**Report on:**

# Quantization Effects in Nanodevices and the Need for SCHRED

Dragica Vasileska

Professor

Ira A. Fulton School of Engineering

Department of Electrical Engineering

Arizona State University

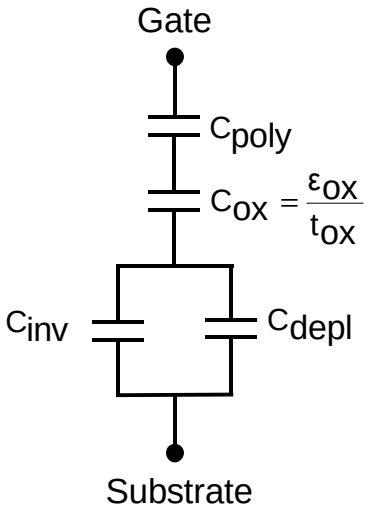
Tempe, AZ 85287-5706, USA

June 2008

## 1. Quantization Effects in Nano-Devices

In the past, quantum effects have been known to dominate the operation of resonant tunneling diodes [1], quantum cascade lasers [2], etc. Tunneling through the gate oxide [3], source to drain tunneling and space-quantization effects are expected to be important in nano-scale MOSFETs and will require solution of the one-dimensional (1D) Schrödinger-Poisson problem. Solutions of the two-dimensional (2D) Schrödinger-Poisson problem are needed, for example, for describing the channel charge in narrow-width MOSFETs. With regard to gate-oxide tunneling, the one-electron effective-mass approximation may not be sufficiently accurate and *ab initio* calculations will most probably be needed.

Note that successful scaling of MOSFETs towards shorter channel lengths requires thinner gate oxides and higher doping levels to achieve high drive currents and minimized short-channel effects [4,5]. For these nanometer devices it was demonstrated a long time ago that, as the oxide thickness is scaled 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 (see Figure 1). As a consequence, the device transconductance is degraded relative to the expectations of the scaling theory [6].




$$C_{tot} = \frac{C_{ox}}{1 + \frac{C_{ox}}{C_{poly}} + \frac{C_{ox}}{C_{inv} + C_{depl}}} \xrightarrow{\text{metal gates}} \frac{C_{ox}}{1 + \frac{C_{ox}}{C_{inv}}}$$

1) Long channel devices, $t_{ox}$ large
$C_{ox}$ small, $C_{ox}/C_{inv} \rightarrow 0$, $C_{tot} = C_{ox}$
2) Nano-scale devices, $t_{ox}$ small
$C_{ox}$ large, $C_{ox}/C_{inv}$ finite, $C_{tot} < C_{ox}$

Note that $C_{inv}$ is always large because the thickness of the inversion layer is small

Figure 1. 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.

The inversion layer capacitance was also identified as being the main cause of the second-order thickness dependence of MOSFET's *IV*-characteristics [7]. The finite inversion layer thickness was estimated experimentally by Hartstein and Albert [8]. The high levels of substrate doping, needed in nano-devices

to prevent the punch-through effect, that lead to 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 this effect was proposed by van Dort and co-workers [9,10]. Later on, Vasileska and Ferry [11] confirmed these findings by investigating the doping dependence of the threshold voltage in MOS capacitors. The experimental data for the doping dependence of the threshold voltage shift and our simulation results from Ref. [11] are shown in Figure 2.

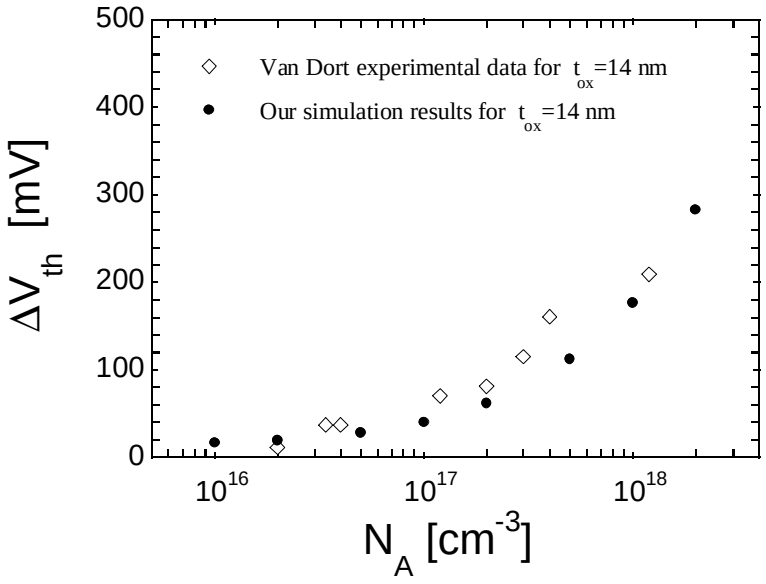



Figure 2. SCHRED simulation data for the shift in the threshold voltage compared to the experimental values provided by van Dort and co-workers.

