Quantum Corrections for Monte Carlo Simulation

Umberto Ravaioli

Beckman Institute and
Dept. of Electrical and Computer Engineering

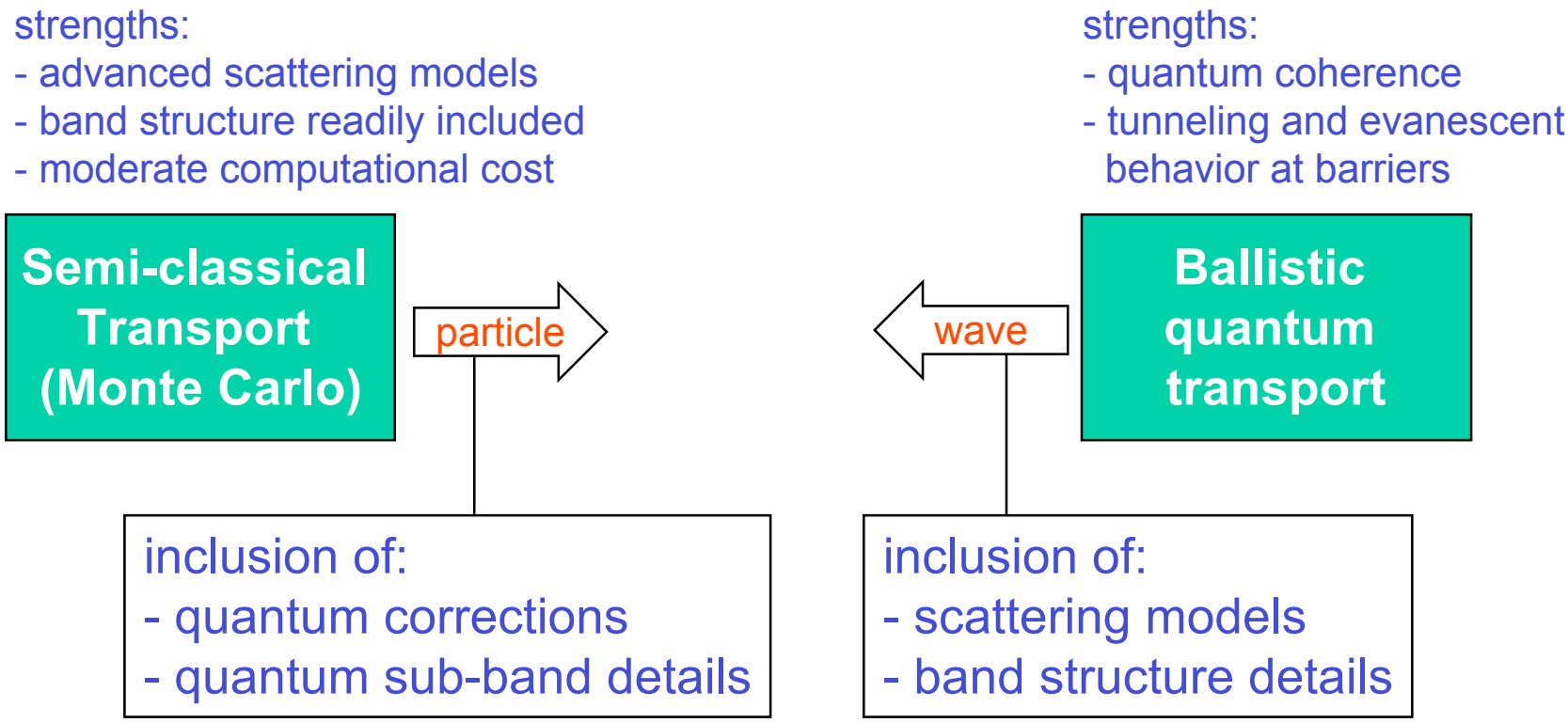
University of Illinois at Urbana-Champaign
USA

Outline

- Quantum corrections for size-quantization
  - Effective potential correction
  - Wigner equation correction
  - Bohm equation correction
  - Schrödinger equation correction
- Extending Schrödinger-based correction to device simulation
- Examples

Introduction

- **There is still a considerable gap between semi-classical particle simulation and quantum models in terms of physical details**



Motivation for Quantum Corrections

- Full-quantum transport is often impractical
- Goal is to extend the validity of semi-classical Monte Carlo to ≈10-nm regime
- Quantum corrections can extend the validity of semi-classical Monte Carlo in a practical way
  - Mixed quantum/classical effects are treated in a unified fashion
  - Little extra computational overhead is added in both 2D and 3D

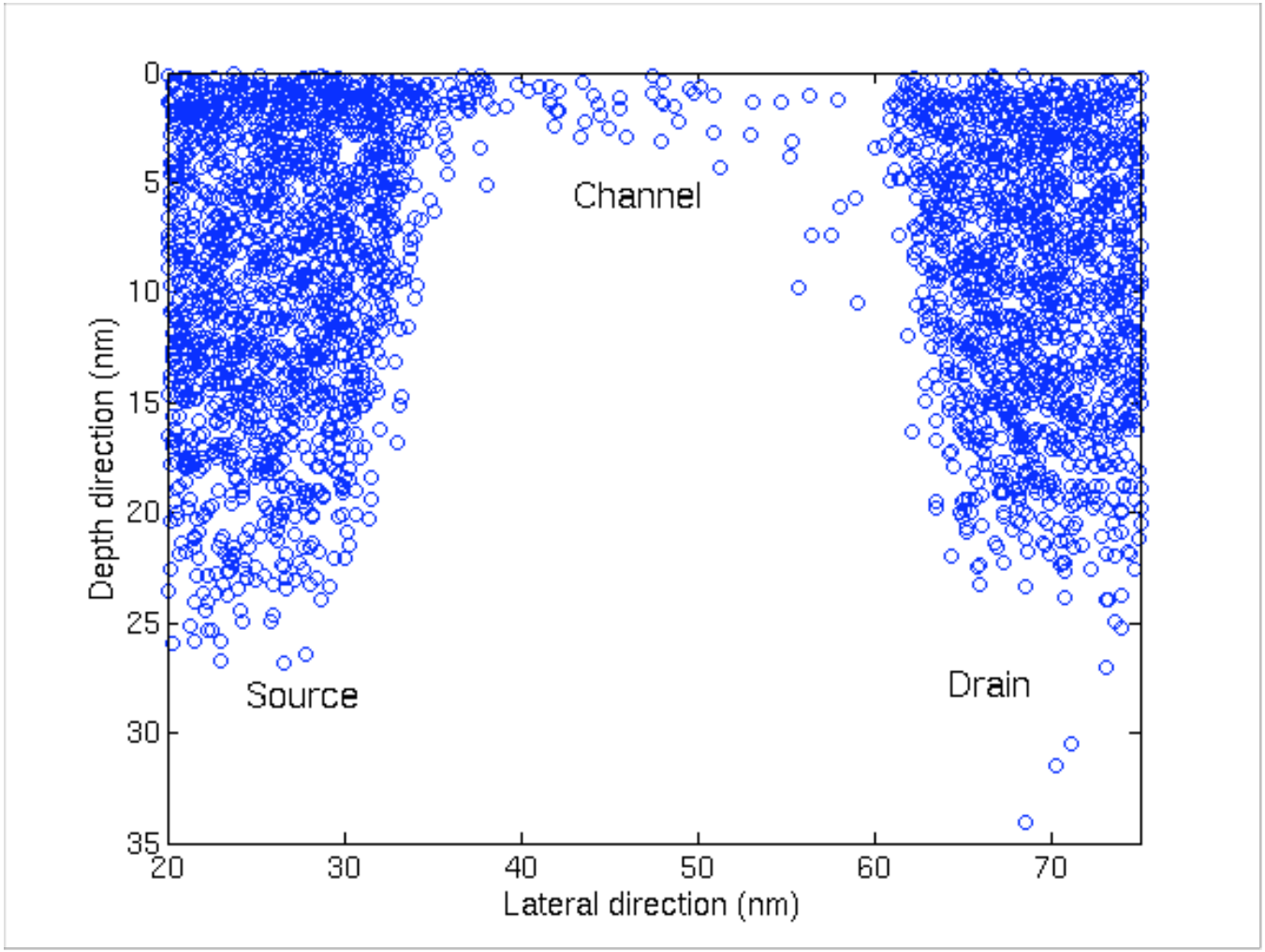
The “expensive” MC alternative [1,2]

- Resolve complete transverse subbands
- Describe particle dynamics within subbands
- Implement intra-subband, inter-subband and possibly 2-D to 3-D scattering if a continuum of states is used at high energies
- If a continuum of states is used in the contacts, one needs to treat the dis-continuity of band edge to subbands at the channel entrance.
- Applications have been mainly limited to uniform 2DEG channels.

Role of Quantum Corrections

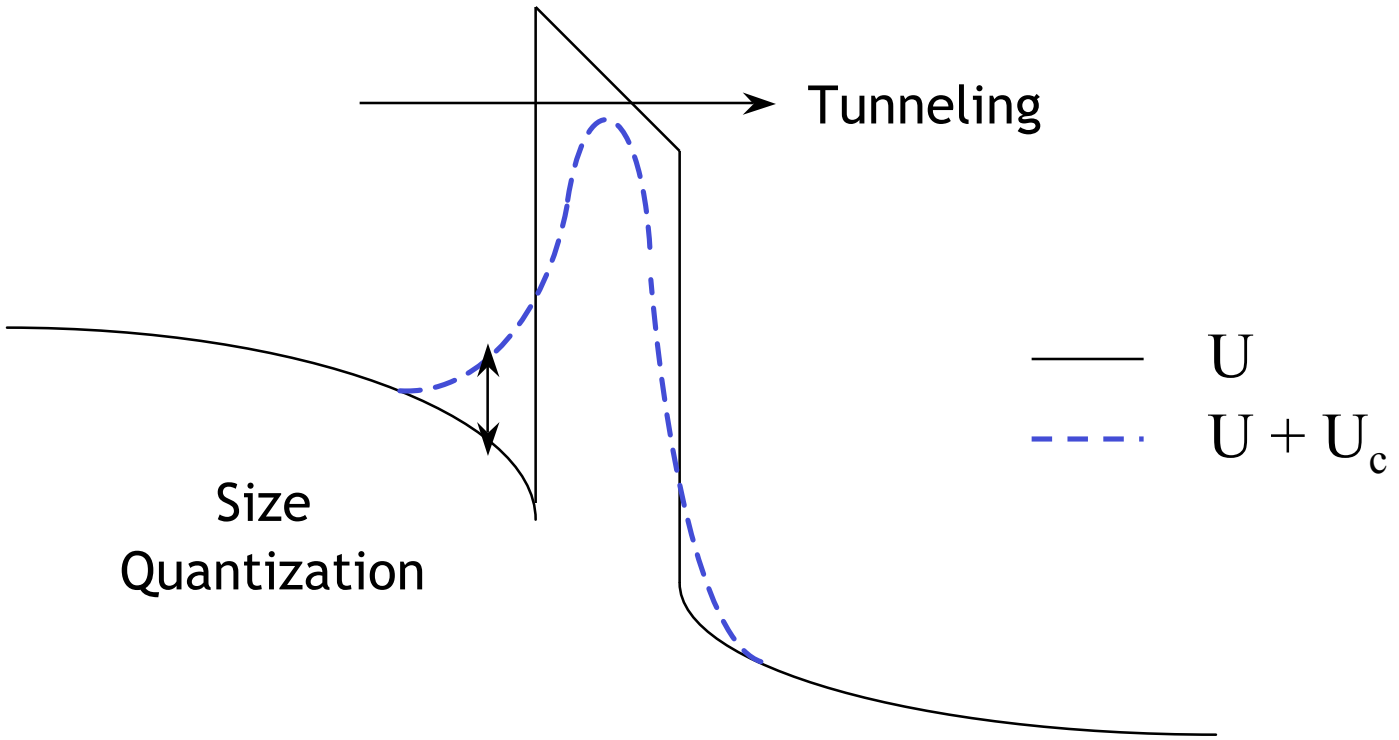
- Monte Carlo particles represent the motion of wave packets centroids in the crystal.
- Monte Carlo does not account for interference effects due to rapidly varying applied fields or hetero-junctions.
- Quantum corrections can capture non-coherent interference effects.
- Coherent transport effects are small for pure silicon devices above 10-nm regime.

Monte Carlo Snapshot of a MOSFET

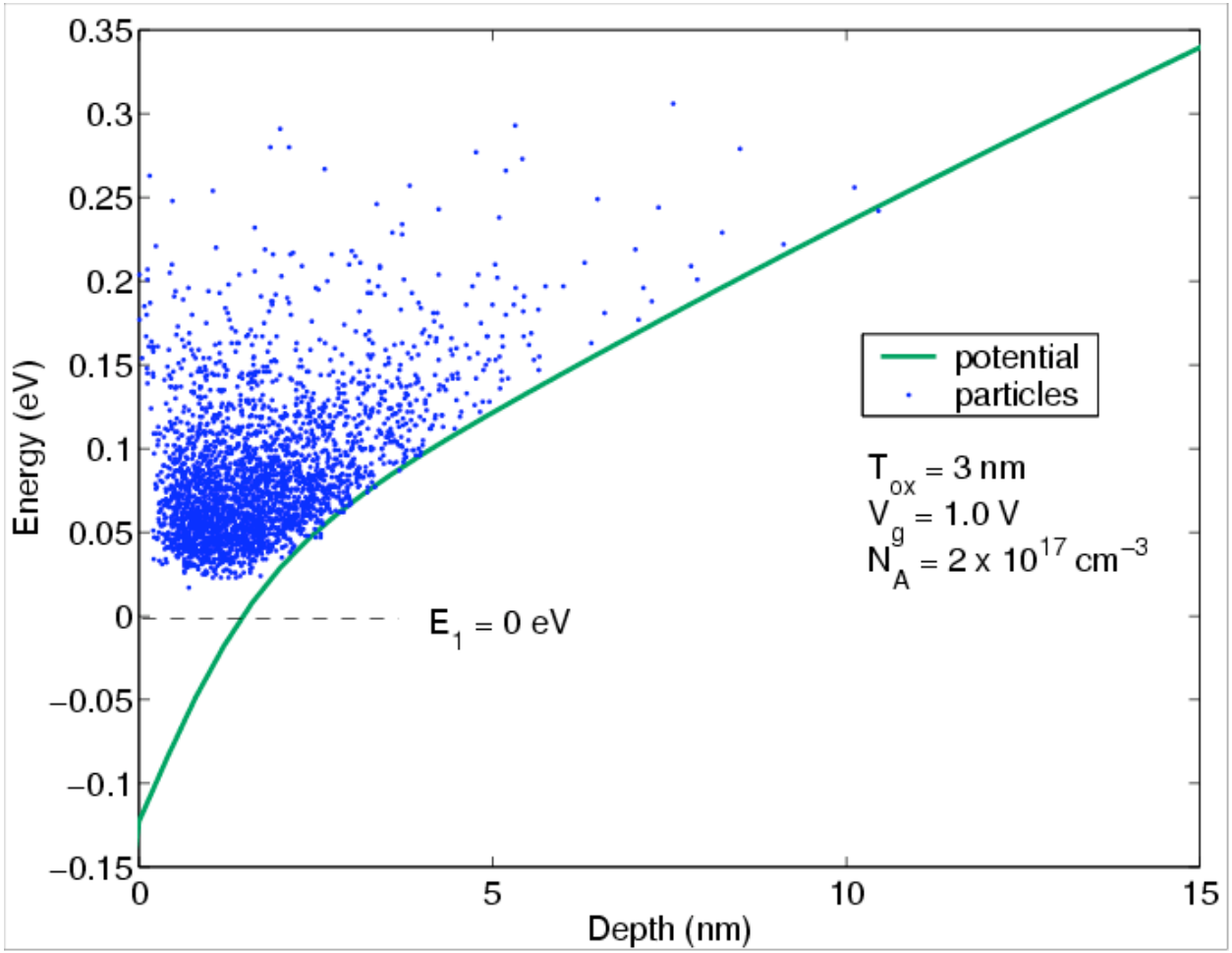


Corrections in Monte Carlo

Quantum effects can be approximated in Monte Carlo by correcting the classical potential



