Modeling and optimization of Tunnel-FET architectures exploiting carrier gas dimensionality

THÈSE N^{O} 7587 (2017)

PRÉSENTÉE LE 24 MARS 2017
À LA FACULTÉ DES SCIENCES ET TECHNIQUES DE L'INGÉNIEUR
LABORATOIRE DES DISPOSITIFS NANOÉLECTRONIQUES
PROGRAMME DOCTORAL EN MICROSYSTÈMES ET MICROÉLECTRONIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Cem ALPER

acceptée sur proposition du jury:

Dr G. Boero, président du jury
Prof. M. A. Ionescu, Prof. P. Palestri, directeurs de thèse
Dr U. E. Avci, rapporteur
Prof. M. Luisier, rapporteur
Dr J.-M. Sallese, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE
Suisse
2017

Acknowledgements

This thesis marks the end my PhD journey which took a bit more than four years but felt like eternity. Indeed, working on the same subject for such a long time sometimes got a bit monotonous and frustrating. What keeps you going is the moments when you overcome tough intellectual challenges that makes you feel extremely proud at the end. Overall, I am more than happy to say that the long work paid off big time!

First, my gratitude and appreciation go to my thesis advisor Prof. Adrian M. Ionescu for accepting me into his lab and allowing me to work on a very fascinating subject. His extremely smart observations and his ability to see the 'big picture' helped me set my course back on track when I got lost. Many thanks to Prof. Pierpaolo Palestri of University of Udine who acted as my co-advisor. I was amazed to see his keen attention to detail and his willingness to get his hands dirty that saved my back so many times. It is truly an honor to witness our collaboration, which started as a pseudo-random encounter, develop into a full-fledged student-tutor relationship (thank you, Skype). Thanks are due to Dr. Jean-Michel Sallese, Dr. Giovanni Boero, Prof. Mathieu Luisier, Dr. Uygar Avci for kindly accepting to be in my PhD jury and their valuable remarks on my thesis.

Having an ideal lab environment is essential when doing a PhD which can be daunting at times. I consider myself extremely lucky in this aspect, as I was surrounded by very nice and smart people. I thank all my lab members (past and present) including Livio, Luca, Jose, Nilay, Arnab, Emanuele and Maneesha for our collaborations and Wolfgang and Clara for our chitchats and for letting me use the couch when I was too bored or tired to work! Many thanks to Isabelle and Karin for their positive attitudes and their help with the administrative work.

When I first came to Lausanne almost seven years ago for master studies, I must admit that I mainly regarded Lausanne as a stepping stone to the US for the PhD. Indeed my plan seemed to work fine as I got a very compelling offer from a world-renown university in the US as a PhD. However, by the time I was finished with my masters, I was so happy here that I decided to stay. Still, it was the hardest decision I made in my life so far (so, lucky me?). I cannot express enough how happy I am with the outcome. I thank especially Berker, Kerem and Ece for their support during the tough times.

Berker also deserves a very special mention as he is a constant source of calmness, someone to talk to whenever I needed; or, in short, a brother to me. Many many thanks to Kerem for being such a good friend and also an excellent lab/project partner for a good deal of courses since our undergraduate. The Lausanne Turkish gang deserves the highest regards, the Akçok family (Başak, İsmail and Alpin), Nariye, Meriç, Can, Buğra, Gizem, Gizem, Ece and Emre, you made

my time in Lausanne worthwhile. I will never forget the nights at our "MacLaren's" (i.e., Bar Tabac). Special thanks to Dilan, who has been an excellent and caring friend and an essential part of our domestic trio.
I cannot express my feelings for my beloved Handan with words . She is an amazing woman who truly understands me. She is patient enough to put up with my erratic behavior at times. I am so lucky to have you by my side. I know I would be lost without you. Thank you.
Finally, I want to thank my parents Birsen & Yusuf, my brother and sister-in-law Emre & Belgin and my extended family for their constant support. I have not received anything but love, warmth and sympathy from them. I am ever so grateful for making me who I am and providing a perfect life for me as well as some (too) long nights playing video games like teenagers.

Lausanne, 17 February 2017 Cem Alper

Abstract

The semiconductor industry, governed by the Moore's law, has achieved the almost unbelievable feat of exponentially increasing performance while lowering the costs for years. The main enabler for this achievement has been the scaling of the CMOS transistor that allowed the manufacturers to pack more and more functionality into the same chip area. However, it is now widely agreed that the happy days of scaling are well over and we are about to reach the physical limits of the CMOS concept. One major, insurmountable limit of CMOS is the so-called thermionic emission limit which dictates that the switching slope of the transistor cannot go below 60mV/dec at room temperature. This makes it impossible to scale down the supply voltage for CMOS transistor without dramatically increasing the static power consumption. To address this issue, a novel transistor concept called Tunnel FET (TFET) which utilizes the quantum mechanical band-to-band tunneling (BTBT) has been proposed. TFETs possess the potential to overcome the thermionic emission limit and therefore allow for low supply voltage operation.

This thesis aims at investigating the performance of TFETs with alternative architectures exploiting quantized carrier gases through quantum mechanical simulations. To this end, 1D and 2D self-consistent Schrödinger-Poisson solvers with closed boundaries are developed along with the phonon-assisted and direct BTBT models implemented as a post-processing step. Moreover, we propose an efficient method to incorporate the quantization along the transverse direction which enables us to simulate different dimensionality combinations. The implemented models are calibrated against experimental and more fundamental quantum mechanical simulation methods such as $k \cdot p$ and tight-binding NEGF using tunneling diode structures.

Using these tools, we simulate an advanced TFET architecture called electron-hole bilayer TFET (EHBTFET) which exploits BTBT between 2D electron and hole gases electrostatically induced by two separate oppositely biased gates. The subband-to-subband tunneling is first analyzed with the 1D simulator where the device working principle is demonstrated. Then, non-idealities of the EHBTFET operation such as the lateral tunneling and corner effects are investigated using the 2D simulator. The origin of the lateral leakage and techniques to reduce it are analyzed in detail.

A parameter space analysis of the EHBTFET is performed by simulating a wide range of channel materials, channel thickness and oxide thicknesses. Our results indicate the possibility of having 2D-2D and 3D-3D tunneling for the EHBTFET, depending on the parameters chosen.

A novel digital logic scheme utilizing the independent biasing property of the EHBTFET n-and p-gates is proposed and verified through quantum-corrected TCAD simulations. The performance benchmarking against a 28nm FD-SOI CMOS technology is performed as well. The results indicate that the EHBTFET logic can outperform the CMOS counterpart in the low supply voltage (subthreshold) regime, where it can offer significantly higher drive current due to its steep switching slope.

We also compare the different dimensionality cases and highlight important differences between the face and edge tunneling devices in terms of their dependence on the device parameters (channel material, channel thickness and EOT).

Key words: Tunnel FET, quantum simulation, steep slope switch, circuit benchmarking, band-to-band tunneling, analytical modeling, numerical simulation, finite element method, carrier gas dimensionality, density of states (DOS), 2D-2D tunneling, counterdoping, hetero-gate, electron-hole bilayer

Sommario

Basandosi sulla legge di Moore, l'industria dei semiconduttori è stata in grado di mettere in atto un'impresa quasi incredibile : aumentare esponenzialmente le prestazioni dei dispositivi, congiuntamente alla riduzione dei costi. Questo è stato possibile grazie al processo di miniaturizzazione del transistore CMOS, che ha permesso di aumentare sempre di più la densità di dispositivi per unità d'area. Tuttavia, è ormai comunemente accettato il fatto che questo processo stia giungendo a un termine, essendo al punto di raggiungere i limiti fisici della tecnologia CMOS. Ad esempio, un limite insormontabile relativo a questa tecnologia è dovuto alla cosiddetta emissione termoionica, che fa sì che la pendenza di sottosoglia non possa essere inferiore a 60 mV/dec a temperatura ambiente. Per questo motivo, risulta impossibile ridurre la tensione di alimentazione per i transistori CMOS senza aumentare esponenzialmente il consumo di potenza statico. Per risolvere questo problema, è stato proposto un transistore alternativo chiamato Tunnel FET (TFET), basato sul meccanismo di tunneling quantistico banda-banda. I TFET offrono la possibilità di superare il limite dell'emissione termoionica e così permettere di ridurre la tensione di alimentazione.

L'obiettivo di questa tesi consiste nello studio delle prestazioni di TFET con architetture alternative, basate su gas di portatori di carica quantizzati. Lo studio è stato effettuato attraverso simulazioni numeriche, includendo effetti di meccanica quantistica. A tal fine, sono stati sviluppati programmi per risolvere in maniera autoconsistente problemi Schrödinger-Poisson 1D e 2D, congiuntamente all'uso di modelli di tunneling quantistico banda-banda diretto o assistito da fononi, applicati in una fase successiva. Inoltre, in questa tesi proponiamo un metodo per considerare la quantizzazione nella direzione trasversa, che ci permette di simulare efficientemente diverse combinazioni di dimensionalità. I modelli implementati sono stati calibrati avendo come termine di paragone dati sperimentali e metodi di simulazione alternativi basati su simulazioni di meccanica quantistica fondamentale, come il $k \cdot p$ e il tight-binding NEGF basato su strutture di diodo a effetto tunneling.

Tramite questi strumenti, abbiamo simulato un'architettura TFET avanzata denominato TFET a doppio strato elettrone-lacuna (electron-hole bilayer TFET o EHBTFET), che si basa sul tunneling banda-banda tra gas bidimensionali di elettroni e lacune, indotti in modo elettrostatico da due terminali di gate disposti oppositamente. Prima di tutto abbiamo analizzato il tunneling da sottobanda a sottobanda con un simulatore monodimensionale, al fine di dimostrare il principio di funzionamento del dispositivo. Successivamente il simulatore 2D è stato impiegato per studiare gli aspetti non-ideali dell'EHBTFET, come il tunneling laterale e gli effetti agli angoli. Le cause del tunneling laterale e le tecniche per ridurlo sono state

analizzate in dettaglio.
Abbiamo effettuato uno studio dei parametri di progetto dell'EHBTFET simulando una vasta gamma di materiali per il canale, e di spessori per il canale e per il dielettrico. I nostri risultati dimostrano che a seconda dei parametri l'EHBTFET può esibire tunneling sia 2D-2D che 3D-3D. Inoltre abbiamo proposto e verificato attraverso simulazioni TCAD con correzioni quantistiche uno schema logico digitale innovativo basato sulle proprietà di polarizzazione indipendente dei gate di tipo n e p. dell'EHBTFET. Inoltre abbiamo confrontato le prestazioni dell'EHBTFET rispetto a quelle offerte dalla tecnologia FD-SOI CMOS a 28 nm. I risultati mostrano che una logica basata sull'EHBTFET può fornire migliori prestazioni rispetto alla logica CMOS in regime di bassa tensione di alimentazione, dove la corrente può essere più alta a causa della maggiore pendenza sottosoglia. Abbiamo anche confrontato diversi casi di dimensionalità e sottolineato importanti differenze tra dispositivi basati su tunneling di superficie o di spigolo in quanto a dipendenza su parametri di progetto (materiale e spessore del canale, spessore effettivo del dielettrico).

Parole chiave : tunnel FET, simulazione quantistica, interruttore ad alta pendenza sottosoglia, analisi comparativa di circuiti, tunnneling banda-banda, modellizzazione analitica, simulazione numerica, metodo ad elementi finiti, dimensionalità di gas di portatori di carica, densità di stati (DOS), tunneling 2D-2D, controdopaggio, eterogate, doppio strato elettrone-lacuna.

List of Publications

Parts of this thesis have been published in the papers listed below:

1. C. Alper, J. L. Padilla, P. Palestri, A. M. Ionescu. "A novel reconfigurable sub-0.25V digital logic family using the electron-hole bilayer TFET," *IEEE Transactions on Electron Devices*, In preparation.
2. D. Esseni, P. Palestri, M. Pala, C. Alper[1], T. Rollo. "A review of selected topics in physics based modelling of small slope transistors based on band-to-band-tunnelling," *Semiconductor Science and Technology*, In preparation.
3. C. Alper, J. L. Padilla, P. Palestri, A. M. Ionescu. "Impact of device geometry of the fin Electron-Hole Bilayer Tunnel FET," *ESSDERC 2016 - 46th European Solid State Device Research Conference*, Sept. 12-15, 2016.
4. C. Alper, P. Palestri, J. L. Padilla, A. M. Ionescu. "The Electron-Hole Bilayer TFET: Dimensionality Effects and Optimization," *IEEE Transactions on Electron Devices*, 63(6), pp. 2603 - 2609, 2016.
5. C. Alper, P. Palestri, J. L. Padilla, A. M. Ionescu. "Underlap counterdoping as an efficient means to suppress lateral leakage in the electron-hole bilayer tunnel FET," *Semiconductor Science and Technology*, 31(4), pp.045001, 2016.
6. C. Alper, M. Visciarelli, P. Palestri, J. L. Padilla, A. Gnudi, E. Gnani, A. M. Ionescu. "Modeling the imaginary branch in III-V tunneling devices: Effective mass vs $k \cdot p$," *SISPAD 2015*, Sept. 9-11, 2015.
7. C. Alper, P. Palestri, J.L. Padilla, A. Gnudi, R.Grassi, E. Gnani, M. Luisier and A. M. Ionescu. "Efficient quantum mechanical simulation of band-to-band tunneling," *EUROSOI-ULIS 2015*, Jan. 26-28, 2015.
8. A. M. Ionescu, C. Alper, J. L. Padilla, L. Lattanzio, P. Palestri. "Electron-hole bilayer deep subthermal electronic switch: Physics, promise and challenges," *2014 SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S)*, Oct. 6-9, 2014.
9. C. Alper, P. Palestri, L. Lattanzio, J. L. Padilla and A. M. Ionescu. "Two dimensional quantum mechanical simulation of low dimensional tunneling devices," *Solid State Electronics*, 113, pp. 167-172, 2015.

---

[1] Final author ordering TBD.

10. C. Alper, L. De Michielis, L. Lattanzio, P. Palestri, L. Selmi and A. M. Ionescu. "Quantum mechanical simulation of the germanium electron-hole bilayer TFET," *IEEE Transactions on Electron Devices*, 60(9), pp. 2754-2760, 2013.

A complete list of publications can be found at the end of the thesis.

Contents

List of Figures

List of Tables

1 Towards the End of CMOS Scaling & Motivation of the Thesis

Ever since the advent of integrated circuits (IC) based on CMOS technology, the semiconductor industry experienced an era of constant increase in performance and scaling of transistors, which occurred at a neck-break pace. Much of this outcome owes to a key observation made by Gordon Moore, one of the founders of Intel, in 1965 which stated that the number of transistors in a given area will double in approximately 24 months and the performance will double too [17] . Over decades, this observation transformed into a *self-fulfilling prophecy* [18], pushing the semiconductor companies to pack more and more transistors into chips operating at consistently higher speeds. This rapid rate of advancement is indeed phenomenal when you compare it against other industries such as automotive etc. Imagine that your car would consume half the fuel to go the same amount of distance in two years' time (and twice as fast)! Constant scaling of transistors throughout the 'happy scaling era' [19] is what enabled the Moore's law to hold more or less until now and keep the steam going for the semiconductor industry.

However, this fast growth came at a price: another relatively less mentioned aspect of Moore's law is the exponential increase in investments for semiconductor foundries. Fig. 1.1(Left) depicts the investment capital costs for exposure systems for different line widths, which we take as a general indicator of skyrocketing investment costs overall [1]. Looking from an economic perspective, semiconductor industry is fast becoming a mature market like automotive industry. We have already observed some big companies such as IBM and AMD sell their foundries (IBM even paid Globalfoundries to take over its fabs! [20]) simply due to huge operating costs incurred in investments and R&D. In fact, the average growth of the semiconductor industry between 1988-2012 was 10.5%, whereas it is estimated to be 2.72% for 2013-2018 [21]. Overall, we can argue that the Moore's law applies on the economic side of business as well as the technical ones.

One might wonder why it has become so expensive to develop the next CMOS node. The answer is that we are approaching the physical limits of CMOS and Silicon technology. We are going to need much higher resolution photo-lithography, very good process control, good reliability etc. Fig. 1.1 (Right) shows the physical transistor half-pitch length over each

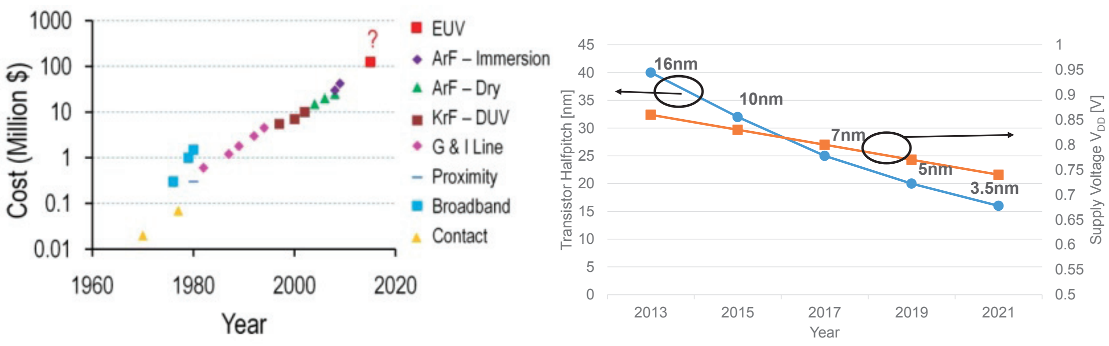


Figure 1.1 – (Left) Exposure system costs over time (taken from [1]). (Right) The transistor half-pithches and supply voltages over year and indication of the 'technology node' name given by companies (labels in bold) [2].

technology node. We can already see that the node naming has started to lose its physical meaning as the naming and the physical length are no more the same [2], indicating how hard it has become to physically shrink the transistors.

Besides this problem which could potentially be overcome with a lot of engineering such as EUV as the light source for lithography [22], a much more pressing issue is at hand. No matter how much we try to perfect the CMOS transistors, there is an inherent physical limit. Supply voltage scaling is no more possible, as seen in Fig. 1.1(Right), where we can indeed observe that the supply voltage reduction is limited to about 30mV per node. The reason for this stall in supply voltage scaling is that the OFF state current in MOSFET increases exponentially as the supply voltage is lowered. The reason for this exponential increase is the fact that MOSFET is a thermionic device [23], i.e., it relies on the movement of thermally activated carriers which obey the Fermi-Dirac distribution. Fig. 1.2 depicts a scenario where the supply voltage V_{DD} is lowered to $V_{\mathrm{DD}'}$ via e.g., gate metal workfunction engineering where the transfer characteristic curve is rigidly shifted to the left, but the switching slope remains approximately the same.

More specifically, the reason that spelled the doom for the conventional scaling of the supply voltage is the subthreshold current that grows at an exponential rate as the supply voltage is shrunk. The physical basis for this is the fact that sub-threshold current for MOSFETs is dominantly caused by diffusion in weak inversion, for which the drain current is given as [24]:

$$I_{\mathrm{DS}} \sim \exp \frac{V_{\mathrm{GS}}}{\eta V_{\mathrm{T}}} \tag{1.1}$$

where η is given as $\eta = 1 + \frac{C_{\mathrm{S}}}{C_{\mathrm{G}}}$, C_{S} being the semiconductor capacitance and C_{G} the gate oxide capacitance. The logarithmic slope of the transfer characteristics is defined as the *subthreshold*

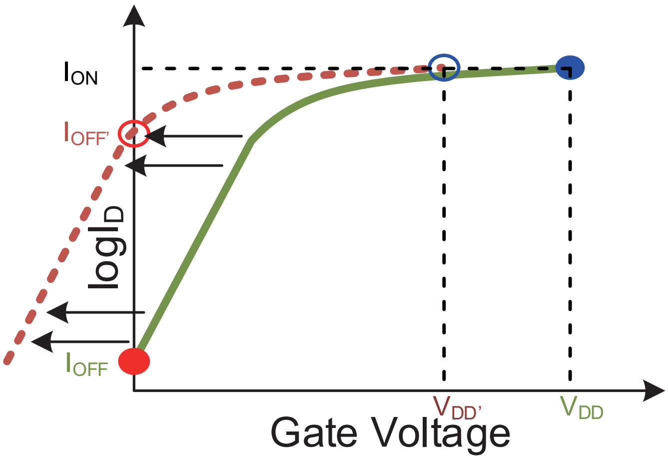


Figure 1.2 – Depiction of MOSFET transfer ($I_\mathrm{D}-V_\mathrm{GS}$) characteristics. Given the same MOSFET, shifting the transfer curves by modifying the gate workfunction to reduce the supply voltage from V_DD to $V_{\mathrm{DD}'}$ with the same ON current I_ON results in an exponential increase in the OFF current (I_OFF to $I_{\mathrm{OFF}'}$).

swing (SS). Mathematically, SS is given as [25]:

$$SS = \frac{\partial V_\mathrm{GS}}{\partial \log_{10} I_\mathrm{DS}} = \underbrace{\frac{\partial V_\mathrm{GS}}{\partial \phi_\mathrm{S}}}_{\eta}\underbrace{\frac{\partial \phi_\mathrm{S}}{\partial \log_{10} I_\mathrm{DS}}}_{n} \tag{1.2}$$

It is evident that in order to improve the SS, ratio $\frac{C_\mathrm{Si}}{C_\mathrm{G}}$ must be minimized such that η is minimized. In CMOS, η will tend to 1 as the capacitance ratio is minimized. In general terms, it corresponds to the case with excellent electrostatic control. However, even with this perfect control, n factor is dependent on the rate of increase in the number of carriers that have energy larger than the barrier height in the channel and this quantity is governed by the Fermi-Dirac distribution; therefore n has a lower bound of $\sim \frac{kT}{q}\ln 10 \simeq 60\frac{\mathrm{mV}}{\mathrm{dec}}$ at room temperature. This limitation of the n factor has dramatic consequences for voltage scaling. It imposes a limit on the ON/OFF ratio. For instance, if an ON/OFF ratio of 10^5 is required, a minimum supply voltage of 300mV is required. Therefore, we can argue that using CMOS in a low supply voltage scenario is nearly impossible and overcoming this limit requires a paradigm shift in the device design. It means that we will either need to rethink the concept of transistor by considering different conduction mechanisms, or somehow combine the conventional CMOS with some technological boosters that will allow it to overcome this thermionic barrier.

Looking at Eq. 1.2, one can observe that we have essentially two options: either make $\eta < 1$ or $n < 60\frac{\mathrm{mV}}{\mathrm{dec}}$. In the literature, novel device concepts were proposed which can be classified to the either case. For example, ferro-electric MOSFETs which can achieve internal voltage amplification thanks to the ferro-electric dielectric material in the gate stack that allows for $\eta < 1$ have been demonstrated experimentally [26]. Phase change materials have been used in a similar fashion [27]. Electro-mechanical FETs [28] relies on the mechanical movement of the gate electrode which gets in and out of contact with the gate stack depending on the gate voltage whereas NEMS relays [29]) or swtiches [30] makes use of moving beams/cantilevers that connects the source and drain electrodes.

Another way is to alter the conduction mechanism. In this branch, we can see that a plethora of devices such as impact ionization FETs which make use of avalanche multiplication that amplifies the current once a critical field is exceeded [31]. Another idea is the so-called 'feedback FET' which utilizes extra charges in a p-i-n diode that alters the potential profile which results in positive feedback [32].

In this thesis, however, we will investigate a type of transistors called 'tunnel FET' that exploits the quantum mechanical phenomenon called band-to-band tunneling (BTBT). TFET has emerged as one of the most promising new device concepts that can overcome the thermionic barrier by filtering out the exponential tail of the Fermi-Dirac distribution [33]. We will go into more detail about the working principle of TFETs in the next chapter.

1.1 Thesis Outline

The thesis outline with the highlights of each chapter is given below:

Chapter 2: Tunnel Field Effect Transistors: An Introduction and State-of-the-Art

This chapter introduces the concept of band-to-band tunneling, its history of exploitation as well as a basic intuitive physical picture. Then the concept of tunnel FET is introduced and different variants of tunnel FETs proposed in the literature are presented. Then a survey of direct and phonon-assisted band-to-band tunneling models for bulk and quantized structures is given. Finally, we detail the Vandenberghe model that we have utilized in the thesis to estimate phonon-assisted band-to-band tunneling rate.

Chapter 3: Quantum Mechanical Models for Band-to-band Tunneling Devices

This chapter describes the quantum mechanical numerical simulator developed for this thesis. The Schrödinger and Poisson equations that govern the electrostatics and quantum states in the device are introduced. The numerical discretization schemes employed for the 1D and 2D variants of the simulators are given in detail. The solution of the self-consistent system of equations and the algorithmic approach for coupling the two equations is detailed. The non-parabolicity corrections are then introduced to extend the range of validity of the effective mass approximation employed in the 1D code and two different non-parabolicity models are discussed and compared. We then discuss the models that we have modified and implemented in our code to estimate direct BTBT in quantized carrier gases. We highlight the differences compared to the bulk case, such as the direction dependence of the BTBT coupling element.

Two different direct BTBT models are compared and benchmarked through quantum well diode simulations. Following this, we introduce a general and efficient method to account for transverse quantization which can be utilized to estimate the impact of carrier dimensionality on the BTBT current. We then propose an efficient way to account for the "anti-crossing" effect that results in a drastic asymmetry of the effective masses in the real and imaginary branches of the heavy hole dispersion relation. We compare our results and calibrate our simulator using InAs and Ge tunnel diode characteristics simulated by full-quantum approaches $k \cdot p$ and TB-NEGF respectively. For Si and Ge diodes, we also provide a comparison of our results against experimentally measured tunnel diode data found in the literature.

Chapter 4: Dimensionality Effects in Tunneling Devices: Exploitation and Device Optimization

This chapter presents the results obtained by the quantum mechanical simulators developed in the previous chapter. The bulk of the chapter is devoted to the *electron-hole bilayer tunnel FET* (EHBTFET). We first present our initial results obtained with our 1D simulator, where we show that EHBTFET is actually a subband-to-subband tunneling device, that differs from other (bulk) TFETs. We then report our simulation results of the EHBTFET using the 2D simulator and we uncover that a serious non-ideality, lateral tunneling, exists in the EHBTFET that cripples the promising switching slope estimates of the 1D model. We identify wavefunction penetration from the drain into the underlap region as the source of this non-ideality. We then propose two different solution methods to overcome this leakage. Then, we perform a parametric analysis of the EHBTFET using our 1D code. Our results highlight that the EHBTFET can operate both as a 2D-2D tunneling device as well as a (pseudo-)3D-3D one. We identify the distinctions between the two cases and explain the physical basis of the characteristics of the 2D-2D tunneling through an analytical model. Moreover, we estimate the performance upper bounds for the EHBTFET with various channel materials. We also simulate the fin version of the EHBTFET and discover another non-ideality particular to the fin version related to the corner effect. We propose different solution techniques and compare their effectiveness in suppressing this non-ideality. We then move on to the circuit level and propose a novel digital logic family for the EHBTFET that can potentially offer significant reductions in the transistor count. We verify our idea through quantum-corrected TCAD simulations and benchmark the EHBTFET logic against a state-of-the-art 28nm fully depleted silicon-on-insulator (FD-SOI) process using various test circuits. We then compare 2D-2D face and edge tunneling devices against each other and their 1D counterparts (i.e., 1D-1D face and edge tunneling) using our code with transverse quantization and pinpoint the differences between each case.

Chapter 5: Conclusions & Future Perspectives

This chapter summarizes the main accomplishments presented in the preceding chapters. Also, possible future directions in terms of continuation of this work as well as broader general research ideas are provided.

2 Tunnel Field Effect Transistors: An Introduction and State-of-the-Art

This chapter reviews the working principle of the Tunnel FETs (TFET), along with a literature survey about the current progress in TFET design and optimization. Moreover, a brief survey about the proposed numerical models for describing band-to-band tunneling is provided.

2.1 Band-to-band Tunneling

TFETs utilize a quantum mechanical phenomenon called band-to-band tunneling (BTBT) for the modulation of the conductance. BTBT occurs when the valence and conduction band edges of a semiconductor are positioned such that an energy window ΔE opens (see Fig. 2.1) that allows for tunneling from the valence to conduction band (or vice versa). BTBT is typically characterized by very abrupt current changes since it is directly affected by the alignment of the bands. Moreover, it does not rely on thermionic emission of carriers which is governed by the Fermi-Dirac distribution. This property allows for steep switching slope values that are independent of the 'Boltzmann tyranny' [34]. The phenomenon is well-known and utilized in Zener diodes to achieve soft-breakdown at relatively lower voltages (compared to avalanche diodes) and also in Esaki diodes [35] to obtain negative differential resistance (NDR), useful in microwave applications [36].

Despite being a complex quantum mechanical effect, band-to-band tunneling at a p-n junction can be modeled fairly well using the WKB approach. The transmission probability through an energy barrier is given by [37]:

$$T = \exp\left(-\int_{0}^{t_c} 2\left|\kappa\left(x\right)\right| dx\right) \tag{2.1}$$

where $x = 0$ and $x = t_\mathrm{c}$ are the classical turning points and $\kappa(x) = \frac{\sqrt{2m^*(E-V(x))}}{\hbar}$ is the imaginary wave-vector. In the case of the p-n diode, we can approximate the electric field as constant in the depletion region. So, making use of the familiar triangular potential barrier approximation (see. Fig. 2.1), the transmission probability of an electron through this barrier is given by the

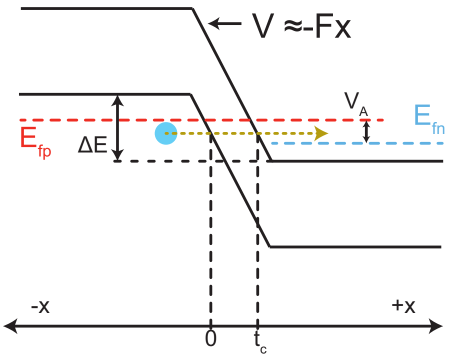


Figure 2.1 – BTBT at the p-n junction approximated as a transmission of electrons through a triangular potential barrier with an applied voltage V_{A}.

following analytical expression [38]:

$$T = \exp\left(-\frac{4\sqrt{2m^*}E_{\mathrm{G}}^{\frac{3}{2}}}{3q\hbar F}\right) \tag{2.2}$$

Looking at the above equation, we can see that the transmission probability depends on the barrier height (i.e., the bandgap), the effective mass of electrons and the electric field.

In this approximation, the effective mass of electron is assumed to be constant through its entire trajectory. A slightly more nuanced approach is to note that the effective mass will tend to the hole mass as it gets closer to the valence band end of its trajectory. Such an effect can be modeled as series of two triangular barriers where the region closer to the conduction band is assumed as electron tunneling and the other region as hole tunneling [39] which is closer to the actual physical picture. In section 2.3 and Chapter 3, we will show more sophisticated models which provide a more accurate picture of the BTBT phenomenon.

2.2 TFETs

The idea of utilizing band-to-band tunneling as a conduction mechanism in a transistor have been around since the 1980s when the first interband tunneling devices were proposed which commonly made use of III-V hetero-structures [40] and were deemed as a variant of the then-popular resonant tunneling diodes [41]. The main emphasis back then was to obtain the NDR behavior, whereas digital logic applications was of secondary importance [42].

The first proposition of the conventional TFET as we know today dates back to as early as 1992 [43] when Baba proposed the p-i-n gated diode structure and the working principle for the conventional TFETs. However, the idea did not seem to gain traction within the community when it was first proposed. Around the beginning of the 2000s, the landscape for the emerging device research had changed dramatically as CMOS scaling started facing serious challenges.

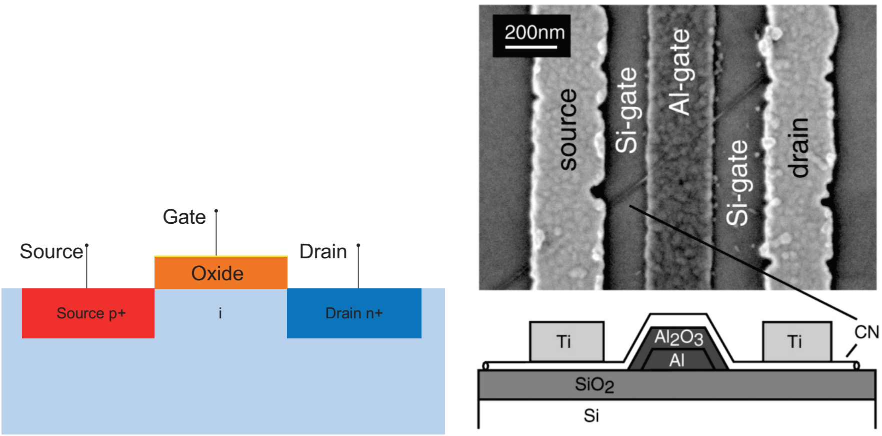


Figure 2.2 – (Left) Planar n-type TFET structure utilizing a gated p-i-n diode with the BTBT region indicated. Note that in theory, the same structure could be operated as a p-type device as well by reversing the voltage polarities and swapping the source and the drain electrodes. (Right) The carbon nanotube TFET (CNT TFET) structure fabricated by IBM (image taken from [3]).

Within this new context, band-to-band tunneling devices started garnering considerable attention as a next generation device that could potentially be more resilient against the short channel effects [44]. The modern concept of Tunnel FETs for low power applications emerged when IBM demonstrated the first TFET utilizing carbon nanotubes and experimentally shown that sub-thermal ($SS < 60\frac{\text{mV}}{\text{dec}}$) conduction is possible [3]. These works have laid out the basic structure that was extensively utilized by the *first generation* TFETs. The layouts of these transistors were very similar to those of MOSFET, except for the use of a p-i-n doping profile, Fig. 2.2(Left). The overall structure, therefore, is a gated p-i-n diode.

The principle of operation of the n-type TFET of Fig. 2.2(Left) is sketched in Fig. 2.3. The idea is as follows: the gate electrode controls the band profile inside the channel region where it shifts the bands according to the applied gate voltage. BTBT is allowed when the conduction band edge at the channel and the valence band edge at the source are aligned. The BTBT barrier at the source-channel interface is modulated by the applied gate bias which in turn determines the BTBT current. One major downside of this structure is the low ON current levels caused by the fact that BTBT is concentrated around a very narrow region around the source-channel interface and the fact that gate control over the lateral electric field around the tunnel junction is relatively poor. These types of structures are commonly referred as *point tunneling* devices [45]. Moreover, as seen in Fig. 2.3(Right), TFETs are prone to ambipolar conduction [46], which is caused by the formation of a tunneling energy window at the drain side when the transistor is biased in opposite polarity compared to the ON state. To overcome this issue, drain underlaps [47] as well as utilizing heterojunctions at the source-channel interface have been proposed [48].

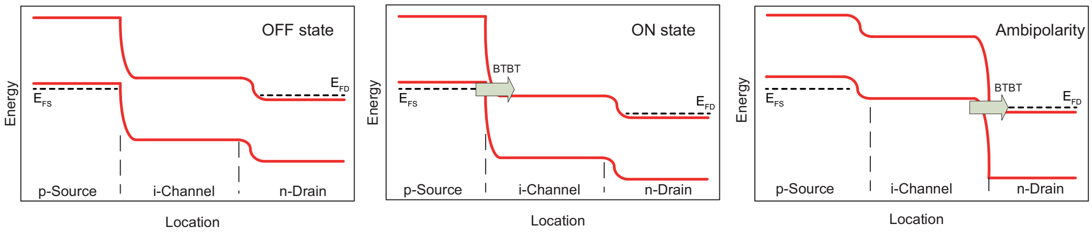


Figure 2.3 – Sketch of n-type TFET device operation: (Left) When $V_{GS} \sim 0$ the TFET is in the OFF state, since no tunneling energy window is present. (Middle) $V_{GS} > 0$ the band profile in the channel is pushed down such that a BTBT energy window is formed at the source-channel interface. (Right) When $V_{GS} < 0$ ambipolar conduction occurs at the drain-channel interface. $V_{DS} > 0$ in all sketches.

To address the issue of low ON current, a second generation of TFETs is proposed which utilize the so-called *line tunneling* [45], in which the tunneling direction is aligned with the electric field induced by the gate voltage. This is achieved by overlapping the gate stack with the source region, which creates an inversion layer inside the source region through which BTBT takes place. Fig. 2.4 depicts one example of the line tunneling device concept as proposed by de Michielis et al. [4]. A drastic improvement in the ON current was indeed observed when gate-source overlap is introduced since the current is now proportional to the overlap area. Line tunneling concept was experimentally proven by Schmidt et al. [49] where a linear relationship was observed between the ON current and the tunneling area. The 'green transistor' concept proposed in [50] and also [51] are also other examples of such devices.

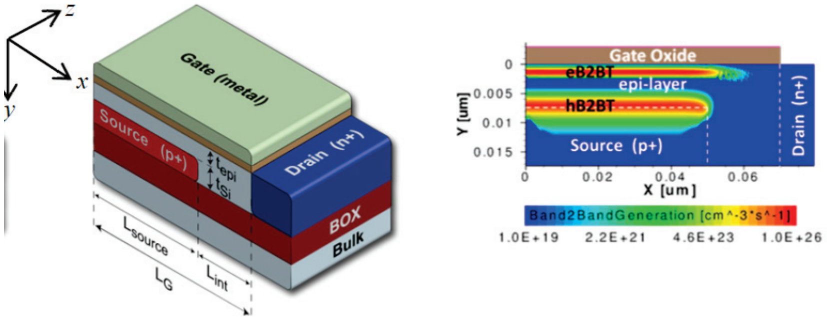


Figure 2.4 – (Left) Line tunneling TFET proposed by de Michielis et al. [4]. The inversion layer is formed in the epi-layer grown on top of the source region by overlapping the gate and the source. (Right) Simulation of the generation plot of electron and hole pairs (images taken from [4]).

As a natural extension to the line tunneling devices, Lattanzio et al. [5] proposed the so-called *electron hole bilayer TFET* (EHBTFET) in which both electron and hole gases are induced through electrostatics (see Fig. 2.5). In this device, a positive (negative) gate bias is applied to the n-gate (p-gate) to induce 2D electron (hole) gases. The gate stacks are misaligned in order to prevent lateral tunneling. Initial semi-classical simulations have shown very promising results with a near-ideal SS $10\frac{\text{mV}}{\text{dec}}$ [5] and encouraging initial results for circuit operation [52]. However, due to the thin semiconductor layer employed and the anti-symmetric biasing of the gates, size and field induced quantization is expected to be very strong and therefore a quantum mechanical that can take such effects duly into account is necessary. In fact, the quantum mechanical study, optimization and exploitation of the EHBTFET will be the bulk of this thesis.

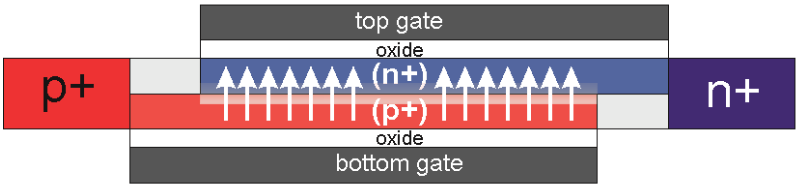


Figure 2.5 – The EHBTFET structure proposed by Lattanzio et al. [5]. Image taken from [5].

Simultaneously with these development, the concept of *density of states (DOS) switch* was proposed by Agarwal & Yablonovitch [15]. The idea is to exploit the different characteristics of the DOS functions of different dimensions in TFETs. More specifically, DOS switch proposes to exploit the step-like and spiking (i.e., $1/\sqrt{E}$) behaviors of 2D and 1D DOS functions respectively which would ideally allow for significant number of states right after subband alignment and therefore improve the switching slope compared to the 3D (bulk) case for which the $DOS \propto \sqrt{E}$ (see Fig. 2.6). We will go into more detail on this concept in Chapter 4.

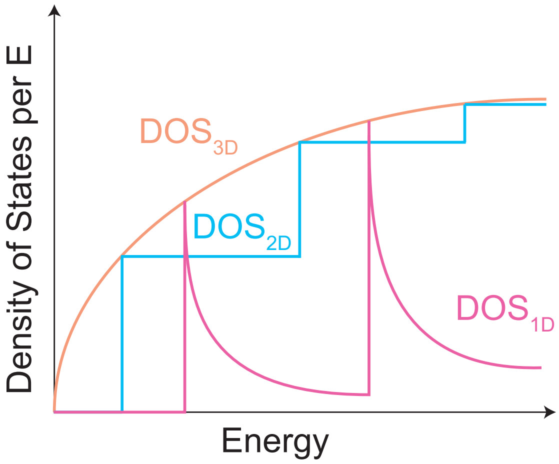


Figure 2.6 – Qualitative behavior of DOS functions with different carrier dimensionalities.

2.3 Survey of the BTBT Tunneling Models

The BTBT can be classified into two types of transitions: direct and phonon-assisted. The first one denotes a process when the involved extrema of the electron and hole bands are at the same $\vec{k}$ point in the Brillouin zone (for example in InAs). Since BTBT is an elastic process, the total energy and transverse momentum are conserved. For the phonon-assisted BTBT, the extrema of the electron and hole bands reside at different $\vec{k}$-points (for example in Si and Ge). Since they are at different points in the Brillouin zone, the transverse momentum offset is compensated by the absorption or the emission of a phonon corresponding to the difference in the crystal momentum of the connected band extrema. Fig. 2.7 depicts the two types of transitions.

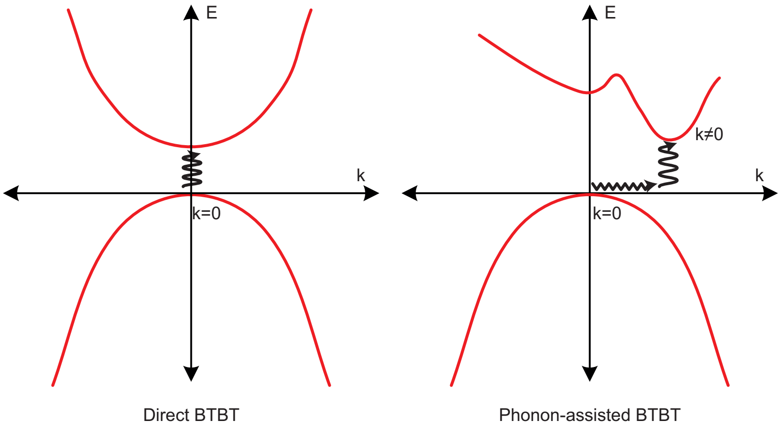


Figure 2.7 – Different types of band-to-band tunneling: (Left) direct tunneling, where tunneling takes place between the bands around the same $\vec{k}$-point. (Right) phonon-asssited tunneling, where the bands at different $\vec{k}$-points corresponding to their energy extrema are connected by means of a phonon which compensates the discrepancy in the momentum.

In this section, we review some of the prominent band-to-band tunneling models for both direct and phonon-assisted transition. The aim is to highlight the similarities and differences between different approaches. The models for BTBT in quantized gases developed during the PhD will be described in Chapter 3.

2.3.1 Direct Band-to-band Tunneling in Bulk Structures

Among the various models for direct BTBT in the literature, we can see that there are essentially two types of models, based on the calculation of the tunneling rate spectrum. In the first group the tunneling rate is calculated with transmission coefficient obtained with the WKB approximation then integrated over the total energy and transverse momentum of the particles. The second type makes use of the Fermi's Golden Rule, which estimates the tunneling rate by calculating the magnitude of the interband coupling coefficient M_{CV}.

Kane Theory of Tunneling under Uniform Fields

Kane proposed his band-to-band tunneling theory in his seminal paper in 1959 [9]. The paper made use of the perturbation theory to write the Schrödinger equation as a sum of interband and intraband coupling elements. He further illustrated his model using a simple two-band model to explain tunneling phenomenon in InSb Zener diodes. Here we provide a different, simpler proof of his results based on Landauer conduction formula and WKB approximation. Note that the model assumes a uniform electric field, which make it impossible to use it as a 'nonlocal' model. 'Local' in this context means that the model does not take into account the trajectory of the electron during which it could experience drastic changes in the magnitude and the direction of the electric field which makes it suitable only for regular devices with relatively uniform electric field profiles such as p-n diodes.

For the derivation, we start by writing the general expression of the current density using the Landauer's conduction formula [53]:

$$J = \frac{q}{\pi \hbar A} \int \sum_{\vec{k}_\perp} T(E, \vec{k}_\perp) \left[f_\mathrm{v} - f_\mathrm{c} \right] dE \tag{2.3}$$

where A is the area of the device in the direction normal to tunneling, $k_\perp$ is the transverse wave-vector, $T(E, k_\perp)$ is the transmission probability and $f_\mathrm{v} - f_\mathrm{c}$ is the Fermi supply function. We make the usual discrete sum to integral transform for a 2D k-space:

$$J = \frac{q}{4\pi^3 \hbar} \iint T(E, k_\perp) \left[f_\mathrm{v} - f_\mathrm{c} \right] d\vec{k}_\perp^2 \, dE \tag{2.4}$$

To follow Kane's approach, we set $f_\mathrm{v} = 1$ and $f_\mathrm{c} = 0$, although it is now seen essential to include the Fermi supply term in BTBT tunneling models due to the fact that non-zero current is predicted even when zero bias is applied at the junction [54]. One should note here that, within the semi-classical framwork, BTBT has been considered as a generation process and therefore incorporated into the continuity equations whereas till now the expressions we deal with are spectral current densities (i.e., current flux through an energy bin ΔE, see Fig. 2.8). Under constant field, we convert this spectral density to a generation rate by using the relation starting from the continuity equation (see Fig. 2.8):

$$\frac{dJ}{dx} = -qG_\mathrm{BTBT} \tag{2.5}$$

$$\frac{dJ}{dx}\frac{dx}{dV} = -q\frac{dx}{dV}G_\mathrm{BTBT} \tag{2.6}$$

$$\frac{dJ}{dE} = \frac{1}{|F|}G_\mathrm{BTBT} \tag{2.7}$$

Plugging in the expression in Eq. 2.4 into Eq. 2.7, switching to polar coordinates in $|k_\perp|$ and

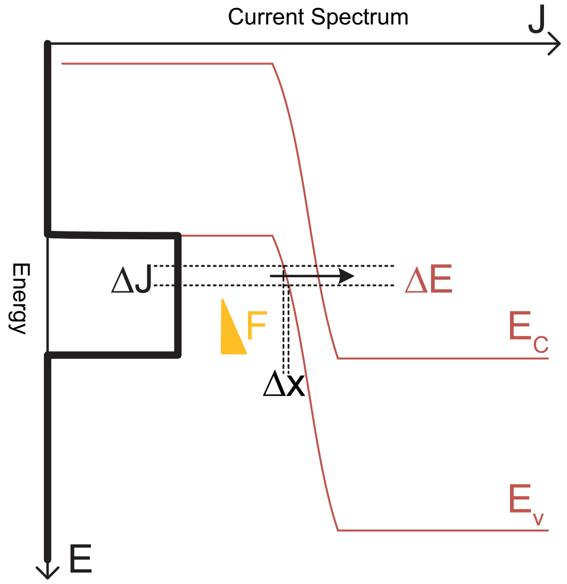


Figure 2.8 – Depiction of BTBT under uniform field. The overall generation rate G_{BTBT} at a position x is given as an integral of spectral BTBT current densities ΔJ over all the energy bins ΔE.

assuming T being independent of the k-vector direction, we obtain:

$$G_{\mathrm{BTBT}} = \frac{q|F|}{4\pi^3\hbar}2\pi\int T(E,k_\perp)|k_\perp|d|k_\perp| \tag{2.8}$$

where the transmission rate is given by the WKB approximation:

$$T = \frac{\pi^2}{9}\exp\left(-2\int \mathrm{Im}(k_x)dx\right) \tag{2.9}$$

It is evident that we need an $E-k$ relation valid into the gap (i.e., describes $\mathrm{Im}(k_x)$) which connects the valence band to the conduction band to perform the integral in Eq. 2.9. Moreover, since direct BTBT is an elastic process total transverse momentum and energy needs to be conserved. Among various possibilities such as Flietner's dispersion [55], we utilize, in the following, Kane's two-band dispersion relation [56]:

$$E_\pm = \frac{E_G}{2} + \frac{\hbar^2k^2}{2m_0} \pm \frac{1}{2}\sqrt{E_G^2 + \frac{E_G\hbar^2k^2}{2m_r}} \tag{2.10}$$

where $\frac{1}{m_r} = \frac{1}{m_e} + \frac{1}{m_h}$ and $E_{+(-)}$ denotes the electron (hole) band. Assuming the electric field F is uniform and along x-direction, k_x inside the bandgap (which is imaginary) can be written as:

$$\mathrm{Im}(k_x) = \sqrt{\frac{m_\mathrm{r}}{E_\mathrm{G}\hbar^2}}\sqrt{E_\mathrm{G}^2 + E_\mathrm{G}\frac{\hbar^2|k_\perp|^2}{m_\mathrm{r}} - 4\left(q|F|x - \frac{E_\mathrm{G}}{2}\right)^2} \tag{2.11}$$

where $m_\mathrm{r} = \frac{m_\mathrm{e}m_\mathrm{h}}{m_\mathrm{e}+m_\mathrm{h}}$ and we set $x = 0$ at $E = 0$ (the classical turning point). Once determining the classical turning points of the conduction and valence bands, it remains to integrate the WKB transmission coefficient over the x-direction:

$$T = \frac{\pi^2}{9}\exp\left(-\pi\frac{\sqrt{m_\mathrm{r}}E_\mathrm{G}^{\frac{3}{2}}}{2\hbar q|F|}\right)\exp\left(-\frac{\pi\hbar\left|\vec{k}_\perp\right|^2}{2q|F|}\sqrt{\frac{E_\mathrm{G}}{m_\mathrm{r}}}\right) \tag{2.12}$$

Putting the resulting expression into Eq. 2.8 and setting $|k_\perp| = 0$, we finally obtain the well-known expression:

$$G = \frac{q^2|F|^2\sqrt{m_\mathrm{r}}}{18\pi\hbar^2\sqrt{E_\mathrm{G}}}\exp\left(-\pi\frac{\sqrt{m_\mathrm{r}}E_\mathrm{G}^{\frac{3}{2}}}{2\hbar q|F|}\right) \tag{2.13}$$

Non-Local Band-to-band Tunneling Model

The local models such as the ones presented above have the problem of grossly overestimating the current when the electric field rapidly changes over relatively small distances. In order to overcome this issue within the semiclassical framework, the so-called non-local models were introduced [57, 58]. The particular aim in these models was to introduce the creation of carriers at either side of the tunneling barrier, in contrast to the local models which unphysically assign electron and hole generation at the same position.

As an illustrative example, a non-local direct BTBT model based on Kane's two band dispersion could be derived by obtaining an analytical expression for the imaginary dispersion κ inside the forbidden gap, by writing the inverse of Eq. 2.10:

$$\kappa = \frac{1}{\hbar}\sqrt{m_\mathrm{r}E_\mathrm{G}\left(1-\alpha^2\right)} \tag{2.14}$$

$$\alpha = -\frac{m_0}{2m_\mathrm{r}} + 2\sqrt{\frac{m_0}{2m_\mathrm{r}}\left(\frac{E-E_\mathrm{V}}{E_\mathrm{G}} - \frac{1}{2}\right) + \frac{m_0^2}{16m_\mathrm{r}^2} + \frac{1}{4}}$$

We assume that the tunneling path connects the points x_i and x_f which denotes the initial and the final positions of the carrier and $E_\mathrm{V}(x_i) = E_\mathrm{C}(x_f)$. Making use of Eqs. 2.9 and noting that only small $k_\perp$ provide high contribution to the tunneling, we can write $\mathrm{Im}(k_x) = \sqrt{\kappa^2 + k_\perp^2} \simeq$

$\kappa + \frac{|k_\perp|^2}{2\kappa}$. We therefore find:

$$T = \frac{\pi^2}{9} \exp\left(-2\int_{x_i}^{x_f} \mathrm{Im}(k_x)\,dx\right) \tag{2.15}$$

$$= \frac{\pi^2}{9} \exp\left(-2\int_{x_i}^{x_f} \kappa\,dx\right) \exp\left(-|k_\perp|^2 \int_{x_i}^{x_f} \frac{dx}{\kappa}\right) \tag{2.16}$$

To keep the computation tractable, the tunneling path is computed for $k_\perp = 0$. Putting the expression into Eq. 2.8 and performing the $k_\perp$ integral, we have:

$$G_{\mathrm{BTBT}}(x_i) = |\nabla E_{\mathrm{V}}(x_i)|\,\frac{q}{36\hbar}\left(\int_{x_i}^{x_f}\frac{dx}{\kappa}\right)^{-1}\left[1-\exp\left(-k_{\mathrm{m}}^2\int_{x_i}^{x_f}\frac{dx}{\kappa}\right)\right]\exp\left(-2\int_{x_i}^{x_f}\kappa\,dx\right)\big(f_{\mathrm{c}}(E)-f_{\mathrm{v}}(E)\big) \tag{2.17}$$

with $E_{\mathrm{V}}(x_i) = E_{\mathrm{C}}(x_f) = E$ which is precisely the formula used in the commercial simulator Sdevice [57].

In this model, the non-locality is achieved by attributing a trajectory to the carriers within the classically forbidden region. Different schemes have been proposed such as the carrier following the gradient of the energy band [57], or using a predefined tunneling direction [59]. The problem with this approach is that the results depend on scheme employed [59] and it is not obvious which trajectory provides best approximation of this quantum phenomenon.

Pan's Model

Pan et al. [60, 61] generalized the results obtained by Kane for bulk semiconductors (i.e., two dimensional transverse k-space) to 2D and 1D cases (i.e., one dimensional and zero dimensional k-space, respectively). They obtain the following expression where d is the number of dimensions of the device:

$$J = \frac{q}{\pi\hbar}\int \frac{dk_\perp^{d-1}}{(2\pi)^{d-1}}\int dE\,T(E,k_\perp)\left[f_{\mathrm{c}}(E)-f_{\mathrm{v}}(E)\right] \tag{2.18}$$

$$= A(F)\int_0^{\Delta E} dE\exp(-\frac{B}{F})\left[f_{\mathrm{c}}(E)-f_{\mathrm{v}}(E)\right] \tag{2.19}$$

with a modified $B = B_{\mathrm{bulk}}\left(\frac{E_{\mathrm{g,QC}}}{E_{\mathrm{G}}}\right)^2$ parameter due to inclusion of non-parabolicity and higher-order band coupling effects. Moreover, the A parameter is calculated for each dimensionality by properly integrating along the transverse k-space:

$$A_{2D} = \frac{q\sqrt{m_{\mathrm{r}}E_{g,QC}}}{18\hbar^2}\left[\sqrt{\frac{q\pi F}{4B}} + 0.185\left(\frac{qF}{B}\right)^{1.2}\right] \tag{2.20}$$

$$A_{1D} = \frac{2q\pi^2}{9h} \tag{2.21}$$

Note that the model presented here is a local model, but the same modifications can be carried over to non-local models as well.

Carrillo-Nuñez Model

Carrillo-Nuñez et al. [62] also proposed a BTBT model based on multi-band envelope function formalism for bulk, 2D and 1D gases. Unlike the approach we will follow in this thesis, they utilize a Landauer-based formula (similar to Eq. 2.3) with a tranmission coefficient that is obtained from the coupling constant M_{CV}, which is equivalent to our Fermi's Golden Rule based expressions [62]. The overall BTBT current is then calculated using the Landauer formula (similar to Eq. 2.4). Note that this model considers the anisotropic coupling of the conduction and valence bands which introduces a dependence of the band coupling element on the wave-vector direction (more details will be provided in section 3.5.3).

2.3.2 Phonon-Assisted Band-to-Band Tunneling Models

Phonon-assisted tunneling occurs in indirect gap materials whose conduction band minima do not lie at $\vec{k} = 0$ such as Si or Ge. Keldysh [63] first proposed an expression describing the phonon-assisted BTBT by making use of the perturbation theory. Later on, one of the first non-local BTBT models has been proposed by Tanaka for phonon-assisted (and direct) tunneling [58]. In this model, the interband coupling due to electron-phonon interaction in deformable ion model is used to interband coupling elements of the Wannier formula. The wavefunctions are obtained patching the plane wave solutions (classically allowed region) with decaying components (classically forbidden region) using the WKB approximation.

Vandenberghe Model

A recent model that is applicable for non-uniform and 1D, 2D and 3D potential profiles considering quantization has been proposed by Vandenberghe et al. [64, 65, 66]. The approach makes use of the diagonal entries of the spectral functions that contain information about the energy spectrum and the envelope function information of the states and are defined as:

$$A_{c\alpha}(\vec{r}, \vec{r}, E) = 2\pi \sum_k \Psi^*_{k\alpha}(\vec{r})\delta(E - E_{k\alpha})\Psi_{k\alpha}(\vec{r}) \tag{2.22}$$

where $\Psi_{k\alpha}$ is the wavefunction, $E_{k\alpha}$ is energy eigenvalue corresponding to $\Psi_{k\alpha}$, k is the subband index (quantum number), α is the valley index.

