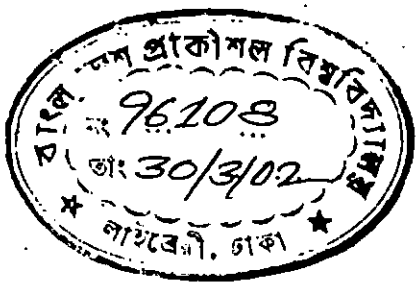
Self-Consistent Calculation of Direct Tunneling Gate Leakage Current in Deep Submicron n- and p-MOSFETs in Presence of Inelastic Scattering

A thesis submitted to
the department of Electrical and Electronic Engineering
of
Bangladesh University of Engineering and Technology
in partial fulfillment of the requirement
for the degree of
MASTER OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING



by
Khairul Alam

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY
2002

The thesis titled " **Self-Consistent Calculation of Direct Tunneling Gate Leakage Current in Deep Submicron n- and p-MOSFETs in Presence of Inelastic Scattering**" Submitted by Khairul Alam, Roll No.: 040006244F, Session: April 2000 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of MASTER OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING on March 13, 2002.

BOARD OF EXAMINERS

1. _____ 13/3/2002
Dr. Anisul Haque — **Chairman**
Associate Professor — (Supervisor)
Department of Electrical and Electronic Engineering
BUET, Dhaka-1000, Bangladesh.

2. _____ 13/3/02
Dr. M. Rezwan Khan — **Member**
Professor
Department of Electrical and Electronic Engineering
BUET, Dhaka-1000, Bangladesh.

3. _____
Dr. M. M. Shahidul Hassan — **Member**
Professor and Head — (Ex-officio)
Department of Electrical and Electronic Engineering
BUET, Dhaka-1000, Bangladesh.

4. _____
Dr. M. Ali Asgar — **Member**
Professor — (External)
Department of Physics
BUET, Dhaka-1000, Bangladesh.

Declaration

I hereby declare that this thesis or any part of it has not been submitted elsewhere for the award of any degree or diploma.

Signature of the candidate

(Khairul Alam)

Dedication

To My Parents.

Contents

List of Figures

List of Tables

Acknowledgement

I wish to convey my heartiest gratitude and profound respect to my supervisor Dr. Anisul Haque, Associate Professor, Department of Electrical and Electronic Engineering (EEE), Bangladesh University of Engineering and Technology (BUET), Bangladesh, for his continuous guidance, suggestions and wholehearted supervision throughout the progress of this work, without which this thesis never be materialized. I am grateful to him for acquainting me with the world of advanced research.

I am grateful to Mr. Abul Husain, Associate Professor and Head, Dept. of EEE, Ahsanullah University of Science and Technology (AUST), who provided with all the facilities of VLSI laboratory of the department and cooperation to complete the work. I want to thank S. Zaman, Lecturer, Dept. of EEE , AUST, M. M. Chowdhury, Lecturer, Dept. of EEE, AUST and I. B. Chowdhury, former Lecturer, Dept. of EEE, AUST for their continuous suggestions, inspiration, encouragement and many fruitful discussions. In this regard, I also like to express thanks and gratitude to Dr. M. M. Shahidul Hassan, Professor and Head, Dept. of EEE, BUET.

I want to thank my friends Yunus, Ifthekar and many others, who were directly or indirectly related to this work, for their support and encouragement. I also thank all the personnel at the departmental library, BUET reference library and xerox section for providing me with the valuable journals and thesis papers to complete this work.

Abstract

Direct tunneling gate leakage current in both n- and p-MOS devices with ultra-thin gate-oxide is studied. The effects of inelastic scattering of the inversion carriers, tunneled into the gate-oxide region, on the direct tunneling gate current are investigated. Coupled Schrödinger's and Poisson's equations are solved self-consistently. Open boundary conditions, taking into account the wave function tail inside the gate-oxide region, are used to solve Schrödinger's equation within the self-consistent loop. Also, the fraction of inversion charges inside the gate-oxide region is taken into account in the solution of Poisson's equation. Inelastic scattering is represented by a collision time which appears as an imaginary potential in Schrödinger's equation. Simulated direct tunneling currents are compared with published experimental results. Inelastic scattering effects on direct tunneling current are found to be significant in devices with oxide thickness $\geq$ 2 nm and at lower gate voltages. Therefore, the existing mismatch at lower gate voltages between experimental and simulated direct tunneling gate currents can be explained in terms of inelastic scattering effects. However, for accurate modeling of direct tunneling current, appropriate spatial and gate bias dependence of collision time needs to be taken into account. It is also found that carrier effective mass in gate-oxide region is not a constant, rather a function of the applied gate bias. Moreover, electron and hole effective masses in gate-oxide region show opposite dependence on gate bias. Physical reasons for this behaviour are not yet known. Since the same expression of spatial and gate bias dependent collision time accurately simulates the direct tunneling gate current in both n- and p-MOSFETs, we believe that our empirical expression contains the essential physics of inelastic trap scattering. The contribution of split-off holes to direct tunneling gate current in p-MOSFETs, particularly at higher gate voltages, is not negligible. Consequently this contribution should not be neglected as done in some recent studies.

Chapter 1

Introduction

Advancement of semiconductor device technology is continuing the scaling down of Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) feature sizes. According to [1], for sub-100 nm devices, gate-oxide thickness will be around 2 nm. When such deep submicron devices are biased by applying a voltage to the gate-electrode, band bending occurs near oxide-semiconductor interface, that results in a quantum well. Therefore, energy quantization of inversion carriers occurs and the distribution of inversion carriers can no longer be modeled by semi-classical approaches [2], rather a complete quantum mechanical treatment is necessary. Also, the barrier height for inversion carriers is finite (3.1 eV for electrons and 4.8 eV for holes). Therefore, direct tunneling of carriers occurs from the Si inversion layer to the gate-electrode through the gate-oxide which results in a gate leakage current. To model such devices, accurate determination of device parameters such as gate leakage current, gate capacitance, oxide thickness, etc. is necessary by taking into account the quantization of inversion carriers.

1.1 Literature Review

It is known that Quantum Mechanical (QM) effects arise in MOS inversion layer when a steep quantum well is formed near the oxide-semiconductor interface due to band bending [3]. As the gate-oxide thickness reduces to around 2 nm for sub-100 nm MOS devices [1], a substantial direct tunneling current flows from the gate-electrode to the channel, leading to a gate leakage current that increases exponentially with decreasing oxide thickness [4]. This gate leakage current is a severe problem in device operation, especially in terms of standby power consumption. Therefore, a quantitative understanding of direct tunneling gate leakage current is necessary. A significant

amount of research has already been conducted to study direct tunneling gate current. A partial review of these researches is presented in this section.

Rana *et al.* [5] have solved Poisson's and Schrödinger's equations self-consistently for accumulation layers in MOS devices and have applied to the calculation of tunneling current for very thin oxides. They have solved Schrödinger's equation assuming wave function tail inside the gate-oxide and zero at the gate-electrode/SiO_2 interface and at a distance deep inside the semiconductor. In the calculation of tunneling current, they have taken into account both the bulk extended states and the quasi-bound states near the Si/SiO_2 interface. They have calculated the tunneling current in the direct regime and have obtained almost the same tunneling currents using both the semi-classical and self-consistent models. This agreement in current between the models has been explained in their work by showing a larger potential drop in the semiconductor with self-consistent solutions. Because larger drop in the semiconductor means smaller electric field inside the SiO_2 barrier which reduces the tunneling rates from both the extended and quasi-bound states. But electron energies of the quasi-bound states are higher in the self-consistent solutions, that increase the tunneling rates. Therefore, the current predicted by both the models has almost the same value.

Lo *et al.* [4] have studied the direct tunneling current from the quantized inversion layer of ultra-thin-oxide n-MOS devices. They have solved the coupled effective-mass Schrödinger's and Poisson's equations self-consistently and have calculated the lifetimes of quasi-bound states using a transverse-resonant method. They have found that more than 90% of the total current density comes from the lowest two eigenstates of twofold-degenerate and fourfold-degenerate valleys in the conduction band. It has been found that at a gate voltage of 1.5V, the current density increases by ten orders of magnitude as the oxide thickness decreases from 3.6 nm to 1.5 nm.

Register *et al.* [6] have developed an analytical model for direct tunneling gate current in MOS devices as a function of surface electric field using a modified WKB approximation both for n- and p-MOS transistors. They have properly modeled the field dependencies of sheet charge, electron impact frequency on the interface and mean tunneling probability that results in somewhat accurate modeling of the low-field roll-off in the current.

Lo, Buchanan and Taur have reviewed their work on direct tunneling gate current in Ref. [7]. They have developed a QM model and have studied the C-V and I-V characteristics of n- and p-MOS structures. They have calculated the eigenenergies using transverse-resonant method of wave guides and transmission probability using transfer matrix approach. According to their study, for a chip of standby power $\leq$ 0.1 W per chip, direct tunneling current can be tolerated down to an oxide thickness of 1.5-2 nm. The

transconductance reduction due to polysilicon depletion and finite inversion layer thickness effects becomes more severe for thinner oxides. The finite thickness effects degrade the gate capacitance by 13% or more for an oxide thickness of 2.5 nm or less. Their model predicts higher threshold voltage than the semi-classical model, and the difference increases with the electric field strength at the Si/SiO_2 interface.

Yang *et al.* [8] have studied direct tunneling gate current and C-V characteristics of MOS devices by solving silicon substrate inversion layer quantum mechanically and calculating transmission probability with a modified WKB approximation. They have also investigated the effects of substrate doping concentration and polysilicon doping concentration on direct tunneling current. In their study, the QM calculation in the substrate assumes an infinite barrier height at the Si/SiO_2 interface, while the transmission probability calculation assumes wave function continuity across a limited barrier height at this interface. Therefore, the whole procedure itself is not entirely self-consistent. They have found that at high bias conditions, the lowest subband can be used alone for current calculation with reasonable accuracy. Also, the substrate band bending from QM calculation is considerably larger than the semi-classical results at high bias. They have showed that direct tunneling current decreases as the substrate doping level increases for the same polysilicon doping level and gate voltage. Same thing is true for polysilicon doping level.

Mudanai *et al.* [9] have studied direct tunneling current through different gate dielectrics by solving Schrödinger's and Poisson's equations self-consistently. However, in the self-consistent loop, they have solved Schrödinger's equation assuming that the wave function goes to zero at dielectric-silicon interface. Once self-consistent potential has been obtained, they have again solved Schrödinger's equation assuming wave function penetration into the gate dielectric. They have calculated the real parts of eigenenergies of the quasi-bound states assuming that the wave function goes to zero at the oxide/gate-electrode interface. The lifetimes of the quasi-bound states have been calculated using first order perturbation approach. This technique requires the determination of complex eigenenergies of a non-Hermitian Hamiltonian. Therefore, this technique is numerically time consuming and inefficient. According to their simulated results, the current through the pure oxide dielectric is much higher than the current through the high-K gate stack structure. Also, the current through the high-K gate stack structure shows oscillations at high gate bias. But the oscillations were observed to decrease and eventually vanish as the barrier height was increased.

Yang *et al.* [10] have developed a model for the hole direct tunneling gate current accounting for heavy and light hole subbands in the quantized

inversion layer. They have extracted physical oxide thickness, t_{ox}, using QM-corrected C-V fitting based on van Dort's model. According to their study, the hole direct tunneling is responsible for source/drain current while the valence electron direct tunneling constitutes bulk current. They have also showed that the hole direct tunneling dominates over the valence band electron direct tunneling in gate voltage of 0V to a certain crossover point, and its magnitude increases with decreasing oxide thickness.

Cassan *et al.* [11] have investigated the charge control and gate leakage in MOS structures and heterojunction-MOS (H-MOS) structures with ultra-thin oxide using both semi-classical and wave-mechanical calculations. In the semi-classical approach, they have determined the direct tunneling gate current using the transmission probability formalism whereas the notion of quasibound state lifetime has been applied in the wave-mechanical model. They have found that threshold voltage significantly depends on the applied model for conventional MOS structures, but the gate leakage shows an excellent agreement between both approaches provided the threshold voltage correction is taken into account. For H-MOS, they have found that the gate leakage has been reduced by at least two orders of magnitude compared with the conventional MOS design.

Hou *et al.* [12] have reported the calculation of hole direct tunneling current from the inversion layer in a p-MOSFET. They have included the effect of valence band mixing on hole quantization by an improved one-band effective mass approximation [13]. A modified WKB approximation has been used to calculate the hole transmission probability with a more accurate dispersion equation in the oxide gap. Their simulated results with parabolic dispersion in the oxide show a good agreement with the experimental data for thinner oxide thickness and at lower gate voltages. However, at high gate voltage and thick oxide thickness, the simulated results show a significant deviation from the experimental data. They have used a dispersion form proposed by Freeman and Dahlke in the oxide and have obtained a significant improvement in the simulated results, that closely match with the experimental data for thicker oxide at high gate voltage.

Wang *et al.* [14] have developed a model for calculating direct tunneling gate current by solving Schrödinger's and Poisson's equations self-consistently in the whole metal-oxide-semiconductor structure. They have solved Schrödinger's equation using a modified Airy function technique. They have showed that in case of strong electric field (F_{ox} = 12MV/cm), the tunneling current for thicker oxides ($\geq$ 2.5 nm) does not decrease exponentially as oxide thickness increases, because it enters the Fowler-Nordheim tunneling regime. Also, electron concentration at the Si/SiO_2 interface is notably larger for self-consistent solution than that obtained from non self-consistent solution.

It is well known that due to the presence of impurities, defects, interface states and phonons, electrons tunneling through the oxide experience phase-breaking or inelastic scattering. Although phonon-assisted and inelastic trap-assisted tunneling have been considered to describe stress-induced leakage current (SILC) and oxide tifetimes [15, 16, 17, 18, 19], very little work has been done on the effects of inelastic scattering on the direct tunneling current. It may be mentioned that although the simulated direct tunneling currents in n-MOS devices show a good agreement with the measured data at high bias and thinner oxides, at low bias and thicker oxides, the experimental current is much higher than the simulated current [4, 6, 7, 8, 20]. Similar mismatch between measured and simulated direct tunneling gate currents has also been observed in p-MOSFETs [12]. No satisfactory explanation for such mismatch has so far been proposed.

Städele *et al.* [21] have studied the influence of elastic impurities, such as neutral oxygen vacancies, on the magnitude of tunneling current through ultra-thin gate-oxide of a prototypical metal-oxide field-effect transistor. Electron distribution has been determined from full band Monte-Carlo transport simulation and transmission coefficients from three-dimensional semi-emperical tight-binding calculations. The positions of the atoms in the junction have been determined by first-principles density-functional calculation. They have found that gate current increases significantly (typically one to three orders of magnitude) in the presence of impurities having a density around $10^{12}/cm^2$, provided that the resonant energy levels lie less than 1 eV above the Si conduction band edge.

Fu *et al.* [22] have studied carrier transmission through thin oxide layer in the presence of an ionized impurity in the oxide layer. The Coulumb potential of the ionized impurity provides extra conduction channels, that increases tunneling current. They have found that the ratio between the conduction current through the extra conduction channels and the direct tunneling current can be as large as 1.9. The extra value of the ratio depends on the location of the charge in the oxide layer.

When the substrate is heavily doped, a thermionic barrier is present that opposes the direct tunneling of gate electrons when the applied gate voltage is between 0 V and flatband voltage in the accumulation region. In such conditions, the measured gate current cannot be explained by direct tunneling alone but features an additional, dominant component [23]. Ghetti *et al.* [23] have investigated the low voltage tunneling in ultra-thin oxides. They have found that the extra component is very sensitive to temperature and is due to gate electrons tunneling into anode interface states. Their simulation model has been built on by solving Schrödinger's and Poisson's equations self-consistently taking into account the quantization effects of both elec-

trons and holes. The transmission probability has been computed through an exact solution of Schrödinger's equation in terms of Airy's function. They have showed that tunneling of gate electrons into anode interface states is very sensitive to the interface state distribution and to electrical stress.

Zaman [24] has investigated the effects of inelastic scattering on the direct tunneling gate leakage current for the first time. Although his work qualitatively explains the observed mismatch between the experimental and simulated direct tunneling currents, a quantitative agreement with experimental results could not be obtained in his study. He has found that inelastic scattering has pronounced effect on the direct tunneling current at lower gate voltages in the devices with oxide thickness equal to 2 nm or higher. When the oxide thickness is reduced below 2 nm, the inelastic scattering has no significant effect on the direct tunneling current.

