Quantum Mechanical Carrier Transport and Nano-scale MOS Modeling

Zhiping Yu and Dawei Zhang
Institute of Microelectronics
Tsinghua University, Beijing, China
yuzhip@tsinghua.edu.cn

2nd Hiroshima Int'l Workshop
Jan. 30, 2004

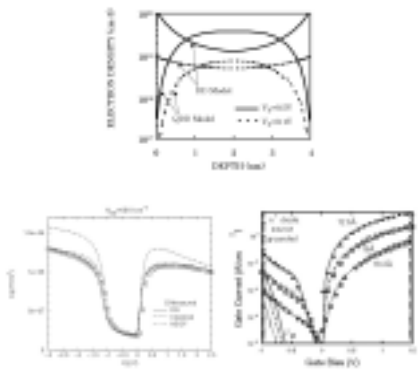
Outline

- Introduction
- Schrödinger Equation and Wigner Function
- Macroscopic Modeling Approach
  - Quantum Hydrodynamic (QHD)
  - Density Gradient (DG)
  - Effective Potential (EP)
- Microscopic Modeling Approach
  - QTBM (Quantum Transmitting Boundary Method)
  - QDAME (Quan. Dev. Ana. by Modal Evaluation)
  - NEGF (Non-equilibrium Green's Function)
- Numerical Model for 3D FinFETs
- Ballistic MOS Model (BMM)

Introduction

Manifestation of quantum mechanical effects on MOS characteristics:

- Increasing effective gate oxide thickness with consequence of
  - Smaller gate capacitance
  - Larger threshold voltage
- Tunneling current
  - through gate oxide
  - through barrier between S/D along the channel
- Ballistic transport
- Subthreshold slope



Modeling approaches:

- **Macroscopic: particle (solve Shockley semiconductor eqs.)**
  - **Quantum hydrodynamic model (QHD)**
  - **Density gradient drift-diffusion (QDD)**
  - **Effective potential approach (EP)**
- **Microscopic: wavefunction (solve Schrödinger eq.)**
  - **Quantum transmitting boundary method (QTBM)**
  - **Modified QTBM – QDAME**
  - **Green's function approach – NEGF**
- **Compact circuit model**
  - **Changing threshold voltage**
  - **Model ballistic transport (transmission theory)**

Schrödinger Equation and Wigner Function

- Schrödinger Equation

$$\left[-\frac{\hbar^2}{2m^*}\nabla^2 + U(\mathbf{r})\right]\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \quad (1)$$

- Wigner Function

$$f_W(\mathbf{r},\mathbf{p}) = \frac{1}{(\pi\hbar)^3}\int \Psi^*(\mathbf{r}+\mathbf{r}')\Psi(\mathbf{r}-\mathbf{r}')e^{2i\mathbf{p}\cdot\mathbf{r}'/\hbar}d\mathbf{r}' \quad (2)$$

Momentum-shifted Wigner fuction

$$f_W^{(2)}(\mathbf{x},\mathbf{p}) = \mathcal{A}\frac{e^{-\beta E}}{(2\pi\hbar)^3}\left\{1+\hbar^2\left[-\frac{\beta^2}{8m^*}\nabla^2 U + \frac{\beta^3}{24m^*}(\nabla U\cdot\nabla U) + \frac{\beta^3}{24m^*}(\mathbf{p}'\cdot\nabla)^2 U\right]\right\} \quad (3)$$

where $\mathbf{p}' = \mathbf{p} - m^*\mathbf{u}$ with $\mathbf{u}$, the macroscopic fluid (or drift) velocity, and $E = p'^2/2m^* + U$.

- Wigner-Boltzmann Equation

$$\frac{\partial f_W}{\partial t} + \frac{\mathbf{p}\cdot\nabla f_W}{m^*} - \frac{2}{\hbar}\sin\left(\frac{\hbar}{2}\nabla^U\cdot\nabla_{\mathbf{p}}\right)U(\mathbf{r})f_W(\mathbf{r},\mathbf{p}) = 0 \quad (4)$$

Or use the integral form by defining the nonlocal potential energy

$$\hat{U}(\mathbf{r},\mathbf{p}) = 2\int \sin\left(\frac{\mathbf{p}\cdot\mathbf{y}}{\hbar}\right)\left[U(\mathbf{r}+\mathbf{y}/2) - U(\mathbf{r}-\mathbf{y}/2)\right]d\mathbf{y} \quad (5)$$

The Wigner-Boltzmann eq. then takes form of

$$\frac{\partial f_W}{\partial t} + \frac{\mathbf{p}\cdot\nabla f_W}{m^*} - \frac{2}{\hbar}\int \frac{1}{2\pi\hbar}\hat{U}(\mathbf{r},\mathbf{p}-\mathbf{p}')f_W(\mathbf{r},\mathbf{p}')d\mathbf{p}' = 0 \quad (6)$$

