Simulation of Non-Volatile Memory Cells by Accounting for Inelastic Trap-Assisted Tunneling Current

A. Gehring, F. Jiménez-Molinos[†], A. Palma[†], F. Gámiz[†], H. Kosina, and S. Selberherr

Institute for Microelectronics, TU Vienna, Gusshausstr. 27–29, A-1040 Vienna, Austria
Phone +43-1-58801-36016, FAX +43-1-58801-36099, e-mail: gehring@iue.tuwien.ac.at

[†] Departamento de Electrónica, Facultad de Ciencias, Universidad de Granada
Avenida Fuentenueva s/n 18071-Granada, Spain

Abstract

A new model for the simulation of physical inelastic trap-assisted tunneling has been implemented in the device simulator MINIMOS-NT. It is based on a recently published model for multi-phonon transitions between trapped and detrapped electronic states. We show that the numerical solution of the Schrödinger equation can be approximated by analytical expressions, avoiding the need for numerical integrations. We calibrate the model to measurement data from recent literature and find excellent agreement. The model is used to calculate the decharging characteristics of EEPROM devices. Furthermore, an expression for transient tunneling has been developed.

1 Introduction

It is generally accepted that stress-induced leakage current (SILC) is mainly responsible for retention time degradation in state-of-the-art EEPROM device[1][2]. Inelastic tunneling incorporating a charge loss via phonon transitions is identified to be the main source of this leakage current. This is supported by the relatively small voltage dependence of SILC and by energy-loss measurements [3]. However, models for inelastic trap-assisted tunneling tend to be computationally expensive and literature still lacks a physics-based model suitable for implementation in two- or even three-dimensional device simulators, where the tunneling process has to be evaluated in every iteration for the sake of self-consistency.

2 Tunneling Model

The new tunneling model is based on the model presented in [4]. It describes a two-step tunneling process via traps in the oxide, incorporating energy loss via phonon relaxation. Band bending and Fermi energies are provided via the solution of the drift-diffusion transport model of MINIMOS-NT. The trap-assisted tunneling current is given as

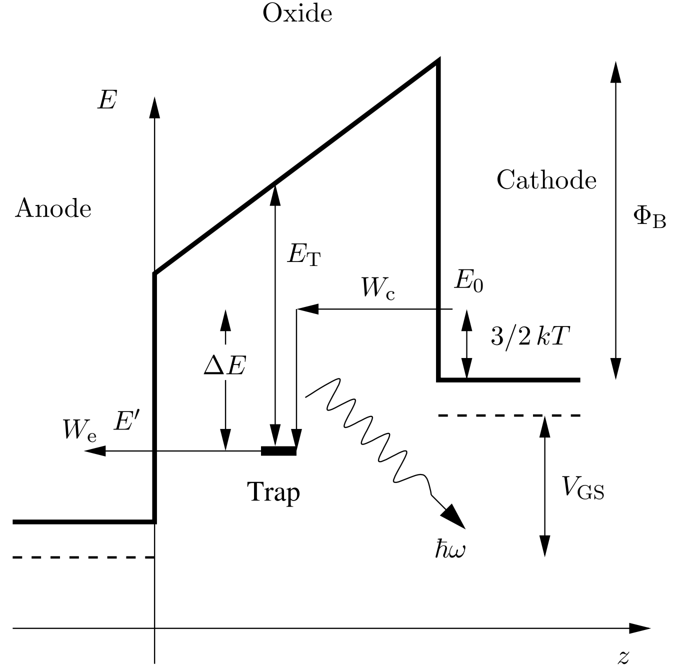

Figure 1: Trap-assisted tunneling transition.

