CARRIER TRANSPORT IN ULTRA-SCALED DEVICES

A Thesis
Submitted to the Faculty
of
Purdue University
by
Kaushik Balamukundhan

In Partial Fulfillment of the
Requirements for the Degree
of
Master of Science in Electrical and Computer Engineering

August 2008
Purdue University
West Lafayette, Indiana
I would like to express sincere thanks and gratitude to my major adviser Prof. Gerhard Klimeck, for providing me the opportunity to work at the Network for Computational Nanotechnology and supervising my Masters studies at Purdue University and guiding me through a tremendous learning experience, that has not only sharpened my understanding of the subject but also shaped my personality through the rigors of a Master’s degree.

Many thanks to Prof. Mark Lundstrom for his comments and suggestions towards writing this report, and Prof. Supriyo Datta, who provided the foundation for my work through his graduate level course Quantum Transport. I also appreciate the assistance provided by other members in the group, in particular, Samarth Agarwal, Neerav Kharche and Abhijeet Paul for helping me understand extremely difficult concepts and providing valuable suggestions for writing the thesis. I also thank Dr. Neophytos Neophytou for giving me permission to use his code for developing the 1D Heterostructure simulation tool. It is also my duty to acknowledge the funding support provided by FCRP MSD towards the development of the tool. Also, the use of nanoHUB.org computing resources, operated by Network for Computational Nanotechnology funded by NSF, is acknowledged.

Finally, many thanks to my family for their moral support and constant encouragement, a vital tonic to rejuvenate my spirit in moments of despair.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>vi</td>
</tr>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>viii</td>
</tr>
<tr>
<td>1</td>
<td>ELECTRONIC BANDSTRUCTURE</td>
<td>1</td>
</tr>
<tr>
<td>1.1</td>
<td>Electronic Bandstructure</td>
<td>2</td>
</tr>
<tr>
<td>1.1.1</td>
<td>The E-k Dispersion</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Brillouin zone</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>Drawbacks of Effective Mass model</td>
<td>5</td>
</tr>
<tr>
<td>1.4</td>
<td>Tight-binding description of bandstructure</td>
<td>7</td>
</tr>
<tr>
<td>1.5</td>
<td>Summary</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>BANDSTRUCTURE EFFECTS ON ULTRA THIN BODY STRUCTURES</td>
<td>12</td>
</tr>
<tr>
<td>2.1</td>
<td>From Bulk to UTB</td>
<td>12</td>
</tr>
<tr>
<td>2.2</td>
<td>UTB Hamiltonian</td>
<td>12</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Hardwall boundary: Passivation of surface states</td>
<td>13</td>
</tr>
<tr>
<td>2.3</td>
<td>UTB Bandstructure variation</td>
<td>13</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Ballistic nanoscale MOSFET</td>
<td>15</td>
</tr>
<tr>
<td>2.4</td>
<td>Bandstructure effects on Ge n-MOSFETs: Previous work</td>
<td>16</td>
</tr>
<tr>
<td>2.5</td>
<td>InSb material system</td>
<td>17</td>
</tr>
<tr>
<td>2.5.1</td>
<td>Motivation</td>
<td>18</td>
</tr>
<tr>
<td>2.5.2</td>
<td>Bandstructure variation with channel thickness</td>
<td>19</td>
</tr>
<tr>
<td>2.5.3</td>
<td>Bandstructure variation with applied bias</td>
<td>21</td>
</tr>
<tr>
<td>2.6</td>
<td>Summary</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>SEMICLASSICAL Vs QUANTUM CHARGE</td>
<td>24</td>
</tr>
<tr>
<td>3.1</td>
<td>Semiclassical Electrostatics</td>
<td>24</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>3.2</td>
<td>Simulation of a MOS capacitor</td>
<td>26</td>
</tr>
<tr>
<td>3.3</td>
<td>Self-consistent calculation of Charge and Potential in UTB</td>
<td>28</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Numerical E-k relationship</td>
<td>30</td>
</tr>
<tr>
<td>3.4</td>
<td>Schottky Barrier MESFETs</td>
<td>31</td>
</tr>
<tr>
<td>3.5</td>
<td>Top of the barrier model</td>
<td>34</td>
</tr>
<tr>
<td>3.6</td>
<td>Application of top of the barrier model to InSb and Si MOSFETs</td>
<td>37</td>
</tr>
<tr>
<td>3.6.1</td>
<td>High-κ Dielectric materials</td>
<td>39</td>
</tr>
<tr>
<td>3.7</td>
<td>Summary</td>
<td>40</td>
</tr>
<tr>
<td>4</td>
<td>MODFET DEVICE SIMULATION</td>
<td>41</td>
</tr>
<tr>
<td>4.1</td>
<td>MODFET Theory</td>
<td>41</td>
</tr>
<tr>
<td>4.2</td>
<td>Modern day HEMT structure</td>
<td>41</td>
</tr>
<tr>
<td>4.3</td>
<td>From UTB to Heterostructure</td>
<td>44</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Heterogeneous InSb quantum wells</td>
<td>44</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Simulation approach for a Heterostructure</td>
<td>44</td>
</tr>
<tr>
<td>4.4</td>
<td>Summary</td>
<td>47</td>
</tr>
<tr>
<td>5</td>
<td>1D HETEROSTRUCTURE SIMULATION TOOL</td>
<td>48</td>
</tr>
<tr>
<td>5.1</td>
<td>Simulation of MOScap</td>
<td>48</td>
</tr>
<tr>
<td>5.2</td>
<td>Simulation of Heterostructure</td>
<td>50</td>
</tr>
<tr>
<td>5.3</td>
<td>Simulation options</td>
<td>52</td>
</tr>
<tr>
<td>6</td>
<td>Summary and Future work</td>
<td>55</td>
</tr>
<tr>
<td>6.1</td>
<td>Summary</td>
<td>55</td>
</tr>
<tr>
<td>6.2</td>
<td>Future work</td>
<td>57</td>
</tr>
<tr>
<td>LIST OF REFERENCES</td>
<td></td>
<td>58</td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td></td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Energy functional</td>
<td>4</td>
</tr>
<tr>
<td>1.2 Extended and Reduced zone representation</td>
<td>5</td>
</tr>
<tr>
<td>1.3 Non-parabolic effects in bandstructure</td>
<td>6</td>
</tr>
<tr>
<td>1.4 Tight-binding bandstructure of common semiconductors</td>
<td>10</td>
</tr>
<tr>
<td>2.1 Schematic diagram of UTB</td>
<td>14</td>
</tr>
<tr>
<td>2.2 UTB Double-Gate Mosfet</td>
<td>15</td>
</tr>
<tr>
<td>2.3 2D DOS and I-V characteristics of 16nm Ge UTB</td>
<td>16</td>
</tr>
<tr>
<td>2.4 2D DOS and I-V characteristics of 4nm Ge UTB</td>
<td>17</td>
</tr>
<tr>
<td>2.5 Performance of III-V devices</td>
<td>18</td>
</tr>
<tr>
<td>2.6 Effect of ultra-scaling of InSb UTB on bandstructure</td>
<td>20</td>
</tr>
<tr>
<td>2.7 Lowest conduction sub-bands of a 16nm (dotted) and a 4nm (solid) InSb UTB with zero gate bias. Simulation performed with a new tool [14].</td>
<td>21</td>
</tr>
<tr>
<td>2.8 Band-gap variation of InSb UTB with thickness</td>
<td>22</td>
</tr>
<tr>
<td>2.9 Effect of applied bias on InSb UTB bandstructure</td>
<td>23</td>
</tr>
<tr>
<td>3.1 Band-edge diagram of a MOScap. Application of $V_G$ causes band-bending near the oxide-semiconductor interface.</td>
<td>26</td>
</tr>
<tr>
<td>3.2 Semiclassical calculation of a MOScap</td>
<td>27</td>
</tr>
<tr>
<td>3.3 Simulation domains in a MOScap</td>
<td>28</td>
</tr>
<tr>
<td>3.4 MOScap calculations</td>
<td>29</td>
</tr>
<tr>
<td>3.5 2D DOS of InSb UTB</td>
<td>32</td>
</tr>
<tr>
<td>3.6 Illustration of the self-consistent charge calculation using UTB bandstructure charge and Poisson’s equation.</td>
<td>33</td>
</tr>
<tr>
<td>3.7 InSb UTB self-consistent calculation</td>
<td>34</td>
</tr>
<tr>
<td>3.8 Schottky barrier MESFETs operation</td>
<td>35</td>
</tr>
<tr>
<td>3.9 Electron density Vs device cross-section for Schottky barrier</td>
<td>35</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>3.10 Top of the barrier ballistic model</td>
<td>36</td>
</tr>
<tr>
<td>3.11 Injection velocity Vs Gate-voltage for InSb and Si</td>
<td>38</td>
</tr>
<tr>
<td>3.12 Carrier density at the top of the barrier</td>
<td>38</td>
</tr>
<tr>
<td>3.13 Source-Drain current at the top the barrier</td>
<td>39</td>
</tr>
<tr>
<td>3.14 Role of high-$\kappa$ material</td>
<td>40</td>
</tr>
<tr>
<td>4.1 Single-channel Hemt device representation</td>
<td>42</td>
</tr>
<tr>
<td>4.2 Band-profile of HEMT</td>
<td>42</td>
</tr>
<tr>
<td>4.3 Conduction band profile of modern day HEMT</td>
<td>43</td>
</tr>
<tr>
<td>4.4 InSb Quantum well transistor</td>
<td>44</td>
</tr>
<tr>
<td>4.5 Effective Mass Bands of AlInSb/InSb system. Simulation performed with a new tool [14].</td>
<td>45</td>
</tr>
<tr>
<td>4.6 Different simulation domains in a heterostructure</td>
<td>46</td>
</tr>
<tr>
<td>4.7 Self-consistent tight binding calculations</td>
<td>46</td>
</tr>
<tr>
<td>4.8 Output characteristics of QWFET</td>
<td>47</td>
</tr>
<tr>
<td>5.1 Rappture output screenshots for MOScap</td>
<td>49</td>
</tr>
<tr>
<td>5.2 DOS of GaAs MOScap</td>
<td>50</td>
</tr>
<tr>
<td>5.3 AllInSb/InSb quantum well FET</td>
<td>51</td>
</tr>
<tr>
<td>5.4 Rappture output screenshots of MODFET</td>
<td>52</td>
</tr>
<tr>
<td>5.5 Algorithmic details</td>
<td>53</td>
</tr>
<tr>
<td>5.6 Input and Simulation options</td>
<td>54</td>
</tr>
</tbody>
</table>
ABSTRACT