Moment Approach – QHD

- Retaining terms up to $\hbar^2$ in the serious expansion of Eq. (4) and take first three moments in $\mathbf{p}$-space: $1, \mathbf{p}, \mathbf{p}\cdot\mathbf{p}/2m^*$.
- Quantum Hydrodynamic Equations

$$\frac{\partial n}{\partial t} + \frac{1}{m^*}\nabla\cdot\mathbf{\Pi} = 0 \quad (7)$$

$$\frac{\partial}{\partial t}\Pi_i + \nabla\cdot(\mathbf{u}\Pi_i) - \sum_{j=1}^{3}\frac{\partial P_{ji}}{\partial x_j} = -n\frac{\partial U}{\partial x_i} - \frac{1}{\tau_p}\Pi_i,\ i = 1,2,3 \quad (8)$$

$$\frac{\partial W}{\partial t} + \nabla\cdot(\mathbf{u}W - \hat{P}\mathbf{u} + \mathbf{q}) = -\frac{1}{m^*}\mathbf{\Pi}\cdot\nabla U - \frac{W - W_0}{\tau_w} \quad (9)$$

for variables $n, \mathbf{u}$, and energy density W, where $\mathbf{\Pi} = nm^*\mathbf{u}$ is the momentum density and $\hat{P}$ the stress tensor.

- Quantum corrections

  The gradient of carrier density is a manifestation of non-locality, which is the essence of quantum mechanics.

– Stress tensor

$$\hat{P} = -nTI + \frac{\hbar^2 n}{12m^*}(\nabla\nabla)\ln n \tag{10}$$

– Energy density

$$W = \frac{3}{2}nT + \frac{1}{2}m^* n u^2 - \frac{\hbar^2 n}{24m^*}\nabla^2 \ln n \tag{11}$$

– A compact form

$$\nabla\cdot(\nabla \ln n) = -\frac{1}{n^2}\nabla n\cdot\nabla n + \frac{1}{n}\nabla^2 n \tag{12}$$

$$2\frac{\nabla^2\sqrt{n}}{\sqrt{n}} = -\frac{1}{2}\frac{1}{n^2}\nabla n\cdot\nabla n + \frac{1}{n}\nabla^2 n \tag{13}$$

Density Gradient (DG) Drift-Diffusion Model

$$V_n = V + Q_n \tag{14}$$
$$V_p = V - Q_p \tag{15}$$

where quantum potential

$$Q_n = 2b_n\frac{\nabla^2\sqrt{n}}{\sqrt{n}}, \quad b_n = \frac{\hbar^2}{12qm_n^*} \tag{16}$$
$$Q_p = 2b_p\frac{\nabla^2\sqrt{p}}{\sqrt{p}}, \quad b_p = \frac{\hbar^2}{12qm_p^*} \tag{17}$$

V_n and V_p are used in the place of V for conventional DD formulation, e.g.,

$$n = n_i e^{(V_n-\phi_n)/V_T} \tag{18}$$
$$p = n_i e^{(\phi_p-V_p)/V_T} \tag{19}$$

The carrier fluxes are expressed as

$$\mathbf{F}_n = -D_n\nabla n + \mu_n n\nabla V_n \tag{20}$$
$$\mathbf{F}_p = -D_p\nabla p - \mu_p p\nabla V_p \tag{21}$$

Solving Shockley semiconductor equations for n, p, and V

$$\nabla\cdot(\epsilon\nabla V) = -q(p - n + N_D^+ - N_A^-) \tag{22}$$
$$\nabla\cdot\mathbf{F}_n + r = 0 \tag{23}$$
$$\nabla\cdot\mathbf{F}_p + r = 0 \tag{24}$$

Solution approaches

(main challenge: the robustness of the method)

- Trade number of variables to the order of derivative
- Five variables: $V, S_n = \sqrt{n}, S_p = \sqrt{p}, \phi_n$, and ϕ_p.

Equation set:

$$\nabla\cdot(\epsilon\nabla V) + q(p - n + N_D^+ - N_A^-) = 0 \tag{25}$$
$$\nabla\cdot(b_n\nabla S_n) + \frac{S_n}{2}\left(V - \frac{k_B T}{q}\ln\frac{n}{n_i} - \phi_n\right) = 0 \tag{26}$$
$$\nabla\cdot(b_p\nabla S_p) - \frac{S_p}{2}\left(V + \frac{k_B T}{q}\ln\frac{p}{n_i} - \phi_p\right) = 0 \tag{27}$$
$$\nabla\cdot(\mu_n n\nabla\phi_n) + \frac{\partial n}{\partial t} + r = 0 \tag{28}$$
$$\nabla\cdot(\mu_p p\nabla\phi_p) - \frac{\partial p}{\partial t} - r = 0 \tag{29}$$

The disadvantage of this scheme is when $b \to 0$, the Helmhotz equations, Eqs. (26-27) becomes singular. The convergence behavior is poor, often needs to keep the bias step small to prevent negative $S_n(\sqrt{n})$ and $S_p(\sqrt{p})$.