$$J_{TAT} = q \int_0^{t_{ox}} \frac{N_T}{\tau_c(z) + \tau_e(z)}\, dz \tag{1}$$

where N_T denotes the trap density, $\tau_c(z)$ is the capture time and $\tau_e(z)$ the emission time of a trap located at position z (see Fig. 1). The capture and emission times are derived from

$$\tau_c^{-1}(z) = \int_{E'}^{\infty} N_c(E) f_c(E) W_c(z, E', E)\, dE \tag{2}$$

$$\tau_e^{-1}(z) = \int_{E'}^{\infty} N_a(E)(1 - f_a(E)) W_e(z, E', E)\, dE \tag{3}$$

where E' is the energy position of the trap, N_c and N_a are the density of states, and f_c and f_a the Fermi func-

tions in the cathode and the anode, respectively. The capture probability per time unit $W_\mathrm{c}(z, E', E)$ for an electron transition between E and E' is given as

$$W_\mathrm{c}(z, E', E) = \frac{\pi}{\hbar^2\omega} |V_\mathrm{e}|^2 \, S \left(1 - \frac{p}{S}\right)^2 \cdot I_p(\xi) \cdot \exp\left[-(2n+1)\,S + \frac{\Delta E}{2k_\mathrm{B}T}\right]. \quad (4)$$

In this expression, S is the Huang-Rhys factor, $\hbar\omega$ is the phonon energy, $p = \Delta E/\hbar\omega$ the number of emitted phonons, $\xi = 2S[n(n+1)]^{1/2}$, $\Delta E = E - E'$, and the population of phonons given by the Bose-Einstein statistics:

$$n = \left[\exp(\hbar\omega/kT) - 1\right]^{-1}. \quad (5)$$

The transition matrix element $|V_\mathrm{e}|^2$ is computed by an integration over the trap cube (Fig. 2)

$$|V_\mathrm{e}|^2 = 5\pi S(\hbar\omega)^2 \frac{a_\mathrm{T}^2}{V} \int\limits_{z_0 - z_\mathrm{T}/2}^{z_0 + z_\mathrm{T}/2} |\zeta(z)|^2 \, dz \quad (6)$$

where the trap radius a_T and the corresponding trap cube side length z_T is

$$a_\mathrm{T} = \frac{\hbar}{\sqrt{2m_\mathrm{ox}E_\mathrm{T}}} \qquad z_\mathrm{T} = a_\mathrm{T}\left(\frac{4\pi}{3}\right)^{1/3}. \quad (7)$$

The volume V is cancelled out by the respective term of the density of states, and $\zeta(z)$ denotes the wave function in the oxide. Since the numerical evaluation of the wave functions in the oxide degrades the computational efficiency of a multi-purpose device simulator where simulation speed is a critical topic, we used an analytic approximation for the wave function. We assumed plane waves and replaced the triangular or trapezoidal barriers by a series of rectangular barriers as shown in Fig. 2. The integration in (6) yields two analytical expressions valid for the direct (trapezoidal barrier) and Fowler-Nordheim (triangular barrier) region. Furthermore, the integrations in (2) and (3) can be avoided by two assumptions. For the capture process it is assumed that all electrons depart from the same energy level $E_0 = E_\mathrm{c} + 3/2kT$, therefore

$$\tau_c^{-1}(z) = W_c(z, E', E_0) \cdot \int\limits_{E_C}^{\infty} N_c(E) f_c(E) dE \quad (8)$$

where the right integral equals the carrier concentration which is supplied by the transport model. For the emission time, zero-phonon transitions have been assumed. The electrons are thus emitted from the trap without absorption of phonons. With this approximation the emission probability is $W_e(z, E', E') = W_c(z, E', E')$. Finally, the approximation $f_a(E) \simeq 0$ leads to

$$\tau_e^{-1}(z) = W_e(z, E', E') \cdot N_a(E'). \quad (9)$$

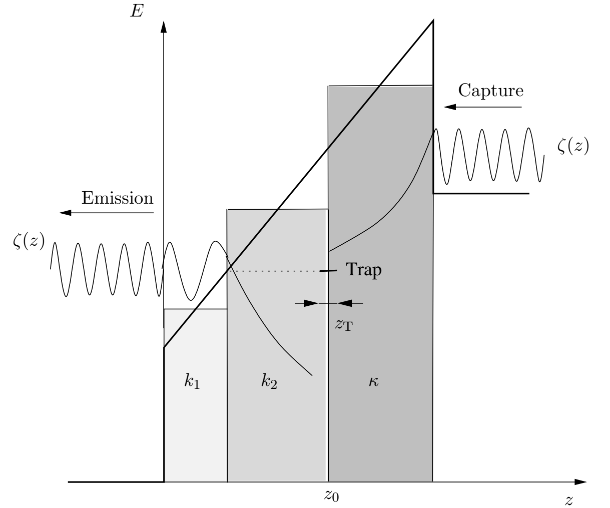

Figure 2: Approximation of the barrier in the Fowler-Nordheim region.

