# On the Tunneling Energy within the Full–Band Structure Approach

F. M. Bufler*† and A. Schenk*†
* Institut für Integrierte Systeme, ETH Zürich
Gloriastrasse 35, CH-8092 Zürich, Switzerland
E-mail: bufler@iis.ee.ethz.ch
† Synopsys Schweiz GmbH
Affolternstrasse 52, CH-8050 Zürich, Switzerland

***Abstract*— The aim of this paper is to generalize the well-known formula used for the computation of the Fowler-Nordheim tunneling current of cold electrons to the case of hot-electron full-band transport. The same framework of approximations is employed which was successful in the simulation of the erasing mechanism in flash memory cells. For the WKB tunneling probability, an approximation for the tunneling energy within the full-band structure approach is proposed which consists of weighting the total electron energy by the ratio between the square of the group velocity component normal to the interface to the square of the total group velocity. The difference of the gate current when using either the total or the "perpendicular" energy for tunneling is about two orders of magnitude with decreasing tendency for higher gate voltage, similar to previous results with a more elaborated model in the literature.**

## I. Introduction

The number of applications for nonvolatile memories is strongly increasing. The major part of this market is accounted for by flash memory cells where a single cell is electrically programmable and where a block of cells is electrically erasable [1]. In the NOR flash architecture, the floating gate of the transistor cell in Fig. 1 is usually written by channel *hot* electron (CHE) injection and erased via Fowler–Nordheim (FN) tunneling of *cold* electrons. While tunneling of *cold* electrons is well described by the FN current [2], the lucky–electron model [3] or the carrier injection model of Fiegna [4] normally employed in drift–diffusion (DD) and hydrodynamic (HD) simulations for the injection of *hot* electrons lead to differences in the gate current of about eight orders of magnitude as can be seen in Fig. 2. A rigorous treatment of hot–electron transport is possible by full–band Monte Carlo (FBMC) simulation. However, the energy associated with the momentum parallel to the oxide/silicon interface is not available for transmission across the potential barrier so that the total electron energy recorded during MC simulation has to be reduced accordingly when computing the transmission probability. In contrast to the case of parabolic band structures with a unique electron mass [2], the analogous determi-

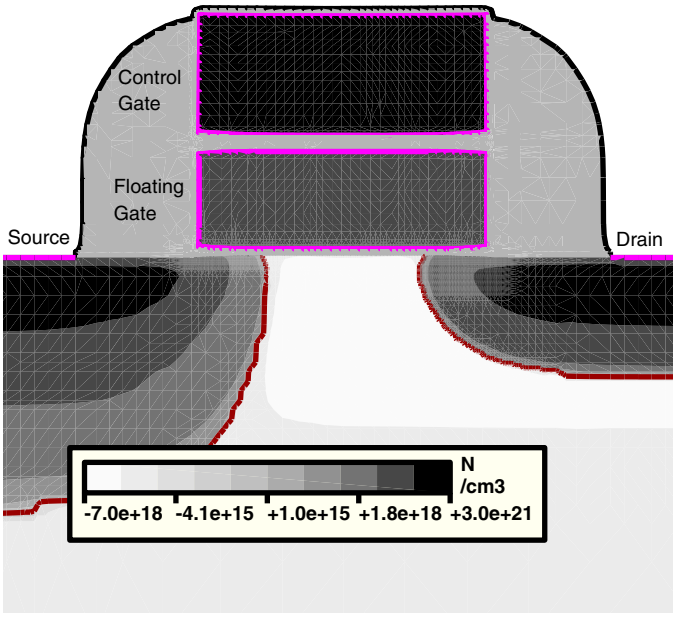


Fig. 1. Schematic cross section of a flash memory cell consisting of an nMOSFET with a floating gate and a control gate.