Snapshot from MOS Capacitor



Effective Potential [3]

- Feynman developed the effective potential in the 1960s and applied it to quantum corrections in statistical mechanics
- Particles feel nearby potential due to quantum fluctuations around classical path of least action
- V_{eff} is a non-local function of the nearby potential

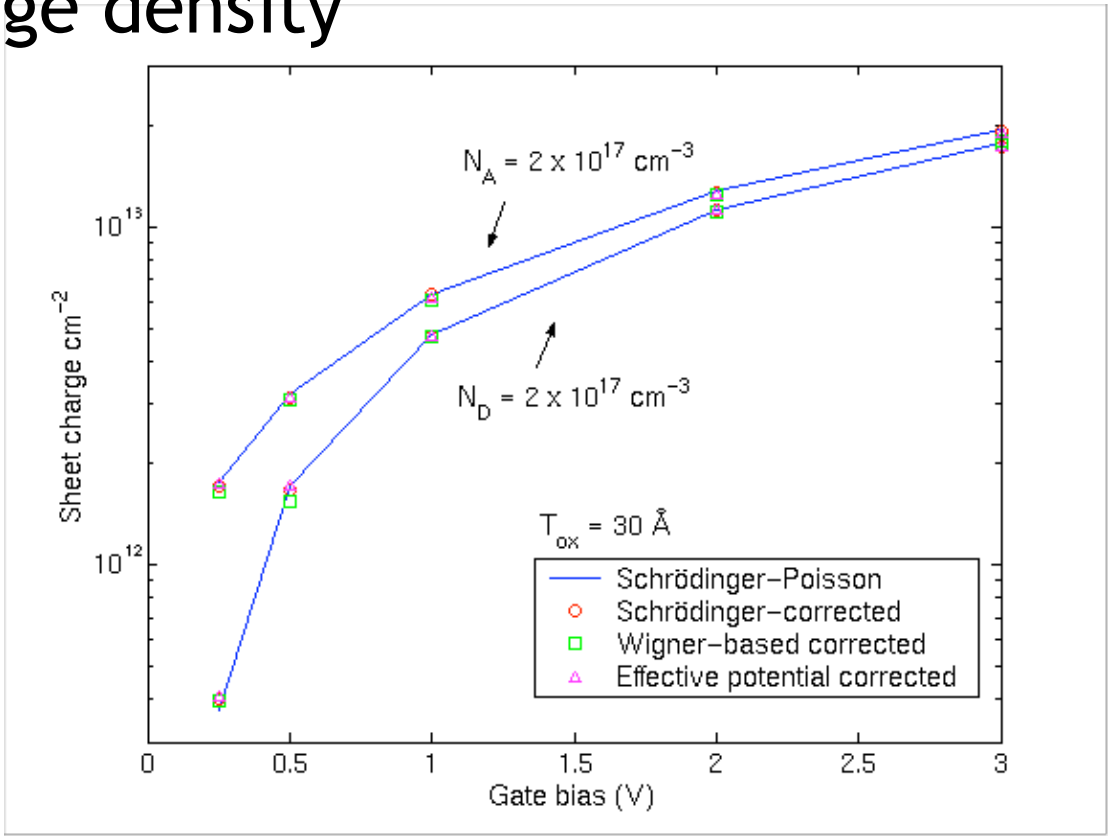
$$V_{eff}(x) = \int V(x) e^{-\frac{(x-x')^2}{2a^2}} dx' \quad , \quad a = \frac{\hbar^2}{12mkT}$$

Properties of Effective Potential

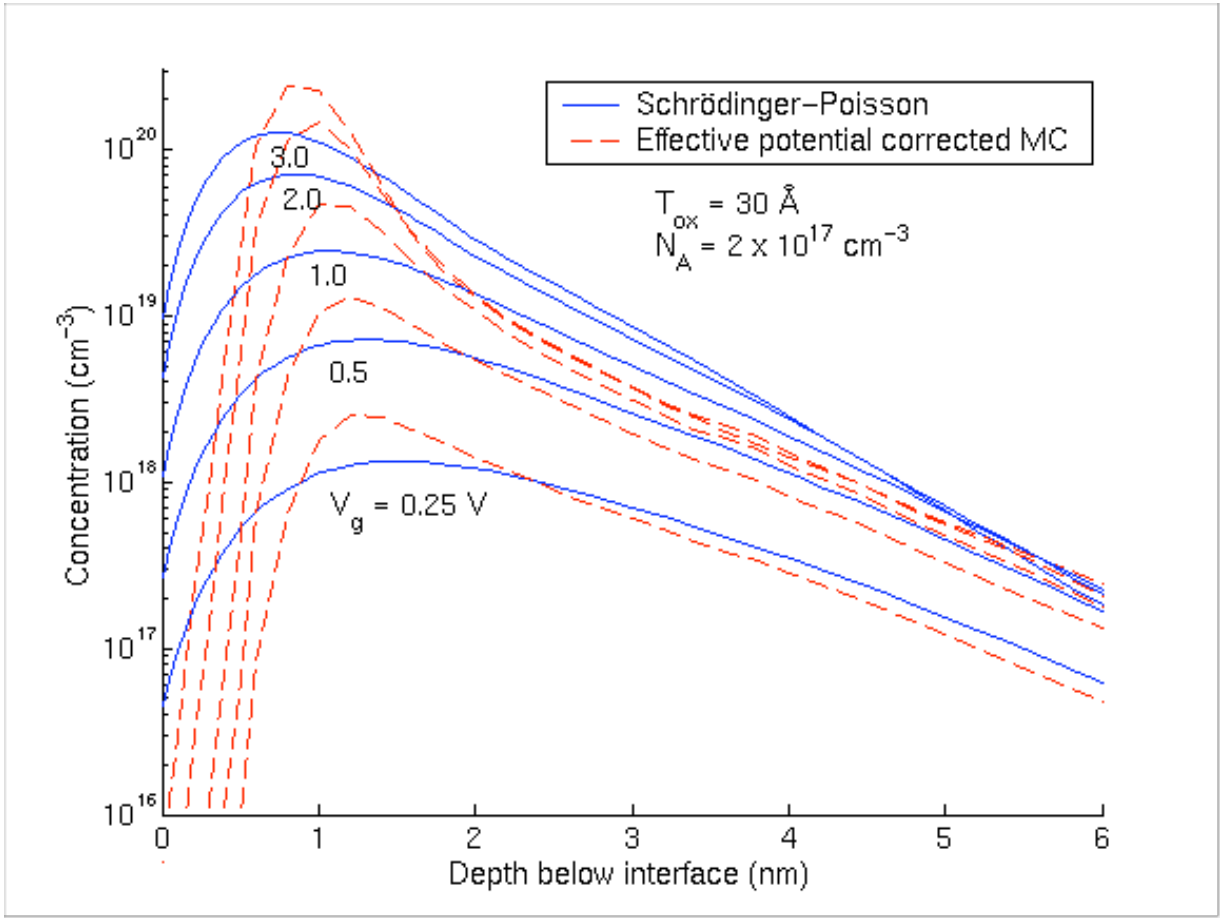
- Simple to implement and to calculate
- Not sensitive to noise from Monte Carlo
- Works best for smooth, symmetric potentials
- *a* can be treated as a fitting parameter describing the "size" of the particle
- Detailed solution near large heterojunctions is typically incorrect and cannot be fit

Effective potential for MOS

- Effective size *a* can be tuned to match sheet charge density



Effective potential for MOS



Parameter-free effective potential [4]

Recently, a more general form of the effective potential has been introduced (Ahmed, Ringhofer and Vasileska, 2005) which depends explicitly on the wave vector *k* and does not contain fitting parameters:

$$V^{Q}(x,k) = \frac{1}{(2\pi)^3}\int \frac{2m^*}{\beta \mathrm{h}^2 k\cdot\xi}\sinh\left(\frac{\beta \mathrm{h}^2 k\cdot\xi}{2m^*}\right)$$
$$\times \exp\left(-\frac{\beta \mathrm{h}^2}{8m^*}|\xi|^2\right) V(y)\, e^{i\xi\cdot(x-y)}\, dy\, d\xi$$

Here, *V(y)* is the sum of the barrier discontinuity at the interface and the solution of Poisson equation. The standard effective potential is a particular case.

Wigner Formulation [5]

- The Wigner method was developed in the 1930s for quantum correction to statistical mechanics
- Wigner formulation of quantum mechanics formally separates the quantum and classical contributions to the equation of motion

$$\frac{\partial f_w}{\partial t} = -\frac{\partial H}{\partial p}\frac{\partial f_w}{\partial x} + \frac{\partial H}{\partial x}\frac{\partial f_w}{\partial p}$$

← Classical Boltzmann equation

Quantum Contribution →

$$-\frac{\hbar^2}{24}\frac{\partial^3 H}{\partial x^3}\frac{\partial^3 f_w}{\partial p^3} + \frac{\hbar^4}{1920}\frac{\partial^5 H}{\partial x^5}\frac{\partial^5 f_w}{\partial p^5} + \ldots$$

Wigner formulation for quantum corrections

- **We start from the general Wigner function representation of quantum transport**

$$f_w(\vec{r},p)=\frac{1}{h^3}\int dy\; e^{j\vec{p}\cdot y/h}\rho\left(x+y/2,x-y/2\right)$$

where

$$r\left(x,x'\right)$$

is the density matrix.

- **The Wigner function is the quantum equivalent of the distribution function in the semi-classical Boltzmann equation.**

Wigner formulation for quantum corrections

- The quantum transport equation of the Wigner function has the form (parabolic bands, ballistic)

$$\frac{\partial f_w}{\partial t} + \frac{\vec{p}}{m^*} \cdot \nabla f_w = \left(\frac{1}{\pi \hbar^2} \right)^3 \iint d\vec{s}\, d\vec{P}\, K(s,P)\, f_w(r, p+P, t)$$

$$K(\vec{s},\vec{P}) = \left[U\left(r + \frac{\vec{s}\hbar}{2} \right) - U\left(r - \frac{s}{2} \right) \right] \sin(s \cdot P)$$

- Direct solution of the Wigner transport equation is still a considerable numerical challenge.
- We are interested in determining a truncated expansion of the quantum equation, that resembles the standard Boltzmann equation, so that the standard Monte Carlo technique can be applied with minor modifications.

Wigner formulation for quantum corrections

- **The complete Wigner transport equation, inclusive of collision terms, can be reformulated to resemble Boltzmann equation**

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \nabla_{\vec{r}} f - \frac{1}{\hbar} \nabla_{\vec{r}} U \cdot \nabla_{\vec{k}} f$$

$$+ \sum_{\alpha=1}^{\infty} \frac{(-1)^{\alpha+1}}{\hbar 4^{\alpha} (2\alpha+1)!} \left(\nabla_{\vec{r}} U \cdot \nabla_{\vec{k}} f \right)^{2\alpha+1} = \left(\frac{\partial f}{\partial t} \right)_C$$

- **At first order, we truncate considering only $\alpha = 1$.**

Wigner formulation for quantum corrections

- The truncated equation has a form resembling Boltzmann equation

$$\frac{\partial f}{\partial t} + \vec{\mathbf{v}} \cdot \nabla_{\mathbf{r}}^{\vec{}} f - \frac{1}{\hbar} F^{qc} \cdot \nabla_{\mathbf{k}}^{\vec{}} f = \left(\frac{\partial f}{\partial t} \right)_C$$

  where F^{qc} contains the quantum correction to the forces.

- With the modified forces, the particles move as if under the influence of a classical potential, but following equivalent quantum trajectories.
- The quantum correction essentially modifies the potential energy felt by the particles.

Wigner formulation for quantum corrections

- **The corrected forces can be evaluated by making assumptions on distribution function and bandstructure. For**

$$f = \exp\left\{-\frac{1}{k_B T}\left[E_{\mathrm{k}-\bar{\mathrm{k}}} + U(\mathrm{r}) - E_f\right]\right\}$$

Displaced Maxwellian

$$E_{\mathrm{k}-\bar{\mathrm{k}}} = \sum_{i=x,y,z} \frac{\hbar^2 \left(k_i - \bar{k}_i\right)^2}{2m_i}$$

$$\bar{k} = \left(\bar{k}_x, \bar{k}_y, \bar{k}_z\right) = \text{momentum centroid}$$

Parabolic bands

$$F_x^{qc} = -\frac{\partial}{\partial x}\left(U - \frac{1}{24}\left\{\left[\gamma_x^2\left(k_x - \bar{k}_x\right)^2 - 3\gamma_x\right]\frac{\partial^2 U}{\partial x^2} + 3\left[\gamma_y^2\left(k_y - \bar{k}_y\right)^2 - \gamma_y\right]\frac{\partial^2 U}{\partial y^2}\right\}\right)$$

$$F_y^{qc} = -\frac{\partial}{\partial y}\left(U - \frac{1}{24}\left\{\left[\gamma_y^2\left(k_y - \bar{k}_y\right)^2 - 3\gamma_y\right]\frac{\partial^2 U}{\partial y^2} + 3\left[\gamma_x^2\left(k_x - \bar{k}_x\right)^2 - \gamma_x\right]\frac{\partial^2 U}{\partial x^2}\right\}\right)$$

with $\gamma_x = \frac{\hbar^2}{m_x k_B T}$; $\gamma_y = \frac{2}{m_y k_B T}$

Wigner formulation for quantum corrections

- The forces obtained have still some practical problems in regions like sharp interfaces, where quantum effects are prominent.
- To obtain a smooth potential, we can use approximate relations obtained by integrating the displaced Maxwellian distribution with the momentum.
- We obtain these alternative expressions for the second order derivatives of the potential

$$\frac{\partial^2 U}{\partial x^2} ;; -k_B T \frac{\partial^2 (\ln n)}{\partial x^2} ; \qquad \frac{\partial^2 U}{\partial y^2} \quad -k_B T \frac{\partial^2 (\ln n)}{\partial y^2}$$

Wigner formulation for quantum corrections

- **The "smooth" version of the quantum corrected forces is**

$$F_x^{qc} = -\frac{\partial}{\partial x}\left(U - \frac{k_B T}{24}\left\{ \left[\gamma_x^2 \left(k_x - \bar{k}_x \right)^2 - 3\gamma_x \right] \frac{\partial^2 \ln n}{\partial x^2} + 3\left[\gamma_y^2 \left(k_y - \bar{k}_y \right)^2 - \gamma_y \right] \frac{\partial^2 \ln n)}{\partial y^2} \right\} \right)$$

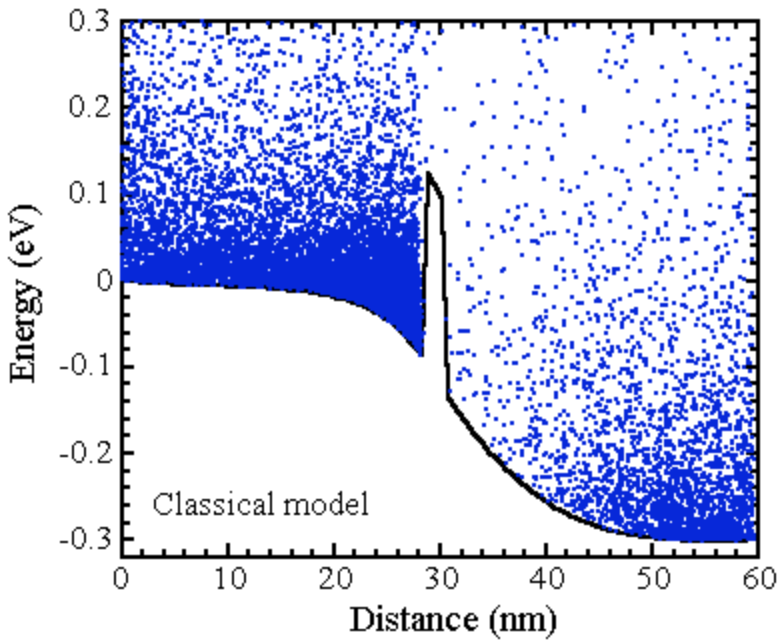
$$F_y^{qc} = -\frac{\partial}{\partial y}\left(U - \frac{k_B T}{24}\left\{ \left[\gamma_y^2 \left(k_y - \bar{k}_y \right)^2 - 3\gamma_y \right] \frac{\partial^2 \ln n}{\partial y^2} + 3\left[\gamma_x^2 \left(k_x - \bar{k}_x \right)^2 - \gamma_x \right] \frac{\partial^2 \ln n}{\partial x^2} \right\} \right)$$

