MOSFET Hot-Carrier Induced Gate Current Simulation by Self-Consistent Silicon/Oxide Monte Carlo Device Simulation

Andrea Ghetti
ST Microelectronics, Via Olivetti 2, 20041 Agrate Brianza, Italy
E-mail: andrea.ghetti@st.com

Abstract—**Hot electron transport in MOS transistors is investigated by means of coupled silicon/oxide Monte Carlo (MC) simulation. First, a new MC simulator able to handle self-consistently different materials is developed. Then, the impact of oxide transport on the gate current (I_G) is analyzed comparing different injection models with experiments. It is shown that oxide transport plays an important role on I_G when the gate voltage is below the drain voltage ($V_{GS} < V_{DS}$). In this condition, coupled silicon/oxide (SI+OX) simulation is important to quantitatively assess I_G. It is also shown that oxide scattering in the image force potential well does not significantly reduce I_G. Furthermore, we propose a new injection model that empirically accounts for oxide scattering and that provides the same I_G of the SI+OX model, but with the simulation of the silicon channel only, thus enabling a significant reduction of simulation time.**

I. Introduction

Although the continuous reduction of physical dimensions and applied voltage will make MOS transistors work in the direct tunnel regime [1], there is still a large class of MOS devices with oxide thickness (t_{ox}) larger than 3nm, e.g. non volatile memories, in which the gate current (I_G) is due to hot electrons. These electrons are mostly emitted over the energy barrier (ϵ_B) at the silicon/oxide interface, and travel across the oxide layer before reaching the gate electrode. In these conditions oxide transport can affect I_G. Nevertheless, its effect has not been thoroughly investigated yet.

In most cases, I_G Monte Carlo (MC) simulations adopt the ballistic approach (BAL): only transport in the silicon channel is taken into account, and I_G is evaluated by weighting each electron hitting the interface with a local transmission probability (P_T), computed assuming carriers to be ballistic [2], [3], i.e. neglecting oxide transport (Fig. 1).

Recently, coupled oxide/silicon transport has been included in the MC model [4]–[6]. However, in Ref. [4] silicon and oxide transport were simulated in sequence (non self-consistently), and only in MOS capacitors (1D case). References [5], [6] did not pay particular attention to the effect of the 2D field profile present in MOS transistors, and report opposite results about the effect of oxide scattering in the image force potential well. In [5] it is suggested that oxide scattering in the image force potential well reduces I_G of orders of magnitude, while in [6] it is reported that I_G is reduced at most by a factor of two. Therefore, a clarification on this issue is needed.

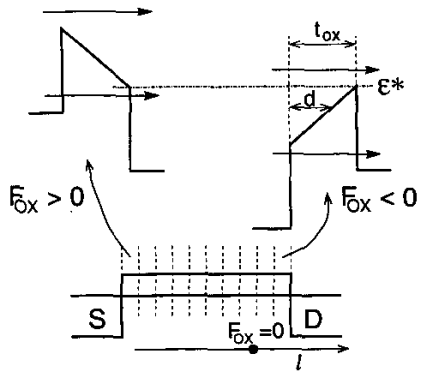

Fig. 1. Schematic illustration of the ballistic injection model.

In this paper, we report an in depth investigation of the effect of oxide transport on MOSFET I_G for realistic 2D field profiles, and analyze different injection models. By comparing measurements and simulations it is shown that oxide transport impacts appreciably hot electron I_G when the oxide field at the drain end of the channel is repulsive ($F_{ox} < 0$). We found that self-consistent silicon/oxide MC simulation (SI+OX) well reproduces experimental data. On the contrary, the ballistic model is accurate enough only when $F_{ox} > 0$. Moreover, a new injection model is proposed. This new model empirically accounts for oxide scattering and provides the same I_G of SI+OX model but with a significant reduction of simulation time. Furthermore, it is reported that oxide scattering in the image force potential well does not significantly reduce I_G.