Results

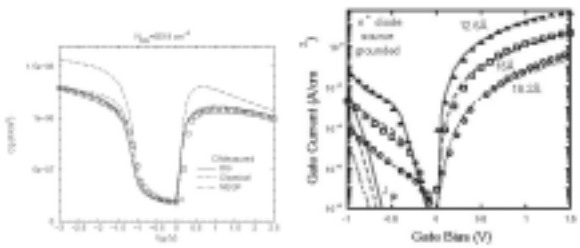


Figure 1: Measured and simulated C-V for $t_{\text{ox}} = 2.1\,\text{nm}$ and tunneling through MOS gate structure using DG.

- Use variables V, u_n, u_p, ϕ_n, and ϕ_p (Prof. Shinji Odanaka of Osaka Uni.) where

$$u_n = \frac{V_n - \phi_n}{2}, \quad u_p = \frac{\phi_p - V_p}{2} \tag{30}$$

The advantage with variable u's is that the carrier concentration is guaranteed to be positive, e.g.,

$$S_n = \sqrt{n} = \sqrt{n_i} e^{u_n/V_T} \tag{31}$$

The eq. for u_n is

$$-b_n \nabla \cdot (S_n \nabla u_n) + S_n u_n = \frac{S_n}{2}(V - \phi_n) \tag{32}$$

This eq. is of Sturm-Liouville type of problem for u_n given S_n, V, and ϕ_n.

Results:

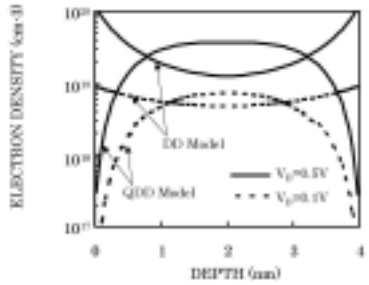


Figure 2: Carrier distribution in double gate MOSFET ($t_{\text{Si}} = 5\,\text{nm}$, $t_{\text{ox}} = 1.5\,\text{nm}$, $L_g = 25\,\text{nm}$.

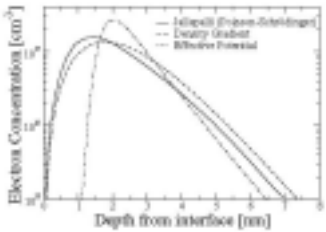
The advantage: robustness – The drain bias can jump to 0.5 V in one step.

Effective Potential (EP) Model

- Considering electron wavepacket as finite in size

$$V_{\text{eff}} = \int V_{\text{cl}}(\mathbf{x} + \mathbf{y}) G(\mathbf{y}, a_0) d\mathbf{y} \tag{33}$$

- Comparison between DG and EP models



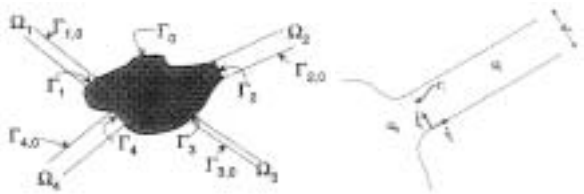
- Similarity with DG

$$V_{\text{eff}}(x) = \frac{1}{\sqrt{2\pi}a_0}\int V(x+\xi)e^{-\xi^2/2a_0^2}d\xi$$
$$\approx V(x) - \frac{2a_0^2}{\beta}\frac{\partial^2 \ln\sqrt{n/n_0}}{\partial x^2} + \cdots \quad (34)$$

Poor accuracy of EP may be due to the missing of another term with the same order as $\hbar^2$.

Microscopic QM Transport

- Quantum Transmitting Boundary Model (QTBM)



Find $\Psi_0(\mathbf{r}) \in C^2(\Omega_0)$, knowing the form of $\Psi_i(\xi_i,\eta_i) \in C^2(\Omega_i)$, $i = 1,\ldots,n$. The boundary condition for Ψ_0 in Ω_0 is

$$\Psi_0 = 0 \text{ on } \Gamma_0 = \partial\Gamma_0 - \sum_{i=1}^{n}\Gamma_i \quad (35)$$
$$\Psi_0 = \Psi_i \text{ on } \Gamma_i \quad (36)$$
$$\nabla\Psi_0\cdot\mathbf{n}_{\Gamma_i} = \nabla\Psi_i\cdot\mathbf{n}_{\Gamma_i} \text{ on } \Gamma_i \quad (37)$$