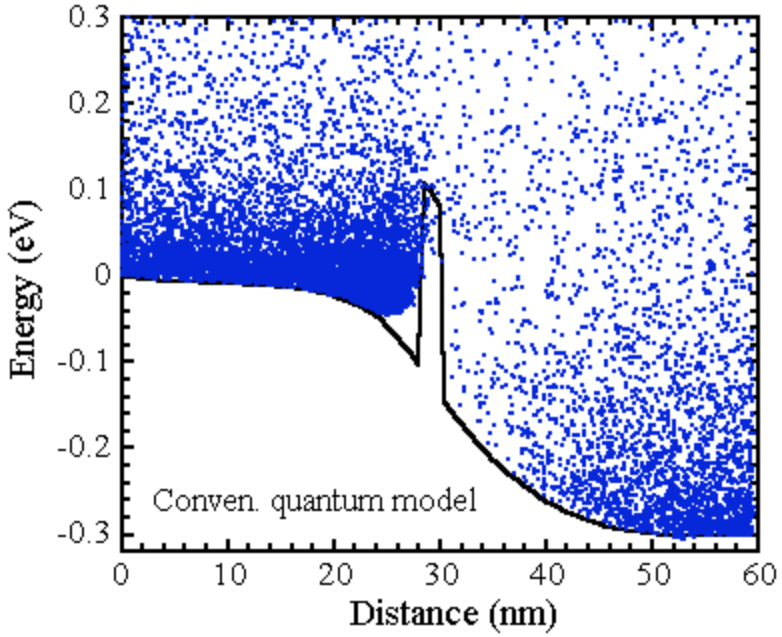
- **This formulation has explicit momentum dependence and improves upon previous results in the literature where the momentum terms were evaluated with the thermal energy**

$$\frac{\mathrm{h}^2 \left(k_i - \bar{k}_i \right)^2}{2m_i} = \frac{k_B T}{2}$$

Wigner formulation for quantum corrections

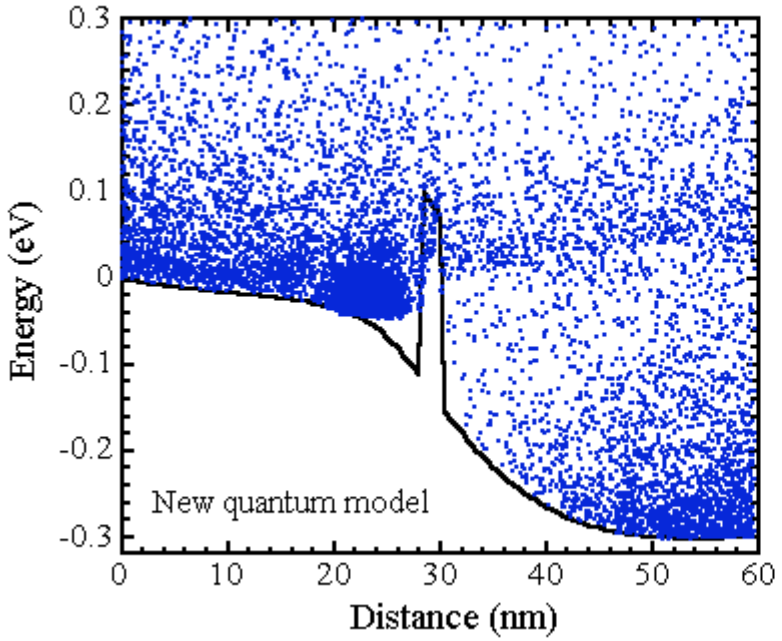
- **Test of the quantum corrections in Monte Carlo:**
  Single GaAs/AlGaAs/GaAs barrier with a fixed potential (1)

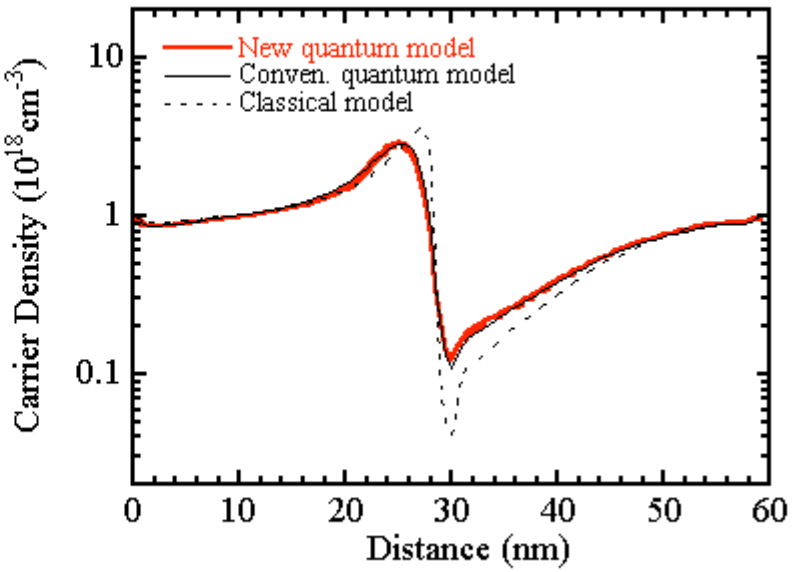




Wigner formulation for quantum corrections

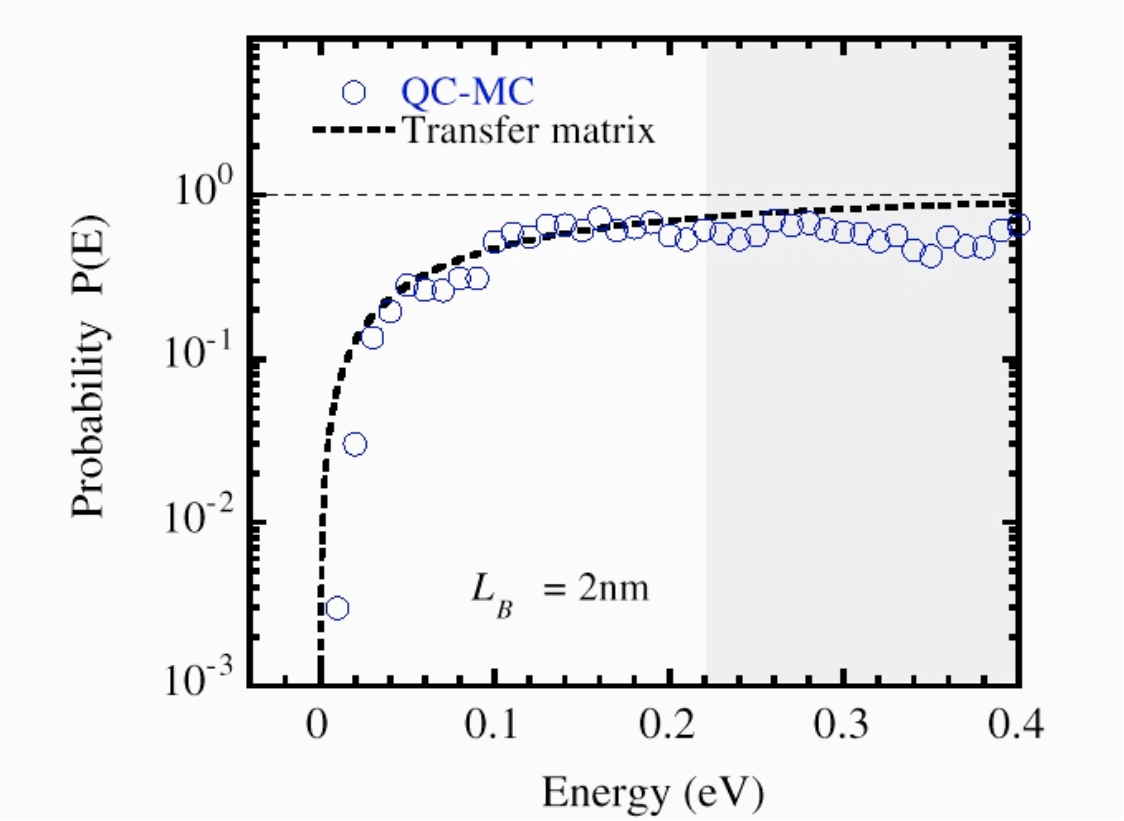
- **Single GaAs/AlGaAs/GaAs barrier with a fixed potential (2)**





Wigner formulation for quantum corrections

- **Single GaAs/AlGaAs/GaAs barrier with a fixed potential (3)**

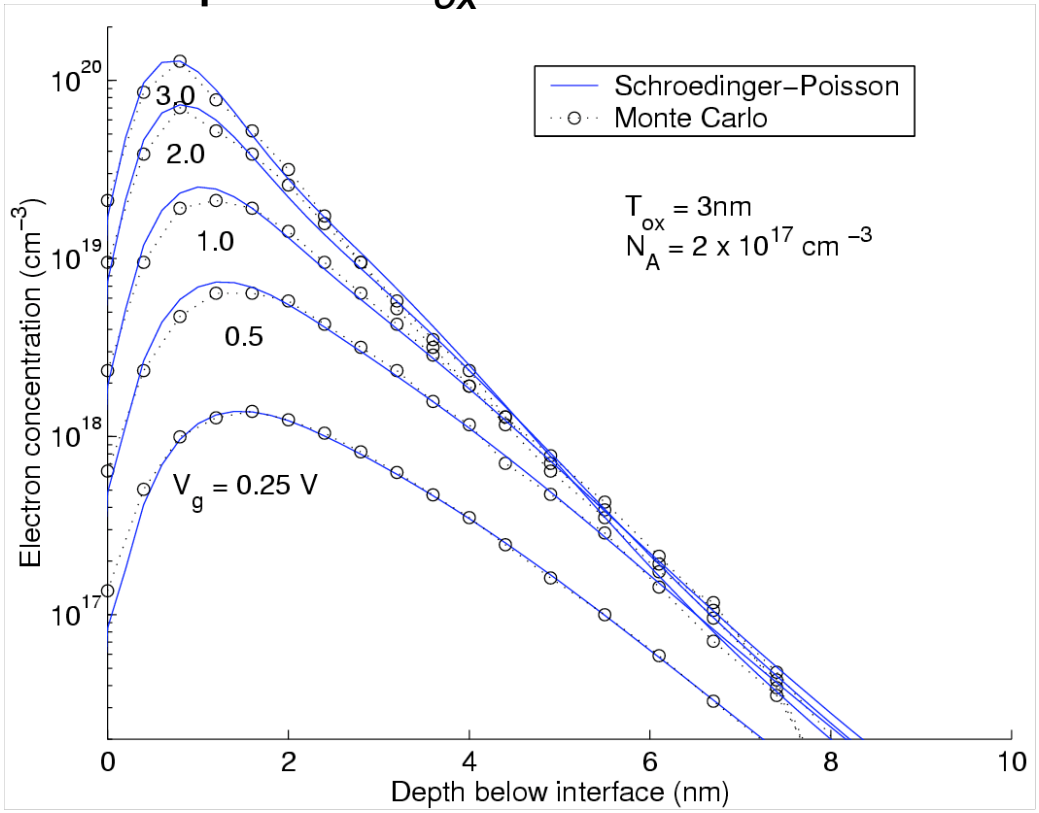


Properties of Wigner Correction

- Requires long run times due to Monte Carlo noise in $\nabla^2 \ln(n)$ and restricts grid spacing
- Works well in drift-diffusion where noise is not an issue
- Unlike effective potential, $\nabla^2 \ln(n)$ is *local*
- For MOS, a single fitting parameter was found to adjust the correction at the oxide interface
- Requires no fitting in the silicon region

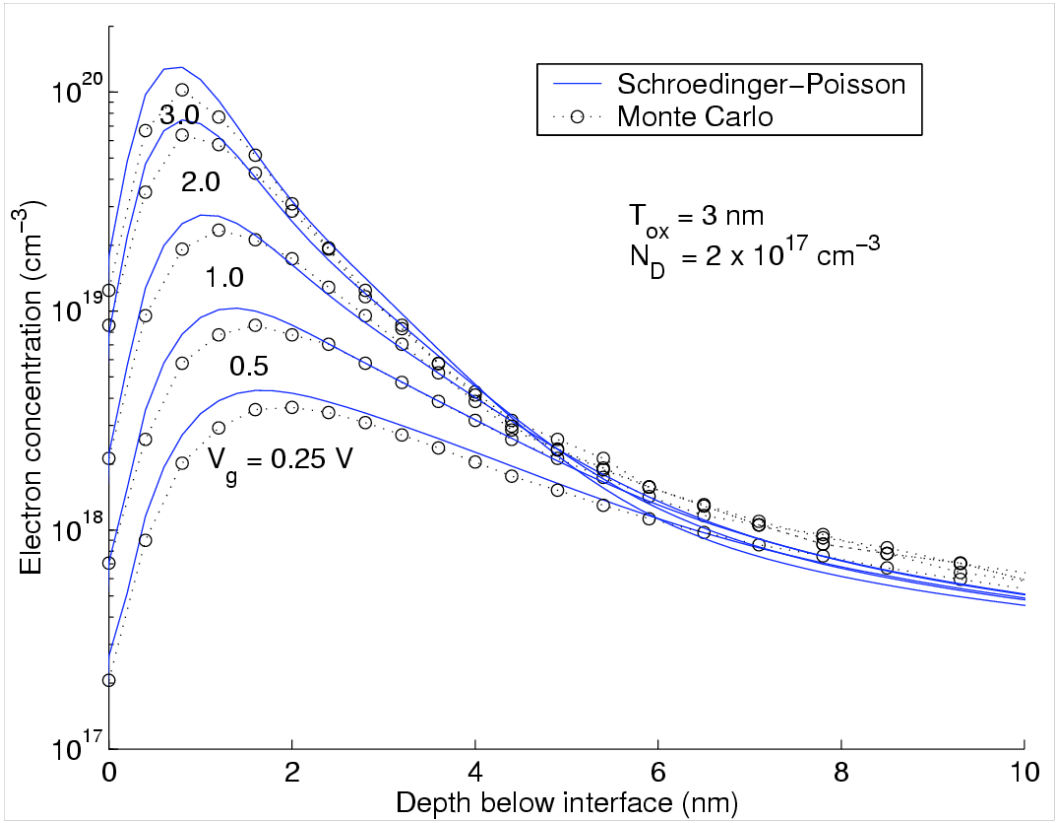
Wigner-corrected MOS Inversion

Wigner correction is accurate across range of biases—empirical $n_{ox} = 10^{15}$ cm^{-3} to fit interface



Wigner-corrected MOS Accumulation

$n_{ox} = 10^{15}$ cm^{-3}



Bohm potential correction [6]

$$\psi = R \cdot \exp\left(i\, S / \mathrm{h}\right)$$

$$E = V\left(\vec{r}\right) - \frac{\mathrm{h}\hbar}{4m}\left[\frac{\nabla^2 P}{P} - \frac{\left(\nabla P\right)^2}{2P^2}\right] = V\left(r\right) - \frac{^2}{2m}\frac{\nabla^2 R}{R}$$

$$E = \text{ eigenenergy}$$

$$P\left(\vec{r}\right) = R^2\left(r\right) = \text{probability density } \propto n\left(r\right)$$

The eigenenergy plays the role of the effective total potential, or effective conduction band edge

$$V^*\left(\vec{r}\right) \approx V\left(r\right) - \frac{\mathrm{h}^2}{2m}\frac{\nabla^2\sqrt{n}}{\sqrt{n}} = V\left(r\right) + V_B^Q\left(r\right)$$

Bohm potential correction

Using $n \propto \exp\left(-V^* / k_B T\right)$ in 1D

$$\frac{d^2 V^*}{d y^2} - \frac{1}{2 k_B T}\left(\frac{d V^*}{d y}\right)^2 = \frac{4 m^* k_B T}{\mathrm{h}^2}\left(V^* - V(y)\right)$$

Effective conduction band edge equation

(valid under thermal equilibrium)

The equation represents a first-order quantum correction to the semi-classical BTE by taking into account the effect of carriers only occupying the quantized ground state. Usually, where the quantum correction is needed, density is not too high (Fermi level is below the first excited state).

