Title
Electrical Transport in Nanoscale Semiconductors: A Quantum Transport Approach

Author
Kim, Sangwook

Publication Date
2016

Peer reviewed|Thesis/dissertation

Electrical Transport in Nanoscale Semiconductors: A Quantum Transport Approach

by

Sangwook Kim

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Materials Science and Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ramamoorthy Ramesh, Chair
Professor Sayeef Salahuddin
Professor Daryl C. Chrzan

Summer 2017

Electrical Transport in Nanoscale Semiconductors: A Quantum Transport Approach

Abstract

Electrical Transport in Nanoscale Semiconductors: A Quantum Transport Approach

by

Sangwook Kim

Doctor of Philosophy in Engineering - Materials Science and Engineering

University of California, Berkeley

Professor Ramamoorthy Ramesh, Chair

As predicted by the International Roadmap for Semiconductors, III-V metal-oxide semiconductor field-effect transistors(MOSFETs) are the prime candidates for the 7-nm node and beyond owing to their high mobilities. However, several challenges need to be overcome before III-V materials can replace silicon in extremely scaled devices. As the size of semiconductor devices enters into the nanoscale regime, it is important to understand quantum mechanical effect in electron transport because quantum mechanical effects such as microscopic scattering, quantum mechanical tunneling, carrier confinement and interference effects play a crucial role. We have investigated quantum transport properties of III-V and silicon in nano-scale devices using the non-equilibrium Greens function (NEGF) formalism coupled with a 20 orbital $sp^3d^5s^*$-SO tight-binding model. A mode space NEGF approach combined with the full-band WKB model has been introduced to calculate band-to-band tunneling current under the valence band where an atomistic full-band NEGF approach provides a zero current. Self-energy functions for carrier scattering have been implemented within the mode space NEGF formalism to instigate the influence of microscopic scattering effects on quantum transport. This approach has been benchmarked by comparing the scattering rates obtained from Fermis golden rule which is widely used in traditional semi-classical calculations. The scattering model is used to study the role of phonon scattering and surface roughness scattering on the carrier transport characteristics of Si and III-V channel devices. In the presence of the electron-phonon interactions, the drain current decreases compared with its ballistic limit, and the current reduction ratio increase as the channel length increases. Effects of indium mole fraction and source/drain doping density (N_{SD}) on the performance of nanoscale III-V MOSFETs are explored and their performance is compared with that of Si transistors. For III-V's, we have found an optimum indium mole fraction and N_{SD} that maximize the Ion/Ioff ratio and Ion by balancing injection velocity and short channel effect.

To my family

Contents

List of Figures

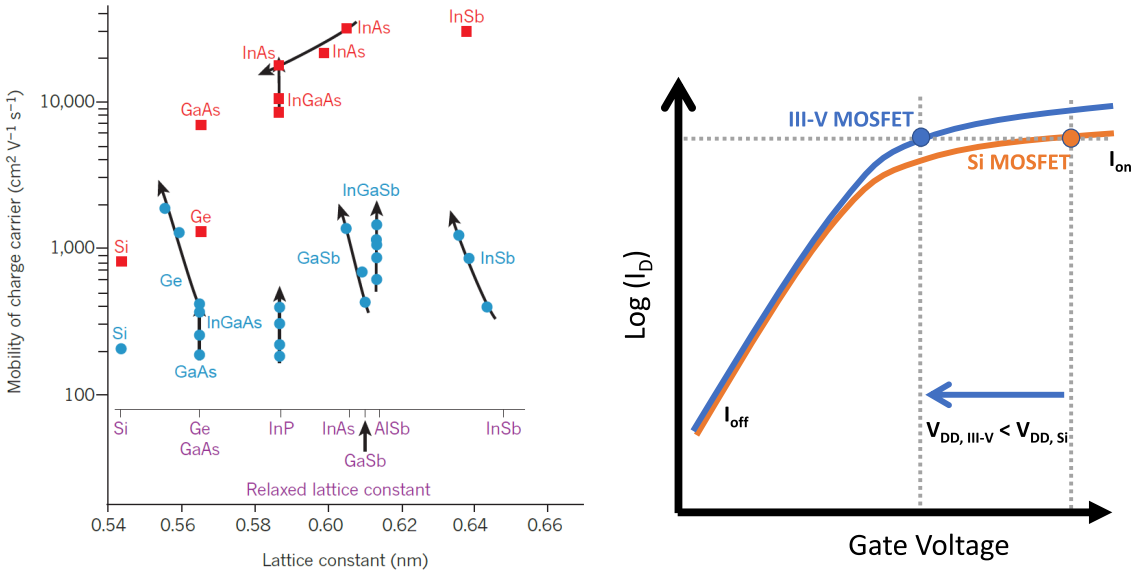
List of Tables

Chapter 1

Introduction

1.1 Overview

Outstanding progress in information technology is attributed to the development of the silicon-based complementary metal oxide semiconductor (CMOS) field-effect transistor(FET) technology [1]. For more than 40 years, continuous transistor scaling has continued pushing towards smaller and smaller devices in an effort to obtain higher computing power and lower power consumption of integrated circuits. However, as scaling of metal-oxide-semiconductor FETs (MOSFETs) approaches the 10 nm-node, it faces fundamental technical issues regarding power consumption, fabrication (especially lithography), physics and chip costs [2, 3]. Among them, high power consumption owing to high leakage current and saturated supply voltage at around 0.8 - 0.9 V is the largest problem of Si based CMOS technology [4]. In order to continue the scaling of MOSFETs and maintain the historic progress in information processing and transmission, new strategies such as innovative device structures, new materials and new device operation mechanisms are required. Among the various candidate material for the next-generation CMOS technology, III-V compound semiconductors have found significant interest due to their high mobility [5, 6] (Fig. 1.1a). Since the On-current for nanoscale FETs is determined by the product of the injection velocity and the charge density, the high injection velocity of III-V materials could provide high on-current at a lowered supply voltage compared Si FETs (Fig. 1.1b). Since experimental research of various III-V materials in nanoscale MOSFETs is extremely vast and complex, modeling and simulation could help predict important design guidelines and understanding. So, the objective of this dissertation is to develop a simulation framework capable of handling quantum transport to understand and exploit the electrical transport properties of nanoscale semiconductor devices. We also include effects of scattering in the simulation framework to study the influence of microscopic scattering mechanisms. We compare the electrical performance of III-V materials with Si MOSFETs for future technology nodes. In section 1.2, the motivation for III-V devices, current research and challenges are presented. In Sec. 1.3, quantum transport models based on Non-equilibrium Greens function (NEGF) approach are reviewed. In Sec.



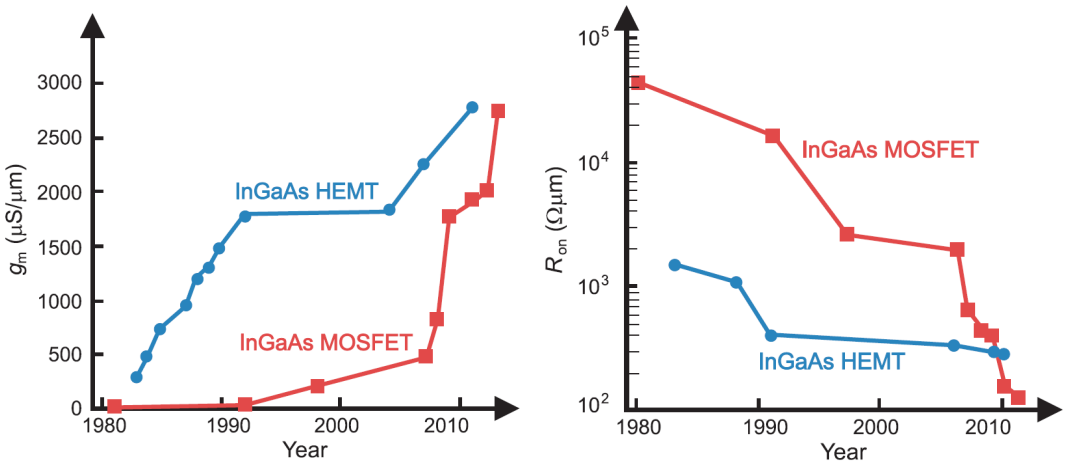
(a) (b)

Figure 1.1: (a) Room temperature mobility of electrons (red) and holes (blue) in inversion layers and quantum wells is shown as a function of the actual semiconductor lattice constant. Ref. [5]. (b) Qualitative comparison of the influence of the channel materials. To achieve the same I_{on}, the III-V MOSFET can be operated at a lower $V_{DD,III-V}$ than $V_{DD,Si}$ of the supply voltage for Si MOSFET.

1.4, the outline of this thesis is presented.

1.2 Nanometer-scale III-V MOSFETs

For many years, III-V compound semiconductors have been used as channel materials for high electron-mobility transistors (HEMTs) for high-speed and high-frequency applications [5, 7]. For logic applications, however, the mature HEMTs intrinsically suffer from high gate leakage because they use modulation doping in heterojunctions instead of using the metal-oxide-semiconductor structure [8, 9]. On the other hand, there are two crucial issues of InGaAs MOSFETs to achieve required performance: Fermi-level pinning and the high parasitic resistance. The high interface trap density (D_{it}) caused by non-ideal bonding or other defects at the semiconductor-oxide interface results in 'Fermi-level pinning' which prohibits modulation of the electrostatic potential inside the semiconductor and generates stability issues [10, 11]. So, research on III-V MOSFETs has been focused on improving the channel-dielectric interface. Unlike Si-SiO_2 interfacial states that can be passivated by hydrogen, compound semiconductor surface states are not affected by such an approach. Especially, oxidation of GaAs in an air environment generates near mid-gap interface state.



(a) (b)

Figure 1.2: Performance comparison of inversion-type InGaAs MOSFETs and HEMTs versus year. (a) Transconductance g_m and (b) ON-resistance R_{on}. Ref. [7].

Density functional theory (DFT) simulations showed that surface oxidation in GaAs can create non-ideal bonding such as As-As bonds, As dangling bonds, Ga vacancies resulting in increased D_{it} [12, 13]. In 1995, Passlack, M. *et al.* showed that a low D_{it} was achieved by in situ deposition of $Ga_2O_3(Gd_2O_3)$ on GaAs with the molecular beam epitaxy (MBE) technique [14]. This result suggested that III-V MOSFETs with unpinned Fermi level were indeed possible and led many researchers to fabricate high quality dielectric/III-V interfaces and understand Fermi-level pinning. Major progress occured in 2003 by using the atomic layer deposition (ALD) technique. Ye, P. D. *et al.* showed that *ex situ* Al_2O_3 deposition on GaAs with ALD also could reduce D_{it} [15]. This result was unexpected because *ex situ* method usually creates a low-quality native oxide when GaAs exposed to the air. It was found that most native surface oxides are eliminated during the early stages of the ALD process, which is so-called "self-cleaning effect" [16]. Furthermore, subsequent high-k dielectric deposition does not regrow the GaAs oxides. However, according to computational simulation and analytical characterization, the interface between GaAs and ALD high-k dielectrics still had several interfacial defects such as As-As dimer and Ga dangling bonds, and their high defect density deteriorated device performance [13,17,18]. In order to deal with this problem, there were various approaches such as pre- or post- deposition treatments, the use of interfacial layers, modifications of the deposition chemistry and *in situ* ALD [19–22]. Another challenge to further improve III-V MOSFET performance is to reduce parasitic resistance. First, silicide-like contacts have been investigated because they are promising for device integration. Ni, Co, or Pd metals were alloyed with InGaAs at relatively low temperature [23, 24]. However, this approach showed inferior contact resistance values when compared with silicide contacts. When InAs composition of InGaAs is higher than 70 %, the

Fermi level location pins close to the conduction band edge resulting in very good contacts with small contact resistances [25, 26]. The ohmic contacts to n+ InGaAs using a refractory metal such as Pd, Mo, and W also show low contact resistivity of around 1-2 $\Omega\mu m^2$. By the efforts mentioned above, it has been observed that the InGaAs MOSFET characteristics improved significantly as shown in Fig. 1.2 [7].

1.3 Quantum transport based on NEGF

Electron transport in semiconductors has received much attention since semiconductor devices operate by controlling the flow of the electrons and holes through a device. Classically, electrons and holes are treated as semi-classical particles with an effective mass in solids, and transport properties of electrons in semiconductors were understood by the drift-diffusion (DD) equation which is represented by $J = nq\mu_n E + qD_n\nabla n$ [27]. When the device size is smaller than 1 μm, the DD equation cannot predict non-local and hot-carrier effects owing to rapid change of large electric field inside the semiconductor. So, it is necessary to use the hydrodynamic model, which includes an additional balance equation for the average carrier energy and modifies the current expression to be proportional to the gradient of the carrier temperature. When the device size is smaller than 100 nm, moment-based solutions to Boltzmann transport equation (BTE) start to lose validity with increase in electric field strength and could result in spurious velocity peaks due to truncation of moments [28]. Therefore, we have to rely on the microscopic transport models. Within semi-classical approaches, the particle-based Monte Carlo method or solving the full-blown BTE itself are

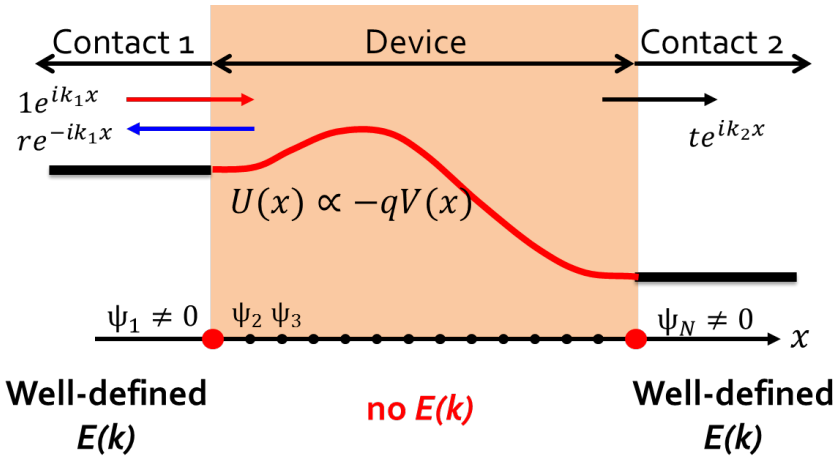


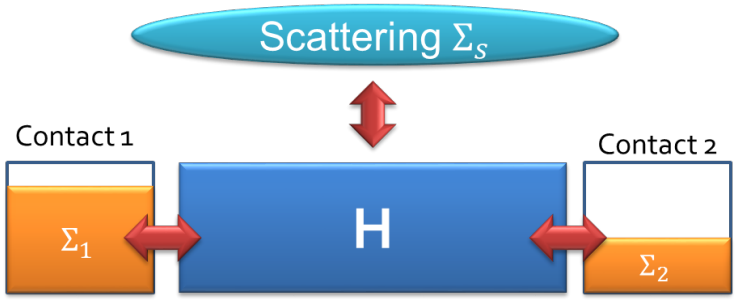
Figure 1.3: Schematic of a one-dimensional device connected to two semi-infinite contacts. Since the periodicity is broken by potential changes in the device, the wave function in the device should be obtained by solving through NEGF formalism instead of eigenvalue problem posed by the Schrödinger equation.

the most popular approaches [29]. However, with the device size close to the thermal de Broglie wavelength ($\lambda_B = h/\sqrt{(3m^* k_B T)} \approx 10\ nm$ for Si), the Boltzmann transport equation, which treats electrons as semi-classical particles, is generally not valid anymore because estimating the particle's position exactly is impossible. Furthermore, in order to understand the electronic transport problem properly, transport models should have the ability to capture quantum effects such as interference effects, quantum mechanical tunneling and microscopic scattering. Therefore, accurate and efficient simulation tools based on a quantum mechanical formalism are necessary to interpret experimental results. All quantum transport models are based on the Schrödinger equation, but directly solving the Schrödinger equation for a many particle system is almost impossible. So, several methods have been introduced to address this problem [30]. Among them, the Non-equilibrium Greens function method appears to be appropriate for treating quantum transport in nanoscale semiconductors. The NEGF formalism was developed in the 1960s by the works of Keldysh [31] and others [32–34] using the methods of many body perturbation theory (MBPT) to solve nonequilibrium problems in statistical physics. In the early 1970s, the general NEGF formalism for device current was first presented [35, 36]. The partitioning of an infinite system into device, and contacts, and the derivation of the open boundary self-energies was presented in Ref. [37]. Datta, S. introduced a different approach starting with the one-electron Schrödinger equation, which is used by many in the nanoelectronics community [38, 39]. Within his approach [40], the NEGF approach starts from the time-independent Schrödinger equation $E\{\psi\} = H\{\psi\}$ and add two terms representing the outflow and inflow from the contacts as $E\{\psi\} = H\{\psi\} + \Sigma\{\psi\} + \{S\}$. These two terms arise from imposing open boundary conditions on the Schrödinger equation with an incident wave from the contact. Using this modified Schrödinger equation, the wave function can be written as $\{\psi\} = [EI - H - \Sigma]^{-1}\{S\}$. The Greens functions can be defined as $G = [EI - H - \Sigma]^{-1}$. Device Hamiltonian H consists of kinetic energy (H_0), reperesented in some basis of choice and potential energy (U). The Greens function gives the response of a system to a perturbation $\{S\}$ in the Schrödinger equation. So, this NEGF approach is identical to solving the Schrödinger equation for open quantum system.

In the NEGF formalism, the microscopic scattering processes can be taken into account by self-energies which describe the potential felt by a carrier due to interaction with the phonon bath [39]. Fig. 1.4 shows the schematic NEGF model with coupling to the phonon bath. In the self-consistent field (SCF) regime, U in the Greens function can be calculated through the self-consistent solution of the NEGF equations and Poisson's equation. The calculation procedure of this type of simulation using the NEGF formalism is summarized in Fig. 1.5. From converged U, charge density and electrical current relations can be obtained from the Greens function and self-energy of the contacts by using relations of $\rho \sim \{\psi\}\{\psi\}^\dagger$ and $I = q\dfrac{dN}{dt} = qTr\left[\dfrac{d}{dt}\psi\psi^\dagger\right]$ [41].

1.4 Outline of the dissertation

This dissertation is organized as follows. Chapter 2 describes necessity of atomistic modeling for bandstructure and certain aspects related to using the tight-binding method including the history of model development and the assumptions. The choice of an appropriate basis set to accurately predict the full bandstructure for nanoscale semiconductors is discussed. We show how the bandstructure changes when we confine a system within an ultra-thin-body (UTB) geometry. Then, the non-parabolic effect of III-V semiconductors in the nanoscale structure is explored. In Chapter 3, we show that commonly used quantum transport calculations using the NEGF method could significantly underestimate the band-to-band tunneling current in III-V MOSFETs. We show how failing to account for proper boundary conditions could lead to insufficient filling of the body in a MOSFET under a typical Off condition. We propose a numerically efficient method to include this leakage current. In chapter 4, we present a 2D simulation framework for quantum transport including phonon and surface roughness scattering. The scattering-NEGF framework is benchmarked with Fermi's golden rule. Device simulation results of InAs UTB MOSFETs using NEGF with scattering models for different device sizes are presented. The scattering effect on quantum transport in III-V MOSFETs is analyzed and the scattering-limited low-field mobility and mean free path for different scattering mechanisms are calculated. Chapter 5 looks into the bandstructure effect and the effect of phonon scattering in Si UTB MOSFETs. We also obtain the phonon-limited low field mobility by calculating total resistance with respect to different gate lengths. In chapter 6, we investigate $In_xGa_{1-x}As$ short channel MOSFETs within the scattering-NEGF framework. The effects of indium mole fraction and source/drain doping concentration are explored. We also perform benchmarking against the same size Si MOSFET. A summary and possible future work are discussed in chapter 7.



$$G = [EI - (H_0 + \Sigma_s + U) - \Sigma_1 - \Sigma_2]^{-1}$$

Figure 1.4: Schematic of NEGF model with coupling to the phonon bath.

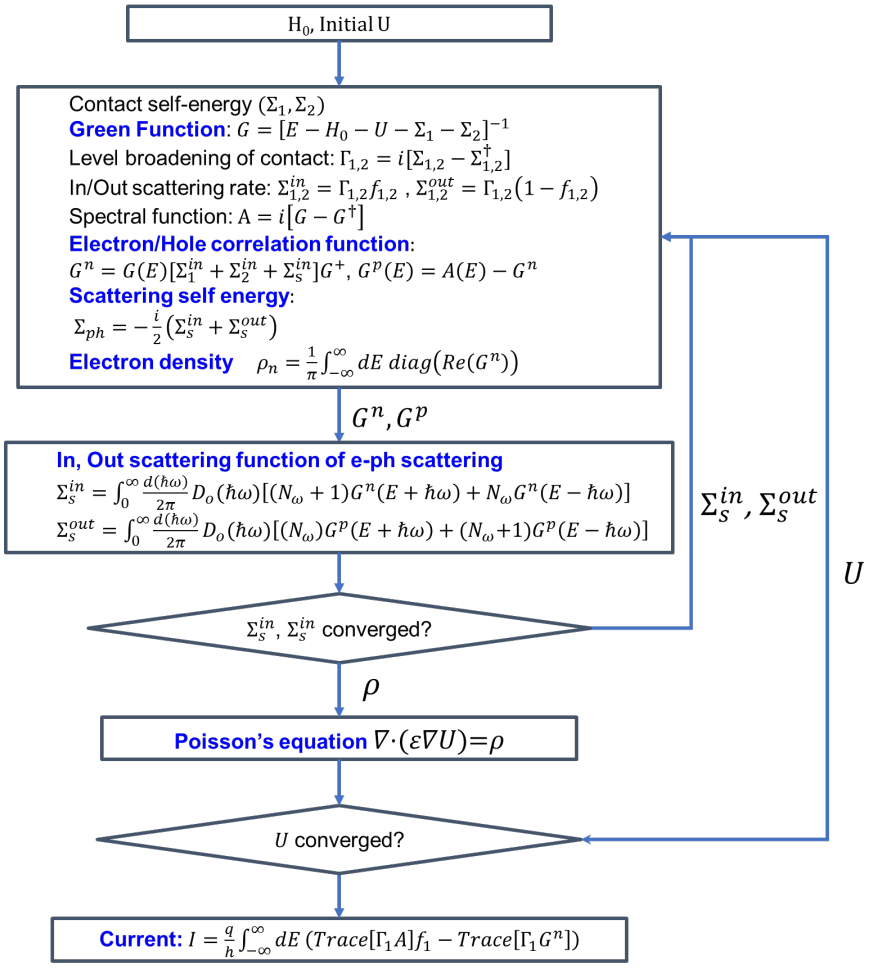


Figure 1.5: Flow chart illustrating the simulation procedure of the quantum transport including scattering.

Chapter 2

Electrical bandstructure of nanoscale semiconductors

2.1 Introduction

The microscopic behavior of carriers in semiconductors is governed inherently by their bandstructure. Real part of bandstructure provides information on how the carriers are distributed in the energy and momentum space and how the carriers react to the external electric fields. The imaginary part of bandstructure gives the rate at which wavefunction of an electron decays within the bandgap [16]. The effective mass approximation (EMA) with parabolic E-k relationship has been widely employed both in the semiclassical transport model [42] [43] and in some studies employing a quantum transport treatment [44] [45] due to the computational efficiency. However, in treating nanoscaled structure such as thin films and nanowires, the traditional EMA starts to lose validity. The quantum size effect along the confined direction generates discrete subbands with higher energies than the bulk band edges and the bandstructure non-parabolicities increase the effective mass with decreasing channel thickness (especially for the III-V semiconductor). Furthermore, the single band model does not model the imaginary part of bandstructure that helps us to understand the tunneling process of carriers between the conduction and valence bands. Thus, more accurate electronic bandstructure calculation methods should be utilized to satisfy several requirements to accurately capture nanoscale device physics. Among the variety of bandstructure calculation methods, we used a semi-empirical atomistic tight-binding approach with 20 orbitals consisting of an $sp^3d^5s^*$ basis with spin-orbit coupling. This model is a powerful atomistic model which can provide information about atomistic and orbital distribution of charge in the semiconductor, non-parabolicity, subband energy levels as functions of quantization, accurate E-k dispersion with different transport and growth orientation,and complex bandstructures of nanoscale semiconductors [46] [47].

In this chapter, we will describe the features of the $sp^3d^5s^*$ nearest neighbor TB model and show how the bandstructure changes when we confine a system within thin film. Then

we will study the non-parabolic effect of III-V semiconductor in the nanoscale structure. Finally, we discuss spin-orbit coupling and its effect on the energy levels in semiconductors

2.2 Semi-empirical Tight-binding model

There are two main categories of electronic band structure calculation methods: ab-initio methods and semi-empirical methods. While the ab-initio methods such as Hartree-Fock or Densitiy Functional Theory (DFT) are first principle calculations that do not need empirical fitting parameter, the main drawback of ab-initio method is that they are computationally expensive for even simple properties such as the electronic band gap. In contrast to ab-initio method, the semi-empirical methods involve empirical parameters to fit experimental data. They are easier to implement but their accuracy depends upon the number of fitting parameters. The most popular methods of semi-empirical bandstructure calculation are the empirical pseudopotential method, the k·p method and the tight-binding (TB) method. Among the variety of semi-empirical bandstructure calculation methods, the TB method has received considerable attention because it offers an intuitively simple accurate description of the bandstructure in terms of chemical bonds and is appropriate for treating disorder and strain at the atomic level. Contrary to pseudopotential approach, TB method starts with assumption that the electrons are tightly bound to their nuclei as in the atoms. When separations between atoms are close enough in solid, wave functions of electrons overlap and can be approximated by linear combinations of the atomic wave functions. So, the eigenstates of Hamiltonian can be expanded in a linear combination of atomic-like orbitals (LCAO). In this method, the bandstructure can be defined in terms of overlap parameters and these overlap parameters represent interactions between electrons on adjacent atoms [48]. When we expand the crystal Hamiltonian in the Bloch sum basis, there are three types of Hamiltonian matrix elements: 1. On-site matrix element: Both orbitals and the potential are located on the same atom. 2. Two-center matrix element (hopping integral): Two orbitals located on different atoms and the potential is on one of these two atoms. 3. Three-center matrix element: Each of the two orbitals and the potential are located at different atoms. Slater and Koster [49] showed that general form of matrix element can be simply expressed with two-center overlap energies for σ, π, or δ type and direction cosines of the displacement vector between nearest atoms by assuming three-center matrix is much smaller than two-center matrix element (two-center integral assumption).

In this work, we will not look at the mathematical details for the formalism of the tight-binding theory. Readers interested in gaining further knowledge on the TB theory should consult Ref. [39] and [48] for the detailed mathematics. Eventually, from Bloch's theorem, considering any particular lattice point n connected to its neighboring point m with a coupling matrix $[H_{nm}]$, the bandstructure can be calculated by solving a matrix eigenvalue equation of the form

$$E(\phi_0) = [h(\vec{k})](\phi_0) \tag{2.1}$$

where

$$[h(\vec{k})] = \sum_{m} [H_{nm}] e^{i\vec{k}\cdot(\vec{d}_m - \vec{d}_n)} \tag{2.2}$$

The size of matrix $[h(\vec{k})]$ is $(b \times b)$, where b is the number of basis orbitals per a primitive unit cell. The summation over m in Eq.2.2 runs over all neighboring lattice points including itself. For a simple example, we calculate the band structure of graphene sheet. The carbon atoms of graphene layer are arranged in a hexagonal pattern as shown in the Fig. 2.1. The two carbon atoms A and B are the crystal basis of primitive unit cell of graphene. The periodicity of each lattice point can be described by $\vec{R} = m'\vec{a}_1 + n'\vec{a}_2$ where m'and n'are integers, $\vec{a}_1$ and $\vec{a}_2$ are the unit vectors which are defined as

$$\vec{a}_1 = a\left(\frac{3}{2}, \frac{\sqrt{3}}{2}\right), \vec{a}_2 = a\left(\frac{3}{2}, -\frac{\sqrt{3}}{2}\right) \tag{2.3}$$

where $a = 2.46\mathring{A}$ is the lattice constant of graphene.

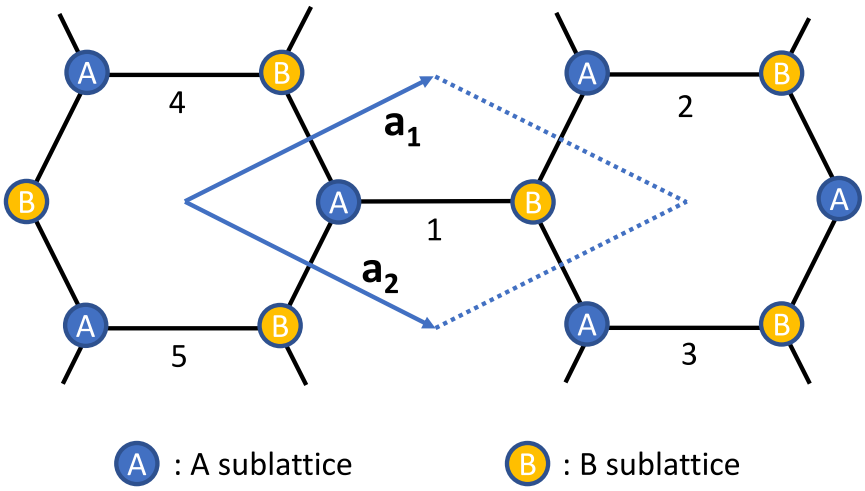


Figure 2.1: Schematic of a graphene's hexagonal lattice. Each Bravais lattice unit cell includes two nonequivalent sites, which are denoted by A and B. For the unit cell 1, there are four neighboring unit cells (from 2 to 5).

Then, $[h(\vec{k})]$ is determined from Eq.2.2 as

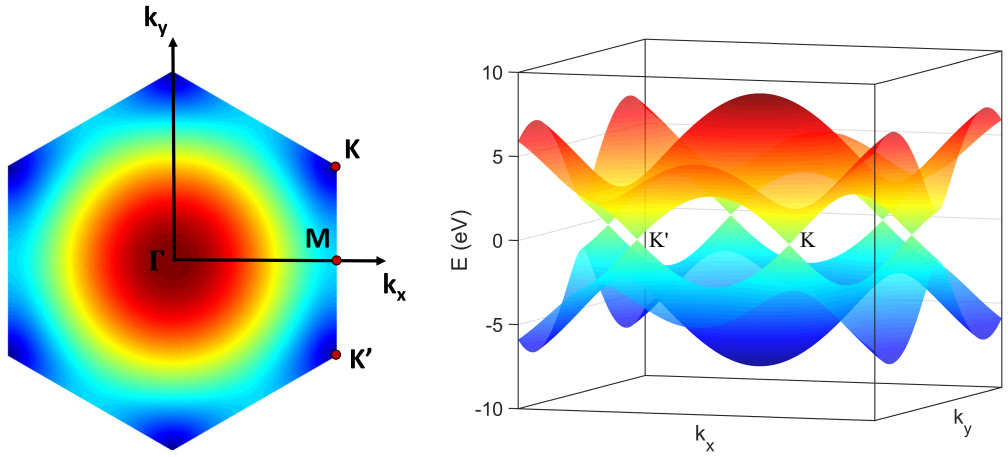
$$[h(\vec{k})] = H_{11} + H_{12}e^{i\vec{k}\cdot\vec{a}_1} + H_{13}e^{i\vec{k}\cdot\vec{a}_2} + H_{14}e^{i\vec{k}\cdot(-\vec{a}_2)} + H_{15}e^{i\vec{k}\cdot(-\vec{a}_1)} \tag{2.4}$$

where H_{11} is the on-site block Hamiltonian of the unit cell 1 and H_{1x} are the coupling block Hamiltonian between the unit cell 1 and x. Since the levels involving $2s$, $2p_x$, $2p_y$ orbitals

are far away from the Fermi energy level due to the sp^2 hybridization, the conduction and valence band states can be described only with one orbital ($2p_z$ orbital) [50]. Then, the Hamiltonian matrix can be simply written down with the overlapping energy of p_z ($-t$),

$$[h(\vec{k})] = \begin{bmatrix} 0 & -t \\ -t & 0 \end{bmatrix} + \begin{bmatrix} 0 & -t \\ 0 & 0 \end{bmatrix} e^{i\vec{k}\cdot\vec{a}_1} + \begin{bmatrix} 0 & -t \\ 0 & 0 \end{bmatrix} e^{i\vec{k}\cdot\vec{a}_2} + \begin{bmatrix} 0 & 0 \\ -t & 0 \end{bmatrix} e^{i\vec{k}\cdot(-\vec{a}_2)} + \begin{bmatrix} 0 & 0 \\ -t & 0 \end{bmatrix} e^{i\vec{k}\cdot(-\vec{a}_1)} \quad (2.5)$$

By calculating eigenvalues of $[h(\vec{k})]$ with different 2D momentum(k_x, k_y), the energy dispersion of graphene can be shown as Fig. 2.2.



(a) (b)

Figure 2.2: (a) Conduction bandstructure through the whole region of the Brillouin zone. The coordinates of high symmetry points are $\Gamma = (0, 0)$, $\mathrm{K} = \frac{2\pi}{3a}(1, \frac{1}{\sqrt{3}})$, and $\mathrm{M} = \frac{2\pi}{3a}(1, 0)$. (b) Energy dispersion of graphene. The lower and the upper surfaces denote the valence π and the conduction π^* energy bands, respectively.