General solution procedure:

– The solution in the lead i has the form of

$$\psi_i(\xi_i,\eta_i) = \sum_{m=1}^{N_i}\left[a_m^i\chi_m^i(\xi_i)e^{-jk_m^i\eta_i} + b_m^i\chi_m^i(\xi_i)e^{jk_m^i\eta_i}\right]$$
$$+ \sum_{m=N_i+1}^{\infty} b_m^i\chi_m^i(\xi_i)e^{-jk_m^i\eta_i} \quad (38)$$

– Eliminating the explicit dependence on Ψ_i, a mixed BC for Ψ_0 on Γ_i is obtained as

$$\nabla\Psi_0\cdot\mathbf{n}_{\Gamma_i}\big|_{\xi_i} = \sum_{i=1}^{N^i} ik_m^i\chi_m^i(\xi_i)\left(-2a_m^i + \int_{\Gamma_i}\chi_m^i(\Gamma_i)\Psi_0(\Gamma_i)d\Gamma_i\right)$$
$$- \sum_{m=N^i+1}^{\infty} k_m^i\chi_m^i(\xi_i)\int_{\Gamma_i}\chi_m^i(\Gamma_i)\Psi_0(\Gamma_i)d\Gamma_i \quad (39)$$

- Quantum Device Analysis by Mode Evaluation (QDAME)

Based on QTBM and has the following features:

– Able of evaluating I-V characteristics

– Discretely sample a device' s density of states using standing wave boundary conditions, decomposing the standing waves into traveling waves injected from the contacts to assign occupancies.

Solution procedure:

1. Find standing wavefunctions by imposing (at Γ_i)

$$\text{sine}: \Psi_n^{(s)} = 0 \implies E_n^{(s)}$$
$$\text{cosine}: \nabla\Psi_n^{(c)}\cdot\mathbf{n}_{\Gamma_i} = 0 \implies E_n^{(c)}$$

2. Decompose the normal mode into traveling components

$$\Psi_n^{(s,c)}(\mathbf{r}) = \sum_i \Phi_{n,i}^{(s,c)}(\mathbf{r}) \tag{40}$$

3. Given the normal mode energy and the injection coefficients, the traveling components are computed using a modified version of the QTBM.

4. The electron density $n(\mathbf{r})$ is the sum of electron densities from the sine sampling $n^{(s)}(\mathbf{r})$ and cosine sampling $n^{(c)}(\mathbf{r})$, found by summing over all traveling eigen-component densities multiplied by a thermal occupation factor

$$n(\mathbf{r}) = \sum_{s,c} n^{s,c}(\mathbf{r}) = \sum_{s,c} \left\{ \sum_{n,i} |\Phi_{n,i}^{s,c}(\mathbf{r})|^2 \sum_p c_{p,n,i}^{s,c} \rho(k_D^i, T, E_n^{s,c}) \right\} \tag{41}$$

where the thermal occupancy factor is the sum of occupancy factor $\rho(k_D^i, T, E_n^{s,c})$ and coefficients $c_{p,n,i}^{s,c}$ representing the fractional weight of each mode p in lead i making up the traveling eigen-component.

Drifted Fermi-Dirac occupancy factor

$$\rho(k_{Di}, T, E) = \left(\frac{8m_z^i k_B T}{h^2}\right)^{1/2} \times F_{-1/2}\left(\frac{E_{Fi} - E - \frac{\hbar^2}{2m_\eta^i} k_D^i \left(k_D^i - 2k_\eta^{p,i}\right)}{k_B T}\right) \tag{42}$$

where the drift momentum $\hbar k_D^i$ in lead i is found from the current continuity requirement between lead and device.

- **Non-equilibrium Green's Function** (**NEGF**)

NEGF provides a microscopic theory for quantum transport including dissipative interaction.

Green's function, $G^R(\mathbf{r}, \mathbf{r}')$, is obtained from solving Schrödinger equation with boundary conditions

$$[E - H_d]G^R(\mathbf{r}, \mathbf{r}') - \int \Sigma^R(\mathbf{r}, \mathbf{r}_1) G^R(\mathbf{r}_1, \mathbf{r}') d\mathbf{r}_1 = \delta(\mathbf{r} - \mathbf{r}') \tag{43}$$

which can be viewed as the wavefunction at the point $\mathbf{r}$ due to a unit excitation at $\mathbf{r}'$.

Without the source term, the above Green's function becomes the Schrödinger-like equation with boundary condition built-in it:

$$E\Psi(\mathbf{r}) = H_d \Psi(\mathbf{r}) + \int \Sigma^R(\mathbf{r}, \mathbf{r}_1) \Psi(\mathbf{r}_1) d\mathbf{r}_1 \tag{44}$$

which describes the dynamics of an electron inside the device region.

