

Quantum-mechanical modeling towards trap-assisted tunneling in semiconductor devices

Mazharuddin Mohammed

Supervisor:
Prof. dr. ir. Guido Groeseneken

Dissertation presented in partial
fulfillment of the requirements for the
degree of Doctor of Engineering
Science (PhD): Electrical Engineering

August 2018

Quantum-mechanical modeling towards trap-assisted tunneling in semiconductor devices

Mazharuddin MOHAMMED

Examination committee:

Prof. dr. ir. Jean Berlamont, chair
Prof. dr. ir. Guido Groeseneken, supervisor
Prof. dr. ir. Wim Dehaene
Prof. dr. ir. Bart Sorée
Prof. dr. ir. Michel Houssa
Dr. ir. Anne S. Verhulst, daily advisor
Dr. ir. Robin Degraeve

Prof. dr. ir. William G. Vandenberghe
(The University of Texas at Dallas, USA)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Electrical Engineering



August 2018

© 2018 KU Leuven – Faculty of Engineering Science
Uitgegeven in eigen beheer, Mazharuddin Mohammed, Kasteelpark Arenberg 10 box 2440, B-3001 Leuven
(Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag worden 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.

Preface

I would like to take an opportunity to thank all people who helped and supported me during my PhD from June 2014 till August 2018 at imec and KU Leuven. I must admit that it was my greatest chance that I was able to pursue my PhD at one of the renowned and prestigious organizations in the world. It gave me the greatest opportunities in my career to improve my personal and technical capabilities. I learned the discipline and self-confidence by being a researcher at TCAD group in imec and learned how the most famous researchers in semiconductor industry develop novel concepts. At the same time I enjoyed being a student at ESAT, KU Leuven.

First and foremost, I would like to thank Prof. Guido Groeseneken, my promoter, for his unwavering support and for giving me the opportunity to pursue PhD under his supervision. Furthermore, I would like to express my sincere gratitude to Dr. Anne S. Verhulst, my daily advisor, who is a tremendous mentor. I sincerely thank her for providing motivating, encouraging, and enthusiastic environment during my PhD. It was her steady support and inspiring guidance that brought this work to function. It was a great pleasure for me to conduct my work under her supervision. I thank her especially for thoroughly checking my weird derivations, revising the manuscripts and adding value to its content. I am grateful to have had the opportunity to work under her guidance and direction. I will always reflect on the numerous discussions I had with her.

I would like to thank Prof. Bart Sorée for familiarizing me with the ingredients to derive the formalism. Many thanks also for proofreading the manuscripts and for all the valuable scientific discussions. I am also thankful to Dr. Robin Degraeve and Dr. Ben Kaczer for always strongly supporting my work and providing constructive feedback. I am also grateful to Prof. Wim Magnus, Dr. Maarten L. Van de Put and Dr. Devin Verreck for their valuable advises during the derivations and helping me in making a sense out of the corresponding fields. Furthermore, I would like to thank all the Jury members for critically reviewing the thesis. I would also like to thank Prof. Jean Berlamont for agreeing to chair the examination committee. I am also

grateful to Prof. Wim Dehaene, Prof. Michel Houssa and Prof. William G. Vandenberghe for helpful comments and for agreeing to be the jury members in my examination committee.

I would also like to extend my sincere gratitude to the current and former members of the TCAD group and management for the great collaboration and the pleasant work environment: Alireza Alian, An De Keersgieter, Anda Mocuta, Aaron Thean, Anne Vandooren, Arturo Sibaja-Hernandez, Ashish Dabral, Eddy Simoen, Fabian Bufler, Geoffrey Pourtois, Geert Eneman, Jasper Bizindavyi, Nadine Collaert, Philippe Matagne, Quentin Smets and many others. I am also thankful to Betty Dennis, Myriam Janowski and Claire Declerck for arranging all kinds of administrative practicalities. I would also like to thank the friends and colleagues from imec: Goutham Arutchelvan, Subhali Subhechha, Siva Ramesh, Tarun Kumar Agarwal, Abhitosh Vais, Amey Walke, Abhinav Gaur, Yashwant Balaji, Naga Sruti Avasarala, Surya Gurunarayanan, Shibesh Dutta, Finub James Shirley and many others for being good company and having interesting conversations about science, religion, politics, economic policies, agriculture, etc. I am also grateful to my friends from Germany: Abdur Rehman Jalil, Senan Aube, Mohamad Khodashenaz and many others for their support and guidance.

Last but certainly the most important, on the personal side, I am indebted to my wife, Asiya. I can not thank her enough for the unconditional support. My parents deserve special thanks, for their unwavering love, support and encouragement. In the end, I would like to dedicate this work to my sister Sofia Tabassum, graduated in Mathematics, whom we lost in 2011, and who has inspired/encouraged me to pursue my PhD.

Mazharuddin Mohammed
Leuven, August 2018.

Abstract

For the last few decades, the persistent downscaling of the metal oxide semiconductor field-effect transistor (MOSFET) has been the main reason for the enormous increase in the computational performance of everyday electronics. Currently, the fundamental 60 mV/decade subthreshold swing (SS) limits the downscaling of the MOSFET supply voltage which in turn makes it difficult to further downscale the dimensions, because the power density of integrated circuits can not be kept constant or lowered. Of all novel solutions to replace the MOSFET, the tunnel field-effect transistor (TFET) promises to surmount the SS limitation, because of its inherent band-to-band tunneling (BTBT) mechanism for carrier transport. However, there is a substantial discrepancy in theoretical predictions and experimental results of TFET performance. In particular, the latter exhibits a degraded SS and considers trap-assisted tunneling (TAT) as one of the probable reasons for SS degradation. Simultaneously, the TAT increases the unwanted stress-induced leakage current and gate-induced drain leakage current in the MOSFET. Owing to such importance, the accurate modeling of TAT in semiconductor devices is essential to assess the device performance. We therefore present an approach towards an all-inclusive, fully quantum-mechanical simulator of TAT in semiconductor devices.

In the first part of the thesis, we investigate the electric field effects on the traps in semiconductors and oxides, with as aim to present the significance of quantum effects in determining TAT and to justify the need for a quantum-mechanical based TAT model. In particular, we determine the impact of high electric field on the characteristic trap energy level. We find that the traps in semiconductors exhibit a field-induced level shift and broadening at TFET-relevant electric fields. We further find that the field-effects of traps in oxides depend on the trap position, trap charge state, tunneling barrier type and the host material parameters. We study the impact of electric field on the TAT-relevant emission rates in semiconductors by including the field-induced level shift and broadening in the existing semi-classical emission rate formalism. To determine the implications of field-effects on the oxides emission rates, we formulate a semi-classical based emission rate based on the standard Wenzel-

Kramer-Brillouin approximation for the tunneling probability through a finite barrier. We find that the field-induced quantum effects can increase the emission rates of a semiconductor trap level at high electric field, while the impact for the oxide traps is much smaller due to the higher effective mass and barrier heights.

Since the fundamental mechanism underlying TAT in a semiconductor device is phonon-assisted tunneling (PAT), we present a formalism to calculate the PAT current for bi-dimensional semiconductor devices. We formulate the PAT current equation by revising an existing approach of calculating Zener tunneling in indirect-bandgap devices. We apply a quantum transmitting boundary method approximation and calculate the electron-phonon coupling (EPC) strength for homostructure and heterostructure devices, while using the envelope function approximation of electron wavefunctions. To allow for a faster calculation of PAT current, we use the physics based lattice approach and apply the low phonon wavevector approximation to eliminate the basis functions in the overlap integrals of the EPC strengths.

To allow for simulations of the PAT current in homostructure III-V devices, we implement the PAT current formalism in an existing direct-BTBT simulator (Pharos) and study the PAT in $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diodes with a 2 and 15-band implementation of the formalism. We find that there is a limited difference for the calculated PAT current densities between the 2 and 15-band model of our formalism. We discover that the electron-phonon coupling strength is inefficient across the tunneling junction due to the phase-shift between the envelope functions injected from bands corresponding to basis function with different parity. We find that the PAT current densities depend on the minimum allowed phonon wavevector and on the device length along the transport direction because of the PAT contributions in the near-tunneling regions. We also find that the combined effects of the near-tunneling region length and the effective tunneling-energy interval determine the observed doping dependence of the PAT current density. Finally, we present the future framework to calculate multiphonon assisted tunneling to enable the accurate predictions of TAT currents in semiconductor devices.

Beknopte samenvatting

De aanhoudende schaling van de metaal-oxide-halfgeleider-veldeffecttransistor (MOSFET) is de laatste decennia de belangrijkste reden geweest voor de enorme toename van de computationele prestaties van alledaagse elektronica. Op dit moment beperkt de fundamentele 60mV/decade subthreshold swing (SS) de verlaging van de MOSFET voedingsspanning, wat het op zijn beurt moeilijk maakt om de dimensies verder te verkleinen omdat de vermogensdichtheid van geïntegreerde circuits niet kan worden constant gehouden of verlaagd. Van alle nieuwe oplossingen om de MOSFET te vervangen, belooft de tunnel-veldeffecttransistor (TFET) de SS-beperking te overwinnen, vanwege het inherente band-tot-band tunneling (BTBT) mechanisme voor ladingsdrager transport. Er is echter een aanzienlijke discrepantie tussen theoretische voorspellingen en experimentele resultaten van TFET prestaties. In het bijzonder vertoont de laatste een gedegradeerde SS en defect-geassisteerd tunnelen (TAT) wordt als een van de meest waarschijnlijke redenen voor SS degradatie beschouwd. Tegelijkertijd verhoogt TAT de ongewenste stress-geïnduceerde lekstroom en gate-geïnduceerde drain lekstroom in de MOSFET. Vanwege dit belang is het nauwkeurig modelleren van TAT in halfgeleider componenten essentieel voor het beoordelen van de prestaties van de component. We presenteren daarom een eerste stap naar een allesomvattende, volledig kwantummechanische simulator van TAT in halfgeleider componenten.

In het eerste deel van het proefschrift onderzoeken we de elektrische veldeffecten op defecten in halfgeleiders en oxides, met als doel de impact van kwantumeffecten te presenteren bij het bepalen van TAT en om de noodzaak van een kwantummechanisch gebaseerd TAT-model te rechtvaardigen. In het bijzonder bepalen we de impact van een hoog elektrisch veld op het karakteristieke defect-energieniveau. We merken dat de defecten in halfgeleiders een veldgeïnduceerde niveauverschuiving en verbreding vertonen bij voor TFET relevante elektrische velden. We vinden verder dat de veldeffecten van defecten in oxides afhankelijk zijn van de defect positie, defect ladingstoestand, tunnelingbarrière en de parameters van het halfgeleider materiaal. We bestuderen de impact van het elektrisch veld op de

TAT-relevante emissiesnelheden in halfgeleiders door de veldgeïnduceerde niveaoverschuiving en verbreding in het bestaande semiklassieke emissiesnelheid formalisme op te nemen. Om de implicaties van veldeffecten op de oxide emissiesnelheden te bepalen, formuleren we een semiklassiek gebaseerde emissiesnelheid op basis van de standaard Wenzel-Kramer-Brillouinbenadering voor de tunnelwaarschijnlijkheid door een eindige barrière. We constateren dat de veldgeïnduceerde quantumeffecten de emissiesnelheden van een halfgeleider defect niveau bij een hoog elektrisch veld kunnen verhogen, terwijl de impact voor de oxide defecten veel kleiner is vanwege de hogere effectieve massa en barrièrehoogtes.

Aangezien het fundamentele mechanisme dat ten grondslag ligt aan TAT in een halfgeleider component fonon-geassisteerde tunneling (PAT) is, presenteren we een veralgemeend formalisme om de PAT-stroom voor bi-dimensionale halfgeleider componenten te berekenen. We formuleren de PAT-stroomvergelijking door een bestaand formalisme voor het berekenen van Zener-tunneling in indirecte-bandgap-componenten te herzien. We passen een quantum-transmitterende grensmethodebenadering toe en berekenen de elektron-fononkoppeling(EPC) sterkte voor homostructuur en heterostructuur componenten, terwijl de enveloppe functie benadering voor elektronengolffuncties wordt gebruikt. Om een snellere berekening van de PAT-stroom mogelijk te maken, gebruiken we de op fysica gebaseerde rooster benadering en passen we de lage fonon golfvectorbenadering toe om de basisfuncties in de overlap-integralen van de EPC-sterktes te elimineren.

Om simulaties van de PAT-stroom in homostructuur III-V-componenten mogelijk te maken, implementeren we het PAT-stroom formalisme in een bestaande directe BTBT-simulator (Pharos) en bestuderen we de PAT in $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ pn-diodes met een 2- en 15-band implementatie van het formalisme. We vinden dat er een beperkt verschil is voor de berekende PAT-stroomdichtheden tussen het 2- en 15-banden model van ons formalisme. We ontdekken dat de elektronen-fononkoppeling sterkte inefficiënt is over de tunneling-junctie als gevolg van de faseverschuiving tussen de enveloppe functies die worden geïnjecteerd uit banden die overeenkomen met basisfunctie met verschillende pariteit. We vinden dat de PAT-stroomdichtheden afhankelijk zijn van de minimaal toegestane fonon-golfvector en van de lengte van de component langs de transportrichting vanwege de PAT-bijdragen in de nabij-tunnelinggebieden. We vinden ook dat de gecombineerde effecten van de lengte van het nabije tunnelinggebied en het effectieve tunneling-energie-interval de waargenomen doperingsafhankelijkheid van de PAT-stroomdichtheid bepalen. Ten slotte presenteren we het toekomstige raamwerk voor het berekenen van door meerdere fotonen ondersteunde tunneling om de voorspellingen van TAT-stromen in halfgeleider componenten mogelijk te maken.

List of Abbreviations

Acronym	Description
BTBT	Band-to-Band Tunneling
CBM	Conduction Band Minimum
CC	Configuration Coordinate
CMOS	Complementary Metal oxide Semiconductor
CW	Coulomb Well (trap potential)
DFPT	Density Functional Perturbation Theory
DFT	Density Functional Theory
DLTS	Deep Level Transient Spectroscopy
EF	Envelope Function
EMA	Effective Mass Approximation
EPC	Electron-Phonon Coupling
FWHM	Full-Width Half-Maximum
GIDL	Gate-Induced Drain Leakage
IC	Integrated Circuit
IVP	Initial Value Problem
LDOS	Local Density of States
LRME	Lattice Relaxation Multiphonon Emission
MOSCAP	Metal-Oxide-Semiconductor-Capacitor
MOSFET	Metal Oxide Semiconductor Field-Effect Transistor

Acronym	Description
MTM	Modified Transfer Matrix (method)
NBTI	Negative Bias Temperature Instability
NEGF	Non-Equilibrium Green Function
NEMFET	Nano-Electro-Mechanical Field-Effect Transistor
NMP	Non-radiative Multiphonon
PAT	Phonon-Assisted Tunneling
QM	Quantum-Mechanical
QTBM	Quantum Transmitting Boundary Method
RTN	Random Telegraph Noise
SCBA	Self-Consistent Born Approximation
SILC	Stress-Induced Leakage Currents
SRH	Shockley-Read-Hall
SS	Subthreshold Swing
SW	Square Well (trap potential)
TAT	Trap-Assisted Tunneling
TDDS	Time Dependent Defect Spectroscopy
TFET	Tunnel Field-Effect Transistor
WKB	Wenzel-Kramer-Brillouin

List of Symbols

The list describes several symbols that will be later used in the thesis

Ai	Airy function
$\langle e_{nF} \rangle$	Average field-dependent electron emission rate
$\langle v_{th} \rangle$	Average thermal energy
$E_{c,v}$	Band edge energy of conduction, valence band
U_n	Basis function corresponding to band n
U_n^l	Basis function of layer l corresponding to band n
$S_n^{k \rightarrow l}$	Basis function transformation of layer k to layer l
$f_{c/t}^{\mathbf{k}_\perp}$	Bloch theory envelope function characterizing electrons in the conduction band, trap
k_B	Boltzmann constant
v	Bose-Einstein distribution function
μ_{ph}	Chemical potential of the free phonons system
$\hat{c}_j^\dagger, \hat{c}_j$	Creation, annihilation operator for an electron in the j^{th} -state
$\hat{a}_{\mathbf{q}}^\dagger, \hat{a}_{\mathbf{q}}$	Creation, annihilation operator of a phonon
V_c	Crystal lattice potential energy
$\hat{\varrho}_0$	Density matrix operator for the unperturbed system of electrons and phonons
$\hat{\varrho}_{\alpha,\beta}$	Density matrix operator of electrons in the left, right contact
$\hat{\varrho}_{el}$	Density matrix operator of electrons
$\hat{\varrho}_{ph}$	Density matrix operator of phonons
$\hat{\varrho}$	Density matrix operator
C_{dep}	Depletion layer capacitance
I_{DS}	Drain-source current
N_{DOS}^C	Effective density of state in the conduction band
ω_0	Effective phonon frequency
F	Electric-field strength
\mathbf{F}	Electric-field
$\mu_{\alpha,\beta}$	Electro-chemical potential in the left, right contact

μ	Electro-chemical potential
n_t, p_t	Electron and hole densities in the trap region
c_n	Electron capture rate
$m_{e,l}^*$	Electron effective mass of layer l
m_e^*	Electron effective mass
e_n	Electron emission rate
$\hat{\psi}$	Electron field operator
$\hat{N}_{\alpha,\beta}$	Electron Number operator for the left, right contact
\hat{N}_{el}	Electron Number operator
\mathbf{k}_{\perp}	Electron wavevector along y, z -direction
\mathbf{k}	Electron wavevector
$g_{\alpha\beta}^{k \rightarrow l}$	Electron-phonon coupling strength for heterostructures
$g_{\alpha\beta\mathbf{q}\nu}$	Electron-phonon coupling strength for homostructures
$\hat{\mathcal{H}}_{\text{int}}$	Electron-phonon interaction Hamiltonian operator
$\phi_{\alpha\beta\mathbf{q}}^n$	Electron-phonon overlap function
e_q	Elementary charge
E	Energy
$F_{\alpha n}$	Envelope function corresponding to state α and band n
$F_{\alpha n}^l$	Envelope function of layer l corresponding to state α and band n
f_{FD}	Fermi-Dirac distribution function
E_i^{F}	Field dependent energetic distance of trap energy level w.r.t CBM
c_{nF}	Field-dependent electron capture rate
e_{nF}	Field-dependent electron emission rate
c_{pF}	Field-dependent hole capture rate
e_{pF}	Field-dependent hole emission rate
$\Delta\mathcal{V}$	First order electron-lattice interaction potential
$\hat{\mathcal{H}}_{\text{ph}}$	Free phonons Hamiltonian operator
m_0	Free-electron mass
V_{GS}	Gate-source voltage
S	Gibbs entropy of a system
$\mathcal{Z}_{\alpha,\beta}$	Grand canonical partition function of electrons in the left, right contact
\mathcal{Z}_{el}	Grand canonical partition function of electrons
\mathcal{Z}_{ph}	Grand canonical partition function of free phonons
$\hat{\mathcal{H}}_0$	Hamiltonian operator for the unperturbed system of electrons and phonons
ε_{∞}	High frequency dielectric constant of the host material
c_p	Hole capture rate
e_p	Hole emission rate
S	Huang-Rhys factor
ν	Index specifying the longitudinal-transversal branch as well as the acoustic or optical phonons

$\tilde{\mathcal{H}}_{\text{int}}$	Interaction-picture electron-phonon interaction Hamiltonian operator
\mathcal{J}_{sp}	Jost matrix
\mathbf{R}	Lattice vector
β	Left-traveling electron state (in the right contact)
$L_{x,y,z}$	length along the x, y, z -direction
$E_{\text{t}}^{\text{L,U}}$	Lower, upper limit of trap energy level spectrum
$\hat{\mathcal{H}}_{\text{el}}$	Many-electrons Hamiltonian operator
Ψ	Many-electrons wavefunction
\mathcal{I}_p	Modified-Bessel function of order p
\mathcal{M}	Modified-transfer matrix of the total system
C_{ox}	Oxide capacitance
J_{ph}	PAT current density
I_{ph}	PAT current
\widehat{M}	Phonon Number operator
\mathbf{q}	Phonon wavevector
$\epsilon_{\mathbf{q}\nu}$	Polarization vector
\mathbf{r}	Position vector
\mathbf{G}	Reciprocal lattice vector
\hbar	Reduced Planck constant
α	Right-traveling electron state (in the left contact)
$\Delta^2\mathcal{V}$	Second order electron-lattice interaction potential
$g_{\alpha\beta\mathbf{q}\nu\mathbf{q},\nu}$	Second order electron-phonon coupling strength
$\tilde{\mathcal{H}}_{\text{int}}^{(2)}$	Second order electron-phonon interaction Hamiltonian operator in interaction picture
$\widehat{\mathcal{H}}_{\text{int}}^{(2)}$	Second order electron-phonon interaction Hamiltonian operator
Ω	Semiconductor device volume
H_{el}	Single-electron Hamiltonian
\hat{n}_{ϑ}	Single-electron state ϑ occupation number operator
$\psi_{\alpha,\beta}$	Single-electron wavefunction of state α, β
ψ	Single-electron wavefunction
\hat{m}_{λ}	Single-phonon state λ occupation number operator
ϵ_0	Static dielectric constant
$\psi_n^{(0)}$	Stationary state electron wavefunction
p_{λ}	Statistical weight of lattice initial states
J_{TAT}	TAT current density
T	Temperature
V_{th}	Threshold voltage
N_T	Trap density
$E_{\text{t}}^{\text{Max}}$	Trap energy level w.r.t maximum probability density
V_{tr}	Trap potential
ψ_t	Trap state wavefunction
w_{if}	Trap-assisted transition rate

m_t^*	Tunneling mass
\mathbf{R}_u	Unit cell vector
Ω_c	Unit cell volume
Θ	Unit step-function
$k_{x,y,z}$	Wave number in x, y, z -direction
\mathcal{N}_d	Wavefunction normalization factor
w	Weight function
\mathcal{W}	Wronskian
σ_{n0}	Zero-field electron capture cross section
e_n^∞	Zero-field electron emission rate coefficient
e_{n0}	Zero-field electron emission rates
E_i	Zero-field energetic distance of trap energy level w.r.t CBM
E_t	Zero-field trap energy level

Contents

Abstract	iii
Beknopte samenvatting	v
List of Abbreviations	viii
List of Symbols	xii
Contents	xiii
List of Figures	xvii
1 Introduction	1
1.1 MOSFET and scaling limitation	2
1.2 TFET: A promising low-power device	3
1.3 Parasitics limiting the TFET performance	5
1.4 Literature review of trap-assisted tunneling in semiconductor devices	6
1.5 Modeling perspective of TAT	8
1.5.1 Thermally-assisted tunneling models	11
1.5.2 Structural relaxation based TAT models	13
1.5.3 Limitations of the models	20

1.6	Goal of the thesis	22
1.7	Organization of the thesis	23
2	Impact of high electric field on traps in semiconductors and oxides	25
2.1	Trap energy level extraction	26
2.1.1	Modified transfer matrix method	27
2.1.2	Wavefunction normalization	33
2.2	Numerical procedure	41
2.3	Electric field effects: Level shift and broadening	42
2.3.1	Semiconductor traps	42
2.3.2	Oxide traps	46
2.4	Impact of field-effects on the emission rate	51
2.5	Conclusions	59
3	Phonon-assisted tunneling current formalism	61
3.1	A many-particle description of electrons and phonons	62
3.1.1	System of non-interacting electrons	63
3.1.2	System of free phonons	67
3.2	Derivation of the PAT current equation	68
3.3	Electron-phonon coupling strength calculation	77
3.3.1	EPC strength for direct-bandgap homostructures	78
3.3.2	EPC strength for direct-bandgap heterostructures	83
3.4	PAT current density for a bi-dimensional-semiconductor structure	87
3.5	Numerical implementation	89
3.6	Conclusions	90
4	Application of PAT current formalism	91

4.1	Homostructure $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode	92
4.2	Comparison between the 2- and 15-band PAT current density	94
4.3	Inefficient electron-phonon coupling across the tunneling junction	95
4.4	PAT dependence on the device length	97
4.5	Doping dependence of PAT current density	99
4.6	Conclusions	102
5	Conclusions and outlook	105
5.1	Conclusions	105
5.2	Outlook	107
5.2.1	Impact on experimental work	107
5.2.2	Future research towards enhanced models	108
	Appendix A: Fundamentals of NMP theory	115
	Appendix B: Fundamentals of equilibrium statistical mechanics	121
B.1	Fermi-Dirac statistics	121
B.2	Bose-Einstein statistics	124
	Bibliography	127
	Curriculum Vitae	137
	List of publications	139

List of Figures

1.1	Schematic representations of (a) an n-MOSFET, in which orange and green colors show n-type and p-type semiconductors, respectively. (b) Fermi-Dirac statistics at the source region, (c) Band diagram at the indicated cross-section $z - z'$ of Fig. 1.1(a) and in steps from OFF-to-ON state, (d) the transfer-characteristics of the n-MOSFET. The dual-rounded orange arrows show the $I - V$ data points corresponding to high energy Fermi-tails, preventing the abrupt switching of the transistor.	2
1.2	Schematic illustrations of (a) an n-TFET, in which orange, grey and green colors show n-type, intrinsic and p-type semiconductors, respectively. (b) Fermi-Dirac statistics at the source region, (c) Band diagram at the indicated cross-section $z - z'$ of Fig. 1.2(a) and in steps from OFF-to-ON state, (d) the transfer-characteristics of the n-MOSFET (dash-orange) and n-TFET (blue). The dual-rounded orange arrows represent the $I - V$ data points corresponding to high energy Fermi-tails. The transfer-characteristics of an n-MOSFET (dash-orange) is shown for comparison with the n-TFET.	4
1.3	Schematic diagrams of (a) n-TFET, in which orange, grey and green colors show n-type, intrinsic and p-type semiconductors, respectively. (b) Band diagram at the indicated cross-section $z - z'$ of Fig. 1.3(a) illustrating a TAT process, where the blue and red lines correspondingly represent the ON- and OFF-state of the TFET. (c) the transfer $I - V$ characteristics showing the qualitative impact of TAT on SS-degradation.	6

- 1.4 Schematic diagrams of (a) n-MOSFET, in which orange and green colors show n-type and p-type semiconductors, respectively. (b) Band diagram at the indicated cross-section $x - x'$ of Fig. 1.4(a) illustrating the TAT impact on GIDL, where the red arrows represent the additional contribution of the TAT, (c) the transfer characteristics at high V_{DS} showing qualitatively the increase in OFF-state leakage currents due to the impact of TAT on GIDL. (d) Band diagram at the indicated cross-section $x_1 - x'_1$ of Fig. (1.4)(a) describing the effect of the TAT process on SILC, where DT refers to ballistic tunneling process. (e) Band diagram at the indicated cross-section $x_2 - x'_2$ of Fig. (1.4)(a) describing the impact of the TAT process on the drain-substrate tunneling. 7
- 1.5 Graphical illustrations of (a) n-TFET, in which orange, grey and green colors show n-type, intrinsic and p-type semiconductors, respectively. The trap is presented as a circle at the source junction. (b) Band diagram with a trap state at the indicated cross-section $z - z'$ of Fig. 1.5(a), with a focused region representing, (c) two ballistic-trap tunneling transitions, (d) ballistic tunneling followed by phonon-assisted tunneling, (e) phonon-assisted tunneling followed by ballistic tunneling, (f) phonon-assisted tunneling and phonon-assisted tunneling, (g) NMP-based recombination, followed by ballistic tunneling and (h)SRH-generation, from the valence band to the conduction band via a trap state. 9
- 2.1 An example of a one-dimensional potential of a finite length along the x -axis, which corresponds to the conduction band edge of a MOSCAP biased in accumulation. The inset shows a zoomed-out version of the oxide trap implemented as a square well potential. The asymptotic regions are indicated by α and β , respectively. 28
- 2.2 The electrostatics of a MOSCAP biased in accumulation shown with an inset of a zoomed-in version of the oxide trap configuration. The wavefunctions are shifted along the y -axis for the illustrative purpose of showing the behavior of the oxide trap wavefunction in the trap region and in the asymptotic regions. The unphysical exponential growth of the wavefunction (red) obtained with the delta normalization is indicated by a magenta circle, where as the wavefunction obtained with the constant density normalization is indicated with green. 39

- 2.3 The flowchart depicting the numerical procedure of finding the trap level and FWHM spectral range. 40
- 2.4 The electrostatic potential (solid black), the characteristic level ($E_t \approx -0.002$ eV) and the corresponding wavefunction of a square well 2 nm-wide trap configuration in a 30 nm long $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ region at 100 kV/cm electric field strength. E_t^F is the energetic difference taken between E_C and E_t at the trap center. 42
- 2.5 (a) The electrostatic potential and the characteristic sub-levels specifying the lower and the upper bounds of full-width half-maximum spectral range ($E_t^{\text{Max}} \approx -1.003$ eV, $E_t^L \approx -1.034$ eV and $E_t^U \approx -0.962$ eV), (b) The wavefunctions corresponding to the trap levels of (a), for a 2 nm-wide SW trap in a 30 nm long section of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ 43
- 2.6 Trap level shift and broadening determined as a function of electric field strength for the 2 nm-wide SW of Fig. (2.4). The closed and the open symbols represents the localization and the non-localization of the trap-level E_t in the SW trap region. . . 44
- 2.7 Electric field-induced level shift and broadening for the 2 nm-wide SW, 1 nm-wide SW and the CW trap configurations. The open symbols represents the non-local trap states. The inset depicts the described trap configurations at zero-field condition and the wavefunction probability densities for the indicated trap levels E_t . 45
- 2.8 A PolySi/SiO₂/p⁻Si MOSCAP with an oxide trap in accumulation with an inset of a zoomed-in view of the conduction band energy. The oxide trap is configured as a 0.5 nm wide and 3.55 eV deep SW and is located at 0.25 nm from the SiO₂/p⁻Si interface. The trap exhibits broadening as shown with the varying probability densities in the trap region. 46
- 2.9 A PolySi/SiO₂/p⁻Si MOSCAP with an oxide trap in inversion with an inset of a zoomed-in view of the conduction band energy. The oxide trap is configured as a 0.5 nm wide and 3.55 eV deep SW and is located at 0.25 nm from the SiO₂/p⁻Si interface. The trap state is a fully localized bound state. The higher amplitudes of the wavefunction, right before entering into the conduction band, could probably be related to the interference process occurring between the incident and the reflected elements of the wavefunction. 47

- 2.10 Level shift (solid) and broadening (dashed) as a function of oxide electric field strengths for the 0.5 nm-wide SW oxide-trap configuration in the PolySi/SiO₂/Si MOSCAP system. The traps are configured with varying well-depth and are positioned at 1 nm from the oxide-substrate interface. 48
- 2.11 Level shift (solid) and broadening (dashed) as a function of oxide electric field strengths for the 0.5 nm-wide 3.05 eV deep SW oxide-trap configuration in the PolySi/SiO₂/Si MOSCAP system. The trap-level is at around 1 eV above the substrate conduction band edge at flat-band voltage condition. The traps are positioned at varying distance from the oxide-substrate interface. The traps near to the oxide-substrate interface show significant level broadening. The black dashed box indicates the transition from quasi-bound to bound systems. 49
- 2.12 Level shift (solid) and broadening (dashed) as a function of oxide electric field strengths for the 0.5 nm-wide SW oxide-trap configuration in PolySi/SiO₂/Si and Al/HfO₂/InGaAs MOSCAP systems. The traps are configured with 3.05 eV and 3.5 eV well-depths and are positioned at a fixed distance of 1 nm from the oxide-substrate interface, respectively. The trap-levels in both MOSCAP systems are at 1 eV above the substrate conduction band edge in flat-band voltage condition. The trap in Al/HfO₂/InGaAs exhibits substantial broadening compared to PolySi/SiO₂/Si MOSCAP systems. The quasi-bound to bound transition region is indicated with a black dashed box 50
- 2.13 Graphical illustrations of (a) n-TFET, in which orange, grey and green colors show n-type, intrinsic and p-type semiconductors, respectively. The trap is presented as a circle at the source junction. (b) Band diagram with a trap state at the indicated cross-section $z - z'$ of Fig. 2.13(a), with the focused regions representing the impact of electric field on the TAT in the off-state of a simple p-i-n TFET for (c) the thermally-assisted TAT models and (d) the structural relaxation based TAT models, respectively. 52
- 2.14 (a) The Cr acceptor type trap configured as a SW and CW, (b) Replication of the emission rates outlined in literature on a different scale, along with the derived intrinsic impact of the field on the emission rates. The field effects are reflected as the increase in the average emission rates as can be seen from the level shift-only (dot) and the spectrum based solid lines. 55

2.15	Graphical illustration of the field enhanced emission mechanism in a finite-width barrier type structure. The trapezoidal barrier is treated as a combination of a fixed-width type and a triangular-type barrier with a common thermal ionization energy E_r	56
2.16	Electric field dependent average emission rates at 300K as a function of oxide electric field strengths for different instances of the 0.5 nm-wide SW oxide-trap configuration in the PolySi/SiO ₂ /Si and the Al/HfO ₂ /In _{0.53} Ga _{0.47} As MOSCAP systems. The increase in the average emission rates for MOSCAP systems is negligible in comparison with that of semiconductors.	58
3.1	A schematic representation of a QTBM approximation.	70
3.2	Schematic depiction of the injection of a single mode γ_α (open circle) into a diode at an injection energy E_α , whereas the injection of modes γ_β at the same energy are shown for illustrative purpose.	81
4.1	Simulated homostructure In _{0.53} Ga _{0.47} As p-n diode configuration. Dopant profiles are abrupt and uniform. Translational invariance is assumed in the y -direction. The black-dashed line refers to a cross-section along the center of the diode.	92
4.2	The direct-BTBT and PAT current densities of the homostructure In _{0.53} Ga _{0.47} As p-n diode of Fig. (4.1) as a function of applied voltage. The diode is 60-nm long and 20-nm wide and is uniformly doped with a concentration of 5×10^{18} [at/cm ³]. FB and RB stand for the forward-bias and reverse-bias, respectively.	94
4.3	Band diagram along the indicated cross-section of the In _{0.53} Ga _{0.47} As p-n diode illustrated in Fig. (4.1). The solid-black lines correspond to valence band maximum and conduction band minimum energies. The grey-dashed vertical lines limit the region where tunneling across the junction takes place at the given energies. The pink lines correspond to the quasi-Fermi energy levels E_{fp} and E_{fn}	95

- 4.4 The 2-band envelope function densities averaged over the unit cell at specific energies E_α and E_β , indicated in Fig. (4.3). The orange (blue)-solid line corresponds to an electron injected at E_α from the left contact (at E_β from the right contact), respectively. The brown-solid line represents the coupling density, where we chose the phonon plane wave with the directional components $(q_x, q_y, q_z) = (\pi/L_x, 0, 0)$. The limits of the tunneling junction region at (E_α, E_β) are indicated with the grey-dashed vertical lines and the drop of the EPC in this region is highlighted with a grey-dashed circle. The contribution of the PAT process in the near-tunneling regions is indicated with grey-dashed boxes. 96
- 4.5 The real and imaginary components of each band of the EFs for an electron injected from both contacts. The orange (blue)-solid (dotted) lines correspond to the valence (conduction)-band real (imaginary) contributions of the EF of an electron injected from the left contact. The red (dark yellow)-solid (dot) lines refer to the valence (conduction)-band real (imaginary) contributions of the EF of an electron injected from the right contact. 97
- 4.6 Dependence of direct-BTBT and PAT current densities on the device length L_x along the transport direction and on the minimum phonon wavevector. The simulated device is an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode with varying L_x and a constant 20nm width, uniformly doped with a concentration of 5×10^{18} [at/cm³]. 98
- 4.7 Band diagrams for the different doping concentrations along the indicated cross-section of the p-n diode of Fig. (4.1). The solid-(dot-, dashed-) lines correspond to the band-edge energies for 5×10^{19} (5×10^{18} , 5×10^{17}) [at/cm³] doping concentrations, respectively. The orange (blue)-solid line refers to the energies E_α (E_β). 99
- 4.8 The orange-(blue-) solid (dot,dashed) lines indicate the envelope function density averaged over the unit cell in a 2-band model and for doping concentrations of 5×10^{19} (5×10^{18} , 5×10^{17}) [at/cm³] corresponding to the band diagram of Fig. (4.7). The tunneling regions are marked with the horizontal arrows " L_t ". 100

- 4.9 The brown- solid (dot,dashed) lines refer to the coupling density of a 2-band model and for 5×10^{19} (5×10^{18} , 5×10^{17}) [at/cm³] doping concentrations corresponding to the band diagram of Fig. (4.7). The coupling density is determined by considering the phonon plane wave with the directional components $(q_x, q_y, q_z) = (\pi/L_x, 0, 0)$. The tunneling regions are marked with the horizontal arrows " L_t ". 101
- 4.10 Comparison of the direct-BTBT and PAT current densities for different doping concentrations. The orange-(blue-) solid (dot, dashed) lines represent the direct-BTBT (PAT) current densities for 5×10^{19} (5×10^{18} , 5×10^{17}) [at/cm³] doping concentrations. The cross-over between the direct-BTBT and PAT current density curves is indicated by a grey circle. 102
- A.1 Graphical illustration of configuration coordinate (CC) diagram of (a) a system of a delocalized electron and the lattice (defect) site prior to capture of an electron and (b) a system when an electron is captured at the lattice (defect) site. 119

Chapter 1

Introduction

The first technical use of a semiconductor copper-oxide as a rectifier in 1926 by Grondahl and in photocells by Lang in 1932, marked the beginning of the exploration of semiconductors [1]. The quest for a solid-state-based electronic switching element to replace vacuum tubes initiated the research of elemental Silicon and Germanium semiconductors, which turned out to be decisive for the development of semiconductor devices. Some highlights were the discovery of the p-n junction in 1939 by Russell Ohl [1] followed by the invention of the Germanium-based bipolar transistor in 1947 by W.Shockley [2, 3] and the realization of the Silicon-based metal oxide field effect transistor (MOSFET) in 1964 by Fairchild and RCA [4]. Meanwhile, the advent of Integrated Circuits (IC) by Jack Kilby and Robert Noyce (1958-1959) laid the foundation for the creation of microprocessors and supported the substantial growth of semiconductor industry.

Apart from the efficient processing of an IC, the scaling down of MOSFETs in size and the subsequent increase in transistor density on a chip while improving power efficiency and speed, positioned MOSFETs as the dominant electronic switching technology over the last four decades. Therefore, we start by briefly discussing the working principle and scaling challenges of MOSFETs followed by exploring a novel device, the Tunnel Field-Effect Transistor (TFET) and the parasitics limiting the optimum TFET performance. Since trap-assisted tunneling (TAT) is seen as one of the many reasons for TFET performance degradation, an extensive discussion of TAT is presented. Finally, the objective of this thesis is defined.

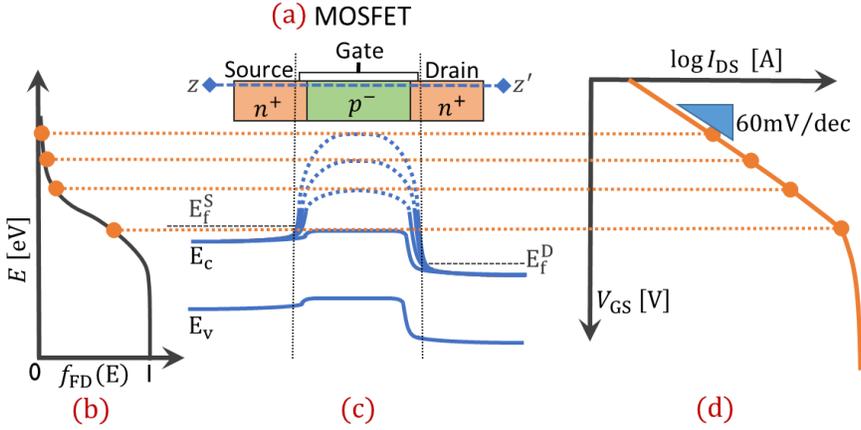


Figure 1.1: Schematic representations of (a) an n -MOSFET, in which orange and green colors show n -type and p -type semiconductors, respectively. (b) Fermi-Dirac statistics at the source region, (c) Band diagram at the indicated cross-section $z - z'$ of Fig. 1.1(a) and in steps from OFF-to-ON state, (d) the transfer-characteristics of the n -MOSFET. The dual-rounded orange arrows show the $I - V$ data points corresponding to high energy Fermi-tails, preventing the abrupt switching of the transistor.

1.1 MOSFET and scaling limitation

Gordon Moore noted in 1965 that the scaling down of MOSFETs lead to the doubling of the number of transistors on a chip every year, which became Moore's law [5]. However, the failure to follow the criteria of constant field scaling resulted in the modification of Moore's law, which suggested the doubling of transistor density on a chip every two years [6]. According to Dennard, to enforce the constant field scaling, the scaling of voltage is necessary by the same factor used for the scaling of the device dimensions [7]. However, the scaling down of gate voltage requires the scaling of threshold voltage (V_{th}), to maintain constant overdrive voltages. Consequently, assuming an ideally scaled on-current, this increases the static power dissipation of a transistor due to the exponential increase of subthreshold currents as the threshold voltage decreases [8, 9]. Moreover, the leakage currents associated with tunneling through the gate dielectric, which are more pronounced for smaller devices due to reduced dielectric thickness, resulted in severe static power dissipation and self-heating problems. It further limited the scaling of the threshold voltage and, in turn, of the operating voltage [7].

A better understanding of the scaling restrictions is attained by investigating the characteristics of an ideal MOSFET at room temperature, which is illustrated graphically in Fig. (1.1). In an ideal MOSFET (Fig. (1.1) (a)), the gate efficiently controls the electrostatic potential in the channel, since there are no presumed trap states present at the oxide substrate interface. In the subthreshold region of the transfer characteristics (Fig. (1.1) (d)), the gate-controlled channel barrier (Fig. (1.1) (c)) is large, limiting the entering of source region electrons into the channel. The electrons and holes in the source region obey Fermi-Dirac statistics (Fig. (1.1) (b)), and therefore have a non-zero occupational probability for electrons in high energy states (thermionic tails). The electrons in these states can surmount the channel-barrier which results in appreciable diffusion currents in the subthreshold operation of the MOSFET. The process of such thermionic emission is characterized by a well known Subthreshold Swing (SS) parameter given by [8]

$$SS = \left(\ln(10) \frac{k_B T (C_{ox} + C_{dep})}{e_q C_{ox}} \right) \quad (1.1)$$

where C_{dep} stands for the depletion layer capacitance per unit area, C_{ox} is the oxide capacitance, e_q is the elementary charge, k_B is the Boltzmann constant and T is the temperature. For an ideal MOSFET, the SS approximately equals $\ln(10)k_B T/e_q$ and at room temperature ($T = 300K$), this is 60mV/dec. This is the fundamental limit of SS for an ideal MOSFET, and it implies a minimum of 0.3V of threshold voltage to attain five orders of magnitude for the I_{ON}/I_{OFF} ratio. For gate voltages greater than the threshold voltage, a strong inversion layer of electrons is created in the channel (Fig. (1.1)c). As a result, the electrons can drift towards the drain.

1.2 TFET: A promising low-power device

In order to overcome the fundamental limit of the MOSFET, a novel device is needed based on either new materials or unconventional switching mechanisms. The potential devices which can replace or complement the MOSFET are classified into three categories [10]. The first category of devices deals either with improving the gate control over the channel electrostatics (nanowires, FinFETs) or with enhancing the field dependent transport properties of the channel material (high mobility III-V and SiGe). The second category is based on novel techniques to introduce and control charge carriers in the channels, such as negative-capacitance and Nano-Electro-Mechanical Field-Effect Transistor (NEMFET) devices. The last category of novel devices are non-charge based devices which use electron-spin states as the variables (Spintronics). Of all

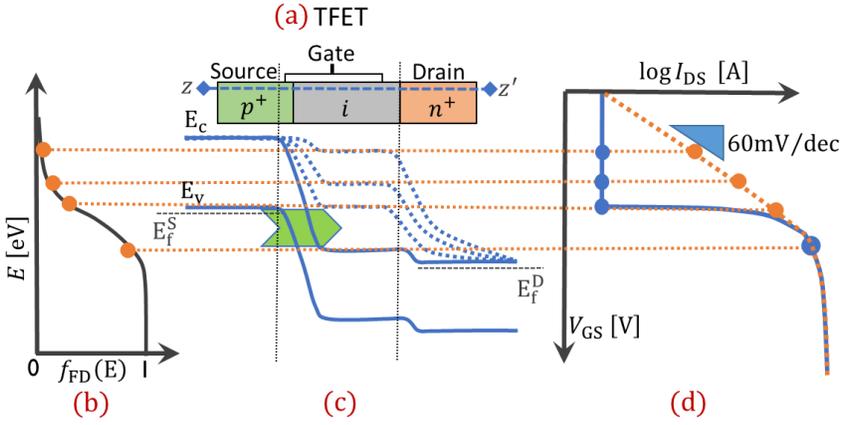


Figure 1.2: Schematic illustrations of (a) an *n*-TFET, in which orange, grey and green colors show *n*-type, intrinsic and *p*-type semiconductors, respectively. (b) Fermi-Dirac statistics at the source region, (c) Band diagram at the indicated cross-section $z - z'$ of Fig. 1.2(a) and in steps from OFF-to-ON state, (d) the transfer-characteristics of the *n*-MOSFET (dash-orange) and *n*-TFET (blue). The dual-rounded orange arrows represent the $I - V$ data points corresponding to high energy Fermi-tails. The transfer-characteristics of an *n*-MOSFET (dash-orange) is shown for comparison with the *n*-TFET.

these categories, TFET is considered as one of the Steep-Subthreshold Swing (Steep-SS) devices using novel transport mechanism, which could theoretically allow to achieve sub-60mV/dec SS. It is seen as the most promising alternative for the MOSFET for low power applications, as it can easily be integrated in the existing CMOS technology to fabricate the devices.

In TFETs, the transport of charge carriers is determined by transmission through a classically forbidden bandgap region, specifically by a quantum mechanical (QM) phenomenon called Band-to-Band tunneling (BTBT) [11]. Historically, the Esaki-tunnel diode was the first device demonstrating BTBT as the transport mechanism [12]. The working principle of the TFET and the possibility to overcome the fundamental SS-limit of the MOSFET are schematically shown in Fig. (1.2)(a)-(d). The major difference between TFET and MOSFET is that BTBT in TFET mostly determines the transport of the charge carriers, whereas the transport in MOSFET is determined by the classical drift-diffusion of charge carriers along the channel. The structure of a *n*-type TFET is graphically shown in Fig. (1.2)(a), in which the source is degenerate *p*-type, the channel is lowly doped intrinsic and the drain is highly doped *n*-type

semiconductor. The electrons in the source region obey Fermi-Dirac statistics Fig. (1.2)(b), where the probability of finding an electron in the high energy states is non-zero. However, for the TFET subthreshold gate voltages, these high energy states are filtered by the effective bandgap of the intrinsic channel Fig. (1.2)(c). Subsequently, negligible currents are observed in the transfer characteristics Fig. (1.2)(d) and the device is still in the OFF-state. When the gate voltage is higher than a certain onset voltage, the valence and the conduction band overlap and a steep onset is expected (blue line in Fig. (1.2)(d)) in the transfer characteristics which can attain sub-60mV/dec SS as compared to MOSFET (orange dashed line in Fig. (1.2)(d)).

1.3 Parasitics limiting the TFET performance

As the transport mechanism is determined by BTBT, the on-currents I_{ON} in TFETs are low compared to MOSFET, particularly for silicon TFETs due to the large effective indirect-bandgap [13]. This resulted in an ongoing extensive research in pursuit of alternative device structures and materials to boost the low I_{ON} currents [14, 15]. For instance, a line-TFET configuration is proposed for group-IV semiconductors, which is different from the conventional TFET (also known as point TFET) in its structure that the tunneling is oriented orthogonal to the gate [16, 17]. Another example for group-IV semiconductors is the introduction of localized doping “pocket” regions to boost I_{ON} or to reduce the impact of field-induced quantum confinement [18, 19]. Additionally, recent advances predicted substantial improvement in both steep onset and I_{ON} current by using heterostructure configurations of direct-bandgap III-V materials [20].