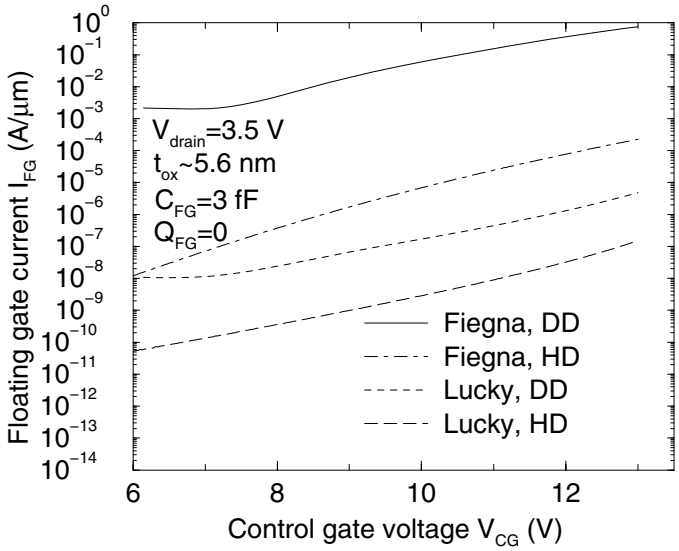


Fig. 2. Quasi–stationary gate currents as a function of the control gate voltage according to drift–diffusion (DD) and hydrodynamic (HD) simulations based on the lucky–electron and the hot–carrier injection model of Fiegna, respectively.

nation of such a perpendicular tunneling energy is not clear for a full-band structure which is at the same time mandatory in the hot–electron regime. It is the goal of this work to propose an approximate expression for a tunneling energy to be used in the calculation of the transmission probability and to demonstrate the sensitivity of the gate current to the choice of the tunneling energy by full–band MC simulation of a flash memory cell.

## II. Fowler-Nordheim tunneling of cold electrons

The detailed physics of electron injection through the oxide into the gate is very complicated and still not fully understood, e.g. concerning the band structure of the amorphous silicon dioxide or the presence and concrete form of an image force [5]. It is not our aim to address these issues. Instead, our reference point is the calculation of the tunneling current density by Simmons [2]: The current density of electrons being in Fig. 3 transmitted through the trapezoidal potential barrier to the polysilicon gate is

$$J = \frac{2}{(2\pi)^3}\int\int\int dk_x\,dk_y\,dk_z\;v_x(k_x)\times \\ \times D(E_x(k_x))\,f_{\rm eq}(E(\mathbf{k})) \quad (1)$$

where

$$E(\mathbf{k}) = \frac{\hbar^2k_x^2}{2m}+\frac{\hbar^2k_y^2}{2m}+\frac{\hbar^2k_z^2}{2m} \\ \equiv E_x(k_x)+E_y(k_y)+E_z(k_z) \quad (2)$$

is the electron energy, $E_x$ the associated perpendicular tunneling energy,

$$D(E_x) = \exp\left(-\frac{2}{\hbar}\int_{s_1}^{s_2}dx\;\sqrt{2m(V(x)-E_x)}\right) \quad (3)$$

the transmission probability in the WKB approximation, which depends on the perpendicular tunneling energy $E_x$, $v_x=\hbar^{-1}dE_x/dk_x$ the $x$ component of the group velocity, $m$ the electron mass and $f_{\rm eq}(E)$ the equilibrium electron distribution function. Approximating the Fermi–Dirac distribution with the step function leads for high oxide fields to the Fowler–Nordheim tunneling current [2]

$$J = \frac{2.2e^3F^2}{8\pi h\varphi_0}\exp\left(-\frac{8\pi}{2.96heF}\sqrt{2m}\,\varphi_0^{1,5}\right) \quad (4)$$

where $\varphi_0$ is the barrier height and $F$ the oxide field. This formula has proved very successful in describing the dependence of the FN tunneling current of *cold* electrons on oxide thickness and gate voltage, in particular for the erase operation in silicon flash memory cells.

## III. Transmission of hot electrons

Motivated by the success of Eq. (4), it is the aim of this paper to generalize this approach for arbitrary nonequilibrium conditions where the corresponding distribution function $f(\mathbf{k})$ can be obtained from full–band Monte Carlo simulation. Hence, the goal is to describe in addition the dependence of the gate current on (i) the drain voltage associated with *hot* electrons and (ii) the anisotropic behavior of the group velocity $v_x(\mathbf{k})$ to discriminate e.g. between situations where the mean current flow in the channel is either parallel (channel hot electrons) or perpendicular (substrate hot electrons) to the oxide interface. The problem is then to determine in the expression for the gate current

$$J = \frac{2}{(2\pi)^3}\int\int\int dk_x\,dk_y\,dk_z\;v_x(\mathbf{k})\times \\ \times D(E_T(\mathbf{k}))\,f(\mathbf{k}) \quad (5)$$