This paper is arranged as follows. Section II describes the investigated and adopted physical models, whereas results are reported and discussed in Sec. III. Section IV presents the new injection model. Section V addresses the issue of scattering in the image force potential well. Finally, Sec. VI draws some conclusions.

II. Monte Carlo Model

Our Monte Carlo program (named FURBO) features the anisotropic (full) band structure of silicon [7] computed with the nonlocal pseudopotential method [8], and the most important scattering mechanisms, namely phonon, impact ionization (II), ionized impurity and surface scattering, as well as electron-electron and plasmon-electron scattering. We

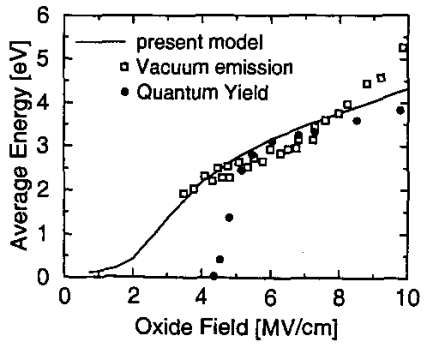

Fig. 2. Calculated average energy of emitted electrons (line) in comparison with experimental data from [12].

adopted the impact ionization scattering rate from full band calculation of Ref. [9], while phonon scattering was adjusted to reproduce a large number of experimental data sensitive to both the low and high energy part of the electron distribution such as drift velocity, universal mobility curve, impact ionization coefficient, quantum yield, homogeneous injection probability, gate current when oxide transport is not important (see Sec. III). The adopted rates [10] are in good agreement with the most recent determinations [6], [11].

For oxide simulation, a parabolic band with $m_{ox} = 0.5m_0$ is assumed for electrons. Polar optical phonon scattering (63- and 153-meV modes) has been implemented as in [12], while inelastic acoustic phonon and II scattering have been treated as in [13]. Scattering parameters have been chosen to reproduce the scattering rates of Ref. [13], therefore the present model provides the same energy distributions and average energy of Ref. [13], that compare favorably with experimental results (Fig. 2).

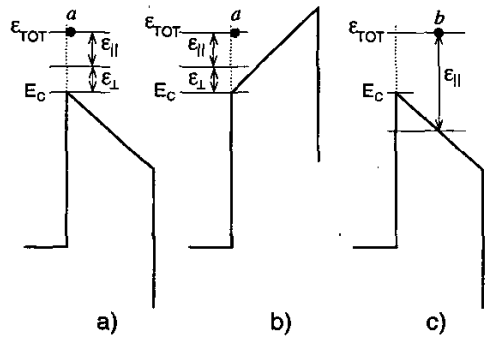

Fig. 3. Schematic illustration of the coupled silicon/oxide injection model.

In this work, we investigate different injection models. The first one is the aforementioned ballistic (BAL) model [2], [3]. The second one is the coupled self-consistent silicon/oxide MC model (SI+OX). When a particle hits the interface the new perpendicular momentum ($k_\perp$) in the oxide is computed conserving total energy (ϵ_{TOT}) and parallel momentum ($k_\parallel$)

$$\frac{\hbar^2 k_\perp^2}{2m_{ox}} = \epsilon_\perp = \epsilon_{TOT} - E_C - \frac{\hbar^2 k_\parallel^2}{2m_{ox}} \qquad (1)$$

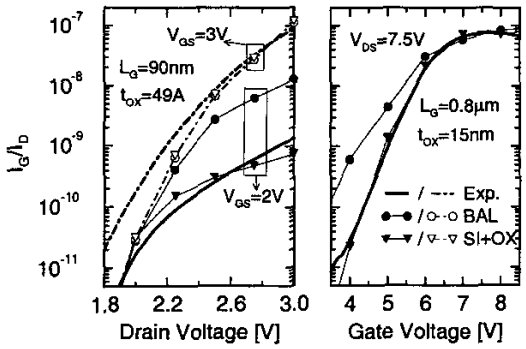

Fig. 4. Comparison of BAL (circles) and SI+OX (triangles) simulation results with experimental data (lines) on two different devices: the "Well Tempered" transistor [14] (left) and an high voltage transistor (right). SI+OX model reproduces satisfactorily the experiments. BAL model largely overestimates I_G when $V_{GS} < V_{DS}$, while agrees with SI+OX for $V_{GS} > V_{DS}$.