Balamukundhan, Kaushik M.S.E.C.E., Purdue University, August, 2008. Carrier transport in Ultra-Scaled devices. Major Professor: Gerhard Klimeck.

As devices shrink to a few nanometers it becomes essential to model quantum-mechanical effects and to capture the device description at the atomic level. This concept, though extremely important, is rarely appreciated. With rapid device scaling, the role of computational nano-physicists becomes vital. Sophisticated modeling techniques need to be employed to explain various phenomena observed in experiments. In this work, we will look at a generic 1D-modeling tool that computes carrier transport in MOSFETs and MODFETs. We shall start with the effective mass bandstructure description of the channel materials and enlist its drawbacks with ultra-scaling of device thickness. We will then see how nearest neighbor empirical tight-binding model (sp3d5s*) accurately describes the material properties and device geometry at the atomic level. We then extend the tight-binding model for simulating Ultra Thin Bodies (UTB). Appropriate boundary conditions are applied to model the quantum confinement effects. We will then see how bandstructure is sensitive to variation in device parameters like thickness and bias, and how this can significantly alter the carrier transport in the nanoscale regime. Once a proper material description is obtained, the quantum-mechanical charge is computed in regions where there is a rapid variation in electrostatic potential. In regions where potential is flat, we stick to the semi-classical technique of the charge calculation. This reduces the simulation time considerably without compromising on the accuracy. We will then investigate different material systems and compute the I-V and other important device characteristics.
1. ELECTRONIC BANDSTRUCTURE

Since their invention nearly five decades ago [1], Si Metal Oxide Field Effect Transistors (MOSFETs) are still preferred over other devices for creating logic components in memories, processors etc. With a constant, exponential downsizing, physical gate length of Si MOSFETs have scaled down from 100\(\mu\)m to 35nm [2]. A discussion on the MOSFET history is provided in [3].

Due to their steady scaling, MOSFETs have become the primary choice for high performance, low-power logic applications. Gordon Moore in the year 1965, was the first to observe the exponential scaling down of transistor sizes [4, 5], that resulted in exponential increase in transistor count in integrated circuits. His prediction, which later became famous as Moore’s law, is still applicable to present day MOSFETs.

Although Si bulk MOSFETs with sub-10nm channel lengths have been cited in literature [6], they are not a viable option for integrated circuits due to their high off-state leakage currents. As a result, materials like Ge, III-V materials like GaAs, InAs, InSb etc. are being investigated as channel materials for the next generation transistor. Moreover, devices like HEMTs [7] and Heterostructures are also being actively pursued due to their high channel mobilities. Novel InSb-Quantum Well FETs (QWFETs) have triggered tremendous interest due to extremely high electron mobility in the channel, giving rise to excellent transport properties [8]. InSb QWFETs with physical gate lengths of 85nm have already been cited in literature [9].

In this work, we will first investigate the tight-binding description of material bandstructure and the important effects that impact the device performance when it is scaled to a few nanometers. We will then discuss the development of a 1D Heterostructure tool that handles different material systems and is capable of performing self-consistent calculations of device electrostatics.
1.1 Electronic Bandstructure

In this chapter, we will discuss how the Schrödinger Equation is applied for solving quantum mechanical problems. Using an effective mass Hamiltonian we obtain a parabolic E-k relation, also known as the dispersion relation. We understand the concept of Brillouin zone and explain why it is significant in modeling bandstructure. We then see the effective mass model for computing the bandstructure and what are its limitations. We then discuss the semi-empirical tight-binding method for modeling the correct bandstructure. Bandstructures of different semiconductors are shown. Finally, we briefly discuss the concept of spin-orbit coupling and how it affects the bandstructure.

1.1.1 The E-k Dispersion

Most crystals that interest experimentalists exhibit a periodic arrangement of atoms. Atoms when in isolation have discrete energy levels depending on their electronic structure. However, when these atoms are brought together in a regular arrangement to form a crystal, they form bands of energies, called the conduction band (empty states) and the valence band (filled states), which facilitate the movement of electrons and holes respectively in semiconductors. There is a region between the two band edges called the band-gap with attenuated states, known as imaginary bands. The concept of bandstructure is used to identify energy levels of such a solid. The starting point for computing the bandstructure is the time-independent Schrödinger Equation [1]:

\[ -\frac{\hbar^2}{2m^*}\nabla^2 \Psi + U_n(x, y, z) \Psi = E \Psi \]  

(1.1)

where \( m^* \) is the effective mass of the particle, \( \Psi \) is the wavefunction, \( U_n \) is the potential energy of the n atoms, \( \Psi \) is a function of position, \( \Psi^2 \) gives the probability of finding the particle at a particular point in space, and \( E \) gives the total energy of the system. In 1D we can write the Schrödinger equation [2] as:
\[
\begin{bmatrix}
-\hbar^2 \frac{\partial^2}{\partial x^2} & + U_n(x) \\
2m^* \frac{\partial^2}{\partial x^2}
\end{bmatrix} \psi(x) = E\psi(x) \tag{1.2}
\]

The term inside the bracket is known as the Hamiltonian of system. The Hamiltonian is a matrix whose diagonal elements represent the on-site energies of the atoms and the off-diagonal elements gives the value of the coupling of the atom with the nearby atoms. The Hamiltonian \[\text{[?]}\] can be described by the following matrix:

\[
\begin{bmatrix}
2t_0 + U_1 & -t_0 & 0 & 0 \\
-t_0 & 2t_0 + U_2 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 2t_0 + U_{N-1} \\
0 & 0 & 0 & -t_0 & 2t_0 + U_N
\end{bmatrix} \tag{1.3}
\]

Recall equation (1.2). Assuming that \(U_n(x) = U_0\), a constant, an exponential is a possible solution for the above equation. Thus, we get:

\[\Psi \propto e^{ikx}\]

which is the plane wave ansatz. The parameter \(k\) is called the Bloch wavevector, which when multiplied by \(\hbar\) gives the crystal momentum. This solution will work only when:

\[E = U_0 + \frac{\hbar^2 k^2}{2m^*} \tag{1.4}\]

Using (1.4), we obtain a \textit{parabolic E-k relation}. Thus, knowing the effective mass and the \(k\) points we can compute the energies at those points.

1.2 Brillouin zone

Let us assume a chain of \(N\) atoms, each separated by distance \(a\). If we assume a periodic potential within the crystal, then according to Bloch’s theorem,

\[\psi(x + a) = e^{ika}\psi(x)\]
For a given $E$, values of $k$ differing by a multiple of $2\pi/a$ give rise to one and the same wavefunction solution. Therefore, a complete set of distinct $k$-values is limited to a $2\pi/a$ range. Thus, the different range of $k$-values is $-n\pi/a$ to $+n\pi/a$ where $n = 1, 2, 3, ...$. It is common practice to choose the $k$ range from $-\pi/a$ to $+\pi/a$ as the first Brillouin Zone, $-2\pi/a$ to $+2\pi/a$ as the second Brillouin zone and so on.

Performing an approximate one-dimensional analysis using the Kronig-Penny model, we can identify the allowed energy values for different $k$ points. We see from this model that there are discontinuities in energy values at the edges of the Brillouin Zone. These discontinuities represent the band-gap of the solid between various bands, also known as the forbidden gap. When all the bands are folded back into the first Brillouin zone, the diagram is called a reduced-zone representation.