The calculation of the band structure of other semiconductors proceeds through a process almost similar to the band structure of the graphene described above. The Hamiltonian matrix for the bandstructure should be carefully constructed according to a primitive unit cell and the crystalline translation symmetry depending on crystal structure and broken periodicity with different crystal dimensionality. For physically confined structures such as thin films and nanowires, the surface states should be passivated with the sp^3 hybridization scheme, otherwise dangling bonds will create spurious energy states in the bandgap of the device as an artifact. The sp^3 hybridization proceeds in the following order [51]

1. Taking the s & p components of the onsite Hamiltonian for atoms on the edges
2. Rotating to the basis of the sp^3 hybridized molecular orbitals
3. Adding a large energy (~30eV) to the orbitals corresponding to dangling bonds
4. Rotating back to the atomic orbital basis

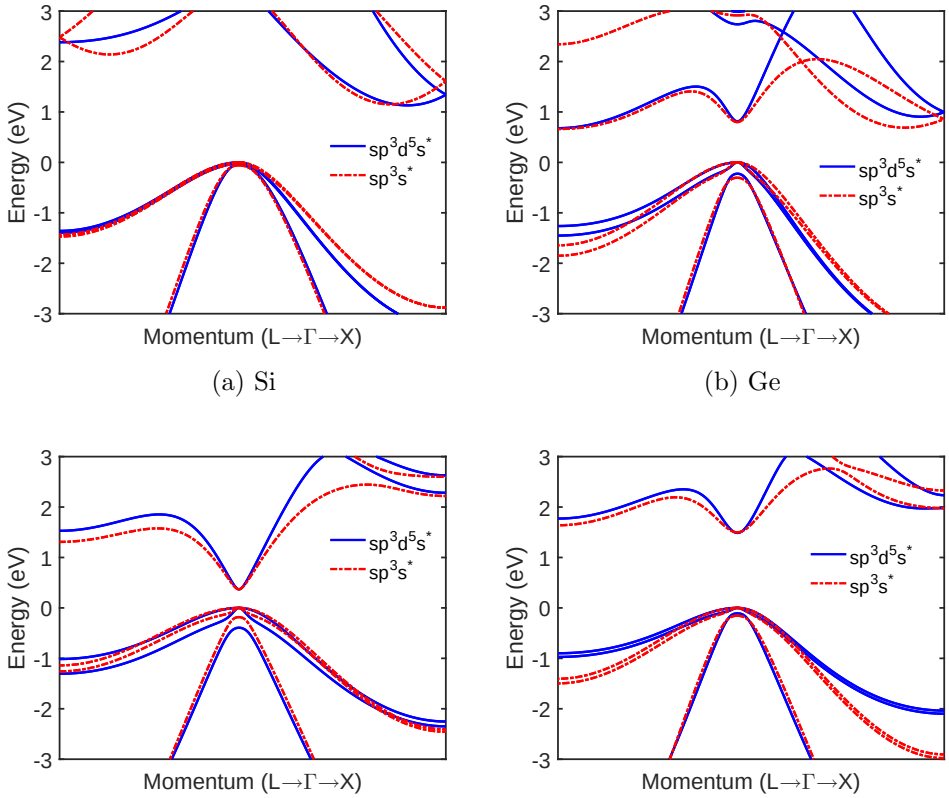
2.3 The choice of orbitals

$sp^3d^5s^*$

The choice of right orbital basis is important due to trade-off relationship between computation efficiency and accuracy. Over the last decades, there have been a series of improvements to the pioneering work of Slater and Koster to achieve a more accurate band structure. Slater and Koster [49] describe energy dispertion with a minimal sp^3 basis. Their model is acceptable for describing the valence band. However, sp^3 basis model is not enough to reproduce the indirect gap of semiconductors correctly, especially at the X point. Contrary to a TB description in the sp^3 basis, several theoretical studies showed that the lowest conduction state at X is not entirely anti-bonding and unoccupied atomic d orbitals contribute the charge densities of the first conduction band at the X, and L points [46] [52] [53]. To mimic the influence of the excited d states, Vogl et al. added the excited s-like orbital, s*, which yielded better reproduction of the lowest conduction band of diamond and zinc blende semiconductors at X and L [54]. However, transverse masses at these points and the second conduction band show poor agreement with experimental measurement [55]. Boguslawsky et al. showed the necessity of including the full d symmetry near the X point by comparing pseudopotential calculations with TB models [56]. In Ref. [46], Jancu et al. reproduced superior transverse masses of the indirect bands by adding the full excited d orbitals($sp^3d^5s^*$). Fig. 2.3 describes the bulk bandstructure of Si, Ge, InAs and GaAs. The $sp^3d^5s^*$ model with spin-orbit coupling is plotted in blue solid line and compared to sp^3s^* model with spin-orbit coupling which is plotted in red dashed lines. The material parameters for $sp^3d^5s^*$ were obtained from Ref. [54] and the parameters of $sp^3d^5s^*$ model for IV group (Si and Ge) and III-V group (InAs and GaAs) semiconductors were obtained from Ref. [57] and Ref. [58], respectively. While the sp^3s^* and $sp^3d^5s^*$ models show relatively good agreement between their energies and curvatures at the Γ point, the deviation between the two models become significant away from the Γ point. In addition, the sp^3s^* model fail to reproduce the conduction band energies for indirect valleys as well as the transverse effective mass at those point because of d symmetry. The inaccuracies of these sp^3 models becomes more pronounced in the nanoscale structure as shown in Fig. 2.4. Similar to the case of bulk, the sp^3 model thin film does not implement the band energy or mass in the remaining region well except the gamma point. Except for direct small bandgap semiconductor,such as InAs and InSb, whose conduction band minima at Γ point is significantly lower than other valleys, the use of sp^3s^* basis bandgapstructure leads to generate serious error.

Spin-orbit coupling

The spin-orbit coupling, which is an electromagnetic interaction between the electron's spin and the magnetic field generated by the electron's orbit around the nucleus, leads to spin splitting of energy bands. Since spin-orbit interaction has a small effect on electronic bands, the bandstructure calculation with only $sp^3d^5s^*$ orbitals is reasonably accurate for conduc-



(a) Si (b) Ge

(c) InAs (d) GaAs

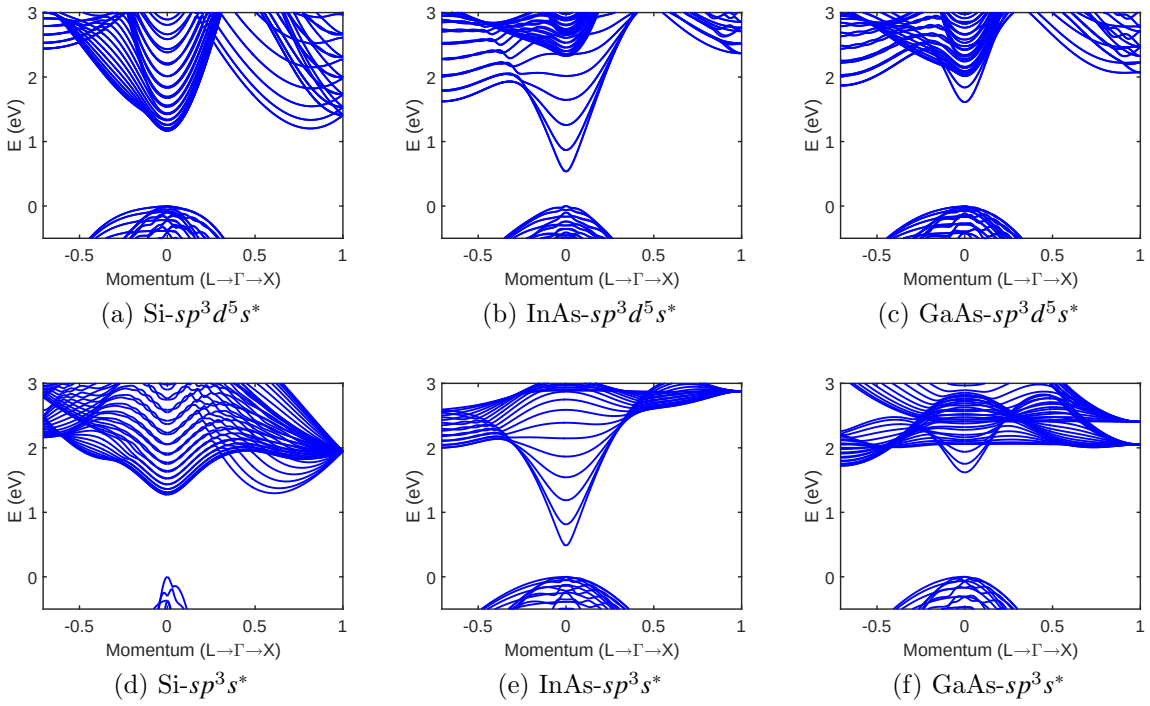
Figure 2.3: Bulk bandstructure of Si, Ge, InAs, and GaAs computed with the sp^3s^* (red dashed line) and $sp^3d^5s^*$(blue solid line).

tion band. However, in order to describe the top of the valence band precisely, it is necessary to include spin-orbit coupling. The matrix form of time-independent Schrödinger equation without spin-orbit coupling is expressed as

$$E(\psi) = [H_{op}](\psi) \tag{2.6}$$

where $H_{op} = -\frac{\hbar^2}{2m}\nabla^2 + U(\vec{r})$. Since Schrödinger equation is a non-relativisitc equation, eigenvalues of the Hamiltonian operator H_{op} are the energy levels having spin degeneracy, spin-up and spin-down. In order to split the spin degeneracy with spin-orbit coupling, the nonrelativistic schrödinger equation should be replace by

$$E(\psi)\begin{pmatrix}\psi\\ \psi\end{pmatrix} = \begin{bmatrix}H_{op} & 0\\ 0 & H_{op}\end{bmatrix}\begin{pmatrix}\psi\\ \psi\end{pmatrix} + [H_{so}]\begin{pmatrix}\psi\\ \psi\end{pmatrix} \tag{2.7}$$



(a) Si-$sp^3d^5s^*$ (b) InAs-$sp^3d^5s^*$ (c) GaAs-$sp^3d^5s^*$

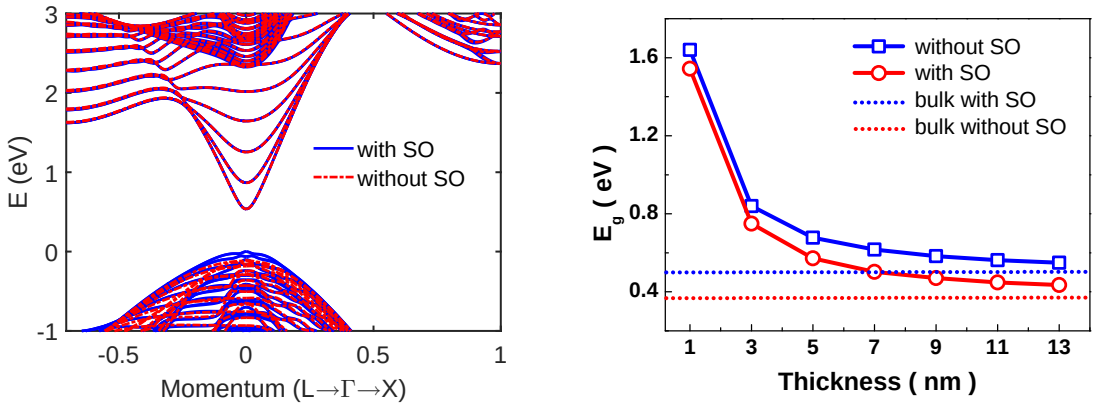
(d) Si-sp^3s^* (e) InAs-sp^3s^* (f) GaAs-sp^3s^*

Figure 2.4: Bandstructure of 6nm thin film (100) for Si, Ge, InAs, and GaAs using the $sp^3d^5s^*$ (first row) and sp^3s^* (second row).

where ψ and $\overline{\psi}$ are the up-spin component and the down-spin component of the electronic wavefunction, respectively. The spin-orbit Hamiltonian H_{so} is written as [59]

$$H_{so} = \frac{q\hbar}{4m^2c^2}\vec{\sigma}\cdot(\vec{E}\times\vec{p}) \tag{2.8}$$

where $\vec{E}$ is the nuclear electric field, $\vec{p}$ is momentum operator, and $\vec{\sigma}$ is the Pauli spin matrices. A detailed derivation for matrix elements of H_{so} can be found in Ref. [60]. By including the spin-orbit interaction, the size of Hamilonian matrix becomes twice the original size. For example, the size of Hamiltonian matrix for bulk with $sp^3d^5s^*$ with spin-orbit coupling increase from 20×20 to 40×40. Fig. 2.5 shows the bandstructure of InAs(100) thin films with and without spin-orbit coupling. In the case of conduction band calculations, the spin-orbit coupling can be ignored for computational efficiency. However, the $sp^3d^5s^*$ model including spin-orbit coupling parameters can correctly reproduce the band gap and effective masses of valence band only if those are included in the calculation. The neglection of this effect leads to an increase of the band gap so that the Off current of transistor would be underestimated.



(a) (b)

Figure 2.5: (a) Bandstructure of 6nm InAs (100) thin film using the $sp^3d^5s^*$ tight-binding model with spin-orbit coupling (blue) and without spin-orbit coupling (red). (b) Comparison of the bandgap with spin-orbit coupling and one without spin-orbit coupling as a function of the film thickness for InAs at Γ point. The dashed lines show the lower limit of the bandgap. The bulk bandgaps using $sp^3d^5s^*$-SO basis and $sp^3d^5s^*$ basis are 0.37 eV and 0.50 eV, respectively.

2.4 Bandstructure effects in nanoscale structures

In this section, we will discuss the bandstructure effects in nanoscale semiconductor. Fig. 2.6 illustrates the band structure and the density of states of InAs and GaAs thin films with thickness of 3nm. The transport and thin film growth directions are along the $\langle 100 \rangle$ and $\langle 001 \rangle$ axes respectively. For computational efficiency, the conduction bands and density of states are calculated using the 10 orbitals consisting of an $sp^3d^5s^*$ basis and shifted to the conduction band minimum energy calculated with $sp^3d^5s^*$ basis with spin-orbit coupling. The density of states is compared with the effective mass approximation using bulk effective masses. For the effective mass approach, the transverse and longitudinal effective masses of InAs and GaAs at Γ, X and L valleys are obtained from the bulk bandstructure using $sp^3d^5s^*$-SO TB model and are listed in Table 2.1. The effective masses of transport (m_x), transverse (m_y) and confinement (m_z) are calculated from the relation listed in Table 2.2 [61]. From the parabolic effective mass model, the energy dispersion of electron with confinement in z-direction is given by

$$E = E_{c,bulk} + \Delta E_n + \frac{\hbar^2 k_x^2}{2m_x^*} + \frac{\hbar^2 k_y^2}{2m_y^*} \tag{2.9}$$

where $\Delta E_n = \dfrac{\hbar^2 k_z^2}{2m_z^*} = \dfrac{\hbar^2 n^2 \pi^2}{2m_z^* L^2}$ $\quad n = 1, 2, \ldots$

The minimum energies of first subband for each valley are listed in Table 2.3

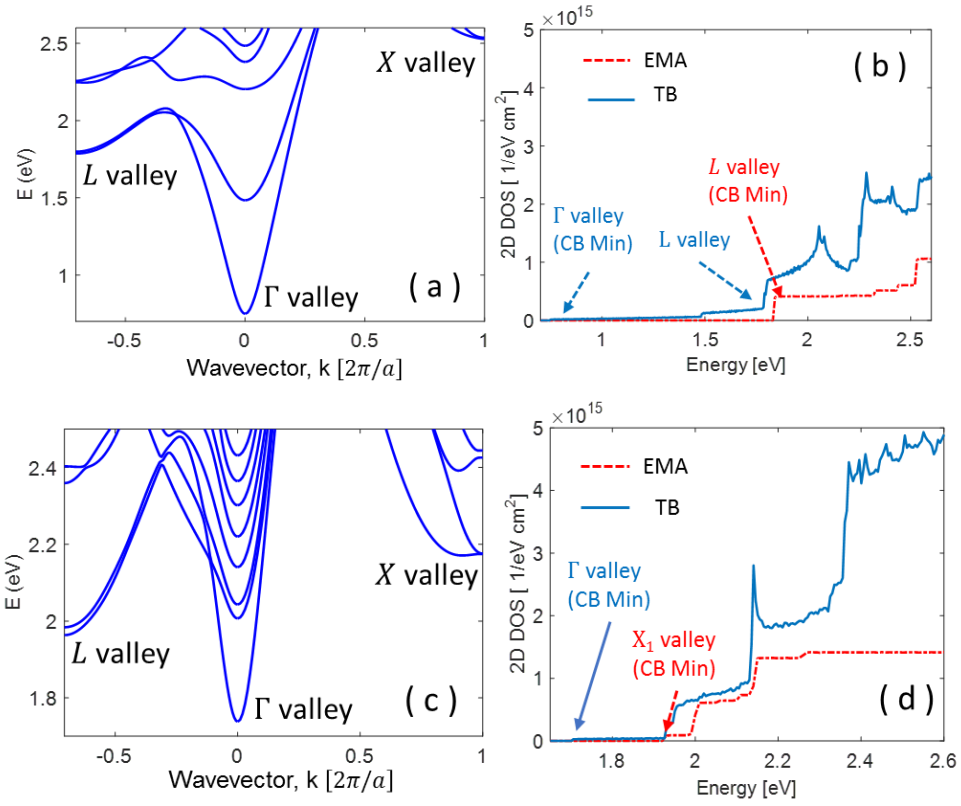


Figure 2.6: Bandstructure and DOS of InAs (a, b) and GaAs (c, d) thin films with thickness of 3nm. While the bulk EMA shows that L valley and X valley are the lowest in energy for InAs and GaAs, respectively. The TB model shows that Γ valley is still the lowest in energy.

	InAs	GaAs
m_Γ	0.0235	0.0657
$m_{X,l}$	1.126	1.881
$m_{X,t}$	0.175	0.175
$m_{L,l}$	1.540	1.728
$m_{L,t}$	0.0941	0.0967

Table 2.1: Effective masses (in units of the free-electron mass) for InAs and GaAs. All masses are computed at the respective extrema.

Valley	Degeneracy	m_x	m_y	m_z
Γ	1	m_Γ	m_Γ	m_Γ
X_1	1	$m_{X,t}$	$m_{X,t}$	$m_{X,l}$
X_2	1	$m_{X,t}$	$m_{X,l}$	$m_{X,t}$
	1	$m_{X,l}$	$m_{X,t}$	$m_{X,t}$
L	4	$m_{L,t}\frac{2m_{L,l}+m_{L,t}}{2m_{L,t}+m_{L,l}}$	$\frac{m_{L,l}+2m_{L,t}}{3}$	$\frac{3m_{L,l}m_{L,t}}{2m_{L,l}+m_{L,t}}$

Table 2.2: Transport, transverse and confinement effective masses and subband degeneracies of (100)/⟨100⟩ III-V thin film from the bulk principal effective masses for each valley. Ref. [61].

InAs				GaAs			
Valley	$E_{c,bulk}^{valley}$	ΔE_1	$E_{c1,EMA}^{valley}$	Valley	$E_{c,bulk}^{valley}$	ΔE_1	$E_{c1,EMA}^{valley}$
Γ	0.3699	1.7781	2.1480	Γ	1.4159	0.6356	2.0515
X_1	2.2827	0.0371	2.3198	X_1	1.9016	0.0222	1.9238
X_2	2.2827	0.2388	2.5215	X_2	1.9016	0.2384	2.1400
L	1.5287	0.3051	1.8338	L	1.7012	0.2961	1.9973

Table 2.3: Minimum energy of bulk conduction band, confinement-induced energy shift and minimum energy of thin film subband for each valley.

The bulk EMA predicts that the lowest energy valley is L for InAS and X for GaAs because the energy shift of Γ valley subband is lager than other valleys due to significantly small confinement effective mass. Contrary to the bulk EMA, the TB model shows that both semiconductors still have the lowest energy at the Γ valley. The discrepancy between the TB and bulk EMA results is due to the non-parabolicity of the Γ valley in GaAs and InAs. The nonparabolicity arises from the interaction between the wave functions of the two bands without band gap. This $k \cdot p$ interaction remains even when the crystal potentials open the gap between the two bands and decreases as the band-gap increases [62]. In the small bandgap semiconductor such as InSb and InAs, the strong $k \cdot p$ interaction across the narrow direct gap between the conduction and valence bands attribute the nonparabolicity in

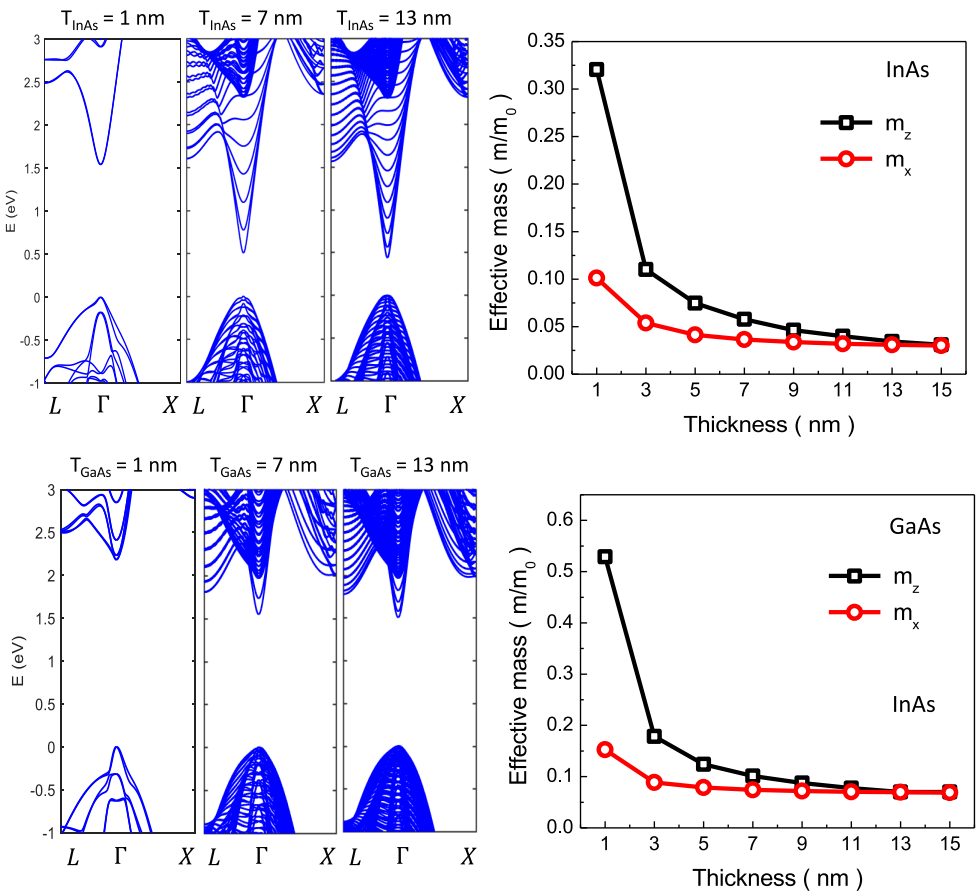


Figure 2.7: Band structure with different thickness for InAs and GaAs (left). Confinement (m_z) and transport effective mass (m_x) of the Γ valley as a function of channel thickness (right).

the Γ valley. On the other hand, the bulk effective mass approach that does not include the nonparabolicity underestimates the importance of the Γ valley, which result in different turn-on voltages and On-current densities for FET devices. Fig. 2.7 presents the bandstructure of InAs and GaAs thin film with different channel thickness from 1 nm to 13 nm. The Γ valley have the lowest energy for even 1nm (7 atomic layer) thickness. The Γ valley confinement effective masses (m_z) and transport effective mass (m_x) and bandgap are plotted as a function of film thickness in Fig. 2.7. The m_z is obtained from the energy difference between the bulk conduction band minimum and the lowest subband at the Γ point. m_x are calculated at Γ valley minima by fitting the E-k dispersion with a parabola. The effective masses increase as film thickness decrease significantly due to quantum confinement effect. It can be seen that the Γ valley bandgap increases while the band curvature decreases with decrease in film thickness. These quantum confinement results in a considerable distinction in transport properties of electrons in semiconductors such as band-to-band tunneling probability, the

carrier injection velocity and density of states.

2.5 Summary

In this chapter, we have discussed the necessity of atomistic modeling for bandstructure and certain aspects related to using tight-binding method. In confined semiconductor structures, the quantum confinement effect along the thickness direction increase bandgap and effective mass. Since a correct description of the bandstructure is crucial for accurately calculating the subband energies and their dispersions at the nanoscale, we used a semi-empirical atomistic tight-binding approach with the nearest-neighbor $sp^3d^5s^*$ orbital basis with spin-orbit coupling. This model is a powerful tool to predict full bandstructure of nanoscale semiconductors accurately.

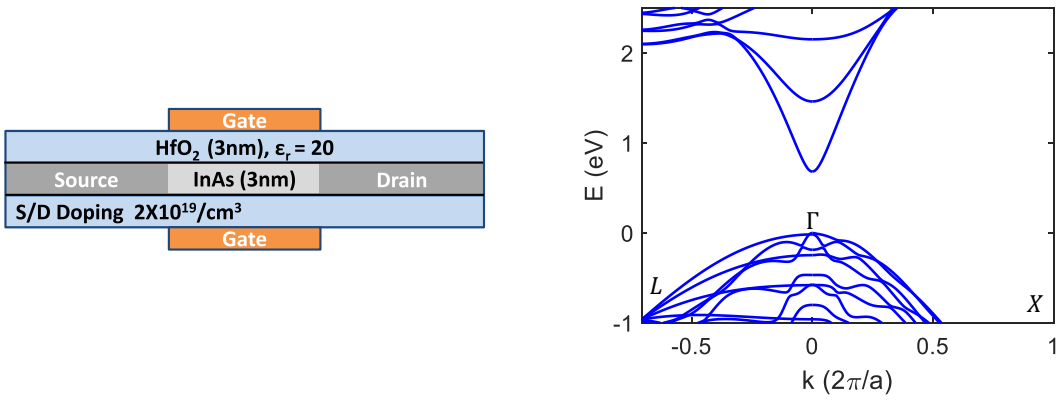
Chapter 3

Off state leakage current due to Band-to-band tunneling in III-V MOSFETs

3.1 Introduction

III-V metal oxide semiconductor field-effect transistor (MOSFET) devices have found significant interest for next generation transistors due to their high mobility [7, 63, 64], which could enable higher performance in the ultra-scaled nodes. But at the same time the higher mobility could also enable reaching a specific ON current at a smaller voltage. Due to this latter reason, the III-V devices could also become important for Internet of Things (IOT) applications. However, III-V materials usually have a direct energy bandgap. Due to this reason, the band-to-band tunneling (BTBT) current in these transistors are significantly larger than indirect bandgap materials such as Si. This could severely limit their application because of increased standby leakage. Therefore, to properly assess the potential of III-V transistors, an appropriate understanding of the band-to-band tunneling current, especially in the off-state, is necessary.

Modeling of the BTBT current in III-V transistors is not new. Approaches such as the Non-Equilibrium Green's Function (NEGF) method provides a rigorous framework for treating quantum transport in nanoscale devices and has been used successfully for many different device structures and modeling systems [65–67]. As such the NEGF method is the most accurate formalism, within a single particle picture, to model BTBT phenomenon. To this end, many previous work have modeled current-voltage characteristic of III-V transistors in the presence of band-to-band tunneling [68–71]. The objective of this dissertation is not to propose a new model for calculating such tunneling currents, but to point out a specific issue that has often been ignored in the quantum transport calculations and that can significantly contribute to BTBT induced leakage current. We shall call it 'Inadequate body filling by the source'. In what follows next, we shall discuss this problem within the context of an



(a) (b)

Figure 3.1: (a) Schematic structure of the simulated double-gate InAs FET. (b) Bandstructure of 3nm InAs(100) thin film using the sp^3s^* tight-binding model.

ultra-thin body (UTB) InAs transistor.

3.2 Simulation Results

As a model system a double gate (DG), UTB InAs MOSFET is simulated following the NEGF approach. The schematic cross-section of the double gate (DG) ultrathin body (UTB) InAs MOSFET is shown in Fig. 3.1a. The device is composed of an intrinsic channel (gate length = 10 nm) and n-doped source and drain with a doping density of $2\times10^{19}\text{cm}^{-3}$. The channel thickness is 3 nm. The gate dielectric is HfO_2 ($\varepsilon_r = 20$) with a thickness of 3 nm. The transport direction is aligned with the [100] crystal axis. First, a real space, ballistic transport within NEGF is solved self-consistently with the Poisson's equation. In this case, we use a sp^3s^* tight-binding (TB) Hamiltonian. The sp^3s^* model for InAs is accurate enough to capture the tunneling currents appropriately [68]. The TB parameters of Ref. [72] were used and passivation of the surface dangling bonds was assumed and implemented in the Hamiltonian with appropriate potentials. The bandstructure of 3nm thick InAs thin film (20 atomic layers) is show in Fig. 3.1b. A minimum band gap of E_g is 0.70 eV at the Γ point and the extracted transport effective mass (m_x) is $0.041m_0$. For simplicity, hence forth we shall call the real space, self-consistent, NEGF solution as the 'Full Band NEGF'.

Fig. 3.2 shows electron correlation function and current spectrum within Full Band NEGF at a typical off condition ($V_{GS} = -0.1$ V and $V_{DS} = 0.9$ V). There are two energy regions of interest: (i)E_1; the two ends of which are cut off by the band edges and (ii)E_2; which is the region underneath the valence band in the body region but above E_1. Note from Fig. 3.2a that the region E_2 has no electrons. This is somewhat surprising but can be understood by looking at how the electrons are being injected from the source. The valence

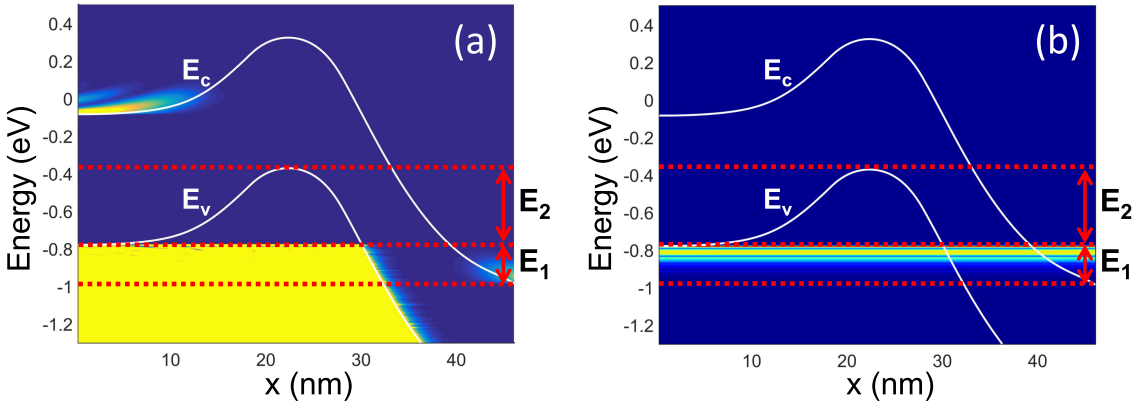


Figure 3.2: (a) Electron correlation function, G^n, which shows electron distribution in energy and position and (b) Energy-resolved current spectrum at $V_{DS} = 0.9$ V and $V_{GS} = -0.1$ V.

band on the source side cuts off the Fermi distribution. No electrons above the edge of the valence band at the source side can be injected, leading to the empty region in the body. *This is the main issue that we want to bring attention to.* In practice, such empty region underneath the body does not exist. The abundant valence electrons continuously go through election-electron interaction at a rate ($\tau_{ee} \approx 10^{-13}$ s) that is much faster than the tunneling rate at the drain side. Therefore, for all practical purposes, the distribution of the electrons inside the body will build up according to the Fermi function and electrons will fill up the empty region E_2. This is what we have named 'Inadequate body filling by the source'. Due to the fact that it is numerically intractable to solve electron-electron interactions in an already demanding approach such as the NEGF, this inadequate filling of the body has commonly been ignored in device simulations. The central result of this dissertation is to point out that ignoring the filling leads to significant errors in the calculation of off-state leakage current. The only way to accurately model this leakage current is to use a boundary condition that appropriately captures the filling of the body. This poses a challenge as the only two well-defined Fermi levels in a quantum transport problem are the source and the drain Fermi functions. Hence, an approximation has to be made. The approach we take is to calculate the ballistic current first and then do a second calculation which is essentially a two terminal current calculation between body and drain by defining a Fermi distribution at the body. In this two-terminal problem, we use the self-consistent potential previously calculated from the three-terminal problem and calculate the BTBT. Note that some of these issues have been discussed within the context of resonant tunneling diodes [73]. One could do another NEGF calculation for this two terminal current. However, we have previously shown [74] that when the same potential is used, BTBT current calculated from the Wentzel-Kramers-Brillouin (WKB) approximation closely matches that of the NEGF approach. Given that WKB is numerically a much lighter approach than NEGF, we have used the WKB.

