### Title

Analytic modeling of tunnel Field-Effect-Transistors and experimental investigation of GaN High-Electron-Mobility-Transistors




### Author

Wu, Jianzhi



### Publication Date

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

# Analytic modeling of tunnel Field-Effect-Transistors and experimental investigation of GaN High-Electron-Mobility-Transistors

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Electrical Engineering (Applied Physics)

by

Jianzhi Wu

Committee in charge:

Professor Paul K. L. Yu, Co-Chair
Professor Yuan Taur, Co-Chair
Professor Peter M. Asbeck
Professor Prabhakar Bandaru
Professor Yi Chen

2016

The dissertation of Jianzhi Wu is approved, and it is acceptable in quality and form for publication on microfilm:

_____

_____

_____

_____
Co-Chair

_____
Co-Chair

University of California, San Diego

2016

# DEDICATION

This dissertation is dedicated to my parents for their great support and unbounded love.

# TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

# ACKNOWLEDGEMENTS

I have a great fortune to be surrounded by so many nice people who helped me throughout my graduate academic career. Here I would like to acknowledge all of them for their enormous support. First and foremost, I would like to express my deepest gratitude to my advisor, Professor Paul K. L. Yu, who provided tremendous support and enlightening guidance throughout my research and life. He trained me to operate MOCVD, fabricate microelectronic devices, and most importantly how to think scientifically. Besides his academic supervision, he is also a great mentor of personality who taught me to become mature and professional. Without his guidance and encouragement throughout many hard times, this dissertation would not have been possible.

I am deeply indebted to my co-advisor, Professor Yuan Taur, who has significant influence on me with his in-depth knowledge, his sincere attitude and enthusiasm towards scientific research. Prof. Taur is the top notch in the field of semiconductor device, and more importantly he is a great mentor. In addition to his patient guidance throughout my research, he also imparts me how to think deeper. This is more valuable for the knowledge itself, and it benefits me for my whole life.

I would also like to express my appreciation to my co-advisor, Prof. Peter M. Asbeck, for his incisive advice and guidance in my graduate study. I am greatly impressed by his solid and broad knowledge as well as his great patience during his discussions with students.

Tons of thanks should be given to my committee members, Prof. Prabhakar Bandaru, and Prof. Yi Chen for taking their time to serve as committee members, to review my dissertation and give me valuable comments.

I am most thankful for Dr. Steve Yao and Mary Fang for their enormous encouragement and kind support for my career as well as my life.

At UCSD, I was fortunate to grow intellectually and professionally with so many talented people. Their inspirations and insightful ideas helped me tremendously during my research. My deep thankfulness goes to all my colleagues at USCD, namely, Dr. Wei Lu, Mr. Art Clawson, Dr. Dingbo Chen, Jie Min, Siyuan Gu, Dr. Bo Yu, Dr. Lingquan (Dennis) Wang, Dr. CK Sun, Dr. Steve Pappart, Dr. Han-Ping Chen, Hui Wang, Conway Wang, Po-Han (Peter) Wang, Prof. Wen Qiao, Dr. Yanjin Kuang, Prof. Qing Gu, Dr. Chenhui Niu, Dr. David Keogh, Dr. Winnie Chen, Dr. Dylan Lu, Dr. Zhelin Sun, Dr. Zack Chen, Dr. Feifei Wei, Dr. Yuzhe Xiao, Dr. Yi Jing, Dr. Kangmu Lee, Dr. Yuchun Zhou, Alex Zhang, Amanda Kerr, Guobi Zhao, Huanhui Zhan.

My life in San Diego would not have been such beautiful and fulfilling without many of my friends, especially those who played basketball with me. Thanks to Dr. Yaoguang Ma, Dr. Yu Wang, Prof. Shenggao Zhou, Jinxing Li, Sha He, Yuanyuan Han, Qian Ma, Rui La, Tiantian Zhang, Dr. Zhehua Li, Danyi Liu, Alice Yu, and Jamie Huang. Those friendship has meant a lot to me.

I also thank to the support from the staffs of ECE department, Graduate Division and international center, especially to Ms. Kacy Vega, Ms. Van Lee-Yamamoto and Ms. Gabriela Hoffmann.

Finally, and most importantly, I would like to thank my parents for their endless love and support. They sacrificed so much to help me get my education. Also my aunt Ms. Gao and uncle Dr. Zhu have been greatly supporting me getting through numerous life difficulties. This dissertation is dedicated to them.

The text of Chapter 2, in part, is a reprint of the material as it appears in "An analytic model for heterojunction tunnel FETs with exponential barrier" by Yuan Taur, Jianzhi Wu and Jie Min, IEEE Transaction on Electron Devices, May 2015. The dissertation author was a co-author of this paper.

The text of Chapter 2, in part, is a reprint of the material as it appears in "Short channel effects in tunnel FETs" by Jianzhi Wu, Jie Min and Yuan Taur, IEEE Transaction on Electron Devices, Sept. 2015. The dissertation author was the primary investigator and author of this paper.

The text of Chapter 3, in part, is a reprint of the material as it appears in "Dimensionality dependence of TFET performance down to 0.1 V supply voltage" by Yuan Taur, Jianzhi Wu and Jie Min, IEEE Transaction on Electron Devices, Feb. 2016. The dissertation author was a co-author of this paper.

The text of Chapter 3, in part, is a reprint of the material as it appears in "Analysis of source doping effect in tunnel FETs with staggered bandgap" by Jie Min, Jianzhi Wu and Yuan Taur, IEEE Electron Device Letters, Oct. 2015. The dissertation author was a co-author of this paper.

The text of Chapter 3, in part, is a reprint of the material as it appears in "An analytic model for heterojunction tunnel FETs with exponential barrier" by Yuan Taur, Jianzhi Wu and Jie Min, IEEE Transaction on Electron Devices, May 2015. The dissertation author was a co-author of this paper.

The text of Chapter 4, in part, is a reprint of the material as it appears in "Short channel effects in tunnel FETs" by Jianzhi Wu, Jie Min and Yuan Taur, *IEEE Transaction on Electron Devices,* Sept. 2015. The dissertation author was the primary investigator and

author of this paper.

The text of Chapter 5, in part, is a reprint of the material as it appears in "Examination of two-band $E(k)$ relations for band-to-band tunneling" by Yuan Taur and Jianzhi Wu, IEEE Transaction on Electron Devices, Feb. 2016. The dissertation author was a co-author of this paper.

The text of Chapter 6, in part, is a reprint of the material as it appears in "A continuous, semi-analytic current model for DG and NW TFETs" by Jianzhi Wu and Yuan Taur, IEEE Transaction on Electron Devices, Feb. 2016. The dissertation author was the primary investigator and author of this paper.

The text of Chapter 6, in part, is a reprint of the material as it appears in "Reduction of TFET off current and subthreshold swing by lightly doped drain" by Jianzhi Wu and Yuan Taur, IEEE Transaction on Electron Devices, August 2016. The dissertation author was the primary investigator and author of this paper.

The text of Chapter 7, in part, is a reprint of the material as it appears in "A continuous, semi-analytic current model for DG and NW TFETs" by Jianzhi Wu and Yuan Taur, IEEE Transaction on Electron Devices, Feb. 2016. The dissertation author was the primary investigator and author of this paper.

The text of Chapter 8, in part, is a reprint of the material as it appears in "Thermal Resistance Extraction of AlGaN/GaN Depletion-Mode HEMTs on Diamond" by Jianzhi Wu, Jie Min, Wei Lu and Paul. K. L. Yu, *Journal of Electronic Materials*, vol. 44, no. 5, pp. 1275–1280, May 2015. The dissertation author was the primary investigator and author of this paper.

The text of Chapter 9, in part, is a reprint of the material as it appears in "Normally-

OFF AlGaN/GaN MOS-HEMT with a Two−Step Gate Recess" by Jianzhi Wu, Wei Lu and Paul. K. L. Yu, *IEEE International Conference on Electron Devices and Solid-State Circuits* (*EDSSC*) (pp. 594-596), Singapore, 2015. The dissertation author was the primary investigator and author of this paper.

# VITA

EDUCATION

2016 Doctor of Philosophy, University of California, San Diego, La Jolla, CA, U.S.A.

2012 Master of Science, University of Pennsylvania, Philadelphia, PA, U.S.A.

2010 Bachelor of Science, Southeast University, Nanjing, China.

PUBLICATIONS

[1] **Jianzhi Wu** and Yuan Taur, "An all-region I-V model for 1D short-channel MOSFETs", submitted 2016.

[2] **Jianzhi Wu** and Yuan Taur, "Reduction of TFET off current and subthreshold swing by lightly doped drain", *IEEE Transactions on Electron Devices*, vol. 63, no. 8, pp. 3342-3345, Aug. 2016.

[3] Yuan Taur, **Jianzhi Wu** and Jie Min, "A short channel I-V model for 2D MOSFETs", *IEEE Transactions on Electron Devices*, vol. 63, no. 6, pp. 2550-2555, June 2016.

[4] **Jianzhi Wu** and Yuan Taur, "A continuous, semianalytic current model for DG and NW TFETs", *IEEE Transactions on Electron Devices*, vol. 63, no. 2, pp. 841-847, Feb. 2016.

[5] Jie Min, Lingquan Wang, **Jianzhi Wu** and Peter Asbeck, "Analysis of Temperature Dependent Effects on I-V Characteristics of Heterostructure Tunnel Field Effect Transistors", *IEEE Journal of the Electron Devices Society*, vol. 4, no. 6, pp. 416-423, Nov. 2016.

[6] Yuan Taur, **Jianzhi Wu** and Jie Min, "Dimensionality dependence of TFET performance down to 0.1 V supply voltage", *IEEE Transactions on Electron Devices*, vol. 63, no. 2, pp. 877-880, Feb. 2016.

[7] Yuan Taur and **Jianzhi Wu**, "Examination of two-band $E(k)$ relations for band-to-band tunneling", *IEEE Transactions on Electron Devices*, vol. 63, no. 2, pp. 869-872, Feb. 2016.

[8] **Jianzhi Wu**, Jie Min, and Yuan Taur, "Short channel effects in tunnel FETs", *IEEE Transactions on Electron Devices*, vol. 62, no. 9, pp. 3019-3024, Sept. 2015.

[9] **Jianzhi Wu**, Jie Min, Jingwei Ji and Yuan Taur, "An analytic model for heterojunction and homojunction tunnel FETs with 3D density of states", in *Proc. of 73rd Device Research Conference* (*DRC*) (pp. 249-250), Ohio, USA, June, 2015.

[10] **Jianzhi Wu**, Wei Lu and Paul. K. L. Yu, "Normally-OFF AlGaN/GaN MOS-HEMT with a Two-Step Gate Recess", *IEEE International Conference on Electron Devices and Solid-State Circuits* (*EDSSC*) (pp. 594-596), Singapore, 2015.

[11] Yuan Taur, **Jianzhi Wu** and Jie Min, "An analytic model for heterojunction tunnel FETs with exponential barrier" *IEEE Transactions on Electron Devices*, vol. 62, no. 5, pp. 1399-1404, May 2015.

[12] Jie Min, **Jianzhi Wu** and Yuan Taur, "Analysis of source doping effect in tunnel FETs with staggered bandgap", *IEEE Electron Device Letters*, vol. 36, no. 10, pp. 1094-1096, Oct. 2015.

[13] **Jianzhi Wu**, Jie Min, Wei Lu and Paul. K. L. Yu, "Thermal Resistance Extraction of AlGaN/GaN Depletion-Mode HEMTs on Diamond", *Journal of Electronic Materials*, vol. 44, no. 5, pp. 1275–1280, May 2015.

[14] **Jianzhi Wu**, Wei Lu, Jie Min and Paul. K. L. Yu, "AlGaN/GaN Depletion-Mode HEMT on Diamond", Oral presentation in *56$^{th}$ Electronic Materials Conference* (*EMC*), Santa Barbara, CA, USA, 25-27 June, 2014

[15] A. J. Kerr, E. Chagarov, S. Gu, T. Kaufman-Osborn, S. Madisetti, **J. Wu**, P. M. Asbeck, S. Oktyabrsky, and A. C. Kummel. "Preparation of gallium nitride surfaces for atomic layer deposition of aluminum oxide." *The Journal of Chemical Physics* 141, no. 10 (2014).

FIELDS OF STUDY

Major Field: Electrical Engineering

Focused Field: Applied Physics

# ABSTRACT OF THE DISSERTATION

Analytic modeling of tunnel Field-Effect-Transistors and experimental investigation of GaN High-Electron-Mobility-Transistors

by

Jianzhi Wu

Doctor of Philosophy in Electrical Engineering (Applied Physics)

University of California, San Diego, 2016

Professor Paul K. L. Yu, Co-Chair

Professor Yuan Taur, Co-Chair

High density and lower power drive the aggressive scaling down of CMOS transistors. Yet, the scaling of Si bulk MOSFETs are approaching physical limits, suffering from poor electrostatic control due to short channel effects, gate leakage current caused by gate oxide tunneling, and most importantly the non-scaled supply voltage imposed by thermionic emission limitation. Tunnel FETs (TFETs) based on band-to-band tunneling

current injection mechanism, have emerged as promising candidates to deliver steep turn-off slopes, thus enables a sharp reduction of supply voltage to below 0.5 V.

This dissertation is primarily devoted to develop an accurate analytic model for TFETs with a double-gate structure, providing physical insights to the design principles. At the core of the model is a gate-controlled channel potential that satisfies the source and drain boundary conditions. The potential is of an exponential profile with a characteristic scale length given by the device thickness. Both the source-to-channel tunneling and source-to-drain tunneling are developed and included in the model. It has been verified by numerical simulations for a wide range of bandgaps and channel lengths. Also incorporated in the model are the short-channel effect, source doping effect, ambipolar effect, and de-bias of gate voltage by channel charge. Based on these, the guidelines for scaling TFETs to sub-10-nm channel lengths are brought forth. The model is continuous, physical and predictive in the sense that there is no need for ad hoc fitting parameters.

For high-power and high-frequency applications, GaN high-electron-mobility-transistors (HEMTs) stand out as promising candidate devices for achieving high breakdown voltage, high output current and high transconductance characteristics. Yet, the performance of GaN HEMTs suffers from mobility degradation due to poor thermal dissipation of conventional epitaxial substrates. This dissertation also experimentally demonstrates the GaN HEMTs fabricated on diamond substrate with extraordinary thermal management capability. The self-heating induced current droop is effectively absent in the saturated $I_{ds}$-$V_{ds}$ characteristics of the resulting devices, thus paving the way for enhancing the energy conversion efficiency.

# Chapter 1 Introduction

## 1.1 Review of CMOS scaling

Over the past few decades, the electronic industry has made great strides to keep pace with the Moore's law, and doubles the number of transistors on a chip every two years [1.1]. In Fig. 1.1, the technology trends for memory product function/chips were summarized by International Technology Roadmap for Semiconductors (ITRS) in 2011 [1.2]. The Fig. 1.1 predicts that in years after 2016, the scaling rate of on-chip transistors for memory products [DRAM ~2×/2.5 yrs] tends to fall below the pace of the average "Moore's Law."

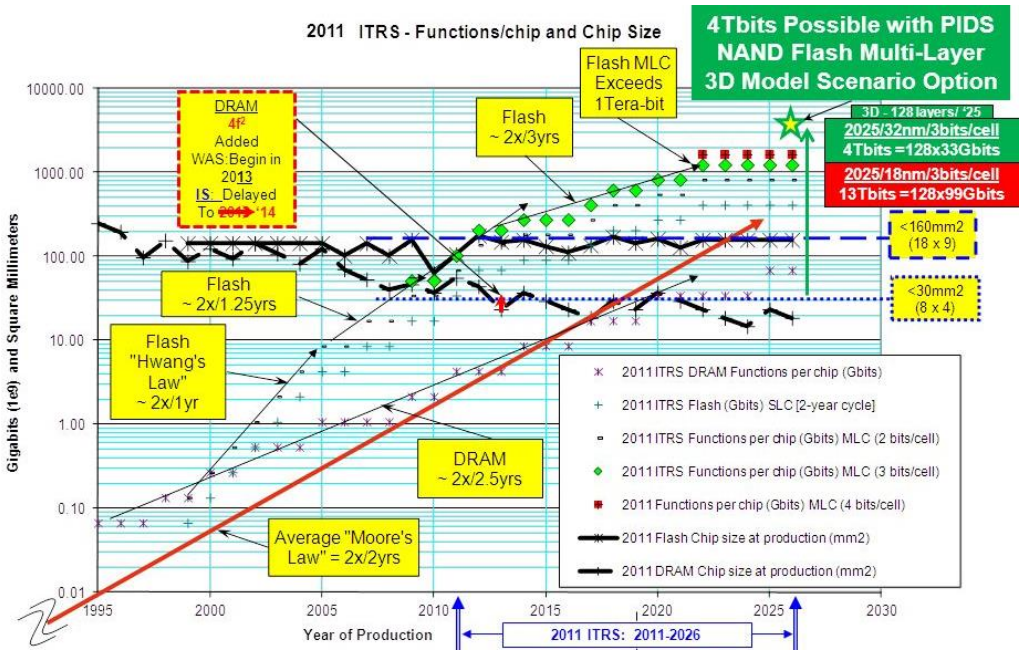



Figure 1.1: 2011 ITRS product technology trends: Memory product functions/chip and industry average "Moore's Law" and chip size trends [1.2].

An inherent problem resulted from the CMOS scaling is the power consumption in modern microelectronic circuits. Fig. 1.2 shows when the gate length is shrunk down to a

few tenth nanometer scale, the passive power density dominates over the active power density. Unfortunately, the passive power density (standby leakage power) is unusable, which eventually heats up the on-chip transistors, thus degrading the performance of the chip.

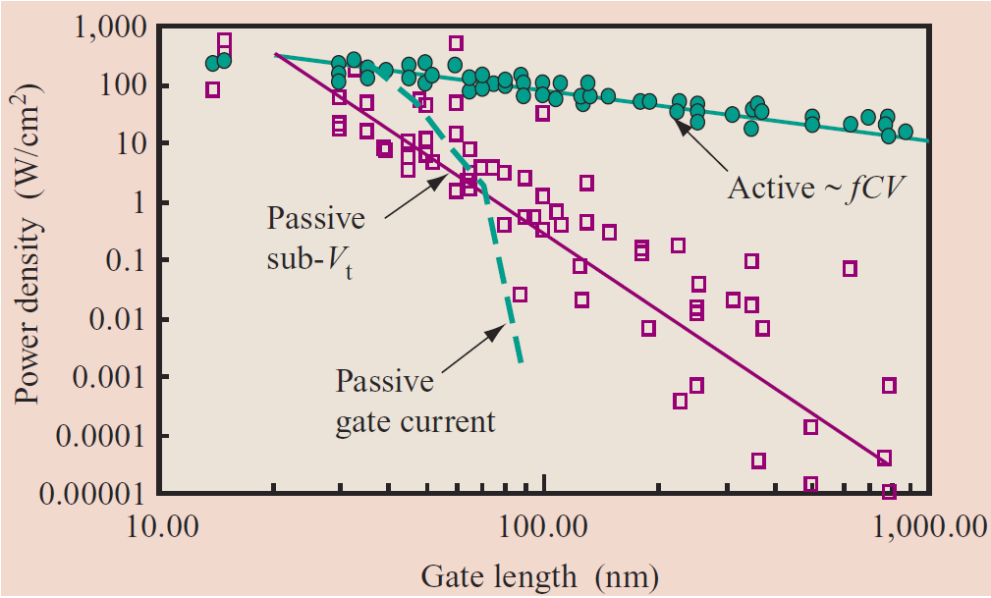



Figure 1.2: Both active and passive power density vs. gate length [1.3]

The leakage power is proportional to the supply voltage. Therefore the innovation of novel designs of transistors are in urgent demand, to reduce the supply voltage thus the leakage power.

## 1.2 Voltage non-scaling

MOSFET is the basic integrated circuit element, its off-current ($I_{off}$) is defined as the source-to-drain leakage current when the gate-to-source voltage ($V_{gs}$) is biased in the off-state, while the drain-to-source voltage ($V_{ds}$) is biased at supply voltage $V_{dd}$. When the $V_{gs}$ is biased at subthreshold voltage, the $I_{off}$ is expressed as [1.4]:

$$I_{off} = I_{ds(Vt)}e^{-qVt/mkT} \quad (1.1)$$

where $V_t$ is the threshold voltage, and $I_{ds(Vt)}$ is the source-to-drain current at threshold. Therefore, with the acceptable $I_{off}$ defined as three-order of magnitude less than the threshold current $I_{ds(Vt),}$ it requires the minimum threshold voltage $V_t$ to be ~ 6.9 $mkT/q$ to turn off the device.

Reduction of the standby leakage power, $P_{leak} = I_{off} \bullet V_{dd}$, requires the down scaling of supply voltage $V_{dd}$. However, the lower bound of $V_{dd}$ is imposed by the on-current and $CV/I$ propagation delay. The propagation delay is vital since it directly affects the speed at which a digital device can operate, consequently the speed of memory chips and microprocessors. The Fig. 1.3 shows an example of the CMOS performance, expressed in terms of inverse of propagation delay, versus the normalized threshold voltage $V_t/V_{dd}$. It exhibits the scenario when the $V_t/V_{dd}$ is increased from 0.2 to 0.3, the $CV/I$ delay will then lose (0.68-0.5)/0.68 or ~ 26% performance. Hence, it is desirable to maintain $V_t/V_{dd} < 0.25$ for high performance CMOS circuits, due to the delay sensitivity.

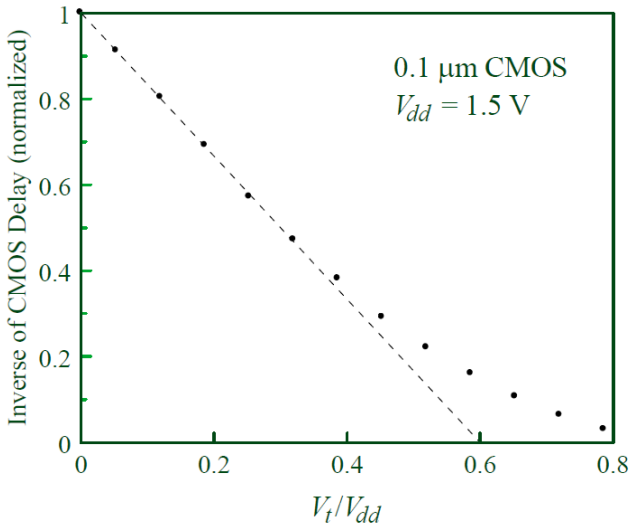



Figure 1.3: The inverse of CMOS delay vs. $V_t/V_{dd}$ from [1.5]. The dots are SPICE simulations results. The dashed line is a fitting proportional to 0.6-$V_t/V_{dd.}$

## 1.3 Beyond CMOS technologies

Towards the end of the traditional CMOS scaling, steep subthreshold swing transistors emerge as promising candidates to enable aggressive scaling of supply voltage, thus extending Moore's Law in the roadmap. The subthreshold swing determines the amount of supply voltage required to gain a ten-fold change (a decade) in drain current. In the conventional MOSFETs, the subthreshold swing is limited by the temperature dependent occupancy probability to 60 mV/dec at room temperature. Various transistor structures have been demonstrated to deliver steeper than 60 mV/dec turn-off slope, based on different switching mechanisms. Investigations of tunnel FETs [1.4], Negative capacitance FETs [1.5], impact-ionization FETs [1.6] and Nano-electro-mechanical FETs [1.7] have been carried out. Among them, tunnel FET (TFET) stands out as the most promising candidate, while other candidates suffer from the delayed positive feedback

problems [1.4]. Furthermore, tunnel FETs are attractive due to its CMOS-compatibility and mass manufacturability. Hence the efforts of this dissertation focus on the analytic modeling of tunnel FETs.

## 1.4 Organization of this work

The outline of this dissertation is as follows. In Chapter 2, the analytic potential solutions of double gate and nanowire TFET are derived and validated by numerical simulations. Chapter 3 discusses the modeling of ON-current of TFETs. Chapter 4 discusses the modeling of OFF-current. Chapter 5 discusses the examinations of two-band $E(k)$ relations for valence-band-to-conduction-band tunneling. Chapter 6 discusses the TFET scaling guidelines. Chapter 7 discusses the hardware and model correlations. Chapter 8 demonstrates the GaN-on-diamond high electron mobility transistor technology. Chapter 9 investigates Normally-OFF AlGaN/GaN MOS-HEMT technology.

## 1.5 References

[1.1] G. Moore, “Progress in Digital Integrated Electronics,” IEDM Tech. Dig., 11-13, 1975.

[1.2] ITRS Overview from Korea Winter Public Conference. [On-line] http://www.ewh.ieee.org/r6/scv/eds/announcements/2012-02-15-EDS-SCV-DL-Allan.pdf

[1.3] W. Haensch, E. J. Nowak, R. H. Dennard, P. M. Solomon, A. Bryant, O. H. Dokumaci, A. Kumar, X. Wang, J. B. Johnson, and M. V. Fischetti. "Silicon CMOS devices beyond scaling." IBM Journal of Research and Development 50.4.5 (2006): 339-361.

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

[1.5] S. Salahuddin and S. Datta, “Use of negative capacitance to provide voltage amplification for low power nanoscale devices,” *Nano Lett*., vol. 8, no. 2, pp. 405–410, 2008.

[1.6] K. Gopalakrishnan, P. B. Griffin, and J. D. Plummer, “I-MOS: A novel semiconductor device with a subthreshold slope lower than kT/q,” in *Proc. Int. Electron Devices Meeting*, 2002, pp. 289–292.

[1.7] H. Kam, D. T. Lee, R. T. Howe, and T.-J. King, “A new nano-electro-mechanical field effect transistor (NEMFET) design for low-power electronics,” *in Proc. Int. Electron Devices Meeting*, 2005, pp. 463–4

# Chapter 2 Potential profile in a Double-Gate (DG) or Nanowire (NW) TFET

## 2.1 Early TFET models

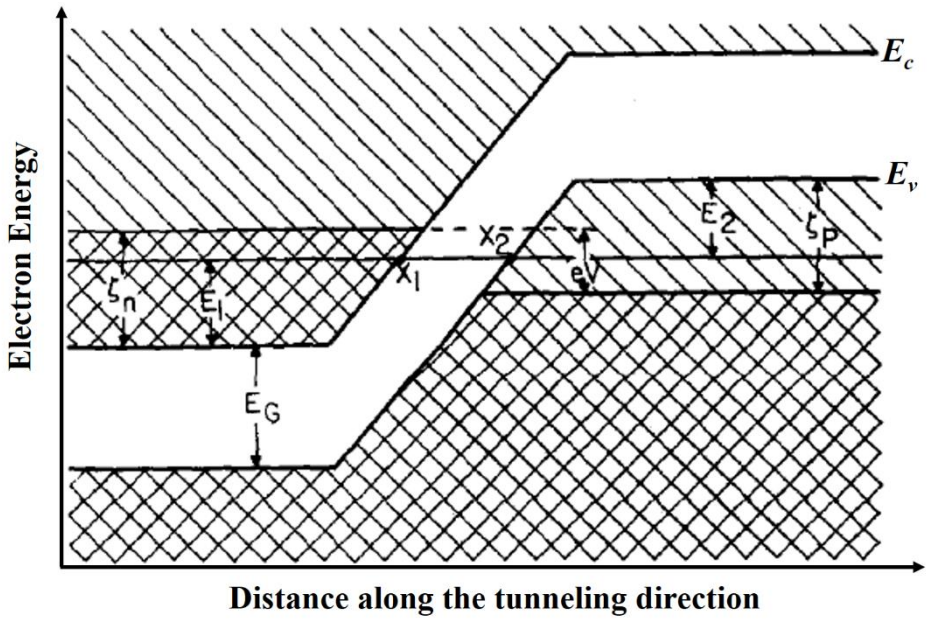



Figure 2.1 Constant field model of a *p*-*n* junction from [2.1].

The first tunnel FET (TFET) model refers to E. O. Kane's work in 1960s, which has been widely cited in a variety of literatures and textbooks [2.1]. Fig. 2.1 shows that Kane's model is based on the assumption of constant electrical field for a *p*-*n* junction model, which yields the tunneling probability $T$ for a carrier tunneling from the source valence band ($E_v$) to the channel conduction band ($E_c$), spatially across the bandgap at the same energy:

$$T = \frac{\pi^2}{9}\exp\left\{\frac{-\pi m^{*1/2} E_G^{\,3/2}}{2\sqrt{2}\hbar F}\right\}\exp\left\{-2E_\perp / \overline{E}_\perp\right\} \tag{2.1}$$

where the parameter $\overline{E}_{\perp} = \sqrt{2}\hbar F / \pi m^{*1/2} E_G^{1/2}$. $E_G$ is the bandgap of the junction material, $F$ is the electrical field, and $E_{\perp}$ is the energy associated with the momentum perpendicular to the tunneling direction.

In fact, neither constant electric field nor *p-n* junction model applies to the tunnel FET which is a three terminal device. Furthermore, Eq. (2.1) is limited to the homojunction case only, in which the tunneling probability is independent of the incident carrier energy along the tunneling direction, and across the bandgap at the same energy. It requires a high electrical field $F$ to achieve a high transmission rate $T$. Yet, many recent literatures report that the staggered heterojunctions deliver high tunneling probability, with no need for applying high electrical field near the tunneling junction [2.2]-[2.4]. This would necessitate an accurate model for heterojunction TFETs to better assess the performance of TFETs.

## 2.2 Modeling of 2-D potential in TFETs

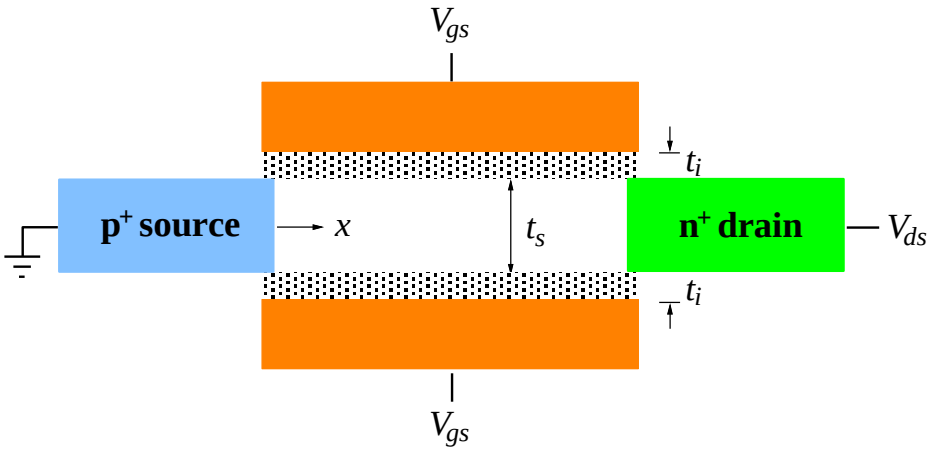



Figure 2.2: A schematic of double-gate (DG) TFET (top and bottom gate)

2-D potential has been solved analytically for a conventional Double-Gate (DG) as well as Nanowire (NW) MOSFETs in the absence of mobile charges [2.5]-[2.8]. The solution can be adopted by TFETs with a straightforward change of the source boundary condition from $n^+$ to $p^+$. The assumption of negligible mobile charges is justified when the Fermi level is below the conduction band of channel. The bias condition in which mobile charges has a significant effect on the potential will be addressed later. We focus on the mathematically simpler DG TFETs, as the solution of NW structure consists of Bessel functions. By solving Poisson's equation in 2D boundary value problems [2.3], we can express the analytic potential in the semiconductor as a series of eigenfunctions (Eq. 2.3) with discrete eigenvalues $\lambda$ satisfying the equation

$$\tan\left(\frac{\pi t_i}{\lambda}\right)\tan\left(\frac{\pi t_s}{2\lambda}\right)=\frac{\varepsilon_i}{\varepsilon_s} \tag{2.2}$$

where $t_i$, $t_s$ are defined in Fig. 2.2 and $\varepsilon_i$, $\varepsilon_s$ are the permittivities of the insulator and the semiconductor, respectively. To gain a physical insight for $\lambda$, consider the simple case of $\varepsilon_i = \varepsilon_s$, in which the solutions to Eq. (2.2) are $\lambda_n = t_s + 2t_i$, $(t_s + 2t_i)/3$, $(t_s + 2t_i)/5$, … etc. The longest $\lambda = t_s + 2t_i$ is called the scale length, which is simply the vertical distance between the two gates. It should be noted that there is a different kind of scale lengths in the literature: those derived from the "polynomial potential" models [2.9]-[2.11]. They take some general form in terms of $\sqrt{t_s t_i}$ . The polynomial-potential based models have been criticized for not satisfying the 2-D Poisson's equation in the entire semiconductor region, and for neglecting the lateral field in the insulator [2.12]. The scale lengths of the $\sqrt{t_s t_i}$

type led to an incorrect asymptotic behavior in the limits of either $t_s >> t_i$ or $t_s << t_i$. Note that with the correct scale length, $\lambda = t_s + 2t_i$, $\lambda$ cannot be smaller than $t_s$ or $2t_i$, whichever is larger.

The full 2-D potential is the sum of the long channel term, $V_{gs} - \Delta\phi$, and a series of eigenfunctions with the $x$ dependence,

$$\psi(x,y) = V_g - \Delta\phi + \sum_{n=1}^{\infty}\left[\frac{b_n \sinh\left[\pi(L-x)/\lambda_n\right] + c_n \sinh(\pi x/\lambda_n)}{\sinh(\pi L/\lambda_n)} \bullet \sin\left(\frac{n\pi}{2} + \frac{\pi y}{\lambda_n}\right)\right] \quad (2.3)$$

stemming from the source and drain boundary conditions [2.7]. Here, $L$ is the channel length and $\Delta\phi$ is the gate work function. $b_n$ and $c_n$ are constants expressed in terms of the boundary conditions and the film thickness. Fig. 2.3 shows an example of the solutions for two $V_{gs}$ values. The analytic solutions, consisting of four terms of the eigenfunctions ($\lambda_1$, $\lambda_3$, $\lambda_5$, $\lambda_7$), are validated by the Sentaurus simulation [2.13]. Note that a higher $V_{gs}$ causes a thinner barrier as well as a wider tunneling window (energy range allowing carriers to tunnel) from the source to the channel.

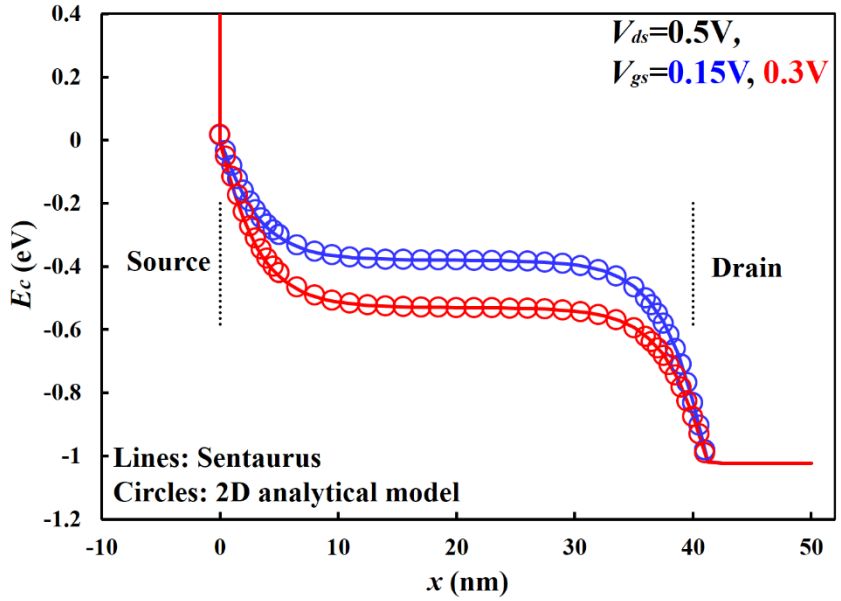



Figure 2.3: Conduction band energy at the center of film for an example of $t_s$ = 5 nm, $t_i$ = 2 nm, with $\varepsilon_i = \varepsilon_s$. The conduction band of the source is above the scale due to the band offset of heterojunction.

## 2.3 First order approximation

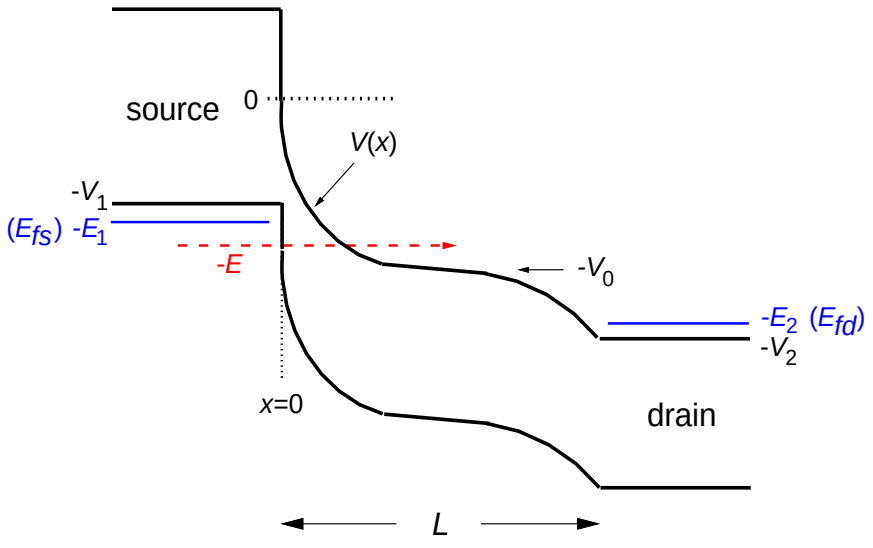



Figure 2.4: Band diagram of a heterojunction TFET with $p^+$ source and $n^+$ drain.

Fig. 2.4 depicts the band diagram of a staggered heterojunction TFET. The conduction band energy of the channel at the heterojunction boundary ($x$ = 0+) is defined as the zero energy reference. The diagram is for a turned-on TFET biased in saturation. $V_1$ is the minimum barrier height at $x$ = 0 for valence band electrons in the source to tunnel to the conduction band of channel. $V_0$ is mainly controlled by the gate voltage that determines the tunneling window. $V_2$ is related to the drain bias. $L$ is the channel length. The potential function $V(x)$ holds the key to the TFET current.

For the function $V(x)$ in the TFET model depicted in Fig. 2.4, we approximate the full 2D potential solution, Eq. (2.3), by a combination of only the $n$ = 1 term in Eq. 2.3 and the constant term,

$$V(x) = V_0 \frac{\sinh[\pi(L-x)/\lambda]}{\sinh(\pi L/\lambda)} - V_0 - (V_2 - V_0)\frac{\sinh(\pi x/\lambda)}{\sinh(\pi L/\lambda)} \tag{2.4}$$

Note that the coefficients, different from $b_1$ and $c_1$ in Eq. (2.3), are chosen to satisfy the depth independent source and drain boundary conditions in Fig. 2.2, $V(0) = V_1$ and $V(L) = -V_2$. The three terms of $V(x)$ in Eq. (2.4) represent the effects of source, gate, and drain on the channel potential. Fig. 2.5 compares Eqs. (2.3) and (2.4) for the example of $\lambda = 9$ nm ($t_s$ = 5 nm, $t_i$ = 2 nm) for $L$ = 40, 20, and 10 nm. It shows that the dual sinh function of Eq. (2.4) adequately captures the channel length dependence of the full 2D potential solution.

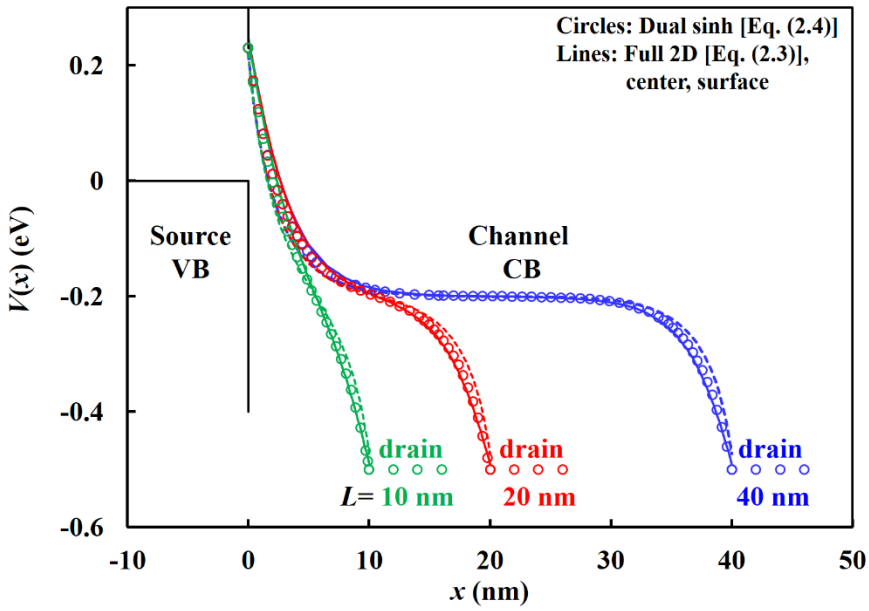



Figure 2.5: Channel potential without source depletion. Circles are the dual sinh function of Eq. (2.4) with $V_1$ = 0.23 eV, $V_0$ = 0.43 eV, $V_2$ = 0.5 eV, and $\lambda$ = 9 nm. Solid lines are $-q\psi(x, 0)$ and dashed lines $-q\psi(x, t_s/2)$, both from Eq. (2.3) with $n$ = 1, 3, 5, 7 terms.

Yet, the dual sinh function of Eq. (2.4) cannot be integrated in closed form. For TFETs biased in saturation, the current is mainly determined by the tunneling barrier near the source. In Fig. 2.6, we zero in on the potential solution close to the source, where the eigenfunctions are dominated by the term $\propto \sinh[\pi(L-x)/\lambda]/\sinh[\pi L/\lambda] \approx \exp(-\pi x/\lambda)$. Depending on the film thickness, the potential has a slight variation between the surface and the center of the semiconductor. To enable an analytic model for TFET, we approximate both the center and the surface potentials with a single exponential function, $\exp(-\pi x/\lambda) - 1$, where $\lambda = t_s + 2t_i$ is the scale length [2.14]-[2.16]. The approximation of uniform potential in the depth direction is more valid for the case of relatively thin $t_s$ and thick $t_i$. A similar function works for NW TFET as well, with $\lambda = \pi(r_s + t_i)/\alpha$ for the case

$\varepsilon_i = \varepsilon_s$, where $r_s$ is the NW radius and $\alpha = 2.405$ is the first zero of the zeroth order Bessel function [2.8].

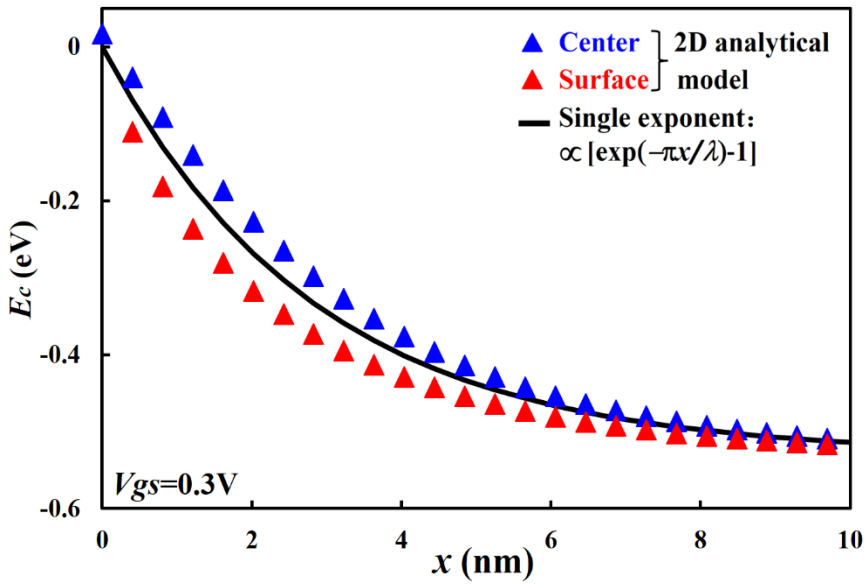



Figure 2.6: Approximation of the center and surface potential near the source with a single exponential function. The scale length is $\lambda = 9$ nm in this example.