Schrödinger-Based Correction [7]

- Treats quantum effects in the direction perpendicular to transport
- Accurate
- No fitting parameters
- Not sensitive to noise in the Monte Carlo concentration estimator
- Efficient, additional computation time is small

Applying Schrödinger Correction

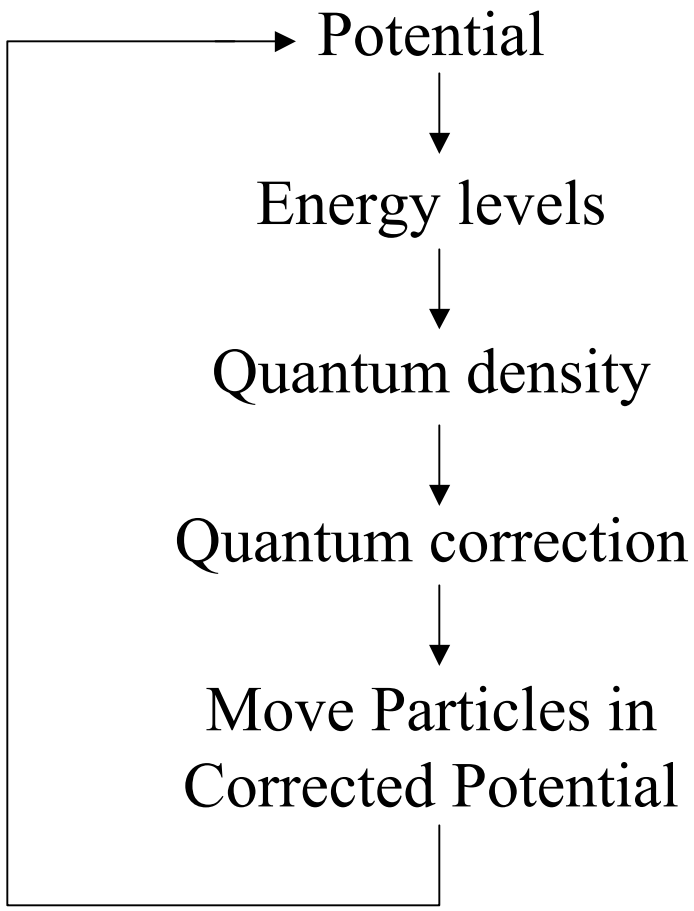
- Schrödinger equation is solved along 1D slices of the 2-D domain
- Self-consistent Monte Carlo potential is the input to Schrödinger and quantum density, n_q, is output
- Concentration is linked to the correction with a Boltzmann dependence

$$n_q(z) \propto e^{-\{V_p(z) + V_{qc}(z)\}/kTt}$$

Consistent with Non-equilibrium Transport

- Schrödinger energy levels/wavefunctions are filled on a Boltzmann distribution
- Within each slice, the correction forces the ***shape*** of the quantum density onto Monte Carlo
- No Fermi level is required
- Relative concentration between the slices is determined naturally by Monte Carlo transport

Quantum-corrected Simulation Flow



$$\nabla \cdot \varepsilon \nabla V_p = -\rho$$

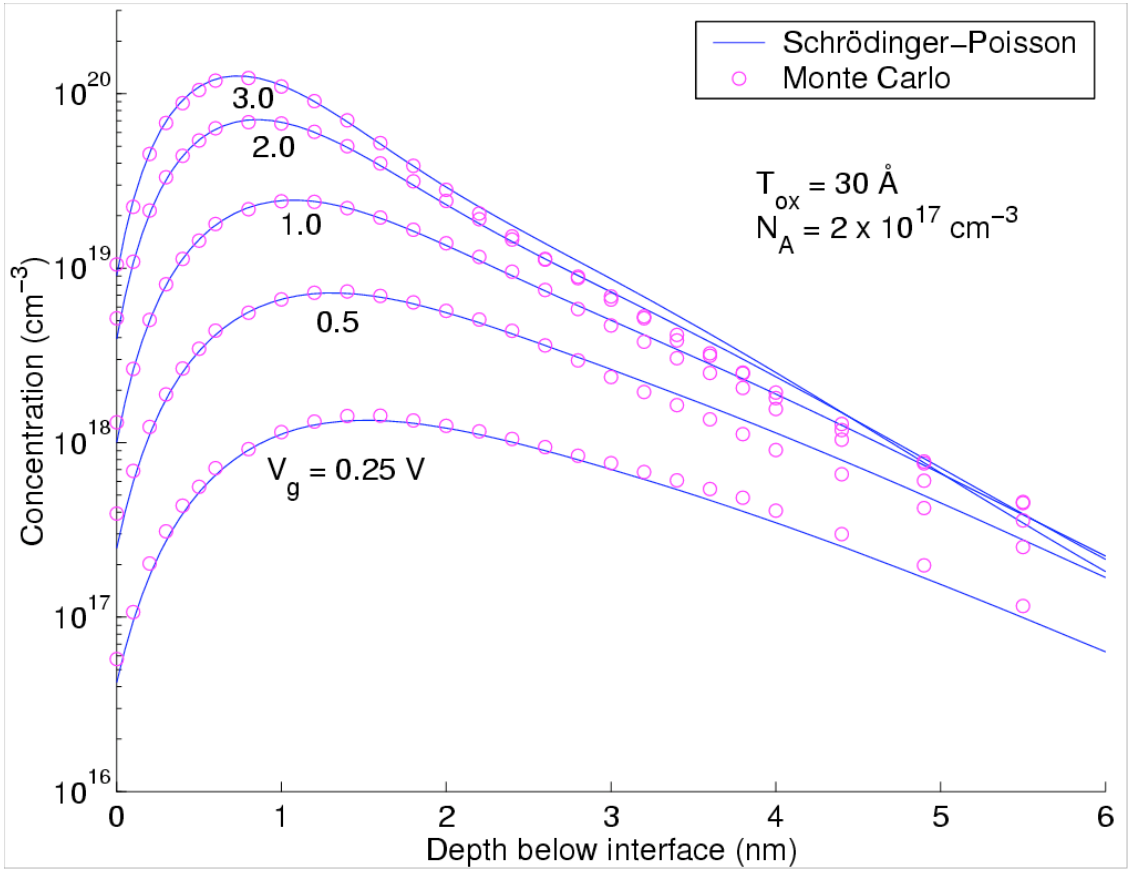
$$-\frac{\mathrm{h}^2}{2}\frac{d}{dz}\left(\frac{1}{m^*}\frac{d\psi}{dz}\right) + V_p\psi = E\psi$$

$$n_q(z) \propto \sum_E \left|\psi(E,z)\right|^2 e^{-E/kT_t}$$

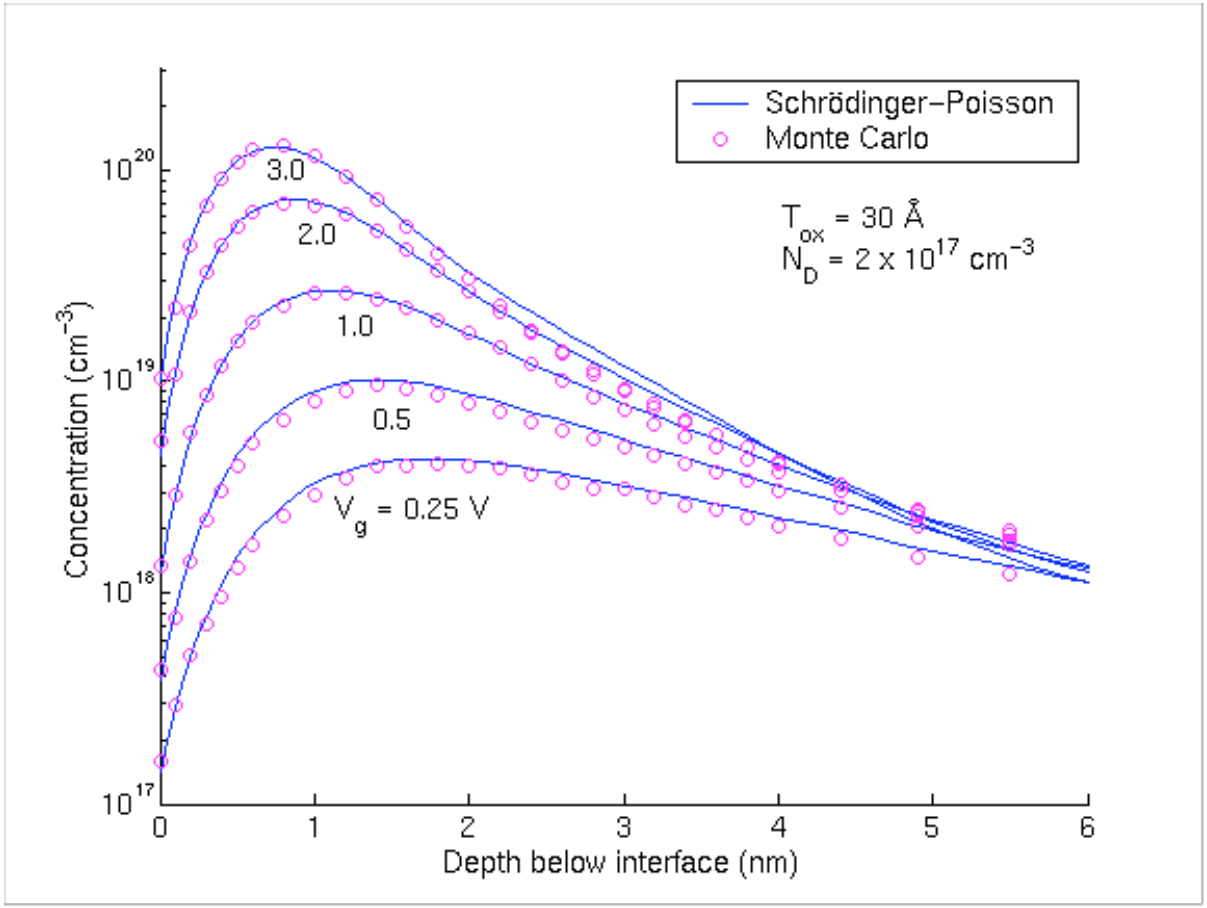
$$-kT_t \log\left(n_q(z)\right) - V_p(z) + V_o$$

$$\frac{d\vec{k}}{dt} = \frac{q}{\hbar}\nabla_{\vec{r}}\left(V_p + V_{qc}\right)$$