where E_C is the oxide conduction band edge. If $k_\perp$ is real, the particle is put just inside the oxide (point a in Fig. 3.a,b), otherwise P_T is evaluated, a new particle is created in the position corresponding to $k_\perp = 0$ (point b in Fig. 3.c) with statistical weight $w_{ox} = w_{si}P_T$, and the particle in the silicon is reflected back with w_{si} reduced by w_{ox}. Moving in the oxide the new particle can reach the gate/oxide interface (in this case it is removed from the simulation and I_G is updated), or be scattered back in the silicon.

The third model is an original model. It explicitly considers only silicon, but, differently from BAL, it also includes an analytical oxide scattering correction (ASC). It will be discussed in detail in Sec. IV.

III. Results and Discussion

This Section shows simulation results that have been obtained assuming a trapezoidal (TRPZ) barrier (as in Fig. 3). Effect of image force barrier lowering will be discussed later in Sec. V.

Figure 4 compares BAL and SI+OX I_G simulations to experimental data for thin (left) and thick (right) oxide devices. In both cases, SI+OX simulation well reproduces experiments for any bias condition. In particular, SI+OX is able to reproduce the change of slope of I_G for $V_{GS} = V_{DS}$. In contrast, BAL model provides an higher I_G for $V_{GS} < V_{DS}$ ($F_{ox} < 0$), while it coincides with SI+OX for $V_{GS} \geq V_{DS}$. This observation suggests the importance of oxide transport when the oxide field (F_{ox}) at the end of the channel pushes the electrons back in the substrate ($F_{ox} < 0$).

In order to understand why BAL model overestimates I_G when $F_{ox} < 0$, even if it properly accounts for the increased energy barrier due to the repulsive field, Fig. 5 shows the ratio (R_J) of the flux of electrons reaching the gate electrode (J_{OUT}) and the flux of electrons entering the oxide from the channel (J_{IN}) with total energy (ϵ_{TOT}) higher than the energy barrier (ϵ^* in Fig. 1). This represents the ratio between particles that actually reach the gate and those that would have reached it according to the BAL model ($R_J = J_{OUT}(\epsilon_{TOT} >$

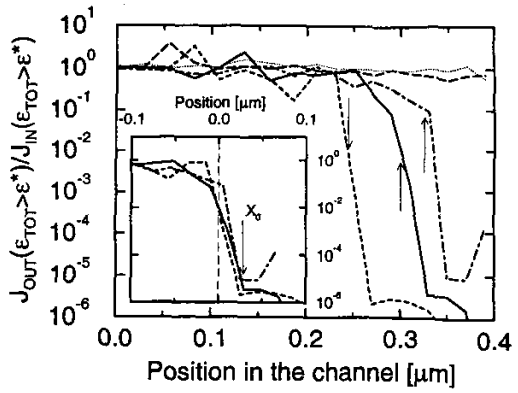

Fig. 5. Ratio of the OUT and IN fluxes with $\epsilon_{TOT} > \epsilon^*$ for the high voltage device of Fig. 4.right at $V_{DS} = 7.5V$ and different V_{GS}: 2V (dashed line), 4V (solid line), 5V (dot-dashed line), 8V (long dashed line), 10V (dotted line). Arrows mark the position where $F_{ox} = 0$: $F_{ox} > 0$ at the left, $F_{ox} < 0$ at the right. *Inset:* same curves for $V_{GS} < V_{DS}$, but with x coordinate relative to the field inversion position ($F_{ox} = 0$ for $x = 0$).