These results clearly demonstrate the influence of quantum-effects on the operation of nano-scale MOSFETs in both the off- and the on-state. The two physical origins of the inversion layer capacitance, due to the finite density of states and due to the finite inversion layer thickness, were demonstrated experimentally by Takagi and Toriumi [12]. 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 PICSEC simulator [13]. The influence of the image and many-body exchange-correlation effects on the inversion layer and the total gate capacitance was studied by Vasileska *et al*. [14]. It was also pointed out that the depletion of the poly-silicon gates considerably affects the magnitude of the total gate capacitance [15].

The above examples outline the advances during the two decades of research on the influence of quantum-effects on the operation on nano-devices. The conclusion is that any state-of-the-art device simulator must take into consideration the quantum-mechanical nature of the carrier transport and the

poly-depletion effects to correctly predict the device off- and on-state behavior. As noted by many of these authors, to account for the quantum-mechanical effects, one in principle has to solve the 2D/3D Schrödinger-Poisson problem in conjunction with an appropriate transport kernel. (For devices in which velocity overshoot is strongly pronounced, minimum that one can do is to solve the Boltzmann transport equation using the Ensemble Monte Carlo (EMC) technique.)

## 2. Quasi-2D Electron Gas

An important system where quantum effects have been observed is a quasi-two-dimensional electron gas (Q2DEG). There are two basic systems where 2DEG has been studied. As already noted, one of them is Si MOSFET. A very good review of such systems is given in Ref. [16]. A typical device is shown in Figure 3. A (100) Si surface serves as a substrate while $SiO_2$ layer behaves as an insulator. 2DEG is induced electrostatically by application a positive voltage $V_G$. The sheet density of 2DEG can be described as

$$N_s = \frac{\varepsilon_{ox}}{ed_{ox}}\left(V_G - V_T\right) \tag{1}$$

where $V_T$ is the threshold voltage for the barrier's creation.

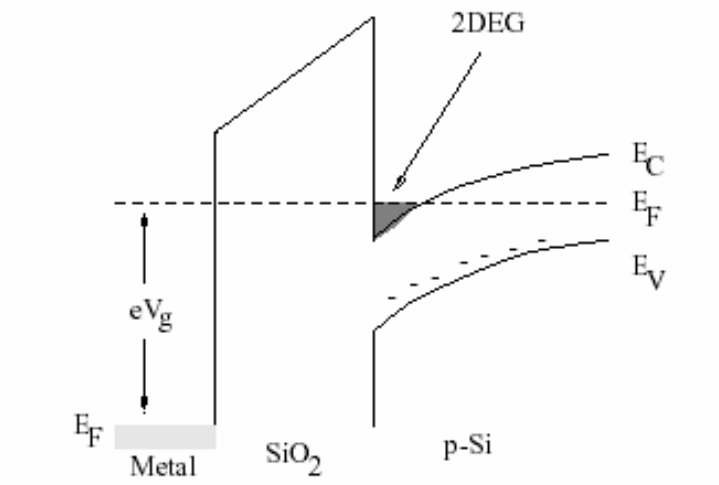



Figure 3. Energy Band diagram showing conductance band $E_C$, valence band $E_V$ and quasi-Fermi level $E_F$. A 2DEG is formed at the interface between the oxide ($SiO_2$) and *p*-type silicon substrate as a consequence of the gate voltage $V_g$.

Another important system with Q2DEG involves modulation-doped GaAs-AlGaAs heterostructures. The bandgap in AlGaAs is wider than in GaAs. By variation of doping it is possible to

move the Fermi level inside the forbidden gap. When the materials are put together, a unified level of chemical potential is established, and an inversion layer is formed at the interface.