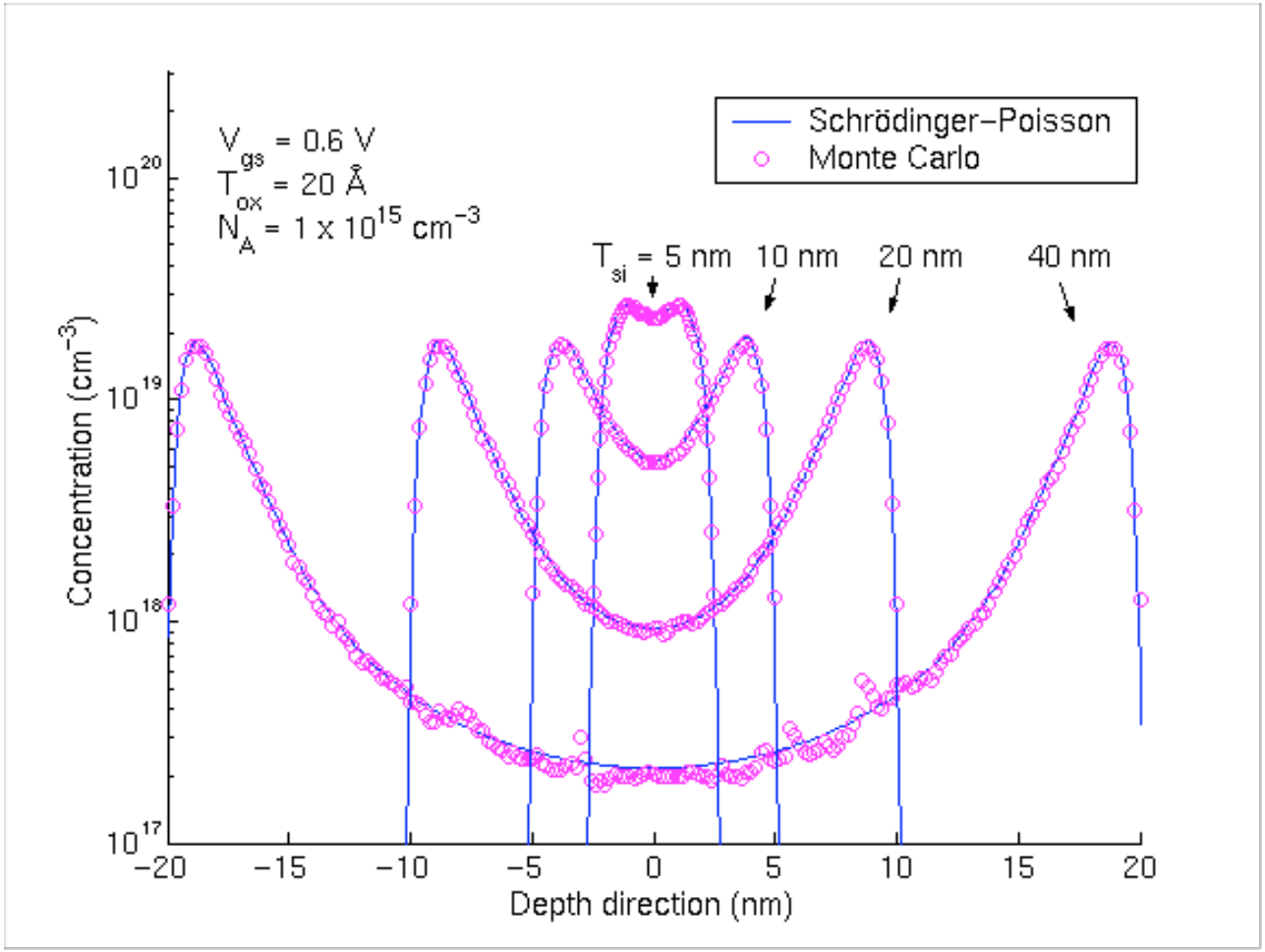
Schrödinger-corrected MOS Inversion



Schrödinger-corrected MOS Accumulation

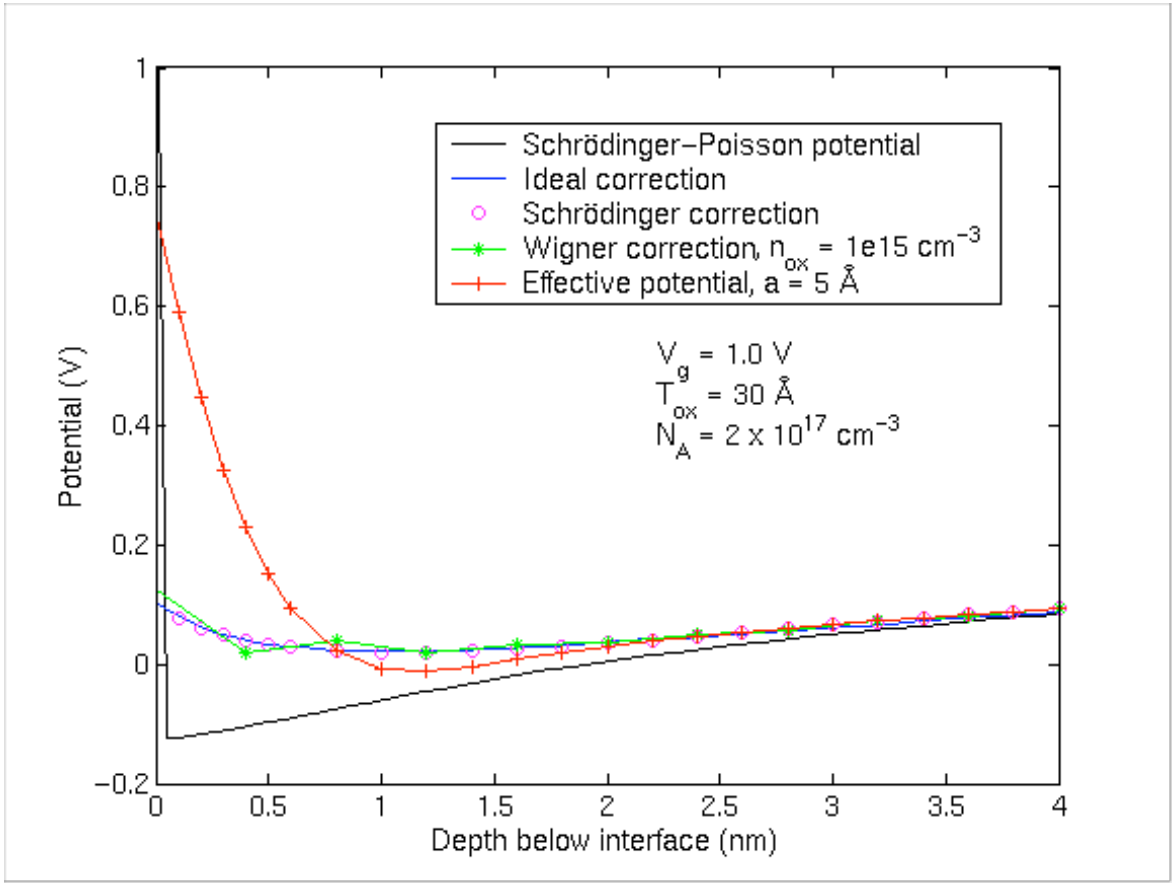


Schrödinger-corrected Double-gate

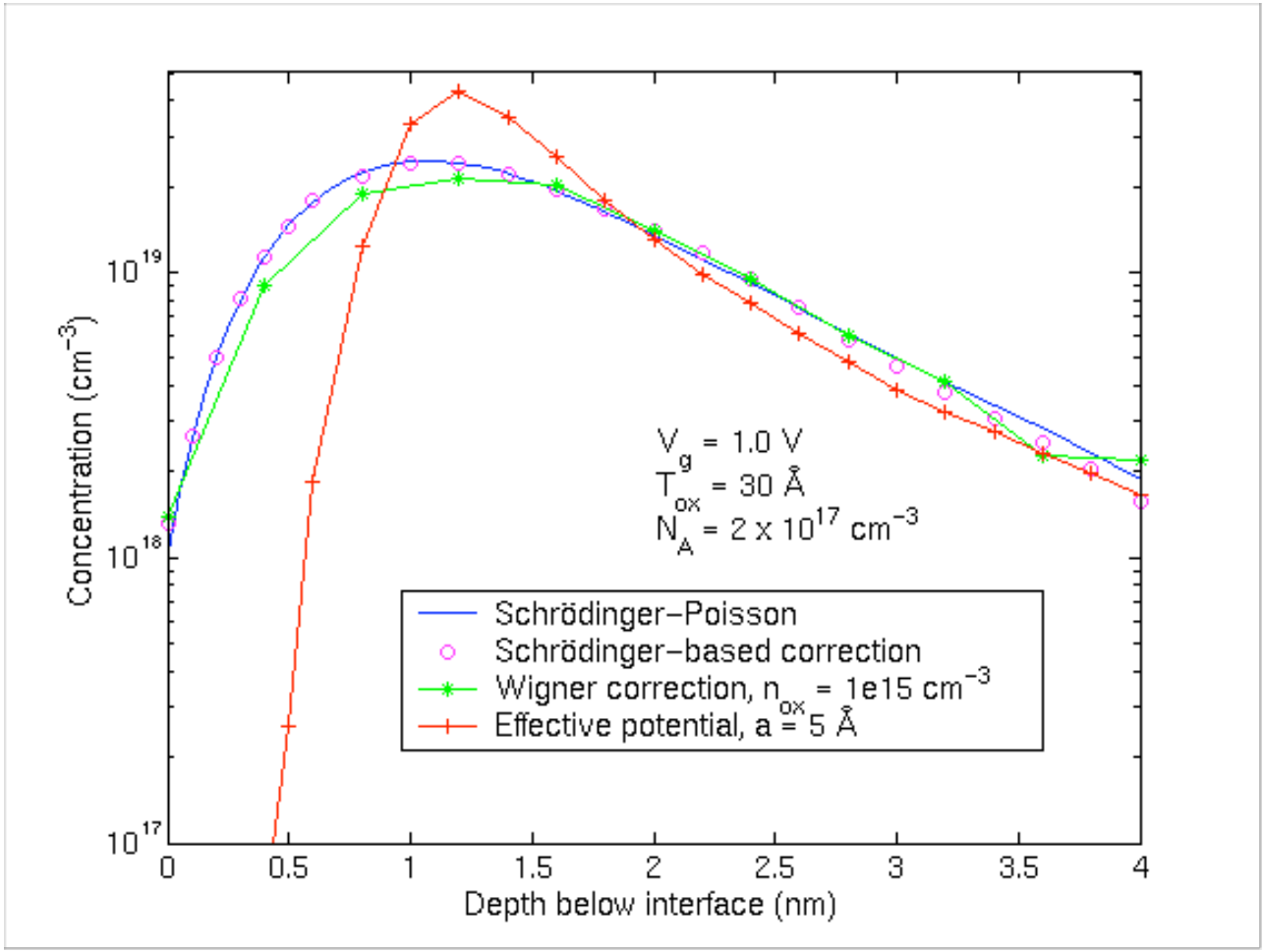


Schrödinger vs Other Corrections for MOS

- Typical behavior of different corrections for MOS



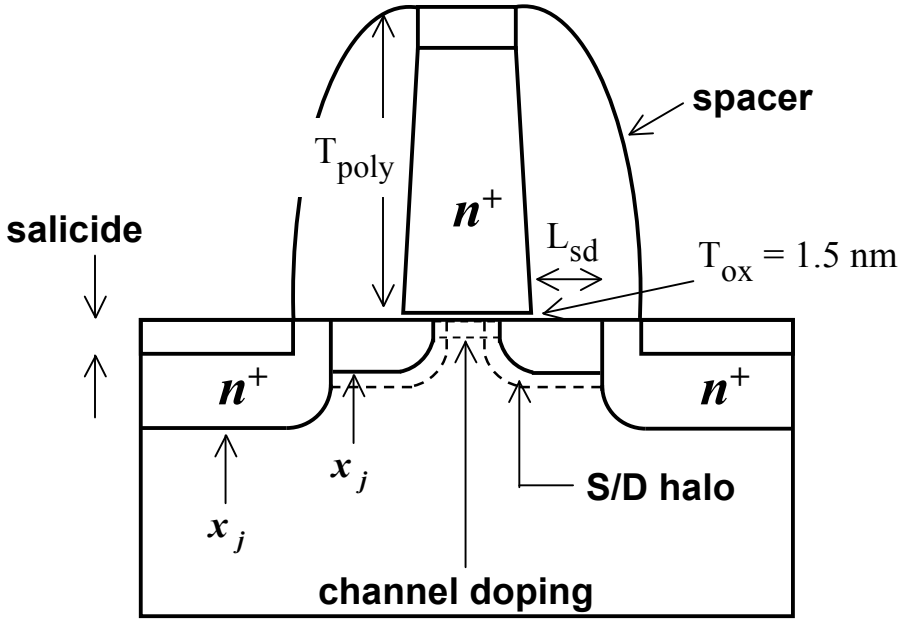
Schrödinger vs Other Corrections for MOS



Extending Schrödinger Correction to Devices

- Heating occurs in the direction $\perp$ to transport
- Cannot make use of an electron "temperature" because it is not well-defined for non-equilibrium
- Define a "transverse" temperature T_t to describe the variation of the concentration with potential in the $\perp$ direction
- T_t is validated if a single temperature at each point along the transport path accurately describes the variation

25 nm Well-tempered MOSFET – 2D Monte Carlo simulation

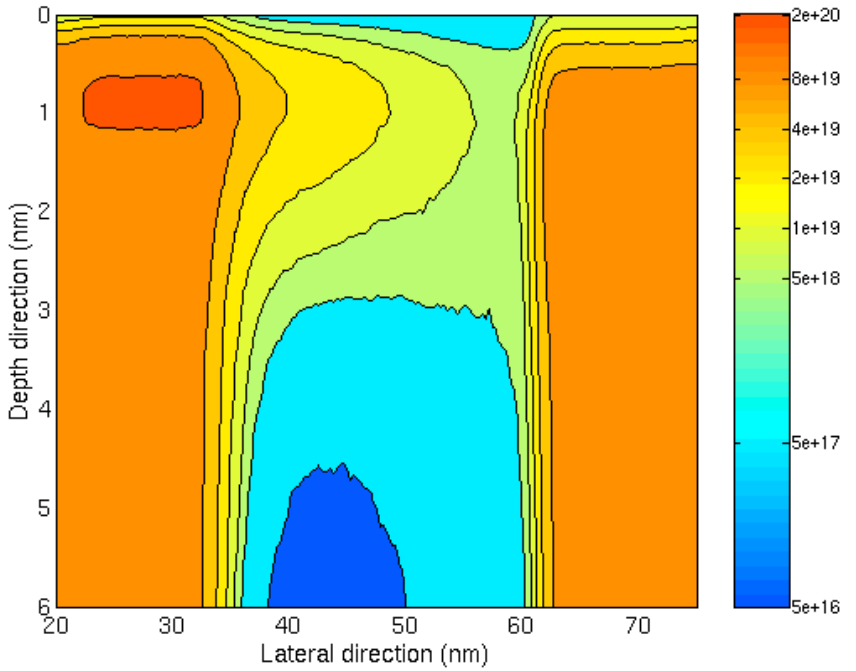


L_{eff} = 25 nm T_{OX} = 15 Å

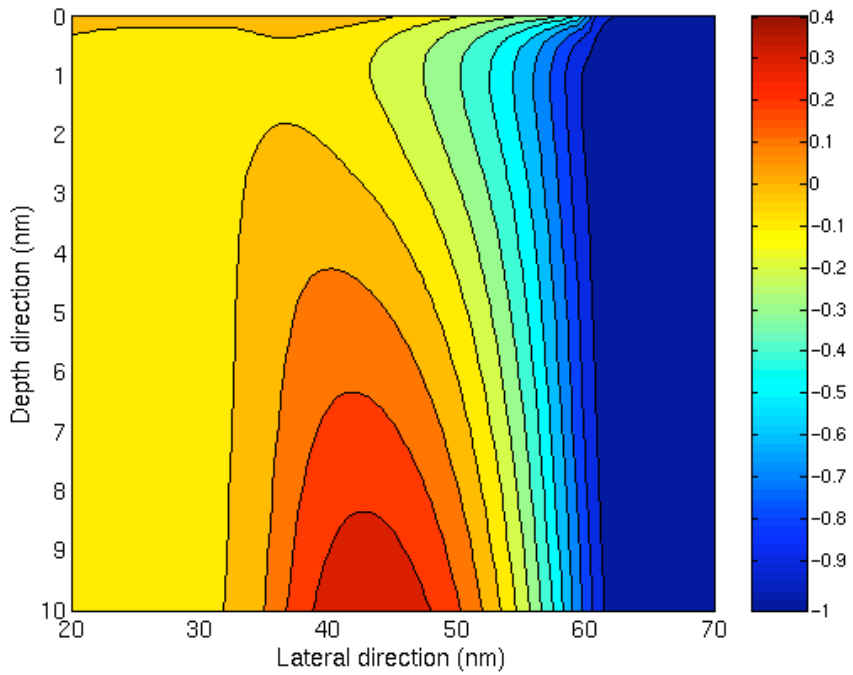
V_G = 1.0 V V_D = 1.0 V

25 nm Well-tempered MOSFET

electron concentration



corrected potential energy

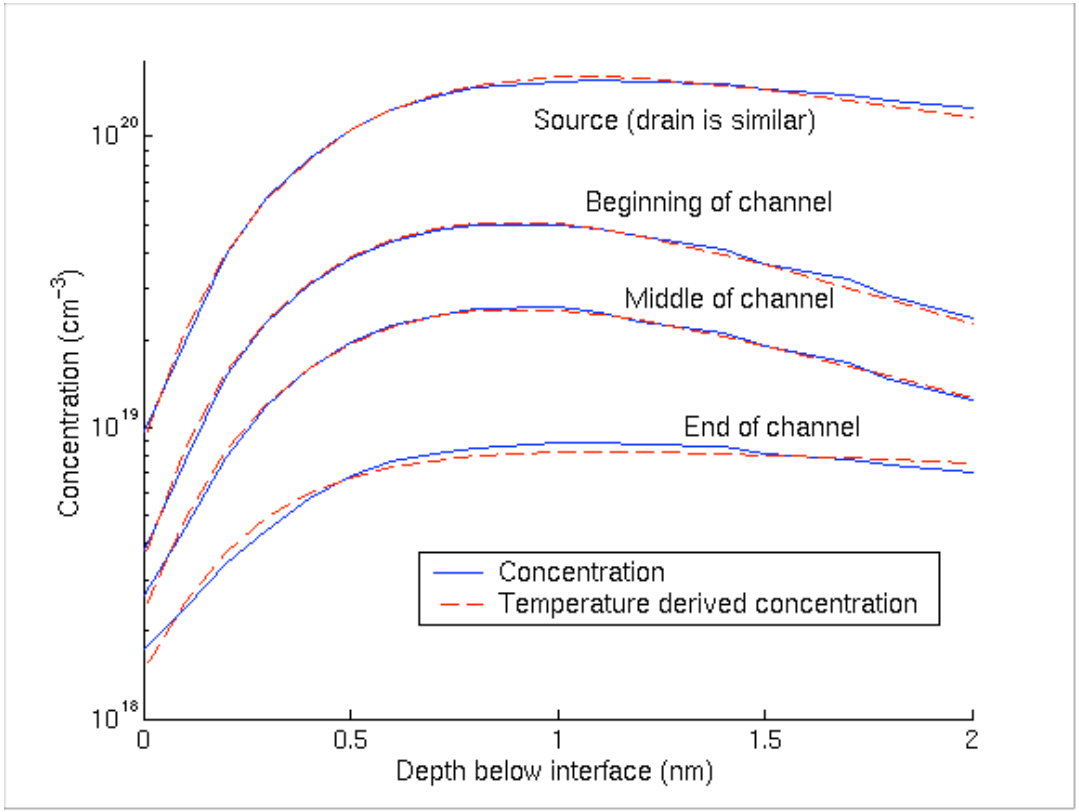


L_{eff} = 25 nm T_{OX} = 15 Å

V_G = 1.0 V V_D = 1.0 V

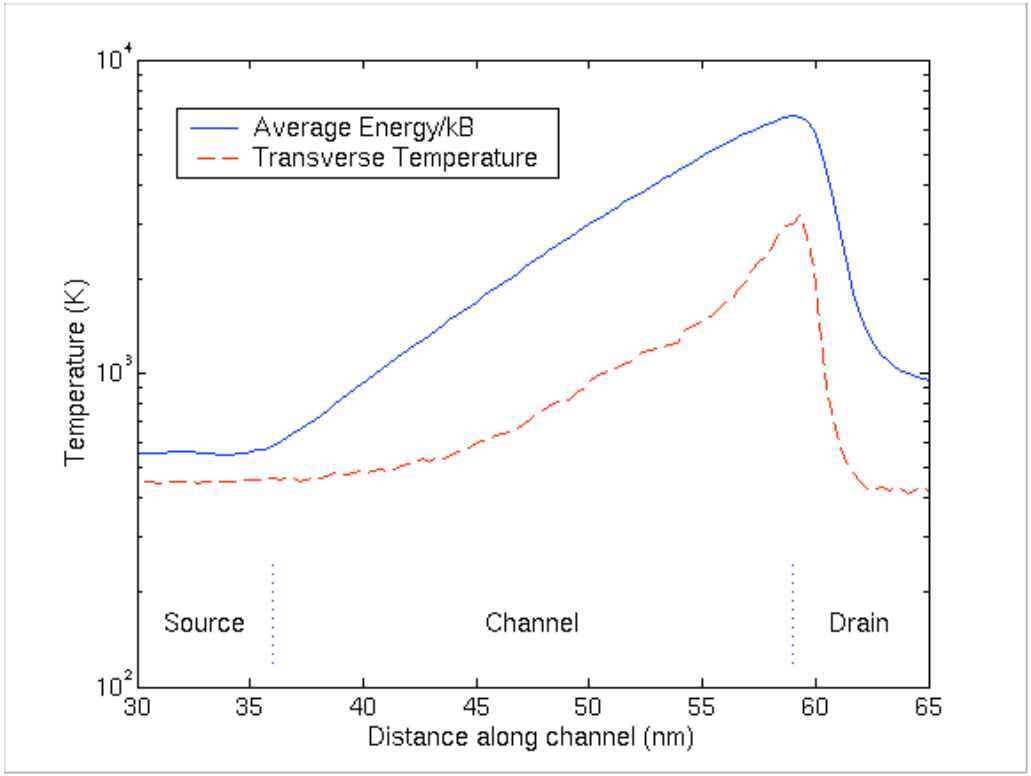
Validating Transverse Temperature

$T_t(z)$ accurately describes potential→concentration for highly non-equilibrium transport in 25nm FET