$\epsilon^*)/J_{IN}(\epsilon_{TOT} > \epsilon^*)$). As it can be seen, whenever $F_{ox} > 0$, R_J is ≈ 1 (apart from statistical noise), i.e. most particles with $\epsilon_{TOT} > \epsilon^*$ reach the gate. In contrast, when $F_{ox} < 0$, only a small fraction of the particles with $\epsilon_{TOT} > \epsilon^*$ reaches the gate, while most of them is back-scattered. Notice the smooth transition between the two regimes (around $x = 0$ in inset of Fig. 5), indicating that a non negligible number of particles injected into the oxide with $\epsilon_{TOT} > \epsilon^*$ where F_{ox} is still slightly positive is pushed laterally where $F_{ox} < 0$ and back-scatters towards the substrate.

Interestingly enough, $R_J(x)$ features the same position dependence with respect to the point where $F_{ox} = 0$, regardless of the bias conditions, as shown in the inset of Fig. 5.

IV. New Injection Model

The new injection model (ASC) tries to mimic the above mentioned behavior simply by correcting P_T by R_J if $F_{ox} < 0$ somewhere along the channel. To this purpose R_J is approximated with the following expressions: $1 - (1 - exp(-t_{ox}/\lambda))exp(l/\gamma)$ if $l < 0$, $exp(-d/\lambda)$ if $l > X_0$, while an exponential interpolation between the extreme values of the previous two expressions is used for $0 < l < X_0$, where l is the lateral distance from the point where $F_{ox} = 0$ (see Fig. 1), X_0 is defined in the inset of Fig. 5, d is the path length in the oxide conduction band (energy dependent). λ, γ and X_0 are t_{ox}-dependent fitting parameters chosen to match R_J given by SI+OX simulation.

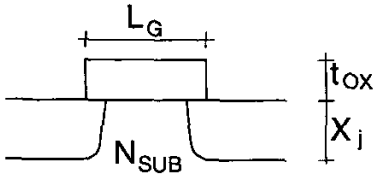

Fig. 6. Schematic structure of the test device.

$L_G[\mu m]$	0.25	0.5	0.75
$t_{OX}[nm]$	4	8	12
$X_j[nm]$	100	200	400
$N_{SUB}[cm^{-3}]$	5×10^{17}	2.5×10^{17}	1.5×10^{17}
$V_{DS}[V]$	2.5	5	7.5

Table I. Main parameters of test devices based on the ITRS.

In order to find general expressions for λ, γ and X_0, we simulated three case-study MOS transistors (Fig. 6), whose main characteristics and applied V_{DS} are reported in Table I, and fitted SI+OX results with ASC simulation. The best fit is shown in Fig. 7, and was achieved with the following values: $\lambda = 0.1t_{ox}+3$, $\gamma = Max(12.5t_{ox}-500, 0)$, and $X_0 = 2.5t_{ox}$ (all quantities in Å).

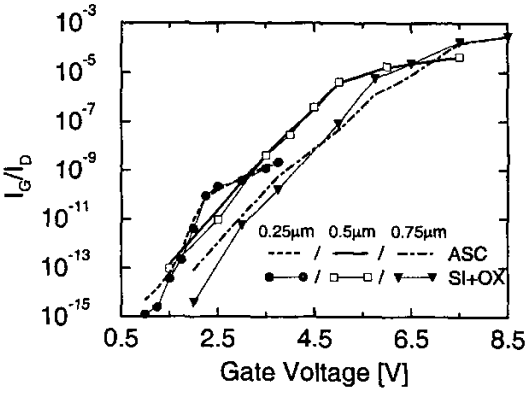

Fig. 7. Comparison of the simulated I_G provided by the SI+OX and ASC models. V_{DS} as reported in Table I. Parameters of ASC model were chosen to fit SI+OX results and are reported in the text.

