2D Analysis of Source-to-Drain Tunneling in Decananometer MOSFETs with the Density-Gradient Model

Andreas Schenk* and Andreas Wettstein**

* Swiss Federal Institute of Technology, Integrated Systems Laboratory
Gloriastr. 35, CH-8092 Zürich, Switzerland, schenk@iis.ee.ethz.ch
** ISE Integrated Systems Engineering AG, Balgriststr. 102, CH-8008 Zürich,
Switzerland, wettstae@ise.ch

ABSTRACT

The density gradient method is able to reproduce the quantum-mechanical charge density in CMOS devices. Its ability to describe gate tunneling currents is still a matter of dispute. This paper presents the first 2-dimensional application of the density gradient model to decananometer MOSFETs. By shrinking the effective channel length to zero it is found that the degradation of the sub-threshold swing due to source-to-drain tunneling is weak and nearly independent of the channel length. It is shown that the presence of the abrupt oxide potential barrier pins the height of the source-drain barrier and limits tunneling of the confined electrons in the channel. As a result, thermionic emission determines the off-state current at 300 K even for vanishing channel length. It is concluded that 1D calculations of source-to-drain tunneling are inadequate, since they neglect the dominant influence of the Si-SiO_2 potential barrier on the transport in channel direction.

Keywords: density gradient method, decananometer MOSFETs, deep sub-micron devices, source-to-drain tunneling, TCAD

1 INTRODUCTION

It is often argued that source-to-drain tunneling represents a fundamental limitation to CMOS scaling. Kawaura et al [1] measured the sub-threshold current of a MOSFET with 8 nm gate length. Its temperature-independence below 77 K was taken as indication for direct tunneling. One-dimensional calculations of the source-to-drain tunneling current were performed by several groups [1] - [3] with the evident result of an exponential dependence of the off-state current on barrier width and height. The question arises to which extent the strong quantum confinement perpendicular to the Si-SiO_2 interface will modify those results. We show by means of self-consistent two-dimensional device simulation with the density gradient model that source-to-drain tunneling of 2D-like channel electrons differs qualitatively from the naive 1D picture and that the off-state leakage at 300 K will be dominated by thermionic emission even in the limit of zero gate length.

2 DENSITY GRADIENT MODEL

The density gradient model used here adds a quantum potential Λ of the form

$$\Lambda = -\frac{\gamma\hbar^2}{12m}\left\{\nabla^2(\beta E_{\mathrm{F},n} - \beta\bar{\Phi}) + \frac{1}{2}\left[\nabla(\beta E_{\mathrm{F},n} - \beta\bar{\Phi})\right]^2\right\}$$

to the classical potential in the density formula, where $\bar{\Phi}$ is the smoothed potential $\bar{\Phi} = E_c + \Phi_m + \Lambda$, containing the band edge E_c (including electrostatic potential), a mass driving term Φ_m (resulting from DOS discontinuities), and the quantum potential Λ itself. Extensive comparisons to the more accurate Schrödinger/Poisson method have shown that the fit parameter γ does not depend on oxide thickness, channel doping, or device temperature [4]. The quantum-mechanical (QM) charge density in MOS devices is well reproduced if $\gamma = 3.6$ in silicon regions and $\gamma \approx 1$ in oxide regions.

The main effect of the quantum potential Λ is to smooth out rapid changes of the potential on a length scale of the thermal de Broglie wave length. Gate tunneling currents are the result of a strongly reduced oxide potential barrier. Fig. 1 shows the smoothed potential

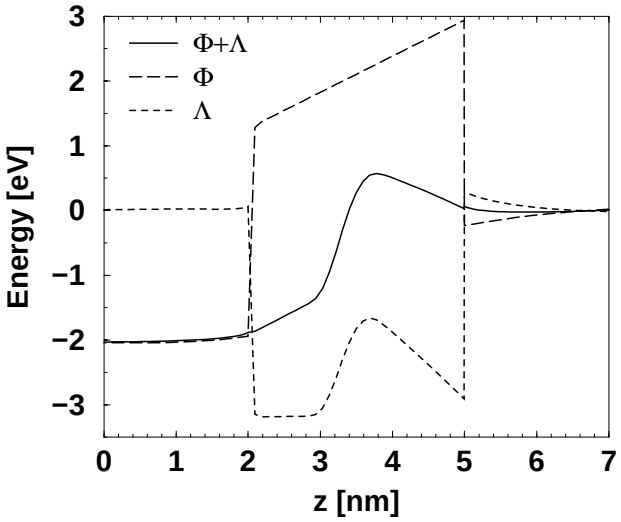

Figure 1: Smoothed potential (solid), quantum potential (dotted), and conduction band (dashed) of a MOS structure with $t_{ox} = 3$ nm and $V_G = 2$ V.

across a MOS structure with 3 nm oxide thickness and 2 V gate voltage. Note that on the silicon side of the Si-SiO_2 interface ($z \geq 5$ nm) the smoothed potential is raised compared to the classical conduction band energy. This effect is essential for the source-to-drain tunneling (transport parallel to the interface).

