only the electron effective mass in InAs and the hole effective mass in GaSb are needed.

Transfer characteristics of the TFET simulated by S-Device using the DNLP model are compared with those obtained from the full-band $\mathrm{sp^3d^*}$ tight-binding quantum transport simulator OMEN [30]

in Fig. 2.2. The good agreement between the two confirms that the DNLP BTBT model can correctly calculate the BTBT rate. Such a match is not unexpected for the following reasons. The TCAD model uses Wentzel-Kramers-Brillouin (WKB) approximation and performs an integration over the imaginary dispersion relation obtained using Kane's model in the calculation of the tunnel probability. The imaginary dispersion obtained using tight-binding calculation in OMEN yields excellent agreement with that obtained by Kane's two-band model, particularly in small band gap III-V semiconductors [31]. In this way, the effects of reduced tunnel gap at the heterojunction, non-uniformity of electric field, etc. are well accounted. Also, the effect of geometrical confinement is covered by changing the masses and the band gaps.

The transfer characteristics of the p-channel TFET show interesting shape with the device turning off at two points. Near zero bias, BTBT at the hetero-interface is suppressed. Residual tunneling happens at the drain-side edge of the gate which results in the ambipolar leakage current. As the negative gate bias increases, holes accumulate in the i-GaSb channel and BTBT begins at the InAs/GaSb interface. With the increasing gate bias, the tunnel rate increases and then saturates at $V_{GS} = -1\,V$. A further increase of the gate bias results in volume inversion of the heavily doped InAs source region overlapped with the gate. As a result, the electric field at the heterojunction gradually decreases which quenches BTBT at the hetero-interface. Consequently, the drain current is reduced on increasing the negative gate bias. Thus, the TFET exhibits a negative trans-conductance region. This effect has also been observed in experimental TFETs as described in Chapter 4.

2.2 Modeling the Effect of Channel Quantization

In a TFET, BTBT can take place normal to the gate which is often called line tunneling. Line tunneling begins when the device is under strong inversion. Channel quantization results in a delay in the onset of line tunneling as well as a reduced strength of tunneling. The delay in the onset is due to the effective increase of the tunnel gap in the

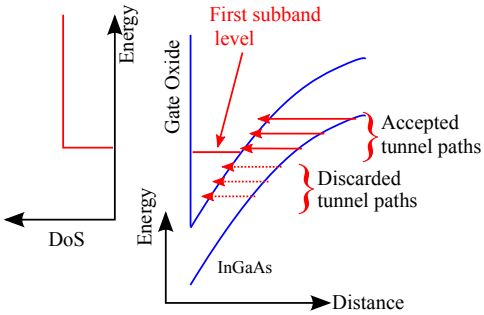


Figure 2.3: Schematic representation of the semiclassical model for channel quantization. The band diagram along a line perpendicular to the gate is shown along with the possible tunnel paths. The effect of the step-like DOS in the channel is modeled by rejecting tunnel paths below the first sub-band level.

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 farther away from the oxide. As a result, the active tunnel paths are longer than the forbidden ones which reduces the strength of the tunnel current. This effect can be taken into account in TCAD simulations by the following models.

2.2.1 Modification of the conduction band edge

A simple method to model the quantization effect within the semiclassical framework has been proposed in Ref. [32]: Tunnel paths with an energy above the lowest sub-band level are accepted while those with a lower energy are rejected, as illustrated in Fig. 2.3. The above technique is implemented in S-Device by using the quantum potential correction available in S-Device. It is possible to find a quantum potential that mimics the quantization effect by creating an effective

band gap in the inversion layer. In this model using the apparent band-edge shift Physical Model Interface (PMI) in S-Device, the value of the quantum potential adjacent to the gate dielectric is calculated as

$$E_{\mathrm{QM}} = \begin{cases} (\frac{q^2F^2\hbar^2}{2m_e})^{(1/3)}|a_1| - d_{\mathrm{ox}} \cdot F & \text{if } E_{\mathrm{QM}} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

where a_1 is the first zero of the Airy function, q is the magnitude of the elementary charge, m_e is the electron effective mass, $F(x, y)$ is the local electric field perpendicular to the nearest oxide interface, and d_{ox} is the distance from the oxide interface. The quantum potential is set to zero when its value is negative. It is then added to the CB edge which results in an effective CB edge

$$E_{\mathrm{CB}}^{\mathrm{eff}} = E_{\mathrm{CB}} + E_{\mathrm{QM}}. \tag{2.4}$$

It is obvious that the almost horizontal effective CB edge is equivalent to an explicit rejection of all energy levels below the lowest sub-band energy. As the contributions of higher sub-bands are small because of their weak population and a larger tunnel gap, they are neglected. In the case of a constant electric field, this approach is exactly the same as the technique suggested in Ref. [32]. In real devices, the vertical component of the electric field is never constant in the channel region. Hence, a moderate variation of the CB edge profile near the gate oxide interface is observed.

Application of the model

The above model is employed in the simulation of a double-gate GOS TFET with $In_{0.53}Ga_{0.47}As$ as a source material and InP as a channel material using S-Device [33]. The effect of the above described model on the transfer characteristics of the TFET is captured in Fig. 2.4(b). When the channel quantization model is deactivated, line and point tunneling begin at nearly the same gate bias point owing to the delay in the point tunneling due to the high CB offset [34]. After applying the model, line tunneling is delayed due to the channel quantization. Additionally, the strength of line tunneling is reduced which reduces the on-current of the TFET.

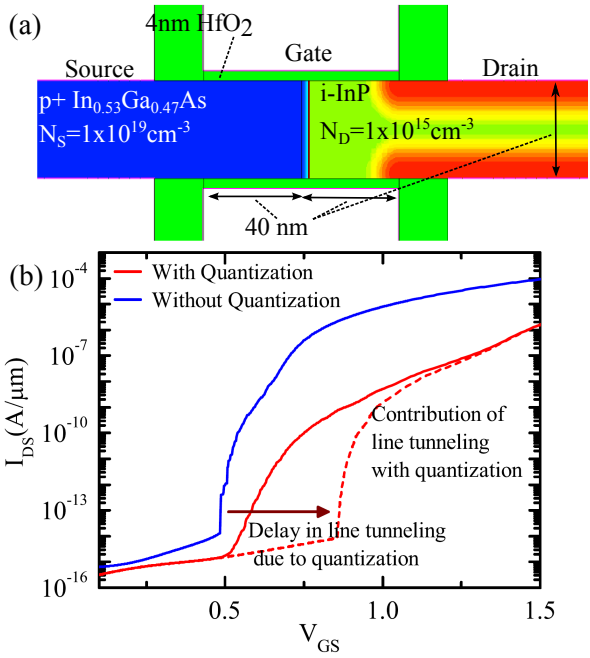


Figure 2.4: (a) Device geometry of the simulated n-type TFET. (b) Transfer characteristics of the TFET with and without channel quantization ($V_{\mathrm{DS}} = 0.5\,\mathrm{V}$). The dashed red line shows the contribution of line tunneling to the total current when the effect of quantization is considered. The difference between solid and dashed red lines is the contribution of point tunneling.

Note, that the alteration of the band edge to mimic channel quantization results in a redistribution of electrons which affects the electrostatics of the device. To avoid this effect, it is recommended to use the non-local model based on the path-rejection method described below.

2.2.2 Non-local model based on the path-rejection method

To account for quantization in the channel without affecting the device electrostatics, a non-local model based on the path-rejection method is developed. The model dynamically determines the tunnel path that starts from the VB edge, continues along the direction of the electric field, and ends at the CB edge. The tunnel path is then extended beyond the CB edge till it intersects the oxide/semiconductor interface. The component of the electric field parallel to the tunnel direction at the intersection point is used to calculate the energy of the first sub-band relative to the bottom of the triangular well at the interface using the expression,

$$E_0 = \left(\frac{\hbar^2 e^2 F_s^2}{m_e} \right)^{\frac{1}{3}}. \tag{2.5}$$

Here, F_s is electric field at the oxide/semiconductor interface, the remaining symbols have the same meaning as in Eq. (2.3). If the sub-band level is lower than the energy of the tunneling electron, the tunnel path is accepted and the generation rate of electron-hole pairs is introduced in the drift-diffusion equation. Otherwise, the tunnel path is rejected. The above description of the model applies for the case of channel quantization in n-channel TFETs and is illustrated in Fig. 2.3. In the case of p-channel TFETs, the tunnel path is extended along the tunnel direction beyond the VB edge till it intersects the oxide/semiconductor interface. Then, the algorithm proceeds as above. Whether the given tunnel path refers to a p-channel or a n-channel TFET can be determined “on the go” by searching for the band edge where the extension is finite. The equivalence of this simple model with a rigid quantum-mechanical calculation of the tunnel current based on Liouville’s equation was shown in [32].

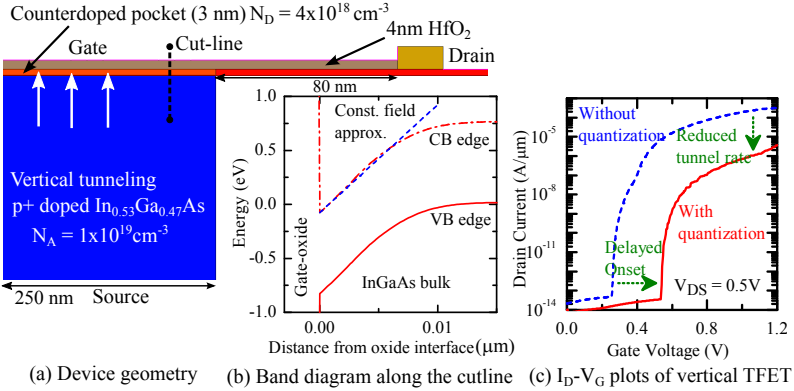


Figure 2.5: (a) InGaAs vertical TFET with counter-doped pocket. (b) Band edge diagram at $V_{GS} = 0.625\,\mathrm{V}$ along the cut line perpendicular to the channel. (c) Impact of channel quantization on the transfer characteristics of the vertical n-channel TFET.

A vertical TFET is simulated to study the impact of channel quantization. The device consists of p^+-doped $In_{0.53}Ga_{0.47}As$ bulk with a doping concentration of $1 \times 10^{19}\,\mathrm{cm}^{-3}$. A 3 nm thick counter-doped layer with n^+-doping of $4 \times 10^{18}\,\mathrm{cm}^{-3}$ is introduced in the channel to assist channel formation. The chosen geometry and doping scheme (shown in Fig. 2.5(a)) 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, the model based on the path-rejection method described above is employed. The sub-band energy is calculated using the triangular well approximation. The applicability of the latter is confirmed by comparing the linearized CB edge with the one obtained from self-consistent simulations at the onset of tunneling (see in Fig. 2.5(b)). The model is implemented in S-Device using the "nonlocal recombination PMI".

A comparison of the transfer characteristics of the above-described TFET including and excluding channel quantization is shown in Fig. 2.5(c). Inclusion of channel quantization delays the onset of line tunneling and reduces its strength in the on-state as expected.

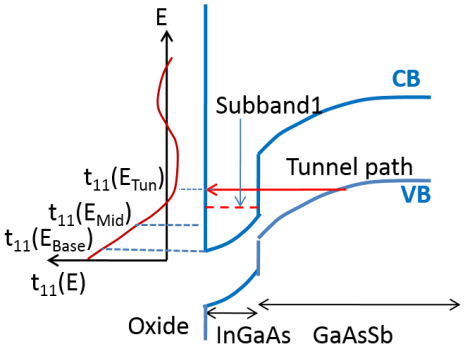


Figure 2.6: Schematic of the proposed semiclassical model to take into account the geometric confinement effect. The model is based on the path rejection technique and uses the transfer matrix method to check whether the tunnel path lies above the first sub-band level or not.

2.2.3 Transfer-matrix based model

In the non-local model based on the path-rejection method described in the previous subsection, the sub-band energy from the lowest of the triangular-like well is calculated by the triangular well approximation. This model is not applicable in the presence of a thick counter-doped layer, as the triangular well approximation may not be valid. In certain device geometries, confinement arising from hetero-junctions adjacent to the gate is superimposed on the triangular-like well. This changes the energy level of the first sub-band. The transfer matrix based model described below can account for the effects of quantization due to an arbitrary band-edge profile.

The implementation of the transfer matrix based quantization model for a one-dimensional GaAsSb/InGaAs hetero-junction vertical TFET is schematically given in Fig. 2.6. Tunnel paths which begin at the VB edge and end at the CB edge are determined from the band edge profile. The tunnel paths are then extended till the oxide interface. The transfer matrix is evaluated along the line between the starting point of the tunnel path and the oxide/semiconductor interface. It is evaluated at three energy points, namely, the tunnel path energy,

the energy of the CB edge at the oxide/semiconductor interface (i.e. bottom of the well), and at the middle of these two energy points. If the transfer matrix element $t_{11}(E)$ changes its sign between consecutive energy points, it implies that the sub-band level lies below the tunnel path energy (see Fig. 1(b)). In that case, the tunnel path is accepted. It is rejected otherwise. The above model requires the evaluation of the transfer matrix element at three energy points. Therefore, the model is computationally slightly more expensive than the one which uses the triangular well approximation. The above-described model is similar to the path rejection method except that the sub-band level obtained by the transfer matrix method is used instead of the one calculated using the triangular well approximation. Therefore, it can accurately model core-shell nanowire TFETs or bilayer TFETs.

In Ref. [35], the tunnel current of the pseudo-one-dimensional hetero-junction vertical TFET is obtained by the above-described transfer matrix based model and compared with the quantum mechanical calculation of the tunnel current by the Landauer formalism. The good agreement between the semiclassical and the quantum mechanical calculations confirms the validity of the model.

Two-dimensional simulations

In order to implement the semiclassical model based on the transfer matrix method for 2D device simulations, one needs to convert the 2D simulation problem to a number of 1D problems. This is achieved by extracting the tunnel paths from the 2D band edge diagram of the TFET. Each tunnel path starts from the CB edge, continues in direction of the electric field, and ends at the VB edge. This tunnel path is then extended by up to 8 nm beyond its end-point. If the extension intersects the oxide/semiconductor interface, the tunnel path is counted as a "line tunnel path". Otherwise it is considered as a "point tunnel path". The tunnel path is also extended by up to 8 nm before its starting point. This extension-length of the tunnel path is used-defined. Here, it is set to 8 nm, because the thickness of the inversion layer is believed to be in that range. Any tunnel path starting/ending more than 8 nm away from the oxide/semiconductor interface is certainly above the first sub-band in the channel. Therefore, it must always be accepted. This technique of thresholding of the extension-length avoids

the computational overhead of calculating the transfer matrix. The channel quantization model is not applied on point tunnel paths as point tunneling is not affected by channel quantization here. On line tunnel paths it is applied as follows: If the extension of the tunnel path intersects the oxide/semiconductor interface, the band diagram and effective masses along the tunnel path are extracted. These quantities are then used to evaluate the transfer matrix at the three energy values as explained earlier. Based on whether $t_{11}(E)$ changes its sign, the tunnel path is either selected or rejected. If the tunnel path is selected, the tunnel current is evaluated using Eq. (2.2). For the calculation of the imaginary dispersion relations, Flietner's two-band model [36] is used.

In order to reduce the computational burden, the channel quantization model based on the path rejection method is applied only in those regions where the tunnel paths are likely to intersect the oxide-semiconductor interface. The dimensions of these regions have to be specified by the user.

The above-described implementation of the channel quantization model based on transfer matrix method is employed for the semiclassical simulation of InAs/Si bilayer TFETs in Figs. 2.7(a-1), 2.7(b-1), and 2.7(c-1). The same TFETs had been simulated using the full-band quantum transport simulator OMEN in an earlier work [37]. There, the electrostatic potential, the effective masses, and the band gap were obtained from OMEN simulations at each bias point [37]. The same electrostatic potential is used here to extract semiclassical tunnel paths in the above devices to confirm that the transfer matrix based quantization model works in 2D as well. That is, instead of using S-Device to self-consistently simulate the devices and to calculate the generation current, the electrostatic potential obtained from OMEN is directly used. The CB edge and the VB edge throughout the TFET is obtained and tunnel paths are extracted from it. In the next step, the transfer matrix based channel quantization model is applied on these paths to calculate the BTBT rate by

$$G_{\mathrm{cv}}(x, y, E_{\mathrm{tun}}) = G_{\mathrm{btb}} \cdot \Theta(E_{\mathrm{tun}} - E_0). \tag{2.6}$$

Here, the generation rate G_{btb} is calculated at a point (x, y) using Eq. (2.2), $\Theta(x)$ denotes the Heaviside function, and E_0 is energy of

the first sub-band. The additional factor $\Theta(x)$ in the above equation implies that the generation rate is calculated only if the tunnel path lies above the first sub-band energy. This factor is equivalent to the path-rejection based on transfer matrix method described above. The generation rate at each vertex is summed over the entire device area to calculate the tunnel current. Only the tunnel paths with energy above the first sub-band level are selected in the integration. This results in,

$$I_{\mathrm{D}}(V_{\mathrm{GS}}) = \int_0^T \int_0^W G_{\mathrm{cv}}(x, y, E_{\mathrm{V}}(x, y))dxdy, \tag{2.7}$$

where, I_{D} is the drain current per unit width of the device. The above procedure implicitly assumes ballistic transport of the carriers in the device. The same approximation had been used in the full-band OMEN simulations.

Transfer characteristics obtained from OMEN and the semiclassical simulation are presented in Figs. 2.7(a-2), 2.7(b-2), and 2.7(c-2). Contributions from line and point tunneling have been separated in the semiclassical simulations and are also shown along with the total drain current. The good agreement with the OMEN suggests that the proposed model is able to take into account the effect of channel quantization and geometric confinement within a semiclassical framework. The comparison of individual contributions of line and point tunneling implies that point tunneling is the major contributor in TFET-A, although the contribution of line tunneling cannot be ignored. In TFET-B, point tunneling is the major contributor to the drain current in the subthreshold region while line tunneling begins to dominate at $V_{\mathrm{GS}} \approx 0.7\,\mathrm{eV}$ resulting in a kink in the transfer characteristics. Such a kink is also observed in the OMEN results at nearly the same gate bias. In TFET-C, tunneling is exclusively line tunneling, due to geometric restrictions on the tunnel direction. Note, that the blue circles at the bottom-left corner of Fig. 2.7(c-2) denote contribution by point tunneling. Also, the dashed-blue line denoting contribution of line tunneling is missing, since it merges with the solid-blue line.

In the subthreshold region, semiclassical and OMEN transfer characteristics do not match well as observed in Fig. 2.7(a-2) and Fig. 2.7(b-2). In this region, the current is dominated by point

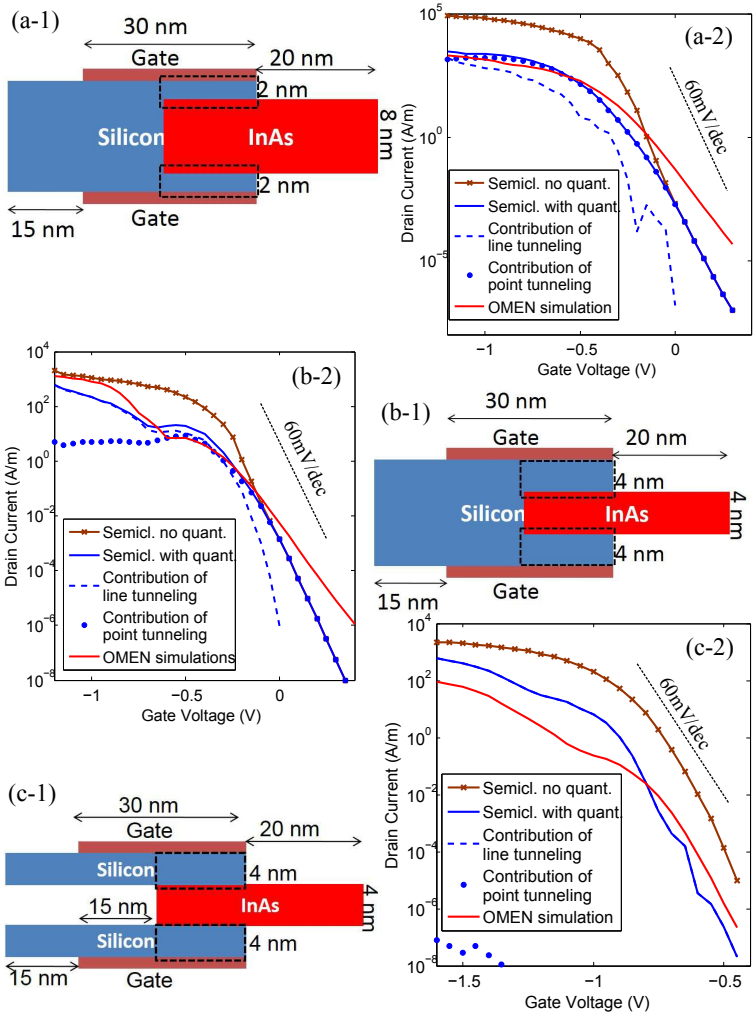


Figure 2.7: (a-1, b-1, c-1) 2D device geometries of InAs/Si bilayer TFETs simulated with both the full-band quantum transport solver OMEN and the semiclassical model in S-Device. The box in each figure shows the region where the channel quantization model is activated. (a-2, b-2, c-2) Comparison of transfer characteristics simulated by OMEN with the ones simulated using the semiclassical model based on the transfer matrix method. ($V_{DS} = 0.5\,\mathrm{V}$)

tunneling for which the proposed semiclassical model is not applicable. A possible explanation for the discrepancy could be the following. Only the Γ-point is used in the OMEN simulations of TFET-A and TFET-B to save CPU time. This simplification results in an over-estimation of the drain current in the subthreshold regime of these devices. In the OMEN simulation of TFET-C, a multiple k-point grid is used giving a smaller swing which better agrees with the semiclassical curve. Furthermore, the self-consistent electrostatics from OMEN is used in the semiclassical model. Using the electrostatics from S-Device (Poisson and drift-diffusion equations) instead would lead to some small differences.

2.3 Impact of Surface Roughness on TFETs

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. A TCAD analysis of the effect of interface roughness in quantum mechanical as well as in semiclassical simulations by a spatial variation in the z-position of the infinitely high potential wall is presented here. This spatial variation is accompanied by a random shift of the sub-band level which smoothens the otherwise staircase-shaped DOS of the Two-dimensional Electron Gas (2DEG) in the inversion layer. Band tail states appear below the sub-band energy as shown in Fig. 2.8(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 [38]. According to Ref. [38], the DOS of a 2D system in the presence of a random potential $U(r)$ is given by

$$\varrho(E-E_0) = \frac{m^*}{\pi\hbar^2}\left\{\frac{1}{2}\left[1+\mathrm{erf}\left(\frac{E-E_0}{\sqrt{2}\eta}\right)\right] - \frac{\hbar^2\mathcal{F}^2}{24\sqrt{2\pi}m^*\eta^3}\left[1-\frac{(E-E_0)^2}{\eta^2}\right]\exp\left[-\frac{(E-E_0)^2}{2\eta^2}\right]\right\}, \quad (2.8)$$

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

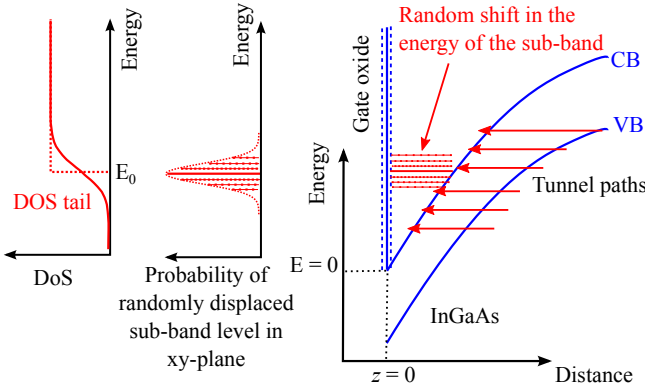


Figure 2.8: Interface roughness results in a random shift of the sub-band level which on integration gives rise to DOS tails. A semiclassical model of roughness-induced DOS tails on tunneling is schematically shown. All the tunnel paths are accepted and the BTBT rate along each path is multiplied with the DOS factor.

strength of the field associated with the random potential ($\nabla U(r)$). These quantities are related to the potential by

$$W(r-r') = \langle U(r)\cdot U(r')\rangle$$
$$\eta^2 = \langle U(r)^2\rangle = W(r-r')|_{r=r'} \tag{2.9a}$$
$$\mathcal{F}^2 = \langle(\nabla U(r))^2\rangle = \nabla_r\nabla_{r'}W(r-r')|_{r=r'}. \tag{2.9b}$$

In Eqs. (2.9a), (2.9b), r and r' are vectors in the xy-plane. To adapt the expression in Eq. (2.8) for modeling the DOS tails arising from interface roughness, it is necessary to determine the random potential associated with the roughness. It is obtained from Ando's formalism [39] and is given by

$$U(r) := \nu_{\mathrm{SR}}(r) = \int dz\ \zeta_0^*(z)\cdot V_{\mathrm{SR}}(z,r)\cdot\zeta_0(z)\ . \tag{2.10}$$

Here, $\zeta_0(z)$ is the wave function of the ground state in the triangular well and $V_{\mathrm{SR}}(z,r)$ the perturbation operator associated with interface

roughness as defined in Ref. [39]. If the interface roughness is assumed to have Gaussian correlation, then

$$\langle \Delta(r) \cdot \Delta(r') \rangle = \Delta^2 \exp\left[-\frac{(r-r')^2}{L^2} \right] \tag{2.11}$$

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 interface roughness potential ($\nu_{\mathrm{SR}}(r)$) is derived in Ref. [39]:

$$\nu_{\mathrm{SR}}(r) = \int dk\ \Delta(k)\Gamma(k) \exp(ik \cdot r)\ . \tag{2.12}$$

$\Delta(k)$ 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, 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:

$$\Gamma(k) = \frac{4\pi e^2}{\kappa_{\mathrm{s}}} \left(N_{\mathrm{depl}} + \frac{1}{2} N_{\mathrm{inv}} \right) \approx F_{\mathrm{eff}}\ . \tag{2.13}$$

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. (2.13) in Eq. (2.12) one obtains for the interface roughness random potential

$$\nu_{\mathrm{SR}}(r) = F_{\mathrm{eff}}\, \Delta(r)\ . \tag{2.14}$$

Inserting this into Eqs. (2.9a) and (2.9b) gives the following results for η and $\mathcal{F}$:

$$\eta = F_{\mathrm{eff}}\, \Delta \tag{2.15a}$$

$$\mathcal{F} = F_{\mathrm{eff}}\, \frac{\Delta}{L}\ . \tag{2.15b}$$

Here, Δ is the amplitude and L the auto-correlation length in the Gaussian correlation function describing interface roughness (Eq. (2.11)). By definition, the effective electric field F_{eff} is calculated as

$$F_{\text{eff}}^{\text{e}}(r) = \frac{\int_0^{z_{\max}} dz\; n(r,z) F_{\perp}(r,z)}{\int_0^{z_{\max}} dz\; n(r,z)}\;, \tag{2.16}$$

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 $F_{\text{eff}}^{\text{e/h}}$ has been used in the numerical implementation of the model for DOS tails as described below. It has to be noted that $F_{\text{eff}}^{\text{e/h}}$ is bias-dependent. Thus, η and $\mathcal{F}$ defined in Eqs. (2.15a) and (2.15b) are also bias-dependent. Their values must be updated at each bias point during the simulation.

2.3.1 Quantum mechanical calculations

In this sub-section, the impact of interface roughness on the steepness of transfer characteristics of the vertical TFET is analyzed by quantum mechanical calculations. As mentioned before, the roughness of the interface (assumed to be in the xy-plane) causes a random shift of the infinite potential wall of the triangular well along the z-axis. This results in a random shift of the first sub-band level around its mean energy as shown in Fig. 2.10(a). Therefore, energy of the first sub-band level is different at different locations in the xy-plane. If the sub-system at each location in the xy-plane forms a 2D system, the DOS of the quantized electronic states is step-like. The average DOS of the total system is obtained by taking the weighted average of the DOS of each of such localized sub-systems at each location in the xy-plane. The weights are the probability of existence of a sub-band level at a given energy. For the rough interface with Gaussian correlation, the spread of the sub-band levels is a Gaussian with the center at the mean energy of the first sub-band as shown in Fig. 2.10(a). Averaging of the DOS as described above with a Gaussian as a weighting function results in the average DOS with the shape of an error-function as illustrated in Fig. 2.10(b). The aggregate effect of this random shift of the sub-band level is the formation of band tail states below the mean

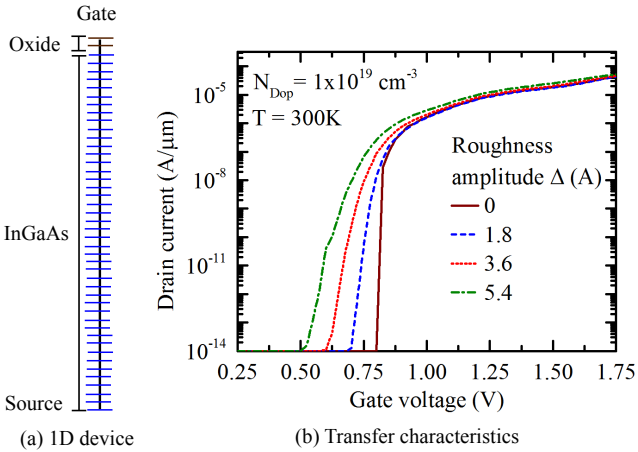


Figure 2.9: (a) One-dimensional TFET used for quantum mechanical calculations. (b) Comparison of transfer characteristics of the uni-dimensional TFET for different values of the roughness amplitude.

energy. Quantum mechanical calculations are performed to calculate the tunnel current for sub-band edges at different energies adjacent to the energy of the first sub-band. The drain current is calculated by an ensemble average of these tunnel currents as described below.

The device geometry of the vertical TFET shown in Fig. 2.5(a) favors line tunneling and, therefore, it is strongly influenced by surface roughness. Since quantum mechanical calculations of such bulk-like TFETs are computationally expensive, a pseudo-one dimensional TFET along the cut-line in Fig. 2.5(a), also shown in Fig. 2.9(a), is simulated using a self-consistent 1D k·p-Poisson solver available in the S-band package [40]. For each gate bias, the energy levels and the quantized electron wave functions of the first sub-band are obtained from the solver. The Numerov method is used to calculate the VB wave functions in the bulk for all $E^{\perp}$ in the tunnel window. The tunnel probability is calculated for each $E^{\perp}$ using Fermi's golden rule. The total drain current per unit gate area is then calculated by using the Landauer Formalism. The entire procedure is described below.

The triangular potential well in the channel results in the quantization of the CB states. These states are obtained by solving the following envelope equation numerically using the 1D k·p-Poisson solver [40]:

$$\left(-\frac{\hbar^2}{2m_c}\cdot\frac{d^2}{dx^2}+U(z)\right)\cdot\chi_{c,n}(z)=E_n\cdot\chi_{c,n}(z) \tag{2.17}$$

In Eq. (2.17) it is assumed that the device is homogenous in the xy-plane. Hence, x- and y-dependent components of the envelope function are given by plane waves. This adds an additional energy term to the total energy. The total energy thus becomes $E_{n,\perp}=E_n+E^{\perp}$.

Since VB states are not quantized, an envelope function exists for each available value of $E^{\perp}$. The envelope functions of the VB states were obtained solving the following equation using the Numerov algorithm:

$$\left(-\frac{\hbar^2}{2m_v}\cdot\frac{d^2}{dx^2}+(U(z)-E-E')\right)\cdot\chi_v(z)=0. \tag{2.18}$$

Hard-wall boundary conditions at the oxide-semiconductor interface ($\chi_v(z=0,E^{\perp})=0$) are used for the wave function as required for the Numerov algorithm.

