DESIGN GUIDELINES FOR HIGH EFFICIENCY PHOTOVOLTAICS AND
LOW POWER TRANSISTORS USING QUANTUM TRANSPORT

A Dissertation
Submitted to the Faculty
of
Purdue University
by
Samarth Agarwal

In Partial Fulfillment of the
Requirements for the Degree
of
Doctor of Philosophy

December 2010
Purdue University
West Lafayette, Indiana
To my parents.
ACKNOWLEDGMENTS

I wish to acknowledge the efforts of Kyle Montgomery for the work in Chapter 4. I would also like to thank Prof. Jerry Woodall and Prof. Timothy Boykin whose guidance made this work possible.

I am intellectually indebted to Dr. Mathieu Luisier whose research has often been the starting point for my work. He has provided deep technical insights on various topics and made the work in Chapter 5 possible. I would also like to thank him for the support, in the form of different quantum transport simulators, that has allowed me to contribute to a number of projects.

I have benefited immensely from my association with Dr. Michael Povolotskyi and Dr. Tillmann Kubis who have helped me through the work in Chapter 6. I have been fortunate to have been associated with Dr. Neerav Kharche, Dr. Rajib Rahman, Abhijeet Paul, Dr. Neophytos Neophytou, Kaushik Balamukundan and Tuna Toksoz. I wish to thank all students at the Network for Computational Nanotechnology for valuable scientific discussions.

I want to thank Prof. Dragica Vasileska, Prof. Norbert Neumeister, Prof. Andrew Hirsch, Prof. Mark Haugan and Prof. Nicholas Giordano who have helped me at different points of my graduate studies. I feel the need to mention the efforts of Cheryl Haines and Vicki Johnson from the Network for Computational Nanotechnology and Sandy Formica from the Physics office. They have handled all the paperwork for me at various stages which allowed me to remain focused on my research.

I am thankful to Prof. Supriyo Datta and Prof. Sergei Savikhin for being on my advisory committee. I have also benefitted from their advice during coursework at Purdue.
I wish to thank Prof. Ronald Reifenberger who has supported me throughout my graduate studies. His advice has helped me whenever I have struggled intellectually or otherwise during my doctoral work.

Finally, I want to express my gratitude to Prof. Gerhard Klimeck who has given me the support, the resources and the expertise to conduct research. I am indebted to him not only for the technical insights but also for his tireless efforts at trying to improve my professional skills.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>vii</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xiii</td>
</tr>
<tr>
<td>1  INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Efficiencies in photovoltaics</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Power consumption in Transistor technology</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Need for Atomistic Models and the tight binding approach</td>
<td>4</td>
</tr>
<tr>
<td>1.4 Associated Numerical techniques</td>
<td>6</td>
</tr>
<tr>
<td>1.5 Brief Outline</td>
<td>7</td>
</tr>
<tr>
<td>2  NUMERICAL METHODS AND ASSOCIATED TOPICS</td>
<td>9</td>
</tr>
<tr>
<td>2.1 Recursive Green’s Function: Algorithm and implementation</td>
<td>9</td>
</tr>
<tr>
<td>2.2 Resonance finding using the Shift and invert non-symmetric Lanczos algorithm and Newton’s method: Idea and implementation</td>
<td>11</td>
</tr>
<tr>
<td>2.3 Resonance refinement using Inverse Iteration: Algorithm and implementation</td>
<td>14</td>
</tr>
<tr>
<td>2.4 In-homogeneous Energy grid creation through analysis of spectral features: Grid mappings</td>
<td>16</td>
</tr>
<tr>
<td>2.5 Adaptive Energy grid creation based on charge, current and density of states</td>
<td>21</td>
</tr>
<tr>
<td>2.6 Boundary Conditions for multi-band tight-binding</td>
<td>22</td>
</tr>
<tr>
<td>2.6.1 Hard-Wall, Passivation of Surface States</td>
<td>22</td>
</tr>
<tr>
<td>2.6.2 Open Boundary Conditions</td>
<td>25</td>
</tr>
<tr>
<td>2.7 Resonance finder and RGF results for resonant inter-band tunneling diodes using open boundary conditions</td>
<td>25</td>
</tr>
<tr>
<td>2.7.1 Interest in resonant inter-band tunneling diodes</td>
<td>26</td>
</tr>
<tr>
<td>2.7.2 Non-intuitive behavior of hole bands in GaSb</td>
<td>26</td>
</tr>
<tr>
<td>2.7.3 Analysis of GaSb based BRIT structure using transmission</td>
<td>28</td>
</tr>
<tr>
<td>2.7.4 Analysis of GaSb based RIT structure using transmission</td>
<td>31</td>
</tr>
<tr>
<td>2.7.5 Analysis of GaSb based BRIT and RIT structures using resonant line-widths</td>
<td>31</td>
</tr>
<tr>
<td>3  1D HETEROSTRUCTURE TOOL</td>
<td>35</td>
</tr>
<tr>
<td>3.1 Methodology</td>
<td>35</td>
</tr>
<tr>
<td>3.2 Tool Usage and Capabilities</td>
<td>37</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Device detail of the default structure: Heterojunction FET</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Formation of the well and comparison with analytical solutions</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Charge profile and the Fermi-level</td>
</tr>
<tr>
<td>3.2.4</td>
<td>Effect of delta doped layers</td>
</tr>
<tr>
<td>3.2.5</td>
<td>Comparison with semi-classical charge</td>
</tr>
<tr>
<td>3.2.6</td>
<td>Effect of the Fermi level on charge</td>
</tr>
<tr>
<td>3.3</td>
<td>Tool limitations</td>
</tr>
<tr>
<td>3.4</td>
<td>Algorithmic details</td>
</tr>
<tr>
<td>4</td>
<td>DESIGN GUIDELINES FOR TRUE GREEN LEDS AND HIGH EFFICIENCY PHOTOVOLTAICS USING ZnSe/GaAs DIGITAL ALLOYS</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction and approach</td>
</tr>
<tr>
<td>4.2</td>
<td>Methodology</td>
</tr>
<tr>
<td>4.3</td>
<td>Results and Discussion</td>
</tr>
<tr>
<td>4.4</td>
<td>Conclusion</td>
</tr>
<tr>
<td>4.5</td>
<td>Supplementary topic: Parameter fitting</td>
</tr>
<tr>
<td>5</td>
<td>LEAKAGE REDUCTION DESIGN CONCEPTS FOR LOW POWER VERTICAL TUNNELING FIELD-EFFECT TRANSISTORS</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>5.2</td>
<td>Simulation Approach and Results</td>
</tr>
<tr>
<td>5.3</td>
<td>Conclusion</td>
</tr>
<tr>
<td>5.4</td>
<td>Supplementary topic: Active energy range in the gFET and the modified design</td>
</tr>
<tr>
<td>6</td>
<td>ADAPTIVE QUADRATURE FOR SHARPLY SPIKED INTEGRANDS</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>6.2</td>
<td>Methodology: 5 point adaptive scheme</td>
</tr>
<tr>
<td>6.3</td>
<td>Results and Discussion: Accuracy Vs Cost</td>
</tr>
<tr>
<td>6.4</td>
<td>Conclusion</td>
</tr>
<tr>
<td>6.5</td>
<td>Supplementary topic: Adding 4 Vs 1</td>
</tr>
<tr>
<td>6.5.1</td>
<td>Analysis for a broad resonance</td>
</tr>
<tr>
<td>6.5.2</td>
<td>Analysis for a sharp resonance</td>
</tr>
<tr>
<td>6.5.3</td>
<td>Remarks: 4pts Vs 1pt addition</td>
</tr>
<tr>
<td>7</td>
<td>SUMMARY</td>
</tr>
<tr>
<td></td>
<td>LIST OF REFERENCES</td>
</tr>
<tr>
<td></td>
<td>VITA</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Bandgaps and lattice constants of materials used in high efficiency photovoltaics.</td>
<td>2</td>
</tr>
<tr>
<td>3.1 Layer table for the default structure.</td>
<td>39</td>
</tr>
<tr>
<td>4.1 Nearest neighbor $sp^3s^*$ ZnSe parameters(eV). Labels denote matrix elements. $E$:onsite, $V$:coupling, $\lambda$:spin-orbit coupling. Subscripts denote the basis. $s,p,x,y$: orbital symmetries. $a,c$: anion or cation [48]</td>
<td>58</td>
</tr>
<tr>
<td>4.2 Experimental data for ZnSe and the achieved values using the re-parametrized nearest neighbor tight-binding parameters.</td>
<td>63</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>1.1</td>
<td>Supply voltage ($V_{DD}$) of transistors as a function of technology node size [2].</td>
</tr>
<tr>
<td>1.2</td>
<td>The localized conduction band edge for a GaAs/AlAs/GaAs structure. Along the Γ direction the band edge presents a barrier while in the X direction it looks like a well. The bulk band structures have been generated using tight-binding. This figure is based on the discussion in [7].</td>
</tr>
<tr>
<td>1.3</td>
<td>Difference at higher energies between the parabolic bulk conduction band of GaAs and a non-parabolic band generated using tight-binding. This figure is based on the discussion in [7].</td>
</tr>
<tr>
<td>2.1</td>
<td>Schematic representation of forward and backward recursion in the Recursive Green’s Function algorithm and the subsequent calculation of the first and last column. The diagonal blocks of $G^R$ give the local density of states. The first and last columns are required to compute the charge.</td>
</tr>
<tr>
<td>2.2</td>
<td>Local Density of states for a double barrier GaAs/AlGaAs/GaAs structure with transmission using a single band effective mass scheme [21]. The RGF algorithm is used to compute the density of states and the transmission.</td>
</tr>
<tr>
<td>2.3</td>
<td>Local Density of states for a double barrier GaAs/AlGaAs/GaAs structure with resonances as found by the resonance finder using a single band effective mass scheme [21]. The resonances are compared with the analytical solutions of a particle in a box.</td>
</tr>
<tr>
<td>2.4</td>
<td>Local Density of states for a double barrier GaAs/AlGaAs/GaAs structure with wave-function magnitude squared found using inverse iteration in a single band effective mass scheme [21]. The variation of the wavefunction matches well with the local density of states.</td>
</tr>
<tr>
<td>2.5</td>
<td>Contribution to spectral features: Lorentzian.</td>
</tr>
<tr>
<td>2.6</td>
<td>Contribution to spectral features: Principal value Integral.</td>
</tr>
<tr>
<td>2.7</td>
<td>Contribution to spectral features: Peak Enhancement.</td>
</tr>
<tr>
<td>2.8</td>
<td>Contribution to spectral features: Density of States.</td>
</tr>
<tr>
<td>2.9</td>
<td>Contribution to spectral features: Fermi function.</td>
</tr>
<tr>
<td>2.10</td>
<td>Contribution to spectral features: Fermi function turn-ons.</td>
</tr>
</tbody>
</table>
Figure | Page
--- | ---
2.11 Local Density of states for a double barrier GaAs/AlGaAs/GaAs structure calculated on an energy grid generated using the grid mapping scheme using a single band effective mass scheme [21]. | 20
2.12 Local Density of states for a double barrier GaAs/AlGaAs/GaAs structure demonstrating the adaptive energy grid using a single band effective mass scheme [21]. | 21
2.13 Bulk HH dispersion for GaSb in the $k_x - k_y$ plane | 27
2.14 Bulk LH dispersion for GaSb in the $k_x - k_y$ plane | 27
2.15 Schematic band profile of a BRIT structure. | 29
2.16 Transmission and E-k plots for the BRIT structure. | 29
2.17 Schematic band profile of a RIT structure. | 30
2.18 Transmission and E-k plots for the RIT structure. | 30
2.19 Resonant line-width for the BRIT structure. | 32
2.20 Resonant line-width for the RIT structure. | 32
3.1 The band profile of the default H-FET structure without any electrostatics. | 38
3.2 The doping profile of the default H-FET structure. | 38
3.3 The band profile of the default structure with electrostatics determined by a charge self-consistent calculation. | 40
3.4 The quantum mechanical charge profile of the default structure with electrostatics determined by a charge self-consistent calculation. | 40
3.5 Analytically computed wave-function magnitude squared of the ground state of a GaAs triangular potential well with a $10^6 V/m$ electric field. | 41
3.6 Analytically computed wave-function magnitude squared of the first excited state of a GaAs triangular potential well with a $10^6 V/m$ electric field. | 42
3.7 The wave-function magnitude squared of the eigen-energies found in the triangular well of the default structure. | 42
3.8 The band profile of the default structure modified to include a delta doped layer (Density=$10^{12}/cm^2$) in the center of the AlGaAs layer with electrostatics determined by a charge self-consistent calculation. | 44
3.9 The lowest eigen-energy as a function of bias for the default structure and the modified structure with the delta doping. | 45
<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.10</td>
<td>The semi-classical charge profile of the default structure with electrostatics determined by a charge self-consistent calculation.</td>
</tr>
<tr>
<td>46</td>
<td></td>
</tr>
<tr>
<td>3.11</td>
<td>The sheet density in the well region for the default structure as a function of bias for different Fermi levels.</td>
</tr>
<tr>
<td>48</td>
<td></td>
</tr>
<tr>
<td>3.12</td>
<td>The band profile of the default structure with a very high doping 10^{19}/cc in the AlGaAs region. Though the tool shows the band profile as converged it is physically incorrect.</td>
</tr>
<tr>
<td>50</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>AM 1.5 solar spectrum with the shaded region indicating photons absorbed by a 2.4eV solar cell (constituting 21.1% of total irradiance 7.5mA/cm² total current density).</td>
</tr>
<tr>
<td>55</td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>General concepts for incorporating the ZnSe/GaAs SL</td>
</tr>
<tr>
<td>57</td>
<td></td>
</tr>
<tr>
<td>4.3</td>
<td>Band diagram of the ZnSe/GaAs SL: The band offsets have been indicated with arrows. The dashed band edge indicates that the structure is periodic. The dotted lines indicate confined states that penetrate into the barrier region.</td>
</tr>
<tr>
<td>59</td>
<td></td>
</tr>
<tr>
<td>4.4</td>
<td>Ground state energy in eV of the states formed in the conduction band well for different monolayers of ZnSe and GaAs.</td>
</tr>
<tr>
<td>60</td>
<td></td>
</tr>
<tr>
<td>4.5</td>
<td>Difference in the ground state energy in the conduction band well and the highest state in the valence band barrier in eV for different monolayers of ZnSe and GaAs.</td>
</tr>
<tr>
<td>61</td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>(a) Lateral p-i-n single-gate TFET. (b) Conduction and valence band edges of the p-i-n device in its OFF- (dashed lines) and ON- (solid lines) state along the horizontal dashed line in (a). (c) Single-gate gFET structure as proposed in Ref. [86]. (d) Conduction and valence band edges of the gFET in its OFF- and ON- state along the vertical dashed line in (c). The buried oxide extends from -50 nm to 0 nm, but only the 20 nm close to the semiconductor channel are shown here.</td>
</tr>
<tr>
<td>68</td>
<td></td>
</tr>
<tr>
<td>5.2</td>
<td>(a) Transfer characteristics ( I_d-V_{gs} ) at ( V_{ds}=0.5 ) V of a lateral p-i-n TFET (solid line) and a vertical gFET (line with symbols) with ( L_g=40 ) nm and EOT=0.5 nm. (b) Spatial distribution of the gFET ON-current. High current density regions appear darker. (c) Same as (b), but for the OFF-current. For clarity, a different color scale is used. (d) CB and VB edges of the gFET OFF-state on the horizontal line below the ( n^+ ) pocket at ( y=4 ) nm in (c).</td>
</tr>
<tr>
<td>69</td>
<td></td>
</tr>
</tbody>
</table>
(a) Modified gFET design. The buried oxide is raised to the height of the pocket in the drain and in the lightly doped channel region. (b) Comparison of the transfer characteristics $I_d-V_{gs}$ at $V_{ds}=0.5$ V for the "conventional" (line with symbols) and the modified gFET (solid line) design.