![Energy Functional Vs Energy](image)

**Fig. 1.1.** Energy Functional Vs Energy. The horizontal lines represent the allowed energies. Simulations were performed using the Periodic Potential Lab [10]

Figure 1.2 illustrates the bandstructure as an extended-zone and a reduced-zone representation of allowed $E - k$ states in a one-dimensional crystal using the Kronig-Penny model. The black curve indicates the dispersion of a free particle and blue curve represents that actual dispersion of a particle within a crystal.

The Brillouin zone shown in the figure were described using a one-dimensional or scalar $k$. However, for bandstructure description of 3-D lattices, the wavevector $k$
Fig. 1.2. Left: Extended zone representation, Right: Reduced zone representation. We can see how the bands are folded back into the first Brillouin zone. Simulations were performed using Periodic Potential Lab [10].

becomes a vector and Brillouin zones become volumes. Thus, a 3-D Brillouin zone is the volume in $k$ space enclosing the set of $k$ values associated with a given energy band.

1.3 Drawbacks of Effective Mass model

We can calculate the E-k dispersion using a readymade formula that gives a parabolic relationship between Energy and $k$ points. So far we have used a term Effective mass to describe the mass of an electron or a hole when it is in motion inside a crystal. Ignoring the potential energy term in (1.4) we can see that

$$m^* = \frac{1}{\hbar^2 \frac{d^2 E}{dk^2}}$$

Thus, Effective mass is the curvature of the E-k relation at the Gamma ($k=0$) point.

We have assumed a parabolic dispersion in our discussion so far. Realistically however, the bands are non-parabolic in nature. The effective mass model is useful for a few $k$ points close the $\Gamma$ valley where the bands tend to remain parabolic. Figure 1.3 shows how the bandstructure becomes non-parabolic as we move in a
particular direction inside the crystal. At higher \( k \) points, there is significant difference in the prediction of energy states. The effective mass model predicts a higher energy, especially at higher \( k \) points. Through [11–13], we have developed a new tool [14] for performing these simulations.

The simulation was performed using a new tool [14]

![Fig. 1.3. Effective mass (parabolic) and 10 band tight-binding (non-parabolic) bandstructure. As we move away from the Γ point, the bands become highly non-parabolic. Simulation performed with a new tool [14].](image)

When the devices are scaled to few a nanometers, we observe significant changes in the bandstructure as a result of quantum confinement. The device properties can be altered due to ultra-scaling. Effective mass model fails to capture these effects.

For some materials such as AlSb, AlAs etc., the conduction band minima does not occur at the Γ point. There are valleys outside the Γ point along \( X \) (100) and \( L \) (111) directions inside the crystal. In material systems like AlAs-GaAs quantum wells, these \( X \) valley states [15] become important at high bias values. The single band effective mass model fails to explain these so called indirect band-gap semiconductor bandstructure.
The conduction and valence bands are coupled by an imaginary band [16] called complex band that gives the rate at which wavefunction of an electron decays within the band-gap. The single band model does not model this wrapping of complex band. Moreover, the valence bandstructure, especially the heavy hole band, cannot described by a parabolic dispersion due to its inherent anisotropy.

Effective mass is thus not the preferred nor the correct way to model nanoelectronic devices.

1.4 Tight-binding description of bandstructure

The effective mass model discussed in the previous section greatly reduces the computational effort required to model nanoscale semiconductors. But its failure to model the correct bandstructure has led to the use of more sophisticated models like tight-binding, a formulation for modeling realistic bandstructures [17, 18]. As the device geometry gets smaller (a few tens of nanometer thick), the quantum mechanical effects can be modeled using semi-empirical tight binding technique. In this section we will discuss the model briefly and look at how it models bandstructure accurately.

Slater and Koster [19] originally proposed a semi-empirical tight binding model using Linear Combination of Atom(like) Orbitals (LCAO). In this work, we will not look at the formalism of the tight-binding theory but just see how this model captures all the bandstructure effects accurately.

All common semiconductors (like GaAs) belong to the diamond structure, which has a unit cell consisting of two atoms; a cation and an anion. For each atom we use orbitals $s, p, d$ and $s^*$. We choose this model since $s$ and $p$ orbitals are enough to model the valence band, and the excited $d$ orbitals are necessary to model the conduction band properly [20]. We know from atomic theory that each orbital has a different orientation. For example, the $s$ and $s^*$ orbitals are spherical; the $p$ orbital has $p_x, p_y, p_z$ orientations; the $d$ orbitals have $d_{xy}, d_{yz}, d_{zx}, d_{x^2-y^2}, d_{z^2}$ orientations which gives us a total of 10 orbitals for anion and 10 for the cation. Each orbital belonging
to the anion of one atom has an onsite energy represented by $E_i^a$ where $i$ represents the type of the orbital. Similarly, we have onsite energy for the orbitals belonging to the cation $E_i^c$. These onsite energies form the diagonal elements of a matrix for the anion denoted by $[H_{aa}]$.

\[
[H_{aa}] = \begin{pmatrix}
E_s^a & & & \\
& E_p^a & & \\
& & E_p^a & \\
& & & E_p^a
\end{pmatrix}
\]

Similarly, $[H_{cc}]$ describes the on-site elements for the cation. Now the atomic orbitals on the anion and the cation interact with each other. Those parameters are described by another block denoted by $[H_{ac}]$ and $[H_{ca}]$.

\[
[H_{ac}] = \begin{pmatrix}
H_{sp^3s^*-sp^3s^*}^{ac} & H_{sp^3s^*-d^5}^{ac} \\
H_{d^5-sp^3s^*}^{ac} & H_{d^5-d^5}^{ac}
\end{pmatrix}
\]

Using these we get the complete Hamiltonian

\[
[H] = \begin{pmatrix}
[H_{aa}] & [H_{ac}] \\
[H_{ca}] & [H_{cc}]
\end{pmatrix}
\]

(1.5)

Since each anion and cation onsite block is of size 10 X 10, the resulting size of the hamiltonian of equation (1.5) is 20 X 20. Thus, from a simple, single band effective mass model we have now gone into the 10 band tight-binding model, also known as $sp^3s^*d^5$ model. We immediately see that computationally the 10 band model is slightly more challenging than the single band model, but its not that big a price
to pay for a more accurate representation of electronic bandstructure. Figure 1.4 shows the 10 band tight-binding bandstructure for Si, GaAs and InSb. The material parameters were obtained from [21]. We can see that there are conduction band valleys outside the Γ point. The curvature of these valleys gives the effective mass. For Si, an indirect band-gap semiconductor, the conduction band minima is at the X point. For GaAs and InSb, the conduction band minima occurs at the Γ point. Also, the valence bands of all the materials are highly non-parabolic.

The bandstructure we have obtained is reasonably accurate, but it does not describe the top of the valence band very well. Moreover the 10-band Hamiltonian should be understood as doubly degenerate, one for up-spin and one for down-spin. To obtain the correct bandstructure it is necessary to include spin-orbit coupling. As a result, after including the spin-orbit interaction there will be twice as many bands. The spin-orbit Hamiltonian [22] is explicitly added to the tight-binding Hamiltonian \([H]\), and is often written as described in Equation (1.6). A detailed derivation of the approach for modeling spin-orbit coupling into the Hamiltonian can be found in [23].

\[
H_{so} = \frac{\hbar q}{4m^2c^2} \vec{\sigma}.(\vec{E} \times \vec{p})
\] (1.6)

where the Pauli spin matrices \(\vec{\sigma}\) are defined as

\[
\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}
\]

Figure 1.4 describes the bandstructure of Si, GaAs and InSb, with and without spin-orbit coupling. Using the tight-binding model with spin-orbit interaction, we can see that the hole bandstructure is quite different from the one without the spin-orbit term telling us that spin-orbit effects are extremely pronounced for holes. The split-off band shifts down in energy. Also, the hole band split into a heavy-hole and a light-hole band. Splitting in conduction sub-bands ais also visible.
Fig. 1.4. Bulk bandstructure of Si, GaAs and InSb using $sp^3s^*d^5$ basis. Left: Spin-orbit coupling term not included in the Hamiltonian. Right: The bandstructure with spin-orbit coupling. Splitting of hole bands into heavy-hole and light-hole bands is visible. Split-off band shifts down in energy. Splitting of conduction sub-bands is also visible. Simulation performed with a new tool [14].
1.5 Summary

In this chapter we described the dispersion relation using the Schrödinger equation. We explained the effective mass formalism of bandstructure, listed the various drawbacks of such an approach and described how these are overcome by using tight-binding method.
2. BANDSTRUCTURE EFFECTS ON ULTRA THIN BODY STRUCTURES

2.1 From Bulk to UTB

In this chapter, we will see how the bandstructure changes when we confine a system and limit its thickness to a few atomic layers. Once such a system is obtained it becomes essential to treat each atom separately, resulting in a 2D Hamiltonian that looks quite different from a bulk Hamiltonian. Treatment of the so-called Ultra Thin Body (UTB) is both complicated and computationally expensive as we need to compute the eigenvalues of the matrix that is typically 1000 by 1000 for a 10nm UTB.

