

# The Influence of Space-Quantization Effects and Poly-Gate Depletion on the Threshold Voltage, Inversion Layer and Total Gate Capacitances in Scaled Si-MOSFETs

Dragica Vasileska

Department of Electrical Engineering

Arizona State University, Tempe AZ 85287-5706, USA

vasilesk@imap2.asu.edu

## ABSTRACT

We investigate the influence of space-quantization effect and poly-gate depletion on the inversion layer capacitance  $C_{inv}$ , total gate capacitance  $C_{tot}$  and threshold voltage  $V_T$  in scaled Si-MOSFETs. We also present an analytical expression for the total gate capacitance  $C_{tot}$  that uses classical charge description and takes into account the depletion of the poly-silicon gates. Our simulation results suggest that poly-gate depletion influences more the magnitude of  $C_{tot}$  than the quantum-mechanical charge description. For the threshold voltage  $V_T$  the situation is more complicated; it is affected by both poly-gate depletion and the quantum effects in the channel.

**Keywords:** scaled Si-MOSFETs, space-quantization, threshold voltage, inversion layer capacitance, total gate capacitance, poly-gate depletion, exchange-correlation effects.

## 1. INTRODUCTION

Since the invention of the transistor in 1947, solid state electronics has developed with a breathtaking pace, irreversibly transforming the information and computer technology. The continued growth of the silicon-based VLSI technology over time has led to orders of magnitude of improvement in performance, device density and cost. The Semiconductor Industry Association (SIA) projects a steady rate of improvement in most key parameters up to year 2012 when leading edge devices will employ 0.05  $\mu\text{m}$  gate lengths. This will enable more than 20 million logic gates to be integrated on a single chip, DRAMs to have capacity of over 16 billion bits, whereas SRAMs will be capable of storing more than a billion bits.

The successful scaling of MOSFETs toward shorter channel lengths requires thinner gate oxides and higher doping levels in order to have high drive currents and minimized short-channel effects [1,2]. For these state-of-the art devices, it was demonstrated a long time ago that, as the gate oxide thickness is scaled down to 10 nm and below, the total gate capacitance is smaller than the oxide capacitance due to the comparable values of the oxide and the inversion layer capacitances. As a consequence, the device transconductance is degraded relative to the expectations of the scaling theory [3]. The inversion layer capacitance was also identified as being the main cause of the second-order thickness-dependence of MOSFET's  $I$ - $V$  characteristic [4]. The finite inversion layer thickness was estimated experimentally by Hartstein and Albert [5]. The high levels of substrate doping needed in deep-submicrometer MOS devices, that lead to a pronounced quasi-two-dimensional (Q2D) nature of the carrier transport were found responsible for the increased threshold voltage and decreased channel mobility, and a simple analytical model that accounts for these effects was proposed [6,7]. The two physical origins of the inversion layer capacitance (due to finite density of states and due to finite inversion layer thickness) were demonstrated experimentally by Takagi and Toriumi [8]. A computationally efficient three-subband model, that predicts both the quantum-mechanical effects in the electron inversion layers and the electron distribution within the inversion layer, was proposed and implemented into the PISCES simulator [9]. The influence of the image and exchange-correlation effects on the inversion layer and total gate capacitance was studied by Vasileska *et al.* [10]. Very recently, it was also pointed out that the depletion of the poly-silicon gates considerably influences the magnitude of  $C_{tot}$  [11]. However, nobody has yet examined the influ-

ence of the quantum-mechanical space-quantization effect and poly-gate depletion on the threshold voltage or on the inversion layer capacitance, two very important issues that we focus on here.

## 2. NUMERICAL CALCULATION OF $C_{inv}$ AND $C_{tot}$

It has been known for many years that carriers in the inversion layer of a Si MOSFET are confined by the barrier between the semiconductor-oxide interface on one side and the band-bending of the conduction band on the other side. Since the average thickness of the inversion layer is comparable to the de Broglie wavelength of the electrons, this confinement is sufficient to produce quantization in the direction normal to the oxide-semiconductor interface. The space quantization effect is very important in determining the number of carriers in the inversion layer for devices with very high substrate doping, representative of the state-of-the art technology.

From the above discussion it is obvious that an accurate description of the charge in the inversion layer of deep-submicrometer devices and, therefore, the magnitude of the total gate capacitance  $C_{tot}$  requires a self-consistent solution of the 1D Poisson

$$\frac{\partial}{\partial z} \left[ \varepsilon(z) \frac{\partial \varphi}{\partial z} \right] = -e \left[ N_D^+(z) - N_A^-(z) + p(z) - n(z) \right], \quad (1)$$

and the 1D Schrödinger equation

$$\left[ -\frac{\hbar^2}{2m_i^\perp} \frac{\partial^2}{\partial z^2} + V_{eff}(z) \right] \Psi_{ij}(z) = E_{ij} \Psi_{ij}(z). \quad (2)$$

In (1) and (2),  $\varphi(z)$  is the electrostatic potential,  $\varepsilon(z)$  is the spatially dependent dielectric constant,  $N_D^+(z)$  and  $N_A^-(z)$  are the ionized donor and acceptor concentrations,  $n(z)$  and  $p(z)$  are the electron and hole densities,  $V_{eff}(z)$  is the effective potential energy term,  $m_i^\perp$  is the effective mass normal to the semiconductor-oxide interface of the  $i$ -th valley, and  $E_{ij}$  and  $\Psi_{ij}(z)$  are the energy level and the corresponding wavefunction of the electrons residing in the  $j$ -th subband from the  $i$ -th valley. The effective potential energy term  $V_{eff}(z)$  in the 1D Schrödinger equation equals the sum of the Hartree [ $V_H = -e\varphi(z)$ ], image [ $V_{im}(z)$ ] and exchange-correlation [ $V_{exc}(z)$ ] terms. The Hartree term represents

the solution of the 1D Poisson equation. The image term, which arises because of the different dielectric constants of the semiconductor and the oxide, is calculated from

$$V_{im}(z) = \frac{e^2}{16\pi\epsilon_{sc}z} \frac{\epsilon_{sc} - \epsilon_{ox}}{\epsilon_{sc} + \epsilon_{ox}}, \quad (3)$$

where  $\epsilon_{sc}$  and  $\epsilon_{ox}$  are the semiconductor and the oxide dielectric constants, respectively. When evaluating the exchange-correlation corrections to the chemical potential, we have relied on the validity of the density functional theory (DFT) of Hohenberg and Kohn [12], and Kohn and Sham [13]. According to DFT, the effects of exchange and correlation can be included through a one-particle exchange-correlation term  $V_{exc}[n(z)]$ , defined as a functional derivative of the exchange-correlation part of the ground-state energy of the system with respect to the electron density  $n(z)$ . In the local density approximation (LDA), one replaces the functional  $V_{exc}[n(z)]$  with a function  $V_{exc}[n(z)] = \mu_{exc}[n_0 = n(z)]$ , where  $\mu_{exc}$  is the exchange-correlation contribution to the chemical potential of a homogeneous electron gas of density  $n_0$ , which is taken to be equal to the local electron density  $n(z)$  of the inhomogeneous system. In our model, we use the LDA and approximate the exchange-correlation potential energy term  $V_{exc}(z)$  by an interpolation formula developed by Hedin and Lundqvist [14]

$$V_{exc}(z) = -\frac{e^2}{8\pi\epsilon_{sc}b} \left[ 1 + 0.7734x \ln\left(1 + \frac{1}{x}\right) \right] \left( \frac{2}{\pi\alpha r_s} \right), \quad (4)$$

which is accurate over a large density range. In (4),  $\alpha = (4/9\pi)^{1/3}$ ,  $x = x(z) = r_s/21$ ,  $r_s = r_s(z) = [4\pi b^3 n(z)/3]^{-1/3}$ , and  $b = 4\pi\epsilon_{sc} \hbar^2 / m^* e^2$ . Exchange and correlation effects tend to lower the total energy of the system, and as discussed later, lead to non-uniform shift of the energy levels and repopulation of the various subbands. The enhancement of the exchange-correlation contribution to the energy predominantly affects the ground subband of the occupied valley; the unoccupied subbands of the same valley are essentially unaffected [15]. As a result, noticeable increase in the energy of the inter-subband transitions can be observed at high electron densities.

In all the calculations presented here, we assume that the  $SiO_2/Si$  interface is parallel to the [100] plane. For this particular case, the six equivalent minima of the bulk silicon conduction band split into

two sets of subbands (Fig. 1). The first set ( $\Delta_2$ -band) consists of the two equivalent valleys with in-plane effective mass  $m_{\parallel}=0.19m_0$  and perpendicular effective mass  $m_{\perp}=0.91m_0$ . The second set ( $\Delta_4$ -band) consists of the four equivalent valleys with  $m_{\parallel}=0.42m_0$  and  $m_{\perp}=0.19m_0$ . The energy levels associated with the  $\Delta_2$ -band comprise the so-called *unprimed* ladder of subbands, whereas those associated with the  $\Delta_4$ -band comprise the *primed* ladder of subbands.