Finally, the so-calibrated ASC model has been applied to the simulation of the same devices of Fig. 4 (results are shown in Fig. 8). The ASC model agrees with experiments much better than BAL, and reproduces the main features of more accurate SI+OX simulation with a significant reduction of the CPU time. As matter of fact, the normalized speed (CPU seconds needed to move one particle for one picosecond) for simulations like the ones in Fig. 8.right is 0.074 for ASC against 0.27 for SI+OX (where only 1/4 of the particles are in the oxide). This is because oxide MC simulation requires a much shorter time step due to the higher scattering rate.

V. Image Force

Figure 9 shows the energy profiles considered to study the effect of oxide scattering in the image force potential well. For the BL profile, electrons experience an high scattering rate because of the high energy [5] and, if they fall below the top of the barrier, they are reflected back in the silicon. In contrast, with the SIMPLE profile electrons undergo a reduced scattering rate (because of the smaller energy) and $F_{ox} = 0$ makes it easier to reach the gate electrode anyway [5].

Figure 10 compares simulations and measurements of injection probability (P_{inj}) for the simple experimental con-

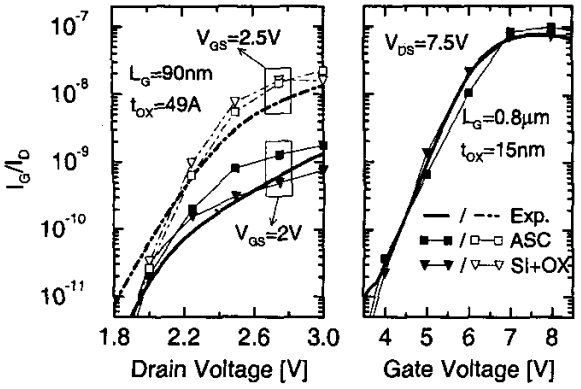

Fig. 8. Comparison of ASC (squares) and SI+OX (triangles) simulations with experimental data (lines) on the same devices of Fig. 4. The ASC model reproduces the main features of SI+OX.

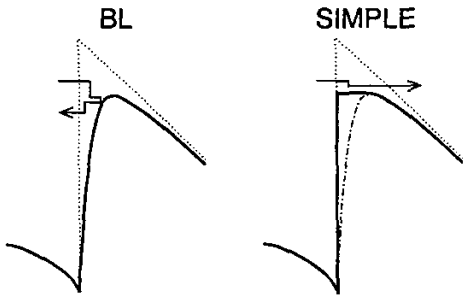

Fig. 9. Left: full image force correction (BL). Right: simplified correction (SIMPLE).

dition of homogeneous substrate hot electron injection [15]. BAL and SI+OX model with trapezoidal barrier well agree with experiments. SI+OX simulations with the image force corrections of Fig. 9 overestimate P_{inj} for $V_{TOT} = V_{SB} + 2\Phi_F > \epsilon_B/q$ because of the reduced barrier, while they slightly underestimate P_{inj} for low V_{TOT} because tunneling was not included in the simulation to single out only the effect of oxide transport. However, BL and SIMPLE profiles provide essentially the same results, demonstrating that oxide scattering in the image force potential well does not significantly change I_G [6].

To check this conclusion we repeated the same exercise of Fig. 5 computing the ratio of the OUT and IN fluxes with ϵ_{TOT} larger than the top of the barrier. This ratio turned out to be at least 0.5 [6]. If we assume that the effect of oxide scattering in the region where $F_{ox} < 0$ simply reduces I_G by $exp(-d/\lambda_\epsilon)$, where d is the path length in the oxide conduction band and λ_ϵ is some effective energy relaxation length, this result and the ones in Fig. 5 are consistent with $\lambda_\epsilon \approx 1 - 2nm$, in agreement with experimental and theoretical results of photo emission [6]. Moreover, this value of λ_ϵ is also consistent with the energy relaxation time $\tau_\epsilon \approx 20fs$ given by SI+OX simulations if λ_ϵ is estimated as $\lambda_\epsilon = v_d\,\tau_\epsilon$, where $v_d \approx 10^7 cm/s$ is the electron drift velocity in the oxide.