Hence, the single exponential function is a valid approximation to model the TFET on-current in the long channel limit. It can be integrated in closed form thus giving mathematical simplifications, which will be elaborated in Chapter 3.

The text of Chapter 2, in part, is a reprint of the material as it appears in "An analytic model for heterojunction tunnel FETs with exponential barrier" by Yuan Taur, Jianzhi Wu and Jie Min, IEEE Transaction on Electron Devices, May 2015. The dissertation author was a co-author of this paper.

The text of Chapter 2, in part, is a reprint of the material as it appears in "Short channel effects in tunnel FETs" by Jianzhi Wu, Jie Min and Yuan Taur IEEE Transaction on Electron Devices, Sept. 2015. The dissertation author was the primary investigator and

author of this paper.

## 2.4 References

[2.1] E. O. Kane, "Theory of tunneling." *Journal of Applied Physics* 32, no. 1 (1961): 83-91.

[2.2] L. Wang, E. Yu, Y. Taur, and P. Asbeck, "Design of tunneling field-effect transistors based on staggered heterojunctions for ultralow-power applications," *Electron Device Letters, IEEE* 31.5 (2010): 431-433.

[2.3] Y. Taur, J. Wu and J. Min, "An analytic model for heterojunction tunnel FETs with exponential barrier," *IEEE Trans. Electron Devices*, vol.62, no.5, pp. 1399-1404, May 2015.

[2.4] J. Wu, J. Min, J. Ji, and Y. Taur. "An analytic model for heterojunction and homojunction tunnel FETs with 3D density of states." In *Device Research Conference (DRC), 2015 73rd Annual*, pp. 249-250. IEEE, 2015.

[2.5] D. J. Frank, Y. Taur, and H.-S. P. Wong, "Generalized Scale Length for Two-Dimensional Effects in MOSFETs," *IEEE Electron Device Letters*, Vol. 19, No. 10, pp. 385-387, Oct. 1998.

[2.6] S.-H. Oh, D. Monroe, and J. M. Hergenrother, "Analytic description of short-channel effects in fully-depleted double-gate and cylindrical, surrounding-gate MOSFETs," *IEEE Electron Device Letters*, Vol. 21, pp. 445-447, 2000.

[2.7] X. P. Liang and Y. Taur, "A 2-D analytical solution for SCEs in DG MOSFETs," *IEEE Trans. Electron Devices*, vol. 51, no. 9, pp. 1385–1391, Sep. 2004.

[2.8] B. Yu, L. Wang, Y. Yuan, P. M. Asbeck, and Y. Taur, "Scaling of nanowire transistors," *IEEE Trans. Electron Devices*, vol. 55, pp. 2846-2858, Nov. 2008.

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

[2.10] K. Suzuki, T. Tanaka, Y. Tosaka, H. Horie, and Y. Arimoto, "Scaling theory for double-gate SOI MOSFETs," *IEEE Trans. Electron Devices*, vol. 40, no. 12, pp. 2326–2329, Dec. 1993.

[2.11] C. P. Auth and J. D. Plummer, "Scaling theory for cylindrical, fully-depleted,

surrounding-gate MOSFETs," *IEEE Electron Device Letters*, Vol. 18, pp. 74-76, 1997.

[2.12] Q. Xie, J. Xu, and Y. Taur, "Review and Critique of Analytic Models of MOSFET Short-Channel Effects in Subthreshold," *IEEE Trans. Electron Devices*, vol. 59, pp. 1569-1579, June 2012.

[2.13] *TCAD Sentaurus Device Manual* (Synopsys, Inc., Mountain View, CA, 2010).

[2.14] P. M. Solomon, D. J. Frank and S. O. Koswatta, "Compact model and performance estimation for tunneling nanowire FET," in *Device research conference (DRC)*, 2011 69th annual; 2011. p. 197–8. doi:10.1109/DRC.2011.5994495.

[2.15] L. Liu, D. Mohata, and S. Datta, "Scaling length theory of double-gate interband tunnel field-effect transistors," *Electron Devices, IEEE Transactions on* 59.4 (2012): 902-908.

[2.16] B. Rajamohanan, D. Mohata, A. Ali, and S. Datta, "Insight into the output characteristics of III-V tunneling field effect transistors," *Applied Physics Letters* 102, no. 9 (2013): 092105.

# Chapter 3 Modeling of ON-current

## 3.1 Source-to-channel tunneling

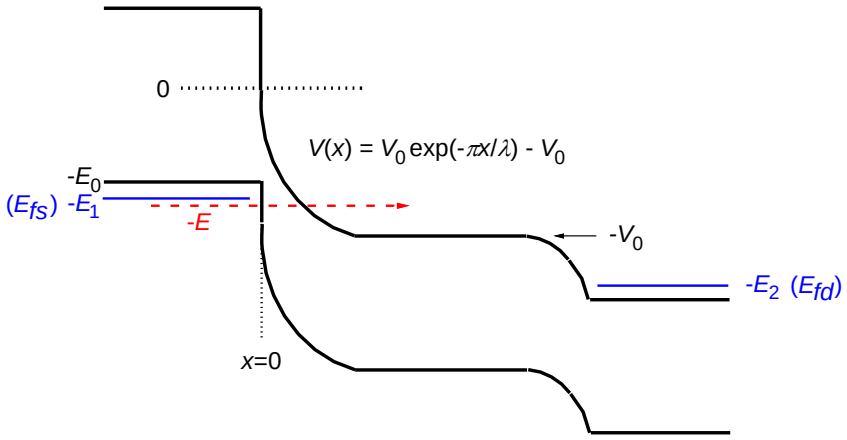



Figure 3.1: Band diagram of a heterojunction TFET biased in saturation.

The band diagram of a staggered heterojunction TFET is shown in Fig. 3.1. The zero energy reference is chosen to be the conduction band energy of the channel at the heterojunction boundary. It staggers below the conduction band energy of the source by the band offset. The diagram is for the TFET biased in turn-on and in saturation. Using the single exponent approximation, the energy barrier for electrons tunneling from the valence band of the source to the conduction band of the channel takes the form

$$V(x) = V_0 \exp(-\pi x / \lambda) - V_0 \tag{3.1}$$

where $V_0$ is mainly controlled by the gate voltage. For electrons at energy $-E$ in the valence band, the tunneling probability is given by the WKB integral as

$$T(E) = \exp\left\{-\frac{2\sqrt{2m}}{\hbar}\int_0^d \sqrt{V(x)+E}\,dx\right\} = \exp\left\{-\frac{2\sqrt{2m}}{\hbar}\int_0^d \sqrt{V_0 e^{-\pi x/\lambda} - V_0 + E}\,dx\right\} \tag{3.2}$$

where $V(d) + E = 0$. The integral can be carried out analytically to yield

$$T(E) = \exp\left\{-\frac{4\lambda\sqrt{2m}}{\pi\hbar}\left[\sqrt{E} - \sqrt{V_0 - E}\sin^{-1}\sqrt{E/V_0}\right]\right\} \tag{3.3}$$

For a 1-D ballistic TFET, the current is given by the Landauer equation,

$$\begin{aligned} I_{ds} &= \frac{2q}{h}\int_{E_0}^{V_0}(f_s - f_d)T(E)dE \\ &= \frac{2q}{h}\int_{E_0}^{V_0}\left[\frac{1}{1+e^{(E_1-E)/kT}} - \frac{1}{1+e^{(E_2-E)/kT}}\right]\exp\left\{-\frac{4\lambda\sqrt{2m}}{\pi\hbar}\left[\sqrt{E} - \sqrt{V_0 - E}\sin^{-1}\sqrt{E/V_0}\right]\right\}dE \end{aligned} \tag{3.4}$$

where $E_0$, $E_1$, $E_2$ are positive quantities defined in Fig. 3.1. The $f_s$ and $f_d$ are the occupancies of states at source and drain. Note that $V_{ds} = (E_2 - E_1)/q$. $V_{gs}$ (in the unit of volt) is defined such that $V_{gs} = 0$ corresponds to $V_0 = E_0$ (in the unit of Joule), i.e., where the tunneling window (energy range allows carrier to tunnel) starts to open up. This definition has the merit that the off condition is maintained at the same $V_{gs}$ for different designs, but it implies a choice of gate work function dependent on the band offset ($E_0$) and source degeneracy ($E_1$).

For energies where $f_s - f_d \approx 1$, the current is proportional to the area under $T(E)$ from $E_0$ to $V_0$. Fig. 3.2 considers an example of $m = 0.1m_0$, $\lambda = 9$ nm ($t_s = 5$ nm, $t_i = 2$ nm, $\varepsilon_i = \varepsilon_s$), and $E_0 = 0.15$ eV for several values of $V_0$ ($= qV_{gs} + E_0$). It is clear that as $V_0$ or $V_{gs}$ increases, most of the current (area) gain comes from thinning of the barrier, rather than from expanding the tunneling window. It is also clear that a smaller $E_0$ resulting from a larger band offset would significantly raise the tunneling current. Too low an $E_0$ (< 0.1 eV) or broken gap ($E_0 < 0$) designs, however, have been reported to result in subthreshold swing (SS) > 60 mV/decade [3.1]. In addition, the subthreshold swing will further degrade due to the presence of band-tail states.

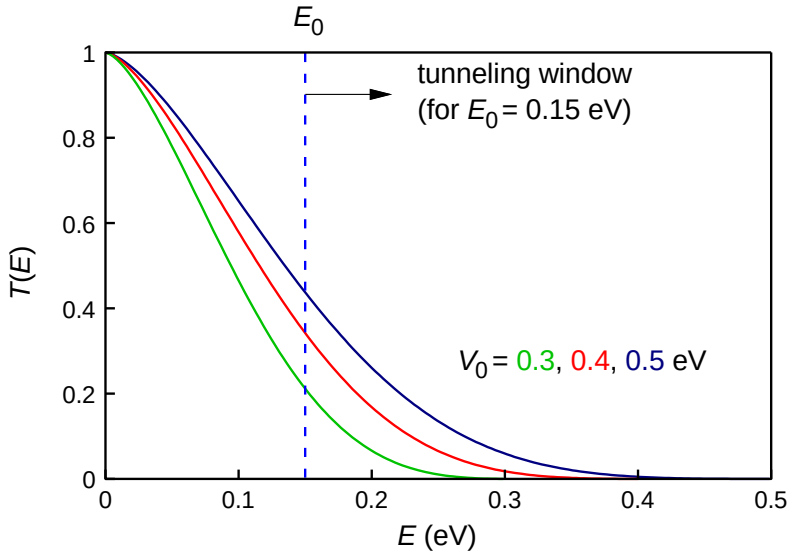



Figure 3.2: Tunneling probability versus carrier energy. The tunneling window is constrained by the density of states to $E_0 < E < V_0$.

For a given material and doping, the band offset $E_0$ and source degeneracy $E_1$ are fixed. Continuous $I_{ds}(V_{gs}, V_{ds})$ characteristics can be generated from Eq. (3.4) once the parameters $V_0$ and $E_2$ are expressed as continuous functions of $V_{gs}$ and $V_{ds}$. It is straightforward to show that $E_2 = E_1 + qV_{ds}$. For $V_0$, we first consider saturation in which the drain Fermi level is below the conduction band of channel, i.e., $E_2 > V_0$ in Fig. 3.1, and $V_0$ is controlled only by $V_{gs}$. With our choice of reference, $V_0 = E_0 + qV_{gs}$.

In Fig. 3.3, we plot $I_{ds}$-$V_{gs}$ characteristics for several values of the source degeneracy (the energy difference between source fermi level and valence band energy), $d_1 = E_1 – E_0$. $d_1$ is adjusted independent of source doping which is assumed to be high enough that there is negligible depletion in the source region. By our definition, $I_{ds} = 0$ at $V_{gs} = 0$ where $V_0 = E_0$. In reality, there is a current floor set by the source-to-drain tunneling or by minority carrier generation and recombination which is not considered here. The heterojunction TFET example in Fig. 3.3(a) is that of an AlGaAsSb source and InGaAs channel with $E_0$

= 0.23 eV. It is well known that the TFET turn-off slope is degraded by source degeneracy because of the *kT* tail of the Fermi-Dirac distribution below the valence band edge. Here, we also observe that the on current is degraded by higher source degeneracy. The underlying reason can be seen in Fig. 3.1. For a given $E_0$, electrons at the top of the valence band see the lowest barrier. Source degeneracy pushes the electron population below the valence band edge, hence increases their barrier height for tunneling. Such effect is verified qualitatively by the Sentaurus simulations shown in Fig. 3.3(b) [3.2]. Too low a source doping of course leads to reduced field at the junction, longer tunneling path, and decreased TFET current. There is an optimum source doping level in the range of mid-$10^{19}$ $cm^{-3}$.

A maximum TFET current can be derived from the analytic model in the limit of $E_0$ = 0 (zero effective bandgap) and $f_s - f_d = 1$. For $E/V_0 << 1$, Eq. (3.3) is approximated as

$$T(E) \approx \exp\left\{-\frac{4\lambda\sqrt{2m}}{\pi\hbar}\frac{E^{3/2}}{3V_0}\right\} \tag{3.5}$$

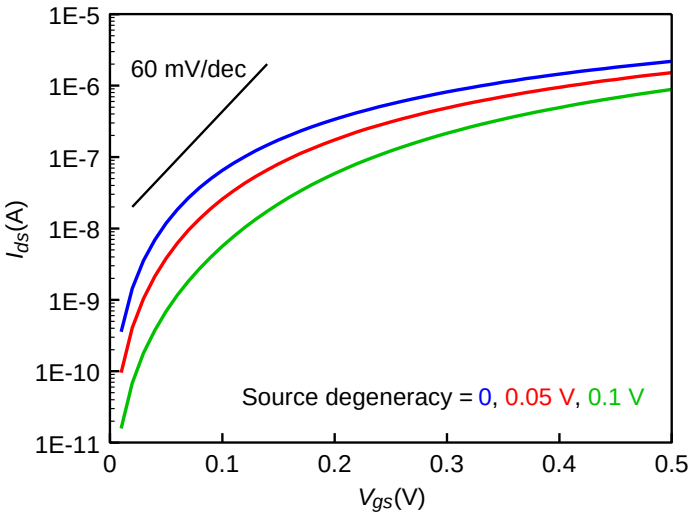
Eq. (3.4) then gives

$$I = \frac{2q}{h}\int_0^{V_0}\exp\left\{-\frac{8\lambda\sqrt{2m}}{3hV_0}E^{3/2}\right\}dE \approx \frac{2q}{h}\left[\frac{3hV_0}{8\lambda\sqrt{2m}}\right]^{2/3}\int_0^{\infty}\exp(-y^{3/2})dy \tag{3.6}$$

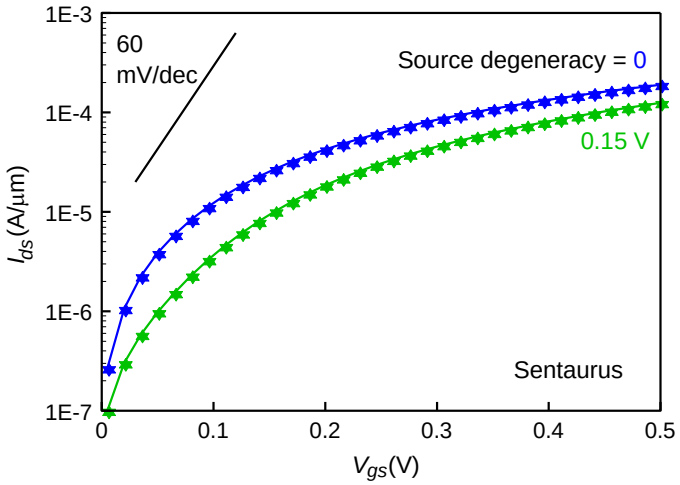
The numerical integral ≈ 0.9 and $V_0 = qV_{gs}$. Therefore,

$$I_{\max} = 0.9q\left[\frac{3qV_{gs}}{4\lambda\sqrt{mh}}\right]^{2/3} \tag{3.7}$$

For $m = 0.1m_0$, $\lambda$ = 9 nm, $V_{gs}$ = 0.5 V, $I_{\max} \approx 13$ μA. This value is in line with the non-equilibrium Green's function simulated results in [3.21].




(a)




(b)

Figure 3.3: (a) $I_{ds}$ versus $V_{gs}$ for source degeneracies, $d_1 = E_1 - E_0 = 0$, 0.05 eV, and 0.1 eV (top to bottom). (b) Sentaurus simulation, in which the source degeneracy is varied by adjusting the valence band effective density of states with no change in the source doping.

## 3.2 Debias effects in the linear region

When the TFET is biased in the linear region, $E_2$ approaches $E_1$ and $f_d \neq 0$. Moreover, since the Fermi level in the channel is near or above the conduction band edge, there is a

de-biasing effect on $V_0$ due to the channel inversion charge [3.3][3.4]. Instead of $V_0 = E_0 + qV_{gs}$ as in the high $V_{ds}$ case, $V_0$ is reduced to $E_0 + q(V_{gs} - Q_{inv}/C_{ox})$, where $Q_{inv}/C_{ox}$ is the potential drop across the gate insulator. This is an electrostatic effect unrelated to the transport. For a given $V_{gs} - V_{ds}$, $Q_{inv}/C_{ox}$ can be calculated from a continuous, analytic solution of Poisson's equation with mobile charge for DG MOSFETs [3.5]:

$$Q_{inv} = \frac{4kT\varepsilon_s}{qt_s}\beta\tan\beta \tag{3.8}$$

where the intermediary parameter $\beta$ is solved from an implicit equation [3.5],

$$\frac{q(V_{gs}-V_{ds}-d_1)}{2kT} - \ln\left[\frac{2}{t_s}\sqrt{\frac{2\varepsilon_s kT}{q^2 N_c}}\right] = \ln\beta - \ln[\cos\beta] + \frac{2\varepsilon_s t_i}{\varepsilon_i t_s}\beta\tan\beta \tag{3.9}$$

Here, $N_c$ is the effective density of states of the conduction band and $d_1$ is the source degeneracy. They play a key role on the onset of de-biasing versus $V_{ds}$, and therefore on the linear region characteristics. Fig. 3.4 shows three de-biasing curves, one for silicon-like and two for InGaAs with different source degeneracies. The silicon-like case assumes the $N_c$ of silicon, with everything else the same as the AlGaAsSb/InGaAs heterojunction example considered in Fig. 3.3. The high $N_c$ of silicon results in a significant de-bias as soon as $V_{ds}$ is below $V_{gs}$ + 0.1 V. The de-bias for InGaAs does not start until $V_{ds}$ is below $V_{gs}$ – 0.05 V if no source degeneracy, and below $V_{gs}$ – 0.15 V if there is a source degeneracy of 0.1 eV.

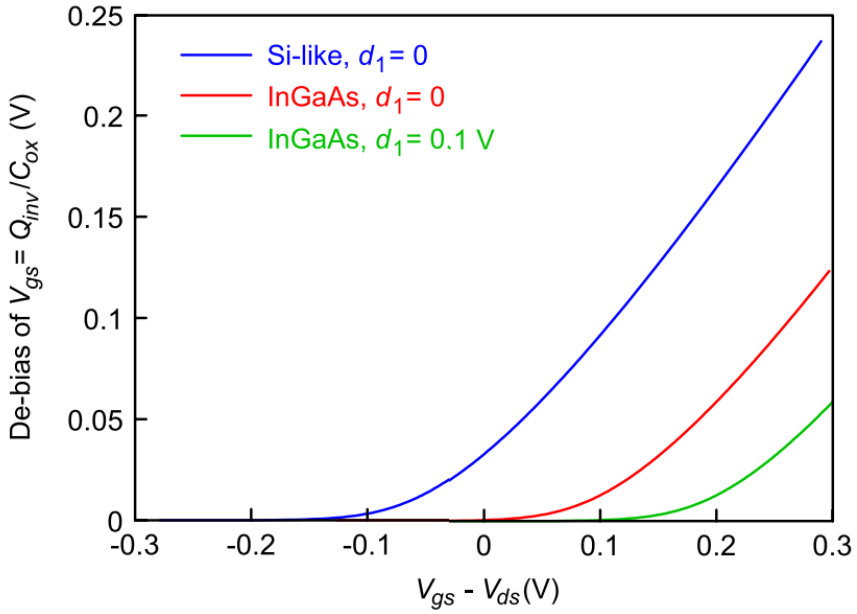



Figure 3.4: Reduction of $V_0$ in the linear region by $Q_{inv}$ in the channel. $N_c = 3\times10^{19}$ cm$^{-3}$ for Si-like, $N_c = 8.7\times10^{16}$ cm$^{-3}$ for InGaAs. $\varepsilon_s = \varepsilon_i = 14.6\varepsilon_0$ for all cases.

To incorporate the de-bias effect in the generation of continuous $I_{ds}(V_{gs}, V_{ds})$ characteristics from Eq. (3.4), we take an extra step to first calculate $Q_{inv}$ from Eqs. (3.9) and (3.8) for given $V_{gs}$ and $V_{ds}$. Then $V_0$ is set to $E_0 + q(V_{gs} - Q_{inv}/C_{ox})$ in the current integral. At a fixed $V_{gs}$, when $V_{ds}$ becomes high enough, $V_{gs} - V_{ds}$ in Fig. 3.4 goes negative and $Q_{inv} \rightarrow 0$. The corresponding $I_{ds}$ makes a smooth transition to the saturation value for that $V_{gs}$. Fig. 3.5(a) shows the model generated $I_{ds}$-$V_{ds}$ characteristics for InGaAs, $d_1 = 0.1$ eV with and without de-bias. In the no de-bias case, the only $V_{ds}$ dependent factor in Eq. (3.4) is $f_d$. $I_{ds}$ saturates quickly when $E_{fd}$ is ~ 0.15 eV below $E_{fs}$ and the current becomes source injection limited. The effect of de-bias is to reduce the linear region current and push $V_{dsat}$ higher with no impact on $I_{dsat}$. For the case of InGaAs with $d_1 = 0$ in Fig. 3.5(b), the de-bias effect is more pronounced, resulting in higher $V_{dsat}$. But the magnitude of $I_{dsat}$ is significantly higher than that of $d_1 = 0.1$ eV, for reasons given earlier with Fig. 3.3. The

most severe de-bias happens with the silicon-like TFET in Fig. 3.5(c). The high $N_c$ of silicon gives rise to the super-linear $I_{ds}$-$V_{ds}$ characteristics. These trends are all confirmed qualitatively by Sentaurus simulations, as well as by published hardware data in the literature [3.6]-[3.9].

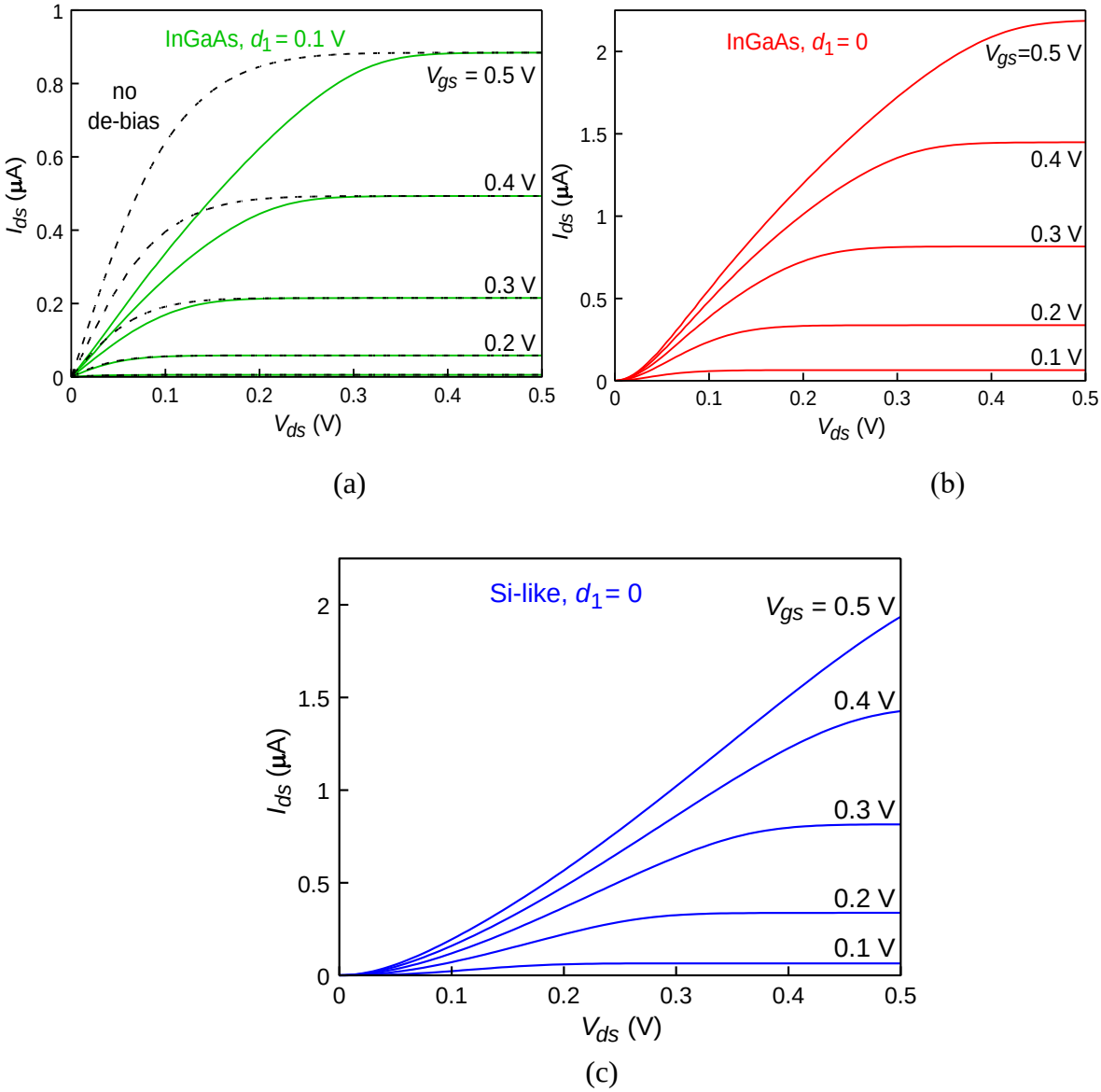



Figure 3.5: Model generated $I_{ds}$-$V_{ds}$ characteristics for the three de-biasing conditions in Fig. 3.4. The dashed curves in (a) are for no de-biasing.

The $I_{ds}$-$V_{ds}$ characteristics in Fig. 3.5 are generated by Eq. (3.4) of the model with a modification that the low end of the tunneling window is limited by the $E_c$ of the channel or the $E_c$ of the drain, whichever is higher. In other words, the upper bound of the integral in Eq. (3.4) is given by $V_0$ or $E_0 + qV_{ds} + d_1 + d_2$, whichever is lower. Here $d_2$ is the drain degeneracy. In practice, this makes only a very slight difference in $I_{ds}$ at $V_{ds}$ below 0.1 V, because in that energy range, the tunneling path is long and both $f_s$ and $f_d \approx 1$. There is very little contribution to the tunneling current.

### 3.3 Dimensionality dependence of TFET performance

With the recent emergence of 1D and 2D semiconductors, this section assesses the performance of tunnel FETs made in semiconductors of different dimensionality.

The TFET currents with 1D, 2D, or 3D density of states are given by [3.10]:

$$I_{1D} = \frac{2q}{h}\int (f_s - f_d)T(E)dE \tag{3.10}$$

$$I_{2D} = \frac{2q}{h}\frac{\sqrt{2\pi m}}{h}\int\int (f_s - f_d)T(E+E_\perp)\frac{dE_\perp}{\sqrt{\pi E_\perp}}dE \tag{3.11}$$

$$I_{3D} = \frac{2q}{h}\frac{2\pi m}{h^2}\int\int (f_s - f_d)T(E+E_\perp)dE_\perp dE \tag{3.12}$$

The WKB integral in the tunneling probability can be analytically evaluated for the exponential barrier to model the on-current, which has been explained in the Chapter 2:

$$T(E) = \exp\left\{-(2\sqrt{2m}/\hbar)\int_0^l \sqrt{V(x)+E}\,dx\right\} = \exp\left\{-(8\lambda\sqrt{2m}/h)\left[\sqrt{E+V_1} - \sqrt{V_0-(E+V_1)}\sin^{-1}\sqrt{(E+V_1)/V_0}\right]\right\} \quad (3.13)$$

By expanding $T(E + E_\perp)$ to the first order of $E_\perp$ [3.11], the $E_\perp$ integration can be executed to consolidate all TFET currents to a single integral of the following general form (in proper units of A, A/m, and A/m$^2$ for $I_{1D}$, $I_{2D}$, and $I_{3D}$),

$$I_{n\mathrm{D}} = (2q/h)\int_0^{V_0-V_1}(f_s - f_d)T_{n\mathrm{D}}(E)\left[2\pi\, mE_t/h^2\right]^{(n-1)/2}dE \quad (3.14)$$

where $n$ = 1, 2, 3. Here, $T_{1D}(E) = T(E)$, $T_{2D}(E) = T(E)$ erf$[(E/E_t)^{1/2}]$, $T_{3D}(E) = T(E)[1 - \exp(-E/E_t)]$, with $E_t$ given by:

$$E_t = \left\{-(\sqrt{2m}/\hbar)\int_0^l dx/\sqrt{V(x)+E}\right\}^{-1} = h\sqrt{V_0-(E+V_1)}\,/\left[4\lambda\sqrt{2m}\sin^{-1}\sqrt{(E+V_1)/V_0}\right] \quad (3.15)$$

The gate and drain voltage dependence of $I_{n\mathrm{D}}$ comes through $V_0 = V_1 + qV_{gs}$ (saturation) and $f_d = \{1 + \exp[(-E + qV_{ds})/kT]\}^{-1}$. $f_s = [1 + \exp(-E/kT)]^{-1}$.

Fig. 3.6 plots $T_{1D}$, $T_{2D}$, and $T_{3D}$ versus $E$. For 1D, $T(E)$ is highest at $E$ = 0. For 2D and 3D TFETs, however, because of the density of states factors due to $E_\perp$, the $T_{2D}$ and $T_{3D}$ products start at 0 at $E$ = 0 then rise to a peak at $E > 0$, thus missing the contribution at low $E$ where the tunneling probability is highest.

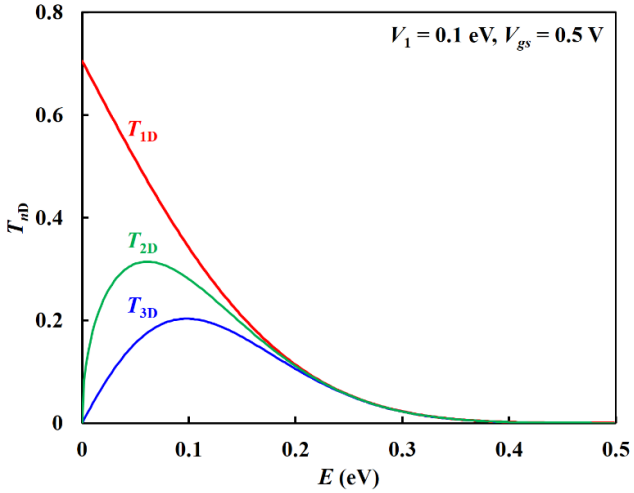



Figure 3.6: $T_{n\mathrm{D}}$ versus $E$. $m = 0.1m_0$ and $\lambda = 9$ nm are assumed throughout this section

In Fig. 3.7, we plot model generated $I_{ds}$-$V_{gs}$ curves for 3D TFETs with fixed $V_{ds}$ = 0.5 V and several $V_1$. The model results are consistent with Sentaurus simulations [3.2]. They quantify the gains in current from the narrower bandgap of the heterojunctions. The highest current is obtained in the case of zero effective bandgap or $V_1 \rightarrow 0$. $V_1 < 0$ or broken gap TFETs are vulnerable to subthreshold swings > 60 mV/decade [3.1].

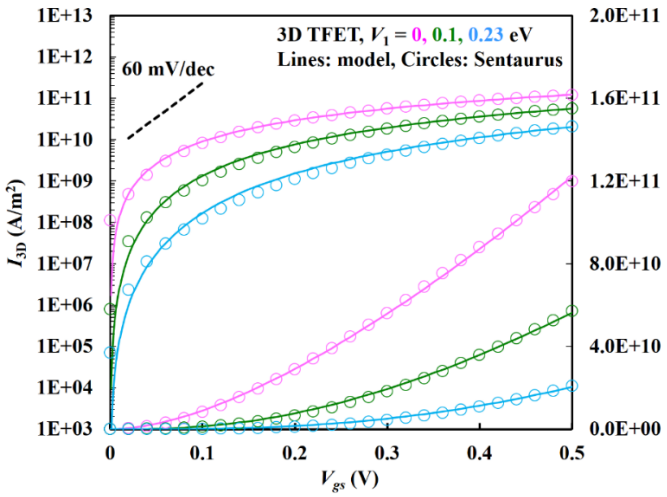



Figure 3.7: $I_{3\mathrm{D}}$-$V_{gs}$ for three different bandgaps. Circles: Sentaurus simulation. The semiconductor is 5 nm thick, with 2 nm thick insulator on each side. $L$ = 40 nm, $m = 0.1m_0$, $V_{ds}$ = 0.5 V. The same set of parameters is used in Sentaurus.

Expressions for the maximum currents of 1D, 2D, and 3D heterojunction TFETs can be derived in the limit of $V_1 \to 0$. With $f_s - f_d$ set to 1 under the condition $V_{ds} \geq V_{gs} >> kT/q$, the current integrals are numerically fitted to various powers of $V_{gs}$:

$$I_{1D}(V_1 = 0) \approx 0.74q\left(\frac{q^2}{mh\lambda^2}\right)^{1/3} V_{gs}^{\,2/3} \tag{3.16}$$

$$I_{2D}(V_1 = 0) \approx 0.58\frac{q^2 V_{gs}}{\lambda h} \tag{3.17}$$

$$I_{3D}(V_1 = 0) \approx 0.47\frac{q^2\sqrt{mq}V_{gs}^{\,3/2}}{\lambda h^2} \tag{3.18}$$

The $V_{gs}$ dependence can be summarized as $I_{n\mathrm{D}} \propto V_{gs}^{\,k}$ where $k = (3/2)^{n-2}$ for all $n$ = 1, 2, 3. Note that the $V_{gs}^{3/2}$ dependence of $I_{3D}$ is the same as the saturation current of a ballistic MOSFET [3.12]. All maximum currents are $\propto 1/\lambda$ or nearly so, thus improve with scaling of the film thickness. The dependence on effective mass $m$ (assumed isotropic) is mixed. $I_{1D}$ goes up with lighter $m$ due to the tunneling mass in the WKB integral. $I_{3D}$ increases with $m$ where the density of states is more important. $I_{2D}$ is independent of $m$ as the two effects cancel (Table 3.1).

Table 3.1: Contrast of 1D, 2D, and 3D TFET parameters

| | $T_{nD}$ ($E$=0) | d$T_{nD}$/d$E$ ($E$=0) | units of $I_{nD}$ | $I_{nD}$-$m$ dependence | $I_{nD}$-$V_{gs}$ dependence |
|---|---|---|---|---|---|
| 1D | > 0 max. | < 0 | A | $\propto m^{-1/3}$ | $\propto V_{gs}^{\,2/3}$ |
| 2D | 0 | $\infty$ | A/m | Indep. of $m$ | $\propto V_{gs}$ |
| 3D | 0 | > 0 | A/m$^2$ | $\propto m^{1/2}$ | $\propto V_{gs}^{\,3/2}$ |

By employing the debias model in the last section, Fig. 3.8 plots the model generated $I_{ds}$-$V_{ds}$ curves for 1D, 2D, and 3D TFETs. The sub-linear, linear, and super-linear $V_{gs}$ dependence of the saturation currents is evident.

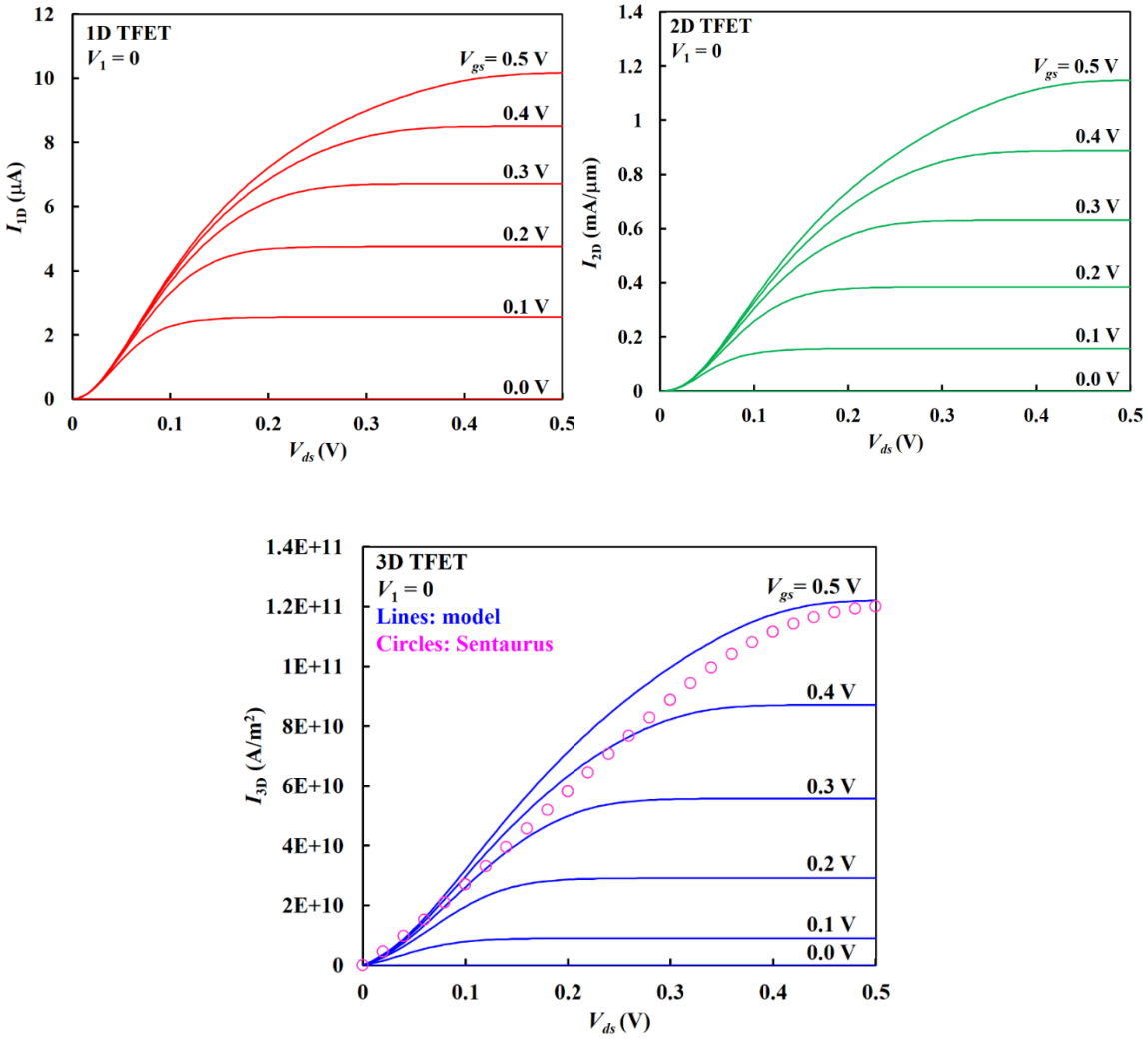



Figure 3.8: $I_{1D}$-$V_{ds}$, $I_{2D}$-$V_{ds}$, and $I_{3D}$-$V_{ds}$ for zero staggered bandgap. Sentaurus curve is shown for $V_{gs}$ = 0.5 V.

Fig. 3.9 plots $CV_{dd}/I_{ds}(V_{gs}=V_{ds}=V_{dd})$ calculated from Eq. (3.14) versus $V_{dd}$ for $n$ = 1, 2, 3. For 2D, $I_{ds}$ is taken to be $I_{2D}$ in A/m or mA/μm. For 3D, $I_{ds}$ is $I_{3D}$ times 5 nm, the semiconductor thickness. For 1D, $I_{ds}$ is $I_{1D}$ divided by 5 nm, assuming that one piece of 1D semiconductor can be placed per 5 nm width. $C$ is assumed to be a constant, 2 fF/μm. It is based on $C \sim 2\varepsilon L/t_{ox}$, where $\varepsilon \sim 0.1$ fF/μm, $L/t_{ox} \sim 10$, and a factor of 2 is for CMOS circuits. The exact value of $C$ does not affect the main point here. For 3D TFETs, $CV/I$ goes up as $V_{dd}$ decreases, similar to the conventional MOSFETs. For 2D, $CV/I$ is more or less flat. But for 1D, $CV/I$ ***decreases*** as $V_{dd}$ is reduced, implying that instead of delay-power tradeoff, both the delay and the power improve at lower voltages. A minimum $CV/I$ is reached at $V_{dd} \approx 0.15$ V for 1D TFETs, shown in more detail in Fig. 3.10. Below that $CV/I$ goes up sharply because of the $kT$ transition width of the Fermi-Dirac distribution function [3.13][3.14]. If the temperature is reduced to 150 K, the $V_{dd}$ for minimum $CV/I$ is also reduced by 2× to ≈ 0.08 V. The zero temperature curve, for which $f_s - f_d = 1$, keeps on decreasing until reaching the quantum conductance limit of $I/V = 2q^2/h$. Thus the lower limit is $V_{dd} \approx 6kT/q$ for 1D TFETs without losing performance. The ultimate voltage scaling is achieved through scaling of dimension.

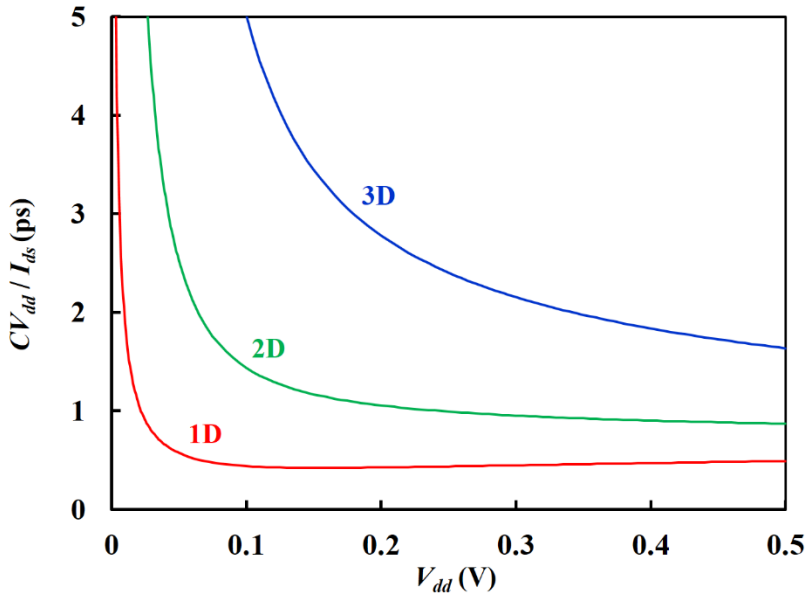



Figure 3.9: *CV*/*I* versus $V_{dd}$ for 1D, 2D, and 3D TFETs. $C$ = 2 fF/μm, $V_1$ = 0.