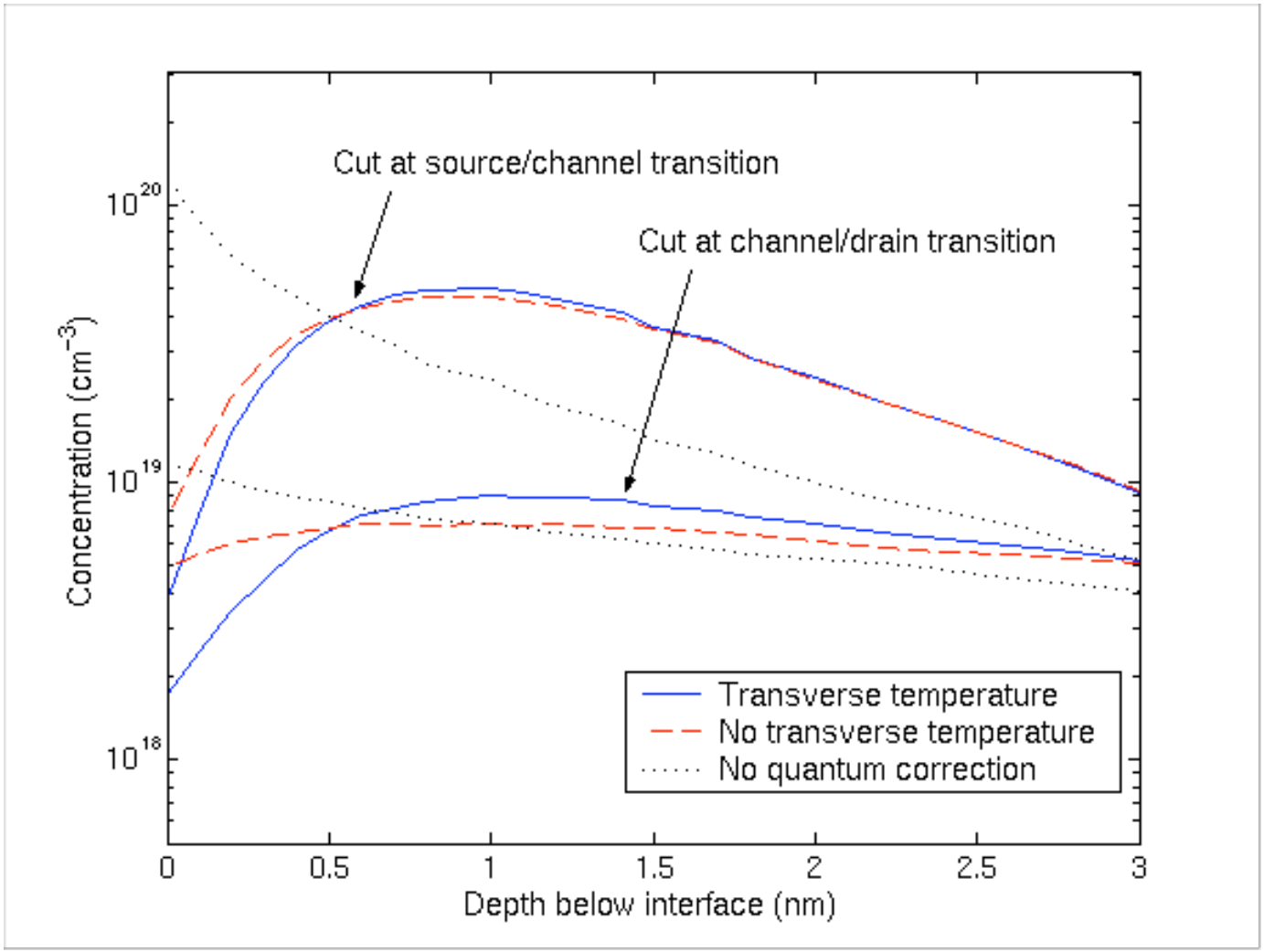


Typical Transverse Temperature

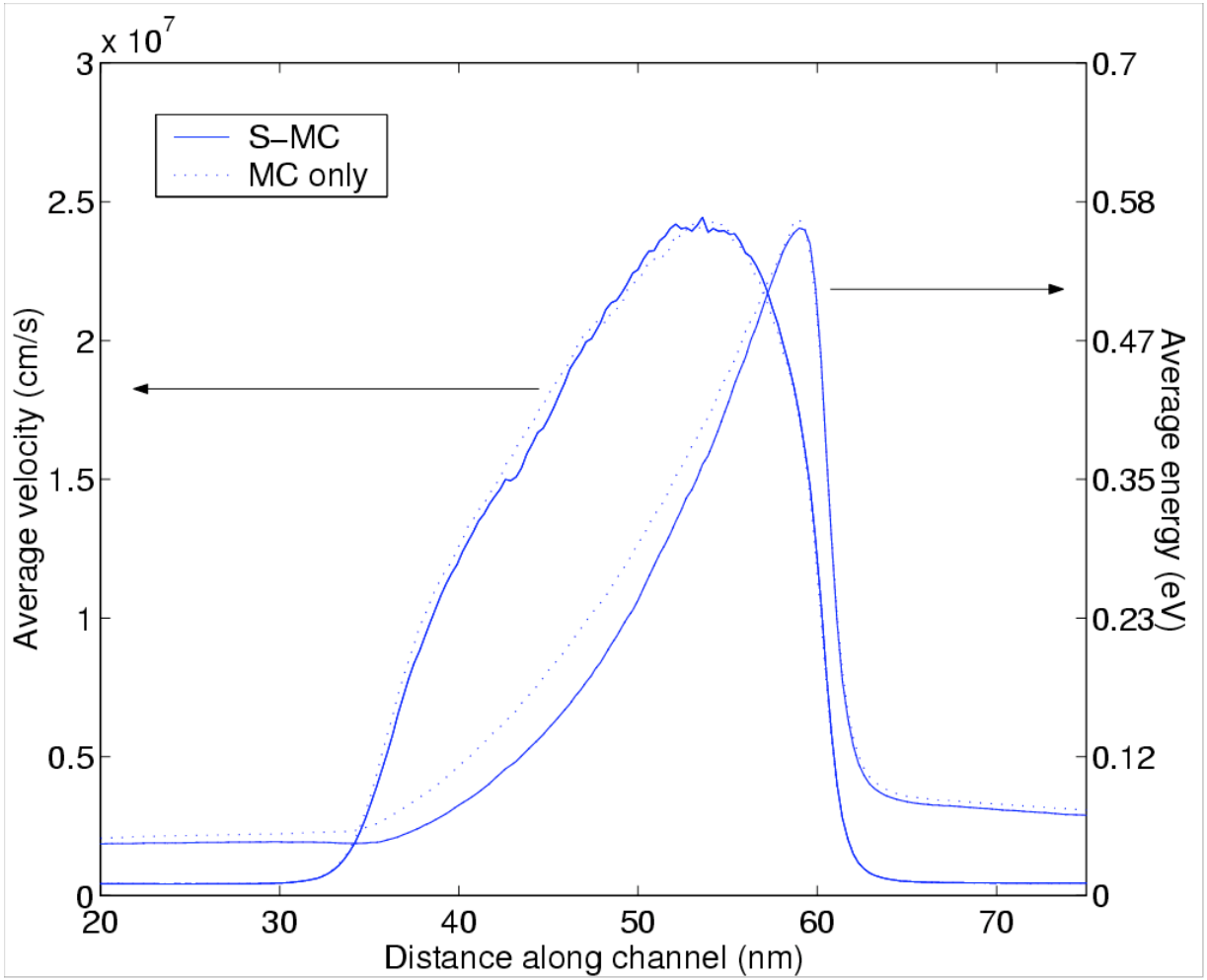
- Transverse temperature for a 25-nm MOSFET in saturation bias



Effect of Heating in the ⊥ Direction



Properties along transport path

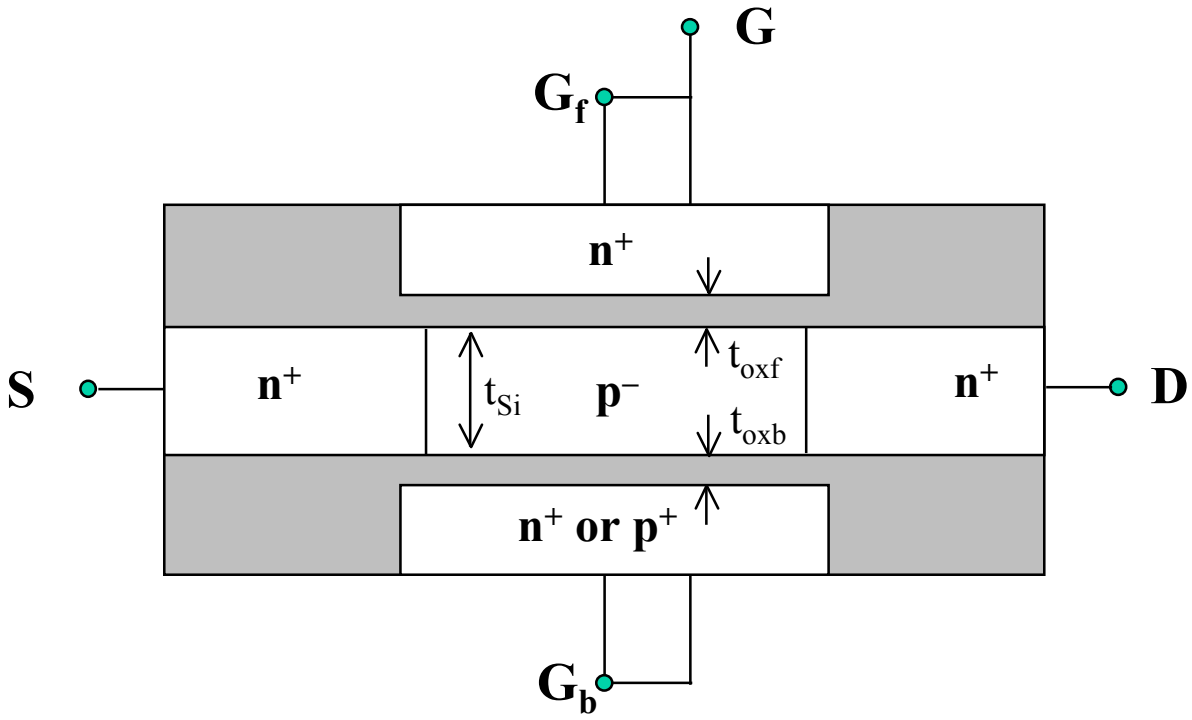


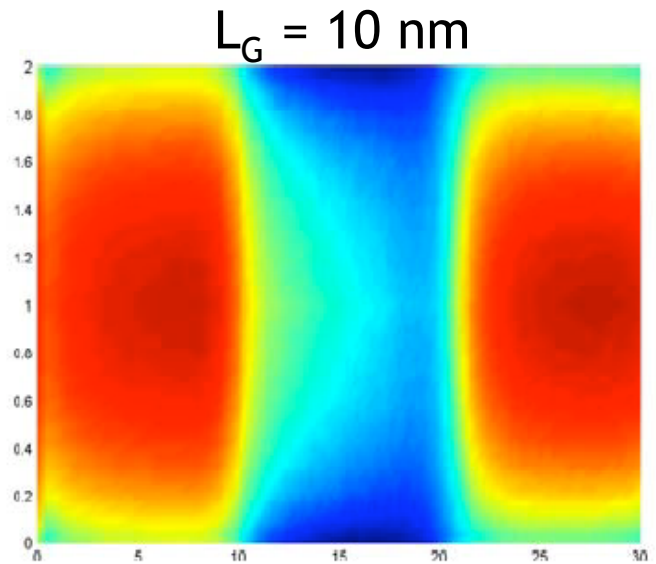
MOS SCALING

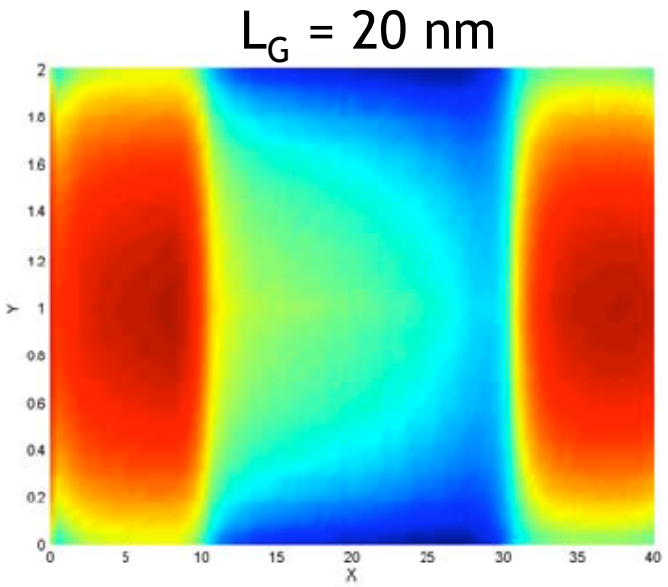
New structures are considered to replace the bulk MOSFET:

- **Silicon-on-insulator (SOI)**
- **Double-gate MOSFET**
- **Schottky-barrier MOSFET (silicide contacts)**
- **FinFET**
- **Quantum wire MOSFET**

