



**KATHOLIEKE UNIVERSITEIT LEUVEN**  
Faculty of Engineering  
Department of Electrical Engineering  
Kasteelpark Arenberg 10, B-3001 Leuven, Belgium

**QUANTUM TRANSPORT  
IN TUNNEL FIELD-EFFECT TRANSISTORS  
FOR FUTURE NANO-CMOS APPLICATIONS**

Proefschrift ingediend ter  
verkrijging van de graad van  
Doctor in de ingenieurswetenschappen

door

**William VANDENBERGHE**

August 2012





KATHOLIEKE UNIVERSITEIT LEUVEN  
Faculty of Engineering  
Department of Electrical Engineering  
Kasteelpark Arenberg 10, B-3001 Leuven, Belgium

# QUANTUM TRANSPORT IN TUNNEL FIELD-EFFECT TRANSISTORS FOR FUTURE NANO-CMOS APPLICATIONS

Jury:

Prof. dr. ir. Y. Willems, voorzitter  
Prof. dr. ir. G. Groeseneken, promotor  
Prof. dr. W. Magnus, promotor  
(Universiteit Antwerpen)  
Prof. dr. ir. B. Sorée  
(Universiteit Antwerpen)  
Prof. dr. ir. M. Heyns  
Prof. dr. ir. W. Dehaene  
Dr. ir. A. Verhulst (imec)  
Prof. dr. M. Fischetti  
(University of Texas at Dallas, USA)

Proefschrift ingediend ter  
verkrijging van de graad van  
Doctor in de ingenieurswetenschappen

door

**William VANDENBERGHE**

In collaboration with:



August 2012

© 2012 Katholieke Universiteit Leuven, Groep Wetenschap & Technologie,  
Arenberg Doctoraatschool, W. de Croylaan 6, 3001 Heverlee, België

Alle rechten voorbehouden. Niets uit deze uitgave mag vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

Legal Depot Number D/2012/7515/94

ISBN 978-94-6018-560-1

*To my wife*



# Preface

Five years ago, while finishing my master in electrical engineering, I found my scientific knowledge to be lacking as there were many things I did not fully understand yet. As a solution to my lack of understanding, I decided to embark on a journey to achieve the highest academic title: Doctor of Philosophy (PhD). To increase my understanding, the choice for a theoretical topic was obvious and after talking to Marc, Bart, Wim and Geoffrey, I decided that Bart's proposal to do quantum mechanical modeling of the tunnel field-effect transistor at imec was the perfect vehicle to increase my knowledge.

It soon became clear that prof. Wim Magnus would be my promotor and after some time, it was decided that prof. Guido Groeseneken would be my promotor as well and I would be a K.U.Leuven PhD student. I also had to apply for a grant with the IWT and try to convince them my topic was of great importance to the Flemish industry. My first bid for a grant in 2007 was not successful but during my second attempt I was awarded a grant covering 4 years of salary and expenses.

During my PhD, I traveled around the world to scientific conferences in France, the United States and Japan. I attended a summer school in Italy, I did an internship at U.T.Dallas with prof. Massimo (Max) Fischetti and I must thank prof. Wachutka and prof. Selmi for inviting me to give a lecture at their institute.

I must thank my promotors Guido and Wim. Guido for timely proofreading my manuscripts and being forgiving when I forgot to teach at university or when I was not ready to compromise. Wim for always being supportive of my work and painstakingly correcting and improving my manuscripts. I must thank Bart for being a good daily supervisor, first of all for not literally daily supervising my work and giving me enough freedom to explore on my own. Further to always aim to do science at the highest level and to be fair and honest with me.

The members of my jury each deserve their piece of credit in the work. Anne

Verhulst, for having great interest in my work, asking me questions, encouraging me to write papers, co-authoring some of my papers and granting me the privilege of being a co-author on many of her own papers. Prof. Wim Dehaene, for always strongly supporting the quantum cause despite operating in a different world himself. Prof. Marc Heyns, for always being enthusiastic and showing great respect for science. Finally, I must pay my respect to Max. For having me over for an internship at U.T.Dallas, for teaching me the ways of the pseudopotential and presenting me with a very interesting representation of phonons. And probably most importantly, for convincing me to stay in the scientific community after my PhD by making a respectful offer to do a post-doc at U.T. Dallas the coming two years.

On a personal level I am indebted to many. First of all I must thank my parents Antoon and Belinda, for passing on a fine selection of their genes, for their continued financial support but most of all for their unconditional love and support. I want to thank my sisters Marlies and Renate, my grandmothers Hermina and Jenny and my in-laws Eric, Marijke, Thibaut and Barbara for their support as well.

During my PhD I was also able to convince a fantastic woman, Fleur, to marry me. Being my wife she had to deal with my at times eccentric behaviour, kick me out of bed in the morning while giving me the freedom to attend scientific conferences and to do an internship in the US. I can not thank her enough for her unwavering support.

I must thank my friends Alexander, Joram, Bram (2x), Karen, Pieter, Joris, Roel, Lieboud and Rogier for inviting me to the many evenings or afternoons out, supporting me and asking the dreaded question “when will you finish your PhD?”. Geert for being a good friend at imec, accompanying me to conferences and for inviting me to the occasional cocktail party. I must thank my friends Liesbeth, Thomas, Annelies and Alexandre who set the good example and fled Belgium before I decided to do so. I must thank my imec colleagues Anne, Frank, Quentin, Maarten, Philippe, Cedric, Geoffrey for being good company, my master thesis students Maarten and Sergej for being good minions and Karim, Chris, Rudy and Ivan for sitting next to me for a couple of years and having interesting conversations about brothels, unions and low-k material.

In the true spirit of the academy, I also used my five years of PhD to broaden my knowledge of philosophy, economics and politics to finally settle the question for myself whether the socialist or the free society is superior. I must thank David, Kristof, Dieter, Wouter, Boudewijn Bouckaert and many more helping me in my quest.

William Vandenberghe

Leuven, August 2012

# Abstract

After decades of scientific and technological development to fabricate ever smaller, faster and more energy efficient MOSFETs, reducing MOSFET power consumption is becoming increasingly difficult. As a possible successor to the MOSFET, the tunnel field-effect transistor (TFET) has been proposed. The topic of this thesis is to study the working principle of the TFET and to go beyond the semiclassical models towards a fully quantum mechanical modeling of the TFET which has band-to-band tunneling (BTBT) as its working principle.

Semiclassical analytical models are developed to describe the two different tunneling components in a TFET: point tunneling and line tunneling.

The Kane model for BTBT in direct semiconductors is extended to non-uniform fields and the envelope function formalism is extended towards heterostructures. Graphene based semiconductors are shown to have similar BTBT behavior to III-V based semiconductors. Theoretically calculating the BTBT parameters for germanium and silicon shows that direct tunneling dominates in germanium. For all direct and indirect semiconductor TFETs, it is shown there is an optimal doping concentration and material's band gap.

We develop a general formalism for BTBT in indirect semiconductors capable of calculating BTBT current under a one-, two- and threedimensional potential profile. We provide a convenient expression to calculate the BTBT current based on spectral functions and prove that in the uniform field limit, the formalism reduces to the existing BTBT models. Calculating the BTBT current in the line tunneling configuration shows the impact of field-induced quantum confinement: a large shift in onset voltage and a deteriorated subthreshold slope compared to semiclassical models. The model for field-induced quantum confinement is extended to account for penetration into the gate oxide. A modified semiclassical model for BTBT is proposed and shown to be capable of taking quantum confinement into account. Two-dimensional modeling of BTBT in indirect semiconductors is carried out and it is shown that also in a two-dimensional configuration, the quantum mechanical effects are important. In particular, the subthreshold slope and on-current are degraded with respect

to the semiclassical simulation. We use two-dimensional quantum mechanical simulations to show that the use of counter-doped pockets can alleviate the effects of quantum confinement.

# Korte Inhoud

Na jarenlange wetenschappelijke en technologische ontwikkeling om immer kleinere, snellere en energie-efficiëntere MOSFETs te fabriceren, wordt het verder schalen van MOSFET vermogenverbruik alsmaar moeilijker. Als een mogelijke opvolger voor de MOSFET werd de tunnel-veld-effect transistor (TFET) voorgesteld. Het onderwerp van deze thesis is het werkingsprincipe van de TFET bestuderen en de semiklassieke modellen overstijgen met een volledige kwantummechanische modelering van de TFET die band-tot-band tunneling (BTBT) als werkingsprincipe heeft.

Semiklassieke modellen worden ontwikkeld om de twee verschillende tunneling componenten van de TFET te beschrijven: punttunneling en lijntunneling.

Het Kane model voor BTBT in directe halfgeleiders wordt uitgebreid naar niet uniforme velden en het envelopfunctieformalisme wordt uitgebreid naar heterostructuren. Er wordt aangetoond dat op grafeen gebaseerde halfgeleiders gelijkaardig BTBT gedrag vertonen als III-V halfgeleiders. Door het theoretisch berekenen van BTBT parameters voor silicium en germanium tonen we aan dat directe tunneling overheerst in germanium. Voor alle directe en indirecte halfgeleider TFETs wordt aangetoond dat er een optimale doperingsconcentratie en materiaalbandkloof bestaat.

We ontwikkelen een algemeen formalisme voor BTBT in indirecte halfgeleiders dat in staat is om BTBT stroom te berekenen in een één-, twee- en driedimensionaal potentiaalprofiel. We verschaffen een gerieflijke uitdrukking voor het berekenen van de BTBT stroom gebaseerd op spectraalfuncties en tonen aan dat in de limiet van het uniform veld, het formalisme zich herleidt tot de bestaande semiklassieke modellen. Door het berekenen van de stroom in de lijntunneling-configuratie tonen we de impact van veld-geïnduceerde kwantumopsluiting: een grote verschijving van de aanzetspanning en een slechtere sub-drempel helling vergeleken met de semiklassieke modellen. Het model voor veld-geïnduceerde kwantumopsluiting wordt uitgebreid om penetratie in het gate oxide in rekening te brengen. Een aangepast semiklassiek model voor BTBT wordt voorgesteld waarvan wordt getoond dat dit de effecten van kwantumopsluiting in rekening

kan brengen. We tonen een twee-dimensionale modelering van BTBT in indirecte halfgeleiders en tonen aan dat ook in een twee-dimensionale configuratie, de kwantummechanische effecten belangrijk zijn. In het bijzonder verslechteren de sub-drempel helling en de aan-stroom vergeleken met de semiklassieke simulaties. We gebruiken twee-dimensionale kwantummechanische simulaties om aan te tonen dat het aanbrengen van tegen-gedopeerde gebieden, de effecten van kwantumopsluiting kan verlichten.

# Glossary

## Mathematical symbols

$\nabla$	Del operator
$a^\pm, b^\pm, c^\pm$	Coefficients for incoming and reflected waves
$a_n(\mathbf{k})$	Coefficients of Bloch functions
$\alpha$	Scaling factor
$A_{v,c}(\mathbf{r}, \mathbf{r}')$	Single particle conduction and valence band spectral functions
<b>A</b>	Vector potential
$\hat{A}$	Second quantized operator
$\beta$	Inverse of thermal energy
$c\mathcal{R}$	Eigenvalue to lattice translation operator
$\chi(x)$	One-dimensional wave function or envelope function
$\chi_{v,cl}(\mathbf{r})$	Valence and conduction band wavefunctions
$C_g$	Gate capacitance
$\delta_{nn'}$	Kronecker delta
$\delta(\mathbf{k} - \mathbf{k}')$	Dirac delta distribution
$E$	Energy
$E_g$	Semiconductor band gap
$E_{fn,p}$	Electron and hole quasi Fermi levels
$E_{v,c}$	Valence and conduction band
$E^{0 \rightarrow V_{dd}}$	Energy required to switch from 0 to $V_{dd}$
$E_{n\mathbf{k}}$	Energy eigenvalues of the Bulk Hamiltonian
$E_n^0$	Band edge energies
$E_{v,cl}$	Energy of valence and conduction band wavefunctions
$\epsilon_{0,ox,s}$	Vacuum, oxide and semiconductor dielectric constant
$\mathcal{E}$	Electric field
$f(E)$	Fermi-Dirac distribution
$f_{L,R}(E)$	Fermi-Dirac distribution associated with left and right running waves
$f_{v,c}(E)$	Fermi-Dirac distribution associated with valence and conduction band
$f_n(\mathbf{r})$	Envelope functions

$f(\mathbf{r}, \mathbf{k})$	Wigner distribution function
$\tilde{f}(\mathbf{k})$	Fourier transformed envelope functions
$F$	Magnitude of the force
$\bar{F}$	Average force along tunnel path
$\mathbf{F}$	Force acting on a particle
$G_{\text{BTBT}}$	Band-to-band generation rate
$G_{\text{Kane}}$	Generation rate in the Kane model
$G(\mathbf{r})$	Position dependent generation rate
$G^{\lessgtr R,A}(\mathbf{r}, \mathbf{r}'; E)$	Lesser, greater, retarded and advanced Green's functions
$\hbar$	Reduced Planck constant
$H_{\text{nm}}(z)$	Heterostructure layer-dependent Hamiltonian matrix elements
$\hat{H}_0$	Second quantized non-interacting Hamiltonian
$\hat{H}_{\text{int}}$	Interaction Hamiltonian
$\hat{H}_{\text{E}}, \hat{H}_{\text{EP}}, \hat{H}_{\text{P}}$	Second quantized electron, electron-phonon and phonon Hamiltonian
$I_{\text{ds}}$	Drain-source current
$I_{\text{g,dleak}}$	Gate and drain leakage current
$J_{\text{I,T}}$	Current carried by incoming and transmitted wave
$\mathbf{J}_{n,p}(\mathbf{r})$	Electron and hole current density
$k$	Wave vector
$k_{\text{B}}$	Boltzmann constant
$\kappa$	Imaginary part of the wave vector
$\mathbf{k}$	Wavevector in three-dimensional reciprocal space
$\mathbf{K}$	Wavevector in two-dimensional reciprocal space
$l_0$	Characteristic tunneling length
$l_{\text{tun}}$	Tunneling distance
$\ell$	Set of quantum numbers
$L$	Gate length
$m_0$	Free electron mass
$\mu$	Fermi level
$\mu_{n,p}$	Electron and hole mobility
$M'_{\mathbf{k}_0}$	Electron-phonon interaction strength
$n, m$	Discrete quantum numbers
$n_{\text{i}}$	Intrinsic carrier concentration
$n_{\text{eq,bulk}}$	Bulk electron concentration in equilibrium
$n(\mathbf{r})$	Electron concentration
$\nu(E)$	Bose-Einstein distribution
$\hat{N}$	Number operator
$\Omega$	Region inside semiconductor
$p$	Momentum matrix element in the 2-band model
$\mathbf{p}_{nm}$	Momentum matrix element from $\mathbf{k} \cdot \mathbf{p}$ theory
$\phi_{\text{s}}$	Surface potential energy
$\psi(\mathbf{r})$	Wave function

$\psi_{n\mathbf{k}}(\mathbf{r})$	Bloch function
$\hat{\psi}(\mathbf{r}), \hat{\psi}^\dagger(\mathbf{r})$	Field operators
$\hat{\psi}(\mathbf{r}t)$	Field operator in the Heisenberg picture
$P_{\text{static,active}}$	Static and active power consumption
$q$	Elementary charge
$Q_g$	Gate charge
$\rho_{\text{DOS}}(\mathbf{r})$	Position dependent density of states
$\mathbf{r}$	Three dimensional position vector (x,y,z)
$\rho_{\text{DOS,bulk}}$	Bulk density of states
$\hat{\rho}$	Many-particle density matrix
$R(\mathbf{r})$	Recombination rate
$\mathbf{R}$	Two-dimensional position vector (y,z)
$\mathcal{R}$	Lattice vector
$\Sigma^{\leq \text{R,A}}(\mathbf{r}, \mathbf{r}'; E)$	Lesser, greater, retarded and advanced Self energies
$t_{\text{ox}}$	Oxide thickness
$\theta(t)$	Heaviside step function
$T$	Temperature
$T_{\text{Kane}}$	Tunneling probability in the Kane model
$T_{\mathcal{R}}$	Lattice translation operator
$T(E)$	Tunneling probability
$T_v^{\text{abs,em}}(E)$	Probability to have tunneling from the valence band under absorption or emission of a phonon
$u_{n0}(\mathbf{r})$	Periodic basis functions
$u_{n\mathbf{k}}(\mathbf{r})$	Periodic part of a Bloch wave function
$\tilde{u}_{nm\mathbf{k}}$	Components of $u_{n\mathbf{k}}(\mathbf{r})$ in reciprocal space
$U_0$	Rectangular potential barrier height
$U(\mathbf{r})$	Potential energy
$U_{\text{lat}}(\mathbf{r})$	Lattice potential energy
$U_{\text{ext}}(\mathbf{r})$	External potential energy
$v_{\text{I,T}}$	Velocity of the incoming/transmitted wave
$V_{\text{dd}}$	Supply voltage
$V_{\text{FB}}$	Flat-band voltage
$V_{\text{GS}}$	Gate-source voltage
$V_\epsilon$	Electromotive force
$V(\mathbf{r})$	Electrostatic potential
$X_{nn'}(\mathbf{k})$	Interband matrix elements
$\mathcal{Z}$	Grand Canonical Partition function

**Acronyms and Abbreviations**

BGN	Band-Gap-Narrowing
BTBT	Band-to-Band Tunneling
CNT	Carbon Nanotube
DFT	Density Functional Theory
DIBL	Drain Induced Barrier Lowering
FCC	Face-Centered Cubic
FET	Field-Effect transistor
FIQC	Field-Induced Quantum Confinement
GAA	Gate-All-Around
GIDL	Gate-Induced Drain Leakage
I-MOS	Impact Ionization MOS
KKB	Keldysh-Kadanoff-Baym
MLDA	Modified Local-Density Approach
MOSFET	Metal-Oxide Field-Effect Transistor
NEGF	Non-Equilibrium Green's Functions
SOI	Silicon on Insulator
SS	Subthreshold Slope
TCAD	Technology Computer-Aided Design
TFET	Tunnel Field-Effect Transistor
WKB	Wentzel-Kramer-Brillouin

# Contents

Preface	i
Abstract	iii
Korte Inhoud	v
Glossary	vii
Contents	xi

## I Introduction

<b>1 Introduction</b>	<b>3</b>
1.1 Goal and organization of the thesis . . . . .	3
1.2 The MOSFET . . . . .	5
1.2.1 The “miracle” of scaling . . . . .	5
1.2.2 MOSFET working principle . . . . .	8
1.3 The tunnel field-effect transistor . . . . .	14
1.3.1 TFET as a future low-power device . . . . .	14
1.3.2 Basic working principle . . . . .	15
1.3.3 Semiclassical simulation of TFET . . . . .	17

1.3.4	TFET configurations . . . . .	18
1.3.5	Alternative sub-60mV subthreshold slope devices . . . . .	23
1.4	Band-to-band tunneling . . . . .	24
1.4.1	Tunneling in a single band . . . . .	25
1.4.2	Semiconductor bulk bandstructure . . . . .	28
1.4.3	Envelope function formalism . . . . .	31
1.4.4	The Kane model: BTBT in direct semiconductors . . . . .	33
1.5	Quantum transport . . . . .	35
1.5.1	The Non-Equilibrium problem . . . . .	36
1.5.2	Tunneling current in the single-particle picture . . . . .	37
1.5.3	Wigner Distribution Functions . . . . .	38
1.5.4	Non-equilibrium Green's functions . . . . .	39
1.5.5	Balance equations . . . . .	40
1.5.6	Considerations with respect to band-to-band tunneling . . . . .	41
1.6	Thesis contributions . . . . .	43
1.7	Chapters and publications overview . . . . .	54
1.A	Appendix: band structure calculation . . . . .	56
1.B	Appendix: many-particle systems . . . . .	60
1.B.1	Describing a quantum system . . . . .	60
1.B.2	Statistical Mechanics . . . . .	63
	Bibliography . . . . .	66

## II Studying the TFET using analytical models

<b>2</b>	<b>Analytical model for line tunneling</b>	<b>75</b>
2.1	Introduction . . . . .	77
2.2	Analytical model . . . . .	78

2.2.1	Device structure . . . . .	78
2.2.2	Potential profile . . . . .	79
2.2.3	Current . . . . .	80
2.3	Approximating formula . . . . .	83
2.3.1	Derivation . . . . .	83
2.3.2	Interpretation . . . . .	85
2.4	Comparison with simulation results . . . . .	86
2.5	Conclusion . . . . .	88
2.A	Parameters used . . . . .	88
2.A.1	Medici parameters used for Fig. 2.5 . . . . .	88
2.A.2	Medici parameters used for Fig. 2.6 . . . . .	88
2.A.3	Device parameters used for Table 2.1 . . . . .	89
	Bibliography . . . . .	89
<b>3</b>	<b>Analytical model for point tunneling</b>	<b>91</b>
3.1	Introduction . . . . .	93
3.2	Current in a TFET . . . . .	94
3.3	Point tunneling . . . . .	95
3.4	Line tunneling . . . . .	98
3.5	Conclusion . . . . .	99
	Bibliography . . . . .	103

**III Quantum mechanical modeling of BTBT in direct semiconductors**

<b>4</b>	<b>BTBT in direct semiconductors in non-uniform fields</b>	<b>107</b>
4.1	Introduction . . . . .	109
4.2	Envelope function approximation . . . . .	110

4.3	Two band model . . . . .	111
4.4	Uniform potential profile . . . . .	112
4.5	Piecewise uniform field . . . . .	114
4.6	Current . . . . .	117
4.6.1	Example: abrupt p-n diode . . . . .	119
4.7	Conclusions . . . . .	119
	Bibliography . . . . .	124
<b>5</b>	<b>BTBT in heterostructures</b>	<b>127</b>
5.1	Introduction . . . . .	129
5.2	Heterostructure Envelope Functions . . . . .	130
5.3	Transformation of Basis Functions . . . . .	131
5.4	Transformation Elements via $\mathbf{k} \cdot \mathbf{p}$ . . . . .	132
5.5	Transformed Matrix Elements . . . . .	133
5.6	Interpretation in Different Layers . . . . .	134
5.7	Non-lattice matched heterostructures . . . . .	134
5.8	Restriction of Band Set . . . . .	136
5.9	Conclusion . . . . .	137
	Bibliography . . . . .	138
<b>6</b>	<b>BTBT in graphene based semiconductors</b>	<b>141</b>
6.1	Introduction . . . . .	143
6.2	Graphene based semiconductors: band structure . . . . .	144
6.3	Tunnelling probability . . . . .	146
6.4	Conclusions . . . . .	148
	Bibliography . . . . .	148
<b>7</b>	<b>BTBT in Germanium</b>	<b>151</b>

7.1	Introduction . . . . .	153
7.2	Theoretical BTBT calculation . . . . .	154
7.2.1	Parameter Calculation . . . . .	154
7.2.2	BTBT Generation Rate Calculation . . . . .	156
7.3	Simulations and discussion . . . . .	159
7.3.1	Simulation Settings . . . . .	159
7.3.2	Results of Semiclassical Simulations . . . . .	159
7.3.3	Impact of Field-Induced Quantum Confinement . . . . .	164
7.4	Conclusions . . . . .	167
	Bibliography . . . . .	167
<b>8</b>	<b>Optimal doping and bandgap for TFETs</b>	<b>177</b>
	Bibliography . . . . .	188

**IV Quantum mechanical modeling of BTBT in indirect semiconductors**

<b>9</b>	<b>BTBT in indirect semiconductors</b>	<b>193</b>
9.1	Introduction . . . . .	195
9.2	Envelope function approximation for an indirect semiconductor	197
9.3	Phonons . . . . .	201
9.4	Non-equilibrium density matrix . . . . .	201
9.4.1	Equilibrium . . . . .	201
9.4.2	The quantum Liouville equation . . . . .	202
9.5	The steady state current . . . . .	203
9.5.1	The phonon-assisted tunneling current . . . . .	204
9.5.2	Evaluation of the current using spectral functions . . . . .	205
9.6	A one-dimensional potential profile . . . . .	207

9.6.1	Bulk . . . . .	208
9.6.2	The 1D Uniform field . . . . .	209
9.6.3	1D Non-uniform fields: the WKB approximation . . . . .	211
9.6.4	Non-uniform fields: numerical evaluation . . . . .	212
9.7	Discussion . . . . .	213
9.7.1	Generation rate . . . . .	213
9.7.2	Comparing the different models for a $p - n$ diode . . . . .	215
9.8	Conclusion . . . . .	218
9.A	Appendix: Tunneling probability in a uniform field . . . . .	219
9.B	Appendix: Tunneling probability in a non-uniform field (WKB) . . . . .	220
	Bibliography . . . . .	221
<b>10</b>	<b>Field-induced quantum confinement in TFETs</b>	<b>227</b>
	Bibliography . . . . .	234
<b>11</b>	<b>Modified semiclassical model: quantum confinement</b>	<b>237</b>
11.1	Introduction . . . . .	239
11.2	Quantum mechanical modeling of BTBT . . . . .	241
11.3	Semiclassical models . . . . .	242
11.3.1	The existing semiclassical model . . . . .	244
11.3.2	Modified semiclassical model . . . . .	244
11.3.3	Further limitations to the semiclassical model . . . . .	246
11.4	Conclusion . . . . .	246
	Bibliography . . . . .	247
<b>12</b>	<b>Two-dimensional modeling of BTBT</b>	<b>251</b>
12.1	Introduction . . . . .	253
12.2	Formalism . . . . .	253

12.3	BTBT current in a double gate silicon TFET . . . . .	256
12.3.1	The first device . . . . .	256
12.3.2	The second device . . . . .	257
12.4	Conclusion . . . . .	259
12.5	Acknowledgements . . . . .	262
	Bibliography . . . . .	262
<b>13</b>	<b>Study of counter-doped pockets</b>	<b>265</b>
13.1	Introduction . . . . .	267
13.2	Device structure and physical models . . . . .	268
13.3	Basic properties of the GoSo TFET . . . . .	269
13.3.1	Source Doping Concentration $N_s$ . . . . .	269
13.3.2	Gate Alignment . . . . .	270
13.3.3	Fringing Field . . . . .	271
13.3.4	Body Thickness . . . . .	271
13.4	Counter-doped Pocket . . . . .	272
13.4.1	Impact on the Input Characteristics . . . . .	272
13.4.2	Improved Immunity against Gate Alignment Variability	277
13.4.3	Larger Physical Oxide Thickness Allowed . . . . .	277
13.5	GoSo Ge TFET . . . . .	277
13.6	Conclusions . . . . .	281
	Bibliography . . . . .	281
<b>V</b>	<b>Conclusions</b>	
<b>14</b>	<b>Conclusions and Suggestions for Future Research</b>	<b>287</b>
14.1	Summary and Conclusions . . . . .	287

14.1.1	Part II: Studying the TFET using analytical models . . .	287
14.1.2	Part III: Quantum mechanical modeling of BTBT in direct semiconductors . . . . .	288
14.1.3	Part IV: Quantum mechanical modeling of BTBT in indirect semiconductors . . . . .	289
14.1.4	General conclusion . . . . .	290
14.2	Suggestions for Future Research . . . . .	291
14.2.1	TFET modeling . . . . .	291
14.2.2	BTBT modeling . . . . .	291
14.2.3	General quantum transport . . . . .	292
	Bibliography . . . . .	293
	<b>Publication List</b>	<b>295</b>
	International Journal Papers . . . . .	295
	International Conference Papers . . . . .	296
	International Patents . . . . .	299
	<b>Curriculum Vitae</b>	<b>301</b>

## **Part I**

# **Introduction**



# Chapter 1

## Introduction

As technology and scientific insight move forward, new possibilities open themselves to mankind. In this thesis, we investigate the working principle of an electronic switch exploiting the quantum mechanical tunneling process. The electronic switch under investigation in this work is called the tunnel field-effect transistor (TFET) and could provide a big improvement over existing metal-oxide-semiconductor field-effect transistor (MOSFET) technology.

### 1.1 Goal and organization of the thesis

At the onset of this thesis, only semiclassical calculations using commercial device simulators were available to calculate the current of a TFET device. This contrasts with the TFET working principle which is based on tunneling, a fundamentally quantum mechanical mechanism. Moreover, tunneling is a sensitive process so any quantitative calculations of tunnel currents must be based on accurate models. The goal of this thesis is to go beyond the semiclassical picture and to build a fully quantum mechanical model of the TFET which rigorously accounts for interband transitions induced by non-uniform electric fields, as well as size quantization effects.

The improved models developed in this thesis will help in understanding the working principle of the TFET and this in turn can be used to improve TFET performance. Not only the discovery of quantum mechanical effects which are beneficial but also those which are harmful is important. The design of a TFET then consists of exploiting the beneficial effects while mitigating or avoiding the harmful effects.

Since the TFET drive current is determined by band-to-band tunneling (BTBT), an accurate modeling of the TFET requires an accurate modeling of BTBT. The BTBT models we develop are of course not limited in application to the TFET but can also be used to model BTBT in MOSFET devices. In MOSFETs, improved BTBT models aid in the understanding of the unwanted gate-induced drain leakage (GIDL) current. And with the possible introduction of smaller band gap III-V and germanium MOSFETs in future nodes, modeling of BTBT and GIDL will become more important.

The thesis consists of 5 parts.

Part I contains the introduction (chapter 1). In section 1.2, we introduce the current workhorse of the semiconductor industry: the MOSFET. In section 1.3, the topic of this thesis, the TFET is introduced and an overview of existing literature is given. In section 1.4, we give a proper introduction to BTBT, the working principle of the TFET. In section 1.5, we give an overview of some of the popular quantum transport methods. In section 1.6, a condensed overview and summary of the contributions included in this thesis in parts II-IV is given. Finally, section 1.7 is a publication overview.

In part II, a description of the TFET using semiclassical models is given. The semiclassical models, despite being inaccurate, can provide good insight into the operation and the electrostatics which govern the TFET. We present two analytical models describing the tunneling in the TFET in chapters 2-3.

In part III, quantum mechanical modeling of BTBT in direct semiconductors is described. We extend the existing models for uniform fields to cover the effects of non-uniform fields (chapter 4), we extend the envelope function formalism to heterostructures (chapter 5). We explore the impact of different materials such as graphene (chapter 6) and show that direct tunneling is dominant in germanium (chapter 7). Furthermore, we demonstrate that there is an optimal doping concentration and material's band gap for the TFET (chapter 8).

In part IV, quantum mechanical modeling of BTBT in indirect semiconductors is discussed. We develop a general theory to calculate BTBT in indirect semiconductors under one-, two- and three-dimensional potential profiles (chapter 9). Important quantum mechanical effects for tunnel field-effect transistors such as field-induced quantum confinement are demonstrated and investigated (chapter 10). A modified semiclassical model is introduced capable of accounting for quantum confinement (chapter 11). The modeling of BTBT in a two-dimensional device structure is demonstrated (chapter 12). The two-dimensional modeling of BTBT is used to investigate the effect of counter-doped pockets on TFET performance (chapter 13).

Part V contains the conclusion of our thesis and suggestions for future research in chapter 14.

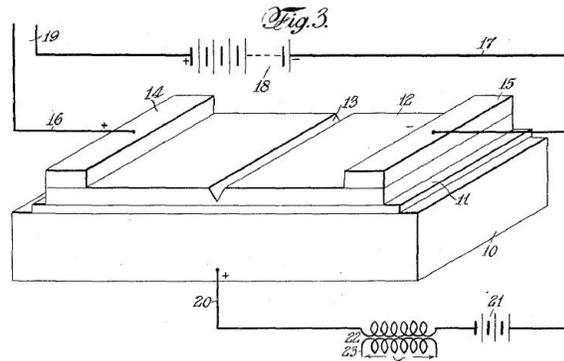


Figure 1.1: Illustration from J.E. Lilienfeld's patent.

## 1.2 The MOSFET

In this section, we describe the workhorse of the semiconductor industry: the MOSFET. We discuss the advantages of miniaturizing the MOSFET, we explain the MOSFET working principle and the semiclassical equations used to describe the MOSFET. And finally, we show how the MOSFET subthreshold slope prohibits a further reduction of the MOSFET supply voltage.

### 1.2.1 The “miracle” of scaling

In 1928, a patent was filed for a “device for controlling current” shown in Fig. 1.1 [1], a variant of a device that is now known as the metal oxide semiconductor field-effect transistor (MOSFET). Because of the difficulty to get a good oxide-semiconductor interface, it took until 1959 before the first MOSFET was fabricated [2]. But now, after many years of research and development a single chip now contains up to 6.8 billion of these devices. Being only 80 years old, the semiconductor industry and its products are already a landmark of human civilization.

At the heart of the success story behind the MOSFET is the concept of scaling. For most commodities, making a commodity smaller means it takes fewer natural resources to produce and this generally translates into a lower price. If we take the example of everyday commodities such as cars or houses: a smaller car or house is generally cheaper than a bigger car or house. The MOSFET is no different in this regard, a smaller MOSFET is cheaper than a bigger one. At the same time ordinary commodities suffer from the paradigm “bigger is better” such that performance goes down with smaller size. Smaller cars or houses can not transport or house the same number of people with the same comfort. In this respect the MOSFET is fundamentally different: smaller MOSFETs are

faster and have improved energy efficiency!

The rise of the MOSFET goes hand in hand with the rise of digital computers. Computers can be programmed to perform a pre-assigned computation based on a given set of inputs facilitating many repetitive or complex tasks. A single MOSFET with a drain-source bias applied can already be considered a computer but not with much functionality: if the gate voltage is high, the drain-source current is high, otherwise it is low. Using two MOSFETs, an inverter can be built: if the inverter input voltage is high, the output voltage is low and vice versa.

This constitutes the basis of binary logic: computation is performed based on a set of input voltages. The exact voltage level is not important only whether the voltage is high or low. By assigning a low voltage to a 0 digit and a high voltage to a 1 digit, computation based on binary numbers can be performed. As an example, a circuit which adds two binary numbers can be built using 10 transistors for each digit of the binary number. By using more transistors, much more functionality such as multipliers, dividers, selectors and memory can be created. The combination of all this functionality creates the computers we know today.

Computer performance is measured by the complexity of the computation it can perform in a given time. In a computer composed of transistors, performance is fundamentally limited by the number of transistors that are available and the speed at which they operate. In most applications, the power consumption of the computer is also a major limiting factor, making transistor power consumption a third very important parameter. The limitation on power consumption can stem from the limited energy supply in the case of a battery-operated computer or from a limited cooling capacity for the computer. So to improve computers we are looking for more transistors which switch faster and dissipate less power.

Scaling MOSFETs results in smaller transistors which operate faster and dissipate less. For this reason, the process of scaling has been very successful and lead the semiconductor industry to a dramatic growth up to a point where it now directly represents 2.5% of the world economy. At the same time, further scaling is desired and research spending at leading semiconductor companies represents more than 10% of company revenue which makes the semiconductor research industry the biggest in the world [3].

In numbers, commercial MOSFET technology has been scaled from 300  $\mu\text{m}$  devices down to 22 nm in 2011. Projections for scaling down to dimensions of 7,4 nm have been made [4] but it is not clear if and how these projections will materialize. The worries for future scaling and finding a way to enable further scaling are the main motivators for this thesis.

Device dimensions	$1/\alpha$
Doping concentration	$\alpha$
Voltage	$1/\alpha$
Current	$1/\alpha$
Capacitance	$1/\alpha$
Delay time/circuit	$1/\alpha$
Power dissipation/circuit	$1/\alpha^2$
Power density	1

Table 1.1: Scaling laws: device performance after scaling device dimensions with a factor  $1/\alpha$ .

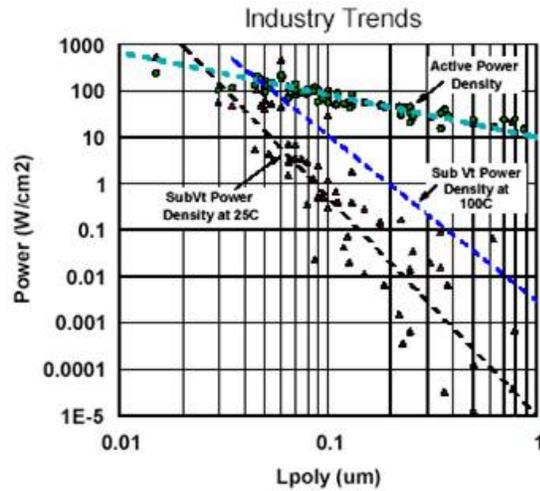


Figure 1.2: Illustration of increasing importance of the static power as scaling progresses.

### MOSFET scaling laws

In 1974, Robert Dennard proposed a set of scaling laws for the MOSFET given in Table 1.1 [5]. Dimensions, voltage, current, capacitance, delay and power consumption/circuit decrease while doping concentrations increase. Dennard's rules have been a good guide for scaled MOSFET performance but nevertheless, computers have witnessed an increase in both dynamic (active) and static (standby) power density as shown in Fig. 1.2.

The increase in power density is due to the inability to further reduce the supply voltage as well as an increased static power consumption. Dennard's scaling law for supply voltage was based on the voltage that is required to obtain a sufficiently high drive current. But with ever decreasing voltages, the threshold

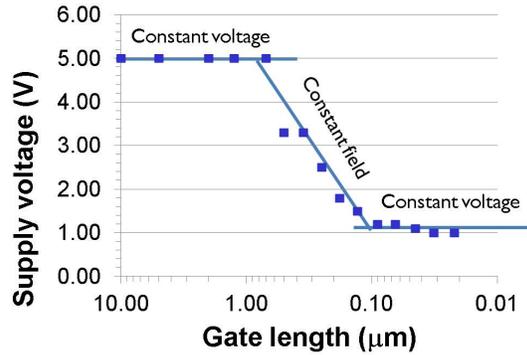


Figure 1.3: Illustration of the end of voltage scaling as dimension scaling progresses.

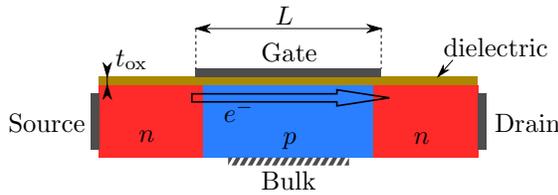


Figure 1.4: Illustration of an  $n$ MOS device indicating source, drain and gate contact as well as the optional bulk contact. If a high gate-source bias is applied electrons flow from the source to the drain (large arrow).

voltage has been decreased and this has led to an increase in off-state current. As a result scaling of supply voltage was forced to continue at a greatly reduced pace as illustrated in Fig. 1.3.

## 1.2.2 MOSFET working principle

The MOSFET has three terminals: a source, a gate and a drain and an optional fourth bulk terminal. The working principle of the MOSFET is quite simple: the voltage between gate and source determines whether there is a high or a low electrical resistance between source and drain. The channel is made of a semiconductor where the amount of charge can be easily modulated by the gate. The gate electrode is a metal or a highly conductive semiconductor and an oxide between the gate and the channel prevents current flow from the gate to the source or the drain. The typical MOSFET structure is shown in Fig. 1.4.

There are two types of MOSFET devices,  $n$ - and  $p$ MOS devices, the distinction between both devices is made depending on which carrier is mediating the transport in the channel. In an  $n$ MOS device, electrons mediate the transport

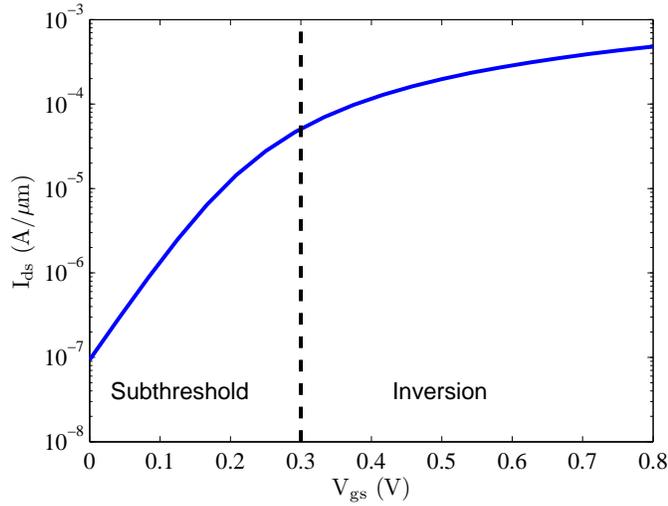


Figure 1.5:  $n$ MOSFET current input characteristics with the subthreshold and inversion voltage range illustrated.

in the channel while in a  $p$ MOS, holes mediate transport. As a result, an  $n$ MOS device turns on at positive gate-source bias while a  $p$ MOS turns on for negative gate-source bias. To explain the working principle in more detail, we study an  $n$ MOS device which has  $n$ -type dopants such as arsenide in its source and drain contacts and  $p$ -type dopants such as boron in the channel.

In an  $n$ MOS, with increasing gate bias ( $V_{gs}$ ) the potential energy for electrons in the semiconductor ( $-qV_{gs}$ ) is decreased and a depletion layer appears in the  $p$ -type doped channel. As the channel is being depleted from its holes, a small number of electrons appear at the interface as well. When the potential energy is further reduced, the electron concentration at the interface exceeds the doping concentration and a layer of electrons, called the inversion layer, appears at the oxide-channel interface. The inversion charge can easily conduct electrons from source to drain. The voltage range preceding the formation of the inversion layer is called the subthreshold voltage range. The subthreshold voltage range is illustrated in a typical MOSFET input characteristic in Fig. 1.5.

### Semiclassical drift-diffusion equations

To gain more understanding into the MOSFET working principle, the physics underlying the electronic transport must be studied. The most popular method to simulate MOSFET operation is using the drift-diffusion equations. In this section we give a quick derivation of the drift-diffusion formalism.

In the semiconductor, there are obviously many electrons and devices are normally studied around room temperature. The behavior of an ensemble of particles at finite temperature is described by statistical mechanics. A key result of statistical mechanics is that in equilibrium the distribution of the non-interacting electrons, which are fermions, over the available energy levels is given by the Fermi-Dirac distribution

$$f(E) = (1 + \exp(\beta(E - \mu)))^{-1} \quad (1.1)$$

where  $E$  is the energy of the state under consideration while  $\mu$  is the Fermi level and  $\beta = 1/(k_B T)$  with  $k_B$  the Boltzmann constant and  $T$  the absolute temperature. The total equilibrium number of charge carriers in a bulk semiconductor is obtained by integrating over all the states:

$$n_{\text{eq,bulk}} = \int_{-\infty}^{\infty} dE \rho_{\text{DOS,bulk}}(E) f(E) \quad (1.2)$$

where  $\rho_{\text{DOS,bulk}}(E)$  gives the bulk density of states.

If we want to study transport in a semiconductor, we must consider electrons in a non-equilibrium configuration. In the drift-diffusion approximation, the non-equilibrium system is described by a position dependent quasi-Fermi level for electrons and holes ( $E_{fn,p}(\mathbf{r})$ ). The concentration of electrons and holes is now determined as a function of position

$$n(\mathbf{r}) = \int_{-\infty}^{\infty} dE \rho_{\text{DOS}}(E, \mathbf{r}) (1 + \exp(\beta(E - E_{fn}(\mathbf{r}))))^{-1}. \quad (1.3)$$

For the density of states, a local approximation is made based on the electrostatic potential ( $V(\mathbf{r})$ ):

$$\rho_{\text{DOS}}(E, \mathbf{r}) = \rho_{\text{DOS,bulk}}(E + qV(\mathbf{r})). \quad (1.4)$$

For non-degenerate semiconductors, only the exponential tail of the Fermi-Dirac distribution is important and the Fermi-Dirac distribution can be approximated by the Maxwell-Boltzmann distribution

$$f(E) \approx \exp(-\beta(E - \mu)). \quad (1.5)$$

Accordingly, the electron charge can be simplified to

$$n(\mathbf{r}) \approx n_i \exp(\beta(E_{fn}(\mathbf{r}) + qV(\mathbf{r}))) \quad (1.6)$$

where  $n_i$  is the intrinsic carrier concentration accounting for the density of states.

Apart from an equation determining the charge density based on the quasi-Fermi levels (Eq. (1.6)), we also need an equation determining the current

based on the quasi-Fermi levels. In the drift-diffusion formalism, the relation between current and quasi-Fermi levels relies on the concept of mobility: to first order, the electrical current is proportional to the applied electric field  $\mathcal{E}$

$$\mathbf{J} = qn\mu_n\mathcal{E} \quad (1.7)$$

with  $q$  the electron charge and  $\mu_n$  the electron mobility. Applying the concept of mobility to the non-equilibrium system described by the quasi-Fermi levels results in an equation for current in the system

$$\mathbf{J}_n(\mathbf{r}) = n(\mathbf{r})\mu_n\nabla E_{fn}(\mathbf{r}). \quad (1.8)$$

Finally, using Eq. (1.6) the expression for the current can be rewritten as

$$\mathbf{J}_n(\mathbf{r}) = -qn(\mathbf{r})\mu_n\nabla V(\mathbf{r}) + \mu_n kT\nabla n(\mathbf{r}). \quad (1.9)$$

The first term of Eq. (1.9) is proportional to the local electric field and is called the drift component while the second term of Eq. (1.9) is proportional to the gradient of the electron concentration and is called the diffusion component. For this reason, the framework employing the quasi-Fermi levels is called the drift-diffusion formalism.

The complete set of semiclassical equations is given in [6]

$$n(\mathbf{r}) = n_i \exp(\beta(E_{fn}(\mathbf{r}) + qV(\mathbf{r}))) \quad (1.10)$$

$$p(\mathbf{r}) = n_i \exp(-\beta(E_{fp}(\mathbf{r}) + qV(\mathbf{r}))) \quad (1.11)$$

$$\nabla^2 V(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon} \quad (1.12)$$

$$\rho(\mathbf{r}) = q(p(\mathbf{r}) - n(\mathbf{r}) + N_D(\mathbf{r}) - N_A(\mathbf{r})) \quad (1.13)$$

$$-\nabla \cdot \mathbf{J}_n(\mathbf{r}) + q(R(\mathbf{r}) - G(\mathbf{r})) = 0 \quad (1.14)$$

$$\nabla \cdot \mathbf{J}_p(\mathbf{r}) + q(R(\mathbf{r}) - G(\mathbf{r})) = 0 \quad (1.15)$$

$$\mathbf{J}_n(\mathbf{r}) = \mu_n n(\mathbf{r})\nabla E_{fn}(\mathbf{r}) \quad (1.16)$$

$$\mathbf{J}_p(\mathbf{r}) = -\mu_p p(\mathbf{r})\nabla E_{fp}(\mathbf{r}) \quad (1.17)$$

The electron and hole concentrations ( $n(\mathbf{r}), p(\mathbf{r})$ ), the electrostatic potential ( $V(\mathbf{r})$ ), the electron and hole quasi-Fermi levels ( $E_{fn}(\mathbf{r}), E_{fp}(\mathbf{r})$ ), the generation and recombination rates  $G(\mathbf{r})$  and  $R(\mathbf{r})$  and the electron and hole current ( $\mathbf{J}_{n,p}(\mathbf{r})$ ) are determined self-consistently.

Most commercial MOSFET simulators are based on the drift-diffusion equations [7, 8].

## MOSFET power consumption

There are two forms of power consumption in a circuit. A first form of power consumption is static power consumption because of leakage through the transistor. A second form of power consumption is active power consumption because of the switching of the transistor.

### Static power consumption

Considering the static power consumption, the leakage current in a MOSFET can originate from gate-source and gate-drain leakage ( $I_{g,\text{leak}}$ ) and from drain-source and drain-bulk leakage in the off-state ( $I_{d,\text{leak}}$ ) resulting in a total power consumption

$$P_{\text{static}} = I_{\text{leak}}V_{\text{dd}} = (I_{g,\text{leak}} + I_{d,\text{leak}})V_{\text{dd}} \quad (1.18)$$

with  $V_{\text{dd}}$  the supply voltage.

Gate-source and gate-drain leakage is caused by tunneling through the gate oxide, and the magnitude of the gate leakage current is determined by the gate oxide barrier height and the gate oxide thickness. Historically, silicon dioxide was the gate dielectric of choice but because of the reduction of gate oxide thickness associated with MOSFET scaling, gate leakage has increased over the years until it became unacceptably high at the 90 nm node. As a result, silicon dioxide was abandoned as a gate oxide material and replaced by hafnium oxide [9].

Drain leakage is the result of three components: subthreshold leakage between source and drain, direct source-to-drain tunneling [10] and gate induced drain leakage (GIDL) from drain to bulk [11]

$$I_{\text{ds,off}} = I_{\text{ds,subthreshold}} + I_{\text{ds,direct tunneling}} + I_{\text{ds,GIDL}} \quad (1.19)$$

Subthreshold leakage originates from the remaining surface charge in the channel in the subthreshold region. The charge concentration and the off-state subthreshold current is determined by the relation

$$I_{\text{ds,subthreshold}} \propto n_{\text{subthreshold}} \propto \exp(\beta\phi_s) \quad (1.20)$$

where  $\phi_s$  is the potential energy inside the channel at the channel-oxide interface. The subthreshold slope (SS) is defined as the derivative of the drain-source current with respect to the applied gate bias

$$\text{SS} \equiv \left( \frac{d \log_{10} I_{\text{ds}}}{dV_{\text{gs}}} \right)^{-1}. \quad (1.21)$$

If the MOSFET has perfect channel control, i.e.  $\phi_s = V_{\text{gs}} + \text{cte}$ , the MOSFET subthreshold slope attains its best possible value of  $\log(10)kT/q = 60 \text{ mV}$  at room temperature.

The presence of the subthreshold slope implicates that for every 60 mV voltage reduction, the subthreshold current decreases with a factor of 10 as a best case scenario. So if the off-state current has to be 5 orders of magnitude below the on-state current, the threshold voltage has to be at least  $5 \times 60 \text{ mV} = 0.3 \text{ V}$ .

Direct source-to-drain tunneling occurs when the channel becomes very short and the electrons tunnel through the channel despite the potential barrier. Direct tunneling is not important at the current technology nodes but is expected to become important with ever decreasing channel lengths.

GIDL is the result of the large electric fields near the drain when the transistor is switched off. Indeed in an *n*MOS, the gate voltage is low while the drain voltage is high in the off-state resulting in a large gate-drain voltage  $|V_{gd}| = V_{dd}$ . The large gate-drain bias causes large electric fields at the drain and this results in band-to-band tunneling (BTBT) causing electrons to tunnel from the drain into the substrate.

### Active power consumption

Active power consumption is related to the switching of the transistors. When the gate voltage is switched on, charge has to be transported into the gate electrode, the semiconductor channel and the source/drain contacts. When the gate voltage is switched off again, the charge has to be evacuated again. This constitutes an irreversible process and induces power consumption.

The energy associated with the charging and discharging of the MOSFET, is related to the gate capacitance

$$C_g = \frac{dQ_g}{dV_g} \quad (1.22)$$

and in the inversion regime, the gate capacitance can be approximated as a parallel plate capacitor

$$C_g \approx C_{ox} = WL \frac{\epsilon_{ox}}{t_{ox}} \quad (1.23)$$

with  $W$  and  $L$  the gate width and length.

Upon switching, the power supply provides charge to the transistor. Defining  $[Q_g]_0^{V_{dd}}$  as the charge that has to be supplied to the transistor to switch from low to high bias, a total energy

$$E^{0 \rightarrow V_{dd}} = V_{dd} [Q_g]_0^{V_{dd}} \quad (1.24)$$

is drawn from the power supply to change the bias from 0 to  $V_{dd}$ . If the capacitance is constant and equal to  $C_{ox}$ , the total energy is given by

$$E^{0 \rightarrow V_{dd}} \approx C_{ox} V_{dd}^2. \quad (1.25)$$

If the switching frequency  $f$  is known, the corresponding active power consumption can be calculated:

$$P_{\text{active}} = fC_{\text{ox}}V_{\text{dd}}^2. \quad (1.26)$$

Eq. (1.26) shows that in order to reduce the active power consumption, a lower supply voltage is required. A simple way to reduce the supply voltage is to reduce the threshold voltage but reducing the supply voltage leads to a higher subthreshold current which leads to a higher static power consumption.

## 1.3 The tunnel field-effect transistor

In this section, we introduce the topic of this thesis, a transistor which does not suffer from the subthreshold limitation: the tunnel field-effect transistor (TFET). We present the TFET working principle, we show how the TFET can be simulated using semiclassical device simulations and we show different TFET device configurations and their advantages. At the end of this section, we also give an overview of alternative devices which do not have the subthreshold slope limitation.

### 1.3.1 TFET as a future low-power device

Given the great importance of power consumption for transistor performance, it is a major nuisance that the scaling of the supply voltage has halted. As shown in previous section, the supply voltage scaling problem boils down to the presence of the MOSFET subthreshold slope. This motivates the search for alternative transistor devices which do not suffer from a subthreshold slope limitation.

In the 1980s and 1990s, transistor action was observed in gated diode structures [12–16]. These gated diodes consist of a highly doped  $p$ -region and a highly doped  $n$ -region with an undoped or intrinsic ( $i$ ) region in between and are called  $p$ - $i$ - $n$  diodes. The underlying mechanism of these gated  $p$ - $i$ - $n$  diode transistors was shown to be Zener tunneling [17] also known as Band-to-Band tunneling (BTBT) or interband tunneling. For this reason, the device is called the tunnel field-effect transistor (TFET). Like the MOSFET, the TFET has three terminals: source, gate and drain but in a TFET, no fourth contact towards the “bulk” is available. The basic TFET configuration is shown in Fig. 1.6.

Whereas the MOSFET operating principle is the modulation of charge in the channel, the TFET operating principle is the modulation of the potential in the device enabling or disabling tunneling from the valence to the conduction

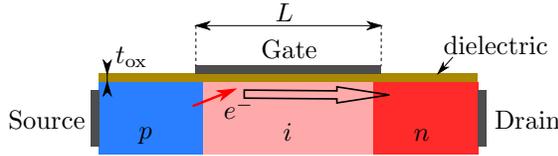


Figure 1.6: Illustration of an  $n$ TFET device indicating source, drain and gate contact. If a high gate-source bias is applied electrons tunnel from the valence band of the source into the conduction band in the channel (solid arrow) and then continue to flow towards the drain (hollow arrow).

band. The MOSFET subthreshold off-current is related to the small remaining charge in the channel while the TFET off-current is related to the tunnel paths which are available in the off-state. As a consequence, the TFET does not have a 60 mV/decade limit on its subthreshold slope.

If a TFET with an average slope steeper than 60 mV/decade is demonstrated, the TFET will enable the reduction of the supply voltage while maintaining a low off-current. Maintaining low off-current results in a low static power consumption while the lower supply voltage results in a lower active power consumption. Finally resulting in a TFET which is a better low-power transistor than the MOSFET.

As a final goal, TFETs with supply voltages below 0.5V while maintaining an off-current in the pA/ $\mu\text{m}$  range are aimed at [4].

### 1.3.2 Basic working principle

In this section, we explain the working principle of the TFET based on the semiclassical picture. We study an  $n$ TFET device which has a  $p$ -doped source and  $n$ -doped drain and turns on for high gate bias.

First of all, the TFET is a reverse biased  $p$ - $i$ - $n$  diode. In general, a reverse biased diode only carries a small current and current densities are of the order of  $10^{-12} - 10^{-15}$  A/cm<sup>2</sup>. For a device with a thickness of 10 nm, this corresponds to a current of  $10^{-16} - 10^{-19}$  A/ $\mu\text{m}$  which is orders of magnitude below the required off-state current.

Since we are interested in transistor action, the reverse biased diode must also have a state where it conducts a large current. There are three different ways in which reverse biased diodes can carry a large current [18]: thermal instability, band-to-band tunneling (BTBT) and impact ionization. The use of thermal instability will not result in a low-power device and the use of impact ionization for transistor action is discussed in section 1.3.5. The TFET exploits BTBT for its drive current.

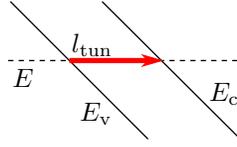


Figure 1.7: Illustration of the tunnel distance in the semiclassical picture.

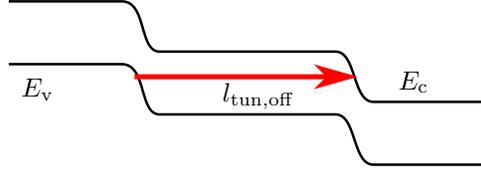


Figure 1.8: Illustration of the TFET in the off-state.

The efficiency of BTBT depends on the tunnel distance ( $l_{\text{tun}}$ ) which is defined as the distance between valence and conduction band for a given energy ( $E$ ) in the semiclassical picture as illustrated in Fig. 1.7. And as a first order approximation, the BTBT generation rate is exponentially dependent on  $l_{\text{tun}}$

$$G_{\text{BTBT}} \propto \exp(-l_{\text{tun}}/l_0) \quad (1.27)$$

where  $l_0$  is a characteristic length for a given semiconductor material.

In the absence of any field-effect from the gate, the TFET behavior resembles that of a  $p$ - $i$ - $n$  diode and the tunnel distance is large and proportional to the length of the intrinsic region as illustrated in Fig. 1.8. The amount of BTBT is negligible and the TFET is in the off-state.

To study the impact of the gate voltage, we first define a flat-band voltage ( $V_{\text{FB}}$ ). When the flat-band voltage is applied between the gate and the source contact ( $V_{\text{gs}} = V_{\text{FB}}$ ), there is no band bending inside the source due to the gate voltage. When a gate voltage higher than the flat-band voltage is applied ( $V_{\text{gs}} > V_{\text{FB}}$ ), a depletion region will appear in the source while for lower gate bias ( $V_{\text{gs}} < V_{\text{FB}}$ ), an accumulation layer will appear.

In the intrinsic region in the semiconductor right under the gate dielectric, the gate electrode exerts great electrostatic control over the potential energy at the surface

$$\phi_s \approx -q(V_{\text{gs}} - V_{\text{FB}}). \quad (1.28)$$

For very high gate bias:  $qV_{\text{gs}} > qV_{\text{FB}} + E_g \gg qV_{\text{FB}}$ , the surface potential  $\phi_s$  will exceed the band gap  $E_g$  and a path from valence to conduction band will emerge near the source which is illustrated in Fig. 1.9. The tunnel distance is now much smaller than the length of the intrinsic region and a dramatic

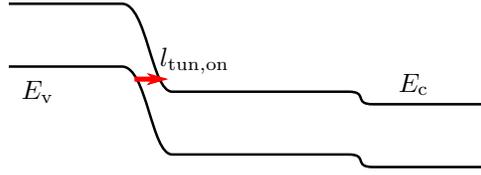


Figure 1.9: Schematic illustration of the reduced tunnel distance in a TFET in the on-state.

increase in BTBT occurs. Further increasing the gate bias further decreases the tunneling distance and increases the drain-source current.

### 1.3.3 Semiclassical simulation of TFET

To be able to account for Gate-Induced Drain Leakage (GIDL) which is a BTBT-based leakage current [11], semiclassical MOSFET simulators were modified to be able to account for BTBT. In the drift-diffusion formalism, there is no immediate way to account for quantum mechanical effects but BTBT can be introduced by an additional BTBT generation and recombination term in Eqs. (1.14-1.15). In most situations, BTBT is a generation process so the term is generally referred to as the BTBT generation rate ( $G_{\text{BTBT}}$ ) while a recombination process is modeled as a negative generation rate.

The first generation of semiclassical BTBT models which were implemented are so-called local models. Electrons and holes are generated in the entire simulated structure and the generation rate is calculated from the magnitude of the local electric field [19–24], i.e. the local gradient of the electrostatic potential

$$\mathbf{F} = q\mathcal{E} = -q\nabla V(\mathbf{r}) \quad (1.29)$$

with  $\mathbf{F}$  the force and  $\mathcal{E}$  the electric field. Furthermore, the generation rate is weighed with an occupation factor taking the number of occupied and empty electron and hole states into account. The Hurkx model for BTBT gives us the expression [25]

$$G_{\text{BTBT}} = \frac{np - n_i^2}{(n + n_i)(p + n_i)} AF^D \exp(-B/F) \quad (1.30)$$

with  $A$  and  $B$  material parameters,  $D = 2$  for direct semiconductors and  $D = 2.5$  for indirect semiconductors and  $F$  is the magnitude of the force  $|\mathbf{F}|$ . If the product of the electron and hole concentration exceeds the square of the intrinsic concentration, the generation rate is negative, otherwise it is positive.

More advanced models use so-called non-local models which define tunnel paths inside the semiconductor structure. The tunnel paths used in the non-local

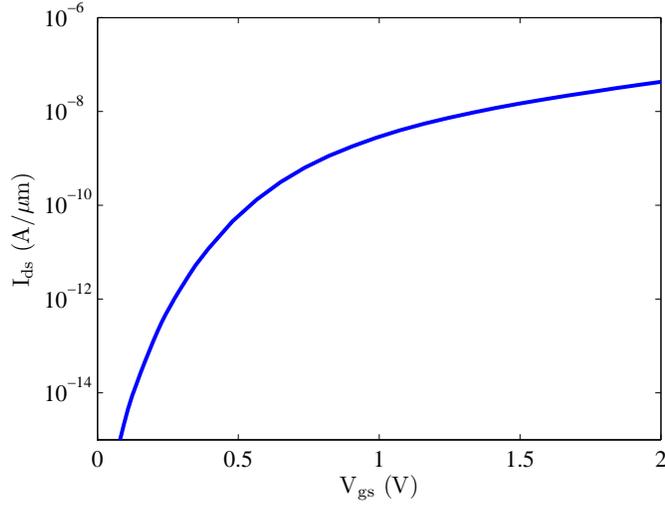


Figure 1.10: TFET current input characteristics.

model start at the valence band and end at the conduction band at equal energy. Rather than using the local electric field, the magnitude of the average electric field ( $\bar{\mathcal{E}}$ ) or the average force ( $\bar{F}$ ) over the tunnel path is used. The average field is calculated from the ratio of the band gap and the tunnel path length  $q\bar{\mathcal{E}} = \bar{F} = E_g/l_{\text{tunn}}$ . An alternative non-local approach is to calculate the tunneling probability based on the Wentzel-Kramer-Brillouin (WKB) approximation along the tunnel path. The generation rate corresponding to the thus determined electric field is included in the drift-diffusion equations in the generation rate. The holes are generated at the start of the tunnel path while the electrons are generated at the end of the tunnel path.

In order to understand the working principle of the TFET, a straightforward approach is to use the semiclassical device simulators used for MOSFETs and apply them to the TFET structure. It turns out that the local BTBT models are useless to describe the TFET current as they can not capture the sudden onset of BTBT when the conduction and valence band align. The non-local BTBT models capture the basic TFET working principle but can not capture some important quantum effects as we will show in this thesis.

### 1.3.4 TFET configurations

In the following section, we describe various TFET configurations and briefly give the advantages of the different structures.

### Ambipolar behavior and the short gate concept

The TFET transistor action is determined by the potential at the source. But if we consider the influence of the gate on the potential at the drain which has  $n$ -type doping, exactly the analog TFET behavior is observed for negative bias. As a result, the TFET switches on, both for negative gate bias (drain side) and for positive gate bias (source side), and this is called ambipolar behavior. From an application perspective, ambipolar behavior is generally unwanted as it increases the off-current.

A solution to the problem of ambipolar behavior was proposed by Verhulst et al. by the introduction of the short gate configuration [26]. In the short gate configuration, the gate is not over the entire intrinsic region but is limited to the region next to the source. By not extending the gate to the drain, the gate loses control in the region close to the drain while the gate retains its control over the region close to the source thus suppressing the ambipolar behavior. An alternative solution to suppress ambipolar behavior is the use of lower doping concentration at the drain [27].

### Vertical tunnel field-effect transistor

MOSFET technology has evolved on silicon wafers where source and drain are fabricated inside the silicon wafer and where the gate stack is deposited on top of the wafer surface. Devices like this, where the active region is inside the silicon are called planar devices.

To improve MOSFET performance, research was done towards methods of improving the control of the gate over the channel. A first way to improve gate control is to reduce the thickness of the channel which was done in silicon-on-insulator (SOI) technology. A better way to improve gate control is to introduce a second gate making a double-gate structure where the channel is sandwiched between both gates. The best gate control is obtained with the introduction of a cylindrical channel where the gate is all-around.

To fabricate double-gate and gate all-around structures, the planar configuration can no longer be used and the channel must be created by a growing or an etching process. This leads to the introduction of vertical MOSFET structures [28] and so-called FinFET structures [29]. In a vertical MOS, the channel is first created by a growing or an etching process. After the creation of the channel, a gate dielectric and gate metal is deposited around the channel. The source and drain contact are located at the top and bottom of the vertical structure.

For the TFET, the improved channel control when using a double-gate structure (Fig. 1.11) results in improved TFET characteristics [30]. Additional improvement can be obtained by the gate-all around configuration illustrated

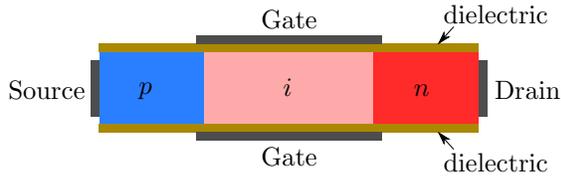


Figure 1.11: Illustration of double gate TFET.

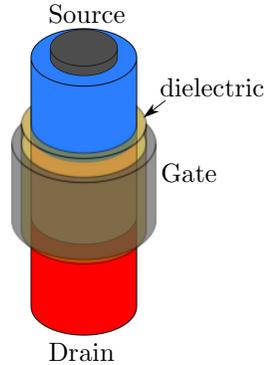


Figure 1.12: Illustration of vertical TFET with the gate all-around the channel.

in Fig. 1.12. The use of these vertical structures presents additional possible advantages [31]. A grown channel allows the possibility to dope the semiconductor during growth which should result in a better control of the doping at the source-channel interface which is very important for TFET operation. Furthermore, a grown channel facilitates the introduction of a heterostructure.

A disadvantage of a vertical process is that the gate is no longer self-aligned with the source and drain. But for a TFET this problem is alleviated as the alignment of the gate with the drain is not critical because the short gate concept can be used [26].

### Carbon nanotube tunnel field-effect transistor

With the discovery of carbon nanotubes [32] and graphene [33], carbon has become available as a semiconductor material. Graphene is a single sheet of carbon atoms, arranged in a hexagonal lattice. A graphene sheet which is rolled up is called a carbon nanotube (CNT). A small sheet of graphene is called a graphene nanoribbon and exhibits similar properties as CNTs. We will limit the following discussion to CNT's but similar behavior is expected for graphene nanoribbons.

While bulk graphene is a semi-metal and shows no band gap, in a CNT, the

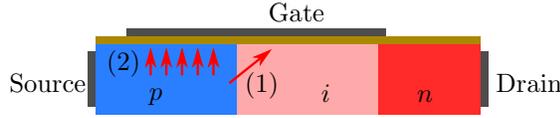


Figure 1.13: Illustration of the point tunneling contribution (1) and line tunneling contribution (2) in a TFET with the gate over the source.

translation symmetry is broken and a chiral and a translation vector can be identified. Because the graphene is rolled up, the chiral vector is no longer continuous but is now quantized. Depending on the direction of the chiral vector and the diameter of the CNT, the CNT is a semimetal or a semiconductor.

The CNTs with a band gap can be used to fabricate semiconductor devices. Furthermore, the semiconducting CNTs also show BTBT and as a result, a gated  $p-i-n$  CNT structure operates as a TFET [34]. The operational principle of the CNT TFET is the same as for the ordinary TFET, except that the gate is wrapped around the carbon nanotube [35]. The gate controls the intrinsic channel and increases the electric field which leads to improved BTBT and transistor action.

CNTs are an interesting material for TFET research as the CNT band gap depends on its diameter and making a TFET from a small band gap CNT will result in high TFET on-currents. On the other hand, CNTs are not yet considered a viable successor for silicon logic due to the lack of control over the CNT chirality and diameter during fabrication.

### Point and line tunneling

When a TFET is constructed with its gate extending over the  $p$ -type source and the intrinsic channel of the  $p-i-n$  diode, two different components of tunneling can be distinguished and both components are illustrated in Fig. 1.13 [36]. The first component tunnels from the source towards the channel, this component is referred to as point tunneling or lateral tunneling. The second component tunnels inside the source, i.e. the tunneling starts in the valence band of the source and tunnels towards the gate dielectric where it tunnels into the conduction band near the gate dielectric interface. This component is referred to as line tunneling or vertical tunneling. Point tunneling generally sets in at lower voltages than line tunneling while line tunneling can result in higher tunneling currents. The line tunneling current will be proportional to the length of the overlap of the gate over the source.

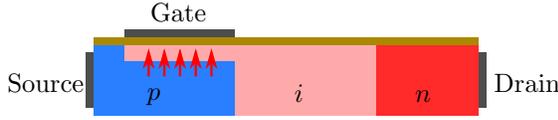


Figure 1.14: Illustration of a TFET with the gate over the source only and with an undoped pocket also known as the hybrid TFET. Promoting line tunneling and suppressing point tunneling will lead to an improved subthreshold slope.

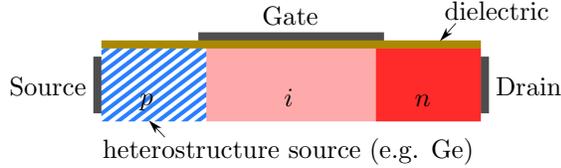


Figure 1.15: Illustration of a heterostructure TFET where the source and channel material are different.

### TFET with the gate over the source only

In the “standard” configuration, the TFET is a gated  $p$ - $i$ - $n$  diode with its gate over the intrinsic region. In Fig. 1.13, it is shown that the TFET can also have the gate over the source and this leads to the line tunneling component. Taking the idea of the short gate one step further, the overlap of the gate with the intrinsic region can be limited altogether such that the gate is exclusively located on top of the source [37].

Eliminating the overlap of the gate over the channel suppresses the point tunneling contribution and can lead to a steeper TFET onset. Further promoting line tunneling with respect to point tunneling by introducing an undoped pocket or a counter-doped pocket will further improve the subthreshold slope [38, 39]. An illustration of the TFET with the gate over the source only and with an undoped pocket also known as the hybrid TFET is shown in Fig. 1.14.

### Heterostructure TFET

In order to improve the TFET on-current, heterostructures have been proposed [40] where the source material is different from the channel material as illustrated in Fig. 1.15. Introducing a smaller band gap material with a beneficial band alignment can greatly reduce the tunnel distance as illustrated in Fig. 1.16. As a result of the reduced tunneling distance, heterostructure TFETs will have an increased the tunnel efficiency and a higher drive current.

Using a germanium-silicon heterostructure, silicon can be retained as a channel

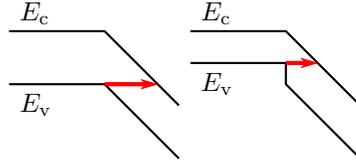


Figure 1.16: Illustration of the tunnel path in a homostructure (left) and in a heterostructure (right). By choosing a heterostructure with a beneficial band alignment, the tunneling distance from valence band to conduction band will greatly reduce the tunnel distance and improve the TFET on-current.

material while the source consists of germanium. Because of the germanium source, the tunneling efficiency is greatly improved since the distance from the germanium valence band to the conduction band is much shorter than in a comparable all-Si TFET. Complementary to the germanium source  $n$ TFET, we introduce an  $\text{In}_x\text{Ga}_{1-x}\text{As}$  heterostructure in Ref. [41], and show a  $p$ TFET performance complementary to that of the germanium-source TFET.

### 1.3.5 Alternative sub-60mV subthreshold slope devices

#### Impact-ionization MOS

An alternative device capable of beating the 60 mV/decade subthreshold slope is the impact-ionization MOS (I-MOS) [42, 43]. The operating principle of the I-MOS is the process of impact ionization also known as avalanching. During impact ionization, electrons in the conduction band gain an amount of energy which is larger than the band gap, over a small distance. The electrons transfer the energy they have gained to electrons in the valence band which are excited from the valence to the conduction band. In turn, these electrons are accelerated by the electric field gaining energy and exciting more electrons from the valence to the conduction band.

Triggering the avalanche process with a gate electrode, I-MOS devices can be fabricated which switch steeper than 60 mV/decade. Nevertheless, the I-MOS can not be considered as a successor of the MOSFET since the impact-ionization process can only take place if the drain-source bias exceeds the band gap. Furthermore, the process of impact ionization produces many hot carriers which are very detrimental to device reliability.

#### Ferroelectric gate dielectric

In the derivation of the limit for the subthreshold slope, we have assumed that the best case scenario for the MOSFET is that the gate potential perfectly modulates the surface potential ( $\phi_s = V_{gs} + \text{cte}$ ) and as a result  $\partial\phi_s/\partial V_{gs} = 1$ .

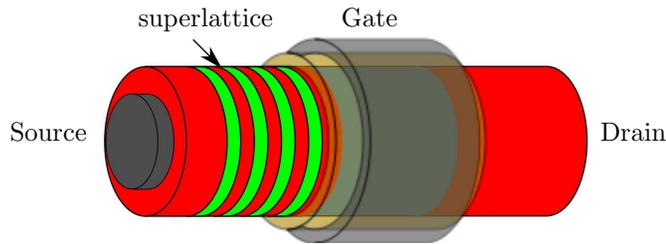


Figure 1.17: Illustration of the superlattice source or steep slope nanowire FET. The source is replaced by a superlattice introducing a forbidden gap in the conduction band enabling sub-60 mV/decade operation.

It turns out that the use of ferroelectric materials (materials with an internal electrical polarization) inside the gate dielectric can result in a “negative” gate capacitance. When a gate dielectric has a negative capacitance, this means the surface potential can change faster than the applied gate bias and a better than 60 mV/decade subthreshold slope can be realized [44].

There is active research into MOSFETs with a ferroelectric gate dielectric but with current ferroelectric materials, gate stacks including 400nm ferroelectric are suggested making them impractical for scaled devices [45].

### Superlattice source FET

The superlattice source or steep slope nanowire FET is a MOSFET which has a superlattice, i.e. a heterostructure with a given periodicity, in its source as illustrated in Fig. 1.17 [46]. Since the superlattice forms a periodic lattice, the destructive or constructive Bragg reflection of electrons will introduce a forbidden gap inside the conduction band. The forbidden gap inside the conduction band will filter out the hot electrons in the source and as a result, sub-60 mV/decade operation can be obtained.

## 1.4 Band-to-band tunneling

In this section, we discuss Band-to-Band Tunneling (BTBT). We start by describing tunneling within a single band and show how the tunneling probability can be calculated from the wavefunctions and, in particular, how the tunneling probability can be approximated using the WKB approximation. We proceed by describing how the semiconductor lattice potential leads to a band structure with multiple bands and how breaking the lattice symmetry leads to a complex band structure and enables tunneling from one band to another. Next, we describe the envelope function formalism which is capable of describing a wave-

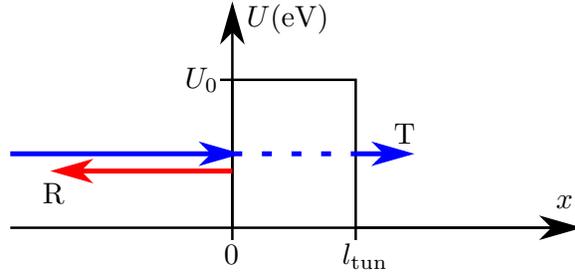


Figure 1.18: Rectangular potential used to illustrate the tunneling process. An incoming electron is illustrated and in the quantum mechanical picture, there is a finite tunneling probability to transmit (T) through the barrier.

function in the presence of an external potential superimposed on the lattice potential. Finally, at the end of the section, we show the Kane model which describes BTBT in a uniform field.

### 1.4.1 Tunneling in a single band

Tunneling is a quantum mechanical phenomenon where a particle traverses spatial areas which are not accessible according to classical mechanics, e.g. potential barriers. Tunneling stems directly from the particle-wave duality introduced in quantum mechanics. In everyday life tunneling is rarely directly observed. Classical mechanics gives a good description of objects which are large and heavy in comparison to the elementary particles. But for the electron, which is an elementary particle, tunneling effects can become very important.

#### Rectangular barrier

A simple model explaining the tunneling process is found by solving the one-electron Schrödinger equation in one dimension in the presence of a rectangular barrier

$$-\frac{\hbar^2}{2m_0} \frac{d^2}{dx^2} \psi(x) + U(x)\psi(x) = E\psi(x) \quad (1.31)$$

with  $m_0$  the electron mass. The rectangular barrier with its barrier height exceeding the particle energy ( $U_0 > E$ ) is given by

$$U(x) = \begin{cases} 0 & \text{if } x < 0 \\ U_0 & \text{if } 0 \leq x \leq l_{\text{tunn}} \\ 0 & \text{if } l_{\text{tunn}} < x \end{cases} \quad (1.32)$$

and the potential profile and the tunneling process are illustrated in Fig. 1.18.

The solutions to the Schrödinger equation providing the eigenfunctions and energy eigenvalues for a uniform potential are plane waves of the form  $\exp(\pm ikx)$ . Based on the uniform potential solutions, the solution to the Schrödinger equation with the rectangular barrier is given by

$$\psi(x) = \begin{cases} a^+ \exp(ikx) + a^- \exp(-ikx) & \text{if } x < 0 \\ b^+ \exp(\kappa x) + b^- \exp(-\kappa x) & \text{if } 0 \leq x \leq l_{\text{tun}} \\ c^+ \exp(ikx) + c^- \exp(-ikx) & \text{if } l_{\text{tun}} < x \end{cases} \quad (1.33)$$

with

$$k = \sqrt{\frac{2m_0 E}{\hbar^2}} \quad \text{and} \quad \kappa = \sqrt{\frac{2m_0(U_0 - E)}{\hbar^2}}. \quad (1.34)$$

Continuity of the wavefunction and its first derivative at  $x = 0$  and  $x = l_{\text{tun}}$  results in the requirements

$$\begin{bmatrix} 1 & 1 \\ ik & -ik \end{bmatrix} \begin{bmatrix} a^+ \\ a^- \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \kappa & -\kappa \end{bmatrix} \begin{bmatrix} b^+ \\ b^- \end{bmatrix} \quad (1.35)$$

and

$$\begin{bmatrix} e^{\kappa l_{\text{tun}}} & e^{-\kappa l_{\text{tun}}} \\ \kappa e^{\kappa l_{\text{tun}}} & -\kappa e^{-\kappa l_{\text{tun}}} \end{bmatrix} \begin{bmatrix} b^+ \\ b^- \end{bmatrix} = \begin{bmatrix} e^{ikl_{\text{tun}}} & e^{-ikl_{\text{tun}}} \\ ik e^{ikl_{\text{tun}}} & -ik e^{-ikl_{\text{tun}}} \end{bmatrix} \begin{bmatrix} c^+ \\ c^- \end{bmatrix}. \quad (1.36)$$

To calculate the tunneling probability, take a wavefunction which only has a transmitting wave for  $x > l_{\text{tun}}$ , i.e.  $c^- = 0$ . The tunneling probability now equals the ratio of the current of the transmitted wave to the current of the incoming wave

$$T = \frac{J_{\text{T}}}{J_{\text{I}}} = \frac{|c^+|^2 v_{\text{T}}}{|a^+|^2 v_{\text{I}}} \quad (1.37)$$

with  $v_{\text{I,T}}$  the velocity of the incoming and the transmitted wave respectively. In the case of the rectangular barrier the velocity of the incoming wave is the same as that of the transmitted wave  $v_{\text{I}} = v_{\text{T}} = \hbar k / m_0$ . After some algebra, the transmission coefficient can be evaluated [47]:

$$T = \left( 1 + \frac{U_0^2 \sinh(\kappa l_{\text{tun}})^2}{4E(U_0 - E)} \right)^{-1} \quad (1.38)$$

and for large  $l_{\text{tun}}$  and  $E \ll U_0$ , this can be approximated by

$$T \approx \left( \frac{4E(U_0 - E)}{U_0^2} \right) e^{-2\kappa l_{\text{tun}}}, \quad (1.39)$$

and for an energy  $E = U_0/2$ , the tunneling probability attains a value of

$$T \approx e^{-2\kappa l_{\text{tun}}}. \quad (1.40)$$

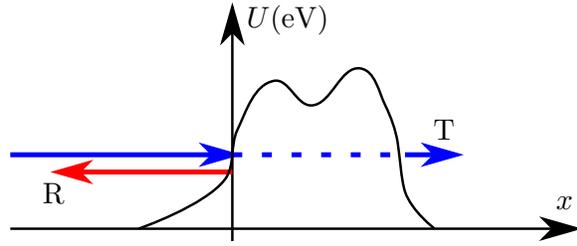


Figure 1.19: Illustration of a non-rectangular potential barrier. The tunneling probability can be calculated from a numerical solution of the Schrödinger equation or using the WKB approximation.