The total tunneling current is derived as the sum of trap-assisted tunneling current and direct or Fowler-Nordheim tunneling current. Since in the direct tunneling regime (low voltages) the trap-assisted tunneling component exceeds the direct tunneling component by orders of magnitude, only the Fowler-Nordheim regime is critical for which we used the expression

$$J_\mathrm{FN} = \frac{q^3}{8\pi h \Phi_B} \cdot F_{ox}^2 \cdot \exp\left(-\frac{4\sqrt{2m\Phi_B^3}}{3q\hbar F_{ox}}\right)$$

with F_ox being the electric field in the oxide and Φ_B the barrier height[5]. The total tunneling current is the sum of J_TAT and J_FN.

3 Simulation Results

The model has been implemented in the device simulator MINIMOS-NT in a self-consistent manner. The gate oxide is divided into one-dimensional slices with constant potential and material properties, and the total current is found by a summation over all slices. The input-deck parameters for the model comprise the Huang-Rhys factor, the trap concentration, the electron masses in the oxide, anode and cathode, and the trap energy level. Electrons which tunnel through the oxide are considered in the continuity equation in the substrate to assure self-consistency.

3.1 Capture and Emission Times

Fig. 3 and Fig. 4 show the capture end emission times as a function of the z-coordinate in the oxide for the direct tunneling regime ($V_\mathrm{GS} = 3\,\mathrm{V}$) and the Fowler-Nordheim regime ($V_\mathrm{GS} = 7\,\mathrm{V}$). In these figures, the analytical solu-

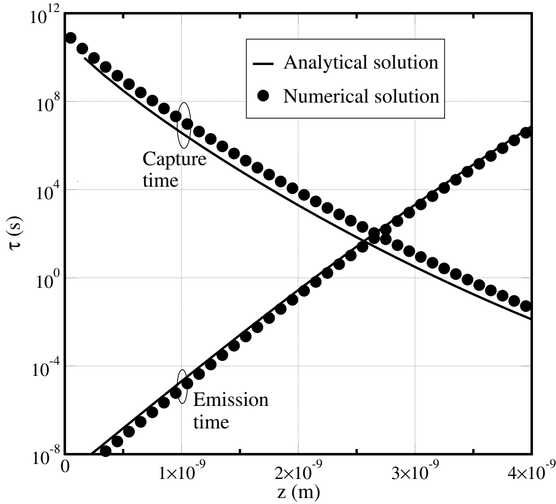

Figure 3: Capture and emission times for analytical and numerical solution at $V_{\mathrm{GS}} = 3\,\mathrm{V}$.

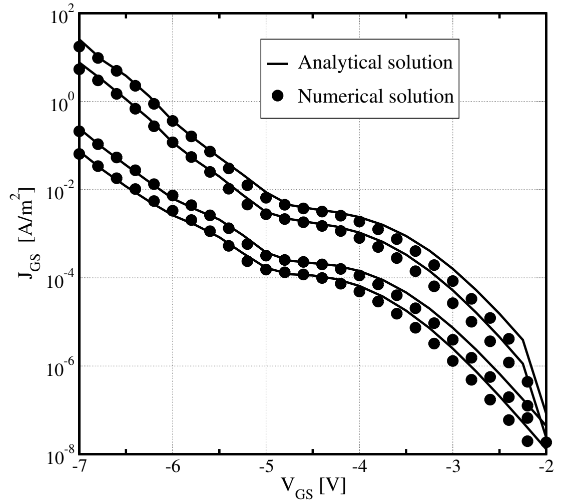

Figure 5: Comparison of the analytical and numerical solution for the trap-assisted component of the tunneling current.