Double-Gate MOSFET - Volume inversion effect



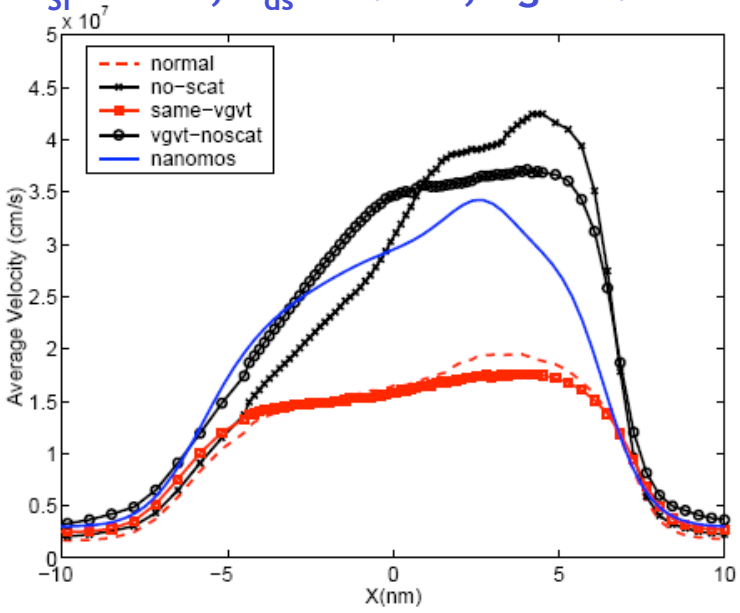




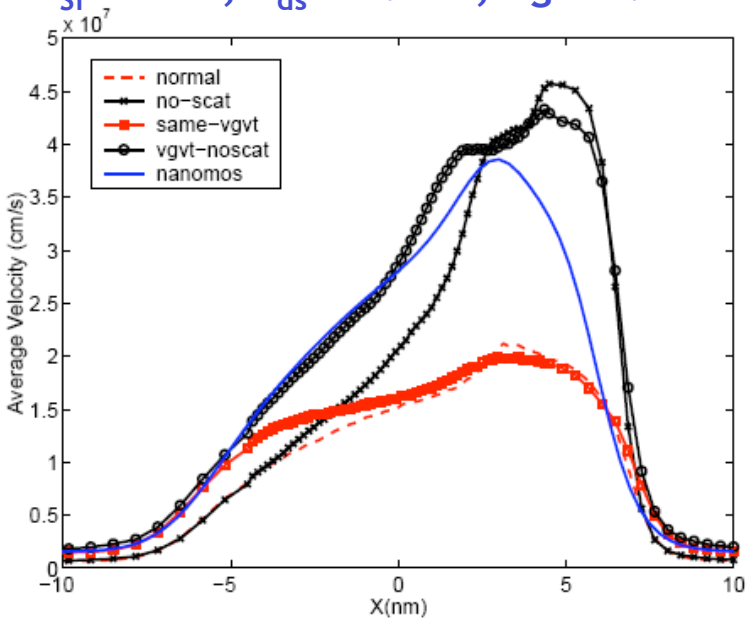
Rule of thumb: Comparisons with NEGF simulations (Purdue nanoMOS) indicate that quantum corrected Monte Carlo is accurate for layers of thickness in the range t_{Si} ~ 3nm.

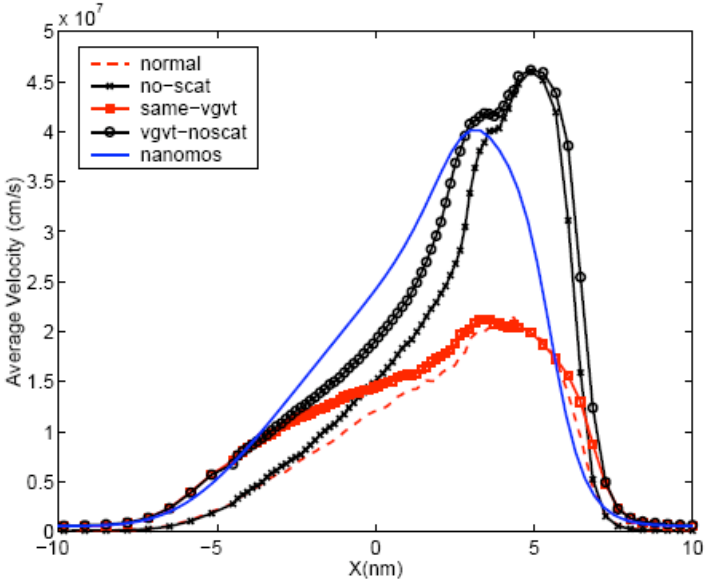
Comparison with nanoMOS

t_{Si} = 4nm, V_{ds} = 0.50V, Vg = 0.50V



t_{Si} = 4nm, V_{ds} = 0.50V, Vg = 0.35V





t_{Si} = 4nm, V_{ds} = 0.50V, Vg = 0.20V

FinFET

3D Monte Carlo simulation

L_G = 20 nm

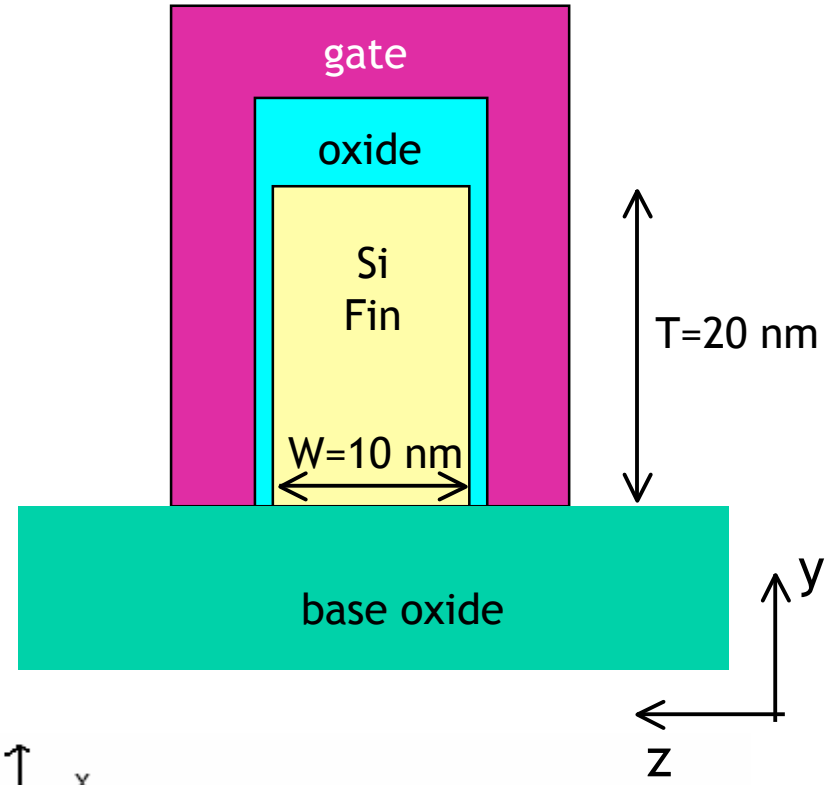
T_{Si} = 20 nm

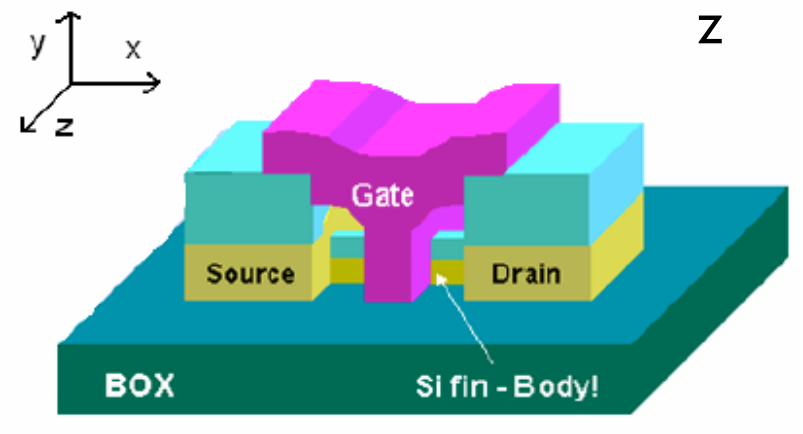
W = 10 nm

t_{ox} = 1 nm

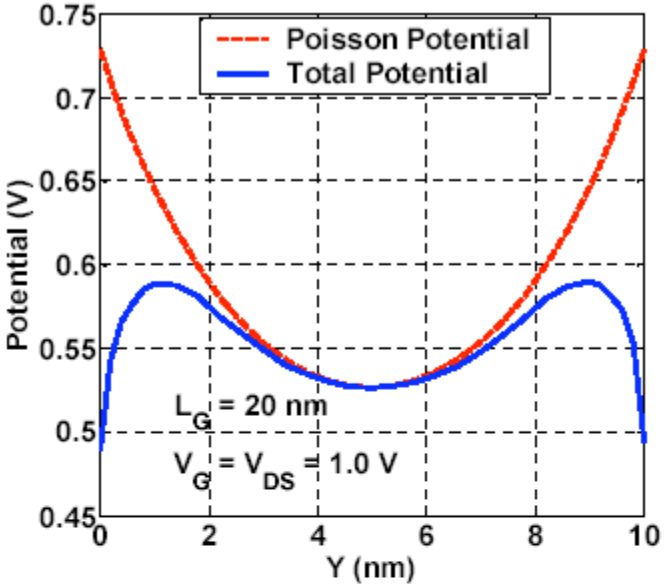
$N_D = 10^{20}\ cm^{-3}$

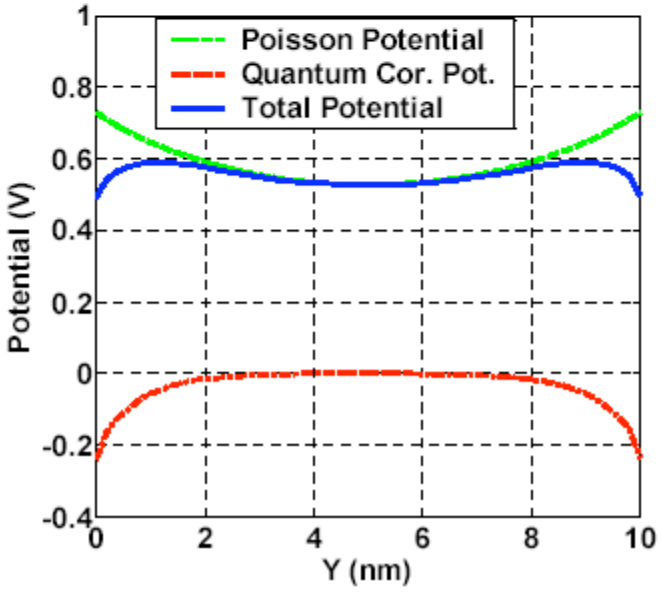
$N_A = 10^{16}\ cm^{-3}$





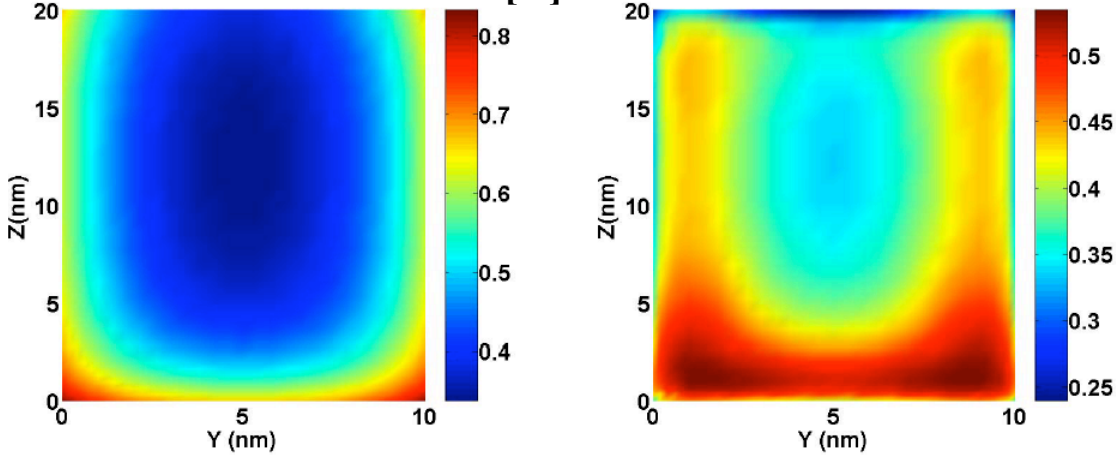
FinFET simulation





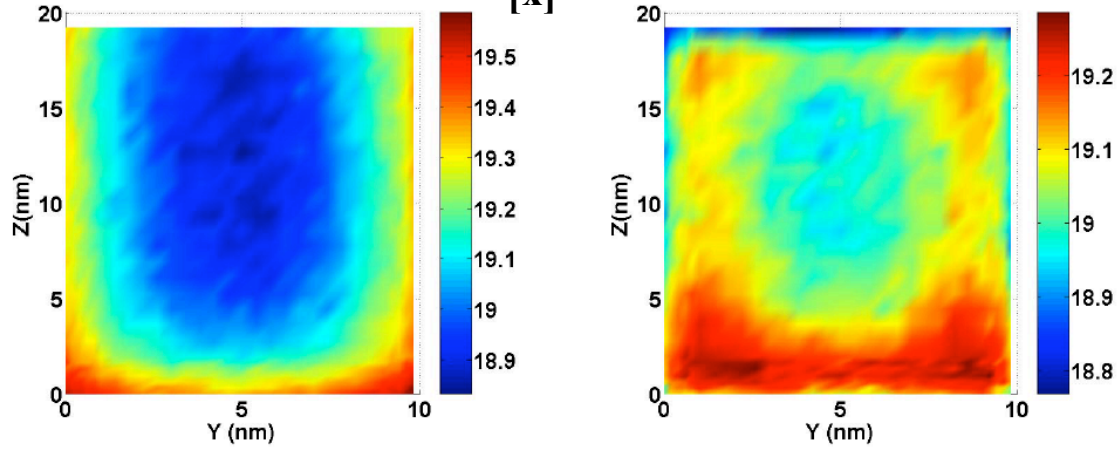
Potential

[V]



Charge density $n=10^x$

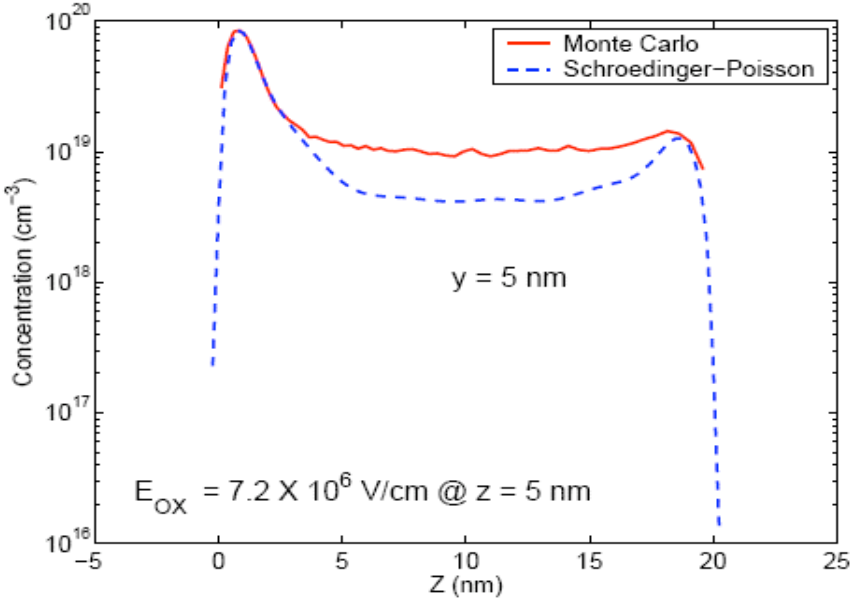
[x]