The self-consistent solution of the 1D Schrödinger-Poisson problem is obtained in the following way [16]: We start with some initial guess for the electrostatic potential and use it to solve the 1D Schrödinger equation numerically [17]. For each iteration (guess for the energy), the 1D Schrödinger equation is integrated leftward ( $\psi^<$ ) and rightward, and the two solutions are matched at the turning point. An eigenvalue is indicated by the continuity of the logarithmic derivative at the matching point, i.e. when the matching tolerance is less than some predetermined value. After we determine the eigenfunctions and the eigenvalues that characterize the electrons in the inversion layer, the inversion layer electron density appearing in the 1D Poisson equation is obtained by summing over all subbands to get

$$n(z) = \sum_{i,j} N_{ij} \psi_{ij}^2(z) = \sum_{i,j} g_i \frac{m_i^{\parallel} k_B T}{\pi \hbar^2} \ln \left[ 1 + \exp \left( \frac{E_F - E_{ij}}{k_B T} \right) \right] \psi_{ij}^2(z). \quad (5a)$$

In (5a),  $E_F$  is the Fermi level,  $k_B$  is the Boltzmann constant,  $T$  is the temperature,  $m_i^{\parallel}$  is the in-plane effective mass of the  $i$ -th band,  $N_{ij}$  is the sheet-charge density corresponding to the  $j$ -th subband from the  $i$ -th band, and  $g_i$  ( $g_1=2$  for the  $\Delta_2$ -band, and  $g_2=4$  for the  $\Delta_4$ -band) is the band degeneracy. It is important to note that the inversion layer electrons are treated quantum-mechanically only when confined by the surface-field. If otherwise, or if we relied on the validity of the classical description of the inversion layer electron density, we skip the solution of the 1D Schrödinger equation and use

$$n(z) = N_C F_{1/2} \left[ \frac{E_F - E_C(z)}{k_B T} \right], \quad (5b)$$

where  $N_C$  is the effective density of states of the conduction band. For holes, which are always treated classically, we use

$$p(z) = N_V F_{1/2} \left[ \frac{E_C(z) - E_G - E_F}{k_B T} \right], \quad (6)$$

where  $N_V$  is the effective densities of states of the valence band, and  $E_G$  is the semiconductor bandgap. For the evaluation of the Fermi-Dirac integrals, which appear in (5b) and (6), we use the analytical approximation due to Bednarczyk and Bednarczyk [18]

$$F_{1/2}(x) = \left[ e^{-x} + 3\sqrt{\pi}/4\nu^{3/8} \right]^1, \quad (7a)$$

where

$$\nu(x) = x^4 + 50 + 33.6x \left\{ 1 - 0.68 \exp \left[ -0.17(x+1)^2 \right] \right\}. \quad (7b)$$

In both our numerical and our analytical model (described in the next section), the poly-silicon gates are modeled as heavily-doped single-crystal silicon. Both the electrons and holes are treated classically using (5b) and (6), and assuming general Fermi-Dirac statistics, valid for degenerate semiconductors.

After we update the electron and hole concentrations in the semiconductor and/or the poly-silicon gates, we resolve numerically the 1D Poisson equation for the electrostatic potential using finite-difference discretization scheme and LU decomposition method. We then solve the 1D Schrödinger equation to find the updated values for the electron density at each mesh point, and repeat the above-described procedure until self-consistent solution is found. It is important to note that the potential energy profile for the next iteration is obtained by using fixed-convergence factor scheme for the first 10 iterations and the extrapolated convergence-factor scheme thereafter. The error criterion for the convergence of the self-consistent field iterations is that the absolute value of the difference between the input and output potentials at each mesh point is less than 0.01 mV.

At self-consistency, i.e. once we determine the self-consistent results for the variation of the charge distribution on the semiconductor side of the MOS capacitor as a function of the gate voltage  $V_G$ , we proceed with the calculation of the total gate capacitance  $C_{tot}$  (Fig. 2). We determine  $C_{tot}$  by differentiating the total induced charge density in the channel with respect to  $V_G$ . In contrast to some previous studies [3,19], where  $C_{inv}$  was approximated with  $\epsilon_{sc} / \langle z \rangle_{av}$ , where  $\langle z \rangle_{av}$  is the centroid of the electron

density distribution, here we calculate the inversion layer capacitance by differentiating the total sheet charge density

$$N_s = \sum_{i,j} N_{ij} \quad (8)$$

with respect to the surface potential [10]. The depletion-layer capacitance  $C_{depl}$  and the poly-gate capacitance  $C_{poly}$  are evaluated in an analogous manner.

### 3. ANALYTICAL MODEL

In our extended analytical model, to find the poly-gate capacitance  $C_{poly}$ , we essentially follow the approach in [20]. Briefly, for a given oxide field  $E_{ox}$ , we first solve for the surface potential  $\varphi_p$  by solving

$$E_{ox} = \frac{\epsilon_{sc}}{\epsilon_{ox}} \sqrt{\frac{2eN_D}{\epsilon_{sc}}} G(\varphi_p), \quad (9)$$

where

$$G(\varphi_p) = \left[ \varphi_p + \frac{N_C}{N_D} \int_0^{-\varphi_p} d\varphi F_{1/2} \left( \frac{e\varphi - \Delta E_{CP}}{k_B T} \right) - \frac{N_V}{N_D} \int_0^{-\varphi_p} d\varphi F_{1/2} \left( \frac{\Delta E_{CP} - E_G - e\varphi}{k_B T} \right) \right]^{1/2}. \quad (10)$$

We then calculate the poly-gate capacitance from

$$C_{poly} = \frac{\sqrt{eN_D \epsilon_{sc}}}{\sqrt{2}G(\varphi_p)} \left[ 1 - \frac{N_C}{N_D} F_{1/2} \left( -\frac{e\varphi_p + \Delta E_{CP}}{k_B T} \right) + \frac{N_V}{N_D} F_{1/2} \left( \frac{\Delta E_{CP} - E_G + e\varphi_p}{k_B T} \right) \right]. \quad (11)$$

In (9-11),  $N_D$  is the doping of the poly-silicon gates, and  $\Delta E_{CP} = (E_C - E_F)_{poly}$  is found from the charge neutrality condition in the poly-silicon away from the interface. The corresponding gate voltage is found from

$$V_G = \varphi_{sc} + \varphi_p + (\Delta E_{CP} - \Delta E_{CB})/e + V_{ox}, \quad (12)$$

where  $\varphi_{sc}$  is the surface potential in the semiconductor side of the semiconductor/oxide interface,  $\Delta E_{CB} = (E_C - E_F)_{sc}$  is found from the charge neutrality condition in the semiconductor bulk region, and  $V_{ox} = E_{ox} t_{ox}$

(where  $t_{ox}$  is the oxide thickness) is the voltage drop in the oxide. The total gate capacitance per unit area is then evaluated as a series combination of  $C_{poly}$ ,  $C_{ox}=\epsilon_{ox}/t_{ox}$  and  $C_{inv}$ .

## 4. RESULTS AND DISCUSSION

### 4.1. Inversion Layer, Poly-Gate and Total Gate Capacitance

To demonstrate the existence of the two physical origins of the inversion layer capacitance  $C_{inv}$ , discussed in Refs. [8] and [10], in Fig. 3 we show the variation of  $C_{inv}$  with inversion charge density  $N_s$  in the channel of a MOS capacitor with substrate doping  $N_A=5\times 10^{17}$  cm<sup>-3</sup>, oxide thickness  $t_{ox}=4$  nm and metal gates. Exchange-correlation and image contributions to the effective potential energy term  $V_{eff}(z)$ , appearing in the 1D Schrödinger equation, have not been included in these simulations.

The pronounced double-slope behavior of the quantum-mechanically calculated  $C_{inv}$  comes from the fact that the total inversion layer capacitance can be represented as a series capacitance of two contributions. The first contribution is classical and comes from the finite density of states, i.e. due to the fact that a finite change in the surface potential is always necessary to increase  $N_s$  (inset of Fig. 3), which, in turn, leads to finite value for  $C_{inv}$ . This term dominates at low values of  $N_s$  (low gate voltages). The second contribution to  $C_{inv}$  is due to the finite inversion layer thickness, which effectively increases the oxide thickness in terms of the total gate capacitance, thus providing an additional capacitance component. As shown in Fig. 4, the origin of this contribution is the quantum-mechanical space quantization effect in the inversion layer that leads to larger inversion layer thickness and larger displacement of the charge from the interface when compared to what SC calculations give (inset of Fig. 4). This term dominates at large gate voltages, where the inversion charge density  $N_s$  significantly influences the band bending and leads to a steeper rise of the conduction band near the  $SiO_2/Si$  interface.