Eq. (1.39) shows that in quantum mechanics, there is a small but finite probability to transmit through a barrier even if the energy of the incoming particle does not exceed the barrier height. It is important to note that the tunneling process does not alter the barrier in any way and tunneling does not induce any damage. A second incoming electron will see the exact same barrier and have the same tunneling probability as the first electron striking the barrier.

To put things into perspective, we can calculate the probability for an electron to tunnel through a sheet of paper. If the sheet of paper presents a barrier with a height of the order of 1 eV and a thickness of  $100\ \mu\text{m}$ , the tunneling probability is of the order of  $1/10^{445000}$ . For all practical purposes  $1/10^{445000}$  equals zero. This means that tunneling is not directly observable at dimensions which we are used to in everyday life.

### Non-rectangular barrier: the WKB approximation

The rectangular barrier presents a textbook example of tunneling but in a realistic situation, the tunneling barrier is not rectangular as illustrated in Fig. 1.19.

A numerical solution of the Schrödinger equation is required to calculate the tunneling probability. But when the potential barrier is sufficiently smooth, an approximate way to calculate the tunneling probability is the Wentzel-Kramer-Brillouin (WKB) approximation. In the WKB approximation the tunneling probability is given by

$$T = \exp\left(-2 \int_0^{l_{\text{tun}}} dx \kappa(x)\right) \quad (1.41)$$

where  $\kappa(x)$  is the complex wavevector inside the tunneling barrier

$$\kappa(x) = \sqrt{\frac{2m_0(U(x) - E)}{\hbar^2}}. \quad (1.42)$$

In the WKB approximation, the wavefunction is approximated by

$$\psi(x) \approx \exp\left(\int_0^x dx' \kappa(x')\right). \quad (1.43)$$

Substituting Eq. (1.43) into the Schrödinger equation yields

$$-\frac{\hbar^2 \kappa(x)^2}{2m_0} - \frac{\hbar^2 \kappa(x)}{2m_0} \frac{d\kappa(x)}{dx} + U(x) = E. \quad (1.44)$$

The error induced by the approximation is given by the second term and shows that the WKB approximation is valid in the limit of  $d\kappa(x)/dx \rightarrow 0$ . So when the second term in Eq. (1.44) is small, the WKB approximation is valid.

It can be shown that the WKB approximation corresponds to ignoring the reflections due to the potential changes. As a result, for the rectangular barrier presented in the preceding section, the error induced by the WKB approximation is given by the prefactor of Eq. (1.39) and is not always negligible. But for more smoothly changing potentials, a smaller error can be expected.

## 1.4.2 Semiconductor bulk bandstructure

In a crystalline semiconductor, the periodic nature of the lattice creates a set of forbidden and allowed energies and BTBT is the process of electrons traversing the forbidden regions. To describe which energies are allowed and disallowed, the Schrödinger equation for the bulk semiconductor must be solved and this results in the description of the bulk bandstructure. We proceed without a formal introduction to the periodic crystal problem but the formal derivation of the Bloch theorem is available in appendix 1.A together with an overview of the methods to calculate semiconductor bandstructures.

### Valence band, conduction band and band gap

In a semiconductor crystal, the crystal potential energy  $U_{\text{lat}}(\mathbf{r})$  is periodic, the Hamiltonian is given by

$$\left[-\frac{\hbar^2 \nabla^2}{2m_0} + U_{\text{lat}}(\mathbf{r})\right] \psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (1.45)$$

and the wavefunctions can be written as Bloch functions

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r}) u_{n\mathbf{k}}(\mathbf{r}) \quad (1.46)$$

where  $\mathbf{k}$  is a reciprocal wavevector,  $n$  is the band index and  $u_{n\mathbf{k}}(\mathbf{r})$  are periodic functions.

The bandstructure is given by the eigenvalues of the Hamiltonian, i.e. the eigenenergies  $E_{n\mathbf{k}}$  associated with the Bloch functions  $\psi_{n\mathbf{k}}(\mathbf{r})$  describe the

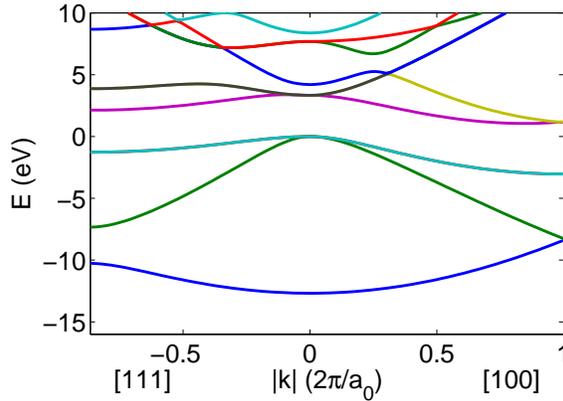


Figure 1.20: Band structure for silicon determined using the local empirical pseudopotential method.

bandstructure. To visualize the bandstructure, the energy eigenvalues are usually plotted as a function of the wavevector along the high symmetry axes of the system. As an example, the silicon band structure obtained from pseudopotential calculations [48] is plotted in Fig. 1.20.

In Fig. 1.20, the band gap which extends from  $E = 0$  eV to  $E = 1.12$  eV in silicon can be observed. There are states available with an energy lower than 0 eV and states available for energies higher than 1.12 eV but there are no states with an energy inside the band gap. The bands below the band gap are called the valence bands and the bands above the band gap are called the conduction bands. BTBT is defined as tunneling through the band gap.

### Complex band structure

The bandstructure as illustrated in Fig. 1.20 describes the allowed energies for electrons inside the semiconductor assuming translational invariance. But in a device, contacts, dielectrics and applied biases will break the translational symmetry and if translational invariance is broken, the bulk band structure no longer describes the semiconductor. Nevertheless, even when the symmetry is broken, many of the features of revealed by the bulk bandstructure such as the band gap can still be observed and the WKB approximation can be extended towards semiconductors using the complex band structure.

Consider an electron with a given energy  $E$  which is injected into the semiconductor. If the energy of the electron is within the band gap then propagation of the electron through the semiconductor is not possible as there are no bulk states inside the semiconductor with the energy  $E$ . If a thin slab of semicon-

ductor is considered, the wavefunction of the electron will decay exponentially but there will still be a small probability of transmission through the semiconductor slab because of tunneling. So in order to study tunneling and the decay of the wavefunction, we have to study crystals where translational symmetry is broken.

To be able to calculate the complex band structure, translational symmetry must be broken. If we take a thin slab of semiconductor where the lattice potential inside the slab is the same as for the full crystal, the same Schrödinger equation as for the bulk system applies in the slab

$$\left[ -\frac{\hbar^2 \nabla^2}{2m_0} + U_{\text{lat}}(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad : x \in \text{semiconductor} \quad (1.47)$$

but the potential energy no longer has the translational invariance in the  $x$ -direction.

In a given subspace inside the semiconductor, the potential is invariant under a selected set of translations, i.e.

$$T_{\mathcal{R}} U_{\text{lat}}(\mathbf{r}) = U_{\text{lat}}(\mathbf{r}) \quad : \mathbf{r} \in \Omega \quad (1.48)$$

with  $\Omega$  denoting a region inside the semiconductor,  $\mathcal{R}$  a lattice vector and  $T_{\mathcal{R}}$  corresponds to a translation such that  $\Omega + \mathcal{R}$  is still within the semiconductor. The set of translation operators under which the Hamiltonian is invariant in a given subspace  $\Omega$  no longer forms a group since this set of translation operators does not satisfy the closure requirement.

Within the subspace  $\Omega$ , the Hamiltonian commutes with the translation vector. So it is possible to find eigensolutions to the Hamiltonian within the subspace  $\Omega$  which are also eigensolutions to the translation operator in the subspace  $\Omega$ :

$$\psi(\mathbf{r}) = c_{\mathcal{R}} \psi(\mathbf{r} + \mathbf{R}) \quad : \mathbf{r} \in \Omega. \quad (1.49)$$

Contrary to the bulk crystal, the eigenvalues of the translation eigenvalue  $c_{\mathcal{R}}$  no longer have to be a representation of a group and can have any value. For this reason, the solutions of the Hamiltonian in the subspace can be taken to have the same Bloch form as for the bulk system

$$\psi(\mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r}) u_{n\mathbf{k}}(\mathbf{r}) \quad (1.50)$$

with  $u_{n\mathbf{k}}(\mathbf{r})$  still periodic but now  $\mathbf{k}$  can be any complex vector.

The complex band structure is described by all complex  $\mathbf{k}$ -values for a given energy  $E$  contrasting with the real bandstructure giving all energy eigenvalues for a given  $\mathbf{k}$ . Using the  $\mathbf{k} \cdot \mathbf{p}$  Hamiltonian, the complex  $\mathbf{k}$  are calculated by rewriting the problem from an energy eigenvalue problem to a second order eigenproblem in  $k_x$  for a given  $k_y, k_z$  and  $E$ :

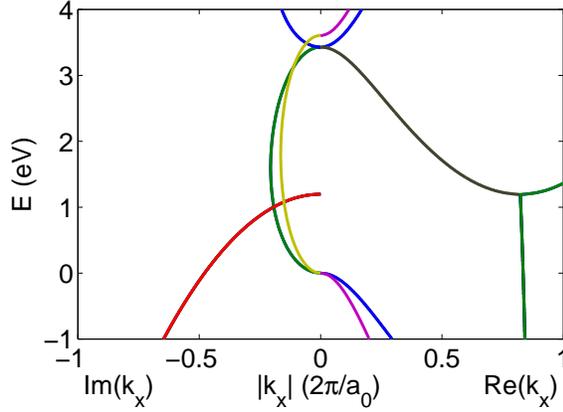


Figure 1.21: Complex bandstructure in the [100] direction for silicon using the  $\mathbf{k} \cdot \mathbf{p}$  method for  $\mathbf{K} = 0$ . The minimum of the conduction band can be observed at  $k_x = 0.852\pi/a_0$ .

$$\frac{\hbar^2 |\mathbf{K}|^2}{2m_0} - E + \sum_{m'} \frac{\hbar^2}{m_0} \mathbf{K} \cdot \mathbf{P}_{mm'} + k_x \sum_{m'} \frac{\hbar^2}{2m_0} (2p_{xmm'} + k_y + k_z) + k_x^2 \frac{\hbar^2}{2m_0} = 0 \quad (1.51)$$

where  $\mathbf{K}$  denotes the 2D-vector  $(k_y, k_z)$ ,  $\mathbf{P}_{mm'}$  is the corresponding vector for the momentum matrix element  $(p_{ymm'}, p_{zmm'})$  and  $p_{xmm'}$  is the x-component of the momentum matrix element. For an introduction to the  $\mathbf{k} \cdot \mathbf{p}$  method to calculate bandstructures see Appendix 1.A.

In Fig. 1.21 we show the complex bandstructure as a function of energy for  $\mathbf{K} = (0, 0)$ .

The complex band structure enables the extension of the WKB approximation from ordinary tunneling to tunneling in semiconductors. The tunneling probability in the WKB approximation is still given by Eq. (1.41) but in a semiconductor the complex wavevector  $\kappa(x)$  describing the decay of the wavefunction has to be computed from the complex band structure.

### 1.4.3 Envelope function formalism

The complex band structure provides a way to extend the WKB approximation from tunneling within a single band to band-to-band tunneling. However, to calculate the tunneling probability beyond the WKB approximation, the Schrödinger equation must be solved under the application of external biases. In this section, we describe a formalism capable of solving the Schrödinger

equation in the presence of a lattice and an external potential: the envelope function formalism.

When an external bias is applied to the semiconductor, the potential is no longer determined by the core ions and electrons alone. The potential energy is now taken to be  $U_{\text{lat}}(\mathbf{r}) + U_{\text{ext}}(\mathbf{r})$  where  $U_{\text{lat}}(\mathbf{r})$  is the potential energy in the absence of an external potential and  $U_{\text{ext}}(\mathbf{r})$  the potential due to the external bias. The external potential does not have lattice periodicity and now presents a new Schrödinger equation to solve:

$$\left[ -\frac{\hbar^2 \nabla^2}{2m_0} + U_{\text{lat}}(\mathbf{r}) + U_{\text{ext}}(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (1.52)$$

A first method to solve the Schrödinger equation with the external potential, is just solving the problem as is. This is possible if an explicit expression is available for  $U_{\text{lat}}(\mathbf{r})$ , for example from pseudopotentials. In the case of pseudopotentials, the Schrödinger equation is solved by constructing a new (bigger) unit cell, composed of many primitive unit cells and  $U_{\text{ext}}(\mathbf{r})$  is approximated by a periodic function on the new unit cell. In the pseudopotential approach, the problem is solved in reciprocal space and the Hamiltonian is dense. With the present-day computational resources, solving the problem for relatively large two-dimensional structures in a short time is not feasible.

An easier alternative to solving the problem directly is to use envelope functions. In the case of envelope functions, we expand the solutions of the Schrödinger equation with the external potential based on the known solutions of the bulk Schrödinger equation without the external potential. More specifically, the wavefunction is written as

$$\psi(\mathbf{r}) = \sum_n f_n(\mathbf{r}) u_{n0}(\mathbf{r}), \quad (1.53)$$

where  $u_{n0}(\mathbf{r})$  are the periodic basis functions and  $f_n(\mathbf{r})$  are the envelope functions. The periodic basis functions  $u_{n0}(\mathbf{r})$  are usually taken from the Bloch solutions around a high symmetry point in reciprocal space. In Fourier space

$$\tilde{\psi}(\mathbf{k}) = \sum_n \sum_m \tilde{u}_{nm} \tilde{f}_n(\mathbf{k} - \mathcal{G}_m) \quad (1.54)$$

with  $\tilde{u}_{nm}$  the components of the Fourier series described by  $u_{n0}(\mathbf{r})$  and  $\tilde{f}_n(\mathbf{k})$  the Fourier transform of the envelope functions. Furthermore, to have a unique expansion of a wavefunction as a sum of envelope functions and basis functions, the Fourier transform of the envelope functions  $\tilde{f}_n(\mathbf{k})$  is restricted to the first Brillouin zone [49].

Substituting Eq. (1.53) into Eq. (1.52) yields

$$\begin{aligned} \sum_n f_n(\mathbf{r}) \left[ \frac{-\hbar^2 \nabla^2}{2m_0} + U_{\text{lat}}(\mathbf{r}) \right] u_{n0}(\mathbf{r}) + \sum_n (\nabla f_n(\mathbf{r})) \cdot \frac{\hbar^2 \nabla}{m_0} u_{n0}(\mathbf{r}) \\ + \sum_n u_{n0}(\mathbf{r}) \left[ \frac{\hbar^2 \nabla^2}{2m_0} + U_{\text{ext}}(\mathbf{r}) \right] f_n(\mathbf{r}) = E \sum_n u_{n0}(\mathbf{r}) f_n(\mathbf{r}). \end{aligned} \quad (1.55)$$

If the Fourier transform of  $f_n(\mathbf{r})$  is restricted to the first Brillouin zone,  $\nabla f_n(\mathbf{r})$  and  $\nabla^2 f_n(\mathbf{r})$  will also be restricted to the first Brillouin zone. At the same time,  $u_{n0}(\mathbf{r})$ ,

$$\nabla u_{n0}(\mathbf{r}) = \sum_{n'} p_{nn'} u_{n'0}(\mathbf{r}) \quad (1.56)$$

and

$$\left[ U_{\text{lat}}(\mathbf{r}) + \frac{\hbar^2 \nabla^2}{2m_0} \right] u_{n0}(\mathbf{r}) = E_n^0 u_{n0}(\mathbf{r}) \quad (1.57)$$

have lattice periodicity. Furthermore, the basis functions are orthogonal

$$\int d^3r u_{n'0}^*(\mathbf{r}) e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} u_{n0}(\mathbf{r}) = \delta_{nn'} \delta(\mathbf{k}-\mathbf{k}') \quad (1.58)$$

if  $\mathbf{k}$  is within the first Brillouin zone and the integral runs over the entire space under consideration.

The envelope function approximation proceeds with the assumption that the last term  $U_{\text{ext}}(\mathbf{r})f_n(\mathbf{r})$  is also within the first Brillouin zone. It then follows from the orthogonality of the basis functions that the set of envelope functions have to satisfy

$$E_n^0 f_n(\mathbf{k}) + \sum_{n'} \frac{\hbar}{m_0} \mathbf{p}_{nn'} \cdot (\nabla f_n(\mathbf{r})) - \frac{\hbar^2 \nabla^2}{2m_0} f_n(\mathbf{r}) + U_{\text{ext}}(\mathbf{r}) f_n(\mathbf{r}) = E f_n(\mathbf{r}). \quad (1.59)$$

The assumption that  $U_{\text{ext}}(\mathbf{r})f(\mathbf{r})$  is restricted to the first Brillouin zone is a good approximation for slowly varying envelope functions and external potentials such that their Fourier components outside of the Brillouin zone are negligible.

#### 1.4.4 The Kane model: BTBT in direct semiconductors

A very popular model to describe band-to-band tunneling in direct semiconductors is using the Kane model. In the following paragraphs we will give a brief overview of the derivation of the Kane model [19].

Take the external potential to be a uniform potential

$$U_{\text{ext}}(\mathbf{r}) = -Fx. \quad (1.60)$$

Expanding the wavefunctions in the Bloch functions which are the solutions to the bulk Hamiltonian

$$\psi(\mathbf{r}) = \sum_n \int d^3k a_n(\mathbf{k}) \exp(i\mathbf{k} \cdot \mathbf{r}) u_{n\mathbf{k}}(\mathbf{r}), \quad (1.61)$$

Kane rewrites the Schrödinger equation in terms of the coefficients  $a_n(\mathbf{k})$ ,

$$E_n(\mathbf{k})a_n(\mathbf{k}) + F \frac{d}{dk} a_n(\mathbf{k}) + F \sum_{n'} X_{nn'}(\mathbf{k}) a_{n'}(\mathbf{k}) = E a_n(\mathbf{k}) \quad (1.62)$$

with

$$X_{nn'}(\mathbf{k}) = \int_{\Omega} d^3r u_{n\mathbf{k}}^*(\mathbf{r}) \frac{d}{dk_x} u_{n'\mathbf{k}}(\mathbf{r}). \quad (1.63)$$

Kane uses Fermi's golden rule to calculate the transition probability. This means Eq. (1.62) is first solved without the interband term  $X_{nn'}(\mathbf{k})$ :

$$a_n^{(0)}(\mathbf{k}) = \frac{1}{\sqrt{F}} \exp\left(\frac{i}{F} \int_0^{k_x} dk'_x (E - E_n(\mathbf{k}'))\right) \exp(i\mathbf{K} \cdot \mathbf{R}) \quad (1.64)$$

while the interaction term  $X_{nn'}(\mathbf{k})$  is treated as a perturbation. The interband transition probability can then be calculated from Fermi's golden rule:

$$T_{\text{Kane}} = \left| \int_{-\infty}^{\infty} dk_x X_{nn'}(\mathbf{k}) \exp\left(\frac{i}{F} \int_0^{k_x} dk'_x (E_n(\mathbf{k}') - E_{n'}(\mathbf{k}'))\right) \right|^2. \quad (1.65)$$

Using a 2-band  $\mathbf{k} \cdot \mathbf{p}$  model, the interaction between both bands is described by

$$X_{cv}(\mathbf{k}) = \frac{i\hbar E_g p}{m_0(E_g^2 + 4\hbar^2 |\mathbf{k}|^2 p^2 / m_0^2)} \quad (1.66)$$

with  $p$  the interband momentum matrix element between valence and conduction band. The distance between conduction and valence band is

$$E_c(\mathbf{k}) - E_v(\mathbf{k}) = \sqrt{E_g^2 + 4\hbar^2 |\mathbf{k}|^2 p^2 / m_0^2}. \quad (1.67)$$

All elements are available to evaluate Eq. (1.65) and Kane proceeds by performing a complex contour integration. For  $\mathbf{K} = (0, 0)$ , the tunneling probability obtained by Kane is:

$$T_{\text{Kane}} = \frac{\pi^2}{9} \exp\left(-\frac{\pi E_g^2 m_0}{4\hbar p F}\right). \quad (1.68)$$

The transmission coefficient  $T_{\text{Kane}}$  gives the probability for an electron with given energy  $E$  and  $\mathbf{K} = (0, 0)$  to transmit through the potential. Integrating Eq. (1.68) over all perpendicular  $\mathbf{K}$  results in tunneling probability per unit

area. Furthermore, the uniform field gives us the possibility to rewrite the transmission probability to a transition probability per unit distance and per unit time:

$$G_{\text{Kane}, \mathbf{K}=(0,0)} = 2 \frac{FT_{\text{Kane}}}{2\pi\hbar} \quad (1.69)$$

where the factor of 2 accounts for spin. The momentum matrix element can be written using the relation

$$p = \frac{m}{2} \sqrt{\frac{E_g}{m_r}} \quad (1.70)$$

with  $m_r^{-1} = m_v^{-1} + m_c^{-1}$  the reduced effective mass. Finally, this results in the final Kane generation rate per unit volume

$$G = \frac{F^2 m_r^{\frac{1}{2}}}{18\pi\hbar^2 E_g^{\frac{1}{2}}} \exp\left(-\frac{\pi E_g^{\frac{3}{2}} m_r^{\frac{1}{2}}}{2\hbar F}\right). \quad (1.71)$$

Eq. (1.71) gives the rate of carrier generation per unit volume. Because the Kane model assumes a uniform field, the force can be related to a tunneling distance  $F = E_g/l_{\text{tun}}$ . Corresponding to the tunneling formula for ordinary single-band tunneling, the current is exponential with respect to the tunneling distance. The prefactor for the generation rate for direct tunneling has a field dependence  $F^2$ . A first factor of  $F$  can be traced down to the number of states per unit area while the second factor of  $F$  relates to the number of states per unit length in the tunneling direction.

Apart from a model for direct semiconductors, a model for indirect semiconductors is also described by Kane and Keldysh in [20–22].

The biggest limitation of the Kane model is that it starts from the assumption of a uniform field while uniform fields never present themselves in a real device. To enable the use of the Kane model in non-uniform fields, the generation rate is calculated using a semiclassical approach. Defining a semiclassical tunnel path running from the valence to the conduction band, the average electric field along the tunnel path can be used to calculate the generation rate. Integrating the generation rate over the entire device, an estimate of the BTBT current can be calculated. Nevertheless, the Kane model does not provide an estimate on how large the error due to the introduction of the semiclassical approach is, motivating the search for improved models.

## 1.5 Quantum transport

The goal of this thesis is to study quantum transport in TFETs but none of the existing quantum transport formalisms were developed with the purpose of

describing BTBT in an efficient way. To illustrate the difficulties or capabilities of the existing quantum transport formalisms to describe BTBT, we provide a brief introduction and overview to a number of the existing formalisms in this section. But this section should by no means be considered as an exhaustive list of all quantum transport formalisms.

Some of the formalisms described in this section make use of quantum field theory and statistical mechanics. We refer readers who are not familiar with field theoretical concepts such as a field operator  $\hat{\psi}(\mathbf{r})$ , second quantized operators  $\hat{A}$  or a grand canonical ensemble density matrix  $\hat{\rho}$  to Appendix 1.B for an introduction.

At the end of this section, we compare the presented formalisms with respect to the calculation of BTBT.

### 1.5.1 The Non-Equilibrium problem

The biggest challenge in developing a quantum transport formalism is dealing with the non-equilibrium character of transport. A force drives the transport and heat is dissipated inside the device. This leads to a steady state situation when the amount of dissipated heat equals the amount of energy put into the system. Nevertheless, once the driving force is removed the system is expected to relax back to its original equilibrium state.

In equilibrium, the time evolution of a system is described by a Hamiltonian  $\hat{H}_0$  and the system exhibits time reversal symmetry. In a grand-canonical ensemble, i.e. allowing for the exchange of energy and particles with the outside world, in equilibrium the system is described by a density matrix

$$\hat{\rho} = \exp(-\beta(\hat{H}_0 - \mu\hat{N}))/\mathcal{Z} \quad (1.72)$$

with  $\hat{N}$  the number operator and  $\mathcal{Z}$  the partition function.

If we introduce a perturbation into the system at  $t = 0$ , we modify the Hamiltonian to become

$$\hat{H} = \hat{H}_0 + \theta(t)\hat{H}_{\text{int}} \quad (1.73)$$

with  $\theta(t)$  the Heaviside step function and  $\hat{H}_{\text{int}}$  the Hamiltonian associated with the perturbation of the system. The equilibrium is disturbed and if the interaction exhibits time reversal symmetry, the system will evolve towards a new equilibrium

$$\hat{\rho} = \exp(-\beta(\hat{H}_0 + \hat{H}_{\text{int}} - \mu\hat{N})). \quad (1.74)$$

The evolution from the first equilibrium towards the second will not be instantaneous and during the transition period, the system will be in non-equilibrium.

For infinitely large systems with a local change in the Hamiltonian, the transition time will be infinitely long and the system will exhibit a non-equilibrium

steady state situation. When describing current through transistors, it is generally assumed that the description of steady state satisfies to describe the behavior of the transistor.

### 1.5.2 Tunneling current in the single-particle picture

In this section, we show how to calculate the tunneling current in the case where the potential changes along one dimension and under the assumption of a small tunneling probability. We define the tunneling direction as the  $x$  direction and  $U(x)$  denotes the potential describing the tunnel barrier extending from  $x = 0$  to  $x = l_{\text{tun}}$ .

Tunneling can generally be treated as a single-particle phenomenon and interaction with phonons and impurities can be considered as second order effects. An exception to the single particle picture is phonon-assisted tunneling which dominates in indirect semiconductors because the large difference in wavevector between the top of the valence band and the bottom of the conduction band prohibits direct transitions.

Single-particle behavior is described by the single-particle three-dimensional Schrödinger equation

$$\left(-\frac{\hbar^2}{2m_0}\nabla^2 + U(x)\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}), \quad (1.75)$$

the solutions form subbands and are of the form

$$\psi(\mathbf{r}) = e^{i\mathbf{K}\cdot\mathbf{R}}\chi^\pm(x; E, \mathbf{K}) \quad (1.76)$$

where  $\mathbf{R}$  and  $\mathbf{K}$  are the coordinates and the wavevector in the  $y$  and  $z$  direction respectively. The subband wavefunctions  $\chi^\pm(x; E)$  have to satisfy the one-dimensional Schrödinger equation

$$\left(\frac{\hbar^2}{2m_0}\frac{d^2}{dx^2} + \frac{\hbar^2\mathbf{K}^2}{2m_0} + U(x)\right)\chi^\pm(x; E, \mathbf{K}) = E\chi^\pm(x; E, \mathbf{K}). \quad (1.77)$$

For an infinite system, the energy spectrum is continuous and the solutions that carry current are doubly degenerate for an energy  $E$  because of time reversal symmetry. As a basis for the doubly degenerate state, we define the wavefunction with a component transmitting towards the right as  $\chi^+(x; E, \mathbf{K})$  and the wavefunction with a transmitting component to the left as  $\chi^-(x; E, \mathbf{K})$ .

In the limit of an infinitely high barrier, the probability to tunnel from left to right becomes zero and vice versa. The left and the right part of the system are separated and when the two systems are separated, it is possible to assign to each system a separate Fermi level,  $\mu_L$  and  $\mu_R$ . Correspondingly, there are

two Fermi-Dirac distributions  $f_L(E)$  and  $f_R(E)$  and the total charge density in the system can be computed from

$$\rho(x) = \int \frac{dE}{2\pi} \int \frac{d^2K}{(2\pi)^2} (f_L(E)|\chi^+(x; E, \mathbf{K})|^2 + f_R(E)|\chi^-(x; E, \mathbf{K})|^2) \quad (1.78)$$

when the wavefunctions are normalized according to

$$\int dx \chi^{\pm*}(x; E, \mathbf{K})\chi^{\pm}(x; E', \mathbf{K}) = \delta(E - E'). \quad (1.79)$$

When there is a finite tunneling barrier, the tunneling probability becomes finite but generally remains small. As a first order approximation, we can still take Eq. (1.78) to hold since

$$|\chi_L^+(x; E, \mathbf{K})|^2 \approx |\chi_{L, \text{infinite barrier}}^+(x; E, \mathbf{K})|^2 \quad \text{for } x < 0 \quad (1.80)$$

$$|\chi_L^+(x; E, \mathbf{K})|^2 \propto T \approx 0 \quad \text{for } x > l_{\text{tun}}, \quad (1.81)$$

but contrary to the infinite barrier, the system now carries current.

In the non-interacting electron picture, the tunneling current per unit area is calculated as

$$J = \frac{2q}{\hbar} \int \frac{dE}{2\pi} \int \frac{d^2K}{(2\pi)^2} (f_L(E) - f_R(E))T(E, \mathbf{K}). \quad (1.82)$$

where the prefactor 2 accounts for spin and the tunneling transmission probability can be calculated by

$$T(E) = \frac{|c^T|^2 v_T}{|c^I|^2 v_I} \quad (1.83)$$

where  $c^{T,I}$  are the amplitudes of the transmitted and incoming waves respectively and  $v_{T,I}$  is the group velocity of the transmitted/incoming wave.

### 1.5.3 Wigner Distribution Functions

Going beyond the single-particle picture, the Wigner distribution function gives a tractable description of the system under consideration and is related to the density matrix by the relation

$$f(\mathbf{r}, \mathbf{k}) = \int d^3s \exp(i\mathbf{k} \cdot \mathbf{s}) \text{Tr}(\hat{\psi}^\dagger(\mathbf{r} - \mathbf{s}/2)\hat{\psi}(\mathbf{r} + \mathbf{s}/2)\hat{\rho}) \quad (1.84)$$

where  $\mathbf{s}$  is the position vector used for a Fourier transform.

The kinetic equation for the Wigner function is given by

$$\frac{\partial}{\partial t} f(\mathbf{r}, \mathbf{k}) + \frac{\mathbf{p}}{m} \cdot \nabla_{\mathbf{r}} - \int d^3k' V_w(\mathbf{r}, \mathbf{k} - \mathbf{k}') f(\mathbf{r}, \mathbf{k}', t) - \left( \frac{\partial f_w}{\partial t} \right)_{\text{coll}} = 0 \quad (1.85)$$

where  $V_w(\mathbf{r}, \mathbf{k})$  is the nonlocal kernel given by

$$V_w(\mathbf{r}, \mathbf{k}) = \frac{1}{i\hbar} \int \frac{d^3s}{(2\pi)^3} \left( V\left(\mathbf{r} + \frac{\mathbf{s}}{2}\right) - V\left(\mathbf{r} - \frac{\mathbf{s}}{2}\right) \right) \exp(-i\mathbf{k} \cdot \mathbf{s}). \quad (1.86)$$

The collision term  $(\partial f_w / \partial t)_{\text{coll}}$  has to be used to account for scattering events due to the electron-phonon or electron-electron interactions.

Transport is described by considering the steady state solution under the application of boundary conditions corresponding to a set of leads. In one dimension with open boundaries at  $x = 0$  and  $x = l$ , the boundary conditions of the Wigner distribution are given by

$$f(0, k)|_{k>0} = f_L(k) \quad (1.87)$$

$$f(l, k)|_{k<0} = f_R(k) \quad (1.88)$$

where  $f_{L,R}(k)$  give the distribution of the electrons inside the left and right contact.

Finally, the current is computed from

$$\mathbf{J} = \int d^3k \frac{\hbar \mathbf{k}}{m} f(\mathbf{r}, \mathbf{k}). \quad (1.89)$$

### 1.5.4 Non-equilibrium Green's functions

In equilibrium, many-body systems can be studied using Green's functions. For non-equilibrium problems an approach with similar structure exists called the non-equilibrium Green's function (NEGF) formalism or also called the Keldysh formalism or the Keldysh-Kadanoff-Baym (KKB) formalism [50]. In NEGF, four Green's functions are used

$$G^>(\mathbf{r}t, \mathbf{r}'t') = -\frac{i}{\hbar} \langle \hat{\psi}(\mathbf{r}t) \hat{\psi}^\dagger(\mathbf{r}'t') \rangle \quad (1.90)$$

$$G^<(\mathbf{r}t, \mathbf{r}'t') = -\frac{i}{\hbar} \langle \hat{\psi}(\mathbf{r}t)^\dagger \hat{\psi}(\mathbf{r}'t') \rangle \quad (1.91)$$

$$G^R(\mathbf{r}t, \mathbf{r}'t') = \theta(t - t') (G^>(\mathbf{r}t, \mathbf{r}'t') - G^<(\mathbf{r}t, \mathbf{r}'t')) \quad (1.92)$$

$$G^A(\mathbf{r}t, \mathbf{r}'t') = \theta(t' - t) (G^<(\mathbf{r}t, \mathbf{r}'t') - G^>(\mathbf{r}t, \mathbf{r}'t')) \quad (1.93)$$

which are labeled the greater, lesser, retarded and advanced Green's function respectively. The field operators are in the Heisenberg picture and the expectation value  $\langle \rangle$  corresponds to the trace over the density matrix describing the system.

In steady state, the Green's functions become independent of the difference between the two time-arguments  $t - t'$ . Fourier transforming with respect to  $t - t'$  results in the energy-dependent Green's functions  $G^{\leq R, A}(\mathbf{r}, \mathbf{r}'; E)$ .

The equations of motion for the Green's functions are given by [51]

$$(E - H_0(\mathbf{r})) G^{\text{R}}(\mathbf{r}, \mathbf{r}'; E) - \int d^3 r_1 \Sigma^{\text{R}}(\mathbf{r}, \mathbf{r}_1; E) G^{\text{R}}(\mathbf{r}_1, \mathbf{r}'; E) = \delta(\mathbf{r} - \mathbf{r}') \quad (1.94)$$

$$G^{\lessgtr}(\mathbf{r}, \mathbf{r}') = \int d^3 r_1 d^3 r_2 G^{\text{R}}(\mathbf{r}_1, \mathbf{r}_2; E) \Sigma^{\lessgtr}(\mathbf{r}_1, \mathbf{r}_2; E) G^{\text{A}}(\mathbf{r}_2, \mathbf{r}'; E) \quad (1.95)$$

where the retarded self-energy  $\Sigma^{\text{R}}$  is related to the lesser and greater self-energies  $\Sigma^{\lessgtr}$ . The self-energies are usually either ignored (ballistic transport) or treated in a self-consistent Born approximation so that they become dependent on  $G^{\lessgtr}(\mathbf{r}, \mathbf{r}')$ . In the self-consistent Born approximation, the relation between  $\Sigma^<$  and  $G^<$  is based on the interactions under consideration.

The applied bias is modeled by introducing so-called contact self-energy  $\Sigma^{\text{RC}}(E)$  which corresponds to imposing transmitting boundary conditions. A non-equilibrium situation is obtained by using an equilibrium Fermi-Dirac distribution in different contacts  $f_{\text{C}}(E)$  and modifying the contact self-energies accordingly

$$\Sigma^{\text{C}}(E) = f_{\text{C}}(E)(\Sigma^{\text{RC}}(E) + \Sigma^{\text{AC}}(E)). \quad (1.96)$$

The non-equilibrium Green's function formalism then proceeds by solving the Green's functions  $G^{\text{R}}(\mathbf{r}, \mathbf{r}'; E)$  and  $G^{\text{C}}(\mathbf{r}, \mathbf{r}'; E)$  self-consistently with the self-energies  $\Sigma^{\text{R}}(\mathbf{r}, \mathbf{r}'; E)$  and  $\Sigma^{\text{C}}(\mathbf{r}, \mathbf{r}'; E)$ .

The current can be calculated from the retarded Green's function

$$\mathbf{J} = \frac{-q\hbar}{4\pi m} \int dE \lim_{\mathbf{r} \rightarrow \mathbf{r}'} (\nabla - \nabla') G^{\text{C}}(\mathbf{r}, \mathbf{r}'; E) \quad (1.97)$$

while simplified equations exist for the ballistic regime.

### 1.5.5 Balance equations

The Balance equation approach departs from the idea that in every quantum system, there are certain quantities which are conserved such as momentum and energy [52]. To change the value of a conserved quantity, interference from the external world is required and in steady state the momentum and energy added to the system must equal the momentum and energy removed from the system.

In the case of quantum transport, the interference of the external world onto the system is the application of a power supply. If we take the power supply to provide its energy through an electromagnetic field according to

$$\mathcal{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}, \quad (1.98)$$

the electrons gain energy in the circuit from the electric field ( $\mathcal{E}$ ). For each time the electron goes around the circuit, it gains an energy

$$q \int_{\Gamma} \mathcal{E} \cdot d\mathbf{r} = qV_{\epsilon} \quad (1.99)$$

where  $\Gamma$  is a closed path going around the circuit once and  $V_{\epsilon}$  is the electromotive force applied to the circuit. If the circuit carries a current  $I$ , the power put into the system by the power supply is given by

$$P = IV_{\epsilon}. \quad (1.100)$$

On the other hand, the energy dissipated inside the system can be calculated from the energy transferred from the electrons to the phonons

$$\frac{\partial}{\partial t} \text{Tr}(\hat{H}_{\text{E}}\hat{\rho}) = \frac{\text{Tr}(\hat{H}_{\text{E}}[\hat{\rho}, \hat{H}_{\text{E}} + \hat{H}_{\text{EP}} + \hat{H}_{\text{P}}])}{i\hbar} \quad (1.101)$$

$$= \frac{\text{Tr}([\hat{H}_{\text{E}}, \hat{H}_{\text{EP}}]\hat{\rho})}{i\hbar} \quad (1.102)$$

where  $\hat{H}_{\text{E}}$ ,  $\hat{H}_{\text{EP}}$ ,  $\hat{H}_{\text{P}}$  denotes the Hamiltonian for the electrons, the electron-phonon interaction and the phonons.

The final energy balance equation reads

$$IV_{\epsilon} = \frac{i}{\hbar} \lim_{t \rightarrow \infty} \text{Tr}([\hat{H}_{\text{E}}, \hat{H}_{\text{EP}}]\hat{\rho}). \quad (1.103)$$

Apart from the energy, also the momentum has to be conserved in the circuit. It can be shown that the momentum gained from the field is

$$\mathbf{F} = \int_{\Omega} d\tau \left[ \varrho(\mathbf{r})(\mathbf{E} - \frac{1}{q} \nabla U + \mathbf{J} \times \mathbf{B}) \right] \quad (1.104)$$

resulting in the balance equation

$$\int_{\Omega} d\tau \left[ \rho(\mathbf{r})(\mathbf{E} - \frac{1}{q} \nabla U + \mathbf{J} \times \mathbf{B}) \right] = \frac{i}{\hbar} \lim_{t \rightarrow \infty} \text{Tr}([\mathbf{P}, \hat{H}_{\text{EP}}]\hat{\rho}). \quad (1.105)$$

### 1.5.6 Considerations with respect to band-to-band tunneling

Although all methods can deal with band-to-band tunneling in principle, and some calculations have been done using the NEGF formalism for small structures [53–55], practical considerations make the use of most methods for large structures and using limited computational power impractical. Wigner distribution functions and Green's functions are a function of two arguments, this

means their memory and computational time scale proportional to the square of the number of mesh points. Given the rapid changes of the wavefunctions in the tunneling regions, many mesh points are required and for a two-dimensional structure all these methods are computationally very expensive.

Balance equations depend on the balance between the dissipation process and the energy fed into the system. When considering tunneling however, the tunnel barrier and not the exact dissipation process is the most important in determining the tunneling current.

In summary, the calculation of BTBT in two-dimensional structures is computationally very expensive using the existing quantum transport formalisms.

## 1.6 Thesis contributions

### Part II: Studying the TFET using analytical models

#### Chapter 2: Analytical model for line tunneling

Chapter 2 introduces the TFET with the gate over the source only. We develop an analytical model for line tunneling (section 1.3.4), which facilitates understanding the operation of the TFET. The electrostatics are described by using the Shockley approximation and are illustrated in Fig 1.22. The tunneling current is calculated from the semiclassical generation rate from Kane. It is shown that a higher doping concentration in the source leads to a higher onset voltage while a thinner gate dielectric reduces the onset voltage. Smaller band gap materials result in a lower onset voltage and a higher on-current.

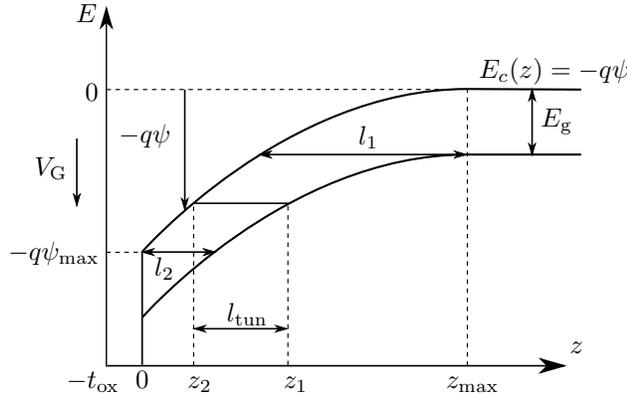


Figure 1.22: Band diagram illustrating the potential along the tunnel paths in the line tunneling configuration as shown in chapter 2.

#### Chapter 3: Analytical model for point tunneling

In this chapter, we briefly summarize the analytical model for line tunneling and proceed by presenting an analytical model for point tunneling. The model is derived under the assumption that the source is very highly doped such that depletion in the doped region can be neglected. Similar to the line tunneling configuration, a lower band gap results in a lower onset voltage and a higher tunneling current. The sensitivity of the TFET current towards the gate oxide thickness is much more pronounced in the case of point tunneling compared to the case of line tunneling.

## Part III: Quantum mechanical modeling of BTBT in direct semiconductors

### Chapter 4: BTBT in direct semiconductors in non-uniform fields

Kane's model for BTBT starts from the assumption that the external potential is described by a uniform field (section 1.4.4) but in real devices, the potential is never uniform. Using a 2-band  $\mathbf{k} \cdot \mathbf{p}$  model and the envelope function formalism (section 1.4.3), we calculate the transmission coefficient in a non-uniform field. For the case of a stepwise uniform field, the Kane result is obtained apart from a factor  $\pi^2/9$  when electric fields are not too strong and the total potential drop in the studied region well exceeds the band gap. In Fig. 1.23, we show the calculated transmission coefficient for a non-uniform field as a function of electric field strength. For intermediately strong electric fields, the non-uniform potential gives a higher transmission probability while for very high fields the Kane result is too optimistic. For the case of a  $p-n$  diode, the WKB approximation gives reasonable results but the current can be seen to be higher for intermediate doping levels and lower for very high doping levels.

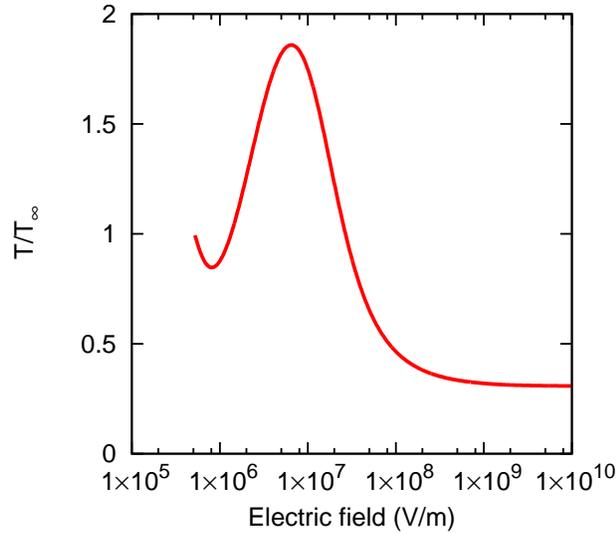


Figure 1.23: Calculation of the transmission probability in InSb compared with the transmission probability calculated by Kane as a function of the electric field for a constant potential difference between the contacts as shown in chapter 4.

## Chapter 5: BTBT in heterostructures

In chapter 5, we present an extension of the envelope function formalism to heterostructures resulting in the heterostructure envelope function equation

$$\sum_m H_{nm}(z)f_m(z) - \frac{\hbar^2}{2m_0} \frac{d^2 f_n(z)}{dz^2} - \frac{i\hbar}{m} \sum_m p_{nm} \frac{df_m(z)}{dz} + U_{\text{ext}}(z)f_n(z) = E f_n(z) \quad (1.106)$$

where the band edge energies  $E_n^0$  from Eq. (1.59) are replaced by a layer-dependent Hamiltonian matrix elements  $H_{nm}(z)$ . To calculate the Hamiltonian matrix elements, we define transformations between the different zone-centre bulk Hamiltonian solutions belonging to different layers and show how to calculate these transformations from the  $\mathbf{k} \cdot \mathbf{p}$  bulk parameters. The formalism enables the extension of the results obtained in chapter 4 to heterostructures [56].

## Chapter 6: BTBT in graphene based semiconductors

In this chapter, we study BTBT in graphene based semiconductors (section 1.3.4). We construct a 2-band  $\mathbf{k} \cdot \mathbf{p}$  Hamiltonian at the  $K$ -point based on symmetry considerations of the basis functions. We show that the  $\mathbf{k} \cdot \mathbf{p}$  Hamiltonian obtained in this way corresponds to the widely used tight binding Hamiltonian. Having written the Hamiltonian for the graphene based semiconductors as a 2-band  $\mathbf{k} \cdot \mathbf{p}$  Hamiltonian, describing BTBT proceeds in the same way as for ordinary semiconductors. We show that contrary to previous claims by Jena et al. [57], there is a smaller tunnel probability in graphene based semiconductors compared to zinc-blende semiconductors for the same band gap (Fig. 1.24). On the other hand, for the same effective mass, tunneling is more probable in graphene based semiconductors. This makes graphene based semiconductors not well-suited for MOSFETs but they are a good material for TFETs.

## Chapter 7: BTBT in Germanium

Germanium is a popular semiconductor material for MOSFET and TFET devices. The germanium band gap is indirect and has a minimum of the  $L$  valleys 0.66 eV above the valence band. Studies of BTBT in germanium generally only consider the indirect gap but the direct gap is only 0.8 eV. In this chapter, we calculate the  $A$  and  $B$  parameters for the direct and the indirect Kane model for silicon, germanium and unstrained  $\text{Si}_{1-x}\text{Ge}_x$ . We show that direct tunneling dominates over indirect tunneling in germanium for uniform fields. In a TFET, both the indirect and the direct component will appear due to the earlier onset of the indirect BTBT. The different contributions are studied using TCAD simulations. The impact of field-induced quantum confinement is estimated based on a back-of-the-envelope calculation.

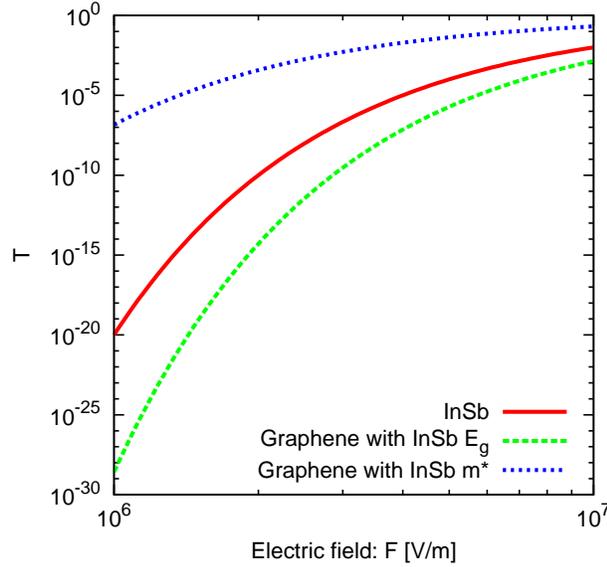


Figure 1.24: Comparison between the BTBT probability of InSb and a graphene based semiconductor with the InSb band gap and a graphene based semiconductor with the InSb effective mass as shown in chapter 6.

## Chapter 8: Optimal doping and bandgap for TFETs

In chapter 2, an analytical model for point tunneling was proposed and the current was calculated based on the Kane generation rate. In chapter 8, we improve over the Kane generation by using a WKB approach (section 1.4.1). The tunneling probability is calculated based on the 2-band  $\mathbf{k}\cdot\mathbf{p}$  model in the WKB approximation and the perpendicular momentum is incorporated by increasing the band gap. We show that the limited density of states results in an optimal doping concentration and also in an optimal band gap to obtain the highest TFET on-current at a given supply voltage (Fig. 1.25). We confirm that our results also hold when taking quantum confinement into account by performing 2-dimensional quantum mechanical simulations for silicon and germanium.

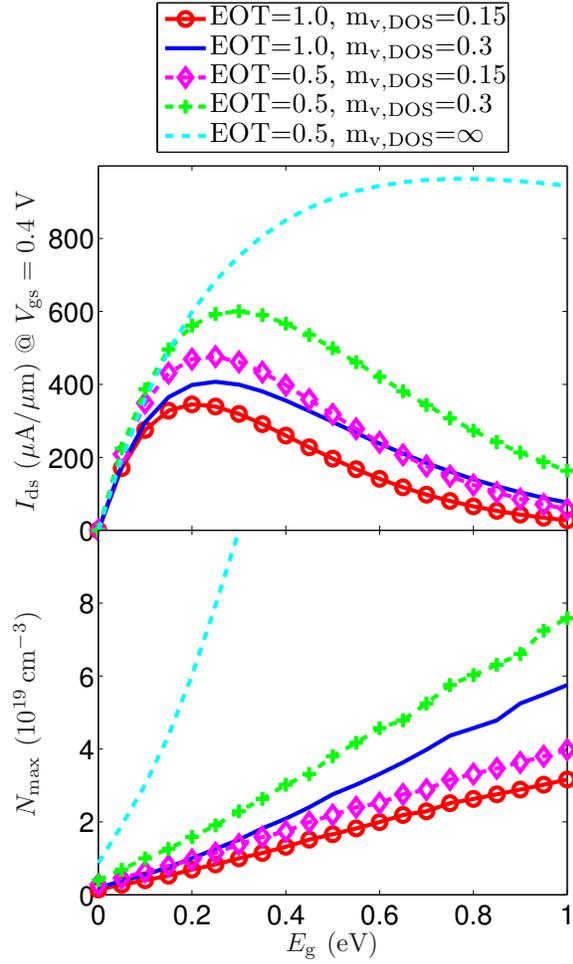


Figure 1.25: Line tunneling current at optimal doping concentration (top) and optimal doping concentration (bottom) as a function of the band gap of the source material at  $V_{gs} = 0.4$  V. Values of EOT and the valence band density of states mass  $m_{v,DOS}$  are given in the legend in the units of nm and  $m_0$ , respectively as shown in chapter 8.

## Part IV: Quantum mechanical modeling of BTBT in indirect semiconductors

### Chapter 9: BTBT in indirect semiconductors

In chapter 9, we develop a general formalism to calculate BTBT in indirect semiconductors, which is then used in the rest of part IV. The development of a new formalism suited for calculating BTBT in indirect semiconductors addresses the difficulty of the existing quantum transport methods to deal with BTBT (section 1.5.6). We start from a semiconductor where the electron-phonon coupling between the valence and conduction band is disabled. As a result, we can assume that the valence and conduction band are in equilibrium with two separate Fermi levels and two Fermi-Dirac distributions for valence and conduction band:  $f_v(E)$  and  $f_c(E)$ . Enabling the electron-phonon interaction connecting the valence and the conduction band, the BTBT current can be calculated from the rate of change of the number of valence or conduction band electrons.

The evaluation of the BTBT current can be simplified by assuming that all phonons which participate in the indirect tunneling process have the same energy  $\hbar\omega_0$  and have the same interaction strength  $M'_{\mathbf{k}_0}$ . After simplification, the current can be evaluated as an integral over energy

$$\begin{aligned}
 I = -\frac{q}{\hbar} \int \frac{dE}{2\pi} & \left( (f_v(E)(1 - f_c(E - \hbar\omega_{\mathbf{k}_0}))(\nu(\hbar\omega_{\mathbf{k}_0}) + 1) \right. \\
 & - f_c(E - \hbar\omega_{\mathbf{k}_0})(1 - f_v(E))\nu(\hbar\omega_{\mathbf{k}_0})) T_v^{\text{em}}(E) \\
 & + (f_v(E)(1 - f_c(E + \hbar\omega_{\mathbf{k}_0}))\nu(\hbar\omega_{\mathbf{k}_0}) \\
 & \left. - f_c(E + \hbar\omega_{\mathbf{k}_0})(1 - f_v(E))(\nu(\hbar\omega_{\mathbf{k}_0}) + 1)) T_v^{\text{abs}}(E) \right) \quad (1.107)
 \end{aligned}$$

where  $\nu(E)$  is the Bose-Einstein distribution and the probability of exciting an electron from the valence band to the conduction band with emission or absorption of a phonon has been introduced:

$$T_v^{\text{abs,em}}(E) = \Omega |M'_{\mathbf{k}_0}|^2 \int d^3r A_v(\mathbf{r}, \mathbf{r}; E) A_c(\mathbf{r}, \mathbf{r}; E \pm \hbar\omega_{\mathbf{k}_0}) \quad (1.108)$$

where  $\Omega$  is the total volume of the system and  $A_v(\mathbf{r}, \mathbf{r}; E)$  and  $A_c(\mathbf{r}, \mathbf{r}; E)$  are the valence and conduction band spectral functions defined as

$$A_{v,c}(\mathbf{r}, \mathbf{r}'; E) = 2\pi \langle \mathbf{r} | \delta(E - H_{v,c}) | \mathbf{r}' \rangle \quad (1.109)$$

$$= 2\pi \sum_{\ell} \chi_{v,c\ell}(\mathbf{r}) \delta(E - E_{v,c\ell}) \chi_{v,c\ell}^*(\mathbf{r}') \quad (1.110)$$

where  $\ell$  runs over all quantum numbers and  $\chi_{v,c\ell}(\mathbf{r})$  are the valence and conduction band wavefunctions with energies  $E_{v,c\ell}$ .

Eq. (1.107) consists of four terms which have a direct physical interpretation: the first term describes tunneling from the valence band to the conduction band while emitting a phonon, the second term describes tunneling from the conduction band to the valence band absorbing a phonon, etc.

Applying the formalism to uniform field case, we come to an expression improving over the existing semiclassical Kane generation rate for indirect BTBT but which agrees with the Kane model in the low-field limit. A semiclassical model for non-uniform fields is derived using the WKB approximation. The formalism is used to study the tunneling in a  $p-n$  junction and is compared with the semiclassical models as illustrated in Fig. 1.26. It is shown that the uniform field model is the best semiclassical model compared to a WKB approximation but a numerical quantum mechanical treatment is required at low bias conditions.

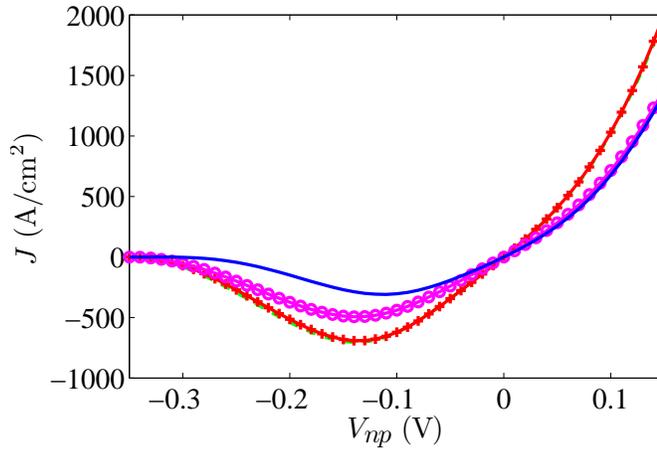


Figure 1.26: Calculation of the current for a symmetrically doped Si diode with a doping level of  $1.5 \times 10^{20} \text{ cm}^{-3}$ . The different current curves are obtained using the semiclassical Kane model (solid line with +), the Uniform field model, (solid line with o) and the WKB model (dashed line, almost coinciding with the Kane model) with respect to the direct numerical calculation (solid line).

## Chapter 10: Field-induced quantum confinement in TFETs

Using the model from chapter 9, we evaluate the tunneling current in a TFET with the gate over the source so that only line tunneling is important. In the line tunneling configuration, the band-to-band tunneling problem becomes a one-dimensional problem which facilitates the implementation and interpretation. The electrostatic potential is solved self-consistently with the quantum mechanical charge density presented by all different conduction and valence band valleys. Calculating the quantum-mechanical phonon-assisted current and comparing with semiclassical results, a large discrepancy can be observed in Fig. 1.27. The main source of the discrepancy is the quantum confinement due to the high electric fields imposed by the gate bias and the barrier presented by the gate dielectric. For high doping concentrations, tunneling to the different conduction band valleys can be distinguished. The electrons with a heavy mass in the tunneling direction have smaller current but do not suffer from quantum confinement as much.

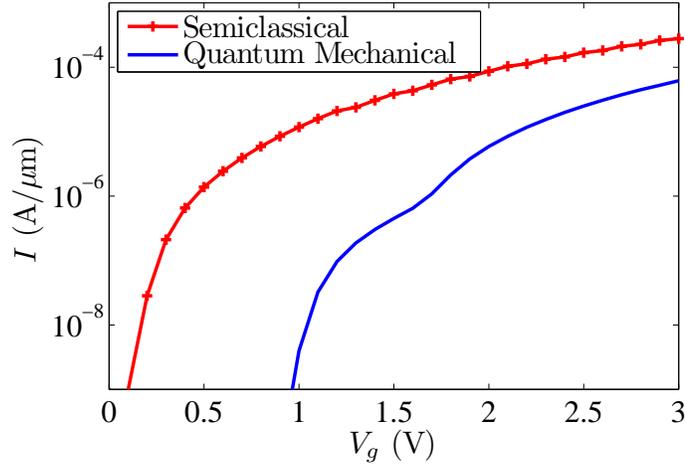


Figure 1.27: Line tunneling current as a function of applied gate bias for a silicon TFET: full quantum mechanical result (solid) and semiclassical Kane model (+). The inflection point of the solid curve around  $V_{gs} = 1.7$  V indicates when transitions to electron states with the lower effective mass in the tunneling direction become predominant as shown in chapter 10.

## Chapter 11: Modified semiclassical model accounting for quantum confinement

In chapter 10, the quantum confinement is studied while modeling the gate dielectric as an ideal hard wall, i.e. the wavefunctions can not penetrate the gate dielectric. But in real gate dielectrics there is some penetration into the gate dielectric. We have modeled the effect of the penetration in the gate dielectric by introducing a decay length. The TFET current when taking the penetration into the gate dielectric into account can be seen to be higher compared to the hard wall case because the quantum confinement effect is slightly alleviated. Nevertheless, the quantum confinement effects remain important. In chapter 10 we also propose a modified semiclassical model rejecting a set of tunnel paths and show that this modified semiclassical model can account for quantum confinement as illustrated in Fig. 1.28.

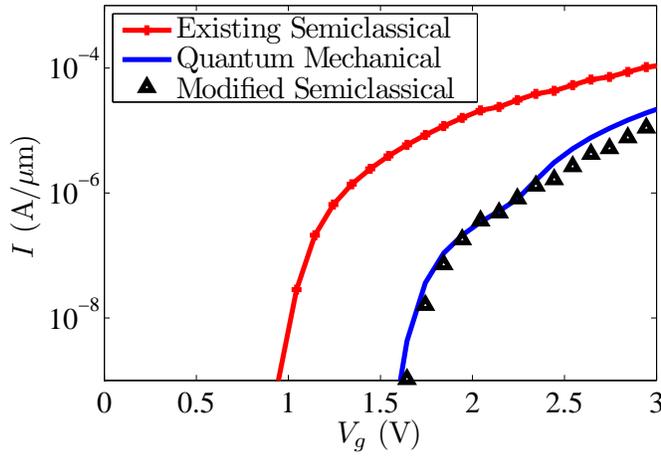


Figure 1.28: Full quantum mechanical current and the existing semiclassical current for a source doping  $N_a = 10^{20}$  at/cm<sup>-3</sup> compared with the modified semiclassical model. Wavefunction penetration is taken into account. Figure as shown in chapter 11.

## Chapter 12: Two-dimensional modeling of BTBT

This chapter shows the implementation of the general formalism to calculate quantum confinement developed in chapter 9 for the case of a two-dimensional double gate TFET structure. While the popular non-equilibrium Green's function formalism employs transmitting boundary conditions, we use open boundary conditions facilitating the diagonalization of the Hamiltonian. We further introduce a quantum mechanical generation rate facilitating the interpretation of the obtained results. The formalism is applied to two double gate TFET structures whose structure is illustrated in Fig. 1.29. The first is an "ordinary" TFET: a double-gate  $p-i-n$  diode with a 10 nm body thickness, high source and drain doping and the gate over the intrinsic region. The second TFET is a TFET where line tunneling is expected: a double-gate  $p-i-n$  diode with a 20 nm body thickness, slightly lower doped source and drain and with the gate extending over the source. The quantum mechanically calculated charge (Fig. 1.30) is solved self-consistently with the potential and the current is calculated by integrating the quantum mechanical generation rate (Fig. 1.31). Comparing with semiclassical results, large differences can be seen and the predicted on-current and subthreshold slopes are worse for the quantum mechanical result due to quantum confinement. And while the semiclassical simulations predict a better performance for the second device compared to the first, the quantum mechanical results show that the opposite is true. This highlights the importance of quantum mechanical simulations when studying TFETs.

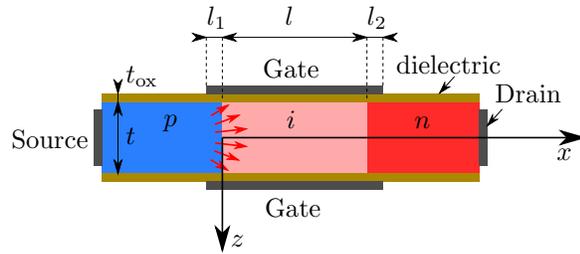


Figure 1.29: Illustration of simulated double gate  $n$ TFET where the red arrows symbolize the tunneling as shown in chapter 12.

## Chapter 13: Study of counter-doped pockets

The two-dimensional quantum mechanical simulation of BTBT in a double gate TFET shown in chapter 12 is used to study a TFET with the gate over the source only. By performing semiclassical and fully quantum mechanical simulations, it is shown that the introduction of counter-doped pockets under the gate reduces the sensitivity towards gate misalignment and oxide thickness

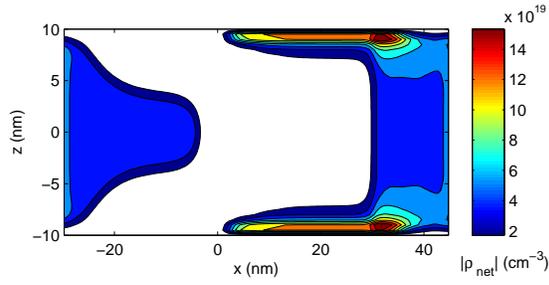


Figure 1.30: Net charge carrier density for device 2 for  $V_{gs} = 2$  V and  $V_{ds} = 0.6$  V as shown in chapter 12.

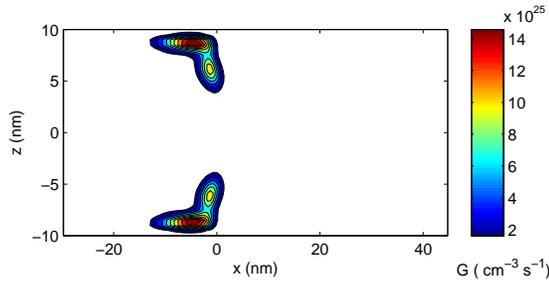


Figure 1.31: Localized tunneling rate for  $V_{gs} = 2$  V and  $V_{ds} = 0.6$  V as shown in chapter 12.

variations. Furthermore, the counter-doped pocket significantly reduces the effect of field-induced quantum confinement.

## 1.7 Chapters and publications overview

The following publications are included in this thesis:

### Part II: Studying the TFET using analytical models

#### Chapter 2: Analytical model for line tunneling

**W. Vandenberghe**, A. Verhulst, G. Groeseneken, B. Sorée, and W. Magnus. Analytical model for a tunnel field-effect transistor. In *Proceedings of the 14th IEEE Mediterranean Electrotechnical Conference - MELECON*, pages 923–928, Ajaccio (France), May 2008.

#### Chapter 3: Analytical model for point tunneling

**W. Vandenberghe**, A.S. Verhulst, G. Groeseneken, B. Sorée, and W. Magnus. Analytical model for point and line tunneling in a tunnel field-effect transistor. In *Simulation of Semiconductor Processes and Devices, 2008. SISPAD 2008. International Conference on*, pages 137–140, sept. 2008.

### Part III: Quantum mechanical modeling of BTBT in direct semiconductors

#### Chapter 4: BTBT in direct semiconductors in non-uniform fields

**W. Vandenberghe**, B. Sorée, W. Magnus, and G. Groeseneken. Zener tunneling in semiconductors under nonuniform electric fields. *Journal of Applied Physics*, 107(5):054520, 2010.

#### Chapter 5: BTBT in heterostructures

M. Van de Put, **W. G. Vandenberghe**, B. Sorée, and G. Groeseneken. An Envelope Function Formalism for Heterostructure Semiconductors. Publishable summary of master thesis 2011-2012.

#### Chapter 6: BTBT in graphene based semiconductors

**W. Vandenberghe**, B. Sorée, W. Magnus, and G. Groeseneken. Zener tunneling in graphene based semiconductors - the k.p method. In *Proceedings of EDISON 16: 16th International Conference on Electron Dynamics in Semiconductors, Optoelectronics and Nanostructures*, pp. 012111, Montpellier (France), August 2009.

#### Chapter 7: BTBT in Germanium

K.-H. Kao, A. S. Verhulst, **W. G. Vandenberghe**, B. Sorée, G. Groeseneken, and K. De Meyer. Direct and indirect band-to-band tunneling in germanium-based tfets. *Electron Devices, IEEE Transactions on*, 59(2):292–301, feb. 2012.

**Chapter 8: Optimal doping and bandgap for TFETs**

**W. G. Vandenberghe**, A. S. Verhulst, K.-H. Kao, K. De Meyer, B. Sorée, W. Magnus, and G. Groeseneken. A model determining optimal doping concentration and material's band gap of tunnel field-effect transistors. *Applied Physics Letters*, 100:193509, 2012.

**Part IV: Quantum mechanical modeling of BTBT in indirect semiconductors****Chapter 9: BTBT in indirect semiconductors**

**W. Vandenberghe**, B. Sorée, W. Magnus, and M. V. Fischetti. Generalized phonon-assisted zener tunneling in indirect semiconductors with non-uniform electric fields: A rigorous approach. *Journal of Applied Physics*, 109(12):124503, jun 2011.

**Chapter 10: Field-induced quantum confinement in TFETs**

**W. G. Vandenberghe**, B. Sorée, W. Magnus, G. Groeseneken, and M. V. Fischetti. Impact of field-induced quantum confinement in tunneling field-effect devices. *Applied Physics Letters*, 98(14):143503, 2011.

**Chapter 11: Modified semiclassical model accounting for quantum confinement**

**W.G. Vandenberghe**, B. Sorée, W. Magnus, G. Groeseneken, A.S. Verhulst, and M.V. Fischetti. Field induced quantum confinement in indirect semiconductors: Quantum mechanical and modified semiclassical model. In *Simulation of Semiconductor Processes and Devices (SISPAD), 2011 International Conference on*, pages 271 –274, sept. 2011.

**Chapter 12: Two-dimensional modeling of BTBT**

**W. G. Vandenberghe**, B. Sorée, W. Magnus, M. V. Fischetti, A. S. Verhulst, and G. Groeseneken. Two-dimensional quantum mechanical modeling of band-to-band tunneling in indirect semiconductors. In *Electron Devices Meeting (IEDM), 2011 IEEE International*, pages 5.3.1 –5.3.4, dec. 2011.

**Chapter 13: Study of counter-doped pockets**

K.-H. Kao, A. S. Verhulst, **W.G. Vandenberghe**, B. Sorée, W. Magnus, D. Leonelli, G. Groeseneken, and K. De Meyer. Optimization of Gate-on-Source-only Tunnel FETs with Counter-doped Pockets. *Electron Devices, IEEE Transactions on*, accepted.

## 1.A Appendix: band structure calculation

In this appendix, we give a formal derivation of the Bloch theorem and an overview of the methods to obtain the band structure of a crystal with a periodic lattice.

### Crystal potential

The semiconductors we study for the TFET are crystal lattices of group IV or a combination of group III and group V elements such as silicon (Si) or gallium arsenide (GaAs) respectively. The semiconductor atoms arrange themselves into a periodic face-centered cubic (FCC) lattice with two atoms in each unit cell. Most of the electrons are tightly bound to a nucleus, these electrons called the core electrons are of no interest for electronic transport. Some of the electrons however are mobile and they can carry current through the semiconductor, we call these electrons the shell electrons.

Using the Born-Oppenheimer approximation and ignoring magnetic effects, the behavior of the shell electrons is governed by a scalar potential determined by the nuclei and the core electrons:  $V_{\text{lat}}(\mathbf{r})$ . Ignoring relativistic effects, the available states to the shell electrons are determined by solving the Schrödinger equation:

$$H_{\text{lat}}\psi(\mathbf{r}) = \left[ -\frac{\hbar^2\nabla^2}{2m_0} + U_{\text{lat}}(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (1.111)$$

with  $U_{\text{lat}} = -qV_{\text{lat}}(\mathbf{r})$  the potential energy for an electron in the crystal potential.

Since we discuss crystalline semiconductors, the nuclei form a periodic lattice and as a consequence the crystal potential energy is also periodic. Denoting the primitive vectors of the crystal as  $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$  and the group of lattice translations as  $T$ , each group element  $T_{\mathcal{R}}$  translates corresponding to a vector  $\mathcal{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$  where  $n_1, n_2$  and  $n_3$  are integers. The periodicity of the potential energy can be formally written as

$$T_{\mathcal{R}}[U_{\text{lat}}(\mathbf{r})] = U_{\text{lat}}(\mathbf{r} + \mathcal{R}) = U_{\text{lat}}(\mathbf{r}). \quad (1.112)$$

### Bloch Theorem

The crystal Hamiltonian is the sum of the kinetic energy and the crystal potential energy. Since the kinetic energy term in the Hamiltonian ( $-\hbar^2\nabla^2/(2m_0)$ ) is invariant under any translation and the crystal potential energy is invariant under lattice translation, the total crystal Hamiltonian is invariant under lattice translation.

The invariance of the Hamiltonian under lattice translation is formally described by the commutation of the Hamiltonian with the group of lattice translation operators:  $H_{\text{lat}}T_{\mathcal{R}} = T_{\mathcal{R}}H_{\text{lat}}$ . The commutation of the Hamiltonian with the lattice translation operators implies that a set of eigenfunctions can be found which are both an eigenfunction of the Hamiltonian and of the lattice translation operators.

The eigenfunctions of the translation operator are defined as

$$T_{\mathcal{R}}[\psi(\mathbf{r})] = c_{\mathcal{R}}\psi(\mathbf{r}). \quad (1.113)$$

The translation eigenvalues  $c_{\mathcal{R}}$  must be a representation of the group of lattice translations  $T$  which forms an infinite cyclic group and has representations of the form  $c_{\mathcal{R}} = \exp(i\mathbf{k}\cdot\mathcal{R})$ . The wavefunctions diagonalizing the lattice Hamiltonian can be chosen to satisfy

$$\psi(\mathbf{r} + \mathcal{R}) = \exp(i\mathbf{k}\cdot\mathcal{R})\psi(\mathbf{r}) \quad (1.114)$$

and Eq. (1.114) is called the Bloch Theorem. Alternatively, the wavefunctions satisfying Eq. (1.114) can be written under the form of so-called Bloch functions

$$\psi(\mathbf{r}) = \exp(i\mathbf{k}\cdot\mathbf{r})u_{n\mathbf{k}}(\mathbf{r}) = \psi_{n\mathbf{k}}(\mathbf{r}) \quad (1.115)$$

where  $u_{n\mathbf{k}}(\mathbf{r})$  is a function with lattice periodicity such that  $\psi(\mathbf{r})$  satisfies Eq. (1.114).

A Bloch function  $\psi_{n\mathbf{k}}(\mathbf{r})$  which is an eigenfunction of the Hamiltonian with an energy  $E_{n\mathbf{k}}$  is characterized by two quantum numbers: the wave vector ( $\mathbf{k}$ ) and the band index ( $n$ ). The wave vector  $\mathbf{k}$  is a vector in the reciprocal lattice formed by the primitive reciprocal vectors

$$(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = \frac{2\pi}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} (\mathbf{a}_2 \times \mathbf{a}_3, \mathbf{a}_3 \times \mathbf{a}_1, \mathbf{a}_1 \times \mathbf{a}_2) \quad (1.116)$$

and a reciprocal lattice vector is denoted by  $\mathcal{G} = u_1\mathbf{b}_1 + u_2\mathbf{b}_2 + u_3\mathbf{b}_3$ . In order to have a unique wavefunction associated with each set of quantum numbers  $n\mathbf{k}$ , the wavevector ( $\mathbf{k}$ ) has to be restricted to the first Brillouin zone and a index  $n$  is added to each reciprocal lattice vector  $\mathcal{G}_n$ . The first Brillouin zone is defined as the set of points in reciprocal lattice space which are closer to the reciprocal lattice vector at the origin ( $\mathcal{G}_0 = \mathbf{0}$ ) than to any other reciprocal lattice vector.

### Calculating the band structure: methods

To calculate the bandstructure in a semiconductor different methods can be used.

#### 1. The pseudopotential method

The crystal potential energy is expanded in Fourier components

$$U_{\text{lat}}(\mathbf{r}) = \sum_m \tilde{U}_{\text{lat}}(\mathbf{g}_m) \exp(i\mathbf{g}_m \cdot \mathbf{r}). \quad (1.117)$$

where the tilde indicates functions defined in reciprocal space.

In practical implementations of the pseudopotential method, only the smaller Fourier components are retained and the pseudopotential does not capture the rapid changes of the real lattice potential. This is where the method derives its name from as it is not the real lattice potential but a pseudopotential. The number of Fourier components that need to be known to describe the pseudopotential can be further reduced based on the rotational symmetries of the lattice.

To solve the Hamiltonian, rewrite the periodic part of the Bloch function  $u_{n\mathbf{k}}(\mathbf{r})$  as a Fourier series

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r}) \sum_m \tilde{u}_{n\mathbf{k}m} \exp(i\mathbf{g}_m \cdot \mathbf{r}) \quad (1.118)$$

where  $\tilde{u}_{n\mathbf{k}m}$  are the coefficients of the Fourier series. With the Bloch functions and the pseudopotential given in reciprocal space, Eq. (1.111) can be rewritten as

$$\frac{\hbar^2 |\mathbf{k} + \mathbf{g}_m|^2}{2m_0} u_{n\mathbf{k}m} + \sum_{m'} \tilde{U}_{\text{lat}}(\mathbf{g}_m - \mathbf{g}_{m'}) u_{n\mathbf{k}m'} = E_{n\mathbf{k}} u_{n\mathbf{k}m}. \quad (1.119)$$

The values of the parameters for the pseudopotential can be determined empirically by calibrating the resulting band structure to experimental data, this method is called the local empirical pseudopotential method. An alternative is to determine the pseudopotentials by determining the lattice pseudopotential and the charge density in the semiconductor in a self-consistent way on which we will not elaborate further here. For silicon, a good description of the bandstructure can already be obtained using empirical pseudopotentials with 3 parameters. Eq. (1.119) represents an eigenvalue problem which can be solved numerically and the resulting silicon band structure is shown in Fig. 1.20.

## 2. The $\mathbf{k} \cdot \mathbf{p}$ method

According to the  $\mathbf{k} \cdot \mathbf{p}$  method, Eq. (1.111) is expanded in wavefunctions of Bloch form based on a given set of periodic functions  $u_{m0}(\mathbf{r})$ :

$$\psi(\mathbf{r}) = \sum_m a_m \phi_{m\mathbf{k}}(\mathbf{r}) = \sum_m a_m \exp(i\mathbf{k} \cdot \mathbf{r}) u_{m0}(\mathbf{r}). \quad (1.120)$$

Substituting Eq. (1.120) into Eq. (1.111) results in the  $\mathbf{k} \cdot \mathbf{p}$  Hamiltonian

$$\frac{\hbar^2 |\mathbf{k}|^2}{2m_0} a_m + \sum_{m'=0}^M \frac{\hbar}{m_0} \mathbf{k} \cdot \mathbf{p}_{mm'} a_{m'} + E_m^0 a_m = E a_m \quad (1.121)$$

with the momentum matrix elements

$$\mathbf{p}_{mm'} = -i\hbar \int_{\Omega} d^3r u_{m'0}^*(\mathbf{r}) \nabla u_{m0}(\mathbf{r}) \quad (1.122)$$

and the basis functions at  $\mathbf{k} = 0$  are taken to be eigenfunctions of the Hamiltonian

$$\left[ -\frac{\hbar^2 \nabla^2}{2m_0} + U_{\text{lat}}(\mathbf{r}) \right] u_{m0}(\mathbf{r}) = E_m^0 u_{m0}(\mathbf{r}). \quad (1.123)$$

For numerical computation, a finite number of bands  $M$  has to be selected. The description of the bandstructure will be better if more bands are included and remote bands will contribute less than nearby bands. Selecting 15 bands based on the free electron band structure and making use of symmetry rules to reduce the number of unknown matrix elements, the band structure can be determined by fitting to experimental data [58].

For computational purposes the reduction of the size of the basis is often desired but the resulting inaccuracy of the bandstructure is unwanted. To have improved accuracy with a smaller basis set, the influence of the remote bands can be taken into account in an approximate way. This can be done using Lödwin's perturbation theory which takes the effect of the remote bands into account using first order perturbation theory. In our thesis we have not used this approach apart from the effective mass approximation which reduces the basis to a single band. In the effective mass approximation, the bandstructure is approximated by

$$E_{n\mathbf{k}} = E_n^0 \pm \frac{\hbar^2 |\mathbf{k}|^2}{2m_n^*} \quad (1.124)$$

where  $E_n^0$  is the valence band maximum or the conduction band minimum and the band has positive curvature in the case of the conduction band and a negative curvature in the case of the valence band.

### 3. Tight binding approximation

According to the tight binding approximation, the Hamiltonian is expanded in a basis which is localized at every lattice site. Next, the resulting Hamiltonian is truncated such that only a number of nearest neighbor terms are retained. When extending this method beyond bulk structures, a mesh point is required at every atom in the structure introducing a heavy computational burden. For this reason we have not used the tight binding method in this thesis.

### 4. Density functional theory

Density functional theory (DFT) software is commercially available and widely used to calculate band structures of chemical compounds but due to their inability to accurately calculate band gap values, DFT is rarely used to calculate electronic transport in semiconductors.

## 1.B Appendix: many-particle systems

### 1.B.1 Describing a quantum system

An  $N$ -particle quantum mechanical system can be described by its wavefunction  $\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t)$  and the probability to find the first particle at  $\mathbf{r}_1$ , the second at  $\mathbf{r}_2$ , etc. . . is given by the square of the modulus of the wavefunction  $|\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t)|^2$ .

When the wavefunction describes indistinguishable particles such as electrons, it is no longer possible to discriminate between the first and the second electron. This means that the probability of finding the first and second particle at  $\mathbf{r}_1$  and  $\mathbf{r}_2$  respectively has to equal the probability of finding the first and second at  $\mathbf{r}_2$  and  $\mathbf{r}_1$  respectively. In other words,  $|\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t)|^2 = |\psi(\mathbf{r}_2, \mathbf{r}_1, \dots, \mathbf{r}_N, t)|^2$ .

A wavefunction which describes indistinguishable particles therefore has two possibilities, either the wavefunction is symmetric with respect to permutation of the particle coordinate or the wavefunction is anti-symmetric or formally:

$$\begin{aligned} \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) &= \psi(\mathbf{r}_2, \mathbf{r}_1, \dots, \mathbf{r}_N, t) && \text{bosons} \\ \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) &= -\psi(\mathbf{r}_2, \mathbf{r}_1, \dots, \mathbf{r}_N, t) && \text{fermions.} \end{aligned} \quad (1.125)$$

All indistinguishable particles can be categorized based on whether their wavefunction is symmetric or antisymmetric under permutation of the coordinates and particles are called bosons when the wavefunction is symmetric and fermions when the wavefunction is antisymmetric.

When studying electronic transport, the particles under study are electrons and electrons are fermions. A well-known property of fermions is that they must satisfy the Pauli exclusion principle which states that no two-electrons can occupy the same state. The Pauli exclusion principle follows immediately from the antisymmetric property of the wavefunction, i.e.  $\psi(\mathbf{r}_1, \mathbf{r}_1, \dots, \mathbf{r}_N, t) = 0$ .