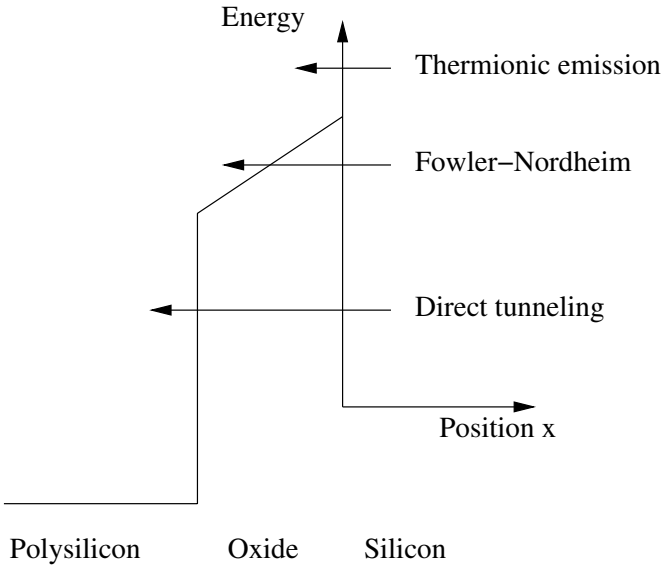


Fig. 3. Classification of the transmission mechanisms across an oxide into the polysilicon gate according to the electron's energy component perpendicular to the interface: (i) thermionic emission, (ii) Fowler–Nordheim tunneling and (iii) direct tunneling.

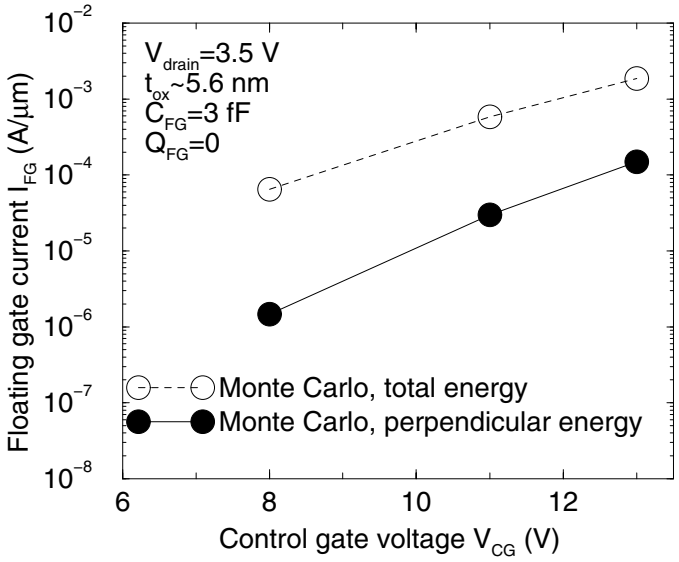


Fig. 4. Quasi–stationary gate currents as a function of the control gate voltage according to full–band Monte Carlo simulation when (i) the total energy and (ii) only the energy component perpendicular to the interface is used for the calculation of the transmission probability across the trapezoidal potential barrier in Fig. 3.

the perpendicular tunneling energy $E_T$ in the context of a general full–band structure $E(\mathbf{k})$. Here, we propose the expression

$$E_T(\mathbf{k}) \equiv E(\mathbf{k}) \times \frac{v_x^2(\mathbf{k})}{v_x^2(\mathbf{k}) + v_y^2(\mathbf{k}) + v_z^2(\mathbf{k})} \qquad (6)$$

which reduces in the limit of parabolic band structures to the perpendicular energy $E_x$ in Eq. (2). Note that apart from the generalization to a hot-electron and anisotropic formulation, the framework of approximations is still the same as in Ref. [2], involving for example the ignorance of gate isolator material properties beyond the barrier height (there is e.g. no "oxide mass" in Eqs. (1)-(3)).

The transmission probability $D(E_T)$ in Eq. (3) is taken to be 1, if the perpendicular energy is above the barrier height and the direct tunneling regime is also included via the WKB approximation, if the perpendicular electron energy is below the triangle in Fig. 3, so that all three transmission mechanisms in Fig. 3 are included. The tunneling mass $m$ is considered to be a fitting parameter. Its actual value will therefore be influenced by the approximations involved in the present model and may differ between different device structures. Hence, it does not necessarily coincide with an oxide mass the concept of which is anyway unclear due to the amorphous nature of the $SiO_2$ gate oxide. In this respect, the approximation for the tunneling energy $E_T$ in Eq. (6) has the advantage not to depend on the tunneling mass $m$. Thus it permits in particular to compare consistently the dependence of the gate current on the drain voltage also for different device structures such as floating gate flash memory cells or SONOS (semiconductor-oxide-nitride-oxide-semiconductor) memories.

