Extraction of Exact Layer Thickness of Ultra-thin Gate Dielectrics in Nanoscaled CMOS under Strong Inversion

Munmun Dey and Sanatan Chattopadhyay

Abstract—The impact of surface quantization on device parameters of a Si metal oxide semiconductor (MOS) capacitor has been analyzed in the present work. Variation of conduction band bending, position of discrete energy states, variation of surface potential, and the variation of inversion carrier concentration at charge centroid have been analyzed for different gate voltages, substrate doping concentrations and oxide thicknesses. Oxide thickness calculated from the experimental C-V data of a MOS capacitor is different from the actual oxide thickness, since such data include the effect of surface quantization. A correction factor has been developed considering the effect of charge centroid in presence of surface quantization at strong inversion and it has been observed that the correction due to surface quantization is crucial for highly doped substrate with thinner gate oxide.

Index Terms—Charge centroid, CMOS, eigen states, surface quantization, ultra-thin gate dielectric

I. INTRODUCTION

Silicon has been emerged as the sole contender in semiconductor industry over the last 50 years [1]. Throughout this period of time, Integrated Circuit (IC) technology has evolved through the methodology of down scaling of individual MOS devices. For few decades, scaling was proved to be the most reliable methodology on the way to successful transition from one technology node to another [2] in order to achieve desired device performance. Scaling involves gradual miniaturization of device dimensions such as gate dielectric thickness (t_{ox}), gate length (L_G) and higher level of channel doping. As the device dimension enters into deep submicron (DSM) regime ($L_G \leq 0.25$ μm), application of gate bias results in high transverse electric field at SiO_2/Si interface [3]. Presence of such a strong electric field at the interface gives rise to bending of conduction band (E_C) at semiconductor side. At this situation a potential well is formed just beneath the SiO_2/Si interface and conduction band at the surface splits into discrete energy subbands and quantum mechanical effects (QME) become apparent [4]. The inversion carrier distribution and transport properties of carriers under QME are quite different from that of classical influence. Classically, most of the inversion carriers reside at the oxide interface, whereas, quantum mechanically, inversion carrier distribution has its peak or centroid few Angstrom deep into the substrate. The region between the interface and the inversion layer centroid has negligible carrier concentrations and hence behaves as a virtually depleted region.

Efficient circuit designing requires accurate device modeling. Two terminal MOS structure has been extensively studied [5] for capacitance-voltage (C-V) modeling. As we go for successive technology nodes the dielectric layer thickness of MOS devices has become few atomic layers thick and this is comparable and even thinner than inversion layer (< 5 nm). Classical models are inadequate at DSM and nano-metric dimensions and application of such models will lead to inaccurate modeling of capacitance and extraction of erroneous gate oxide thickness. Therefore, the accurate C-V modeling of such devices must include surface quantisation effect [6-7]. Presence of charge centroid effectively widens the gate dielectric layer and as a result capacitance value

decreases. Carrier concentration in the inversion layer is contributed by the occupation probability of the discrete energy states formed within the potential well. Generally the ground state electrons take part in the transport mechanism. Individual carrier distribution of each state altogether gives rise to average carrier distribution across the channel.

In the present work, a Si MOS capacitor with ultra-thin gate (dielectric) oxide thickness has been simulated using both classical and quantum mechanical models and the results are compared. Variation in quantisation effect is observed by applying a range of gate bias. Dependence of QME on substrate doping concentration (N_a) and gate oxide thickness are also analysed. A correction factor for the extraction of exact thickness of such ultra-thin oxides is developed by taking into account the impact of surface quantization effect underneath the oxide layer.

II. SIMULATION OF THE MOS CAPACITOR

Simulation of the MOS capacitors with different ultra-thin gate oxide thicknesses are performed using the commercially available simulation package, SILVACO [8]. The typical device structure considered for the current simulation is shown in Fig. 1. A single gate MOS capacitor on a p-Si substrate with Al (Aluminum) as the gate material has been considered. The Schrödinger and Poisson's equations along with Fermi-Dirac distributions were solved self consistently in one dimension. Quantum calculations are performed for five eigen states for both longitudinal (m_l) and transverse effective masses (m_t) of the electrons. Devices with three sets of substrate doping concentrations (N_a) such as, 1×10^{17} cm^{-3}, 5×10^{17} cm^{-3}, and 1×10^{18} cm^{-3}, and four different oxide thicknesses (t_{OX}) such as, 1.5 nm, 2.5 nm, 3.5 nm and 5 nm have been simulated. Values for all the physical parameters used in the present simulations are given in Table 1.