Just as in the case of a single particle, the time-evolution of a many-particle wavefunction is described by a Hamiltonian. For example, for a set of non-interacting identical particles in a potential profile  $U(\mathbf{r})$ , the many-particle Schrödinger equation reads:

$$\begin{aligned} i\hbar \frac{\partial \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t)}{\partial t} = \\ \left[ -\frac{\nabla_1^2}{2m_0} - \frac{\nabla_2^2}{2m_0} - \dots - \frac{\nabla_N^2}{2m_0} + U(\mathbf{r}_1) + U(\mathbf{r}_2) + \dots + U(\mathbf{r}_N) \right] \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) \end{aligned} \quad (1.126)$$

where  $\nabla_i$  is the nabla operator with respect to the  $i$ -th coordinate  $\mathbf{r}_i$ .

## Second Quantization

The many-particle wavefunction is relatively easy to understand but it has a number of important disadvantages. First of all, the wavefunction does not take the properties of the indistinguishable particles into account and the symmetric or antisymmetric properties have to be imposed explicitly. And in conjunction with that, operators such as the Hamiltonian need a term for each particle in the system. On the practical side, it is impossible to solve the many-body wavefunction exactly for a large number of particles as the problem scales exponentially with the number of particles in the system. And there is no hope that we will one day be able to solve the many-body wavefunction exactly as the many-particle wavefunction is so rich in information that even storing the wavefunction in a physical memory will never be possible.

For this reason, many-body states are generally described in second quantization [59, 60]. In second quantization, the states are described using an abstract state vector,  $|\psi(t)\rangle$  in our example. Given that an explicit wavefunction in general can never be calculated, representing a state by an abstract state vector does not present a real disadvantage with respect to an ordinary many-body wavefunction. On the other hand, the second quantized wavefunction has the advantage that the symmetric or anti-symmetric nature of the particles under study is taken into account automatically based on a set of commutation rules for second quantized operators. We denote all second quantized operators with a hat ( $\hat{A}$ ).

The set of state vectors corresponding to all 0-particle, 1-particle, ... infinite-particle states spans the total space relevant for state vectors, this space is also known as Fock Space. A general state vector in Fock Space is however not limited to states with a definite number of particles, a linear combination of two state vectors results in a new state vector in Fock Space.

To make the concept of second quantization more concrete, we introduce the second quantized state vector corresponding to our example of an  $N$ -particle wavefunction [52]:

$$|\psi(t)\rangle = \frac{1}{\sqrt{N!}} \int d^3r_1 d^3r_2 \cdots d^3r_N \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) \hat{\psi}^\dagger(\mathbf{r}_1) \hat{\psi}^\dagger(\mathbf{r}_2) \cdots \hat{\psi}^\dagger(\mathbf{r}_N) |0\rangle. \quad (1.127)$$

$|0\rangle$  denotes the vacuum state or 0-particle state.  $\hat{\psi}^\dagger(\mathbf{r})$  is called a field operator and creates a particle at position  $\mathbf{r}$ .

From the antisymmetric nature of the fermion wavefunction, it follows that the field operators have the property

$$\{\hat{\psi}^\dagger(\mathbf{r}), \hat{\psi}^\dagger(\mathbf{r}')\} = \hat{\psi}^\dagger(\mathbf{r})\hat{\psi}^\dagger(\mathbf{r}') + \hat{\psi}^\dagger(\mathbf{r}')\hat{\psi}^\dagger(\mathbf{r}) = 0 \quad (1.128)$$

where  $\{X, Y\} = XY + YX$  is the anticommutator. The commutation relations automatically account for the fermionic character of the wavefunction.

Apart from creation operators, there are also annihilation operators  $\hat{\psi}(\mathbf{r})$  which are the Hermitian conjugate of the creation operators

$$\left(\hat{\psi}^\dagger(\mathbf{r})\right)^\dagger = \hat{\psi}(\mathbf{r}) \quad (1.129)$$

The annihilation operators also satisfy the anticommutation rule

$$\{\hat{\psi}(\mathbf{r}), \hat{\psi}(\mathbf{r}')\} = 0 \quad (1.130)$$

and the creation and annihilation operator satisfy the anticommutation rule

$$\{\hat{\psi}(\mathbf{r}), \hat{\psi}^\dagger(\mathbf{r}')\} = \delta(\mathbf{r} - \mathbf{r}'). \quad (1.131)$$

Apart from the states, also operators can be written in second quantization often making the many-body problem more tractable. Expectation values of an observable are obtained by calculating the expectation value with the corresponding second quantized operator

$$\langle A(t) \rangle = \langle \psi(t) | \hat{A} | \psi(t) \rangle. \quad (1.132)$$

Some popular operators are, the local electron density operator

$$\hat{n}(\mathbf{r}) = \hat{\psi}^\dagger(\mathbf{r})\hat{\psi}(\mathbf{r}), \quad (1.133)$$

and the number operator, which is just the integral over the local electron density operator over the complete volume

$$\hat{N} = \int d^3r \hat{\psi}^\dagger(\mathbf{r})\hat{\psi}(\mathbf{r}). \quad (1.134)$$

The operator associated with a scalar potential  $U(\mathbf{r})$  is given by

$$\hat{U} = \int d^3r U(\mathbf{r})\hat{\psi}^\dagger(\mathbf{r})\hat{\psi}(\mathbf{r}), \quad (1.135)$$

while a potential describing two-particle interaction  $V(\mathbf{r}_1, \mathbf{r}_2)$  is given by

$$\hat{V} = \frac{1}{2} \int d^3r_1 d^3r_2 V(\mathbf{r}_1, \mathbf{r}_2) \hat{\psi}^\dagger(\mathbf{r}_1) \hat{\psi}^\dagger(\mathbf{r}_2) \hat{\psi}(\mathbf{r}_2) \hat{\psi}(\mathbf{r}_1). \quad (1.136)$$

In second quantization, the Hamiltonian from Eq. (1.126) reads

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = \hat{H} |\psi(t)\rangle = [\hat{T} + \hat{U}] |\psi(t)\rangle \quad (1.137)$$

with  $\hat{T}$  the second quantized kinetic energy operator.

In this section, we have introduced second quantization based on field operators ( $\hat{\psi}(\mathbf{r})$ ) which take a real space coordinate as an argument. The choice for the

real space coordinate is however arbitrary and it is equally permissible to use a different coordinate set which spans the entire Hilbert space. Any linear combination of field operators also results in a second quantized operator. As a result, any unitary transform of the space under study presents an equivalent set of second quantized operators.

If  $\phi_{\mathbf{k}}(\mathbf{r})$  is a complete orthonormal set of wavefunctions

$$\int d^3r \phi_{\mathbf{k}}(\mathbf{r})\phi_{\mathbf{k}'}(\mathbf{r}) = \delta(\mathbf{k} - \mathbf{k}'), \quad (1.138)$$

$$\int d^3k \phi_{\mathbf{k}}(\mathbf{r})\phi_{\mathbf{k}}(\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \quad (1.139)$$

a new set of second quantized operators is given by

$$\hat{c}_{\mathbf{k}} = \int d^3r \phi_{\mathbf{k}}(\mathbf{r})\hat{\psi}_{\mathbf{k}}(\mathbf{r}). \quad (1.140)$$

If the set of wavefunctions  $\phi_{\mathbf{k}}(\mathbf{r})$  diagonalize the first quantized Hamiltonian of the system with an eigenenergy  $E_{\mathbf{k}}$

$$H\phi_{\mathbf{k}}(\mathbf{r}) = E_{\mathbf{k}}\phi_{\mathbf{k}}(\mathbf{r}) \quad (1.141)$$

the second quantized Hamiltonian becomes diagonal and is given by

$$\hat{H} = \int d^3k E_{\mathbf{k}}\hat{c}_{\mathbf{k}}^\dagger\hat{c}_{\mathbf{k}}. \quad (1.142)$$

## 1.B.2 Statistical Mechanics

What we have described up to now are so-called pure states and pure states are the perfect quantum mechanical description of a system. In practice however, it is not always possible to perfectly determine or control the state of a system. Interactions with the outside world can not be eliminated completely. For this reason, the concept of an ensemble is introduced, different pure states can occur with a given probability. To deal with the mathematics of ensembles and calculate the expectation value of operators, the density matrix is introduced

$$\hat{\rho} = \sum_i p_i |i\rangle\langle i| \quad (1.143)$$

where  $i$  is a label for the pure state and the sum runs over all pure states in the system and  $p_i$  is the probability for a system to be in that pure state.

The definition of the density matrix is chosen such that the expectation value of an observable with operator  $\hat{A}$  on a system described by the density matrix  $\hat{\rho}$  is given by

$$\langle A \rangle = \sum_i p_i \langle i | \hat{A} | i \rangle = \sum_{i,j} p_i \langle i | j \rangle \langle j | \hat{A} | i \rangle \quad (1.144)$$

$$= \sum_j \langle j | \sum_i \hat{A} | i \rangle p_i \langle i | j \rangle \quad (1.145)$$

$$= \text{Tr}(\hat{A}\hat{\rho}) \quad (1.146)$$

and the time dependence of the density matrix is determined by the quantum Liouville equation

$$i\hbar \frac{\partial}{\partial t} \hat{\rho} = [\hat{H}, \hat{\rho}] \quad (1.147)$$

with  $[X, Y] = XY - YX$  the commutator.

### The Grand Canonical Ensemble

In this section we determine the density matrix describing a general system which has a known expectation value of energy and number of particles. The system determined in this way is called a grand canonical ensemble.

Without any knowledge about a system, there is no preference with regard to the occupation of one state compared to another and each state is assumed to be occupied by a particle with equal probability:  $p_i = p_j$ . On the other hand, an isolated system is subject to the conservation of energy and we can determine the energy of the total system. This presents us with a constraint on the states which can be occupied. The absence of preference of one state over another is measured by the entropy. Describing a system based on a limited knowledge of the system is done by maximizing the entropy while satisfying the constraints imposed on the system.

Formally, the Gibbs entropy of a system can be obtained from its density matrix by the relation [61]

$$S = -k_B \text{Tr}(\hat{\rho} \log(\hat{\rho})). \quad (1.148)$$

The density matrix always has the property

$$\text{Tr}(\hat{\rho}) = 1 \quad (1.149)$$

and furthermore, the grand canonical ensemble imposes an expectation value on the energy of the system

$$\text{Tr}(\hat{H}\hat{\rho}) = \langle E \rangle \quad (1.150)$$

and an expectation value on the number of particles in the system

$$\text{Tr}(\hat{N}\hat{\rho}) = \langle N \rangle. \quad (1.151)$$

Describing the grand canonical density matrix now amounts to maximizing a problem with constraints, this can be done by introducing Lagrange multipliers for the different constraints. The density matrix which maximizes the

entropy has to satisfy

$$\delta \text{Tr} \left( \alpha_0 \hat{\rho} + \alpha_E \hat{H} \hat{\rho} + \alpha_N \hat{N} \hat{\rho} - k_B \hat{\rho} \log(\hat{\rho}) \right) = 0, \quad (1.152)$$

$$\text{Tr} \left( \left( \alpha_0 \hat{I} + \alpha_E \hat{H} + \alpha_N \hat{N} - k_B \log(\hat{\rho}) - k_B \hat{I} \right) \delta \hat{\rho} \right) = 0 \quad (1.153)$$

for any  $\delta \hat{\rho}$ , which requires

$$\alpha_0 \hat{I} + \alpha_E \hat{H} + \alpha_N \hat{N} - k_B \log(\hat{\rho}) - k_B \hat{I} = 0 \quad (1.154)$$

which results in the Grand Canonical Ensemble in its usual form

$$\hat{\rho} = \exp \left( -\beta(\hat{H} - \mu \hat{N}) \right) / \mathcal{Z} \quad (1.155)$$

where the Langrange multipliers have been relabeled:  $\beta = \alpha_E/k_B$ ,  $\mu = \alpha_N/\alpha_E$  and the remaining Lagrange multiplier  $\alpha_0$  ensures that  $\text{Tr}(\hat{\rho}) = 1$  which can be accomplished by introducing the partition function

$$\mathcal{Z} = \text{Tr} \left( \exp \left( -\beta(\hat{H} - \mu \hat{N}) \right) \right). \quad (1.156)$$

If the Hamiltonian is diagonalized by the operators  $\hat{c}_{\mathbf{k}}$ , the Grand Canonical Ensemble is given by

$$\hat{\rho} = \int d^3k \exp \left( -\beta(E_{\mathbf{k}} - \mu) \hat{c}_{\mathbf{k}}^\dagger \hat{c}_{\mathbf{k}} \right) / \mathcal{Z}. \quad (1.157)$$

The probability that a state  $\mathbf{k}$  is occupied is calculated from

$$f(\mathbf{k}) = \text{Tr}(\hat{c}_{\mathbf{k}}^\dagger \hat{c}_{\mathbf{k}} \hat{\rho}) \quad (1.158)$$

and since  $\hat{c}_{\mathbf{k}} \hat{\rho} = \exp(-\beta(E_{\mathbf{k}} - \mu)) \hat{\rho} \hat{c}_{\mathbf{k}}$ ,

$$f(\mathbf{k}) = e^{-\beta(E_{\mathbf{k}} - \mu)} \frac{\text{Tr}(\hat{c}_{\mathbf{k}}^\dagger \hat{\rho} \hat{c}_{\mathbf{k}})}{\mathcal{Z}}. \quad (1.159)$$

Using the anti-commutation property of fermions and the invariance of the trace under cyclic permutation, the Fermi-Dirac distribution is obtained

$$f(\mathbf{k}) = e^{-\beta(E_{\mathbf{k}} - \mu)} (1 - f(\mathbf{k})) = \frac{1}{e^{\beta(E_{\mathbf{k}} - \mu)} + 1}. \quad (1.160)$$

### Heisenberg and interaction picture

Up to now, we have described quantum mechanics in the Schrödinger picture where the operators do not have time dependence while the states have time dependence. When investigating the reaction of a system to a perturbation, it

is often more convenient to use time-dependent operators than time-dependent state vectors. To get time-dependent operators, the Schrödinger picture has to be abandoned and the Heisenberg picture or the interaction picture has to be introduced.

The Heisenberg picture of quantum mechanics reverses the Schrödinger picture. The states are stationary and the operators gain time dependence. States are transformed

$$|\hat{\psi}_H\rangle = e^{i\hat{H}t/\hbar}|\hat{\psi}_S(t)\rangle \quad (1.161)$$

so that they lose their time dependence while operators are transformed according to

$$\hat{A}_H(t) = e^{i\hat{H}t/\hbar}\hat{A}_S e^{-i\hat{H}t/\hbar}. \quad (1.162)$$

The expectation values are the same in both pictures

$$\langle\hat{\psi}_H|\hat{A}_H(t)|\hat{\psi}_H\rangle = \langle\hat{\psi}_S(t)|\hat{A}_S|\hat{\psi}_S(t)\rangle \quad (1.163)$$

so the Heisenberg and the Schrödinger picture are equivalent.

An alternative picture in which both the state vectors and the operators gain time dependence is the interaction picture. Associated with an interaction Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{H}_{\text{int}}, \quad (1.164)$$

the interaction picture is defined in such a way that the state vectors have no time dependence in the absence of the interaction Hamiltonian  $\hat{H}_{\text{int}}$

$$|\hat{\psi}(t)\rangle_I = e^{i\hat{H}_0 t/\hbar}|\hat{\psi}(t)\rangle_S \quad (1.165)$$

while the operators are

$$\hat{A}_I(t) = e^{i\hat{H}_0 t/\hbar}\hat{A}_S e^{-i\hat{H}_0 t/\hbar}. \quad (1.166)$$

The time evolution of a state vector is governed by the interaction Hamiltonian

$$i\hbar\frac{\partial}{\partial t}|\hat{\psi}(t)\rangle_I = \hat{H}_{\text{int},I}(t)|\hat{\psi}(t)\rangle_I. \quad (1.167)$$

## Bibliography

- [1] J. E. Lilienfeld, "Device for controlling electric current," U.S. Patent 1 900 018, Mar 7, 1933.
- [2] D. Kahng and M. M. Atalla, "Silicon-silicon dioxide field induced surface devices," in *IRE Solid State Device Research Conference, Pittsburgh, PA*, 1960.

- [3] “The McClean report,” IC Insights, Tech. Rep., 2011.
- [4] *International Technology Roadmap for Semiconductors - Process Integration, Devices, and structures*, 2010.
- [5] R. H. Dennard, F. H. Gaensslen, V. L. Rideout, E. Bassous, and A. R. LeBlanc, “Design of ion-implanted mosfet’s with very small physical dimensions,” *IEEE Journal of Solid-State Circuits*, vol. 9, no. 5, pp. 256–268, 1974.
- [6] P. Heremans, K. Maex, R. Mertens, and R. Van Overstraeten, *Semiconductor Devices*. Cursusdienst VTK vzw, 2002.
- [7] *Sentaurus Device*, Synopsis, March 2010.
- [8] *ATLAS User Manual*, Silvaco International, Santa Clara, CA.
- [9] K. Mistry, C. Allen, C. Auth, B. Beattie, D. Bergstrom, M. Bost, M. Brazier, M. Buehler, A. Cappellani, R. Chau, C.-H. Choi, G. Ding, K. Fischer, T. Ghani, R. Grover, W. Han, D. Hanken, M. Hattendorf, J. He, J. Hicks, R. Huessner, D. Ingerly, P. Jain, R. James, L. Jong, S. Joshi, C. Kenyon, K. Kuhn, K. Lee, H. Liu, J. Maiz, B. McIntyre, P. Moon, J. Neiryneck, S. Pae, C. Parker, D. Parsons, C. Prasad, L. Pipes, M. Prince, P. Ranade, T. Reynolds, J. Sandford, L. Shifren, J. Sebastian, J. Seiple, D. Simon, S. Sivakumar, P. Smith, C. Thomas, T. Troeger, P. Vandervoorn, S. Williams, and K. Zawadzki, “A 45nm logic technology with high-k+metal gate transistors, strained silicon, 9 cu interconnect layers, 193nm dry patterning, and 100% pb-free packaging,” in *Electron Devices Meeting, 2007. IEDM 2007. IEEE International*, dec. 2007, pp. 247–250.
- [10] H. Kawaura, T. Sakamoto, and T. Baba, “Observation of source-to-drain direct tunneling current in 8 nm gate electrically variable shallow junction metal-oxide-semiconductor field-effect transistors,” *Applied Physics Letters*, vol. 76, no. 25, pp. 3810–3812, jun 2000.
- [11] T. Y. Chan, J. Chen, P. K. Ko, and C. Hu, “The impact of gate-induced drain leakage current on mosfet scaling,” in *Proc. Int. Electron Devices Meeting*, vol. 33, 1987, pp. 718–721.
- [12] S. Banerjee, W. Richardson, J. Coleman, and A. Chatterjee, “A new three-terminal tunnel device,” *Electron Device Letters, IEEE*, vol. 8, no. 8, pp. 347–349, Aug. 1987.
- [13] E. Takeda, H. Matsuoka, Y. Igura, and S. Asai, “A band to band tunneling mos device (b/sup 2/t-mosfet) - a kind of ‘si quantum device’,” in *Proc. Int. Electron Devices Meeting IEDM ’88. Technical Digest*, 1988, pp. 402–405.

- [14] T. Baba, "Proposal for surface tunnel transistors," *Japanese Journal of Applied Physics*, vol. 31, no. Part 2, No. 4B, pp. L455–L457, 1992. [Online]. Available: <http://jjap.jsap.jp/link?JJAP/31/L455/>
- [15] W. M. Reddick and G. A. J. Amaratunga, "Silicon surface tunnel transistor," *Applied Physics Letters*, vol. 67, no. 4, pp. 494–496, July 1995.
- [16] J. Koga and A. Toriumi, "Negative differential conductance at room temperature in three-terminal silicon surface junction tunneling device," *Applied Physics Letters*, vol. 70, no. 16, pp. 2138–2140, 1997.
- [17] C. Zener, "A theory of the electrical breakdown of solid dielectrics," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 145, no. 855, pp. 523–529, July 1934.
- [18] S. M. Sze, *Physics of Semiconductor Devices*. Wiley, 1981.
- [19] E. O. Kane, "Zener tunneling in semiconductors," *Journal of Physics and Chemistry of Solids*, vol. 12, pp. 181–188, 1959.
- [20] —, "Theory of tunneling," *Journal of Applied Physics*, vol. 32, no. 1, pp. 83–91, 1961.
- [21] L. Keldysh, "Behavior of non-metallic crystals in strong electric fields," *Sov. Phys. JETP*, vol. 6, p. 763, 1958.
- [22] —, "Influence of the lattice vibrations of a crystal on the production of electron-hole pairs in a strong electric field," *Sov. Phys. JETP*, vol. 7, p. 665, 1959.
- [23] A. Schenk, "Rigorous theory and simplified model of the band-to-band tunneling in silicon," *Solid-State Electronics*, vol. 36, no. 1, pp. 19–34, 1993.
- [24] A. Heigl, A. Schenk, and G. Wachutka, "Correction to the schenk model of band-to-band tunneling in silicon applied to the simulation of nanowire tunneling transistors," in *Computational Electronics, 2009. IWCE '09. 13th International Workshop on DOI - 10.1109/IWCE.2009.5091099*, 2009, pp. 1–4. [Online]. Available: [10.1109/IWCE.2009.5091099](https://doi.org/10.1109/IWCE.2009.5091099)
- [25] G. Hurkx, D. Klaassen, and M. Knuvers, "A new recombination model for device simulation including tunneling," *Electron Devices, IEEE Transactions on*, vol. 39, no. 2, pp. 331–338, feb 1992.
- [26] A. S. Verhulst, W. G. Vandenberghe, K. Maex, and G. Groeseneken, "A tunnel field-effect transistor without gate-drain overlap," *Applied Physics Letters*, vol. 91, no. 053102, July 2007.

- [27] J. Appenzeller, Y.-M. Lin, J. Knoch, Z. Chen, and P. Avouris, "Comparing carbon nanotube transistors - the ideal choice: a novel tunneling device design," *Electron Devices, IEEE Transactions on*, vol. 52, no. 12, pp. 2568 – 2576, dec. 2005.
- [28] K. De Meyer, M. Caymax, N. Collaert, R. Loo, and P. Verheyen, "The vertical heterojunction MOSFET," *Thin Solid Films*, vol. 336, no. 1 - 2, pp. 299 – 305, 1998. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S004060909801284X>
- [29] D. Hisamoto, W.-C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T.-J. King, J. Bokor, and C. Hu, "Finfet-a self-aligned double-gate mosfet scalable to 20 nm," *Electron Devices, IEEE Transactions on*, vol. 47, no. 12, pp. 2320 – 2325, dec 2000.
- [30] A. S. Verhulst, B. Sorée, D. Leonelli, W. G. Vandenberghe, and G. Groeseneken, "Modeling the single-gate, double-gate, and gate-all-around tunnel field-effect transistor," *Journal of Applied Physics*, vol. 107, no. 2, p. 024518, 2010. [Online]. Available: <http://link.aip.org/link/?JAP/107/024518/1>
- [31] K. K. Bhuiwarka, S. Sedlmaier, A. K. Ludsteck, C. Tolksdorf, J. Schulze, and I. Eisele, "Vertical tunnel field-effect transistor," *IEEE Transactions on Electron Devices*, vol. 51, no. 2, pp. 279–282, 2004.
- [32] S. Iijima *et al.*, "Helical microtubules of graphitic carbon," *nature*, vol. 354, no. 6348, pp. 56–58, 1991.
- [33] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov, "Two-Dimensional Gas of Massless Dirac Fermions in Graphene," *Nature*, vol. 438, pp. 197–200, Sep. 2005. [Online]. Available: <http://dx.doi.org/10.1038/nature04233>
- [34] S. O. Koswatta, D. E. Nikonov, and M. S. Lundstrom, "Computational study of carbon nanotube p-i-n tunnel fets," in *Proc. IEDM Technical Digest Electron Devices Meeting IEEE Int*, 2005, pp. 518–521.
- [35] J. Appenzeller, Y.-M. Lin, J. Knoch, and P. Avouris, "Band-to-band tunneling in carbon nanotube field-effect transistors," *Phys. Rev. Lett.*, vol. 93, p. 196805, Nov 2004. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevLett.93.196805>
- [36] P.-F. Wang, "Complementary tunneling-fets (CTFET) in CMOS technology," Ph.D. dissertation, Technische Universität München, 2003.
- [37] W. G. Vandenberghe, A. S. Verhulst, G. Groeseneken, B. Sorée, and W. Magnus, "Analytical model for a tunnel field-effect transistor," in *IEEE Melecon*, 2008.

- [38] C. Hu, D. Chou, P. Patel, and A. Bowonder, "Green transistor - a vdd scaling path for future low power ics," in *VLSI Technology, Systems and Applications, 2008. VLSI-TSA 2008. International Symposium on*, april 2008, pp. 14–15.
- [39] D. Leonelli, A. Vandooren, R. Rooyackers, A. Verhulst, C. Huyghebaert, S. De Gendt, M. Heyns, and G. Groeseneken, "Novel architecture to boost the vertical tunneling in tunnel field effect transistors," in *SOI Conference (SOI), 2011 IEEE International*, oct. 2011, pp. 1–2.
- [40] O. Nayfeh, C. Chleirigh, J. Hennessy, L. Gomez, J. Hoyt, and D. Antoniadis, "Design of tunneling field-effect transistors using strained-silicon/strained-germanium type-ii staggered heterojunctions," *Electron Device Letters, IEEE*, vol. 29, no. 9, pp. 1074–1077, sept. 2008.
- [41] A. Verhulst, W. Vandenberghe, K. Maex, S. De Gendt, M. Heyns, and G. Groeseneken, "Complementary silicon-based heterostructure tunnel-fets with high tunnel rates," *Electron Device Letters, IEEE*, vol. 29, no. 12, pp. 1398–1401, dec. 2008.
- [42] K. Gopalakrishnan, P. B. Griffin, and J. D. Plummer, "Impact ionization mos (i-mos)-part i: device and circuit simulations," *IEEE Transactions on Electron Devices*, vol. 52, no. 1, pp. 69–76, 2005.
- [43] K. Gopalakrishnan, R. Woo, C. Jungemann, P. B. Griffin, and J. D. Plummer, "Impact ionization mos (i-mos)-part ii: experimental results," *IEEE Transactions on Electron Devices*, vol. 52, no. 1, pp. 77–84, 2005.
- [44] S. Salahuddin and S. Datta, "Use of negative capacitance to provide voltage amplification for low power nanoscale devices," *Nano Letters*, vol. 8, no. 2, pp. 405–410, 2008, pMID: 18052402. [Online]. Available: <http://pubs.acs.org/doi/abs/10.1021/nl071804g>
- [45] A. I. Khan, C. W. Yeung, C. Hu, and S. Salahuddin, "Ferroelectric negative capacitance mosfet: Capacitance tuning & antiferroelectric operation," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, 2011.
- [46] E. Gnani, P. Maiorano, S. Reggiani, A. Gnudi, and G. Baccarani, "Performance limits of superlattice-based steep-slope nanowire fets," in *Electron Devices Meeting (IEDM), 2011 IEEE International*, dec. 2011, pp. 5.1.1–5.1.4.
- [47] L. E. Ballentine, *Quantum Mechanics: a Modern Development*. World Scientific Publishing, 1998.
- [48] J. R. Chelikowsky and M. L. Cohen, "Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors," *Phys. Rev. B*, vol. 14, pp. 556–582, Jul 1976. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevB.14.556>

- [49] M. G. Burt, “The justification for applying the effective-mass approximation to microstructures,” *Journal of Physics: Condensed Matter*, vol. 4, no. 32, p. 6651, 1992. [Online]. Available: <http://stacks.iop.org/0953-8984/4/i=32/a=003>
- [50] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, “Single and multi-band modeling of quantum electron transport through layered semiconductor devices,” *Journal of Applied Physics*, vol. 81, no. 12, pp. 7845–7869, 1997.
- [51] R. Lake, “Application of the keldysh formalism to quantum device modeling and analysis,” Ph.D. dissertation, Purdue University, 1992.
- [52] W. Magnus and W. Schoenmaker, “On the use of a new integral theorem for the quantum mechanical treatment of electric circuits,” *Journal of Mathematical Physics*, vol. 39, no. 12, pp. 6715–6719, 1998. [Online]. Available: <http://link.aip.org/link/?JMP/39/6715/1>
- [53] G. Fiori and G. Iannaccone, “Ultralow-voltage bilayer graphene tunnel fet,” *Electron Device Letters, IEEE*, vol. 30, no. 10, pp. 1096–1098, oct. 2009.
- [54] M. Luisier and G. Klimeck, “Atomistic full-band design study of inas band-to-band tunneling field-effect transistors,” *Electron Device Letters, IEEE*, vol. 30, no. 6, pp. 602–604, june 2009.
- [55] S. Koswatta, M. Lundstrom, and D. Nikonov, “Performance comparison between p-i-n tunneling transistors and conventional mosfets,” *Electron Devices, IEEE Transactions on*, vol. 56, no. 3, pp. 456–465, march 2009.
- [56] M. Van de Put, “Modelling of band-to-band tunneling in heterostructure semiconductors,” Master’s thesis, Katholieke Universiteit Leuven, 2012.
- [57] D. Jena, T. Fang, Q. Zhang, and H. Xing, “Zener tunneling in semiconducting nanotube and graphene nanoribbon p-n junctions,” *Applied Physics Letters*, vol. 93, no. 11, p. 112106, 2008.
- [58] M. Cardona and F. H. Pollak, “Energy-Band Structure of Germanium and Silicon: The k-p Method,” *Phys. Rev.*, vol. 142, pp. 530–543, 1966.
- [59] A. L. Fetter and J. D. Walecka, *Quantum theory of many-particle systems*. McGraw-Hill Book Company, 1971.
- [60] G. D. Mahan, *Many-particle physics*, 2nd ed. Plenum Press, 1990.
- [61] L. E. Reichl, *A modern Course in Statistical Physics*. Wiley-VCH, 2009.



## Part II

# Studying the TFET using analytical models



## Chapter 2

# Analytical model for line tunneling

### Analytical Model for a Tunnel Field-Effect Transistor

William G. Vandenberghe, Anne S. Verhulst, Guido Groeseneken,  
Bart Sorée and Wim Magnus

Published in *The 14th IEEE Mediterranean Electrotechnical Conference (MELECON), 2008*, pp. 923 - 928 , May 2008.

©2008 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

### **Contributions of first author**

- Conception of the idea to make an analytical model for the TFET
- Development of the analytical model
- Performing device simulations
- Text redaction and editing

## Abstract

The tunnel field-effect transistor (TFET) is a promising candidate for the succession of the MOSFET at nanometer dimensions. Due to the absence of a simple analytical model for the TFET, the working principle is generally not well understood. In this paper a new TFET structure is introduced and using Kane's model, an analytical expression for the current through the TFET is derived. Furthermore, a compact expression for the TFET current is derived and conclusions concerning TFET design are drawn. The obtained analytical expressions are compared with results from a 2D device simulator and good agreement at low gate voltages is demonstrated.

## 2.1 Introduction

As MOSFETs reach nanometer dimensions, power consumption becomes a major bottleneck for further scaling. The continued reduction of the MOSFET size is leading to an increased leakage current due to short channel effects, such as Drain Induced Barrier Lowering (DIBL), and the power supply voltage cannot be reduced any further because of the subthreshold slope being limited to 60 mV/decade at room temperature. In this view, the exploration of alternative devices which possibly outperform the MOSFET at these nanometer dimensions is required.

A promising alternative for the MOSFET, which does not suffer from these limitations, is the tunneling field-effect transistor (TFET). Throughout the rest of this paper the term TFET will not refer to any specific implementation form, but the TFET is defined as "a semiconductor device in which the gate controls the source-drain current through modulation of Band-to-Band Tunneling (BTBT)". Band-to-Band Tunneling is a process in which electrons tunnel from the valence band through the semiconductor bandgap to the conduction band or vice versa.

TFETs with a subthreshold slope lower than 60 mV/decade have already been demonstrated [1, 2] and due to their built-in tunnel barrier, Si TFETs are expected to maintain low off-currents for channel lengths down to 10 nm [3].

An advantage of TFETs compared to other alternative device concepts is that their fabrication is compatible with standard CMOS processing since they can be implemented as a reverse biased gated *p-i-n* diode. Moreover, compared to for example the I-MOS [4], TFETs do not rely on high energetic processes like impact ionization, which are known to be detrimental to reliability.

Contrary to the MOSFET and the bipolar transistor, the TFET does not have

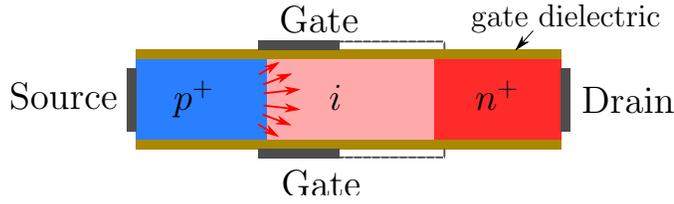


Figure 2.1: Double gate TFET with full (solid) and short (dashed) gate respectively as shown in [5], the direction of BTBT in the semiconductor is indicated by arrows

a simple analytical model. This hampers a clear understanding of the TFET working principle.

In this paper, an analytical model for a TFET is developed. First a new TFET device structure is presented for which the potential profile can be determined straightforwardly. As a result, the total current through the device can be calculated analytically, as shown in Section 2.2. In order to obtain a closed expression for the current, an approximation regarding the behavior of the current around the BTBT onset voltage needs to be made, as discussed in Section 2.3. The calculations originating from the analytical expression are validated in Section 2.4 by comparing them with device simulator results.

## 2.2 Analytical model

### 2.2.1 Device structure

The TFET is generally described as a gated  $p-i-n$  diode. Investigations [5] have shown that the gate does not need to cover the entire intrinsic region but can be restricted to the area close to the source. A double gate TFET is shown in Fig. 2.1 illustrating the full and the short gate concept. Neither of these structures allow for a simple analytical expression for the electric field and the electrostatic potential throughout the device. This hampers an analytical treatment that would give insight into the working principle of the device.

In this paper a new TFET configuration is presented such that the gate is located fully on top of the source as illustrated in Fig. 2.2. The device can be regarded as an extreme case of the short gate TFET as the gate does not even cover a part of the channel. It is important to note that in this device, the BTBT occurs in the direction orthogonal to the gate. The one-dimensional nature of the BTBT enables an approximate determination of the potential profile in the tunneling region.

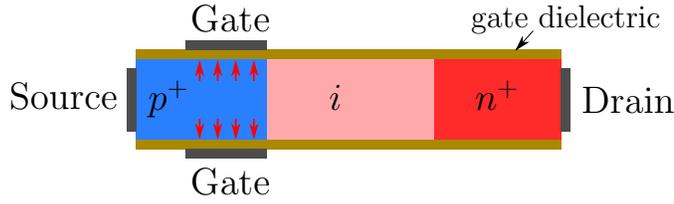


Figure 2.2: New device concept which enables derivation of an analytical model, the direction of BTBT in the semiconductor is indicated by arrows

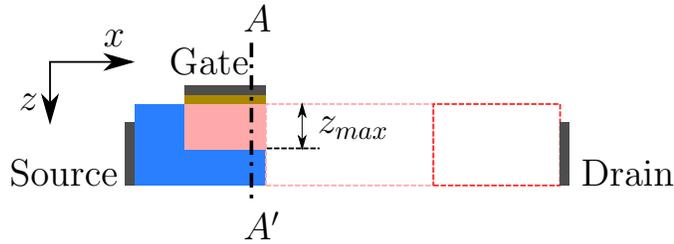


Figure 2.3: TFET from Fig. 2.2 with positive gate bias applied such that a depletion region with thickness  $z_{\max}$  exists (only upper half shown)

The derivation of the current formula in this paper is restricted to the case of an  $n$ TFET, i.e. the gate is located on top of the  $p$ -type source. An analogous calculation can be made for a  $p$ TFET with an  $n$ -type source.

### 2.2.2 Potential profile

Whereas the gate strictly controls the electrostatic potential in the underlying semiconductor, the influence of the drain voltage is found to be weak and can be neglected in the region directly under the gate.

Under the assumption that there is no substantial potential variation in the direction parallel to the gate, the potential profile can be considered one-dimensional. Adopting the depletion layer approximation, the electric field and the electrostatic potential in the semiconductor are written in terms of the acceptor doping concentration of the source ( $N_a$ ) [6]:

$$\psi(z) = \frac{qN_a}{2\epsilon_s} (z - z_{\max})^2 \quad (2.1)$$

$$E_z(z) = -\frac{qN_a}{\epsilon_s} (z - z_{\max}) \quad (2.2)$$

where  $z_{\max}$  is the length of the depletion region, which is a function of the applied gate voltage.  $q$  is the elementary electric charge and  $\epsilon_s$  is the permittivity of the semiconductor.

### 2.2.3 Current

Since a *p-i-n* diode in reverse bias is considered, the current is small if no BTBT is present, in which case the current is referred to as the off-current. As soon as BTBT occurs however, the resulting on-current will dominantly exceed the off-current.

In semi-classical simulators, BTBT is modelled by the introduction of an extra generation term ( $G$ ) in the drift-diffusion equation. When the BTBT current contribution is dominating, the TFET current can be computed as the sum over all charge generated in the device:

$$|I| = q \int G dV = qWL \int G dz \quad (2.3)$$

with  $dV$  an elementary volume in the device,  $L$  and  $W$  the gate length and width respectively and  $G$  the generation rate expressed in number of carriers per unit volume per unit time. In Eq. 3 translational invariance in the  $y$ -direction is assumed and the variation of the generation rate in the  $x$ -direction is neglected.

The most popular model to calculate the generation is Kane's Model [7]. Kane's model is derived for a direct semiconductor in a uniform electric field and is given by:

$$G(E) = A \frac{E^D}{\sqrt{E_g}} \exp\left(-BE_g^{3/2}/E\right) \quad (2.4)$$

where  $E$  is the electric field and  $E_g$  the bandgap, while  $A$  and  $B$  are parameters depending on the effective mass of valence and conduction bands and  $D$  takes a default value of 2, but for the sake of generality it is left unspecified as an adjustable parameter.

The effect of the different regions arising when a positive gate voltage is applied, is discussed in the following subsections.

#### Neutral region

The neutral region is characterized by the absence of an electric field and does not contribute to the BTBT current.

#### Depletion region

In this section, the current is calculated assuming that there is only a depletion region. Instead of relying on the local electric field as in Refs. [8, 9], we insert the average electric field over the tunnel path into the Kane generation rate.

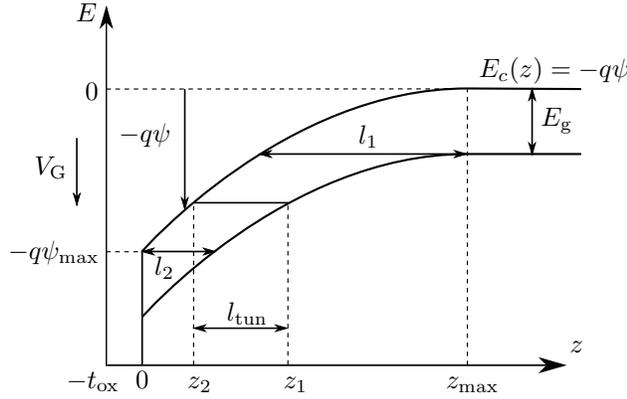


Figure 2.4: Band diagram of the cross-section A-A' from Fig. 2.3 indicating  $l_{\text{tun}}$

A similar strategy is followed by some commercial device simulators [10]. Accordingly, a more accurate estimate of the current can be made, and the tunnel current disappears in a natural way when no tunnel path is available due to the average electric field being zero.

The tunnel path is defined as the physical path between two points corresponding to equal energy for the conduction and valence band respectively. The average field in the  $z$ -direction ( $\bar{E}_z$ ) is then given by:

$$\bar{E}_z = \frac{E_g}{ql_{\text{tun}}} \quad (2.5)$$

with  $E_g$  the bandgap energy and  $l_{\text{tun}}$  the length of the tunnel path.

To calculate  $l_{\text{tun}}$ , the  $z$ -coordinate of points of equal potential in valence and conduction band are considered. Writing,

$$\psi_v(z_1) = \frac{qN_a}{2\epsilon_s}(z_1 - z_{\text{max}})^2 + \frac{E_g}{q} \quad (2.6)$$

$$\psi_c(z_2) = \frac{qN_a}{2\epsilon_s}(z_2 - z_{\text{max}})^2 \quad (2.7)$$

from which the tunnel path length can be determined:

$$l_{\text{tun}} = z_1 - z_2 \quad (2.8)$$

$$\psi_v(z_1) = \psi_c(z_2) \quad (2.9)$$

Writing  $z_2$  as a function of  $l_{\text{tun}}$ ,

$$z_2 = z_{\text{max}} - \frac{1}{2} \frac{l_{\text{tun}}^2 + 2E_g\epsilon_s/(q^2N_a)}{l_{\text{tun}}} \quad (2.10)$$

this yields following expression for  $dz$ :

$$dz = -\frac{1}{2} \left( 1 - \frac{2E_g \epsilon_s}{q^2 N_a} \frac{1}{l_{\text{tun}}^2} \right) dl_{\text{tun}} \quad (2.11)$$

Substituting Eqs. (2.4), (2.5) and (2.11) into Eq. (2.3):

$$I = \frac{qWLA}{2} \int_{l_1}^{l_2} \frac{E_g^{D-\frac{1}{2}}}{q^D l_{\text{tun}}^D} e^{-Bq\sqrt{E_g}l_{\text{tun}}} \left( 1 - \frac{2E_g \epsilon_s}{q^2 N_a} \frac{1}{l_{\text{tun}}^2} \right) dl_{\text{tun}} \quad (2.12)$$

$l_1$  and  $l_2$  respectively denote the maximal and minimal length of the tunnel path in the depletion region:

$$l_1 = \sqrt{\frac{2E_g \epsilon_s}{q^2 N_a}} \quad (2.13)$$

$$l_2 = \sqrt{\frac{2\epsilon_s}{qN_a}} \left( -\sqrt{\psi_{\text{max}} - \frac{E_g}{q}} + \sqrt{\psi_{\text{max}}} \right) \quad (2.14)$$

$\psi_{\text{max}}$  is the potential at the end of the depletion region as indicated in Fig. 2.4.

Integration by parts reduces the integral in (2.12) to a logarithmic integral, that cannot be evaluated analytically.

### Inversion layer

The gate voltage specifying the onset of inversion is given by the well-known expression:

$$V_{\text{GS}} = V_{\text{FB}} + 2\phi_f + \Phi_n + \frac{t_{\text{ox}}}{\epsilon_{\text{ox}}} \sqrt{2qN_a \epsilon_s} \sqrt{\Phi_n + 2\phi_f} \quad (2.15)$$

where  $\Phi_n$  is the electron quasi-Fermi level relative to the source voltage,  $V_{\text{FB}}$  is the flatband voltage,  $\phi_f = \ln(N_a/n_i) kT/q$  is the surface potential,  $t_{\text{ox}}$  is the oxide thickness and  $\epsilon_{\text{ox}}$  is the dielectric constant of the oxide.

The respective positions of the quasi-Fermi levels also determine the ratio between BTBT generation and BTBT recombination. In the previous treatment for the depletion region the respective positions of the quasi-Fermi levels were ignored and BTBT was entirely attributed to generation. This is an acceptable approximation as long as the valence band is filled and the conduction band is empty.

In the case of a non-degenerate semiconductor, the inversion regime sets in before the electron quasi-Fermi level and the conduction band meet. In the

depletion region, only the tail of the Fermi-Dirac distribution occupies the conduction band which is negligible compared to the almost fully occupied valence band. In the inversion layer however, the conduction band is filled and BTBT recombination can no longer be neglected.

However, the inversion layer, being small in size and having only a small generation rate, only gives rise to a small contribution to the total current and therefore its contribution is neglected in this paper.

### Total current

Using the previous approach, the total current in the device equals the current generated in the depletion region. The influence of the gate voltage is straightforward, namely increasing the depletion region size and increasing the current accordingly. The drain voltage determines the onset of inversion and therefore imposes an upper limit to the depletion region size through Eq. (2.15).

In a degenerate semiconductor, the limits  $l_1, l_2$  of the integration (2.12) need to be altered such that the integration only extends over the region of tunneling where the conduction band is (approximately) empty and the valence band is (approximately) full.

## 2.3 Approximating formula

### 2.3.1 Derivation

In order to further elaborate on Eq. (2.12) it is assumed that the exponential terms change much more rapidly than the polynomial terms under a variation of  $l_{\text{tun}}$ . Approximating the integral and inserting Eq. (2.13), an expression for the current as a function of  $l_2$  is derived:

$$I \approx -\frac{WLA E_g^{D-1}}{2Bq^D} \left( \frac{1}{l_2^D} - \frac{2E_g \epsilon_s}{q^2 N_a} \frac{1}{l_2^{D+2}} \right) e^{-Bq\sqrt{E_g}l_2} \quad (2.16)$$

Next,  $\psi_{\text{max}}$  is calculated as a function of the applied gate voltage  $V_G$ .  $V_G$  can be written as the sum of the electrostatic potential at the interface ( $\psi_{\text{max}}$ ) and the potential difference over the oxide,

$$V_{\text{GS}} - V_{\text{FB}} = \psi_{\text{max}} + t_{\text{ox}} E_{\text{ox}} \quad (2.17)$$

Since no potential change along the  $x$ -direction is considered, the electric field has only a component in the  $z$ -direction. The oxide electric field ( $E_{\text{ox}}$ ) can

be calculated from the semiconductor electric field ( $E_s$ ) by making use of the boundary condition at the interface:  $\epsilon_{\text{ox}}E_{\text{ox}} = \epsilon_s E_s$ .

The electric field changes linearly in the depletion region while the electrostatic potential changes quadratically according to Eqs. (2.1) and (2.2). Writing the gate voltage as a function of  $\psi_{\text{max}}$ :

$$V_{\text{GS}} - V_{\text{FB}} = \psi_{\text{max}} + 2t_{\text{ox}} \frac{\epsilon_s}{\epsilon_{\text{ox}}} \sqrt{\frac{qN_a}{2\epsilon_s} \psi_{\text{max}}} \quad (2.18)$$

Combining Eqs. (2.18), (2.16) and (2.14) is useful for comparison of the analytical model with simulation results. The resulting formula however, is rather involved and still hampers a straightforward conclusion. We may however gain more insight by calculating the current in the neighborhood of the TFET onset voltage  $V_{\text{onset}}$  which marks the onset of BTBT in the depletion region, or equivalently, the band bending being equal to the bandgap,  $\psi_{\text{max}} = E_g/q$ . Hence,

$$V_{\text{onset}} = V_{\text{FB}} + \frac{E_g}{q} \left( 1 + 2t_{\text{ox}} \frac{\epsilon_s}{\epsilon_{\text{ox}}} \sqrt{\frac{q^2 N_a}{2E_g \epsilon_s}} \right) \quad (2.19)$$

Calculating the current given by Eq. (2.16) for  $\psi_{\text{max}} = E_g/q + \delta\psi$  where  $\delta\psi$  is a perturbation of the electrostatic potential, we arrive at:

$$I \approx \frac{WLA\sqrt{q}}{BE_g^{3/2}} \left( \frac{E_g N_a}{2\epsilon_s} \right)^{D/2} e^{Bq\sqrt{2\epsilon_s E_g}/\sqrt{qN_a}(\sqrt{\delta\psi} - \sqrt{E_g/q})} \sqrt{\delta\psi} \quad (2.20)$$

Writing  $\delta\psi$  as a function of the gate voltage:

$$\delta\psi = (V_{\text{GS}} - V_{\text{onset}}) / \gamma \quad (2.21)$$

$$\gamma = 1 + t_{\text{ox}} \frac{\epsilon_s}{\epsilon_{\text{ox}}} \sqrt{\frac{q^2 N_a}{2E_g \epsilon_s}} \quad (2.22)$$

we obtain a closed formula for  $D = 2$ ,

$$I \approx WLT e^{S\sqrt{V_{\text{GS}} - V_{\text{onset}}}} \sqrt{V_{\text{GS}} - V_{\text{onset}}} \quad (2.23)$$

with

$$T = q \frac{A}{Bq^{3/2}} \frac{qN_a}{2\epsilon_s} \sqrt{\frac{1}{E_g \gamma}} e^{-BqE_g\sqrt{2\epsilon_s}/\sqrt{q^2 N_a}} \quad (2.24)$$

$$S = Bq \sqrt{\frac{2E_g \epsilon_s}{qN_a} \frac{1}{\gamma}} \quad (2.25)$$

### 2.3.2 Interpretation

Eq. (2.23) provides the on-current in the TFET as a function of gate voltage when the device parameters are known. Clearly, the square root dependence indicates the absence of a 60 mV/decade subthreshold slope.

The influence of the different parameters on the prefactor  $T$  (Eq. (2.24)) and the onset voltage  $V_{\text{onset}}$  (Eq. (2.19)) is illustrated in table 2.1 and will be discussed next. The prefactor  $T$  in front of the exponential rather than the exponential itself is considered, as it determines the current in the nearest neighborhood of the onset voltage.

The bandgap is one of the most important parameters when considering a TFET, since it determines the “barrier” between valence and conduction band. The on-current increases with decreasing bandgap as indicated by Eq. (2.24), while also  $V_{\text{onset}}$  is reduced since the voltage required to create a path from valence to conduction band is directly proportional to the bandgap. A small bandgap is therefore beneficial for a large TFET on-current and is desirable as far as it does not jeopardize the TFET off-current.

It is however important to keep in mind that Eq. (2.23) is derived using Kane’s model which applies to a direct semiconductor in a uniform field. For an indirect semiconductor a more precise treatment is required. In addition, quantization effects due to the small size of the depletion region that were not considered in this work, should also be taken into account.

A higher doping level of the source increases the on-current. This can be understood by realizing that the doping level determines the curvature of the potential in the depletion region. A larger doping level will thus decrease the tunnel distance and increase the current. An upper limit on the onset voltage or on the voltage drop over the oxide will limit the doping level.

The oxide thickness has little impact on the on-current while its influence on the onset voltage is straightforward. I.e. a smaller oxide thickness reduces the voltage drop over the oxide and reduces the onset voltage. A small oxide thickness is required to enable the gate to adequately control the source region and it will therefore improve the validity of the approximations.

Finally it is important to notice that Eq. (2.23) has been derived for the structure shown in Fig. 2.2. The conclusions drawn in this section cannot be directly applied to a general TFET structure, although the working principle remains the same.

Table 2.1: Example values for  $T$  and  $V_{\text{onset}}$ 

Doping level ( $\text{cm}^{-3}$ )	Bandgap (eV)	$T$ ( $\text{Acm}^{-2}\text{V}^{-1/2}$ )	$V_{\text{onset}}$ (V)
$10^{19}$	1	$1.3 \cdot 10^{-4}$	1.2
$10^{19}$	0.5	66	0.64
$2 \cdot 10^{19}$	1	0.5	1.3
$2 \cdot 10^{19}$	0.5	$5.6 \cdot 10^3$	0.7
$10^{20}$	1	$4.9 \cdot 10^4$	1.6
$10^{20}$	0.5	$3.9 \cdot 10^6$	0.95

## 2.4 Comparison with simulation results

To obtain Eq. (2.12) and the formulas thereafter a number of assumptions had to be made: the TFET current was restricted to the dominating BTBT contribution in the depletion region under the source. It is however not proven that the BTBT contribution in the channel or the inversion layer contribution are negligible.

In order to verify the above approximations, the derived formulas (2.12) and (2.23) are calculated numerically and compared with simulation results.

In Fig. 2.5 and Fig. 2.6 the current calculated by the analytical model and the current calculated by Medici [10], both in a self-consistent and a non-selfconsistent way, are shown. The non-selfconsistent current calculation implies a self-consistent solution of the drift-diffusion equations without taking BTBT into account. Based on the potential profile calculated in this way, the total amount of BTBT is calculated which is then identified with the total current.

The analytical results agree very well with the non-selfconsistent calculations by Medici. For large gate voltages, the (more accurate) self-consistent calculation of the current shows a difference, because the electron quasi-Fermi level under the gate has been assumed constant and equal to the drain voltage.

The electrons generated due to BTBT are bound to flow towards the drain, which may cause a significant drop of the electron quasi-Fermi level between the position where the generation takes place and the drain, and can be thought of as a series resistance. The latter depends on the mobility, thickness and length of the channel. The series resistance interpretation can be verified in Fig. 2.6 where simulations for different oxide thicknesses are compared.

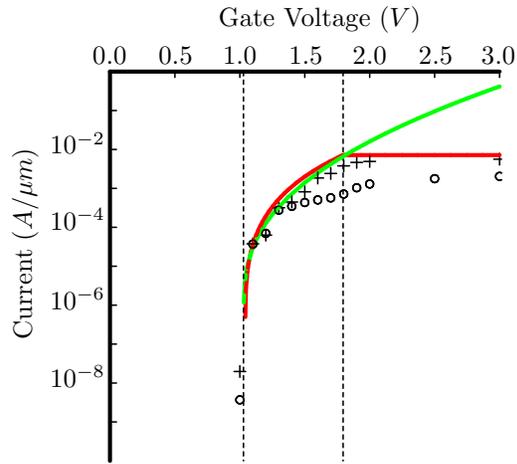


Figure 2.5: Self-consistent simulations ( $\circ$ ) and non-selfconsistent simulations ( $+$ ) compared with analytical formulas (2.12) ( $-$ ) and (2.23) ( $-$ ) with the gate voltage of the former limited according to Eq. (2.15).

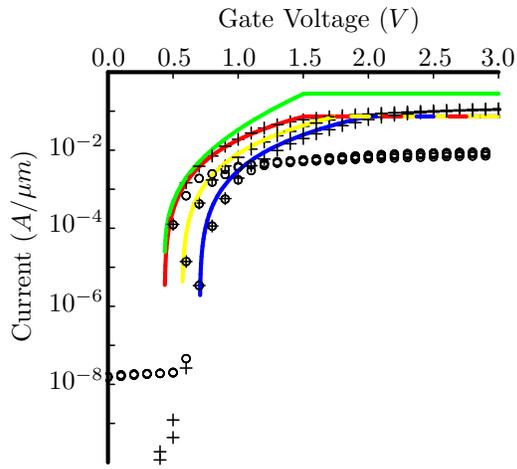


Figure 2.6: Self-consistent simulations ( $\circ$ ) and non-selfconsistent simulations ( $+$ ) compared with analytical formula (2.12) for an oxide thickness of 2 nm ( $-$ ), 3 nm ( $-$ ) and 4 nm ( $-$ ) and formula (2.23) for 2 nm ( $-$ ) with the gate voltage limited according to Eq. (2.15).

## 2.5 Conclusion

By considering a TFET structure with a gate on top of the source, it is possible to obtain an analytical description of the potential profile in the TFET. Using this potential profile and adopting Kane's Model, we derived an analytical expression for the on-current. The approximations accompanying this derivation are justified by comparison with numerical simulation results.

The compact expression for the current shows the impact of the doping level, bandgap and oxide thickness on the TFET on-current.

For the design of a TFET as introduced in this paper, a small bandgap and a high doping level of the source are beneficial to achieving a high on-current. But the bandgap cannot be reduced indefinitely because of the off-current whereas the doping level is also limited by the TFET onset voltage. A small oxide thickness improves the validity of the approximations which were made.

## 2.A Parameters used

### 2.A.1 Medici parameters used for Fig. 2.5

Semiconductor material: Silicon  
 Gate dielectric: HfO<sub>2</sub>  $t_{ox} = 2$  nm  
 Source doping: uniform  $p$ -type  $10^{20}$  cm<sup>-3</sup>  
 Channel doping: uniform  $n$ -type  $10^{13}$  cm<sup>-3</sup>  
 Drain doping: uniform  $n$ -type  $10^{19}$  cm<sup>-3</sup>  
 Gate workfunction: (Si) Neutral  
 Models: BTBT BT.MODEL=3 BGN (self-consistent)  
 Models: ^BTBT (non-selfconsistent)  
 Gate length ( $L$ ): 24 nm  
 Drain voltage ( $V_{DS}$ ): 0.5 V  
 Channel length: 8 nm

### 2.A.2 Medici parameters used for Fig. 2.6

Semiconductor material: Semicond 0.5 V (0.38 V after BGN),  
 mobility  $10^4$  cm<sup>2</sup>/(Vs)  
 Gate dielectric: HfO<sub>2</sub>  $t_{ox} = 2/3/4$  nm  
 Source doping: uniform  $p$ -type  $5 \cdot 10^{19}$  cm<sup>-3</sup>  
 Channel doping: uniform  $n$ -type  $10^{13}$  cm<sup>-3</sup>  
 Drain doping: uniform  $n$ -type  $10^{19}$  cm<sup>-3</sup>  
 Gate workfunction: (Si) Neutral  
 Models: BTBT BT.MODEL=3 BGN (self-consistent)

Models:  $\hat{\text{BTBT}}$  (non-selfconsistent)  
 Gate length ( $L$ ): 24 nm  
 Drain voltage ( $V_{\text{DS}}$ ): 0.5 V  
 Channel length: 45 nm

### 2.A.3 Device parameters used for Table 2.1

Gate dielectric:  $t_{\text{ox}} = 2$  nm  
 Flatband Voltage:  $V_{\text{FB}} = 0$  V  
 BTBT parameters:  $A = 3.5 \cdot 10^{21} (\text{eV})^{1/2} / (\text{cm} \cdot \text{s} \cdot \text{V}^2)$ ,  
 $B = 22.5 \cdot 10^6 \text{ V} / (\text{cm} \cdot (\text{eV})^{3/2})$   
 Dielectric constants:  $\epsilon_{\text{s}} = 11.8\epsilon_0$ ,  $\epsilon_{\text{s}} = 21\epsilon_0$

## Acknowledgements

This work was supported by IMEC's Industrial Affiliation Program. Anne S. Verhulst acknowledges the support of a Marie Curie International Reintegration Grant within the 6th European Community Framework Programme, as well as a postdoctoral fellowship of the Fund for Scientific Research - Flanders.

## Bibliography

- [1] J. Appenzeller, Y.-M. Lin, J. Knoch, Z. Chen, and P. Avouris, "Comparing carbon nanotube transistors - the ideal choice: a novel tunneling device design," *Electron Devices, IEEE Transactions on*, vol. 52, no. 12, pp. 2568–2576, Dec. 2005.
- [2] W. Y. Choi, B.-G. Park, J. D. Lee, and T.-J. K. Liu, "Tunneling field-effect transistors (TFETs) with subthreshold swing (SS) less than 60 mV/dec," *Electron Device Letters, IEEE*, vol. 28, pp. 743–745, Aug. 2007.
- [3] K. K. Bhuvalka, J. Schulze, and I. Eisele, "Scaling the vertical tunnel fet with tunnel bandgap modulation and gate workfunction engineering," *IEEE Transactions on Electron Devices*, vol. 52, no. 5, May 2005.
- [4] K. Gopalakrishnan, P. Griffin, and J. Plummer, "Impact ionization MOS (I-MOS)-Part I: device and circuit simulations," *Electron Devices, IEEE Transactions on*, vol. 52, no. 1, pp. 69–76, January 2005.
- [5] A. S. Verhulst, W. G. Vandenberghe, K. Maex, and G. Groeseneken, "A tunnel field-effect transistor without gate-drain overlap," *Applied Physics Letters*, vol. 91, no. 053102, July 2007.

- [6] S. Sze, *Physics of Semiconductor Devices*, 2nd ed. John Wiley & Sons, 1981.
- [7] E. O. Kane, "Zener tunneling in semiconductors," *Journal of Physics and Chemistry of Solids*, vol. 12, pp. 181–188, 1959.
- [8] T. Chan, J. Chen, P. Ko, and C. Hu, "The impact of gate-induced drain leakage current on MOSFET scaling," *Electron Devices Meeting, 1987 International*, vol. 33, pp. 718–721, 1987.
- [9] A. Qin Zhang; Wei Zhao; Seabaugh, "Low-subthreshold-swing tunnel transistors," *Device Letters, IEEE*, vol. 27, no. 4, pp. 297–300, April 2006.
- [10] *Medici User Guide*, Synopsis, March 2007.

## Chapter 3

# Analytical model for point tunneling

### Analytical Model for Point and Line Tunneling in a Tunnel Field-Effect Transistor

William G. Vandenberghe, Anne S. Verhulst, Guido Groeseneken,  
Bart Sorée and Wim Magnus

Published in *Proceedings of the International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, 2008, pp. 137 - 140 , September 2008.

©2008 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

**Contributions of first author**

- Conception and development of the analytical model for point tunneling
- Performing device simulations
- Creating figures
- Writing the manuscript

## Abstract

The tunnel field-effect transistor (TFET) is a promising candidate for the succession of the MOSFET at nanometer dimensions. In general, the TFET current can be decomposed into two components referred to as point tunneling and line tunneling. In this paper we derive a compact analytical model for the current due to point tunneling complementing the previously derived analytical model for line tunneling. We show that the derived analytical expression for point tunneling provides a more consistent estimate of the TFET current than a commercial device simulator. Both the line and point tunneling current do not show a fixed subthreshold-slope. Three key parameters for design of a TFET are: bandgap, dielectric thickness and source doping level. A small bandgap is beneficial for a high TFET on-current and a low onset voltage. Point tunneling and line tunneling show a strong dependence on gate dielectric thickness and doping concentration respectively.

## 3.1 Introduction

As MOSFETs reach nanometer dimensions, exploration of alternative devices that possibly outperform the MOSFET at the nanometer scale is required. A promising alternative, which does not suffer from the fundamental subthreshold-slope limitation, is the tunnel field-effect transistor (TFET) in which the gate modulates the Band-to-Band Tunneling (BTBT) current between source and drain [1, 2]. Contrary to the MOSFET and the bipolar transistor, there is no simple analytical model available for the current of the general TFET configuration and this hampers a clear understanding of its working principle.

Wang discovered that the TFET current is composed of two components [3]. The first component "point tunneling" occurs at the source-channel interface and its dominant contribution is localized in a small area. The second component is located in the part of the source region overlapped by the gate. Because the area where BTBT starts from resembles a line, this component is called "line tunneling". An analytical model for line tunneling was recently derived yielding the total TFET current when the gate only covers the source area and point tunneling is negligible [4, 5]. In the general case with the gate positioned over a part of the channel, point tunneling must also be considered.

In this paper we derive an analytical expression for the point tunneling current contribution (Section 3.3) and briefly repeat the treatment for line tunneling (Section 3.4). In this way a compact expression is derived for the total TFET current.

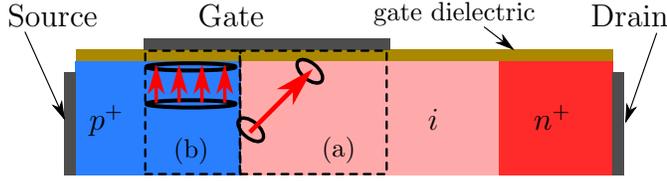


Figure 3.1: General  $n$ TFET configuration with the region of point tunneling (a) and line tunneling (b) schematically indicated. The direction of BTBT in the semiconductor is indicated by arrows and the regions of highest tunneling efficiency are circumscribed by an ellipsoid. The support of the device (not shown) can either be bulk, insulator or a second gate as long as its influence on the electrostatic potential is negligible.

### 3.2 Current in a TFET

The TFET is generally described as a gated reverse biased  $p$ - $i$ - $n$  diode where the gate can be restricted to the area close to the source as illustrated in Fig. 3.1. In our calculations line and point tunneling are treated separately and their contributions are added together to compute the total current.

For the calculation of the current we follow a strategy similar to that of a commercial device simulator [6]. First we determine the potential profile  $\psi(\mathbf{r})$  in the area of interest. Next we determine the distance  $l(\mathbf{r})$  an electron must travel to tunnel from valence to conduction band. Knowing the tunnel distance, the generation rate can be calculated using Kane's model as a function of the average electric field ( $\mathcal{E} = E_g/l$ ) over the tunnel path:

$$G = A \frac{\mathcal{E}^D}{\sqrt{E_G}} \exp\left(-BE_G^{3/2}/\mathcal{E}\right) \quad (3.1)$$

$$= A \frac{E_G^{D-1/2}}{q^D l^D} \exp\left(-Bq\sqrt{E_G}l\right) \quad (3.2)$$

with  $G$  the generation rate expressed in number of carriers per unit volume per unit time,  $E_G$  the bandgap,  $q$  the elementary charge and  $A$ ,  $B$  and  $D$  material dependent parameters of Kane's model [7].

Finally the total TFET current is computed by integrating the generation rate and ignoring the  $p$ - $i$ - $n$  diode leakage current:

$$|I| = q \int G dV \quad (3.3)$$

with  $dV$  an elementary volume in the device.

The drain is typically located some distance away from the source and generally has little direct impact on the electrostatics of the tunnel barrier. As a

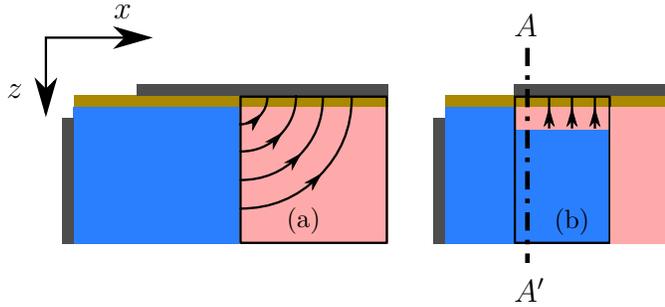


Figure 3.2: Illustration of the electric field lines (black lines with arrows) in the channel (a) and under the gate in the depleted region (b) for the structure from Fig. 2.2. A two-dimensional potential profile is present in the intrinsic region and at high gate bias a significant depletion region will emerge under the gate.

consequence only the gate and source contact determine the potential in the following treatment, obscuring the influence of the drain voltage. For low drain voltages the actual device current will be lower than our calculated results if the conduction band is occupied, this can be mediated by restricting the area of integration. For degenerately doped semiconductors, the device current will be lower because the valence band is not completely occupied. At high currents, the lower actual device current can be interpreted as the consequence of a resistance in series with the tunnel barrier. The main reason for an overestimation of the actual device current is the absence of a self-consistent determination of the potential in our calculations.

### 3.3 Point tunneling

In this section we consider the two-dimensional tunneling at the source-channel interface.

When a positive gate voltage is applied, a depletion region at the source-channel interface is formed, the channel potential profile changes and a limited amount of charge will flow into the channel from the drain. In the case of a highly doped source – which is beneficial to achieving a high on-current – a small depletion region is enough to provide a significant electric field in the channel.

To describe the potential, we adopt the  $x$  and  $z$  direction as shown in Fig. 3.2 and assign the coordinate  $x = 0$  to the source-channel interface and the coordinate  $z = 0$  to the gate-gate dielectric interface.

To determine the potential profile we make four assumptions. A first assump-

tion is that the potential drop due to depletion in the source is negligible and the potential at  $x = 0$  equals the source potential. Secondly, we assume the influence of the charge in the channel on the channel potential can be neglected. A third assumption is the absence of charge in the gate dielectric and a fourth assumption is that gate dielectrics with identical electrical thickness will result in an identical potential profile in the intrinsic region. For the remainder of our treatment, we use a semiconductor equivalent thickness:

$$t'_{\text{ox}} = t_{\text{ox}} \frac{\epsilon_{\text{s}}}{\epsilon_{\text{ox}}} \quad (3.4)$$

with  $\epsilon_{\text{s}}$  and  $\epsilon_{\text{ox}}$  the semiconductor and gate dielectric constant respectively and  $t_{\text{ox}}$  is the physical dielectric thickness.

The potential in the channel and in the gate dielectric satisfy Poisson's equation

$$\nabla^2 \psi = -\frac{\rho}{\epsilon_{\text{s}}} \quad (3.5)$$

with  $\rho$  the charge density.

As a consequence of our second, third and fourth assumption,  $\rho = 0$  in the intrinsic semiconductor and gate dielectric region and the equation reduces to Laplace's equation. Moreover, the electrostatic potential is to obey the boundary conditions:

$$\psi(0, z) = \psi_{\text{s}} \quad , \quad \psi(x, 0) = \psi_{\text{g}} \quad (3.6)$$

where  $\psi$  is the potential taken at the semiconductor valence band edge and  $\psi_{\text{g}}$  and  $\psi_{\text{s}}$  are the potential in the source and gate respectively. For convenience, we take  $\psi_{\text{s}} = 0$ , the gate potential is then related to the applied gate voltage  $V_{\text{gs}}$  by:

$$\psi_{\text{g}} = V_{\text{gs}} - V_{\text{FB}} \quad (3.7)$$

with  $V_{\text{FB}}$  the flatband voltage.

Assuming the semiconductor extends towards infinity for ( $x > 0$  and  $z > 0$ ), the solution for the potential is most easily given in polar coordinates ( $x = r \sin \theta$  and  $z = r \cos \theta$ ):

$$\psi(x, z) = \psi_{\text{g}} \frac{2}{\pi} \theta \quad ; \quad 0 \leq \theta \leq \frac{\pi}{2} \quad (3.8)$$

In the case of our two-dimensional potential, the tunnel path length is computed as the length of an arc along an electric field line:

$$l = r \theta_0 \quad (3.9)$$

with  $\theta_0 = \pi E_{\text{G}} / (2q\psi_{\text{g}})$  the angle between two equipotential lines with a potential difference equal to  $E_{\text{g}}/q$ .

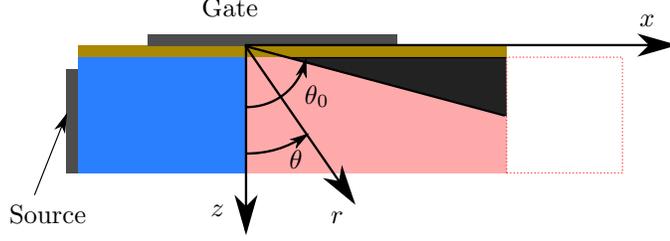


Figure 3.3: Area over which the generation rate must be integrated for point tunneling (black) where the source-channel interface is shown on the left (blue-pink) and the gate-gate dielectric interface on top (grey-brown). The structure is assumed to extend towards infinity at the bottom and towards the right.

We calculate the total current by integrating the generated charge over the entire area where tunneling occurs as shown in Fig. 3.3:

$$I = qW \int_{r_0}^{\infty} \int_{\theta_0}^{\arccos(t'_{\text{ox}}/r)} G(r)r \, d\theta \, dr \quad (3.10)$$

$$= WA \int_{r_0}^{\infty} (\arccos(t'_{\text{ox}}/r) - \theta_0) \frac{E_G^{D-\frac{1}{2}} e^{-Bq\sqrt{E_G}r\theta_0}}{q^{D-1}r^{D-1}\theta_0^D} \, dr \quad (3.11)$$

with  $r_0 = t'_{\text{ox}}/\cos(\theta_0)$ .

Making a first order Taylor expansion around  $r = r_0$ , we integrate by parts and neglecting  $\cos(\theta_0)$  with respect to unity to get:

$$I \approx \frac{WAE_G^{D-\frac{1}{2}}}{q^{D-1}r_0^{D-1}\theta_0^D} \int_{r_0}^{\infty} \frac{t'_{\text{ox}}(r-r_0)e^{-Bq\sqrt{E_G}r\theta_0}}{\sqrt{1-(t'_{\text{ox}}/r_0)^2r_0^2}} \, dr \quad (3.12)$$

$$\approx \frac{WAE_G^{D-\frac{3}{2}}t'_{\text{ox}}}{q^{D-1}B^2} \frac{1}{\theta_0^{D+2}r_0^{D+1}} e^{-Bq\sqrt{E_G}r_0\theta_0} \quad (3.13)$$

To come to a polynomial pre-factor we approximate  $r_0$  as a function of the gate potential

$$r_0 = t'_{\text{ox}}/\sin\left(\frac{\pi(q\psi_g - E_G)}{2q\psi_g}\right) \approx t'_{\text{ox}} \frac{2q\psi_g}{\pi(q\psi_g - E_G)} \quad (3.14)$$

Substituting Eq. (3.14), the value of  $\theta_0$  and the default value of  $D = 2$  into Eq. (3.13):

$$I = WT \frac{q\psi_g}{E_G} \left(\frac{q\psi_g}{E_G} - 1\right)^3 e^{S/(q\psi_g/E_G - 1)} \quad (3.15)$$

with  $\psi_g$  given by Eq. (3.7) and

$$T = \frac{2A\sqrt{E_G}}{\pi q^3 B^2 t_{\text{ox}}^2} \quad , \quad S = -Bq\sqrt{E_G} t_{\text{ox}}' \quad (3.16)$$

A comparison between Eq. (3.15) and device simulation results for the structure of Fig. 2.2 and its  $p$ TFET equivalent is shown in Fig. 3.5 and Fig. 3.6 respectively. Around the onset of tunneling an asymmetry between the  $n$ TFET and the  $p$ TFET can be seen from the device simulations. This indicates that the device simulator overestimates the  $n$ TFET current and underestimates the  $p$ TFET current. Since our calculated results are consistent, they can be considered more accurate in this region of interest.

Eq. (3.15) reveals that the on-current due to point tunneling has a strong dependence on the bandgap and dielectric thickness. A smaller bandgap will lead to a smaller onset voltage  $V_{\text{onset}} = V_{\text{FB}} + E_g$ . A high doping level is assumed in our approximations and will also be beneficial for a reduction of the onset voltage as a higher doping level will change the metal-semiconductor flatband voltage.

### 3.4 Line tunneling

Line tunneling occurs in the source region in the direction normal to the gate and will be important at high gate voltages. When applying a positive gate voltage, a depletion region will be formed and a one dimensional treatment of the potential can be performed. At higher gate voltages, an inversion layer can be formed but as the inversion layer electrons significantly occupy the conduction band, we may ignore their contribution to the BTBT generation current and evaluate the latter for the depletion region only. This treatment leads to an analytical formula for the line tunneling current [5] which is briefly repeated here.

The total line tunneling current  $I$  can be computed as the sum over all charge generated by BTBT per unit time:

$$I = \frac{qWLA}{2} \int_{l_1}^{l_2} \frac{E_G^{D-\frac{1}{2}}}{q^D l^D} e^{-Bq\sqrt{E_G}l} \left( 1 - \frac{2E_G\epsilon_s}{q^2 N_a} \frac{1}{l^2} \right) dl \quad (3.17)$$

where  $l$  denotes the tunnel path length ranging from  $l_1$  to  $l_2$  in the depletion region as shown in Fig. 3.4.  $L$  is the gate length and  $N_a$  is the doping concentration.

Assuming that the exponential terms change much more rapidly than the polynomial terms under a variation of  $l$  and substituting  $D$  by its default value of

2, we obtain closed expressions for the current  $I$  and onset voltage  $V_{\text{onset}}$  as a function of the gate voltage  $V_{\text{gs}}$ :

$$I \approx WLT e^S \sqrt{V_{\text{gs}} - V_{\text{onset}}} \sqrt{V_{\text{gs}} - V_{\text{onset}}} \quad (3.18)$$

$$V_{\text{onset}} = V_{\text{FB}} + \frac{E_{\text{G}}}{q} \left( 1 + 2t_{\text{ox}} \frac{\epsilon_{\text{s}}}{\epsilon_{\text{ox}}} \sqrt{\frac{q^2 N_{\text{a}}}{2E_{\text{G}} \epsilon_{\text{s}}}} \right) \quad (3.19)$$

with

$$T = q \frac{A}{Bq^{3/2}} \frac{qN_{\text{a}}}{2\epsilon_{\text{s}}} \sqrt{\frac{1}{E_{\text{G}} \gamma}} e^{-BqE_{\text{G}} \sqrt{2\epsilon_{\text{s}}} / \sqrt{q^2 N_{\text{a}}}}, \quad (3.20)$$

$$S = Bq \sqrt{\frac{2E_{\text{G}} \epsilon_{\text{s}}}{qN_{\text{a}}}} \frac{1}{\gamma}, \quad \gamma = 1 + t_{\text{ox}} \frac{\epsilon_{\text{s}}}{\epsilon_{\text{ox}}} \sqrt{\frac{q^2 N_{\text{a}}}{2E_{\text{G}} \epsilon_{\text{s}}}} \quad (3.21)$$

To obtain Eq. (3.17), we made the assumption that the TFET current equals the generation current in the depletion region under the source. We validate this assumption by comparing the derived formulas with Medici simulation results in Fig. 3.7. The analytical results agree very well with the non-selfconsistent calculations by Medici. For large gate voltages, the channel resistance in series with the tunnel barrier is no longer negligible and the self-consistently calculated current saturates.

Eq. (3.19) provides the line tunneling current of the TFET as a function of gate voltage when the device parameters are known. Clearly, the square root dependence indicates the absence of a 60 mV/decade subthreshold-slope. The on-current increases with decreasing bandgap and higher doping level. A small bandgap also reduces the onset voltage and will be aimed at as far as it does not jeopardize the TFET off-current. An upper limit on the onset voltage or on the voltage drop over the dielectric will limit the doping level. Finally, enabling the gate to adequately control the source region, a small dielectric thickness will improve the validity of the approximations.

## 3.5 Conclusion

We have shown an approximating analytical description of the TFET potential profile in the channel and in the region under the gate. Using this potential profile and adopting Kane's Model, we derived an approximating compact formula for point tunneling current complementing the analytical model for line tunneling. The approximations accompanying this derivation are justified by observing agreement with device simulation results in the case of line tunneling. In the case of point tunneling, the current calculated by the device simulator

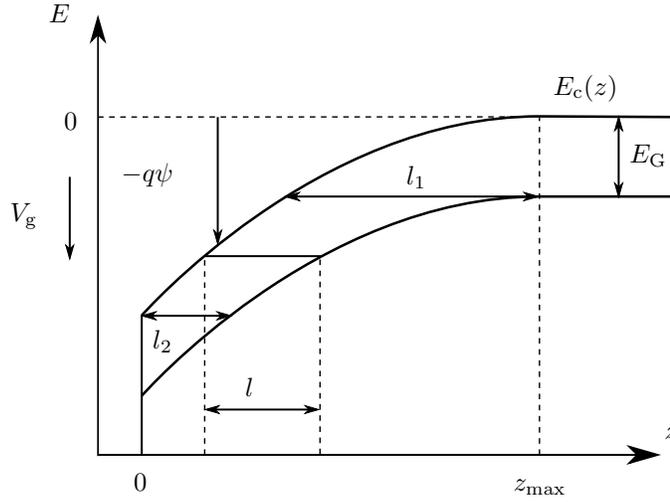


Figure 3.4: Band diagram of the cross-section A-A' from Fig. 3.2 indicating the tunneling distance  $l$ .  $\psi$  is the electrostatic potential,  $E_c$  is the conduction band energy and  $V_g$  the gate voltage. (Line tunneling)

Medici shows a discrepancy with our model due to an over- and underestimation in the case of the  $n$ - and  $p$ TFET respectively and our result can be considered more accurate.

The three key parameters for the TFET current derived from our formulas are: bandgap, dielectric thickness and source doping level. In both the case of point and line tunneling a lower bandgap will lead to an improved on-current and a lower onset voltage. A low bandgap will therefore be aimed at for a TFET as long as the deterioration of the off-current is acceptable. A reduced dielectric thickness will reduce the onset voltage of line tunneling and will significantly improve the current of point tunneling. An increased doping concentration reduces the point tunneling onset voltage while it increases the line tunneling onset voltage but at the same time the line tunneling on-current increases.

## Acknowledgements

The authors would like to thank An De Keersgieter for support with the Medici simulations. This work was supported by IMEC's Industrial Affiliation Program. Anne S. Verhulst acknowledges the support of a Marie Curie International Reintegration Grant within the 6th European Community Framework Programme, as well as a postdoctoral fellowship of the Fund for Scientific Research - Flanders.

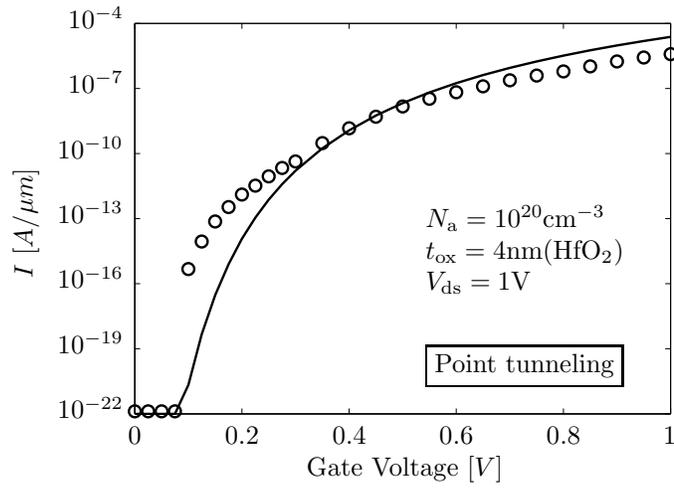


Figure 3.5: Comparison between analytical model for point tunneling (solid line) and Medici calculations ( $\circ$ ) for a silicon  $n$ TFET with a structure as shown in Fig. 2.2.