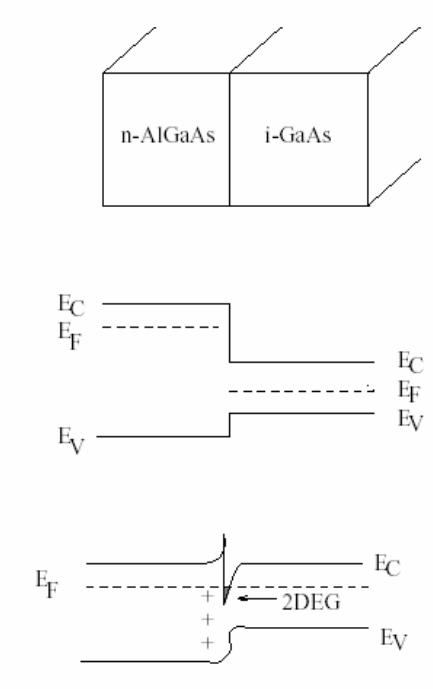



Figure 4. Band structure of the interface between *n*-AlGaAs and intrinsic GaAs, (a) before and (b) after the charge transfer.

The 2DEG created by a modulation doping can be squeezed into narrow channels by selective depletion in spatially separated regions. The simplest lateral confinement technique is to create split metallic gates in a way shown in Figure 5.

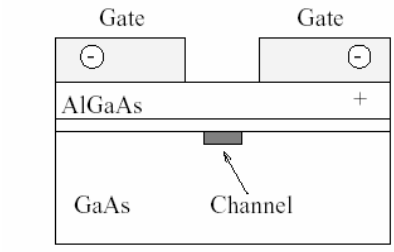



Figure 5. On the formation of a narrow channel by a split gate.

Important quantity characterizing any system of interest is the density of states (DOS) function. The density of states $g(E)$ is defined as the number of states per energy interval (*E*, *E* + *dE*). It is clear that

$$g(E) = \sum_{\alpha} \delta(E - E_{\alpha}) \ , \tag{2}$$

where α is the set of quantum numbers characterizing the states. In the present case, it includes the subband quantum number *n*, spin quantum number σ, valley quantum number υ and the in-plane quasi-momentum **k**. If the spectrum is degenerate with respect to spin and valleys, one can define the spin degeneracy $\nu_s$ and the valley degeneracy $\nu_\upsilon$ to get

$$g(E) = \frac{\nu_s \nu_\upsilon}{(2\pi)^d} \sum_n \int d^d k \delta\left(E - E_n\right) . \tag{3}$$

Here we calculate the number of states per unit volume, *d* being the dimension of the space. For 2D case, we obtain easily

$$g(E) = \frac{\nu_s \nu_\upsilon m}{2\pi\hbar^2} \sum_n \Theta\left(E - E_n\right) . \tag{4}$$

Within a given subband the 2D density of states function is energy independent. Since there can exist several subbands in the confining potential, the total density of states can be represented as a set of steps, as shown in Figure 6. At low temperature ($k_B T << E_F$) all the states are filled up to the Fermi level. Because of energy-independent density of states, the sheet electron density is linear in the Fermi energy, namely

$$N_s = N \frac{\nu_s \nu_\upsilon m E_F}{2\pi\hbar^2} . \tag{5}$$

The Fermi momentum in each subband can be determined as

$$k_{Fn} = \frac{1}{\hbar}\sqrt{2m\left(E_F - E_n\right)} . \tag{6}$$

In Eq. (5), *N* is the number of transverse modes having the edges $E_n$ below the Fermi energy. The situation is more complicated if the gas is confined in a narrow channel, say, along the *y*-axis. In a similar way, the in-plane wave function can be decoupled as a product

$$\psi\left(\mathbf{r}\right) \propto \eta(y) e^{ik_x x} , \tag{7}$$

the corresponding energy being

$$E_{n,s,k} = E + E_s(k_x) + \frac{\hbar^2 k_x^2}{2m} . \tag{8}$$

In Eq. (8), $E_{ns} = E_n + E_s$ characterizes the energy level in the potential confined in both (*z* and *y*) directions. For square-box confinement, the terms are

$$E_s = \frac{\left(s\pi\hbar\right)^2}{2mW^2} , \tag{9}$$

where $W$ is the channel width, while for the parabolic confinement $U(y) = (1/2)m\omega_0^2 y^2$ (typical for split gate structures)

$$E_s = \left(s - 1/2\right)\hbar\omega_0 \quad . \tag{10}$$

For such system, the total density of states is

$$g(E) = \frac{\nu_s \nu_v \sqrt{m}}{2^{3/2}\pi\hbar} \sum_{ns} \frac{\Theta(E - E_{ns})}{\sqrt{E - E_{ns}}} . \tag{11}$$

The energy dependence of the density of states is shown in 7.