However, the theoretical predictions tend to provide promising TFET performance while experimentally obtained results tend to fall short [21]. One reason for such discrepancy is the overlooking of parasitic effects in theoretical calculations. These parasitics can considerably impact the SS and I_{ON} of TFETs. Some of the parasitics detrimental for TFET performance are phonon-assisted tunneling (PAT) [22], trap-assisted tunneling (TAT) [23, 24, 25, 26], heavy doping induced band tail effects [27], Auger recombination [28], the adverse effects from the traps at the semiconductor-oxide interface and surface roughness [29]. One such parasitic effect is graphically described in Fig. (1.3)(a)-(c), where a trap (red dot) is indicated near the tunneling junction. A comparison is made between the On-state and Off-state TFET electrostatics in Fig. (1.3)(b). In the Off-state, an electron can gain sufficient energy from the heat bath and can advance through the indicated trap into the conduction band, as opposed to non-existent direct tunneling transmission. This transition through a trap state

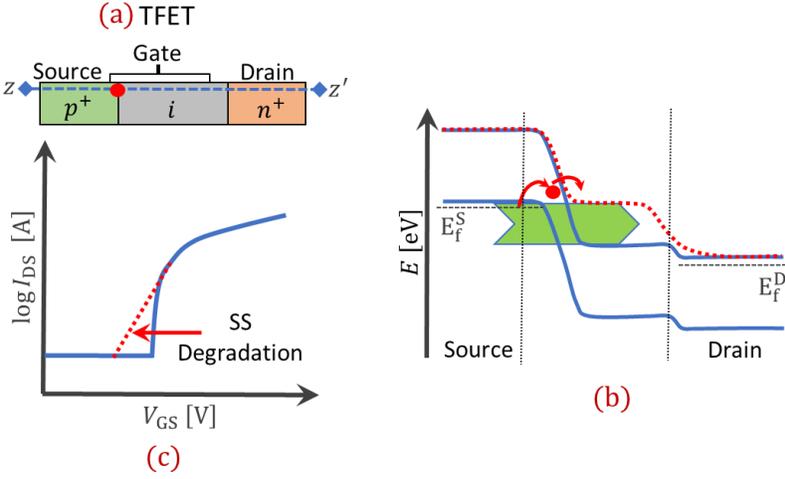


Figure 1.3: Schematic diagrams of (a) n -TFET, in which orange, grey and green colors show n -type, intrinsic and p -type semiconductors, respectively. (b) Band diagram at the indicated cross-section $z - z'$ of Fig. 1.3(a) illustrating a TAT process, where the blue and red lines correspondingly represent the ON- and OFF-state of the TFET. (c) the transfer $I - V$ characteristics showing the qualitative impact of TAT on SS-degradation.

is known as the trap-assisted tunneling (TAT) mechanism. Consequently, the combinations of all such transitions can add up to the transfer characteristics in the form of leakage currents and thereby deteriorating TFET performance (SS degradation).

1.4 Literature review of trap-assisted tunneling in semiconductor devices

The initial step in pursuit of adding non-idealities to TFETs is to identify and understand the dominant mechanism responsible for the observed SS-degradation. In experimental TFETs, the trap assisted tunneling is observed as one of the most probable reasons for SS-degradation [30]. Apart from the standard thermionic and BTBT currents, the leakage currents associated with the TAT process are found as temperature dependent in the subthreshold voltage range of TFET [30]. In fact, the TAT current is first observed as an

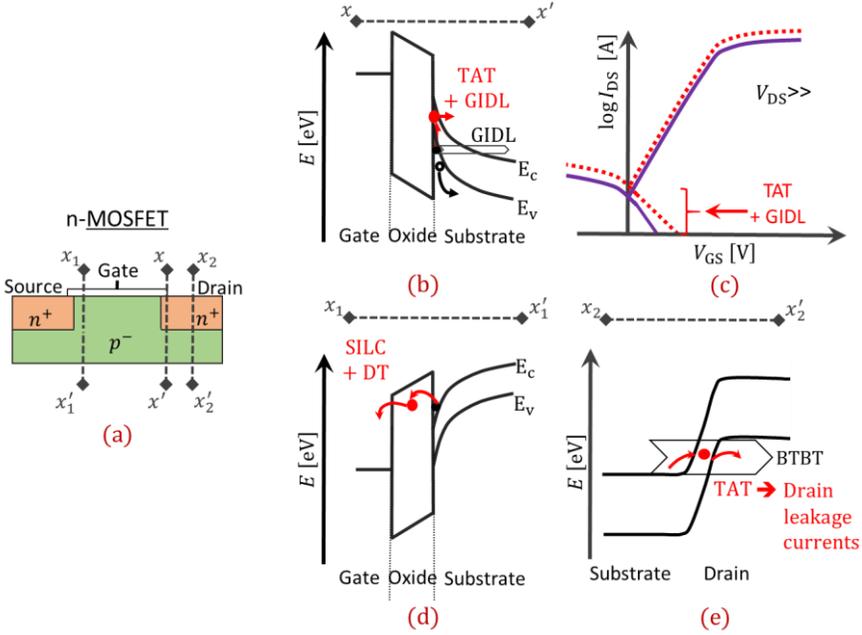


Figure 1.4: Schematic diagrams of (a) n-MOSFET, in which orange and green colors show n-type and p-type semiconductors, respectively. (b) Band diagram at the indicated cross-section $x - x'$ of Fig. 1.4(a) illustrating the TAT impact on GIDL, where the red arrows represent the additional contribution of the TAT, (c) the transfer characteristics at high V_{DS} showing qualitatively the increase in OFF-state leakage currents due to the impact of TAT on GIDL. (d) Band diagram at the indicated cross-section $x_1 - x'_1$ of Fig. (1.4)(a) describing the effect of the TAT process on SILC, where DT refers to ballistic tunneling process. (e) Band diagram at the indicated cross-section $x_2 - x'_2$ of Fig. (1.4)(a) describing the impact of the TAT process on the drain-substrate tunneling.

excess current during the forward-bias characteristics of a P-N tunnel diode (Esaki) [31, 32]. The excess currents calculated for these diodes without the inclusion of intermediate states (the trap states near the space charge region), are found as negligible compared to the experimentally observed currents, even though the currents are taken into account due to the interactions of an electron with photons, phonons, electrons and the Auger recombination [33]. However, the calculations with the trap states are found in good agreement with the experiments, thereby leading to the advent of TAT theory [34].

After a prolonged disinterest in tunnel diode leakage currents [35, 36], the recent progress in the scaling of CMOS devices has revived and promoted the interest in the TAT phenomena. In MOSFETs, the TAT increases the ballistic tunneling based gate induced drain leakage (GIDL) [37] currents in MOSFETs, as it provides an additional transmission across the drain-tunneling region. Therefore, the TAT in MOSFETs is responsible for the increase in OFF-state currents. Moreover, the TAT process is seen as the most probable reason for the generation of stress-induced leakage currents (SILC) [38] compared to ballistic tunneling based SILC for the moderately thick gate dielectric MOSFETs, whereas the ballistic tunneling process is observed as dominant over TAT for SILC in ultra-thin gate dielectric MOSFETs (<3nm oxide thickness) [38]. Fig. (1.4) graphically illustrates the impact of TAT on GIDL and the resultant increase in Off-state leakage currents at high V_{DS} . Moreover, Fig. (1.4) illustrates the impact of TAT on SILC and across substrate-drain tunneling region, whereby TAT can correspondingly increase the gate and drain leakage currents. On top of that, the impact of high oxide electric fields (upon aggressive scaling) and the elevated temperatures (due to compact device integration) on the MOSFET characteristics, particularly a shift in the threshold voltage and a change in SS, is partially assumed to be due to the TAT process in the form of capture of either a conduction or valence band electron into an oxide trap state through an interface trap [34, 39].

Moreover, the strong temperature dependence of TAT, observed in the reverse bias of trap limited diodes suggests a phonon-assisted TAT process[40]. In these diodes, reducing the temperature decreases the current by orders of magnitude, thereby suppressing the thermal part of the TAT process. Owing to such significance of TAT in semiconductor devices, the next section is devoted to a theoretical description and the limitations of the existing approaches to model the TAT process.

1.5 Modeling perspective of TAT

Over the years, numerous approaches have been proposed to describe the TAT process in a semiconductor device. The differences among these approaches is based on the methods to describe the underlying fundamental mechanisms of TAT such as the ballistic tunneling and thermal transitions of charge carriers into/from a trap state. An accurate description of these fundamental mechanisms and the associated sub-processes would manifest a good TAT model. Although the ballistic tunneling mechanism is similar among these models, it is described either by a semi-classical or a quantum mechanical method. Additionally, there exists a discrepancy in describing

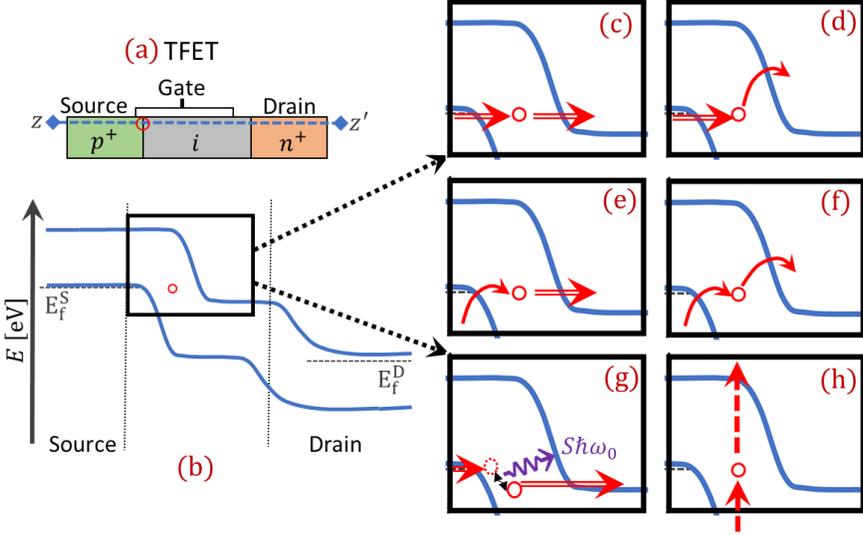


Figure 1.5: Graphical illustrations of (a) *n*-TFET, in which orange, grey and green colors show *n*-type, intrinsic and *p*-type semiconductors, respectively. The trap is presented as a circle at the source junction. (b) Band diagram with a trap state at the indicated cross-section $z - z'$ of Fig. 1.5(a), with a focused region representing, (c) two ballistic-trap tunneling transitions, (d) ballistic tunneling followed by phonon-assisted tunneling, (e) phonon-assisted tunneling followed by ballistic tunneling, (f) phonon-assisted tunneling and phonon-assisted tunneling, (g) NMP-based recombination, followed by ballistic tunneling and (h) SRH-generation, from the valence band to the conduction band via a trap state.

the thermal part of the TAT process among these models. In particular, the thermal transition of TAT is mostly described either by a thermally-excited tunneling [40], phonon-assisted tunneling [26] or non-radiative multiphonon (NMP) based recombination [41, 42].

The sub-processes associated with the ballistic and phonon-assisted tunneling via trap states, which exemplify the possible transitions required to be included in the accurate formulation of a TAT model, are graphically shown for a TFET in Fig. (1.5)(c)-(h). Assuming that the charge state of trap is neutral (also applicable for acceptor-type trap), the first sub-process of ballistic tunneling from the valence band state to a conduction band state via a trap (Fig. (1.5)(c)), also known as the ballistic-trap tunneling transition, is similar

to the resonant tunneling mechanism. However, the ballistic-trap tunneling transition cannot explain the temperature dependence in the subthreshold region of direct-bandgap TFETs as there are no thermal excitations involved in this process. The second sub-process of ballistic tunneling to a trap state (capture) followed by a phonon-assisted tunneling transition (emission) is presented in Fig. (1.5)(d), whereas the third sub-process (Fig. (1.5)(e)) is the result of phonon-assisted tunneling of an electron from the valence band state into a trap state (capture) followed by ballistic tunneling into a conduction band state (emission). The fourth sub-process is the phonon-assisted tunneling of an electron from the valence band to the conduction band via a trap state (see Fig. (1.5)(f)). The fifth sub-process is different in a particular way: the electron capture in a trap state is a combination of two sub-processes, specifically, ballistic tunneling (capture) to a trap state which is in the excited state of the vibrational subsystem and the subsequent lattice relaxation (also known as structural relaxation) upon an electron capture [41, 42]. In particular, the electron capture requires readjustment of the bonding of the lattice ions with their nearest neighbor ions such that the lattice site is polarized in the vicinity of the captured electron. During the electron capture, the trap is assumed not to change position yet. Next, during the subsequent lattice relaxation, the new equilibrium position of the trap is reached with the simultaneous emission of multiple phonons (Fig. (1.5)(g)). In this particular recombination process, which is based on non-radiative multiphonon (NMP) theory and also known as Lattice Relaxation Multiphonon Emission (LRME) [43], the ballistic tunneling and the thermal transitions are coupled such that the capture process of TAT is mediated by multiphonon emission due to the structural relaxation upon an electron capture into the trap state. One of such transition together with the ballistic tunneling (emission process of TAT) from the trap is illustrated in Fig. (1.5)(g). Note that the emission process of lattice relaxation based TAT (emission of a trapped electron into the conduction band) can also be mediated by multiphonon absorption, which is not shown in Fig. (1.5), such that the trap is thermally excited to one of the high energy states of the vibrational subsystem prior to the emission process [41]. The final sub-process of Fig. (1.5)(h) is a standard Shockley-Read-Hall (SRH) generation in which an electron is thermally excited from a valence band state to a trap-state followed by its thermal excitation (emission) into a conduction band state. The SRH based leakage currents can easily be dominant in a tunnel diode, if there is a sufficiently large trap concentration [44].

Based on the possible TAT sub-processes in Fig. (1.5)(c)-(h), the existing modeling approaches of TAT can be generally classified into two categories, namely thermally-assisted tunneling models [40, 26, 45] and structural relaxation based models [41, 42]. The first four TAT sub-processes (Fig. (1.5)(c)-(f)) are included in the thermally-assisted tunneling models, such as Hurkx's

TAT [40] and Non-Equilibrium Green's Function (NEGF) based phonon-assisted tunneling [26, 45] models (as will be discussed in Section 1.5.1), in which the ballistic trap transition (Fig. (1.5)(c)) is intrinsically involved, making these models complete. However, Schenk's structural relaxation based TAT [41] and the extended-NMP (eNMP) [42] models (as discussed in Section 1.5.2) are formulated based on the fifth TAT sub-process (Fig. (1.5)(g)), which does not account for phonon-assisted tunneling of carriers into/out of a trap state (Fig. (1.5)(d)-(f)). Although good experimental agreement has been achieved for specific semiconductor structures, a better understanding of these models is essential to address their limitations. In the next subsections, we therefore discuss these models in detail and explain the limitation of these models.

1.5.1 Thermally-assisted tunneling models

In this section, we discuss the first category of TAT models described by the combination of direct tunneling and thermally-assisted transitions of charge carriers into (or from) the trap states. In particular, we begin the discussion with the semi-classical TAT model of Hurkx followed by the quantum mechanical NEGF based phonon-assisted tunneling approach.

Hurkx's TAT model

The underlying mechanism in Hurkx's TAT model is based on the combination of ballistic tunneling and the thermally-excited transitions (in energy steps of $k_B T$) of charge carriers from the trap into either conduction or valence band. The ballistic tunneling part of TAT is included through the calculation of the carrier densities in the trap region whereas the thermal transitions are manifested in the form of electric field enhanced emission rates of trapped (captured) carriers into the conduction or valence bands.

In more detail, in this statistical based phenomenological model, TAT is calculated as the net recombination rate determined by the detailed balance between the net capture rate of the holes and that of electrons, which is governed by the expression [40]:

$$R_{\text{TAT}} = N_T \frac{c_{pF} c_{nF} n_t p_t - e_{nF} e_{pF}}{c_{nF} n_t + c_{pF} p_t + e_{nF} + e_{pF}} \quad (1.2)$$

where N_T refers to the trap density, n_t and p_t are the captured electron and hole densities in the trap (across the depletion region) with the field-dependent capture rates c_{nF} and c_{pF} , respectively. Similarly, e_{nF} and e_{pF} are the electron and hole emission rates. The total carrier densities inside the trap or in the

conduction band at the trap location x_t in Eq. (1.2) are determined by solving the 1D-Schrödinger equation, while assuming a homogeneous electric field across the tunneling junction, with an analytical Airy solution and are given by [46]

$$n_t(x_t) = \left[n(x_t) + \int_0^{x_t} dx_i \left(-\frac{dn(x_t)}{dx_t} \right) \Big|_{x_t=x_i} \frac{\text{Ai}^2((2e_q F m_e^* / \hbar^2)^{1/3} (x_t - x_i))}{\text{Ai}^2(0)} \right] \quad (1.3)$$

where $n(x_t)$ is the conventional electron density in the conduction band, x_i is the starting (or ending) position of a tunnel path from the conduction band through a triangular electrostatic potential and ending at the trap location x_t , \hbar is the reduced Planck constant, F is the electric field strength and m_e^* is the electron effective mass and the lower limit of the integral (“0”) is the left-most allowed onset of a tunnel path into the trap. Note that the second term in Eq. (1.3) refers to the ballistic tunneling contribution of TAT. The electric field enhanced emission rate (representing thermally-excited transitions) is also determined by solving the 1D-Schrödinger equation for an electron subjected to a triangular electrostatic potential, which is obtained with either an analytical solution (in the form of Airy functions [40]) or a solution based on the Wenzel-Kramer-Brillouin (WKB) approximation [47]. Such field enhanced emission rate is given by [40]

$$\frac{e_{nF}}{e_{n0}} = \left[1 + \frac{1}{k_B T} \int_0^{E_i} d\xi \exp\left(\frac{\xi}{k_B T}\right) \frac{\text{Ai}^2((\sqrt{2m_e^*}/e_q \hbar F)^{2/3} \xi)}{\text{Ai}^2(0)} \right] \quad (1.4)$$

where e_{n0} is the zero-field emission rate and E_i is energetic difference between the measured zero-field trap level (E_t) and the conduction band minimum (which is determined from the thermal ionization of the trap). Moreover, the capture rates are determined by the detailed balance equation, which states that each capture process is balanced by its reverse emission process, and are related to emission rates by [44]:

$$\frac{e_{nF}(E)}{c_{nF}(E)} = \frac{c_{pF}(E)}{e_{pF}(E)} = \exp\left(\frac{E_t - E}{k_B T}\right) \quad (1.5)$$

Due to the similar form of expressions for the carrier densities (Eq. (1.3)) and emission rates (Eq. (1.4)), the field enhancement for these quantities are related by a common factor, which further simplifies the net recombination rates for TAT (Eq. (1.2)).

Since the trapped carrier densities, the capture and the emission rates are formulated for a triangular electrostatic barrier, the TAT rates determined with Hurkx’s model are inaccurate when there is an electron capture process into

the trap from the valence band or a hole capture process into the trap from the conduction band, because the charge carriers in this case are subjected to a trapezoidal electrostatic potential and the Eqs. (1.3) and (1.4) must accordingly be modified to account for this potential. Note that the structural (lattice) relaxation upon either the capture or emission of charge carriers into the trap state is not considered in this model.

NEGF based phonon-assisted tunneling approach

The inevitable limitations of semi-classical models in determining the exact description of TAT lead to the development of quantum-mechanical based numerical methods [26, 45]. One such method is based on solving the 3D Schrödinger equation using an 8-band \mathbf{k}, \mathbf{p} electron basis and Non-Equilibrium Green's Functions (NEGF) [26]. A 1nm^3 cubic trap potential is arbitrarily added to the device electrostatics, specifically at an oxide-semiconductor interface. The acoustic and optical phonon scattering are included in the form of self-energies and the Schrödinger equation is solved in the framework of the self-consistent Born approximation (SCBA) [48], thereby resulting in the local-density of energy states (LDOS) for a given trap configuration. The phonon-assisted tunneling through these LDOS at low gate voltages can then explain the observed SS-degradation, in a regime where there is no tunneling window for direct-BTBT. To complete the study, the spatial and energetic depths of the trap are varied to determine the impact on the SS-degradation of a theoretical InAs-based TFET [26].

Although most of the aspects of TAT are considered in this approach (Fig. (1.5)(c)-(f)), it lacks in accounting for the structural relaxation (Fig. (1.5)(g)) which is still required to comply with an experimental device configuration. Note that since the model is based on SCBA, it typically demands an enormous amount of computational resources.

1.5.2 Structural relaxation based TAT models

Initially, the experimentally observed large capture cross sections of a shallow impurity level lead to the theory of cascade-phonon emission, whereby a carrier is captured through ballistic tunneling into an electronically excited state and loses energy by cascading down the ladder in steps of the energetic distance between the excited levels to the ground state [49]. However, multi-transitions among non-uniformly spaced excited levels require the separation between the excited levels to be less than the phonon energy. This therefore implies an improbable single phonon transition from the first excited state to

the ground state and this transition requires instead a multiphonon process due to its energetic separation larger than the single phonon energy. These restrictions however could not explain the large cross sections measured at room temperatures [49]. Note that the cascade-phonon emission is defined as a reversible process, whereby the release of captured carriers is characterized by multi-absorption of single phonon energies.

Meanwhile, the NMP theory is also being developed to explain the experimentally observed large capture cross sections of a deep trap state in III-V semiconductors under zero-field conditions [50]. The large cross sections of a trap measured by Deep Level Transient Spectroscopy (DLTS) are found in agreement with the calculations made with NMP theory [41]. Due to such agreement with experimental results, TAT models based on multiphonon recombination processes are formulated, which are fundamentally based on the NMP theory [51]. The major distinction between the cascade-phonon and NMP theory is that the former is based on ballistic tunneling and multi-transitions (with single phonon emission or absorption) for capture or release of the carriers through the different energy levels of a trap, whereas the latter accounts for ballistic tunneling and an intrinsic single thermal transition (with multiphonon emission or absorption) because of structural (lattice) relaxation upon the capture (or release) of carriers into (or from) the trap state. Owing to the importance of NMP theory, the underlying fundamental microscopic processes (multiphonon assisted transitions and structural relaxation) are detailed in the Appendix A. The different approaches to apply NMP theory to model TAT are briefly discussed in the next subsections.

NMP recombination model (Andreas Schenk)

Since the multiphonon recombination model by Schenk [41] accounts for an inhomogeneous electric field across the tunneling junction during a TAT process, it is the most widely used model to describe TAT in semiconductor devices. The microscopic process defining Schenk's TAT recombination model is the combination of an NMP process [51] and the theory of Bloch electrons under large electric fields [52]. In general, the trap-assisted transition rate is derived by using the following Fermi's Golden rule with electric field and electron-phonon interactions as the transition inducing operators:

$$w_{if} = \frac{2\pi}{\hbar} \sum_{n,n'} \sum_{\lambda,\lambda'} p_{\lambda} \left| \left\langle \Phi_{n,\lambda}^{(1)} \psi_n^{(0)} \left| [-e_q \mathbf{F} \cdot \mathbf{r} + V_{\text{el-ph}}] \left| \psi_{n'}^{(0)} \Phi_{n',\lambda'}^{(1)} \right. \right\rangle \right|^2 \delta(\mathcal{E}_{n,\lambda} - \mathcal{E}_{n',\lambda'}) \quad (1.6)$$

where $V_{\text{el-ph}}$ is the electron-phonon coupling potential, \mathbf{F} is the electric field, $\Phi_{n,\lambda}^{(1)}$ is the first-order perturbed lattice vibrational function due to the electron-

phonon interaction, $\psi_{n'}^{(0)}$ is the stationary state eigenfunction of the one-electron Hamiltonian, $\mathcal{E}_{n,\lambda}$ is the total energy of the system, n, λ represent the electron and phonon quantum numbers corresponding to the electronic and vibrational subsystems, p_λ is the statistical weight of the initial lattice states and where the summation of electronic states extends over initial occupied and final empty states [53]. With the static approach (explained in Appendix A) of adiabatic theory (Eqs. (A.1)-(A.15) of Appendix A), the coupled-transition matrix of Eq. (1.6) is transformed into the (separate) product of electronic and vibrational transition matrix elements [53].

As mentioned in Appendix A, the ballistic tunneling (electronic transition) part of the NMP theory can be described by using the wavefunctions of the electronic subsystem Eq. (A.14). Therefore, the stationary state electron-Hamiltonian in Eq. (A.14) is described by including the terms corresponding to the externally applied inhomogeneous electric field and the quasi- δ trap potential ($V_{\text{tr}}(\mathbf{r}) = V_{\text{tr}}^0 \delta(\mathbf{r}) [1 + \mathbf{r} \cdot \nabla_{\mathbf{r}}]$) [53]. The electronic subsystem of Eq. (A.14) is then transformed into

$$[H_0 + V_{\text{tr}}^0 \delta(\mathbf{r}) [1 + \mathbf{r} \cdot \nabla_{\mathbf{r}}] - e_q \mathbf{F} \cdot \mathbf{r}] \psi_n^{(0)}(\mathbf{r}) = E_n^{(0)} \psi_n^{(0)}(\mathbf{r}) \quad (1.7)$$

where H_0 is the electron Hamiltonian for the bulk semiconductor, V_{tr}^0 is the zero-field trap potential strength with one bound state in the bandgap and described in terms of the binding energy. The electronic wavefunctions are expanded using envelope functions $f_n^{\mathbf{k}_\perp}(x)$ in the framework of Bloch theory, while assuming a constant potential along the (y, z) -directions, orthogonal to the electron transport in the x -direction. With the following definition for the electronic wavefunctions in the framework of the 2-band ($n = (c, v)$ in Eq. (1.7)) effective mass approximation (EMA) [52],

$$\psi_{c,v}^{(0)}(\mathbf{r}) = \sqrt{\frac{2}{\Omega}} \sum_{\mathbf{k}} \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp) u_{c,v}(\mathbf{r}) f_{c,v}^{\mathbf{k}_\perp}(x) \quad (1.8)$$

where $u_{c,v}(\mathbf{r})$ are the Bloch functions corresponding to the conduction and valence bands. The electronic subsystem of Eq. (1.7), after changing to spectral coordinates, is simplified to [53]

$$[E_c(\mathbf{k} - \mathbf{k}_0) - E + ie_q \mathbf{F} \cdot \nabla_{\mathbf{k}}] A_c^{\mathbf{k}_\perp}(k_x) = 0 \quad (1.9)$$

$$[E_v(\mathbf{k}) - E + ie_q \mathbf{F} \cdot \nabla_{\mathbf{k}}] A_v^{\mathbf{k}_\perp}(k_x) + \sum_{\mathbf{k}'} \langle \mathbf{k} | V_{\text{tr}}(\mathbf{r}) | \mathbf{k}' \rangle A_v^{\mathbf{k}'_\perp}(k'_x) = 0 \quad (1.10)$$

Note that the trap potential is defined in the valence band part of the electron Hamiltonian relative to its band maximum. Further note that \mathbf{k}_0 in Eq. (1.9) is equated to zero for direct-bandgap materials. The spectral functions of

Eqs. (1.9) and (1.10) are determined analytically, without the field-induced effects on the trap-states (neglecting the electric field in Eq. (1.10)) [53]. Using Eqs. (1.8)-(1.10), the initial and the final electronic states' wavefunctions are determined. The electronic transition matrix elements (ballistic tunneling part of TAT) of an approximated form of Eq. (1.6) are then determined using the initial and final electronic states' wavefunctions and result in the field enhanced density of band-states [53].

Now the electronic transition matrix elements are known, the next step is to determine the vibrational transition matrix elements related to multiphonon absorption or emission during structural relaxation, which are calculated using Eq.(A.15) and the electronic wavefunctions (Eq.(1.8)) constructed with the spectral functions of Eqs. (1.9) and (1.10). In order to make the calculation of vibrational transition probabilities feasible with Eq. (A.15), the Einstein model [54, 55] is applied for the vibrational subsystem, which neglects the dispersion of the phonon and which considers only one effective local-mode phonon energy ($\hbar\omega_0$).

Additionally, a linear approximation (similar to Eq. (A.18)) of the electron-phonon coupling is applied in Eq. (A.15) [54, 55], in which the electron-lattice interactions are assumed to be linear in the displacement Q of the host atoms. In second quantization, it is described by

$$\left\langle \psi_n^{(0)} \left| V_{\text{el-ph}}(\mathbf{r}, \mathbf{Q}) \right| \psi_n^{(0)} \right\rangle = g_\lambda^n \hat{b}_n^\dagger \hat{b}_n \left(\hat{a}_\lambda + \hat{a}_{-\lambda}^\dagger \right) \quad (1.11)$$

where g_λ^n represents the electron-phonon coupling strength, $\hat{b}_n^\dagger, \hat{a}_\lambda^\dagger$ are the single particle creation operators and $\hat{b}_n, \hat{a}_\lambda$ are the single particle annihilation operators for the electrons and phonons, respectively. With these approximations, the vibrational transition matrix elements are solved in the framework of first order time-dependent perturbation theory while using the equation of motion for the lattice vibrations (Eq. (A.15)) and yields the following characteristic line shape function (LSF) of phonons [55, 56], which describes the frequency spread of a thermal transition process mediated by phonon (also known as vibrational broadening),

$$L_{sf}(E) = 2\pi\hbar \exp[-S(2v(\hbar\omega_0) + 1)] \sum_{p=-\infty}^{+\infty} \left(\frac{v(\hbar\omega_0) + 1}{v(\hbar\omega_0)} \right)^{p/2} \\ \times \mathcal{I}_p \left(2S\sqrt{v(\hbar\omega_0)(v(\hbar\omega_0) + 1)} \right) \delta(E + p\hbar\omega_0) \quad (1.12)$$

and subsequently resulting in the following multiphonon transition probability [41],

$$M_{c,v}(p) = \frac{(p \mp S)}{S} \exp[-S(2v(\hbar\omega_0) + 1)] \\ \times \exp\left(\frac{p\hbar\omega_0}{2k_B T}\right) \mathcal{I}_p\left(2S\sqrt{v(\hbar\omega_0)(v(\hbar\omega_0) + 1)}\right) \quad (1.13)$$

with the following definition for the Huang-Rhys factor S [54, 55]:

$$S\hbar\omega_0 = \sum_f |g_i^n - g_f^n|^2 \quad (1.14)$$

where g_f and g_i are the final and initial state's electron-phonon coupling strengths. In Eq. (1.12), $v(E)$ is the Bose-Einstein probability distribution function for phonons and \mathcal{I}_p is the modified-Bessel function of order p . Note that there exists a line shape function for the phonon absorption process similar to Eq. (1.12) which is included in the calculation of multiphonon transition probability of Eq. (1.13). Moreover, the Huang-Rhys factor S in Eqs. (1.12) and (1.13) acts as a measure of the difference in electron and phonon coupling strengths between the initial and final system states and determines the lattice relaxation energy. These lattice relaxation energies ($S\hbar\omega_0$) are often assumed to be smaller in semiconductors [57] than in the oxides [58]. Such assumption can be attributed to the fact that traps in oxides exhibit greater lattice relaxation (due to the amorphous nature of these materials) than in semiconductors, when an electron is captured or emitted into/from the defect site [42].

With both the electronic and vibrational transition matrices, the final solution of Eq. (1.6) results in the superposition of the field and temperature-dependent trap-assisted transition rates, which is described in the terms of spectral capture and emission rates. These spectral rates are manifested in terms of carrier lifetimes, which in turn determine the SRH-based net recombination rates of Eq. (1.2) [41]. The integration of these SRH recombination rates over real-space results in TAT currents.

Extended-NMP (eNMP) recombination model (Wolfgang Gös)

In recent years, a recombination model based on an extended NMP theory is able to successfully explain the essential features of Negative Bias Temperature Instability (NBTI) degradation curves and is able to reproduce the experimental findings of Time Dependent Defect Spectroscopy (TDDS) related to single defects. Although this model is formulated based on the exchange of carriers

between the substrate and oxide defects, which is different from a TAT process, the underlying microscopic processes are identical to those of TAT. The emission τ_e and capture τ_c time constants determined from this eNMP approach can be applied in the following expression to determine TAT currents [59]:

$$J_{\text{TAT}} = e_q \int_0^{t_b} dx \frac{N_T(x)}{\tau_c(x) + \tau_e(x)} \quad (1.15)$$

where N_T denotes the trap concentration and t_b is the barrier thickness. The electron capture and emission time constants of Eq. (1.15) in the eNMP approach are defined as

$$\frac{1}{\tau_c(x)} = \frac{\int_{E_c}^{\infty} dE c_n(E) D_n(E) f_{\text{FD}}(E)}{\int_{E_c}^{\infty} dE D_n(E) f_{\text{FD}}(E)} n(x) \quad (1.16)$$

$$\frac{1}{\tau_e(x)} = \frac{\int_{E_c}^{\infty} dE e_n(E) D_n(E) f_{\text{FD}}(E)}{\int_{E_c}^{\infty} dE D_n(E) f_{\text{FD}}(E)} n(x) \quad (1.17)$$

where E_c is the conduction band energy, c_n is the electron capture rate, e_n is the electron emission rate, D_n is the density of electronic states in the conduction band, f_{FD} is the Fermi-Dirac statistics for electrons and $n(x)$ is the conventional electron density in the conduction band. In the eNMP approach, the electron capture and emission rates of Eqs. (1.16) and (1.17) are determined from

$$c_n(E) = x_t A_{yz} r_{n,t}(E) L_{sf}^{n,t}(E) \quad (1.18)$$

$$e_n(E) = x_t A_{yz} r_{t,n}(E) L_{sf}^{t,n}(E) \quad (1.19)$$

where $r_{n,t}$ is the electronic transition rate from a band state to a trap state, $L_{sf}^{n,t}$ is the corresponding line shape function (LSF) of multiphonon theory, x_t is the position of the trap with respect to the oxide-substrate interface and A_{yz} is the planar area of the trap orthogonal to the transport of an electron. Similarly, $r_{t,n}$ is the electronic transition rate from a trap state to a band state.

The interesting aspect of this model is the calculation of electronic transition rates (related to the ballistic tunneling part of TAT), which is based on the Fermi's Golden rule in the framework of Bardeen's approximation [60, 61]. This approximates the total system Hamiltonian as the linear combination of independent and partial-system Hamiltonians. For instance, the system of a

semiconductor substrate and an oxide, with the conduction band interacting with an oxide defect is treated as a separate conduction band and an oxide subsystem based on the assumption that both of these subsystems are weakly coupled. In other words, the transmission probability of a band state electron is negligible at the defect. In this approach, the electronic transition rates from a band state to a trap state at a given energy, for one-dimensional systems, are [62]

$$r_{n,t} = \sum_n \frac{2\pi}{\hbar} |M_{n,t}|^2 \delta(E_t - E_n^{(0)})$$

$$M_{n,t} = \frac{\hbar^2}{2m_n^*} \int dx \int dy \left[\psi_t^* \frac{\partial \psi_n^{(0)}}{\partial z} \Big|_{z=z_{if}} - \left(\psi_n^{(0)} \right)^* \frac{\partial \psi_t}{\partial z} \Big|_{z=z_{if}} \right] \quad (1.20)$$

where $\psi_n^{(0)}(\mathbf{r})$ and $\psi_t(\mathbf{r})$ are stationary state band-state and δ -type trap state wavefunctions, respectively, m_n^* is the effective mass of band n and z_{if} is the position of the interface between oxide and substrate. Note that in the calculation of the transition matrix $M_{n,t}$, the initial band-state and final-trap state wavefunctions are the eigenfunctions of the two independent subsystems, which are the semiconductor substrate with an infinitely long oxide (without an oxide trap) subsystem and an infinitely long oxide subsystem with a trap. This is entirely different from the conventional Fermi's Golden rule (Eq. (1.6)), where the initial and final states electronic wavefunctions are eigenfunctions of the unperturbed Hamiltonian (Eq. (A.14)).

The eigenfunctions $\psi_n^{(0)}(\mathbf{r})$ and $\psi_t(\mathbf{r})$ in Eq. (1.20), characterizing the behavior of electrons in the semiconductor and trap regions, are determined using the Wenzel-Kramer-Brillouin (WKB) method as [62]

$$\psi_n^{(0)}(\mathbf{r}) = \frac{\sqrt{k_x} \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp)}{\sqrt{2\Omega K_{x,n}(x)}} \exp\left(-\int_{x_{if}}^{x_t} dx' K_{x,n}(x')\right) \quad (1.21)$$

$$\psi_t(\mathbf{r}) = \frac{K_{x,t}(x_{if}) \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp)}{\sqrt{2A_{yz} K_{x,t}(x)}} \exp\left(-\int_{x_{if}}^{x_t} dx' K_{x,t}(x')\right) \quad (1.22)$$

with,

$$K_{x,nt}(x) = \sqrt{2m_{n,t}^* \left[E_n^{(0)}(x) - E_x \right] / \hbar^2}; \quad k_x = \sqrt{2m_n^* E_x / \hbar^2}; \quad E_x = E - \frac{\hbar^2 |\mathbf{k}_\perp|^2}{2m_n^*} \quad (1.23)$$

where m_t^* is the tunneling mass and $E_n^{(0)}(x)$ is the respective band edge energy. With Eqs. (1.21)-(1.23), the electronic transitional rates of Eq. (1.20) are determined, while assuming no phonons are involved in the tunneling process [42]. In this model, these tunneling transitional rates are encoded as the temperature-independent capture cross sections (the first factor in the integral of the denominator of Eqs. (1.16) and (1.17) equals $\sigma_n v_{th,n}$, with σ_n being the capture cross section and $v_{th,n}$ the thermal velocity) in either the capture or emission rates of the carrier lifetimes (specifically in $r(E)$ of Eqs. (1.18) and (1.19) [42]).

Another striking feature of this model is that it accounts for the exact configuration of an oxide defect, based on the local environment of bonding of the defect site as well as the charge state of the defect, and configures the defect site into a combination of stable state and metastable state sub-configurations, thereby determining the temporal behavior of defects [42]. Due to the inclusion of such a defect configuration, this approach is known as extended-NMP model. In this model, the defect state configurations are determined from the adiabatic Eqs. (A.7) and (A.8) under zero-field conditions, however, in the framework of density functional theory (DFT). This results in a thermodynamic trap level (the trap energy level calculated after the structural relaxation) based on whether the charge carrier is in the channel of the MOSFET or in an oxide defect.

Since the underlying approach to determine the capture and emission rates is similar to Eq. (1.6) (specifically, Condon-approach of the adiabatic theory [63]), the convoluted transition rates of Eq. (1.6) are simplified as the product of the electronic and vibrational (structural relaxation based) transition rates, where the latter results are in the form of LSF of multiphonon theory (similar to Eq. (1.12)) [51].

1.5.3 Limitations of the models

Most of the discussed models employ certain approximations in agreement with either a specific device or an experimental configuration. For instance, unidirectional tunneling is assumed in Hurkx's and Schenk's model to explain TAT in sufficiently large $p-n$ diodes. As another example, in the eNMP model the set up is assumed to comply with TDDS experiments in which the source and drain of MOSFETs are connected to ground. Hence, it is essential to assess the limitations of each of these models before they can efficiently be applied to determine TAT.

The major drawback of Hurkx's model is the fact that transitions for trap levels which are below the minima/maxima of the conduction/valence

band are disregarded. This means that in TFETs, Hurkx's model would underestimate the prediction of TAT before the onset of BTBT. Apart from this, the tunneling into a defect is determined by either analytical Airy functions or WKB approximations. In the former case, the solutions of the Schrödinger equation cannot adequately replicate the exponential tails into the dielectric due to the approximation that the discontinuity at an interface is modeled as an infinitely high barrier. At the classical turning points, where the electron enters/leaves the forbidden gap, and hence where the electron energy equals the band edge energy, the WKB approximation breaks down in determining the wavefunctions, as the wavefunctions diverge at these edge points [64]. Therefore, neither of the approaches are efficient in describing the tunneling aspect of TAT in Hurkx's model. Additionally, the electron is presumed as fully localized in the trap state irrespective of the barrier type.

In Schenk's TAT model, the determination of the electron envelope function assumed the wavefunction to be a plane wave in two directions orthogonal to the tunneling barrier, implying that the electrostatic potential is constant for these directions. This assumption is incorrect for gated devices, where the electrostatics vary in other directions and thereby prevents the application of this model to gated devices. Additionally, the orthogonal direction plane wave description is also inaccurate for describing very small devices, where size confinement effects are crucial. This model is derived for a homostructure, hence it is inadequate for heterostructures where the effective mass and the effective bandgap are position dependent. The initial state of the trap is assumed to be neutral and this decreases the accuracy of the application of this model to acceptor, donor and extended defects, where the initial-state of the trap is no longer neutral. The aforementioned limitation in Hurkx's model that the electron is presumed as fully localized in the trap state irrespective of the barrier type is equally applicable to Schenk's model. However, Schenk's model can account for such high electric field effects in the form of changes in local trap density of states and level shifts, while retaining the electric field term in Eq. (1.10). These effects can also be incorporated in Hurkx's model by modifying the emission term accordingly, which will be discussed in the next chapter.

The aforementioned limitations related to the plane wave definition of electronic wavefunctions are equivalently applicable to Gös model. Additionally, these wavefunctions are determined by the splitting of the system Hamiltonian into two weakly coupled partial subsystems of the isolated trap and substrate (semiconductor without trap) configurations, presuming that the tunneling is inefficient across the common region of these subsystems (barrier). This approximation is applicable for barriers of which the thickness is sufficiently large. However, for thin and ultra thin oxides with very light effective mass,

these approximations are inaccurate owing to a strong coupling of the subsystems, which will be seen in the next chapter.

It is evident from the discussion of models and from Fig. (1.5), that each model can only be applied to certain TAT processes. For instance, Hurkx's model can describe the TAT process (c)-(f) of Fig. (1.5), while it cannot calculate TAT current based on structural relaxation (Fig. (1.5)(g)) assuming that the structural relaxation is negligible for the traps in semiconductors. Apart from this, it is also clear that the thermal aspect of TAT lacks a unified approach.

1.6 Goal of the thesis

It is clear from the previous section that there is no unified approach to model TAT in semiconductor devices. There is inaccuracy in either determining tunneling or thermal transitions of TAT due to simplifications made in the models. Therefore, the major objective of this thesis is to develop an approach towards the all-inclusive, fully quantum-mechanical modeling of TAT in semiconductor devices. The following are the key sub-goals of this thesis:

1) Investigate the impact of the electric field on traps and on the TAT-emission rates:

- Develop a quantum-mechanical based approach to quantify the high-field effects on the semiconductor traps and on the traps in MOS capacitor systems.
- Derive an emission rate including the field-induced effects on traps and investigate how the field effects can impact the device performances.

2) Develop a quantum-mechanical phonon-assisted tunneling formalism to simulate phonon-assisted transport:

- Derive a QM formalism to simulate phonon-assisted tunneling currents for bi-dimensional semiconductor device potentials (with an infinite third dimension).
- Derive the electron-phonon coupling strength equations for direct bandgap materials including the dispersion of phonons.
- Implement the QM formalism to enable the PAT-current calculations in direct-bandgap semiconductors such that it is inline with the existing

direct-BTBT simulator and optimize the implementation for efficient simulation times of the PAT current calculation.

3) Investigate the impact of PAT on the device characteristics:

- Determine whether the phonon-assisted tunneling is behaving similar to direct-BTBT.
- Determine the impact of the device dimensions on the PAT current calculations.
- Investigate the dependencies of the PAT current density on the doping concentration of p-n diodes.

4) Outline the framework towards TAT calculations in semiconductor devices:

- Present the future framework to calculate multiphonon assisted tunneling to enable TAT current calculation, where an additional structural relaxation step can be invoked.

1.7 Organization of the thesis

After the discussion of limitations of the existing TAT models in this chapter 1, we investigate the importance of field-induced effects on traps in chapter 2 and explain why it is crucial to model TAT using QM based methods. Chapter 2 starts with a brief description of the QM based method to extract trap levels in semiconductor and MOS capacitor systems. We describe the normalization procedure to attain constant probability amplitudes for trap wavefunction in the contacts. We present a numerical approach to quantify the field-induced effects on traps and explain the importance of such effects on TAT events. We derive an analytical expression for the trap emission rates including the field-induced effects. We calculate the emission rates for semiconductor and oxide trap systems.

In Chapter 3, as a first step towards a QM-based TAT formalism we develop the QM-based PAT current formalism and discuss in detail the determination of the electron-phonon coupling (EPC) strengths. We derive the PAT current equation based on the time-dependent perturbation theory using the number operator approach. We determine the EPC in the framework of envelope function theory and apply the low-wavevector approximation for local-based

Fröhlich interactions in direct-bandgap materials. The PAT current density equation is further simplified for bi-dimensional semiconductor devices. We present the numerical procedure to calculate the PAT current densities based on the existing direct-BTBT simulator, while making use of parallel determination of electron envelope functions. We discuss different implementations and outline their respective shortcomings.

Chapter 4 illustrates the application of the PAT current formalism to different direct-bandgap homostructure device configurations. It starts with the in-depth analysis of the electron-phonon coupling across tunneling junctions using the 2- and 15-band model of our formalism. We also compare the full PAT current densities with both models. We explain the impact of device dimensions on the PAT current characteristics of p-n diodes, thereby describing a shortcoming of our formalism. We explain the doping dependence of PAT current densities to determine whether PAT dominates direct-BTBT current densities in certain diode configurations.

The thesis concludes with the major findings and outlines the framework towards the implementation of TAT calculations using a QM approach. In particular, we present how a two-phonon assisted transition can be added to our formalism and can determine the associated currents. We further present a possible pathway of including the lattice-based structural relaxation into our formalism.

Chapter 2

Impact of high electric field on traps in semiconductors and oxides

In the previous chapter, we presented a brief discussion on the challenges in the scaling of MOSFET and introduced TFET as a potential candidate to replace MOSFET as a low power device. We briefly mentioned that TAT, in addition to other parasitic effects, restricts the capability of TFETs as it degrades the sub-60mV/dec SS. We also reviewed the existing semi-classical TAT models in literature and discussed their limitations, which in turn compromise accurate predictions of TAT currents. One limitation is that these models do not account for the intrinsic trap energy level shift with the externally applied electric field, despite the rigorous treatment of the other field contributions on TAT currents.

Since the electric fields in TFETs and across the gate oxides in MOSFETs can typically reach beyond 1MV/cm, we expect it could impact substantially the intrinsic trap characteristics and subsequently the corresponding TAT currents. In this chapter, we therefore present the significance of such field effects on intrinsic trap characteristics and determine the impact on trap emission rates, which itself serves as an initial step to understand the possible impact on the TAT process.

The structure of this chapter is as follows. We start with the theory to extract the trap energy level, which is based on the existing modified transfer matrix (MTM) method of determining spectral states, namely the bound and

the quasi-bound energy states. Although the exact theory of MTM is outlined in Ref.[65], the normalization for the wavefunction is not explicitly mentioned. We therefore rewrite a part of the derivation of MTM followed by the normalization procedure in Section 2.1, which is essential for comparing the probability densities of the wavefunctions at different energies. Section 2.2 details the numerical implementation to extract trap energy levels and to quantify the possible electric field effects. Such field effects on the semiconductor and on the oxide trap configurations are discussed in Section 2.3. Section 2.4 considers the field-induced impact on the emission rates of semiconductor and oxide traps. Finally, we conclude this chapter in Section 2.5.

2.1 Trap energy level extraction

The presence of traps establishes a band discontinuity in the electrostatics, which resembles that of heterostructures. However, such discontinuity is on the sub-microscopic scale. Therefore, the trap system in semiconductor devices can conveniently be constructed as a quantum well in the electrostatic potential. Similar to the determination of spectral states of semiconductor heterostructures, which are the allowed energy states irrespective of their occupation by charge carriers, we will determine the characteristic trap energy levels. Note that the trap energy level is fundamental to the study of TAT. In general, there are three distinguishable spectral (energy) states based on their wavefunction behavior in the infinitely long asymptotic contact regions, namely bound, quasi-bound and free states. The bound energy states are characterized by decaying wavefunctions whose amplitudes are zero in both of the asymptotic contact regions. The quasi-bound states' wavefunctions decay in either of the contacts, whereas the free states' wavefunctions are non-decaying plane wave-like functions in both of the contact regions.