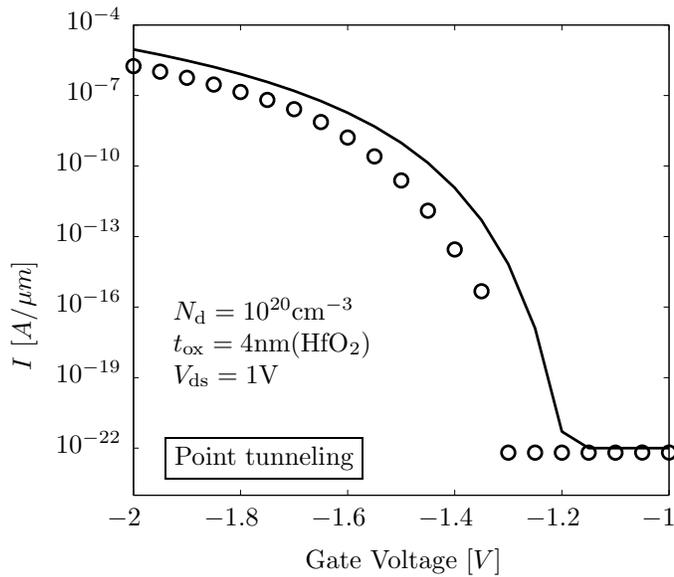


Figure 3.6: Comparison between analytical model for point tunneling (solid line) and Medici calculations ( $\circ$ ) for a silicon  $p$ TFET with a structure as shown in Fig. 2.2.

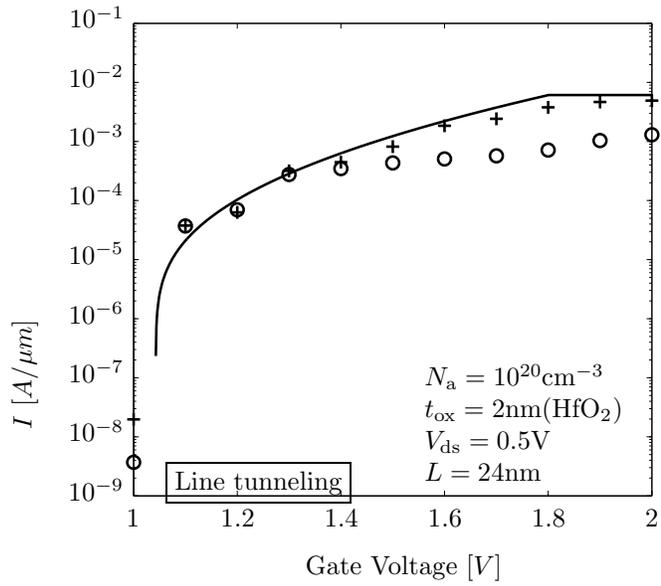


Figure 3.7: Analytical formula (3.18) (solid) with the gate voltage limited to the inversion onset voltage compared with self-consistent device simulations (○) and non-self-consistent device simulations (+) for a Si TFET with the gate only on top of the source compared with (as shown in [5]).

## Bibliography

- [1] J. Appenzeller, Y.-M. Lin, J. Knoch, Z. Chen, and P. Avouris, “Comparing carbon nanotube transistors - the ideal choice: a novel tunneling device design,” *Electron Devices, IEEE Transactions on*, vol. 52, no. 12, pp. 2568–2576, Dec. 2005.
- [2] W. Y. Choi, B.-G. Park, J. D. Lee, and T.-J. K. Liu, “Tunneling field-effect transistors (TFETs) with subthreshold swing (SS) less than 60 mv/dec,” *Electron Device Letters, IEEE*, vol. 28, pp. 743–745, Aug. 2007.
- [3] P.-F. Wang, “Complementary tunneling-fets (CTFET) in CMOS technology,” Ph.D. dissertation, Technische Universität München, 2003.
- [4] A. S. Verhulst, W. G. Vandenberghe, K. Maex, and G. Groeseneken, “A tunnel field-effect transistor without gate-drain overlap,” *Applied Physics Letters*, vol. 91, no. 053102, July 2007.
- [5] W. Vandenberghe, A. Verhulst, G. Groeseneken, B. Soree, and W. Magnus, “Analytical model for point and line tunneling in a tunnel field-effect transistor,” in *Simulation of Semiconductor Processes and Devices, 2008. SISPAD 2008. International Conference on*, sept. 2008, pp. 137–140.
- [6] *Medici*, Synopsis, March 2007.
- [7] E. O. Kane, “Zener tunneling in semiconductors,” *Journal of Physics and Chemistry of Solids*, vol. 12, pp. 181–188, 1959.



## Part III

# Quantum mechanical modeling of BTBT in direct semiconductors



## Chapter 4

# BTBT in direct semiconductors in non-uniform fields

Zener tunneling in semiconductors under  
non-uniform electric fields

William Vandenberghe, Bart Sorée, Wim Magnus and  
Guido Groeseneken

Published in *Journal of Applied Physics*, 107 (5), art. nr. 054520,  
March 2010

The following article appeared in *Journal of Applied Physics*, 107 (5), art. nr.  
054520, March 2010 and may be found at <http://link.aip.org/link/doi/10.1063/1.3311550>.

©2010 American Institute of Physics. This article may be downloaded for  
personal use only. Any other use requires prior permission of the author and  
the American Institute of Physics.

### **Contributions of first author**

- Conception and development of the theory
- Development of the code
- Creating figures
- Text writing and editing

## Abstract

Recently, a renewed interest in Zener tunneling has arisen because of its increasing impact on semiconductor device performance at nanometer dimensions. In this paper we evaluate the tunnel probability under the action of a non-uniform electric field using a two-band model and arrive at significant deviations from the commonly used Kane's model, valid for weak uniform fields only. A threshold on the junction bias where Kane's model for Zener tunneling breaks down is determined. Comparison with Kane's model particularly shows that our calculation yields a higher tunnel probability for intermediate electric fields and a lower tunnel probability for high electric fields. When performing a current calculation comparing to the WKB approximation for the case of an abrupt  $p - n$  junction significant differences concerning the shape of the I-V curve are demonstrated.

## 4.1 Introduction

Zener tunneling in semiconductors involves transitions between the valence and the conduction band induced by appreciable electric fields. It is also widely referred to as interband tunneling or Band-to-Band Tunneling (BTBT) and together with avalanche breakdown caused by impact ionization, it is considered the working principle of the so-called Zener diodes [1]. More recently, BTBT is proposed as a basic mechanism providing the drive current in tunneling field-effect transistors or tunnel FETs [2, 3]. Moreover, also the performance of various "conventional" devices, including the metal-oxide-semiconductor field-effect transistors with nanometer-sized gate lengths, can be seriously deteriorated by Zener tunneling as it may cause the occurrence of leakage currents such as the so-called Gate Induced Drain Leakage (GIDL) [4, 5]. Therefore, a systematic study of BTBT in nano-sized semiconductor structures and, particularly, its relation to the presence of strong and localized electric fields, is in order.

Over forty years ago, a theoretical investigation of Zener tunneling was provided by Kane [6, 7], who treated BTBT for direct gap semiconductors subjected to a uniform electric field. In order to obtain a compact formula he used a two-band  $\mathbf{k} \cdot \mathbf{p}$  model to quantify the band structure, while the transition probabilities were calculated from Fermi's golden rule. In spite of a few mistakes in the original derivation by Kane [8] the transmission coefficient, which can also be obtained from WKB theory [9] was found to be correct. Argyres [10] showed that there is an additional oscillatory term in the tunneling probability, omitted by Kane, which describes the effect of Stark quantization in the case of a

uniform field. In addition however, he pointed out that in the case of Zener tunneling under strongly non-uniform fields, Stark quantization does not apply, although Zener current exhibiting an oscillatory dependence on the applied field can still persist.

For indirect bandgap semiconductors and structures acted upon by non-uniform electric fields, Kane's model and the WKB approximation have led to reasonable fits with recent experimental data [11] but, in general, the corresponding class of structures and its related devices call for an improved and extended BTBT theory, accounting also for the effect of scattering mechanisms. An important example of the latter is phonon-assisted tunneling, which plays a significant role in the case of silicon and other materials with an indirect bandgap [12]. In this paper, however, we treat the problem of BTBT under non-uniform fields only for the case of direct-gap semiconductors.

First, we determine the solution to the Schrödinger equation using the envelope function approximation. Next we obtain the BTBT probability for an electron described by a plane wave entering a junction region that is governed by a local electric field, and we observe a substantial discrepancy between our results and Kane's model. Finally we calculate the tunnel current as a function of the applied bias for a degenerate as well as a non-degenerate Esaki diode.

## 4.2 Envelope function approximation

In order to evaluate the tunneling probability under an applied external potential which is incommensurable with the lattice periodicity, the one-electron Schrödinger equation

$$\left[ -\frac{\hbar^2 \nabla^2}{2m} + U_{\text{lat}}(\mathbf{r}) + U_{\text{ext}}(\mathbf{r}) \right] \psi = E\psi \quad (4.1)$$

must be solved. Following the lines of conventional  $\mathbf{k}\cdot\mathbf{p}$  theory [13], the solution to Eq. (4.1) is expanded in a complete, orthogonal set of lattice-periodic basis functions  $u_n(\mathbf{r})$ ,

$$\psi(\mathbf{r}) = \sum_n f_n(\mathbf{r}) u_n(\mathbf{r}) \quad (4.2)$$

where the expansion coefficients  $f_n(\mathbf{r})$  are envelope functions varying slowly on the scale of a Bravais lattice [14, 15] unit cell and  $n$  is the band index.

For the sake of convenience, the basis functions  $u_n(\mathbf{r})$  are taken to be solutions of the bulk Schrödinger equation corresponding to a particular symmetry point in  $\mathbf{k}$ -space, e.g. the  $\Gamma$ -point ( $\mathbf{k} = \mathbf{0}$ ). Denoting the corresponding energy eigenvalues by  $E_n^0$ , substituting the expansion (4.2) into the one-electron Schrödinger equation and assuming  $U_{\text{ext}}(\mathbf{r})f_n(\mathbf{r})$  only has components within

the first Brillouin zone, we obtain a set of eigenvalue equations for  $f_n(\mathbf{r})$  :

$$-\frac{\hbar^2}{2m}\nabla^2 f_n(\mathbf{r}) - \frac{i\hbar}{m} \sum_{n'} \mathbf{p}_{nn'} \cdot \nabla f_{n'}(\mathbf{r}) + E_n^0 f_n(\mathbf{r}) + U_{\text{ext}}(\mathbf{r}) f_n(\mathbf{r}) = E f_n(\mathbf{r}) \quad (4.3)$$

where  $m$  and  $\mathbf{p}_{nn'}$  denote the free electron mass and the momentum matrix elements respectively with

$$\mathbf{p}_{nn'} = -i\hbar \int_{\text{unit cell}} d^3r u_n^*(\mathbf{r}) \nabla u_{n'}(\mathbf{r}). \quad (4.4)$$

As Eq. (4.1) has real coefficients,  $u_n(\mathbf{r})$  can be taken real and the diagonal momentum matrix elements  $p_{nn}$  are zero. When complex  $u_n(\mathbf{r})$  are taken the diagonal momentum matrix elements remain zero for a crystal with inversion symmetry.

In practice, the hierarchy of equations (4.3) is truncated and their accuracy is determined by the number of remaining equations  $N$  and the spatial profile of the external perturbation  $U_{\text{ext}}(\mathbf{r})$  including the electrostatic potential energy due to the charge distribution and the externally applied voltages. On one hand,  $U_{\text{ext}}(\mathbf{r})$  has to change rapidly on the length scale of the active device area in order to induce measurable BTBT effects, on the other hand it should vary moderately on the atomic scale for equations (4.3) to hold.

The current carried by a given eigenfunction  $\psi(\mathbf{r})$  is obtained from

$$\mathbf{J}(\mathbf{r}) = \frac{iq\hbar}{2m} (\psi^* \nabla \psi - \psi \nabla \psi^*) \quad (4.5)$$

and within the  $\mathbf{k}\cdot\mathbf{p}$  envelope function formalism, the current density can be rewritten as [14]

$$\mathbf{J}(\mathbf{r}) = \frac{q}{m} \left[ \text{Im} \left( -\hbar \sum_{n=1}^N f_n^*(\mathbf{r}) \nabla f_n(\mathbf{r}) \right) + \sum_{n,n'=1}^N f_n^*(\mathbf{r}) \mathbf{p}_{nn'} f_{n'}(\mathbf{r}) \right]. \quad (4.6)$$

### 4.3 Two band model

From hereof, we focus on the case where only two bands are relevant and the interband coupling is restricted to the  $z$ -direction, i.e.  $N = 2$  and  $\mathbf{p}_{nn'} = (1 - \delta_{nn'}) (0, 0, p)$ . The momentum matrix element  $p$  is available as a real parameter for a large set of materials [16], through the quantity

$$E_P = \frac{2p^2}{m}. \quad (4.7)$$

Since the structure is assumed to be translationally invariant in the  $x$  and  $y$  direction, the envelope functions can be factorized as

$$\begin{aligned} f_1(\mathbf{r}) &= \chi_1(z)e^{i\mathbf{K}\cdot\mathbf{R}} \\ f_2(\mathbf{r}) &= \chi_2(z)e^{i\mathbf{K}\cdot\mathbf{R}} \end{aligned} \quad (4.8)$$

where  $\mathbf{K}$  and  $\mathbf{R}$  are two-dimensional vectors in the  $(x, y)$  plane. The hierarchy of equations (4.3) can be written as

$$\begin{bmatrix} U_{\text{ext}} - E + \frac{E_g}{2} - \frac{\hbar^2}{2m} \frac{d^2}{dz^2} & -\frac{i\hbar p}{m} \frac{d}{dz} \\ -\frac{i\hbar p}{m} \frac{d}{dz} & U_{\text{ext}} - E - \frac{E_g}{2} - \frac{\hbar^2}{2m} \frac{d^2}{dz^2} \end{bmatrix} \begin{bmatrix} \chi_1 \\ \chi_2 \end{bmatrix} = 0 \quad (4.9)$$

for  $\mathbf{K} = 0$  and the corresponding eigensolutions describe a continuous energy spectrum. For the sake of brevity, the  $z$  dependence of  $U_{\text{ext}}(z)$ ,  $\chi_1(z)$  and  $\chi_2(z)$  is omitted. The case where  $\mathbf{K} \neq 0$  will be of importance in the last section where a device current is calculated and a pragmatic approach where  $U_{\text{ext}}$ ,  $E_g$  and  $p$  have a  $\mathbf{K}$  dependence is used. In this way we can proceed with Eq. (4.9).

## 4.4 Uniform potential profile

For a uniform potential energy profile

$$U_{\text{ext}}(z) = U_{\text{ext}} \quad (4.10)$$

it turns out that Eq. (4.9) has solutions of the form:

$$\begin{bmatrix} \chi_1(z) \\ \chi_2(z) \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} e^{ikz} \quad (4.11)$$

while the corresponding eigenvalue equation for  $E' = E - U_{\text{ext}}$  reads

$$\left( \frac{\hbar^2 k^2}{2m} - E' \right)^2 - \left( \frac{\hbar k p}{m} \right)^2 - \left( \frac{E_g}{2} \right)^2 = 0. \quad (4.12)$$

As expected, a non-parabolic dispersion relation emerges,

$$E' = \frac{\hbar^2 k^2}{2m} \pm \sqrt{\left( \frac{\hbar k p}{m} \right)^2 + \left( \frac{E_g}{2} \right)^2}, \quad (4.13)$$

as was already derived by Kane [6]. For a given eigenenergy ( $E$ ), two physical and two spurious solutions can be identified for the longitudinal wave number  $k$ , as illustrated in Fig. 4.1:

$$\frac{\hbar^2 k^2}{2m} = E' + \frac{p^2}{m} - \sqrt{\left( E' + \frac{p^2}{m} \right)^2 - \left( E'^2 - \left( \frac{E_g}{2} \right)^2 \right)} \quad (4.14)$$

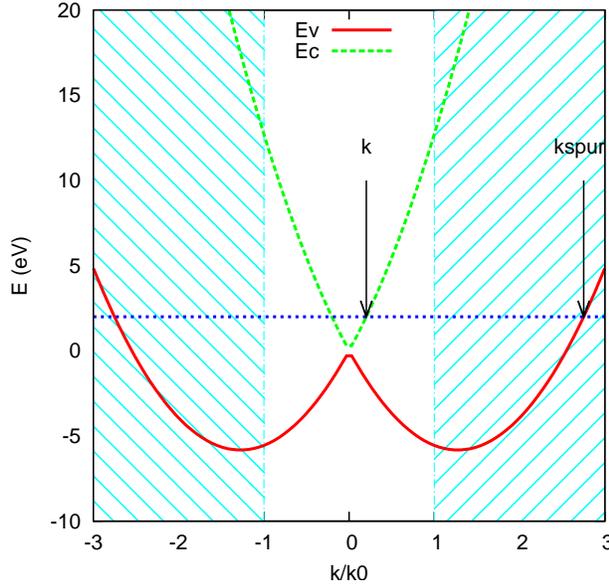


Figure 4.1: Band structure illustrating spurious solution outside the Brillouin zone for a given eigenenergy.  $k_0 = 2\pi/a$  and  $a$  is the lattice constant.

$$\frac{\hbar^2 k_{\text{spur}}^2}{2m} = E' + \frac{p^2}{m} + \sqrt{\left(E' + \frac{p^2}{m}\right)^2 - \left(E' - \left(\frac{E_g}{2}\right)^2\right)}. \quad (4.15)$$

Spurious solutions located outside of the Brillouin zone typically appear when only a finite number of basis functions is taken into account. As the corresponding wave functions clearly do not correspond to physical states, they can be eliminated provided  $|E'| \ll p^2/m$  and  $E_g/2 \ll p^2/m$  where the latter condition is fulfilled in most common semiconductors. When  $|E'|$  is very large the 2 band approximation inevitably breaks down and even both solution  $k$  and  $k_{\text{spur}}$  can be located outside the Brillouin zone and not have physical significance.

Under the assumption that the spurious solutions can be eliminated, the two physical solutions correspond to a right and a left running wave when  $|E'| > E_g/2$ :

$$\begin{bmatrix} \chi_{1,\text{cst}}^\pm(z) \\ \chi_{2,\text{cst}}^\pm(z) \end{bmatrix} = \begin{bmatrix} \sqrt{E' - (\hbar^2 k^2)/(2m) + E_g/2} \\ \pm \sqrt{E' - (\hbar^2 k^2)/(2m) - E_g/2} \end{bmatrix} e^{\pm ikz} \quad (4.16)$$

with  $k$  the positive root of Eq. (4.14).  $\chi_{1,2,\text{cst}}$  are not normalized and the ratio  $(\chi_{2,\text{cst}}/\chi_{1,\text{cst}})^2$  equals the ratio of the diagonal elements of the symmetric matrix (4.9) whereas the sign can be determined by inspection. The current carried by these two states is

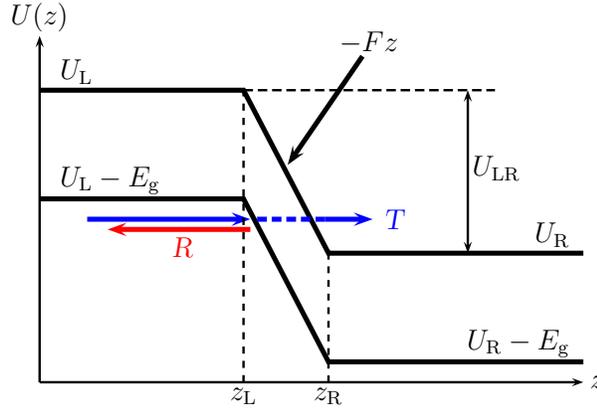


Figure 4.2: Potential energy of a piecewise uniform field. A contact is present on the left and right while the potential has a constant slope in the junction.

$$J(E') \propto \pm \frac{q}{m} \left( \hbar k \left( E' - \frac{\hbar^2 k^2}{2m} \right) + p \sqrt{\left( E' - \frac{\hbar^2 k^2}{2m} \right)^2 - \left( \frac{E_g}{2} \right)^2} \right). \quad (4.17)$$

## 4.5 Piecewise uniform field

In this section we determine the transmission coefficient by solving the envelope function equation for a semiconductor structure consisting of a left and right contact with a junction in between. A state composed of an incident, reflected and transmitted component in the contacts can be constructed. Deriving the corresponding incident and reflected current densities  $J_I$  and  $J_R$  we determine the probabilities for electron transmission ( $T$ ) and reflection ( $R$ ) by

$$T = |J_T/J_I|, \quad R = |J_R/J_I|, \quad (4.18)$$

with  $T + R = 1$ .

In order to proceed with a numerical calculation, we transform Eq. (4.3) to a set of 4 differential equations where we make the choice  $\mathbf{p}_{nn'} = (-1)^n (1 - \delta_{nn'}) (ip, 0, 0)$  such that  $p_{12} = -p_{21}$  is imaginary and the coefficients are real:

$$\frac{d}{dz} \begin{bmatrix} \chi_1(z) \\ \chi_1'(z) \\ \chi_2(z) \\ \chi_2'(z) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \left( U_{\text{ext}}(\mathbf{r}) + \frac{E_g}{2} - E \right) \frac{2m}{\hbar^2} & 0 & 0 & -\frac{2p}{\hbar} \\ 0 & 0 & 0 & 1 \\ 0 & \frac{2p}{\hbar} & \left( U_{\text{ext}}(\mathbf{r}) - \frac{E_g}{2} - E \right) \frac{2m}{\hbar^2} & 0 \end{bmatrix} \begin{bmatrix} \chi_1(z) \\ \chi_1'(z) \\ \chi_2(z) \\ \chi_2'(z) \end{bmatrix}. \quad (4.19)$$

To evaluate the transmission coefficient, an initial condition corresponding to a positive running wave is imposed in the right contact. The solution is integrated through the junction according to Eq. (4.19) up to the left contact. In the left contact ( $z \leq z_L$ ) the solution is decomposed in a positive and negative running wave and two spurious waves:

$$\chi_{1,2} = c_L^+ \chi_{1,2}^+ + c_L^- \chi_{1,2}^- + c_L^{\text{spur}+} \chi_{1,2}^{\text{spur}+} + c_L^{\text{spur}-} \chi_{1,2}^{\text{spur}-} \quad (4.20)$$

where for the sake of brevity, the  $z$  dependence of the envelope functions was omitted. The spurious contributions are discarded which is justified as long as  $|E_{L,R}| \ll p^2/m$  with

$$\begin{aligned} E_L &= E - U_{\text{ext}}(z) & : z < z_L \\ E_R &= E - U_{\text{ext}}(z) & : z > z_R \end{aligned} \quad (4.21)$$

The transmission coefficient is then determined by Eq. (4.18) or more specifically

$$T = (c_R^+)^2 J(E_R) / ((c_L^+)^2 J(E_L)) \quad (4.22)$$

where  $J(E)$  is proportional to the current carried by a positive running wave and given by Eq. (4.17) and  $c_{R,L}^+$  is the amplitude of the positive running wave in the right and left contact respectively. The result of this calculation for a piecewise uniform field (Fig. 4.2) is shown in Fig. 4.3.

In the case of a piecewise uniform field as illustrated in Fig. 4.2, straightforward comparison with Kane's model

$$T_{\text{Kane}} = \frac{\pi^2}{9} \exp \left( -\frac{\pi E_g^2 m}{4\hbar p F} \right) \quad (4.23)$$

which is derived for two bands and a uniform field is possible. More generally, we may compare with the transmission coefficient determined by the WKB approximation

$$T_{\text{WKB}} = \exp \left( -2i \int k(z) dz \right). \quad (4.24)$$

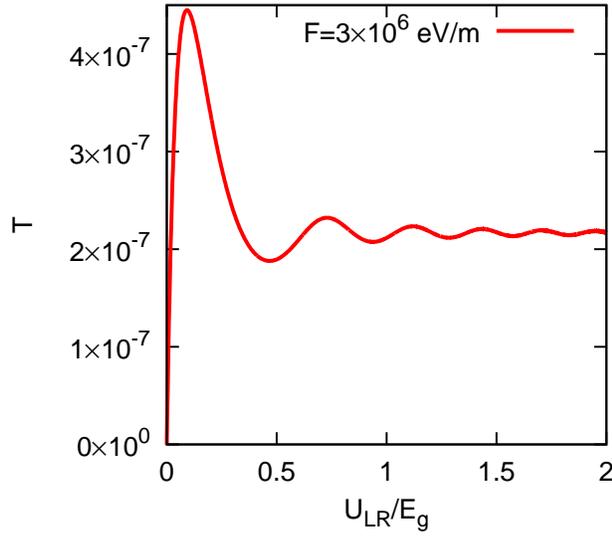


Figure 4.3: Numerical calculation of the transmission probability in InSb for fixed  $F$  and  $E = 0$  while changing  $U_{LR}$  as indicated in Fig. 4.2.

where  $k(z)$  is the imaginary wave vector corresponding to a modified eigenenergy  $E'(z) = E - U_{\text{ext}}(z)$  within the bandgap and can be determined from Eq. (4.14).

In the Appendix, the limiting value and the oscillatory behavior is estimated based on an approximate analytical calculation of the transmission coefficient. Fig. 4.4 shows the transmission coefficient calculated numerically for a piecewise uniform potential versus

$$\tau = -\frac{\log(T_\infty)}{2\pi^2} \left( \left( \frac{U_{LR}}{E_g} \right)^2 - 1 \right). \quad (4.25)$$

It can be seen that the transmission coefficient tends to the asymptotic value

$$T_\infty = \exp \left( -\frac{\pi E_g^2 m}{4\hbar p F} \right) = 1.1 T_{\text{Kane}} \quad (4.26)$$

for

$$U_{LR} > E_g \sqrt{1 - \frac{\pi^2}{\log(T_\infty)}}. \quad (4.27)$$

Eq. (4.27) enables us to assess the validity of the Kane formula for the applied electric fields shown in Fig. 4.5. For small fields ( $|\log(T_\infty)| \gg 0$ ) and the application of a not too small  $U_{LR}$ , the Kane formula proves to be a good

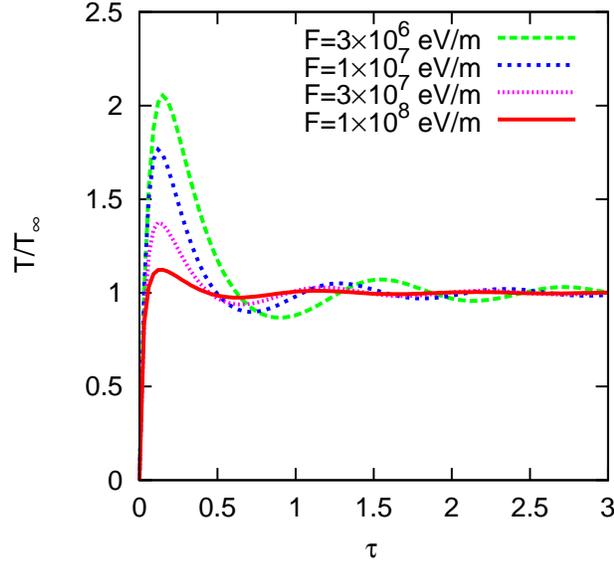


Figure 4.4: Numerical calculation of the transmission probability in InSb (for  $E = 0$ ) relative to the limiting value ( $T_\infty$ ) where each curve is scaled versus  $\tau$  on the horizontal axis. The figure demonstrates that the oscillation frequency roughly equals unity and the Kane transmission probability is approximately valid when  $\tau > 1/2$  or  $U_{LR}$  satisfies Eq. (4.27).

approximation. For intermediate electric fields, the peak in the transmission coefficient as shown in Fig. 4.4 results in an increased tunneling efficiency. For high electric fields and the associated efficient tunneling ( $T_\infty \approx 1$ ), even the application of a significant potential difference between the contacts will not suffice to reach the Kane limit and the current is smaller than predicted by Kane's model.

## 4.6 Current

The tunneling current for a semiconductor with a left and right contact equals [17]:

$$J = \frac{2e}{h} \int_{-\infty}^{\infty} dE (f_R(E) - f_L(E)) \int \frac{d^2 k_\perp}{(2\pi)^2} T(E, \mathbf{K}) \quad (4.28)$$

where  $f_{L,R}(E)$  represents the occupation probability in the left and right contact and  $\mathbf{K}$  is the crystal momentum related to the remaining lattice symmetry in the direction perpendicular to tunneling as introduced in Eq. (4.8).

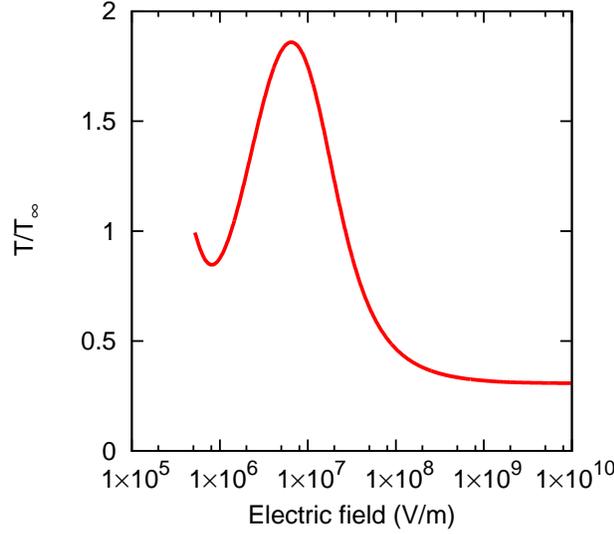


Figure 4.5: Calculation of the transmission probability in InSb compared with the transmission probability calculated by Kane as a function of the electric field for a constant potential difference between the contacts ( $U_{L,R}$ ) and  $E = 0$ .

Since we have chosen the valence and conduction band such that the matrix element is in the direction of transport, there is no interaction between the two basis states in the perpendicular direction. We could employ the two band  $\mathbf{k}\cdot\mathbf{p}$  formalism to describe the conduction band structure in the perpendicular direction if a degenerate valence band state is present but we limit ourselves to an isotropic effective mass approximation for the valence and conduction band where  $m_{v\perp}^*$  and  $m_{c\perp}^*$  are the respective effective masses.

The effect of the perpendicular crystal momentum is taken into account by a modified bandgap

$$E_g(\mathbf{K}) = E_g + \hbar^2|\mathbf{K}|^2/(2m_{v\perp}^*) + \hbar^2|\mathbf{K}|^2/(2m_{c\perp}^*). \quad (4.29)$$

The interaction matrix element ( $p$ ) will only change due to interaction with higher/lower bands which we have chosen to ignore in the current  $\mathbf{k}\cdot\mathbf{p}$  treatment, the matrix element therefore remains unchanged for any  $\mathbf{K}$ :  $p(\mathbf{K}) = p$ .

The occupation probabilities are taken to be the equilibrium Fermi-Dirac distributions in both contacts:

$$f_{L,R}(E) = \frac{1}{1 + \exp((E - E_{FL,R})/kT)} \quad (4.30)$$

where  $E_{FL}$  and  $E_{FR}$  are the equilibrium Fermi level in the left and right contact respectively.

### 4.6.1 Example: abrupt p-n diode

We evaluate the current numerically in the case of an abrupt p-n diode. Using the Shockley approximation, the potential energy is described by:

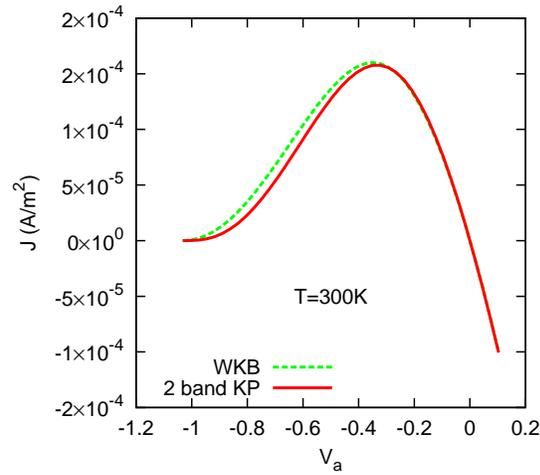
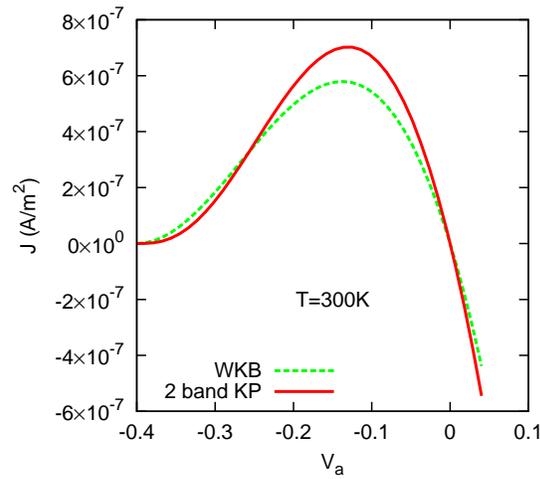
$$U_{\text{ext}}(z) = \begin{cases} -\frac{q^2 N_d}{2\epsilon_s} z_L^2 & : z < z_L \\ \frac{q^2 N_d}{2\epsilon_s} z (z - 2z_L) & : z_L \leq z \leq 0 \\ -\frac{q^2 N_a}{2\epsilon_s} z (z - 2z_R) & : 0 \leq z \leq z_R \\ \frac{q^2 N_a}{2\epsilon_s} z_R^2 & : z_R < z \end{cases} \quad (4.31)$$

and  $z_L$  and  $z_R$  are determined by the condition  $U_{LR} = U_{\text{ext}}(z_L) - U_{\text{ext}}(z_R) = q(V_{bi} + V_a)$ . In the case of a diode, a large field implies a large doping concentration and large doping concentration is associated with degeneracy ( $qV_{bi} > E_g$ ). A diode with a degenerately doped  $p$  and  $n$  contact is an Esaki diode [18] which is known to show a negative differential resistance at forward bias. The latter due to a reduction in the number of states that can take part in the tunneling with increasing applied voltage.

In Fig. 4.6 the result from the previously outlined method of calculation is shown and compared with a WKB approach. The main difference between the two calculations is the shape of the calculated current-voltage characteristic, certainly at the onset of tunneling. Concerning the magnitude of the current, the same trends observed for the transmission probability are present for the calculated device current: a higher current than predicted by the WKB approximation appears for intermediate doping levels (Fig. 4.6b and Fig. 4.6c) while a decreased current is observed for very high doping levels (Fig. 4.6a).

## 4.7 Conclusions

For large applied biases or low fields the Kane formula or the WKB approximation provide a good approximation to the actual tunnel probability. When larger fields or very small biases are applied and qualitative results must be obtained or the shape of the current-voltage characteristic is of importance, the Kane formula or WKB approximation breaks down. For intermediate fields an increased transmission probability can be expected while at even higher fields the transmission is suppressed. For the case of a piecewise uniform field an analytical expression based on the parabolic cylinder function can be used but a numerical evaluation of the tunnel probability is more convenient for more general potential profiles that need to be determined self-consistently.

(a)  $p$  and  $n$ -type doping levels of  $5 \cdot 10^{18} \text{cm}^{-3}$ (b)  $p$  and  $n$ -type doping levels of  $5 \cdot 10^{17} \text{cm}^{-3}$  and  $5 \cdot 10^{18} \text{cm}^{-3}$

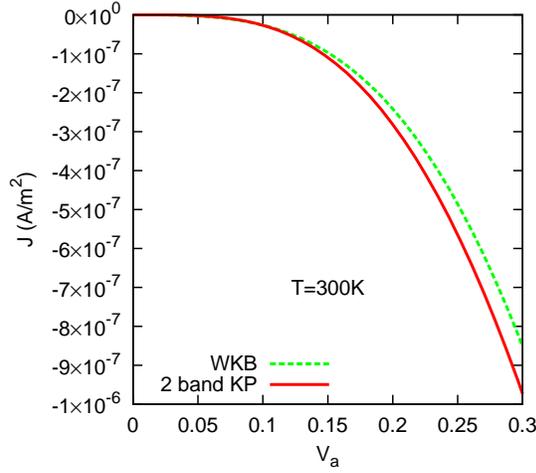
(c)  $p$  and  $n$ -type doping levels of  $5 \cdot 10^{17} \text{cm}^{-3}$ 

Figure 4.6: Calculation of the current in an InSb diode at  $T = 300\text{K}$  compared to a WKB calculation. Our calculation shows a different shape and a current reduction/increase compared to the WKB result.

## Appendix

Considering a potential energy profile as shown in Fig. 4.2

$$U_{\text{ext}}(z) = \begin{cases} -Fz_L & : z < z_L \\ -Fz & : z_L \leq z \leq z_R \\ -Fz_R & : z_R < z \end{cases} \quad (4.32)$$

it is possible to obtain an analytical expression for the wavefunction if we omit the second derivatives in Eq. (4.9). A justification for omitting the second derivatives is given by expanding Eq. (4.13) for small  $k$

$$E' = \pm E_g + \left(1 \pm \frac{2p^2}{E_g m}\right) \frac{\hbar^2 k^2}{2m}. \quad (4.33)$$

The first term within the brackets (1), corresponding to the second derivatives can be omitted for the case  $E_g/2 \ll p^2/m$  which is generally satisfied [16].

First, we determine the solution in the junction where the potential energy profile is

$$U_{\text{ext}}(z) = -Fz = U_{\text{lin}}(z). \quad (4.34)$$

To obtain analytical solutions, we perform a linear transformation on the original envelope functions, yielding sum and difference envelope functions:

$$\chi_s(z) = \frac{1}{\sqrt{2}}(\chi_1(z) + \chi_2(z)) \quad (4.35)$$

$$\chi_d(z) = \frac{1}{\sqrt{2}}(\chi_1(z) - \chi_2(z)). \quad (4.36)$$

Rewriting Eq. (4.9) in this new basis and omitting the second derivatives, we arrive at

$$\begin{bmatrix} U_{\text{ext}} - E - \frac{i\hbar p}{m} \frac{d}{dz} & \frac{E_g}{2} \\ \frac{E_g}{2} & U_{\text{ext}} - E + \frac{i\hbar p}{m} \frac{d}{dz} \end{bmatrix} \begin{bmatrix} \chi_s \\ \chi_d \end{bmatrix} = 0, \quad (4.37)$$

with the  $\mathbf{k} \cdot \mathbf{p}$  term now in the main diagonal such that they can be transformed into Eqs. 19.6.1-2 from Abramowitz and Stegun [19]

$$\begin{bmatrix} \frac{d}{dz} + \frac{1}{2}z & a + \frac{1}{2} \\ 1 & \frac{d}{dz} - \frac{1}{2}z \end{bmatrix} \begin{bmatrix} U(a, z) \\ U(a+1, z) \end{bmatrix} = 0 \quad (4.38)$$

where  $U(a, z)$  is the parabolic cylinder function.

As a result the sum and difference envelope functions may be expressed as:

$$\begin{aligned} \chi_{s,\text{lin}}(z) &= -U\left(i\frac{E_g^2 m}{8p\hbar F} - \frac{1}{2}, -\sqrt{\frac{2mF}{i\hbar}} \left(z + \frac{E}{F}\right)\right) \\ \chi_{d,\text{lin}}(z) &= \sqrt{\frac{i8p\hbar F}{E_g^2 m}} U\left(i\frac{E_g^2 m}{8p\hbar F} + \frac{1}{2}, -\sqrt{\frac{2mF}{i\hbar}} \left(z + \frac{E}{F}\right)\right). \end{aligned} \quad (4.39)$$

The simplified solutions in the junction area were determined by omitting the second derivative. The solutions in the contacts which are characterized by a uniform potential must be determined under this same approximation. It turns out that the omission of the second derivative effectively eliminates the spurious solutions ( $k_{\text{spur}}$ ) and the two remaining solutions are:

$$\begin{bmatrix} \chi_{s,\text{cst}}^\pm(z) \\ \chi_{d,\text{cst}}^\pm(z) \end{bmatrix} = \begin{bmatrix} \sqrt{E' + E_g/2} \pm \sqrt{E' - E_g/2} \\ \sqrt{E' + E_g/2} \mp \sqrt{E' - E_g/2} \end{bmatrix} e^{\pm ikz} \quad (4.40)$$

with

$$k = \frac{m}{\hbar p} \sqrt{E'^2 - (E_g/2)^2} \quad (4.41)$$

and the current carried by these two states is

$$J(E') \propto \pm(qp/m) \sqrt{E'^2 - (E_g/2)^2}. \quad (4.42)$$

Solutions (4.39) and (4.40) enable the calculation of the transmission coefficient but are not directly useful since the parabolic cylinder functions do not allow for a more efficient computation than the numerical method outlined in the body of this article. Nevertheless, we can use the analytical solution to investigate the asymptotic behavior of  $T$  for  $U_{\text{LR}} \rightarrow \infty$  and fixed  $F$ . Selecting  $E = 0$  such

that  $E_L = -E_R = U_{LR}/2$  and  $-z_L = z_R \rightarrow \infty$ , the asymptotic behavior of the envelope equations in the middle section is governed by [19]

$$U(a, x) \asymp x^{-a-1/2} e^{-x^2/4}. \quad (4.43)$$

Using the parameters and arguments for our envelope functions given by Eq. (4.39) and only considering the amplitude, it can be seen that

$$|\chi_{s,\text{lin}}(z_R)| \propto 1 \quad , \quad |\chi_{d,\text{lin}}(z_R)| \propto 0 \quad (4.44)$$

matches with a positive running wave in the right contact since for  $|E'| = |U_{LR}|/2 \rightarrow \infty$ , Eq. (4.40) shows

$$|\chi_{s,\text{cst}}^+| \propto 1 \quad , \quad |\chi_{d,\text{cst}}^+| \propto 0. \quad (4.45)$$

From Eq. (4.42) it follows that  $J(E_L) = J(E_R)$ , the transmission coefficient given by Eq. (4.22) therefore becomes

$$T = |\chi_{s,\text{lin}}(z_L)^2 / \chi_{s,\text{lin}}(z_R)^2|. \quad (4.46)$$

To calculate  $\chi_{s,\text{lin}}(z_L) = \chi_{s,\text{lin}}(-z_R)$ , we use Eq. 19.4.6 from [19]

$$U(a, -x) = -e^{-i\pi(a-\frac{1}{2})} U(a, x) - \frac{\sqrt{2\pi} e^{-i\pi(\frac{a}{2}-\frac{1}{4})}}{\Gamma(a+1/2)} U(-a, ix) \quad (4.47)$$

and for the calculation of  $\chi_{s,\text{lin}}(-z_R)$ , the second term vanishes for large  $z_R$  due to Eq. (4.43) where the parameter ( $a$ ) and the argument ( $x$ ) are determined from Eq. (4.39). The limit for the transmission coefficient is

$$T_\infty = \exp\left(-\frac{\pi E_g^2 m}{4\hbar p |F|}\right) \quad (4.48)$$

which equals the Kane transmission coefficient apart from a factor  $\pi^2/9$ .

We can also use the analytical solution to investigate the oscillation of the transmission coefficient with respect to the potential energy drop over the junction ( $U_{LR}$ ) for fixed  $F$  and  $E = 0$ . The period of this oscillation is related to the phase of the envelope functions. We can estimate this behavior by considering the asymptotic behavior of the phase for  $z \rightarrow \infty$

$$\arg(\chi_{s,\text{lin}}(z)) \approx \frac{mF}{2\hbar p} z^2 + \frac{E_g^2 m}{16\hbar p F} \log\left(\frac{2mF}{\hbar p} z^2\right) - \frac{\pi}{2}. \quad (4.49)$$

To obtain an order of magnitude of the oscillation frequency, we consider only the first term and for fixed  $F$  we expect an oscillation of the transmission coefficient as a function of

$$\tau = -\frac{\log(T_\infty)}{2\pi^2} \left( \left( \frac{U_{LR}}{E_g} \right)^2 - 1 \right) \quad (4.50)$$

where we used Eq. (4.48) and  $z = U_{\text{LR}}/(2F)$ . This expression not only determines the period of oscillation but, for  $\tau \ll 1$ , Kane's formula will fail and a more rigorous treatment like the above presented numerical calculation is necessary.

A last analytical result is given by the high-field limit for the transmission coefficient. For  $F \rightarrow \infty$  and  $U_{\text{LR}}$  fixed, the junction region becomes negligibly small ( $z_{\text{L}} \rightarrow 0$  and  $z_{\text{R}} \rightarrow 0$ ) and we can discard the evolution of the wavefunction in the junction:  $\chi_{\text{s,d}}(z_{\text{L}}) = \chi_{\text{s,d}}(z_{\text{R}})$ . Matching a right running wave in the right contact with a linear combination of a left and right running wave in the left contact

$$\begin{pmatrix} \chi_{\text{s,L}}^+ & \chi_{\text{s,L}}^- \\ \chi_{\text{d,L}}^+ & \chi_{\text{d,L}}^- \end{pmatrix} \begin{pmatrix} c_{\text{L}}^+ \\ c_{\text{L}}^- \end{pmatrix} = c_{\text{R}}^+ \begin{pmatrix} \chi_{\text{s,R}}^+ \\ \chi_{\text{d,R}}^+ \end{pmatrix}, \quad (4.51)$$

combining Eq. (4.51) with Eqs. (4.22), (4.42) and (4.40) where  $E' = -Fz_{\text{L}} = E_{\text{L}}$  and  $E' = -Fz_{\text{R}} = E_{\text{R}}$  for  $\chi_{\text{s,L}}^{\pm}$  and  $\chi_{\text{s,R}}^{\pm}$  respectively, we obtain

$$T = 2 \left( 1 - \frac{\left( E_{\text{L}} E_{\text{R}} - \left( \frac{E_{\text{g}}}{2} \right)^2 \right)}{\sqrt{\left( E_{\text{L}}^2 - \left( \frac{E_{\text{g}}}{2} \right)^2 \right) \left( E_{\text{R}}^2 - \left( \frac{E_{\text{g}}}{2} \right)^2 \right)}} \right)^{-1} \quad (4.52)$$

$$\approx 2 \sqrt{\frac{2E_{\text{L}}}{E_{\text{g}}} - 1} \sqrt{\frac{2E_{\text{R}}}{E_{\text{g}}} - 1}. \quad (4.53)$$

## Bibliography

- [1] C. Zener, "A theory of the electrical breakdown of solid dielectrics," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 145, no. 855, pp. 523–529, July 1934.
- [2] W. M. Reddick and G. A. J. Amaratunga, "Silicon surface tunnel transistor," *Applied Physics Letters*, vol. 67, no. 4, pp. 494–496, July 1995.
- [3] J. Appenzeller, Y.-M. Lin, J. Knoch, Z. Chen, and P. Avouris, "Comparing carbon nanotube transistors - the ideal choice: a novel tunneling device design," *Electron Devices, IEEE Transactions on*, vol. 52, no. 12, pp. 2568–2576, Dec. 2005.
- [4] T. Chan, J. Chen, P. Ko, and C. Hu, "The impact of gate-induced drain leakage current on MOSFET scaling," *Electron Devices Meeting, 1987 International*, vol. 33, pp. 718–721, 1987.

- [5] T. Lim and Y. Kim, "Effect of band-to-band tunnelling leakage on 28 nm mosfet design," *Electronics Letters*, vol. 44, no. 2, pp. 157–158, 17 2008.
- [6] E. O. Kane, "Zener tunneling in semiconductors," *Journal of Physics and Chemistry of Solids*, vol. 12, pp. 181–188, 1959.
- [7] —, "Theory of tunneling," *Journal of Applied Physics*, vol. 32, no. 1, pp. 83–91, 1961.
- [8] M. Takayanagi and S. Iwabuchi, "Theory of band-to-band tunneling under nonuniform electric fields for subbreakdown leakage currents," *Electron Devices, IEEE Transactions on*, vol. 38, no. 6, pp. 1425–1431, 1991.
- [9] S. Sze, *Physics of Semiconductor Devices*. 2nd ed. John Wiley & Sons, 1981.
- [10] P. N. Argyres, "Theory of tunneling and its dependence on a longitudinal magnetic field," *Phys. Rev.*, vol. 126, no. 4, pp. 1386–1393, May 1962.
- [11] Q. Zhang, W. Zhao, and A. Seabaugh, "Low-subthreshold-swing tunnel transistors," *Electron Device Letters, IEEE*, vol. 27, no. 4, pp. 297–300, 2006.
- [12] A. Schenk, "Rigorous theory and simplified model of the band-to-band tunneling in silicon," *Solid-State Electronics*, vol. 36, no. 1, p. 1934, 1993.
- [13] M. Cardona and F. H. Pollak, "Energy-band structure of germanium and silicon: The  $\mathbf{k}\cdot\mathbf{p}$  method," *Physical Review B*, vol. 142, no. 2, pp. 530–543, 1966.
- [14] M. G. Burt, "An exact formulation of the envelope function method for the determination of electronic states in semiconductor microstructures," *Semiconductor Science and Technology*, vol. 2, no. 7, p. 460, 1987.
- [15] —, "The justification for applying the effective-mass approximation to microstructures," *Journal of Physics: Condensed Matter*, vol. 4, no. 32, p. 6651, 1992.
- [16] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, "Band parameters for iii–v compound semiconductors and their alloys," *Journal of Applied Physics*, vol. 89, no. 11, pp. 5815–5875, 2001. [Online]. Available: <http://link.aip.org/link/?JAP/89/5815/1>
- [17] C. Duke, *Tunneling in solids*, ser. Solid State Physics, Supp. 10. Academic Press, 1969.
- [18] L. Esaki, "New phenomenon in narrow germanium  $p-n$  junctions," *Phys. Rev.*, vol. 109, no. 2, pp. 603–604, Jan 1958.

- [19] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, ninth dover printing, tenth gpo printing ed. New York: Dover, 1964.

## Chapter 5

# BTBT in heterostructures

An Envelope Function Formalism for  
Heterostructure Semiconductors

Maarten Van de Put, William G. Vandenberghe, Bart Sorée and  
Guido Groeseneken

Publishable summary of master thesis 2011-2012

### Contributions of second author

- Supervising the work
- Co-conception and co-development of the theory
- Text correcting

## Abstract

We develop an envelope function formalism capable of continuously describing electron behaviour in heterostructures, including interband transitions. For this purpose we define transformations between the different zone-centre bulk Hamiltonian solutions belonging to separate layers. The transformation matrix elements are obtained via  $\mathbf{k}\cdot\mathbf{p}$  interband momentum matrix elements. Furthermore, we use a coordinate transformation to recover periodicity for non-lattice matched heterostructures.

## 5.1 Introduction

Quantum mechanical modelling of heterostructures becomes increasingly important as heterostructures are used in more and more novel devices. For homostructures, the envelope function method provides an elegant way to numerically model microstructures, without the high computational cost of solving the one-electron Schrödinger equation. The envelope function method originates in 1955 when Luttinger and Kohn described equations to determine electron behaviour in a perturbed periodic structure [1]. It was not until 1987 that Burt proved the validity of the envelope function method and the effects of various approximations [2].

From literature, we obtain two different approaches for the calculation of envelope functions in heterostructures. The first uses the effective mass approximation, solving the envelope functions separately in each layer, and using BenDaniel-Duke boundary conditions [3] in between to ensure continuity, one such example is the work of Bastard [4]. The second method uses the so called multi band envelope functions with interband momentum matrix elements, this gives access to more than one band and the interactions between the bands. Because different materials have a different electronic structure, either the envelope functions are calculated for each layer independently with appropriate boundary conditions such as in the work of Rodina et al. [5], or the materials are assumed to be only a weak perturbation of a common reference like in the work of Pötz et al. [6] and more recently Foreman et al. [7].

There seems to be no generally applicable method available for continuous calculation of electron behaviour in non-lattice matched heterostructures with widely varying properties. We develop an envelope function based formalism, capable of describing the wave function behaviour in heterostructure using one continuous set of equations. We apply the slowly varying potential approximation to retain local envelope function equations.

To simplify notation, we describe the one dimensional case. In section 5.2, we

provide a set of continuous envelope functions for lattice matched heterostructures. In section 5.3 we determine the form of a basis transformation, followed by the calculation of its elements in section 5.4. In section 5.5 we calculate the matrix functions using the new basis transformations. We then provide a method for interpretation of the envelope functions in the different layers in section 5.6. In section 5.7 we will consider the changes for a non-lattice matched heterostructure. Finally, in section 5.8 we make an adjustment to better accommodate the restriction of a finite band set. In section 5.9 we present the conclusions of the paper.

## 5.2 Heterostructure Envelope Functions

We describe the heterostructures by assuming perfect bulk crystal behaviour through all layers. Under these circumstances, the crystal potential  $V_c(z)$  becomes a piecewise combination of the periodic potentials  $V_c^l(z)$  of each layer  $l$ ,

$$V_c(z) = \sum_l \theta(z \in \Omega_l) V_c^l(z), \quad (5.1)$$

$\theta(z \in \Omega_l)$  is a logical step function, zero everywhere except in layer  $l$  where it equals to one.

The heterostructure wave function is described by the Schrödinger equation, containing the crystal potential  $V_c(z)$  and an additional external potential  $V_e(z)$ ,

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(z)}{dz^2} + V_c(z)\psi(z) + V_e(z)\psi(z) = E\psi(z). \quad (5.2)$$

We expand the wave functions according to Burt's method,

$$\psi(z) = \sum_n F_n(z) U_n(z), \quad (5.3)$$

where the  $U_n(z)$  are any complete set of lattice periodic basis functions and the  $F_n(z)$  are the envelope functions.

We use the slowly varying potential approximation, meaning we assume  $V_e(z)F_n(z)$  has no components outside the first Brillouin zone. In heterostructures, we justify this as follows; the most abrupt change of the external potential is over the course of one unit cell, if we assume a linear change we get a  $\sin(k)/k$ -like distribution in reciprocal space, where the majority of components lie within the first Brillouin zone. Applying the expansion of equation (5.3) in the Schrödinger equation then yields

$$-\frac{\hbar^2}{2m} \frac{d^2 F_n(z)}{dz^2} - \frac{i\hbar}{m} \sum_m p_{nm} \frac{dF_m(z)}{dz} + \sum_m H_{nm}(z) F_m(z)$$

$$+ V_e(z)F_n(z) = E F_n(z), \quad (5.4)$$

with the interband momentum matrix element  $p_{nm}$  as

$$p_{nm} = \int U_n^*(z')(-i\hbar)\frac{dU_m(z')}{dz'}dz', \quad (5.5)$$

and the bulk Hamiltonian matrix element  $H_{nm}(z)$  position dependent due to the crystal potential of equation (5.1),

$$H_{nm}(z) = \sum_l \theta(z \in \Omega_l) H_{nm}^l \quad (5.6)$$

$$= \sum_l \theta(z \in \Omega_l) \int U_n^*(z') \hat{H}^l(z') U_m(z') dz'. \quad (5.7)$$

To determine the matrix elements, we have to define a set of basis functions  $U_n(z)$ . Zone-centre bulk Hamiltonian solutions are preferred because the matrix elements can then be obtained from empirical  $\mathbf{k} \cdot \mathbf{p}$  fitting. However, the set of zone-centre bulk Hamiltonian solutions in one layer are generally not the same as the zone-centre solutions in a different layer. We denote the set of zone-centre bulk Hamiltonian solutions of layer  $k$  as the set of basis functions  $U_n^k(z)$  of layer  $k$ .

We select the zone-centre bulk Hamiltonian solutions of a reference layer  $k$  as the global set of basis functions  $U_n(z) = U_n^k(z)$ , with associated envelope functions  $F_n^k(z)$ . The envelope function equations using this reference are

$$\begin{aligned} -\frac{\hbar^2}{2m} \frac{d^2 F_n^k(z)}{dz^2} - \frac{i\hbar}{m} \sum_m p_{nm}^k \frac{dF_m^k(z)}{dz} \\ + \sum_m H_{nm}^{l,k}(z) F_m^k(z) + V_e(z) F_n^k(z) = E F_n^k(z), \end{aligned} \quad (5.8)$$

### 5.3 Transformation of Basis Functions

The chosen basis functions  $U_n^k(z)$  diagonalise the bulk Hamiltonian in layer  $k$ , the  $H_{nm}^k$  are then determined by the known band energies at  $\mathbf{k} = \mathbf{0}$  as  $H_{nm}^k = \delta_{n,m} E_n^k$ . However, in any other layer  $l \neq k$  the bulk Hamiltonian matrix elements, expressed in the chosen basis set  $U_n^k(z)$ , are no longer diagonal and do not reduce to generally known matrix elements. In order to calculate them, we define a unitary transformation operator  $\hat{S}^{l \rightarrow k}$  which transforms basis functions from layer  $l$  to layer  $k$ ,

$$|U_n^k\rangle = \hat{S}^{l \rightarrow k} |U_n^l\rangle. \quad (5.9)$$

The transformation is unitary because of the completeness, orthogonality and equal periodicity of the two sets of basis functions  $|U_n^l\rangle$  and  $|U_n^k\rangle$ , and satisfies  $\hat{S}^{-1}\hat{S} = \hat{S}^\dagger\hat{S} = \hat{I}$ , where  $\hat{I}$  is the identity operator and  $\hat{S}^\dagger$  is the Hermitian transpose of  $\hat{S}$ .

We associate transformation matrix elements  $S_{nm}^{l\rightarrow k}$  with the transformation operator  $\hat{S}^{l\rightarrow k}$  by projecting a basis function of layer  $k$  onto a basis function  $l$  and using equation (5.9),

$$\begin{aligned}\langle U_n^l | U_m^k \rangle &= \langle U_n^l | \hat{S}^{l\rightarrow k} | U_m^l \rangle \\ &= \langle U_n^k | \hat{S}^{l\rightarrow k} | U_m^k \rangle = S_{nm}^{l\rightarrow k}.\end{aligned}\quad (5.10)$$

Note that the transformation matrix elements can be expressed in the same way in both sets of basis functions.

## 5.4 Transformation Elements via $\mathbf{k} \cdot \mathbf{p}$

Direct calculation of the transformation matrix elements with equation (5.10) means we need an explicit expression for the basis functions which is not readily available. In the following section we outline a method to deduce the transformation elements from the interband momentum matrix elements of both materials.

The quantum mechanical momentum operator for a free electron under lattice periodic boundary conditions has quantised eigenvalues  $\hbar\mathbf{G}_n$  and associated plane wave eigenfunctions  $\langle \mathbf{r} | \mathbf{G}_n \rangle = e^{i\mathbf{G}_n \cdot \mathbf{r}}$ , with  $\mathbf{G}_n$  a reciprocal lattice vector. The interband momentum matrix elements for such a set of plane waves are then

$$\mathbf{p}_{nm}^{pw} = \langle \mathbf{G}_n | \hat{\mathbf{p}} | \mathbf{G}_m \rangle = \hbar\mathbf{G}_n\delta_{n,m}.\quad (5.11)$$

The plane waves form a complete lattice periodic set of basis functions that conveniently diagonalises the momentum operator. We define a transformation from this set of plane waves (pw) to the set of basis functions of layer  $l$ ,

$$|U_n^l\rangle = \hat{S}^{pw\rightarrow l} | \mathbf{G}_n \rangle = \hat{S} | \mathbf{G}_n \rangle.\quad (5.12)$$

For readability we write  $\hat{S}^{pw\rightarrow l}$  as  $\hat{S}$ . We express the interband momentum matrix elements of layer  $l$  in the set of plane waves,

$$\mathbf{p}_{nm}^l = \langle U_n^l | \hat{\mathbf{p}} | U_m^l \rangle = \langle \mathbf{G}_n | \hat{\mathbf{S}}^\dagger \hat{\mathbf{p}} \hat{\mathbf{S}} | \mathbf{G}_m \rangle.\quad (5.13)$$

Inserting two completeness relations left and right of the momentum operator and using the diagonal plane wave momentum matrix elements of equation (5.11) yields

$$\mathbf{p}_{nm}^l = \sum_i \langle \mathbf{G}_n | \hat{\mathbf{S}}^\dagger | \mathbf{G}_i \rangle \hbar\mathbf{G}_i \langle \mathbf{G}_i | \hat{\mathbf{S}} | \mathbf{G}_m \rangle.\quad (5.14)$$

Equation (5.14) gives a relation between the momentum matrix elements in the plane wave basis set and the momentum eigenvalues of layer  $l$ . Using matrix notation we recognise the eigendecomposition of the interband momentum matrix,

$$\mathbf{P}^l = \mathbf{S}^\dagger \text{diag}(\hbar\mathbf{G}) \mathbf{S}. \quad (5.15)$$

As can be expected, the eigenvalues are those of the momentum operator under the same periodic conditions  $\hbar\mathbf{G}_n$ . The normalised eigenvectors form the transformation matrix  $\mathbf{S}^{\text{pw} \rightarrow l}$  from plane waves to the zone-centre bulk Hamiltonian solutions of layer  $l$ . We obtain the transformation matrix elements from plane wave to a layer  $l$  by diagonalising the interband momentum matrix elements  $p_{nm}^l$  of the layer  $l$ , which are readily available parameters for a whole range of materials and number of bands [8–12].

To transform between two layers  $l$  and  $k$ , we perform two subsequent transformations, first from the basis set of layer  $l$  to a common set of plane waves, then from these plane waves to the basis set of layer  $k$ ,

$$|U_n^k\rangle = \hat{S}^{\text{pw} \rightarrow k} \left[ \hat{S}^{\text{pw} \rightarrow l} \right]^\dagger |U_n^l\rangle \quad (5.16)$$

## 5.5 Transformed Matrix Elements

The interband momentum matrix elements  $p_{nm}^k$  are independent of position, they depend solely on the reference basis set. The position dependent Hamiltonian matrix elements  $H_{nm}(z)$ , however, are layer dependent because of the changing crystal potential.

The local bulk Hamiltonian matrix elements in layer  $l$ , using the basis functions  $U_n^k(z)$  are

$$H_{nm}^{l,k} = \langle U_n^k | \hat{H}^l | U_m^k \rangle = \langle U_n^k | -\frac{\hbar^2}{2m} \frac{d^2}{dz^2} + V_c^l(z) | U_m^k \rangle. \quad (5.17)$$

If  $k = l$ , we previously discussed we simply obtain the band energies at  $\mathbf{k} = \mathbf{0}$ ,

$$H_{nm}^{k,k} = \delta_{n,m} E_n^k \quad (5.18)$$

For  $l \neq k$ ,  $\hat{H}^l$  is no longer diagonalisable within the basis set of layer  $k$  and  $H_{nm}^{l,k}$  has to be calculated using the transformation between layers. We diagonalise the local bulk Hamiltonian using eigendecomposition,

$$H_{nm}^{l,k} = \langle U_n^k | H^l | U_m^k \rangle = \langle U_n^k | U^l E^l [U^l]^\dagger | U_m^{(k)} \rangle. \quad (5.19)$$

$U^l$  is a matrix with the zone-centre bulk Hamiltonian solutions of layer  $l$  as columns. The diagonal matrix  $E^l$  contains the band edge energies of layer  $l$ .

We transform the  $U_n^k$  basis functions to the basis functions of layer  $l$  and obtain

$$H_{nm}^{l,k} = \langle U_n^l | [S^{l \rightarrow k}]^\dagger U^l E^l [U^l]^\dagger S^{l \rightarrow k} | U_m^l \rangle. \quad (5.20)$$

We conveniently express these matrix elements in their full matrix notation,

$$H^{l,k} = [S^{l \rightarrow k}]^\dagger E^l S^{l \rightarrow k}, \quad (5.21)$$

where the components of  $H^{l,k}$  are the bulk Hamiltonian matrix elements  $H_{nm}^{l,k}$ . The complete position dependent bulk Hamiltonian matrix is then

$$H^k(z) = \sum_l \theta(z \in \Omega_l) [S^{l \rightarrow k}]^\dagger E^l S^{l \rightarrow k}. \quad (5.22)$$

## 5.6 Interpretation in Different Layers

In layer  $k$  the envelope functions of equation (5.8) represent the mesoscopic variation of the wave function corresponding to the electronic band  $n$ , however, for  $z$  in a layer  $l \neq k$  interpretation of the envelope functions  $F_n^k(z)$  is not that straightforward. We have to transform the envelope function  $F_n^k(z)$  of the reference layer  $k$  to the envelope functions  $F_n^l(z)$  of the layer to be interpreted  $l$ , using

$$F_n^l(z) = \sum_m \langle U_n^l | S^{l \rightarrow k} | U_m^l \rangle F_m^k(z) = \sum_m S_{nm}^{l \rightarrow k} F_m^k(z). \quad (5.23)$$

This represents a projection of all envelope functions defined against the basis set of  $k$  onto basis function  $n$  of layer  $l$ . Physically interpretable solutions can thus be obtained by transforming every envelope function in each layer to their own set of zone-centre basis functions.

## 5.7 Non-lattice matched heterostructures

Up to now, we assumed the heterostructure contained only lattice matched crystals so the heterostructure lattice is periodic, for most material combinations there is usually a mismatch leading to a non-periodic structure. To regain periodicity we use a coordinate transformation,

$$\tilde{z} = \tilde{z}(z) = \int \sum_l \theta(z \in \Omega_l) \frac{a_{\text{ref}}}{a_l} dz. \quad (5.24)$$

The coordinate transformation is piecewise linear such that each layer is continuously scaled to the same reference lattice constant  $a_{\text{ref}}$ . In this way, the heterostructure is again periodic in the coordinate  $\tilde{z}$ .

We transform the Hamiltonian to the transformed coordinates  $\tilde{z}$ ,

$$\hat{H}_{\tilde{z}} = -\frac{\hbar^2}{2m} \frac{d^2}{dz^2} + V(z) \quad (5.25)$$

$$= -\frac{\hbar^2}{2m} \frac{d\tilde{z}}{dz} \frac{d}{d\tilde{z}} \left( \frac{d\tilde{z}}{dz} \frac{d}{d\tilde{z}} \right) + \tilde{V}(\tilde{z}) \quad (5.26)$$

where  $\tilde{V}(\tilde{z})$  is the transformed version of the total potential  $\tilde{V}(\tilde{z}(z)) = V(z)$ . We assign an operator  $\hat{D}$  to  $\frac{d\tilde{z}}{dz} \frac{d}{d\tilde{z}}$  to simplify notation,

$$\hat{H}_{\tilde{z}} = -\frac{\hbar^2}{2m} \hat{D} \hat{D} + \tilde{V}(\tilde{z}). \quad (5.27)$$

We also transform the envelope function expansion to  $\tilde{z}$  coordinates,  $\psi(z)$ ,  $U_n(z)$  and  $F_n(z)$  are transformed similar to  $V(z)$ ,

$$\tilde{\psi}(\tilde{z}) = \sum_n \tilde{F}_n(\tilde{z}) \tilde{U}_n(\tilde{z}). \quad (5.28)$$

The envelope function equations in the transformed coordinate  $\tilde{z}$  are similar to the direct  $z$  case, using the transformed functions and replacing  $\frac{d}{dz}$  with  $\hat{D}$ ,

$$\begin{aligned} -\frac{\hbar^2}{2m} \hat{D} \hat{D} \tilde{F}_n(\tilde{z}) - \frac{i\hbar}{m} \sum_m \tilde{p}_{nm}(\tilde{z}) \hat{D} \tilde{F}_m(\tilde{z}) \\ + \sum_m \tilde{H}_{nm}(\tilde{z}) \tilde{F}_m(\tilde{z}) + \tilde{V}_e(\tilde{z}) \tilde{F}_n(\tilde{z}) = E \tilde{F}_n(\tilde{z}). \end{aligned} \quad (5.29)$$

The interband momentum matrix elements  $\tilde{p}_{nm}$  in transformed coordinates  $\tilde{z}$  use the  $\hat{D}$  operator,

$$\tilde{p}_{nm} = \int \tilde{U}_n^*(\tilde{z}') (-i\hbar) \hat{D} \tilde{U}_m(\tilde{z}') d\tilde{z}'. \quad (5.30)$$

We relate them to their direct counterparts by transforming back to  $z$  coordinates,

$$\tilde{p}_{nm} = \int U_n^*(z') (-i\hbar) \frac{d\tilde{z}'}{dz'} \frac{dU_m(z')}{d\tilde{z}'} \frac{d\tilde{z}'}{dz'} dz'. \quad (5.31)$$

The matrix element is defined using an integration over one unit cell, and because  $\tilde{z}(z)$  is always linear in a unit cell, its derivative is a constant and we can remove it from the integration, yielding a layer dependent matrix element,

$$\tilde{p}_{nm}(\tilde{z}) = \sum_l \theta(\tilde{z} \in \tilde{\Omega}_l) \tilde{p}_{nm}^l. \quad (5.32)$$

with the local interband momentum matrix elements in transformed coordinates  $\tilde{z}$ ,

$$\tilde{p}_{nm}^l = \left. \frac{d\tilde{z}}{dz} \right|_{z \in \Omega_l} p_{nm}. \quad (5.33)$$

Applying the same procedure to  $\tilde{H}_{nm}^l$ , we find a similar relation

$$\tilde{H}_{nm}^l = \left. \frac{d\tilde{z}}{dz} \right|_{z \in \Omega_l} H_{nm}^l. \quad (5.34)$$

The constants  $\left. \frac{d\tilde{z}}{dz} \right|_{z \in \Omega_l}$  describe the influence of the scaling, and are equal to  $a_{\text{ref}}/a_l$ .

The envelope function equations in transformed coordinates  $\tilde{z}$  are then completely known,

$$\begin{aligned} -\frac{\hbar^2}{2m} \hat{D} \hat{D} \tilde{F}_n(\tilde{z}) - \frac{i\hbar}{m} \frac{d\tilde{z}}{dz} \sum_m p_{nm}(\tilde{z}) \hat{D} \tilde{F}_m(\tilde{z}) \\ + \frac{d\tilde{z}}{dz} \sum_m H_{nm}(\tilde{z}) \tilde{F}_m(\tilde{z}) + \tilde{V}_e(\tilde{z}) \tilde{F}_n(\tilde{z}) = E \tilde{F}_n(\tilde{z}). \end{aligned} \quad (5.35)$$

This equation can usually be directly applied in numerical calculations without expanding operator  $\hat{D} \hat{D} = \frac{d\tilde{z}}{dz} \frac{d}{d\tilde{z}} \left( \frac{d\tilde{z}}{dz} \frac{d}{d\tilde{z}} \right)$  any further.

## 5.8 Restriction of Band Set

For numerical evaluation the considered band set must be restricted to a finite number of bands. For notational purposes we again consider a lattice matched system. For a finite number of bands the set of basis functions is no longer complete and the transformation is not exact. It is then advantageous to use a position dependent transformed interband momentum matrix element,

$$P^k(z) = \sum_l \theta(z \in \Omega_l) [S^{l \rightarrow k}]^\dagger P^l S^{l \rightarrow k}. \quad (5.36)$$

We prove that using the interband momentum matrix of equation (5.36) is indeed better by describing the envelope functions of layer  $l$  in the basis set of another layer  $k$ . For  $z$  in layer  $l$  we can write the limited band set envelope function equations as the following matrix equation:

$$\begin{aligned} -\frac{\hbar^2}{2m} \frac{d^2}{dz^2} \mathbf{F}^k(z) - \frac{i\hbar}{m} P^k \frac{d}{dz} \mathbf{F}^k(z) + [S^{l \rightarrow k}]^\dagger E^l S^{l \rightarrow k} \mathbf{F}^k(z) \\ + V_e(z) \mathbf{F}^k(z) = E \mathbf{F}^k(z) \text{ for } z \in \Omega_l, \end{aligned} \quad (5.37)$$

where  $\mathbf{F}^k(z)$  is a vector whose components are the envelope functions  $F_n^k(z)$ . In this case interpretation in layer  $l$  is provided by

$$\mathbf{F}^l(z) = S^{l \rightarrow k} \mathbf{F}^k(z). \quad (5.38)$$

Using equation (5.38), we transform equation (5.37) to the basis set of layer  $l$ . Using the linearity of derivation, and left-multiplying both sides with  $S^{l \rightarrow k}$ , we obtain

$$-\frac{\hbar^2}{2m} \frac{d^2}{dz^2} \mathbf{F}^l(z) - \frac{i\hbar}{m} S^{l \rightarrow k} P^k [S^{l \rightarrow k}]^\dagger \frac{d}{dz} \mathbf{F}^l(z) + E^l \mathbf{F}^l(z) + V_e(z) \mathbf{F}^l(z) = E \mathbf{F}^l(z) \text{ for } z \in \Omega_l. \quad (5.39)$$

When the transformation is not exact the term  $S^{l \rightarrow k} P^k [S^{l \rightarrow k}]^\dagger$  does not reduce to  $P^l$  and we do not retrieve the envelope function equations in layer  $l$ .

If we had used the position dependent interband momentum matrix of equation (5.36), the transformation matrices cancel and we get the right envelope function equations in layer  $l$ ,

$$-\frac{\hbar^2}{2m} \frac{d^2}{dz^2} \mathbf{F}^l(z) - \frac{i\hbar}{m} P^l \frac{d}{dz} \mathbf{F}^l(z) + E^l \mathbf{F}^l(z) + V_e(z) \mathbf{F}^l(z) = E \mathbf{F}^l(z) \text{ for } z \in \Omega_l. \quad (5.40)$$

Using the position dependent interband matrix element thus allows for a better description in the case of a restricted band set.

## 5.9 Conclusion

The formalism we developed can be generally applied to heterostructures, even if they are non-lattice matched and have widely different basis functions. Our proposed formalism uses no extra parameters and the approximation is limited to assuming a slowly varying potential, which is justified in most practical cases, even for abrupt heterostructures. Continuity is incorporated in the envelope function equations and no internal boundary conditions need to be set. The formalism is easily implemented numerically and produces little overhead because the transformation matrices can be pre-calculated.

## Acknowledgements

William Vandenberghe gratefully acknowledges the support of a Ph.D. stipend from IWT-Vlaanderen.

## Bibliography

- [1] J. M. Luttinger and W. Kohn, "Motion of electrons and holes in perturbed periodic fields," *Phys. Rev.*, vol. 97, pp. 869–883, Feb 1955. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRev.97.869>
- [2] M. G. Burt, "An exact formulation of the envelope function method for the determination of electronic states in semiconductor microstructures," *Semiconductor Science and Technology*, vol. 2, no. 7, p. 460, 1987. [Online]. Available: <http://stacks.iop.org/0268-1242/2/i=7/a=012>
- [3] D. J. BenDaniel and C. B. Duke, "Space-charge effects on electron tunneling," *Phys. Rev.*, vol. 152, pp. 683–692, Dec 1966. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRev.152.683>
- [4] G. Bastard, "Superlattice band structure in the envelope-function approximation," *Phys. Rev. B*, vol. 24, pp. 5693–5697, Nov 1981. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevB.24.5693>
- [5] A. V. Rodina, A. Y. Alekseev, A. L. Efros, M. Rosen, and B. K. Meyer, "General boundary conditions for the envelope function in the multiband  $\mathbf{k}\cdot\mathbf{p}$  model," *Phys. Rev. B*, vol. 65, p. 125302, Feb 2002. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevB.65.125302>
- [6] W. Pötz, W. Porod, and D. K. Ferry, "Theoretical study of subband levels in semiconductor heterostructures," *Phys. Rev. B*, vol. 32, pp. 3868–3875, Sep 1985. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevB.32.3868>
- [7] B. A. Foreman, "First-principles envelope-function theory for lattice-matched semiconductor heterostructures," *Phys. Rev. B*, vol. 72, p. 165345, Oct 2005. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevB.72.165345>
- [8] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, "Band parameters for iii–v compound semiconductors and their alloys," *Journal of Applied Physics*, vol. 89, no. 11, pp. 5815–5875, 2001. [Online]. Available: <http://link.aip.org/link/?JAP/89/5815/1>
- [9] S. Richard, F. Aniel, and G. Fishman, "Energy-band structure of ge, si, and gaas: a thirty-band  $\mathbf{kp}$  method," *Phys. Rev. B*, vol. 70, pp. 235–204, Dec 2004. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevB.70.235204>
- [10] —, "Energy-band structure of ge, si, and gaas: a thirty-band  $\mathbf{kp}$  method," *Phys. Rev. B*, vol. 70, p. 235204, Dec 2004. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevB.70.235204>

- [11] M. Cardona and F. H. Pollak, "Energy-band structure of germanium and silicon: The k·p method," *Phys. Rev.*, vol. 142, pp. 530–543, Feb 1966. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRev.142.530>
- [12] S. B. Radhia, N. Fraj, I. Saidi, and K. Boujdaria, "The eight-level k·p model for the conduction and valence bands of inas, inp, insb," *Semiconductor Science and Technology*, vol. 22, no. 4, p. 427, 2007. [Online]. Available: <http://stacks.iop.org/0268-1242/22/i=4/a=024>



## Chapter 6

# BTBT in graphene based semiconductors

### Zener Tunnelling in Graphene Based Semiconductors - the k·p Method

William G. Vandenberghe, Bart Sorée, Wim Magnus, and Guido Groeseneken

In *Proceedings of EDISON 16: 16th International Conference on Electron Dynamics in Semiconductors, Optoelectronics and Nanostructures*, pp. 012111, Montpellier (France), August 2009.

The following article appeared in *Journal of Physics: Conference Series*, 193, pp. 012111, August 2009 and may be found at <http://dx.doi.org/10.1088/1742-6596/193/1/012111>.

©2009 IOP Publishing Limited. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the IOP Publishing Limited.

### **Contributions of first author**

- Conception and development of the theory
- Creating figures
- Text writing and editing

## Abstract

The carbon nanotube and graphene nanoribbon band structure is derived using the two-band  $\mathbf{k}\cdot\mathbf{p}$  method and shown to have a similar band structure as a III-V semiconductor. Contrary to a previous claim, it is shown that the tunnelling probability is lower for a graphene based semiconductor than for a III-V semiconductor with the same bandgap. Considering the relation between the bandgap and the effective mass we conclude that a graphene based semiconductor is not well-suited for a classical device which suffers from Zener tunnelling, but is rather promising for a device which has its working principle based on Zener tunnelling.

## 6.1 Introduction

Graphene based materials such as carbon nanotubes and graphene nanoribbons provide an interesting research topic since they have a band structure which is different from traditional semiconductors. Graphene has a zero bandgap at the K-points in the Brillouin zone, but some carbon nanotubes or graphene nanoribbons exhibit a non-zero bandgap which, in principle, enables the fabrication of semiconductor devices. These semiconducting carbon nanotubes and graphene nanoribbons will be referred to as graphene based semiconductors throughout this paper.

Graphene based semiconductors generally have a small bandgap and hence Zener tunnelling, also called band-to-band tunnelling (BTBT), cannot be neglected in possible graphene based semiconductor devices. A recent publication [1] shows a calculation of the Zener tunnelling current using the WKB method and concludes that it exceeds the current of group IV and III-V semiconductors with identical bandgaps.

In this paper, we first develop the band structure within the  $\mathbf{k}\cdot\mathbf{p}$  scheme for bulk graphene as well as for nanoribbons or nanotubes with a non-zero bandgap. Next, we consider the tunnelling probability and refute the claim that tunnelling is more efficient in a graphene based semiconductor compared to a III-V semiconductor with the same bandgap, and we explain the discrepancy with the previous publication [1].

## 6.2 Graphene based semiconductors: band structure

For the  $\mathbf{k}\cdot\mathbf{p}$  expansion of the bulk graphene band structure around the  $K$ -point we use two bands: the  $\pi^*$  conduction band and the  $\pi$  valence band which are degenerate at the  $K$ -point due to symmetry [2]. Using these states to construct the  $\mathbf{k}\cdot\mathbf{p}$  Hamiltonian, we need to determine the momentum matrix elements

$$\mathbf{p}_{11} = \int u_1^*(\mathbf{r}, z) \mathbf{p} u_1(\mathbf{r}, z) d^3r, \quad (6.1)$$

$$\mathbf{p}_{12} = \int u_1^*(\mathbf{r}, z) \mathbf{p} u_2(\mathbf{r}, z) d^3r, \quad (6.2)$$

$$\mathbf{p}_{22} = \int u_2^*(\mathbf{r}, z) \mathbf{p} u_2(\mathbf{r}, z) d^3r, \quad (6.3)$$

appearing as integrals over the 2D graphene unit cell in the plane and over the entire space in the  $z$ -direction perpendicular to the plane,

$$\int d^3r = \int_{\text{unitcell}} d^2r \int_{-\infty}^{\infty} dz. \quad (6.4)$$

Here,  $u_1(\mathbf{r}, z)$  and  $u_2(\mathbf{r}, z)$  are two Bloch functions which are orthogonal solutions to the Schrödinger equation with eigen energy equal to zero:

$$\left[ -\frac{\hbar^2 \nabla^2}{2m} + U_{\text{lat}}(\mathbf{r}, z) \right] u_{1,2}(\mathbf{r}, z) = 0 \quad (6.5)$$

$U_{\text{lat}}(\mathbf{r}, z)$  is the potential due to the graphene lattice which is periodic with respect to  $\mathbf{r}$ ,  $m$  is the free electron mass and  $\hbar$  is the reduced Planck constant.