The variation of the inversion layer capacitance  $C_{inv}$  with gate voltage  $V_G$  is shown in Fig. 5. From the results shown, it is obvious that the quantum-mechanical charge description, which leads to approximately three times larger average displacement of the inversion layer electrons from the  $SiO_2/Si$  interface (inset of Fig. 4), gives significantly smaller values for the inversion layer capacitance. The inclusion of

poly-gate depletion only slightly modifies the magnitude of  $C_{inv}$ . Contrary to  $C_{inv}$ , which increases with increasing  $V_G$  due to the closer confinement of the carriers to the interface, the poly gate capacitance (inset of Fig. 5) decreases due to the increase of the width of the depletion region. It is important to note that for  $V_G = 2.5$  V,  $C_{poly}$  is approximately two times smaller than  $C_{inv}$ . This, as shown in Fig. 6, leads to a larger reduction of the total gate capacitance in strong inversion. For instance, for capacitors with metal gates and applied gate bias  $V_G = 2.5$  V, the SC (QM) models predict that the total gate capacitance  $C_{tot}$  is 95% (91%) of  $C_{ox}$ . For capacitors with poly-gates, the magnitude of  $C_{tot}$  is reduced down to 80% (78%) of  $C_{ox}$  when using SC (QM) description of the charge in the channel. These observations imply that, even though space-quantization effect considerably reduces the magnitude of  $C_{inv}$ , it leads to no more than 10% degradation of  $C_{tot}$  in strong inversion. Therefore, dominant degradation mechanism for the total gate capacitance in strong inversion is the depletion of the poly-silicon gates.

From the results shown in Fig. 6, one can also argue that the poly-gate capacitance has a significantly larger influence on the inversion charge density  $N_s$  in the channel than the quantum-mechanical space-quantization effect. Simulation results shown in Fig. 7 clearly show that this is not the case. Both the quantum-mechanical space-quantization effect and the poly-gate depletion increase the required gate voltage to achieve a certain inversion charge density  $N_s$  in the channel. The shift in the gate voltage increases with increasing  $N_s$ . For example, when using a classical model for the charge description and metal gates (SCNP), gate voltage  $V_G = 2.02$  V is needed to achieve  $N_s = 8 \times 10^{12} \text{ cm}^{-2}$  in the inversion layer. When using quantum-mechanical model (QMNP), gate voltage  $V_G = 2.21$  V is required to achieve the same value of  $N_s$  (Higher gate voltage is required because of the smaller density of states function of the Q2D system). The addition of poly-gate depletion (QMWP) further increases  $V_G$  to 2.43 V, and this is a consequence of the finite voltage drop across the depletion layer in the poly-silicon gates. This last value is approximately 0.41 V higher than what the classical model with no poly-gate depletion would give.

The excellent agreement between our self-consistent (numerical) results for  $C_{inv}$ ,  $C_{poly}$  and  $C_{tot}$  with those obtained with our extended analytical model (symbols in Figs. 5 and 6) suggest that our analytical model can be quite successfully used in predicting the errors which will result when neglecting the

effect of poly-gate depletion on the terminal device characteristics. However, in contrast to the QM space-quantization effect, which in strong inversion leads to almost constant deviation of  $C_{tot}$  from  $C_{ox}$ , the depletion of the poly-silicon gates is a bias-dependent phenomenon, and one cannot use a single correction factor to accurately predict  $C_{tot}$  over the entire gate-voltage range. For devices with metal gates, a single correction factor can be used to describe the deviation of the total gate capacitance from the oxide capacitance in strong inversion. This can further be used in correctly estimating the oxide thickness from the low-frequency CV measurements. If this correction factor is not taken into account, the electrically inferred oxide thickness will always be larger than the physical oxide thickness. As shown in Fig. 8, the discrepancy between the two can be as large as 9% for the MOS capacitors with  $t_{ox} = 3$  nm. If the same device is driven into weak inversion, the relative error would have been about 12 % [10]. The results shown in Fig. 8 were obtained in the following manner: for each MOS capacitor, we first calculated the total gate capacitance in strong inversion conditions. Then, the relative error between the electrically inferred oxide thickness and the physical oxide thickness was calculated using

$$\frac{t_{ox,eff} - t_{ox}}{t_{ox}} = \frac{C_{ox}}{C_{tot}} - 1 . \quad (13)$$

In Fig. 9, we show simulated  $C_{tot}$  to oxide capacitance  $C_{ox}$  for metal/*p*-substrate and *n*+poly/*p*-substrate MOS capacitors, as a function of the physical oxide thickness  $t_{ox}$  and the doping of the polysilicon gates  $N_D$ , assuming  $V_G = 3$  V. The high value for  $V_G$ , used here, may overestimate the severity of the bias dependent attenuation for thinner oxides, but a consistent value for  $V_G$  is useful for the purpose of tabulating the simulated results. The results shown clearly demonstrate that classical charge model and Maxwell-Boltzmann (non-degenerate) statistics are clearly inadequate for oxide thickness below 10 nm. Even use of Fermi-Dirac statistics in the classical charge description can lead to significant errors in the estimate of the total gate capacitance for devices with metal gates and oxide thickness less than 5 nm, due to the higher surface fields and, therefore, pronounced quantum-mechanical size-quantization effect in the channel. For example, the classical model that uses Maxwell-Boltzmann (Fermi-Dirac) statistics predicts that, for the device with  $t_{ox} = 1$  nm,  $C_{tot}/C_{ox} = 0.983$  (0.882). On the other hand, the quantum-

mechanical model predicts that  $C_{tot}/C_{ox}=0.795$ , which leads to relative error of 23.65 (10.94) %. As previously noted, the depletion of the poly-silicon gates will further degrade the total gate capacitance.

Finally, in Fig. 10 we show the influence of the image and exchange-correlation effects on the magnitude of the inversion layer and total gate capacitances. The MOS capacitors considered have  $N_A = 10^{18} \text{ cm}^{-3}$ ,  $t_{ox} = 1.5 \text{ nm}$  and metal (aluminum) gates. We find that the inclusion of image and many-body exchange-correlation effects increases  $C_{inv}$  (Fig. 10a), and this increase is quite large at low gate voltages (weak inversion condition). For example, for  $V_G = 0.5 \text{ V}$ , simulation results for  $C_{inv}$  are higher by approximately 39.4% when these two effects are included in the calculation of the subband structure. The increase of  $C_{inv}$  is related to the fact that exchange-correlation effects lead to overall reduction of the total energy of the system and, therefore, shift all the energy levels downward. Since the shift is the largest for the lower-lying subbands, for which the average displacement of the carriers from the interface is the smallest, this will lead to larger values for the inversion layer capacitance  $C_{inv}$ . However, their influence is found to decrease proportionally with increasing the gate voltage  $V_G$  (increasing the surface field) when most of the carriers will populate the lower-lying subbands anyway. The influence of the image effect is found to be relatively constant with  $V_G$ . What concerns the total gate capacitance (Fig. 10b), we find that the inclusion of image and exchange-correlation effects improves the situation, i.e. it leads to smaller degradation of  $C_{tot}$  in weak inversion. Again, the influence of the image and many-body exchange-correlation effects decreases with increasing  $V_G$ .

## 4.2 Threshold voltage

Lowering supply voltages in deep-submicron devices must be accompanied by lowering of the threshold voltage to maintain optimum circuit performance. A careful investigation of the influence of the depletion of the poly-silicon gates and space-quantization effects on the magnitude of the threshold voltage  $V_{th}$  is thus mandatory. To accomplish this task, we do a systematic study of the shift in the threshold voltage due to space quantization effects in the channel region of the device and the depletion of the poly-silicon

gates. A range of substrate doping densities and doping of the poly-silicon gates, representative of the deep-submicron technology, is considered in this study.

The linear region threshold voltage shift between the QM and SC predictions for a device with  $N_A=5\times 10^{17} \text{ cm}^{-3}$  and  $t_{ox}=4 \text{ nm}$  (device from Fig. 5) as a function of the doping of the poly-silicon gate is shown in Fig. 11. The threshold voltage  $V_{th}$  equals the gate voltage for which  $Q_{inv}=10^{-3}Q_{depl}$ . As expected, the QM description of the charge in the channel increases  $V_{th}$  and the shift in the threshold voltage is about 74 mV. This is due to the fact that the QM picture differs from the SC one in two ways: First, the energy spectrum is not continuous, but consists of discrete energy levels which, in turn, reduces the DOS function. Second, the energy of the ground subband from the unprimed ladder of subbands does not coincide with the bottom of the conduction band (Fig. 12), and the energy difference  $\Delta E = E_{11} - E_C$  increases with increasing substrate doping (not shown on this figure). The depletion of the poly-silicon gate, due to insufficient doping, further increases the threshold voltage. The additional shift in the threshold voltage due to the inclusion of the poly-gate depletion can be as large as 68 mV for  $N_D = 10^{19} \text{ cm}^{-3}$ , and drops down to about 18 mV for  $N_D = 2\times 10^{20} \text{ cm}^{-3}$ .