The density gradient and similar models have been used in the literature to simulate tunneling currents. For example, in Ref. [5] the Quantum Hydrodynamical model was used. Ref. [6] uses an extended Density-Gradient model specifically designed to handle tunneling. This variant of the model distinguishes forward- and backward-tunneling particles; unfortunately, this makes the model intrinsically one-dimensional, and a generalization to higher dimension is not known. We use the density-gradient model in it's simplest form, where the electron drift-diffusion current density is calculated according to

$$\mathbf{j}_n = \mu_n \nabla n k_\mathrm{B} T + \mu_n n \nabla (E_\mathrm{c} + \Phi_m + \Lambda) \ .$$

This is equivalent to

$$\mathbf{j}_n = \mu_n n \nabla E_{\mathrm{F},n} \ ,$$

since the density reads (in case of Boltzmann statistics)

$$n = N_\mathrm{c} \exp\left[\beta(E_{\mathrm{F},n} - E_\mathrm{c} - \Phi_m - \Lambda)\right] = n_\mathrm{cl} \exp\left[-\beta\Lambda\right] \ .$$

As the starting point of the quantum drift-diffusion theory is the equilibrium density matrix, the proper generalization of the quantum potential to a situation far from equilibrium is not obvious and in principle questionable. Fig. 2 shows the current across a MOS structure with 2 nm oxide thickness and doping levels of $10^{20}\,\mathrm{cm}^{-3}$ in the poly and $10^{18}\,\mathrm{cm}^{-3}$ in the p-substrate, respectively. In addition to the quantum potential parameters $\gamma = 3.6$ for silicon and $\gamma = 1$ for oxide, a mobility $\mu_{n,\mathrm{ox}} = 0.05\,\mathrm{cm}^2/\mathrm{Vs}$ in the oxide barrier region was used (note, that a drift-diffusion current across an insulating barrier requires to describe this region by a wide-bandgap semiconductor material with parameters of the insulator). Compared to the physically correct current resulting both from a semi-analytical transmission-coefficient model and from a more accurate Bardeen-Schrödinger simulation, the density gradient approach reproduces the current *only* in the vicinity of $V_\mathrm{g} = 0$. Whereas at larger positive bias (tunneling from "channel" to the poly gate) the density gradient current might still be acceptable for TCAD applications, at strong negative bias (tunneling from the poly gate into the substrate) an unphysical minimum of the current appears around flat-band voltage. This artifact is related to the electron depletion on the substrate side and disappears, if both sides are doped symmetrically. The same artifact can be observed in the simulation of RTD's with the DG method, when the doping in the

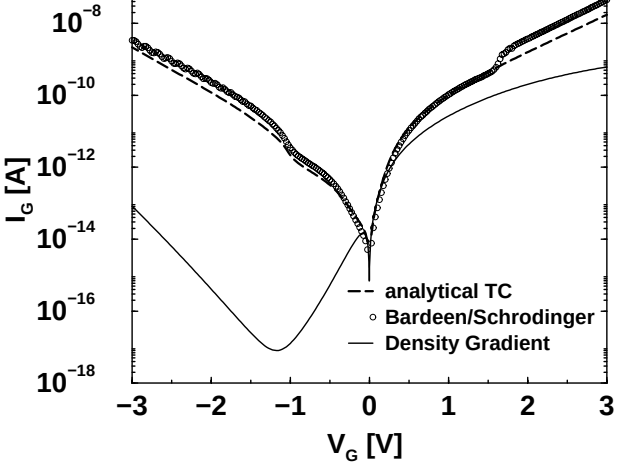

Figure 2: Simulated gate current for a MOS structure with $t_\mathrm{ox} = 2$ nm. Comparison of transmission coefficient approach (dashed), Bardeen-Schrödinger method (symbols), and density gradient model (solid).

quantum well region differs from the doping in the contact regions. The apparent RTD-like IV-curve with its typical hump (sometimes taken as indication for resonant tunneling) immediately turns into a monotonous curve when the doping in the quantum well region is made equal to the doping in the contact regions.