We shall first start with the discussion of Hamiltonian construction for an Ultra Thin Body (UTB). We will understand what kind of boundary conditions are applied for simulating "real" devices. Finally we will look at some of the parameters that significantly alters the bandstructure of the material and how this affects the device performance.

2.2 UTB Hamiltonian

In this section we look at the specific case of constructing a tight-binding Hamiltonian for an Ultra Thin Body (UTB). In an UTB the X and Y directions are large enough for us to assume periodic boundary conditions, compared to the Z direction which is finite. The Hamiltonian is constructed on a atomic basis in the Z direction. Along X and Y directions we still assume plane wave states. The Hamiltonian is in general a function of the $k_x$ and $k_y$ wave-vectors. The wave-vector $k_z$ does not have
any physical significance. It should be noted that the $k_x$ and $k_y$ dependence comes from the off-diagonal coupling blocks. The Hamiltonian takes the following form,

$$
[H(k_x, k_y)] = \begin{pmatrix}
H_{aa} & H_{ac}(k_x, k_y) \\
H_{ac}(k_x, k_y)^\dagger & H_{cc} & H_{ac}(k_x, k_y) \\
& H_{ac}(k_x, k_y)^\dagger & H_{aa} & H_{ac}(k_x, k_y) \\
& & & \ddots & \ddots & \ddots \\
& & & & & \ddots \\
\end{pmatrix}
$$

2.2.1 Hardwall boundary: Passivation of surface states

Let us assume the material to be confined in $z$ direction and infinitely long in the other directions. This means that we can apply periodic BC’s in $x$ and $y$ directions. In the $z$ however, we apply non-periodic BC’s, where hard-wall or closed BC’s are imposed for passivation of dangling bonds that arise at the beginning and the end of UTB chain. 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. The boundary conditions are extensively discussed in [24].

2.3 UTB Bandstructure variation

In this section, we will discuss some important effects on the bandstructure of UTB with changes in device parameters. We will look at different UTB materials and point out why these bandstructure variations play a very important role in modifying device performance.
Fig. 2.1. Schematic diagram of UTB
2.3.1 Ballistic nanoscale MOSFET

There has been a lot of interest in exploring new channel materials, especially materials like Ge and III-V binary compounds like InSb, InAs etc. [25]. As devices continue to shrink, channel thickness is in the order of few tens of nanometers, and it will not be long before it enters the sub-10nm regime. In this realm of operation bandstructure effects start to play a very significant role in device performance. These bandstructure effects are accurately captured using an atomistic approach based on the \( sp^3s^* \) tight-binding model.

Figure 2.2 shows the schematic of a Double-gate MOSFET. Here the channel is a few nanometers in thickness (\( z \) direction) and has gate terminals on the top and bottom. The gate and the channel are separated by a thin oxide layer. The source and the drain contact establish a current flow within the device.

![UTB Double-Gate Mosfet](image)

The thickness of the channel is a few nanometers. The device is treated using UTB bandstructure. In this analysis, we shall neglect the effect of scattering as our primary objective is to understand the role of bandstructure. Hardwall boundary conditions are used at the top and bottom semiconductor-gate interfaces preventing the wavefunction from penetrating into the oxide.
2.4 Bandstructure effects on Ge n-MOSFETs: Previous work

We will now look at the bandstructure effects on Ge (100) UTB DG-MOSFETs. Ge is an indirect band-gap semiconductor, with \( L \) valley forming the lowest conduction sub-band. The inset of Figure 2.3 [26] shows the bandstructure of 16nm Ge (100) UTB. In fact, the first three \( L \) valleys are lower than \( \Gamma \) and \( X \) valleys. As a result, the transport at lower biases would occur only through the \( L \) valley states. Figure 2.3 [26] also shows the 2D Density Of States (DOS) for using Effective Mass (EM) and Tight-Binding (TB) models. For low bias values, the EM and TB DOS agree very well. This is because the \( \Gamma \) and \( X \) valley states are higher in energy than \( L \) valley states, where most of the charge population takes place. However, for higher bias values the DOS using TB approach is much higher than EM, as for these bias values the higher \( \Gamma \) and \( X \) valleys contribute to the DOS, which is not accounted for using EM approach.

![Figure 2.3](image)

(a) Fig. 2.3. (a): 2D DOS of 16nm Ge (100) UTB using EM (\( L \) valley effective mass is used) and TB approach (Inset: Bandstructure of 16nm Ge UTB. (b): \( I_d - V_G \) comparison using EM and TB approach. Similar current values are predicted as only \( L \) valleys contribute to the charge. Figures were obtained from [26].

When we scale the device down to 4nm, we start to see some interesting bandstructure effects. First of all, \( \Gamma \) and \( X \) valley states have moved down in energy than the second \( L \) valley. Moreover, the \( L \) valley states suffer from severe non-parabolic
effects, which is manifested in the $I_d - V_G$ characteristics in the device as a threshold voltage ($V_T$) shift. As a result, the value of current using EM approach is significantly less than that using TB (Figure 2.4).

Fig. 2.4. (a): 2D DOS of 4nm Ge (100) UTB using EM ($L$ valley effective mass is used) and TB. Γ and $X$ valleys have moved down in energy and contribute towards DOS. Non-parabolicity effects in $L$ valley affect the I-V characteristics. (b): $I_d - V_G$ comparison using EM and TB approach. $V_T$ shift due non-parabolic $L$ valleys. EM underestimates the current as contribution from Γ and $X$ valleys not taken into account. Figures were obtained from [26].

2.5 InSb material system

III-V semiconductors, like InAs, InSb etc. are being actively pursued as channel materials for next generation nanotransistors. InSb based Quantum Well Field Effect Transistors (QWFET) offer great promise as an ultra-fast, very low power digital logic technology. InSb QWFETs with 85nm physical gate length ($L_g$) have been demonstrated [9]. AlInSb/InSb modulation-doped QWFETs offer very good transport properties [27]. Benchmarking against Si MOSFETs, InSb QWFETs achieve equivalent high-speed performance with 5-10 times lower power dissipation.
2.5.1 Motivation

InSb has a very high mobility and saturation velocity, and is therefore a promising candidate for ultra-high speed FETs operating at very low supply voltages. When benchmarked against the state-of-the-art Si MOSFETs, these devices exhibited improvements in energy-delay product, which makes them suitable for low power applications. Figure 2.5 shows the variation of gate-delay and energy-delay product with physical gate length for III-Vs, Si MOSFETs and CNTFETs. It can be seen that III-Vs like InSb clearly outperform other devices in this regard.

![Fig. 2.5. Comparison of gate-delay and energy-delay product Vs physical gate lengths of III-V semiconductors with Si MOSFETs and CNTFETs [28]](image)

However, there are certain drawbacks of choosing an InSb material system. In spite of its high mobility and saturation velocity, InSb has a very small band-gap (approximately 0.2 eV). This results in Band To Band Tunneling (BTBT), especially when operating in reasonably high bias values, which can deteriorate the performance of the device.

Due to their small Γ valley effective mass, III-V materials like InSb are being investigated as high mobility channel materials for high performance NMOS. Under ballistic conditions, however, the main advantage of a semiconductor with a small transport effective mass is the high injection velocity. But these materials have a much
lower DOS, which lowers the inversion charge and hence reduce the drive current. The effects of BTBT have been investigated in III-V DG nMOSFETs [29]. InSb is shown to have a comparatively high BTBT current, which limits its performance. However, the tunneling currents in III-V materials are substantially reduced because of significant increase in band-gap due to quantization in ultra-small devices.

The scalability of InSb and III-V semiconductors needs to be demonstrated. Application of these devices as high-speed logic components must be investigated for physical gate lengths of 22nm. Moreover, there are significant challenges [30] prior to the implementation of III-V materials in logic, one of which is their heterogeneous integration with Si substrate.

2.5.2 Bandstructure variation with channel thickness

In this section, we will look at how the bandstructure is affected by scaling the thickness of the channel. Different channel materials show different behavior once the thickness of the UTB is varied. The work by Rahman et. al. [26] showed why tight-binding calculations are important when devices are scaled to such limits. When we look at Ge conduction bands with 113 atomic layers (approx. 16nm), we see that the lowest valleys are formed at the Γ point. Moreover, the energy difference between the Γ and X valleys very small. Hence, not only the Γ valley but also the X valleys contribute towards the Density of States. Also, we see that the TB conduction band splits into sub-bands due to finite body thickness. Using a single band effective mass model, we cannot explain the contribution of higher valleys as they have different curvature and hence different effective mass.

However, when we scale the device to 31 atomic layers (approx. 4nm), we start to see some dramatic changes in the bandstructure. Due to this ultra scaling we see an appreciable splitting in the sub-bands.

When we look at the energy bandstructure of [100] InSb, we see similar results. For a 16nm InSb, the conduction sub-bands are closely spaced. But with the re-
duction of channel thickness to 4nm, the conduction bands split appreciably. Both structures exhibit a direct band-gap. But band-gap of 4nm InSb is greater than that of 16nm InSb due to stronger confinement in the thinner structure. The splitting of the conduction sub-bands is also clearly visible. Figure 2.6 illustrates these effects. Figure 2.7 shows the lowest conduction sub-bands of a 16nm and a 4nm InSb UTB. We see that when the channel thickness is reduced, the conduction sub-bands are shifted up due to severe quantum confinement.