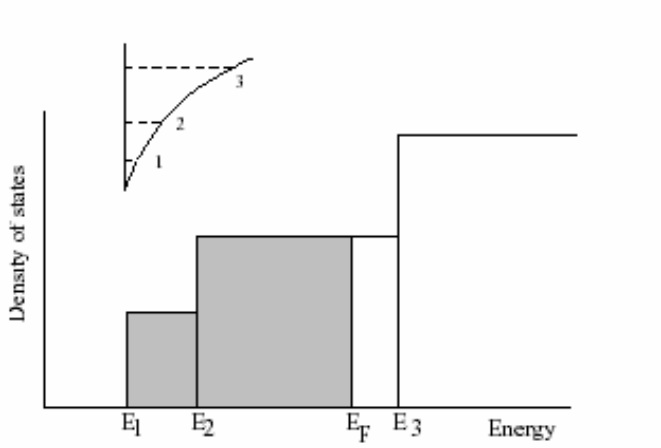



Figure 6. Density of states for a quasi-2D system.

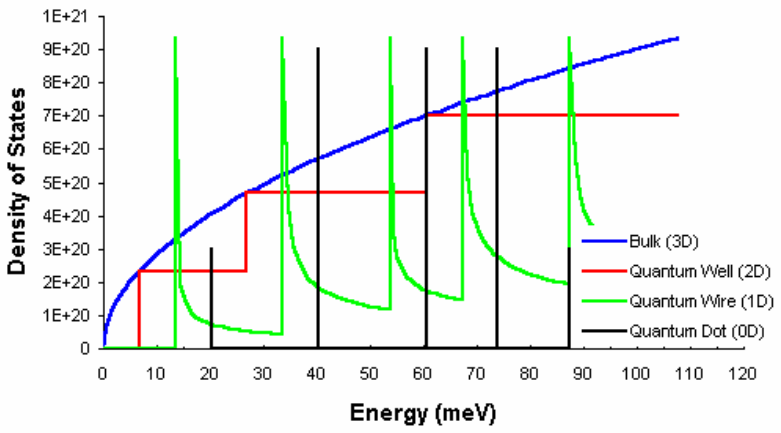



Figure 7. Density of states for bulk (3D blue), quantum well (2D red), quantum wire (1D green) and quantum Dot (0D black)

## 3. Description of SCHRED

The periodic crystal potential in the bulk of semiconducting materials is such that, for a given energy in the conduction band, the allowed electron wavevectors trace out a surface in **k**-space. In the effective-mass approximation for silicon, these constant energy surfaces can be visualized as six equivalent ellipsoids of revolution (Figure 8), whose major and minor axes are inversely proportional to the effective masses. A collection of such ellipsoids for different energies is referred to as a valley.