The envelope functions for the CB and VB states obtained by the above method are exact solutions of the Schrödinger envelope equation. The inter-band matrix element as a function of $E^{\perp}$ can be obtained by treating the field-dependent inter-band coupling term as perturbation,

$$M_{n,v}(E_{n,\perp},E')=\int_0^{\infty}\chi_{c,n}(x)\cdot\frac{\hbar\cdot P}{m_0\cdot E_g}\nabla U(x)\cdot\chi_v(x,E')dx\ . \tag{2.19}$$

Here, P is the momentum matrix element and E_g the band gap. The tunnel probability can then be calculated using Fermi's golden rule,

$$T(E')=\frac{2\pi}{\hbar}\int_{E_n}^{E_{max}}|M_{n,v}(E_{n,\perp},E')|^2\delta(E_{n,\perp}-E')dE_{n,\perp}\ . \tag{2.20}$$

Note that although E_n is discrete, $E_{n,\perp}$ is continuous as it includes the transverse energies. The above integral implies that $E'>E_n$ for

tunneling to take place. The total electron current can be calculated by integrating the tunnel probability over all available transverse energies,

$$j_{\text{tun}} = \int_0^{E_{max}} T(E'^{\perp})\rho(E'^{\perp})dE'^{\perp}. \tag{2.21}$$

The current density (j_{tun}) obtained by this approach can be multiplied by the gate area to obtain the drain current. In this way, the drain current is calculated per unit area of one of the sub-systems in the xy-plane with the sub-band energy of E_0.

Quantum mechanical calculations of a 2D electronic system in the presence of an arbitrary random potential in 2D yield an error-function-like DOS which is in agreement with the classical DOS in Eq. (2.8) described above. The standard deviation (η) appearing in the Gaussian distribution and the error-function is calculated by inserting Ando's roughness potential [39] for the arbitrary random potential in [38]. It is given by Eq. (2.15a). Assuming that the tunnel rate can be independently calculated for each localized sub-system, the total tunnel current in the presence of surface roughness becomes,

$$I_{DS} = \frac{A}{\eta\sqrt{\pi}} \int_{-\infty}^{E_{max}} j_{\text{tun}}(E) \cdot \exp\left(-\frac{(E-E_0)^2}{2\eta^2}\right) dE. \tag{2.22}$$

Here, $j_{\text{tun}}(E)$ is the current calculated by the quantum mechanical treatment described in Eqs. (2.17)-(2.21) for a system with the sub-band edge lying at energy E. E_0 is mean energy of the first sub-band averaged over all the sub-systems in the xy-plane. In this way, the total drain current is calculated as ensemble average of the currents arising from the sub-bands at each of the energies E ($-\infty < E < E_{max}$) around the mean sub-band level.

A roughness amplitude Δ of 1.8 Å and a correlation length L of 1.9 nm was experimentally extracted for a InAs/GaSb heterojunction by Feenstra et al. [41]. Also, a roughness amplitude ranging between 5 Å and 11 Å has been reported for thin $In_{0.75}Ga_{0.25}As$ pHEMT structures [42]. Taking these values as guidelines, quantum mechanical calculations were performed with $\Delta = 1.8\,\text{Å}, 3.6\,\text{Å}, \text{and}\, 5.4\,\text{Å}$. The simulation results are presented in Fig. 2.9(b). As expected, an increasing roughness amplitude degrades the subthreshold swing.

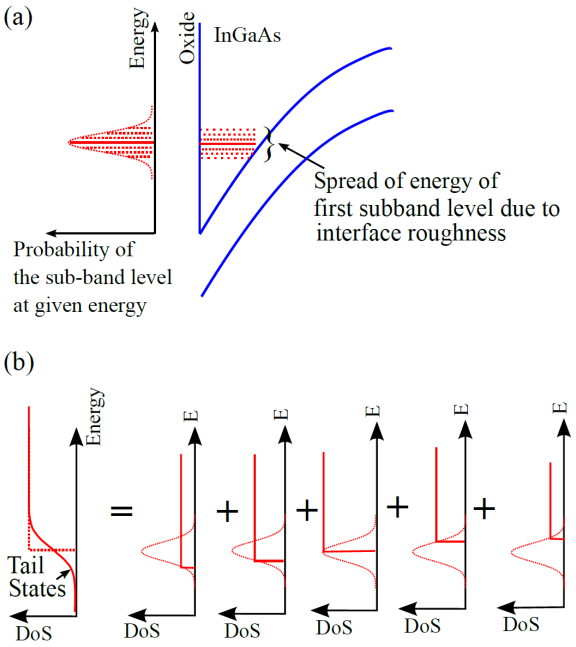


Figure 2.10: (a) Schematic showing Gaussian spread of the energy of the first sub-band level. (b) Sketch of the weighted average of the DOS of all the randomly shifted first sub-band levels resulting in an error-function-like DOS with the tail states below the mean sub-band level.

2.3.2 Semiclassical model

The channel quantization model presented in Section 2.2 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. 2.3). 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 (energy of the lowest sub-band) using the triangular well approximation, the effective electric field is computed for each tunnel path using Eq. (2.16) 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 points along the tunnel path. This effective electric field is the input for the calculation of the DOS parameters η and $\mathcal{F}$ for each tunnel path. From these parameters and the sub-band energy, one can calculate the normalized DOS factor $\overline{\varrho} = \frac{\pi\hbar^2}{m^*}\varrho$ with ϱ from Eq. (2.8). The tunnel rate for a tunnel path with energy E based on the implemented Kane model in Ref. [33] is multiplied by this normalized DOS factor calculated at the same energy:

$$G_{\mathrm{vc}}(E, x = x_{\mathrm{v}}) = \overline{\varrho}(E - E_0) eF(x_{\mathrm{v}}) \frac{g}{72\hbar} \times \frac{1 - \exp\left(-k_m^2 \int_{x_{\mathrm{v}}}^{x_{\mathrm{c}}} \frac{dx}{\kappa}\right)}{\int_{x_{\mathrm{v}}}^{x_{\mathrm{c}}} \frac{dx}{\kappa}} \cdot \exp\left(-2 \int_{x_{\mathrm{v}}}^{x_{\mathrm{c}}} \kappa dx\right) (f_{\mathrm{c}} - f_{\mathrm{v}}) \,. \tag{2.23}$$

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 amplitude of the imaginary wave vector, 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

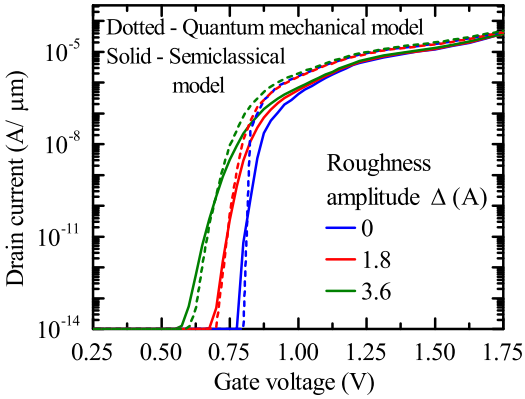


Figure 2.11: Comparison of the transfer characteristics of the one-dimensional device obtained by quantum mechanical calculations with the ones obtained using the semiclassical model.

the beginning and the end of the tunnel path, respectively. The above semiclassical model is solely phenomenological as an energetic tunnel rate for bulk states is averaged by a 2D DOS. To further validate the approach, quantum mechanical calculations are performed on the (pseudo-) 1D device in Fig. 2.9. The transfer characteristics of the 1D device obtained by quantum mechanical calculations are compared with those from the semiclassical model described above in Fig. 2.11. The good agreement between the transfer characteristics obtained by the quantum mechanical calculations and the semiclassical approach supports the applicability of the semiclassical model.

In order to perform simulations of the 2D vertical TFET in Fig. 2.5(a), the semiclassical model needs to be implemented in a 2D device simulator. The model is implemented in S-Device [33] via the "nonlocal recombination PMI". The impact of surface roughness on the performance of the vertical TFET is analyzed using this implementation. Simulations are performed with the roughness amplitude $\Delta = 1.8\,\text{Å}, 3.6\,\text{Å}$ and the correlation length $L = 1.9\,\text{nm}$. A comparison of the transfer characteristics for different parameter values is presented in Fig. 2.12.

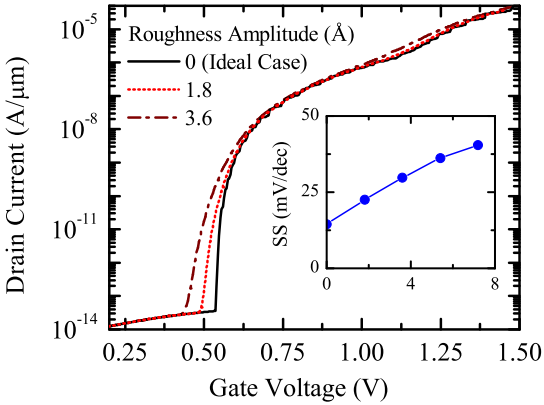


Figure 2.12: Transfer characteristics of the vertical TFET for different values of the roughness amplitude Δ. Increasing Δ degrades the subthreshold swing averaged over three decades of drain current above the off-current (plotted in the inset).

The DOS tails resulting from interface roughness smoothen the onset of line tunneling. An increasing roughness amplitude degrades the subthreshold swing. This is a consequence of the direct proportionality of η and $\mathcal{F}$ to the roughness amplitude Δ. The simulations show that a decreasing auto-correlation length L degrades the subthreshold swing. Eqs. (2.15a), (2.15b) as well as the simulations suggest that an increase in the roughness amplitude or a reduction in the auto-correlation length increase the swing of TFETs which are dominated by vertical BTBT.

It must be noted that the above models for channel quantization and interface roughness are applicable for vertical BTBT only. Lateral tunneling is expected to remain unaffected by the roughness of the oxide/semiconductor interface in a bulk-like TFET. In a narrow nanowire/ultra-thin-body TFET, surface roughness would cause a random localized change in the radius/ultra-thin-body-thickness of the channel, which could give rise to band tails in the source, channel, and the drain. A model described in the next section may be used to model the effect of such band tails on lateral tunneling.

2.4 Modeling the Impact of Density of States (DOS) tails

DOS tails, also called band tails, can originate from random dopant placement or the presence of defects, which gives rise to electronic states in the forbidden gap close to the band edges. Random placement of the dopant atoms creates, on ionization, randomly placed charge centers in the crystal lattice. This forms the band tail states which are highly localized in the warped region, but are strongly coupled to the nearest band edge. The origin of band tails has been studied in detail by many authors using different approaches [43–45].

In this section, a quantum-mechanical treatment is described to analytically calculate the tunnel rate of electrons from VB tail states into the CB. The analytical results are then used to modify a semiclassical BTBT rate which is implemented in the TCAD simulator. The impact of various parameters of the model on the TFET performance is studied for the case of a nanowire transistor.

2.4.1 DOS tail model

According to Kane's theory of band tail states [43], the DOS of a semiconductor in the presence of random dopant fluctuations takes the form

$$\varrho_{\mathrm{t,e/h}}(E) = \frac{(2m_{\mathrm{t,e/h}})^{3/2}}{2\pi^2\hbar^3}\sqrt{\eta}\, Y_{\mathrm{g/e}}(E/\eta)\ , \tag{2.24}$$

where, $m_{\mathrm{t,e/h}}$ is the effective mass of a carrier in the tail state, $\hbar$ is the reduced Planck's constant, and η is the characteristic energy of the band tail. In Eq. (2.24), the energy E counts from the respective band edge. In the case of Gaussian tails the function $Y(E/\eta)$ is given by

$$Y_{\mathrm{g}}(x) = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{x} d\zeta\, \sqrt{x-\zeta}\, e^{-\zeta^2} \tag{2.25}$$

which in the limit $x \to \infty$ turns into

$$Y_{\mathrm{g}}(x) \to \frac{1}{2^{5/2}x^{3/2}}\, e^{-x^2}\ . \tag{2.26}$$

Therefore, if $E \gg \eta$, $Y_{\mathrm{g}}(E/\eta)$ can be approximated by Eq. (2.26) and the tail DOS becomes

$$\varrho_{\mathrm{t,e/h}}(E) = \frac{(m_{\mathrm{t,e/h}})^{3/2}}{4\pi^2\hbar^3}\frac{\eta^2}{E^{3/2}}\exp\left[-(E/\eta)^2\right]\ . \tag{2.27}$$

Note that the approximation (2.27) cannot be used if $|E| < \eta$ or if E is an energy in the band. In this case Eq. (2.24) has to be applied which requires a numerical integration, or Eq. (2.26) is empirically modified such that it approximates $Y_{\mathrm{g}}(x)$ up to $x = 0$. A simple, but efficient modification is given by

$$Y_{\mathrm{g}}(x) \to \frac{e^{-x^2}}{2^{5/2}(x^{3/2}+s)} \tag{2.28}$$

with $s = 0.566$. The modified expression for the tail DOS then reads

$$\varrho_{\mathrm{t,e/h}}(E) = \frac{(m_{\mathrm{t,e/h}})^{3/2}}{4\pi^2\hbar^3}\frac{\sqrt{\eta}}{[(E/\eta)^{3/2}+s]}\exp\left[-(E/\eta)^2\right]\ . \tag{2.29}$$

In the case of exponential tails, the Gaussian function in the integrand of Eq. (2.25) is replaced by an exponential function and $Y(E/\eta)$ becomes

$$Y_{\mathrm{e}}(x) = \frac{1}{2}\int_{-\infty}^{x} d\zeta\, \sqrt{x-\zeta}\, e^{-|\zeta|} \tag{2.30}$$

which yields the correct ideal DOS in the limit $\eta \to 0$. For $x < 0$, the integral can be calculated exactly:

$$Y_{\mathrm{e}}(x) = \frac{1}{4\sqrt{\pi}} e^{-|x|}, \qquad x < 0. \tag{2.31}$$

For $x > 0$ (energies in the band), Eq. (2.30) has to be solved numerically. In Eq. (2.30), $Y_{\mathrm{e}}(x) \to \sqrt{x}$ for $x \to +\infty$.

All functions defining Gaussian and exponential tails are compared in Fig. 2.13(a). In the following, analytical models for the rate of tail-to-band tunneling in a constant electric field are derived based on the degree of localization of the tail states. They are denoted model-0 in the case of strongly localized tail states neglecting their field broadening, and model-1 in the case of field-broadened strongly localized tail states.

2.4.2 Tunneling from/into tail states ignoring their life time broadening

In this subsection tunneling from tail states at the VB into the ideal CB is considered. For the rate derivation it is assumed that they are sufficiently localized such that the pseudo-δ-potential model [46] can be used which enables to express the transition rate between localized states near the VB (energy E_{t}, spatial density $1/(2\pi r_0^3)$) and a CB state by [47]

$$D_{\mathrm{t,c}}(E, E', E_{\mathrm{t}}) = 8\pi \frac{\mathcal{P}}{\tilde{E}^2} E_{\mathrm{t}}^2 r_0^3\; \varrho_{\mathrm{t}}(\tilde{E}, E_{\mathrm{t}})\; \varrho_{\mathrm{c}}(E')\ . \tag{2.32}$$

Here, $\mathcal{P}$ is the Cauchy principal value of integrals over $\tilde{E} = E + E_{\mathrm{g}}$, and $\varrho_{\mathrm{t}}(\tilde{E}, E_{\mathrm{t}})$ denotes the density of the localized single-level states

$$\varrho_{\mathrm{t}}(\tilde{E}, E_{\mathrm{t}}) = \frac{1}{2\pi r_0^3}\delta(\tilde{E} - E_{\mathrm{t}})\ . \tag{2.33}$$

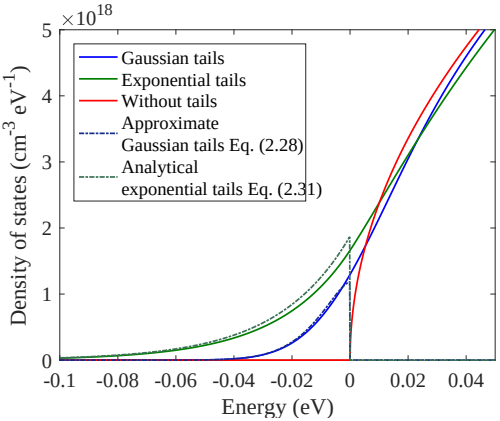


Figure 2.13: Comparison of the conduction band DOS in InAs with Gaussian and exponential tails calculated by the exact expression (2.24) and by various approximations.

The energy level E_t of the tail state is measured from the VB edge. The different energy variables in Eq. (2.32) and their meaning in the derivation below are presented in Fig. 2.14.

For Eq. (2.33) the field effect on the localized state was neglected [47]. It will be included in the following subsection. Note that Eq. (2.32) is a special case of Eq. (2.50) which is derived in Appendix A of Ref. [48]. The "effective" mass m_t of the localized electron is simply related to its localization radius r_0 by

$$E_t = \frac{\hbar^2}{2m_t r_0^2} \ , \tag{2.34}$$

which can be viewed as fitting of the localization radius r_0. The mass m_t appears as result of the single-band envelope method applied to the localized state, a method limited to a single, parabolic and isotropic band. This parameter takes account of the presence of heavy and light holes and the anisotropy of their bands. Deeper tail states might even be affected by the conduction bands. Hence, m_t becomes a parameter of the analytical model.

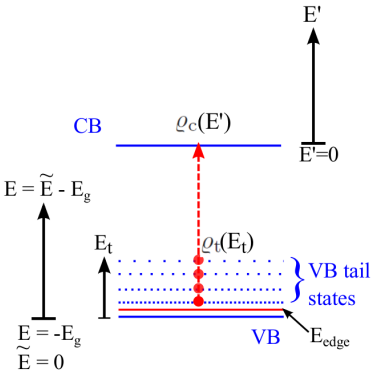


Figure 2.14: Representation of the energy vaiables used in Eq. (2.32) and thereafter.

The tail DOS Eq. (2.24) (with E replaced by $-\tilde{E}$) is assumed to be a dense ladder of single-level DOS Eq. (2.33) and will be composed with weight functions $w(E_\mathrm{t})$ as

$$\varrho_{\mathrm{t,h}}(\tilde{E}) = \int_{E_\mathrm{edge}}^{E_\mathrm{g}} dE_\mathrm{t} w(E_\mathrm{t}) \varrho_\mathrm{t}(\tilde{E}, E_\mathrm{t}) \ . \tag{2.35}$$

The integration over tail states is restricted to $E_\mathrm{edge} < \tilde{E} < E_\mathrm{g}$ (see Fig. 2.14). The lower limit E_edge separates localized states from continuum states, and it is assumed that $0 < E_\mathrm{edge} < \eta$. Thus, E_edge plays the same role as the 'mobility edge' [49] in transport. The mass $m_{\mathrm{t,h}}$ from Kane's DOS model could be set to the hole mass m_h, as done by Kane [50], or to m_t which is preferred here (see discussion above). The model has then just *one* fitting parameter to account for the unknown electronic structure of the tail states. However, it's value cannot differ vastly from that of a hole mass.

The weight function $w(E_\mathrm{t})$ immediately follows from inserting the DOS expressions into above equation:

$$w(E_\mathrm{t}) = \frac{\sqrt{\eta}}{\pi E_\mathrm{t}^{3/2}} Y\left(E_\mathrm{t}/\eta\right) \ . \tag{2.36}$$

In the case of Gaussian tails it can be simplified using (2.28) to

$$w(E_\mathrm{t}) = \frac{\sqrt{\eta}\exp\left(-E_\mathrm{t}^2/\eta^2\right)}{2^{5/2}\pi E_\mathrm{t}^{3/2}\left[(E_\mathrm{t}/\eta)^{3/2}+s\right]} \ . \tag{2.37}$$

The transition rate between a tail state with energy E and a CB state with energy E' becomes

$$\begin{aligned} D_{\mathrm{t,c}}(E,E') &= \int_{E_\mathrm{edge}}^{E_\mathrm{g}} dE_\mathrm{t} w(E_\mathrm{t}) D_{\mathrm{t,c}}(E,E',E_\mathrm{t}) \\ &= 8\pi\left(\frac{\hbar^2}{2m_\mathrm{t}\tilde{E}}\right)^{\frac{3}{2}} \varrho_{\mathrm{t,h}}(\tilde{E})\varrho_\mathrm{c}(E') \end{aligned} \tag{2.38}$$

with

$$\varrho_{\mathrm{t,h}}(\tilde{E}) = \left(\frac{2m_\mathrm{t}}{\hbar^2}\right)^{\frac{3}{2}} \frac{\sqrt{\eta}}{2\pi^2} Y\left(\tilde{E}/\eta\right) \ . \tag{2.39}$$

In a constant electric field F and assuming a parabolic dispersion for the CB, the CB DOS $\varrho_\mathrm{c}(E')$ has the form [51]

$$\begin{aligned} \varrho_\mathrm{c}(E') &= \frac{\sqrt{8m_c^3}}{4\pi\hbar^3}\sqrt{\hbar\theta_\mathrm{c}}\,\mathcal{F}\left(-\frac{E'}{\hbar\theta_\mathrm{c}}\right) \\ \text{with } \mathcal{F}(x) &= Ai'(x)^2 - xAi(x)^2 \\ \text{and } \hbar\theta_\mathrm{c} &= \left(\frac{e^2\hbar^2F^2}{2m_\mathrm{c}}\right)^{1/3} . \end{aligned} \tag{2.40}$$

Here, e is the electron charge and $Ai(x)$ denotes the Airy function. The total emission rate from all tail states with energy E into the CB is given by

$$G_\mathrm{tc} = \frac{(eF)^2 z_\mathrm{cv}^2}{\hbar} \int\limits_{E_\mathrm{edge}-E_\mathrm{g}}^{0} dE \int\limits_{-\infty}^{\infty} dE'\, D_{\mathrm{t,c}}(E,E')\delta(E-E') \ , \tag{2.41}$$

where z_{cv} is the interband transition matrix element [28] $z_{\mathrm{cv}}^2 = \hbar^2/(4m_{\mathrm{r}}E_{\mathrm{g}})$ with the reduced effective mass $m_{\mathrm{r}} = m_{\mathrm{c}}m_{\mathrm{v}}/(m_{\mathrm{c}}+m_{\mathrm{v}}) = m_{\mathrm{c}}m_{\mathrm{v}}/m_{\Sigma}$. For the completion of the band-to-band process it is assumed that the thermionic emission step from the VB to the tail state is very fast and, therefore, not rate-limiting. Then, after inserting (2.38) with (2.27) and (2.40) into Eq. (2.41), the generation rate via Gaussian tail states becomes

$$G_{\mathrm{tc}} = \frac{(eF)^2\sqrt{\eta}\sqrt{\hbar\theta_c}}{\sqrt{2}\pi^2\hbar^2 E_{\mathrm{g}}}\frac{m_{\mathrm{c}}^{3/2}}{m_{\mathrm{r}}}\times \int_{E_{edge}}^{E_g}\frac{d\tilde{E}}{(\tilde{E})^{3/2}}Y\left(-\frac{\tilde{E}}{\eta}\right)\mathcal{F}\left(\frac{E_{\mathrm{g}}-\tilde{E}}{\hbar\theta_c}\right)\ . \tag{2.42}$$

Since $E_{\mathrm{g}} \gg \hbar\theta_{\mathrm{c}}$, the function $\mathcal{F}$ can be replaced by its asymptotic limit for large positive arguments $\mathcal{F}(x) \to \exp(-4x^{3/2}/3)/(8\pi x)$. If Y is replaced by approximation (2.28), Eq. (2.42) simplifies to

$$G_{\mathrm{tc}} = \frac{(eF)^3\sqrt{\eta}}{64\sqrt{2}\pi^3\hbar E_{\mathrm{g}}^2}\frac{m_{\mathrm{c}}}{m_{\mathrm{r}}}\times \times \int_{E_{\mathrm{edge}}}^{E_{\mathrm{g}}} d\tilde{E}\frac{\exp\left[-\frac{\tilde{E}^2}{\eta^2}-\frac{4}{3}\left(\frac{E_{\mathrm{g}}-\tilde{E}}{\hbar\theta_{\mathrm{c}}}\right)^{\frac{3}{2}}\right]}{\tilde{E}^{3/2}\left[(\tilde{E}/\eta)^{3/2}+s\right]}. \tag{2.43}$$

In order to demonstrate what distinguishes the tunnel rate (2.43) from the rate of BTBT, the remaining integral is calculated analytically. The integrand is dominated by the overlap of the steep DOS tail of the VB and the so-called Franz-Keldysh tail of the CB. Since the product of both results in a sharply bell-shaped curve, one can determine its peak position $\tilde{E} = \Delta$ approximately and map the integrand to a Gaussian bell curve. This leads to

$$\int_{E_{\mathrm{edge}}}^{E_{\mathrm{g}}} d\tilde{E}\frac{e^{-\frac{\tilde{E}^2}{\eta^2}-\frac{4}{3}\left(\frac{E_{\mathrm{g}}-\tilde{E}}{\hbar\theta_{\mathrm{c}}}\right)^{\frac{3}{2}}}}{\tilde{E}^{3/2}\left[(\tilde{E}/\eta)^{3/2}+s\right]} \approx \approx \frac{e^{-\frac{\Delta^2}{\eta^2}-\frac{4}{3}\left(\frac{E_{\mathrm{g}}-\Delta}{\hbar\theta_{\mathrm{c}}}\right)^{\frac{3}{2}}}}{\Delta^{3/2}\left[(\Delta/\eta)^{3/2}+s\right]}\int_{-\infty}^{\infty} d\epsilon\ e^{-\frac{\epsilon^2}{\gamma^2}} \tag{2.44}$$

with

$$\Delta \approx \frac{\sqrt{E_\mathrm{g}}\eta^2}{(\hbar\theta_\mathrm{c})^{3/2}} \qquad \text{and} \qquad \gamma \approx \eta \; . \tag{2.45}$$

These expressions rely on the assumption $\eta \ll E_\mathrm{g}$. The generation rate finally takes the form

$$\begin{aligned} G_\mathrm{tc} = & \frac{(eF)^3\,\eta^{3/2}}{64\sqrt{2}\pi^{\frac{5}{2}}\hbar E_\mathrm{g}^2}\,\frac{m_\mathrm{c}}{m_\mathrm{r}} \times \\ & \times \frac{\exp\left[-\frac{\Delta^2}{\eta^2} - \frac{4}{3}\left(\frac{E_\mathrm{g}-\Delta}{\hbar\theta_\mathrm{c}}\right)^{\frac{3}{2}}\right]}{\Delta^{3/2}\left[(\Delta/\eta)^{3/2}+s\right]} \; . \end{aligned} \tag{2.46}$$

When the one-band effective-mass approximation and the WKB limit are applied to compute the generation rate of direct BTBT (with ideal DOS), one obtains [51]

$$G_\mathrm{BTB} = \frac{(eF)^3}{64\pi\hbar E_\mathrm{g}^2}\exp\left[-\frac{4}{3}\left(\frac{E_\mathrm{g}}{\hbar\theta_\mathrm{r}}\right)^{\frac{3}{2}}\right]. \tag{2.47}$$

Fig. 2.15 compares the full model Eq. (2.42) with the analytical approximation Eq. (2.46) and the BTBT rate (Eq. (2.47)) as function of electric field. Band gap and effective mass of bulk InAs have been used. In the low-field range the contribution of tail states to the total BTBT is negligible if the life time broadening of their energy levels is ignored. In this range the analytical expression agrees well with the full model as deep tail states are dominant. With increasing field, shallower states become more and more important and the saddle-point method based on deep Gaussian tails fails at the band edge, causing a growing deviation from the full model.

2.4.3 Tunneling from/into shallow tail states including their life time broadening

The life time broadening is approximately given by $\Delta\tau = \theta_\mathrm{t}^{-1}$ which is a measure of the tunnel probability out of the localized tail state. At $F = 1\times10^5\,\mathrm{V/cm}$ the characteristic energy $\hbar\theta_\mathrm{t}$ ranges from 20 meV to 12 meV for m_t between $0.5\,m_0$ and $2\,m_0$. This is of the same order

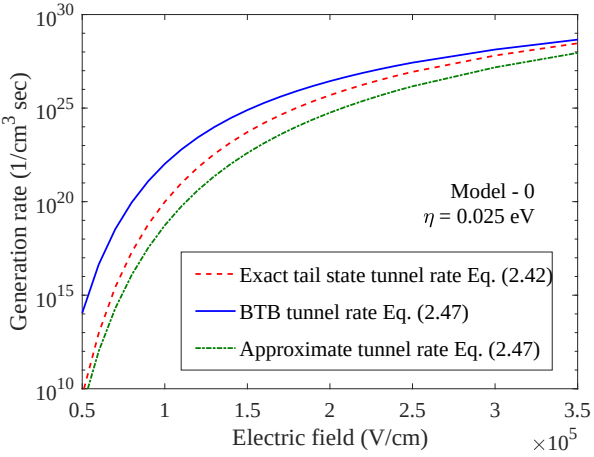


Figure 2.15: Generation rates due to strongly localized tail states without field broadening calculated with exact Gaussian tail DOS (Eq. (2.42)) in comparison to the analytical approximation (Eq. (2.46)) and the BTBT rate (Eq. (2.47)).

as the energies E_t in the tail. Thus, the zero-field DOS of the sharp single-level, Eq. (2.33), is to be replaced by a properly broadened delta-function. This can be done thanks to an analytical solution of the Schrödinger equation for an effective potential which is the sum of the pseudo-δ-potential and the electrostatic potential in a constant electric field [52]. The DOS Eq. (2.33) is then replaced by (see Appendix A of Ref. [48], and Ref. [47])

$$\varrho_\mathrm{t}(\tilde{E}, E_\mathrm{t}) = \frac{1}{4\pi^3 r_0^3 \sqrt{E_\mathrm{t}\hbar\theta_\mathrm{t}}} \frac{\mathcal{F}(\frac{\tilde{E}}{\hbar\theta_\mathrm{t}})}{\mathcal{F}^2(\frac{\tilde{E}}{\hbar\theta_\mathrm{t}}) + \left[\mathcal{G}(\frac{\tilde{E}}{\hbar\theta_\mathrm{t}}) + \frac{1}{\pi}\sqrt{\frac{E_\mathrm{t}}{\hbar\theta_\mathrm{t}}}\right]^2} \ , \quad (2.48)$$

which is a Lorentzian-like function of E_t with the property that it approaches the delta-function Eq. (2.33) for $\hbar\theta_\mathrm{t} \to 0$. The peak position is determined by the zero of $\mathcal{G}(\frac{\tilde{E}}{\hbar\theta_\mathrm{t}}) + \frac{1}{\pi}\sqrt{\frac{E_\mathrm{t}}{\hbar\theta_\mathrm{t}}}$. The function $\mathcal{G}(\frac{\tilde{E}}{\hbar\theta_\mathrm{t}})$ is given by

$$\begin{aligned} \mathcal{G}(x) =& Ai'(x)Bi'(x) - xAi(x)Bi(x) \quad (2.49) \\ &\to -\frac{\sqrt{x}}{\pi}\left(1 - \frac{0.03123}{x^3}\right) \quad \text{for large } x > 0 \end{aligned}$$

where $Bi(x)$ denotes the Airy function of the second kind. The second term in braces is proportional to F^2 and results in a shift of the resonance peak to larger energies (quadratic Stark effect). The peak hight is determined by the inverse of the function $\mathcal{F}(\frac{\tilde{E}}{\hbar\theta_\mathrm{t}})$ given in Eq. (2.40).

The transition probability from a field-broadened localized tail state to a field-dependent state in the opposite band is derived in Appendix A of Ref. [48]. A closed-form expression can be obtained for the cases of strong and very weak localization, respectively. No analytical solution is possible in the general case. The critical parameter is $a^3 = m_\mathrm{c}/(m_\mathrm{c} + m_\mathrm{t}) = \mu/m_\mathrm{t}$, where μ denotes the reduced effective mass $m_\mathrm{c}m_\mathrm{t}/(m_\mathrm{c} + m_\mathrm{t})$. Strong localization is defined by the condition $a^3 \ll 1$, whereas very weak localization by $a^3 \approx 1$.