We begin by solving the NEGF within the mode-space approach [75]. Note that, because

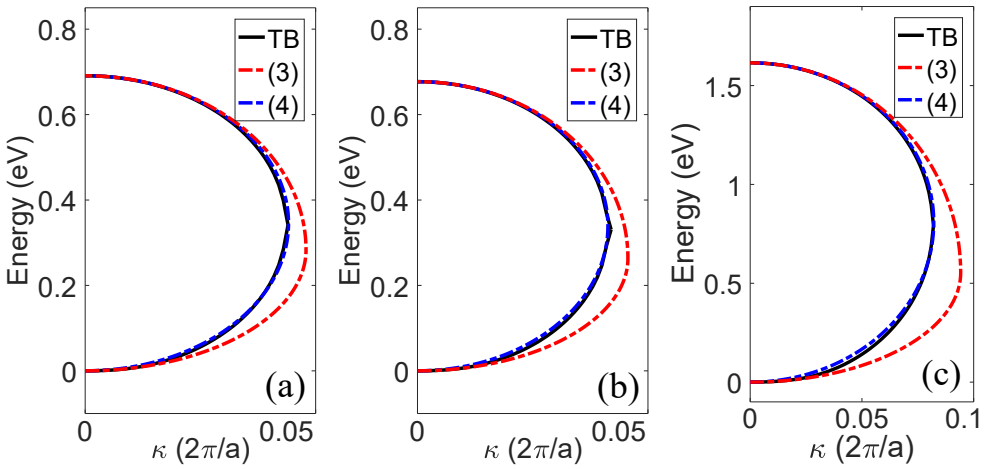


Figure 3.3: Complex band structure along the [100] direction of InAs and GaAs UTBs calculated using the TB approach (black) and the analytical equations in Eq. (3.3) (red) and Eq. (3.4) (blue). (a) 3nm InAs - sp^3s^*, (b) 5nm InAs - $sp^3d^5s^*$, (c) 5nm GaAs - $sp^3d^5s^*$.

here the individual modes are decoupled, this approach cannot directly account for BTBT. Next, we perform a two terminal current calculation between the body and the drain following the WKB method, using the potential self-consistently converged from the mode-space calculation. Since we are interested in the off-state current, there is not too much charge in the channel and therefore, the self-consistent potential is not expected to change significantly even in the presence of the BTBT current. The tunneling probability within the WKB approach can be written as:

$$T_{WKB}(E, k_\perp) = \exp\left(-2\int_{x_1(E)}^{x_2(E)} \kappa(x, k_\perp)\, dx\right) \tag{3.1}$$

where $\kappa(x, k_\perp)$ is the imaginary wavevector as a function of position (x) at a given transverse momentum $k_\perp$ and x_1 and x_2 are positions of valence band and conduction band at E in the band-diagram. From T_{WKB}, the BTBT current (I_{BTBT}) can be obtained from

$$I_{BTBT} = \frac{q}{\hbar}\int \frac{dE}{2\pi} \sum_{k_\perp \in 1^{st}BZ} T_{WKB}(E, k_\perp)\,[f_S(E) - f_D(E)] \tag{3.2}$$

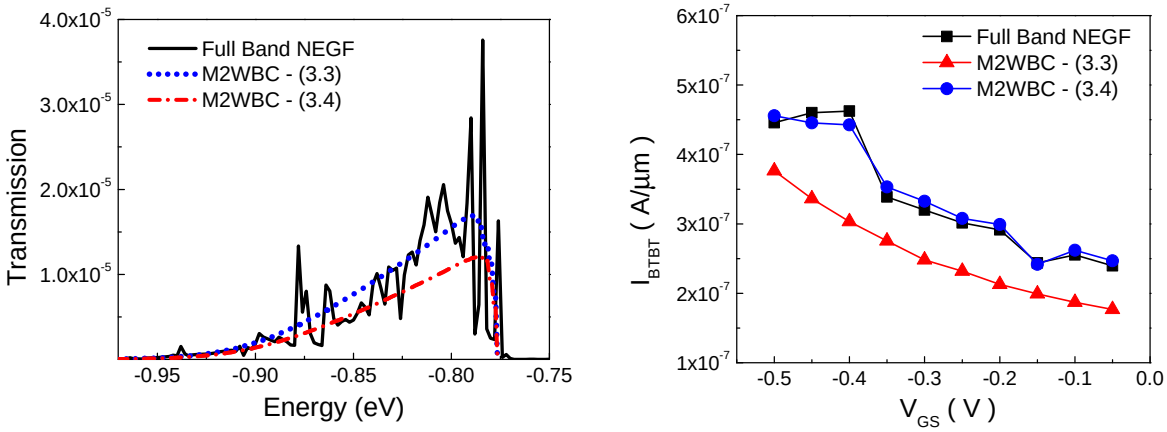
where f_S and f_D are the equilibrium Fermi functions of the source and drain, respectively. For the imaginary wavevector (κ), instead of calculating κ within the TB method [76], an analytical equation from [77] was used:

$$\begin{aligned} \kappa(E) &= \sqrt{\tfrac{2m_h}{\hbar^2}E\left(1 - \tfrac{E}{2E_q}\right)}, && 0 < E < E_q \\ \kappa(E) &= \sqrt{\tfrac{2m_e}{\hbar^2}(E_g - E)\left(1 - \tfrac{E_g - E}{2(E_g - E_q)}\right)}, && E_q \le E < E_g \end{aligned} \tag{3.3}$$

where E_g is the bandgap of the semiconductor, m_e and m_h are the effective masses of electrons and holes respectively and $E_q = E_g \frac{m_e}{m_e+m_h}$ is the branch point at which the conduction and valence branches join each other. When electrons and holes have different effective masses, decaying behavior are attributed to hole effective mass when E < Eq and to electron effective mass when Eq ≤ E. In Fig. 3.3, we compare the complex dispersion of InAs and GaAs along the transport direction calculated using the analytical equation (dashed lines) and the tight-binding approach (black solid line) at the Γ point. We find that using m_e instead of m_h for hole branch predicts a more accurate complex band structure. This result implies that transition between the hole branch and the electron branch occurs in the vicinity of the valence band maximum and charges follow characteristics of electrons during tunneling. Substituting m_e to m_h in (3.3), κ can be obtained with a simplified equation as,

$$\kappa(E) = 1/\hbar\sqrt{2m_eE\left(1-\frac{E}{E_g}\right)}, \quad 0 < E < E_g \tag{3.4}$$

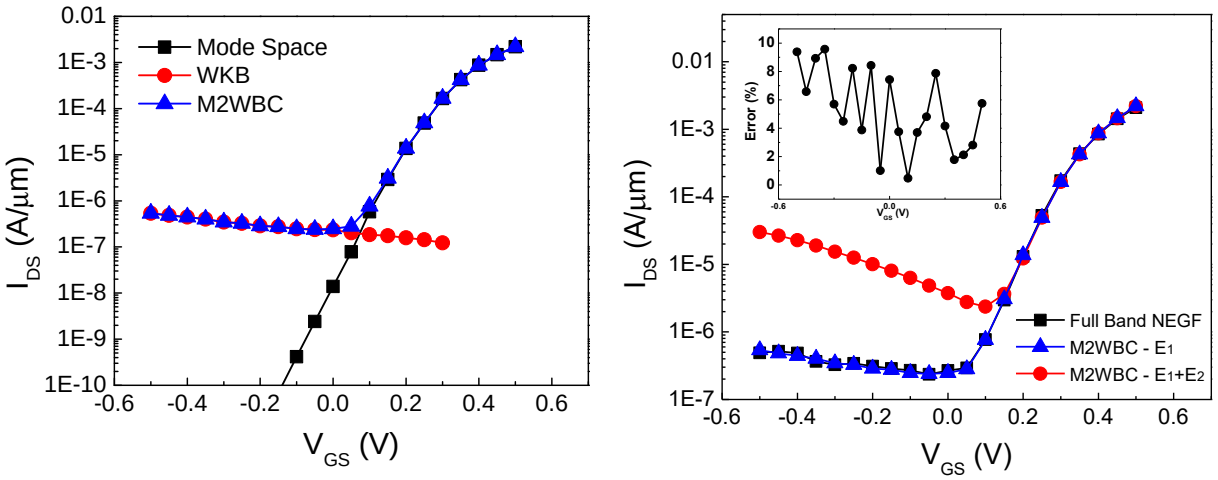
Note that, within the region E_1, a conventional real-space approach will accurately calculate the BTBT current. It is only in the region E_2 where the conventional approach is inaccurate. Therefore, to estimate how our 'mode-space + 2 terminal WKB with appropriate boundary condition' approach (named M2WBC for convenience) compares with the Full Band NEGF approach. The transmission probabilities calculated using the method presented in the dissertation (M2WBC) with Eq. 3.3 and Eq. 3.4 and Full Band NEGF for the region E_1 are compared in Fig. 3.4a. For the WKB approximation, we included only one dominant imaginary band connecting the VB maximum and CB minimum since the transmission coefficient decreases exponentially with the integrated value of imaginary part of the wave



(a) (b)

Figure 3.4: (a) Transmission probability and (b) BTBT current calculated from Full Band NEGF (black) and the M2WBC approach using Eq. (3.3) (red) and Eq. (3.4) (blue) within the tunneling window E1. (V_{GS} = −0.1 V and V_{DS} = 0.9 V).

vector. Even though tunneling probabilities calculated from the WKB approximation do not exhibit the quantum resonances as seen in the quantum transport approach, both approaches show the same mean contour. We compare the BTBT current of the 3nm InAs UTB FET obtained from the Full Band NEGF to the ones calculated with M2WBC approach in Fig. 3.4b. The average value of tunneling probability increases as the gate voltage is decreased, because the width of tunneling barrier decreases as a higher electric field is applied between the gate and source. According to the results shown in Fig. 3.4b, the Eq. (3.4) seems to describe the change in κ with respect to E, accurately. On the other hand, Eq. (3.3) underestimates the tunneling current because the attribution of higher hall effective mass predicts a higher decay rate κ. The transfer characteristics at $V_{DS} = 0.9$ V are shown in Fig. 3.5. For the M2WBC model, first, the currentvoltage characteristics are calculated using the mode-space NEGF formalism. Then, using the electrostatic potential calculated by the mode-space approach, the tunneling current is calculated using the WKB approximation within the tunneling window E_1. The tunneling current calculated using the tunneling probability from the WKB approximation is added to the current from the mode-space approach. Fig. 3.5b shows the comparisons between the Full Band NEGF simulations and the M2WBC approach described above. For E_1, the difference in current calculated by the Full Band NEGF and M2WBC approach is less than 10 % as shown in the inset of Fig. 3.5b, establishing the feasibility of M2WBC for calculating BTBT current. The main distinction is in the region E_2, where naive application of NEGF ignores the appropriate filling, but



(a) (b)

Figure 3.5: (a) Simulated $I_{DS} - V_{GS}$ characteristics for the mode space approach (black), WKB approximation (red) and the M2WBC approach (blue) at $V_{DS} = 0.9$ V. (b) Simulated $I_{DS} - V_{GS}$ characteristics for the Full Band NEGF (black), M2WBC within E_1 only (blue) and M2WBC within E_1 and E_2 (red). Percentage difference in the current between Full Band NEGF and M2WBC within E_1 (inset).

the M2WBC will account for it. As shown in Fig. 3.5b, when the current flowing in E_2 is included, the actual drain current goes up almost two orders of magnitude compared to the Full Band NEGF.

Furthermore, the computation times for the the Full Band NEGF and the M2WBC model are compared in Fig. 3.6. The M2WBC approach is about 20 times faster than the Full Band NEGF method for one iteration of self-consistency in case of 3 nm thick InAs UTB DG MOSFETs with an sp^3s^* TB Hamiltonian description. This difference would become larger as the size of the Hamiltonian increases.

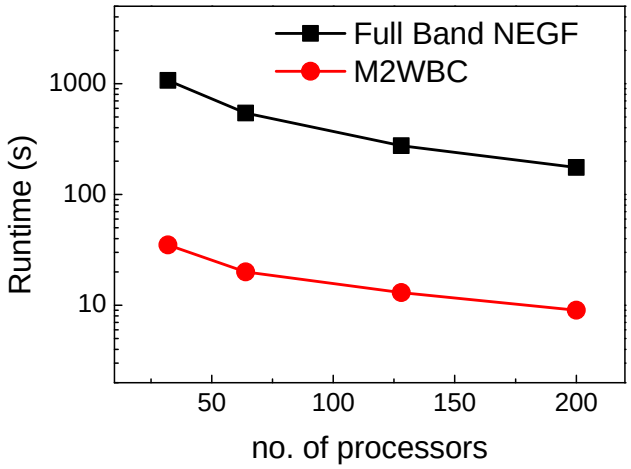


Figure 3.6: Comparison of running times between the Full Band NEGF and the M2WBC.

3.3 Conclusions

To summarize, we have discussed how failing to account for proper boundary conditions could lead to insufficient filling of the body in a MOSFET under a typical OFF condition. This inadequate filling of the body leads to significant error in the calculation of off-state leakage current that originates from band-to-band tunneling. Using a UTB InAs transistor as a model system, we show that the leakage current can be underestimated by almost two orders of magnitude. We have also shown a numerically efficient method to calculate this current. Especially, for direct bandgap materials, properly calculating the off-state leakage is critical to assess their potential for application.

Chapter 4

Effect of phonon scattering on quantum transport in III-V MOSFETs

4.1 Introduction

As the channel length of modern field-effect transistors enters the nanoscale regime, it is important to understand the effects of quantum mechanics on electron transport because quantum mechanical effects such as quantum mechanical tunneling, carrier confinement, and interference effects play a crucial role. Therefore, there have been many researches on nanoscale devices using the Non-Equilibrium Green's Function (NEGF), which provides the most rigorous framework for treating quantum transport. However, the most common approximation in NEGF calculations is to ignore the effects of all scattering processes in order to reduce simulation time [65–69]. Even if the device length, which is the distance traveled by a moving carrier between two thermally equivalent reservoirs (source and drain), is the same or smaller than the mean free path, certain amount of scattering exists and influences the transport of the free carrier. Therefore, an accurate and efficient device simulation method based on quantum mechanical formalism that treats the dissipative process rigorously is necessary to interpret experimental results.

In the NEGF approach, scattering can be incorporated through a self-energy term, i.e., the perturbative Hamiltonian, which describes the interaction between the electron and scattering source. As a phenomenological approach, Büttikerr probe model can implement scattering by replacing individual physical scattering mechanisms with a global energy and momentum absorbing self-energy. Though this model is numerically efficient, it yields only a macroscopic and semi-classical description of the scattering [78]. The microscopic electron scattering mechanisms can be taken into account within the self-consistent Born approximation. In this chapter, we present a two-dimensional (2D) simulation framework capable of handling dissipative transport in the nanoscale MOSFETs based on the NEGF formal-

ism. Instead of treating transverse modes independently and integrating them analytically, we coupled them. Furthermore, we also couple each longitudinal energy to include energy relaxation by inelastic scattering. The mode space approach with an effective mass approximation is used because of a full NEGF treatment, which uses the atomistic electron and phonon energy spectra, is almost impossible to implement in 2D. Note that we use effective masses, which depend on each momentum (k_x and k_y), in order to include the nonparabolicity of the electron band instead of approaching a general mode space that uses only one effective mass for all momentums.

In Section II, we provide a brief review of the theoretical framework for including electronphonon scattering with the deformation potential theory and the self-consistent Born approximation. To verify the obtained self-energy functions, we compare the scattering rates obtained by NEGF formalism to the ones calculated by the Fermi's golden rule (FGR), which is used in classical Boltzmann transport theory. Simulation of the double gate ultrathin body (UTB) MOSFET device with phonon scattering is presented in Section III. Device characteristics are calculated from the self-consistent solution of the 2D Poisson's equation and the NEGF equations including scatterings. InAs was used as channel material. We also include surface roughness (SR) scattering, since SR is the dominant scattering mechanism in UTB films. SR produces fluctuations in the confining potential along the channel, which leads to fluctuation in the conduction band minimum levels that cause electron scattering. Since the channel region of modern day devices is commonly undoped to increase mobility and reduce random dopant fluctuation, ionized impurity scattering is ignored in this study.

4.2 Simulation methodology

A comprehensive description of the use of the mosde space NEGF formalism and the self-consistent Born approximation to consider electron-phonon scattering can be found in the literature [75, 79, 80]. For clarity, however, we will give a brief overview of the technique used to build a 2D mode-space Hamiltonian and how to obtain the self-energies for acoustic phonon (AP), optical phonon (OP) and polar optical phonon (POP) scattering in order to include phonon scattering in the NEGF formalism.

Mode space approach

First, we begin by solving mode-space approach. Here we defined that the transport direction is x-axis, the confined direction perpendicular to the gate is z-axis and the unconfined direction along the width of the device is y-axis. When we assume that there is no variation of potential in y direction and the device is wide enough, the three-dimensional (3D) wave function $\Psi(x, y, z)$ can be expanded in terms of 2D orthonormal basis and the plane wave as $\frac{1}{\sqrt{W}}\Phi(x, y)e^{ik_y y}$, where k_y is the quantum number and W is the device width. The wave

function $\Phi(x,z)$ is obtained from 2D Schrödinger equation as

$$-\frac{\hbar^2}{2m_x^*}\frac{\partial^2}{\partial x^2}\Phi(x,z) - \frac{\hbar^2}{2m_z^*}\frac{\partial^2}{\partial z^2}\Phi(x,z) - qV(x,z)\Phi(x,z) = \left[E - E_{k_y}\right]\Phi(x,z). \tag{4.1}$$

where E_{k_y} corresponds to the eigenenergy of plane wave and $m^*_{x,y,z}$ are the effective mass of each direction. Furthermore, in case of UTB structure, quantum confinement in the thickness direction introduces modes (subbands) and the wave function can be expanded in an one-dimensional (1D) orthonormal basis as

$$\Phi(x,z) = \sum_m \delta(x-x')\xi_m(x',z)\phi_m(x') \tag{4.2}$$

where $\phi_m(x')$ are the expansion coefficients and $\xi_m(x,z)$ are the eigenfunctions of subbands associated confinement in the z-direction.

By using the orthonormal basis and assuming that the shape of the confined mode changes smoothly along the transport direction(x-axis), the 2D Schrödinger equation can be transformed into 1D as

$$-\frac{\hbar^2}{2m_x^*}\frac{\partial^2\phi_m(x)}{\partial x^2} + E_m(x)\phi_m(x) = (E - E_{k_y})\phi_m(x) \tag{4.3}$$

Using the mode space approach, the 3D simulation domain has been split into the several sets of 1D transport problem. Note that there are infinite quantum number in the unconfined direction (y) while a few one in confined direction (z). Since only a few subbands are occupied in UTB devices, the computational burden can be significantly reduced by using a uncoupled mode space method [75, 81]. From NEGF formalism, the retarded Green's function for each mode (α) can be written as

$$G(E,k_y,\alpha) = \left[EI - H(E,k_y,\alpha) - \Sigma_1(E,k_y,\alpha) - \Sigma_2(E,k_y,\alpha) - \Sigma_S(E,k_y,\alpha)\right]^{-1} \tag{4.4}$$

where H is the effective mass Hamiltonian of device, Σ_1 and Σ_2 are the self-energy matrices representing the open boundary conditions of source and drain, respectively and Σ_S is the retarded self-energy matrix representing the interaction between the electron and scattering source. The matrix elements of a single band Hamiltonian (H) and the self-energy of contacts can be written as [39, 75]

$$\begin{aligned} H_{i,j} &= -t_x\delta_{i-1,j} + [2t_x + E_c(k_y) + U(i)]\delta_{i,j} - t_x\delta_{i+1,j} \\ \Sigma_1(i,j) &= -t_x e^{ik_x x}\delta_{1,j}\delta_{1,i} \\ \Sigma_2(i,j) &= -t_x e^{ik_x x}\delta_{N,j}\delta_{N,i} \end{aligned} \tag{4.5}$$

where $t_x = \dfrac{\hbar^2}{2m_x^* a^2}$ with node spacing (a), E_c is minimum energy of conduction-band and $U(i)$ is electrostatic potential at node i.

The electron density for each mode (α) is obtained from NEGF formalism as [79],

$$n^{\alpha}(x) = \frac{1}{W}\sum_{k_y} 2\int \frac{dE}{2\pi} diag\left[G_{\alpha}\Sigma_{\alpha}^{<}G_{\alpha}^{\dagger})\right] \tag{4.6}$$

where $\Sigma_{\alpha}^{<}$ is in-scattering self-energy which is sum of in-scattering self-energy of scattering and contact L and R ($\Sigma_{\alpha}^{<} = \Sigma_{\alpha,L}^{<} + \Sigma_{\alpha,R}^{<} + \Sigma_{\alpha,S}^{<}$).
The 1D electron density from Eq. 4.6 can be distributed in 2D with wave functions in confined direction, $\chi_{\alpha}(x,z)$, as,

$$n(x,z) = \sum_{\alpha} |\chi_{\alpha}(x,z)|^2 \, n^{\alpha}(x) \tag{4.7}$$

$|\chi_{\alpha}(x,z)|^2$ are obtained by using the eigenfunction of TB Hamiltonian which is essentially wave functions in confined direction. The electron density from NEGF is fed back to the Poisson solver until self-consistency is achieved. After convergence is achieved, current are calculated from the converged potential profile as,

$$I = \frac{q}{\hbar}\sum_{\alpha}\frac{1}{W}\sum_{k_y} 2\int \frac{dE}{2\pi} Trace\left[A_{\alpha}\Sigma_{L,\alpha}^{<} - G_{\alpha}^{<}\Gamma_{L,\alpha}\right] \tag{4.8}$$

where A is spectral function, $G^{<}$ is electron correlation function, $\Sigma_{L}^{<}$ is in-scattering function of contact L and Γ_L is level broadening of contact L. Each matrix functions can be obtained from NEGF formalism as

$$\begin{aligned} A &= i\left[G - G^{\dagger}\right] \\ G^{<} &= G\Sigma^{<}G^{\dagger} \\ \Gamma_L &= i\left[\Sigma_L - \Sigma_L^{\dagger}\right] \\ \Sigma_L^{<} &= \Gamma_L f_L \end{aligned} \tag{4.9}$$

In order to approach with the mode-space method, the self-energy functions of scattering must also be transformed into the mode-space from real space. We will derive them rigorously in next subsection.

self-energy for scattering in mode-space

In NEGF formalism, the microscopic scattering processes can be taken into account by self-energies. The lesser and greater self-energy functions for the electron-phonon interactions can be obtained from the self-consistent Born approximation

$$\Sigma_S^{\lessgtr} = D^{\lessgtr}(X_1, X_2)G^{\lessgtr}(X_1, X_2) \tag{4.10}$$

where the argument X = $\{R, t\}$ represent the spatial coordinates and time, respectively. [82–84]. $G^{\lessgtr}$ is the greater and lesser electron Green functions is defined by

$$\begin{aligned} G^{>}(X_1, X_2) &= (-i/\hbar)\left\langle \psi(X_1)\psi^{\dagger}(X_2)\right\rangle, \\ G^{<}(X_1, X_2) &= (i/\hbar)\left\langle \psi^{\dagger}(X_2)\psi(X_1)\right\rangle \end{aligned} \tag{4.11}$$

where ψ is the electron field operator. The $D^{\lessgtr}$ is the greater and lesser phonon Green's functions which is defined in thermal equilibrium as

$$\begin{aligned} D^{<}(X_1, X_2) &= \langle H_{ep}(X_1)H_{ep}(X_2)\rangle, \\ D^{<}(X_1, X_2) &= \langle H_{e-p}(X_2)H_{e-p}(X_1)\rangle \end{aligned} \tag{4.12}$$

The electron-phonon interaction Hamiltonian in a general form can be written as [82, 83, 85]

$$H_{ep} = \frac{1}{\sqrt{V}}\sum_{q} M_q(a_q e^{-i\omega_q t + iq\cdot R} + a_q^{\dagger}e^{i\omega_q t - iq\cdot R}) \tag{4.13}$$

Where V is the volume of the sample, M_q is the electron-phonon matrix element, a_q and $a_q^{\dagger}$ are annihilation and creation operators for a phonon in mode q. In this equations, the spatial coordinates R is 3D real space coordinate (x, y, z). Since the averages of phonon operator products at thermal equilibrium can be written as below relation,

$$\langle a_{q'}^{\dagger}a_q\rangle = \delta_{q'q}n_q,\ \langle a_{q'}a_q^{\dagger}\rangle = \delta_{q'q}(n_q+1),\ \langle a_{q'}a_q\rangle = 0,\ \langle a_{q'}^{\dagger}a_q^{\dagger}\rangle = 0 \tag{4.14}$$

by inserting (4.13) into (4.12) we obtain

$$\begin{aligned} D^{<}(X_1, X_2) &= \frac{1}{V}\sum_{q}|M_q|^2[\ (n_q+1)e^{(i\omega_q(t_1-t_2)+iq\cdot(R_2-R_1))} + n_q e^{(i\omega_q(t_2-t_1)-iq\cdot(R_1-R_2))}]\ , \\ D^{>}(X_1, X_2) &= \frac{1}{V}\sum_{q}|M_q|^2[\ (n_q+1)e^{(i\omega_q(t_2-t_1)+iq\cdot(R_1-R_2))} + n_q e^{(i\omega_q(t_1-t_2)-iq\cdot(R_2-R_1))}] \end{aligned} \tag{4.15}$$

where n_q is the average phonon number in mode q which, in thermal equilibrium, is Bose-Einstein factor, $n_q = 1/(exp(\hbar\omega_q/k_BT) - 1)$. Inserting Eq. (4.15) into Eq. (4.10) and Fourier transforming relative time interval t_2-t_1, we can obtain the energy-dependent self-energy function as

$$\begin{aligned} \Sigma^{<}(R_1, R_2, E) = \frac{1}{V}\sum_{q}|M_q|^2[\ & e^{iq\cdot(R_1-R_2)}(n_q+1)G^{<}(R_1, R_2, E+\hbar\omega_q) \\ & + e^{iq\cdot(R_2-R_1)}n_q G^{<}(R_1, R_2, E-\hbar\omega_q)]\ , \\ \Sigma^{>}(R_1, R_2, E) = \frac{1}{V}\sum_{q}|M_q|^2[\ & e^{iq\cdot(R_2-R_1)}(n_q+1)G^{>}(R_1, R_2, E-\hbar\omega_q) \\ & + e^{iq\cdot(R_1-R_2)}n_q G^{<}(R_1, R_2, E+\hbar\omega_q)] \end{aligned} \tag{4.16}$$

The first term in each equations corresponds to the emission of a phonon, and the second term corresponds to the absorption of a phonon. Since we assume that the system is invariant along the y-axis and all quantities depend only on the difference coordinate $y_1 - y_2$, the unconfined transverse dimension (here, y-axis) can be transformed to the momentum representation by expanding with the momentum states or vice versa [86] [80]. Then $G^<$ and $\Sigma^<$ can be expressed as

$$G^<(R_1, R_2, E) = N^{-1} \sum_{k'_{y1}, k'_{y1}} e^{i(k'_{y_1} y_1 - k'_{y_2} y_2)} G^<(r_1, r_2, k'_{y1}, k'_{y2}, E), \tag{4.17}$$

$$\Sigma_S^<(r_1, r_2, k_{y_1}, k_{y_2} E) = N^{-1} \sum_{r_1, r_2} e^{i(k_{y_2} y_2 - k_{y_1} y_1)} \Sigma^<(R_1, R_2, E) \tag{4.18}$$

where N is the number of grid cells, i.e., $N = \dfrac{L_y}{a_y}$. Substituting Eq.(4.17) into Eq.(4.16) and then inserting into Eq. (4.18), the lesser and greater self-energy of scattering in unconfined transverse momentum space can be obtained with the lesser and greater green function in the same space as

$$\begin{aligned}
\Sigma_S^<(r_1, r_2, k_y, E) = \frac{1}{V} \sum_{q_x, q_y, q_z} |M_q|^2 [\ & e^{iq_x(x_1 - x_2) + iq_z(z_1 - z_2)} (n_q + 1) G^<(r_1, r_2, k_y + q_y, E + \hbar\omega_q) \\
& + e^{iq_x(x_2 - x_1) + iq_z(z_2 - z_1)} n_q G^<(r_1, r_2, k_y - q_y, E - \hbar\omega_q)] \ , \\
\Sigma_S^>(r_1, r_2, k_y, E) = \frac{1}{V} \sum_{q_x, q_y, q_z} |M_q|^2 [\ & e^{iq_x(x_1 - x_2) + iq_z(z_1 - z_2)} (n_q + 1) G^>(r_1, r_2, k_y - q_y, E - \hbar\omega_q) \\
& + e^{iq_x(x_2 - x_1) + iq_z(z_2 - z_1)} n_q G^>(r_1, r_2, k_y + q_y, E + \hbar\omega_q)] \ ,
\end{aligned} \tag{4.19}$$

where the spatial r is 2D real space coordinate (x,z). This self-energy for electron-phonon scattering is for interaction between unconfined bulk phonon and unconfined bulk electrons. However, for the confined device structure, the self-energy for scattering should be transformed to include interaction between 3D phonon and 2D electron. Assumption of 3D bulk phonon is still valid because confining phonon in 2D is extremely difficult unlike electron. Since the real space Green's function can be expanded in a eigenfunction space with eigenvectors which form a orthogonal basis with mode space approch [80, 87, 88], the lesser and

greater self energy of scattering in real space can be transformed as

$$\begin{aligned}\Sigma_S^<(x_1,x_2,\alpha,k_y,E) = \frac{1}{A}\sum_{q_x,q_y,\beta}|M_q|^2[\ & e^{iq_x(x_1-x_2)}(n_q+1)G^<(x_1,x_2,\beta,k_y+q_y,E+\hbar\omega_q)\\ & +e^{iq_x(x_2-x_1)}n_qG^<(x_1,x_2,\beta,k_y-q_y,E-\hbar\omega_q)]\ F(\alpha,\beta)\\ \Sigma_S^>(x_1,x_2,\alpha,k_y,E) = \frac{1}{A}\sum_{q_x,q_y,\beta}|M_q|^2[\ & e^{iq_x(x_1-x_2)}(n_q+1)G^>(x_1,x_2,\beta,k_y-q_y,E-\hbar\omega_q)\\ & +e^{iq_x(x_2-x_1)}n_qG^>(x_1,x_2,\beta,k_y+q_y,E+\hbar\omega_q)]\ F(\alpha,\beta)\end{aligned} \tag{4.20}$$

where F is The form-factor account for transition between subbands α and β.

$$F(\alpha,\beta) = \int_{-\infty}^{\infty}\psi_\alpha^*(z)\psi_\beta^*(z)\psi_\alpha(z)\psi_\beta(z)dz \tag{4.21}$$

For the infinite well approximation, since the envelope functions are given by

$$\psi_\alpha(z) = \sqrt{\frac{2}{L_z}}sin(\frac{\alpha\pi}{L_z}z) \tag{4.22}$$

the form factor is easily calculated as

$$\begin{aligned}F(\alpha,\beta) &= \frac{3}{2L_z}, \quad when \quad \alpha=\beta\\ F(\alpha,\beta) &= \frac{1}{L_z}, \quad when \quad \alpha\neq\beta\end{aligned} \tag{4.23}$$

The electron-phonon matrix element, $|M_q|^2$, for several phonon mechanisms has been obtained from the deformation potential theory. The magnitude of oscillating potential is multiplication of the amplitude of the oscillation ($|A_q|$) and the deformation potential ($|K_q|$), i.e.,$|M_q|^2 = V|A_q|^2|K_q|^2$. The amplitude for one phonon in bulk is $|A_q| = \sqrt{\frac{\hbar}{2\rho V\omega_q}}$, where ρ is density and ω_q is the phonon frequency.

For four specific cases of phonon scattering, each one has the following deformation potential as

$$\begin{aligned}|K_q|^2 &= q^2D_A^2 && (Acoustic\ Phonon),\\ |K_q|^2 &= D_o^2 && (Optical\ Phonon),\\ |K_q|^2 &= \left(\frac{ee_{PZ}}{\epsilon_S\epsilon_0}\right)^2 && (Piezo-electric),\\ |K_q|^2 &= \rho e^2\omega_0^2\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon_0}\right)\frac{q^2}{(q^2+q_0^2)^2} && (Polar\ Optical\ Phonon)\end{aligned} \tag{4.24}$$

We can simplify the Eq. (4.20), when $|M_q|^2\left(n_q + \frac{1}{2} \pm \frac{1}{2}\right)$ is independent of q_x, because the sum over longitudinal wave vector produces the delta function, i.e., $\sum_{q_x} e^{iq_x(x_1-x_2)} = \frac{L_x}{a_x}\delta(x_1 - x_2)$. Therefore, the self-energy functions of scattering can be expressed only with the local site terms. This on-site local scattering significantly reduces computational burden because diagonal matrices of self-energy and lesser/greater-scattering function allow the recursive inversion method instead of full matrix inversion. Within above four phonon scattering, acoustic and optical phonon can be approximated as independent of q_x.
For acoustic phonon, scattering functions can be simplified by the two approximation. First, we can rewrite the number of acoustic phonon as $n_q \approx n_q + 1 \approx \frac{kT}{\hbar\omega_q}$ when acoustic phonon energy is much less than kT. Second, frequency of acoustic phonon is proportional to sound velocity and moment, i.e., $\omega_q \approx v_s q$. Then, $|M_q|^2(n_q + \frac{1}{2} \pm \frac{1}{2})$ can be rewritten as $\frac{D_A^2 kT}{2\rho v_s^2}$ which is independent of the wave vector of phonon. Therefore, the self-energy functions for acoustic phonon is obtained as

$$\begin{aligned}
\Sigma^<_{AP}(x,\alpha,k_y,E) &= \frac{2}{a_x a_y}\frac{D_A^2 kT}{2\rho v_s^2}\sum_\beta \int \frac{dq_y}{2\pi} G^<(x,\beta,k_y+q_y,E)F(\alpha,\beta) \\
\Sigma^>_{AP}(x,\alpha,k_y,E) &= \frac{2}{a_x a_y}\frac{D_A^2 kT}{2\rho v_s^2}\sum_\beta \int \frac{dq_y}{2\pi} G^>(x,\beta,k_y+q_y,E)F(\alpha,\beta)
\end{aligned} \tag{4.25}$$

For optical phonon, $E - q$ relation can be assumed as a fixed-energy mode, i.e., $\omega_q \approx \omega_0$ and deformation potential is independent of the wave vector of phonon. Therefore, electron-phonon matrix element can be regarded as constant with respect to phonon momentum. Then similar with acoustic phonon, the summation over q_x produces a delta function. The self-energy functions for optic phonon can be rewritten as

$$\begin{aligned}
\Sigma^<_{OP}(x,\alpha,k_y,E) = \frac{1}{a_x a_y}\frac{\hbar D_o^2}{2\rho\omega_o}\sum_\beta \int \frac{dq_y}{2\pi}[&(n_q+1)G^<(x,\beta,k_y+q_y,E+\hbar\omega_o) \\
&+n_q G^<(x,\beta,k_y-q_y,E-\hbar\omega_o)]\ F(\alpha,\beta), \\
\Sigma^>_{OP}(x,\alpha,k_y,E) = \frac{1}{a_x a_y}\frac{\hbar D_o^2}{2\rho\omega_o}\sum_\beta \int \frac{dq_y}{2\pi}[&(n_q+1)G^>(x,\beta,k_y-q_y,E-\hbar\omega_o) \\
&+n_q G^>(x,\beta,k_y+q_y,E+\hbar\omega_o)]\ F(\alpha,\beta),
\end{aligned} \tag{4.26}$$

For POP and PZ, electron-phonon matrix element have phonon wave vector dependent term and self-energy functions require off-diagonal term to include nonlocal scattering which be represented by summing with respect to q_x.