Transfer characteristics $I_d-V_{gs}$ of different gFET designs. (a) Variation of the pocket length $L_{pocket}$ at a constant gate length $L_g=40$ nm. (b) Variation of the pocket doping $N_{pock}$. (c) Variation of the body thickness $T_{body}$. (d) Comparison of the optimized InAs gFET (line with circles) with the original gFET design (solid line) and the lateral $p-i-n$ TFET (dashed line).

Transmission Vs. Energy for the gFET and the modified design in its OFF state corresponding to $V_{GS}=-0.1$V in Fig. 5.3(b). The dotted lines denote the quantity $T(f_L-f_R)$ for the corresponding device.

Transmission Vs. Energy for the gFET and the modified design in its ON state corresponding to $V_{GS}=0.5$V in Fig. 5.3(b). The dotted lines denote the quantity $T(f_L-f_R)$ for the corresponding device.

Successive iterations of the adaptive refinement algorithm in the 5 point scheme are illustrated. Vertical solid lines denote the nodes at the beginning of each iteration. Vertical dotted lines denote nodes added at each iteration.

(a) Transmission as a function of energy for a Resonant tunneling device with thick barriers giving rise to very sharp features in transmission. Similar devices have been modeled in [22] (b) The corresponding distribution of nodes generated using the 5 point adaptive grid scheme running with $\epsilon = 0.1$.

The relative error as a function of node count for a normalized Lorentzian with FWHM $=10^{-3}$. The different data points for a given algorithm were obtained by varying the convergence criterion ($\epsilon$) for the 5 point scheme and the error tolerance for the remaining algorithms.

The relative error as a function of node count for a normalized Lorentzian with FWHM $=10^{-7}$. The different data points for a given algorithm were obtained by varying the convergence criterion ($\epsilon$) for the 5 point scheme and the error tolerance for the remaining algorithms.

Scatter plot showing the nodes added at each iteration for a broad resonance (FWHM=$10^{-3}$ centered at 0.3583). The 4pt and 1pt schemes are operating at a 10% convergence criterion and the 1pt log scheme is at 1%. 
Figure | Page
---|---
6.6 Difference in energy points for a broad resonance (FWHM=10^{-3} centered at 0.3583). The 4pt and 1pt schemes are operating at a 10% convergence criterion and the 1pt log scheme is at 1%. | 86
6.7 Relative Error as a function of node count for a broad resonance (FWHM=10^{-3} centered at 0.3583) | 87
6.8 Scatter plot showing the nodes added at each iteration for a sharp resonance (FWHM=10^{-7} centered at 0.3583). The 4pt and 1pt schemes are operating at a 10% convergence criterion and the 1pt log scheme is at 1%. | 88
6.9 Difference in energy points for a sharp resonance (FWHM=10^{-7} centered at 0.3583). The 4pt and 1pt schemes are operating at a 10% convergence criterion and the 1pt log scheme is at 1%. | 89
6.10 Relative Error as a function of node count for a sharp resonance (FWHM=10^{-10} centered at 0.3583) | 89
Energy consumption has become an issue that needs to be addressed on various fronts. Semiconductor technology plays a vital role in different areas where the matter needs attention. In the field of photovoltaics the use of a high band-gap material capable of targeting the high energy photons in the so called "green gap" of the solar spectrum could take the efficiency of multi-junction solar cells beyond 50%. Though InGaN shows some potential, it suffers from phase separation of InN and a high defect density. A material system capable of targeting the "green gap" and alleviating the difficulties in fabrication is required. Equally important is the issue of power consumption in transistors. The supply voltage scaling of transistors which has a direct bearing on the power consumption, has not followed the dimensional scaling of transistors. This is mainly because MOSFETs are fundamentally limited to a 60mV/dec subthreshold slope. Alternate channel materials or new devices are required to allow for a more aggressive supply voltage scaling.

In this work, using an atomistic full-band quantum description of nanoscale devices, design guidelines that seek to address each of these issues are suggested. For multi-junction solar cells the ZnSe/GaAs system has been proposed as a viable alternative. Using atomistic tight-binding the band-gap of the ZnSe/GaAs (001) superlattice as a function of the constituent monolayers is investigated. The possibility of engineering a range of bandgaps with the same material system, with a view to achieving the optimum value for solar cells and light emitting diodes, is proposed. For the purpose of supply voltage scaling, Tunnel Field Effect Transistors (TFETs) are promising devices because there is no lower limit to the subthreshold slope. Tra-
ditional TFET geometries suffer from low ON currents. Using an atomistic full-band quantum transport solver, a vertical TFET geometry that offers larger tunneling area and hence larger ON currents is investigated. The benefits of a global tunneling model for band to band tunneling over other approaches is highlighted. Finally, a design modification that reduces the lateral leakage current in the vertical TFET geometry making it suitable for low power logic applications, is proposed.
1. INTRODUCTION

1.1 EFFICIENCIES IN PHOTOVOLTAICS

The photovoltaics industry has been striving to bring down costs either by increasing the efficiency of the cells or by decreasing the cost per unit area of photovoltaic cells. If the cost of a system increases significantly with increasing area the problem of increasing the efficiency of solar cells assumes great importance. A number of approaches including multi-junction solar cells, thin-film solar cells and solar cells based on crystalline silicon are in use.

Multi-junction photo-voltaic cells consist of many layers of semi-conductors or thin films with each having a different band-gap. The idea is that each layer can efficiently absorb only a portion of the incident spectrum, such that taken together a major portion of the solar spectrum can be efficiently converted to electricity. From a cost perspective, multi-junction solar cells have become viable with concentrated photovoltaic technology where sunlight is focussed into a smaller area thereby reducing the area of the solar cell. Multi-junction solar cells on GaAs, Ge and InP substrates have been manufactured with record efficiencies around 40%.

Thin film solar cells are fabricated by depositing thin films of photovoltaic materials on a substrate. Amorphous Si, CdTe, Copper Inidium Gallium Selenide (CIGS) are some of the photovoltaic materials used. In addition to that dye sensitized solar cells [1] and organic solar cells have also been used. The best thin film solar cell efficiencies are around 20%.

In addition to that monocrystalline Si based solar cells are also being fabricated with record efficiencies around 25%.

Considerable effort in the field of multi-junction solar cells has been to improve the efficiency of solar cells by using a high band-gap material that can target the
Table 1.1
Bandgaps and lattice constants of materials used in high efficiency photovoltaics.

<table>
<thead>
<tr>
<th>Material</th>
<th>Bandgap (eV)</th>
<th>Lattice Constant (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ge</td>
<td>0.67</td>
<td>0.564</td>
</tr>
<tr>
<td>Si</td>
<td>1.11</td>
<td>0.543</td>
</tr>
<tr>
<td>InP</td>
<td>1.34</td>
<td>0.587</td>
</tr>
<tr>
<td>GaAs</td>
<td>1.42</td>
<td>0.565</td>
</tr>
<tr>
<td>CdTe</td>
<td>1.44</td>
<td>0.648</td>
</tr>
<tr>
<td>ZnSe</td>
<td>2.7</td>
<td>0.567</td>
</tr>
</tbody>
</table>

2.3-2.5eV range. This is crucial to take the efficiency of multi-junction solar cells beyond 50%. A lot of work in this area is being conducted with InGaN. The problem with InGaN is that it suffers from the phase separation of InN which does not allow the band-gap to remain in the desired range. InGaN green/blue-green LEDs have been demonstrated though solar cells are not efficient because of the high defect density. Even though SiC has a band-gap close to the desired range, because of its indirect nature it is also not considered a feasible option. In this regard, ZnSe/GaAs ($E_{\text{gap, ZnSe}} = 2.7\text{eV} \quad E_{\text{gap, GaAs}} = 1.42\text{eV}$) digital alloys using a superlattice have been proposed as a viable solution. The digital alloy approach might be more desirable compared to physical alloying given that maintaining control on doping densities in the quaternary would be difficult, though it is possible to controllably dope pure ZnSe and pure GaAs. The band-gaps of the ZnSe/GaAs superlattice, as a function of its period, will give an insight into the usefulness of such a system, should it be possible to fabricate a superlattice with the correct period. Current fabrication techniques have nearly atomic layer precision, which makes it possible to imagine superlattices with periods of only a few atomic monolayers. Therefore the electronic structure calculation of such structures requires that details at the atomic level be taken into account.
1.2 POWER CONSUMPTION IN TRANSISTOR TECHNOLOGY

Power consumption has been a long standing issue for transistor technology. Historically, the supply voltage of a transistor, which has a direct bearing on the power consumption of the device, has not followed the dimensional scaling of the device as shown in Fig. 1.1. This has resulted in an increasing power consumption per chip with each generation. The primary purpose of dimensional scaling was to shrink the size of the integrated circuits so that more and more could be packed in the same area, thereby reducing the cost. Since the number of circuits on a chip has gone up and because the frequency of operation increases with every technology generation, the power consumption per chip has increased. The power consumption has been kept under control only because of a decreasing supply voltage ($V_{DD}$), but clearly the proportionality between the supply voltage and technology node is not being followed.

Figure 1.1. Supply voltage ($V_{DD}$) of transistors as a function of technology node size [2].
Though supply voltage scaling in current MOSFET technology is possible, it will come with severe design compromises. Reducing the supply voltage of MOSFETs would require a corresponding reduction in threshold voltage in order to maintain circuit speeds. This in turn would result in an increase in the subthreshold leakage current, which is not desirable.

MOSFETs are fundamentally limited to a subthreshold slope of 60mV/dec, though there is no lower limit for a Tunnel-FET. The electrons in a MOSFET have to overcome a potential barrier. Given that the carrier distribution in the source is Boltzmann, an order of magnitude change in current can only be facilitated by a 60mV change in the height of the potential barrier.

As an alternative to the MOSFET, for the purpose of supply voltage scaling, the Tunnel-FET seems to be a promising candidate because of the possibility of achieving a subthreshold slope lower than 60mV/dec. Though subthreshold slopes lower than 60mV/dec have been demonstrated experimentally, the biggest problem remains that of low ON currents. A new vertical tunneling approach that seeks to increase the tunneling area for band to band tunneling current is investigated that could potentially solve the problem of low ON currents. To simulate band to band tunneling current through such structures an approach that not only takes into account the coupling between the conduction bands and the valence bands, but also band-structure effects due to geometrical and electrostatic confinement, is required.

1.3 NEED FOR ATOMISTIC MODELS AND THE TIGHT BINDING APPROACH

As device features scale to tens of nanometers, a quantum mechanical treatment becomes imperative to capture the physics. While this issue is often addressed, what is not fully appreciated is the need to capture device details at the nanometer level using an atomistic treatment. As depicted in Fig. 1.2, the effective mass approximation has proven to be inadequate to explain the effects of bandstructure [3, 4]. In devices
Figure 1.2. The localized conduction band edge for a GaAs/AlAs/GaAs structure. Along the Γ direction the band edge presents a barrier while in the X direction it looks like a well. The bulk bandstructures have been generated using tight-binding. This figure is based on the discussion in [7].

Figure 1.3. Difference at higher energies between the parabolic bulk conduction band of GaAs and a non-parabolic band generated using tight-binding. This figure is based on the discussion in [7].

which have geometrical or electrostatic confinement of the order of nanometers, the bandstructure differs significantly from that of bulk. Atomistic models on the other hand, are capable of capturing band-non-parabolicities, as shown in Fig. 1.3, band-to-band mixing [5,6] and crystal orientations, all of which assume a lot of significance as we investigate devices a few nano-meters in size.
It is important that for a quantum mechanical treatment the Hamiltonian be written in an appropriate basis. The tight-binding approach uses basis functions that have atomic orbital like symmetries. Depending on the requirement of the problem and the material system involved more atomic orbitals might be included to model the bandstructure accurately. In addition to that spin-orbit effects should be taken into account to model the imaginary band coupling the conduction band to the valence band. An accurate description of the bandstructure ensures that the details of the electronic properties of the system are included.

1.4 ASSOCIATED NUMERICAL TECHNIQUES

Considerable numerical and computational effort is required in the atomistic simulation of nanoscale devices at different stages. Starting from getting the optimized tight-binding parameter sets using either genetic algorithms [8] or least square fitting techniques. The parameter sets are optimized such that they are able to generate the experimentally known band gaps and effective masses of the bulk bandstructure of a given material. The tight binding Hamiltonian is constructed using these parameter sets. Since each atom is accounted for in the Hamiltonian, a device only a few nanometers in size could have a very large Hamiltonian. The next challenge is to calculate the open boundary conditions that need to be added to the Hamiltonian before any quantum transport calculations can be done. The algorithms that compute open boundary conditions could be iterative [9], which require the repeated inversion of dense or full matrices, or based on scattering boundary methods that require the solution of generalized or normal eigenvalue problems [10, 11]. A quantum transport solver based on the Non-equilibrium Green’s function method [12, 13] or the wave function method [14] is used to calculate charge and current. Some techniques like the Recursive Green’s function [15] are used to efficiently invert large matrices to find the Green’s function. Features in energy or momentum space can be extremely sharp depending on the geometry and materials used. For the efficient integration of such
features, resonance finding techniques [16] are employed. The output of these techniques can be used to generate optimized grids [17] to integrate over. Alternatively various methods in adaptive quadrature can be used to efficiently integrate over such features.

1.5 BRIEF OUTLINE

In Chapter 2 associated numerical techniques that go into quantum transport calculations are discussed and results are shown for single band and some multi-band cases also. In Chapter 3 the 1D Heterostructure tool is discussed which allows the investigation of the distribution of charge in heterostructures. Chapter 4 addresses the issue of efficiencies in multi-junction solar cells. In Chapter 5 the problem of reducing the power consumption in transistors is discussed. A new technique for adaptive quadrature is discussed in Chapter 6 which aims to reduce the cost of integrating a sharply spiked integrand while still maintaining acceptable accuracies. Finally the thesis is summarised in Chapter 7.
2. NUMERICAL METHODS AND ASSOCIATED TOPICS

2.1 RECURSIVE GREEN'S FUNCTION: ALGORITHM AND IMPLEMENTATION

The Recursive Green’s Function technique [18] should be thought of as a tool for inverting block tri-diagonal matrices. Here it might be mentioned that a nearest neighbor Hamiltonian is block tri-diagonal. But a next nearest neighbor Hamiltonian which is penta-diagonal can be thought of as a tri-diagonal system if the block size is increased [15]. The technique is particularly useful in our case since even after the addition of the self-energy matrix [19] the Hamiltonian still retains its block tri-diagonal nature. From the computing perspective we can stop the RGF technique at any stage we choose depending on the quantities we are interested in, i.e. diagonal entries, columns etc. This section outlines the detailed steps required for such an implementation. The discussion is based on the NEMO 1D documentation and Ref. [20].