Apart from these problems related to non-equilibrium, it is still an open question, to which extent the density gradient model covers the tunnel effect (e.g. compared to the WKB approximation).

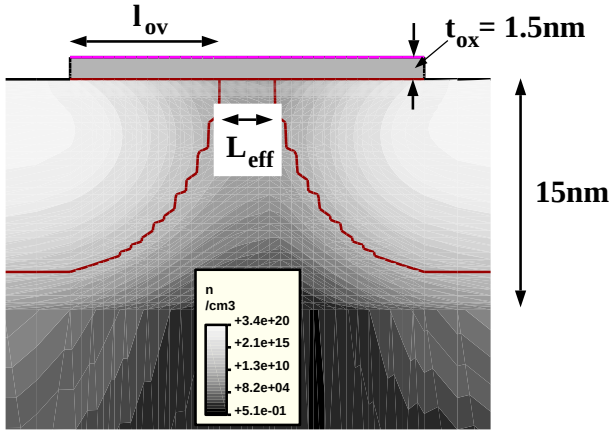

Figure 3: 2D cross section of the nMOSFET with retrograde profile and $L_\mathrm{eff} = 4$ nm. Shown is the electron density at V_G =-1 V and V_D =0.1 V

3 SIMULATION RESULTS

We performed numerical experiments varying the effective channel length $L_{\rm eff}$ (defined as the distance between the metallurgical junctions) of an nMOSFET with $t_{\rm ox} = 1.5\,$nm between 30 nm and 0 nm (Fig. 3). The width of the gate and the gate oxide thickness were kept fixed for all those simulations. The distance l_{ov} between junction at the surface and gate oxide corner (gate overlap) was first fixed to 10 nm (Figs. 4 – 7) and then decreased to 3 nm. A retrograde profile with a low-

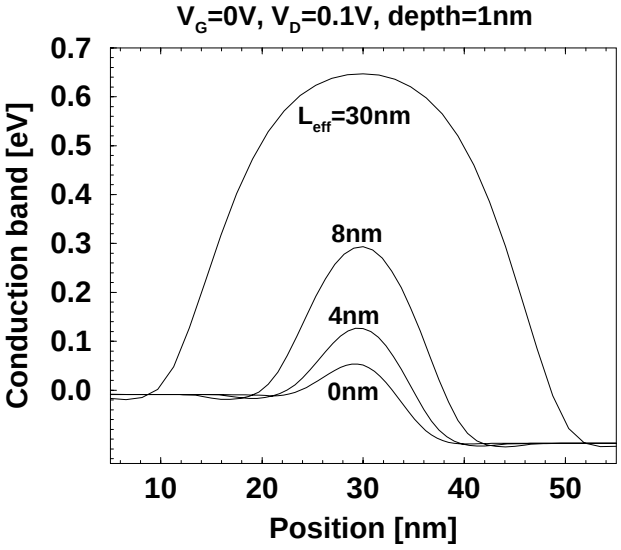

Figure 4: Cut through the potential barrier formed by the conduction band edge 1 nm below the interface for various $L_{\rm eff}$.

high transition at 15 nm depth was assumed to suppress short-channel effects. The resulting source-to-drain potential barriers at 1 nm depth are shown in Fig. 4. Due to the large Debye length at 300 K, the barrier width for electrons at the bottom of the conduction band is by roughly 8 nm larger than $L_{\rm eff}$.

The calculated transfer characteristics for the classical case (standard drift-diffusion with calibrated models) and for the QM case (including the equation for Λ with $\gamma_{\rm Si} = 3.6$ and $\gamma_{\rm ox} = 1$) are presented in Fig. 5. The classical degradation of the sub-threshold swing is due to the rapidly decreasing barrier height (increased thermal leakage). The device with $L_{\rm eff} = 30\,$nm exhibits the usual QM $V_{\rm T}$-shift of roughly 70 meV but no QM degradation of the slope. As $L_{\rm eff}$ approaches zero, the QM degradation of the slope becomes visible which is attributed to source-to-drain tunneling. However, the QM degradation is weak and does not markedly accelerate when $L_{\rm eff}$ decreases from 8 nm to 0 nm. Considering the large difference between the potential barriers, this result is remarkable from the viewpoint of standard 1D tunneling expressions. The explanation is given by