For piezo-electric phonons, the lesser/greater self-energies are

$$\begin{aligned}\Sigma^{<}_{PZ}(x,\alpha,k_y,E) = \frac{K_{PZ}}{a_x a_y}\sum_{\beta}\iint \frac{dq_x}{2\pi}\frac{dq_y}{2\pi}\frac{1}{q^2}\Big[e^{iq_x(x_1-x_2)}G^{<}(x_1,x_2,\beta,k_y+q_y,E)\\ +e^{iq_x(x_2-x_1)}G^{<}(x,\beta,k_y-q_y,E)\Big]F(\alpha,\beta),\\ \Sigma^{>}_{PZ}(x,\alpha,k_y,E) = \frac{K_{PZ}}{a_x a_y}\sum_{\beta}\iint \frac{dq_x}{2\pi}\frac{dq_y}{2\pi}\frac{1}{q^2}\Big[e^{iq_x(x_1-x_2)}G^{>}(x_1,x_2,\beta,k_y-q_y,E)\\ +e^{iq_x(x_2-x_1)}G^{>}(x,\beta,k_y+q_y,E)\Big]F(\alpha,\beta),\end{aligned} \tag{4.27}$$

where $K_{PZ} = \left(\frac{e e_{PZ}}{\epsilon_S \epsilon_0}\right)^2 \frac{k_B T}{2\rho v_s^2}$.
For polar optical phonon,the lesser/greater self-energies are

$$\begin{aligned}\Sigma^{<}_{POP}(x,\alpha,k_y,E) = \frac{K_{POP}}{a_x a_y}\sum_{\beta}\iint \frac{dq_x}{2\pi}\frac{dq_y}{2\pi}\frac{q^2}{(q^2+q_0^2)^2}\Big[e^{iq_x(x_1-x_2)}G^{<}(x_1,x_2,\beta,k_y+q_y,E+\hbar\omega_q)\\ +e^{iq_x(x_2-x_1)}G^{<}(x,\beta,k_y-q_y,E-\hbar\omega_q)\Big]F(\alpha,\beta),\\ \Sigma^{>}_{POP}(x,\alpha,k_y,E) = \frac{K_{POP}}{a_x a_y}\sum_{\beta}\iint \frac{dq_x}{2\pi}\frac{dq_y}{2\pi}\frac{q^2}{(q^2+q_0^2)^2}\Big[e^{iq_x(x_1-x_2)}G^{>}(x_1,x_2,\beta,k_y-q_y,E-\hbar\omega_q)\\ +e^{iq_x(x_2-x_1)}G^{>}(x,\beta,k_y+q_y,E+\hbar\omega_q)\Big]F(\alpha,\beta),\end{aligned} \tag{4.28}$$

where $K_{POP} = \frac{e^2\hbar\omega_o}{2}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon_0}\right)$ and q_0 denotes the inverse screening length. This nonlocal scattering terms increases the complexity of NEGF calculations significantly and require huge amount of memory to couple momentum and energy in different position. Since self-energy of nonlocal scattering decease exponentially with respect to distance, we ignore nonlocal scattering to reduce calculation burden.

For polar optical phonons of local scattering, the lesser/greater self-energies are

$$\begin{aligned}\Sigma^{<}_{POP}(x,\alpha,k_y,E) = \frac{1}{a_x a_y}K_{POP}\sum_{\beta}\int \frac{dq_y}{2\pi}\frac{q^2}{(q^2+q_0^2)^2}[(n_q+1)G^{<}(x,\beta,k_y+q_y,E+\hbar\omega_q)\\ +n_q G^{<}(x,\beta,k_y-q_y,E-\hbar\omega_q)]\,F(\alpha,\beta),\\ \Sigma^{>}_{POP}(x,\alpha,k_y,E) = \frac{1}{a_x a_y}K_{POP}\sum_{\beta}\int \frac{dq_y}{2\pi}\frac{q^2}{(q^2+q_0^2)^2}[(n_q+1)G^{>}(x,\beta,k_y-q_y,E-\hbar\omega_q)\\ +n_q G^{>}(x,\beta,k_y+q_y,E+\hbar\omega_q)]\,F(\alpha,\beta),\end{aligned} \tag{4.29}$$

Following procedure is used to determine q^2 in Eq.(4.27) Eq.(4.28)
1) Determine k_x with given E and k_y.
2) Determine $k_{x_1}^{'}$ with given $E + \hbar\omega_q$ and $k_y + q_y$ for emission
and $k_{x_2}^{'}$ at given $E - \hbar\omega_q$ and $k_y - q_y$ for absorption
3) Determine from $k_z = \dfrac{(\beta - \alpha)\pi}{L}$
4) For emission $q^2 = \left(k_{x_1}^{'} - k_x\right)^2 + q_y^2 + k_z^2$, for absorption $q^2 = \left(k_{x_2}^{'} - k_x\right)^2 + q_y^2 + k_z^2$

NEGF versus Fermi's golden rule

To estimate the physical implication of self-energies for phonon scattering, we compare the scattering rate obtained from mode-space NEGF calculation with the FGR, i.e., first-order perturbation theory, since scattering time has the relationship with the broadening function(Γ) of NEGF as:

$$\Gamma_S = \Sigma_S^{<} + \Sigma_S^{>} = \frac{\hbar}{\tau} \tag{4.30}$$

where τ is the scattering time.
As the self-energys of scattering ($\Sigma_S, \Sigma_S^{\lessgtr}$) are interdependent with the Green's functions($G, G^{\lessgtr}$), we determine them self-consistently as following procedure.
1) Initially, determine G , $G^{\lessgtr}$ in ballistic regime, i.e. , $\Sigma_S = \Sigma_S^{\lessgtr} = 0$
2) Determine $\Sigma_S^{\lessgtr}$ using Eq.(4.25) (4.29) and then calculate Σ_S with below equations

$$\Sigma_S(E) = -\frac{i}{2}[\Sigma_S^{<}(E) + \Sigma_S^{>}(E)] \tag{4.31}$$

3) Build new G using Eq.(4.4). then, $G^{\lessgtr}$ are obtained by

$$G^{<}(E) = G\Sigma^{<}G^{\dagger}, G^{>}(E) = G\Sigma^{>}G^{\dagger} \tag{4.32}$$

4) Repeat steps 2) and 3) until $\Sigma_S(new) - \Sigma_S(old)$ converge.
In Eq. (4.31), $P\int \dfrac{dE'}{2\pi}\dfrac{\Gamma_S(E')}{E - E'}$, which is responsible for the renormalization of energy, is neglected since it does not have any significant effect on the results when the density of state change continuously. Note that all of Green's functions and self-energies are a function of the total energy E and k_y as shown in Fig. 4.1. Thus, the integrals over the Brillouin zone with respect to the q_y in Eq. 4.25, 4.26 and 4.29 require additional conditions that self-energy and Green's functions should be zero when total energy is less than the minimum of conduction band at k_y. Without this constraint, $\Sigma^{\lessgtr}$ have values in the region where there is no electron sate (the navy areas in Fig. 4.1), which causes additional broadenings because of the coupling between the physically meaningless self-energy and Green's functions.
In order to make calculations within the relatively small energy bandwidth, we use a imaginary semiconductor with a single band with $m_x^* = m_y^* = 1$. The other important parameters for scattering rate were InAs values and listed on Table. 4.1.

D_A (eV)	D_o (eV/m)	ρ (kg/m^3)	v_s (m/s)	a_x, a_y (nm)	ω_o (eV)	ε_∞	ε_0
4.9	5.2×10^{10}	5567	4280	0.303	0.3	12.75	15.15

Table 4.1: Simulated parameters for phonon scattering.

Fig. 4.1 shows the distribution of the scattering rate in the k_y-E space calculated by the NEGF method for the AP, OP, and POP cases. From the E-k relation, which follows $E(k_x, k_y) = E_0 + 2t_x - 2t_x cos(k_x a) + 2t_y - 2t_y cos(k_y a)$, the scattering rate are converted into $k_x - k_y$ space and is compared with the scattering rate obtained from FGR in Fig. 4.2. Notably the same coupling constants for acoustic and optical phonons have been used for NEGF and FGR. Since the AP is elastic scattering and the matrix element of scattering is independent of the wave vector of phonon, the scattering rate is proportional to density of states (DOS). The DOS has the largest value near $E = 2$ for this imaginary material, the scattering rate also shows the largest value in the vicinity. The scattering rate of OP is almost similar to that of AP, but it has a small value near the conduction band minimum and maximum because of the energy-momentum conservation. In other words, the scattering rate shows much lower values near $k_x = k_y = 0$ because only electrons whose energy exceed $\hbar\omega_o$ can emit optical phonons. The energy-dependent scattering rate is shown in Fig. 4.3. Importantly, for both acoustic and optical phonons, the NEGF calculated scattering rates closely match those calculated from FGR in both functional dependence with respect to k_x and k_y and also in amplitude. This means that one should be able to use the well-calibrated coupling constants from FGR for NEGF calculations. For polar optical phonons, there are slight differences between NEGF and FGR. Although there is some deviation due to the local scattering approximation, the qualitative trends with respect to energy and screening length remain intact.

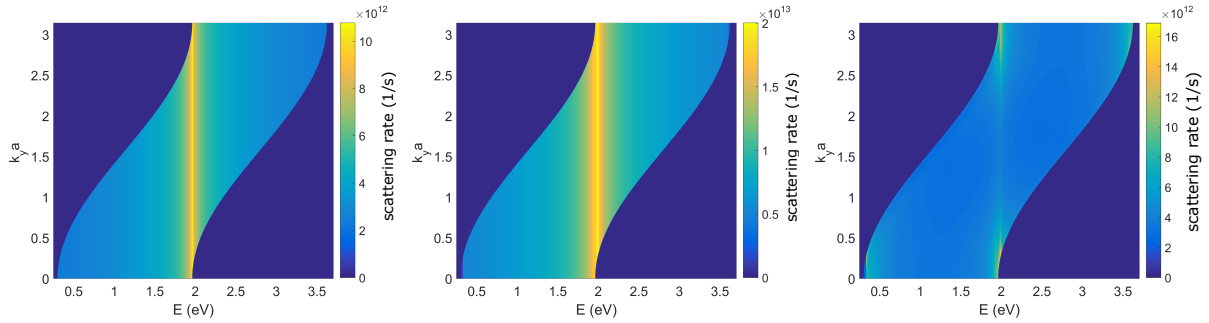


Figure 4.1: Scattering rates of NEGF simulation as a function of electron energy and $k_y a$ for (a) Acoustic phonon, (b) Optical phonon (c) Polar optical phonon.

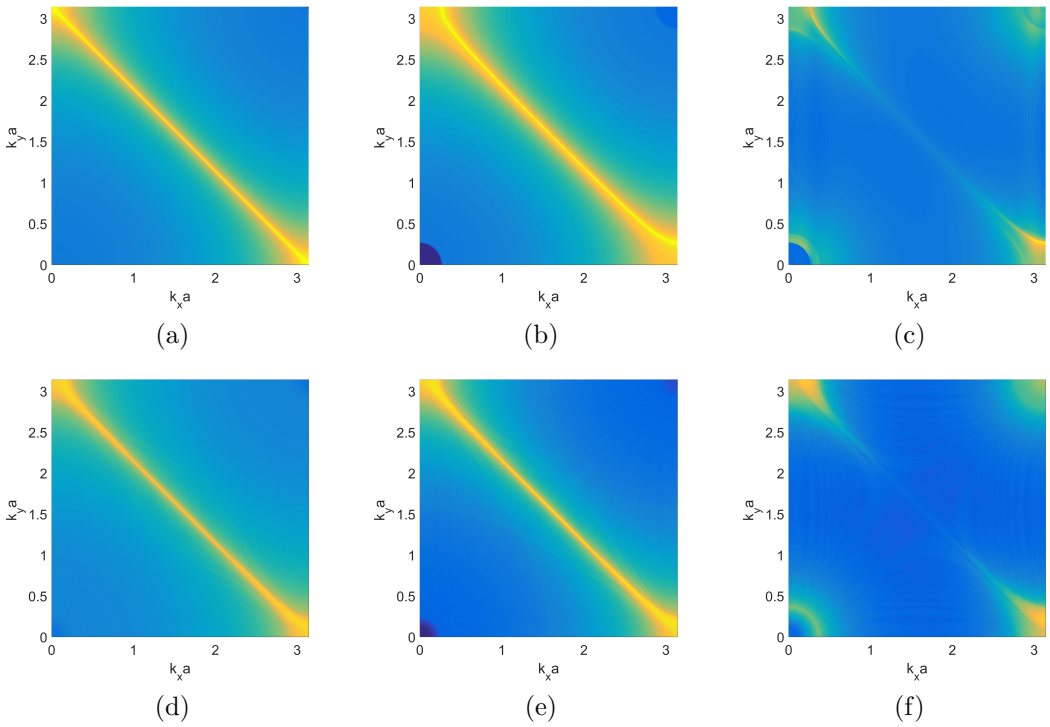


Figure 4.2: Scattering rate as a function of $k_x a$ and $k_y a$. (a),(b) and(c)are obtained with NEGF. (d),(e) and(f)are obtained with FGR. (a),(d): Acoustic phonon, (b),(e): Optical Phonon (c),(f): Polar Optical Phonon.

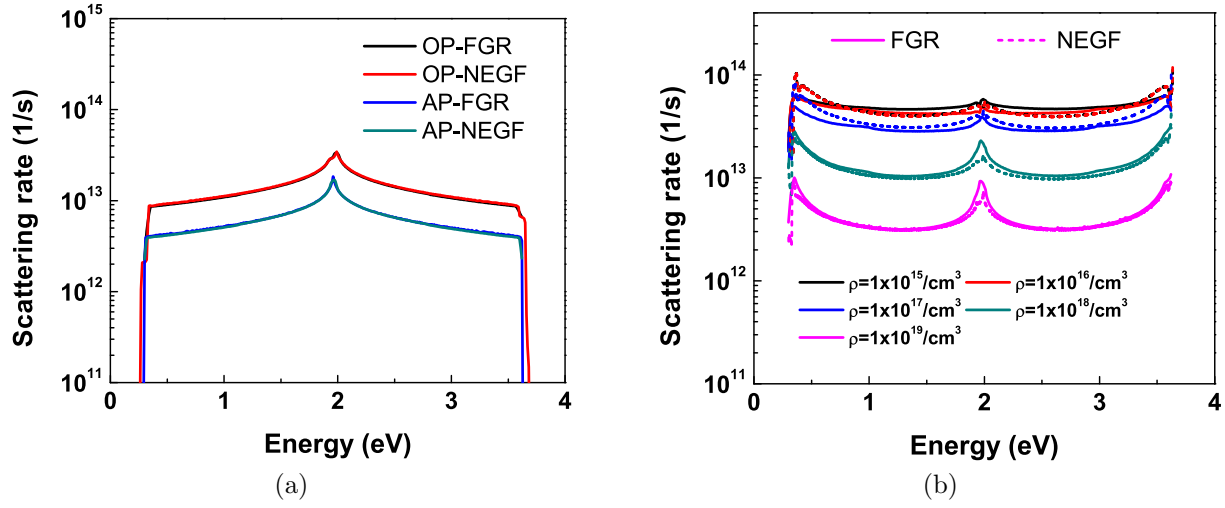
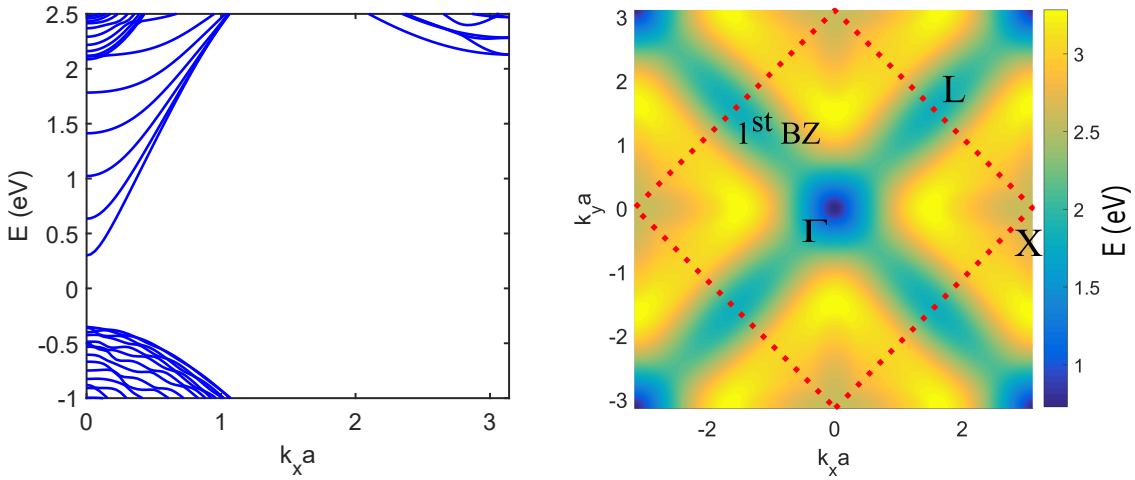


Figure 4.3: Scattering rate with respect to energy. (a) Acoustic phonon and optical phonon, (b) Polar optical phonon.

Multi-mass mode space approach

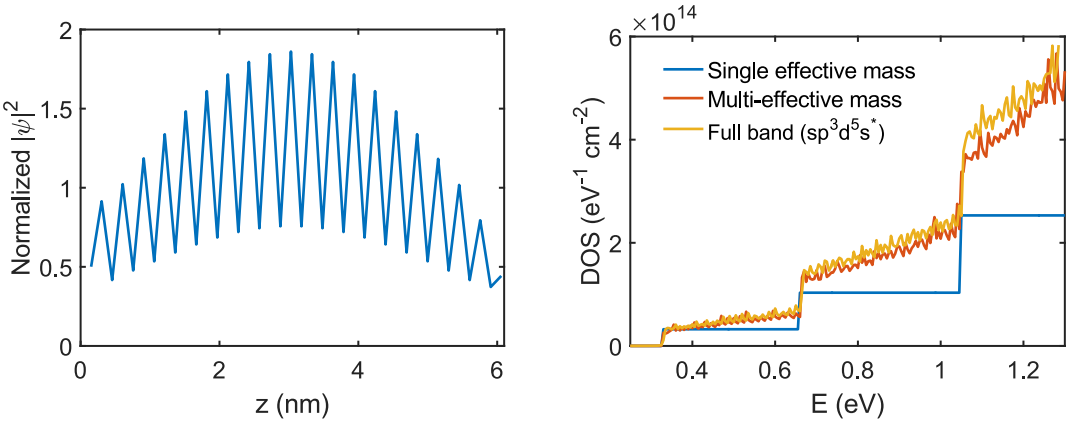
In deeply scaled devices, the quantum effect along the thickness direction forms discrete subbands with energies far from the bulk band edges. Because of nonparabolicity, the use of bulk effective masses in a III-V UTB MOSFET do not provide a good description of these subband energies and their dispersions [61] [89]. Since a correct description of the bandstructure is crucial to predict and explain current flow, we use a semi-empirical atomistic tight-binding approach with 10 orbitals consisting of an $sp^3d^5s^*$ basis without spin-orbit coupling. This model is a powerful tool that captures nonparabolicity and band-to-band coupling and generate accurate charge density on atomic sites because Hamiltonians of TB is built with the basis set based on local combination of atomic orbitals (LCAO). The tight-binding parameters of Ref. [58] are used and passivation of the surface dangling bonds is applied. The $E-k$ dispersion relation at ky = 0 and the lowest conduction band energy of electrons within the 2D Brillouin zone are shown in Fig. 4.4.



(a) (b)

Figure 4.4: (a) $sp^3d^5s^*$ TB bandstructures of geometrically-confined (along [100]), 6 nm thick InAs(100) at ky = 0. (b) Lowest electron subband of 6 nm thick InAs(100). Red dashed line shows the 2D Brillouin zone.

There were several parabolic band approaches using a single effective mass extracted from the TB bandstructure of the confined structure for each unconfined direction [75, 90]. Within the parabolic band approximation, the conduction band energy (E_c) and transport wavevector (k_x) at a certain total energy (E), transverse wavevector (k_y), and subband (α)



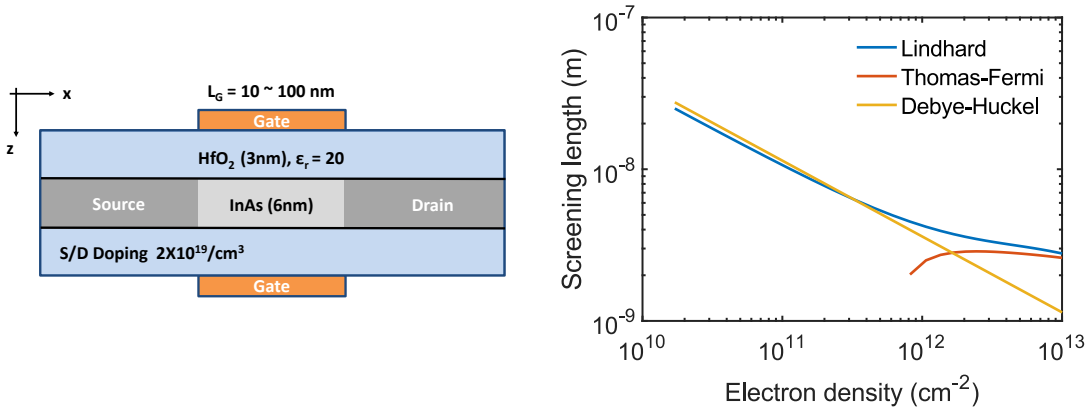
(a) (b)

Figure 4.5: (a) Normalized probability density of first subband of 6 nm InAs. (b) Density of state of 6 nm InAs.

are obtained as

$$\begin{aligned} E_c(k_y,\alpha) &= E_{c_0} + \frac{\hbar^2 k_y^2}{2m_y^*(\alpha)} \\ k_x(E,k_y,\alpha) &= cos^{-1}\left(1 - \frac{E - U - E_c(k_y,\alpha)}{2t_x(\alpha)}\right) \end{aligned} \tag{4.33}$$

where, $t_x(\alpha) = \dfrac{\hbar^2}{2m_x^*(\alpha)a^2}$. In ballistic transport, the parabolic band approximation provides enormous savings in computational burden because the integral over transverse energy can be done analytically and their contribution is taken into account in the Fermi–Dirac integral. Furthermore, for treating scattering, there are several approaches that solve transverse momentum integrals analytically that further simplify the simulation just like the method of obtaining the current and charge density in the ballistic transport [91–93]. While this approach shows an acceptable result for the acoustic and optic phonon scattering in non-polar crystals by transforming the NEGF equations in terms of integrated transverse momentum modes, it leads to unphysical results for optical phonon scattering because scattering potentials significantly depend on the phonon wavenumber (q). For this study, instead of using parabolic band approximation, we directly extract $E_c(k_y,\alpha)$, $m_x^*(E,k_y,\alpha)$ and $k_x^*(E,k_y,\alpha)$ from the TB bandstructure (named multi-effective mode space for convenience). Those values are used for phonon scattering to phonon self-energy functions to take care of energy momentum conservations. When $E < U + E_c(k_y,\alpha)$, since it is difficult to extract physical value from the TB method, values at the conduction band minima at k_y are used. For 6 nm InAs, the first subband eigenfunction from TB is shown in Fig. 4.5a. It shows zig-zag charge distribution because cation-sites have greater probability density than anion-sites.



(a) (b)

Figure 4.6: (a) Schematic structure of the simulated double-gate ultra-thin body MOSFETs. (b) Screening length calculated from different model as function of electron density.

Fig. 4.5b shows a comparison between the DOS of the 6-nm InAs UTB extracted from the parabolic band approximation (named single effective mass for convenience), multi-effective mass mode, and full band using $sp^3d^5s^*$ basis. While the DOS of a mode space using a single mass has a stepwise distribution due to constant effective mass, the DOS of multi-effective mode space, which increases with energy, agrees well with the DOS of the full-band NEGF because it uses the effective masses, which increases with E and k_y due to the nonparabolic effect. Fig. 4.6a shows the device structure of the 2-D double gate (DG) UTB MOSFETs. Gate length is varied from 10 to 100nm. Gate voltages are imposed as Dirichlet boundary conditions in the Poisson solver. The channel material is the 6-nm-thick InAs(100), whose transport direction is [100]. The doping level of the channel is zero and that of source/drain is 2×10^{19} cm^{-3}. The length of source/drain is 18nm. The 3-nm thick HfO_2 ($\epsilon_r = 20$) is used as the dielectric material. Periodic boundary conditions are used along the width and the transverse momentum modes are summed numerically in the calculation of charge densities and current. NEGF equations with the scattering are solved iteratively along with Poisson's equation until a self-consistency between charge and electrostatic potential is achieved. Given that our intent is to investigate the effective mobility, we have ignored band-to-band tunneling current which becomes important at large negative gate bias. For self-energies of the phonon scattering, only POP are included because the mobility of the III-V bulk semiconductor is dominated by the polar optical phonon scattering at room temperature. For POP scattering self energy, we compute the inverse screening length, q_0, at different position. The screen length can be obtained from the electron density at each point. The Lindhard screening theory gives the screening length at each point as a function of the charge density at that point (Fig. 4.6b).

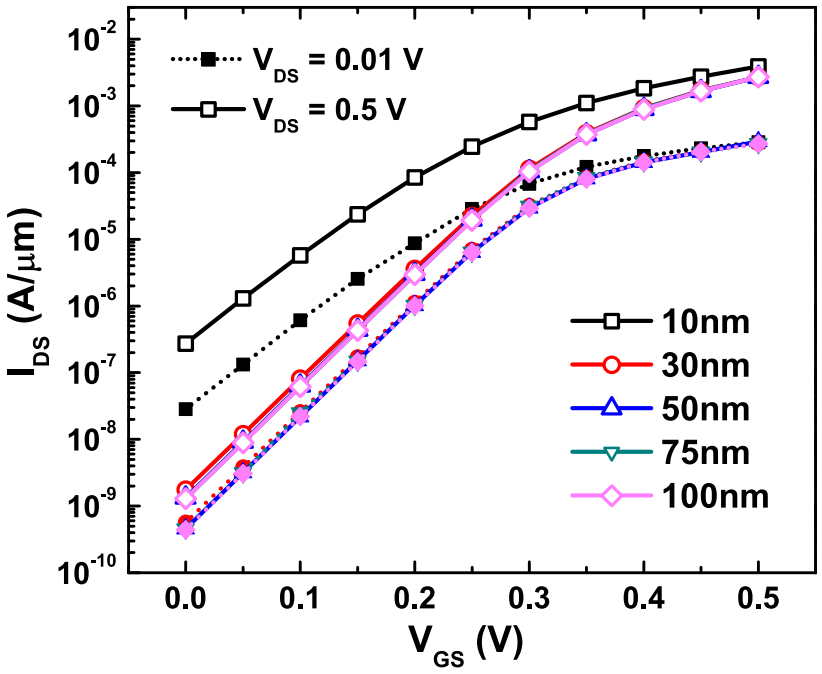


Figure 4.7: Ballistic transfer curves (I_{DS} vs. V_{GS}) for double-gated InAs MOSFETs with different gate lengths at V_{DS} = 0.01 (dashed lines) and 0.5 V(solid lines).

4.3 Results and discussion

Ballistic transport

From ballistic NEGF simulations, the transfer characteristics of of InAs MOSFETs is compared in Fig. 4.7 at V_{DS} = 0.01 V and 0.5 V. The 10-nm-long MOSFET shows deteriorated Subthreshold Swing (SS) and drain-induced barrier lowering (DIBL) than the MOSFET with a longer gate length. Fig. 4.8 show the energy-resolved current spectra at V_{GS} = 0.1 V and V_{DS} = 0.5 V for the ballistic simulation of the MOSFETs with L_G = 10 nm and L_G = 30 nm. White dashed line shows the profile of the conduction band minimum. The source-drain tunneling is a major source of the short channel effect of the MOSFET with L_G = 10nm, since the majority current flows through the electrical potential barrier in contrast to the MOSFET with L_G=30nm. From the $I_{DS} - V_{GS}$ curve in the linear region, we can calculate the ballistic mobility (μ_B) from

$$\frac{I_{DS}}{W} = Q_i(0)\mu_B \frac{V_{DS}}{L_G} \tag{4.34}$$

where $Q_i(0)$ is the sheet electron density at the beginning of the channel. Here, ballistic mobility is the non-physical mobility to express the conventional $I_{DS} - V_{DS}$ equation, which is the function of L_G. Furthermore, according to the Natoris theory of ballistic MOSFET [94], current can be expressed as a function of the injection velocity at the virtual source, which

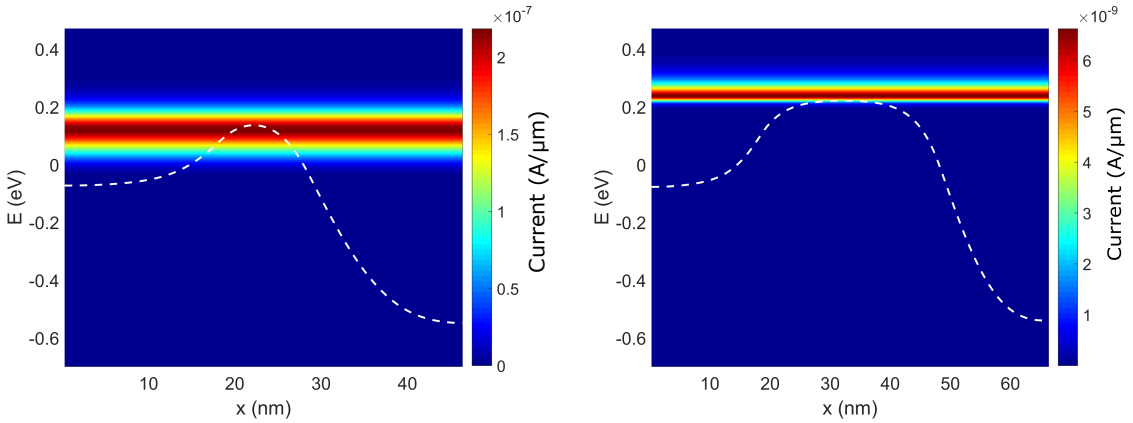


Figure 4.8: Energy-resolved current spectrum of ballistic transport at V_{GS} = 0.1 V and V_{DS} = 0.5 V for (a) L_G = 10 nm and (b) L_G = 30 nm.

is independent of the length of the device. Ballistic mobility can be written as [95, 96]

$$\mu_B = \frac{v_T q L_G}{2k_B T} \frac{\mathscr{F}_{-1/2}(\eta)}{\mathscr{F}_0(\eta)} \tag{4.35}$$

where $v_T = \sqrt{2k_B T/\pi m^*}$ is the unidirectional thermal velocity of the non-degenerate electrons and $\eta = (E_F - E_C)/k_B T$. Fermi-Dirac correction term$\left(\frac{\mathscr{F}_{-1/2}(\eta)}{\mathscr{F}_0(\eta)}\right)$ counts for the impact of degeneracy of the electrons. Since the carrier density in a semiconductor is a function of the Fermi level (E_F), which can be obtained by integrating the product of the density of states with a probability density function, as shown in Fig. 4.9a, the Fermi-Dirac correction term is obtained with respect to charge density in Fig. 4.9b. Fig. 4.10 shows the results of two different μ_B calculations with two different method from Eq. (4.34) and Eq. (4.35). Except L_G = 10 nm, Both μ_B is constant when V_{GS} is smaller than threshold voltage owing to $\mathscr{F}_{-1/2}(\eta) = \mathscr{F}_0(\eta) = e^{\eta}$ under nondegenerate condition. On the other hand, μ_B for L_G = 10 nm decreases as the V_{GS} decreases even in the nondegenerate limit. This is because electrons with relatively low energy contribute to the current through intra-band tunneling and the ratio of tunneling current to thermionic emission current increases as V_{GS} decreases.

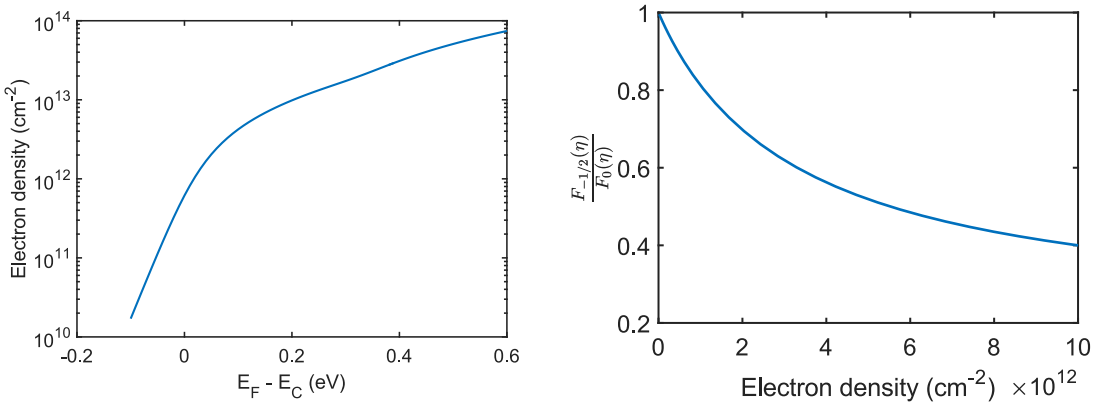


Figure 4.9: (a) Carrier density as a function of Fermi energy $E_F - E_C$ for 6 nm InAs UTB. (b) Fermi-Dirac correction term as a function of carrier density for 6 nm InAs UTB.

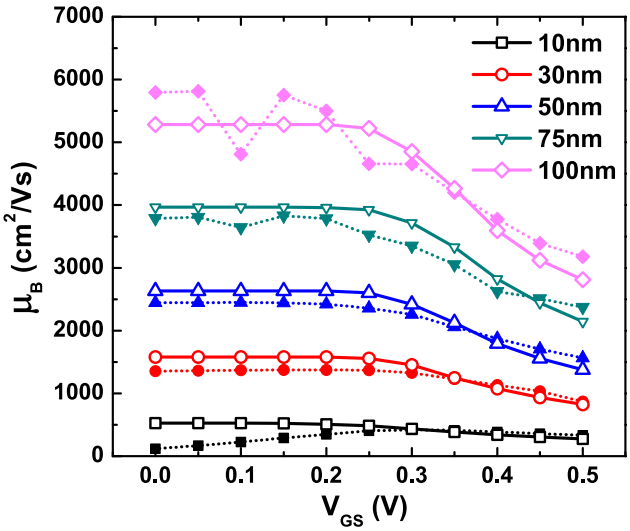


Figure 4.10: Ballistic mobility as a function of gate voltage (V_{GS}) for different gate lengths. The dashed lines are from Eq. (4.35), and the solid lines are from Eq. (4.34).