The linear region threshold voltage shift for the device with  $t_{ox}=4 \text{ nm}$ ,  $N_D=10^{20} \text{ cm}^{-3}$ , and different substrate doping is shown in Fig. 13. Also shown in this figure are the van Dort *et al.* [6] experimental data for a device with metal gates and oxide thickness  $t_{ox}=14 \text{ nm}$ . Very close agreement between the experimentally derived threshold voltage shifts and our simulation results for the device with 14 nm thick oxide can be observed. A major difference from the results shown in Fig. 11 is that the inclusion of both the QM effects in the channel and poly-gate depletion leads to strong dependence of the threshold voltage shift upon the substrate doping  $N_A$ . For example, for a device with  $t_{ox} = 4 \text{ nm}$ ,  $N_A = 10^{18} \text{ cm}^{-3}$  and  $N_D = 10^{20} \text{ cm}^{-3}$ , the inclusion of the quantum-mechanical space-quantization effect leads to a threshold voltage shift of about 106 mV. The addition of poly-gate depletion leads to a further shift in the threshold voltage of about 34 mV. This observation, together with the results shown in Fig. 13, suggests that both a QM description of the charge density distribution in the channel and poly-gate depletion must be accounted for if accurate results for the threshold voltage are desired.

## **5. CONCLUSIONS**

In conclusion, we investigated the influence of the quantum-mechanical space-quantization effect and the poly-gate depletion on the inversion layer and total gate capacitances, as well as their influence on the threshold voltage. We have demonstrated that for devices with thin gate oxides, poly-gate depletion is the dominant degradation mechanism of the total gate capacitance. An analytical model for the calculation of the poly-gate capacitance, based on a SC charge description in the channel, was also presented, and the results obtained by using this model are in excellent agreement with our numerical self-consistent results. We have also demonstrated that image and exchange-correlation effects increase the inversion layer capacitance when the device is operating in weak and moderate inversion conditions, which leads to higher values for the total gate capacitance (smaller degradation). Their influence is found to be insignificant in strong inversion. Our simulation results for the threshold voltage imply that the omission of the quantum-mechanical space-quantization effect can lead to erroneous predictions for the threshold voltage for deep submicrometer devices.

## **ACKNOWLEDGEMENTS**

This work is supported in part by the National Science Foundation under Contract No. ECS-9802596. The author would like to thank Dr. D.K. Ferry and Dr. D.K. Schroder for the valuable discussions during the preparation of this manuscript.

## REFERENCES

1. R.H. Dennard, F.H. Gaensslen, H.-N. Yu, V.L. Rideout, E. Bassous, and A.R. leBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE J. Solid-State Circuits*, Vol. 9, pp. 256-268, 1974.
2. J.R. Brews, W. Fichtner, E.H. Nicollian, and S.M. Sze, "Generalized guide for MOSFET miniaturization," *IEEE Electron Dev. Lett.*, Vol. 1, pp. 2-4, 1980.
3. G. Baccarani, and M.R. Worderman, "Transconductance degradation in thin-oxide MOSFET's," in *Proceedings of the IEDM*, pp. 278-281, 1982.
4. M.-S. Liang, J.Y. Choi, P.-K. Ko, and C. Hu, "Inversion-layer capacitance and mobility of very thin gate-oxide MOSFET's," *IEEE Trans. Electron Devices*, Vol. 33, pp. 409-413, 1986.
5. A. Hartstein and N.F. Albert, "Determination of the inversion-layer thickness from capacitance measurements of metal-oxide-semiconductor field-effect transistors with ultrathin oxide layers," *Phys. Rev. B*, Vol. 38, pp. 1235-1240, 1988.
6. M.J. van Dort, P. H. Woerlee , A.J. Walker, C.A.H. Juffermans, and H. Lifka, "Influence of high substrate doping levels on the threshold voltage and the mobility of deep-submicrometer MOSFET's," *IEEE Trans. Electron Devices*, Vol. 39, pp. 932-938, 1992.
7. M. J. van Dort, P. H. Woerlee and A. J. Walker, "A simple model for quantization effects in heavily-doped silicon MOSFET's at inversion conditions," *Solid-State Electronics*, Vol. 37, pp. 411-414, 1994.
8. S. Takagi and A. Toriumi, "Quantitative understanding of inversion-layer capacitance in Si MOSFET's," *IEEE Trans. Electron Devices*, Vol. 42, pp. 2125-2130, 1995.
9. S. A. Hareland, S. Krishnamurthy, S. Jallepalli, C.-F. Yeap, K. Hasnat, A. F. Tasch, Jr. and C. M. Maziar, "A computationally efficient model for inversion layer quantization effects in deep submicron *n*-channel MOSFET's," *IEEE Trans. Electron Devices*, Vol. 43, pp. 90-96, 1996.

10. D. Vasileska, D. K. Schroder and D. K. Ferry, "Scaled silicon MOSFET's: Degradation of the total gate capacitance," *IEEE Trans. Electron Devices*, Vol. 44, pp. 584-587, 1997.
11. K. S. Krisch, J. D. Bude and L. Manchanda, "Gate capacitance attenuation in MOS devices with thin gate dielectrics," *IEEE Electron Device Letters*, Vol. 17, pp. 521-524, 1996.
12. P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas," *Phys. Rev.*, Vol 136, pp. B864-B871, 1964.
13. W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects," *Phys. Rev.*, Vol 140, pp. A1133-A1140, 1965.
14. L. Hedin and B. I. Lundqvist, "Explicit Local Exchange-Correlation Potentials," *J. Phys. C: Solid St. Phys.*, Vol. 4, pp. 2064-2083, 1971.
15. D. Vasileska, P. Bordone, T. Eldridge, and D.K. Ferry, "Calculation of the average interface field in inversion layers using zero-temperature Green's function formalism," *J. Vac. Sci. Technol. B*, Vol. 13, pp. 1841-1847, 1995.
16. Our 1D Schrödinger-Poisson solver, called SCHRED, has been installed on the Purdue Semiconductor Simulation Hub ([www.ecn.purdue.edu/labs/punch](http://www.ecn.purdue.edu/labs/punch)) and is now accessible through the Web.
17. S.E. Koonin and D.C. Meredith, *Computational Physics* (Addison-Wesley Publishing Company, New York, 1990).
18. J.S. Blakemore, "Approximations for Fermi-Dirac integrals, especially the function  $F_{1/2}(\eta)$  used to describe electron density in a semiconductor," *Solid-State Electron.*, Vol. 25, pp. 1067-1076, 1982.
19. S. Luryi, "Quantum capacitance devices," *Appl. Phys. Lett.*, Vol. 52, pp. 501-503, 1988.
20. K. Chen, G. Zhang, J. Duster, C. Hu, J. Huang, Z. Liu and P. K. Ko, "MOSFET inversion layer capacitance model based on Fermi-Dirac statistics for wide temperature range," *Solid-State Electronics*, Vol. 41, pp. 507-509, 1997.

## FIGURE CAPTIONS

Figure 1: (a) Constant energy surfaces in Si together with the description of the  $\Delta_2$ - and  $\Delta_4$ -bands. Also shown are the appropriate transverse and in-plane masses for the two equivalent bands. (b) Schematics of the band-bending in MOS capacitors. Also shown are the energy levels belonging to the unprimed and primed ladder of subbands corresponding to the  $\Delta_2$ - and  $\Delta_4$ -bands, respectively. The first index describes the band (=1 for the  $\Delta_2$ -band, and =4 for the  $\Delta_4$ -band), whereas the second one refers to the appropriate energy level within the band.

Figure 2: Equivalent circuit that shows the various contributions to the total gate capacitance in a MOS capacitor. The effect of interface traps has been omitted in the present analysis. If included, it would lead to an additional capacitance component in parallel to the inversion layer and depletion layer capacitances.

Figure 3: Variation of the inversion layer capacitance with inversion charge density at T=300 K. In the inset we show the self-consistent results for the variation of the surface potential with  $N_s$  when using both SC and QM description of the electron density in the inversion layer. The surface potential is calculated using  $\varphi_s = (E_F - E_i)_{sc/ox} - (E_F - E_i)_{bulk}$ , where  $E_i$  is the intrinsic energy level.

Figure 4: Classically and quantum-mechanically calculated charge distribution in the inversion layer for the device from Fig. 3, for  $V_G=2.5$  V and T=300 K. In the inset, we show the variation of the centroid of the electron density distribution in the inversion layer when using SC and QM description of the inversion layer electrons.

Figure 5: Variation of  $C_{inv}$  with  $V_G$  when using SC and QM description of the charge in the channel, with (WP) and without (NP) the inclusion of the poly-gate depletion. We use  $N_A=5 \times 10^{17}$  cm<sup>-3</sup>,  $t_{ox}=4$  nm and  $N_D=5 \times 10^{19}$  cm<sup>-3</sup>. Also shown here are the semiclassical results obtained with

our analytical model (symbols). In the inset we show the variation of  $C_{poly}$  with  $V_G$  obtained by using the numerical model (solid lines) and our analytical model (symbols).