Table 1. Summary of parameter values used in the current simulations

Symbol	Description	Value
E_g(Si)	Band gap of Si at 300 K	1.08 eV
χ_{Si}	Electron affinity of Si	4.17 eV
ε_{SiO_2}	Permittivity of SiO$_2$	3.9

III. RESULTS AND DISCUSSION

The schematic of the simulated structure and its band alignment at strong inversion is shown in Fig. 1. A 1 μm × 1 μm MOS capacitor with 100 nm epitaxial Si layer and different doping concentrations (1×10^{17} cm^{-3}, 5×10^{17} cm^{-3}, 1×10^{18} cm^{-3}) and oxide thicknesses (t_{ox} = 1.5 nm, 2.5 nm, 3.5 nm, and 5 nm) are simulated.

At the bulk $E_i - E_F = \Psi_B$ and the value of surface potential Ψ_S (i.e, Ψ_B at the surface) is 0. The measure of surface potential at any point within the semiconductor is the difference of energy between E_i in the bulk and the said point. At a point where E_i touches E_F, the concentrations of electron and hole are similar. Inversion of the surface begins when E_i just crosses E_F. The region, where $\Psi_B < \Psi_S < 2 \Psi_B$, is called weak inversion and, the strong inversion occurs for a condition $\Psi_S > 2 \Psi_B$.

Classically, maximum number of carriers reside at the surface at strong inversion and its concentration is determined by the value of surface potential (Ψ_S) [5] as,

$$n_S = n_0 \left[\exp\left(\frac{q(\Psi_S - \Psi_B)}{kT} \right) \right] \tag{1}$$

where, n_0 is the intrinsic carrier concentration, k is Boltzmann constant and T is the temperature in absolute scale. The variation of potential from the bulk towards surface could be calculated by solving Poisson's equation [9],

$$\frac{d^2\Psi}{dz^2} = -\frac{\rho}{\varepsilon_s} \tag{2}$$
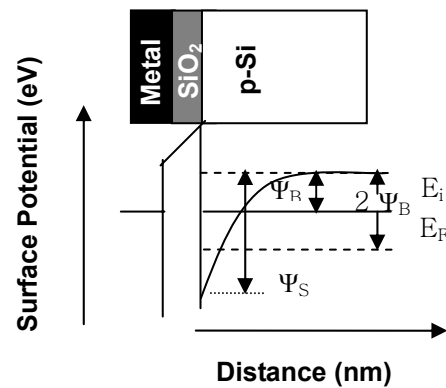

Fig. 1. Schematic of the Si MOS capacitor with its band alignment at strong inversion.

where, ρ is the charge density, and ε_s is the semiconductor dielectric constant.

The expression for electric field variation can be obtained by solving equation (2) as,

$$\frac{d\Psi}{dz} = \sqrt{\frac{2p_{p0}kT}{\varepsilon_s}\left[\left(e^{-\beta\Psi} + \beta\Psi - 1\right) + \frac{n_{p0}}{p_{p0}}\left(e^{\beta\Psi} - \beta\Psi - 1\right)\right]} \quad (3)$$

Application of positive bias (for p-type Si substrate) at the gate terminal causes bending of conduction band. The conduction band (E_C) under strong inversion forms a quantum well (QW) with SiO$_2$/Si interface, which can be approximated as a triangular barrier potential well. Under such a condition, the conduction band is no more continuous and splits into discrete energy sub-bands. The electrons reside in these sub-bands within the well and are allowed to move in two directions only. Movement of the carriers in perpendicular direction to the plane of interface gets confined. The energy position of ith sub-band can be determined by solving Schrödinger equation,

$$\frac{d^2\psi_i}{dz^2} + \frac{2m_z}{\hbar^2}\left[E_i + q\phi(z)\right]\psi_i = 0 \quad (4)$$