In the case of strong localization ($a^3 \ll 1$) one obtains (see Appendix A of Ref. [48])

$$D_{\mathrm{t,c}}(E, E', E_\mathrm{t}) = 8\pi \frac{m_\mathrm{c}}{\mu} \frac{\mathcal{P}}{\tilde{E}^2} E_\mathrm{t}^2 r_0^3 \ \varrho_\mathrm{t}(\tilde{E}, E_\mathrm{t}) \ \varrho_\mu(E') \ . \quad (2.50)$$

Obviously, Eq. (2.32) is the special case in which the field-effect on the localized states is neglected. The essential difference to Eq. (2.32) is that the CB DOS $\varrho_{\rm c}(E')$ given by Eq. (2.40) is replaced by the *joint* DOS

$$\varrho_\mu(E') = \frac{\sqrt{8\mu^3}}{4\pi\hbar^3}\sqrt{\hbar\theta_\mu}\mathcal{F}\left(-\frac{E'}{\hbar\theta_\mu}\right) \tag{2.51}$$

which contains the reduced effective mass μ instead of the CB mass $m_{\rm c}$. This leads to an increased tunnel probability. Note, that this result is non-trivial. The occurrence of a reduced effective mass for a transition from a localized state into a Bloch state is due to the assumption that the localized state is built from a single, parabolic band with "effective" mass $m_{\rm t}$. Note also, that $\varrho_{\rm t}(\tilde{E}, E_{\rm t})$ is exactly the same as defined in Eq. (2.48).

With (2.48) the generation rate is given by

$$\begin{aligned} G_{\rm tc} = &\frac{(eF)^2\sqrt{\eta}}{2^{3/2}\pi^4\hbar^2 E_{\rm g}}\frac{m_{\rm c}\sqrt{\mu}}{m_{\rm r}}\sqrt{\frac{\hbar\theta_\mu}{\hbar\theta_{\rm t}}}\int_{E_{edge}}^{E_g} dE_t\, Y\left(-\frac{E_t}{\eta}\right) \\ &\int_{E_{edge}}^{E_g}\frac{d\tilde{E}}{\tilde{E}^2}\frac{\mathcal{F}(\frac{E_g-\tilde{E}}{\hbar\theta_\mu})\mathcal{F}(\frac{\tilde{E}}{\hbar\theta_t})}{\mathcal{F}^2(\frac{\tilde{E}}{\hbar\theta_t}) + \left[\mathcal{G}(\frac{\tilde{E}}{\hbar\theta_{\rm t}}) + \frac{1}{\pi}\sqrt{\frac{E_{\rm t}}{\hbar\theta_{\rm t}}}\right]^2} \end{aligned} \tag{2.52}$$

which after inserting the function Y for Gaussian tails becomes the triple integral

$$\begin{aligned} G_{\rm tc} = &\frac{(eF)^2}{2^{3/2}\pi^{9/2}\hbar^2 E_{\rm g}}\frac{m_{\rm c}\sqrt{\mu}}{m_{\rm r}}\sqrt{\frac{\hbar\theta_\mu}{\hbar\theta_{\rm t}}}\frac{1}{\eta}\times \\ &\times\int_0^\infty d\zeta\sqrt{\zeta}\int_{E_{\rm edge}}^{E_{\rm g}}\frac{d\tilde{E}}{\tilde{E}^2}\mathcal{F}(\frac{E_{\rm g}-\tilde{E}}{\hbar\theta_\mu})\mathcal{F}(\frac{\tilde{E}}{\hbar\theta_{\rm t}})\times \\ &\times\int_{E_{\rm edge}}^{E_{\rm g}} dE_{\rm t}\frac{\exp\left[-\frac{(\zeta+E_{\rm t})^2}{\eta^2}\right]}{\mathcal{F}^2(\frac{\tilde{E}}{\hbar\theta_{\rm t}}) + \left[\mathcal{G}(\frac{\tilde{E}}{\hbar\theta_{\rm t}}) + \frac{1}{\pi}\sqrt{\frac{E_{\rm t}}{\hbar\theta_{\rm t}}}\right]^2}\,. \end{aligned} \tag{2.53}$$

For an analytical treatment one can use the fact that even if $\hbar\theta_{\rm t} \approx \eta$, the Lorentzian describing the trap DOS has a sharp maximum at the resonance energy defined by $\mathcal{G}(\frac{\tilde{E}}{\hbar\theta_{\rm t}}) + \frac{1}{\pi}\sqrt{\frac{E_{\rm t}}{\hbar\theta_{\rm t}}} = 0$. At $F = 1\times10^5$ V/cm

a mass as small as $m_\mathrm{t} = 0.003\,m_0$ results in a broadening of only $\sim$1 meV. Hence, one can replace $\mathcal{F}/(\mathcal{F}^2+\hat{\mathcal{G}}^2) \to 2\pi^2\sqrt{E_\mathrm{t}\hbar\theta_\mathrm{t}}\,\delta(\tilde{E}-E_\mathrm{t})$. With the asymptotic limit for $\mathcal{F}(\frac{E_\mathrm{g}-\tilde{E}}{\hbar\theta_\mu})$ this leads to

$$G_\mathrm{tc} = \frac{(eF)^3}{16\pi^3\hbar E_\mathrm{g}^2}\frac{m_\mathrm{c}}{m_\mathrm{r}} \int\limits_{E_\mathrm{edge}/\eta}^{E_\mathrm{g}/\eta} \frac{d\epsilon}{\epsilon^{3/2}} \exp\left[-\frac{4}{3}\left(\frac{E_\mathrm{g}-\epsilon\,\eta}{\hbar\theta_\mu}\right)^{3/2}\right] \times$$
$$\times \frac{1}{\sqrt{\pi}} \int\limits_0^\infty dt\sqrt{t}\exp\left[-(t+\epsilon)^2\right] \ . \tag{2.54}$$

Note, that the last line is just $Y_\mathrm{g}(-\epsilon)$. In the low-field range only deeper tail states contribute and the saddle-point method can be applied as before. Using approximation (2.28) for the Gaussian tail DOS yields

$$G_\mathrm{tc} = \frac{(eF)^3\,\eta^{3/2}}{64\sqrt{2}\pi^{\frac{5}{2}}\hbar E_\mathrm{g}^2}\frac{m_\mathrm{c}}{m_\mathrm{r}}\left(\frac{\mu}{m_\mathrm{c}}\right)^{\frac{1}{3}} \times$$
$$\times \frac{\exp\left[-\frac{\Delta^2}{\eta^2} - \frac{4}{3}\left(\frac{E_\mathrm{g}-\Delta}{\hbar\theta_\mu}\right)^{\frac{3}{2}}\right]}{\Delta^{3/2}\left[(\Delta/\eta)^{3/2}+s\right]} \tag{2.55}$$

where θ_c has to be replaced by θ_μ in the expression (2.45) for Δ. Hence, basically Eq. (2.46) is recovered, with the reduced effective mass μ instead of the CB effective mass. At high fields (large $\hbar\theta_\mu$) or for small η the most contributing energies are close to the VB edge. The factor of the integrand in the first line of Eq. (2.54) is a relatively smooth function here and can be taken out at the characteristic tail energy η ($\epsilon = 1$) or at some fraction $\epsilon^* = E^*/\eta$ of it ($E_\mathrm{edge}/\eta < \epsilon^* < 1$). The integration limits of the ϵ-integral can be approximately changed such that the remaining double integral becomes

$$\int\limits_0^\infty d\epsilon \int\limits_0^\infty dt\sqrt{t}\exp\left[-(t+\epsilon)^2\right] = \gamma = 0.302\ . \tag{2.56}$$

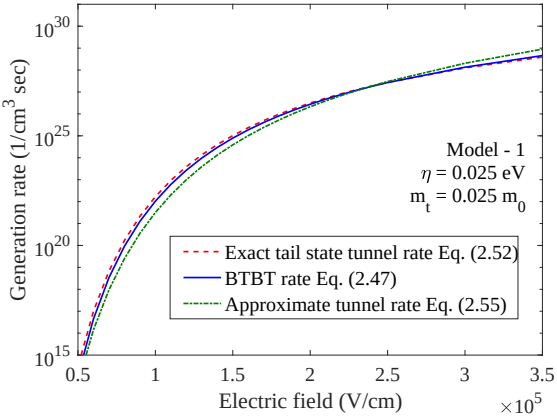


Figure 2.16: Generation rates due to strongly localized tail states with $m_\mathrm{t} = 0.025\, m_0$ including their field broadening calculated with the full model (Eq. (2.52)) in comparison to the analytical approximation (Eq. (2.55)) and the BTBT rate (Eq. (2.47)).

The error compared to Eq. (2.54) vanishes with the ratio E_edge/η. The result is

$$G_\mathrm{tc} = \frac{\gamma\,(eF)^3}{16\pi^{7/2}\hbar E_\mathrm{g}^2}\,\frac{m_\mathrm{c}}{m_\mathrm{r}}\left(\frac{\eta}{E^*}\right)^{\frac{3}{2}} \exp\left[-\frac{4}{3}\left(\frac{E_\mathrm{g}-E^*}{\hbar\theta_\mu}\right)^{\frac{3}{2}}\right]. \qquad (2.57)$$

In the case of exponential tails, the factor γ has the value 0.785.

The generation rate obtained for the field-broadened tail states (Eq. (2.52)) is compared with the analytical approximation (Eq. (2.55)) and the rate of BTBT for reference in Figs. 2.16 and 2.17. If the effective mass of the localized state m_t is set to the light-hole mass, the generation rate due to tail states almost coincides with the BTBT rate. For strong localization ($m_\mathrm{t} = 0.41\, m_0$) the effect of field broadening almost disappears and model-1 gives a similar curve as model-0.

In Ref. [48], an analytical model is derived also for the case of very weak localization of the wave functions of the tail states. However, simulations yield unphysical results when this model is applied in the

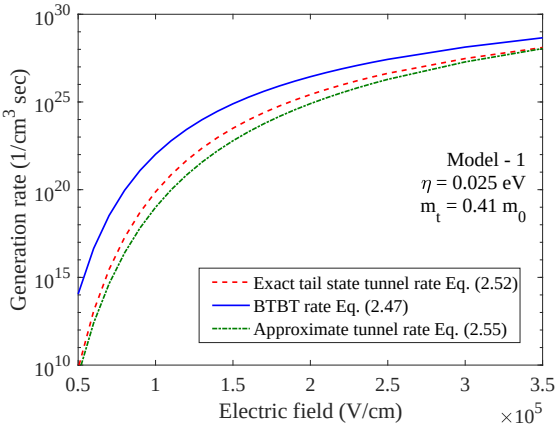


Figure 2.17: Generation rates due to strongly localized tail states with $m_t = 0.41\, m_0$ including their field broadening calculated with the full model (Eq. (2.52)) in comparison to the analytical approximation (Eq. (2.55)) and the BTBT rate (Eq. (2.47)).

simulation of realistic TFETs. Therefore, it is not described here. Interested readers may refer to [48].

2.4.4 Dynamic non-local path algorithm

The analytical forms of the generation rates for tunneling between VB tail states and CB states have been derived for a long semiconductor region with constant electric field. Using them as local models in TCAD would result in an over-estimation of the generation current in TFETs. This is illustrated in Fig. 2.18. A tunnel path starting from the CB edge at $x = x_1$ ends at the VB edge. Therefore, electrons from all tail states throughout the band gap ($E_{\mathrm{edge}} < E_{\mathrm{t}} < E_{\mathrm{g}}$) tunnel to the CB at x_1. However, a tunnel path starting at $x = x_2$ on the CB edge can only be used by tail states with an energy above mid-gap ($E_{\mathrm{t}} > E_{\mathrm{g}}/2$). As a result, the tail states above mid-gap ($\mathrm{E_{edge}} < \mathrm{E}_t < \mathrm{E}_g/2$) are active. A calculation of the electron generation at x_1 using the expressions in the previous sub-section would yield correct results. However, employing

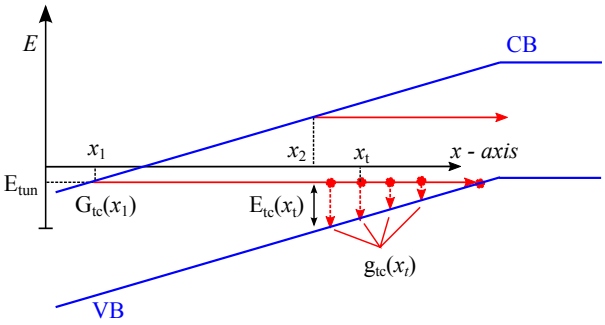


Figure 2.18: Schematic representation of two generic types of tunnel paths occurring in a TFET. The path beginning at x_2 does not come across tail states with $E_\mathrm{t} < E_\mathrm{mid}$. Various variables used in Eqs. (2.59)–(2.61) are shown.

the same expressions to calculate the generation rate at x_2 would overestimate the tunnel current. Therefore, the numerical implementation must be generalized to inhomogeneous electric fields. This is done in S-Device [33] by the DNLP algorithm which calculates the tunnel rate by numerical integration over the action along dynamically extracted tunnel paths. Thus, the expressions derived in the previous section need to be adapted to the DNLP algorithm.

To do so, all energetic tunnel rates are generally transformed into position-dependent rates defined along a tunnel path. The adaptation of Eq. (2.42) proceeds as follows. Each tunnel path connects the CB edge and the VB edge. At any point along the tunnel path located at a distance $x_\mathrm{t} \in \{x_1, x_1 + L\}$, $\tilde{E} \to E_\mathrm{tun} - E_\mathrm{v}(x_\mathrm{t})$ and $d\tilde{E} \to eF(x_\mathrm{t})\Delta x_\mathrm{t}$. In this way, Eq. (2.42) becomes locally defined at x_1:

$$G_\mathrm{tc}(x_1) = \sum_{x_\mathrm{t}=x_1}^{x_\mathrm{t}=x_1+L} g_\mathrm{tc}(x_\mathrm{t}) \tag{2.58}$$

$$g_\mathrm{tc}(x_\mathrm{t}) = \frac{\left(eF_\mathrm{av}(x_\mathrm{t})\right)^2 \sqrt{\eta}\sqrt{\hbar\theta_\mathrm{c}} m_\mathrm{c}^{\frac{3}{2}}}{\sqrt{2}\pi^{5/2}\hbar E_\mathrm{g}^2\, m_\mathrm{r}} \frac{eF(x_\mathrm{t})\Delta x_\mathrm{t}}{\left(E_\mathrm{tun} - E_\mathrm{v}(x_\mathrm{t})\right)^{\frac{3}{2}}} \times$$
$$\times\, Y\left(-\frac{E_\mathrm{tun} - E_\mathrm{v}(x_\mathrm{t})}{\eta}\right) \mathcal{F}\left(\frac{E_\mathrm{c}(x_\mathrm{t}) - E_\mathrm{tun}}{\hbar\theta_\mathrm{c}}\right) \times$$

$$\times \left[f_{\mathrm{n}}(x_1) - f_{\mathrm{p}}(x_{\mathrm{t}}) \right] \ . \tag{2.59}$$

Here, L is the length of the tunnel path, $x_{\mathrm{t}} \in \{x_1, x_1 + L\}$ the location of the tail state along the tunnel path, Δx_{t} the discretization interval, $F_{\mathrm{av}}(x_{\mathrm{t}}) = \frac{1}{|x_{\mathrm{t}} - x_1|} \int_{x_1}^{x_{\mathrm{t}}} F(x) dx$ the electric field averaged over the segment of the tunnel path between $x = x_1$ and $x = x_{\mathrm{t}}$, and E_{tun} the CB energy at the beginning of the tunnel path. $E_{\mathrm{c}}(x_{\mathrm{t}})$, $E_{\mathrm{v}}(x_{\mathrm{t}})$, and $F(x_{\mathrm{t}})$ are, respectively, the CB edge, the VB edge, and the electric field at the location of the tail state, g_{tc} is the generation rate at x_{t}, and G_{tc} is the total generation rate at x_1. The function $f_{\mathrm{n/p}}(x) = \left[\exp\left(E_{\mathrm{tun}} - E_{\mathrm{F,n/p}}(x) \right) / k_{\mathrm{B}} T + 1 \right]^{-1}$ represents the Fermi distribution at x. The value of $Y(-\frac{E_{\mathrm{tun}} - E_{\mathrm{v}}(x_{\mathrm{t}})}{\eta})$ is calculated at each x_{t} using Eq. (2.25) or Eq. (2.31) for Gaussian or exponential tails, respectively.

If field broadening of the tail states is considered, the generation rate due to tail states is given by Eq. (2.52) ($a^3 = \frac{m_{\mathrm{c}}}{m_{\mathrm{c}} + m_{\mathrm{t}}} \ll 1$) which is modified by making the substitutions $E_{\mathrm{t}} \to E_{\mathrm{tun}} - E_{\mathrm{v}}(x_{\mathrm{t}}) = E_{\mathrm{t}}(x_{\mathrm{t}})$ and $dE_{\mathrm{t}} \to eF(x_{\mathrm{t}}) \Delta x_{\mathrm{t}}$. If $\tilde{E}$ in the second integral is replaced by $\epsilon\, \hbar\theta_{\mathrm{t}}$, the generation rate $g_{\mathrm{tc}}(x_{\mathrm{t}})$ becomes

$$\begin{aligned} g_{\mathrm{tc}}(x_{\mathrm{t}}) = & \frac{\left(eF_{\mathrm{av}}(x_{\mathrm{t}}) \right)^2 \sqrt{\eta}}{2^{3/2} \pi^{9/2} \hbar^2 E_{\mathrm{g}}} \frac{m_{\mathrm{c}} \sqrt{\mu}}{m_{\mathrm{r}} \hbar \theta_{\mathrm{t}}} \sqrt{\frac{\theta_\mu}{\theta_{\mathrm{t}}}} eF(x_{\mathrm{t}}) \Delta x_{\mathrm{t}} \times \\ & \times Y\left(-\frac{E_{\mathrm{t}}(x_{\mathrm{t}})}{\eta} \right) \mathcal{H}\left(\frac{E_{\mathrm{g}}}{\hbar\theta_\mu}, \frac{E_{\mathrm{t}}(x_{\mathrm{t}})}{\hbar\theta_{\mathrm{t}}} \right) \times \\ & \times \left[f_{\mathrm{n}}(x_1) - f_{\mathrm{p}}(x_{\mathrm{t}}) \right] \end{aligned} \tag{2.60}$$

with

$$\mathcal{H}(p, q) = \int_{E_{\mathrm{edge}}/\hbar\theta_t}^{E_g/\hbar\theta_t} \frac{d\epsilon}{\epsilon^2} \frac{\mathcal{F}(p - a\,\epsilon) \mathcal{F}(\epsilon)}{\mathcal{F}^2(\epsilon) + \left[\mathcal{G}(\epsilon) + \frac{1}{\pi} \sqrt{q} \right]^2} \ . \tag{2.61}$$

The total generation rate at x_1 is calculated by the sum (2.58).

The functions Y and $\mathcal{H}$ can be implemented in the form of look-up tables which are evaluated at the beginning of the simulation run, once the values of $m_{\mathrm{c}}, m_{\mathrm{v}}$, and m_{t} are known. They are computed at any given input by interpolating between pre-evaluated values using cubic

splines. For the integrals to be accurate, the input $q = E_{\mathrm{t}}(x_{\mathrm{t}})/\hbar\theta_{\mathrm{t}}$ in Eq. (2.61) needs to be less than the upper limit of the integral, i.e. $E_{\mathrm{t}}(x_{\mathrm{t}}) < E_{\mathrm{g}}$. For look-up tables it is necessary that the integration limits are fixed. For the lower limit, $E_{\mathrm{edge}}/\hbar\theta_t = 0.03$ is used and $E_{\mathrm{g}}/\hbar\theta_t = 10$ for the upper. With the fixed lower limit all contributions to the integral are safely embraced, and the divergence at $E_{\mathrm{edge}} = 0$ is excluded. In this way, the double/triple integrals in Eqs. (2.59) and (2.60) can be transformed into a single integral along the tunnel path. The equations have been implemented in S-Device using the PMI nonlocal generation-recombination. The original DNLP BTBT model requires the effective tunnel barrier and the electron/hole effective masses as input parameters. In addition to these parameters, for the new DNLP tail-to-band tunneling model one has to provide the effective mass m_{t} of a hole (or electron for the complementary process) localized in the tail state as well as the characteristic energy η of the DOS tail.

2.4.5 TCAD Implementation of the model

A semiconductor may exhibit both CB and VB tail states with different characteristic energies η. This general case is approximated here by the sum of the rates for VB-tail-to-CB tunneling and CB-tail-to-VB tunneling. The expressions for the latter are straightforwardly obtained by obvious changes in the notation of parameters occurring in the former, e.g. $m_{\mathrm{c}} \to m_{\mathrm{v}}, \theta_{\mathrm{c}} \to \theta_{\mathrm{v}}$. The contribution from tail-to-tail tunneling is neglected due to the small probability when both initial and final states are localized.

As in the original DNLP model for BTBT, the proposed model involves searching for active tunnel paths. A tunnel path starting at the VB edge must have an end point *at the CB edge* at the same energy. If no such point is found, the path is discarded. Once all the active tunnel paths at a given bias voltage are found, the tail-to-band tunnel rates are calculated at each discretization point along the tunnel path using one of the models described above and summed up to obtain the rate at the starting point. Tunneling of an electron from the tail state generates a hole at the same location since it is implicitly assumed that thermionic emission into the VB continuum is very rapid and thus not rate-limiting. The densities of generated holes and electrons

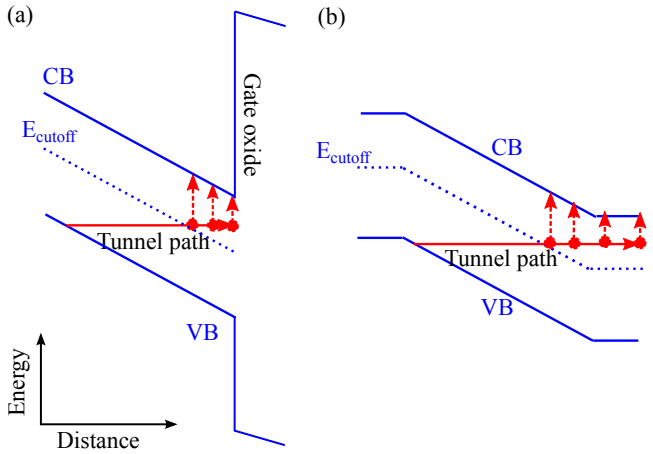


Figure 2.19: Treatment of special cases - (a) The tunnel path intersects the semiconductor-oxide interface instead of the CB edge. (b) The tunnel path reaches a maximum length without intersecting the CB edge. In both cases, the tunnel path is accepted if, at its end point, the energy difference between CB edge and tunnel energy is smaller than a cut-off energy.

enter the Poisson equation and self-consistently impact the solution of the drift-diffusion equation system.

In a TFET two cases occur that need special treatment. Fig. 2.19(a) presents the case in which a tunnel path encounters an insulator interface instead of the CB edge. In this situation, the tunnel path is accepted if the energy difference between the tunnel energy (ϵ) and the CB edge at the end point is smaller than a cut-off energy (E_{cutoff}). The integration of the generation rate is performed over all tail states along the path up to the intersection point. The second special case (Fig. 2.19(b)) occurs when the tunnel path exceeds a maximum length (set to 100 nm) but is still fully contained in the semiconductor. Then the cut-off energy is set to five times the characteristic tail energy (η) or to half of the gap, whichever is smaller. The choice of the band tail shape (Gaussian or exponential) rests upon the user.

2.4.6 Application to hetero-junction TFETs

In a hetero-TFET, the tunnel gap is reduced due to the staggered band alignment. The corresponding reduction in the tunnel path length may be included in the model as follows.

Within Kane's two-band $k \cdot p$ formalism, the BTBT rate with WKB approximation is calculated by integrating from the CB edge to the VB edge over the imaginary dispersion relation. It is given by, $\exp\left(-2\int_0^L \kappa(x)dx\right)$. In the one-band formalism described above, the BTBT rate with the WKB approximation is given by $8\pi p\mathcal{F}(p) \approx \exp\left[-\frac{4}{3}p^{\frac{3}{2}}\right]$ where $p = \frac{E_{\text{tun}}}{\hbar\theta_\mu}$. Comparing the BTBT rates resulting from the two formalisms gives us,

$$p = \left[\frac{3}{2}\int_0^L \kappa(x)dx\right]^{\frac{2}{3}}.$$

The above equivalence may be used to replace $\frac{E_g-\tilde{E}}{\hbar\theta_{\mu/c}}$ in the Eqs. (2.42) and (2.52) by,

$$p_{\mu/c} = \left[\frac{3}{2}\int_0^{x_t} \kappa_{\text{tb/ob}}(x)dx\right]^{\frac{2}{3}}, \tag{2.62}$$

where x_t is the location of the tail state along the tunnel path that begins from the CB edge ($x = 0$), $\kappa_{\text{tb}}(x)$ and $\kappa_{\text{ob}}(x)$ are the values of imaginary wave number (κ) calculated using Kane's expressions for two band and one band dispersion relations, respectively. Similarly, Eqs. (2.59) and (2.60) may also be modified to account for the reduction in the tunnel gap by replacing $\mathcal{F}(\frac{E_C(x_t)-E_{\text{tun}}}{\hbar\theta_{\mu/c}})$ in the integrals of the equations with $\mathcal{F}(p)$ where p is given by Eq. (2.62).

2.4.7 Simulation results

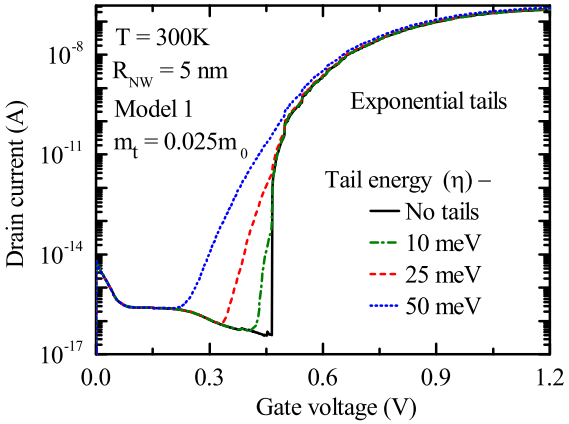
In order to analyze the impact of band tails on the transfer characteristics of TFETs, a representative sample is simulated with the TCAD tool. The structure is a radially symmetric InAs NW with a diameter of 10 nm and a gate with 80 nm length. Its radial cross section is shown in Fig. 1.1(a) of Chapter 1. The gate is slightly overlapped with the heavily doped source (p-type, $N_A = 1 \times 10^{19}\,\text{cm}^{-3}$) making it a

GOS NW TFET. The electron and light-hole effective mass values of InAs are set to $0.023\,\mathrm{m}_0$ and $0.026\,\mathrm{m}_0$, respectively [53]. As a result of geometrical confinement the InAs band gap increases from 0.36 eV to 0.76 eV [54]. Band Gap Narrowing (BGN) due to heavy doping is neglected in the simulations. The inclusion of BGN would reduce the tunnel gap for both BTBT as well as tail-to-band tunneling thereby scaling up both rates. To first order, this does not much change the relative strength of both generation mechanisms. BTBT between CB and heavy-hole band is also ignored.

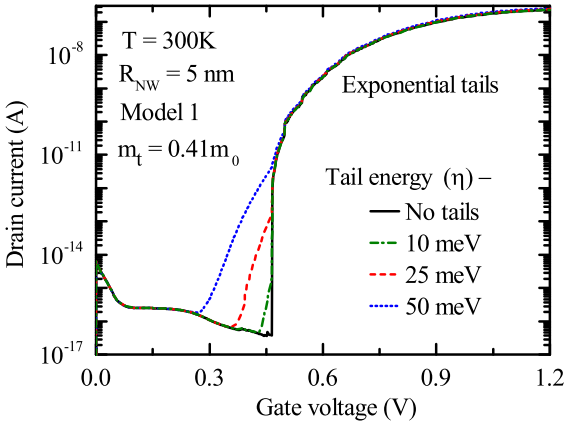
In the following, the effect of various model parameters such as the characteristic energy of the tail η, the tail shape, and the effective mass m_t on the transfer characteristics will be analyzed. The model developed in an earlier sub-section which includes the effect of field-induced broadening of tail states is applied to the device shown in Fig. 1.1(a).

The characteristic energy η of the tail determines how deep the latter penetrates into the gap, although the penetration depth is different for Gaussian and exponential tails. The value of η is varied over a feasible range to analyze its impact. The simulated transfer characteristics in the case of exponential tails are presented in Figs. 2.20(a) and 2.20(b) for two values of the parameter m_t. One can distinguish two distinct branches originating from tail-to-band tunneling and BTBT, respectively. This distinction is more pronounced for smaller η or larger m_t. As observed, a decreasing η steepens the transfer characteristics. For the range of the effective mass m_t an interval between the light-hole mass and the VB DOS mass is assumed. An increasing m_t reduces the drain current arising from tail-to-band tunneling as consequence of its reduced rate. The almost constant drain current prior to the onset of tunneling arises from SRH generation of electron-hole pairs in the depletion region. Carrier lifetimes of 10^{-9} sec are used in the simulation.

The average subthreshold swing is calculated for different values of η by averaging the inverse slope of the transfer curve over the range from 10^{-15} A to 10^{-11} A in the drain current. The result is shown in Fig. 2.21 as function of η for the two values of m_t. Increasing η degrades the average swing, but the degradation tends to saturate at larger η. Reducing m_t from $0.41\,m_0$ to $0.025\,m_0$ results in an only small increase of the swing. A possible reason for this relative insensitivity of



(a)



(b)

Figure 2.20: Impact of the characteristic energy η of an exponential DOS tail on the transfer characteristics of the TFET.

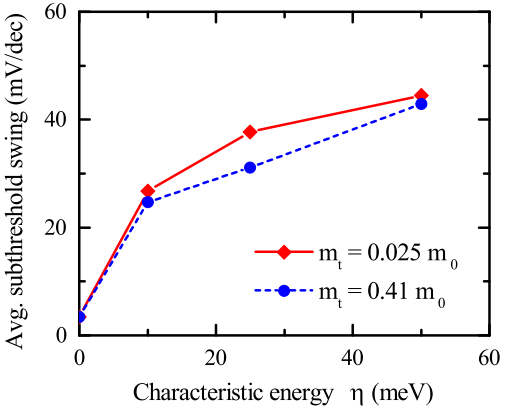


Figure 2.21: Degradation of the subthreshold swing swing with characteristic energy η for different values of the effective mass m_t.

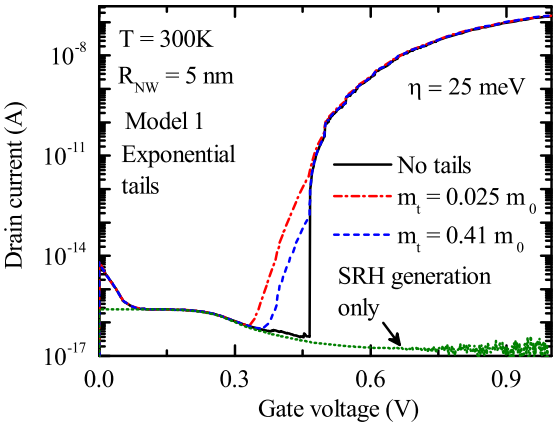


Figure 2.22: Comparison of the transfer characteristics obtained using model-1 with different values of the effective mass m_t.