- We need to compute the following

\[ G_R(E) = [EI - H_{tight-binding} - \Sigma]^{-1} \]  
(2.1)

\[ G_R(E) = \begin{bmatrix}
D_1 & t_{1,2} & & & & & \\
t_{2,1} & D_2 & t_{2,3} & & & & \\
& t_{3,2} & D_3 & t_{3,4} & & & \\
& & & \ddots & \ddots & \ddots & \\
& & & & t_{N-1,N} & & \\
t_{N,N-1} & & & & & D_N & \\
\end{bmatrix}^{-1} \]  
(2.2)

Since the Hamiltonian is itself dependent on \( \vec{k} \), one pass through the RGF technique is for one \( E \) and one \( \vec{k} \) value. It follows, that all quantities, i.e., density
Figure 2.1. Schematic representation of forward and backward recursion in the Recursive Green's Function algorithm and the subsequent calculation of the first and last column. The diagonal blocks of $G^R$ give the local density of states. The first and last columns are required to compute the charge.

of states, charge etc., are valid for the given $E$ and $\vec{k}$ value. In this section this dependence has been suppressed to avoid tedious notation.

- Forward Recursion: Unconnected Green function

\[
g_{i,1}^{left,R} = (D_1)^{-1} \tag{2.3}
\]

\[
g_{i,i}(i=2,3,...N) = (D_i - t_{i,i-1}g_{i-1,i-1}^{left,R}t_{i-1,i})^{-1} \tag{2.4}
\]

or

\[
g_{N,N}^{right,R} = (D_N)^{-1} \tag{2.5}
\]

\[
g_{i,i}(i=N-1,N-2,...1) = (D_i - t_{i,i+1}g_{i+1,i+1}^{right,R}t_{i+1,i})^{-1} \tag{2.6}
\]
Depending on the quantities required to be computed, either or both of \( g^{left,R} \) and \( g^{right,R} \) might be needed.

- **Backward Recursion: Connected Green Function diagonal block**

  \[
  G_{N,N}^R = (D_N - t_{N,N-1}g_{N-1,N-1}^{left,R})^{-1} = g_{N,N}^{left,R} \tag{2.7}
  \]

  \[
  G_{i,i(i=N-1,N-2,...1)}^R = g_{i,i}^{left,R} + g_{i,i}^{left,R}t_{i,i+1}G_{i+1,i+1}^R + t_{i,i+1}g_{i,i}^{left,R} \tag{2.8}
  \]

  or

  \[
  G_{1,1}^R = (D_1 - t_{1,2}g_{2,2}^{right,R})^{-1} = g_{1,1}^{right,R} \tag{2.9}
  \]

  \[
  G_{i,i(i=2,3,...N)}^R = g_{i,i}^{right,R} + g_{i,i}^{right,R}t_{i,i-1}G_{i-1,i-1}^R + t_{i,i-1}g_{i,i}^{right,R} \tag{2.10}
  \]

- **General Expression for off-diagonal blocks**

  \[
  G_{i,j(i<j)}^R = G_{j,i(i<j)}^R = -g_{i,i}^{left,R}t_{i,i+1}G_{i+1,j}^R \tag{2.11}
  \]

  \[
  G_{i,j(i>j)}^R = G_{j,i(i>j)}^R = -g_{i,i}^{right,R}t_{i,i-1}G_{i-1,j}^R \tag{2.12}
  \]

As shown schematically in Fig. 2.1, the trace of the diagonal blocks of \( G^R \) will give us the Local Density of States [10]. The first and last columns of \( G^R \) are required to compute the charge [22]. Another important quantity is the transmission which can be computed with the quantities already given by the RGF technique. Further the transmission is necessary in the calculation of current [15]. Fig. 2.2 shows the local density of states and the transmission computed using the RGF.

### 2.2 RESONANCE FINDING USING THE SHIFT AND INVERT NON-SYMMETRIC LANCZOS ALGORITHM AND NEWTON’S METHOD: IDEA AND IMPLEMENTATION

Even though the tight-binding Hamiltonian is itself not a function of energy and depends only on \( \vec{k} \), the addition of the self-energy matrix makes it energy dependent.
This makes the problem of finding the eigenvalues of the total Hamiltonian or more specifically the poles of the retarded Green’s function, non-linear.

\[ |EI - H - \Sigma(E)| = \left| G^R(E)^{-1} \right| \]  \hspace{1cm} (2.13)

It is expensive to compute the self-energy matrix for each energy point. Therefore a shift and invert non-symmetric (SINS) Lanczos algorithm is used that computes a subspace that accurately represents the region of the eigen-spectrum which is of interest. The approach is laid out as follows.
Figure 2.3. Local Density of states for a double barrier GaAs/AlGaAs/GaAs structure with resonances as found by the resonance finder using a single band effective mass scheme [21]. The resonances are compared with the analytical solutions of a particle in a box.

- The energy domain is divided into segments. The width of the segment should be decided keeping in mind the time that a user is willing to spend on resonance finding. The trade-off is missing eigenvalues altogether.

- In each of these segments the SINS Lanczos algorithm [16] in combination with the partial LR algorithm [23] is used to find approximate eigenvalues. The boundary conditions are calculated at the centre of each segment.
The Newton’s method is used to refine these eigenvalues. For these energies

\[ |EI - H - \Sigma(E)| = 0 \]  

(2.14)

and they correspond to the poles of \( G^R(E) \).

Fig. 2.3 shows how the resonances compare with the local density of states and the analytical results of a particle in a box. There is agreement with the local density of states, though the discrepancy with the analytical results can be explained on the basis of an increase in the effective length of the well due to barrier penetration. That is why the results for a particle in a box are always higher in energy.

### 2.3 RESONANCE REFINEMENT USING INVERSE ITERATION: ALGORITHM AND IMPLEMENTATION

A further refinement of eigenvalues might be necessary to accurately locate a resonance on the real axis and be sure of its width, which is the imaginary part of the eigenvalue. Having already found the eigenvalues to some accuracy using the SINS Lanczos algorithm and Newton’s method, this method converges to an even better estimate. The method also lets us estimate the eigenvector which tells us where the resonance is present spatially in the system. This fact is of great importance in creating energy grids, since not all eigenvalues are important. Only the resonances inside the device should contribute towards the generation of an inhomogeneous energy grid. The details of the algorithm are as follows [24].

- Start with a normalized random column vector with the same dimension as that of the Hamiltonian.
Figure 2.4. Local Density of states for a double barrier GaAs/AlGaAs/GaAs structure with wave-function magnitude squared found using inverse iteration in a single band effective mass scheme [21]. The variation of the wavefunction matches well with the local density of states.

- Go through the following iteration

\[ \text{Solve} : (H + \Sigma(E_m) - E_m I)x_{m+1} = x_m \]  \hspace{1cm} (2.15)

\[ \delta E = \|x_{m+1} . x_m\|^{-1} \]  \hspace{1cm} (2.16)

\[ E_{m+1} = E_m + \delta E \]  \hspace{1cm} (2.17)

\[ x_{m+1} = x_{m+1} / \|x_{m+1}\| \]  \hspace{1cm} (2.18)

- This process goes on till \( \delta E \) satisfies a convergence criterion.

- \( E_m \) is the converged eigenvalue and \( x_m \) is the associated eigenvector.
The converged eigenvalue is a better estimate than what is obtained at the end of
Newton refinement. It should be noted that this method converges fast only if we
already have a good estimate of the eigenvalue. The associated eigenvector should
then be used to find the position of the resonance inside the device. Fig. 2.4 shows
the calculation for a double barrier structure. The variation in the local density of
states is in agreement with the wave-functions.

2.4 IN-HOMOGENEOUS ENERGY GRID CREATION THROUGH ANAL-
YSIS OF SPECTRAL FEATURES: GRID MAPPINGS

For many reasons, it is often desirable to work with fixed energy grids. Adaptive
grids that are generated by the addition of nodes within the integration range of
interest are often based on a relative error analysis and do not take into account the
overall contribution of the added node to the integral. This section deals with the
creation of fixed energy grids optimized to have an in-homogeneous distribution of
energy points. The distribution is determined by the variation in the density of states,
( which is heavily dependent on the position of resonances) and Fermi-functions [17].
Assuming that we already know the eigenvalues(both real and imaginary parts) of the
Hamiltonian that describes our quantum device we proceed to model each resonance
as follows.

- Each resonance is given by $E_r - i\Gamma$ where $E_r$ denotes the location and $\Gamma$ the
  width of the resonance.
- Close to the resonance we can model the density of states as a Lorentzian located
  at $E_r$ with a width of $\Gamma$. The integral over this function gives us the contribution
  called $m_{IM}(E)$.
- The magnitude of the principal value integral gives us $m_{RE}(E)$. 
• Since the real part of $G^R$ is doubly peaked at $E_r \pm \Gamma/2$, two more Lorentzian functions are placed at $E_r \pm \Gamma/2$ with a width of $(\sqrt{3}-1)\Gamma/2$. The integral over these gives us $m_{RE-IM}(E)$.

Figure 2.5. Contribution to spectral features: Lorentzian

Figure 2.6. Contribution to spectral features: Principal value Integral

Though the resonances have been taken care of, there are other spectral features that must be considered as well.
The density of states around the conduction band edge is modeled as a constant number below the conduction band edge and as $1/\sqrt{E}$ above it. The integral of this gives us $m_{\text{band-edge}}$.

- The Fermi occupation factor is taken into account with $m_{\text{Fermi}} = 1 - f$

- The turn-ons of the Fermi function are modeled as the fourth derivative of $f$. The contribution is $m_{\text{Fermi'}}$ which is the integral of that function, is again a
monotonic function. This contribution captures the energy domain in which the Fermi occupation is important.

Taking a weighted average of these we get,

\[ m(E) = x_1 m_{IM} + x_2 m_{RE} + x_3 m_{RE-IM} + x_4 m_{band-edge} + x_5 m_{fermi} + x_6 m_{fermi'} \quad (2.19) \]
Figure 2.11. Local Density of states for a double barrier GaAs/AlGaAs/GaAs structure calculated on an energy grid generated using the grid mapping scheme using a single band effective mass scheme [21].

, which is a monotonic function of energy. It must be kept in mind that weights of the contributions from the resonances must reflect the number of resonances. Fig. 2.11 shows the energy grid using the mapping scheme based on spectral features for a double barrier structure.
2.5 ADAPTIVE ENERGY GRID CREATION BASED ON CHARGE, CURRENT AND DENSITY OF STATES

Given the significant overhead of resonance finding imposed by in-homogeneous energy grid algorithms, it might be feasible to create energy grids adaptively as well. Successive partitioning of the integration range, where integrands vary rapidly, adds more nodes to the existing grid. The process is continued till the relative contribution of two adjoining nodes is within a specified percentage (typically 1 percent or even smaller) of the total contribution.
Further a baseline energy value is defined which does not allow the addition of nodes closer than the given energy difference. There are a few issues that need consideration.

- The addition of a node improves the value of the integral only in the energy segment where the node is added. For the overall error of the integral this might not be important at all.

- There is always a chance that the adaptive grid gets caught in a "noisy" node and keeps on adding nodes in that energy segment. The baseline energy helps to reduce this problem.

Fig. 2.12 shows the energy grid generated using the adaptive grid procedure to calculate the local density of states for a double barrier structure.

2.6 BOUNDARY CONDITIONS FOR MULTI-BAND TIGHT-BINDING

Boundary conditions are perhaps the most critical components of any geometry and differential equation. Much of the underlying calculation is determined by the type of boundary conditions applied to the system, in solving the Schrödinger equation. In this section we discuss both closed and open boundary conditions in a multi-band framework.

2.6.1 HARD-WALL, PASSIVATION OF SURFACE STATES

Imposing hard-wall or closed boundary conditions involves passivation of the dangling bonds that arise at the beginning and end of the UTB chain [25]. We need to transform from the atomic orbital basis to the $sp^3$ hybridized basis, since these bonds
are primarily formed by $sp^3$ hybridization. For an anion to its nearest neighbors we can write the following,

\begin{align*}
|sp^3⟩_{[111]} &= \frac{1}{2}(|s⟩ + |p_x⟩ + |p_y⟩ + |p_z⟩) \\
|sp^3⟩_{[\bar{1}11]} &= \frac{1}{2}(|s⟩ - |p_x⟩ - |p_y⟩ + |p_z⟩) \\
|sp^3⟩_{[1\bar{1}1]} &= \frac{1}{2}(|s⟩ + |p_x⟩ - |p_y⟩ - |p_z⟩) \\
|sp^3⟩_{[\bar{1}1\bar{1}]} &= \frac{1}{2}(|s⟩ - |p_x⟩ + |p_y⟩ - |p_z⟩)
\end{align*}

If the chain of atoms is constructed in the [001] direction then the bonds [111] and [\bar{1}11] are above the [001] plane and the bonds [1\bar{1}1] and [\bar{1}1\bar{1}] are below it. Therefore the transformation can be written as,

\begin{align*}
\begin{bmatrix}
|sp^3⟩_{[111]} \\
|sp^3⟩_{[\bar{1}11]} \\
|sp^3⟩_{[1\bar{1}1]} \\
|sp^3⟩_{[\bar{1}1\bar{1}]}
\end{bmatrix} &= \frac{1}{2}
\begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & -1 & -1 & 1 \\
1 & 1 & -1 & -1 \\
1 & -1 & 1 & -1
\end{bmatrix}
\begin{bmatrix}
|s⟩ \\
|p_x⟩ \\
|p_y⟩ \\
|p_z⟩
\end{bmatrix}
\end{align*}

(2.24)

When the top surface is made of anions, corresponding to the bonds [111] and [\bar{1}11] we want to add the following,

$$[H]_{hybrid} =
\begin{bmatrix}
\delta_{sp^3} & 0 & 0 & 0 \\
0 & \delta_{sp^3} & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
$$

(2.25)

to the corresponding element in the Hamiltonian in the hybridized basis. This corresponds to adding,

$$[U]^\dagger[H]_{hybrid}[U] = \frac{1}{2}
\begin{bmatrix}
\delta_{sp^3} & 0 & 0 & \delta_{sp^3} \\
0 & \delta_{sp^3} & \delta_{sp^3} & 0 \\
0 & \delta_{sp^3} & \delta_{sp^3} & 0 \\
\delta_{sp^3} & 0 & 0 & \delta_{sp^3}
\end{bmatrix}
$$

(2.26)
in the original basis where,

\[
[U] = \frac{1}{2} \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & -1 & -1 & 1 \\
1 & 1 & -1 & -1 \\
1 & -1 & 1 & -1
\end{bmatrix}
\] (2.27)