As the  $z$ -axis stands perpendicular to the graphene plane, both the  $\pi$  and  $\pi^*$  bands are odd under a reflection in the  $xy$ -plane and as a result the 3 selected matrix elements have no  $z$  component. The  $y$  direction is chosen such that  $u_1$  and  $u_2$  are respectively odd and even under reflection in the  $yz$ -plane. This requires  $\mathbf{p}_{11}$  and  $\mathbf{p}_{22}$  to be in the  $y$  direction and  $\mathbf{p}_{12}$  in the  $x$  direction. At the  $K$ -point the valence and conduction band are degenerate due to an additional rotational symmetry over  $120^\circ$ . This adds the requirement  $|\mathbf{p}_{11}| = |\mathbf{p}_{12}| = |\mathbf{p}_{22}| = p$  and  $\mathbf{p}_{11} = -\mathbf{p}_{22}$ .

According to the  $\mathbf{k}\cdot\mathbf{p}$  method a state with wavevector  $\mathbf{k}$  relative to a  $K$ -point can be written as

$$\psi(\mathbf{r}, z) = e^{i\mathbf{k}\cdot\mathbf{r}} (a_1 u_1(\mathbf{r}, z) + a_2 u_2(\mathbf{r}, z)) \quad (6.6)$$

and its energy eigenvalues  $E_{\pm}(\mathbf{k})$  are determined by the 2-band  $\mathbf{k}\cdot\mathbf{p}$  Hamiltonian

$$H = \begin{pmatrix} \frac{\hbar^2}{2m}(k_x^2 + k_y^2) + \frac{\hbar p}{m} k_x & \frac{\hbar p}{m} k_y \\ \frac{\hbar p}{m} k_y & \frac{\hbar^2}{2m}(k_x^2 + k_y^2) - \frac{\hbar p}{m} k_x \end{pmatrix} \quad (6.7)$$

yielding

$$E_{\pm}(\mathbf{k}) = \frac{\hbar^2 |\mathbf{k}|^2}{2m} \pm \frac{\hbar p}{m} |\mathbf{k}|, \quad |\mathbf{k}| = \sqrt{k_x^2 + k_y^2}. \quad (6.8)$$

To first order, this corresponds to the nearest neighbour tight binding expression

$$E_{\pm}(\mathbf{k}) \approx \pm \hbar v_F |\mathbf{k}| \quad (6.9)$$

and the matrix element  $p$  can be determined as  $p = mv_F$  with  $v_F$  the graphene Fermi velocity.

Note that an alternative choice of basis functions  $u_1$  and  $u_2$  and omission of the second derivative transforms Eq. (6.7) into a more familiar formulation of the Hamiltonian

$$H = \frac{\hbar}{m} p \begin{pmatrix} 0 & k_x - ik_y \\ k_x + ik_y & 0 \end{pmatrix}, \quad (6.10)$$

the resulting Schrödinger equation being algebraically identical to the 2D Weyl equation [3].

Returning to Eq. (6.7) for the case of a semiconducting nanoribbon or nanotube we use the zone folding technique to determine the band structure. The  $\mathbf{k}$  vector is continuous in one direction and quantized in the perpendicular direction. For the sake of convenience we select a zig-zag nanoribbon or nanotube such that  $k_x$  is quantized and  $k_y$  is continuous. The following derivation of the band structure is valid however for an arbitrary orientation of the tube or ribbon since Eq. (6.8) is invariant under rotations.

Denoting  $k_0$  by the smallest allowed value of  $k_x$  we obtain the bandgap of the graphene based semiconductor

$$E_g = 2 \frac{\hbar p}{m} k_0 \quad (6.11)$$

where size quantization allows for  $k_0 \neq 0$  in the case of a semiconducting nanotube or ribbon, while Eq. (6.7) turns into

$$H = \begin{pmatrix} E_0 + \frac{\hbar^2 k_y^2}{2m} + \frac{E_g}{2} & \frac{\hbar p}{m} k_y \\ \frac{\hbar p}{m} k_y & E_0 + \frac{\hbar^2 k_y^2}{2m} - \frac{E_g}{2} \end{pmatrix} \quad (6.12)$$

with  $E_0 = \hbar^2 k_0^2 / 2m$  and the energy dispersion relation for the lowest subband reads

$$E_{\pm} = E_0 + \frac{\hbar^2 k_y^2}{2m} \pm \sqrt{\frac{E_g^2}{4} + \frac{\hbar^2 p^2}{m^2} k_y^2}. \quad (6.13)$$

Within the 2 band approximation Eq. (6.12) is formally identical to the  $\mathbf{k} \cdot \mathbf{p}$  Hamiltonian of a III-V semiconductor with a bandgap  $E_g$ , a matrix element  $p$  and an energy offset  $E_0$  [4, 5]. More specifically, in a III-V semiconductor,

the two-band Hamiltonian describes the interaction between the light valence band and the conduction band. The energy dispersion relation of a graphene based semiconductor is therefore the same as that of a III-V semiconductor apart from the absence of a heavy and a split-off valence band.

### 6.3 Tunnelling probability

As a graphene based semiconductor and a III-V semiconductor share a similar band structure, the expression of the Zener tunnelling probability for an electron in a uniform field will also be the same:

$$T = \exp\left(\frac{-\pi E_g^2 m}{4\hbar p q F}\right) \quad (6.14)$$

with  $F$  the applied uniform electric field and  $q$  the unit charge [5]. To compare different materials, we consider the matrix element  $p$  which is generally expressed as the energy unit  $E_p = 2p^2/m$  taking the value  $E_p = 2mv_F^2 = 11.4\text{eV}$  for graphene and  $E_p \approx 20\text{eV}$  for III-V semiconductors, such as GaAs, InAs, InSb, ... [6].

For conventional devices, such as the metal-oxide-semiconductor field-effect transistor (MOSFET), BTBT is a detrimental effect since it increases the off-state current. To make a fair comparison concerning the amount of BTBT we must consider MOSFETs with a common figure of merit such as the transport effective mass. Making use of the relation between the effective mass and the bandgap,  $m^* \approx mE_g/E_p$  and the definition of  $E_p$ , the transmission coefficient (6.14) can be written as

$$T \approx \exp\left(\frac{-\pi E_p^{3/2} m^{*2}}{2\sqrt{2}m^{3/2}\hbar q F}\right). \quad (6.15)$$

Since  $E_{p,\text{graphene}} < E_{p,\text{III-V}}$ , Eq. (6.15) shows that for the same effective mass, the tunnelling probability will be much greater for a graphene based MOSFET compared to a III-V based MOSFET as shown in Fig. 6.1. Or, in other words a graphene based MOSFET with the same on-current characteristics as a III-V based MOSFET will show a higher leakage current due to increased BTBT.

In the case of a III-V semiconductor, the effective mass ( $m^*$ ) refers to the conduction band mass or the light hole mass. In the case of a graphene based semiconductor, there is only one direction with a continuous wavevector defining single valence and conduction bands. The absence of a heavy hole and a split-off band in graphene band structure could still allow graphene to outperform a similar III-V device but certainly for nMOS devices we do not expect an improvement over III-V MOSFETs.

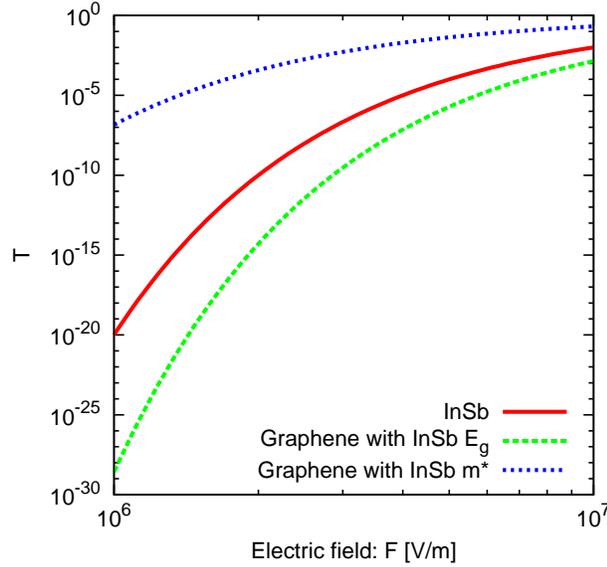


Figure 6.1: Comparison between the Zener tunnelling probability of InSb and a graphene based semiconductor with the InSb bandgap (Eq. 6.14) and a graphene based semiconductor with the InSb effective mass (Eq. 6.15).

For devices operating on the basis of BTBT such as the tunnel FET [7], large tunnelling probabilities are required. Since the matrix element  $p \propto \sqrt{E_p}$  is smaller for graphene than for III-V semiconductors, graphene based semiconductors have a smaller tunnel probability for the same bandgap as shown in Fig. 6.1. To obtain the same performance as a III-V tunnel FET, a graphene based semiconductor tunnel FET must have a smaller bandgap. But since for graphene based semiconductors the bandgap can be made smaller by selecting a different size of nanoribbon or nanotube, this needs not be a disadvantage if graphene based semiconductors can be fabricated reliably.

Finally, we report on a discrepancy between our results and the work by Jena et al. [1] who came to the opposite conclusion concerning the tunnelling probability. To our opinion, the disagreement stems from the assumption by Jena et al. that traditional semiconductors have a "parabolic bandgap". Obviously, the parabolic nature of the band structure in any semiconductor only holds in a limited energy range and does not cover the entire Brillouin zone. As we have shown in our previous section, there is no formal difference in the band structure between a graphene based semiconductor and a III-V semiconductor except for the presence of an additional heavy hole and split-off band in the latter.

The direct cause of the disagreement lies in the use of the transmission coefficient [8]

$$T = \exp\left(-\frac{4\sqrt{2m^*}E_g^{3/2}}{3\hbar F}\right) \quad (6.16)$$

by Jena et al. In the case of Zener tunnelling, the "parabolic bandgap" approximation is too crude, comparing Eq. (6.16) with Eq. (6.14) and using  $p = m/\sqrt{2}\sqrt{E_g/m^*}$  we observe a deviation of a factor  $3\pi/16 = 0.59$  in the exponential. Correcting for this factor the conclusion from the previous publication that graphene has a larger tunnel probability compared to a III-V with the same bandgap is reversed and agrees with our current publication, claiming that the III-V material will show the larger tunnel probability for the same bandgap.

## 6.4 Conclusions

The band structure of the valence and the conduction band in a graphene based semiconductor can be described in the same way as the light valence and the conduction band structure in a III-V semiconductor.

Graphene based semiconductors have a lower Zener tunnelling probability than III-V semiconductors for the same bandgap.

Semiconductor devices where the effective mass is a figure of merit and where leakage (off-current) due to Zener tunnelling is unwanted such as the MOS-FET will show higher leakage for the same on-current in the case of graphene based semiconductors compared to III-V semiconductors. On the other hand, graphene based semiconductors are well suited for devices which operate on the basis of Zener tunnelling such as the tunnel FET.

## Acknowledgements

William Vandenberghe gratefully acknowledges the support of a Ph.D. stipend from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen).

## Bibliography

- [1] D. Jena, T. Fang, Q. Zhang, and H. Xing, "Zener tunneling in semiconducting nanotube and graphene nanoribbon p - n junctions,"

- Applied Physics Letters*, vol. 93, no. 11, p. 112106, 2008. [Online]. Available: <http://link.aip.org/link/?APL/93/112106/1>
- [2] J. M. Stéphanie Reich, Christian Thomsen, *Carbon nanotubes: basic concepts and physical properties*. Wiley-VCH, 2004.
- [3] D. P. DiVincenzo and E. J. Mele, “Self-consistent effective-mass theory for intralayer screening in graphite intercalation compounds,” *Phys. Rev. B*, vol. 29, no. 4, pp. 1685–1694, Feb 1984.
- [4] W. Vandenberghe, B. Sorée, W. Magnus, and G. Groeseneken, “Zener tunneling in semiconductors under non-uniform fields,” 2009, submitted for publication to *Phys. Rev. B*.
- [5] E. O. Kane, “Zener tunneling in semiconductors,” *Journal of Physics and Chemistry of Solids*, vol. 12, pp. 181–188, 1959.
- [6] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, “Band parameters for iii-v compound semiconductors and their alloys,” *Journal of Applied Physics*, vol. 89, no. 11, pp. 5815–5875, 2001. [Online]. Available: <http://link.aip.org/link/?JAP/89/5815/1>
- [7] W. Vandenberghe, A. Verhulst, G. Groeseneken, B. Sorée, and W. Magnus, “Analytical model for point and line tunneling in a tunnel field-effect transistor,” in *Simulation of Semiconductor Processes and Devices, 2008. SISPAD 2008. International Conference on*, Sept. 2008, pp. 137–140.
- [8] S. Sze, *Physics of Semiconductor Devices*. 2nd ed. John Wiley & Sons, 1981.



## Chapter 7

# BTBT in Germanium

Direct and Indirect Band-to-Band Tunneling in  
Germanium-based TFETs

Kuo-Hsing Kao, Anne S. Verhulst, William G. Vandenberghe,  
Bart Sorée, Guido Groeseneken, and Kristin De Meyer

Published in *IEEE Transactions on Electronic Devices*, 59 (2),  
pp.292-301, February 2012

©2012 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

### **Contributions of third author**

- Discovery of importance of direct tunneling in germanium
- Description of the electron-phonon interaction
- Text correcting and editing

## Abstract

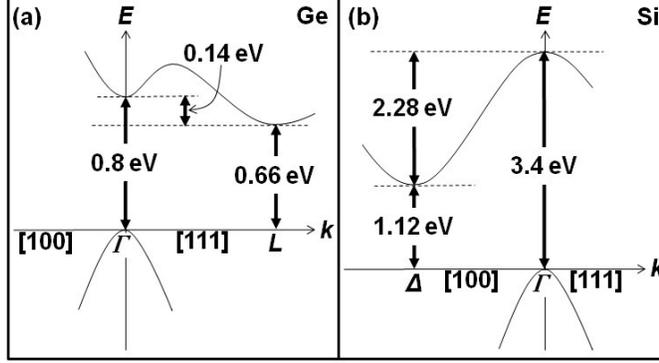
Germanium is a widely-used material for tunnel FETs because of its small band gap and compatibility with silicon. Typically only the indirect band gap of Ge at 0.66 eV is considered. However, direct band-to-band tunneling in Ge should be included in tunnel FET modeling and simulations since the energy difference between the Ge conduction band edges at the L and  $\Gamma$  valleys is only 0.14 eV at room temperature. In this paper, we theoretically calculate the parameters  $A$  and  $B$  of Kane's direct and indirect band-to-band tunneling models at different tunneling directions ([100], [110] and [111]) for Si, Ge and unstrained  $\text{Si}_{1-x}\text{Ge}_x$ . We highlight how the direct band-to-band tunneling component becomes more important as the Ge mole fraction increases. The calculation of the band-to-band generation rate in the uniform electric field limit reveals that direct tunneling always dominates over indirect tunneling in Ge. The impact of the direct transition in Ge on the performance of two realistic TFET configurations is illustrated with TCAD simulations. The influence of field-induced quantum confinement is included in the analysis based on a back-of-the-envelope calculation.

## 7.1 Introduction

The tunnel field-effect transistor (TFET) is one of the devices that can reach sub-60 mV/decade subthreshold swing (SS) since the injection mechanism of carriers from the source is based on band-to-band tunneling (BTBT) [1–7]. However poor SS and low on-current have been seen in all-Si TFETs due to the large indirect energy band gap ( 1.12 eV). Therefore a Si-based heterojunction TFET has been proposed to reduce the effective tunneling barrier height by replacing the Si source with a smaller band gap material, e.g. germanium for n-channel TFETs (nTFETs) [5–15]. Also all-Ge TFETs are being studied [16,17]. So far most research on Ge-based TFETs has merely considered the indirect transitions [6–17] and the default parameter file of the TCAD simulation tool [18] only specifies the Ge parameters for the indirect BTBT transition. At room temperature, however, the energy difference in the conduction band edges at the L and  $\Gamma$  valleys of Ge is only 0.14 eV (see Fig. 7.1) [19]. Moreover, direct transitions are experimentally observed to dominate the BTBT current in Ge diodes [20,21]. The appropriate BTBT parameters for  $\text{Si}_{1-x}\text{Ge}_x$  with varying Ge mole fraction have not been investigated.

In this paper, we start with the theoretical Si, Ge and unstrained  $\text{Si}_{1-x}\text{Ge}_x$  parameter calculations of Kane's direct and indirect BTBT model in Section

Figure 7.1: Schematic band structure of (a): Ge and (b): Si at room temperature.



II. Based on these calculated parameters we discuss the BTBT generation rate as a function of the electric field in the uniform electric field limit. In Sec. III we present the impact of the direct band gap of Ge on two different Ge-based TFETs by simulations, and discuss the influence of quantum confinement.

## 7.2 Theoretical BTBT calculation

### 7.2.1 Parameter Calculation

In this section, we theoretically calculate the  $A_{\text{dir}}/B_{\text{dir}}$  and  $A_{\text{ind}}/B_{\text{ind}}$  parameters of Kane's direct and indirect BTBT models in different tunneling directions ([100], [110] and [111]) for Si, Ge and unstrained  $\text{Si}_{1-x}\text{Ge}_x$ , where the minimum band gap as well as the direct band gap at the  $\Gamma$  point are considered, as shown in Fig. 7.1.

In the uniform electric-field limit, the widely-used Kane's model to determine the BTBT generation rate  $G$  per unit volume is given by [18, 22]:

$$G = A \left( \frac{F}{F_0} \right)^P \exp\left(-\frac{B}{F}\right) \quad (7.1)$$

where  $F_0 = 1 \text{ V/cm}$ ,  $P = 2$  and  $2.5$  for the direct and indirect BTBT, respectively,  $A$  and  $B$  are the Kane parameters and  $F$  is the electric field. The prefactor  $A$  and exponential factor  $B$  for direct and indirect transitions can be expressed by [18, 22, 23]:

$$A_{\text{dir}} = \left( \frac{g\pi m_r^{1/2} (qF_0)^2}{9\hbar^2 (E_g^\Gamma)^{1/2}} \right) \frac{9}{\pi^2} \quad (7.2)$$

$$B_{\text{dir}} = \frac{\pi^2 m_r^{1/2} (E_g^\Gamma)^{3/2}}{qh} \quad (7.3)$$

$$A_{\text{ind}} = \frac{g(m_c m_v)^{3/2} (1 + 2N_{\text{TA}}) D_{\text{TA}}^2 (qF_0)^{5/2}}{2^{21/4} h^{5/2} m_r^{5/4} \rho \epsilon_{\text{TA}} E_g^{7/4}} \quad (7.4)$$

$$B_{\text{ind}} = \frac{2^{7/2} \pi m_r^{1/2} E_g^{3/2}}{3qh} \quad (7.5)$$

where  $g$  is a degeneracy factor detailed later in this section,  $m_r$  is the reduced tunneling mass,  $q$  is the elementary charge,  $h$  is Planck's constant,  $E^\Gamma$  is the direct band gap at the  $\Gamma$  point,  $m_v$  ( $m_c$ ) is the valence (conduction) band density of states effective mass,  $N_{\text{TA}} = 1/[\exp(\epsilon_{\text{TA}}/kT) - 1]$  is the occupation number of the transverse acoustic phonon at temperature  $T$ , where  $k$  is the Boltzmann constant,  $D_{\text{TA}}$  is the deformation potential of transverse acoustic phonons,  $\rho$  is the mass density,  $\epsilon_{\text{TA}}$  is the transverse acoustic phonon energy and  $E_g$  is the minimum band gap. Note that we are using the same formula (Eqs. 7.2-7.5) as given by SentaaurusDevice [18] except for a correction to the prefactor  $A_{\text{dir}}$ , which is here multiplied by  $(9/\pi^2)$  according to Ref. [23]. For the indirect model, a more complete formula, fully incorporating the impact of the emitted or absorbed phonon is described in Ref. [24].

Table 7.1 lists all parameters of unstrained  $\text{Si}_{1-x}\text{Ge}_x$  with various Ge mole fraction  $x$ . The mass density  $\rho$  is a function of the Ge mole fraction  $x$  as given by  $\rho = 2.329 + 3.493x - 0.499x^2$  [25]. The estimated deformation potential of transverse acoustic phonons,  $D_{\text{TA}}$ , for Si is  $2.45 \times 10^{10}$  eV/m [26] and by performing an analogous calculation the value for Ge is  $0.8 \times 10^{10}$  eV/m [27]. The transverse acoustic phonon energy  $\epsilon_{\text{TA}}$  for Si (19 meV) and Ge (8.6 meV) are extracted from Ref. [8]. The  $\epsilon_{\text{TA}}$  of  $\text{Si}_{1-x}\text{Ge}_x$  with various Ge concentrations is determined by linear interpolation since the variation of  $\epsilon_{\text{TA}}$  is almost linear with increasing Ge mole fraction [28]. Note that we only take the transverse acoustic phonons into account because they have the highest phonon occupation number and the smallest phonon energy and therefore provide the main contribution in Si and Ge tunnel diodes [29]. For indirect transitions, the conduction and valence band density of state effective mass  $m_c$  and  $m_v$  are respectively given by  $(m_L m_T^2)^{1/3}$  and  $(m_{\text{lh}}^{3/2} + m_{\text{hh}}^{3/2})^{2/3}$ , where  $m_L$  and  $m_T$  are the electron longitudinal and transverse effective masses and  $m_{\text{lh}}$  and  $m_{\text{hh}}$  are the light and heavy hole effective masses, respectively [30].

Table 7.2 lists the degeneracy factor  $g$ , the reduced tunneling mass  $m_r$  and the lightest electron mass  $m_e$  for direct and indirect BTBT in different tunneling directions. The  $m_r$  is appropriate to the light hole  $m_{\text{lh}}$  and the lightest electron effective mass  $m_e$  along a certain tunneling direction ( $m_r = \frac{m_e m_{\text{lh}}}{m_e + m_{\text{lh}}}$ ) [10, 22] because the lightest carriers have the dominant contribution to the tunneling current. For direct BTBT,  $m_r$  is determined by assuming that the electron effective mass is the same as light-hole effective mass according to  $k \cdot p$  theory [31]. The degeneracy factor  $g$  is expressed as

$$g = 2g_v g_c \quad (7.6)$$

which consists of the electron spin degeneracy factor 2, and the valence and conduction band valley degeneracy. As mentioned above, since we only consider the lightest carriers,  $g_v$  is equal to 1 (light hole band) and  $g_c$  depends on the tunneling direction (see Table 7.2) [32].

Table 7.3 lists the theoretically-calculated  $A_{\text{ind}}$ ,  $B_{\text{ind}}$ ,  $A_{\text{dir}}$  and  $B_{\text{dir}}$  by Eqs. (7.2)-(7.5) for unstrained  $\text{Si}_{1-x}\text{Ge}_x$  with various Ge mole fractions  $x$ . The deviation between our theoretical value  $B_{\text{dir}}$  ( 6.04 MV/cm) for direct transitions in Ge and the prediction of Ref. [33] ( 5.3 MV/cm) and the experimental calibration of Ref. [21] ( 5.7 MV/cm) is small, namely about 14% and 6%, respectively. This deviation is expected to originate mainly from the uncertainty on the value of the reduced mass  $m_r$ . The existing parameters  $A_{\text{exi}}$  and  $B_{\text{exi}}$  provided in the simulation tool are  $4 \times 10^{14} \text{ cm}^{-3}\text{s}^{-1}$  and  $19 \text{ MVcm}^{-1}$  for Si and  $9.1 \times 10^{16} \text{ cm}^{-3}\text{s}^{-1}$  and  $4.9 \text{ MVcm}^{-1}$  for Ge, respectively [18]. The existing parameters of Si are extracted from experimental data of a p-n tunneling diode [34]. The calibration for Ge has been done for the indirect BTBT transitions based on p-n diodes and theoretical calculations [8, 33] although the procedure used by the TCAD software to extract the existing parameters is not fully transparent. In the next subsection, the BTBT generation rate based on existing and theoretically-calculated parameters is compared.

## 7.2.2 BTBT Generation Rate Calculation

The BTBT generation rate as a function of the uniform electric field is plotted in Fig. 7.2 for the indirect and direct band gap transitions in pure Si and pure Ge according to Eq. (7.1). It is expected that the BTBT generation rate for direct band gap transitions dominates over the rate for indirect bandgap transitions, if the energy values of the direct and indirect band gap are the same and if all other relevant parameters, such as the electron and hole masses, are the same. This is because a transition from the valence band to the conduction band which requires interaction with a phonon as in an indirect band gap material is less likely than such a transition which occurs without interaction with a phonon, as in a direct band gap material. The BTBT generation rate of Si for indirect transitions determined by our theoretically calculated parameters and by the existing parameters match very well, while our theoretically calculated direct BTBT rate is negligible compared to the indirect BTBT. This suggests that the indirect BTBT process dominates in Si. The BTBT generation rate determined by our theoretically calculated  $A_{\text{dir}}$ ,  $B_{\text{dir}}$  of Ge is always larger than the rate determined by  $A_{\text{ind}}$ ,  $B_{\text{ind}}$ . This implies that when the band bending is such that direct transitions are allowed in Ge, the direct transitions will dominate over the indirect transitions being in agreement with the experimental data [20, 21]. Fig. 7.2 also shows that for Ge, the BTBT generation rate determined by our

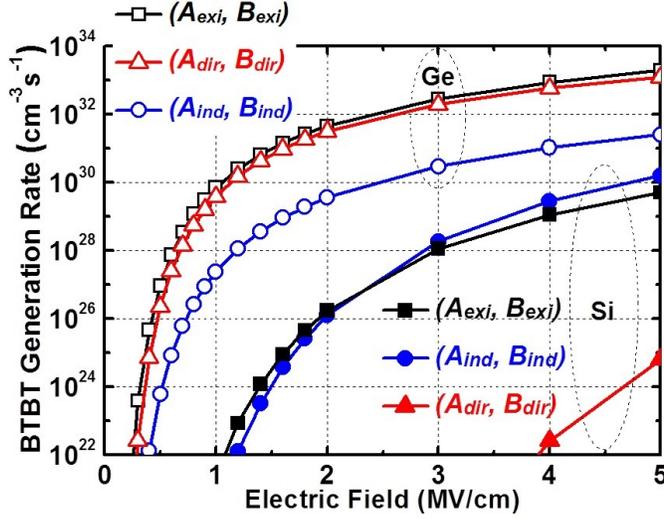


Figure 7.2: BTBT generation rates of Si and Ge as a function of uniform electric field with theoretically calculated  $A_{\text{dir}}$ ,  $B_{\text{dir}}$  and  $A_{\text{ind}}$ ,  $B_{\text{ind}}$  for direct and indirect models and with existing  $A_{\text{exi}}$ ,  $B_{\text{exi}}$  for indirect model. The tunneling direction is along [100].

theoretically calculated  $A_{\text{dir}}$ ,  $B_{\text{dir}}$  matches very well with the rate determined by the existing Ge parameters used in an indirect BTBT setup. This implies that an indirect BTBT model has been widely used by many researchers and by the simulation software, predicting BTBT generation rates with an intensity corresponding to a direct transition, even when the band bending is such that only indirect transitions are allowed.

Fig. 7.3 presents the theoretically calculated BTBT generation rate as a function of uniform electric field for unstrained  $\text{Si}_{1-x}\text{Ge}_x$  with various Ge mole fractions for direct and indirect transitions. At  $x = 0.5$ , the direct BTBT is significantly smaller than the indirect BTBT rate up to 2 MV/cm. At  $x = 0.8$ , however, the direct and indirect BTBT rate have the same order of magnitude, despite the fact that the direct band gap of 1.37 eV is still significantly larger than the indirect band gap of 0.86 eV (see Table 7.1). Note that a relatively pronounced gain of the indirect BTBT generation rate is observed between  $\text{Si}_{0.2}\text{Ge}_{0.8}$  and 100% Ge. This is attributed to an abrupt decrease of the indirect band gap while the  $\text{Si}_{1-x}\text{Ge}_x$  band structure transforms from a Si-like to a Ge-like band structure at  $x = 0.85$  [35], as shown in Fig. 7.4.

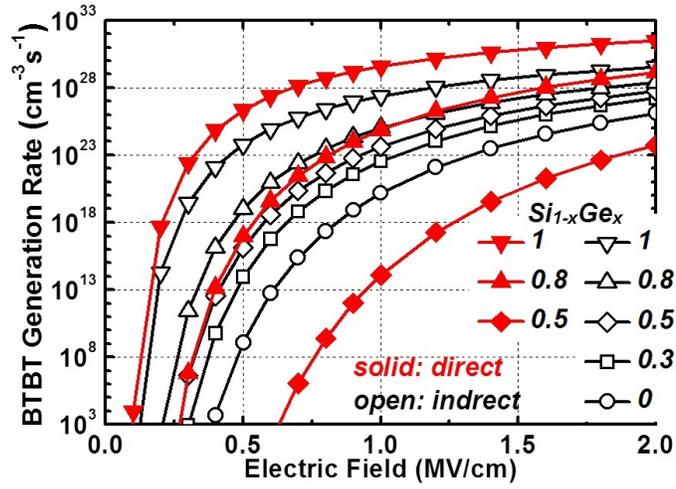


Figure 7.3: BTBT generation rates of  $\text{Si}_{1-x}\text{Ge}_x$  as a function of uniform electric field with theoretically calculated parameters listed in Table 7.3. The tunneling direction is along [100].

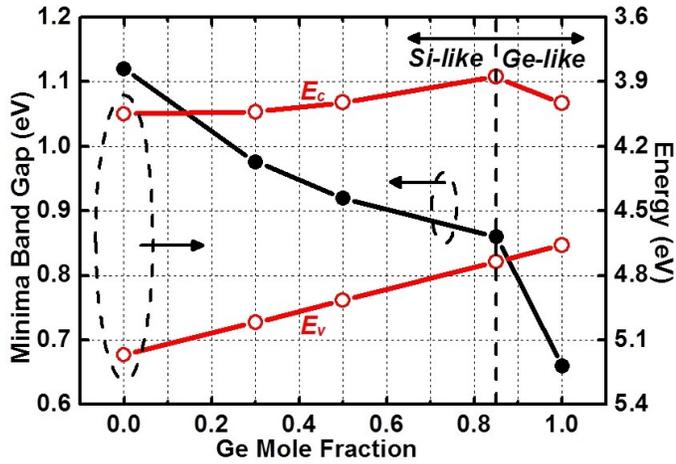


Figure 7.4: The minimum energy band gap and the highest valence and the lowest conduction band of unstrained  $\text{Si}_{1-x}\text{Ge}_x$  as a function of Ge mole fraction.

## 7.3 Simulations and discussion

In this section the impact of a correct use of parameters on performance predictions is illustrated based on device simulations of two realistic TFET configurations. The direct and indirect BTBT models are activated for Si and Ge in the simulations. The calculated  $A_{\text{ind}}$ ,  $B_{\text{ind}}$  and  $A_{\text{dir}}$ ,  $B_{\text{dir}}$  are specified for Ge. For Si, the calculated  $A_{\text{dir}}$ ,  $B_{\text{dir}}$  are specified for the direct BTBT model, but the existing  $A_{\text{exi}}$ ,  $B_{\text{exi}}$  are used for the indirect BTBT model since the deviation from the theoretically calculated parameters is very small, as shown in Fig. 7.2.

### 7.3.1 Simulation Settings

Self-consistent device simulations are performed with SentaaurusDevice [18] using the Fermi-Dirac statistics model, the drift-diffusion carrier transport model, a doping-dependent mobility model, the high field velocity saturation model, the Auger and Shockley-Read-Hall generation/recombination models, the doping-dependent band-gap-narrowing (BGN) model and the dynamic non-local path BTBT model at 300 K. Note that the BTBT model allows tunneling across up to three different band gaps simultaneously. This feature is used to simultaneously determine tunneling across the direct and indirect band gaps. Since the carrier density at the semiconductor-oxide interface is typically quite high, we also include the interface-orientation-dependent modified local-density approximation (MLDA) model to consider the carrier redistribution due to surface quantization effects. At the end of this section, we will also adjust the data to qualitatively include the impact of field-induced quantum confinement [36].

The first simulated structure, a Ge-source Si-TFET, is schematically presented in Fig. 7.5(a) and all device parameters are indicated. An N-type doped drain concentration of  $1 \times 10^{20} \text{ cm}^{-3}$ , a p-type doped source concentration  $N_s$  of  $1 \times 10^{19} \text{ cm}^{-3}$ , a n-doped channel of  $1 \times 10^{15} \text{ cm}^{-3}$  with 40 nm channel length, a 50 nm body thickness, a 0.6 nm effective oxide thickness (EOT), a 30 nm gate-source overlap and a 4.05 eV gate workfunction are specified in the simulations. Abrupt and uniform doping profiles are used. The second simulated structure, an all-Ge TFET (see Fig. 7.8(a)), is identical to the configuration of Fig. 7.5(a), except for a replacement of the channel and drain material with Ge.

### 7.3.2 Results of Semiclassical Simulations

Fig. 7.5(b) shows the input characteristics of the Ge-source nTFET with the architecture shown in Fig. 7.5(a). The large gate-source overlap results in a lateral tunneling current at onset transitioning into a vertical tunneling current as the gate-source voltage increases. This is illustrated in Fig. 7.6 with the generation rate contours: at small gate-source voltage (Fig. 7.6(a)), the

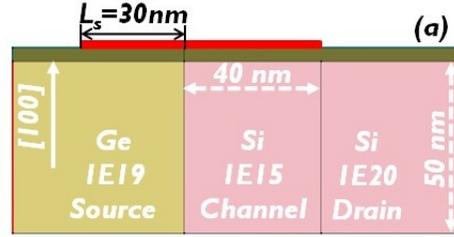


Figure 7.5: (a): The simulated heterostructure device. (b): The input characteristics of (a) obtained by the indirect model with existing parameters (dashed) and theoretically calculated  $A_{\text{ind}}$ ,  $B_{\text{ind}}$  (circle), and by direct and indirect models with  $A_{\text{ind}}$ ,  $B_{\text{ind}}$  and  $A_{\text{dir}}$ ,  $B_{\text{dir}}$  (solid). (c): Decomposition of the solid curve shown in (b) into three components as indicated in the legend and simulated with calculated parameters. (V: vertical, L: lateral)

tunneling is lateral, at intermediate gate-source voltage (Fig. 7.6(b)), vertical tunneling builds up, and at large gate-source voltage (Fig. 7.6(c)), the vertical tunneling is dominant. Fig. 7.5(b) shows that at large gate-source voltages, the indirect tunneling current (circle) is smaller than the direct tunneling current (solid), as expected from Fig. 7.2. However, the direct tunneling has a later onset by nearly 0.2 V because the direct band gap is 0.14 eV larger than the indirect band gap. Comparing to the indirect BTBT current with the existing parameters (dashed), the tunneling current based on a correct use of parameters is therefore significantly reduced at the small gate-source voltages when only indirect BTBT is possible. Furthermore the final tunneling current is also slightly smaller due to the corresponding smaller BTBT generation rate in the uniform field limit (see Fig. 7.2) and the smaller BTBT region resulting from the larger direct band gap. The latter is observable in Fig. 7.6(c) showing a higher BTBT generation rate but smaller BTBT region associated with direct transitions having a longer tunneling length. To visualize this more clearly, Fig. 7.7 presents the band diagram and the BTBT generation rate along a 1-D cut perpendicular to the oxide interface at the middle of the source-gate overlap of Fig. 7.6(c). There is a clear shoulder on the hole generation rate curve (dash) with a shorter tunneling length associated with a smaller band gap: the hole generation area is therefore larger for the indirect band gap transitions. A less pronounced shoulder can also be seen in the electron generation curve (solid) indicating a larger electron generation area for the direct band gap transitions.

In Fig. 7.5(c), the input characteristic with direct and indirect models corresponding to the solid curve of Fig. 7.5(b) is split in its different components: indirect lateral tunneling, indirect vertical tunneling and direct vertical tunneling. The indirect lateral tunneling curve is obtained with a structure identical to Fig. 7.5(a) except for the gate-source overlap which is reduced to 0 nm. In this simulation, only BTBT across the indirect bandgap is allowed. The

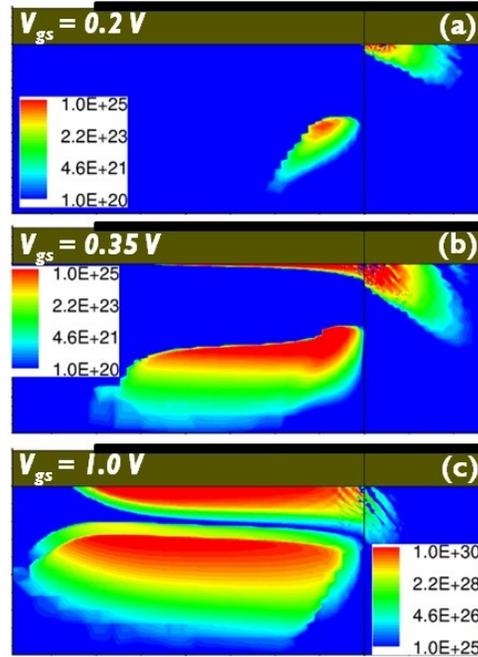


Figure 7.6: The electron and hole BTBT generation rate contours ( $\text{cm}^{-3} \text{ s}^{-1}$ ) of the n-channel device of Fig. 7.5 (a) are shown in (a) for  $V_{gs} = 0.2 \text{ V}$ , in (b) for  $V_{gs} = 0.35 \text{ V}$ , and in (c) for  $V_{gs} = 1 \text{ V}$  resulting from simulations with direct and indirect BTBT models corresponding to the solid curve in Fig. 7.5 (b). The drawing, which is a segment of the full TFET, is to scale with the 30 nm gate-source overlap as size reference.

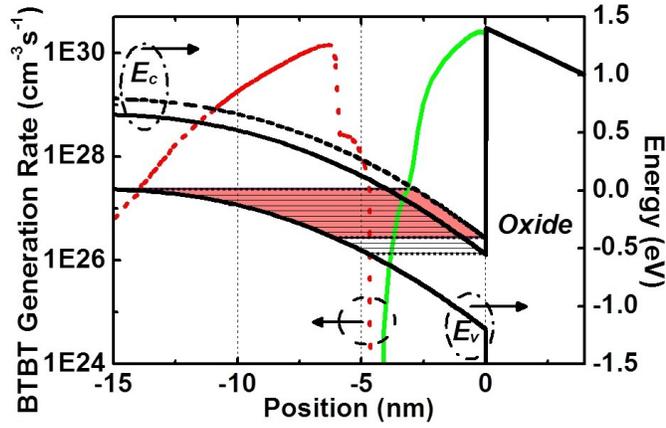


Figure 7.7: Energy band diagram (right vertical axis) and BTBT generation rate (left vertical axis) of electrons (solid) and holes (dash) of the configuration of Fig. 7.6(c) along a 1-D cut perpendicular to the oxide interface at the middle of the source-gate overlap. The (in)direct conduction band is denoted with a dash (solid) curve. The shaded (striped) region shows the (in)direct tunneling with longer (shorter) tunnel lengths and narrower (wider) generation region.



Figure 7.8: (a): The simulated all-Ge TFET. (b): The input characteristics of the all-Ge TFET obtained by the indirect model with existing parameters (dashed) and theoretically calculated  $A_{ind}$ ,  $B_{ind}$  (circle), and by direct and indirect models with  $A_{ind}$ ,  $B_{ind}$  and  $A_{dir}$ ,  $B_{dir}$  (solid). (c): Decomposition of the solid curve shown in (a) into four components as indicated in the legend and simulated with calculated parameters. (V:vertical, L:lateral)

(in)direct vertical tunneling curve is obtained with a structure identical to Fig. 7.5(a) except for the gate-channel overlap which is reduced to 0 nm. In this simulation, only BTBT across the (in)direct bandgap is allowed. The sum of the subcomponents (not shown) is very close to the solid curve. This implies that the modified gate electrode and the tunneling current itself only have a limited impact on the electrostatic profile and therefore result in a very limited modification of the device current. The latter is in agreement with previous work, where the tunneling current prediction based on a non-selfconsistent analytical model is in very good agreement with the selfconsistently simulated current as long as the current is not very high [37].

Note that at a gate-source voltage of about 0.3 V, there is a small kink in the dashed and solid curves of Fig. 7.5(b), which results from the transition from lateral tunneling leakage from source to channel into vertical tunneling in the source perpendicular to the gate, as can be observed from the decomposition in Fig. 7.5(c). This lateral leakage current is tunneling across a heterojunction. Since the simulator is limited to either take a direct BTBT model or an indirect BTBT model along the entire tunneling path [18], and since the direct band gap of Si is too high to play a role, the lateral tunneling current is based on the indirect BTBT parameters in both materials. More insight in heterojunction tunneling is needed to determine whether it would not be more meaningful to use the direct-band-gap Ge parameters for the heterojunction tunneling.

Fig. 7.8(b) shows the input characteristics of the all-Ge nTFET with the structure shown in Fig. 7.8(a). Similar results to Fig. 7.5(b) are observed at high gate-source voltages. At low gate-source voltages, the current is larger due to the larger drain-source leakage current and the increased lateral tunneling which are both due to the smaller band gap of Ge together with the availability of direct BTBT transitions into channel and drain. Fig. 7.8(c) shows a split of the solid curve of Fig. 7.8(b) in its different components: direct and indirect lateral tunneling and direct and indirect vertical tunneling. The different curves are obtained in a similar way to the curves of Fig. 7.5(c). This split-up shows that the kink at 0.5 V in the solid curve is associated with the transition from direct lateral BTBT to direct vertical BTBT. Direct lateral tunneling is dominant over indirect lateral tunneling, even near onset of lateral tunneling. The earlier onset by about 0.14 V of the indirect component is present at very low current levels but not observable in the configuration with 40 nm channel length. For a device with 80 nm channel length, the leakage currents are smaller, and it can be seen that the indirect lateral tunneling has an onset which is about 0.14 V earlier than the direct lateral tunneling (not shown). Note that the predictions of the lateral tunneling component in Fig. 7.5(c) and Fig. 7.8(c) should not be compared quantitatively, since the direct Ge bandgap can not contribute to the lateral tunneling in the heterostructure due to a limitation of the simulation software as discussed before and since the stress associated with the Ge-Si heterostructure is not included in the analysis [6, 38, 39]. The latter

choice is made because stress is not properly implemented in the BTBT model and an in-depth study of the impact of the stress is beyond the scope of this article. In particular, the simulation software uses the specified parameters  $A$  and  $B$  of the BTBT model and these parameters are constants along the tunneling paths despite the fact that the band gap and effective mass are varying in a heterostructure configuration due to the non-uniform stress.

The delayed onset of direct BTBT versus indirect BTBT with about 0.14 V is also expected in the presence of non-abrupt doping profiles in the source (e.g. one-dimensional doping gradient along the gate dielectric interface). Note that this might not be observable as clear kinks in the characteristics due to the degradation of the subthreshold slope and the averaging effect associated with the graded doping profile.

### 7.3.3 Impact of Field-Induced Quantum Confinement

The modified local-density approximation (MLDA) model is included in simulations to consider quantum confinement near an oxide interface, which reduces the local carrier density and decreases the tunneling current. In fact field-induced quantum confinement (FIQC) also creates subbands and gradually raises the subband energy with increasing gate bias, which has been shown to significantly delay the onset of vertical tunneling [36]. We will now include FIQC in our analysis. According to Ref. [36], the ground-state subband energy of the  $\alpha$  valley  $E_{\text{sub}}^{\alpha}$  in the triangular well approximation is given by

$$E_{\text{sub}}^{\alpha} = -a_0 \left( \frac{\hbar^2 F_{\text{ox},\alpha}}{8\pi^2 m_{\alpha}^*} \right)^{1/3} \quad (7.7)$$

where  $a_0 \approx -2.3381$  is the first zero of the Airy function,  $m_{\alpha}^*$  is the electron effective mass in the  $\alpha$  valley. The force acting on the electron at the semiconductor-oxide interface  $F_{\text{ox},\alpha}$  at the onset of vertical tunneling is

$$F_{\text{ox},\alpha} = \sqrt{\frac{2q^2 N_s}{\epsilon_s} (E_g^{\alpha} + E_{\text{sub}}^{\alpha})} \quad (7.8)$$

where  $\epsilon_s$  is the semiconductor permittivity and  $E_g^{\alpha}$  is the tunneling band gap between the valence band and the  $\alpha$  valley of the conduction band. Fig. 7.9(a) shows the subband energy for two different electron effective masses at the onset of vertical tunneling, this is when the subband energy level aligns with the valence band level along the [100] as a function of  $E_g^{\alpha}$ . The subband energy decreases with decreasing  $E_g^{\alpha}$  due to the weaker electric field at the onset of vertical tunneling. The subband energy increases with decreasing electron effective mass as a result of stronger quantum confinement. Note that

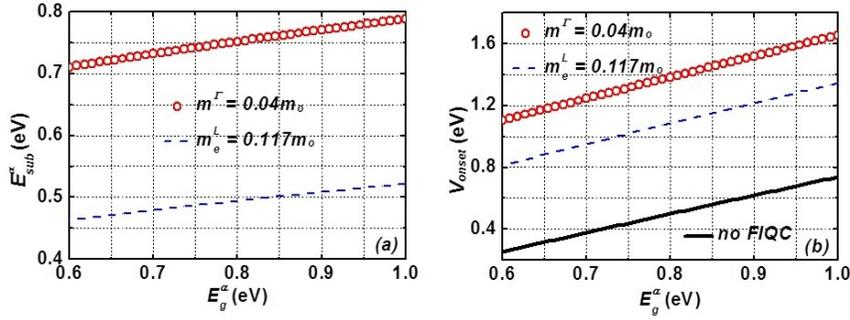


Figure 7.9: (a): The subband energy at the onset of vertical tunneling and (b): the onset voltage of vertical tunneling with different electron effective masses as a function of energy band gap.

$m_e^\Gamma = 0.044m_0$  and  $m_e^L = 0.117m_0$  are the electron effective mass at the  $\Gamma$  valley (direct band gap) and electron effective mass in the  $L$  valley of Ge along the [100] tunneling direction, respectively (see Table 7.2).

For the device operation, the onset voltage also needs to be determined (see Fig. 7.9(b)). The onset voltage  $V_{onset}^\alpha$  of the vertical tunneling can be written as

$$V_{onset}^\alpha = E_g^\alpha + E_{sub}^\alpha + F_{ox,\alpha} T_{ox} \frac{\epsilon_s}{\epsilon_{ox}} - V_{FB} \quad (7.9)$$

where  $V_{FB}$  is the flat band voltage between the gate workfunction and the Fermi level of the source region. The third term in Eq. (7.9) is the potential drop across the oxide, in which  $T_{ox}$  and  $\epsilon_{ox}$  are the physical oxide thickness and oxide permittivity, respectively. As shown in Fig. 7.9(b), the onset voltage of vertical tunneling is decreased with decreasing  $E_g^\alpha$  and increasing electron effective mass. The differences in onset voltage when including FIQC are  $\Delta V_{onset} = 0.563V$  at  $E_g^L = 0.66eV - \Delta E_{BGN} = 0.645eV$ , and  $\Delta V_{onset}^\Gamma = 0.883V$  at  $E_g^\Gamma = 0.8eV - \Delta E_{BGN} = 0.785eV$ . (In the  $\Gamma$  and  $L$  valleys,  $E_g^\alpha$  are extracted from simulations, which includes doping-induced band gap narrowing effect  $\Delta E_{BGN} = 0.015eV$  for  $N_s = 1E19cm^{-3}$ .)

Manually shifting the vertical direct and indirect tunneling components shown in Fig. 7.5(c) by , Fig. 7.10(a) shows the input characteristics of the Ge-source Si TFET considering FIQC. The dashed line with existing parameters is the same as in Fig. 7.5(b) and is added to highlight the difference in predictions when incorrect assumptions are made. The solid line is the sum of the lateral tunneling and the shifted vertical tunneling from the valence band to the first subband determined by the light electron effective mass  $\frac{3m_L m_\Gamma}{2m_L + m_\Gamma}$  in [100] direction (see Table 7.2) in the  $L$  valley (indirect BTBT) and the first subband in the  $\Gamma$  valley (direct BTBT). It is clear that predictions are significantly af-

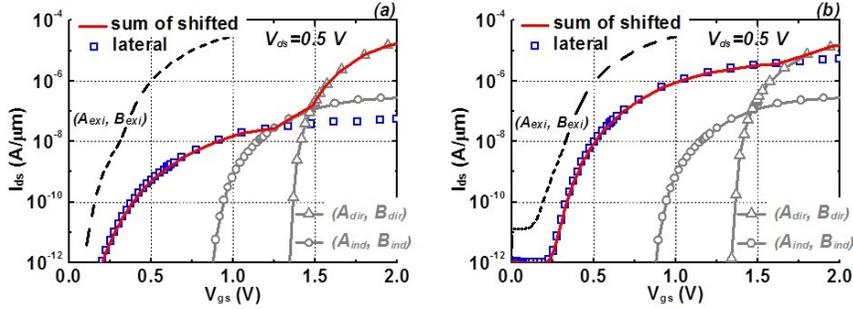


Figure 7.10: Input characteristics including the shifted vertical tunneling components according to the back of the envelope calculation of FIQC (see Fig. 7.9) for (a): Ge-source Si TFETs and (b): all Ge TFETs.

ected by correctly including the direct band gap of Ge. The onset of direct tunneling is delayed by two effects, firstly the delay of about 0.2 V due to the larger band gap as discussed above, and secondly the relative delay compared to the indirect tunneling of about 0.3 V due to the smaller effective tunneling mass at  $\Gamma$  resulting in a larger subband energy derived in Fig. 7.9(b). This last relative shift allows to observe pure indirect tunneling in Ge over a larger range of electric fields than is possible in a Ge diode, and therefore, calibration of the indirect BTBT in Ge should be more straightforward with the TFET configuration.

The same manual shifting has been applied to the all-Ge TFET as shown in Fig. 7.10(b). The vertical indirect BTBT associated with the 1st subband in the L valley is lower than the lateral direct BTBT. The transition from lateral to vertical BTBT can therefore only be observed for the 1st subband in the  $\Gamma$  valley.

Note that the orientation of the lateral tunneling path with respect to the gate oxide varies with gate bias. Lateral tunneling starts with a long tunneling path which is rather parallel to the oxide-channel interface at low gate bias, such that less FIQC is expected. With increasing gate bias, the tunneling direction tilts more towards the gate oxide, and FIQC is expected to increase. Given that a varying impact of FIQC is not realistic with the available software [40], we have assumed for the calculations in Fig. 7.10 that there is no FIQC for the lateral tunneling.

## 7.4 Conclusions

We theoretically calculate the parameters  $A$  and  $B$  of Kane's direct and indirect band-to-band tunneling models for different tunneling directions ([100], [110] and [111]) for Si, Ge and unstrained  $\text{Si}_{1-x}\text{Ge}_x$ . A relatively pronounced gain in the indirect BTBT generation rate between  $\text{Si}_{0.2}\text{Ge}_{0.8}$  and Ge is attributed to an abrupt decrease of band gap while  $\text{Si}_{1-x}\text{Ge}_x$  transforms from Si-like to Ge-like material at  $x = 0.85$ . The direct BTBT contribution in unstrained  $\text{Si}_{1-x}\text{Ge}_x$  becomes equally important to the indirect BTBT when the Ge concentration is about 80%. For 100% Ge, the calculation of BTBT generation rate in the uniform electric field limit reveals that direct tunneling always dominates. We show that less optimistic predictions for a Ge-source TFET upon using the correct parameters and including the Ge direct band gap. The onset of direct versus indirect BTBT is delayed by about 0.5 V in the vertical tunneling component due to the larger direct band gap and the larger direct subband energy resulting from field-induced quantum confinement. The Ge-source TFET should therefore be a more appropriate vehicle to calibrate the strength of the indirect Ge BTBT process than Ge diodes, where calibration of the indirect branch is limited to a 0.14 V voltage window.

## Bibliography

- [1] W. Y. Choi, B.-G. Park, J. D. Lee, and T.-J. K. Liu, "Tunneling field-effect transistors (TFETs) with subthreshold swing (SS) less than 60 mv/dec," *Electron Device Letters, IEEE*, vol. 28, no. 8, pp. 743–745, aug. 2007.
- [2] F. Mayer, C. Le Royer, J.-F. Damlencourt, K. Romanjek, F. Andrieu, C. Tabone, B. Previtali, and S. Deleonibus, "Impact of soi, si1-xgexoi and geoi substrates on cmos compatible tunnel fet performance," in *Electron Devices Meeting, 2008. IEDM 2008. IEEE International*, dec. 2008, pp. 1–5.
- [3] K. Boucart and A. M. Ionescu, "A new definition of threshold voltage in tunnel FETs," *Solid-State Electronics*, vol. 52, no. 9, pp. 1318–1323, 2008, papers Selected from the 37th European Solid-State Device Research Conference - ESSDERC'07. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038110108001196>
- [4] D. Leonelli, A. Vandooren, R. Rooyackers, A. S. Verhulst, S. D. Gendt, M. M. Heyns, and G. Groeseneken, "Performance enhancement in multi gate tunneling field effect transistors by scaling the fin-width," *Japanese Journal of Applied Physics*, vol. 49, no. 4, p. 04DC10, 2010. [Online]. Available: <http://jjap.jsap.jp/link?JJAP/49/04DC10/>

Table 7.1: Physical parameters used in theoretical calculations of  $A_{\text{dir}}$ ,  $B_{\text{dir}}$  and  $A_{\text{ind}}$ ,  $B_{\text{ind}}$  for  $\text{Si}_{1-x}\text{Ge}_x$  at various Ge mole fractions. All effective masses are in the unit of the free electron mass  $m_0$ . Note that the electron effective masses  $m_{\text{T}}$  and  $m_{\text{L}}$  are constant in  $0 \leq x \leq 0.8$  and the hole effective masses  $m_{\text{hh}}$  and  $m_{\text{lh}}$  decrease linearly with increasing  $x$ .

$x$	$\rho$ ( $\text{kg}/\text{m}^3$ )	$D_{\text{TA}}$ ( $\text{eV}/\text{m}$ )	$\epsilon_{\text{TA}}$ (meV)	$E_{\text{g}}$ (eV)	$E_{\text{g}}^{\text{I}}$ (eV)	$m_{\text{L}}$	$m_{\text{T}}$	$m_{\text{lh}}$	$m_{\text{hh}}$	$m_{\text{c}}$	$m_{\text{v}}$
0	2329	$2.45 \times 10^{10}$	19	1.12	3.4	0.98	0.19	0.16	0.49	0.328	0.549
0.3	3332	$1.96 \times 10^{10}$	15.88	0.976	2.79	0.98	0.19	0.125	0.442	0.328	0.485
0.5	3951	$1.63 \times 10^{10}$	13.8	0.92	2.19	0.98	0.19	0.102	0.41	0.328	0.443
0.8	4804	$1.13 \times 10^{10}$	10.68	0.86	1.37	0.98	0.19	0.067	0.362	0.328	0.381
1	5323	$0.80 \times 10^{10}$	8.6	0.66	0.8	1.6	0.08	0.044	0.33	0.217	0.341
Ref.	[25]	[26,27]	[8,28]	[35]	[19,35]	[30,41]	[30,41]	[41,42]	[41,42]	[30]	[30]

Table 7.2: Degeneracy factor  $g$ , electron effective mass  $m_e$  and the reduced tunneling mass  $m_r$  used in theoretical calculations of  $A_{\text{dir}}$ ,  $B_{\text{dir}}$  and  $A_{\text{ind}}$ ,  $B_{\text{ind}}$  for  $\text{Si}_{1-x}\text{Ge}_x$  in different tunneling directions. Note that band structure of bulk  $\text{Si}_{1-x}\text{Ge}_x$  exhibits a Si-like (Ge-like) conduction band minima at  $\Delta$  (L) with six (eight) equivalent valleys when the Ge mole fraction,  $x$ , is less (more) than 0.85. All effective masses are in unit of the free electron mass  $m_0$ .

$x$	Direct BTBT		Indirect BTBT											
	$g$	$m_r$	[100]			[110]			[111]					
			$g$	$m_e$	$m_r$	$g$	$m_e$	$m_r$	$g$	$m_e$	$m_r$			
0	2	0.08	8	$m_T$	0.087	4	$m_T$	0.087	12	$\frac{3m_L m_T}{2m_L + m_T}$	0.099	12	$\frac{3m_L m_T}{2m_L + m_T}$	0.099
0.3	2	0.063	8	$m_T$	0.075	4	$m_T$	0.075	12	$\frac{3m_L + m_T}{2m_L + m_T}$	0.084	12	$\frac{3m_L + m_T}{2m_L + m_T}$	0.084
0.5	2	0.051	8	$m_T$	0.066	4	$m_T$	0.066	12	$\frac{3m_L m_T}{2m_L + m_T}$	0.073	12	$\frac{3m_L m_T}{2m_L + m_T}$	0.073
0.8	2	0.034	8	$m_T$	0.050	4	$m_T$	0.050	12	$\frac{3m_L m_T}{2m_L + m_T}$	0.053	12	$\frac{3m_L m_T}{2m_L + m_T}$	0.053
1	2	0.022	8	$\frac{3m_L m_T}{2m_L + m_T}$	0.032	4	$m_T$	0.028	6	$\frac{9m_L m_T}{8m_L + m_T}$	0.029	6	$\frac{9m_L m_T}{8m_L + m_T}$	0.029

Table 7.3: The theoretically calculated and existing parameters  $A$  and  $B$  for nonlocal path BTBT model for unstrained  $\text{Si}_{1-x}\text{Ge}_x$  at various Ge mole fractions.  $A$  and  $B$  are in unit of  $\text{cm}^{-3}\text{s}^{-1}$  and  $\text{MVcm}^{-1}$ , respectively.

$x$	Direct BTBT		Indirect BTBT						$A_{\text{exi}}$	$B_{\text{exi}}$
	$A_{\text{dir}}$	$B_{\text{dir}}$	[100]		[110]		[111]			
			$A_{\text{ind}}$	$B_{\text{ind}}$	$A_{\text{ind}}$	$B_{\text{ind}}$	$A_{\text{ind}}$	$B_{\text{ind}}$		
0	$1.35 \times 10^{20}$	101	$3.29 \times 10^{15}$	23.8	$1.64 \times 10^{15}$	23.8	$4.19 \times 10^{15}$	25.5	$4 \times 10^{14}$	19
0.3	$1.33 \times 10^{20}$	66.5	$2.61 \times 10^{15}$	18.1	$1.30 \times 10^{15}$	18.1	$3.40 \times 10^{15}$	19.1		
0.5	$1.35 \times 10^{20}$	41.6	$2.27 \times 10^{15}$	15.5	$1.13 \times 10^{15}$	15.5	$3.01 \times 10^{15}$	16.3		
0.8	$1.39 \times 10^{20}$	16.8	$1.92 \times 10^{15}$	12.1	$9.62 \times 10^{14}$	12.1	$2.64 \times 10^{15}$	12.6		
1	$1.46 \times 10^{20}$	6.04	$1.67 \times 10^{15}$	6.55	$9.7 \times 10^{14}$	6.17	$1.39 \times 10^{15}$	6.29	$9.1 \times 10^{16}$	4.9

- [5] T. Krishnamohan, D. Kim, C. Jungemann, Y. Nishi, and K. Saraswat, "Strained-si, relaxed-ge or strained-(si)ge for future nanoscale p-mosfets?" in *VLSI Technology, 2006. Digest of Technical Papers. 2006 Symposium on*, 0-0 2006, pp. 144 –145.
- [6] O. Nayfeh, C. Chleirigh, J. Hennessy, L. Gomez, J. Hoyt, and D. Antoniadis, "Design of tunneling field-effect transistors using strained-silicon/strained-germanium type-ii staggered heterojunctions," *Electron Device Letters, IEEE*, vol. 29, no. 9, pp. 1074 –1077, sept. 2008.
- [7] S. H. Kim, H. Kam, C. Hu, and T.-J. K. Liu, "Germanium-source tunnel field effect transistors with record high ion/loff," in *VLSI Technology, 2009 Symposium on*, june 2009, pp. 178 –179.
- [8] M. Fischetti, T. O'Regan, S. Narayanan, C. Sachs, S. Jin, J. Kim, and Y. Zhang, "Theoretical study of some physical aspects of electronic transport in nmosfets at the 10-nm gate-length," *Electron Devices, IEEE Transactions on*, vol. 54, no. 9, pp. 2116 –2136, sept. 2007.
- [9] C. Hu, D. Chou, P. Patel, and A. Bowonder, "Green transistor - a vdd scaling path for future low power ics," in *VLSI Technology, Systems and Applications, 2008. VLSI-TSA 2008. International Symposium on*, april 2008, pp. 14 –15.
- [10] A. S. Verhulst, W. G. Vandenberghe, K. Maex, and G. Groeseneken, "Boosting the on-current of a n-channel nanowire tunnel field-effect transistor by source material optimization," *Journal of Applied Physics*, vol. 104, no. 6, p. 064514, 2008. [Online]. Available: <http://link.aip.org/link/?JAP/104/064514/1>
- [11] A. Verhulst, W. Vandenberghe, K. Maex, S. De Gendt, M. Heyns, and G. Groeseneken, "Complementary silicon-based heterostructure tunnel-fets with high tunnel rates," *Electron Device Letters, IEEE*, vol. 29, no. 12, pp. 1398 –1401, dec. 2008.
- [12] S. Mookerjea and S. Datta, "Comparative study of si, ge and inas based steep subthreshold slope tunnel transistors for 0.25v supply voltage logic applications," in *Device Research Conference, 2008*, june 2008, pp. 47 –48.
- [13] S. H. Kim, S. Agarwal, Z. Jacobson, P. Matheu, C. Hu, and T.-J. Liu, "Tunnel field effect transistor with raised germanium source," *Electron Device Letters, IEEE*, vol. 31, no. 10, pp. 1107 –1109, oct. 2010.
- [14] G. Han, P. Guo, Y. Yang, L. Fan, Y. S. Yee, C. Zhan, and Y.-C. Yeo, "Source engineering for tunnel field-effect transistor: Elevated source with vertical silicon-germanium/germanium heterostructure," *Japanese Journal of Applied Physics*, vol. 50, no. 4, p. 04DJ07, 2011. [Online]. Available: <http://jjap.jsap.jp/link?JJAP/50/04DJ07/>

- [15] G. Han, P. Guo, Y. Yang, C. Zhan, Q. Zhou, and Y.-C. Yeo, "Silicon-based tunneling field-effect transistor with elevated germanium source formed on (110) silicon substrate," *Applied Physics Letters*, vol. 98, no. 15, pp. 153 502 –153 502–3, apr 2011.
- [16] Q. Zhang, S. Sutar, T. Kosel, and A. Seabaugh, "Fully-depleted ge interband tunnel transistor: Modeling and junction formation," *Solid-State Electronics*, vol. 53, no. 1, pp. 30 – 35, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038110108003213>
- [17] Y. Yang, P.-F. Guo, G.-Q. Han, C.-L. Zhan, L. Fan, and Y.-C. Yeo, "Drive current enhancement with invasive source in double gate tunneling field-effect transistors," in *International Conference on Solid State Devices and Materials*, 2010, pp. 798–799.
- [18] *Sentaurus Device*, Synopsis, version D-2010.03.
- [19] O. Madelung, *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology*. Springer-Verlag, 1982.
- [20] A. G. Chynoweth, W. L. Feldmann, C. A. Lee, R. A. Logan, G. L. Pearson, and P. Aigrain, "Internal field emission at narrow silicon and germanium  $p - n$  junctions," *Phys. Rev.*, vol. 118, pp. 425–434, Apr 1960. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRev.118.425>
- [21] M. S. Tyagi, "Determination of effective mass and the pair production energy for electrons in germanium from zener diode characteristics," *Japanese Journal of Applied Physics*, vol. 12, no. 1, pp. 106–108, 1973. [Online]. Available: <http://jjap.jsap.jp/link?JJAP/12/106/>
- [22] E. O. Kane, "Zener tunneling in semiconductors," *Journal of Physics and Chemistry of Solids*, vol. 12, pp. 181–188, 1959.
- [23] W. Vandenberghe, B. Soree, W. Magnus, and G. Groeseneken, "Zener tunneling in semiconductors under nonuniform electric fields," *Journal of Applied Physics*, vol. 107, no. 5, pp. 054 520 –054 520–7, mar 2010.
- [24] W. Vandenberghe, B. Soree, W. Magnus, and M. V. Fischetti, "Generalized phonon-assisted zener tunneling in indirect semiconductors with non-uniform electric fields: A rigorous approach," *Journal of Applied Physics*, vol. 109, no. 12, pp. 124 503 –124 503–12, jun 2011.
- [25] M. E. Levinshtein, S. L. Rumyantsev, and M. S. Shur, *Properties of Advanced Semiconductor Materials: GaN, AlN, InN, BN, SiC, SiGe*. Wiley, 2001.
- [26] C. Rivas, R. Lake, G. Klimeck, W. R. Frensley, M. V. Fischetti, P. E. Thompson, S. L. Rommel, and P. R. Berger, "Full-band simulation of

- indirect phonon assisted tunneling in a silicon tunnel diode with delta-doped contacts,” *Applied Physics Letters*, vol. 78, no. 6, pp. 814–816, feb 2001.
- [27] C. Rivas, “Full bandstructure modeling of indirect phonon assisted transport in tunnel devices with silicon contacts,” Ph.D. dissertation, The University of Texas at Dallas, 2004.
- [28] R. A. Logan, J. M. Rowell, and F. A. Trumbore, “Phonon spectra of ge-si alloys,” *Phys. Rev.*, vol. 136, pp. A1751–A1755, Dec 1964. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRev.136.A1751>
- [29] A. G. Chynoweth, R. A. Logan, and D. E. Thomas, “Phonon-assisted tunneling in silicon and germanium esaki junctions,” *Phys. Rev.*, vol. 125, pp. 877–881, Feb 1962. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRev.125.877>
- [30] S. M. Sze, *Physics of Semiconductor Devices*. Wiley, 1981.
- [31] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors: Physics and Materials Properties*. Springer-Verlag, 2001.
- [32] A. Rahman, M. S. Lundstrom, and A. W. Ghosh, “Generalized effective-mass approach for n-type metal-oxide-semiconductor field-effect transistors on arbitrarily oriented wafers,” *Journal of Applied Physics*, vol. 97, no. 5, pp. 053 702–053 702–12, mar 2005.
- [33] D. Kim, T. Krishnamohan, L. Smith, H.-S. Wong, and K. Saraswat, “Band to band tunneling study in high mobility materials : Iii-v, si, ge and strained sige,” in *Device Research Conference, 2007 65th Annual*, june 2007, pp. 57–58.
- [34] G. Hurkx, D. Klaassen, and M. Knuvers, “A new recombination model for device simulation including tunneling,” *Electron Devices, IEEE Transactions on*, vol. 39, no. 2, pp. 331–338, feb 1992.
- [35] R. Braunstein, A. R. Moore, and F. Herman, “Intrinsic optical absorption in germanium-silicon alloys,” *Phys. Rev.*, vol. 109, pp. 695–710, Feb 1958. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRev.109.695>
- [36] W. G. Vandenberghe, B. Soree, W. Magnus, G. Groeseneken, and M. V. Fischetti, “Impact of field-induced quantum confinement in tunneling field-effect devices,” *Applied Physics Letters*, vol. 98, no. 14, pp. 143 503–143 503–3, apr 2011.
- [37] A. S. Verhulst, D. Leonelli, R. Rooyackers, and G. Groeseneken, “Drain voltage dependent analytical model of tunnel field-effect transistors,” *Journal of Applied Physics*, vol. 110, no. 2, pp. 024 510–024 510–10, jul 2011.

- [38] K. Boucart, W. Riess, and A. Ionescu, "Lateral strain profile as key technology booster for all-silicon tunnel fets," *Electron Device Letters, IEEE*, vol. 30, no. 6, pp. 656–658, june 2009.
- [39] P. Solomon, I. Lauer, A. Majumdar, J. Teherani, M. Luisier, J. Cai, and S. Koester, "Effect of uniaxial strain on the drain current of a heterojunction tunneling field-effect transistor," *Electron Device Letters, IEEE*, vol. 32, no. 4, pp. 464–466, april 2011.
- [40] W. Vandenberghe, B. Soree, W. Magnus, G. Groeseneken, A. Verhulst, and M. Fischetti, "Field induced quantum confinement in indirect semiconductors: Quantum mechanical and modified semiclassical model," in *Simulation of Semiconductor Processes and Devices (SISPAD), 2011 International Conference on*, sept. 2011, pp. 271–274.
- [41] M. V. Fischetti and S. E. Laux, "Band structure, deformation potentials, and carrier mobility in strained si, ge, and sige alloys," *Journal of Applied Physics*, vol. 80, no. 4, pp. 2234–2252, aug 1996.
- [42] S. Chun and K. Wang, "Effective mass and mobility of holes in strained si<sub>1-x</sub>ge<sub>x</sub> layers on (001) si<sub>1-y</sub>ge<sub>y</sub> substrate," *Electron Devices, IEEE Transactions on*, vol. 39, no. 9, pp. 2153–2164, sep 1992.

### List of corrections and clarifications

- The voltages in Eq. (7.9) must be converted into energies by multiplying by the unit charge  $q$  such that Eq. (7.9) should read

$$qV_{\text{onset}}^{\alpha} = E_{\text{g}}^{\alpha} + E_{\text{sub}}^{\alpha} + F_{\text{ox},\alpha} T_{\text{ox}} \frac{\epsilon_{\text{s}}}{\epsilon_{\text{ox}}} - qV_{\text{FB}}. \quad (7.10)$$



## Chapter 8

# Optimal doping and bandgap for TFETs

A model determining optimal doping concentration and material's band gap of tunnel field-effect transistors

William G. Vandenberghe, Anne S. Verhulst, Kuo-Hsing Kao,  
Kristin De Meyer, Bart Sorée, Wim Magnus and  
Guido Groeseneken

Published in *Applied Physics Letters*, 100, art nr. 193509, May 2012.

The following article appeared in *Applied Physics Letters*, 100, art. nr. 193509, May 2012 and may be found at <http://link.aip.org/link/doi/10.1063/1.4714544>.

©2012 American Institute of Physics. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the American Institute of Physics.

### **Contributions of first author**

- Conception and development of the theory
- Development of the code
- Co-creating figures
- Text co-writing and editing

## Abstract

We develop a model for the tunnel field-effect transistor (TFET) based on the Wentzel-Kramer-Brillouin approximation which improves over existing semi-classical models employing generation rates. We hereby introduce the concept of a characteristic tunneling length in direct semiconductors. Based on the model, we show that a limited density of states results in an optimal doping concentration as well as an optimal material's band gap to obtain the highest TFET on-current at a given supply voltage. The observed optimal-doping trend is confirmed by 2-dimensional quantum-mechanical simulations for silicon and germanium.

Because of their ability to have a steeper than 60mV/decade subthreshold swing, TFETs have been put forward as future low-power devices [1–5]. Significant advances have been made in understanding and designing TFETs. However, up to now, it has been assumed that the highest doping and the smallest material's band gap will result in the largest gate-controlled tunneling current [6, 7].

We have upgraded our TFET model with Fermi-Dirac statistics and we have linked the available states for tunneling to the effective tunneling gap by replacing the generation rate approach with a Wentzel-Kramer-Brillouin (WKB) approach [8, 9]. Based on this model, we show the existence of an optimal source doping concentration as well as an optimal material's band gap, which are both strongly dependent on the effective density-of-states. We limit the investigation to an  $n$ TFET, for  $p$ TFET qualitative similar conclusions are expected. Transport through the ungated channel, trap-assisted tunneling currents or other leakage mechanisms are not considered. Our model does not take quantum confinement (QC) [10] into account either, but our results are not impacted by the QC-induced voltage shift and at the end of the paper we support our conclusions with quantum mechanical simulations.

The device of Fig. 8.1 with the gate on top of the source is studied. In the on-state, the source is depleted and tunneling occurs perpendicular to the gate dielectric. This process is referred to as line tunneling [8]. Fig. 8.2 shows a band diagram along the cross-section A-A'. Tunneling sets in when the surface potential energy  $\phi_s$  exceeds the band gap  $E_g$ .

In the WKB approximation [11], the direct band-to-band tunneling probability at an energy level  $E$  is described by

$$T(E) = \exp\left(-2 \int_{x_1(E)}^{x_2(E)} \kappa(x, E) dx\right) \quad (8.1)$$

where  $x_1(E)$  and  $x_2(E)$  are the starting and ending points of the tunnel path determining the tunnel path length  $l(E) = |x_2(E) - x_1(E)|$  and  $\kappa(x, E)$  is the imaginary part of the wavevector inside the band gap. To determine  $\kappa(x, E)$  in a direct semiconductor, the two band  $\mathbf{k} \cdot \mathbf{p}$  model [7, 12, 13] can be used which presents an improvement over the triangular well approximation used in Ref. [14] [15]. In the presence of an external potential  $U(x)$  and neglecting the small difference between valence and conduction band curvature in the tunneling direction,  $\kappa$  is given by:

$$\begin{aligned} \kappa(x, E) &= \kappa(U(x) - E) \\ &= \frac{m_0}{2\hbar p} \sqrt{E_g^2 - 4 \left( U(x) - E + \frac{E_g}{2} \right)^2} \end{aligned} \quad (8.2)$$

with  $m_0$  the free electron mass and with the top of the valence band corresponding to  $E = U(x)$  [13]. The momentum matrix element  $p$  is usually specified in energy units according to  $E_p = 2p^2/m_0$ . For all semiconductors with the zincblende lattice structure  $E_p \approx 20$  eV [16] and for graphene based semiconductors  $E_p = 11.4$  eV [15].

To model the tunneling correctly, the momentum  $\mathbf{K}_\perp$  in the direction perpendicular to the tunneling path and the corresponding increase in effective band gap need to be considered. This increase in band gap corresponds to the kinetic energy  $E_{r\perp} = \hbar^2 |\mathbf{K}_\perp|^2 / (2m_{r\perp})$  associated with the perpendicular momentum of the electron and hole participating in the tunneling process, with  $m_{r\perp}^{-1} = m_{v\perp}^{-1} + m_{c\perp}^{-1}$  and  $m_{v,c\perp}$  the perpendicular valence and conduction band mass.

Under the assumption of a uniform field  $F = E_g/l$  along the tunnel path and taking the perpendicular momentum into account, the tunneling probability is given by

$$T(E, E_{r\perp}) = \exp(-l(E, E_{r\perp})/l_0(E_{r\perp})) \quad (8.3)$$

where

$$l_0(E_{r\perp}) = \frac{E_g + E_{r\perp}}{2} / \left( \int_0^{E_g + E_{r\perp}} \kappa(U, E_{r\perp}) dU \right). \quad (8.4)$$

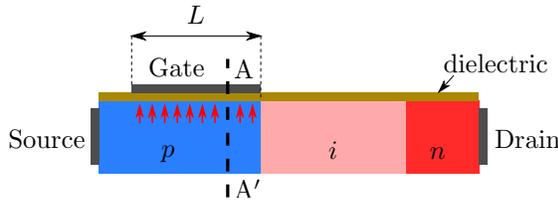


Figure 8.1: Illustration of a line-tunneling  $n$ TFET

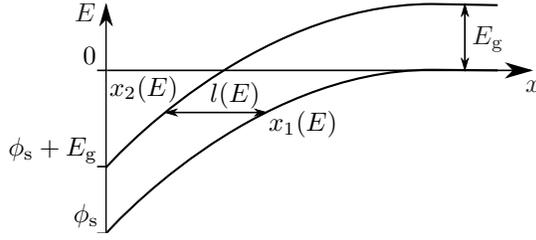


Figure 8.2: Band diagram along the A-A' cross-section.

The length  $l_0$  represents a characteristic tunneling length for the semiconductor under consideration, in particular the length corresponding to a 35% tunneling efficiency. Using the two-band model, the characteristic tunneling length is

$$l_0(E_{r\perp}) = \frac{4\hbar p}{\pi(E_g + E_{r\perp})m_0}. \quad (8.5)$$

which equals  $l_0(0) = 5.10$  nm for InSb and  $l_0(0) = 2.76$  nm for InAs.

The tunneling current  $J$  ( $\text{A}/\mu\text{m}^2$ ) in the TFET of Fig. 8.1 is obtained by summing over all states in the system [11]:

$$J \approx \frac{2qm_{r\perp}}{\hbar^3} \int_0^{|\phi_s + E_g|} \frac{dE_{r\perp}}{2\pi} \int_{\phi_s + E_g + E_{r\perp} m_{r\perp}/m_{c\perp}}^{-E_{r\perp} m_{r\perp}/m_{v\perp}} \frac{dE}{2\pi} (f_v(E) - f_c(E)) T(E, E_{r\perp}). \quad (8.6)$$

The occupation of the electrons available for tunneling is given by the Fermi-Dirac distribution

$$f_{v(c)}(E) = (1 + \exp((E - E_{F_{S(D)}})/(kT)))^{-1}, \quad (8.7)$$

with  $E_{F_{S(D)}}$  the Fermi level in the source (drain). The  $E_{F_S}$  can be determined from the doping density ( $N_S$ ):

$$2 \int \frac{dE}{2\pi} f(E) \left( \frac{2m_{v,\text{DOS}}}{\hbar^2} \right)^{3/2} \frac{\sqrt{E}}{2\pi} = N_S, \quad (8.8)$$

with  $m_{v,\text{DOS}}$  the density-of-states mass in the valence band, and the  $E_{F_D}$  from  $E_{F_S} - E_{F_D} = qV_{ds}$  [9].

In the case of line tunneling with a uniformly doped source, the potential profile is parabolic in the depletion region. The tunneling distance illustrated in Fig. 8.2 is given by

$$l(E, E_{r\perp}) = \sqrt{\frac{2\epsilon_s}{q^2 N_S}} \left( \sqrt{E_g + \frac{E_{r\perp} m_{r\perp}}{m_{v\perp}}} - E - \sqrt{-E - \frac{E_{r\perp} m_{r\perp}}{m_{c\perp}}} \right). \quad (8.9)$$

The relation of  $\phi_s$  to the applied gate bias  $V_{gs}$  is given by [9]

$$\phi_s = \left( \sqrt{\frac{t_{ox}^2 q^2 N_S \epsilon_s}{2\epsilon_{ox}^2} + qV_{gs} + E_{F_S} - WF} - \sqrt{\frac{t_{ox}^2 q^2 N_S \epsilon_s}{2\epsilon_{ox}^2}} \right)^2 \quad (8.10)$$

with  $\epsilon_{ox}$  and  $\epsilon_s$  the oxide and semiconductor permittivity respectively, and with WF the gate workfunction which is "chosen" such that  $\phi_s = -E_g$  @  $V_{gs} = 0V$ . The presented model is an improved version of the models shown in Refs. [8] and [9], because of the introduction of Fermi-Dirac statistics, which allows for a study of degenerate doping in a TFET, and because of the linking of the effective band gap to the perpendicular momentum. The improved model allows for an accurate comparison of different material band gaps for TFET.

We calculate the tunneling current for the TFET of Fig. 8.1 with an effective oxide thickness  $EOT = 1$  nm, a gate length  $L = 10$  nm and  $V_{ds} = 0.4$  V based on the presented model. The conduction band is modeled as a spherical valley with mass  $m_{c\perp}$ . For the valence band, the light hole mass will determine the tunneling probability while the heavy hole mass will determine the kinetic energy associated with the momentum in the perpendicular direction ( $E_{r\perp} m_{r\perp} / m_{v\perp}$ ), for this reason we model the valence band as an elliptic valley with the light hole mass ( $m_{lh}$ ) in the tunneling direction and the heavy hole mass ( $m_{hh}$ ) in the perpendicular direction. The resulting valence band density of states mass  $m_{v,DOS} = (m_{hh}^2 m_{lh})^{1/3}$ . We show the InAs TFET input characteristics in Fig. 8.3 for two doping concentrations. Since only the line tunneling component is considered, a steep onset is observed and for the specified parameters, the InAs TFET with  $10^{19} \text{ cm}^{-3}$  doping concentration has a current of about  $100 \mu\text{A}/\mu\text{m}$  for  $V_{gs} = 0.4$  V.