Phonon scattering

Fig. 4.11a shows the transfer characteristics in the linear region (V_{DS} = 0.01 V) under either ballistic or dissipative transport with POP scattering. The impact of the phonon scattering on different sizes of the devices have been explored on both linear and logarithmic plot and are shown to be compared in both the above-threshold and the subthreshold region. It is apparent that as the channel length increases, the enhanced probability of scattering causes

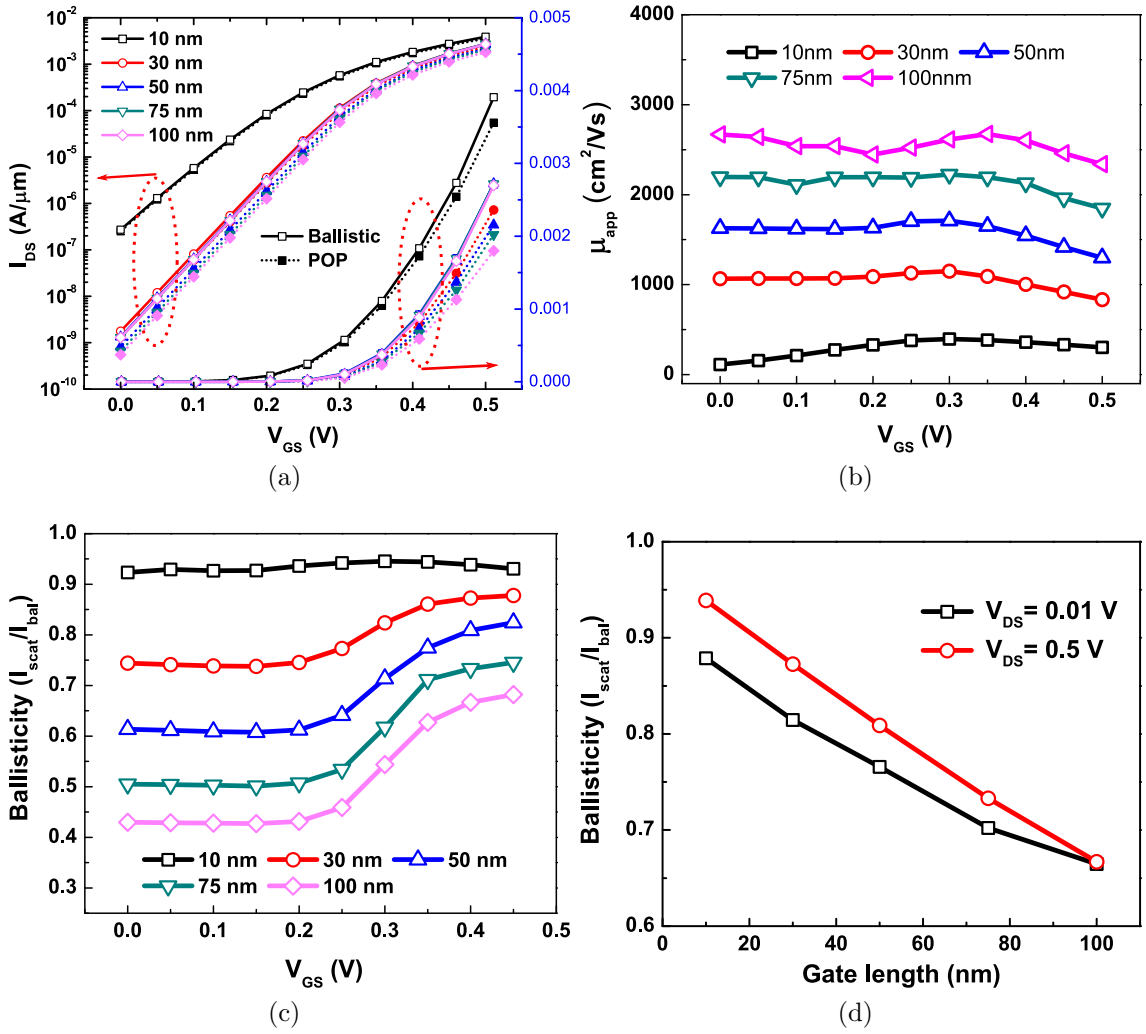


Figure 4.11: (a) I_{DS} – V_{GS} characteristics at V_{DS} = 0.5 V with different gate lengths on logarithmic (left axis) and linear scales (right axis). Sold lines are from ballistic simulation and dashed lines are from dissipative simulation (POP). (b) Apparent mobility from dissipative simulation with POP scattering as function of gate voltage (V_{GS}) for different gate lengths. (c) Ballisticity (I_{scat}/I_{bal}) as a function of gate voltage at V_{DS} = 0.5 V with POP scattering. (d) Ballisticity (I_{scat}/I_{bal}) as a function of gate length at V_{DS} = 0.01 V and V_{DS} = 0.5 V with POP scattering.

a greater decrease in the current. Furthermore, the effect of phonon scattering of POP also relies on the gate voltage and the drain voltage. For a better understanding of voltage dependence, we have plotted the ballisticity curve, which is the ratio of the scattering and ballistic current (I_{scat}/I_{bal}). Fig. 4.11c shows that the impact of POP is reduced as the gate voltage increases. The increased electron density, which is induced by gate voltage,

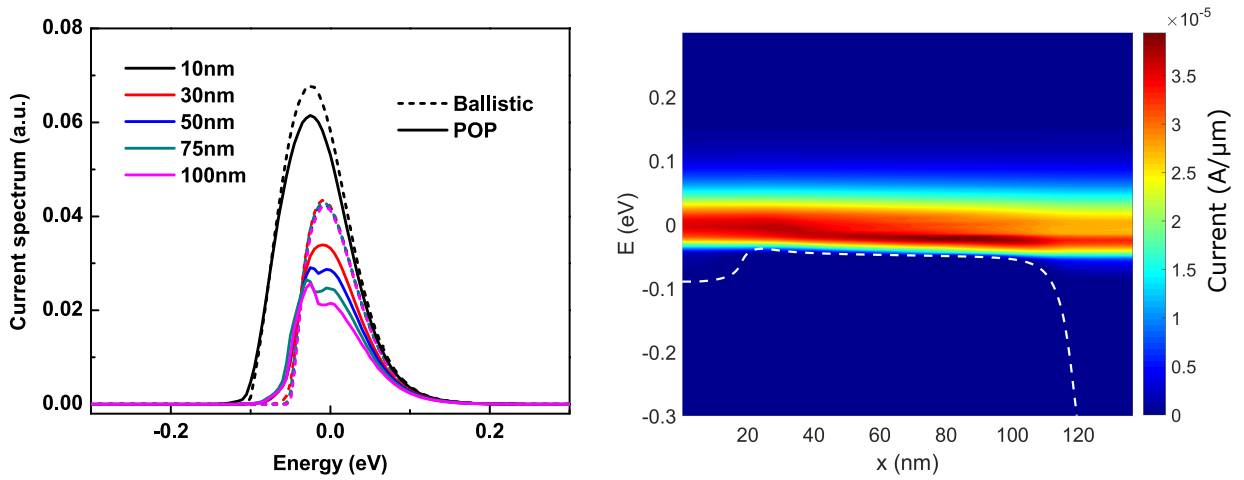


Figure 4.12: (a) Energy-resolved current with different gate length under ballistic (dashed line) and dissipative transport with POP (solid line). Calculations are performed at the drain side end at V_{GS} = 0.4 V and V_{DS} = 0.5 V. (b) Energy-resolved current spectrum at V_{GS} = 0.4 V and V_{DS} = 0.5 V with POP scattering.

results in a better screening of deformation potentials of POP, and leads to less efficient scattering by the dipole field of ion. Ballisticity also depends on the drain voltage. Fig. 4.11d shows that the number of scattering events decrease as the drain voltage increases. The impacts of drain bias become negligible in the long channel device (100 nm). Since the average velocity of the electron along the device increases as the drain bias increases, the time for electrons to stay in the device becomes smaller, and therefore, the probability of the scattering event relatively decreases. Furthermore, once the energy of the electron is relaxed by POP at a high V_{DS}, it is highly unlikely that electrons can travel back to the source, which increases the negative current, because of the insufficient energy to overcome the top of the potential barrier. In Fig. 4.12a, we have shown that the energy resolved current in the drain for a different gate length at V_{GS} = 0.4 V and V_{DS} = 0.5 V. While the ballistic simulation shows an almost identical current distribution regardless of the length of the device, except for very short channel MOSFET (L_G = 10 nm), the scattering simulation with POP clearly shows device length dependence of the current spectrum. When electrons propagate from the source to the drain, scattering with POP results in the broadening of the drain current spectrum as electrons can occupy certain energy levels, which are forbidden under the ballistic regime. Furthermore, the energy relaxation by emitting one or a few inelastic phonon produces a separation of the peak of the current spectrum. As the number of scattering events is proportional to the channel length, energy relaxation is enhanced as the channel length increases. This impact of energy relaxation caused by POP is clearly observed in the energy-resolved current spectrum, as shown in Fig. 4.12b.

As the channel length of an MOSFET is comparable with mean free path, the conven-

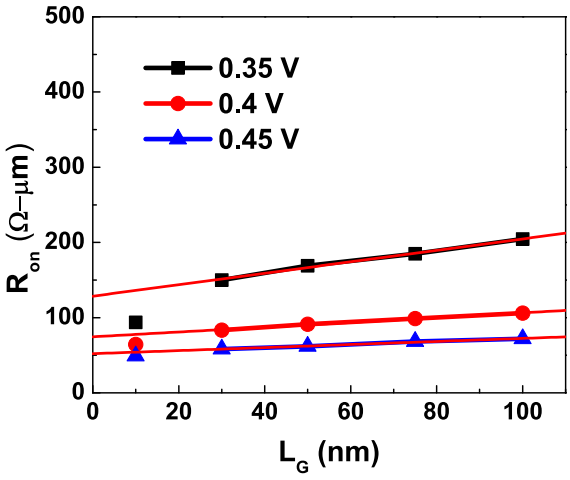


Figure 4.13: R_{on} versus L_G under different V_{GS} of 6 nm InAs MOSFETs with POP scattering. The slope of linear curve is proportional to inverse of μ_{eff} and y-axis intercept is $R_{on}(0) = R_{S/D} + R_B$.

tional mobility from Drudes model should be treated carefully. In the quasi-ballistic regime, the effective mobility in a long-channel equations should be replaced by the apparent channel mobility , which is governed by Matthiessen's law.

$$\frac{1}{\mu_{app}} = \frac{1}{\mu_B} + \frac{1}{\mu_{eff}} \tag{4.36}$$

Here, μ_{eff} is the diffusive mobility in the scattering-dominated long channel. From the conventional current equation at the low field, the on-resistance for a given length, L_G, is given as

$$\begin{aligned} R_{on}(L_G) =& R_{S/D} + \frac{V_{DS}}{I_{on}(L_G)} \\ =& R_{S/D} + \frac{1}{\mu_{app}} \frac{L_G}{Q_i(0)} \\ =& R_{S/D} + \left(\frac{1}{\mu_B} + \frac{1}{\mu_{eff}} \right) \frac{L_G}{Q_i(0)} \end{aligned} \tag{4.37}$$

where $R_{S/D}$ is the resistance of the S/D region. Since $\frac{1}{\mu_B} \frac{L_G}{Q_i(0)} = R_B$ is independent of L_G, the effective mobility is extracted from the slope of the linear $R_{on} - L_G$ curve. Since R_B has been calculated from the ballistic simulation, $R_{S/D}$ can be obtained from the y-intercept of linear $R_{on} - L_G$ curve using $R_{on}(0) = R_{S/D} + R_B$. The carrier mean free path (λ_{mfp}) is obtained from the scattering limited mobility equation, which is generally given as

$$\mu_{eff} = \frac{v_T q \lambda_{mfp}}{k_B T} \frac{\mathscr{F}_{-1/2}(\eta)}{\mathscr{F}_0(\eta)} \tag{4.38}$$

Fig. 4.13 shows a linear dependence of R_{on} versus L_G under different V_{GS}. The R_{on} value from L_G = 30 nm to 100 nm is well located on the linear fitting line. However, the R_{on} value of L_G = 10 nm is smaller than the linear fitting and the V_{GS} is smaller and the deviation is larger. In the short channel (L_G = 10 nm), as the V_{GS} increases, the ratio of the current flowing under the potential barrier increases, resulting in a R_{on} value less than the linear fit. The extracted results are summarized in Table. 4.2. As V_{GS} increases, μ_{eff} and λ_{mfp} increase together. However, we note that λ_{mfp} continues to increase while the mobility saturates. This is because the degeneracy of the electrons continues to decrease as the electron density increases.

V_{GS} (V)	Slope (Ω)	$Q_i(0)$ ($\#/cm^2$)	μ_{eff} ($cm^2/V\cdot s$)	R_B ($\Omega\cdot\mu m$)	$R_{S/D}$ ($\Omega\cdot\mu m$)	λ_{mfp} (nm)
0.35	761	1.12E+12	7323	128.5	6.1	171
0.4	319	2.22E+12	8814	74.5	4.6	243
0.45	203	3.49E+12	8810	52.0	3.3	280

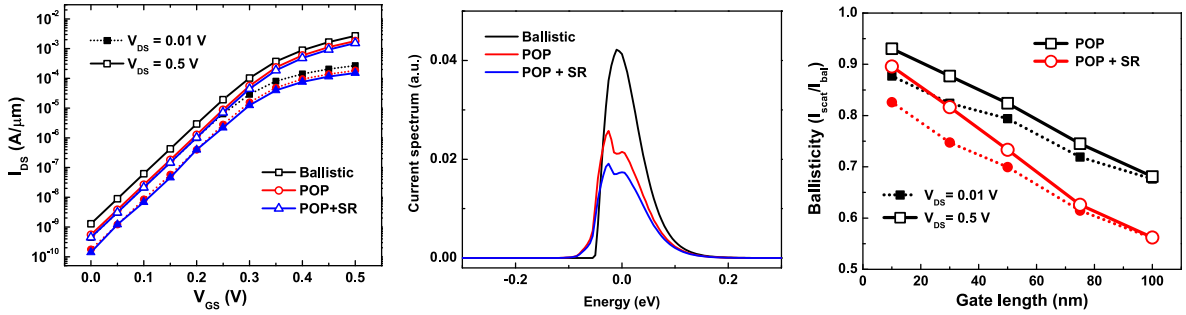
Table 4.2: Extracted characteristics of POP scattering for 6 nm InAs MOSFETs with different V_{GS}.

Surface roughness scattering

In contrast to phonon scattering, which is an intrinsic scattering mechanism, SR is an extrinsic mechanism, which highly depends on current fabrication technology and has possibility that further research may reduce the contributions. However, in UTB films, especially short channel device that apply high normal electric fields, the charge move very close to the channel-insulator interface, so SR scattering becomes a dominant scattering source. Here, we examine the effect of SR scattering on electron transport and adjust the parameter of SR to match the mobility from the experimental data. SR induces fluctuation of the potential along the channel and the subband wave function, of the electron charge, and of the interface polarization charge. For this study, we take into account the modification of the wavefunction due to SR, which is expressed with the Prange-Nee term and can be implemented in NEGF via an appropriate self-energy. From the semi-classical expression for surface roughness scattering, the square of the matrix element can be obtained as

$$\begin{aligned}\Sigma^{<}_{SR}(x,\alpha,k_y,E) = \frac{1}{a_x a_y}|M_{SR}|^2 \sum_{\beta}\int \frac{dq_y}{2\pi} Y(q)[G^{<}(x,\beta,k_y+q_y,E) \\ +G^{<}(x,\beta,k_y-q_y,E)]\ F(\alpha,\beta), \\ \Sigma^{>}_{SR}(x,\alpha,k_y,E) = \frac{1}{a_x a_y}|M_{SR}|^2 \sum_{\beta}\int \frac{dq_y}{2\pi} Y(q)[G^{>}(x,\beta,k_y-q_y,E) \\ +G^{>}(x,\beta,k_y+q_y,E)]\ F(\alpha,\beta),\end{aligned} \tag{4.39}$$

Here, Y(q) is the momentum dependence of the power spectrum, which is obtained from the Fourier transform of the roughness correlation, and $|M_{SR}|^2$ is the strength of scattering. We use the exponential autocorrelation function, i.e., $Y(q) = (1 + (\Lambda^2 q^2)/2)^{-3/2}$, and the square of the matrix element for the small quantum well, i.e., $|M_{SR}|^2 = \pi\Delta^2\Lambda^2 \dfrac{\pi^4\hbar^4}{m^{*2}T_s^6}$, where Λ is the correlation length, Δ is the roughness r.m.s., and T_s is film thickness. The values of Λ and Δ are chosen to fit the low-field mean free path of InAs, which is experimentally extracted as 150 nm ($\Lambda = 0.2\ nm$, $\Delta = 6\ nm$). In Fig. 4.14a, we compare the $I - V$ characteristics for InAs MOSFET under ballistic, with only POP scattering and both POP and SR scattering. In Fig. 4.14b and 4.14c, the result of simulation under POP and SR together exhibits a smaller energy-resolved current and ballisticity than the result obtained under only POP scattering because it have higher scattering probability. However, SR scattering has little impact on the shape of current spectrum and tendency of ballisticity because SR scattering is an elastic scattering, which does not include energy relaxation.



(a) (b) (c)

Figure 4.14: (a) I_{DS} - V_{GS} characteristics of 100 nm transistor at V_{DS} = 0.01 and 0.5 V under ballisic and dissipative transport. (b) Energy-resolved current at the drain side end at V_{GS} = 0.4 V and V_{DS} = 0.5 V. (c) Ballisticity (I_{scat} / I_{bal}) as a function of channel length at two different drain voltages and under two different scattering mechanisms (POP and POP + SR) at V_{GS} = 0.45 V.

In Fig. 4.15, we show a linear dependence of R_{on} versus L_G for InAs MOSFET computed under SR and POP scattering and the extracted results are summarized in Table. 4.2. As in the case of a stand-alone POP, μ_{eff} and λ_{mfp} with both SR and POP increase as V_{GS} increases. Mobility saturation appears to take place at higher voltages.

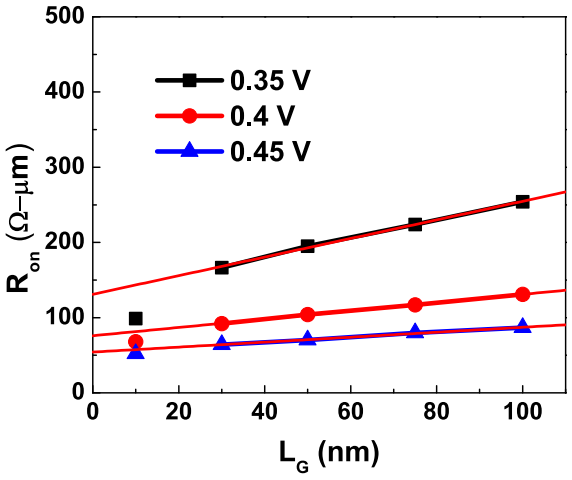
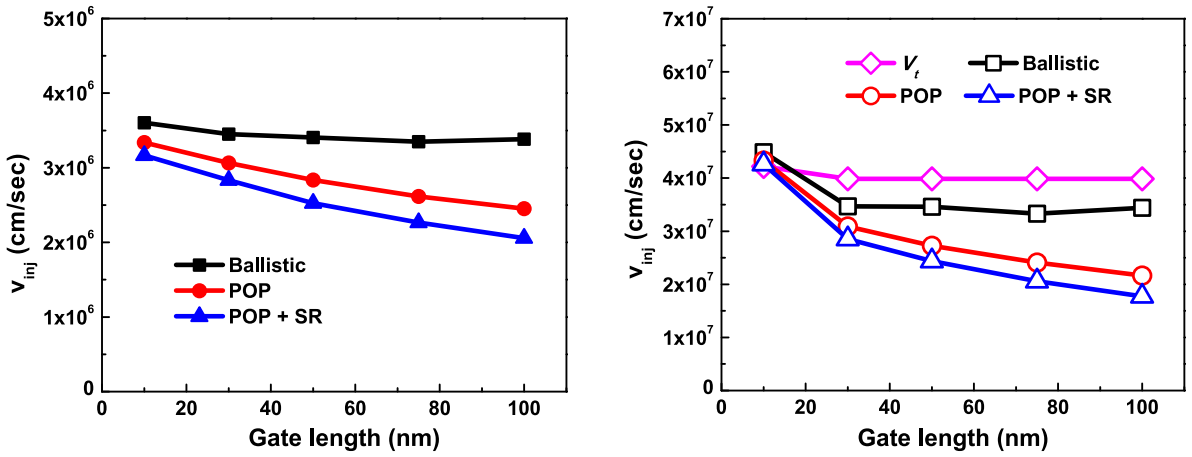


Figure 4.15: R_{on} versus L_G under different V_{GS} of 6 nm InAs MOSFETs with POP + SR scattering. The slope of linear curve is proportional to inverse of μ_{eff} and y-axis intersect is $R_{on}(0) = R_{S/D} + R_B$.

V_{GS} (V)	Slope (Ω)	$Q_i(0)$ $(\#/cm^2)$	μ_{eff} $(cm^2/V{\cdot}s)$	R_{on} $(\Omega{\cdot}\mu m\ )$	$R_{S/D}$ $(\Omega{\cdot}\mu m\ )$	λ_{mfp} (nm)
0.35	1237	1.11E+12	4513	131.1	9.2	108
0.4	550	2.20E+12	5115	75.9	6.5	136
0.45	330	3.48E+12	5413	54.2	5.5	160

Table 4.3: Extracted characteristics of SR scattering for 6nm InAs MOSFETs with different V_{GS}.

In Fig. 4.16, we compare the average injection velocity as a function of channel length under ballistic and dissipative transport simulation at high V_{DS} and low V_{DS}. For dissipative transport, the average v_{inj} tends to decrease with an increasing length. This is because the number of electrons traveling in the + x direction decreases as the number of backward electrons increases because of the scattering. This difference decreases with decreasing length, but v_{inj} at low V_{DS} still shows a difference by the scattering mechanism even at 10 nm because there are enough chance that scattering occurs during transport. This tendency is the same even when V_{DS} = 0.5 V. However, v_{inj} at the high V_{DS} for the 10-nm device converges to the unidirectional thermal velocity (v_T) value and shows a negligible difference by scattering (less than 10 %).



(a) (b)

Figure 4.16: (a) Average injection velocity at $V_{GS} = 0.45$ V and $V_{DS} = 0.01$ V. (b) Average injection velocity at $V_{GS} = 0.45$ V and $V_{DS} = 0.5$ V.

4.4 Conclusion

In this chapter, we present a 2D simulation framework for electronic transport incorporating electronphonon scattering based on the NEGF formalism. A mode space approach is adopted to reduce the computational complexity and the local self-energy functions for the electronphonon interactions have been rigorously derived within the self-consistent Born approximation. We verify the obtained self-energy functions by comparing the scattering rates obtained by NEGF formalism and FGR. We study the quantum transport of the double gate III-V UTB MOSFETs in the ballistic limit and in the presence of electronphonon interactions and surface roughness scattering. We also obtain the phonon-limited low field mobility and mean free path in the long channel limit by extracting the apparent mobility and the ballistic mobility for different gate lengths. Within POP scattering, which is defined by intrinsic material parameters, the mobility of InAs UTB MOSFETs was about 8800 $cm^2/V \cdot s$. For SR scattering, the correlation length and the roughness r.m.s. are chosen to fit the low-field mobility of InAs, which is experimentally extracted as approximately 5000 $cm^2/V \cdot s$. It is found that the electronphonon interactions reduce the drain current, and broaden the local density of states. For a 10-nm InAs UTB transistor, the on-sate current with only POP is 93.0 % of ballistic limit and reduced to 89.6 % with POP and SR scattering.

Chapter 5

Effect of phonon scattering on quantum transport in Si MOSFETs

5.1 Introduction

As the Internet of Things (IOT) applications become increasingly important, ultra-thin body metal oxide semiconductor field effect transistor (UTB MOSFET) devices are seeing a resurgence. While the UTB MOSFET devices cannot deliver the same ON current as FinFETs, their OFF state performance is significantly better, making them particularly interesting for low stand-by power applications. To properly assess the potential of the UTB MOSFET devices at the ultra scaled dimensions, two aspects need to be considered. Firstly, various types of tunneling could significantly reduce the ON/OFF ratio. Secondly, a rigorous description of carrier scattering is critical to understand how the current voltage characteristic will be in the quasi-ballistic regime. This means that one needs a quantum transport calculation, for example, following the Non-Equilibrium Green's Function (NEGF) formalism, including scattering to appropriately account for necessary physics. However, such a calculation is significantly complicated, especially for UTB MOSFETs where scattering couples the transverse modes, thereby restricting the ability for an analytical summation. Consequently, most of the existing work in this context have investigated device behavior only from ballistic calculations [66–68] or based on Buttiker probes [78, 92] or ignored the coupling between transverse modes [97].

Luisier [69] reported NEGF calculations where transverse modes are appropriately coupled and compared ON currents of GaSb, strained Si and $In_{0.53}Ga_{0.47}As$ channels. In this work, we use a similar approach to investigate effective mobility in UTB Si MOSFET devices. We have focused on the effective mobility [98–102] as this is a parameter that can be directly measured from experiments. Furthermore, we show that the scattering rates calculated using NEGF are very similar to those calculated using Fermi's Golden Rule (FGR). Given that FGR has been very successfully used to model scattering limited transport in semiconductor materials [103–105], the similarity between NEGF and FGR provides a pathway to choose

the coupling constants by fitting it to large channel length devices and then using those same constants to predict behavior for short channel length devices.

This chapter is organized in the following way. Section II discusses the simulation methodology including a brief discussion of how the scattering mechanisms are included in the calculation, how the scattering within the NEGF method compares with Fermi's Golden rule, how the device is set up and how the effects of band structure is captured. Section III discusses the main results in terms of current voltage characteristic and how the scattering mechanisms affect it. The effective mobility is calculated in this section and a framework is presented to show the mobility can be compared with experimental data. Finally, section IV summarizes the results.

5.2 Simulation methodology

Following NEGF formalism, we can write the retarded Green's function for each mode (α) at a given transverse momentum (k_y) as [79, 80, 106–109]

$$G(E, k_y, \alpha) = [\ EI - H - \Sigma_1 - \Sigma_2 - \Sigma_S]^{-1} \tag{5.1}$$

where H is the device Hamiltonian, Σ_1 and Σ_2 are the self-energy matrices representing the open boundary conditions of source and drain, respectively and Σ_S is the retarded self-energy matrix representing the interaction between the electron and scattering source. Within the self-consistent Born approximation, the electron-phonon interactions can be included in the lesser and greater self-energy functions in the following manner.
For the acoustic phonon (AP) scattering:

$$\Sigma^{\lessgtr}_{AP}(x, \alpha, k_y, E) = \frac{1}{a_x a_y} \frac{D_A^2 k_B T}{\rho v_s^2} \sum_{\beta} \int \frac{dq_y}{2\pi} G^{\lessgtr}(x, \beta, k_y + q_y, E) F(\alpha, \beta) \tag{5.2}$$

, where ρ is the density, v_s is the sound velocity in silicon, a_x and a_y are the discretization mesh in the x and y direction, D_A is the AP deformation potential and $F(\alpha, \beta)$ is the form-factor account for transition between subbands α and β.

$$F(\alpha, \beta) = \int_{-\infty}^{\infty} \psi^*_{\alpha}(z) \psi^*_{\beta}(z) \psi_{\alpha}(z) \psi_{\beta}(z) dz \tag{5.3}$$

For the optical phonon (OP) scattering:

$$\begin{aligned} \Sigma^{\lessgtr}_{OP}(x, \alpha, k_y, E) = & \frac{1}{a_x a_y} \frac{\hbar D_o^2}{2\rho\omega_o} \sum_{\beta} \int \frac{dq_y}{2\pi} [(n_q + 1) G^{\lessgtr}(x, \beta, k_y \pm q_y, E \pm \hbar\omega_o) \\ & + n_q G^{\lessgtr}(x, \beta, k_y \mp q_y, E \mp \hbar\omega_o)] F(\alpha, \beta) \end{aligned} \tag{5.4}$$

, where D_o is the OP deformation potential, ω_o is the angular frequency and $n_q = [exp(\hbar\omega_o / k_B T) - 1]^{-1}$ is the phonon occupation number.

The imaginary part of the scattering self-energy is,

$$\Sigma_S^i(E) = -\frac{i}{2}\Gamma_S(E) = -\frac{i}{2}[\Sigma_S^< + \Sigma_S^>] \tag{5.5}$$

The real part of the self-energy can be computed by the Hilbert transform of the level broadening of scattering (Γ). However, for materials with reasonably continuous density of states, the real part has negligible effect [85]. Therefore, we have ignored it in our calculations. The electron density obtained using the NEGF formula including the scattering is fed back to the Poisson solver until self-consistency is achieved. Once the electrostatic self-consistency is obtained, the current is calculated from

$$I = \frac{q}{\hbar}\sum_{\alpha}\frac{1}{W}\sum_{k_y} 2\int \frac{dE}{2\pi} Trace\left[A_\alpha \Sigma_{L,\alpha}^< - G_\alpha^< \Gamma_{L,\alpha}\right] \tag{5.6}$$

where A is the spectral function, $G^<$ is the electron correlation function, $\Sigma_L^<$ is the in-scattering function of contact L and Γ_L is the level broadening of contact L.

5.3 Results and discussion

An overview of quantum transport within 2D mode-space NEGF including scattering processes was presented in previous section. Now we will discuss the effect of scattering on the electron transport and the electrical characteristics in nanoscale devices. Fig.5.1 shows the simulated device structure. It has a double gate configuration. The channel consists of Si(100) with 6 nm thickness and transport direction is chosen to be along [100]. Gate

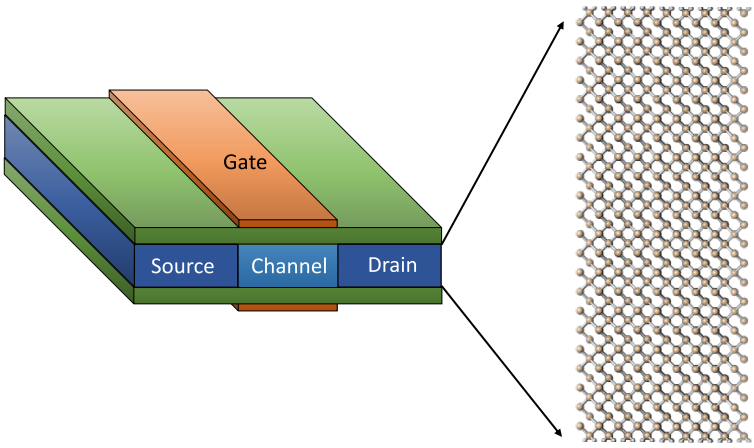
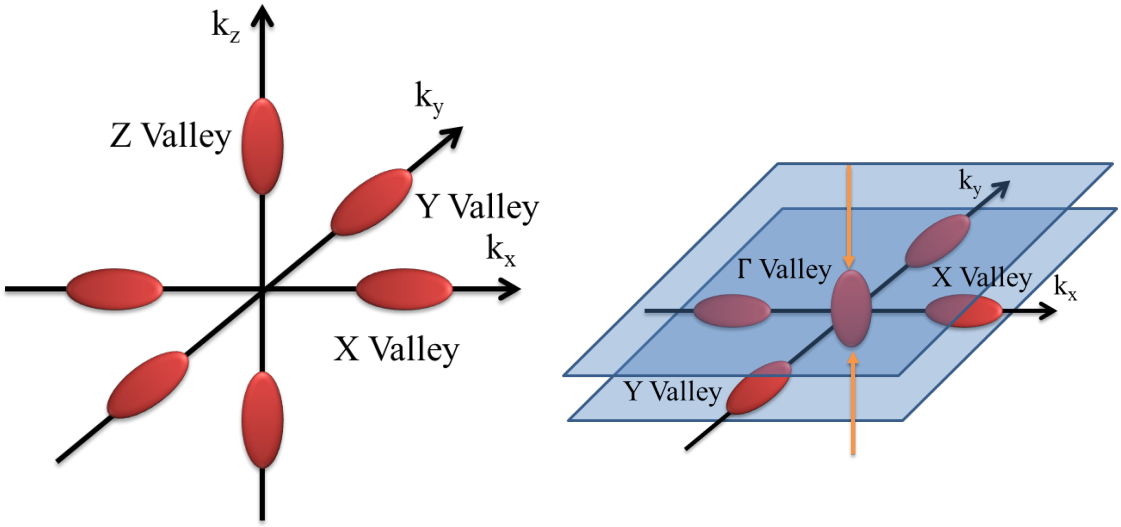


Figure 5.1: UTB DG MOSFET as a model device and the underlying atomic structure. A 20 orbital $sp^3d^5s^* - SO$ TB model is used to calculate the 2D subband dispersion for Si. There are 45 atomic layers of Si ($\approx$ 6nm).

length is varied from 10 to 100 nm. The doping level of the channel is zero and those of source/drain are of 4×10^{19} cm^{-3}. The length of source/drain region is 20 nm. A 3-nm-thick HfO_2 ($\varepsilon_r = 20$) is used as the dielectric material ($EOT = 0.585\ nm$). NEGF equations are solved in the mode space (discussed in more detail in the following) self-consistently with a two-dimensional finite element Poisson solver. Given that our intent is to investigate the effective mobility, we have ignored band-to-band tunneling current which becomes important at large negative gate bias. Also, impurity scattering has been ignored as the channel is assumed to be undoped.

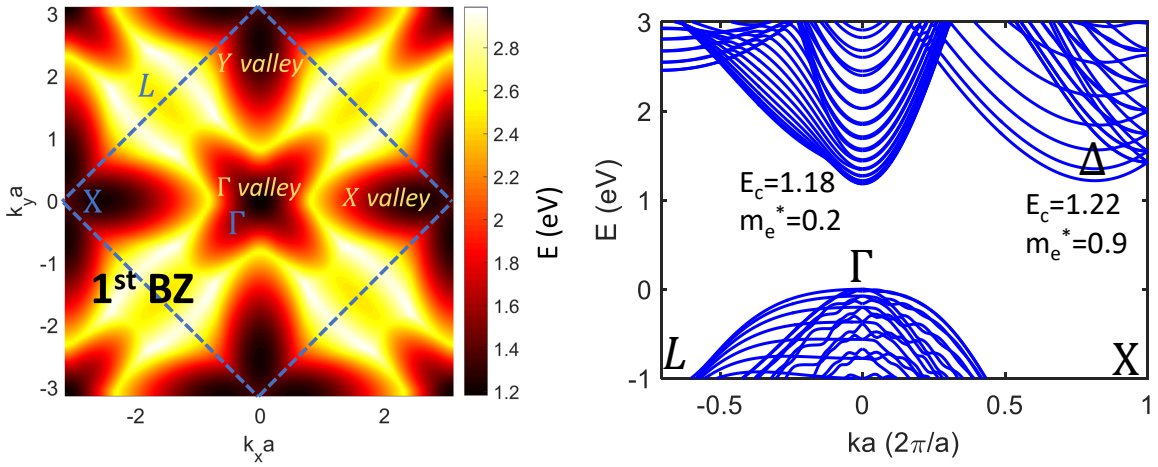
Bandstructure

For calculating the band structure in Si channel, we use a $sp^3d^5s^*$-SO basis set using the tight-binding parameters from Ref. [58]. Passivation of the surface dangling bonds is assumed. The channel is confined in the z direction. For (001) Si , the bulk Δ valleys along k_z in three-dimensional (3D) Brillouin Zone are projected at Γ point (Fig. 5.2). For other surface orientations, conduction band edges are not projected onto the Γ point because there are nonzero components of the in-plane wavevectors. Fig. 5.3a illustrates the lowest subband energy of 6-nm-thick Si thin film in the two-dimensional (2D) reciprocal space. The L point is the projection of 3D bulk L point. The dotted square represents the 2D Brillouin Zone. The X and Y valleys have a small confinement mass and a large transport mass. On the other hand, Γ valley has a large confinement mass and a small transport mass. Consequently, conduction band edges at the Γ point are less up-shifted by confinement effects than other



(a) (b)

Figure 5.2: The direct bandgap transition from (a) 3D bulk silicon to (b) (001) UTB Si. The Δ valleys on z-axis (Z valley) are projected on to the origin of the first BZ (Γ valley).