1.2 Objective of the Work

The objective of this research is to calculate the direct tunneling gate leakage current in deep submicron MOSFETs for both n-MOS and p-MOS devices by solving Poisson's and Schrödinger's equations self-consistently in the presence of inelastic scattering. The commonly used boundary conditions to solve Schrödinger's equation are that the wave function goes to zero at the Si/SiO_2 interface and at a point deep inside the bulk Si. Therefore, the wave function tail inside the oxide is ignored, which is the cause of direct tunneling. We will solve Schrödinger's equation using Green's function formalism with open boundary conditions [25, 26]. Therefore, the wave function tail in the oxide will be calculated in our model. The eigenenergies will be calculated by locating the peaks of one-dimensional (1D) density-of-state(DOS) and the lifetimes of the inversion carriers will be calculated from the full-width-at-half-maximum(FWHM) of the energy broadened DOS. Therefore, our model is numerically efficient because it eliminates the need for solving non-Hermitian Hamiltonian matrices. The present model will shed light on the role of inelastic scattering processes on direct tunneling gate current. The inelastic scattering effects will be included in the model following the theory developed in [24]. We will extend the work of Ref. [24] by

(i) solving Schrödinger's and Poisson's equations self-consistently,

(ii) choosing more realistic expressions for collision time,

(iii) using improved models for effective masses of carriers within the gate-oxide region and

(iv) calculating gate currents in n-MOS as well as in p-MOS devices.

As a result, we expect to obtain a quantitative agreement with experimental results.

1.3 Thesis Layout

This thesis consists of four chapters. Chapter one gives an introduction followed by literature review and objective of the work.

Chapter two deals with the existing self-consistent technique along with a detailed discussion on the procedures followed to develop the proposed model. Also, the theory to include the effects of inelastic scattering is presented in this chapter.

The simulated results for both n-and p-MOSFETs are presented in chapter three. A comparison between the measured gate current and the simulated gate current is also presented in this chapter.

Conclusive remarks and discussions are presented in chapter four. A recommendation for future study is also suggested here.

Chapter 2

Quantum Mechanical Calculation of Silicon Inversion Layer

Quantization of silicon inversion carriers in MOSFETs occurs when the electric field, caused by the gate bias, confines the carriers in a narrow region near oxide-semiconductor interface. The electric field causes band bending and a potential well is formed near the Si/SiO_2 interface. The quantum mechanical calculation of inversion carriers requires the solution of Poisson's and Schrödinger's equations self-consistently. Since the potential varies along the z direction only (see Fig. 2.1) and does not vary along x and y directions, the wave functions can be written as free particle wave functions along x and y directions. Therefore, one-dimensional (1D) solutions of Poisson's and Schrödinger's equations are required. This chapter presents the existing self-consistent technique, as well as the calculation approaches of our model. A detailed discussion on the calculation of direct tunneling gate leakage current in the presence of inelastic scattering is also presented in this chapter.

2.1 Existing Self-Consistent Technique

The numerical self-consistent calculation for Si inversion layer was first carried out by Stern [3]. Self-consistent calculation made by Stern [3] was based on three major approximations:

(i) The effective-mass approximation is valid, so that periodic potential can be neglected and the effective mass and dielectric constant of the perfect crystal can be used.

(ii) The envelope wave function vanishes at the Si/SiO_2 interface.

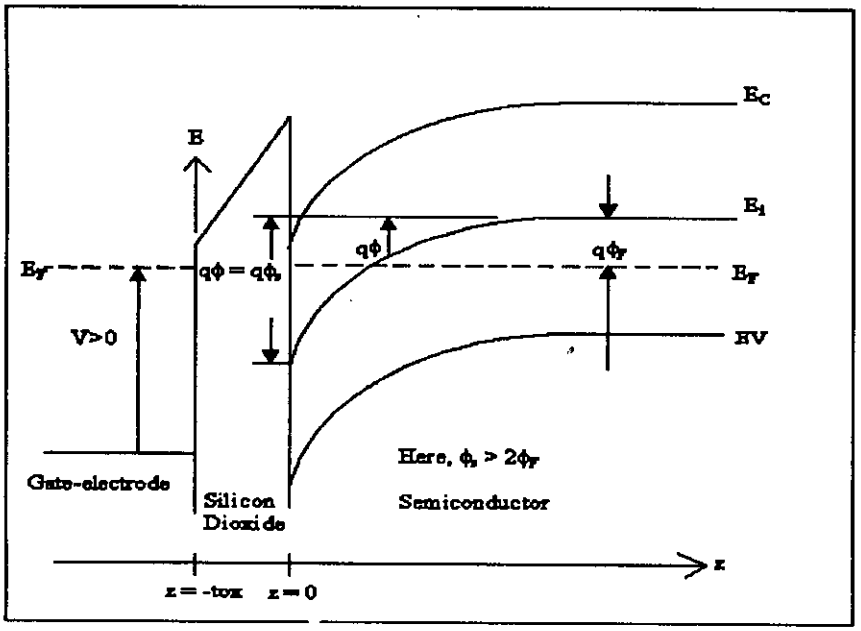


Figure 2.1: A typical band diagram of an n-channel MOSFET under strong inversion condition.

(iii) Surface states are neglected and the effect of any charges in the oxide adjacent to the semiconductor is replaced by an equivalent electric field.

The band bending, $q\Phi$, as shown in Fig. 2.1, at a semiconductor surface can be characterized by an electrostatic potential $\Phi(z)$. Within the effective-mass approximation, the electronic wave function for the jth subband in the ith valley is the product of the Bloch function at the bottom of the conduction band and an envelope function

$$\psi_{ij}(x, y, z) = \zeta_{ij}(z) e^{\theta z} e^{i(k_1 x + k_2 y)}, \tag{2.1}$$

where k_1 and k_2 are measured relative to the band edge, θ depends on k_1 and k_2, and $\zeta_{ij}(z)$ is the solution of

$$\frac{d^2\zeta_{ij}(z)}{dz^2} + \frac{2m_3}{\hbar^2}\left[E_{ij} + q\Phi(z)\right]\zeta_{ij}(z) = 0. \tag{2.2}$$

Here, m_3 is the carrier effective mass in the direction perpendicular to the surface, E_{ij} is the eigenenergy of the jth subband in the ith valley in the same direction and q is the charge of electron. Boundary conditions used

by [3] for the solution of Eq. (2.2) are $\zeta_{ij}(\infty) = 0$ and $\zeta_{ij} = 0$ at the oxide-semiconductor interface (z=0). Each eigenvalue, E_{ij}, found from the solution of Eq. (2.2) is the bottom of a continuum of levels called the subband, with energy levels given by

$$E_{ij}(\vec{k}) = E_{ij} + \frac{\hbar^2 k_1^2}{2m_1} + \frac{\hbar^2 k_2^2}{2m_2}, \tag{2.3}$$

where m_1 and m_2 are the principal effective masses for motion parallel to the surface, which can be obtained in a straightforward way from the bulk masses [3]. There can be as many as three values of m_3 for a given surface orientation because the conduction band of Si has six valleys along the $\langle 100 \rangle$ directions of the Brillouin zone. Under the effective-mass approximation, the valleys are degenerate in pairs.

The potential $\Phi(z)$ used in Eq. (2.2) is the solution of Poisson's equation

$$\frac{d^2\Phi(z)}{dz^2} = -\frac{\rho(z)}{\epsilon_o \epsilon(z)}, \tag{2.4}$$

where

$$\rho(z) = \rho_{depl}(z) + q \sum_{ij} N_{ij} \mid \zeta_{ij}(z) \mid^2, \tag{2.5}$$

ϵ_o is the permittivity of free space and $\epsilon(z)$ is the dielectric constant in different regions (gate-electrode, oxide and semiconductor). N_{ij}, used in the calculation of $\rho(z)$, is the carrier concentration in the jth subband of the ith valley given by

$$N_{ij} = n_{vi} m_{di} \frac{KT}{\pi\hbar^2} \ln\left[1 + \exp\left(\frac{E_F - E_{ij}}{KT}\right)\right], \tag{2.6}$$

where n_{vi} and m_{di} are the valley degeneracy and density-of-states effective mass per valley and E_F is the Fermi energy. $\rho_{depl}(z)$ is the charge density of the depletion layer given by

$$\rho_{depl}(z) = -q(N_A - N_D), \quad 0 < z < z_d \tag{2.7}$$

$$\rho_{depl}(z) = 0, \qquad z > z_d \tag{2.8}$$

where z_d is the width of the depletion layer given by

$$z_d = \sqrt{\frac{2\epsilon_{si}\epsilon_o \Phi_d}{q(N_A - N_D)}}. \tag{2.9}$$

Here, ϵ_{si} is the dielectric constant of Si and Φ_d is the effective band bending from the bulk to the surface, apart from the contribution of the inversion layer itself; its value is given by [3]

$$\Phi_d = \Phi_s - \frac{KT}{q} - \frac{qN_{inv}z_{av}}{\epsilon_{si}\epsilon_o}, \tag{2.10}$$

where Φ_s is the total band bending from the bulk to the surface, $N_{inv} = \sum_{ij} N_{ij}$ is the inversion carrier density per unit area and

$$z_{av} = \frac{\sum_{ij} N_{ij} \int z \mid \zeta_{ij}(z) \mid^2 dz}{\sum_{ij} N_{ij} \int \mid \zeta_{ij}(z) \mid^2 dz} \tag{2.11}$$

is the average penetration of the inversion carriers into the semiconductor from the oxide-semiconductor interface. The boundary conditions used to obtain the solution of Eq. (2.4) are

$$\frac{d\Phi}{dz} = 0 \quad \text{for} \quad z = \infty$$
$$\frac{d\Phi}{dz} = -F_s \quad \text{at} \quad z = 0,$$

where

$$F_s = \frac{q(N_{inv} + N_{depl})}{\epsilon_o \epsilon_{si}} \tag{2.12}$$

and

$$N_{depl} = z_d(N_A - N_D) \tag{2.13}$$

is the carrier density per unit area in the depletion layer. In the self-consistent calculation made by Stern [3], Eqs. (2.2)-(2.13) are solved iteratively for a given value of N_{inv} until the profile converges within a specified limit. The calculation is started with a small value of N_{inv} and then increased gradually to larger values, taking the results of each case to construct the starting potential for the next.

2.2 Green's Function Formalism

According to [26], Green's function can be used to calculate the wave function and DOS. $Z(z, z_1; E)$, the logarithmic derivative of retarded Green's function, G^R, is defined by

$$Z(z, z_1; E) = \frac{2\hbar}{jm_3}\left[\frac{\partial G_R(z, z_1; E)}{\partial z} / G^R(z, z_1; E)\right]. \tag{2.14}$$

Here, $G^R(z, z_1; E)$ is the solution of the equation

$$\left(E - q\Phi(z) + j\eta + \frac{\hbar^2}{2m_3}\frac{\partial^2}{\partial z^2}\right) G^R(z, z_1; E) = \delta(z - z_1), \tag{2.15}$$

where $\delta(z - z_1)$ is a unit excitation applied at $z = z_1$ and η is an infinitesimal positive energy. G^R has two important properties:

(i) it is continuous everywhere and

(ii) its first derivative is discontinuous at $z = z_1$ by $2m_3/\hbar^2$.

Therefore, $Z(z, z_1; E)$ has a discontinuity at $z = z_1$ and we need two boundary conditions to calculate $Z(z, z_1; E)$. If we assume that the potential profile is flat sufficiently far from the oxide-semiconductor interface in both directions, the boundary conditions (called asymptotic boundary conditions) are [25]

$$Z(z \to \infty, z_1; E) = Z_o(\infty) \quad z > z_1 \tag{2.16}$$

$$Z(z \to -\infty, z_1; E) = -Z_o(-\infty) \quad z < z_1, \tag{2.17}$$

where

$$Z_o(z) = \frac{2\hbar\gamma(z)}{jm_3} \tag{2.18}$$

and

$$\gamma(z) = j\sqrt{\frac{2m_3}{\hbar^2}(E - q\Phi(z) + j\eta)}. \tag{2.19}$$

The use of two boundary conditions reveals an interesting property of the function $Z(z, z_1; E)$;

$$Z(z, z_1; E) = Z_{ir}(z; E) \quad \text{for all } z > z_1, \tag{2.20}$$

$$Z(z, z_1; E) = Z_{il}(z; E) \quad \text{for all } z < z_1. \tag{2.21}$$

Here,

$$Z_i = Z_o\frac{Z_l\cosh(\gamma l) - Z_o\sinh(\gamma l)}{Z_o\cosh(\gamma l) - Z_l\sinh(\gamma l)}, \tag{2.22}$$

and the subscripts ir and il refer to the quantum-mechanical impedances looking to the right and left, respectively. The method for the calculation of Z_{ir} and Z_{il} has been discussed elsewhere [27]. It has been shown in [28] that

$$G^R(z, z; E) = \frac{4}{j\hbar}\frac{1}{Z_{ir}(z; E) - Z_{il}(z; E)}. \tag{2.23}$$

Now the DOS and wave function can be calculated using G^R. The local DOS is defined by the diagonal elements of G^R as

$$N(z; E) = -\frac{1}{\pi}\Im m\left[G^R(z, z; E)\right]. \tag{2.24}$$

Taking the imaginary parts of Eq. (2.23) on both sides and putting that value in Eq. (2.24) one obtains

$$N(z;E) = \frac{4}{\pi\hbar}\Im m\left[\frac{j}{Z_{ir}(z;E) - Z_{il}(z;E)}\right]. \tag{2.25}$$

It is well known that G^R can be expressed in terms of a complete set of eigenfunctions:

$$G^R(z,z_1;E) = \sum_i \frac{\zeta_i(z)\zeta_i^*(z_1)}{E - E_i + j\eta}. \tag{2.26}$$

If $E_{i+1} - E_i >> \eta$ for all values of i, only one term in the series dominates when $E \rightarrow E_n$, since the discrete eigenenergies in one-dimension are non-degenerate. For the diagonal elements of G^R, we obtain

$$G^R(z,z;E \rightarrow E_n) = \frac{\mid \zeta_n(z) \mid^2}{E - E_n + j\eta}. \tag{2.27}$$

Taking the imaginary parts of both sides of Eq. (2.27), substituting $E = E_n$ and putting the value of $G^R(z,z;E)$ from Eq. (2.23) we obtain for the wave function

$$\mid \zeta_n(z) \mid^2 = \frac{4\eta}{\hbar}\Im m\left[\frac{j}{Z_{ir}(z;E_n) - Z_{il}(z;E_n)}\right]. \tag{2.28}$$

2.3 Calculation Approaches Followed to Develop The Model

The self-consistent model is developed by solving Eqs. (2.2)-(2.13), as discussed in Section 2.1, iteratively. In our model, eigenenergies and wave functions are calculated in a different way than the conventional one [29]. Using the fact that at eigenenergies, DOS is maximum, we calculate the eigenenergies by locating the peaks of DOS, starting from the bottom of the conduction band (n-MOS) and peak of valence band (p-MOS) at a position within the quantum well (2.5 nm deep inside the semiconductor from the oxide-semiconductor interface). For split-off band of holes, 44 meV is added to the calculated values to obtain the correct eigenenergies [29]. Once the eigenenergies are found, the wave functions are calculated using Eq.(2.28). In the development of Eq.(2.28), asymptotic boundary conditions are used [25]. Therefore, our model includes the wave function tail in the oxide region, which is a crucial factor to calculate the direct tunneling gate leakage current. In our model, Poisson's equation is also solved considering the fraction of the inversion charges within the gate-oxide region and the Fermi energy is

calculated using Eq. (2.6) taking into account the number of inversion carriers residing within the gate-oxide region.

The calculation is started with a small value of N_{inv}. Then Poisson's equation is solved and eigenenergies, wave functions, Fermi energy and inversion charge density are calculated. However, to solve Poisson's equation we need ρ_{inv} and F_s, where Φ_d is needed to calculate F_s. For the starting value of N_{inv}, to find the initial potential profile, we assume $\rho_{inv} = 0$ and solve Poisson's equation analytically. The effective band bending, Φ_d, is required to calculate the depletion layer width, z_d. But Φ_d can be calculated if the potential profile and z_{av} are known. For every starting value of N_{inv} we calculate Φ_d using a classical formula given by [2]

$$Q_{inv} = \sqrt{2q\epsilon_o\epsilon_{si}N_A}\left(\sqrt{\Phi_d + V_T e^{(\Phi_d-\Phi_F)/V_T}} - \sqrt{\Phi_d}\right), \tag{2.29}$$

where

$$\Phi_F = V_T \ln(\frac{N_A}{n_i}), \tag{2.30}$$

V_T is the thermal potential and Q_{inv} is the inversion charge density per unit area. Once Poisson's and Schrödinger's equations are solved for the starting value of N_{inv}, the program solves these two equations self-consistently by increasing the values of N_{inv} in steps up to the highest specified value of N_{inv}.