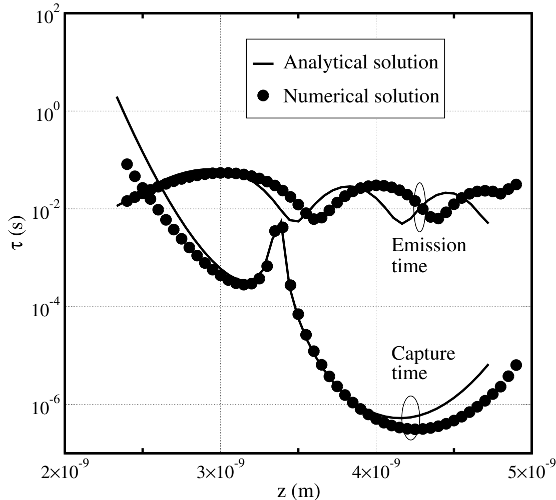

Figure 4: Capture and emission times for analytical and numerical solution at $V_{\mathrm{GS}} = 7\,\mathrm{V}$.

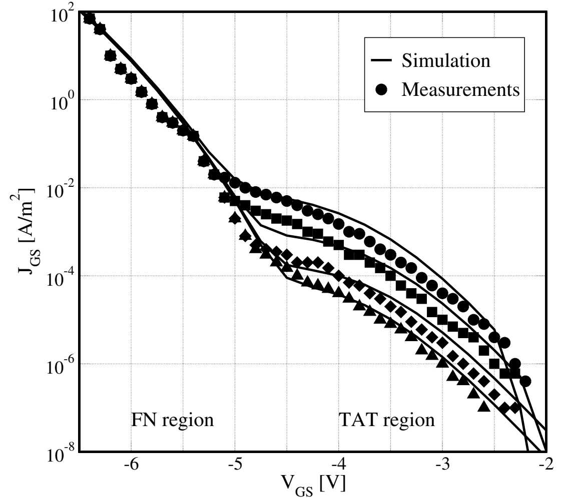

Figure 6: Comparison to measurement data. The values are taken from [1].

tion of the Schrödinger equation is compared to the numerical solution using the simulation method described in [4]. Electrons are captured from the right and emitted to the left. It can be seen that our analytical model fits almost perfectly in the direct tunneling regime, while it shows some inaccuracy in the Fowler-Nordheim region. The oscillations in this regime are due to the fact that the phase of the electron wave function is not constant but depends on the voltage. It is interesting to note that the capture time shows a peak in the middle of the oxide.

Despite the slight deviations in the position-dependent emission time as compared to the numerical solution, the resulting values for J_{TAT} are nearly exactly the same as shown in Fig. 5. This is due to the assumption of a position-independent trap concentration N_{T}. The total current only depends on the integral of the capture and emission times, and not on their exact shape. Thus, the approximations for the wave functions cause no significant deterioration of the results.

3.2 Steady-state tunneling

The model is compared to measurement values taken from [1] for a MOS capacitor with an oxide thickness of 5 nm and stressing times of 0.1 s, 1 s, 10 s, and 100 s. We used values of $E_{\mathrm{T}} = 2.6...2.8\,\mathrm{eV}$, $S = 65...85$, and $N_{\mathrm{T}} = 2 \times 10^{19}...4 \times 10^{18}\,\mathrm{cm}^{-3}$ to fit the measurements. These numbers correspond well to the values reported in [4], which justifies the assumptions described above. The model was then used to simulate charge loss in a fully charged EEPROM device. Fig. 7 shows the decharging characteristic of an EEPROM device simulated before (upper curve) and after stressing with different stress times. The device is shown in the inset of Fig. 7. The

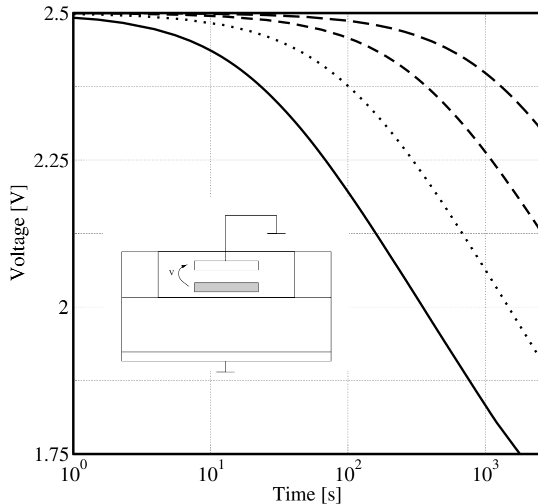

Figure 7: Decharging curve of an EEPROM for the stressing conditions as in Fig. 6.

