

A New Gate Current Model Accounting for a Non-Maxwellian Electron Energy Distribution Function

A. Gehring, T. Grasser, H. Kosina, and S. Selberherr

Institute for Microelectronics, TU Vienna
Gusshausstrasse 27-29, Vienna, Austria
Email: Gehring@iue.tuwien.ac.at

Abstract – We report on a new formulation to describe hot electron injection through gate dielectrics. It is based on an expression which accounts for the non-Maxwellian shape of the electron energy distribution function. We use the first three even moments of the Boltzmann equation n , T_n , and β_n found by the solution of a six moments transport model to describe the shape of the distribution function. Excellent agreement with results from rigorous Monte Carlo simulations and measurements is achieved.

I. INTRODUCTION

In state-of-the-art sub-quartermicron devices with gate oxide thicknesses around or below 2 nm the proper simulation of gate oxide tunneling currents is indispensable. Approaches to describe tunneling effects usually assume a Fermi-Dirac or cold Maxwellian shape of the electron energy distribution function (EED). This assumption is clearly violated in the channel of contemporary MOS devices where the deviation from the Maxwellian shape may reach several orders of magnitude due to a pronounced high energy tail. A correct modeling of the shape of the distribution function is thus necessary to reliably predict gate currents due to hot electron injection.

II. TUNNELING MODEL

Following [1] the hot electron gate current density can be derived from the Tsu-Esaki expression as

$$J_g = q \cdot \int_0^{\infty} f(\mathcal{E}) \cdot g(\mathcal{E}) \cdot v_{\perp}(\mathcal{E}) \cdot T(\mathcal{E}) d\mathcal{E} \quad (1)$$

where $f(\mathcal{E})$ is the electron energy distribution, $g(\mathcal{E})$ the density of states, $v_{\perp}(\mathcal{E})$ the electron velocity perpendicular to the interface, and $T(\mathcal{E})$ the tunneling probability. The integration is performed starting from the conduction band edge which serves as reference energy. The velocity perpendicular to the interface can be derived from the dispersion relation using the expression $v_{\perp}(\mathcal{E}) = 1/4 \cdot \partial \mathcal{E} / \partial p$ [1]. This approach offers the possi-

bility to explicitly take the non-Maxwellian shape of the electron energy distribution function into account.

A simple model for the transmission coefficient can be derived using the WKB approximation for trapezoidal and triangular barriers:

$$T(\mathcal{E}) = \exp \left\{ -4 \frac{\sqrt{2m_{ox}}}{3\hbar q F_{ox}} \cdot \phi(\mathcal{E}) \right\} \quad (2)$$

with F_{ox} being the electric field and m_{ox} the electron mass in the oxide. The function $\phi(\mathcal{E})$ is defined as

$$\phi(\mathcal{E}) = \begin{cases} (\Phi - \mathcal{E})^{3/2} & \Phi_0 < \mathcal{E} < \Phi \\ (\Phi - \mathcal{E})^{3/2} - (\Phi_0 - \mathcal{E})^{3/2} & \mathcal{E} < \Phi_0 \end{cases} \quad (3)$$

where Φ and Φ_0 are the upper and lower barrier heights, as shown in the inset in Fig. 1. The value of Φ_0 is calculated from $\Phi_0 = \Phi - q \cdot F_{ox} \cdot t_{ox}$ where q is the electron charge and t_{ox} denotes the gate oxide thickness. The transmission coefficient for different oxide thicknesses is depicted in Fig. 1 for silicon dioxide and several oxide thicknesses.

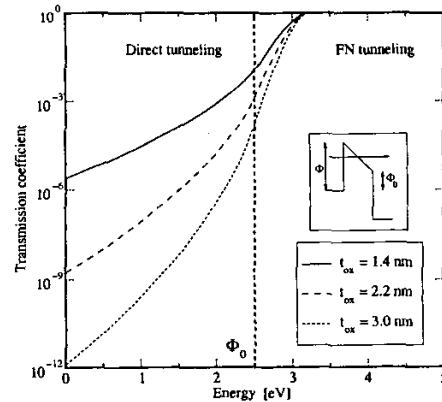


Figure 1: Transmission coefficient for different oxide thicknesses with $\Phi_0 = 2.5$ eV.