As N_{inv} is increased in steps, the profile obtained for the previous value of N_{inv} is taken as the starting profile for the next N_{inv}. The convergence is checked on Fermi energy and first eigenenergy. Also, a minimum number of iterations is set, so that the program can not get out of the self-consistent loop before that number of iterations. This is done for increasing reliability on the convergence criteria.

2.4 Calculation of Direct Tunneling Gate Leakage Current

It is well known that for an isolated system, the Hamiltonian matrix **H** is Hermitian because the wave function is zero at the boundaries and the system has bound states only. The eigenenergy values are then real. But for an open system, the Hamiltonian matrix **H** is no longer Hermitian and the system has quasi-bound states with complex eigenenergy values. The real part of the complex eigenenergy gives the energy of resonance and the imaginary

part, Γ_{ij}, is related to the carrier lifetime, τ_{ij}, of the jth eigenstate in the ith valley as [4],

$$\tau_{ij} = \frac{\hbar}{2\Gamma_{ij}}. \tag{2.31}$$

Therefore, carrier lifetime becomes finite, and DOS broadens around the eigenenergy rather than becoming a delta function at that energy. In our model, the real part of complex eigenenergy is calculated by the technique discussed in Section 2.3, and the imaginary part is calculated by measuring the FWHM of the energy broadened DOS.

The gate leakage current due to quantum mechanical tunneling of carriers from the quasi-bound states of Si inversion layer to the gate-electrode through the gate-oxide can be calculated from the carrier concentration and lifetime of carriers at all the eigenstates using the following relationship [4]

$$J_l = \sum_{ij} \frac{qN_{ij}}{\tau_{ij}}. \tag{2.32}$$

Here, J_l is the coherent gate current. Gate voltage, V_g, is calculated using the following relation

$$V_g = d_m F_s + \Phi_{ms} + t_{ox} F_{ox} + \Phi_s, \tag{2.33}$$

where d_m is the depletion layer width of polysilicon (used as gate-electrode), t_{ox} is the gate-oxide thickness, $F_{ox} = \epsilon_{si} F_s / \epsilon_{ox}$ is the electric field in the oxide and Φ_{ms} is the work function difference between the gate electrode and semiconductor.

The attractive feature of our model is the inclusion of inelastic scattering in the gate-oxide region. We assume that inelastic trap scattering is the dominant inelastic scattering mechanism in the gate-oxide region [30]. Inelastic scattering in the oxide can be represented by a position and bias dependent collision time $\tau_{coll}(z; F_{ox})$, where the position and bias dependence is determined by the nature and the distribution of the traps present in the gate-oxide region. In Schrödinger's equation, inelastic scattering can be included by an imaginary potential term jV_I, where

$$V_I = \frac{\hbar}{2\tau_{coll}(z; F_{ox})}. \tag{2.34}$$

It is known that inelastic scattering also causes broadening of DOS around eigenenergies [26]. The total width of the broadened DOS is equal to the

partial widths for the jth state in the ith valley as [31]

$$\Gamma_{tij} = \Gamma_{lij} + \Gamma_{sij}, \tag{2.35}$$

where Γ_t is total broadening, Γ_l is broadening due to coherent leakage only and Γ_s is broadening due to inelastic scattering only.

Therefore, electrons are lost from the Si inversion layer due to inelastic scattering at the rate of $\sum_{ij} qN_{ij}/\tau_{sij}$ [24] where

$$\tau_{sij} = \frac{\hbar}{2\Gamma_{sij}}. \tag{2.36}$$

If the lost electrons from the inversion layer face an inelastic scatterer during their journey inside the oxide, a fraction of the scattered electrons will travel in the same direction but lose their phase memory. These electrons are incoherent with respect to those that did not undergo any inelastic scattering. However, the incoherent electrons will again accumulate equal phases and will be in phase with one another as long as they do not undergo an additional scattering. It is shown in [28] that all the electrons moving in a particular direction at any z, travel with the same group velocity [$v_g^+(z; E_{ij})$ or $v_g^-(z; E_{ij})$] regardless of the location of last inelastic collision, where

$$v_g^+(z; E_{ij}) = \frac{Re[Z_{ir}(z; E_{ij})]}{2}, \tag{2.37}$$

$$v_g^-(z; E_{ij}) = \frac{Re[Z_{il}(z; E_{ij})]}{2}, \tag{2.38}$$

and $v_g^+(z; E_{ij}) > 0$ and $v_g^-(z; E_{ij}) < 0$. Here, $+(-)$ denotes positive(negative) moving electron. Now the current due to inelastic scattering only, at energy E_{ij}, can be decomposed into two oppositely flowing currents as follows:

$$J_{sij}(z; E_{ij}) = J_{sij}^+(z; E_{ij}) + J_{sij}^-(z; E_{ij}). \tag{2.39}$$

Assuming that current from any eigenenegy is not coupled to that of another eigenenergy via inelastic scattering, the left hand side of Eq. (2.39) becomes position independent from current conservation requirement. The current continuity equations for $J_{sij}^{\pm}$ can be written as [28]

$$\frac{dJ_{sij}^+(z; E_{ij})}{dz} = -\alpha^+(z; E_{ij}) J_{sij}^+(z; E_{ij}) +$$
$$r^+(z; E_{ij}) \left[\alpha^+(z; E_{ij}) J_{sij}^+(z; E_{ij}) + \alpha^-(z; E_{ij}) J_{sij}^-(z; E_{ij})\right], \tag{2.40}$$

$$\frac{dJ^{-}_{sij}(z;E_{ij})}{dz} = -\alpha^{-}(z;E_{ij})J^{-}_{sij}(z;E_{ij}) + r^{-}(z;E_{ij})\left[\alpha^{+}(z;E_{ij})J^{+}_{sij}(z;E_{ij}) + \alpha^{-}(z;E_{ij})J^{-}_{sij}(z;E_{ij})\right], \quad (2.41)$$

where $\alpha^{\pm}(z;E_{ij}) = 1/\left(\tau_{coll}v^{\pm}_{g}(z;E_{ij})\right)$ describes the rate of inelastic scattering events per unit length and

$$r^{\pm}(z;E_{ij}) = \frac{\left[|v^{\pm}_{g}(z;E_{ij})|\right]}{\left[|v^{+}_{g}(z;E_{ij})| + |v^{-}_{g}(z;E_{ij})|\right]} \quad (2.42)$$

represents the fraction of the scattered electrons that travel to the right and to the left, respectively. The boundary conditions needed to solve Eq. (2.40) and Eq. (2.41) are [28]

$$J^{+}_{sij}(-t_{ox};E_{ij}) = 0 \text{ and}$$
$$J^{-}_{sij}(0;E_{ij}) = \frac{qN_{ij}}{\tau_{sij}}.$$

Here, $J^{+}_{sij}(-t_{ox};E_{ij}) = 0$ because electrons are injected only from the oxide-semiconductor interface. Using these relations and defining the following quantities:

$$\beta(z;E_{ij}) = \frac{1}{\tau_{coll}}\left[\frac{1}{v^{+}_{g}(z;E_{ij})} + \frac{1}{v^{-}_{g}(z;E_{ij})}\right], \quad (2.43)$$

$$A(z,0;E_{ij}) = e^{-\int_{z}^{0}\beta(z;E_{ij})dz}, \quad (2.44)$$

$$\xi(z;E_{ij}) = -\frac{1}{\tau_{coll}}\left[\frac{v^{+}_{g}(z;E_{ij})}{\left[v^{+}_{g}(z;E_{ij}) - v^{-}_{g}(z;E_{ij})\right]v^{-}_{g}(z;E_{ij})}\right], \quad (2.45)$$

$$B(-t_{ox},0;E_{ij}) = \int_{-t_{ox}}^{0}\xi(z;E_{ij})A(z,0;E_{ij})dz, \quad (2.46)$$

we finally get the following expression [28]

$$J_{sij}(E_{ij}) = \frac{J^{+}_{sij}(-t_{ox};E_{ij})A(-t_{ox},0;E_{ij}) + J^{-}_{sij}(0;E_{ij})}{1 + B(-t_{ox},0;E_{ij})}. \quad (2.47)$$

Now putting the value of $J^{+}_{sij}(-t_{ox};E_{ij})$ and $J^{-}_{sij}(0;E_{ij})$ from the boundary conditions we get

$$|\, J_{sij}(E_{ij})\,| = \frac{qN_{ij}/\tau_{sij}}{1 + B(-t_{ox},0;E_{ij})} \quad (2.48)$$

and the total current associated with inelastic scattering

$$J_s = \sum_{ij} \mid J_{sij}(E_{ij}) \mid . \tag{2.49}$$

Here, B in Eq. (2.48) is a dimensionless quantity, that represents the effects of backscattering on J_s. The total gate leakage current, J, due to direct tunneling in the presence of inelastic scattering in the oxide is

$$J = J_l + J_s \tag{2.50}$$

where J_s indicates the effect of inelastic scattering on direct tunneling current.

Chapter 3

Results and Discussions

Simulated results both for n-MOS and p-MOS devices, based on the formalism discussed in the previous chapter, are presented in this chapter. The validity of the model is checked by comparing the simulated results with other published simulated results and experimental data. The simulation program has been developed using MATLAB software and the program flow-chart is given in Appendix A. Fig. 2.1 shows the schematic band diagram of the device under our consideration.

3.1 Simulated Results for n-MOS Devices

The self-consistent analysis has been done for $\{100\}$ Si surface at room temperature. Si has six valleys in the conduction band. For $\{100\}$ Si surface, it has two-fold (longitudinal or lower valleys) and four-fold (transverse or higher valleys) degeneracy. The quantization effective masses and the density-of-states effective masses for electrons used in the model are shown in Table 3.1. The barrier height due to conduction band discontinuity of Si and SiO_2 has been taken to be 3.1 eV and the gate-electrode is n-type polysilicon with doping concentration of $N_{poly} = 1 \times 10^{20}/cm^3$. The substrate doping concentration is taken to be $N_A = 5 \times 10^{17}/cm^3$.

Surface	$\{100\}$	
Valleys	Longitudinal	Transverse
Degeneracy n_v	2	4
Quantization effective mass m_3/m_o	0.916	0.190
Density-of-states effective mass m_d/m_o	0.190	0.417

Table 3.1: Effective masses of electrons in different valleys.

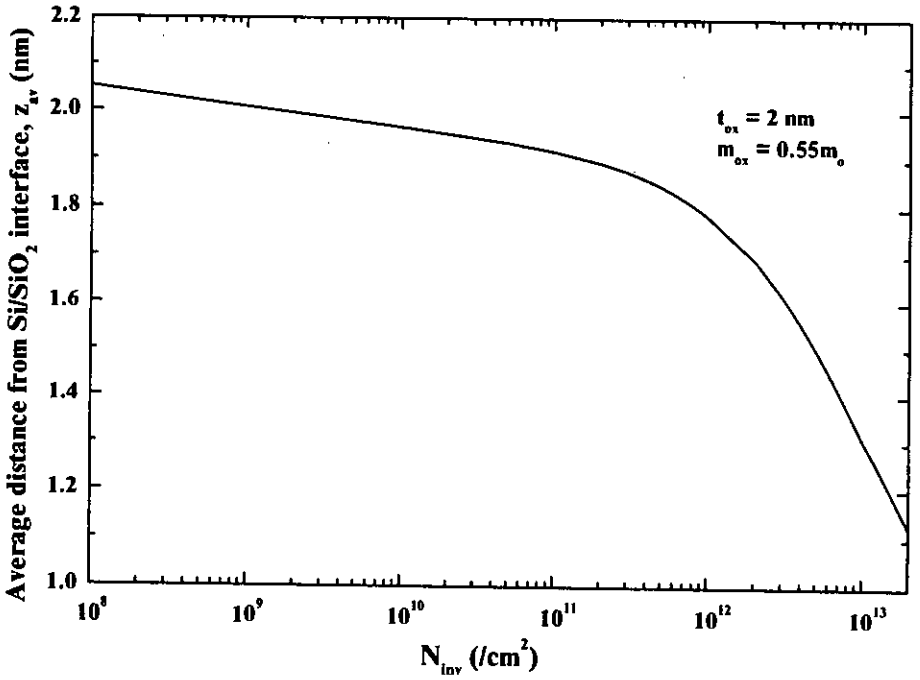


Figure 3.1: The average distance of the inversion electrons from Si/SiO_2 interface, z_{av}, as a function of inversion electron density, N_{inv}. Oxide thickness is $t_{ox} = 2$ nm and electron effective mass in oxide is $m_{ox} = 0.55m_o$.

The average distance of the inversion electrons from Si/SiO_2 interface, z_{av}, as a function of inversion electron density, N_{inv}, is shown in Fig. 3.1. The result closely matches with the simulated result reported in [32]. In Fig. 3.1, we see that the value of z_{av} decreases slowly at lower values of N_{inv} but rapidly at higher values. We can say that the centroid of inversion carrier distribution has the tendency to move closer toward Si/SiO_2 interface as the gate voltage or inversion carrier density increases. The substrate band bending, Φ_s, calculated semi-classically and quantum-mechanically are shown in Fig. 3.2 as a function of surface electric field, F_s. The difference between the two calculations is negligible in the sub-threshold region (lower F_s), but in strong inversion, Φ_s, calculated quantum-mechanically is significantly higher than the semi-classical results. This indicates the importance of extra band bending due to QM effects, which is significant in devices with ultra-thin gate-oxide. Similar trend of band bending is also reported in [8]. Band bending due to inversion charge only, Φ_i, is shown in Fig. 3.3. We observe that Φ_i has very small value at small N_{inv} but it increases rapidly after a certain value of N_{inv}, which indicates the achievement of strong inversion. A non-zero value of Φ_i at strong inversion is due to the shift of the

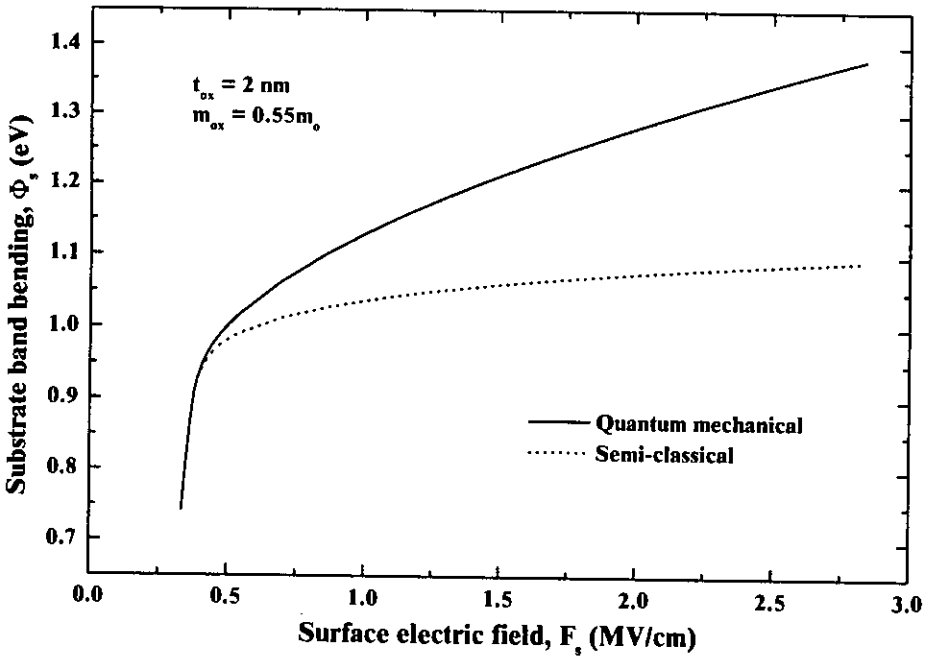


Figure 3.2: The substrate band bending, Φ_s, as a function of surface electric field, F_s. Oxide thickness is $t_{ox} = 2$ nm and electron effective mass in oxide is $m_{ox} = 0.55m_o$.