The source doping strongly affects the input characteristics. In Fig. 8.4, the current is plotted for an InAs and InSb TFET at a fixed overdrive  $V_{gs} = 0.4$  V for different doping concentrations and an optimal  $N_S$  is observed at approximately  $10^{19} \text{ cm}^{-3}$  and  $4 \times 10^{18} \text{ cm}^{-3}$ , respectively. With increasing doping, it is known that the electric fields in the semiconductor and hence the tunneling probabilities increase. However, larger doping also results in an increased voltage drop over the oxide, and hence smaller changes in  $\phi_s$ , as well as a more degenerately doped semiconductor with a reduced amount of carriers available for tunneling. As a result, there is an optimal doping concentration for a given band gap. In Fig. 8.4, the curves with no degeneracy ( $m_{v,DOS} = \infty$ ) show that the impact of degeneracy on the optimal doping concentration is larger than the impact of the oxide thickness. The optimal  $N_S$  for InAs and

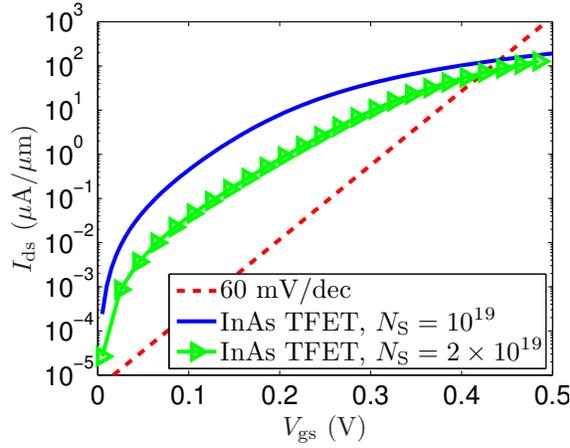


Figure 8.3: Line tunneling current as a function of gate bias for InAs with  $N_S = 10^{19} \text{ cm}^{-3}$  and  $N_S = 2 \times 10^{19} \text{ cm}^{-3}$  where  $E_g = 0.417 \text{ eV}$ ,  $m_{v\perp} = m_{hh} = 0.33m_0$ ,  $m_{v,\text{DOS}} = (m_{hh}^2 m_{lh})^{1/3} = .144m_0$ ,  $E_p = 21.5 \text{ eV}$ ,  $m_{c\perp} = 0.026m_0$  are taken from Ref. [16] and  $\epsilon_s = 12.3\epsilon_0$ .

InSb is respectively a factor of 6 and 7 larger than the respective density of states value

$$N_{v,\text{DOS}} = 2 \left( \frac{2\pi m_{v,\text{DOS}} kT}{h^2} \right)^{3/2}. \quad (8.11)$$

The material's band gap also strongly affects the input characteristics. In Fig. 8.5, we show the current at fixed overdrive for a given band gap at its optimal  $N_S$ . The isotropic conduction band mass is calculated from  $m_{c\perp} = E_g/E_p m_0$ , a formula valid in the two band  $\mathbf{k} \cdot \mathbf{p}$  model, while  $m_{v\perp}$  is taken as the one from InAs and  $m_{v,\text{DOS}}$  is given in the legend. With decreasing band gap, it is known that tunneling probability increases. However, with decreasing band gap, the density of states in both valence and conduction band decreases, the latter being responsible for the reduction of the tunneling current to 0 at  $E_g = 0 \text{ eV}$ .

In the bottom part of Fig. 8.5, we show the optimal  $N_S$  corresponding to the optimal current shown in the upper part of Fig. 8.5. An increasing and nearly linear dependence between optimal  $N_S$  and  $E_g$  is observed. The increase is due to the increasing difference with increasing  $E_g$  between the characteristic length  $l_0$  and the path length at onset  $l(E = 0)$ . As a consequence, for larger  $E_g$ , improvements in  $T(E)$  resulting from increased  $N_S$  are bigger and hence the degradation due to a smaller occupation probability is less effective, resulting in larger optimal  $N_S$ . For larger  $m_{v,\text{DOS}}$  or smaller EOT, the optimal  $N_S$  is larger (see Fig. 8.4), and due to the larger increases of  $N_S$  with increasing  $E_g$

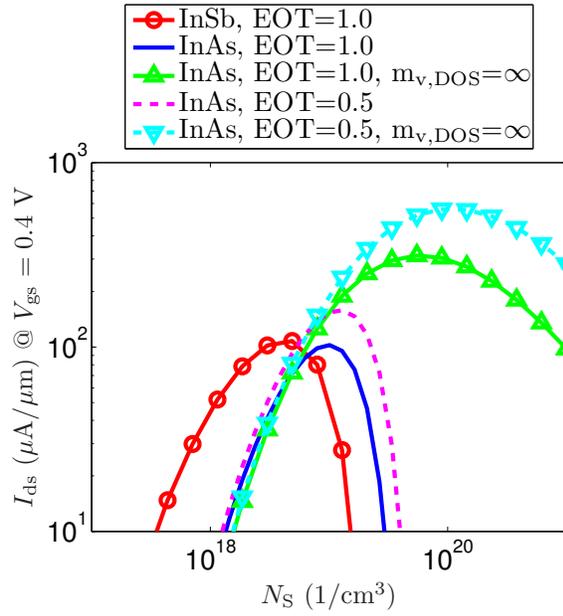


Figure 8.4: Line tunneling current at  $V_{gs} = 0.4 \text{ V}$  versus doping concentration in an InAs and InSb line TFET. For InAs, the same parameters as in Fig. 8.3 are used, for InSb  $E_g = 0.235 \text{ eV}$ ,  $m_{v\perp} = m_{hh} = 0.263m_0$ ,  $m_{v,DOS} = (m_{hh}^2 m_{lh})^{1/3} = 0.107m_0$ ,  $E_p = 23.3 \text{ eV}$ ,  $m_{c\perp} = 0.0135m_0$  and  $\epsilon_s = 16.8\epsilon_0$ . Values of EOT in the legend are given in units of nm.

(see Fig. 8.5, bottom part), the optimal  $E_g$  increases. If  $m_{v,\text{DOS}} = (m_{\text{hh}}^2 m_{\text{lh}})^{1/3}$  would have been used with  $m_{\text{lh}} = E_g/E_p m_0$ , the optimal  $E_g$  would be slightly larger than in Fig. 8.5.

Quantum mechanical simulations are performed for a Si and Ge TFET using a two-dimensional simulator which solves the Poisson-Schrödinger equation self-consistently and calculates the non-local phonon-assisted band-to-band tunneling (see Fig. 8.6) [17]. In agreement with the presented model, an optimal doping is observed and the curves have a similar shape to the ones of Fig. 8.4. The value for the Si and Ge TFET is respectively a factor 40 and 25 larger than the respective  $N_{v,\text{DOS}}$  values. This trend is in agreement with the data of Fig. 8.5. Note that the on-current is not expected to follow the same trend as in Fig. 8.5, since phonon-assisted tunneling is dependent on the electron-phonon interaction parameters as well.

A point tunneling TFET, for which the gate doesn't cover the source and whereby tunneling occurs at the source-channel interface, also has an optimal  $N_S$ . The optimum is less pronounced and shifted to higher  $N_S$  values (see Fig. 8.6) due to the inherently less steep onset and smaller current of point tunneling TFETs such that the impact of  $N_{v,\text{DOS}}$  on the input characteristics is mostly at current levels below an observable current of about 1 pA/ $\mu\text{m}$ . If smaller off-currents are considered, the optimum  $N_S$  is closer to the line tunneling optimum. For smaller band gap material TFETs, the impact of degenerate doping is expected to become more observable.

The identification of an optimal value for band gap, as well as the strong impact of  $m_{\text{DOS}}$ , is very relevant for the identifications of the optimal TFET configuration. Our study suggests that stressed hetero-interface TFETs, which might have a small effective band gap, but which will also have a small  $m_{\text{DOS}}$ , may never be able to achieve the same performance as staggered band gap heterostructure TFETs [18], consisting of 2 materials with large  $m_{\text{DOS}}$  values.

In conclusion, we present an improved model to calculate the line tunneling current in tunnel field-effect transistors. We show the existence of an optimal source doping concentration to Fermi level degeneracy and due to a voltage drop across the gate oxide at very high doping levels. The optimal doping is shown to be at least a factor of 5 larger than the density of states value, with the difference increasing with increasing band gap. We also highlight the existence of an optimal band gap value, around 0.3 eV, due to a small density of states associated with very small band gap materials. Our study therefore suggests to further explore staggered gap heterostructure TFETs, consisting of 2 large band-gap materials.

William Vandenberghe gratefully acknowledges the support of a Ph.D. stipend from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). This work was supported by IMEC's

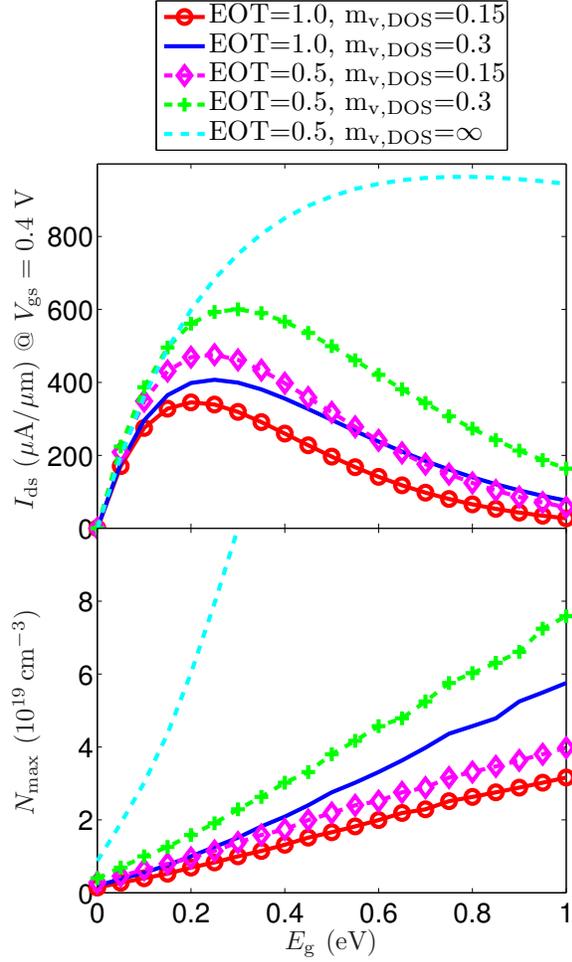


Figure 8.5: Line tunneling current at optimal  $N_S$  (top) and optimal  $N_S$  (bottom) as a function of the band gap of the source material at  $V_{gs} = 0.4$  V with  $m_{v\perp} = 0.33m_0$ ,  $m_{v,DOS}$  as in legend,  $E_p = 21.5$  eV and  $m_{c\perp} = E_g/E_p m_0$ . Values of EOT and  $m_{v,DOS}$  are given in the legend in the units of nm and  $m_0$ , respectively.

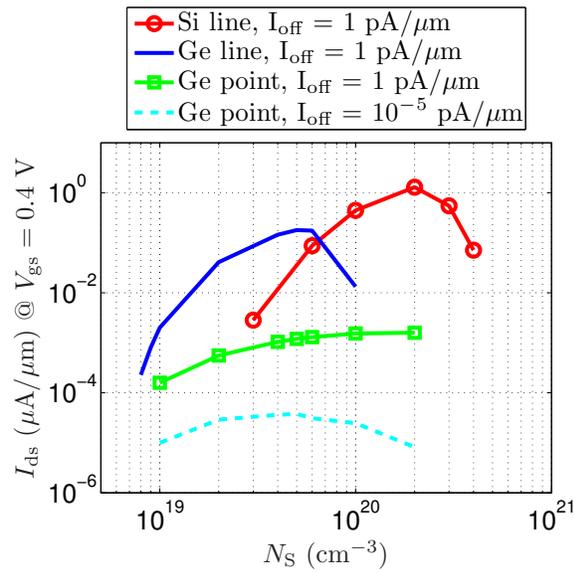


Figure 8.6: Two-dimensional quantum-mechanical calculation of tunneling current versus doping concentration at  $V_{\text{gs}} = 0.4$  V beyond  $V_{\text{onset}}$  @  $1$  pA/ $\mu\text{m}$  in a Si and Ge line TFET with  $L = 20$  nm (solid) and a Ge point TFET (dashed) with EOT = 0.66 nm. For Ge, only the indirect band gap is considered.

Industrial Affiliation Program.

## Bibliography

- [1] A. S. Verhulst, W. G. Vandenberghe, K. Maex, and G. Groeseneken, "A tunnel field-effect transistor without gate-drain overlap," *Applied Physics Letters*, vol. 91, no. 053102, July 2007.
- [2] K. Boucart and A. M. Ionescu, "Double-gate tunnel fet with high- $\kappa$  gate dielectric," *IEEE Transactions on Electron Devices*, vol. 54, no. 7, pp. 1725–1733, 2007.
- [3] S. H. Kim, H. Kam, C. Hu, and T.-J. K. Liu, "Germanium-source tunnel field effect transistors with record high ion/ioff," in *Proc. Symp. VLSI Technology*, 2009, pp. 178–179.
- [4] S. Mookerjee, D. Mohata, T. Mayer, V. Narayanan, and S. Datta, "Temperature-dependent i-v characteristics of a vertical  $\text{In}_{0.53}\text{Ga}_{0.47}$  As tunnel fet," *IEEE Electron Device Letters*, vol. 31, no. 6, pp. 564–566, 2010.
- [5] J. Wan, C. Le Royer, A. Zaslavsky, and S. Cristoloveanu, "A tunneling field effect transistor model combining interband tunneling with channel transport," *Journal of Applied Physics*, vol. 110, no. 10, p. 104503, 2011.
- [6] P.-F. Wang, T. Nirschl, D. Schmitt-Landsiedel, and W. Hansch, "Simulation of the esaki-tunneling fet," *Solid-State Electronics*, vol. 47, pp. 1187–1192, 2003.
- [7] E. O. Kane, "Zener tunneling in semiconductors," *Journal of Physics and Chemistry of Solids*, vol. 12, pp. 181–188, 1959.
- [8] W. G. Vandenberghe, A. S. Verhulst, G. Groeseneken, B. Sorée, and W. Magnus, "Analytical model for a tunnel field-effect transistor," in *IEEE Melecon*, 2008.
- [9] A. S. Verhulst, D. Leonelli, R. Rooyackers, and G. Groeseneken, "Drain voltage dependent analytical model of tunnel field-effect transistors," *Journal of Applied Physics*, vol. 110, no. 2, p. 024510, 2011. [Online]. Available: <http://link.aip.org/link/?JAP/110/024510/1>
- [10] W. G. Vandenberghe, B. Sorée, W. Magnus, G. Groeseneken, and M. V. Fischetti, "Impact of field-induced quantum confinement in tunneling field-effect devices," *Applied Physics Letters*, vol. 98, no. 14, p. 143503, 2011. [Online]. Available: <http://link.aip.org/link/?APL/98/143503/1>

- [11] C. Duke, *Tunneling in solids*, ser. Solid State Physics, Supp. 10. Academic Press, 1969.
- [12] E. O. Kane, "Theory of tunneling," *Journal of Applied Physics*, vol. 32, no. 1, pp. 83–91, 1961.
- [13] W. Vandenberghe, B. Sorée, W. Magnus, and G. Groeseneken, "Zener tunneling in semiconductors under nonuniform electric fields," *Journal of Applied Physics*, vol. 107, no. 5, p. 054520, 2010. [Online]. Available: <http://link.aip.org/link/?JAP/107/054520/1>
- [14] A. Seabaugh and Q. Zhang, "Low-voltage tunnel transistors for beyond cmos logic," *Proceedings of the IEEE*, vol. 98, no. 12, pp. 2095–2110, dec. 2010.
- [15] W. Vandenberghe, B. Sorée, W. Magnus, and G. Groeseneken, "Zener tunnelling in graphene based semiconductors the kp method," *Journal of Physics: Conference Series*, vol. 193, no. 1, p. 012111, 2009. [Online]. Available: <http://stacks.iop.org/1742-6596/193/i=1/a=012111>
- [16] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, "Band parameters for iii–v compound semiconductors and their alloys," *Journal of Applied Physics*, vol. 89, no. 11, pp. 5815–5875, 2001. [Online]. Available: <http://link.aip.org/link/?JAP/89/5815/1>
- [17] W. G. Vandenberghe, B. Sorée, W. Magnus, M. V. Fischetti, A. S. Verhulst, and G. Groeseneken, "Two-dimensional quantum mechanical modeling of band-to-band tunneling in indirect semiconductors," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, 2011.
- [18] J. Knoch and J. Appenzeller, "Modeling of high-performance p-type iii–v heterojunction tunnel fets," *IEEE Electron Device Letters*, vol. 31, no. 4, pp. 305–307, 2010.



## Part IV

# Quantum mechanical modeling of BTBT in indirect semiconductors



## Chapter 9

# BTBT in indirect semiconductors

Generalized phonon-assisted Zener tunneling in  
indirect semiconductors with non-uniform  
electric fields : a rigorous approach

William G. Vandenberghe, Bart Sorée, Wim Magnus and  
Massimo V. Fischetti

Published in *Journal of Applied Physics*, 109 (12), art.nr. 124503,  
June 2011

The following article appeared in *Journal of Applied Physics*, 109 (12), art. nr.  
124503, June 2011 and may be found at <http://link.aip.org/link/doi/10.1063/1.3595672>.

©2011 American Institute of Physics. This article may be downloaded for  
personal use only. Any other use requires prior permission of the author and  
the American Institute of Physics.

### **Contributions of first author**

- Conception and development of the theory
- Development of the code
- Creating figures
- Text redaction and editing

## Abstract

A general framework to calculate the Zener current in an indirect semiconductor with an externally applied potential is provided. Assuming a parabolic valence and conduction band dispersion, the semiconductor is in equilibrium in the presence of the external field as long as the electron-phonon interaction is absent. The linear response to the electron-phonon interaction results in a non-equilibrium system. The Zener tunneling current is calculated from the number of electrons making the transition from valence to conduction band per unit time. A convenient expression based on the single particle spectral functions is provided, enabling the evaluation of the Zener tunneling current under any three-dimensional potential profile. For a one dimensional potential profile an analytical expression is obtained for the current in a bulk semiconductor, a semiconductor under uniform field and a semiconductor under a non-uniform field using the WKB approximation. The obtained results agree with the Kane result in the low field limit. A numerical example for abrupt  $p-n$  diodes with different doping concentrations is given, from which it can be seen that the uniform field model is a better approximation than the WKB model but a direct numerical treatment is required for low bias conditions.

## 9.1 Introduction

The continued device scaling in the semiconductor industry has enabled the fabrication of devices with nanometer sizes. The scaling of the dimensions was however not accompanied by a similar reduction in operating voltages and, as a consequence, large electric fields are present in these nanometer-sized semiconductor device structures. Accordingly, tunneling from valence to conduction band, also referred to as Zener tunneling, produces a significant current, while further reduction of the device dimensions will enlarge the role of Zener tunneling even more. In the end, Zener tunneling could be made so efficient that it can provide significant drive current for a novel type of transistor [1].

Zener tunneling was first described by Zener [2] and the first calculations of the Zener tunneling rate in a uniform electric field were made by Keldysh [3] and Kane [4] for a direct semiconductor based on a two band model. In the presence of a non-uniform field however, the WKB approximation can be used provided the applied bias voltage is generating slowly varying potential profiles inside the active device area [5]. The problem of Zener tunneling in an indirect semiconductor was treated earlier by Keldysh [6] and Kane [7], whereas Schenk

[8] formulated a model for the tunneling probability in a uniform field using the Kubo formalism [9, 10]. Another noteworthy approach is the calculation by Rivas [11] based on a Green's function formulation of Fermi's golden rule to calculate the current through a silicon tunnel diode.

Since most of the available models are one-dimensional, the currently most popular approach to calculate Zener current in a device is to choose a set of one-dimensional tunnel paths and to determine the tunneling probability along these paths within the WKB approximation [12, 13]. However, as all one-dimensional models ignore the pronounced two-dimensional shape of a realistic potential profile, the development of a more comprehensive calculation of Zener tunneling in semiconductors is in order. Furthermore, Zener tunneling should not only be modeled accurately for two-dimensional potential profiles but should as well incorporate phonon-assisted tunneling phenomena.

Various general methods for dealing with electronic transport in semiconductors taking quantum effects into account have been developed such as the Wigner transport equation [14–16], the non-equilibrium Green's function formalism [17], the Pauli master equation [18] and the quantum balance equations [19]. But all methods come with their own limitations and no comprehensive generally applicable scheme for electronic transport is available due to the absence of a general non-equilibrium thermodynamic framework. Furthermore all methods can become computationally very expensive when multiple bands have to be included as required by Zener tunneling.

Being mediated by phonon scattering, the process of Zener tunneling in an indirect semiconductor strongly contrasts with conventional single-band transport where phonon scattering acts as a mere dissipative mechanism limiting the drive current. This conceptual difference in the role of phonon scattering jeopardizes the useability of the above mentioned methods to model Zener tunneling [12] even further.

In this paper we provide a framework to calculate the current due to phonon-assisted tunneling in indirect semiconductors. In section 9.2, we discuss the valence and conduction band electrons in the indirect semiconductor under the application of an external potential. In section 9.3, we establish the Hamiltonian of the phonons and the electron-phonon interaction. In section 9.4 we construct a non-equilibrium system starting from an equilibrium system consisting of three non-interacting ensembles: valence band electrons, conduction band electrons and free phonons, all being characterized by a grand canonical ensemble density matrix of the Gibbs form. We use perturbation theory to calculate the non-equilibrium density matrix up to first order in the electron – phonon interaction, the latter mediating the band-to-band tunneling events. Then, we extract an expression for the steady state phonon-assisted tunneling current from the non-equilibrium density matrix as well as the net rate of electrons making the transition from valence to conduction band. In particu-

lar, for the sake of computational efficiency, we rewrite the tunneling current in terms of properly chosen spectral density functions. Using the proposed method, an analytical expression is derived for tunneling in the presence of a one-dimensional potential profile in three cases: 1) no external electric field, 2) a uniform electric field and 3) a non-uniform 1D field within the WKB approximation. Finally, we discuss both the analytical results and the numerical calculations.

In a recent paper [20] we apply the presented formalism to the case of a tunnel field-effect transistor and show that the use of semi-classical model introduces very large errors due to quantum confinement near the interface.

## 9.2 Envelope function approximation for an indirect semiconductor

In this section we consider the valence and conduction band electrons in an indirect semiconductor subjected to an external bias voltage. We introduce the single-electron energies and eigenfunctions required to construct the non-interacting electron Hamiltonian and some related second-quantized operators. In section 9.4, the obtained operators are used to construct the non-equilibrium density matrix by introducing the interaction with the phonons.

In a semiconductor an electron is subjected to a periodic potential energy  $U_{\text{lat}}(\mathbf{r})$  caused by the ions and electrons in the lattice. In a semiconductor device, doping and bias voltages externally applied at the contacts induce an additional non-periodic potential energy  $U_{\text{ext}}(\mathbf{r})$ . The single-electron Schrödinger equation reads

$$\left( -\frac{\hbar^2 \nabla^2}{2m} + U_{\text{lat}}(\mathbf{r}) + U_{\text{ext}}(\mathbf{r}) \right) \psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (9.1)$$

In the absence of the non-periodic potential, the periodicity of the potential results in Bloch wavefunctions

$$\left( -\frac{\hbar^2 \nabla^2}{2m} + U_{\text{lat}}(\mathbf{r}) \right) \exp(i\mathbf{k} \cdot \mathbf{r})u_{n\mathbf{k}}(\mathbf{r}) = E_{n\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{r})u_{n\mathbf{k}}(\mathbf{r}). \quad (9.2)$$

where  $n$  is a band index,  $\mathbf{k}$  is a wavevector and  $E_{n\mathbf{k}}$  describes the energy dispersion relation. For semiconductor device applications,  $n$  is essentially running over  $c$  and  $v$ , respectively denoting the conduction and valence band separated by the bandgap, while the states that participate in the conduction process are located near the Fermi level. In a degenerate semiconductor, as often employed when studying Zener tunneling, the Fermi level will be located inside the conduction or valence band.

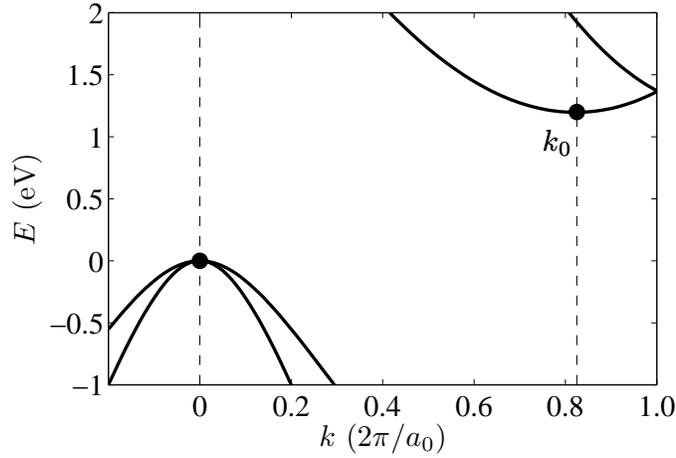


Figure 9.1: The Si bandstructure along the [100] direction [21] around the valence band maximum and conduction band minimum indicating the minimum of the conduction band at  $\mathbf{k} = \mathbf{k}_0$

In the indirect semiconductor under consideration, we assume a single valence band maximum occurring at  $\mathbf{k} = 0$  and one of the conduction band minima being located at  $\mathbf{k} = \mathbf{k}_0$ , as shown in Fig. 9.1. The energy dispersion can be approximated around the conduction band minimum and the valence band maximum using the effective mass approximation:

$$E_{v\mathbf{k}} \approx E_{v0} - \frac{\hbar^2}{2m_v^*} |\mathbf{k}|^2 \quad \text{for } \mathbf{k} \approx 0, \quad (9.3)$$

$$E_{c\mathbf{k}} \approx E_{c\mathbf{k}_0} + \frac{\hbar^2}{2m_c^*} |\mathbf{k} - \mathbf{k}_0|^2 \quad \text{for } \mathbf{k} \approx \mathbf{k}_0. \quad (9.4)$$

When an external potential is added, the wavefunctions are no longer Bloch functions and no energy dispersion can be defined. However, the Bloch states solving Eq. (9.2) still provide a complete orthogonal set of basis states in which each one-electron eigenstate  $\psi(\mathbf{r})$  can be expanded as

$$\psi(\mathbf{r}) = \sum_{n,\mathbf{k}} a_{n\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{r}) u_{n\mathbf{k}}(\mathbf{r}) \quad (9.5)$$

On the other hand, we may as well expand  $\psi(\mathbf{r})$  around a high-symmetry point within the first Brillouin zone, such as the  $\Gamma$ -point ( $\mathbf{k} = 0$ ), writing

$$\psi(\mathbf{r}) = \sum_n \chi_{n0}(\mathbf{r}) u_{n0}(\mathbf{r}) \quad (9.6)$$

where Eq. (9.6) defines the envelope functions  $\chi_{n0}(\mathbf{r})$ .

For the conduction band electrons, the envelope functions defined in Eq. (9.6) are rapidly oscillating since the conduction band minimum has a large wavevector. In order to end up with smoothly varying envelope functions, we may alternatively define the set  $\chi_{n\mathbf{k}_0}$  by

$$\psi(\mathbf{r}) \equiv \sum_n \chi_{n\mathbf{k}_0}(\mathbf{r}) \exp(i\mathbf{k}_0 \cdot \mathbf{r}) u_{n\mathbf{k}_0}(\mathbf{r}). \quad (9.7)$$

For a given eigen energy  $E$ , the dominant contribution in Eq. (9.7) comes from the component with its eigen energy  $E_{n\mathbf{k}_0}$  closest to  $E$ . By restricting the summation of the index to the conduction band, the wavefunction of a conduction band electron can be written as

$$\psi_c(\mathbf{r}) = \chi_c(\mathbf{r}) \exp(i\mathbf{k}_0 \cdot \mathbf{r}) u_{c\mathbf{k}_0}(\mathbf{r}) \quad (9.8)$$

with a slowly varying envelope function, *i.e.* the Fourier components  $a_{n\mathbf{k}}$  are only significant for  $\mathbf{k} \approx \mathbf{k}_0$ . Since we consider the conduction band to be a non-degenerate valley at  $\mathbf{k}_0$ , we conclude that  $u_{c\mathbf{k}} \approx u_{c\mathbf{k}_0}$  for  $\mathbf{k} \approx \mathbf{k}_0$  and

$$\chi_c(\mathbf{r}) \approx \sum_{\mathbf{k}} a_{c\mathbf{k}} \exp(i(\mathbf{k} - \mathbf{k}_0) \cdot \mathbf{r}) \quad (9.9)$$

Moreover, adopting the effective mass approximation, we may approximate the kinetic plus the periodic term of Eq. (9.1) using Eq. (9.2) to get:

$$\left( -\frac{\hbar^2 \nabla^2}{2m} + U_{\text{lat}}(\mathbf{r}) \right) \chi_c(\mathbf{r}) \exp(i\mathbf{k}_0 \cdot \mathbf{r}) u_{c\mathbf{k}_0}(\mathbf{r}) \quad (9.10)$$

$$= \sum_{\mathbf{k}} a_{c\mathbf{k}} E_{c\mathbf{k}} \exp(i(\mathbf{k} - \mathbf{k}_0) \cdot \mathbf{r}) \exp(i\mathbf{k}_0 \cdot \mathbf{r}) u_{c\mathbf{k}_0}(\mathbf{r}) \quad (9.11)$$

$$\approx \left( \left( E_{c\mathbf{k}_0} - \frac{\hbar^2 \nabla^2}{2m_c^*} \right) \chi_c(\mathbf{r}) \right) \exp(i\mathbf{k}_0 \cdot \mathbf{r}) u_{c\mathbf{k}_0}(\mathbf{r}) \quad (9.12)$$

The envelope functions are determined by combining Eq. (9.1), Eq. (9.8) and Eq. (9.12), which results in the effective Schrödinger equation

$$\left( E_{c0} - \frac{\hbar^2}{2m_c^*} \nabla^2 + U_{\text{ext}}(\mathbf{r}) \right) \chi_c(\mathbf{r}) = E \chi_c(\mathbf{r}). \quad (9.13)$$

For valence band states the interaction between the different degenerate valence bands must be taken into account. Nevertheless, we will restrict ourselves to the effective mass approximation also for the valence band which is a common approach [8]. To go to a more realistic valence band structure, 6-band  $\mathbf{k} \cdot \mathbf{p}$  theory could be employed without compromising the following approach to calculate the current.

A major improvement of going beyond the effective mass approximation will be a more accurate description of the complex band structure in the bandgap

which affects the decay of the wavefunctions penetrating the bandgap. Inspection of the complex band structure for the case of silicon reveals that the effective mass approximation overestimates the decay. Hence, the real tunneling probabilities will be higher than those predicted by the effective mass approximation. A second consequence of taking non-parabolicity into account is a different density of states, which will greatly influence the valence band because of its anisotropy and the presence of the split-off band.

In the effective mass approximation the wave functions for valence band electrons are

$$\psi_v(\mathbf{r}) = \chi_v(\mathbf{r})u_{v0}(\mathbf{r}) \quad (9.14)$$

where the envelope functions are determined from

$$\left(E_{v0} + \frac{\hbar^2}{2m_v^*} \nabla^2 + U_{\text{ext}}(\mathbf{r})\right) \chi_v(\mathbf{r}) = E \chi_v(\mathbf{r}). \quad (9.15)$$

We can now construct the electron basis set consisting of the valence and conduction band eigenstates

$$\phi_{v\ell}(\mathbf{r}) = \chi_{v\ell}(\mathbf{r})u_{v0}(\mathbf{r}) \quad (9.16)$$

$$\phi_{c\ell}(\mathbf{r}) = \chi_{c\ell}(\mathbf{r})u_{c\mathbf{k}_0}(\mathbf{r}) \exp(i\mathbf{k}_0 \cdot \mathbf{r}), \quad (9.17)$$

where  $\chi_{v\ell}(\mathbf{r})$  and  $\chi_{c\ell}(\mathbf{r})$  are the envelope eigenfunctions of Eq. (9.13) and Eq. (9.15) with eigenvalues  $E_{v\ell}$  and  $E_{c\ell}$  respectively.  $\ell$  is a subband index running through the set of relevant quantum numbers.

The one-electron states allow us to write the electron Hamiltonian in second quantization

$$\hat{H}_{\text{el}} = \sum_{\ell} E_{v\ell} \hat{b}_{\ell}^{\dagger} \hat{b}_{\ell} + E_{c\ell} \hat{c}_{\ell}^{\dagger} \hat{c}_{\ell} \quad (9.18)$$

where  $\hat{b}_{\ell}$  and  $\hat{c}_{\ell}$  annihilate an electron in the states  $\phi_{v\ell}(\mathbf{r})$  and  $\phi_{c\ell}(\mathbf{r})$  respectively. Furthermore, it proves convenient to define electron field operators

$$\hat{\psi}_v(\mathbf{r}) = \sum_{\ell} \phi_{v\ell}(\mathbf{r}) \hat{b}_{\ell}, \quad (9.19)$$

$$\hat{\psi}_c(\mathbf{r}) = \sum_{\ell} \phi_{c\ell}(\mathbf{r}) \hat{c}_{\ell}. \quad (9.20)$$

to represent the electron density and other position dependent quantities. As explained in the next section, the tunneling current is related to the expectation value of the valence and conduction band electron number operators

$$\hat{N}_v = \sum_{\ell} \hat{b}_{\ell}^{\dagger} \hat{b}_{\ell} = \int d^3r \hat{\psi}_v^{\dagger}(\mathbf{r}) \hat{\psi}_v(\mathbf{r}), \quad (9.21)$$

$$\hat{N}_c = \sum_{\ell} \hat{c}_{\ell}^{\dagger} \hat{c}_{\ell} = \int d^3r \hat{\psi}_c^{\dagger}(\mathbf{r}) \hat{\psi}_c(\mathbf{r}). \quad (9.22)$$

## 9.3 Phonons

In general, one may want to consider various agents other than phonons that are proven to assist Zener tunneling, such as traps and defects, but in this work we have investigated the role of phonons only.

The free phonons are described by the Hamiltonian

$$\hat{H}_{\text{PH}} = \sum_{\mathbf{q}} \hbar\omega_{\mathbf{q}} \hat{a}_{\mathbf{q}}^{\dagger} \hat{a}_{\mathbf{q}} \quad (9.23)$$

where  $a_{\mathbf{q}}$  annihilates a phonon of mode  $\mathbf{q}$  with energy  $\hbar\omega_{\mathbf{q}}$ .

As this work focuses on indirect semiconductors, we specifically consider the interaction between electrons and short-wavelength phonons bridging the gap between the top of the valence band and the bottom of the conduction band located at  $\mathbf{k}_0$ . The corresponding interaction hamiltonian reads [22]

$$\hat{H}' = \sum_{\ell\ell'\mathbf{q}} \left( g_{v\ell c\ell'\mathbf{q}} \hat{b}_{\ell}^{\dagger} \hat{c}_{\ell'} \left( \hat{a}_{\mathbf{q}} + \hat{a}_{-\mathbf{q}}^{\dagger} \right) + \text{h.c.} \right) . \quad (9.24)$$

The coupling strengths  $g_{v\ell c\ell'\mathbf{q}}$  incorporate all possible interband transitions and typically reflect the features of the envelope wave functions. Alternative interaction mechanisms capable of assisting interband transitions such as localized trapped electrons are not considered in this work.

## 9.4 Non-equilibrium density matrix

In this section we construct the non-equilibrium density matrix by tracing the time evolution of the density matrix up to the lowest order in the electron-phonon interaction. Mediating electron transitions between the valence and conduction band, the non-equilibrium density matrix carries a non-zero tunneling current.

### 9.4.1 Equilibrium

In the absence of interactions we may consider the valence band and conduction band electrons as well as the phonons three separate entities that are individually characterized by their equilibrium density matrices. The density matrix governing the entire, uncoupled system therefore reduces to the direct product of the three density matrices, which can be written in its grand-canonical form

$$\hat{\rho}_0 = \hat{\rho}_v \otimes \hat{\rho}_c \otimes \hat{\rho}_{\text{PH}} \quad (9.25)$$

$$= \frac{\exp\left(-\beta(\hat{H}_{\text{el}} - \mu_{\text{v}}\hat{N}_{\text{v}} - \mu_{\text{c}}\hat{N}_{\text{c}} + \hat{H}_{\text{PH}})\right)}{\mathcal{Z}} \quad (9.26)$$

where  $\beta = 1/(kT)$  and  $\mathcal{Z}$  denotes the partition function ensures that  $\text{Tr}(\hat{\rho}_0) = 1$ . The chemical potentials of the conduction and valence band electrons  $\mu_{\text{v}}, \mu_{\text{c}}$  will be determined by imposing charge neutrality in the device contacts whereas the phonon system has zero chemical potential.

The net charge density of the equilibrium system can be calculated from

$$\rho_{\text{net}}(\mathbf{r}) = -e\text{Tr}\left((\hat{\psi}_{\text{c}}^{\dagger}(\mathbf{r})\hat{\psi}_{\text{c}}(\mathbf{r}) - \hat{\psi}_{\text{v}}(\mathbf{r})\hat{\psi}_{\text{v}}^{\dagger}(\mathbf{r}))\hat{\rho}_0\right). \quad (9.27)$$

In order to facilitate both numerical or analytical evaluations, we define the valence and conduction band spectral functions based on the envelope functions:

$$A_{\text{v,c}}(\mathbf{r}, \mathbf{r}'; E) = 2\pi\delta(E - H_{\text{v,c}}) \quad (9.28)$$

$$= 2\pi \sum_{\ell} \chi_{\text{v,c}\ell}(\mathbf{r})\delta(E - E_{\text{v,c}\ell})\chi_{\text{v,c}\ell}^*(\mathbf{r}') \quad (9.29)$$

The net charge density averaged over a unit cell can now be evaluated as

$$\rho_{\text{net}}(\mathbf{r}) = -e \int \frac{dE}{2\pi} \left( f_{\text{c}}(E)A_{\text{c}}(\mathbf{r}, \mathbf{r}; E) + (1 - f_{\text{v}}(E))A_{\text{v}}(\mathbf{r}, \mathbf{r}; E) \right). \quad (9.30)$$

where the valence and the conduction band distribution function is a Fermi-Dirac distribution

$$f_{\text{v,c}}(E) = \frac{1}{\exp(\beta(E - \mu_{\text{v,c}})) + 1}. \quad (9.31)$$

### 9.4.2 The quantum Liouville equation

Taking  $\hat{H}_0$  to be the Hamiltonian of the non-interacting electrons and phonons, *i.e.*

$$\hat{H}_0 = \hat{H}_{\text{el}} + \hat{H}_{\text{PH}}, \quad (9.32)$$

we express the time dependent density matrix in the interaction picture as

$$\tilde{\rho}(t) = \exp(i\hat{H}_0 t/\hbar)\hat{\rho}(t)\exp(-i\hat{H}_0 t/\hbar) \quad (9.33)$$

to calculate its time evolution from

$$i\hbar \frac{d}{dt} \tilde{\rho}(t) = [\tilde{H}'(t), \tilde{\rho}(t)] \quad (9.34)$$

where  $\tilde{H}'(t)$  denotes the electron – phonon interaction, expressed in the interaction picture,

$$\tilde{H}'(t) = \exp(i\hat{H}_0 t/\hbar)\hat{H}'\exp(-i\hat{H}_0 t/\hbar). \quad (9.35)$$

Integrating Eq. (9.34) and keeping only terms of first order in  $\tilde{H}'(t)$ , we obtain

$$\tilde{\varrho}^{(1)}(t) = \hat{\varrho}_0 - \frac{i}{\hbar} \int_0^t d\tau [\tilde{H}'(\tau), \hat{\varrho}_0] \quad (9.36)$$

since  $\hat{\varrho}_0$  commutes with  $\hat{H}_0$  and is unchanged in the interaction picture.

Since  $\tilde{H}'$  represents a phonon creation or annihilation event and  $\hat{\varrho}_0$  is diagonal in the phonon states, the diagonal elements of the density matrix are not affected to first order. Hence, the electron density remains unaltered and can still be calculated from Eq. (9.30).

## 9.5 The steady state current

The quantity we want to calculate is the current  $I$  carried by the interacting electron – phonon system at steady state, *i.e.* when  $t \rightarrow \infty$ . As a definition, the tunneling current basically represents the rate of change in the number of electrons in the conduction band per unit time:

$$I_v = -e \lim_{t \rightarrow \infty} \frac{d}{dt} \text{Tr} \left( \tilde{N}_v \tilde{\varrho}(t) \right), \quad (9.37)$$

$$I_c = -e \lim_{t \rightarrow \infty} \frac{d}{dt} \text{Tr} \left( \tilde{N}_c \tilde{\varrho}(t) \right), \quad (9.38)$$

where  $I_v + I_c = 0$  holds due to particle conservation.

Since the valence band number operator  $\hat{N}_v$  commutes with  $\hat{H}_0$ ,  $\hat{N}_v$  does not gain any time dependence in the interaction picture. Making use of the Liouville equation and Eq. (9.37), we obtain

$$I_v = -\frac{e}{\hbar^2} \lim_{t \rightarrow \infty} \int_0^t d\tau \text{Tr} \left( \hat{N}_v [\tilde{H}'(t), [\tilde{H}'(\tau), \hat{\varrho}_0]] \right). \quad (9.39)$$

Exploiting the identity

$$\text{Tr}(A[B, [C, D]]) = \text{Tr}([A, [B, C]]D) \quad (9.40)$$

which follows from the invariance under cyclic permutations, we rewrite Eq. (9.39),

$$I_v = \frac{e}{\hbar^2} \lim_{t \rightarrow \infty} \int_0^t d\tau \text{Tr} \left( [[\hat{N}_v, \tilde{H}'(t)], \tilde{H}'(\tau)] \hat{\varrho}_0 \right). \quad (9.41)$$

### 9.5.1 The phonon-assisted tunneling current

For the phonons, the interaction Hamiltonian in the interaction picture is given by

$$\begin{aligned} \tilde{H}'(t) = & \sum_{\ell\ell'\mathbf{q}} g_{v\ell c\ell'\mathbf{q}} \hat{b}_{\ell}^{\dagger} \hat{c}_{\ell'} \left( \hat{a}_{\mathbf{q}} \exp(i(E_{v\ell} - E_{c\ell'} - \hbar\omega_{\mathbf{q}})t/\hbar) \right. \\ & \left. + \hat{a}_{-\mathbf{q}}^{\dagger} \exp(i(E_{v\ell} - E_{c\ell'} + \hbar\omega_{\mathbf{q}})t/\hbar) \right) + \text{h.c.} \end{aligned} \quad (9.42)$$

under the assumption  $\omega_{-\mathbf{q}} = \omega_{\mathbf{q}}$ .

The commutator  $[\hat{N}_v, \tilde{H}'(t)]$  can be evaluated straightforwardly to yield

$$\begin{aligned} [\hat{N}_v, \tilde{H}'(t)] = & \sum_{\ell\ell'\mathbf{q}} g_{v\ell c\ell'\mathbf{q}} \hat{b}_{\ell}^{\dagger} \hat{c}_{\ell'} \left( \hat{a}_{\mathbf{q}} \exp(i(E_{v\ell} - E_{c\ell'} - \hbar\omega_{\mathbf{q}})t/\hbar) \right. \\ & \left. + \hat{a}_{-\mathbf{q}}^{\dagger} \exp(i(E_{v\ell} - E_{c\ell'} + \hbar\omega_{\mathbf{q}})t/\hbar) \right) - \text{h.c.} \end{aligned} \quad (9.43)$$

Since  $\varrho_0$  is diagonal in the phonon operators, only terms connecting a phonon annihilation operator in  $[\hat{N}_v, \tilde{H}'(t)]$  with a creation operator in  $\tilde{H}'(\tau)$  and vice versa do contribute to the average of the double commutator appearing in Eq. (9.41). The resulting expression for the current therefore simplifies to

$$\begin{aligned} I_v = & \frac{e}{\hbar^2} \lim_{t \rightarrow \infty} \int_0^t d\tau \sum_{\ell\ell'\mathbf{q}} \text{Tr} \left( \right. \quad (9.44) \\ & \left[ g_{v\ell c\ell'\mathbf{q}} \hat{b}_{\ell}^{\dagger} \hat{c}_{\ell'} \hat{a}_{\mathbf{q}}, \left( g_{v\ell c\ell'\mathbf{q}} \hat{b}_{\ell}^{\dagger} \hat{c}_{\ell'} \hat{a}_{\mathbf{q}} \right)^{\dagger} \right] 2 \cos((E_{v\ell} - E_{c\ell'} - \hbar\omega_{\mathbf{q}})(t - \tau)/\hbar) \hat{\varrho}_0 \\ & \left. + \left[ g_{v\ell c\ell'\mathbf{q}} \hat{b}_{\ell}^{\dagger} \hat{c}_{\ell'} \hat{a}_{-\mathbf{q}}^{\dagger}, \left( g_{v\ell c\ell'\mathbf{q}} \hat{b}_{\ell}^{\dagger} \hat{c}_{\ell'} \hat{a}_{-\mathbf{q}}^{\dagger} \right)^{\dagger} \right] 2 \cos((E_{v\ell} - E_{c\ell'} + \hbar\omega_{\mathbf{q}})(t - \tau)/\hbar) \hat{\varrho}_0 \right). \end{aligned} \quad (9.45)$$

Bearing in mind that the initial equilibrium state describes uncoupled electrons and phonons, we may employ Eq. (9.25) to simplify the evaluation of the traces. For example,

$$\begin{aligned} \text{Tr} \left( \hat{b}_{\ell}^{\dagger} \hat{c}_{\ell'} \hat{a}_{\mathbf{q}} \hat{a}_{\mathbf{q}'}^{\dagger} \hat{c}_{r'}^{\dagger} \hat{b}_r \hat{\varrho}_0 \right) &= \text{Tr} \left( \hat{b}_{\ell}^{\dagger} \hat{b}_r \hat{\varrho}_v \right) \text{Tr} \left( \hat{c}_{\ell'} \hat{c}_{r'}^{\dagger} \hat{\varrho}_c \right) \text{Tr} \left( \hat{a}_{\mathbf{q}} \hat{a}_{\mathbf{q}'}^{\dagger} \hat{\varrho}_{\text{PH}} \right) \\ &= \delta_{\ell r} \delta_{\ell' r'} \delta_{\mathbf{q} \mathbf{q}'} f_v(E_{v\ell}) \left( 1 - f_c(E_{c\ell'}) \right) \left( 1 + \nu(\hbar\omega_{\mathbf{q}}) \right), \end{aligned} \quad (9.46)$$

where  $\nu(E)$  denotes the Bose-Einstein distribution

$$\nu(E) = \frac{1}{\exp(\beta E) - 1}. \quad (9.47)$$

Taking the limit  $t \rightarrow \infty$  and using the identity

$$\lim_{t \rightarrow \infty} \int_0^t d\tau \cos \omega \tau = \pi \delta(\omega), \quad (9.48)$$

we derive the steady-state current

$$\begin{aligned} I_v = & -\frac{2\pi e}{\hbar} \sum_{\ell\ell' \mathbf{q}} |g_{v\ell c\ell' \mathbf{q}}|^2 \\ & \times \left( f_v(E_{v\ell})(1 - f_c(E_{c\ell'}))(\nu(\hbar\omega_{\mathbf{q}}) + 1)\delta(E_{v\ell} - E_{c\ell'} - \hbar\omega_{\mathbf{q}}) \right. \\ & - f_c(E_{c\ell'})(1 - f_v(E_{v\ell}))\nu(\hbar\omega_{\mathbf{q}})\delta(E_{v\ell} - E_{c\ell'} - \hbar\omega_{\mathbf{q}}) \\ & + f_v(E_{v\ell})(1 - f_c(E_{c\ell'}))\nu(\hbar\omega_{\mathbf{q}})\delta(E_{v\ell} - E_{c\ell'} + \hbar\omega_{\mathbf{q}}) \\ & \left. - f_c(E_{c\ell'})(1 - f_v(E_{v\ell}))(\nu(\hbar\omega_{\mathbf{q}}) + 1)\delta(E_{v\ell} - E_{c\ell'} + \hbar\omega_{\mathbf{q}}) \right). \quad (9.49) \end{aligned}$$

The four terms contributing to the current in Eq. (9.49) can be interpreted in a similar way as in Ref. [11]. For instance, the first term refers to an electron being excited from the valence to the conduction band while emitting a phonon, etc.

### 9.5.2 Evaluation of the current using spectral functions

Alternatively, the tunneling current may be written as an energy integral

$$\begin{aligned} I = & -\frac{e}{\hbar} \int \frac{dE}{2\pi} \sum_{\mathbf{q}} \left( (f_v(E)(1 - f_c(E - \hbar\omega_{\mathbf{q}}))(\nu(\hbar\omega_{\mathbf{q}}) + 1) \right. \\ & - f_c(E - \hbar\omega_{\mathbf{q}})(1 - f_v(E))\nu(\hbar\omega_{\mathbf{q}})T_v^{\text{em}}(E, \mathbf{q}) \\ & + (f_v(E)(1 - f_c(E + \hbar\omega_{\mathbf{q}}))\nu(\hbar\omega_{\mathbf{q}}) \\ & \left. - f_c(E + \hbar\omega_{\mathbf{q}})(1 - f_v(E))(\nu(\hbar\omega_{\mathbf{q}}) + 1))T_v^{\text{abs}}(E, \mathbf{q}) \right) \quad (9.50) \end{aligned}$$

where the property of the delta function

$$\delta(E_{v\ell} - E_{c\ell'} - \hbar\omega_{\mathbf{q}}) = \int_{-\infty}^{+\infty} dE \delta(E - E_{c\ell'} - \hbar\omega_{\mathbf{q}}) \delta(E - E_{v\ell})$$

has been used and the probability of exciting an electron from the valence band to the conduction band with emission or absorption of a phonon has been introduced:

$$T_v^{\text{abs,em}}(E, \mathbf{q}) = (2\pi)^2 \sum_{\ell\ell'} |g_{v\ell c\ell' \mathbf{q}}|^2 \delta(E - E_{v\ell}) \delta(E - E_{c\ell'} \pm \hbar\omega_{\mathbf{q}}). \quad (9.51)$$

Equally, it would be possible to write Eq. (9.50) in terms of the reverse transition by using

$$T_c^{\text{em}}(E, \mathbf{q}) = T_v^{\text{abs}}(E - \hbar\omega_{\mathbf{q}}, -\mathbf{q}), \quad (9.52)$$

$$T_c^{\text{abs}}(E, \mathbf{q}) = T_v^{\text{em}}(E + \hbar\omega_{\mathbf{q}}, -\mathbf{q}). \quad (9.53)$$

Further numerical processing of the current formula requires an explicit representation of the scattering matrix elements. The latter are related to the bulk coupling strength  $M_{\mathbf{q}}$  through

$$g_{v\ell c\ell'\mathbf{q}} = M_{\mathbf{q}} \int d^3r \phi_{v\ell}^*(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}) \phi_{c\ell'}(\mathbf{r}). \quad (9.54)$$

In this work, we have borrowed  $M_{\mathbf{q}}$  from the deformation potential interaction [22]

$$M_{\mathbf{q}} = D|\mathbf{q}| \sqrt{\frac{\hbar}{2\rho_s\omega_{\mathbf{q}}\Omega}} \quad (9.55)$$

where  $\rho_s$ ,  $D$  and  $\Omega$  respectively represent the semiconductor mass density, the deformation potential and the total volume. In order to express the matrix elements in terms of the envelope functions, we assume that the basis functions are normalized on the lattice unit cell  $\Omega_c$  and define

$$M'_{\mathbf{q}} = M_{\mathbf{q}} \int_{\Omega_c} d^3r u_{v0}^*(\mathbf{r}) u_{c\mathbf{k}_0}(\mathbf{r}). \quad (9.56)$$

Since the envelope functions are slowly varying in space, they can be taken constant along each unit cell and the matrix element from Eq. (9.54) can be approximated as

$$g_{v\ell c\ell'\mathbf{q}} = M'_{\mathbf{q}} \int d^3r \chi_{v\ell}^*(\mathbf{r}) \exp(i(\mathbf{q} + \mathbf{k}_0) \cdot \mathbf{r}) \chi_{c\ell'}(\mathbf{r}). \quad (9.57)$$

Again, within the perspective of numerical evaluation, we may use the single-particle spectral functions  $A(\mathbf{r}, \mathbf{r}'; E)$  defined in Eq. (9.29) to rewrite the tunneling probability

$$T_v^{\text{abs,em}}(E, \mathbf{q}) = |M'_{\mathbf{q}}|^2 \int d^3r \int d^3r' \left( \exp(i(\mathbf{q} + \mathbf{k}_0) \cdot (\mathbf{r}' - \mathbf{r})) \right. \\ \left. \times A_v(\mathbf{r}, \mathbf{r}'; E) A_c(\mathbf{r}', \mathbf{r}; E \pm \hbar\omega_{\mathbf{q}}) \right). \quad (9.58)$$

As the matrix elements  $g_{v\ell c\ell'\mathbf{q}}$  turn out to be strongly peaked for  $\mathbf{q} = -\mathbf{k}_0$ , we can assume negligible dispersion for the phonons assisting the tunneling processes, *i.e.*  $\omega_{\mathbf{q}} \approx \omega_{-\mathbf{k}_0} = \omega_{k_0}$ . Hence, the occupation probabilities in Eq. (9.50) become independent of  $\mathbf{q}$  and we can define effective transition probabilities by

$$T_v^{\text{abs,em}}(E) \equiv \sum_{\mathbf{q}} T_v^{\text{abs,em}}(E, \mathbf{q}), \quad (9.59)$$

to write the tunneling current as follows:

$$\begin{aligned}
I = -\frac{e}{\hbar} \int \frac{dE}{2\pi} & \left( (f_v(E)(1 - f_c(E - \hbar\omega_{\mathbf{k}_0}))(\nu(\hbar\omega_{\mathbf{k}_0}) + 1) \right. \\
& - f_c(E - \hbar\omega_{\mathbf{k}_0})(1 - f_v(E))\nu(\hbar\omega_{\mathbf{k}_0}) T_v^{\text{em}}(E) \\
& + (f_v(E)(1 - f_c(E + \hbar\omega_{\mathbf{k}_0}))\nu(\hbar\omega_{\mathbf{k}_0}) \\
& \left. - f_c(E + \hbar\omega_{\mathbf{k}_0})(1 - f_v(E))(\nu(\hbar\omega_{\mathbf{k}_0}) + 1)) T_v^{\text{abs}}(E) \right). \quad (9.60)
\end{aligned}$$

Furthermore, neglecting the dispersion of  $M'_{\mathbf{q}}$  in Eq. (9.58) for analogous reasons, we may carry out the sum over the wave vectors, yielding

$$\sum_{\mathbf{q}} \exp(i(\mathbf{k}_0 - \mathbf{q}) \cdot (\mathbf{r}' - \mathbf{r})) \rightarrow \Omega \int \frac{d^3q}{(2\pi)^3} \exp(i(\mathbf{k}_0 - \mathbf{q}) \cdot (\mathbf{r}' - \mathbf{r})) = \Omega \delta(\mathbf{r} - \mathbf{r}'). \quad (9.61)$$

In turn, the tunneling probability of Eq. (9.59) simplifies to

$$T_v^{\text{abs,em}}(E) = \Omega |M'_{\mathbf{k}_0}|^2 \int d^3r A_v(\mathbf{r}, \mathbf{r}; E) A_c(\mathbf{r}, \mathbf{r}; E \pm \hbar\omega_{\mathbf{k}_0}). \quad (9.62)$$

involving only the diagonal terms of the spectral density.

## 9.6 A one-dimensional potential profile

To demonstrate the outlined method, we consider the case where the external voltage gives rise to a one-dimensional potential profile, say in the  $x$ -direction, while translational invariance exists in the  $y$ - and  $z$ -directions.

The valence and conduction band envelope Schrödinger equations read

$$\begin{aligned}
\left( \frac{\hbar^2}{2m_v} \frac{d^2}{dx^2} - E_v^\perp(\mathbf{K}) + U_{\text{ext}}(x) \right) \chi_v(\mathbf{r}; \mathbf{K}, E) \\
= E \chi_v(x; \mathbf{K}, E), \quad (9.63)
\end{aligned}$$

$$\begin{aligned}
\left( E_g - \frac{\hbar^2}{2m_c} \frac{d^2}{dx^2} + E_c^\perp(\mathbf{K}) + U_{\text{ext}}(x) \right) \chi_c(\mathbf{r}; \mathbf{K}, E) \\
= E \chi_c(x; \mathbf{K}, E). \quad (9.64)
\end{aligned}$$

Here, the valence band maximum and the conduction band minimum respectively correspond to the energy values  $E = 0$  and  $E = E_g$ , whereas  $\mathbf{K}$  denotes the two-dimensional wave vector reciprocal to  $\mathbf{R} = (y, z)$ . Representing the

kinetic energies in the transverse directions,  $E_{v,c}^\perp(\mathbf{K})$  can be expressed in the effective mass approximation as

$$E_{v,c}^\perp(\mathbf{K}) = \frac{\hbar^2 |\mathbf{K}|^2}{2m_{v,c}}. \quad (9.65)$$

Based on the solutions of the Schrödinger equation, the diagonal part of the spectral densities read

$$A_{v,c}(\mathbf{r}, \mathbf{r}; E) = \int_{-\infty}^{\infty} \frac{d^2 K}{(2\pi)^2} |\chi_{v,c}(\mathbf{r}; \mathbf{K}, E)|^2 \quad (9.66)$$

provided that the wavefunctions are delta normalized

$$\int_{-\infty}^{+\infty} d^3 r \chi_{v,c}^*(\mathbf{r}; \mathbf{K}, E) \chi_{v,c}(\mathbf{r}; \mathbf{K}', E') = (2\pi)^3 \delta(E - E') \delta(\mathbf{K} - \mathbf{K}'). \quad (9.67)$$

Changing integration variables in Eq. (9.66) from  $\mathbf{K}$  to  $E_{v,c}^\perp$  and the polar angle of  $\mathbf{K}$ , we obtain

$$A_{v,c}(\mathbf{r}, \mathbf{r}; E) = \frac{m_v}{\hbar^2} \int_0^\infty \frac{dE_v^\perp}{2\pi} |\chi_{v,c}(\mathbf{r}; E_{v,c}^\perp, E)|^2 \quad (9.68)$$

assuming  $|\chi_{v,c}(\mathbf{r}; \mathbf{K}, E)|^2$  only depends on the magnitude of  $\mathbf{K}$ . Note that we have assumed an isotropic mass. In case the mass is anisotropic with a mass  $m_{v,x}, m_{v,y}, m_{v,z}$  in the  $x, y$  and  $z$  direction respectively,  $m_v$  in Eq. (9.68) has to be replaced with  $\sqrt{m_{v,y} m_{v,z}}$  while  $m_v$  in Eq. (9.63) has to be replaced with  $m_{v,x}$ .

### 9.6.1 Bulk

As a first example, consider a bulk semiconductor where no external field is applied,

$$U_{\text{ext}}(\mathbf{r}) = 0. \quad (9.69)$$

Clearly, no Zener tunneling is expected in the absence of an electric field but the treatment of an unbiased bulk semiconductor illustrates the consistency of the formalism.

The normalized solutions to the envelope Schrödinger equation for  $E < E_v^\perp(\mathbf{K})$  are

$$\chi_v^\pm(\mathbf{r}; \mathbf{K}, E) = \frac{e^{\pm ix \sqrt{-2m_v/\hbar^2(E + E_v^\perp(\mathbf{K}))}}}{\sqrt[4]{-\frac{2\hbar^2}{m_v}(E + E_v^\perp(\mathbf{K}))}} e^{i\mathbf{K}\cdot\mathbf{R}}, \quad (9.70)$$

$$\chi_c^\pm(\mathbf{r}; \mathbf{K}, E) = \frac{e^{\pm ix \sqrt{2m_c/\hbar^2(E - E_g - E_c^\perp(\mathbf{K}))}}}{\sqrt[4]{\frac{2\hbar^2}{m_c}(E - E_g - E_c^\perp(\mathbf{K}))}} e^{i\mathbf{K}\cdot\mathbf{R}} \quad (9.71)$$

where the superscript  $\pm$  denoting right and left running waves had to be added to complete the basis. The valence band spectral function for  $E < 0$  is

$$\begin{aligned} A_v(\mathbf{r}, \mathbf{r}; E) &= \frac{m_v}{\hbar^2} \int_0^E \frac{dE_v^\perp}{2\pi} (|\chi_v^+(\mathbf{r}; E_v^\perp, E)|^2 + |\chi_v^-(\mathbf{r}; E_v^\perp, E)|^2) \\ &= \frac{m_v}{\hbar^2 2\pi} \sqrt{\frac{m_v}{2\hbar^2}} \int_0^E dE_v^\perp \frac{2}{\sqrt{-(E + E_v^\perp)}} \\ &= \frac{m_v^{\frac{3}{2}}}{\hbar^3 \pi} \sqrt{-2E}. \end{aligned} \quad (9.72)$$

The conduction band spectral function for  $E > E_g$  is obtained by substituting  $v$  with  $c$  and  $E$  with  $(E_g - E)$  in the envelope functions

$$A_c(\mathbf{r}, \mathbf{r}; E) = \frac{m_c^{\frac{3}{2}}}{\hbar^3 \pi} \sqrt{2(E - E_g)}. \quad (9.73)$$

The transition probability for  $E_g - \hbar\omega < E < 0$  is

$$T_v^{\text{em}}(E) = 0, \quad (9.74)$$

$$T_v^{\text{abs}}(E) = \Omega^2 |M'_{\mathbf{k}_0}|^2 \frac{2(m_v m_c)^{\frac{3}{2}}}{(\hbar^3 \pi)^2} \sqrt{-E(E + \hbar\omega - E_g)}. \quad (9.75)$$

Eq. (9.74) reflects that no transitions from valence band to the conduction band by phonon emission are possible in bulk semiconductors ( $E_g > 0$ ).

In Eq. (9.75) the condition for transitions to occur is that the phonon energy exceeds the bandgap which is generally not the case in semiconductors. As expected, no phonon-assisted tunneling is present in the absence of an external potential.

### 9.6.2 The 1D Uniform field

A situation more relevant to the problem of Zener tunneling is the application of a uniform electric field or a uniform force  $F$  to an indirect semiconductor

$$U_{\text{ext}}(x) = -Fx. \quad (9.76)$$

The solutions to the effective mass Schrödinger equations are written in terms of the Airy function defined as the bounded solution of the differential equation [23]

$$\frac{d^2}{dx^2} \text{Ai}(x) = x \text{Ai}(x) \quad (9.77)$$

and  $\text{Ai}(0) = 3^{-2/3}/\Gamma(2/3)$ .

The delta normalized solutions to the valence band Schrödinger equation are

$$\chi_v(\mathbf{r}; \mathbf{K}, E) = \sqrt{\frac{2\pi}{Fx_v^2}} \text{Ai} \left( \frac{x}{x_v} + \frac{E + E_v^\perp(\mathbf{K})}{Fx_v} \right) e^{i\mathbf{K}\cdot\mathbf{R}}, \quad (9.78)$$

$$\chi_c(\mathbf{r}; \mathbf{K}, E) = \sqrt{\frac{2\pi}{Fx_c^2}} \text{Ai} \left( -\frac{x}{x_c} - \frac{E - E_g - E_c^\perp(\mathbf{K})}{Fx_v} \right) e^{i\mathbf{K}\cdot\mathbf{R}} \quad (9.79)$$

with

$$x_{v,c}^3 = \frac{\hbar^2}{2m_{v,c}F} \quad (9.80)$$

and where the orthogonality of the Airy functions is taken from Aspnes [24]

$$\int_{-\infty}^{\infty} du \text{Ai}(x+u)\text{Ai}(y+u) = \delta(x-y). \quad (9.81)$$

The spectral functions are

$$A_v(\mathbf{r}, \mathbf{r}; E) = \frac{m_v}{\hbar^2} \frac{1}{x_v} \int_0^\infty du \text{Ai}^2 \left( \frac{x}{x_v} + \frac{E}{Fx_v} + u \right), \quad (9.82)$$

$$A_c(\mathbf{r}, \mathbf{r}; E) = \frac{m_c}{\hbar^2} \frac{1}{x_c} \int_0^\infty du \text{Ai}^2 \left( -\frac{x}{x_c} + \frac{E - E_g}{Fx_c} + u \right). \quad (9.83)$$

Since the field is uniform, the transition probability is independent of the energy  $T_v^{\text{abs,em}}(E) = T_v^{\text{abs,em}}(0)$ . The integration of the spectral functions yielding the transition probability is straightforward using Eq. (34B) from Aspnes [24] and given in the appendix, the result is

$$T_v^{\text{abs,em}}(E) = \Omega |M'_{\mathbf{k}_0}|^2 \frac{A(m_c m_v)^{\frac{3}{2}} F^{\frac{1}{3}}}{8\pi \hbar^4 \hbar^{\frac{2}{3}} \bar{m}^{\frac{2}{3}}} \text{Ai}_3 \left( \frac{2\bar{m}^{\frac{1}{3}}(E_g \mp \hbar\omega)}{F^{\frac{2}{3}} \hbar^{\frac{2}{3}}} \right) \quad (9.84)$$

with  $A$  the device area in  $y$  and  $z$  direction,  $1/\bar{m} = 1/m_v + 1/m_c$  and  $\text{Ai}_3(x)$  is defined as

$$\text{Ai}_3(x) = \frac{1}{2} (\text{Ai}(x) + x\text{Ai}'(x) + x^2\text{Ai}_1(x)) \quad (9.85)$$

with

$$\text{Ai}_1(x) = \int_x^\infty du \text{Ai}(u). \quad (9.86)$$

The transition probability can be approximated for large fields using Eq. (9) from Heigl et al. [25]

$$\text{Ai}_3(x) \approx \frac{x^{-7/4}}{2\sqrt{\pi}} \exp \left( -\frac{2}{3} x^{\frac{3}{2}} \right) \quad (9.87)$$

resulting in the Kane expression:

$$T_{\text{v}}^{\text{abs,em}}(E) \approx \frac{\Omega |M'_{\mathbf{k}_0}|^2 A(m_{\text{c}} m_{\text{v}})^{\frac{3}{2}} F^{\frac{3}{2}}}{2^{\frac{23}{4}} \pi^{\frac{3}{2}} \hbar^{\frac{7}{2}} (E_{\text{g}} \mp \hbar\omega)^{\frac{7}{4}} \bar{m}^{\frac{5}{4}}} \exp\left(-\frac{4}{3} \frac{\sqrt{2\bar{m}}(E_{\text{g}} \mp \hbar\omega)^{\frac{3}{2}}}{F\hbar}\right). \quad (9.88)$$

### 9.6.3 1D Non-uniform fields: the WKB approximation

Next, we calculate the phonon-assisted tunneling current in a one-dimensional device structure with external potential

$$U_{\text{ext}}(\mathbf{r}) = U_{\text{ext}}(x) \quad (9.89)$$

using the WKB approximation [26]. The envelope functions within the WKB approximation are

$$\chi_{\text{v}}(\mathbf{r}; \mathbf{K}, E) = \frac{\exp\left(-\int_{x_{\text{tv}}}^x dx' \sqrt{\frac{2m_{\text{v}}}{\hbar^2}(E - U_{\text{ext}}(x') + E_{\text{v}}^{\perp})}\right)}{\sqrt[4]{\frac{2\hbar^2}{m_{\text{v}}}(E - U_{\text{ext}}(x) + E_{\text{v}}^{\perp})}} e^{i\mathbf{K}\cdot\mathbf{R}}, \quad (9.90)$$

$$\chi_{\text{c}}(\mathbf{r}; \mathbf{K}, E) = \frac{\exp\left(\int_{x_{\text{tc}}}^x dx' \sqrt{-\frac{2m_{\text{c}}}{\hbar^2}(E - E_{\text{g}} - U_{\text{ext}}(x') - E_{\text{c}}^{\perp})}\right)}{\sqrt[4]{-\frac{2\hbar^2}{m_{\text{c}}}(E - E_{\text{g}} - U_{\text{ext}}(x') - E_{\text{c}}^{\perp})}} e^{i\mathbf{K}\cdot\mathbf{R}} \quad (9.91)$$

where  $x_{\text{tv,c}}$  indicate the location of the turning points as illustrated in Fig. 9.2. In the spirit of the WKB approximation, the normalization can be verified by comparing to the uniform field case. The turning points are determined by the equations

$$\begin{aligned} E - U_{\text{ext}}(x_{\text{tv}}) + E_{\text{v}}^{\perp} &= 0, \\ E_{\text{g}} + U_{\text{ext}}(x_{\text{tc}}) - E + E_{\text{c}}^{\perp} &= 0. \end{aligned} \quad (9.92)$$

In appendix B, an approximate expression for the spectral functions and the tunneling probability is obtained. The results are

$$A_{\text{v}}(\mathbf{r}, \mathbf{r}; E) \approx \frac{m_{\text{v}}}{4\pi\hbar^2} \frac{\exp\left(-2 \int_{x_{\text{tv}0}}^x dx' \kappa_{\text{v}}(x'; E)\right)}{\kappa_{\text{v}}(x; E) \int_{x_{\text{tv}0}}^x dx' / \kappa_{\text{v}}(x'; E)}, \quad (9.93)$$

$$A_{\text{c}}(\mathbf{r}, \mathbf{r}; E) \approx \frac{m_{\text{c}}}{4\pi\hbar^2} \frac{\exp\left(2 \int_{x_{\text{tc}0}}^x dx' \kappa_{\text{c}}(x'; E)\right)}{\kappa_{\text{c}}(x; E) \int_{x_{\text{tc}0}}^x dx' / \kappa_{\text{c}}(x'; E)} \quad (9.94)$$

and

$$T_{\text{v}}^{\text{abs,em}}(E) \approx \frac{A\Omega |M'_{\mathbf{k}_0}|^2 (m_{\text{v}} m_{\text{c}})^{\frac{1}{2}}}{2^{\frac{19}{4}} \pi^{\frac{3}{2}} \hbar^{\frac{3}{2}} (E \mp \hbar\omega)^{\frac{3}{4}} \bar{m}^{\frac{1}{4}} \sqrt{U'_{\text{ext}}(x_{\text{max}}^{\pm})}}$$

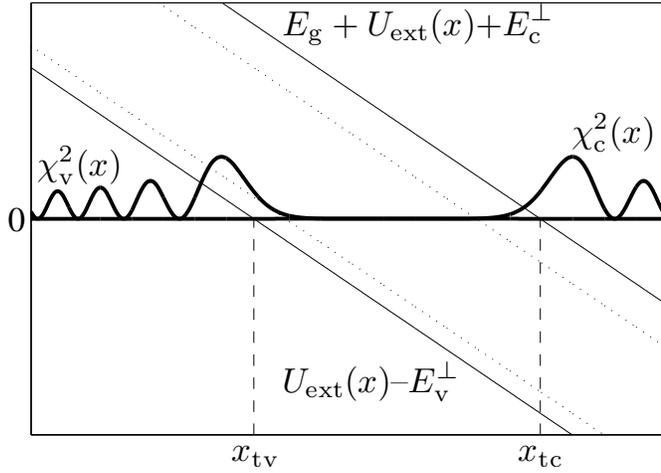


Figure 9.2: Illustration of the turning points  $x_{tv}$  and  $x_{tc}$  for  $E = 0$  appearing as the solutions to  $U_{\text{ext}}(x_{tv}) - E_v^\perp = 0$  and  $E_g + U_{\text{ext}}(x_{tc}) + E_c^\perp = 0$  respectively.

$$\times \frac{\exp\left(-2 \int_{x_{tv0}}^{x_{\text{max}}} dx' \kappa_c(x'; E)\right) \exp\left(2 \int_{x_{tc0}}^{x_{\text{max}}} dx' \kappa_v(x'; E \pm \hbar\omega_{\mathbf{k}_0})\right)}{\int_{x_{tv0}}^{x_{\text{max}}} dx' / \kappa_v(x'; E)} \frac{\int_{x_{tc0}}^{x_{\text{max}}} dx' / \kappa_c(x'; E \pm \hbar\omega_{\mathbf{k}_0})}{\int_{x_{tc0}}^{x_{\text{max}}} dx' / \kappa_c(x'; E \pm \hbar\omega_{\mathbf{k}_0})} \quad (9.95)$$

where  $x_{tv,c0}$  are the turning points  $x_{tv,c}$  for  $E_{v,c}^\perp = 0$ .  $x = x_{\text{max}}^\pm$  denotes the point where the imaginary wave vector of the valence and conduction band equal each other:  $\kappa_v(x_{\text{max}}^\pm; E) = \kappa_c(x_{\text{max}}^\pm; E \pm \hbar\omega_{\mathbf{k}_0})$ .  $U'_{\text{ext}}(x_{\text{max}}^\pm)$  denotes the first derivative of the external potential at  $x = x_{\text{max}}^\pm$ . The imaginary wavevectors  $\kappa_{v,c}$  are defined as

$$\kappa_v(x; E) = \sqrt{\frac{2m_v}{\hbar^2} (E - U_{\text{ext}}(x))}, \quad (9.96)$$

$$\kappa_c(x; E) = \sqrt{\frac{2m_c}{\hbar^2} (E_g + U_{\text{ext}}(x) - E)}. \quad (9.97)$$

Taking the limit of the tunneling probability for small uniform fields  $U_{\text{ext}}(x) = -Fx$ , the WKB expression approaches the uniform field expression obtained in the previous section.

#### 9.6.4 Non-uniform fields: numerical evaluation

In a real semiconductor under a one dimensional external potential, the external potential can be taken uniform in two contact regions on the right and left hand side. The discretized wavefunction can be determined by discretizing the Schrödinger equation and applying so-called transmitting boundary conditions [27]. In a Zener tunnel junction, the wave will not be transmitted but

completely reflected by the potential barrier. The wavefunction will only have a decaying component in the contact opposite to the side where the wave is injected.

The incoming component of the wavefunction is given by the bulk wavefunction from Eq. (9.70). The spectral functions are obtained by integrating the wavefunctions with respect to energy according to Eq. (9.66).

## 9.7 Discussion

### 9.7.1 Generation rate

Rather than the transition probability  $T_v^{\text{abs,em}}$  used in this paper, the so-called Band-to-Band Tunneling (BTBT) generation/recombination rate per unit volume  $G$  is usually presented to calculate the Zener tunneling rate. The current as a function of generation rate is given by

$$I = I_v^{\text{em}} + I_v^{\text{abs}} \quad (9.98)$$

with

$$\begin{aligned} I_v^{\text{abs,em}} = -e \int d^3r G_v^{\text{abs,em}}(\mathbf{r}) \\ \left( f_v(U_{\text{ext}}(\mathbf{r})) (1 - f_c(U_{\text{ext}}(\mathbf{r}) \pm \hbar\omega_{\mathbf{k}_0})) \left( \frac{1}{2} \mp \frac{1}{2} + \nu(\hbar\omega_{\mathbf{k}_0}) \right) \right. \\ \left. - f_c(U_{\text{ext}}(\mathbf{r}) \pm \hbar\omega_{\mathbf{k}_0}) (1 - f_v(U_{\text{ext}}(\mathbf{r}))) \left( \frac{1}{2} \pm \frac{1}{2} + \nu(\hbar\omega_{\mathbf{k}_0}) \right) \right). \quad (9.99) \end{aligned}$$

$G_v^{\text{abs,em}}(\mathbf{r})$  denotes the generation rate for an electron with a tunnel path starting at  $\mathbf{r}$  and ending at  $\mathbf{r} + \mathbf{l}^{\text{abs,em}}(\mathbf{r})$  such that  $U_{\text{ext}}(\mathbf{r}) = E_g \mp \hbar\omega_{\mathbf{k}_0} + U_{\text{ext}}(\mathbf{r} + \mathbf{l}^{\text{abs,em}}(\mathbf{r}))$ .

Comparing the semi-classical equation with the expression for the current, the relation between the transition probability and the generation rate is given by

$$G_v(x) = \frac{U'_{\text{ext}}(x)}{2\pi\hbar A} T_v^{\text{abs,em}}(U_{\text{ext}}(x)). \quad (9.100)$$

In the low field limit (Kane/Keldysh result) and assuming an anisotropic effective mass,

$$G_{v,\text{Kane}}^{\text{abs,em}} = G_0^{\text{abs,em}} F^{\frac{5}{2}} \exp\left(-\frac{4}{3} \frac{\sqrt{2\bar{m}_x} (E_g \mp \hbar\omega)^{\frac{3}{2}}}{F\hbar}\right) \quad (9.101)$$

with

$$G_0^{\text{abs,em}} = Ag_0\Omega |M'_{\mathbf{k}_0}|^2 \frac{\sqrt{m_{v,x}m_{v,y}m_{v,z}m_{c,x}m_{c,y}m_{c,z}}}{2^{\frac{27}{4}} \pi^{\frac{5}{2}} \hbar^{\frac{7}{2}} (E_g \mp \hbar\omega)^{\frac{7}{4}} \bar{m}_x^{\frac{5}{4}}} \quad (9.102)$$

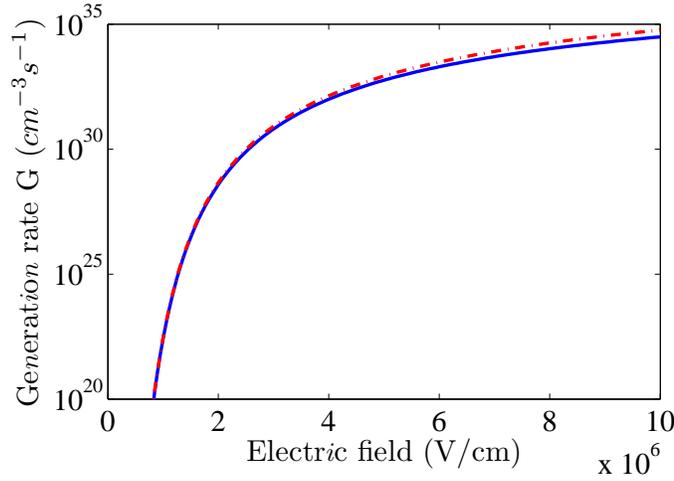


Figure 9.3: Generation rate calculated using the uniform field model (solid line) and the Kane limit (dashed line) for Si with  $E_g = 1.12$  eV,  $m_v = (.16, .49, .49)m_0$  and  $m_c = (.1905, .9163, .9163)m_0$  for the electronic parameters and  $D\mathbf{k}_0 = 6 \times 10^8$  eV/cm,  $\hbar\omega = 57.6 \times 10^{-3}$  eV,  $\rho_{\text{Si}} = 2.328$  cm $^{-3}$  for the phonon interaction [11]. Assuming tunneling along [100], the total degeneracy prefactor is 16.

where  $g_0$  is a degeneracy factor. The factor of 2 for spin degeneracy has to be accounted for in  $g_0$  together with the appropriate valley degeneracy and the phonon degeneracy. The generation rate in the uniform field case is

$$G_{v,U}^{\text{abs,em}} = G_0^{\text{abs,em}} F^{\frac{5}{2}} I_2 \left( \frac{2\bar{m}_x^{\frac{1}{3}}(E_g \mp \hbar\omega)}{(\hbar F)^{\frac{2}{3}}} \right) \quad (9.103)$$

with

$$I_2(x) = 2\sqrt{\pi}x^{\frac{7}{4}}\text{Ai}_3(x). \quad (9.104)$$

In Fig. 9.3, the generation rate is plotted as a function of the electric field for Si using the uniform field model and the low field limit. The degeneracy factor is 2 due to spin, 4 due to the four X valleys with their transversal mass in the [100] direction and 2 due to TO phonon degeneracy. Since the tunneling probability is strongly dependent on the effective mass the tunneling to the two X valleys with their longitudinal mass in the [100] direction can be neglected. Similarly the tunneling from the heavy hole band can be neglected. At high fields, the error of the Kane model due to the use of the asymptotic expression for  $\text{Ai}_3(x)$  can be observed. For the maximum field used in the plot (10MV/cm), the difference between both models is a factor of 1.86.