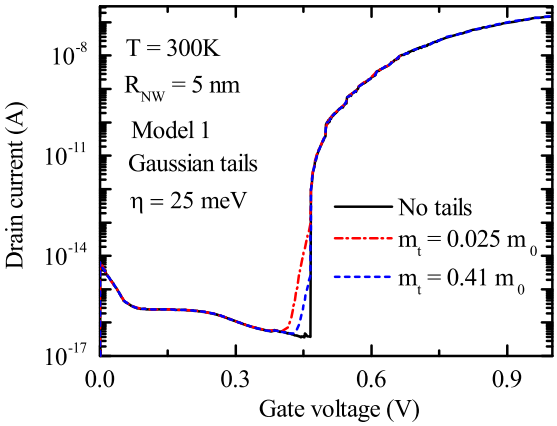


Figure 2.23: Impact of the effective mass m_{t} in the case of a Gaussian DOS. Model-1 which accounts for lifetime broadening of the tail states is used.

the swing to a variation of m_{t} could be the following. The steep onset of BTBT in the considered TFET results from the energetic alignment of the DOS in the source with the DOS in the channel at a certain gate bias. The onset is gradual due to the presence of tail states in the heavily doped source which smoothen the band edge. Therefore, the swing depends only on η which defines the degree of smearing. The parameter m_{t} merely determines the tunnel rate between the tail state and the CB state. The transfer characteristics for different m_{t} presented in Fig. 2.22 show that an increasing m_{t} scales down the tail-to-band current but does not significantly change the swing.

The transfer characteristics of the TFET for different values of m_{t} resulting from a Gaussian shape of the DOS tail are plotted in Fig. 2.23. Again, the drain current is reduced with increasing m_{t}. A comparison of the transfer characteristics of the device with exponential and Gaussian DOS tails, respectively, implies that the exponential shape of the DOS tails degrades the TFET performance more severely than the Gaussian shape. This is due to the fact that, for any given η, a band edge with Gaussian smoothing is sharper compared to a band

edge with exponential smoothing. Photoluminescence measurements infer the worst case, i.e. the presence of exponential DOS tails.

2.5 Effect of Traps and TAT

Trap-assisted tunneling (TAT) takes place via excitation of electrons from the valence band to the trap state and subsequent emission of the trapped electrons to the conduction band or vice-versa. The occupancy of a trap state is determined by the principle of detailed balance taking into account all trapping and de-trapping processes. The total trap occupancy at any interface determines the charge density at the respective interface which changes the electrostatics of the device self-consistently. Traps at the gate-oxide/semiconductor interface introduce an additional capacitance in parallel with the depletion layer capacitance which degrades the electrostatic coupling between the gate and the channel [55]. The reduced gate coupling degrades the subthreshold swing. This effect is similar to what is observed in MOSFETs in the presence of oxide interface traps. The degradation of the electrostatic control of the channel due to traps is observed irrespective of whether the traps are taking part in TAT or not.

2.5.1 TCAD modeling of TAT

An approximate expression to calculate the tunnel rate between the CB states and the trap state with an energy E_{t}^{0} and the trap density of N_{t} is derived below.

The density of localized single-level trap state is given by,

$$\varrho_{\mathrm{t}}(\tilde{E}, E_{\mathrm{t}}) = N_{\mathrm{t}}\delta(\tilde{E} - E_{\mathrm{t}}), \tag{2.63}$$

where the different energy variables have the same meaning as in Fig. 2.14. The energetic density of traps at each location is given by

$$w(E_{\mathrm{t}}) = \delta(E_{\mathrm{t}} - E_{\mathrm{t}}^{0}). \tag{2.64}$$

The trap DOS, Eq. (2.35), can now be calculated using Eqs. (2.63) and (2.64). One obtains,

$$\varrho_{\mathrm{t,h}}(\tilde{E}) = N_{\mathrm{t}}\delta(\tilde{E} - E_{\mathrm{t}}^{0}). \tag{2.65}$$

Eq. (2.32) in the previous section is used to calculate the transition rate between the localized trap level and the CB:

$$\begin{aligned} D_{\mathrm{t,c}}(E,E') &= \int_{-\infty}^{\infty} dE_{\mathrm{t}} w(E_{\mathrm{t}}) D_{\mathrm{t,c}}(E,E',E_{\mathrm{t}}) \\ &= 8\pi r_0^3 \varrho_{\mathrm{t,h}}(\tilde{E}) \varrho_{\mathrm{c}}(E'), \end{aligned} \tag{2.66}$$

where the trap DOS $\varrho_{\mathrm{t,h}}$ is given by Eq. (2.65) and the CB DOS $\varrho_{\mathrm{c}}(E')$ is obtained from Eq. (2.40) in the previous section.

The emission rate from the trap state to the CB becomes,

$$\begin{aligned} G_{\mathrm{tc}} =& \frac{(eF)^2 z_{\mathrm{cv}}^2}{\hbar} \int\limits_{E_{\mathrm{edge}}-E_{\mathrm{g}}}^{0} dE \int\limits_{-\infty}^{\infty} dE'\, D_{\mathrm{t,c}}(E,E') \delta(E-E') \\ =& \frac{(eF)^2 z_{\mathrm{cv}}^2}{\hbar} 8\pi r_0^3 N_{\mathrm{t}} \frac{\sqrt{8m_{\mathrm{c}}^3}}{4\pi\hbar} \sqrt{\hbar\theta_c} \int_{-E_{\mathrm{g}}}^{0} dE\, \delta(E+E_{\mathrm{g}}-E_{\mathrm{t}}^0) \times \\ & \times \int_{-\infty}^{\infty} dE'\, \delta(E-E') \mathcal{F}\left(-\frac{E'}{\hbar\theta_c}\right) \\ =& \frac{\sqrt{2}(eF)^2 m_{\mathrm{c}}^{3/2}}{4\hbar^2 m_{\mathrm{r}} E_{\mathrm{g}}} r_0^3 N_{\mathrm{t}} \sqrt{\hbar\theta_c} \mathcal{F}\left(\frac{E_{\mathrm{g}}-E_{\mathrm{t}}^0}{\hbar\theta_c}\right). \end{aligned} \tag{2.67}$$

where N_{t}, E_{t}^0, and r_0 represent, respectively, the trap density, the trap energy, and the spatial spread of each individual trap wave function. The remaining symbols carry the same meaning as in Eq. (2.41). The occupation factors of the trap state as well as the CB states are omitted in the above calculations. The total emission rate in Eq. (2.67) must be multiplied with the Fermi distribution $f\left(\frac{E_{\mathrm{F,n}}-E_{\mathrm{t}}^0}{kT}\right)$ to obtain the electron generation rate in the CB.

In S-Device, the direct tunnel process between the trap level and the CB edge is modeled by a non-local model which calculates the tunnel rate in WKB approximation with numerical integration over the imaginary dispersion [56] as follows:

$$\begin{aligned} G_{\mathrm{tc}} =& N_{\mathrm{t}} \frac{\sqrt{8m_{\mathrm{c}}} m_0}{\hbar^4\pi} V_{\mathrm{T}} [E_{\mathrm{C}}(z_0) - E_{\mathrm{t}}^0]^2 \Theta[E_{\mathrm{t}}^0 - E_{\mathrm{C}}(0)] \times \\ & \times \sqrt{E_{\mathrm{t}}^0 - E_{\mathrm{C}}(0)} \exp\left(-\int_0^{z_0} \kappa(z) dz\right) f\left(\frac{E_{\mathrm{f,n}}-E_{\mathrm{t}}^0}{kT}\right). \end{aligned} \tag{2.68}$$

This expression is similar to the one given by Eq. (2.67). Here, the localization volume of the trap wave function is denoted by V_{T} instead of r_0^3. The physical interpretation of the trap volume is the effective interaction volume. The trap volume is roughly equal to the volume enclosed by the equi-probability contour of the InAs/Si interface trap wave function shown in Fig. 5.4(b) of Chapter 4. The point $z = 0$ denotes the beginning of the tunnel path whereas $z = z_0$ is the location of the trap. $E_{\mathrm{C}}(z)$ denotes the CB edge along the non-local mesh and E_{t} is the trap energy level.

In the simulations, the energetic distribution of the traps at each vertex at the interface is discretized to create one hundred uniformly distributed trap levels which carry appropriate trap densities. Each trap level is connected by a tunnel path to the nearest point on the nonlocal line specified by the user. This is shown schematically in Fig. 2.24(a). All the vertices at the two interfaces are connected to the vertices on the nonlocal lines NL1 and NL2 as denoted by the dashed arrows in Fig. 2.24(a). Note, that the vertices at the InAs/Si interface are connected to the two nonlocal lines simultaneously. Each trap level at the oxide-interface vertex on the cut-line is now connected to the vertex on the nonlocal line by a tunnel path (see Fig. 2.24(b)). Integration over the imaginary dispersion relation is performed along the tunnel path to calculate the capture rate using Eq. (2.68). Since the traps at the InAs/Si interface are connected to the two nonlocal lines, capture rates from both the CB and the VB are calculated. Since the above TAT model is a built-in model in S-Device, it is used for the calculation of the TAT rate. Interested readers are requested to refer to the S-Device manual [33] for more information.

The generation rates given by Eqs. (2.67) and (2.68) have two fitting parameters, namely trap volume V_{T} (or r_0^3) and trap density N_{t} apart from the band structure quantities such as effective mass and band gap. The TAT rate is proportional to the product $V_{\mathrm{T}} \cdot N_{\mathrm{t}}$ only. On the other hand, the traps are charged when occupied, and the trap charge (proportional to N_{t}) affects the solution of the Poisson equation. The trap volume has no effect on the electrostatics. This allows one to adjust the two parameters wherever the measured trap density is not available.

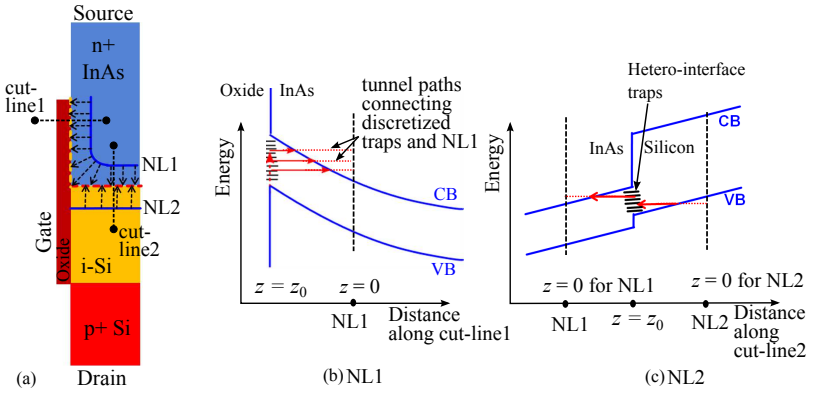


Figure 2.24: (a) Schematic of the InAs/Si hetero-junction TFET showing the locations of non-local lines as well as traps. (b) and (c) Band edge diagrams along cut-line1 and cut-line2, respectively, depicting how the nonlocal tunnel paths join the nonlocal line and the traps.

2.5.2 TCAD modeling of multi-phonon excitation

The multi-phonon excitation of electrons is modeled by the phenomenological V-model which assumes temperature-independent capture cross sections and the thermal velocity to calculate the carrier capture-emission rate. Capture (c_p) and emission (e_p) rates of a hole to/from the trap state are given by,

$$c_\mathrm{p} = \sigma_\mathrm{p} \cdot v_\mathrm{th}^\mathrm{p} \cdot p \tag{2.69}$$

$$e_\mathrm{p} = \sigma_\mathrm{p} \cdot v_\mathrm{th}^\mathrm{p} \cdot p_1. \tag{2.70}$$

Here, σ_p is the hole capture cross-section, $v_\mathrm{th}^\mathrm{p} = \sqrt{\frac{kT}{2m_\mathrm{h}}}$ is the thermal velocity of holes, p is the hole density, while p_1 is given by $p_1 = N_\mathrm{V} \exp\left[(E_\mathrm{V} - E_\mathrm{t}^0)/kT\right]$. The deeper the trap level the smaller is the occupation probability, and therefore, the smaller is the emission rate of a hole from the trap level. This is captured by the V-section model. Since multiple trap energy levels are located at each vertex at the InAs/oxide, the capture and the emission rates are calculated separately for each trap level.

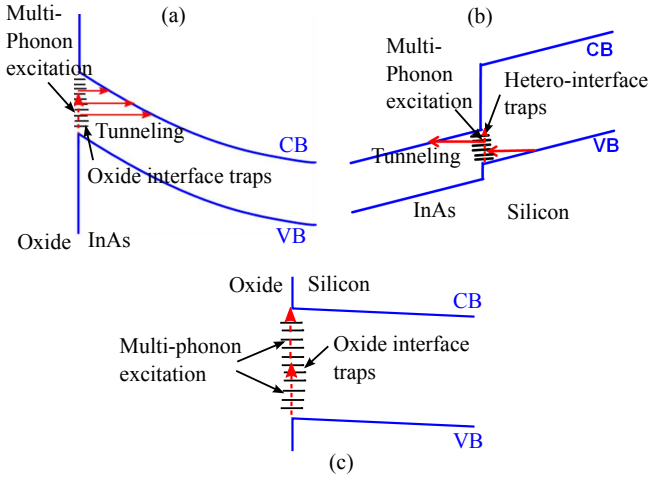


Figure 2.25: Schematic showing TAT at (a) oxide/semiconductor interface under surface inversion, (b) semiconductor hetero-interface, and (c) oxide/semiconductor interface under accumulation or volume inversion.

TAT at the oxide/semiconductor interface and at the hetero-junction occur via strikingly different mechanisms as outlined below.

2.5.3 TAT at the oxide/semiconductor interface

Device in surface inversion mode

Due to surface inversion, a triangular potential well is formed in the semiconductor adjacent to this interface. In the following description, a hole inversion layer is assumed. Due to the triangular well, TAT takes place in two steps. In the first step, an electron is captured by a trap state from the VB of the semiconductor by a multi-phonon excitation process. The trapped electron is emitted into the CB via either multi-phonon excitation, or phonon-assisted or direct tunneling (Fig. 2.25(a)). The multi-phonon excitation of electrons is modeled by the V-model described above. Phonon-assisted and direct tunnel processes are modeled by the non-local model described above. The model is activated on the non-local mesh that is constructed perpendicular to the semiconductor/oxide interface.

Device in accumulation mode

In the accumulation mode of an n-channel/p-channel TFET, electrons/holes accumulate in the entire channel region without significant band bending. In the absence of band bending, TAT does not occur. Instead, an electron is captured via multi-phonon excitation from the VB and is subsequently emitted to the CB via another multi-phonon process (Fig. 2.25(c)). This kind of generation is modeled using the phenomenological V-model described above. This is also referred to as surface SRH generation. Due to the absence of tunneling to/from trap states, this mechanism is weaker compared to the other mechanisms which involve TAT.

2.5.4 TAT at the semiconductor hetero-interface

In this process, an electron from the VB of the first semiconductor (Silicon, in this case) can be trapped via both a multi-phonon excitation and a (direct or phonon-assisted) tunneling process (Fig. 2.25(b)). Similarly, a captured electron can be emitted to the CB of the other

semiconductor (in this case - InAs) via either of these two processes. Similar to the case of the semiconductor/oxide interface, multi-phonon excitation of electrons is modeled by the phenomenological V-model, whereas phonon-assisted and direct tunnel processes are modeled by the non-local model. The non-local model is activated on the non-local mesh that runs on either side of the semiconductor hetero-interface and that is constructed normal to the interface.

The V-model for multi-phonon excitation has two parameters – capture cross section and thermal velocity, while the non-local model for direct and phonon-assisted tunneling contains three parameters – trap interaction volume, Huang-Rhys factor, and phonon energy. The parameter *trap interaction volume* is proportional to the spatial spread of the trap wave function and acts as a scaling factor of the nonlocal TAT generation rate. The parameter *Huang Rhys factor* is considered as a measure of the coupling strength of the diagonal electron-phonon coupling of the trap levels. This factor is important only in the phonon-assisted tunneling process. This process is dominant only for electronic transitions between trap state and the indirect CB edge. In all the TFETs studied in this work, TAT may take place between the trap and the direct CB edge. Thus, the TFET characteristics have been found to be much less sensitive to Huang-Rhys factor as well as phonon energy. Therefore, these factors are not used in the fitting process. Phonon energy is set to transverse-optical (TO) phonon energy of Silicon in the simulations. The values of these parameters for both types of interfaces are provided in Table 2.1. Both models calculate capture and emission rates for each discrete trap level in the band gap. The rates are proportional to the trap density. The non-local model requires effective electron and hole masses to perform integration over the imaginary dispersion relations. Effective electron and hole masses in respective bulk semiconductors have been used for the non-local TAT model in this work.

Table 2.1: Values of trap-assisted tunneling parameters required for local and non-local TAT models.

Parameter	Unit	Semiconductor hetero-interface	semiconductor /oxide
Non-local TAT model			
Huang-Rhys factor (S)	-	3	3
Phonon energy ($\hbar\omega$)	eV	0.06	0.06
Trap Volume	$Å^3$	50	10
Local V-section model			
Cross-section	$Å^2$	10	10
Thermal velocity	cm/sec	2.042×10^7	7.57×10^7

Chapter 3

Non-idealities in InAs/Si TFETs

To deliver a high on-current, the TFET must have a small band gap semiconductor at the source/channel junction. On the contrary, suppression of ambipolar current necessitates the use of a large band gap semiconductor at the drain/channel junction. Both requirements are fulfilled by a hetero-junction TFET in which a large-gap material is used for the channel and the drain, while a small-gap semiconductor is used as source. An InAs/Si heterojunction p-channel TFET utilizes this technique. Due to the large indirect band gap of 1.11 eV of Si, the tunnel rate between the channel and the drain is small which results in a small ambipolar leakage current. The InAs with a direct gap of 0.36 eV ensures a tunnel rate between the source and the channel. Due to the staggered band alignment at the InAs/Si heterojunction, the effective tunnel gap for electrons tunneling from the VB of Si to the CB of InAs is further reduced enhancing the on-current. Thus, an InAs/Si heterojunction is well suited for TFETs. From a technology point of view, the InAs/Si TFET technology can be built on a well advanced Si Complementary Metal-Oxide Semiconductor (CMOS) technology. This would avoid creating a new manufacturing ecosystem from scratch and can significantly bring down the cost of introducing TFETs. Also, it can enable the co-existence of both TFET and MOSFET blocks

on the same integrated circuit chip. For the above reasons InAs/Si TFETs are sought after by the industry.

In this chapter, a detailed TCAD analysis of vertical as well as lateral NW InAs/Si heterojunction TFETs is presented. The TFETs were fabricated at IBM Research-Zürich and the data was kindly shared with us. The analysis is an attempt to understand why fabricated InAs/Si p-channel TFETs are not able to deliver sub-60mV/dec operation. Thus, it is primarily diagnostic in nature rather than exploratory.

3.1 InAs/Si Lateral Nanowire TFETs

In this section, the measured transfer characteristics of lateral InAs/Si nanowire heterojunction TFETs presented in [57] are analyzed. These devices exhibit a best swing value of 67 mV/decade, fairly close to the thermionic limit. By simulation, the role played by TAT, BTBT, and thermionic transport in the TFET operation is untangled. Based on the analysis, an estimate for the D_{it} which would allow a sub-thermal swing is provided.

3.1.1 Simulation set-up

Device geometry

TFETs were made from InAs/Si NWs with a square-shaped cross section of 30 nm × 30 nm laterally placed on buried oxide. Two of the fabricated devices are selected for the simulation study, namely, FF44 and FF41 (for more details, see [57]). An intrinsic underlap region between i-Si channel and p+-Si drain is present in FF44, while no such underlap exists in FF41. The InAs region is n^+-doped with a concentration of $2 \times 10^{18}\,\mathrm{cm}^{-3}$. The section of Si in the channel and the underlap region are intrinsic. The overlap between the InAs source with the gate is 270 nm long while that between Si and the gate is 730 nm long. The underlap region has an extension of 100 nm. The gate oxide consists of HfO_2/Al_2O_3 with an EOT of 1.75 nm. A vertical cross section of the device at the mid-point along the width of the NW (Fig. 3.1(a)) is simulated using S-Device [33]. Since the channel orientation is along the ⟨100⟩-direction while the InAs/Si interface is a

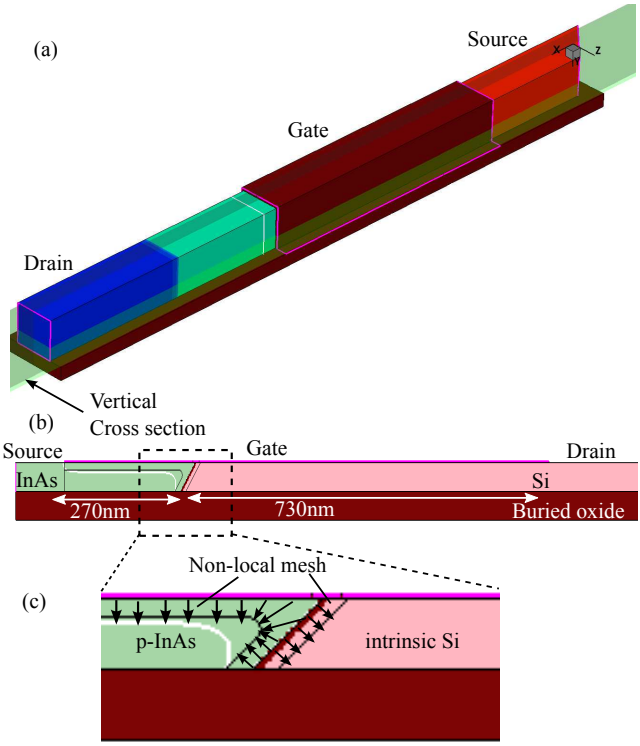


Figure 3.1: (a) 3D simulation domain of the InAs/Si lateral NW TFET along with the plane of vertical cross section. (b) 2D cross section of the device with various lengths. (c) Non-local mesh generated to model TAT at the InAs/Oxide and InAs/Si interfaces.

(111)-plane, the InAs/Si interface in the cross section is slanted making an angle of 45^0 with the buried oxide. It must be noted that the device does not possess any axis of symmetry. A vertical cross section is selected for 2D simulations to lower the computational burden, but some 3D simulations are also done for validation purposes.

Miscellaneous physical models

The dynamic non-local path BTBT model in S-Device [33] is used for the simulation of tunneling. This model requires the CB effective mass in the Γ-valley and the light hole (LH) effective mass, the direct band gap, and the degeneracy factor to simulate BTBT between the Γ-valley and the VB. Table 3.1 lists the values of various band structure quantities used in the simulation [58]. The temperature dependence of the band gap is modeled following Varshni [59] with default parameters in S-Device. The temperature dependence of the effective mass is ignored. Tunneling from the Si LH band to the InAs CB is the primary BTBT process in this device. The degeneracy factor ($g = 2 \times g_C \times g_V$) for this process is 2. Therefore, the degeneracy factor for BTBT in Si is set to 2, although the factor is 12 for intra-material tunneling in Si. Owing to the large effective mass of heavy holes (HH), tunneling between the Γ-valley of the CB and the HH band is ignored. The VB offset between InAs and Si is chosen to be 130 meV referring to the experimental data in Ref. [60]. Due to the small band gap of InAs, SRH generation is strong. It is modeled with constant lifetimes of 1×10^{-9} s. These values provided the best fit to the temperature dependence of reverse IV-curves of unintentionally doped InAs/Si hetero NW diodes [61]. The gate work function is set to 5.05 eV and the source-drain bias is 500 mV in the simulations presented below.

Modeling trap-assisted tunneling

As explained in Section 2.5, TAT takes place via excitation of electrons from the VB to the trap state and subsequent emission of the trapped electrons to the CB or vice versa. The occupancy of a trap state is determined by the principle of detailed balance taking into account all trapping and de-trapping process. The total trap occupancy at any interface determines the charge density at the respective interface

Table 3.1: Values of various band structure parameters required for the BTBT model.

Parameter	Unit	InAs	Silicon
		Direct gap	Indirect gap
E_g	eV	0.36	1.11
ΔE_C	eV	0.88	
ΔE_V	eV	0.13	
$m_C\langle 100\rangle$	m_0	0.023	0.19
$m_{LH}\langle 100\rangle$	m_0	0.026	0.147
Degeneracy (g)	1	2	2

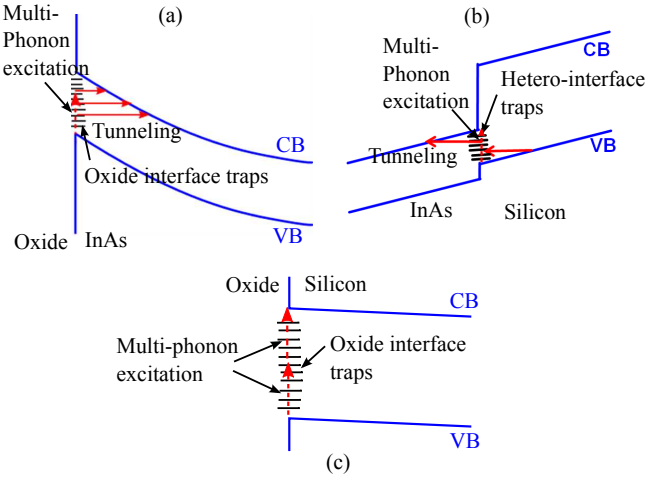


Figure 3.2: Schematic showing TAT at (a) oxide/InAs interface, (b) InAs/Si interface, and (c) Oxide/Si interface.

which changes the electrostatics of the device self-consistently. All the three generation processes described in Section 2.5 are present in the given TFET. The parameter values of the TAT and surface generation model are provided below.

Due to unavailability of measured data of the energetic trap distribution at the InAs/oxide interface, it is assumed to be uniform throughout the InAs band gap. The trap density at the InAs/Oxide interface is set to its measured value of $1 \times 10^{13}\,\mathrm{cm}^{-2}\mathrm{eV}^{-1}$ taken from Ref. [62]. At the InAs/Si interface, the energetic distribution of traps is assumed to have Gaussian shape with a peak concentration of $6 \times 10^{13}\,\mathrm{cm}^{-2}\mathrm{eV}^{-1}$, a Full Width Half Maximum (FWHM) of 0.22 eV, and the peak position at the VB edge of InAs. For the simulation of FF41 TFETs, the trap density at the Si/oxide interface is assumed to be uniform throughout the band gap of Si. Its value is set to $4 \times 10^{12}\mathrm{cm}^{-2}\mathrm{eV}^{-1}$ after fitting the experimental data. The parameters used for the V-section model and for the nonlocal TAT model are provided in Table 2.1 of Chapter 2. Note that, traps at the Si/oxide interface were not included in the simulation of FF44 devices as their impact is negligible. However, they had to be included in the simulation of the FF41 devices, possibly due to a high concentration of traps at Si/oxide interface due to process variations.

3.1.2 Results and discussion

Temperature dependence

The FF44 TFET (see Fig. 3.1(b)) is simulated by including traps at the InAs/Si and the InAs/oxide interfaces and using the model parameters outlined in the previous subsection. A comparison of simulated and measured transfer characteristics is presented in Fig. 3.3. The used model parameters are mostly experimental data, only Huang-Rhys factor, trap interaction volume, and trap density at the InAs/Si interface, which are related to the TAT model, have been fitted to obtain a good match to the experimental data. It is observed that the variation of the Huang-Rhys factor has only little effect on the transfer characteristics. The trap interaction volume merely acts as a scaling factor in the expression for the capture rate. A small change in the trap interaction volume will not cause a significant shift in the transfer

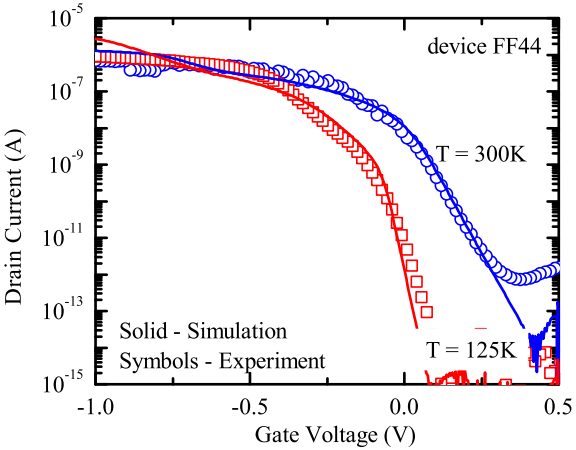


Figure 3.3: Comparison of experimental and simulated transfer characteristics of the FF44 TFETs for two different temperatures ($V_{\mathrm{DS}} = 0.5\,\mathrm{V}$).

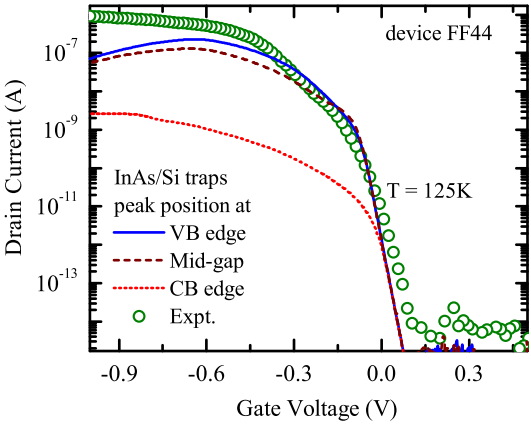


Figure 3.4: Variation of the peak position of the Gaussian D_{it} in the band gap of InAs at the InAs/Si interface. Only TAT at the InAs/Si interface is enabled in the simulation ($V_{\mathrm{DS}} = 0.5\,\mathrm{V}$).

characteristics on log scale. Therefore, the transfer characteristics only weakly depend on the uncertainty associated with these two parameters.

The energetic trap distribution at the InAs/Si interface affects the electrostatics of the device as well as the electron capture and emission rates. At 125 K, the contribution from oxide traps to the transfer characteristics is negligible. Therefore, the low-temperature characteristics can be used to study the distribution of traps at the InAs/Si interface. A Gaussian shape is assumed for the D_{it}, and the position of the peak is varied throughout the InAs band gap keeping the FWHM fixed to 0.22 eV. Fig. 3.4 shows transfer characteristics of the TFET at 125 K for three different peak positions in the band gap of InAs along with the experimental data. Best agreement is obtained for a peak position close to the VB edge of InAs.

The FF41 TFET is also simulated using the same simulation set-up. Fig. 3.5 shows the comparison of simulated transfer characteristics with experimental data at different temperatures. Unlike in the analysis of

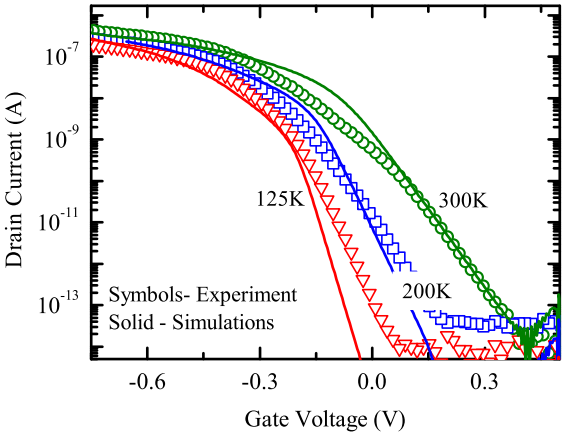


Figure 3.5: Measured and simulated transfer characteristics of the FF41 TFETs for three different temperatures ($V_{\mathrm{DS}} = 0.5\,\mathrm{V}$).

the FF44 TFET, traps at the Si/oxide interface have to be included for a good fit. Because of the large band gap of Si, multi-phonon excitation of electrons via traps at the Si/oxide interface is a less important process. Instead, the charging of these interface traps impacts the transfer characteristics by degrading the coupling between the gate electrode and the Si channel as explained below.