## IV. Simulation results

Figure 4 shows the gate current computed by our self–consistent full–band Monte Carlo simulator SPARTA [6], if either the total electron energy $E$ or the perpendicular energy $E_T$ is used for the calculation of the transmission probability. It can be seen that taking the total energy leads to a gate current which is larger by about two orders of magnitude with decreasing tendency at higher gate voltages. The same order of magnitude and the same tendency was found in a similar comparison in Fig. 19 of Ref. [5] by Fischetti *et al.* who investigated the injection probability as a function of the oxide field with a much more sophisticated approach where electron transport in the oxide is considered or in the case of tunneling the transfer-matrix method is employed. There, the consideration of the parallel

energy reducing the available transmission energy turned out to be in good agreement with measurements. Of course, this is not yet a proof for the general applicability of our perpendicular energy concept, since e.g. the device structure studied in Ref. [5] is not the same as in our investigation. However, it is in any case a significant improvement over the use of the total energy for the transmission process [4], [7]–[9] and the gate current estimation scheme in Eq. (5) is much more efficient for Monte Carlo device simulation [10] than the elaborated procedure in Ref. [5]. A confirmation of the validity of the presented approach still requires, of course, comprehensive comparisons with gate current measurements in corresponding test structures.

## V. Conclusions

In conclusion, an approximation for the "perpendicular" tunneling energy within the WKB formula has been proposed. Full-band Monte Carlo simulations of a floating-gate memory cell have shown that the gate current is larger by about two orders of magnitude when using the total energy for tunneling instead, similar to previous results with a more elaborated model in the literature. The presented formulation is well suited for studying hot-electron injection in nonvolatile memory cells.

## Acknowledgment

We thank Y. Saad, P. Pfäffli and A. Erlebach for useful discussions and collaboration. Financial support by Fujitsu Laboratories Ltd. (Akiruno, Japan) is gratefully acknowledged.

## References

[1] P. Pavan, R. Bez, P. Olivo, and E. Zanoni, "Flash memory cells — An overview," *Proc. IEEE*, vol. 85, pp. 1248–1271, 1997.

[2] J. G. Simmons, "Generalized formula for the electric tunnel effect between similar electrodes separated by a thin insulating film," *J. Appl. Phys.*, vol. 34, pp. 1793–1803, 1963.

[3] K. Hasnat, C.-F. Yeap, S. Jallepalli, W.-K. Shih, S. A. Hareland, V. M. Agostinelli, Jr., A. F. Tasch, Jr., and C. M. Maziar, "A pseudo–lucky electron model for simulation of electron gate current in submicron NMOSFET's," *IEEE Trans. Electron Devices*, vol. 43, pp. 1264–1273, 1996.

[4] C. Fiegna, F. Venturi, M. Melanotte, E. Sangiorgi, and B. Riccò, "Simple and efficient modeling of EPROM writing," *IEEE Trans. Electron Devices*, vol. 38, pp. 603–610, 1991.

[5] M. V. Fischetti, S. E. Laux, and E. Crabbé, "Understanding hot–electron transport in silicon devices: Is there a shortcut?" *J. Appl. Phys.*, vol. 78, pp. 1058–1087, 1995.

[6] F. M. Bufler, C. Zechner, A. Schenk, and W. Fichtner, "Single–particle approach to self–consistent Monte Carlo device simulation," *IEICE Trans. Electron.*, vol. E86–C, pp. 308–313, 2003.

[7] K. Hasnat, C.-F. Yeap, S. Jallepalli, S. A. Hareland, W.-K. Shih, V. M. Agostinelli, Jr., A. F. Tasch, and C. M. Maziar, "Thermionic emission model of electron gate current in submicron NMOSFETs," *IEEE Trans. Electron Devices*, vol. 44, pp. 129–138, 1997.

[8] A. Ghetti, L. Selmi, and R. Bez, "Low–voltage hot electrons and soft–programming lifetime prediction in nonvolatile memory cells," *IEEE Trans. Electron Devices*, vol. 46, pp. 696–702, 1999.

[9] D. R. Nair, S. Mahapatra, S. Shukuri, and J. D. Bude, "Drain disturb during CHISEL programming of NOR flash EEPROMs — Physical mechanisms and impact of technological parameters," *IEEE Trans. Electron Devices*, vol. 51, pp. 701–707, 2004.

[10] F. M. Bufler, A. Schenk, and W. Fichtner, "Efficient Monte Carlo device modeling," *IEEE Trans. Electron Devices*, vol. 47, pp. 1891–1897, 2000.