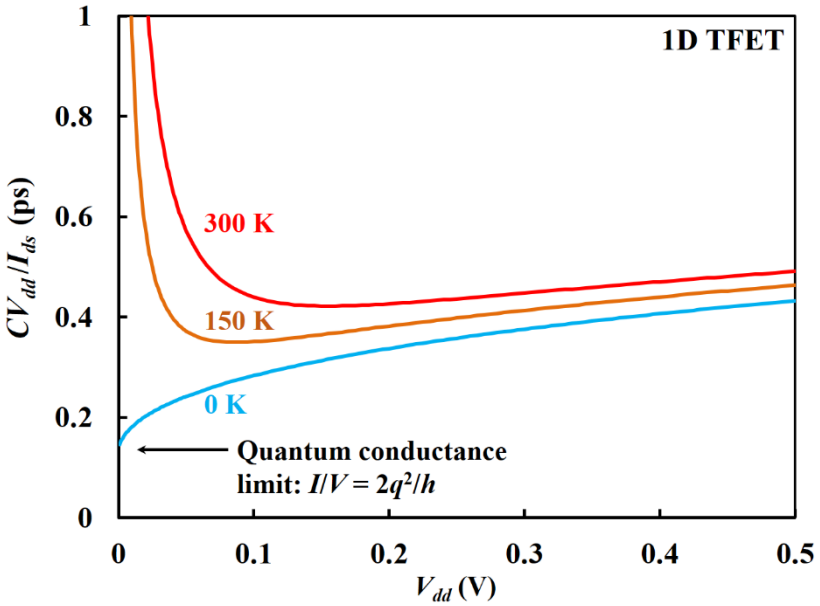



Figure 3.10: *CV*/*I* versus $V_{dd}$ for 1D TFETs at three different temperatures. $V_1$ = 0.

## 3.4 Source doping effects

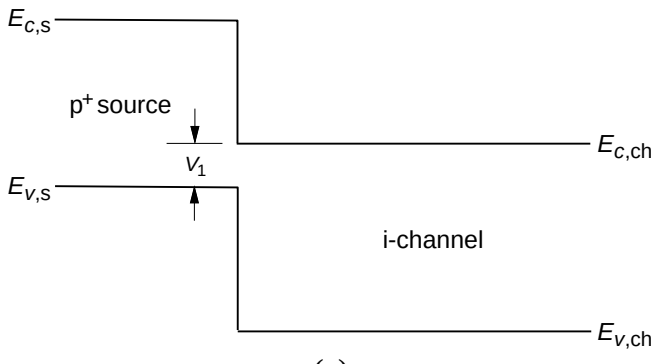
One of the key design parameters for a tunnel MOSFET is the source doping concentration. Too low a doping will cause depletion in the source junction and increase

the barrier height and tunneling distance. Too high a doping will give rise to source degeneracy that softens the TFET turn-on characteristics [3.1][3.15]-[3.17]. This section investigates the effect of source doping on heterojunction TFETs by analytically modeling the channel potential profile along with source depletion effects.

### 3.4.1 Depletion approxiamtion

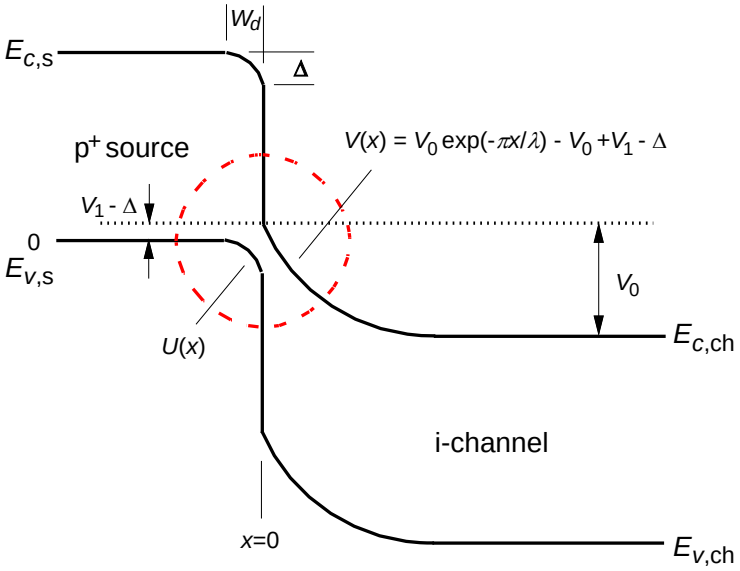
Fig. 3.11(a) shows the source-channel band diagram of a heterojunction TFET with a staggered bandgap $V_1$. Both the source and channel bandgaps are assumed to be much wider than $V_1$ that only $E_{v,s}$ and $E_{c,ch}$ are relevant as far as the tunneling current is concerned. Fig. 3.11(b) shows that the TFET is turned on by gate modulation of the channel potential. The effect of gate voltage goes into $V_0 + \Delta$, the bending of the channel bands and the source bands. The latter extends over a depletion width $W_d$.

By choosing the valence band edge of source (far from the heterointerface) as the zero energy reference, the conduction band of channel is expressed as (note that $V$ is energy in joule):

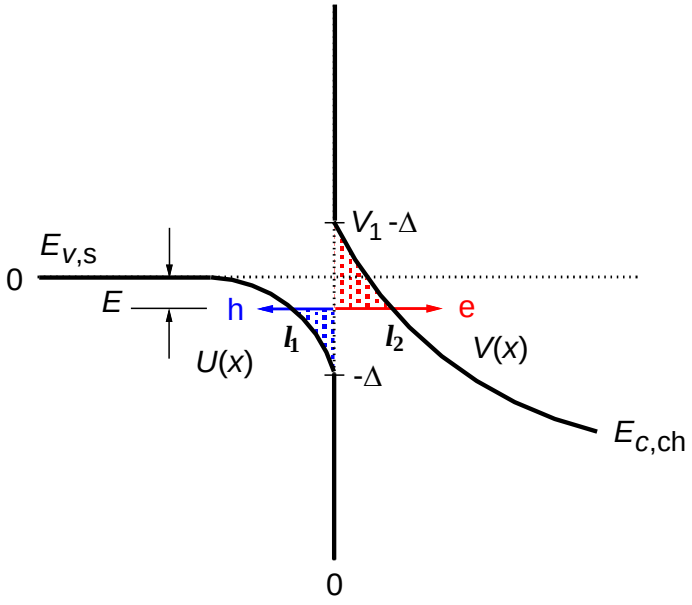



(a)

Figure 3.11: (a) Band diagram of staggered heterojunction under flatband. (b) Band diagram of turned on TFET. The circled tunneling region is magnified in (c). The shaded areas depict the conduction band and valence band barriers for electron and hole tunneling respectively




(b)




(c)

Figure 3.11: continued (b) Band diagram of turned on TFET. The circled tunneling region is magnified in (c). The shaded areas depict the conduction band and valence band barriers for electron and hole tunneling respectively

$$V(x) = V_0 \exp(-\pi x / \lambda) - V_0 + V_1 - \Delta \tag{3.19}$$

where $x = 0$ is at the heterojunction boundary and $\lambda$ is the scale length solved from 2D Poisson's equation as a function of the physical dimension of the gate insulator and semiconductor and their permittivities. For TFETs with equal permittivities, $\lambda = t_s + 2t_i$ for DG and $\lambda = \pi(r_s + t_i)/2.405$ for NW [3.18], where $t_s$ is the semiconductor film thickness, $t_i$ is the gate insulator thickness, and $r_s$ is the NW radius. Here we assume that the gate length is over $2\lambda$ so the drain effect can be neglected.

By applying the condition that the field is continuous from one side of the heterojunction to the other (assuming no change of permittivity and no mobile charge at the interface), we have

$$\frac{1}{q}\left|\frac{dV}{dx}\right|_{x=0} = \frac{\pi V_0}{q\lambda} = \frac{qN_aW_d}{\varepsilon_s} \tag{3.20}$$

where the depletion approximation is employed with source doping $N_a$. We can then write

$$W_d = \frac{\pi\varepsilon_s V_0}{q^2\lambda N_a} \tag{3.21}$$

and

$$\Delta = \left(\frac{qN_aW_d}{\varepsilon_s}\right)\left(\frac{W_d}{2}\right)q = \frac{\pi^2\varepsilon_s V_0^{\,2}}{2q^2\lambda^2 N_a} \tag{3.22}$$

The function describing the bending of the source valence band in Fig. 3.13(b) is therefore

$$U(x) = -\frac{q^2N_a}{2\varepsilon_s}(x + W_d)^2 = -\frac{q^2N_a}{2\varepsilon_s}\left[x + \frac{\pi\varepsilon_s V_0}{q^2\lambda N_a}\right]^2 \tag{3.23}$$

Note that the TFET starts to turn on when $V_0 + \Delta = V_1$. If we define this condition to be $V_{gs} = 0$, then

$$qV_{gs} = V_0 + \Delta - V_1 = V_0 + \frac{\pi^2 \varepsilon_s V_0^{\,2}}{2q^2 \lambda^2 N_a} - V_1 \tag{3.24}$$

This is valid under saturation, or $V_{ds} > V_{gs}$, where there is negligible mobile charge in the channel and the gate direct modulates the channel potential. For a given $V_{gs}$, $V_0$ is solved by the above quadratic equation:

$$V_0 = \frac{q^2 \lambda^2 N_a \left[ \sqrt{1 + 2\pi^2 \varepsilon_s (V_1 + qV_{gs}) / (q^2 \lambda^2 N_a)} - 1 \right]}{\pi^2 \varepsilon_s} \tag{3.25}$$

An example of the analytically calculated band from the source to channel is shown in the inset to Fig. 3.13, compared with TCAD simulations.

The region of band-to-band tunneling is magnified in Fig. 3.11(c). Consider tunneling at an energy $-E$ ($E > 0$) in the valence band of source. For electron energies lying within the staggered bandgap, i.e., $\Delta - V_1 < E < \Delta$, the process consists of hole tunneling ($m_h$) to the left of the heterojunction and electron tunneling ($m_e$) to the right of the heterojunction. The total tunneling probability is given by

$$T(E) = \exp\left\{ -\frac{2\sqrt{2}}{\hbar} \left[ \sqrt{m_h} \int_{l_1}^{0} \sqrt{-E - U(x)}\, dx + \sqrt{m_e} \int_{0}^{l_2} \sqrt{V(x) + E}\, dx \right] \right\} \tag{3.26}$$

where $E + U(l_1) = 0$ ($l_1 < 0$), and $E + V(l_2) = 0$. With $V(x)$ of (3.19) and $U(x)$ of (3.23), both integrals can be carried out analytically:

$$\int_{l_1}^{0}\sqrt{-E-U(x)}dx=\sqrt{\frac{\varepsilon_s}{2q^2N_a}}\left[\sqrt{\Delta(\Delta-E)}-E\ln\left(\sqrt{\frac{\Delta}{E}}+\sqrt{\frac{\Delta-E}{E}}\right)\right] \tag{3.27}$$

and

$$\int_{0}^{l_2}\sqrt{V(x)+E}dx=\frac{2\lambda}{\pi}\left[\sqrt{E+V_1-\Delta}-\sqrt{V_0+\Delta-E-V_1}\sin^{-1}\sqrt{\frac{E+V_1-\Delta}{V_0}}\right] \tag{3.28}$$

For $E > \Delta$, there is only electron tunneling given by (3.28). For $E < \Delta - V_1$ (if $\Delta > V_1$), there is only hole tunneling given by (3.27). With the bandgap of both the source and the channel much wider than $V_1$, the two-band $E(k)$ effects (to be discussed in Chapter 5) can be neglected.

With the analytically solved $T(E)$, the current density of a ballistic TFET with 3D density of states is calculated from [3.19]

$$j=\frac{qm}{2\pi^2\hbar^3}\int_0^{V_0+\Delta-V_1}(f_s-f_d)\left[\int_0^E T(E,E_\perp)dE_\perp\right]dE \tag{3.29}$$

where

$$T(E,E_\perp)=\exp\left\{-\frac{2\sqrt{2}}{\hbar}\left[\sqrt{m_h}\int_{l_1}^{0}\sqrt{-U(x)-(E-E_\perp)}dx+\sqrt{m_e}\int_0^{l_2}\sqrt{V(x)+E+E_\perp}dx\right]\right\} \tag{3.30}$$

and $f_s$, $f_d$ are the source and drain occupation factors with Fermi energies $-d_1$ and $-d_1 - qV_{ds}$, in terms of the source degeneracy $d_1$ and the drain voltage $V_{ds}$. The upper limit of integration in (3.29) is for the saturation region where the tunneling window is bounded by the channel conduction band (Fig. 3.11(b)).

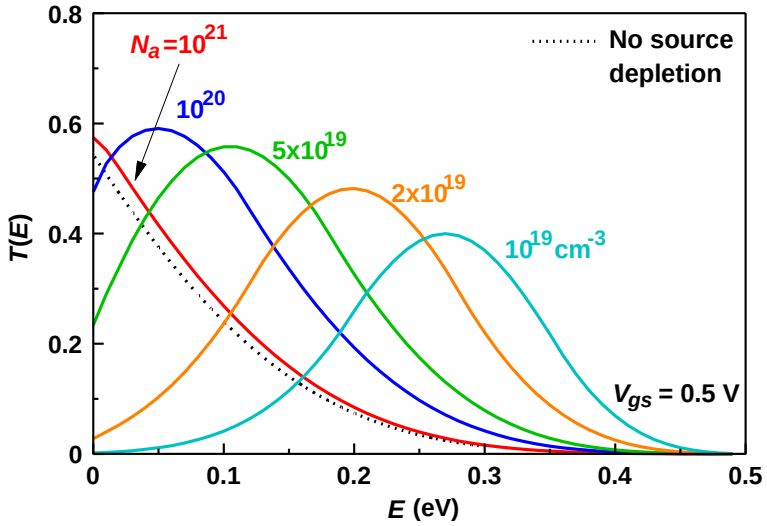



Figure 3.12: Tunneling probability versus energy from the analytic model for a heterojunction TFET biased at $V_{gs}$ = 0.5 V. The parameters are $V_1$ = 0.15 eV, $m_e = m_h$ = $0.1m_0$, $\lambda$ = 9 nm, and $\varepsilon_s$ = 11.7.

An example of $T(E)$ [or $T(E, E_\perp = 0)$] is shown in Fig. 3.12 for a range of source doping levels. For $N_a = 10^{21}$ cm$^{-3}$ or higher, source depletion is negligible. $T(E)$ is all due to electron tunneling of barrier $V_1$, the same as that from the no source depletion model [3.20]. For $N_a$ between $10^{20}$ and $10^{19}$ cm$^{-3}$, however, $T(E)$ exhibits a peak at $E > 0$, i.e., at an energy $-E$, below the valence band of source. This is because hole tunneling comes into play when the source depletion is significant. As can be seen from Fig. 3.11(c), the probability of hole tunneling with respect to the valence band barrier increases with $E$, in contrary to electron tunneling. The total $T(E)$ therefore first increases then decreases with $E$. Fig. 3.12 also shows that some degree of source depletion can help, as the total area under the $T(E)$ curve for $N_a \sim$ high $10^{19}$ cm$^{-3}$ is significantly larger than that of no depletion.

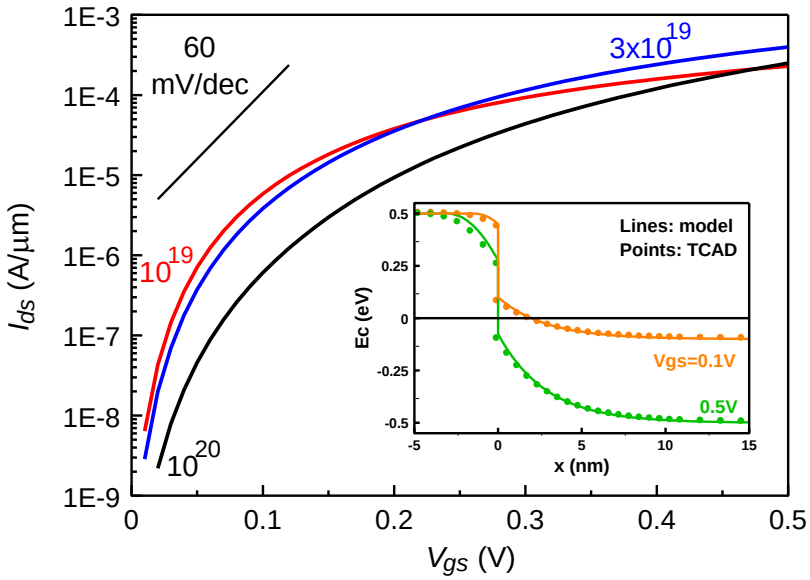



Figure 3.13: $I_{ds}$-$V_{gs}$ characteristics from the analytic model for a range of source doping assuming $N_v = 2\times10^{19}$ cm$^{-3}$. $V_{ds}$ = 0.5 V. The rest of the parameters are the same as in Fig. 3.12. The inset shows an example of the model calculated conduction band energy from the source to the channel, validated by Sentaurus simulation. $N_a = 3.2\times10^{19}$ cm$^{-3}$.

Fig. 3.13 shows the $I_{ds}$-$V_{gs}$ characteristics of an example of DG TFET generated from the analytic model for $V_{ds}$ = 0.5 V (saturation). The effective density of states of the source valence band is fixed at $N_v = 2\times10^{19}$ cm$^{-3}$ as $N_a$ is varied. The source degeneracy $d_1$ is calculated for each $N_a$ using Fermi integrals, namely, from $F_{1/2}(d_1/kT) = (\pi^{1/2}/2)(N_a/N_v)$. By earlier definition, $I_{ds} = 0$ when $V_{gs} = 0$ where $V_0 + \Delta = V_1$. The gate work function is allowed to vary for each $N_a$ to maintain this condition. The current rises up more sharply for lighter source doping with a smaller $d_1$. The on current at $V_{gs}$ = 0.5 V, however, is highest at some intermediate doping between $10^{19}$ and $10^{20}$ cm$^{-3}$. This is more clearly shown in Fig. 3.14 where the on current for $N_v = 2\times10^{19}$ cm$^{-3}$ peaks at a source doping of $N_a = 3.5\times10^{19}$ cm$^{-3}$.

Also shown is the Sentaurus [3.2] validation of the model result. For materials with lower $N_v$, the degeneracy factor increases. The peak current decreases and shifts to a lighter $N_a$.

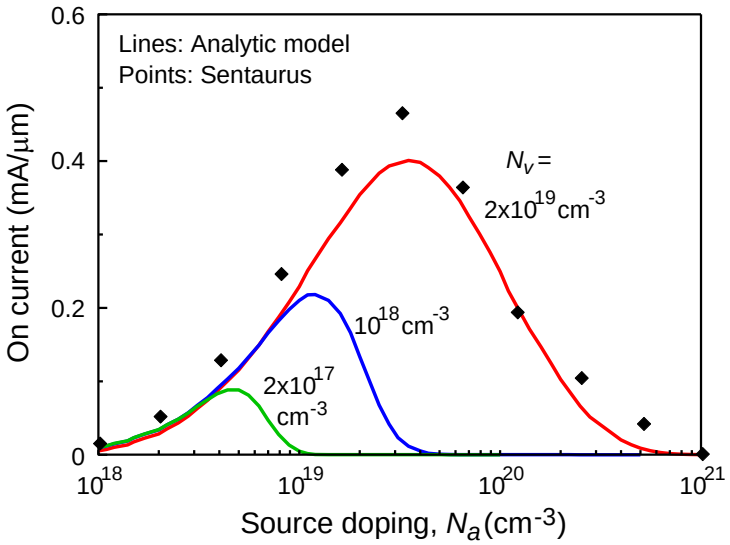



Figure 3.14: On current (at $V_{gs} = V_{ds} = 0.5$ V) of heterojunction TFETs versus source doping concentration for three values of $N_v$ (effective density of states). The rest of the parameters are the same as in Fig. 3.13. $N_v$ is set to $2\times10^{19}$ cm$^{-3}$ for the Sentaurus simulation.

A different value of $V_1$ will change the magnitude of peak current as expected, but not the $N_a$ value where the current peaks. If $m_h$ is changed to $10m_e$ with the same $m_e$ (= $0.1m_0$), the current peak (for $N_v = 2\times10^{19}$ cm$^{-3}$) becomes ~40% lower and shifts to a higher $N_a$ ($7\times10^{19}$ cm$^{-3}$).

### 3.4.2 Fermi-Dirac integral (no depletion approximation)

Poisson's equation can be written as:

$$\frac{d^2E_v}{dx^2} = \frac{q^2}{\varepsilon_s}\left[p - N_a\right] \tag{3.31}$$

where $E_v$ is the source valence band energy, $p$ is the hole concentration in the source, and $N_a$ is the density of ionized dopants. In section 3.4.1, depletion approximation was employed by assuming $p - N_a$ as a step function, thus $p - N_a = -N_a$ as soon as $E_v < 0$. Strictly, the $p$ needs to be calculated from the fermi integral:

$$p = \frac{2}{\sqrt{\pi}} N_v F_{1/2}\left(\frac{E_v - E_f}{kT}\right) \tag{3.32}$$

where the $E_f$ is solved from the Fermi integrals $F_{1/2}(-E_f/kT) = \pi^{1/2}N_a/(2N_v)$, in terms of the source doping $N_a$, and the effective density of states $N_v$.

By multiplying $dE_v/dx$ on both sides, Eq. (3.31) can be re-written as:

$$\frac{1}{2}\frac{d}{dx}\left(\frac{dE_v}{dx}\right)^2 = \frac{q^2}{\varepsilon_s}\frac{d}{dx}\int\left[p - N_a\right]dE_v \tag{3.33}$$

Integrating Eq. (3.33) from $-\infty$ to $x$ yields:

$$\left(\frac{dE_v}{dx}\right)^2 = \frac{2q^2}{\varepsilon_s}\int_0^{E_v}\left[p - N_a\right]dE_v \tag{3.34}$$

The Eq. (3.34) is numerically evaluated to solve for $E_v(x)$, with $E_v = -\Delta$ at $x = 0$.

Fig. 3.15 shows that the model curve from Eq.(3.34) is consistent with Sentaurus simulations.

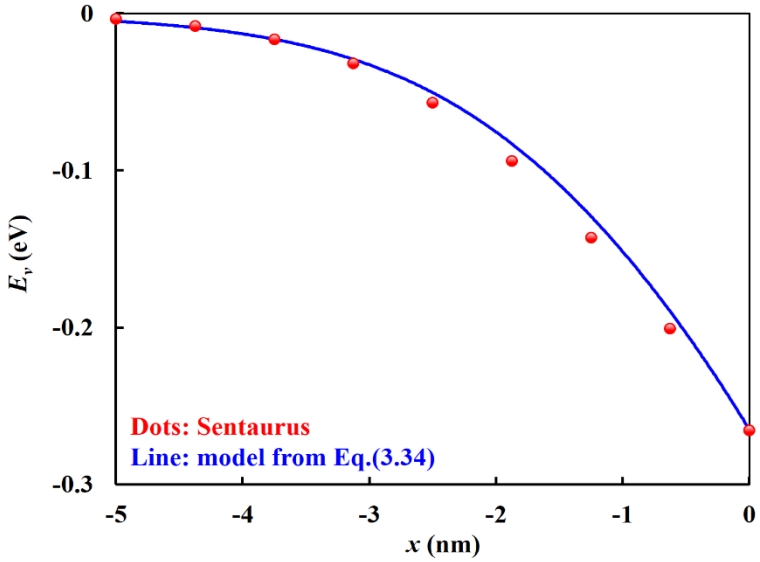



Figure 3.15: Valence band energy from bulk source to the source-to-channel junction with $N_a = 3\times10^{19}$ cm$^{-3}$, $N_v = 2\times10^{19}$ cm$^{-3}$, $V_{gs}$ = 0.5V, $V_{ds}$ = 0.1V. The dots are from Sentaurus simulations extracted at the center of the film.

Fig. 3.16 shows the contrast between the depletion approximation from Eq. (3.23) and the Fermi-Dirac integral from Eq. (3.34). It is clear that the depletion approximation slightly underestimates the depletion width. Yet, the depletion approximation still gives reasonable accuracy and is very efficient for computation. Hence it is employed throughout the rest of this work.

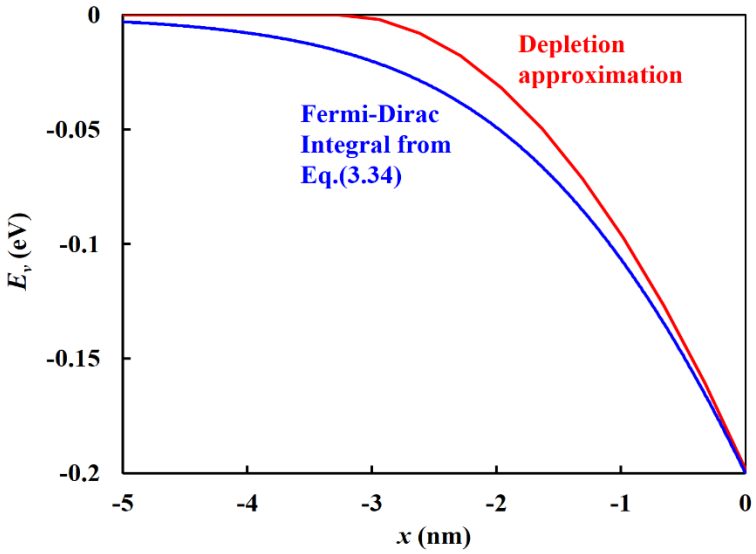



Figure 3.16: Source valence band energy plot generated from depletion approximation versus Fermi-Dirac integral. The rest of the parameters are the same as in Fig. 3.15.

The text of Chapter 3, in part, is a reprint of the material as it appears in "Dimensionality dependence of TFET performance down to 0.1 V supply voltage" by Yuan Taur, Jianzhi Wu and Jie Min, IEEE Transaction on Electron Devices, Feb. 2016. The dissertation author was a co-author of this paper.

The text of Chapter 3, in part, is a reprint of the material as it appears in "Analysis of source doping effect in tunnel FETs with staggered bandgap" by Jie Min, Jianzhi Wu and Yuan Taur, IEEE Electron Device Letters, Oct. 2015. The dissertation author was a co-author of this paper.

The text of Chapter 3, in part, is a reprint of the material as it appears in "An analytic model for heterojunction tunnel FETs with exponential barrier" by Yuan Taur, Jianzhi Wu and Jie Min, IEEE Transaction on Electron Devices, May 2015. The dissertation author was a co-author of this paper.

## 3.5 References

[3.1] J. Knoch and J. Appenzeller, “Modeling of high performance p-type III-V heterojunction tunnel FETs,” *IEEE Electron Device Letters*, Vol. 31, pp. 305-307, April 2010.

[3.2] *TCAD Sentaurus Device Manual* (Synopsys, Inc., Mountain View, CA, 2010).

[3.3] P. M. Solomon, D. J. Frank and S. O. Koswatta, “Compact model and performance estimation for tunneling nanowire FET,” in *Device research conference (DRC)*, 2011 69th annual; 2011. p. 197–8. doi:10.1109/DRC.2011.5994495.

[3.4] B. Rajamohanan, D. Mohata, A. Ali, and S. Datta, "Insight into the output characteristics of III-V tunneling field effect transistors," *Applied Physics Letters* 102, no. 9 (2013): 092105.

[3.5] Y. Taur, X. Liang, W. Wang, and H. Lu, “A continuous, analytic drain-current model for DG-MOSFETs,” *IEEE Electron Device Letters*, Vol. 25, No. 2, pp. 107-109, Feb. 2004.

[3.6] O. M. Nayfeh, Judy L. Hoyt, and D. A. Antoniadis. “Strained-Band-to-Band Tunneling Transistors: Impact of Tunnel-Junction Germanium Composition and Doping Concentration on Switching Behavior." *IEEE Trans. Electron Devices*, vol. 56, no. 10, pp. 2264–2269, Sep. 2009.

[3.7] A. Vandooren, D. Leonelli, R. Rooyackers, A. Hikavyy, K. Devriendt, M. Demand, R. Loo, G. Groeseneken, and C. Huyghebaert. "Analysis of trap-assisted tunneling in vertical Si homo-junction and SiGe hetero-junction Tunnel-FETs." *Solid-State Electronics*, vol. 83, pp. 50-55, May 2013.

[3.8] G. Zhou, Y. Lu, R. Li, Q. Zhang, Q. Liu, T. Vasen, H. Zhu, J. M. Kuo, T. Kosel, M. Wistey, P. Fay, A. Seabaugh and H. Xing. "InGaAs/InP Tunnel FETs With a Subthreshold Swing of 93 mV/dec and Ratio Near." *IEEE Electron Device Letters*, Vol. 33, No. 6, pp. 782-784, Jun. 2012.

[3.9] R. Li, Y. Lu, G. Zhou, Q. Liu, S. D. Chae, T. Vasen, W. S. Hwang， Q. Zhang, P. Fay, T. Kosel, M. Wistey, H. Xing and A. Seabaugh. "AlGaSb/InAs tunnel field-effect transistor with on-current of 78 at 0.5 V." *IEEE Electron Device Letters*, Vol. 33, No. 3, pp. 363-365, Mar. 2012.

[3.10] D. Jena, "Tunneling transistors based on graphene and 2-D crystals," *Proceedings of the IEEE* 101.7 (2013): 1585-1601.

[3.11] J. Wu, J. Min, J. Ji, and Y. Taur. "An analytic model for heterojunction and homojunction tunnel FETs with 3D density of states." In *Device Research Conference (DRC), 2015 73rd Annual*, pp. 249-250. IEEE, 2015.

[3.12] Y. Taur and T. H. Ning, *Fundamentals of modern VLSI devices*, Cambridge, U.K.: Cambridge University Press, 2009, p.592.

[3.13] T. Fang, A. Konar, H. Xing, and D. Jena, "Carrier statistics and quantum capacitance of graphene sheets and ribbons," *Appl. Phys. Lett.*, 092109, 2007.

[3.14] A. C. Seabaugh and Q. Zhang, "Low-voltage tunnel transistors for beyond CMOS logic," *Proceedings of the IEEE* 98.12 (2010): 2095-2110.

[3.15] C. Sandow, J. Knoch, C. Urban, Q.-T. Zhao and S. Mantl "Impact of electrostatics and doping concentration on the performance of silicon tunnel field-effect transistors", *Solid-State Electronics*, vol.53, (2009): pp. 1126-1129

[3.16] J. Appenzeller, J. Knoch, M.T. Björk, H. Riel, H. Schmid and W. Riess "Toward Nanowire Electronics", *IEEE Trans. Electron Devices*, vol.55, no.11, (2008): pp. 2827-2845

[3.17] L. De Michielis, L. Lattanzio, and A. M. Ionescu, "Understanding the superlinear onset of tunnel-FET output characteristic." *Electron Device Letters, IEEE* 33, no. 11 (2012): 1523-1525.

[3.18] B. Yu, L. Wang, Y. Yuan, P. M. Asbeck, and Y. Taur, "Scaling of nanowire transistors," *IEEE Trans. Electron Devices*, vol. 55, pp. 2846-2858, Nov. 2008.

[3.19] E. O. Kane, "Theory of tunneling," *J. Applied Physics*, vol. 32, pp. 83-91 (1961).

[3.20] Y. Taur, J. Wu and J. Min "An analytic model for heterojunction tunnel FETs with exponential barrier", *IEEE Trans. Electron Devices*, vol.62, no.5, pp. 1399-1404, May 2015.

[3.21] S. O. Koswatta, M. S. Lundstrom and D. E. Nikonov, "Performance comparison between p-i-n tunneling transistors and conventional MOSFETs", *IEEE Trans. Electron Devices*, vol. 56, no. 3, pp. 456-465, 2009

# Chapter 4 Modeling of OFF-current

## 4.1 Dual sinh potential

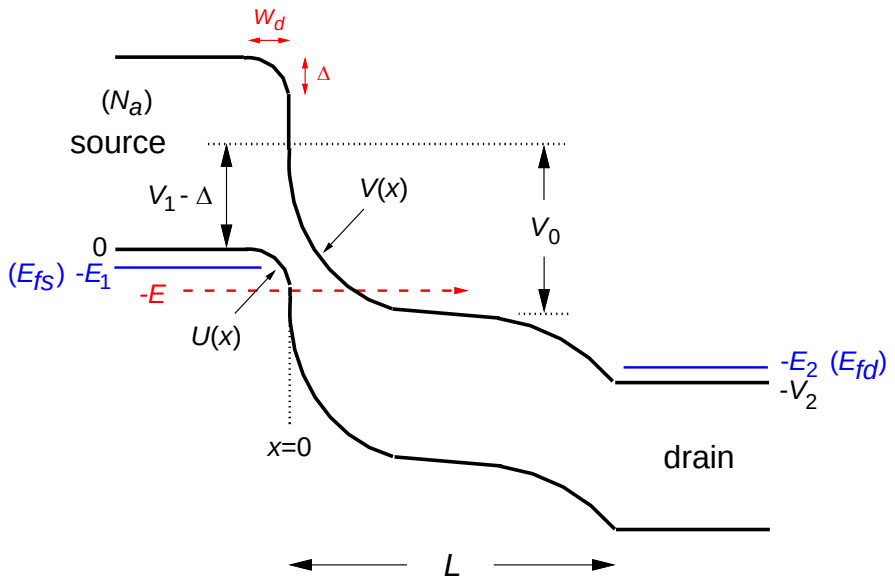



Figure 4.1: Band diagram of a heterojunction TFET with $p^+$ source and $n^+$ drain

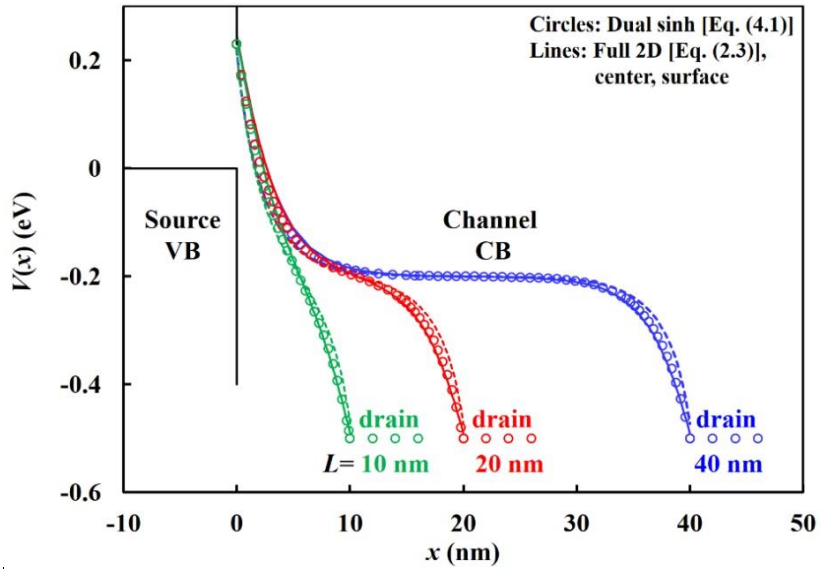
Fig. 4.1 shows the band diagram of a staggered heterojunction double gate (DG) TFET. The zero energy reference is taken to be the valence band edge of the bulk source region far from the heterointerface. The diagram depicts the situation of a turned-on TFET biased in saturation. $V_1$ is the effective energy barrier between the conduction band of channel and the valence band of source at the heterojunction boundary ($x$ = 0). Δ and $W_d$ are the band bending and the depletion width in the source region doped at a density $N_a$. $V_0$ represents the gate control of the channel conduction band that determines the tunneling window. $V_2$ is related to the drain bias. $L$ is the channel length. The conduction band function $V(x)$ holds the key to the TFET current.

For the function $V(x)$ in the short-channel TFET model in Fig. 4.1, we approximate

the full 2D potential solution of Eq. (2.3), by a combination of only the $n = 1$ *sinh* factors and the constant term,

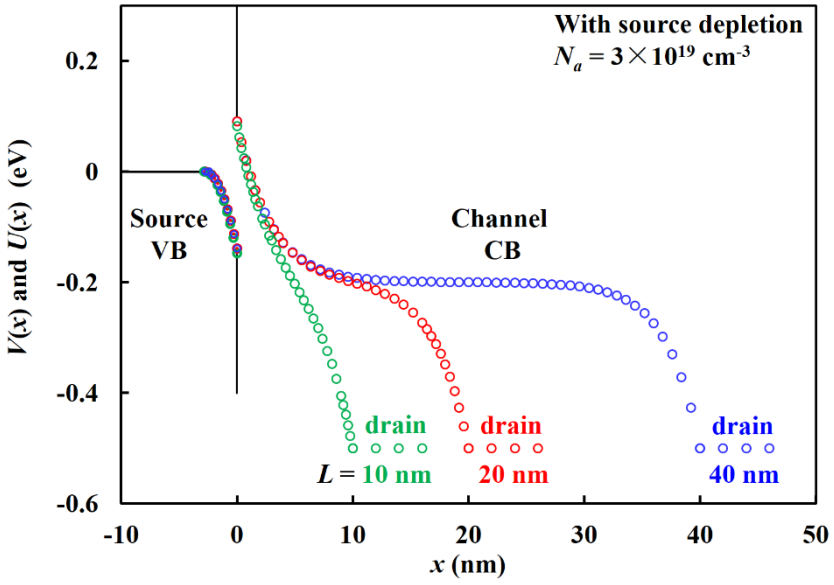
$$V(x) = V_0 \frac{\sinh[\pi(L-x)/\lambda]}{\sinh(\pi L/\lambda)} - (V_0 - V_1 + \Delta) - (V_2 - V_0 + V_1 - \Delta)\frac{\sinh(\pi x/\lambda)}{\sinh(\pi L/\lambda)} \qquad (4.1)$$

Note that the coefficients, different from $b_1$ and $c_1$ in Eq. (2.3), are chosen to satisfy the depth independent source and drain boundary conditions in Fig. 4.1, $V(0) = V_1 - \Delta$ and $V(L) = -V_2$. The three terms of $V(x)$ in Eq. (4.1) represent the effects of source, gate, and drain on the channel potential. We choose a material system of AlGaAsSb source and InGaAs channel that gives an effective bandgap of $V_1 = 0.23$ eV. Fig. 4.2(a) compares Eqs. (2.3) and (4.1) for the example of $\lambda = 9$ nm ($t_s = 5$ nm, $t_i = 2$ nm) and $\Delta = 0$ for $L = 40$, 20, and 10 nm. It shows that the dual *sinh* function of Eq. (4.1) adequately captures the channel length dependence of the full 2D potential solution. The 10 nm potential curve reveals a thinning of the source barrier by the proximity of the drain—an indication of short-channel effect.




(a)

Figure 4.2: (a) Channel potential without source depletion. Circles are the dual *sinh* function of Eq. (4.1) with $\Delta = 0$, $V_1 = 0.23$ eV, $V_0 = 0.43$ eV, $V_2 = 0.5$ eV, and $\lambda = 9$ nm. Solid lines are $-q\psi(x, 0)$ and dashed lines $-q\psi(x, t_s/2)$, both from Eq. (2.3) with $n = 1, 3, 5, 7$ terms.




(b)

Figure 4.2: (b) Channel and source potential with source depletion of doping $N_a = 3\times10^{19}$ cm$^{-3}$. Circles are from Eqs. (4.1) and (4.4) coupled by Eq. (4.2). The same $V_1$, $V_2$, and $V_0 + \Delta$ (= 0.43 eV) as in (a) are assumed.

By applying the condition that the field is continuous from one side of the heterojunction to the other (assuming no change of permittivity), we get

$$\frac{1}{q}\left|\frac{dV}{dx}\right|_{x=0} = \left(\frac{\pi}{q\lambda}\right)\frac{V_0\cosh(\pi L/\lambda)+(V_2-V_0+V_1-\Delta)}{\sinh(\pi L/\lambda)} = \frac{2\Delta}{qW_d} = \sqrt{\frac{2N_a\Delta}{\varepsilon_s}} \tag{4.2}$$

where we employed the depletion approximation for the source, $W_d = [2\varepsilon_s\Delta/(q^2N_a)]^{1/2}$. $V_2$ is determined by the drain voltage, $V_{ds}$. Note in Fig. 4.1 that the tunneling window starts to open when $V_0 = V_1 - \Delta$. If we define this condition to be $V_{gs} = 0$, then $qV_{gs} = V_0 - (V_1 - \Delta)$. This is valid under saturation conditions, or $V_{ds} > V_{gs}$, when the mobile charge in the channel is negligible so the gate directly modulates the channel potential. Implicitly assumed is a choice of the gate work function dependent on the band offset, the barrier height $V_1$ at the heterojunction, and the source doping. Substituting the above relation in Eq. (4.2) allows $V_0$ to be solved for given $V_{gs}$:

$$\left(\frac{\pi}{q\lambda}\right)\frac{V_0\cosh(\pi L/\lambda)+V_2-qV_{gs}}{\sinh(\pi L/\lambda)} = \sqrt{\frac{2N_a(qV_{gs}+V_1-V_0)}{\varepsilon_s}} \tag{4.3}$$

With that, $\Delta$ and $W_d$ are also determined and the valence band function in the source depletion region in Fig. 4.1 is

$$U(x) = -\frac{q^2N_a}{2\varepsilon_s}(x+W_d)^2 \tag{4.4}$$

Fig. 4.2(b) shows the calculated $V(x)$ and $U(x)$ with a source doping $N_a$ of $3\times10^{19}$ cm$^{-3}$, for the same bias conditions and channel lengths as those in Fig. 4.2(a).

## 4.2 Source-to-drain tunneling

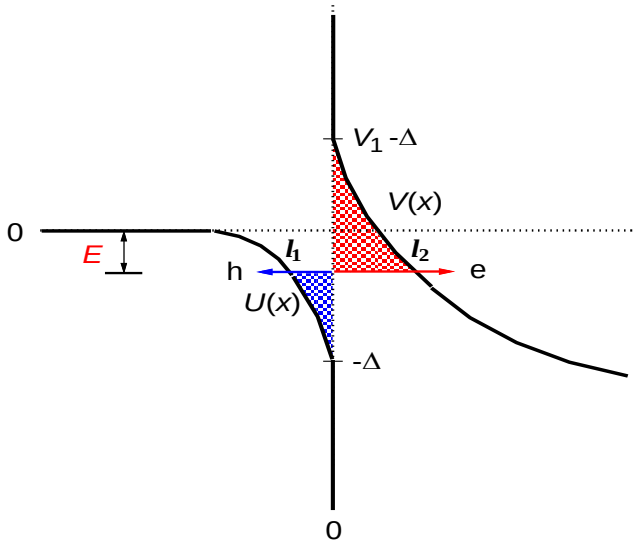



Figure 4.3: A magnified view of the tunneling region. The shaded areas depict the conduction band and valence band barriers for electron and hole tunneling respectively.