Figure 6: Simulated low-frequency  $CV$ -curves for the device from Fig. 5, which illustrate the change in the total gate capacitance  $C_{tot}$  due to finite  $C_{inv}$  and finite poly-gate capacitance  $C_{poly}$ . Symbols have the same meaning as in Fig. 5.

Figure 7: Variation of inversion layer electron density  $N_s$  with  $V_G$  for the device from Fig. 5.

Figure 8: Relative error between the electrically inferred oxide thickness and the physical oxide thickness for different values of  $t_{ox}$  for devices with metal gates and substrate doping  $N_A=5\times 10^{17}$   $\text{cm}^{-3}$ .

Figure 9: Simulated  $C_{tot}$  to oxide capacitance  $C_{ox}$  for metal/ $p$ -substrate and  $n+$ -poly/ $p$ -substrate MOS capacitors, as a function of the physical oxide thickness  $t_{ox}$  and the doping of the polysilicon gates  $N_D$ . We use  $V_G=3$  V.

Figure 10: (a) Inversion layer  $C_{inv}$ , and (b) total gate capacitance  $C_{tot}$  versus gate voltage  $V_G$  for the MOS capacitor with  $N_A = 10^{18}$   $\text{cm}^{-3}$ ,  $t_{ox} = 1.5$  nm, and metal (aluminum) gates.

Figure 11: Linear region threshold voltage shift between the QM and the SC predictions versus  $N_D$ . We use  $N_A=5\times 10^{17}$   $\text{cm}^{-3}$  and  $t_{ox}=4$  nm.

Figure 12: Variation of the energies of the lowest six subbands (four from the  $\Delta_2$ -band and two from the  $\Delta_4$ -band) and the position of the Fermi level with inversion charge density in the channel for the device from Fig. 5. All energies are measured from the bottom of the conduction band at the  $\text{SiO}_2/\text{Si}$  interface.

Figure 13: Linear region threshold voltage shift between the QM and the SC predictions versus  $N_A$ .