Individual contributions of various TAT processes

At 300 K the drain current originates from three sources: BTBT, TAT at the InAs/oxide interface, and TAT at the InAs/Si interface. Contributions of the individual mechanisms are plotted in Fig. 3.6 for the case of device FF44. Note that the remaining components had to be disabled which neglects the effect of the self-consistent coupling of all sources. TAT at both interfaces begins at a lower gate bias compared to BTBT as the effective tunnel gap for TAT is much smaller than that for BTBT. The drain current component due to TAT at the InAs/oxide interface increases much more gradually and has a

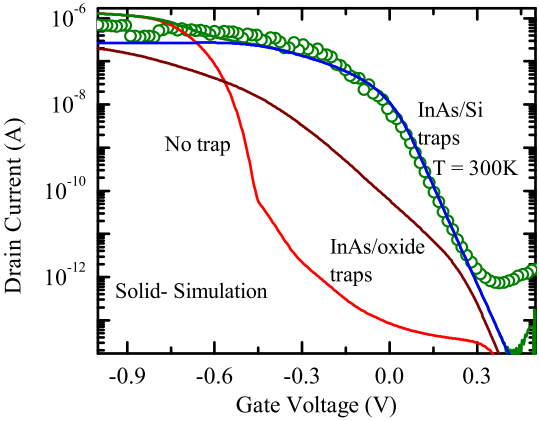


Figure 3.6: Individual contributions of BTBT and the two TAT processes at 300 K in the simulated transfer characteristics of FF44 TFETs. TAT at the InAs/Si interface is dominant in the subthreshold region ($V_{\mathrm{DS}} = 0.5\,\mathrm{V}$).

larger subthreshold swing compared to the one originating from traps at the InAs/Si interface. The reason is the weak bias dependence of the multi-phonon excitation step at the InAs/oxide interface. On the other hand, the drain current component from TAT at the InAs/Si interface yields a subthreshold swing close to thermal limit (60 mV/dec) at 300 K. This astonishing coincidence will be explained later.

The onset voltage of the TFET is fully dominated by TAT. The contribution of BTBT is negligible at the onset. However, the on-current is determined by BTBT, as it becomes the dominant mechanism at high gate bias. The over-estimation of the simulated on-current at very high gate bias (see Fig. 3.3) is due to the neglect of quantization effects [17]. Inclusion of quantization effects will merely delay tunneling perpendicular to the gate (line tunneling) as explained in Chapter 2. It will not affect point tunnel paths across InAs/Si hetero-interface. It will reduce on-current only, leaving onset of BTBT unaffected which will improve agreement with the experimental data.

Impact of inclined InAs/Si interface on BTBT

The simulated transfer characteristics in the absence of traps are unusual as they exhibit a very weak slope at the onset. As can be seen in Fig. 3.6, the drain current gradually increases at a lower bias ($-0.3\,\mathrm{V} < V_{\mathrm{GS}} < 0.3\,\mathrm{V}$) which can be attributed to the slanted InAs/Si interface. As demonstrated in Fig. 3.7(a), the tunnel path connecting the electron and hole generation centers is long for a TFET with slanted InAs/Si interface. Also the gate control is weaker as the electron generation center is far away from the gate and shielded by the carriers in the inversion layer. Both factors result in a gradual increase of the drain current. Note, that the same behavior is also observed in the transfer characteristics obtained by 3D simulations with the exact device geometry (see Fig. 3.7) which suggests that it is not an artifact of simulating a 2D cross section only. A vertical InAs/Si interface would shift the center of electron generation closer to the gate and would reduce the tunnel length, as shown in Fig. 3.7(b). This greatly improves the transfer characteristics as shown Fig. 3.8. For the given TFET with slanted InAs/Si interface, only the increase of the InAs source doping can improve the BTBT current slope, as demonstrated for a value of $8 \times 10^{18}\,\mathrm{cm}^{-3}$. This is a result of the

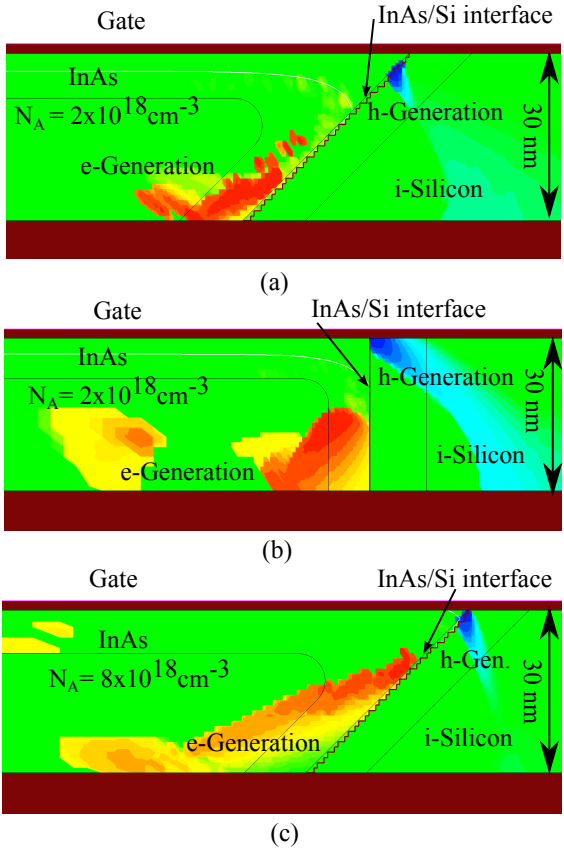


Figure 3.7: Color-mapped diagrams of electron and hole generation rates at (a) inclined InAs/Si interface, (b) vertical InAs/Si interface with moderate source doping, and (c) inclined InAs/Si interface with high source doping. In all cases the gate voltage is $V_{GS} = 0.2\,V$ and $V_{DS} = 0.5\,V$.

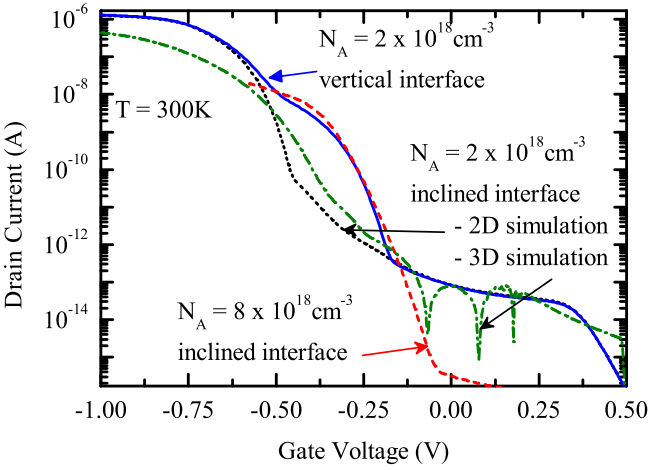


Figure 3.8: Simulated transfer characteristics of an ideal lateral InAs/Si NW TFET without traps for different device geometries and source doping levels. A vertical InAs/Si interface or high source doping can result in a sharp onset of the drain current ($V_{\mathrm{DS}} = 0.5\,\mathrm{V}$).

reduced tunnel length (see Fig. 3.7(c)). However, in the real TFETs the drain current component due to BTBT is completely dominated by the TAT current at the InAs/Si junction. Therefore, the above alterations will improve the subthreshold characteristics only after suppression of TAT.

Mechanism of operation

The swing extracted from the measured characteristics of the TFET is 70 mV/dec and 35 mV/dec at 300 K and 125 K, respectively. Contrary to the weak temperature dependence of an ideal TFET, the swing is roughly proportional to the temperature here, which is a characteristic of thermionic emission. Fig. 3.9 explains the reason for this anomaly observed in the experiments and the simulations. The band diagram along the channel of the TFET (Fig. 3.9(a)) for different gate bias values clearly reveals a large thermionic barrier in the Si channel region. This barrier arises from the high D_{it} at the InAs/Si interface, which creates a large sheet charge when occupied. This sheet charge results in an upward band bending at the InAs/Si interface. The entire set-up results in a two-step current conduction process: TAT at the InAs/Si interface and subsequent transport of holes to the drain by thermionic emission. The barrier is lowered with increasing negative gate bias. In the subthreshold regime the thermionic barrier is high, and thermionic emission of holes becomes the rate-determining step. Therefore, the slope of the drain current is typical for a thermal swing. At a higher gate bias, when the barrier is sufficiently low, TAT at the InAs/Si interface becomes the bottleneck and the drain current becomes determined by TAT. The two regimes in the transfer characteristics are clearly distinguishable at 125 K as shown in Fig. 3.9(b). In short, the TAT leakage current is blocked by the thermionic barrier which gradually releases the blockade with increasing gate bias. This results in a swing dominated by the thermionic emission mechanism. In this way, the presence of a high trap density at the InAs/Si interface causes the TFET to operate like a MOSFET with the intrinsic Si region as gated channel.

Note that the VB offset at the InAs/Si interface is 130 meV although it is not visible in the band edge diagram in Fig. 3.9(a). This is due to the discretization at the interface, which linearizes the otherwise

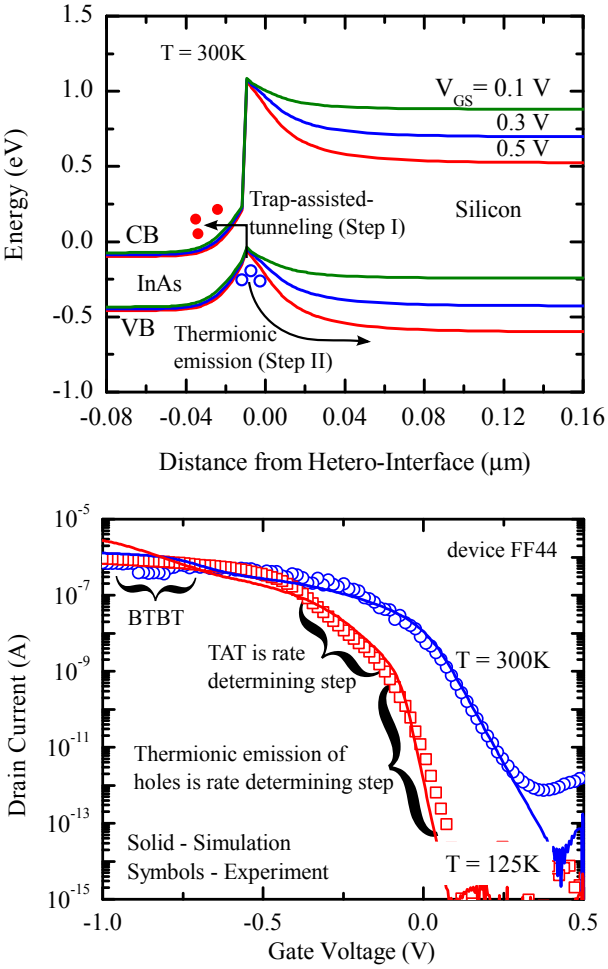


Figure 3.9: (a) Band diagram along the channel for three different gate bias values near the onset of the TFET. Transport of carriers happens via two subsequent steps, viz. TAT and thermionic emission. (b) Two separate branches in the transfer characteristics with dominant mechanisms in respective intervals ($V_{\mathrm{DS}} = 0.5\,\mathrm{V}$).

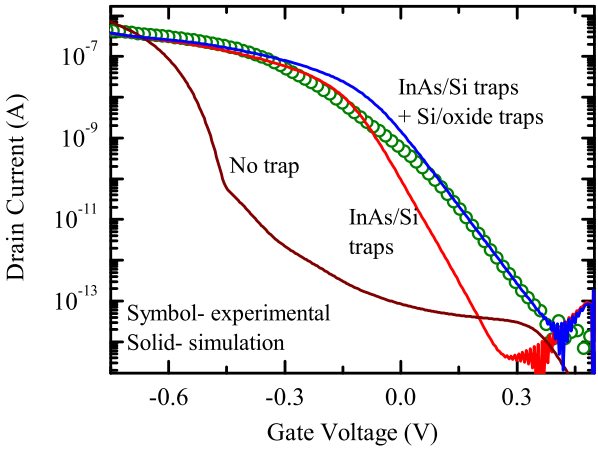


Figure 3.10: Effect of Si/oxide traps on the transfer characteristics of FF41 TFETs. Si/oxide traps reduce the electrostatic coupling between gate and channel. This degrades the swing similar to what is observed in a MOSFET ($V_{DS} = 0.5\,V$).

abrupt band offset, and due to the high electric field, which gives rise to a sharp gradient at the interface making it difficult to distinguish the band offset.

Effect of traps at the Si/oxide interface

In the analysis of FF44 TFETs, traps at the Si/oxide interface could be ignored. However, such traps seem to have a strong impact in FF41 TFETs, possibly due to a higher concentration as result of process variations. The effect of Si/oxide traps is highlighted in Fig. 3.10 which compares the transfer characteristics of FF41 TFETs with and without Si/oxide interface traps. It can be explained as follows. When Si/oxide traps are included, the slope of the transfer characteristics changes in the region where it is dominated by thermionic emission. Traps at the oxide/Si interface introduce an additional capacitance

in parallel with the depletion layer capacitance which degrades the electrostatic coupling between the gate and the intrinsic Si channel [55]. The reduced gate coupling increases the swing from its ideal value to 81 mV/dec for the above devices at 300 K. This effect is similar to what is observed in MOSFETs in the presence of oxide interface traps. Generation rate of carriers by multi-phonon excitation is weak at this interface due to large band gap of Si. Thus, the degradation of the swing due to Si/oxide traps can be entirely attributed to the reduced gate coupling.

It must be noted that a gate work function of 4.9 eV is used in the simulation of FF41 devices as compared to 5.05 eV for the FF44 devices. This change could be due to trapped oxide charges.

Prediction for D_{it} limits

Having achieved good agreement between the experimental and simulated TFET characteristics, the above simulation set-up can be utilized to study the impact of a given D_{it} on the swing. D_{it} at the hetero-interface is assumed to have a Gaussian energetic profile with the peak located at the VB edge of InAs. Simulations are carried out by varying the uniform trap density at the InAs/oxide interface and the peak trap density at the InAs/Si interface. A source doping of $8 \times 10^{18}\,\mathrm{cm}^{-3}$ is used in these simulations, as it is optimal for the given device geometry. The resulting transfer characteristics are shown in Fig. 3.11. It turns out that the device characteristics will degrade even if either of the two interfaces has a trap concentration above $1 \times 10^{12}\,\mathrm{cm}^{-2}\mathrm{eV}^{-1}$. Such a value has also been found as a minimum requirement for vertical InAs/Si NW TFETs with larger diameters as explained in the next section. As a more conservative estimate, the D_{it} at both the InAs/Si and InAs/oxide interfaces needs to be close to $5 \times 10^{11}\,\mathrm{cm}^{-2}\mathrm{eV}^{-1}$ to enable a swing $< 60\,\mathrm{mV/dec}$.

3.2 InAs/Si Vertical Nanowire TFETs

In this section, the role of interface traps and channel quantization in the degradation of, respectively, the swing and the on-current of vertical nanowire (diameter = 100 nm) InAs/Si pTFETs are analyzed.

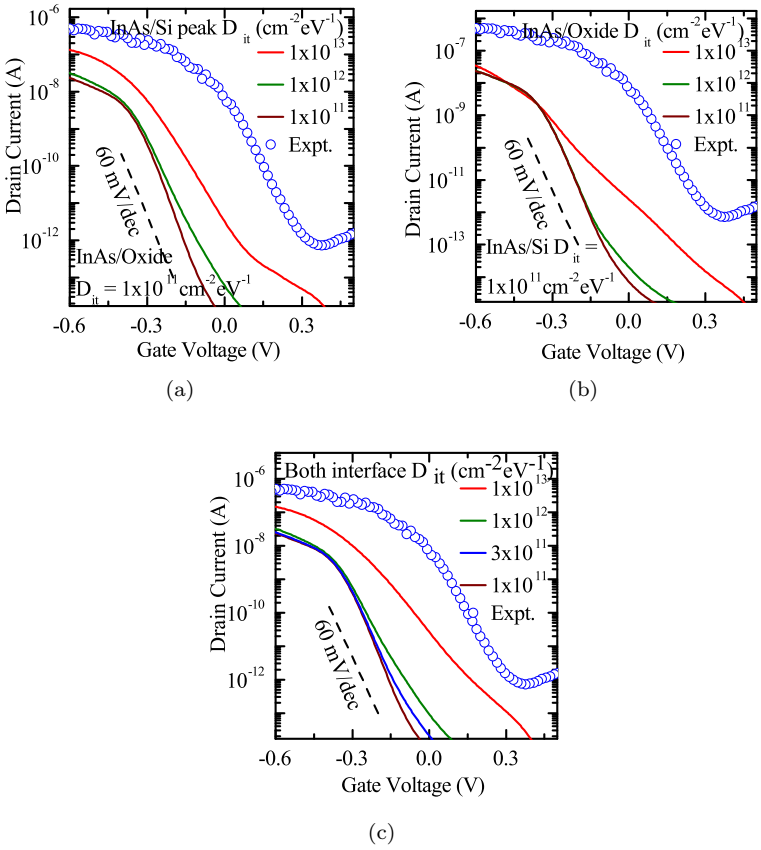


Figure 3.11: Predictive simulations to study the impact of changing D_{it} at (a) InAs/Si interface, (b) InAs/oxide interface, and (c) both InAs/Si and InAs/oxide interface.

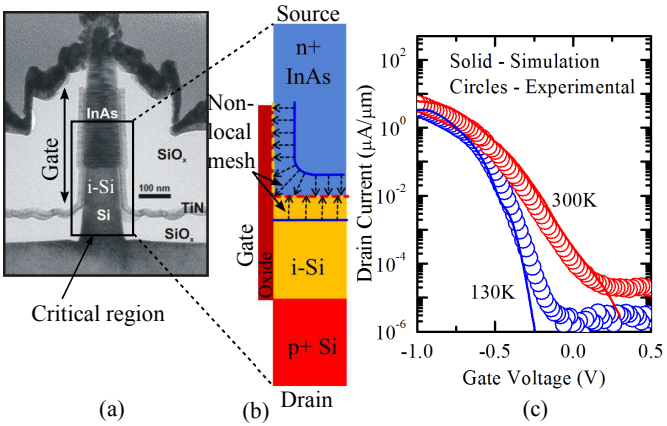


Figure 3.12: (a) TEM image of the InAs/Si vertical nanowire TFET (taken from [62]). (b) Radial cross section of the critical region for TCAD simulation. (c) Comparison of experimental and simulated transfer characteristics.

Results are compared with measured transfer characteristics, and the simulation set-up is employed to determine the device geometry which is least susceptible to TAT, i.e. trap-tolerant.

3.2.1 Simulation set-up

The experimental InAs/Si vertical nanowire TFETs consist of a p+ doped Si drain ($N_A = 3 \times 10^{19}\text{cm}^{-3}$), 100 nm long intrinsic Si channel, and a 500 nm long n+ doped InAs ($N_D = 2 \times 10^{18}\text{cm}^{-3}$) as shown by the TEM image in Fig. 3.12(a). The gate stack is Al_2O_3/HfO_2 with an EOT of 1.3 nm. The critical region of the device is simulated with S-Device [33] (see Fig. 3.12(b)). Length of GOS region in the vertical NW TFET shown in the TEM image in Fig. 3.12(a) is approx. 220 nm. In the simulated TFETs, this length is set to 50 nm to reduce computation time. This change in the device geometry would merely scale down the contribution of line tunneling to the total drain current leaving point tunneling unaffected. As explained in the subsequent subsections, line tunneling does not contribute to the

total on-current due to delay in the line tunneling due to channel quantization. Therefore, shrinking the length of the GOS region does not affect the results. Band gaps in Si and InAs are set to 1.1 eV and 0.36 eV, respectively, and the VB offset at the InAs/Si hetero-junction to 130 meV [61]. The gate metal work function is set to 4.8 eV. To account for the effect of quantization in a triangular potential well in the channel on tunneling, the BTBT model based on the path rejection method as described in Chapter 2 is implemented using the PMI in the simulator. Size quantization is ignored in Si due to the large effective mass. Both BTBT models require electron and hole effective masses as input parameters. They are set to their values in bulk. The band structure parameters are listed in Table 3.1. Since the CB of InAs is at the Γ-point, direct tunneling is expected to be the dominant tunneling mechanism between the CB of InAs and the VB of Si at moderate electric field [31]. Thus, phonon-assisted tunneling is not considered. The drain current is scaled by the circumference of the nanowire to compare the IV-plots at various diameters.

As shown in Fig. 3.13(a), TAT at the InAs/oxide interface involves multi-phonon excitation of electrons from the VB to the traps and subsequent tunneling to the CB. At the InAs/Si interface, traps can mediate a direct or phonon-assisted tunneling process between the Si VB and the InAs CB as shown in Fig. 3.13(b). Steady-state occupation of traps results from the principle of detailed balance and affects the electrostatics self-consistently. TAT at interface traps is modeled with the nonlocal model as described in the previous section. Traps at the InAs/Oxide interface are assumed to be uniformly distributed in the band gap. A maximum D_{it} of $1 \times 10^{13}\,\mathrm{cm}^{-2}\mathrm{eV}^{-1}$ is considered in accordance with Ref. [62]. Traps at the InAs/Si hetero interface are located close to the VB edge as observed in the study of the lateral InAs/Si nanowire TFETs. The energetic distribution of D_{it} at the InAs/Si interface is assumed to be Gaussian with its peak at the VB edge following the outcome of the simulations of lateral InAs/Si TFETs, in particular, Fig. 3.4. A field-induced shift in the trap energy level [63, 64] is ignored in the analysis. As seen from the calculations reported in Ref. [63], the shift of the trap energy level is less than 20 meV for electric fields up to 1 MV/cm. Such a shift is not expected to alter the results significantly. The peak D_{it} value is set to $1 \times 10^{13}\,\mathrm{cm}^{-2}\mathrm{eV}^{-1}$ and the FWHM to 220 meV which, on integration

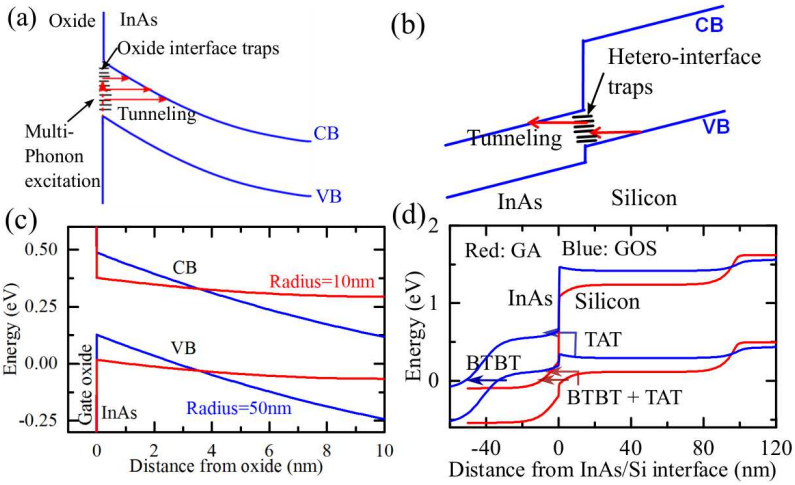


Figure 3.13: Schematic band edge diagrams showing TAT at (a) the oxide/InAs interface and (b) Si/InAs interface. (c) Extracted band edge diagrams normal to the oxide/InAs interface of gate-overlapped-source (GOS) TFET at the on-state $V_{GS} = -0.75$ V. A triangular well is reduced in the NW with $d = 20$ nm suppressing TAT. (d) Extracted band edge diagrams along the axis of the $d = 20$ nm InAs/Si NW pTFET with GOS and gate-aligned (GA) geometries at V_{GS} = -0.5 V.

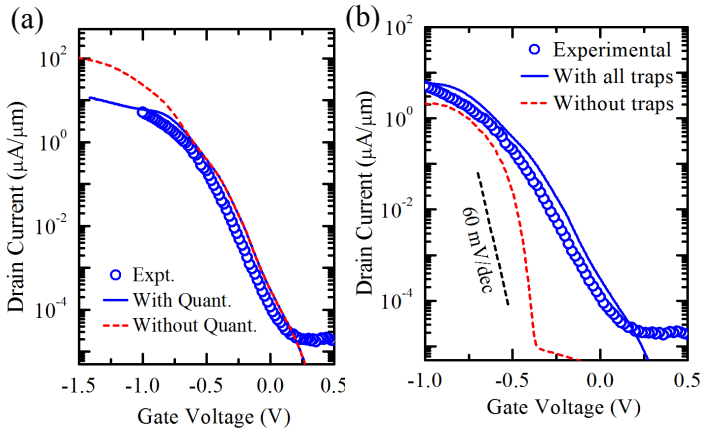


Figure 3.14: Comparison of the simulated transfer characteristics (a) with and without channel quantization, and (b) with and without interface traps. Channel quantization lowers the on-current, whereas interface traps degrade the subthreshold swing in TFETs.

over energy, gives a trap density of $2.7 \times 10^{12}\,\mathrm{cm}^{-2}$. This corresponds to approximately one trap per circular disk with a radius of 3.4 nm. TEM analysis of the InAs/Si interface reported by Tomioka et al. [65] reveals a periodic arrangement of dislocations with $\approx 3\,\mathrm{nm}$ spacing. If each dislocation contributes one trap level, the calculated D_{it} agrees well with the value used here. The parameters for the TAT model and the multi-phonon excitation model are listed in Table 2.1 of Chapter 2.

3.2.2 Results and discussion

The comparison of simulated and measured transfer characteristics of the fabricated TFET is shown in Fig. 3.12(c). Fig. 3.14(a) presents the effect of channel quantization, revealing a severe lowering of the on-current. Only the inclusion of channel quantization can match simulated and measured curves, which provides some confidence that the quantization effect is indeed present in these devices. The reduction of the on-current in the presence of channel quantization is due to

the rejection of shorter tunnel paths which leads to an increase of the average tunnel length and a reduced BTBT rate. Fig. 3.14(b) depicts the impact of TAT on the swing at 300 K. In the subthreshold region, the drain current is entirely dominated by TAT with negligible contribution from BTBT. The I_{D}-V_{GS} curve at lower gate voltage is solely degraded by surface SRH generation. With increasing voltage, i.e. sufficient band bending, tunneling from trap levels into the CB (Fig. 3.13(a)) starts to become effective. Since the tunnel rate between trap and CB states is much larger, multi-phonon excitation becomes the bottleneck. As multi-phonon excitation exhibits a small field enhancement compared to zero-phonon trap-to-band tunneling, the drain current increases sluggishly resulting in a strong swing degradation. At 300 K, the contributions of InAs/Si and InAs/oxide TAT to the total subthreshold drain current are of similar magnitude (see Fig. 3.15(b)). On the other hand, at 130 K the contribution of InAs/oxide traps becomes negligibly small due to the suppression of multi-phonon excitation of electrons from the VB as shown in Fig. 3.15(a). As a result, hetero-interface TAT is dominant in the subthreshold region which reduces the swing at lower temperatures. Since both the InAs/oxide and InAs/Si traps contribute to the total current at 300 K, the D_{it} at both interfaces must be suppressed to achieve a sub-thermal swing.

Parameters such as capture cross section, trap volume, and the product of Huang-Rhys factor and phonon energy are adjusted to obtain a good match to the experimental data. As mentioned in the analysis of InAs/Si lateral TFETs, small changes of these parameters do not strongly alter the transfer characteristics. The effect of trap volume and capture cross section variations on the transfer characteristics of the vertical TFETs is shown in Fig. 3.16. Changing both parameters over an order of magnitude merely scales up the total drain current. It does not change the shape of the transfer characteristics. Fig. 3.17 shows the impact of varying the Huang-Rhys factor by an order of magnitude. Since BTBT occurs between the the Γ-valley of InAs and the Light Hole (LH) band of Si, direct tunneling dominates the total tunnel rate. Similarly, TAT takes place between Γ-valley of InAs and the interface traps, or between LH band of Si and the interface traps. The direct tunneling is predominant in TAT as well. Therefore, the simulated drain current is insensitive to a variation of

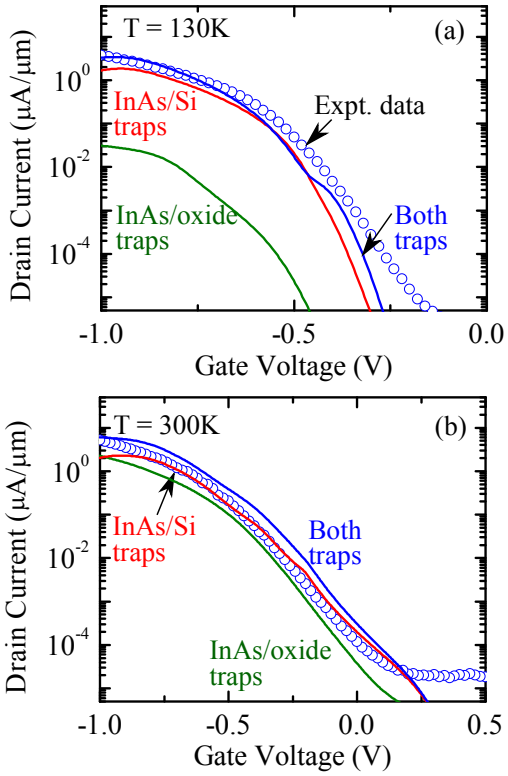


Figure 3.15: Transfer characteristics of the vertical NW TFET showing individual contribution of InAs/Si and InAs/oxide interface traps to the total drain current at two different temperatures ($V_{DS} = 0.5\,V$). Total current (blue) in (a) is slightly different from the sum of the current due to each individual trap-type (red and green). This is because of the change in the electrostatics after combining both the traps.

the Huang-Rhys factor. Thus, the simulations confirm that adjusting the above mentioned parameters only slightly changes the slope but does not alter the conclusions.

3.2.3 Design of trap-tolerant TFETs

Diameter scaling

The simulation of NW TFETs using different D_{it} at the InAs/Si interface confirms that passivation of traps at the InAs/Si interface alone would not yield a sub-thermal swing (see Fig. 3.18(a)). Therefore, other techniques to suppress TAT must be explored. Although it is tough (but not impossible) to passivate traps at the InAs/oxide interface, their contribution to the drain current via TAT can be reduced by decreasing the NW diameter d to 20 nm. This results in the steep I_{D}-V_{GS} plots shown in Fig. 3.18(b). The NW TFET with d = 100 nm undergoes surface inversion (see Fig. 3.13(c)) which creates a triangular-like potential well at the InAs/oxide interface. TAT is significantly enhanced at such a well due to tunneling between traps and the CB edge. On the other hand, the TFET with $d = 20$ nm undergoes volume inversion resulting in flat band conditions at the InAs/oxide interface (see Fig. 3.13(c) for comparison). In the latter case, only multi-phonon excitation can take place, TAT is inhibited due to insufficient band bending. Note that, due to the nearly flat bands at the InAs/oxide interface, the effect of channel quantization is absent in the TFET with $d = 20$ nm. Therefore, the default BTBT model is used instead, and the energy of the first sub-band is added to the CB edge. Electrostatic screening of the channel by the oxide interface traps [55] is still present. However, it does not degrade the swing as much as TAT. In this way, diameter scaling can bring down the minimum point swing from 157 mV/dec ($d = 100\,\mathrm{nm}$) to 68 mV/dec ($d = 20\,\mathrm{nm}$).

Gate alignment

Although the diameter scaling suppress TAT at the oxide interface, TAT at the InAs/Si hetero-junction continues to happen creating a leakage basin of the TAT current at the onset of BTBT in the pTFET.

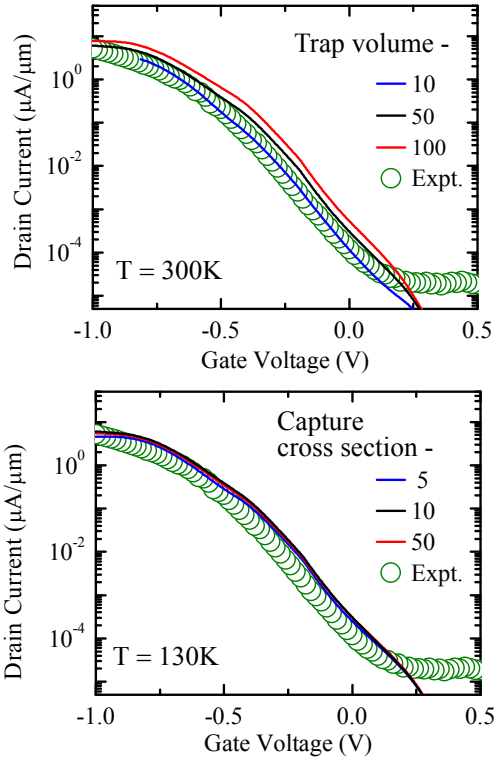


Figure 3.16: Effect of small variation of (a) trap volume (in Å^3) and capture cross section (in Å^2) on the transfer characteristics ($V_{\text{DS}} = 0.5\,\text{V}$).