Matrix form of Green's function

$$G^R = \left[EI - H_d - \Sigma^R\right]^{-1} \tag{45}$$

The use of Green's functions

– DOS: G^R and its conjugate transpose $G^A = [G^R]^\dagger$ represent the density of states in the energy space.

$$N(E) = \frac{1}{2\pi} \mathrm{Tr}\left\{ i[G^R(E) - G^A(E)] \right\} \tag{46}$$

Or the position-dependent density of states

$$\rho(\mathbf{r}, E) = -\frac{1}{\pi} \mathrm{Im}\left[G^R(\mathbf{r}, \mathbf{r}; E)\right] \tag{47}$$

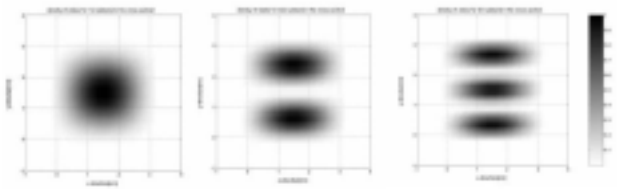


Figure 3: DOS on the cross-section (3 nm×3 nm) of a FinFET

– Carrier concentration: can be expressed using the so-called correlation function G^n:

$$G^n(\mathbf{r},\mathbf{r}') = \iint G^R(\mathbf{r},\mathbf{r}_1)\Sigma^{\text{in}}(\mathbf{r}_1,\mathbf{r}_1')[G^R(\mathbf{r}',\mathbf{r}_1')]^* d\mathbf{r}_1 d\mathbf{r}_1' \quad (48)$$

where Σ^{in} is the inscattering function related to Σ^R and the quasi-Fermi level on the contact as in

$$\Sigma_p^{\text{in}}(\mathbf{r},\mathbf{r}';E) = f_p(E,\mu_p)\Gamma_p(\mathbf{r},\mathbf{r}';E) \quad (49)$$

where p represents the contact (or the lead) and μ_p is the quasi-Fermi level for that contact and $\Gamma = i\left[\Sigma^R - \{\Sigma^R\}^\dagger\right]$.

$$\Sigma^R = \Sigma_\varphi^R + \sum_p \Sigma_p^R \quad (50)$$

$$\Sigma_p^R(i,j;E) = -\frac{\hbar^2}{2m^* a^2}\sum_{m\in p}\chi_m(p_i)e^{ik_m a}\chi_m(p_j) \quad (51)$$

The carrier concentration is

$$n(\mathbf{r}) = 2\int \frac{1}{2\pi}G^n(\mathbf{r},\mathbf{r};E)dE \quad (52)$$

where the pre-factor 2 is for the spin degeneration.

– Current density: The current density inside the device is

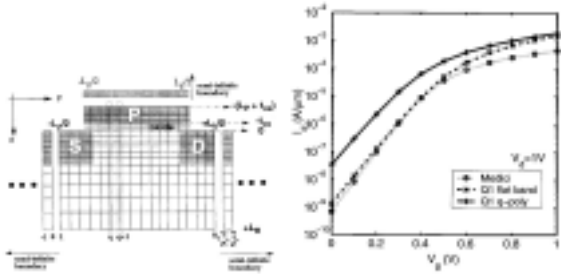
$$\mathbf{j}(\mathbf{r};E) = -\frac{iq\hbar}{2m^*}\left[(\nabla-\nabla')G^n(\mathbf{r},\mathbf{r}';E)\right]_{\mathbf{r}=\mathbf{r}'} \quad (53)$$

where the gradient operator ∇' means that it only acts on $\mathbf{r}'$. The terminal current per unit energy is

$$i(E) = \int \mathbf{j}(\mathbf{r},E)\cdot d\mathbf{S} \quad (54)$$

2D MOS Simulation using NEGF

– NASA's simulation of 2D MOSFET with NEGF, including a self-consistent treatment of 2D gate oxide tunneling.



Comparison with DG:

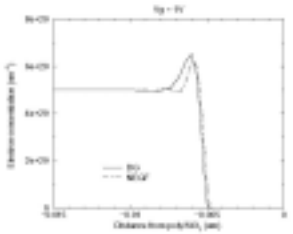


Figure 4: Electron distribution in the poly-gate region using DG and NEGF for MIT *wtm* 25 ($L_{\text{ch}} = 25\,\text{nm}$).