(a) (b)

Figure 5.3: (a) Lowest subband of Si (100) thin film for approximately T =6 nm within the 2D Brillouin zone (b) Band structure of 6 nm (100) thin film using the $sp^3d^5s^*$ tight-binding model.

	Valley	E_g	m_x
1st	Γ	1.183	0.201
2nd	Γ	1.216	0.202
3rd	Γ	1.269	0.203
4th	Γ	1.344	0.205
5th	Γ	1.439	0.208
6th	X	1.218	0.895
7th	X	1.352	0.913
8th	Y	1.218	0.224
9th	Y	1.352	0.325
10th	Y	1.427	0.564

Table 5.1: Band gap (E_g) and effective masses of transport (m_x) of each subbands using $sp^3d^5s^*$ TB model.

valleys since the conduction band edges at the Γ point have a higher confinement effective mass [110]. As a consequence, Si (001) thin films have a direct band gap (Fig. 5.3b).

Number of subbands

For current voltage calculation, one must also determine the number of subbands that should be included in the calculation. Here, we have included all subbands within 0.3 eV of the conduction band minimum. As we show later, subbands outside this range contribute negligibly to the total current. Within this energy range, here are 10 modes in total: 5 for Γ valley, 2 for X valley, and 3 for Y valley. X valleys and Y valleys have different effective masses (m_e^* at X valley ≈ 0.90 m_0, m_e^* at Y valley ≈ 0.22 m_0) in the transport direction and should be included separately. The effective masses of transport (m_x^*) and band gap (E_g) of each subbands, calculated from $sp^3d^5s^*$ Hamiltonian, are listed in Table. 5.1. The effective mass is calculated by fitting a parabola dispersion at the minimum of the conduction band in each valley. Furthermore, the probability densities for 10 subbands of 6 nm Si can be obtained from the eigenfunction of TB calculations, as shown in Fig. 5.4. Those probability densities $|\psi_\alpha(x,z)|^2$ are used to redistribute the electron density to x-z plane.

In Fig. 5.5, we compare the density of states of 6 nm Si UTB calculated using the mode-space NEGF and the real-space NEGF approach. The density of states can be obtained by integrating the spectral function, with respect to k_y under the flat band condition. Mode-

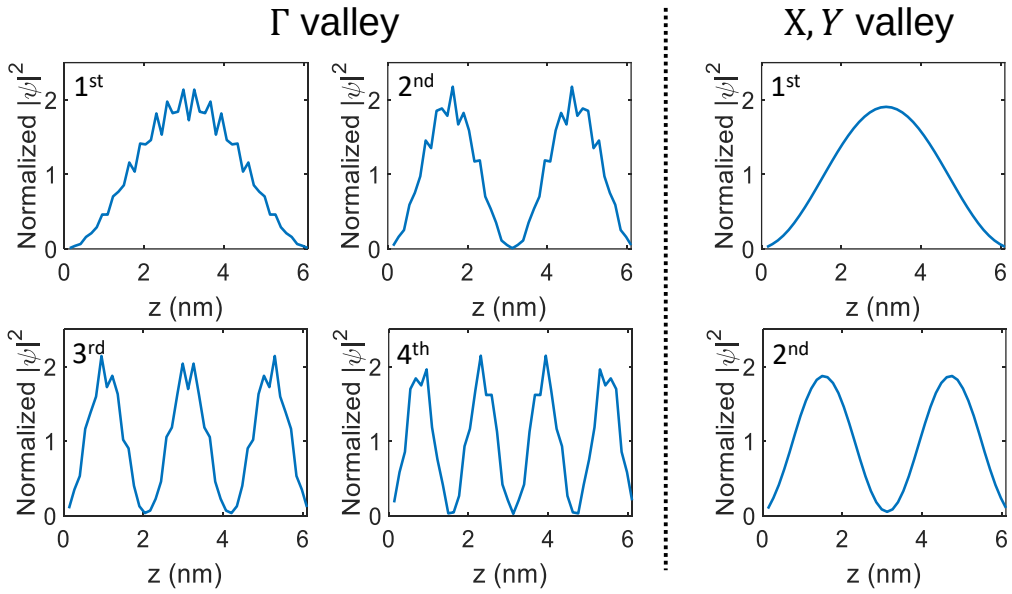


Figure 5.4: Normalized probability density of subbands at each valley for 6 nm Si. The probability density at Y valley. The eigenfunctions of Y valley have the same distribution as of X valley. The fifth subband at Γ valley and the third subband of Y valley are not shown here.

space using 10 subbands is in very close agreement with the real-space NEGF results within 0.3 eV from the conduction band edge. We note that the initial very small step near the conduction band edge represents the 1st subband of Γ valley and the second step is the combined density of states from the 2nd subband of Γ valley and the 1st subbands of X and Y valleys.

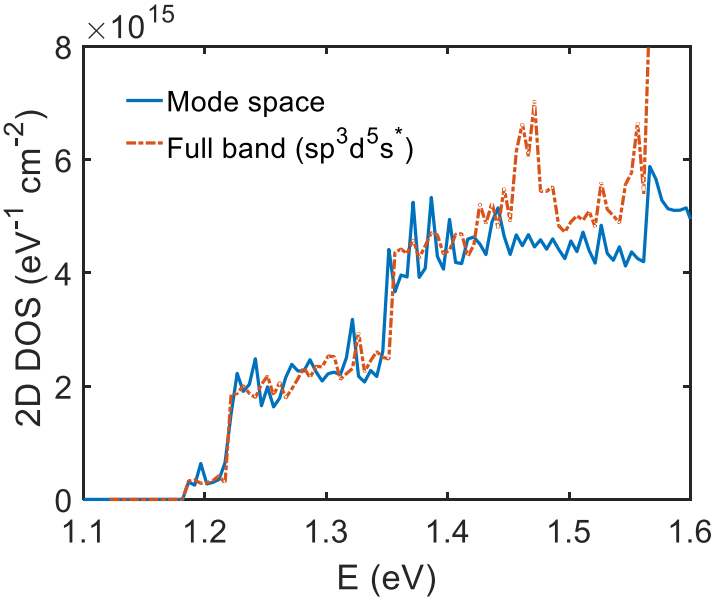


Figure 5.5: The density of states for 6 nm Si UTB using mode-space NEGF and full-band NEGF approach.

Ballistic transport

We start with a ballistic calculation to understand two aspects: (i)how impactful is the direct source drain tunneling in the context of the total current and (ii) how does the current distribute itself among different subbands. This second aspect will help understand the effect of coupling between the modes when we bring in the scattering in the next section. Fig. 5.6 shows the $I_{DS} - V_{GS}$ characteristics for different gate lengths under ballistic transport for V_{DS} of 0.01 V and 0.5 V. Significant short channel effects for small channel lengths are clearly evident from the current-voltage characteristic. The 10 nm length MOSFET shows a deteriorated subthreshold swing (SS) and a drain-induced barrier lowering (DIBL) in comparison with longer gate length MOSFET.

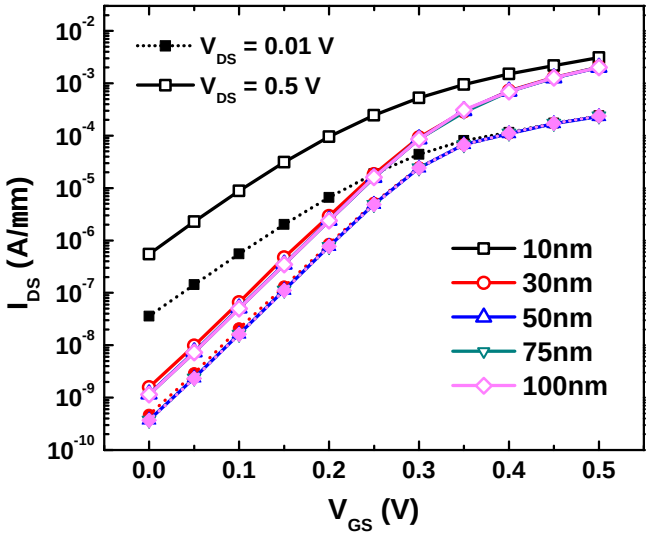


Figure 5.6: Ballistic Transfer curves (I_{DS} vs. V_{GS}) for double-gated Si MOSFETs with different gate lengths at V_{DS} = 0.01 (dashed lines) and 0.5 V(solid lines).

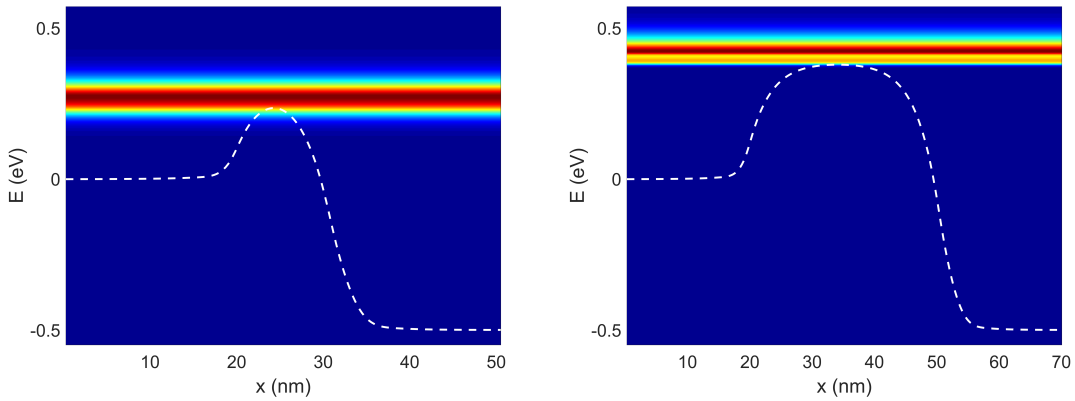


Figure 5.7: Energy-resolved current spectrum of ballistic transport at V_{GS} = 0 V and V_{DS} = 0.5 V for (a) L_G = 10 nm and (b) L_G = 30 nm.

Fig. 5.7 shows the energy-resolved current spectra at V_{GS} = 0.0 V and V_{DS} = 0.5 V for ballistic simulation of L_G = 10 nm and L_G = 30 nm MOSFETs. White dashed line shows profile of conduction band minimum. Fig. 5.8a shows the off-state (V_{GS} = 0 V) band diagram of first subband for different gate lengths at low drain (V_{DS} = 0.01 V) and high drain bias (V_{DS} = 0.5 V). The potential energy barrier for electrons in the channel is lowered both by

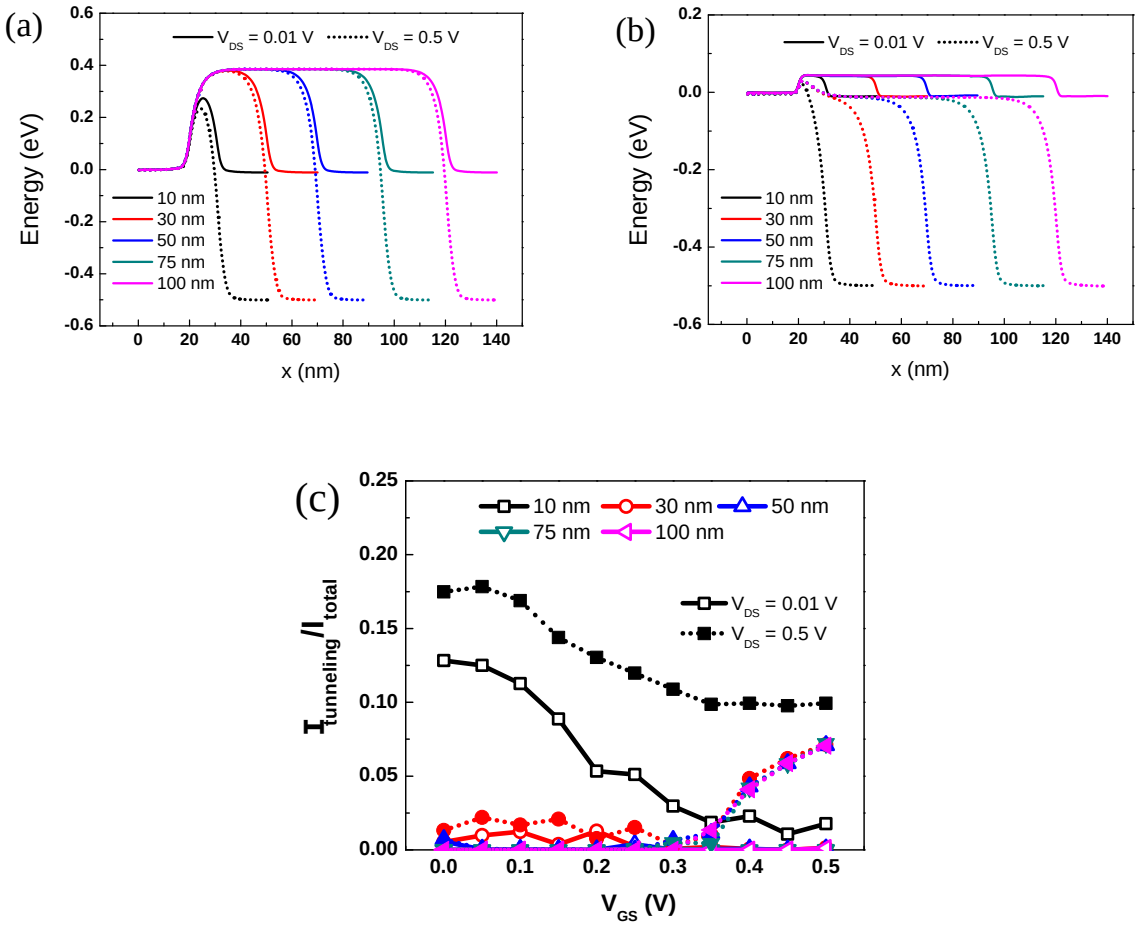
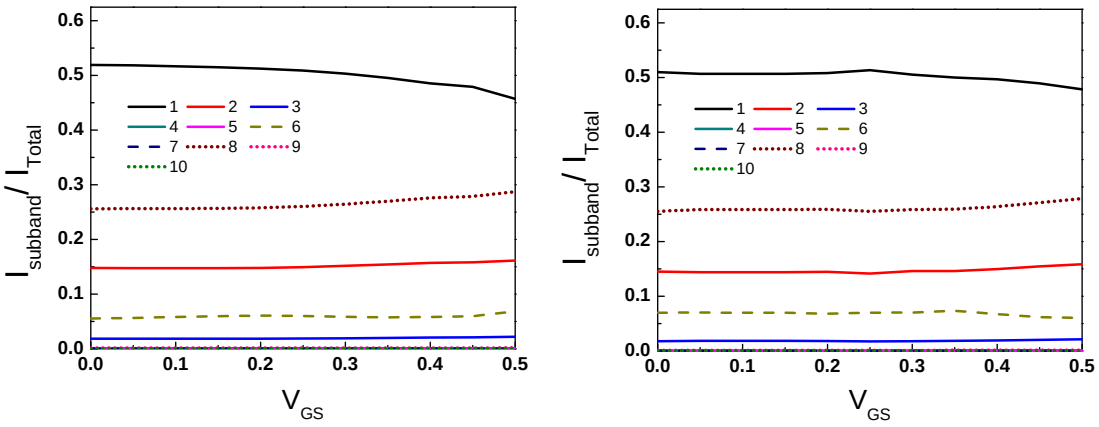


Figure 5.8: (a) First subband E_C profile for different gate lengths at V_{GS} = 0 V. (b) First subband E_C profile for different gate lengths at V_{GS} = 0.45 V. (c) Percentage of BTBT current for different gate lengths as a function of V_{GS}.

decreasing gate length (V_{th} roll-off) and increasing drain voltage (DIBL). Both phenomena stem from the relatively increased charge-sharing effect between the channel depletion region and source/drain depletion regions when compared to the long-channel device case. This charge-sharing effect makes a transistor require less gate voltage to deplete and makes V_{th} decrease. Furthermore, the source-drain tunneling (SDT) current for off-state also increases with gate length reduction. Fig.5.8c shows the percentage tunneling for different gate lengths as function of gate voltage at high drain and low drain bias. The off-state percentage tunneling of short channel device at high drain bias is higher than the ones at low drain bias. We note that the SDT current is also an important source of current for the On-state and is relatively less sensitive to channel length due to the narrowed top-of-barrier (TOB) width. The off-state SDT current for L_G = 10 nm and L_G = 30 nm are 17.5 % and 1.3 % of

the total current, respectively. The On-state SDT current for L_G = 10 nm and L_G = 30 nm are 9.9 % and 7.1 % of the total current, respectively. Fig. 5.9 (a) and (b) show Ratio of 10 subbands current to the total current (I_{sub}/I_{Total}) as function of V_{GS} at V_{DS} = 0.5 V for L_G = 10 nm and L_G = 100 nm, respectively. For ballistic simulations, the percentage current for the lowest first subband is about 51 %. The percentage current for the first subband of Y valley (8th subbands), the second subband of Γ valley (2nd subbands) and the first subband of X valley (6th subbands) are 25.5 %, 14.5 % and 7.0 %, respectively. The three subbands have almost the same E_g but have different percentage currents due to the different m_x and m_y. Subband with small m_x and large m_y carry lager currents.

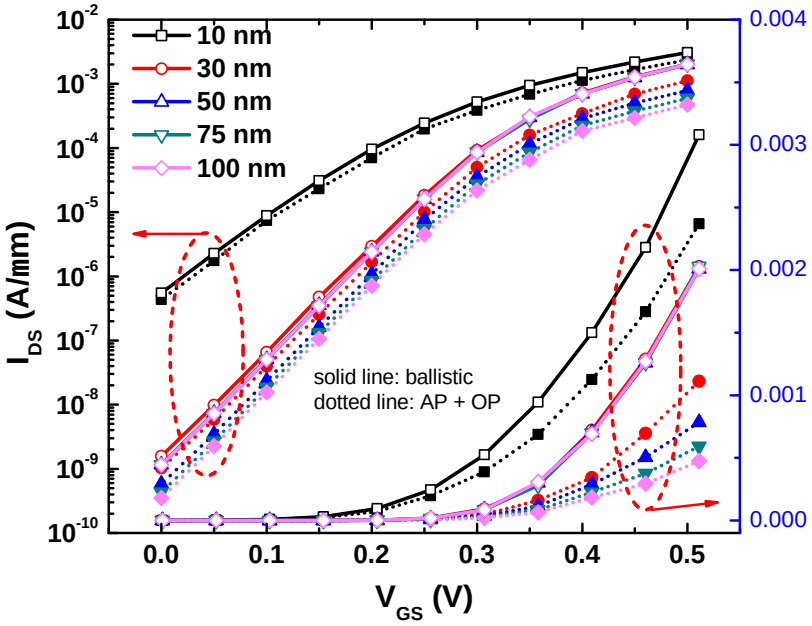


(a) (b)

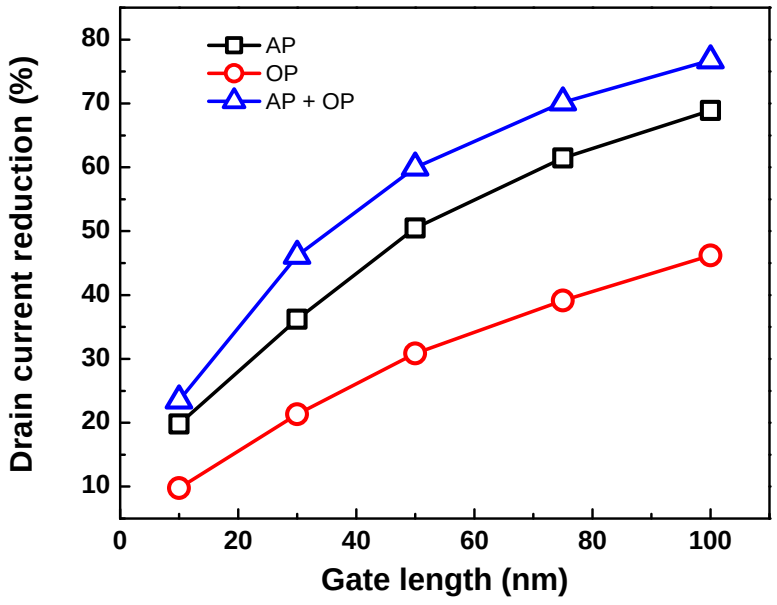
Figure 5.9: Ratio of 10 subbands current to the total current ($I_{subband}/I_{Total}$) as function of V_{GS} at V_{DS} = 0.5 V for (a) L_G = 10 nm and (b) L_G = 100 nm.

Electron-Phonon Scattering

We now include the electron phonon scattering as discussed in Section 5.2. For this, the values for the deformation potential for acoustic phonons (D_A) and optical phonons (D_o) are fitted to match the phonon-limited Si effective mobility ($\mu_{eff} \approx 1000 cm^2/V \cdot s$) [111]. These values are 3.05 eV for acoustic phonons and of 2.56×10^8 eV/cm for optical phonons. In addition, $\rho = 2329 kg/m^3$, $v_s = 9040 m/s$ and $a_x = a_y = 0.2715$ nm are used. Fig.5.10a shows the $I_{DS} - V_{GS}$ characteristics in the saturation region (V_{DS} = 0.5 V) for both ballistic and dissipative conditions. Fig. 5.10b shows percentage of reduction in the drain current($(I_{ball} - I_{scatt})/I_{ball} \times 100$) due to various scattering mechanisms as a function of the gate length at V_{GS} = 0.45 V. It is apparent that as the channel length increases, the enhanced probability of scattering decreases the current, because a portion of scattered carriers returns to the source and cannot reach the drain. AP scattering is found to be twice as effective as OP. Interestingly, the effect of phonon scattering is a function of both V_{GS} and V_{DS}. Fig. 5.11



(a)



(b)

Figure 5.10: (a) I_{DS} – V_{GS} characteristics at V_{DS} = 0.5 V with different gate lengths in logarithmic (left axis) and linear scales (right axis). Sold lines are from ballistic simulations and dashed lines are from dissipative simulation (AP + OP). (b) Percentage reduction in the on-current caused by scattering as a function of gate length at high drain bias, V_{DS} = 0.5 V.

shows the same percentage reduction in the drain current as in Fig. 5.10b but as a function of V_{GS} for the L_G = 50 nm device at V_{DS} = 0.01 V and V_{DS} = 0.5 V. The impact of phonon scattering increases as the gate bias increases. This can be explained by the increase in the density of states with respect to increasing energy from the conduction band edge (E_c) . At a low V_{GS}, the current flows mainly near the conduction band edge where the carriers encounter a smaller density of states and hence less scattering. In contrast, at a high V_{GS}, a considerable fraction of carriers flow well above the E_c where density of states is higher than that at the conduction band edge and the scattering rate increases (Fig. 5.5). We also find that OP scattering is more effective at high drain voltages. This is expected as the carriers need a threshold energy to emit an optical phonon. Fig.5.12 shows the energy-resolved

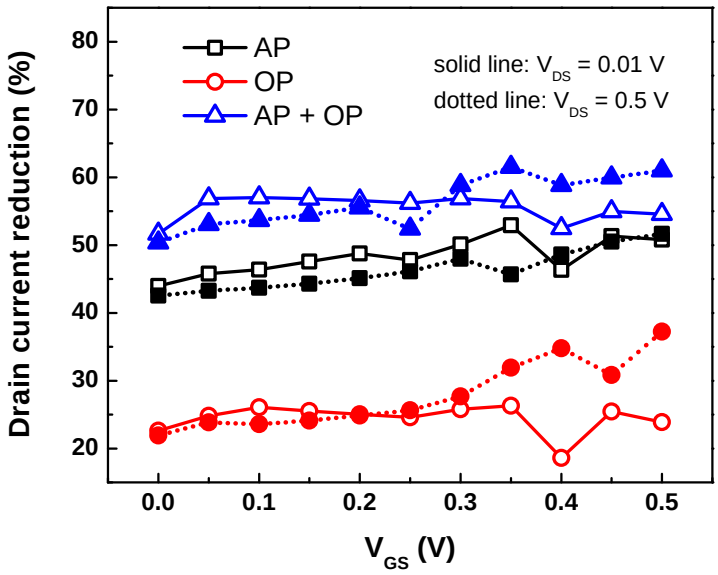


Figure 5.11: Percentage reduction in the on-current caused by scattering as a function of gate length at V_{DS} = 0.01 V and V_{DS} = 0.5 V.

current spectrum at the OFF condition (V_{GS} = 0 V and V_{DS} = 0.5 V) for the L_G = 50 nm for ballistic as well as dissipative conditions. The coupling between various energies for the OP phonon scattering and therefore the spreading of the electrons to various energies can be clearly seen. To better understand this, we have plotted the current spectrum at the drain end in Fig.5.13a. The spectrum exhibits two peaks with an energy spacing of approximately 0.3 eV. The first peak (P_1) is the current from the first subband of the Γ valley and the second peak (P_2) is the current from the sum of the first subbands of the X valley and Y valley and the first and second subbands of the Γ valley. In ballistic transport, the second peak is higher because of the larger density of states, even though it is further away from the conduction band edge. Under dissipative transport, however, this is not the case. With scattering the height of P_1 increases and that of P_2 decreases. This is result of the different

subtends and k_y modes getting coupled that distributes carriers more evenly among various energy levels. A different way to understand this effect is to look at the proportion of total current carried by the first subband. This is shown in Fig. 5.13b. At the drain where significant space opens up in the lower energy ranges, the electrons spread out reducing the ratio of the current carried by the first subband.

We have performed the same study at the ON condition ($V_{GS} = 0.45\ V$ and $V_{DS} = 0.5\ V$) as shown in Figures 5.14 and 5.13. As the drain voltage now slants the barrier profile from source to drain, energy states become available even in the middle of the device and current spreading and therefore reduction in the current carried by the subband can be seen far away from the drain.

Fig. 5.16a shows the ratio of first subband current to the total current (I_{1stsub}/I_{Total}) as function of V_{GS} for ballistic (dashed lines) and scattering (solid lines) simulations for $V_{DS} = 0.5\ V$. This value of the ratio is obtained by averaging the values inside the entire device. In the case of ballistic transport, the ratio of first subband current continues to decrease as V_{GS} increases, whereas in the case of dissipative transport, the ratio increases in the sub-threshold region and decreases in the above-threshold region as V_{GS} increases. The ratio of first subband at top of the barrier is almost independent of V_{GS}. However, in the sub-threshold region, the ratio in the source/drain region increases as V_{GS} increases, and in the above-threshold region, the ratio decreases because of the tilted potential in the channel region. The increase in the ratio with increasing channel length in the sub-threshold region is due to the relatively low contribution of the source/drain regions. On the other hand, in the above-threshold region, the length of the channel with slanted potential increases with increasing L_G, thus resulting in a further reduction of the fraction of the first subband current. These results support the interpretation that as V_{GS} increases, the fraction of carriers flowing well above the E_c increases, resulting in higher scattering rate due to higher density of states (Fig. 5.11).

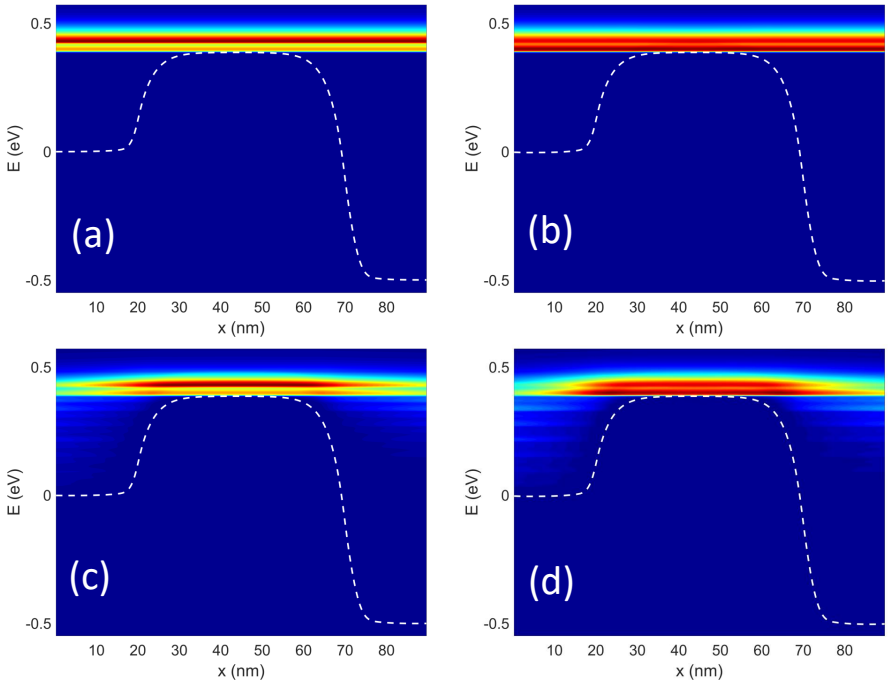
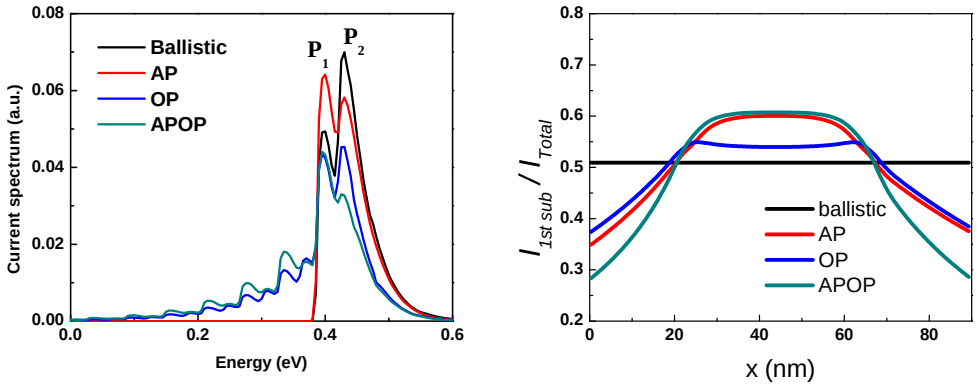


Figure 5.12: Energy-resolved current spectrum at $V_{GS} = 0\ V$ and $V_{DS} = 0.5\ V$ under (a) ballistic transport and dissipative transport with (b) AP, (c)OP, and (d) AP + OP.



(a) (b)

Figure 5.13: (a) Energy-resolved current with different phonon scattering mechanisms. Values are obtained at the drain end at $V_{GS} = 0\ V$ and $V_{DS} = 0.5\ V$. (b) Ratio of first subband current to the total current (I_{1stsub}/I_{Total}) as function of position.

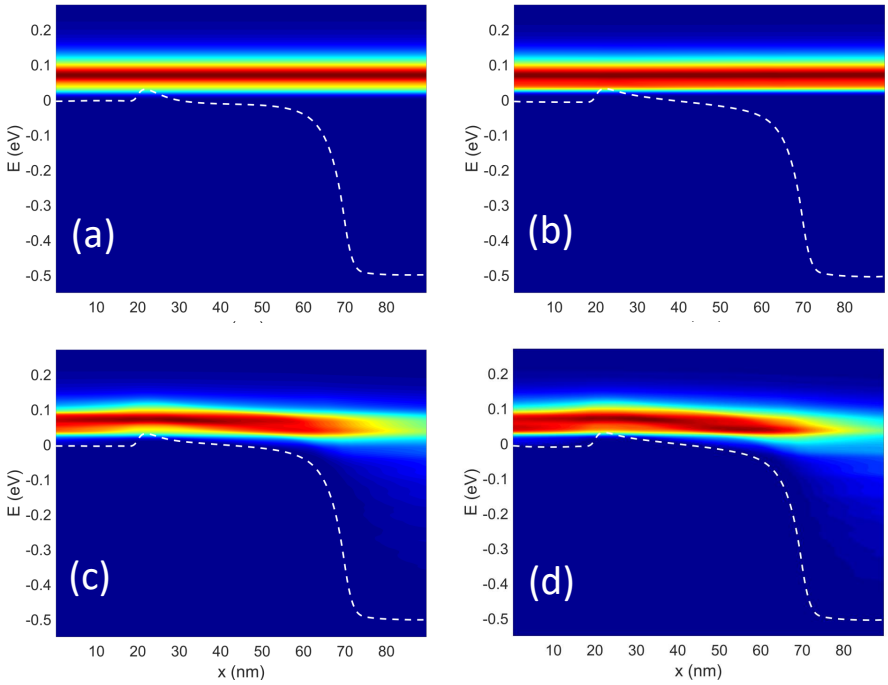
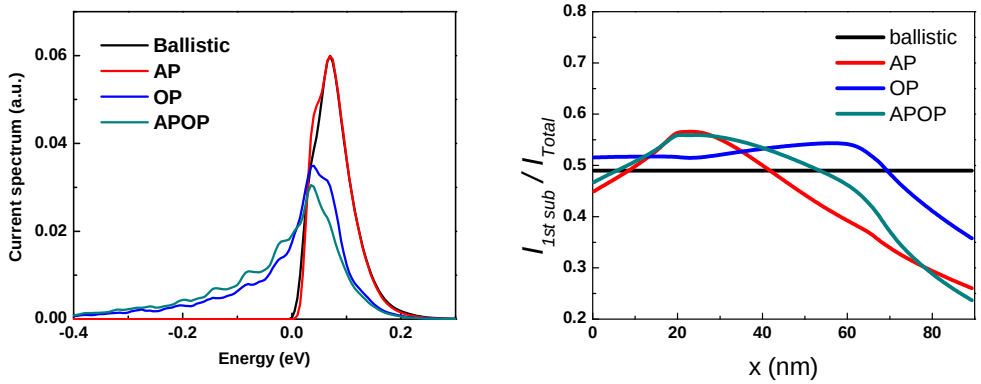
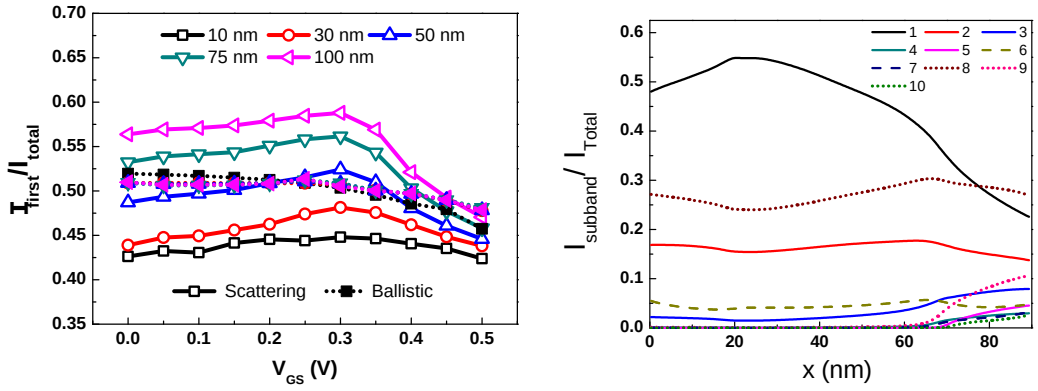


Figure 5.14: Energy-resolved current spectrum at V_{GS} = 0.45 V and V_{DS} = 0.5V under (a) ballistic transport and dissipative transport with (b) AP, (c)OP, and (d) AP + OP.