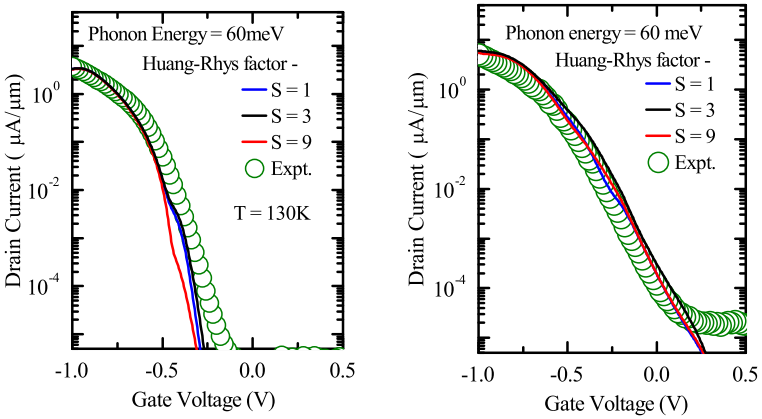


Figure 3.17: Effect of variation of the product $S\hbar\omega$ (in meV) on the transfer characteristics at (a) 130 K and (b) 300 K ($V_{\mathrm{DS}} = 0.5\,\mathrm{V}$).

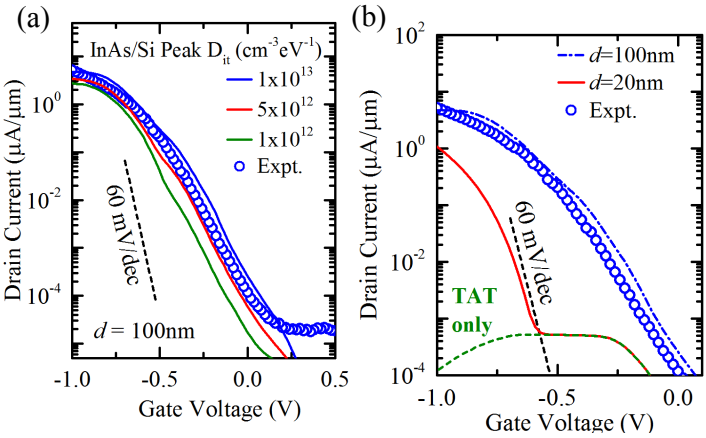


Figure 3.18: (a) Effect of passivating the InAs/Si hetero-interface in the InAs/Si pTFET ($d = 100\,\text{nm}$). The swing remains large despite reducing the D_{it}. (b) Simulated transfer characteristics of the InAs/Si pTFET with $d = 100\,\text{nm}$ and $20\,\text{nm}$. The individual contribution of hetero-junction TAT ($d = 20\,\text{nm}$) is plotted in dashed green.

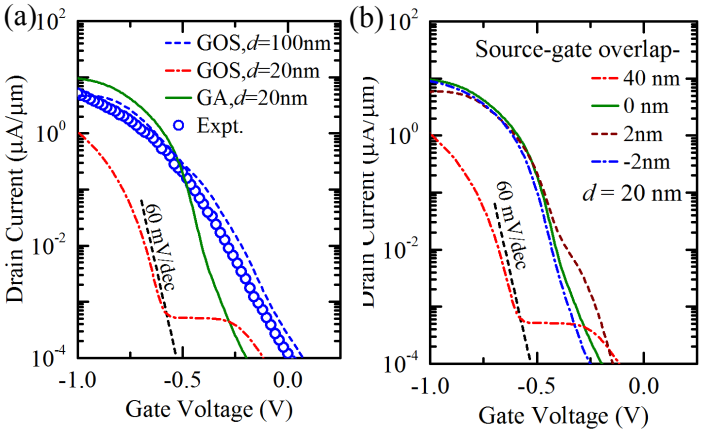


Figure 3.19: (a) Simulated transfer characteristics of the Gate-overlapped-Source (GOS) and Gate-aligned (GA) TFETs. The leakage floor disappears in case of the GA TFET. (b) Effect of a gate-source misalignment of ± 2 nm in the GA TFET. Gate-source overlap is set to 50 nm in the GOS geometry.

Since the lateral electric field at the hetero-interface saturates in the GOS TFET when volume inversion begins, the hetero-TAT current also saturates at 10 nA/μm (see Fig. 3.18(b)). This high value makes the TFET incapable for low-power application. To overcome this drawback, the gate must be aligned with the InAs/Si hetero-interface. Simulation results for the NW TFET (d = 20nm) with GA and with GOS geometry are compared in Fig. 3.19(a). It may be inferred from the I_{D}-V_{GS} plots that the gate alignment advances the onset of BTBT in the TFET. As seen from the band edge diagram along the axis of the two TFETs in Fig. 3.13(d), BTBT takes place at the source end of the gate in the GOS TFET, but at the InAs/Si hetero-interface in the GA TFET. Therefore, in the GOS geometry, BTBT begins only when the entire GOS region undergoes volume inversion which happens at a higher gate bias. This results in a delayed onset of BTBT relative to TAT which causes the leakage basin. On the contrary, in the GA geometry, BTBT begins as soon as accumulation starts in the i-Si channel. Hence, the onset of hetero-junction TAT coincides with the onset of BTBT (see Fig. 3.13(d)). The simultaneous onset of TAT and BTBT in GA TFETs improves the point swing to 59 mV/dec (from 69 mV/dec in GOS geometry). More importantly, it results in a lower leakage floor. Additionally, the GA geometry yields a higher on-current compared to the GOS geometry. In the GA TFET, BTBT takes place at the hetero-interface which offers a lower tunnel barrier due to the staggered band alignment at the interface. The transfer characteristics in Fig. 3.19(b) imply that already a small gate-source overlap degrades the swing. The swing is almost unaffected if an underlap is present between gate and source.

In summary, since the suppression of oxide- and hetero-interface traps is difficult, "trap-tolerant" TFET geometries need to be explored. In the case of an InAs/Si NW TFET, diameter scaling down to 20 nm would minimize TAT at the oxide-interface and gate alignment would avoid the leakage basin arising from the hetero-TAT when the defects are confined at the hetero-interface. The above approach may be applicable to TFETs with other material systems as well.

Chapter 4

Non-idealities in all-III-V Hetero-junction TFETs

The on-current of the TFET is strongly affected by the band gap of the semiconductor and the effective masses. Due to the small effective masses and the direct band gap, III-V materials are well suited for TFET application. The hetero-junction in the TFET can further improve the on-current by reducing the effective tunnel gap. Thus, III-V heterojunction TFETs can deliver both high on-current and low subthreshold swing. Experimental data of InGaAsSb/InAs NW TFETs presented in [66, 67] confirm the theoretical predictions.

In this chapter, the experimental data of InGaAsSb/InAs hetero-junction TFETs [66] are analyzed using TCAD to understand whether the non-idealities discussed in Chapter 2 degrade the subthreshold swing of all-III-V TFETs, which non-idealities affect these TFET the most, and what is the best performance the given TFETs can achieve.

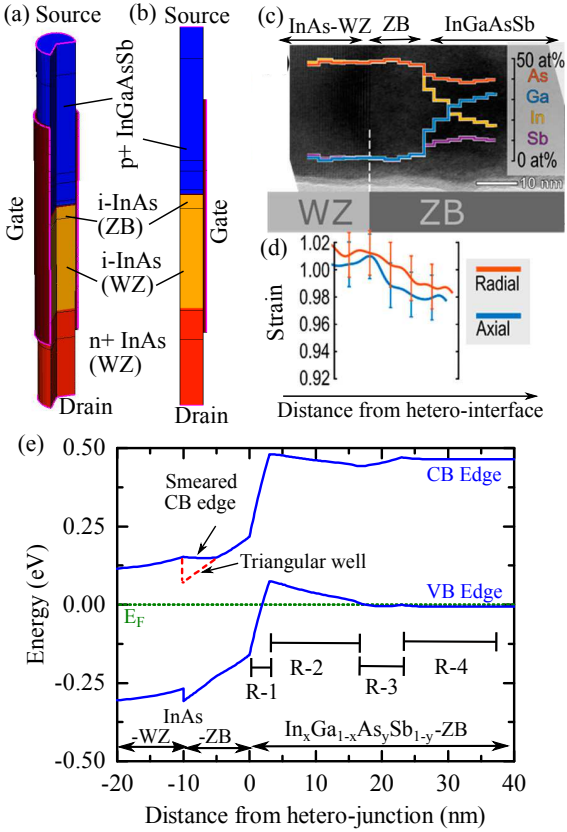


Figure 4.1: (a) Schematic of the simulated InGaAsSb/InAs NW TFET, (b) axial cross section, (c) TEM image of the hetero-interface, (d) lattice measurements, (e) band edge diagram along the axis of the NW at $V_{DS} = V_{GS} = 0$ V. A triangular-like quantum well is present at the InAs-ZB/InAs-WZ interface. Its effect on the band structure is mimicked by gradually changing the band gap of the material. The compositional changes in atomic % for respective element are shown in colored curves with legend and scale to the right, overlaying the TEM image of the transition.

4.1 Simulation Set-up

The vertical NW TFET consists of an intrinsically doped InAs channel, a p+ doped InGaAsSb source and a n+ doped InAs drain. The axial cross section of the NW is shown in Fig. 4.1(a). Cylindrical coordinate system is used in the simulation of the device. Using cylindrical coordinates enables to simulate a radially symmetric 3-dimensional device using its axial cross-section in 2D. This is done initially to speed-up the simulations. However, it cannot simulate the effect of a "single trap" (i.e. a spatially localized trap) in the bulk or a hetero-junction, if the trap is not located along the axis of the nanowire. This is because, the cylindrical coordinate system transforms the single trap into a co-axial ring. Therefore, true 3D simulations are performed later to assess the impact of single traps and their location on the device characteristics. The gate oxide is an Al_2O_3/HfO_2 bilayer with an EOT of 1.4 nm. The InAs/InGaAsSb/GaSb NW with a diameter of 40 nm was grown on the substrate using Metal Oxide Vapor Phase Epitaxy from Au seed particles which were patterned on a Si (111) substrate by electron beam lithography. The InAs/InGaAsSb segment was later trimmed to a diameter of 20 nm before depositing high-k dielectric. A detailed description of the device fabrication can be found in Ref. [66]. TEM analysis was performed on the NW to obtain information on the variation of material composition as well as its crystal structure along the NW axis [66]. The analysis reveals that the channel consists of a Wurtzite (WZ) segment of InAs followed by a 10 nm Zinc-blende (ZB) segment which is followed by the quaternary alloy InGaAsSb. The InGaAsSb segment, which forms the source, is *in-situ* doped with $N_A = 1 \times 10^{19} \mathrm{cm}^{-3}$. The InAs segment in the drain is *in-situ* doped with $N_D = 1 \times 10^{19} \mathrm{cm}^{-3}$. The gate overlaps with the 60 nm long InGaAsSb segment making it a GOS TFET. The TEM image of the hetero-interface is shown in Fig. 4.1(c) along with the plot of the composition of InGaAsSb. The variation of the uni-axial strain along the NW axis can be inferred from the lattice measurements reported in Fig. 4.1(d).

4.1.1 Band structure modeling

Adjacent to the hetero-interface, the composition of InGaAsSb changes rapidly along the axis which results in a variation of the band gap as well as the electron and hole effective masses. Since BTBT is strongly affected by the band gap and the effective masses, it is necessary to consider the variation of the material composition in the simulation. The InGaAsSb segment of the NW is divided into four regions, R-1,..., R-4. In region R-1, the alloy composition varies from InAs (to the left) to $In_{0.7}Ga_{0.3}As_{0.84}Sb_{0.16}$ (to the right), denoted as $In_{1\rightarrow0.7}Ga_{0\rightarrow0.3}As_{1\rightarrow0.84}Sb_{0\rightarrow0.16}$. Region R-2 consists of $In_{0.7\rightarrow0.44}Ga_{0.3\rightarrow0.56}As_{0.84\rightarrow0.72}Sb_{0.16\rightarrow0.28}$. Region R-3 is composed of $In_{0.44\rightarrow0.32}Ga_{0.56\rightarrow0.68}As_{0.72}Sb_{0.28}$ whereas R-4 consists of $In_{0.32}Ga_{0.68}As_{0.72}Sb_{0.28}$. In each segment, the band gap and the effective masses are set to the experimental values taken from Ref. [58]. To obtain the values for intermediate compositions, an interpolation formula suggested by Adachi [68] is used. The material composition is assumed to vary linearly within each region. This doesn't necessarily imply a linear variation of band gap and effective masses, since bowing of these quantities for intermediate compositions is taken into account in the calculations. The energetic position of the CB edge is set to the electron affinity of the semiconductor/alloy taken from Ref. [53]. Thus, the band alignment at all the semiconductor interfaces follows Anderson's rule. The abrupt change in the crystal structure of InAs from WZ to ZB during the growth is taken into account by using the band structure parameters of the respective allotropes. In addition to the composition variation, uniaxial compressive strain is present at the hetero-interface. The position-dependent uniaxial stress along the axis is extracted from the high-resolution TEM images [66] and plotted in Fig. 4.1(c). The effect of strain on band gap and band alignments has been modeled with the model-solid theoretical approach by Van de Walle [69]. The temperature dependence of the band gap is modeled using Varshni's law [59]. The parameters α and β of Varshni's model are taken from the literature [58] and interpolated using Adachi's interpolation formula for quaternary alloys. In this way, any explicit fitting of band structure quantities and band offsets has been avoided. The band diagram at zero bias extracted along the NW axis of the

Table 4.1: Composition-dependent band structure parameters at the end of each segment in the NW.

InGaAsSb Composition	Strain	E_g (including strain effect)	m_C	m_V
InAs(WZ)	0%	0.42	0.042	0.084
InAs(ZB)	-2%	0.376	0.026	0.026
$In_{0.7}Ga_{0.3}As_{0.84}Sb_{0.16}$	-1%	0.405	0.036	0.038
$In_{0.44}Ga_{0.56}As_{0.72}Sb_{0.28}$	0%	0.432	0.043	0.048
$In_{0.32}Ga_{0.68}As_{0.72}Sb_{0.28}$	0%	0.47	0.048	0.077

TFET is plotted in Fig. 4.1(e). Band structure quantities used at the beginning/end of each of the above segments are listed in Table 4.1.

In addition to the above variations, quantum confinement may also alter the device characteristics. The inability to account for quantum-confinement effects is a drawback of semiclassical simulations. An abrupt change in the crystal structure from InAs-WZ to ZB results in a band offset at the interface which gives rise to a triangular-like potential well. Calculations suggest that the quantization of the CB states in the well results in the formation of only one bound state close to the top of the finite barrier. This effectively smears out the otherwise steep well at the interface. Simulations confirmed that, ignoring this effect causes a strong discrepancy between simulated and measured IV-characteristics. Therefore, quantization has been modeled by introducing a pseudo-grading of the CB edge at the interface thus making it continuous. The small effects related to quantum confinement normal to the transport direction are ignored.

4.1.2 Trap distribution and modeling of TAT

The analysis of the TEM images reveals that the composition of InGaAsSb changes sharply in region R-1 which may induce defect states in this region. The degradation of the TFET performance due to TAT at the hetero-interface is simulated using the nonlocal TAT model. The nonlocal TAT models are activated on a non-local mesh

constructed at the hetero-interface. Since the exact location of the traps in the segment R-1 is unknown, a constant density of bulk traps (D_{bt}) in this region is initially assumed. Choosing $D_{\mathrm{bt}} = 1.6 \times 10^{18}\,\mathrm{cm}^{-3}$ gives, on spatial integration, approximately one trap in region R-1. The variation of the defect energy level in the band gap leads to best agreement with the experimental IV-data when it resides 100 meV above the VB edge. The trap interaction volume is set to $50\,\mathrm{Å}^3$ to match the simulated current level to the experimental one. This parameter is related to the volume of the localized wave function. Parameters used in phonon-assisted TAT model, such as Huang-Rhys factor and phonon energy are set to 3 and 7 meV, respectively. This value of phonon energy corresponds to the transverse-acoustic phonon in InAs.

In addition to traps near the hetero-junction, traps at the InAs/oxide interface may degrade the TFET performance. Donor-like traps are introduced at this interface. Their energetic distribution is extracted from capacitance-voltage (CV) measurements of InAs NWs [23] and adapted here. Traps may also exist in the band gap at the InGaAsSb/oxide interface. Since the Fermi level in InGaAsSb is located below the VB edge, these interface traps are empty and hence electrostatically inactive.

The gate work function (WF) is obtained by horizontally shifting the simulated I_D-V_{GS} plots for the best match with the experimental data. The WF is found to vary linearly from 4.86 eV to 4.92 eV as the temperature decreases from 298 K to 223 K. This variation could be due to the freezing of charges at certain oxide traps causing a shift of the WF. A similar WF shift with temperature in InAs NW MOSFETs had been observed earlier [70]. The NW TFET is simulated with S-Device [33] using the above-described simulation set-up.

4.2 Simulation Results and Discussion

4.2.1 Temperature dependence

A comparison of simulated and measured temperature-dependent transfer characteristics is shown in Fig. 4.2(a). The good agreement confirms the validity of the simulation set-up. Note that only the

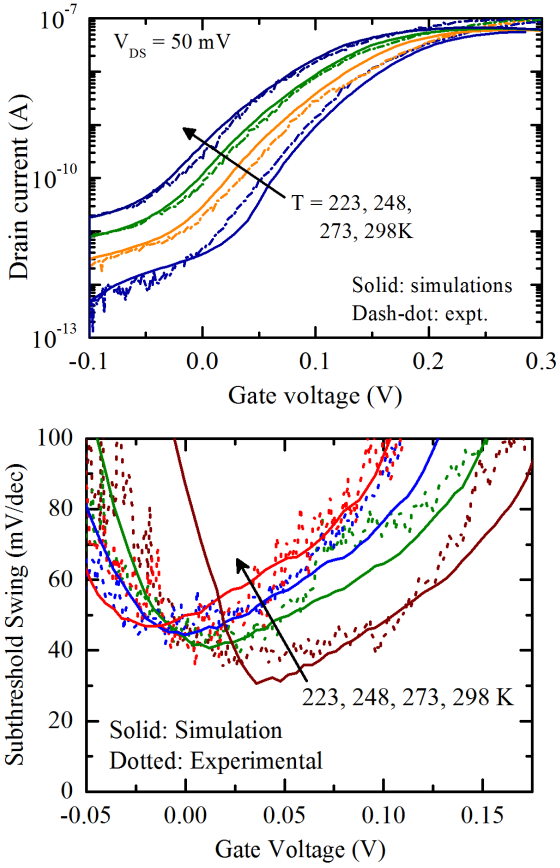


Figure 4.2: (a) Comparison of simulated temperature-dependent transfer characteristics of the TFET with the experimental data. (b) Comparison of the subthreshold swing calculated from the simulated and the experimental transfer characteristics. A reduction of the minimum swing with decreasing temperature is visible.

density of traps in region R-1 is fitted to the IV-curves. The remaining simulation parameters are taken from published experimental data. The good agreement also validates the above approach to obtain band gap, band alignment and effective masses by interpolating their experimental values to the extracted position-dependent material compositions.

The simulated on-current is about 30% smaller than the measured one. This difference could be due to an inevitable error introduced in determining the composition of InGaAsSb from the TEM image. It could also result from neglecting quantum confinement in radial direction or from neglecting BGN in the source/drain region. To assess the influence of sub-band formation due to quantum confinement in radial direction, the TFET is simulated by adding energy of the first electron and hole subband (obtained from quantum mechanical calculations) to the CB and the VB edges, respectively. This increases the effective tunnel gap at the hetero-interface which reduces the on-current as shown in Fig. 4.4. When the impact of BGN is included in the simulations, the VB edge in the source region is shifted upward due to BGN which shrinks the effective tunnel gap at the hetero-interface thus increasing the total current to the level of the experimental data (see Fig. 4.4). In this way, errors caused by neglect of quantum confinement and BGN cancel each other. Both the effects are ignored in the simulations presented below.

The temperature-dependent swing is obtained from the transfer characteristics and plotted in Fig. 4.2(b). Within an interval of about 100 mV, the swing lies below the thermal limit of 60 mV/dec. The minimum point slope exhibits a weak temperature dependence only. It stems from the temperature dependence of the band gap in the rate of direct BTBT in III-V materials. This can be inferred from the $I_{\mathrm{D}} \propto$ WKB probability $\propto \exp(-\frac{\sqrt{2m_{\mathrm{red}}E_{\mathrm{g}}^3}}{3q\hbar F})$. Assuming that the local electric field is proportional to the gate voltage, the slope of I_{D} as a function of $V_{\mathrm{GS}}(= \frac{d(\log(I_{\mathrm{D}}))}{dV_{\mathrm{GS}}})$ is proportional to $E_{\mathrm{g}}^{3/2}$. Hence, at a given V_{GS}, the swing is expected to decrease with the slight widening of the gap caused by the decreasing temperature.

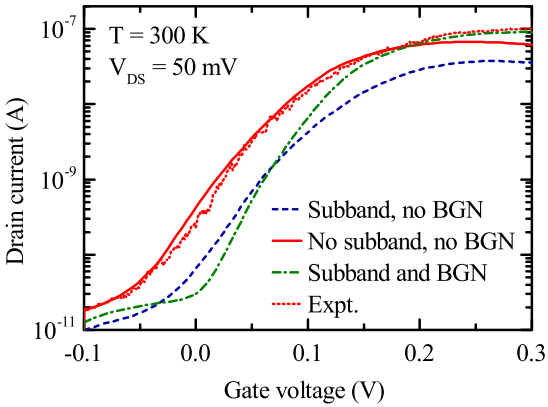


Figure 4.3: Effect quantum confinement in the radial direction (which results in subband formation) and band-gap narrowing (BGN) on the transfer characteristics of the TFET.

4.2.2 Effect of traps

InAs/oxide traps and bulk traps adjacent to the hetero-interface affect the performance of the TFET by strikingly different mechanisms. One can study them separately by activating each trap type individually. The transfer characteristics simulated by instantiating each of the two trap types are plotted in Fig. 4.4(a).

Tunneling at bulk traps takes place near the hetero-interface in two steps. An electron either tunnels (with or without phonon assistance) from the VB to the trap or undergoes a multi-phonon excitation to the trap level creating a hole in the VB. The trapped electron then undergoes another tunneling or multi-phonon excitation step to excite to the CB. Influence of multi-phonon excitation can be assessed by comparing the transfer characteristics with and without multi-phonon excitation (i.e. by setting capture cross-section to 0 $Å^2$). Similarly, impact of phonon assistance to the tunneling process between the trap state and each of the CB and the VB can be determined by deactivating the phonon-assisted tunneling process (by setting Huang-Rhys factor

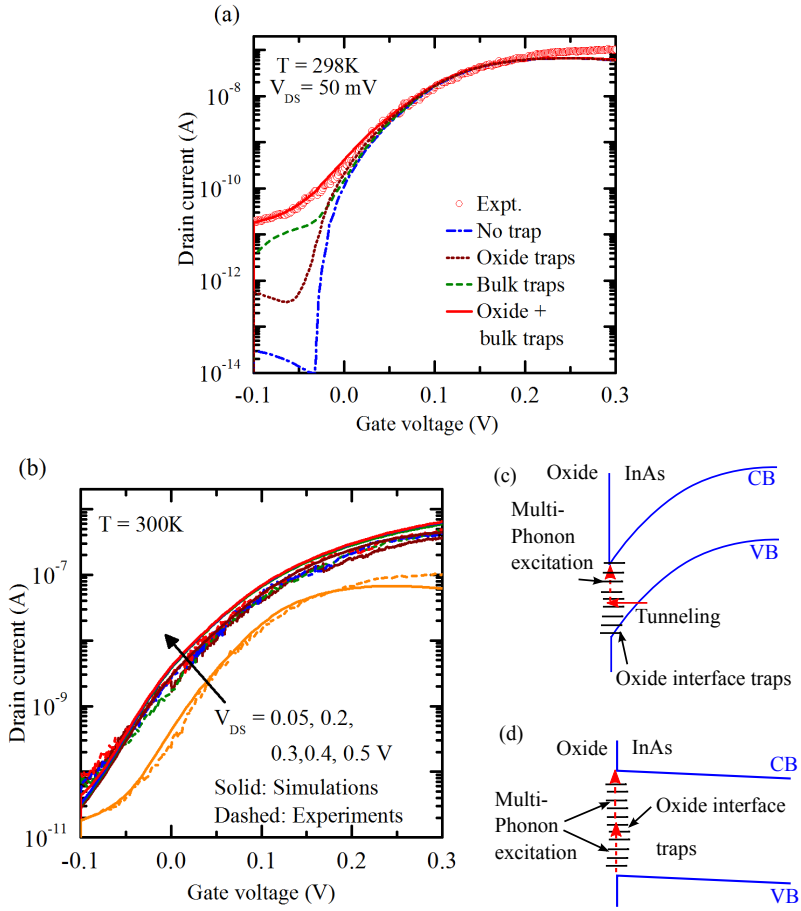


Figure 4.4: (a) Transfer characteristics of the TFET taking into account oxide and bulk traps separately. (b) Threshold voltage shift as observed in experiment with increasing V_{DS} is reproduced in the simulation. (c) Schematic diagram of trap-assisted tunneling and (d) multi-phonon excitation processes at the oxide/InAs interface.

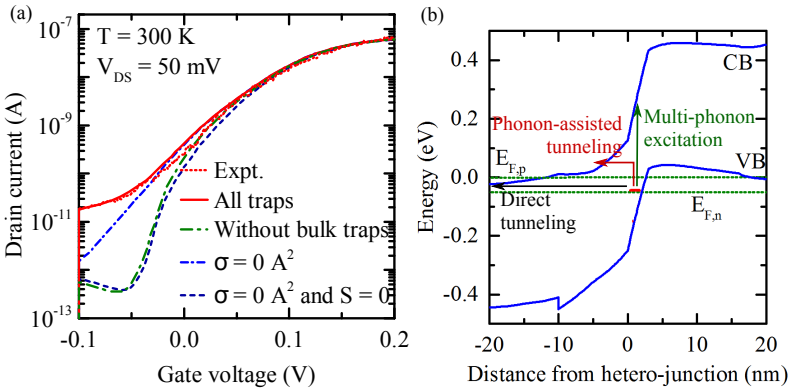


Figure 4.5: (a) Contribution of multi-phonon excitation and phonon-assistance in the TAT process at InAs/InGaAsSb hetero-interface. (b) Sketch of different processes by which an electron can get excited from the trap state to the CB at the hetero-interface in the given TFET. Band diagram along the channel at $V_{GS} = -0.1\,V$ and $V_{DS} = 50\,mV$ is used for the sketch.

to 0). Transfer characteristics obtained by deactivating each of the above processes are plotted in Fig. 4.5(a). The figure shows that, strong off-state leakage arises from multi-phonon excitation at the hetero-interface. On the other hand, phonon-assisted TAT causes degradation of the subthreshold swing. Direct tunneling is nearly absent in the off-state due to the long tunnel path necessary for the tunneling (see Fig. 4.5(b)). In the present device, the hetero-junction TAT results in a severe degradation of the subthreshold swing. It also gives rise to a large leakage current which could render the TFET less attractive for low-power applications. Suppressing bulk traps near the hetero-junction will significantly improve the TFET characteristics.

At the oxide/semiconductor interface, the TAT mechanism is different. When a triangular-like well is formed at the oxide interface, an electron can tunnel from the VB of the semiconductor to the interface trap creating a hole in the VB. This is followed by a multi-phonon transition of the electron to the CB as shown in Fig. 4.4(c).

Thus, band bending at the oxide interface is necessary for the TAT process to take place. In the absence of sharp band bending, TAT at the oxide/semiconductor interface is inhibited. Then, both the capture of an electron from the VB and its emission to the CB take place by multi-phonon excitation which is a much weaker generation process. In the present TFET, the InAs channel is intrinsically n-doped. As the gate bias is ramped up, the narrow InAs segment enters the accumulation region. Therefore, the necessary band bending does not happen and the triangular well is not formed at the oxide/InAs interface which inhibits tunneling. Still, oxide interface traps result in surface SRH generation of electron-hole pairs by multi-phonon excitation, which give rise to a weak leakage current ($\approx 1\,\mathrm{pA}$ as seen in Fig. 4.4(a)). Oxide/InAs traps change the electrostatics of the TFET and cause a slight degradation of the swing as observed in Fig. 4.4(a). This is due to electrostatic screening of the InAs channel by the oxide interface traps which reduces the coupling between gate and channel [55]. A variation of the threshold voltage with V_{DS} is observed experimentally. This variation is reproduced in the simulations as shown in Fig. 4.4(c). The threshold voltage is shifted by nearly 60 mV when V_{DS} is increased from 50 mV to 200 mV. For $V_{\mathrm{DS}} > 200\,\mathrm{mV}$, the threshold voltage shift is negligible in the experimental as well as simulated transfer characteristics. This is a result of the shift in the Fermi level with V_{DS} in the channel region which changes the occupancy of the oxide/InAs interface traps and alters the interface charge density. A rising V_{DS} increases the positive charge density of the oxide/InAs interface which acts as a positive gate charge and shifts the onset of tunneling to a lower gate bias. In this way, the threshold voltage is changed. This V_{T} shift as function of V_{DS} vanishes as soon as the oxide/InAs interface traps are deactivated in the simulation, confirming that the effect has no other reason.

4.2.3 Variability induced by traps

In the above analysis, a spatially uniform distribution of bulk traps in the segment R-1 had been assumed. In reality, the trap wave function is expected to be strongly localized at the defect location. To understand how the strong spatial localization of the trap charge changes the picture, a single trap with an energy level of 100 meV

above the VB edge is introduced in the segment R-1. The location of the trap is varied along the axis of the NW and also within a circular disk 2 nm away from the hetero-interface. The corresponding transfer characteristics are shown in Figs. 4.6(b) and 4.6(c). Since the trap energy is kept constant at 100 meV above the VB edge, moving the trap away from the hetero-interface results in a shift of the energy level relative to the Fermi energy (see Fig. 4.6(a)). This has the consequence that the trap energy level approaches the "tunnel window" (the energy interval at the hetero-junction in which the BTBT rate is maximum) when the trap is 1 nm away from the hetero-interface. Shifting the trap further to a location 2 nm away from the interface moves the trap energy level out of the tunnel window. Therefore, the TAT current is high when the trap is located 1 nm away from the interface, but reduces when the trap approaches a distance of 2 nm from the interface, as can be observed in Fig. 4.6(b). In a similar way, as the trap is displaced from the axis along the radius, the trap level moves deeper into the tunnel window which results in an increasing TAT. Ref. [71] reports a statistical analysis of a number of NW TFETs fabricated using the same process flow. The variation of the swing observed in these TFETs is in the same range as the simulated data. This is a strong indication that the subthreshold swing variations are caused by the random location of a single trap in the segment R-1.