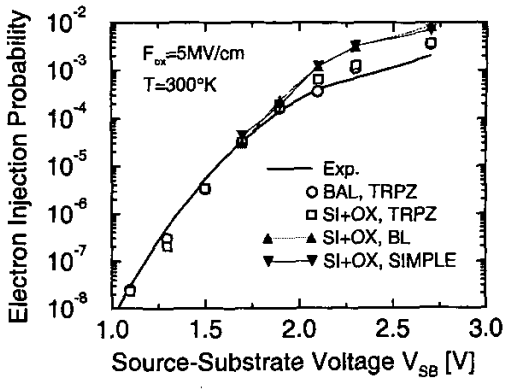

Fig. 10. Homogeneous injection probability provided by different injection models (symbols) in comparison with experimental data (line) from [15].

VI. Conclusions

In this paper we have investigated the impact of oxide transport on the gate current of MOS transistors by means of coupled self-consistent silicon/oxide Monte Carlo simulation, also comparing different injection models.

It has been shown that oxide transport is not important when $F_{ox} > 0$. In this case, the widely used ballistic assumption is accurate enough. On the contrary, when a portion of the channel experiences a negative oxide field, oxide transport plays an important role that must be taken into account for accurate evaluation of the hot carrier induced gate current. This observation holds true even for thin oxide devices as long as I_G is made of hot electrons emitted in the oxide conduction band.

In order to save the long CPU time required by oxide MC simulation, we have proposed a new injection model providing the same degree of accuracy as SI+OX, but requiring only the simulation of silicon.

In addition, it has also been shown that oxide scattering in the image force potential well does not significantly change I_G.

References

[1] E. Cassan *et al.*, *Journal of Appl. Phys.*, vol. 86, no. 7, p. 3804, 1999.
[2] A. Duncan *et al.*, *IEEE Trans. on Elec. Dev.*, vol. 45, no. 4, p. 867, 1998.
[3] A. Ghetti *et al.*, *IEEE Trans. on Elec. Dev.*, vol. 46, p. 696, Apr. 1999.
[4] P. Palestri *et al.*, in *Proc. SISPAD Conference*, p. 38, 2000.
[5] T. Ezaki *et al.*, in *IEDM Technical Digest*, p. 485, 2001.
[6] M. Fischetti *et al.*, *Journal of Applied Physics*, vol. 78, p. 1058, 1995.
[7] F. Venturi *et al.*, in *Int. Conf. on Simulation of Semi. Proc. and Dev. (SISPAD)*, (Piscataway, NJ), p. 343, 1997. IEEE.
[8] M. Fischetti *et al.*, in *Monte Carlo device simulation: full band and beyond*, (K. Hess, ed.), ch. 5, p. 123. Norwell, MA: Kluwer Accademic Publishers, 1991.
[9] J. D. Bude *et al.*, *IEEE Electron Device Letters*, vol. 16, p. 439, 1995.
[10] A. Ghetti, *Applied Physics Letters*, vol. 80, no. 11, p. 1939, 2002.
[11] M. Fischetti *et al.*, *IEEE on line TCAD Journal*, `http://www.ieee.org/products/online/journal/tcad/accepted/fischetti-feb97/`, no. 3, 1997.
[12] M. Fischetti *et al.*, *Phys. Rev. B*, vol. 31, no. 12, p. 8124, 1985.
[13] D. Arnold *et al.*, *Phys. Rev. B*, vol. 49, no. 15, p. 10278, 1994.
[14] D. Antoniadis *et al.*, `http://www-mtl.mit.edu/Well/`, 1999.
[15] B. Fischer *et al.*, *IEEE Trans. on Elec. Dev.*, vol. 44, p. 288, Feb. 1997.