The total phonon-assisted tunneling current is evaluated by first calculating the transmission coefficient over energy:

$$T_v^{\text{abs,em}}(E) = \Omega|M'_{k_0}|^2 \sum_{\alpha,\alpha'} \int A_{v\alpha'}(\vec{r},E) A_{c\alpha}(\vec{r},E \pm E_{\text{ph}}) d^3 r \tag{2.23}$$

where

$$\Omega|M'_{k_0}|^2 = \frac{D_{\text{ph}}^2 \hbar^2}{2\rho E_{\text{ph}}} \tag{2.24}$$

Here, D_{ph}, $E_{\text{ph}} = \hbar\omega_{k_0}$ is the deformation potential and the energy of the phonon corresponding to the crystal momentum offset of conduction and valence bands [64, 65, 66] considered respectively and ρ is the density of the material. Then, the phonon-assisted BTBT current is calculated as the summation of phonon emission and absorption terms:

$$\begin{aligned} J_{\text{ph}} = & \frac{2q}{\hbar} \int \frac{dE}{2\pi} \cdot \big(\{ f_{\text{v}}(E)[1 - f_{\text{c}}(E - \hbar\omega_{k_0})][\nu(\hbar\omega_{k_0}) + 1] \\ & - f_{\text{c}}(E - \hbar\omega_{k_0})[1 - f_{\text{v}}(E)]\nu(\hbar\omega_{k_0})\} T_v^{\text{em}}(E) \\ & + \{ f_{\text{v}}(E)[1 - f_{\text{c}}(E + \hbar\omega_{k_0})]\nu(\hbar\omega_{k_0}) \\ & - f_{\text{c}}(E + \hbar\omega_{k_0})[1 - f_{\text{v}}(E)][\nu(\hbar\omega_{k_0}) + 1]\} T_v^{\text{abs}}(E) \big) \end{aligned} \tag{2.25}$$

where $f_{c,v}$ are the Fermi-Dirac distributions of the conduction and valence bands, respectively, $\nu(\hbar\omega_{k_0})$ is the number of phonons calculated as evaluating the Bose-Einstein distribution function at $\hbar\omega_{k_0}$. In this thesis, we have utilized Vandenberghe model for our simulations which will be presented in the subsequent chapters.

3 Quantum Mechanical Models for Band-to-band Tunneling Devices

This chapter describes the numerical simulator built for the quantum mechanical simulation of tunnel FETs. Different variants of the code, their features and limitations are explained in detail.

At the early stages of the semiconductor industry, models used by the device engineers were mostly analytical derived from the underlying fundamental physical relations. With the increased computational power brought by computers during the 1960s [67], first numerical simulators emerged based on the drift-diffusion equations [68]. This framework served the industry very well during the ’happy scaling era’ [19] where quantum effects and off-equilibrium transport were mostly negligible and the semiclassical approach remained valid.

However, as the device dimension scaling continued, the semiclassical approach commonly employed by commercial TCAD tools started to breakdown at narrow dimensions (i.e., thin semiconductor films) and under high electric fields. To address this issue, the wave property of the electrons needed to be taken into account. Today, this is almost always done by solving directly the Schrödinger equation, which is the wave equation for electrons and holes. To obtain the correct electrostatics within the device, Schrödinger equation is coupled with the usual Poisson equation that governs the relationship between the potential profile and carrier concentrations. Methods such as tight-binding NEGF [14], DFT-NEGF [69], $k \cdot p$-NEGF [70], wavefunction formalism [71] and Wigner function formalism [72] are wide spread and popular frameworks for quantum transport.

Simultaneously with the necessity of quantum mechanical tools that go beyond the semiclassical approximation, the need for accurate band-to-band tunneling models is also amplified. This is justified by two pressing needs: the first one is that in ulta-scaled modern MOSFETs, band-to-band tunneling start to have a significant impact on the OFF state currents. The second one is the emergence of band-to-band tunneling devices such as TFETs as next generation low supply voltage switches due to incompressibility of the switching slope of MOSFETs.

3.1 Effective Mass Approximation: Different Boundary Conditions

While the full-quantum methods mentioned above are very accurate, they come at a costly price. The computational requirement for the simulation of a single device become prohibitive for parameter analyses etc. unless you have access to a high performance computer with 1000s of CPUs. A compromise can be achieved between accuracy and simulation time by solving the Schrödinger equation using the effective mass approximation, in which a parabolic dispersion relation is usually assumed. Quantum mechanical methods that are based on the single-band effective mass approximation can be considered as a 'sweet spot' in terms of accuracy and computational requirements (see Fig. 3.1).

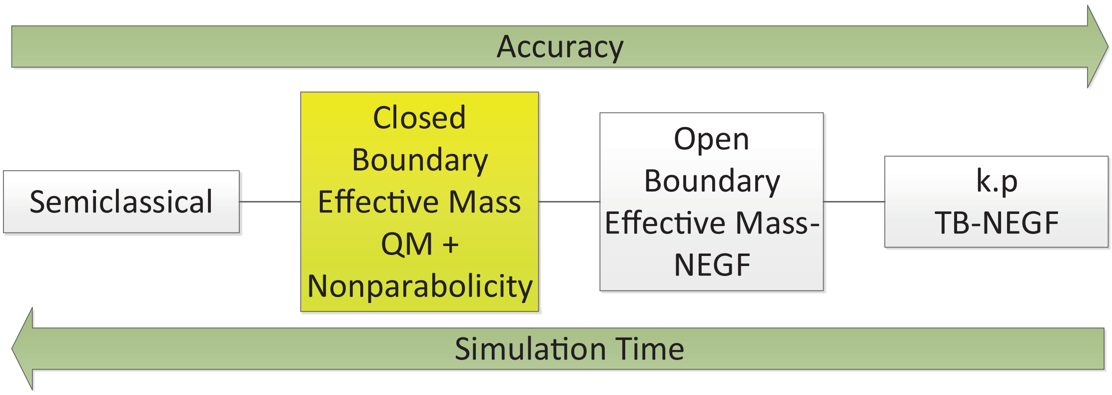


Figure 3.1 – The spectrum of computational load and accuracy of different simulation approaches. While TB-NEGF atomistic approach provides very good accuracy, it is also very costly in terms of memory and CPU power.

Another important question concerns the boundary conditions employed when simulating the device. Essentially, we can have open or closed boundaries. Non-equilibrium Green's function (NEGF) formalism [73] is an extremely popular approach that considers open boundaries for the Schrödinger equation. Another prominent approach is the wavefunction formalism (also known as quantum transmitting boundary method [74]). Here, the open boundary means that certain boundaries of the device (i.e., the contacts) are assumed to be connected to semi infinite leads that have constant potential along the wave propagation direction. Under this assumption, the wave function inside the device is patched up to the outside plane wave solutions. Employing open boundaries requires the solution of the system of equations for a *continuum* of energies of the incoming wave, which needs to be discretized and solved (preferably in parallel). This, however, generally results in a high computational burden stemming from this solution procedure that needs to be repeated for a large number of incoming wave energies. This concern is partially addressed by employing more complicated techniques such as mode space solutions [75] and using algorithms that allows for massive parallelism [76].

The approach that we will follow in this thesis, on the other hand, assumes closed boundaries.

This assumption results in a closed system, where the energies are now quantized. The eigenvalues of the Hamiltonian matrix now gives the quantized energies of the subbands and the eigenvectors are the wave functions [77]. Hence, the equation system is solved only once as a standard eigenvalue problem. It must be noted that in theory the transport in an electronic switch must be modeled with an open boundary model, since closed systems by definition cannot account for carrier transport. However, we can get a fairly good approximation of the open boundary results by employing artificially extended source and drain regions to imitate the semi-infinite nature of the contacts. Fig. 3.2 compares the charge distribution and the electrostatics of a nano-scale Si MOSFET simulated with open (using NEGF following [6]) and closed boundaries. Note the artificial bending of the bands and peaks of charge density around at the source and drain electrode regions, caused by the vanishing wavefunction at the boundary, but the effect decays very quickly and does not have any impact whatsoever on the electrostatics in the active device regions. Further comparison in terms of BTBT current will be shown for various instances in the following sections.

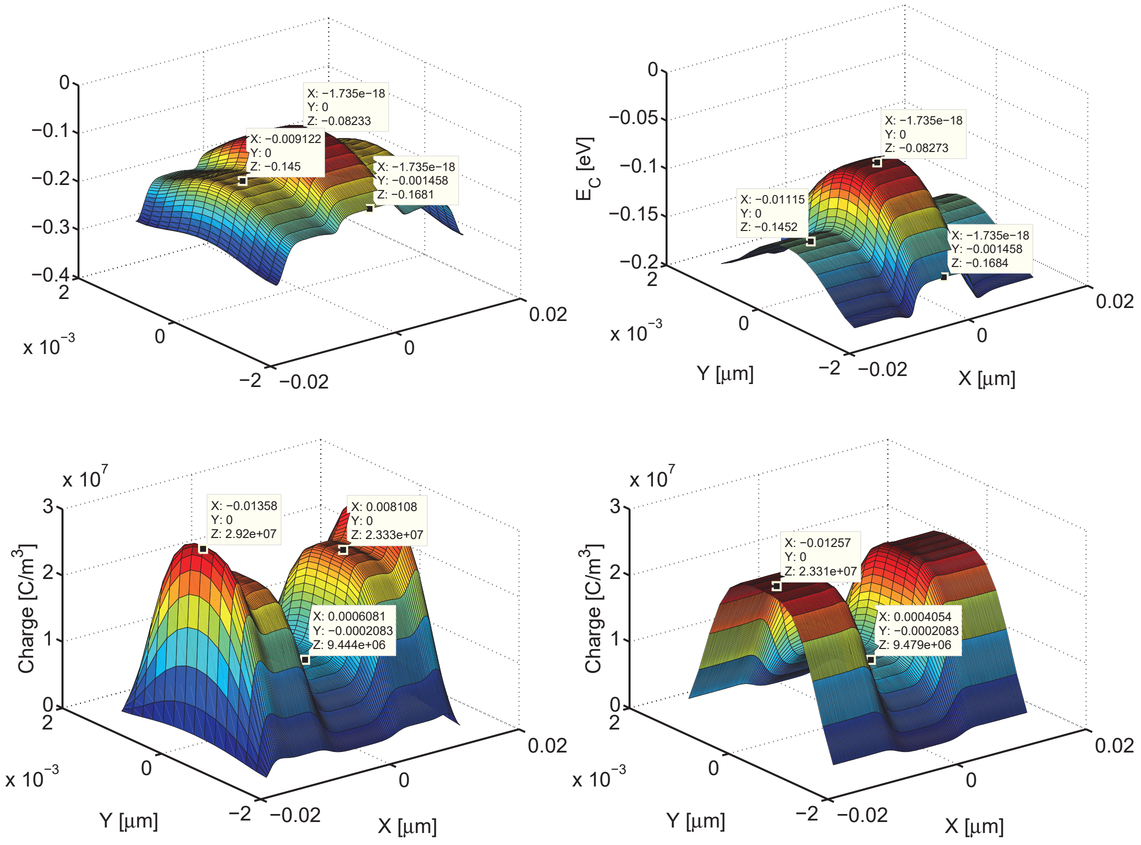


Figure 3.2 – Conduction band (upper) and Charge density profiles (bottom) for a Silicon DG-MOS. (Left) Closed boundary system (Right) Open boundary system implemented following [6]. For $V_{\mathrm{GS}} = 1\mathrm{V}$ and $V_{\mathrm{GS}} = 0\mathrm{V}$.

In this following sections, two different versions of the effective mass Schrödinger-Poisson solver will be described in detail. The 1D version solves the device electrostatics on a 1D line, thus assuming 1D quantization and 2D carrier gases. It is used in the thesis as a tool for extensive parameter space simulations and optimization of intrinsic device performance. It also incorporates non-parabolicity corrections which extend the region of validity of the effective mass approximation. The 2D version, on the other hand, is based on the parabolic bands and requires more computational power. It is able to capture 2D quantum effects that cannot be adequately captured by the commercially available TCAD simulators [57], however lacks the advanced non-parabolicity correction models used in the 1D simulator (due to the difficulties to be mentioned in the section 3.4). Due to this, it is not particularly suitable for simulations utilizing lower bandgap materials such as InAs which exhibit very strong non-parabolic effects around the conduction band minimum.

We note that phenomena such as trap-assisted tunneling [78] and band tails in the energy gap [39], that may have a significant impact on the experimental TFET devices are not included at present in the simulation model. Trap-assisted tunneling due to the traps on the semiconductor-oxide interface as well as on the hetero-junction interface are recently shown to be critical to achieve optimum device performance [79]. Band tails occur due to the dislocations found in the crystalline structure and doping [80].

3.2 1D Quantization Model

Our 1D implementation solves the effective mass Schrödinger and Poisson equation self-consistently in a one dimensional slice of semiconductor for both electrons and holes. A simulation flow is given in Fig. 3.3. The quantization direction is assumed to be the z-direction, with the transverse xy-plane assumed to be extending to infinity with a potential profile assumed to be constant along the xy-plane i.e., $V(\vec{R}) = V(z)$. The differential equations are discretized using the Finite Difference Method (FDM) which is a well-established method [81] that allows the treatment of non-homogenous systems (e.g., heterojunctions). Since we assume constant potential along the xy-plane, we can separate the 3D wavefunction $\Psi(\vec{R})$ as $\Psi(\vec{R}) = \psi(z)e^{i\vec{k}\cdot\vec{r}}$ where $\vec{r}$ is a vector in the xy-plane. Therefore, we are left with the solution of the envelope equation along the z-direction. In the quantization direction z, the envelope equation for the kth wavefuction residing in the band and valley α reads:

$$\left[\mp\frac{\hbar^2}{2}\nabla\cdot\frac{1}{m_{z,\alpha}(z)}\nabla + E_{(\mathrm{C,V})}(z)\right]\psi_{l\alpha}(z) = E_{l\alpha}\psi_{l\alpha}(z) \tag{3.1}$$

where $m_{z,\alpha}$ is the effective mass in the quantization direction, $E_{(\mathrm{C,V})}(z)$ are the conduction and valence band profiles respectively. The effective mass Schrödinger equation is solved for each electron valley, heavy hole and light hole bands. The wavefunctions $\psi_{l\alpha}(z)$ and subband

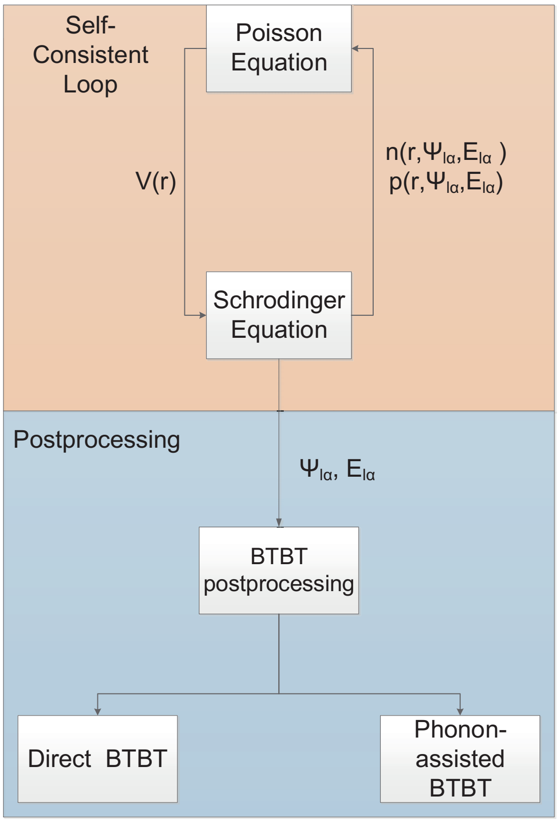


Figure 3.3 – Simulation flowchart for the self-consistent Schrödinger - Poisson loop and the BTBT post-processing step.

energies $E_{l\alpha}$ are then used to obtain the carrier concentrations as [77]:

$$
\begin{aligned}
n(z) &= \sum_{\alpha\in c}\sum_{l}\psi^*_{l\alpha}(z)\psi_{l\alpha}(z)n_{l\alpha}\\
p(z) &= \sum_{\alpha\in v}\sum_{l}\psi^*_{l\alpha}(z)\psi_{l\alpha}(z)p_{l\alpha}
\end{aligned}
\tag{3.2}
$$

where $n_{l\alpha}$ and $p_{l\alpha}$ terms are the electron and hole occupation terms of their respective lth subbands, which can be calculated as a zeroth order Fermi integral:

$$
\begin{aligned}
n_{l\alpha} &= \frac{g_\alpha m_{xy,\alpha}}{\pi\hbar^2}\int_{E_{l\alpha}}^{\infty}\frac{1}{1+e^{\frac{E-E_{\mathrm{Fn}}}{k_{\mathrm{B}}T}}}dE\\
&= \frac{g_\alpha m_{xy,\alpha}}{\pi\hbar^2}\mathscr{F}_0\left(\frac{E_{Fn}-E_{l\alpha}}{k_{\mathrm{B}}T}\right) = \frac{g_\alpha m_{xy,\alpha}k_{\mathrm{B}}T}{\pi\hbar^2}\ln\left(1+e^{\frac{E_{\mathrm{Fn}}-E_{l\alpha}}{k_{\mathrm{B}}T}}\right)\\
p_{l\alpha} &= \frac{g_\alpha m_{xy,\alpha}k_{\mathrm{B}}T}{\pi\hbar^2}\ln\left(1+e^{\frac{E_{l\alpha}-E_{\mathrm{Fp}}}{k_{\mathrm{B}}T}}\right)
\end{aligned}
\tag{3.3}
$$

In the above equation, $m_{xy,\alpha}$ is the density of states effective mass in the transverse direction (xy-plane) and g_α is the valley degeneracy factor. $m_{xy,\alpha}$ is related to the principal masses of the equi-energy ellipsoids with: $m_{xy,\alpha} = \sqrt{m_{x,\alpha} m_{y,\alpha}}$ [82]. Note that Eq. 3.3 assumes parabolic bands, which may not be suitable for semiconductors that exhibit strong non-parabolicity in their conduction band such as InAs. We will detail in Section 3.4 the non-parabolicity corrections to the DOS and wavefunctions which will have a considerable impact on the device characteristics. We also note that we do not consider the complex warped band structure of holes but utilize instead parabolic bands with hole density of state masses.

The calculated quantum mechanical charge densities are then fed into the right hand side of the Poisson equation, which reads:

$$\begin{aligned} -\nabla \cdot \epsilon(z) \nabla V(z) &= \rho(z) \\ &= -qn(z) + qp(z) - qN_\mathrm{A} + qN_\mathrm{D} \end{aligned} \tag{3.4}$$

where $\epsilon(z)$ is the dielectric permittivity of the material, $\rho(z)$ is the total charge density, N_A and N_D are acceptor and donor concentration respectively. $V(z)$ is then updated by solving 3.4 with the newly calculated charge density. This loop is continued until the convergence is reached. The solution of the Poisson equation is achieved through a non-linear Newton's method that has the property of global convergence [83] which was indeed observed to be almost independent on the initial guess in the simulations. The electrostatic potential is updated in each iteration through the following relation:

$$\vec{V}^{(n+1)} = \vec{V}^{(n)} + T_\mathrm{K} \overrightarrow{\delta_V} \tag{3.5}$$

where $\vec{V}^{(n)}$ and $\vec{V}^{(n+1)}$ are the n and $n+1$st iteration vectors containing the electrostatic potential values for each position in the grid, T_K is a damping parameter to improve convergence which is updated in each iteration according to the algorithm [83] and $\overrightarrow{\delta_V}$ is the update vector in the electrostatic potential. As usual in Newton's method, $\overrightarrow{\delta_V}$ is obtained by solving the Jacobian:

$$\overrightarrow{\delta_V} = -\mathbf{J}^{-1} \vec{H}^{(n)} = -\left(\mathbf{A} - \boldsymbol{\delta}_\mathrm{RHS}\right)^{-1} \left(\mathbf{A} \vec{V}^{(n)} - \overrightarrow{\mathrm{RHS}}\right) \tag{3.6}$$

here, $\mathbf{J}$ is the Jacobian matrix containing the first order partial derivatives of the error term $\left(-\nabla \cdot \epsilon(z) \nabla V(z) - \rho(z)\right)$ w.r.t. electrostatic potential in each grid position. $\mathbf{A}$ is the tridiagonal matrix for the Laplacian operator approximated in FDM, $\boldsymbol{\delta}_\mathrm{RHS}$ is the derivative matrix of the right-hand side of the Poisson's equation, $\overrightarrow{\mathrm{RHS}}$ is the right-hand side of Poisson equation. Finally $\vec{H}^{(n)}$ is the error term between the left- and right- hand sides of the Poisson equation. The diagonal of the $\boldsymbol{\delta}_\mathrm{RHS}$ matrix is calculated from Eq. 3.2 (the other terms are 0):

$$\begin{aligned} \mathrm{diag}\left(\boldsymbol{\delta}_\mathrm{RHS}\right) = \delta_\mathrm{RHS}(z) = \frac{\partial \overrightarrow{\mathrm{RHS}}}{\partial V} = &-\sum_{\alpha \in c} \sum_{l} \frac{q g_\alpha m_{xy,\alpha}}{\pi \hbar^2} \psi_{l\alpha}^{*}(z) \psi_{l\alpha}(z) \frac{1}{1 + e^{\frac{E_\mathrm{Fn} - E_{l\alpha}}{k_\mathrm{B} T}}} - \\ &\sum_{\alpha \in v} \sum_{l} \frac{q g_\alpha m_{xy,\alpha}}{\pi \hbar^2} \psi_{l\alpha}^{*}(z) \psi_{l\alpha}(z) \frac{1}{1 + e^{\frac{E_{l\alpha} - E_\mathrm{Fp}}{k_\mathrm{B} T}}} \end{aligned} \tag{3.7}$$

Since the current is computed in a post-processing step and not included in the self-consistent loop, it is implicitly assumed that electron and hole distributions are at equilibrium with the n-region and p-region contacts, respectively, which translates into $E_{\mathrm{Fn}} = -qV_{\mathrm{D}}$ and $E_{\mathrm{Fp}} = -qV_{\mathrm{S}}$. It is thus safely assumed that, due to the low current levels inside the device, tunneling generation does not appreciably alter the charge distribution in the low voltage regime.

For the numerical discretization, we define our device structure as a finite number of discrete points (say M) along the z axis, denoted as $z_1, z_2 \dots z_m \dots z_M$. We call the spacing between each sequential point as $\Delta_1, \Delta_2 \dots \Delta_m \dots \Delta_{M-1}$ and we also discretize $h(z)$ as $h_1, h_2 \dots h_m \dots h_M$. The Laplacian terms $\nabla \cdot h(z) \nabla$ discretized using the three-point stencil have the familiar tridiagonal shape for both Schrödinger and Poisson equations [77]:

$$\mathbf{A}_{\mathrm{m,m-1}} = \frac{h_{m-1} + h_m}{\Delta_{m-1}\left(\Delta_{m-1} + \Delta_m\right)}, \qquad \mathbf{A}_{\mathrm{m,m+1}} = \frac{h_{m+1} + h_m}{\Delta_m\left(\Delta_{m-1} + \Delta_m\right)}, \qquad \mathbf{A}_{\mathrm{m,m}} = -\mathbf{A}_{\mathrm{m,m-1}} - \mathbf{A}_{\mathrm{m,m+1}} \tag{3.8}$$

where we replace $h(z)$ with $\epsilon(z)$ and $\frac{1}{m_z(z)}$ for Poisson and Schrödinger equations, respectively.

Another peculiar aspect while coupling Schrödinger and Poisson equations is the inherent instability of the linear loop which severely degrades convergence [84]. So, we have utilized the non-linear predictor-corrector coupling scheme proposed by Trellakis et al. [84], which makes the charge density and its derivative dependent on the voltage step between the successive solutions of the electrostatic potential. Doing so provides significant stability as well as a dramatic increase in convergence rate [84]. While the scheme is extremely effective, it requires very little modification to the original self-consistent scheme. We simply replace the right-hand side by:

$$\begin{aligned} n_{l\alpha} &= \frac{g_\alpha m_{xy,\alpha} k_{\mathrm{B}} T}{\pi \hbar^2} \ln\left(1 + e^{\frac{E_{\mathrm{Fn}} - \Delta E - E_{l\alpha}}{k_{\mathrm{B}} T}}\right) \\ p_{l\alpha} &= \frac{g_\alpha m_{xy,\alpha} k_{\mathrm{B}} T}{\pi \hbar^2} \ln\left(1 + e^{\frac{E_{l\alpha} + \Delta E - E_{\mathrm{Fp}}}{k_{\mathrm{B}} T}}\right) \end{aligned} \tag{3.9}$$

where $\Delta E = E_{\mathrm{C}}^{(n+1)}(z) - E_{\mathrm{C}}^{(n)}(z) = E_{\mathrm{V}}^{(n+1)}(z) - E_{\mathrm{V}}^{(n)}(z)$. The derivative of the right-hand side is also modified as follows:

$$\begin{aligned} \delta_{\mathrm{RHS}}(z) = \frac{\partial \overrightarrow{\mathrm{RHS}}}{\partial V} &= -\sum_{\alpha \in c} \sum_{l} \frac{q g_\alpha m_{xy,\alpha}}{\pi \hbar^2} \psi_{l\alpha}^*(z) \psi_{l\alpha}(z) \frac{1}{1 + e^{\frac{E_{\mathrm{Fn}} - \Delta E - E_{l\alpha}}{k_{\mathrm{B}} T}}} - \\ &\sum_{\alpha \in v} \sum_{l} \frac{q g_\alpha m_{xy,\alpha}}{\pi \hbar^2} \psi_{l\alpha}^*(z) \psi_{l\alpha}(z) \frac{1}{1 + e^{\frac{E_{l\alpha} + \Delta E - E_{\mathrm{Fp}}}{k_{\mathrm{B}} T}}} \end{aligned} \tag{3.10}$$

Once the self-consistency is achieved, the band-to-band tunneling current is calculated as a post-processing step using the numerically calculated wavefunctions. We use the model proposed by Vandenberghe et al. (described in the section 2.3.2 in the previous chapter) for phonon-assisted BTBT. In the specific case of 1D quantization, the diagonal elements of the spectral functions (Eq. 2.22) are given as:

$$\begin{aligned} A_{c\alpha}(z,E) &= \frac{g_\alpha m_{xy,\alpha}}{\hbar^2}\sum_l \psi^*_{l\alpha}(z)\psi_{l\alpha}(z)\Theta(E-E_{l\alpha}) \\ A_{v\alpha}(z,E) &= \frac{g_\alpha m_{xy,\alpha}}{\hbar^2}\sum_l \psi^*_{l\alpha}(z)\psi_{l\alpha}(z)\Theta(E_{l\alpha}-E) \end{aligned} \tag{3.11}$$

where z is the quantization direction, $\Theta(E)$ is the step function given by the step-like 2D DOS. Since wavefunctions ψ are assumed $k_\perp$ independent, the spectral functions essentially reduce into summations of various subband 2D DOS functions weighted by the probability distribution of carriers $\psi^*_{l\alpha}(z)\psi_{l\alpha}(z) = \left|\psi_{l\alpha}(z)\right|^2$. Once the diagonal elements of the spectral functions are obtained, the phonon-assisted BTBT current is obtained by calculating Eqs. 2.23 - 2.25. For direct BTBT, we utilize Schenk or Bigelow models (will be detailed in the following sections). We also implemented a gate leakage model based on NEGF which is detailed in Appendix C.

Additionally, due to being a 1D model, only tunneling along the 1D slice is accounted for. The inability to incorporate 2D effects such as lateral tunneling, the short channel effects etc. limits the use of the 1D model to a few specific devices that are essentially 1D in nature. These devices include p-n tunneling diodes in the reverse bias region. The electron hole bilayer TFET can also be simulated using this model to estimate the vertical tunneling onset voltage and the ON current levels but one must be aware of the limitations of the model.

3.3 2D Quantization Model

The 2D version of the code follows the same program flow as in the 1D version (Fig. 3.3). The 2D code simulates a 2D cross section of the device which is assumed to be infinite in the width direction y. The difference compared to 1D case is that since the simulation domain is 2D, the potential is $V(x,y,z) = V(x,z)$ i.e., the potential is assumed uniform along only the y direction. Similarly, we separate the overall 3D wavefunction as $\Psi(\vec{R}) = \psi(\vec{r})e^{-i\vec{k}_y\cdot\vec{y}}$. The wavefunctions are found as the eigenvectors of the time independent, effective mass Schrödinger equation for both holes and electrons:

$$\left[\mp\nabla\cdot\frac{\hbar^2}{2}[m_\alpha]^{-1}\nabla + E_{(c,v)}(\vec{r})\right]\psi_{l\alpha}(\vec{r}) = E_{l\alpha}\psi_{l\alpha}(\vec{r}) \tag{3.12}$$

where α, l and $[m_\alpha]^{-1}$ indicates the valley index, ladder index and the inverse effective mass tensor, respectively. The anisotropy of the conduction band valleys is taken into account through the 2×2 inverse effective mass tensor [85, 86] with the elements to describe the equi-energy ellipsoids of the conduction band [87]. While our code can handle the cases where the device coordinate domain does not align with the principal axes of the ellipsoids (in which case the mass tensor contains non-diagonal terms), we utilize the method proposed in [87] to diagonalize the mass tensor. The quantum mechanical carrier distribution is then calculated as:

$$n_q(\vec{r}) = \sum_{\alpha\in c}\sum_{l} g_\alpha \sqrt{\frac{2m_{\alpha,\perp}k_{\mathrm{B}}T}{\pi^2\hbar^2}}\,\psi_{l\alpha}(\vec{r})\psi^*_{l\alpha}(\vec{r})\mathscr{F}_{-1/2}\left(\frac{E_{Fn}-\Delta E-E_{l\alpha}}{k_{\mathrm{B}}T}\right) \tag{3.13}$$

$$p_q(\vec{r}) = \sum_{\alpha\in v}\sum_{l} g_\alpha \sqrt{\frac{2m_{\alpha,\perp}k_{\mathrm{B}}T}{\pi^2\hbar^2}}\,\psi_{l\alpha}(\vec{r})\psi^*_{l\alpha}(\vec{r})\mathscr{F}_{-1/2}\left(\frac{E_{l\alpha}+\Delta E-E_{Fp}}{k_{\mathrm{B}}T}\right) \tag{3.14}$$

where $\mathscr{F}_{-1/2}$ is the complete Fermi-Dirac integral of order $^{-1}\!/_{2}$ and $m_{\alpha,\perp}$ is the effective mass in the y direction since the transverse k-space is now 1D. Similar to the 1D case, we define $\Delta E = E_{\mathrm{C}}^{(n+1)}(\vec{r}) - E_{\mathrm{C}}^{(n)}(\vec{r}) = E_{\mathrm{V}}^{(n+1)}(\vec{r}) - E_{\mathrm{V}}^{(n)}(\vec{r})$. We use the same globally convergent Newton's method algorithm for the solution of coupled partial differential equations [83]. For the solution of the nonlinear system of equations, the iteration scheme proposed by Trellakis et al. [84] has been used as done in the 1D version. We use COMSOL [88] to generate the device structure and export the first order triangular mesh as an ASCII file, which we import to MATLAB to assemble the mass and stiffness matrices (details below), compute the right hand sides [89] and then solve the linear system of equations. The δ_{RHS} term is given as [84]:

$$\begin{aligned}\delta_{\mathrm{RHS}}(\vec{r}) = \frac{\partial\overrightarrow{\mathrm{RHS}}}{\partial V} = &-\sum_{\alpha\in c}\sum_{l} g_\alpha q \sqrt{\frac{2m_{\alpha,\perp}}{\pi^2\hbar^2 k_{\mathrm{B}}T}}\,\psi_{l\alpha}(\vec{r})\psi^*_{l\alpha}(\vec{r})\mathscr{F}_{-3/2}\left(\frac{E_{Fn}-\Delta E-E_{l\alpha}}{k_{\mathrm{B}}T}\right) - \\ &\sum_{\alpha\in v}\sum_{l} g_\alpha q \sqrt{\frac{2m_{\alpha,\perp}}{\pi^2\hbar^2 k_{\mathrm{B}}T}}\,\psi_{l\alpha}(\vec{r})\psi^*_{l\alpha}(\vec{r})\mathscr{F}_{-3/2}\left(\frac{E_{l\alpha}+\Delta E-E_{Fp}}{k_{\mathrm{B}}T}\right)\end{aligned} \tag{3.15}$$

where $\mathscr{F}_{-3/2}$ is the complete Fermi-Dirac integral of order $^{-3}\!/_{2}$. Similar to the 1D version, we calculate BTBT as a post-processing step using the 2D versions of Vandenberghe (phonon-assisted BTBT) and the direct BTBT models that will be detailed in section 3.5, respectively. For the Vandenberghe method, the spectral functions in the case of 2D quantization are given as:

$$\begin{aligned} A_{c\alpha}(\vec{r},E) &= \frac{g_\alpha}{2\pi}\sqrt{\frac{m_{\perp,\alpha}}{2\pi^2\hbar^2}}\sum_{l}\psi^*_{l\alpha}(\vec{r})\psi_{l\alpha}(\vec{r})\frac{1}{\sqrt{E-E_{l\alpha}}}\Theta(E-E_{l\alpha}) \\ A_{v\alpha}(\vec{r},E) &= \frac{g_\alpha}{2\pi}\sqrt{\frac{m_{\perp,\alpha}}{2\pi^2\hbar^2}}\sum_{l}\psi^*_{l\alpha}(\vec{r})\psi_{l\alpha}(\vec{r})\frac{1}{\sqrt{E_{l\alpha}-E}}\Theta(E_{l\alpha}-E)\end{aligned} \tag{3.16}$$

similar to the 1D version, once the diagonal elements of the spectral functions are obtained, the phonon-assisted BTBT current is obtained by means of Eqs. 2.23 & 2.25.

We used finite element method (FEM) to discretize the partial differential equations, the reason being that FEM can naturally account for non-rectangular, complicated geometries which will be critical in our analyses in the following chapters. We provide here a brief outline of how PDEs can be formulated in FEM. We first define the Laplacian operator by $\Delta u = \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y}$.

A generic PDE is given in its strong form by [89]:

$$\begin{aligned}
&-\Delta u + cu = f && \text{in} \quad \Omega \\
&u = g_0 && \text{in} \quad \Gamma_D \\
&\hat{\mathbf{n}} \cdot \nabla u = g_1 && \text{in} \quad \Gamma_N
\end{aligned} \tag{3.17}$$

where Ω denotes the region where the function u is a free quantity, $\Gamma_{D,N}$ denotes the boundaries with Dirichlet and von Neumann conditions, respectively. We note the possibility of having multiple boundary regions with different Dirichlet and Neumann conditions. The weak formulation [90] of the same problem is then given by:

$$\int_\Omega \nabla u \cdot \nabla v + c \int_\Omega uv = \int_\Omega f v + \int_{\Gamma_N} g_1 v \tag{3.18}$$

where v is a testing function that is 0 at Γ_D. We then discretize the PDE domain with a finite number (say N) of non-overlapping triangles (other elements such as rectangles or a combination is certainly possible). We then approximate the integrals in the weak formulation with the help of a local test function that defined for each point i in the triangulation φ_i which has the property of being 1 at point i and linearly decreasing to 0 to each of the connected neighboring points. We also replace the continous function u with its discretized approximation using the basis function φ: $\tilde{u} = \sum_i \varphi_i u_i$ where $u_i = u$ at point i. The discretization of the weak formulation for each point i is then [89]:

$$\sum_{j\in\Omega} \left(\int_\Omega \nabla\varphi_j \cdot \nabla\varphi_i + c \int_\Omega \varphi_j \varphi_i \right) u_j = \underbrace{\int_\Omega f\varphi_i + \int_{\Gamma_N} g_1 \varphi_i}_{b_i} - \sum_{j\in\Gamma_D} \left(\int_\Omega \nabla\varphi_j \cdot \nabla\varphi_i + c \int_\Omega \varphi_j \varphi_i \right) g_0(j) \tag{3.19}$$

We note that the above set of equations can be written as a linear system of equations where each of the terms are computed per element and all the contributions from each of the elements for a point i are added up. This process is called *assembly* [91] where we define the $\mathbf{W}_{ij} = \int_\Omega \nabla\varphi_j \cdot \nabla\varphi_i$ as the stiffness matrix and $\mathbf{M}_{ij} = \int_\Omega \varphi_j \varphi_i$ as the mass matrix [89]. The overall system of equations is:

$$\sum_{j\in\Omega} \left(\mathbf{W}_{ij} + c\mathbf{M}_{ij} \right) u_j = b_i - \sum_{j\in\Gamma_D} \left(\mathbf{W}_{ij} + c\mathbf{M}_{ij} \right) g_0(j) \tag{3.20}$$

where $\mathbf{W}$ and $\mathbf{M}$ both form matrices of size $N \times N$. Note for the case of the Poisson equation, $\mathbf{M} = 0$. For the eigenvalue problem (Eq. 3.12 in the same form as Eq. 3.17), the discretized set of equations with eigenvalues λ are:

$$\sum_{j\in\Omega} \left(\mathbf{W}_{ij} + c\mathbf{M}_{ij} \right) u_j = \lambda \sum_{j\in\Omega} \mathbf{M}_{ij} u_j \tag{3.21}$$

$$\mathbf{W}\vec{u} = \lambda\mathbf{M}\vec{u} \tag{3.22}$$

which is a generalized eigenvalue equation. Note that in our case, c does not have to be a constant and therefore needs to be included inside the mass matrix integral, which we did not in our description for the sake of clarity. **W** and **M** are symmetric matrices, so the eigenvalues of this system are real, corresponding to the quantized energies of the Schrödinger equation.

It should also be noted that care must be taken with the quasi-bound states that are also the solutions of the closed-boundary Schrödinger equation. Typically, they occur in channel regions under relatively high biases and they are not connected to the source or drain electrodes because their energy is lower than the first mode entering/exiting into/from the device. In the steady-state operation, they are stationary hence do not contribute to the current [92], however they are occupied due to their subband energy levels and therefore they need to be included when calculating the charge distribution controlling the electrostatics. We overcome this problem by discarding the states which has less than 10% of their probability distribution (i.e, $|\psi|^2$) in the source/drain regions in the BTBT postprocessing step, or expressed mathematically (for electrons, dual expression applies for holes):

$$\frac{\int\limits_{\text{DRAIN}} |\psi|^2 \, d^2\vec{r}}{\int\limits_{\text{TOTAL}} |\psi|^2 \, d^2\vec{r}} < 0.1 \tag{3.23}$$

3.4 Non-parabolicity Corrections

While approximating the dispersion relation around the conduction minima with a parabolic approximation holds valid for relatively larger bandgap materials (e.g., Si) with bulk structures [93], the strength of non-parabolicity becomes non-negligible for low bandgap materials (such as InAs) and deep nano-scale devices due to energy quantization. In these cases, the energy dispersion relation starts deviating from the parabolic relation. While higher order models such as $k \cdot p$ can account naturally for these effects [94], non-parabolicity corrections can also be applied to the effective mass approximation to extend its range of validity [95]. Analytically, this band relationship can be written (for isotropic valleys) as [96]:

$$E(1 + \alpha_{\text{NP}}E) = \frac{\hbar^2 k^2}{2m^*} \tag{3.24}$$

where α_{NP} is the non-parabolicity coefficient (unit inverse energy $\text{eV}^{-1}, \text{J}^{-1}$). Qualitatively, a higher α_{NP} corresponds to lower quantization energies and higher density of states.

More specifically, non-parabolicity corrections affect the following four quantities in tunneling devices:

1. decrease of quantized energies compared to parabolic approximation, that lowers the

voltage required for subband alignment in structures such as the EHBTFET [97],

2. increase of the transverse DOS, hence the carrier occupation of the valley,
3. ‘heavier’ effective masses [98], that decrease the amount of WF overlap between the electron and hole states,
4. increase in (similar to the point 2) the ‘joint density of states’ available for tunneling that goes as $\frac{1}{\frac{dE_\mathrm{e}}{dk}+\frac{dE_\mathrm{h}}{dk}}$ (evaluated at the transverse wave-vector k_T which fulfills simultaneous conservation of total energy and transverse momentum, we will go into more details for this in section 3.5.4).

In this thesis, two different non-parabolicity correction methods have been implemented and utilized. Jin’s model applies corrections only to the subband energies, whereas Troger’s method is a more comprehensive one that also provides non-parabolicity corrections to the wavefunctions, which is critical when estimating the coupling between the conduction and valence band states for BTBT current estimation. We utilize Troger’s method to obtain the self-consistent solution (i.e., nonparabolic WF computation and charge calculation) and use Jin’s method in the postprocessing step to obtain the nonparabolic transverse dispersion relations required to compute the BTBT rate.

Jin’s Method

Jin proposed his correction expression as a phenomenological solution to the problem of obtaining a tractable solution to the non-parabolic subband energies [99]. The expression for subband minimum energy (i.e., $k = 0$) is given as:

$$E_\mu^\mathrm{NP} = U_\mu + \frac{-1+\sqrt{1+4\alpha_\mathrm{NP}\left(E_\mu^\mathrm{P}-U_\mu\right)}}{2\alpha} \tag{3.25}$$

where U_μ is the ’expected value of potential energy’ [99] given as $U_\mu = \int \left|\psi_\mu^\mathrm{P}\right|^2 U(z)dz$ and the E_μ^P is the eigenvalues of the parabolic effective mass Hamiltonian. This expected value calculation makes use of the fact that quantized states are normalized, therefore can be interpreted as probability density functions for the physical quantities related to the electron (in this case the potential energy ’felt’ by the electron).

The model also provides a formula for non-parabolic inverse dispersion relation, given as [99]:

$$k(E) = \hbar^{-1}\sqrt{2m_\perp\left[E-E_\mu^\mathrm{P}+\alpha_\mathrm{NP}\left(E-U_\mu\right)^2\right]} \tag{3.26}$$

One should note that the model is fairly efficient in terms of computational requirements, as it only requires calculation of the expected values and a few simple mathematical operations.

Since the model modifies the subband energy and provides an analytical E-k relation, it can cover the points 1, 2, and 4 of the list above. The modification of the wavefunctions cannot be estimated by this method which we will demonstrate below to be a critical shortcoming.

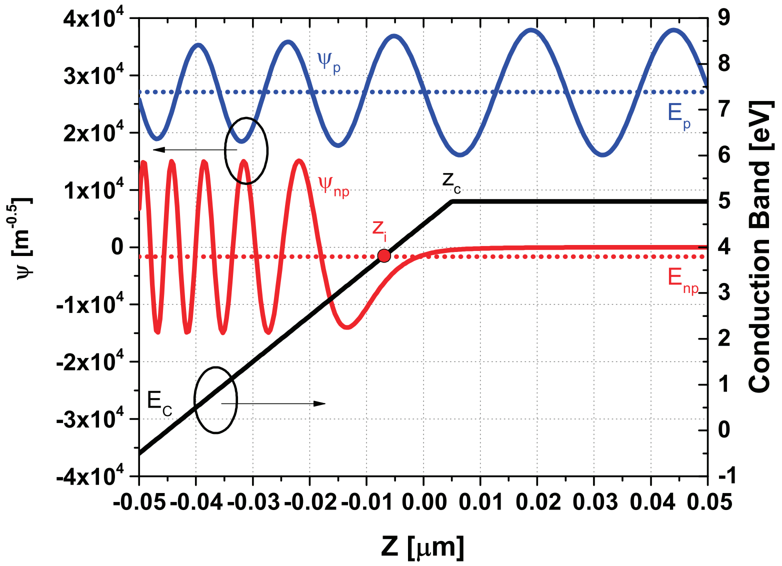


Figure 3.4 – Example of a case where non-parabolicity corrections dramatically alters the wavefunction in InAs. The black lines indicate the potential profile whereas blue (red) lines indicate the corresponding energy levels and wavefunctions for parabolic (non-parabolic) case. z_i indicates the classical turning point of the quasi-bound state, beyond which the non-parabolic wavefunction decays exponentially. Note that the potential profile is artificial for illustrative purposes and the non-parabolicity parameter α_{NP} was set to 5eV^{-1} in this simulation to amplify the separation of the quantized energy levels.

We should note here that while Jin's method is efficient and accurate enough for materials with relatively lower conduction band non-parabolicity, the fact that it only provides corrections to the quantized energies and not the wavefunctions prohibits its usage in materials such as InAs. To illustrate this, let us consider a toy example where we would like to find the quantized energies and wavefunctions of a triangular potential well from the dielectric/semiconductor interface up to a position z_{c} (see Fig. 3.4). We start from the quantized conduction band states obtained from the parabolic Hamiltonian which (in this example) predicts an unbound state that extends into the region $z > z_{\mathrm{c}}$ (ψ_p, E_p pair, blue curves Fig. 3.4). If we now calculate the non-parabolicity corrected energy levels, we find that the corresponding subband is actually quasi-bound (ψ_{np}, E_{np} pair, red curves Fig. 3.4). Jin's method, however, takes into account only the corrections to the subband energy (i.e., it only calculates E_{np}) whereas it does not change ψ_p. One can easily note the dramatic difference of the parabolic (does not decay in the forbidden region $z > z_i$) and actual non-parabolic wavefunctions (decays in $z > z_i$). In this case, using non-parabolicity corrections only to the subband energies would result in utilizing an unbounded state instead of a quasi-bound one and therefore result in a gross error when evaluating physical quantities related to the wavefunction (charge and current).

We detail another method that is able to calculate the non-parabolic wavefunctions in the following section.

Troger's Method

A more complete non-parabolicity correction method was proposed by Troger et al [100]. It is based on a basis transformation of the real space wavefunctions onto another space whose basis vectors are the eigenvectors of the parabolic kinetic energy operator:

$$\hat{T} = \frac{\hbar^2 k_\perp}{2m_\perp} + \frac{d}{dz}\frac{1}{m_z}\frac{d}{dz} \tag{3.27}$$

The basis vectors are found as the eigenvectors of the sparse matrix **T** [100]:

$$\mathbf{T} u_{m,k_\perp} = \Gamma_{m,k_\perp} u_{m,k_\perp} \tag{3.28}$$

where m is the eigenvector index, $u_{m,k_\perp}$ is the mth eigenvector and $\Gamma_{m,k_\perp}$ is its corresponding eigenvalue. Then, non-parabolicity is incorporated as a perturbation and the 'spectrum' of the non-parabolic wavefunction and the non-parabolic eigenvalue $E_{n,(k_\perp)}$ is found as the solution of another (dense) eigenvalue equation [100]:

$$\left(\mathbf{T}_{k_\perp} + \mathbf{V}\right) a_{n,k_\perp} = E_{n,(k_\perp)} a_{n,k_\perp} \tag{3.29}$$

where $a_{n,k_\perp}$ is the spectral amplitudes of u_m components in the non-parabolic wavefunction. Here the matrices **T** and **V** are the non-parabolic kinetic energy operator and the potential energy operator transformed into the new basis space, whose elements are given by [100]:

$$\mathbf{T_{n,m}} = \frac{2\Gamma_{n,k_\perp}}{1+\sqrt{1+4\alpha_{\mathrm{NP}}\Gamma_{n,k_\perp}}}\delta_{n,m} \tag{3.30}$$

$$\mathbf{V_{n,m}} = \int u_{m,k_\perp}(z) V(z) u_{n,k_\perp}(z)\, dz \tag{3.31}$$

Finally, the obtained spectral amplitudes are reverted by to the real space by:

$$\psi_{n,k_\perp}(z) = \sum_m a_{m,k_\perp} u_{m,k_\perp}(z) \tag{3.32}$$

In comparison to Jin's method, Troger's method does not provide a closed form expression for the inverse dispersion relation $k(E)$. Instead, the $k(E)$ is obtained by solving Eq. 3.29 for different $k_\perp$ as an input to obtain the $E_{n,k_\perp}(k)$ (see Fig. 3.5). Consequently, Troger's method is more costly in terms of computational requirements compared to Jin's method. Fig. 3.6 compares the dispersion relations obtained by Jin's and Troger's methods and a general

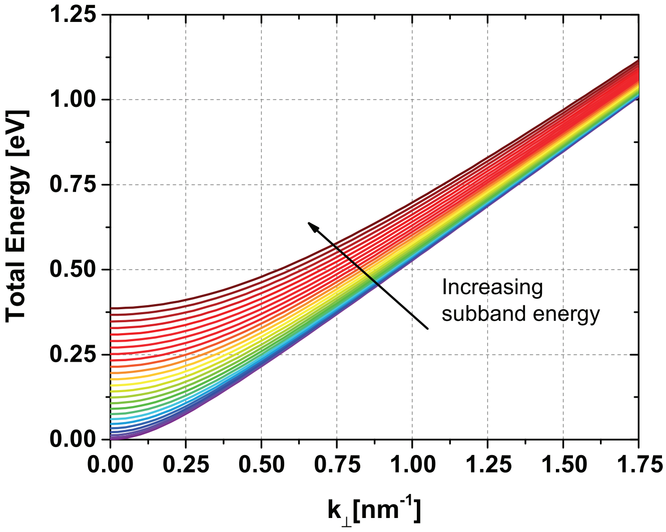


Figure 3.5 – E-k relations obtained by Troger's method for different subbands for an InAs infinite square well (35nm well width, $m_e^* = 0.025m_0$, $\alpha_{\rm NP} = 3.6{\rm eV}^{-1}$).

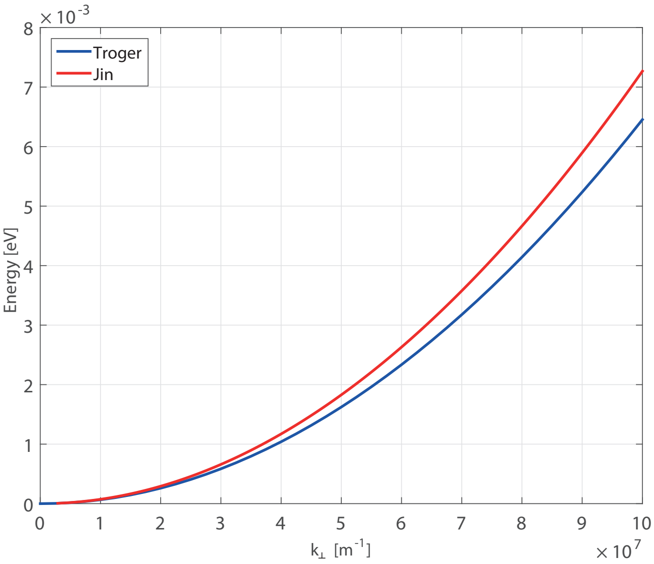


Figure 3.6 – Comparison of transverse dispersion relations obtained by Jin and Troger's methods for an $In_{0.53}Ga_{0.47}As$ quantum well (10nm well width $m_e^* = 0.042m_0$, $\alpha_{\rm NP} = 1.5{\rm eV}^{-1}$).

agreement between the two results is observed. Since the non-parabolic wavefunctions can also be obtained using Eq. 3.32, all aspects of non-parabolicity corrections listed above are covered by Troger's method.

One critical problem arises when applying Troger's method to 2D structures in the specific

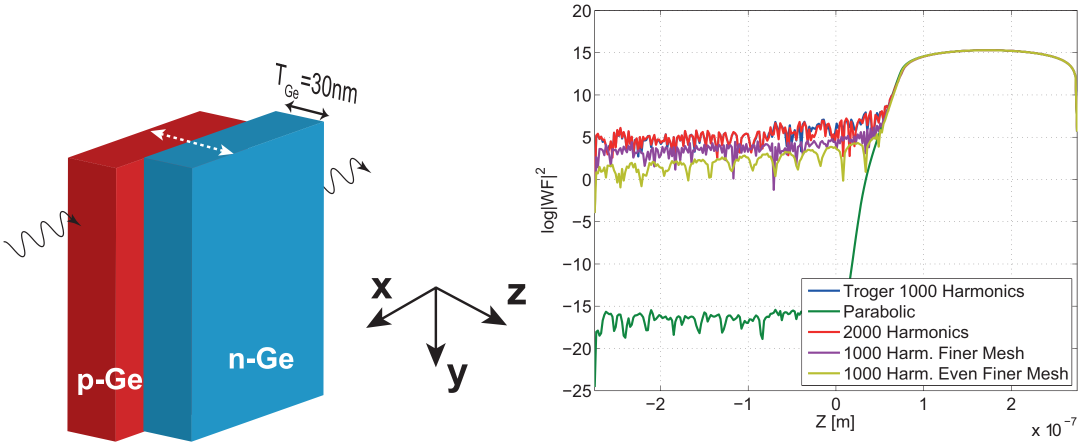


Figure 3.7 – (Left) Quantum well diode structure used for real-space wavefunction calculations in 2D. (Right) The construction of the real space WF using different number of harmonics and mesh sizes in the Troger's method. Note that the quantity plotted here is the 1D cut through a 2D quantum well diode on the left subfigure.

case of TFETs (or any other BTBT device). In 2D, the number of basis vectors required to obtain the real space wavefunctions becomes large. However, the amplitudes of higher order basis vectors which determine the function tails in the forbidden region decrease quickly. In the end, the tails of the constructed real space wavefunction is determined by a summation of many basis vectors with very low amplitude and a few lower order ones with very high amplitude i.e., it involves adding/subtracting very large numbers with very small ones. Since we use finite-precision arithmetic, this results in severe loss of significance [101]. Hence, the accuracy drops off rapidly and reconstruction of the tail of the real space wavefunction is not possible. To illustrate this point, in Fig. 3.7(Right) we plot the case where we deliberately set $\alpha_{NP} = 0$ and compare the 1D cuts of WFs obtained by direct solution of the parabolic Schrödinger equation (i.e., Eq. 3.1) and reproduced WFs using Troger's method in the 2D code. While an excellent agreement is obtained for the classically allowed region, it is evident that Troger's method quickly runs into a very high 'noise floor' in the exponentially decaying tail, thereby failing to reproduce the full WF.

3.5 Direct Tunneling in Quantized Gases

In the presence of quantization, different considerations need to be made depending on the orientation of the simulation domain and the quantization directions. In our specific case of closed boundary models, we always assume our simulation domain is bounded from all simulated directions (i.e., we assume 1D quantization for the 1D simulator and 2D quantization for the 2D one). For most of the figures in the thesis, the transverse direction is assumed to extend to infinity. However, we also propose an extension of the code in which the transverse direction has finite length in section 3.6.

For both 1D and 2D cases, we can write the following general expression of the direct BTBT current density using the Fermi's Golden Rule [37, 102, 15]:

$$J = \frac{4\pi q}{\hbar}\sum_{l\in\Gamma}\sum_{\alpha=\{\mathrm{HH},\mathrm{LH}\},l'}\left|M_{\mathrm{CV}}^{ll'\alpha}\right|^2 \mathrm{JDOS}(E_\mathrm{T})\left[f_\mathrm{c}(E_\mathrm{T})-f_\mathrm{v}(E_\mathrm{T})\right] \tag{3.33}$$

where the summation runs over all possible electron-hole subband pairs. Note that simultaneous conservation of total energy and transverse momentum (since direct BTBT is an elastic process) results in a single tunneling energy E_T and a single transverse k-vector magnitude k_T for each pair (see Fig. 3.8) rather than an integral over the entire tunneling energy window as in phonon-assisted BTBT. Eq. 3.33 reveals that BTBT has three main components that are beneficial to underline:

- **The interband coupling element** M_CV: the strength of coupling between electron and hole states. The main determinant for this quantity is the *overlap* of electron and hole wavefunctions which, in turn, is determined by material properties (bandgap and effective mass) as well as the device geometry and biasing. To a first order, this term is independent of the dimensionality of the carriers; although, indirectly, improvements in electrostatics provided by gate-all-around structures using nanowires etc. might have an impact,
- **The joint density of states** $\mathrm{JDOS}(E_\mathrm{T})$: the density of states of electrons and holes combined that obeys the energy conservation during the tunneling event. Note that the JDOS slightly differs from the usual DOS of a single carrier type that we use in charge density calculations. The expressions depend on the dimensionality of the carrier gases. More specifically, the dimensionality of the transverse k-space,
- **Fermi supply term** $f_\mathrm{c}(E_\mathrm{T}) - f_\mathrm{v}(E_\mathrm{T})$: This determines the occupancy of the electron and hole states and it is critical to consider this factor, as neglecting it will result in unnatural results such as having a non-zero current when $V_\mathrm{DS} = 0$ [16].

In the following subsections, we will dwell on two different methods implemented to calculate the interband coupling coefficient M_CV[1] in the presence of quantization.

---

[1] Note that we dropped the $ll'\alpha$ indices for clarity in the remainder of the thesis.

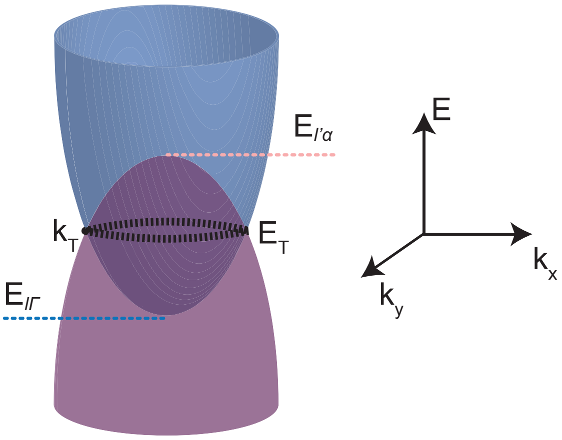


Figure 3.8 – Tunneling in the presence of 1D quantization (2D carrier gas) in which the conduction and valence bands are quantized into subbands along the tunneling direction with energies $E_z^{\mathrm{e(h)}}$. Tunneling then occurs at a single tunneling energy E_{T} that corresponds to the transverse wavevector k_{T} that is the intersection of the transverse dispersion relations for the electron and holes.

Another important point to note for tunneling in quantized gases is the gross asymmetry between the effective masses of real and imaginary subband dispersions of the holes. While quantization typically favors heavy hole subbands to contribute to the current, the imaginary dispersion relation calculations [103] show that the effective mass for the imaginary band is actually much closer to the light hole mass, which means that the interband coupling matrix magnitude under quantization is close to the one in the bulk case [60, 61]. An efficient method to account for this asymmetry will be given in section 3.7 for the 1D model.

Note that the models we describe here consider BTBT as a post-processing step after the electrostatics is determined by the self-consistent solutions of Schrödinger and Poisson equations, which is a reasonable assumption as long as the BTBT rate is not too high. This is different with respect to the case of NEGF or full quantum transport formalisms, where transport and electrostatics are inherently coupled.