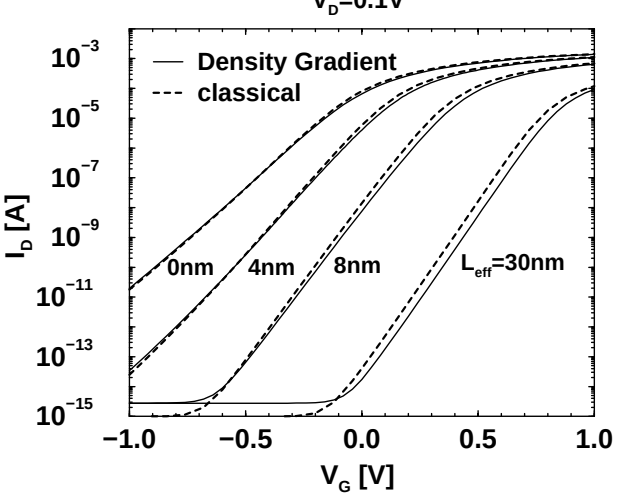

Figure 5: Simulated $I_{\rm D}$-$V_{\rm G}$ curves for the classical (dashed) and QM case (solid). The gate overlap was l_{ov} =10 nm.

Figs. 6 and 7, where for the case $L_{\rm eff} = 4\,$nm the classical barrier (l.h.s.) is compared with the corresponding smoothed potential $\bar{\Phi}$ (r.h.s.). At threshold ($V_{\rm G} = 0\,$V, Fig. 6) the smoothed potential is raised compared to the classical case because of the presence of the dominant oxide barrier (compare Fig. 1). This is the cause of the QM $V_{\rm T}$-shift. In the off-state ($V_{\rm G} = -1\,$V, Fig. 7) the height of $\bar{\Phi}$ is only slightly smaller compared to the classical counterpart, again due to the dominance of the Si-SiO_2 potential step. A strong reduction of the source-to-drain barrier (similar to the oxide barrier, Fig. 1) cannot take place. Thus, at 300 K thermionic emission determines the off-state current even in the limit $L_{\rm eff} = 0$.

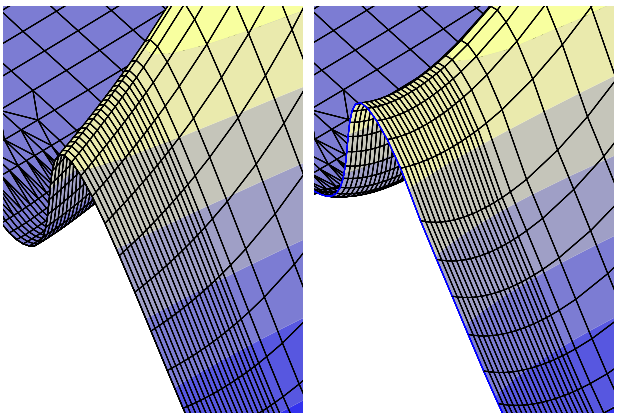

Figure 6: 2D potential barrier at threshold ($V_{\rm G} = 0\,$V). Left: classical conduction band, right: smoothed potential.

Decreasing the gate overlap from 10 nm to 3 nm, the difference between DG and classical simulations remains

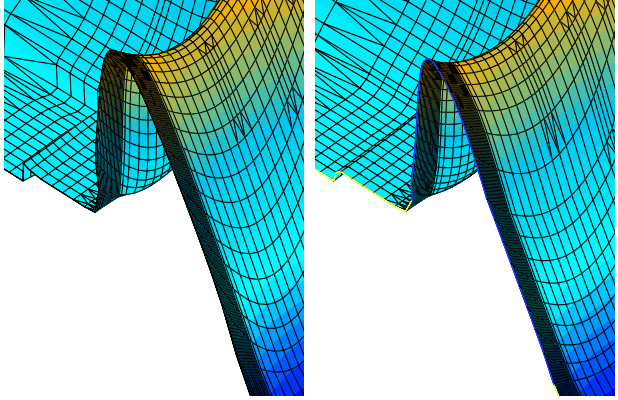

Figure 7: 2D potential barrier in the off-state (V_{G} =-1 V). Left: classical conduction band, right: smoothed potential.

unaffected (see Fig. 8), but the shorter gate degrades the sub-threshold slope and lowers the on-current for $L_{\mathrm{eff}} \leq 8$ nm (Fig. 9).