4.2.4 Impact of band tails

Band tails may exist in the InGaAsSb source segment due to heavy doping and crystal defects. The effect of band tails on TFET characteristics is simulated using the semiclassical model developed in Chapter 2. This model has been implemented in S-Device via the nonlocal PMI. The modified form of model-1 described in Chapter 2.4 which is recommended for the simulation of hetero-junction TFETs is used here. It has two input parameters: the characteristic energetic width of the band tail (denoted by η) and the effective mass of an electron in the tail state (denoted by m_{t}). VB tails are introduced in the source region of the TFET under study. The parameter η is varied from 0 to 100 meV and m_t is varied from 0.5 m_0 to 0.025 m_0. None of the values resulted in a reasonable match. Therefore, m_{t} is set to m_{lh} (a lower limit defined by the light-hole mass). A characteristic energy of $\approx 25\,\mathrm{meV}$ has been

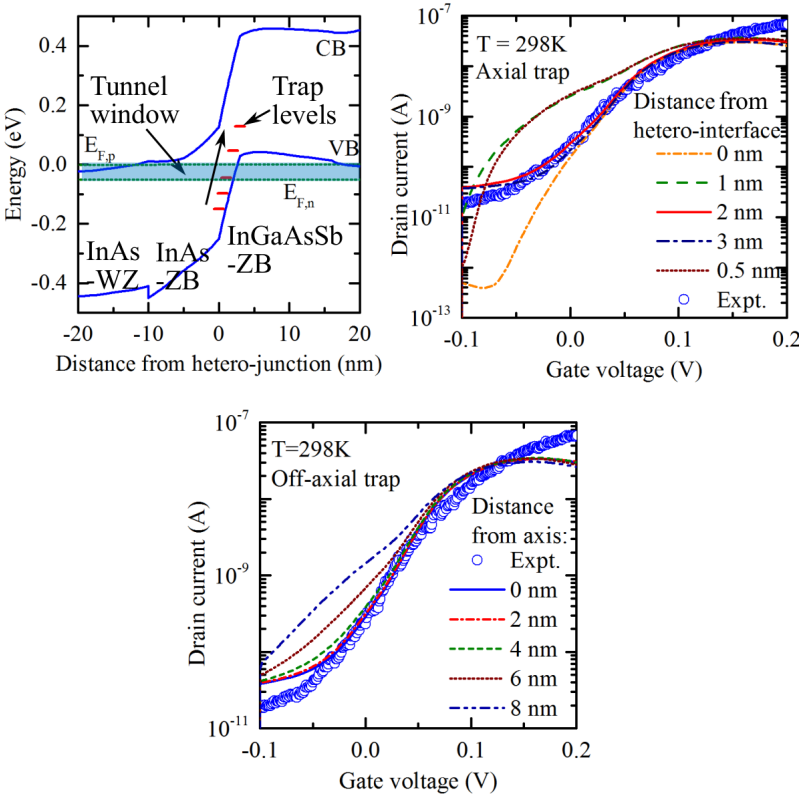


Figure 4.6: (a) Energetic and spatial location of a single trap leading to the I_D-V_{GS} characteristics by 3D simulations of the NW TFET shown in Fig. 1(a). (b) Variation of the location of a single trap along the axis of the NW in segment R-1. (c) Variation of the location of the trap along the radius, at an axial distance of 2 nm.

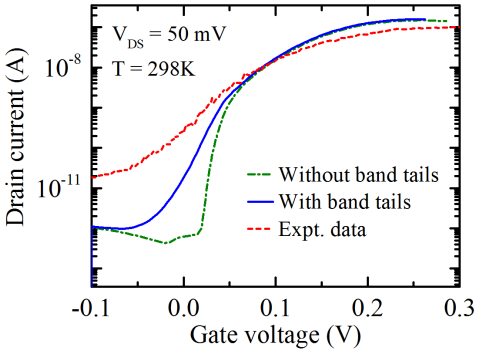


Figure 4.7: Impact of VB tails in the heavily doped source and in i-InAs on the transfer characteristics of the NW TFET. Parameters: $\eta = 25$ meV, $m_t = m_{lh}$.

reported in the literature for band tails in InSb [72] and GaAs [73]. Therefore, η is set to 25 meV. Bulk traps are deactivated during the simulation. The transfer characteristics obtained by simulating the TFET with the band tail model are presented in Fig. 4.7. VB tails in the source certainly degrade the swing, but not to the extent as done by bulk traps.

It may be inferred from the above analysis, that the measured subthreshold swing and leakage floor can only be reproduced by bulk traps in the segment R-1. This does not rule out the occurrence of band tails. The latter may become relevant if bulk traps are absent in the segment R-1.

4.2.5 Origin of negative trans-conductance

The measured transfer characteristics at $V_{DS} = 50$ mV show a peculiar reduction in the drain current with increasing gate bias above $V_{GS} \approx 0.3$ V as shown in Fig. 4.8(a). The simulated transfer characteristics (also shown in Fig. 4.8(a)) correctly reproduce this "negative trans-conductance region" although the simulated peak current is about

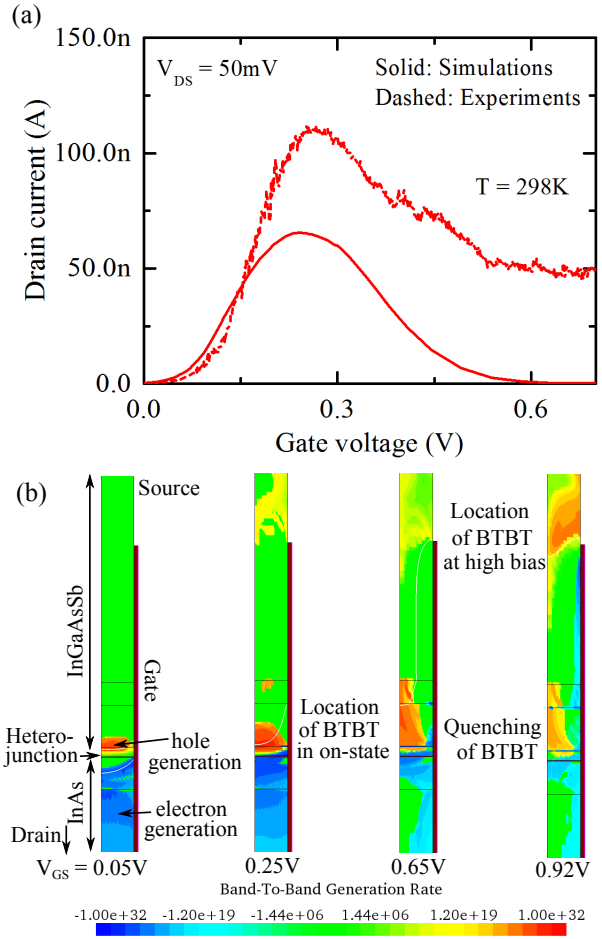


Figure 4.8: (a) Measured and simulated transfer characteristics over an extended gate voltage range. (b) Color-mapped diagram of the BTBT rate in the axial cross section of the NW TFET at different V_{GS} values. Quenching of BTBT generation at the hetero-junction causes negative trans-conductance.

30% lower than the measured one. The origin of this negative trans-conductance is explained as follows. At low gate bias, electrons accumulate in the i-InAs channel while the p+ doped InGaAsSb source is under weak inversion. As a consequence, BTBT starts at the InGaAsSb/InAs hetero-interface. With increasing gate bias, BTBT at the hetero-interface becomes stronger and reaches its maximum at $V_{GS} \approx 0.25\,V$. Beyond this voltage, volume inversion begins in the entire source overlapped by the gate. This leads to a gradual quenching of the BTBT generation at the hetero-interface till $V_{GS} \approx 1\,V$. This causes the current to reduce gradually giving rise to the negative trans-conductance region. The stages of building up the BTBT generation rate at the hetero-junction and its subsequent quenching are discernible in Fig. 4.8(b) and confirm the above hypothesis. The possibility of obtaining a negative trans-conductance regime in narrow NW TFETs is suggested in Refs. [74, 75]. The transfer characteristics of the present TFET give an evidence of this phenomenon.

4.2.6 Improving the device geometry

In the study of vertical InAs/Si TFETs presented in Chapter 3, gate alignment is found to improve the swing of the TFET even in the presence of high trap concentrations at the hetero-interface. In the above TFET, bulk traps adjacent to the hetero-junction degrade the swing. To assess whether gate alignment can mitigate their detrimental effect, the device is simulated with a gate-aligned InAs/InGaAsSb hetero-junction. Transfer characteristics with and without gate alignment are presented in Fig. 4.9(a). The swing as well as the on-current improve as soon as the gate is aligned with the pn-junction (which is also the hetero-interface). In this situation, the electric field at the junction is superimposed on the strong field at the source edge of the gate. This amplification of the field improves the swing as well as the on-current. Note that both the distributions of bulk traps and oxide/InAs traps are kept the same as before.

Although the gate alignment improves the TFET characteristics, a small misalignment will degrade swing and on-current significantly. The transfer characteristics of the device with a 5 nm overlap on the InGaAsSb source and the one with an underlap region of 5 nm between the source and the gate are shown in Fig. 4.9(b). As soon as

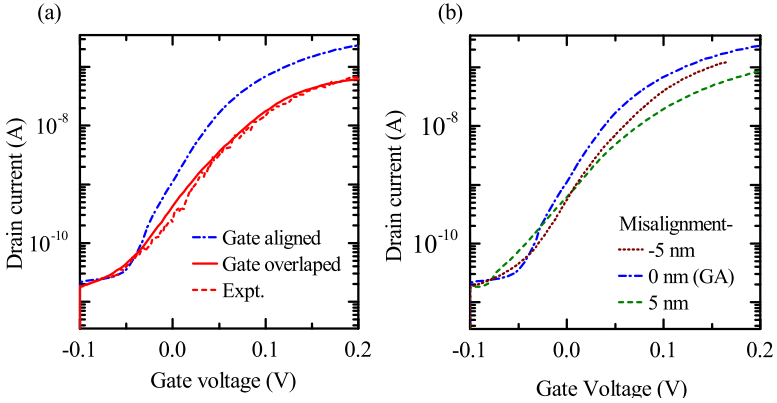


Figure 4.9: (a) Transfer characteristics of the NW TFET with the gate aligned to the InAs/InGaAsSb hetero-junction and with the gate overlapped on the source (original device geometry). (b) Impact of an overlap of 5 nm and an underlap of 5 nm on the transfer characteristics compared to the gate-aligned TFET.

the overlap region is introduced, the improvement over the original TFET geometry is reversed. On the other hand, an underlap of 5 nm results in a less severe degradation of swing and on-current.

In the study on vertical NW InAs/Si TFETs, it turned out that gate alignment advances the onset of BTBT with the result that it coincides with that of hetero-interface TAT thus eliminating the leakage floor caused by the latter. This is not observed here. The reason is the nearly broken-gap band alignment at the InAs/InGaAsSb hetero-interface as opposed to a staggered-gap band alignment at the InAs/Si interface. If the 3 nm long R-1 segment is ignored, the InAs-ZB/InGaAsSb interface exhibits a broken gap. As a consequence, both TAT and BTBT take place at the hetero-interface, at the same gate bias, even in the presence of gate overlap. Hence, BTBT cannot be advanced any further by the band alignment.

Chapter 5

Ab-initio Modeling of the InAs/Si Interface

The technological advancement of Si CMOS technology is now saturating due to the intrinsic physical constraints of Si as a semiconductor. To circumvent these limitations, the integration of various III-V materials on Si has been proposed [76, 77]. Furthermore, monolithic integration of opto-electronic devices such as LASERs and LEDs on ICs for on-chip photonics may be achieved by the growth of III-V on Si [78]. This creates a III-V/Silicon hetero-junction which is known to exhibit a high defect density due to the large lattice mismatch between Si and many III-V semiconductors. These defects give rise to interface traps which degrade the device performance. In Chapter 3, we have shown that traps at the InAs/Si interface degrade their performance and inhibit sub-thermal operation of the TFET. The investigation into origin and nature of such interface traps using *ab-initio* atomistic modeling is a useful first step in mitigating their detrimental impact.

The physics-based TCAD simulation of InAs/Si heterojunction NW TFETs presented in Chapter 3 yielded good agreement with experimental temperature-dependent I_{D}-V_{GS}-plots [57] only when the peak energy of the InAs/Si D_{it} is chosen in the lower half of the InAs band gap. This led to the presumption that the degradation of the TFET characteristics originates from the D_{it} in the lower half of the

gap. In the following, an attempt is made to understand the origin of traps at the InAs/Si interface by performing electronic structure calculations using Density Functional Theory (DFT). In the first step, structural minimization is performed on a relaxed InAs/Si slab to determine the atomic positions at the interface in a minimum energy configuration. This atomic structure is used in the next step in which electronic structural calculations are performed to find the origin of localized trap states at the interface.

5.1 Simulation Set-up

5.1.1 Si(111) surface reconstruction during growth

Typically, Si(111) is used for the growth of III-V semiconductors due to the possibility of generating an atomistically flat Si(111) surface by chemical etching. This technique had been employed for the growth of InAs on Si in the fabrication of the TFETs mentioned above. The Si(111) surface exhibits three types of reconstruction, namely, 1x1, 1x2, and 7x7 reconstruction [79]. A surface treatment just before the growth typically involves Hydrogen Fluoride etching of a thin native oxide layer on the Si(111) surface. This and the wet hydroxide etching of Si results in a hydrogen-passivated Si(111) surface [80]. Therefore, quasi-ideal 1x1 reconstructed H-passivated Si(111) is most likely to be present during the growth phase. Although the 7x7-reconstructed Si(111) surface is the most stable one among all the reconstructions up to 850°C, it involves a tremendous rearrangement of Si atoms at the surface. Therefore, it is not formed during wet etching of Si. Cleaving the Si crystal along (111) usually yields a metastable 2x1 reconstruction at room temperature. To the best of the author's knowledge, there are no reports confirming the occurrence of a 2x1 of Si during wet etching of Si or HF etching of oxide. Therefore, Si(111) with 1x1 reconstruction is assumed to be present at the InAs/Si interface.

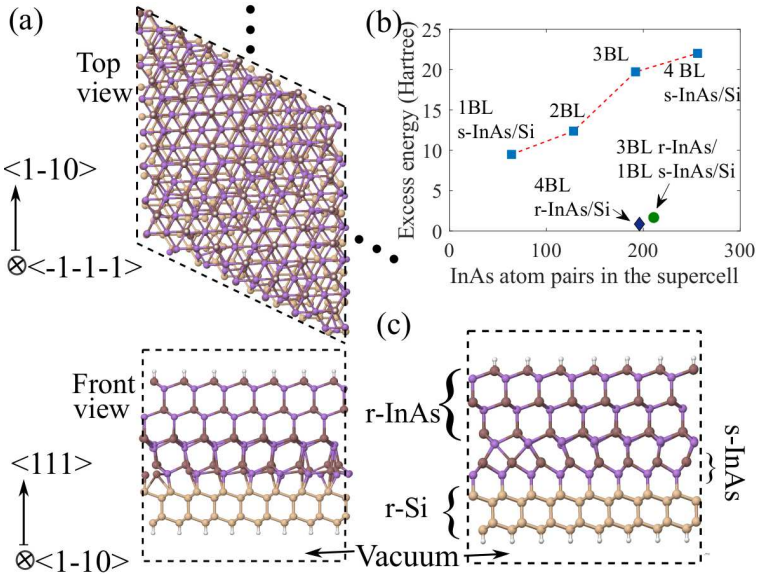


Figure 5.1: (a) Unit cell of the InAs/Si slab after structural minimization. Top view shows the hexagonal shape of the unit cell with periodic boundaries while the front view is the ⟨110⟩ view. (b) Excess energy in strained and relaxed InAs/Si slabs vs. InAs atom pairs. The number of Si atoms is kept constant among all the structures. (c) Unit cell of 3 bilayers (BL) of relaxed-InAs/1BL strained-InAs/Si slab after structural minimization (⟨110⟩ view).

5.1.2 Structural minimization of the InAs/Si interface

Due to the large lattice mismatch between InAs and Si (experimental value $\approx 11\%$), the pseudomorphic growth of lattice-matched InAs layers on Si substrate is highly unfavorable. Therefore, in the atomistic simulations, defect creation is assumed to relax the strain in the first bilayer of InAs. This defect creation distorts the otherwise uniform placement of atoms at the interface. In the first step, structural minimization is performed to find the most favorable atomic structure at the InAs/Si interface using the Mixed Gaussian-Plane-Wave (GPW) [81] based DFT simulation package CP2K [82, 83]. Perdew-Burke-Ernzerhof (PBE) pseudopotentials are used along with a double-zeta valence + one polarization (DZVP) basis set [84] for In, As, and Si atoms for modeling the structure. Structural minimization is performed on an InAs/Si slab consisting of 4 bilayers of relaxed InAs(111) stacked over 2 bilayers of relaxed Si(111) in the growth direction ($\langle 111 \rangle$). Both InAs and Si are assumed to have zinc-blende/diamond crystal structure. The unit cell of the (111) slab of zinc-blende (ZB) lattice is hexagonal with two (111) faces on top, bottom and four (110) faces on four vertical sides of the supercell. The lattice constants of InAs and Si are set to 6.247 Å and 5.478 Å, respectively, after structural minimization of the unit cell of the individual elements. Using the above lattice constants, 7 unit cells of InAs in $\langle 110 \rangle$ are nearly lattice matched to 8 unit cells of Si in $\langle 110 \rangle$ direction parallel to the interface. This lattice-matching results in 0.2% biaxial tensile strain in InAs instead of $\approx 12.5\%$ compressive strain that would be present if InAs is pseudomorphically grown on Si. Hence, a hexagonal supercell consisting of 7 and 8 unit cells of InAs and Si, respectively, along $\langle 110 \rangle$ (see Fig. 5.1) is simulated with periodic boundary conditions along the three directions. With this, the slab is padded with vacuum at the top and the bottom to ensure that periodic images of the InAs/Si slab along $\langle 111 \rangle$ do not interact. Growth of InAs on Si may take place by either an As termination at the InAs/Si interface or As incorporation in the monolayer of Si at the surface. The former mechanism leads to In-rich growth facets while the latter one leads to As-rich growth facets along $\langle 111 \rangle$. Here, the InAs(111) slab is assumed to be As-terminated at the interface. The top (111) surface of InAs and the bottom (111) surface of Si

are hydrogen-passivated to avoid surface states. Initially, the relative position of the InAs slab with respect to the Si slab is determined by shifting the InAs slab along the three axes to minimize the energy of the structure. This position is then used for structural minimization. All atoms except those in the first two bilayers of InAs adjacent to the interface are constrained during structural minimization. Since bilayers of InAs are deposited over Si during the growth, InAs atoms are not expected to significantly alter the positions of Si atoms in the substrate. The root mean square (RMS) step as well as that of the RMS gradient and the maximum gradient are chosen as convergence criteria. The convergence limits are set to their default values in CP2K.

5.1.3 Critical layer thickness of InAs on Si

In order to confirm the hypothesis that the mismatch strain is relaxed in the first bilayer of InAs, calculations are performed to find the critical height of pseudomorphic InAs on Si using the theory developed by People *et al.* [85]. Taking a slip distance (b - magnitude of Burger's vector) of 4 Å it is found that the force on the dislocation line due to misfit stress (F_{H}) of $\approx 12.5\%$ is greater than the tension in the dislocation line (F_{D}) for all values of the pseudomorphic layer thickness (h). This implies that a critical layer thickness (h_{c}) for which $F_{\mathrm{H}} = F_{\mathrm{D}}$ does not exist and a defect formation at the interface is more favorable for any h.

Additionally, structural minimization of lattice-matched strained InAs on relaxed Si substrate is performed for up to 4 bilayers of InAs. The excess energy in the strained InAs (s-InAs)/Si system calculated by subtracting the energy of relaxed InAs (r-InAs), Si atoms in bulk, and H_2 in vacuum from the energy of the strained structure is plotted in Fig. 5.1(b). The systems with s-InAs have a higher excess energy compared to the r-InAs/Si system suggesting that the latter is more favorable. Furthermore, structural minimization of the r-InAs/1BL s-InAs/Si system is performed and the excess energy of the system is plotted. The atomic structure of the optimized system is shown in Fig. 5.1(c). The excess energy of this system is larger than that of r-InAs/Si. With more bilayers of strained InAs, the excess energy further increases. This validates the hypothesis that r-InAs/Si is the most favorable structure.

5.1.4 Electronic structural calculations of InAs/Si slab

In the next step, electronic structural calculations are performed using the DFT simulation package CP2K on an InAs/Si slab to find the origin of localized states at the InAs/Si interface. Since both the Si and InAs layers in a TFET are very thick, i.e. bulk-like, an extremely thin slab such as the one used before cannot reliably represent the electronic states at the InAs/Si interface. On the other hand, DFT simulations of a bulk-like slab are computationally expensive. The InAs and Si layers in the slab are clipped to 16 bilayers and 11 bilayers, respectively. The geometrically optimized atomic structure of the InAs/Si interface obtained above is pasted at the interface of this clipped InAs/Si slab. The hexagonal supercell obtained this way is shown in Fig. 5.2(a) along with the cell boundaries. It is well known that the PBE pseudopotentials do not open a band gap in bulk InAs and underestimate the gap in Si. Therefore, the pseudopotential developed by Tao et al. (Tao-Perdew-Staroverov-Scuseria - TPSS) is used [86]. Scaling factor for exchange as well as correlation part of the TPSS functional is set to 1.9 to reproduce the experimental band gap in bulk InAs. Fermi-Dirac smearing is activated, and the temperature is set to 300 K. Initially, the energy minimization is performed with PBE pseudopotentials to obtain the total wave function of the structure. This wave function is taken as initial guess in the next step to perform the energy minimization using the TPSS technique.

5.2 Simulation Results and Discussions

5.2.1 Atomic structure at the InAs/Si interface

The converged configuration of atoms at the r-InAs/Si interface is shown in Fig. 5.1(a). The front view of the supercell in the (110) plane reveals that the periodic structure of InAs atoms at the interface remains more or less intact, except for the last unit cell in which the atoms undergo significant restructuring. From a closer look at the supercell one finds that As atoms in the first monolayer are aligned to the atoms of the adjacent Si layer at the center of the cell. This alignment gradually impairs for As atoms away from the center. At

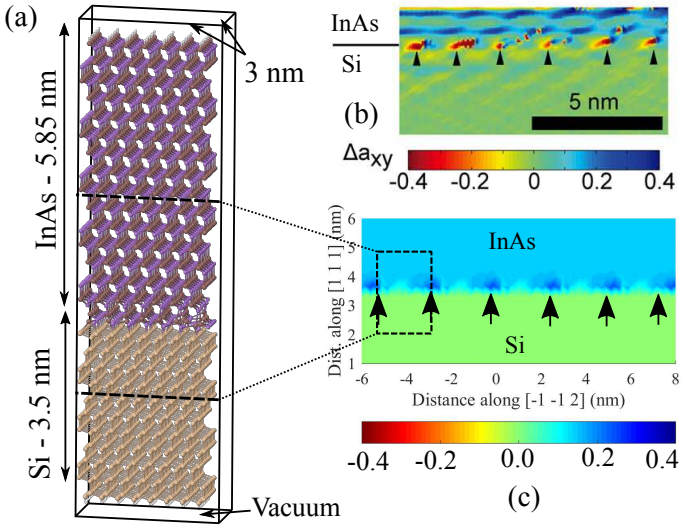


Figure 5.2: (a) Unit cell of the InAs/Si slab used for electronic structure calculations. (b) Filtered TEM image of the InAs/Si interface showing the difference between the predicted lattice constant along [0 0 1] and the one obtained from the analysis of the TEM image (taken from Ref. [65]). (c) Difference between the atomic position at the InAs/Si interface before and after geometric optimization projected along [0 0 1] direction.

the edge of the cell, the As and Si atoms are maximally misaligned, resulting in a non-uniform structure. Such a periodic disorder has been observed in the TEM images reported by Tomioka *el al.* [65]. The filtered TEM image of the difference between the predicted and the observed lattice constant in (110) at each atomic site as a fraction of the lattice constant of Si is shown in Fig. 5.2(b). For a quick comparison with this image, the difference between the atomic positions in the (110) plane before and after structural minimization of the simulated supercell is projected along the [0 0 1] direction and plotted in Fig. 5.2(c). The calculated atomic displacements show a good match with the filtered TEM image at the interface. As the atomic positions in bulk are constrained during structural minimization, the color is uniform in InAs and Si regions. The calculated positions correctly predict the periodic nature of the defects as well as the period. Thus, the derived atomic structure of the InAs/Si interface is a viable model to study the defects at the interface.

5.2.2 Atom-resolved DOS at InAs/Si interface

The projection of the wave function of each energy level on spherical harmonics at each atomic position gives the atom-resolved DOS of that energy level. The atom-resolved DOS of all the energy levels at each atom is temperature-broadened with a Gaussian broadening ($\sigma = 50\,\text{meV}$) and summed up to compute the energy-dependent DOS on that atom. The atom-resolved DOS of all atoms in each monolayer in the (111) plane is averaged and plotted as a ribbon plot in Fig. 5.3 vs. energy. The ribbons are arranged in an increasing order of their distance from the bottom of the supercell. The VB and the CB states separated by a band gap of 1.1 eV are visible in the InAs slab. Similarly, CB and VB states are separated in the Si slab by a band gap of 3.5 eV. The strong increase in the band gap compared to the experimental bulk value is due to both quantum confinement in $\langle 111 \rangle$ and the use of the semi-empirical TPSS technique which overestimates the gap of Si. Also, the structure erroneously predicts type-I band alignment at the InAs/Si interface which may be attributed to the unequal shift of the VB edge due to confinement along $\langle 111 \rangle$ in the two slabs. The CB and the VB edges show a slope along $\langle 111 \rangle$ which

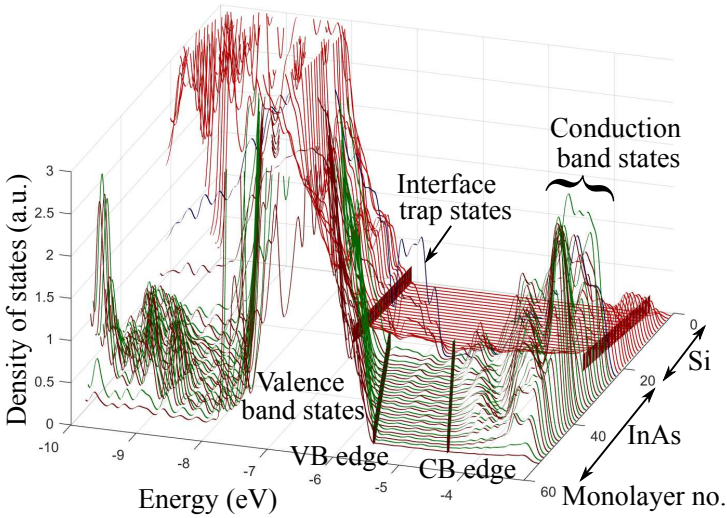


Figure 5.3: Ribbon plot depicting the DOS in each monolayer of InAs and Si in the (111) plane. The DOS on In atoms is plotted in brown, on As atoms in green, and on Si atoms in red. The DOS of the As layer at the interface is plotted in blue to highlight the traps. The band edges are marked by the dark lines running across the slab.

is due to the residual electric field induced by the localized charges at the interface.

The DOS of As atoms at the InAs/Si interface is plotted as a blue ribbon in Fig. 5.3. It exhibits an unusual shape with a large DOS in the band gap of InAs as compared to the Si or InAs ribbons in the bulk of the slab. This could be attributed to the trap energy levels at the interface. Among the adjacent monolayers of In, As, and Si at the interface, the As layer exhibits the highest DOS close to the VB edge suggesting that the trap states are primarily located on the As atomic layer. The wave function of one representative trap level is shown as a contour plot in Fig. 5.4(b). It is highly localized on the labeled As atom at the interface. Additionally, the inspection of these atoms in Fig. 5.4 reveals that both atoms have one unsaturated orbital (dangling bond). Examination of the wave functions of other trap levels at the interface leads to the conclusion that all the trap levels are highly localized on one or a few As atoms at the interface. All the As atoms hosting trap levels have at least one dangling bond. These atoms are colored in blue in Fig. 5.4(a) along with their labels. The above analysis suggests that the InAs/Si interface trap levels originate from dangling bonds present on As atoms at the interface [87].

Band edge diagram along the growth direction is plotted in Fig. 5.3 with the CB and the VB edges represented as dark lines. The band diagram shows downward (towards negative energy) band bending at InAs/Si interface. On the contrary, band edge diagram plotted in Fig. 3.9(a) shows upward band bending giving rise to a triangular peak at the hetero-interface. Reason for this discrepancy is as follows. The DFT simulations described above are equilibrium simulations. Hence electrons/holes in the heterostructure will not be swept away by the source-drain bias as in the TCAD simulations. On the contrary, TCAD simulations are non-equilibrium simulations. This difference between equilibrium conditions in DFT and steady state non-equilibrium conditions in TCAD is responsible for the observed discrepancy.

5.2.3 Passivation of the interface traps by S and H

Interface traps must be passivated to minimize the trap-induced degradation of a hetero-junction device such as a TFET. Since the interface traps originate from the unsaturated orbitals present on As

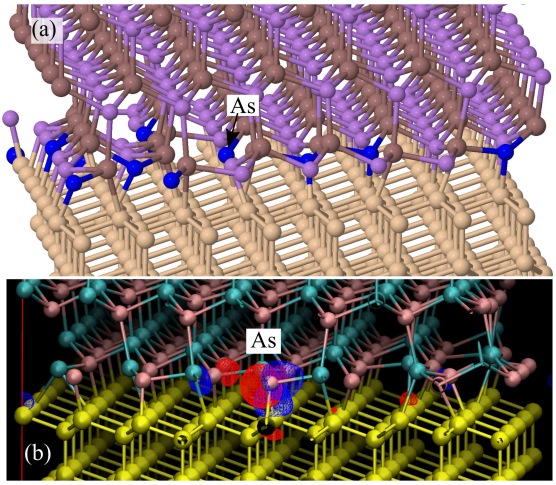


Figure 5.4: (a) Atoms which contribute most to the trap levels are highlighted. All the As atoms hosting the trap levels have at least one dangling bond. (b) Contour plot of the wave function of the representative trap level at E = -6.253 eV.

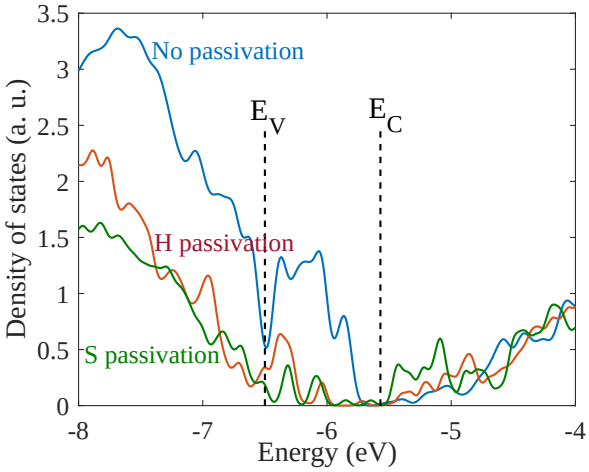


Figure 5.5: Comparison of the DOS at an As monolayer at the interface with and without passivation.

atoms, Hydrogen or Sulfur atoms may be used to saturate the dangling bonds which will reduce the D_{it}. The above developed atomistic model of the InAs/Si interface is used to study the effectiveness of the two passivating materials. To study H passivation, Hydrogen atoms are attached to the interface As atoms having unsaturated orbitals. Their exact position is determined by performing structure minimization of the InAs/Si slab by constraining all the atoms except the added H atoms. Electronic structural calculations of the converged structure are performed to obtain the ribbon plot (not shown) similar to that of Fig. 5.3. The atom-resolved DOS of the As interface atoms is found to be reduced after the passivation implying the effectiveness of Hydrogen passivation. The same study is performed by passivating the unsaturated As interface atoms by Sulfur, except that the first two bi-layers of InAs are included in the structural minimization for determining the optimum position of the passivating S atoms. This is necessary since S atoms, due to a larger covalent radius, are expected to alter the atomic configuration in their vicinity.

The DOS at As interface atoms with and without passivation are compared in Fig. 5.5. Passivation by both H and S significantly reduces the D_{it} at the hetero-interface. The VB DOS at As interface atoms is large for the unpassivated interface compared to the passivated one. As a result, the trap states capture/emit an electron and acquire a charge. This attracts electronic states in the VB towards the interface which, in turn, increases the atom-resolved DOS in the VB as seen in Fig. 5.5. When the traps are passivated, the local charge density at the interface is reduced which decreases the DOS at As interface atoms. The areal density of H or S atoms required for passivating the traps is close to $10^{14}\,\mathrm{atoms/cm^2}$. Introducing such a large number of H or S atoms at the interface *after* the growth is technologically challenging.