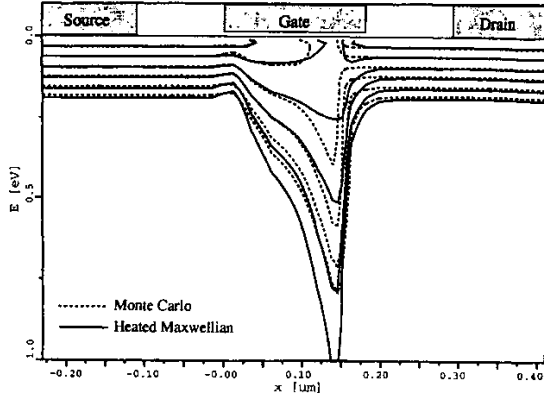


Figure 2: Heated Maxwellian EED compared to Monte Carlo results for a 180 nm MOSFET.

III. DISTRIBUTION FUNCTION

Various research deals with the problem of distribution function modeling for hot carriers in the channel region of a MOSFET [2, 3]. The problem arises from the fact that the assumption of a cold Maxwellian distribution function

$$f(\mathcal{E}) = A \cdot \exp\left(-\frac{\mathcal{E}}{k_B \cdot T_L}\right), \quad (4)$$

with T_L being the lattice temperature and A a normalization constant accounting for the Fermi energy, underestimates the high-energy tail of the electron energy distribution near the drain region. The straightforward approach is to use a heated Maxwellian distribution function

$$f(\mathcal{E}) = A \cdot \exp\left(-\frac{\mathcal{E}}{k_B \cdot T_n}\right) \quad (5)$$

where the lattice temperature T_L is simply replaced by the electron temperature T_n calculated from a suitable transport model.

We applied a Monte Carlo simulator employing analytical non-parabolic bands to check the validity of the heated Maxwellian approximation. The effect of electron-electron interaction, which increases the population of the high energy tail [2], was neglected in this study. Fig. 2 shows the contour lines of the heated Maxwellian EED in comparison to Monte Carlo results for a MOSFET device with a gate length of $L_g = 180$ nm at $V_{DS} = V_{GS} = 1$ V. It can be clearly seen that the heated Maxwellian distribution (full lines) yields only poor agreement with the Monte Carlo results (dashed lines). The heated Maxwellian distribution overestimates the high-energy tail in the channel by several orders of magnitude. Furthermore, at the drain end of the channel hot electrons mix with cold electrons supplied from the drain region, which leads to an additional population of cold electrons which cannot be reproduced by this model.

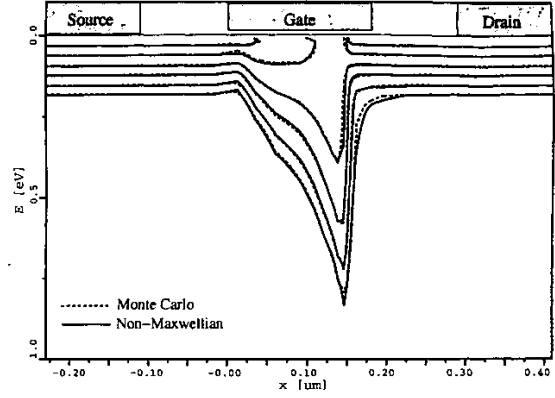


Figure 3: Non-Maxwellian EED compared to Monte Carlo results for a 180 nm MOSFET.

A correct expression for the EED has to account for both effects, namely for the slope of the high energy tail along the channel and for the emerging population of cold electrons near the drain contact. Such an expression was proposed by Sonoda *et al.*[4], and an improved model has been suggested by Grasser *et al.*[5]:

$$f(\mathcal{E}) = A \left\{ \exp\left[-\left(\frac{\mathcal{E}}{a}\right)^b\right] + c \cdot \exp\left[-\frac{\mathcal{E}}{k_B T_L}\right] \right\}. \quad (6)$$

Here the pool of cold carriers in the drain region is correctly modeled by an additional cold Maxwellian subpopulation which leads to a reduced high-energy tail. The values of a , b , and c are derived from the solution variables of a six moments transport model using the procedure described in Ref. [5]. In Fig. 3 expression (6) is compared to Monte Carlo results showing excellent agreement all along the channel.

Note that only a six moments transport model can provide information about the kurtosis of the distribution function. The electron concentration, electron temperature and kurtosis are derived from

$$n = \langle 1 \rangle \quad (7)$$

$$\frac{3k_B T_n}{2} = \langle \mathcal{E} \rangle \quad (8)$$

$$\frac{5\beta_n}{3} = \frac{\langle \mathcal{E}^2 \rangle}{\langle \mathcal{E} \rangle^2} \quad (9)$$

where the moments of the distribution function are defined as

$$\langle \Phi \rangle = \int_0^\infty \Phi f(\mathcal{E}) g(\mathcal{E}) d\mathcal{E}. \quad (10)$$

The value of the relative kurtosis in the bulk is used to locate the regions where the cold Maxwellian part of the distribution function emerges.