Quasi-3D Numerical Model for FinFETs

- Simulation method:
  - Schrödinger eq. solver on 2D cross-section
  - NEGF along the channel
  - 3D Poisson's solver
- Device structure

Figure 5: FinFET structure and layout

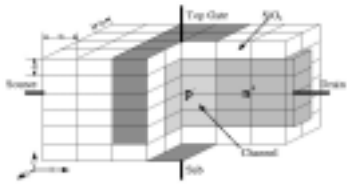


Figure 6: Simulation region and mesh

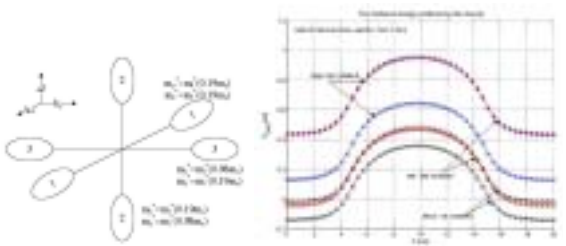
- Simulation details
  - Separation of variables and WKB theory on x-direction

$$\Psi(x,y,z) = X(x)\varphi(y,z) = e^{ik_x x}\varphi(y,z) \tag{55}$$

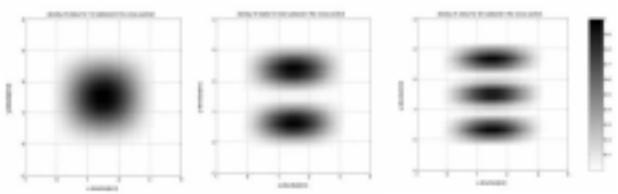
  - 2D Schrödinger eq. on y-z cross-section

$$\left(-\frac{\hbar^2}{2m_y^*}\frac{\partial^2}{\partial y^2} - \frac{\hbar^2}{2m_z^*}\frac{\partial^2}{\partial z^2}\right)\Psi_x^i(y,z) - qV_x(y,z)\Psi_x^i(y,z) = E_{t,x}^i\Psi_x^i(y,z) \tag{56}$$

  - Bounded states affected by different effective mass



– 2D density of states (DOS)



– NEGF solution of 1D Schrödinger eq. with open BC for density matrix in energy representation

$$H\,X(x) = \left[-\frac{\hbar^2}{2m_x^*}\frac{\partial^2}{\partial x^2} + E_t(x)\right] X(x) = E\,X(x) \qquad (57)$$

$$G(E) = [EI - H - \Sigma(E)]^{-1} \qquad (58)$$

where the self-energy matrix

$$\Sigma = \begin{bmatrix} \Sigma_{\text{source}} & 0 & \cdots & 0 & 0 \\ 0 & \Sigma_s(1) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \Sigma_s(n) & 0 \\ 0 & 0 & \cdots & 0 & \Sigma_{\text{drain}} \end{bmatrix} \qquad (59)$$

For ballistic transport (neglecting scattering), $\Sigma_s(i) = 0, i = 1, \cdots, n$, and

$$\Sigma_{\text{contact}}(E) = -\frac{\hbar^2}{2m_x^*}\frac{d^2}{dx^2}e^{ik_l x} \qquad (60)$$

where k_l can be solved from

$$E = E_{t,\text{contact}} + \frac{\hbar^2}{2m_x^*}\frac{d^2}{dx^2}(1 - \cos k_l x) \qquad (61)$$

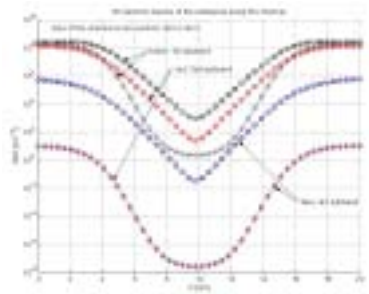
$E_{t,\text{contact}}$ is the eigen-energy determined by the lead (contact).

– The density matrix at energy E is

$$\rho(E) = \frac{1}{\pi}\frac{d}{dx}[F(\mu_s - E)A_S + F(\mu_d - E)A_D] \qquad (62)$$

$$A_S = G[i(\Sigma_s - \Sigma_s^+)]G^+, \quad A_D = G[i(\Sigma_d - \Sigma_d^+)]G^+ \qquad (63)$$

– 2D electron density for each subband



– 3D electron density on the cross-section and along the channel.