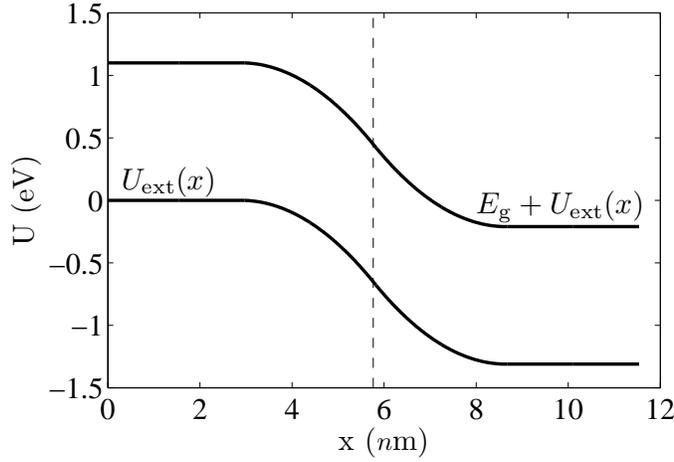


Figure 9.4: Parabolic potential as obtained in a abrupt  $p-n$  diode using the Shockley approximation. Donor and acceptor doping levels are  $10^{20} \text{ cm}^{-3}$ , the dashed line indicates the junction.

### 9.7.2 Comparing the different models for a $p-n$ diode

In an abrupt  $p-n$  diode, the Shockley approximation can be used to determine the potential profile consisting of two parabolic sections as shown in Fig. 9.4. The fermi levels are determined by imposing charge neutrality in the contacts where the carrier concentration can be calculated from the bulk spectral functions given in Eq. (9.82) and Eq. (9.83).

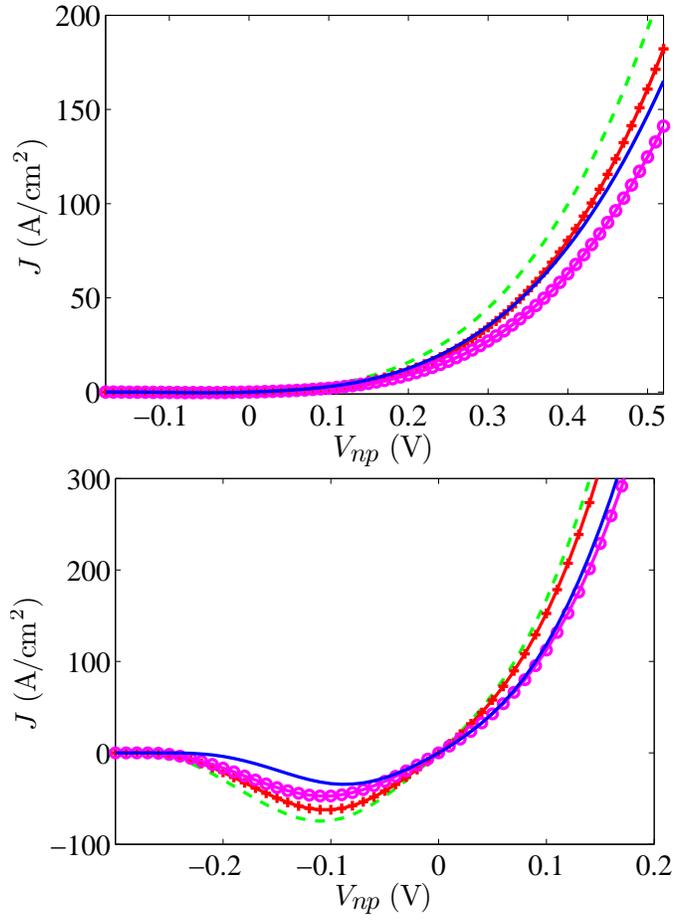
A comparison between the numerical results and the uniform field model, the limit of the uniform field model and the non-uniform model using the WKB approximation is given in Fig. 9.5. For the uniform field model in the presence of a non-uniform field, the usual approximation

$$F \approx E_g/l_{\text{tun}} \quad (9.105)$$

with  $U_{\text{ext}}(x) = U_{\text{ext}}(x + l_{\text{tun}}(x)) + E_g$  was used. The non-uniform model from this paper is

$$G_{\text{v,NU}}^{\text{abs,em}} = G_0^{\text{abs,em}} \frac{\hbar^2 (E_g \mp \hbar\omega_{\mathbf{k}_0}) \bar{m}_x}{m_{\text{v},x} m_{\text{c},x} \sqrt{U'_{\text{ext}}(x_{\text{max}})}} \times \frac{\exp\left(-2 \int_{x_{\text{tv}0}}^{x_{\text{max}}} dx' \kappa_{\text{c}}(x'; E)\right) \exp\left(2 \int_{x_{\text{tc}0}}^{x_{\text{max}}} dx' \kappa_{\text{v}}(x'; E \pm \hbar\omega_{\mathbf{k}_0})\right)}{\int_{x_{\text{tv}0}}^{x_{\text{max}}} dx' / \kappa_{\text{v}}(x'; E) \int_{x_{\text{tc}0}}^{x_{\text{max}}} dx' / \kappa_{\text{c}}(x'; E \pm \hbar\omega_{\mathbf{k}_0})} \quad (9.106)$$

The Schenk model [8,25] has a different dependence on the field strength ( $F^{9/2}$ ) compared to the Kane/Keldysh/WKB result and for this reason we will not compare with the Schenk model.



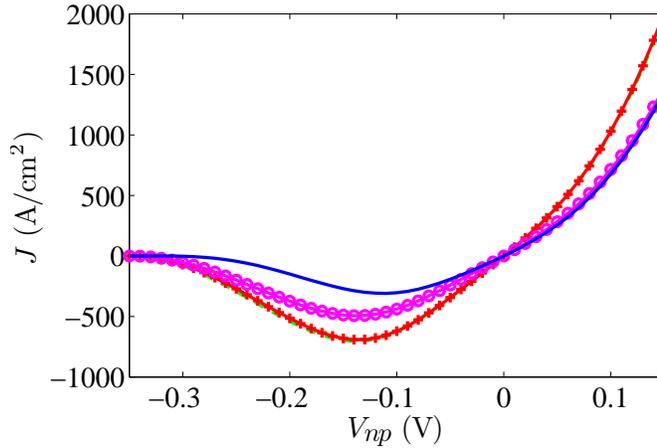


Figure 9.5: Calculation of the current for a symmetrically doped Si diode with a doping level of  $5 \times 10^{19} \text{cm}^{-3}$ ,  $10^{20} \text{cm}^{-3}$  and  $1.5 \times 10^{20} \text{cm}^{-3}$ . The same Si parameters were used as in Fig. 9.3. The different current curves are obtained using the semi-classical Kane (Eq. 9.101, solid line with +), Uniform (Eq. (9.103), solid line with o) and WKB (Eq. (9.106), dashed line) model with respect to the direct numerical calculation of Eq. (9.62) (solid line). In the case of the highest doping level the Kane and WKB model coincide.

In Fig. 9.5 it can be seen that of the three semi-classical models the uniform field model ( $G_{v,U}$ ) is the best. Despite disregarding the non-uniformities, the uniform field model is a better approximation than the WKB model and this can be explained as follows. At low bias ( $V_{pn} \approx 0$ ), taking the integration limit  $|\mathbf{K}| \rightarrow \infty$  is too optimistic resulting in an overestimation of the current in all semi-classical models. But on the other hand the uniform field model is pessimistic about the tunnel probability compared to the WKB approximation and this partly compensates for the overestimation of the integration limits. In the high bias regime ( $V_{pn} \gg 0$ ), the approximation of the potential relevant for tunneling by a uniform field is reasonable since the field does not change much in the region where the field is the highest. This can be verified by the similarity of the Kane model and the WKB models. The difference between the uniform field model and the Kane model is that the integrals over  $x$  and  $\mathbf{K}$  are treated exactly in the uniform field case based on the properties of the Airy functions.

Nevertheless, for the degenerately doped semiconductors, none of the semiclassical models are a good approximation in the forward bias regime. When an accurate estimation of the current magnitude or shape for the forward bias or the low bias regime is required, a direct numerical procedure is required.

## 9.8 Conclusion

A general formalism to calculate the phonon-assisted Zener tunneling current in non-uniform fields using spectral functions was presented. The expression using the spectral function enables the calculation of the tunneling probability to two- or three-dimensional potential profiles. The evaluation of the current can be done using the expression for the current (Eq. (9.60)) and the expression for the transition probability (Eq. (9.62)).

In the presence of a one-dimensional external potential, an analytical expression for the tunneling can be obtained for uniform fields using Airy functions and non-uniform fields using the WKB approximation. It is shown that in the weak and uniform field limit, the formalism from this paper reduces to the indirect Kane result. Furthermore, an improved uniform field model is derived and an expression using the WKB approximation are obtained. Comparing the different approximations in the case of a  $p - n$  junction, the improved uniform field model is shown to be a better approximation than the model using the WKB approximation. But for low bias conditions, no approximate model accurately describes the tunneling current and a direct numerical treatment is essential even in the one-dimensional case.

## Acknowledgements

William Vandenberghe gratefully acknowledges the support of a Ph.D. stipend from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen).

## Appendix A: Tunneling probability in a uniform field

In this appendix, we derive the transition probability for the uniform field and a non-uniform field starting from Eq. (9.62). Using the translation invariance in  $y$  and  $z$  direction, the integration in  $y$  and  $z$  corresponds to a simple multiplication by the area  $A$ :

$$T_v^{\text{abs,em}}(E) = \Omega |M'_{\mathbf{k}_0}|^2 A \int_{-\infty}^{\infty} dx A_v(\mathbf{r}, \mathbf{r}; 0) A_c(\mathbf{r}, \mathbf{r}; \pm \hbar \omega)$$

## 9.A Tunneling probability in a uniform field

Introducing the spectral functions for the uniform field from Eq. (9.82) and Eq. (9.83) the tunneling probability reads

$$T_v^{\text{abs,em}}(E) = \Omega |M'_{\mathbf{k}_0}|^2 \frac{m_c m_v A}{\hbar^4 x_v x_c} \int_{-\infty}^{\infty} dx \int_0^{\infty} du_1 \int_0^{\infty} du_2 \text{Ai}^2(x/x_v + u_1) \times \text{Ai}^2\left(-\left(x + \frac{\pm \hbar \omega_{\mathbf{k}_0} - E_g}{F}\right)/x_c + u_2\right). \quad (9.107)$$

The integral over  $x$  can be performed using Eq. (B34b) from Aspnes [24]

$$T_v^{\text{abs,em}}(E) = \Omega |M'_{\mathbf{k}_0}|^2 \frac{m_c m_v A}{\hbar^4 x_v x_c} \quad (9.108)$$

$$\times \int_0^{\infty} du_1 \int_0^{\infty} du_2 \frac{x_v}{4\pi \sqrt{x_v/x_c}} \text{Ai}_1\left(\frac{2^{\frac{2}{3}}(x_v/x_c u_1 - \frac{\pm \hbar \omega_{\mathbf{k}_0} - E_g}{F x_c} + u_2)}{(1 + (x_v/x_c)^3)^{\frac{1}{3}}}\right) \\ = \Omega |M'_{\mathbf{k}_0}|^2 \frac{m_c m_v A}{4\pi \hbar^4 \sqrt{x_v x_c}} \quad (9.109)$$

$$\times \int_0^{\infty} du_1 \int_0^{\infty} du_2 \text{Ai}_1\left(\frac{2^{\frac{2}{3}}(x_v u_1 + x_c u_2 + \frac{E_g \mp \hbar \omega_{\mathbf{k}_0}}{F})}{(x_v^3 + x_c^3)^{\frac{1}{3}}}\right). \quad (9.110)$$

Defining the integral

$$\text{Ai}_3(x) = \int_0^{\infty} du_1 \int_0^{\infty} du_2 \text{Ai}_1(x + u_1 + u_2) \\ = \int_0^{\infty} \frac{d(u_1 + u_2)}{\sqrt{2}} \int_{-(u_1 + u_2)}^{u_1 + u_2} \frac{d(u_1 - u_2)}{\sqrt{2}} \text{Ai}_1(x + u_1 + u_2) \\ = \int_0^{\infty} du u \text{Ai}_1(x + u) \\ = \frac{1}{2} (\text{Ai}(x) + x \text{Ai}'(x) + x^2 \text{Ai}_1(x)) \quad (9.111)$$

where the last equality comes from Schenk [8], the transition probability can be written as

$$T_v^{\text{abs,em}}(E) = \Omega |M'_{\mathbf{k}_0}|^2 \frac{m_c m_v A}{4\pi \hbar^4 \sqrt{x_v x_c}} \frac{(x_v^3 + x_c^3)^{\frac{2}{3}}}{x_c x_v 2^{\frac{4}{3}}} \text{Ai}_3\left(\frac{2^{\frac{2}{3}}(E_g \mp \hbar \omega_{\mathbf{k}_0})}{F(x_v^3 + x_c^3)^{\frac{1}{3}}}\right). \quad (9.112)$$

Substituting the values for  $x_c$  and  $x_v$ , Eq. (9.112) reads

$$T_v^{\text{abs,em}}(E) = \Omega |M'_{\mathbf{k}_0}|^2 \frac{A(m_c m_v)^{\frac{3}{2}} F^{\frac{1}{3}}}{8\pi \hbar^4 \hbar^{\frac{2}{3}} \bar{m}^{\frac{2}{3}}} \text{Ai}_3\left(\frac{2\bar{m}^{\frac{1}{3}}(E_g \mp \hbar \omega_{\mathbf{k}_0})}{F^{\frac{2}{3}} \hbar^{\frac{2}{3}}}\right). \quad (9.113)$$

with  $1/\bar{m} = 1/m_v + 1/m_c$ .

## 9.B Tunneling probability in a non-uniform field using the WKB approximation

In the WKB approximation, the integral over  $E_{v,c}^\perp$  yielding the spectral functions is generally a nonelementary integral. To obtain a simplified expression, approximate the spectral functions by making the first order Taylor expansion of the argument of the exponential in Eq. (9.90) at  $E_v^\perp = 0$ ,

$$\begin{aligned} & \int_{x_1}^x dx' \sqrt{\frac{2m_v}{\hbar^2} (E - U_{\text{ext}}(x') + E_v^\perp)} \\ & \approx \int_{x_1}^x dx' \sqrt{\frac{2m_v}{\hbar^2} (E - U_{\text{ext}}(x'))} \\ & + \frac{E_v'}{2} \int_{x_1}^x dx' \sqrt{\frac{2m_v}{\hbar^2}} \frac{1}{\sqrt{E - U_{\text{ext}}(x')}}. \end{aligned} \quad (9.114)$$

Introduce the imaginary wave vector

$$\kappa_v(x; E) = \sqrt{\frac{2m_v}{\hbar^2} (E - U_{\text{ext}}(x))}, \quad (9.115)$$

and take the variation of the exponential to be much faster than that of the denominator in Eq. (9.90). The spectral function is approximated as

$$A_v(\mathbf{r}, \mathbf{r}; E) \approx \frac{m_v}{4\pi\hbar^2} \frac{\exp(-2 \int_{x_{tv0}}^x dx' \kappa_v(x'; E))}{\kappa_v(x; E) \int_{x_{tv0}}^x dx' / \kappa_v(x'; E)} \quad (9.116)$$

where  $x_{tv,c0}$  are the turning points  $x_{tv,c}$  for  $E_{v,c}^\perp = 0$ . The conduction band spectral function is given by

$$A_c(\mathbf{r}, \mathbf{r}; E) \approx \frac{m_c}{4\pi\hbar^2} \frac{\exp(2 \int_{x_{tc0}}^x dx' \kappa_c(x'; E))}{\kappa_c(x; E) \int_{x_{tc0}}^x dx' / \kappa_c(x'; E)} \quad (9.117)$$

with

$$\kappa_c(x; E) = \sqrt{\frac{2m_c}{\hbar^2} (E_g + U_{\text{ext}}(x) - E)}. \quad (9.118)$$

The product of the spectral functions appearing under the integral in the expression for the tunneling probability

$$T_v^{\text{abs,em}}(E) = \Omega |M'_{\mathbf{k}_0}|^2 A \int_{-\infty}^{\infty} dx A_v(\mathbf{r}, \mathbf{r}; 0) A_c(\mathbf{r}, \mathbf{r}; \pm\hbar\omega) \quad (9.119)$$

can be rewritten as a prefactor  $f(x)$  and an exponential  $e^{g(x)}$

$$A_v(\mathbf{r}, \mathbf{r}; 0)A_c(\mathbf{r}, \mathbf{r}; \pm\hbar\omega) = f(x)e^{g(x)}. \quad (9.120)$$

Observing the exponential terms in Eq. (9.116) and Eq. (9.117),

$$g(x) = 2 \int_{x_{tc}}^x dx' \kappa_c(x'; E \pm \hbar\omega_{k_0}) - 2 \int_{x_{tv}}^x dx' \kappa_v(x'; E). \quad (9.121)$$

The exponential will decay rapidly from its maximum which is found at  $x = x_{\max}$ , which is determined by the condition  $dg(x)/dx = 0$ :

$$\begin{aligned} \kappa_v(x_{\max}^{\pm}; E) &= \kappa_c(x_{\max}^{\pm}; E \pm \hbar\omega_{k_0}) = \kappa_{\max}^{\pm} \\ &= \sqrt{\frac{2\bar{m}}{\hbar^2}(E_g \mp \hbar\omega_{k_0})}. \end{aligned} \quad (9.122)$$

The argument of the exponential can be expanded to second order at  $x_{\max}$

$$\begin{aligned} g(x) \approx 2 \int_{x_{tc0}}^{x_{\max}^{\pm}} dx' \kappa_c(x'; E \pm \hbar\omega_{k_0}) - 2 \int_{x_{tv0}}^{x_{\max}^{\pm}} dx' \kappa_v(x'; E) \\ - \frac{x^2}{2} \frac{2(m_v + m_c)}{\hbar^2} \frac{U'_{\text{ext}}(x_{\max}^{\pm})}{\kappa_{\max}^{\pm}} \end{aligned} \quad (9.123)$$

The prefactor  $f(x)$  can be taken to be slowly varying and its value at  $x_{\max}$  can be used. Integrate the exponentials using the gaussian integral  $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$  to obtain the final expression for the WKB tunneling probability,

$$\begin{aligned} T_v^{\text{abs,em}}(E) \approx & \frac{A\Omega|M'_{\mathbf{k}_0}|^2(m_v m_c)^{\frac{1}{2}}}{2^{\frac{19}{4}} \pi^{\frac{3}{2}} \hbar^{\frac{3}{2}} (E \mp \hbar\omega)^{\frac{3}{4}} \bar{m}^{\frac{1}{4}} \sqrt{U'_{\text{ext}}(x_{\max}^{\pm})}} \\ & \times \frac{\exp\left(-2 \int_{x_{tv0}}^{x_{\max}} dx' \kappa_c(x'; E)\right)}{\int_{x_{tv0}}^{x_{\max}} dx' / \kappa_v(x'; E)} \\ & \times \frac{\exp\left(2 \int_{x_{tc0}}^{x_{\max}} dx' \kappa_v(x'; E \pm \hbar\omega_{k_0})\right)}{\int_{x_{tc0}}^{x_{\max}} dx' / \kappa_c(x'; E \pm \hbar\omega_{k_0})} \end{aligned} \quad (9.124)$$

## Bibliography

- [1] W. M. Reddick and G. A. J. Amaratunga, "Silicon surface tunnel transistor," *Applied Physics Letters*, vol. 67, no. 4, pp. 494–496, July 1995.

- [2] C. Zener, "A theory of the electrical breakdown of solid dielectrics," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 145, no. 855, pp. 523–529, July 1934.
- [3] L. Keldysh, "Behavior of non-metallic crystals in strong electric fields," *Sov. Phys. JETP*, vol. 6, p. 763, 1958.
- [4] E. O. Kane, "Zener tunneling in semiconductors," *Journal of Physics and Chemistry of Solids*, vol. 12, pp. 181–188, 1959.
- [5] W. Vandenberghe, B. Sorée, W. Magnus, and G. Groeseneken, "Zener tunneling in semiconductors under nonuniform electric fields," *Journal of Applied Physics*, vol. 107, no. 5, p. 054520, 2010. [Online]. Available: <http://dx.doi.org/10.1063/1.3311550>
- [6] L. Keldysh, "Influence of the lattice vibrations of a crystal on the production of electron-hole pairs in a strong electric field," *Sov. Phys. JETP*, vol. 7, p. 665, 1959.
- [7] E. O. Kane, "Theory of tunneling," *Journal of Applied Physics*, vol. 32, no. 1, pp. 83–91, 1961.
- [8] A. Schenk, "Rigorous theory and simplified model of the band-to-band tunneling in silicon," *Solid-State Electronics*, vol. 36, no. 1, pp. 19–34, 1993.
- [9] R. Kubo, "Statistical-mechanical theory of irreversible processes. i," *Journal of the Physical Society of Japan*, vol. 12, p. 570, 1957.
- [10] R. Enderlein and K. Peuker, "On the theory of the electric conductivity of solids in a strong electric field," *phys. stat. sol. (b)*, vol. 48, no. 1, pp. 231–241, 1971. [Online]. Available: <http://dx.doi.org/10.1002/pssb.2220480122>
- [11] C. Rivas, R. Lake, G. Klimeck, W. R. Frensley, M. V. Fischetti, P. E. Thompson, S. L. Rommel, and P. R. Berger, "Full-band simulation of indirect phonon assisted tunneling in a silicon tunnel diode with delta-doped contacts," *Applied Physics Letters*, vol. 78, no. 6, pp. 814–816, 2001. [Online]. Available: <http://dx.doi.org/10.1063/1.1343500>
- [12] M. Fischetti, T. O'Regan, S. Narayanan, C. Sachs, S. Jin, J. Kim, and Y. Zhang, "Theoretical study of some physical aspects of electronic transport in nmosfets at the 10-nm gate-length," *Electron Devices, IEEE Transactions on DOI - 10.1109/TED.2007.902722*, vol. 54, no. 9, pp. 2116–2136, 2007. [Online]. Available: [10.1109/TED.2007.902722](http://dx.doi.org/10.1109/TED.2007.902722)

- [13] A. S. Verhulst, B. Sorée, D. Leonelli, W. G. Vandenberghe, and G. Groeseneken, “Modeling the single-gate, double-gate, and gate-all-around tunnel field-effect transistor,” *Journal of Applied Physics*, vol. 107, no. 2, p. 024518, 2010. [Online]. Available: <http://dx.doi.org/10.1063/1.3277044>
- [14] W. Magnus and W. Schoenmaker, *Quantum transport in submicron devices: a theoretical introduction*. Springer series series on Solid-State Science, 2002.
- [15] M. Nedjalkov, S. Kosina, S. Selberherr, C. Ringhofer, and D. K. Ferry, “Unified particle approach to wigner- boltzmann transport in small semiconductor devices,” *Physical review B*, vol. 70, pp. 115 319–115 335, 2004.
- [16] F. Brosens and W. Magnus, “Newtonian trajectories: A powerful tool for solving quantum dynamics,” *Solid-State Communications*, vol. 150, pp. 2102–2105, 2010.
- [17] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, “Single and multiband modeling of quantum electron transport through layered semiconductor devices,” *J. Appl. Phys.*, vol. 81, no. 12, pp. 7845–7869, Jun. 1997. [Online]. Available: <http://link.aip.org/link/?JAP/81/7845/1>
- [18] M. V. Fischetti, “Master-equation approach to the study of electronic transport in small semiconductor devices,” *Phys. Rev. B*, vol. 59, no. 7, pp. 4901–, Feb. 1999. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevB.59.4901>
- [19] B. Sorée, W. Magnus, and W. Schoenmaker, “Energy and momentum balance equations: An approach to quantum transport in closed circuits,” *Phys. Rev. B*, vol. 66, no. 3, pp. 035 318–, Jul. 2002. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevB.66.035318>
- [20] W. G. Vandenberghe, B. Soree, W. Magnus, G. Groeseneken, and M. V. Fischetti, “Impact of field-induced quantum confinement in tunneling field-effect devices,” *Applied Physics Letters*, vol. 98, no. 14, pp. 143 503 – 143 503–3, apr 2011.
- [21] M. Cardona and F. H. Pollak, “Energy-band structure of germanium and silicon: The k.p method,” *Phys. Rev.*, vol. 142, no. 2, pp. 530–, Feb. 1966. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRev.142.530>
- [22] C. Kittel, *Introduction to Solid State Physics*. John-Wiley and Sons, 1976.
- [23] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, ninth dover printing, tenth gpo printing ed. New York: Dover, 1964.

- [24] D. E. Aspnes, “Electric-field effects on optical absorption near thresholds in solids,” *Phys. Rev.*, vol. 147, no. 2, pp. 554–, Jul. 1966. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRev.147.554>
- [25] A. Heigl, A. Schenk, and G. Wachutka, “Correction to the schenk model of band-to-band tunneling in silicon applied to the simulation of nanowire tunneling transistors,” in *Computational Electronics, 2009. IWCE '09. 13th International Workshop on DOI - 10.1109/IWCE.2009.5091099*, 2009, pp. 1–4. [Online]. Available: [10.1109/IWCE.2009.5091099](http://dx.doi.org/10.1109/IWCE.2009.5091099)
- [26] F. W. J. Olver, *Asymptotics and special functions*. Academic Press, 1974.
- [27] C. S. Lent and D. J. Kirkner, “The quantum transmitting boundary method,” *Journal of Applied Physics*, vol. 67, no. 10, pp. 6353–6359, 1990. [Online]. Available: <http://dx.doi.org/10.1063/1.345156>

### List of corrections and clarifications

- The operator representation of Eq. (9.28) should read

$$A_{v,c}(\mathbf{r}, \mathbf{r}') = 2\pi \langle \mathbf{r} | \delta(E - H_{v,c}) | \mathbf{r}' \rangle. \quad (9.125)$$

- Eq. (9.65) does not represent the kinetic energy of an electron in the valence band but its magnitude. The equations used in the text are correct using this convention.



## Chapter 10

# Field-induced quantum confinement in TFETs

Impact of field-induced quantum confinement in  
tunneling field-effect devices

William G. Vandenberghe, Bart Sorée, Wim Magnus,  
Guido Groeseneken, and Massimo V. Fischetti

Published in *Applied Physics Letters*, 98 (14), art nr. 143503,  
April 2011.

The following article appeared in *Applied Physics Letters*, 98 (14), art. nr. 143503, April 2011 and may be found at <http://link.aip.org/link/doi/10.1063/1.3573812>.

©2011 American Institute of Physics. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the American Institute of Physics.

### Contributions of first author

- Development of the theory
- Development of the code
- Creating figures
- Text writing and editing

## Abstract

Being the working principle of a tunnel field-effect transistor, band-to-band tunneling is given a rigorous quantum mechanical treatment to incorporate confinement effects, multiple electron and hole valleys, and interactions with phonons. The model reveals that the strong band bending near the gate dielectric, required to create short tunnel paths, results in quantization of the energy bands. Comparison with semi-classical models reveals a big shift in the onset of tunneling. The effective mass difference of the distinct valleys is found to reduce the sub-threshold swing steepness.

Tunnel field-effect transistors (TFETs) are considered potential candidates to break the limit of the 60 mV/decade subthreshold swing in a metal-oxide-semiconductor field-effect transistor (MOSFET) operating at room temperature. In principle TFETs do not suffer from this limit, since any increase of the gate voltage also increases the band-to-band tunneling (BTBT) probability while, in contrast, the barrier seen by majority carriers in a conventional MOSFET is lowered. Next to the fabrication of TFET devices [1–3], various simulations have been performed [4–6] to understand and improve the TFET performance but up to now, the effect of quantum confinement on BTBT in TFETs has not been investigated.

All TFET configurations that have been proposed so far seem to share the gated reverse-biased  $p-i-n$  diode as a basic building block [7]. In a  $n$ -TFET the  $p^{++}$ -doped region, the intrinsic region and the  $n^{++}$ -doped region respectively play the role of source, channel and drain. We have chosen a device with a gate on top of the source region as shown in Fig. 10.1. When the gate voltage is gradually increased in this configuration, the source region underneath the gate is depleted until an inversion layer appears. At high gate bias, a tunneling current emerges, whereas the corresponding tunneling mechanism is referred to as line tunneling [8]. Under the gate, the potential profile can be assumed to vary only in the direction perpendicular to the dielectric-semiconductor interface which we take to be the  $z$  direction. The current is now proportional to the length of the gate  $L$  and the device width  $W$  [8]. An experimental indication of line tunneling can be found in Ref. 3.

The calculation of BTBT in semi-classical models is commonly based on the band diagram of the device. An electron tunneling from valence to conduction band is mimicked as a classical particle disappearing at the valence band edge and reappearing at the conduction band edge, as shown in Fig. 10.2. Correspondingly, the tunnel current equals the integral of an appropriate tunnel

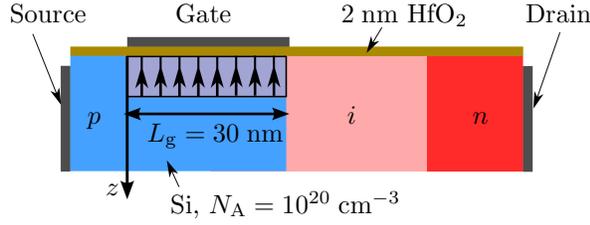


Figure 10.1: Illustration of a TFET with the gate over the source as investigated in this letter

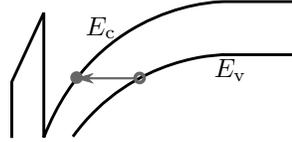


Figure 10.2: Semi-classical picture of band-to-band tunneling

generation rate  $G$  over the entire device,

$$I_{\text{semiclassical}} = q \int G dV \quad (10.1)$$

where  $G$  can be obtained from Kane's model [9,10]. Though looking attractive at first glimpse, the semi-classical model does not account for important quantum phenomena such as subband quantization due to the narrow potential well. Particularly, the discreteness of the subband energy levels will drastically affect the BTBT transition probabilities and, hence the tunneling current. Whereas in Kane's model the energy spectra of both the valence and the conduction band are considered continuous, we are now confronted with transitions from a valence band with a continuous spectrum to a conduction band consisting of discrete subbands.

Going beyond the semi-classical picture, we have calculated the electron and hole wavefunctions for a silicon TFET. Adopting the effective mass approximation for both carrier types, we have considered all six [100] oriented conduction band valleys as well as the two heavy and the light hole valleys of the valence band. The Schrödinger equation for holes (electrons) reads

$$\left( E_{0v(c)} \pm \frac{\hbar^2}{2} \left( \nabla \cdot \left[ m_{v(c)\alpha}^* \right]^{-1} \nabla \right) + U_{\text{ext}}(z) \right) \chi_{v(c)\alpha;\ell}(\mathbf{r}) = E_{v(c)\alpha;\ell} \chi_{v(c)\alpha;\ell}(\mathbf{r}). \quad (10.2)$$

$E_{0v(c)}$  denotes the bulk valence (conduction) band edge such that their difference equals the bandgap  $E_{0c} - E_{0v} = E_g$ , while  $\left[ m_{v(c)\alpha}^* \right]^{-1}$  denotes the inverse effective mass tensor for the valence (conduction) band and  $\alpha$  is a valley in-

dex.  $E_{v(c)\alpha;\ell}$  denotes the energy eigenvalue corresponding to the wavefunction  $\chi_{v(c)\alpha;\ell}(\mathbf{r})$  labeled by a quantum number  $\ell$ .

Defining the spectral functions as

$$A_{v(c)\alpha}(\mathbf{r}, \mathbf{r}'; E) = 2\pi \sum_{\ell} \chi_{v(c)\alpha;\ell}(\mathbf{r}) \delta(E - E_{v(c)\alpha;\ell}) \chi_{v(c)\alpha;\ell}^*(\mathbf{r}') \quad (10.3)$$

where, assuming translational symmetry in the  $x$ - and  $y$ -directions the diagonals of the spectral function depend only on  $z$ . The net charge density is given by

$$\rho_{\text{net}}(z) = -qN_a + 2q \int \frac{dE}{2\pi} \left( (1 - f_v(E)) \sum_{\alpha} A_{v\alpha}(z, z; E) - f_c(E) \sum_{\alpha} A_{c\alpha}(z, z; E) \right) \quad (10.4)$$

with  $-qN_a$  the charge due to the ionized acceptors.  $f_v(E)$  and  $f_c(E)$  denote the Fermi-Dirac functions for valence and conduction band determined by the source and drain bias respectively. The potential energy  $U_{\text{ext}}(z)$  reflecting all bias voltages can be determined by solving the 1D-Poisson equation self-consistently with the wavefunctions.

In Fig. 10.3, the distinct components of the net charge density are shown. Note that, due to confinement, the electrons are repelled from the oxide-semiconductor interface at  $z = 0$ . Moreover, only the valleys with the longitudinal mass in the  $z$  direction are significantly occupied as they are heavier and confinement does not affect them as strongly as the valleys having the lighter transverse mass in the  $z$  direction. Ignoring leakage currents for the sake of simplicity, we have imposed Dirichlet boundary conditions  $\chi(0) = 0$  at the oxide-semiconductor interface  $z = 0$ . The potential energy at the oxide-semiconductor interface is determined by the gate voltage applied at the gate contact:  $U_{\text{ext}}(0) + U'_{\text{ext}}(0)t_{\text{ox},s} = -q(V_{\text{gs}} - V_{\text{FB}})$  where  $V_{\text{FB}}$  is the flat band voltage,  $t_{\text{ox},s} = t_{\text{ox}}\epsilon_s/\epsilon_{\text{ox}}$  is the equivalent oxide thickness and  $U'_{\text{ext}}(z)$  denotes the derivative of  $U_{\text{ext}}(z)$ . The parameters used in our example are:  $E_g = 1.12416$  eV,  $m_{c,\text{long}}^* = 0.9163m_0$ ,  $m_{c,\text{trans}}^* = 0.1905m_0$ ,  $m_{v,\text{heavy}}^* = 0.49m_0$ ,  $m_{v,\text{light}}^* = 0.16m_0$ ,  $\epsilon_s = 11.5\epsilon_0$ ,  $\epsilon_{\text{ox}} = 15\epsilon_0$ ,  $t_{\text{ox}} = 2$  nm,  $N_a = 10^{20}$  cm $^{-3}$ ,  $V_{\text{FB}} = -2$  V,  $T = 300$  K,  $L = 30$  nm.

Taking the interaction with the phonons into account, we compute the phonon-assisted current [11] from

$$\begin{aligned} I = -\frac{2e}{\hbar} \int \frac{dE}{2\pi} & \left( (f_v(E)(1 - f_c(E - \hbar\omega_{\mathbf{k}_0}))(\nu(\hbar\omega_{\mathbf{k}_0}) + 1) \right. \\ & - f_c(E - \hbar\omega_{\mathbf{k}_0})(1 - f_v(E))\nu(\hbar\omega_{\mathbf{k}_0})) T_v^{\text{em}}(E) \\ & + (f_v(E)(1 - f_c(E + \hbar\omega_{\mathbf{k}_0}))\nu(\hbar\omega_{\mathbf{k}_0}) \\ & \left. - f_c(E + \hbar\omega_{\mathbf{k}_0})(1 - f_v(E))(\nu(\hbar\omega_{\mathbf{k}_0}) + 1)) T_v^{\text{abs}}(E) \right) \quad (10.5) \end{aligned}$$

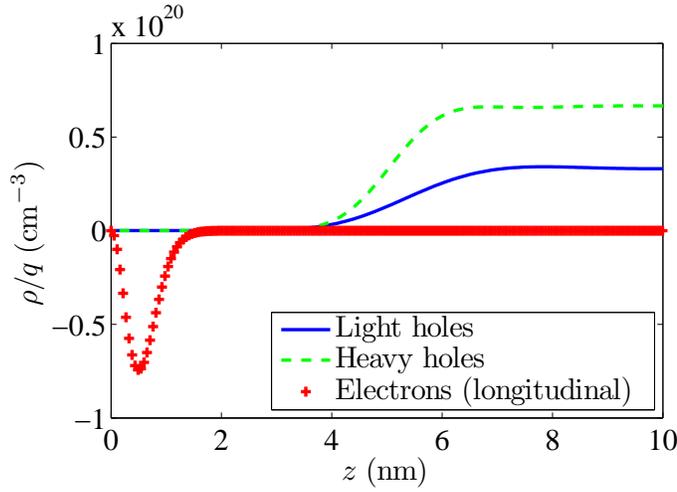


Figure 10.3: Different contributions to the net charge density: charge due to the light hole band (solid), the two heavy hole bands (dashed) and the two electron bands with longitudinal mass in the  $z$ -direction (+). The 4 electron bands with their transverse mass in the  $z$ -direction are not shown since they are not occupied.  $V_{\text{gs}} = 1.5$  V and  $V_{\text{ds}} = 0.2$  V.

with

$$T_{\nu}^{\text{abs,em}}(E) = \Omega |M'_{\mathbf{k}_0}|^2 \sum_{\alpha, \alpha'} \int d^3r A_{\nu\alpha}(\mathbf{r}, \mathbf{r}; E) A_{c\alpha'}(\mathbf{r}, \mathbf{r}; E \pm \hbar\omega_{\mathbf{k}_0}) \quad (10.6)$$

where  $\Omega |M'_{\mathbf{k}_0}|^2$  measures the electron-phonon interaction strength and  $\hbar\omega_{\mathbf{k}_0}$  is the energy of the phonon assisting a transition to a conduction band valley located at  $\mathbf{k}_0$ , while  $\nu(E)$  is the Bose-Einstein distribution function describing the occupation of the phonon modes. For the evaluation of the current, the parameters are [12]:  $\hbar\omega_{\mathbf{k}_0} = 18.4$  meV and  $\Omega |M'_{\mathbf{k}_0}|^2 = 4.86 \cdot 10^{-25}$  eV<sup>2</sup> cm<sup>3</sup>.

The quantum mechanical current as a function of gate voltage is compared with the corresponding semi-classical result obtained from Kane's model [9] in Fig. 10.4. Strikingly, the onset of the tunneling current occurs at a much larger voltage, compared to the semi-classical result, which can be attributed to the confinement of the electrons in the potential well near the oxide-semiconductor interface. A rough estimate of the difference in onset voltage can be obtained by considering a triangular potential well near  $z = 0$ . Taking  $F_{\text{ox}} = -U'_{\text{ext}}(0)$  to be the electric force acting on the electron at  $z = 0$ , the ground-state subband approximately lies at

$$E_{\text{sub},\alpha} = -a_0 \left( \frac{\hbar^2 F_{\text{ox}}^2}{2m_{c\alpha}^*} \right)^{\frac{1}{3}} \quad (10.7)$$

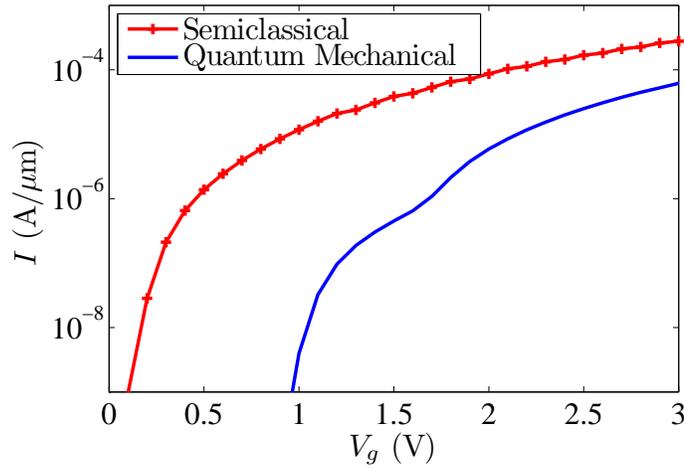


Figure 10.4: Line tunneling current as a function of applied gate bias for a silicon TFET: full quantum mechanical result (solid) and semiclassical Kane model (+). The inflection point of the solid curve around  $V_{gs} = 1.7$  V indicates when transitions to electron states with the lower effective mass in the  $z$ -direction become predominant.

where  $a_0 \approx -2.3381$  is the first zero of the Airy function.

Taking the potential energy to be parabolic in the depletion region and linear in the gate oxide region, the force at the surface is

$$F_{\text{ox}} = \sqrt{\frac{2q^2 N_a}{\epsilon_s} (U_{\text{ext}}(\infty) - U_{\text{ext}}(0))}. \quad (10.8)$$

Taking the confinement into account, the onset of tunneling corresponds to  $E_g + E_{\text{sub},\alpha} = U_{\text{ext}}(\infty) - U_{\text{ext}}(0)$  or, equivalently, to the solution of a cubic equation in  $F_{\text{ox}}^{2/3}$ . The latter yields  $F_{\text{ox}} = 8.623$  MeV/cm when the electron transverse light mass is in the  $z$  direction  $F_{\text{ox}} = 7.513$  MeV/cm otherwise. This corresponds respectively to  $V_{\text{onset,long}} = 0.9456$  V and  $V_{\text{onset,tran}} = 1.6852$  V instead of the semi-classical onset voltage  $V_{\text{onset,sc}} = 0.0362$  V, which is in accordance with the shift seen in Fig. 10.4. Taking conduction band non-parabolicity into account will slightly change the quantitative values of the shift but will not affect the general picture. A similar conclusion can be drawn when the Dirichlet boundary condition (making all wave functions vanish at  $z=0$ ) is relaxed. An estimate of the impact of the effect can be made based on the electric fields calculated before and assuming a penetration of 1-2 Å which results in a reduction of the subband energy by about 80-160 meV.

Fig. 10.5 shows the current as a function of applied gate voltage for different drain voltages. Reduction of the latter induces a build-up of inversion

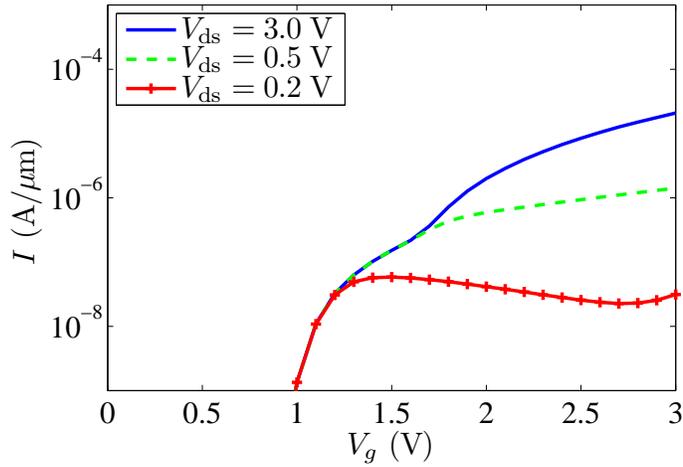


Figure 10.5: Line tunneling current as a function of applied gate bias in a silicon TFET for different drain voltages.

charge further limiting the increase of the electric field in the semiconductor. Apart from the reduced current, the shifted onset and the negative differential transconductance for the lowest drain bias, the agreement between the fully quantum mechanical and the semi-classical calculations is only qualitative.

In conclusion, quantum confinement critically determines the onset of line tunneling in TFETs and predicts a significant reduction of tunneling current as compared to semi-classical models, which erroneously predict non-zero current before the first subband is established. Appropriate compact models should therefore incorporate the subband energy. If the conduction band valleys are hosting electrons with different effective masses, the heavy mass electrons tunnel at lower gate bias compared to those of higher masses. However, the heavy masses decrease the electron tunneling probability which leads to a deterioration of the subthreshold swing.

The authors acknowledge Anne Verhulst for useful discussions. William Vandenberghe gratefully acknowledges the support of a Ph.D. stipend from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen).

## Bibliography

- [1] S. Banerjee, W. Richardson, J. Coleman, and A. Chatterjee, “A new three-terminal tunnel device,” *Electron Device Letters, IEEE*, vol. 8, no. 8, pp.

347 – 349, Aug. 1987.

- [2] W. M. Reddick and G. A. J. Amaratunga, “Silicon surface tunnel transistor,” *Applied Physics Letters*, vol. 67, no. 4, pp. 494–496, July 1995.
- [3] D. Kazazis, P. Jannaty, A. Zaslavsky, C. L. Royer, C. Tabone, L. Clavelier, and S. Cristoloveanu, “Tunneling field-effect transistor with epitaxial junction in thin germanium-on-insulator,” *Applied Physics Letters*, vol. 94, no. 26, p. 263508, 2009. [Online]. Available: <http://link.aip.org/link/?APL/94/263508/1>
- [4] K. K. Bhuiwarka, J. Schulze, and I. Eisele, “A simulation approach to optimize the electrical parameters of a vertical tunnel FET,” *IEEE Transactions on Electron Devices*, vol. 52, no. 7, July 2005.
- [5] A. S. Verhulst, W. G. Vandenberghe, K. Maex, and G. Groeseneken, “A tunnel field-effect transistor without gate-drain overlap,” *Applied Physics Letters*, vol. 91, no. 053102, July 2007.
- [6] K. Boucart and A. Ionescu, “Double-gate tunnel fet with high gate dielectric,” *Electron Devices, IEEE Transactions on*, vol. 54, no. 7, pp. 1725–1733, 2007.
- [7] A. S. Verhulst, B. Sorée, D. Leonelli, W. G. Vandenberghe, and G. Groeseneken, “Modeling the single-gate, double-gate, and gate-all-around tunnel field-effect transistor,” *Journal of Applied Physics*, vol. 107, no. 2, p. 024518, 2010. [Online]. Available: <http://link.aip.org/link/?JAP/107/024518/1>
- [8] W. G. Vandenberghe, A. S. Verhulst, G. Groeseneken, B. Sorée, and W. Magnus, “Analytical model for a tunnel field-effect transistor,” in *IEEE Melecon*, 2008.
- [9] E. O. Kane, “Theory of tunneling,” *Journal of Applied Physics*, vol. 32, no. 1, pp. 83–91, 1961.
- [10] W. Vandenberghe, B. Sorée, W. Magnus, and G. Groeseneken, “Zener tunneling in semiconductors under nonuniform electric fields,” *Journal of Applied Physics*, vol. 107, no. 5, p. 054520, 2010. [Online]. Available: <http://link.aip.org/link/?JAP/107/054520/1>
- [11] W. Vandenberghe, B. Sorée, W. Magnus, and M. V. Fischetti, , 2011, unpublished.
- [12] C. Rivas, R. Lake, G. Klimeck, W. R. Frensley, M. V. Fischetti, P. E. Thompson, S. L. Rommel, and P. R. Berger, “Full-band simulation of indirect phonon assisted tunneling in a silicon tunnel diode with delta-doped contacts,” *Applied Physics Letters*, vol. 78, no. 6, pp. 814–816, 2001. [Online]. Available: <http://dx.doi.org/10.1063/1.1343500>



## Chapter 11

# Modified semiclassical model accounting for quantum confinement

Field Induced Quantum Confinement  
in Indirect Semiconductors: Quantum Mechanical  
and Modified Semiclassical Model

William G. Vandenberghe, Bart. Sorée, Wim Magnus,  
Guido Groeseneken, Anne S. Verhulst and Massimo V. Fischetti

Published in *Simulation of Semiconductor Processes and Devices  
(SISPAD)*, 2011, pp. 271 - 174 , September 2011.

©2011 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

### Contributions of first author

- Conception and development of the theory
- Development of the code
- Creating figures
- Text writing and editing

## Abstract

Going beyond the existing semiclassical approach to calculate band-to-band tunneling (BTBT) current we have developed a quantum mechanical model incorporating confinement effects and multiple electron and hole valleys to calculate the tunnel current in a tunnel field-effect transistor. Comparison with existing semiclassical models reveals a big shift in the onset of tunneling due to energy quantization. We show that the big shift due to quantum confinement is slightly reduced by taking penetration into the gate dielectric into account. We further propose a modified semiclassical model capable of accounting for quantum confinement.

## 11.1 Introduction

On the one hand, nanosized MOSFETs suffer from gate-induced drain leakage (GIDL) which deteriorate their off-current [1]. On the other hand, tunnel field-effect transistors (TFETs) are considered potential candidates to overcome the limit of the 60 mV/decade subthreshold swing in a MOSFET operating at room temperature [2]. In both instances the Band-to-Band tunneling (BTBT) process is responsible for the tunneling current as illustrated in Fig. 11.1. An accurate theory of BTBT is therefore highly desirable.

The calculation of BTBT current is traditionally based on semiclassical models using band diagrams inside the device [3,4]. An electron tunneling from valence to conduction band is mimicked as a classical particle disappearing at the valence band edge and reappearing at the conduction band edge, as shown in Fig. 11.2. Correspondingly, the tunnel current equals the integral of a tunnel generation rate  $G$  over the entire device,

$$I_{\text{semiclassical}} = q \int G dV \quad (11.1)$$

where  $G$  can be obtained from Kane's model for direct [5] or indirect semiconductors [6–8].

To compare semi-classical results with quantum mechanical results, we study a TFET device with a large gate-source overlap as shown in Fig. 11.3. The potential can be taken to vary only in the direction perpendicular to the semiconductor-dielectric interface ( $z$  direction) [9] facilitating modeling and interpretation. In this paper, we treat tunneling in an indirect semiconductor and make a quantum mechanical calculation of the so-called line tunneling current component, which is proportional to the gate-source overlap. In section

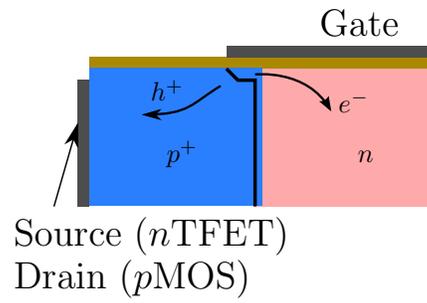


Figure 11.1: Picture of GIDL/TFET working principle showing hole and electron generation due to BTBT

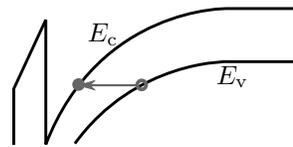


Figure 11.2: Semiclassical picture of band-to-band tunneling

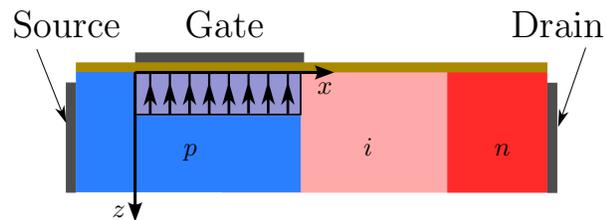


Figure 11.3: Illustration of a TFET with the gate over the source

11.2, we outline the quantum mechanical framework to calculate current similar to [8] but now also including penetration into the gate dielectric. In section 11.3, we propose a modification of existing semiclassical models to account for quantum confinement and discuss some further limitations of the semiclassical models.

## 11.2 Quantum mechanical modeling of BTBT

To obtain a quantum mechanical estimate of the BTBT in indirect semiconductors, we use the method outlined in [10]. First, the conduction and valence band electron wavefunctions have to be determined. Adopting the effective mass approximation for both carrier types, the Schrödinger equations read:

$$\left( E_{0v(c)} \pm \frac{\hbar^2}{2} \left( \nabla \cdot \left[ m_{v(c)\alpha}^* \right]^{-1} \nabla \right) + U_{\text{ext}}(z) \right) \chi_{v(c)\alpha;\ell}(\mathbf{r}) = E_{v(c)\alpha;\ell} \chi_{v(c)\alpha;\ell}(\mathbf{r}). \quad (11.2)$$

We consider all six [100] oriented conduction band valleys as well as the three valence band valleys. Solving the Schrödinger equation results in a complete set of wavefunctions  $\chi_{v(c)\alpha;\ell}(\mathbf{r})$  and the corresponding energy eigenvalues  $E_{v(c)\alpha;\ell}$  where  $\ell$  denotes the set of quantum numbers and  $\alpha$  is the valley index.

Defining the spectral functions as

$$A_{v(c)\alpha}(\mathbf{r}, \mathbf{r}'; E) = 2\pi \sum_{\ell} \chi_{v(c)\alpha;\ell}(\mathbf{r}) \delta(E - E_{v(c)\alpha;\ell}) \chi_{v(c)\alpha;\ell}^*(\mathbf{r}'), \quad (11.3)$$

the different contributions to the charge density are determined by weighing the spectral functions with the Fermi-Dirac distribution functions  $f_{v(c)}(E) = 1/(1 + \exp((E - \mu_{v(c)})/(kT)))$ :

$$\rho_{\text{net}}(z) = -qN_a + 2q \int \frac{dE}{2\pi} \left( (1 - f_v(E)) \sum_{\alpha} A_{v\alpha}(z, z; E) - f_c(E) \sum_{\alpha} A_{c\alpha}(z, z; E) \right) \quad (11.4)$$

with  $N_a$  the doping concentration. The potential energy  $U_{\text{ext}}(z)$  reflecting all bias voltages can be determined by solving the 1D-Poisson equation self-consistently with the wavefunctions.

Taking the interaction with the phonons into account, we compute the phonon-assisted current from

$$I = -\frac{2e}{\hbar} \int \frac{dE}{2\pi} \left( (f_v(E)(1 - f_c(E - \hbar\omega_{\mathbf{k}_0}))(\nu(\hbar\omega_{\mathbf{k}_0}) + 1) \right)$$

$$\begin{aligned}
& - f_c(E - \hbar\omega_{\mathbf{k}_0})(1 - f_v(E))\nu(\hbar\omega_{\mathbf{k}_0})T_v^{\text{em}}(E) \\
& + (f_v(E)(1 - f_c(E + \hbar\omega_{\mathbf{k}_0}))\nu(\hbar\omega_{\mathbf{k}_0}) \\
& - f_c(E + \hbar\omega_{\mathbf{k}_0})(1 - f_v(E))(\nu(\hbar\omega_{\mathbf{k}_0}) + 1))T_v^{\text{abs}}(E) \Big) \quad (11.5)
\end{aligned}$$

with

$$T_v^{\text{abs,em}}(E) = \Omega |M'_{\mathbf{k}_0}|^2 \times \sum_{\alpha, \alpha'} \int d^3r A_{v\alpha}(\mathbf{r}, \mathbf{r}; E) A_{c\alpha'}(\mathbf{r}, \mathbf{r}; E \pm \hbar\omega_{\mathbf{k}_0}). \quad (11.6)$$

At the gate dielectric ( $z = 0$ ), a boundary condition for the wavefunctions is required. A first approximation is to use Dirichlet boundary conditions:  $\chi(0) = 0$  [8]. This corresponds to modeling the gate dielectric as a hard wall potential that strictly confines the electrons to the device region.

However, real devices are found to suffer from wavefunction penetration into the dielectric. In general the penetration leads to an unwanted gate leakage current but for the field induced quantum confinement, the penetration of the gate dielectric will make the impact of confinement slightly less pronounced. To account for penetration into the dielectric, we assume the dielectric is infinitely thick and that the wavefunction decays with a given decay length  $l_{\text{dec}}$  inside the dielectric, i.e.  $\chi(z) \propto \exp(z/l_{\text{dec}})$  for  $z < 0$ . The boundary condition for the wavefunction is now:

$$\left. \frac{d\chi(z)}{dz} \right|_{z=0} = \frac{\chi(0)}{l_{\text{dec}}}. \quad (11.7)$$

The value of  $l_{\text{dec}}$  can be determined from the complex band structure which is about 3 Å for  $\text{HfO}_2$  [11].

In Fig. 11.4, the quantum mechanical current is compared with the semiclassical current for two different doping concentrations. The big onset shift between the semiclassical and the quantum mechanical calculation is due to the quantum confinement of the electrons near the interface. The shift is bigger for the device with the larger doping concentration as larger fields and stronger carrier confinement is present. The two different effective masses (transversal and longitudinal) in the  $z$  direction give rise to different energy levels, the signatures of which can be clearly observed as a cusp in the current-voltage characteristic for the TFET with high doping concentration. The shift due to confinement can be seen to be smaller when the penetration into the gate dielectric is taken into account.

### 11.3 Semiclassical models

In this section we propose a modification to the semiclassical model to account for quantum confinement.

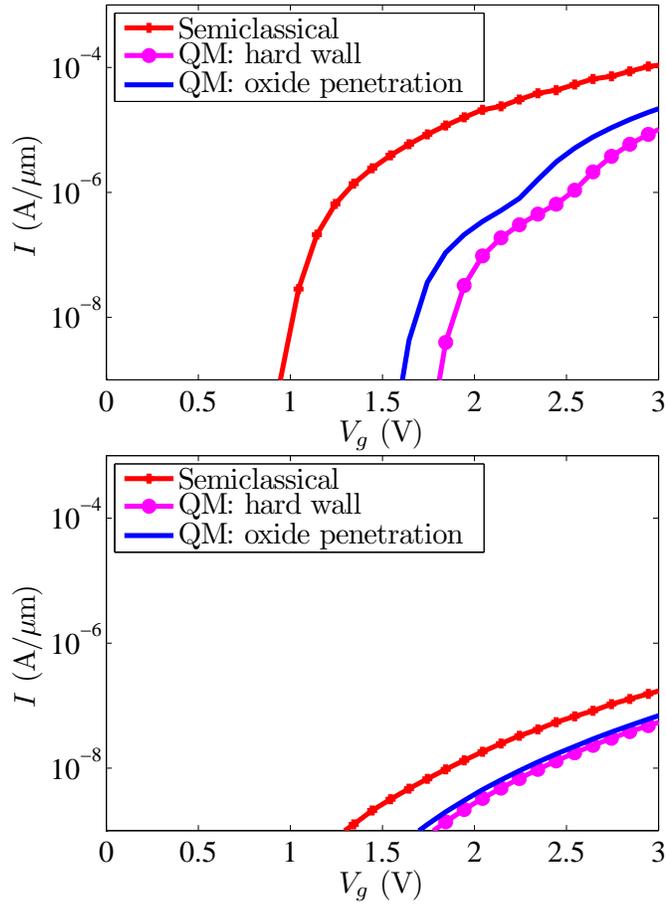


Figure 11.4: Comparison of full quantum mechanical current with semiclassical current for source doping  $N_a = 10^{20} \text{ cm}^{-3}$  (top) and  $N_a = 10^{19} \text{ cm}^{-3}$  (bottom) revealing the big shift due to quantum confinement for the hard wall condition and taking penetration into the gate dielectric into account. Lower doping results in a reduced effect of quantum confinement. Additional parameters used for the calculation are given in the Appendix.

### 11.3.1 The existing semiclassical model

The semiclassical model we have used for our comparison in Fig. 10.4 is similar to that used in Ref. [10] and defines two tunnel paths starting at the valence band and ending at the conduction band: one corresponding to a path bridging a gap  $E_g + \hbar\omega_{\mathbf{k}_0}$  and the other bridging a gap  $E_g - \hbar\omega_{\mathbf{k}_0}$ . The first path corresponds either to an electron going from valence to conduction band emitting a phonon or an electron going in the other direction absorbing a phonon. The second path describes the dual processes.

Based on the tunnel path start point ( $z = z_1$ ) and end point ( $z = z_2$ ), illustrated in Fig. 11.5, the path length ( $l_{\text{tun}} = z_2 - z_1$ ) is determined and a generation rate is calculated according to the average force  $\bar{F} = (U_{\text{ext}}(z_2) - U_{\text{ext}}(z_1))/l_{\text{tun}}$  along the tunnel path. To have a fair comparison between the quantum mechanical and the semiclassical model we have taken the same parameters for the electron-phonon interaction strength, effective masses and bandgap and substituted them in the indirect Kane model [7, 10]. Only the tunneling of the light holes to the electrons with their transversal mass in the tunneling direction has been taken into account as this process dominates over the other tunneling processes.

### 11.3.2 Modified semiclassical model

As observed in Fig. 10.4, the major discrepancy between the semiclassical and the quantum mechanical input characteristics is the big shift in onset voltage due to the absence of states to tunnel into due to quantum confinement. Based on this observation we propose a modified semiclassical model which proceeds as follows: 1) start a tunnel path at an initial position  $z_1$  in the valence band, 2) determine the location of the intersection of the tunnel path with the conduction band  $z_2$ , 3) determine the depth of the conduction band well at the intersection with the dielectric interface  $\Delta E_c$ , 4) determine a minimum well depth  $\Delta E_{\text{min},c}$  required to accommodate a particle, 5) if  $\Delta E_c > \Delta E_{\text{min},c}$  calculate generation rate in normal fashion, otherwise reject the tunnel path. An illustration of an accepted and a rejected path is given in Fig. 11.5.

To estimate the minimal well depth required to accommodate a particle, calculate the energy level of the first state in the triangular well approximation

$$E_{0c} = -a_0 \left( \frac{\hbar^2 (F(z_2))^2}{2m_{c,z}^*} \right)^{1/3} \quad (11.8)$$

with  $a_0 \approx -2.3381$  the first zero of the Airy function,  $m_{c,z}^*$  is the effective mass in the  $z$ -direction and where

$$F(x_2) = -\frac{d}{dz} U_{\text{ext}}(z) \Big|_{z=z_2} \quad (11.9)$$

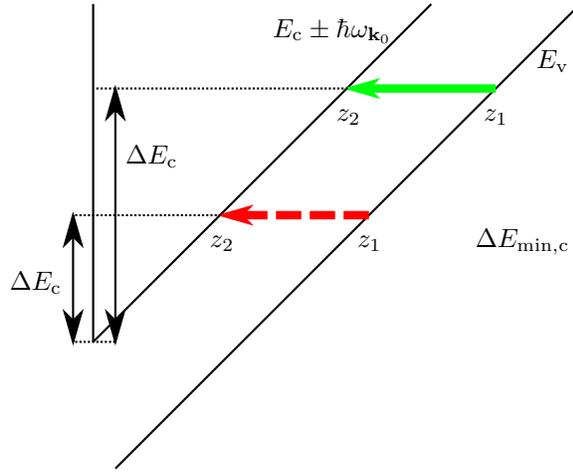


Figure 11.5: Illustration of how a semiclassical path starting at  $z_1$  going to  $z_2$  is defined. For the modified semiclassical model the well depth  $\Delta E_c$  at the interface has to be determined. Tunneling either proceeds in the normal fashion if  $\Delta E_c > \Delta E_{\min,c}$  (solid path) or the tunnel path is rejected  $\Delta E_c \leq \Delta E_{\min,c}$  (dashed path).

is the force on the electron at  $z = z_2$ . In the modified semiclassical picture, the penetration into the gate dielectric can be taken into account by reducing  $E_0$  by the energy the electron is expected to gain over the distance of the decay length:

$$\Delta E_{\min,c} = E_{0c} - F(x_2)l_{\text{dec}}. \quad (11.10)$$

We have implemented the modified semiclassical model outlined above and compared it with the previously calculated current-voltage characteristics and show the result in Fig. 11.6. One more change that was made in the modified semiclassical model shown in Fig. 11.6 is that tunneling of both the transversal and the longitudinal electrons was taken into account. The tunneling probability is much lower for the longitudinal electrons but due to reduced impact of quantum confinement, their contribution is now also important. The modified semiclassical model manages to capture the effect of the quantum confinement and gives reasonably good agreement with the quantum mechanical result.

In our case we have studied a  $n$ TFET where the conduction band electrons are confined. In case confinement is present for the valence band electrons, it must be verified if the well for the valence band electrons is deep enough.

### 11.3.3 Further limitations to the semiclassical model

We have shown that semiclassical models can be adapted to incorporate field induced quantum confinement. Nevertheless, we believe it is important to remind the reader that there are still some other unsolved issues with the use of semiclassical models we have not discussed in this paper.

First of all, in a two-dimensional potential profile, tunnel paths are chosen in the device according to straight lines [12], along the electric field lines [13] or along Newtonian trajectories [14]. For a potential with pronounced two-dimensional features, none of these methods can be justified on theoretical grounds and certainly not for the case of phonon-assisted tunneling where the interaction with the phonon has to be accounted for.

Secondly, the electron/hole generation is assumed to take place at the edge of the conduction and valence band when introduced in the drift-diffusion equations. Again no theoretical framework on how the generation should be incorporated is available.

## 11.4 Conclusion

In a semiconductor device with high electric fields near the gate-dielectric, taking quantum confinement effects into account when calculating BTBT is paramount and existing semiclassical models fail. Penetration of the wavefunction into the gate dielectric slightly reduces the quantum confinement effect compared to semiclassical results. We have also demonstrated a modified semiclassical model which amounts to a small correction of the existing models and which is capable of capturing the effect of field-induced quantum confinement.

## Acknowledgments

William Vandenberghe gratefully acknowledges the support of a Ph.D. stipend from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). This work was supported by imec's Industrial Affiliation Program.

## Appendix

Parameters used:

Bandgap: 1.12 eV

Effective masses:  $m_{c,l} = 0.9163 m_0$ ,  $m_{c,t} = 0.1905 m_0$ ,  $m_{v,lh} = 0.16 m_0$  and  $m_{v,hh} = 0.490 m_0$

Dielectric constants:  $\epsilon_s = 11.5\epsilon_0$  and  $\epsilon_{ox} = 15\epsilon_0$

Electron-phonon parameters: 18.4 meV,  $\Omega|M'_{\mathbf{k}_0}|^2 = 4.8634 \times 10^{-31} \text{ eV}^2 \text{ m}^3$

## Bibliography

- [1] T. Chan, J. Chen, P. Ko, and C. Hu, "The impact of gate-induced drain leakage current on MOSFET scaling," *Electron Devices Meeting, 1987 International*, vol. 33, pp. 718–721, 1987.
- [2] S. Banerjee, W. Richardson, J. Coleman, and A. Chatterjee, "A new three-terminal tunnel device," *Electron Device Letters, IEEE*, vol. 8, no. 8, pp. 347 – 349, Aug. 1987.
- [3] K. K. Bhuwalka, J. Schulze, and I. Eisele, "A simulation approach to optimize the electrical parameters of a vertical tunnel FET," *IEEE Transactions on Electron Devices*, vol. 52, no. 7, July 2005.
- [4] A. S. Verhulst, W. G. Vandenberghe, K. Maex, and G. Groeseneken, "A tunnel field-effect transistor without gate-drain overlap," *Applied Physics Letters*, vol. 91, no. 053102, July 2007.
- [5] E. O. Kane, "Zener tunneling in semiconductors," *Journal of Physics and Chemistry of Solids*, vol. 12, pp. 181–188, 1959.
- [6] L. Keldysh, "Influence of the lattice vibrations of a crystal on the production of electron-hole pairs in a strong electric field," *Sov. Phys. JETP*, vol. 7, p. 665, 1959.
- [7] E. O. Kane, "Theory of tunneling," *Journal of Applied Physics*, vol. 32, no. 1, pp. 83–91, 1961.
- [8] W. Vandenberghe, B. Sorée, W. Magnus, and M. V. Fischetti, "Generalized phonon-assisted zener tunneling in indirect semiconductors with non-uniform electric fields: A rigorous approach," *Journal of Applied Physics*, vol. 109, no. 12, p. 124503, 2011. [Online]. Available: <http://link.aip.org/link/?JAP/109/124503/1>
- [9] W. G. Vandenberghe, A. S. Verhulst, G. Groeseneken, B. Sorée, and W. Magnus, "Analytical model for a tunnel field-effect transistor," in *IEEE Melecon*, 2008.
- [10] W. G. Vandenberghe, B. Sorée, W. Magnus, G. Groeseneken, and M. V. Fischetti, "Impact of field-induced quantum confinement in tunneling

- field-effect devices,” *Applied Physics Letters*, vol. 98, no. 14, p. 143503, 2011. [Online]. Available: <http://link.aip.org/link/?APL/98/143503/1>
- [11] F. Sacconi, J. Jancu, M. Povolotskyi, and A. Di Carlo, “Full-band tunneling in high- $\kappa$  oxide mos structures,” *Electron Devices, IEEE Transactions on*, vol. 54, no. 12, pp. 3168–3176, dec. 2007.
- [12] *Sentaurus Device*, Synopsis, March 2010.
- [13] W. G. Vandenberghe, A. S. Verhulst, G. Groeseneken, B. Sorée, and W. Magnus, “Analytical model for point and line in a tunnel field-effect transistor,” in *Simulation of Semiconductor Processes and Devices, 2008 International Conference on*, Sept. 2008.
- [14] C. Shen, L.-T. Yang, G. Samudra, and Y.-C. Yeo, “A new robust non-local algorithm for band-to-band tunneling simulation and its application to tunnel-fet,” *Solid-State Electronics*, vol. 57, no. 1, pp. 23–30, 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038110110003539>

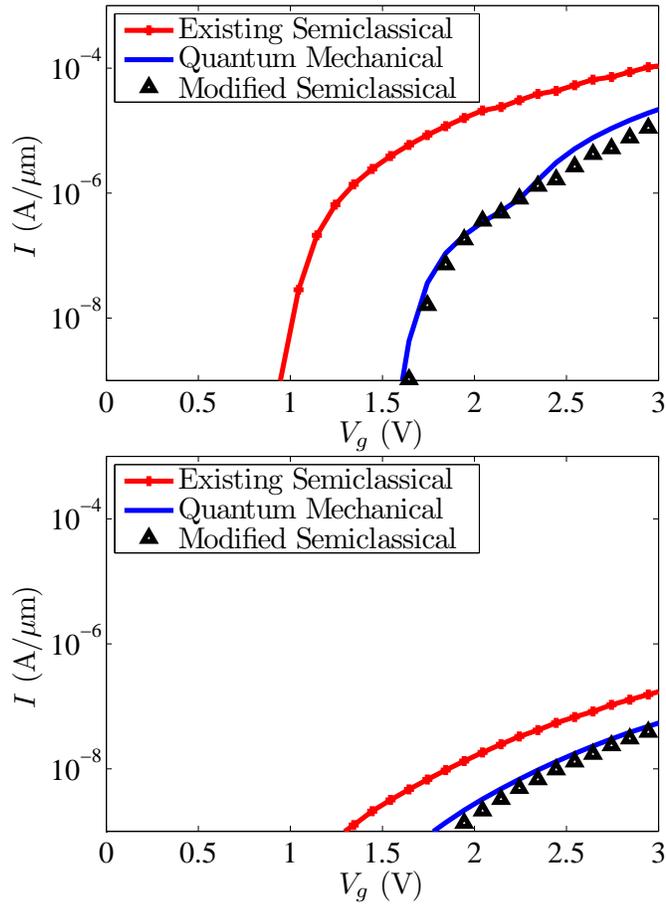


Figure 11.6: Full quantum mechanical current and the existing semiclassical current for source doping  $N_a = 10^{20}$   $\text{at}/\text{cm}^{-3}$  (top) and  $N_a = 10^{20}$   $\text{at}/\text{cm}^{-3}$  (bottom) as given in Fig. 10.4 compared with the modified semiclassical model. Wavefunction penetration is taken into account.



## Chapter 12

# Two-dimensional modeling of BTBT

Two-Dimensional Quantum Mechanical  
Modeling of Band-to-Band Tunneling in Indirect  
Semiconductors

William G. Vandenberghe, Bart Sorée, Wim Magnus, Massimo V.  
Fischetti, Anne S. Verhulst and Guido Groeseneken

Published in *IEEE International Electron Devices Meeting  
(IEDM) 2011*, pp. 5.3.1-5.3.4, December 2011.

©2011 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

### Contributions of first author

- Conception and development of the theory
- Development of the code
- Creating figures
- Text writing and editing

## Abstract

A quantum mechanical procedure to calculate phonon-assisted tunneling current in indirect semiconductors in a two-dimensional structure is demonstrated. Applying the procedure to two types of double-gate silicon tunnel field-effect transistor (TFET) structures, it is observed that semiclassical predictions strongly overestimate the subthreshold swing and the device current. Furthermore, while the semiclassical simulation suggests a higher current for one of the investigated TFET devices, a proper quantum mechanical calculation comes to the opposite conclusion. This result, which is expected to apply to direct semiconductors as well, proves the importance of correct quantum mechanical tunneling models for performance predictions of novel devices such as TFETs.

## 12.1 Introduction

As scaling continues, device dimensions are reduced while supply voltages are not scaled accordingly, and thus electric fields are increasing. As a consequence band-to-band tunneling (BTBT) is gaining importance. For example BTBT is responsible for gate induced drain leakage in MOSFETs [1]. Moreover, attempts to overcome the 60 mV/decade subthreshold swing in MOSFETs have spawned research of alternative devices such as the tunnel field-effect transistor (TFET), whose operation is based on BTBT [2]. Semiclassical models are often employed to study tunneling in MOSFETs or TFETs by defining one-dimensional tunnel paths inside the structure [3]. This approach, however, lacks theoretical justification and until now no proper treatment of the problem was available for indirect semiconductors [4].

In this paper, we describe a new procedure to calculate phonon-assisted tunneling current in indirect semiconductors in a two-dimensional structure and apply it to two types of double-gate silicon TFETs.

## 12.2 Formalism

The formalism we employ and its physical models are outlined in [5], where it is assumed that there is no tunneling in the absence of electron-phonon interaction. A flowchart of the algorithm used in this paper is shown in Fig. 12.1. First, the Schrödinger equation has to be solved for all relevant conduction and valence band valleys ( $\alpha$ ). We have treated the valence and conduction bands

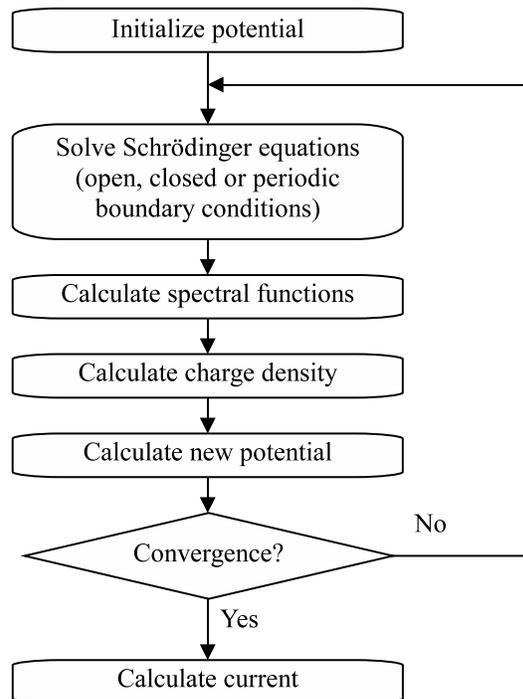


Figure 12.1: Flow chart of procedure to calculate current

in the effective mass approximation, so that the Schrödinger equation reads:

$$\left( E_{0v(c)} \pm \frac{\hbar^2}{2} \left( \nabla \cdot [m_{v(c)\alpha}^*]^{-1} \nabla \right) + U_{\text{ext}}(\mathbf{R}) \right) \chi_{v(c)\alpha;\ell}(\mathbf{r}) = E_{v(c)\alpha;\ell} \chi_{v(c)\alpha;\ell}(\mathbf{r}), \quad (12.1)$$

where  $\ell$  runs over the complete set of quantum numbers and

$$\mathbf{R} = (x, z) \quad , \quad \nabla_{\mathbf{R}} = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial z} \right).$$

In the contacts, contrary to a Non-Equilibrium Green's Formalism approach, transmitting boundary conditions are not required. Instead, Dirichlet, Neumann, or periodic boundary conditions can be used. As a consequence, the wavefunctions are solutions of a Hermitian eigenvalue problem, no explicit energy discretization is required, and, most importantly, localized states can be taken into account in a straightforward way. The occupation of valence and conduction band states is determined by two separate Fermi levels corresponding to - initially - uncoupled valence and conduction band electron gases. This corresponds to using a Fermi-Dirac distribution for valence and conduction band electrons  $f_{v,c}(E)$  and the difference between the corresponding Fermi levels is determined by the applied bias. The charge density

$$\begin{aligned} \rho_{\text{net}}(\mathbf{R}) &= -qN_{\text{a}}(\mathbf{R}) + qN_{\text{d}}(\mathbf{R}) \\ &+ 2q \int \frac{dE}{2\pi} \left( (1 - f_v(E)) \sum_{\alpha} A_{v\alpha}(\mathbf{r}, \mathbf{r}; E) - f_c(E) \sum_{\alpha} A_{c\alpha}(\mathbf{r}, \mathbf{r}; E) \right) \end{aligned} \quad (12.2)$$

is determined from the spectral functions

$$A_{v(c)\alpha}(\mathbf{r}, \mathbf{r}'; E) = 2\pi \sum_{\ell} \chi_{v(c)\alpha;\ell}(\mathbf{r}) \delta(E - E_{v(c)\alpha;\ell}) \chi_{v(c)\alpha;\ell}^*(\mathbf{r}'). \quad (12.3)$$

Solving the Poisson equation

$$(\nabla_{\mathbf{R}} \cdot \epsilon(\mathbf{R}) \nabla_{\mathbf{R}}) U_{\text{ext}}(\mathbf{R}) = -q\rho_{\text{net}}(\mathbf{R}) \quad (12.4)$$

self-consistently, the current mediated by the electron-phonon interaction can be calculated using (5-7),

$$\begin{aligned} T_{\text{v}}^{\text{abs,em}}(\mathbf{R}; E) &= \Omega |M_0|^2 \sum_{\alpha, \alpha'} A_{v\alpha}(\mathbf{r}, \mathbf{r}; E) A_{c\alpha'}(\mathbf{r}, \mathbf{r}; E \pm \hbar\omega_0) \\ &(\Omega = \text{total volume}), \end{aligned} \quad (12.5)$$

$$G(\mathbf{R}) = -\frac{2}{\hbar} \int \frac{dE}{2\pi} \left( (f_v(E)(1 - f_c(E - \hbar\omega_0))(\nu(\hbar\omega_0) + 1) \right.$$

$$\begin{aligned}
& - f_c(E - \hbar\omega_0)(1 - f_v(E))\nu(\hbar\omega_0)T_v^{\text{em}}(\mathbf{R}; E) \\
& \quad + (f_v(E)(1 - f_c(E + \hbar\omega_0))\nu(\hbar\omega_0) \\
& \quad - f_c(E + \hbar\omega_0)(1 - f_v(E))(\nu(\hbar\omega_0) + 1))T_v^{\text{abs}}(\mathbf{R}; E), \quad (12.6)
\end{aligned}$$

$$I_{\text{ds}} = qW \int d^2R G(\mathbf{R}). \quad (12.7)$$

$T_v^{\text{em,abs}}(\mathbf{R}; E)$  describes the valence band tunnel probability under phonon emission or absorption. The generation rate  $G(\mathbf{R})$  consists of four terms with a clear physical interpretation, as an example the first term describes electrons going from valence to conduction band emitting a phonon.

In the existing semiclassical approach an equation similar to (7) is used to determine the current, but the generation rate is determined from the semiclassical tunnel paths running from valence to conduction band [3].

## 12.3 BTBT current in a double gate silicon TFET

Fig. 12.2 shows the structure of a double gate tunnel field-effect transistor. Two devices are studied with different device parameters given in Table I. Electron-phonon interaction strengths ( $M_0 = M_{\mathbf{k}_0}$ ) and phonon energies ( $\hbar\omega_0 = \hbar\omega_{\mathbf{k}_0}$ ) are taken from [6]. Dirichlet boundary conditions are used at the gate dielectric semiconductor interface confining the carriers to the semiconductor. At the contact electrodes, an arbitrary choice for Neumann boundary conditions is made.

### 12.3.1 The first device

The first device under study is a TFET with a 10 nm body thickness and high source doping of  $10^{20}\text{cm}^{-3}$ . The self-consistent potential energy and net carrier density at 2 V gate bias are shown in Figs. 12.3-12.4. As expected, an inversion layer repelled by the confining gate dielectric can be seen under the gate. In the semiclassical picture, the shortest tunnel paths and the highest generation rates are expected near the gate dielectric. In the quantum mechanical picture, however, at low bias (Fig. 12.5) the generation rate is centered in the device due to quantum confinement. With increasing gate bias the electric fields near the semiconductor-dielectric interface increase. As a consequence, the localized contribution to the tunnel current for high gate bias (Fig. 12.6) can be seen to move towards the interface with increasing gate bias.

The calculated current is shown in Fig. 12.7. The quantum mechanically determined TFET average subthreshold swing and saturation current are worse

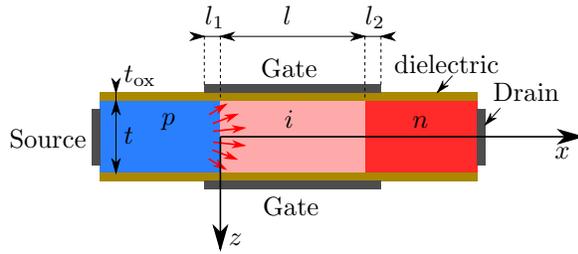


Figure 12.2: Illustration of double gate nTFET where the red arrows symbolize the tunneling

Device #	1	2
Doping (at/cm <sup>-3</sup> )	10 <sup>20</sup>	5 × 10 <sup>19</sup>
l <sub>1</sub> /l/l <sub>2</sub>	5/30/5	20/30/5
t/t <sub>ox</sub>	10/3	20/3
ε <sub>ox,r</sub> /ε <sub>s,r</sub>	11.9/11.9	
Gate workfunction (eV)	4.05	

Table 12.1: Summary of values used for the two devices under study

when compared to the semiclassical model. As was already observed in [7], quantum confinement becomes more important with increasing bias and this explains the increasing discrepancy between semiclassical and quantum mechanical current with increasing gate bias.