where, z is the depth of the device from SiO$_2$/Si interface. For simplification, $\phi(z)$ in Eq. (4) is replaced by $-F_S z$ for $z > 0$, and by an infinite barrier (∞) for $z < 0$, which is the so-called triangular potential approximation. Solution of such equation includes the well-known Airy function [10] which can be expressed as:

$$\psi_i(z) = Ai\left[\left(\frac{2m_z q F_S}{\hbar^2}\right)^{\frac{1}{3}}\left(z - \left(\frac{E_i}{qF_S}\right)\right)\right] \quad (5)$$

where,

$$E_i = \left(\frac{\hbar^2}{2m_z}\right)^{\frac{1}{3}}\left[\frac{3}{2}\pi q F_S\left(i + \frac{3}{4}\right)\right]^{\frac{2}{3}} \quad (6)$$

Average separation of carriers residing in the ith state from SiO$_2$/Si interface is given as [10],

$$z_i = \frac{\int z\psi_i^2(z)dz}{\int \psi_i^2(z)dz} \quad (7)$$

The triangular potential well approximation works well under depletion condition when there is no or little charge in the inversion layer compared to the depletion charge. When only one subband is occupied then a variational approach gives better result for the lowest sub-band, and the corresponding trial eigen function can be approximated as [10],

$$\psi_0(x) = \left(\frac{1}{2}b^3\right)^{\frac{1}{2}} z \exp\left(\frac{-bz}{2}\right) \quad (8)$$

where the value of parameter b is calculated by minimizing the energy of the system using the eigen function in Eq. (8). The ground state sub-band energy of such a system is given by [10],

$$E_0 = \left(\frac{\hbar^2 b^2}{8m_z}\right) + \left(\frac{3q^2}{\in_s b}\right)\left[N_{depl} + \frac{11}{16}N_{inv} - \left(\frac{2}{b}\right)\left(N_a - N_d\right)\right] \quad (9)$$

and eigen energies of the higher sub-bands are obtained to be [10]:

$$E_i = E_{i,d} - \frac{q^2 F_{depl} F_{inv} z_0^2}{4E_{i,d}} - \frac{4E_{i,d}^2}{15qF_{depl}z_d} + qF_{inv}z_o \quad (10)$$

where, q is the electronic charge, F_{depl} is the depletion field, F_{inv} is the field in the inversion layer, z_0 is the average distance of the charge density associated with variational wave function and $E_{i,d}$ is the eigen values of the energy states found from triangular potential barrier approximation (Eq. 6).

Fig. 2 shows the variation of conduction band potentials at strong inversion along the distance from the SiO$_2$/Si interface for MOS structures having different substrate doping concentrations (1 x 10^{17} cm^{-3}, 5 x 10^{17} cm^{-3}, 1 x 10^{18} cm^{-3}), oxide thickness, t$_{ox}$ = 1.5 nm and gate bias, V$_G$ = 2 V. The mutual energy separations of three ground states, as seen in Fig. 2, are very small. Thus, these states have nearly equal electron occupation probabilities. However, the occupation probabilities for higher energy states (E$_1$ and E$_2$) differ much, in accordance
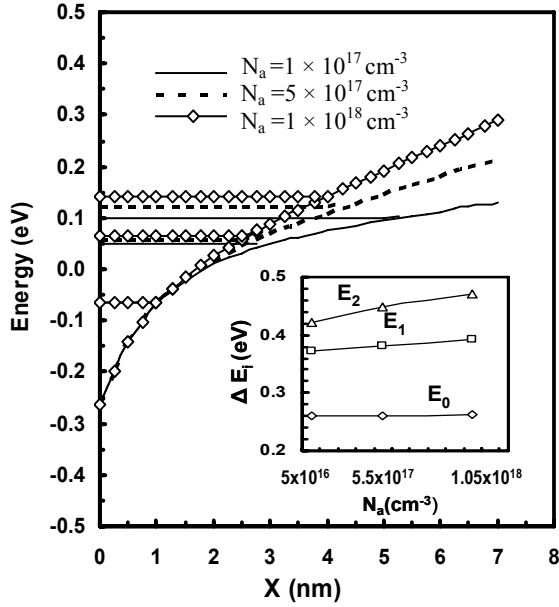
Fig. 2. Plots of simulated lowest eigen states for different substrate doping concentrations for the oxide thickness, t_{ox} = 1.5 nm and gate bias, V_G = 2 V. The plots of variation of ΔE_i with substrate doping concentration are shown in the inset.