If the bottom layer is also anionic then we need to add,

\[
[U]^\dagger \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & \delta_{sp^3} & 0 \\
0 & 0 & 0 & \delta_{sp^3}
\end{bmatrix} [U] = \frac{1}{2} \begin{bmatrix}
\delta_{sp^3} & 0 & 0 & -\delta_{sp^3} \\
0 & \delta_{sp^3} & -\delta_{sp^3} & 0 \\
0 & -\delta_{sp^3} & \delta_{sp^3} & 0 \\
-\delta_{sp^3} & 0 & 0 & \delta_{sp^3}
\end{bmatrix}
\] (2.28)

to the corresponding element in the Hamiltonian in the original basis. If we repeat the procedure for cations at the top and the bottom the transformation changes to,

\[
[U] = \frac{1}{2} \begin{bmatrix}
1 & -1 & 1 & 1 \\
1 & 1 & -1 & 1 \\
1 & -1 & -1 & -1 \\
1 & 1 & 1 & -1
\end{bmatrix}
\] (2.29)

So for the top cationic interface we will add

\[
[U]^\dagger \begin{bmatrix}
\delta_{sp^3} & 0 & 0 & 0 \\
0 & \delta_{sp^3} & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix} [U] = \frac{1}{2} \begin{bmatrix}
\delta_{sp^3} & 0 & 0 & \delta_{sp^3} \\
0 & \delta_{sp^3} & \delta_{sp^3} & 0 \\
0 & \delta_{sp^3} & \delta_{sp^3} & 0 \\
\delta_{sp^3} & 0 & 0 & \delta_{sp^3}
\end{bmatrix}
\] (2.30)

and the bottom cationic interface,

\[
[U]^\dagger \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & \delta_{sp^3} & 0 \\
0 & 0 & 0 & \delta_{sp^3}
\end{bmatrix} [U] = \frac{1}{2} \begin{bmatrix}
\delta_{sp^3} & 0 & 0 & -\delta_{sp^3} \\
0 & \delta_{sp^3} & -\delta_{sp^3} & 0 \\
0 & -\delta_{sp^3} & \delta_{sp^3} & 0 \\
-\delta_{sp^3} & 0 & 0 & \delta_{sp^3}
\end{bmatrix}
\] (2.31)

to the corresponding elements of the Hamiltonian in the original basis.
2.6.2 OPEN BOUNDARY CONDITIONS

A crucial part of the Non-equilibrium Green’s function technique is the computation of the open boundary conditions, that is the self-energy matrix based on the Dyson equation treatment used by many authors [4, 26–31]. A lot of the computational effort goes into the computation of the open boundary conditions. The method proposed by Luisier et al [11, 32], has been employed which has been shown to outperform other methods in the literature [9]. It starts from a scattering boundary ansatz and works for an arbitrary transport direction. The method can be thought of as a band-structure calculation of an infinite structure with exchanged input and output variables. In a band-structure calculation the aim is to find all the eigenvalues at a given \( \vec{k} \) point. In computing the open boundary conditions the aim is to find all the wave-vectors \( k \) for a given energy. Further depending on the sign and value of the imaginary part of the wave-vector and whether it is in the left or right reservoir, one can determine whether the state corresponds to transmission, reflection, decaying transmission or decaying reflection. Having computed the eigen-fucntions of the states transmitted or reflected inside the device, the surface Green’s function on either side of the device (left and right) is found. Finally the surface Green’s function [16] gives the boundary self-energy, which is readily used in the formulation laid out in Eq. (2.1).

2.7 RESONANCE FINDER AND RGF RESULTS FOR RESONANT INTER-BAND TUNNELING DIODES USING OPEN BOUNDARY CONDITIONS

In this sections results for the resonant inter-band tunneling devices using multi-band tight binding have been shown. The reasons for choosing resonant inter-band tunneling diodes are discussed. Numerical techniques mentioned in the previous section are used in conjunction and results are shown to be in agreement.
2.7.1 INTEREST IN RESONANT INTER-BAND TUNNELING DIODES

Here we deal with nearly lattice matched InAs/GaSb and InAs/GaSb/AlSb material systems. These systems exploit the various band alignments based on the offset values. All $sp^3s^*d^5$ tight binding parameter sets taken from [33] are based on the assumption that the valence band maximum is at zero. The band offsets for these materials are incorporated based on the values given in [34]. These devices are of interest because of the high current densities they have shown. In addition to that, large peak-to-valley current ratios have been observed making them an attractive option for digital and analog applications. Very high peak to valley current ratios have been reported [35] at room temperature as well as at low temperature. Various Interband devices have been studied [36–42], showing a great potential in device application. Though the primary transport mechanism arises from the coupling of the InAs conduction band states to the GaSb light hole states [43, 44], it is possible that hole mixing effects in device structures containing GaSb quantum wells could lead to coupling of InAs conduction band states and GaSb heavy hole states, which changes the current-voltage characteristics [45]. Studies to date have been based on either the $\vec{k}.\vec{p}$ model, that does not couple light hole, heavy hole and split-off bands at $k = 0$ [45, 46], or on the $sp^3s^*$ nearest neighbor or next nearest neighbor model that does not give a satisfactory fit of the valence band dispersion [47–50]. Specifically, the inability of these models to model shifts of the heavy hole from the zone centre has been reported [51]. Since these features have been reported for GaSb without a complete treatment of spin-orbit coupling [52, 53], here we use the $sp^3s^*d^5$ nearest neighbor model and include spin-orbit interaction to all orders which ensures that the valence band dispersions are captured correctly.

2.7.2 NON-INTUITIVE BEHAVIOR OF HOLE BANDS IN GaSb

Light Hole and heavy hole band anisotropy has been discussed for GaAs [54]. Here we show similar trends for GaSb. Bulk dispersion have been calculated using
Figure 2.13. Bulk HH dispersion for GaSb in the $k_x - k_y$ plane

Figure 2.14. Bulk LH dispersion for GaSb in the $k_x - k_y$ plane
the $sp^3s^*d^5$ model in the nearest neighbor formulation. The model represents hole anisotropy better than the $sp^3s^*$ nearest neighbor formulation.

Here we show a light hole and heavy hole bulk dispersion in the $kx - ky$ plane. As can be seen in Fig. 2.13, heavy hole dispersion indicates a heavier effective mass in the [110] direction compared to the [100] direction. The ridge like feature clearly indicates a high degree of anisotropy. The light hole dispersion in Fig. 2.14 appears to be isotropic, but the non-parabolicity is still present.

2.7.3 ANALYSIS OF GaSb BASED BRIT STRUCTURE USING TRANSMISSION

We proceed to analyze hetero-structures where there is the added complication of the coupling of the conduction and valence bands. First, we consider a Barrier-less Resonant Interband Tunneling Structure (BRIT) as shown in Fig. 2.15. The structure gets its name from the band profile. The device consists of 7nm of GaSb with InAs on either side. The valence bands of GaSb at 0.83eV couple with the conduction band of InAs at 0.68eV. Since there is no conventional barrier, the structure is called a BRIT. The accompanied Fig. 2.16 shows the transmission for two different momentum ($k = 0$ and $k = 0.04$). Two sharp peaks can be seen with a broad peak. The coupling of the GaSb LH band is strong with the conduction band of InAs resulting in a broad peak. The HH coupling is weaker which gives sharper peaks. Next, we try to connect these diagrams using the energy dispersion along the transverse momentum. It should be noted that the transmission at two different momenta are not energy shifted variations of each other. The $E - k$ diagram is not intuitive and can move up or down with momentum.
Figure 2.15. Schematic band profile of a BRIT structure.

Figure 2.16. Transmission and E-k plots for the BRIT structure.
Figure 2.17. Schematic band profile of a RIT structure.

Figure 2.18. Transmission and E-k plots for the RIT structure.
2.7.4 ANALYSIS OF GaSb BASED RIT STRUCTURE USING TRANSMISSION

For the Resonant Interband Tunneling Structure (RIT) in Fig. 2.17 we simulate 7nm of GaSb with 1.5nm of AlSb on either side, followed by InAs on either side. Since the valence band of AlSb is at 0.3eV and the conduction band is at 2.75eV, it serves as a barrier. We see the same features repeated in Fig. 2.18, two sharp peaks corresponding to the weak HH coupling with the conduction band and the broad peak corresponding to the strong LH coupling with the conduction band. The dispersion relation connecting the two transmission plots is again not intuitive and one cannot obtain one transmission plot by shifting the other. In fact, an added feature, the splitting of the lowest band on the dispersion relation is seen. This is also verified by the double peaks in the transmission at an energy close to 0.75eV for $k = 0.04$.

2.7.5 ANALYSIS OF GaSb BASED BRIT AND RIT STRUCTURES USING RESONANT LINE-WIDTHS

Finally, we show the resonant line-widths for the dispersion relations computed before. The resonant line-widths correspond to the value of the imaginary part of the eigenvalue found in determining the dispersion relations. Since we are dealing with open systems, all eigenvalues in general are complex and the imaginary part is indicative of the life-time of the state it corresponds to. As can be seen in Fig. 2.19 and Fig. 2.20 the broad LH resonance has the largest line-width corresponding to the wide peak we saw on the transmission plots, for both the BRIT and RIT structures. For the HH resonances the line-widths become sharper but also show some non-monotonic variations. In the case of the BRIT structure the first HH peak just gets sharper close to $k = 0$. The second HH peak is the sharpest at a non-zero momentum. Both the peaks get broader by 2-3 orders of magnitude with momentum. This fact is one of the many factors that contribute to off-zone center hole transport [55].
Figure 2.19. Resonant line-width for the BRIT structure.

Figure 2.20. Resonant line-width for the RIT structure.
For the RIT structure the second HH peak becomes sharper closer to $k = 0$. The first HH peak splits into two, indicating spin-splitting. The two split bands not only have different line-widths but also the variations are different.

The preceding analysis suggests that hole band-structure is more complicated than electron band-structure. An explicit integration over momentum is required to compute quantities like charge and current to capture all the features discussed here. In addition to that spin-orbit coupling should be included to all orders, since it is crucial to the correct modeling of the hole band-structure.
3. 1D HETEROSTRUCTURE TOOL

An essential component of understanding semiconductor structures is to understand how charge is distributed both in energy and in space. As devices become smaller and smaller the focus shifts from classical physics to quantum physics. Quantum mechanics predicts discrete states at a definite energy rather than a continuum of states that are obtained from classical physics. Heterostructures or layers of different semi-conducting materials with different doping concentrations built on top of each other, present an opportunity to investigate such quantum effects. In this chapter, the 1D Heterostructure Tool is used to understand how these states are distributed in a given heterostructure both in space as well as in energy. In particular, the problem is of interest to someone interested in the operation of such a heterostructure, so as to understand the flow of electrons inside it. The tool allows the control of external potential by the application of voltages at selected terminals that can control the working of such a device. The entire discussion would not be complete without a comprehension of the internal electrostatics of a given structure. Finding the correct electrostatic potential, such that the charge calculated using it is in agreement with the doping of the structure, is necessary to get physically meaningful results. The tool takes this fact into account and shows how the electrostatics have a profound impact on the working of any heterostructure.

3.1 METHODOLOGY

To solve any problem quantum mechanically, or equivalently to solve the Schrödinger equation, it is required that the Hamiltonian of the system, in this case the closed
system be written. In the finite-difference scheme the Hamiltonian using just an effective mass description can be written as,

$$-rac{\hbar^2}{2m} \frac{(\psi_{i+1} - 2\psi_i + \psi_{i-1})}{\delta^2} + V(x_i)\psi_i$$  \hspace{1cm} (3.1)$$

where \(m\) is the effective mass, \(V\) is the electrostatic potential and \(\delta\) is the spacing between the nodes. Using the effective mass here implies that a continuum description of the device has been employed and as such the spacing between nodes can be as small as possible. However in other descriptions such as multi-band tight binding [3], the meshing has to be atomistic. In the matrix form the effective mass Hamiltonian without the electrostatic potential looks like,

$$H = \begin{bmatrix}
2t_o & -t_o & & & \\
-t_o & 2t_o & -t_o & & \\
& -t_o & 2t_o & -t_o & \\
& & & \ddots & \ddots & \ddots \\
& & \ddots & \ddots & \ddots & -t_o \\
& & & -t_o & 2t_o &
\end{bmatrix}$$  \hspace{1cm} (3.2)$$

The electrostatic potential and the localized conduction band edge are added to the diagonal entries. Here \(t_o = \frac{\hbar^2}{2m\delta^2}\) where \(m\) is the effective mass of the semi-conductor at the \(i\)th site. The meshing is done such that points are on either side of an interface and never at it. Assuming continuity of the wavefunction and its derivative across the interface the diagonal terms can be written as,

$$d_i = \frac{\hbar^2}{\delta^2} \left( \frac{1}{m_{i-1} + m_i} + \frac{1}{m_i + m_{i+1}} \right) + V_i$$  \hspace{1cm} (3.3)$$

and off diagonal,

$$s_i = \frac{\hbar^2}{\delta^2} \frac{1}{m_{i-1} + m_i}$$  \hspace{1cm} (3.4)$$

In this case it turns out that the Hamiltonian is Hermitian. This implies that the eigenvalues are real. Had there been a complex part to the eigenvalues it would have meant that the states that corresponds to the complex eigenvalue could not have held charge forever, that is, had a finite lifetime. This lifetime is inversely proportional to
the complex part of the eigenvalue. For the case of real eigenvalues it implies that the state has the capability of holding charge for an infinite time. Using these eigenvalues and eigenvectors of this Hamiltonian the charge for the heterostructure needs to be calculated as follows,

\[ n(x) = \frac{mk_B T}{\pi \hbar^2} \sum_i |\phi_i(x)|^2 \log(1 + \exp\left(\frac{E_F - \epsilon_i}{k_B T}\right)) \]  

(3.5)

where \( T \) is the temperature, \( E_F \) is the Fermi level and \( \phi_i, \epsilon_i \) are the \( i \)th eigenvector and eigenvalue respectively. A discussion of how this equation is obtained can be found in [3]. Further the solution to the Poisson equation is required to obtain charge self-consistent results such that the electrostatic potential is in agreement with the charge in the structure. A discussion of how this is achieved and numerical issues associated with it can be found in [56]. Since this system is closed it follows that no current flows across the layers of the heterostructure.

3.2 TOOL USAGE AND CAPABILITIES

In this section we discuss how the tool can be used to study charge distribution and electronic states in heterostructures. We start with the default device in the tool.

3.2.1 DEVICE DETAIL OF THE DEFAULT STRUCTURE: HETEROJUNCTION FET

The band profile of the default device is shown in fig. 3.1. The device is basically GaAs indented with AlGaAs. It is commonly referred to as a Heterojunction Field Effect Transistor or a High electron mobility transistor (HEMT). The layer specification starts with the substrate which is typically a very long region. Subsequent layers are added below the substrate layer. Specifying a layer requires the input of the material it is made of, the mole fraction if it is an alloy, the doping concentration and the thickness which can be entered in nanometers or as the number of monolayers. In this case the AlGaAs region has a high doping (10^{18}/cc) whereas the GaAs
Figure 3.1. The band profile of the default H-FET structure without any electrostatics.

Figure 3.2. The doping profile of the default H-FET structure.
Table 3.1
Layer table for the default structure.