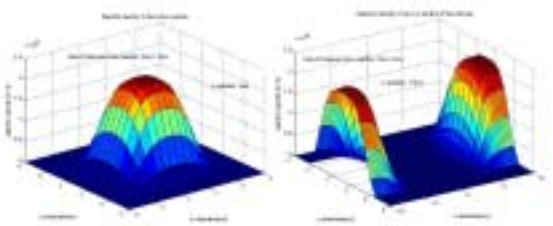
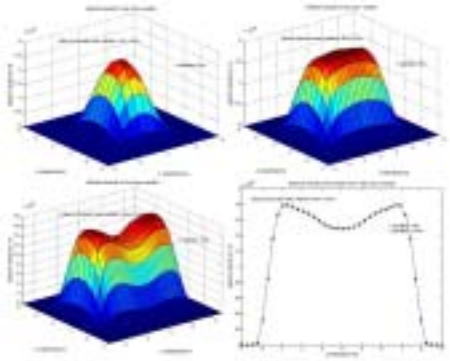


Figure 7: Cross-section: 3 nm×3 nm

– Size effect



– I-V characteristics

$$\begin{aligned} I(E) &= \tfrac{2q}{\hbar}[F(\mu_s - E) - F(\mu_d - E)]\mathrm{Tr}[\Gamma_s A_D] \\ &= \tfrac{2q}{\hbar}[F(\mu_s - E) - F(\mu_d - E)]\mathrm{Tr}[\Gamma_d A_S] \end{aligned} \tag{64}$$

Where $\Gamma_s = i[\Sigma_s - \Sigma_s^+]$ and $\Gamma_d = i[\Sigma_d - \Sigma_d^+]$.

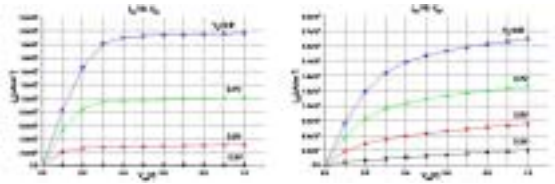


Figure 8: $L_{\mathrm{ch}} = 10, 5\,\mathrm{nm}$, cross-section: $3\,\mathrm{nm}\times 3\,\mathrm{nm}$, and $t_{\mathrm{ox}} = 1\,\mathrm{nm}$

Ballistic MOS Model (BMM) with 2D QM Corrections

Features:

- Ballistic carrier transport along the channel

$$\frac{I_D}{W} = Q(0)\frac{1-r}{1+r}\left[v_T\frac{\mathcal{F}_{1/2}(\eta)}{\mathcal{F}_0(\eta)}\right]\frac{1-\frac{\mathcal{F}_{1/2}(\eta-U_D)}{\mathcal{F}_{1/2}(\eta)}}{1+\frac{1-r}{1+r}\frac{\mathcal{F}_0(\eta-U_D)}{\mathcal{F}_0(\eta)}} \tag{65}$$

where v_T is the thermal velocity, $U_D = V_{DS}/V_T$, $\eta = (E_{FS} - E_{\max})/k_BT$, and back scattering coefficient

$$r = \frac{l}{l+\lambda},\ l = L\left(V_T\frac{\beta}{V_{DS}}\right)^{\alpha},\ \lambda = V_T\frac{2\mu}{v_T}\frac{\mathcal{F}_0^2(\eta)}{\mathcal{F}_{-1}(\eta)\mathcal{F}_{1/2}(\eta)}$$

- QM corrected threshold voltage

$$Q(0) = C_{ox}V_{g,\mathrm{eff}} \tag{66}$$

where $C_{ox} = \epsilon_{ox}/t_{ox}$ is the gate-oxide capacitance per unit surface area and $V_{g,\mathrm{eff}}$ is the effective gate voltage,

$$V_{g,\mathrm{eff}} = \frac{2sV_T\ln\left[1+\exp\left(\frac{V_{od}}{2sV_T}\right)\right]}{1+2sC_{ox}\sqrt{\frac{2\psi_S}{q\epsilon_0\epsilon_{\mathrm{si}}N_{\mathrm{sub}}}}\exp\left[\frac{V_{od}-2(V_{GS}-V_{th}-V_{\mathrm{off}})}{2sV_T}\right]} \tag{67}$$

where $V_{od} = V_{GS} - V_{FB} - \phi_S - Q_{\mathrm{dep}}/C_{ox}$, s swing factor, and V_{off} a parameter. The key of the modeling is to find right expression for the threshold voltage

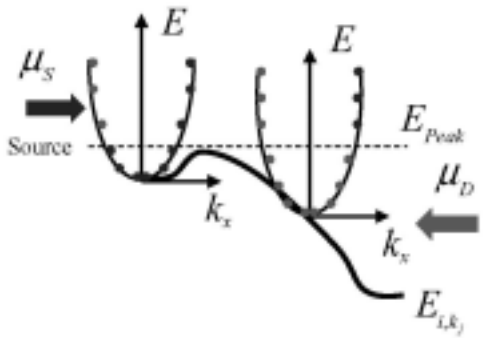
$$\boxed{V_{th} = V_{th,\mathrm{cl}} + \Delta V_{1\mathrm{D\_qm}} + \Delta V_{\mathrm{pg\_qm}} + \Delta V_{2\mathrm{D\_qm}}} \tag{68}$$

According to Natori,

$$Q(0) \approx 2.5\frac{qkT}{2\pi\hbar^2}\sqrt{m_t m_l}\left[\ln\left(1+e^{\eta}\right) + \ln\left(1+e^{\eta-U_D}\right)\right] \tag{69}$$

from which one can find η, given bias.