![Graph](image)

(a) 16nm InSb UTB  
(b) 4nm InSb UTB

Fig. 2.6. Left: 16nm InSb UTB bandstructure. Splitting of conduction sub-bands as a result of channel confinement is evident. The higher Γ valleys have greater curvature, hence higher effective mass. Right: 4nm InSb UTB bandstructure. The bands separated out in energy and their curvature is increased. The lowest conduction sub-bands also move higher in energy. Simulation performed with a new tool [14].

Figure 2.8 shows how the band-gap for InSb UTB varies with channel thickness. We see that as the thickness becomes large and is approx. 15nm, the energy band-gap
approaches that of the bulk device. However, as the thickness is less than 8nm, the band-gap rises significantly.

![Graph](image)

Fig. 2.7. Lowest conduction sub-bands of a 16nm (dotted) and a 4nm (solid) InSb UTB with zero gate bias. Simulation performed with a new tool [14].

### 2.5.3 Bandstructure variation with applied bias

In this section, we will see how bandstructure is modified with applied voltage. From basic quantum theory, we understand that when a bias is applied to a channel the energy levels of the channel shift down to accumulate more charge. This phenomena is no different in UTB, however not that simple to understand. Figure 2.9 shows how the bandstructure is modified with applied bias.

Recall the Hamiltonian construction of the effective mass Schrödinger Equation Eq. ???. We added the potential from the equation into the diagonal elements of the
Fig. 2.8. Variation of band-gap with thickness for InSb UTB. As the thickness is increased the band-gap approaches the band-gap of bulk InSb (0.22 eV). Simulations performed using [14].

Hamiltonian. When we apply a gate bias to a UTB, it modifies the potential inside the channel. This potential is calculated using Poisson’s equation (Eq. 2.1),

$$\nabla^2 \Phi = -\frac{\rho}{\epsilon}$$ \hspace{1cm} (2.1)

The potential $\Phi$ is added to the diagonal elements of the Hamiltonian. Thus, the eigen values of the new Hamiltonian (with potential added) are different from the original one. This modifies the energy levels inside the channel resulting in a different bandstructure. In Figure 2.9, the top figure shows bandstructure of a 4nm InSb UTB without bias and the bottom shows the bandstructure with a bias of 0.5V. The line at energy zero is the Fermi level for the device. We see that when there is no bias, the bandstructure is considerably higher in energy. This leads to a lesser population of the conduction band states by electrons resulting in lower charge density lower current inside the device. But with the application the bands shift down below the Fermi level and they are populated by electrons, resulting in greater current flow inside the device.
Fig. 2.9. InSb 4nm UTB bandstructure variation with bias Top: $V_G = 0V$. Bottom:$V_G = 0.4V$. When the gate bias is increased the bands shift down to accumulate the charge. The amount of charge inside the device can be determined by the Fermi level. Simulation performed with a new tool [14].

2.6 Summary

In this chapter we extended the idea of tight-binding for modeling the bandstructure of ultra thin bodies. We investigated the bandstructure of different materials and showed how the scaling down of device thickness splits the conduction sub-bands. Also there is a variation in the curvature of the bands which results in different effective mass. Also the sub-bands shift down when higher bias is applied that results in populating more states using the Fermi level.
3. SEMICLASSICAL VS QUANTUM CHARGE

So far we have discussed why it is essential to represent the correct bandstructure and how it is modeled using tight-binding technique. Most of the phenomenon within the semiconductor can be explained using bandstructure. In this section we will see how the energy levels in a device that are derived from the bandstructure, are used to calculate the density of states using which we can compute the charge inside the device.

We will start this section with semiclassical semiconductor theory and explain carrier transport using the drift-diffusion model. This model is particularly useful when the size of the device is large of the order of few tens of nanometers. We will then apply ballistic quantum-mechanical transport equations to certain sections of the device where quantum confinement plays a very significant role. We will see at the end of this chapter why it is important to include semiclassical and quantum mechanical theories side by side and the improvements it offers in terms of computational speed and accuracy while modeling semiconductor devices.

3.1 Semiclassical Electrostatics

Let us now shift our focus from Quantum Mechanics and discuss about semiclassical theory for modeling semiconductor devices. The energy-band diagram is the starting step for all our electrostatics. The energy-band diagram and E-k diagram should not be confused with one another. E-k diagram is a pure quantum mechanical concept that gives the value of energy at every $k$ point for a particular length of the device. Whereas, the energy-band diagram is a representation of the conduction and valence bands of semiconductors.
When materials with different band-gaps are brought together, the resulting device is called a semiconductor Heterostructure. When such materials with different doping and band edges are brought together such that their fermi levels are aligned, the electrons (or holes) diffuse from a region of higher concentration to a region of lower concentration until their population on either side of the interface becomes equal. This is the fundamental principle behind the formation of $p-n$ junction diode and much more complicated devices like Field-Effect Transistors (FET).

Once the energy bands are formed, it is now our task to identify how many energy states are to be found at any given energy in the conduction and valence bands. This energy distribution of states, or Density Of States (DOS), as it is more commonly known, is an essential component in determining carrier distributions and concentrations. Equations (3.1) and (3.2) describe the density of states of the conduction band and valence bands respectively [31].

\[
g_c(E) = \frac{m^*_n \sqrt{2m^*_n (E - E_c)}}{\pi^2 \hbar^3}, E \geq E_c \tag{3.1}
\]

\[
g_v(E) = \frac{m^*_n \sqrt{2m^*_p (E_v - E)}}{\pi^2 \hbar^3}, E \leq E_v \tag{3.2}
\]

Whereas the density of states tells one how many states exist at a given energy $E$, the Fermi function $f(E)$ specifies how many of those existing states at the energy $E$ will be filled with an electron. Fermi function is simply a probability distribution function. It is given by

\[
f(E) = \frac{1}{1 + e^{(E-E_F)/kT}} \tag{3.3}
\]

After identifying the Fermi function, the electron (or hole density) is simply the product of fermi function and conduction band (or valence band) DOS given by (3.1) and (3.2).

\[
n = \int_{E_c}^{E_{top}} g_c(E) f(E) dE \tag{3.4}
\]
\[ p = \int_{E_{\text{bottom}}}^{E_v} g_v(E)[1 - f(E)]dE \] (3.5)

3.2 Simulation of a MOS capacitor

In this section, we will simulate a MOS capacitor (MOScap). Figure 3.1 illustrates the band-edge diagram of a MOScap. When a gate is applied to a MOScap, the potential inside the device changes resulting in band-bending at the semiconductor-oxide interface. The charge is calculated using the equation described above.

Fig. 3.1. Band-edge diagram of a MOScap. Application of \( V_G \) causes band-bending near the oxide-semiconductor interface.

The above approach is a pure semiclassical one. Figure 3.2 shows the semiclassical charge that is computed self-consistently with potential. It shows the self-consistent electrostatic potential profile and electron densities for different gate voltages. As the gate voltage is increased, there is more band bending resulting in more accumulation of charge.

The semiclassical method overestimates the carrier density at the interface. In regions where there is a rapid variation in potential, it is imperative to include quantum mechanical treatment due to confinement of charge. However, quantum mechanical calculations are extremely time consuming as it deals with matrices that are extremely
Fig. 3.2. Semiclassical calculation of a MOScap using drift-diffusion model. When the bias is high, the band bending is more pronounced. More charges are accumulated at the oxide-semiconductor interface. Simulation performed with a new tool [14].

huge in size. It is advisable to divide the simulation into a semiclassical domain and a quantum mechanical domain. Figure 3.3 illustrates such a scheme.

Once we have identified our simulation domains we then need to find out whether we can treat the quantum mechanical region using effective mass Schrödinger equa-
Fig. 3.3. The region close to the oxide-semiconductor interface is treated quantum mechanically. Semiclassical calculations are applied in the regions where the potential is flat.

or do we need to use tight-binding calculations. For instance, if we choose the quantum domain as 20nm, then effective mass Schrödinger equation would provide reasonably accurate results. If we use tight-binding and treat the region using UTB bandstructure then the simulations are be extremely time consuming. However, for ultra-scaled devices it is essential to use tight-binding calculations due to severe quantum confinement.

The figure below shows the effective mass calculation for Si MOScap device. Both self-consistent potential and electron density plots are shown. We see that the electron density does not pile up near the interface as in semiclassical calculation. This is due to quantum confinement. Also, when there is greater band-bending more carriers accumulate near the semiconductor-oxide interface.