The region of band-to-band tunneling is magnified in Fig. 4.3. Considering tunneling at an energy $-E$ ($E > 0$) in the valence band of source, for electron energies above $-\Delta$, the process consists of hole tunneling to the left of the heterojunction and electron tunneling to the right of the heterojunction. Assuming $m_e = m_h = m$, the tunneling probability given by Eq. (3.26) can be expressed as:

$$T(E) = \exp\left\{-\frac{2\sqrt{2m}}{\hbar}\left[\int_{l_1}^{0}\sqrt{-E-U(x)}\,dx + \int_{0}^{l_2}\sqrt{V(x)+E}\,dx\right]\right\} \tag{4.5}$$

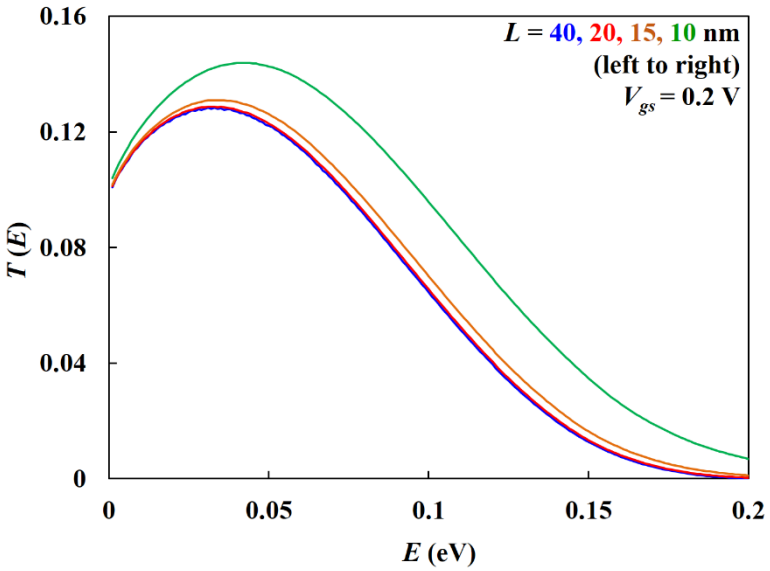
where $E + U(l_1) = 0$ ($l_1 < 0$), and $E + V(l_2) = 0$. With the $U(x)$ of (4.4), the first integral can be carried out analytically [4.1]:

$$\int_{l_1}^{0}\sqrt{-E-U(x)}\,dx = \sqrt{\frac{\varepsilon_s}{2q^2N_a}}\left[\sqrt{\Delta(\Delta-E)} - E\ln\left(\sqrt{\frac{\Delta}{E}}+\sqrt{\frac{\Delta-E}{E}}\right)\right] \tag{4.6}$$

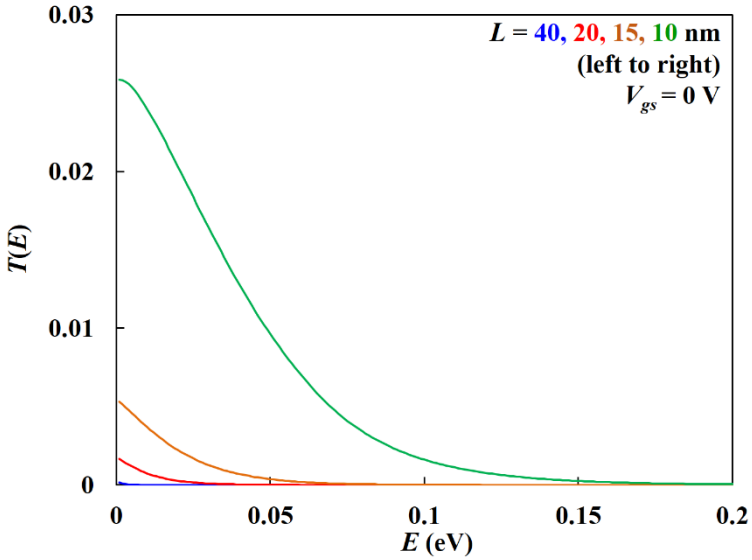
The second integral needs to be evaluated numerically with $V(x)$ of (4.1). For electron energies below $-\Delta$, there is no tunneling in the source and the first integral of Eq. (4.5) goes away.

The barrier heights in Eq. (4.5) and Fig. 4.3 are justified by the assumption that both the source and the channel bandgaps are wide enough to leave the $E_c$ of source and the $E_v$ of channel out of the picture. In such cases, both the one-band [4.2] and the two-band [4.3] $E(k)$ models give the same tunneling barriers as in Eq. (4.5).

Fig. 4.4 shows an example of $T(E)$ calculated for several different $L$, with a source doping of $N_a = 3\times10^{19}$ cm$^{-3}$. The tunneling window in Fig. 4.1 covers an energy range $0 \le E \le V_2$. The total flux of tunneling is proportional to the area under $T(E)$. When $qV_{gs} = V_0 - (V_1 - \Delta) = 0.2$ eV in Fig. 4.4(a), the conduction band of channel is well within the tunneling window and the area is only slightly sensitive to the channel length. Note that $T(E)$ tends to peak at an energy $E = \Delta - V_1/2$, where the electron and hole tunneling barriers are about equal. However, when $V_{gs} = 0$ in Fig. 4.4(b), $T(E = 0)$ and beyond consist only of electron tunneling from the source to the drain. The area under $T(E)$ is very sensitive to $L$.




(a)




(b)

Figure 4.4: Tunneling probability versus (–) carrier energy for different channel lengths. The window for source to channel tunneling is 0.2 eV for (a) and 0 for (b). Other parameters are: $N_a$=3×10$^{19}$ cm$^{-3}$, $V_1$ = 0.23 eV, $V_2$ = 0.5 eV, $\lambda$ = 9 nm, and $m$ = 0.1$m_0$.

## 4.3 Channel length dependence

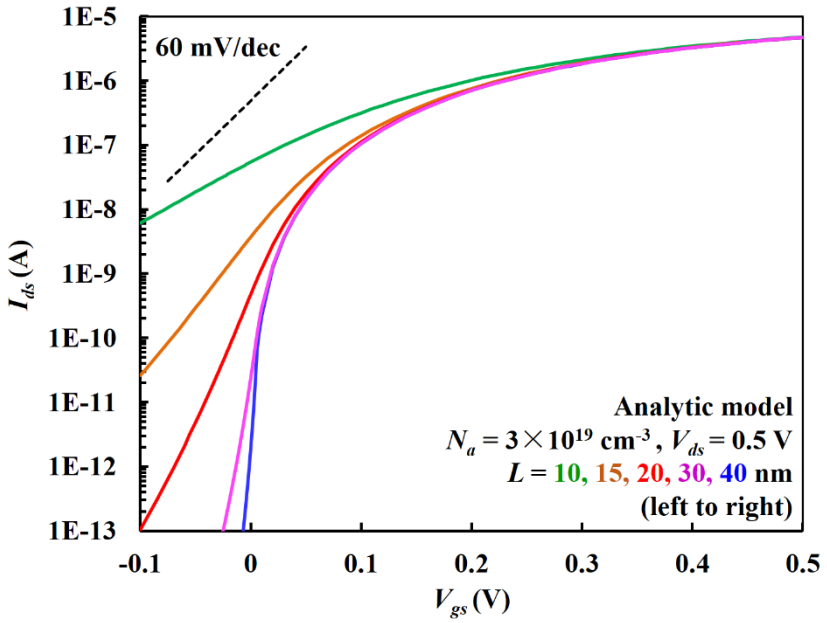
The current for a 1D ballistic TFET is given by the Landauer equation [4.4],

$$I_{ds} = \frac{2q}{h}\int_0^{V_2}(f_s - f_d)T(E)dE$$
$$= \frac{2q}{h}\int_0^{V_2}\left[\frac{1}{1+e^{(d_1-E)/kT}} - \frac{1}{1+e^{(d_1+qV_{ds}-E)/kT}}\right]T(E)dE \quad (4.7)$$

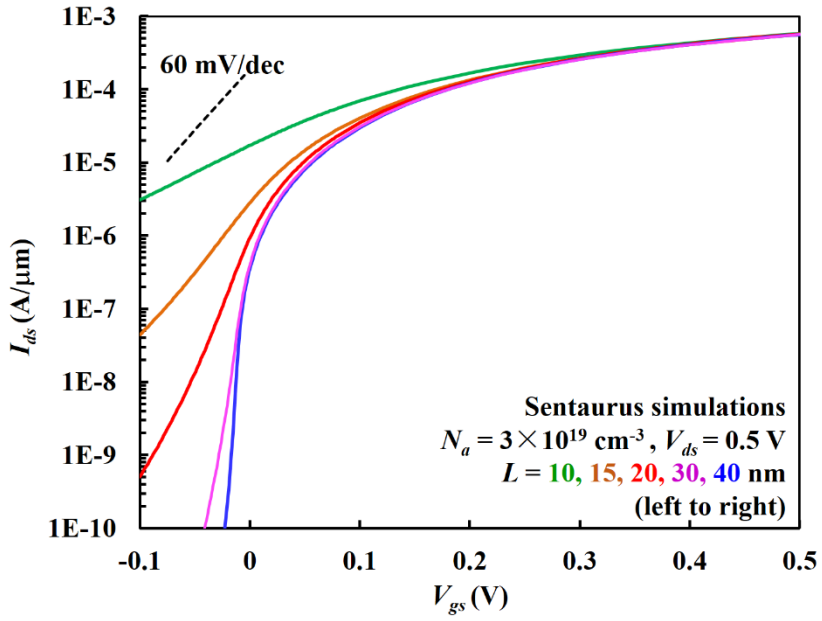
where $T(E)$ is given by Eq. (4.5), $d_1 = -E_{fs}$ is the source degeneracy. The current integral covers the source-to-channel tunneling for $E \in (0, V_0 - V_1 + \Delta)$ and the source-to-drain tunneling for $E \in (V_0 - V_1 + \Delta, V_2)$. The parameter $V_2$ in Eq. (4.1), which is also the upper bound of the current integral, can be expressed as $V_2 = qV_{ds} + d_1 + d_2$ where $d_2$ is the drain degeneracy.

$I_{ds}$-$V_{gs}$ characteristics generated by Eq. (4.7) with $V_{ds} = 0.5$ V are plotted in Fig. 4.5(a) for several different $L$. The same source doping, $N_a = 3\times10^{19}$ cm$^{-3}$, as in Fig. 4.4 is assumed. $d_1$ is calculated to be 0.024 eV from the Fermi integral $F_{1/2}(d_1/kT) = (\pi^{1/2}/2)(N_a/N_v)$ with an effective density of states $N_v = 2\times10^{19}$ cm$^{-3}$. For long channel TFETs, the current is not sensitive to $L$. Below $L \sim 2\lambda = 18$ nm, however, both the subthreshold slope and the off current, $I_{ds}(V_{gs} = 0)$, degrade rapidly. This is more clearly shown in Fig. 4.6 by plotting $I_{ds}(V_{gs} = 0)$ and $I_{ds}(V_{gs} = 60$ mV) versus $L$. $I_{ds}(V_{gs} = 0)$ comes only from source-to-drain tunneling, which keeps on increasing toward shorter channel lengths. $I_{ds}(V_{gs} = 60$ mV) consists of both source-to-channel tunneling and source-to-drain tunneling but is dominated by the former. It is insensitive to $L$ until $L \leq 2\lambda$ where thinning of the source-to-channel barrier sets in. The subthreshold current slope at any given $L$ can be read from the ratio of the two currents in Fig. 4.6. Below $L \sim 1.5\lambda$, the subthreshold swing can no longer beat 60 mV/decade, the $kT/q$ limit.

Fig. 4.5(b) shows the $I_{ds}$-$V_{gs}$ curves from Sentaurus simulations [4.5] under a similar set of parameters. They generally agree with the model results in Fig. 4.5(a).




(a)




(b)

Figure 4.5: (a) Model generated high drain-bias $I_{ds}$-$V_{gs}$ characteristics for different values of $L$. Parameters $N_a$, $V_1$, $\lambda$, $m$ are the same as Fig. 4.4. $d_2$ = 0. (b) Sentaurus simulations with the same set of parameters.

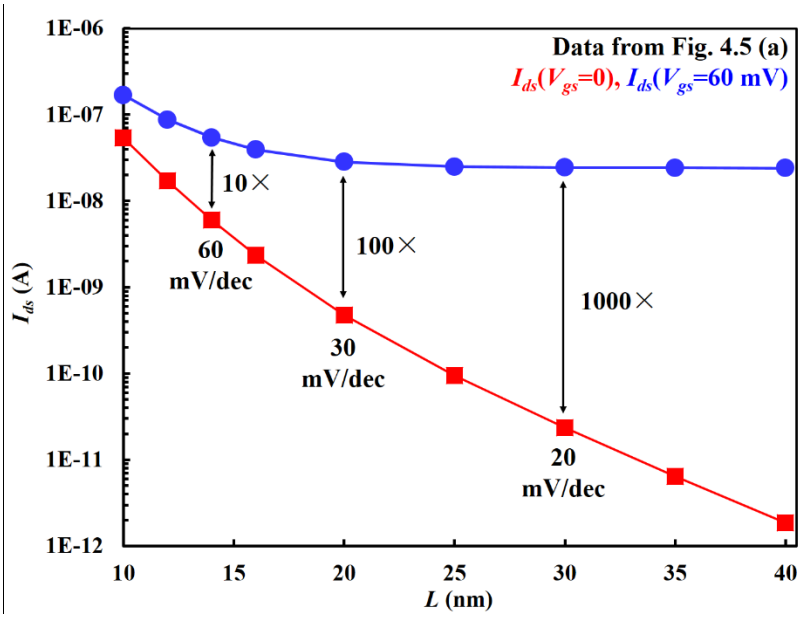



Figure 4.6: Currents at $V_{gs}$ = 0 and 60 mV from the data of Fig. 4.5(a) versus $L$.

There is a distinct difference between the short channel effects (SCE) of TFETs and that of MOSFETs. In MOSFETs, the first-order effect of SCE is a lower threshold voltage which shifts the entire $I_{ds}$-$V_{gs}$ curve negatively in a parallel fashion that both the off- and the on-currents go up [4.6]. In TFETs, SCE mainly degrades the slope of $I_{ds}$-$V_{gs}$ near $V_{gs}$ = 0, hence the off-current goes up sharply, while the on-current at large $V_{gs}$ is hardly affected.

When the TFET is biased in the linear region, i.e., when $V_{ds} < V_{gs}$, the parameter $V_0$ is no longer solved by Eq. (4.3) as derived from $qV_{gs} = V_0 - (V_1 - \Delta)$. Fig. 4.7(a) shows that while the $V(x)$ curve of Eq. (4.1) for $V_{gs} = V_{ds} = 0.5$ V is consistent with that of Sentaurus simulations, the curve for $V_{gs} = 0.5$ V and $V_{ds} = 0.1$ V is far off. This is because of the de-biasing effect of inversion charge [4.7]-[4.9] when the Fermi level of the drain is close to or above the conduction band of channel. As far as the effect on channel potential

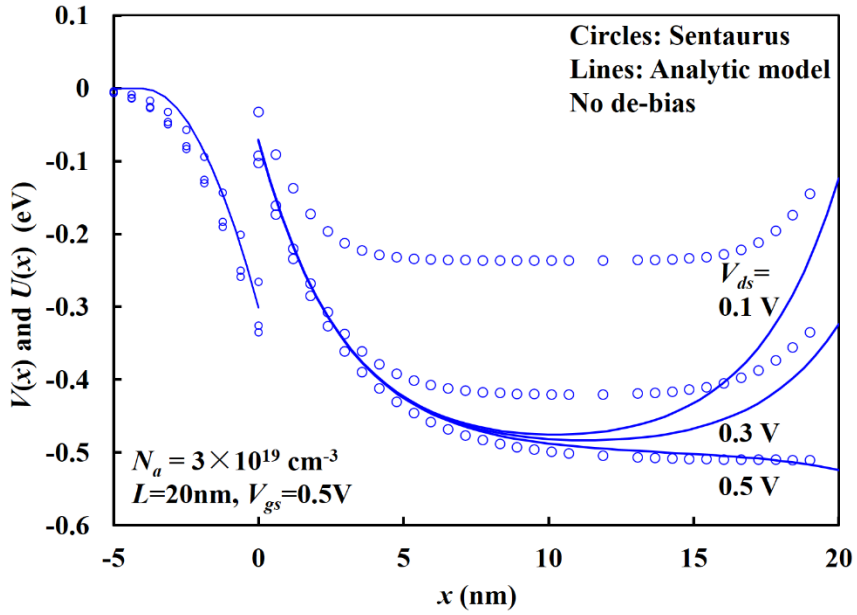
is concerned, $V_{gs}$ is degraded to $V_{gs} - Q_{inv}/C_{ox}$, where $Q_{inv}/C_{ox}$ is the potential drop across the gate insulator. For given $V_{gs}$ and $V_{ds}$, $Q_{inv}$ can be calculated from a continuous, analytic solution of Poisson's equation with mobile charge for DG MOSFETs [4.10]:

$$Q_{inv} = \frac{4kT\varepsilon_s}{qt_s}\beta\tan\beta \tag{4.8}$$

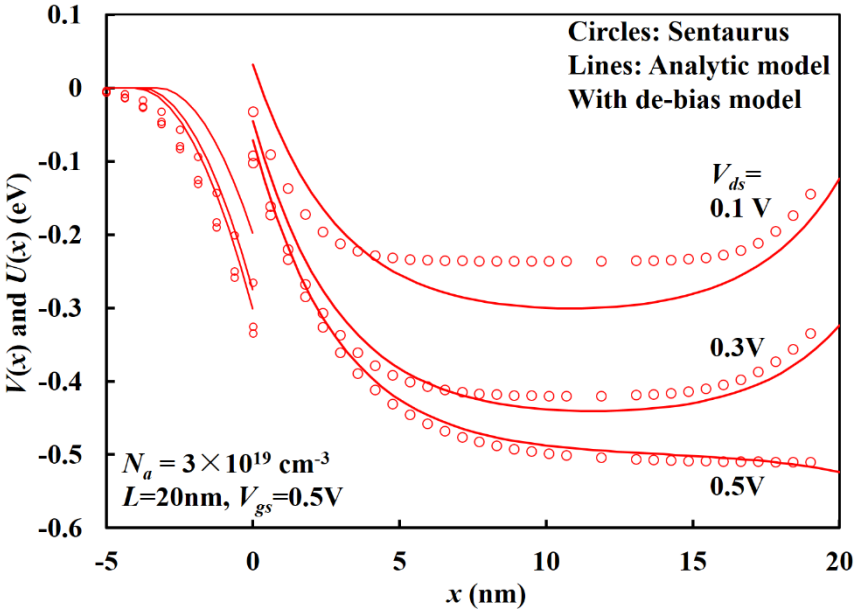
where the intermediary parameter $\beta$ is solved from

$$\frac{q(V_{gs} - V_{ds} - d_1)}{2kT} - \ln\left[\frac{2}{t_s}\sqrt{\frac{2\varepsilon_s kT}{q^2 N_c}}\right] = \ln\beta - \ln[\cos\beta] + \frac{2\varepsilon_s t_i}{\varepsilon_i t_s}\beta\tan\beta \tag{4.9}$$

Here, $N_c$ is the effective density of states of the conduction band. With $V_{gs}$ replaced by $V_{gs} - Q_{inv}/C_{ox}$ in Eq. (4.3), the de-biased $V_0$, $\Delta$, and $W_d$ are solved. $V(x)$ of Eq. (4.1) and $U(x)$ of Eq. (4.4) with the de-biased parameters are plotted in Fig. 4.7(b). Note that there is less source depletion in the $V_{ds}$ = 0.1 V case because of the lower field at the junction due to de-bias. The model curves are generally consistent with those of Sentaurus taken at the center of the semiconductor film. At the surface, the Sentaurus potential at mid-channel closely matches that of the analytic model. But there is more source depletion in the surface potential of Sentaurus due to the effect of gate fringe field not considered in the analytic model.




(a)

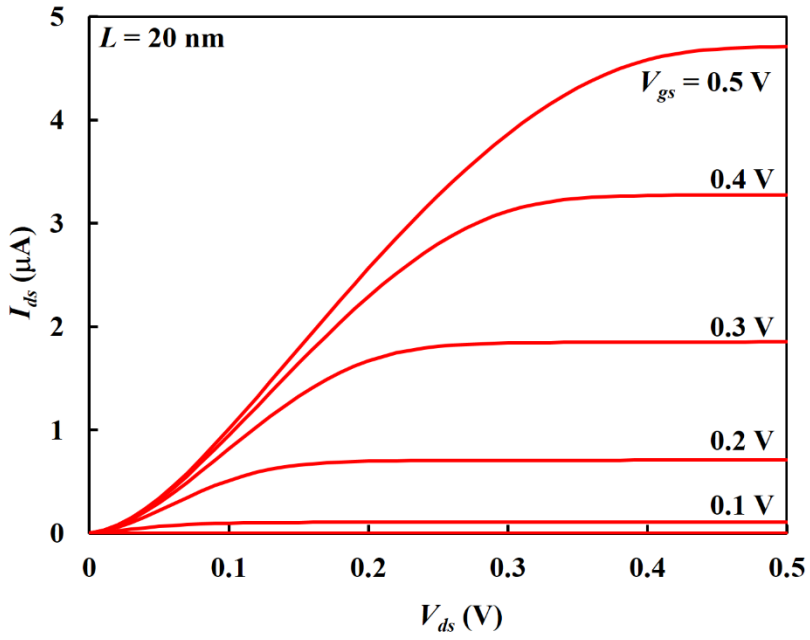



(b)

Figure 4.7: Conduction band energy of channel from source to drain and valence band energy of source for a 20 nm TFET. $V_{gs}$ = 0.5 V. Solid lines are calculated from (a) no de-bias and (b) de-bias model. $N_c$ = 8.7×10$^{16}$ cm$^{-3}$ (InGaAs) is assumed in Eq. (4.9). $d_2$ = 0. Circles are from Sentaurus simulations taken at the center of film [both (a) and (b)].

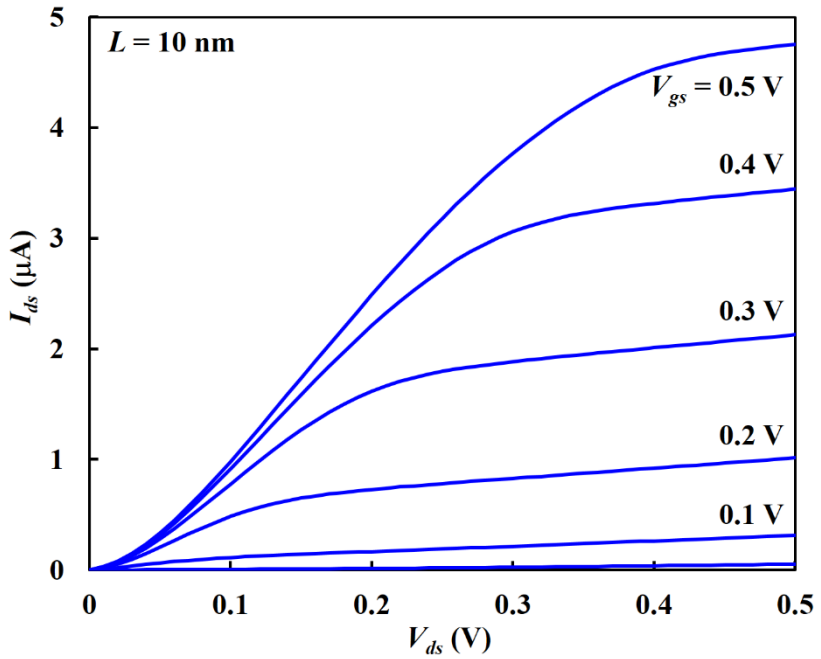
By using the de-biased $V(x)$ and $U(x)$ in $T(E)$ of the current integral, Eq. (4.7), continuous $I_{ds}$-$V_{ds}$ characteristics are generated for $L$ = 20 nm and 10 nm TFETs, as shown in Fig. 4.8. Similar to short-channel MOSFETs, finite output conductance appears in the saturation region of the 10 nm TFET. It is due to the drain effect on the source-to-channel barrier, as noted before with the $L$ = 10 nm curve (green) in Fig. 4.2. This effect does not become significant until $L \approx \lambda$.

With the de-bias model, low drain voltage (50 mV) $I_{ds}$-$V_{gs}$ characteristics are calculated and plotted in Fig. 4.9 along with high drain characteristics. The drain bias has no effect on the subthreshold current of a 40 nm TFET. For a 15 nm TFET, however, both the off current and the subthreshold slope are sensitive to the drain voltage. Unlike MOSFETs, there is no region of parallel shift between the high drain and low drain curves.




(a)

Figure 4.8: $I_{ds}$-$V_{ds}$ characteristics for (a) $L$ = 20 nm




(b)

Figure 4.8: continued (b) 10 nm TFETs generated by the analytic model with de-bias. $N_c$ and $d_2$ are the same as in Fig. 4.7.

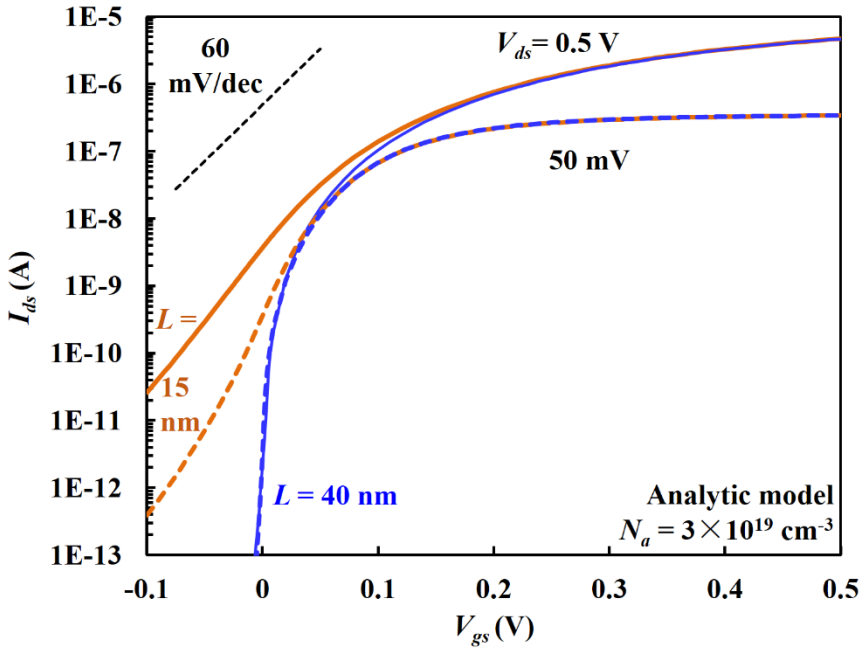



Figure 4.9: Model generated high-drain and low-drain $I_{ds}$-$V_{gs}$ characteristics for long channel (40 nm) and short channel (15 nm) TFETs

The text of Chapter 4, in part, is a reprint of the material as it appears in “Short channel effects in tunnel FETs” by Jianzhi Wu, Jie Min and Yuan Taur, *IEEE Transaction on Electron Devices,* Sept. 2015. The dissertation author was the primary investigator and author of this paper.

## 4.4 References

[4.1] J. Min, J. Wu and Y. Taur, "Analysis of source doping effect in tunnel FETs with staggered bandgap", *IEEE Electron Device Lett.*, vol. 36, no. 10, pp. 1094-1096, 2015.

[4.2] L. De Michielis, N. Dağtekin, 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* 103, no. 12 (2013): 123509.

[4.3] L. Liu, D. Mohata, and S. Datta, "Scaling length theory of double-gate interband tunnel field-effect transistors," *Electron Devices, IEEE Transactions on* 59 no.4 (2012): 902-908.

[4.4] M. Lundstrom (2000) *Fundamentals of Carrier Transport*, Cambridge University Press.

[4.5] *TCAD Sentaurus Device Manual* (Synopsys, Inc., Mountain View, CA, 2010).

[4.6] X. P. Liang and Y. Taur, “A 2-D analytical solution for SCEs in DG MOSFETs,” *IEEE Trans. Electron Devices*, vol. 51, no. 9, pp. 1385–1391, Sep. 2004.

[4.7] P. M. Solomon, D. J. Frank and S. O. Koswatta, “Compact model and performance estimation for tunneling nanowire FET,” in *Device research conference (DRC)*, 2011 69th annual; 2011. p. 197–8. doi:10.1109/DRC.2011.5994495.

[4.8] B. Rajamohanan, D. Mohata, A. Ali, and S. Datta, "Insight into the output characteristics of III-V tunneling field effect transistors," *Applied Physics Letters* 102, no. 9 (2013): 092105.

[4.9] Y. Taur, J. Wu and J. Min, “An analytic model for heterojunction tunnel FETs with exponential barrier,” *IEEE Trans. Electron Devices*, vol.62, no.5, pp. 1399-1404, May 2015.

[4.10] Y. Taur, X. Liang, W. Wang, and H. Lu, “A continuous, analytic drain-current model for DG-MOSFETs,” *IEEE Electron Device Letters*, Vol. 25, No. 2, pp. 107-109, Feb. 2004.

# Chapter 5 Examinations of twoband $E(k)$ relations for band-to-band tunneling

Thus far, one-band $E(k)$ relation has been assumed in the previous chapter. To more accurately model the current in TFETs, imaginary $E(k)$ relationship in the forbidden bandgap needs to be thoroughly investigated. In this chapter, four different two-band $E(k)$ relations for band-to-band tunneling in the literature are examined. Three of them are continuous functions, from Franz, Kane, and Flietner, respectively. One is piecewise, consisting of two linear lines from the band edges in a $k^2$ versus $E$ plot.

## 5.1 Introduction

Imaginary $E(k)$ relationship in the forbidden bandgap plays an important role in the current model of tunnel FETs (TFETs). Several two-band $E(k)$ relations have been published in the literature since the 1950s. They have commonalities as well as distinctions. No detailed comparison has been made on the different mathematical behavior of these models in various limits. This chapter examines four different two-band $E(k)$ models in terms of their $k^2$-$E$ plots. A power-$n$ model is introduced which is a general form of two of the published models.

## 5.2 Model descriptions

The tunneling probability is expressed by the WKB integral,

$$T(E) = \exp\left[-2\int |k(E)|dx\right] = \exp\left[-\frac{2\sqrt{2m}}{\hbar}\int \sqrt{\phi_B(E)}dx\right] \tag{5.1}$$

where $k$ is the imaginary wave vector, $\phi_B \equiv \hbar^2|k|^2/2m$ is the energy barrier height. The carrier energy $E$ is conserved in the tunneling process. The $x$-dependence comes from the bands: conduction ($E_c$) and valence ($E_v$). At the two ends of the tunneling path (i.e. in Fig. 4.1), $\phi_B(E = E_c, E_v) = 0$. Each $E(k)$ model provides a relation between $k$ and $E$, but cannot be separated into $m$ and $\phi_B$ in a rigorous way, as the imaginary wave vector for conduction band $k_c$ and that of valence band $k_v$ are coupled together. In the following description, we express the models in terms of the mass-barrier product $[m\phi_B]$, or $\hbar^2|k|^2/2$.

For the semiconductors with both heavy hole (HH) and light hole (LH) bands, the tunneling probability is the sum of the tunneling components from each hole band to the CB. While the LH band has a lower tunneling mass, its density of states is smaller than that of the HH band. For thin film TFETs, the quantization effect is stronger with the LH band, thereby moving it farther from the CB.

### 5.2.1 Franz's model

Franz was the first to publish a continuous two-band $E(k)$ model for band-to-band tunneling, as early as 1952 [5.1]. The initial model, Eq. (5.19) below, applies only for the simplest mathematical case of equal mass. Later on, it evolved to the more general form for unequal masses [5.2][5.3]. The mass×barrier height in terms of the electron energy $E$ in the bandgap, or $E_v \le E \le E_c$, is

$$[m\phi_B](Franz) = \left[\frac{1}{m_c(E_c - E)} + \frac{1}{m_v(E - E_v)}\right]^{-1} \tag{5.2}$$

Here, $m_c$ and $m_v$ are the effective masses of electrons and holes in the tunneling direction.

It can be shown that the maximum mass-barrier product is

$$[m\phi_B]_{\max}(Franz) = \frac{E_g}{\left[\frac{1}{\sqrt{m_c}} + \frac{1}{\sqrt{m_v}}\right]^2} \tag{5.3}$$

at an electron energy

$$E_{\max}(Franz) = E_i - \frac{E_g}{2} \times \frac{\sqrt{m_v} - \sqrt{m_c}}{\sqrt{m_v} + \sqrt{m_c}} \tag{5.4}$$

where $E_g = E_c - E_v$, and $E_i = (E_c + E_v)/2$.

### 5.2.2 Kane's model

Kane is well known for his pioneering work on band to band tunneling [5.4]. His result has been incorporated in many textbooks. He derived the following $E(k)$ relation from a perturbation solution to Schrodinger's eq. under the assumption of a constant field.

$$[m\phi_B](Kane) = \frac{2\left[\sqrt{E_g\left(\frac{E - E_v}{m_c^{\,2}} + \frac{E_c - E}{m_v^{\,2}}\right)} - \left(\frac{E - E_v}{m_c} + \frac{E_c - E}{m_v}\right)\right]}{\left(\frac{1}{m_c} - \frac{1}{m_v}\right)^2} \tag{5.5}$$

It has a maximum mass-barrier of

$$[m\phi_B]_{\max}(Kane) = \frac{E_g}{2\left(\frac{1}{m_c} + \frac{1}{m_v}\right)} \tag{5.6}$$

at an energy

$$E_{\max}(Kane) = E_i - \frac{E_g}{4} \times \frac{m_v - m_c}{m_v + m_c} \tag{5.7}$$

### 5.2.3 Flietner's model

Flietner published his model in 1972 [5.5]. It has been frequently cited in recent TFET publications.

$$[m\phi_B](Flietner) = \frac{E_g}{\dfrac{E_c - E}{m_v(E - E_v)} + \dfrac{E - E_v}{m_c(E_c - E)} + \dfrac{2}{\sqrt{m_c m_v}}} \tag{5.8}$$

The maximum value is

$$[m\phi_B]_{\max}(Flietner) = \frac{E_g\sqrt{m_c m_v}}{4} \tag{5.9}$$

at the same energy as that of Franz's model, Eq. (5.4).

### 5.2.4 Piecewise linear $k^2$-$E$ model

This model simply takes the parabolic $E(k)$ relations, or the linear $k^2$-$E$ relations, at the conduction band edge and the valence band edge and linearly extrapolate them into the forbidden gap until they intersect with each other. While the $k^2$-$E$ function is piecewise continuous, the derivative at the point of intersection is not. Note that this model has been associated with indirect tunneling where a phonon is involved [5.6][5.7].

The mass-barrier product or $\hbar^2|k|^2/2$ is

$$[m\phi_B](piecewise) = m_v(E - E_v) \qquad \text{for} \quad E \le E_{\max} \tag{5.10}$$

and

$$[m\phi_B](piecewise) = m_c(E_c - E) \qquad \text{for} \quad E \geq E_{\max} \tag{5.11}$$

where

$$E_{\max}(piecewise) = E_i - \frac{E_g}{2} \times \frac{m_v - m_c}{m_v + m_c} \tag{5.12}$$

The maximum mass×barrier is

$$[m\phi_B]_{\max}(piecewise) = \frac{E_g}{\frac{1}{m_c} + \frac{1}{m_v}} \tag{5.13}$$

### 5.2.5 Generalized Power-n model

By mimicking (i.e. taking a mean power) a mathematical expression from the MOSFET velocity saturation model [5.10], one can consolidate Franz's model and the piecewise linear model into the following generalized equation:

$$[m\phi_B](generalized) = \left\{ \left[ \frac{1}{m_c(E_c - E)} \right]^n + \left[ \frac{1}{m_v(E - E_v)} \right]^n \right\}^{-1/n} \tag{5.14}$$

$n = 1$ gives Franz's model, while $n = \infty$ yields the piecewise linear model. $n = 10$ gives a close approximation to the piecewise model with a $k^2$-$E$ function continuous and differentiable at every $E$.

## 5.3 Model comparison

The first general observation of all two-band models is that they are all symmetric with respect to CB and VB. In other words, the model remains invariant under the exchange of $m_c \leftrightarrow m_v$ and $(E_c - E) \leftrightarrow (E - E_v)$.

From Eqs. (5.6) and (5.13), $[m\phi_B]_{max}$(Kane) = 0.5$[m\phi_B]_{max}$(piecewise). Furthermore, it can be shown that

$$[m\phi_B]_{\max}(Kane) \le [m\phi_B]_{\max}(Franz) \le [m\phi_B]_{\max}(Flietner) \tag{5.15}$$

It can also be shown that

$$[m\phi_B]_{\max}(Kane) \le [m\phi_B]_{\max}(Franz) \le [m\phi_B]_{\max}(piecewise) \tag{5.16}$$

However, the order of $[m\phi_B]_{max}$(Flietner) and $[m\phi_B]_{max}$(piecewise) depends on $m_v/m_c$, as discussed below.

### 5.3.1 Near the band Edges

All five models share the same asymptotic behavior near $E_c$ and $E_v$, namely,

$$[m\phi_B] = m_v(E - E_v) \qquad \text{for} \quad E \approx E_v \tag{5.17}$$

and

$$[m\phi_B] = m_c(E_c - E) \qquad \text{for} \quad E \approx E_c\,. \tag{5.18}$$

This is expected of the continuity of parabolic $E(k)$ from the CB and VB into the bandgap.

### 5.3.2 Equal masses

If $m_c = m_v$, the three continuous models, Franz, Kane, Flietner, yield identical $E(k)$ or $k^2$-$E$ (Fig. 5.1),

$$[m\phi_B] = m_c \frac{(E_c - E)(E - E_v)}{E_g} \tag{5.19}$$

The maximum mass×barrier height is $[m\phi_B]_{max} = m_c E_g/4$ at $E = (E_c + E_v)/2$ or the midgap.

This is the simple parabolic function Franz first came up with in 1952. Fig. 5.1 shows that the piecewise linear model has a maximum mass×barrier twice as high, $m_c E_g/2$, at the midgap.

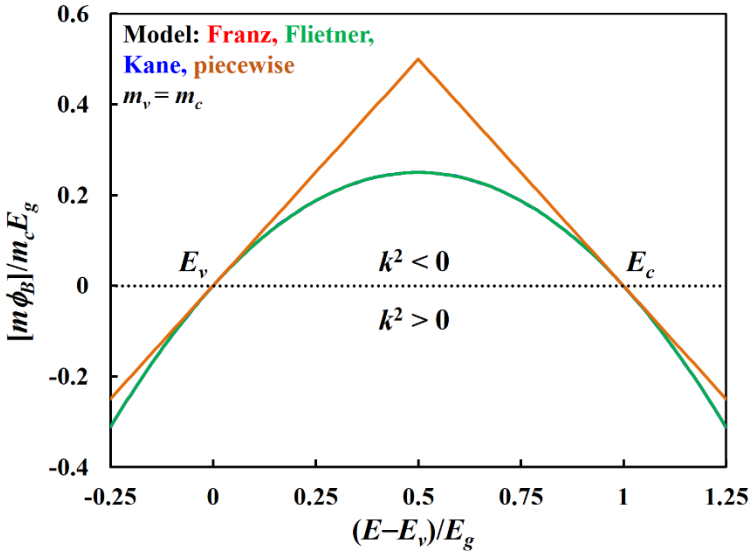



Figure 5.1: $[m\phi_B]/m_c$ or $-\hbar^2k^2/2m_c$ versus $E$ in units of $E_g$ for $m_c = m_v$. All three continuous models are described by the same curve.

### 5.3.3 Mismateched masses

In the case of highly mismatched $m_c$ and $m_v$, $[m\phi_B]_{max}$ is controlled by the lighter of the two masses in four of the five models. The only exception is Flietner's model in Eq. (5.8), in which the heavier mass plays as important a role in $[m\phi_B]_{max}$ as the lighter mass.

Here, we designate $m_c$ to be the lighter of the two, as usually is the case. Fig. 5.2 plots all four models for $m_v/m_c = 10$. Consider the slope $d[m\phi_B]/dE$ from $E_v$ to $E_c$. All slopes start with $m_v$ at $E = E_v$, then decrease to zero at the peak (for the three continuous models) and continue to negative values until reaching $-m_c$ at $E_c$. For Franz's and Kane's models, $d[m\phi_B]/dE$ decreases monotonically. However, for Flietner's model, the curve goes outside

the piecewise curve between $E_{max}$ and $E_c$. This means that the slope reaches a value more negative than $-m_c$ before coming back to $-m_c$ at $E_c$. This becomes more evident in Fig. 5.3 for $m_v/m_c = 100$.

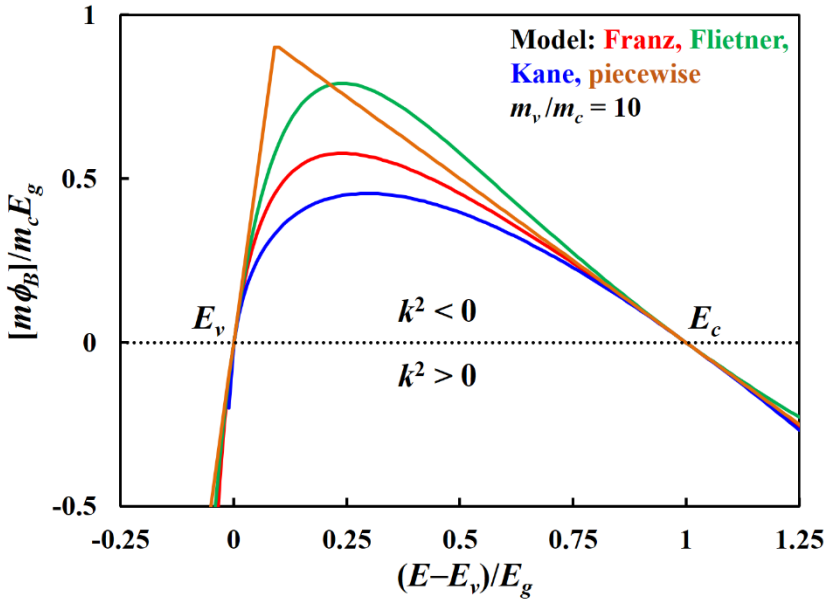



Figure 5.2: $[m\phi_B]/m_c$ or $-\hbar^2k^2/2m_c$ versus $E$ in units of $E_g$ for $m_v/m_c = 10$.

Fig. 5.3 also shows that in the limit of $m_v/m_c >> 1$, both Franz's model and the piecewise linear model converge to a mass-barrier product of $[m\phi_B] = m_c(E_c - E)$, independent of $m_v$. This can simply be interpreted as a one-band model with mass $m_c$ and barrier height $E_c - E$. Kane's model also becomes independent of $m_v$ in the same limit, but with $[m\phi_B]_{max} = m_cE_g/2$—only half of Franz's or the piecewise model.

Flietner's model behaves very differently in the limit of $m_v/m_c >> 1$. The mass-barrier product or $|k|^2$ keeps rising with $m_v$, with $[m\phi_B]_{max} = m_cE_g(m_v/m_c)^{1/2}/4$. It is not clear how to physically interpret the mass and barrier separately. Mathematically, Flietner's model

yields a maximum $|k|^2$ higher than that of the piecewise linear model when the $m_v/m_c$ ratio exceeds ~14.

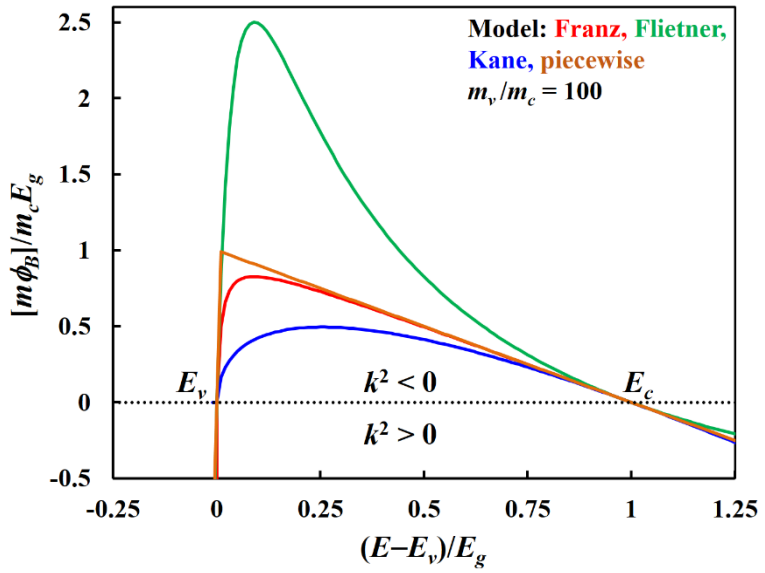



Figure 5.3: $[m\phi_B]/m_c$ or $-\hbar^2k^2/2m_c$ versus $E$ in units of $E_g$ for $m_v/m_c = 100$

### 5.3.4 Beyond band Edges

As far as band-to-band tunneling is concerned, only the $E(k)$ relation within the bandgap matters. Nevertheless, a reasonable degree of continuity of the mathematical functions is expected when extrapolating back into the valence band ($E < E_v$) and the conduction band ($E > E_c$). The piecewise linear model of course just continues its trends outside $E_g$. For $m_c = m_v$, the common function of the first three models, Eq. (5.19), extends well into the bands, as seen in Fig. 5.1. For $m_v/m_c = 10$ in Fig. 5.2, however, there are a few adverse effects of the models. Kane's model turns imaginary when $E$ is only $0.01E_g$ below $E_v$. Franz's and Flietner's models both exhibit a singularity when $E$ is a fraction of $E_g$ below $E_v$. The singularity is avoided in the $n = 2$ generalized model.