<table>
<thead>
<tr>
<th>Layer Name</th>
<th>Material System</th>
<th>X-Mole Fraction</th>
<th>Dop. Conc. $(cm^{-3}$ or $-2)$</th>
<th>Thick. (nm)</th>
<th>Mono layers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substr.</td>
<td>GaAs</td>
<td>0.0</td>
<td>$10^{14}$</td>
<td>1000.06</td>
<td>1769</td>
</tr>
<tr>
<td>L01</td>
<td>GaAs</td>
<td>0.0</td>
<td>$10^{14}$</td>
<td>149.81</td>
<td>265</td>
</tr>
<tr>
<td>L02</td>
<td>AlGaAs</td>
<td>0.3</td>
<td>$10^{18}$</td>
<td>29.96</td>
<td>53</td>
</tr>
<tr>
<td>L03</td>
<td>GaAs</td>
<td>0.0</td>
<td>$10^{14}$</td>
<td>19.79</td>
<td>35</td>
</tr>
</tbody>
</table>

layers have a low value of $10^{14}/cc$ as shown in fig. 3.2. In fig. 3.1 the AlGaAs region is shown as a barrier because of the relative band edges of the materials that make up the device. This is the case when it is assumed that the electrostatic potential is zero everywhere in the heterostructure.

The motivation for making such a structure is the high mobility of electrons in the GaAs region just after the AlGaAs region. The n-doped AlGaAs region contributes electrons to the undoped GaAs region where electrons move faster. In the AlGaAs region the dopants that contribute the electrons would have themselves slowed them down through scattering. The electrons would thus travel in the well region (discussed in the next section) in GaAs. Across the different layers of the heterostructures, no current flows. A voltage applied at $x = 0$ modifies the conductivity of the electrons in the well region. The tool this presents an opportunity to study and design similar devices and improve on the performance.

3.2.2 FORMATION OF THE WELL AND COMPARISON WITH ANALYTICAL SOLUTIONS

This device is then taken through the process of Hamiltonian construction, charge calculation and self-consistency (Poisson equation) to arrive at the band edge profile as shown in fig. 3.3 and the charge as shown in fig. 3.4. This profile is a consequence
Figure 3.3. The band profile of the default structure with electrostatics determined by a charge self-consistent calculation.

Figure 3.4. The quantum mechanical charge profile of the default structure with electrostatics determined by a charge self-consistent calculation.
of what doping and charge is present in the structure. The states formed in the well can be seen. That they are enclosed by the AlGaAs barrier on the left and the bent GaAs potential profile on the right makes these states confined or bound states. The energy of these states corresponds to the eigenvalues of the Hamiltonian.

The shape of the well is quite similar to that of a triangular well. Analytical solutions to the Schrödinger equation exist for the case of a triangular potential. They are the Airy functions. The analytical solutions of the ground and first excited state for a triangular potential well in GaAs with a $10^6 \, V/m$ electric field are shown in fig. 3.5 and fig. 3.6. As had been mentioned before, to calculate the charge, the Hamiltonian of the system has to be solved for the eigenvalues and eigenvectors. The magnitude squared of these eigenvectors is shown in fig. 3.7, with the position on the energy scale being given by the corresponding eigenvalue. The variation in these eigenvectors matches well with Airy functions. One must account for the fact that the well is not perfectly triangular in shape. Analytical solutions to the Schrödinger equation exist only for a few cases. Hence there is the need to solve a general problem.
Figure 3.6. Analytically computed wave-function magnitude squared of the first excited state of a GaAs triangular potential well with a $10^6 \, V/m$ electric field.

Figure 3.7. The wave-function magnitude squared of the eigen-energies found in the triangular well of the default structure.
numerically in order to get an accurate picture of how the eigenvalues and eigenvectors are distributed in a well with an arbitrary shape.

3.2.3 CHARGE PROFILE AND THE FERMI-LEVEL

Fig. 3.4 shows the charge calculated using the eigenvalues and eigenvectors found. The key point to note here is that finding the charge is an iterative process, since it has to be done such that it is consistent with the electrostatic potential which itself is used in the calculation of the charge. Therefore the charge self-consistency proceeds until there is agreement between the charge and the electrostatic potential. Though the charge profile is a sum total of contributions from all the states distributed in energy and as such should have the spatial effect of all eigenvectors, the contributions from different states might be different. The charge profile in this case resembles the eigenvector corresponding to the lowest eigenvalue in the well. This is because the lowest state is closest to and above the Fermi level which in this case had been set around the mid-gap region of the semi-conductor. The carrier statistics which give the contribution at every energy are governed by the Fermi distribution. States above the Fermi level will give higher contributions if they are closer to it. Thus the charge profile has the maximum contribution from the lowest state in the well, with a decreasing contribution from states higher up.

3.2.4 EFFECT OF DELTA DOPED LAYERS

A modification is now introduced into the default structure that has been discussed until now. A delta doped layer is introduced in the center of the AlGaAs layer with a density of $10^{12}/cm^2$. Delta-doping is a common technique that introduces a small region with a high doping in the device. The advantages of this technique can be understood by considering that ionized impurity scattering is reduced by an ordered array of dopants. Charges can be transported through the structure with reduced ionized impurity scattering using delta doping. This technique can also be
Figure 3.8. The band profile of the default structure modified to include a delta doped layer (Density=$10^{12}/cm^2$) in the center of the AlGaAs layer with electrostatics determined by a charge self-consistent calculation.
viewed as a way to get improved control over device performance by modifying the band profile. The band profile of the device with the delta doped layer is shown in fig. 3.8. The difference from fig. 3.3 is evident in the amount of band bending in the AlGaAs region at the point of the delta doping. From a numerics points of view, delta doping presents a problem as far as achieving convergence with the Poisson solver is concerned. Sometimes the doping is smeared over a small length rather than being present at a single point to achieve an acceptable convergence.

The real purpose of the delta doping can be demonstrated using the data shown in fig. 3.9. The variation of the lowest eigenvalue in the well with the voltage applied on the left side (0 nm) with and without the delta doped layer is shown here. The delta doped layer has the tendency to isolate the well region from the site where the voltage has been applied thereby reducing the effect of the applied bias on any feature in and beyond the well region. For the default case we see a much greater lowering of the states in the well compared to the case when the delta doped layer is present. This suggest that it is possible to tune the effect of the applied voltage on the states in the well region by using a delta doped layer of appropriate density, in the structure.
3.2.5 COMPARISON WITH SEMI-CLASSICAL CHARGE

As had been mentioned before classical mechanics predicts that in energy charge should be distributed continuously because the states holding it are continuous. This is in sharp contrast with what is suggested by quantum mechanics, discrete states present at specific energies in regions that form a potential well. Nevertheless it is instructive to study how a calculation that is actually semi-classical in nature compares with a quantum mechanical calculation. Charge is again calculated using the default structure, but this time semi-classically using the Thomas-Fermi formulation where an effective density of states is assumed. The requirement of the charge being calculated self-consistently with the electrostatic potential remains.

As can be seen in fig. 3.10, the results of such a calculation suggest that the potential profile is not too different from the case when charge was calculated quantum mechanically. The band bending and the formation of the well are in good agreement with fig. 3.4. However the charge profile differs from the quantum mechanical case. The peak of the charge is almost double in the semi-classical case compared to the
value in the quantum mechanical case. Also the peak in fig. 3.10 is spiked compared to the rounded peak in fig. 3.4. The charge profile actually begins to decrease inside the well region close the barrier for a quantum mechanical case and might still be non-zero inside the AlGaAs barrier for a small distance. In the semi-classical case the changes are more abrupt and coincide with the interface of the well and the AlGaAs barrier. In the semi-classical calculation of charge, the states are not only continuous but also spread above the well. Though the effective mass used to construct the Hamiltonian and calculate the semi-classical charge is the same, it turns out that the semi-classical charge is higher than the quantum mechanical charge.

3.2.6 EFFECT OF THE FERMI LEVEL ON CHARGE

For a device to have charge it must provide a favorable distribution of states that have the capability to hold charge and an equally favorable carrier distribution which ascertains which of these states is filled. The carrier distribution in such devices is governed by the Fermi distribution. Therefore a critical part of device design is the position of the Fermi level. Though the Fermi level in equilibrium conditions is governed by the doping in the device, in the tool the specification of the Fermi level has been kept independent of any other device detail. The logic behind this is the following. As had been mentioned before closed boundary conditions are applied to the Hamiltonian of the one dimensional device. This means that no current flows across the layers of the heterostructure that make up the device in that dimension. Instead for an actual device the current flow is perpendicular to this dimension, something that is not included in the Hamiltonian. The Fermi level would be determined by details of the device and phenomena in the dimension perpendicular to the one modelled here. At this point it must be mentioned that the usefulness of modeling a closed device in which no current flows, remains, as the details of the band profile in that dimension will determine current flow in the dimension perpendicular to it.
Figure 3.11. The sheet density in the well region for the default structure as a function of bias for different Fermi levels.
To see what difference the Fermi level has on the charge we analyse the charge in the device for different Fermi levels. As can be seen in fig. 3.11 the sheet density in the well region is shown as a function of applied bias for different Fermi levels. The sheet density increases with bias for a given Fermi level because the well is pushed lower and it gets closer and closer to the Fermi level. At a given bias the higher the Fermi level higher the sheet density because of its proximity to the states in the well. A relatively flat region can be seen around 0.15 V for $E_F = 0.824 \, eV$. This is because of the formation of another well region on the left side of the AlGaAs barrier at around 20 nm. The Poisson equation tries to balance the total charge in the device. Since an additional well is available which can hold some charge, the charge in the well on the right does not increase as much as the other cases.

### 3.3 TOOL LIMITATIONS

While a lot of effort is made to understand the capabilities of a tool, it is equally important to understand the limitations both in the theory supporting the tool and in the numerics that are used. It is also important to realize that for many structures which from a fabrication perspective are unrealistic, it might still be possible to get final converged results. The meaning of these results should be interpreted carefully, keeping in mind their physical implications and correctness.

Perhaps the most important consideration in this tool is the convergence of the electrostatic potential. Care should be taken to assess whether each bias point has converged to an acceptable value. Unusually high doping concentrations are mainly responsible for the convergence failing. In certain cases even though convergence is achieved the results do not make sense, simply because of excessive band bending. As shown in fig. 3.12, where the default structure was modified to include a very high doping of $10^{19}/cc$ in the AlGaAs region, the band bending is comparable to the band gap.
Figure 3.12. The band profile of the default structure with a very high doping $10^{19}/cc$ in the AlGaAs region. Though the tool shows the band profile as converged it is physically incorrect.
Further, it must be noted that the Hamiltonian is written only for the electronic charge, hole are treated semi-classically. Structures in which holes play a dominant role should not be modeled with this tool. Another important control in the tool is the resonance energy range which effectively controls the search space for resonance about the Fermi level. If resonances are far away from the Fermi level in an intermediate self-consistent iteration then convergence might be hard to achieve.

### 3.4 ALGORITHMIC DETAILS

Like boundary conditions are imposed on the Schrödinger equation, which is essential to calculate the correct charge profile, it is also necessary to impose boundary conditions on the Poisson equation to calculate the correct electrostatic potential. The boundary conditions could be Neumann, which means that the value of the derivative of the potential is specified or Dirichlet which specify the value of the electrostatic potential. For the terminal on the left side (0 nm) a voltage is directly applied to the device. It follows that Dirichlet boundary conditions are needed here.

\[ V(x = 0) = - \text{Applied bias} \quad (3.6) \]

On the right hand side however the device extends into the substrate which might or might not be part of the Hamiltonian. It is assumed that the electrostatic potential becomes flat deep into the substrate since there is no variation in material properties or the geometry of the device. Therefore the Neumann boundary conditions of setting the electric field to zero is,

\[ \frac{dV}{dx} \bigg|_{x_{\text{max}}} = 0 \quad (3.7) \]

However a Dirichlet boundary condition can also be applied which is based on the doping of the substrate. It is possible to calculate what the exact potential would be, given the doping concentration and the effective mass of the material that forms the substrate.

\[ V(x_{\text{max}}) = E_F + k_B T \log\left(\frac{N_A}{N_V}\right) \quad (3.8) \]
where $N_A$ is the acceptor doping and $N_V$ is the usual effective density of states in the valence band.

The methods to solve the Poisson equation have been discussed in detail in [56]. The Newton’s method can be used which uses the Newton-Raphson method to find the solution to an equation. The method can be understood as follows. In the finite difference scheme the Poisson equation can be written as,

$$\frac{(V_{i+1} - 2V_i + V_{i-1})}{\delta^2} + \frac{\rho(V_i)}{\epsilon} = 0 \quad (3.9)$$

Using the idea of the Newton-Raphson method, the update in the potential $\Delta V$ can be related as,

$$\frac{(-\Delta V_{i+1} + 2\Delta V_i - \Delta V_{i-1})}{\delta^2} - \frac{1}{\epsilon \, \partial V_i} \frac{\partial \rho}{\partial V_i} \Delta V_i = \frac{(V_{i+1} - 2V_i + V_{i-1})}{\delta^2} + \frac{\rho(V_i)}{\epsilon} \quad (3.10)$$

By defining the right hand side as $R_i$ and,

$$M_i = \frac{2}{\delta^2} - \frac{1}{\epsilon \, \partial V_i} \frac{\partial \rho}{\partial V_i} \quad (3.11)$$

in matrix form this can be expressed as,

$$\begin{bmatrix}
M_1 & -\frac{1}{\delta^2} & & & & & \\
-\frac{1}{\delta^2} & M_2 & -\frac{1}{\delta^2} & & & & \\
-\frac{1}{\delta^2} & M_3 & -\frac{1}{\delta^2} & & & & \\
& & \ddots & \ddots & \ddots & & \\
& & & -\frac{1}{\delta^2} & M_{N-1} & -\frac{1}{\delta^2} & \\
& & & & -\frac{1}{\delta^2} & M_N & \\
\end{bmatrix}
\begin{bmatrix}
\Delta V_1 \\
\Delta V_2 \\
\Delta V_3 \\
\vdots \\
\Delta V_{N-1} \\
\Delta V_N \\
\end{bmatrix}
= 
\begin{bmatrix}
R_1 \\
R_2 \\
R_3 \\
\vdots \\
R_{N-1} \\
R_N \\
\end{bmatrix}
\quad (3.12)$$

Factorisation of this tri-diagonal matrix is used to solve the system of equations and generate an update $\Delta V$ at each location [57]. The Newton’s method converges fast only if the initial guess is close enough to the true solution.
4. DESIGN GUIDELINES FOR TRUE GREEN LEDs AND HIGH EFFICIENCY PHOTOVOLTAICS USING ZnSe/GaAs DIGITAL ALLOYS

4.1 INTRODUCTION AND APPROACH

In the fields of solid state lighting and high efficiency solar photovoltaics (PV), a need still exists for a material system that can target the 2.3-2.5eV energy range. The ZnSe/GaAs system is shown to have great potential. The digital alloy approach can be utilized as a well-ordered design alternative to the disordered alloyed systems. The effective band-gap of the ZnSe/GaAs(001) superlattice has been studied, as a function of the constituent monolayers using tight binding. The possibility of engineering a range of band-gaps with the same material system, to achieve the optimum value for solar PV and LED applications, has been proposed.