with their relative energy separation. Energy separation between the band-edge and sub-band is denoted as ΔE_i. The inset of Fig. 2 shows the variations of ΔE_i with substrate doping concentrations, for lowest three eigen states. Variation of ΔE_i, at strong inversion, for ground state is almost independent of N_a, whereas, ΔE_i increases significantly with N_a for the excited states. As ΔE_i increases, i^{th} state energetically moves upward, which indicates the formation of steeper QW just beneath the SiO$_2$/Si interface. Thus, ΔE_i can be considered as the measure of quantization. Upward shift of energy states reduce the electron occupation probability. For the MOS structures having fairly high substrate doping (10^{18} cm^{-3}), mutual separation of energy states, are higher (inset of Fig. 2) which gives an indication that the excited states of highly doped substrate are less populated compared to that of their lightly doped counterpart. It is observed from Fig. 2 that for a given gate bias, the effective electric field inside the semiconductor is almost independent of the doping concentration within a distance of 1.5 nm from the SiO$_2$/Si interface. Although, the overall shape of the QW formed below the gate dielectric is much dependent on the value of N_a. In weak inverted and depleted regions, for a fixed value of V_G, the potential profile of E_C has higher value for higher doping concentration. Accordingly, the crossover point

of quasi Fermi level (E_F) and E_i is shifted towards the SiO$_2$/Si interface leading to a narrower inversion layer for highly doped substrates.

Carrier concentration in the ith sub-band is given by [10],

$$N_i = \left(\frac{n_{vi} m_{di} kT}{\pi \hbar^2} \right) F_0 \left[\frac{E_F - E_i}{KT} \right] \qquad (11)$$

where, n_{vi} is the valley degeneracy, m_{di} is the density of states effective mass per valley, k is the Boltzmann constant., $\hbar$ is the reduced Plank's constant, and the location of averege charge centroid is given as [10],

$$z_{av} = \frac{\sum_i N_i z_i}{N_{inv}} \qquad (12)$$

As discussed above, the relative energy positions of the first eigen states for different doping concentrations are identical and those for higher eigen states are separated. Therefore it is expected that the inversion carrier distribution in the first eigen states will be equal and that at higher eigen states different, and their relative energy values will dictate the carrier distributions. Fig. 3 shows the distribution of inversion carriers as a function of distance, in the ground state (E_0) and first excited state (E_1) for three different substrate doping. Fig. 3 depicts the situation at strong inversion. Distributions of inversion
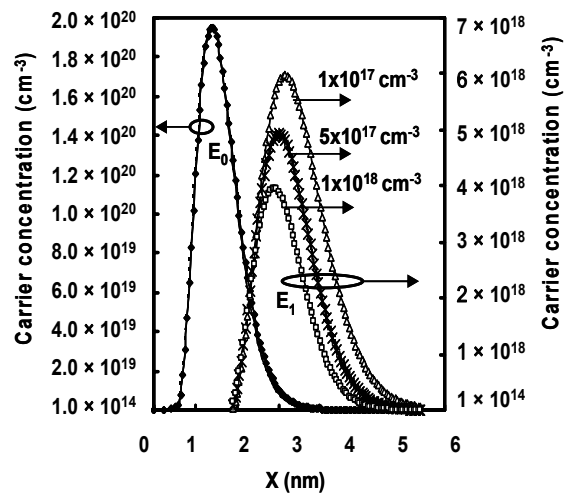

Fig. 3. Variation of electron concentration in two lowest eign states as a function of distance for different doping concentrations.

carriers, as seen from Fig. 3, in ground states are almost same, although the distributions for first excited states differ from each other. Increase in doping results in reduced electron concentration but the peak position of the distribution shifts towards SiO_2/Si interface.