### 5.3.5 $T(E)$ under uniform field

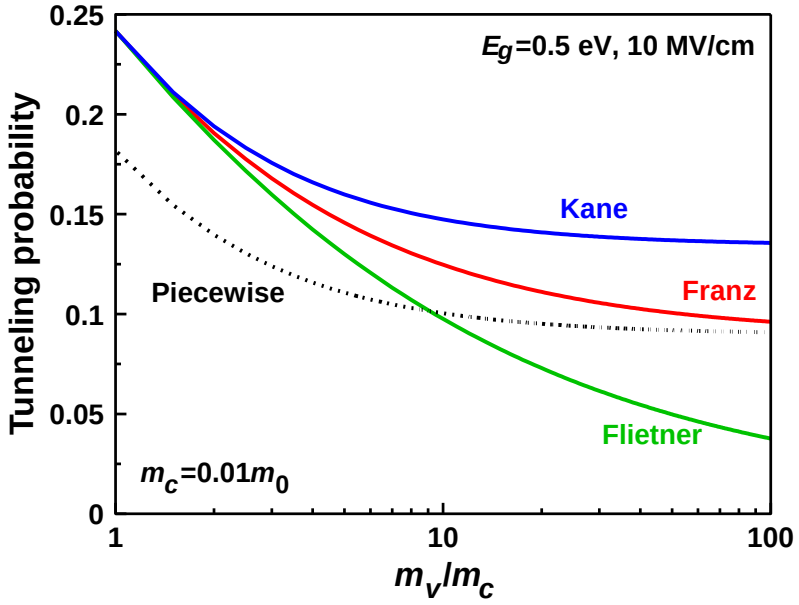



Figure 5.4: Tunneling probability under a constant field versus mass ratio.

To illustrate the effect of two-band $E(k)$ models on the tunneling probability of Eq. (5.1), we consider a uniform field case in Fig. 5.4. For a uniform field, the tunneling distance and therefore $T(E)$ is independent of $E$. The tunneling probability is plotted as a function of $m_v/m_c$ ratio with $m_c$ set to $0.01m_0$. At $m_v/m_c = 1$, all three continuous two-band models give the same $T$ somewhat higher than the piecewise model as noted before. In the limit of $m_v/m_c >> 1$, the piecewise model approaches the well-known expression for a triangular barrier,

$$T = \exp\left(-\frac{4\sqrt{2m_c}E_g^{3/2}}{3q\hbar\xi}\right) \tag{5.20}$$

where the tunneling is electron-like throughout the bandgap. Franz's model also approaches the same limit. However, Flietner's model continues to go down as $m_v/m_c$

increases. Kane's model, on the other hand, settles to a value higher than the piecewise model. This suggests that Franz's model is the model of choice for the widest range of $m_v$ and $m_c$.

## 5.4 Summary

By examining the four published two-band $E(k)$ relations, including a piecewise case, we found that both Franz's model and the piecewise linear model behave reasonably in all asymptotic limits. They are both reduced to a one-band model when one mass is much lighter than the other, as one might expect. Kane's model under-predicts the maximum barrier height when the masses are highly mismatched. Flietner's model, on the other hand, can give mass-barrier height products or $k^2$ larger than that of the piecewise linear model in the same limit. Franz's model is the model of choice for the widest range of $m_v$ and $m_c$.

The text of Chapter 5, in part, is a reprint of the material as it appears in "Examination of two-band $E(k)$ relations for band-to-band tunneling" by Yuan Taur and Jianzhi Wu, IEEE Transaction on Electron Devices, Feb. 2016. The dissertation author was a co-author of this paper.

## 5.5 References

[5.1] W. Franz, *Ann. Phys.* (Germany) (6) 11, 17 (1952).

[5.2] Y-Y. Fan, S. P. Mudanai, W. Chen, L. F. Register, and S. K. Banerjee. "High-k Gate Dielectric Materials Integrated Circuit Device Design Issues." In High Dielectric Constant Materials, pp. 570. Springer Berlin Heidelberg, 2005.

[5.3] *TCAD Sentaurus Device Manual* (Synopsys, Inc., Mountain View, CA, 2013).

[5.4] E. O. Kane, "Zener tunneling in semiconductors." *Journal of Physics and Chemistry of Solids* 12, no. 2 (1960): 181-188.

[5.5] H. Flietner, "The E (k) Relation for a Two‐Band Scheme of Semiconductors and the Application to the Metal‐Semiconductor Contact." *Physica Status Solidi* (b) 54, no. 1 (1972): 201-208.

[5.6] L. V. Keldysh, "Behavior of non-metallic crystals in strong electric fields." *Soviet Journal of Experimental and Theoretical Physics* 6 (1958): 763.

[5.7] P. M. Solomon, J. Jopling, D. J. Frank, C. D'Emic, O. Dokumaci, P. Ronsheim, and W. E. Haensch. "Universal tunneling behavior in technologically relevant P/N junction diodes." *Journal of Applied Physics* 95, no. 10 (2004): 5800-5812.

[5.8] http://www.tf.uni-kiel.de/matwis/amat/semi_en/kap_2/backbone/r2_3_1. html

[5.9] http://www.ioffe.ru/SVA/NSM/Semicond/

[5.10] Y. Taur and T. H. Ning, *Fundamentals of modern VLSI devices*, Cambridge, U.K.: Cambridge University Press, 2009, pp.187-191.

# Chapter 6 A continuous and semianalytic I-V model for DG and NW TFETs

## 6.1 Model descriptions

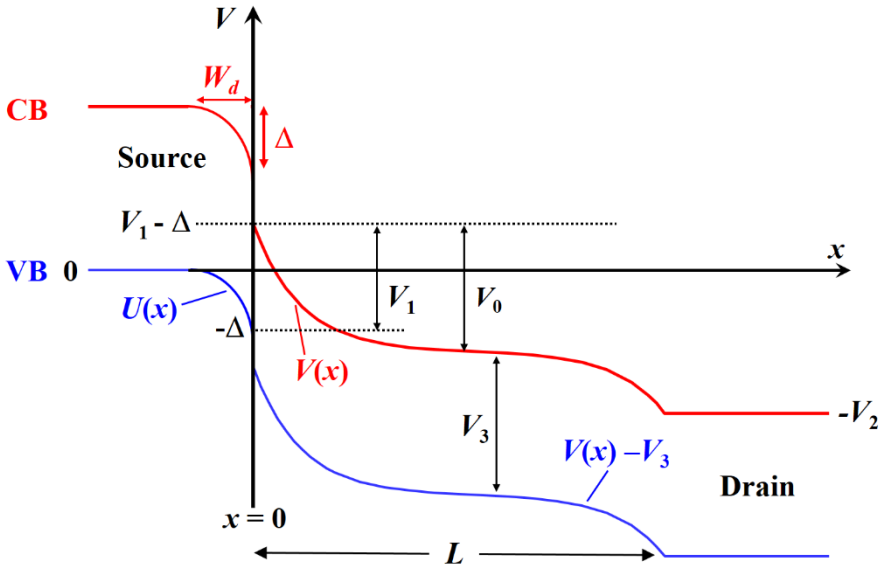



Figure 6.1: Schematic band diagram of a TFET.

The band diagram of a DG or NW TFET with $p^+$ source and $n^+$ drain is shown in Fig. 6.1. The material dependent parameters are $V_1$: the staggered bandgap of the source to channel heterojunction, and $V_3$: the channel bandgap. The bias dependent parameters are $V_0$: controlled by the gate voltage, $V_2$: controlled by the drain voltage, and $\Delta$: band bending in the source region of doping density $N_a$. The zero energy reference is chosen to be the top of the source valence band. From an approximate solution to 2D Poisson's eq., the channel conduction band is expressed as [6.1]:

$$V(x) = V_0 \frac{\sinh[\pi(L-x)/\lambda]}{\sinh(\pi L/\lambda)} - (V_0 - V_1 + \Delta) - (V_2 - V_0 + V_1 - \Delta)\frac{\sinh(\pi x/\lambda)}{\sinh(\pi L/\lambda)} \qquad (6.1)$$

It satisfies the source and drain boundary conditions, $V(0) = V_1 - \Delta$ and $V(L) = -V_2$. Here,

$\lambda$ is the scale length given by the device dimensions in the direction perpendicular to the gate. Consider the simplest case of the same permittivity between the semiconductor and the gate insulator, $\lambda = t_s + 2t_i$ for DG [6.2] and $\lambda = \pi(r_s + t_i)/\alpha$ for NW where $t_s$ is the semiconductor thickness, $r_s$ is the nanowire radius, $t_i$ is the insulator thickness, and $\alpha$ = 2.405 is the first zero of the zeroth order Bessel function [6.3].

By employing the depletion approximation, the source valence band is expressed as:

$$U(x) = -\frac{q^2 N_a}{2\varepsilon_s}(x + W_d)^2 \tag{6.2}$$

where $W_d = [2\varepsilon_s\Delta/(q^2N_a)]^{1/2}$ is the depletion width. Both $V_0$ and $\Delta$ increase with the gate voltage $V_{gs}$. We assume a gate work function such that at $V_{gs} = 0$, $V_0 = V_1 - \Delta$ and the tunneling window starts to open. When the TFET is biased in saturation, or when $V_2 > V_0 - (V_1 - \Delta)$, the channel potential is directly modulated by the gate voltage so $qV_{gs} = V_0 - (V_1 - \Delta)$. Substituting this relation into $dV/dx\big|_{x=0} = dU/dx\big|_{x=0}$, the condition for continuity of field at the heterojunction, yields

$$\left(\frac{\pi}{\lambda}\right)\frac{V_0\cosh(\pi L/\lambda) + (V_2 - qV_{gs})}{\sinh(\pi L/\lambda)} = \sqrt{\frac{2q^2N_a(qV_{gs} - V_0 + V_1)}{\varepsilon_s}} \tag{6.3}$$

Here, $V_2$ is given by the source-drain voltage: $V_2 = qV_{ds} + d_1 + d_2$, with $d_1$ and $d_2$ defined as the source and drain degeneracies. They are given by the Fermi integrals, $F_{1/2}(d_1/kT) = (\pi^{1/2}/2)(N_a/N_v)$ and $F_{1/2}(d_2/kT) = (\pi^{1/2}/2)(N_d/N_c)$, in terms of the source and drain doping $N_a$, $N_d$ and the effective density of states $N_v$, $N_c$. For given $V_{gs}$ and $V_{ds}$, $V_0$ is solved from Eq. (6.3), a quadratic equation [6.4].

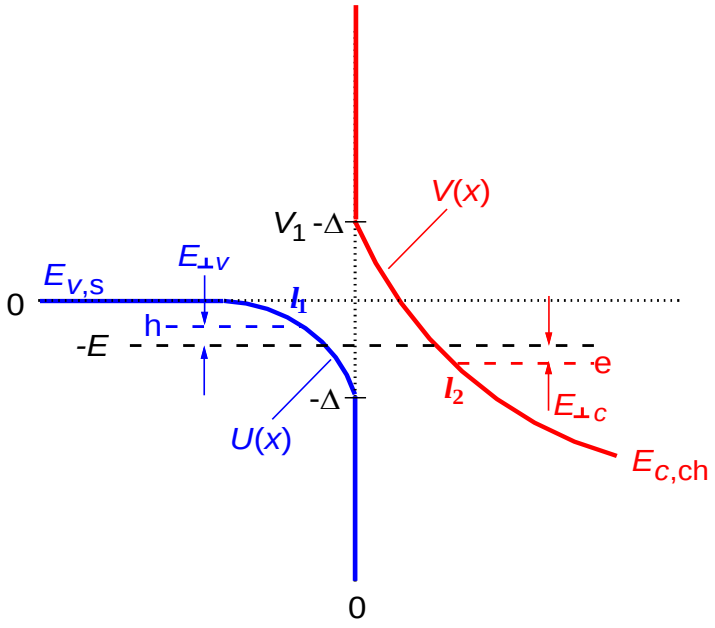



Figure 6.2: A magnified view of the tunneling region in Fig. 6.1. $E_{\perp v}$ and $E_{\perp c}$ are the carrier kinetic energies in the direction perpendicular to the tunneling path.

For electrons at energy $-E$ with perpendicular kinetic energy of $E_{\perp v}$ and $E_{\perp c}$ in Fig. 6.2, the probability of band-to-band tunneling is given by the WKB integrals:

$$T(E,E_{\perp v}) = \exp\left\{-2\left[\int_{l_1}^{0} k_{is}(E,E_{\perp v})dx + \int_{0}^{l_2} k_{im}(E,E_{\perp v})dx\right]\right\} \tag{6.4}$$

where $k_{is}$ and $k_{im}$ are the imaginary wavevectors for tunneling in the source and in the channel, respectively. If $m_v$ and $m_c$ are the effective masses associated with the valence band and the conduction band, conservation of perpendicular momentum requires $m_v E_{\perp v} = m_c E_{\perp c}$. By adopting Franz's two-band $E(k)$ relation [6.5]-[6.8],

$$k_{im}(E,E_{\perp v}) = \frac{\sqrt{2}}{\hbar}\left(\frac{1}{m_v[V_3 - V(x) - E + E_{\perp v}]} + \frac{1}{m_c[V(x) + E + E_{\perp c}]}\right)^{-1/2} \tag{6.5}$$

Note that $V(x) - V_3$ is the valence band energy and $m_c E_{\perp c}$ can be replaced by $m_v E_{\perp v}$. The

end of the tunneling path in the conduction band of channel is such that $V(l_2) + E + E_{\perp c} = 0$.

For tunneling in the source, a one-band $E(k)$ model is used under the assumption that the source bandgap is much wider than $V_1$:

$$k_{is}(E, E_{\perp v}) = \frac{\sqrt{2m_v}}{\hbar}\sqrt{-U(x) - (E - E_{\perp v})} \tag{6.6}$$

The end of the tunneling path in the valence band of source, $l_1 < 0$, is given by $U(l_1) + E - E_{\perp v} = 0$. This WKB integral can be done analytically,

$$\int_{l_1}^{0} k_{is}(E, E_{\perp v})dx = \frac{\sqrt{\varepsilon_s m_v}}{q\hbar\sqrt{N_a}}\left[\sqrt{\Delta(\Delta - E + E_{\perp v})} - (E - E_{\perp v})\ln\left(\sqrt{\frac{\Delta}{E - E_{\perp v}}} + \sqrt{\frac{\Delta}{E - E_{\perp v}} - 1}\right)\right] \tag{6.7}$$

The WKB integral of $k_{im}$ [2nd in Eq. (6.4)] must be evaluated numerically. For electron energies below the staggered bandgap, or $E > \Delta$, there is only one WKB integral, that of electron tunneling in the channel. For energies further below $V_1 - \Delta - V_3$, the lower limit of the 2nd integral is changed to $l_0 > 0$ where $V(l_0) - V_3 + E - E_{\perp v} = 0$.

The current for a 3D ballistic TFET is given by [6.9]:

$$j = \frac{qm_v}{2\pi^2\hbar^3}\int_0^{V_2}(f_s - f_d)\left[\int_0^{E_{\perp m}} T(E, E_{\perp v})dE_{\perp v}\right]dE \tag{6.8}$$

where $f_s(E)$ and $f_d(E)$ are the state occupancy factors at the source and the drain. Since $m_v E_{\perp v} = m_c E_{\perp c}$, the upper limit of $E_{\perp v}$ can be imposed by the kinetic energy in the valence band $E_{\perp v} \le E$, or by $E_{\perp c} \le V_2 - E$ in the conduction band. $E_{\perp m}$ is then the smaller of $E$ and $(m_c/m_v)(V_2 - E)$.

For heterojunction TFETs, $m_v$ of the channel may be different from the $m_v$ of source. If we designate $m_v'$ and $E_{\perp v}'$ as those for the channel, then $m_v E_{\perp v} = m_v' E_{\perp v}' = m_c E_{\perp c}$. The denominator of the first term in the bracket of Eq. (6.5) becomes $m_v'[V_3 - V(x) - E + E_{\perp v}']$ $= m_v'[V_3 - V(x) - E] + m_v E_{\perp v}$. In other words, the only change is $m_v$ to $m_v'$ in front of the $[V_3 - V(x) - E]$ factor in Eq. (6.5). The model can also be generalized to accommodate anisotropic mass in the source or channel bands by straightforward modifications of Eqs. (6.5) and (6.6):

$$k_{im}(E, E_{\perp v}) = \frac{\sqrt{2}}{\hbar}\left(\frac{1}{m_{//v}'[V_3 - V(x) - E] + (m_{//v}'/m_{\perp v}')m_{\perp v}E_{\perp v}} + \frac{1}{m_{//c}[V(x) + E] + (m_{//c}/m_{\perp c})m_{\perp v}E_{\perp v}}\right)^{-1/2} \quad (6.9)$$

where $m_{//v}'$, $m_{\perp v}'$ are the effective masses for channel material along two orthogonal directions, which are different from $m_{//v}$, $m_{\perp v}$ for the source material.

Eq. (6.8) is a triple integral for current calculation. A standard approach is to expand $T(E, E_{\perp v})$ to the first order of $E_{\perp v}$ so the $E_{\perp v}$ integration can be carried out analytically [6.10] to reduce the current to a double integral. We found that while this approach gives accurate results when $m_c \geq m_v$ in the absence of source depletion, the error is generally not acceptable in other cases.

## 6.2 Scaling guideline

With $V_1 = 0$, $V_3 = 1.0$ eV, and $V_{ds} = 0.5$ V, Fig. 6.3 shows the $I_{ds}$-$V_{gs}$ characteristics generated by the model for several channel lengths. Both the source doping and the effective density of states are set at $5\times10^{18}$ cm$^{-3}$. The on current is essentially independent of $L$, as its contribution is mostly from source to channel tunneling, or $E \in (0, V_0 - V_1 + \Delta)$

in the current integral. But the off current, mainly from source to drain tunneling over the energy range $E \in (V_0 - V_1 + \Delta, V_2)$, is very sensitive to $L$.

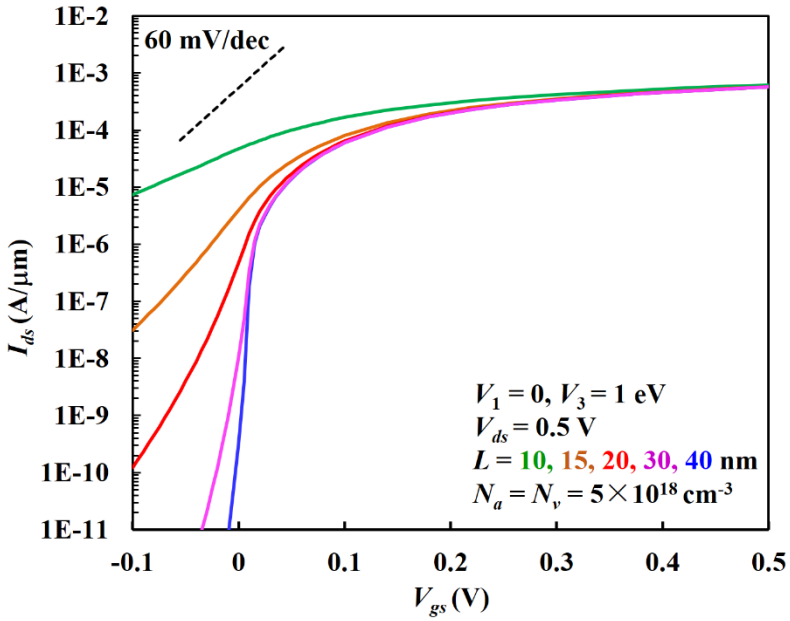
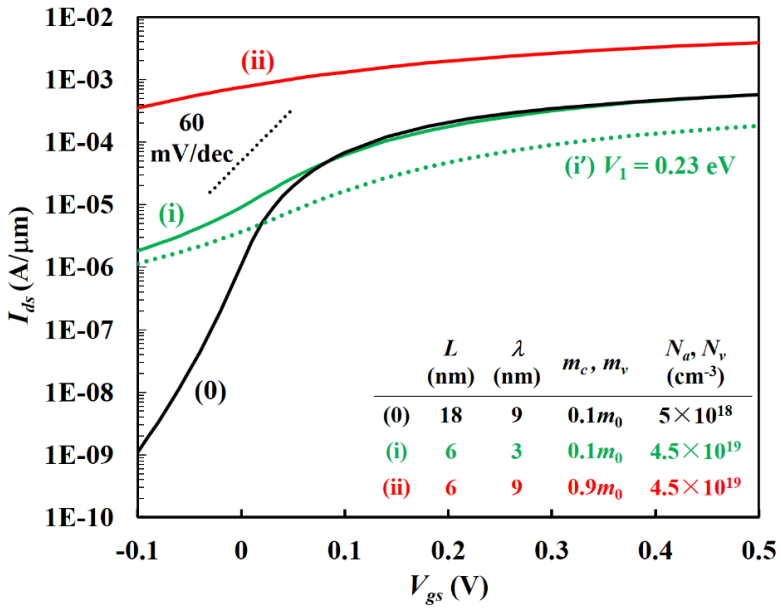



Figure 6.3: Model generated $I_{ds}$-$V_{gs}$ characteristics for varying $L$. The drain degeneracy is $d_2 = 0.3$ eV based on $N_d = 3\times10^{19}$ cm$^{-3}$ and $N_c = 10^{18}$ cm$^{-3}$. The rest of the parameters are $\lambda = 9$ nm, $t_s = 5$ nm, $m_c = m_v = 0.1m_0$.

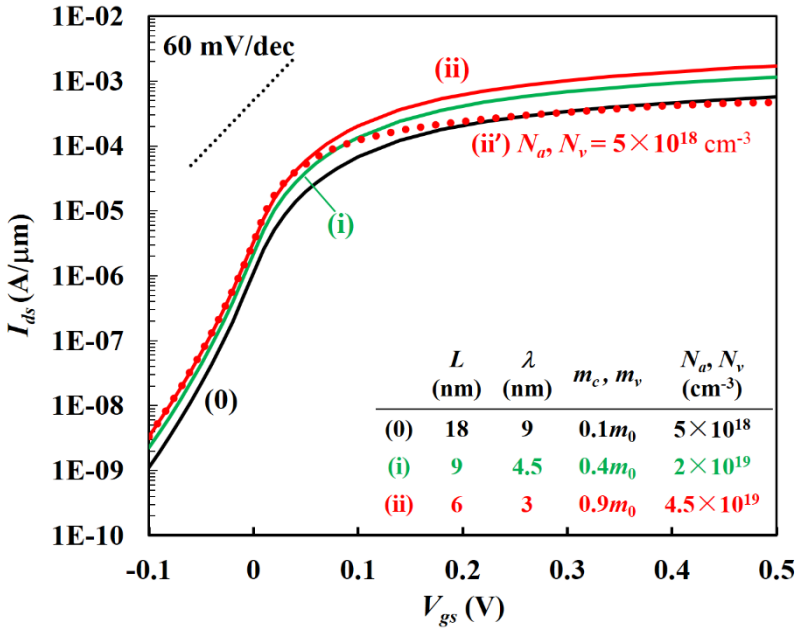
Depending on the exact specification for $I_{off}$, the minimum acceptable channel length in Fig. 6.3 is $L \approx 18$ nm, or about $2\lambda$. To further scale the TFET length for circuit density, $\lambda$ needs to be reduced. Fig. 6.4(a) considers a 3× scaling scenario to $L = 6$ nm with $\lambda = 3$ nm. Curve (i) shows that although the geometric scaling of $\lambda$ preserves the 2D field pattern, it is insufficient to contain the short channel off currents. Raising $V_1$ to 0.23 eV [curve (i′)] does not help either. This is because the WKB integral, $-2\int k_{im}dx$, is proportional to $-\lambda\sqrt{m}$ where $m$ stands for both $m_c$ and $m_v$. $T(E)$ goes up by an exponential factor when $\lambda$ is reduced by 3×. To keep the WKB integral invariant under scaling, both $m_c$ and $m_v$ need to be

increased by 9× to 0.9$m_0$. On the other hand, just scaling $m_c$ and $m_v$ to 0.9$m_0$ with the same $\lambda$ = 9 nm [curve (ii) in Fig. 6.4(a)] does not work either. Mathematically, the proper scaling relations derived from the model equations are: $\lambda \rightarrow \lambda/\alpha$, $m_c \rightarrow \alpha^2 m_c$, $m_v \rightarrow \alpha^2 m_v$, $N_a \rightarrow \alpha^2 N_a$, $N_v \rightarrow \alpha^2 N_v$, resulting in $I_{ds} \rightarrow \alpha I_{ds}$. In Fig. 6.4(b), we show that for $\alpha$ = 2 and $\alpha$ = 3, the entire $I_{ds}$-$V_{gs}$ characteristics simply shift up by the scaling factor. If $N_a$ and $N_v$ are not scaled up, $I_{off}$ is not affected, but $I_{on}$ does not increase by the scaling factor [curve (ii′) in Fig. 6.4(b)]. While the above rules apply to constant voltage scaling, $V_{dd}$ can be reduced at any node by trading off $I_{on}$. For example, with a requirement of $I_{off}$ = 1 nA/μm, the on current at $V_{dd}$ = 0.3 V is ~ 0.45 of that at $V_{dd}$ = 0.5 V from the solid curves in Fig. 6.4(b). The CV/I delay is then (0.3/0.5)/0.45 or ~ 30% longer at $V_{dd}$ = 0.3 V, while the active power, I×V ~ 0.45×(0.3/0.5), is ~ 3.7× less [6.11].




(a)

Figure 6.4: (a) 3× scaling of (0) $L$ = 18 nm TFET to $L$ = 6 nm, with (i) $\lambda$ scaling only, (ii) $m$ scaling only, and the dotted curve (i′) same as (i) with $V_1$ = 0.23 eV. All solid curves are for $V_1$ = 0 as in Fig. 5.




(b)

Figure 6.4: continued (b) 2× and 3× scaling from the same (0) as in (a) to (i) $L$ = 9 nm and (ii) $L$ = 6 nm. The dotted curve (ii′) is the same as (ii) but without $N_a$, $N_v$ scaling.

Of course, physics and material availability will limit how far TFETs can be scaled. In one study, an empirical $m_c/m_0 \approx E_g$(eV)/20 relationship was found for III-V semiconductors [6.12]. Another study took the effect of quantum confinement into account and found that Si or Ge nanowire TFETs can deliver the highest $m_c/m_0$ [6.13]. The 3× scaling scenario depicted in Fig. 6.4(b) may not be realizable without material innovation.

## 6.3 Ambipolar effects

Thus far we assumed a wide channel bandgap, $V_3$ = 1.0 eV, so there is no ambipolar effect in the range of $V_{gs}$ and $V_{ds}$ studied. Fig. 6.5 shows that ambipolar effect comes in when $V_3$ < 0.5 eV. It causes unacceptably high $I_{off}$ and subthreshold swing. This stems from

the channel-to-drain tunneling depicted in Fig. 6.6 [6.14]. The tunneling window opens when $qV_{gs} < V_2 - V_3$, or $qV_{gs} < qV_{ds} + d_1 + d_2 - V_3$. For the case of $V_3 = V_2 = 0.5$ eV in Fig. 6.5, the tunneling probability near the off condition is plotted in Fig. 6.7 as a function of $E$. At $V_{gs} = 0$, there is only a minimum current due to source-to-drain tunneling very close to $E = 0$. A slightly positive $V_{gs}$ opens the window for source-to-channel tunneling. At negative $V_{gs}$, channel-to-drain tunneling starts to show up at $E = 0.5$ eV, then spreads to lower $E$. To avoid the increase of $I_{off}$ at $V_{gs} = 0$ due to ambipolar effect, the design guideline is to use $V_3 > qV_{dd} + d_1 + d_2$.

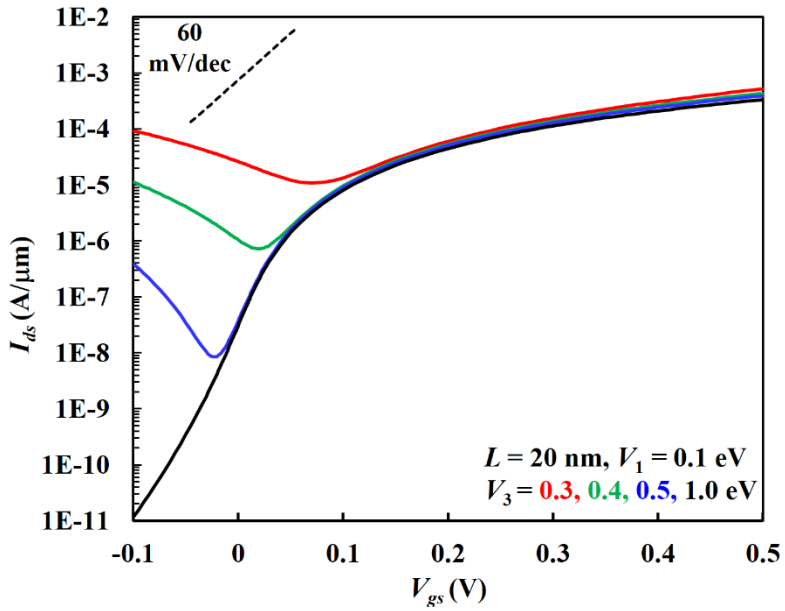



Figure 6.5: Model generated $I_{ds}$-$V_{gs}$ characteristics for varying $V_3$. Fixed are $V_1 = 0.1$ eV and $V_2 = 0.5$ eV. The rest of the parameters are the same as in Fig. 6.3.

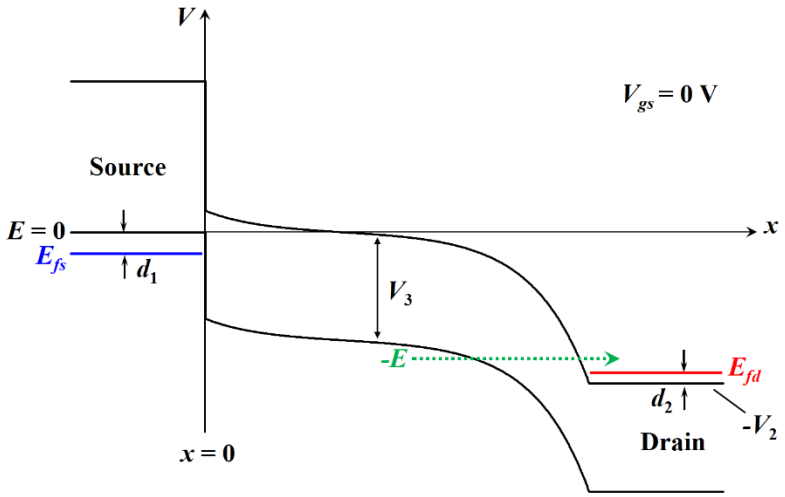



Figure 6.6: Schematic band diagram of TFET at $V_{gs}$ = 0, showing channel-to-drain tunneling or ambipolar effect.

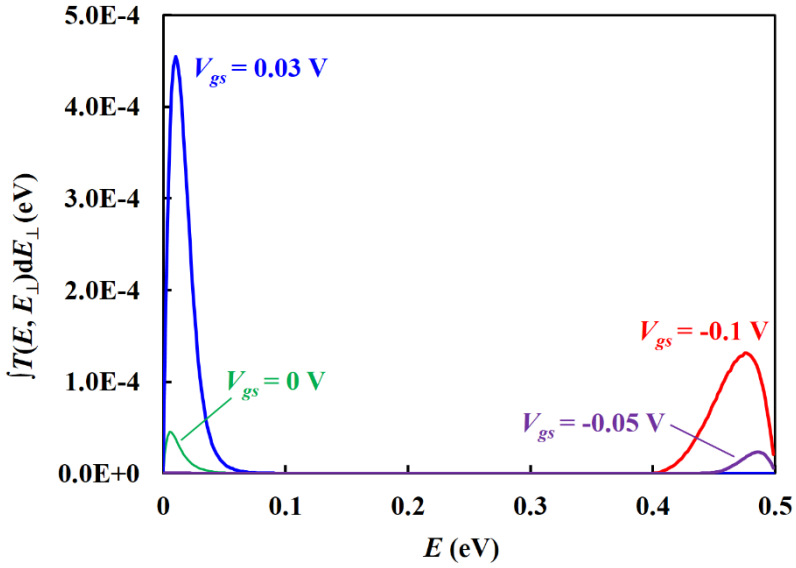



Figure 6.7: Tunneling probability versus energy at several $V_{gs}$ near zero for the $V_3$ = 0.5 eV case in Fig. 6.5.

## 6.4 Drain doping effects

For TFETs with sub-10 nm channel length, both the off-current ($I_{off}$) and the

subthreshold swing (SS) go up sharply because of source to drain tunneling [6.4]. A number of TFET designs early on have exercised moderately doped drain [6.15]-[6.17]. In this section, we developed a comprehensive model on the use of lightly doped drain to suppress the $I_{off}$ and SS of short-channel TFETs.

Fig. 6.8 shows the band diagram of a DG TFET with depletion in the source and the drain. $V_1$ is the staggered bandgap of the source to channel heterojunction. $V_3$ is the channel bandgap. $V_0$ expresses the effect of gate voltage, and $V_2$ the drain voltage. The potential profile of the device consists of three regions, channel, source, and drain:

$$V(x) = V_0 \frac{\sinh[\pi(L-x)/\lambda]}{\sinh(\pi L/\lambda)} - V_0 + V_1 - \Delta - (V_2 - V_0 + V_1 - \Delta - \Delta') \frac{\sinh(\pi x/\lambda)}{\sinh(\pi L/\lambda)} \quad (0 \le x \le L) \text{ (6.10)}$$

$$U(x) = -\frac{\Delta}{(W_d)^2}(x + W_d)^2 \quad (x < 0) \text{ (6.11)}$$

$$W(x) = -V_2 + \frac{\Delta'}{(W_d')^2}(x - L - W_d')^2 \quad (x > L) \text{ (6.12)}$$

Here, $\lambda$ is the scale length given by the device thickness [6.4]. $L > \lambda$ is assumed so higher order terms can be left out of $V(x)$. $\Delta$ and $\Delta'$ are the band bending in the source and drain regions. If their doping densities are $N_a$ and $N_d$, then the depletion widths for $U(x)$ and $W(x)$ are $W_d = [2\varepsilon_s\Delta/(q^2N_a)]^{1/2}$ and $W_d' = [2\varepsilon_s\Delta'/(q^2N_d)]^{1/2}$, respectively. The gate work function is chosen such that at $V_{gs} = 0$, $V_0 = V_1 - \Delta$ and the tunneling window starts to open. For $V_{gs} > 0$, $qV_{gs} = V_0 - (V_1 - \Delta)$ when biased in saturation. Continuity of field at the source and drain junctions, $dV/dx|_{x=0} = dU/dx|_{x=0}$ and $dV/dx|_{x=L} = dW/dx|_{x=L}$, yields

$$\frac{\pi}{q\lambda}\times\frac{(qV_{gs}-\Delta+V_1)\cosh(\pi L/\lambda)+(V_2-qV_{gs}-\Delta')}{\sinh(\pi L/\lambda)}=\sqrt{\frac{2N_a\Delta}{\varepsilon_s}} \tag{6.13}$$

$$\frac{\pi}{q\lambda}\times\frac{(qV_{gs}-\Delta+V_1)+(V_2-qV_{gs}-\Delta')\cosh(\pi L/\lambda)}{\sinh(\pi L/\lambda)}=\sqrt{\frac{2N_d\Delta'}{\varepsilon_s}} \tag{6.14}$$

which allow $\Delta$ and $\Delta'$ to be solved for given $V_{gs}$ and $V_2$. Note that $V_2 = qV_{ds} + d_1 + d_2$, with $d_1$ and $d_2$ the source and drain degeneracies given by the Fermi integrals, $F_{1/2}(d_1/kT) = (\pi^{1/2}/2)(N_a/N_v)$ and $F_{1/2}(d_2/kT) = (\pi^{1/2}/2)(N_d/N_c)$, where $N_v$, $N_c$ are the effective density of states of the source and the drain.

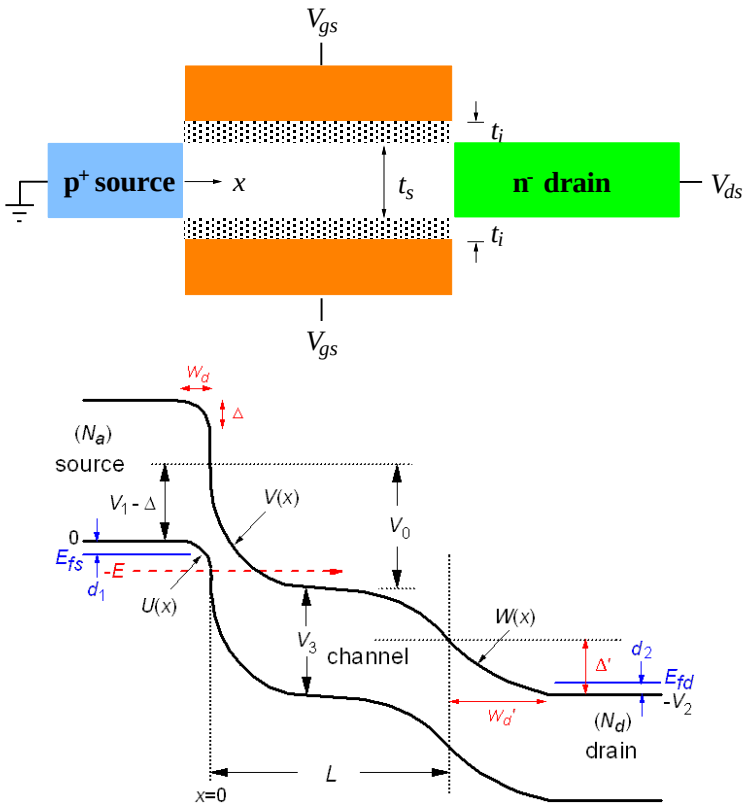



Figure 6.8: Top: TFET cross-section. The gates are assumed to cover an intrinsic channel of length $L$, with no overlap of the source and drain. Bottom: Band diagram of a heterojunction TFET. $U(x)$ is the valence band energy of the source, $V(x)$ and $W(x)$ the conduction band of channel and drain.

The band-to-band tunneling probability is calculated from the sum of up to three WKB integrals:

$$T(E,E_{\perp v})=\exp\left\{-2\left[\int_{l_0}^{0}k_s(E,E_{\perp v})dx+\int_{l_1}^{l_2}k_{ch}(E,E_{\perp v})dx+\int_{L}^{l_3}k_d(E,E_{\perp v})dx\right]\right\} \tag{6.15}$$

where $k_s$, $k_{ch}$ and $k_d$ are the imaginary wavevectors for tunneling in the source, channel and drain, respectively. $E_{\perp v}$ and $E_{\perp c}$ are the carrier kinetic energies perpendicular to the tunneling direction. If $m_c$ and $m_v$ are the effective masses of the conduction and the valence band, conservation of momentum requires $m_vE_{\perp v} = m_cE_{\perp c}$. Franz's two-band $E(k)$ relation is employed for tunneling in the channel region [6.5].

$$k_{ch}(E,E_{\perp v})=\frac{\sqrt{2}}{\hbar}\left(\frac{1}{m_v[V_3-V(x)-E+E_{\perp v}]}+\frac{1}{m_c[V(x)+E+E_{\perp c}]}\right)^{-\frac{1}{2}} \tag{6.16}$$

One-band $E(k)$ models are used for tunneling in the source and the drain,

$$k_s(E,E_{\perp v})=\frac{\sqrt{2m_v}}{\hbar}\sqrt{-U(x)-(E-E_{\perp v})} \tag{6.17}$$

$$k_d(E,E_{\perp c})=\frac{\sqrt{2m_c}}{\hbar}\sqrt{W(x)+E+E_{\perp c}} \tag{6.18}$$

The $k_s$ and $k_d$ integrals in Eq. (6.16) can be done analytically [6.18], but the $k_{ch}$ integral needs be evaluated numerically. The lower limits of the WKB integrals, $l_0$ and $l_1$, are solved from the intercepts of $-E + E_{\perp v}$ with $U(x)$ and $V(x) - V_3$. The upper limits, $l_2$ and $l_3$, are solved from the intercepts of $-E - E_{\perp c}$ with $V(x)$ and $W(x)$, respectively. Note that $l_0 < 0$, $l_1 \geq 0$, and $l_2 \leq L$. Not all three integrals are present in $T(E, E_{\perp v})$ of Eq. (6.16). They depend

on the tunneling energies and the bias conditions. For $-E + E_{\perp v} < -\Delta$ and $-E - E_{\perp c} > -V_2 + \Delta'$, only the $k_{ch}$ integral is in the tunneling path.

The 3D ballistic TFET current is given by [6.9]:

$$j = \frac{q m_v}{2\pi^2 \hbar^3} \int_0^{V_2} \left( f_s - f_d \right) \left[ \int_0^{E_{\perp m}} T(E, E_{\perp v}) dE_{\perp v} \right] dE \tag{6.19}$$

where the $f_s(E)$ and $f_d(E)$ are the state occupancy factors at source and drain. $E_{\perp m}$ is the smaller of $E$ and $(m_c/m_v)(V_2 - E)$. This model has been found to be in agreement with the atomistic simulations shown in Fig. 7.4 of Chapter 7.

Fig. 6.9 plots the model generated $I_{ds}$-$V_{gs}$ characteristics for several channel lengths. The parameters assumed are generally in the range of those for an AlGaAsSb source and InGaAs channel. The on current is essentially independent of $L$, since its contribution is mainly from source-to-channel tunneling, or $E \in (0, V_0 - V_1 + \Delta)$. But the off current, dominated by source-to-drain tunneling for $E \in (V_0 - V_1 + \Delta, V_2)$ is very sensitive to $L$. The $L$ = 10 nm TFET exhibits unacceptably high $I_{off}$ and SS for a high drain doping of $N_d$ = $3\times10^{19}$ cm$^{-3}$. By decreasing the drain doping to $2\times10^{18}$ and $10^{17}$ cm$^{-3}$, it is shown in Fig. 6.10 that the added depletion width $W_d$' to the source-drain tunneling path greatly reduced the off current in the 10 nm TFET with very little impact to the on current. The SS in the $N_d = 10^{17}$ cm$^{-3}$ case is also improved to ~ 28 mV/decade. The inset to Fig. 6.10 shows very high output resistance in the saturation region of the 10 nm TFET, indicating diminished drain induced barrier lowering (DIBL).

The series resistance of the lightly doped drain is not a serious issue if its length is limited to ~ 10 nm or so, by backing it up with a $n^+$ contact. For InGaAs, for example, the resistivity is 0.007 Ω-cm for n-type doping of $10^{17}$ $cm^{-3}$. A region 5 nm thick and 10 nm long adds a resistance of ~ 140 Ω-μm on the drain side, with little impact on the saturation current.