In recent years, numerous methods have been proposed to determine the spectral states of heterostructures. These methods can generally be classified as “large matrix” methods, based on a full system Hamiltonian [66, 67] and “small matrix” methods, involving the elementary transfer-matrices [68]. Among the small matrix methods, the modified transfer matrix (MTM) method is selected for our study of the electric field effect on the trap level in a one dimensional structure [69]. The preference for the MTM method is based on its efficient numerical implementation and handling of the abrupt electrostatic potential transition at a heterointerface, which we will use to configure the trap. Note that the extension of the MTM method to determine the trap level in higher dimensional (2D or 3D) heterostructures is uncertain[70]. For such applications,

the ‘‘large matrix’’ methods can be employed which however require enormous computational resources.

2.1.1 Modified transfer matrix method

In this method, the 1D heterostructure is presumed to consist of l segments (not necessarily of equal lengths) of piecewise constant effective mass $m_{e,l}^*$ as shown in Fig. (2.1). For numerical simplicity, we assume that the wavevector component $\mathbf{k}_\perp(k_y, k_z) = 0$ orthogonal to the transport x -direction at the extreme contact regions, which are also known as the asymptotic regions. Note that $\mathbf{k}_\perp(k_y, k_z) \neq 0$ would provide a different solution depending on the materials used. For the asymptotic regions, the symbols α and β are used (see Fig. (2.1), where the injected right-traveling electron in segment α and the injected left-traveling electron in segment β are indicated with arrows).

Given the previous assumption that $\mathbf{k}_\perp(k_y, k_z) = 0$, the motion of a charged particle through a 1D-heterostructure is described by the following effective-mass Schrödinger equation

$$\left[-\frac{\hbar^2}{2} \frac{\partial}{\partial x} \frac{1}{m_e^*(x)} \frac{\partial}{\partial x} + V(x) \right] \psi(x) = E\psi(x) \quad (2.1)$$

where $V(x)$ represent the potential energy, which includes the external potential in addition to lattice and trap potential, and the effective mass m_e^* changes with the x -coordinate. For a given segment l of constant effective mass $m_{e,l}^*$, the Schrödinger Eq. (2.1) is simplified as:

$$\left[\frac{\hbar^2}{2m_{e,l}^*} \frac{d^2}{dx^2} + E - V_l(x) \right] \psi_l(x, E) = 0 \quad (2.2)$$

where $V_l(x)$ represents the potential energy of the l^{th} segment. As a second order differential equation, Eq. (2.2) can possess two linearly independent basic functions composing the so-called fundamental system of solutions, whereby any physical solution (wavefunction) can be constructed as the linear combination of these basic functions. With the basic functions $\varphi_{l,(1,2)}(x, E)$, the wavefunction over the interval of the l^{th} segment is written as:

$$\psi_l(x, E) = C_{l,1}\varphi_{l,1}(x, E) + C_{l,2}\varphi_{l,2}(x, E); \quad x \in [x_{l-1}, x_l] \quad (2.3)$$

where $C_{l,(1,2)}$ denote the unknown coefficients. With the above definition of wavefunction (Eq. (2.3)), the equation of motion for each segment is solved using the variation of parameter approach, whereby each basic function is constructed

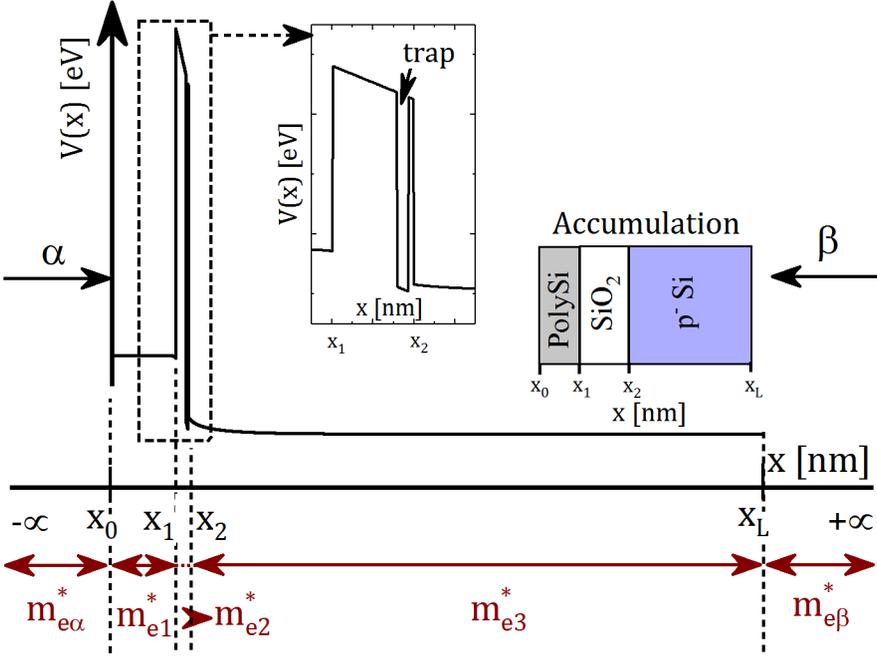


Figure 2.1: An example of a one-dimensional potential of a finite length along the x -axis, which corresponds to the conduction band edge of a MOSCAP biased in accumulation. The inset shows a zoomed-out version of the oxide trap implemented as a square well potential. The asymptotic regions are indicated by α and β , respectively.

as linear combinations of right and left traveling plane-wave type functions and written as:

$$\varphi_{l,\eta}(x, E) = e^{ik_l^0 x} F_{l,\eta}^{(+)}(x, E) + e^{-ik_l^0 x} F_{l,\eta}^{(-)}(x, E); \quad \eta = 1, 2 \quad (2.4)$$

where $F_{l,(1,2)}^{\pm}$ denote the right-traveling (+) and the left-traveling (-) unknown functions and k_l^0 represents the wavenumber of a electron state at energy E ($k_l^0 = \sqrt{2m_{e,l}^* E/\hbar^2}$). With this method, the solution of Schrödinger Eq. (2.2) requires the determination of four unknown functions ($F_{l,(1,2)}^{\pm}$) compared to the initial two unknown basic functions. The following approach in the variation of parameters, also known as the Lagrange condition (also known as the Lagrange method of undetermined multipliers), relates the four unknown functions ($F_{l,(1,2)}^{\pm}$) with the assumption that the first derivative of the basic

functions can be determined from only the derivatives of the exponential factors in Eq. (2.4) [65, 71]. In the framework of the MTM method, this condition using Eq. (2.4) results in:

$$e^{ik_i^0 x} \frac{\partial}{\partial x} F_{l,\eta}^{(+)}(x, E) + e^{-ik_i^0 x} \frac{\partial}{\partial x} F_{l,\eta}^{(-)}(x, E) = 0; \quad \eta = 1, 2 \quad (2.5)$$

With Eqs. (2.2), (2.4) and (2.5), the Lagrange condition is simplified and results in the following set of coupled differential equations for the unknown functions $F_{l,(1,2)}^{(\pm)}$,

$$\begin{aligned} \frac{\partial}{\partial x} F_{l,\eta}^{(+)}(x, E) &= \frac{m_{e,l}^*}{ik_l^0 \hbar^2} e^{-ik_i^0 x} V_l(x) \left[e^{ik_i^0 x} F_{l,\eta}^{(+)}(x, E) + e^{-ik_i^0 x} F_{l,\eta}^{(-)}(x, E) \right]_{\eta=1,2} \\ \frac{\partial}{\partial x} F_{l,\eta}^{(-)}(x, E) &= -\frac{m_{e,l}^*}{ik_l^0 \hbar^2} e^{ik_i^0 x} V_l(x) \left[e^{ik_i^0 x} F_{l,\eta}^{(+)}(x, E) + e^{-ik_i^0 x} F_{l,\eta}^{(-)}(x, E) \right]_{\eta=1,2} \end{aligned} \quad (2.6)$$

Eq. (2.6) represent an initial value problem (IVP).

Note that the above choice of basic functions ($\varphi_{l,(1,2)}(x, E)$) is not unique unlike the basis function in a Hilbert space. Therefore, the following set of boundary conditions, which also determine the nature of wavefunction in the asymptotic regions for a given spectral state, are defined as [65],

$$\begin{aligned} \varphi_{l,1}(x_j, E) = 0 \quad ; \quad \frac{d}{dx} (\varphi_{l,1}(x_j, E)) = 1 \\ ; x_j \in [x_{l-1}, x_l] \quad (2.7) \\ \varphi_{l,2}(x_j, E) = 1 \quad ; \quad \frac{d}{dx} (\varphi_{l,2}(x_j, E)) = 0 \end{aligned}$$

thereby satisfying the following criterion [65],

$$\mathcal{W}(\varphi_{l,1}, \varphi_{l,2}) = -1 \quad (2.8)$$

where \mathcal{W} is the Wronskian of two basic functions of Eq. (2.2) and ensures the linear independence of the two chosen basic functions for all the position points over the interval $(-\infty, +\infty)$ [65].

The initial conditions of the IVP of Eq. (2.6) are determined based on the linearly independent condition of basic solutions (Eq. (2.7)), while applying boundary conditions analogous to the quantum transmitting boundary method (QTBM), which result in the following set of equations:

$$\begin{aligned} F_{l,1}^{(+)}(x_j, E) = \frac{e^{-ik_i^0 x_j}}{2ik_l^0} \quad ; \quad F_{l,1}^{(-)}(x_j, E) = -\frac{e^{ik_i^0 x_j}}{2ik_l^0} \\ F_{l,2}^{(+)}(x_j, E) = \frac{e^{-ik_i^0 x_j}}{2} \quad ; \quad F_{l,2}^{(-)}(x_j, E) = \frac{e^{ik_i^0 x_j}}{2} \end{aligned} \quad (2.9)$$

With the above defined initial boundary conditions (Eqs. (2.9)) and with Eq. (2.6), the unknown functions $\left(F_{l,(1,2)}^{(\pm)}\right)$ of Eq. (2.4) in the l^{th} segment are determined, followed by the calculation of the basic functions itself (Eq. (2.4)). Then, the wavefunctions at an interface x_l between two adjacent segments $(l, l+1)$ are made continuous with the following continuity condition:

$$\psi_l(x, E)|_{x=x_l} = \psi_{l+1}(x, E)|_{x=x_l} \quad (2.10)$$

and their derivatives are related by the so-called Ben-Daniel-Duke-Bastard continuity condition, specifically for heterostructures, defined by

$$\frac{1}{m_{e,l}^*} \frac{d}{dx} (\psi_l(x, E)) \Big|_{x=x_l} = \frac{1}{m_{e,l+1}^*} \frac{d}{dx} (\psi_{l+1}(x, E)) \Big|_{x=x_l} \quad (2.11)$$

Matching the wavefunction in any two adjacent segments $(l, l+1)$ through these two conditions of Eqs. (2.10) and (2.11), a relation between the coefficient pairs $C_{l,\eta}$ of adjacent segments can be determined by transforming the corresponding set of equations into self-contained segment matrices as:

$$\begin{aligned} & \begin{bmatrix} \varphi_{l,1}(x_l, E) & \varphi_{l,2}(x_l, E) \\ \frac{1}{m_{e,l}^*} \varphi'_{l,1}(x_l, E) & \frac{1}{m_{e,l}^*} \varphi'_{l,2}(x_l, E) \end{bmatrix} \begin{bmatrix} C_{l,1} \\ C_{l,2} \end{bmatrix} \\ &= \begin{bmatrix} \varphi_{l+1,1}(x_l, E) & \varphi_{l+1,2}(x_l, E) \\ \frac{1}{m_{e,l+1}^*} \varphi'_{l+1,1}(x_l, E) & \frac{1}{m_{e,l+1}^*} \varphi'_{l+1,2}(x_l, E) \end{bmatrix} \begin{bmatrix} C_{l+1,1} \\ C_{l+1,2} \end{bmatrix} \end{aligned} \quad (2.12)$$

Rewriting Eq. (2.12) as

$$\begin{bmatrix} C_{l,1} \\ C_{l,2} \end{bmatrix} = \mathcal{K}_{l,l+1} \begin{bmatrix} C_{l+1,1} \\ C_{l+1,2} \end{bmatrix} \quad (2.13)$$

results in the so-called segment transfer matrix $\mathcal{K}_{l,l+1}$ described by,

$$\begin{aligned} \mathcal{K}_{l,l+1} &= \begin{bmatrix} \varphi_{l,1}(x_l, E) & \varphi_{l,2}(x_l, E) \\ \frac{1}{m_{e,l}^*} \varphi'_{l,1}(x_l, E) & \frac{1}{m_{e,l}^*} \varphi'_{l,2}(x_l, E) \end{bmatrix}^{-1} \\ &\times \begin{bmatrix} \varphi_{l+1,1}(x_l, E) & \varphi_{l+1,2}(x_l, E) \\ \frac{1}{m_{e,l+1}^*} \varphi'_{l+1,1}(x_l, E) & \frac{1}{m_{e,l+1}^*} \varphi'_{l+1,2}(x_l, E) \end{bmatrix} \end{aligned} \quad (2.14)$$

The elements of the segment transfer matrix are given by Eq. (2.4) and the first derivative of Eq. (2.4) is given by

$$\varphi'_{l,\eta}(x, E) = ik_l^0 \left[e^{ik_l^0 x} F_{l,\eta}^{(+)}(x, E) - e^{-ik_l^0 x} F_{l,\eta}^{(-)}(x, E) \right]; \quad \eta = 1, 2 \quad (2.15)$$

The transfer-matrix, which relates the combination coefficients of the left-most and right-most segments of the physical structure is then described by,

$$\begin{bmatrix} C_{\alpha,1} \\ C_{\alpha,2} \end{bmatrix} = \mathcal{M} \begin{bmatrix} C_{\beta,1} \\ C_{\beta,2} \end{bmatrix} \quad (2.16)$$

where \mathcal{M} represents the modified transfer matrix of the total system and is obtained by the product of the segment matrices as can be understood by a recursive expansion of Eq. (2.13),

$$\mathcal{M} = \mathcal{K}_{\alpha,1} \mathcal{K}_{1,2} \mathcal{K}_{2,3} \dots \mathcal{K}_{L,\beta} \quad (2.17)$$

As mentioned earlier, the spectral states (or quasi-bound states) are characterized by energy values (spectral points) in the complex plane at which the wavefunction has only outgoing waves in the extreme asymptotic regions. Therefore, for the spectral points, the amplitudes of the incoming waves at the far left-most and at the far right-most segments of the potential (see Fig. (2.1)) are zero, which results in the following set of equations:

$$\begin{aligned} C_{\alpha,1} F_{\alpha,1}^{(+)}(-\infty, E) + C_{\alpha,2} F_{\alpha,2}^{(+)}(-\infty, E) &= 0 \\ C_{\beta,1} F_{\beta,1}^{(-)}(+\infty, E) + C_{\beta,2} F_{\beta,2}^{(-)}(+\infty, E) &= 0 \end{aligned} \quad (2.18)$$

From Eq. (2.16), the combination coefficients of the left-most segment are related to the right-most segment of the physical structure by the following expression,

$$C_{\alpha,\eta} = \mathcal{M}_{\eta,1} C_{\beta,1} + \mathcal{M}_{\eta,2} C_{\beta,2}; \quad \eta = 1, 2 \quad (2.19)$$

Using Eq. (2.19), Eq. (2.18) is transformed into matrix form as

$$\mathcal{J}_{\text{sp}}(E) \begin{bmatrix} C_{\beta,1} \\ C_{\beta,2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (2.20)$$

where \mathcal{J}_{sp} represent the Jost matrix, which can determine the physical solution of the structure from the two independent solutions of the asymptotic regions:

$$\mathcal{J}_{\text{sp}}(E) = \begin{bmatrix} \mathcal{J}_{11} & \mathcal{J}_{12} \\ \mathcal{J}_{21} & \mathcal{J}_{22} \end{bmatrix} \quad (2.21)$$

with the following matrix elements:

$$\begin{aligned}
\mathcal{J}_{11} &= \mathcal{M}_{1,1}F_{\alpha,1}^{(+)}(-\infty, E) + \mathcal{M}_{2,1}F_{\alpha,2}^{(+)}(-\infty, E) \\
\mathcal{J}_{12} &= \mathcal{M}_{1,2}F_{\alpha,1}^{(+)}(-\infty, E) + \mathcal{M}_{2,2}F_{\alpha,2}^{(+)}(-\infty, E) \\
\mathcal{J}_{21} &= F_{\beta,1}^{(-)}(+\infty, E) \\
\mathcal{J}_{22} &= F_{\beta,2}^{(-)}(+\infty, E)
\end{aligned} \tag{2.22}$$

From the earlier definition of spectral states and using Eq. (2.20), the spectral points, which also represent the required trap energy level E_t , are the complex roots of the following equation:

$$||\mathcal{J}_{\text{sp}}(E)|| = 0 \tag{2.23}$$

Once the energy E_t is extracted, the corresponding wavefunction can also be determined. Therefore, the state at the edge of an asymptotic region is fixed by choosing the amplitudes of the left-most right-traveling and the right-most left-traveling components of the wavefunction. In general, these amplitudes are defined by

$$\begin{aligned}
C_{\alpha,1}F_{\alpha,1}^{(+)}(x_0, E) + C_{\alpha,2}F_{\alpha,2}^{(+)}(x_0, E) &= I_{\alpha} \\
C_{\beta,1}F_{\beta,1}^{(-)}(x_L, E) + C_{\beta,2}F_{\beta,2}^{(-)}(x_L, E) &= I_{\beta}
\end{aligned} \tag{2.24}$$

Using Eq. (2.19), this condition can be expressed in matrix form as:

$$\mathcal{J}_{\text{in}}(E) \begin{bmatrix} C_{\beta,1} \\ C_{\beta,2} \end{bmatrix} = \begin{bmatrix} I_{\alpha} \\ I_{\beta} \end{bmatrix} \tag{2.25}$$

where the Jost-matrix \mathcal{J}_{in} is identical to \mathcal{J}_{sp} . In Eq. (2.25), the injection of a state at a specific asymptotic region is described by setting the amplitude of the other asymptotic region to zero. For instance, the left-injection type is defined by ($I_{\alpha} = 1, I_{\beta} = 0$) in Eq. (2.25) and vice versa for the right-injection type. Note that the bound state energy E_t extraction is independent of the choice of the injection type, which means the trap energy level E_t is identical for both left- and right-injected wavefunctions. Moreover, once the energy E_t is known, the right-injected wavefunction (in this thesis) is determined by first constructing the Jost-matrix \mathcal{J}_{in} with the substitution of $E = E_t$ in the set of Eqs. (2.22), followed by determining the coefficients ($C_{\beta,1}, C_{\beta,2}$) for the right-most segment β from Eq. (2.25) (with $I_{\alpha} = 0, I_{\beta} = 1$). Subsequently, the coefficients for the next left segments are determined, while using Eq. (2.13), all the way to the left-most segment α . Meanwhile, the unknown functions $F_{l,\eta}^{(+)}(x, E)$ of each l^{th}

segment are determined (with the substitution of $E = E_t$) from Eq. (2.6) by using the boundary conditions of Eq. (2.9), followed by determining the shape of the basic functions (from Eq. (2.4)). Finally, the complete wavefunction at energy E_t is constructed using the formerly determined coefficients and basic functions. So far, we showed that the solution of Eq. (2.25) at the energy E_t determined from Eq. (2.23) with the condition ($I_\alpha = 0, I_\beta = 1$) gives the wavefunction assuming injection from the right contact. However, Eq. (2.23) can only be properly solved for a bound state. Hence a normalization procedure is required, which is presented in the next subsection.

2.1.2 Wavefunction normalization

The next essential step required for the characterization of the trap level is the wavefunction normalization, which will allow to compare wavefunction amplitudes at different energies, and which will be needed to determine the broadening of the trap level E_t . We will start with imposing the delta normalization condition as it conforms with the theory of the MTM method [65]. We make use of the asymptotic conditions of the MTM theory, where we assume that the potential profile in the contacts is a segment of constant potential energy $V_{\alpha,\beta}$. For each of these regions, the momentum is determined, while considering the constant values of the potential and the effective mass, by

$$k_l = \frac{\sqrt{2m_{e,l}^*(E - V_l)}}{\hbar}; l = \begin{cases} \alpha & x \in (-\infty, x_0) \\ \beta & x \in [x_L, +\infty) \end{cases} \quad (2.26)$$

with x_L the start of the right-most segment. Here we repeat the procedure outlined in Section 2.1.1. Using Eq. (2.26), the Schrödinger Eq. (2.2) of each segment in the outer-most asymptotic regions is simplified as

$$\left[\frac{d^2}{dx^2} + k_l^2 \right] \psi_l(x, E) = 0; l = \alpha, \beta \quad (2.27)$$

where Eq. (2.27) is a second order differential equation which can be solved by considering the wavefunction as a linear combination of two independent basic solutions (similar to Eq. (2.3)).

$$\psi_l(x, E) = C_{l,1}\varphi_{l,1}(x, E) + C_{l,2}\varphi_{l,2}(x, E); l = \alpha, \beta \quad (2.28)$$

where, in alignment with the asymptotic conditioning of Eq. (2.3), the basic solutions in Eq. (2.28) are defined by,

$$\varphi_{l,\eta}(x, E) = e^{ik_l x} F_{l,\eta}^{(+)}(E) + e^{-ik_l x} F_{l,\eta}^{(-)}(E); l = \alpha, \beta; \eta = 1, 2 \quad (2.29)$$

Note the introduction of $F_{l,\eta}^{(\pm)}$ in Eq. (2.29) as position independent in comparison with the unknown functions of Eq. (2.4), which is in alignment with the MTM theory. In particular, when the potential in the asymptotic regions is unvarying, the $F_{l,\eta}^{(\pm)}$ in Eq. (2.6) become constant as their partial derivative is zero, thereby resulting in energy-dependent constant factors. The linear independence of the basic solutions in Eq. (2.28) is ensured by similar conditions as in Eq. (2.7),

$$\begin{aligned} F_{l,1}^{(+)}(E) &= \frac{e^{-ik_l x_j}}{2ik_l} & ; & \quad F_{l,1}^{(-)}(E) = -\frac{e^{ik_l x_j}}{2ik_l} \\ & & & ; x_j = \begin{cases} x_0, & l = \alpha \\ x_L, & l = \beta \end{cases} \end{aligned} \quad (2.30)$$

$$F_{l,2}^{(+)}(E) = \frac{e^{-ik_l x_j}}{2} & ; & \quad F_{l,2}^{(-)}(E) = \frac{e^{ik_l x_j}}{2}$$

To this end, the system of equations for the outer-most asymptotic regions, which are characterized by a constant potential and effective mass are defined. In the subsequent subsections, the normalization methods essential to obtain physical wavefunctions (non-spurious) are outlined.

Delta normalization method

Due to the plane wave nature of the wavefunction in the asymptotic regions, the method of delta normalization is imposed, which is described by the following condition:

$$\int_{-\infty}^{+\infty} dx \psi^*(x, E') \psi(x, E) = 2\pi\delta(E' - E) \quad (2.31)$$

where for a given energy the specific asymptotic region wavefunction is described by the type of injection (see Eq. (2.25)). For instance, for a left-injection type state, the asymptotic region wavefunction is a linear combination of the right/left traveling left-most (α) and the right-traveling right-most (β) basic functions. In general, such description of the wavefunction is given by

$$\begin{aligned} \psi(x, E) = \mathcal{N}_d \sum_{\eta=1}^2 \Theta_l(x) C_{l,\eta} \left[e^{ik_l x} F_{l,\eta}^{(+)}(E) + e^{-ik_l x} F_{l,\eta}^{(-)}(E) \right] \\ + \Theta_{l'}(x) C_{l',\eta} e^{\pm ik_{l'} x} F_{l',\eta}^{(\pm)}(E) \end{aligned} \quad (2.32)$$

where \mathcal{N}_d is the normalization constant, which we are trying to determine. In the last term of Eq. (2.32) (which defines the transmission part of the wavefunction),

the upper (+) sign indicates the left-type injection whereas the lower (−) sign specifies the right-type injection. There must either be an upper (+) or lower (−) sign related term (correspondingly with either $+k_{l'}$ or $-k_{l'}$ in the exponential function) present in the wavefunction description Eq. (2.32) depending on the type of injection, which is described by the following notations:

$$\begin{aligned} l = \alpha &\iff l' = \beta; && \text{left injection} \\ l = \beta &\iff l' = \alpha; && \text{right injection} \end{aligned} \quad (2.33)$$

In Eq. (2.32), $\Theta_{\alpha,\beta}$ are the unit step-functions defined in accordance with the piecewise-description of the wavefunction in Eq. (2.32) and are given by,

$$\Theta_{\alpha}(x) = \begin{cases} 1, & x \in (-\infty, x_0] \\ 0, & x > x_0 \end{cases} \quad ; \quad \Theta_{\beta}(x) = \begin{cases} 1, & x \in [x_L, +\infty) \\ 0, & x < x_L \end{cases} \quad (2.34)$$

Substitution of Eq. (2.32) into the left hand side of Eq. (2.31) yields,

$$\begin{aligned} \int_{-\infty}^{+\infty} dx \psi^*(x, E') \psi(x, E) &= |\mathcal{N}_d|^2 \int_{-\infty}^{+\infty} dx \left[\sum_{\eta, \eta'} (\Theta_l(x))^2 C_{l, \eta'}^* C_{l, \eta} \right. \\ &\times \left\{ F_{l, \eta'}^{(+)*}(E') F_{l, \eta}^{(+)}(E) e^{i(k_i - k'_i)x} + F_{l, \eta'}^{(+)*}(E') F_{l, \eta}^{(-)}(E) e^{i(-k_i - k'_i)x} \right. \\ &\quad \left. + F_{l, \eta'}^{(-)*}(E') F_{l, \eta}^{(+)}(E) e^{i(k_i + k'_i)x} + F_{l, \eta'}^{(-)*}(E') F_{l, \eta}^{(-)}(E) e^{i(k'_i - k_i)x} \right\} \\ &\left. + (\Theta_{l'}(x))^2 C_{l', \eta'}^* C_{l', \eta} F_{l', \eta'}^{(\pm)*}(E') F_{l', \eta}^{(\pm)}(E) e^{\pm i(k_{l'} - k'_{l'})x} \right] \quad (2.35) \end{aligned}$$

In Eq. (2.35), the cross terms related to $\Theta_l(x) \Theta_{l'}(x)$ vanish while making use of the definitions of the step function in Eq. (2.34). With the following approximation for the x -integral, where we assumed that the contribution of $\int_{x_0}^{x_L} dx$ is negligible in determining the normalization factor, and subsequently

using the property of the delta function,

$$\int_{-\infty}^{+\infty} dx (\Theta_l(x))^2 e^{i(k_l - k'_l)x} \approx \left[\int_{-\infty}^{x_0} dx + \int_{x_L}^{+\infty} dx \right] (\Theta_l(x))^2 e^{i(k_l - k'_l)x} \quad (2.36)$$

$$\equiv \int_{-\infty}^{x_0} dx e^{i(k_l - k'_l)x} \approx \pi \delta(k_l - k'_l) \quad (2.37)$$

$$\int_{-\infty}^{+\infty} dx (\Theta_{l'}(x))^2 e^{\pm i(k_{l'} - k'_{l'})x} \approx \left[\int_{-\infty}^{x_0} dx + \int_{x_L}^{+\infty} dx \right] (\Theta_{l'}(x))^2 e^{\pm i(k_{l'} - k'_{l'})x} \quad (2.38)$$

$$\equiv \int_{x_L}^{-\infty} dx e^{\pm i(k_{l'} - k'_{l'})x} \approx \pi \delta(k_{l'} - k'_{l'}) \quad (2.39)$$

Eq. (2.35) reduces to

$$\begin{aligned} \int_{-\infty}^{+\infty} dx \psi^*(x, E') \psi(x, E) = \\ \pi |\mathcal{N}_d|^2 \left[\sum_{\eta, \eta'} C_{l, \eta'}^* C_{l, \eta} \left\{ F_{l, \eta'}^{(+)*}(E') F_{l, \eta}^{(+)}(E) \delta(k_l - k'_l) \right. \right. \\ \left. \left. + F_{l, \eta'}^{(+)*}(E') F_{l, \eta}^{(-)}(E) \delta(-k_l - k'_l) + F_{l, \eta'}^{(-)*}(E') F_{l, \eta}^{(+)}(E) \delta(k_l + k'_l) \right. \right. \\ \left. \left. + F_{l, \eta'}^{(-)*}(E') F_{l, \eta}^{(-)}(E) \delta(k'_l - k_l) \right\} + C_{l', \eta'}^* C_{l', \eta} F_{l', \eta'}^{(\pm)*}(E') F_{l', \eta}^{(\pm)}(E) \delta(k_{l'} - k'_{l'}) \right] \quad (2.40) \end{aligned}$$

It should be noted that the wavefunction definition (Eq. (2.32)) is unreasonable over the device length (from x_0 to x_L) as the potential energy is no longer constant due to the externally applied electric field (see Fig. (2.1)). We therefore disregard the wavefunction normalization over the device length (from x_0 to x_L) in Eqs. (2.36) and (2.38). Remembering from Eqs. (2.29) and (2.32) that the independent wavenumbers $k_l > 0$ and $k'_l > 0$. Therefore, the terms with $-k_l - k'_l$ and $k_l + k'_l$ vanish due to:

$$\delta(-k_l - k'_l) = \delta(k_l + k'_l) = 0 \quad (2.41)$$

The above result brings Eq. (2.40) to

$$\int_{-\infty}^{+\infty} dx \psi^*(x, E') \psi(x, E) = \pi |\mathcal{N}_d|^2 \left[\sum_{\eta, \eta'} C_{l, \eta'}^* C_{l, \eta} \left\{ F_{l, \eta'}^{(+)*}(E') F_{l, \eta}^{(+)}(E) + F_{l, \eta'}^{(-)*}(E') F_{l, \eta}^{(-)}(E) \right\} \delta(k'_l - k_l) + C_{l', \eta'}^* C_{l', \eta} F_{l', \eta'}^{(\pm)*}(E') F_{l', \eta}^{(\pm)}(E) \delta(k'_{l'} - k_{l'}) \right] \quad (2.42)$$

The independent wavenumbers $k_{l, l'}$ and $k'_{l, l'}$ are energy dependent and therefore the delta functions in Eq. (2.42) can be rewritten as

$$\delta(k'_{l, l'} - k_{l, l'}) = \delta(k_{l, l'}(E') - k_{l, l'}(E)) \quad (2.43)$$

By using the following property of the delta function,

$$\delta(k_{l, l'}(E') - k_{l, l'}(E)) = \frac{\delta(E' - E)}{\left| \frac{\partial k_{l, l'}(E')}{\partial E'} \right|_{E'=E}} = \delta(E' - E) \frac{\hbar^2 k_{l, l'}}{m_{e, (l, l')}^*} \quad (2.44)$$

The probability density of a wavefunction is determined by

$$\int_{-\infty}^{+\infty} dx \psi^*(x, E') \psi(x, E) = \pi \delta(E' - E) |\mathcal{N}_d|^2 \left[\sum_{\eta, \eta'} C_{l, \eta'}^* C_{l, \eta} \left\{ F_{l, \eta'}^{(+)*}(E) F_{l, \eta}^{(+)}(E) + F_{l, \eta'}^{(-)*}(E) F_{l, \eta}^{(-)}(E) \right\} \frac{\hbar^2 k_l}{m_{e, l}^*} + C_{l', \eta'}^* C_{l', \eta} F_{l', \eta'}^{(\pm)*}(E) F_{l', \eta}^{(\pm)}(E) \frac{\hbar^2 k_{l'}}{m_{e, l'}^*} \right] \quad (2.45)$$

Comparing the right hand side of Eq. (2.45) with that of the normalization condition Eq. (2.31), the normalization factor is thus obtained by

$$|\mathcal{N}_d| = \frac{1}{\sqrt{\frac{1}{2} \sum_{\eta, \eta'} \left[C_{l, \eta'}^* C_{l, \eta} \left\{ F_{l, \eta'}^{(+)*}(E) F_{l, \eta}^{(+)}(E) + F_{l, \eta'}^{(-)*}(E) F_{l, \eta}^{(-)}(E) \right\} \frac{\hbar^2 k_l}{m_{e,l}^*} \right.} + C_{l', \eta'}^* C_{l', \eta} F_{l', \eta'}^{(\pm)*}(E) F_{l', \eta}^{(\pm)}(E) \frac{\hbar^2 k_{l'}}{m_{e,l'}^*} \left. \right]} \quad (2.46)$$

Note that the above normalization method of wavefunction works well for bound states. For quasi-bound states, this normalization can not resolve the problem of unphysical solutions appearing due to the imaginary component of the calculated energies. We therefore propose an alternative approach to circumvent this problem, which is described in the subsequent section.

Constant density normalization method

In the previous section, we determined the normalization of the trap wavefunction. However, for quasi-bound or free states the energies associated with the trap are points in a complex plane, where the complex components are related to the life times and determine the energy-width of the line shape function of a resonant state. However, this complex component can result in an unphysical exponential modulation of the plane wave amplitude of the wavefunction in the asymptotic regions. We therefore, propose a constant density normalization to mitigate such unphysical solutions.

In this normalization approach, we first artificially replace the imaginary component of the determined trap energies with zero, which will circumvent the exponential modulation due to the imaginary component. Consequently, we assume that the $E - k$ relation in the asymptotic region is linear in the range of trap levels of interest, to obtain a reasonable comparison of the probability densities at different energies, which is needed for quantifying the broadening of a trap state discussed in the next section. With the previous assumption that $\mathbf{k}_\perp(k_y, k_z) = 0$, and the typical proportionality $|E - E_n| \sim k^2$ (with E_n a band extremum), this hypothesis suggests that the energy window of interest is far away from the material's band edge in the lower-potential asymptotic region. Under these assumptions, the state density at different energies is constant in the lower-potential contact region. Therefore, the wavefunction densities are normalized to this constant state density and we term such normalization as constant density normalization method. Later, it has been verified that when using this approach, the broadness of the peaks varies smoothly and continuously, in a physically consistent way, which justifies the procedure used. Note that

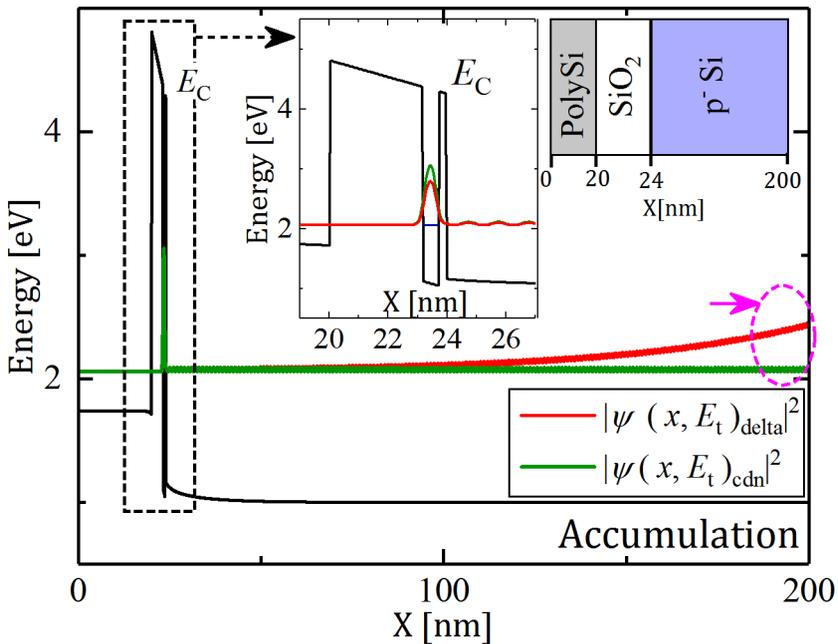


Figure 2.2: The electrostatics of a MOSCAP biased in accumulation shown with an inset of a zoomed-in version of the oxide trap configuration. The wavefunctions are shifted along the y-axis for the illustrative purpose of showing the behavior of the oxide trap wavefunction in the trap region and in the asymptotic regions. The unphysical exponential growth of the wavefunction (red) obtained with the delta normalization is indicated by a magenta circle, where as the wavefunction obtained with the constant density normalization is indicated with green.

this constant density normalization does not provide an absolute normalization, but rather a relative comparison between wavefunctions at energies sufficiently far from the band edge.

Application of the normalization methods

The differences in both of these normalization methods is depicted in Fig. (2.2), which compares the wavefunctions normalized with the aforementioned methods. For a delta-type normalized wavefunction, the unphysical exponential rise in the amplitude at the extreme contacts (asymptotic

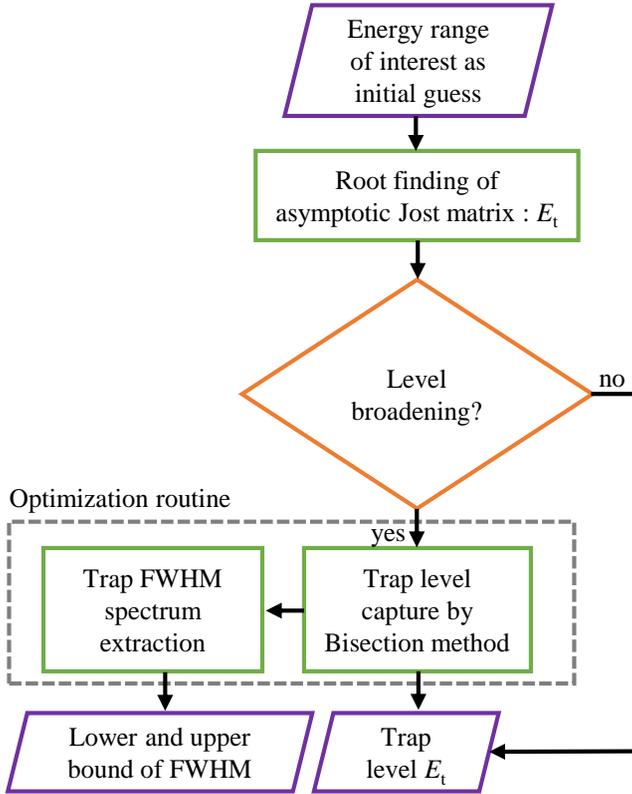


Figure 2.3: The flowchart depicting the numerical procedure of finding the trap level and FWHM spectral range.

regions), indicated by the magenta circle, is due the imaginary component in the trap energy E_t . However, the constant density normalization is able to circumvent such unphysical solutions, which can be seen with a constant wavefunction (green colored) amplitude at the contacts, particularly in the lower potential region of the far right contact.

2.2 Numerical procedure

In the previous section, we described the theory of determining the trap energy level and the normalized trap wavefunctions. To estimate these physical solutions, we propose a numerical implementation procedure described by the flowchart in Fig. (2.3). We start by a discretization of the 1D heterostructure using a finite difference (FD) scheme, in which the potential energy is replaced by a combination of the trap potential $V_t(x)$ and the electric field F ($V(x) = V_t(x) - e_q F(x)x$). In our study, we limit ourselves to the different cases of a semiconductor trap, which is configured with a one dimensional single quantum well having the same effective mass as the surrounding semiconductor, while a MOS capacitor (MOSCAP) system is used for analyzing oxide traps. The traps are assumed to be invariant planar structures in the other two (y, z)-dimensions. The initial set of conditions for the unknown functions $F_{t,(1,2)}^\pm$ are found in Eqs. (2.9). These initial conditions are used to determine the basic functions, while making use of the IVP of Eqs. (2.6), by applying the Runge-Kutta method of solving coupled differential equations. Subsequently, the basic functions are used to construct the modified transfer (\mathcal{M}) and the spectral Jost (\mathcal{J}_{sp}) matrices.

The energy range of interest, which covers the entire range of the electrostatic potential of a trap configuration, is an input to the solver. The characteristic roots of the asymptotic Jost matrix system (\mathcal{J}_{sp}) can efficiently be determined from Eq. (2.23) using the iterative Newton-Raphson method and represent the trap energy level E_t (see Figs. (2.2) and (2.4)). The correctness of the proposed numerical scheme is in agreement with the spectral states and their corresponding wavefunctions (without applying any normalizations) of the heterostructures described in literature [66, 69].

The broadening of a spectral state at high electric field is anticipated. The solver is therefore extended with an optimization routine to determine the broadening of the trap energy state at high fields. The solver requests for this routine only if the maximum probability density of the wavefunction within the trap region has reduced with less than 50% at 1 μeV from the extracted energy E_t .

In the optimization routine, a comparison is made between the wavefunction probability density peak inside the trap regions. The energy interval corresponding to full-width half-maximum (FWHM) is determined. During this optimization routine, the constant density normalization method is employed to compare the probability densities for a quasi-bound system. For a bound system, the delta normalization is used.

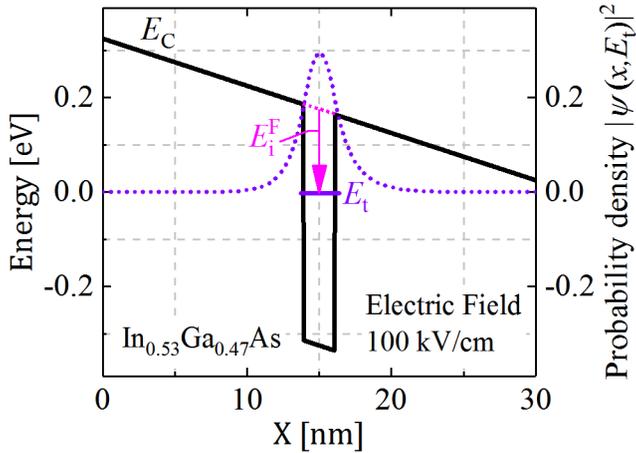


Figure 2.4: The electrostatic potential (solid black), the characteristic level ($E_t \approx -0.002$ eV) and the corresponding wavefunction of a square well 2 nm-wide trap configuration in a 30 nm long $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ region at 100 kV/cm electric field strength. E_i^F is the energetic difference taken between E_C and E_t at the trap center.

2.3 Electric field effects: Level shift and broadening

In this section, we apply the aforementioned method of determining the trap level (see Sections 2.1 and 2.2) in different cases of semiconductor and oxide traps. Further, we examine the electric field-induced effects of a trap level.

2.3.1 Semiconductor traps

The effect of a TFET-like source-channel junction field on the trap is examined with different quantum well configurations in a uniform electric field. The characteristic bound state and the corresponding wavefunction of the square well (SW) trap structure at low uniform electric field ($F = 100$ kV/cm) is illustrated in Fig. (2.4). A 30 nm long section of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ is considered. The arbitrarily chosen well depth and width are 0.5 eV and 2 nm, at zero-field conditions, respectively. The first bound state of the system is at 169 meV from

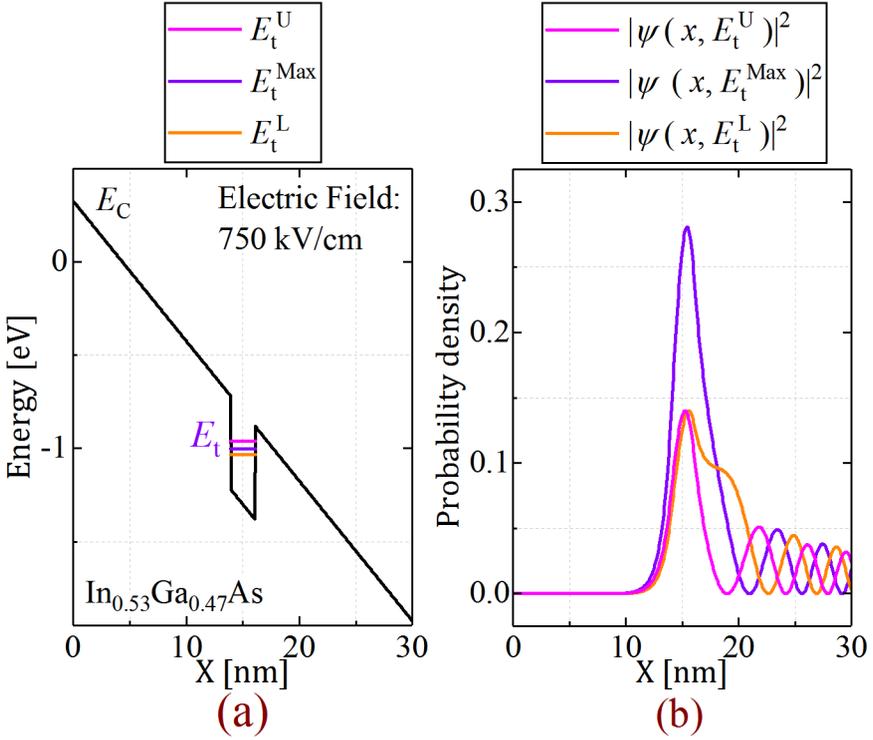


Figure 2.5: (a) The electrostatic potential and the characteristic sub-levels specifying the lower and the upper bounds of full-width half-maximum spectral range ($E_t^{\text{Max}} \approx -1.003$ eV, $E_t^L \approx -1.034$ eV and $E_t^U \approx -0.962$ eV), (b) The wavefunctions corresponding to the trap levels of (a), for a 2 nm-wide SW trap in a 30 nm long section of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$.

the top of the conduction band ($E_i=169$ meV) and could describe a shallow acceptor-like trap.

The sharp low-field energy level depicted in Fig. (2.4) spreads into a spectral range $[E_t^L, E_t^U]$ at 750 kV/cm field strength as illustrated in Fig. (2.5)(a). The densities in correspondence with the levels E_t^{Max} , E_t^L and E_t^U are shown in Fig. (2.5)(b). It is apparent from the comparison of Figs. (2.4) and (2.5) that this spread in spectral states can be associated with the tunneling of the trap wavefunction into the adjacent lower potential region (right side in Fig. (2.5)).

The electric field is further varied from 0 to 3 MV/cm for the SW trap configuration in Fig. (2.4). The resultant plots show the trap level shifts against

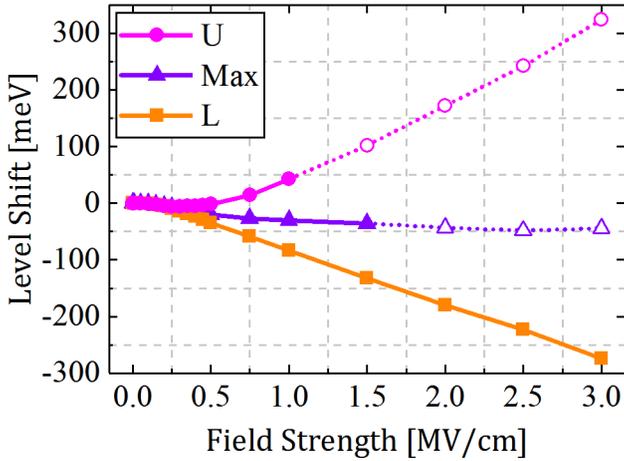


Figure 2.6: Trap level shift and broadening determined as a function of electric field strength for the 2 nm-wide SW of Fig. (2.4). The closed and the open symbols represents the localization and the non-localization of the trap-level E_t in the SW trap region.

the respective fields in Fig. (2.6), where the level shift is calculated as the difference between the zero-field trap level E_i and the non-zero field trap level E_i^F . From Fig. (2.4), the trap level E_i^F is the energetic difference between E_C and E_t , determined at the trap center. The shift of the upper bound of the FWHM spectrum, most probable level (Max) and the lower bound of the FWHM spectrum (L) and (U) are also shown in Fig. (2.6).

The level broadening for the above trap configuration increases with electric field (see Fig. (2.6)), which is attributed to a reduction in potential barrier and therefore an increase in tunneling-of trap wavefunction into the continuum states of the lower potential region. Moreover, the curve symbols change from closed to open when the probability density of the level becomes marginally peaked (less than or equal to) in the trap SW compared to the probability density peaks outside of the trap SW, which occurs when no energy barrier is left at one side of the trap and hence, the wavefunction becomes non-local. Note that the maximum density energy level need not be at the center of its FWHM spectrum, as it is apparent from Fig. (2.6).