centroid of average carriers from the Si/SiO_2 interface. It is Φ_i that causes the difference between the values in Φ_s, when calculated semi-classically and quantum-mechanically. Fig. 3.4 shows the variation of Fermi energy level and the three lowest subband energy levels of inversion electrons as a function of surface electric field, F_s. Here, the energies are measured from the bottom of the quantum well at Si/SiO_2 interface. From the figure, we observe that the energies increase with increasing F_s, and at a certain value of F_s, the Fermi energy crosses the first eigenenergy of the longitudinal valleys. This is due to the lower value of valley digeneracy of longitudinal valleys. Similar trend is also observed in the numerical results reported in [8]. Fig. 3.5 shows the relative occupation factor, N_i/N_{inv}, for the lowest four eigen states. The occupation factor of the first eigen state of the longitudinal valley and the first eigen state of the transverse valley shows reverse trend with variation in F_s. Similar trend is observed in Fig. 3.4, where the difference between Fermi energy and the energy of first eigen state of the longitudinal valley decreases rapidly and that between Fermi energy and the energy of first eigen state of the transverse valley increases rapidly up to 1.1 MV/cm at which Fermi energy crosses the energy of the first eigen state of longitudinal valley. How-

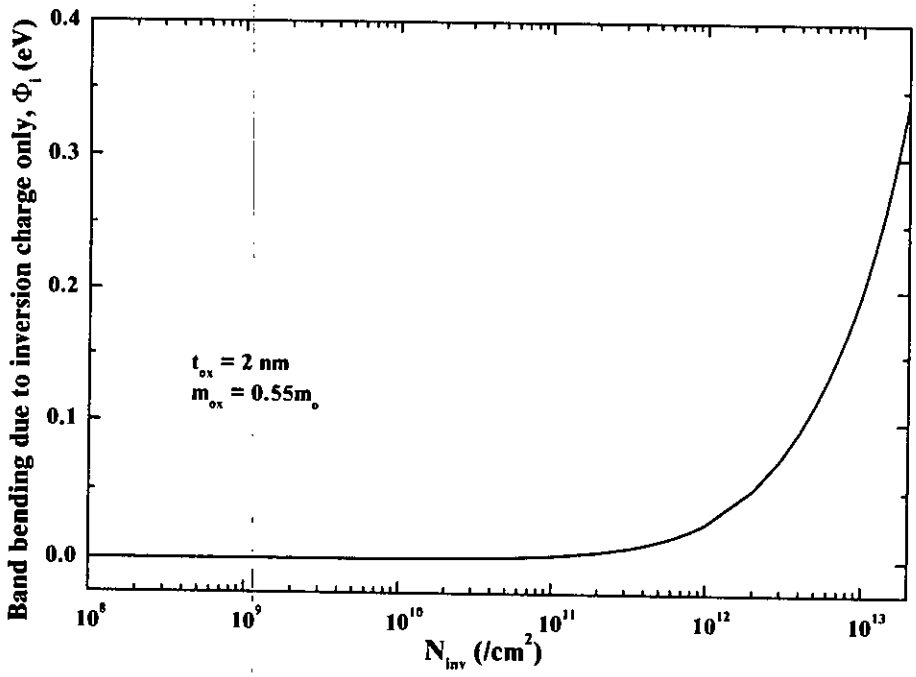


Figure 3.3: Band bending due to inversion charge only, Φ_i, as a function of inversion electron density N_{inv}. Oxide thickness is $t_{ox} = 2$ nm and electron effective mass in oxide is $m_{ox} = 0.55m_o$.

ever, total occupation of first eigen states of the longitudinal and transverse valleys together increases with surface electric field and this factor is over 95% at a field higher than 0.41 MV/cm. Therefore, we can conclude that only the first eigen states of the longitudinal and transverse valleys can be used at high bias condition with reasonable accuracy. A good agreement of our numerical calculations presented in Figs. 3.1-3.5 with published simulated results [8] demonstrates the accuracy of our numerical model.

Now we present results of our calculation related to estimation of the direct tunneling current. Similar qualitative results have been presented earlier in Ref. [24]. However, our calculations are different from Ref. [24] because we have

(i) solved Schrödinger's and Poisson's equations self-consistently,

(ii) chosen more realistic expressions for collision time,

(iii) used improved models for effective masses of carriers within the gate-oxide region and

(iv) calculated the direct tunneling gate currents in n-MOS as well as in p-MOS devices.

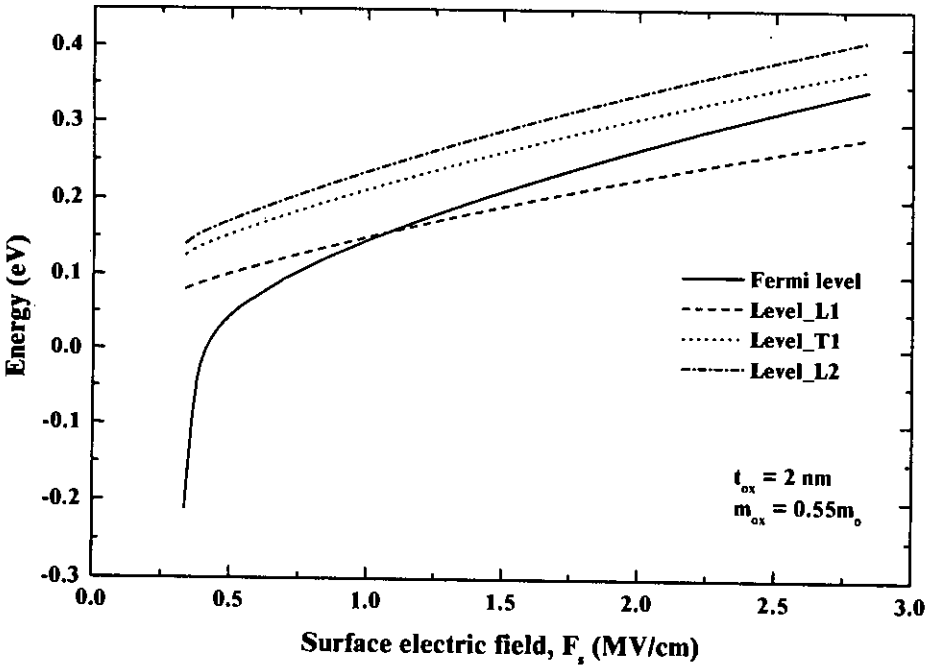


Figure 3.4: Fermi energy level and three lowest subband energy levels of inversion electrons as a function of surface electric field, F_s. Here, L1 and L2 are the first and second eigenstates of longitudinal valley and T1 is the first eigenstate of transverse valley. Oxide thickness is $t_{ox} = 2$ nm and electron effective mass in oxide is $m_{ox} = 0.55m_o$.

Fig. 3.6 shows the calculated lifetimes, τ_n, of inversion electrons as a function of eigenstate number, n, for a given surface electric field, F_s and four different values of t_{ox}. Experimental direct tunneling currents for these values of oxide thickness have been reported in Refs. [7] and [8]. No inelastic scattering is considered in Fig. 3.6. Fig. 3.6(a) is for the longitudinal valley and Fig. 3.6(b) is for the transverse valley. It is seen that τ_n exhibits rather unusual behavior with n. It does not decrease monotonically, rather starts increasing for the higher eigenstates. An explanation for this phenomena is provided in Ref. [33]. Also, lifetimes are found to depend almost exponentially on oxide thickness. We find that electrons in the transverse valley, due to their smaller quantization effective mass [3], have shorter lifetimes. Fig. 3.7 is the plot of τ_n for the lowest four states as a function of F_s for $t_{ox} = 2$ nm. The lifetimes decrease nearly exponentially with increasing F_s. This is because of the fact that with increasing surface electric field, the oxide barrier becomes more triangular and narrower at higher energies (Fig. 2.1) and the tunneling transparency of the barrier increases.

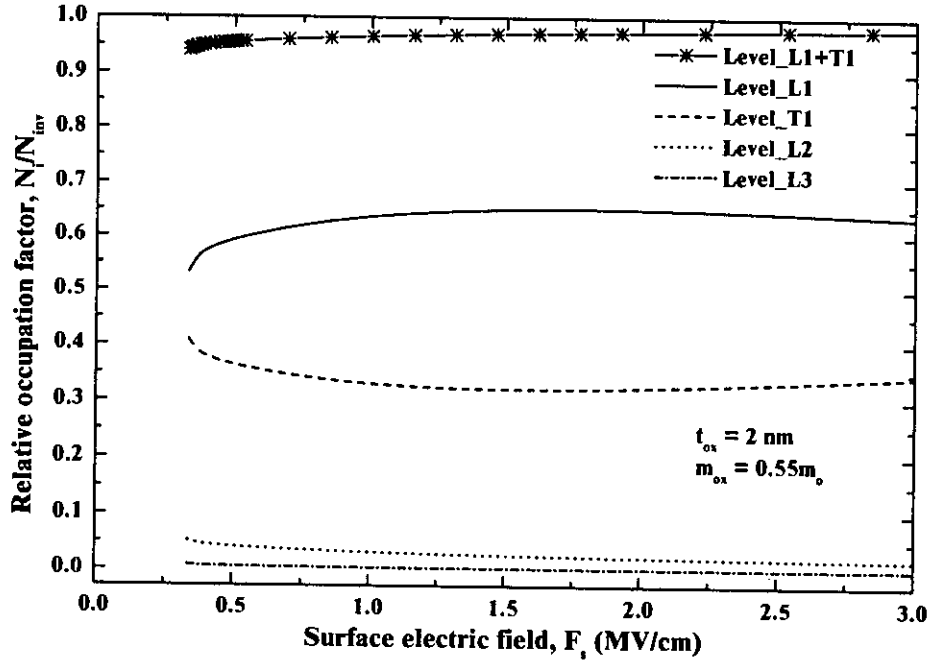


Figure 3.5: Relative occupation factor of inversion electrons, N_i/N_{inv}, for the lowest four subbands as a function of surface electric field, F_s. Here, L1, L2 and L3 are the lowest three eigenstates of longitudinal valley and T1 is the first eigenstate of transverse valley. Oxide thickness is t_{ox} = 2 nm and electron effective mass in oxide is $m_{ox} = 0.55m_o$.

Next, the effects of inelastic scattering on direct tunneling are studied. We first numerically investigate the validity of Eq. (2.35), on which our model is based. We calculate Γ_{ln} for a structure with a finite t_{ox} that contains no inelastic scattering, Γ_{sn} for the same F_s is obtained from a device in which inelastic scattering is present, but $t_{ox} \rightarrow \infty$, so that there is no phase coherent direct tunneling. Γ_{tn} is determined from the actual device conditions, *i.e.*, t_{ox} is finite and the oxide region contains inelastic scattering processes. Tables 3.2 and 3.3 present the results for three lowest quasi-bound states for t_{ox} = 2 nm. In these calculations, we have used two different values of F_s (0.49 MV/cm and 1.62 MV/cm) and a constant value of τ_{coll} with no spatial or bias dependence (10^{-6} sec). Other values and distributions for τ_{coll} has also been used (for example, Eq. (3.3)), and we have verified that the qualitative nature of our conclusions does not depend on the actual distribution of τ_{coll}. For all the states (results for higher states not presented), Eq. (2.35) is found to hold with negligible error. We have also numerically checked that Eq. (2.35) holds for all the devices under our consideration. It

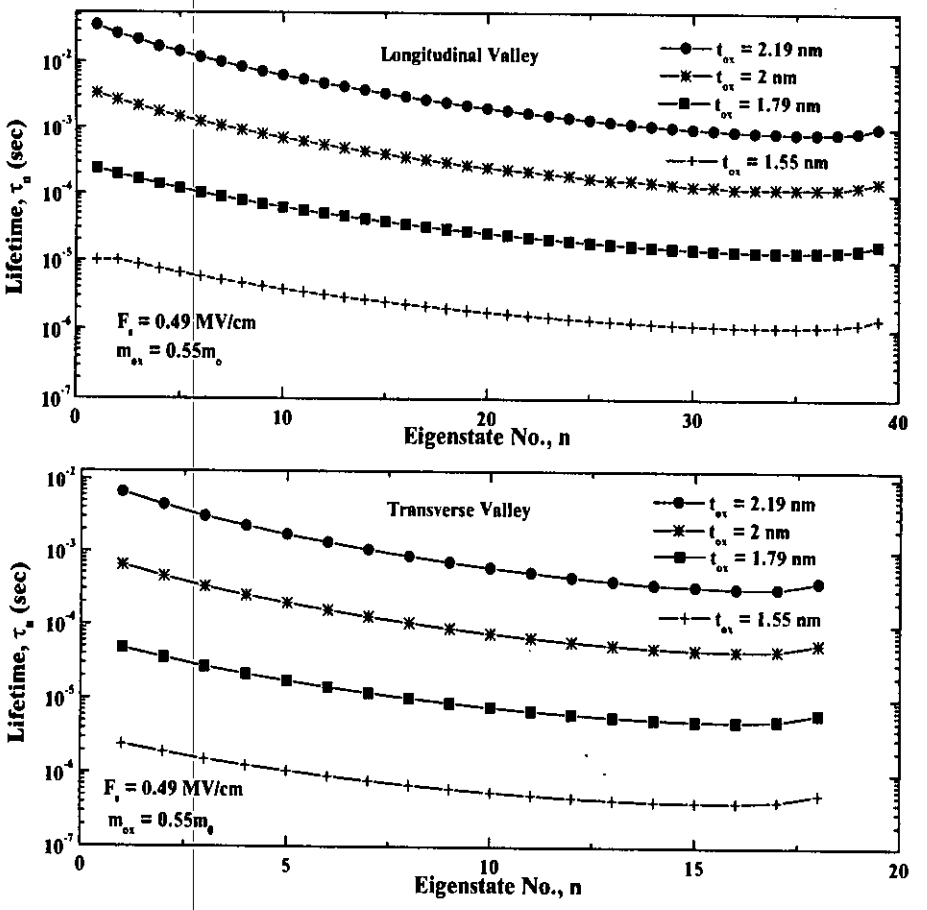


Figure 3.6: Variation of lifetimes, τ_n, with eigenstate number, n, for a given surface electric field, F_s. Four different values of t_{ox} are considered. (a) is for the longitudinal valley, and (b) is for the transverse valley.

is expected that Eq. (2.35) will remain valid as long as the total broadening is negligible compared to the separation between the adjacent eigenenergies. From these tables, it is evident that as F_s increases from 0.49 MV/cm to 1.62 MV/cm, the broadening due to coherent leakage increases by more than one order of magnitude, implying a corresponding decrease in lifetimes associated with coherent tunneling. On the other hand, the broadening due to inelastic scattering increases relatively slowly with increasing F_s. Thus Γ_s plays an important role only at lower gate voltages. Γ_s has been found to be quite sensitive to the quantization effective mass of electrons in silicon. Owing to their smaller effective mass, electrons in the transverse valley has a

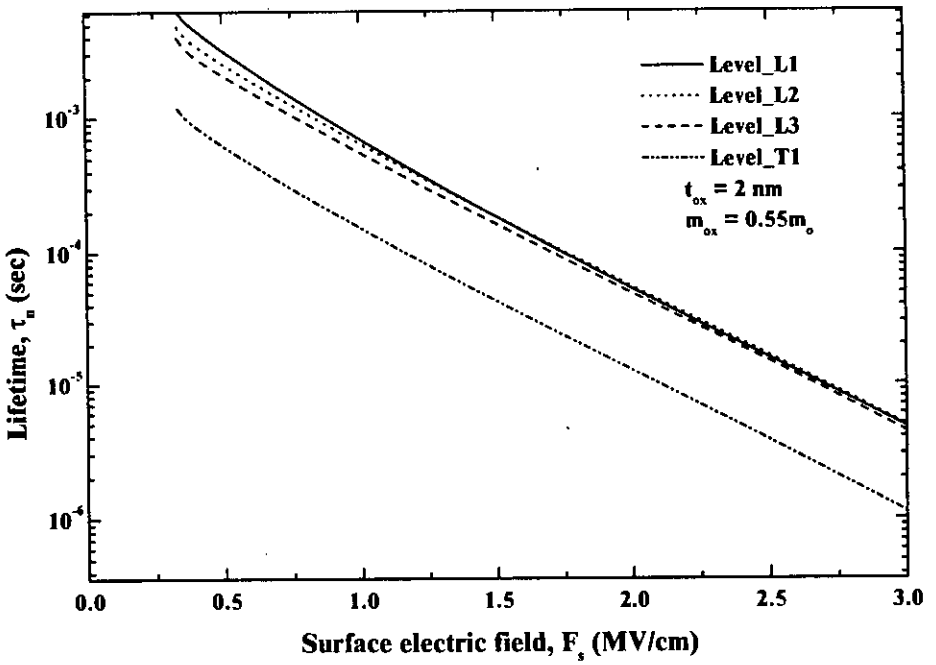


Figure 3.7: Variation of lifetimes, τ_n, with surface electric field, F_s, for lowest three longitudinal valley states (L1, L2, L3) and the lowest transverse valley state (T1).