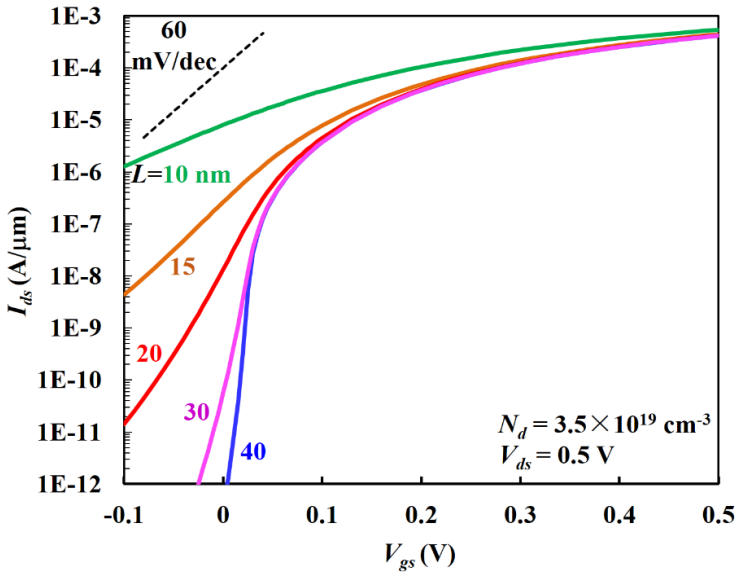



Figure 6.9: $I_{ds}$-$V_{gs}$ characteristics for varying gate lengths with a high drain doping of $N_d$ = $3.5\times10^{19}$ $cm^{-3}$. Other parameters are $V_1$ = 0.15 eV, $V_3$ =1 eV, $\lambda$ = 9 nm, $m_c = m_v = 0.1m_0$. For the source, $N_a = 3.5\times10^{19}$ $cm^{-3}$, $N_v = 2\times10^{19}$ $cm^{-3}$. $N_c = 10^{17}$ $cm^{-3}$ for the drain.

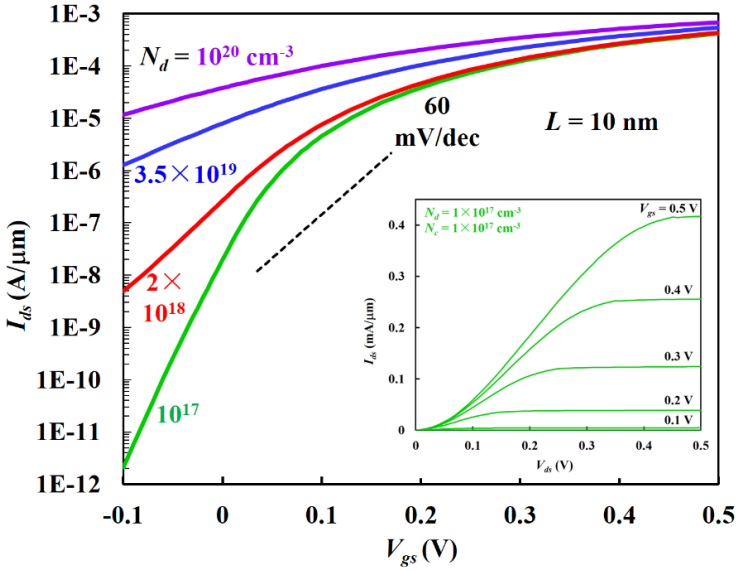



Figure 6.10: $I_{ds}$-$V_{gs}$ characteristics for a fixed $L$ = 10 nm, versus a range of drain doping. Other parameters are the same as those in Fig. 6.9. The inset shows $I_{ds}$-$V_{ds}$ for the $N_d = 10^{17}$ cm$^{-3}$ case.

When the channel bandgap $V_3$ is below $qV_{dd} + d_1 + d_2$, ambipolar effect or channel-to-drain tunneling comes into play at $V_{gs} \leq 0$, as shown in Fig. 6.11. Lightly doped drain is very effective in suppressing ambipolar effect as it increases the tunneling distance from the channel valence band to the drain conduction band. Note that ambipolar effect is insensitive to channel length because it takes place near the drain junction. This is unlike the off current, for which a longer $L$ would help, at the price of density in a way similar to the lightly-doped drain. But a longer $L$ does not reduce ambipolar effect [dotted curve in Fig. 6.11], while the lightly-doped drain does.

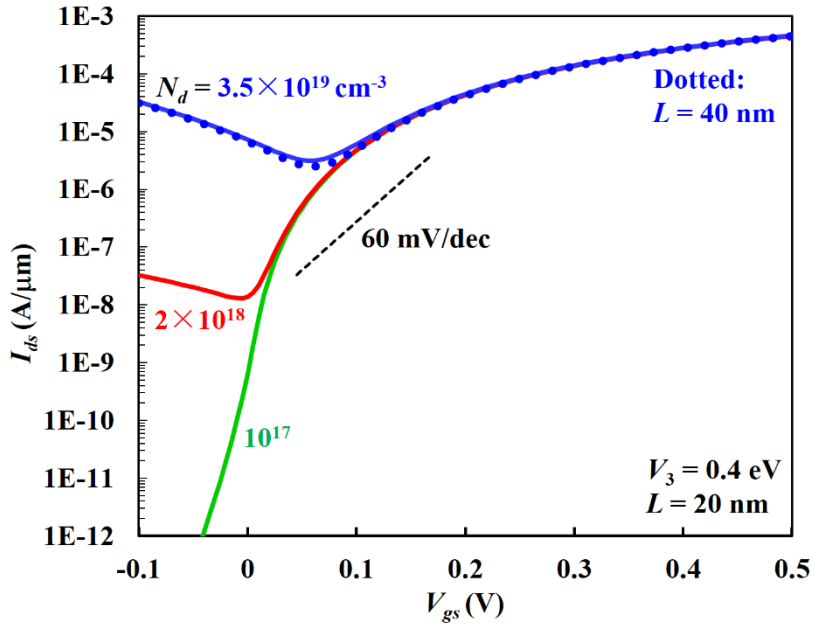



Figure 6.11: $I_{ds}$-$V_{gs}$ characteristics for varying $N_{d.}$ with fixed $V_3$ = 0.4 eV. All solid curves are for $L$ = 20 nm, while the dotted curve is for $L$ = 40 nm. Other parameters are the same as in Fig. 6.9.

To generate the $I_{ds}$-$V_{ds}$ characteristics in the linear region, a de-bias model is invoked to account for the retardation of gate effect by inversion charge, namely, through $q(V_{gs} - Q_{inv}/C_{ox}) = V_0 - (V_1 - \Delta)$ [6.4]. When $-V_2 \geq -V_0 + V_1 - \Delta$, i.e., when the $E_c$ of drain is higher than the $E_c$ of channel, there is no depletion charge in the drain. In that case, $\Delta'$ is set to zero and $\Delta$ is simply solved from Eq. (6.13).

A possible drawback of the lightly doped drain is non-degeneracy if the density of states of the drain is not low enough. Since the tunneling window is bounded by $V_2 = qV_{ds} + d_1 + d_2$, a negative $d_2$ could cause a delayed opening of the tunneling window until $V_{ds} > -(d_1 + d_2)/q$. This is shown in Fig. 6.12(a) where $d_2$ = –0.06 eV for $N_d = 10^{17}$ cm$^{-3}$ and $N_c = 10^{18}$ cm$^{-3}$. The delayed rise of $I_{ds}$ with $V_{ds}$ causes superlinearity hence lower currents in

the linear region. The problem can be cured by using a lower density of states, $N_c = 10^{17}$ $cm^{-3}$, in Fig. 6.12(b). Here, the channel is assumed to be the same material as the drain, so the extent of de-bias is also reduced [6.2]. A good choice of the material is InGaAs, for which $N_c = 8.7\times10^{16}$ $cm^{-3}$.

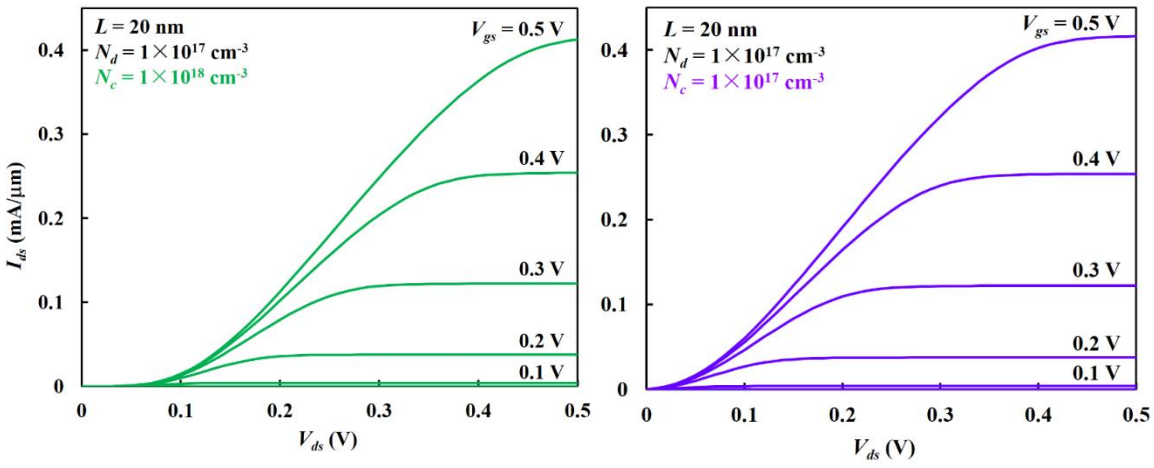



Figure 6.12: $I_{ds}$-$V_{ds}$ characteristics for the same drain doping, $N_d = 10^{17}$ $cm^{-3}$, but different $N_c$, the effective density of states of the drain, resulting in $d_2 = -0.06$ eV for (a) and 0.01 eV for (b). For both (a) and (b), $d_1 = 0.03$ eV.

The text of Chapter 6, in part, is a reprint of the material as it appears in "A continuous, semi-analytic current model for DG and NW TFETs" by Jianzhi Wu and Yuan Taur, IEEE Transaction on Electron Devices, Feb. 2016. The dissertation author was the primary investigator and author of this paper.

The text of Chapter 6, in part, is a reprint of the material as it appears in "Reduction of TFET off current and subthreshold swing by lightly doped drain" by Jianzhi Wu and Yuan Taur, IEEE Transaction on Electron Devices, August 2016. The dissertation author

was the primary investigator and author of this paper.

## 6.5 Reference

[6.1] J. Wu, J. Min, and Y. Taur, “Short channel effects in tunnel FETs”, *IEEE Transactions on Electron Devices,* vol. 62, no. 9, pp. 3019-3024, September 2015.

[6.2] Y. Taur, J. Wu and J. Min, “An analytic model for heterojunction tunnel FETs with exponential barrier,” *IEEE Trans. Electron Devices*, vol.62, no.5, pp. 1399-1404, May 2015.

[6.3] B. Yu, L. Wang, Y. Yuan, P. M. Asbeck, and Y. Taur, “Scaling of nanowire transistors,” *IEEE Trans. Electron Devices*, vol. 55, no. 11, pp. 2846–2858, Nov. 2008.

[6.4] J. Wu, J. Min, and Y. Taur, “Short channel effects in tunnel FETs”, *IEEE Transactions on Electron Devices,* vol. 62, no. 9, pp. 3019-3024, September 2015.

[6.5] W. Franz, *Ann. Phys.* (Germany) (6) 11, 17 (1952).

[6.6] Y-Y. Fan, S. P. Mudanai, W. Chen, L. F. Register, and S. K. Banerjee. "High-k Gate Dielectric Materials Integrated Circuit Device Design Issues." In High Dielectric Constant Materials, pp. 570. Springer Berlin Heidelberg, 2005.

[6.7] *TCAD Sentaurus Device Manual* (Synopsys, Inc., Mountain View, CA, 2013).

[6.8] Y. Taur and J. Wu, "Examination of two-band E(k) relations for band-to-band tunneling", *IEEE Trans. Electron Devices*, vol. 63, no. 2, pp. 869-872, Feb. 2016.

[6.9] E. O. Kane, "Theory of tunneling." *Journal of Applied Physics* 32, no. 1 (1961): 83-91.

[6.10] J. Wu, J. Min, J. Ji, and Y. Taur. "An analytic model for heterojunction and homojunction tunnel FETs with 3D density of states." In *Device Research Conference (DRC), 2015 73rd Annual*, pp. 249-250. IEEE, 2015.

[6.11] Y. Taur, J. Wu and J. Min, "Dimensionality dependence of TFET performance down to 0.1 V supply voltage", *IEEE Trans. Electron Devices* vol. 63, no. 2, pp. 877-880, Feb. 2016.

[6.12] Q. Zhang, Y. Lu, C. A. Richter, D. Jena and A. Seabaugh, "Optimum bandgap and supply voltage in tunnel FETs", *IEEE Trans. Electron Devices*, vol. 61, no. 8, pp. 2719-2724, 2014

[6.13] S. S. Sylvia, H.-H. Park, M. A. Khayer, K. Alam, G. Klimeck and R. K. Lake, "Material selection for minimizing direct tunneling in nanowire transistors", *IEEE Trans. Electron Devices*, vol. 59, no. 8, pp. 2064-2069, 2012

[6.14] E. Gnani, A. Gnudi, S. Reggiani, and G. Baccarani. "Drain-conductance optimization in nanowire TFETs by means of a physics-based analytical model." *Solid-State Electronics* 84 (2013): 96-102.

[6.15] K. Boucart and A. M. Ionescu, "Double-gate tunnel FET with high-k gate dielectric ", *IEEE Trans. Electron Devices*, vol. 54, no. 7, pp. 1725-1733, 2007.

[6.16] T. Krishnamohan , D. Kim , S. Raghunathan and K. C. Saraswat, "Double gate strained-Ge heterostructure tunneling FET (TFET) with record high drive current and < 60 mV/dec subthreshold slope", *IEDM*, pp. 947-949, 2008.

[6.17] E.-H. Toh, G.H. Wang, G. Samudra and Y.-C. Yeo, "Device physics and design of germanium tunneling field-effect transistor with source and drain engineering for low power and high performance applications", *J. Appl. Phys.*, vol. 103, no. 10, pp. 104 504, 2008.

[6.18] J. Min, J. Wu and Y. Taur, "Analysis of Source Doping Effect in Tunnel FETs With Staggered Bandgap," *IEEE Elec. Dev. Lett.* , vol. 36, pp. 1094-1096, 2015.

# Chapter 7 Survey of hardware data and model correlations

## 7.1 Hardware data of 14nm FinFET (Natarajan et al, IEDM'14)

In IEDM 2014, Intel announced the 14 nm logic technology with 2nd generation FinFET transistors [7.1]. The $I_{ds}$-$V_{gs}$ characteristics are shown in Fig. 7.1. In this section, we compare our TFET model predicted $I$-$V$ performance with that of the 14 nm FinFETs.

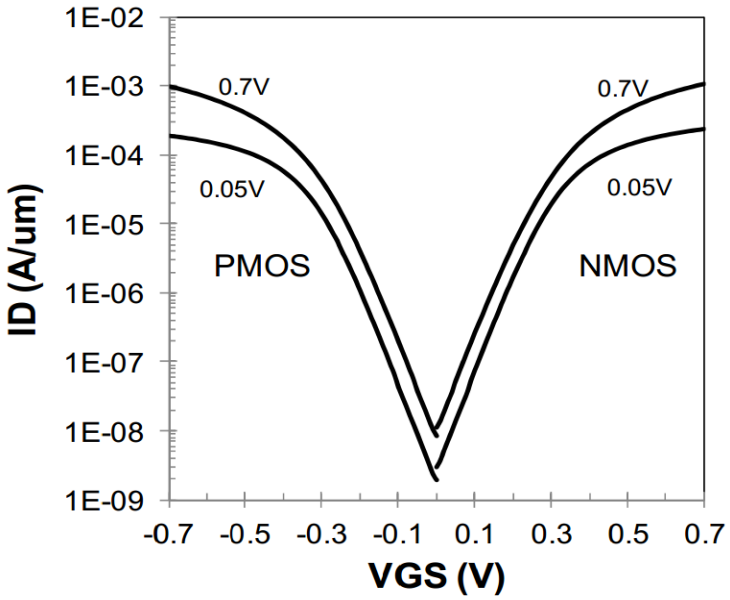
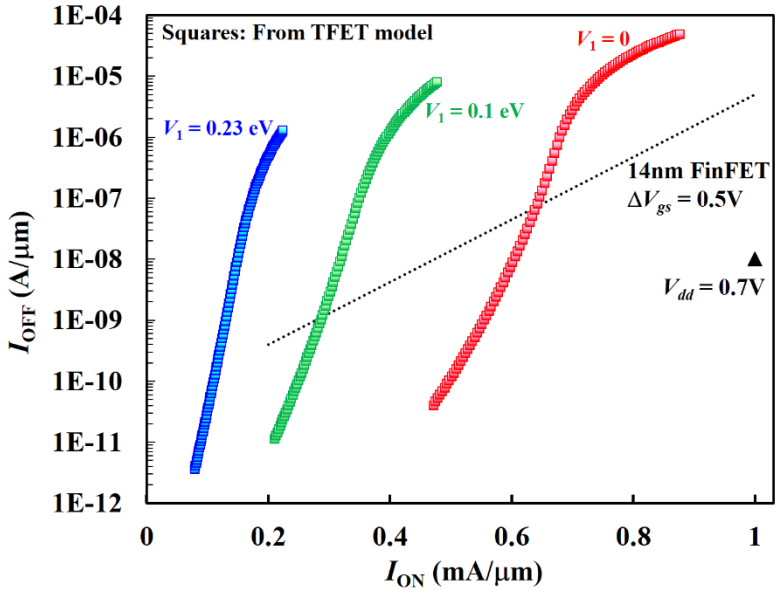



Figure 7.1: $I_{ds}$-$V_{gs}$ characteristics of the 14 nm 2nd generation FinFET. The device dimensions are $W_{fin}$ = 8 nm, Height = 42 nm. Adapted from [7.1].

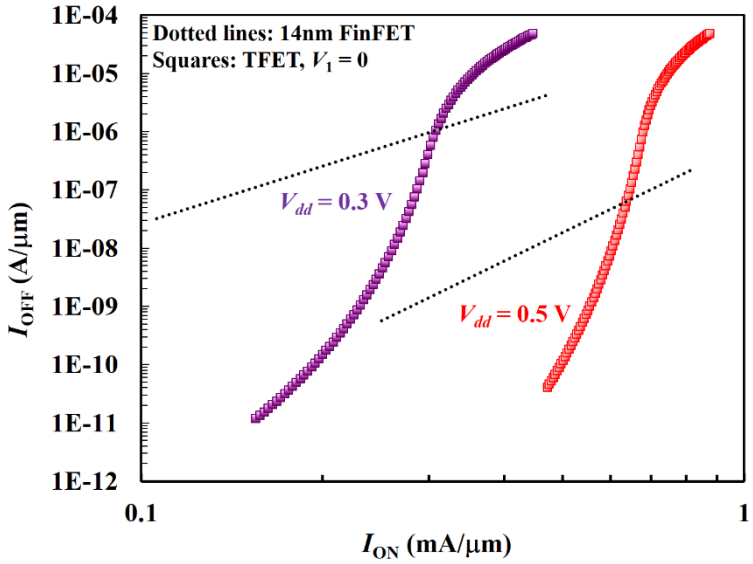
Fig. 7.2(a) shows the $I_{off}$-$I_{on}$ plot generated by our TFET model, along with a benchmark of the $I_{off}$-$I_{on}$ data from 14 nm FinFETs [7.1]. Note that no model or threshold voltage is needed for the $I_{off}$-$I_{on}$ plot. The tradeoff between $I_{on}$ and $I_{off}$ is self-generated by the data. Fig. 7.2(a) makes it clear that the narrower the effective bandgap $V_1$, the higher the TFET performance. For a standard specification of $I_{off}$ = 1 nA/μm, $I_{on}$ = 0.13, 0.28, 0.56 mA/μm for $V_1$ = 0.23, 0.1, 0 eV, respectively. When the $I_{off}$-$I_{on}$ curve of a TFET intersects

with the dotted MOSFET line, the TFET performs better below the crossover, whereas about it the MOSFET is better. The steeper slopes of the TFET curves are indication that their subthreshold swings (SS) are smaller than MOSFETs. When the supply voltage is reduced from 0.5 V to 0.3 V, as shown in Fig. 7.2(b), TFETs are far superior to MOSFETs. In fact, the 60 mV/decade SS limits the MOSFET $I_{on}/I_{off}$ ratio to $10^5$ for a $V_{dd}$ of 0.3 V. With an $I_{off}$ specification of 1 nA/μm, the maximum $I_{on}$ MOSFETs can have is 0.1 mA/μm.




(a)

Figure 7.2: (a) $I_{off}$-$I_{on}$ plots generated from the TFETs $I_{ds}$-$V_{gs}$ characteristics for $V_{dd}$ = 0.5 V. The $V_{dd}$ = 0.7 V point is from published data of 14 nm FinFET [7.1]. The dotted line is generated by sliding a $\Delta V_{gs}$ = 0.5 V window across the FinFET data with $V_{ds}$ = 0.7 V.




(b)

Figure 7.2: continued (b) $V_{dd}$ = 0.5 V line and curve are the same as in (a). $V_{dd}$ = 0.3 V TFET curve is from the model with $V_{ds}$ = 0.3 V and $\Delta V_{gs}$ = 0.3 V. The $V_{dd}$ = 0.3 V dotted line is from the same FinFET data, but with $\Delta V_{gs}$ = 0.3 V.

## 7.2 Atomstic simulations of GaSb-InAs TFET (Luisier et al, IEDM'09)

The performance of GaSb-InAs broken gap double-gated tunnel FETs has been reported by Luisier et al [7.2] utilizing a ballistic, atomistic, full-band, quantum transport simulator.

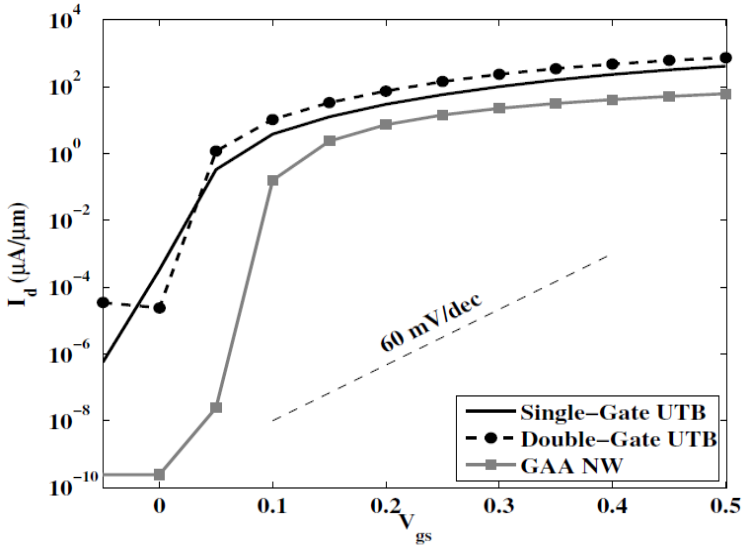



Figure 7.3: $I_{ds}$-$V_{gs}$ characteristics for a GaSb-InAs broken gap TFET. $L_g$ = 40 nm, $N_a$ = 4 $\times 10^{19}$ cm$^{-3}$, $N_d$ = 1$\times 10^{18}$ cm$^{-3}$, $V_{dd}$ = 0.5 V. Adapted from [7.2].

This is a broken gap system in bulk form. When the semiconductor film is 5 nm thick in a DG configuration, quantization effect raises the electron subband energy above the bulk $E_c$ so the TFET is close to the $V_1$ = 0 condition. Quantization also mitigates the ambipolar effect. The low drain doping in [7.2] effectively adds a depleted region to their 40 nm gate length so the source-to-drain distance $L$ is more like 60 nm. This is mainly for the purpose of $I_{off}$ reduction as $I_{on}$ is totally insensitive to $L$. Fig. 7.4 shows that the $I_{ds}$-$V_{gs}$ characteristics calculated from our analytic model are largely consistent with those from the full band quantum simulation [7.2].

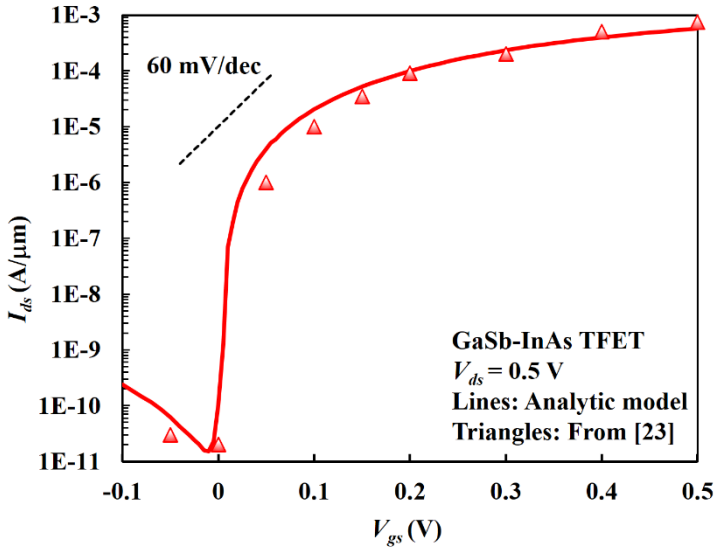



Figure 7.4: Comparison of GaSb-InAs TFET characteristics generated by the analytic model and by full-band quantum simulations [7.2].

## 7.3 Hardware data of $I_{ds}$-$V_{ds}$ characteristics (superlinear turn-on)

Fig. 7.5 and 7.6 show the experimentally measured $I_{ds}$-$V_{ds}$ characteristics for a strained SiGe/Si TFET and a B-$Si_{1-x}Ge_x$ heterojunction n-TFET, respectively. Both of them exhibit the superlinear onset in the linear region output characteristics. Hence, these experimental results qualitatively validate our model predicted superlinear behavior due to de-bias effects, as explained in chapter 3.

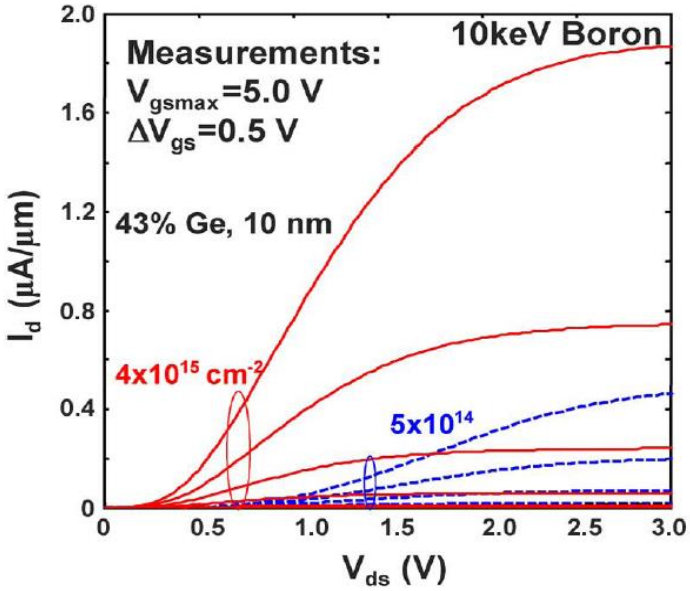



Figure 7.5: Experimentally measured $I_{ds}$-$V_{ds}$ characteristics for a strained SiGe/Si TFET with varying implant dose (10-keV boron). Adapted from [7.3].

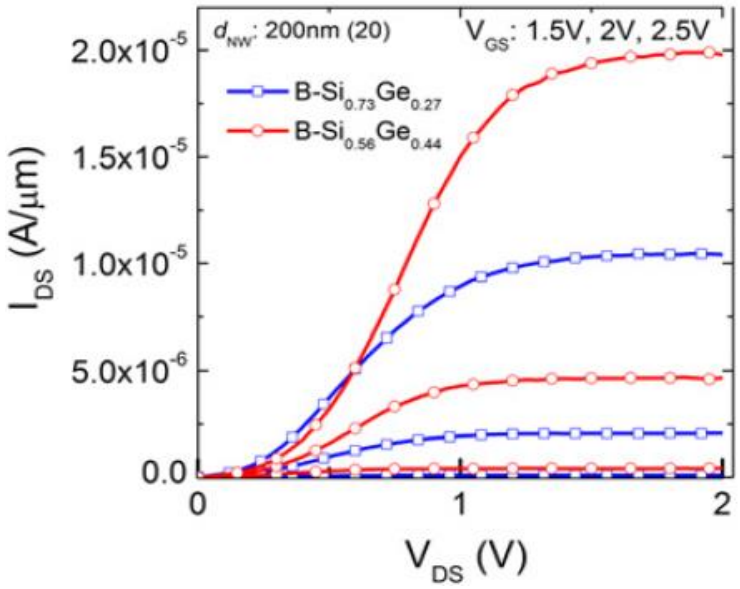



Figure 7.6: Experimental $I_{ds}$-$V_{ds}$ characteristics for a B-$Si_{1-x}Ge_x$ heterojunction n-TFET featuring an array of 20 nanowires of $d_{NW}$ = 200 nm. Adapted from [7.4].

## 7.4 Homojunction TFETs

In [7.5], Choi et al experimentally demonstrated, for the first time, a silicon-based TFET featuring sub-60-mV/dec subthreshold swing at room temperature.

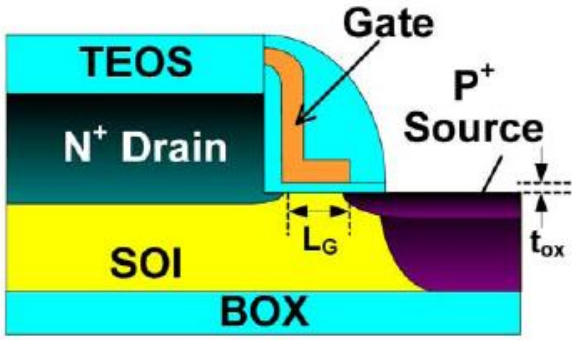



Figure 7.7: Schematic illustration of the n-channel Si TFET. Adapted from [7.5].

The schematic device structure of the Si TFET is shown in Fig. 7.7. A gated p-i-n structure is fabricated on a silicon-on-insulator (SOI) substrate.

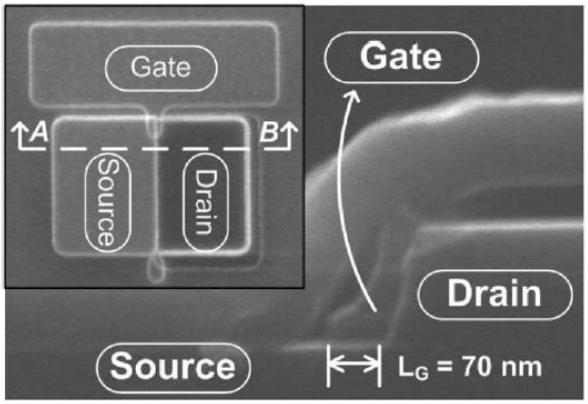



Figure 7.8: SEM image (cross-section) of the fabricated Si TFET. The gate length is measured as 70 nm. Adapted from [7.5].

The gate length of the device is 70 nm, oxide thickness is 2 nm, and SOI substrate is 70 nm as shown in Fig. 7.8.

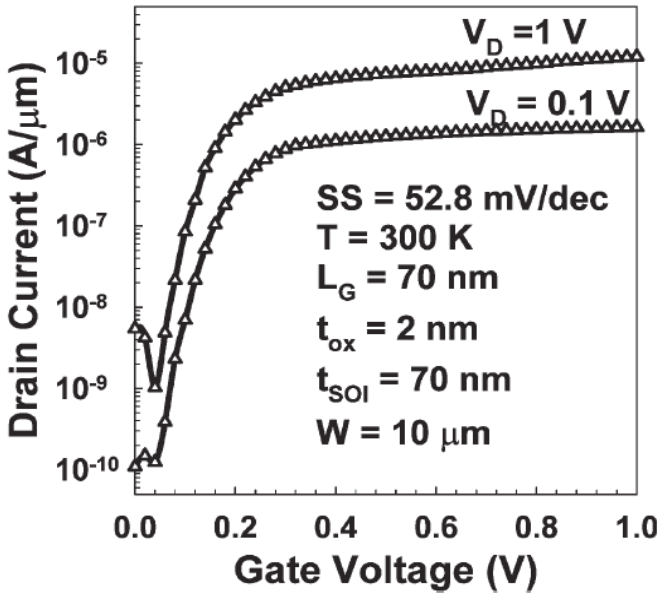



Figure 7.9: $I_{ds}$-$V_{gs}$ characteristics of the fabricated TFET. Adapted from [7.5].

Fig. 7.9 shows the measured $I_{ds}$-$V_{gs}$ characteristics. The observed minimum subthreshold swing (S.S.) value is 52.8 mV/dec at room temperature. Yet, the on-current is only 12.1 μA/μm at $V_{dd}$ = 1.0 V operation, which is far too low compared to that of conventional MOSFETs. The on-current may be further improved by optimizing the source doping as illustrate in Chapter 3. But the inherent reason for this Si-TFET suffering from low on-current is its indirect and wide bandgap, and the heavy tunneling mass. Therefore, III-V materials, e.g. GaSb, InAs, InGaAs, outstand for their narrower, direct bandgaps and light tunneling mass, which are very promising to achieve high tunneling probability.

In [7.6], a vertical InGaAs based TFET is reported by Zhao et al as shown in Fig. 7.10.

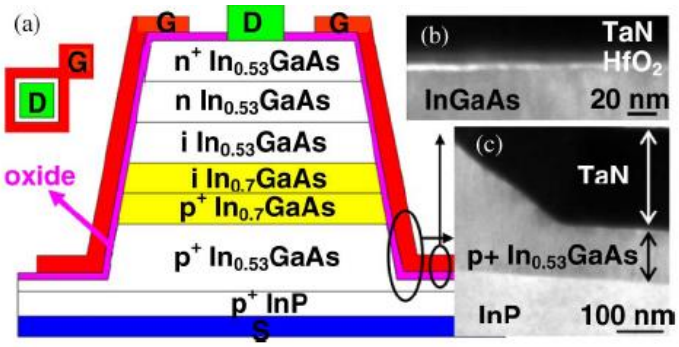



Figure 7.10: (a) Schematic view of the InGaAs vertical TFET with TaN gate and $HfO_2$ gate dielectric (b) TEM image of the TaN/$HfO_2$/InGaAs interface. (c) TEM image of the sidewall. Adapted from [7.6].

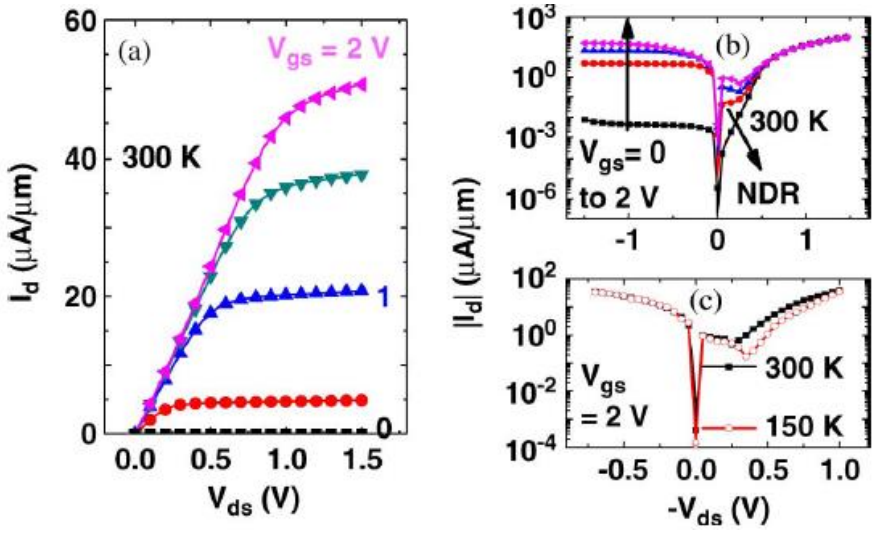



Figure 7.11: (a) $I_{ds}$-$V_{ds}$ characteristics for InGaAs TFET at 300K. (b) log$I_{ds}$-$V_{ds}$ characteristics for varying $V_{gs}$ at 300K. (c) log$I_{ds}$-$V_{ds}$ characteristics of the same device for varying temperature. Device dimension: $W$ = 560 μm, $L$ = 100 nm. Adapted from [7.6].

The measured $I_{ds}$-$V_{ds}$ characteristics are shown in Fig. 7.11. When the P-N diode is forward biased ($-V_{ds} > 0$), the Esaki diode behavior [7.13] is exhibited as negative differential resistance (NDR) in Fig. 7.11 (b) and (c). Thus the band-to-band tunneling

mechanism is confirmed. Furthermore, the on-current is 20 μA/μm at $V_{dd}$ = 1V, which is 65% higher than the Si-TFET discussed before in [7.5]. Therefore the benefits of having a narrower, direct bandgap and smaller tunneling mass for InGaAs are evident from the boosted on-current.

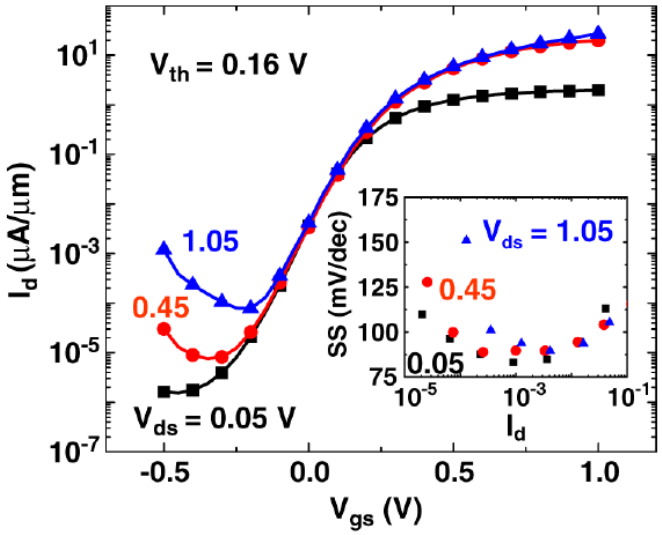



Figure 7.12: $I_{ds}$-$V_{gs}$ characteristics of InGaAs TFET for varying $V_{ds}$. The inset figure shows the SS as a function of $I_{ds}$ at different $V_{ds}$. Adapted from [7.6].

The $I_{ds}$-$V_{gs}$ characteristics for the fabricated InGaAs TFET with equivalent oxide thickness (EoT) = 1.2 nm is shown in Fig. 7.12. The minimum SS values are extracted as 86 mV/dec and 93 mV/dec at $V_{ds}$ = 0.05 and 1.05 V, respectively. Unfortunately, those SS values are still no steeper than the thermionic limit 60 mV/dec. This is probably caused by the interface trap assisted tunneling and the subsequent thermal emission. The conductance method extracted midgap interface trap density for the $HfO_2$/InGaAs interface is around 4 $\times 10^{12}$ /eV/cm$^2$. Therefore the oxide/III-V material interface needs to be improved to achieve the sub-60 mV/dec SS.

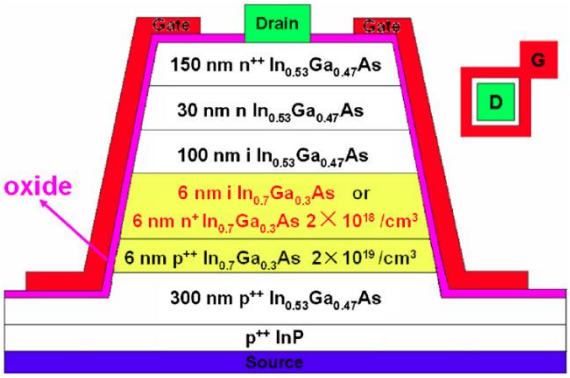



Figure 7.13: Cross-section of InGaAs TFET with p++/i or p++/n+ tunneling junction. Gate electrode is TaN and gate dielectric is $HfO_2$. Adapted from [7.7].

To further improve the on-current, Zhao et al demonstrated to use a tunneling junction formed between p++ layer and a thin, fully depleted n+ layer as shown in Fig.7.13 [7.7].

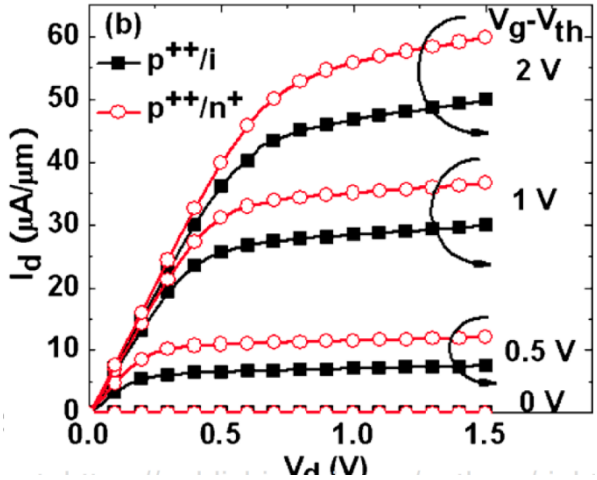



Figure 7.14: $I_{ds}$-$V_{ds}$ characteristics of the InGaAs TFET with p++/i and p++/n+ tunneling junctions. Adapted from [7.7].

The resultant $I_{ds}$-$V_{ds}$ contrast between the p++/n+ and p++/i junctions are shown in Fig. 7.14. The p++/n+ junction boosts the on-current by 61% and 20% at $V_{gs}$-$V_t$ = 0.5 V, and 2 V, respectively. Such a current enhancement is believe to attribute to the n+ layer near the tunneling junction, which increases the electric field near the junction. It confirms our model finding that high junction field gives shorter tunneling distance, thus yields higher tunneling probability in Chapter 3.

## 7.5 Staggered bandgap TFETs

As summarized in section 7.4, steep switching slope has been observed for Si TFETs in [7.5], while high drive current is achieve by InGaAs TFETs [7.7], yet not simultaneously. In [7.8], Rajamohanan et al demonstrated a staggered bandgap InGaAs/GaAsSb TFET featuring both high on-current and steep subthreshold swing.

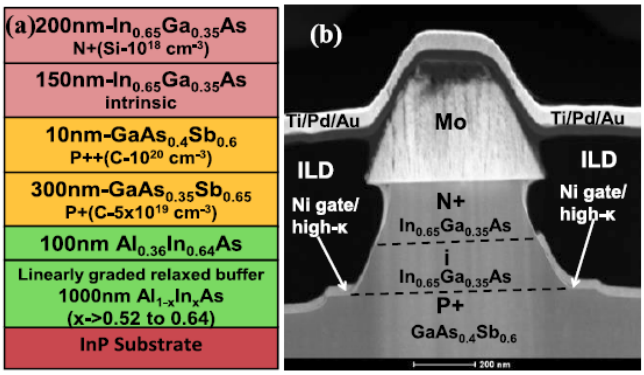



Figure 7.15: (a) Cross-section view of the staggered bandgap TFET. (b) TEM image of the fabricated staggered bandgap TFET. Adapted from [7.8].

The layer structure of the staggered-gap InGaAs/GaAsSb TFET is shown in Fig. 7.15. The effective bandgap of the InGaAs/GaAsSb heterojunction is estimated to be ~ 0.31 eV. Three types of gate stacks are employed in the fabrication of TFETs. (1) 4 nm $HfO_2$/Pd gate (2) 4 nm HfO2/Ni gate and (3) 3 nm HfO2/Ni gate. The Pd gate metal of (1)

is deposited by e-beam evaporator, while the Ni electrode of both (2) and (3) are formed by thermal evaporation.