The comparison of the level shift lines for a reduced width SW and a Coulomb well (CW, which represents the configuration of the trap system in semiconductor devices by a Coulomb potential) with the previously defined 2nm

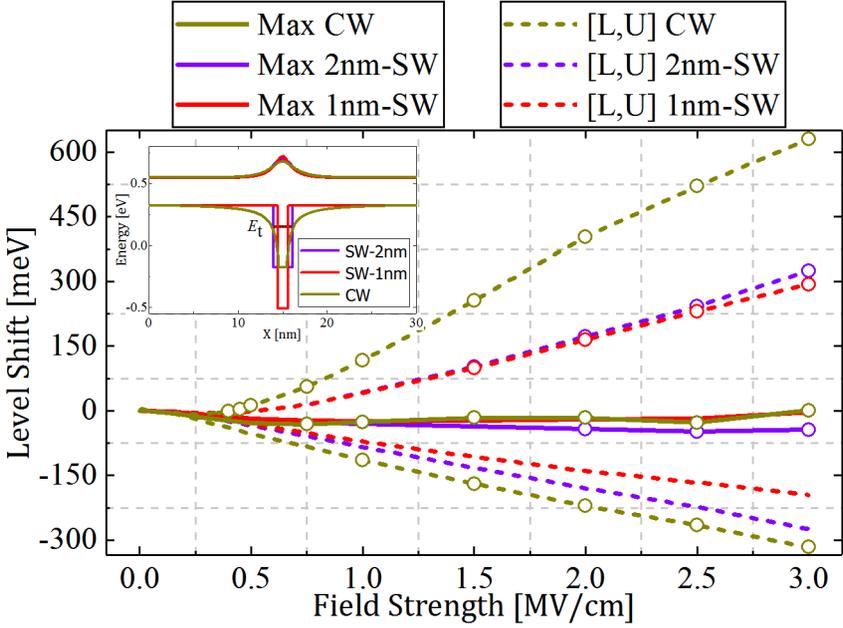


Figure 2.7: *Electric field-induced level shift and broadening for the 2 nm-wide SW, 1 nm-wide SW and the CW trap configurations. The open symbols represents the non-local trap states. The inset depicts the described trap configurations at zero-field condition and the wavefunction probability densities for the indicated trap levels E_t .*

wide SW (Fig. (2.4)) is illustrated in Fig. (2.7). Note that the energetic depth E_t^F increases (see Fig. (2.7)), when both the depth and the width of the trap well increases. The energetic distance from the quantized energy level E_t (in the inset of Fig. (2.7)) to the bottom of the well increases with decreasing trap width and increasing trap depth. However, the different trap configurations in Fig. (2.7) are defined such that their zero-field trap energy levels E_t are identical (see the inset of Fig. (2.7)). It is evident that the level shift and the level broadening depend on the barrier type. The 1 nm-wide trap has an increased effective barrier width and therefore, a smaller level shift is noticed than the 2 nm-wide trap. The Coulomb trap encounters a hyperbolic barrier which is smoother than the triangular barrier of SWs in an electric field. This results in a reduction of the effective barrier and hence, it shows a larger level broadening than that of a SW. The impact of the well-known barrier lowering (Poole-Frenkel effect) [47] can also be noticed as the early out-shift (open symbols) of the CW trap level compared to the SW level, which is apparent from Fig. (2.7). Therefore, the

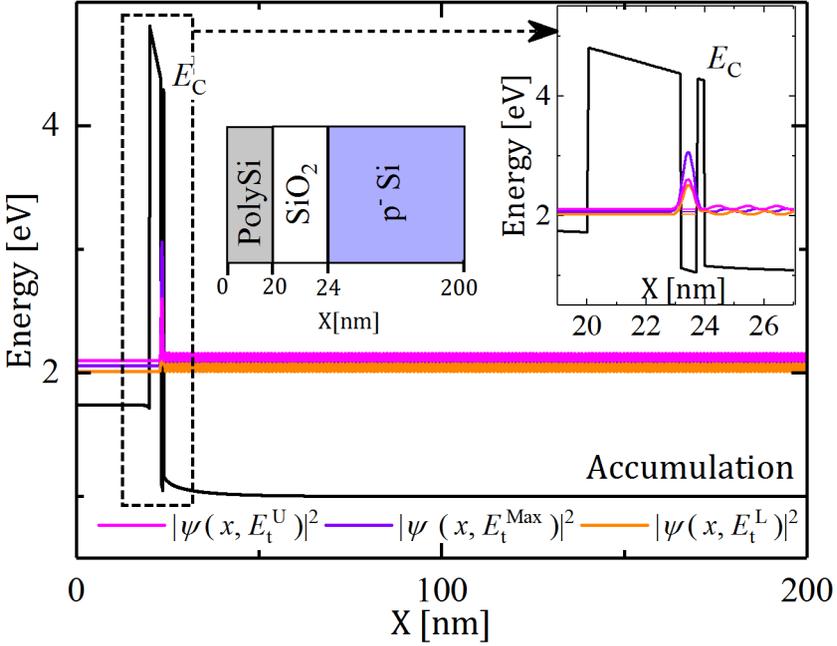


Figure 2.8: A PolySi/SiO₂/p⁻ Si MOSCAP with an oxide trap in accumulation with an inset of a zoomed-in view of the conduction band energy. The oxide trap is configured as a 0.5 nm wide and 3.55 eV deep SW and is located at 0.25 nm from the SiO₂/p⁻ Si interface. The trap exhibits broadening as shown with the varying probability densities in the trap region.

specific choice of the 1nm³ wide SW trap potential [72] in device electrostatics (see NEGF based TAT model in Section 1.5.1) may have a quantitative impact on the predicted TAT current.

2.3.2 Oxide traps

The numerical procedure of the trap spectrum extraction (Fig. (2.3)) is now employed to find the intrinsic characteristics of oxide traps in a MOSCAP system in the presence of an electric field. In particular, we examine the traps in PolySi/SiO₂/p⁻ Si and Al/HfO₂/p⁻ In_{0.53}Ga_{0.47}As MOS heterostructures. In these structures, the substrate is doped with 10¹⁷ at/cm³ p-type concentration, while the substrate depth and the oxide thickness are 175 nm and 4 nm, respectively (see Figs. (2.8) and (2.9)). The trap in the oxide is configured with

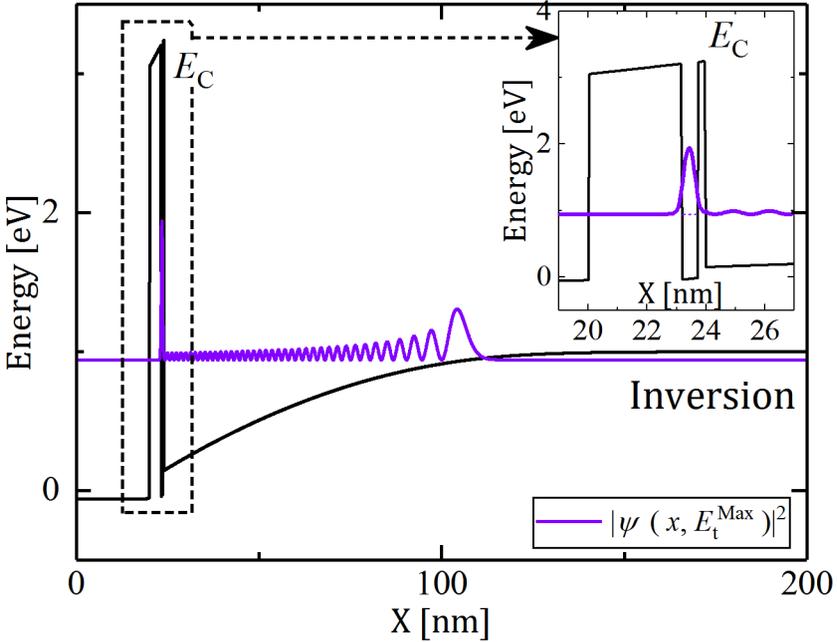


Figure 2.9: A PolySi/SiO₂/p⁻Si MOSCAP with an oxide trap in inversion with an inset of a zoomed-in view of the conduction band energy. The oxide trap is configured as a 0.5 nm wide and 3.55 eV deep SW and is located at 0.25 nm from the SiO₂/p⁻Si interface. The trap state is a fully localized bound state. The higher amplitudes of the wavefunction, right before entering into the conduction band, could probably be related to the interference process occurring between the incident and the reflected elements of the wavefunction.

a constant 0.5 nm wide SW. These MOS structures are biased from accumulation (Fig. (2.8)) to inversion (Fig. (2.9)).

In the PolySi/SiO₂/p⁻Si MOSCAP system, the 0.5 nm wide SW type trap configuration is located at 1 nm from the SiO₂/p⁻Si interface in the SiO₂ (see Fig. (2.8)). The depth of the well is varied from 2.05 eV to 3.55 eV with reference to the SiO₂ conduction band edge. This combination of the MOS and the oxide trap structure is subjected to external applied biases. The results of field-induced trap level shifts and broadenings are plotted against the oxide field strength in Fig. (2.10). It is apparent from this figure that the level shifts and level broadenings are low in comparison to those of the semiconductor traps of Fig. (2.7). This is due to the higher effective mass and barrier heights in

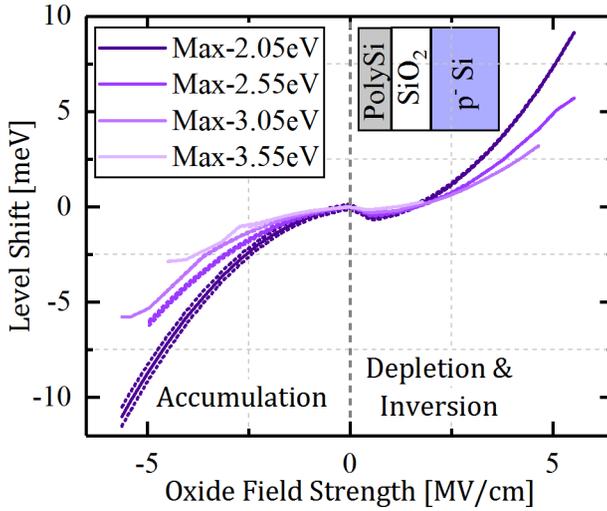


Figure 2.10: *Level shift (solid) and broadening (dashed) as a function of oxide electric field strengths for the 0.5 nm-wide SW oxide-trap configuration in the PolySi/SiO₂/Si MOSCAP system. The traps are configured with varying well-depth and are positioned at 1 nm from the oxide-substrate interface.*

the oxide ($E_i > 2$ eV, $m_{\text{SiO}_2}^* \approx 0.5m_0$) [73] compared to the semiconductor trap configurations ($E_i = 0.17$ eV, $m_{\text{InGaAs}}^* \approx 0.043m_0$).

The levels shift in opposite direction to inversion, when the MOSCAP is biased in accumulation. The combined effects of the lowering of the substrate conduction band minimum (CBM) and lowering of the effective barrier height determine the negative level shift with increasing accumulation, which is shown in Fig. (2.8). In contrast, the rise of the substrate CBM and the barrier height (see Fig. (2.9)), results in the positive level shift of the trap level with increasing depletion. The data in Fig. (2.10) also show that the level broadening exhibits an inverse dependence with the depth of the trap level, which is obvious since deeper traps allow for less leakage.

The impact of the location of the trap with respect to the oxide-substrate interface is presented in Fig. (2.11). The trap is configured with a constant 0.5 nm wide and 3.05 eV deep SW inside SiO₂. The position of the trap is varied in the range of 1.75 nm-0.25 nm from the SiO₂/p-Si interface. It is apparent from Fig. (2.11) that the level shifts and broadenings depend on the tunneling of its wavefunction through the tunneling barrier: these field-induced quantum effects are more pronounced when the trap is positioned closer to the

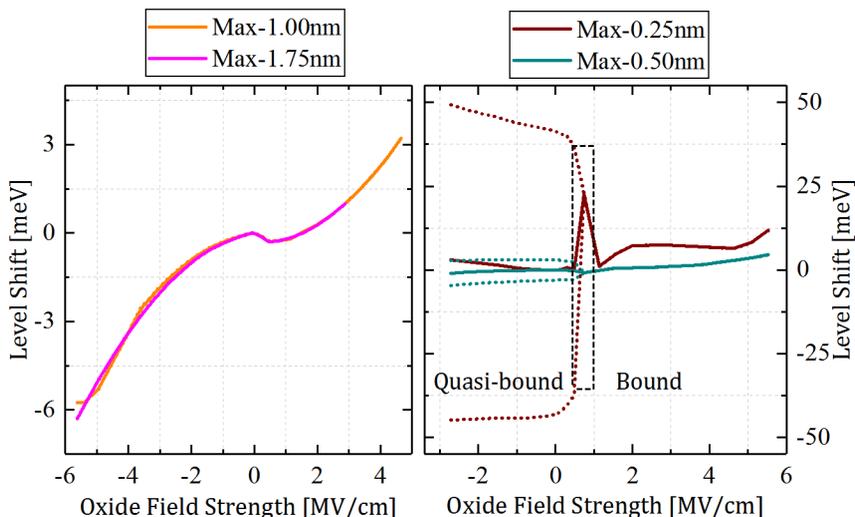


Figure 2.11: Level shift (solid) and broadening (dashed) as a function of oxide electric field strengths for the 0.5 nm-wide 3.05 eV deep SW oxide-trap configuration in the PolySi/SiO₂/Si MOSCAP system. The trap-level is at around 1 eV above the substrate conduction band edge at flat-band voltage condition. The traps are positioned at varying distance from the oxide-substrate interface. The traps near to the oxide-substrate interface show significant level broadening. The black dashed box indicates the transition from quasi-bound to bound systems.

SiO₂/p⁻Si interface.

Note that these traps are configured such that the trap level is fixed to be around 1 eV above the substrate CBM at zero-field condition and could correspond to trap energy levels related to the neutral-type three fold-Silicon defect pair found in SiO₂ [74]. As the bias brings the MOSCAP more in depletion, the trap level moves closer to and eventually below the substrate CBM, as can be seen from Fig. (2.9). From the latter bias condition on, the trap level is fully localized (from a quasi-bound to a bound state as indicated in Fig. (2.11)) and the level broadening disappears. The level values in the quasi-bound to bound transition region indicated with a black dashed box are not reliable, as the approximation of fixed state density (constant density normalization at the lower potential contact region) no longer holds close to the conduction band edge. Hence, the comparison between the trap energy levels is

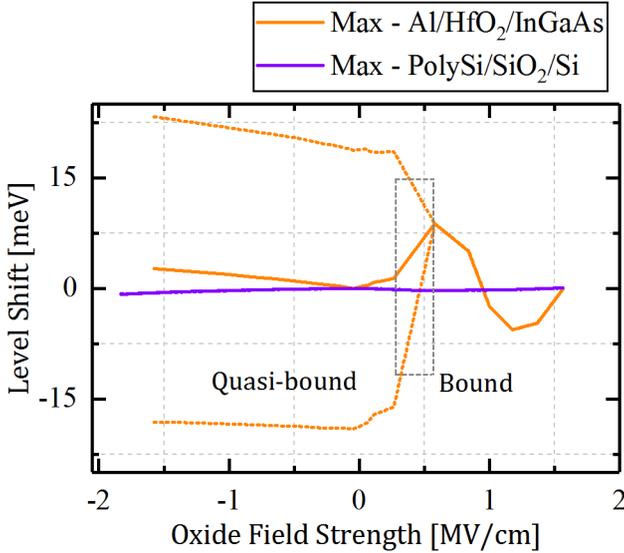


Figure 2.12: *Level shift (solid) and broadening (dashed) as a function of oxide electric field strengths for the 0.5 nm-wide SW oxide-trap configuration in PolySi/SiO₂/Si and Al/HfO₂/InGaAs MOSCAP systems. The traps are configured with 3.05 eV and 3.5 eV well-depths and are positioned at a fixed distance of 1 nm from the oxide-substrate interface, respectively. The trap-levels in both MOSCAP systems are at 1 eV above the substrate conduction band edge in flat-band voltage condition. The trap in Al/HfO₂/InGaAs exhibits substantial broadening compared to PolySi/SiO₂/Si MOSCAP systems. The quasi-bound to bound transition region is indicated with a black dashed box*

no longer straightforward. The kinks observed in the depletion regime of the 0.25 nm spectral line, are due to the coupling between the quantized-inversion states and the localized trap state in the substrate.

The oxide trap level characteristics also depend on the oxide material parameters such as the effective mass. One such example is illustrated in Fig. (2.12). With the 1 nm position away from the oxide-semiconductor interface of a 0.5 nm wide SW trap, the trap well-depths for both MOSCAP systems in Fig. (2.12) are varied such that the energetic distance of the trap levels from the corresponding substrate conduction band edge are identical at zero-field condition. Note that the trap level in the Al/HfO₂/InGaAs MOSCAP is arbitrarily fixed at 1 eV (above the substrate conduction band edge under flat-band voltage conditions) to allow for a relative comparison of the field-

induced effects with respect to a similar trap level (related to neutral-type three fold-Silicon defect pair [74]) in the PolySi/SiO₂/Si MOSCAP system.

The substantial level shifts and broadenings of the Al/HfO₂/In_{0.53}Ga_{0.47}As MOSCAP system can be attributed to the lighter electron effective mass ($m_{\text{HfO}_2}^* \approx 0.11m_0$, $m_{\text{InGaAs}}^* \approx 0.043m_0$) than in PolySi/SiO₂/p⁻Si ($m_{\text{SiO}_2}^* \approx 0.5m_0$, $m_{\text{Si}}^* \approx 0.09m_0$) [73]. The gradual disappearance of the level broadening in the depletion region and the corresponding inter-state coupling (between the bound trap and the bound substrate states) for Al/HfO₂/In_{0.53}Ga_{0.47}As MOSCAP system can also be noticed in Fig. (2.12). Moreover, Figs. (2.11) and (2.12) suggest that the field-induced trap level broadening (a few 100meV) is smaller in magnitude than the defect-bands (a few 1000meV) formed due to the intrinsic variability in the local environment of a defect site in amorphous (or poly-crystalline) oxides [74].

However, it is also evident from the Figs. (2.8), (2.11) and (2.12) that the system of substrate and oxide trap potentials are strongly coupled, which is visible in the form of an increased tunneling of the trap wavefunction into the substrate and the substantial broadening of the trap states. This implies that the Bardeen approximation [60, 61] of decoupling the total system into independent partial systems of substrate and trap potentials is no longer accurate not only for the trap potentials near the oxide interface in PolySi/SiO₂/p⁻Si MOSCAP system but also for the Al/HfO₂/In_{0.53}Ga_{0.47}As MOSCAP system. As a result, the tunneling transition rate in the eNMP recombination model, which is based on the Bardeen approximation (see eNMP based TAT model in Section 1.5.2), requires an appropriate correction.

2.4 Impact of field-effects on the emission rate

The field-induced level shift and broadening could impact the prediction of trap related capture and emission rates in semiconductors and oxides, as well as the calculation of TAT currents. One example is shown in Fig. (2.13)(a)-(d) for both thermally-assisted (Fig. (2.13)(c)) and structural-relaxation (Fig. (2.13)(d)) based TAT models, which also graphically depicts the TFET electrostatics during the off-state. For the indicated mid-gap trap, the high doping concentration in the TFET source exhibits a strong band bending and thereby can result in an apparent trap spectrum (Fig. (2.13)(a)-(b)) as opposed to the presumed sharp trap level in literature [75, 46] (E_t). As illustrated in Fig. (2.13)(c), the lower bound of the spectrum can be responsible for larger TAT currents in TFETs, this is, as the carriers can advance through this spectrum with higher probability since capture requires less thermal energy for a TAT event ($n < m$),

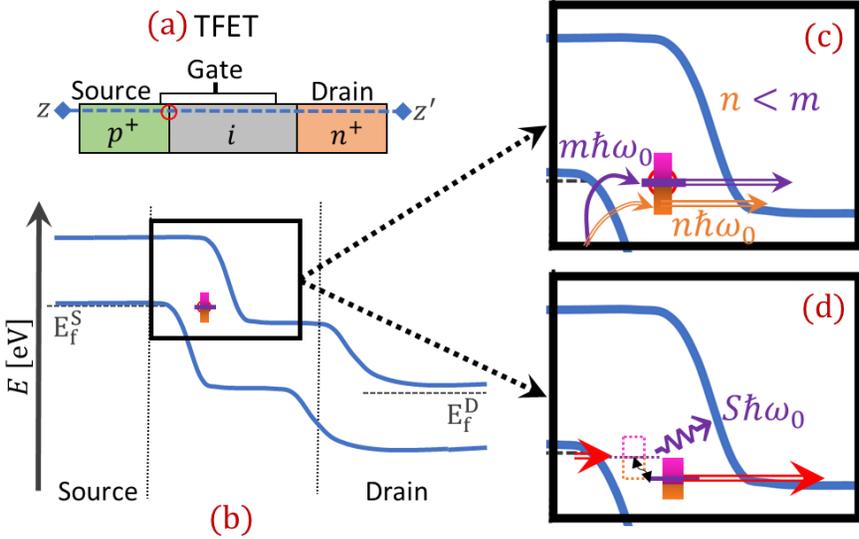


Figure 2.13: Graphical illustrations of (a) *n*-TFET, in which orange, grey and green colors show *n*-type, intrinsic and *p*-type semiconductors, respectively. The trap is presented as a circle at the source junction. (b) Band diagram with a trap state at the indicated cross-section $z - z'$ of Fig. 2.13(a), with the focused regions representing the impact of electric field on the TAT in the off-state of a simple *p-i-n* TFET for (c) the thermally-assisted TAT models and (d) the structural relaxation based TAT models, respectively.

where (m, n) are the number of phonon required for a TAT process in the thermally-assisted TAT models (described in Section 1.5.1).

For the structural-relaxation based models, the field-induced broadening will result in a spectrum of the configuration coordinate (CC) representation of a trap (illustrated in Fig. (A.1) and in Fig. (2.13)(d) with level broadening) rather than a discrete trap CC representation (see Fig. (1.5)(g)), whereas the trap level shift will result in a correction of the energetic distance of the trap's non-equilibrium CC relative to its equilibrium CC representation. The combination of these two field-effects can result in the correction of capture and emission barrier heights in relaxation based TAT models and eventually can lead to increases in capture and emission rates. However, these assumptions would require an efficient tunneling process of the TAT event. We expect that these assumptions, which are based on the field-induced effects of traps, could probably result in higher TAT currents than those calculated with the existing

semi-classical TAT models.

In this section, we present initial steps towards a TAT calculation by determining the semi-classical average emission rate from a trap in the presence of an electric field. In TFETs, the TAT current associated with a specific trap state is proportional to its emission rate. We follow the work in literature [47] describing the field-dependent emission rates, as it is used in determining TAT currents in the field enhanced thermal ionization model (the Hurkx's TAT model described in Section 1.5.1). In this Section, the phonon-assisted emission rate formula for the SW and the CW are modified to account for the field-induced quantum effects on the trap level. In particular, the broadening of the trap level is implemented as a summation of emission rates weighted with probability density and normalized with the total sum of the probability densities of all sub-levels in the spectrum, whereby the spectrum is linearly discretized.

For the SW type semiconductor trap configuration, this results in the following average emission rate $\langle e_{nF} \rangle$ in the presence of field-induced quantum effects (see also Fig. (2.14)(a) for symbols use)

$$\frac{\langle e_{nF} \rangle}{e_{n0}} = \frac{\sum_{\mu=L}^U \chi_{\mu} \exp \left\{ \frac{\delta E_{\mu}}{k_B T} \right\}}{\sum_{\mu} \chi_{\mu}} \times \left[1 + \int_0^{E_{\mu}/k_B T} d\xi \exp \left\{ \xi - \xi^{3/2} \left(\frac{4\sqrt{2m_e^*} (k_B T)^3}{3e_q \hbar F} \right) \right\} \right] \quad (2.47)$$

Where,

$$e_{n0} = e_n^{\infty} \exp \left\{ \frac{-E_i}{k_B T} \right\}; \chi_{\mu} = |\psi(x_p, E_{\mu})|^2; \xi = \frac{E_{\mu} - E_{th}}{k_B T}$$

and where e_n^{∞} ($= \sigma_{n0} \langle v_{th} \rangle N_{DOS}^C$) is the zero-field emission rate coefficient, σ_{n0} is the zero-field capture cross section, $\langle v_{th} \rangle$ is the average thermal velocity of the carrier, N_{DOS}^C is the effective density of states in the conduction band, χ_{μ} is the probability density peak value for the trap sub-level E_{μ} (x_p being the position of the peak), E_{μ} is the trap sub-level at the trap center measured from E_C , δE_{μ} ($= E_i - E_{\mu}$) is the trap sub-level shift compared to the zero-field value, E_{th} is the net thermal energy which the carrier uses to emit from the trap, k_B is the Boltzmann constant and T is the temperature. In Eq. (2.47), the exponential pre-factor ($\exp \{ \delta E_{\mu} \}$) reflects the dependence of the emission rates on the level shift and the summation over μ denotes the impact of broadening.

Similarly, the average emission rate for the CW type trap can be extended as

$$\begin{aligned} \frac{\langle e_{nF} \rangle}{e_{n0}} = & \frac{\sum_{\mu=L}^U \chi_{\mu} \exp \left\{ \frac{\delta E_{\mu}}{k_B T} \right\}}{\sum_{\mu} \chi_{\mu}} \left[\exp \left\{ \frac{\delta E_i^{PF}}{k_B T} \right\} \right. \\ & \left. + \int_{\frac{\delta E_i^{PF}}{k_B T}}^{E_{\mu}/k_B T} d\xi \exp \left\{ \xi - \xi^{3/2} \left(\frac{4\sqrt{2m_e^*(k_B T)^3}}{3e_q \hbar F} \right) \left[1 - \left(\frac{\delta E_i^{PF}}{\xi k_B T} \right)^{5/3} \right] \right\} \right] \end{aligned} \quad (2.48)$$

In Eq. (2.48), the barrier lowering $\left(\delta E_i^{PF} = e_q \sqrt{e_q F / \pi \epsilon_r \epsilon_0} \right)$ [47], responsible for the Poole-Frenkel mechanism is unvaried with respect to the described effects. In the expression of δE_i^{PF} , ϵ_r and ϵ_0 represent the relative permittivity of the host material at high frequency and the vacuum permittivity, respectively.

The average emission rates of Eqs. (2.47) and (2.48) are applied to the Cr acceptor impurity [47] with $\sigma_{n0} = 10^{-14} \text{cm}^2$ in GaAs (see Fig. (2.14)(a)). This trap is configured as SW and CW $(-Ze_q^2/4\pi\epsilon_r\epsilon_0|x|)$ with $Z = 1$, $\epsilon_r^{\text{GaAs}} = 12.9\epsilon_0$. The ionization energy of this impurity in GaAs is 0.8 eV and is considered to be the ground state of an isolated trap ($E_i = 0.8 \text{ eV}$) [47]. Therefore, the SW is described with an approximately one lattice constant (0.6 nm) wide and 2.4 eV deep configuration, while the CW is configured as a 0.6 nm wide at the bottom and 2.8 eV deep hydrogenic well. Fig. (2.14) replicates the results of phonon-assisted emission rates [47] at room temperature (300K), superimposed with the average emission rates of Eqs. (2.47) and (2.48). Considering only the impact of the level shift first (replacing the summation of μ with Max in Eq. (2.47)), the average emission rates are marginally impacted compared to the emission rates of the fixed Cr level for the SW (dotted violet and green emission line in Fig. (2.14)). The combined effect of level shift and spectral broadening can further enhance the average emission rates for both trap configurations. This increase in average emission rates is most pronounced in the high-field region and is negligible below 1 MV/cm.

The enhanced average emission rates above 1 MV/cm field strengths in Fig. (2.14) suggest that the approximation of constant field-enhanced factor between the emission rates and the carrier density of states, used in the Hurkx's TAT model discussed in Section 1.5.1, can no longer be accurate in determining TAT currents, because the field enhanced carrier density of state equation has a similar formula as the field dependent emission rate[47], when it does not account

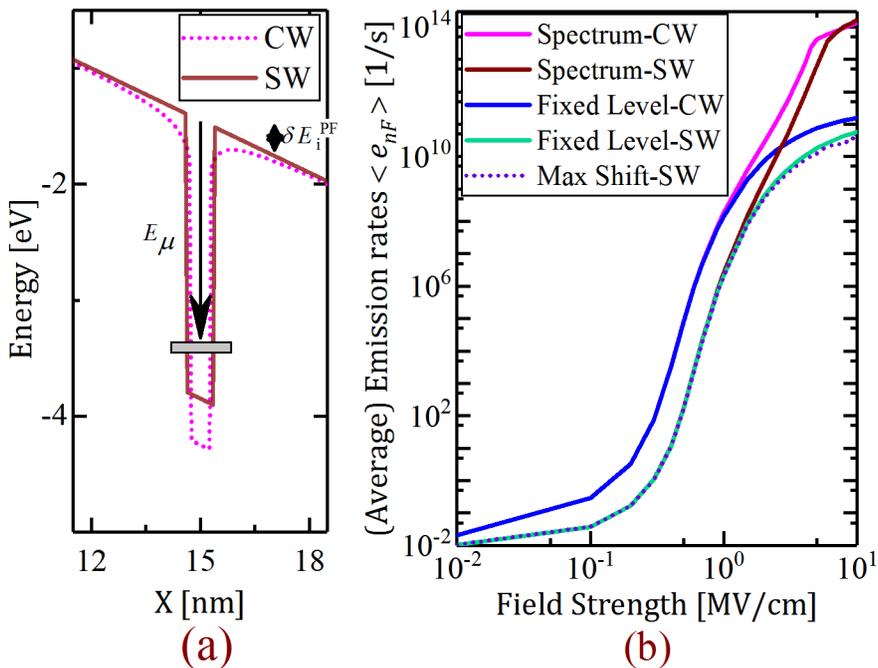


Figure 2.14: (a) The Cr acceptor type trap configured as a SW and CW, (b) Replication of the emission rates outlined in Ref.[47] on a different scale, along with the derived intrinsic impact of the field on the emission rates. The field effects are reflected as the increase in the average emission rates as can be seen from the level shift-only (dot) and the spectrum based solid lines.

for the field-induced broadening and level shift. However, in Schenk's TAT model discussed in Section 1.5.2, there is a provision to include field-induced level shift and broadening by retaining the electric field term in the description of the trap state Eq. (1.10), which could result in a local density of trap states in contrast to the assumed single bound state.

The emission rate equations Eqs. (2.47) and (2.48) are limited to the triangular and hyperbolic barriers, respectively. In addition to this, the effective density of states into which the trap states can leak is assumed to be an abundant continuum of free states. However, the oxide-traps in a MOS capacitor system experience finite-width trapezoidal-like barrier structures. This example is illustrated in Fig. (2.15). For a SW-type oxide trap configuration of Fig. (2.15), the trapezoidal barrier can be treated as the combination of a triangular-type barrier and a fixed-width barrier. This results in the following emission rate

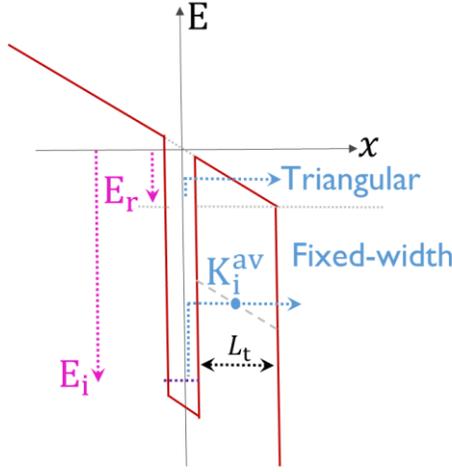


Figure 2.15: Graphical illustration of the field enhanced emission mechanism in a finite-width barrier type structure. The trapezoidal barrier is treated as a combination of a fixed-width type and a triangular-type barrier with a common thermal ionization energy E_r .

expression for a fixed trap level [47]:

$$\begin{aligned} \frac{e_{nF}}{e_n^\infty} = & \exp\left\{\frac{-E_i}{k_B T}\right\} \left[1 + \int_0^{E_r/k_B T} d\xi \exp\left\{\xi - \xi^{3/2} \left(\frac{4\sqrt{2m_e^*}(k_B T)^3}{3e_q \hbar F}\right)\right\} \right] \\ & + \exp\left\{\frac{-E_i}{k_B T}\right\} \left[\int_{E_r/k_B T}^{E_i/k_B T} d\xi \exp\left\{\xi - \left(\xi + \frac{e_q F L_t}{2k_B T}\right)^{1/2} \left(\frac{2L_t \sqrt{2m_e^* k_B T}}{\hbar}\right)\right\} \right] \end{aligned} \quad (2.49)$$

where L_t is the tunneling length for a fixed-width barrier, E_r is the energy at which the triangular barrier changes into a fixed-width barrier (see Fig. (2.15)) measured from the oxide conduction band edge and at the center of the trap. The first integral of Eq. (2.49) corresponds to the emission through a triangular-type barrier. The emission through a fixed-width barrier is given by the second integral term of Eq. (2.49). This fixed-width barrier emission is based on the standard WKB-approximation for the tunneling probability through a finite barrier. In this integral term, the tunneling through a barrier is assumed to occur with an average imaginary wavevector (K_i^{av} in Fig. (2.15)) corresponding

to the value of the imaginary wavevector at the center of the barrier. The field-induced level shift and the broadening is now added into Eq. (2.49) and results in the following average emission rate:

$$\frac{\langle e_{nF} \rangle}{e_n^\infty} = \frac{\sum_{\mu=L}^U \chi_\mu \exp \left\{ \frac{-E_\mu}{k_B T} \right\}}{\sum_{\mu} \chi_\mu} \times \left[\left(1 + \int_0^{E_r/k_B T} d\xi \exp \left\{ \xi - \xi^{3/2} \left(\frac{4\sqrt{2m_e^* (k_B T)^3}}{3e_q \hbar F} \right) \right\} \right) + \left(\int_{E_r/k_B T}^{E_\mu/k_B T} d\xi \exp \left\{ \xi - \left(\xi + \frac{e_q F L_t}{2k_B T} \right)^{1/2} \left(\frac{2L_t \sqrt{2m_e^* k_B T}}{\hbar} \right) \right\} \right) \right] \quad (2.50)$$

Note that Eq. (2.50) is used for finding the oxide trap emission rates which exhibit level broadening at zero-field condition, whereas Eqs. (2.47) and (2.48) assume a sharp trap level for the semiconductor traps at zero-field condition.

The field dependent average emission rates based on Eq. (2.50) for the oxide traps of PolySi/SiO₂/p⁻Si and Al/HfO₂/In_{0.53}Ga_{0.47}As MOSCAP structures with $\sigma_{n0} = 10^{-15} \text{cm}^2$ at 300K are shown in Fig. (2.16). The lower and the upper branch of each of the emission rate lines in Fig. (2.16) correspond to respectively the depletion and the accumulation electrostatic domains of the considered MOSCAP structures and reflects the impact of the effective barrier width towards the substrate. The different cases of the trap configuration are identical to the ones of Figs. (2.10), (2.11) and (2.12). Comparison with the fixed-trap level emission rate (determined with Eq. (2.49)) shows negligible difference (which is not visible on the order of magnitude of Y-axis limits), implying that the level shift and broadening have a negligible impact in the considered parameter space.

From Fig. (2.16), the typical decrease of average emission rate with the increase in trap energetic depth can be noticed. For the case of varying distance of the trap level from the SiO₂/p⁻Si interface, the increase in emission rates with the decrease in barrier width is quite apparent in this figure. The effect of the decrease in tunneling length tends to decrease the split between the accumulation and depletion branches of emission lines. Additionally, it is found that an around 5% increase in average emission rate for the HfO₂ trap compared to that of SiO₂, which can be related to the correspondingly lower effective

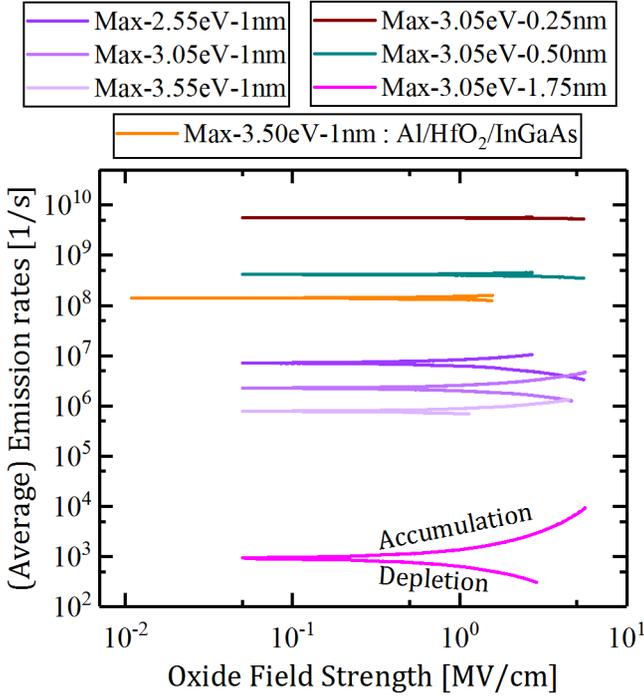


Figure 2.16: Electric field dependent average emission rates at 300K as a function of oxide electric field strengths for different instances of the 0.5 nm-wide SW oxide-trap configuration in the PolySi/SiO₂/Si and the Al/HfO₂/In_{0.53}Ga_{0.47}As MOSCAP systems. The increase in the average emission rates for MOSCAP systems is negligible in comparison with that of semiconductors.

mass in Eq. (2.50). Also for the HfO₂ MOSCAP configuration, the differences in the trap level shift and broadening (see Fig. (2.12)) have a negligible impact on the emission rates.

Even though the emission rate model (Eqs. (2.49) and (2.50)) has the potential to include any trap-type configuration, it lacks the ability to account for the relaxation effects [75, 46, 73] associated with the trapping dynamics. However, this study provides a reference framework for the expected level shifts and broadenings, which could be incorporated into the effective capture and emission barrier heights in models including the relaxation effects [75, 46, 73].

2.5 Conclusions

In this chapter, we explored the implications of high electric fields on the characteristic trap energy level. We defined the normalization method to circumvent the problem of unphysical wavefunction amplitudes at the trap location due to the complex part of the spectral energy states. For the planar semiconductor and oxide traps in a one-dimensional structure, we outlined the numerical procedure to capture the field-induced level broadening, which is typically hundreds of meV for the former and tens of meV for the latter at electric fields of 2 MV/cm. The amount of broadening is sensitive to the chosen trap configuration and may have a quantitative impact on the calculation of the TAT current in the existing NEGF-based TAT models.

The field effects are implemented in the existing semi-classical emission rate formalism. We found that the field-induced quantum effects can increase the emission rates of a semiconductor trap level at high electric field, while the impact for the oxide traps is much smaller due to the higher effective mass and higher barrier heights than those in the former. The field-induced quantum effects for the trap level and the associated emission rates enhancement in a semiconductor device depend on the combined influence of the field strength, trap charge state (neutral, donor or acceptor trap), trap position, tunneling barrier type and the host material parameters. It is expected that the broadening predicted in our chapter will impact the TAT current calculations of the existing semi-classical TAT models and therefore, a more rigorous QM based calculation of TAT is required.

The key findings of this chapter have been published in Journal of Applied Physics [1].

Chapter 3

Phonon-assisted tunneling current formalism

In the previous Chapter 2, we presented a comprehensive study of the impact of electric field-induced quantum effects on the traps and on their emission rates. We determined that the traps in semiconductors and in oxides exhibit a trap level shift, which depends on the barrier type. We further presented that the traps exhibit a trap-level broadening due to the tunneling of the wavefunctions into the continuum states of the lower potential region. Consequently, these effects increase the emissions rate of a semiconductor trap, specifically at operating electric fields of TFETs, and suggest for the modification of existing TAT models to accurately calculate the TAT currents. Owing to such importance of quantum effects on traps, a quantum-mechanical based model of TAT is required. Since phonon-assisted tunneling (PAT) is an essential element in the TAT process in a semiconductor devices, it can be studied separately, while assuming the absence of traps in the devices studied. As an initial step towards an accurate description of TAT in semiconductor devices, we therefore present in this chapter, a formalism to calculate the phonon-assisted tunneling (PAT) current, specifically for bi-dimensional semiconductor devices having two finite dimensions and an infinite third dimension.

In literature, there are different frameworks of determining quantum transport in semiconductor devices, such as the Landauer-Büttiker method, Fermi's Golden rule, the density matrix approach and others. In the Landauer-Büttiker method, the transport is determined with the effective transmission probability of a carrier injected from an initial to a transmitted final state, whereas Fermi's Golden rule determines the transport with the effective

transition rates (from an initial to a final state). Although both of these frameworks are similar, they are incomplete as these frameworks need not necessarily include all the transition (interaction/scattering) processes. However, these frameworks can be made complete by invoking the remaining transition processes. For instance, these frameworks would require the additional terms to account for the opposite PAT currents from final to initial state, and in turn these frameworks would require the prior knowledge of all transition inducing processes. In the density matrix approach, the transport (an observable quantity) is determined from the statistical averages of the corresponding quantum-mechanical operator, whereby the operator intrinsically includes all transition inducing processes and makes the framework complete. Among all the aforementioned frameworks, we prefer the density matrix approach in formulating the PAT current density equation, due to its advantageous complete framework and possible future extensions. In this chapter, we therefore present the formulation of PAT current density based on the density matrix approach.

The structure of this chapter is as follows: we begin with the introductory theory of many-particle physics to determine the statistical average of an observable quantity in Section 3.1, which is essential in the formulation of the PAT current equation in the subsequent sections of this chapter. In Section 3.2, we derive the PAT current in the framework of second quantization, while applying the quantum transmitting boundary method (QTBM) approximation. Section 3.3 details the calculation of the electron-phonon coupling strength for direct-bandgap materials in the framework of the envelope function description of wavefunctions, in which we apply the low phonon wavevector approximation for direct-bandgap materials. We determine the PAT current density equation for bi-dimensional semiconductor structures in Section 3.4. Section 3.5 briefly discusses the numerical implementation of the formalism followed by conclusions in Section 3.6.

3.1 A many-particle description of electrons and phonons

Since we will formulate the PAT current equation while using the QTBM approximation, in particular the injection of electrons will be considered from the semi-infinite contacts into the active region, it is necessary to determine the statistical average of the single-electron or single-phonon state occupation number operator in these contacts. For this purpose, we present the system of non-interacting electrons and free phonons in this section.

Some of the formalism described in this chapter make use of the equilibrium

statistical mechanics. We refer readers who are not familiar with the theoretical concepts of equilibrium statistical mechanics, such as Fermi-Dirac and Bose-Einstein statistics in the framework of grand canonical ensemble, to Appendix B for a brief introduction.

3.1.1 System of non-interacting electrons

To determine the statistical average of a single-electron state occupation number in the contacts, we first characterize these contacts as a system of N electrons. In general, the equation of motion of a N -electron system is described by the following N -electron Schrödinger equation [76],

$$\begin{aligned} \mathcal{H}_{\text{el}}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi_{\vartheta_1, \vartheta_2, \dots, \vartheta_N}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \\ = E \Psi_{\vartheta_1, \vartheta_2, \dots, \vartheta_N}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \end{aligned} \quad (3.1)$$

where \mathcal{H}_{el} denotes the N -electron Hamiltonian operator, Ψ represents the N -electron eigenfunction and $\mathbf{r}_j, \vartheta_j$ are the position and the quantum (energy) state of j^{th} electron, respectively. The probability distribution as a function of the above representation of the N -electron wavefunction describes that an electron, which is in the quantum state ϑ_1 , is probably located at position \mathbf{r}_1 and so on. In Dirac's bra-ket notation, the N -electron wavefunction is defined by [76],

$$\Psi_{\vartheta_1, \vartheta_2, \dots, \vartheta_N}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \langle \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N | \vartheta_1, \vartheta_2, \dots, \vartheta_N \rangle \quad (3.2)$$

Here, the N -electron wavefunction obeys the principle of indistinguishability of identical fermions, in which the wavefunction is antisymmetric under particle exchange (by interchanging either the two indices of position or of quantum state). It also obeys the Pauli exclusion principle, in which no two electrons can occupy the same state. In QTBM, where the contacts are usually described by a constant potential and the electrons in these contacts are characterized by plane wave like wavefunctions, we assume that the electrons under these conditions can be treated as non-interacting particles. For such a non-interacting electron system, the N -electron Hamiltonian splits into single-electron operators as [76]:

$$\mathcal{H}_{\text{el}}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \sum_{j=1}^N H_{\text{el}}(\mathbf{r}_j) \quad (3.3)$$

whereby H_{el} denotes the single-electron Hamiltonian operator. Correspondingly, the N -electron wavefunction is simplified as the following product of single-

electron eigenfunctions of H_{el} .

$$\Psi_{\vartheta_1, \vartheta_2, \dots, \vartheta_N}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \prod_{j=1}^N \psi_{\vartheta_j}(\mathbf{r}_j); E = \sum_{j=1}^N E_{\vartheta_j} \quad (3.4)$$

where $\psi_{\vartheta_j}(\mathbf{r}_j)$ is the single-electron eigenfunction of H_{el} . From Eq.(3.4), the total energy is determined as the sum of the eigen energies of H_{el} . Note that the above representation of the N -electron wavefunction (Eqs. (3.2)-(3.4)), which is known as the state representation or first quantization formalism, is rather cumbersome to work with in practice. An alternative and succinct way of representing the N -electron wavefunction, known as the occupation number representation or second quantization formalism, incorporates the essential information of many-electron states and is practically convenient to solve the N -electron system in question. We make a clear distinction between these two representations by introducing a ‘‘hat’’ over the operators in the occupation number representation. Therefore, the N -electron system can be equivalently represented by the following occupation number representation, in which the occupation number of the single-electron states rather than the quantum states themselves are used to describe the system [76]:

$$|\vartheta_1, \vartheta_2, \dots, \vartheta_N\rangle \equiv |n_1, n_2, \dots\rangle \quad (3.5)$$

Note that the ordering of single-state occupation numbers in Eq. (3.5) is always arranged in accordance with the ascending order of the quantum states. Moreover, to conserve the total number of electrons, Eq. (3.5) must fulfill the following condition,

$$\sum_{\vartheta=1}^{\infty} n_{\vartheta} = N \quad (3.6)$$

where N is the number of electrons and n_{ϑ} is the occupation number of a particular single-electron state ϑ , which can either be 0 or 1 depending on whether the state is unoccupied or occupied (from Pauli’s principle). Note that in contrast to Eqs. (3.3) and (3.4), the index in Eq. (3.6) counts over all single-electron states, and no longer over all electrons. In occupation number representation, the number of electrons N in a given system can be described by [76]

$$\widehat{N}_{\text{el}} |n_1, n_2, \dots\rangle = N |n_1, n_2, \dots\rangle \text{ with } N = \sum_{\vartheta=1}^{\infty} n_{\vartheta} \quad (3.7)$$

Since the occupation number n_{ϑ} of a particular single-electron state ϑ in Eq. (3.7) is also an observable quantity, it can be defined by the occupation number operator by the following equation [76],

$$\widehat{n}_{\vartheta} |n_1, n_2, \dots\rangle = n_{\vartheta} |n_1, n_2, \dots\rangle \text{ with } n_{\vartheta} = 0, 1 \quad (3.8)$$

Subsequently, the equation of motion of the non-interaction system of electrons (Eq. (3.1)) can be described in the following occupation number representation,

$$\widehat{H}_{\text{el}} |n_1, n_2, \dots\rangle = E |n_1, n_2, \dots\rangle \text{ with } E = \sum_{\vartheta=1}^{\infty} E_{\vartheta} n_{\vartheta} \quad (3.9)$$

Note that the occupation number vector $|n_1, n_2, \dots\rangle$ is a simultaneous eigenvector of Eqs. (3.7)-(3.9) and must satisfy the following orthonormality condition, as the two state vectors are identical if and only if all the occupation numbers n_{ϑ} are equal.

$$\langle n'_1, n'_2, \dots | n_1, n_2, \dots \rangle = \delta_{n'_1, n_1} \delta_{n'_2, n_2} \dots \quad (3.10)$$

Until now, we described the single-electron state occupation number operator. To find the statistical average of a single-electron state occupation number, which is presented in Appendix B, we first describe the method of determining the statistical average of any observable quantity in the next subsection.

Statistical average of an observable

In this subsection we first present the approach of finding the statistical average of any observable quantity in state representation followed by the detailed description in the occupation number representation. In quantum statistics and considering a system of non-interacting electrons, the statistical average of any observable quantity, described by the operator O in state representation, is determined from the following average of its quantum mechanical expectation values weighted with the probabilities [76] $\varrho_{\{\vartheta_j\}, \{\vartheta'_j\}}$:

$$\langle O \rangle = \sum_{\vartheta_1, \vartheta_2, \dots, \vartheta_N} \sum_{\vartheta'_1, \vartheta'_2, \dots, \vartheta'_N} \varrho_{\{\vartheta_j\}, \{\vartheta'_j\}} \langle \vartheta'_1, \vartheta'_2, \dots, \vartheta'_N | O | \vartheta_1, \vartheta_2, \dots, \vartheta_N \rangle \quad (3.11)$$

where the set $\{\vartheta_j\}$ accounts for all N -electron states. In Eq. (3.11), the density matrix elements $\varrho_{\{\vartheta_j\}, \{\vartheta'_j\}}$ represent the probabilities with which the quantum mechanical expectation values ($\langle \vartheta'_1, \vartheta'_2, \dots, \vartheta'_N | O | \vartheta_1, \vartheta_2, \dots, \vartheta_N \rangle$) contribute to the statistical average and are determined from “ ϱ ” (also known as statistical operator, ϱ will be determined at the end of this subsection) [76],

$$\varrho_{\{\vartheta_j\}, \{\vartheta'_j\}} = \langle \vartheta_1, \vartheta_2, \dots, \vartheta_N | \varrho | \vartheta'_1, \vartheta'_2, \dots, \vartheta'_N \rangle \quad (3.12)$$

Substituting Eq. (3.12) into Eq. (3.11) and using the following completeness property of eigen functions of Eq. (3.4) (also known as resolution of identity),

$$I = \sum_{\vartheta'_1, \vartheta'_2, \dots, \vartheta'_N} | \vartheta'_1, \vartheta'_2, \dots, \vartheta'_N \rangle \langle \vartheta'_1, \vartheta'_2, \dots, \vartheta'_N | \quad (3.13)$$

the statistical average of an observable is simplified as the diagonal sum of the QM expectation values of the observable quantity correspondingly multiplied with the statistical probabilities, which is as follows:

$$\langle O \rangle = \sum_{\vartheta_1, \vartheta_2, \dots, \vartheta_N} \langle \vartheta_1, \vartheta_2, \dots, \vartheta_N | \rho O | \vartheta_1, \vartheta_2, \dots, \vartheta_N \rangle \equiv \mathbf{Tr}(\rho O) \quad (3.14)$$

In Eq. (3.14), the diagonal sum of expectation values is identified with a Trace operator.

In occupation number representation, the statistical average of an observable can similarly be derived and simplified using Eqs. (3.5), (3.11)-(3.14). This results in:

$$\langle \hat{O} \rangle = \sum_{n_1, n_2, \dots=0}^1 \langle n_1, n_2, \dots | \hat{\rho} \hat{O} | n_1, n_2, \dots \rangle \equiv \mathbf{Tr}(\hat{\rho} \hat{O}) \quad (3.15)$$

where $\left(\sum_{n_1, n_2, \dots=0}^1 \right)$ represents the abstract notation for $\left(\sum_{n_1=0}^1 \sum_{n_2=0}^1 \dots \right)$. In general, the Gibbs entropy of a system is obtained from the density matrix operator $\hat{\rho}$ by,

$$S = -k_B \mathbf{Tr}(\hat{\rho} \log \hat{\rho}) \quad (3.16)$$

where k_B is the Boltzmann constant. In a grand canonical ensemble, the total number of electrons is not fixed, and instead it can vary by the exchange of particles with an external reservoir characterized by a constant chemical potential μ . In this framework, the density matrix (in occupation number representation) is obtained by maximizing the entropy of Eq. (3.16) [76] with the following constraints:

$$\begin{aligned} \langle N_{\text{el}} \rangle &= \mathbf{Tr}(\hat{\rho} \hat{N}_{\text{el}}) \\ \langle E \rangle &= \mathbf{Tr}(\hat{\rho} \hat{H}_{\text{el}}) \end{aligned} \quad (3.17)$$

such that the average energy and the average particle number have a fixed value under constant temperature T and chemical potential μ , respectively. With these constraints, the density matrix operator can be written as:

$$\hat{\rho}_{\text{el}} = \frac{\exp\left\{-\Gamma\left(\hat{H}_{\text{el}} - \mu \hat{N}_{\text{el}}\right)\right\}}{\mathcal{Z}_{\text{el}}} \quad (3.18)$$

whereby $\Gamma = 1/k_B T$, T is the temperature and \mathcal{Z} is the grand canonical partition function which ensures the conservation of the statistical probabilities

($\text{Tr}(\hat{\rho}) = 1$) described by

$$\mathcal{Z}_{\text{el}} = \text{Tr} \left(\exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_{\text{el}} - \mu \hat{N}_{\text{el}} \right) \right\} \right) \quad (3.19)$$

Therefore, the statistical average of an observable, in occupation number representation, can be obtained by using Eqs. (3.15), (3.18) and (3.19), which is fully described in Appendix B.

3.1.2 System of free phonons

In the framework of the grand canonical ensemble, the statistical average of a single-phonon (energy) state occupation number is determined in a similar way as it is for electrons. For the free phonons system, the equation of motion is described in the occupation number representation as [76]

$$\hat{\mathcal{H}}_{\text{ph}} |m_1, m_2, \dots\rangle = E_{\text{ph}} |m_1, m_2, \dots\rangle \quad (3.20)$$

where E_{ph} is the eigen energy of the free phonons system whereas the Hamiltonian (of the free phonons system) splits into the summation over the single-phonon Hamiltonian operators, which is described in the state representation as follows:

$$\mathcal{H}_{\text{ph}}(\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_M) = \sum_{j=1}^M H_{\text{ph}}(\mathbf{Q}_j) \quad (3.21)$$

whereby \mathbf{Q}_j denotes the normal coordinate representing an independent vibrational mode of the lattice with wavenumber j , which is known as a normal mode. The eigen energy (E_{ph}) of the free phonons system is simplified as the summation of the single-phonon eigen energies, which is described in the following equation. The occupation number of a single-phonon state λ counts over all the natural numbers, which is due to the fact that the free phonons wavefunction satisfy the principle of indistinguishability of identical bosons, in which the wavefunction remains symmetric (unchanged) upon the exchange of any two phonon states [76]. In other words, the principle describes that more than one phonon can occupy the same single-phonon state.

$$E_{\text{ph}} = \sum_{\lambda=1}^{\infty} E_{\lambda} m_{\lambda}; M = \sum_{\lambda=1}^{\infty} m_{\lambda}; m_{\lambda} = 0, 1, 2, \dots \quad (3.22)$$

Similar to electrons, we are interested in determining the statistical average of a single-phonon state occupation number. It can be determined with Eq. (3.15), while using Eqs. (3.20) and (3.22), which is presented in Appendix B.

3.2 Derivation of the PAT current equation

In this section, we first present the system of electrons and phonons in the framework of second quantization followed by the application of the QTBM approximation. To allow for simulations of PAT current in III-V devices, we follow the framework of Zener tunneling in indirect-bandgap semiconductors, proposed by Vandenberghe *et al.* [77], to formulate the PAT current equation. However, there are some essential differences between our formalism and the formalism of Vandenberghe [77]. Vandenberghe works with uncoupled valence and conduction bands, this is: the effective mass approximation has been used. Since we work with a coupled-band system, it is not possible to use the same split-up in valence and conduction band electrons, as Vandenberghe did. Instead, we split the system of non-interacting electrons into subsystems of the left and the right contact electrons, while imposing QTBM in the contacts. Moreover, we include the phonon dispersion in formulating the PAT current equation (later in Section 3.3), in contrast to the presumed negligible phonon dispersion in Ref. [77].

For a system of non-interacting and boosted electrons, which is a system of non-interacting electrons under the influence of externally applied voltage, the many-electron Hamiltonian in the framework of second quantization is defined by [78]

$$\widehat{\mathcal{H}}_{\text{el}} = \int d\mathbf{r} \widehat{\psi}^\dagger(\mathbf{r}) H_{\text{el}} \widehat{\psi}(\mathbf{r}) \quad (3.23)$$

whereby $\widehat{\mathcal{H}}_{\text{el}}$ denotes the second quantization description of the many-electron Hamiltonian \mathcal{H}_{el} and H_{el} represents a single electron Hamiltonian, which includes external electrostatic potential profiles. Note that we are not introducing a different description of the many-electron Hamiltonian, as the second quantization description of Eq. (3.23) is equivalent to the many-electron Hamiltonian of Eqs. (3.3) and (3.4). An operator that creates an electron at a particular point in space \mathbf{r}_j for the system of non-interacting electrons (Eq. (3.3)) is defined in second quantization as [78]

$$\widehat{\psi}^\dagger(\mathbf{r}_j) = \sum_{\vartheta_j} \hat{c}_{\vartheta_j}^\dagger \psi_{\vartheta_j}^*(\mathbf{r}_j) \quad (3.24)$$

where the operator $\widehat{\psi}^\dagger(\mathbf{r}_j)$ is also known as electron field operator and represents a many-electron wavefunction. Similarly, an operator that annihilates an electron at a particular point in space \mathbf{r}_j is defined as follows [78]:

$$\widehat{\psi}(\mathbf{r}_j) = \sum_{\vartheta_j} \psi_{\vartheta_j}(\mathbf{r}_j) \hat{c}_{\vartheta_j} \quad (3.25)$$

where $\hat{c}_{\vartheta_j}^\dagger$, \hat{c}_{ϑ_j} in Eqs. (3.24) and (3.25) correspond to the creation and annihilation operators for an electron in the ϑ_j^{th} state, respectively, and are defined by the inverse transformation of Eqs. (3.24) and (3.25) as:

$$\begin{aligned}\hat{c}_{\vartheta_j}^\dagger &= \int d\mathbf{r}_j \psi_{\vartheta_j}(\mathbf{r}_j) \widehat{\psi}^\dagger(\mathbf{r}_j) \\ \hat{c}_{\vartheta_j} &= \int d\mathbf{r}_j \psi_{\vartheta_j}^*(\mathbf{r}_j) \widehat{\psi}(\mathbf{r}_j)\end{aligned}\quad (3.26)$$

where $\psi_{\vartheta_j}(\mathbf{r}_j)$ refers to the ϑ_j^{th} state single-electron wavefunction. Note that the electron creation and annihilation operators must fulfill the following anti-commutation relations, which manifest themselves the antisymmetric nature of the many-electron wavefunction which results from fulfilling Pauli's exclusion condition and the principle of identical fermions indistinguishability [78].

$$\left\{ \hat{c}_{\vartheta_i}, \hat{c}_{\vartheta_j}^\dagger \right\} = \delta_{\vartheta_i \vartheta_j}; \quad \left\{ \hat{c}_{\vartheta_i}, \hat{c}_{\vartheta_j} \right\} = 0; \quad \left\{ \hat{c}_{\vartheta_i}^\dagger, \hat{c}_{\vartheta_j}^\dagger \right\} = 0 \quad (3.27)$$

where the indices ϑ_i, ϑ_j can denote any single-electron state. Using Eqs. (3.7), (3.8) and (3.26), the second quantization description of the electron number operator defined in Eq. (3.7) is given by [78]

$$\widehat{N}_{\text{el}} = \sum_{\vartheta_j} \hat{c}_{\vartheta_j}^\dagger \hat{c}_{\vartheta_j} \equiv \sum_{\vartheta_j} \hat{n}_{\vartheta_j} \quad (3.28)$$

where \hat{n}_{ϑ_j} represents the second quantization definition for the single-electron state occupation number.

The second quantization Hamiltonian for the system of free phonons is described by [78]

$$\widehat{\mathcal{H}}_{\text{ph}} = \sum_{\mathbf{q}\nu} \hbar\omega_{\mathbf{q}\nu} \hat{a}_{\mathbf{q}\nu}^\dagger \hat{a}_{\mathbf{q}\nu} \quad (3.29)$$

where $\mathbf{q}\nu$ and $\omega_{\mathbf{q}\nu}$ are the phonon wavevector and frequency of a particular ν -branch, respectively. The index ν in Eq. (3.29) specifies the polarization vector (longitudinal: $\boldsymbol{\epsilon}_{\mathbf{q}\nu} \parallel \mathbf{q}$) or (transversal: $\boldsymbol{\epsilon}_{\mathbf{q}\nu} \perp \mathbf{q}$) as well as the phonon branch (acoustic or optical). The phonon creation and annihilation operators in Eq. (3.29) satisfy the canonical commutation relations, which are defined as [78]:

$$\left[\hat{a}_{\mathbf{q}_1\nu_1}, \hat{a}_{\mathbf{q}_2\nu_2}^\dagger \right] = \delta_{\mathbf{q}_1\mathbf{q}_2} \delta_{\nu_1\nu_2}; \quad \left[\hat{a}_{\mathbf{q}_1\nu_1}, \hat{a}_{\mathbf{q}_2\nu_2} \right] = 0; \quad \left[\hat{a}_{\mathbf{q}_1\nu_1}^\dagger, \hat{a}_{\mathbf{q}_2\nu_2}^\dagger \right] = 0 \quad (3.30)$$

Similar to Eq. (3.28), the second quantization description of the phonon number operator is [78]

$$\widehat{M} = \sum_{\mathbf{q}\nu} \hat{a}_{\mathbf{q}\nu}^\dagger \hat{a}_{\mathbf{q}\nu} \equiv \sum_{\mathbf{q}\nu} \widehat{m}_{\mathbf{q}\nu} \quad (3.31)$$

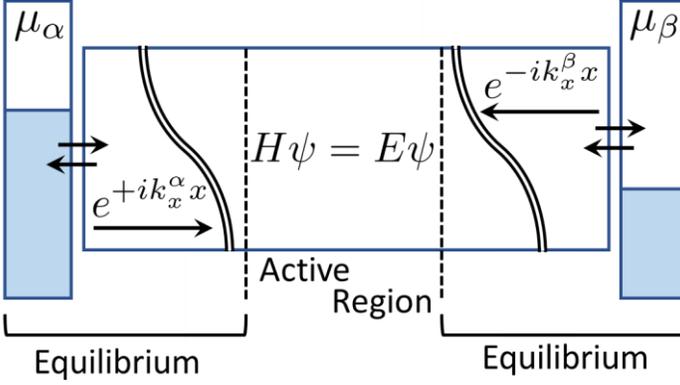


Figure 3.1: A schematic representation of a QTBM approximation, redrawn from Ref. [79].

where $\hat{m}_{q\nu}$ denotes the second quantization definition for the single-phonon state occupation number. We so far presented the description of a system of non-interacting boosted electrons and free phonons in the context of second quantization. In the next subsection, we formulate the PAT current equation by applying the QTBM approximation.

A QTBM approximation

In QTBM [80], an electron (energy) state consists of a mode which is injected from either of the two semi-infinite contacts into the active region (refer to Fig. (3.1)) and which results in the reflection and the transmission of the injected electron. Therefore, the state also comprises all the reflected and the transmitted modes corresponding to the injected mode. However, we identify a single-electron state with the wavevector \mathbf{k} of the injected mode and then split up the complete set of the electron states of the non-interacting electron system in an α -set and a β -set, which correspond to the electron states identified with the eigenstates of the left and right contacts, respectively. These eigenstates therefore correspondingly determine the right-traveling (\mathbf{k}_α) and the left-traveling (\mathbf{k}_β) injected electron plane waves together with all the corresponding reflected and transmitted modes. Substituting Eq. (3.24) in Eq. (3.23) and including the possible combinations of creation and annihilation operators for both contacts, while keeping in mind that all the electron states

identified with eigenstates in the same contact are orthonormal ($\langle \alpha | \alpha' \rangle = \delta_{\alpha\alpha'}$) [81] and in different contacts are orthogonal ($\langle \alpha | \beta \rangle = 0$), results in the following Hamiltonian description of the system of the non-interacting electrons:

$$\widehat{\mathcal{H}}_{\text{el}} = \sum_{\alpha} E_{\alpha} \widehat{c}_{\alpha}^{\dagger} \widehat{c}_{\alpha} + \sum_{\beta} E_{\beta} \widehat{c}_{\beta}^{\dagger} \widehat{c}_{\beta} = \widehat{\mathcal{H}}_{\alpha} + \widehat{\mathcal{H}}_{\beta} \quad (3.32)$$

where E_{α} and E_{β} in Eq. (3.32) correspond to the solutions of a single-electron Schrödinger equation in the left and the right contact, respectively. The electrons in each of the two contacts are exchanged with an external reservoir, which is in equilibrium, thereby fixing the electro-chemical potential of these contacts [80, 79].

The second quantization Hamiltonian for the system of both non-interacting electrons and free phonons is described by a linear combination of the system of non-interacting electrons and free phonons [78],

$$\widehat{\mathcal{H}}_0 = \widehat{\mathcal{H}}_{\text{el}} + \widehat{\mathcal{H}}_{\text{ph}} \quad (3.33)$$

Since we are interested in the statistical average of the PAT currents resulting from a time-dependent perturbation, as will be discussed later, we also describe the density matrix of the system of electrons-phonons in the framework of the grand canonical ensemble. As mentioned in Section 3.1.1 (Eq. (3.14)), the density matrix allows to readily determine the statistical average of an observable. For the combined system of non-interacting electrons and free phonons, while $\mu_{\text{ph}} \widehat{M}_{\text{ph}}$ is approximately zero for the system of free phonons, the density matrix in the framework of the grand canonical ensemble can be determined from the generalization of Eq. (3.18) and results as:

$$\widehat{\varrho} = \frac{\exp \left\{ -\Gamma \left(\widehat{\mathcal{H}}_{\text{el}} - \mu \widehat{N}_{\text{el}} + \widehat{\mathcal{H}}_{\text{ph}} \right) \right\}}{\text{Tr} \left(\exp \left\{ -\Gamma \left(\widehat{\mathcal{H}}_{\text{el}} - \mu \widehat{N}_{\text{el}} + \widehat{\mathcal{H}}_{\text{ph}} \right) \right\} \right)} \quad (3.34)$$

With the QTBM approximation, where we split the system of non-interacting electrons into subsystems of the left and the right contact electrons (Eq. (3.32)), and with the application of the particle number conservation, the total electron number operator (described by \widehat{N}_{el}) can also be separated into the combination of the left ($\widehat{N}_{\text{el},\alpha}$) and the right contact ($\widehat{N}_{\text{el},\beta}$) electron number operators while characterized by the respective density matrices with electro-chemical potentials μ_{α} and μ_{β} . With these approximations, the density matrix of the combined system of non-interacting electrons and free phonons is given by [77],

$$\widehat{\varrho}_0 = \frac{\exp \left\{ -\Gamma \left(\widehat{\mathcal{H}}_{\alpha} - \mu_{\alpha} \widehat{N}_{\text{el},\alpha} + \widehat{\mathcal{H}}_{\beta} - \mu_{\beta} \widehat{N}_{\text{el},\beta} + \widehat{\mathcal{H}}_{\text{ph}} \right) \right\}}{\mathcal{Z}} \quad (3.35)$$

where μ_α and μ_β represent the electro-chemical potential of the left contact and the right contact, respectively, and the electron number operators $\hat{N}_{\text{el},(\alpha,\beta)}$ are defined in accordance with Eq. (3.28) by replacing ($\vartheta_j = \alpha, \beta$). In Eq. (3.35), while applying the assumption of non-interacting electrons and phonons, the partition function is defined by [77]

$$\begin{aligned} \mathcal{Z} &= \mathbf{Tr} \left(\exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_\alpha - \mu_\alpha \hat{N}_{\text{el},\alpha} + \hat{\mathcal{H}}_\beta - \mu_\beta \hat{N}_{\text{el},\beta} + \hat{\mathcal{H}}_{\text{ph}} \right) \right\} \right) \\ &= \mathbf{Tr} \left(\exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_\alpha - \mu_\alpha \hat{N}_{\text{el},\alpha} \right) \right\} \otimes \exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_\beta - \mu_\beta \hat{N}_{\text{el},\beta} \right) \right\} \right. \\ &\quad \left. \otimes \exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_{\text{ph}} \right) \right\} \right) \end{aligned} \quad (3.36)$$

Subsequently, the density matrix of the combined system of non-interacting electrons and phonons reduces to the direct product of the density matrices of electrons and phonons, while using the identity $\mathbf{Tr} (A_1 \otimes A_2 \otimes A_3) = \mathbf{Tr} (A_1) \mathbf{Tr} (A_2) \mathbf{Tr} (A_3)$, as follows:

$$\begin{aligned} \hat{\varrho}_0 &= \left[\frac{\exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_\alpha - \mu_\alpha \hat{N}_{\text{el},\alpha} \right) \right\}}{\mathbf{Tr} \left(\exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_\alpha - \mu_\alpha \hat{N}_{\text{el},\alpha} \right) \right\} \right)} \right] \\ &\otimes \left[\frac{\exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_\beta - \mu_\beta \hat{N}_{\text{el},\beta} \right) \right\}}{\mathbf{Tr} \left(\exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_\beta - \mu_\beta \hat{N}_{\text{el},\beta} \right) \right\} \right)} \right] \otimes \left[\frac{\exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_{\text{ph}} \right) \right\}}{\mathbf{Tr} \left(\exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_{\text{ph}} \right) \right\} \right)} \right] \end{aligned} \quad (3.37)$$

In abstract form, it is written as

$$\hat{\varrho}_0 = \hat{\varrho}_{\text{el}} \otimes \hat{\varrho}_{\text{ph}} \equiv \hat{\varrho}_\alpha \otimes \hat{\varrho}_\beta \otimes \hat{\varrho}_{\text{ph}} \quad (3.38)$$

In the framework of perturbation theory, the electron-phonon interactions are introduced by adding the electron-phonon interaction Hamiltonian to the combined system of Eq. (3.33), which is defined in second quantization by [77, 78]

$$\hat{\mathcal{H}}_{\text{int}} = \sum_{\substack{\alpha\beta \\ \mathbf{q}\nu}} \left[g_{\alpha\beta\mathbf{q}\nu} \hat{c}_\beta^\dagger \hat{c}_\alpha \left(\hat{a}_{\mathbf{q}\nu} + \hat{a}_{-\mathbf{q}\nu}^\dagger \right) \right] + \left[g_{\alpha\beta\mathbf{q}\nu}^* \left(\hat{a}_{\mathbf{q}\nu}^\dagger + \hat{a}_{-\mathbf{q}\nu} \right) \hat{c}_\alpha^\dagger \hat{c}_\beta \right] \quad (3.39)$$

where $g_{\alpha\beta\mathbf{q}\nu}$ denotes the electron-phonon coupling (EPC) strength related to the interband transitions mediated by electron-phonon interactions (see Section 3.3).

The steady-state phonon-assisted current can be calculated by taking the statistical average of the rate of change of the number of electrons in either of the contacts. As mentioned earlier, the number of electrons is given by the number

operator ($\vartheta_j = \alpha, \beta$ in Eq. (3.28)). In the framework of the density matrix approach, the statistical average of the PAT current is therefore determined by [77],

$$I_{\text{ph}} = -e_q \lim_{t \rightarrow \infty} \left[\frac{d}{dt} \left\{ \text{Tr} \left(\tilde{N}_\alpha(t) \tilde{\varrho}(t) \right) \right\} \right] \quad (3.40)$$

with “ $\tilde{\varrho}(t)$ ” representing the time-dependent density matrix and where we use tilde over N and ϱ as the notation to specify the time dependence of the number and density matrix operators.

In Eq. (3.40), the electron-number operator $\tilde{N}_\alpha(t)$ is independent of time as it commutes with the Hamiltonian $\hat{\mathcal{H}}_0$ and hence it reduces to \hat{N}_α ($\vartheta_j = \alpha$ in Eq. (3.28)). The density matrix is constructed by a recursive expansion of the following von-Neumann equation [82], combined with iterative time integrations:

$$\frac{d}{dt} \tilde{\varrho}(t) = -\frac{i}{\hbar} \left[\tilde{\mathcal{H}}_{\text{int}}(t), \tilde{\varrho}(t) \right] \quad (3.41)$$

The perturbative approximation consists of taking the first-order expansion of “ $\tilde{\varrho}(t)$ ” from Eq. (3.41) and is of the following form [83, 77]

$$\tilde{\varrho}^{(1)}(t) \approx \hat{\varrho}_0 - \frac{i}{\hbar} \int_0^t dt_1 \left[\tilde{\mathcal{H}}_{\text{int}}(t_1), \hat{\varrho}_0 \right] \quad (3.42)$$

where $\tilde{\mathcal{H}}_{\text{int}}(t)$ represents the time-dependent electron-phonon Hamiltonian in the interaction picture. Formally, an operator in the interaction picture, which is a formalism often convenient for investigating the response of the system to a perturbation (in our case, the electron-phonon interactions) and in which both states and the operators gain time dependence, is defined by [77, 78]

$$\tilde{O}(t) = e^{\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \hat{O} e^{-\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \quad (3.43)$$

where \mathcal{H}_0 is the time-independent Hamiltonian for the combined system of electrons and phonons of Eq. (3.33). From this general definition, the time-dependent electron-phonon Hamiltonian in the interaction picture is described by [77, 78],

$$\tilde{\mathcal{H}}_{\text{int}}(t) = e^{\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \hat{\mathcal{H}}_{\text{int}} e^{-\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \quad (3.44)$$

To determine the current (Eq. (3.40)), first the density matrix (Eq. (3.42)) has to be known, and for this, the interaction Hamiltonian (Eq. (3.44)) has to be known. We start with the latter. Using Eq. (3.39), Eq. (3.44) is rewritten as

$$\tilde{\mathcal{H}}_{\text{int}}(t) = \sum_{\substack{\alpha\beta \\ \mathbf{q}\nu}} g_{\alpha\beta\mathbf{q}\nu} e^{\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \left[\hat{c}_\beta^\dagger \hat{c}_\alpha \left(\hat{a}_{\mathbf{q}\nu} + \hat{a}_{-\mathbf{q}\nu}^\dagger \right) \right] e^{-\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} + \text{h.c.} \quad (3.45)$$

Making use of the identity property of the exponential function of operators ($\widehat{\mathbb{I}} = \exp(\widehat{O})\exp(-\widehat{O})$), Eq. (3.45) is rewritten in the following form:

$$\begin{aligned} \widetilde{\mathcal{H}}_{\text{int}}(t) = \sum_{\substack{\alpha\beta \\ \mathbf{q}\nu}} g_{\alpha\beta\mathbf{q}\nu} e^{\frac{i\widehat{\mathcal{H}}_0 t}{\hbar}} \left[\begin{array}{ccc} -i\widehat{\mathcal{H}}_0 t & i\widehat{\mathcal{H}}_0 t & -i\widehat{\mathcal{H}}_0 t \\ \widehat{c}_\beta^\dagger e^{\frac{\phantom{-i\widehat{\mathcal{H}}_0 t}}{\hbar}} & e^{\frac{\phantom{i\widehat{\mathcal{H}}_0 t}}{\hbar}} & \widehat{c}_\alpha e^{\frac{\phantom{-i\widehat{\mathcal{H}}_0 t}}{\hbar}} \\ & & \end{array} \right. \\ \left. e^{\frac{i\widehat{\mathcal{H}}_0 t}{\hbar}} \left(\widehat{a}_{\mathbf{q}\nu} + \widehat{a}_{-\mathbf{q}\nu}^\dagger \right) \right] e^{\frac{-i\widehat{\mathcal{H}}_0 t}{\hbar}} + \text{h.c.} \quad (3.46) \end{aligned}$$

From the definition of an operator in the interaction picture (Eq. (3.43)), Eq. (3.46) reduces to the following form, where the creation and annihilation operators obtain time dependence.

$$\widetilde{\mathcal{H}}_{\text{int}}(t) = \sum_{\substack{\alpha\beta \\ \mathbf{q}\nu}} g_{\alpha\beta\mathbf{q}\nu} \left[\widetilde{c}_\beta^\dagger(t) \widetilde{c}_\alpha(t) \left(\widetilde{a}_{\mathbf{q}\nu}(t) + \widetilde{a}_{-\mathbf{q}\nu}^\dagger(t) \right) \right] + \text{h.c.} \quad (3.47)$$

Formally, the equation of motion for the creation operator (in fact, any operator) is described in the interaction picture by the first line of Eq. (3.49). Making use of the interaction description of Eq. (3.43), subsequently by Eqs. (3.33), (3.32), (3.29), the time-dependent creation operator in Eq. (3.47) is rewritten as follows [78]:

$$\begin{aligned} i\hbar \frac{d}{dt} \widetilde{c}_\beta^\dagger(t) &= \left[\widetilde{c}_\beta^\dagger(t), \widetilde{\mathcal{H}}_0(t) \right] \\ &= e^{\frac{i\widehat{\mathcal{H}}_0 t}{\hbar}} \left[\widehat{c}_\beta^\dagger, \widehat{\mathcal{H}}_0 \right] e^{\frac{-i\widehat{\mathcal{H}}_0 t}{\hbar}} \\ &= e^{\frac{i\widehat{\mathcal{H}}_0 t}{\hbar}} \left[\widehat{c}_\beta^\dagger, \sum_{\alpha'} E_{\alpha'} \widehat{c}_\alpha^\dagger \widehat{c}_{\alpha'} + \sum_{\beta'} E_{\beta'} \widehat{c}_{\beta'}^\dagger \widehat{c}_{\beta'} + \sum_{\mathbf{q}'\nu'} \hbar\omega_{\mathbf{q}'\nu'} \widehat{a}_{\mathbf{q}'\nu'}^\dagger \widehat{a}_{\mathbf{q}'\nu'} \right] \\ &\quad \times e^{\frac{-i\widehat{\mathcal{H}}_0 t}{\hbar}} \quad (3.48) \end{aligned}$$

Using the canonical commutator relations of Eq. (3.27) and Eq. (3.30), followed by applying Eq. (3.43), Eq. (3.48) is simplified as follows [78]:

$$\begin{aligned}
i\hbar \frac{d}{dt} \hat{c}_\beta^\dagger(t) &= e^{\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \sum_{\beta'} E_{\beta'} \left[\hat{c}_\beta^\dagger, \hat{c}_{\beta'}^\dagger, \hat{c}_{\beta'} \right] e^{-\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \\
&= e^{\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \sum_{\beta'} -E_{\beta'} \hat{c}_{\beta'}^\dagger \delta_{\beta\beta'} e^{-\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \\
&= -E_\beta e^{\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \hat{c}_\beta^\dagger e^{-\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \\
&= -E_\beta \hat{c}_\beta^\dagger(t)
\end{aligned} \tag{3.49}$$

The solution of the above differential equation is simply given by [78]

$$\frac{d}{dt} \hat{c}_\beta^\dagger(t) = \frac{iE_\beta}{\hbar} \hat{c}_\beta^\dagger(t) \Rightarrow \hat{c}_\beta^\dagger(t) = e^{\frac{iE_\beta t}{\hbar}} \hat{c}_\beta^\dagger(0) \equiv e^{\frac{iE_\beta t}{\hbar}} \hat{c}_\beta^\dagger \tag{3.50}$$

Therefore, the electron creation operator in the interaction picture (Eq. (3.50)) retains its character as a single-electron operator, with the time dependence as only the additional phase factor. Similar to Eqs. (3.49) and (3.50), the rest of the time-dependent single-particle operators of Eq. (3.47) can be derived, and are [78]

$$\tilde{c}_\beta(t) = e^{-\frac{iE_\beta t}{\hbar}} \hat{c}_\beta; \tilde{a}_{-\mathbf{q}\nu}^\dagger(t) = e^{\frac{i\hbar\omega_{-\mathbf{q}\nu} t}{\hbar}} \hat{a}_{-\mathbf{q}\nu}^\dagger; \tilde{a}_{\mathbf{q}\nu}(t) = e^{-\frac{i\hbar\omega_{\mathbf{q}\nu} t}{\hbar}} \hat{a}_{\mathbf{q}\nu} \tag{3.51}$$

Substituting Eqs. (3.50) and (3.51) in Eq. (3.47) for the time-dependent single particle operators results in the following expression for the time-dependent electron-phonon interaction Hamiltonian in the interaction picture [77, 78]:

$$\begin{aligned}
\tilde{\mathcal{H}}_{\text{int}}(t) &= \sum_{\substack{\alpha\beta \\ \mathbf{q}\nu}} \left[g_{\alpha\beta\mathbf{q}\nu} \hat{c}_\beta^\dagger \hat{c}_\alpha \left(\hat{a}_{\mathbf{q}\nu} \exp\left(\frac{i(E_\beta - E_\alpha - \hbar\omega_{\mathbf{q}\nu}) t}{\hbar}\right) \right. \right. \\
&\quad \left. \left. + \hat{a}_{-\mathbf{q}\nu}^\dagger \exp\left(\frac{i(E_\beta - E_\alpha + \hbar\omega_{\mathbf{q}\nu}) t}{\hbar}\right) \right) \right] + \text{h.c.} \tag{3.52}
\end{aligned}$$

After the derivation of the above time-dependent electron-phonon interaction Hamiltonian, we return to the PAT current equation formulation. We first

replace the time derivative of the density matrix operator in the PAT current equation Eq. (3.40) with the von-Neumann Eq. (3.41). We then insert Eq. (3.42) for the time-dependent density matrix operator, followed by applying the linear mapping of the Trace. Subsequently, we remove the first term from the resulting equation as it corresponds to zero-PAT current, and we finally apply cyclic permutations under the Trace operation after rearranging the remaining term. This results in [77]:

$$I_{\text{ph}} = \frac{e_q}{\hbar^2} \lim_{t \rightarrow \infty} \int_0^t dt_1 \text{Tr} \left(\left[\left[\widehat{N}_\alpha, \widetilde{\mathcal{H}}_{\text{int}}(t) \right], \widetilde{\mathcal{H}}_{\text{int}}(t_1) \right] \widehat{\varrho}_0 \right) \quad (3.53)$$

Eq. (3.53) is simplified in a similar way as in ref. [77]. For example, the double commutator in Eq. (3.53) results in numerous terms associated with various interband transitions, which can be brought together efficiently after identifying the non-zero terms with the help of Wick's theorem [84]. The trace of one such combination is $\text{Tr}(\widehat{c}_{\beta_1}^\dagger \widehat{c}_{\alpha_1} \widehat{a}_{\mathbf{q}_1 \nu_1} \widehat{a}_{\mathbf{q}_2 \nu_2}^\dagger \widehat{c}_{\alpha_2}^\dagger \widehat{c}_{\beta_2} \widehat{\varrho}_0)$ and it is evaluated by making use of Eq. (3.38), while keeping in mind that the density matrix $\widehat{\varrho}_0$ describes the system of noninteracting electrons and phonons, and subsequently applying sets of Eqs. [(B.5), (3.28), (B.12)] and [(B.5), (3.31), (B.20)] [77],

$$\text{Tr}(\widehat{c}_{\beta_1}^\dagger \widehat{c}_{\alpha_1} \widehat{a}_{\mathbf{q}_1 \nu_1} \widehat{a}_{\mathbf{q}_2 \nu_2}^\dagger \widehat{c}_{\alpha_2}^\dagger \widehat{c}_{\beta_2} \widehat{\varrho}_0) = \text{Tr}(\widehat{c}_{\beta_1}^\dagger \widehat{c}_{\beta_2} \widehat{\varrho}_\beta) \text{Tr}(\widehat{c}_{\alpha_1} \widehat{c}_{\alpha_2}^\dagger \widehat{\varrho}_\alpha) \text{Tr}(\widehat{a}_{\mathbf{q}_1 \nu_1} \widehat{a}_{\mathbf{q}_2 \nu_2}^\dagger \widehat{\varrho}_{\text{ph}}) \quad (3.54)$$

$$= \delta_{\beta_1 \beta_2} f_{\beta_1}(E_{\beta_1}) \delta_{\alpha_1 \alpha_2} (1 - f_{\alpha_2}(E_{\alpha_2})) \\ \times \delta_{\mathbf{q}_1 \nu_1 \mathbf{q}_2 \nu_2} (v(\hbar\omega_{\mathbf{q}_1}) + 1) \quad (3.55)$$

The individual Trace operations in Eq. (3.54) are the statistical averages of the single particle occupation number described in Sections 3.1.1 and 3.1.2, which are characterized by Fermi-Dirac statistics for electrons and Bose-Einstein statistics for phonons, respectively. Under steady state condition, the time-dependent exponential factors in the PAT current equation are approximated with an energy-conserving delta function, resulting in the following expression [77, 85].

$$I_{\text{ph}} = \frac{2\pi e_q}{\hbar} \sum_{\alpha\beta} \sum_{\mathbf{q}\nu} |g_{\alpha\beta\mathbf{q}\nu}|^2 \\ \times \begin{bmatrix} f_\alpha(E_\alpha) & (1 - f_\beta(E_\beta)) & v(\hbar\omega_{\mathbf{q}\nu}) & \delta(E_\beta - E_\alpha - \hbar\omega_{\mathbf{q}\nu}) \\ - f_\beta(E_\beta) & (1 - f_\alpha(E_\alpha)) & (v(\hbar\omega_{\mathbf{q}\nu}) + 1) & \delta(E_\beta - E_\alpha - \hbar\omega_{\mathbf{q}\nu}) \\ + f_\alpha(E_\alpha) & (1 - f_\beta(E_\beta)) & (v(\hbar\omega_{\mathbf{q}\nu}) + 1) & \delta(E_\beta - E_\alpha + \hbar\omega_{\mathbf{q}\nu}) \\ - f_\beta(E_\beta) & (1 - f_\alpha(E_\alpha)) & v(\hbar\omega_{\mathbf{q}\nu}) & \delta(E_\beta - E_\alpha + \hbar\omega_{\mathbf{q}\nu}) \end{bmatrix} \quad (3.56)$$

The first and last term in Eq. (3.56) correspond to the current contribution due to the excitation of an electron from an α state in the left contact to a β state in the right contact, mediated by phonon absorption. The second and third term refer to the phonon-emission current contribution running in opposite direction between the same left and right state contacts. Therefore, Eq. (3.56) represents the PAT current equation in the framework of the QTBM approximation. An essential parameter of Eq. (3.56) is the electron-phonon coupling (EPC) strength $g_{\alpha\beta\mathbf{q}\nu}$, which will be calculated for the direct-bandgap materials in the next section by using the envelope function theory for an electron wavefunction.

3.3 Electron-phonon coupling strength calculation

Formally, the electron-phonon coupling (EPC) strength $g_{\alpha\beta\mathbf{q}\nu}$ for a given α and β electron state, which are correspondingly identified with the (electron) wavevector \mathbf{k}_α and \mathbf{k}_β of the injected modes, is given in Dirac's bra-ket notation by [86]:

$$g_{\alpha\beta\mathbf{q}\nu} = \langle \mathbf{k}_\beta | \Delta\mathcal{V} | \mathbf{k}_\alpha \rangle \equiv \langle \mathbf{k}_\alpha + \mathbf{q} | \Delta\mathcal{V} | \mathbf{k}_\alpha \rangle \quad (3.57)$$

where the phonon state is identified with its wavevector \mathbf{q} and $\Delta\mathcal{V}$ represents the first-order electron-lattice interaction potential defined by the first-order change in the crystal lattice potential (V_c) due to the displacement ($\Delta\mathbf{R}_{u,i}^{j,\nu} = \mathbf{R}_{u,i}^{j,0} - \mathbf{R}_{u,i}^{j,\nu}$) of a j -th atom in the i -th unit-cell from its equilibrium position $\mathbf{R}_{u,i}^{j,0}$ along the ν -direction, which is denoted as [87, 86]:

$$\Delta\mathcal{V} = \sum_{i,j,\nu} \frac{\partial H_{\text{el}}}{\partial \mathbf{R}_{u,i}^{j,\nu}} \Delta\mathbf{R}_{u,i}^{j,\nu} \approx \sum_{i,j,\nu} \frac{\partial V_c}{\partial \mathbf{R}_{u,i}^{j,\nu}} \Delta\mathbf{R}_{u,i}^{j,\nu} \quad (3.58)$$

In wavefunction approach, the generalized form of the electron-phonon coupling strength equivalent to Eq. (3.57), while including the acoustic and the optical phonon branch, for the homostructure is defined by [86]

$$g_{\alpha\beta\mathbf{q}\nu} = M_{\mathbf{q}\nu} \int d\mathbf{r} \psi_\beta^*(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} \psi_\alpha(\mathbf{r}) \quad (3.59)$$

where $M_{\mathbf{q}\nu}$ can either represent the deformation potentials (corresponding to contributions from the longitudinal ($\boldsymbol{\epsilon}_{\mathbf{q}\nu} \parallel \mathbf{q}$) as well as the transverse ($\boldsymbol{\epsilon}_{\mathbf{q}\nu} \perp \mathbf{q}$) branch of both acoustic and optical phonons) or Fröhlich model of interaction potential (with dominant contributions from the longitudinal ($\boldsymbol{\epsilon}_{\mathbf{q}\nu} \parallel \mathbf{q}$) branch of the optical phonons). The exact calculation of Eq. (3.59) would require the detailed knowledge of the electron wavefunctions, which in turn would substantially increase the computational time. We therefore use the envelope

function theory for the electron wavefunctions and subsequently apply the low-phonon wavevector approximation for the direct-bandgap homostructure device in Section 3.3.1. In Section 3.3.2, we present the derivation to calculate EPC strength for the direct-bandgap heterostructure device, while using the approximations similar to the homostructure calculation.

3.3.1 EPC strength for direct-bandgap homostructures

As mentioned earlier, the EPC strength in Eq. (3.56) for a given α and β electron state is defined by Eq. (3.59), where $M_{\mathbf{q}\nu}$ refers to the bulk electron-phonon coupling strength coefficient. Since the materials under study are III-V semiconductors, the Fröhlich-interaction is the dominant electron-phonon interaction. Therefore, $M_{\mathbf{q}\nu}$ in Eq. (3.59) is described by the following bulk polar coupling strength:

$$M_{\mathbf{q}\nu} = \frac{1}{|\mathbf{q}|} \sqrt{\frac{2\pi e_q^2 \hbar \omega_{\mathbf{q}\nu}}{\Omega} \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0} \right)} \quad (3.60)$$

where \mathbf{q} is the phonon wavevector, e_q is the elementary charge, \hbar is the reduced Planck constant, $\omega_{\mathbf{q}\nu}$ is the polar optical phonon frequency with respect to either the longitudinal or transverse branch, Ω is the semiconductor device volume, ε_∞ and ε_0 are the high frequency and static dielectric constants, respectively [88]. To make the EPC strength calculation computationally feasible, we first apply the envelope function theory of electron wavefunctions followed by the low-phonon wavevector approximation in accordance with the larger EPC strengths at low wavevectors of the polar coupling (Eq. (3.60)).

Envelope function theory of electron wavefunction

In this approach, the α and β state's wavefunctions of Eq. (3.59) are expanded with the following envelope function approximation,

$$\psi_\beta^*(\mathbf{r}) = \sum_m U_m^*(\mathbf{r}) F_{\beta m}^*(\mathbf{r}); \quad \psi_\alpha(\mathbf{r}) = \sum_m F_{\alpha m}(\mathbf{r}) U_m(\mathbf{r}) \quad (3.61)$$

whereby $F_{\alpha m}(\mathbf{r})$ are the slowly varying envelope functions, whose Fourier components lie within the first Brillouin zone (1BZ), and $U_m(\mathbf{r})$ constitute a complete set of orthonormal basis functions [89], where m represents the band index. Substituting Eq. (3.61) in Eq. (3.59) results in the following EPC strength expression.

$$g_{\alpha\beta\mathbf{q}\nu} = M_{\mathbf{q}\nu} \sum_{mn} \int d\mathbf{r} U_n^*(\mathbf{r}) F_{\beta n}^*(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} F_{\alpha m}(\mathbf{r}) U_m(\mathbf{r}) \quad (3.62)$$

Replacing the position vector \mathbf{r} by a linear combination of the unit cell vector \mathbf{R}_u and the lattice vector \mathbf{R} , transforms the volume integral in the EPC strength (Eq. (3.62)) into

$$\mathbf{r} = \mathbf{R}_u + \mathbf{R} \Rightarrow \int d\mathbf{r} = \sum_{\mathbf{R}} \int_{\Omega_c} d\mathbf{R}_u \quad (3.63)$$

whereby Ω_c is the volume of the unit cell. Note that the above transformation is only valid for crystalline materials. The EPC strength of Eq. (3.62) can therefore be written as

$$g_{\alpha\beta\mathbf{q}\nu} = M_{\mathbf{q}\nu} \sum_{mn} \sum_{\mathbf{R}} \int_{\Omega_c} d\mathbf{R}_u U_n^*(\mathbf{R}_u + \mathbf{R}) F_{\beta n}^*(\mathbf{R}_u + \mathbf{R}) e^{i\mathbf{q}\cdot(\mathbf{R}_u + \mathbf{R})} F_{\alpha m}(\mathbf{R}_u + \mathbf{R}) U_m(\mathbf{R}_u + \mathbf{R}) \quad (3.64)$$

By making use of the lattice periodicity of the basis functions and assuming that the envelope functions are constants over the unit cell, the basis and envelope functions in Eq. (3.64) can be simplified in the following way

$$\begin{aligned} U_n^*(\mathbf{R}_u + \mathbf{R}) &= U_n^*(\mathbf{R}_u) \quad ; \quad U_m(\mathbf{R}_u + \mathbf{R}) = U_m(\mathbf{R}_u) \\ F_{\beta n}^*(\mathbf{R}_u + \mathbf{R}) &\approx F_{\beta n}^*(\mathbf{R}) \quad ; \quad F_{\alpha m}(\mathbf{R}_u + \mathbf{R}) \approx F_{\alpha m}(\mathbf{R}) \end{aligned} \quad (3.65)$$

The above simplification leads to the separation of envelope functions from basis functions in Eq. (3.64) and the EPC equation can be rewritten as

$$g_{\alpha\beta\mathbf{q}\nu} = M_{\mathbf{q}\nu} \sum_{mn} \sum_{\mathbf{R}} F_{\beta n}^*(\mathbf{R}) e^{i\mathbf{q}\cdot\mathbf{R}} F_{\alpha m}(\mathbf{R}) \int_{\Omega_c} d\mathbf{R}_u U_n^*(\mathbf{R}_u) e^{i\mathbf{q}\cdot\mathbf{R}_u} U_m(\mathbf{R}_u) \quad (3.66)$$

Note that the phonon wavevector in Eq. (3.66) is not necessarily restricted to the first Brillouin zone.

A low- phonon wavevector approximation

Since the polar EPC strengths of Eq. (3.59) are significantly larger for the smaller phonon wavevectors (Eq. (3.60)), we apply a low-phonon wavevector approximation. This approximation further allows to fully exploit the efficient use of the envelope function approximation of electron wavefunction [89], as the initial and final electron states' wavevectors will be nearly identical and the Fourier components of their envelope functions will always lie within the 1BZ.