### 12.3.2 The second device

The second device has a lower source/drain doping concentration of 5 × 10<sup>19</sup> cm<sup>-3</sup>, a thicker 20 nm body thickness and an increased gate-source overlap of 20 nm

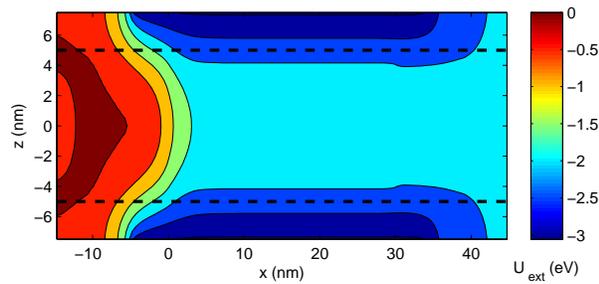


Figure 12.3: Potential energy profile in device and gate dielectric region for device 1 at V<sub>gs</sub> = 2 V and V<sub>ds</sub> = 0.6 V

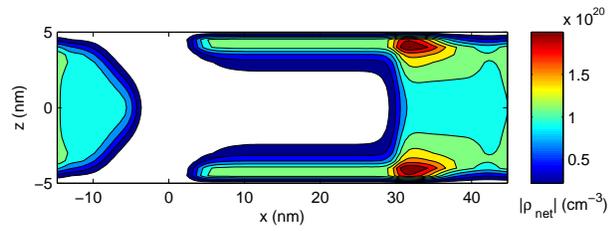


Figure 12.4: Net charge carrier density for device 1 at  $V_{gs} = 2$  V and  $V_{ds} = 0.6$  V

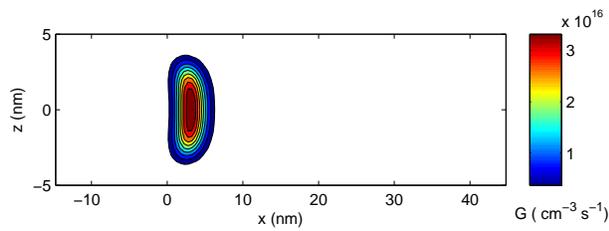


Figure 12.5: Localized tunneling rate for device 1 at  $V_{gs} = 0$  V and  $V_{ds} = 0.6$  V

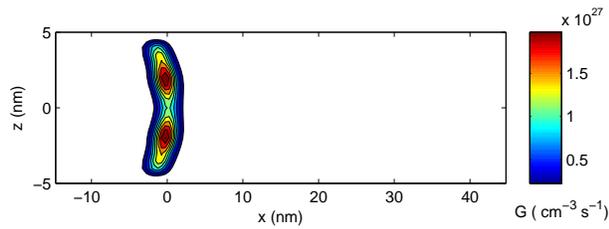


Figure 12.6: Localized tunneling rate for device 1 at  $V_{gs} = 2$  V and  $V_{ds} = 0.6$  V

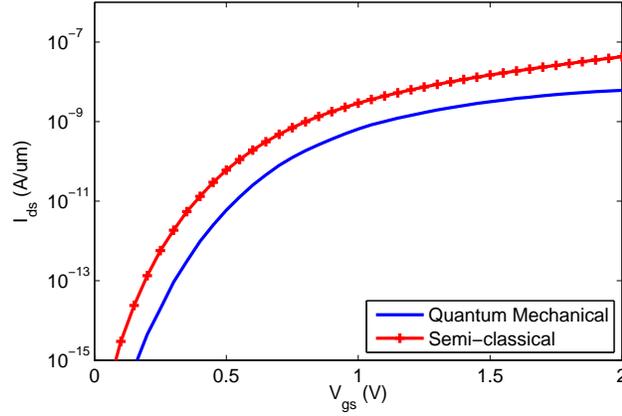


Figure 12.7: Current calculated using quantum mechanical and semi-classical approach [3] for device 1 for  $V_{ds} = 0.6$  V

in total (Table I), so that tunneling perpendicular to the gate dielectric is expected at higher gate voltages [8] thus giving rise to higher currents in the semiclassical picture. The charge carrier distribution is shown in Fig. 12.8 and is again similar to the semiclassical result apart from the quantum confinement which repels the charge away from the gate dielectric. The depletion region reaches a depth of around 5 nm under the gate ( $x=-20$  to  $x=0$ ). Even though the device now has a 20 nm thickness, the dominant localized tunneling rate is still in the center of the device at 0 V gate bias (Fig. 12.9). At 0.4 V the localized tunneling rate splits into two distinct regions (Fig. 12.10). From a gate bias of 1.8 V on the tunneling component under the gate can be clearly distinguished (Fig. 12.11). However, the tunneling area does not extend over the total gate length of 20 nm because of quantum confinement in the x-direction on top of the confinement in the z-direction. Similar to device 1, the quantum confinement leads to a deterioration of the subthreshold swing in Fig. 12.12. And while the semiclassical simulation suggests a higher current for device 2 compared to device 1, a proper quantum mechanical calculation comes to the opposite conclusion.

## 12.4 Conclusion

A two-dimensional quantum mechanical modeling of phonon-assisted BTBT has been demonstrated. The current predicted by the two-dimensional quantum mechanical approach has a worse subthreshold swing and a strongly reduced current compared to semiclassical predictions in the double gate TFETs

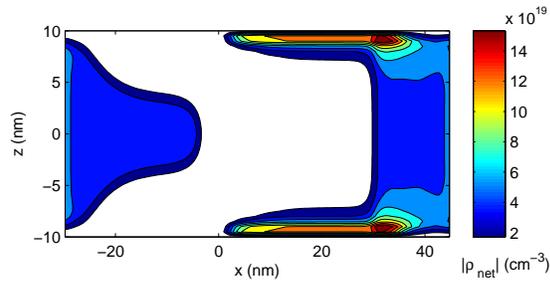


Figure 12.8: Net charge carrier density for device 2 for  $V_{gs} = 2$  V and  $V_{ds} = 0.6$  V

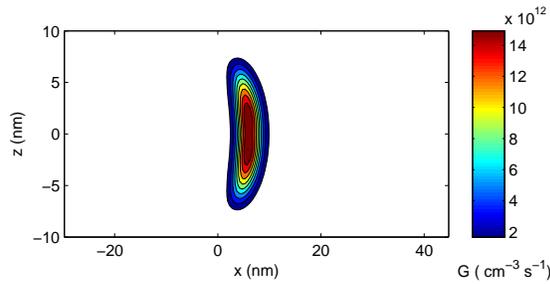


Figure 12.9: Localized tunneling rate for device 2 for  $V_{gs} = 0.4$  V and  $V_{ds} = 0.6$  V

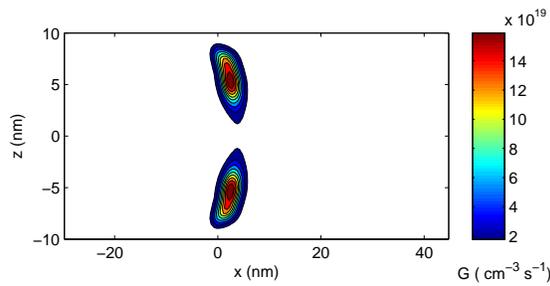


Figure 12.10: Localized tunneling rate for  $V_{gs} = .4$  V and  $V_{ds} = 0.6$  V

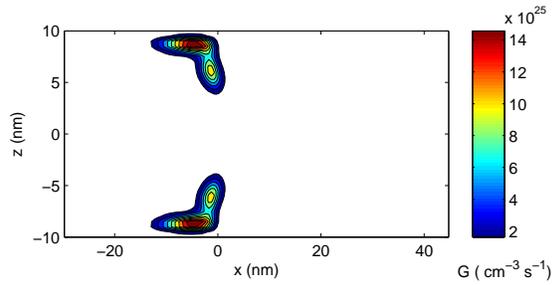


Figure 12.11: Localized tunneling rate for  $V_{gs} = 2$  V and  $V_{ds} = 0.6$  V

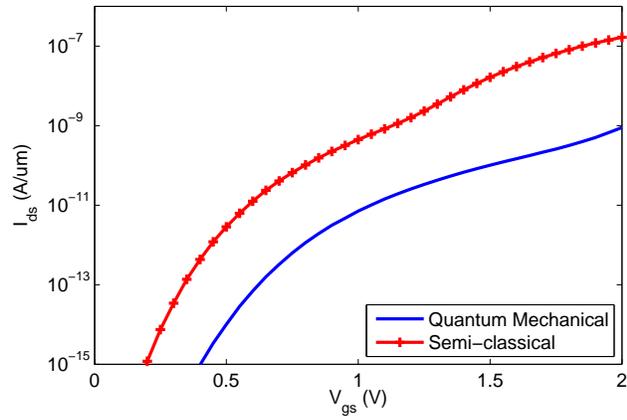


Figure 12.12: Current calculated using quantum mechanical and semi-classical approach [3] for device 2 for  $V_{ds} = 0.6$  V

under study. These results highlight the importance of having more accurate quantum mechanical models for the prediction of the performance of novel devices such as TFETs. Direct semiconductors were not studied but also in direct semiconductors, the effects of quantum confinement are important and similar conclusions are expected.

## 12.5 Acknowledgements

William Vandenberghe gratefully acknowledges the support of a Ph.D. stipend from IWT-Vlaanderen. This work was supported by imec's Industrial Affiliation Program.

## Bibliography

- [1] X. Yuan, J.-E. Park, J. Wang, E. Zhao, D. Ahlgren, T. Hook, J. Yuan, V. Chan, H. Shang, C.-H. Liang, R. Lindsay, S. Park, and H. Choo, "Gate-induced-drain-leakage current in 45-nm cmos technology," *Device and Materials Reliability, IEEE Transactions on*, vol. 8, no. 3, pp. 501–508, sept. 2008.
- [2] C. Hu, P. Patel, A. Bowonder, K. Jeon, S. H. Kim, W. Y. Loh, C. Y. Kang, J. Oh, P. Majhi, A. Javey, T.-J. K. Liu, and R. Jammy, "Prospect of tunneling green transistor for 0.1v cmos," in *Electron Devices Meeting (IEDM), 2010 IEEE International*, dec. 2010, pp. 16.1.1–16.1.4.
- [3] *Sentaurus Device*, Synopsis, March 2010.
- [4] M. Luisier and G. Klimeck, "Performance comparisons of tunneling field-effect transistors made of insb, carbon, and gasb-inas broken gap heterostructures," in *Electron Devices Meeting (IEDM), 2009 IEEE International*, dec. 2009, pp. 1–4.
- [5] W. Vandenberghe, B. Soree, W. Magnus, and M. V. Fischetti, "Generalized phonon-assisted zener tunneling in indirect semiconductors with non-uniform electric fields: A rigorous approach," *Journal of Applied Physics*, vol. 109, no. 12, pp. 124503–124503–12, jun 2011.
- [6] C. Rivas, R. Lake, G. Klimeck, W. R. Frensley, M. V. Fischetti, P. E. Thompson, S. L. Rommel, and P. R. Berger, "Full-band simulation of indirect phonon assisted tunneling in a silicon tunnel diode with delta-doped contacts," *Applied Physics Letters*, vol. 78, no. 6, pp. 814–816, 2001. [Online]. Available: <http://link.aip.org/link/?APL/78/814/1>

- [7] W. G. Vandenberghe, B. Sorée, W. Magnus, G. Groeseneken, and M. V. Fischetti, “Impact of field-induced quantum confinement in tunneling field-effect devices,” *Applied Physics Letters*, vol. 98, no. 14, p. 143503, 2011. [Online]. Available: <http://link.aip.org/link/?APL/98/143503/1>
- [8] W. Vandenberghe, A. Verhulst, G. Groeseneken, B. Soree, and W. Magnus, “Analytical model for point and line tunneling in a tunnel field-effect transistor,” in *Simulation of Semiconductor Processes and Devices, 2008. SISPAD 2008. International Conference on*, sept. 2008, pp. 137–140.

### **List of corrections and clarifications**

- In table 12.1, the same dielectric constant is used because the Poisson equation is solved assuming a single dielectric constant for gate and semiconductor.

## Chapter 13

# Study of counter-doped pockets

### Optimization of Gate-on-Source-only Tunnel FETs with Counter-doped Pockets

Kuo-Hsing Kao, Anne S. Verhulst, William G. Vandenberghe, Bart Sorée, Wim Magnus, Daniele Leonelli, Guido Groeseneken, and Kristin De Meyer

Accepted for publication in *IEEE Transactions on Electronic Devices*, 2012

©2012 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

### **Contributions of third author**

- Development of code for calculation of indirect tunneling current
- Providing user-friendly interface for current calculation
- Text correcting and editing

## Abstract

We investigate a promising tunnel FET configuration having a gate on the source only, which is simultaneously exhibiting steeper subthreshold slope and higher on-current than the lateral tunneling configuration with a gate on the channel. Our analysis is performed based on a recently developed two-dimensional quantum mechanical simulator calculating band-to-band tunneling and including quantum confinement. It is shown that the two disadvantages of the structure, namely the sensitivity to gate alignment and physical oxide thickness, are mitigated by placing a counter-doped parallel pocket underneath the gate-source overlap. The pocket also significantly reduces the field-induced quantum confinement. The findings are illustrated with all-Si and all-Ge GoSo TFET simulations.

## 13.1 Introduction

The tunnel field-effect transistor (TFET) is one of the devices that can reach sub-60 mV/decade subthreshold swing (SS) since the injection mechanism is based on band-to-band tunneling (BTBT) instead of thermal diffusion [1–6]. Various architectures have been proposed to enhance the TFET performance, e.g. heterostructure TFETs [3–10], encroaching source TFETs [11–14], a TFET with a short gate electrode [15] and TFETs with a large gate-source overlap [8–10, 16–18]. For the latter TFETs, the improved performance is due to the enhanced vertical tunneling component which has higher current and steeper onset than the lateral component [17, 18]. To eliminate the lateral tunneling component, the gate-channel overlap can be removed, resulting in a gate-on-source-only (GoSo) TFET as shown in Fig. 13.1(a) for a double-gate configuration. In this article, we present an in-depth quantum-mechanical study of this promising TFET configuration, taking into account the field-induced quantum confinement (FIQC) effects [19, 20] and focusing on minimization of the device variability.

We start in Sec. 13.2 with a description of the device structure and the physical models used in the quantum mechanical and semi-classical simulations. In Sec. 13.3 we detail the fundamental properties of the BTBT current in a double gate GoSo TFET, in particular the impact of source doping concentration, gate alignment, physical oxide thickness and body thickness. In Sec. 13.4, we show that the variability due to gate misalignment and fringing fields in the GoSo configuration can be reduced by counter-doped pockets underneath the gate-on-source region. In Sec. 13.6, we illustrate the generality of the findings with all-Ge GoSo TFET simulations.

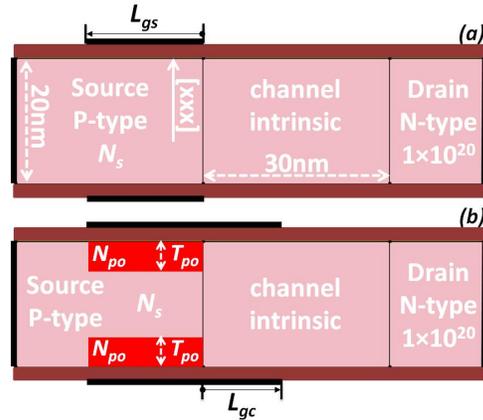


Figure 13.1: (a): The simulated GoSo TFET structure with a gate-source overlap  $L_{gs} = 20$  nm and a source doping concentration  $N_s$ . The solid arrow indicates the orientation of vertical tunneling, which is chosen to be [100] or [111]. (b): the same structure as (a) except including  $T_{po}$ -thick counter-doped pockets (n-type) with a concentration of  $N_{po}$  and a gate-channel overlap  $L_{gc}$ .

## 13.2 Device structure and physical models

The simulated GoSo n-channel TFET (nTFET) is schematically presented in Fig 1(a). In all-Si TFET simulations, an n-type doped drain with a concentration of  $1 \times 10^{20} \text{ cm}^{-3}$ , an intrinsic 30 nm long channel, a 20 nm thick body, a 4.05 eV gate workfunction (WF), and a gate-source overlap of  $L_{gs} = 20$  nm are specified. Abrupt and uniform doping profiles are used. Quantum-mechanical (QM) simulations are performed by using a recently developed two-dimensional simulator solving the Poisson-Schrödinger equation self-consistently, after which the nonlocal phonon-assisted BTBT current is calculated [20]. This simulator captures FIQC, both for vertical and lateral BTBT components. It captures size quantization effects and uses Fermi-Dirac statistics. A limitation of the quantum mechanical simulator is that a single dielectric constant is specified for the entire structure, as a result, the dielectric constant of the oxide  $\epsilon_{ox}$  and semiconductor materials  $\epsilon_s$  are the same (for Si:  $\epsilon_{ox} = \epsilon_s = 11.9$  and for Ge:  $\epsilon_{ox} = \epsilon_s = 16$ ). Hence in the quantum mechanical simulator, structures with different EOT are investigated by taking a different physical oxide thickness. Another limitation is that only the BTBT current is calculated (no intravalley scattering), which reduces the accuracy of the predictions for high currents. Therefore, this QM simulation is mainly utilized for small currents to accurately determine the onset of BTBT being strongly affected by FIQC [20]. On the other hand, semi-classical simulations [21] considering the drift-diffusion effect of carriers are used as well to complement the quantum-mechanical pre-

dictions. The semi-classical simulations are also based on Fermi-Dirac statistics and non-local phonon-assisted BTBT models, whereby the theoretically-calculated coefficients of pure Si and pure Ge along a specific direction are used [22]. Semi-classical simulations do not take into account the split up of the subbands due to FIQC, but using the modified local-density approximation (MLDA) model [23, 24], a redistribution of the carriers near the gate oxide is included. Since a dopant-induced band gap narrowing model is not included in the quantum mechanical simulation, no band gap narrowing model is included in the semi-classical simulation for a fair comparison.

### 13.3 Basic properties of the GoSo TFET

In this section we discuss the basic BTBT performance of a double-gate GoSo TFET in terms of the source doping concentration, gate alignment, fringing field and body thickness.

#### 13.3.1 Source Doping Concentration $N_s$

Fig. 13.2 (a) and (b) respectively present the basic input characteristics of all-Si GoSo nTFETs (see Fig. 13.1(a)) based on quantum mechanical and semi-classical simulations. Fig. 13.2(a) highlights the promise of GoSo configuration: for the  $N_s = 9 \times 10^{19} \text{ cm}^{-3}$  case, an average SS of 45 mV/dec over 4 decades is predicted for this all-Si TFET with tunneling orientation along [111]. With increasing  $N_s$ , the onset voltage  $V_{\text{onset}}$  of vertical BTBT increases because the voltage drop across the gate oxide increases in a semi-classical picture [25]:

$$V_{\text{onset}} = V_{\text{FB}} + \frac{E_g}{q} + \frac{1}{C_{\text{ox}}} \sqrt{2N_s \epsilon_s E_g} \quad (13.1)$$

with  $V_{\text{FB}}$  flatband voltage,  $E_g$  band gap,  $q$  elementary charge and  $C_{\text{ox}}$  gate oxide capacitance per unit area. With increasing  $N_s$ , the band bending in the vertical direction (source-oxide) is sharper and therefore the FIQC effect results in larger subband energies  $E_{\text{sub}}$ , which further increases  $V_{\text{onset}}$ . This is illustrated in the inset of Fig. 13.2(b), which displays the difference of  $V_{\text{onset}}$ , between the quantum mechanical and the semi-classical simulations at  $I_{\text{ds,off}} = 50 \text{ pA}/\mu\text{m}$ . Note that the trend in Fig. 13.2 is opposite to that of lateral BTBT, where the source depletion at onset is limited, such that the voltage drop across the oxide is rather small, and where the impact of increased electric field and sharper band bending in the lateral direction (source-channel) with higher  $N_s$  dominates, resulting in higher BTBT and a smaller  $V_{\text{onset}}$  [26, 27]. With increasing  $N_s$ , the average SS from  $V_{\text{onset}}$  to  $V_{\text{onset}} + V_{\text{dd}}$ , with  $V_{\text{dd}}$  the supply voltage window, improves due to steeper band bending and corresponding larger electric field. Note that this trend is the same for both vertical and lateral tunneling.

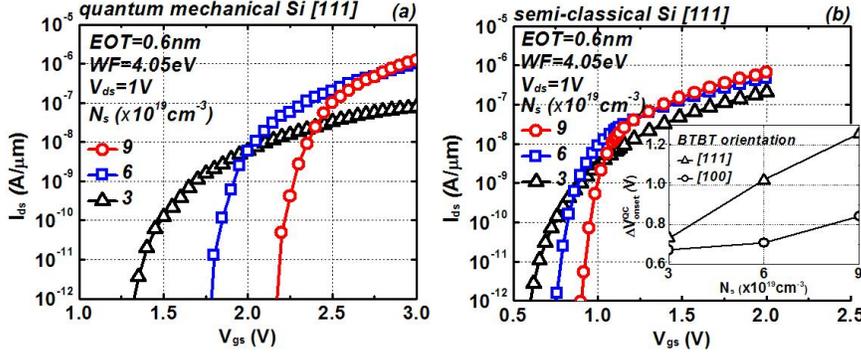


Figure 13.2: The input characteristics of all-Si GoSo nTFETs as shown in Fig. 13.1(a) with  $L_{gs} = 20\text{ nm}$  and different source doping concentration  $N_s$  based on (a) quantum mechanical and (b) semi-classical simulations. The inset of (b) shows the shift in  $V_{onset}$  due to FIQC as a function of  $N_s$ .

### 13.3.2 Gate Alignment

Fig. 13.3 (a) and (b) show the input characteristics of TFETs with various gate-channel overlap  $L_{gc}$  based on quantum mechanical and semi-classical simulations, respectively. Upon increasing  $L_{gc}$ , the steep SS of the vertical tunneling is deteriorated by an increase of the lateral tunneling. The first few nm gate-channel overlap have the largest impact, which implies that variability caused by gate misalignment is a serious issue for the basic all-Si GoSo configuration. A GoSo TFET with a gate-source underlap ( $L_{gc} = -1\text{ nm}$ , see Fig. 13.3(a)) only causes a minor shift in onset voltage compared to the default GoSo, since both configuration have no lateral tunneling.

The gate-channel overlap also affects the on-current. Since the quantum-mechanical simulations do not include scattering or drift-diffusion, the on-current is discussed based on the semi-classical simulations. As shown in Fig. 13.3(b), a lower on-current at  $V_{gs} = 2\text{ V}$  is observed for  $L_{gc} = 0\text{ nm}$  compared to  $L_{gc} \neq 0\text{ nm}$ . This can be explained by a band diagram parallel and  $0.5\text{ nm}$  away from the oxide interface. As the solid lines ( $L_{gc} = 0\text{ nm}$ ) in the inset of Fig. 13.3(b) show a potential barrier between the source and the ungated channel region limiting the on-current. A limited gate-channel overlap (dashed lines,  $L_{gc} = 3\text{ nm}$ ) is sufficient to push this potential barrier far into the channel, such that the barrier is at an energy level below the level into which the carriers tunnel, and the BTBT current is no longer affected.

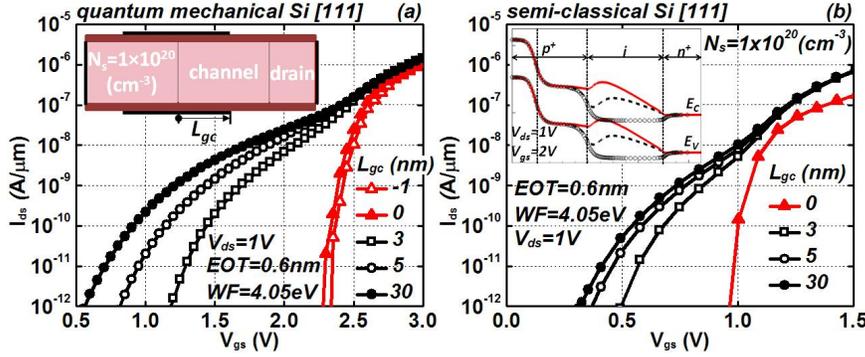


Figure 13.3: The input characteristics of all-Si TFETs based on (a): quantum mechanical and (b): semi-classical simulations with different gate-channel overlap  $L_{gc}$  (see the inset in (a)). The inset in (b) shows a band diagram along and 0.5 nm away from the semiconductor-oxide interface with  $L_{gc} = 0$  nm (solid), 3 nm (dashed) and 30 nm (circle). The gate length is equal to  $L_{gs} + L_{gc}$ , in which  $L_{gs} = 20$  nm.

### 13.3.3 Fringing Field

The GoSo TFET configuration is sensitive not only to the effective oxide thickness (EOT), but also to the physical oxide thickness, as shown in Fig. 13.4(a). With increasing the physical oxide thickness, the lateral BTBT is enhanced, which is illustrated in Fig. 13.4(b) and (c) presenting the BTBT generation rate for the GoSo TFETs with 3 nm and 6 nm physical oxide thicknesses at  $V_{gs} = 0.9$  V and  $V_{ds} = 1$  V. For a given gate bias, a physically thicker oxide extends the gate potential further via a larger fringing field, and therefore affects the electrostatic potential in the channel near the gate-oxide interface over a larger area, which enhances the lateral BTBT. These results are obtained based on the semi-classical simulations, due to the fact that the quantum-mechanical simulator does not allow different dielectric constants in the semiconductor and oxide. The impact of the quantum confinement is therefore not included, but this is not expected to affect the trend qualitatively. The necessity to use a physically-thin gate dielectrics puts serious limitations to the scaling of the basic all-Si GoSo configuration, as the gate leakage current will be large.

### 13.3.4 Body Thickness

The body thickness has little impact on the TFET performance, as long as it is sufficiently thick. As shown in Fig. 13.5, the TFETs with body thickness  $T_B = 20$  nm and 15 nm show almost the same input characteristics. When  $T_B$  is decreased to 5 nm, the SS and on-current are degraded. This is because the source region underneath the gate has become fully depleted, vertical tunneling

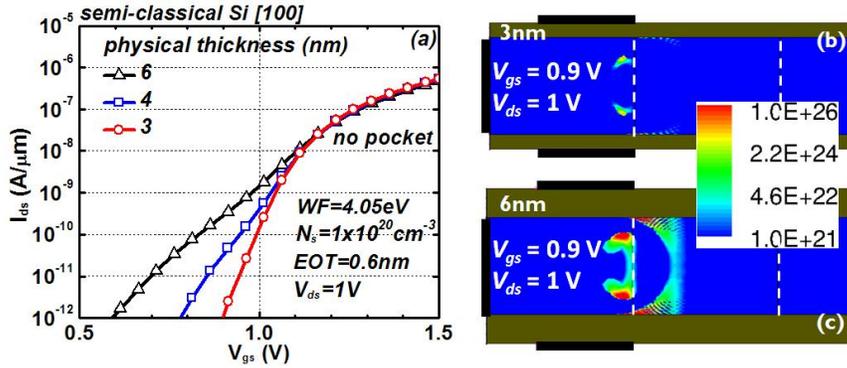


Figure 13.4: (a): Input characteristics of GoSo all-Si TFETs as shown in Fig. 13.1(a) with  $L_{gs} = 20\text{ nm}$  and different physical oxide thickness  $t_{ox}$  but retaining the same effective oxide thickness ( $EOT = 0.6\text{ nm}$ ) based on semi-classical simulations. (b) and (c) show the semi-classical BTBT hole and electron generation rates of GoSo n-channel TFETs.

no longer occurs and the current results from lateral tunneling from the ungated source region into the depleted source region underneath the gate electrode.

## 13.4 Counter-doped Pocket

In this section, the impact of parallel counter-doped pockets on the GoSo TFET performance is analyzed in detail and shown to be beneficial.

### 13.4.1 Impact on the Input Characteristics

Fig. 13.1(b) shows the TFET structure with counter-doped pockets (n-type pockets in the p-type source of an nTFET) underneath the gate-on-source electrode. Fig. 13.6 (a) and (b) present the input characteristics of GoSo TFETs without and with pockets based on quantum mechanical simulations where the vertical tunneling occurs in the [111] and [100] orientation, respectively. The thicker and more heavily counter-doped pockets decrease the onset voltage of vertical BTBT more strongly. The reason is twofold. First, there is an inherent band bending at the junction between the p-type source region and the n-type pocket. This implies that the conduction band at the semiconductor-oxide interface is energetically closer to the valence band edge deep in the source region, from where tunneling starts, than for the GoSo TFET without pocket. The second reason can be visualized by the offset in FIQC-induced onset as shown in Fig. 13.6(c). As can be seen, the quantum confinement effect is strongly

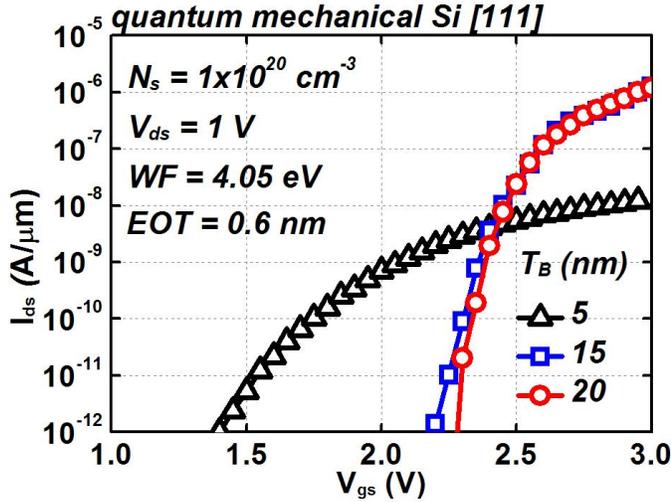


Figure 13.5: The input characteristics of GoSo TFETs as shown in Fig. 13.1(a) with  $L_{gs} = 20$  nm and different body thickness  $T_B$  based on quantum mechanical simulations.

reduced by a heavily-doped thick pocket (approaching 0 V).

To understand the impact of the pocket better, Fig. 13.7 schematically illustrates the electrostatic potential energy in the source in a direction perpendicular to the gate dielectric. We consider the special case where the Fermi level in the n-pocket is at the same energy level as the workfunction of the gate and where the pocket thickness – within the Shockley approximation – equals the equilibrium space charge width of the n-side of the p-n junction. In this case, at the semiconductor-oxide interface, there is a zero electric field and the conduction (valence) band edge  $E_c$  ( $E_v$ ) has a parabolic shape inside the n-pocket (solid curve in Fig. 13.7). For the doping levels of  $1 \times 10^{20} \text{ cm}^{-3}$  in both p- and n-region, the pocket thickness is about 2.8 nm and the maximum electric field  $E_{\text{max}}$  at the p-n junction is about 4.3 MV/cm.

According to the parabolic quantum well concept, the first subband energy is given by the second state of the harmonic oscillator at  $3/2\hbar\omega$  in the half potential well, in which  $\hbar$  is the Plank constant and  $\omega$  is given by  $\omega = \sqrt{k/m_e}$ , with  $m_e$  the electron mass in the conduction band (or hole mass if quantum confinement (QC) in the valence band is considered) and  $k$  representing the parabolic shape by [28]:

$$U(y) = \frac{1}{2}ky^2, \quad (13.2)$$

The potential distribution  $U(y)$  in the n-region of the p-n junction is given

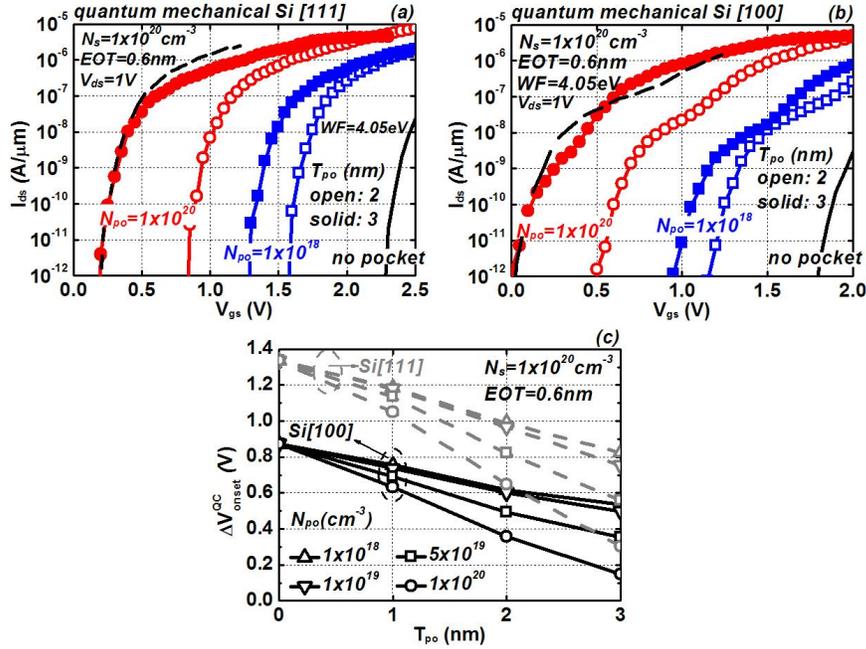


Figure 13.6: Input characteristics of all-Si GoSo TFETs as shown in Fig. 13.1(b) with  $L_{gs} = 20 \text{ nm}$ , different pocket thickness  $T_{po}$  and doping concentration  $N_{po}$  based on quantum mechanical simulations with tunneling along (a): [111] and (b): [100]. The dashed curves are shifted versions of the no-pocket curves to highlight the SS degradation caused by the pockets. (c): The impact of pocket thickness  $T_{po}$  on the FIQC-induced onset voltage shift  $\Delta V_{onset}$  for various  $N_{po}$  and different tunneling orientations.

by [29]

$$U(y) = \frac{1}{2} \frac{qE_{\max}}{W_d} y^2, \quad (13.3)$$

where  $W_d$  is the total depletion width. This implies that:

$$k = \frac{qE_{\max}}{W_d} \quad (13.4)$$

As a result, the first subband energy levels are about 0.12 eV (indicated as a solid arrow to a solid horizontal line in Fig. 13.7) and 0.26 eV for the longitudinal and transverse electron effective masses, respectively. If on the other hand no pocket is present and the same  $E_{\max}$  is reached at  $y = 0$ , a triangular potential well approximation can be used (dotted line in Fig. 13.7). The quantum confinement is stronger due to the deeper potential well. The first subband is at 0.45 (0.78) eV for longitudinal (transverse) electron, as indicated by a dotted arrow to a dotted horizontal line in Fig. 13.7 [19]. The subband level is rather close to the configuration with the pocket, yet the required gate voltage for the onset of tunneling is much larger. Fig. 13.7 therefore explains how a thicker pocket physically increases the width of the parabolic potential well, and accordingly reduces the quantum confinement. Note that both the triangular and parabolic potential approximation ignore the band bending towards the body center (dashed  $E_c$  ( $y \approx 2.8$  nm) shown in Fig. 13.7) and overestimate the subband energies.

The average SS is degraded with increasing the pocket thickness and concentration (see dashed curve and left-most solid curve in Fig. 13.6(a)). This is because at  $V_{\text{onset}}$ , which occurs at a band bending close to the initial crossover of the conduction band with the valence band, the total depletion length is longer in the pocket configuration than in the no-pocket configuration, as can be understood from Fig. 13.7. Therefore, in the pocket configuration, the initial tunnel path is longer and gate voltage increase result in smaller changes of the conduction band edge at the semiconductor-oxide interface. Both effects are degrading SS.

This effect is also present, yet less obvious in Fig. 13.6(b), which displays a kink in the characteristics due to the presence of two subband levels, one associated with the heavier longitudinal and one with the lighter transverse electron mass [19]. The former starts earlier and has smaller BTBT currents than the latter. Due to the reduced FIQC with increasing pocket thickness and concentration,  $V_{\text{onset}}$  of both components becomes more comparable, which can improve the average SS.

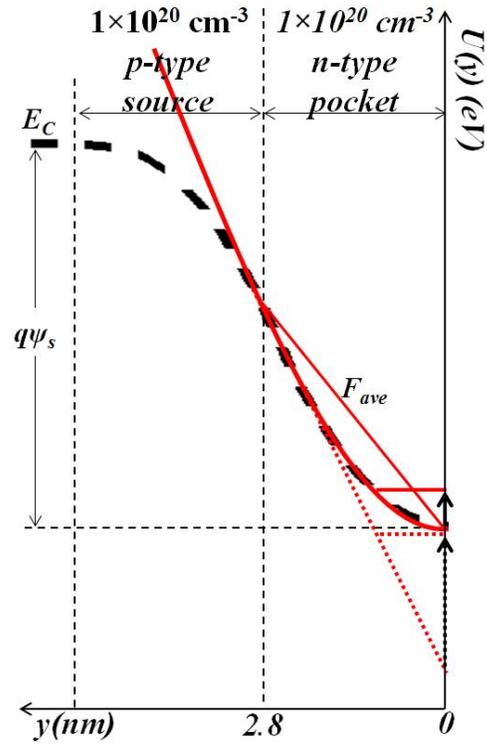


Figure 13.7: Illustration of a parabolic potential energy profile (solid curve) for the conduction band (dashed curve,  $E_c$ ) of a Si p-n junction with a maximum electric field  $E_{\max} = 4.3$  MV/cm at the junction. The dotted straight line forming a triangular potential well at the oxide ( $y = 0$ ) approximates the potential profile in a uniformly doped source region with the same field  $E_{\max}$  at  $y = 0$ . The solid straight line indicates the average uniform electric field  $F_{\text{ave}}$  in the fully-depleted n-type pocket. Oxide potential is assumed to be infinitely high.

### 13.4.2 Improved Immunity against Gate Alignment Variability

As discussed in Sec. 13.3.2, the gate alignment causes a serious variability issue for the GoSo TFET without the counter-doped pocket. However, the difference of  $V_{\text{onset}}$  between devices with  $L_{\text{gc}} = 0$  and 3 nm decreases with introducing a 3-nm-thick pocket and with increasing the pocket concentration from  $1 \times 10^{18}$  to  $1 \times 10^{20} \text{ cm}^{-3}$  (see Fig. 13.8(a) and (b)). This can be attributed to the earlier onset of the vertical BTBT, which reduces the impact of the lateral BTBT tail, and the introduction of an n-i-n junction close to the oxide interface, which disables the lateral BTBT at the interface of the n-type pocket and intrinsic channel. Fig. 13.8(a) shows a higher on-current on the curves with  $L_{\text{gc}} = 3$  nm than  $L_{\text{gc}} = 0$  nm. This is because the additional 3 nm gate-channel overlap reduces the potential barrier at the source-channel interface, as discussed in Sec. 13.3.2. It can therefore be concluded that for the all-Si GoSo TFET with an at least 3 nm thick pocket with a concentration of at least  $1 \times 10^{20} \text{ cm}^{-3}$ , a gate-channel overlap  $L_{\text{gc}} = 3$  nm, and in extension a full gate-channel overlap, results in better performance, in particular a higher on-current without SS penalty, although it is at the expense of an increased gate capacitance.

### 13.4.3 Larger Physical Oxide Thickness Allowed

Fig. 13.8(c) indicates that the SS deterioration resulting from the fringing field of the gate electrode can be suppressed by the counter-doped pockets. The reasons are the same as in the previous paragraph: an earlier onset of vertical versus lateral tunneling as well as a disabling of the lateral BTBT at the interface of the n-pocket and the channel decrease the observable fringing-field induced lateral tunneling component. Heavily-doped pockets, as the result, can not only decrease the  $V_{\text{onset}}$  of vertical BTBT but also improve the immunity of the transistor characteristics against gate misalignment and fringing field degradation. Overall, the GoSo TFET pocket thickness optimization consists of a careful trade-off between the beneficial impacts of improved robustness against variability, the possibility to use realistic workfunctions between 4 and 5 eV and on the other hand the negative impact on the SS. For a gate workfunction of at least 4 eV, the optimal pocket thickness which brings the  $V_{\text{onset}}$  to 0 V for the all-Si TFET is between 3 and 4 nm.

## 13.5 GoSo Ge TFET

The impact of counter-doped pockets in GoSo TFETs is further illustrated with all-Ge TFETs with vertical tunneling along [100] (Fig. 13.9(a)), whereby the direct bandgap transitions in Ge are not activated to make the analysis

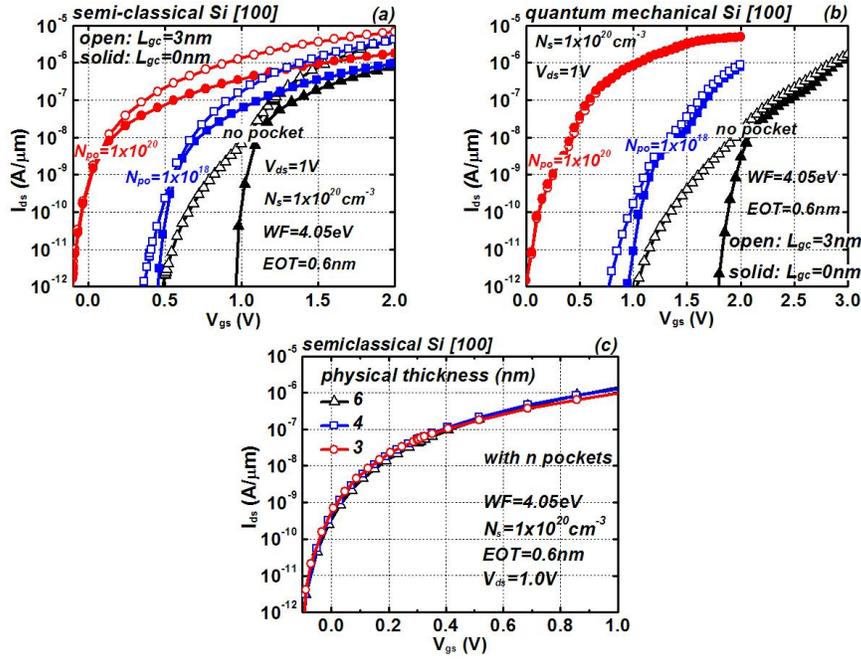
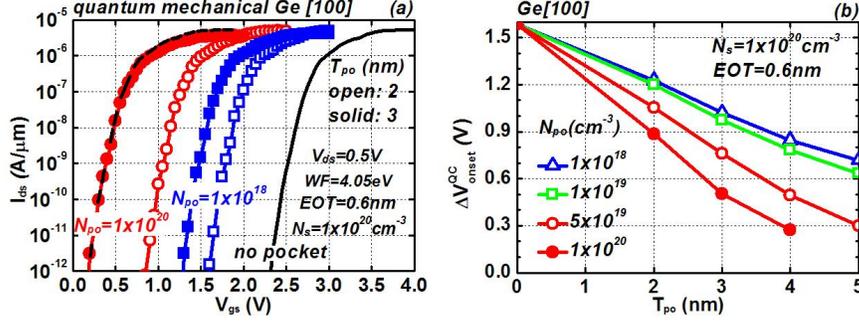


Figure 13.8: Input characteristics of all-Si TFETs as shown in Fig. 13.1(b) with or without 3-nm gate-channel overlap  $L_{gc}$  and 3-nm-thick pockets based on (a): semi-classical and (b): quantum mechanical simulations. (c): Input characteristics of GoSo TFETs with the device structure shown in Fig. 13.1(b),  $L_{gc} = 0$  nm and different physical oxide thickness. The gate length is equal to  $L_{gs} + L_{gc}$ , in which  $L_{gs} = 20$  nm.

Figure 13.9: (a): Input characteristics of GoSo Ge TFETs with  $L_{gs} = 20$  nm, different pocket thickness  $T_{po}$  and doping concentration  $N_{po}$  based on quantum mechanical simulations with tunneling along [100]. The dashed curves are shifted versions of the no-pocket curves. (b): The impact of  $T_{po}$  and  $N_{po}$  on the FIQC-induced onset voltage shift  $\Delta V_{onset}^{QC}$ .



of pocket thickness more transparent [22]. The effective mass for both configurations is given by  $3m_L m_T / (2m_L + m_T)$ , where  $m_L$  and  $m_T$  are electron longitudinal and transverse masses, respectively (see Table I), with a value of  $0.26m_0$  for Si and  $0.12m_0$  for Ge. The simulated all-Ge TFET device uses the same structure of Fig. 13.1(b) except with a 30 nm body thickness, a 50 nm intrinsic length to decrease the source-drain leakage, a 2.5 nm oxide thickness ( $EOT = 0.6$  nm) and  $L_{gc} = 0$  nm. Fig. 13.9(a) shows the input characteristic of all-Ge nTFETs without and with counter-doped pockets and Fig. 13.9(b) shows the delayed  $\Delta V_{onset}^{QC}$  of the BTBT induced by FIQC. Comparing Fig. 13.6(c) and Fig. 13.9(b), the quantum confinement is stronger in Ge TFETs because of the smaller effective mass. In the case without pocket ( $T_{po} = 0$ ), for example, the delayed  $\Delta V_{onset}^{QC}$  induced by FIQC is about 1.6 V in Ge (1.3 V in Si). To reduce the  $\Delta V_{onset}^{QC}$  to about 0.3 V, a pocket with  $N_{po} = 1 \times 10^{20} cm^{-3}$  and  $T_{po} = 4$  nm is needed. The SS degradation in the all-Ge TFET is smaller than in the all-Si TFET of Fig. 13.6(a) (see comparison of dashed curves with the curves for 3 nm pocket with  $N_{po} = 1 \times 10^{20} cm^{-3}$ ). This is due to the smaller total depletion length in Ge related to its smaller bandgap (see Sec. 13.4.1) The optimal pocket thickness which results in  $V_{onset} = 0$  V for a gate workfunction of at least 4 eV for the all-Ge TFET is also between 3 and 4 nm.

Similar to the all-Si GoSo TFET (Fig. 13.8(a)), the impact of gate misalignment is no longer observable for the all-Ge GoSo TFET when a 3-nm-thick pocket with doping concentration  $N_{po} = 1 \times 10^{20} cm^{-3}$  is inserted.

Indirect BTBT				
	[100]	[110]	[111]	
Si	$m_L = 0.98$	$m_t = 0.19$	$m_T = 0.19$	$\frac{3m_L m_T}{2m_L + m_T} = 0.26$
Ge	$\frac{3m_L m_T}{2m_L + m_T} = 0.12$	$\frac{2m_L m_T}{m_L + m_T} = 0.32$	$\frac{3m_L m_T}{m_L + 2m_T} = 0.22$	$m_L = 1.6$   $\frac{9m_L m_T}{8m_L + m_T} = 0.09$

Table 13.1: Electron effective masses of Si and Ge for indirect tunneling along three different technologically important orientations [22]. All effective masses are in the unit of the free electron mass  $m_0$ .

## 13.6 Conclusions

By using two-dimensional quantum mechanical simulations including field-induced quantum confinement together with semi-classical simulations including mobility models, we have investigated that a TFET with a GoSo configuration exhibits steeper SS and higher on-current than the TFET which exhibits only lateral BTBT. A Si GoSo TFET with [111] tunneling orientation and optimized counter-doped pockets shows better performance than tunneling along [100]. The counter-doped pockets underneath the gate-source overlap decrease the quantum confinement, which is a potential source of variability, and prevent the steep SS from lateral BTBT degradation. This reduces the variability associated with the gate alignment as well as the oxide thickness limitation. The optimal gate-channel overlap therefore depends on the specific capacitance and current requirements, with increasing overlap resulting in enhanced on-current in the channel-resistance limited current regime. For a gate workfunction of at least 4 eV, the optimal pocket thickness for [111] Si as well as for [100] Ge is between 3 and 4 nm.

## Bibliography

- [1] D. Leonelli, A. Vandooren, R. Rooyackers, A. S. Verhulst, S. D. Gendt, M. M. Heyns, and G. Groeseneken, "Performance enhancement in multi gate tunneling field effect transistors by scaling the fin-width," *Japanese Journal of Applied Physics*, vol. 49, no. 4, p. 04DC10, 2010. [Online]. Available: <http://jjap.jsap.jp/link?JJAP/49/04DC10/>
- [2] K. Boucart and A. M. Ionescu, "A new definition of threshold voltage in tunnel FETs," *Solid-State Electronics*, vol. 52, no. 9, pp. 1318 – 1323, 2008, papers Selected from the 37th European Solid-State Device Research Conference - ESSDERC'07. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038110108001196>
- [3] F. Mayer, C. Le Royer, J.-F. Damlencourt, K. Romanjek, F. Andrieu, C. Tabone, B. Previtali, and S. Deleonibus, "Impact of soi, si1-xgeoxi and geoi substrates on cmos compatible tunnel fet performance," in *Electron Devices Meeting, 2008. IEDM 2008. IEEE International*, dec. 2008, pp. 1–5.
- [4] T. Krishnamohan, D. Kim, S. Raghunathan, and K. Saraswat, "Double-gate strained-ge heterostructure tunneling FET (TFET) with record high drive currents and  $\sim 60\text{mV}/\text{dec}$  subthreshold slope," in *Electron Devices Meeting, 2008. IEDM 2008. IEEE International*, dec. 2008, pp. 1–3.

- [5] O. Nayfeh, C. Chleirigh, J. Hennessy, L. Gomez, J. Hoyt, and D. Antoniadis, "Design of tunneling field-effect transistors using strained-silicon/strained-germanium type-ii staggered heterojunctions," *Electron Device Letters, IEEE*, vol. 29, no. 9, pp. 1074–1077, sept. 2008.
- [6] S. H. Kim, H. Kam, C. Hu, and T.-J. K. Liu, "Germanium-source tunnel field effect transistors with record high ion/ioff," in *VLSI Technology, 2009 Symposium on*, june 2009, pp. 178–179.
- [7] A. Verhulst, W. Vandenberghe, K. Maex, S. De Gendt, M. Heyns, and G. Groeseneken, "Complementary silicon-based heterostructure tunnel-fets with high tunnel rates," *Electron Device Letters, IEEE*, vol. 29, no. 12, pp. 1398–1401, dec. 2008.
- [8] E.-H. Toh, G. H. Wang, L. Chan, G. Samudra, and Y.-C. Yeo, "Device physics and guiding principles for the design of double-gate tunneling field effect transistor with silicon-germanium source heterojunction," *Applied Physics Letters*, vol. 91, no. 24, p. 243505, 2007. [Online]. Available: <http://link.aip.org/link/?APL/91/243505/1>
- [9] S. H. K. S. H. Kim, S. Agarwal, Z. A. Jacobson, P. Matheu, C. H. C. Hu, and T. J. K. Liu, "Tunnel field effect transistor with raised germanium source," pp. 1107–1109, 2010. [Online]. Available: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=5559328](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5559328)
- [10] G. Han, P. Guo, Y. Yang, C. Zhan, Q. Zhou, and Y.-C. Yeo, "Silicon-based tunneling field-effect transistor with elevated germanium source formed on (110) silicon substrate," *Applied Physics Letters*, vol. 98, no. 15, p. 153502, 2011. [Online]. Available: <http://link.aip.org/link/?APL/98/153502/1>
- [11] K. Jeon, W.-Y. Loh, P. Patel, C. Y. Kang, J. Oh, A. Bowonder, C. Park, C. Park, C. Smith, P. Majhi, H.-H. Tseng, R. Jammy, T.-J. Liu, and C. Hu, "Si tunnel transistors with a novel silicided source and 46mv/dec swing," in *VLSI Technology (VLSIT), 2010 Symposium on*, june 2010, pp. 121–122.
- [12] Y. Yang, P.-F. Guo, G.-Q. Han, C.-L. Zhan, L. Fan, and Y.-C. Yeo, "Drive current enhancement with invasive source in double gate tunneling field-effect transistors," in *International Conference on Solid State Devices and Materials*, 2010, pp. 798–799.
- [13] D. Leonelli, A. Vandooren, R. Rooyackers, A. S. Verhulst, S. D. Gendt, M. M. Heyns, and G. Groeseneken, "Silicide engineering to boost si tunnel transistor drive current," *Japanese Journal of Applied Physics*, vol. 50, no. 4, p. 04DC05, 2011. [Online]. Available: <http://jjap.jsap.jp/link?JJAP/50/04DC05/>

- [14] K.-H. Kao, A. S. Verhulst, W. G. Vandenberghe, B. Sorée, G. Groeseneken, and K. D. Meyer, “Modeling the impact of junction angles in tunnel field-effect transistors,” *Solid-State Electronics*, vol. 69, no. 0, pp. 31 – 37, 2012. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038110111003984>
- [15] A. S. Verhulst, W. G. Vandenberghe, K. Maex, and G. Groeseneken, “Tunnel field-effect transistor without gate-drain overlap,” *Applied Physics Letters*, vol. 91, no. 5, p. 053102, 2007.
- [16] D. Kazazis, P. Jannaty, A. Zaslavsky, C. L. Royer, C. Tabone, L. Clavelier, and S. Cristoloveanu, “Tunneling field-effect transistor with epitaxial junction in thin germanium-on-insulator,” *Applied Physics Letters*, vol. 94, no. 26, p. 263508, 2009. [Online]. Available: <http://link.aip.org/link/?APL/94/263508/1>
- [17] S. Agarwal, G. Klimeck, and M. Luisier, “Leakage-reduction design concepts for low-power vertical tunneling field-effect transistors,” *Electron Device Letters, IEEE*, vol. 31, no. 6, pp. 621 –623, june 2010.
- [18] K. Ganapathi, Y. Yoon, and S. Salahuddin, “Analysis of inas vertical and lateral band-to-band tunneling transistors: Leveraging vertical tunneling for improved performance,” *Applied Physics Letters*, vol. 97, no. 3, p. 033504, 2010. [Online]. Available: <http://link.aip.org/link/?APL/97/033504/1>
- [19] W. G. Vandenberghe, B. Sorée, W. Magnus, G. Groeseneken, and M. V. Fischetti, “Impact of field-induced quantum confinement in tunneling field-effect devices,” *Applied Physics Letters*, vol. 98, no. 14, p. 143503, 2011. [Online]. Available: <http://link.aip.org/link/?APL/98/143503/1>
- [20] W. G. Vandenberghe, B. Soree, W. Magnus, M. V. Fischetti, A. S. Verhulst, and G. Groeseneken, “Two-dimensional quantum mechanical modeling of band-to-band tunneling in indirect semiconductors,” in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, 2011.
- [21] *Sentaurus Device*, Synopsis, March 2010.
- [22] K.-H. Kao, A. Verhulst, W. Vandenberghe, B. Soree, G. Groeseneken, and K. De Meyer, “Direct and indirect band-to-band tunneling in germanium-based tfets,” *Electron Devices, IEEE Transactions on*, vol. 59, no. 2, pp. 292 –301, feb. 2012.
- [23] G. Paasch and H. Übensee, “A modified local density approximation. electron density in inversion layers,” *physica status solidi (b)*, vol. 113, no. 1, pp. 165–178, 1982. [Online]. Available: <http://dx.doi.org/10.1002/pssb.2221130116>

- [24] H. C.-N. nez, W. Magnus, and F. M. Peeters, "A simplified quantum mechanical model for nanowire transistors based on non-linear variational calculus," *Journal of Applied Physics*, vol. 108, no. 6, p. 063708, 2010. [Online]. Available: <http://link.aip.org/link/?JAP/108/063708/1>
- [25] W. G. Vandenberghe, A. S. Verhulst, G. Groeseneken, B. Sorée, and W. Magnus, "Analytical model for a tunnel field-effect transistor," in *IEEE Melecon*, 2008.
- [26] P.-F. Wang, T. Nirschl, D. Schmitt-Landsiedel, and W. Hansch, "Simulation of the esaki-tunneling fet," *Solid-State Electronics*, vol. 47, no. 7, pp. 1187 – 1192, 2003, <http://www.sciencedirect.com/science/article/pii/S0038110103000455>
- [27] A. S. Verhulst, D. Leonelli, R. Rooyackers, and G. Groeseneken, "Drain voltage dependent analytical model of tunnel field-effect transistors," *Journal of Applied Physics*, vol. 110, no. 2, p. 024510, 2011. [Online]. Available: <http://link.aip.org/link/?JAP/110/024510/1>
- [28] J. Davies, *The Physics of Low-Dimensional Semiconductors*. Cambridge University Press, 1998.
- [29] S. M. Sze, *Physics of Semiconductor Devices*. Wiley, 1981.

## Part V

# Conclusions



## Chapter 14

# Conclusions and Suggestions for Future Research

### 14.1 Summary and Conclusions

#### 14.1.1 Part II: Studying the TFET using analytical models

In chapters 2 and 3, we have developed an analytical model for point and line tunneling. The potential profile was determined by solving the Poisson equation and the current was calculated using the semiclassical generation rates. A compact expression was derived for the TFET current and the analytical models were verified to match semiclassical device simulations. A small bandgap improves the TFET current for both line and point tunneling. A reduced dielectric thickness will reduce the onset voltage of line tunneling and will significantly improve the current of point tunneling. An increased doping concentration reduces the point tunneling onset voltage while it increases the line tunneling onset voltage, but at the same time the line tunneling on-current increases.

### 14.1.2 Part III: Quantum mechanical modeling of BTBT in direct semiconductors

In chapter 4, the Kane formula or the WKB approximation were shown to provide a good approximation for the actual tunnel probability for large applied biases or low fields. When larger fields or very small biases are applied and qualitative results must be obtained or the shape of the current-voltage characteristic is of importance, the Kane formula or WKB approximation breaks down. For intermediate fields an increased transmission probability can be expected while at even higher fields the transmission is suppressed. In chapter 5, the envelope function formalism was expanded towards heterostructures enabling the extension of the calculation of BTBT to heterostructures.

In chapter 6, we have shown that the bandstructure of graphene based semiconductors can be described in the same way as the light valence and the conduction band structure in a III-V semiconductor. Graphene based semiconductors have a lower BTBT probability than III-V semiconductors for the same bandgap. III-V materials are better suited for MOSFETs compared to graphene based semiconductors but graphene based semiconductors are well suited for TFETs because of their low bandgap.

In chapter 7, we have theoretically calculated the parameters  $A$  and  $B$  of Kane's direct and indirect band-to-band tunneling models for different tunneling directions ([100], [110] and [111]) for Si, Ge and unstrained  $\text{Si}_{1-x}\text{Ge}_x$ . The direct BTBT contribution in unstrained  $\text{Si}_{1-x}\text{Ge}_x$  becomes equally important to the indirect BTBT when the Ge concentration is about 80%. For 100% Ge, the calculation of BTBT generation rate in the uniform electric field limit reveals that direct tunneling always dominates.

In chapter 8, an improved model for line tunneling was developed using the WKB approximation and Fermi-Dirac statistics. Because of the increasing voltage drop over the gate oxide and the Fermi level degeneracy, an optimal doping concentration resulting in the highest TFET current was demonstrated. And because of the decreasing density of states with decreasing band gap, there is an optimal band gap which was shown to lie around 0.3 eV.

### 14.1.3 Part IV: Quantum mechanical modeling of BTBT in indirect semiconductors

In chapter 9, a general formalism to calculate the phonon-assisted BTBT current in non-uniform fields using one-particle spectral functions was presented. The expression using the spectral function enables the calculation of the tunneling probability to two- or three-dimensional potential profiles. The evaluation of the current uses the expression for the current (Eq. (9.60)) and the expression for the transition probability (Eq. (9.62)). It was shown that in the weak and uniform field limit, the formalism from this work reduces to the indirect Kane result. Furthermore, an improved uniform field model was derived and an expression using the WKB approximation were obtained. Comparing the different approximations in the case of a  $p-n$  junction, the improved uniform field model was shown to be a better approximation than the model using the WKB approximation. But for low bias conditions, no approximate model accurately describes the tunneling current and a direct numerical treatment is essential even in the one-dimensional case.

In chapter 10, the effects of quantum confinement were investigated and it was shown that quantum confinement critically determines the onset of line tunneling in TFETs and predicts a significant reduction of tunneling current as compared to semi-classical models, which erroneously predict non-zero current before the first subband is established. If the conduction band valleys are hosting electrons with different effective masses, the heavy mass electrons tunnel at lower gate bias compared to those of higher masses. However, the heavy masses decrease the electron tunneling probability which leads to a deterioration of the subthreshold swing. In chapter 11, penetration of the wavefunction into the gate dielectric was shown to slightly reduce the quantum confinement effects. A modified semiclassical model which amounts to a small correction of the existing models was presented and shown to be capable of capturing the effect of field-induced quantum confinement.

In chapter 12 a two-dimensional quantum mechanical modeling of phonon-assisted BTBT has been demonstrated. The TFET current predicted by the two-dimensional quantum mechanical approach has a worse subthreshold swing and a strongly reduced current compared to semiclassical predictions in the double gate TFETs under study. The results highlight the importance of having more accurate quantum mechanical models for the prediction of the performance of novel devices such as TFETs. Direct semiconductors were not studied but also in direct semiconductors, the effects of quantum confinement are important and similar conclusions are expected. In chapter 13, the developed two-dimensional quantum mechanical model was used to study a TFET with the gate over the source. It was shown that the gate on source only configuration has a better slope than a TFET which relies on lateral BTBT. Counter-doped pockets under the gate decrease quantum confinement reducing

variability and preventing the degradation of the slope due to lateral BTBT.

#### 14.1.4 General conclusion

In conclusion, we have given an improved description of quantum transport in TFETs by extending the existing Kane model to non-uniform fields and heterostructures. We have studied BTBT in graphene, germanium and direct semiconductors with different bandgaps. We have developed a formalism capable of describing phonon-assisted BTBT in a general structure and a one-dimensional and a full two-dimensional quantum modeling of indirect tunneling in a double gate TFET were demonstrated.

Taking quantum effects, and quantum confinement in particular, into account was shown to be very important for TFET modeling. When making realistic predictions and comparing different TFET structures, quantum effects always have to be accounted for and we have shown it is possible to perform a two-dimensional quantum mechanical simulation of BTBT. Quantum effects often lead to worsened predictions compared to the too optimistic semiclassical results. For now, the TFET structures we have modeled are not yet able to go into competition with scaled MOSFET devices.

## 14.2 Suggestions for Future Research

### 14.2.1 TFET modeling

#### The quest for the ultimate TFET configuration

In the thesis, we have discussed point and line tunneling in TFETs, TFETs with different materials and heterostructures. Nevertheless, the question which of these configuration results in the best TFET characteristics remains an open question.

#### Quantitative experimental verification

While preparing this thesis, we have been in contact with people working on the fabrication of TFETs and diodes. At the same time, there is some literature on experimental devices [1, 2]. Nevertheless, we have published very little on a quantitative experimental verification of our models.

The main difficulty with confirming tunneling with experiments lies with the need for exact device structures to be able to confirm, rather than to fit, a model. I believe confirming and extracting a parameter such as the electron-phonon interaction strength from experiments is a very daunting task.

### 14.2.2 BTBT modeling

#### Two-dimensional tunneling in direct semiconductors: rigorous formalism

In chapter 9, we have introduced the general theory capable of calculating BTBT in an indirect semiconductor and in chapter 12 we have applied the theory to a two-dimensional device structure.

In an indirect semiconductor, the transport is mediated by the electron-phonon interaction and the conduction and valence band Schrödinger equation can be solved separately. In a direct semiconductor, the separation of the valence and the conduction band is not as straightforward because the coupling of the valence and conduction band is both responsible for the curvature and the interaction of the bands.

The use of ballistic NEGF [3] or a solution of the wavefunction with transmitting boundaries [4] employing a bandstructure capable of describing the valence and the conduction band and their interaction is a possibility to calculate BTBT in direct semiconductors. Nevertheless, it would be interesting to have a formalism similar to that developed in chapter 9 for direct semiconductors.

### Improved band structure description

In this thesis we used a description of the bandstructure with a small number of bands, the effective mass approximation in the case of indirect semiconductors and the 2 band  $\mathbf{k}\cdot\mathbf{p}$  model for the case of direct semiconductors.

The extension to a description with more bands [5–7] will result in more accurate results than those described in the thesis. At the same time, taking more bands into account will mean that the computational burden will increase.

### Discrete dopants and surface roughness

In our work, we have always assumed that the specified doping concentration introduces a corresponding background charge which has to be compensated by the mobile carriers in the semiconductor. In reality, the concentration of doping refers to the probability of finding a dopant in a given volume. And the operation of a semiconductor device depends on the exact location of the doping atoms.

When considering small volumes and when small changes in the doping position are important, the effects of the statistics on the dopants must be investigated. One possible approach is to generate a set of dopant distributions and then calculate the current for each dopant distribution [8].

A second statistical effect is the lack of smoothness of the gate-channel interface also known as surface roughness. The effect is well-known in MOSFETs but has not been studied yet in TFETs.

### Trap-assisted tunneling

For the on-current of the TFET, we have only considered the BTBT which describes the tunneling from valence to conduction band. However, in the presence of traps, current can also flow because of trap-assisted tunneling. In an indirect semiconductor, the interaction with traps can replace the electron-phonon interaction in providing the momentum difference between valence and conduction band. An alternative process of trap-assisted tunneling is hopping transport where the electron first tunnels from the valence band to the trap and at some later time, the electron tunnels from the trap to the conduction band.

## 14.2.3 General quantum transport

### Microscopic determination of the electron-phonon interaction strength

The electron-phonon interaction strengths we have used when describing the electron-phonon interaction in indirect semiconductors is based on very crude

assumptions. A better approach would be to calculate the electron-phonon interaction in a more rigorous way, for example starting from the rigid-ion approximation [9].

### **Developing an efficient method to calculate electron transport in nanostructures**

The study of transport in nanostructures is in many ways different from transport in microstructures. For this reason, former modeling techniques no longer apply today and new ones have to be developed. Unfortunately, resolving the evolution of a general many-body system is computationally impossible and we have to resort to approximations. Furthermore, contrary to equilibrium systems where the use of Green's functions is common place, the description of non-equilibrium quantum ensembles is not well-established.

### **Incorporating more many-body effects**

In our thesis we have introduced a number of many-body effects such as the electron-phonon interaction to describe BTBT in indirect semiconductors and the self-consistent solution of the Poisson and Schrödinger equation.

Solving the potential and charge self-consistently is known as the Hartree approximation and is only a first step in incorporating many-body effects. A first possible improvement on the Hartree approximation would be the introduction of the exchange interaction to solve the problem in the Hartree-Fock approximation.

## **Bibliography**

- [1] R. A. Logan and A. G. Chynoweth, "Effect of degenerate semiconductor band structure on current-voltage characteristics of silicon tunnel diodes," *Phys. Rev.*, vol. 131, pp. 89–95, Jul 1963. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRev.131.89>
- [2] 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*, vol. 95, no. 10, pp. 5800–5812, 2004. [Online]. Available: <http://link.aip.org/link/?JAP/95/5800/1>
- [3] D. Mamaluy, M. Sabathil, and P. Vogl, "Efficient method for the calculation of ballistic quantum transport," *Journal of Applied Physics*, vol. 93, no. 8, pp. 4628–4633, 2003. [Online]. Available: <http://link.aip.org/link/?JAP/93/4628/1>

- [4] C. S. Lent and D. J. Kirkner, "The quantum transmitting boundary method," *Journal of Applied Physics*, vol. 67, no. 10, pp. 6353–6359, 1990. [Online]. Available: <http://link.aip.org/link/?JAP/67/6353/1>
- [5] J. R. Chelikowsky and M. L. Cohen, "Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors," *Phys. Rev. B*, vol. 14, pp. 556–582, Jul 1976. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevB.14.556>
- [6] M. Cardona and F. H. Pollak, "Energy-Band Structure of Germanium and Silicon: The  $k$  [MIDDLE DOT]  $p$  Method," *Phys. Rev.*, vol. 142, pp. 530–543, 1966.
- [7] P. Vogl, H. P. Hjalmarson, and J. D. Dow, "A semi-empirical tight-binding theory of the electronic structure of semiconductors?" *Journal of Physics and Chemistry of Solids*, vol. 44, no. 5, pp. 365 – 378, 1983. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0022369783900641>
- [8] A. Asenov, "Random dopant induced threshold voltage lowering and fluctuations in sub-0.1  $\mu\text{m}$  mosfet's: A 3-d ldquo;atomistic rdquo; simulation study," *Electron Devices, IEEE Transactions on*, vol. 45, no. 12, pp. 2505 –2513, dec 1998.
- [9] J. M. Ziman, *Electrons and phonons : the theory of transport phenomena in solids*. Oxford : Clarendon, 1979.

# Publication List

## International Journal Papers

- [1] **W. Vandenberghe**, B. Sorée, W. Magnus, and G. Groeseneken. Zener tunneling in semiconductors under nonuniform electric fields. *Journal of Applied Physics*, 107(5):054520, 2010.
- [2] **W. G. Vandenberghe**, B. Sorée, W. Magnus, G. Groeseneken, and M. V. Fischetti. Impact of field-induced quantum confinement in tunneling field-effect devices. *Applied Physics Letters*, 98(14):143503, 2011.
- [3] **W. Vandenberghe**, B. Sorée, W. Magnus, and M. V. Fischetti. Generalized phonon-assisted zener tunneling in indirect semiconductors with non-uniform electric fields: A rigorous approach. *Journal of Applied Physics*, 109(12):124503, jun 2011.
- [4] **W. G. Vandenberghe**, A. S. Verhulst, K.-H. Kao, K. De Meyer, B. Sorée, W. Magnus, and G. Groeseneken. A model determining optimal doping concentration and material's band gap of tunnel field-effect transistors. *Applied Physics Letters*, 100:193509, 2012.
- [5] A. Verhulst, **W. Vandenberghe**, K. Maex, and G. Groeseneken. Tunnel field-effect transistor without gate-drain overlap. *Applied Physics Letters*, 91(5):053102, 2007.
- [6] A. Verhulst, **W. Vandenberghe**, K. Maex, S. De Gendt, M. Heyns, and G. Groeseneken. Complementary silicon-based heterostructure tunnel-fets with high tunnel rates. *Electron Device Letters, IEEE DOI - 10.1109/LED.2008.2007599*, 29(12):1398–1401, 2008.
- [7] A. Verhulst, **W. Vandenberghe**, K. Maex, and G. Groeseneken. Boosting the on-current of a n-channel nanowire tunnel field-effect transistor by source material optimization. *Journal of Applied Physics*, 104(6):064514, 2008.

- [8] A. Verhulst, B. Sorée, D. Leonelli, **W. Vandenberghe**, and G. Groeseneken. Modeling the single-gate, double-gate, and gate-all-around tunnel field-effect transistor. *Journal of Applied Physics*, 107(2):024518, 2010.
- [9] J. Zhuge, A. S. Verhulst, **W. G. Vandenberghe**, W. Dehaene, R. Huang, Y. Wang, and G. Groeseneken. Digital-circuit analysis of short-gate tunnel fets for low-voltage applications. *Semiconductor Science and Technology*, 26(8):085001, 2011.
- [10] B. Sorée, W. Magnus, and **W. Vandenberghe**. Low-field mobility in ultrathin silicon nanowire junctionless transistors. *Applied Physics Letters*, 99(23):233509, 2011.
- [11] K.-H. Kao, A. S. Verhulst, **W. G. Vandenberghe**, B. Sorée, G. Groeseneken, and K. De Meyer. Direct and indirect band-to-band tunneling in germanium-based tfets. *Electron Devices, IEEE Transactions on*, 59(2):292–301, feb. 2012.
- [12] K.H. Kao, A. S. Verhulst, **W. G. Vandenberghe**, B. Sorée, G. Groeseneken, and K. De Meyer. Modeling the impact of junction angles in tunnel field-effect transistors. *Solid-State Electronics*, Volume 69, March 2012, Pages 31-37.
- [13] K.H. Kao, A. S. Verhulst, **W. G. Vandenberghe**, B. Sorée, G. Groeseneken, and K. De Meyer. Optimization of Gate-on-Source-Only Tunnel FETs With Counter-Doped Pockets. *Electron Devices, IEEE Transactions on*, Accepted.
- [14] K.H. Kao, A. S. Verhulst, **W. G. Vandenberghe**, and K. De Meyer. Counter-Doped Pocket Thickness Optimization of Gate-on-Source-only Tunnel FETs. *Electron Devices, IEEE Transactions on*, Submitted.
- [15] H. Carrillo Nuñez, W. Magnus, **W. G. Vandenberghe**, B. Sorée, and F. Peeters. Band-to-Band Tunneling in the Framework of the Modified Local Density Approximation. Submitted.
- [16] H. Carrillo Nuñez, W. Magnus, **W. G. Vandenberghe**, B. Sorée, and F. Peeters. Phonon-assisted Zener Tunneling in a p-n Diode Silicon Nanowire. In preparation.

## International Conference Papers

- [1] **W. Vandenberghe**, A. Verhulst, G. Groeseneken, B. Sorée, and W. Magnus. Analytical model for a tunnel field-effect transistor. In *Proceedings of the 14th IEEE Mediterranean Electrotechnical Conference - MELECON*, pages 923–928, Ajaccio (France), May 2008.

- [2] **W. Vandenberghe**, A.S. Verhulst, G. Groeseneken, B. Sorée, and W. Magnus. Analytical model for point and line tunneling in a tunnel field-effect transistor. In *Simulation of Semiconductor Processes and Devices, 2008. SISPAD 2008. International Conference on*, pages 137–140, sept. 2008.
- [3] **W. Vandenberghe**, B. Sorée, W. Magnus, and G. Groeseneken. Zener tunneling in graphene based semiconductors - the k.p method. In *Proceedings of EDISON 16: 16th International Conference on Electron Dynamics in Semiconductors, Optoelectronics and Nanostructures*, pp. 012111, Montpellier (France), August 2009.
- [4] **W.G. Vandenberghe**, B. Sorée, W. Magnus, G. Groeseneken, A.S. Verhulst, and M.V. Fischetti. Field induced quantum confinement in indirect semiconductors: Quantum mechanical and modified semiclassical model. In *Simulation of Semiconductor Processes and Devices (SISPAD), 2011 International Conference on*, pages 271–274, sept. 2011.
- [5] **W. G. Vandenberghe**, B. Sorée, W. Magnus, M. V. Fischetti, A. S. Verhulst, and G. Groeseneken. Two-dimensional quantum mechanical modeling of band-to-band tunneling in indirect semiconductors. In *Electron Devices Meeting (IEDM), 2011 IEEE International*, pages 5.3.1–5.3.4, dec. 2011.
- [6] A. Verhulst, **W. Vandenberghe**, D. Leonelli, F. Kao, R. Rooyackers, A. Vandooren, J. Zhuge, B. Sorée, W. Magnus, M. Fischetti, G. Pourtois, R. Huang, Y. Wang, K. De Meyer, W. Dehaene, C. Huyghebaert, M. Heyns, and G. Groeseneken. Tunnel field-effect transistors for low-power nano-electronics. In *Tsukuba Nanotechnology Symposium - TNS*, 2009.
- [7] A. Verhulst, **W. Vandenberghe**, S. De Gendt, M. Heyns, and G. Groeseneken. Tunnel-fets for future low-power nano-electronics. In *Proceedings of the NODE Dissemination Workshop*, 2009.
- [8] A. Verhulst, **W. Vandenberghe**, D. Leonelli, R. Rooyackers, A. Vandooren, S. De Gendt, M. Heyns, and G. Groeseneken. Tunnel field-effect transistors for future low-power nano-electronics. In *ECS Transactions*, volume 25 of 7, pages 455–462, Vienna, October 2009.
- [9] A. Verhulst, **W. Vandenberghe**, S. De Gendt, K. Maex, and G. Groeseneken. Boosting the on-current of silicon nanowire tunnel-fets. In *ECS Transactions*, volume 33 of 6, pages 363–372, 2010.
- [10] A. Verhulst, **W. Vandenberghe**, D. Leonelli, R. Rooyackers, A. Vandooren, J. Zhuge, F. Kao, B. Sorée, W. Magnus, M. Fischetti, G. Pourtois, C. Huyghebaert, R. Huang, Y. Wang, K. De Meyer, W. Dehaene,

- M. Heyns, and G. Groeseneken. Si-based tunnel field-effect transistors for low-power nano-electronics. In *69th Device Research Conference*, 2011.
- [11] A. Vandooren, R. Rooyackers, D. Leonelli, F. Iacopi, S. De Gendt, A. Verhulst, M. Heyns, E. Kunnen, D. Nguyen, M. Demand, P. Ong, W. Lee, J. Moonens, O. Richard, **W. Vandenberghe**, and G. Groeseneken. A 35nm diameter vertical silicon nanowire short-gate tunnelfet. In *Nanotechnology Workshop*, 2009.
- [12] M. Heyns, F. Bellenger, G. Brammertz, M. Caymax, S. De Gendt, B. De Jaeger, A. Delabie, G. Eneman, G. Groeseneken, M. Houssa, D. Leonelli, D. Lin, K. Martens, C. Merckling, M. Meuris, J. Mitard, J. Penaud, G. Pourtois, M. Scarrozza, E. Simoen, S. Van Elshocht, **W. Vandenberghe**, A. Vandooren, A. Verhulst, and W. Wang. High mobility channel materials and novel devices for scaling of nanoelectronics beyond the si roadmap. In *MRS Fall Meeting Symposium A: High-k Dielectrics on Semiconductors with High Carrier Mobility*, 2009.
- [13] M. Heyns, F. Bellenger, G. Brammertz, M. Caymax, M. Cantoro, S. De Gendt, B. De Jaeger, A. Delabie, G. Eneman, G. Groeseneken, G. Hellings, M. Houssa, F. Iacopi, D. Leonelli, D. Lin, W. Magnus, K. Martens, C. Merckling, M. Meuris, J. Mitard, J. Penaud, G. Pourtois, M. Scarrozza, E. Simoen, B. Sorée, S. Van Elshocht, **W. Vandenberghe**, A. Vandooren, P. Vereecken, A. Verhulst, and W. Wang. Shaping the future of nanoelectronics beyond the si roadmap with new materials and devices. In *Optical Microlithography XXIII*, 2010.
- [14] B. Sorée, W. Magnus, M. Szepieniec, **W. Vandenberghe**, A. Verhulst, G. Pourtois, G. Groeseneken, S. De Gendt, and M. Heyns. Novel device concepts for nanotechnology: the nanowire pinch-off fet and graphene tunnel fet. In *ECS Transactions*, volume 28 of 5, pages 15–26, 2010.
- [15] F. Kao, A. Verhulst, **W. Vandenberghe**, G. Groeseneken, and K. De Meyer. The impact of junction angle on tunnel fets. In *Proceedings of the 12th International Conference on Ultimate Integration on Silicon - ULIS*, 2011.
- [16] J. Zhuge, A. Verhulst, **W. Vandenberghe**, W. Dehaene, R. Huang, Y. Wang, and G. Groeseneken. Short gate tunnel-fets for low voltage digital applications. In *China Semiconductor Technology International Conference - CSTIC*, 2011.
- [17] H. Carrillo Nuñez, W. Magnus, **W. G. Vandenberghe**, B. Sorée, and F. Peeters. On the calculation of phonon-assisted band-to-band tunneling currents in Si nanowires. In *IWCE*, 2012.

## International Patents

- [1] A. S. Verhulst, **W. G. Vandenberghe**. TUNNEL EFFECT TRANSISTORS BASED ON ELONGATE MONOCRYSTALLINE NANOSTRUCTURES HAVING A HETEROSTRUCTURE. *US Patent App.* 20,080/067,607 (2007), *EP Patent* 1,901,355 (2008).
- [2] **W. G. Vandenberghe**, A. S. Verhulst. TUNNEL FIELD-EFFECT TRANSISTOR WITH GATED TUNNEL BARRIER. *US Patent* 20,120,115,296 (2012), *EP Patent* 1,901,354 (2008).



# Curriculum Vitae



William Vandenberghe was born in Tielt, Belgium in 1985. He received the M.Sc. degree (magna cum laude) in Electrical Engineering from the Katholieke Universiteit Leuven (K.U. Leuven), Belgium in 2007. Since 2007, he was pursuing the Ph.D. degree at the Department of Electrical Engineering at the K.U. Leuven under the supervision of prof. Guido Groeseneken (K.U. Leuven) and prof. Wim Magnus (Universiteit Antwerpen). In 2010, he was a visiting researcher at the University of Texas at Dallas under the supervision of prof. Massimo Fischetti.

He was laureate of the Flemish Chemistry Olympiad (2001, 2002), the Flemish Physics Olympiad (2002) and the Flemish Mathematics Olympiad (2001, 2002) and was member of the Belgian International Mathematical Olympiad team in 2002. In 2007, he received the AMI semiconductor thesis prize. In 2008, he received the second prize in the IEEE Region 8 Student Paper Contest. From 2008 to 2012, he received a Ph.D. grant from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). He has authored or co-authored over 30 publications in international journals and conference proceedings.

His research interest lies in the modeling of nano-electronic devices. More specifically, he investigates the theory of electron transport starting from the laws of quantum mechanics, the numerical methods involved in modeling nano-electronic devices and he studies realistic device structures using theoretical modeling. During his Ph.D., he focused on the study of electronic transport in tunnel field-effect transistors.