Why ZnSe/GaAs? The solar cell story. Currently, no material system is well tuned for the conversion of high energy photons in the range of 2.3 to 2.5eV, as depicted in Fig.4.1. In the field of multijunction, or tandem stack, solar cell design, this high energy range is of crucial importance for reaching a combined cell efficiency greater than 50%. In particular, a material system with an energy gap of 2.4eV would be ideal for the top-most cell in a vertically integrated multijunction stack [58,59]. Additionally, the need still exists for a highly efficient true green light emitting diode (LED) in the wavelength range of 555 to 560nm. InGaN is able to achieve high brightness in green/blue-green LEDs (around 532nm), while GaP and AlGaInP

---

work best in the yellow-green range (around 567nm), leaving a gap in between (the so-called ”green gap”). Most work being conducted in this area is with InGaN. However, indium-rich compositions suffer severely from phase separation of InN [60], resulting in detuning from the desired spectral range. In addition, all existing InGaN growth processes result in defect densities that are too high for efficient solar cells. As an alternative, the system between GaAs and ZnSe is particularly appealing for investigation. Given that they are both direct band gap and lattice matched (within 0.27%) semiconductors, the possibility exists of engineering materials both in physical and/or digital alloy form over the full range of band gaps from 1.42eV (GaAs) to 2.7eV (ZnSe). Shen et. al have previously shown calculations on this superlattice (SL) system, though no details were given on our targeted range of band gaps [61]. Our work seeks to guide the experimentalist to engineer ZnSe/GaAs digital alloys (DAs) based on barrier and well thickness.

**Digital Alloys vs. Physical Alloys** From a materials engineering perspective, there are two possible methods for fabricating materials with band gaps between that of GaAs and ZnSe: 1) physical alloying and 2) digital alloying using a SL. It might be indicated that the DA technique has an advantage because of the fact that the density of states for a quantum well like structure has a stair-case form which translates into a non-zero value of density of states even at the minimum(maximum) energy for the conduction(valence) band [62]. In this paper, we investigate the DA technique [63] to provide a well-ordered design alternative to the disordered physical alloyed systems. Due to their miscibility, ZnSe and GaAs could be formed as a physical quaternary alloy. However, the use of a DA is preferable given the heterovalency of the ZnSe/GaAs alloy. As zinc and selenium will each dope GaAs, as well as the reverse case with gallium and arsenic in ZnSe being true, maintaining control on doping densities in the quaternary would be difficult. However, given that one can maintain control in doping pure ZnSe and GaAs, the ability to fabricate doped DAs of ZnSe/GaAs is possible. Given that ZnSe is lattice matched to GaAs, periodic
structures on the monolayer scale can be grown using molecular beam epitaxy (MBE). Previous work by Qian et al has shown the ability to create ZnSe/GaAs interfaces with low defect densities [64]. Additionally, Kobayashi et al have demonstrated ZnSe/GaAs SLs grown by migration-enhanced epitaxy (MEE) as an alternative growth method [65,66]. We do note, however, that growth of GaAs on ZnSe does present a difficulty due to growth temperatures, but experiments can be done using low-temperature-grown GaAs [67].

**Device Descriptions**  By utilizing the DA technique for device engineering, one has more freedom in narrowing in on a specific need, such as optimal light absorption
in a solar cell or tuned light emission in an LED. Fig.4.2 shows each of these devices in configurations that could be used for incorporation of the ZnSe/GaAs SL in each. These designs are by no means intended to be novel, but provide a basis from which to build upon. For use in a solar cell, a p-i-n structure would be well suited where primary absorption takes place in the intrinsic SL region. Ideally, one would expect to incorporate the necessary number of SL periods to absorb an adequate number of photons (99%). In pure ZnSe this is around 500nm [68] and around 1500nm for pure GaAs [69]. The experimentalist could estimate from these what may be acceptable for the ZnSe/GaAs SL, taking into account the greater difficulty and time involved in growing thicker SL layers. From a cost perspective, the increased cost to fabricate such SL-based solar cells may turn out to be worthwhile in the scheme of single or dual-axis concentrator systems. For an LED, a similar design could be used except for the desire to have doped SL layers on either side of an intrinsic SL layer. Making the intrinsic layer very thin should help promote carrier confinement for increasing radiative recombination. A ZnSe substrate would need to be used for light transmission through the bottom, or a GaAs substrate could be used for growth and then etched off. The p ZnSe needs to be contacted directly due to the insulating nature of ZnSe substrates.

**Need for atomistic models:** The need to capture properties at the atomic level for device detail that varies at the nanometer level is often not fully appreciated. It has been recognized for several years now that effective mass models cannot treat band non-parabolicities properly. From the point of view of the SL it is imperative that a quantum mechanical approach that includes confinement effects is used. Effects of band non-parabolicities, material variations and confinement are readily captured in multi-band tight binding models. In this paper we employ the $sp^3s^*_{\text{nearest neighbor}}$ tight binding scheme to model the SL. The scheme is capable of replicating both the conduction and valence bands close to the Γ valley. Since X-minima are not in the energy range of interest we do not need to model them extremely accurately and we
therefore do not include 'd' orbitals or second nearest neighbor interactions in this tight binding approach.

4.2 METHODOLOGY

Hamiltonian Construction and Boundary Conditions: The nearest neighbor tight binding Hamiltonian is constructed with the well region (GaAs) surrounded by the barrier region (ZnSe) on either side. Further periodic boundary conditions are incorporated to repeat the structure indefinitely.

Supplementary topic: Parameter sets and Band offset treatment: Based on the experimental data for ZnSe [70–72], the effective masses in the seminal work of Vogl et al. [48] turn out to be inaccurate. In addition to that the Vogl parameters fit the low temperature gaps whereas we need to model room-temperature gaps. Finally the Vogl [48] parameter sets do not incorporate spin-orbit coupling which is essential.
Table 4.1

Nearest neighbor \( sp^3s^* \) ZnSe parameters (eV). Labels denote matrix elements. E: onsite, V: coupling, \( \lambda \): spin-orbit coupling. Subscripts denote the basis. s,p,x,y: orbital symmetries. a,c: anion or cation [48]

| \( E_{sa} \) | -12.6921 | \( V_{s,s} \) | -6.3967 |
| \( E_{pa} \) | 1.5072  | \( V_{x,x} \) | 3.1784 |
| \( E_{sc} \) | 0.0183  | \( V_{x,y} \) | 5.3489 |
| \( E_{pc} \) | 6.0298  | \( V_{sa,pc} \) | 3.498 |
| \( E_{ssa} \) | 7.5872  | \( V_{sc,pa} \) | 7.3683 |
| \( E_{ssc} \) | 8.9928  | \( V_{ssa,pc} \) | 2.5891 |
| \( \lambda_a \) | 0.16    | \( V_{pa,ssc} \) | 3.9533 |
| \( \lambda_c \) | 0.03    |                     |        |

to model the imaginary band linking the light-hole and conduction band [49]. We re-parametrize the Vogl ZnSe parameter set [Table 4.1] based on the Landolt-Bornstein tables using the analytical expressions for effective masses in Ref. [49]. For GaAs the \( sp^3s^* \) parameters from Boykin et al. [49] are used. These parameter sets assume the valence band for both ZnSe and GaAs to be at 0eV. Various authors have reported the experimentally measured Valence band offset (VBO) and Conduction Band offset (CBO) for the ZnSe/GaAs(001) and ZnSe/GaAs(110) SLs. Raman Scattering [73], Electrical [74], Optical [75], and XPS [76,77] measurements put the VBO at 0.9-1.1eV for ZnSe/GaAs(110) and 0.7-0.9eV for ZnSe/GaAs(001) hetero-junctions. We assume it to be close to 0.96eV [78]. This leads us to the band diagram as depicted in Fig. 4.3. We have also verified that by using the Vogl parameters for both ZnSe/GaAs, similar results are obtained, though electron and hole states have to be treated separately to get the correct band offsets.
4.3 RESULTS AND DISCUSSION

**Effective Conduction band of the SL:** The Hamiltonian with the correct conduction band offset is constructed such that the ZnSe and GaAs conduction band edges are at 2.68eV and 2.38eV (1.42eV+0.96eV) respectively. The ground state eigenvalues are found for varying SL periods. As is expected, in Fig. 4.4, when the ZnSe(GaAs) content is much larger than the GaAs(ZnSe) content, the ground state energy approaches the bulk value of 2.68eV(2.38eV) as shown in Fig.4.3. For a given thickness of ZnSe as the thickness of GaAs is decreased, an increase in the ground
state energy can be seen. This trend is attributed to confinement effects. Alternatively, for a given thickness of GaAs, the ground state energy starts from the bulk value of GaAs(2.38eV) for low ZnSe content and gradually increases until the thickness of ZnSe is enough such that the eigen-states in adjoining quantum wells have no significant effects on each other. This increase can be understood in terms of the coupled quantum well picture. If the coupling between two quantum wells is reduced by increasing the thickness of the barrier material, the energy of the lower bonding state increases and that of the higher anti-bonding state decreases. The plot for the effective valence bands shows equivalent trends and can be explained using similar arguments.

Figure 4.4. Ground state energy in eV of the states formed in the conduction band well for different monolayers of ZnSe and GaAs.
Effective band gap of the SL: Finally in Fig. 4.5 the effective band gap of the SL is estimated as the difference in the lowest conduction band and the highest valence band state for different SL periods. The tendency to approach the appropriate bulk value is evident again when the percentage of one material is much greater than the other. We recognize that as the thickness of the well regions becomes smaller, the effects from the (Ga, Se) compounds at the interfaces [79] may affect these calculations, though including such effects is outside the scope of this work.
4.4 CONCLUSION

Range of band gaps can be fabricated: We have used the tight-binding technique to predict the effective band gap of the ZnSe/GaAs SL. The calculation potentially paves the way for obtaining a range of band gaps from 1.7-2.5\,eV in the same material system using DAs. A band gap slightly greater than 2.4\,eV can be achieved with 20 monolayers of ZnSe interspersed with 2 monolayers of GaAs. This will enable solar cell to target the spectrum around 2.3-2.5\,eV and thus increase efficiency. Though experimental measurements are required to verify the accuracy of these numbers, the variations and trends discussed here will provide a guideline.

4.5 SUPPLEMENTARY TOPIC: PARAMETER FITTING

It has been mentioned that the effective masses in the seminal work of Vogl et al. [48] are inaccurate on the basis of experimental data [70–72]. The Vogl parameters fit low temperature gaps, whereas our requirement is of room temperature gaps. Finally the addition of spin-orbit coupling is essential to model the imaginary band linking the light-hole and the conduction band, which is not present in the Vogl parameter set.

A least square fitting approach was used to calculate the new parameter set for ZnSe in Table 4.1. The experimental data from Ref. [70–72] as listed in Table 4.2 were used as targets and the $sp^3s^*$ parameters from Vogl et al. [48] were used as the starting guess. The effective masses and band edges were calculated analytically based on the analytical expression in Ref. [49].
Table 4.2

Experimental data for ZnSe and the achieved values using the re-parametrized nearest neighbor tight-binding parameters.

<table>
<thead>
<tr>
<th>Targets</th>
<th>Landolt-Bornstein</th>
<th>Re-parameterized Set</th>
<th>% Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electron mass</td>
<td>0.16</td>
<td>0.1478</td>
<td>7.6</td>
</tr>
<tr>
<td>Light hole mass</td>
<td>-0.145</td>
<td>-0.1591</td>
<td>9.7</td>
</tr>
<tr>
<td>Heavy hole mass</td>
<td>-0.78</td>
<td>-0.78</td>
<td>0</td>
</tr>
<tr>
<td>CB minima (Γ)</td>
<td>2.68 eV</td>
<td>2.68 eV</td>
<td>0</td>
</tr>
<tr>
<td>VB maxima (Γ)</td>
<td>0 eV</td>
<td>0 eV</td>
<td>0</td>
</tr>
<tr>
<td>CB minima (X)</td>
<td>4.54 eV</td>
<td>4.54 eV</td>
<td>0</td>
</tr>
<tr>
<td>CB minima (L)</td>
<td>3.96 eV</td>
<td>4.11 eV</td>
<td>3.8</td>
</tr>
</tbody>
</table>
5. LEAKAGE REDUCTION DESIGN CONCEPTS FOR LOW POWER VERTICAL TUNNELING FIELD-EFFECT TRANSISTORS

Using an atomistic full-band quantum transport solver, we investigate the performances of vertical band-to-band tunneling FETs (TFETs) whose operation is based on the enhancement of the gate induced drain leakage (GIDL) mechanism of MOSFETs and we compare them to lateral $p$-$i$-$n$ devices. Although the vertical TFETs offer larger tunneling areas and therefore larger ON-currents than their lateral counterparts, they suffer from lateral source-to-drain tunneling leakage away from the gate contact. We propose a design improvement to reduce the OFF-current of the vertical TFETs, maintain large ON-currents, and provide steep subthreshold slopes\(^1\).

5.1 INTRODUCTION

Band-to-band TFETs are emerging as an attractive alternative to MOSFETs to reduce the power consumption of integrated circuits. While the subthreshold slope ($SS$) of conventional MOSFETs is fundamentally limited to 60 mV/dec at room temperature, in TFETs, the cold injection of valence band (VB) electrons from a source contact into the conduction band (CB) of a drain contact does not impose any lower limit on the $SS$. Though $SS$ lower than 60 mV/dec have been demonstrated for lateral TFETs based on carbon nanotube [80], silicon [81], and strained germanium [82] devices, all these approaches suffer from low-ON currents, typically a few $\mu$A/$\mu$m, when more than 1,000 $\mu$A/$\mu$m are required [83].

Atomistic full-band transport simulations indicate that lateral \textit{p-i-n} TFETs made of a homogeneous material are not capable to provide large enough ON-currents, even with a low direct band gap such as InSb and a perfect electrostatic control of the channel as gate-all-around nanowires \cite{84,85}. The ON-current limitation of lateral TFETs originates from their small tunneling area limited to the device cross section. By switching the tunneling mechanism to a vertical approach characterized by an enhancement of the gate induced drain leakage (GIDL) present in MOSFETs, it is possible to increase the tunneling area of TFETs proportionally to their gate length and to obtain large ON-currents. Such a vertical TFET design has been recently proposed \cite{2,86,87} and named “green FET” (gFET) for its potential to reduce the supply voltage of transistors.

Using the atomistic simulation approach described in Section 5.2, we show in this paper that the gFETs offer larger ON-currents than lateral devices, but suffer from lateral source-to-drain tunneling leakage away from the gate contact. This effect has never been analyzed before although it increases the gFET OFF-current by several orders of magnitude and makes any steep $SS$ impossible. We therefore propose a design modification to suppress the source-to-drain tunneling leakage in vertical TFETs without affecting their ON-current. We further demonstrate using an InAs device that very good performances can be obtained by optimizing the gFET structure.