Figure 1: Vasileska

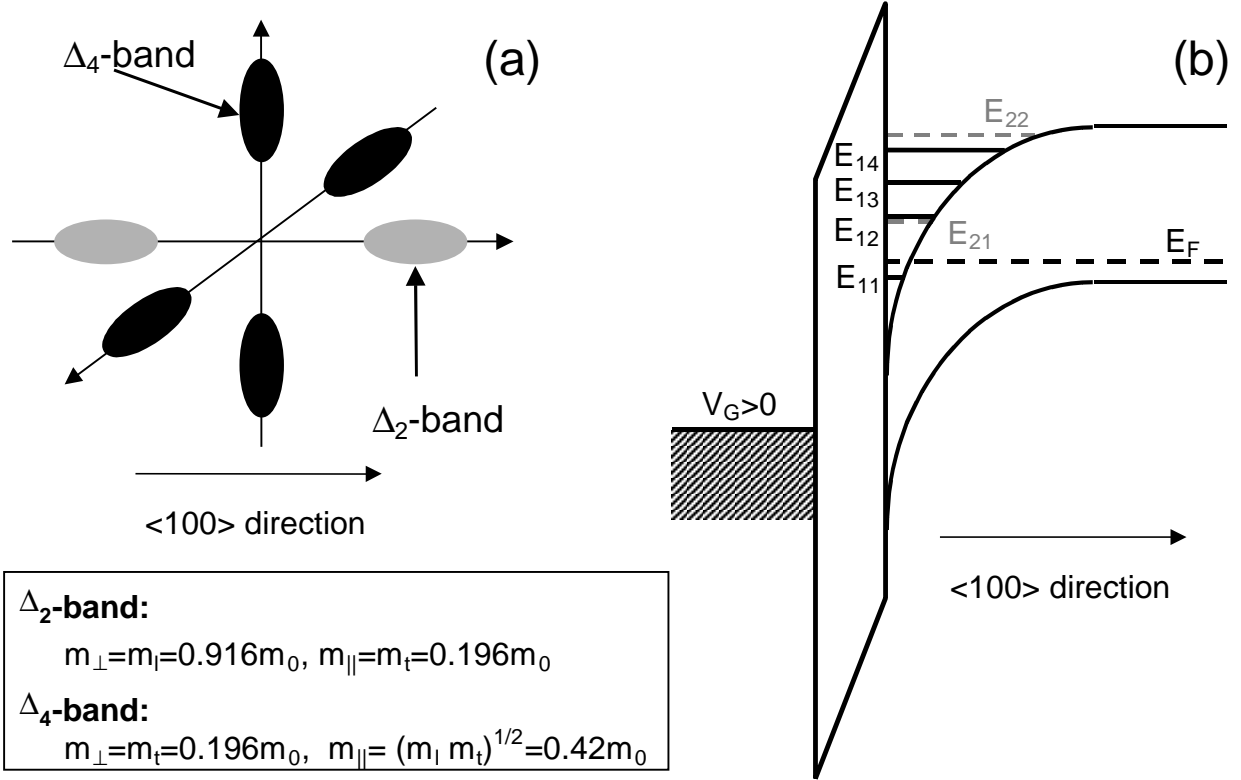


Figure 2: *Vasileska*

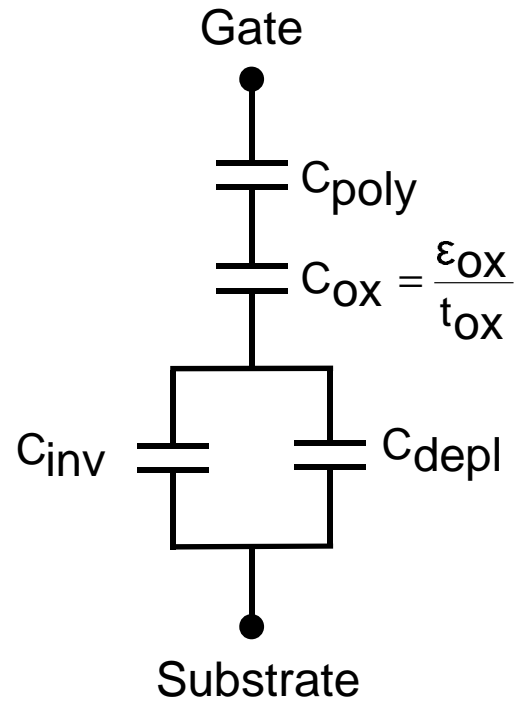


Figure 3: Vasileska

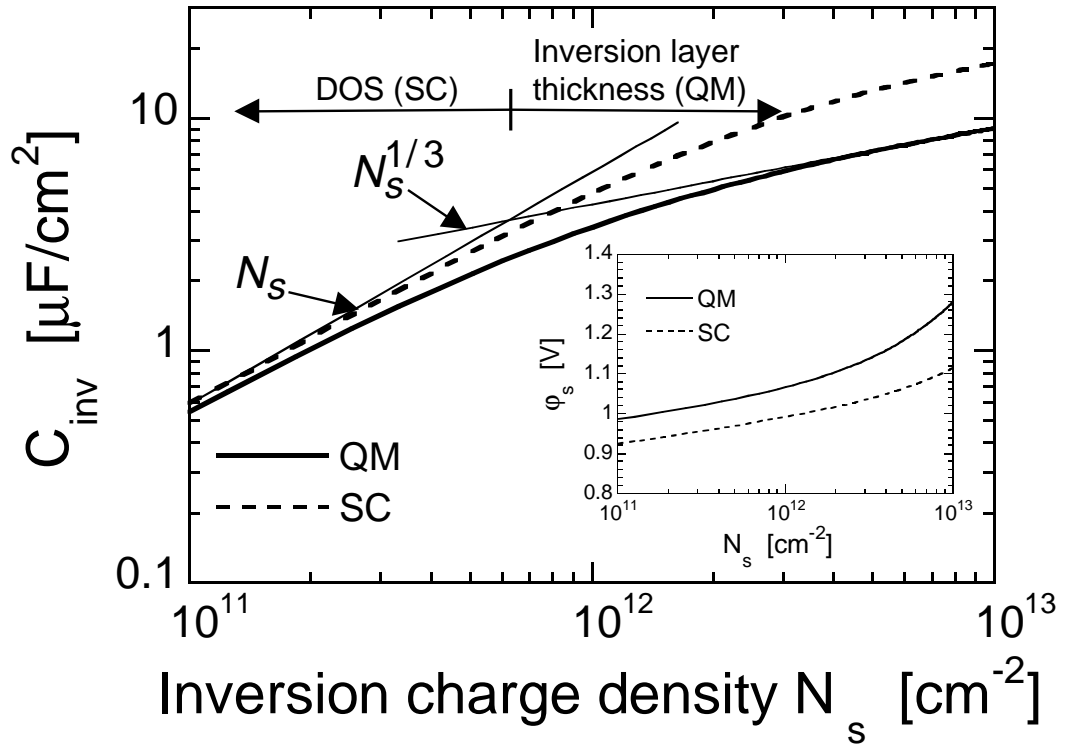


Figure 4: *Vasileska*

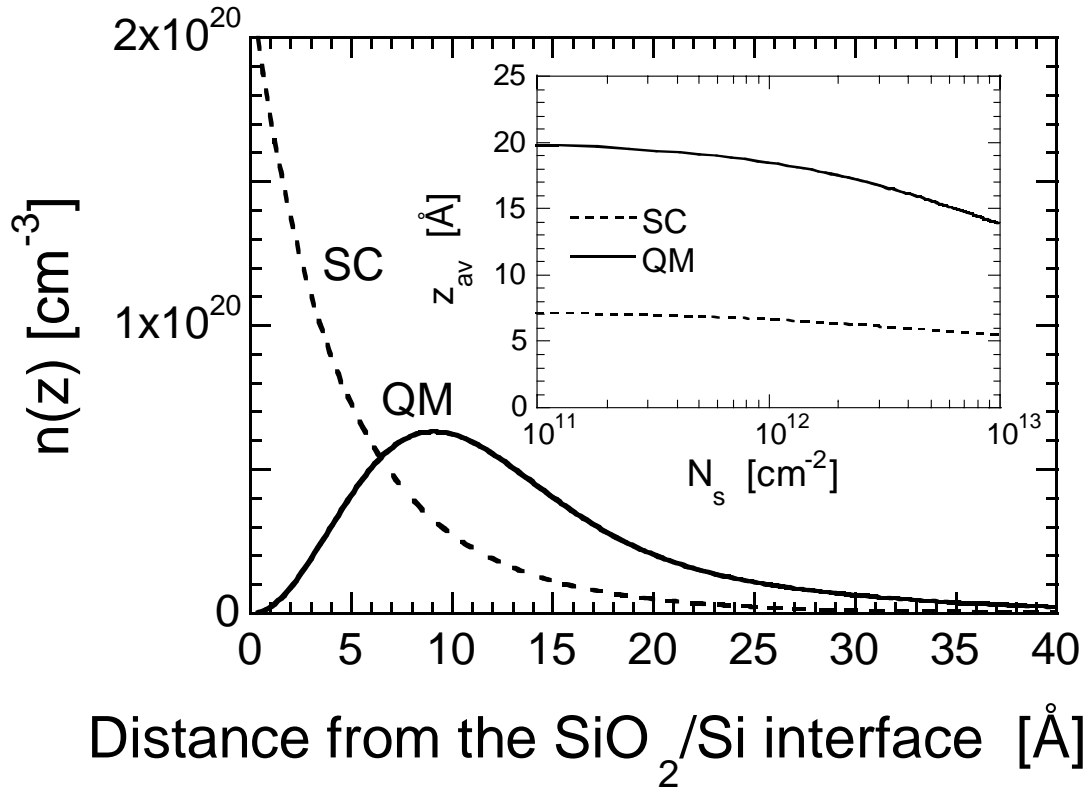


Figure 5: Vasileska