Surface potential (Ψ_S) is related with the gate voltage (V_G) as,

$$\Psi_s = V_G - V_{FB} + \frac{(Q_{depl} + Q_{inv})}{C_{ox}}$$

$$= V_G - (\phi_m - \chi_S - \frac{E_g}{2} - \frac{kT}{q} \ln \frac{N_a}{n_0}) + \frac{(Q_{depl} + Q_{inv}) \, t_{ox}}{\varepsilon_{ox}} \quad (13)$$

Fig. 4 shows the variation of Ψ_S as a function of gate voltage for different substrate doping. V_G is varied between 0 to 2 V. Initially, up to 0.75 V, Ψ_S varies almost linearly with V_G. In this linear region the plots for three different doping concentrations are distinct from each other. When $V_G > 0.75$ V, Ψ_S no longer maintains the linearity with V_G and above 1 V three different plots merge with each other. As mentioned, surface potential is the measure of energy difference between the intrinsic and quasi Fermi levels of the system at the interface. Therefore, the merging of surface potentials above 1 V (after strong inversion) indicates that the relative positions of energy eigen values will not vary significantly and hence, the inversion carrier distribution will be identical for different doping concentrations. Thus the plots of Fig. 4 support the observation of Fig. 2. For a given gate voltage, when the surface is in strong
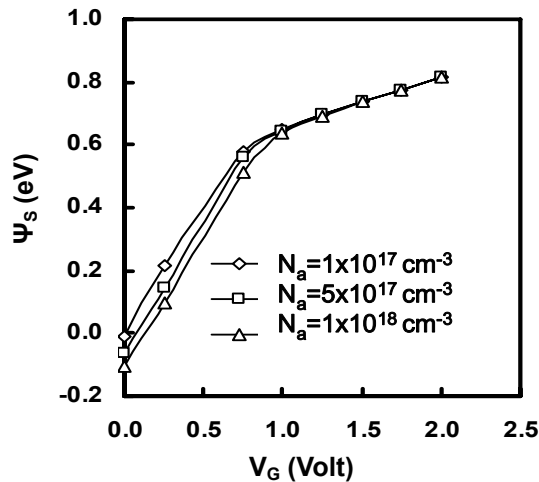
inversion then the region few nanometers below the surface is in moderate or weak inversion. At strong inversion, the conduction bands are merged up to 1.5 nm (see Fig. 2), whereas these are separated after 1.5 nm from the surface where the region is in moderate to weak inversion.

Fig. 5 shows the variation of centroid position with the corresponding inversion carrier concentration at the centroid (N_z) for three different N_a. Position of centroids for different N_a are far apart from each other at lower value of N_z. As N_z increases, centroids shift towards the SiO_2/Si interface and at strong inversion centroid positions become almost independent of substrate doping. At lower V_G, when E_C is not bent enough, less number of inversion carriers are present in the channel. At this situation, bending of E_C and the width of the QW is strongly dependent on substrate doping. Accordingly, ground states (E_0) and excited states for different value of N_a are far apart from each other and average centroid positions are also dependent upon the value of N_a.

For lightly doped substrate, inversion width and thus the potential well is much wider than that of highly doped substrate. As a result, the centroid is formed much deeper into the substrate for lower value of N_a. Increase in V_G will cause further bending of E_C and under strong inversion, induced potential at the surface will become similar irrespective of N_a leading to partial merging of the quantum wells, as shown in Fig. 2. As a result, the energy separation of respective energy states for different N_a will be reduced. Thus, the position of average charge
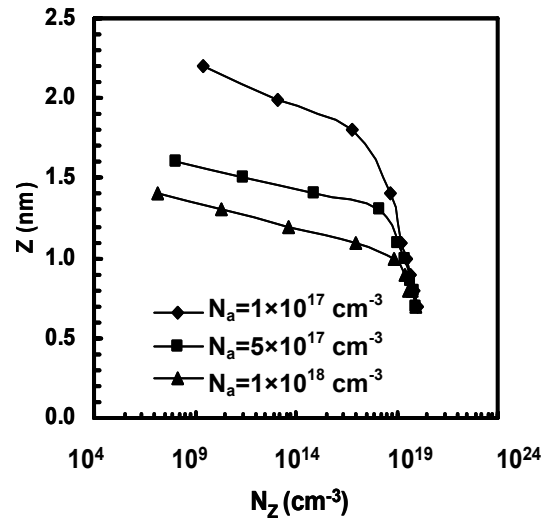

Fig. 4. Plots of variation of surface potential with applied gate voltage for different doping concentrations.