The value of the normalization constant A is calculated from the carrier concentration. This assures consistency of the model, in contrast to the normalization used by Hasnat *et al.* [6] which is independent of the carrier concentration and inevitably leads to erroneous results for both the carrier concentration and the electron temperature. Additionally, the population of cold carriers near the drain side of the channel can not be reproduced using Hasnat's model.

IV. RESULTS

For the evaluation of the tunneling model we apply our non-Maxwellian distribution function to the simulation of MOS transistors with varying gate lengths and oxide thicknesses. We simulated nMOS devices in on-state with $V_{GS} = 1$ V and $V_{DS} = 1$ V. Gate lengths of 350 nm, 250 nm, 180 nm and 100 nm with gate oxide thicknesses of 3.4 nm, 3.0 nm, 2.6 nm, 2.2 nm, 1.8 nm, 1.4 nm, and 1.0 nm have been assumed. Gaussian source and drain doping peaks of 10^{20} cm^{-3} with LDD extensions were used. In the following figures the results from Monte Carlo simulations will serve as reference.

Fig. 4 shows the integrand of expression (1) as a function of the electron energy for the case of a heated Maxwellian distribution and the non-Maxwellian distribution (expressions (5) and (6)). The simulated device has a gate length of 100 nm and a gate oxide thickness of 3 nm. While at low energies the difference between the non-Maxwellian distribution function and the heated Maxwellian distribution seems to be negligible, the amount of overestimation of the incremental gate current density for the heated Maxwellian distribution reaches several orders of magnitude at 1 eV and peaks when the electron energy exceeds the barrier height. This spurious effect is clearly more pronounced for points at the drain end of the channel where the electron temperature is high.

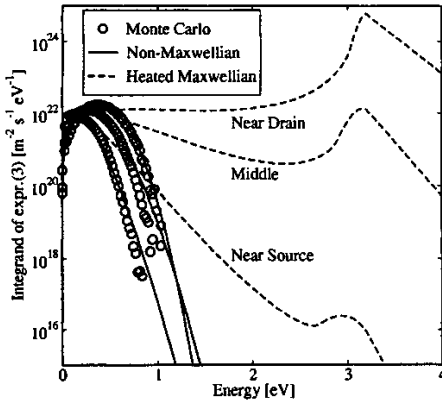


Figure 4: Integrand of (1) at different points in the channel of a 100 nm MOSFET.

The peak in the integrand for the heated Maxwellian EED results in an increased gate current density near the drain side of the gate contact. Fig. 5 shows the gate current density along the channel of a 180 nm gate length device for different oxide thicknesses. The heated Maxwellian EED heavily overestimates the gate current density near the drain contact, while the non-Maxwellian expression (6) correctly reproduces the Monte Carlo results.

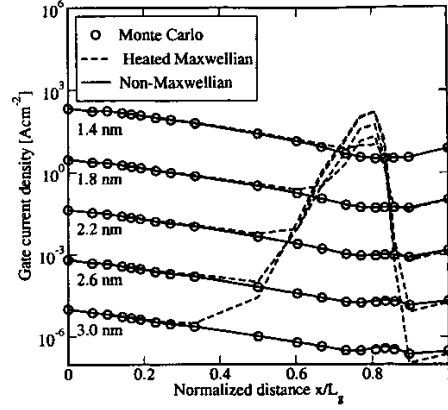


Figure 5: Gate current density along the channel of a 180 nm MOSFET with different oxide thicknesses.

In Fig. 6 the gate current density is depicted for the 100 nm and 180 nm device for an oxide thickness of 2.6 nm. While the non-Maxwellian distribution correctly reproduces the Monte Carlo results, the cold Maxwellian distribution leads to a sound underestimation reaching one order of magnitude, and the heated Maxwellian distribution predicts a much too high gate current density near the drain side of the channel. Due to the increasing electron temperature in the channel this effect is even more pronounced for smaller gate lengths.

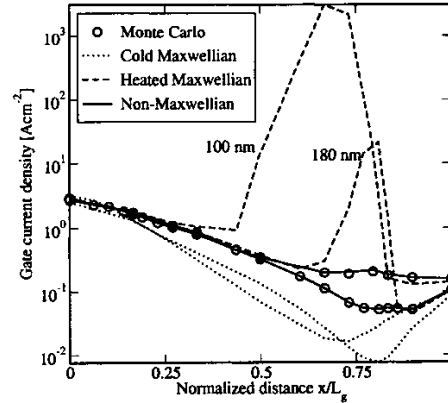


Figure 6: Gate current density along the channel of a 180 nm and a 100 nm MOSFET.