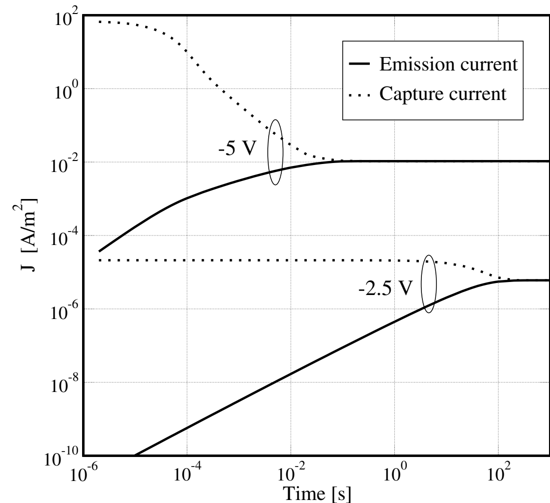

Figure 8: Transient capture and emission currents, starting from flatband condition at $t = 0\,s$.

values for trap concentration, S, and trap energy were set equal to that of Fig. 6.

3.3 Transient tunneling

Transient effects must be taken into account for the proper simulation of trap-assisted tunneling current. The change rate of the concentration of occupied traps $n_\mathrm{T}(z)$ with respect to time is given by

$$\frac{d\,n_\mathrm{T}(z)}{d\,t} = N_\mathrm{T}\,[1 - f_\mathrm{T}(z)]\,\tau_\mathrm{c}^{-1}(z) - N_\mathrm{T} f_\mathrm{T}(z)\tau_\mathrm{e}^{-1}(z). \tag{11}$$

For steady state current the condition $d\,n_\mathrm{T}(z)/d\,t = 0$ is used to derive expression (1). However, for transient charging effects capture and emission currents are not equal. We simulated a trap charging characteristic by applying a voltage step on the cathode starting from the flat band condition. Fig. 8 shows the charging characteristic. It can be seen that the charging current exceeds the steady-state current by orders of magnitude, and the time after which a steady state is reached depends heavily on the applied voltage, as reported in [3].

4 Conclusion

We showed the simulation of inelastic trap-assisted tunneling current using the device simulator MINIMOS-NT. A recently published model has been adapted to be feasible for a device simulator, namely avoiding the exact numerical solution of the Schrödinger equation and computationally expensive integrations. While capture and emission times can only be fit to a reasonable accuracy, the total tunneling current agrees well with measurements. We thus identify this model as a reasonable compromise between simulation speed and accuracy, being well suited for implementation in two- and three-dimensional device simulators.

Acknowledgment

Part of this work has been carried out within the framework of Research Project No. PB97-0815, supported by the Spanish Government.

References

[1] B. Ricco, G. Gozzi, and M. Lanzoni, “Modeling and Simulation of Stress-Induced Leakage Current in Ultrathin SiO2 Films,” *IEEE Trans.Electron Devices*, vol. 45, pp. 1554–1560, July 1998.

[2] W. J. Chang, M. P. Houng, and Y. H. Wang, “Simulation of Stress-Induced Leakage Current in Silicon Dioxides: A Modified Trap-Assisted Tunneling Model considering Gaussian-Distributed Traps and Electron Energy Loss,” *J.Appl.Phys.*, vol. 89, June 2001.

[3] D. Ielmini, A. S. Spinelli, and M. A. R. A. L. Lacaita, “Modeling of SILC Based on Electron and Hole Tunneling - Part I: Transient Effects,” *IEEE Trans.Electron Devices*, vol. 47, pp. 1258–1265, June 2000.

[4] F. Jimenez-Molinos, A. Palma, F. Gamiz, J. Banqueri, and J. A. Lopez-Villanueva, “Physical Model for Trap-Assisted Inelastic Tunneling in Metal-Oxide-Semiconductor Structures,” *J.Appl.Phys.*, vol. 90, no. 7, pp. 3396–3404, 2001.

[5] M. Lenzlinger and E. Snow, “Fowler-Nordheim Tunneling into Thermally Grown SiO_2,” *J.Appl.Phys.*, vol. 40, no. 1, pp. 278–283, 1969.