3.5.1 Bigelow Model

In the case of quantization, additional considerations compared to the bulk case (e.g., model described in section 2.3.1) are necessary since the carrier gases are not three dimensional anymore. In the case of 2D gases (quantum wells), Bigelow et al. [102] developed a model that is based on Bardeen's Transfer Hamiltonian approach [104]. This approach exploits the fact that at either sides of a potential barrier, the Hamiltonian and the wavefunction of the state will tend to the 'isolated' cases [15]. Based on this, the coupling element M_{CV} is expressed as:

$$M_{\mathrm{CV}} = -\frac{\hbar^2}{2m}\int [\phi_{\mathrm{e}}\nabla\phi_{\mathrm{h}} - \phi_{\mathrm{h}}\nabla\phi_{\mathrm{e}}]\cdot d\vec{S} \tag{3.34}$$

where $\vec{S}$ is normal to the plane that separates the electron and hole systems and $\phi_{\mathrm{e,h}}$ are the electron and hole wavefunctions including the Bloch terms. Note that the necessity to separate the overall device into two regions makes this model problematic to use in TFET simulations, where the potential profile is often complicated and the procedure to divide the device into two regions is not straightforward.

Using $k \cdot p$, theory Bigelow et al. [102] estimated the coupling matrix element as:

$$M_{\mathrm{CV}}^2 = \hbar^2 \frac{E_{\mathrm{G}}}{2\overline{m}} \delta_{k_\perp, k'_\perp} \psi_{\mathrm{e}}^2 \psi_{\mathrm{h}}^2 C_0(\theta) \tag{3.35}$$

where $\overline{m} = \frac{2m_{\mathrm{e}}^* m_{\mathrm{h}}^*}{m_{\mathrm{e}}^* + m_{\mathrm{h}}^*}$. The product of slowly-varying envelope functions $|\psi_{\mathrm{e}}\psi_{\mathrm{h}}|$ is computed at its point of maximum which is a somewhat arbitrary choice. An important difference compared to the bulk case is that inclusion of the $C_0(\theta)$ which accounts for the polarization dependence of the coupling element in the presence of quantization [105] that we will discuss later.

Using parabolic bands, the two dimensional joint density of states (JDOS) that is available for tunneling that preserves the tunneling energy E_{T} is given by[2] [106]:

$$\mathrm{JDOS}_{\mathrm{2D}}(E_{\mathrm{T}}) = \frac{\overline{m}}{4\pi\hbar^2} \Theta(E_{l'\alpha} - E_{l\Gamma}) \quad \text{with} \quad E_{\mathrm{T}} = E_{l\Gamma} + \frac{\hbar^2 k_\perp^2}{2m_{\mathrm{e}}} = E_{l'\alpha} - \frac{\hbar^2 k_\perp^2}{2m_{\mathrm{h}}} \tag{3.36}$$

so the overall expression using Bigelow's coupling element for the 1D model is:

$$I_{\mathrm{Bigelow}}^{\mathrm{1D}} = \frac{qE_{\mathrm{G}} L_G W}{2\hbar} \sum_{l} \sum_{\alpha, l' \in v} \psi_{l\Gamma}^2 \psi_{l'\alpha}^2 C_{0,\alpha}(\theta) \left(f_v(E_{\mathrm{T}}) - f_c(E_{\mathrm{T}}) \right) \Theta(E_{l'\alpha} - E_{l\Gamma}) \tag{3.37}$$

where L_{G}, W are the gate length (along direction x) and device width (along the direction y) respectively. Note that the superscript 1D in Eq. 3.37 denotes the 1D model and not the 1D current density.

The original Bigelow method was a 1D model. Here we propose a method to extend it to 2D. The idea is to find the optimum contour line C between the electron and hole regions such that it maximizes the overlap line integral $\mathrm{g}(C')$ of the electron and hole wavefunctions:

$$C = \operatorname{argmax}\left(\mathrm{g}(C')\right) = \operatorname{argmax}\left(\left(\int_{C'} \psi_{l'\alpha} \psi_{l\Gamma} \, dr \right)^2 \right) \tag{3.38}$$

---

[2] Strictly speaking, JDOS is a function of both the electron and hole subband energies $E_{\mathrm{l\Gamma}}$ and $E_{\mathrm{l'}\alpha}$and not E_{T}. However, we prefer to use $\mathrm{JDOS}\left(\mathrm{E_T}\right)$ since it better describes the physical picture.

So the coupling constant in 2D is then given as:

$$M_{\mathrm{CV}}^2 = \hbar^2 \frac{E_{\mathrm{G}}}{2\overline{m}} \delta_{k_\perp, k'_\perp} \left(\int_C \psi_{l'\alpha} \psi_{l\Gamma} dr \right)^2 C_0(\theta) \tag{3.39}$$

The JDOS of electrons and holes combined available for tunneling is given by (for parabolic bands) [107]:

$$\mathrm{JDOS}_{\mathrm{1D}}(E_{\mathrm{T}}) = \frac{1}{2\pi} \sqrt{\frac{\overline{m}}{\hbar^2 (E_{l'\alpha} - E_{l\Gamma})}} \Theta(E_{l'\alpha} - E_{l\Gamma}) \quad \text{with} \quad E_{\mathrm{T}} = E_{l\Gamma} + \frac{\hbar^2 k_\perp^2}{2m_{\mathrm{e}}} = E_{l'\alpha} - \frac{\hbar^2 k_\perp^2}{2m_{\mathrm{h}}} \tag{3.40}$$

Combining everything together, the 2D tunneling current expression is given by:

$$I_{\mathrm{Bigelow}}^{\mathrm{2D}} = \sum_{l} \sum_{\alpha, l' \in v} \frac{W q E_{\mathrm{G}}}{\sqrt{\overline{m}(E_{l'\alpha} - E_{l\Gamma})}} \left(\int_C \psi_{l'\alpha} \psi_{l\Gamma} dr \right)^2 C_0(\theta) (f_{\mathrm{v}}(E_{\mathrm{T}}) - f_{\mathrm{c}}(E_{\mathrm{T}})) \Theta(E_{l'\alpha} - E_{l\Gamma}) \tag{3.41}$$

where W, E_{G}, $\overline{m}$, $E_{l\Gamma}$ are the device width, the bandgap of the semiconductor, the reduced effective mass and the lth Γ electron subband energy, respectively. E_{T} is the tunneling energy that conserves both energy and momentum.

One very big problem about the Transfer Hamiltonian method in 2D is the non-trivial task of separating the device into electron and hole regions for the devices such as the EHBTFET which exhibit very complex potential profile, carrier density profiles and wavefunction shape being dependent on the bias applied to the two gates. Other devices, such as line tunneling TFETs which make use of the Van der Waals gap created by stacking two 2D materials [108], could be more ideal for this model where the electron and hole regions are essentially separated by the hetero-interface due to band offset between two materials.

Due to this reason, we have implemented another model that does not require this partitioning which we will detail in the following section.

3.5.2 Schenk Model

To overcome the limitation of the Kane's model described in section 2.3.1, Schenk proposed a more generalized version [109], which could be utilized in non-homogeneous and strong fields. The model estimates the transition probability using Fermi's Golden Rule making use of Kane's two-band model. The model treats the electrostatic potential as a perturbation that couples the conduction and valence band states.

In order to simulate quantized gases, we do the following: we take the coupling element M_{CV}

calculated by Schenk, include the polarization dependence term to account for tunneling in quantized gases and calculate the tunneling current using the general expression given in Eq. 3.33.

In his paper [109], Schenk calculates the coupling element as follows:

$$M_{\mathrm{CV}} = q x_{\mathrm{CV}}(0)\, \delta_{k_\perp, k'_\perp} \int \psi^*_{l\Gamma}(z) |F(z)| \psi_{l'\alpha}(z) dz \tag{3.42}$$

where $x_{\mathrm{CV}}(0) = \sqrt{\frac{\hbar^2}{4\overline{m} E_G^\Gamma}}$ is the momentum element of the Bloch functions given by Kane's two band model [109] and $F(z)$ is the electric field.

In the case of quantization, to include the directionality dependence of the matrix element that we mentioned earlier for Bigelow model and will discuss in detail in section 3.5.3, we incorporate the C_0 term into Eq. 3.42 as [7]:

$$M_{\mathrm{CV}} = q x_{\mathrm{CV}}(0)\, \delta_{k_\perp, k'_\perp} \int \psi^*_{l\Gamma}(z) \sqrt{C_0(\theta)}\, |F(z)| \psi_{l'\alpha}(z) dz \tag{3.43}$$

The overall current in 1D, similar to Bigelow's method, is then given by [7]:

$$I^{\mathrm{1D}}_{\mathrm{Schenk}} = \frac{4\pi W L_G q}{\hbar} \sum_l \sum_{\alpha=\{\mathrm{HH},\mathrm{LH}\},l'} |M_{\mathrm{CV}}|^2 \frac{\overline{m}}{4\pi\hbar^2} \Theta(E_{l'\alpha} - E_{l\Gamma}) \left(f_{\mathrm{c}}(E_{\mathrm{T}}) - f_{\mathrm{v}}(E_{\mathrm{T}})\right) \tag{3.44}$$

An important difference of the Schenk model compared to the approach of Bigelow model is that it can be relatively easily extended to cover the cases where the electric field is non-uniform in more than one dimension (i.e., planar TFETs [7]) or quantization is present in more than one dimension (i.e., nanowire TFETs [62]). The main idea for extension to 2D (and the same goes even for 3D as well) is to incorporate the direction dependence of the coupling element into the electric field factor in Schenk's M_{CV} and change the JDOS expression to the one for 1D transverse k-space (Eq. 3.40). The general expression for the current in 2D is given as:

$$I^{\mathrm{2D}}_{\mathrm{Schenk}} = \frac{4\pi q W}{\hbar} \sum_l \sum_{\alpha, l' \in v} |M_{\mathrm{CV}}|^2 \frac{1}{2\pi} \sqrt{\frac{\overline{m}}{\hbar^2 (E_{l'\alpha} - E_{l\Gamma})}}\, \Theta(E_{l'\alpha} - E_{l\Gamma}) \left(f_c(E_T) - f_v(E_T)\right) \tag{3.45}$$

where the coupling element is defined as:

$$M_{\mathrm{CV}} = q x_{\mathrm{CV}}(0) \delta_{k_\perp, k'_\perp} \int \psi^*_{l'\alpha}(\vec{r})\, |F(\vec{r})|\, \psi_{l\Gamma}(\vec{r}) d^2\vec{r} \tag{3.46}$$

Here $|F(\vec{r})| = \sqrt{F_z(\vec{r})^2 C_{0z}(\theta,\phi) + F_x(\vec{r})^2 C_{0x}(\theta,\phi)}$ is the modified magnitude of the electric field where $F_{x,z}(\vec{r})$ are the x, z components of the electric field respectively. Since the integral in Eq. 3.43 is an area integral over the entire device, there is no need to partition the device into two regions unlike Bigelow model. Moreover, compared to Bigelow's model (where 1D definition of C_0 i.e., $C_0(\theta)$ was used in both 1D and 2D versions), 2D form factors $C_{0x(z)}$ can be easily incorporated into Schenk's coupling element.

Comparing against the model proposed by Carrillo-Nuñez et al. [62], for the 2D and 1D gases, the polarization dependence (to be explained in detail in section 3.5.3) of the coupling elements are treated exactly the same way as our model (i.e., C_0 terms used to 'weigh' the electric field components in each direction), with the exception of the determination of the ϕ angle. We make use of the concept of directional energies (the expected value of the kinetic energy along the principal axes of simulation domain), whereas Carrillo-Nuñez et al. fixes the angle at $\frac{\pi}{4}$ and takes the average with respect to ϕ. Let us now give some detail about the C_0 form factors found in both models.

3.5.3 Polarization Dependent Coupling: C_0 Factor

For bulk structures the interband coupling element M_{CV} is isotropic [110]. Here we briefly prove this starting from the atomic-like orbital elements (see Appendix B, Eqs. B.2 - B.8 for the definitions). Suppose that we would like to calculate coupling along a certain direction. In the case of bulk, the k wavevector of the electron will not have any constraints. Therefore the average of all possible directions needs to be taken. Say we calculate the coupling between conduction band and heavy hole band in the x direction which is given by $\left|\vec{x}\cdot\vec{p}_{\text{cv}}\right|^2 = \left\langle\left|\vec{x}\cdot M^2_{\text{c-hh}}\right|\right\rangle$ [110]. In spherical coordinates, we take the average on the surface of a sphere with radius r for the calculation of the average:

$$\left\langle\left|\vec{x}\cdot M^2_{\text{c-hh}}\right|\right\rangle = \frac{1}{4\pi r^2}\int_0^{\pi}\int_0^{2\pi} \left|\vec{x}\cdot M_{\text{c-hh}}\right|^2 r^2 \sin\theta\, d\phi\, d\theta \tag{3.47}$$

Since the spin is conserved in direct BTBT, the coupling between spin up and spin down states will be zero. So in total, for a given spin we need to calculate only one single coupling element: $\left|\langle iS\uparrow|p_x|\frac{3}{2},\frac{3}{2}\rangle\right|$ or $\left|\langle iS\downarrow|p_x|\frac{3}{2},-\frac{3}{2}\rangle\right|$ (for heavy holes only) depending on which spin we take (the result will be exactly the same no matter which one is chosen). It should be noted that only the $|X\rangle$ parts will be non-zero due to the vanishing p_x elements caused by oddness of p-like states. Overall, we would have:

$$\left\langle|\vec{x}\cdot M^2_{\text{c-hh}}|\right\rangle = \frac{1}{4\pi r^2}\int_0^{\pi}\int_0^{2\pi}\left|\langle iS\uparrow|p_x|\frac{3}{2},\frac{3}{2}\rangle\right|^2 r^2 \sin\theta\, d\phi\, d\theta \tag{3.48}$$

$$\left|\langle iS\uparrow|p_x|\frac{3}{2},\frac{3}{2}\rangle\right| = \left|\langle iS|p_x|X\rangle\right|^2 \frac{\left(\cos^2\theta\cos^2\phi+\sin^2\phi\right)}{2} \tag{3.49}$$

the $\left|\langle iS|p_m|I\rangle\right| \equiv P_x$ (where $m = \{x, y, z\}$) term is related to Kane parameter P through $P =$

$P_x\frac{\hbar}{m_0}$ [110]. Now we perform the integral:

$$\begin{aligned}\left\langle |\vec{x}\cdot M^2_{\text{c−hh}}|\right\rangle &= \frac{1}{4\pi r^2}\left|\langle iS|\,p_x\,|X\rangle\right|^2 \int_0^{\pi}\int_0^{2\pi} \frac{\left(\cos^2\theta\cos^2\phi+\sin^2\phi\right)}{2} r^2\sin\theta\, d\phi\, d\theta \\ &= \frac{1}{3}P_x^2 \equiv M_b^2\end{aligned} \tag{3.50}$$

M_b is defined as the bulk momentum matrix element per spin. Moreover, both heavy-hole and light-hole valleys have exactly the same contribution as well as being isotropic ($\left|\vec{e}\cdot M^2_{cv}\right|$ is equal for all $\vec{e}$). We now repeat the same for the directions y and z, for both heavy and light holes. We have the following:

$$\begin{aligned}\left\langle |\vec{y}\cdot M^2_{\text{c−hh}}|\right\rangle &= \frac{1}{4\pi r^2}\int_0^{\pi}\int_0^{2\pi} \left|\langle iS|\,p_y\,|Y\rangle\right|^2 \frac{\left(\cos^2\theta\sin^2\phi+\sin^2\phi\right)}{2} r^2\sin\theta\, d\phi\, d\theta \\ &= \frac{1}{3}P_x^2 \equiv M_b^2\end{aligned} \tag{3.51}$$

$$\begin{aligned}\left\langle |\vec{z}\cdot M^2_{\text{c−hh}}|\right\rangle &= \frac{1}{4\pi r^2}\int_0^{\pi}\int_0^{2\pi} \left|\langle iS|\,p_z\,|Z\rangle\right|^2 \frac{\left(\sin^2\theta\right)}{2} r^2\sin\theta\, d\phi\, d\theta \\ &= \frac{1}{3}P_x^2 \equiv M_b^2\end{aligned} \tag{3.52}$$

$$\left\langle |\vec{y}\cdot M^2_{\text{c−lh}}|\right\rangle = \tag{3.53}$$

$$\frac{1}{4\pi r^2}\int_0^{\pi}\int_0^{2\pi} \left|\langle iS|\,p_y\,|Y\rangle\right|^2 \left(\frac{\left(\cos^2\theta\sin^2\phi+\cos^2\phi\right)}{6}+\frac{2\sin^2\theta\sin^2\phi}{3}\right) r^2\sin\theta\, d\phi\, d\theta \tag{3.54}$$

$$= \frac{1}{3}P_x^2 \equiv M_b^2$$

$$\begin{aligned}\left\langle |\vec{z}\cdot M^2_{\text{c−lh}}|\right\rangle &= \frac{1}{4\pi r^2}\int_0^{\pi}\int_0^{2\pi} \left|\langle iS|\,p_z\,|Z\rangle\right|^2 \left(\frac{\sin^2\theta}{6}+\frac{2\cos^2\theta}{3}\right) r^2\sin\theta\, d\phi\, d\theta \\ &= \frac{1}{3}P_x^2 \equiv M_b^2\end{aligned} \tag{3.55}$$

The observation that interband coupling depends on the direction of the k-vector of the wavefunction under quantization was first made by Yamanishi and Suemune [105] in quantum well structures. This phenomenon comes from the symmetry properties of the p-like orbital functions with respect to direction and due to the fact that the averaging process to estimate the coupling element magnitude is now done in two dimensions. In such a case the dipole moment elements between conduction and valence band states depend on the wavevector direction of the state [111], denoted by the angle θ. Here, we consider the two quantization cases: quantization in one and two dimensions. The main problem to be solved is to determine the direction of the k-vector of the wavefunction, which modulates the strength of coupling.

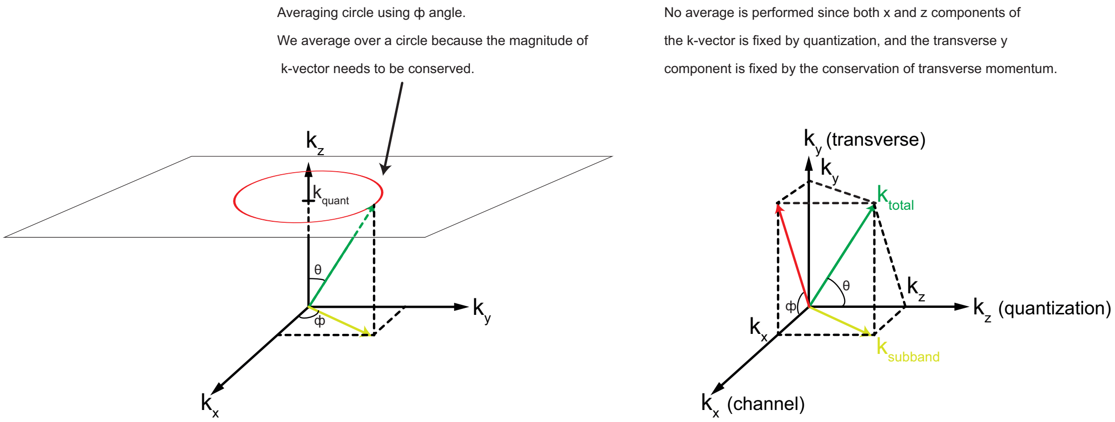


Figure 3.9 – (Left) Depiction of averaging for 1D definition of C_0. (Right) Visualization of 2D quantization and angle definitions.

Using a spherical angle notation, the direction of the k-vector is determined by the angles θ and ϕ. For 1D quantization (2D carrier gas), the transverse k-space is two dimensional on which we take the average (by sweeping the angle ϕ, see Fig. 3.9(Left)) for a transverse k-vector of a given magnitude that is determined by the conservation of energy and momentum between the electron and hole states. It should be noted that only the direction and not the magnitude of the k-vector matters. Since we perform the averaging along the angle ϕ, the form factors we obtain do not depend on the angle ϕ. θ, on the other hand, can be estimated as the ratio between the kinetic energy component along the tunneling direction z (E_z) and the transverse kinetic energy ($E_\perp$) [102]:

$$\cos^2\theta = \frac{k_z^2}{k_\perp^2 + k_z^2} = \frac{E_z}{E_\perp + E_z} \tag{3.56}$$

where $E_{z,(\perp)}$ is an effective kinetic energy along the quantization (transverse) direction. $C_0(\theta)$ for quantization along the z direction is then derived as:

$$\begin{aligned} C_0(\theta) = \frac{\langle|\vec{z}\cdot M^2_{\text{c-hh}}|\rangle}{M_b^2} &= \frac{1}{2\pi r M_b^2}\int_0^{2\pi} \left|\langle iS|\, p_z\,|Z\rangle\right|^2 \frac{(\sin^2\theta)}{2} r\,d\phi \\ &= \frac{\sin^2\theta}{2}\frac{P_x^2}{M_b^2} = \frac{3\sin^2\theta}{2} \end{aligned} \tag{3.57}$$

$$\begin{aligned} \langle|\vec{z}\cdot M^2_{\text{c-lh}}|\rangle &= \frac{1}{2\pi r}\int_0^{2\pi} \left|\langle iS|\, p_z\,|Z\rangle\right|^2 \left(\frac{\sin^2\theta}{6} + \frac{2\cos^2\theta}{3}\right) r\,d\phi \\ &= \left(\frac{\sin^2\theta}{6} + \frac{2\cos^2\theta}{3}\right)\frac{P_x^2}{M_b^2} = \left(\frac{\sin^2\theta}{2} + 2\cos^2\theta\right) \end{aligned} \tag{3.58}$$

therefore, the interband coupling element M_{CV} gains a polarization dependence with the following form factors in the case of electric field aligned with the quantization direction on the z axis:

$$C_0(\theta) = \begin{cases} 2\cos^2\theta + \frac{\sin^2\theta}{2} & \text{for LH} \\ \frac{3\sin^2\theta}{2} & \text{for HH} \end{cases} \tag{3.59}$$

Note that the expressions above correspond to the TM mode in a quantum well laser [105] where the quantization direction is aligned with the electric field on the z direction. Different expressions apply for the cases where the two directions are perpendicular (i.e., TE) mode.

In two dimensions, the situation becomes more complicated since we do not perform any averaging and we need to calculate two form factors in each direction of the simulation domain (C_{0x} and C_{0z}). To calculate $C_{0x(z)}\left(\theta,\phi\right)$, an estimation of the wavevector direction is necessary for each subband, but this time it will depend on the two angles θ and ϕ. The two k-vector components (along x and z directions) are estimated from the wavefunctions that are obtained as the solution of the Schrödinger equation, whereas the transverse component is determined again by the simultaneous convservation of total energy and momentum. The determination of these angles is done by finding the expectation value of the so-called estimated *(directional) kinetic energies* E_X, E_Z, obtained by applying the quantum mechanical (partial) kinetic energy operator on the wavefunctions along the principal axes:

$$E_{X(Z)} = -\frac{\hbar^2}{2}\frac{1}{m^*_{x(z)}}\left\langle\psi\right|\frac{\partial^2}{\partial x^2(z^2)}\left|\psi\right\rangle \tag{3.60}$$

We note that partial derivative is not a good quantum operator (i.e., does not commute with the Hamiltonian), so it does not exactly correspond to the actual kinetic energy along that direction, unless the wavefunction can be separated into pure components along the x and z directions, as in the case of a 2D infinite well. Having approximately separated the kinetic energy into two directions, the direction of the wavevector is estimated using:

$$\cos^2\theta = \frac{E_Z}{E_X+E_Y+E_Z} \tag{3.61}$$

$$\cos^2\phi = \frac{E_X}{E_X+E_Y} \tag{3.62}$$

where E_Y (kinetic energy component extending out of the simulation plane) is $E_Y = E_T - E_{l\Gamma}$ and ultimately, the $C_0\left(\theta,\phi\right)$ is given as [111]:

$$C_{0x}\left(\theta,\phi\right) = \begin{cases} 2\sin^2\theta\cos^2\phi + \frac{\cos^2\theta\cos^2\phi+\sin^2\phi}{2} & \text{for LH} \\ \frac{3(\cos^2\theta\cos^2\phi+\sin^2\phi)}{2} & \text{for HH} \end{cases} \tag{3.63}$$

$$C_{0z}\left(\theta,\phi\right) = \begin{cases} 2\cos^2\theta + \frac{\sin^2\theta}{2} & \text{for LH} \\ \frac{3\sin^2\theta}{2} & \text{for HH} \end{cases} \tag{3.64}$$

3.5.4 Benchmarking of Different Direct BTBT Models

To verify that the new 2D models are consistent with the available 1D ones and one another, we first use a fairly uniform structure, which can be described as a quantum well diode, as in Fig. 3.10(Left). Compared to a bulk diode, the difference of the structure in Fig. 3.10(Left) is the narrow width of the n- and p- doped regions, resulting in significant size-induced quantization. The p- and n-sides are both highly doped ($N_\mathrm{D} = N_\mathrm{A} = 10^{20}\mathrm{cm}^{-3}$) with a thickness of 10nm each. The uniformity of the device allows us to directly compare the results obtained by 1D and 2D models since the potential profiles are almost overlapping and the quantized energy levels are very close between 1D and 2D simulators. Furthermore, it allows us to choose just a slice oriented along the x-direction as the 1D integral contour C for Bigelow's model.

The comparison of the I-V curves for the device in Fig. 3.10(Left) predicted by different models is given in Fig. 3.10(Right). As can be seen from the plot, all the four models predict the same curvature and the same order of magnitudes for the tunneling current, which means that eqs. 3.41 and 3.45 are consistent with the 1D models in the case of uniform structures. However it should be reminded that we are merely emulating the unquantized nature of the wavefunctions in the y-direction by taking a large width ($L = 500$nm). Furthermore, for this simulation only, we are taking the Γ electron effective mass $m^*_\Gamma = m^*_\mathrm{hh}$. The reason is that heavier carriers are less affected by quantization effects compared to lighter mass ones, hence we can keep L relatively small, saving simulaton time, but still emulating a 1D quantization case.

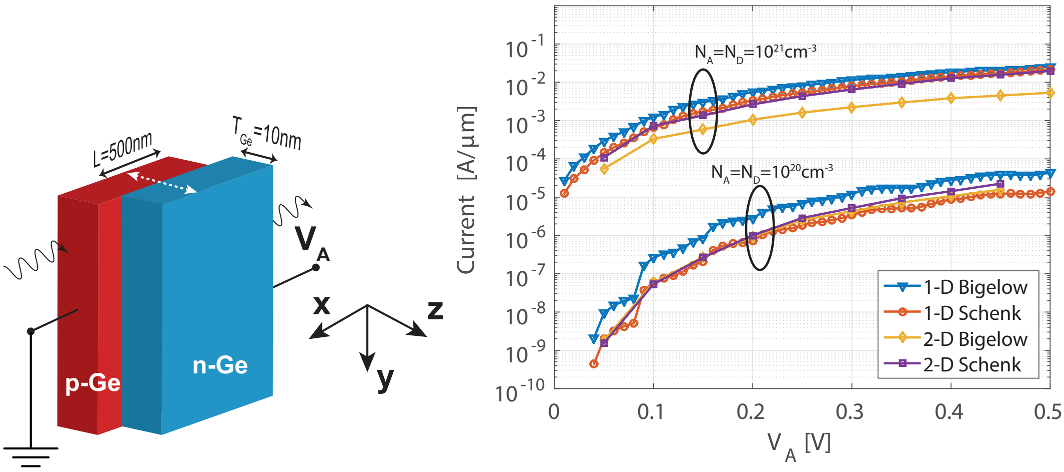


Figure 3.10 – Comparison of the two direct tunneling models (Schenk and Bigelow) for Germanium quantum well diodes for different dopings [7]. A general agreement is seen between all the cases.

Looking at the Fig. 3.10(Right), it can be deduced that Schenk model is the most robust one in complex 2D structures and the one providing closer agreement between 1D and 2D models in the QW diodes, therefore it seems to be the feasible choice to be used in both 1D and 2D simulations. We should note that since both models depend on solutions of one-band effective mass Schrödinger equation, band mixing effects which may alter the HH effective mass [112] are neglected. We will see later an empirical correction for the so-called 'anti-crossing' [113] that causes asymmetry in the real and imaginary branches of the heavy hole dispersion relation.

Non-parabolicity Corrections for Joint DOS Calculation of Tunneling States

The next step is to include the non-parabolicity corrections into the DOS calculation for counting the available states for tunneling. We assume here 2D quantization where the transverse k-space is 1D, but a similar derivation follows for 1D quantization (2D transverse k-space) as well. We go back to the δ function definition of the DOS over the two transverse k-spaces for electron and hole gases (spin excluded) and require energy and transverse k conservation:

$$g(E_{\mathrm{h}}, E_{\mathrm{e}}) = \frac{1}{\pi}\int_0^\infty\int_0^\infty \delta\left(E_{\mathrm{h}} - E_{\mathrm{e}} - \left(E_\perp^{\mathrm{h}}(k') + E_\perp^{\mathrm{e}}(k'')\right)\right)\delta\left(k'' - k'\right)dk'dk'' \tag{3.65}$$

$$g(E_{\mathrm{h}}, E_{\mathrm{e}}) = \frac{1}{\pi}\int_0^\infty \delta\left(E_{\mathrm{h}} - E_{\mathrm{e}} - \left(E_\perp^{\mathrm{h}}(k') + E_\perp^{\mathrm{e}}(k')\right)\right)dk' \tag{3.66}$$

we define $E = E_\perp^{\mathrm{h}}(k) - E_\perp^{\mathrm{e}}(k)$ and switch to energy integral:

$$\begin{aligned} g(E_{\mathrm{h}}, E_{\mathrm{e}}) &= \frac{1}{\pi}\int_0^\infty \delta(E_{\mathrm{h}} - E_{\mathrm{e}} - E)\frac{dk}{dE}dE \\ &= \frac{1}{\pi}\frac{dk}{dE}\bigg|_{E=E_{\mathrm{h}}-E_{\mathrm{e}}}\Theta(E_{\mathrm{h}} - E_{\mathrm{e}}) \end{aligned} \tag{3.67}$$

which means that we simply need to find the derivative of the inverse dispersion relation (inverse dispersion relation defined as the functional inverse of $E(k) = E_\perp^{\mathrm{h}}(k) + E_\perp^{\mathrm{e}}(k)$ and evaluate it at $E = E_{\mathrm{h}} - E_{\mathrm{e}}$. Note the change of notation from $E_{l,\Gamma}$ and $E_{l',\alpha}$ to E_{e} and E_{h} respectively, for clarity. The approach is valid for any dispersion relation. We now derive analytically the expression for the parabolic dispersion relations. We note that $\frac{dk}{dE} = \frac{1}{2}\sqrt{\frac{\overline{m}}{\hbar^2 E}}$ and we plug this in the above expression and evaluate it at $E = E_{\mathrm{h}} - E_{\mathrm{e}}$, we have the familiar JDOS expression (Eq. 3.40) [107]:

$$g(E_{\mathrm{h}}, E_{\mathrm{e}}) = \frac{1}{2\pi}\sqrt{\frac{\overline{m}}{\hbar^2(E_{\mathrm{h}} - E_{\mathrm{e}})}}\Theta(E_{\mathrm{h}} - E_{\mathrm{e}}) \tag{3.68}$$

We now want to generalize this approach to any kind of dispersion relations. In order to do so, we need to find a way to evaluate $\frac{dk}{dE}$. We do not assume any closed form expression for the E-k relation. It is only assumed that we have a set of E and corresponding k values for electron and hole bands separately (for instance, in our case we can obtain them by evaluating Eq. 3.26

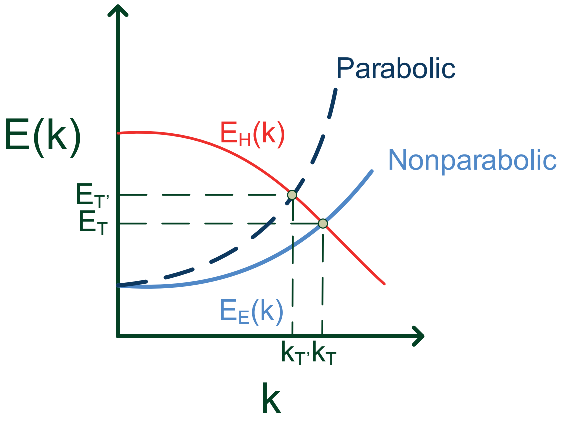


Figure 3.11 – The impact of non-parabolicity correction on JDOS.

in a certain range of E for electrons and use the parabolic E-k relation for holes). The following algorithm is proposed:

1. Solve the equation $E_\mathrm{e} + E_\perp^\mathrm{e}(k_\mathrm{T}) = E_\mathrm{h} - E_\perp^\mathrm{h}(k_\mathrm{T})$ for k_T. Note that since we only have a set of E-k values, the accuracy of the solution will depend on the spacing of our data set. To solve the equation, an iterative method such as Newton's method, for instance can be used, by calculating the derivative of the E-k relation numerically if no analytical form exists.

2. Calculate $\frac{dk}{dE}$. Since here the k involved is the inverse dispersion relation, corresponding to the *joint* E-k relation it cannot be done directly (except for the parabolic case). However, we know the value of $E = E_\perp^\mathrm{h}(k) + E_\perp^\mathrm{e}(k)$ which is the addition of transverse energies of the electron and hole bands separately, which we have already calculated. Then we make use of the following property to express the derivative at a certain point in terms of the derivative of the inverse of the function:

$$\left(f^{-1}(a)\right)' = \frac{1}{f'\left(f^{-1}(a)\right)} \tag{3.69}$$

putting in our functions we have:

$$\left.\frac{dk}{dE}\right|_{E=E_\mathrm{h}-E_\mathrm{e}} = \frac{1}{\left.\frac{dE}{dk}\right|_{k=k_\mathrm{T}}} = \frac{1}{\left.\frac{dE_\perp^\mathrm{e}}{dk}\right|_{k=k_\mathrm{T}} + \left.\frac{dE_\perp^\mathrm{h}}{dk}\right|_{k=k_\mathrm{T}}} \tag{3.70}$$

now it remains to evaluate the two derivatives around k_T numerically.

3. Then Eq. 3.67 can easily be calculated.

3.6 Efficient Modeling of Transverse Quantization

The models described so far assume an infinite extension of the device along the transverse direction which means that the wavefunctions along the transverse direction are assumed to be plane waves $\psi_{\perp} \sim e^{\vec{k_{\perp}} \cdot \vec{r_{\perp}}}$ in order to study different dimensionality combinations such as 1D-1D face tunneling [15], 1D-1D edge tunneling etc., we can change our assumption of the device extending to infinity in the transverse direction and assume that the device width is finite (say T_{W}). Then, very simply, we can modify our simulation scheme in order to account for this by doing the following changes:

1. Assume that the potential profile in the transverse direction is constant and therefore the potential profile is an infinite potential well along the transverse direction with well width T_{W},

2. The transverse k-space dimensionality is now reduced and the quantities such as carrier densities are now given as a discrete summation of different subbands whose potential profiles are the rigidly shifted versions of the actual band profiles. For example, the conduction band profile of the 1st electron subband is given as $E_{\mathrm{C}}^{*} = E_{\mathrm{C}} + E_1$ where E_1 is the 1st quantized energy level for the transverse subband and E_{C} is the conduction band profile obtained from the self-consistent solution of the Schrödinger and Poisson equations in the simulation domain. A similar formula applies for holes as well.

We note that the first assumption (potential is constant along the transverse direction) is rather crude. In fact, for narrow structures, the wavefunction will be repelled from the boundaries which will reduce the charge distribution near them. This reduction will cause bending in the potential profile and may result in significant divergence from the constant potential approximation.

We mention specifically the modifications for 1D and 2D version of the codes in the following subsections. We note here that all the indexes and symbols introduced for the models without the transverse quantization (i.e, sections 3.2, 3.3 and 3.5) bear the same meanings here.

1D Quantization with 1D Transverse k-space

In this case, we modify the 1D version of the code to incorporate a 1D transverse k-space instead of the previous 2D. Since we assume a uniform potential along the finite transverse direction (say y, with width T_{W}), the y wavefunction components $\phi_m(y)$ are given by $\phi_m(y) = \sqrt{\frac{2}{T_{\mathrm{W}}}} \sin\left(\frac{m\pi}{T_W} y\right)$. As known, the solutions of the infinite square well potential form a complete basis set and ϕ_m are orthonormal i.e., $\langle \phi_m | \phi_{m'} \rangle = \delta_{m,m'}$ [37]. The charge densities are now given by (remember that z is the other quantization direction, whereas along x the wavefuction are plane waves):

$$n_q(z) = \sum_{\alpha\in c}\sum_{m}\sum_{l}\sqrt{\frac{2m_{\alpha,\perp}k_{\rm B}T}{\pi^2\hbar^2T_{\rm W}^2}}\left|\psi_{l\alpha}(z)\right|^2\mathscr{F}_{-1/2}\left(\frac{E_{\rm Fn}-(E_{l\alpha}+E_m)}{k_{\rm B}T}\right) \tag{3.71}$$

$$p_q(z) = \sum_{\alpha\in v}\sum_{m}\sum_{l}\sqrt{\frac{2m_{\alpha,\perp}k_{\rm B}T}{\pi^2\hbar^2T_{\rm W}^2}}\left|\psi_{l\alpha}(z)\right|^2\mathscr{F}_{-1/2}\left(\frac{(E_{l\alpha}-E_m)-E_{\rm Fp}}{k_{\rm B}T}\right) \tag{3.72}$$

where $E_m = \frac{\pi^2m^2\hbar^2}{2m_{\alpha,\perp}T_{\rm W}^2}$ for parabolic bands and the solution of $E_m(1+\alpha_{\rm NP}E_m) = \frac{\pi^2m^2\hbar^2}{2m_{\alpha,\perp}T_{\rm W}^2}$ for non-parabolic conduction bands. As far as the BTBT expression goes, the JDOS used is the same as Eq. 3.40. The other modification is the inclusion of the orthogonality condition $\delta_{m,m'}$ which imposes the quantum numbers m, m' of the electron and hole wavefunctions, respectively, to be equal [37]:

$$I^{\rm 1D}_{\perp-\rm lim} = \frac{4\pi qWL_G}{\hbar T_{\rm W}}\sum_{m,m'}\sum_{l\in\Gamma}\sum_{\alpha,l'\in v}|M_{\rm CV}|^2\times$$
$$\frac{1}{2\pi}\sqrt{\frac{\overline{m}}{\hbar^2(E_{l'\alpha m'}-E_{l\Gamma m})}}\delta_{m,m'}\Theta(E_{l'\alpha m'}-E_{k\Gamma m})\left(f_{\rm c}(E_{\rm T})-f_{\rm v}(E_{\rm T})\right) \tag{3.73}$$

where $E_{l\Gamma m}$ and $E_{l'\alpha m'}$ are given as $E_{l\Gamma m} = E_{l\Gamma}+E_m$ and $E_{l'\alpha m'} = E_{l'\alpha}-E_{m'}$ and they represent the quantized energy levels for electron and holes for a transverse mode m and m' respectively.

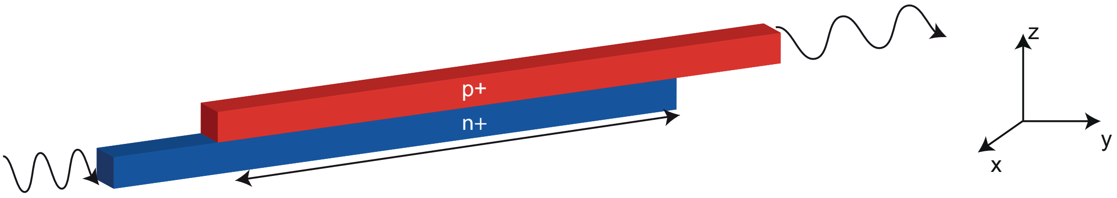


Figure 3.12 – A 1D-1D face tunneling diode.

Within the classification of Agarwal et al. [15], this modified code could be used to simulate 1D-1D face tunneling devices (see Fig. 3.12) in which the tunneling direction is aligned with the quantized direction.

2D Quantization with 0D Transverse k-space

In this case, the transverse k-space reduces to 0D, i.e., the density of states is 1 (excluding spin). Then, the charge densities are simply given by:

$$n_q(\vec{r}) = \frac{2}{T_{\rm W}}\sum_{\alpha\in c}\sum_{l}\sum_{m}\left|\psi_{l\alpha}(\vec{r})\right|^2\left(\frac{1}{1+e^{\frac{(E_{l\alpha}+E_m)-E_{\rm Fn}}{k_{\rm B}T}}}\right) \tag{3.74}$$

$$p_q(\vec{r}) = \frac{2}{T_{\mathrm{W}}} \sum_{\alpha\in v} \sum_{l} \sum_{m} \left|\psi_{l\alpha}(\vec{r})\right|^2 \left(\frac{1}{1+e^{\frac{E_{\mathrm{Fp}}-(E_{l\alpha}-E_m)}{k_{\mathrm{B}}T}}} \right) \tag{3.75}$$

where $\vec{r} = (x, z)$ and we assume uniform charge distribution along y. The JDOS, on the other hand, reduces into a δ-function:

$$I^{\mathrm{2D}}_{\perp-\mathrm{lim}} = \frac{4\pi q}{\hbar T_{\mathrm{W}}} \sum_{m,m'} \sum_{l} \sum_{\alpha,l'\in v} |M_{\mathrm{CV}}|^2 \delta_{m,m'} \delta\left(E_{l'\alpha m'} - E_{k\Gamma m}\right) \left(f_{\mathrm{c}}(E_{\mathrm{T}}) - f_{\mathrm{v}}(E_{\mathrm{T}})\right) \tag{3.76}$$

Within the classification of Agarwal et al. [15], this modified code could be used to simulate 1D-1D edge tunneling devices as well as 1D-1D face tunneling ones. Examples could include ultra narrow cross-section TFETs such as in [114], or EHBTFETs (briefly described in section 2.2 and detailed in the next chapter) with a very narrow width that results in quantization along the transverse direction. We will make use of this extension in the next chapter to compare the different dimensionality effects for two different TFET architectures.

Finally, we note that the idea presented here could also be extended further for other combinations such as 1D Quantization with 0D transverse k-space which would correspond to e.g, quantum dot p-n junctions.

3.7 Modeling the Imaginary Dispersion Effective Mass Discrepancy

Models based on the effective mass approximation (EMA) are more efficient than other approaches such as $k\cdot p$ from a numerical point of view, can handle large devices and can include phonon assisted tunneling with a small additional effort [115]. They are however based on separate (i.e., single band) solutions of the electron and hole Schrödinger equations (see section 3.2). The coupling between bands is added in post processing assuming a dispersion relationship in the gap that connects the electron and hole branches (section 3.5.1 & 3.5.2).

In such approaches, the curvature of the E-k relationship is symmetric in both real and imaginary branches which means that the effective mass is the same for both real and imaginary k. This assumption, however, implies that size-induced quantization should essentially kill the tunneling current. The logic is as follows: the BTBT in bulk materials is dominantly caused by light holes in the bulk structure due to their very low effective mass which increases greatly the wavefunction overlap of the electron and hole states and typically heavy hole contribution is virtually negligible. Contrary to this rationale, NEGF simulations [70] show significant BTBT current even in quantum well and nanowires, where LH bands are not expected to contribute to current due to strong quantization.

This evident contradiction is caused by the fact that the tunneling (imaginary) effective mass is closer to the LH mass even when HH subbands are dominant. This phenomenon is called "anti-crossing" [113] of the light-hole (LH) and heavy-hole (HH) branches and it is not natively included in the EMA approach, which means that the curvature of the E-k relationship is symmetric in both real and imaginary branches. With "anti-crossing" we mean that, the top of the valence band in III-V semiconductors is connected with the conduction band through an imaginary branch with low mass (the LH one), even in the presence of strong quantization that makes the HH branch dominate. *The goal here is to show that the EMA-NP model using the LH mass as tunneling mass for the HH subband fairly reproduces* $k \cdot p$ *results.*

To obtain the tunneling parameters, we have fitted the imaginary dispersion relation (see Fig. 3.13 for the real and imaginary dispersion relations obtained) from the 4-band $k \cdot p$ simulator developed by the University of Bologna [8]. The effective masses obtained from the fitting are reported in Table 3.1. For InAs, the fitted tunneling effective mass ($m_{\mathrm{HH,imag}}$) is drastically different than the effective mass of the real branch (m_{HH}). Note that the fitted value almost coincides with the bulk LH mass, thus confirming the results of [60, 61] and highlighting the anti-crossing [113]. The electron dispersion, on the other hand, rather exhibits a symmetry between the real and imaginary branches.

Table 3.1 – Effective masses and conduction band non-parabolicity factor extracted for InAs slabs of different thicknesses (see Fig. 3.13).

Slab thickness	$m^*_{\mathrm{HH,real}}$	$m^*_{\mathrm{HH,imag}}$	m^*_{e}	$\alpha_{\mathrm{NP}}\,[\mathrm{eV}^{-1}]$
5nm	0.33	0.037	0.037	3.6
10nm	0.33	0.03	0.026	
15nm	0.33	0.023	0.025	

This finding, however, brings up the question of using adequate effective masses for the classically allowed ($E_{\mathrm{h}} < E_{\mathrm{V}}$) and forbidden regions ($E_{\mathrm{h}} > E_{\mathrm{V}}$), since the effective mass in the forbidden region strongly determines the amplitude of the wavefunction tail, and therefore the spatial overlap between the electron and hole states in the bandgap (Eq. 3.42). To overcome this issue, we re-calculate the WFs using the quantized energies obtained by the 1D Schrödinger equation. We employ WKB approximation using single-band approximation, in which we modify the effective mass in the forbidden region e.g., with the value extracted from the imaginary dispersion calculated by the 4-band $k \cdot p$ simulator (table 3.1) or the light hole mass. The dispersion relation is given by:

$$k_z(z) = \begin{cases} \frac{\sqrt{2m^*_{\mathrm{HH,real}}(E_{\mathrm{V}}(z)-E_{\mathrm{h}})}}{\hbar} & E_{\mathrm{h}} < E_{\mathrm{V}} \\ \frac{\sqrt{2m^*_{\mathrm{HH,imag}}(E_{\mathrm{V}}(z)-E_{\mathrm{h}})}}{\hbar} & E_{\mathrm{h}} > E_{\mathrm{V}} \end{cases} \tag{3.77}$$

The WF is then calculated using the familiar expressions given by the WKB approximation [37]:

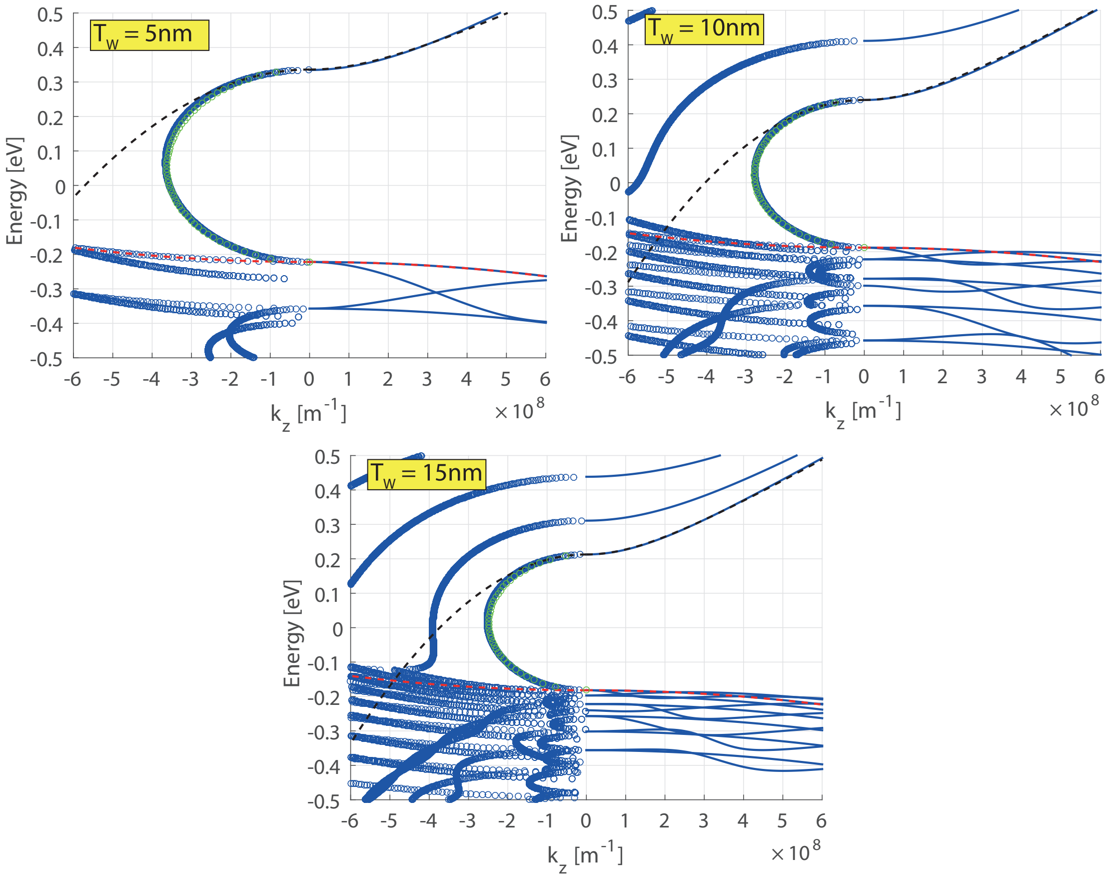


Figure 3.13 – Real ($k_z > 0$) and imaginary ($k_z < 0$) dispersion for (Up left) 5nm (Up right) 10nm (Bottom) 15nm thick InAs slabs from $k \cdot p$ [8]. The red and black curves indicate the real and imaginary relations predicted by single band approximation for HH (with mass $m^*_{\mathrm{HH,real}}$) and conduction band (with mass m^*_{e}), respectively. The green curve is the fit obtained by the Kane's two-band dispersion relation [9] (Eq. 2.14). Parameters from the fitting are given in Table 3.1.

$$\psi_h(z) = \begin{cases} \frac{2C}{\sqrt{|k_z(z)|}} \sin\left(\int\limits_{z}^{z_c} k_z\left(z'\right) dz' + \frac{\pi}{4} \right) & E_{\mathrm{h}} < E_{\mathrm{V}} \\ \frac{C}{\sqrt{|k_z(z)|}} e^{-\int\limits_{z_c}^{z} |k_z(z')| dz'} & E_{\mathrm{h}} > E_{\mathrm{V}} \end{cases} \tag{3.78}$$

where C is a normalization constant and z_c is the classical turning point (i.e, $E_{\mathrm{h}} = E_{\mathrm{V}}$).

3.8 Results Comparison & Calibration

As a first check, we compare the self-consistent solutions obtained by our EMA-NP code and the $k \cdot p$ results for a bulk InAs tunnel diode. The parameters used for the bulk diode simulations are given in Table 3.2. The band profiles for the bulk InAs diode (Fig. 3.14) show excellent matching between NP-EMA and $k \cdot p$ for the n-side. The slight difference in the p-region comes from the warped bandstructure of holes not captured by the parabolic approximation employed in our model. Fig. 3.16(Left) shows the current densities obtained by the EMA-NP model for a bulk diode with different doping levels. A general agreement with the $k \cdot p$ results, better at low doping levels, is observed.

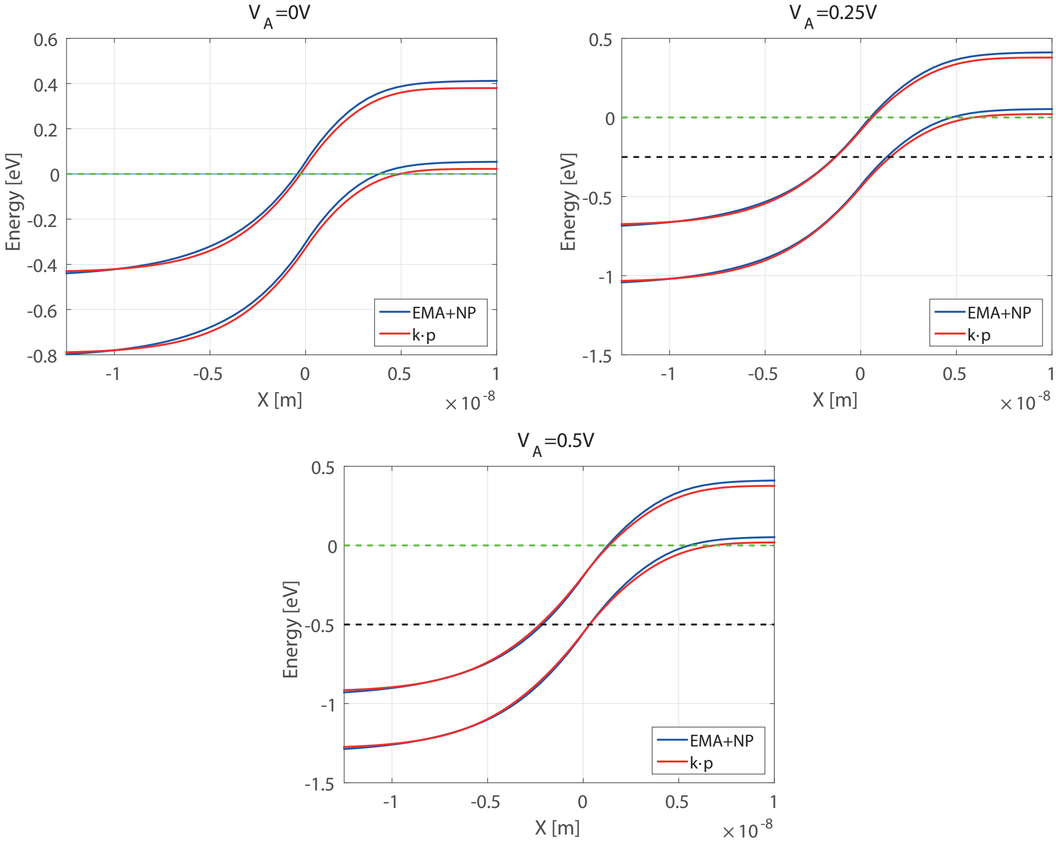


Figure 3.14 – Comparison of the potential profiles calculated for an InAs tunnel diode with $N_A = N_D = 2 \times 10^{19}\text{cm}^{-3}$ for applied reverse biases of (up, middle, down) 0V, 0.25V, 0.5V. The quasi-Fermi levels for electrons (holes) are indicated as black (green) dashed lines. 1D EMA-NP model described in previous sections is compared against $k \cdot p$ calculations from [8].

Going further, we now investigate the impact of transverse quantization, using our 1D code with transverse quantization. We simulate InAs quantum well diodes with finite transverse width T_W. For all the I-V characteristics results presented here on, we take the potential profile calculated by the $k \cdot p$ simulator as input and run our BTBT models as a post-processing step, in order to better isolate the influence of the tunneling parameters from the influence of device electrostatics.

Fig. 3.16(Right) compares the $k \cdot p$ and EMA-NP results for QW diodes (Fig. 3.15) with $T_W = 5$nm

Table 3.2 – Parameters used by EMA+NP model for Ge, InAs and Si for the calibration and comparison study. The transport (tunneling) is assumed to occur along the [100] direction in all cases. The Γ valley is assumed to be isotropic. In the anisotropic cases (L and Δ valleys) the effective masses are given as simulation-mass/transverse-mass. The degeneracy of each valley is indicated in parantheses.

Material	m^*_{L}	m^*_{Γ}	$\alpha_{\mathrm{NP},\Gamma}$[eV^{-1}]	m^*_{Δ}	$\alpha_{\mathrm{NP,L}}$[eV^{-1}]	$\alpha_{\mathrm{NP},\Delta}$[eV^{-1}]	m^*_{hh}	m^*_{lh}
Ge	0.117/0.292 (4)	0.043	-	0.85	0.33	-	0.33	0.043
InAs	-	0.022	-	3.6	-	-	0.41	0.025
Si	-	-	-	0.9163/0.1905 (4) 0.1905/0.416 (2)	-	0.5	0.49	0.16

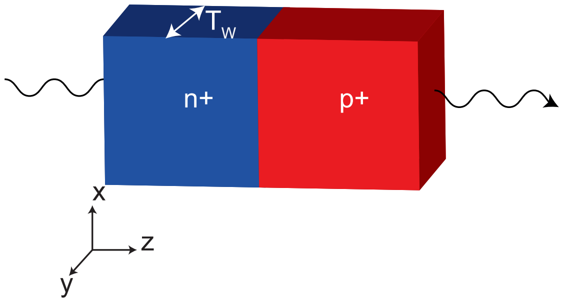


Figure 3.15 – The InAs diode structure considered in this study. For bulk diode T_{W} is assumed infinite, whereas in quantum well diode simulations T_{W} is assumed finite (T_{W} values are specified where necessary for each case).

and different doping levels. The difference between the models seems to be increasing with increasing doping levels, which results from states states with higher k being able to contribute to the current, where the EMA description starts to break down.

Fig. 3.17(Left) compares the models for different T_{W}: both models show the same trend of current increase as T_{W} increases. However, comparing Figs. 3.16(Left) & 3.16(Right), it is seen that the current density remains nearly the same, which signals that in both QW and bulk diodes, it is the same imaginary path that connects the valence and conduction bands with very similar effective mass values. Of course, a slight degradation is observed for the quantized cases compared to the bulk ones, due to the transverse quantization which increases the effective bandgap. To assess the impact of the large asymmetry in mass of the real and imaginary branches, we compare the cases using the real or imaginary branch masses in the forbidden region: Fig. 3.17(Right) shows a dramatic decrease in current when the real HH mass is used for tunneling.