Number-of state	Γ_t (10^{-12} eV)	Γ_l (10^{-12} eV)	Γ_s (10^{-12} eV)	$\Gamma_l + \Gamma_s$ (10^{-12} eV)
L1	0.533191	0.100044	0.433146	0.533190
T1	2.288970	0.514319	1.774651	2.288970
L2	0.525233	0.125737	0.399495	0.525232

Table 3.2: Relationship among calculated Γ_t, Γ_l and Γ_s for the lowest two states in the longitudinal valley (L1, L2) and the lowest state in the transverse valley (T1). Here, $F_s = 0.49$ MV/cm, $\tau_{coll} = 10^{-6}$ sec and $t_{ox} = 2$ nm.

much larger Γ_s. Dependence of backscattering parameter, B, on F_s for four different t_{ox} for the lowest two states is shown in Fig 3.8. $\tau_{coll} = 10^{-6}$ sec in Fig. 3.8(a) and for Fig. 3.8(b), τ_{coll} is given by Eq.(3.3). It is observed that the effects of backscattering are much more significant in devices with

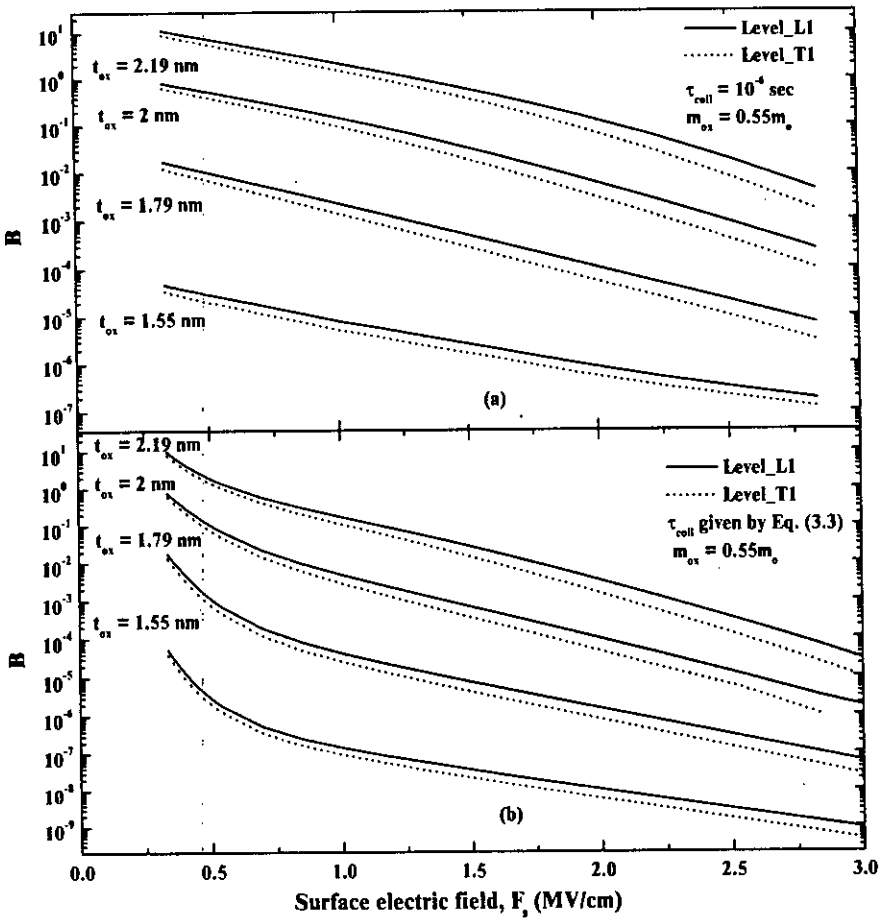


Figure 3.8: Backscattering parameter B vs F_s for the lowest longitudinal and transverse valley states for a given τ_{coll} and four different values of t_{ox}. (a) is for $\tau_{coll} = 10^{-6}$ sec with no spatial or bias dependence. In (b) τ_{coll} is given by Eq. (3.3).

thicker gate oxides and is almost negligible in the structures with sub-2 nm t_{ox}. Fig. 3.8 also shows that as the electric field or the gate voltage is increased, backscattering becomes less important. τ_n ($\tau_n = \tau_{ln}\tau_{sn}/(\tau_{ln} + \tau_{sn})$) as a function of n in the presence of inelastic scattering for a given F_s is presented in Fig. 3.9. The lifetimes of the lower states, which are responsible for the tunneling current, are more severely affected by inelastic scattering. An interesting feature observed in Fig. 3.9 is that inelastic scattering reduces the decreasing nature of τ_n with n for lower values of n, and in the presence of sufficiently strong inelastic scattering, τ_n even starts increasing with in-

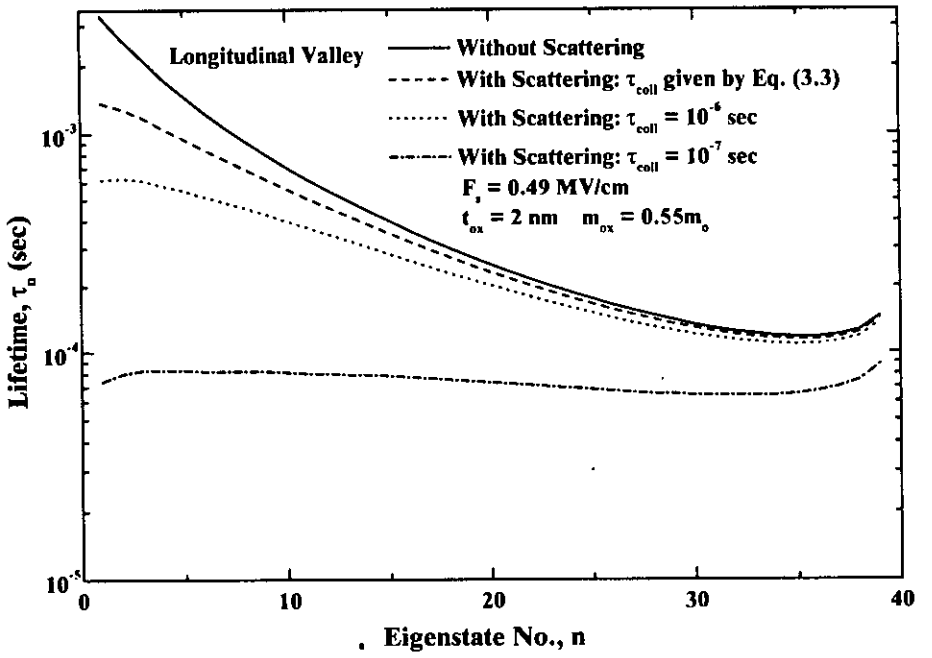


Figure 3.9: Effects of inelastic scattering on the lifetimes of quasi-bound states, τ_n, in the longitudinal valley. Here, τ_{coll} is given by Eq. (3.3). States in the transverse valley also show similar behavior.

Number-of state	Γ_t (10^{-12} eV)	Γ_l (10^{-12} eV)	Γ_s (10^{-12} eV)	$\Gamma_l + \Gamma_s$ (10^{-12} eV)
L1	3.085451	2.381057	0.704403	3.085460
T1	13.777935	10.297753	3.480181	13.777934
L2	3.288287	2.698016	0.590262	3.288278

Table 3.3: Relationship among calculated Γ_t, Γ_l and Γ_s for the lowest two states in the longitudinal valley (L1, L2) and the lowest state in the transverse valley (T1). Here, $F_s = 1.62$ MV/cm, $\tau_{coll} = 10^{-6}$ sec and $t_{ox} = 2$ nm.

creasing n for lower states. This implies that in a device with strong inelastic scattering in the gate-oxide region, an inversion electron with higher energy will find it *more* difficult to tunnel out of the inversion layer.

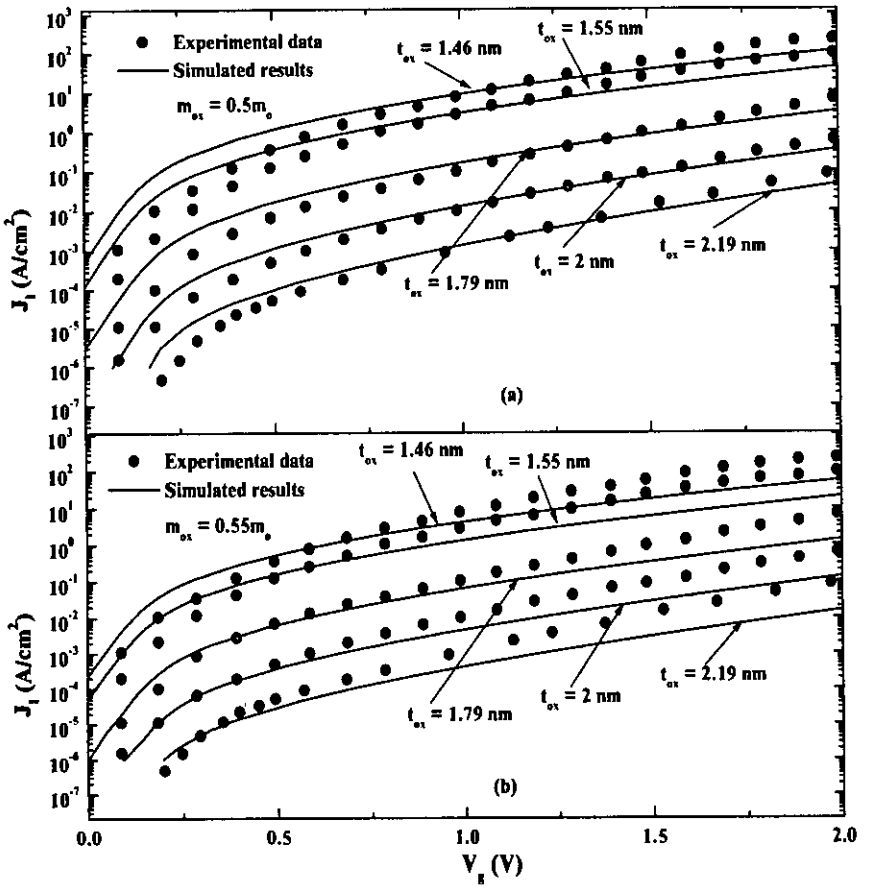


Figure 3.10: Measured and simulated direct tunneling gate leakage current as a function of gate voltage for two different values of effective masses in oxide, m_{ox}. No inelastic scattering is considered here. Experimental data for $t_{ox} = 2.19$ nm has been taken from [7], and for other devices has been taken from [8].

Fig. 3.10 shows direct tunneling gate current calculated with two different values of m_{ox}. No inelastic scattering is considered here. It is found that for a lower value of m_{ox}, gate current at higher gate voltages agree well with experimental results (experimental data for $t_{ox} = 2.19$ nm has been taken from Ref. [7] and the data for other devices has been taken from Ref. [8].), but the current is highly over-estimated at lower gate voltages. On the other hand, when a higher m_{ox} is used, the agreement at lower gate voltages is improved, but the current is under-estimated at higher gate voltages. It is

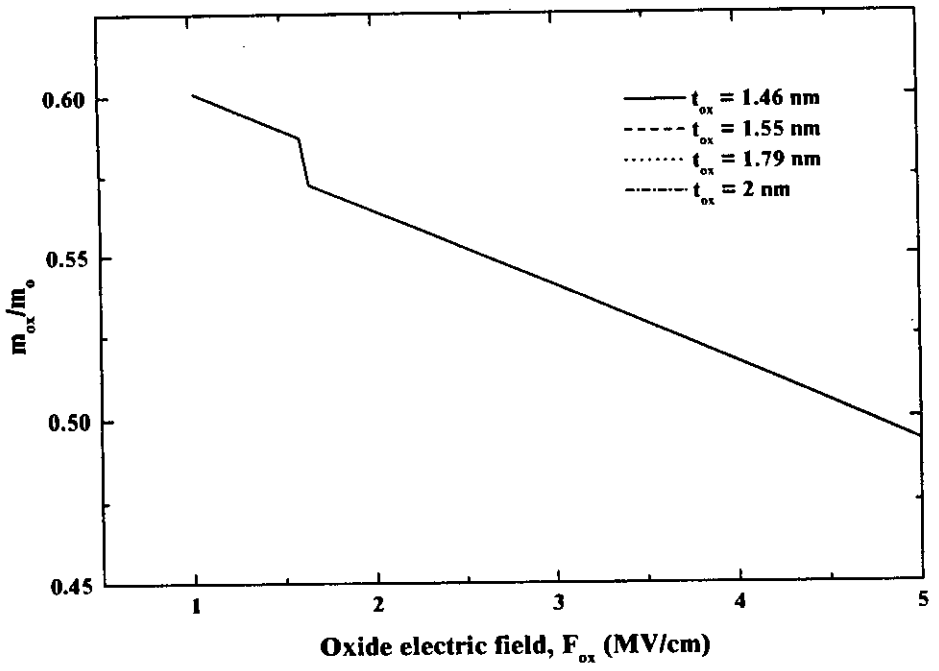


Figure 3.11: Variation of m_{ox} with F_{ox} for electrons. Here, four devices are considered.

also observed that particularly at lower gate voltages, no single value of m_{ox} can fit the experimental data for all the devices well. Fig. 3.10 leads us to conclude from comparison with experiment that m_{ox} is not a constant, rather it is a oxide electric field (or gate voltage) dependent function. Hou and Li [13] have also proposed a field dependent effective mass for holes in Si in p-MOSFETs. In their study, field dependence comes from the non-parabolicity and mixing of the valence bands. We, too, argue that because of the non-parabolicity of the dispersion relationship in gate-oxide region (only a few mono-layers thick), a constant effective mass cannot adequately describe electrons in gate-oxide. From different simulation runs, we come up with the following expression of m_{ox} as a function of F_{ox}.

$$m_{ox} = m_{k1}\left[1 - K_1\left(\frac{F_{ox}}{K_2} - K_3\right)\right] \quad \text{for} \quad F_{ox} \leq 1.6. \tag{3.1}$$

$$m_{ox} = m_{k2}\left[1 - K_1\left(\frac{F_{ox}}{K_2} - K_4\right)\right] \quad \text{for} \quad F_{ox} > 1.6. \tag{3.2}$$

Here, $m_{k1} = 0.6m_o$, $m_{k2} = 0.59m_o$, $K_1 = 0.04$, $K_2 = 1$ MV/cm, $K_3 = 1.06$, $K_4 = 0.9$ and F_{ox} is in MV/cm. The dependence of m_{ox} for electrons on F_{ox} is shown graphically in Fig. 3.11. However, for device with $t_{ox} = 2.19$

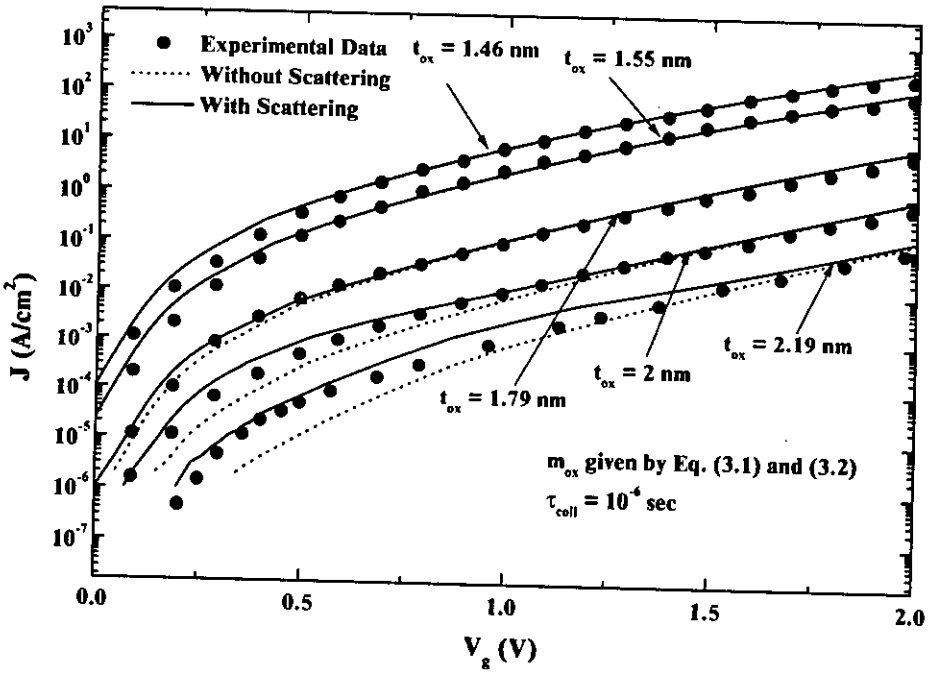


Figure 3.12: Direct tunneling gate current, J, as a function of gate voltage, V_g, calculated with and without considering inelastic scattering. $\tau_{coll} = 10^{-6}$ sec and m_{ox} is function of F_{ox} given in Eqs. (3.1, 3.2).