### 3.3 Self-consistent calculation of Charge and Potential in UTB

In the previous section we computed the charge and potential self-consistently for a bulk device that was described by an effective mass hamiltonian. However, when
Fig. 3.4. Simulations for a Si MOScap with different simulation domains. The width of the quantum domain is 20nm and modeled using effective mass Schrödinger equation. Substrate doping of $1 \times 10^{18} cm^{-3}$ was used; $T_{oxide} = 1$nm. Simulations were performed using Schred tool in www.nanohub.org [32].

the quantum confined region is really small (a few nanometers), we need to treat the region using UTB bandstructure described in chapter 2.
### 3.3.1 Numerical E-k relationship

The "dispersion" described by the effective mass hamiltonian gives a parabolic relation between energy and reciprocal space. This is called analytical E-k relationship. However, when we use tight-binding approach for calculating the E-k relation, we obtain a non-parabolic relation that cannot be described by an analytical formula. In such cases, we perform a Numerical Integration of Energy.

Recalling basic quantum transport equation, we know that carrier density is the product of Density of states (DOS) and the fermi function and a particular energy (equation (3.6)).

\[
N = \int_{-\infty}^{+\infty} f_0(E - \mu)D(E) \, dE \tag{3.6}
\]

where the 2D DOS is given by the expression [33]

\[
D(E) = \frac{S}{4\pi^2} \int_{E_{\text{min}}}^{E_{\text{max}}} \delta(E - E(k_x, k_y)) \, dk_x dk_y \tag{3.7}
\]

The real power of the expression is that it is valid regardless of whether we have an analytical expression for energy. In order to calculate the DOS we broaden energies using a Lorentzian distribution given by equation (3.8).

\[
\delta(E - \epsilon_\alpha) \rightarrow \frac{\gamma/2\pi}{(E - \epsilon_\alpha)^2 + (\gamma/2)^2} \tag{3.8}
\]

Energy resonance has a finite life-time, which means that it has a specific width. In order to compute the DOS, we first need to create an energy grid that enables us perform the numerical integration of energy (equation (3.7)). The grid needs to be extremely fine in order to be able to capture the energy resonance. The fineness of the grid is decided by the width of the resonance \(\gamma\), used in equation (3.8). Also, we must ensure that the resonances are not broadened to the extent that they overlap with each other, which will result in an incorrect DOS value. In figure 3.5 we show the 2D DOS plot for a 16nm and a 4nm InSb UTB using a Lorentzian width \(\gamma\) of 0.5 meV. The energy grid was 1 meV. When the device thickness is fairly large then the
subbands are closely spaced to one another. In this case, the DOS more or less looks continuous due to close spacing of conduction sub-bands. However, for smaller systems the DOS shows a step-wise increments at every valley in the bandstructure.

Figure 3.5 shows how the different valleys contribute towards the DOS.

After computing the DOS, charge is calculated (equation (3.6)) using the numerical E-k relationship that is obtained from the UTB bandstructure. The computed charge is the fed into the Poisson solver which computes the potential. This updated potential is then added to the UTB hamiltonian to calculate the new bandstructure. This is the self-consistent calculation using the UTB Bandstructure. Figure 3.6 illustrates this self-consistency.

Appreciable changes in bandstructure can be seen when self-consistent calculations are performed. Figure 3.7 shows how the InSb bandstructure in [100] direction is altered when self consistent calculations are performed for a zero gate bias. We can see that the band has shifted down when we perform a self-consistent calculation. At the Γ point the conduction band is shifted down by approximately 100meV. This appreciable shift in energy can affect the device characteristics.

3.4 Schottky Barrier MESFETs

A Schottky barrier is formed at the semiconductor-metal interface when the work-function of the metal is higher than the semiconductor workfunction and such a device is called Metal-Semiconductor FET (MESFET). The fundamental difference between a MOSFET and a MESFET is that in the latter instead of the oxide layer the metal is directly grown on the semiconductor substrate.

Under equilibrium conditions, the fermi levels of the metal ($E_{FM}$) and semiconductor ($E_{FS}$) align themselves. Due to the diffusion of electrons from semiconductor to the metal, the energy bands are pulled down. Applying a positive gate bias pulls $E_{FM}$ down resulting in a reduced barrier for the semiconductor electrons ($V_b$), because of which more electrons can cross the barrier and go into the metal (Figure 3.8).
Fig. 3.5. 2D DOS for 16nm (top) and 4nm (bottom) InSb UTB. The DOS plot shows a step-wise increment when each conduction subband valley is encountered. The energy resonances were broadened using $\gamma = 0.5\, meV$ for the Lorentzian.
Figure 3.9 shows the comparison of electron density for a 10nm Si UTB DG device with schottky barriers on the top and bottom gates. The workfunction of the metal was chosen to be more than that of the semiconductor. We see for high gate bias more carriers are accumulated near the metal-semiconductor interface. Schottky barrier typically act as an ohmic contact for positive gate bias and behaves in a non-ohmic manner when a negative bias is applied. When $V_G$ is negative, the barrier height is too high for the electrons to surpass. In that case, the leakage current inside the device is practically negligible.
3.5 Top of the barrier model

MOSFET channel lengths continue to shrink rapidly toward the sub-10 nm dimensions called for by the International Technology Roadmap for Semiconductors [36]. With the use of high mobility materials and nanometer channel lengths, the device
Fig. 3.8. Left: $V_g=0.2\text{V}$. Right: $V_g=0.4\text{V}$. Applying more higher gate bias reduces $V_b$ resulting in more electrons crossing the metal-semiconductor barrier. $\Phi_M > \Phi_S$.

Fig. 3.9. Electron density with device cross-section for a 10nm Si UTB. UTB was doped with dopant impurity concentration of $N_d = 1e17/cm^3$. $\Phi_M > \Phi_S$. $\Phi_B = 0.5eV$ under equilibrium conditions. Simulation performed with a new tool [14].
essentially operates in near-ballistic limits. In this section we will explain the theory behind ballistic transistors and apply it for modeling Source-Drain (S/D) current in a MOSFET. A detailed derivation of the analytical formula is discussed in [37,38].

Figure 3.10 illustrates the top of the barrier ballistic model. $E_f1$ and $E_f2$ represent the source and drain fermi levels and the potential profile is along the source-drain direction. What top of the barrier essentially tells us is: If we have an E-k relationship (Fig. 3.10 shows an effective mass dispersion) that describes the bandstructure of the material, then we can populate these negative and positive k-states using the source and drain fermi levels. The net current within the device is due to the different between these positive and negative k-states. Choosing the bandstructure at the top of the barrier assumption makes the device ballistic in nature. At the top of the barrier there will not be any reflection of the carriers from barrier and hence they can go from source to drain and vice versa ballistically.
Table 3.1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{ox}$ [nm]</td>
<td>0.9 (EOT)</td>
</tr>
<tr>
<td>$t_{body}$ [nm]</td>
<td>2.5</td>
</tr>
<tr>
<td>$L_g$ [nm]</td>
<td>22</td>
</tr>
<tr>
<td>$I_{off}$ [uA/um]</td>
<td>0.71</td>
</tr>
<tr>
<td>$I_{sat}$ [uA/um]</td>
<td>1513</td>
</tr>
<tr>
<td>$V_{dd}$ [V]</td>
<td>1</td>
</tr>
</tbody>
</table>

3.6 Application of top of the barrier model to InSb and Si MOSFETs

The top of the barrier model can be used to compute current through a MOSFET. The channel material in a MOSFET can be described using an UTB bandstructure. Once the bandstructure is calculated we can use that E-k relation at the top of the barrier and compute the source and drain currents. The net current flowing through the device is the difference between these two currents.

In order to compare the I-V characteristics of InSb and Si MOSFETs, we first adjust their off-currents ($I_{off}$). This is done by modifying the fermi level of the source so that both devices give $I_{off}=0.71 \mu A/\mu m$. Table 3.1 gives the parameters of 2007 edition of ITRS roadmap for high performance logic technology requirements.

Figure 3.11 shows the injection velocities for InSb and Si. We can see that InSb has a much higher injection velocity than Si, especially at high bias points. As a result the carrier mobility in InSb is high. Figure 3.12 shows that Si has a much higher carrier density than InSb. This is because of the higher DOS for Si than for
InSb [39]. Current is the product of the carriers and velocity at the top of the barrier. We see InSb has a larger carrying capacity than Si mainly due to its higher injection velocity (Figure 3.13).

![Graph showing comparison of injection velocities for InSb and Si for $V_D = 0.8V$. InSb, due to its lower effective mass, has about 5 times larger injection velocity than Si. Simulations performed with a new tool [14].](image1)

**Fig. 3.11.** Comparison of injection velocities for InSb and Si for $V_D = 0.8V$. InSb, due to its lower effective mass, has about 5 times larger injection velocity than Si. Simulations performed with a new tool [14].

![Graph showing carriers at the top of the barrier for InSb and Si for $V_D = 0.8eV$. Si has almost twice the number of carriers due to its higher DOS than InSb. Simulation performed with a new tool [14].](image2)

**Fig. 3.12.** Carriers at the top of the barrier for InSb and Si for $V_D = 0.8eV$. Si has almost twice the number of carriers due to its higher DOS than InSb. Simulation performed with a new tool [14].
Fig. 3.13. $I_d - V_d$ for 2.5nm DG InSb and Si MOSFETs for $V_G = 0.6V$. $I_d$ is the product of injection velocity and carrier density at the top of the barrier. InSb has a much higher current than Si due to its high injection velocity even though Si has a higher DOS. The fermi level of the source was adjusted to achieve $I_{off} = 0.71 \mu A/\mu m$ for both devices. Simulation performed with a new tool [14].