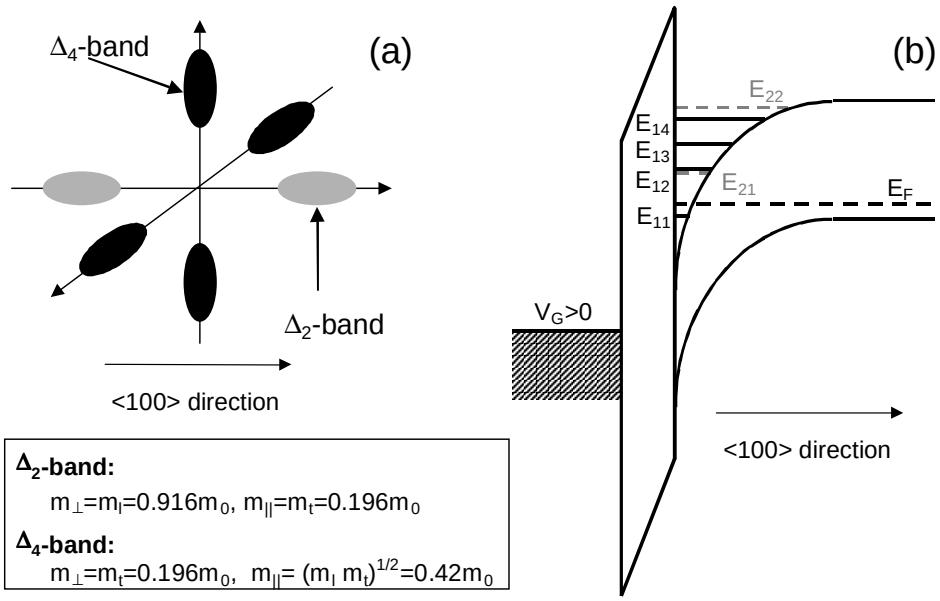



Figure 8. Right panel - Potential diagram for inversion of *p*-type semiconductor. In this first notation $E_{ij}$ refers to the j-th subband from either the $\Delta_2$-band (i=1) or $\Delta_4$-band (i=2). Left panel - Constant-energy surfaces for the conduction-band of silicon showing six conduction-band valleys in the <100> direction of momentum space. The band minima, corresponding to the centers of the ellipsoids, are 85% of the way to the Brillouin-zone boundaries. The long axis of an ellipsoid corresponds to the longitudinal effective mass of the electrons in silicon, $m_l = 0.916 m_o$, while the short axes correspond to the transverse effective mass, $m_t = 0.190 m_o$. For <100> orientation of the surface, the $\Delta_2$-band has the longitudinal mass ($m_l$) perpendicular to the semiconductor interface and the $\Delta_4$-band has the transverse mass ($m_t$) perpendicular to the interface. Since larger mass leads to smaller kinetic term in the Schrodinger equation, the unprimed lader of subbands (as is usually called), corresponding to the $\Delta_2$-band, has the lowest ground state energy. The degeneracy of the unprimed ladder of subbands for <100> orientation of the surface is 2. For the same reason, the ground state of the primed ladder of subbands corresponding to the $\Delta_4$-band is higher that the lowest subband of the unprimed ladder of subbands, The degeneracy of the primed ladder of subbands for (100) orientation of the interface is 4.

In this framework, the bulk Hamiltonian for an electron, residing in one of these valleys is of the form

$$H_o(\mathbf{R}) = -\left( \frac{\hbar^2}{2m_x^*}\frac{\partial^2}{\partial x^2} + \frac{\hbar^2}{2m_y^*}\frac{\partial^2}{\partial y^2} + \frac{\hbar^2}{2m_z^*}\frac{\partial^2}{\partial z^2} \right) + V_{eff}(z) = H_{o\parallel}(\mathbf{r}) + H_{o\perp}(z) \ , \tag{12}$$

where $\mathbf{R} = (\mathbf{r}, z)$, $V_{eff}(z) = V_H(z) + V_{exc}(z)$ is the effective potential energy profile of the confining potential, $V_H(z)$ is the Hartree potential which is nothing more but a solution of the 1D Poisson equation introduced later in the text, $V_{exc}(z)$ is the exchange-correlation potential also discussed later in the text, $H_{o\parallel}$ is the parallel part of $H_o$, and the transverse part is defined as

$$H_{o\perp}(z) = -\frac{\hbar^2}{2m_z^*}\frac{\partial^2}{\partial z^2} + V_{eff}(z) \ . \tag{13}$$

The basis-states of the unperturbed Hamiltonian are assumed to be of the form

$$\Psi_n(\mathbf{R}) = \frac{1}{\sqrt{A}} e^{i\mathbf{k}\cdot\mathbf{r}} \psi_n(z) \ , \tag{15}$$

where $\mathbf{k}$ is a wavevector in the *xy*-plane and $A$ is the area of the sample interface. The subband wavefunctions satisfy the one-dimensional Schrödinger equation

$$H_{o\perp}(z)\psi_n(z) = \varepsilon_n \psi_n(z) \tag{16}$$

subject to the boundary conditions that $\psi_n(z)$ are zero for $z = 0$ and approach zero as $z \to \infty$. In Eq, (16), $\varepsilon_n$ is the subband energy and $\psi_n(z)$ is the corresponding wavefunction. In the parabolic band approximation, the total energy of the electrons is given by

$$E_n(\mathbf{k}) = \frac{\hbar^2\mathbf{k}^2}{2m_{xy}^*} + \varepsilon_n = \varepsilon_{\mathbf{k}} + \varepsilon_n \ , \tag{17}$$