nm [7] m_{k1} changes from $0.6m_o$ to $0.62m_o$, m_{k2} from $0.59m_o$ to $0.58m_o$ and Eq. 3.1 is valid up to 3 MV/cm of F_{ox}. It may be mentioned that although the existing simulated direct tunneling currents show a good agreement with the measured data at high bias and thinner oxides, at low bias and thicker oxides, the experimental current is much higher than the simulated current [4, 6, 7, 8, 20]. No satisfactory explanation for this mismatch has so far been proposed. In our study, we include the effects of inelastic scattering on direct tunneling gate current to explain the observed mismatch. Direct tunneling gate currents, J, calculated with and without considering inelastic scattering as a function of gate voltage, V_g, for five different t_{ox} are presented in Fig. 3.12. Here, a constant value of $\tau_{coll} = 10^{-6}$ sec is used. Simulated results are also compared with experimental data of Refs. [7, 8] shown in Fig. 3.10. We find that inelastic scattering process has no significant effect in sub-2 nm t_{ox} devices and good agreement with experiment for these devices over the entire range of the gate voltage is obtained without incorporating inelastic scattering. These results can be explained in terms of Tables 3.2-3.3 and Figs. 3.8 and 3.9. The results for $t_{ox} = 2$ nm and 2.19 nm show that the mismatch between simulated and experimental currents

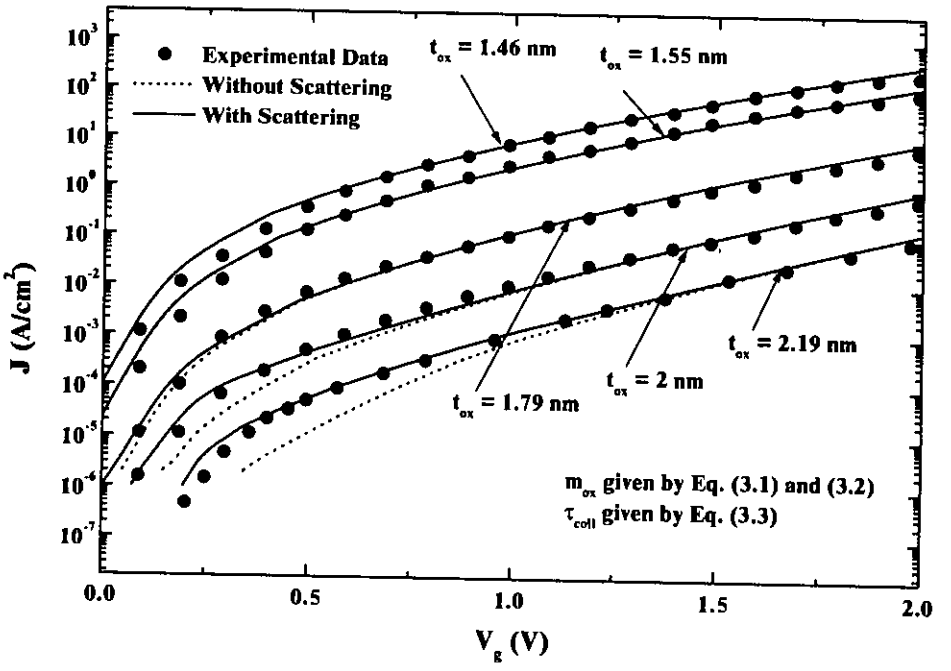


Figure 3.13: Direct tunneling gate current, J, vs gate voltage, V_g. Here, m_{ox} is function of F_{ox} given by Eqs. (3.1, 3.2) and τ_{coll} is given by Eq. (3.3). Excellent agreement with experimental data [7, 8] is achieved for all the devices over the entire range of the gate voltage in the presence of inelastic scattering.

at low gate voltages are reduced by incorporating inelastic scattering effects. However, overall match between model and data is not very good. The gate current for these two devices is over-estimated for intermediate gate voltages. Consequently, we conclude that although inelastic trap scattering can explain the under-estimation of the gate current in structures with $t_{ox} \geq 2$ nm at low gate voltages, the representation of inelastic trap scattering by a position and bias independent τ_{coll} is not adequate for accurate modeling of direct tunneling gate current.

It has been shown recently that the spatial distribution of traps is not uniform throughout the gate-oxide region. Rather the traps, contributing to inelastic scattering, are mostly concentrated near silicon/gate-oxide interface [23, 34]. A detailed discussion on the spatial distribution of traps is given in Ref. [34]. Comparison of modeled and experimental data in Fig. 3.12 also indicates that actually the effects of inelastic scattering on the tunneling current decreases with increasing bias at a rate faster than that dictated

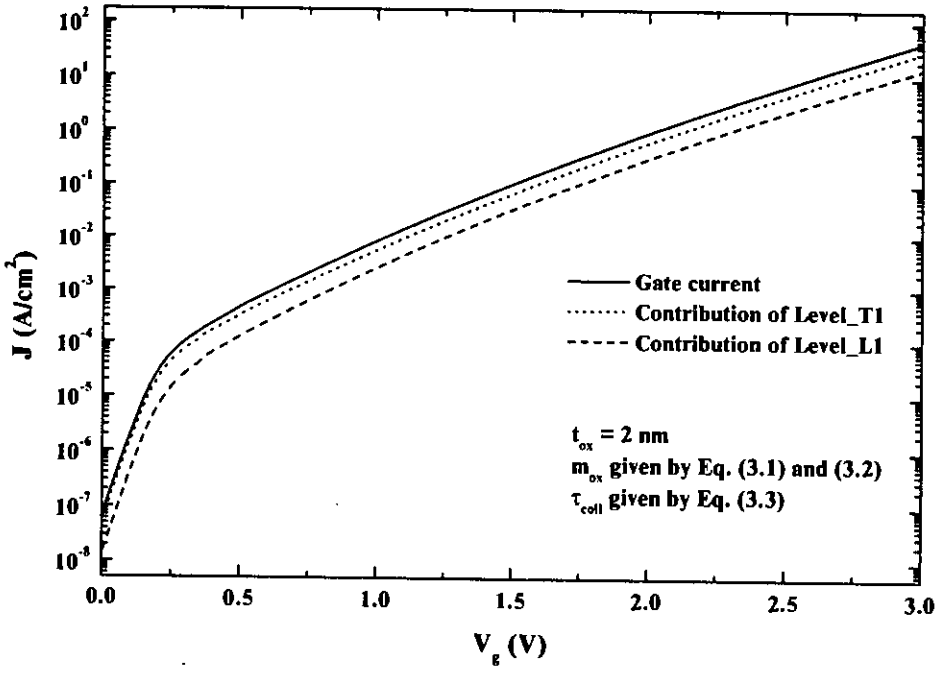


Figure 3.14: Gate current and contribution to gate current of the first eigenstates (L1, T1) of longitudinal and transverse valleys. Here, m_{ox} is function of F_{ox} given by Eqs. (3.1, 3.2) and τ_{coll} is function of position and field given by Eq. (3.3).

by a constant τ_{coll}. We have found from our numerical calculations (results not shown) that good agreement with experimental results cannot be realized when either the spatial dependence or the bias voltage (oxide electric field) dependence of inelastic scattering is ignored. We propose the following expression for the inelastic scattering rate considering both the factors:

$$\frac{1}{\tau_{coll}} = \frac{1}{\tau_0}\left(1 + \left(\frac{F_0}{F_{ox}}\right)^n\right)\exp[z/l], \quad -t_{ox} \leq z \leq 0, \tag{3.3}$$

where τ_0, F_0 and l are three empirical parameters. The oxide electric field or bias dependence may be explained in the following way. At lower F_{ox} (lower gate voltage), many of the traps are unfilled and contribute to inelastic scattering. As F_{ox} is increased, more and more traps become filled and cannot participate in inelastic scattering process. As a result, the scattering rate decreases at higher F_{ox} (higher gate voltage). Fig. 3.13 shows calculated tunneling currents with inelastic scattering rate determined by Eq. (3.3). Excellent agreement between simulated and measured results are obtained even for devices with $t_{ox} \geq 2$ nm over the entire gate voltage range. $\tau_0 = 10^{-5}$

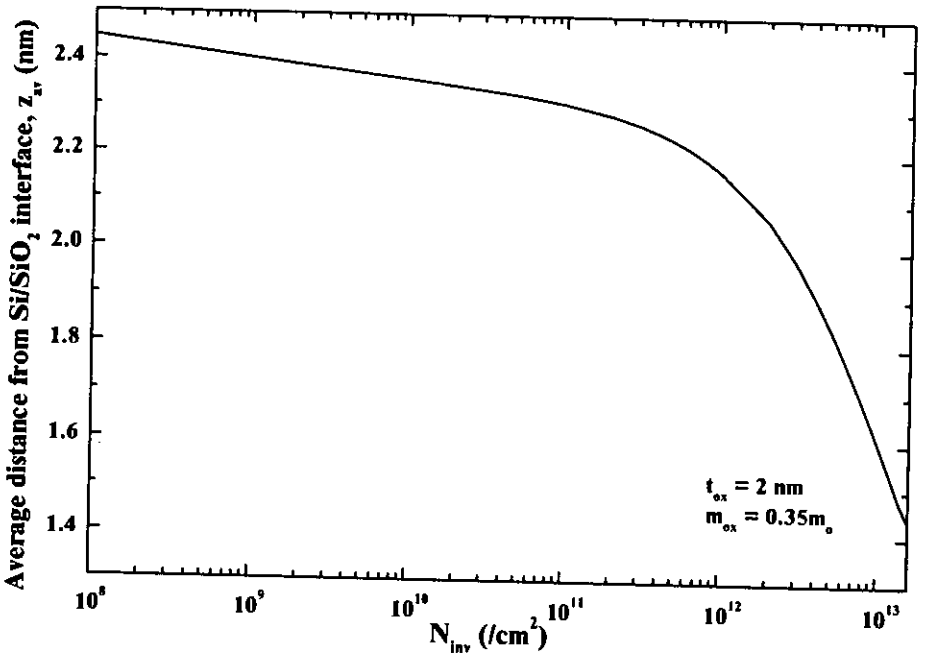


Figure 3.15: The average distance of the inversion holes from Si/SiO_2 interface, z_{av}, as a function of inversion hole density, N_{inv}. Oxide thickness is $t_{ox} = 2$ nm and hole effective mass in oxide is $m_{ox} = 0.35m_o$.

sec, $F_0 \simeq 1.7$ MV/cm, $l = t_{ox}/5$ and $n = 4$ have been used in our calculations. These results explain the existing mismatch between measured and modeled direct tunneling gate currents in devices with $t_{ox} \geq 2$ nm in terms of inelastic trap scattering. Fig. 3.14 shows gate current and contributions to gate current of the lowest two eigenstates (L1, T1) of longitudinal and transverse valleys. Here, m_{ox} is function of F_{ox} given in Eqs. (3.1, 3.2), τ_{coll} is function of position and field given in Eq. (3.3) and $t_{ox} = 2$ nm. Here, we see that majority percent of total current comes from the first eigenstates of longitudinal and transverse valleys. Also, the contribution of first eigenstate of transverse valley is higher than that of first eigenstate of longitudinal valley. This is because τ for transverse valley is much smaller than that for longitudinal valley (see Figs. 3.6, 3.7)

3.2 Simulated Results for p-MOS Devices

Results for p-MOS devices are presented in this section. Again, the analysis has been done for $\{100\}$ Si surface at room temperature. In our analysis,

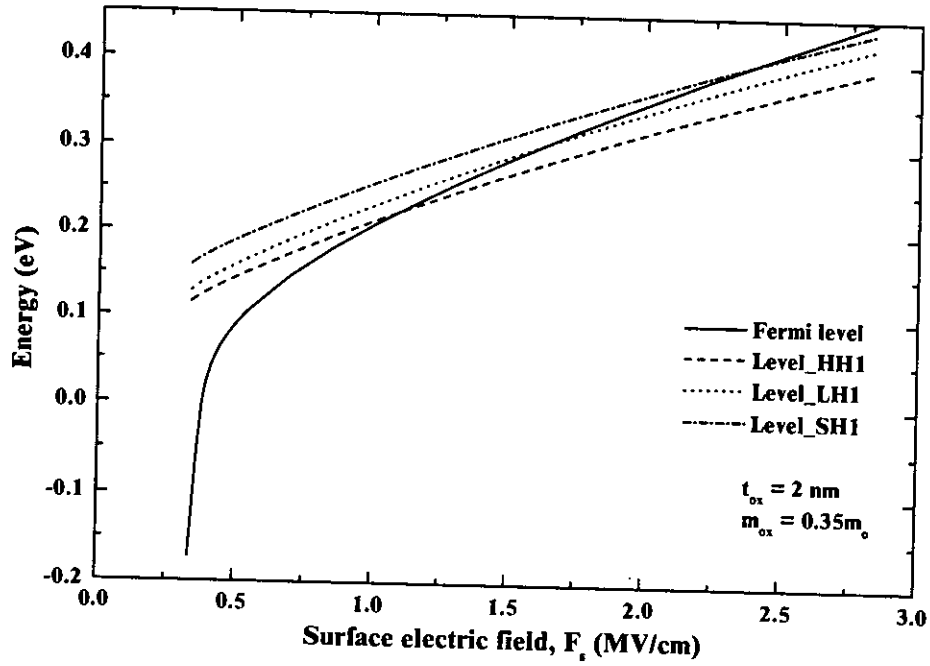


Figure 3.16: Fermi energy level and the first eigenstates of heavy, light and split-off holes as a function of surface electric field, F_s. Here, HH1, LH1 and SH1 are the first eigenstates of heavy, light and split-off hole bands respectively. Oxide thickness is $t_{ox} = 2$ nm and hole effective mass in oxide is $m_{ox} = 0.35m_o$.

Bands	m_3/m_o	m_d/m_o
Heavy hole	0.29	0.433
Light hole	0.20	0.169
Split-off hole	0.29	0.29

Table 3.4: Effective masses for different types of holes.

we have considered three types of hole bands, *i.e.*, heavy, light and split-off hole bands. The quantization effective masses, m_3 and the density-of-states effective masses, m_d for holes used in the model are shown in Table 3.4. The barrier height due to valence band discontinuity of Si and SiO_2 has been taken to be 4.8 eV and the gate-electrode is p-type polysilicon with doping concentration of $N_{poly} = 1 \times 10^{20}/cm^3$. The substrate doping concentration is taken to be $N_D = 5 \times 10^{17}/cm^3$.