(a) (b)

Figure 5.15: (a) Energy-resolved current with different phonon scattering mechanism. Values are obtained at the drain end at V_{GS} = 0.45 V and V_{DS} = 0.5 V. (b) Ratio of first subband current to the total current (I_{1stsub}/I_{Total}) as function of position



(a) (b)

Figure 5.16: (a) Ratio of first subband current to the total current (I_{1stsub}/I_{Total}) as function of V_{GS} for ballistic (dashed lines) and scattering (solid lines) simulations at $V_{DS} = 0.5\ V$ (b) Ratio of 10 subbands current to the total current (I_{1stsub}/I_{Total}) as function of position for a $L_G = 50\ nm$ device at $V_{GS} = 0.5\ V$ and $V_{DS} = 0.5\ V$

Phonon-limited mobility

Finally, we are at a position to investigate the phonon limited mobility. For this, we follow the approach prescribed in [98–102]. Conventionally, the drain current at low electric fields can be expressed as

$$I_{DS} = \frac{W}{L_G}\mu Q_i(0)V_{DS} \tag{5.7}$$

where $Q_i(0)$ is the sheet electron density at the top of the barrier, μ is the channel mobility, V_{DS} is the drain voltage and W and L_G are the device width and length, respectively.

On the other hand, current under ballistic conditions can be written as [94, 112]

$$I_{DS} = W\left[\frac{v_T q}{k_B T}\frac{\mathscr{F}_{-1/2}(\eta)}{\mathscr{F}_0(\eta)}\right]Q_i(0)V_{DS} \tag{5.8}$$

where $v_T = \sqrt{2k_BT/\pi m^*}$ is the unidirectional thermal velocity of nondegenerate electrons, the function $\mathscr{F}_i$ is the Fermi-Dirac integral and $\eta = (E_F - E_{C0})/k_BT$, where E_F is the source Fermi level and E_{C0} is the first subband level for electrons. To relate the channel length independent ballistic current to an expression similar to conventional devices (5.7), Shur [95] defined a quantity called ballistic mobility μ_B as

$$\mu_B = \frac{v_T q L_G}{2k_B T}\frac{\mathscr{F}_{-1/2}(\eta)}{\mathscr{F}_0(\eta)} \tag{5.9}$$

In the quasi-ballistic regime, where the device length is of the order of the mean free path (mfp) of electron phonon scattering, the μ in (5.7) is replaced by the apparent channel

mobility (μ_{app}) [98],

$$\frac{1}{\mu_{app}} = \frac{1}{\mu_B} + \frac{1}{\mu_{eff}} \tag{5.10}$$

where μ_{eff} is the effective mobility of the quasi-ballistic transistor.

Now, at small drain voltage, the total resistance of the device, $R_{total} = I_{DS}/V_{DS}$, can be written as

$$R_{total} = R_{ch} + R_{S/D} + R_B \tag{5.11}$$

where R_{ch} is the resistance of the channel and a function of channel length. On the other hand, $R_{S/D}$ is the series resistance of the source/drain region and R_B is the ballistic resistance. Neither $R_{S/D}$ nor R_B is a function of the channel length. From (5.11), one can write,

$$R_{total} = R_{S/D} + \frac{1}{\mu_{app}} \frac{L_G}{Q_i(0)} = R_{S/D} + \left(\frac{1}{\mu_B} + \frac{1}{\mu_{eff}} \right) \frac{L_G}{Q_i(0)} \tag{5.12}$$

Inserting (5.9) into (5.12), we obtain

$$R_{total} = R_{S/D} + \frac{2k_B T}{v_T q Q_i(0)} \frac{\mathscr{F}_0(\eta)}{\mathscr{F}_{-1/2}(\eta)} + \frac{1}{\mu_{eff}} \frac{L_G}{Q_i(0)} \tag{5.13}$$

Therefore, the effective mobility can be obtained from the slope of the R_{total} vs. L_G plot.

The carrier mean free path (λ_{mfp}) can be obtained from the scattering-limited mobility equation, which is generally given as [113]

$$\mu_{eff} = \frac{v_T q \lambda_{mfp}}{2k_B T} \frac{\mathscr{F}_{-1/2}(\eta)}{\mathscr{F}_0(\eta)} = \lambda_{mfp} \mu_B \tag{5.14}$$

We calculate $I_{DS} - V_{GS}$ characteristics in the linear region ($V_{DS} = 0.01\ V$) under ballistic and dissipative transport with AP and OP scattering (Fig. 5.17). The $I_{DS} - V_{GS}$ curves have been plotted on both linear and logarithmic scale to compare the above-threshold and the sub-threshold regions. Under the ballistic transport, currents at the same gate voltage have the same value regardless of the length, except for the increase of current in short channel devices in the sub-threshold region owing to short channel effect. From I_{DS} and $Q_i(0)$ for ballistic and dissipative transport, the ballistic mobility and the apparent mobility are obtained as shown in Fig. 5.18. Ballistic mobility increases linearly with device length as shown in (5.9), while the apparent mobility is similar to the ballistic mobility for a length of 10 nm, but its value saturates as the length increases, because the apparent mobility is governed by the Matthiessen's rule (5.10).

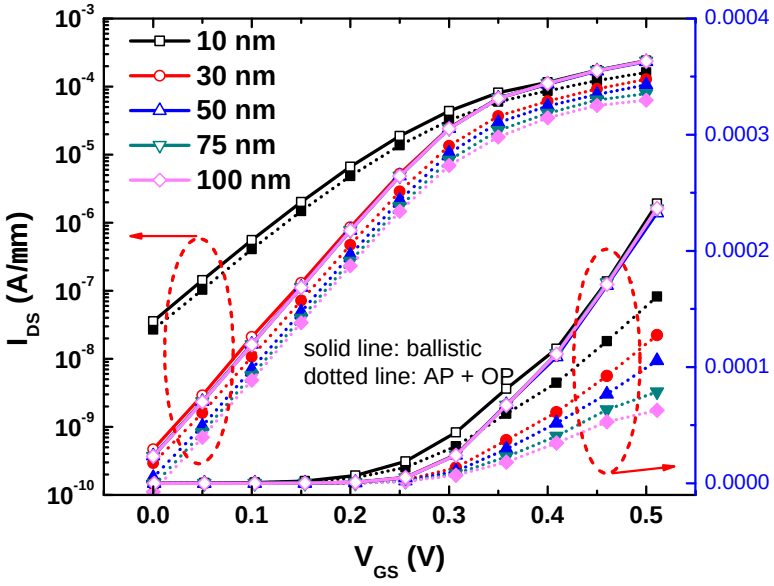
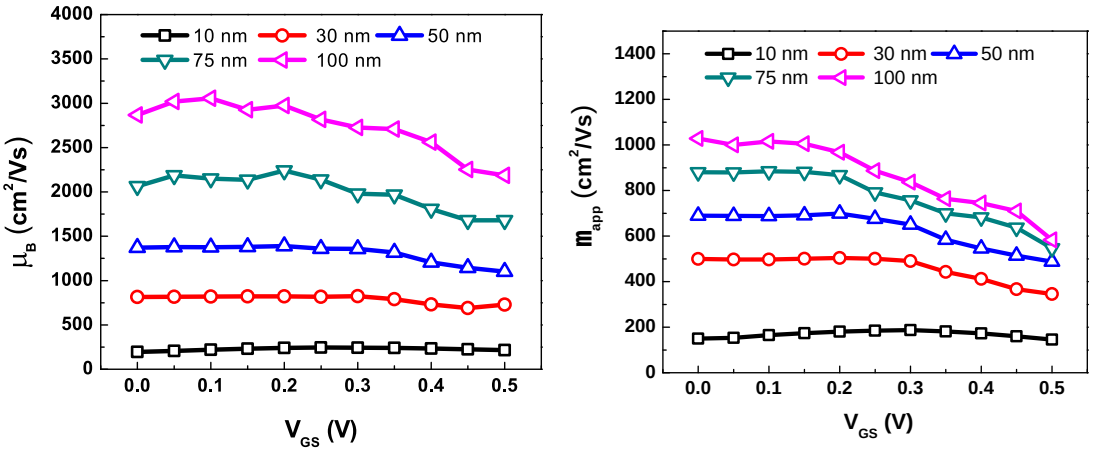
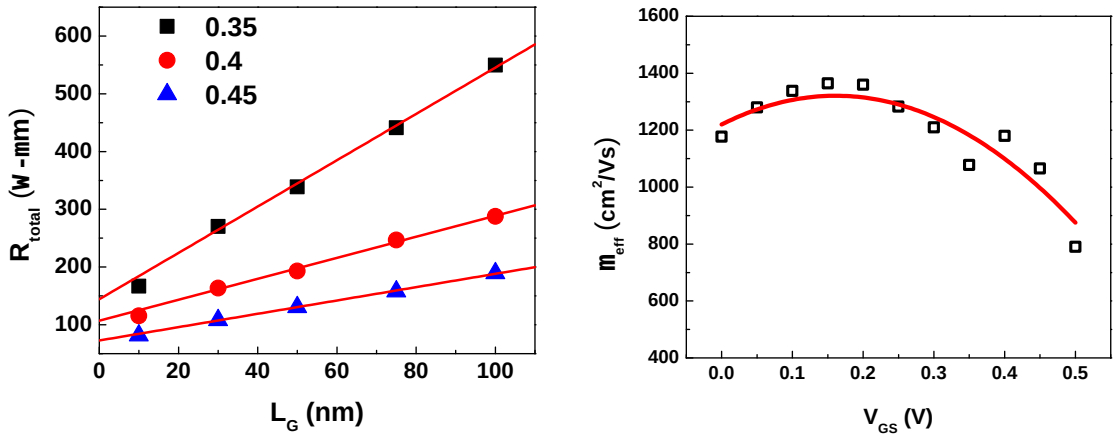


Figure 5.17: Ballistic and dissipative ID-VG transfer curves (I_{DS} vs. V_{GS}) for double-gated Si MOSFETs with different gate lengths at V_{DS} = 0.01.



(a) (b)

Figure 5.18: (a) Ballistic mobility and (b) Apparent mobility as a function of gate voltage (V_{GS}) for different gate lengths.



(a) (b)

Figure 5.19: (a) R_{total} versus L_G for different V_{GS} of a 6 nm Si MOSFETs with AP and OP scattering. The slope of the linear curve is proportional to the inverse of μ_{eff} and the y-axis intersect is $R_{total}(0) = R_{S/D} + R_B$. (b) Effective mobility as a function of gate voltage (V_{GS}).

Fig. 5.19a shows dependence of R_{total} versus L_G for different V_{GS}. The R_{total} value from 30 nm to 100 nm is well fitted by a straight line, but the R_{total} value of 10 nm is smaller than the linear fitting and the V_{GS} are smaller and the deviation is larger. This is because the potential barrier in the short channel ($L_G = 10\ nm$) is lowered by the drain bias (DIBL) in the sub-threshold region, resulting in a R_{total} value less than for the long channel. The extracted results are summarized in Table. 5.2. For low carrier concentration (low V_{GS}), λ_{mfp} is 48 nm and μ_{eff} is 1300 $cm^2/V \cdot s$. As V_{GS} increases, μ_{eff} and λ_{mfp} decrease together. When $Q_i(0)$ is $6.8 \times 10^{12}\ cm^{-2}$ at $V_{GS} = 0.5\ V$, λ_{mfp} is down to 36 nm and μ_{eff} is 790 $cm^2/V \cdot s$. Notably μ_{eff} decreases at high gate voltage (Fig. 5.19b). This is due to increased scattering with increasing density of states as we discussed before.

Figures 5.19a and 5.19b provide the connection to experimental measurement. In principle, a similar R vs. L_G plot can be drawn from experimental data and value for μ_{eff} can be extracted. The simulation results can then be compared against the experimental measurements by using the acoustic and optical phonon scattering coupling constants as fitting parameters much the same way it was done for Monte Carlo simulations of long channel MOSFETs. Once this is done, predictive simulation can be performed for shorter channel devices where the $R - L_G$ may vary from a linear trend due to DIBL (as discussed in the preceding paragraph) and where source-drain direct tunneling becomes a significant part of the total current (Fig. 5.8c).

V_{GS} (V)	Slope (Ω)	$Q_i(0)$ $(\#/cm^2)$	μ_{eff} $(cm^2/V{\cdot}s)$	R_B $(\Omega{\cdot}\mu m\ )$	$R_{S/D}$ $(\Omega{\cdot}\mu m\ )$	λ_{mfp} (nm)
0.35	3672	1.58E+12	1078	167.0	30	41
0.4	1785	2.96E+12	1180	91.5	15	49
0.45	1280	4.66E+12	1065	58.0	11	47
0.5	1172	6.76E+12	787	42.7	8	36

Table 5.2: Extracted characteristics of phonon scattering for 6nm Si MOSFETs with different V_{GS}.

5.4 Conclusion

To summarize, we have studied the effects of electron-phonon scattering in double Si UTB MOSFETs by using a mode space approach within the NEGF formalism. Due to coupling between subbands and transverse modes that open up significant phase space at larger gate voltages, the effective mobility of electrons in an undoped Si channel could be reduced from a high of 1400 $cm^2/V \cdot s$ to almost half of that value at high voltages. We presented a framework that can be used to directly compare experimentally obtained effective mobility with simulation results. Such comparison at channel lengths such as 50-100 nm, where direct source-to-drain tunneling is not important can be used to calibrate the model parameters. This calibrated model can then be used to predict device behavior for ultra scaled MOSFETs where both tunneling and short channel effects become important. For a complete description, one would also need to include surface roughness scattering. However, since this is an elastic scattering process, this can be added in the same was as acoustic phonons [80] and therefore the general framework presented here can still be applied.

Chapter 6

Impact of indium mole fraction on the quantum transport of ultra-scaled InGaAs MOSFET

6.1 Introduction

InGaAs has been actively researched as an n-channel material for future highly scaled metal-oxide-semiconductor field-effect transistors (MOSFETs) due to its high mobility. The high mobility (or high injection velocity, v_{inj}) could improve electrical performance characteristics, such as on-current and delay. Since the high injection velocity comes from the light carrier effective mass m^*, high indium content devices continue to attract attention because of their low m^*. However, there are two main physical issues with high indium content channel materials. The first one is the so-called density-of-states (DOS) bottleneck problem which limits the on-current of III-V MOSFETs with thin oxide thickness. For MOS capacitors, the total gate capacitance (C_G) is the series combination of the oxide capacitance and the quantum capacitance of channel is

$$\frac{1}{C_G} = \frac{1}{C_{OX}} + \frac{1}{C_Q} \tag{6.1}$$

and is dominated by the smaller of two. The quantum capacitance, (C_Q), is a function of density of states. The small DOS coming from the light effect mass degrades C_G and therfore on-current for thin oxides. The second is the increase of off-state leakage current due to the effective mass and small band-gap. A light effective mass increases leakage current through direct source-to-drain tunneling and body-to-drain tunneling. On the other hand, N_{SD} also has fundamental effects on the carrier transport within the device. At the DOS bottleneck point, since quantum capacitance is proportional to the density of states around the electrochemical potential μ, high N_{SD} increases the C_Q by shifting the electrochemical potential level from conduction band edge. However, high N_{SD} also has drawbacks because

it pushes down the S/D potential and makes the off-state potential barrier narrower. So, as N_{SD} increases, significant current can flow under the top of the barrier (source-to-drain tunneling). In this paper, we theoretically calculate the performance of ultra-scaled InGaAs n-MOSFETs using the Non-Equilibrium Greens Function (NEGF) formalism, which provides a rigorous framework for treating quantum transport. We explore the effects of N_{SD} on the on and off states for different $In_xGa_{1-x}As$ composition. We also perform benchmarking against silicon.

6.2 Simulation approach

Fig. 6.1 illustrates the device structure of a 2-D double gate (DG) ultra-thin body (UTB) MOSFET. The gate length is 10 nm and the length of source/drain regions is 20 nm. The 3 nm thick HfO_2 ($\epsilon_r = 20$) is used as the dielectric material (EOT = 0.585 nm).The channel is undoped. We vary the doping level of the source/drain from 1×10^{19} cm^{-3} to 4×10^{19} cm^{-3}. The thickness of InGaAs is fixed to 6 nm and the transport direction is [100]. For this work, strain effects are not considered, and we focus on the intrinsic material properties. To obtain

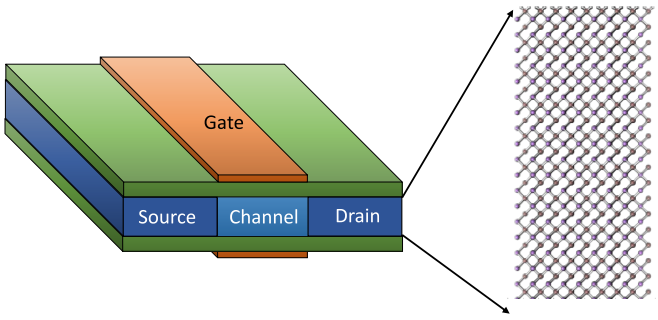


Figure 6.1: Schematic of DG $In_xGa_{1x}As$ UTB MOSFET with atomic arrangements of anion and cation layers

the current-voltage characteristics, mode-space NEGF within an effective mass approximation is employed. The effective mass of $In_xGa_{1-x}As$ UTB materials are interpolated from the $E-k$ dispersion which was obtained from the atomistic tight-binding (TB) model using the $sp^3d^5s^*$ basis set with spin-orbit (SO) coupling. SO coupling should be included for small band gap materials such as InAs to reproduce the proper band gap energy (E_g) which significantly affects the body-to-drain band-to-band tunneling (BTBT) leakage current. For example, when SO coupling is included, the band gap of the InAs is reduced by about 20 %. TB parameters for GaAs and InAs are extracted from Ref. [114] and the TB parameters are modified to include a bowing parameter for precise InGaAs alloy bandstructure [114]. Fig. 6.2 shows the band gap energy and effective mass for 6nm thick $In_xGa_{1-x}As$ UTB as function of In concentration x. Bulk values are also presented for comparison. The increasing indium mole fraction leads to a reduction in both the band gap energy and the electron

effective mass. We assume that the dimension of the devices is large along the width, and

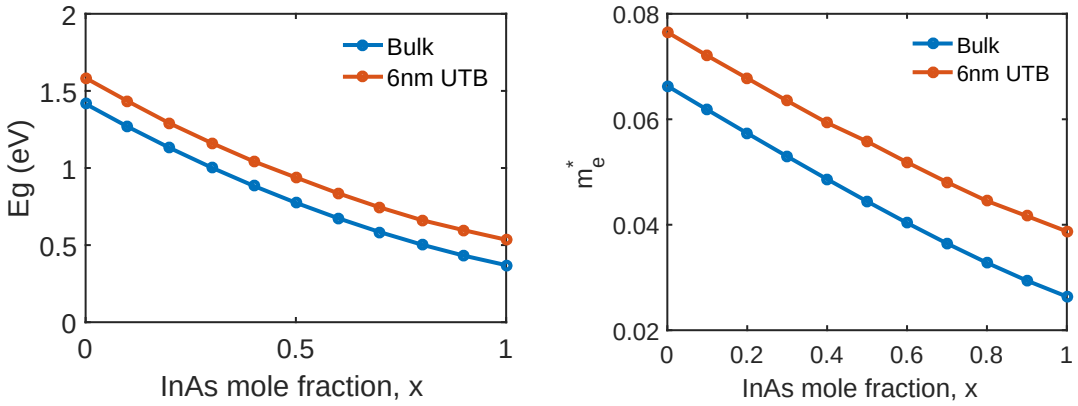
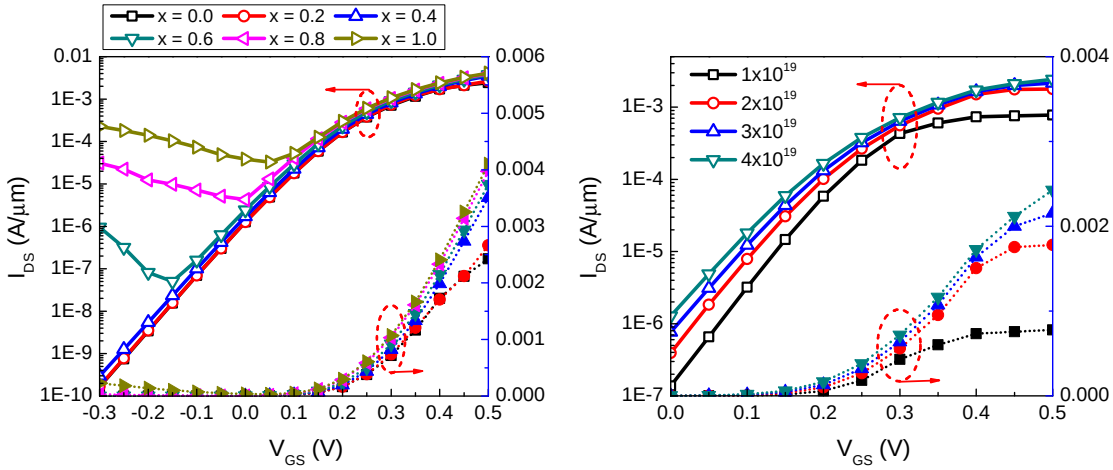


Figure 6.2: (a) Band gap energy (b) transport effective mass for bulk and the 6nm $In_xGa_{1-x}As$ UTB as function of indium mole fraction x

hence we use a periodic boundary condition in that direction. The transverse momentum modes are summed numerically in the calculation of charge density and current. The NEGF equations are solved iteratively along with Poisson's equation until self-consistency between charge and electrostatic potential is achieved. The microscopic electron scattering mechanisms have taken into account within the self-consistent Born approximation [79, 80]. For an InGaAs MOSFET, polar optical phonon (POP) and surface roughness (SR) scattering are considered. For SR scattering, we use the exponential autocorrelation function with the correlation length (Λ) and the r.m.s. roughness (Δ) chosen to fit the low-field mobility of InAs, which is experimentally extracted as approximately 5000 $cm^2/V \cdot s$ [113] ($\Lambda = 0.2\ nm$, $\Delta = 6\ nm$). For a Si MOSFET, acoustic phonon and optical phonon scattering are considered. The values for the deformation potential for acoustic phonons(D_A) and optical phonons(D_o) are fitted to match the phonon-limited Si effective mobility ($\mu_{eff} \approx 1000\ cm^2/V \cdot s$) which was calculated using a Monte Carlo method [111] ($D_A = 3.05\ eV$, $D_o = 2.56 \times 10^8\ eV/cm$). Dirichlet boundary conditions are imposed at the source/drain contacts. Note that the Neumann boundary condition is not suitable to observe the DOS bottleneck phenomenon because under this condition E_C-E_F at high V_G significantly changes in order to maintain the electrostatics as the channel charge density becomes comparable to that at the source. With the same potential profile from the self-consistent solution, the tunneling leakage current at high V_{DS} is obtained by using the 'mode-space + 2 terminal WKB with appropriate boundary conditions (M2WBC) approach. A detailed description of the M2WBC modeling is described in Chapter 3.

6.3 Results and discussion

Fig. 6.3 presents simulation results with scattering for $In_xGa_{1-x}As$ n-MOSFETs. We plot the transfer characteristics in both linear and logarithmic scales to explore the performance in both the off-state and the on-state. The $I_{DS}-V_{GS}$ characteristics for $In_xGa_{1-x}As$ UTB for high drain voltage (V_{DS} = 0.5 V) with $N_{SD} = 4\times10^{19}$ are shown in Fig. 6.3a. First, in the case of the off characteristic, increasing the indium mole fraction (x) increases the leakage current. When x is higher than 0.6, the channel valence band and the drain conduction band overlap at the drain junction so that gate-induced drain leakage (GIDL) occurs. In fact, $In_xGa_{1-x}As$ MOSFETs with an indium content greater than 0.8 cannot meet the off-state criteria of $0.1\mu A/\mu m$ due to GIDL. As the consequence, the on-off ratio of the devices is less than 10^3. Second, on-state current is enhanced due to the decrease in the effective mass as the indium content is increased. The values of the charge density at top-of-barrier (TOB) are about $6.5\times10^{12}\ cm^{-2}$ and the difference according to the composition is within 5 %. Since the DOS near the Fermi energies of the doping levels are similar, the influence of the on-current due to the DOS change according to the composition is not significant. Therefore, high indium content, having low effective mass, shows higher on-current due to high injection velocity (v_{inj}) and more tunneling current under the TOB. With POP and SR scattering, the on-state current for x = 0 is reduced by 10.3 % in terms of drain reduction percentage ($(I_{ball}-I_{scatt})/I_{ball}\times100$). The drain reduction percentage for x = 1 is increased up to 12.9 % due to higher deformation potential for POP scattering. Fig. 6.3b shows transfer characteristics for GaAs MOSFETs for different N_{SD} at V_{DS} = 0.5 V. In the on-

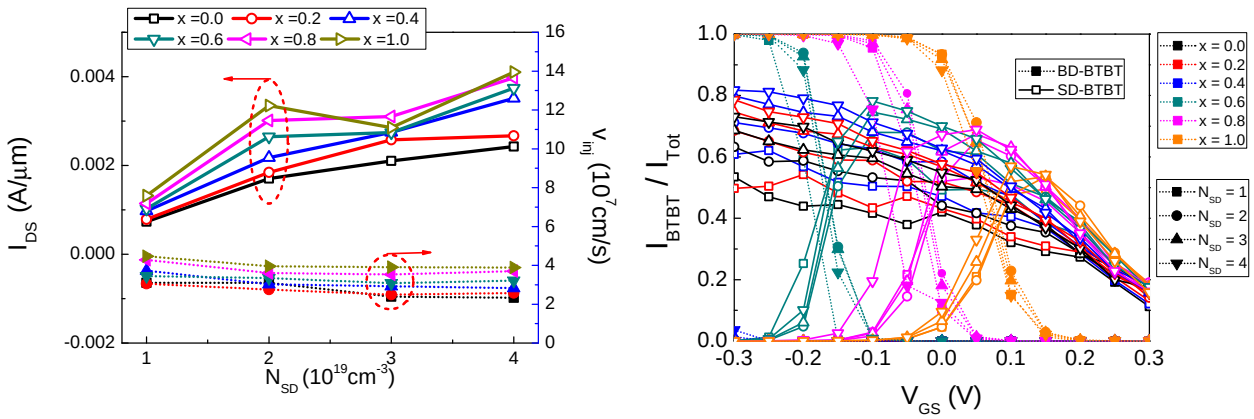


(a) (b)

Figure 6.3: (a) Simulation results for I_{DS} versus V_{GS} at V_{DS} = 0.5 V for $In_xGa_{1-x}As$ UTB MOSFETs. $N_{SD} = 4\times10^{19}$. (b) Simulation results for I_{DS} versus V_{GS} at V_{DS} = 0.5 V for GaAs UTB MOSFETs for various source/drain doping level.

state, low N_{SD} MOSFETs show degraded I_{DS}. This is well known as 'source exhaustion', where. the source with low doping level cannot supply enough electrons when the channel carrier density becomes comparable or higher than that of source [115, 116]. It causes high series resistance and transconductance degradation at high V_{GS}.

Fig. 6.4a shows I_{DS} and v_{inj} at $V_{GS} = 0.5\ V$ and $V_{DS} = 0.5\ V$ for various indium mole fractions and various N_{SD}'s. In general, the higher the indium content and the higher the N_{sd}, the greater the on-current. N_{sd} dependence of on-current is a consequence of the difference in charge density at the TOB owing to source exhaustion, and composition dependence of on-current is due to the difference in injection velocity (Fig. 6.4a). Whereas the injection velocity decreases as N_{SD} increases. This is because the unidirectional non-degenerate velocity of electrons, v_{T0} is reduced because of the impact of electron degeneracy as $v_T = v_{T0}\mathscr{F}_{-1/2}(\eta)/\mathscr{F}_0(\eta)$, where $v_{T0} = \sqrt{2k_BT/\pi m^*}$, the function $\mathscr{F}_i$ is the Fermi-Dirac integral and $\eta = (E_F - E_{C0})/k_BT$, where E_F is the source Fermi level and E_{C0} is the first subband level for electrons. On the other hand, however, increasing N_{SD} degrades the performance in the subthreshold region, giving leakage current and the deteriorated subthreshold swing (SS). For the GaAs MOSFET with $N_{SD} = 1 \times 10^{19}$, the SS is 72.3 mV/decade and increases to 79.8 for $N_{SD} = 4 \times 10^{19}$. This is mainly due to source-drain tunneling. To understand the impact of the change in effective mass and N_{SD} on leakage current, Fig.6.4b shows the the ratio of tunneling current to the total current (I_{BTBT}/I_{Total}) as a function of V_{GS} for various indium content fractions and N_{SD}'s. In this plot, we include two-types of BTBT mechanism, i.e., intra-band source-drain (SD) tunneling and inter-band body-drain



(a) (b)

Figure 6.4: (a) I_{DS} at $V_{GS} = 0.5\ V$ and $V_{DS} = 0.5\ V$ for various indium mole fractions and N_{SD}'s. (b) The ratio of tunneling current to the total current (I_{BTBT}/I_{Total}) as function of V_{GS} for various indium content and N_{SD}'s. Two-types of BTBT mechanism are compared. One is intra-band source-drain tunneling (SD-BTBT) and another is inter-band body-drain tunneling (BD-BTBT)

(BD) tunneling. As we explained before, BD BTBT is the dominant leakage source for high indium mole fractions. Also, the increase of indium content enhances the SD BTBT due to light effective mass. Furthermore, as N_{SD} is increased, total leakage current is increased by the enhanced SD BTBT rate. This is because the potential barrier is narrowed as high doping levels push down the potential at the source/drain region.

In order to determine the trade-off between on and off performance as a function of indium mole fraction and N_{SD}, we plot the on current as a function of On/Off ratio as shown in Fig. 6.5. The On/Off ratio is obtained by changing the bias window of 0.5 V for gate voltage along the transfer characteristics. This could be achieved in practice by engineering the workfunction of the gate metal [117]. The performance of 6nm thickness Si (100) DG MOSFETs is compared with that of $In_xGa_{1-x}As$ devices. $In_xGa_{1-x}As$ gives higher on-current than Si. This is in contrast to the prediction that III-V materials will show a lower current due to the DOS bottleneck. This can be explained by two causes. First, at the low V_{DD}, the on-current is mainly determined by the injection velocity and the charge density at the TOB is less affected by DOS. Second, scattering effects for Si are stronger than for $In_xGa_{1-x}As$. The drain reduction percentage of Si device for on-state is 23.1 % due to a higher electron-phonon scattering rate. We set our performance target as $I_{On} > 1000\ \mu A/\mu m$ and $I_{On}/I_{Off} > 10^5\ \mu A/\mu m$. Fig. 6.5b shows the I_{On} versus I_{On}/I_{Off} curves meeting these conditions. The best on-off performance is obtained when the In content is 0.2 and the N_{SD} is 2×10^{19}.

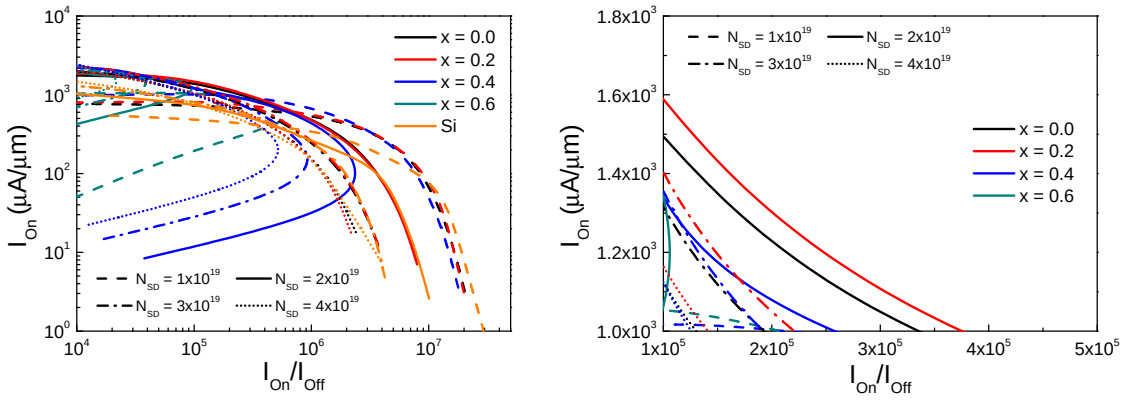


Figure 6.5: (a) ON-current as a function of ON/OFF ratios at $V_{DD} = 0.50$ V for Si and InGaAs MOSFETs for various N_{SD}'s . (b) Same Simulation results between our performance target ($I_{on} > 1000\ \mu A/\mu m$ and $I_{on}/I_{off} > 10^5$)

6.4 Conclusion

To summarize, $In_xGa_{1-x}As$ transistors with ultra-scaled gate length (10 nm) were studied using a quantum transport model within the NEGF formalism including scattering effects. We explored the effect of indium mole fraction and S/D doping in $In_xGa_{1-x}As$ DG MOSFETs. High indium content (higher than 0.6) shows high GIDL current due to the direct and low band-gap. Increased of In content and N_{SD} has a trade-off relationship between increased on-current and decreased off-current. By calculating $I-V$ characteristic for various compositions and N_{SD}, we find that the optimum values for N_{SD} and Indium mole fraction are 2×10^{19} and $x = 0.2$, respectively. This condition could provide an on-current of 1590 $\mu A/\mu m$ at 10^5 On/Off ratio.