Band diagram along the channel



– 1D-QM correction

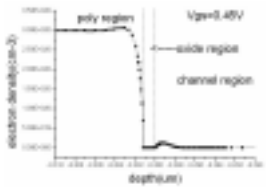
$$\Delta V_{\text{1D\_qm}} = 1 + \frac{1}{2C_{\text{ox}}}\sqrt{\frac{2q\epsilon_0\epsilon_{\text{si}}N_{\text{sub}}}{a_1 + b_1\overline{N}_{\text{sub}} + c_1\overline{N}_{\text{sub}}^2}} \tag{70}$$

$\overline{N}_{\text{sub}} = N_{\text{sub}}/10^{18}$, fitting parameters $a_1 = 1.02411, b_1 = 0.04606$, and $c_1 = -0.00206$.

– QM in poly gate

$$\Delta V_{\text{pg\_qm}} = a_p + b_p\overline{N}_{\text{sub}} = c_p\overline{N}_{\text{sub}} \tag{71}$$

$a_p = 0.01967, b_p = 0.01133$, and $c_p = \frac{1}{3} \times 10^{-3}$



– 2D Corrections

The solution to the 3D Schrödinger equation using the separation of variables as follows.

$$\Psi(x,y,z) = \underbrace{\frac{A}{\sqrt{p(x)}}\exp\left(\frac{i}{\hbar}\int^x p(\xi)d\xi\right)}_{X(x)} \cdot \underbrace{\sqrt{\frac{2}{W}}\sin\left(\frac{n_y\pi}{W}y\right)}_{Y(y)} \cdot \underbrace{\varphi_{n_z}(z)}_{Z(z)} \tag{72}$$

With WKB theory,

$$p(x) = \sqrt{2m_x[E - U(x,y,z)]} \tag{73}$$

$$\frac{i}{\hbar}\int_a^b \sqrt{2m_x[E_d - U(x)]}dx = -4 \tag{74}$$

for $e^{-4} = 1.83\%$ can be considered as negligibly small. The correction to the threshold voltage due to 2D QM effects can

be expressed as

$$\Delta V_{\text{2D-QM}} = \frac{E_d - E_{\text{peak}}}{q}\left(1 + \frac{1}{2C_{\text{ox}}}\sqrt{\frac{\epsilon_0\epsilon_{\text{Si}}q^2N_{\text{sub}}}{k_BT\ln(N_{\text{sub}}/n_0)}}\right) \tag{75}$$

The results by applying the 2D QM correction to the simulation of transfer characteristics.

Analytical approach: approximate $E_{n_z}(x)$ with a parabola.

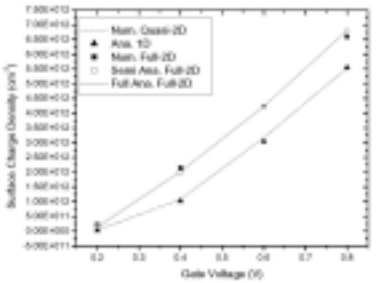
$$E_{n_z}(x) \approx E_{\text{peak}} - \sigma(x - x_{\max})^2 \tag{76}$$

And obtain

$$E_d - E_{\text{peak}} = \frac{8\hbar}{\pi}\sqrt{\frac{\sigma}{2m_x}} \tag{77}$$

The fitting parameter σ is less sensitive to t_{ox} and N_{sub}, but strongly depends on the channel length L. An empirical formula is

$$\sigma = 4.73 \times 10^{-3}e^{-L/16} \tag{78}$$



where L in units nm.

Simulation Examples using BMM

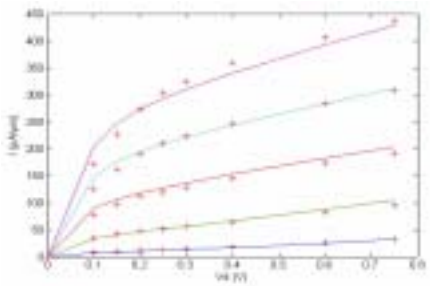


Figure 9: Modeling of 14 nm MOSFET (A. Hokazono *et al.*, 2002) with BMM, $V_{gs} = 0.25 - 0.65$ V of step 0.1 V.

Conclusions

- **QM effects become first-order ones in nano-scale MOSFETs**
- **Robust macroscopic model(s) now exists for device simulation**
- **Wavefunction based approach may find broad application in nanoelectronic devices**
- **The built-in of QM effects in compact model likely will take the similar path to the "surface potential" approach.**