5.2 SIMULATION APPROACH AND RESULTS

The InAs lateral and vertical TFETs considered in this work are simulated using an atomistic, full-band quantum transport solver based on the $sp^3s^*$ nearest-neighbor tight-binding method and a wave function approach \cite{11,84}. Transport is treated in the ballistic limit and electron-phonon scattering is not included. The InAs tight-binding parameters are taken from Ref. \cite{88}. For computational reasons, spin-orbit
coupling is neglected, which does not alter our conclusions [84]. Gate leakage currents are also neglected in this work.

The tight-binding model accurately and simultaneously describes the conduction and the valence bands of most semiconductor materials, as well as the imaginary bands coupling them which is responsible for tunneling. Contrary to standard TFET simulators where tunneling is treated as a perturbation in the WKB approximation [89] and only exists in pre-defined regions, our approach is characterized by a global tunneling model (tunneling is present everywhere by default). Hence, it is not necessary to specify the regions where tunneling is expected to take place and no crucial tunneling path is omitted.

TFETs are usually based on lateral band-to-band tunneling of electrons from a $p^+$ source into a $n^+$ drain as shown in Fig. 5.1(a). An increase of the gate voltage pushes down the CB edge of the device channel below the VB edge of the source and opens a bias-dependent tunneling window at the source/channel interface as illustrated in Fig. 5.1(b). The fact that the electrostatic control of the channel diminishes deep into the device body and does not go beyond a few nano-meters puts a restriction on the cross-sectional area available for tunneling and on the highest achievable ON-current.

The vertically tunneling gFET structure, as depicted in Fig. 5.1(c), keeps the $p^+$ source and $n^+$ drain of the lateral device, but the source extends close to the drain contact and both regions are separated only by a lightly $n$-doped layer. A thin and highly $n$-doped region called “pocket” is implanted just below the gate contact in the $p^+$ extended source. Tunneling occurs under the gate region, between the VB electrons of the $p^+$ extended source and the available confined CB states of the $n^+$ pocket as shown in Fig. 5.1(d). A gate voltage increase moves down the CB edge of the pocket below the VB edge of the extended source and opens a vertical tunneling channel whose area is proportional to the gate length.

Fig. 5.2(a) compares the transfer characteristics $I_d-V_{gs}$ at $V_{ds}=0.5$ V of a single-gate InAs lateral $p-i-n$ TFET and a single-gate InAs gFET. Both devices have the same gate length $L_g=40$ nm, equivalent oxide thickness EOT=0.5 nm, supply voltage
Figure 5.1. (a) Lateral p-i-n single-gate TFET. (b) Conduction and valence band edges of the p-i-n device in its OFF- (dashed lines) and ON- (solid lines) state along the horizontal dashed line in (a). (c) Single-gate gFET structure as proposed in Ref. [86]. (d) Conduction and valence band edges of the gFET in its OFF- and ON-state along the vertical dashed line in (c). The buried oxide extends from -50 nm to 0 nm, but only the 20 nm close to the semiconductor channel are shown here.

$V_{DD}=0.5$ V, source and drain doping ($N_A=2\times10^{19}\text{cm}^{-3}$, $N_D=2\times10^{19}\text{cm}^{-3}$). The lateral device has a body thickness of 5 nm to maintain a good electrostatic control of the channel while the gFET has $T_{body}=10$ nm, $T_{pock}=5$ nm, $N_{pock}=2\times10^{19}\text{cm}^{-3}$, and the n doping of the 5 nm layer between the source and the drain is $10^{18}\text{cm}^{-3}$.

As expected, the ON-current ($I_{ON}=I_d$ at $V_{gs}=V_{ds}=V_{DD}$) is much larger for the gFET (180 $\mu\text{A}/\mu\text{m}$) than for the p-i-n lateral TFET (14 $\mu\text{A}/\mu\text{m}$), but it is still
Figure 5.2. (a) Transfer characteristics $I_d-V_{gs}$ at $V_{ds}=0.5$ V of a lateral $p$-$i$-$n$ TFET (solid line) and a vertical gFET (line with symbols) with $L_g=40$ nm and EOT=0.5 nm. (b) Spatial distribution of the gFET ON-current. High current density regions appear darker. (c) Same as (b), but for the OFF-current. For clarity, a different color scale is used. (d) CB and VB edges of the gFET OFF-state on the horizontal line below the $n^+$ pocket at $y=4$ nm in (c).

well below the ITRS requirement. The spatial distribution of the gFET ON-current is given in Fig. 5.2(b). It can be observed that vertical band-to-band tunneling occurs between the extended source and the pocket, but the tunneling current is not homogeneously distributed and two main channels (arrows) can be distinguished. This is due to the profiles of the CB and VB edges that vary along the $x$-axis while they should ideally remain constant.
The detrimental characteristics of the gFET is its very high OFF-current \( I_{OFF} = I_d \) at \( V_{gs} = 0 \) V and \( V_{ds} = V_{DD} \) that has never been reported before \([2, 86, 87]\) and whose spatial distribution is shown in Fig. 5.2(c). A lateral leakage current flowing between the \( n^+ \) pocket and the buried oxide can be identified. Figure 5.2(d) describes the band diagram of the region where the OFF-current flows. It looks like a lateral \( p-i-n \) structure that the gate cannot control due to its distance. VB electrons laterally tunnels from the extended source into the drain contact and increase \( I_{OFF} \). The gFET as depicted in Fig. 5.1(c) is therefore unsuitable for low power logic applications.

Figure 5.3. (a) Modified gFET design. The buried oxide is raised to the height of the pocket in the drain and in the lightly doped channel region. (b) Comparison of the transfer characteristics \( I_d - V_{gs} \) at \( V_{ds} = 0.5 \) V for the “conventional” (line with symbols) and the modified gFET (solid line) design.

To address the large OFF-current issue, we propose a modification of the gFET design in Fig. 5.3(a). The spatial distribution of the OFF-current in Fig. 5.2(c) shows that the VB electrons originating from the extended source are collected in a portion of the drain contact that is not active in the ON-state plotted in Fig. 5.2(b). The buried oxide below the drain and the lightly \( n \)-doped channel is therefore pushed up to the height of the pocket in the drain and lightly \( n \)-doped channel such that it blocks any lateral tunneling current pathway. The transfer characteristics of the
“original” and modified gFETs are compared in Fig. 5.3(b), demonstrating that $I_{OFF}$ is reduced by several orders of magnitude without affecting $I_{ON}$.

Figure 5.4. Transfer characteristics $I_d-V_{gs}$ of different gFET designs. (a) Variation of the pocket length $L_{pocket}$ at a constant gate length $L_g=40$ nm. (b) Variation of the pocket doping $N_{pock}$. (c) Variation of the body thickness $T_{body}$. (d) Comparison of the optimized InAs gFET (line with circles) with the original gFET design (solid line) and the lateral $p$-$i$-$n$ TFET (dashed line).

The gFET ON-current and $SS$ must also be improved by optimizing the body thickness $T_{body}$, pocket length $L_{pocket}$, width $T_{pock}$, and doping $N_{pock}$ in Fig. 5.3(a). Figure 5.4(a) shows the effect of the pocket length $L_{pocket}$ on the gFET $I_d-V_{gs}$. A decrease in $L_{pocket}$ and increase in $L_{td}$ ($L_g=L_{pocket}+L_{td}$ remains constant) leads to a further reduction of the gFET OFF-current and of the $SS$. Although the raising
of the buried oxide below the drain contact suppressed most of the source-to-drain leakage currents, some diagonal paths joining the extended source and the drain contact persisted. Increasing \(L_{ld}\) makes these tunneling paths longer and reduces \(I_{OFF}\) and \(SS\). However, the tunneling ON-current, which is roughly proportional to \(L_{pocket}\), slightly decreases (about 10% reduction for \(L_{pocket}\) going from 35 to 25 nm).

Increasing the pocket doping \(N_{pock}\), as described in Fig. 5.4(b), induces larger vertical electric fields so that the tunneling length between the \(p^+\) extended source and the \(n^+\) pocket becomes smaller, the current magnitude larger, and the \(SS\) steeper. Hence, \(SS\) reduces here from 118 mV/dec for \(N_{pock}=10^{19}\text{cm}^{-3}\) to 47 mV/dec for \(N_{pock}=4\times10^{19}\text{cm}^{-3}\) while \(I_{ON}\) increases from 143 \(\mu\text{A}/\mu\text{m}\) to 220 \(\mu\text{A}/\mu\text{m}\). Similarly, a reduction \(T_{pock}\) improves \(I_{ON}\) and \(SS\).

The influence of the body thickness is investigated in Fig. 5.4(c). As can be seen from Fig. 5.1(d), an increase of the gate voltage not only pushes down the bands close to the gate contact, in the pocket region, as desired (energy shift labeled \(\Delta V_p\)), but also deeper in the \(p^+\) extended source (\(\Delta V_s\)). Of course, \(\Delta V_s < \Delta V_p\) since the gate control decreases as function of the distance, but ideally, \(\Delta V_s\) should be as small as possible to maximize the electric field between the pocket and the extended source and the tunneling current. Increasing \(T_{body}\) from 10 to 15 nm helps reduce \(\Delta V_s\) and increase \(I_{ON}\). A further increase to 20 nm does not improve \(I_{ON}\).

Finally, Fig. 5.4(d) shows the best transfer characteristics \(I_d-V_{gs}\) at \(V_{ds}=0.5\text{V}\) that could be obtained by optimizing the InAs gFET structure (\(T_{body}=15\text{nm}, L_{pocket}=25\text{nm}, T_{pock}=3\text{nm}, N_{pock}=4\times10^{19}\text{cm}^{-3}, L_g=40\text{nm},\) all the other parameters remain the same). An intrinsic \(I_{ON}=667\ \mu\text{A}/\mu\text{m}\) (no contact series resistance), \(I_{OFF}=1\times10^{-3}\ \mu\text{A}/\mu\text{m},\) and \(SS=36\ \text{mV/dec}\) are reported for the optimized gFET, much better than the original gFET design, and the lateral \(p-i-n\) TFET.
5.3 CONCLUSION

We have investigated InAs vertical TFETs known as gFETs and compared them to lateral p-i-n devices. Simulation results suggest that the gFET gives a significant gain in the ON-current over the lateral tunneling geometries due to an increase of the tunneling area. Still the original gFET design is unsuitable for low power logic applications because of very high OFF-currents caused by lateral source-to-drain tunneling. A structure modification and an optimization of the device parameters allow for a reduction of $I_{OFF}$ by several order of magnitudes, an increase of $I_{ON}$ close to the ITRS requirement, and a $SS$ below the 60 mV/dec limit of MOSFETs.

5.4 SUPPLEMENTARY TOPIC: ACTIVE ENERGY RANGE IN THE gFET AND THE MODIFIED DESIGN

The purpose of the design modification to the gFET was to alleviate the problem of leakage current due to lateral tunneling in its OFF state. Though this modification does demonstrate that the OFF current is reduced by several orders of magnitude while leaving the ON current virtually unaffected, it should not be expected that the active region in energy remains the same either for the OFF state or the ON state.

Fig. 5.5 shows the transmission (solid lines) for the gFET and the modified design in the OFF state. The dotted lines denote the quantity $T(f_L - f_R)$ for each device. The quantity $T(f_L - f_R)$ begins to show a significant reduction from the transmission at almost the same energy for the two designs. This suggests that the active energy range of the device is not very different and the primary difference in OFF state current comes about because of the difference in magnitude of the transmission.

Fig. 5.6 shows the transmission (solid lines) for the gFET and the modified design in the ON state. Since the reduction in the quantity $T(f_L - f_R)$ from the transmission comes about at the same energy, the active energy range for both the gFET and the modified gFET is similar. That the ON state currents are of the same magnitude can be seen from the fact that transmissions also remain at comparable values.
Figure 5.5. Transmission Vs. Energy for the gFET and the modified design in its OFF state corresponding to $V_{GS}=-0.1V$ in Fig. 5.3(b). The dotted lines denote the quantity $T(f_L - f_R)$ for the corresponding device.

In both the figures, 5.5 and 5.6 the energy range over which the transmission has an appreciable value is greater for the original gFET structure. Since the active energy range is primarily determined by the Fermi levels in the left and right contacts, any noticeable change in final current is because of the magnitude of the transmission.
Figure 5.6. Transmission Vs. Energy for the gFET and the modified design in its ON state corresponding to $V_{GS}=0.5\text{V}$ in Fig. 5.3(b). The dotted lines denote the quantity $T(f_L - f_R)$ for the corresponding device.
6. ADAPTIVE QUADRATURE FOR SHARPLY SPIKED INTEGRANDS

A new adaptive quadrature algorithm that places a greater emphasis on cost reduction while still maintaining an acceptable accuracy is demonstrated. The different needs of science and engineering applications are highlighted as the existing algorithms are shown to be inadequate. The performance of the new algorithm is compared with the well known adaptive Simpson, Gauss-Lobatto and Gauss-Kronrod methods. Finally, scenarios where the proposed algorithm outperforms the existing ones are discussed.

6.1 INTRODUCTION

With the increasing complexity of problems in science and engineering, more and more computational tasks require the use of massively parallel simulations running for hours to days. A reduction in the computational burden could translate into several hours saved and far lesser CPU time for parallel simulations. Adaptive quadrature is an essential technique to minimize the cost of integrating or simply resolving features in a function. When the cost of evaluating a function is high or if adaptive quadrature has to be used repetitively, the efficiency of the algorithm being employed becomes critical.

In the field of quantum transport it is often required to resolve sharp features in the transmission probability of a structure at different energies [3]. To convey the enormity of the problem, it is comparable to finding and integrating a Lorentzian with a full-width at half maximum (FWHM) of $10^{-10}$ eV over a 0.5 eV range. Numerical techniques other than adaptive quadrature have been developed [16,17] that attempt

---

to reduce the cost. Though these techniques are beneficial they do add a significant overhead to the computational task, since some cost is incurred by these techniques as well. In this regard any efficient adaptive quadrature method is desirable since the cost incurred by the adaptive quadrature method itself is in general minimal.

To understand why the existing adaptive quadrature methods are inadequate for such applications, the purpose and the intent behind these methods must be highlighted. Problems in computing and numerical analysis require high precision, possibly 7-8 digits. In engineering and scientific applications it is enough to have a lower precision since there are several input parameters to the simulation, that are not known to such high accuracies anyway. In device engineering a general procedure is to guide a design by many simulations that provide design trends. Absolute numbers with perfect accuracy are not required but rapid computational turn around is needed. To meet the requirements of such applications, accuracy must be traded for a reduction in cost. Further, some implementations of adaptive quadrature methods throw away function evaluations from previous iterations in the quest for accuracy, a scenario that is not desirable, given that it can be very costly to evaluate a function at every single point.

6.2 METHODOLOGY: 5 POINT ADAPTIVE SCHEME

Fig. 6.1 illustrates the procedure for the 5 point adaptive scheme for a Lorentzian placed asymmetrically in the range of integration. The Lorentzian is to be resolved and integrated as accurately as possible. After sampling the function on a widely spaced homogeneous grid, nodes are added in successive iterations. The addition of nodes in a given region, like other adaptive schemes is based on two estimates of the integral, one of which is more accurate than the other. In this case the estimates are $I_5$ and $I_3$ over a region which is typically only a small part of the entire range to be integrated over. $I_5$ and $I_3$ refer to the number of points used in evaluating the integral over a small subsection of the total range. These integrals can be performed
using the mid-point rule or by fitting a quadratic function to a set of 3 points. The scheme is not very sensitive to the quadrature rule used to find $I_5$ and $I_3$ since the addition of nodes is based on the relative error of the two and not the actual values. The scheme is implemented in the following manner,

- **Step 1:** The function to be integrated is sampled on a widely spaced homogeneous grid.