where $\varepsilon_{\mathbf{k}}$ is the kinetic energy and $m_{xy}^*$ is the density of states mass along the *xy*-plane. 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

$$\frac{\partial}{\partial z}\left[ \varepsilon(z)\frac{\partial \varphi}{\partial z} \right] = -e\left[ N_D^+(z) - N_A^-(z) + p(z) - n(z) \right], \tag{18}$$

and the 1D Schrödinger equation

$$\left[ -\frac{\hbar^2}{2m_i^\perp}\frac{\partial^2}{\partial z^2} + V_{eff}(z) \right]\psi_{ij}(z) = E_{ij}\psi_{ij}(z) \, . \tag{19}$$

In (18) and (19), $\varphi(z)$ is the electrostatic potential [the Hartree potential $V_H(z)$=-e $\varphi(z)$], $\varepsilon(z)$ 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 that equals the sum of the Hartree and exchange-correlation corrections to the ground state energy of the system, $m_i^{\perp}$ 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 electron-density is calculated using

$$n(z) = \sum_{i,j} N_{ij} \psi_{ij}^2(z)$$

where $N_{ij}$ is the sheet electron concentration in the *i*-th subband from the j-th valley is given by

$$N_{ij} = g_i \frac{m_{xy}^*}{\pi\hbar^2} k_B T \ln\left\{1 + \exp\left[\left(E_F - E_{ij}\right)/k_B T\right]\right\} \tag{20}$$

where $g_i$ is the valley degeneracy factor and $E_F$ is the Fermi energy. 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 [17], and Kohn and Sham [18]. 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 [19]

$$V_{exc}(z) = -\frac{e^2}{8\pi\varepsilon_{sc} b}\left[1 + 0.7734 x \ln\left(1 + \frac{1}{x}\right)\right]\left(\frac{2}{\pi\alpha r_s}\right), \tag{21}$$

which is accurate over a large density range. In (4), $\alpha = (4/9\pi)^{1/3}$, $x = x(z) = r_s/21$, $r_s = r_s(z) = \left[4\pi b^3 n(z)/3\right]^{-1/3}$, and $b = 4\pi\varepsilon_{sc}\hbar^2/m^* e^2$. 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. As a result, noticeable increase in the energy of the inter-subband transitions can be observed at high electron densities.

Similarly, the valence band is represented by the heavy hole band and light hole band, the spit-off band is ignored because the spit-off energy is large enough to exclude any hole staying there. In treating holes quantum mechanically, the same effective mass based Schrodinger equation is solved with the masses quoted from references (C. Hu, S. Banerjee, k. Sadra, B.G. Streetman and R. Sivan, "Quantization Effects in Inversion Layers of PMOSFET's on Si (100) Subatrates," IEEE Electron Dev. Lett., Vol. 17, No. 6, pp.276-278, June 1996 and S. Takagi, M. Takayanagi, and A. Toriumi, "Characterization of Inversion-Layer Capacitance of Holes in Si MOSFET's," IEEE Trans. Electron Devices 46, pp. 1446-1450, 1999).Due to their different perpendicular masses, the heavy holes form the first set of energy levels which are relatively low, and the light holes form the second set with higher confined energies. Schred 2.1 also has the capability of treating the electron/hole density in the inversion layer classically by using either Maxwell-Boltzmann or Fermi-Dirac statistics.

In doing bulk structure quantum mode simulation, the version 2.1 can not only solve the effective mass based Schrodinger equation for inversion layer carriers, but also can solve the equation for accumulation layer carriers, for example, if the bulk is p type silicon, in the inversion range, electrons are treated quantum mechanically, whereas in the accumulation range, holes are treated quantum mechanically. This is a feature that many other simulators do not offer.

In doing SOI quantum mode simulation, both electrons and holes are treated quantum mechanically at the same time. This is because in most cases, the SOI bodies are undoped or lightly doped, and the two dielectric gates confine the carriers in both inversion and accumulation regimes, therefore, the quantum effects can be equally important for both electrons and holes at low biases.