3.6.1 High-$\kappa$ Dielectric materials

Since the inception of MOS technology, $SiO_2$ has been the used as the dielectric material. However, with reduction in transistor size it is becoming increasingly hard to reduce the oxide thickness. Reduction of oxide thickness causes an increased leakage current [40] that reduces the performance of the circuit. Materials with high dielectric constants [41] are being actively pursued to replace $SiO_2$ as the oxide material in nanoscale transistors. Figure 3.14 shows how current in the device is reduced when thickness of the oxide is increased. When we use materials with a high dielectric constant [42] the capacitance of the oxide increases which increases the current in the device. As a result, we can continue to work with higher oxide thickness which reduces leakage in the device.

$$I_{D,\text{Sat}} = \frac{W}{L} \mu C_{\text{inv}} \left( V_G - V_{th} \right)^2$$  \hspace{1cm} (3.9)
Figure 3.14 shows that the saturation current \((I_{d,sat})\) is almost 1.5 times higher for a high-\(\kappa\) material.

\[I_{d,sat}\]

\[V_D [\text{Volt}]\]

\[I_D [\mu \text{A/}\mu \text{m}]\]

\[V_G = 0.4 \text{V}\]

\[t_{ox} = 0.9 \text{nm}\]

\[SiO_2\]

\[\text{High-}\kappa\]

Fig. 3.14. \(I_d - V_G\) characteristics of a 2.5nm InSb UTB DG MOSFET. \(t_{ox} = 0.9\)nm, \(N_d = 1e18\)cm\(^{-3}\). Simulation performed with a new tool [14].

### 3.7 Summary

In this chapter, we initially explained the semiclassical technique for calculating charge. With ultra scaling of devices the quantum mechanical effects start appearing, especially near the oxide-semiconductor interface. For fairly large devices quantum domain can be restricted to a few nanometers from the interface, and the remaining region can be treated semiclassically. However, when the device thickness is extremely small, resulting in UTB DG-MOSFET’s we treat the entire region using tight-binding. Schottky barrier MESFET’s were briefly discussed. Current was calculated in these devices using the top of the barrier model. We analyzed the performance of InSb and Si based MOSFET’s and concluded that InSb based devices have a better current carrying capacity due to their higher mobility.
4. MODFET DEVICE SIMULATION

4.1 MODFET Theory

MODFET stands for MODulation Doped Field Effect Transistor. It is more popularly known as High Electron Mobility Transistor (HEMT). Figure 4.1 shows the schematic of an early HEMT structure. It consists of a barrier and a channel material. Due to the lower band-gap of the channel material, a triangular quantum well is created at the barrier-channel interface. Charges are induced due to quantum confinement.

An n-channel MODFET is created when semiconductor materials with different conduction band energies are arranged in such a manner that the material with a lower conduction band edge is sandwiched between materials with higher conduction band edge. By selectively doping the heterojunction, very high electron mobility is achieved in the channel. Moreover, the current inside the channel in a MODFET is considerably higher than a corresponding mosfet due to increased charge concentration. Using this device, high-speed circuits operating at relatively lower supply voltages can be realized that can be used for computation applications.

Figure 4.2 shows the conduction band profile of a HEMT for low and a high bias. The barrier material is doped with donor impurities. A 2-Dimensional Electron Gas (2DEG) is induced in the channel due to quantum confinement.

4.2 Modern day HEMT structure

Modulation doping is an essential characteristic of present-day MODFET. Figure 4.3 illustrates an InSb quantum well transistor with InSb channel sandwiched between AlInSb barriers. Modulation doping is achieved by selectively doping a few
atomic layers of the top AlInSb barrier with a donor impurity and leaving the remaining portion intrinsic. In literature this is often referred to as delta-doping. The region of AlInSb depleted of electrons forms a positive space charge region which is balanced by the electrons confined at the AlInSb-InSb interface. The resulting electric field perpendicular to the interface causes a severe band bending.

The electrons diffuse from the doped AlInSb region to InSb layer, where they are confined by the energy barrier creating a 2D electron gas (2DEG) at the channel-
barrier interface, and the concentration gradually reduces as we move deeper inside the well. Even though the ionized dopant atoms and electrons are spatially separated by the spacer layer, their close proximity allows an electrostatic interaction called Coulomb scattering. By setting the donors away from the interface, Coulombic scattering by donors can be reduced. In this work, we will limit the discussion to a ballistic device, and model the device without scattering.

The height of the barrier can be modulated by varying the concentration of Al in AlInSb. Higher concentration of Al creates a higher barrier. However, for higher Al concentrations there will be a lattice mismatch between AlInSb and InSb layers, resulting in a strained material. The bandstructure of a system under strain is quite different. It then becomes necessary to include strain effects in the calculations. Delta-doping and higher barriers prevent the problem of parallel conduction in the device. In this work, we will ignore the effects of strain in bandstructure.

Fig. 4.3. Left: Conduction band profile of present-day HEMT. The left AlInSb is delta-doped with donor impurities. The region between the delta-doped layer and the channel is called spacer region. Right: SEM image of a HEMT device.
4.3 From UTB to Heterostructure

4.3.1 Heterogeneous InSb quantum wells

InSb has a very high electron mobility and saturation velocity, and it is promising material for ultra-high speed field effect transistors that operate at very low supply voltages. Benchmarking InSb and other III-V semiconductors against Si MOSFET, we find that InSb has a lower gate delay and energy delay product [43]. Fabrication of InSb quantum wells though remains an important issue as seamless integration of the heterostructure with Si substrate [44]. Figure 4.4 shows the schematic diagram of InSb quantum well transistor. It consists of AlInSb barriers on either side. The entire device is grown on a semi-insulating GaAs substrate.

Fig. 4.4. InSb Quantum well transistor

4.3.2 Simulation approach for a Heterostructure

The most important concept that we discussed while simulating an UTB was the correct description of bandstructure. A heterostructure is an extension of an UTB by adding materials with higher bandgaps on either side. Usually the thickness of these layers are fairly large in the orders of few hundreds of nanometers. As a result, it is not really essential nor feasible to perform tight-binding calculation in these regions.
These large regions can be treated semiclassically with UTB region treated using tight-binding.

The figure 4.5 below shows the conduction band profile of a MODFET. The band-edges were obtained through semiclassical calculation.

![Graph showing effective mass bands and conduction band profile](image)

**Fig. 4.5.** Effective Mass Bands of AlInSb/InSb system. Simulation performed with a new tool [14].

Let us now look at the step by step procedure for simulating a heterostructure using semiclassical and tight-binding methods simultaneously.

- Perform a semiclassical calculation in all the regions of the device self-consistently with Poisson’s equation. Obtain the converged electrostatic potential. This is the initial guess for potential.
- Identify the tight-binding domain and compute UTB bandstructure in that domain alongwith the potential obtained from the semiclassical calculation.
- Calculate charge using UTB bandstructure self-consistently with Poisson’s equation. Current is calculated using top of the barrier model.

Figure 4.6 shows an illustration of different simulation domains in a heterostructure. Figure 4.7 shows the band-edges using tight-binding calculations.

Figure 4.8 and ?? represent the transfer and output characteristics of the device.
Fig. 4.6. Different simulation domains in a heterostructure

Fig. 4.7. Top: Band-edges for 5nm InSb channel sandwiched between AlInSb barriers. Bottom: Electron density Vs device cross-section. Simulation performed with a new tool [14].
Fig. 4.8. Output \((I_d - V_d)\) characteristics of 5nm InSb channel sandwiched between AlInSb barriers. The top barrier was \(\delta\) doped with a sheet doping of \(1e12cm^{-2}\). The other regions in the device are intrinsic. The \(I_{off}\) was fixed at 0.71 \(\mu A/\mu m\) by adjusting the source fermi level [36]. As the gate voltage is increased, more carriers are transferred into the channel and the source-drain current increases. Simulation performed with a new tool [14].

4.4 Summary

In this chapter, we applied the important concepts learnt in the previous chapters for simulating a 1D Heterostructure device. The device under investigation was AlInSb/InSb single-channel MODFET. Semiclassical and tight-binding charge calculation were performed in relevant regions inside the device by breaking down the entire device into structure into different simulation domains. We computed the output and transfer characteristics of the device.
5. 1D HETEROSTRUCTURE SIMULATION TOOL

In this chapter, we will look at the features of the 1D Heterostructure simulation tool [14], that should prove beneficial for experimentalists and theoreticians alike. We use the $sp^3s^*$ tight-binding model to model the bandstructure of the materials. The tool offers capability for simulating MOSFETs and Heterostructures. The support provided by nanoHUB.org computing resources operated by Network for Computational Nanotechnology, funded by NSF, are acknowledged.