Fig. 3.15 shows the average distance of inversion hole distribution, z_{av},

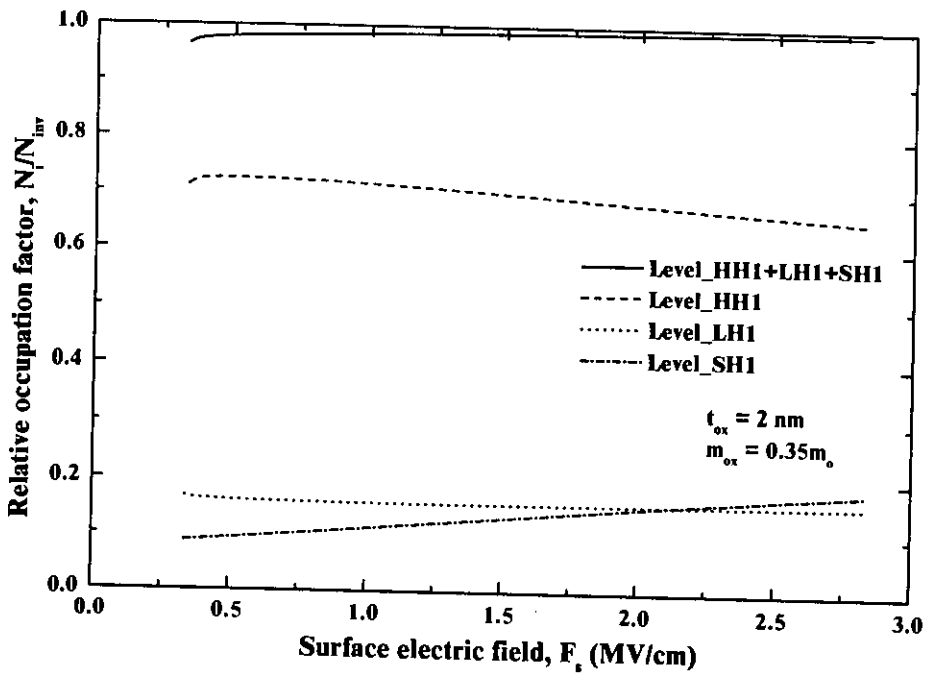


Figure 3.17: Relative occupation factor of inversion holes, N_i/N_{inv}, for the first eigenstates of heavy, light and split-off holes as a function of surface electric field, F_s. Here, HH1, LH1 and SH1 are the first eigenstates of heavy, light and split-off hole bands respectively. Oxide thickness is $t_{ox} = 2$ nm and hole effective mass in oxide is $m_{ox} = 0.35m_o$.

from Si/SiO_2 interface as a function of inversion particle density, N_{inv}. Comparing Fig. 3.15 with Fig. 3.1 we see that the centroid of inversion hole distribution lies at a larger distance from Si/SiO_2 than that for inversion electrons. This is due to the lower effective mass of holes [32]. Fig. 3.16 shows the variation of Fermi energy level and the first eigenstates of heavy, light and split-off holes as a function of surface electric field, F_s. Here, the energies are measured from the bottom of the quantum well at Si/SiO_2 interface. From the figure, we observe that the energies increase with increasing F_s, and at certain values of F_s, the Fermi energy crosses the first eigenenergy of heavy, light and split-off holes respectively. This is because there is no valley degeneracy of heavy, light and split-off holes. However, although the quantization effective masses of heavy and split-off holes are same, the Fermi level crosses level_SH1 at higher F_s because the first eigenstate of split-off hole is 44 meV higher than that of heavy hole, which causes carrier density of this state to become smaller. Similar trend is also observed in the numerical results reported in [32]. Fig. 3.17 shows the relative occupation

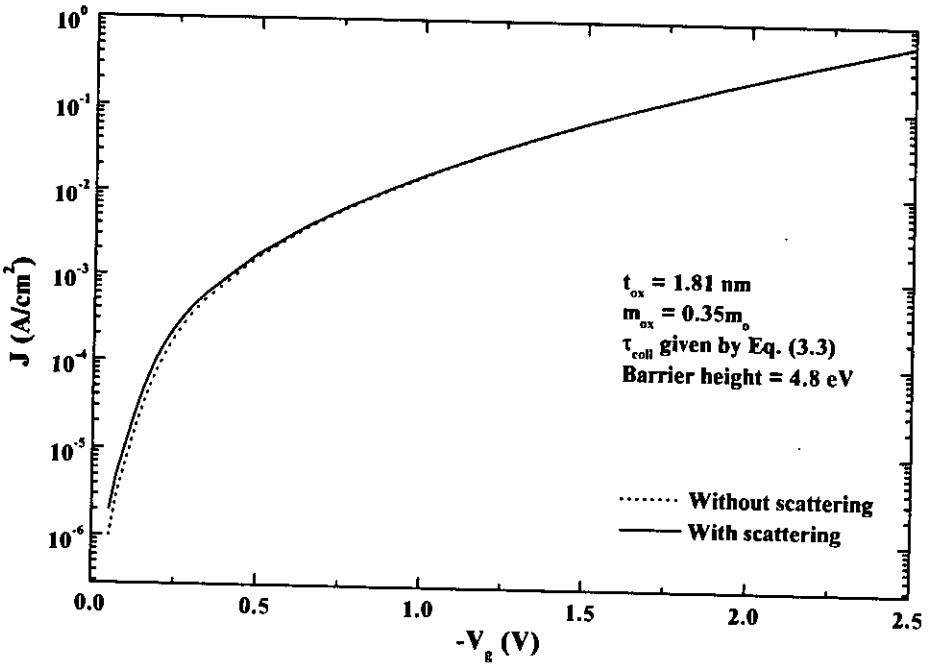


Figure 3.18: Gate current, J, vs gate voltage, V_g, for $t_{ox} = 1.81$ nm. Here, $m_{ox} = 0.35m_o$, barrier height = 4.8 eV and τ_{coll} is given by Eq. (3.3).

factor, N_i/N_{inv}, of the first eigenstates of heavy, light and split-off holes. Here, Level_HH1, Level_LH1 and Level_SH1 refer to the first eigenstates of heavy, light and split-off holes respectively. Also, sum of N_i/N_{inv} for the first eigenstates of three different types of holes is shown in Fig. 3.17. We see that more than 98% of the inversion holes reside in the first eigenstates of three hole bands even at lower surface electric field 0.34MV/cm. Therefore, we can say that only the first eigenstates of three hole bands are sufficient for numerical calculation with reasonable accuracy even at lower surface electric field. Also, relative occupation of split-off holes crosses that of light holes at a certain surface electric field. So, inclusion of only heavy and light holes in the calculation of hole direct tunneling would result in a significant error at higher F_s as done in Ref. [10].

In the preceding discussion we see that with respect to direct tunneling calculation, behaviour of gate-oxide in p-MOS devices differ from n-MOS devices in two respects. The barrier height of the quantum well and carrier effective mass in the oxide of p-MOS devices are not the same as those of n-MOS devices. Now we study the effects of barrier height of the quantum well and m_{ox} on inelastic scattering of direct tunneling carriers. Figs. 3.18,

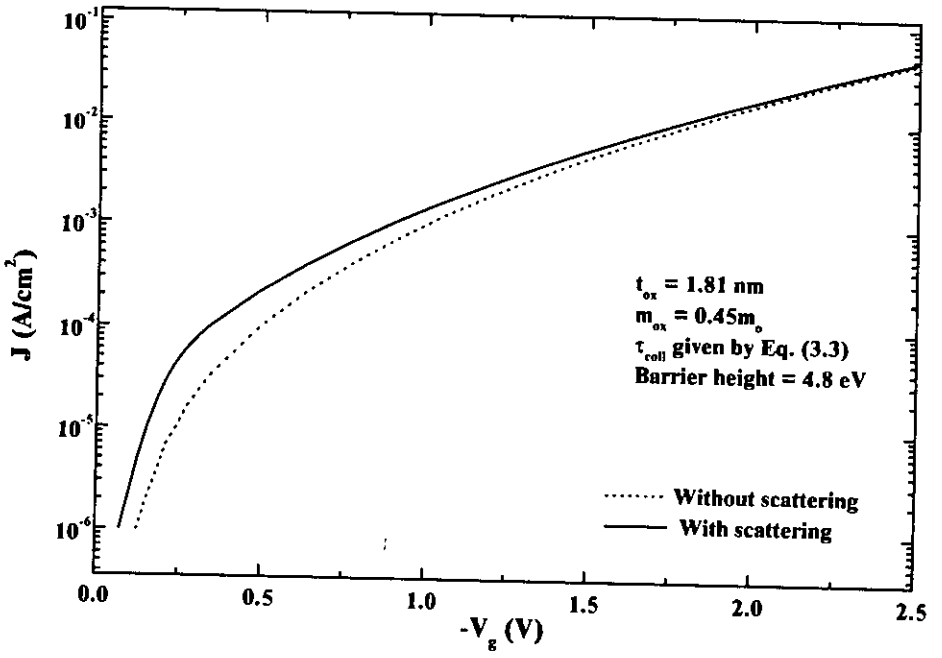


Figure 3.19: Gate current, J, vs gate voltage, V_g, for $t_{ox} = 1.81$ nm. Here, $m_{ox} = 0.45m_o$, barrier height = 4.8 eV and τ_{coll} is given by Eq. (3.3).

3.19 and 3.20 show J vs V_g for two different values of m_{ox} and barrier height. Only $t_{ox} = 1.81$ nm is considered here and the expression of collision time is the same as used for n-MOS devices. Comparing Fig. 3.18 with Fig. 3.19, we see that inelastic scattering effects increase with the increase of m_{ox}. Also, the gate current increases with a reduction in m_{ox}. Again, comparing Fig. 3.18 with Fig. 3.20, we see that inelastic scattering effects decrease in quantum wells with lesser barrier height, although the direct tunneling current increases with reduced barrier height. However, the effects of inelastic scattering on direct tunneling is more sensitive to changes in m_{ox} than changes in barrier height.

We present the simulated hole direct tunneling gate current and compare the results with experimental data. The experimental data is taken from [12]. Again, m_{ox} is not a constant, rather it is a oxide electric field or gate voltage dependent function. From different simulation runs, we come up with the following expression of m_{ox} as a function of F_{ox}.

$$m_{ox} = m_k \left[1 + K_1 \left(\frac{F_{ox}}{K_2} - K_3\right)\right] \tag{3.4}$$

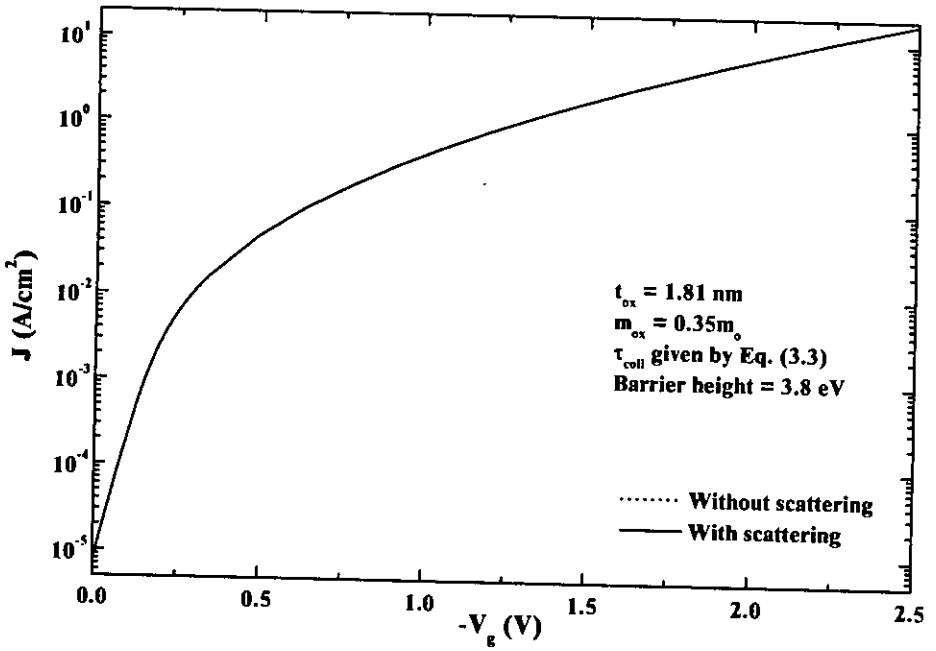


Figure 3.20: Gate current, J, vs gate voltage, V_g, for $t_{ox} = 1.81$ nm. Here, $m_{ox} = 0.35m_o$, barrier height = 3.8 eV and τ_{coll} is given by Eq. (3.3).

Here, $m_k = 0.34m_o$, $K_1 = 0.014$, $K_2 = 1$MV/cm and $K_3 = 0.95$. F_{ox} is in MV/cm. Eq. 3.4 is graphically shown in Fig. 3.21. Note that for holes, m_{ox} increases with increasing F_{ox}, while for electrons, m_{ox} decreases with increasing F_{ox} (Fig. 3.11). Moreover, a single expression can describe hole m_{ox} over the entire range of F_{ox}. The physical reasons behind the opposite F_{ox} dependence of electron and hole m_{ox} are not yet known. It may be due to the different nature of interactions between lower and upper valley electrons and among heavy, light and split-off holes in the gate-oxide region. Fig. 3.22 show gate current, J, vs gate voltage, V_g for three different values of t_{ox} with and without inelastic scattering. Calculated results are compared with the measured data for the same oxide thickness as obtained from [12]. We see that J calculated with inelastic scattering shows excellent agreement with the experimental data for all the devices for the entire range of the gate voltage, even though the model of Ref. [12] underestimates the gate current at low gate voltages. Again, inelastic scattering effects are found to be significant at lower gate voltages and in devices with gate-oxide thickness ≥ 2 nm. It may be mentioned that in our calculations, unlike Ref. [12], we have obtained close match with measurement using bulk values for hole effective masses in Si. To account for non-parabolicity and mixing of valence bands, modified

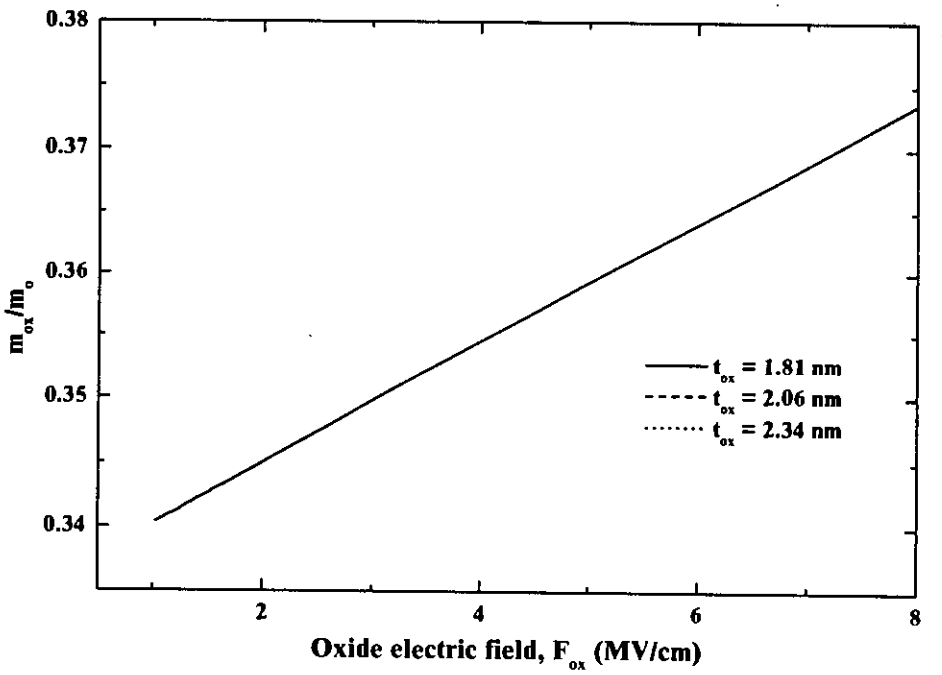


Figure 3.21: Variation of m_{ox} with F_{ox} for holes. Here, three devices are considered.

hole effective masses were used in the model of Ref. [12]. However, the values of these effective masses are derived under the assumption of zero inversion layer charge [13]. Consequently, its validity in strong inversion region cannot be justified. An important feature of our model is the expression of the collision time (Eq. 3.3). The simulated direct tunneling gate currents in the presence of inelastic scattering for both n- and p-MOS devices show excellent agreement with experimental data using Eq. 3.3 to represent collision rate. Therefore, we argue that Eq. 3.3 contains the essential physics of inelastic trap scattering of inversion electrons and holes in gate-oxide region. Fig 3.23 shows the contributions of the lowest states of heavy, light and split-off holes to the hole direct tunneling gate leakage current. Here field dependent m_{ox} (Eq. 3.4) and position and field dependent τ_{coll} (Eq. 3.3) are used. We see that heavy holes have the highest contribution and the contribution of split-off holes is non-negligible. Also, the contribution of split-off holes increases with the increase of gate voltage. So, inclusion of only heavy and light holes in the calculation of hole direct tunneling current, as done in Ref. [10], will introduce a significant error.

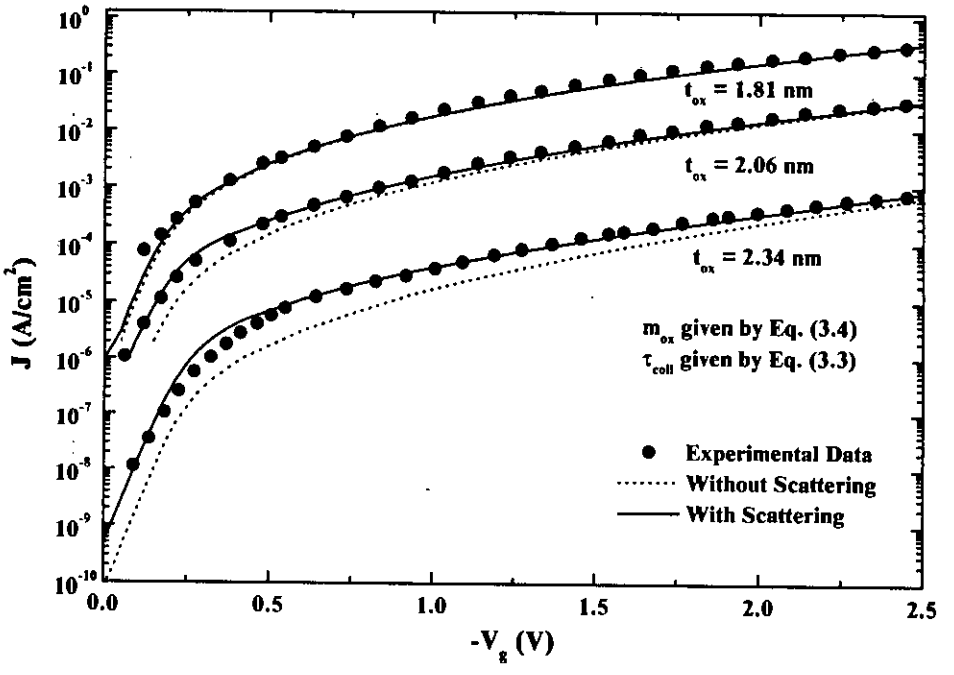


Figure 3.22: Measured and simulated gate current, J, vs gate voltage, V_g, for three different values of t_{ox}. Here, m_{ox} is given by Eq. (3.4) and τ_{coll} is given by Eq. (3.3).