The simulation results have been compared to data reported in recent publications. For a MOSFET with zero drain-source voltage, the cold Maxwellian, heated Maxwellian, and non-Maxwellian model deliver of course the same results which are shown in Fig. 7. The measurement values were taken from Ref. [7]. The electron mass in the oxide was used as fitting parameter and an excellent fit for all oxide thicknesses was achieved with $m_{ox} = 0.65 m_0$. Note that the result is independent of the gate length since no drain-source bias was applied.

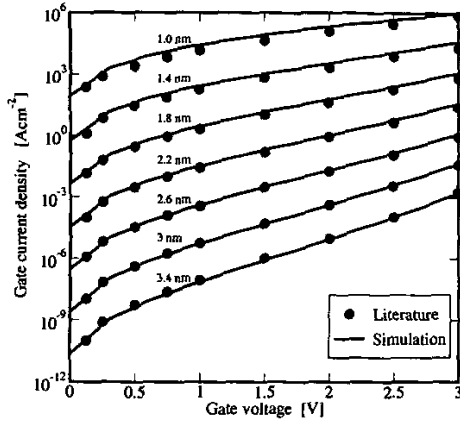


Figure 7: Gate current density for different MOS capacitors.

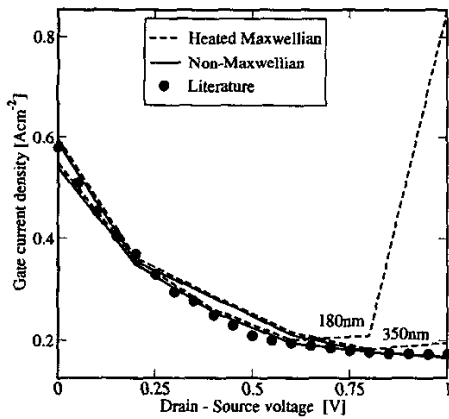


Figure 8: Gate current density for a 350 nm and a 180 nm device.

For the case of hot carriers, however, the heated Maxwellian distribution fails to reproduce even the qualitative behavior observed in measurements. Fig. 8 shows the gate current density as a function of the drain-source voltage for a 350 nm gate length, 1.8 nm oxide thickness MOSFET compared to measurement data presented in [8]. Perhaps due to differences in gate oxide thickness determination and the measurement setup, the gate oxide thickness had to be increased to give the same values for

zero drain voltage as in Fig. 7. For increasing drain voltage, the electric field in the oxide is reduced which leads to a reduced transmission coefficient and lower gate current. The heated Maxwellian distribution overestimates the gate current density especially for high bias. While this effect is not so strong for the 350 nm device, it is clearly visible for a 180 nm device and, due to the increase of the electric field, it will even be higher for increased drain voltages. The non-Maxwellian model, however, correctly reproduces the measurements and shows reasonable results for the 180 nm device.

V. CONCLUSION

We presented a model to describe the hot-electron gate tunneling current by taking the non-Maxwellian shape of the electron energy distribution function into account. Conventional models which are based on the assumption of a cold or heated Maxwellian distribution fail to reproduce hot carrier tunneling in contemporary short-channel devices. The heated Maxwellian assumption delivers a much too high gate current density due to an overestimation of the high energy tail. Our model makes use of a recently developed non-Maxwellian expression for the distribution function which is based on a six-moments transport model. Using the new expression we can accurately reproduce Monte Carlo results and measurement data of a turned-on MOSFET.

REFERENCES

- [1] C. Fiegna, F. Venturi, M. Melanotte, E. Sangiorgi, and B. Riccò, *IEEE Trans. Electron Devices* **38**, 603 (1991).
- [2] A. Abramo and C. Fiegna, *J. Appl. Phys.* **80**, 889 (1996).
- [3] D. Cassi and B. Ricco, *IEEE Trans. Electron Devices* **37**, 1514 (1990).
- [4] K.-I. Sonoda, M. Yamaji, K. Taniguchi, and C. Hamaguchi, *J. Appl. Phys.* **80**, 5444 (1996).
- [5] T. Grasser, H. Kosina, C. Heitzinger, and S. Selberherr, *J. Appl. Phys.* **91**, 3869 (2002).
- [6] K. Hasnat, C.-F. Yeap, S. Jallepalli, S. A. Hareland, W.-K. Shih, V. M. Agostinelli, A. F. Tasch, and C. M. Maziar, *IEEE Trans. Electron Devices* **44**, 129 (1997).
- [7] A. Shanware, J. P. Shiely, and H. Z. Massoud, in *Intl. Electron Devices Meeting* (1999), pp. 815–818.
- [8] S. Schwantes and W. Krautschneider, in *European Solid-State Device Research Conference* (Nuremberg, Germany, 2001), pp. 471–474.