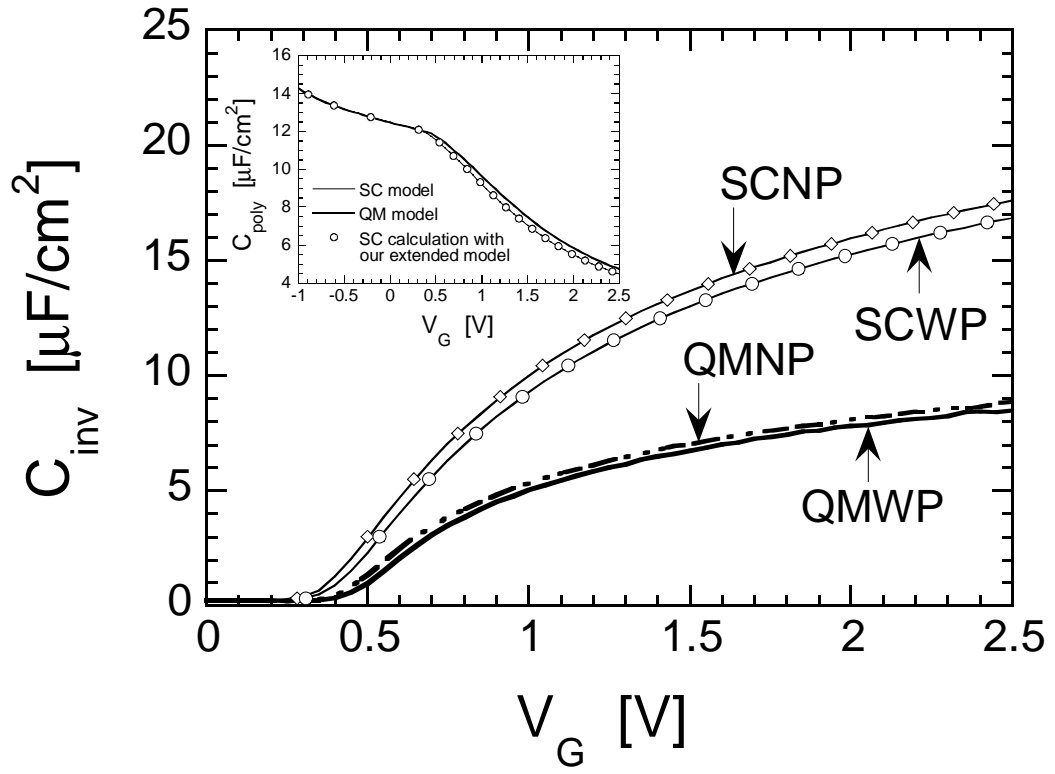


Figure 6: *Vasileska*

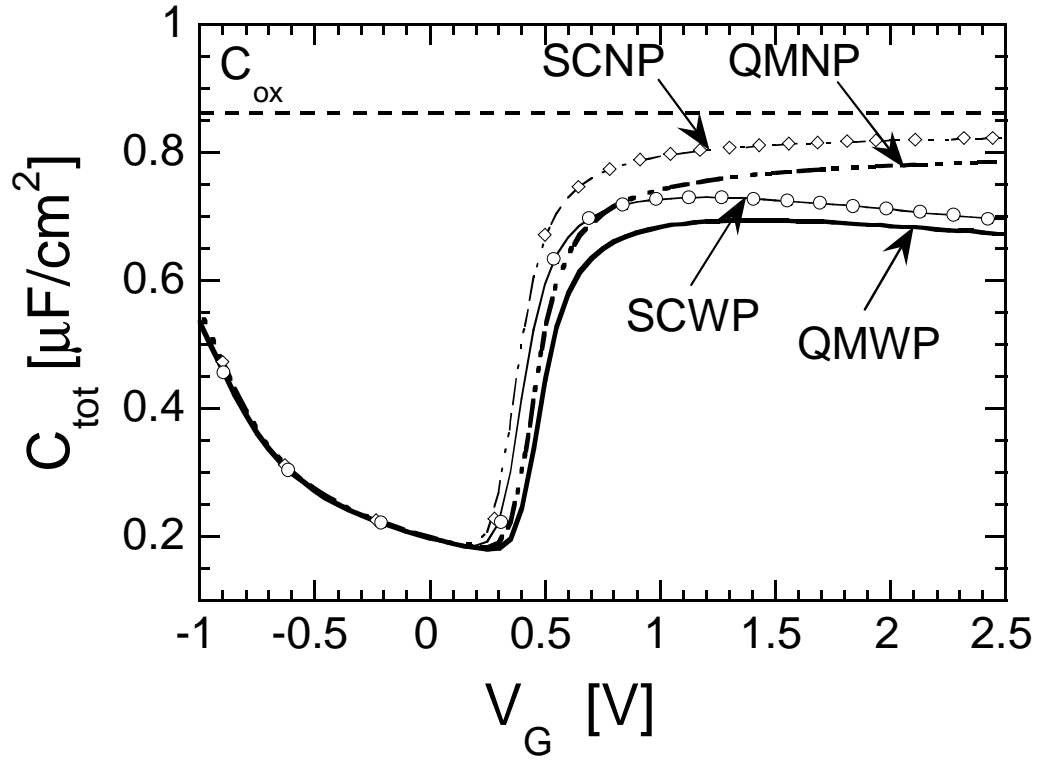


Figure 7: *Vasileska*

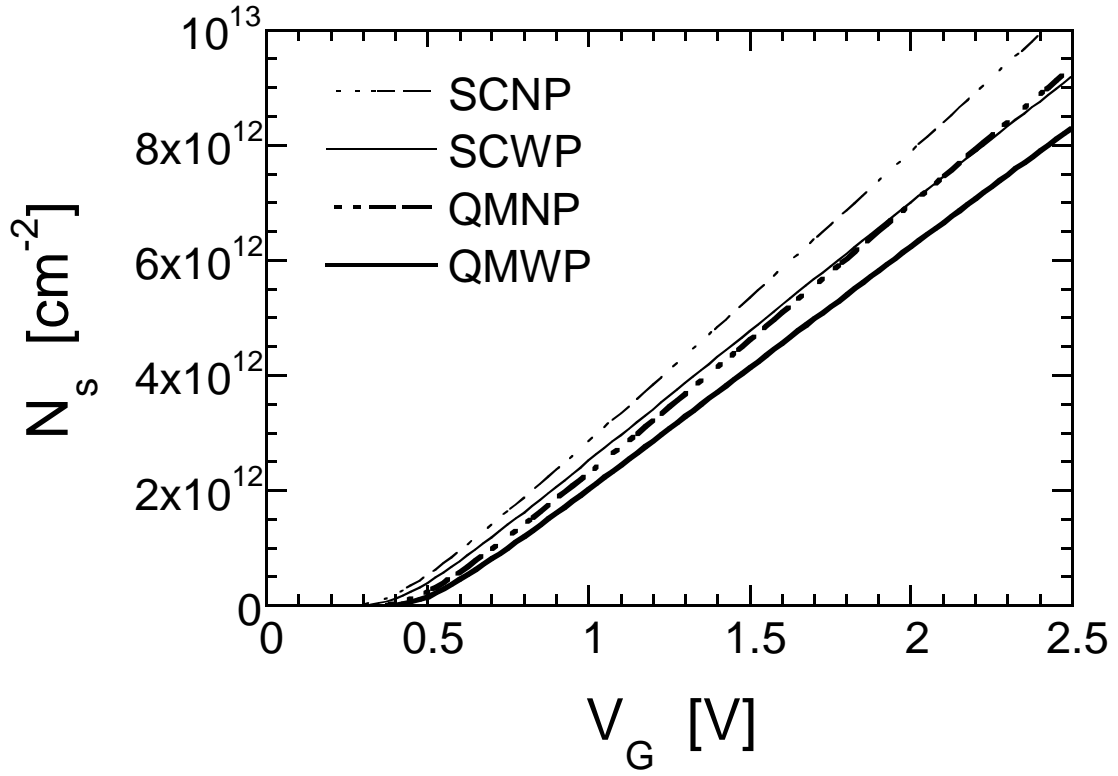


Figure 8: *Vasileska*

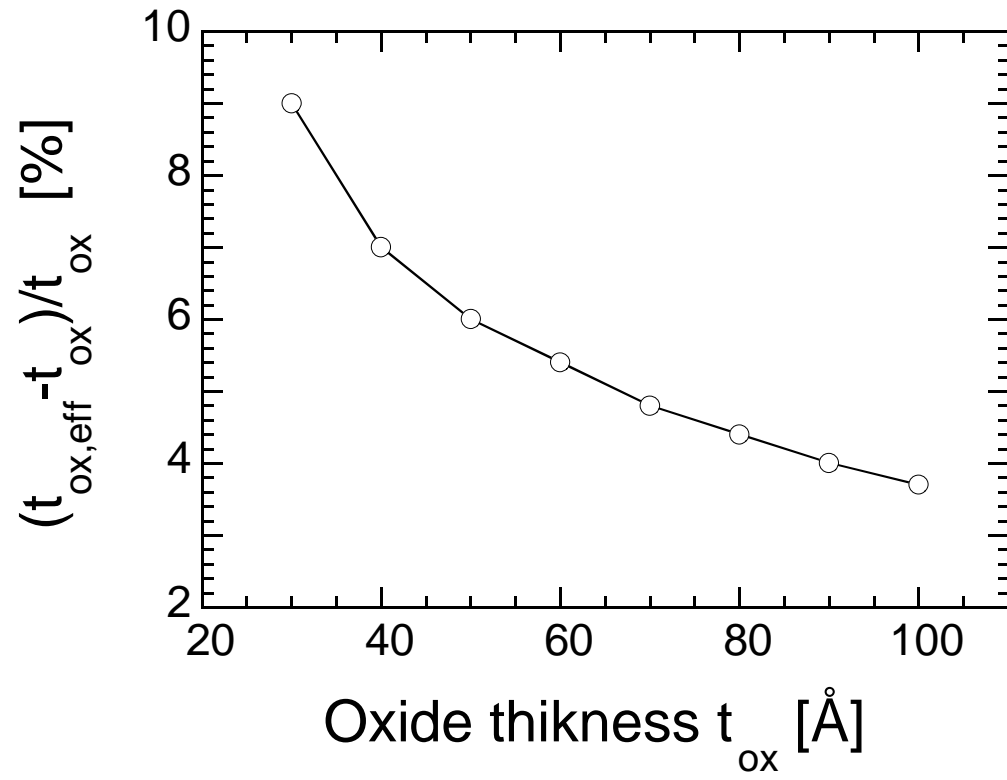


Figure 9: Vasileska

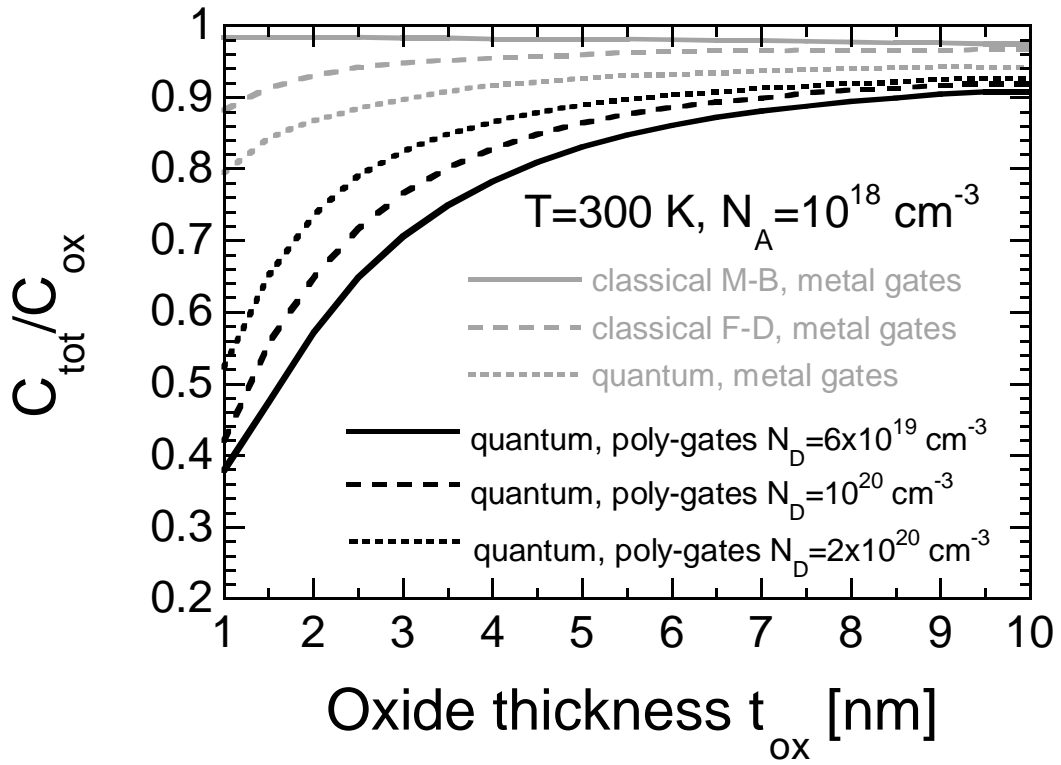