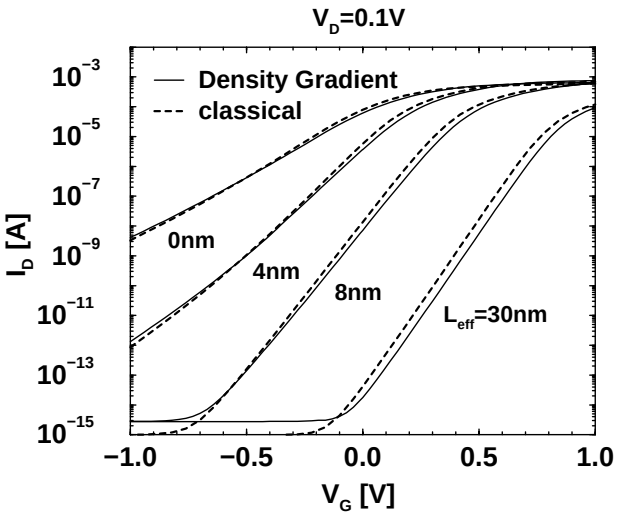

Figure 8: Simulated I_{D}-V_{G} curves for the classical (dashed) and QM case (solid). The gate overlap was l_{ov} =3 nm.

4 DISCUSSION

The physical reason for the increase of the smoothed potential in the channel is the 3-dimensional wave nature of the particles. In particular, the wave functions of particles tunneling through the source-drain barrier are extended in direction perpendicular to the silicon-oxide interface. The particles cannot be assumed to be fully located in the channel; due to their wave nature they also feel the 'repelling' oxide barrier. Therefore, the presence of the oxide increases the tunneling barrier. In other words, particles in confined states tunnel worse

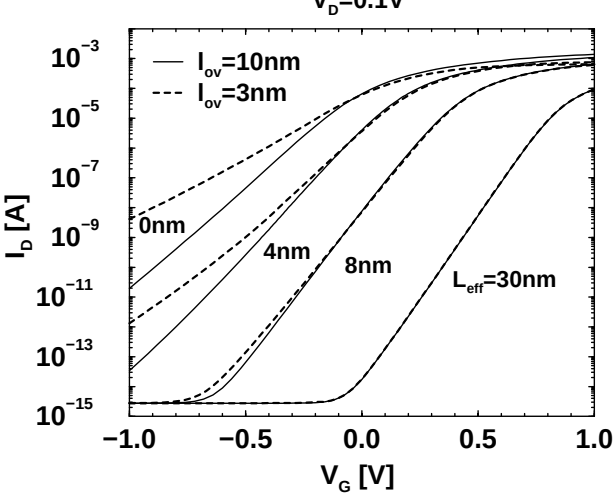

Figure 9: Density gradient simulation of I_{D}-V_{G} for different gate overlaps. l_{ov} =10 nm (solid) and l_{ov} =3 nm (dashed).

than plane waves. Part of their energy is contained in the kinetic energy for the motion perpendicular to the interface and is thus lost for tunneling.

While the ability of the density gradient method to describe tunneling quantitatively is still a matter of dispute, this physical argument shows that the observed behavior is qualitatively correct. Hence, we conclude that a 1D description of source-to-drain tunneling, neglecting the increase of the effective barrier by the gate oxide, is inadequate.

There is a close analogy to quantum dot devices, where constrictions in the 'leads', either formed geometrically or electro-statically, raise the sub-bands and define tunneling barriers that enable the functionality of these devices. Despite the possible shortcomings in describing tunneling, the density gradient model provided physical insight to the problem, and thereby has proven as an invaluable research tool.

REFERENCES

[1] H. Kawaura, T. Sakamoto, and T. Baba, Appl. Phys. Lett. 76 (25), 3810 (2000).

[2] Y. Naveh and K. K. Likharev, IEEE Electron Device Lett. 21 (5), 242 (2000).

[3] L. Chang and C. Hu, Superlattices and Microstructures 28 (5/6), 351 (2000).

[4] A. Wettstein, A. Schenk, and W. Fichtner, IEEE Trans. Electron Devices 48 (2), 279 (2000).

[5] C. L. Gardner, SIAM J. Appl. Math. 54 (2), 409 (1994).

[6] M. G. Ancona, Z. Yu, R. W. Dutton, P. J. V. Voorde, M. Cao, and D. Vook, IEEE IEEE Trans. Electron Devices 47 (12), 2310 (2000).