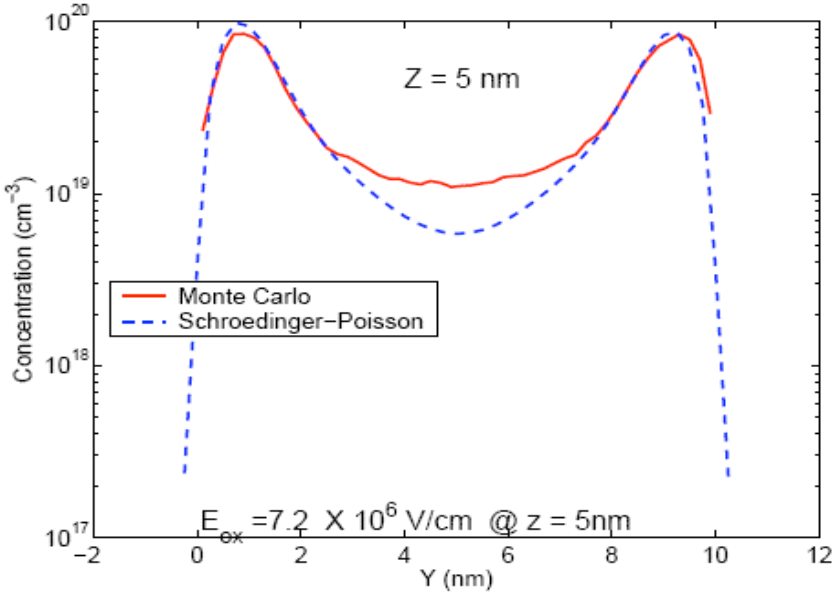


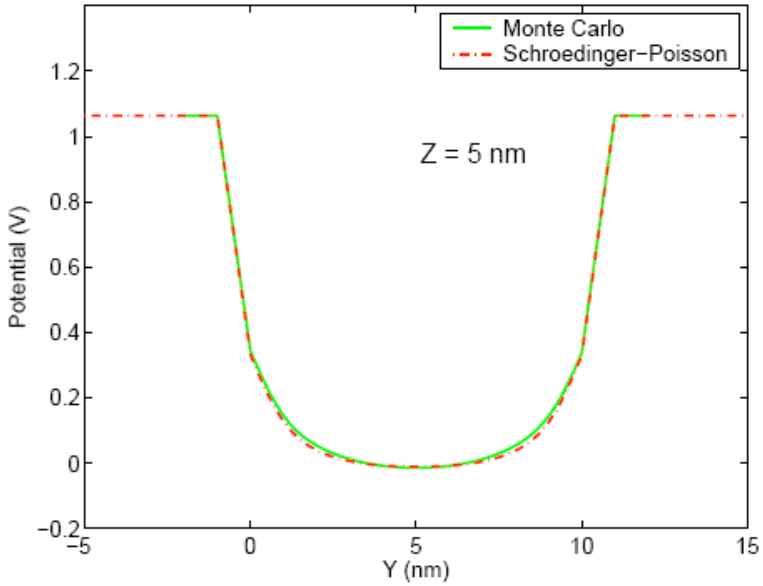
Classical

Quantum corrected

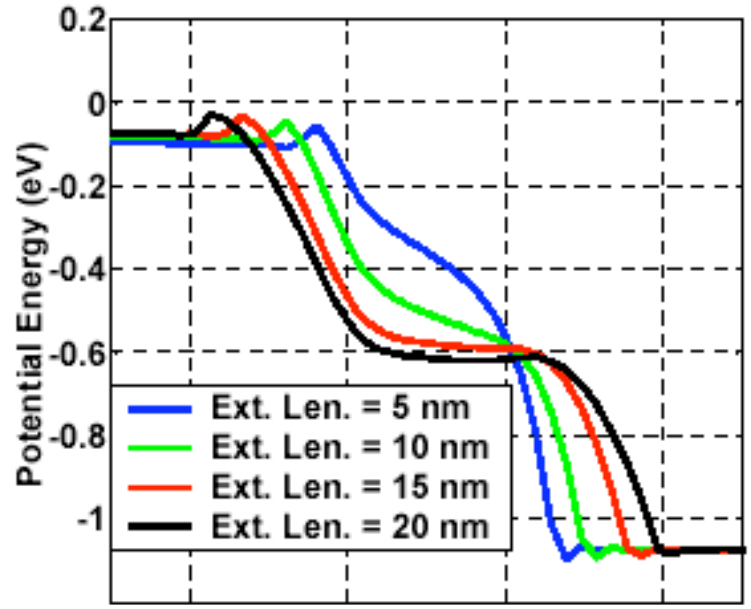
Comparisons with Schrödinger Solver

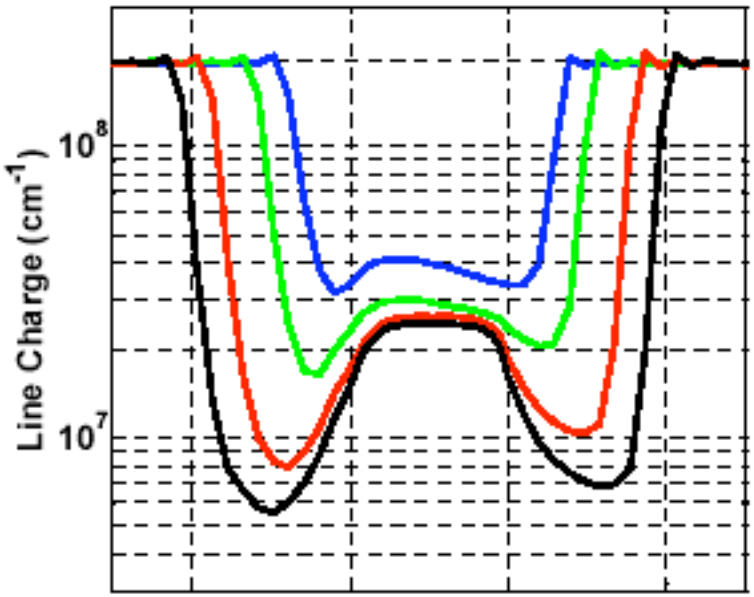


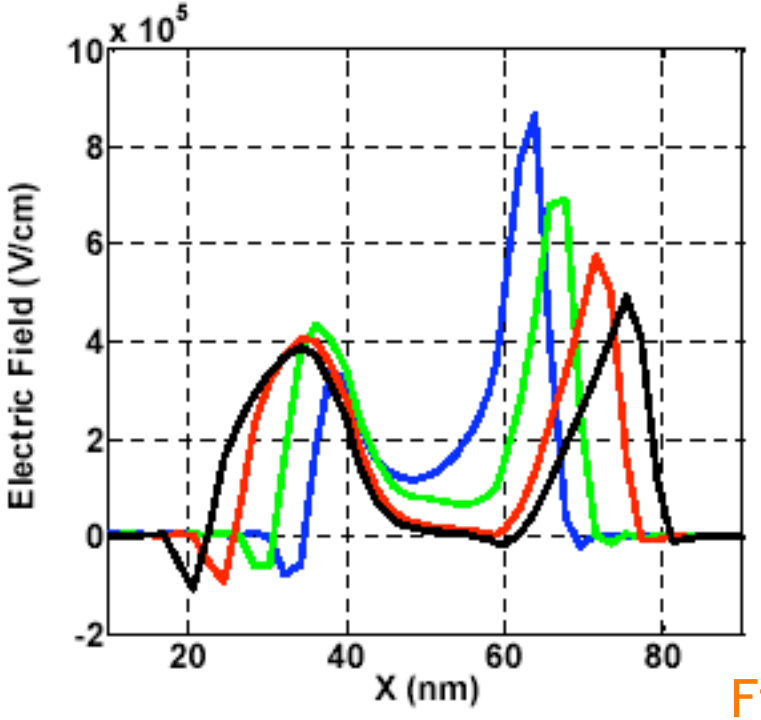


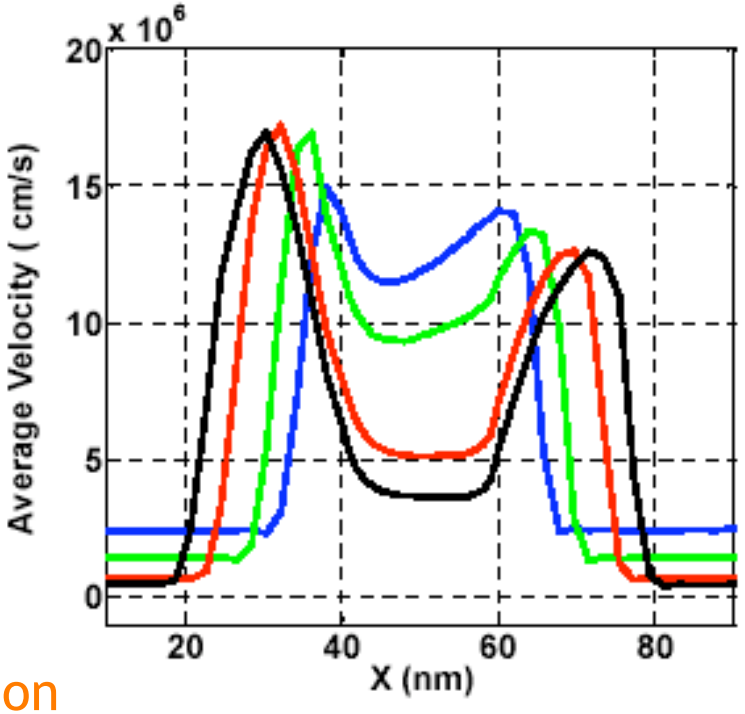


FinFET simulation



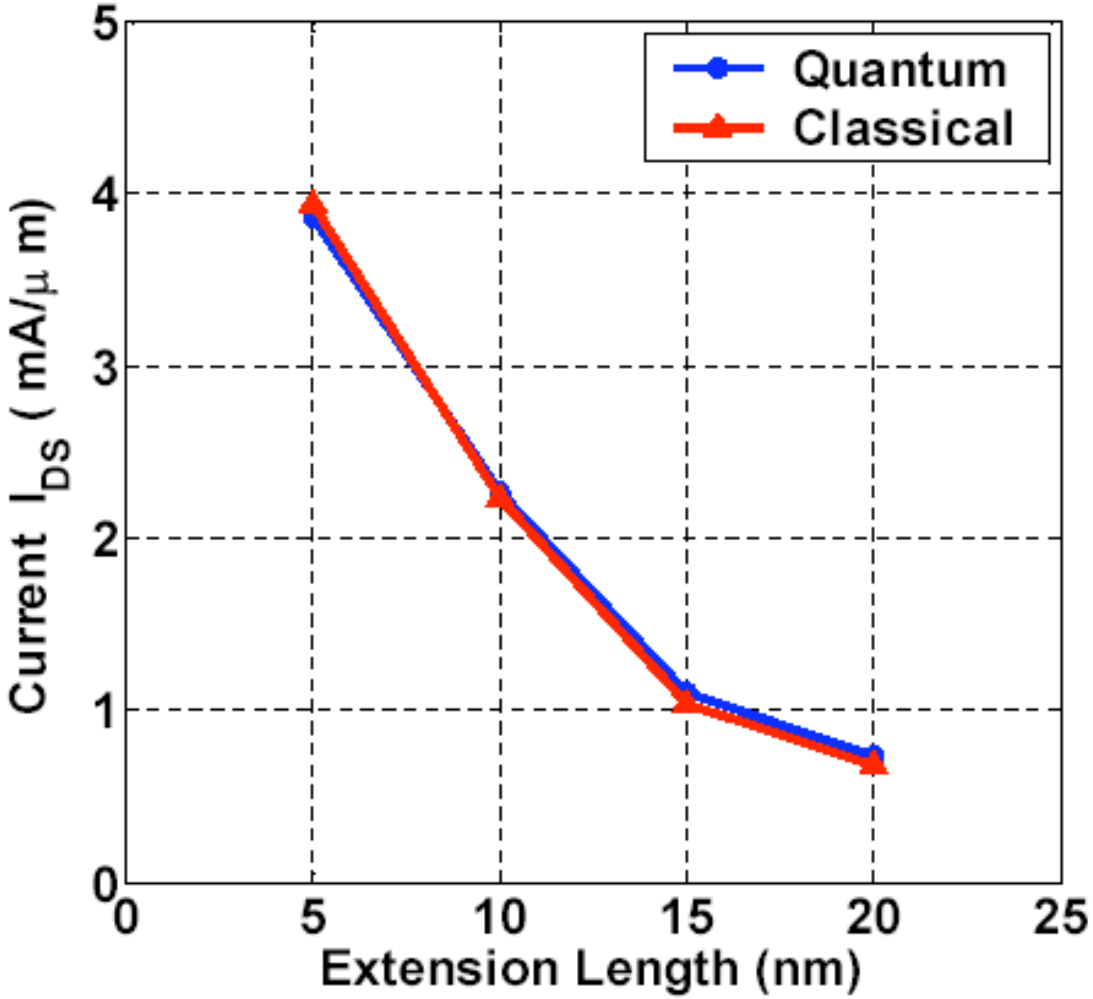






FinFET simulation

FinFET simulation



Transverse Temperature improvements [8]

- In a structure lacking a substrate reference (e.g. double gate MOSFET) the definition of transverse temperature may be problematic.
- One may replace $k_B T$ with

$$\bar{U}_{yy} = \overline{\left\langle v_y \mathrm{h}\, k_y \right\rangle}$$

the average value of the stress tensor along each transverse direction, to account for the variation of the electron temperature along the longitudinal direction.

Conclusions

- Quantum corrections can be used to extend the validity of Monte Carlo device simulation to the 10-nm regime
- Wigner-based corrections
  - accurate; momentum-dependence is interesting
  - somewhat impractical due to noise in Monte Carlo
- Bohm correction
  - improves noise with respect to Wigner correction
- Effective potential
  - simple and fast
  - accurate for small heterojunctions
- Schrödinger-based correction
  - accurate and efficient with no fitting parameters
  - requires eigenvalue solver but not too expensive in MC
  - possibly best choice for size quantization effects

Some references

Complete Subband treatment of MOS channel transport:

1. M.V. Fischetti and S.F. Laux, "Monte Carlo study of electron transport in silicon inversion layers", Physical Review B, vol. 48, pp. 2244-2274, 1993.
2. S. Yamakawa, H. Ueno, K. Taniguchi, C. Hamaguchi, K. Miyatsuji, K. Masaki and U. Ravaioli, "Study of interface roughness dependence of electron mobility in Si inversionlayers using the Monte Carlo method", Journal of Applied Physics, vol. 79, pp 991-916, 1996.

Quantum corrections:

3. D.K. Ferry, R. Akis and D. Vasileska, "Quantum effects in MOSFETs: Use of an effective potential in 3D Monte Carlo simulation of ultra-short channel devices", IEDM Tech. Digest, pp. 287-290, 2000.
4. S. Ahmed, C. Ringhofer and D. Vasileska, "Parameter-Free Effective Potential Method for Use in Particle-Based Device Simulations", IEEE Transactions on Nanotechnology, vol. 4, pp. 465-471, 2005.
5. H. Tsuchiya and U. Ravaioli, "Particle Monte Carlo simulation of quantum phenomena in semiconductor nanostructures", Journal of Applied Physics, vol. 89, pp. 4023-4029, 2001.
6. T.-W. Tang and B. Wu, "Quantum correction for the Monte Carlo simulation via the effective conduction-band edge equation", Semiconductor Science and Technology vol. 19, pp. 54-60, 2004.
7. B. Winstead and U. Ravaioli, "A quantum correction based on Schrödinger equation applied to Monte Carlo device simulation", IEEE Transactions on Electron Devices, vol. 50, pp. 440-446, 2003.
8. B. Wu and T.-W. Tang, "The effective conduction-band edge method of quantum correction to the Monte Carlo device simulation", 10th International Workshop on Computational Electronics, Purdue University, October 24-27, 2004.
9. K. Kalna and A. Asenov, "Quantum corrections in the Monte Carlo simulation of scaled PHEMTs with multiple delta doping, Journal of Computational Electronics, vol. 1, pp. 257-261, 2002.

Questions and Answers