5.2.4 Limitations of the analysis

There are shortcomings and assumptions made in the above analysis. The lattice constants of Si and InAs differ from their measured values. The use of experimental numbers would result in the relaxation of the interface during the structural minimization by high hydrostatic strain, not by the misalignment. To avoid it, the lattice constants of InAs and Si are relaxed beforehand. The bulk of InAs and Si slabs is assumed to be devoid of any defect during structural minimization of the interface. The growth of a polar semiconductor such as InAs on a nonpolar substrate such as Si may result in the formation of anti-phase domains. By tuning the growth conditions, an InAs nanowire can be grown from a single nucleation, and anti-phase domains can be avoided.

Chapter 6

Conclusion

Since non-idealities play an important role in degradation of the subthreshold swing and other performance metrics of the TFETs, physics-based TCAD models are developed to account for various non-idealities such as channel quantization, surface roughness, and DOS tails. Using these newly developed models as well as existing TCAD models for TAT, a good reproduction the experimental transfer characteristics of InAs/Si and all-III-V TFETs is achieved. A summary of the results presented in this thesis is provided below.

6.1 Summary

Physics-based TCAD models, which accurately account for the influence of various non-idealities on TFETs, are described in Chapter 2. Three different models for channel quantization are developed and implemented in the TCAD simulator using the PMI for non-local recombination. The band-edge correction method, despite being computationally the most efficient one, severely alters the electrostatics in the channel of the device, thus may not be suitable for some TFETs. The other two models based on the path rejection method do not alter the electrostatics, and are, therefore, recommended. The latter model which is based on the transfer matrix method is applicable also for any arbitrary band edge profile adjacent to the channel, hence suitable

for any kind of TFET, in particular for hetero-junction core-shell NW TFETs. The simulations using all the above models confirm that, channel quantization delays the onset of vertical tunneling and reduces the strength of tunneling which results in the reduction of the on-current.

To capture the effect of DOS tails caused by oxide/semiconductor interface roughness on vertical tunneling, an analytical model for the DOS of a 2D quantum-mechanical system in the presence of an arbitrary random field was adapted. The resulting model is implemented in the TCAD simulator using the nonlocal PMI. The simulations of a vertical TFET using realistic parameters for roughness amplitude and auto-correlation length have shown that the subthreshold swing degrades due to interface roughness.

TCAD model to account for the effect of DOS tails is implemented in S-Device and the impact of band tails on a NW Gate-All-Around TFET is studied. It is observed that, as a consequence of the smoothed band edge, the energetic overlap of initial and final states occurs gradually at the onset of tunneling. This increases the subthreshold swing of the TFET. Exponential tails have a much stronger impact here than Gaussian tails. The localization parameter m_{t} hardly changes the swing, but determines the magnitude of the generation rate. The developed model is a "continuum model", i.e. it implicitly assumes a proper average over random disorder caused by doping. Aggressive geometrical scaling of NW TFETs leads to a countable number of doping atoms in the source region. Then, the occurrence of tails becomes questionable and only an atomistic approach, like full-band tight-binding NEGF, is able to correctly simulate the effect.

In Chapter 3, InAs/Si lateral as well as vertical NW GOS TFETs are analyzed taking into account not only BTBT, but also TAT processes at the InAs/Si and InAs/oxide interfaces. It is confirmed that the TAT current begins at a lower gate bias compared to BTBT in both vertical as well as lateral TFETs. Therefore, the interface traps determine the TFET characteristics in the subthreshold regime. In the vertical NW TFETs, it is observed that channel quantization is responsible for the degradation of the on-current while traps at the oxide/InAs and InAs/Si hetero-interface are responsible for the degradation of the subthreshold swing. In lateral TFETs, which are found to exhibit a high trap concentration at the InAs/Si interface, current conduction takes

place in two steps. Electron-hole pairs are generated at this interface by a TAT process which is followed by thermionic emission of holes to the drain. In the subthreshold regime, due to the presence of a high thermionic barrier, thermionic emission of holes is the rate-determining step. At higher gate bias, when the thermionic barrier is sufficiently lowered, TAT becomes the rate-determining step. In addition to these mechanisms, traps at the Si/oxide interface further degrade the subthreshold swing by reducing the coupling between the gate and the Si channel. Predictive simulations with different D_{it} values have shown that $D_{\mathrm{it}} = 5 \times 10^{11}\mathrm{cm}^{-2}\mathrm{eV}^{-1}$ is the maximum allowable value that still would result in a sub-thermal swing.

In Chapter 4, measured InAs/InGaAsSb NW TFETs with high on-current and sub-thermal subthreshold swing are studied by calibrated TCAD to understand how different non-idealities such as bulk traps, interface traps, and band tails affect their performance. The position-dependent composition of the quaternary alloy $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{As}_y\mathrm{Sb}_{1-y}$ extracted from TEM and other physical characterizations was used to determine the values of band gap, band alignment, and effective masses, thus avoiding a fit of these parameters. A good match between simulated and measured IV-characteristics at different temperatures could be obtained which validates the approach. The analysis showed that bulk traps present adjacent to the hetero-junction are primarily responsible for the degradation of the swing. Band tails with reasonable parameters are less likely responsible for the increase of the swing.

6.2 Outlook

6.2.1 Selection between line and point tunneling

If the band alignment at the hetero-interface is of type-I (for example $\mathrm{In}_{0.53}\mathrm{Ga}_{0.47}\mathrm{As}/\mathrm{InP}$), then line tunneling is more efficient than point tunneling due to smaller effective tunnel gap in the case of line tunneling compared to that of point tunneling. Also, the gate field is aligned with the tunnel direction in the case of a vertical TFET, thus further enhancing it. On the other hand, if band alignment is of type-II, then point tunneling is more efficient and yields steeper characteristics due to a smaller effective tunnel gap for point tunneling compared to that

of line tunneling. In the case of type-I heterojunction TFETs, the counter-doped pocket needs to be introduced under the gate in order to advance the onset of line tunneling ahead of point tunneling.

Full-band quantum transport simulations of InAs/Si (type-II heterojunction) TFETs presented in [37] showed that for ultra-thin body (UTB) bilayer TFETs with the body thickness < 10 nm, point tunneling begins before line tunneling which yields a better swing compared to line tunneling. Similar results are obtained when the same device is simulated using the channel quantization model based on transfer matrix method. For bulk-like TFETs, the onset of line tunneling can be advanced by introducing a counter-doped pocket under the gate which improves the swing. In this case, line tunneling can yield a higher on-current and steeper onset.

The above findings are obtained without taking into account TAT taking place at the oxide/semiconductor interface. Line tunneling occurs normal to the gate only when a triangular well is formed adjacent to the gate. The formation of the triangular well would inevitably result in TAT at the oxide/semiconductor interface which would degrade subthreshold characteristics. Thus, TFETs in which line tunneling is the dominant mechanism may be preferred only when the oxide/semiconductor interface is trap-free. Otherwise, point tunneling must be preferred in TFETs.

6.2.2 Trap-tolerant device design

Although simulations of ideal TFETs suggest that they can deliver the sub-thermal swing, the presence of various non-idealities such as oxide interface traps, hetero-interface traps, band tails, surface roughness etc. degrade the subthreshold characteristics. Interface traps have the strongest impact. Predictive simulations suggest that an oxide- as well as a hetero-interface trap density below $5 \times 10^{11}\,\mathrm{cm}^{-2}\mathrm{eV}^{-1}$ are necessary to achieve sub-thermal swing if gate-overlapped-source TFETs with thicker nanowires are used. It is challenging to achieve such a low D_{it} throughout the band gap of III-V semiconductors at the oxide/III-V interface. Therefore, it is necessary to search for a TFET geometry which is least sensitive to TAT.

In the case of an InAs/Si NW TFET, diameter scaling down to 20 nm would minimize TAT at the oxide-interface and gate alignment

would reduce the leakage basin caused by hetero-interface TAT. Note that, this approach may also be applicable in TFETs with other material systems.

In this thesis, this was demonstrated for the fabricated InAs/InGaAsSb TFETs. One may infer from the temperature independence of the subthreshold swing of these TFETs, that TAT at the InAs/oxide as well as the InGaAsSb/oxide interfaces is reduced. This is possible due to the narrow NW dimensions. Simulations performed using the developed simulation set-up suggest that an improvement in both on-current and subthreshold swing may be achieved by aligning the gate with the InAs/InGaAsSb hetero-junction. However, a misalignment of only 5 nm will completely reverse this advance.

Appendix A

Notation and Acronyms

Acronyms

2DEG	Two-dimensional Electron Gas
BGN	Band Gap Narrowing
BTBT	Band-to-band Tunneling
CB	Conduction Band
CMOS	Complementary Metal-Oxide Semiconductor
MOSFET	Metal-Oxide Semiconductor Field Effect Transistor
DFT	Density Functional Theory
WKB	Wentzel-Kramers-Brillouin
DNLP	Dynamic Non-local Path
DOS	Density of States
FWHM	Full Width Half Maximum

GA	Gate-aligned
GOS	Gate-overlapped Source
LH	Light Hole
NW	nano-wire
PMI	Physical Model Interface
SRH	Schockley-Read-Hall
TAT	Trap-assisted Tunneling
TCAD	Technology Computer Aided Design
TEM	Transmission Electron Microscopy
TFET	Tunnel Field Effect Transistor
VB	Valence Band
WZ	Wurtzite
ZB	Zinc-blende

Bibliography

[1] ITR Committee, "International Technology Roadmap for Semiconductors 2.0," Tech. Rep., 01 2015.

[2] R. H. Dennard, F. H. Gaensslen, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE Journal of Solid-State Circuits*, vol. 9, no. 5, pp. 256–268, Oct 1974.

[3] A. M. Ionescu and H. Riel, "Tunnel field-effect transistors as energy-efficient electronic switches," *Nature*, vol. 479, p. 329, 2011.

[4] A. C. Seabaugh and Q. Zhang, "Low-Voltage tunnel transistors for beyond CMOS Logic," *Proc. IEEE*, vol. 98, p. 2095, 2010.

[5] H. Lu and A. C. Seabaugh, "Tunnel Field-Effect Transistors: State-of-the-Art," *IEEE J. Electron Devices Soc.*, vol. 2, p. 44, 2014.

[6] M. T. Björk, O. Hayden, H. Schmid, H. Riel, and W. Riess, "Vertical surround-gated silicon nanowire impact ionization field-effect transistors," *Applied Physics Letters*, vol. 90, no. 14, p. 142110, 2007.

[7] O. Y. Loh and H. D. Espinosa, "Nanoelectromechanical contact switches," *Nature Nanotechnology*, vol. 7, p. 283, 2012.

[8] G. Zhou, Y. Lu, R. Li, Q. Zhang, Q. Liu, T. Vasen, H. Zhu, J. M. Kuo, T. Kosel, M. Wistey *et al.*, "InGaAs/InP Tunnel FETs With

a Subthreshold Swing of 93 mV/dec and I_{ON}/I_{OFF} Ratio Near 10^6," *IEEE Electron Device Letters*, vol. 33, no. 6, pp. 782–784, June 2012.

[9] R. Asra, M. Shrivastava, K. V. R. M. Murali, R. K. Pandey, H. Gossner, and V. R. Rao, "A Tunnel FET for V_{DD} Scaling Below 0.6 V With a CMOS-Comparable Performance," *IEEE Transactions on Electron Devices*, vol. 58, no. 7, pp. 1855–1863, July 2011.

[10] G. Dewey, B. Chu-Kung, J. Boardman, J. M. Fastenau, J. Kavalieros, R. Kotlyar, W. K. Liu, D. Lubyshev, M. Metz, N. Mukherjee *et al.*, "Fabrication, characterization, and physics of III-V heterojunction tunneling Field Effect Transistors (H-TFET) for steep sub-threshold swing," in *2011 International Electron Devices Meeting*, Dec 2011, pp. 33.6.1–33.6.4.

[11] 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 Letters*, vol. 31, no. 6, pp. 564–566, June 2010.

[12] A. M. Walke, A. Vandooren, R. Rooyackers, D. Leonelli, A. Hikavyy, R. Loo, A. S. Verhulst, K. H. Kao, C. Huyghebaert, G. Groeseneken *et al.*, "Fabrication and Analysis of a $Si/Si_{0.55}Ge_{0.45}$ Heterojunction Line Tunnel FET," *IEEE Transactions on Electron Devices*, vol. 61, no. 3, pp. 707–715, March 2014.

[13] H. Riel, K. E. Moselund, C. Bessire, M. T. Björk, A. Schenk, H. Ghoneim, and H. Schmid, "InAs-Si heterojunction nanowire tunnel diodes and tunnel FETs," in *2012 International Electron Devices Meeting*, Dec 2012, pp. 16.6.1–16.6.4.

[14] S. Richter, S. Blaeser, L. Knoll, S. Trellenkamp, A. Schäfer, J. M. Hartmann, Q. T. Zhao, and S. Mantl, "SiGe on SOI nanowire array TFETs with homo- and heterostructure tunnel junctions," in *2013 14th International Conference on Ultimate Integration on Silicon (ULIS)*, March 2013, pp. 25–28.

[15] A. Schenk, R. Rhyner, M. Luisier, and C. Bessire, "Simulation study of nanowire tunnel FETs," in *70th Device Research Conference*, June 2012, pp. 201–202.

[16] M. Luisier and G. Klimeck, "Simulation of nanowire tunneling transistors: From the Wentzel-Kramers-Brillouin approximation to full-band phonon-assisted tunneling," *Journal of Applied Physics*, vol. 107, no. 8, p. 084507, 2010.

[17] W. G. Vandenberghe, B. Soree, W. Magnus, G. Groeseneken, and M. V. Fischetti, "Impact of field-induced quantum confinement in tunneling field-effect devices," *Applied Physics Letters*, vol. 98, no. 14, p. 143503, 2011.

[18] S. Sant and A. Schenk, "Modeling the Effect of Interface Roughness on the Performance of Tunnel FETs," *IEEE Electron Device Letters*, vol. 38, no. 2, pp. 258–261, Feb 2017.

[19] A. Schenk, S. Sant, K. Moselund, and H. Riel, "III-V-based hetero tunnel FETs: A simulation study with focus on non-ideality effects," in *2016 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS)*, Jan 2016, pp. 9–12.

[20] S. Madisetti, V. Tokranov, A. Greene, M. Yakimov, M. Hirayama, S. Oktyabrsky, S. Bentley, and A. P. Jacob, "Growth of strained InGaSb quantum wells for p-FET on Si: Defects, interfaces, and electrical properties," *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena*, vol. 32, no. 5, p. 051206, 2014.

[21] K. Tomioka, Y. Kobayashi, J. Motohisa, S. Hara, and T. Fukui, "Selective-area growth of vertically aligned GaAs and GaAs/AlGaAs core-shell nanowires on Si(111) substrate," *Nanotechnology*, vol. 20, no. 14, p. 145302, 2009.

[22] C. H. Wang, G. Doornbos, G. Astromskas, G. Vellianitis, R. Oxland, M. C. Holland, M. L. Huang, C. H. Lin, C. H. Hsieh, Y. S. Chang *et al.*, "High-k dielectrics on (100) and (110) n-InAs: Physical and electrical characterizations," *AIP Advances*, vol. 4, no. 4, p. 047108, 2014.

[23] J. Wu, A. S. Babadi, D. Jacobsson, J. Colvin, S. Yngman, R. Timm, E. Lind, and L.-E. Wernersson, "Low Trap Density in InAs/High-k Nanowire Gate Stacks with Optimized Growth and Doping Conditions," *Nano Letters*, vol. 16, no. 4, pp. 2418–2425, 2016, pMID: 26978479.

[24] S. Lee, V. Chobpattana, C.-Y. Huang, B. J. Thibeault, W. Mitchell, S. Stemmer, A. C. Gossard, and M. J. W. Rodwell, "Record 5 nm gate length ZrO_2/InAs/InAlAs MOSFETs," in *IEEE Symposium on VLSI Technology Digest of Technical Papers*, June 2014, pp. 1–2.

[25] E. O. Kane, "Theory of tunneling," *Journal of Applied Physics*, vol. 32, no. 1, pp. 83–91, 1961.

[26] A. Schenk, "Rigorous theory and simplified model of the band-to-band tunneling in silicon," *Solid-State Electronics*, vol. 36, no. 1, pp. 19 – 34, 1993. [Online]. Available: http://www.sciencedirect.com/science/article/pii/003811019390065X

[27] D. R. Fredkin and G. H. Wannier, "Theory of electron tunneling in semiconductor junctions," *Phys. Rev.*, vol. 128, pp. 2054–2061, Dec 1962. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.128.2054

[28] E. O. Kane, "Zener tunneling in semiconductors," *Journal of Physics and Chemistry of Solids*, vol. 12, no. 2, pp. 181 – 188, 1960.

[29] H. Carrillo-Nunez, A. Ziegler, M. Luisier, and A. Schenk, "Modeling direct band-to-band tunneling: From bulk to quantum-confined semiconductor devices," *Journal of Applied Physics*, vol. 117, no. 23, p. 234501, 2015.

[30] 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*, vol. 74, p. 205323, Nov 2006. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.74.205323

[31] H. Carrillo-Nunez, M. Luisier, and A. Schenk, "Analysis of inas-si heterojunction nanowire tunnel fets: Extreme confinement vs. bulk," *Solid-State Electronics*, vol. 113, no. Supplement C, pp. 61 – 67, 2015, selected papers from ESSDERC 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0038110115001513

[32] W. G. Vandenberghe, B. Soree, W. Magnus, G. Groeseneken, A. S. Verhulst, and M. V. Fischetti, "Field induced quantum confinement in indirect semiconductors: Quantum mechanical and modified semiclassical model," in *2011 International Conference on Simulation of Semiconductor Processes and Devices*, Sept 2011, pp. 271–274.

[33] *Sentaurus Device User Guide Version 2015.06*, Synopsys Inc. Mountain View, CA, USA, June 2015.

[34] S. Sant and A. Schenk, "Methods to Enhance the Performance of InGaAs/InP Heterojunction Tunnel FETs," *IEEE Transactions on Electron Devices*, vol. 63, no. 5, pp. 2169–2175, May 2016.

[35] S. Sant, H. Carrillo-Nunez, M. Luisier, and A. Schenk, "Transfer matrix based semiclassical model for field-induced and geometrical quantum confinement in tunnel FETs," in *2016 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, Sept 2016, pp. 77–80.

[36] H. Flietner, "The E(k) Relation for a Two-Band Scheme of Semiconductors and the Application to the Metal-Semiconductor Contact," *physica status solidi (b)*, vol. 54, no. 1, pp. 201–208, 1972. [Online]. Available: http://dx.doi.org/10.1002/pssb.2220540119

[37] H. Carrillo-Nunez, M. Luisier, and A. Schenk, "Analysis of InAs-Si heterojunction double-gate tunnel FETs with vertical tunneling paths," in *2015 45th European Solid State Device Research Conference (ESSDERC)*, Sept 2015, pp. 302–305.

[38] D. N. Quang and N. H. Tung, "A semiclassical approach to the electron gas in two-dimensional semiconductor structures," *physica status solidi (b)*, vol. 207, no. 1, pp. 111–123,

1998. [Online]. Available: http://dx.doi.org/10.1002/(SICI)1521-3951(199805)207:1⟨111::AID-PSSB111⟩3.0.CO;2-P

[39] T. Ando, "Screening effect and quantum transport in a Silicon inversion layer in strong magnetic fields," *Journal of the Physical Society of Japan*, vol. 43, no. 5, pp. 1616–1626, 1977.

[40] *Sentaurus Device Monte Carlo User Guide Version 2015.06*, Synopsys Inc. Mountain View, CA, USA, June 2015.

[41] 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, Apr 1994. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.72.2749

[42] M. Naidenkova, M. S. Goorsky, R. Sandhu, R. Hsing, M. Wojtowicz, T. P. Chin, T. R. Block, and D. C. Streit, "Interfacial roughness and carrier scattering due to misfit dislocations in $In_{0.52}Al_{0.48}As/In_{0.75}Ga_{0.25}As/InP$ structures," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, vol. 20, no. 3, pp. 1205–1208, 2002.

[43] E. O. Kane, "Thomas-Fermi approach to impure semiconductor band structure," *Phys. Rev.*, vol. 131, pp. 79–88, Jul 1963. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.131.79

[44] B. I. Shklovsky and A. L. Efros, "Density-of-states tails in heavily doped semiconductors," *Sov. Phys. Semiconductors*, vol. 4, no. 2, pp. 249–257, 1970.

[45] B. I. Halperin and M. Lax, "Impurity-band tails in the high-density limit. i. minimum counting methods," *Phys. Rev.*, vol. 148, pp. 722–740, Aug 1966. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.148.722

[46] G. Lucovsky, "On the photoionization of deep impurity centers in semiconductors," *Solid State Communications*, vol. 3, no. 9, pp. 299 – 302, 1965. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0038109865900396

[47] A. Schenk, R. Enderlein, and D. Suisky, "Electroabsorption for deep level to band transitions accompanied by multiphonon processes," *physica status solidi (b)*, vol. 131, no. 2, pp. 729–739, 1985. [Online]. Available: http://dx.doi.org/10.1002/pssb.2221310234

[48] S. Sant and A. Schenk, "The effect of density-of-state tails on band-to-band tunneling: Theory and application to tunnel field effect transistors," *Journal of Applied Physics*, vol. 122, no. 13, p. 135702, 2017. [Online]. Available: https://doi.org/10.1063/1.4994112

[49] N. Mott, "The mobility edge since 1967," *Journal of Physics C: Solid State Physics*, vol. 20, no. 21, p. 3075, 1987. [Online]. Available: http://stacks.iop.org/0022-3719/20/i=21/a=008

[50] E. O. Kane, "Band tails in semiconductors," *Solid-State Electronics*, vol. 28, no. 1, pp. 3 – 10, 1985. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0038110185902035

[51] A. Schenk, M. Stahl, and H.-J. Wünsche, "Calculation of interband tunneling in inhomogeneous fields," *physica status solidi (b)*, vol. 154, no. 2, pp. 815–826, 1989. [Online]. Available: http://dx.doi.org/10.1002/pssb.2221540240

[52] E. A. Vinogradov, L. K. Vodopyanov, and G. S. Oleinik, *English translation in - Sov. Phys. Solid State*, vol. 15, no. 2, p. 322, 1973.

[53] *New Semiconductor Materials (NSM) Archives*, St. Petersburg, Russian Federation, Jan. 2001.

[54] M. Luisier and G. Klimeck, "Simulation of nanowire tunneling transistors: From the Wentzel-Kramers-Brillouin approximation to full-band phonon-assisted tunneling," *Journal of Applied Physics*, vol. 107, no. 8, p. 084507, 2010.

[55] S. M. Sze and K. Ng, *Physics of Semiconductor Devices*. Wiley, New York, NY, USA, 2007.

[56] A. Palma, A. Godoy, J. A. Jiménez-Tejada, J. E. Carceller, and J. A. López-Villanueva, "Quantum two-dimensional calculation of time constants of random telegraph signals in metal-oxide/semiconductor structures," *Phys. Rev. B*, vol. 56, pp. 9565–9574, Oct 1997. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.56.9565

[57] K. E. Moselund, D. Cutaia, H. Schmid, M. Borg, S. Sant, A. Schenk, and H. Riel, "Lateral InAs/Si p-Type Tunnel FETs Integrated on Si–Part 1: Experimental Devices," *IEEE Transactions on Electron Devices*, vol. 63, no. 11, pp. 4233–4239, Nov 2016.

[58] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, "Band parameters for III-V compound semiconductors and their alloys," *Journal of Applied Physics*, vol. 89, no. 11, pp. 5815–5875, 2001.

[59] Y. Varshni, "Temperature dependence of the energy gap in semiconductors," *Physica*, vol. 34, no. 1, pp. 149 – 154, 1967. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0031891467900626

[60] C. D. Bessire, M. T. Björk, H. Schmid, A. Schenk, K. B. Reuter, and H. Riel, "Trap-Assisted Tunneling in Si-InAs Nanowire Heterojunction Tunnel Diodes," *Nano Letters*, vol. 11, no. 10, pp. 4195–4199, 2011, pMID: 21875101.

[61] C. D. Bessire, *Semiconducting nanowire tunnel devices: From all-Si tunnel diodes to III-V heterostructure tunnel FETs (Ph.D. Thesis)*. Hartung-Gorre Verlag Konstanz, 2014.

[62] D. Cutaia, K. E. Moselund, M. Borg, H. Schmid, L. Gignac, C. M. Breslin, S. Karg, E. Uccelli, and H. Riel, "Vertical InAs-Si Gate-All-Around Tunnel FETs Integrated on Si Using Selective Epitaxy in Nanotube Templates," *IEEE Journal of the Electron Devices Society*, vol. 3, no. 3, pp. 176–183, May 2015.

[63] M. Mohammed, A. S. Verhulst, D. Verreck, M. V. de Put, E. Simoen, B. Soree, B. Kaczer, R. Degraeve, A. Mocuta, N. Collaert *et al.*, "Electric-field induced quantum broadening

of the characteristic energy level of traps in semiconductors and oxides," *Journal of Applied Physics*, vol. 120, no. 24, p. 245704, 2016.

[64] A. Schenk, K. Irmscher, D. Suisky, R. Enderlein, F. Bechstedt, and H. Klose, *Electric Field Effect on Multiphonon Transitions at Deep Centres*. New York, NY: Springer New York, 1985, pp. 613–616. [Online]. Available: https://doi.org/10.1007/978-1-4615-7682-2_135

[65] K. Tomioka, M. Yoshimura, and T. Fukui, "Sub 60 mV/decade Switch Using an InAs Nanowire-Si Heterojunction and Turn-on Voltage Shift with a Pulsed Doping Technique," *Nano Letters*, vol. 13, no. 12, pp. 5822–5826, 2013, pMID: 24215512.

[66] E. Memisevic, M. Hellenbrand, E. Lind, A. R. Persson, S. Sant, A. Schenk, J. Svensson, R. Wallenberg, and L.-E. Wernersson, "Individual Defects in InAs/InGaAsSb/GaSb Nanowire Tunnel Field-Effect Transistors Operating below 60 mV/decade," *Nano Letters*, vol. 17, no. 7, pp. 4373–4380, 2017, pMID: 28613894.

[67] E. Memisevic, J. Svensson, M. Hellenbrand, E. Lind, and L. E. Wernersson, "Vertical InAs/GaAsSb/GaSb tunneling field-effect transistor on Si with S = 48 mV/decade and I_{on} = 10 A/m for I_{off} = 1 nA/μm at V_{ds} = 0.3 V," in *2016 IEEE International Electron Devices Meeting (IEDM)*, Dec 2016, pp. 19.1.1–19.1.4.

[68] S. Adachi, "Band gaps and refractive indices of AlGaAsSb, GaInAsSb, and InPAsSb: Key properties for a variety of the 2-4 mm optoelectronic device applications," *J. Appl. Phys.*, vol. 61, p. 4869, 1987.

[69] C. G. Van de Walle, "Band lineups and deformation potentials in the model-solid theory," *Phys. Rev. B*, vol. 39, pp. 1871–1883, Jan 1989. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.39.1871

[70] H. A. Nilsson, P. Caroff, C. Thelander, E. Lind, O. Karlström, and L.-E. Wernersson, "Temperature dependent properties of InSb and InAs nanowire field-effect transistors," *Applied Physics Letters*, vol. 96, no. 15, p. 153505, 2010.

[71] E. Memisevic, J. Svensson, E. Lind, and L. E. Wernersson, "InAs/InGaAsSb/GaSb Nanowire Tunnel Field-Effect Transistors," *IEEE Transactions on Electron Devices*, vol. 64, no. 11, pp. 4746–4751, Nov 2017.

[72] S. W. Kurnick and J. M. Powell, "Optical Absorption in Pure Single Crystal InSb at 298K and 78K," *Phys. Rev.*, vol. 116, pp. 597–604, Nov 1959. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.116.597

[73] J. I. Pankove, "Absorption edge of impure Gallium Arsenide," *Phys. Rev.*, vol. 140, pp. A2059–A2065, Dec 1965. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.140.A2059

[74] A. Schenk, S. Sant, K. Moselund, and H. Riel, "Comparative Simulation Study of InAs/Si and All-III-V Hetero Tunnel FETs," *Transactions of Electrochemical Society*, vol. 66, pp. 157–169, 2015.

[75] A. R. Trivedi, K. Z. Ahmed, and S. Mukhopadhyay, "Negative gate transconductance in gate/source overlapped heterojunction tunnel fet and application to single transistor phase encoder," *IEEE Electron Device Letters*, vol. 36, no. 2, pp. 201–203, Feb 2015.

[76] M. Heyns and W. Tsai, "Ultimate Scaling of CMOS Logic Devices with Ge and III-V Materials," *MRS Bulletin*, vol. 34, no. 7, pp. 485–492, 2009.

[77] S. Yadav, K. H. Tan, A. Kumar, K. H. Goh, G. Liang, S. F. Yoon, X. Gong, and Y. C. Yeo, "Monolithic Integration of InAs Quantum-Well n-MOSFETs and Ultrathin Body Ge p-MOSFETs on a Si Substrate," *IEEE Transactions on Electron Devices*, vol. 64, no. 2, pp. 353–360, Feb 2017.

[78] K. Tomioka, K. Ikejiri, T. Tanaka, J. Motohisa, S. Hara, K. Hiruma, and T. Fukui, "Selective-area growth of III-V nanowires and their applications," *Journal of Materials Research*, vol. 26, no. 17, pp. 2127–2141, 2011.

[79] M. Horn-von Hoegen, *Surface and Interface Science: Chapter - Surfaces of Elemental Semiconductors*. Wiley-VCH Verlag GmbH and Co., 2013.

[80] G. S. Higashi, Y. J. Chabal, G. W. Trucks, and K. Raghavachari, "Ideal hydrogen termination of the Si (111) surface," *Applied Physics Letters*, vol. 56, no. 7, pp. 656–658, 1990.

[81] B. G. Lippert, J. Hutter, and M. Parrinello, "A hybrid Gaussian and plane wave density functional scheme," *Molecular Physics*, vol. 92, no. 3, pp. 477–488, 1997.

[82] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, "cp2k: atomistic simulations of condensed matter systems," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 4, no. 1, pp. 15–25, 2014. [Online]. Available: http://dx.doi.org/10.1002/wcms.1159

[83] *CP2K version 5.0 (Development Version)*, The CP2K developers group, May 2016.

[84] S. Goedecker, M. Teter, and J. Hutter, "Separable dual-space gaussian pseudopotentials," *Phys. Rev. B*, vol. 54, pp. 1703–1710, Jul 1996. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.54.1703

[85] R. People and J. C. Bean, "Calculation of critical layer thickness versus lattice mismatch for Ge_xSi_{1-x}/Si strained-layer heterostructures," *Applied Physics Letters*, vol. 47, no. 3, pp. 322–324, 1985.

[86] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, "Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids," *Phys. Rev. Lett.*, vol. 91, p. 146401, Sep 2003. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.91.146401

[87] S. Sant, M. Luisier, and A. Schenk, "DFT-based analysis of the origin of traps at the InAs/Si (111) interface," in *2017 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, Sept 2017, pp. 17–20.

Curriculum Vitae

Saurabh Sant
Birth: 11$^{\text{th}}$ November 1989 in Mumbai, India
Citizenship: India
Languages: English (fluent), German (A1), Marathi (native), Hindi (fluent)

Education:

- February 2014 – April 2018
  Ph. D. in Electrical Engineering at ETH Zürich
  under the supervision of Prof. Andreas Schenk

- September 2012 – December 2013
  Master of Science in Electrical Engineering at ETH Zürich
  Masters Thesis under the supervision of Dr. Heike Riel and Prof. Andreas Schenk on "Template-based fabrication of InGaAs nanowires for TFET devices"

- August 2007 – May 2011
  Bachelor of Technology in Electrical Engineering at IIT Bombay, Mumbai, India.

Work experience:

- June 2017 – August 2017
  Internship on "TCAD Modeling of Photonic and Plasmonic LASERs-On-Chip"
  IBM Research GmbH, Rüschlikon

- July 2011 – August 2012
  Research Assistant on "Mobility Modeling of Unstrained and Strained Ge Channel pMOSFET" Department of Electrical Engineering, IIT Bombay, Mumbai, India.