Chapter 7

Summary and Future work

7.1 Summary

The focus of this thesis was to investigate the performance of III-V and Si channel devices at future technology generation using the non-equilibrium Greens function (NEGF) formalism coupled with a 20 orbital $sp^3d^5s^*$-SO tight-binding model. Chapter 1 served as an introduction to the main themes of this dissertation. This chapter gave an introduction to the III-V devices and quantum transport models based on NEGF approach. I began by discussing the demands on high mobility semiconductors as the alternatives for post-Si logic application. This was followed by a brief overview of current research and challenges of III-V MOSFETs. In Chapter 1.3, we reviewed the NEGF method providing rigorous framework for treating quantum transport in nanoscale semiconductors.

The necessity of atomistic modeling for bandstructure and certain aspects related to using tight-binding method were discussed in Chapter 2. In confined semiconductor structures, the quantum confinement effect along the thickness direction increase bandgap and effective mass. Since effective masses calculated from the bulk do not provide a good description of these subband, rigorous band structure should be used to study electrical transport. A 20 orbital $sp^3d^5s^*$-SO is used to study nanoscale bandstructure. This model is a powerful tool to predict full bandstructure of nanoscale semiconductors accurately.

In Chapter 3, a mode space NEGF approach including band-to-band tunneling with the WKB approximation is proposed for simulating double gate MOSFETs with direct bandgap semiconductor channels. Due to significant transport advantage, high mobility III-V materials are very actively being researched as channel materials for future highly scaled CMOS devices. However, most high mobility materials also have a significantly smaller bandgap compared to Si, leading to very high band-to-band tunneling (BTBT) leakage current. Unexpectedly, we have found that this off-state leakage current could be significantly underestimated in commonly used quantum transport calculations using the NEGF formalism. This is because the conventional NEGF method ignores electron-electron interaction and cannot fill the empty states under valence band which are cut off by the source side band edges. We

have proposed a new method, the coupled mode-space NEGF with two terminal WKB approach, that allows for estimating tunneling current correctly. When the appropriate filling of the body of the transistor is not taken into account, the off-state leakage current for an InAs transistor could be underestimated by almost 2 orders of magnitude.

In Chapter 4, a framework using NEGF formalism capable of handling quantum transport including carrier scattering have developed. Proper treatment of microscopic scattering is necessary to assess the performance of channel materials because a certain amount of scattering exists and influences the transport of the free carriers, even when the device length is smaller than the mean free path. The mode space approach is adopted to reduce the computational complexity and the local self-energy functions for the electron-phonon interactions have been rigorously derived within the self-consistent Born approximation. We verify the obtained the self-energy functions by comparing the scattering rates obtained by NEGF formalism and Fermis golden rule. The effect of carrier scattering on the quantum transport of ultra-thin body transistors is investigated. In the presence of electron-phonon interactions, the drain current decreases compared with the ballistic limit and the current reduction ratio increases as the channel length increases. The phonon-limited low field mobility and mean free path in the long channel limit are obtained by extracting the apparent mobility and the ballistic mobility for different gate lengths. For III-V MOSFETs, surface roughness scattering should be included to predict the performance for current technology since the calculated mobility with POP scattering alone is larger than the mobility extracted from experimental data.

Effect of electron phonon coupling in ultra thin body Si MOFETs with small channel lengths has been explored in Chapter 5. The phonon limited effective mobility is significantly affected by the coupling between subbands and transverse modes. Increased density of states leads to enhanced scattering. It is found that this could lead to almost two times reduction in the effective mobility at large gate voltages. At very small channel lengths, direct source to drain tunneling, Drain Induced Barrier Lowering and electron phonon scattering will combine together to determine the eventual current-voltage characteristic.

In chapter 6, we explored the effects of indium mole fraction and S/D doping in $In_xGa_{1-x}As$ DG MOSFETs with 10 nm gate length thorough the scattering-NEGF method. High indium content InGaAs with direct and low band-gap show high off-state current due to body-drain band-to-band tunneling. The In content and S/D doping has a trade-off relationship between on-current and off-current. We found that there exists an optimum indium mole fraction and S/D doping concentration that maximizes Ion/Ioff ratio and Ion by balancing injection velocity and short channel effect.

7.2 Future work

For the future study of the present work, we suggest several interesting topics as follows:

1. In this dissertation, strain effects are not considered, and we focus on the intrinsic material properties. However, strain effects can significantly alter the bandstructure through

effects such as band splitting and band wrapping and modify carrier effective mass and scattering rate. As process-induced strain engineering is being used as performance booster in today's scaled in Si devices, strain effects can enhance the performance of III-V MOSETs. Furthermore, thin films with a heterostructure can get high stress and strain due to lattice mismatch. Within the tight-binding method, the parameters can be modified to predict bandstructure depending on the magnitude of strain. We can also include the effect of crystal orientation. The crystal orientation effect for III-V semiconductors would not be as large as Si because the bulk conduction band for III-V is located in the Γ point, which is high symmetry point.

2. We can improve the accuracy of the electron-phonon scattering model instead of using several assumptions that we have made for this work. First, our approach used a simplified bulk phonon dispersion for the semiconductors and the electron-phonon scattering is computed with a dispersionless acoustic mode and with fixed-energy optical modes. Accurate full phonon dispersions can be obtained through various approaches such as *ab-initio* methods, the adiabatic bond charge model and the valence force field model. However, since full band electron-phonon interaction is computationally intractable and actual the device structure such as source/drain, gate, insulator and substrate also affect the phonon dispersion, a model that represent actual phonon dispersion is required. Second, we only include local scattering interactions to reduce computational burden. In the self-energies for POP scattering, there are nonlocal objects that depend on the spatial coordinates difference. Non-local scattering will increase the scattering rate and form coupling between the electrons at different positions. Non-local scattering can only be included by full matrix calculation instead of using recursive algorithms which use only use diagonal elements. Third, we can include other scattering mechanisms such as alloy disorder and ionized impurities. Although the self-energy of the two-scattering mechanism already exists for the real space NEGF, an appropriate self-energy for 2D mode-space should be derived. Alternatively, we can include impurity scattering within NEGF by the simulation of an ensemble of devices by adding randomly placed perturbation potentials for impurities. While this approach involves inhomogeneous screening and multiple scattering, it needs many simulations to obtain average values.

3. One of the interesting directions would be the quantum transport in p-type MOS devices. Since the valence band structure shows significant nonparabolicity and anisotropy owing to mixed dispersion between heavy-hole, light-hole, and split-off bands, the full atomistic Hamiltonian with spin-orbit coupling should be used. The situation in nanostructures is even more complicated due to confinement effects. This is why there are few studies on p-type devices. While there are several studies using a simplified six-band k.p model, the accuracy of this model remains unsatisfactory, and its agreement with experimental data is still unclear. Scattering-NEGF with the atomistic Hamiltonian is computationally challenging, but can be simulated for 1D nanowires or thin UTB MOSFETs.

Bibliography

[1] I. Ferain, C. A. Colinge, and J.-P. Colinge, "Multigate transistors as the future of classical metal-oxide-semiconductor field-effect transistors," *Nature*, vol. 479, no. 7373, pp. 310–316, 2011.

[2] B. H. Calhoun, Y. Cao, X. Li, K. Mai, L. T. Pileggi, R. A. Rutenbar, and K. L. Shepard, "Digital circuit design challenges and opportunities in the era of nanoscale cmos," *Proceedings of the IEEE*, vol. 96, no. 2, pp. 343–365, 2008.

[3] R. Chau, B. Doyle, S. Datta, J. Kavalieros, and K. Zhang, "Integrated nanoelectronics for the future," *Nature materials*, vol. 6, no. 11, pp. 810–812, 2007.

[4] K. Bernstein, D. J. Frank, A. E. Gattiker, W. Haensch, B. L. Ji, S. R. Nassif, E. J. Nowak, D. J. Pearson, and N. J. Rohrer, "High-performance cmos variability in the 65-nm regime and beyond," *IBM journal of research and development*, vol. 50, no. 4.5, pp. 433–449, 2006.

[5] J. A. Del Alamo, "Nanometre-scale electronics with iii-v compound semiconductors," *Nature*, vol. 479, no. 7373, pp. 317–323, 2011.

[6] J.-H. Hur and S. Jeon, "Iii–v compound semiconductors for mass-produced nanoelectronics: theoretical studies on mobility degradation by dislocation," *Scientific reports*, vol. 6, 2016.

[7] H. Riel, L.-E. Wernersson, M. Hong, and J. A. Del Alamo, "Iii–v compound semiconductor transistorsfrom planar to nanowire structures," *Mrs Bulletin*, vol. 39, no. 8, pp. 668–677, 2014.

[8] L. D. Nguyen, A. S. Brown, M. A. Thompson, and L. M. Jelloian, "50-nm self-aligned-gate pseudomorphic alinas/gainas high electron mobility transistors," *IEEE Transactions on Electron Devices*, vol. 39, no. 9, pp. 2007–2014, 1992.

[9] L. D. Nguyen, L. E. Larson, and U. K. Mishra, "Ultra-high speed modulation-doped field-effect transistors: a tutorial review," *Proceedings of the IEEE*, vol. 80, no. 4, pp. 494–518, 1992.

[10] W. Spicer, I. Lindau, P. Skeath, and C. Su, "Unified defect model and beyond," *Journal of Vacuum Science and Technology*, vol. 17, no. 5, pp. 1019–1027, 1980.

[11] A. Dimoulas, P. Tsipas, A. Sotiropoulos, and E. Evangelou, "Fermi-level pinning and charge neutrality level in germanium," *Applied Physics Letters*, vol. 89, no. 25, p. 252110, 2006.

[12] M. Scarrozza, G. Pourtois, M. Houssa, M. Caymax, A. Stesmans, M. Meuris, and M. Heyns, "A theoretical study of the initial oxidation of the gaas (001)-β 2 (2$\times$ 4) surface," *Applied physics letters*, vol. 95, no. 25, p. 253504, 2009.

[13] W. Wang, K. Xiong, R. M. Wallace, and K. Cho, "Impact of interfacial oxygen content on bonding, stability, band offsets, and interface states of gaas: Hfo2 interfaces," *The Journal of Physical Chemistry C*, vol. 114, no. 51, pp. 22 610–22 618, 2010.

[14] M. Passlack, M. Hong, J. P. Mannaerts, S. Chu, R. L. Opila, and N. Moriya, "In-situ ga/sub 2/o/sub 3/process for gaas inversion/accumulation device and surface passivation applications," in *Electron Devices Meeting, 1995. IEDM'95., International*. IEEE, 1995, pp. 383–386.

[15] P. Ye, G. Wilk, J. Kwo, B. Yang, H.-J. Gossmann, M. Frei, S. Chu, J. Mannaerts, M. Sergent, M. Hong *et al.*, "Gaas mosfet with oxide gate dielectric grown by atomic layer deposition," *IEEE Electron Device Letters*, vol. 24, no. 4, pp. 209–211, 2003.

[16] S. Oktyabrsky and D. Y. Peide, *Fundamentals of III-V semiconductor MOSFETs*. Springer, 2010.

[17] C. Hinkle, A. Sonnet, E. Vogel, S. McDonnell, G. Hughes, M. Milojevic, B. Lee, F. Aguirre-Tostado, K. Choi, H. Kim *et al.*, "Gaas interfacial self-cleaning by atomic layer deposition," *Applied Physics Letters*, vol. 92, no. 7, p. 071901, 2008.

[18] L. Lin and J. Robertson, "Defect states at iii-v semiconductor oxide interfaces," *Applied Physics Letters*, vol. 98, no. 8, p. 082903, 2011.

[19] É. OConnor, B. Brennan, V. Djara, K. Cherkaoui, S. Monaghan, S. Newcomb, R. Contreras, M. Milojevic, G. Hughes, M. Pemble *et al.*, "A systematic study of (nh 4) 2 s passivation (22%, 10%, 5%, or 1%) on the interface properties of the al 2 o 3/in 0.53 ga 0.47 as/inp system for n-type and p-type in 0.53 ga 0.47 as epitaxial layers," *Journal of Applied Physics*, vol. 109, no. 2, p. 024101, 2011.

[20] Y. Wu, T. Lin, T. Chiang, Y. Chang, H. Chiu, Y. Lee, M. Hong, C. Lin, and J. Kwo, "Engineering of threshold voltages in molecular beam epitaxy-grown al 2 o 3/ ga 2 o 3 (gd 2 o 3)/ in 0.2 ga 0.8 as," *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena*, vol. 28, no. 3, pp. C3H10–C3H13, 2010.

[21] C.-W. Cheng, G. Apostolopoulos, and E. A. Fitzgerald, "The effect of interface processing on the distribution of interfacial defect states and the cv characteristics of iii-v metal-oxide-semiconductor field effect transistors," *Journal of Applied Physics*, vol. 109, no. 2, p. 023714, 2011.

[22] Y.-T. Chen, H. Zhao, Y. Wang, F. Xue, F. Zhou, and J. C. Lee, "Fluorinated hfo 2 gate dielectric engineering on in 0.53 ga 0.47 as metal-oxide-semiconductor field-effect-transistors," *Applied Physics Letters*, vol. 96, no. 10, p. 103506, 2010.

[23] S. Kim, M. Yokoyama, N. Taoka, R. Iida, S. Lee, R. Nakane, Y. Urabe, N. Miyata, T. Yasuda, H. Yamada *et al.*, "Self-aligned metal source/drain inxga1-xas n-metal–oxide–semiconductor field-effect transistors using ni–ingaas alloy," *Applied Physics Express*, vol. 4, no. 2, p. 024201, 2011.

[24] E. Y.-J. Kong, S. Subramanian, Q. Zhou, J. Pan, Y.-C. Yeo *et al.*, "Coingaas as a novel self-aligned metallic source/drain material for implant-less in 0.53 ga 0.47 as n-mosfets," *Solid-State Electronics*, vol. 78, pp. 62–67, 2012.

[25] W. Lu, A. Guo, A. Vardi, and J. A. Del Alamo, "A test structure to characterize nanoscale ohmic contacts in iii-v mosfets," *IEEE Electron Device Letters*, vol. 35, no. 2, pp. 178–180, 2014.

[26] A. Baraskar, A. Gossard, and M. J. Rodwell, "Lower limits to metal-semiconductor contact resistance: Theoretical models and experimental data," *Journal of Applied Physics*, vol. 114, no. 15, p. 154516, 2013.

[27] M. Lundstrom, *Fundamentals of carrier transport*. Cambridge University Press, 2009.

[28] K. Ganapathi, "Tunneling in low-power device-design: A bottom-up view of issues, challenges, and opportunities," Ph.D. dissertation, University of California, Berkeley, 2013.

[29] D. Vasileska, D. Mamaluy, I. Knezevic, H. Khan, S. Goodnick, and H. Nalwa, "Quantum transport in nanoscale devices," *Encyclopedia of Nanoscience and Nanotechnology. American Scientific Publishers, Syracuse*, 2010.

[30] D. Vasileska, S. M. Goodnick, and G. Klimeck, *Computational Electronics: Semiclassical and Quantum Device Modeling and Simulation*. CRC press, 2016.

[31] L. V. Keldysh *et al.*, "Diagram technique for nonequilibrium processes," *Sov. Phys. JETP*, vol. 20, no. 4, pp. 1018–1026, 1965.

[32] L. P. Kadanoff and G. A. Baym, *Quantum statistical mechanics: Green's function methods in equilibrium and nonequilibirum problems*. Benjamin, 1962.

[33] P. C. Martin and J. Schwinger, "Theory of many-particle systems. i," *Physical Review*, vol. 115, no. 6, p. 1342, 1959.

[34] J. Schwinger, "Brownian motion of a quantum oscillator," *Journal of Mathematical Physics*, vol. 2, no. 3, pp. 407–432, 1961.

[35] C. Caroli, R. Combescot, D. Lederer, P. Nozieres, and D. Saint-James, "A direct calculation of the tunnelling current. ii. free electron description," *Journal of Physics C: Solid State Physics*, vol. 4, no. 16, p. 2598, 1971.

[36] R. Combescot, "A direct calculation of the tunnelling current. iii. effect of localized impurity states in the barrier," *Journal of Physics C: Solid State Physics*, vol. 4, no. 16, p. 2611, 1971.

[37] C. Caroli, R. Combescot, P. Nozieres, and D. Saint-James, "Direct calculation of the tunneling current," *Journal of Physics C: Solid State Physics*, vol. 4, no. 8, p. 916, 1971.

[38] S. Datta, *Electronic transport in mesoscopic systems*. Cambridge university press, 1997.

[39] ——, *Quantum transport: atom to transistor*. Cambridge University Press, 2005.

[40] ——, "Non-equilibrium green's function (negf) method: a different perspective," in *Computational Electronics (IWCE), 2015 International Workshop on*. IEEE, 2015, pp. 1–6.

[41] M. Paulsson, "Non equilibrium green's functions for dummies: Introduction to the one particle negf equations," *arXiv preprint cond-mat/0210519*, 2002.

[42] T. Ando, A. B. Fowler, and F. Stern, "Electronic properties of two-dimensional systems," *Reviews of Modern Physics*, vol. 54, no. 2, p. 437, 1982.

[43] J.-L. P. van der Steen, D. Esseni, P. Palestri, L. Selmi, and R. J. Hueting, "Validity of the parabolic effective mass approximation in silicon and germanium n-mosfets with different crystal orientations," *IEEE Transactions on Electron Devices*, vol. 54, no. 8, pp. 1843–1851, 2007.

[44] S. Laux, A. Kumar, and M. Fischetti, "Analysis of quantum ballistic electron transport in ultrasmall silicon devices including space-charge and geometric effects," *Journal of Applied Physics*, vol. 95, no. 10, pp. 5545–5582, 2004.

[45] L. Lucci, P. Palestri, D. Esseni, and L. Selmi, "Multi-subband monte carlo modeling of nano-mosfets with strong vertical quantization and electron gas degeneration," in *Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International*. IEEE, 2005, pp. 617–620.

[46] J.-M. Jancu, R. Scholz, F. Beltram, and F. Bassani, "Empirical spds* tight-binding calculation for cubic semiconductors: General method and material parameters," *Physical Review B*, vol. 57, no. 11, p. 6493, 1998.

[47] M. Luisier, A. Schenk, W. Fichtner, and G. Klimeck, "Atomistic simulation of nanowires in the s p 3 d 5 s* tight-binding formalism: From boundary conditions to strain calculations," *Physical Review B*, vol. 74, no. 20, p. 205323, 2006.

[48] Y. Peter and M. Cardona, *Fundamentals of semiconductors: physics and materials properties*. Springer Science & Business Media, 2010.

[49] J. C. Slater and G. F. Koster, "Simplified lcao method for the periodic potential problem," *Physical Review*, vol. 94, no. 6, p. 1498, 1954.

[50] P. R. Wallace, "The band theory of graphite," *Physical Review*, vol. 71, no. 9, p. 622, 1947.

[51] S. Lee, F. Oyafuso, P. Von Allmen, and G. Klimeck, "Boundary conditions for the electronic structure of finite-extent embedded semiconductor nanostructures," *Physical Review B*, vol. 69, no. 4, p. 045316, 2004.

[52] A. Zunger and M. L. Cohen, "First-principles nonlocal-pseudopotential approach in the density-functional formalism. ii. application to electronic and structural properties of solids," *Physical Review B*, vol. 20, no. 10, p. 4082, 1979.

[53] S. L. Richardson, M. L. Cohen, S. G. Louie, and J. R. Chelikowsky, "Electron charge densities at conduction-band edges of semiconductors," *Physical Review B*, vol. 33, no. 2, p. 1177, 1986.

[54] á. Vogl, H. P. Hjalmarson, and J. D. Dow, "A semi-empirical tight-binding theory of the electronic structure of semiconductors," *Journal of Physics and Chemistry of Solids*, vol. 44, no. 5, pp. 365–378, 1983.

[55] Y.-C. Chang and D. Aspnes, "Theory of dielectric-function anisotropies of (001) gaas (2× 1) surfaces," *Physical Review B*, vol. 41, no. 17, p. 12002, 1990.

[56] P. Boguslawski and I. Gorczyca, "Atomic-orbital interpretation of electronic structure of iii-v semiconductors: Gaas versus alas," *Semiconductor science and technology*, vol. 9, no. 12, p. 2169, 1994.

[57] T. B. Boykin, G. Klimeck, and F. Oyafuso, "Valence band effective-mass expressions in the sp 3 d 5 s* empirical tight-binding model applied to a si and ge parametrization," *Physical Review B*, vol. 69, no. 11, p. 115201, 2004.

[58] T. B. Boykin, G. Klimeck, R. C. Bowen, and F. Oyafuso, "Diagonal parameter shifts due to nearest-neighbor displacements in empirical tight-binding theory," *Physical Review B*, vol. 66, no. 12, p. 125207, 2002.

[59] R. Winkler, "Origin of spin–orbit coupling effects," *Spin–Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems*, pp. 61–68, 2003.

[60] A. Rahman, "Exploring new channel materials for nanoscale cmos devices: A simulation approach," Ph.D. dissertation, Purdue University, 2005.

[61] Y. Liu, N. Neophytou, G. Klimeck, and M. S. Lundstrom, "Band-structure effects on the performance of iii–v ultrathin-body soi mosfets," *IEEE Transactions on Electron Devices*, vol. 55, no. 5, pp. 1116–1122, 2008.

[62] D. K. Ferry, *Semiconductors: Bonds and bands*. IOP Publishing Ltd, 2013.

[63] "International technology roadmap for semiconductors," *Semiconductor Industry Association*, 2015.

[64] M. Huang, Y. Chang, C. Chang, Y. Lee, P. Chang, J. Kwo, T. Wu, and M. Hong, "Surface passivation of III-V compound semiconductors using atomic-layer-deposition-grown Al_2O_3," *Appl. Phys. Lett.*, vol. 87, no. 25, p. 252104, 2005.

[65] S. Datta, "The non-equilibrium green's function (negf) formalism: An elementary introduction," in *Electron Devices Meeting, 2002. IEDM'02. International*. IEEE, 2002, pp. 703–706.

[66] A. Martinez, M. Bescond, J. R. Barker, A. Svizhenko, M. Anantram, C. Millar, and A. Asenov, "A self-consistent full 3-d real-space negf simulator for studying nonperturbative effects in nano-mosfets," *IEEE Transactions on Electron Devices*, vol. 54, no. 9, pp. 2213–2222, 2007.

[67] G. Fiori and G. Iannaccone, "Simulation of graphene nanoribbon field-effect transistors," *IEEE Electron Device Letters*, vol. 28, no. 8, pp. 760–762, 2007.

[68] D. Basu, R. Kotlyar, C. E. Weber, and M. A. Stettler, "Ballistic band-to-band tunneling in the off state in ingaas mosfets," *IEEE Transactions on Electron Devices*, vol. 61, no. 10, pp. 3417–3422, 2014.

[69] M. Luisier, "Performance comparison of gasb, strained-si, and ingaas double-gate ultrathin-body n-fets," *IEEE electron device letters*, vol. 32, no. 12, pp. 1686–1688, 2011.

[70] D. Lizzit, D. Esseni, P. Palestri, P. Osgnach, and L. Selmi, "Performance Benchmarking and Effective Channel Length for Nanoscale InAs, $in_{0.53}ga_{0.47}$as, and ssi n-MOSFETs," *IEEE Trans. Electron Devices*, vol. 61, no. 6, pp. 2027–2034, 2014.

[71] N. Kharche, G. Klimeck, D.-H. Kim, J. A. Del Alamo, and M. Luisier, "Multiscale metrology and optimization of ultra-scaled InAs quantum well FETs," *IEEE Trans. Electron Devices*, vol. 58, no. 7, pp. 1963–1971, 2011.

[72] G. Klimeck, R. C. Bowen, T. B. Boykin, and T. A. Cwik, "sp3s* tight-binding parameters for transport simulations in compound semiconductors," *Superlattices and Microstructures*, vol. 27, no. 5, pp. 519–524, 2000.

[73] G. Klimeck, R. Lake, R. C. Bowen, W. R. Frensley, and T. S. Moise, "Quantum device simulation with a generalized tunneling formula," *Appl. Phys. Lett.*, vol. 67, no. 17, pp. 2539–2541, 1995.

[74] K. Ganapathi and S. Salahuddin, "Zener tunneling: Congruence between semi-classical and quantum ballistic formalisms," *Journal of Applied Physics*, vol. 111, no. 12, p. 124506, 2012.

[75] Z. Ren, R. Venugopal, S. Goasguen, S. Datta, and M. S. Lundstrom, "nanomos 2.5: A two-dimensional simulator for quantum transport in double-gate mosfets," *Electron Devices, IEEE Transactions on*, vol. 50, no. 9, pp. 1914–1925, 2003.

[76] Y.-C. Chang and J. Schulman, "Complex band structures of crystalline solids: An eigenvalue method," *Physical Review B*, vol. 25, no. 6, p. 3975, 1982.

[77] X. Guan, D. Kim, K. C. Saraswat, and H.-S. P. Wong, "Complex band structures: From parabolic to elliptic approximation," *Electron Device Letters, IEEE*, vol. 32, no. 9, pp. 1296–1298, 2011.

[78] M. Büttiker, "Four-terminal phase-coherent conductance," *Physical review letters*, vol. 57, no. 14, p. 1761, 1986.

[79] M. Anantram, M. S. Lundstrom, and D. E. Nikonov, "Modeling of nanoscale devices," *Proceedings of the IEEE*, vol. 96, no. 9, pp. 1511–1550, 2008.

[80] D. Nikonov, H. Pal, and G. Bourianoff, "Scattering in negf: Made simple," http://nanohub.org/resources/7772, 2009.

[81] R. Venugopal, Z. Ren, S. Datta, M. Lundstrom, and D. Jovanovic, "Simulating quantum transport in nanoscale transistors: Real versus mode-space approaches," *Journal of Applied Physics*, vol. 92, no. 7, pp. 3730–3739, 2002.

[82] S. Datta, "A simple kinetic equation for steady-state quantum transport," *Journal of Physics: Condensed Matter*, vol. 2, no. 40, p. 8023, 1990.

[83] G. D. Mahan, "Quantum transport equation for electric and magnetic fields," *Physics Reports*, vol. 145, no. 5, pp. 251–318, 1987.

[84] S. Jin, Y. J. Park, and H. S. Min, "A three-dimensional simulation of quantum transport in silicon nanowire transistor in the presence of electron-phonon interactions," *Journal of Applied Physics*, vol. 99, no. 12, p. 123719, 2006.

[85] S. O. Koswatta, S. Hasan, M. S. Lundstrom, M. Anantram, and D. E. Nikonov, "Non-equilibrium green's function treatment of phonon scattering in carbon nanotube transistors," *Transactions on Electron Devices, IEEE*, vol. 54, no. 9, pp. 2339–2351, 2007.

[86] A. Svizhenko, M. Anantram, T. Govindan, B. Biegel, and R. Venugopal, "Two-dimensional quantum mechanical modeling of nanotransistors," *Journal of Applied Physics*, vol. 91, no. 4, pp. 2343–2354, 2002.

[87] R. Venugopal, S. Goasguen, S. Datta, and M. Lundstrom, "Quantum mechanical analysis of channel access geometry and series resistance in nanoscale transistors," *Journal of Applied Physics*, vol. 95, no. 1, pp. 292–305, 2004.

[88] R. Grassi, A. Gnudi, I. Imperiale, E. Gnani, S. Reggiani, and G. Baccarani, "Mode space approach for tight-binding transport simulations in graphene nanoribbon field-effect transistors including phonon scattering," *Journal of Applied Physics*, vol. 113, no. 14, p. 144506, 2013.

[89] Y. Liu, N. Neophytou, T. Low, G. Klimeck, and M. S. Lundstrom, "A tight-binding study of the ballistic injection velocity for ultrathin-body soi mosfets," *IEEE Transactions on Electron Devices*, vol. 55, no. 3, pp. 866–871, 2008.

[90] A. Pan, G. Leung, and C. O. Chui, "Junctionless silicon and in 0.53 ga 0.47 as transistorspart i: Nominal device evaluation with quantum simulations," *IEEE Transactions on Electron Devices*, vol. 62, no. 10, pp. 3199–3207, 2015.

[91] A. Wacker, A.-P. Jauho, S. Rott, A. Markus, P. Binder, and G. H. Döhler, "Inelastic quantum transport in superlattices: success and failure of the boltzmann equation," *Physical review letters*, vol. 83, no. 4, p. 836, 1999.

[92] R. Venugopal, M. Paulsson, S. Goasguen, S. Datta, and M. Lundstrom, "A simple quantum mechanical treatment of scattering in nanoscale transistors," *Journal of Applied Physics*, vol. 93, no. 9, pp. 5613–5625, 2003.

[93] H. S. Pal, D. E. Nikonov, R. Kim, and M. S. Lundstrom, "Electron-phonon scattering in planar mosfets with negf," in *Silicon Nanoelectronics Workshop (SNW), 2010*. IEEE, 2010, pp. 1–2.

[94] K. Natori, "Ballistic metal-oxide-semiconductor field effect transistor," *Journal of Applied Physics*, vol. 76, no. 8, pp. 4879–4890, 1994.

[95] M. S. Shur, "Low ballistic mobility in submicron hemts," *IEEE Electron Device Letters*, vol. 23, no. 9, pp. 511–513, 2002.

[96] J. Wang and M. Lundstrom, "Ballistic transport in high electron mobility transistors," *IEEE Transactions on Electron Devices*, vol. 50, no. 7, pp. 1604–1609, 2003.

[97] H. S. Pal, D. E. Nikonov, R. Kim, and M. S. Lundstrom, "Electron-phonon scattering in planar mosfets: Negf and monte carlo methods," *arXiv preprint arXiv:1209.4878*, 2012.

[98] M. Lundstrom, "Elementary scattering theory of the si mosfet," *IEEE Electron Device Letters*, vol. 18, no. 7, pp. 361–363, 1997.

[99] S. Rakheja, M. Lundstrom, and D. Antoniadis, "A physics-based compact model for fets from diffusive to ballistic carrier transport regimes," in *Electron Devices Meeting (IEDM), 2014 IEEE International*. IEEE, 2014, pp. 35–1.

[100] M. Lundstrom, S. Datta, and X. Sun, "Emission–diffusion theory of the mosfet," *IEEE Transactions on Electron Devices*, vol. 62, no. 12, pp. 4174–4178, 2015.

[101] C. Jeong, D. A. Antoniadis, and M. S. Lundstrom, "On backscattering and mobility in nanoscale silicon mosfets," *IEEE Transactions on electron devices*, vol. 56, no. 11, pp. 2762–2769, 2009.

[102] M. S. Lundstrom and D. A. Antoniadis, "Compact models and the physics of nanoscale fets," *IEEE Transactions on Electron Devices*, vol. 61, no. 2, pp. 225–233, 2014.

[103] M. V. Fischetti and S. E. Laux, "Monte carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects," *Physical Review B*, vol. 38, no. 14, p. 9721, 1988.

[104] M. Fischetti, "Master-equation approach to the study of electronic transport in small semiconductor devices," *Physical Review B*, vol. 59, no. 7, p. 4901, 1999.

[105] M. V. Fischetti, "Monte carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures. i. homogeneous transport," *IEEE Transactions on Electron Devices*, vol. 38, no. 3, pp. 634–649, 1991.

[106] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, "Single and multiband modeling of quantum electron transport through layered semiconductor devices," *Journal of Applied Physics*, vol. 81, no. 12, pp. 7845–7869, 1997.

[107] T. Kubis and P. Vogl, "Assessment of approximations in nonequilibrium greens function theory," *Physical Review B*, vol. 83, no. 19, p. 195304, 2011.

[108] R. K. Lake and R. R. Pandey, "Non-equilibrium green functions in electronic device modeling," *arXiv preprint cond-mat/0607219*, 2006.

[109] H. S. Pal, "Device physics studies of iii-v and silicon mosfets for digital logic," Ph.D. dissertation, Purdue University, 2010.

[110] M. O. Baykan, S. E. Thompson, and T. Nishida, "Strain effects on three-dimensional, two-dimensional, and one-dimensional silicon logic devices: Predicting the future of strained silicon," *Journal of Applied Physics*, vol. 108, no. 9, p. 093716, 2010.

[111] S.-i. Takagi, J. L. Hoyt, J. J. Welser, and J. F. Gibbons, "Comparative study of phonon-limited mobility of two-dimensional electrons in strained and unstrained si metal–oxide–semiconductor field-effect transistors," *Journal of Applied Physics*, vol. 80, no. 3, pp. 1567–1577, 1996.

[112] A. Rahman and M. S. Lundstrom, "A compact scattering model for the nanoscale double-gate mosfet," *IEEE Transactions on Electron Devices*, vol. 49, no. 3, pp. 481–489, 2002.

[113] J. Lin, Y. Wu, J. A. del Alamo, and D. A. Antoniadis, "Analysis of resistance and mobility in ingaas quantum-well mosfets from ballistic to diffusive regimes," *IEEE Transactions on Electron Devices*, vol. 63, no. 4, pp. 1464–1470, 2016.

[114] M. Luisier and G. Klimeck, "Investigation of in_xga_ {1-x} as ultra-thin-body tunneling fets using a full-band and atomistic approach," in *Simulation of Semiconductor Processes and Devices, 2009. SISPAD'09. International Conference on*. IEEE, 2009, pp. 1–4.

[115] H. S. Pal, T. Low, and M. S. Lundstrom, "Negf analysis of ingaas schottky barrier double gate mosfets," in *Electron Devices Meeting, 2008. IEDM 2008. IEEE International*. IEEE, 2008, pp. 1–4.

[116] N. Neophytos, R. Titash, and M. S. Lundstrom, "Performance analysis of 60-nm gate-length iii-v ingaas hemts: Simulations versus experiments," *IEEE TRANSACTIONS ON ELECTRON DEVICES*, vol. 56, no. 7, p. 1377, 2009.

[117] S. M. Sze and K. K. Ng, *Physics of semiconductor devices*. John Wiley & Sons, 2006.