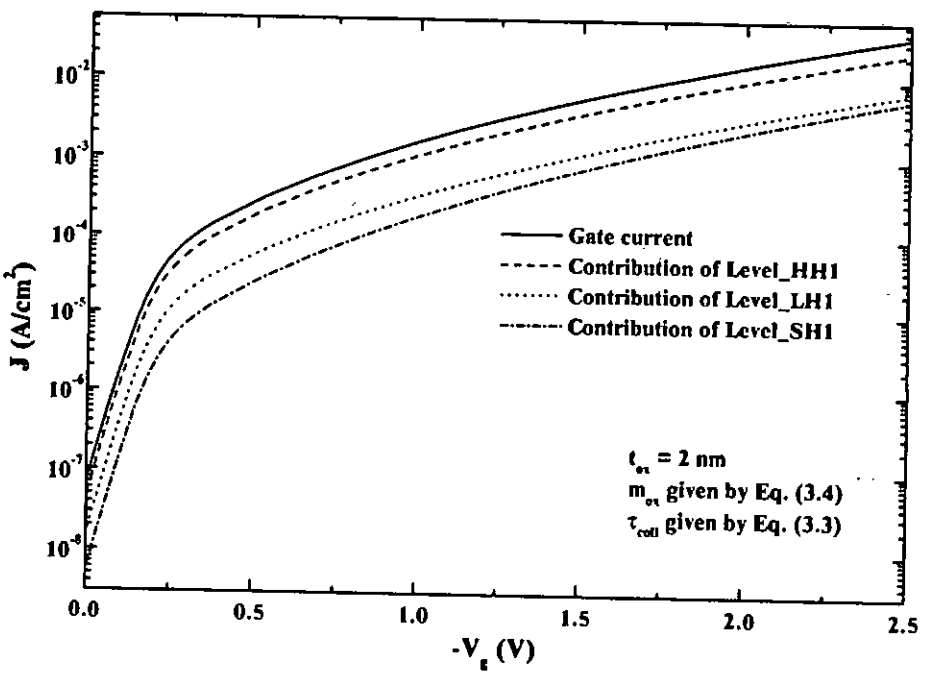


Figure 3.23: Contribution of the first eigenstates of heavy, light and split-off holes to gate current, J, for $t_{ox} = 2$ nm. Here, m_{ox} is given by Eq. (3.4) and τ_{coll} is given by Eq. (3.3).

Chapter 4

Conclusions

An improved and numerically efficient self-consistent model for simulation of direct tunneling gate leakage current incorporating inelastic scattering effects in gate-oxide region has been proposed. This new model is used to study the direct tunneling currents in n- and p-MOS devices by comparing the simulated results with other published simulated results and experimental data.

4.1 Summary

A model of direct tunneling gate leakage current in the presence of inelastic scattering inside gate-oxide region is developed. The one-dimensional Poisson's and Schrödinger's equations are solved self-consistently including the wave function penetration inside the oxide. The real parts of the complex eigenenergies are calculated by locating the peaks of the energy broadened DOS and the imaginary parts are calculated by measuring correcponding FWHM. The wave function and the 1D DOS are calculated using Green's function formalism with open boundary conditions. Therefore, the model is free from the calculation of complex eigenenergies, extensive matrix manipulation and ignoring wave function tail in the oxide. This makes the model numerically efficient and more accurate. The attractive feature of the model is the inclusion of inelastic scattering that occurs inside the oxide. Inelastic trap scattering has been assumed to be the dominant scattering mechanism in gate-oxide region.

Numerical results for n-MOS devices show that electrons in the transverse valley have shorter lifetimes than those in the longitudinal valley. This is due to their smaller quantization effective mass. Also, the carrier lifetimes,

τ_n, exhibits rather unusual behavior with the eigenstate No., n. It does not decrease monotonically, rather starts increasing for the higher eigenstates. Oxide thickness and F_s dependence of lifetimes are found to be almost exponential. Comparing the simulated results with experimental data, it has been found that m_{ox} is not a constant, rather it is a oxide electric field (or gate voltage) dependent function. Oxide field dependence of m_{ox} for electrons is found to be decreasing with the increase of F_{ox}. Also, no single expression of m_{ox} can fit the experimental data for the entire range of gate voltage. The existing mismatch between simulated and measured direct tunneling current at low gate voltages in devices with oxide thickness $\geq$ 2 nm can be explained in term of inelastic scattering. The inelastic scattering effects are found to be significant in devices with oxide thickness $\geq$ 2 nm at low gate voltages. At higher gate voltages, the inelastic scattering effects are insignificant. However, bias and position independent constant collision time cannot explain the existing mismatch at low gate voltages well. Appropriate spatial and bias dependence of collision time needs to be considered for accurate modeling of direct tunneling gate current. The contributions of different valleys to direct tunneling gate leakage current shows that majority percent of the direct tunneling current comes from the lowest state of the transverse valley, although the energy of the lowest state of the longitudinal valley is lower.

Simulated results for p-MOS devices show that the inelastic scattering in gate-oxide is sensitive to the value of m_{ox}. Relative effects of inelastic scattering on direct tunneling current is found to decrease with the reduction of m_{ox}. Also, the inelastic scattering depends on the barrier height of the quantum well. The barrier height dependence of inelastic scattering is not significant. For p-MOS devices, m_{ox} is also found to be a oxide electric field (or gate voltage) dependent function, but the field dependence of m_{ox} for p-MOS devices shows opposite behaviour than that for n-MOS devices. The direct tunneling gate current for p-MOS devices can be accurately modeled over the entire range of gate voltage with the same bias and position dependent collision time as for n-MOS devices. Therefore, the bias and position dependent expression of collision time contains the essential physics of trap scattering. The contribution of split-off holes to direct tunneling current, particularly at higher gate voltages, is not negligible. Therefore, this effect should not be neglected as done in some recent studies.

4.2 Suggestions for Further Work

The model is developed by solving one-dimensional Poisson's and Schrödinger's equations self-consistently assuming that inversion has occurred with zero drain to source voltage. Therefore, a future work can be done for weak inversion, as well as for accumulation with a drain to source voltage applied. When a drain to source voltage is applied in deep submicron devices, two-dimensional Schrödinger's-Poisson's equations will have to be solved. The present work is done for only one gate dielectric material (SiO_2). The effects of different gate dielectrics on the direct tunneling current can be studied. The oxide field dependence of m_{ox} for n- and p-MOS devices shows opposite behaviour. The physical reasons behind this behaviour are not yet known. This can be suggested as a future study. Although the assumed expression of bias and position dependence of collision time can accurately model the direct tunneling gate current over the entire range of gate voltage, the expression for the collision time has been obtained empirically and a quantitative explanation for this expression has not been given in this work. Inelastic scattering phenomena may include energy change of tunneling electrons, which is ignored in this work. This may also be suggested as future study.

Bibliography

[1] "The National Technology Roadmap for Semiconductors: Technology Needs," Semiconductor Industry Association, San Jose, California, 1997.

[2] Y. Tsividis, *Operation and Modeling of the MOS Transistor,* McGraw-Hill, Ch. 3, 1999.

[3] F. Stern, "Self-consistent results for n-type Si inversion layers, " Physical Review B, vol. 5, pp. 4891-4899, 1972.

[4] S.-H. Lo, D. A. Buchanan, Y. Taur and W. Wang, "Quantum-mechanical modeling of electron tunneling current from the inversion layer of ultra-thin-oxide nMOSFET's," IEEE Electron Device Lett., vol. 18, pp. 209-211, 1997.

[5] F. Rana, S. Tiwari and D. A. Buchanan, "Self-consistent modeling of accumulation layers and tunneling currents through very thin oxides," Appl. Phys. Lett., vol. 69, pp. 1104-1106, 1996.

[6] L. F. Register, E. Roesenbaum and K. Yang, "Analytical model for direct tunneling current in polysilicon-gate-metal-oxide-semiconductor devices," Appl. Phys. Lett., vol. 74, pp. 457-459, 1999.

[7] S.-H. Lo, D. A. Buchanan and Y. Taur, "Modeling and characterization of quantization polysilicon depletion, and direct tunneling effects in MOSFETs with ultrathin oxides, " IBM J. Res. Develop., vol. 43, pp. 327-337, 1999.

[8] N. Yang, W. K. Henson, J. R. Hauser and J . J. Wortman, "Modeling study of ultrathin gate oxides using direct tunneling current and capacitance-voltage measurements in MOS devices," IEEE Trans. on Electron Devices, vol. 46, pp. 292-294, 1999.

[9] S. Mudanai, Y.-Y. Fan, Q. Ouyang, Al F.Tasch and S. K. Banerjee, "Modeling of direct tunneling through gate dielectric stacks, " IEEE Trans. on Electron Devices, vol. 47, pp. 1851-1857, 2000.

[10] K.-N. Yang, H.-T. Huang, M.-C. Chang, C.-M. Chu, Y.-S. Chen, M.-J. Chen, Y.-M. Lin, M.-C. Yu, S. M. Jang, D. C. H. Yu, and M. S. Liang, "A physical model for hole direct tunneling current in P^+ poly-gate PMOSFETs with ultrathin gate oxides, " IEEE Trans. on Electron Devices, vol. 47, pp. 2161-2166, 2000.

[11] E. Cassan, P. Dollfus, S. Galdin and P. Hesto, "Semiclassical and wave-mechanical modeling of charge control and direct tunneling leakage in MOS and H-MOS devices with ultrathin oxides," IEEE Trans. on Electron Devices, vol. 48, pp. 715-721, 2001.

[12] Y. T. Hou, M. F. Li, W. H. Lai and Y. Jin, "Modeling and chacterization of direct tunneling hole current through ultrathin gate oxide in p-metal-oxide-semiconductor field-effect transistors," Appl. Phys. Lett., vol. 78, pp. 4034-4036, 2001.

[13] Y. T. Hou and M. F. Li, "Hole quantization effects and threshold voltage shift in pMOSFET-assessed by improved one-band effective mass approximation," IEEE Trans. Electron Devices, vol. 48, pp. 1188-1193, 2001.

[14] J. Wang, Y. Ma, L. Tian, and Z. Li, "Modified Airy function method for modeling of direct tunneling current in metal-oxide-semiconductor structures," Appl. Phys. Lett., vol. 79, pp. 1831-1833, 2001.

[15] K. Sakakibara, N. Ajika, K. Eikyu, K. Ishikawa and H. Miyoshi, "A quantitative analysis of time -decaying reproducible stress-induced leakage current in SiO_2 film," IEEE Trans. Electron Devices, vol. 44, pp. 1002-1008, 1997.

[16] S.-I. Takagi, N. Yasuda and A. Toriumi, "Experimental evidence of inelastic tunneling in stress-induced leakage current," IEEE Trans. Electron Devices, vol. 46, pp. 335-341, 1999.

[17] S.-I. Takagi, N. Yasuda and A. Toriumi, "A new I-V model for stress-induced leakage current including inelastic scattering," IEEE Trans. Electron Devices, vol. 46, pp. 348-354, 1999.

[18] M. Houssa *et al.*, "Trap assisted tunneling in high permittivity gate dielectric stacks," J. Appl. Phys., vol. 87, pp. 8615-8620, 2000.

[19] A. Ghetti, J. Bude and G. Weber, "T_{BD} prediction from low-voltage near-interface trap-assisted tunneling current measurements," IEEE Trans. Electron Devices, vol. 48, pp. 1354-1359, 2001.

[20] L. Larcher, A. Paccagnella and G. Ghidini, "Gate current in ultra-thin MOS capacitors: a new model of tunneling current," IEEE Trans. Electron Devices, vol. 48, pp. 271-278, 2001.

[21] M. Städele, B. Fischer, B. R. Tuttle and K. Hess, "Influence of defects on elastic gate tunneling currents through ultrathin SiO_2 gate oxides: predictions from microscopic models, " Superlattices and Microstructures, vol. 28, pp. 517-524, 2000.

[22] Y. Fu, M. Willander, P. Lungren and E. Aderstedt, "Enhanced tunneling current through thin oxide due to single-defect scattering," Appl. Phys. Lett., vol. 78, pp. 486-488, 2001.

[23] A. Ghetti, E. Sangiorgi, J. Bude, T. W. Sorsch and G. Weber, "Tunneling into interface states as reliability monitor for ultrathin oxides," IEEE Trans. on Electron Devices, vol. 47, pp. 2358-2365, 2000.

[24] Saif Uz Zaman, "Calculation of gate leakage current in deep submicron MOSFET in the presence of electron phase-breaking scattering," M.Sc. Thesis, Bangladesh University of Engineering and Technology, Department of Electrical and Electronic Engineering, 2001.

[25] A. Haque, A. Rahman and I. B. Chowdhury, "On the use of appropriate boundary conditions to calculate the normalized wave functions in the inversion layers of MOSFETs with ultra-thin gate oxides," Solid-State Electronics, vol. 44, pp. 1833-1836, 2000.

[26] A. Haque and A. N. Khondker, "An efficient technique to calculate the normalized wave functions in arbitrary one-dimensional quantum well structures," J. Appl. Phys., vol. 84, pp. 5802-5804, 1998.

[27] A. N. Khondker, M.R. Khan and A.F.M. Anwar, "Transmission line analogy of resonance tunneling phenomena: the generalized impedance concept," J. Appl. Phys., vol. 63, pp. 5191-5193, 1988.

[28] A. N. Khondker and M. A. Alam, "Buttiker-Landauer conductance formulas in the presence of inelastic scattering," Physical Review B, vol. 44, pp. 5444-5452, 1991.

[29] C. Moglestue, "Self-consistent calculation of electron and hole inversion charges at silicon-silicon dioxide interfaces," J. Appl. Phys., vol. 59, pp. 3175-3183, 1986.

[30] L. Larcher, A. Paccagnella and G. Ghidini, "A model of the stress induced leakage current in gate oxides," IEEE Trans. Electron devices, vol. 48, pp. 285-288, 2001.

[31] A. D. Stone and P. A. Lee, "Effects of inelastic processes on resonant tunneling in one dimension," Physical Review Lett., vol. 54, pp. 1196-1199, 1985.

[32] M. Zahed Kauser, "Effects of wavefunction penetration into the gate oxide on self-consistent modeling of deep submicron MOSFETs," M.Sc. Thesis, Bangladesh University of Engineering and Technology, Department of Electrical and Electronic Engineering, 2001.

[33] A. Rahman and A. Haque, "A study into the broading of quantized inversion layer states in deep submicron MOSFETs," Solid-State Electron., vol. 45, pp. 755-760, 2001.

[34] M. M. Chowdhury, S. Zaman, A. Haque and M. R. Khan, "Determination of electron trap distribution in gate-oxide region of deep submicron metal-oxide-semiconductor structure from direct tunneling gate current," Appl. Phys. Lett., to be published, 2002.

[35] S. Jallepalli, J. Bude, W.-K. Shih, M. R. Pinto, C. M. Maziar and A. F. Tasch, "Electron and hole quantization and their impact on deep submicron silicon p- and n-MOSFET characteristics," IEEE Trans. on Electron Devices, vol. 44, pp. 297-302, 1997.

[36] S.-I Tagaki, M. Takayanag and A. Toriumi, "Characterization of inversion-layer capacitance of holes in Si MOSFET's, " IEEE Trans. on Electron Devices, vol. 46, pp. 1446-14450, 1999.

[37] C. A. Richter, A. R. Hefner and E. M. Vogel, "A Comparison of quantum-mechanical capacitance-voltage simulators," IEEE Electron Device Lett., vol. 22, pp. 35-37, 2001.

Appendix A

Program flow-chart: