Multiscale Modeling of Charge Trapping in Molecule Based Flash Memories

Oves Badami
School of Engineering
University of Glasgow
Glasgow, United Kingdom
oves.badami@glasgow.ac.uk

Toufik Sadi
School of Science
Aalto University
Aalto, Finland
toufik.sadi@aalto.fi

Vihar Georgiev
School of Engineering
University of Glasgow
Glasgow, United Kingdom
vihar.georgiev@glasgow.ac.uk

Fikru Adamu-Lema
School of Engineering
University of Glasgow
Glasgow, United Kingdom
fikru.adamu-lema@glasgow.ac.uk

Vasanthan Thirunavukkarasu
School of Engineering
University of Glasgow
Glasgow, United Kingdom
vasanthan.thirunavukkarasu@glasgow.ac.uk

Jie Ding
College of Electrical and Power Engineering
Taiyuan University of Technology,
030024 China
dingjie2015@foxmail.com

Asen Asenov
School of Engineering
University of Glasgow
Glasgow, United Kingdom
asen.asenov@glasgow.ac.uk

***Abstract*—To keep up with the increase in demand for storing data, flash memories have been scaled down dramatically and stacked by the semiconductor industry. Furthermore, processing large data has highlighted the limitations of the von Neumann architecture. To overcome this, different types of memory devices like Resistive Random-Access Memories (RRAMs) have also gained a lot of importance. Hence, carrier dynamics in oxides has gained significant traction in recent years. In this work, we discuss the kinetic Monte Carlo methodology as implemented in our integrated simulation environment NESS (Nano-Electronic Simulation Software) that allows us to study carrier transport in the oxide using accurate physics based models. As an example, we study the retention characteristics in a molecule based flash memory.**

***Index Terms*—Kinetic Monte Carlo, POM molecule, Trap Assisted Tunneling**

I. Introduction

In the last two decades, the SiO_2 based gate oxide in metal-oxide-semiconductor field-effect transistors (MOSFETs) has been replaced by high-κ dielectrics in order to maintain the down-scaling trend of these devices. Nonvolatile memory devices likewise have also undergone remarkable changes to increase their packing density, to keep up with the increase in the demand to store data. Hence, charge trap and nano-crystal based flash memories have attracted substantial attention because of their ability to overcome the intercell capacitance problem of the floating gate based flash memories [1]. Furthermore, oxide-based Resistive Random Access Memory (RRAM) devices have gained recently a huge interest, as they are extremely well-suited for in-memory computing [2], [3]. All these advancements in memory technology necessitate the development of reliable tools for the modeling of carrier dynamics in oxides.

The research is funded by the EPSRC (UK) grants no. EP/S000224/1 and EP/S001131/1. This work is partly supported by NSFC project (Project No. 61604105)

978-1-7281-0940-4/19/$31.00 ©2019 IEEE

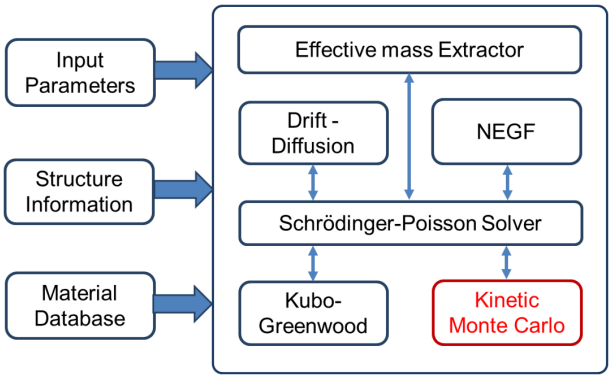

Fig. 1. Capabilities of the Nano-Electronic Simulation Software (NESS) framework.

Fig. 1 illustrates the different modules available in our in-house comprehensive 3D Nano-Electronic Simulation Software (NESS) [4], which can be employed to study different aspects of the carrier transport in modern day transistors. In this work, we report on the extended the capabilities of NESS to simulate charge transport in oxide layers using the kinetic Monte Carlo (kMC) simulation methodology [5]. Here, we illustrate the applicability of the new kMC module by performing a simulation study of the retention time in molecule based flash memory.

II. Methodology

Fig. 2 shows the flowchart of the kMC module implemented in NESS. After reading and initializing the simulation details, such as the trap type (neutral or positive), trap status (filled or empty) at time (0 sec), Poisson's equation is solved self-consistently in the entire 3D domain along with the equilibrium carrier statistics. This is followed by randomly generating the initial electron state (kinetic energy of the electron)

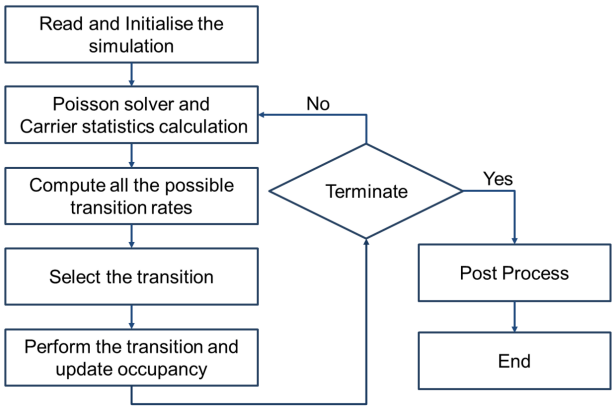

Fig. 2. A flowchart of the implementation of the kinetic Monte Carlo module in NESS.

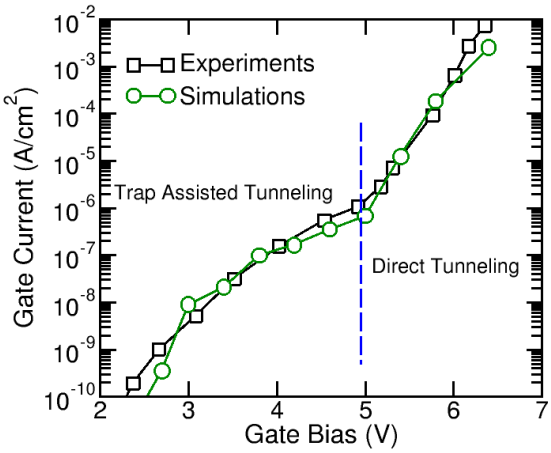

Fig. 3. Comparison of the gate current calculated in this work and experiment. The semiconductor-oxide barrier and electron mass in the oxide were tuned to 3 eV and 0.42 m_0, respectively, to calibrate to the experimental data which agree with those reported in the literature [9].

at the silicon-oxide and metal-oxide interfaces according to the Fermi-Dirac distribution. Following this, all the possible transition rates for the elastic and inelastic trap assisted tunneling, Poole-Frenkel emission, defect-to-defect tunneling mechanisms are computed [5]–[7]. In order to calculate the tunneling rates, the transmission coefficients were evaluated using the Wentzel-Kramers-Brillouin (WKB) approximation assuming a parabolic effective mass. This is followed by randomly selecting a transition, μ, from the cumulative ladder constructed from all the possible tunneling rates according to [5]:

$$\frac{\Sigma_{\mu'=1}^{\mu-1} R_{\mu'}}{R_{TOT}} < r \leq \frac{\Sigma_{\mu'=1}^{\mu} R_{\mu'}}{R_{TOT}}, \tag{1}$$

where r is a uniformly distributed random number between 0 and 1, R_{TOT} is the total transition rate and $R_{\mu'}$ is the transition rate due to the process indexed by μ'. The trap is either populated or depopulated by an electron depending on whether the transition is of an electrode-to-defect nature or vice-versa. The time step, τ, can then be calculated as [5] :

$$\tau = -\frac{ln\,(r')}{R_{TOT}} \tag{2}$$

where r' is a uniformly distributed random number between 0 and 1 and R_{TOT} is the total transition rate for different tunneling processes. After checking the termination condition, Poisson's equation is solved again to update the electrostatic potential, to account for the rearrangement of the charge in the simulation domain.

The implemented kMC module is flexible allowing different looping conditions, such as terminating when all the traps are filled or empty in addition to the conventional check on the total simulation time. This is extremely useful in analyzing retention, programming and erase times for the memory devices.