Fig. 5. Plots of centroid positions as a function of N_Z for three different substrate doping concentrations.

centroids at strong inversion, as shown in Fig. 5, will become almost independent of substrate doping concentrations.

Fig. 6 shows the simulated C-V characteristics for MOS structures with and without the effects of charge centroid for four different oxide thicknesses, such as, 1.5 nm, 2.5 nm, 3.5 nm, and 5 nm for a substrate doping concentration of 5×10^{17} cm^{-3}. It is evident from the plots that the quantum mechanically simulated inversion capacitance values are less than their classical counterpart and this difference increases for MOS structures with thinner dielectric layer. Presence of charge centroid and variation of its position as a function of applied voltage, effectively change the oxide thickness of the MOS structures. As a result, the measured inversion capacitance is a series combination of the inversion capacitance and the capacitance due to the depleted zone. This will depend on the oxide thickness t_{ox}, and substrate doping concentration N_a, since these will alter the distribution of inversion carrier concentration and centroid position as has been seen from Fig. 2 to Fig. 5. Oxide thickness of a MOS structure is extracted from the experimental data by using the expression,

$$t_{ox} = \frac{\varepsilon_{ox}}{C_{ox}} \qquad (14)$$

However, application of Eq. (14) to measured capacitance

data will result in the effective oxide thickness (EOT) instead of its exact value. The effective oxide thickness (EOT) has been considered as the modified thickness in presence of charge centroid apparent due to surface quantization effect.

Therefore, an error will occur during the extraction of the oxide thickness and that can be expressed as,

$$\left(\frac{(t_{ox})_{measured} - t_{ox}}{t_{ox}} \times 100 \right) \qquad (15)$$

Hence, a correction factor to eliminate the effect of charge centroid can be written as,

$$\left(\frac{EOT - t_{ox}}{EOT} \times 100 \right) \qquad (16)$$

Fig. 7 shows the variation of correction factor as a function of oxide thickness for different substrate doping concentrations. The value of correction factor increases with the decrease of oxide thickness and increase of doping concentration. Position of charge centroids for lightly doped MOS structures, as seen from Fig. 5, are shifted deeper into the semiconductor from SiO$_2$/Si interface. Thus, the width of the virtually depleted region is much higher for lightly doped substrate and at the same time, effective carier concentration in this region is higher compared to their highly doped counterparts. Therefore, under strong inversion, for a particular t_{ox},
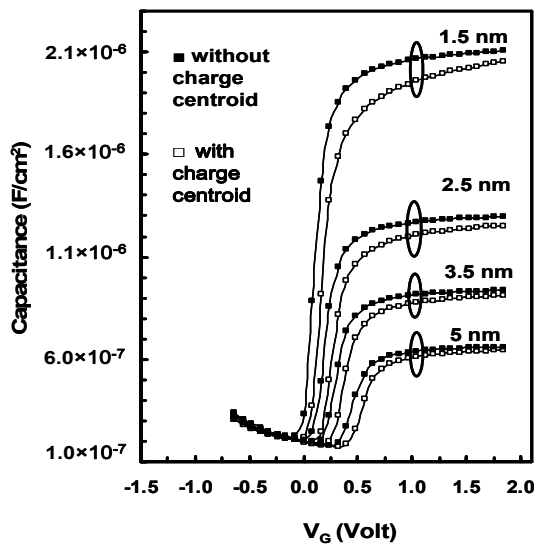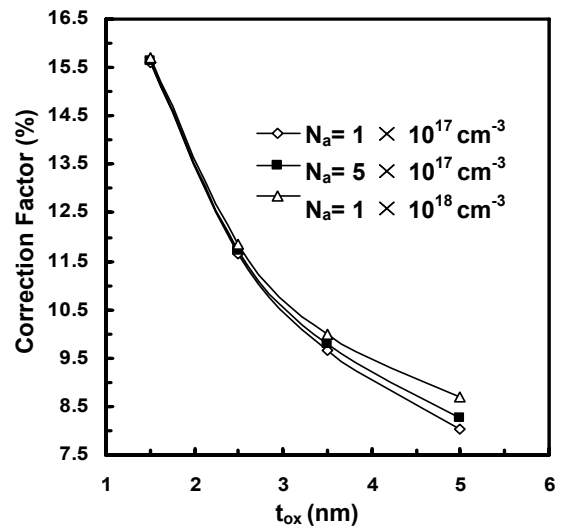

Fig. 6. Variation of inversion capacitance with gate voltage (V_G) for four different oxide thicknesses, with a substrate doping concentration of 5×10^{17} cm^{-3}.