To assess the impact under strong field-induced quantization, we now switch to another device, the EHBTFET [5], which was briefly introduced in the previous chapter. It is expected to have much stronger quantization effects in the EHBTFET compared to the diode due to the fact that n- and p-gates are typically biased in opposite polarities which results in a

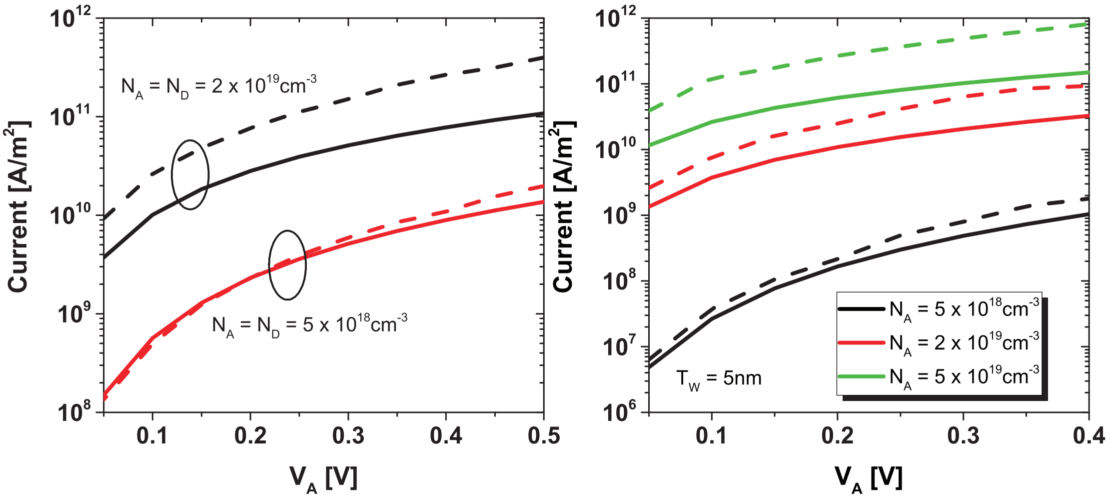


Figure 3.16 – (Left) I-V characteristics of bulk diodes from EMA-NP (dashed) and $k \cdot p$ (solid) for various doping levels. (Right) Comparison of I-V characteristics for the QW diode shown in Fig. 3.15 considering various doping levels. $T_{\mathrm{W}} = 5$nm. EMA-NP (dashed), $k \cdot p$ (solid).

large electric field. We simulate a vertical slice in the overlap region (see the inset in the Fig. 3.18(Left)). Fig. 3.18(Left) compares the wavefunctions obtained using formulas in Eq. 3.78 using real or imaginary masses for heavy holes in the forbidden gap regions. Note the orders of magnitude higher amplitude in function tails (i.e., WF component in the forbidden region) due to much lighter imaginary effective mass.

Looking at Fig. 3.18(Left), one can of course note the spikes at the classical turning point, caused by the fact that WKB approximation breaks down around the classical turning points [37]. While this may seem a serious problem, for our purposes, it can be noted that we are interested in the overlap of wavefunctions (more specifically, the product $\psi_{\mathrm{e}}\psi_{\mathrm{h}}$). We perform a numerical experiment (Fig. 3.18(Right)) in order to quantify the impact of the spikes on the quantities that we are interested in. To do this, we take the WF overlaps obtained by the initial version of our 1D code where the wavefunctions are calculated directly from the numerical solution of the Schrödinger equation and we compare them against our WKB-based correction code with the *real* effective mass in the entire device region (i.e., the anti-crossing effect is turned off). We note that the spikes are far away from the region where the overlap is significant and the difference between the squared coupling elements $|M_{\mathrm{CV}}|^2$ ($\propto$ overlap integral) of numerical and the WKB-based solution is relatively small (~ 5%, see Fig. 3.18 (Right)). Bear in mind that some part of this discrepancy comes from the fact that WKB is an approximate method, independently of the presence of the spike.

One final note is that the WKB-based model to account for the imaginary dispersion relation is implemented only for the 1D version whereas in the 2D version we utilize the standard EMA (i.e., symmetric real and imaginary dispersion relations) since 2D extension of the WKB approximation, to our knowledge, is only possible when the direction of the probability flux is

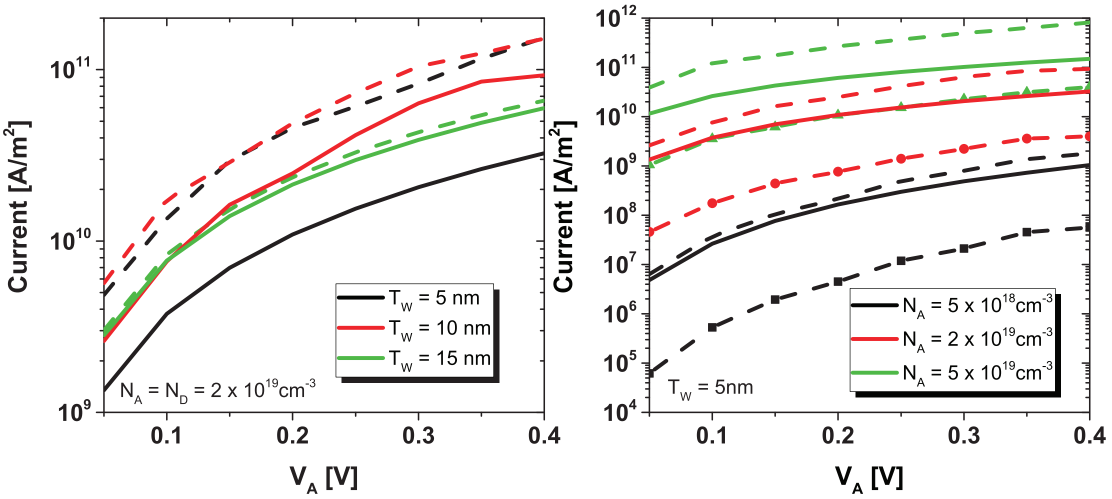


Figure 3.17 – (Left) I-V characteristics for various T_W values for the InAs diodes in Fig. 3.15. $N_A = N_D = 2 \times 10^{19}\text{cm}^{-3}$. EMA-NP (dashed) $k \cdot p$ (solid) (Right) Comparison of I-V characteristics using $k \cdot p$ (solid) and EMA-NP with $m_{\text{tun,HH}} = 0.03m_0$ (Dashed) or with $m_{\text{tun,HH}} = 0.33m_0$ (Dashed with symbols). $T_W = 5\text{nm}$.

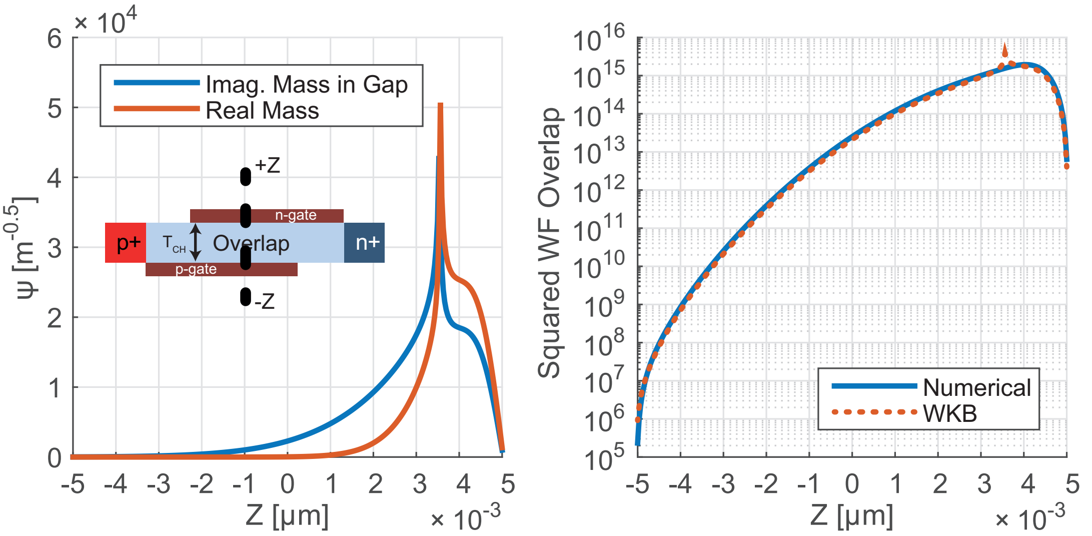


Figure 3.18 – (Left) Comparison of wavefunctions (lowest Γ electron subband) using (blue) imaginary mass (red) real mass listed for T_W = 10nm in table 3.1 in the gap, for an InAs EHBTFET with T_{CH} = 10nm. (Right) Comparison of the wavefunction overlap squared between the lowest electron and hole subbands with (blue) numerically calculated (red, dashed) WKB-based (Eq. 3.78) WFs.

known in advance [116] which is not the case for a general-purpose simulator where different geometries can be simulated.

3.8.1 Phonon-Assisted Band-to-Band Tunneling Calibration & Comparison against Experiments

We now turn our attention to indirect gap semiconductors and compare our results against experimentally reported diode data. We see a general agreement exists between our results obtained using Vandenberghe model explained in section 2.3.2 and the experimental literature [10, 11] in Fig. 3.19(Left) for Silicon. Fig. 3.19(Middle) shows that our model for direct tunneling in Ge, without any adjustment, reproduces the experimental data for Ge diodes collected in [10]. The parameters used for the bulk diode simulations are given in Table 3.2. We observe that direct BTBT domintates over phonon-assisted BTBT in bulk diodes. However, once quantization comes into play we will see that this picture will change dramatically (in Chapter 4).

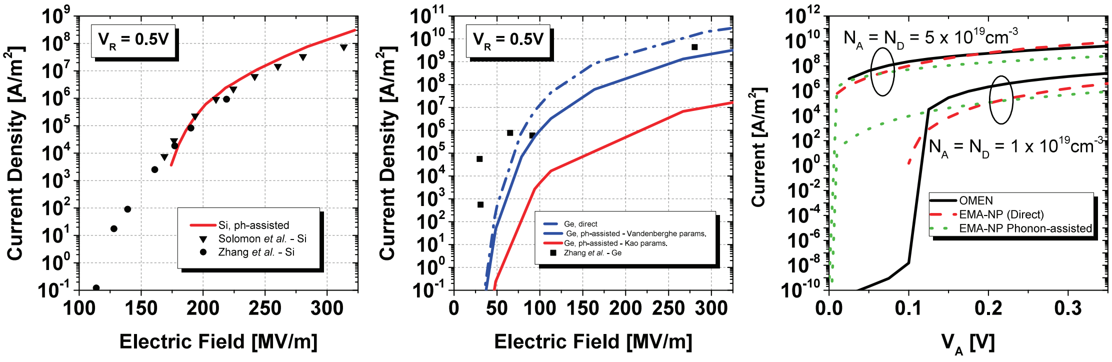


Figure 3.19 – Current versus maximum electric field for (Left) Silicon (Middle) Germanium bulk tunnel diodes under a reverse bias of 0.5V. Symbols indicate experimentally measured data points (taken from [10, 11] for Silicon and [10] for Germanium) whereas lines indicate our simulation results. For Germanium (Middle figure), blue and red solid lines represent the simulations using the parameters given in [12] and in [13] respectively. (Right) Reverse bias characteristics of simulated bulk Ge tunnel diodes for different doping levels by the atomistic simulator OMEN [14] and our code (labeled EMA-NP).

We also note here the dramatic difference of phonon-assisted BTBT rate between the two sets of parameters given in the literature by Kao et al. [13] and Vandenberghe et al. [12] where the latter one is an update to the first. Comparing the results in Fig. 3.19(Right) for the two parameter sets, we can see that the parameters of Vandenberghe et al. estimate about 200x higher BTBT rates than the one of Kao et al. The parameter set given in [12] is obtained using a first principles density functional theory (DFT) code and we consider it as more credible than the set given before in [13] which is a "crude estimate" [12]. Due to this, we will use in the remainder of the thesis the deformation potential $D = 7.8 \times 10^{10}$eV/m and phonon energy $E_{\mathrm{ph}} = 6$meV [12], which are inputs of the Eqs. 2.24 & 2.25 that we use to estimate the phonon-assisted BTBT rate.

Fig. 3.19(Right) compares the results obtained by our tool with the atomistic tight-binding simulator OMEN [14] for Ge tunnel diodes in reverse bias. A good mutual agreement is obtained, validating the proposed model. Since the thermionic current in reverse bias is neglected, very low current levels are observed for the device with $N_\mathrm{A} = N_\mathrm{D} = 1 \times 10^{19}\mathrm{cm}^{-3}$ before the device turns on (i.e., the conduction and valence bands are aligned). OMEN results for the doping $N_\mathrm{A} = N_\mathrm{D} = 5 \times 10^{18}\mathrm{cm}^{-3}$ show very little current at lower V_A due to the fact that electron-phonon interactions (which require a very large amount of computational resources) were turned off for this curve, for which prevents the prediction of the phonon-assisted BTBT current.

3.9 Chapter Summary & Highlights

In this chapter, a brief description of the quantum mechanical simulation schemes used in the following parts of the thesis is provided. The contributions of this chapter includes:

- A simulation scheme for solving closed boundary Schrödinger-Poisson equations self-consistently has been implemented. The 1D EMA-NP solver using FDM can be used for parameter space investigation with more accurate non-parabolicity correction models. 2D solver FEM can be used to investigate 2D effects in tunneling devices.
- Modifications to already existing direct band-to-band tunneling models made to consider tunneling in quantized gases. Expressions for polarization dependent form factors $C_{0x,z}$ are derived. Two different direct BTBT models implemented and benchmarked using a quantum well diode.
- The EMA-NP model has been modified to account for anti-crossing of the valence band and transverse quantization through adequate shifting of energy bands and modification of the transverse JDOS. A general agreement is observed between the EMA-NP and $k \cdot p$ models after proper calibration.
- Variants of the implemented 1D and 2D models including transverse quantization effects are proposed and implemented. These variants enable the simulation of various tunneling dimensionality combinations.

4 Dimensionality Effects in Tunneling Devices: Exploitation and Device Optimization

As discussed in the previous chapters, the concept of *density of states* (DOS) switch may show a great promise since the shape of the DOS of the charge carriers in quantized cases can result in abrupt switching. Agarwal et al. [15] defines various combinations in terms of the dimensionality of the electron and hole gases (see Fig. 4.1). Throughout the chapter, we will follow the classifications and naming conventions made by Agarwal et al. [15] to denote the carrier and tunneling dimensionalities.

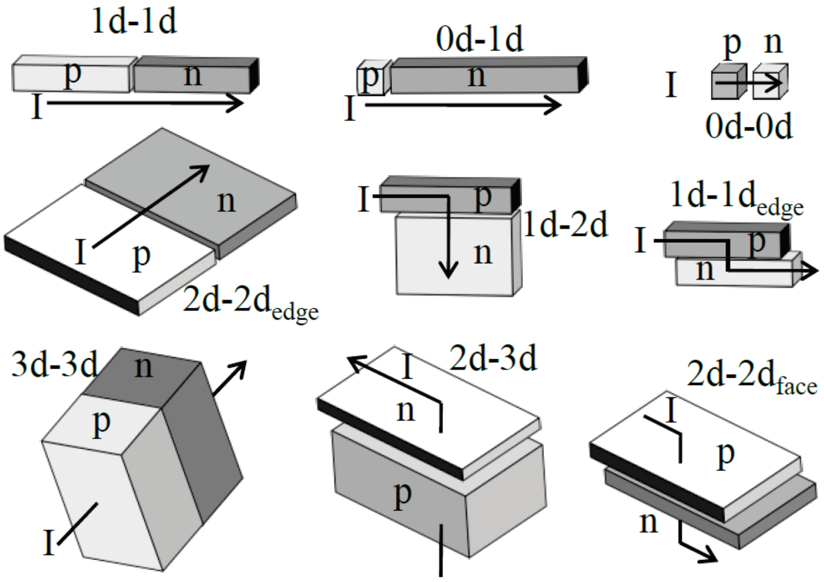


Figure 4.1 – The different dimensionality combinations of the carrier gases in the BTBT devices (taken from [15]).

For tunnel FETs, the case where BTBT occurs between two 2D gases appears to be the most promising for a variety of reasons. First, the 2D DOS function is step-like; therefore, a significant amount of states are available once the electron and hole levels align. This is in contrast to other combinations such as 1D-1D tunneling (nanowires) or 0D-0D tunneling (quantum-dot transistors), where the low DOS limits the amount of BTBT current even though in terms of electrostatics they might be advantageous. Hence it would be necessary to use many of such lower dimensional devices in stacks or arrays which might be a huge challenge in terms of processing. The 2D-2D face tunneling case, instead, offers an optimum point in terms of current drive. Another advantage of 2D-2D tunneling is the contact resistance to access the device. We expect that contact and interconnect resistances will be very critical for lower dimensional devices, whereas 2D-2D tunneling devices will fare better, due to large area available for contacts.

In this chapter, we will mainly investigate a 2D-2D tunneling device called *Electron Hole Bilayer TFET*, that could be a promising candidate for a *sub-thermal switch* which offers switching slopes (SS) much lower than 60mV/dec. We describe the device working principle along with initial quantum mechanical simulations using the tools described in the previous chapter. Then, we focus on some of the non-idealities that can potentially be problematic, such as the lateral leakage. We extensively study the device parameter space and point out key features of 2D-2D face tunneling. Then, we study another incarnation of the EHBTFET using a fin structure and discuss it in detail. We then propose a novel logic family making use of the unique properties of the EHBTFET. In the final section, we discuss a few other dimensionality cases, by using our simulation tool that accounts for transverse quantization, to see if they present advantages with respect to the 2D-2D case.

4.1 Device Working Principle

The EHBTFET is a thin double-gated p-i-n diode (see Fig. 4.2), similar to a conventional TFET. The i-region is gated from both top and bottom to induce electron and hole inversion layers by field effect. The only architectural difference between a conventional double gate (DG) TFET [117] and the EHBTFET is the asymmetric placement and biasing of the gates. The region which is overlapped by the two gates is called 'overlap region', whereas the two regions that are controlled by only one gate are called 'underlap regions'.

The asymmetrically-biased gates induce quantized energy levels in both electron and hole inversion layers in the thin channel. The exact configuration of subbands in terms of quantization energy and occupancy depends on the semiconductor band structure and on the quantization direction. The basic device operation can be summarized as follows: in the OFF state, the minimum subband energy of electrons $E_{\mathrm{e}1}$ is above the maximum subband energy of the holes $E_{\mathrm{h}1}$ therefore no BTBT can take place. When the device enters into the ON state, the subband energies for electrons and holes align and, owing to the step-like behavior of the 2D density of states (DOS), a significant amount of states are available to contribute for tunneling even at low voltages and thus a steep switching behavior can be obtained.

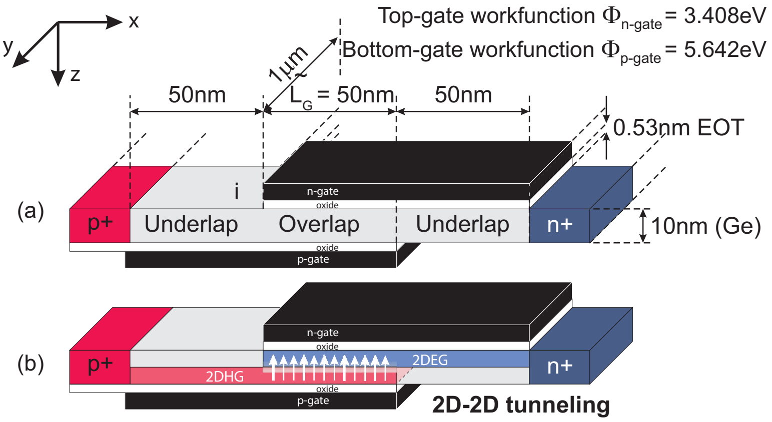


Figure 4.2 – (a) Cross-sectional view of the EHBTFET and device parameters utilized in the initial simulations. (b) Formation of induced electron/hole the bilayer by application of appropriate gate biases. BTBT direction denoted with arrows.

4.1.1 Numerical Simulations of the Intrinsic Device: Ge Channel

We utilize the 1D version of the Schrödinger-Poisson self-consistent solver described in the Chapter 3. We assume Dirichlet boundary conditions for the oxide-semiconductor interface (i.e. we neglect wavefunction penetration into the gate oxide). A Ge EHBTFET with a gate length L_G = 50nm is considered in the simulations (Fig. 4.2). Both gate stacks feature a gate oxide with an equivalent oxide thickness (EOT) of 0.53nm. The i-region is lightly n-doped with $N_D = 10^{15}\text{cm}^{-3}$. Unless otherwise stated, a device with channel thickness T_{CH} = 10nm is considered, and the n- and p-gate workfunctions are $\phi_{n-\text{gate}} = 3.408\text{eV}$ and $\phi_{p-\text{gate}} = 5.642\text{eV}$, respectively. The quantization direction is aligned with the [100] crystal direction of Ge. In this direction the L electron valleys of germanium are fourfold degenerate with quantization effective masses calculated as $m_{z,\alpha} = 0.117m_0$ and transverse effective mass $m_{xy,\alpha} = 0.29m_0$[82]. For the Γ valley, the effective masses of heavy, light holes and electrons are $m_{\text{HH}} = 0.33m_0$, $m_{\text{LH}} = 0.043m_0$ and $m_e = 0.043m_0$, respectively [13]. The electron-phonon interaction term $\Omega\left|M'_{k_0}\right|^2$ and the phonon energy $\hbar\omega_{k_0}$ for Ge are calculated (using Eq. 2.24) and given as: $6.62\times10^{-30}\text{eV}^2\text{m}^3$ and 6meV, respectively, using the parameters from [13] and [12]. Sentaurus simulations have been used for the calculation of the SRH recombination contribution with the following carrier lifetime parameters: $\tau_{\text{max,n}} = 4\cdot10^{-5}\text{s}$ and $\tau_{\text{max,p}} = 4\cdot10^{-5}\text{s}$ [118]. The non-parabolicity factor for the Γ conduction band of germanium is taken as 0.85eV^{-1} [119].

Sentaurus Device with the nonlocal BTBT model has been used for comparison purposes. Quantum mechanical effects on the electrostatics are included in Sentaurus Device via the Density Gradient (DG) correction method [57].

For this study, we only vary the n-gate bias (which controls the electron inversion layer) and keep the p-gate bias at a constant value. Consequently, the p-gate workfunction is adjusted so that the hole inversion layer is already formed in zero bias ($V_{\mathrm{p-gate}} = 0$).

The transfer characteristics are shown in Fig. 4.3 in conjunction with the evolution of the subband energies. The results confirm the working principle described in section 4.1: an OFF region is seen until $V_{\mathrm{GS}} \approx 0.05\mathrm{V}$ due to lack of alignment between any electron and hole subband. As the n-gate bias increases, the electrostatic potential in the channel region controlled by the n-gate increases as well, therefore pushing down the subband energies of the electrons (in a linear fashion in weak inversion). The onset of strong inversion where the charge density starts to reach critical values is seen when the lowest subband approaches the drain quasi-Fermi level $E_{\mathrm{Fn}} = -qV_{\mathrm{DS}}$. A slight effect on the hole energy levels is also observed, since the n-gate voltage influences the portion of the channel populated by the holes as well.

The switching behavior is caused by the alignment of the ground state (from L valley) of electron and heavy hole energy level which are denoted in the plot as e1 and hh1, respectively. The most striking feature of this characteristic is the super steep switching slope, allowing for $SS_{\mathrm{avg}} \approx 10^{\mathrm{mV}}/_{\mathrm{dec}}$ over three decades at $V_{\mathrm{DS}} = 0.5\mathrm{V}$. We note that this represents an ideal case, where all 2D effects which we will show to be critical are assumed to be suppressed. However, above a certain V_{GS} ($\sim 0.3\mathrm{V}$), a plateau of relatively constant current is seen until a second alignment takes place, resulting in an ON current of $\sim 2^{\mu\mathrm{A}}/_{\mu\mathrm{m}}$ at $V_{\mathrm{GS}} = V_{\mathrm{DS}} = 0.5\mathrm{V}$. The reason for that saturation of current is the decrease of the wavefunction overlaps of e1 and hh1 levels (see Fig. 4.4), which is caused by the increased electric field in the confinement well, pushing electron and hole wavefunctions towards the semiconductor/oxide interfaces. Although the energetic window allowing for tunneling is constantly increasing with increasing gate voltage, the reduction in wavefunction overlap can actually counterbalance this gain and may cause a decrease of the drain current for increasing V_{GS}. Further increasing the gate voltage eventually results in overlap of the heavy hole and Γ valley electron subbands, hence allowing for direct BTBT to occur.

Note that the results presented here differ from our already published results [115] due to the change of deformation potential and phonon energy parameters. In [115], we had utilized the parameter set given in [13] whereas we updated our results according to a more recent parameter set given in [12]. The rationale for this is explained in section 3.8.1. The main difference is an increase in phonon-assisted BTBT current by about 200x in the results presented in this thesis compared to [115].

Since we have a bandgap difference of $\Delta E_{\mathrm{G}} = 0.14\mathrm{eV}$ between L and Γ valleys in Ge, the onset of direct tunneling is expected at relatively lower voltages compared to other indirect bandgap materials such as silicon. While this onset is close to ΔE_{G} in structures with no quantization such as bulk diodes [120], the onset in a quantized structure such as the EHBTFET will be shifted to higher voltages. Indeed, we observe a shift of $\Delta V_{\mathrm{GS}} \sim 0.6\mathrm{V}$ between the onset of phonon-assisted and direct BTBT for the EHBTFET under study (see Fig. 4.3).

It is helpful to underline the main determinants of the tunneling current in relation to the

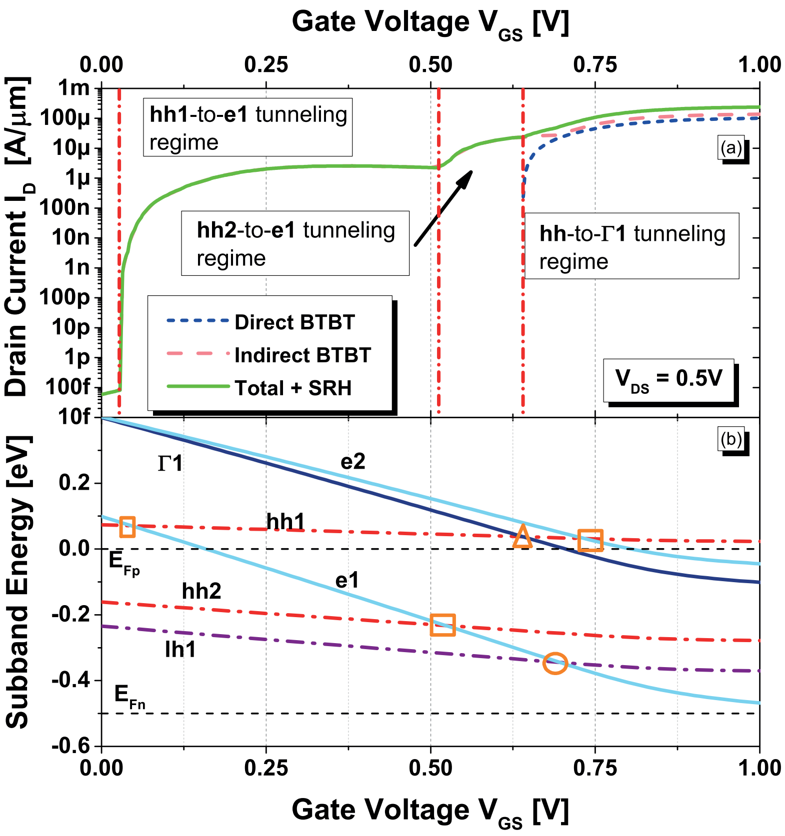


Figure 4.3 – (a) Transfer characteristics ($I_D - V_{GS}$) for the EHBTFET with separate contributions from phonon-assisted and direct BTBT indicated. Results indicate steep switch-on behavior at low V_{GS} and additional contributions due to alignment of higher energy subbands. (b) Evolution of the subband energies for L-valley (e), Γ-valley (Γ) electrons, heavy (hh) and light (lh) holes with respect to the n-gate voltage V_{GS}.

discussion above. A closer look at equations 3.11, 2.25 and 3.35 reveals that the tunneling current is mainly dependent on: i) the energetic position of the quantized levels; ii) the density of states of both electrons and holes; iii) the overlap between the electron and hole wavefunctions. The current spectra given in Fig. 4.5 indicate that, provided subband alignment takes place, light hole contribution is higher than that of the heavy holes. However, low effective mass of the light holes result in considerably higher quantization energies compared to heavy holes (e.g., compare the contributions of hh2 and lh1 in Fig. 4.5(c)) therefore the light holes do not play any role at low voltage levels in the region of switching.

We can also note a spike in the current spectrum in Fig. 4.5(c) around $E = 20$meV corresponding to the only allowed tunneling energy (E_T in eq. (3.33)) that preserves momentum and energy of the tunneling electron. We do not observe such behavior for phonon-assisted BTBT, since the phonons compensate for the momentum difference between the initial and final state, therefore resulting in a continuum of tunneling energy values available between aligned

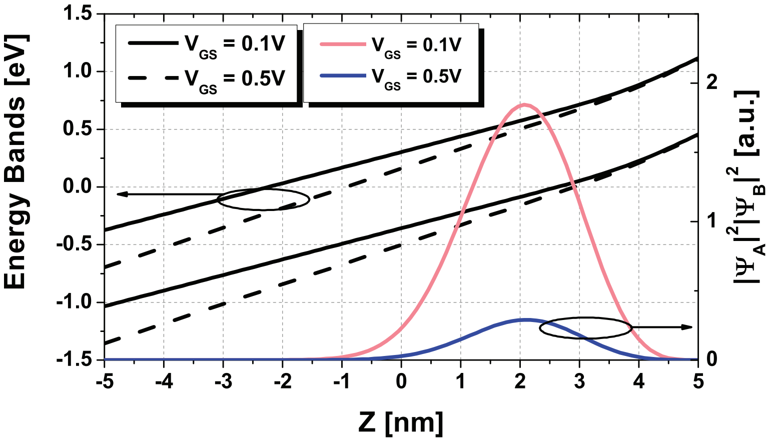


Figure 4.4 – Impact of the gate voltage on the wavefunctions overlap. Energy band diagrams (black lines, left y-axis) and product between the electron and hole wavefunctions squared (right y-axis). A sizable decrease in the WF product when V_{GS} increases is evident stemming from the wavefunctions being pushed towards the oxide of the Ge EHBTFET of Fig. 4.2.

hole and electron subband energies.

One point worth noting with regard to the simulation tools is the difference between the semi-classical Sentaurus Device and the quantum mechanical model simulations (Fig. 4.6). It can be observed that although Sentaurus assumes a bulk, continuous band structure rather than quantized subbands, the semi-classical model is still able to capture an important part of the device operating principle. The reason is that, even in the continuous band case, there is no current conduction if the conduction and valence bands are not aligned. Once the alignment is achieved, tunneling is eventually allowed and a step-like switching behavior is seen. The step-like behavior is caused by the switching from an OFF-state to an ON-state with relatively short tunneling distance. However, since the energy bands are not quantized in Sentaurus, the alignment voltage is drastically different, which is consistent with the results of Vandenberghe et al. [65]. The model parameters for direct and phonon-assisted BTBT in the semi-classical simulations have been taken from [13] and [12] respectively. We reiterate once again that the behavior in Fig. 4.6 depicts an ideal case where lateral leakage effects that we will show to be critical is not included.

Fig. 4.7 presents the transfer characteristics for different V_{DS} voltages. What is most striking is the almost overlap of the curves for $V_{\mathrm{DS}} \geq 0.2\mathrm{V}$. Up to 0.2V, the increase in V_{DS} effectively increases the DOS that can contribute to tunneling. But further increasing V_{DS} does not result in an increase in the tunneling current, unless the increase allows for another alignment between energy subbands. This can be explained by the previously mentioned wavefunction repulsion mechanism, which decreases the tunneling probability albeit the increase in the allowed energy window, resulting in negative trans-conductance at high V_{GS}. In the case of $V_{\mathrm{DS}} = 0.2\mathrm{V}$, we observe an $I_{\mathrm{ON}}/I_{\mathrm{OFF}}$ ratio of 10^7 and an average subthreshold slope of 30$^{\mathrm{mV}}/_{\mathrm{dec}}$

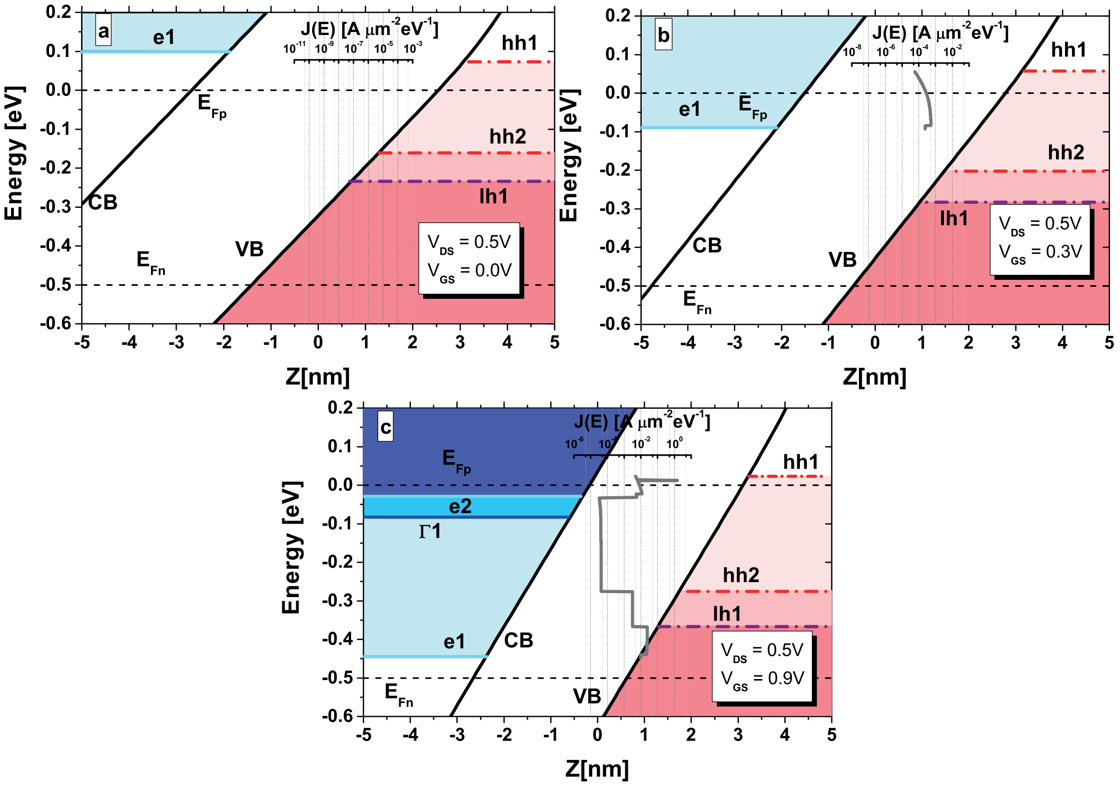


Figure 4.5 – Quantized energy levels for various gate biases for the Ge EHBTFET of Fig. 4.2. The current density spectra are given as insets for each case. Various subband energies for heavy hole (hh), light hole (lh), *L* (e) and Γ (Γ) valley electrons. Spectra reveal that phonon-assisted tunneling between e-lh occurs at a higher rate per energy slice. Note the spike (Dirac-delta) in the current spectrum for $V_{GS} = 0.9$V corresponding to the single tunneling energy in direct 2D-2D tunneling as a result of simultaneous energy and in-plane momentum conservation.

up to V_{DD} that we will see is over optimistic since we are here neglecting the parasitic lateral tunneling paths.

The output characteristics is given in Fig. 4.8 for different V_{GS} voltages. In line with Fig. 4.7, we observe a smaller saturation current for $V_{GS} = 0.5$V than for $V_{GS} = 0.4$V, due to the negative transconductance effect. The quasi-ideal saturation characteristic is a limitation of the model, since it calculates the current using 1D vertical slices of the device and 2D effects are not taken into account. Nevertheless, full 2D Sentaurus simulations have already shown that an extremely good saturation behavior can be achieved [97] and there is no reason to think that vertical quantization would change this property. Therefore, a high output resistance is expected from the EHBTFET in the saturation region. Also, the absence of a superlinear onset [16, 121] at low V_{DS} is a result of subband alignment occurring at energies sufficiently larger than the source Fermi level E_{Fp}.

One important distinction must be made regarding the EHBTFET compared to MOSFET

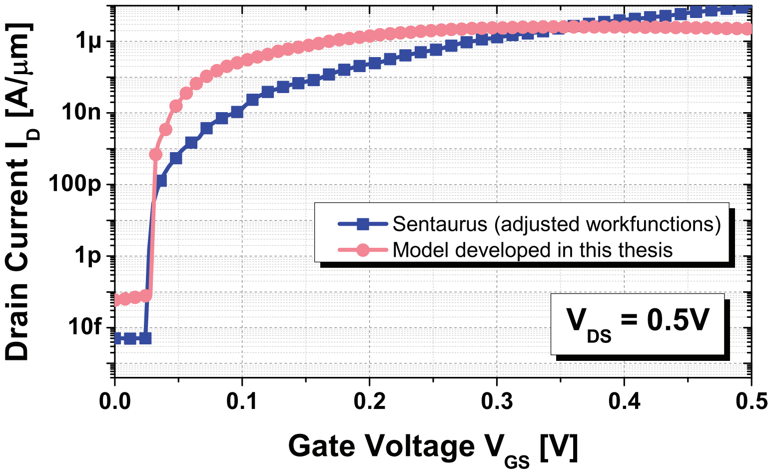


Figure 4.6 – Transfer characteristics of the Ge EHBTFET: comparison between the quantum mechanical model (of section 3.2) and Sentaurus. Sentaurus curve has been rigidly shifted to match the turn on potential calculated by the QM model. Both models predict steep switching behavior, although Sentaurus significantly overestimates the ON current. The workfunctions used for Sentaurus are $\phi_{\text{n-gate}} = 3.955\text{eV}$ and $\phi_{\text{p-gate}} = 4.875\text{eV}$. The workfunctions used for QM model are the same as Fig. 4.3 ($\phi_{\text{n-gate}} = 3.408\text{eV}$ and $\phi_{\text{p-gate}} = 5.642\text{eV}$).

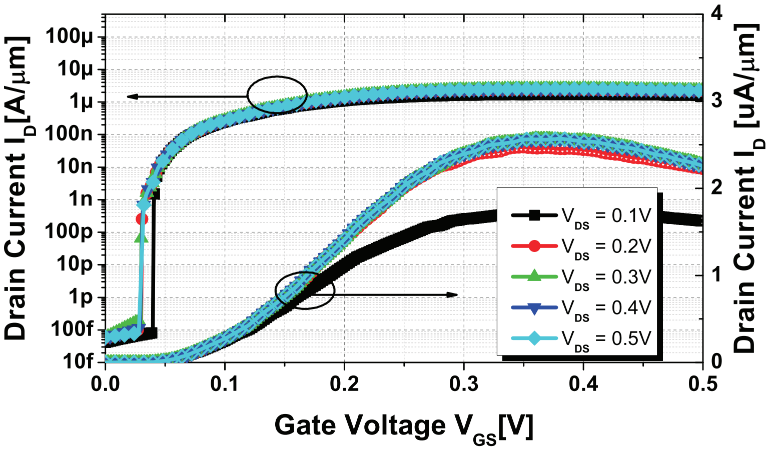


Figure 4.7 – Ge EHBTFET transfer characteristics for different V_{DS} in linear and log scale. A negative transconductance region is visible at high V_{GS} voltages.

regarding the output characteristics under negative V_{DS}. Unlike MOSFET, where the output characteristics is symmetric with respect to the negative and positive V_{DS}, the EHBTFET might suffer from the unidirectional conduction that is shared by all TFETs. The EHBTFET consists of a p-i-n profile along the source-drain direction which is reverse biased in the intended mode of operation (i.e. when $V_{DS} > 0V$). In the case where the junction is forward biased ($V_{DS} < 0V$), the device might enter into the NDR region where the current levels degrade significantly. Fig.

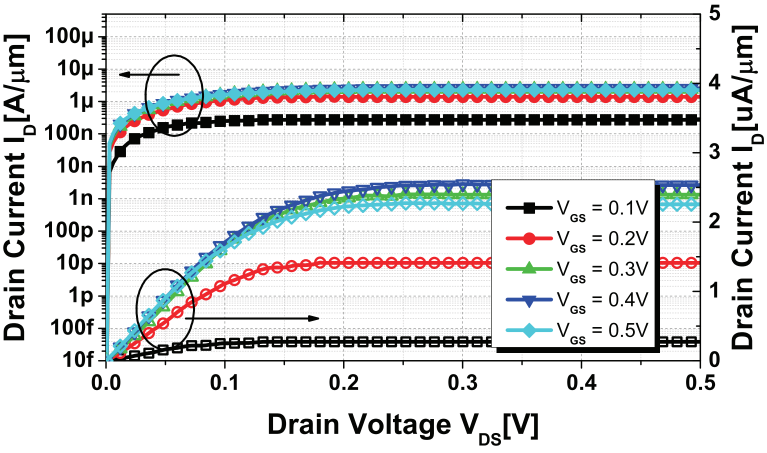


Figure 4.8 – Ge EHBTFET output characteristics for different V_{GS}. Note the absence of the superlinear region at low V_{DS} region, which is commonly seen in TFET output characteristics [16].

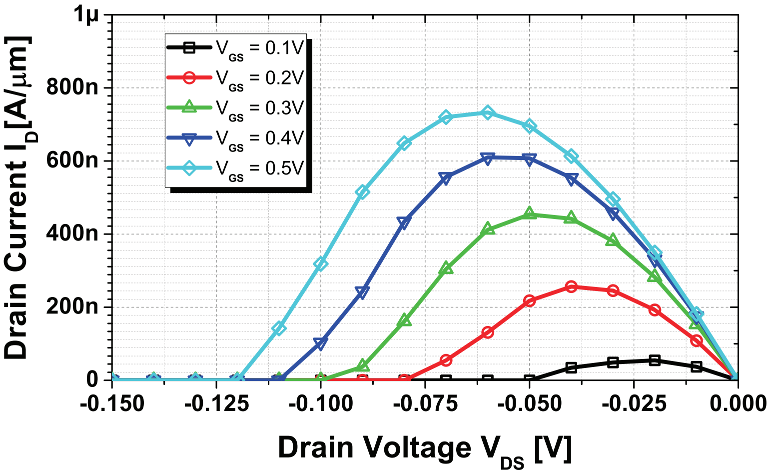


Figure 4.9 – Ge EHBTFET output characteristics for different V_{GS} under negative V_{DS}. The width of the NDR region increases when increasing the gate voltage.

4.9 shows the output characteristics of the EHBTFET under negative V_{DS} where the diode is forward biased. We can clearly observe an increase in the peak voltage and peak current with increasing gate voltage as the gate voltage increases the overlap of the energy bands. It should be noted that the forward diode current of the EHBTFET is not included in our simulator, therefore an exponential increase in current is expected as the V_{DS} is further lowered.

Through 1D quantum mechanical simulations, although reduced ON current levels are obtained compared to CMOS, the EHBTFET is still shown to be a promising candidate for low supply voltage applications due to its steep switching resulting in $SS_{avg} \sim 30^{mV/dec}$ and

$I_{ON}/I_{OFF} \approx 10^7$ ($V_{DD} \approx 0.2$V). Furthermore, it has been demonstrated that, due to the direct (Γ) bandgap being comparable energetically to the indirect (L) bandgap, both phonon-assisted BTBT and direct BTBT can be observed albeit direct BTBT kicks in at much larger voltages compared to the bulk case.

4.2 Two Dimensional Effects on the EHBTFET

As stated in the previous section, the initial results using the 1D model have shown promising results although requiring significantly larger applied gate bias compared to semi-classical simulations [115]. However, 2D effects also need to be investigated. We now turn our attention to 2D simulations of the Ge EHBTFET to identify if any 2D effect is prominent. We again deal with Germanium based devices, in which both direct and phonon-assisted BTBT can be observed due to the fact that the direct Γ valley is separated by only 0.14eV from the indirect L valleys [13]. To this end, we have utilized the models for both direct and phonon-assisted tunneling detailed in Chapter 3.

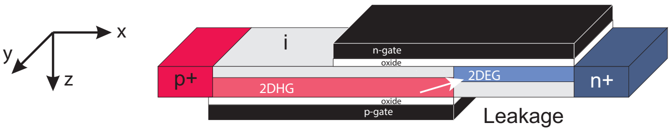


Figure 4.10 – Possible leakage path during OFF state in the EHBTFET of Fig. 4.2.

For the full 2D tunneling simulations, we consider again the EHBTFET structure [122] sketched in Fig. 4.2. The 2D Schenk model is used here for the calculation of the coupling element M_{CV} of direct BTBT, due to the difficulty in separating the hole and electron domains in the Bigelow model for this 2D structure. The source and drain regions are degenerately doped ($N_D = N_A = 10^{20}\text{cm}^{-3}$), whereas the underlap and overlap regions are lightly doped ($N_I = 10^{15}\text{cm}^{-3}$). The n-gate and p-gate workfunctions are $\phi_{n-gate} = 3.434$eV and $\phi_{p-gate} = 5.644$eV, respectively.

Although the tunneling occurring between the 2D gases in the ON state is essentially a 1D flow in the overlap region, it is expected to have some leakage tunneling between the underlap and inversion layers (Fig. 4.10). However, this effect cannot be captured by a 1D model since the electric field direction changes in the spacer regions as a function of device bias. A fully 2D simulator can capture such effects and hence provide more realistic results. We again remind that, unless otherwise stated, we have set the source-drain length to 200nm in order to reduce the impact of artificial quantization caused by the closed boundaries in the Schrödinger equation.

The transfer characteristic of the Ge EHBTFET is given in Fig. 4.11: phonon-assisted (indirect) and direct components are plotted separately. Unless otherwise stated, the p-gate (hole gate) has been fixed to 0V, whereas the n-gate (electron gate) was swept. As can be seen, in the 2D model the SS is degraded compared to the estimate provided by the 1D model, which predicts

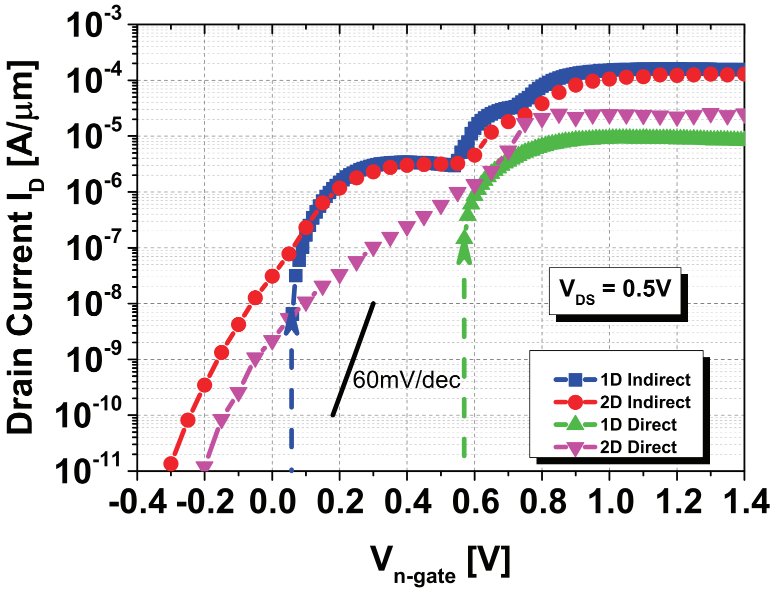


Figure 4.11 – Comparison of direct and indirect components of the Ge EHBTFET of Fig.4.2 with different gate metal workfunctions ($\phi_{\text{n-gate}} = 3.434\text{eV}$ and $\phi_{\text{p-gate}} = 5.642\text{eV}$). It is seen that at high $V_{\text{n-gate}}$, where the device is in ON state, 1D and 2D tunneling models have mutual agreement for both indirect and direct tunneling components. $V_{\text{p-gate}} = 0V, V_{\text{DS}} = 0.5\text{V}, \text{EOT} = 0.53\text{nm}$.

a 'quasi-infinite' point switching slope (indicated by a dashed arrow) due to the fact that it cannot account for non-vertical (leakage) tunneling. This leakage is caused by the penetration of the wavefunctions into the underlap region well before the electron inversion layer has been formed in the overlap region. This penetration acts as a leakage path (see Fig. 4.10) which can be better observed in Fig. 4.12. Fig. 4.12(c) shows the integrand for the electron-hole coupling element M_{CV} integral in eq. 3.46 for the electron-hole wavefunction pair having the highest contribution for OFF state of the EHBTFET. Since the p-gate gate voltage is low enough to attract the wavefunctions deep into the overlap region (see Fig. 4.12(b)), the heavy hole wavefunction is able to penetrate up to the drain side of the channel. However, the Γ electron wavefunction can only penetrate into the underlap region and not into the overlap region, as seen in Fig. 4.12(c). The reason for this behavior is that electric field magnitude is smaller in the underlap region than the overlap region, which results in weaker field-induced quantization in the underlap region and permits wavefunction penetration from the drain to this region.

Next, we turn to investigate the different natures of the ON state and OFF state currents. As a form of classification, OFF state current can be described as *point tunneling* [45], which is the tunneling type seen in conventional TFETs. The fact that the current generation is very narrowly concentrated around the underlap-overlap boundary (see Fig. 4.12) is consistent with this classification.

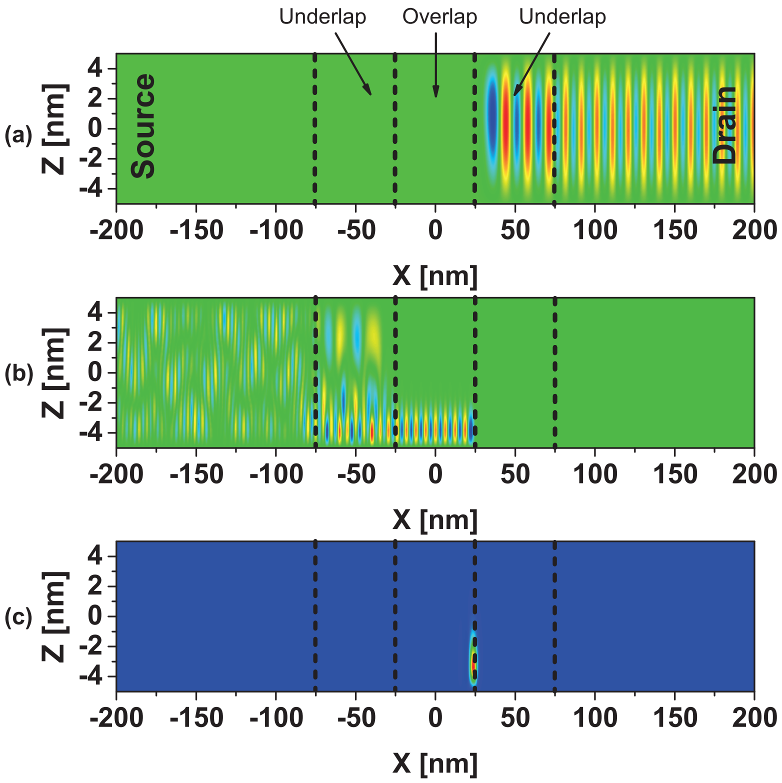


Figure 4.12 – (a) Γ electron wavefunction (b) Heavy hole wavefunction (c) Generation rate for the electron-hole subband pair that contributes the most to the direct BTBT current in the OFF state ($V_{\mathrm{n-gate}} = 0\mathrm{V}$, $V_{\mathrm{p-gate}} = 0\mathrm{V}$, $V_{\mathrm{DS}} = 0.5\mathrm{V}$) for the Ge EHBTFET of Fig. 4.2. Note the penetration of the Γ electron wavefunction into the underlap region in (a). The concentration of the BTBT generation around a narrow peak is a strong indicator of point tunneling.

In contrast, the ON state could be classified as *line tunneling* [45] which is the name given to tunneling aligned to the electric field induced by the gate. By definition line tunneling is proportional to the tunneling area, which is consistent with the findings in Fig. 4.13, where it is seen that the majority of the overlap region contributes to BTBT.

Figs. 4.14 & 4.15 investigate the geometrical dependences of the BTBT in the Ge EHBTFET. Fig. 4.14 plots the transfer characteristics for different EOT at both n-gate and p-gates. It is important to note that in Fig. 4.14 the phonon-assisted BTBT is already in the ON state at about $V_{\mathrm{n-gate}} = 0.1\mathrm{V}$, due to large effective mass of L valley electrons, whereas direct BTBT transition to ON state is observed at relatively higher n-gate voltages. The onset of ON state for direct BTBT corresponds to the inflection points in the I-V characteristics e.g., $V_{\mathrm{n-gate}} \simeq 0.9\mathrm{V}$ for EOT = 0.7nm.

Bearing these in mind, we can qualitatively compare the features of line and point tunneling by comparing the phonon-assisted BTBT (ON state, line tunneling) and direct BTBT (OFF state, point tunneling) curves. The most significant feature of the figures is the drastic difference

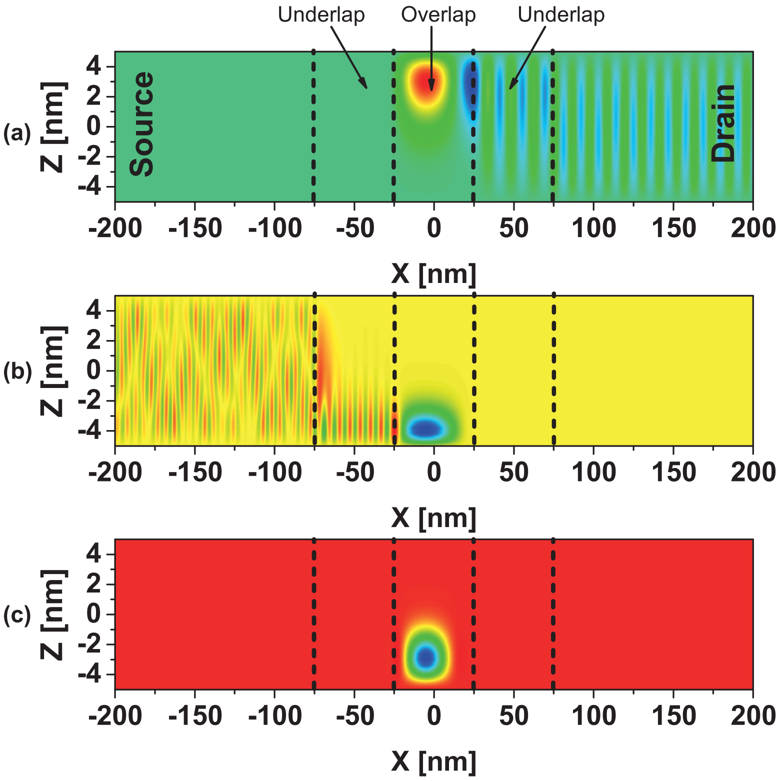


Figure 4.13 – Same as Fig. 4.12 but for the ON state ($V_{\text{n-gate}} = 2\text{V}$, $V_{\text{p-gate}} = 0\text{V}$, $V_{\text{DS}} = 0.5\text{V}$). As expected from line tunneling, the BTBT generation is spread throughout the whole overlap region.

in the dependences of ON state and OFF state. Observing Fig. 4.14, one sees that the EOT does not play an important role in determining the magnitude of the current in the ON state but merely shifts the transfer curves. In the OFF state, however, it is found out that the EOT is effective in determining the magnitude of the current as well as the tunneling onset as we see (in Fig. 4.14) a monotonic increase of the OFF current with decreasing EOT. Decreasing EOT also slightly improves the switching slope.

The role of the channel thickness (Fig. 4.15) seems to be more complex compared to the influence of the oxide thickness. It is seen that the shapes of the transfer characteristics are drastically different due to dramatic change in the subband energies for different channel thicknesses. As for the direct BTBT, it can be said that the $T_{\text{CH}} = 7.5\text{nm}$ case is too extreme in the sense that no direct BTBT ON current is seen. The reason for such behavior is that as the channel thickness is reduced, the subband energies in the overlap region for the Γ electron band increases at a higher rate than those of the L valley due to the low effective mass of Γ electron valley. Consequently, as the gate voltage is increased L valley states in the overlap region becomes occupied and charge density increases which pins the potential profile. Once the potential profile is pinned, the gate induced electric field is screened and

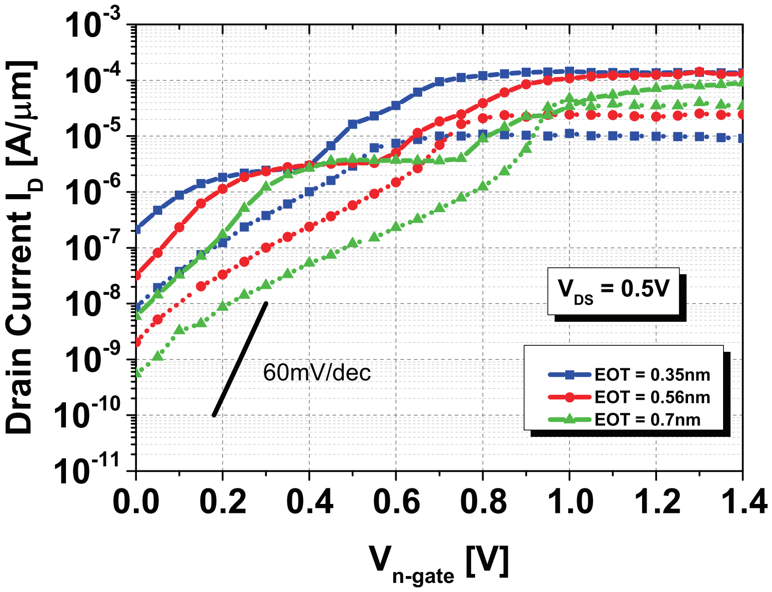


Figure 4.14 – Transfer characteristics of the Ge EHBTFET of Fig. 4.2 varying the gate oxide thicknesses T_{OX} at both gates. Phonon-assisted (direct) BTBT are denoted separately as solid (dotted) lines. $V_{\mathrm{DS}} = 0.5\mathrm{V}$ and $V_{\mathrm{p-gate}} = 0\mathrm{V}$. $T_{\mathrm{CH}} = 10\mathrm{nm}$.