Fig. 7.16 plots the S.S. as a function of drain current for the TFETs with three types of gate stacks. The improved S.S. from 130 mV/dec of gate stack (1) to 105 mV/dec of gate stack (2) is believed to be due to the different gate metallization processes. The e-beam evaporation generated X-ray can damage the oxide/semiconductor interface. The frequency dependent C-V measurement in Fig. 7.17 confirms that the thermal evaporated gate stack (2) results in less frequency dispersion, hence less traps. This explains why the gate stack (2) achieves better SS performance than (1) in Fig. 7.16. Furthermore, it is shown in Fig. 7.16 that the gate stack (3) further improves the SS to 97 mV/dec due to thinner $HfO_2$. This has been well explained by our analytic model (in Chapter 2) that the thinner oxide gives shorter scale length ($\lambda = t_s + 2t_{ox}$), thus the better electrostatic control.

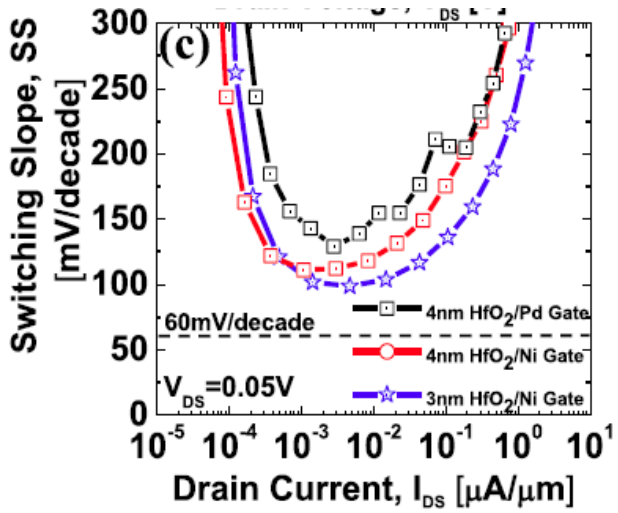



Figure 7.16: Measured SS versus drain current for varying gate stacks. Adapted from [7.8].

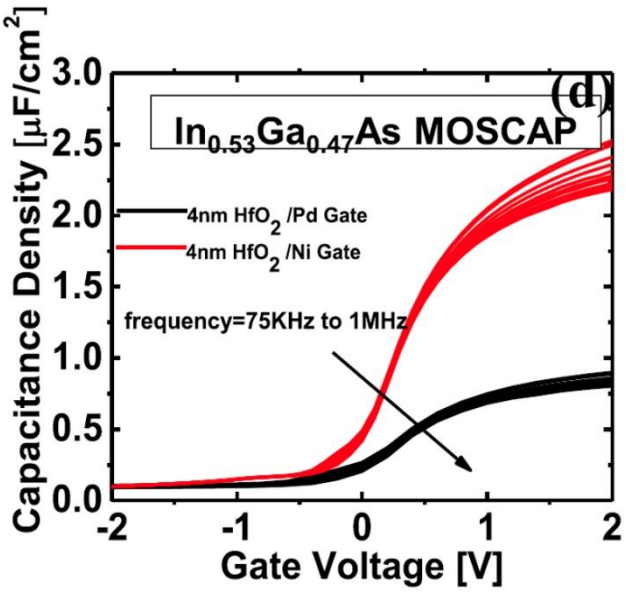



Figure 7.17: Measured C-V characteristics of the InGaAs MOSCAP. Adapted from [7.8].

To further enhance the S.S. by eliminating the trap assisted tunneling, fast $I_{ds}$-$V_{gs}$ measurement is performance and compared with DC *I-V* characteristics as shown in Fig. 7.18.

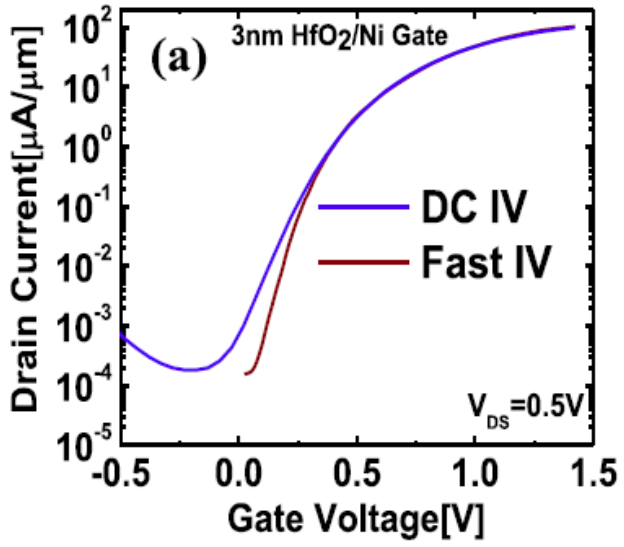
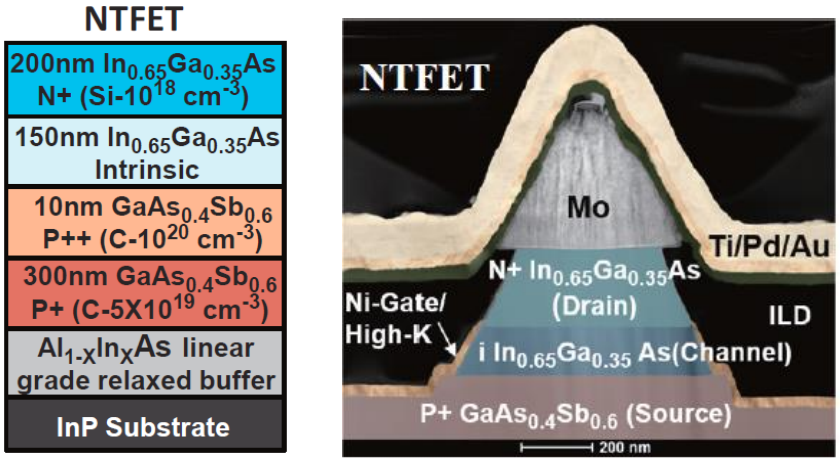



Figure 7.18: $I_{ds}$-$V_{gs}$ characteristics of the TFET measured from fast-IV and DC at room temperature. Adapted from [7.8].

With the gate pulse rise time of 1 $\mu$s, the fast I-V shows a minimum SS of 64 mV/dec over a $I_{ds}$ range between $10^{-3}$ and $2\times10^{-2}$ μA/μm, at $T$ = 300 K. The on-current is measured as high as 8.4 μA/μm at $V_{dd}$ = 0.5V. The improvement of on-current is believed to benefit from the low tunneling barrier of the staggered-gap tunneling junction, which has been covered by our analytic model in Chapter 3.

More recently, in [7.9] Pandey et al reported sub-$kT/q$ switching slope for the n-TFETs with the similar staggered-gap structure as in [7.8].




(a) (b)

Figure 7.19: Cross-section (a) schematic (b) TEM image of the fabricated n-TFET. Adapted from [7.9].

The layer structure of the fabricated n-TFET is shown in Fig. 7.19. The gate dielectric is 4 nm $ZrO_2$ high-k dielectric with CET of 1.1 nm, deposited by ALD with in-situ $N_2$ plasma/TMA clean.

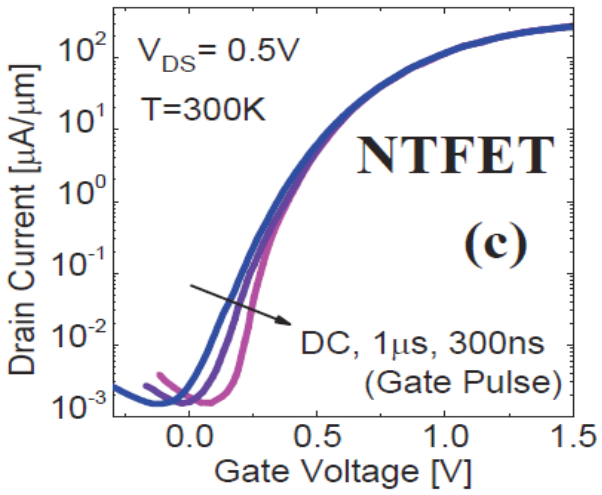



Figure 7.20: Pulsed/DC $I_{ds}$-$V_{gs}$ characteristics of the fabricated n-TFET at room temperature. Adapted from [7.9].

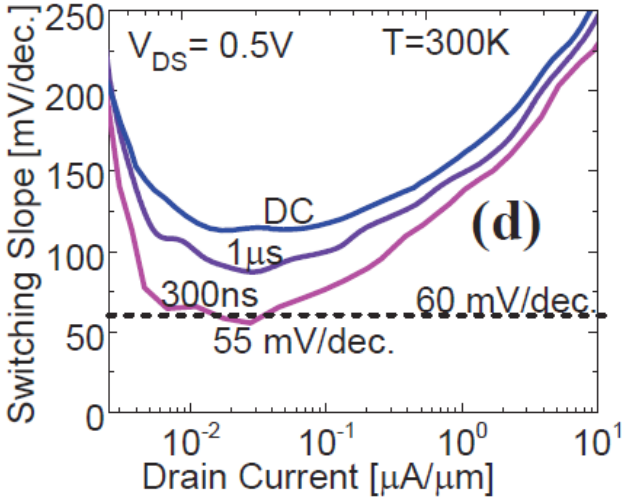



Figure 7.21: Extracted switching slope versus drain current from the measured data in Fig. 7.20. Adapted from [7.9].

The pulsed and DC $I_{ds}$-$V_{gs}$ characteristics is shown in Fig. 7.20. The extracted SS as a function of drain current is then plotted in Fig. 7.21. It shows the switching slope is becoming steeper towards shorter pulsed width, due to the mitigated trap assisted tunneling. The minimum SS achieved for this n-TFET is 55 mV/dec at room temperature, which beats

the *kT*/*q* limit. Such improvement is believed to be owing to the high-quality scaled high-k gate dielectric.

## 7.6 TFETs based on two-dimensional materials

2D semiconductor material is a promising candidate for making nanometer scale TFETs because of its sheet form with atomic layer thickness. This facilities better electrostatic potential control according to the scale length theory covered in Chapter 2. In addition, the 2D material layers can be thinned without increasing the material bandgap, as opposed to their 3D counterparts. Furthermore, the stacked layers of 2D materials are weakly bonded by van der Waals force, hence alleviates the "lattice mismatch" issue while constructing heterojunctions with materials of different lattice constants.

In [7.10], Roy et al reported a dual-gated $MoS_2/WSe_2$ van der Waals tunnel diode.

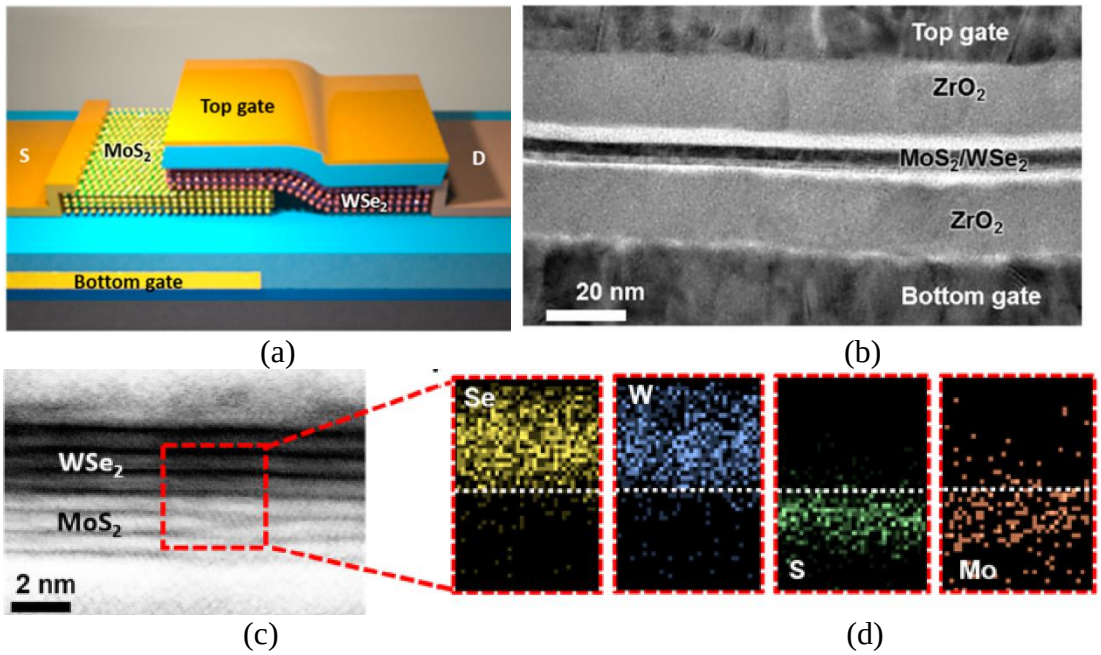



Figure 7.22: (a) Schematic view (b) TEM image of the fabricated $MoS_2/WSe_2$ tunnel diode. (c) High-resolution STEM image (d) EDS mapping of the $MoS_2/WSe_2$ heterostructure. Adapted from [7.10].

The cross-section view of the fabricated $MoS_2/WSe_2$ tunnel diode is shown in Fig. 7.22. The heterostack of $MoS_2/WSe_2$ is formed by dry transfer process [7.11]. Subsequently, Ni electrodes are patterned as electron contacts to $MoS_2$, while Pd electrodes are formed as hole contacts to $WSe_2$. By appropriately tuning the bottom-gate voltage, the $MoS_2$ layer is electrostatically doped as n-type, from strong accumulation of electrons. Likewise, the $WSe_2$ layer is accumulated of holes by biasing the top gate. Thus the device resembles an n+/p+ diode for the $MoS_2/WSe_2$ heterostack.

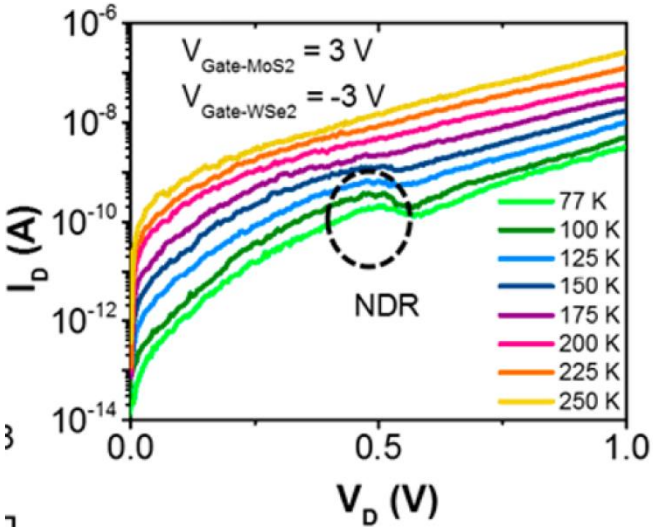
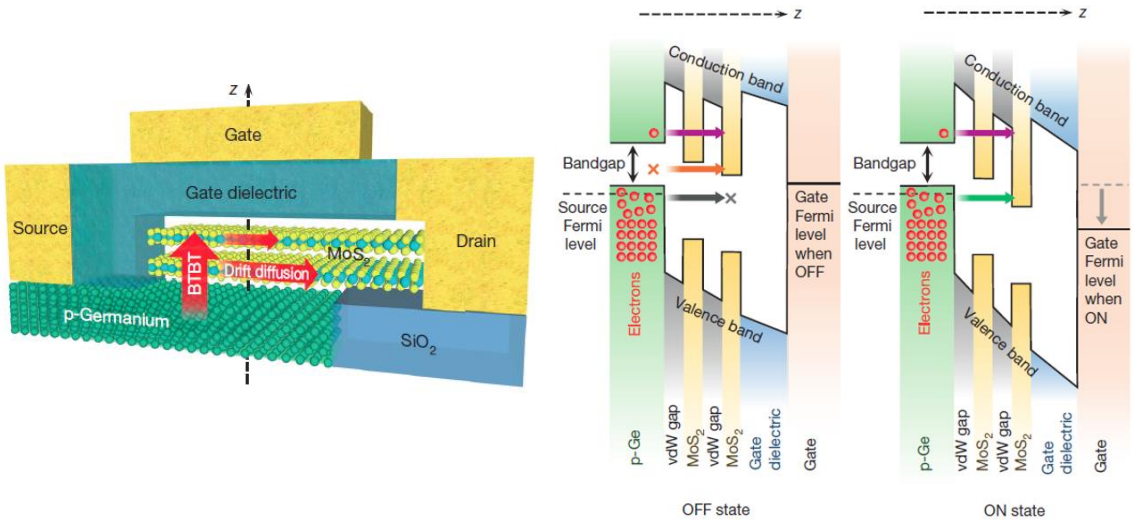



Figure 7.23: Temperature dependent *I-V* characteristics of the $MoS_2/WSe_2$ tunnel diode with 4 layers each in thickness. Adapted from [7.10].

The measured *I-V* characteristics of the $MoS_2/WSe_2$ tunnel diode for varying temperature is shown in the Fig. 7.23. The negative differential resistance (NDR) manifests itself in the forward bias for temperature below 175K. However, for temperatures above 175K, the thermionic current dominates over the tunneling current, thus the Esaki diode like NDR vanishes in the forward *I-V* characteristics. The current gain may need to be further improved by enhancing the gate coupling efficiency.

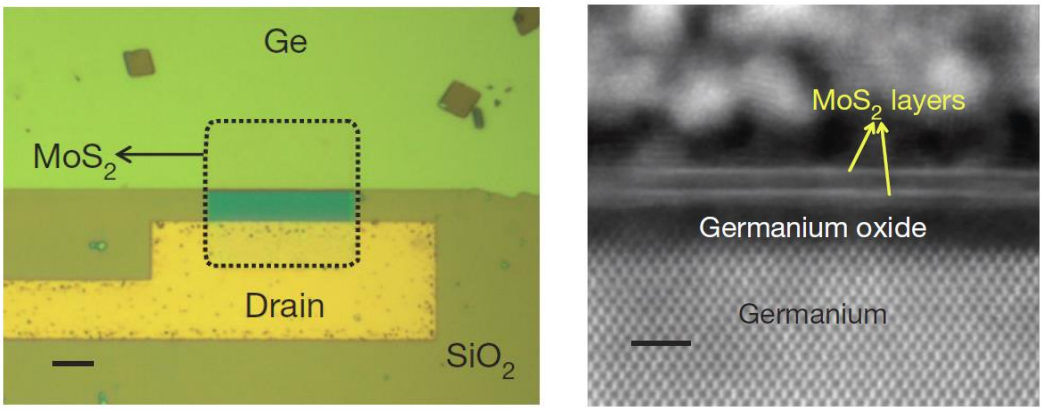
In [7.12], Sarkar et al reported, for the first time, a sub-thermionic tunnel FET with 2D material for channel. The device structure of the atomically thin and layered semiconducting-channel tunnel-FET (ATLAS-TFET) is shown in Fig. 7.24. Highly doped Ge is employed as the p+ source. The channel is formed by a 1.3 nm thick bilayer $MoS_2$. The carriers are supposed to inject from the p+ source to the $MoS_2$ channel through band-to-band tunneling, then be collected by the drain through drift/diffusion process.




(a) (b) (c)

Figure 7.24: (a) Schematic view of the ATLAS-TFET. Band diagram in the (b) OFF and (c) ON state. Adapted from [7.12].

The fabrication process starts with a degenerately p-doped Ge substrate, then etch ~300 nm trenches and subsequently fill with $SiO_2$ dielectric. Ni (20nm)/Au(50nm) source contact is deposited to Ge. The $MoS_2$ is synthesized by CVD and transferred onto the engineered Ge substrate. Solid polymer electrolyte including poly (ethylene oxide) and $LiClO_4$ is utilized as gate dielectric. Y(20nm)/Au(50nm) is formed as the drain contact. The fabricated ATLAS-TFET device is shown in Fig. 7.25.




(a) (b)

Figure 7.25: (a) Optical image of the fabricated ATLAS-TFET (b) TEM image of the Ge/MoS2 heterojunction. Adapted from [7.12].

Without gate yet, Fig. 7.26 shows the measured *I-V* characterisitcics by probing the source and drain contacts only. The trend towards negative differential resistance (NDR) is observed in the forward bias in Fig. 7.26 (b). This confirms the carrier injection mechanism consists of band-to-band tunneling.

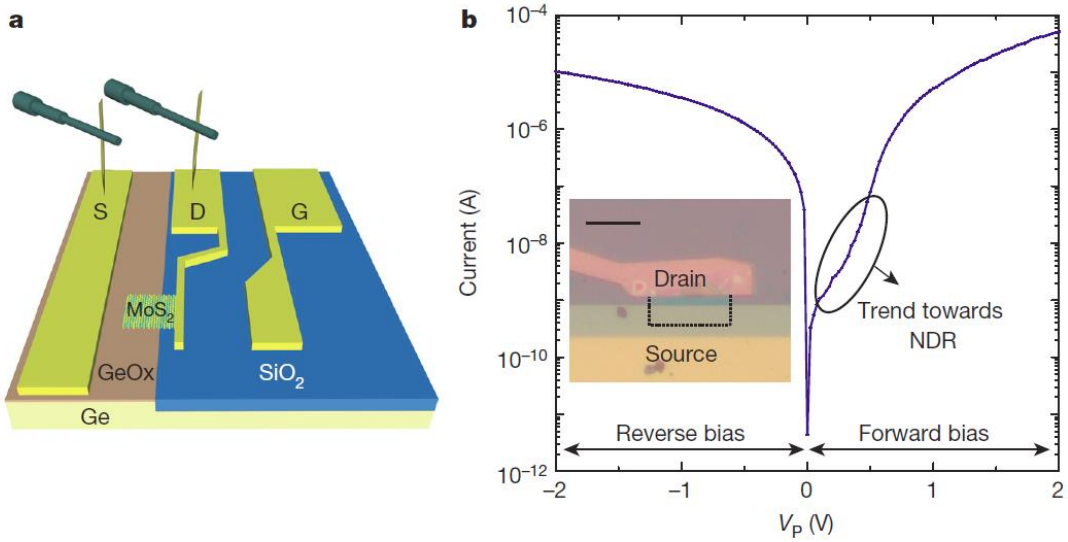



Figure 7.26: (a) Schematic diagram of the probing configuration (b) Measured *I-V* characteristics for the fabricated p-n junction diode at room temperature. Adapted from [7.12].

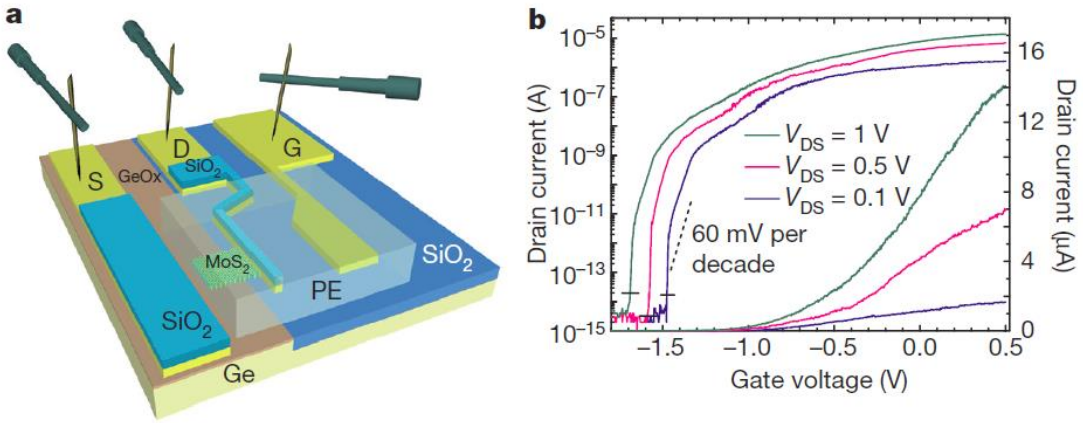



Figure 7.27: (a) Schematic diagram of the probing configuration (b) Measured log (left) and linear (right) $I_{ds}$-$V_{gs}$ characteristics for the fabricated ATLAS-TFET at room temperature. Adapted from [7.12].

Then the measured $I_{ds}$-$V_{gs}$ characteristics at room temperature for the ATLAS-TFET (three-terminal device) is shown in Fig. 7.27. The subthreshold swing below 60 mV/dec is observed to cover ~ 4 decades of current, which go beyond the conventional MOSFET limited by thermionic tail. Thus its operating voltage of $V_{dd}$ = 0.5V is less than $V_{dd}$ = 0.7V of the state-of-the-art MOSFET in [7.1]. However, the on-current for ATLAS-TFET needs to be further boosted to enable fast switching.

## 7.7 Summary

In this chapter, the experimental progress of tunnel FETs has been reviewed. Close to or sub-60 mV/dec turn-off slope has been confirmed for TEFTs with band-to-band tunneling mechanism. The results from the analytic model has been compared to the state-of-the-art FinFET in [7.1] as well as atomistic simulation results of TFETs in [7.2]. The

superlinear onset behavior revealed by analytic model has been qualitatively validated by hardware data in [7.3][7.4]. Our model suggested staggered bandgap design is corroborated by [7.8][7.9], to boost the on-current due to the low effective tunneling barrier. Finally, the recent efforts of TFETs made of 2D semiconductor material has been discussed. Steep subthreshold swing has been observed as is consistent with our model, hence this work paves the way for choosing the appropriate material system for TFETs to achieve high output current while eliminating the undesired superlinear on-set behavior.

The text of Chapter 7, in part, is a reprint of the material as it appears in “A continuous, semi-analytic current model for DG and NW TFETs” by Jianzhi Wu and Yuan Taur, IEEE Transaction on Electron Devices, Feb. 2016. The dissertation author was the primary investigator and author of this paper.

## 7.8 References

[7.1] S. Natarajan, M. Agostinelli, S. Akbar, M. Bost, A. Bowonder, V. Chikarmane, S. Chouksey, A. Dasgupta, K. Fischer, Q. Fu, T. Ghani, M. Giles, S. Govindaraju, R. Grover, W. Han, D. Hanken, E. Haralson, M. Haran, M. Heckscher, R. Heussner, P. Jain, R. James, R. Jhaveri, I. Jin, H. Kam, E. Karl, C. Kenyonv M. Liu, Y. Luo, R. Mehandru, S. Morarka, L. Neiberg, P. Packan, A. Paliwal, C. Parker, P. Patel, R. Patel, C. Pelto, L. Pipes, P. Plekhanov, M. Prince, S. Rajamani, J. Sandford, B. Sell, S. Sivakumar, P. Smith, B. Song, K. Tone, T. Troeger, J. Wiedemer, M. Yang, K. Zhang. "A 14nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588 μm 2 SRAM cell size." *IEEE International Electron Devices Meeting* (*IEDM*), pp. 3-7., 2014.

[7.2] M. Luisier, and Gerhard Klimeck. "Performance comparisons of tunneling field-effect transistors made of InSb, Carbon, and GaSb-InAs broken gap heterostructures." In *IEEE International Electron Devices Meeting* (*IEDM*), pp. 1-4., 2009.

[7.3] O. M. Nayfeh , J. L. Hoyt and D. A. Antoniadis, "Strained- band-to-band tunneling transistors: Impact of tunnel-junction germanium composition and doping concentration

on switching behavior", *IEEE Trans. Electron Devices*, vol. 56, no. 10, pp. 2264-2269, 2009.

[7.4] A. Vandooren, D. Leonelli, R. Rooyackers, A. Hikavyy, K. Devriendt, M. Demand, R. Loo, G. Groeseneken, and C. Huyghebaert, "Analysis of trap-assisted tunneling in vertical Si homo-junction and SiGe hetero-junction Tunnel-FETs," Solid-State Electronics, vol. 83, pp. 50, 2013.

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

[7.6] H. Zhao , Y. Chen , Y. Wang , F. Zhou , F. Xue and J. Lee, "InGaAs tunneling field-effect-transistors with atomic-layer-deposited gate oxides", *IEEE Trans. Electron Devices*, vol. 58, no. 9, pp. 2990-2995, 2011

[7.7] H. Zhao , Y. Chen , Y. Wang , F. Zhou , F. Xue and J. Lee, "Improving the on-current of tunneling field-effect-transistors by p++/n+ tunneling junction", *Appl. Phys. Lett.*, vol. 98, no. 9, pp. 093501, 2011.

[7.8] B. Rajamohanan, R. Pandey, V. Chobpattana, C. Vaz, D. Gundlach, K. P. Cheung, J. Suehle, S. Stemmer, and S. Datta. "0.5 V Supply Voltage Operation of In 0.65 Ga 0.35 As/GaAs 0.4 Sb 0.6 Tunnel FET." *IEEE Electron Device Letters*, 36, no. 1 (2015): 20-22.

[7.9] R. Pandey, H. Madan, H. Liu, V. Chobpattana, M. Barth, B. Rajamohanan, M. Hollander, T. Clark, K. Wang, J.-H. Kim, D. Gundlach, K. Cheung, J. Suehle, R. Engel Herbert, S. Stemmer and S. Datta, "Demonstration of p-type in0.7ga0.3as/gaas0.35sb0. 65 and n-type gaas0.4sb0.6/in0.65ga0.35as complimentary heterojunction vertical tunnel fets for ultra-low power logic", *Proc. VLSI Technol. Symp.*, pp. T206-T207, 2015.

[7.10] T. Roy, M. Tosun, X. Cao, H. Fang, D.-H. Lien, P. Zhao, Y-Z Chen, J. Guo, and A. Javey. "Dual-gated MoS2/WSe2 van der waals tunnel diodes and transistors" ACS Nano, 9 (2) (2015), pp. 2071–2079.

[7.11] T. Roy, "Field-effect transistors built from all two-dimensional material components", *ACS Nano*, vol. 8, no. 6, pp. 6259-6264, 2014.

[7.12] D. Sarkar, X. Xie, W. Liu, W. Cao, J. Kang, Y. Gong, S. Kraemer, P. M. Ajayan, and K. Banerjee, Nature 526,91 (2015).

[7.13] L. Esaki, "New phenomenon in narrow germanium p-n junctions", Phys. Rev., vol. 109, no. 2, pp. 603-604, 1958

# Chapter 8 GaN high electron mobility transistors (HEMTs) on diamond

Besides improving the switching slope, the other option to scale down the supply voltage is to boost the on-current. III-V transistors stand out due to their high electron mobility as well as high saturation velocity. In recent years, GaN power switches attract a great deal of interest for their decent power conversion effecicency at both high power and frequencies. This significantly reduces device complexity and weight, thus leading to lower cost. This chapter primarily demonstrates the experimental results of GaN HEMTs for high power applications.

## 8.1 Background information

Due to the excellent material properties of wide bandgap nitride-semiconductor, high channel electron mobility AlGaN/GaN transistors have demonstrated the potential for high power densities at both low frequency (DC/60Hz) and microwave frequency [8.1]. Although AlGaN/GaN HEMTs with power density exceeding 40 W/mm have been reported for microwave power amplifiers by Y. F. Wu et al [8.2], their performance can be limited by the capability of thermal dissipation of the substrate. When temperature rises, electron mobility and saturated velocity in AlGaN/GaN HEMTs decrease due to optical phonon scattering at the interface of AlGaN/GaN channel [8.3], which limit the maximum output current and power conversion efficiency. SiC, with a high thermal conductivity of ~350 W/$m^{-}K$, is considered as a heterogeneous epitaxial substrate for AlGaN/GaN power HEMTs. An alternate yet promising way of heat management is to replace the substrate with a very high bulk thermal conductivity diamond substrate (>1500 W/m$^{-}$K) [8.4].

AlGaN/GaN HEMT on diamond was firstly demonstrated by G. H. Jessen et al. [8.5]. Due to the superb thermal conductivity of diamond substrates, GaN-on-diamond HEMT was considered to be an excellent candidate for high power operations as compared with GaN-on-SiC HEMTs by J. G. Felbinger et al. [8.6]. More recently, TriQuint reported their GaN-on-diamond transistor technology which enabled a new generation of RF amplifiers with up to three times the power of the present state-of-the-art GaN products [8.7]. Among all these GaN-on-diamond works reported, thermal resistance evaluation is of great interest to researchers. A convenient method for the process-control-measurement (PCM) of thermal resistance of devices on different substrates needs to be developed.

Various methods, such as Three Omega method [8.8]， picosecond time-domain thermoreflectance (TDTR) [8.9] and micro-Raman/infrared thermograph analysis [8.10] have been commonly used for thermal characterizations. In spite of this, direct extraction of thermal resistance from electrical characterization which is fast and does not require special geometry has yet to be reported. In this study, we demonstrated a HEMT device made on diamond substrate and the methodology of thermal resistance extraction based on temperature dependent direct-current (DC) current-voltage (IV) measurements. As a comparison, a HEMT on Si substrate with the same device dimensions was also measured and analyzed. This work extracted thermal resistance of substrate material from GaN-on-diamond HEMT's electrical characteristics and thus quantitatively show how substrates' thermal resistance correlates with the electrical performance of HEMTs. The experimental results are further corroborated by 3D thermal simulation.

## 8.2 Experimental procedures

AlGaN/GaN HEMT device (Sample A) is made from a GaN epitaxial layer test structure on which polycrystalline diamond has been grown on the backside (from Element Six Technologies). The AlGaN/GaN HEMT structure initially includes an AlGaN transition layer during the epi-wafer preparation [8.14]. This low thermal conductivity transition layer is removed in our epi-wafer in order to avoid high thermal resistance between GaN epilayer and the diamond substrate. For comparison, a HEMT structure on a Si (111) substrate from a commercial vendor was also fabricated (Sample B). The schematic diagrams of the epilayer structure and cross-sections of the HEMT of the two samples are shown in Fig. 8.1. They both have similar $Al_{0.25}Ga_{0.75}N$ barriers and thin GaN cap layers to prevent the AlGaN from being oxidized. GaN-on-diamond has a thinner GaN buffer (~ 800 nm) and substrate (~95 μm) than GaN-on-Si wafer's buffer layer (~ 2μm) and substrate (~1mm).

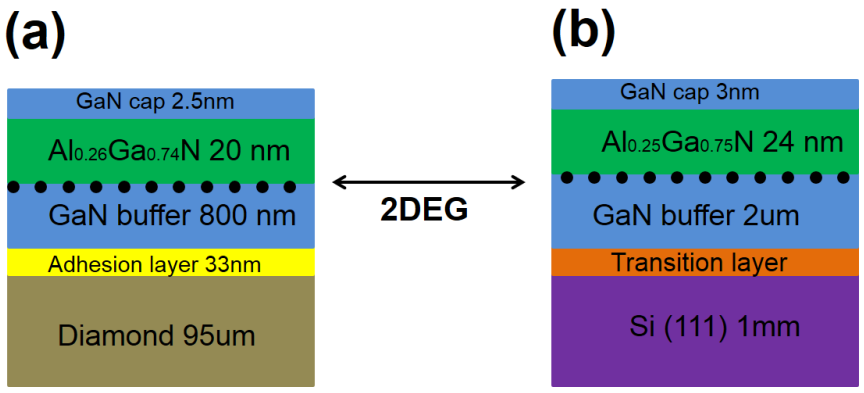



Figure 8.1: Schematic diagrams of the epitaxial layer structure and cross-sections of the HEMTs: (a) sample A and (b) sample B.

Devices from these materials were fabricated. To ease wafer handling, the GaN-on-diamond was mounted to a 1.5mm × 1.5mm carrier using a wax adhesive and it was dismounted before each fabrication processing step that exceeded 200 ºC. Although the GaN-on-diamond wafer used in processing was diced into 1.0mm × 1.0mm squares, the available wafer size is up to 4 inch.

For both material structures, mesa isolation was first achieved with $Ar/Cl_2/BCl_3$ plasma based Reactive-ion etching (RIE) system. Ohmic contact was then formed by Ti/Al/Ti/Au e-beam evaporation followed by 30s rapid thermal annealing at 850 ºC in $N_2$ ambient. After opening the gate windows with photolithography, and prior to the gate metal deposition, diluted HCl solution was used for surface treatment. Then Ni/Au was e-beam evaporated as Schottky gate metals. Finally, Ti/Au was deposited as metal pads for measurement. The final device structure and top-view photo of the device are shown in Fig. 8.2.

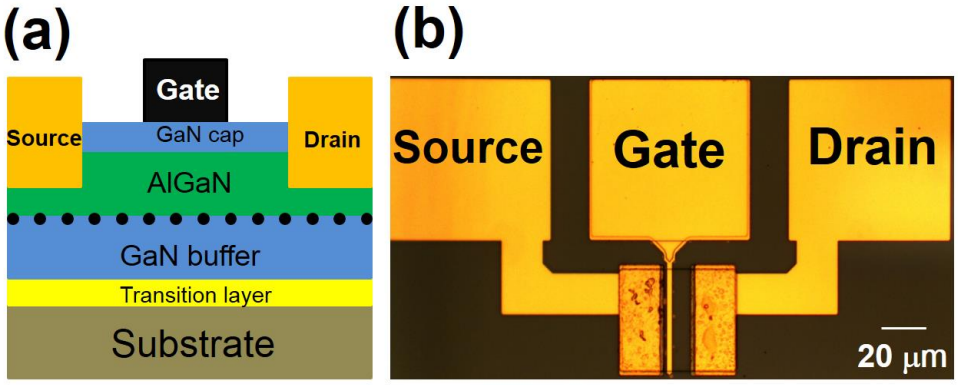



Figure 8.2: (a) Device structure of AlGaN/GaN HEMT (b) Optical image of fabricated HEMT device.

## 8.3 Device characterizations and analysis

Transmission Line Method (TLM) measurement was carried out using an Agilent B1500 system to obtain the contact resistance. The resistances for both devices are derived from the data in Fig. 8.3 and summarized in Table 8.1. The device made on GaN-on-diamond has higher contact resistance since the contacts used here were not optimized for this particular epi-source, as well as from the higher sheet resistance in the GaN-on-diamond HEMT channel.

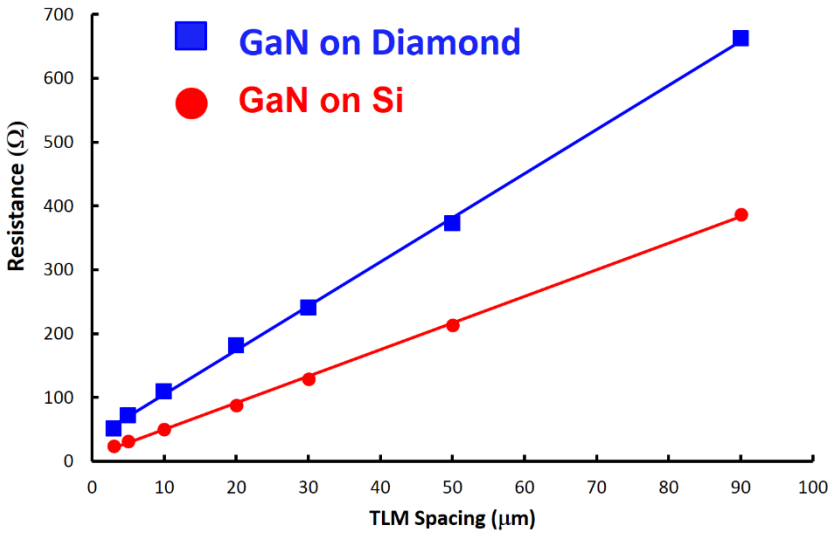



Figure 8.3: Measured resistance of TLM structures on both devices. Blue squares are for GaN-on-diamond HEMT, and red circles for GaN-on-Si.

Table 8.1: Contact resistance and sheet resistance of both devices obtained from TLM measurement.

**Table8.1 Contact resistance and sheet resistance of both devices obtained from TLM measurement**

| | Contact resistance (Ω-mm) | Sheet resistance (Ω/□) | Transfer length ($\mu$m) |
|---|---|---|---|
| GaN-on-diamond | 1.78 | 692 | 2.58 |
| GaN-on-Si | 0.38 | 421 | 0.9 |

DC characterizations of HEMTs from both materials, with gate length of 2um, source-drain spacing of 6um, and gate width of 50um, were performed using the Agilent B1500. During the I-V measurement sweeps, both GaN-on-Si and GaN-on-diamond wafers were placed on a large aluminum stage (diameter ~20cm) which remained at nearly constant room temperature ~25 ºC. We applied drain and source voltage to devices through DC probes (Ni tips) and heat was generated in the channel. Results are shown in Fig. 8.4.

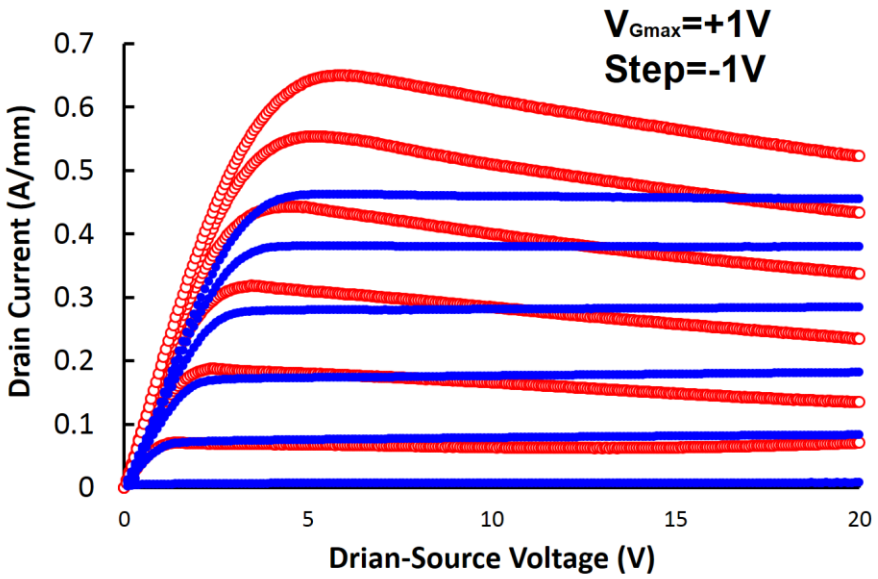



Figure 8.4: DC I-V characteristics for GaN-on-diamond (blue) and GaN-on-Si (red) HEMTs with 2um gate length and 50um width.

The threshold voltage for GaN-on-diamond HEMT is -3V and for GaN-on-Si HEMT is -3.5V as can be seen in Fig. 8.5. The gate voltage was swept from -4V to 1V with a step of 1V. In the I-V characteristics shown in Fig. 8.4, thermally-induced negative output resistance in the saturation region at DC is observed for GaN-on-Si and is negligible for the GaN-on-diamond device at comparable power densities. Similar self-heating induced current droop for HEMTs has also been observed by other groups on various substrates ([8.6], [8.10], [8.11]),.This limits AlGaN/GaN HEMTs from achieving their maximum current and lowers the power conversion efficiencies. In comparison, the current in the saturation region of GaN-on-diamond HEMT showed very little droop at the same level of power dissipation as shown in Fig. 8.4. However, the maximum current for GaN-on-diamond HEMT is lower than GaN-on-Si owing to the higher sheet resistance and contact resistance in our GaN-on-diamond epitaxial structures. The field effect mobility extracted

from the I-V characteristics in Fig. 8.4 is 833 $cm^2/(V \cdot s)$ and 1212 $cm^2/(V \cdot s)$ for GaN-on-diamond HEMT and GaN-on-Si HEMT respectively. The lower field effect mobility for GaN-on-diamond HEMT may due to greater roughness scattering in the GaN-on-diamond epitaxial layers.