Mathematically, this approximation implies $e^{i\mathbf{q}\cdot\mathbf{R}_u} \approx 1$ in Eq. (3.66) which then simplifies the EPC strength to:

$$g_{\alpha\beta\mathbf{q}\nu} = M_{\mathbf{q}\nu} \sum_{mn} \sum_{\mathbf{R}} F_{\beta n}^*(\mathbf{R}) e^{i\mathbf{q}\cdot\mathbf{R}} F_{\alpha m}(\mathbf{R}) \int_{\Omega_c} d\mathbf{R}_u U_n^*(\mathbf{R}_u) U_m(\mathbf{R}_u) \quad (3.67)$$

Making use of the following orthonormality of basis functions,

$$\int_{\Omega_c} d\mathbf{R}_u U_n^*(\mathbf{R}_u) U_m(\mathbf{R}_u) = \Omega_c \delta_{nm} \quad (3.68)$$

the EPC strength is determined by the phonon-mediated overlap of same-band envelope functions

$$g_{\alpha\beta\mathbf{q}\nu} = M_{\mathbf{q}\nu} \Omega_c \sum_n \sum_{\mathbf{R}} F_{\beta n}^*(\mathbf{R}) e^{i\mathbf{q}\cdot\mathbf{R}} F_{\alpha n}(\mathbf{R}) \quad (3.69)$$

By the inverse transformation of the lattice vector $\left(\sum_{\mathbf{R}} \rightarrow \frac{1}{\Omega_c} \int d\mathbf{r}\right)$, we regain the spatial dependence of the slowly varying envelope functions, and the EPC strength becomes:

$$g_{\alpha\beta\mathbf{q}\nu} = M_{\mathbf{q}\nu} \sum_n \int_{\Omega} d\mathbf{r} F_{\beta n}^*(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} F_{\alpha n}(\mathbf{r}) \quad (3.70)$$

Each phonon-wavevector consists of its individual components ($\mathbf{q} \rightarrow q_x, q_y, q_z$). Considering a semiconductor device, with dimensions L_x (along the transport x -direction), L_y (along the translational invariant y -direction), L_z (along the confined z -direction), the envelope functions are of the following form

$$F_{\beta n}^*(\mathbf{r}) = F_{\beta n}^*(x, z) e^{-ik_y^\beta y}; \quad F_{\alpha n}(\mathbf{r}) = e^{ik_y^\alpha y} F_{\alpha n}(x, z) \quad (3.71)$$

Inserting Eq. (3.71) in Eq. (3.70) and solving for the integral along the y -direction, simplifies the EPC strength to

$$g_{\alpha\beta\mathbf{q}\nu} = M_{\mathbf{q}\nu} \frac{\left(e^{i(k_y^\alpha - k_y^\beta + q_y)L_y} - 1\right)}{\left(i(k_y^\alpha - k_y^\beta + q_y)\right)} \sum_n \int_{L_x} dx \int_{L_z} dz F_{\beta n}^*(x, z) e^{i(q_x x + q_z z)} F_{\alpha n}(x, z) \quad (3.72)$$

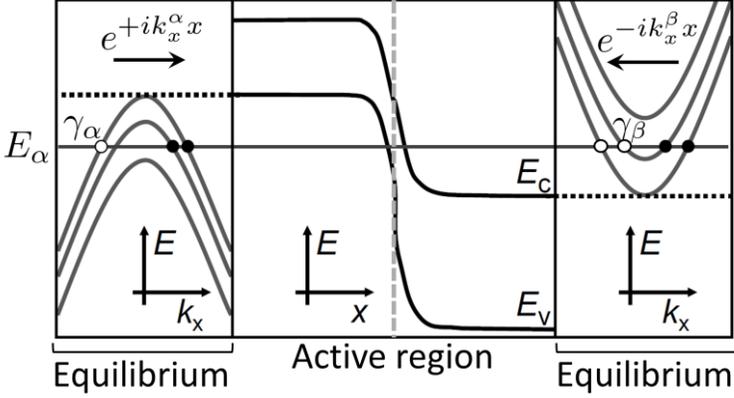


Figure 3.2: Schematic depiction of the injection of a single mode γ_α (open circle) into a diode at an injection energy E_α , whereas the injection of modes γ_β at the same energy are shown for illustrative purpose. Redrawn from Ref. [90].

From Eq. (3.72), the absolute square of the EPC strength is

$$|g_{\alpha\beta\mathbf{q}\nu}|^2 = |M_{\mathbf{q}\nu}|^2 \frac{\sin^2(((k_y^\alpha - k_y^\beta + q_y)/2)L_y)}{((k_y^\alpha - k_y^\beta + q_y)/2)^2} \left[\sum_n \int_{L_x} dx \int_{L_z} dz \phi_{\beta\alpha\mathbf{q}}^n(x, z) \right] \times \left[\sum_m \int_{L_x} dx' \int_{L_z} dz' \phi_{\alpha\beta\mathbf{q}}^n(x', z') \right] \quad (3.73)$$

where $\phi_{\beta\alpha\mathbf{q}}^n(x, z)$ and $\phi_{\alpha\beta\mathbf{q}}^n(x', z')$ denote the overlap functions defined by

$$\phi_{\beta\alpha\mathbf{q}}^n(x, z) = F_{\beta n}^*(x, z) e^{i(q_x x + q_z z)} F_{\alpha n}(x, z) \quad (3.74)$$

$$\phi_{\alpha\beta\mathbf{q}}^n(x', z') = F_{\alpha n}^*(x', z') e^{-i(q_x x' + q_z z')} F_{\beta n}(x', z') \quad (3.75)$$

The PAT current of Eq. (3.56) involves the summation over all possible phonon-wavevectors $\left(\sum_{\mathbf{q}\nu} = \sum_{\mathbf{q}} \sum_{\nu} \right)$. The phonon-wavevector component along the (x, y) -directions can have a continuous range of values within the first Brillouin zone, therefore the summations over (q_x, q_y) are converted into integrals

$\left(\sum_{q_x, q_y} \sum_{\nu} \rightarrow \frac{L_x L_y}{4\pi^2} \sum_{\nu} \int dq_x \int dq_y \right)$. The confinement along the z -direction splits electron sub-bands and retains the discrete nature of q_z , as can be seen from Fig. (3.2). The summation over q_z disappears as the conservation of momentum along the z -direction during the transition from the state α (k_z^α) to the state β (k_z^β) allows for only one value of q_z . With these transformations, Eq. (3.73) can be written as

$$\sum_{\mathbf{q}\nu} |g_{\alpha\beta\mathbf{q}\nu}|^2 = \frac{L_x L_y^2}{4\pi^2} \sum_{\nu} \int dq_x \int dq_y w(k_y^\alpha - k_y^\beta + q_y) G_{\alpha\beta\nu}(q_x, q_y, k_z^\beta - k_z^\alpha) \quad (3.76)$$

where $w(k_y^\alpha - k_y^\beta + q_y)$ is the weight function defined by

$$w(k_y^\alpha - k_y^\beta + q_y) = \left(\frac{\sin(((k_y^\alpha - k_y^\beta + q_y)/2)L_y)}{((k_y^\alpha - k_y^\beta + q_y)/2)} \right) (\text{sinc}(((k_y^\alpha - k_y^\beta + q_y)/2)L_y)) \quad (3.77)$$

whereas, the absolute square of EPC strength is given by

$$G_{\alpha\beta\nu}(q_x, q_y, k_z^\beta - k_z^\alpha) = |M_{\mathbf{q}\nu}|^2 \left[\sum_n \int_{L_x} dx \int_{L_z} dz \phi_{\beta\alpha\mathbf{q}}^n(x, z) \right] \times \left[\sum_m \int_{L_x} dx' \int_{L_z} dz' \phi_{\alpha\beta\mathbf{q}}^m(x', z') \right] \quad (3.78)$$

In the limit $L_y \rightarrow \infty$, the first factor in the weight function “ w ” (Eq. (3.77)) behaves like a delta function $\pi\delta\left(\frac{k_y^\alpha - k_y^\beta + q_y}{2}\right)$. As a consequence, the second cardinal-sine function only matters if $k_y^\alpha - k_y^\beta + q_y = 0$, which results in a value of one. This simplifies the absolute square of the EPC strength in Eq. (3.76) to

$$\sum_{\mathbf{q}\nu} |g_{\alpha\beta\mathbf{q}\nu}|^2 = \frac{L_x L_y^2}{2\pi} \sum_{\nu} \int dq_x \int dq_y \delta(k_y^\alpha - k_y^\beta + q_y) G_{\alpha\beta\nu}(q_x, q_y, k_z^\beta - k_z^\alpha) \quad (3.79)$$

The inner integral over dq_y in Eq. (3.79) is simplified by using the sifting identity of the delta function, thereby resulting in the following summation of absolute square of the EPC strength

$$\sum_{\mathbf{q}\nu} |g_{\alpha\beta\mathbf{q}\nu}|^2 = \frac{L_x L_y^2}{2\pi} \sum_{\nu} \int dq_x G_{\alpha\beta\nu}(q_x, k_y^\beta - k_y^\alpha, k_z^\beta - k_z^\alpha) \quad (3.80)$$

where $G_{\alpha\beta\nu}(q_x, k_y^\beta - k_y^\alpha, k_z^\beta - k_z^\alpha)$ has an absolute square of the following Fröhlich interaction coefficient:

$$M_{\mathbf{q}\nu} = \frac{1}{\sqrt{q_x^2 + (k_y^\beta - k_y^\alpha)^2 + (k_z^\beta - k_z^\alpha)^2}} \sqrt{\frac{2\pi e^2 \hbar \omega_{\mathbf{q}\nu}}{\Omega} \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0} \right)} \quad (3.81)$$

Note that the summation of ν in Eq. (3.80) counts over all the possible longitudinal and transverse branches of the polar optical phonon. Note that the PAT current calculation will be continued in Section 3.4, after the extension made to EPC strengths in the next Section 3.3.2.

3.3.2 EPC strength for direct-bandgap heterostructures

In this section, we present the calculation of an electron-phonon coupling (EPC) strength in heterostructures, because TAT is found to be more pronounced in these structures due to an excess of traps at the hetero-interface. The framework of determining the EPC strength for the direct-bandgap homostructure, presented in the previous Section 3.3.1, can eventually be extended to direct-bandgap heterostructures. In comparison with the direct-bandgap homostructure, the major change in determining the EPC strength for the heterostructure will be due to the material dependent electron basis functions. It is therefore required to derive an expression for the heterostructure EPC strength from the beginning. In this section, we present the detailed derivation of the EPC strength for the lattice-matched heterostructure followed by a brief discussion of extending it to lattice-mismatched heterostructures.

To begin with, the EPC for the lattice-matched heterostructure can be defined by,

$$g_{\alpha\beta\mathbf{q}\nu}^{k \rightarrow l} = M_{\mathbf{q}\nu} \int d\mathbf{r} \psi_\beta^{k,*}(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} \psi_\alpha^l(\mathbf{r}) \quad (3.82)$$

where we arbitrarily assume that the materials in the left (α)- and the right (β)-contacts are made up of l and k direct-bandgap semiconductors, respectively. The α and β state's wavefunctions are expanded, similar to the homostructure derivation, with the following envelope function approximation as

$$\psi_\beta^{k,*}(\mathbf{r}) = \sum_m U_m^{k,*}(\mathbf{r}) F_{\beta m}^{k,*}(\mathbf{r}); \quad \psi_\alpha^l(\mathbf{r}) = \sum_m F_{\alpha m}^l(\mathbf{r}) U_m^l(\mathbf{r}) \quad (3.83)$$

where $F_{\alpha m}^l(\mathbf{r})$ and $F_{\beta m}^k(\mathbf{r})$ are the envelope functions slowly varying over the unit cell of the l and k materials, respectively. $U_m^l(\mathbf{r})$ and $U_m^k(\mathbf{r})$ correspondingly represent the l and k material's electron basis functions. Note that the basis

functions of one material $U_m^l(\mathbf{r})$ are not the eigenfunctions of the other material (k) of the heterostructure. We therefore apply the unitary transformation of $U_m^l(\mathbf{r})$ on to the chosen reference basis set $U_m^k(\mathbf{r})$, developed by Maarten Van de Put *et.al.* [91], so that it allows for employing only one set of basis functions over the whole heterostructure. The unitary transformation is defined as:

$$U_m^l = \sum_j S_{mj}^{k \rightarrow l} U_j^k(\mathbf{r}) \quad (3.84)$$

where $S_{mj}^{k \rightarrow l}$ represent the transformation matrix elements which are defined by [91]

$$S_{mj}^{k \rightarrow l} = \int d\mathbf{r} U_m^{k,*}(\mathbf{r}) U_j^l(\mathbf{r}) \quad (3.85)$$

The transformation of Eq. (3.85) between the basis solutions of bulk Hamiltonians of different materials can be obtained by the common eigenvalue decomposition of inter-band momentum matrices P^l and P^k corresponding to the l and k material, without knowing the exact basis functions [91]. In matrix form, the transformation of Eq. (3.85) is constructed as:

$$S^{k \rightarrow l} = Q^{l\dagger} Q^k \quad (3.86)$$

where Q^l is the matrix containing common eigenvectors of P^l , $Q^{l\dagger}$ its element-wise complex conjugate and Q^k the matrix containing common eigenvectors of P^k . It is apparent from the EPC definition of Eq. (3.82) that it can be convenient to expand the wavefunction for one of the materials in terms of the basis of the other material by using the basis transformation of Eq. (3.84). Therefore, we construct the envelope function expansion of the wavefunction for the left contact in the basis of the right contact's bulk-Hamiltonian solutions, while using Eqs. (3.83) and (3.84), in the following way:

$$\begin{aligned} \psi_\alpha^l(\mathbf{r}) &= \sum_m F_{\alpha m}^l(\mathbf{r}) U_m^l(\mathbf{r}) \\ &= \sum_m F_{\alpha m}^l(\mathbf{r}) \sum_j S_{mj}^{k \rightarrow l} U_j^k(\mathbf{r}) \\ &= \sum_m \sum_j F_{\alpha m}^l(\mathbf{r}) S_{mj}^{k \rightarrow l} U_j^k(\mathbf{r}) \end{aligned} \quad (3.87)$$

Substituting Eqs. (3.87) and (3.83) in Eq. (3.82) and resolving the position vector of the left and the right contact over the same set of lattice and unit-cell vectors assuming the materials in both contacts are lattice-matched, then applying the slowly varying approximation of envelope functions, the periodicity

of basis functions and subsequently the low-phonon wavevector approximation (similar to the homostructure EPC derivation from Eqs. (3.62)-(3.67)), the EPC strength of lattice-matched heterostructures reduces to:

$$g_{\alpha\beta\mathbf{q}\nu}^{k\rightarrow l} = M_{\mathbf{q}\nu} \sum_{nm} \sum_j \sum_{\mathbf{R}} F_{\beta n}^{k,*}(\mathbf{R}) e^{i\mathbf{q}\cdot\mathbf{R}} F_{\alpha m}^l(\mathbf{R}) S_{mj}^{k\rightarrow l} \int_{\Omega_c} d\mathbf{R}_u U_n^{k,*}(\mathbf{R}_u) U_j^k(\mathbf{R}_u) \quad (3.88)$$

Using the following orthonormality condition of basis functions

$$\int_{\Omega_c} d\mathbf{R}_u U_n^{k,*}(\mathbf{R}_u) U_j^k(\mathbf{R}_u) = \Omega_c \delta_{nj} \quad (3.89)$$

and subsequently applying the inverse transformation of lattice-vector into position vector, similar to the simplification from Eq. (3.69) to Eq. (3.70), the EPC strength in lattice-matched heterostructures simplifies to:

$$g_{\alpha\beta\mathbf{q}\nu}^{k\rightarrow l} = M_{\mathbf{q}\nu} \sum_{nm} \int_{\Omega} d\mathbf{r} F_{\beta n}^{k,*}(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} F_{\alpha m}^l(\mathbf{r}) S_{mn}^{k\rightarrow l} \quad (3.90)$$

Note the difference between the EPC strength for homostructures (Eq. (3.70)) and that for heterostructures (Eq. (3.90)) as the latter involves the overlap integrals of the different bands' envelope functions of the left and the right contacts correspondingly multiplied with the transformation matrix elements defined by Eq. (3.86). Similarly, the complex conjugate of the lattice-matched heterostructure EPC strength can be derived as:

$$g_{\alpha\beta\mathbf{q}\nu}^{k\rightarrow l,*} = M_{\mathbf{q}\nu}^* \sum_{nm} \int_{\Omega} d\mathbf{r} [S_{mn}^{k\rightarrow l}]^* F_{\alpha m}^{l,*}(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} F_{\beta n}^k(\mathbf{r}) \quad (3.91)$$

For the bi-dimensional semiconductor system of Fig. (3.2) with the translational invariance along the y -direction, the summation over all possible phonon wavevectors of the absolute square of the EPC strength can also be simplified by resolving the summation over all possible phonon wavevectors into their directional components and using Eqs. (3.71), (3.78) and (3.79), while keeping in mind that the momentum is conserved along the z -direction. It reduces to the following form:

$$\sum_{\mathbf{q}\nu} |g_{\alpha\beta\mathbf{q}\nu}^{k\rightarrow l}|^2 = \frac{L_x L_y^2}{2\pi} \sum_{\nu} \int dq_x G_{\alpha\beta\nu}^{k\rightarrow l}(q_x, k_y^\beta - k_y^\alpha, k_z^\beta - k_z^\alpha) \quad (3.92)$$

where

$$\begin{aligned}
G_{\alpha\beta\nu}^{k\rightarrow l}(q_x, k_y^\beta - k_y^\alpha, k_z^\beta - k_z^\alpha) &= |M_{\mathbf{q}\nu}|^2 \\
&\times \left[\sum_{nm} \int_{L_x} dx \int_{L_z} dz F_{\beta n}^{k,*}(x, z) e^{i(q_x x + q_z z)} F_{\alpha m}^l(x, z) S_{mn}^{k\rightarrow l} \right] \\
&\times \left[\sum_{nm} \int_{L_x} dx' \int_{L_z} dz' [S_{mn}^{k\rightarrow l}]^* F_{\alpha m}^{l,*}(x', z') e^{-i(q_x x' + q_z z')} F_{\beta n}^k(x', z') \right] \quad (3.93)
\end{aligned}$$

where the indices (α, β) of $G_{\alpha\beta\nu}^{k\rightarrow l}(q_x, k_y^\beta - k_y^\alpha, k_z^\beta - k_z^\alpha)$ in Eq. (3.93) refer to the quantum numbers $(k_x^{\alpha,\beta}, k_y^{\alpha,\beta})$. Similar to Eq. (3.80), the summation of ν in Eq. (3.92) counts over the longitudinal and transverse branches of the polar optical phonon. Note that $M_{\mathbf{q}\nu}$ in Eq. (3.93) is assumed to be identical to Eq. (3.81), whereby we assumed the same frequency of the phonons in both materials (through the commutation rules of Eq. (3.30)) as well as the same static and high frequency dielectric constants in both materials. However, a more accurate calculation of EPC strengths must include the mismatch of the aforementioned parameters (phonon frequencies and dielectric constants) between the materials, while adding the interface phonon modes to the system of phonons. In such a calculation, the commutation rules of Eq. (3.30) are no longer valid and the system of phonons (Eq. (3.29)) must be treated in a rigorous way as we did for the system of electrons (Eqs. (3.32) and (3.83)).

For the lattice-mismatched heterostructures, the EPC strength can be determined from Eq. (3.82) by using the basis transformation of Eq. (3.87), then resolving the position vector into two different sets of lattice and unit cell vectors and subsequently applying the coordinate transformation prescribed by Maarten Van de Put *et.al.* [92] for retaining the uniform periodicity in both contacts. The resultant EPC strength and the subsequent PAT current Eq. (3.56) for the lattice-mismatched heterostructure are at least solvable as all required transformations are known. We leave the derivation for the EPC strength in the lattice-mismatched heterostructure for future research. Note that the EPC strength calculation in lattice-mismatched heterostructures would further require higher computational efforts than in the lattice-matched heterostructures, due to an additional coordinate transformation. Moreover, an accurate calculation of EPC strength in lattice-mismatched heterostructures must allow for the inclusion of the additional potential energy term in the system of electrons (Eq. (3.32)), due to the local strain induced by the lattice-mismatched hetero-interface.

3.4 PAT current density for a bi-dimensional-semiconductor structure

Until now, we presented the calculation of EPC strengths for two different cases. Combining the absolute square of EPC strengths of Sections 3.3.1-3.3.2 into the following form, while using Eqs. (3.80) and (3.92),

$$\sum_{\alpha\nu} |g_{\alpha\beta\mathbf{q}\nu}|^2 = \frac{L_x L_y^2}{2\pi} \sum_{\nu} \int dq_x \Lambda_{\alpha\beta\nu}(q_x, k_y^{\beta} - k_y^{\alpha}, k_z^{\beta} - k_z^{\alpha}) \quad (3.94)$$

where

$$\Lambda_{\alpha\beta\nu}(q_x, k_y^{\beta} - k_y^{\alpha}, k_z^{\beta} - k_z^{\alpha}) = \begin{cases} G_{\alpha\beta\nu}(q_x, k_y^{\beta} - k_y^{\alpha}, k_z^{\beta} - k_z^{\alpha}); & \text{Case I} \\ G_{\alpha\beta\nu}^{k \rightarrow l}(q_x, k_y^{\beta} - k_y^{\alpha}, k_z^{\beta} - k_z^{\alpha}); & \text{Case II} \end{cases} \quad (3.95)$$

where Cases I and II refer to the calculation of EPC strengths for the direct-bandgap homostructures and (lattice-matched) heterostructures, respectively. Substituting Eq. (3.94) in Eq. (3.56) and interchanging the order of the summations over $(\alpha\beta)$ and ν , the PAT current equation is of the following form,

$$I_{\text{ph}} = \frac{e_q L_x L_y^2}{\hbar} \sum_{\nu} \sum_{\alpha\beta} \int dq_x \Lambda_{\alpha\beta\nu}(q_x, k_y^{\beta} - k_y^{\alpha}, k_z^{\beta} - k_z^{\alpha}) \times \begin{bmatrix} f_{\alpha}(E_{\alpha}) & (1 - f_{\beta}(E_{\beta})) & v(\hbar\omega_{\mathbf{q}\nu}) & \delta(E_{\beta} - E_{\alpha} - \hbar\omega_{\mathbf{q}\nu}) \\ -f_{\beta}(E_{\beta}) & (1 - f_{\alpha}(E_{\alpha})) & (v(\hbar\omega_{\mathbf{q}\nu}) + 1) & \delta(E_{\beta} - E_{\alpha} - \hbar\omega_{\mathbf{q}\nu}) \\ +f_{\alpha}(E_{\alpha}) & (1 - f_{\beta}(E_{\beta})) & (v(\hbar\omega_{\mathbf{q}\nu}) + 1) & \delta(E_{\beta} - E_{\alpha} + \hbar\omega_{\mathbf{q}\nu}) \\ -f_{\beta}(E_{\beta}) & (1 - f_{\alpha}(E_{\alpha})) & v(\hbar\omega_{\mathbf{q}\nu}) & \delta(E_{\beta} - E_{\alpha} + \hbar\omega_{\mathbf{q}\nu}) \end{bmatrix} \quad (3.96)$$

Due to translational invariance along the y -direction and the plane wave-like envelope functions along the transport x -direction, the summations of the electron-wavevector along these directions are transformed into integrals. The summation of the electron-wavevector along the z -direction is retained due to confinement along this direction (see Fig. (3.2)). These transformations are represented by the followed equation

$$\sum_{\alpha\beta} = \frac{1}{16\pi^4} \int dk_y^{\alpha} \int dk_y^{\beta} \int dk_x^{\alpha} \int dk_x^{\beta} \sum_{k_z^{\alpha} k_z^{\beta}} \quad (3.97)$$

By applying the transformation of Eq. (3.97) and by dividing both sides of Eq. (3.96) by L_y , the PAT current density equation is described as

$$J_{\text{ph}} = \frac{e_q L_x L_y}{16\pi^4 \hbar} \sum_{\nu} \int dk_y^{\alpha} \int dk_y^{\beta} \int dk_x^{\alpha} \left[\int dk_x^{\beta} \delta(E_{\beta} - E_{\alpha} - \hbar\omega_{\mathbf{q}\nu}) \sum_{k_z^{\alpha} k_z^{\beta}} \int dq_x \Lambda'_{\alpha\beta\nu}(q_x, k_y^{\beta} - k_y^{\alpha}, k_z^{\beta} - k_z^{\alpha}) \{f_{\alpha}(E_{\alpha})(1 - f_{\beta}(E_{\beta}))v(\hbar\omega_{\mathbf{q}\nu}) - f_{\beta}(E_{\beta})(1 - f_{\alpha}(E_{\alpha}))v(\hbar\omega_{\mathbf{q}\nu}) + 1\} + \int dk_x^{\beta} \delta(E_{\beta} - E_{\alpha} + \hbar\omega_{\mathbf{q}\nu}) \sum_{k_z^{\alpha} k_z^{\beta}} \int dq_x \Lambda'_{\alpha\beta\nu}(q_x, k_y^{\beta} - k_y^{\alpha}, k_z^{\beta} - k_z^{\alpha}) \{f_{\alpha}(E_{\alpha})(1 - f_{\beta}(E_{\beta}))v(\hbar\omega_{\mathbf{q}\nu}) + 1\} - f_{\beta}(E_{\beta})(1 - f_{\alpha}(E_{\alpha}))v(\hbar\omega_{\mathbf{q}\nu}) \} \right] \quad (3.98)$$

where the indices α and β in $\Lambda'_{\alpha\beta\nu}(q_x, k_y^{\beta} - k_y^{\alpha}, k_z^{\beta} - k_z^{\alpha})$ count over the quantum numbers $(k_x^{\alpha,\beta}, k_y^{\alpha,\beta}, k_z^{\alpha,\beta})$. With the following $k_x^{\alpha,\beta} \rightarrow E_{\alpha,\beta}$ conversion factors,

$$\frac{dk_x^{\alpha}}{dE_{\alpha}} = A_{\alpha}; \quad \frac{dk_x^{\beta}}{dE_{\beta}} = B_{\beta} \quad (3.99)$$

the PAT current density equation is further simplified to:

$$J_{\text{ph}} = \frac{e_q L_x L_y}{16\pi^4 \hbar} \sum_{\nu} \int dk_y^{\alpha} \int dk_y^{\beta} \int dE_{\alpha} A_{\alpha} \left[\int dE_{\beta} \delta(E_{\beta} - E_{\alpha} - \hbar\omega_{\mathbf{q}\nu}) B_{\beta} \sum_{k_z^{\alpha} k_z^{\beta}} \int dq_x \Lambda'_{\alpha\beta\nu}(q_x, k_y^{\beta} - k_y^{\alpha}, k_z^{\beta} - k_z^{\alpha}) \{f_{\alpha}(E_{\alpha})(1 - f_{\beta}(E_{\beta}))v(\hbar\omega_{\mathbf{q}\nu}) - f_{\beta}(E_{\beta})(1 - f_{\alpha}(E_{\alpha}))v(\hbar\omega_{\mathbf{q}\nu}) + 1\} + \int dE_{\beta} \delta(E_{\beta} - E_{\alpha} + \hbar\omega_{\mathbf{q}\nu}) B_{\beta} \sum_{k_z^{\alpha} k_z^{\beta}} \int dq_x \Lambda'_{\alpha\beta\nu}(q_x, k_y^{\beta} - k_y^{\alpha}, k_z^{\beta} - k_z^{\alpha}) \{f_{\alpha}(E_{\alpha})(1 - f_{\beta}(E_{\beta}))v(\hbar\omega_{\mathbf{q}\nu}) + 1\} - f_{\beta}(E_{\beta})(1 - f_{\alpha}(E_{\alpha}))v(\hbar\omega_{\mathbf{q}\nu}) \} \right] \quad (3.100)$$

The inner integrals over E_{β} in Eq. (3.100) are solved by using the following sifting property of the delta function

$$\int dE_{\beta} \delta(E_{\beta} - (E_{\alpha} \pm \hbar\omega_{\mathbf{q}\nu})) \mathcal{D}(E_{\beta}) = \mathcal{D}(E_{\alpha} \pm \hbar\omega_{\mathbf{q}\nu}) \quad (3.101)$$

This reduces the complexity of the PAT current density equation to a single energy integral over E_α . For a given energy E_α and $k_y^{\alpha,\beta}$, the summation over $k_z^{\alpha,\beta}$ counts all sub-band modes $\gamma_{\alpha,\beta}$ corresponding to the confinement along the z -direction. This results in the final form of the PAT current density equation,

$$J_{\text{ph}} = \frac{e_q L_x L_y}{16\pi^4 \hbar} \sum_\nu \int dk_y^\alpha \int dk_y^\beta \left[\begin{aligned} & \int dE_\alpha \sum_{\gamma_\alpha \gamma_\beta} \int dq_x A_{\gamma_\alpha} B_{\gamma_\beta}^+ \Lambda'_{\alpha\beta\nu}(q_x, k_y^\beta - k_y^\alpha, k_z^\beta - k_z^\alpha) \\ & \left\{ f_\alpha(E_\alpha) \left(1 - f_\beta(E_\beta^+) \right) v(\hbar\omega_{\mathbf{q}\nu}) - f_\beta(E_\beta^+) \left(1 - f_\alpha(E_\alpha) \right) (v(\hbar\omega_{\mathbf{q}\nu}) + 1) \right\} \\ & + \int dE_\alpha \sum_{\gamma_\alpha \gamma_\beta} \int dq_x A_{\gamma_\alpha} B_{\gamma_\beta}^- \Lambda'_{\alpha\beta\nu}(q_x, k_y^\beta - k_y^\alpha, k_z^\beta - k_z^\alpha) \\ & \left\{ f_\alpha(E_\alpha) \left(1 - f_\beta(E_\beta^-) \right) (v(\hbar\omega_{\mathbf{q}\nu}) + 1) - f_\beta(E_\beta^-) \left(1 - f_\alpha(E_\alpha) \right) v(\hbar\omega_{\mathbf{q}\nu}) \right\} \end{aligned} \right] \quad (3.102)$$

where

$$E_\beta^+ = E_\alpha + \hbar\omega_{\mathbf{q}\nu}; \quad E_\beta^- = E_\alpha - \hbar\omega_{\mathbf{q}\nu} \quad (3.103)$$

thereby, fixing the conversion factors in the q_x -integral accordingly as,

$$A_{\gamma_\alpha} = \frac{dk_{x,\gamma}^\alpha}{dE_\alpha}; \quad B_{\gamma_\beta}^+ = \frac{dk_{x,\gamma}^\beta}{dE_\beta} \Bigg|_{E_\alpha + \hbar\omega_{\mathbf{q}\nu}}; \quad B_{\gamma_\beta}^- = \frac{dk_{x,\gamma}^\beta}{dE_\beta} \Bigg|_{E_\alpha - \hbar\omega_{\mathbf{q}\nu}} \quad (3.104)$$

Note that the summation over ν in the PAT current density Eq. (3.102) must be carried out in accordance with the selection of $\Lambda'_{\alpha\beta\nu}(q_x, k_y^\beta - k_y^\alpha, k_z^\beta - k_z^\alpha)$ in Eq. (3.95), thereby determining the PAT current densities for the corresponding semiconductor materials. For instance, the summation of ν in Cases I and II counts over the longitudinal and transverse branches of the polar optical phonon.

3.5 Numerical implementation

At the moment of writing, we implement the formalism for the cases of direct-bandgap homostructure and heterostructure (Cases I and II in Eq. (3.95)), as an extension to an existing full-zone quantum-mechanical simulator Pharos [81] which was used earlier for predicting direct-BTBT currents in TFETs. The choice for implementing our formalism in Pharos

is based on its ability to simulate large device structures, while efficiently filtering the spurious solutions through spectral decomposition and therefore reducing the computational demand [81]. The additional feature of Pharos is to calculate direct-BTBT currents in unstrained/strained-heterostructure TFET configurations, offering an eventual extension of our formalism to these structures [90].

The formalism is implemented as a post-processing step in Pharos, where we first calculate the direct-BTBT and save the envelope functions (EFs) for all injected electron energies (E_α, E_β). We retain these EFs to determine the electron-phonon coupling strengths and subsequently calculate the PAT current densities from Eq. (3.102). We prefer the post processing type implementation due to its numerical efficiency. In contrast, the simultaneous calculation of direct-BTBT and PAT would require communication between different CPUs and subsequently increase the computational time.

3.6 Conclusions

In this chapter, as an initial step towards an accurate description of TAT in semiconductor devices, we present a formalism to calculate the phonon-assisted tunneling current, aimed for bi-dimensional semiconductor structures. We formulate the PAT current equation by recycling an existing approach of calculating Zener tunneling in indirect-bandgap devices, while applying a quantum transmitting boundary method approximation.

We calculate the electron-phonon coupling strengths for the direct-bandgap materials in the framework of envelope function theory of electron wavefunctions. We apply the physics-based lattice approach and the low phonon wavevector approximation to eliminate the basis functions in the electron-phonon overlap integrals of both direct-bandgap homostructures and lattice-matched heterostructure devices. We derive the PAT current density equation for the bi-dimensional semiconductor device structures and implement the formalism (for the cases of direct-bandgap homostructures and lattice-matched heterostructures) as a post-processing step in an existing direct-BTBT simulator (Pharos). In the next chapter, we will apply the developed formalism, while using Case I of EPC strength calculation, to different direct-bandgap homostructure device configurations in order to obtain an in-depth understanding of the PAT contributions.

Chapter 4

Application of PAT current formalism

In the previous Chapter 3, we developed a formalism for determining PAT current densities in bi-dimensional semiconductor devices. We determined the electron-phonon coupling (EPC) strength for the direct-bandgap materials, while applying the envelope function theory of electron wavefunctions. In particular, we applied the low-phonon wavevector approximation in determining EPC for the direct-bandgap polar homostructures and heterostructures. For direct-bandgap homostructures, this approximation resulted in the electron-phonon overlap integral of same-band envelope functions. To allow for the use of our formalism, we have implemented the formalism as a post-processing extension in Pharos to determine the PAT current densities in homostructure devices. In this chapter, we will now discuss the results, while using the 2-band and 15-band material description of the formalism for an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode, and examine the nature of the EPC in different diode configurations.

The chapter is organized as follows: We start with the details of the numerical implementation and of the chosen simulation parameters, based on which we further simplify the PAT current density Eq. (3.102), to calculate the PAT current densities in $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diodes, in Section 4.1. Section 4.2 compares the PAT current densities between the 2-band and 15-band material description for an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode. In Section 4.3, we investigate the characteristic EPC across the tunneling junction, specifically for the 2-band model, as there is insignificant difference found between the calculated 2-band and 15-band PAT current densities. In Section 4.4, we discuss the dependence of the PAT current densities on the device's dimensions along the transport

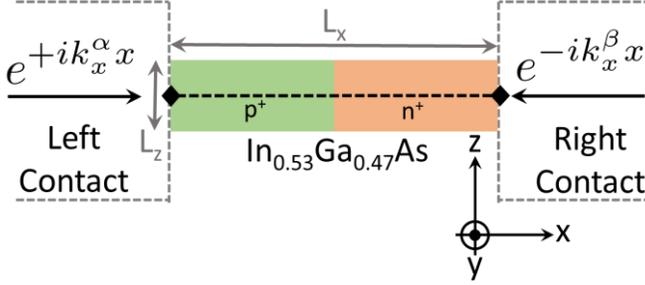


Figure 4.1: Simulated homostructure $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode configuration. Dopant profiles are abrupt and uniform. Translational invariance is assumed in the y -direction. The black-dashed line refers to a cross-section along the center of the diode.

direction, given that we observe a contribution of the PAT process in the near-tunneling region, and on the minimum allowed phonon wavevector. Section 4.5 details the dependence of the PAT current density on the doping concentration in a p-n diode. Finally, we conclude the chapter in Section 4.6.

4.1 Homostructure $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode

In this section, we present the numerical parameters required for an efficient implementation, which is usually determined by a trade-off between the desired accuracy and the computational time. The formalism of the PAT current density for homostructures (using Case I of Eq. (3.95) in Eq. (3.102)) presented in Chapter 3, is applied to an up to 100nm long and 20nm wide $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode as shown in Fig. (4.1). The doping profiles are assumed as abrupt and uniform. The simulated device's spatial mesh sizes are 0.1nm and 0.5nm along the x - and z -directions, respectively. The device is assumed to exhibit translational symmetry along the infinitely large y -direction. The simulation is performed on 40 adaptive energy points, 10 k_y -points to cover all relevant α, β -states and 6 k_z -points depending on the width of the device.

We consider only the dominant Fröhlich interaction, which is based on the longitudinal polar optical branch of the phonons, which physically induces bond stretching of the lattice [93], in a local Fröhlich-based EPC strength with a constant energy ($\hbar\omega_{\mathbf{q}\nu} = \hbar\omega_0$) of 34meV. With this simplification, the

summation of ν in Eq. (3.102) vanishes and the PAT current density of Eq. (3.102) further reduces, while using Case I of Eq. (3.95) in Eq. (3.102), to the following form:

$$J_{\text{ph}} = \frac{e_q L_x L_y}{16\pi^4 \hbar} \int dk_y^\alpha \int dk_y^\beta \left[\begin{aligned} & \int dE_\alpha \sum_{\gamma_\alpha \gamma_\beta} \int dq_x A_{\gamma_\alpha} B_{\gamma_\beta}^+ G'_{\alpha\beta}(q_x, k_y^\beta - k_y^\alpha, k_z^\beta - k_z^\alpha) \\ & \left\{ f_\alpha(E_\alpha) \left(1 - f_\beta(E_\beta^+) \right) v(\hbar\omega_0) - f_\beta(E_\beta^+) \left(1 - f_\alpha(E_\alpha) \right) (v(\hbar\omega_0) + 1) \right\} \\ & + \int dE_\alpha \sum_{\gamma_\alpha \gamma_\beta} \int dq_x A_{\gamma_\alpha} B_{\gamma_\beta}^- G'_{\alpha\beta}(q_x, k_y^\beta - k_y^\alpha, k_z^\beta - k_z^\alpha) \\ & \left\{ f_\alpha(E_\alpha) \left(1 - f_\beta(E_\beta^-) \right) (v(\hbar\omega_0) + 1) - f_\beta(E_\beta^-) \left(1 - f_\alpha(E_\alpha) \right) v(\hbar\omega_0) \right\} \end{aligned} \right] \quad (4.1)$$

where

$$E_\beta^+ = E_\alpha + \hbar\omega_0; \quad E_\beta^- = E_\alpha - \hbar\omega_0 \quad (4.2)$$

and the absolute square of EPC strength is given by,

$$G'_{\alpha\beta}(q_x, k_y^\beta - k_y^\alpha, k_z^\beta - k_z^\alpha) = |M_{\mathbf{q}}|^2 \left[\sum_n \int_{L_x} dx \int_{L_z} dz \phi_{\beta\alpha\mathbf{q}}^n(x, z) \right] \\ \times \left[\sum_m \int_{L_x} dx' \int_{L_z} dz' \phi_{\alpha\beta\mathbf{q}}^m(x', z') \right] \quad (4.3)$$

Additionally, we adaptively determine the q_x -integral with 40 mesh points, starting from π/L_x till the edge of the first Brillouin zone. This defines the criterion to determine the minimum amplitude $|q_{\min}|$ of the phonon wavevector, which is an important parameter since the lower the magnitude of the phonon wavevector, the larger is the Fröhlich-based EPC strength (see Eq. (3.60)). This criterion further ensures that the phonon plane wave is matched to the device length along the transport x -direction. Note that Eqs. (4.1)-(4.3) are inapplicable for determining PAT current in homostructure devices, where the dominant electron-phonon interactions are due to the transverse optical as well as longitudinal and transversal acoustic phonons, as the energy of these phonons is not constant with wavevector \mathbf{q} .

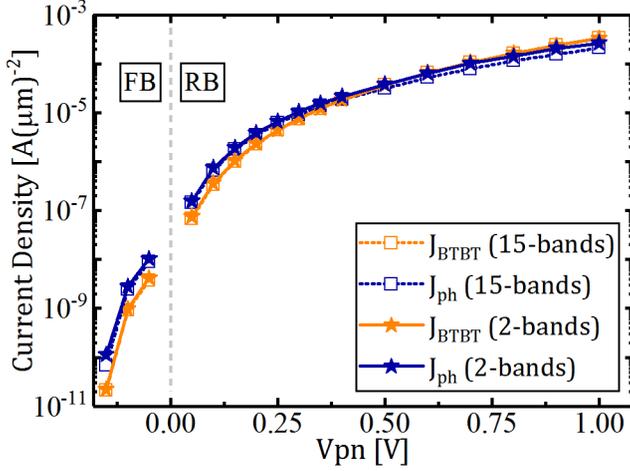


Figure 4.2: The direct-BTBT and PAT current densities of the homostructure $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode of Fig. (4.1) as a function of applied voltage. The diode is 60-nm long and 20-nm wide and is uniformly doped with a concentration of 5×10^{18} [at/cm³]. FB and RB stand for the forward-bias and reverse-bias, respectively.

4.2 Comparison between the 2- and 15-band PAT current density

In this section, the PAT current density calculated from Eq. (4.1) is compared with direct-BTBT [81] for a 20nm wide and 60nm long $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode in Fig. (4.2). The comparison is made for both the 2-band and the 15-band material description of the formalism. As expected, the PAT current density is larger than the direct-BTBT in both models, especially for the forward bias. This is due to the insufficiently available effective energy tunneling window for the direct-BTBT, during which PAT offers an auxiliary transmission path by the gain of phonon energy. Moreover, it is obvious from Fig. (4.2) that the PAT current densities of the 15-band and 2-band model are mostly identical, however with a slight dissimilarity at high reverse bias. Such better agreement between the 15-band and 2-band PAT current densities is obtained using 40 q_x -points in calculating Eq. (4.1). Moreover, the observed PAT and BTBT currents are comparable, which has been reported for several direct-bandgap materials [94]. Due to the faster calculation and a limited difference between the 15-band and the 2-band PAT currents for the given device structure (Fig. (4.1)), specifically at low reverse and in forward bias (the extrapolation of this voltage range to

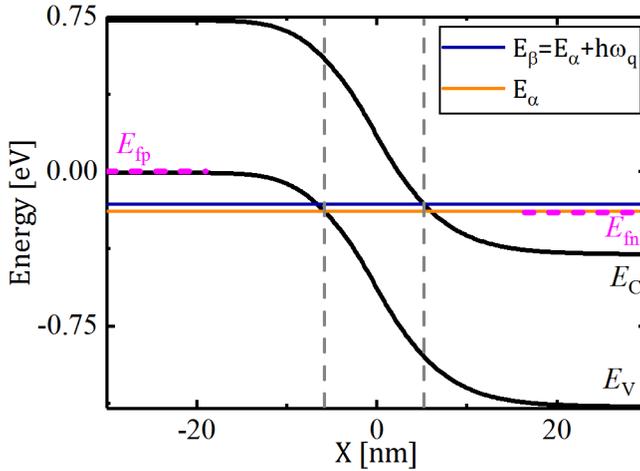


Figure 4.3: Band diagram along the indicated cross-section of the $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode illustrated in Fig. (4.1). The solid-black lines correspond to valence band maximum and conduction band minimum energies. The grey-dashed vertical lines limit the region where tunneling across the junction takes place at the given energies. The pink lines correspond to the quasi-Fermi energy levels E_{fp} and E_{fn} .

TFETs can characterize the impact of PAT on the subthreshold swing), we choose the application of the 2-band model for the remainder of this chapter.

4.3 Inefficient electron-phonon coupling across the tunneling junction

To better understand the PAT current densities of the given device structure, we perform an in-depth analysis of the EPC in this section, particularly by examining the 2-band overlap function density based on Eq. (3.78). For this purpose, the band diagram of a 60nm-long $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode is shown in Fig. (4.3) at the indicated cross-section of Fig. (4.1), where the probability density of the electron wavefunction is maximum. The 2-band envelope function densities ($n=2$ in Eq. (4.3)), averaged over the unit cell, are presented along the transport x -direction in Fig. (4.4). These densities are taken at the indicated energies (E_α, E_β) of Fig. (4.3) and at the center of the z -direction. The overlap function density at the center of the z -direction (see Eq. (3.74)) is shown as

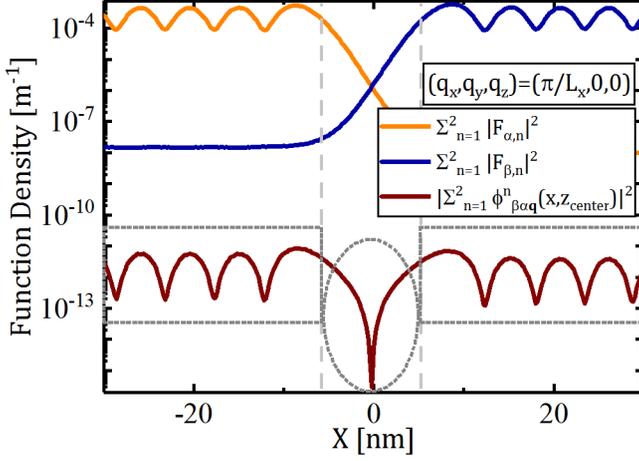


Figure 4.4: The 2-band envelope function densities averaged over the unit cell at specific energies E_α and E_β , indicated in Fig. (4.3). The orange (blue)-solid line corresponds to an electron injected at E_α from the left contact (at E_β from the right contact), respectively. The brown-solid line represents the coupling density, where we chose the phonon plane wave with the directional components $(q_x, q_y, q_z) = (\pi/L_x, 0, 0)$. The limits of the tunneling junction region at (E_α, E_β) are indicated with the grey-dashed vertical lines and the drop of the EPC in this region is highlighted with a grey-dashed circle. The contribution of the PAT process in the near-tunneling regions is indicated with grey-dashed boxes.

a solid-brown line in Fig. (4.4). A drop in overlap function density in the bandgap-region, which is pointed-out by a grey-circle in Fig. (4.4), suggests that the EPC is inefficient across the tunneling junction.

The EPC inefficiency, indicated by a grey-circle in Fig. (4.4), can be explained by examining the envelope functions of both α and β states. The real and the imaginary components of the envelope functions corresponding to each band of a two-band model, at the indicated energies (E_α, E_β) of Fig. (4.3) and at the center of the z -direction, are plotted along the x -direction in Fig. (4.5). It is evident that the conduction band component (band 2) exhibits a 90° -phase shift relative to the valence band component (band 1) of the envelope function, resulting from the different parity of the conduction and valence bands' basis functions ($U_n(\mathbf{r})$). This 90° -phase shift is valid as long as there is limited transmission of electron wavefunction through the forbidden bandgap [95]. This 90° -shift in both α and β states transforms the summation of the individual band's coupling $\phi_{\beta\alpha\mathbf{q}}^n(x, z_{\text{center}})$ into their subtraction and therefore, explains

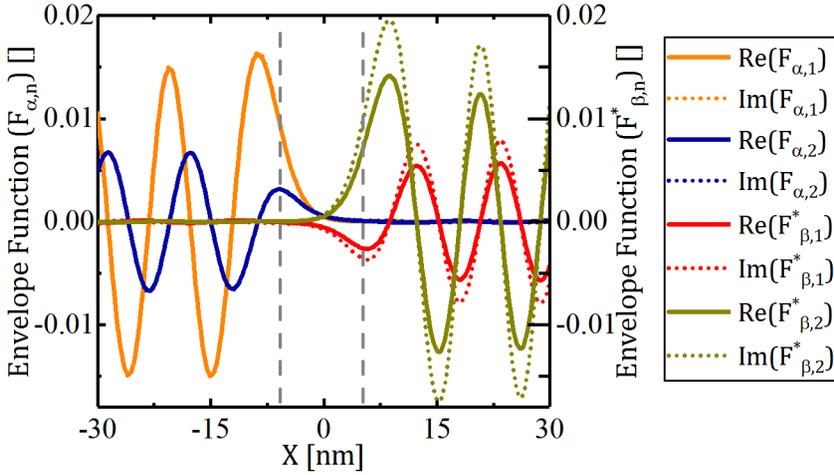


Figure 4.5: The real and imaginary components of each band of the EFs for an electron injected from both contacts. The orange (blue)-solid (dotted) lines correspond to the valence (conduction)-band real (imaginary) contributions of the EF of an electron injected from the left contact. The red (dark yellow)-solid (dot) lines refer to the valence (conduction)-band real (imaginary) contributions of the EF of an electron injected from the right contact.

the drop of overlap function density in Fig. (4.4). Note that the drop in overlap function is also observed in a 15-band model, due to the 90° phase-shift between the conduction (S) and the valence (X) bands.

4.4 PAT dependence on the device length

In this section, we discuss another interesting aspect of Fig. (4.4), which is the finite transmission of an electron's envelope function at the given energies (E_α, E_β) into the region beyond the tunneling junction (indicated by grey rectangles). We describe these regions as “near-tunneling regions”. Consequently, the PAT process in these near-tunneling regions causes a dependence of the PAT current densities on the device's length along the transport direction L_x . Such dependence can be observed from the solid- and dashed-blue lines of Fig. (4.6), which compares the direct-BTBT and PAT current densities for a 100nm and 60nm long $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode, where the 20nm width is constant in both configurations.

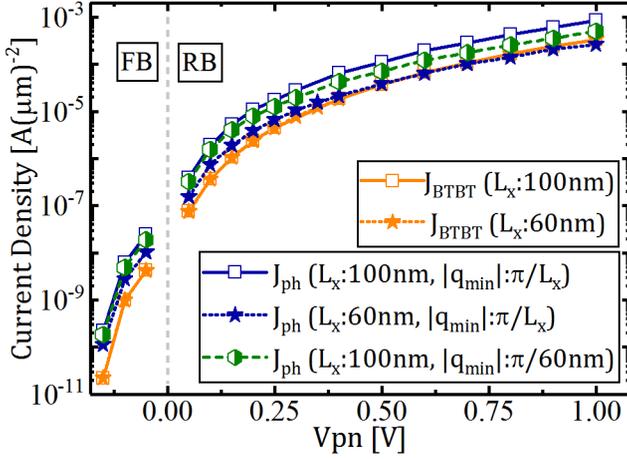


Figure 4.6: Dependence of direct-BTBT and PAT current densities on the device length L_x along the transport direction and on the minimum phonon wavevector. The simulated device is an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p - n diode with varying L_x and a constant 20nm width, uniformly doped with a concentration of $5 \times 10^{18} \text{ [at/cm}^3\text{]}$.

As expected, it is apparent from Fig. (4.6) that the direct-BTBT current densities (orange) do not depend on the device length L_x . However, the PAT current densities exhibit substantial dependence on the device length L_x with higher reverse-bias (see solid- and dashed-blue lines in Fig. (4.6)), which is due to the increased length of the near-tunneling regions. Moreover, the PAT current density also exhibits a dependence on the minimum allowed phonon wavevector, which is evident by comparing the solid-blue and the green-dashed lines in Fig. (4.6). The PAT current density is found to be higher when a larger phonon wavelength and hence smaller phonon wavevector is allowed, which is resulting from the inverse dependence of EPC strengths on the phonon wavevector (see Eqs. (3.59) and (3.60)). A rigorous calculation would accurately limit the region contributing to the PAT current density based on the electron mean-free path. Such restriction is however dependent on the energy of the electron [96] and would increase the mathematical complexity and in-turn the computational efforts. Because of these concerns, we leave it for future research.

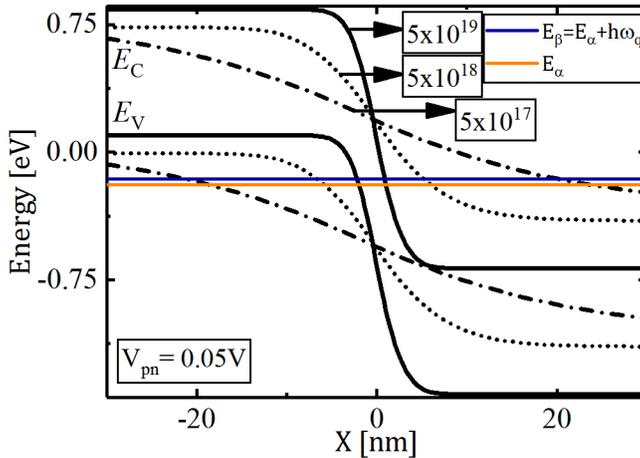


Figure 4.7: Band diagrams for the different doping concentrations along the indicated cross-section of the p-n diode of Fig. (4.1). The solid-(dot-, dashed-) lines correspond to the band-edge energies for 5×10^{19} (5×10^{18} , 5×10^{17}) [at/cm^3] doping concentrations, respectively. The orange (blue)-solid line refers to the energies E_α (E_β).

4.5 Doping dependence of PAT current density

Until now, we have studied the characteristics of the EPC strength, the PAT current densities (both 2-band and 15-band models) and their dependence on the device length for a moderately doped (5×10^{18} [at/cm^3]) p-n diode. The source region in TFETs is typically doped in the range of 1×10^{19} - 5×10^{19} [at/cm^3], which is done to provide a the strong band bending and hence to achieve high on-currents. We therefore examine the dependence of the PAT current densities on the doping concentration in this section, particularly ranging from high to low doping concentrations, with a purpose to determine if PAT can be one of the potential reasons for the SS-degradation in TFETs.

The band diagrams of a 20nm wide and 60nm long $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode, taken at the center along the z -direction, are depicted for different doping concentrations in Fig. (4.7). The increase in tunneling length with decrease in doping concentration is obvious from Fig. (4.7). The electron envelope function densities, taken at the energies (E_α , E_β) of Fig. (4.7), show the decrease in tunneling-transmission in the more lowly doped diodes in Fig. (4.8), which is due to the increase in tunneling length L_t indicated by arrows. Subsequently,

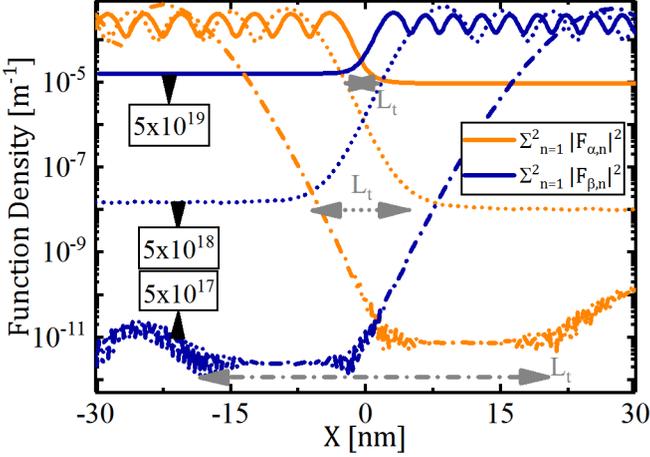


Figure 4.8: The orange-(blue-) solid (dot,dashed) lines indicate the envelope function density averaged over the unit cell in a 2-band model and for doping concentrations of 5×10^{19} (5×10^{18} , 5×10^{17}) [at/cm³] corresponding to the band diagram of Fig. (4.7). The tunneling regions are marked with the horizontal arrows “ L_t ”.

the overlap function densities shown in Fig. (4.9) possess lower amplitudes for the more lowly doped diodes. As a result, the PAT current density dependence on the doping concentrations is illustrated in Fig. (4.10).

It is observed from Fig. (4.10) that the PAT current density partially exceeds the direct-BTBT current density in reverse bias for high doping concentration (5×10^{19} [at/cm³]), while it exceeds for a longer range of reverse bias in moderately doped p-n diodes (5×10^{18} [at/cm³]). For a given voltage (see Fig. (4.7)), the effective direct-BTBT tunneling-energy window, which is determined by doping dependent characteristic Fermi energy levels, is much larger in the 5×10^{19} [at/cm³] diode than in the more lowly doped diodes. This implies that the PAT current in the 5×10^{19} [at/cm³] diode can only be dominant over direct-BTBT at more negative bias where this effective tunneling energy window is sufficiently small compared to the phonon energy. The shift in crossover (grey-circle in Fig. (4.10)) of PAT with respect to direct-BTBT for the 5×10^{18} [at/cm³] diode versus 5×10^{19} [at/cm³] diode can be ascribed to the occurrence of a smaller effective direct-BTBT tunneling window counteracted by the lower contribution of the PAT process in the near-tunneling region. In case of the lowest doped diode (5×10^{17} [at/cm³]), the reduction of the near-tunneling region leads to the smaller PAT relative to direct-BTBT current

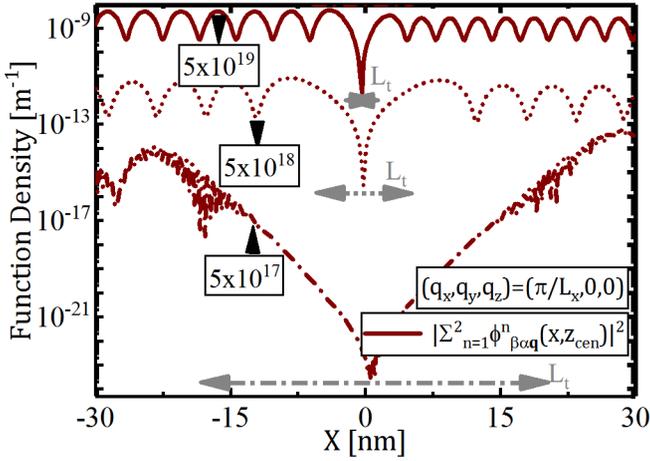


Figure 4.9: The brown- solid (dot, dashed) lines refer to the coupling density of a 2-band model and for 5×10^{19} (5×10^{18} , 5×10^{17}) [at/cm³] doping concentrations corresponding to the band diagram of Fig. (4.7). The coupling density is determined by considering the phonon plane wave with the directional components $(q_x, q_y, q_z) = (\pi/L_x, 0, 0)$. The tunneling regions are marked with the horizontal arrows “ L_t ”.

densities. Hence, the combination of the effective tunneling-energy window and the length of the near-tunneling regions determines the doping dependence of the PAT current density. Due to the large PAT currents, which are comparable to direct-BTBT in Fig. (4.10), we expect that the PAT could be a potential source of SS-degradation in TFETs. However, the quantitative assessment of the impact of PAT on SS degradation in TFETs would require a multi-phonon assisted tunneling current formalism, so that one can gain more insight in the contributions of the underlying PAT processes.

Note that the valence band and the conduction band edges for different doping concentrations in Fig. (4.7) are determined as the band edge energies of the defect free and dopant free diode structures. However, with increasing doping concentrations, the real devices are more prone to random dopant fluctuations and the accurate band edge calculations would require an atomistic based approach to self-consistently determine the device electrostatics [97]. Such a calculation would also require enormous computational resources to determine current densities, even for smaller device dimensions [97]. If realistic electrostatics are known based on such atomistic calculations, our approach could drastically decrease the computational effort in calculating BTBT and

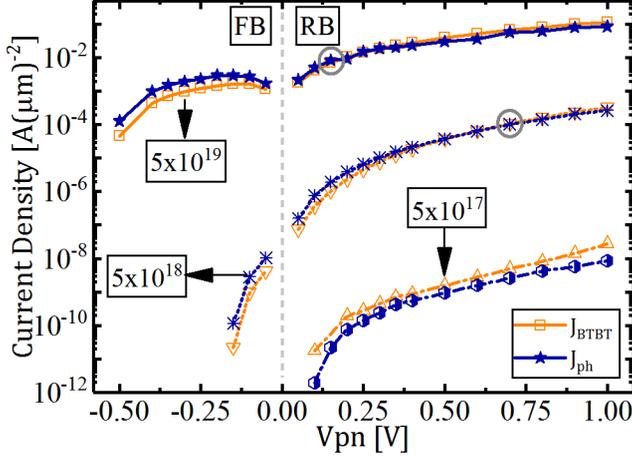


Figure 4.10: Comparison of the direct-BTBT and PAT current densities for different doping concentrations. The orange-(blue-) solid (dot, dashed) lines represent the direct-BTBT (PAT) current densities for 5×10^{19} (5×10^{18} , 5×10^{17}) [at/cm^3] doping concentrations. The cross-over between the direct-BTBT and PAT current density curves is indicated by a grey circle.

PAT current densities if a defect-inclusive electrostatic profile is used.

4.6 Conclusions

In this chapter, we further simplify our PAT current density formalism of the previous Chapter 3, by considering only the longitudinal branch of the polar optical phonon, and apply it to a $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ homostructure p-n diode. We study the PAT current density in up to 100nm long and 20nm wide $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode with a 2 and 15-band implementation of the formalism. We determine that the PAT current density exceeds over direct-BTBT for the forward bias, at which the effective direct-BTBT energy interval is sufficiently small compared to the phonon energy and PAT offers an additional transmission path based on the gain of phonon energy. We observe that the PAT currents are comparable to BTBT. We further find that there is no substantial difference for the calculated PAT current densities in forward (and at low reverse bias) between the 2 and 15-band model of our formalism.

We discover that the electron-phonon coupling strength is inefficient

across the tunneling junction because of the phase-shift between the envelope functions injected from different bands. We find that the PAT current densities depend on the device's length along the transport direction because of the PAT contributions in the near-tunneling regions. We further observe that PAT current densities depend on the minimum allowed phonon wavevector, which is due to the inverse dependence of EPC strengths on the phonon wavevector. We find that the combined effects of effective tunneling-energy interval and near-tunneling region length determine the observed doping dependence of the PAT current density. We expect that our PAT current density formalism can eventually be applied to study the PAT current densities in direct-bandgap heterostructures, as the framework of the implementation (carried out for studying $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diodes) can be reused with a few minor changes. We further expect that our PAT current density formalism can be extended to include multi-phonon assisted tunneling and eventually to assess TAT in devices, which is briefly outlined in the outlook of the next chapter.

The key findings of this chapter have been presented at the APS March meeting 2018 [1] and a detailed report is submitted to the Journal of Applied Physics [2].

Chapter 5

Conclusions and outlook

In this chapter, the conclusions of this thesis are summarized in Section 5.1. Section 5.2 details the extension of the PAT formalism to include multi-phonon assisted tunneling transitions and briefly discusses the prospect of this work, which is eventually the approach to determine TAT in semiconductors.

5.1 Conclusions

The summary and conclusions of this thesis are organized in accordance with the topical goals outlined in Chapter 1. As described in Chapter 1 when discussing the limitations of the existing TAT models, it is clear that there is no unified approach to model TAT in semiconductor devices. We aimed to circumvent the limitations by making use of a quantum mechanical approach to determine TAT.

Therefore, we began with the study of the impact of electric field-induced quantum effects on semiconductor and oxide traps in Chapter 2, with as purpose to determine if quantum effects are crucial in determining TAT. In particular, we explored the implications of high electric fields on the characteristic trap energy level. We defined a normalization method to circumvent the problem of unphysical wavefunctions resulting from the complex part of the spectral energy states. For the planar semiconductor and oxide traps in a one-dimensional structure, we outlined the numerical procedure to capture the field-induced level broadening and found typically hundreds of meV for the former and tens of meV for the latter at TFET-relevant electric fields of 2 MV/cm. The amount

of broadening is sensitive to the chosen trap configuration and may have a quantitative impact on the calculation of the TAT current. In this chapter, the field-induced quantum effects are implemented in the existing semi-classical emission rate formalism. We found that these effects can increase the emission rates of a semiconductor trap level at high electric field, while the impact for the oxide traps is much smaller due to the higher effective mass and barrier heights than those in the former and depend on the combined influence of the field strength, trap charge state (neutral, acceptor or donor trap), trap position, tunneling barrier type and the host material parameters.

In Chapter 3, we presented a formalism to calculate the phonon-assisted tunneling current, aimed at bi-dimensional semiconductor devices. We formulated the PAT current equation by recycling an existing approach of calculating Zener tunneling in indirect-bandgap devices, while applying a quantum transmitting boundary method approximation. We determined the electron-phonon coupling in the framework of envelope function theory of electron wavefunctions. We applied the physics based lattice approach and the low phonon wavevector approximation to eliminate the basis functions in the electron-phonon overlap integrals of both direct-bandgap homostructure and heterostructure devices. We derived the PAT current density equation for bi-dimensional semiconductor structures and implemented the formalism as a post-processing step in an existing direct-BTBT simulator (Pharos).

In Chapter 4, we simplified our PAT current density formalism of the previous chapter and applied it to an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode. We studied the PAT current density in up to 100nm long and 20nm wide $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode with a 2 and 15-band implementation of the formalism. We observed that the PAT currents are comparable to BTBT and found that there is no substantial difference for the calculated PAT current densities between the 2 and 15-band model of our formalism. We discovered that the electron-phonon coupling is inefficient across the tunneling junction because of the phase-shift between the envelope functions injected from the different bands. We found that the PAT current densities depend on the minimum allowed phonon wavevector and on the device's length along the transport direction because of the PAT contributions in the near-tunneling regions. We further determined that the combined effects of the effective tunneling-energy interval and the near-tunneling region length result in the observed doping dependence of the PAT current density.

5.2 Outlook

Based on the chapters presented in this thesis, we provide a few suggestions for future research that improve the presented models (Section 5.2.2) and can result in new applications (Section 5.2.1).

5.2.1 Impact on experimental work

The field-induced effects observed in Chapter 2 for a single-oxide defect can provide a better understanding of the microscopic defect properties, when these observations are correlated with the statistical characterizations of Random Telegraph Noise (RTN) signals related to a single-oxide defect in the ultra scaled devices [98]. We suggest that these field effects can be incorporated (by using the average emission/capture rates which were discussed at the end of Section 2.3.2) in the applied physics-based models for the calibration of trap time constants [98]. This could lead to a distribution of time constants and subsequently, result in the field-dependent decrease of the average time constants related to a single-oxide defect. These effects are mainly expected for border trap states, or trap states close to the oxide-semiconductor interface (0.5nm in Fig. (2.11) and 1nm in Fig. (2.12)).

The PAT currents calculated for different diode configurations in Chapter 4, with the application of our formalism, provide more physical insight into parasitic current contributions observed in experimental devices [99]. For instance, the comparison of the oscillatory conductance obtained from the PAT curves of the calculated and the experimental diodes, could validate the hypothesis that the EPC strengths solely determined with longitudinal optical phonons are indeed dominant in III-V direct bandgap materials [100]. Additionally, these oscillatory conductances obtained from the PAT curves of nearly defect-free tunnel diodes, which are determined at low external forward bias voltages, could provide a rough estimate (after calibration and extrapolation) of the impact of the PAT contribution on TFET performance (such as SS degradation). Note that the observable PAT currents, which are intrinsic to the material, contribute to the total currents of TFETs, whereby it can increase the total currents in subthreshold region of the transfer characteristics, and thereby impact the TFET performance. These PAT contributions, which are dependent on the electron mean free path (see Fig. (4.6)), are also expected to moderately increase the I_{ON} currents in TFETs by 20%. To obtain better estimations, we suggest to compare the calculated PAT current with $I - V$ measurements of defect-free tunnel diodes in forward bias for the calibration of TFET-related parameters, for which the diode dimensions are to be scaled with respect to the transport

direction, with lengths varying up to 100nm in agreement with the mean free path of electrons at room temperature in III-V direct bandgap materials [93].

5.2.2 Future research towards enhanced models

In Chapter 2, we used the MTM approach to determine the trap energy level from the 1D- Schrödinger equation using an effective mass approach and adapted the same approach to qualitatively determine the impact of field-induced effects on TAT based emission rates. However, this approach is too simplistic to quantify these effects in real devices with 2D or 3D traps, as we assumed the traps to be one-dimensional with invariant planar structures in the two (y, z) -dimensions. We suggest that a more accurate calculation of trap levels and the related field-induced effects in higher dimensional systems (with 2D or 3D traps) can be attained using approaches such as NEGF formalism [26, 97], Wigner function method [79] and others.

In Chapter 3, we simplified the calculation of EPC strength in semiconductors based on the properties of a crystalline lattice (Eq. (3.63)). The same approximation is strictly speaking not applicable to amorphous materials like oxides. However, this approximation can still be used to qualitatively determine the PAT currents in oxides, by exemplifying oxides with wide bandgap materials. A more accurate prediction of EPC strength in oxides would require DFT-based approach [101].

Two-phonon assisted tunneling current formalism

Based on the successful implementation of PAT in homostructure and heterostructure direct-bandgap semiconductors, we provide suggestions for future research which eventually will lead to the calculation of TAT not only in TFETs but also in general semiconductor devices. We first present the theory of determining multi-PAT in semiconductors, and conclude the thesis with a brief discussion of the approach towards TAT calculations.

In this thesis, we developed the formalism for single-phonon assisted tunneling current in Chapter 3 and applied it to a homostructure $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-n diode in Chapter 4. However, the trap-assisted tunneling process would typically require a one-step multi-phonon assisted tunneling transition of either a conduction or valence band electron into the trap site as illustrated in Fig. (1.5). For developing such a multi-phonon assisted tunneling model, we found it to be inappropriate to use higher order time dependent perturbation theory, whereby the higher order perturbative approximation is obtained by taking the high-

order expansion of the density matrix “ $\tilde{\varrho}(t)$ ” from Eq. (3.41). Substitution of such a density matrix expansion in the PAT current Eq. (3.40) and solving the resultant commutators, results in the time evolution of a single electron-phonon interaction and eventually leads to a time-dependent multi-step single-PAT current. For steady-state TAT currents, we are not interested in such current calculation as we are looking for a time-independent one-step multi-PAT current.

In this section, we therefore briefly present the procedure to determine the two-phonon assisted tunneling currents, which eventually can be applied to determine steady-state TAT currents. We suggest that the calculation of both single-phonon and two-phonon assisted tunneling currents can be obtained by substituting the time-dependent version, in the interaction picture, of the following time-independent electron-phonon interaction Hamiltonian in the PAT current Eq. (3.53),

$$\hat{\mathcal{H}}_{\text{int}} = \hat{\mathcal{H}}_{\text{int}}^{(1)} + \hat{\mathcal{H}}_{\text{int}}^{(2)} \quad (5.1)$$

where $\hat{\mathcal{H}}_{\text{int}}^{(1)}$ corresponds to the already known first-order electron-phonon interaction Hamiltonian of Eq. (3.39), whereas $\hat{\mathcal{H}}_{\text{int}}^{(2)}$ is the interaction Hamiltonian related to the two-phonon coupling with an electron [86, 102, 101, 103]. The second-order electron-phonon interaction Hamiltonian $\hat{\mathcal{H}}_{\text{int}}^{(2)}$, which was recently reported in a DFT-based calculation of renormalized bulk band structures [101], is defined by [86, 102, 101, 103]

$$\begin{aligned} \hat{\mathcal{H}}_{\text{int}}^{(2)} = & \sum_{\substack{\alpha\beta \\ \mathbf{q}\nu\mathbf{q},\nu}} g_{\alpha\beta\mathbf{q}\nu\mathbf{q},\nu} \hat{c}_{\beta}^{\dagger} \hat{c}_{\alpha} \left(\hat{a}_{\mathbf{q}\nu} + \hat{a}_{-\mathbf{q}\nu}^{\dagger} \right) \left(\hat{a}_{\mathbf{q},\nu} + \hat{a}_{-\mathbf{q},\nu}^{\dagger} \right) \\ & + g_{\alpha\beta\mathbf{q}\nu\mathbf{q},\nu}^* \left(\hat{a}_{\mathbf{q},\nu}^{\dagger} + \hat{a}_{-\mathbf{q},\nu} \right) \left(\hat{a}_{\mathbf{q}\nu}^{\dagger} + \hat{a}_{-\mathbf{q}\nu} \right) \hat{c}_{\alpha}^{\dagger} \hat{c}_{\beta} \quad (5.2) \end{aligned}$$

where $g_{\alpha\beta\mathbf{q}\nu\mathbf{q},\nu}$ represents the coupling strength between an electron and two phonons. The corresponding time-dependent (in the interaction picture) electron-phonon interaction Hamiltonian is

$$\tilde{\mathcal{H}}_{\text{int}}(t) = \tilde{\mathcal{H}}_{\text{int}}^{(1)}(t) + \tilde{\mathcal{H}}_{\text{int}}^{(2)}(t) \quad (5.3)$$

where $\tilde{\mathcal{H}}_{\text{int}}^{(1)}(t)$ is equivalent to Eq. (3.52) and the time-dependent second-order electron-phonon interaction Hamiltonian $\tilde{\mathcal{H}}_{\text{int}}^{(2)}(t)$ is, using Eqs. (5.2) and (3.43)-

(3.51), is defined as:

$$\begin{aligned}
\tilde{\mathcal{H}}_{\text{int}}^{(2)}(t) = & \sum_{\substack{\alpha\beta \\ \mathbf{q}\nu\mathbf{q},\nu}} g_{\alpha\beta\mathbf{q}\nu\mathbf{q},\nu} \hat{c}_{\beta}^{\dagger} \hat{c}_{\alpha} \left[\hat{a}_{\mathbf{q}\nu} \hat{a}_{\mathbf{q},\nu} \exp\left(\frac{i(E_{\beta} - E_{\alpha} - \hbar\omega_{\mathbf{q}\nu} - \hbar\omega_{\mathbf{q},\nu})t}{\hbar}\right) \right. \\
& + \hat{a}_{\mathbf{q}\nu} \hat{a}_{-\mathbf{q},\nu}^{\dagger} \exp\left(\frac{i(E_{\beta} - E_{\alpha} - \hbar\omega_{\mathbf{q}\nu} + \hbar\omega_{-\mathbf{q},\nu})t}{\hbar}\right) \\
& + \hat{a}_{-\mathbf{q}\nu}^{\dagger} \hat{a}_{\mathbf{q},\nu} \exp\left(\frac{i(E_{\beta} - E_{\alpha} + \hbar\omega_{-\mathbf{q}\nu} - \hbar\omega_{\mathbf{q},\nu})t}{\hbar}\right) \\
& \left. + \hat{a}_{-\mathbf{q}\nu}^{\dagger} \hat{a}_{-\mathbf{q},\nu}^{\dagger} \exp\left(\frac{i(E_{\beta} - E_{\alpha} + \hbar\omega_{-\mathbf{q}\nu} + \hbar\omega_{-\mathbf{q},\nu})t}{\hbar}\right) \right] + \text{h.c.} \quad (5.4)
\end{aligned}$$

where ‘‘h.c.’’ stands for the hermitian conjugate. Substituting Eq. (5.3) in Eq. (3.53), the PAT current Eq. (3.53) becomes:

$$\begin{aligned}
I_{\text{ph}} = & \frac{e_q}{\hbar^2} \lim_{t \rightarrow \infty} \int_0^t dt_1 \mathbf{Tr} \left(\left[[\hat{N}_{\alpha}, \tilde{\mathcal{H}}_{\text{int}}^{(1)}(t)], \tilde{\mathcal{H}}_{\text{int}}^{(1)}(t_1) \right] \hat{\rho}_0 \right) \\
& + \frac{e_q}{\hbar^2} \lim_{t \rightarrow \infty} \int_0^t dt_1 \mathbf{Tr} \left(\left[[\hat{N}_{\alpha}, \tilde{\mathcal{H}}_{\text{int}}^{(2)}(t)], \tilde{\mathcal{H}}_{\text{int}}^{(1)}(t_1) \right] \hat{\rho}_0 \right) \\
& + \frac{e_q}{\hbar^2} \lim_{t \rightarrow \infty} \int_0^t dt_1 \mathbf{Tr} \left(\left[[\hat{N}_{\alpha}, \tilde{\mathcal{H}}_{\text{int}}^{(1)}(t)], \tilde{\mathcal{H}}_{\text{int}}^{(2)}(t_1) \right] \hat{\rho}_0 \right) \\
& + \frac{e_q}{\hbar^2} \lim_{t \rightarrow \infty} \int_0^t dt_1 \mathbf{Tr} \left(\left[[\hat{N}_{\alpha}, \tilde{\mathcal{H}}_{\text{int}}^{(2)}(t)], \tilde{\mathcal{H}}_{\text{int}}^{(2)}(t_1) \right] \hat{\rho}_0 \right) \quad (5.5)
\end{aligned}$$

It is obvious from Eq. (5.5) that the second and third terms contribute zero PAT current as these terms lead to zero expectation values, when using Wick’s

theorem [84]. Therefore, Eq. (5.5) is further simplified as:

$$\begin{aligned}
 I_{\text{ph}} = & \frac{e_q}{\hbar^2} \lim_{t \rightarrow \infty} \int_0^t dt_1 \text{Tr} \left(\left[\left[\hat{N}_\alpha, \tilde{\mathcal{H}}_{\text{int}}^{(1)}(t) \right], \tilde{\mathcal{H}}_{\text{int}}^{(1)}(t_1) \right] \hat{\rho}_0 \right) \\
 & + \frac{e_q}{\hbar^2} \lim_{t \rightarrow \infty} \int_0^t dt_1 \text{Tr} \left(\left[\left[\hat{N}_\alpha, \tilde{\mathcal{H}}_{\text{int}}^{(2)}(t) \right], \tilde{\mathcal{H}}_{\text{int}}^{(2)}(t_1) \right] \hat{\rho}_0 \right) \quad (5.6)
 \end{aligned}$$

Note that the first term in Eq. (5.6) corresponds to the PAT current formalism presented in Chapter 3, whereas the two-phonon assisted tunneling currents can be derived by solving the double-commutator of the last term in Eq. (5.6). So far, we proposed to determine the two-phonon assisted tunneling current formalism based on the density matrix framework outlined in Chapter 3, where we used the time-dependent interaction picture for the operators and the second-order electron-interaction Hamiltonian ($\hat{\mathcal{H}}_{\text{int}}^{(2)}$ of Eq. (5.1)) to determine two-PAT currents while maintaining a first-order expansion of the density matrix (Eq. (3.41)).

Note that the crucial element in determining the two-phonon assisted tunneling currents is the calculation of the second-order EPC strength $g_{\alpha\beta\mathbf{q}\nu\mathbf{q},\nu}$. It has been calculated with a DFT-based implementation, which determined the EPC strength for few nanometer wide bulk structures and in the absence of external applied voltages [101]. At the moment of writing this thesis, these DFT-based calculations (of $g_{\alpha\beta\mathbf{q}\nu\mathbf{q},\nu}$) have not been reported yet for realistic device sizes and structures (heterostructures). We therefore present the method to calculate second order EPC strength in the presence of external applied voltages by generalizing the approximations used for the DFT-based calculations. In our method, the effects of the external applied voltages are intrinsically included in the wavefunction description of electrons, which will be used in determining second-order EPC strengths. In the aforementioned bra-ket notation (see Eq. (3.57) of Section 3.3), the second-order EPC strength is defined by the following second-order electron-lattice interaction potential [86],

$$g_{\alpha\beta\mathbf{q}\nu\mathbf{q},\nu} = \frac{1}{2} \langle \mathbf{k}_\beta | \Delta^2 \mathcal{V} | \mathbf{k}_\alpha \rangle \equiv \frac{1}{2} \langle \mathbf{k}_\alpha + \mathbf{q} + \mathbf{q}_\nu | \Delta^2 \mathcal{V} | \mathbf{k}_\alpha \rangle \quad (5.7)$$

whereby the second-order electron-lattice interaction potential is described as the following second-order change in crystal lattice V_c potential due to the atomic displacements $\left(\Delta \mathbf{R}_{u,i}^{j,\nu}, \Delta \mathbf{R}_{u,l}^{m,\nu'} \right)$ [86]

$$\Delta^2 \mathcal{V} = \sum_{i,j,\nu} \sum_{l,m,\nu'} \frac{\partial^2 V_c}{\partial \mathbf{R}_{u,i}^{j,\nu} \partial \mathbf{R}_{u,l}^{m,\nu'}} \Delta \mathbf{R}_{u,i}^{j,\nu} \Delta \mathbf{R}_{u,l}^{m,\nu'} \quad (5.8)$$

The exact calculation of the second order EPC strength (Eq. (5.7)) for bulk-materials is challenging from a computational point of view as it would require the application of second-order density functional perturbation theory (DFPT) [101]. To circumvent this, Allen and Heine [102] proposed the recasting of Eq. (5.8), which is the second-order electron-lattice interaction potential defined in first-order perturbation theory, into the linear combination of two first-order electron-lattice interaction potentials defined in second-order perturbation theory (refer Eqs. 1-4 of [103]), while applying the observed translational invariance of atomic displacements within the harmonic and adiabatic approximations (these approximations are explained in Appendix A). These approximations are coupled with an additional condition that if every atom is further displaced by $\Delta\mathbf{R}_{u,i}^{j,\nu}$, the energy of the total crystal is unchanged which prevents anharmonicity. This method is referred to as “adiabatic Allen-Heine formula” of monoatomic crystals [102], which is later extended to polyatomic unit cells by Allen and Cardona [103]. With these approximations, the second-order EPC is simplified as

$$g_{\alpha\beta\mathbf{q}\nu\mathbf{q},\nu'} = -\frac{1}{2} \sum_{\mathbf{k}_{I\alpha}=\mathbf{k}_\alpha+\mathbf{q},} \frac{\langle \mathbf{k}_\alpha + \mathbf{q} + \mathbf{q}_l | \Delta\mathcal{V}_{u,i}^{j,\nu} | \mathbf{k}_{I\alpha} \rangle \langle \mathbf{k}_{I\alpha} | \Delta\mathcal{V}_{u,l}^{m,\nu'} | \mathbf{k}_\alpha \rangle}{E_\beta - E_I} - \frac{1}{2} \sum_{\mathbf{k}_{I\beta}=\mathbf{k}_\beta-\mathbf{q},} \frac{\langle \mathbf{k}_\alpha + \mathbf{q} + \mathbf{q}_l | \Delta\mathcal{V}_{u,l}^{m,\nu'} | \mathbf{k}_{I\beta} \rangle \langle \mathbf{k}_{I\beta} | \Delta\mathcal{V}_{u,i}^{j,\nu} | \mathbf{k}_\alpha \rangle}{E_\beta - E_I} \quad (5.9)$$

with

$$\Delta\mathcal{V}_{u,i}^{j,\nu} = \sum_{i,j,\nu} \frac{\partial V_c}{\partial \mathbf{R}_{u,i}^{j,\nu}} \Delta\mathbf{R}_{u,i}^{j,\nu}; \quad \Delta\mathcal{V}_{u,l}^{m,\nu'} = \sum_{l,m,\nu'} \frac{\partial V_c}{\partial \mathbf{R}_{u,l}^{m,\nu'}} \Delta\mathbf{R}_{u,l}^{m,\nu'} \quad (5.10)$$

where $|\mathbf{k}_{I\alpha}\rangle$ represents an intermediate state injected from the left contact and identified with $\mathbf{k}_{I\alpha} = \mathbf{k}_\alpha + \mathbf{q}_l$ at energy E_I . Similarly, $|\mathbf{k}_{I\beta}\rangle$ represents an intermediate state identified with $\mathbf{k}_{I\beta} = \mathbf{k}_\beta - \mathbf{q}_l$, while injected from the right contact at energy E_I . Note that the denominator in Eq. (5.9) represents a phonon energy ($E_\beta - E_I = \hbar\omega_{\mathbf{q},\nu'}$).

Using the definition of the first-order EPC strengths from Eq. (3.59), the second-order EPC strength can further be defined in the following form:

$$g_{\alpha\beta\mathbf{q}\nu\mathbf{q},\nu_i} = -\frac{1}{2} \left[\sum_{\mathbf{k}_{I\alpha}=\mathbf{k}_\alpha+\mathbf{q}} \frac{g_{I\alpha\beta\mathbf{q}\nu} * g_{\alpha I_\alpha\mathbf{q},\nu_i}}{E_\beta - E_I} + \sum_{\mathbf{k}_{I\beta}=\mathbf{k}_\beta-\mathbf{q}} \frac{g_{I\beta\mathbf{q}\nu} * g_{\alpha I_\beta\mathbf{q},\nu_i}}{E_\beta - E_I} \right] \quad (5.11)$$

$$\begin{aligned} &= -\frac{1}{2} \sum_{\mathbf{k}_{I\alpha}=\mathbf{k}_\alpha+\mathbf{q}} \frac{M_{\mathbf{q}\nu} M_{\mathbf{q},\nu_i}}{E_\beta - E_I} \\ &\quad \times \int d\mathbf{r} \psi_\beta^*(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} \psi_{I\alpha}(\mathbf{r}) \int d\mathbf{r}' \psi_{I\alpha}^*(\mathbf{r}') e^{i\mathbf{q}\cdot\mathbf{r}'} \psi_\alpha(\mathbf{r}') \\ &- \frac{1}{2} \sum_{\mathbf{k}_{I\beta}=\mathbf{k}_\beta-\mathbf{q}} \frac{M_{\mathbf{q}\nu} M_{\mathbf{q},\nu_i}}{E_\beta - E_I} \\ &\quad \times \int d\mathbf{r} \psi_\beta^*(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} \psi_{I\beta}(\mathbf{r}) \int d\mathbf{r}' \psi_{I\beta}^*(\mathbf{r}') e^{i\mathbf{q}\cdot\mathbf{r}'} \psi_\alpha(\mathbf{r}') \end{aligned} \quad (5.12)$$

where $\psi_{I\alpha}(\mathbf{r})$ represent the intermediate-state's electron wavefunction, injected from the left contact at energy E_I , characterized by either an emission or absorption of a single phonon $\hbar\omega_{\mathbf{q},\nu_i}$, from the initial α state. Similarly, $\psi_{I\beta}(\mathbf{r})$ represent the intermediate-state's electron wavefunction, injected from the right contact at energy E_I , characterized by either an emission or absorption of a single phonon $\hbar\omega_{\mathbf{q},\nu_i}$, from the final β state. Note that the above simplification for the second-order EPC is also applicable for the heterostructure, where the first-order EPC in Eq. (5.11) can be replaced with the definition of Eq. (3.82). With the above simplifications for $g_{\alpha\beta\mathbf{q}\nu\mathbf{q},\nu_i}$, we expect that the calculation of two-phonon assisted tunneling currents will be a straightforward extension of the implementation presented in Chapters 3 and 4.

After determining the multi-phonon assisted tunneling currents, which can generally be extrapolated from the aforementioned combination of single and two-PAT currents, the TAT currents in semiconductors can eventually be determined as

$$J_{\text{TAT}} = J_{\text{ph}}^{\text{Trap}} - J_{\text{ph}}^{\text{noTrap}} \quad (5.13)$$

where J_{TAT} is the desired TAT current density, $J_{\text{ph}}^{\text{Trap}}$ is the multi-PAT current density calculated with the chosen trap potential profile and $J_{\text{ph}}^{\text{noTrap}}$ is the multi-PAT current density determined with a trap-free potential profile. Note that the TAT current Eq. (5.13) does not account for the structural relaxation

discussed in Appendix A. However, we suggest that the structural relaxation process can be included in the formalism described in Section 3.2 of Chapter 3, specifically, by including the trap subsystem Hamiltonian in the unperturbed system Eq. (3.33). In the framework of second quantization, the trap subsystem Hamiltonian can be defined by the so-called impurity models such as the Kondo model [104] and Anderson impurity model [105, 78]. We expect that such addition of models could result in the current equation similar to Eq. (3.56) inclusive of trap energy level and occupational probability.

In addition to the above suggestions, we expect that the extension of the PAT current density formalism to tri-dimensional semiconductor systems such as nanowire-TFETs, by following again the framework described in Chapter 3, could enable us to determine the TAT currents for realistic trap configurations.

Appendix A: Fundamentals of NMP theory

In the NMP theory, the effective Hamiltonian describing the total system of electrons and lattice vibrations [106] is:

$$H(\mathbf{r}, \mathbf{Q}) = H_{\text{el}}(\mathbf{r}) + V_{\text{el-ph}}(\mathbf{r}, \mathbf{Q}) + H_{\text{ph}}(\mathbf{Q}) \quad (\text{A.1})$$

where $H_{\text{el}}(\mathbf{r})$ is the Hamiltonian for the electronic subsystem, $H_{\text{ph}}(\mathbf{Q})$ is the Hamiltonian of the free lattice and $V_{\text{el-ph}}(\mathbf{r}, \mathbf{Q})$ is the electron-phonon coupling potential, \mathbf{r} are electronic coordinates and \mathbf{Q} are normal coordinates representing the displacement of an atom from its equilibrium position. To solve the Schrödinger equation ($H\Psi = \mathcal{E}\Psi$) with the Hamilton-operator (Eq. (A.1)), the delocalized electrons of the system are assumed to follow the motion of defects and of host material atoms instantaneously, which is the well-known Born-Oppenheimer approximation [106]. This approximates the total wavefunction of the system of Eq. (A.1) as the following product of the electron and the lattice vibration wavefunctions with 'n' in $\Phi_{n,\lambda}(\mathbf{Q})$ implying that the lattice vibrational state will generally depend on the state of the electronic subsystem.

$$\Psi_{n,\lambda}(\mathbf{r}, \mathbf{Q}) = \psi_n(\mathbf{r}, \mathbf{Q})\Phi_{n,\lambda}(\mathbf{Q}) \quad (\text{A.2})$$

where n, λ denote the electron and phonon quantum numbers corresponding to the electronic and vibrational subsystems, and where $\psi_n(\mathbf{r}, \mathbf{Q}), \Phi_{n,\lambda}(\mathbf{Q})$ are the solutions of the electronic and vibrational subsystems, respectively. With Eq. (A.2), the Schrödinger equation of the total system becomes

$$\begin{aligned} \Phi_{n,\lambda}(\mathbf{Q})H_{\text{el}}(\mathbf{r})\psi_n(\mathbf{r}, \mathbf{Q}) + V_{\text{el-ph}}(\mathbf{r}, \mathbf{Q}) [\psi_n(\mathbf{r}, \mathbf{Q})\Phi_{n,\lambda}(\mathbf{Q})] \\ + H_{\text{ph}}(\mathbf{Q}) [\psi_n(\mathbf{r}, \mathbf{Q})\Phi_{n,\lambda}(\mathbf{Q})] = \mathcal{E}_{n,\lambda} [\psi_n(\mathbf{r}, \mathbf{Q})\Phi_{n,\lambda}(\mathbf{Q})] \end{aligned} \quad (\text{A.3})$$

Subtracting $\psi_n(\mathbf{r}, \mathbf{Q})H_{\text{ph}}(\mathbf{Q})\Phi_{n,\lambda}(\mathbf{Q})$ while also dividing by the approximated form of the total wavefunction (Eq. (A.2)), Eq. (A.3) of the total system becomes

$$\begin{aligned} & \frac{1}{\psi_n(\mathbf{r}, \mathbf{Q})} H_{\text{el}}(\mathbf{r})\psi_n(\mathbf{r}, \mathbf{Q}) + V_{\text{el-ph}}(\mathbf{r}, \mathbf{Q}) \\ & + \frac{1}{\psi_n(\mathbf{r}, \mathbf{Q})\Phi_{n,\lambda}(\mathbf{Q})} \left\{ \widehat{\mathcal{L}}(\psi_n(\mathbf{r}, \mathbf{Q})\Phi_{n,\lambda}(\mathbf{Q})) \right\} = \mathcal{E}_{n,\lambda} - \frac{1}{\Phi_{n,\lambda}(\mathbf{Q})} H_{\text{ph}}(\mathbf{Q})\Phi_{n,\lambda}(\mathbf{Q}) \end{aligned} \quad (\text{A.4})$$

where, the non-adiabatic operator $\widehat{\mathcal{L}}$ is defined by

$$\begin{aligned} \widehat{\mathcal{L}}(\psi_n(\mathbf{r}\mathbf{Q})\Phi_{n,\lambda}(\mathbf{Q})) &= H_{\text{ph}}(\mathbf{Q})\psi_n(\mathbf{r}, \mathbf{Q})\Phi_{n,\lambda}(\mathbf{Q}) - \psi_n(\mathbf{r}, \mathbf{Q})H_{\text{ph}}(\mathbf{Q})\Phi_{n,\lambda}(\mathbf{Q}) \\ &\equiv [H_{\text{ph}}(\mathbf{Q}), \psi_n(\mathbf{r}, \mathbf{Q})] \Phi_{n,\lambda}(\mathbf{Q}) \end{aligned} \quad (\text{A.5})$$

In the framework of the adiabatic approximation, where the electronic subsystem is assumed to follow the motion of the host material atoms adiabatically, the term with the non-adiabatic operator “ \mathcal{L} ” is zero [106]. With the following definition of the electronic energy in terms of normal coordinates,

$$E_n(\mathbf{Q}) = \mathcal{E}_{n,\lambda} - \frac{1}{\Phi_{n,\lambda}(\mathbf{Q})} H_{\text{ph}}(\mathbf{Q})\Phi_{n,\lambda}(\mathbf{Q}) \quad (\text{A.6})$$

the total system of Eq. (A.4) is separated into the following set of coupled electronic and vibrational subsystems, which allows for the separate determination of electronic and vibrational wavefunctions:

$$[H_{\text{el}}(\mathbf{r}) + V_{\text{el-ph}}(\mathbf{r}, \mathbf{Q})] \psi_n(\mathbf{r}, \mathbf{Q}) = E_n(\mathbf{Q})\psi_n(\mathbf{r}, \mathbf{Q}) \quad (\text{A.7})$$

$$[H_{\text{ph}}(\mathbf{Q}) + E_n(\mathbf{Q})] \Phi_{n,\lambda}(\mathbf{Q}) = \mathcal{E}_{n,\lambda}\Phi_{n,\lambda}(\mathbf{Q}) \quad (\text{A.8})$$

Here, the electron energies $E_n(\mathbf{Q})$, which act as a coupling between the electronic and vibrational subsystems, are additional contributions to the potential of the nuclear oscillations in the lattice equation (Eq. (A.8)).

In the general case, the adiabatic system of Eqs. (A.7) and (A.8) can not be solved exactly. To overcome this complexity the well-known Franck-Condon approximation of the adiabatic wavefunctions ($\psi_n(\mathbf{r}, \mathbf{Q}), \Phi_{n,\lambda}(\mathbf{Q})$) is considered [63] [51], whereby it is assumed that the lattice cannot readjust instantaneously when an electron capture or emission occurs. This results in the electronic wavefunction of Eq. (A.7) to be expanded into an electronic wavefunction independent of the normal coordinates \mathbf{Q} , and perturbation terms using perturbation theory. In the framework of the Condon-approximation, the

approximated total wavefunction of Eq. (A.2) is

$$\Psi_{n,\lambda}(\mathbf{r}, \mathbf{Q}) = \left(\psi_n^{(0)}(\mathbf{r}) + \sum_{n \neq m} \frac{V_{\text{el-ph}}^{(0)}(\mathbf{r}, \mathbf{Q})}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)}(\mathbf{r}) \right) \Phi_{n,\lambda}^{(1)}(\mathbf{Q}) \quad (\text{A.9})$$

$$\approx \psi_n^{(0)}(\mathbf{r}) \Phi_{n,\lambda}^{(1)}(\mathbf{Q}) \quad (\text{A.10})$$

where $\psi_n^{(0)}(\mathbf{r})$ refer to the stationary state eigenfunctions of the electronic Hamiltonian $H_{\text{el}}(\mathbf{r})$ which are assumed to simultaneously diagonalize the total system Hamiltonian $H(\mathbf{r}, \mathbf{Q})$, and where $\Phi_{n,\lambda}^{(1)}(\mathbf{Q})$ refers to the first-order perturbed lattice vibrational function due to the electron-lattice interaction. In order to make the calculations of transition probabilities (probabilities of either a carrier capture or emission event) feasible, the higher order non-diagonal contributions of the electronic wavefunctions in Eq. (A.9) are usually neglected in the conventional NMP theory, which states that the capture or release of a carrier occurs through a combination of ballistic tunneling and multiphonon emission or absorption during structural relaxation of the lattice site. However, these non-diagonal terms can correspond to the phonon-assisted NMP process, which would be a second order effect of the conventional NMP theory (hence ignored) [107], because they represent electron-phonon coupling during the electronic transitions (tunneling transitions of the NMP theory). Note that these assumptions are similar to the static approach of adiabatic theory [107]. The total system of Eq. (A.1), while using the definition of Eq. (A.10) and keeping in mind the identity of stationary electronic states $\left(\hat{I} = \sum_{n'} |\psi_{n'}^{(0)}\rangle \langle \psi_{n'}^{(0)}| \right)$, can be written as

$$\begin{aligned} \sum_{n'} |\psi_{n'}^{(0)}\rangle \langle \psi_{n'}^{(0)}| [H_{\text{el}}(\mathbf{r}) + V_{\text{el-ph}}(\mathbf{r}, \mathbf{Q}) + H_{\text{ph}}(\mathbf{Q})] |\psi_n^{(0)}\rangle \Phi_{n,\lambda}^{(1)}(\mathbf{Q}) \\ = \sum_{n'} |\psi_{n'}^{(0)}\rangle \langle \psi_{n'}^{(0)}| \mathcal{E}_{n,\lambda} |\psi_n^{(0)}\rangle \Phi_{n,\lambda}^{(1)}(\mathbf{Q}) \end{aligned} \quad (\text{A.11})$$

The orthonormality condition of the electronic state eigenfunctions $\left(\langle \psi_{n'}^{(0)} | \psi_n^{(0)} \rangle = \delta_{n,n'} \right)$ brings Eq. (A.11) to

$$\left[E_n^{(0)} + \langle \psi_n^{(0)} | V_{\text{el-ph}}(\mathbf{r}, \mathbf{Q}) | \psi_n^{(0)} \rangle + H_{\text{ph}}(\mathbf{Q}) \right] \Phi_{n,\lambda}^{(1)}(\mathbf{Q}) = \mathcal{E}_{n,\lambda} \Phi_{n,\lambda}^{(1)}(\mathbf{Q}) \quad (\text{A.12})$$

Comparing Eqs. (A.12) and (A.8), the lattice-dependent electronic energy is a linear combination of stationary state electronic energy and the expectation value of the electron-phonon coupling strength, which is

$$E_n(\mathbf{Q}) \approx E_n^{(0)} + \langle \psi_n^{(0)} | V_{\text{el-ph}}(\mathbf{r}, \mathbf{Q}) | \psi_n^{(0)} \rangle \quad (\text{A.13})$$

Combining this approximation with Eq. (A.10), the total system of electronic and lattice vibrations (Eqs. (A.7) and (A.8)) is transformed into the following simplified set of coupled equations:

$$H_{\text{el}}(\mathbf{r})\psi_n^{(0)}(\mathbf{r}) = E_n^{(0)}\psi_n^{(0)}(\mathbf{r}) \quad (\text{A.14})$$

$$\left[H_{\text{ph}}(\mathbf{Q}) + \left\langle \psi_n^{(0)} \left| V_{\text{el-ph}}(\mathbf{r}, \mathbf{Q}) \right| \psi_n^{(0)} \right\rangle + E_n^{(0)} \right] \Phi_{n,\lambda}^{(1)}(\mathbf{Q}) = \mathcal{E}_{n,\lambda} \Phi_{n,\lambda}^{(1)}(\mathbf{Q}) \quad (\text{A.15})$$

Note that the simplifications made from Eq. (A.1) until Eq. (A.15) are equally carried out in formulating the Schenk TAT [41] and the extended-NMP based TAT model [42], however with different approaches to determine the tunneling and thermal transition rates (briefly described in Section 1.5.2 of Chapter 1). In order to make the calculation of thermal transition probabilities feasible with Eq. A.15, the harmonic approximation for the phonon Hamiltonian $H_{\text{ph}}(\mathbf{Q})$ is applied, which states that the lattice displacement during vibrations around their equilibrium position is small, and results in a quadratic dependence on the normal coordinates:

$$H_{\text{ph}}(\mathbf{Q})\Phi_{n,\lambda}^{(1)}(\mathbf{Q}) = A Q^2 \Phi_{n,\lambda}^{(1)}(\mathbf{Q}) \quad (\text{A.16})$$

Remember that the NMP theory is characterized by ballistic tunneling followed by multiphonon emission (or absorption) due to structural relaxation upon the capture (or release) of a carrier. The interpretation of these underlying mechanisms can be explained by analyzing Eqs. (A.6), (A.13), (A.14) and (A.15). First, the ballistic tunneling can be determined from the tunneling transition rates, obtained while applying Fermi's Golden rule, by using the wavefunctions of the electronic subsystem Eq. (A.14).

Second, the structural relaxation is to be understood as follows. The total energy of the system of delocalized electrons and the lattice site (defect) prior to capture of an electron, while using the aforementioned harmonic and adiabatic approximations, can be written from Eqs. (A.6) and (A.13) as

$$\mathcal{E}_{n,\lambda} = E_n(\mathbf{Q}) + \frac{1}{\Phi_{n,\lambda}(\mathbf{Q})} H_{\text{ph}}(\mathbf{Q}) \Phi_{n,\lambda}(\mathbf{Q}) \equiv E_n^{(0)} + A Q^2 \quad (\text{A.17})$$

In this case, we assumed insignificant coupling, $\left\langle \psi_n^{(0)} \left| V_{\text{el-ph}}(\mathbf{r}, \mathbf{Q}) \right| \psi_n^{(0)} \right\rangle \approx 0$, in Eq. (A.13), between the delocalized electron and the lattice site and we further assumed that the system energy is minimal when the lattice is in its undistorted equilibrium ($Q = 0$). The capture of an electron at the lattice site readjusts the bonding of the lattice ions with their nearest neighbor ions. The lattice site is then polarized in the vicinity of the captured electron and subsequently the equilibrium position of the lattice ions changes. This effect is usually

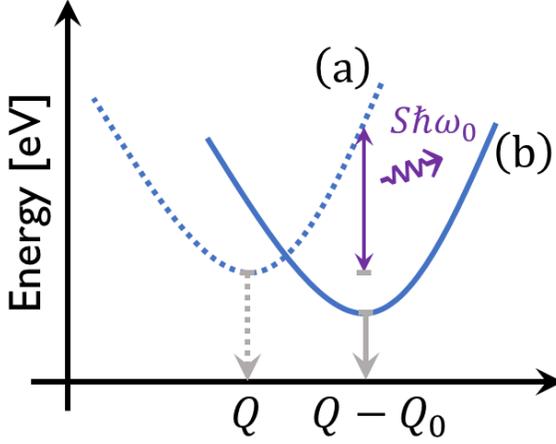


Figure A.1: Graphical illustration of configuration coordinate (CC) diagram of (a) a system of a delocalized electron and the lattice (defect) site prior to capture of an electron and (b) a system when an electron is captured at the lattice (defect) site.

described by $\langle \psi_n^{(0)} | V_{\text{el-ph}}(\mathbf{r}, \mathbf{Q}) | \psi_n^{(0)} \rangle$ in Eq. (A.13). The linear approximation of $\langle \psi_n^{(0)} | V_{\text{el-ph}}(\mathbf{r}, \mathbf{Q}) | \psi_n^{(0)} \rangle$, which states that the lattice deformation linearly modulates the electron energy, is defined as

$$\delta E_n^{(0)} = -BQ = \frac{\partial E_n^{(0)}}{\partial Q} Q \quad (\text{A.18})$$

where B is the deformation potential. Upon electron capture, the total system energy, while using Eqs. (A.6), (A.13) and (A.18), becomes

$$\begin{aligned} \mathcal{E}_{n,\lambda} &= E_n^{(0)} + AQ^2 - BQ \\ &= E_n^{(0)} + A(Q - Q_0)^2 - AQ_0^2 \end{aligned} \quad (\text{A.19})$$

where we defined $Q_0 = B/2A$. Comparing Eqs. (A.17) and (A.19), it is clear that the total system energy, upon the capture of an electron at the lattice site, is decreased with the minimum energy AQ_0^2 and the normal coordinate shifts from Q to $Q - Q_0$. This phenomenon is known as structural relaxation, which is graphically illustrated by the configuration coordinate diagram in Fig. (A.1). The difference in the total system energy between Eqs. (A.17) and (A.19), which is due to structural relaxation, is released in the form of multiphonon. Hence,

the theory is known as non-radiative multiphonon (NMP) recombination. There are two widely used TAT models in literature which are based on the application of the NMP theory. These models are briefly explained in Section 1.5.2, while discussing their limitations in Section 1.5.3 of Chapter 1.

Appendix B: Fundamentals of equilibrium statistical mechanics

In this appendix, the statistical average of the occupation number operator of a single particle state in a grand canonical ensemble, a many-particle system characterized by the exchange of particles with an external reservoir, is determined.

B.1 Fermi-Dirac statistics

As mentioned in the beginning of Section 3.1, we are interested in determining the statistical average of a single-electron state occupation number operator ($\langle \hat{n}_\vartheta \rangle$) from the system of many electrons. It can be determined with Eq. (3.15), while using the single-electron state occupation number operator (Eq. (3.8)) and the many-electron density operator (Eq. (3.18)) and results in [76]

$$\langle \hat{n}_\vartheta \rangle = \mathbf{Tr} (\hat{\varrho}_{\text{el}} \hat{n}_\vartheta) \equiv \frac{\mathbf{Tr} \left(\left(\exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_{\text{el}} - \mu \hat{N}_{\text{el}} \right) \right\} \right) \hat{n}_\vartheta \right)}{\mathcal{Z}_{\text{el}}} \quad (\text{B.1})$$

Using the definition of Trace operator from Eq. (3.14), the grand canonical partition function of electrons in Eq. (B.1), however in occupation number

representation, can be calculated from the following expression:

$$\begin{aligned} \mathcal{Z}_{\text{el}} &= \mathbf{Tr} \left(\exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_{\text{el}} - \mu \hat{N}_{\text{el}} \right) \right\} \right) \\ &= \sum_{n_1, n_2, \dots = 0}^1 \langle n_1, n_2, \dots | \exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_{\text{el}} - \mu \hat{N}_{\text{el}} \right) \right\} | n_1, n_2, \dots \rangle \end{aligned} \quad (\text{B.2})$$

The exponential function of the operator can be interpreted in the following Taylor series expansion around $-\Gamma \left(\hat{\mathcal{H}}_{\text{el}} - \mu \hat{N}_{\text{el}} \right) = 0$:

$$\exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_{\text{el}} - \mu \hat{N}_{\text{el}} \right) \right\} = \sum_{j=0}^{\infty} \frac{\left\{ -\Gamma \left(\hat{\mathcal{H}}_{\text{el}} - \mu \hat{N}_{\text{el}} \right) \right\}^j}{j!} \quad (\text{B.3})$$

Using the fact that an occupation vector $|n_1, n_2, \dots\rangle$ is simultaneously an eigenvector of $\hat{\mathcal{H}}_{\text{el}}$ and \hat{N}_{el} with eigen values from Eqs. (3.9) and (3.7) as E and N , respectively, one obtains $\exp \left\{ -\Gamma (E - \mu N) \right\}$ by applying the above Taylor series on this eigenvector and reconstructing the eigen values back into an exponential function. With the descriptions of a non-interacting N-electrons system and using Eq. (3.10), the grand canonical partition function can further be simplified as:

$$\begin{aligned} \mathcal{Z}_{\text{el}} &= \sum_{n_1, n_2, \dots = 0}^1 \exp \left\{ -\Gamma (E - \mu N) \right\} \langle n_1, n_2, \dots | n_1, n_2, \dots \rangle \\ &= \sum_{n_1, n_2, \dots = 0}^1 \exp \left\{ -\Gamma \sum_{\vartheta=1}^{\infty} (E_{\vartheta} - \mu) n_{\vartheta} \right\} \end{aligned} \quad (\text{B.4})$$

Using the property $\left[\exp \left(\sum_j a_j \right) = \prod_j \exp(a_j) \right]$ of the exponential function and subsequently applying the identity $\exp(ij) = (\exp(i))^j$, the grand canonical partition function can be rewritten as [76]:

$$\begin{aligned} \mathcal{Z}_{\text{el}} &= \sum_{n_1, n_2, \dots = 0}^1 \prod_{\vartheta=1}^{\infty} \exp \left\{ -\Gamma (E_{\vartheta} - \mu) n_{\vartheta} \right\} \\ &= \sum_{n_1, n_2, \dots = 0}^1 \prod_{\vartheta=1}^{\infty} \left(\exp \left\{ -\Gamma (E_{\vartheta} - \mu) \right\} \right)^{n_{\vartheta}} \end{aligned} \quad (\text{B.5})$$

Rewriting Eq. (B.5) and applying the possible single-electron state occupation numbers ($n_k = 0, 1$) simplifies the partition function as

$$\begin{aligned} \mathcal{Z}_{\text{el}} &= \left[\sum_{n_1=0}^1 (\exp \{-\Gamma(E_1 - \mu)\})^{n_1} \right] \left[\sum_{n_2=0}^1 (\exp \{-\Gamma(E_2 - \mu)\})^{n_2} \right] \dots \\ &= \prod_{\vartheta=1}^{\infty} \sum_{n_{\vartheta}=0}^1 [\exp \{-\Gamma(E_{\vartheta} - \mu)\}]^{n_{\vartheta}} \equiv \prod_{\vartheta=1}^{\infty} [1 + \exp \{-\Gamma(E_{\vartheta} - \mu)\}] \quad (\text{B.6}) \end{aligned}$$

Since the partition function is determined, the statistical average of a single-state occupation number can be calculated. From Eq. (B.1), it is rewritten in the occupation number representation as:

$$\langle \hat{n}_{\vartheta} \rangle = \frac{1}{\mathcal{Z}_{\text{el}}} \sum_{n_1, n_2, \dots=0}^1 \langle n_1, n_2, \dots | (\exp \{-\Gamma(\hat{\mathcal{H}}_{\text{el}} - \mu \hat{N})\}) \hat{n}_{\vartheta} | n_1, n_2, \dots \rangle \quad (\text{B.7})$$

With the Taylor series expansion for the exponential function of operators and the fact that a vector $|n_1, n_2, \dots\rangle$ is also an eigenvector of \hat{n}_{ϑ} , the statistical average is simplified by applying Eq. (B.3) and subsequently, by using Eqs. (3.7)-(3.10), which results in the following expression for the statistical average [76]:

$$\langle \hat{n}_{\vartheta} \rangle = \frac{1}{\mathcal{Z}_{\text{el}}} \sum_{n_1, n_2, \dots=0}^1 (\exp \{-\Gamma(E - \mu N)\}) \langle n_1, n_2, \dots | \hat{n}_{\vartheta} | n_1, n_2, \dots \rangle \quad (\text{B.8})$$

$$= \frac{1}{\mathcal{Z}_{\text{el}}} \left[\sum_{n_1, n_2, \dots=0}^1 \left(\exp \left\{ -\Gamma \sum_{\vartheta=1}^{\infty} (E_{\vartheta} - \mu) n_{\vartheta} \right\} \right) n_{\vartheta} \right] \quad (\text{B.9})$$

The square bracket term of Eq. (B.9) can be rewritten, while using the partial derivative with respect to an energy E_{ϑ} of a particular single-electron state ϑ [76], in the following form:

$$\langle \hat{n}_{\vartheta} \rangle = \frac{1}{\mathcal{Z}_{\text{el}}} \left[-\frac{1}{\Gamma} \frac{\partial}{\partial E_{\vartheta}} \left[\sum_{n_1, n_2, \dots=0}^1 \exp \left\{ -\Gamma \sum_{\vartheta=1}^{\infty} (E_{\vartheta} - \mu) n_{\vartheta} \right\} \right] \right] \quad (\text{B.10})$$

Using Eq. (B.4), the statistical average Eq. (B.10) is rewritten as

$$\langle \hat{n}_{\vartheta} \rangle = \frac{1}{\mathcal{Z}_{\text{el}}} \left[-\frac{1}{\Gamma} \frac{\partial}{\partial E_{\vartheta}} \mathcal{Z}_{\text{el}} \right] \equiv \left[-\frac{1}{\Gamma} \frac{\partial}{\partial E_{\vartheta}} (\ln(\mathcal{Z}_{\text{el}})) \right] \quad (\text{B.11})$$

Substituting Eq. (B.6) in Eq. (B.11) for the partition function, the statistical average of a single-electron state occupation number is further simplified as:

$$\langle \hat{n}_{\vartheta} \rangle = \frac{\exp \{-\Gamma(E_{\vartheta} - \mu)\}}{1 + \exp \{-\Gamma(E_{\vartheta} - \mu)\}} \equiv f_{\text{FD}}(E_{\vartheta} - \mu) \quad (\text{B.12})$$

where f_{FD} is the well-known Fermi-Dirac probability distribution function of electrons, which is described in its final form by

$$f_{\text{FD}}(E_{\vartheta} - \mu) = \frac{1}{1 + \exp \{ \Gamma (E_{\vartheta} - \mu) \}} \quad (\text{B.13})$$

Therefore, the statistical average of a single-electron state occupation number in a grand canonical ensemble results in the Fermi-Dirac statistics of the electron.

B.2 Bose-Einstein statistics

In a grand canonical ensemble, the statistical average of a single-phonon state occupation number operator ($\langle \hat{m}_{\lambda} \rangle$) can be determined, in the occupation number representation [76], as follows:

$$\langle \hat{m}_{\lambda} \rangle = \frac{1}{\mathcal{Z}_{\text{ph}}} \sum_{m_1, m_2, \dots=0}^{\infty} \langle m_1, m_2, \dots | \exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_{\text{ph}} - \mu_{\text{ph}} \hat{M} \right) \right\} \hat{m}_{\lambda} | m_1, m_2, \dots \rangle \quad (\text{B.14})$$

Using the general definition of the grand canonical partition function, which is similar to the case of electrons in the previous subsection, the partition function of free phonons in Eq. (B.14) can be described in terms of the occupation number representation, which is further simplified by applying the Taylor series expansion for the exponential function of operators on the eigenvectors $|m_1, m_2, \dots\rangle$ and using Eqs. (3.20) and (3.22), which results in the following equation [76]:

$$\begin{aligned} \mathcal{Z}_{\text{ph}} &= \sum_{m_1, m_2, \dots=0}^{\infty} \langle m_1, m_2, \dots | \exp \left\{ -\Gamma \left(\hat{\mathcal{H}}_{\text{ph}} - \mu_{\text{ph}} \hat{M}_{\text{ph}} \right) \right\} | m_1, m_2, \dots \rangle \\ &= \sum_{m_1, m_2, \dots=0}^{\infty} \exp \{ -\Gamma (E_{\text{ph}} - \mu_{\text{ph}} M) \} \\ &= \sum_{m_1, m_2, \dots=0}^{\infty} \exp \left\{ -\Gamma \sum_{\lambda=1}^{\infty} (E_{\lambda} - \mu_{\text{ph}}) m_{\lambda} \right\} \end{aligned} \quad (\text{B.15})$$

Using the aforementioned properties of the exponential function and similar to the transitions made from Eq. (B.4) till Eq. (B.6), the phonons' grand canonical partition equation can be rewritten by using all possible occupation numbers

for a single-phonon state, in the following forms [76]:

$$\begin{aligned} \mathcal{Z}_{\text{ph}} &= \left[\sum_{m_1=0}^{\infty} (\exp \{-\Gamma (E_1 - \mu_{\text{ph}})\})^{m_1} \right] \left[\sum_{m_2=0}^{\infty} (\exp \{-\Gamma (E_2 - \mu_{\text{ph}})\})^{m_2} \right] \dots \\ &= \prod_{\lambda=1}^{\infty} \sum_{m_{\lambda}=0}^{\infty} (\exp \{-\Gamma (E_{\lambda} - \mu_{\text{ph}})\})^{m_{\lambda}} \end{aligned} \quad (\text{B.16})$$

The summation in Eq. (B.16) represents a geometric series, which can be simplified and has the following value [76]:

$$\sum_{m_{\lambda}=0}^{\infty} (\exp \{-\Gamma (E_{\lambda} - \mu_{\text{ph}})\})^{m_{\lambda}} = \frac{1}{1 - \exp \{-\Gamma (E_{\lambda} - \mu_{\text{ph}})\}} \quad (\text{B.17})$$

Substituting Eq. (B.17) into Eq. (B.16) results in the following simplified expression for the phonons' grand canonical partition function:

$$\mathcal{Z}_{\text{ph}} = \prod_{\lambda=1}^{\infty} \frac{1}{1 - \exp \{-\Gamma (E_{\lambda} - \mu_{\text{ph}})\}} \quad (\text{B.18})$$

Similar to the case of electrons, the statistical average of Eq. (B.14) can be simplified, while using Eq. (B.15), into the following expressions [76]:

$$\begin{aligned} \langle \hat{m}_{\lambda} \rangle &= \frac{1}{\mathcal{Z}_{\text{ph}}} \sum_{m_1, m_2, \dots=0}^{\infty} \left(\exp \left\{ -\Gamma \sum_{\lambda=1}^{\infty} (E_{\lambda} - \mu_{\text{ph}}) m_{\lambda} \right\} \right) m_{\lambda} \\ &= \left[-\frac{1}{\Gamma} \frac{\partial}{\partial E_{\lambda}} (\ln (\mathcal{Z}_{\text{ph}})) \right] \end{aligned} \quad (\text{B.19})$$

which is equal to:

$$\langle \hat{m}_{\lambda} \rangle = \frac{\exp \{-\Gamma (E_{\lambda} - \mu_{\text{ph}})\}}{1 - \exp \{-\Gamma (E_{\lambda} - \mu_{\text{ph}})\}} \equiv f_{\text{BE}}(E_{\lambda} - \mu_{\text{ph}}) \quad (\text{B.20})$$

where f_{BE} is the expected Bose-Einstein probability distribution function for phonons, which is described in its final form by

$$f_{\text{BE}}(E_{\lambda} - \mu_{\text{ph}}) = \frac{1}{\exp \{\Gamma (E_{\lambda} - \mu_{\text{ph}})\} - 1} \quad (\text{B.21})$$

Note that the chemical potential μ_{ph} for the free phonons system remains close to zero until the temperature reaches the Bose-Einstein condensation temperature T_{BE} , at which the occupancy of the single-phonon ground state

becomes vanishingly small [108]. For the temperatures above T_{BE} , the chemical potential μ_{ph} for the free phonons system is negative [108]. In conclusion, the statistical average of a single-phonon state occupation number in a grand canonical ensemble results in Bose-Einstein statistics of the phonons.

In this appendix, we determined the mean occupation of electrons and phonons under equilibrium conditions, which exemplifies the quantum statistics in the semi-infinite contacts of a QTBM-based configuration. We use these statistics in formulating the PAT current equation in Section 3.2 of Chapter 3, while applying QTBM.

Bibliography

- [1] L. Hoddeson. *Chapter 9 - Innovation and basic research in the industrial laboratory: the repeater, transistor and Bell Telephone System*. North-Holland Delta Series. Elsevier, Amsterdam, 1990. ISSN 09275029. 181 - 214 pp.
- [2] W. F. Brinkman, D. E. Haggan, and W. W. Troutman. A history of the invention of the transistor and where it will lead us. *IEEE Journal of Solid-State Circuits*, 32(12):1858–1865, Dec 1997. ISSN 0018-9200.
- [3] Michael Riordan, Lillian Hoddeson, and Conyers Herring. The invention of the transistor. *Rev. Mod. Phys.*, 71:S336–S345, Mar 1999.
- [4] R. G. Arns. The other transistor: early history of the metal-oxide semiconductor field-effect transistor. *Engineering Science and Education Journal*, 7(5):233–240, Oct 1998. ISSN 0963-7346.
- [5] G. E. Moore. Cramming more components onto integrated circuits, reprinted from electronics, volume 38, number 8, april 19, 1965, pp.114 ff. *IEEE Solid-State Circuits Society Newsletter*, 11(3):33–35, Sept 2006. ISSN 1098-4232.
- [6] D. Foty. Perspectives on scaling theory and CMOS technology - understanding the past, present, and future. In *Proceedings of the 2004 11th IEEE International Conference on Electronics, Circuits and Systems, 2004. ICECS 2004.*, pages 631–637, Dec 2004.
- [7] 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*, 9(5):256–268, Oct 1974. ISSN 0018-9200.
- [8] Y. Tsvividis. *Chapter 2 - Innovation and basic research in the industrial laboratory: the repeater, transistor and Bell Telephone System*. WORLD SCIENTIFIC, 2011. 15 - 45 pp.

- [9] M. Bohr. A 30 Year Retrospective on Dennard's MOSFET Scaling Paper. *IEEE Solid-State Circuits Society Newsletter*, 12(1):11–13, Winter 2007. ISSN 1098-4232.
- [10] *Advanced CMOS Devices*. Taylor & Francis Group, 2014.
- [11] S. Banerjee, W. Richardson, J. Coleman, and A. Chatterjee. A new three-terminal tunnel device. *IEEE Electron Device Letters*, 8(8):347–349, Aug 1987. ISSN 0741-3106.
- [12] Leo Esaki. New Phenomenon in Narrow Germanium $p - n$ Junctions. *Phys. Rev.*, 109:603–604, Jan 1958.
- [13] William G. Vandenberghe, Bart Sorée, Wim Magnus, Guido Groeseneken, and Massimo V. Fischetti. Impact of field-induced quantum confinement in tunneling field-effect devices. *Applied Physics Letters*, 98(14):143503, 2011.
- [14] S. O. Koswatta, S. J. Koester, and W. Haensch. On the Possibility of Obtaining MOSFET-Like Performance and Sub-60-mV/dec Swing in 1-D Broken-Gap Tunnel Transistors. *IEEE Transactions on Electron Devices*, 57(12):3222–3230, Dec 2010. ISSN 0018-9383.
- [15] U. E. Avci, D. H. Morris, and I. A. Young. Tunnel Field-Effect Transistors: Prospects and Challenges. *IEEE Journal of the Electron Devices Society*, 3(3):88–95, May 2015. ISSN 2168-6734.
- [16] C. Hu, P. Patel, A. Bowonder, K. Jeon, S. H. Kim, Wei Yip Loh, Chang Yong Kang, Jungwoo Oh, P. Majhi, A. Javey, T. J. K. Liu, and R. Jammy. Prospect of tunneling green transistor for 0.1V CMOS. In *2010 International Electron Devices Meeting*, pages 16.1.1–16.1.4, Dec 2010. ISSN 0163-1918.
- [17] W. G. 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 *2008 International Conference on Simulation of Semiconductor Processes and Devices*, pages 137–140, Sept 2008. ISSN 1946-1569.
- [18] 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. *IEEE Transactions on Electron Devices*, 59(8):2070–2077, Aug 2012. ISSN 0018-9383.

- [19] D. Verreck, A. S. Verhulst, K. H. Kao, W. G. Vandenberghe, K. De Meyer, and G. Groeseneken. Quantum Mechanical Performance Predictions of p-n-i-n Versus Pocketed Line Tunnel Field-Effect Transistors. *IEEE Transactions on Electron Devices*, 60(7):2128–2134, July 2013. ISSN 0018-9383.
- [20] Dheeraj Mohata, Saurabh Mookerjee, Ashish Agrawal, Yuanyuan Li, Theresa Mayer, Vijaykrishnan Narayanan, Amy Liu, Dmitri Loubychev, Joel Fastenau, and Suman Datta. Experimental Staggered-Source and N+ Pocket-Doped Channel III-V Tunnel Field-Effect Transistors and Their Scalabilities. *Applied Physics Express*, 4(2):024105, 2011.
- [21] A. C. Seabaugh and Q. Zhang. Low-Voltage Tunnel Transistors for Beyond CMOS Logic. *Proceedings of the IEEE*, 98(12):2095–2110, Dec 2010. ISSN 0018-9219.
- [22] James Charles, Prasad Sarangapani, Roksana Golizadeh-Mojarad, Robert Andrawis, Daniel Lemus, Xinchun Guo, Daniel Mejia, James E. Fonseca, Michael Povolotskyi, Tillmann Kubis, and Gerhard Klimeck. Incoherent transport in nemo5: realistic and efficient scattering on phonons. *Journal of Computational Electronics*, 15(4):1123–1129, Dec 01, 2016. ISSN 1572-8137.
- [23] Elvedin Memisevic, Markus Hellenbrand, Erik Lind, Axel R. Persson, Saurabh Sant, Andreas Schenk, Johannes Svensson, Reine Wallenberg, and Lars-Erik Wernersson. Individual Defects in InAs/InGaAsSb/GaSb Nanowire Tunnel Field-Effect Transistors Operating below 60 mV/decade. *Nano Letters*, 17(7):4373–4380, 2017. PMID: 28613894.
- [24] A. S. Verhulst, D. Verreck, Q. Smets, K. H. Kao, M. Van de Put, R. Rooyackers, B. Sorée, A. Vandooren, K. De Meyer, G. Groeseneken, M. M. Heyns, A. Mocuta, N. Collaert, and A. V. Y. Thean. Perspective of tunnel-FET for future low-power technology nodes. In *Electron Devices Meeting (IEDM), 2014 IEEE International*, pages 30.2.1–30.2.4, Dec 2014.
- [25] A. Schenk, S. Sant, K. Moselund, H. Riel, E. Memisevic, and L. E. Wernersson. The impact of hetero-junction and oxide-interface traps on the performance of InAs/Si and InAs/GaAsSb nanowire tunnel FETs. In *2017 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pages 273–276, Sept 2017. ISSN 1946-1569.
- [26] M. G. Pala, D. Esseni, and F. Conzatti. Impact of interface traps on the IV curves of InAs Tunnel-FETs and MOSFETs: A full quantum study. In *2012 International Electron Devices Meeting*, pages 6.6.1–6.6.4, Dec 2012. ISSN 0163-1918.

- [27] J. Bizindavyi, A. S. Verhulst, Q. Smets, D. Verreck, N. Collaert, A. Mocuta, B. Sorée, and G. Groeseneken. Calibration of the high-doping induced ballistic band-tails tunneling current with In_{0.53}Ga_{0.47}As Esaki diodes. In *2017 Fifth Berkeley Symposium on Energy Efficient Electronic Systems Steep Transistors Workshop (E3S)*, pages 1–3, Oct 2017.
- [28] James T. Teherani, Sapan Agarwal, Winston Chern, Paul M. Solomon, Eli Yablonovitch, and Dimitri A. Antoniadis. Auger generation as an intrinsic limit to tunneling field-effect transistor performance. *Journal of Applied Physics*, 120(8):084507, 2016.
- [29] H. Lu and A. Seabaugh. Tunnel Field-Effect Transistors: State-of-the-Art. *IEEE Journal of the Electron Devices Society*, 2(4):44–49, July 2014. ISSN 2168-6734.
- [30] S. Mookerjee, D. Mohata, T. Mayer, V. Narayanan, and S. Datta. Temperature-Dependent $I - V$ Characteristics of a Vertical In_{0.53}Ga_{0.47}As Tunnel FET. *IEEE Electron Device Letters*, 31(6):564–566, June 2010. ISSN 0741-3106.
- [31] C. T. Sah. Electronic Processes and Excess Currents in Gold-Doped Narrow Silicon Junctions. *Phys. Rev.*, 123:1594–1612, Sep 1961.
- [32] A. G. Chynoweth, W. L. Feldmann, and R. A. Logan. Excess Tunnel Current in Silicon Esaki Junctions. *Phys. Rev.*, 121:684–694, Feb 1961.
- [33] Evan O. Kane. Theory of Tunneling. *Journal of Applied Physics*, 32(1): 83–91, 1961.
- [34] Horng-Sen Fu and Chih-Tang Sah. Theory and experiments on surface $1/f$ noise. *IEEE Transactions on Electron Devices*, 19(2):273–285, Feb 1972. ISSN 0018-9383.
- [35] J. Y. Wong. Effect of trap tunneling on the performance of long-wavelength Hg_{1-x}Cd_xTe photodiodes. *IEEE Transactions on Electron Devices*, 27(1): 48–57, Jan 1980. ISSN 0018-9383.
- [36] W. W. Anderson and H. J. Hoffman. Field ionization of deep levels in semiconductors with applications to Hg_{1-x}Cd_xTe p-n junctions. *Journal of Applied Physics*, 53(12):9130–9145, 1982.
- [37] Yuki Mori, Kenichi Takeda, and Ren ichi Yamada. Random telegraph noise of junction leakage current in submicron devices. *Journal of Applied Physics*, 107(1):014509, 2010.

- [38] J. Wu, L. F. Register, and E. Rosenbaum. Trap-assisted tunneling current through ultra-thin oxide. In *1999 IEEE International Reliability Physics Symposium Proceedings. 37th Annual (Cat. No.99CH36296)*, pages 389–395, 1999.
- [39] M.J. Kirton and M.J. Uren. Noise in solid-state microstructures: A new perspective on individual defects, interface states and low-frequency (1/f) noise. *Advances in Physics*, 38(4):367–468, 1989.
- [40] G. A. M. Hurkx, D. B. M. Klaassen, M. P. G. Knuvers, and F. G. O’Hara. A new recombination model describing heavy-doping effects and low-temperature behaviour. In *International Technical Digest on Electron Devices Meeting*, pages 307–310, Dec 1989. ISSN 0163-1918.
- [41] A. Schenk. A model for the field and temperature dependence of Shockley-Read-Hall lifetimes in silicon. *Solid-State Electronics*, 35(11):1585 – 1596, 1992. ISSN 0038-1101.
- [42] Gös Wolfgang. *Hole Trapping and the Negative Bias Temperature Instability*. PhD thesis, Fakultät für Elektrotechnik und Informationstechnik, TU Wien, December 2011. 173 + xxii pp. Tibor, Grasser (supervisor).
- [43] D. V. Lang and C. H. Henry. Nonradiative recombination at deep levels in GaAs and GaP by lattice-relaxation multiphonon emission. *Phys. Rev. Lett.*, 35:1525–1528, Dec 1975.
- [44] W. Shockley and W. T. Read. Statistics of the Recombinations of Holes and Electrons. *Phys. Rev.*, 87:835–842, Sep 1952.
- [45] S. Steiger, M. Povolotskyi, H. H. Park, T. Kubis, and G. Klimeck. NEMO5: A Parallel Multiscale Nanoelectronics Modeling Tool. *IEEE Transactions on Nanotechnology*, 10(6):1464–1474, Nov 2011. ISSN 1536-125X.
- [46] G. A. M. Hurkx, D. B. M. Klaassen, and M. P. G. Knuvers. A new recombination model for device simulation including tunneling. *IEEE Transactions on Electron Devices*, 39(2):331–338, Feb 1992. ISSN 0018-9383.
- [47] G. Vincent, A. Chantre, and D. Bois. Electric field effect on the thermal emission of traps in semiconductor junctions. *Journal of Applied Physics*, 50(8):5484–5487, 1979.
- [48] Frederiksen Thomas. Inelastic electron transport in nanosystems, February 2004. Mads, Brandbyge (supervisor).
- [49] Melvin Lax. Giant traps. *Journal of Physics and Chemistry of Solids*, 8: 66 – 73, 1959. ISSN 0022-3697.

- [50] D. V. Lang and C. H. Henry. Nonradiative Recombination at Deep Levels in GaAs and GaP by Lattice-Relaxation Multiphonon Emission. *Phys. Rev. Lett.*, 35:1525–1528, Dec 1975.
- [51] Huang Rhys. Theory of light absorption and non-radiative transitions in F-centres. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 204(1078):406–423, 1950. ISSN 0080-4630.
- [52] A. Schenk, M. Stahl, and H.-J. Wünsche. Calculation of Interband Tunneling in Inhomogeneous Fields. *physica status solidi (b)*, 154(2): 815–826, 1989. ISSN 1521-3951.
- [53] A. Schenk, R. Enderlein, and D. Suisky. Electroabsorption for Deep Level to Band Transitions Accompanied by Multiphonon Processes. *physica status solidi (b)*, 131(2):729–739, 1985. ISSN 1521-3951.
- [54] G. Helms. Zur Theorie der Störstellenelektronen. I Optische Übergänge. *Annalen der Physik*, 452(6-8):356–370, 1956. ISSN 1521-3889.
- [55] G. Helms. Zur Theorie der Störstellenelektronen. II Strahlungslose Übergänge. *Annalen der Physik*, 454(1-2):41–54, 1956. ISSN 1521-3889.
- [56] A. Schenk. An improved approach to the Shockley-Read-Hall recombination in inhomogeneous fields of space-charge regions. *Journal of Applied Physics*, 71(7):3339–3349, 1992.
- [57] S. Sant, K. Moselund, D. Cutaia, H. Schmid, M. Borg, H. Riel, and A. Schenk. Lateral InAs/Si p-type tunnel FETs integrated on si; part 2: Simulation study of the impact of interface traps. *IEEE Transactions on Electron Devices*, 63(11):4240–4247, Nov 2016. ISSN 0018-9383.
- [58] L. Vandelli, A. Padovani, L. Larcher, R. G. Southwick, W. B. Knowlton, and G. Bersuker. A physical model of the temperature dependence of the current through SiO₂/HfO₂ stacks. *IEEE Transactions on Electron Devices*, 58(9):2878–2887, Sept 2011. ISSN 0018-9383.
- [59] Anthony I. Chou, Kafai Lai, Kiran Kumar, Prasenjit Chowdhury, and Jack C. Lee. Modeling of stress-induced leakage current in ultrathin oxides with the trap-assisted tunneling mechanism. *Applied Physics Letters*, 70 (25):3407–3409, 1997.
- [60] J. Bardeen. Tunnelling from a Many-Particle Point of View. *Phys. Rev. Lett.*, 6:57–59, Jan 1961.
- [61] Evan O. Kane. Theory of Tunneling. *Journal of Applied Physics*, 32(1): 83–91, 1961.

- [62] Tewksbury III Theodore L. *Relaxation Effects in MOS Devices due to Tunnel Exchange with Near-Interface Oxide Traps*. PhD thesis, Electrical Engineering and Computer Science, MIT, 1992. 317 + xvii pp. Hae-Seung, Lee (supervisor).
- [63] Melvin Lax. The Franck Condon Principle and Its Application to Crystals. *The Journal of Chemical Physics*, 20(11):1752–1760, 1952.
- [64] Wolfgang Nolting. *Approximation Methods*. Springer International Publishing, 2017. ISBN 978-3-319-63324-4. 190-199 pp.
- [65] S. A. Rakityansky. Unified treatment of bound, scattering, and resonant states in one-dimensional semiconductor nanostructures. *Phys. Rev. B*, 68:195320, Nov 2003.
- [66] Chenjing L. Fernando and William R. Frensley. An efficient method for the numerical evaluation of resonant states. *Journal of Applied Physics*, 76(5):2881–2886, 1994.
- [67] William R. Frensley. Numerical evaluation of resonant states. *Superlattices and Microstructures*, 11(3):347 – 350, 1992. ISSN 0749-6036.
- [68] E. Anemogiannis, E. N. Glytsis, and T. K. Gaylord. Quasi-bound states determination using a perturbed wavenumbers method in a large quantum box. *IEEE Journal of Quantum Electronics*, 33(5):742–752, May 1997. ISSN 0018-9197.
- [69] S. A. Rakityansky. Modified transfer matrix for nanostructures with arbitrary potential profile. *Phys. Rev. B*, 70:205323, Nov 2004.
- [70] Luis L. Sánchez-Soto, Juan J. Monzón, Alberto G. Barriuso, and José F. Cariñena. The transfer matrix: A geometrical perspective. *Physics Reports*, 513(4):191 – 227, 2012. ISSN 0370-1573. The transfer matrix: A geometrical perspective.
- [71] J. Mathews and R. L. Walker. *Chapter 12 - Calculus of Variations*. The Advanced Book Program. Addison-Wesley Publishing Company, Inc, 1970. 322 - 341 pp.
- [72] M. G. Pala, D. Esseni, and F. Conzatti. Impact of interface traps on the IV curves of InAs tunnel-FETs and MOSFETs: A full quantum study. In *Electron Devices Meeting (IEDM), 2012 IEEE International*, pages 6.6.1–6.6.4, Dec 2012. ISSN 0163-1918.
- [73] Tibor Grassler. Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities. *Microelectronics*

- Reliability*, 52(1):39 – 70, 2012. ISSN 0026-2714. 2011 Reliability of Compound Semiconductors (ROCS) Workshop.
- [74] Nathan L. Anderson, Ravi Pramod Vedula, Peter A. Schultz, R. M. Van Ginhoven, and Alejandro Strachan. Defect level distributions and atomic relaxations induced by charge trapping in amorphous silica. *Applied Physics Letters*, 100(17):172908, 2012.
- [75] A. Schenk. A model for the field and temperature dependence of Shockley-Read-Hall lifetimes in silicon. *Solid-State Electronics*, 35(11):1585 – 1596, 1992. ISSN 0038-1101.
- [76] Richard P. Feynmann. *Statistical Mechanics: A Set of Lectures*. Frontiers in Physics. CRC-Press, 1998.
- [77] William Vandenberghe, Bart Sorée, Wim Magnus, and Massimo 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, 2011.
- [78] Gerald D. Mahan. *Many-Particle Physics*. Physics of Solids and Liquids. Springer US, 1998.
- [79] Maarten. L. Van de Put. *Modeling of quantum electron transport with applications in energy filtering nanostructures*. PhD thesis, Universiteit Antwerpen, 2016. pages 115–119.
- [80] Craig S. Lent and David J. Kirkner. The quantum transmitting boundary method. *Journal of Applied Physics*, 67(10):6353–6359, 1990.
- [81] Devin Verreck, Anne S. Verhulst, Maarten Van de Put, Bart Sorée, Wim Magnus, Anda Mocuta, Nadine Collaert, Aaron Thean, and Guido Groeseneken. Full-zone spectral envelope function formalism for the optimization of line and point tunnel field-effect transistors. *Journal of Applied Physics*, 118(13):134502, 2015.
- [82] *Basic Principles*, pages 1–23. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. ISBN 978-3-540-36217-3.
- [83] U. Fano. Note on the quantum theory of irreversible processes. *Phys. Rev.*, 96:869–873, Nov 1954.
- [84] G. C. Wick. The evaluation of the collision matrix. *Phys. Rev.*, 80:268–272, Oct 1950.

- [85] Cristian Rivas, Roger Lake, Gerhard Klimeck, William R. Frensley, Massimo V. Fischetti, Phillip E. Thompson, Sean L. Rommel, and Paul R. Berger. Full-band simulation of indirect phonon assisted tunneling in a silicon tunnel diode with delta-doped contacts. *Applied Physics Letters*, 78(6):814–816, 2001.
- [86] P B Allen and V Heine. Theory of the temperature dependence of electronic band structures. *Journal of Physics C: Solid State Physics*, 9(12):2305, 1976.
- [87] William G. Vandenberghe and Massimo V. Fischetti. Deformation potentials for band-to-band tunneling in silicon and germanium from first principles. *Applied Physics Letters*, 106(1):013505, 2015.
- [88] H. Fröhlich. Electrons in lattice fields. *Advances in Physics*, 3(11):325–361, 1954.
- [89] M G Burt. An exact formulation of the envelope function method for the determination of electronic states in semiconductor microstructures. *Semiconductor Science and Technology*, 3(8):739, 1988.
- [90] Devin Verreck. *Quantum mechanical transport towards the optimization of heterostructure tunnel field-effect transistors*. PhD thesis, KU Leuven, 2017.
- [91] Maarten L. Van de Put, William G. Vandenberghe, Wim Magnus, and Bart Sorée. An envelope function formalism for lattice-matched heterostructures. *Physica B: Condensed Matter*, 470-471:69 – 75, 2015. ISSN 0921-4526.
- [92] Maarten L. Van de Put. Modeling of band-to-band tunneling in heterostructure semiconductors, 2011-2012. Master Thesis, KU Leuven, pages 38–41.
- [93] Te-Huan Liu, Jiawei Zhou, Bolin Liao, David J. Singh, and Gang Chen. First-principles mode-by-mode analysis for electron-phonon scattering channels and mean free path spectra in gaas. *Phys. Rev. B*, 95:075206, Feb 2017.
- [94] A. S. Verhulst, D. Verreck, M. A. Pourghaderi, M. Van de Put, B. Sorée, G. Groeseneken, N. Collaert, and A. V.-Y. Thean. Can p-channel tunnel field-effect transistors perform as good as n-channel? *Applied Physics Letters*, 105(4):043103, 2014.
- [95] Mazharuddin Mohammed, Anne S. Verhulst, and Maarten L. Van de Put. A report on the phase shift between the envelope functions due to band parity. *imec internal reports*, pages 1–15, Apr 2018.

- [96] H. Fröhlich and F. Mott. The mean free path of electrons in polar crystals. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 171(947):496–504, 1939. ISSN 0080-4630.
- [97] A Afzalian, T Vasen, P Ramvall, T-M Shen, J Wu, and M Passlack. Physics and performances of iii–v nanowire broken-gap heterojunction tfets using an efficient tight-binding mode-space negf model enabling million-atom nanowire simulations. *Journal of Physics: Condensed Matter*, 30(25):254002, 2018.
- [98] Tibor Grasser. Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities. *Microelectronics Reliability*, 52(1):39 – 70, 2012. ISSN 0026-2714. 2011 Reliability of Compound Semiconductors (ROCS) Workshop.
- [99] R. N. Hall, J. H. Racette, and H. Ehrenreich. Direct observation of polarons and phonons during tunneling in group 3-5 semiconductor junctions. *Phys. Rev. Lett.*, 4:456–458, May 1960.
- [100] B. C. Cavenett. Electron-phonon interactions in InSb junctions. *Phys. Rev. B*, 5:3049–3057, Apr 1972.
- [101] Feliciano Giustino. Electron-phonon interactions from first principles. *Rev. Mod. Phys.*, 89:015003, Feb 2017.
- [102] P. B. Allen. Solids with thermal or static disorder. i. one-electron properties. *Phys. Rev. B*, 18:5217–5224, Nov 1978.
- [103] P. B. Allen and M. Cardona. Theory of the temperature dependence of the direct gap of germanium. *Phys. Rev. B*, 23:1495–1505, Feb 1981.
- [104] Jun Kondo. Resistance minimum in dilute magnetic alloys. *Progress of Theoretical Physics*, 32(1):37–49, 1964.
- [105] P. W. Anderson. Localized magnetic states in metals. *Phys. Rev.*, 124:41–53, Oct 1961.
- [106] Brian T. Sutcliffe and R. Guy Woolley. On the quantum theory of molecules. *The Journal of Chemical Physics*, 137(22):22A544, 2012.
- [107] R. Pässler. Description of nonradiative multiphonon transitions in the static coupling scheme II. Approximations. *Czechoslovak Journal of Physics B*, 25(2):219–234, Feb 01, 1975. ISSN 1572-9486.
- [108] G. Cook and R. H. Dickerson. Understanding the chemical potential. *American Journal of Physics*, 63(8):737–742, 1995.

Curriculum Vitae

Mazharuddin Mohammed was born in India in Dec-1987. He received the B.E. degree in Electronics and Communication Engineering from the Osmania University (India) in 2009 and the MSc degree in Micro and Nano Systems, from the Technische Universität Chemnitz (Germany) in 2013. Since June-2014, he has been pursuing a PhD degree from the Katholieke Universiteit Leuven, doing his research at imec (Belgium) while being connected to the Department of Electrical Engineering, under the supervision of Dr. Anne S. Verhulst and Prof. Guido Groeseneken.



In 2013, he was a visiting researcher with Advanced Technologies and Standards department of Siemens AG in Nürnberg (Germany), under the supervision of Dr. Dirk Scheibner and Prof. Jan Mehner. In 2012, he was a visiting researcher with Micro Materials Center of Fraunhofer ENAS in Chemnitz (Germany), under the supervision of Dr. Dudek Rainer and Prof. Bernhard Wunderle. In 2012, he was a research assistant with Mess- und Sensortechnik department of TU Chemnitz (Germany), under the supervision of Dr. Uwe Tröltzsch and Prof. Olfa Kanoun. His research interests include the quantum mechanical modeling of parasitic effects in semiconductor devices.

List of publications

Journals:

- [1] **Mazharuddin Mohammed**, Anne S. Verhulst, Devin Verreck, Maarten Van de Put, Eddy Simoen, Bart Sorée, Ben Kaczer, Robin Degraeve, Anda Mocuta, Nadine Collaert, Aaron Thean, and Guido Groeseneken. “Electric-field induced quantum broadening of the characteristic energy level of traps in semiconductors and oxides”. *Journal of Applied Physics*, 120(24):245704, 2016.
- [2] **Mazharuddin Mohammed**, Anne S. Verhulst, Devin Verreck, Maarten Van de Put, Wim Magnus, Bart Sorée, and Guido Groeseneken. “Phonon-assisted tunneling in direct-bandgap semiconductors.” Submitted to *Journal of Applied Physics*.

Conferences:

- [1] **Mazharuddin Mohammed**, Anne S. Verhulst, Devin Verreck, Maarten L. Van de Put, Wim magnus, Bart Sorée, Anda Mocuta, and Guido Groeseneken. “Phonon-Assisted Tunneling in Large-Size Direct-Bandgap Semiconductor Devices.” In *APS March Meeting* 2018, March 2018.
- [2] Anne S. Verhulst, Devin Verreck, William G. Vandenberghe, Quentin Smets, **Mazharuddin Mohammed**, Jasper Bizindavyi, Marc. M. Heyns, Bart Sorée, Nadine Collaert, and Anda Mocuta. “Inherent transmission probability limit between valence-band and conduction-band states and calibration of tunnel-fet parasitics.” In *2017 Fifth Berkeley Symposium on Energy Efficient Electronic Systems Steep Transistors Workshop (E3S)*, pages 1–3, Oct 2017.

FACULTY OF ENGINEERING SCIENCE
DEPARTMENT OF ELECTRICAL ENGINEERING
ESAT-MICAS

Kasteelpark Arenberg 10 box 2440
B-3001 Leuven

mazharuddin.mohammed@student.kuleuven.be

<https://www.esat.kuleuven.be/micas/>