For both simulation modes, (classical or quantum mechanical), if the gate contacts are polysilicon, the charge density on the gates will always be computed classically. The gate dielectric constant can be specified different from $SiO_2$, The new version also allows different dielectrics for the top and bottom gates in a SOI structure. This eases the simulations of effects of exotic insulator materials on device performance. Typical outputs of the solver are the spatial variations of the conduction-band edge and 3D charge density in the body; 2D surface charge density, average distance of the carriers from the interface; inversion layer capacitance $C_{inv}$, depletion layer capacitance $C_{depl}$, total gate capacitance $C_{tot}$ and in the case of capacitors with poly-silicon gates, it also calculates the poly-gate capacitance $C_{poly}$. When choosing quantum-mechanical description of the electron density in thechannel, it also provides the subband energies the subband population, and wavefunction variations in the body.

Schred is written in Fortran 77. The program is efficient compared to other 1D Schrodinger-Poisson self-consistent simulators. On a SPARC-5 workstation, generally, it takes about 10 seconds per bias point in quantum mode calculation, and about 5 seconds per bias point in classical mode calculation. But for bulk accumulation range simulation, it takes a relatively long time--about 2 to 3 minutes for one bias point. This is because in accumulation range, the band potential energy level bends very little, and the subband energies crowds together, so that a large number of subbands need to be included in the calculation in order to accurately account for the contributions from all lowest subbands. A SOI quantum mode simulation with very thick silicon body (thicker than 0.1 micron) can also involve relatively long computation time.

Examples of the application of Schred can be found in D. Vasileska, D. K. Schroder and D.K. Ferry, "Scaled silicon MOSFET's: Part II-Degradation of the total gate capacitance," IEEE Trans. Electron Devices 44, pp. 584-587 (1997), and D. Vasileska, and D.K. Ferry, "The influence of space quantization effects on the threshold voltage, inversion layer and total gate capacitance in scaled Si-MOSFETs," in the Technical Proceedings of the First International Conference on Modeling and Simulation of Microsystems, Semiconductors, Sensors and Actuators, Santa Clara, California, April 6-8, 1998, pp. 408-413. J. Fossum, Z. Ren, K. Kim and M. Lundstrom "Extraordinarily High Drive Currents in Asymmetrical Double-Gate MOSFETs," VLSI symposium June, 2000.

## References


1 R. Tsu and L. Esaki, Appl. Phys. Lett., 22 (1973) 562.

2 J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson and A. Y. Cho, Science, 264 (1994) 553.

3 C.-J. Sheu and S.-L. Jang, Solid-State Electron., 44 (2000) 1819.

4 R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous and A. R. leBlanc, *IEEE J. Solid-State Circuits* **9**, 256 (1974).

5 J. R. Brews, W. Fichtner, E. H. Nicollian and S. M. Sze, *IEEE Electron Dev. Lett.* **1,** 2 (1980).

6 G. Bacarani and M. R. Worderman, in *Proceedings of the IEDM*, 278 (1982).

7 M.-S. Liang, J. Y. Choi, P.-K. Ko and C. Hu, *IEEE Trans. Electron Devices* **33**, 409 (1986).

8 A. Hartstein and N. F. Albert, *Phys. Rev. B* **38**, 1235 (1988).

9 M. J. van Dort, P. H. Woerlee, A. J. Walker, C. A. H. Juffermans and H. Lifka, *IEEE Trans. Electron Dev*. **39**, 932 (1992).

10 M. J. van Dort, P. H. Woerlee and A. J. Walker, *Solid-State Electronics* **37**, 411 (1994).

11 D. Vasileska and D. K. Ferry, in the *Technical Proceedings of the First International Conference on Modeling and Simulation of Microsystems, Semiconductors, Sensors and Actuators*, 408 (1998).

12 S. Takagi and A. Toriumi, *IEEE Trans. Electron Devices* **42**, 2125 (1995).

13 S. A. Hareland, S. Krishnamurthy, S. Jallepali, C.-F. Yeap, K. Hasnat, A. F. Tasch, Jr. and C. M. Maziar, *IEEE Trans. Electron Devices* **43**, 90 (1996).

14 D. Vasileska, D. K. Schroder and D. K. Ferry, *IEEE Trans. Electron Devices* **44**, 584 (1997).

[15] K. S. Krisch, J. D. Bude and L. Manchanda, *IEEE Electron Dev. Lett.* **17**, 521 (1996).
[16] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. 54, 437 (1982).
[17] P. Hohenberg and W. Kohn, Phys. Rev. 136, B864 (1964).
[18] Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965).
[19] L. Hedin and B. I. Lundqvist, J. Phys. C 4, 2064 (1971).