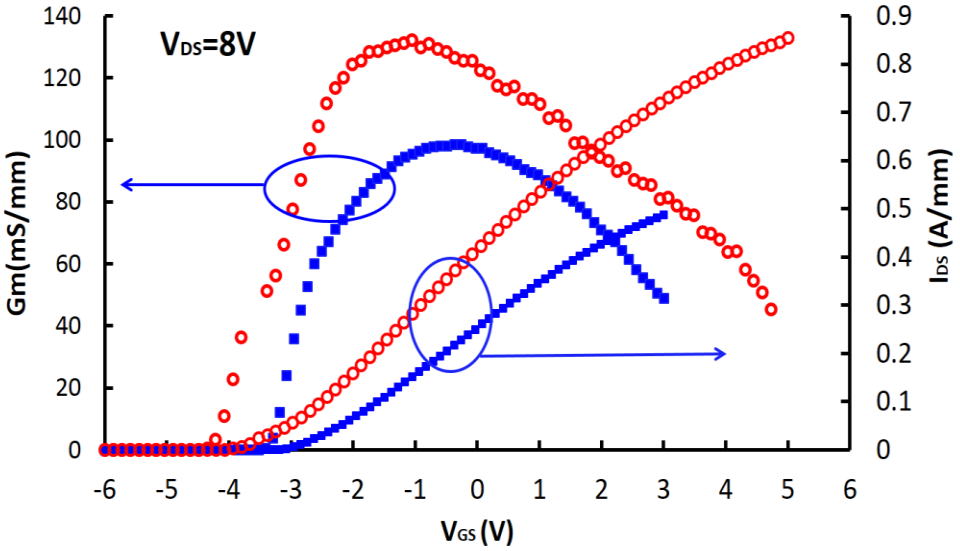



Figure 8.5: Transfer characteristics for GaN-on-diamond (blue) and GaN-on-Si (red) HEMTs with 2um gate length and 50um width.

In order to extract quantitatively the thermal resistance of diamond and Si substrates, I-V characteristics with varied substrate temperature were measured by Agilent B1500 system with an external electric heating source to change the substrate temperature. Both devices measured had the same dimensions (2 μm gate length, 6 μm source-drain spacing and 50 μm gate width) and a gate voltage of 1V was applied.

We define the current droop $\Delta I_{sat}$ as the change in the quantity, $I_{sat} - I_0$, where $I_{sat}$ designates a current value in a particular HEMT working point and $I_0$ denotes a reference

current when the substrate is at 25 ºC Similarly, ΔT is defined as temperature change of $T$-$T_0$, where $T$ is the present substrate temperature and $T_0$ denotes a reference temperature of 25 ºC. We plot the normalized drain current droop as a function of temperature change at a fixed source-drain voltage Vd, (we use Vd=10V in our calculation) and gate bias at 1V as shown in Fig. 8.6. As observed from Fig. 8.6. the current droop ΔIsat has a linear dependence [12] on temperature change ΔT as $\Delta I_{sat} \approx \alpha I_0 \Delta T$ , where α is the thermal coefficient which can be determined from the slope of the curves in Fig. 8.6. $\alpha = 2.27 \times 10^{-3}$ ºC$^{-1}$ for GaN-on-diamond HEMT and $\alpha = 2.66 \times 10^{-3}$ ºC$^{-1}$ for GaN-on-Si HEMT.

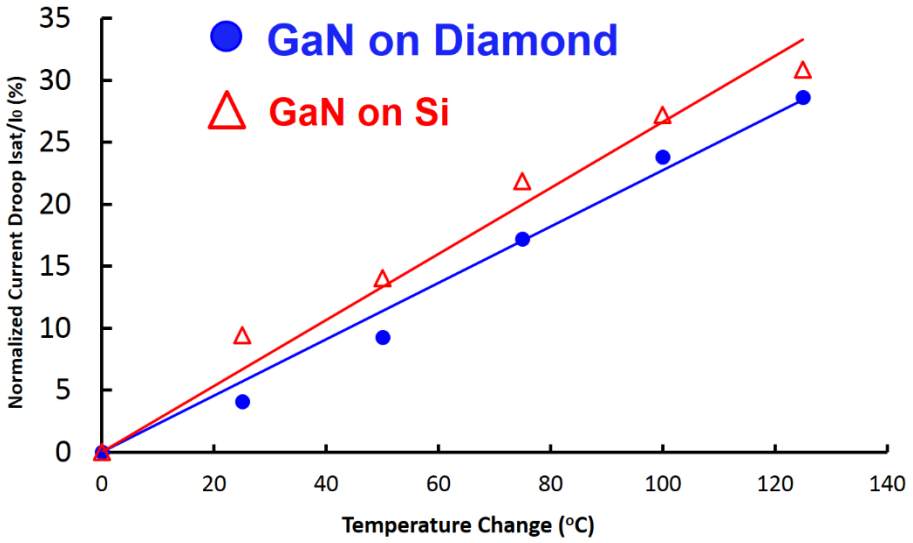



Figure 8.6: AlGaN/GaN HEMT current droop in the satuartion region as a function of temperature change at gate bias of 1V and source-drain voltage of 10V.

The channel temperaure rise $\Delta T_{rise}$ at different drain applied voltage can be estimated by $\Delta T_{rise} = (\Delta I_{droop} / I_0)/\alpha$. Here $\Delta I_{droop}$ is the amount of current droop referred to the value of current at the knee point, and α is the thermal coefficient extracted from Fig. 8.6. The

added power dissipation (ΔP) at the corresponding HEMT working point can be calculated as $\Delta P = P\text{-}P_0$, where $P$ is the particular power dissipation ($P = V_d \cdot I_{sat}$) and $P_0$ is the power dissipation at the knee point ($P_0 = V_{knee} \cdot I_{sat}$). The channel temperature rise versus added power dissipation is ploted in Fig. 8.7. The thermal resistance Rth can be extracted from the slope of the curves in Fig. 8.7 as $R_{th} = \Delta T_{rise} / \Delta P$ [8.12]. Thermal resistances extracted from Fig. 8.7 are 41.1 ºC/W and 156.5ºC/W for diamond and Si substrates, respectively.

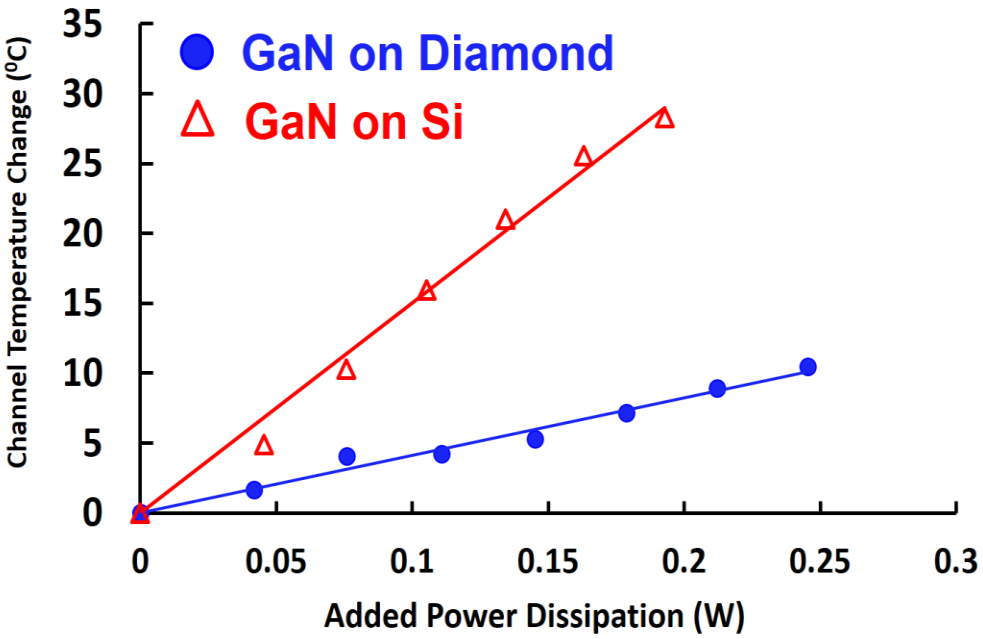



Figure 8.7: Channel temperature change of operating GaN-on-diamond and GaN-on-Si HEMTs.

It should be noted that trapping effects can also cause current collapse leading to current changes that can add to those from self-heating effects [8.13] [8.17] [8.18]. However, current collapse mainly affects the on-resistance at voltages lower than the knee voltage and the saturation region current droop is dominated by temperature rise as

indicated by [8.15]. In our analysis, we assume that the currents at the knee voltage point and at $V_{ds}$ = 10V have very similar trapping effects and the dominant control variable is channel temperature. With the assumption that temperature is the only control variable, the effective thermal resistance can thus be estimated from the self-heating induced current droop.

## 8.4 Simulations

3D thermal simulations were carried out using COMSOL and the simulation results are compared with the experimental results. The geometry of the model is shown in Fig. 8.8. The heat source was considered to be uniform over a region of the size of the HEMT drain extension region which is around 1.5 μm in length and 50 μm in width. The layer structure included an adhesion layer and/or transition layer between the GaN buffer and the substrate material. Boundary conditions applied to the sidewalls were $-\mathbf{n} \cdot (-k\nabla T)=0$, where k is thermal conductivity, T is temperature and **n** is normal vector to the peripheral facets. In other words, the heat flux normal to peripheral facets is zero. A heat sink was defined at the bottom of the structure at constant temperature of 293.15 K.

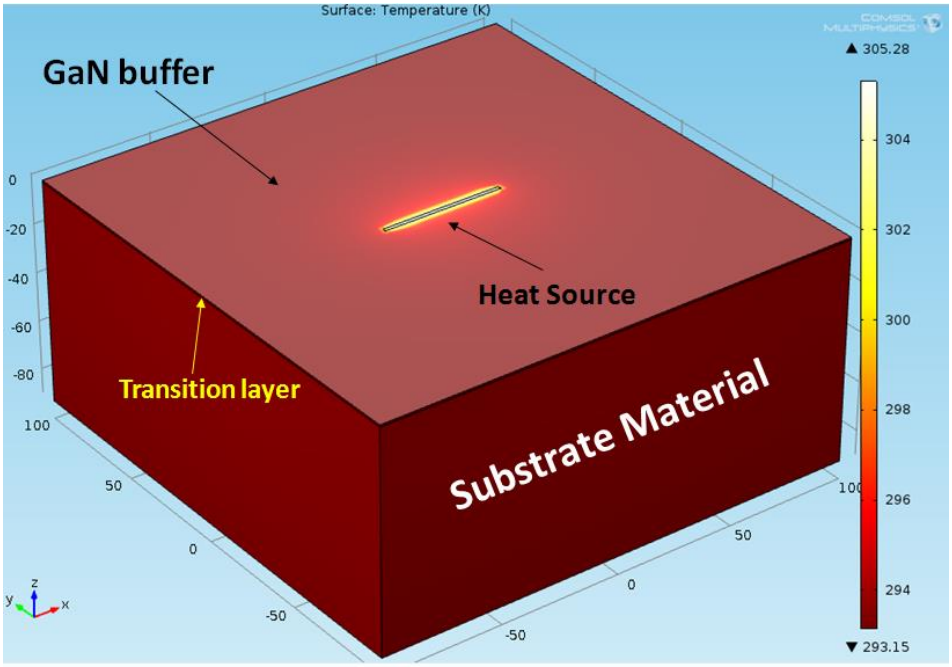



Figure 8.8: 3D thermal model of operating GaN HEMT

The simulated temperature of the GaN was found to vary across the device cross-section. A 1D temperature profile in the FET channel is shown in Fig. 8.9. In order to provide the most accurate comparison with the results of the thermal resistance measurements via drain current variation, we defined an effective temperature $T_{eff}$ that corresponds to the average temperature of the channel in the region which governs the drain current of the device, which we estimate to be the region of the channel between the source side of the gate and the center of the gate.

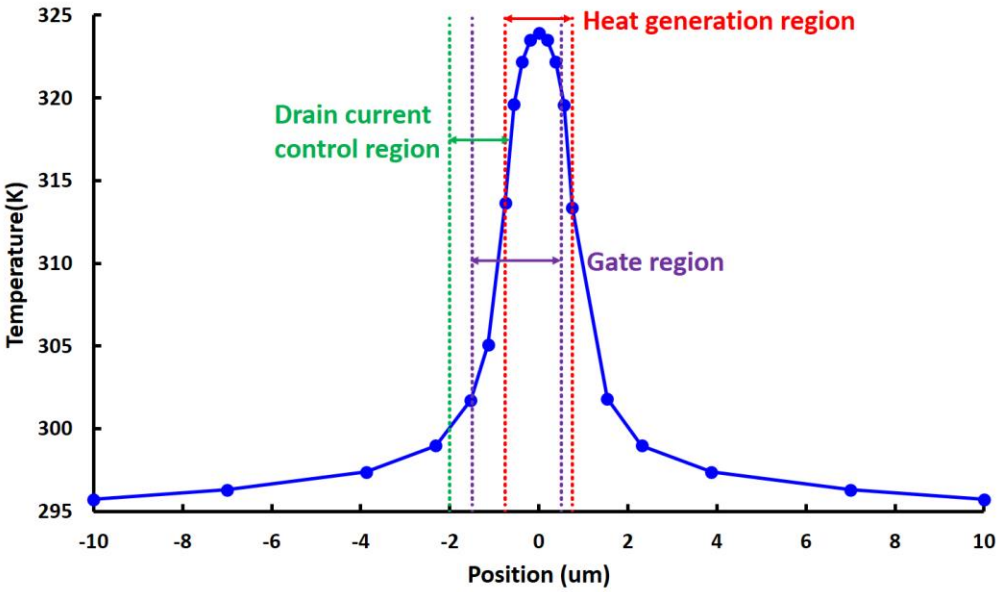



Figure 8.9: 1D temperature profile across the simulated heat generation region

The effective temperature $T_{eff}$ for the drain current control region of the HEMT was calculated by averaging the computed temperature over this area according to $T_{eff} = \frac{1}{N}\sum_{Area}^{N} T_n$, and the temperature rise is $T_{eff}$ - $T_{sink}$, where $T_{sink}$ is the temperature of the heat sink. Fig. 8.9 shows the estimated drain current control region, the gate region and the heat generation region (located on the drain side of the gate). The calculation indicates that the rise in $T_{eff}$ defined in this manner is about 62% of the rise of temperature averaged over the heat generation region (the hottest part of the device).

The theoretical thermal resistance $R_{th}$ is calculated as effective temperature rise divided by input power ($P_{in}$) as $R_{th} = (T_{eff}-T_{sink})/P_{in}$. The thermal resistances for the HEMTs on the diamond, Si and SiC substrates calculated by 3D thermal simulation are 37.1 ºC/W,

163.4 ºC/W and 83.7 ºC/W respectively. These results are consistent with the experimental results for diamond and Si substrates shown in Table 8.2.

Table 8.2: Comparison between experiments and simulations of HEMT thermal resistances on diamond and Si substrates.

**Table 8.2 Comparison between experiments and simulations of HEMT thermal resistances on diamond and Si substrates**

| Substrate | Thermal resistance (°C/W) from experiments | Thermal resistance (°C/W) from simulations |
|---|---|---|
| Diamond | 41.1 | 37.1 |
| Si | 156.5 | 163.4 |

Since the thermal conductivity value for the adhesion layer is unknown to us, and it is difficult to precisely measure its value [8.14], we varied its thermal conductivity in the simulation to understand the effect of the adhesion layer on the overall effective thermal resistance of the FET. Results are shown in Fig. 8.10. The figure shows that when the thermal conductivity of the adhesion layer falls below 10 W/m·K, the effective thermal resistance of the FET starts to show a big increase. However, the thermal resistance of the FET on diamond substrate (41.1 ºC/W) extracted from our experiment suggests that the thermal conductivity of our adhesion layer is within the range of 50 to 150 W/m·K in accordance with Fig. 8.10.

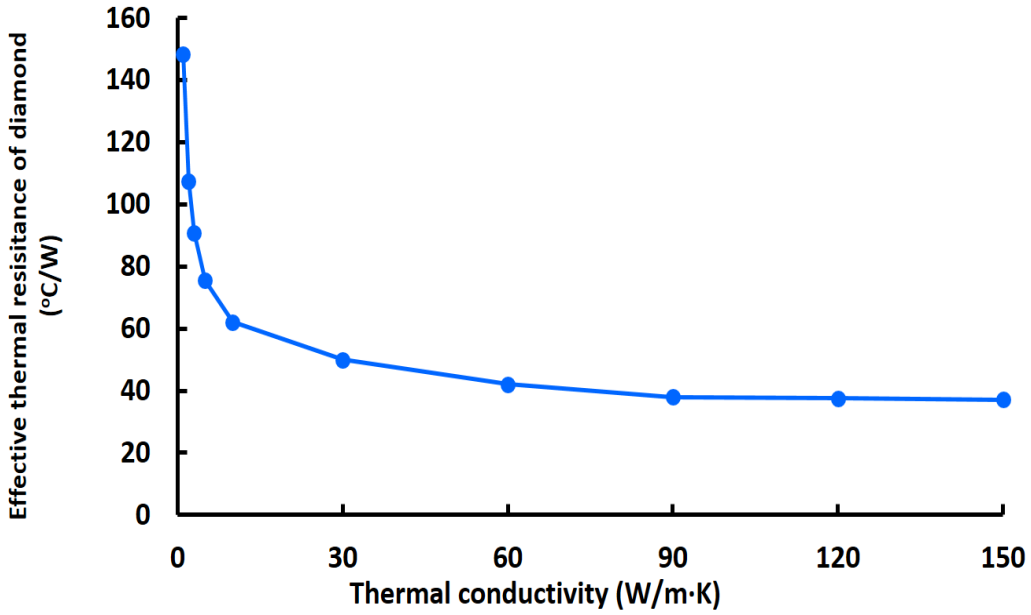



Figure 8.10: Effective thermal resistances of diamond substrate versus thermal conductivity of the adhesion layer.

In addition, the Si substrate (~1mm thick) is much thicker than the diamond substrate (~95 μm thick) as shown in the schematic diagram of epitaxial structure in Fig. 8.1. From the simulation study, the calculated thermal resistance for Si substrate with a thinner substrate ~95 μm is 155.23 W/m-K (only 5% improvement compared with the ~1mm Si substrate case). This suggests that for representative HEMT structures, diamond substrates can provide more than three times lower thermal resistance than Si substrates and two times lower than SiC substrate with the same thickness.

## 8.5 Conclusion

Performance of AlGaN/GaN HEMTs on Si and on diamond substrates have been compared. Self-heating induced negative differential output resistance has been observed

for GaN-on-Si HEMTs in the saturation region, and is negligible for GaN-on-diamond HEMTs. This phenomenon can be attributed to the higher thermal resistance of HEMTs on Si substrates than on diamond substrates. The results have been further confirmed by 3D thermal simulation.

The text of Chapter 8, in part, is a reprint of the material as it appears in "Thermal Resistance Extraction of AlGaN/GaN Depletion-Mode HEMTs on Diamond" by Jianzhi Wu, Jie Min, Wei Lu and Paul. K. L. Yu, *Journal of Electronic Materials*, vol. 44, no. 5, pp. 1275–1280, May 2015. The dissertation author was the primary investigator and author of this paper.

## 8.6 References

[8.1] U. K. Mishra, P. Parikh, and Y. Wu, *in Proc. IEEE*, 90, 1022 (2002).

[8.2] Y. F. Wu, M. Moore, A. Saxler, T. Wisleder, and P. Parikh, *in Proc. IEEE DRC Conf. Dig.*, pp. 151–152 (2006).

[8.3] J. R. Shealy, V. Kaper, V. Tilak, T. Prunty, J. A. Smart, B. Green and L. F. Eastman, *J. Phys. Condens. Matter*, 14, 13 (2002).

[8.4] D. Francis, J. Wasserbauer, F. Faili, D. Babic, F. Ejeckam, W. Hong, P. Specht, and E. R. Weber, *in Proc. CS MANTECH*, pp. 133–136 (2007).

[8.5] G. H. Jessen, J. K. Gillespie, G. D. Via, A. Crespo, D. Langley, J. Wasserbauer, F. Faili, D. Francis, D. Babic, F. Ejeckam, S. Guo, and I. Eliashevich, *in Proc. IEEE Compound Semicond. Integr. Circuit Symp. Tech. Dig.*, pp. 271–274 (2006).

[8.6] J. G. Felbinger, L. F. Eastman, J. Wasserbauer, F. Faili, D. I. Babić, D. Francis, F. Ejeckam, *IEEE Electron Dev. Lett.*, pp. 22-24 (2009).

[8.7] TriQuint Semiconductor, *Microwave Journal*, April 30, 2013.

[8.8] DG Cahill, HE Fischer, T Klitsner, ET Swartz, RO Pohl, *J. Vac. Scie. & Technol. A* 7 (3), 1259-1266, 215, (1989).

[8.9] J. Cho, Z. Li, E. Bozorg-Grayeli, T. Kodama, D. Francis, F. Ejeckam, F. Faili, M. Asheghi, and K. E. Goodson, *Proc. of ITHERM*, pp. 435–439 (2012).

[8.10] F. Berthet, Y. Guhel， H. Gualous, B. Boudart, J.L. Trolet, M. Piccione, C. Gaquiére, *Microelectronics Reliability*, 51(9-11):1796-1800 (2011).
[8.11] M. Kuball, J. M. Hayes, M. J. Uren, T. Martin, J. C. H. Birbeck, R. S. Balmer, and B. T. Hughes, *IEEE Electron Dev. Lett.*, 23, 1 (2002).

[8.12] E.T. Yu and M. O. Manasreh, *III-V Nitride Semiconductors: Applications and Devices*, (CRC Press, September 6, 2002)

[8.13] S. Martin-Horcajo, A. Wang, M-F. Romero, M. J. Tadjer, and F. Calle. *Electron Devices, IEEE Transactions on* 60, no. 12 (2013): 4105-4111.

[8.14] J. W. Pomeroy, M. Bernardoni, D. C. Dumka, D. M. Fanning, and M. Kuball. *Appl. Phys. Lett.* 104, no. 8 (2014): 083513.

[8.15] I. Saidi, M. Gassoumi, H. Maaref, H. Mejri, and C. Gaquière. *J. of Appl. Phys.* 106, no. 5 (2009): 054511.

[8.16] C. Mion, J. F. Muth, E. A. Preble, and D. Hanser. *Superlattices and Microstructures* 40, no. 4 (2006): 338-342.

[8.17] J. Joh, J. A. del Alamo, U. Chowdhury, T.-M. Chou, H.-W. Tserng, and J. L. Jimenez, *IEEE Trans. Electron Devices*, vol. 56, no. 12, pp. 2895–2901, Dec. 2009.

[8.18] J. Joh and J. A. del Alamo, *in Proc. IEDM*, Dec. 2006, pp. 415–418.

[8.19] M. Jamil, E. Irissou, J. R. Grandusky, V. Jindal, and F. Shahedipour-Sandvik. *physica status solidi (c)* 3, no. 6 (2006): 1787-1791

# Chapter 9 Normally-OFF AlGaN/GaN MOS-HEMT with a Two−Step Gate Recess

In addition to the thermal management issue discussed in Chapter 8, conventional AlGaN/GaN HEMTs are operated normally-on [9.1-9.3] due to the presence of high density two-dimensional electron gas in the channel induced by strong spontaneous and piezoelectric polarization charges [9.4]. However, normally-off devices are highly desirable as they are more suitable for reliable power-switching systems and enable the simpler power amplifier circuits using single-polarity voltage supply [9.5]. This chapter introduces the technology to enable normally-OFF operation for GaN HEMTs.

## 9.1 Introduction

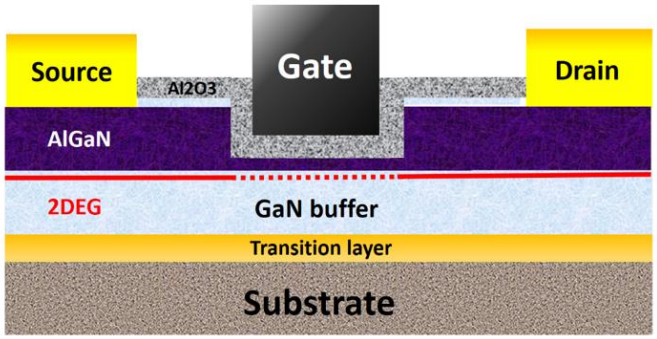
Various fabrication techniques have been investigated for normal-off AlGaN/GaN HEMTs such as fluorine-based plasma treatment [9.6], gate injection transistor [9.7] and gate recess etching [9.8]. In view of long term reliability, the gate recess technique based on ICP etching has been widely adopted [9.8-9.9]. Nonetheless, such approaches always lead to undesirably rough and/or non-uniform surfaces, which can cause the surface roughness scattering [9.10] and thus greatly limit the maximum output current of the device.

We studied a two-step gate recess etching technique which utilizes ICP plasma etching followed by wet chemical etching to realize the MOS-HEMT structure for normally-off operations. *I-V* characteristics of the device fabricated exhibits a high output current.

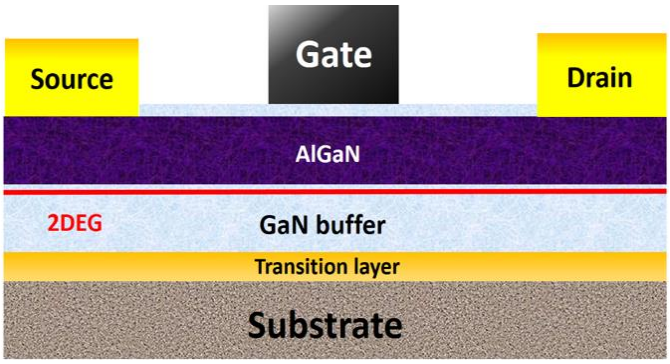
## 9.2 Experimental procedures

The AlGaN/GaN HEMT sample used in this study is grown on Si (111) substrate by a commercial vendor. The epitaxial structure includes a ~2μm unintentionally doped GaN buffer layer, a ~24 nm unintentionally doped $Al_{0.25}Ga_{0.75}N$ barrier layer and ~3 nm GaN cap layer.

Device fabrication starts with $Ar/Cl_2/BCl_3$ plasma based Reactive-ion etching (RIE) system for mesa isolation. Ti/Al/Ti/Au ohmic contact is subsequently deposited by e-beam evaporation followed by 30s rapid thermal annealing in nitrogen ambient. The gate recessed region is defined by optical photolithography with photoresist as etching mask. ICP $BCl_3/Cl_2$ plasma is used to recess the gate region in the first step. The photoresist etch mask is then removed by organic clean reagents. In the second gate recess step, the wafer is dipped into diluted HCl for 2min followed by 2min dip in diluted $NH_4OH$ solution. The gate recess depth profile is examined using Wyko NT1100 Optical Profiling System. Then a 20-nm-thick ALD $Al_2O_3$ layer is deposited as the gate dielectric followed by annealing at 400 $^oC$ for 10 min in forming gas ambient. Ni/Au is then deposited as gate metal by e-beam evaporator and Ti/Au is finally deposited as metal pads for measurements. As control sample, a device with the same mesa isolation, ohmic, gate and pads metallization, but without gate recess process and $Al_2O_3$ deposition, is fabricated in the same processing run for comparison. The two devices structures are shown in Fig. 9.1 (a) and (b) respectively.




(a)




(b)

Figure 9.1: (a) Device structure of gate recessed MOS-HEMT, (b) device structure of (control sample) un-recessed gate AlGaN/GaN HEMT

## 9.3 Results and discussions

The etched depth in the recess region is ~25.7 nm which is close to the thickness of AlGaN barrier (~24 nm) plus the GaN cap (~3 nm) layer, as measured from the optical profiling system (not shown here) and the etched surface is smoothened by the HCl and $NH_4OH$ surface treatment. It is believed that the HCl and $NH_4OH$ wet chemicals can effectively remove the damaged AlGaN residues at the recessed surface caused by the $BCl_3/Cl_2$ ICP etching. Moreover, HCl and $NH_4OH$ surface treatment has been reported to

be critical in the subsequent $Al_2O_3$ ALD process to achieve a low-trap density interface between GaN and $Al_2O_3$ gate dielectric [9.11].

DC *I-V* characteristics of the fabricated MOS-HEMT device and an un-recessed gate AlGaN/GaN HEMT device are measured by Agilent B1500 and displayed in Fig. 9.2.

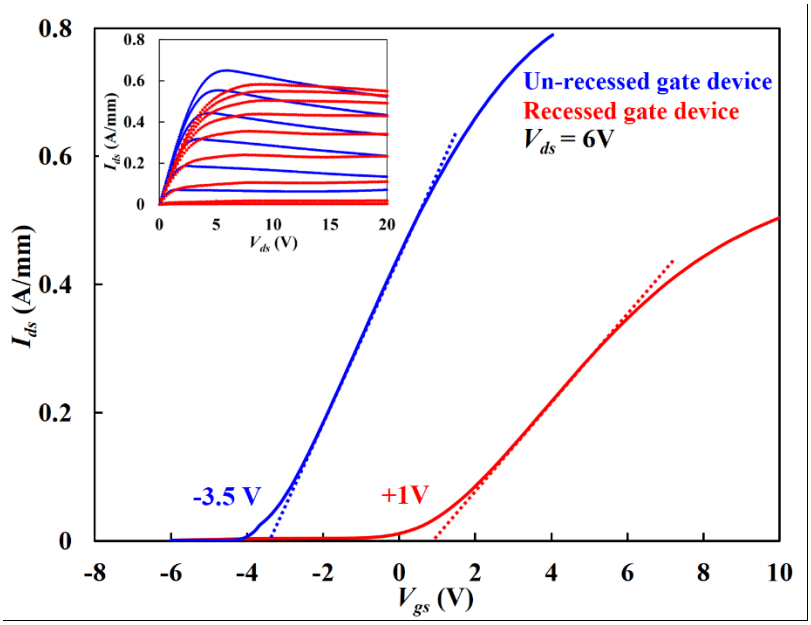



Figure 9.2: Transfer characteristics of the gate recessed MOS-HEMT (red) and un-recessed AlGaN/GaN HEMT (blue) devices. Both are with 2µm gate length, 6µm source-to-drain spacing and 50µm gate width. Inset: $I_{ds}$-$V_{ds}$ characteristics, where the gate swing is $V_{gs}$=-2 to +14V, step=2V and $V_{gs}$=-3 to +2V, step=1V for red and blue curves, respectively.

Fig. 9.2 shows that the gate-recessed MOS-HEMT device achieves a +1V threshold voltage ($V_{th}$) compared with a non-recessed HEMT device with a $V_{th}$ of -3.5V. $V_{th}$ is defined as the gate voltage intercept point of linear extrapolation of the $I_{ds}$-$V_g$ curve at its maximum first derivative point [9.12].

The inset figure of Fig. 9.2 further exhibits the $I_{max}$ for the gate recessed MOS-HEMT device as 0.583 A/mm, which is the highest among the most recently reported $I_{max}$ values for normally-off MIS−HEMT devices as evident from Fig. 9.3. The field−effect mobility

extracted from *I-V* curves is 150 $cm^2$/Vs. $I_{max}$ is expected to be higher with an improved epitaxial layer design.

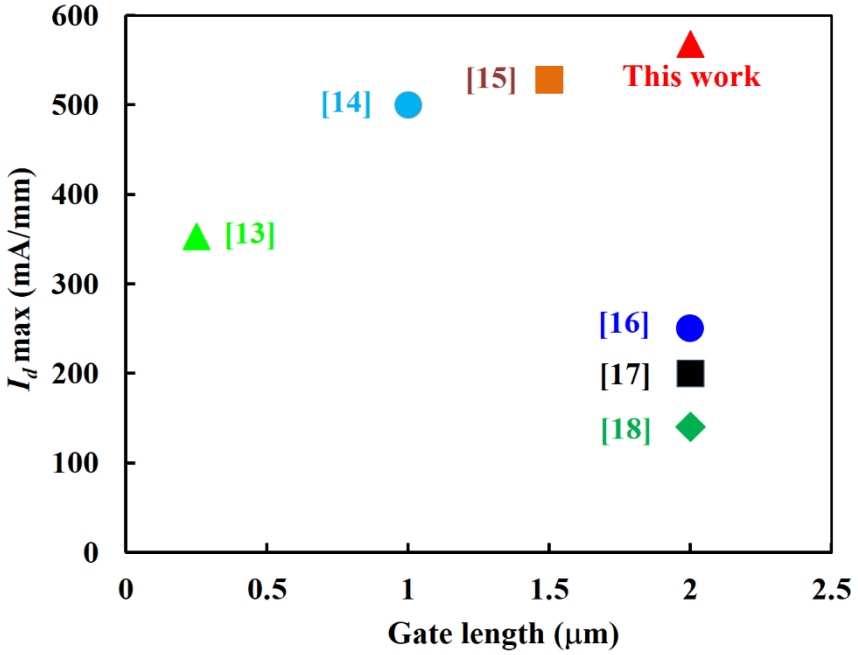



Figure 9.3: Comparison of the device maximum output current density with recently reported data with comparable gate length.

The text of Chapter 9, in part, is a reprint of the material as it appears in "Normally-OFF AlGaN/GaN MOS-HEMT with a Two−Step Gate Recess" by Jianzhi Wu, Wei Lu and Paul. K. L. Yu, *IEEE International Conference on Electron Devices and Solid-State Circuits* (*EDSSC*) (pp. 594-596), Singapore, 2015. The dissertation author was the primary investigator and author of this paper.

## 9.4 References

[9.1] N. Ikeda, Y. Niiyama, H. Kambayashi, Y. Sato, T. Nomura, S. Kato and S. Yoshida. "GaN power transistors on Si substrates for switching applications." *Proceedings of the IEEE,* 98, no. 7 (2010): 1151-1161.

[9.2] U. K. Mishra, L. Shen, T. E. Kazior, and Y. Wu. "GaN-based RF power devices and amplifiers." *Proceedings of the IEEE,* 96, no. 2 (2008): 287-305.

[9.3] T. Hussain, M. Micovic, T. Tsen, M. Delaney, D. Chow, A. Schmitz, P. Hashimoto, D. Wong, J. S. Moon, M. Hu, J. Duvall, D. Mclaughlin. "GaN HFET digital circuit technology for harsh environments." *IET Electronics Letters* 39, no. 24 (2003): 1708-1709.

[9.4] O. Ambacher, J. Smart, J. R. Shealy, N. G. Weimann, K. Chu, M. Murphy, W. J. Schaff, L. F. Eastman, R. Dimitrov, L. Wittmer, M. Stutzmann, W. Rieger and J. Hilsenbeck. "Two-dimensional electron gases induced by spontaneous and piezoelectric polarization charges in N-and Ga-face AlGaN/GaN heterostructures." *Journal of Applied Physics*, 85, no. 6 (1999): 3222-3233.

[9.5] T. Palacios, C-S. Suh, A. Chakraborty, S. Keller, S. P. DenBaars, and U. K. Mishra. "High-performance E-mode AlGaN/GaN HEMTs." *IEEE Electron Device Letters*, 27, no. 6 (2006): 428-430..

[9.6] Y. Cai, Y. Zhou, K. J. Chen, and K. M. Lau. "High-performance enhancement-mode AlGaN/GaN HEMTs using fluoride-based plasma treatment." *IEEE Electron Device Letters*, 26, no. 7 (2005): 435-437.

[9.7] Y. Uemoto, M. Hikita, H. Ueno, H. Matsuo, H. Ishida, M. Yanagihara, T. Ueda, T. Tanaka, and D. Ueda. "Gate injection transistor (GIT)—A normally-off AlGaN/GaN power transistor using conductivity modulation." *IEEE Transactions on Electron Devices*, 54, no. 12 (2007): 3393-3399.

[9.8] Q. Wang, K. Tamai, T. Miyashita, S. Motoyama, D. Wang, J-P Ao, and Y. Ohno. "Influence of Dry Recess Process on Enhancement-Mode GaN Metal–Oxide–Semiconductor Field-Effect Transistors." Japanese *Journal of Applied Physics*, 52, no. 1S (2013): 01AG02.

[9.9] K-W. Kim, S-D. Jung, D-S. Kim, H-S. Kang, K-S. Im, J-J. Oh, J-B. Ha, J-K. Shin, and J.-H. Lee. "Effects of TMAH Treatment on Device Performance of Normally Off MOSFET." *IEEE Electron Device Letters*, 32, no. 10 (2011): 1376-1378

[9.10] B. Liu, Y. W. Lu, G. R. Jin, Y. Zhao, X. L. Wang, Q. S. Zhu, and Z. G. Wang. "Surface roughness scattering in two dimensional electron gas channel." *Applied Physics Letters*, 97, no. 26 (2010): 262111-262111.

[9.11] A. J. Kerr, E. Chagarov, S. Gu, T. Kaufman-Osborn, S. Madisetti, J. Wu, P. M. Asbeck, S. Oktyabrsky, and A. C. Kummel. "Preparation of gallium nitride surfaces for atomic layer deposition of aluminum oxide." *Journal of Chemical Physics,* 141, p. 104702 (2014)

[9.12] S. Sze, Physics of Semiconductor Devices, 2nd ed. New York, NY, USA:Wiley, 1981

[9.13] K-S Im, J-B Ha, K-W Kim, J-S Lee, D-S Kim, S-H Hahm, and J-H Lee. "Normally off GaN MOSFET based on AlGaN/GaN heterostructure with extremely high 2DEG density grown on silicon substrate." *IEEE Electron Device Letters,* 31, no. 3 (2010): 192-194.

[9.14] W. B. Lanford, T. Tanaka, Y. Otoki, and I. Adesida. "Recessed-gate enhancement-mode GaN HEMT with high threshold voltage." *Electronics Letters,* 41, no. 7 (2005): 449-450.

[9.15] Y. Wang, M. Wang, B. Xie, C. P. Wen, J. Wang, Y. Hao, W. Wu, K. J. Chen, and B. Shen. "High-Performance Normally-Off MOSFET Using a Wet Etching-Based Gate Recess Technique." *IEEE Electron Device Letters*, 34, no. 11 (2013): 1370-1372.

[9.16] Z. Xu, J. Wang, Y. Cai, J. Liu, C. Jin, Z. Yang, M. Wang, M. Yu, B. Xie, W. Wu, X. Ma, J. Zhang and Y. Hao. " Enhancement Mode (E-Mode) AlGaN/GaN MOSFET With $10^{-13}$ A/mm Leakage Current and $10^{12}$ ON/OFF Current Ratio." *IEEE Electron Device Letters*, 35, no. 12 (2014): 1200-1202.

[9.17] Z. Xu, J. Wang, Y. Liu, J. Cai, J. Liu, M. Wang, M. Yu, B. Xie, W. Wu, X. Ma and J. Zhang. "Fabrication of normally off AlGaN/GaN MOSFET using a self-terminating gate recess etching technique." *IEEE Electron Device Letters*, 34, no. 7 (2013): 855-857.

[9.18] T-H. Hung, P. S. Park, S. Krishnamoorthy, D. N. Nath, and S. Rajan. "Interface charge engineering for enhancement-mode GaN MISHEMTs." *IEEE Electron Device Letters*, 35, no. 3 (2014): 312-314.

# Chapter 10 Conclusion and future work

## 10.1 Summary

The goal of this dissertation is to develop a predictive and analytic model of tunnel FETs with a double-gate configuration. From the perspective of circuit simulations, it would be necessary to derive an analytic *I-V* model covering all regions of operations, i.e. subthreshold region, linear region and saturation region. Short channel effect, source doping effect, ambipolar effect, charge screening effect are successfully incorporated into the model. A set of TFET scaling rules has been derived that highlights the effective mass as the key to scaling TFETs to sub-10-nm channel lengths. The model is validated by comparison with numerical TCAD simulations for a broad range of channel lengths and bandgaps.

The main contribution of this work, beyond the prior art, is revealing that the tunneling barrier in a TFET is of an exponential nature. It is derived by solving the Poisson's equation in a 2D boundary value problem in chapter 2. Based on that finding, an analytic *I-V* model is formulated for the exponential barrier in closed form. It is concluded from the model that source degeneracy helps the linear region I-V characteristics, but degrades the saturation current. The performance of 1-D, 2-D, and 3-D heterojunction TFETs has been assessed as a function of supply voltage. In the best scenario of zero staggered bandgap, 1-D TFETs can achieve an ultimate lower limit of supply voltage $\approx 6kT/q$, with no loss of performance. Also concluded from chapter 3 is that the correlation between the optimum source doping and the effective density of states is found to be a key factor in TFET design. By investigating the short-channel effects in TFETs in chapter 4, it

is shown that below a minimum channel length about twice the thickness of the device, the off current increases sharply and the subthreshold swing fails to be below 60 mV/decade.

To more accurately model the current TFETs, imaginary $E(k)$ relationship in the forbidden bandgap is thoroughly investigated in chapter 5. Four published two-band $E(k)$ relations was examined, including a piecewise case, showing that both Franz's model and the piecewise linear model behave reasonably in all asymptotic limits. Based on this study, Franz's two-band $E(k)$ relation is employed to derive a continuous and semi-analytic TFET model in chapter 6. Based on which, a set of TFET scaling rules has been derived that highlights the effective mass as the key to scaling TFETs to sub-10 nm channel lengths. Furthermore, hardware data of TFETs are reviewed and compared to the analytic model results in chapter 7. Most published hardware data corroborate with analytic results showing similar superlinear on-set behavior. Also the trend revealed by model that small staggered bandgap design giving high output current is consistent with hardware data.

For high power and high frequency operations, AlGaN/GaN HEMTs on Si and on diamond substrates have been fabricated and compared in chapter 8. Self-heating induced negative differential output resistance has been observed for GaN-on-Si HEMTs in the saturation region, and is negligible for GaN-on-diamond HEMTs. To further change the AlGaN/GaN HEMT device to enhancement mode operations, chapter 9 presents a two-step gate recess technique including ICP $BCl_3$ and $Cl_2$ plasma etching followed by HCl and $NH_4OH$ surface treatment to achieve positive $V_{th}$.

## 10.2 Future work

From the perspective of tunnel FETs modeling work, there are a good number of physics models worth to be explored and incorporated into the present double-gate/nanowire TFET model. For example, it is interesting to explore the temperature dependent effects on I-V characteristics of the tunnel FETs. Trap assisted tunneling effects may need to be embedded to more accurately predict the off-state current. Both of the aforementioned physical mechanisms are supposed to have strong temperature dependence. Besides, noise model, source/drain parasitic resistance model and overlap capacitance model are also worth future investigating. From this dissertation author's point of view, the most challenging issue for tunnel FETs is the band-tailing issues. As indicated from chapter 3, the optimized source doping concentration for TFETs is in the range of a few ~$10^{19}$ $cm^{-3}$. Hence, such a heavy doping implant will very likely cause defects, which results in a considerable number of density-of-states extending to the bandgap (i.e. band-tail states). Consequently, those band-tail states would, in turn, significantly degrade the subthreshold swing and off-state leakage current. As a result, this dissertation author proposes a few approaches to possibly overcome the band-tail issues. 1) Use low-dimensional semiconductor materials (e.g. 1-D material) for source/channel material. They are promising not only for their free of dangling bonds hence no band-tail states, but also for their near atomic thickness which gives better electrostatic control. As shown in chapter 3, TFETs made of 1-D material give the best performance at low $V_{dd}$ compared to their 2-D and 3-D counterparts. 2) Employ III-nitride (e.g. AlN, GaN, InGaN, InN) materials for the TFET source material. They are attractive for their spontaneous and/or piezoelectric

polarization charges with no need for chemical doping. So that a sharp electrical field can be formed at the source-channel junction to boost the on-current, without worrying about band-tail states induced by heavy doping implantations.

From the GaN HEMT devices point of view, although a controllable fabrication process has been developed for normally-OFF GaN MOS-HEMTs, the reliability of the devices may need to be further tested and improved. For example, 1) the pulsed I-V measurement can be done in order to extract the dynamic on-resistance for the enhancement-mode HEMT devices. 2) Better surface treatment process is needed to further reduce the density of interface traps for AlGaN/GaN HEMTs to enable fast switching performance. 3) Multiple field-plate structure would be needed to further improve the breakdown voltage of the HEMT devices. 4) The temperature dependence and stability of the GaN HEMT devices need to be further characterized. Passivation technologies would help enhancing the stability of HEMT devices under high temperature.