Figure 10: Vasileska

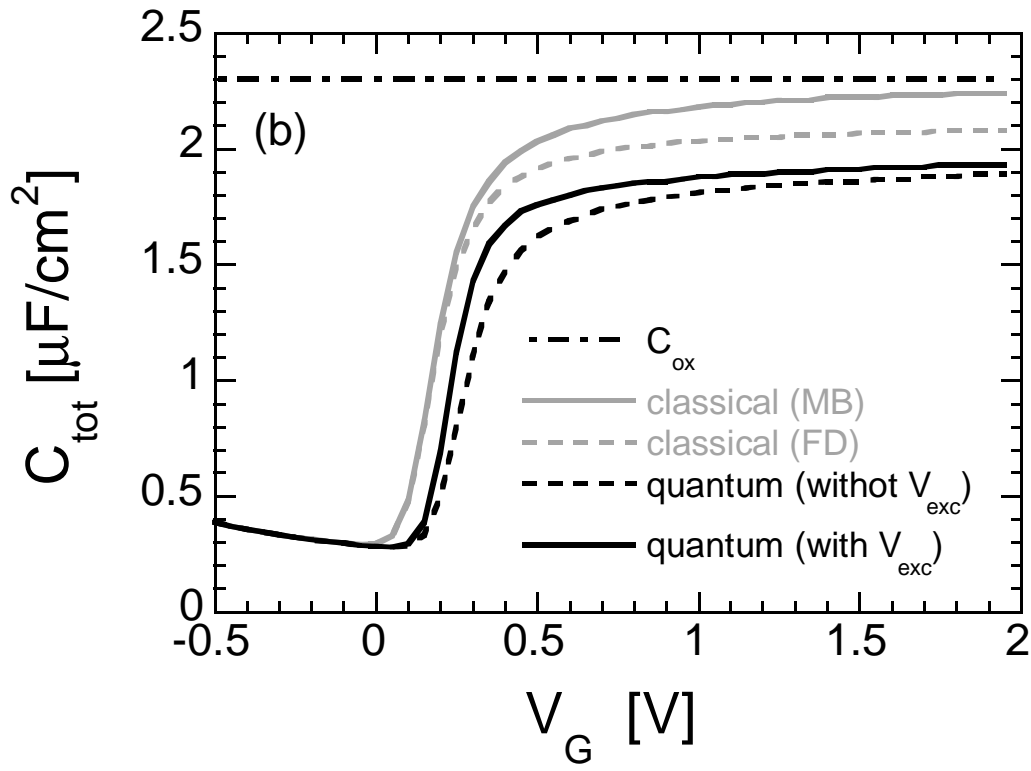
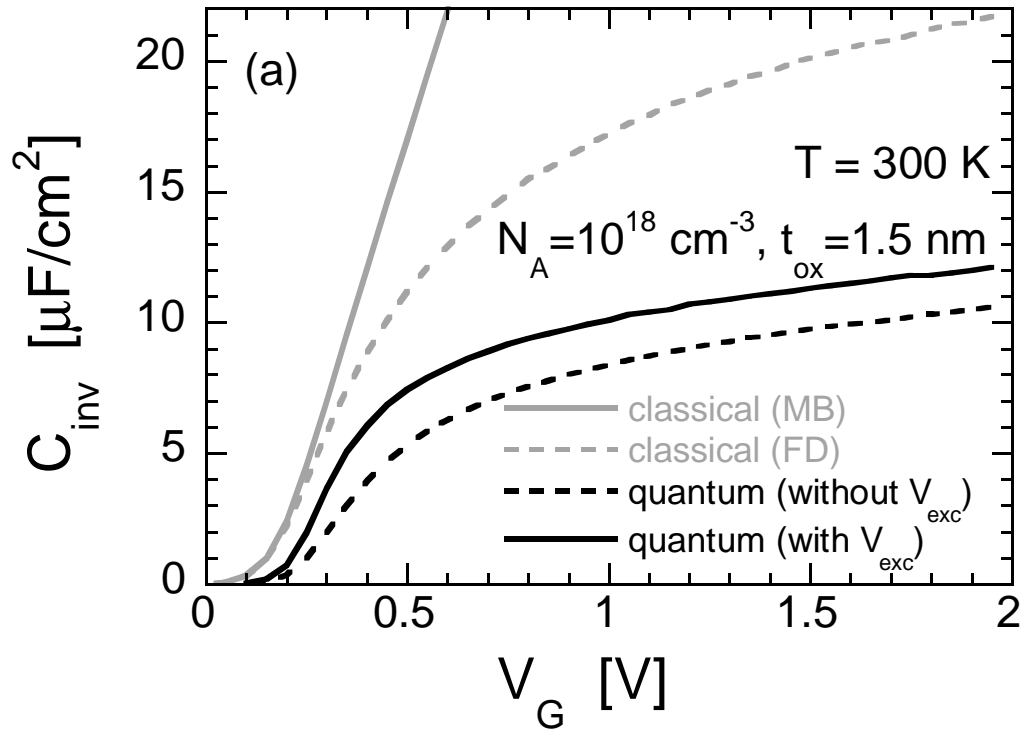


Figure 11. *Vasileska*

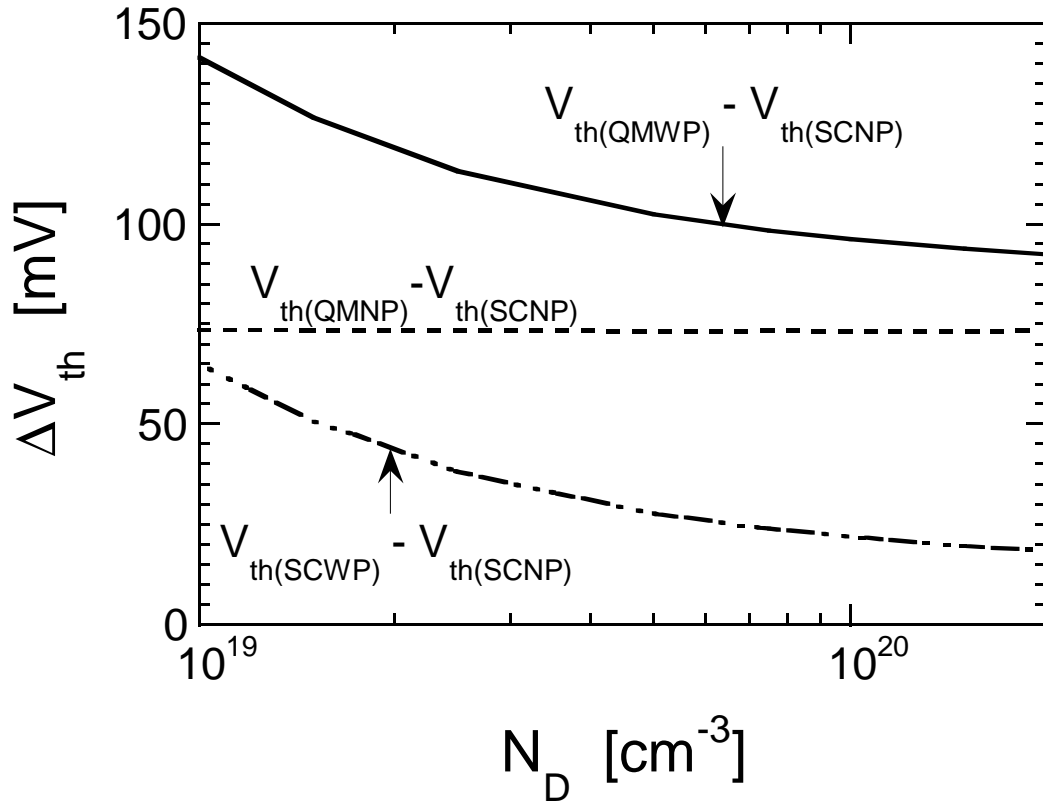


Figure 12: *Vasileska*

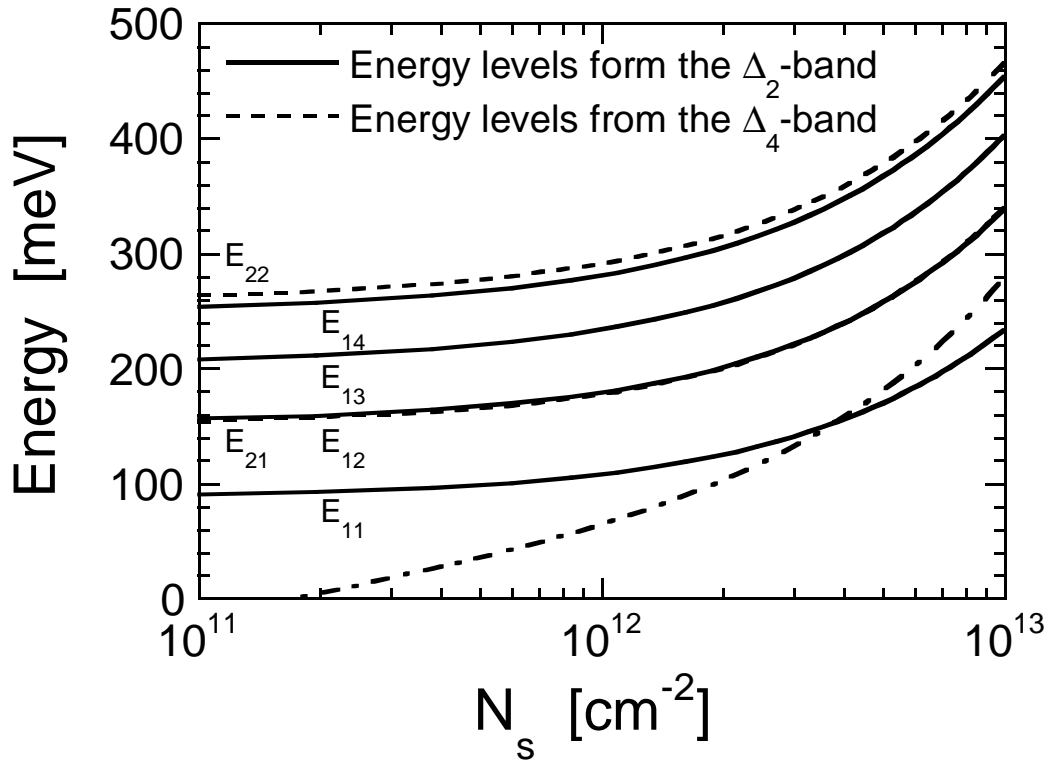


Figure 13: *Vasileska*