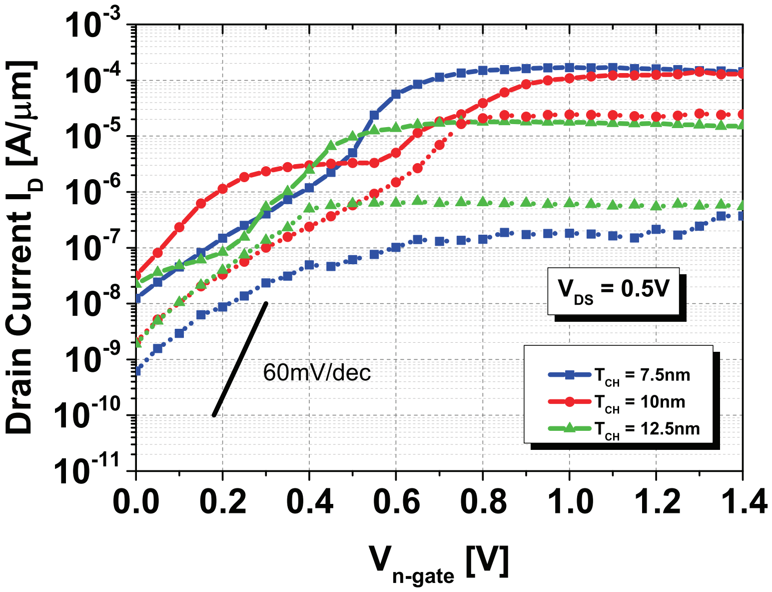


Figure 4.15 – Transfer characteristics of the Ge EHBTFET of Fig. 4.2 varying the channel thicknesses T_{CH}. Phonon-assisted (direct) BTBT are denoted separately as solid (dotted) lines. $V_{\mathrm{DS}} = 0.5\mathrm{V}$ and $V_{\mathrm{p-gate}} = 0\mathrm{V}$. EOT = 0.53nm.

subband efficiency [123] decreases rapidly as seen in Fig. 4.3.

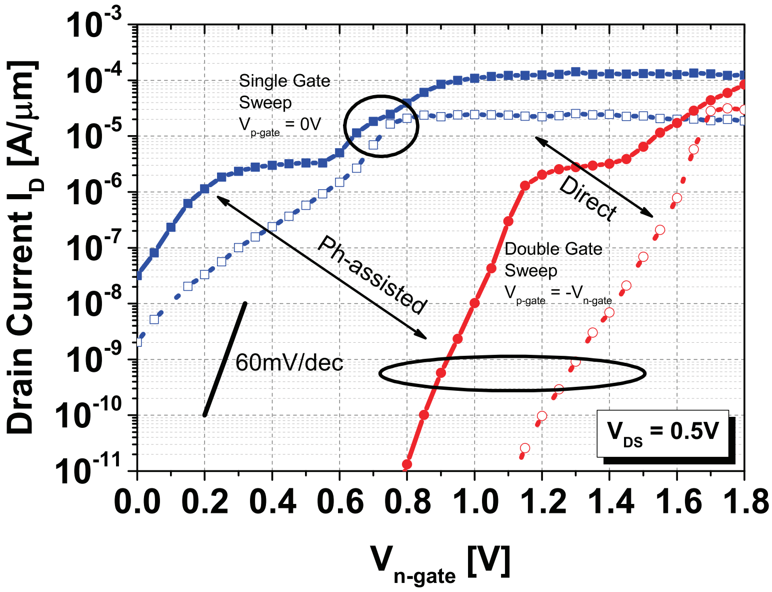


Figure 4.16 – Transfer characteristic for the Ge EHBTFET with different biasing schemes. Direct and phonon-assisted BTBT are denoted separately as lines with filled and hollow symbols respectively. For this simulation the p-gate workfunction is taken to be $\phi_{\text{p-gate}} = 3.64$eV, $V_{DS} = 0.5$V.

In Fig. 4.16 the impact of different biasing schemes is investigated. For this simulation, the n-gate and p-gate electrodes are swept simultaneously but in opposite polarities. Note that for this simulation the p-gate workfunction has been taken as $\phi_{\text{p-gate}} = 3.64$eV, whereas n-gate workfunction was kept at 3.434eV. The current levels rapidly drop off for both direct and phonon-assisted BTBT as $V_{\text{n-gate}}$ is lowered with respect to ON-conditions. This reduces the leakage and improves SS at low currents.

4.3 Device Optimization and Lateral Leakage Suppression

In the previous section, we have shown that a leakage path exists in the EHBTFET due to the penetration of the wavefunction (WF) into the so-called underlap region (see Fig. 4.10). This is due to the field-induced quantization being considerably weaker in the underlap region which is controlled by only one gate. This WF penetration causes significant lateral leakage with subsequent SS degradation. In this section, we propose two different methods [124, 125] to suppress this penetration and hence lateral leakage. Both methods require slight modifications in the device structure and each of them can be chosen depending on what the fabrication process allows. Also, both methods present negligible penalty on the ON current when designed properly.

4.3.1 Counterdoping as a Method for Lateral Leakage Suppression

We first detail the counterdoping method, which introduces dopingin the underlap regions. We simulate an $In_{0.53}Ga_{0.47}As$ EHBTFET oriented along the [100] direction. Since InGaAs is a direct gap material, only direct BTBT is included. The bandgap is 0.81eV and the effective masses are $0.042m_0$, $0.0503m_0$ and $0.465m_0$ for electrons, light holes and heavy holes, respectively [126, 127]. The gate oxide has EOT=1nm at both interfaces, which corresponds to roughly 5.9nm of $HfO_2 (\epsilon = 23)$. We have checked that the magnitude of the electric field inside the oxide is well below the breakdown field for HfO_2 ($E_{bd} \sim 4$MV/cm) reported in [128]. The source and drain regions are doped p-type and n-type respectively with a doping level of 10^{20}cm^{-3}. We have checked the impact of incomplete activation of dopants by simulating a case with 10^{19}cm^{-3} doping in the source and drain regions and seen that the device characteristics are not affected. We assume midgap workfunctions for both n-gate and p-gates. Unless otherwise stated, a $V_{DS} = 0.25$V is used.

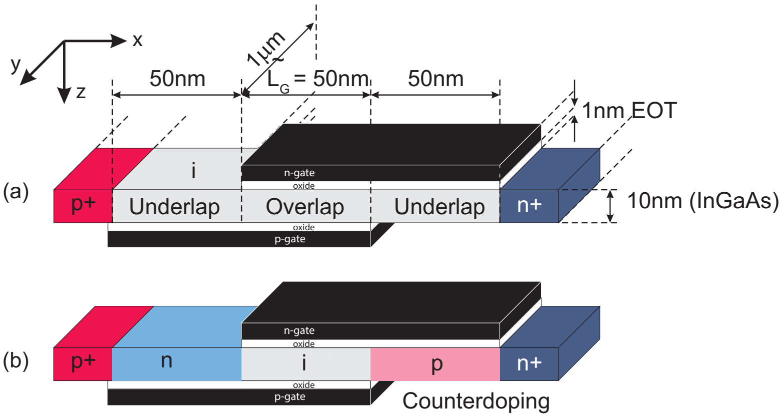


Figure 4.17 – (a) EHBTFET device parameters (b) The proposed solution featuring doped underlap regions (counterdoping).

The simulation and device optimization strategy is as follows: first the 1D version of the quantum mechanical simulator (section 3.2) [115] is utilized to perform a parameter sweep on the channel thickness while fixing the top and bottom oxide thicknesses at 1nm EOT, as given in Fig. 4.18. For this figure, n-gate and p-gate biases are swept simultaneously in opposite polarities (i.e., $V_{p-gate} = -V_{n-gate}$). The figures of merit that are of interest are the subband alignment voltage ($V_{align} = V_{n-gate} - V_{p-gate}$) and the ON current (defined as I_D at $V_{n-gate} = \frac{V_{align}}{2} + 0.25\text{V}$). The figure shows the alignment voltage and ON current levels obtained for different channel thicknesses T_{CH}. It is seen that $T_{CH} = 10$nm seems to offer a reasonable compromise between the alignment voltage (with a required bias of about 3.3V) and the ON current (about 10μA/μm). For the following results, unless otherwise stated, a channel thickness of 10nm is used.

The main idea of using counterdoping in the underlap regions is to partially screen the influence of the gate. Fig. 4.19 reports the conduction band profile cut along the x-direction at

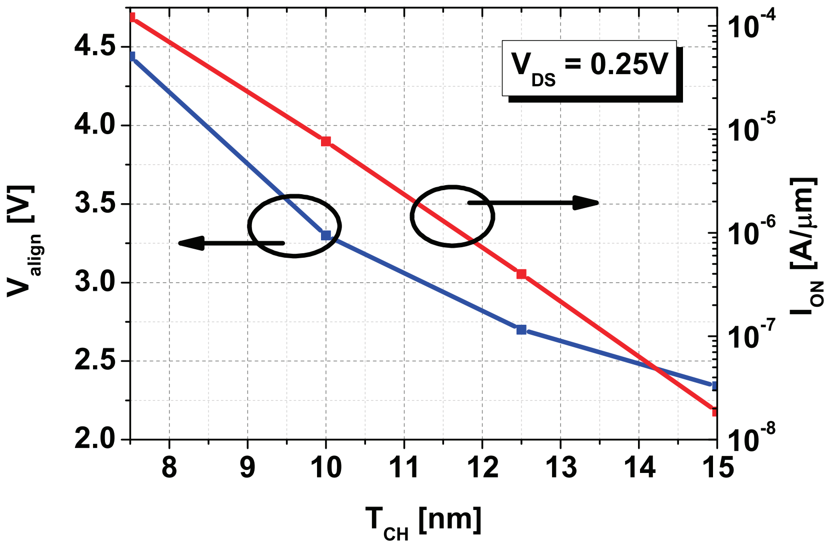


Figure 4.18 – Alignment voltage V_{align} and ON current levels for $In_{0.53}Ga_{0.47}As$ EHBTFETs with different channel thicknesses T_{CH}. EOT = 1nm.

the middle of the semiconductor film. Using counterdoping results in a significant increase of the energy barrier for both holes and electrons in the underlap regions at source and drain sides respectively, which delays the onset of the parasitic lateral tunneling component. Note that counterdoping alleviates the short channel effects (SCE) in the overlap region, by flattening the potential profile around the overlap-underlap boundary. Due to this, counterdoping is also expected to help with the lateral scaling of the EHBTFET by suppressing the SCE. For all the figures shown in this section, the counterdoping is applied only to the underlap region at the drain side (p-doped region in Fig. 4.17(b)) while the underlap region at the source side is left undoped. This is due to the fact that for the devices shown here, we use a p-gate bias that already induces the 2DHG in the overlap region, therefore only the penetration of the electrons from the drain needs to be suppressed.

The effect of the increase in energy barrier is easily observed in the BTBT generation rates given in Fig. 4.20, which indicate orders of magnitude lower rates for counterdoped EHBTFET in the OFF state ($G_{\text{BTBT,max}} \sim 10^{46}\text{m}^{-3}\text{s}^{-1}$ with no counterdoping, compared to $G_{\text{BTBT,max}} \sim 10^{34}\text{m}^{-3}\text{s}^{-1}$ with counterdoping) which shows the efficiency of the method.

Fig. 4.21 shows the transfer characteristics using different underlap doping levels (denoted as N_{UL}), where the p-gate bias is fixed at −1.4V while sweeping the n-gate. In the optimized case, $N_{\text{UL}} = 1.5 \times 10^{19}\text{cm}^{-3}$, an ON current of about 20μA/μm and an average SS of 31mV/dec over eight decades of current are obtained in this particular device. The OFF current is taken to be 0.1pA/μm (at the n-gate voltage denoted as V_{OFF}) while the ON current is the drain current at $V_{\text{n-gate}} = V_{\text{OFF}} + 0.25\text{V}$.

The optimum counterdoping value is expected to depend on the channel thickness, as the

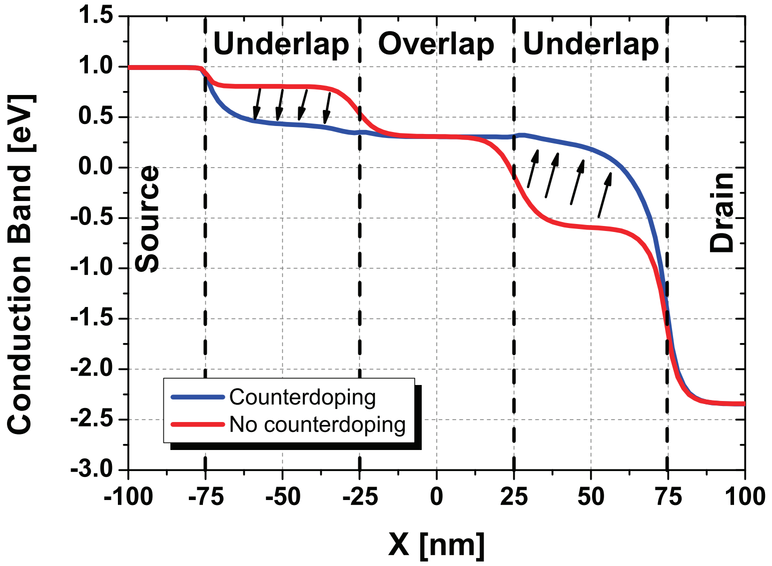


Figure 4.19 – Conduction band profile cut through the middle of the channel along the x-direction for InGaAs EHBTFETs with (blue) and without (red) counterdoping ($N_{\mathrm{UL}} = 1.5 \times 10^{19}\mathrm{cm}^{-3}$ at both underlap regions) in the OFF state ($V_{\mathrm{n-gate}} = 1.4\mathrm{V}$, $V_{\mathrm{p-gate}} = -1.4\mathrm{V}$, $V_{\mathrm{DS}} = 0.25\mathrm{V}$).

latter impacts the degree of quantization. To investigate this point, an InGaAs EHBTFET as in Fig. 4.17 but with $T_{\mathrm{CH}} = 7.5\mathrm{nm}$ is simulated and the transfer characteristics are reported in Fig. 4.22(Left) using different underlap doping levels. For this set of curves, the p-gate bias is fixed at −2.0V while sweeping the n-gate. It is seen that the optimum doping level which suppresses the lateral leakage while preserving the ON current is $N_{\mathrm{UL}} = 3.5 \times 10^{19}\mathrm{cm}^{-3}$, which is drastically higher than the value found for $T_{\mathrm{CH}} = 10\mathrm{nm}$ although the value may slightly differ in different biasing configurations. The optimized configuration in this case yields slightly improved (compared to 10nm) SS and ON current values ($I_{\mathrm{ON}} = 100\mu\mathrm{A}/\mu\mathrm{m}$ and an average SS of 27mV/dec over ten decades of current). Similarly, thicker channels require less doping e.g., the optimum doping is $N_{\mathrm{UL}} = 7.5 \times 10^{18}\mathrm{cm}^{-3}$ for $T_{\mathrm{CH}} = 15\mathrm{nm}$ (4.22(Right)).

The plots in Fig. 4.21 & 4.22 point out the following key insights:

- Too low underlap doping results in ineffective suppression of lateral leakage (e.g., see the case with $N_{\mathrm{UL}} = 1 \times 10^{19}\mathrm{cm}^{-3}$ for $T_{\mathrm{CH}} = 10\mathrm{nm}$)
- Too high underlap doping, in turn, results in the source and drain regions being 'cut off' from the channel due to increased barrier height in the underlap region. This increase causes certain states in the overlap region to become quasi-bound so that they do not contribute to the current. This results in lower ON current.

A peculiar increase in OFF current is observed for $N_{\mathrm{UL}} = 2\times10^{19}\mathrm{cm}^{-3}$ and $N_{\mathrm{UL}} = 4.5\times10^{19}\mathrm{cm}^{-3}$

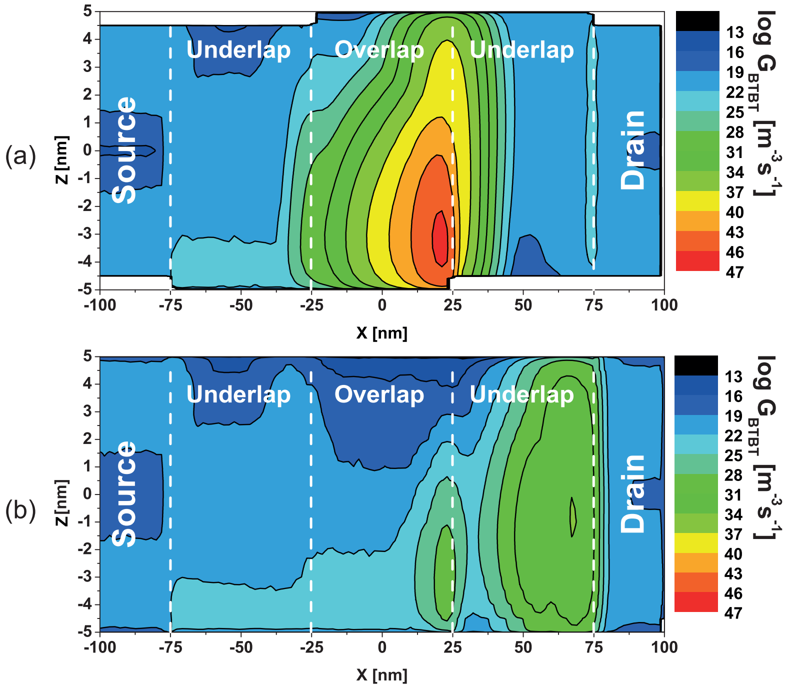


Figure 4.20 – BTBT generation rates for the OFF state ($T_{\mathrm{CH}} = 10\mathrm{nm}$, $V_{\mathrm{n-gate}} = 1.4\mathrm{V}$, $V_{\mathrm{p-gate}} = -1.4\mathrm{V}$, $V_{\mathrm{DS}} = 0.25\mathrm{V}$) of the InGaAs EHBTFET of Fig. 4.17 (a) without counterdoping (b) with counterdoping ($N_{\mathrm{UL}} = 1.5 \times 10^{19}\mathrm{cm}^{-3}$).

for the 10nm and 7.5nm devices, respectively. This is caused by the induced ambipolar conduction that occurs due to WF penetration *from the overlap region* to the underlap region, in contrast to penetration from the drain in the case without counterdoping [129]. As seen in Fig. 4.23, using a too high doping could lower the barrier for hole states in the overlap regions and help them penetrate into the underlap regions controlled by the n-gate only. Moreover, highly doped regions become less sensitive to the gate control. These findings indicate that the control of the doping level is critical for the effectiveness of the counterdoping.

We also study the parameter sensitivity of the counterdoping approach by simulating cases where the border of the counterdoped region is not aligned with the actual border between the underlap and overlap regions. We denote the cases where the counterdoping extends into the overlap region with a positive value of misalignment, whereas a negative value is used for the opposite case. Fig. 4.24 (Left) compares the conduction band profiles along the x-direction at the middle of the channel. One can see indeed that the barrier height is increased for the positive misalignments and vice versa. The results given in Fig. 4.24 (Right) indicate that the counterdoping method retains its effectiveness over a range of misalignments up to -3nm to +3nm. We observe two different detrimental impacts on the device characteristics for negative and positive misalignment cases. For negative misalignments, the lateral leakage is not suppressed as effectively as in the perfectly aligned case; hence resulting in increased OFF current and poorer SS values. For positive misalignments, on the other hand, we observe that

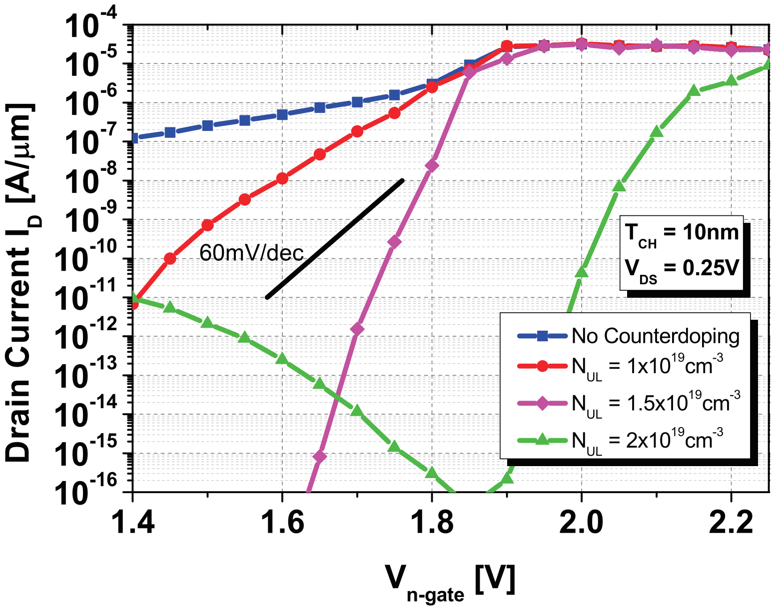


Figure 4.21 – Transfer characteristics for the InGaAs EHBTFET of Fig. 4.17 with different underlap doping levels with T_{CH} = 10nm. V_{DS} = 0.25V, $V_{\mathrm{p-gate}} = -1.4$V. Note the sudden increase in OFF current for the $N_{\mathrm{UL}} = 2 \times 10^{19}\mathrm{cm}^{-3}$ curve, caused by a second leakage path.

the ON current decreases and the SS get deteriorated due to increased barrier height in the counterdoped region as already observed in Fig. 4.24 (Left).

A similar analysis has been performed varying the abruptness of the doping profiles by simulating a non-abrupt junction between the counterdoped underlap and the overlap region with various doping decay rates. As can be seen in Fig. 4.25(Left & Right), for large decay rates (> 5nm/dec), an effect similar to the positive misalignment cases (i.e., as in Fig. 4.24 (Right)) is present. Such non-abrupt doping profiles increase the potential barrier height and therefore reduce the ON current. These findings highlight the need for a very good process control in the alignment and anneal steps to ensure sufficiently abrupt junctions.

These critical aspects of the device processing can be addressed to some extent by the template-assisted selective epitaxy (TASE) growth technique recently developed [130]. This method allows horizontal epitaxial growth of fin structures by defining an oxide template. Such a method can be used to obtain the regions with alternating doping profiles with abrupt junctions that is required by the counterdoped EHBTFET without any implantation steps, which could be used for implementing the counterdoped EHBTFET on a fin.

4.3.2 Hetero-gate Metal for Lateral Leakage Suppression

A similar effect in the band profile inside the channel can also be achieved by utilizing the so-called 'hetero-gate' approach (Fig. 4.26) [125]. The idea is similar to [131] where a combination

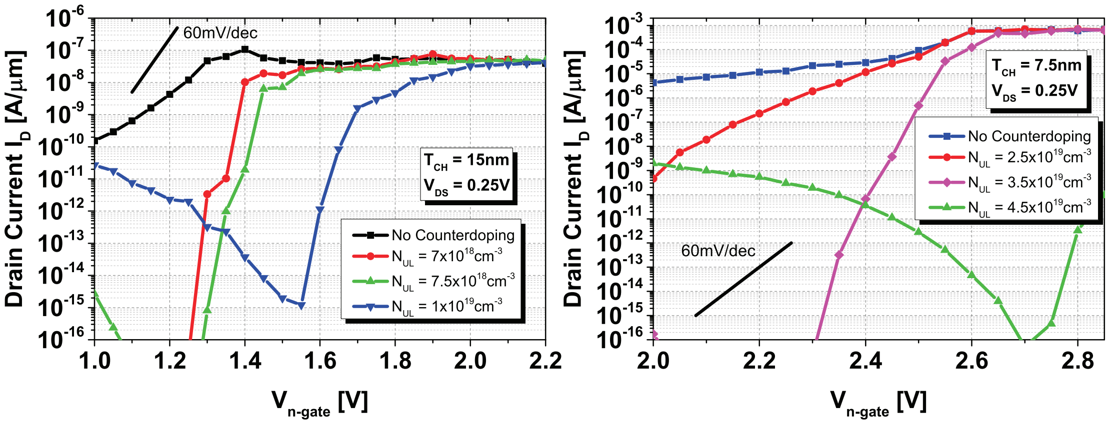


Figure 4.22 – Transfer characteristics for an InGaAs EHBTFET as in Fig. 4.21 but with (Left) $T_{\mathrm{CH}} = 7.5$nm and $V_{\text{p-gate}} = -2$V (Right) $T_{\mathrm{CH}} = 15$nm and $V_{\text{p-gate}} = -1$V with different underlap doping levels. $V_{\mathrm{DS}} = 0.25$V for both figures. An increase in the OFF current (similar to Fig. 4.21) is obverved for $N_{\mathrm{UL}} = 4.5 \times 10^{19}\mathrm{cm}^{-3}$ ($T_{\mathrm{CH}} = 7.5$nm) and $N_{\mathrm{UL}} = 1 \times 10^{19}\mathrm{cm}^{-3}$ ($T_{\mathrm{CH}} = 15$nm).

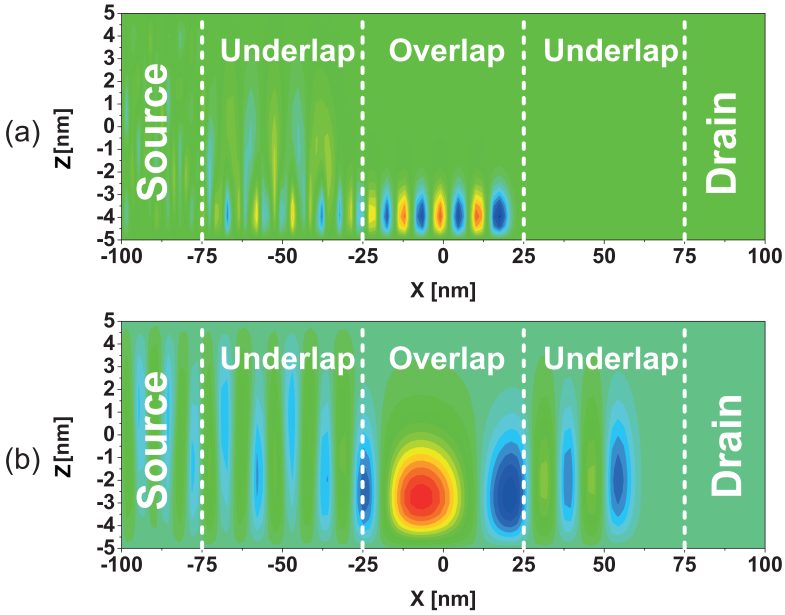


Figure 4.23 – Heavy Hole wavefunction of the subband with highest contribution to the OFF current ($V_{\text{n-gate}} = -V_{\text{p-gate}} = 1.4$V, $V_{DS} = 0.25$V) for the counterdoping levels of (a) $N_{\mathrm{UL}} = 1 \times 10^{19}\mathrm{cm}^{-3}$ (b) $N_{\mathrm{UL}} = 2 \times 10^{19}\mathrm{cm}^{-3}$ in the InGaAs device of Fig. 4.17.

of high-k and low-k gate oxides were used to reduce the gate capacitances to improve the switching performance. Here, however, we are interested in shifting the band profile in the underlaps to prevent early penetration of the wavefunction. To do this, a different gate metal is utilized for the underlap of the n-gate where a more p-like workfunction metal is used.

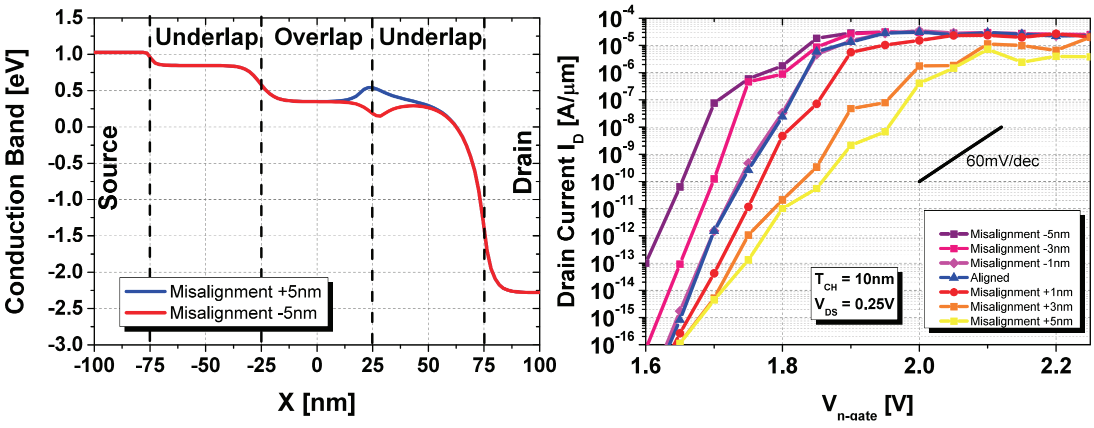


Figure 4.24 – (Left) Conduction band profile cut through the middle of the channel along the x-direction for counterdoped InGaAs EHBTFETs with (blue) +5nm (red) -5nm of underlap doping misalignment ($N_{\mathrm{UL}} = 1.5 \times 10^{19}\mathrm{cm}^{-3}$) in the OFF state ($V_{\mathrm{n-gate}} = 1.4\mathrm{V}$, $V_{\mathrm{p-gate}} = -1.4\mathrm{V}$, $V_{\mathrm{DS}} = 0.25\mathrm{V}$). (Right) Transfer characteristics for the counterdoped InGaAs EHBTFET of Fig. 4.17 ($N_{\mathrm{UL}} = 1.5 \times 10^{19}\mathrm{cm}^{-3}$) with misaligned junctions. ($V_{\mathrm{p-gate}} = -1.4\mathrm{V}$, $V_{\mathrm{DS}} = 0.25\mathrm{V}$). Positive values of misalignment indicate that the counterdoped region extends into the overlap region, whereas the negative values indicate the opposite case.

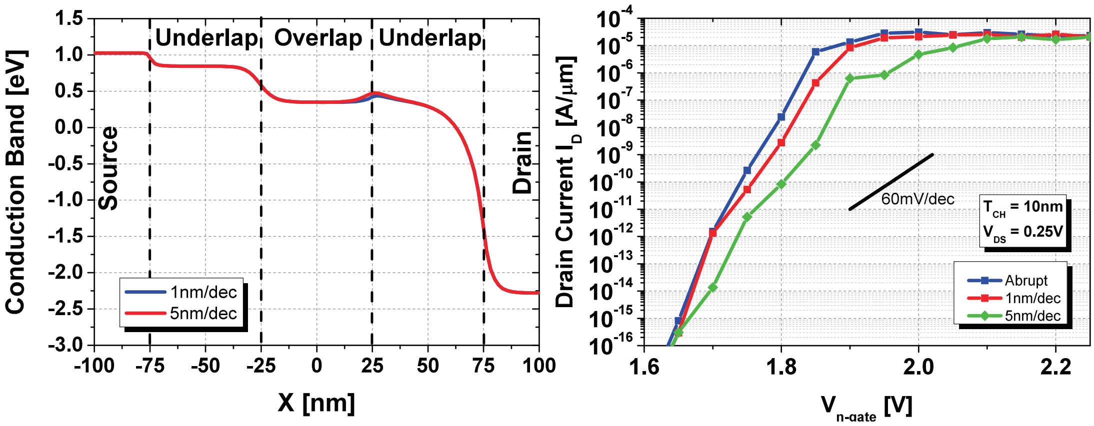


Figure 4.25 – (Left) Conduction band profile cut through the middle of the channel along the x-direction for counterdoped InGaAs EHBTFETs with (blue) 1nm/dec (red) 5nm/dec of underlap doping decay ($N_{\mathrm{UL}} = 1.5 \times 10^{19}\mathrm{cm}^{-3}$) in the OFF state ($V_{\mathrm{n-gate}} = 1.4\mathrm{V}$, $V_{\mathrm{p-gate}} = -1.4\mathrm{V}$, $V_{\mathrm{DS}} = 0.25\mathrm{V}$). (Right) Transfer characteristics for the counterdoped EHBTFET of Fig. 4.17 ($N_{\mathrm{UL}} = 1.5 \times 10^{19}\mathrm{cm}^{-3}$) with nonabrupt junctions. ($V_{\mathrm{p-gate}} = -1.4\mathrm{V}$, $V_{\mathrm{DS}} = 0.25\mathrm{V}$).

Fig. 4.27(Left) depicts the conduction band profiles with and without hetero-gate using our 2D simulator. We can see that indeed an effect similar to the counterdoping solution is present. Fig. 4.27(Right) shows the transfer characteristics for the InGaAs EHBTFET using different

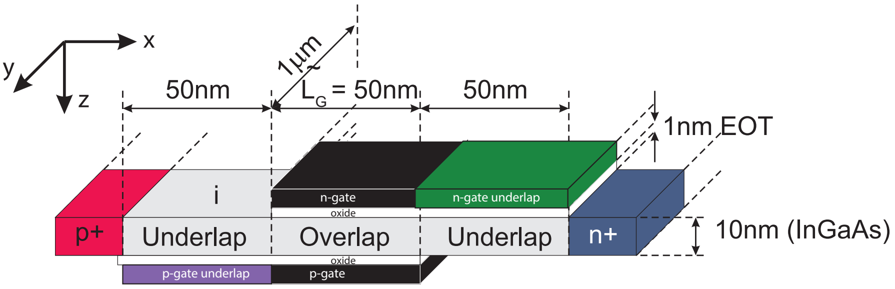


Figure 4.26 – The hetero-gate EHBTFET structure.

workfunction offsets ΔΦ. Noting that we utilize mid-gap metal for the overlap workfunction, it is evident that utilizing a workfunction with an offset of ΔΦ = 1eV is effective in suppressing the lateral leakage. We see the relatively lower offsets are not sufficient to prevent the WF penetration, similar to the lower underlap doping levels for the counterdoping solution. Again, using a too large offset results in a reduction in ON current, due to too much increased potential barrier in the underlap region, effectively cutting off the carriers from the drain region.

In terms of performance, we can argue that both counterdoping and hetero-gate solutions are somewhat equivalent in terms of their average SS and the ON current levels. For both counterdoping and hetero-gate solutions, we see that they entail some additional fabrication complexity. While we have only dealt with the underlap on the drain side for our device-level simulations here, we will see in the following sections regarding the circuit-level simulations that the leakage through the underlap on source side will need to be taken into account. For the counterdoping solution, it will require a doping profile of p-n-i-p-n which is rather complicated, however, possibly achievable in the case of III-V epitaxial growth through *in-situ* doping. On the hetero-gate side, it would entail utilization of four different metal workfunctions and four critical aligment steps which, for a first analysis, might present a more daunting challenge.

4.4 Investigation of the Parameter Space for the EHBTFET

Throughout our 1D and 2D simulations, we have observed some peculiarities on the impact of changes in the effective oxide thickness and the channel thickness of the EHBTFET. In this section, we will tackle this question systematically and propose a simple analytical model to prove our intuitions as well as provide an assessment of the performance of different semiconductors as the channel material for the EHBTFET [132].

As done in the 1D simulations of section 4.1.1, we take a 1D vertical slice in the overlap region and assume translational symmetry along the x-direction (i.e., horizontal potential profile constant in the overlap region). This assumption relies on the fact that the vertical potential profile in the overlap region is very close in both 1D and 2D simulations of section 4.2 and

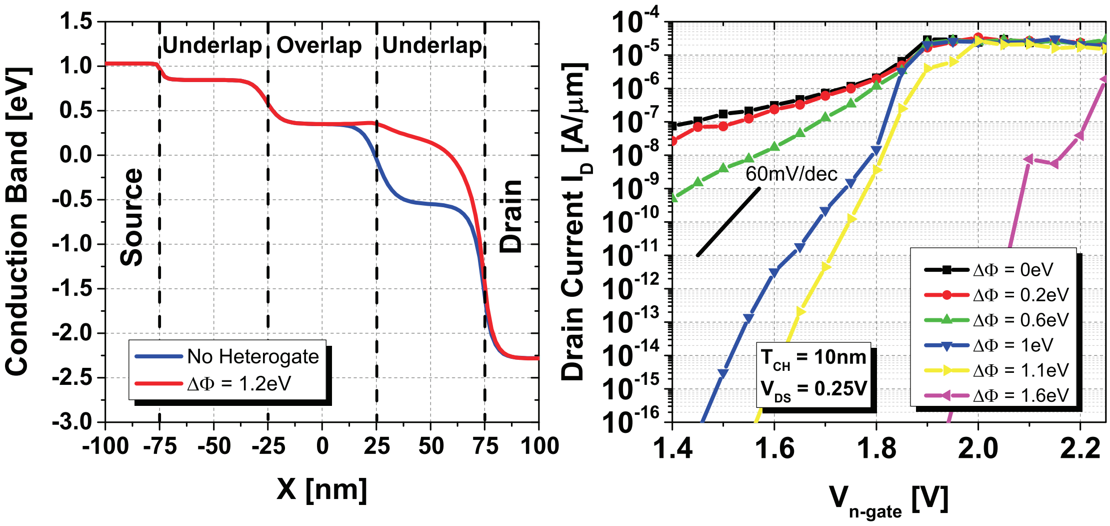


Figure 4.27 – (Left) Conduction band profile cut through the middle of the channel along the x-direction for InGaAs EHBTFETs with (blue) and without (red) hetero-gate ($\Delta\Phi = 1.2$eV) in the OFF state ($V_{\mathrm{n-gate}} = 1.4$V, $V_{\mathrm{p-gate}} = -1.4$V, $V_{\mathrm{DS}} = 0.25$V). (Right) Transfer characteristics for an InGaAs hetero-gate EHBTFET with different underlap gate metal workfunction differences. The overlap region workfunction is the intrinsic level of $\mathrm{In_{0.53}Ga_{0.47}As}$.

lateral effects can be suppressed using hetero-gate configurations (section 4.3.2 [125]) or underlap counterdoping (section 4.3.1 [124]). The parameters used for the simulation are given in Table 4.1. The deformation potential and the phonon energy parameters for Ge are taken from [12]. The parameters for phonon-assisted tunneling (for Si) have been calibrated against experimental data [10, 11] (see Fig. 3.19). The deformation potential for $\mathrm{Si_{0.5}Ge_{0.5}}$ was linearly interpolated from the Si and Ge values.

We perform anti-symmetric gate sweeps (i.e., sweep the n and p gates simultaneously with $V_{\mathrm{n-gate}} = -V_{\mathrm{p-gate}}$), since it was shown to have lower lateral leakage and better subthreshold slopes (see section 4.2 and Fig. 4.16). Midgap gate workfunctions have been considered. Different channel materials (Si, $\mathrm{Si_{0.5}Ge_{0.5}}$, Ge, InAs and $\mathrm{In_{0.53}Ga_{0.5}As}$), film thicknesses ($T_{\mathrm{CH}} = [7.5, 20]$nm) and equivalent oxide thicknesses ($T_{\mathrm{OX}} = [0.5, 3]$nm) are employed. In terms of the channel material, we limit ourselves to homojunction architectures, although significant improvement may be obtained by utilizing staggered and broken gap III-V heterostructures [138, 139].

The device is assumed to be oriented along the [100] direction with [001] surface orientation. For Si, six Δ valleys (four-fold degenerate Δ_4 with transverse mass along the z-direction and two-fold degenerate Δ_2 with longitudinal mass along the z-direction) have been taken into account. Since we are using a mole fraction of 0.5 for SiGe, the effective masses for the Δ valleys are taken the same as in Si [13]. For Ge, four-fold degenerate L valleys and the direct gap Γ

---

[1] Extracted from the $k \cdot p$ results by University of Bologna [8].

Table 4.1 – Bandgaps, effective masses, valley degeneracies, non-parabolicity coefficients along with deformation potentials and phonon energies (phonon-assisted BTBT only) used for numerical simulations. For anisotropic valleys the masses are given as quantization/transverse Mass.

Material (Valley)	E_G[eV]	m_e^*[m$_0$]	m_{hh}^*[m$_0$]	m_{lh}^*[m$_0$]	α_{NP}[eV^{-1}]	D_{ph}[eV/m]	E_{ph}[meV]
Si (Δ_4)	1.12 [13]	0.1905/0.417 [13]	0.49 [13]	0.16 [13]	0.5 [119]	8.16×10^{10} [65]	18.6 [65]
Si (Δ_2)		0.916/0.1905 [133]					
Si$_{0.5}$Ge$_{0.5}$ (Δ_4)	0.91 [134]	0.1905/0.417 [133]	0.41 [13]	0.102 [13]	0.5 [119]	7.98×10^{10} [65, 12]	13.6 [65, 12]
Si$_{0.5}$Ge$_{0.5}$ (Δ_2)		0.916/0.1905[13]					
Ge (L_4)	0.66 [13]	0.117/0.292[13, 135]	0.33 [13]	0.043 [13]	0.33 [119]	7.8×10^{10} [12]	6 [12]
Ge (Γ)	0.8 [13]	0.043 [13]			0.85 [119]	-	-
InAs	0.358 [136]	0.022[136]	0.41[136]	0.025[136]	3.6 [1]	-	-
In$_{0.53}$Ga$_{0.47}$As	0.751 [127]	0.042 [127]	0.457 [127]	0.052 [127]	1.5 [137]	-	-

valley have been considered, so both phonon-assisted and direct BTBT are included. For InAs and In$_{0.53}$Ga$_{0.47}$As, only the conduction band valley in Γ is included since the other valleys sit far apart energetically. The material parameters used in the simulations and calculations are listed in Table 4.1.

4.4.1 From 3D-3D to 2D-2D Face Tunneling

The left subplots in Fig. 4.28 depict the alignment voltage (defined as $V_{\mathrm{align}} = V_{\mathrm{n-gate}} - V_{\mathrm{p-gate}} = 2V_{\mathrm{n-gate}}$) for which the lowest electron and the highest hole subband get aligned, whereas the right plots report the I_{ON} (defined as I_{D} at $V_{\mathrm{n-gate}} = \frac{V_{\mathrm{align}}}{2} + V_{\mathrm{DS}} = -V_{\mathrm{p-gate}}$ with $V_{\mathrm{DS}} = 0.15\mathrm{V}$). Si, Si$_{0.5}Ge_{0.5}$ and Ge results are dominated by phonon-assisted BTBT. Although direct BTBT is included in Ge, it is found not to dominate the current in Ge EHBTFETs for the considered bias, contrarily to what happens in the bulk diodes of Fig. 3.19(Middle & Right) since the quantized energies of the Γ electrons are much higher than those of L valleys. For InAs and In$_{0.53}$Ga$_{0.47}$As EHBTFETs only direct BTBT takes place. Comparing the Ge, In$_{0.53}$Ga$_{0.47}$As and InAs results, it is seen that the overall behavior with respect to parameter changes is similar, except for the drastically higher I_{ON} of the InAs EHBTFET due to small direct bandgap.

About the T_{OX} independence of the ON current levels for Ge, InAs and In$_{0.53}$Ga$_{0.47}$As EHBTFETs, denoted by the almost vertical contour lines, a close inspection on the I-V characteristics reveals that it is caused by the fact that 2D-2D tunneling takes place in those cases which we will discuss in more detail later on. In this context 2D-2D tunneling means that only one electron-hole subband pair gets aligned to contribute to the BTBT current.

The behavior of I_{ON} with respect to the parameters drastically changes for Si EHBTFET (Fig. 4.28 top right) with thicker channel thicknesses ($T_{\mathrm{CH}} > 12.5\mathrm{nm}$). In those cases, the large effective masses of Si Δ_2 valleys and heavy holes allow for the formation of a quasi-3D continuum of states (i.e., many subband pairs get aligned, see Fig. 4.29). This is due to the fact that heavier masses reduce the energy spacing between quantized levels, resulting in (pseudo) 3D-3D tunneling. A similar comment can be made for Si$_{0.5}$Ge$_{0.5}$ EHBTFET as well, due to the similar effective mass values, even though the alignment voltage is relatively smaller due to

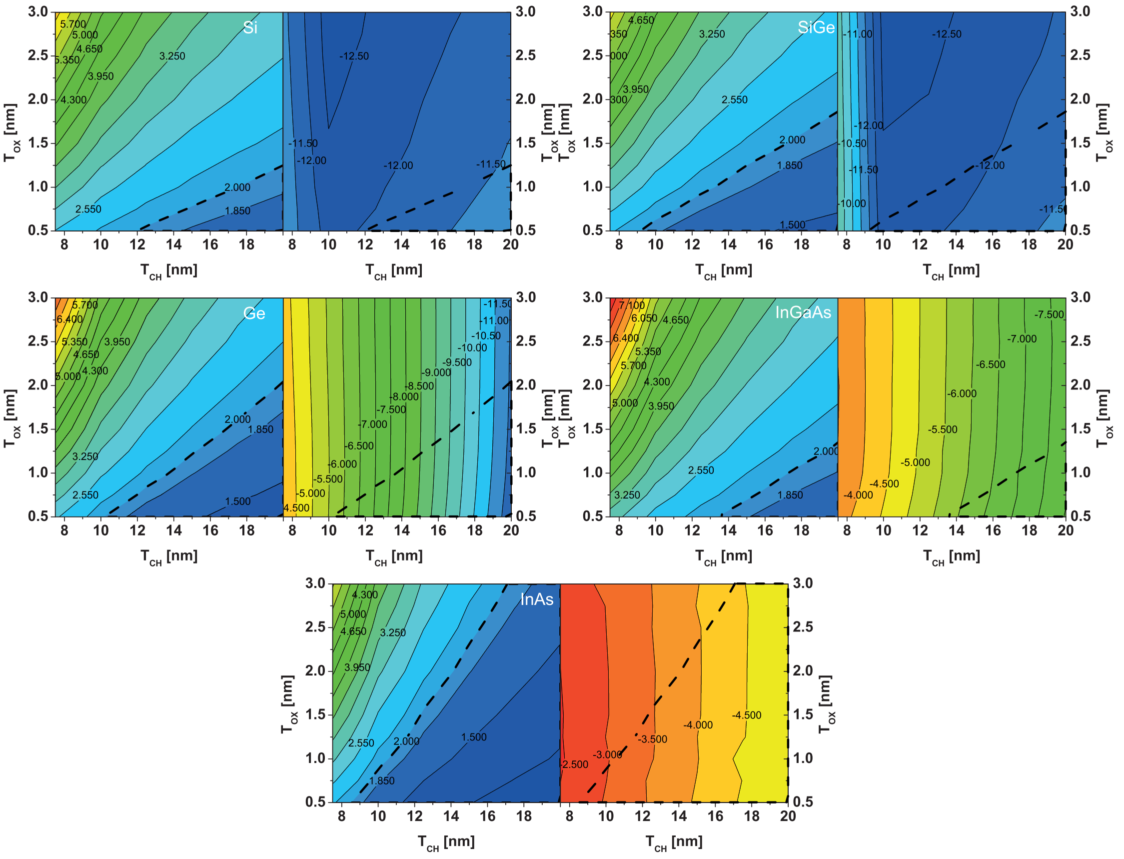


Figure 4.28 – Contour plots of (Left subplots) Alignment voltage V_{align}, (Right subplots) $\log_{10} I_{\text{ON}}\,[\text{A}/\mu\text{m}]$ for (from left top to bottom) Si, $\text{Si}_{0.5}\text{Ge}_{0.5}$, Ge, $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ and InAs EHBT-FETs. The dashed contour line corresponds to $V_{\text{align}} = 2\text{V}$. $V_{\text{DS}} = 0.15\text{V}$.

reduced bandgap. For both Si and SiGe, an interesting feature of the contour plots is that as the channel thickness is reduced from the relatively large values (20nm), a *decrease* in the ON current is seen. This is actually caused by the reduction of the DOS available for tunneling, which is stronger than the competing increase in the tunneling rate brought in by the thinner channel. Eventually, as the channel is further thinned down, the tunneling rate increase starts to counterbalance the effects related to the DOS reduction and the current is enhanced.

This drastic difference between 2D-2D tunneling and 3D-3D tunneling clearly demonstrates that the carrier gas dimensionality plays an important role in determining the influence of device parameters on the drive current. *A limit on the obtainable current is imposed for 2D-2D tunneling due to the reduction of the available states for tunneling, independently of the electrostatic control of the channel.* This is in stark contrast to 3D-3D tunneling, where a better electrostatical control would always yield higher current for the same bias.

This latter point is better visualized by the I-V characteristics given for a 2D-2D tunneling device (i.e., a Ge EHBTFET with $T_{\text{CH}} = 7.5\text{nm}$) versus a pseudo 3D-3D tunneling one (i.e., a Si

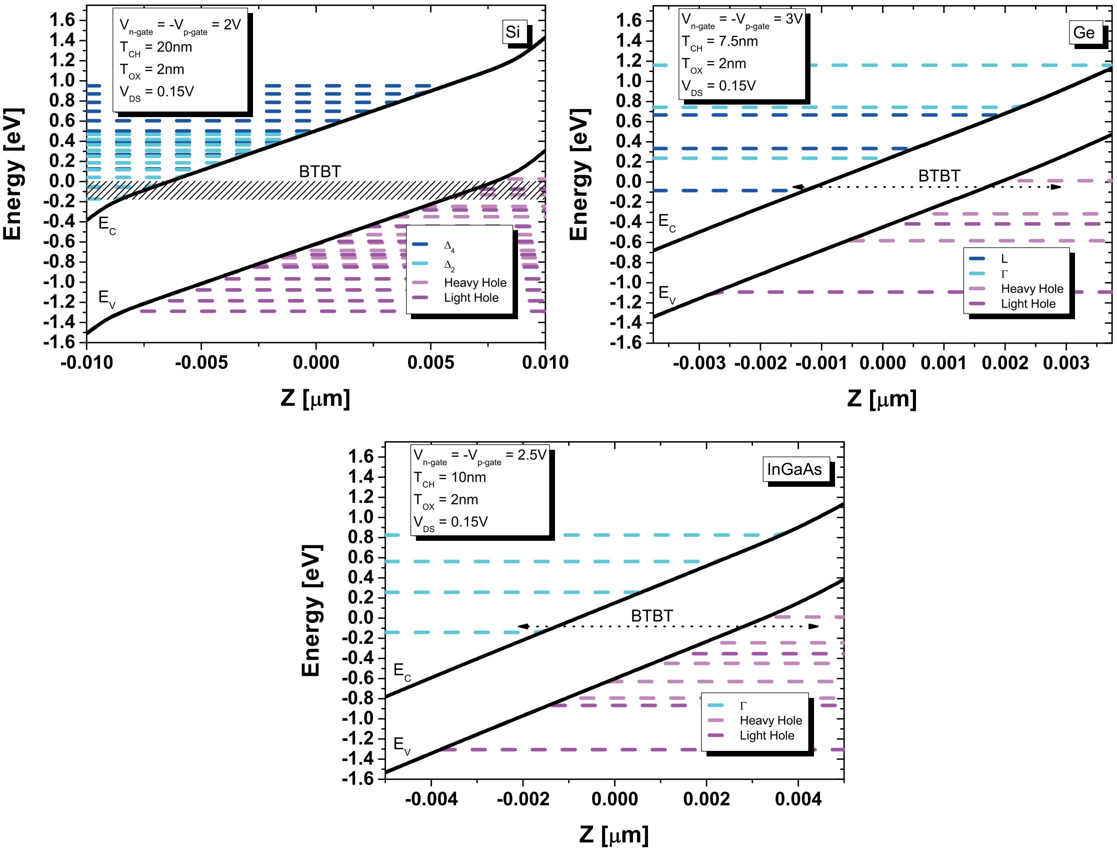


Figure 4.29 – Conduction and valence band profiles and quantized subband energies (dashed lines) for (Left) Si EHBTFET with T_{CH} = 20nm (Middle) Ge EHBTFET with T_{CH} = 7.5nm (bottom) InGaAs EHBTFET with T_{CH} = 10nm. Si EHBTFET exhibits 3D-3D tunneling behavior since multiple subbands get aligned, whereas for Ge and InGaAs EHBTFETs only one subband pair is aligned, hence resulting in 2D-2D tunneling behavior.

EHBTFET with T_{CH} = 15nm) in Fig. 4.30. It is seen that, in 3D-3D tunneling, a thinner oxide increases the ON current and also shifts the onset of tunneling to lower gate biases, whereas only a rigid shift of the onset of tunneling is seen for 2D-2D tunneling.

Another distinguishing feature between the 2D-2D and 3D-3D tunneling is that the I-V characteristics for the former generally exhibit a negative transconductance behavior [115], which stems from the wavefunction repulsion as the electric field increases and from the step-like 2D DOS which does not change with the gate bias. Wavefunction repulsion also occurs for 3D-3D tunneling, but the effect is masked by the addition of new states available for tunneling as the gate bias increases.

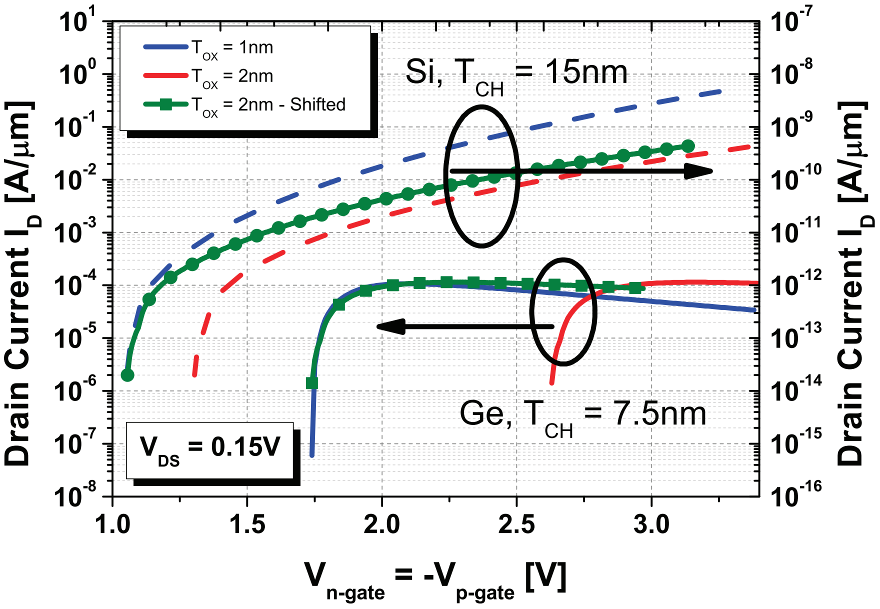


Figure 4.30 – I-V transfer characteristics for (Left y-axis) Ge EHBTFETs with $T_{\mathrm{CH}} = 7.5\mathrm{nm}$ and (Right y-axis) Si EHBTFETs with $T_{\mathrm{CH}} = 15\mathrm{nm}$. (blue) $T_{\mathrm{OX}} = 1\mathrm{nm}$ (red) $T_{\mathrm{OX}} = 2\mathrm{nm}$. $V_{\mathrm{DS}} = 0.15\mathrm{V}$. The green curves are the artificially shifted versions of the $T_{\mathrm{OX}} = 2\mathrm{nm}$ cases to align at the same onset voltage of the $T_{\mathrm{OX}} = 1\mathrm{nm}$ devices.

4.4.2 Analytical Model

To better explain the results of the numerical model and investigate the characteristic features of 2D-2D tunneling, an analytical model has been derived by solving the 1D vertical Poisson equation in the overlap region under the depletion approximation [140]. Neglecting the fixed dopant charges, since a very light doping $N_{\mathrm{A}} = 10^{15}\mathrm{cm}^{-3}$ is used in the simulations, the electrostatic potential inside the device is given by (see Appendix D for the derivation):

$$V(z) = -Fz + \frac{V_{\mathrm{p-gate}} + V_{\mathrm{n-gate}}(1 + \frac{C_{\mathrm{OX}}}{C_{\mathrm{CH}}})}{2 + \frac{C_{\mathrm{OX}}}{C_{\mathrm{CH}}}} \tag{4.1}$$

with:

$$F = \frac{C_{\mathrm{OX}}}{\epsilon_{\mathrm{CH}}} \left(\frac{V_{\mathrm{n-gate}} - V_{\mathrm{p-gate}}}{2 + \frac{C_{\mathrm{OX}}}{C_{\mathrm{CH}}}} \right) \tag{4.2}$$

The capacitances per unit area are defined as $C_{\mathrm{CH,(OX)}} = \frac{\epsilon_{\mathrm{CH,(OX)}}}{T_{\mathrm{CH,(OX)}}}$. Fig. 4.31(a) compares the analytical potential profiles against numerical results: they perfectly match when no inversion charge is induced and a good overall match is observed in general.

Since the potential is (to a very good approximation) linear, the quantized energies $E_{\mathrm{e(h)}}$ and wavefunctions $\psi_{\mathrm{e(h)}}$ are given by the solutions for the triangular potential well, for which

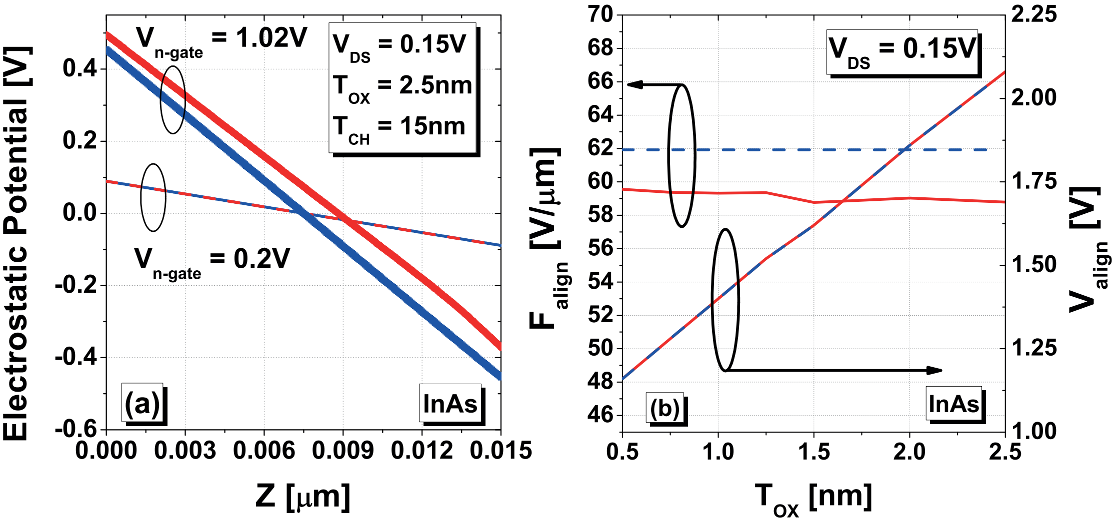


Figure 4.31 – (a) Potential profiles given by the numerical (blue) and analytical (red) models for high (symbols) and low (lines) applied gate voltages for the InAs EHBTFET. $z = 0$ and $z = T_{CH}$ corresponds to n-gate and p-gate oxide-semiconductor interface respectively. (b) Alignment field F_{align} and alignment voltages predicted by the numerical (solid lines) and analytical (dashed lines) models.

analytical expressions exist [135]. The alignment occurs when (considering direct tunneling) $E_e = E_h$ [123], i.e., when:

$$E_C(0) - a_1 \sqrt[3]{\frac{q^2 F_{align}^2 \hbar^2}{2m_e^*}} = E_V(T_{CH}) + a_1 \sqrt[3]{\frac{q^2 F_{align}^2 \hbar^2}{2m_h^*}} \tag{4.3}$$

which can be rewritten as:

$$E_G - a_1 \sqrt[3]{\frac{q^2 F_{align}^2 \hbar^2}{2m_e^*}} - a_1 \sqrt[3]{\frac{q^2 F_{align}^2 \hbar^2}{2m_h^*}} = qF_{align} T_{CH} \tag{4.4}$$

where $a_1 = -2.338$ is the first zero of the Airy function [141] given by the solution of the triangular potential well. Eq. 4.4 holds for direct BTBT process, whereas a similar expression (including the phonon energy) may be derived for phonon-assisted tunneling. Two key insights can be inferred from Eq. 4.4: i) the numerical results are confirmed about the dependences of the I_{ON} in 2D-2D tunneling since the equation has no explicit T_{OX} dependent terms, ii) for a given T_{CH}, the alignment always occurs at a certain magnitude of the electric field F_{align}, which is the solution of Eq. 4.4. The comparison between the numerically calculated and analytical F_{align} values (given in Fig. 4.31(b)), indicates good overall matching (maximum difference ~ 6% for the considered plot).

Now, we turn our attention to the implications of the fixed electric field at alignment on the wavefunctions. Within the triangular approximation, the analytical expressions for the electron and hole wavefunctions at subband alignment are given as [135]:

$$\psi_e(z) = D\mathrm{Ai}\left(\left(\frac{2m_e^*}{\hbar^2}(qF_{\mathrm{align}})\right)^{\frac{1}{3}}\left(z - z^*\right)\right) \tag{4.5}$$

$$\psi_h(z) = D\mathrm{Ai}\left(-\left(\frac{2m_h^*}{\hbar^2}(qF_{\mathrm{align}})\right)^{\frac{1}{3}}(z - z^*)\right) \tag{4.6}$$

where Ai is the Airy function, z^* is the classical turning point at which $E_{\mathrm{C(V)}} = E_{\mathrm{e(h)}}$ and D is a normalization constant. It is easily seen that the wavefunctions are fixed by F_{align}, since it is the only factor governed by the electrostatics. By inspecting Eq. 3.43, the coupling element M_{CV} is thus deduced to be fixed as well. The magnitude of the coupling element thus depends on the channel material properties and on the channel thickness only. The same holds for phonon-assisted BTBT since the wavefunction overlap remains fixed too. We note that the arguments presented above are valid when the triangular potential approximation holds, i.e., in the absence of inversion layers inside the device. In the presence of an inversion layer, detrimental effects such as pinning of the surface potential that decrease in the 'quantum efficiency' [123] and ultimately a decrease in the switching slope might occur.

Even though it does not play a significant role in determining I_{ON}, T_{OX} is actually quite crucial in determining V_{align}, which can also be calculated by the analytical model as:

$$V_{\mathrm{align}} = F_{\mathrm{align}}\left(T_{\mathrm{CH}} + 2\frac{\epsilon_{\mathrm{CH}}}{C_{\mathrm{OX}}}\right) \tag{4.7}$$

As seen in Fig. 4.31(b), Eq. 4.7 predicts V_{align} with excellent accuracy.

The findings above signal that the optimization strategy for the EHBTFET with certain performance constraints is to determine the channel material and T_{CH} first according to I_{ON} requirements, and then, using Eq. 4.7, estimate the required C_{OX} for the desired operating voltages. Moreover, since T_{OX} does not impact I_{ON}, the gate leakage can be mitigated (at the cost of a higher operating voltage or more aggressive gate workfunction engineering) with no penalty on I_{ON}.

4.4.3 Performance Considerations & Impact of Channel Material

In addition to the qualitative and theoretical assessments discussed above, certain inferences can be made about the performance limits (in terms of ON current) of the EHBTFETs with the different channel materials considered here. We report the alignment voltage values, minimum allowed channel thicknesses for a given alignment voltage and the ON current values in Table 4.2. Bearing in mind that EHBTFET is considered mainly for low power and low standby power applications [142, 143], we set an alignment voltage limit of 2V (see Fig. 4.28). This constraint can be achieved through gate workfunction engineering i.e., the 2V obtained

Table 4.2 – Alignment voltages V_{align}, minimum allowed channel thickness for the corresponding alignment voltage $T_{\text{CH,min}}$ and ON currents for the EHBTFETs with various channel materials. EOT = 0.5nm for all cases indicated below.

Channel Material	V_{align} [V]	Min. channel thickness $T_{\text{CH,min}}$ [nm]	ON current [A/μm]
Si	2	12	1×10^{-12}
$\text{Si}_{0.5}\text{Ge}_{0.5}$	2	9	3×10^{-12}
Ge	2	10	3.34×10^{-6}
$\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$	2	13	1×10^{-6}
InAs	2	8.5	1.75×10^{-3}
	1	19	2.6×10^{-5}

here with mid-gap metal work-functions can be achieved with lower applied bias if n- and p-type workfunctions are used for the n- and p-gates, respectively [144]. Within this constraint, InAs seems to offer the best $I_{\text{ON,max}}$ compared to Ge and $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$. Moreover, the InAs EHBTFET can also be operated at an even lower V_{align} of 1V, of course with lower $I_{\text{ON,max}}$, by using larger T_{CH}. Even though the bandgaps are comparable for Ge and $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$, the obtainable ON current is drastically higher for $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ since it is a direct gap material. On the other hand, it is not even possible to obtain 2D-2D tunneling for Si and $\text{Si}_{0.5}\text{Ge}_{0.5}$ within this constraint. It is worthwhile to note that, even if the bulk bandgap is smaller for $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$, the very small effective mass of the Γ electrons result in higher alignment voltages compared to $\text{Si}_{0.5}\text{Ge}_{0.5}$ for a given channel thickness due to stronger size-induced quantization effects.

An overall consistency with [39] is seen in the predicted alignment voltages (see Table 4.3), although our simulations indicate orders of magnitude lower conductance levels ($S_{\text{ON}} = \frac{\partial I_{\text{D}}}{\partial V_{\text{DS}}}$ at first subband alignment) than in [39] for Si and Ge. This is due to the use of a direct tunneling matrix element in Agarwal et al. that largely differs from the phonon assisted tunneling one [66] that we found to be dominant in Si and Ge devices. Even though direct BTBT is dominant in Ge for bulk 3D-3D tunneling (see Fig. 3.19(Right) in the previous chapter), quantization essentially suppresses the direct BTBT in the ON state of the Ge EHBTFET for all the cases considered in this study. Regarding InAs, instead, our results are only 5-10 times lower than the ones in [39] mainly due to the inclusion in our model of the effect of quantization on the matrix element (term $C_0(\theta)$ in Eq. 3.43)[105].

In this section, an extensive analysis for EHBTFETs using different channel thicknesses and materials as well as oxide thicknesses has been performed. Striking differences in geometrical parameter dependences of the I_{ON} is observed between 2D-2D tunneling (InAs, $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ and Ge) and 3D-3D tunneling (Si and $\text{Si}_{0.5}\text{Ge}_{0.5}$). For 2D-2D tunneling with a chosen channel thickness, the alignment always occurs at a certain F_{align}, fixing the overlap between the electron and hole wavefunctions. This exclusive feature of 2D-2D tunneling, along with the step-like characteristic of the 2D DOS, imposes an upper limit on the I_{ON} that can be obtained, contrary to 3D-3D tunneling. This limit depends on the channel thickness and on the material properties, whereas it is independent on electrostatic control of the channel. It is seen that

Table 4.3 – Comparison of calculated conductance values (S_{ON}) and alignment voltages (V_{align}). $T_{\mathrm{OX,EOT}} = 0.8$nm.

		Agarwal et al.		This work	
	T_{CH} [nm]	S_{ON} [S/μm]	V_{align} [V]	S_{ON} [S/μm]	V_{align} [V]
Si	4.3	1×10^{-3}	5.55	5.93×10^{-9}	5.29
	5.9	1×10^{-5}	4.19	3.33×10^{-11}	4.05
Ge	9.7	1×10^{-3}	2.68	1.47×10^{-6}	2.64
	14.9	1×10^{-5}	1.83	4.5×10^{-9}	1.95
InAs	14.6	1×10^{-3}	1.45	1.23×10^{-4}	1.33
	24.2	1×10^{-5}	0.93	2.45×10^{-6}	0.95

InAs is a prominent candidate for the realization of 2D-2D tunneling devices, since it has relatively low alignment voltages and very high ON current levels.

4.5 Fin EHBTFET

The fin EHBTFET is a modified version of the planar EHBTFET [5], where the overlap region [7] is now implemented as a fin-like structure (Fig. 4.32(a)). This structure was also simulated by Zhu et al. [145], however they used a purely semi-classical approach and predicted unrealistically optimistic performance. Here, we will show the necessity of the proper inclusion of the quantum effects as our simulations will show evidence of another non-ideality peculiar to the fin EHBTFET.

In the Fin EHBTFET, exactly in the same manner as in its planar counterpart, 2D electron and hole gases (2DEG and 2DHG, respectively) are induced at the either sides of the fin by applying a positive and negative bias to the n-gate and p-gate electrodes respectively. BTBT occurs in the entire overlap region when the subband energies of the electron and hole states inside the overlap region align. The major advantage of the fin implementation is the relative ease of fabrication compared to the planar version, although several challenges still remain such as effective separation of the n- and p-gates. In this work, we take the substrate region to be a thin semiconductor layer, but a bulk substrate may be possible too, provided substrate leakage effects are duly accounted for. Note that differently from the FinFET structure, the source and drain regions in fin-EHBTFET are along the cross section direction of the fin.

4.5.1 Geometry Effects on the Fin EHBTFET

In this work, we simulate an $In_{0.53}Ga_{0.47}As$ fin EHBTFET with a bandgap of 0.81eV [126] using our 2D parabolic band model (section 3.3). The effective masses for Γ-valley electrons, light holes and heavy holes are $0.042m_0$, $0.0503m_0$ and $0.465m_0$, respectively [127]. Unless otherwise stated, the fin thickness is $T_{\mathrm{CH}} = 10$nm, the substrate thickness $T_{\mathrm{SUB}} = 10$nm, the oxide thickness $T_{\mathrm{OX}} = 3$nm (HfO_2, $\epsilon_{\mathrm{r}} = 22$, corresponding to EOT=0.53nm), the fin height

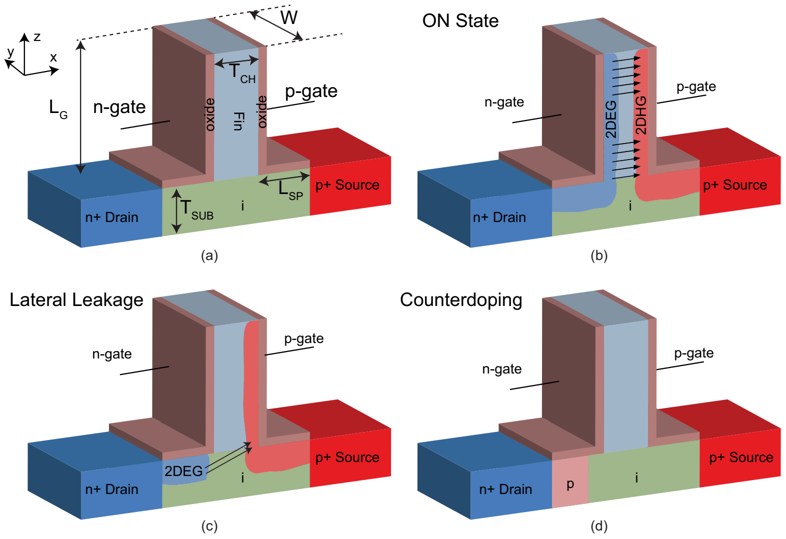


Figure 4.32 – (a) Fin EHBTFET device structure and regions. (b) ON state which occurs when the quantized energies for the 2D electron and hole gases (2DEG and 2DHG, respectively) align; the arrows indicate the BTBT direction. (c) Lateral leakage mechanism at the OFF state. (d) Counterdoping method to prevent lateral leakage.

$L_G = 50$nm and the spacer length $L_{SP} = 50$nm at both sides. Fin width W is taken to be 1μm. The source and drain regions are respectively p and n doped with a concentration of 1×10^{20}cm^{-3}. The drain voltage is $V_{DS} = 0.25$V. Unless otherwise stated, the p-gate is biased at -1V. We have assumed midgap workfunctions for both gates.

Inspecting the transfer characteristics plotted in Fig. 4.33 (Left, blue solid curve) and the BTBT generation rate profile in the OFF state (see Fig. 4.33 (Right)), a leakage component is observed (as sketched in Fig. 4.32(c)) for the fin EHBTFET, since the quantization is less severe in the substrate region and this allows the carriers to penetrate into the underlap region for lower voltages compared to the ones required to penetrate the overlap region, similar what happens in the planar case [7]. To address this issue, we utilize the counterdoping technique discussed in section 4.3.1 (see Fig. 4.32(d)) which delays the onset of the wavefunction penetration into the underlap region and hence recovers the steep slope behavior. It is seen that counterdoping the drain side with a p-type doping of $N_{UL} = 2.1 \times 10^{19}$cm^{-3} results in a drastic improvement of in the SS (red curve in Fig. 4.33 (Left)).

Corner Effect & Solutions to Suppress It

Comparing the ON current levels for the fin EHBTFET in Fig. 4.33 (Left, red curve) with those reported for the planar case with the same channel thickness reported in section 4.3.1 (green curve in Fig. 4.33) quickly reveals that the ON current is orders of magnitude lower for the fin EHBTFET. The reason for that is easily understood when we look at the valence and

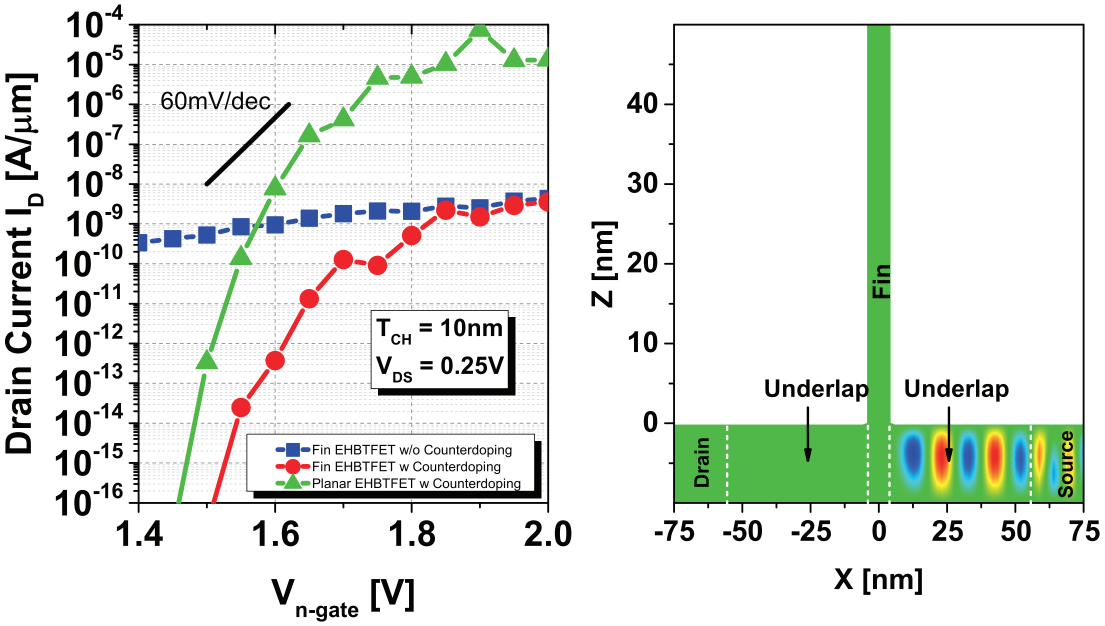


Figure 4.33 – (Left) Transfer characteristics for the (red) Fin EHBTFET without counterdoping, (blue) Fin EHBTFET with counterdoping $N_{\mathrm{UL}} = 2.1 \times 10^{19}\mathrm{cm}^{-3}$, (green) 10nm thick planar EHBTFET with counterdoping $N_{\mathrm{UL}} = 1.9 \times 10^{19}\mathrm{cm}^{-3}$ (see section 4.3.1) using 3nm HfO_2 as the gate stack. (Right) The HH wavefunction for the subband that contributes the most to tunneling at the OFF state (without counterdoping, $V_{\mathrm{n-gate}} = 1.25\mathrm{V}$). The p-gate bias is $-1\mathrm{V}$ for both plots.

conduction band profile cuts at 2Å below the oxide semiconductor interface which indicates sharp energy barriers with heights $\simeq 300-350\mathrm{meV}$ (Fig. 4.34, blue curves). At the corners of the fin at the substrate-fin interface, the *effective* oxide thickness is $\sqrt{2}T_{\mathrm{OX}}$ instead of T_{OX} (see Fig. 4.35(b)). This increase in the oxide thickness that controls the surface potential results in reduced electrostatic control that creates an energy barrier for both electrons and holes. This barrier cuts the connection of the states in the overlap regions from the carrier reservoirs (i.e., source and drain regions), hence resulting in a drastic reduction of the current. Specifically, the presence of the potential barrier effectively results in stationary quasi-bound states that reside in the fin, with little coupling to the source and drain regions [124]. However, scattering to the quasi-bound states from the incoming waves is unlikely, since the phonon energies are essentially insufficient to provide energy for overcoming the barrier. As stated in the preceding chapter, in our simulations, we neglect scattering and we remove such quasi-bound states by filtering out the states with low coupling to the source/drain electrodes as explained at the end of section 3.3.

One solution to alleviate the corner effects could be to smoothen the corners of the fin such that the maximum oxide thickness is always T_{OX} (see Fig. 4.35(b, left)). In this way the loss of electrostatic control could be recovered. However, the conduction band cut along the

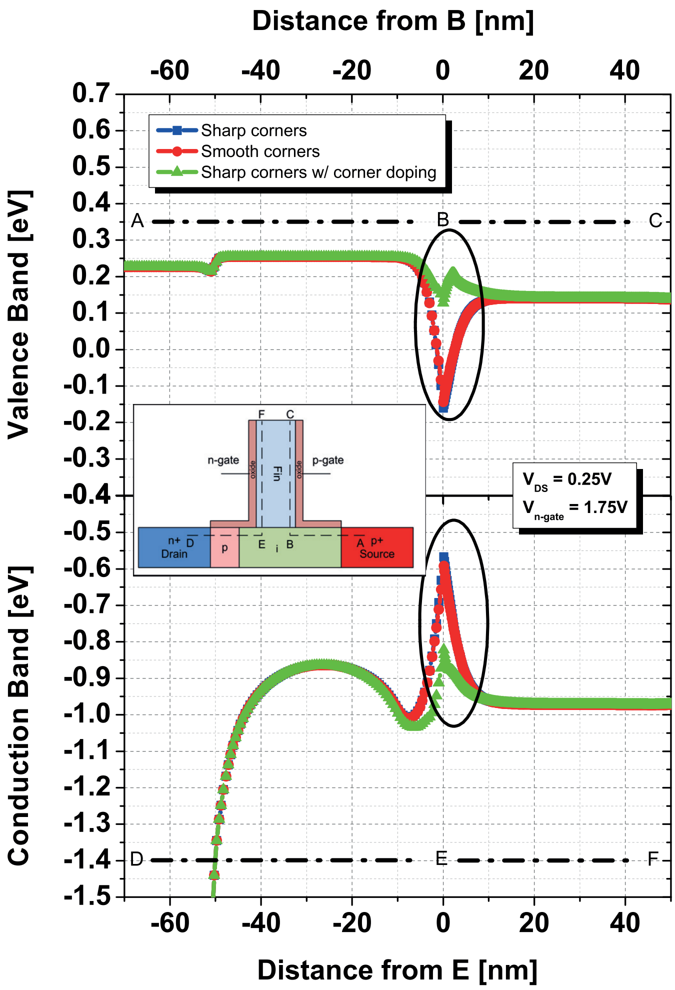


Figure 4.34 – The 1D cut of the valence and conduction bands 2Å below the semiconductor-oxide interface at either sides of the InGaAs fin-EHBTFET with (blue) Sharp corners (red) Smooth corners (green) corner doping ($N_C = 1 \times 10^{20}\text{cm}^{-3}$). Note that red and blue curves are essentially overlapping except very small difference around the corner region.

oxide-semiconductor interface and the transfer characteristics given in Fig. 4.36 indicate that the approach is essentially ineffective, since the change in energy barrier height $\Delta E \simeq 20\text{meV}$ on the surface energy profile is insufficient. This indicates that the main phenomena at the origin of the corner effect is not the increase in the effective oxide thickness, but rather the geometrical shape of the structure. One can illustrate this point by calculating (see Fig. 4.35(b)) the distance between a point K on the surface located exactly on the corner and a point L at the oxide/gate-metal interface that is 1nm apart from a third point M on the oxide/gate-metal interface that is the closest point to K on the oxide/gate-metal interface. Bearing in mind that we utilize a gate oxide of 3nm, the distance between the points K and L are: 5nm (fin with sharp corners) and 4.32nm (fin with corner smoothing). It is evident that corner smoothing only solves the problem for the region closer to the corner but quickly loses its effectiveness as we move away from the corners, as evidenced by the transfer characteristic curve given in Fig. 4.36.

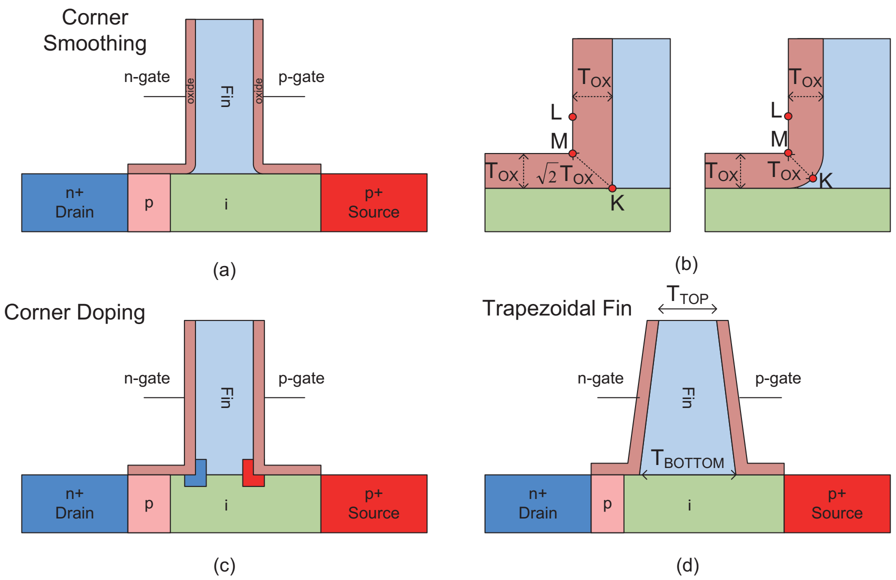


Figure 4.35 – Different methods to alleviate corner effects (a) corner smoothening. (b) zoomed comparison of the oxide configuration around the corners for (Left) no smoothing (Right) smoothing, (c) corner doping, (d) trapezoidal fin structure.

A second solution is to introduce shallow doped regions around the fin corners at the substrate interface (see Fig. 4.35(c)), similar to the technique proposed for the suppression of the corner effects causing earlier onset of conduction in FinFETs [146]. Such a doping profile can be obtained using low-energy tilted implants. Similar to the counterdoped regions, the corner doping modifies the band profile around the corners to attenuate the energy barrier. On Fig. 4.34, it is seen that corner doping is extremely effective; it completely removes the energy barrier. The transfer characteristics for fin EHBTFETs utilizing corner doping are given in Fig. 4.36 where we utilize rectangular doped corner regions of about 6nm × 2nm and 4nm × 3.5nm for n-gate side and p-gate side, respectively. It is seen that the case with $N_C = 1 \times 10^{20}\text{cm}^{-3}$ is able to provide ON current levels similar to the planar version, although the switching slope is not as good at lower current levels. As a quantitative indication of the static device performance, the average slope of the corner doped Fin EHBTFET measures 43mV/dec for over six decades (from 1pA/μm level to 1μA/μm), whereas for the planar case, it is 36mV/dec. An analysis of the OFF states for the cases with and without corner doping indicates that the OFF current is slightly amplified due to corner doping (see e.g., the region where $V_{n-\text{gate}} \simeq 1.5$V in Fig. 4.36). Heavy doping essentially screens the gate control around the corner, which then degrades the OFF state performance. However, we observe this effect only at very low current levels where other sources of leakage, such as SRH or gate leakage, would most likely dominate the current.

As seen in Fig. 4.36(green curve), utilizing a lower doping of $5 \times 10^{19}\mathrm{cm}^{-3}$ delays the onset of line tunneling (i.e., the ON state) and also decreases the subthreshold slope, even though the ON current level reaches the same level as in the $1 \times 10^{20}\mathrm{cm}^{-3}$ case. The reason for such a behavior is again the insufficient suppression of the energy barrier in the underlap region which implies the need of using a higher gate voltage to overcome the barrier.

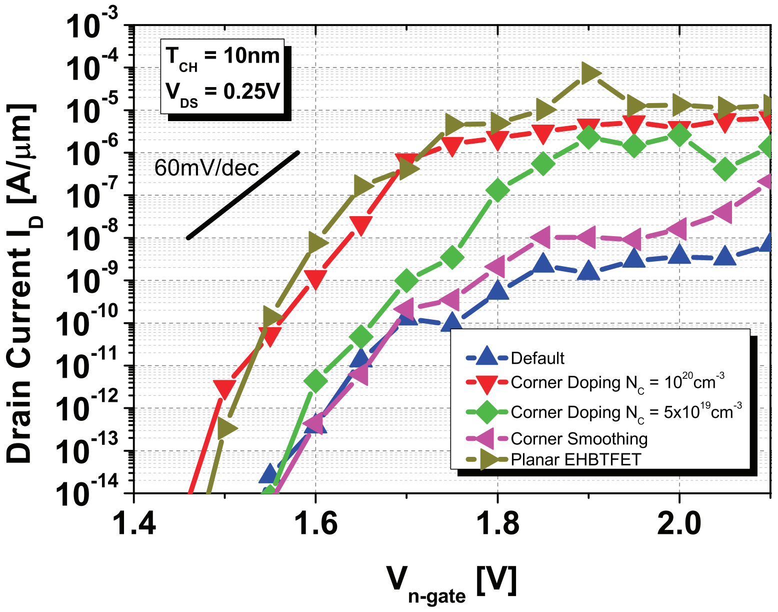


Figure 4.36 – Transfer characteristics for the fin EHBTFET with (blue) no optimization, (red) corner doping $N_C = 1 \times 10^{20}\mathrm{cm}^{-3}$, (green) $N_C = 5 \times 10^{19}\mathrm{cm}^{-3}$, (magenta) corner smoothing, (dark yellow) planar EHBTFET with counterdoping using 3nm HfO_2 as the gate stack.

Another possible solution that could alleviate the corner effect is employing a trapezoidal fin structure (see Fig. 4.35(d)). This method aims at reducing the strength of field-induced quantization around the corners by increasing the fin thickness in those areas. Increasing the fin thickness moves away the n- and p-gates, thereby reducing the electric field. Fig. 4.37 indicates that fins with different trapezoid configurations result in similar characteristics. For all the trapezoidal cases, we see that the corner effect is overcome. However, increased bottom fin thickness results in reduced coupling between electron and hole states in the fin which contributes to the ON current. On the other hand, it also lowers the strength of quantization so that the energy spacing between subbands is reduced. At high $V_{n-\mathrm{gate}}$ (~ 2V), we observe similar ON currents for the trapezoidal cases due to the fact that reduced coupling between the states is compensated by the additional aligned subbands in the fin region with lower current contribution.

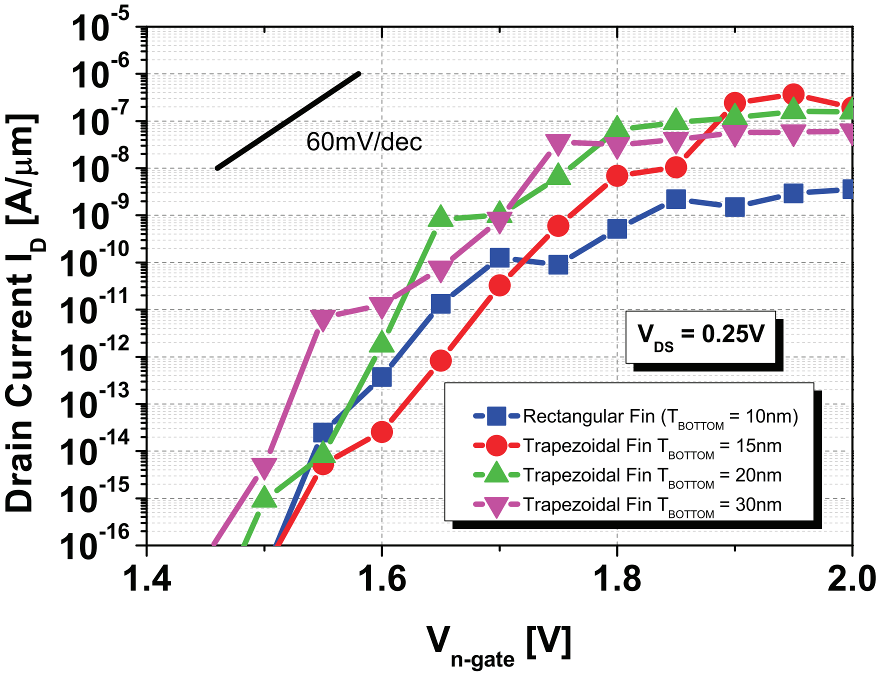


Figure 4.37 – Transfer characteristics for the fin EHBTFET with trapezoidal fin structures (Fig. 4.35(d)). The top thickness of the fins T_{TOP} is kept at 10nm.

4.6 EHBTFET Logic

In this section, we move up from the device-level to the circuit-level and propose a novel logic scheme. By utilizing the independent gate biasing feature of the EHBTFET, one can achieve significant simplifications in circuit design, and could potentially translate into faster operation. We will make use of the fact that a positive (negative) bias needs to be applied to n-gate (p-gate) to induce the electron (hole) subbands in the overlap region. This 'XOR-like' conduction property will be used to reduce the number of transistors required to implement a logic function. We also note here that different configurations are certainly possible with appropriate gate metal workfunction choices. For example, the EHBTFET could be designed to conduct when both gates are logic-0 and the device could be turned off by applying logic-1 to p-gate. This is similar to the proposal in [147] that however assumes TFET architectures different than the EHBTFET.

Another property of the EHBTFET, that is shared by many other TFET structures, is the ambipolarity i.e. the same device can be used as a pull-up (p-type) or a pull-down (n-type) device. However, one should note that the ambipolarity of the EHBTFET is slightly different from the one of conventional TFETs such is often deemed problematic. In the EHBTFET, the source and drain regions do not actively take part in the BTBT as in the conventional TFETS, but only supply the carriers to the tunneling region. Therefore, it could be possible to achieve a complementary functionality by using a single device. Moreover, since the EHBTFET is a symmetric device (i.e. both electrons and holes contribute to conduction mechanism), n-type and p-type static behavior is almost symmetric as well, provided the gate workfunctions are chosen appropriately.

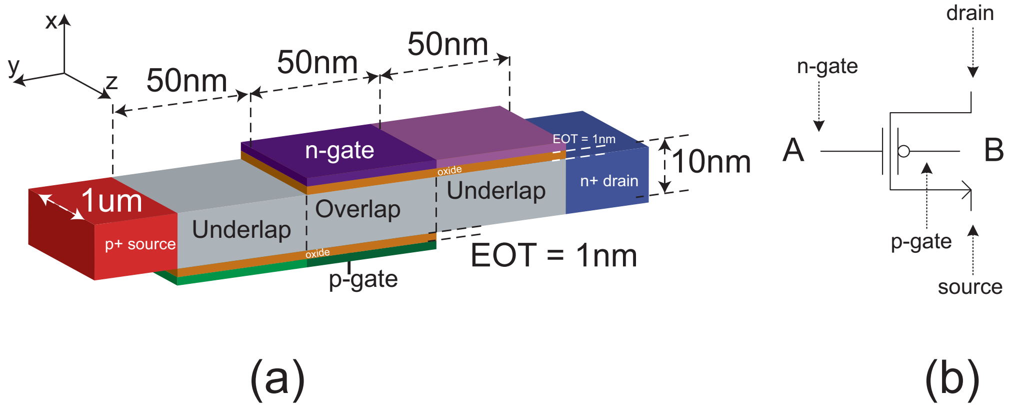


Figure 4.38 – (a) The sketch of the Hetero-Gate (HG)-EHBTFET device (b) The EHBTFET circuit symbol for the EHBTFET denoting the corresponding electrodes.

Different than the previous parts of the thesis, in this section, we demonstrate the operation of various logic circuits using quantum-corrected mixed-mode TCAD simulations. We will use Sentaurus Device to perform the static and transient simulations. We utilize the semi-classical tool due to the fact that it can do transient circuit simulations. However, as shown earlier, a proper treatment of quantum mechanical effect is still a prerequisite. To this end, we make use of a recently proposed method [148] to mimic subband quantization in the semiconductor by using the Physical Model Interface (PMI) of Sentaurus Device [57]. The model modifies the conduction and valence band edges near the semiconductor/oxide interface thus shifting the onset of tunneling in a consistent way with the predictions of the quantum mechanical (QM) simulations [115]. Specifically, the quantum-corrected conduction band edge $\left(E_{\mathrm{C}}^{\mathrm{QM}}\right)$ is given as:

$$E_{\mathrm{C}}^{\mathrm{QM}} = \begin{cases} E_{\mathrm{C}} + a_1 \sqrt[3]{\dfrac{q^2F^2\hbar^2}{2m_e}} - d_{\mathrm{OX}}F & a_1 \sqrt[3]{\dfrac{q^2F^2\hbar^2}{2m_e}} \geq d_{\mathrm{OX}}F, \\ E_{\mathrm{C}} & \text{otherwise} \end{cases}$$

where $a_1 = -2.338$ is the first zero of the Airy function [141], F is the electric field normal to the semiconductor/oxide interface and d_{OX} is the distance from the closest semiconductor/oxide interface. This correction enables the TCAD tool to predict both the vertical (wanted) and the lateral (leakage, unwanted) BTBT current which stems from the fact that quantization is less severe in the underlap region as shown in section 4.2. Using such an approach allows us to include quantum effects properly in the mixed-mode circuit simulation environment. However, it should be noted that the charge densities are still calculated semiclassically. Therefore care must be taken while choosing gate workfunctions to ensure that carrier densities do not reach large values where they will start affecting the curvature of the potential profile, which may possibly cause divergence from the QM simulations.

The device structure, the main dimensions and circuit schematic symbol of the $In_{0.53}Ga_{0.47}As$ EHBTFET is given in Fig. 4.38. The material properties are the ones listed in Table 4.1 for InGaAs. We utilize the heavy hole mass for the PMI quantization model to determine the subband energy. On the other hand, the light hole mass ($m^*_{\mathrm{lh}} = 0.052\mathrm{m}_0$) is used for the dynamic non-local BTBT model [57] since the conduction and heavy hole subbands are connected in the imaginary dispersion with a hole effective mass much closer to the light hole one (see the section 3.7 about anti-crossing). To prevent the lateral leakage that deteriorates the SS, we utilize the hetero-gate configuration [125] with $\phi_{\mathrm{n-gate,OL}} = 3.225\mathrm{eV}$ and $\phi_{\mathrm{n-gate,UL}} = 3.9\mathrm{eV}$ whereas we use $\phi_{\mathrm{p-gate,OL}} = 6.225\mathrm{eV}$ and $\phi_{\mathrm{p-gate,UL}} = 5.75\mathrm{eV}$. Unless otherwise stated, we set $V_{\mathrm{DD}} = 0.25\mathrm{V}$.

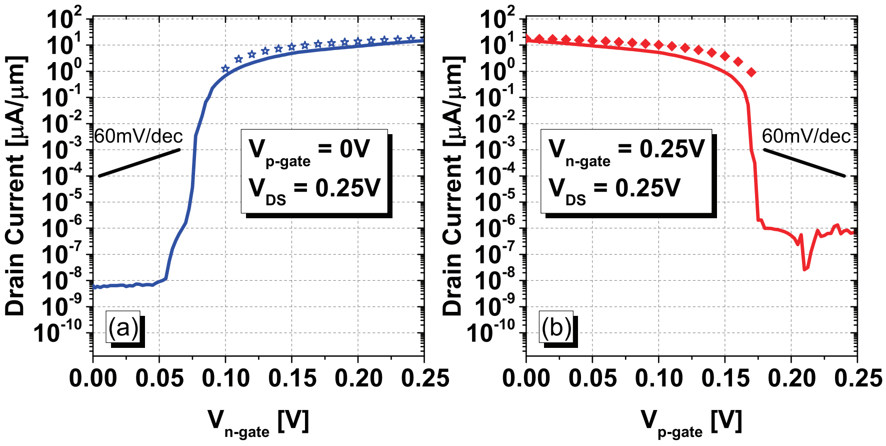


Figure 4.39 – The transfer characteristics of the InGaAs EHBTFET for (a) n-gate sweep with $V_{\mathrm{p-gate}} = 0\mathrm{V}$ and (b) p-gate sweep with $V_{\mathrm{n-gate}} = 0.25\mathrm{V}$. Solid lines indicate the results obtained by the quantum-corrected TCAD tool, whereas the symbols indicate the results obtained by the quantum mechanical model described in section 3.2.

The transfer characteristics from TCAD for the n- and p-gates are given in Fig. 4.39 (a& b). Comparison against the results obtained by the 1D quantum-mechanical model developed in this thesis using the same parameters indicates an excellent match in the current levels and good match in the onset voltages for tunneling (max. difference ~ 20mV). For the n-gate sweep, p-gate is kept at 0V; similarly for the p-gate sweep, n-gate is kept at V_{DD}. The device exhibits few mV/dec switching slope over eight decades of current for both n- and p-gate sweeps. Moreover, ON current, switching slope and the tunneling onset voltages are almost symmetric, signaling that n-type (pull-down) and p-type (pull-up) performances are comparable and the same device can be used for both purposes. The output characteristics for the pull-up (source node charges from 0 to 0.25V) and pull-down (drain node discharges from 0.25V to 0) cases given in Fig. 4.40 (a & b) indicate a good overall match with the QM simulations as well as negligible superlinearity [16]. The pull-up exhibits slightly lower output resistance since the valence band profile in the region controlled by the p-gate is not fully

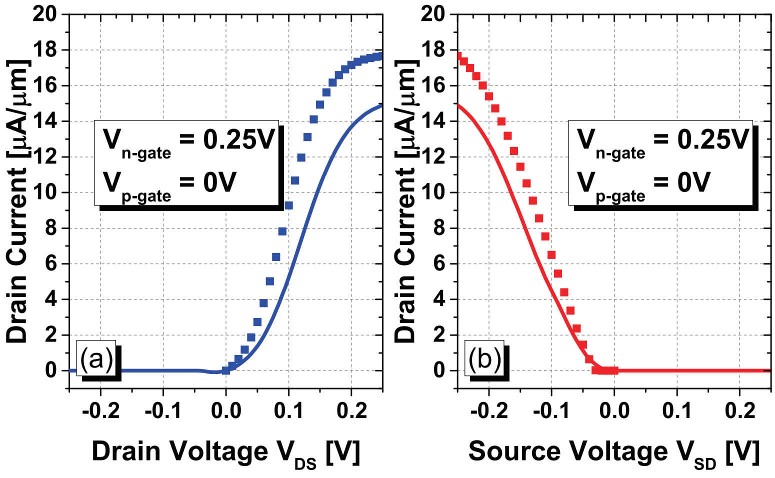


Figure 4.40 – Pull-up (a) and pull-down (b) output characteristics for the EHBTFET. Solid lines indicate the results obtained by the quantum-corrected TCAD tool, whereas the symbols indicate the results obtained by the quantum mechanical model described in section 3.2.

pinned by the inversion charge. These said, it should be noted that the EHBTFET suffers from the partially unidirectional conduction that is shared by all TFETs that, in certain cases, result in excess charging of the circuit nodes [149].

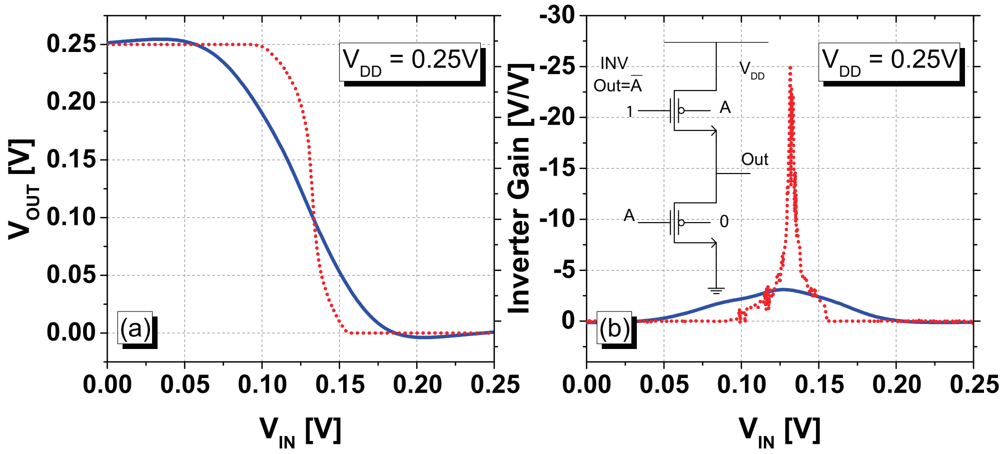


Figure 4.41 – (a) Voltage transfer characteristic (VTC) (b) Inverter gains of the EHBTFET inverter for a supply voltage $V_{\mathrm{DD}} = 0.25\mathrm{V}$ using the workfunctions (WF) given in the manuscript (blue solid line) or using optimized WFs for pull-up and pull-down transistors (red dotted line). Circuit schematic for the inverter using the EHBTFET is given as an inset of (b). The WF values for this latter case are $\Phi_{\mathrm{n-gate,OL}} = 3.3\mathrm{eV}$, $\Phi_{\mathrm{p-gate,OL}} = 6.3\mathrm{eV}$ (pull down) and $\Phi_{\mathrm{n-gate,OL}} = 3.15\mathrm{eV}$, $\Phi_{\mathrm{p-gate,OL}} = 6.15\mathrm{eV}$ (pull up).

Fig. 4.41 (Left) depicts the voltage transfer characteristic (VTC) of an inverter made up of the EHBTFET (Fig. 4.41 (Right inset)) described above used as both pull-up and pull-down, showing sufficient inverter gain and noise margins. We have verified in Fig. 4.41 (Right) that the inverter gain and the noise margins can be vastly improved by utilizing two different EHBTFET devices optimized for pull-up and pull-down behavior respectively. For the remainder of the section, we will present results for the case with same device for pull-up and pull-down.

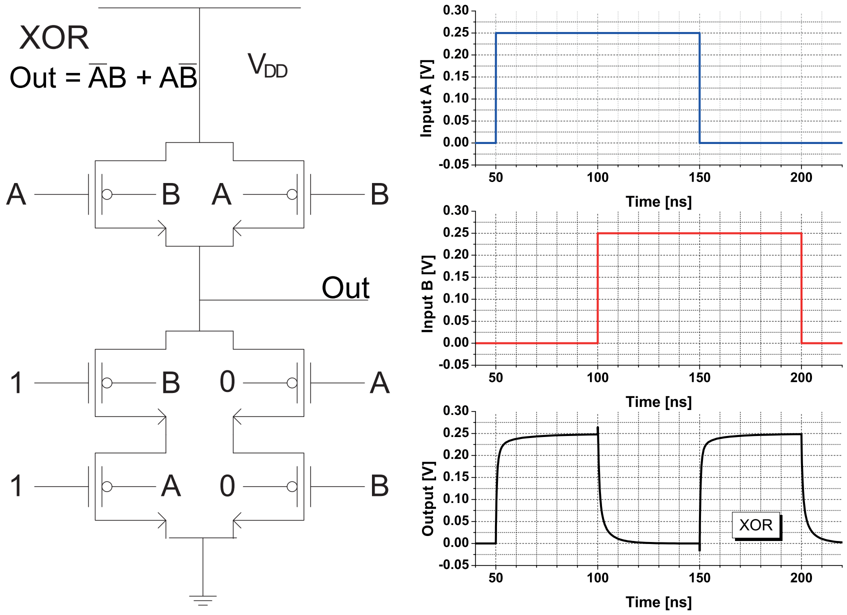


Figure 4.42 – The circuit schematic for the XOR2 gate using EHBTFET logic. Transient simulation input and output waveforms of the XOR2 using $V_{DD} = 0.25\text{V}$. The circuit has 30fF of capacitive loads at their output nodes. The results indicate extremely fast switching (~ 10ps) even at low supply voltage.

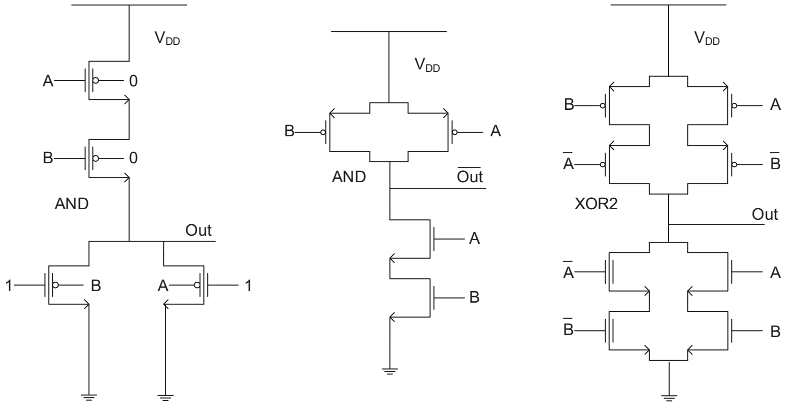
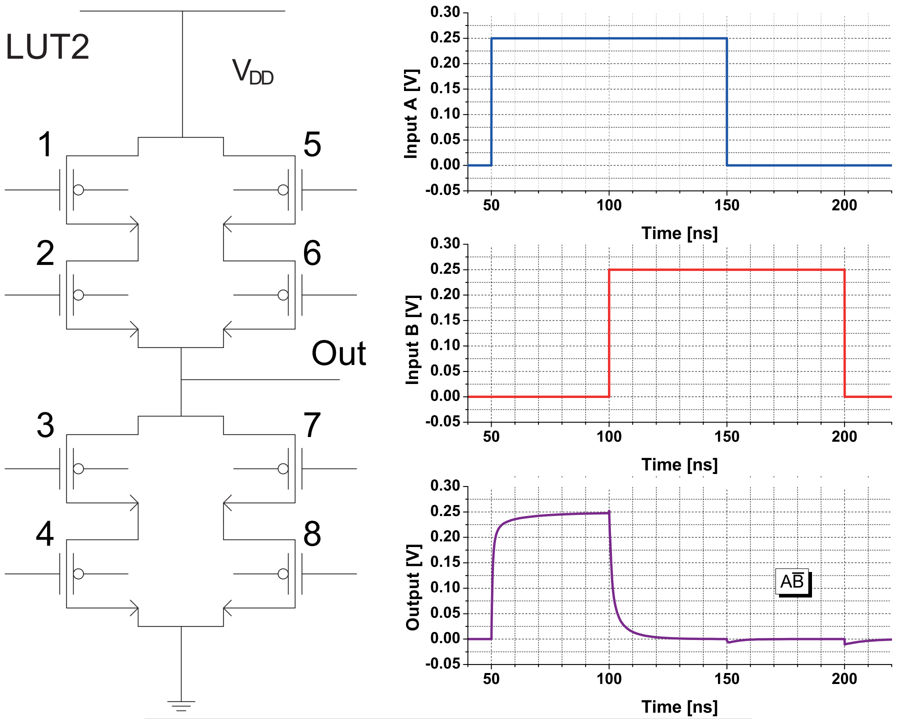


Figure 4.43 – Circuit schematic for (Left) AND2 gate using EHBTFET logic (Middle) AND2 gate using static CMOS logic. Note the circuit schematic drawn here is actually for the NAND2 gate and therefore the output of this stage is labeled as $\overline{\text{Out}}$. This stage needs to be followed by an inverter (i.e., incrementing the transistor count by 2) to obtain AND2 functionality. (Right) XOR2 using standard CMOS logic.



Transistor #	n-gate/p-gate	Transistor #	n-gate/p-gate
1	A/B	5	A/B
2	1/0	6	1/0
3	1/A	7	B/0
4	1/0	8	1/0

Figure 4.44 – The circuit schematic for LUT2 gate implemented with EHBTFET logic. Transient simulation input and output waveforms of LUT2 programmed as $A\overline{B}$ using V_{DD} = 0.25V. The corresponding input signals for LUT2 are indicated in the table above. The circuit has 30fF of capacitive loads at their output nodes.

The main idea behind the EHBTFET logic stems from the observation that the EHBTFET is conducting only when $V_{n-gate} = V_{DD}$ and $V_{p-gate} = 0$. If we denote the input signal at n-gate and p-gate as A and B respectively, the condition for conduction is fulfilled only when $F = A\overline{B}$. This flexibility allows us to implement $A\overline{B}$ terms using one transistor showing n-type (conducts when the n-gate input is 1) and one showing p-type (conducts when the p-gate input is 0) behavior. Fig. 4.42 & Fig. 4.43(b) presents two-input XOR and AND gates implemented with the EHBTFET logic. This flexibility allows for a sizable reduction in the number of transistors for logic gates. For instance, the XOR gate requires 6 transistors in EHBTFET logic compared to 12 transistors in conventional static CMOS logic (including the inverters required to obtain the inverted input signals, see Figs. 4.42 & 4.43(Right)) and AND gate 4 transistors vs 6 transistors (see Figs 4.43(Left) & (Middle)). The EHBTFET logic XOR circuit operation is verified through transient simulations whose waveforms are given in Fig. 4.42.

Fig. 4.44 shows the EHBTFET implementation of the LUT2, a building block for most FPGAs. It allows for implementing any 2-input logic function by multiplexing the inputs A and B as well as ’1’ and ’0’ bits to the gates of the 8 transistors. The same number of transistors is needed also for conventional CMOS, but inverters are needed if $\overline{A}$ and $\overline{B}$ are required. In the EHBTFET case, instead, we can exploit the n- and p-gates to eliminate the inverters. As an example, we show the implementation of the function $A\overline{B}$ using LUT2 circuit with correponding input signals (see table in Fig. 4.44) and resulting transient simulations (Fig. 4.44).

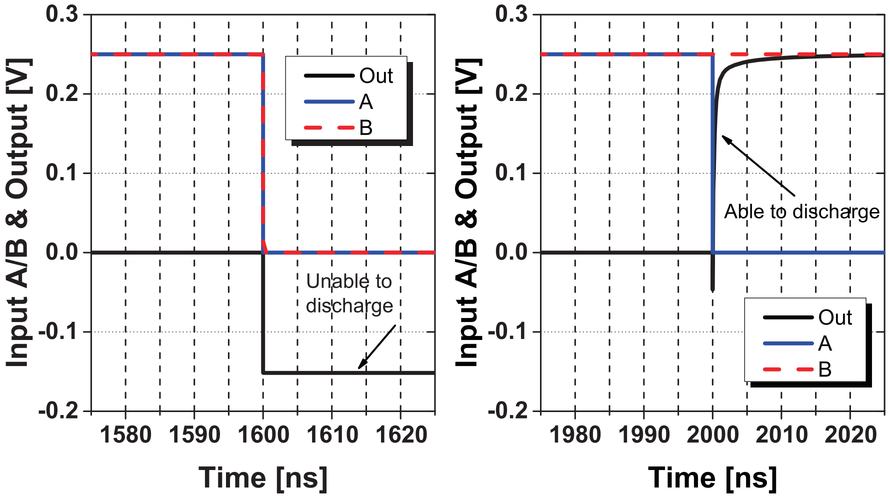


Figure 4.45 – Transient simulations of the XOR2 with a low capacitive load of 300aF. (Left) The overshoot is too much to recover from, resulting in the output node trapped at a negative voltage.(Right) The transistor drive current still remains strong enough to discharge the node.

The aforementioned problem of unipolar conduction presents a possible problem during the switching operation. In the case when the drain somehow gets negatively biased (e.g., voltage undershoots during switching), the p-i-n junction gets positively biased and the current drops significantly due to the fact that the Fermi supply window (i.e., the $f_c - f_v$ term in the BTBT current expression of Esaki [150] or e.g., Eq. 3.33) no longer coincides with the tunneling window. This corresponds to $V_{DS} < 0$ and $V_{SD} > 0$ regions in Fig. 4.40 for pull-up and pull-down devices respectively. This is the familiar negitive differential resistance (NDR) effect seen in Esaki diodes; however in the EHBTFET case, this may cause that the negatively biased drain node cannot be charged back if the junction bias goes beyond the NDR region, where the current only consists of reverse bias current which is typically very low since the built in potential barrier of the junction is not yet overcome. We note here that observation of such behavior depend on the overall capacitive loads seen at each node where the amplitude of the overshoot is inversely proportional to the overall load. Considering the approximate capacitive loads of the next-stage gates which are sizable as well as taking into account the ever-increasing interconnect capacitances [151], we doubt that such issues will be present for properly designed gates. In fact, we were only able to reproduce this behavior by using a very low load capacitance (300aF), in Fig.4.45 and when both inputs A & B are changing. One can also think of a circuit level design guideline to make sure such node voltages do not alter the overall circuit performance and functioning.

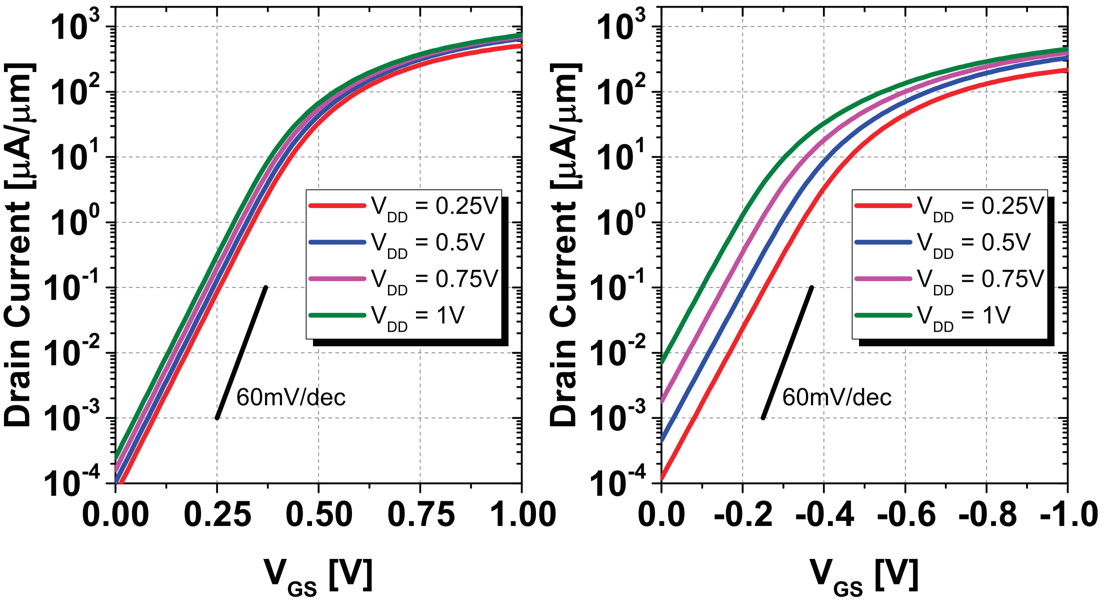


Figure 4.46 – The transfer characteristics for (Left) n-type and (Right) p-type FD-SOI CMOS devices. For both devices, the width is 1μm.

We now look into a quantitative comparison that will allow us to compare the EHBTFET against already existing low-power CMOS technology. In order to do this, we make use of the 28nm fully depleted SOI (FD-SOI) of ST Microelectronics [152]. Fig. 4.46 depicts the transfer characteristics for the n-type and p-type devices. The devices exhibit an average SS of about 80mV/dec. Comparing the transfer characteristics of the EHBTFET (Fig. 4.39) and CMOS (Fig. 4.46) for $V_{DD} = 0.25$V, we see that the EHBTFET is able to deliver dramatically higher currents at lower gate voltages.

Fig. 4.47 compares the worst case delays for varying supply voltages for the EHBTFET logic and 28nm FD-SOI CMOS. We note that we did not utilize the standard cell libraries, but rather designed standard CMOS implementations of the inverter and XOR gates. For fair comparison, we fixed the width of each CMOS transistor at 1μm and varied the gate length for adjusting the driving strength of the transistors for the optimum and balanced rise and fall time behavior. We compare the worst case timings for fan-out-4 [153] inverter delays in Fig. 4.47(Left). As commonly done [143, 154], we simulate an inverter chain where the driving strength quadruples between each stage (see the inset in 4.47(Left)) and we report the worst case delay of the middle stage. It is seen that for both circuits, the EHBTFET outperforms the CMOS implementation for low supply voltages ($V_{DD} < 0.5$V), where it is able to deliver higher drive current than the CMOS counterpart. As the supply voltage is increased, CMOS performance improves dramatically thanks to the increasing ON current of the CMOS transistors.

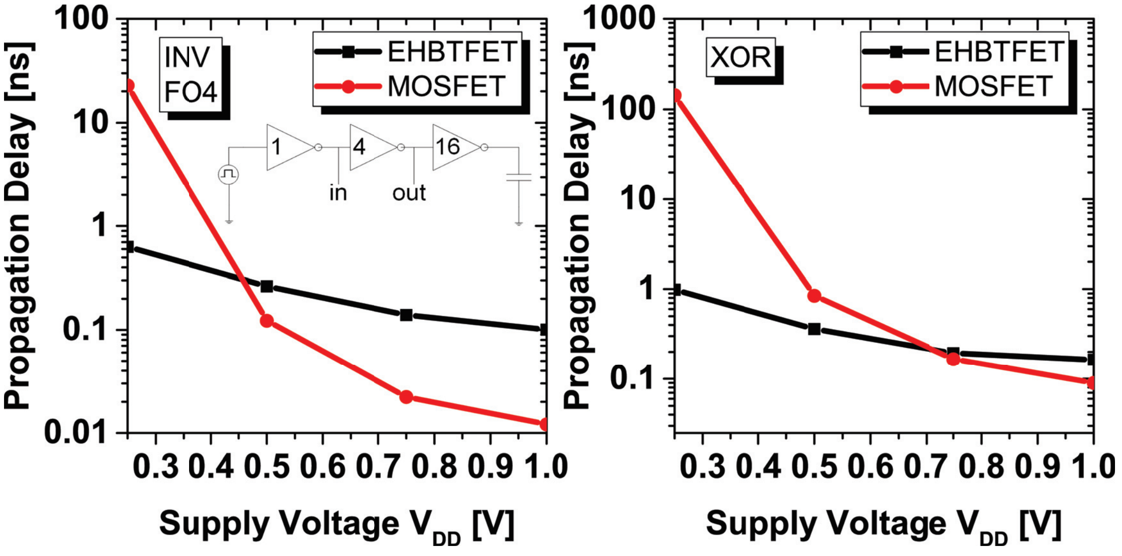


Figure 4.47 – The worst case delays as a function of the supply voltage V_{DD} for (Left) fan-out-4 Inverter (Right) XOR circuits. The black (red) lines indicate EHBTFET logic (CMOS, 28nm FD-SOI from ST Microelectronics) implementations. Both circuits have 30fF of load capacitance.

4.7 Impact of Dimensionality on Tunneling Devices

In the section 4.4.1, we had briefly discussed the impact of carrier dimensionality on the EHBTFET considering the 2D-2D and 3D-3D cases. In this section, we investigate a few other dimensionality cases besides the 2D-2D face tunneling [15] and try to assess the impact of dimensionality effects on the device characteristics.

4.7.1 EHBTFET vs Ultra Thin Body TFET: 2D-2D Face vs 2D-2D Edge Tunneling

The first situation we consider is 2D-2D tunneling. We first aim to compare 2D-2D face and 2D-2D edge tunneling (Fig. 4.1). The 2D EMA simulator developed in section 3.3 allows us to study this aspect considering a planar InGaAs EHBTFET (2D-2D face tunneling) and an InGaAs Ultra Thin Body TFET (UTBTFET, 2D-2D edge tunneling). For a fair comparison, we use the same channel and oxide thicknesses for both devices. We utilize the counterdoping technique explained in section 4.3.1 to suppress the lateral leakage in the EHBTFET.

In the Fig. 4.48(Left), we provide the transfer characteristics for the EHBTFET and UTBTFET using the same EOT = 1nm and the same channel thickness and channel material. We note that we do a double gate sweep (i.e., both gates for the UTBTFET are swept simultaneously), whereas we keep the p-gate gate fixed for the EHBTFET and only sweep the n-gate. We observe that significantly higher current levels and steeper switching slopes can be obtained with 2D-2D face tunneling (i.e., the EHBTFET). This signals that, once isolated from parasitic effects

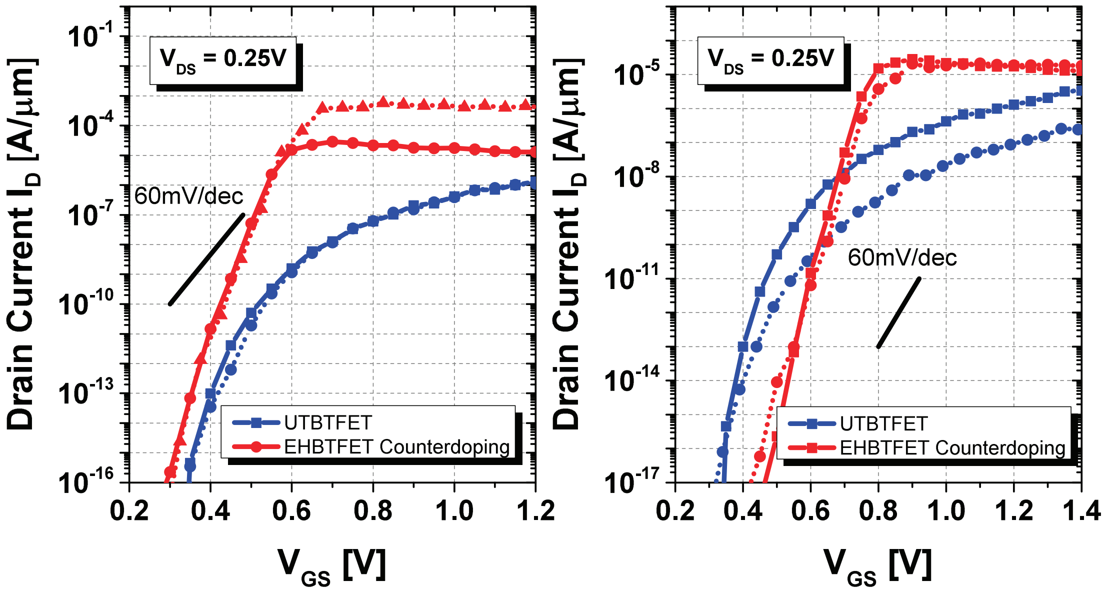


Figure 4.48 – Transfer characteristics comparison for (red) UTBTFET and (blue) EHBTFET with $In_{0.53}Ga_{0.47}As$ channel. Note that the curves are shifted rigidly along the x-axis for ease of viewing together and normalizing the OFF current. (Left) Channel thickness variations: (Solid) T_{CH} = 10nm (Dotted) T_{CH} = 7.5nm, EOT = 1nm. (Right) Effective oxide thickness variations: (Solid) EOT = 1nm (Dotted) EOT = 3nm, T_{CH} = 10nm.

such as the lateral tunneling, utilizing 2D-2D face tunneling could be a viable ON current booster.

Another interesting contrast between the 2D-2D face and 2D-2D edge tunneling is that the channel thickness essentially has no impact on the maximum obtainable current for 2D-2D edge tunneling, as the T_{CH} = 7.5nm and T_{CH} = 10nm cases converge to the same current value at high V_{GS}. This is most likely caused by the fact that the BTBT is concentrated around a very narrow region for the UTBTFET, where the channel thickness has only an indirect impact (i.e., by changing the electrostatics around the tunneling junction). On the other hand, the channel thickness directly determines the distance between the electron and hole gases in the EHBTFET, hence has a much stronger impact.

We also note an important difference between edge-tunneling and face-tunneling devices. Face tunneling devices tend to exhibit a saturation behavior due to the fact that quantization occurs along the direction that is transverse to the direction at which carriers enter into the device. This result alters the relationship between the electrostatics and the DOS involved in tunneling. A monotonous increase of drain current with increasing gate voltage is typically seen for edge tunneling devices, where the carrier injection direction from the source (drain) is aligned with the tunneling direction.

4.7.2 EHBTFET: 2D-2D Face vs 1D-1D Face Tunneling

It should be noted that dimensionality effects are inherently intertwined with the electrostatic effects in TFETs. To put it in clear terms, let us give an example case of a bulk TFET and a nanowire one. It is of course expected that the nanowire TFET outperforms the bulk one. However, it is not at all obvious which fraction of this improvement comes from the electrostatic improvement. The electrostatic control of the nanowire TFET is much better than the bulk one, so it is expected that a significant portion of the performance improvement should result from it. It is however an open question whether the reduction in the carrier dimensionality (3D gas in bulk vs 1D gas in nanowire) is also beneficial for the BTBT current. In the remainder of this section, however, we will utilize the EHBTFET (a 2D-2D tunneling device) rather than a bulk TFET.

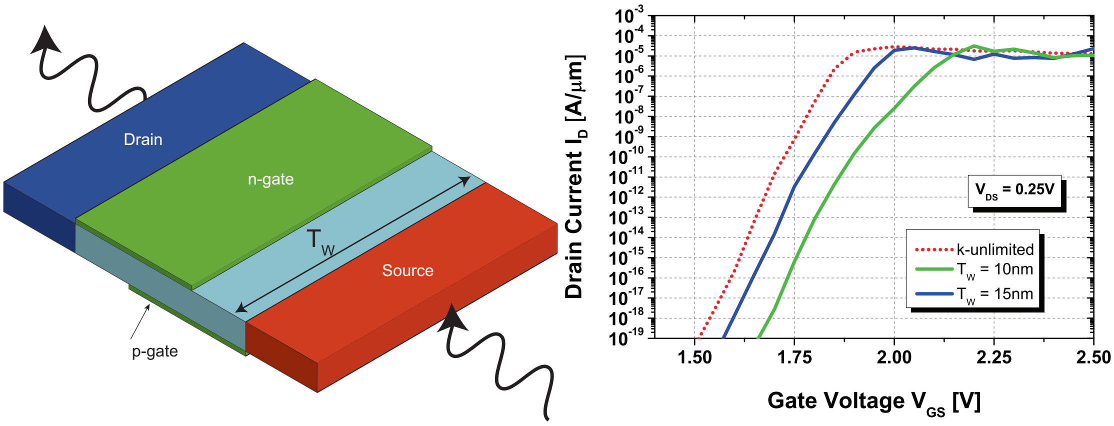


Figure 4.49 – (Left) The EHBTFET structure with transverse direction thickness T_W indicated (Right) Transfer characteristics for InGaAs EHBTFET with counterdoping for varying transverse thicknesses T_W.

We use the 2D simulator with transverse quantization model (section 3.6) to compare 2D-2D and 1D-1D face tunneling cases. We consider an InGaAs EHBTFET with counterdoping to study the impact of limiting the k states in the y-direction (width direction). We simulate the cases with different transverse thicknesses T_W in Fig. 4.49. The first observation is a shift of the onset voltage, which is expected, due to increased band gap associated to the size-induced quantization along the transverse direction y. Most strikingly, it is seen that the ON current density and the slopes are similar, which signals that no visible benefit is gained as the carrier dimensionality switches from 2D to 1D. The slight degradation of the slope for the $T_W = 10$nm case is caused by the formation of inversion layers in the overlap regions at high gate voltages. Since we do not include the transverse quantization while calculating the potential profile, we are essentially neglecting the related quantum capacitance effects that are induced by the electron concentration inside the channel region. Moreover, we remind that we utilize the same band parameters as in the bulk for both 1D and 2D cases. However, these parameters might depend on the exact dimensions of the device and they might deviate from their bulk values. We already observed that in section 3.7, for InAs QW diodes, the effective mass values changed slightly.

4.7.3 UTBTFET: 2D-2D Edge vs 1D-1D Edge Tunneling

Another aspect that can be investigated using the modified EMA code is comparing 2D-2D and 1D-1D edge tunneling. To this end, we simulate and compare an InGaAs UTBTFET with and without transverse k-space limitation. Different transverse thicknesses are considered in Fig. 4.50.

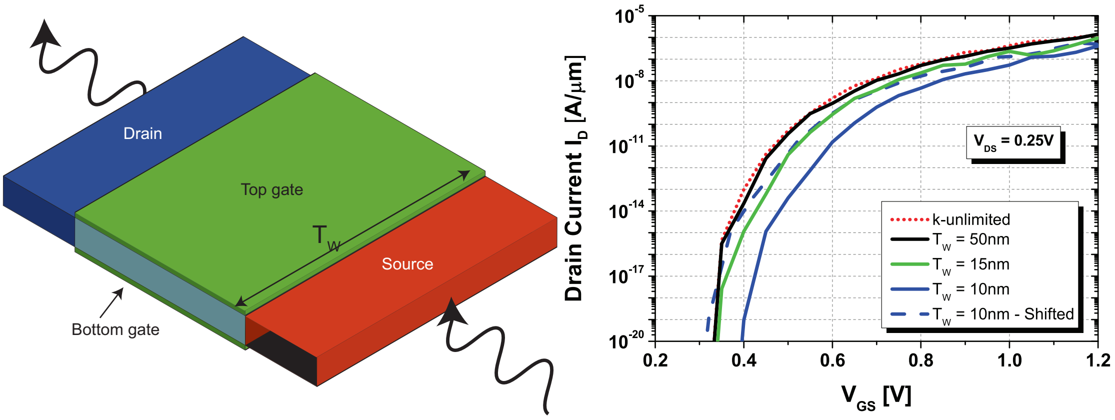


Figure 4.50 – (Left) UTBTFET structure with transverse direction thickness T_W indicated (Right) Transfer characteristics for the InGaAs UTBTFET for different transverse thicknesses T_W.

In contrast to the face tunneling case, switching from 2D-2D edge to 1D-1D edge tunneling results in more evident changes in the transfer characteristics. Similar to the previous case, a shift of tunneling onset is seen as the transverse thickness is decreased as seen in Fig. 4.50. Moreover, differently from face tunneling, a reduction of the ON current and switching slope is observed in edge tunneling when going from 2D-2D to 1D-1D, as seen clearly from the shifted $T_W = 10$nm curve (blue, dashed).

Note that here we do not consider ’anti-crossing’ (section 3.7) since it is not incorporated in the 2D code. Inclusion of this effect might change the picture.

4.8 Chapter Summary & Highlights

The significant contributions in this chapter are:

- It was shown via quantum-mechanical self-consistent simulations that the EHBTFET behaves quite differently compared to a conventional TFET, due to the strong quantization at both sides of the tunneling path. Exploitation of both 2DEG and 2DHG results in a steep switching behavior caused by the staircase-like 2D DOS. Comparison with the semi-classical TCAD simulations shows that TCAD models overestimate the current and also underestimate the alignment voltage since they do not account for the quantized nature of the electron and hole inversion layers.

- SS of the EHBTFET was seen to be degraded by the leakage caused by the penetration of the wavefunction into the underlap region from the source/drain. For this reason, the 2D model of section 3.3 is a powerful tool to optimize the EHBTFET design with the aim to suppress the leakage paths through the underlap region and make the almost 1D tunneling in the overlap region dominate.
- Demonstration of the feasibility of utilizing a counterdoped structure in order to suppress the lateral leakage current using a quantum mechanical simulation method. It is shown that, depending on the body thickness and the considered biasing, there exists an optimum underlap doping value which preserves the ON current and the steep slope. It is seen that a good control of the doping profile in terms of the doping level, alignment and junction abruptness is critical for optimum performance.
- A parameter analysis study of the EHBTFET has been performed. InAs stands out as the most promising material among the considered ones; however, hetero-structures such as InAs-GaSb will most certainly perform better than a homojunction solution as they alleviate the need of a high bias to achieve subband alignment.
- An analytical model providing a very good approximation of the vertical potential profile of the EHBTFET is proposed and verified by comparing it with the numerical findings. The model succinctly explains the distinguishing features of 2D tunneling such as T_{OX} independence.
- A novel logic scheme for the EHBTFET is proposed. It makes use of the independent biasing property of the n- and p-gates and allows for a reduction in the transistor count and possibly performance improvement over state-of-the-art CMOS. The benchmark circuits such as fan-out-4 inverters are simulated for both the EHBTFET logic and CMOS to quantify the advantage of the EHBTFET logic over the CMOS version. It is found that the EHBTFET logic significantly outperforms the CMOS implementation for low supply voltage scenarios.
- Investigation of the geometrical quantization effects on the fin EHBTFETs. It is found that the fin EHBTFET suffers from corner effects caused by a weak electrostatic control around the fin corners. This effect causes the formation of a potential barrier between the substrate and the overlap regions degrading the conductivity for both sharp and smooth corners. Among the possible solutions to suppress this effect, corner doping is the only viable option that can fully recover the full current drive potential of the device. Trapezoidal fins suffer from a reduction of the ON current due to increased tunneling length.

5 Conclusions & Future Perspectives

In chapter 1, we provided a brief overview of the microelectronics industry and introduced issue of limited supply voltage scaling which is the main motivation for the introduction of TFETs. In chapter 2, we explained the working principle of the TFET and of the EHBTFET and provided a brief history/review of the research and development in TFET device design. We also made a literature survey on direct and phonon-assisted BTBT models. Our contributions for the remaining chapters are given as:

- **A closed-boundary Schrödinger-Poisson solver and quantum mechanical BTBT current models.** We introduced our quantum mechanical simulation scheme, which is based on the self-consistent solutions of Schrödinger and Poisson equations. We first detailed our 1D simulation code based on FDM, followed by the 2D variant which is based on FEM. We then proposed/implemented the direct band-to-band tunneling models based on Fermi's Golden Rule. The first one, originally proposed by Bigelow et al. [106] as a 1D model is extended to 2D for uniform structures such as p-n diodes. We also modified the model proposed by Schenk et al. [109] for tunneling under quantization and incorporated polarization dependent coupling elements. This modified model can be also used for non-uniform structures such as the EHBTFET.
- **Modelling the asymmetry of real and imaginary heavy hole branches.** Making use of real and imaginary band calculations from the University of Bologna obtained by the $k \cdot p$ method, we have investigated the impact of quantization on tunneling rates. More specifically, we have noticed the gross asymmetry between effective masses between the real and imaginary branches of the heavy hole subbands. We implemented a WKB-based model to account for this discrepancy and incorporated it in our 1D numerical model.
- **Quantum mechanical simulation of the EHBTFET and analysis of subband-to-subband tunneling.** We first showed that, using our 1D code, the Ge EHBTFET works as a subband-to-subband tunneling device where the BTBT current exhibits a step-like behavior due to the 2D DOS function. Each electron-hole subband alignment contributes

to tunneling which changes the curvature of the transfer characteristics. We also investigated the output characteristics of the EHBTFET and showed that under negative drain bias (forward bias region of the p-i-n diode) the EHBTFET may not conduct current unlike the MOSFET. This is a common characteristic of all TFETs due to the p-i-n diode property.

- **2D simulation of the EHBTFET and identification of lateral leakage.** Through our 2D simulations, we identified the lateral leakage of the EHBTFET as a very critical non-ideality, which was missed out by the previous simulation work based on the semi-classical approximation. We then proposed two different methods to suppress it, namely the counterdoping and the hetero-gate approaches and assessed their effectiveness.

- **Parametric study and analytical model explaining the EHBTFET dependences on geometrical and material properties.** We performed a parametric simulation study, where we quantified the impact of the choice of the channel material, oxide and channel film thicknesses. We have shown that, in 2D-2D tunneling, the step-like behavior causes a saturation in the transfer characteristics that is in contrast with the 3D-3D tunneling. We have shown that depending on the device parameters chosen, the EHBTFET could operate either as a 2D-2D or a 3D-3D tunneling device. We also proposed an analytical model that predicts the alignment voltage. We have shown that in the 2D-2D case the alignment occurs at a fixed electric field magnitude that is independent of the oxide thickness.

- **Identification of the corner effects in the Fin-EHBTFET.** We simulated the fin incarnation of the EHBTFET, and found out another non-ideal effect particular to the fin-EHBTFET, which is the corner effect. We proposed and benchmarked different solution strategies and found that the corner doping seems to offer the best performance, although its practical implementation could pose a significant challenge.

- **EHBTFET logic.** We proposed and simulated a novel logic scheme making use of the independent biasing scheme of the EHBTFET and verified its functionality using quantum-corrected mixed-mode TCAD simulations. This scheme could enable a sizable reduction in the transistor count. We have also compared this logic scheme against 28nm FD-SOI CMOS using fan-out-4 inverter and XOR2 circuits and found significant performance advantage of the EHBTFET logic for low supply voltage scenarios.

- **Comparison of different dimensionality cases.** Using different variants of our code, we analyzed the EHBTFET (2D-2D face tunneling), the ultra thin body TFET (2D-2D edge tunneling) and their 1D counterparts (where the transverse direction has finite width). Comparing the 2D-2D face and edge tunneling cases, we found that the channel thickness has a strong impact on the ON current level for face tunneling, whereas in the UTBTFET case, the impact seems to be more limited; possibly due to the fact that tunneling is concentrated around a narrow region.

5.1 Future Perspectives

5.1.1 Experimental Verification of the EHBTFET

We would like to highlight here that work is underway in our lab within another PhD thesis which aims to fabricate the EHBTFET using the Si-InAs heterostructure system developed in IBM [155]. IBM's approach utilizes the so-called template-assisted selective epitaxy (TASE) which was briefly mentioned in Chapter 4. The method uses hollow templates of SiO_2 with a Si seed layer inside to grow high quality III-V material laterally [130].

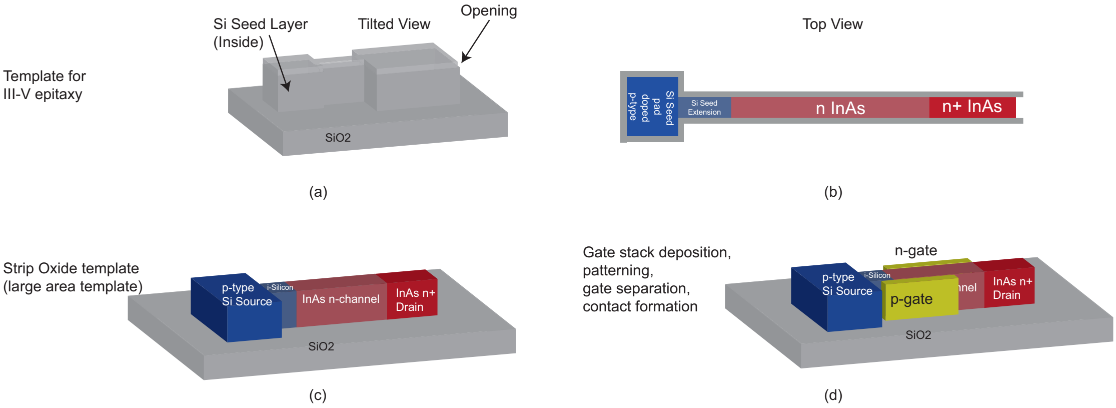


Figure 5.1 – Overall sketch of the EHBTFET fabrication process using TASE.

The idea is to make use of the already demonstrated capabilities of the TASE process developed by IBM to implement the EHBTFET. The channel and the drain region are n and n+ doped InAs, whereas the source is Si. While there exists a huge band offset between the Γ conduction band edges of Si and InAs, the valence band offset is about 130meV [156], which is expected to allow for a good supply of holes into the overlap region. We performed initial simulation and provided feedback to the fabrication efforts. Fig. 5.1 shows the general process flow for the fabrication of the device.

Fig. 5.2 shows the initial simulation results using the 2D code. We used larger electron effective mass ($m_e^* = 0.04\text{m}_0$) for InAs to partially mimic the strong non-parabolicity. The results indicate the device operation and consistent alignment voltage and ON current levels with the 1D parametric simulation results (Fig. 4.28). The initial simulations indicate the device indeed performs as a 2D-2D tunneling device and can be considered as a 'proof of concept' device whose performance is hindered by the lateral tunneling as expected, since no countermeasures proposed in the section 4.3 can be applied with the current status of the fabrication process. We highlight that the results presented here should be taken only as an early indicator of the device characteristics and a more in-depth analysis is warranted.

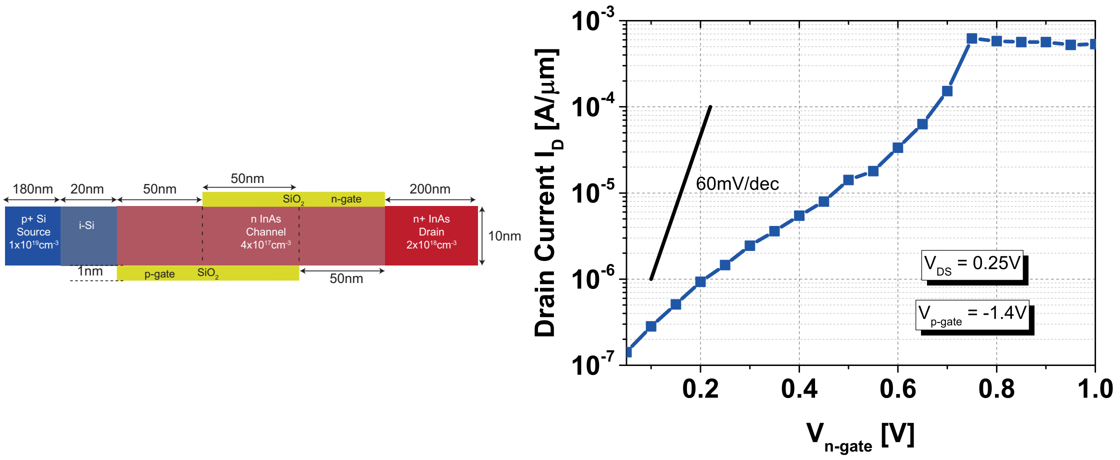


Figure 5.2 – (Left) Simulated device structure. The n- and p-gate metal workfunctions assumed to be midgap energy of InAs. (Right) Transfer characteristics of the Si-InAs EHBTFET. $V_{\text{p-gate}} = -1.4\text{V}$, $V_{\text{DS}} = 0.25\text{V}$.

5.1.2 Possible Future Work

For each of the contributions mentioned above, we can propose possible paths to continue, improve and complement the work done here. Specifically:

- For the simulator, implementing the non-parabolicity corrections for the 2D code would greatly enhance the applicability of the code by enabling the simulation of low bandgap materials such as InAs. Possible remedies for overcoming the numerical difficulty of reproducing the wavefunction tail for orders of magnitude could be include a reformulation of the model in a finite-precision friendly fasion or utilizing an arbitrary precision arithmetic package such as the *Arbirary Precision Toolbox* [157] in MATLAB. Moreover, the series resistance associated with the source/drain electrodes could be critical for low power and low supply voltage devices such as those investigated in this thesis. For this reason, it could be of interest to model the contact resistances in our simulation methodology.

- To better quantify performance of the novel TFET architectures and compare against the CMOS, a circuit-level simulation scheme including the quantum effects would be highly beneficial. The method we utilized in section 4.6 is inherently semi-classical. Right now, such simulations are usually based on lookup tables created by parametric device-level simulations to obtain the small signal AC capacitance matrices and DC current levels [131, 158]. As what can be done one step further, we are left with two options here. The first one is to develop DC and AC compact models for each of these devices, which can easily be coded in Verilog-A or SPICE and utilized in commercial circuit simulators. The other option is to extend the device-level quantum simulators to include also circuit simulation functionality similar to mixed-mode TCAD simulations available in Sentaurus [57].

- Regarding the EHBTFET, first efforts should be concentrated on experimental verification of the device working principle as already mentined above. Any one of the different incarnations that we proposed (the ’planar’ version, the fin EHBTFET or the nanowire one mentioned in the appendix E.2) could be tried. Once this is achieved, the parametric analyses in Chapter 4 will provide guidance for experimental researchers.
- As mentioned multiple times in the thesis, the alignment voltage V_{align} requirements for the EHBTFET could be rather large, due to very strong quantization effects. Improving the effective oxide thickness of the gate stacks would lower the voltage requirements since it would take a smaller bias difference to reach the aligment field F_{align}. To achieve this, some advances need to be made in incorporating extreme high-k insulators (possibly Tantalum- or Zirconium-based [159]). Gate leakage estimation can be performed by using the 1D gate leakage model detailed in Appendix C.
- The optimization study could be extended to include hetero-junctions. Lower alignment voltages could be obtained by creating staggered gap heterojunctions in the channel such as InAs/GaSb or novel 2D semiconducting materials such as MoS_2 or WSe_2 or combining them to form van der Vaals gap structures. Similarly, strain can be utilized to enhance the performance of III-V based TFETs [70] and GeSn alloys can promote direct BTBT [160] compared to the pure Ge considered in this thesis and this may result in a sizable improvement of the ON current.
- The EHBTFET logic idea could be explored further to pinpoint applications where it could have drastic advantages over CMOS implementations.

A Pseudo 2D Model

The method described in the section 3.3 could be computationally demanding compared to the 1D (section 3.2) and semi-classical simulations. Significant computational gains can be obtained by employing the ’pseudo 2D’ method in which the device is dissected into 1D slices and the Schrödinger equation is solved only in slices. In addition to being computationally much more efficient, this method has the advantage of access to many advanced 1D models such as band non-parabolicity [100]; which are not implemented in the full 2D model of section 3.3. Although the aformentioned properties make this approach highly desirable, the question of whether intrinsically 2D effects (such as short channel effects (SCE)) can be properly included remains to be answered.

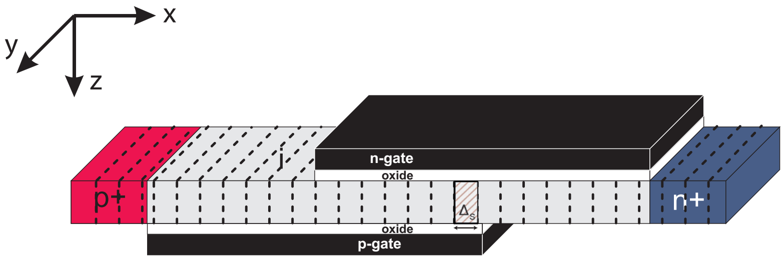


Figure A.1 – EHBTFET structure and the slices in which 1D Schrödinger equation is solved.

For the description of the method, we take the slicing direction as z, the other direction in the simulation domain as x, and the transverse direction (extending normal to the simulation domain) as y (see Fig. A.1). For the self-consistent loop, the quantum mechanical charge density inside a slice is calculated using Eqs. 3.2. Similarly, the 1D BTBT models can readily be used to get the tunneling current inside the slice. It is obvious that the method will only calculate the tunneling in the direction of the slice so it works for the EHBTFET or similar vertical (line [45]) tunneling structures, but not for DG-TFET where the tunneling direction is bias and geometry dependent. The tunneling current will be given as the integral of the current calculated per slice:

$$I_{\mathrm{2D}} = \int_{0}^{\mathrm{L_{x,max}}} I_{\mathrm{1D}}(x)\,dx \tag{A.1}$$

The 2D charge distribution is obtained as an extrapolation of the obtained 1D charge distributions:

$$n_{\mathrm{2D}}(x,z) = \left.\frac{n_{\mathrm{1D}}(z)}{\Delta_{\mathrm{S}}(x)}\right|_{x=x_i} \tag{A.2}$$

where $\Delta_{\mathrm{S}}(x)$ is the slice length along the x-direction at the point x_i.

Fig. A.2 compares the transfer characteristics obtained for the simulation of a Ge EHBTFET introduced in Fig. 2.5. For the pseudo 2D method, the EHBTFET is divided into equally separated vertical slices, since it is expected to have considerable tunneling in the vertical direction. The onset of direct BTBT indicates that pseudo 2D model is unable to account for the lateral tunneling that we have in Chapter 4.

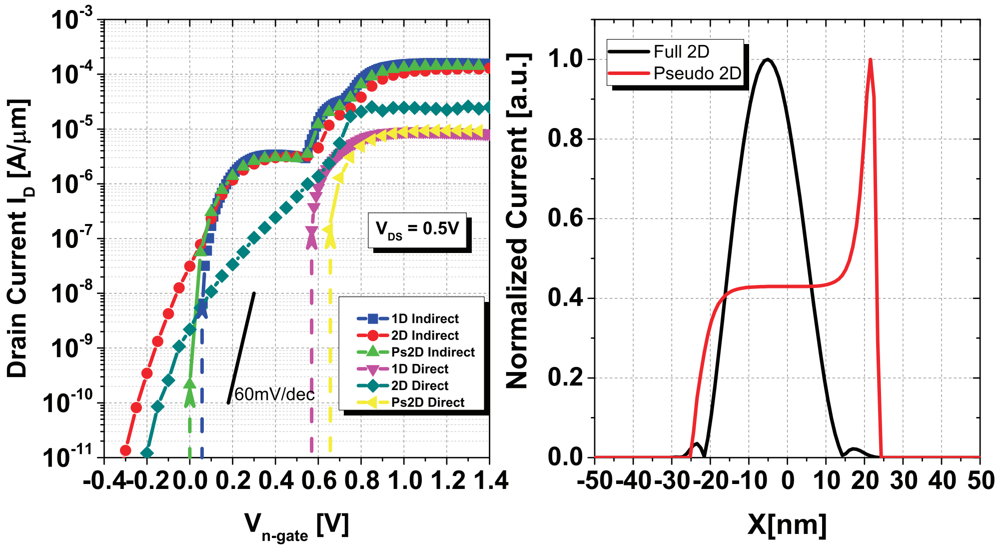


Figure A.2 – (Left) Comparison of different simulation schemes for the Ge EHBTFET described in Fig. 2.5 with gate metal workfunctions $\phi_{\mathrm{n-gate}} = 3.408\mathrm{eV}$ and $\phi_{\mathrm{p-gate}} = 5.642\mathrm{eV}$. It is seen that at high V_{GS} where the device is in the ON state, both phonon-assisted and direct tunneling components of 1D and 2D tunneling models converge to similar values. $V_{\mathrm{p-gate}} = 0\mathrm{V}$, $V_{\mathrm{DS}} = 0.5\mathrm{V}$. (Right) Normalized current densities per slice in the overlap region for 2D and pseudo-2D models.

This being said, it should be noted that the pseudo 2D method is able to capture only some of the 2D effects that are of purely electrostatic origin. Fig. A.2 (Right) shows the the BTBT distribution per vertical slice in the ON state (therefore the majority of tunneling taking place is in vertical direction), it is evident that the pseudo 2D results indicate a reduction of 'effective' gate length in the left overlap-underlap interface, which reduces the tunneling area as well as the higher BTBT rate around the region at the interface on the right. This reduction is caused by the lack of electrostatic control at the interface of underlap and overlap regions due to the fact that underlap regions are controlled by only one gate compared to double gate control in the overlap region. Moreover, absence of doping in the channel increases the characteristic decay length of the potential profile significantly, which further amplifies the SCEs. (see Fig. A.3).

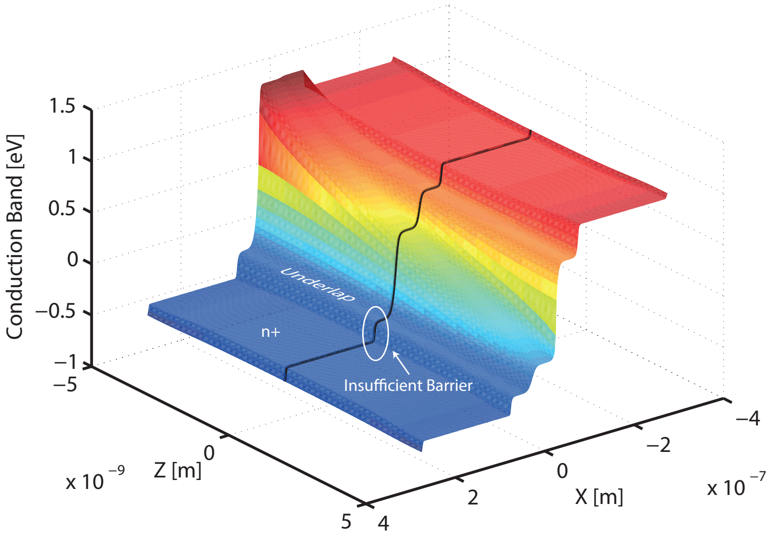


Figure A.3 – Γ conduction band profile of the Ge EHBTFET with T_{CH} = 10nm, $L_{\mathrm{OL}} = L_{\mathrm{UL}}$ = 50nm. The 1D outline (black line) indicates a long characteristic length due to the channel region being undoped (intrinsic).

To conclude, it is seen that pseudo 2D method is unable to determine the leakage current that is present in full 2D simulations, essentially yielding no improvement over the 1D model.

B C_0 Derivation Preliminaries & Definitions

We are following the discussion given in S.L. Chuang's *Physics of Photonic Devices* [110]. We are interested in calculating the coupling element between a conduction band and a valence band state around the Γ extrema $\mathbf{p}_{cv} = \frac{1}{\Omega_0}\int\limits_{\Omega_0} u_c^*(\vec{r})\frac{\hbar}{j}\nabla u_v(\vec{r}) d^3\vec{r}$ where Ω_0 is the unit cell. The band-edge wave basis vectors according to Kane [56] (for any electron wave vector $\vec{k} = k\sin\theta\cos\phi\vec{x} + k\sin\theta\sin\phi\vec{y} + k\cos\theta\vec{z}$) is given as follows:

- **Conduction Band**

$$|iS\downarrow\rangle \tag{B.1}$$

$$|iS\uparrow\rangle \tag{B.2}$$

- **Heavy Hole Band**

$$|\frac{3}{2},\frac{3}{2}\rangle = \frac{-1}{\sqrt{2}}|(\cos\theta\cos\phi - j\sin\phi)X + (\cos\theta\sin\phi + j\cos\phi)Y - \sin\theta Z\rangle|\uparrow\rangle \tag{B.3}$$

$$|\frac{3}{2},-\frac{3}{2}\rangle = \frac{1}{\sqrt{2}}|(\cos\theta\cos\phi + j\sin\phi)X + (\cos\theta\sin\phi - j\cos\phi)Y - \sin\theta Z\rangle|\downarrow\rangle \tag{B.4}$$

- **Light Hole Band**

$$|\frac{3}{2},\frac{1}{2}\rangle = \frac{-1}{\sqrt{6}}|(\cos\theta\cos\phi - j\sin\phi)X + (\cos\theta\sin\phi + j\cos\phi)Y - \sin\theta Z\rangle|\downarrow\rangle \tag{B.5}$$

$$+\sqrt{\frac{2}{3}}|(\sin\theta\cos\phi)X + (\sin\theta\sin\phi)Y + \cos\theta Z\rangle|\uparrow\rangle \tag{B.6}$$

$$|\frac{3}{2},-\frac{1}{2}\rangle = \frac{1}{\sqrt{6}}|(\cos\theta\cos\phi + j\sin\phi)X + (\cos\theta\sin\phi - j\cos\phi)Y - \sin\theta Z\rangle|\uparrow\rangle \tag{B.7}$$

$$+\sqrt{\frac{2}{3}}|(\sin\theta\cos\phi)X + (\sin\theta\sin\phi)Y + \cos\theta Z\rangle|\downarrow\rangle \tag{B.8}$$

where $|S, X, Y, Z\rangle$ are the s, p_x, p_y, p_z-like basis functions, respectively. We now note that p_φ (where φ=x, y or z) is odd in direction φ and even in the other two directions. Without loss of generality, the derivative of $|X\rangle$ will be odd in y and z directions. Due to this symmetry property, the momentum matrix elements $p_{y,z}$ will be 0 for $|X\rangle$.

C 1D Gate Leakage Model

Two different gate leakage models were implemented in the 1D simulator: open boundary and closed boundary models. In the open-boundary model (OBM), gate leakage is estimated using NEGF [161]. The OBM model works as follows: The electrostatics is solved using a 1D self-consistent Schrodinger-Poisson equation solver. Then, the gate tunneling is calculated using open boundary conditions (NEGF), each quasi-bound state (QBS) is detected in the Density of States (DOS) curve, whose full-width-half-maximum (FWHM) gives the ’escape time’ for each QBS [162]. A similar method is proposed in [163].

On the other hand, the closed boundary model (CBM) uses the semiclassical approximation (following [164]), where the electrons are considered as strictly classical particles trapped in a potential well which have T_{WKB} probability to ’escape’ from the well, at an attempt rate calculated using an effective transversal time calculated using well-known classical kinetic energy formula.

Once the escape times are known (with OBM or CBM), the tunneling current is calculated by combining it with the Fermi occupancies and the 2D density of states [165].

C.1 Open Boundary Model Algorithm

1. Solve the electrostatics using closed boundary conditions (oxide penetration included) using the closed boundary model explained in section 3.2.

2. Feed the closed boundary system eigenvalues as ’initial guesses’ to NEGF algorithm. Otherwise, it is computationally very expensive to detect all the peaks using NEGF since it requires a very dense energy mesh, due to the fact that in systems with low contact interation (i.e. thick gate oxide) the life-times tend to be large (see Fig. C.1 for a comparison of DOS peaks for thick and thin oxide stacks), and consequently the DOS peaks are very narrow. In order to keep the computational burden tractable, NEGF is only utilized around the eigenvalues of the closed system, which serve as an initial guess for the QBS levels.

3. Solve the open-boundary Schrodinger Equation using the NEGF. Calculate using [162]: $\mathrm{DOS}(E) = \int \mathrm{LDOS}(z, E)\, dz$

4. Detect the peaks in the DOS curve. Calculate the broadening ΔE by calculating the full width half maximum (FWHM). The 'broadening' of the peaks (FWHM = ΔE) is used to calculate the escape time as τ [73]: $\tau = \frac{\hbar}{\Delta E}$.

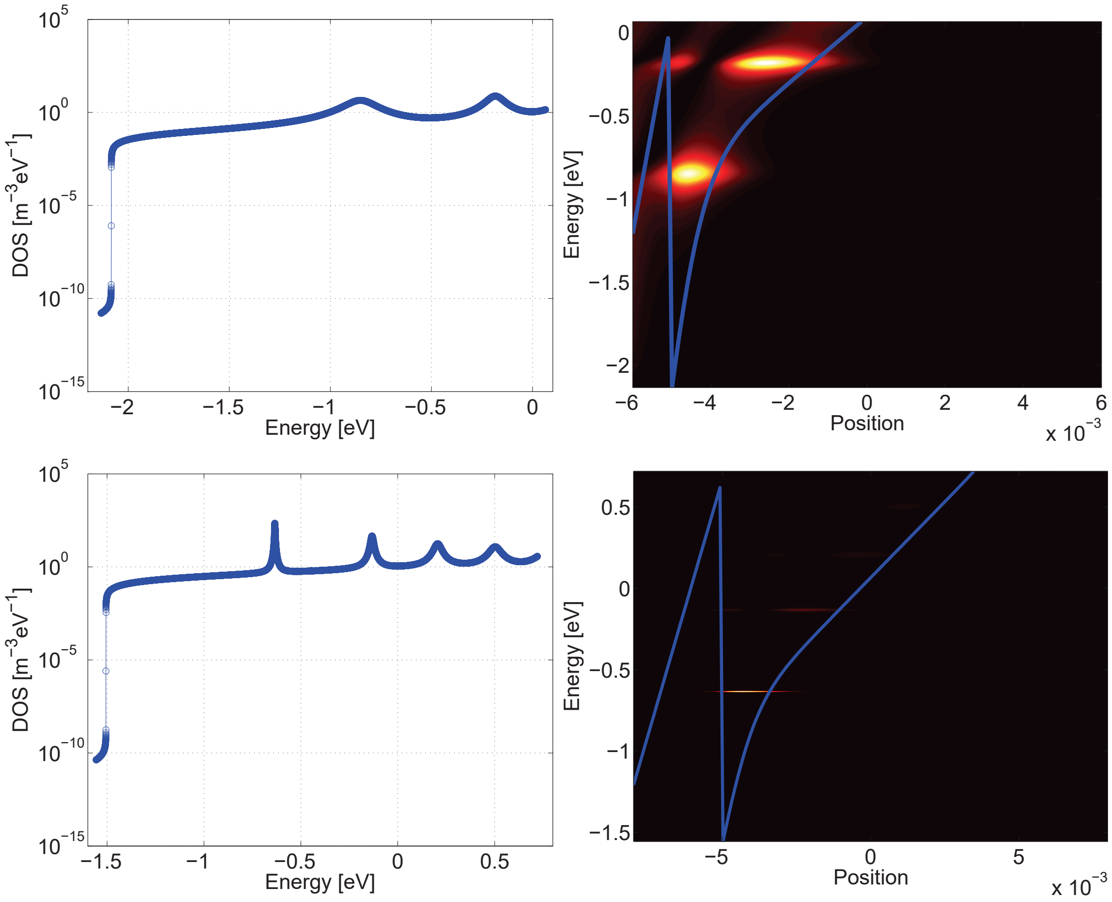


Figure C.1 – (Up left) DOS versus Energy. (Up right) LDOS and Conduction Band near the oxide-semiconductor interface for $T_{\mathrm{OX}} = 1$nm ($\epsilon_r = 22$ corresponding to $\mathrm{HfO_2}$). (Down left, Down right) the same as (a,b) but for $T_{\mathrm{OX}} = 3$nm. Note the larger spread in the DOS and LDOS for the $T_{\mathrm{OX}} = 1$nm case as the coupling increases between the contact and the quasi-bound states.

5. Calculate the error in the broadening ΔE. The error is calculated as:

$$\mathrm{err} = \frac{\mathrm{DOS}(E_{\mathrm{peak}})/2 - \mathrm{DOS}(E_{\mathrm{half\_max}})}{\mathrm{DOS}(E_{\mathrm{peak}})/2} \tag{C.1}$$

where $E_{\mathrm{half\_max}}$ is the energy closest to the FWHM of the peak. 'err' is calculated for both left and right hand sides of the peak and the final error is the average of the two.

Finally the broadening is given as the summation of the left and right side FWHM points: $\Delta E = E_{\text{half\_maxR}} - E_{\text{half\_maxL}}$.

6. For each iteration, the energy spacing dE is updated by $dE_{\text{new}} = dE_{\text{old}}/\zeta$ where ζ is a refinement parameter (which is currently set to 10). NEGF is utilized in the energy windows around the peaks only: $[E_{\text{peak}} - WdE, E_{\text{peak}} + WdE]$
7. Decide whether the half-maximum point has been left outside the energy window in the refinement process. If so, double the W parameter to broaden the energy window (using the same dE). Repeat steps 4, 5, 6 until a 'valid' life-time is obtained (i.e., within 0.5% error tolerance).
8. Once the life-times are calculated, the 2-D gate leakage current is obtained as [165]:

$$I_{\text{G}} = \sum_k \frac{qAL_{\text{G}}N_k}{\tau}\left(\ln(1+e^{\frac{E_{\text{FG}}-E_k}{kT}}) - \ln(1+e^{\frac{E_{\text{FD}}-E_k}{kT}})\right) \tag{C.2}$$

where k is the subband index running through all the QBS, $E_{\text{FG(FD)}}$ is the Fermi level for the gate (drain) contacts, $N_k = \frac{K_B T g_v m^*}{\pi\hbar^2}$ is the 2-D DOS factor, A is area factor, L_{G} is the gate length. $E_{\text{FG}} = -qV_{\text{G}}$ and $E_{\text{FD}} = -qV_{\text{D}}$ respectively.

C.2 Closed-boundary Model

Compared to the OBM, CBM is relatively straightforward to implement. The description of the algorithm is given below:

1. Solve the electrostatics using closed boundary conditions (oxide penetration included).
2. Calculate the WKB transmission probability T_{WKB} through the oxide:

$$T_{\text{WKB}} = e^{-2\int_0^{T_{\text{OX}}} |k(z)|dz} \quad \text{where} \quad k(z) = \frac{\sqrt{2m^*(E_k - E_{\text{C}}(z))}}{\hbar} \tag{C.3}$$

The calculation is done in the region where k is imaginary, and $z = 0$ is set as the left point in the barrier where $E_{\text{C}} = E_k$. However, WKB assumes a sufficiently constant potential profile, which is certainly not the case for our structure (we have a huge discontinuity at the oxide-semiconductor interface!). Therefore, interface reflections should also be taken into account. Some authors have done this via a pre-exponential term in front of the action integral [166].

3. Similarly, the transversal time of the electron t_t is determined as:

$$t_t = \int_0^{z_c} \sqrt{\frac{2m^*}{(E_k - E_{\text{C}}(z))}}dz \tag{C.4}$$

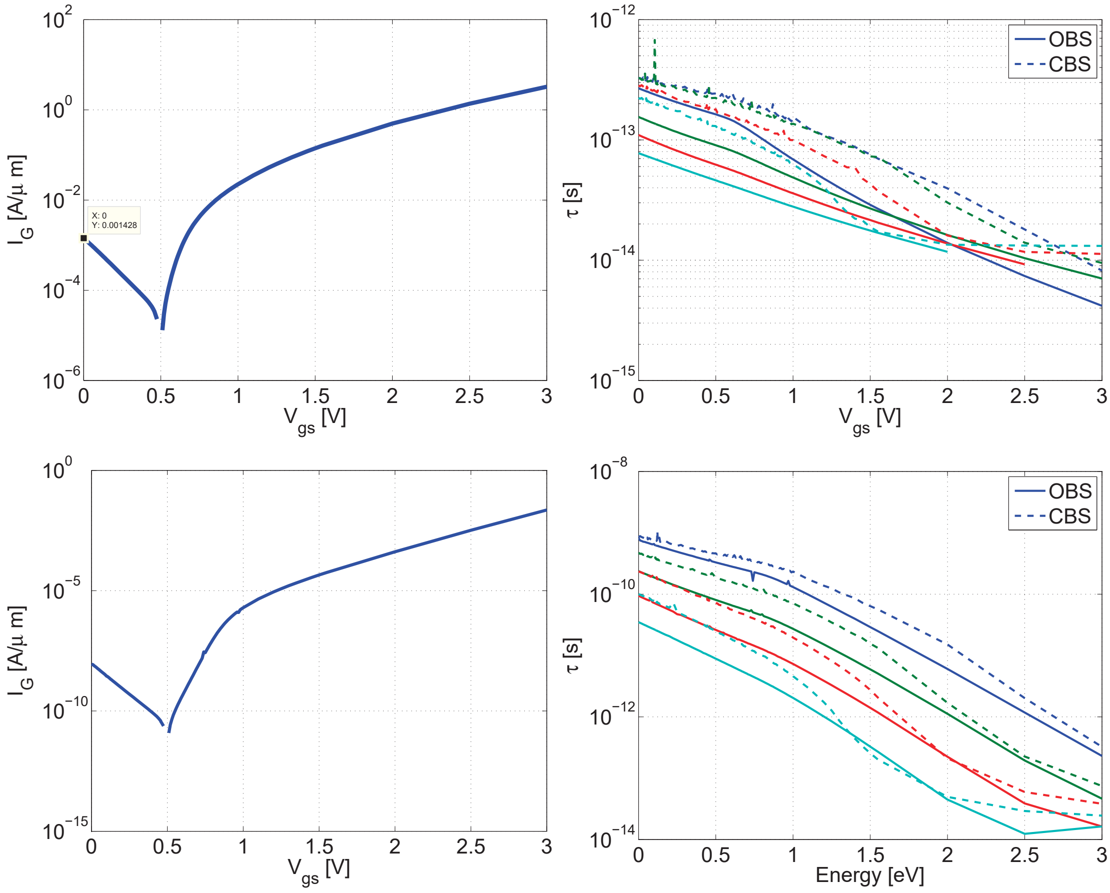


Figure C.2 – (a) I-V curves obtained for a Ge EHBTFET gate leakage for $T_{\mathrm{OX}} = 1$nm ($\epsilon_r = 22$ corresponding to HfO$_2$), $V_{\mathrm{DS}} = 0.5$V using OBC. (b) Electron lifetimes for OBS (solid) and for CBS model (dashed). (c,d) the same as (a,b) but for $T_{\mathrm{OX}} = 3$nm. In this simulations, p-gate is grounded and n-gate is swept (denoted as V_{GS}). $\phi_{\mathrm{n-gate}} = 3.434$eV and $\phi_{\mathrm{n-gate}} = 5.642$eV.

where z_c is the *classical turning point* defined as the point where $E_{\mathrm{C}}(z_c) = E_k$.

4. Finally, the life time in semi-classical approximation is calculated as: $\tau = T_{\mathrm{WKB}}/t_t$.

5. The same equation as OBS (eq. C.2) is then used to calculate the 1D tunneling current.

Comparable time-time values with the OBS system were obtained using the CBS model (see fig. C.2).

The initial simulations indicate significant amounts of gate current for our example of Ge EHBTFET with HfO$_2$ gate stack, especially at lower gate voltages which could cause problems. However, a complete, calibrated analysis for the gate leakage is warranted to give a reliable estimate of the gate leakage. Another concern to take note of is the two dimensional effects that increases the flow of electron around the edges of the gate stack [167], cannot be accounted for in our 1D model. In any case, it seems evident that extreme high-k oxides with adequate band offsets with the channel material is essential for an optimized EHBTFET.

D Derivation of the Analytical Model for the EHBTFET

We start from the 1D Poisson equation under depletion approximation for a 1D vertical slice of the EHBTFET:

$$\frac{\delta^2 V(z)}{\delta z^2} = \frac{qN_i}{\epsilon_{\mathrm{CH}}} \tag{D.1}$$

Integrating the above equation twice yields:

$$V(z) = \frac{qN_i}{2\epsilon_{\mathrm{CH}}} z^2 + Cz + D \tag{D.2}$$

where C and $D = V(0)$ are constants to be determined using the boundary conditions. Since we have an asymmetrically biased double gate structure, we impose two boundary conditions at both top and bottom oxide-semiconductor interface. Specifically, we impose the continuity of the displacement field:

$$\epsilon_{\mathrm{CH}} E_{\mathrm{CH}}(0) = \epsilon_{\mathrm{OX}} E_{\mathrm{OX}} \tag{D.3}$$

$$-\epsilon_{\mathrm{CH}} \frac{\delta V(z)}{\delta z}\bigg|_{z=0} = \frac{V_{\mathrm{n-gate}} - V(0)}{T_{\mathrm{OX}}} \tag{D.4}$$

$$C = \frac{\epsilon_{\mathrm{OX}}}{\epsilon_{\mathrm{CH}}} \left(\frac{V_{\mathrm{n-gate}} - D}{T_{\mathrm{OX}}} \right) \tag{D.5}$$

for the top-gate interface. The capacitance per length is defined as $C_{\mathrm{OX,CH}} = \frac{\epsilon_{\mathrm{OX,CH}}}{T_{\mathrm{OX,CH}}}$. Similarly for the bottom case, we have:

$$\frac{qN_i}{\epsilon_{\mathrm{CH}}} + C = \frac{C_{\mathrm{OX}}}{\epsilon_{\mathrm{CH}}} \left(V_{\mathrm{p-gate}} - \left(\frac{qN_i T_{\mathrm{CH}}^2}{2\epsilon_{\mathrm{CH}}} + CT_{\mathrm{CH}} + D \right) \right) \tag{D.6}$$

solving the Eqs. D.5 & D.6 for C and D yields:

$$C = \frac{C_{\mathrm{OX}}}{\epsilon_{\mathrm{CH}}} \left(D - V_{\mathrm{n-gate}} \right) \tag{D.7}$$

$$D = \frac{V_{\text{p-gate}} - \frac{qN_i T_{\text{CH}}^2}{2\epsilon_{\text{CH}}} + V_{\text{n-gate}}\left(\frac{C_{\text{OX}}}{C_{\text{CH}}} + 1\right) - \frac{qN_i T_{\text{CH}}}{C_{\text{OX}}}}{2 + \frac{C_{\text{OX}}}{C_{\text{CH}}}} \tag{D.8}$$

considering that we have very low doping, we can safely neglect the terms with N_i, which further simplifies the expressions:

$$C = \frac{C_{\text{OX}}}{\epsilon_{\text{CH}}}\left(D - V_{\text{n-gate}}\right) \tag{D.9}$$

$$D = \frac{V_{\text{p-gate}} + V_{\text{n-gate}}\left(\frac{C_{\text{OX}}}{C_{\text{CH}}} + 1\right)}{2 + \frac{C_{\text{OX}}}{C_{\text{CH}}}} \tag{D.10}$$

again neglecting the N_i term in Eq. D.2, plugging in C & D and writing C in terms of electric field $F = -C$:

$$V(z) = -Fz + \frac{V_{\text{p-gate}} + V_{\text{n-gate}}\left(\frac{C_{\text{OX}}}{C_{\text{CH}}} + 1\right)}{2 + \frac{C_{\text{OX}}}{C_{\text{CH}}}} \tag{D.11}$$

with $F = \frac{C_{\text{OX}}}{\epsilon_{\text{CH}}}\left(\frac{V_{\text{n-gate}} - V_{\text{p-gate}}}{2 + \frac{C_{\text{OX}}}{C_{\text{CH}}}}\right)$.

E Possible EHBTFET Fabrication Process and Different Incarnation of the EHBTFET

E.1 Fin EHBTFET Process Flow

One of the main advantages of the fin implementation described in section 4.5 is the relative ease of fabrication with respect to the planar version. In the fin case, having two different gates with asymmetric placement is not required. We sketch in Fig. E.1 two possible process flows for fabrication of the fin EHBTFET.

The two approaches differ from each other only in the fin formation steps, whereas the remaining processing steps remain the same. In the ’Top-Down’ approach the fin is created by first depositing the fin material onto the substrate with a deposition technique (e.g., ALD, MOCVD etc.) and then patterning it into a fin using photo-lithography. For the ’Bottom-Up’ approach, on the other hand, the substrate is first patterned and only the fin regions are left open for epitaxial growth.

The top-down approach begins with a bulk or SOI substrate. The fin material (may or may not be the same material as the substrate) is then epitaxially grown in a blanket fashion. Then the fin is patterned using a hard mask combined with an anisotropic physical etch. The hard mask is then removed.

The bottom-up approach also begins with the same substrate as in the top-down approach. This time, the entire substrate is covered by a hard mask except the fin region to prevent growth outside this region. Then, the fin material is epitaxially grown.