Fig. 7. Plots of correction factor with oxide thickness (t_{ox}) for three different substrate doping concentrations.

EOT increases with substrate doping concentration leading to a higher value of correction factor. Thus, the error factors are observed to be more crucial for highly doped substrates with thinner gate oxides.

IV. CONCLUSIONS

In this paper, a correction factor for the ultra-thin oxide layer of a conventional Si MOS structure has been developed. Variation of charge centroid position has been analysed for different gate voltages. Variation of surface potential, conduction band bending, position of discrete energy states, and the degree of quantization have been studied in detail to analyse the quantization effect in MOS devices with ultra-thin gate oxides. It has been observed that, as a result of severe surface quantization effect, consideration of error factors to extract exact thickness of the gate oxide layer is more crucial for highly doped substrates with ultra-thin gate oxides.

ACKNOWLEDGMENTS

REFERENCES

[1] Maiti CK, Chattopadhyay S, Bera LK. Strained Si Heterostructure Field Effect Devices, Taylor and Francis; 2007.

[2] 2007 Edition of the ITRS, http://public.itrs.net

[3] Stern F, Howard W E. Properties of Semiconductor Surface Inversion Layers in the Electric Quantum Limit. Phy. Rev.1967; 163:816-35.

[4] Lopez-Villanuevap J A, Cartujo-Casinello, Banqueri J, Gamiz F, Rodriquez S. Effects of the inversion layer centroid on MOSFET behavior. IEEE Trans. Electron Devices 1997; 44:1915-22.

[5] Y. Tsividis. Operation and Modeling of the MOS Transistors. 2nd Edition. Oxford University Press; 2004.

[6] Hareland S A, Krishnamrthy S, Jallepalli S, Yeap Choh-Fei , Hasnat K, Tasch Al F , Maziar C M . A Computationally Efficient Model for Inversion Layer Quantization Effects in Deep Submicron N-Channel MOSFET's. IEEE Trans. Electron Devices 1996; 43: 90-6.

[7] Pacelli A, Spinelli A S, Perron L M. Carrier Quantization at Flat Bands in MOS Devices. IEEE Trans. Electron Devices 1999; 46: 383-87.

[8] SILVACO International, SILVACO, USA.

[9] Sze S M. Physics of semiconductor devices. 2nd ed. John Wiley & Sons, Inc; 1985.

[10] Stern F. Self Consistent Results for n-Type Si Inversion Layers. Phy.Rev.1972; 5:4891-99.

Munmun Dey received the B.Sc. degree (Hons.) in Electronics from University of Calcutta, Kolkata, in 2005 and M.Sc. degree in Electronic Science from Department of Electronic Science, University of Calcutta, Kolkata, in 2007, respectively. She is working as a research scholar at the Department of Electronic Science, University of Calcutta since 2007 and joined UPE fellowship from 2008. She is currently pursuing the Ph.D. degree in the Department of Electronic Science, University of Calcutta, Kolkata. Her research focuses modeling and simulation of conventional and strained silicon nano-scaled CMOS devices.

Sanatan Chattopadhyay received the B.Sc. degree (Hons.) in physics and the M.Sc. degree in electronics science from the University of Calcutta, Calcutta, India, in 1992 and 1994, respectively, and the Ph.D. degree from the University of Jadavpur, Calcutta, in 1999. From 1999 to 2001, he was a Post-Doctoral Fellow at the Singapore–Massachusetts Institute of Technology (MIT) Alliance, Singapore, before joining the University of Newcastle-upon-Tyne, Newcastleupon-Tyne, U.K. Currently he is working as a faculty in the Department of Electronic Science, University of Calcutta, India. His current research interests include the fabrication and characterization of strained Si/SiGe MOSFETs, nano-scale CMOS, low resistive silicides/germanides for contact metallization in very large scale integration, and growth and characterization of ultrathin dielectric gate materials.