- **Step 2:** $I_5$ using 5 points and $I_3$ using 3 points are computed over any set of 5 successive nodes. Typically the 2nd and the 4th points in a given set of 5 points are neglected for the purpose of computing $I_3$.

- **Step 3:** If the following condition,

\[
|I_5 - I_3| / I_3 > \epsilon \text{ (convergence criterion)} \tag{6.1}
\]

is found to be true, then 4 new nodes are added in a region spanned by the original 5 successive nodes such that they bisect the 4 already existing intervals symmetrically. Otherwise no new nodes are added.

- **Step 4:** Steps 2-3 are repeated with all the nodes, till the convergence criterion is met over any set of 5 successive nodes.

The integration is done using the composite Simpson’s rule by breaking the range into regions that have the same difference in successive abscissas.

### 6.3 RESULTS AND DISCUSSION: ACCURACY Vs COST

In Fig. 6.2(a) the 5point adaptive scheme is used to resolve features in the transmission through a Resonant Tunneling diode [22] at different energies. Fig. 6.2(b) shows the distribution of nodes in energy which is indicative of the cost of resolving each features in the transmission. As is expected flatter portions of Fig. 6.2(b) correspond to sharper features in Fig. 6.2(a) since more nodes were required to resolve
Start with a Homogeneous Grid

Iteration 1: 4 pts added
Iteration 2: 4 pts added
Iteration 3: 4 pts added

(a)  (b)
(c)  (d)

Figure 6.1. Successive iterations of the adaptive refinement algorithm in the 5 point scheme are illustrated. Vertical solid lines denote the nodes at the beginning of each iteration. Vertical dotted lines denote nodes added at each iteration.

(a)  (b)

Figure 6.2. (a) Transmission as a function of energy for a Resonant tunneling device with thick barriers giving rise to very sharp features in transmission. Similar devices have been modeled in [22] (b) The corresponding distribution of nodes generated using the 5 point adaptive grid scheme running with $\epsilon = 0.1$. 
them. The 5 point adaptive grid scheme is running with $\epsilon = 0.1$ in Eqn. 6.1. A smaller $\epsilon$ would resolve the features even more accurately though it would increase the node count as well.

![Graph](image)

**Figure 6.3.** The relative error as a function of node count for a normalized Lorentzian with FWHM = $10^{-3}$. The different data points for a given algorithm were obtained by varying the convergence criterion ($\epsilon$) for the 5 point scheme and the error tolerance for the remaining algorithms.

The performance of the 5 point adaptive scheme is now compared with other algorithms keeping in mind that the ideal algorithm would be one whose accuracy scales well with an increasing node count, that is, an algorithms that can give low accuracies at a low node count and high accuracies at a high node count. The 5 point adaptive scheme is implemented in MATLAB and is compared with the inbuilt functions quad, quadl and quadgk which correspond to the Adaptive Simpson [90], Gauss-Lobatto [91] and Gauss-Kronrod [92] methods. The performance and implementation of these three algorithms has been discussed in Ref. [93]. A normalized Lorentzian is placed asymmetrically and integrated in the range $[0, 1]$. The relative
error for a Lorentzian with a given FWHM is found with the corresponding node count by varying the convergence criterion for the 5 point adaptive scheme and the error tolerance in the case of the remaining adaptive quadrature methods.

Fig. 6.3 shows how the 5 point adaptive scheme compares with the other algorithms for the case when the Lorentzian is not very sharp (FWHM = 10^{-3}). The performance of the 5 point adaptive scheme is similar to that of the adaptive Simpson’s method with relative errors close to 10^{-5} being achieved with 100 nodes. A relative error close to 1 indicates that the algorithm could not resolve the feature at all, which is the case for the Simpson’s and the Gauss-Lobatto methods around 15 nodes and for Gauss-Kronrod around 200 nodes. One point of distinction could be that the 5 point adaptive scheme does find the Lorentzian with even fewer nodes although with a lower accuracy. It can be argued that the three existing algorithms are missing data points in the low accuracy region (relative error between 10^{-3} and 10^{-5}) as they do not converge, but since the difference in cost between them and the 5 point adaptive scheme is not large enough the discrepancy may not be serious. However, a small node count can reduce the cost of subsequent calculations in terms of memory footprint as well as overall compute time. Accuracies of 10^{-3} may often be acceptable resulting in node reduction by a factor of 2 or more.

In Fig. 6.4, showing the performance for a Lorentzian of FWHM = 10^{-7}, the contrast between the 5 point adaptive scheme and the remaining methods is more evident. The 5 point adaptive scheme shows the ability to find the Lorentzian with a low accuracy (relative error between 10^{-3} and 10^{-6}) and a low node count. The remaining algorithms take larger node counts to find the Lorentzian, although to a higher accuracy. The adaptive Simpson’s, Gauss-Lobatto and Gauss-Kronrod do not have any data points for high relative errors since they do not converge with low node counts.

One advantage that the 5 point adaptive scheme has that at the end of each iteration it is possible to have an estimate of the integral of the function, something that is not possible in the remaining methods. It is also possible to build safeguards
that avoid getting caught in noisy integrands or that treat a certain part of the range differently than others. It has been verified that adding 4 nodes in Step 2 of the algorithm compared to adding only one is more efficient unless there is an analysis of the integrand over the 5 points in question.

### 6.4 CONCLUSION

In this paper the 5 point adaptive quadrature scheme was demonstrated, that placed a greater emphasis on cost reduction, while still maintaining acceptable accuracies. Such a technique was shown to be beneficial for scientific and engineering applications. Existing techniques were shown to be inadequate despite their higher accuracies because of the prohibitive costs involved. A performance comparison was
used to demonstrate the suitability of the 5 point adaptive scheme over the existing algorithms in applications where low accuracies suffice and evaluating functions at each node is expensive.

6.5 SUPPLEMENTARY TOPIC: ADDING 4 VS 1

The 5pt adaptive scheme introduces 4 nodes in a range spanned by 5 nodes if it is found that the 5pt integral and the 3pt integral are not close enough. Though the performance of this scheme compares well with the other more established methods in adaptive quadrature, it might be worthwhile to investigate the accuracy and performance of the 5pt adaptive scheme when less than 4pts are added. In this context we explore three different addition schemes for each region containing 5pts.

If the convergence criterion is not met in a range spanned by 5 nodes then the adaptive scheme will add points using the following rules,

• Add 4pts: Bisect each segment to add 4 new nodes.

• Add 1pt: Bisect the first segment and add only one node.

• Add 1pt: The convergence criterion is computed on the value of the log of the function instead of the function itself. Based on this only one node is added in the first segment.

The addition of 4 nodes based solely on the 5pt and 3pt integrals, without any analysis as to which node might have a greater impact on the overall accuracy can be thought of as a ”brute force” strategy. The motivation behind the other two addition schemes is to obtain an improvement in performance, though no further analysis beyond the 5pt and 3pt integrals is done there as well. We will look further at the performance and the correctness of each of these schemes.
6.5.1 ANALYSIS FOR A BROAD RESONANCE

Fig. 6.5 is a scatter plot showing where the nodes are added at each iteration for the three different addition schemes. A broad lorentzian centered at 0.3583 with FWHM=$10^{-3}$ is adaptively resolved. The 4pt scheme takes fewer iterations to achieve convergence. Both the 1pt and the 1pt log schemes take many more iterations. That the 4pt scheme has converged can be seen by the difference in successive nodes in Fig. 6.6, which is the spectral width of the feature that is being resolved. Both the 1pt and the 1pt log scheme have not converged since the difference in consecutive nodes goes on till $10^{-10}$ which was the stopping condition that was imposed. It can also be seen from the scatter plots in Fig. 6.5 and the energy difference in Fig. 6.6, that the addition of nodes is not quite symmetric for both the 1pt and 1pt log schemes. This is because the addition of nodes itself is not symmetric.
Figure 6.6. Difference in energy points for a broad resonance (FWHM=10$^{-3}$ centered at 0.3583). The 4pt and 1pt schemes are operating at a 10% convergence criterion and the 1pt log scheme is at 1%.

The performance of each of these schemes is compared in Fig. 6.7. The relative error is the deviation of the integrated value of a normalized lorentzian from its analytically calculated value. A very high relative error means that the adaptive grid was unable to resolve the lorentzian. The 4pt scheme gives a lower error at a lower node count compared to the other two schemes. The 1pt log scheme is the most expensive.

6.5.2 ANALYSIS FOR A SHARP RESONANCE

A similar analysis is repeated for a sharper lorentzian with FWHM=10$^{-10}$. In this case as well, the 4pt scheme takes the least number of iterations, as shown in Fig. 6.8. The asymmetry is also seen in the 1pt log scheme. In the 1pt scheme the asymmetry is not that evident because of the sharpness of the spectral feature that
Figure 6.7. Relative Error as a function of node count for a broad
resonance (FWHM=10\(^{-3}\) centered at 0.3583)

is being resolved. This asymmetry can be seen both for the 1pt and 1pt log schemes
in Fig. 6.9. As far as the convergence is concerned, it can be seen in Fig. 6.9 that the
4pt scheme stops at the size of the spectral feature. The 1pt scheme converges at a
slightly lower value. As in the case of the broad resonance, the 1pt log scheme has
not converged since it reaches our stopping condition of 10\(^{-10}\).

The performance of these schemes is shown in Fig. 6.10. The 4pt scheme outper-
forms the other two and the improvement is more evident than the case of the broad
resonance. In this case also the 1pt log scheme has the worst performance.

6.5.3 REMARKS: 4PTS VS 1PT ADDITION

It is clear from the comparison of the different addition schemes that in the absence
of any analysis of the integrand, that can decide on the usefulness of each node that
Figure 6.8. Scatter plot showing the nodes added at each iteration for a sharp resonance (FWHM=10^{-7} centered at 0.3583). The 4pt and 1pt schemes are operating at a 10% convergence criterion and the 1pt log scheme is at 1%.

is added, a 4pt addition scheme is indeed the most correct and efficient, compared to the 1pt and 1pt log schemes.
Figure 6.9. Difference in energy points for a sharp resonance (FWHM=$10^{-7}$ centered at 0.3583). The 4pt and 1pt schemes are operating at a 10% convergence criterion and the 1pt log scheme is at 1%.

Figure 6.10. Relative Error as a function of node count for a sharp resonance (FWHM=$10^{-10}$ centered at 0.3583)
7. SUMMARY

This work tries to address problems in solar cell and transistor technology. While the field of photovoltaics is striving to push the efficiency of the solar cell, MOSFET technology that drives the micro-processor is trying to decrease power consumption. The properties of such semiconductor devices are understood to vary at the nanometer scale. Therefore an atomistic description of devices is seen as the key to capturing the details of quantum transport and electronic structure. Details of bandstructure that are captured by tight binding models are crucial to getting an accurate picture of the electronic properties of the devices.

In the case of photovoltaics, the ZnSe/GaAs (001) superlattice was investigated with a view to finding a high bandgap material capable of driving solar cell efficiencies beyond 50%. Though materials like InGaN and GaP have shown potential, efficiencies remain low. The ZnSe/GaAs system is lattice matched and has the ability to form quaternary alloys with direct bandgaps. Of the two ways of forming alloys (physical alloying and digital alloying using superlattices) with bandgaps between that of ZnSe and GaAs, the digital alloying approach was investigated to provide a well ordered design alternative to the disordered physical alloying approach. Current fabrication techniques give almost atomic layer precision. Atomistic tight binding theory was used to study the bandgaps of the ZnSe/GaAs (001) superlattice as a function of the constituent monolayers. New improved tight binding parameters for ZnSe based on experimental data were calculated since the existing ones were found to be inadequate. The bandgaps for different periods of the superlattice were estimated and the possibility of engineering a range of bandgaps with the same material system was proposed. It was shown that it is possible to make a high bandgap material with a certain superlattice period. The purpose of this study was to guide experiments by providing variations and trends for the material system in question.
To scale the supply voltage of transistors to reduce power consumption, a vertical TFET geometry was investigated. Current transistor technology is struggling to reduce the power consumption by scaling the supply voltage mainly because the MOSFET is fundamentally limited to a subthreshold slope of 60 mV/dec. Tunnel-Field Effect Transistors (TFETs) are seen as a viable alternative for supply voltage scaling since their subthreshold slope does not have a lower limit. The biggest problem impeding the use of TFETs is the low ON state current. It has been identified that the problem of low ON currents arises because conventional TFET geometries do not allow for a large enough tunneling area for band to band tunneling current to flow through. A vertical tunneling geometry provides a large tunneling area proportional to the gate length and also close to the gate. Using an atomistic fullband quantum transport solver, it was shown that the vertical geometry provided a significant gain in ON state current compared to the lateral p-i-n TFET. It was also shown that such vertical geometries suffer from lateral source to drain leakage in the OFF state and are therefore unsuitable for low power logic applications. To alleviate this problem a design modification was proposed that reduced the OFF state leakage current significantly without degrading the ON state current. Further the device design parameters of the modified geometry were optimized to achieve large ON currents and steep subthreshold slopes. The importance of a global tunneling model was emphasized in view of the fact that previous work in this area had missed current pathways because of the lack of such a tunneling model.

Several numerical techniques that are useful in quantum transport simulations have been alluded to in this work. In particular adaptive quadrature necessary to reduces the cost of integrating a function has been described. Sharp spectral features need to be accurately integrated over. Adaptive quadrature is required to achieve this. Since most methods seek to achieve a very high accuracy, at costs that for the purpose of quantum transport simulations might be prohibitive, an alternate method that can keep the costs down is demonstrated. The tradeoff is a reduction in accuracy which is still good enough for the purpose of electronic transport. There are
other quantities like material properties, that are known to a far lesser accuracy which renders integration of functions to a much higher accuracy redundant. A performance comparison of the proposed method with the existing algorithms was shown and the pros and cons associated with each of these were highlighted.
LIST OF REFERENCES
LIST OF REFERENCES


[69] S. Adachi, “Optical dispersion relations for GaP, GaAs, GaSb, InP, InAs, InSb, Al_{x}Ga_{1-x}As, and In_{1-x}Ga_{x}As_{y}P_{1-y},” *Journal of Applied Physics*, vol. 66, no. 12, pp. 6030–6040, 1989.


VITA
VITA

Samarth Agarwal received his Bachelor of Technology degree in Engineering Physics from IIT, Bombay in 2004 and M.S. in Physics from Purdue University in 2006. He joined the doctoral program in Physics at Purdue in August 2004. He has been working for the Network for Computational Nanotechnology under the supervision of Prof. Gerhard Klimeck and Prof. Ronald Reifenberger since June 2005. His work involves band-gap engineering for solar cells and exploring novel devices for reduction in power consumption of transistors.