The remainder of the process is the same for both options. The gate stack materials (i.e., oxide and the n- and p-gate electrodes) are grown conformally using ALD. At this point, the fin is entirely covered by the gate stack. The spacers are deposited using ALD or CVD and etched anisotropically to achieve the desired structure. The separation of n- and p-gates is achieved by etching the metal and gate oxide on top of the fin using anisotropic etch or CMP process. The source and drain regions are obtained by implanting the source and drain regions and covering the remaining parts of the device.

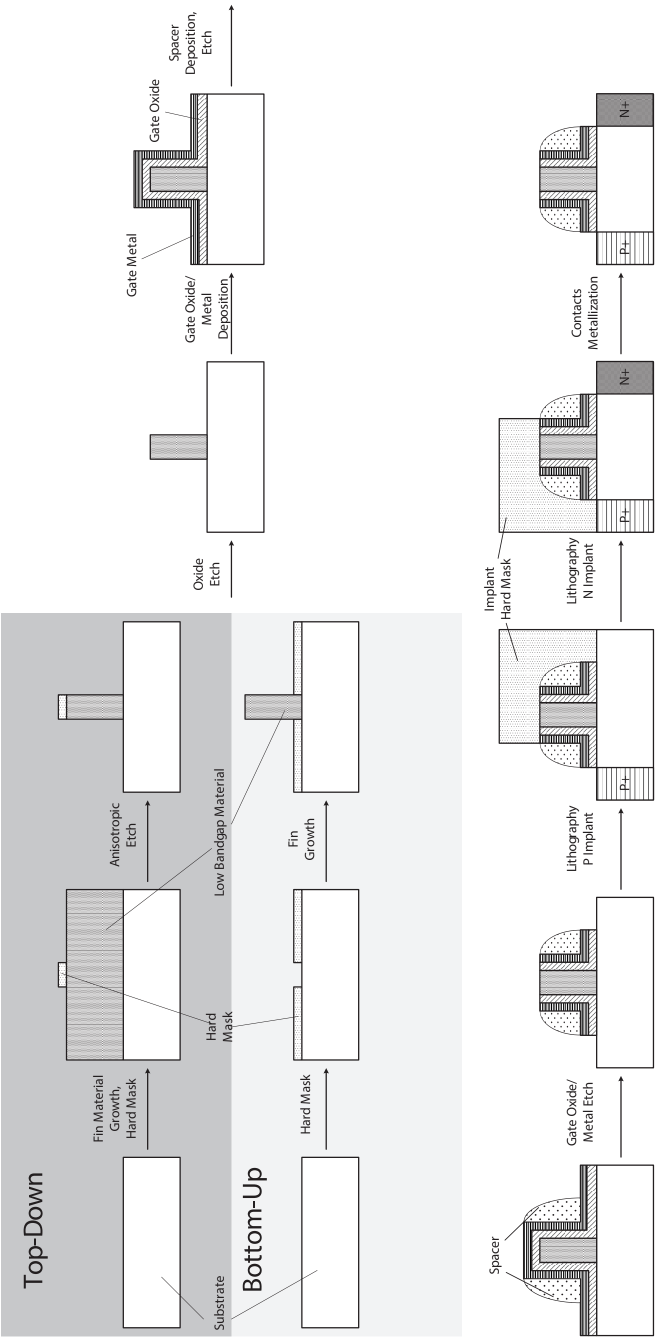


Figure E.1 – The process flow for fin EHBTFET using (i) Top-Down (ii) Bottom-Up approaches.

E.2 Nanowire Implementation of the EHBTFET

Here, we propose another idea that could possibly be used to implement the EHBTFET on a nanowire. It makes use of so-called ’core-shell’ structures for nanowires [168], which employs hollow semiconductor nanowires filled with another material. In the case of nanowire

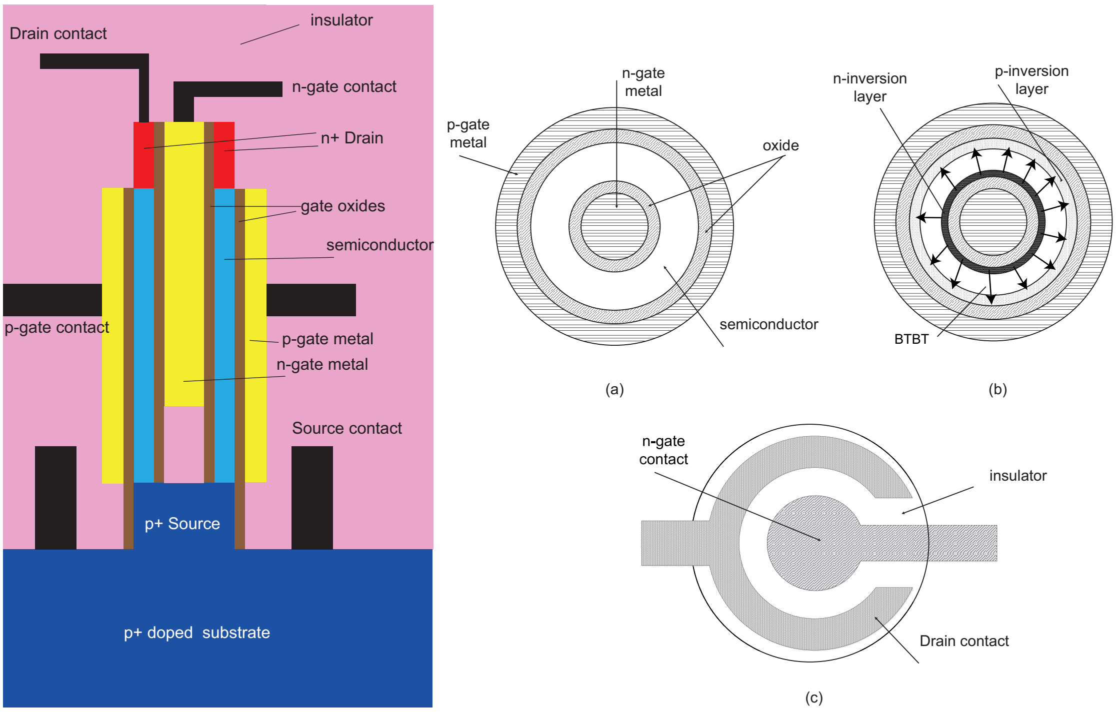


Figure E.2 – (Left) Nanowire cross section for the NW-EHBTFET.(Right) (a) A horizontal cross-section of the nanowire region, indicating different layers. (b) The induced electron and hole gases and BTBT direction in the ON state. (c) The top horizontal view showing the drain and n-gate contact separation.

EHBTFET the gate oxide, an insulating layer to separate the n-gate and the source and the gate metals are deposited conformally both in the hollow core and the outer shell, resulting in a sandwiched layer of semiconductor between two isolated gate stacks, which can effectively be operated as an EHBTFET. The vertical and horizontal cross sections, as well as the device working principle in the ON state is given in Fig. E.2. We have already noted that it is critical to have a thin semiconductor region to achieve high ON currents, so ideally the semiconductor shell should be on the order of 10nm. These small features could indeed be a problem in the top contacts (i.e., n-gate and drain), but we note that the drain region could gradually extend as it grows out of the channel region (possibly by using a different material similar to what is done in NW-TFETS of Lund group [138]).

Bibliography

[1] Z. Or-Bach, "Is the cost reduction associated with IC scaling is over?." http://www.eetimes.com/author.asp?doc_id=1286363, 2012. [Online; accessed 14 November 2016].

[2] B. Pangrle, "A node by any other name." http://semiengineering.com/a-node-by-any-other-name/, 2014. [Online; accessed 21 November 2016].

[3] J. Appenzeller, Y.-M. Lin, J. Knoch, and P. Avouris, "Band-to-band tunneling in carbon nanotube field-effect transistors," *Physical Review Letters*, vol. 93, p. 196805, nov 2004.

[4] L. De Michielis, L. Lattanzio, P. Palestri, L. Selmi, and A. M. Ionescu, "Tunnel-FET architecture with improved performance due to enhanced gate modulation of the tunneling barrier," in *2011 69th Annual Device Research Conference (DRC)*, pp. 111–112, IEEE, jun 2011.

[5] L. Lattanzio, L. De Michielis, and A. M. Ionescu, "Electron-hole bilayer tunnel FET for steep subthreshold swing and improved ON current," in *2011 Proceedings of the European Solid-State Device Research Conference (ESSDERC)*, pp. 259–262, IEEE, sep 2011.

[6] Z. Ren, R. Venugopal, S. Goasguen, S. Datta, and M. S. Lundstrom, "nanoMOS 2.5: A two-dimensional simulator for quantum transport in double-gate MOSFETs," *IEEE Transactions on Electron Devices*, vol. 50, no. 9, pp. 1914–1925, 2003.

[7] C. Alper, P. Palestri, L. Lattanzio, J. Padilla, and A. Ionescu, "Two dimensional quantum mechanical simulation of low dimensional tunneling devices," *Solid-State Electronics*, vol. 113, pp. 167–172, nov 2015.

[8] E. Baravelli, E. Gnani, R. Grassi, A. Gnudi, S. Reggiani, and G. Baccarani, "Optimization of n- and p-type TFETs integrated on the same InAs/$\mathrm{Al_xGa_{1-x}Sb}$ technology platform technology platform," *IEEE Transactions on Electron Devices*, vol. 61, pp. 178–185, jan 2014.

[9] E. Kane, "Zener tunneling in semiconductors," *Journal of Physics and Chemistry of Solids*, vol. 12, pp. 181–188, jan 1959.

[10] Q. Zhang, S. Sutar, T. Kosel, and A. Seabaugh, “Fully-depleted Ge interband tunnel transistor: modeling and junction formation,” *Solid-State Electronics*, vol. 53, pp. 30–35, jan 2009.

[11] P. M. Solomon, “Universal tunneling behavior in technologically relevant P/N junction diodes,” *Journal of Applied Physics*, vol. 95, no. 10, p. 5800, 2004.

[12] W. G. Vandenberghe and M. V. Fischetti, “Deformation potentials for band-to-band tunneling in silicon and germanium from first principles,” *Applied Physics Letters*, vol. 106, no. 1, p. 013505, 2015.

[13] K.-H. Kao, A. S. Verhulst, W. G. Vandenberghe, B. Soree, G. Groeseneken, and K. De Meyer, “Direct and indirect band-to-band tunneling in germanium-based TFETs,” *IEEE Transactions on Electron Devices*, vol. 59, pp. 292–301, feb 2012.

[14] M. Luisier, A. Schenk, W. Fichtner, and G. Klimeck, “Atomistic simulation of nanowires in the sp3d5s* tight-binding formalism: from boundary conditions to strain calculations,” *Physical Review B*, vol. 74, p. 205323, nov 2006.

[15] S. Agarwal and E. Yablonovitch, “Pronounced effect of pn-junction dimensionality on tunnel switch threshold shape,” *Mesoscale and Nanoscale Physics*, vol. 46, pp. fmiii–fmiv, sep 2011.

[16] L. De Michielis, L. Lattanzio, and A. M. Ionescu, “Understanding the superlinear onset of tunnel-FET output characteristic,” *IEEE Electron Device Letters*, vol. 33, pp. 1523–1525, nov 2012.

[17] G. Moore, “Cramming more components onto integrated circuits,” *Proceedings of the IEEE*, vol. 86, pp. 82–85, jan 1998.

[18] “1965: "Moore’s law" predicts the future of integrated circuits.” http://www.computerhistory.org/siliconengine/moores-law-predicts-the-future-of-integrated-circuits/, 2014. [Online; accessed 22 November 2016].

[19] N. Collaert, *CMOS nanoelectronics: innovative devices, architectures, and applications*. Pan Stanford Publishing, 2012.

[20] A. Barinka and I. King, “IBM to pay globalfoundries 1.5 billion to take chip unit.” https://www.bloomberg.com/news/articles/2014-10-19/ibm-agrees-to-pay-globalfoundries-1-5-billion-to-take-chip-unit, 2014. [Online; accessed 22 November 2016].

[21] “Global semiconductor industry revenue growth from 1988 to 2018.” https://www.statista.com/statistics/266976/forecast-revenue-growth-in-the-semiconductor-industry-worldwide/, 2016. [Online; accessed 22 November 2016].

[22] B. Arnold, “Shrinking possibilities.” http://spectrum.ieee.org/semiconductors/design/shrinking-possibilities/, 2009. [Online; accessed 14 November 2016].

[23] T. Ytterdal, Y. Cheng, and T. Fjeldly, *Device modeling for analog and RF CMOS circuit design.* Wiley, 2003.

[24] S. M. Kang and Y. Leblebici, *CMOS digital integrated circuits: analysis and design.* McGraw-Hill, 2003.

[25] A. M. Ionescu and H. Riel, “Tunnel field-effect transistors as energy-efficient electronic switches.,” *Nature*, vol. 479, pp. 329–37, nov 2011.

[26] A. Rusu, G. A. Salvatore, D. Jimenez, and A. M. Ionescu, “Metal-ferroelectric-metal-oxide-semiconductor field effect transistor with sub-60mV/decade subthreshold swing and internal voltage amplification,” in *2010 IEEE International Electron Devices Meeting (IEDM)*, pp. 16.3.1–16.3.4, IEEE, dec 2010.

[27] N. Shukla, A. V. Thathachary, A. Agrawal, H. Paik, A. Aziz, D. G. Schlom, S. K. Gupta, R. Engel-Herbert, and S. Datta, “A steep-slope transistor based on abrupt electronic phase transition.,” *Nature Communications*, vol. 6, p. 7812, 2015.

[28] N. Abele, R. Fritschi, K. Boucart, F. Casset, P. Ancey, and A. Ionescu, “Suspended-gate MOSFET: bringing new MEMS functionality into solid-state MOS transistor,” in *2005 IEEE International Electron Devices Meeting (IEDM)*, pp. 479–481, IEEE, dec 2005.

[29] U. Zaghloul and G. Piazza, “Highly scalable NEMS relays with stress-tuned switching voltage using piezoelectric buckling actuators,” *IEEE Transactions on Electron Devices*, vol. 61, pp. 3520–3528, oct 2014.

[30] D. Grogg, C. L. Ayala, U. Drechsler, A. Sebastian, W. W. Koelmans, S. J. Bleiker, M. Fernandez-Bolanos, C. Hagleitner, M. Despont, and U. T. Duerig, “Amorphous carbon active contact layer for reliable nanoelectromechanical switches,” *Proceedings of the IEEE International Conference on Micro Electro Mechanical Systems (MEMS)*, pp. 143–146, 2014.

[31] K. Gopalakrishnan, P. Griffin, and J. Plummer, “I-MOS: a novel semiconductor device with a subthreshold slope lower than kT/q,” in *2002 IEEE International Electron Devices Meeting (IEDM)*, pp. 289–292, IEEE, 2002.

[32] A. Padilla, C. W. Yeung, C. Shin, C. Hu, and T. J. K. Liu, “Feedback FET: a novel transistor exhibiting steep switching behavior at low bias voltages,” *2008 IEEE International Electron Devices Meeting (IEDM)*, 2008.

[33] J. Knoch and J. Appenzeller, “A novel concept for field-effect transistors - the tunneling carbon nanotube FET,” in *2005 63rd Annual Device Research Conference (DRC)*, vol. 1, pp. 153–156, IEEE, 2005.

[34] V. V. Zhirnov and R. K. Cavin, “Nanoelectronics: negative capacitance to the rescue?,” *Nature Nanotechnology*, vol. 3, pp. 77–78, feb 2008.

[35] Qiang Li, Yu Han, Xing Lu, and Kei May Lau, “GaAs-InGaAs-GaAs Fin-Array Tunnel Diodes on (001) Si Substrates With Room-Temperature Peak-to-Valley Current Ratio of 5.4,” *IEEE Electron Device Letters*, vol. 37, pp. 24–27, jan 2016.

[36] C. H. P. Lorenz, S. Hemour, W. Li, Y. Xie, J. Gauthier, P. Fay, and K. Wu, “Breaking the Efficiency Barrier for Ambient Microwave Power Harvesting With Heterojunction Backward Tunnel Diodes,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, pp. 4544–4555, dec 2015.

[37] D. J. Griffiths, *Introduction to Quantum Mechanics (Second Edition)*. Pearson Education International, 2005.

[38] U. Mishra and J. Singh, *Semiconductor Device Physics and Design*. Springer, 2007.

[39] S. Agarwal, J. T. Teherani, J. L. Hoyt, D. A. Antoniadis, and E. Yablonovitch, “Engineering the electron-hole bilayer tunneling field-effect transistor,” *IEEE Transactions on Electron Devices*, vol. 61, no. 5, pp. 1599–1606, 2014.

[40] L. F. Luo, R. Beresford, and W. I. Wang, “Interband tunneling in polytype GaSb/AlSb/InAs heterostructures,” *Applied Physics Letters*, vol. 55, no. 19, p. 2023, 1989.

[41] M. Tsuchiya, H. Sakaki, and J. Yoshino, “Room temperature observation of differential negative resistance in an AlAs/GaAs/AlAs resonant tunneling diode,” *Japanese Journal of Applied Physics*, vol. 24, pp. L466–L468, jun 1985.

[42] R. Beresford, L. Luo, K. Longenbach, and W. Wang, “Resonant interband tunneling device with multiple negative differential resistance regions,” *IEEE Electron Device Letters*, vol. 11, pp. 110–112, mar 1990.

[43] T. Baba, “Proposal for Surface Tunnel Transistors,” *Japanese Journal of Applied Physics*, vol. 31, pp. L455–L457, apr 1992.

[44] C. Aydin, A. Zaslavsky, S. Luryi, S. Cristoloveanu, D. Mariolle, D. Fraboulet, and S. Deleonibus, “Lateral interband tunneling transistor in silicon-on-insulator,” *Applied Physics Letters*, vol. 84, no. 10, p. 1780, 2004.

[45] W. G. Vandenberghe, A. S. Verhulst, G. Groeseneken, B. Soree, and W. Magnus, “Analytical model for point and line tunneling in a tunnel field-effect transistor,” in *2008 International Conference on Simulation of Semiconductor Processes and Devices*, pp. 137–140, IEEE, sep 2008.

[46] F. Mayer, C. Le Royer, J.-F. Damlencourt, K. Romanjek, F. Andrieu, C. Tabone, B. Previtali, and S. Deleonibus, “Impact of SOI, $\mathrm{Si}_{1-x}\mathrm{Ge}_x$OI and GeOI substrates on CMOS compatible Tunnel FET performance,” in *2008 IEEE International Electron Devices Meeting (IEDM)*, vol. 4, pp. 1–5, IEEE, dec 2008.

[47] A. S. Verhulst, W. G. Vandenberghe, K. Maex, and G. Groeseneken, “Tunnel field-effect transistor without gate-drain overlap,” *Applied Physics Letters*, vol. 91, no. 2007, pp. 2–5, 2007.

[48] S. Richter, S. Trellenkamp, A. Schäfer, J. Hartmann, K. Bourdelle, Q. Zhao, and S. Mantl, “Improved tunnel-FET inverter performance with SiGe/Si heterostructure nanowire TFETs by reduction of ambipolarity,” *Solid-State Electronics*, vol. 108, pp. 97–103, jun 2015.

[49] M. Schmidt, A. Schafer, R. A. Minamisawa, D. Buca, S. Trellenkamp, J.-M. Hartmann, Q.-T. Zhao, and S. Mantl, “Line and point tunneling in scaled Si/SiGe heterostructure TFETs,” *IEEE Electron Device Letters*, vol. 35, pp. 699–701, jul 2014.

[50] C. Hu, P. Patel, A. Bowonder, K. Jeon, S. H. Kim, P. Majhi, A. Javey, T.-J. K. Liu, and R. Jammy, “Prospect of tunneling green transistor for 0.1V CMOS,” in *2010 IEEE International Electron Devices Meeting (IEDM)*, pp. 16.1.1–16.1.4, IEEE, dec 2010.

[51] S. H. Kim, H. Kam, C. Hu, and T.-J. K. Liu, “Germanium-source tunnel field effect transistors with record high $\mathrm{I_{ON}/I_{OFF}}$,” *2009 Symposium on VLSI Technology*, pp. 178–179, 2009.

[52] L. Lattanzio, N. Dagtekin, L. De Michielis, and A. M. Ionescu, “On the static and dynamic behavior of the germanium electron-hole bilayer tunnel FET,” *IEEE Transactions on Electron Devices*, vol. 59, pp. 2932–2938, nov 2012.

[53] J. Knoch and J. Appenzeller, “Tunneling phenomena in carbon nanotube field-effect transistors,” *Physica Status Solidi (a)*, vol. 205, pp. 679–694, apr 2008.

[54] L. De Michielis, N. Dagtekin, A. Biswas, L. Lattanzio, L. Selmi, M. Luisier, H. Riel, and A. M. Ionescu, “An innovative band-to-band tunneling analytical model and implications in compact modeling of tunneling-based devices,” *Applied Physics Letters*, vol. 103, no. 12, p. 123509, 2013.

[55] Y. Taur and J. Wu, “Examination of two-band $E(k)$ relations for band-to-band tunneling,” *IEEE Transactions on Electron Devices*, vol. 63, pp. 869–872, feb 2016.

[56] E. O. Kane, “Band structure of indium antimonide,” *Journal of Physics and Chemistry of Solids*, vol. 1, pp. 249–261, jan 1957.

[57] *Sentaurus Device user guide*, June 2014.

[58] S. Tanaka, “A unified theory of direct and indirect interband tunneling under a nonuniform electric field,” *Solid-State Electronics*, vol. 37, pp. 1543–1552, aug 1994.

[59] L. De Michielis, M. Iellina, P. Palestri, A. M. Ionescu, and L. Selmi, “Tunneling path impact on semi-classical numerical simulations of TFET devices,” in *2011 12th International Conference on Ultimate Integration on Silicon, ULIS 2011*, pp. 146–149, IEEE, mar 2011.

[60] A. Pan and C. O. Chui, “Modeling direct interband tunneling. I. Bulk semiconductors,” *Journal of Applied Physics*, vol. 116, p. 054508, aug 2014.

[61] A. Pan and C. O. Chui, “Modeling direct interband tunneling. II. Lower-dimensional structures,” *Journal of Applied Physics*, vol. 116, p. 054509, aug 2014.

[62] H. Carrillo-Nuñez, 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, p. 234501, jun 2015.

[63] L. Keldysh, “Influence of the lattice vibrations of a crystal on the production of electron-hole pairs in a strong electric field,” *Soviet Journal of Experimental and Theoretical Physics*, vol. 34, no. 4, p. 665, 1958.

[64] W. Vandenberghe, B. Soree, W. Magnus, and M. Fischetti, “Generalized phonon-assisted Zener tunneling in indirect semiconductors with non-uniform electric fields: a rigorous approach,” *Journal of Applied Physics*, vol. 109, no. 12, p. 124503, 2011.

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

[66] W. G. Vandenberghe, B. Soree, W. Magnus, M. Fischetti, A. S. Verhulst, and G. Groeseneken, “Two-dimensional quantum mechanical modeling of band-to-band tunneling in indirect semiconductors,” in *2011 IEEE International Electron Devices Meeting (IEDM)*, pp. 5.3.1–5.3.4, IEEE, dec 2011.

[67] H. Gummel, “A self-consistent iterative scheme for one-dimensional steady state transistor calculations,” *IEEE Transactions on Electron Devices*, vol. 11, pp. 455–465, oct 1964.

[68] M. Lundstrom, “Drift-diffusion and computational electronics - Still going strong after 40 years!,” *International Conference on Simulation of Semiconductor Processes and Devices, SISPAD*, vol. 2015-October, pp. 1–3, 2015.

[69] S.-H. Ke, H. U. Baranger, and W. Yang, “Electron transport through molecules: Self-consistent and non-self-consistent approaches,” *Physical Review B*, vol. 70, p. 085410, aug 2004.

[70] F. Conzatti, M. G. Pala, D. Esseni, E. Bano, and L. Selmi, “Strain-induced performance improvements in InAs nanowire tunnel FETs,” *IEEE Transactions on Electron Devices*, vol. 59, no. 8, pp. 2085–2092, 2012.

[71] E. Polizzi and N. B. Abdallah, “Subband decomposition approach for the simulation of quantum electron transport in nanostructures,” *Journal of Computational Physics*, vol. 202, no. 1, pp. 150–180, 2005.

[72] O. Morandi and F. Schürrer, "Wigner model for quantum transport in graphene," *Journal of Physics A: Mathematical and Theoretical*, vol. 44, p. 265301, jul 2011.

[73] S. Datta, "Nanoscale device modeling: the Green's function method," *Superlattices and Microstructures*, vol. 28, pp. 253–278, oct 2000.

[74] C. S. Lent and D. J. Kirkner, "The quantum transmitting boundary method," *Journal of Applied Physics*, vol. 67, no. 10, p. 6353, 1990.

[75] R. Venugopal, Z. Ren, S. Datta, M. S. Lundstrom, and D. Jovanovic, "Simulating quantum transport in nanoscale transistors: Real versus mode-space approaches," *Journal of Applied Physics*, vol. 92, no. 7, p. 3730, 2002.

[76] S. Cauley, M. Luisier, V. Balakrishnan, G. Klimeck, and C.-K. Koh, "Distributed non-equilibrium Green's function algorithms for the simulation of nanoelectronic devices with scattering," *Journal of Applied Physics*, vol. 110, no. 4, p. 043713, 2011.

[77] I.-H. Tan, G. L. Snider, L. D. Chang, and E. L. Hu, "A self-consistent solution of Schrödinger–Poisson equations using a nonuniform mesh," *Journal of Applied Physics*, vol. 68, no. 8, p. 4071, 1990.

[78] A. M. Walke, A. S. Verhulst, A. Vandooren, D. Verreck, E. Simoen, V. R. Rao, G. Groeseneken, N. Collaert, and A. V. Y. Thean, "Part I: Impact of field-induced quantum confinement on the subthreshold swing behavior of line TFETs," *IEEE Transactions on Electron Devices*, vol. 60, no. 12, pp. 4057–4064, 2013.

[79] S. Sant, K. Moselund, D. Cutaia, H. Schmid, M. Borg, H. Riel, and A. Schenk, "Lateral InAs/Si p-Type tunnel FETs integrated on Si—Part 2: simulation study of the impact of interface traps," *IEEE Transactions on Electron Devices*, vol. 63, pp. 4240–4247, nov 2016.

[80] S. R. Johnson and T. Tiedje, "Temperature dependence of the Urbach edge in GaAs," *Journal of Applied Physics*, vol. 78, no. 9, pp. 5609–5613, 1995.

[81] C. Grossmann, H.-G. Roos, and M. Stynes, *Numerical treatment of partial differential equations*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.

[82] F. Stern and W. Howard, "Properties of semiconductor surface inversion layers in the electric quantum limit," *Physical Review*, vol. 163, pp. 816–835, nov 1967.

[83] R. Bank and D. Rose, "Global approximate Newton methods," *Numerische Mathematik*, pp. 1–16, 1981.

[84] A. Trellakis, A. T. Galick, A. Pacelli, and U. Ravaioli, "Iteration scheme for the solution of the two-dimensional Schrödinger-Poisson equations in quantum structures," *Journal of Applied Physics*, vol. 81, no. 12, p. 7880, 1997.

[85] S. Smirnov, *Physical Modeling of Electron Transport in Strained Silicon and Silicon-Germanium*. PhD thesis, Technische Universität Wien, 2003.

[86] M. Bescond, N. Cavassilas, and M. Lannoo, “Effective-mass approach for n-type semiconductor nanowire MOSFETs arbitrarily oriented,” *Nanotechnology*, vol. 18, p. 255201, jun 2007.

[87] A. Rahman, M. S. Lundstrom, and A. W. Ghosh, “Generalized effective-mass approach for n-type metal-oxide-semiconductor field-effect transistors on arbitrarily oriented wafers,” *Journal of Applied Physics*, vol. 97, no. 5, p. 053702, 2005.

[88] *COMSOL Multiphysics Reference Manual*, 2013.

[89] F.-J. Sayas, “A gentle introduction to the finite element method an introduction.” Technical Report. http://www.math.udel.edu/~fjsayas/documents/anIntro2FEM_2015.pdf [Online; accessed 5 December 2016], 2015.

[90] K. E. Atkinson and W. Han, *Theoretical numerical analysis: a functional analysis framework*. Springer Science+Business Media, 2nd ed., 2005.

[91] L. Chen, “Programming of finite element methods in matlab.” Technical Report. http://www.math.uci.edu/~chenlong/226/Ch3FEMCode.pdf [Online; accessed 5 December 2016], 2011.

[92] M. Sabathil, *Opto-electronic and quantum transport properties of semiconductor nanostructures*. PhD thesis, Technische Universität München, 2004.

[93] D. Rideau, M. Feraille, M. Michaillat, Y. M. Niquet, C. Tavernier, and H. Jaouen, “On the validity of the effective mass approximation and the Luttinger k.p model in fully depleted SOI MOSFETs,” *Solid-State Electronics*, vol. 53, no. 4, pp. 452–461, 2009.

[94] D. Rideau, M. Feraille, L. Ciampolini, M. Minondo, C. Tavernier, H. Jaouen, and a. Ghetti, “Strained Si, Ge, and $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ alloys modeled with a first-principles-optimized full-zone $k \cdot p$ method,” *Physical Review B*, vol. 74, p. 195208, nov 2006.

[95] A. Esposito, M. Frey, and A. Schenk, “Quantum transport including nonparabolicity and phonon scattering: application to silicon nanowires,” *Journal of Computational Electronics*, vol. 8, pp. 336–348, jul 2009.

[96] J. A. Lopez-Villanueva, I. Melchor, P. Cartujo, and J. E. Carceller, “Modified Schrödinger equation including nonparabolicity for the study of a two-dimensional electron gas,” *Physical Review B*, vol. 48, pp. 1626–1631, jul 1993.

[97] L. Lattanzio, L. De Michielis, and A. M. Ionescu, “Complementary germanium electron–hole bilayer tunnel FET for sub-0.5V operation,” *IEEE Electron Device Letters*, vol. 33, pp. 167–169, feb 2012.

[98] L. Wang, P. M. Asbeck, and Y. Taur, “Self-consistent 1-D Schrödinger–Poisson solver for III–V heterostructures accounting for conduction band non-parabolicity,” *Solid-State Electronics*, vol. 54, pp. 1257–1262, nov 2010.

[99] S. Jin, M. Fischetti, and T.-w. Tang, “Modeling of electron mobility in gated silicon nanowires at room temperature: Surface roughness scattering, dielectric screening, and band nonparabolicity,” *Journal of Applied Physics*, vol. 102, no. 8, p. 083715, 2007.

[100] C. Troger, H. Kosina, and S. Selberherr, “Modeling nonparabolicity effects in silicon inversion layers,” in *SISPAD ’97. 1997 International Conference on Simulation of Semiconductor Processes and Devices. Technical Digest*, pp. 323–326, IEEE, 1997.

[101] L. Ridgway Scott, *Numerical Analysis.* Princeton University Press, 2011.

[102] J. Bigelow and J. Leburton, “Self-consistent modeling of resonant interband tunneling in bipolar tunneling field-effect transistors,” *IEEE Transactions on Electron Devices*, vol. 41, no. 2, pp. 125–131, 1994.

[103] C. Alper, M. Visciarelli, P. Palestri, J. L. Padilla, A. Gnudi, E. Gnani, and A. M. Ionescu, “Modeling the imaginary branch in III-V tunneling devices: effective mass vs $k \cdot p$,” in *2015 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 273–276, IEEE, sep 2015.

[104] T. Heinzel, ed., *Mesoscopic Electronics in Solid State Nanostructures.* Weinheim, Germany: Wiley-VCH Verlag GmbH, nov 2006.

[105] M. Yamanishi and I. Suemune, “Comment on polarization dependent momentum matrix elements in quantum well lasers,” *Japanese Journal of Applied Physics*, vol. 23, pp. L35–L36, jan 1984.

[106] J. Bigelow and J. Leburton, “Self-consistent modeling of bipolar tunnel heterostructures with quantum mechanical current,” in *International Technical Digest on Electron Devices*, pp. 767–770, IEEE, 1990.

[107] M. Kuno, *Introductory nanoscience: physical and chemical concepts.* Garland Science, 2011.

[108] A. Szabo, S. J. Koester, and M. Luisier, “Ab-Initio simulation of van der Waals $MoTe_2 - SnS_2$ heterotunneling FETs for low-power electronics,” *IEEE Electron Device Letters*, vol. 36, pp. 514–516, may 2015.

[109] A. Schenk, M. Stahl, and H.-J. Wünsche, “Calculation of interband tunneling in inhomogeneous fields,” *physica status solidi (b)*, vol. 154, pp. 815–826, aug 1989.

[110] S. L. Chuang, *Physics of Photonic Devices.* Oxford: Wiley, 2 ed., 2009.

[111] M. Asada, A. Kameyama, and Y. Suematsu, “Gain and intervalence band absorption in quantum-well lasers,” *IEEE Journal of Quantum Electronics*, vol. 20, pp. 745–753, jul 1984.

[112] I. Suemune and L. Coldren, "Band-mixing effects and excitonic optical properties in GaAs quantum wire structures-comparison with the quantum wells," *IEEE Journal of Quantum Electronics*, vol. 24, pp. 1778–1790, aug 1988.

[113] S. Jin, A.-T. Pham, Woosung Choi, Y. Nishizawa, Young-Tae Kim, Keun-Ho Lee, Youngkwan Park, and Eun Seung Jung, "Performance evaluation of InGaAs, Si, and Ge nFinFETs based on coupled 3D drift-diffusion/multisubband boltzmann transport equations solver," in *2014 IEEE International Electron Devices Meeting (IEDM)*, pp. 7.5.1–7.5.4, 2014.

[114] J. Z. Huang, Y. Wang, P. Long, Y. Tan, M. Povolotskyi, and G. Klimeck, "High-performance complementary III-V tunnel FETs with strain engineering," *arXiv 1605.00955*, pp. 1–6, 2016.

[115] C. Alper, L. Lattanzio, L. De Michielis, P. Palestri, L. Selmi, and A. M. Ionescu, "Quantum mechanical study of the germanium electron–hole bilayer tunnel FET," *IEEE Transactions on Electron Devices*, vol. 60, pp. 2754–2760, sep 2013.

[116] J. Zamastil, "Multidimensional WKB approximation for particle tunneling," *Physical Review A - Atomic, Molecular, and Optical Physics*, vol. 72, no. 2, pp. 1–4, 2005.

[117] K. Boucart and A. Ionescu, "Double gate tunnel FET with ultrathin silicon body and high-k gate dielectric," in *2006 European Solid-State Device Research Conference (ESSDERC)*, pp. 383–386, IEEE, sep 2006.

[118] G. Hellings, G. Eneman, R. Krom, B. De Jaeger, J. Mitard, A. De Keersgieter, T. Hoffmann, M. Meuris, and K. De Meyer, "Electrical TCAD simulations of a germanium pMOSFET technology," *IEEE Transactions on Electron Devices*, vol. 57, pp. 2539–2546, oct 2010.

[119] M. Fischetti, "Monte Carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures. I. Homogeneous transport," *IEEE Transactions on Electron Devices*, vol. 38, pp. 634–649, mar 1991.

[120] J. Morgan and E. Kane, "Observation of direct tunneling in germanium," *Physical Review Letters*, vol. 3, pp. 466–468, nov 1959.

[121] E. Gnani, S. Reggiani, A. Gnudi, and G. Baccarani, "Drain-conductance optimization in nanowire TFETs," in *2012 Proceedings of the European Solid-State Device Research Conference (ESSDERC)*, pp. 105–108, IEEE, sep 2012.

[122] L. Lattanzio, L. De Michielis, and A. M. Ionescu, "The electron–hole bilayer tunnel FET," *Solid-State Electronics*, vol. 74, pp. 85–90, aug 2012.

[123] J. T. Teherani, S. Agarwal, E. Yablonovitch, J. L. Hoyt, and D. Antoniadis, "Impact of quantization energy and gate leakage in bilayer tunneling transistors," *IEEE Electron Device Letters*, vol. 34, pp. 298–300, feb 2013.

[124] C. Alper, P. Palestri, J. L. Padilla, and A. M. Ionescu, "Underlap counterdoping as an efficient means to suppress lateral leakage in the electron–hole bilayer tunnel FET," *Semiconductor Science and Technology*, vol. 31, p. 045001, apr 2016.

[125] J. L. Padilla, C. Alper, F. Gámiz, and A. M. Ionescu, "Assessment of field-induced quantum confinement in heterogate germanium electron–hole bilayer tunnel field-effect transistor," *Applied Physics Letters*, vol. 105, p. 082108, aug 2014.

[126] K.-H. Goetz, D. Bimberg, H. J, J. J. Selders, A. V. Solomonov, G. F. Glinskii, and M. Razeghi, "Optical and crystallographic properties and impurity incorporation of $\mathrm{Ga_xIn_{1-x}As}$ (0.44<x<0.49) grown by liquid phase epitaxy, vapor phase epitaxy, and metal organic chemical vapor deposition," *Journal of Applied Physics*, vol. 54, no. 8, p. 4543, 1983.

[127] M. Levinshtein, S. Rumyantsev, and M. Shur, eds., *Handbook series on semiconductor parameters: volume 2: ternary and quaternary III-V compounds.* World Scientific Publishing Co. Pte. Ltd., 1999.

[128] C. Sire, S. Blonkowski, M. J. Gordon, and T. Baron, "Statistics of electrical breakdown field in HfO_2 and SiO_2 films from millimeter to nanometer length scales," *Applied Physics Letters*, vol. 91, no. 24, p. 242905, 2007.

[129] C. Alper, P. Palestri, L. Lattanzio, J. L. Padilla, and A. M. Ionescu, "Two dimensional quantum mechanical simulation of low dimensional tunneling devices," in *2014 44th European Solid State Device Research Conference (ESSDERC)*, pp. 186–189, IEEE, sep 2014.

[130] H. Schmid, M. Borg, K. Moselund, L. Gignac, C. M. Breslin, J. Bruley, D. Cutaia, and H. Riel, "Template-assisted selective epitaxy of III–V nanoscale devices for co-planar heterogeneous integration with Si," *Applied Physics Letters*, vol. 106, no. 23, p. 233101, 2015.

[131] C. Alper, L. De Michielis, N. Dagtekin, L. Lattanzio, and A. M. Ionescu, "Tunnel FET with non-uniform gate capacitance for improved device and circuit level performance," *2012 Proceedings of the European Solid-State Device Research Conference (ESSDERC)*, vol. 2, pp. 161–164, sep 2012.

[132] C. Alper, P. Palestri, J. L. Padilla, and A. M. Ionescu, "The Electron-Hole Bilayer TFET: Dimensionality Effects and Optimization," *IEEE Transactions on Electron Devices*, vol. 63, pp. 2603–2609, jun 2016.

[133] M. V. Fischetti and S. E. Laux, "Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys," *Journal of Applied Physics*, vol. 80, no. 4, p. 2234, 1996.

[134] S. Adachi, *Properties of semiconductor alloys: group-IV, III-V and II-VI semiconductors.* John Wiley & Sons, 2005.

[135] F. Stern, “Self-consistent results for n-type Si inversion layers,” *Physical Review B*, vol. 5, no. 12, pp. 4891–4899, 1972.

[136] M. Levinshtein, S. Rumyantsev, and M. Shur, *Handbook Series on Semiconductor Parameters: Si, Ge, C (Diamond), GaAs, GaP, GaSb, InAs, InP, InSb.* Handbook Series on Semiconductor Parameters, World Scientific Publishing Co. Pte. Ltd., 1996.

[137] E. Caruso, D. Lizzit, P. Osgnach, D. Esseni, P. Palestri, and L. Selmi, “Simulation analysis of III-V n-MOSFETs: Channel materials, Fermi level pinning and biaxial strain,” in *2014 IEEE International Electron Devices Meeting (IEDM)*, pp. 7.6.1–7.6.4, IEEE, dec 2014.

[138] A. W. Dey, B. M. Borg, B. Ganjipour, M. Ek, K. A. Dick, E. Lind, C. Thelander, and L.-E. Wernersson, “High-current GaSb/InAs(Sb) nanowire tunnel field-effect transistors,” *IEEE Electron Device Letters*, vol. 34, pp. 211–213, feb 2013.

[139] H. Lu and A. Seabaugh, “Tunnel field-effect transistors: state-of-the-art,” *IEEE Journal of the Electron Devices Society*, vol. 2, pp. 44–49, jul 2014.

[140] R. Yan, A. Ourmazd, and K. Lee, “Scaling the Si MOSFET: from bulk to SOI to bulk,” *IEEE Transactions on Electron Devices*, vol. 39, pp. 1704–1710, jul 1992.

[141] F. W. J. Olver, “DLMF: 9.9 Zeros.” http://dlmf.nist.gov/9.9#v, 2015. [Online].

[142] A. R. Trivedi, S. Carlo, and S. Mukhopadhyay, “Exploring tunnel-FET for ultra low power analog applications,” *Proceedings of the 50th Annual Design Automation Conference (DAC)*, no. V, pp. 109:1–109:6, 2013.

[143] U. E. Avci, D. H. Morris, and I. A. Young, “Tunnel field-effect transistors: prospects and challenges,” *IEEE Journal of the Electron Devices Society*, vol. 3, no. 3, pp. 88–95, 2015.

[144] J. L. Padilla, C. Alper, A. Godoy, F. Gamiz, and A. M. Ionescu, “Impact of asymmetric configurations on the heterogate germanium electron-hole bilayer tunnel FET including quantum confinement,” *IEEE Transactions on Electron Devices*, vol. 62, pp. 3560–3566, nov 2015.

[145] Z. Zhu, H. Zhu, M. Xu, J. Zhong, C. Zhao, D. Chen, and T. Ye, “A novel fin electron-hole bilayer tunnel field-effect transistor,” *IEEE Transactions on Nanotechnology*, vol. 13, pp. 1133–1137, nov 2014.

[146] T. Baldauf, A. Wei, R. Illgen, S. Flachowsky, T. Herrmann, T. Feudel, J. Hontschel, M. Horstmann, W. Klix, and R. Stenzel, “Simulation and optimization of tri-Gates in a 22 nm hybrid tri-gate/planar process,” in *ULIS 2011 Ultimate Integration on Silicon*, pp. 1–4, IEEE, mar 2011.

[147] M. Heyns, C. Huyghebaert, A. S. Verhulst, D. Leonelli, R. Rooyackers, and W. Dehaene, “Tunnel transistor, logical gate comprising the transistor, static random-access memory using the logical gate and method for making such a tunnel transistor,” Mar. 2013. US Patent 2013/0064005 A1.

[148] S. Sant and A. Schenk, “Methods to enhance the performance of InGaAs/InP heterojunction tunnel FETs,” *IEEE Transactions on Electron Devices*, pp. 1–7, 2015.

[149] L. Zhang, J. He, and M. Chan, “A compact model for double-gate tunneling field-effect-transistors and its implications on circuit behaviors,” in *2012 IEEE International Electron Devices Meeting (IEDM)*, pp. 6.8.1–6.8.4, IEEE, Dec. 2012.

[150] L. Esaki, “New phenomenon in narrow germanium p-n junctions,” *Physical Review*, vol. 109, pp. 603–604, Jan. 1958.

[151] A. Naeemi, A. Ceyhan, V. Kumar, C. Pan, R. M. Iraei, and S. Rakheja, “BEOL Scaling Limits and Next Generation Technology Prospects,” *Proceedings of the The 51st Annual Design Automation Conference (DAC) ’14*, pp. 1–6, 2014.

[152] “FD-SOI.” http://www.st.com/content/st_com/en/about/innovation---technology/FD-SOI.html, 2016. [Online; accessed 5 December 2016].

[153] D. Harris, R. Ho, G.-Y. Wei, and M. Horowitz, “The fanout-of-4 inverter delay metric,” *Unpublished Manuscript*, pp. 4–5, 2003.

[154] S. Strangio, P. Palestri, M. Lanuzza, F. Crupi, D. Esseni, and L. Selmi, “Assessment of InAs/AlGaSb tunnel-FET virtual technology platform for low-power digital circuits,” *IEEE Transactions on Electron Devices*, vol. 63, pp. 2749–2756, jul 2016.

[155] 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, pp. 4233–4239, nov 2016.

[156] M. T. Björk, H. Schmid, C. D. Bessire, K. E. Moselund, H. Ghoneim, S. Karg, E. Lörtscher, and H. Riel, “Si–InAs heterojunction Esaki tunnel diodes with high current densities,” *Applied Physics Letters*, vol. 97, p. 163501, oct 2010.

[157] “Multiprecision computing toolbox for MATLAB.” http://www.advanpix.com/, 2016. [Online; accessed 5 December 2016].

[158] D. H. Morris, U. E. Avci, R. Rios, and I. A. Young, “Design of low voltage tunneling-FET logic circuits considering asymmetric conduction characteristics,” *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 4, pp. 380–388, dec 2014.

[159] J. Robertson, “High dielectric constant oxides,” *The European Physical Journal Applied Physics*, vol. 28, pp. 265–291, dec 2004.

[160] C. Schulte-Braucks, D. Stange, N. von den Driesch, S. Blaeser, Z. Ikonic, J. M. Hartmann, S. Mantl, and D. Buca, “Negative differential resistance in direct bandgap GeSn p-i-n structures,” *Applied Physics Letters*, vol. 107, no. 4, p. 042101, 2015.

[161] M. Luisier and S. Odermatt, “Simulation of semiconductor lasers part 1 : many-body effects in semiconductor lasers part 2 : carrier transport in quantum structures,” Master’s thesis, Eidgenössische Technische Hochschule (ETH) Zürich, 2002.

[162] S. Datta, *Lessons from Nanoelectronics: A New Perspective on Transport.* World Scientific Publishing Company, 2012.

[163] O. Baumgartner, M. Karner, and H. Kosina, “Modeling of high-k-Metal-Gate-stacks using the non-equilibrium Green’s function formalism,” in *2008 International Conference on Simulation of Semiconductor Processes and Devices*, pp. 353–356, IEEE, sep 2008.

[164] A. Dalla Serra, A. Abramo, P. Palestri, L. Selmi, and F. Widdershoven, “Closed- and open-boundary models for gate-current calculation in n-MOSFETs,” *IEEE Transactions on Electron Devices*, vol. 48, no. 8, pp. 1811–1815, 2001.

[165] A. Campera, G. Iannaccone, and F. Crupi, “Modeling of tunnelling currents in Hf-based gate stacks as a function of temperature and extraction of material parameters,” *IEEE Transactions on Electron Devices*, vol. 54, pp. 83–89, jan 2007.

[166] L. F. Register, E. Rosenbaum, and K. Yang, “Analytic model for direct tunneling current in polycrystalline silicon-gate metal–oxide–semiconductor devices,” *Applied Physics Letters*, vol. 74, no. 3, p. 457, 1999.

[167] M. Luisier and A. Schenk, “Two-dimensional tunneling effects on the leakage current of MOSFETs with single dielectric and high-κ gate stacks,” *IEEE Transactions on Electron Devices*, vol. 55, pp. 1494–1501, jun 2008.

[168] L. J. Lauhon, M. S. Gudiksen, D. Wang, and C. M. Lieber, “Epitaxial core-shell and core-multishell nanowire heterostructures.,” *Nature*, vol. 420, no. 6911, pp. 57–61, 2002.

List of Publications

- **C. Alper**, J. L. Padilla, P. Palestri, A. M. Ionescu. "A novel reconfigurable sub-0.25V digital logic family using the electron-hole bilayer TFET," *IEEE Transactions on Electron Devices*, In preparation.

- D. Esseni, P. Palestri, M. Pala, **C. Alper**[1], T. Rollo. "A review of selected topics in physics based modelling of small slope transistors based on band-to-band-tunnelling," *Semiconductor Science and Technology*, In preparation.

- W. A. Vitale, E. A. Casu, A. Biswas, **C. Alper**, A. Krammer, G. V. Luong, S. Mantl, A. Schuler, and A. M. Ionescu, “PC-TFET: a steep-slope transistor combining phase-change and band-to-band-tunneling to achieve a subthreshold swing less than 10mV/decade and sub-unity body factor,” *Scientific Reports*, under review, 2016.

- E.A. Casu, W.A. Vitale, N. Oliva, T. Rosca, A. Biswas, **C. Alper**, A. Krammer, G. V. Luong, Q. T. Zhao, S. Mantl, A. Schüler, A. Seabaugh, A. M. Ionescu. "Hybrid phase-change – tunnel fet (PC-TFET) switch with subthreshold swing < 10 mv/decade and sub-0.1 body factor: digital and analog benchmarking, " *2016 International Electron Devices Meeting (IEDM)*, Dec. 2-6, 2016.

- **C. Alper**, J. L. Padilla, P. Palestri, A. M. Ionescu. "Impact of device geometry of the fin Electron-Hole Bilayer Tunnel FET," *ESSDERC 2016 - 46th European Solid State Device Research Conference*, Sept. 12-15, 2016.

- **C. Alper**, P. Palestri, J. L. Padilla, A. M. Ionescu. "The Electron-Hole Bilayer TFET: Dimensionality Effects and Optimization," *IEEE Transactions on Electron Devices*, 63(6), pp. 2603-2609, 2016.

- J. L. Padilla, **C. Alper**, F. Gamiz, A. M. Ionescu. "Switching Behavior Constraint in the Heterogate Electron–Hole Bilayer Tunnel FET: The Combined Interplay Between Quantum Confinement Effects and Asymmetric Configurations," *IEEE Transactions on Electron Devices*, 63(6), pp. 2570-2576, 2016.

---

[1] Final author ordering TBD.

- **C. Alper**, P. Palestri, J. L. Padilla, A. M. Ionescu. "Underlap counterdoping as an efficient means to suppress lateral leakage in the electron–hole bilayer tunnel FET," *Semiconductor Science and Technology*, 31(4), pp.045001, 2016.
- J. L. Padilla, **C. Alper**, F. Gamiz, A. M. Ionescu. "Band-to-band tunneling distance analysis in the heterogate electron–hole bilayer tunnel field-effect transistor," *Journal of Applied Physics*, 119(4), pp. 045705, 2016.
- J. L. Padilla, **C. Alper**, F. Gamiz, A. M. Ionescu "Assessment of confinement-induced band-to-band tunneling leakage in the Fin-EHBTFET," *EUROSOI-ULIS 2016*, Jan. 25-27, 2016.
- J. L. Padilla, **C. Alper**, A. Godoy, F. Gamiz, A. M. Ionescu. "Impact of asymmetric configurations on the heterogate germanium electron–hole bilayer tunnel FET including quantum confinement," *IEEE Transactions on Electron Devices*, 62(11), pp. 3560-3566, 2015.
- **C. Alper**, M. Visciarelli, P. Palestri, J. L. Padilla, A. Gnudi, E. Gnani, A. M. Ionescu. "Modeling the imaginary branch in III–V tunneling devices: Effective mass vs $k \cdot p$," *SISPAD 2015*, Sept. 9-11, 2015.
- J. L. Padilla, **C. Alper**, C. Medina-Bailon, F. Gamiz, A. M. Ionescu. "Assessment of pseudo-bilayer structures in the heterogate germanium electron-hole bilayer tunnel field-effect transistor," *Applied Physics Letters*, 106 (26), pp. 262102, 2014.
- **C. Alper**, P. Palestri, J.L. Padilla, A. Gnudi, R.Grassi, E. Gnani, M. Luisier and A. M. Ionescu. "Efficient quantum mechanical simulation of band-to-band tunneling," *EUROSOI-ULIS 2015*, Jan. 26-28, 2015.
- J. L. Padilla, **C. Alper**, F. Gamiz and A. M. Ionescu. "Assessment of field-induced quantum confinement in heterogate germanium electron-hole bilayer tunnel field-effect transistor," *Applied Physics Letters*, 105, 8, pp. 082108, 2014.
- A. M. Ionescu, **C. Alper**, J. L. Padilla, L. Lattanzio, P. Palestri. "Electron-hole bilayer deep subthermal electronic switch: Physics, promise and challenges," *2014 SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S)*, Oct. 6-9, 2014.
- A. Biswas, L. De Michielis, **C. Alper**, A. M. Ionescu. "Conformal mapping based DC current model for double gate tunnel FETs," *ULIS 2014*, Apr. 7-9, 2014.
- A. Biswas, N. Dagtekin, **C. Alper**, L.De Michielis, A. Bazigos, W. Grabinski, A. M. Ionescu. "Compact modeling of homojunction tunnel FETs," *MIXDES 2014*, Jun. 19-21, 2014.
- **C. Alper**, P. Palestri, L. Lattanzio, J. L. Padilla and A. M. Ionescu. "Two dimensional quantum mechanical simulation of low dimensional tunneling devices," *ESSDERC 2014 - 44nd European Solid State Device Research Conference*, Sept. 22-26, 2014.

  - Extended version: **C. Alper**, P. Palestri, L. Lattanzio, J. L. Padilla and A. M. Ionescu. *Solid State Electronics*, 113, pp. 167-172, 2015.

- **C. Alper**, L. De Michielis, L. Lattanzio, P. Palestri, L. Selmi and A. M. Ionescu. "Quantum mechanical simulation of the germanium electron-hole bilayer TFET," *IEEE Transactions on Electron Devices*, 60(9), pp. 2754-2760, 2013.
- **C. Alper**, L. De Michielis, N. Dagtekin, L. Lattanzio and A. M. Ionescu. "Tunnel FET with non-uniform gate capacitance for improved device and circuit level performance," *ESSDERC 2012 - 42nd European Solid State Device Research Conference*, Sept. 17-21, 2012.
  - Extended version: **C. Alper**, L. De Michielis, N. Dagtekin, L. Lattanzio, D. Bouvet, A.M. Ionescu. *Solid State Electronics*, 84, pp. 205-210, 2013.
- A. Biswas, **C. Alper**, L. De Michielis and A. M. Ionescu. "New Tunnel-FET Architecture with Enhanced Ion and Improved Miller Effect for Energy Efficient Switching," *70th Device Research Conference*, June 18-20, 2012.

Patent Applications:

- **C. Alper**, L. Lattanzio, M. A. Ionescu, L. De Michielis, N. Dagtekin. "Semiconductor Tunneling Device," US Patent 2016/0043234 A1.

Cem ALPER

Avenue de l'Eglise-Anglaise 12, 1006 Lausanne SWITZERLAND
+41 76 405 65 86| cm.alper@gmail.com

STRENGTHS

- Expert in semiconductor device modeling
- Quantum mechanical TCAD code development
- Versatile programming background: from low level embedded programming to high performance scientific computing
- Proven team player in an international setting with excellent time management skills

EDUCATION

École Polytechnique Fédérale de Lausanne (EPFL) Lausanne, Switzerland
PhD in Microsystems and Microelectronics 2012 – March 2017 (Expected)

École Polytechnique Fédérale de Lausanne (EPFL) Lausanne, Switzerland
M.Sc in Electrical Engineering 2010 – 2012

Sabanci University Istanbul, Turkey
B.Sc in Electrical Engineering 2006 – 2010
Physics Minor

EXPERIENCE

NANOLAB, EPFL Lausanne, Switzerland
PhD Thesis Title: "Modelling and Optimization of Tunnel-FET architectures exploiting carrier gas dimensionality" 2012 – Present

- Development of 1-D and 2-D quantum mechanical simulators for Tunnel FET simulation in MATLAB. Accurate predication, optimization and design of the energy efficient switch Electron Hole Bilayer TFET for the first time. Implementation of external TCAD models through Physical Model Interface. Mixed-mode quantum-corrected circuit simulations using Sdevice.
- Measurement and statistical characterization of TFETs and MOSFETs fabricated by project partners in room temperature and cryogenic environments.
- Close collaboration with a group of multi-national scientists from academy and industry within the European projects STEEPER and E2SWITCH. Responsibilities included co-preparation of deliverables and milestones as well as presentation of the results to external referees from the European Commission.
- 11 journal, 10 conference articles (see complete list). 1 patent application.

NANOLAB, EPFL Lausanne, Switzerland
Master Thesis Title: "TFET Optimization for Circuit Performance" Sept. 2011 - May 2012

- Mathematical modeling of Tunnel FETs.

- Development of one of the first circuit simulation methodologies using Tunnel FETs making use of Sentaurus TCAD. Benchmarking of sample logic circuits using TFETs and CMOS transistors.

Sabanci University Istanbul, Turkey

Teaching Assistant Spring 2009 & Spring 2010

- Teaching assistant for Electronic Circuits course. Responsibilities included supervising a class of 60 students in various lab sessions as well as grading the homework and lab reports.
- Supervising the lab sessions of approximately 80 students for Microprocessor System Design course. Organization of lab sessions, handling of lab equipment.

Vestek R&D Corp. Istanbul, Turkey

RTL Design Intern Summer 2009

- Worked as an RTL engineer intern in Vestek R&D Corp., a subsidiary of Vestel, a leading Turkish electronics company. Designed the hardware which implements a Chrominance Transient Improvement (CTI) algorithm using VHDL, to be used in the video processing engine Pixellence.

SKILLS

Programming & Hardware Description Languages:

- C/C++, Java (basic), MATLAB, Verilog, VHDL, VHDL-AMS, Python, R, Verilog-A, Tcl

Development Environments:

- Sentaurus TCAD, Microsoft Visual Studio, Eclipse, MATLAB, Octave, Orcad, Cadence, Xilinx ISE, Xilinx Vivado, ModelSim, Altera Quartus II, COMSOL Multiphysics

Measurement & Characterization:

- IC-CAP, Labview

Languages:

- English (Advanced, TOEFL IBT 112/120), French (Intermediate), German (Basic), Turkish (Native)

ACHIEVEMENTS & AWARDS

- Ranked 3rd in *Prix Space Challenge* within Entrepreneurship Day organized by EPFL Innovation Park and Swiss Space Center
- Ranked in ~0.5th percentile in Nationwide University Entrance Exam in 2006
- Certificate of High Honor for 8 Semesters
- Sabanci University Honor Scholarship (Tuition Fee Exemption upon admission).

PERSONAL DETAILS

Nationality: Turkish Swiss Permit: B since 2010

Age: 27 Civil Status: Single

Interests: I like running as a form of meditation, I enjoy playing the piano that I learnt during my undergraduate studies.