5.1 Simulation of MOScap

In this section, we will see the simulation of a MOScap. The tool computes the electrostatics of single-gate and dual-gate MOSFET structures. In charge calculations the use of closed boundary conditions implies that the wave-function should decay at the boundary, while using open boundary conditions allow an exchange of wave-function probability with the contacts. Figure 5.1 shows the potential and charge profile of GaAs MOScap. Applying a positive gate voltage causes band-bending, creating a triangular quantum well near the oxide-semiconductor interface. Quantum mechanical calculations are applied in this region where charge is confined. The flat regions are treated using the semiclassical electrostatics. The distinction comes from the discrete confined states as opposed to continuum of states. The discrete energy states are also depicted in the potential profile.

The charge is calculated from DOS obtained from the Green’s function technique. The levels are populated using a Fermi level determined by the doping in the substrate. Since the gate is present on the left side we use closed boundary conditions such that the wave-function dies out. On the right side we use open boundary conditions to model the at band conditions in the substrate. Figure 5.2 shows confined states lying
(a) Potential profile of a GaAs MOScap with resonant states

(b) Charge Vs Distance

Fig. 5.1. Output screenshots of 30nm GaAs MOScap with an applied bias of 0.5V.

inside the triangular potential well. Features in the DOS above the confined states can also be seen, as waves piling against the semiconductor oxide barrier, which is a signature of continuum states injected from the open boundary on the right. Outside the tight-binding region, it is sufficient to calculate the charge semiclassically since there are no confined states. Finally, the charge calculation has to be done self-consistently with the electrostatic potential which accounts for the free and movable charge. We model the oxide at the oxide semi-conductor interface as a region with a linear drop of potential, since we assume that there is no charge inside the oxide. On the substrate side, we let the potential approach flat band conditions.
5.2 Simulation of Heterostructure

Quantum well heterostructures are created when material with lower band-gap is sandwiched between high band-gap materials. The quantum well holds the charge carriers and is known as the channel. Depending on the number of channels, the device is called single-channel, double-channel and so on. Due to a high quantum confinement inside the channel, discrete energy levels are created which are occupied by the carriers. Application of gate, source and the drain voltages causes current flow inside the device.

In this section, we will look at the simulation of Modulation Doped FET (MOD-FET), a heterostructure in which the barrier material is delta-doped. The doping of the channel is intrinsic in order to prevent scattering of carriers, that can limit the current inside the device. The barrier material induces carriers that are swept
across the channel by the source and drain potentials. Figure 5.3 shows the conduction band profile and charge density of a MODFET structure with different gate voltages. The channel region is treated using tight-binding technique and the barrier materials are treated semiclassically. As the gate bias is increased, the energy bands move down. The fermi level populates the states in the channel. Figure 5.4 shows the characteristics of a double-channel MODFET device.

Charge in the confined states for each well is calculated by constructing a Hamiltonian using the well material with closed boundary conditions on both sides. The bandstructure of the geometry is then calculated based on the eigenvalues of the Hamiltonian. The eigen-vectors give the spatial variation of charge inside the wells. The eigen-values in this case are real as opposed to the case when we use open boundary conditions. In the case of open boundary conditions the eigen-values are complex, where the imaginary part corresponds to the lifetime of the state. The quantum mechanically computed charge replaces the charge which was previously computed.
These calculations are done self-consistently with the electrostatic potential.

### 5.3 Simulation options

To accurately capture the physical quantities in the device at a reasonable computational cost, a number of numerical techniques have been used in the simulator. To integrate over the DOS, an adaptive energy grid generation technique is used. The technique allows the addition of energy nodes in the integration range where the integrand varies rapidly. This is achieved by successive partitioning. This ensures that integration nodes are only added as required to achieve a certain level of accuracy. If the resonances are weakly coupled to continuum reservoir states, their lifetime may be extremely long, resulting in energetically very sharp resonances. It may be extremely expensive to resolve these resonances through adaptive refinement. We, therefore,
also implemented a resonance finder, which determines the eigen-energy and the lifetime or width of the state and therefore allows for the resolution of the resonances with very few energy nodes. The problem of finding resonances is non-linear because of energy dependent open boundary conditions. The idea is to find good starting guess values to resonances using Shift and invert non-symmetric Lanczos [45] for the Newton technique to locate the values accurately.

The simulator gives the user access to a large material database based on effective mass and the $sp^3d^5s^*$ model. The user can design the device in great detail. Figure 5.6 shows the various dimensions necessary to make a MODFET. In addition to that the doping profile inside the device can also be specified.
Fig. 5.6. Input and Simulation options
6. SUMMARY AND FUTURE WORK

6.1 Summary

In this work, we have demonstrated a 1D Heterostructure simulation tool [11–13]. In this first chapter, we discussed about the semiconductor bandstructure. Using the effective mass Schrödinger equation, we obtain a parabolic E-k relation. "Real" devices, however, have non-parabolic bands and states outside the Γ valley along different directions in the crystal. Using tight-binding, we can provide an accurate description of the bandstructure. We briefly explained the Hamiltonian construction using $sp^3d^5s^*$ tight-binding model and discussed the bandstructure of some common semiconductors. The tight-binding model accurately models the bandstructure, especially at the nanometer scale where the material description using effective mass model is incorrect because of the non-parabolic bands. Also, a brief discussion on spin-orbit coupling was presented.

In chapter 2, we applied the concept of bulk tight-binding theory for modeling Ultra Thin Body (UTB) devices. We discussed the construction of the 2D Hamiltonian and the kind of boundary conditions that was applied. By applying Hardwall boundary conditions, it is ensured that the wave-function of the carriers is confined inside the channel. Moreover, due to the hardwall boundary, dangling bonds are present at the surface that are passivated by rising the energy of the surface atoms. We then reviewed the work on Ge n-MOSFETs. It became evident from those results that non-parabolic effects in the bandstructure results in a threshold voltage shift, which leads to a different $I_d - V_G$ characteristics inside the device.

InSb is a promising material for III-V planar transistors due to its high mobility. It offers lower gate delay and energy-delay product than Si MOSFETs and CNTFETs for high-speed logic circuits. It is pursued with great interest in the scientific community.
However, due to its inherent low band-gap, InSb devices suffer from band-to-band-tunneling (BTBT), which increases the leakage current inside the device. Moreover, the scalability of InSb needs to be demonstrated for channel lengths less than 100nm. We investigated some of the important bandstructure effects in InSb. When the channel thickness is scaled, appreciable splitting conduction sub-bands takes place. Also, the bands shift up higher in energy. The curvature of the bands also increases, leading to a higher effective mass. InSb and other III-V materials show promise to outperform Si and Ge MOSFETs due to their high mobility, which could play an important role when the source and drain resistance and other parasitics in ”real” devices become important.

In chapter 3, we saw how confined and continuum charge were treated separately in the device. Due to band-bending in UTB DG n-MOSFETs, triangular quantum well is created at the oxide-semiconductor interface. This confinement creates discrete energy states and are modeled using tight-binding. The flat potential region is treated semi-classically. The charge is calculated self-consistently. Using top of the barrier ballistic model, the source-drain current is calculated. This model is then applied for calculating the $I_d - V_d$ characteristics of InSb and Si UTB n-MOSFETs. InSb MOSFETs, due to their high injection velocity, have a greater current flow inside the device even though DOS for InSb is lower than Si. In chapter 4, HEMT devices were simulated. We briefly discussed about the early HEMT structure, its drawbacks and how it was overcome by using modulation doping in the barrier layers. Here once again, we treat the confined charge (channel) using tight-binding and barrier layers using semi-classical calculation.

Finally, we discussed about a 1D Heterostructure tool that offers the capability for simulating MOSFET and Heterostructures. The bandstructure is modeled using $sp^3d^5s^*$ tight-binding model. Fast algorithms have been implemented in order to compute the resonance width with fewer energy nodes. With a wide range of material systems and advanced simulation models, the tool should prove beneficial to
experimentalists and theoreticians. The tool was deployed on nanoHUB.org operated by Network for Computational Nanotechnology, to serve the scientific community.

6.2 Future work

- **Strain effects in bandstructure:** Strain effects can significantly alter the bandstructure. Moreover, strain can also affect the carrier mobility in the material. Due to lattice mismatch in the layers of a heterostructure, strain effects need to be incorporated into the calculations.

- **Parallelization of the 1D Heterostructure code:** Running tight-binding calculations for large devices can be extremely time consuming even with the use of advanced numerical methods. The tight-binding code can be run in parallel. There are three levels of parallelization: *Energy, k* space and *Bias*.

- **Calculation of leakage current:** Scaling of the supply and threshold voltages of the device, results in an exponential increase in the sub-threshold leakage current. In the current process technology of sub-22nm physical gate length, reduction in device leakage current is a major challenge. Modeling leakage current in these devices is important and challenging.
LIST OF REFERENCES
LIST OF REFERENCES


[13] “Personal communication and collaboration with Neophyto Neophytou.”

[14] “Code to be released in nanoHUB.org named either 1D-HETERO. The code is based on reference [11–13].”


[33] S. Datta, Quantum Transport: From Atom to Transistor. Cambridge University Press.