III. Results

In order to validate the implementation, we compared the gate current calculated using the methodology discussed in Section II along with the experimental data published in the literature [8]. The simulated structure corresponded to a MOS capacitor having an oxide thickness of 5 nm, having a trap density of about 3×10^{19} cm^{-3} which agrees with those reported in the literature [9]. In this work, the traps were assumed to be neutral when empty and uniformly distributed in the oxide. The average energy of trap was assumed to be 2.7 eV which also agrees well with those reported in literature [9]. The substrate doping was taken to be 10^{18} cm^{-3}. The relaxation energy, $E_{REL} = S\hbar\omega_0$ was calibrated to 0.6 eV which agrees well with the value for SiO_2 reported in the literature [10]. The direct tunneling current component was calculated using the Tsu-Esaki formulation [5].

Fig. 4(a) shows a schematic of the charge trap flash memory cell with a $W_{18}O_{54}(SO_3)_2$ molecule (polyoxymetalate, POM) layer. Fig. 4(b) shows the kMC simulation domain adopted here. The Lowest Unoccupied Molecular Orbital (LUMO) energy level was set to 3.67 eV (with respect to the conduction band edge of the SiO_2) as calculated from the first-principle simulations [11]. The POM molecules were approximated by point defects with a cross-section area of $10^{-14}cm^2$ [11]. The p-type doping concentration in the substrate is 10^{18} cm^{-3}. We assume that POM molecules are arranged in a 3×3 configuration embedded in the SiO_2 layer. The total oxide thickness (T_{TUN}+T_{CON}) was set to 14.5 nm. Fig. 5 shows the threshold voltage (V_T) as a function of the number of electrons trapped in the POM molecules. The threshold voltage is evaluated at the center of the device. V_T was defined as the gate bias at which the electron density at the oxide-semiconductor interface becomes 10% of the p-type doping. As expected the presence of the electrons in the POM molecules increases the threshold voltage.

Fig. 6 shows the conduction band edge, E_C, in the plane of the POM molecules at different time instants as the electrons

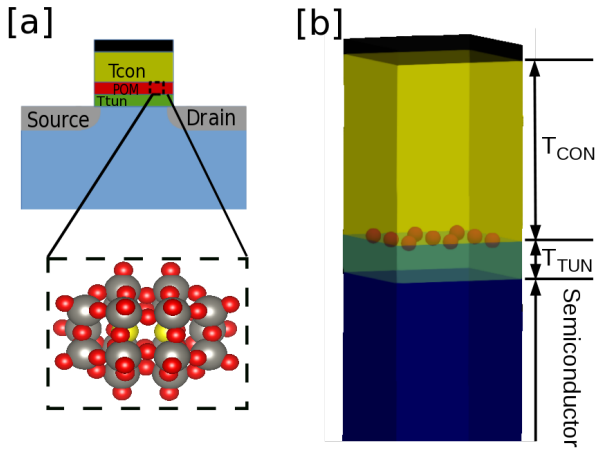

Fig. 4. (a) Illustration of a charge trap flash memory cell with a POM layer. A POM molecule (without counter-cations) is also illustrated in the inset. (b) Schematic of the kMC simulation domain. The 3×3 configuration of the POM molecules is also illustrated (red spheres).

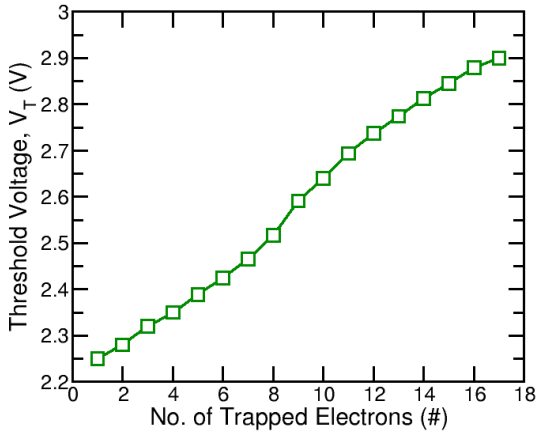

Fig. 5. Calculated threshold voltage (V_T) as a function of the number of electrons trapped in the oxide.

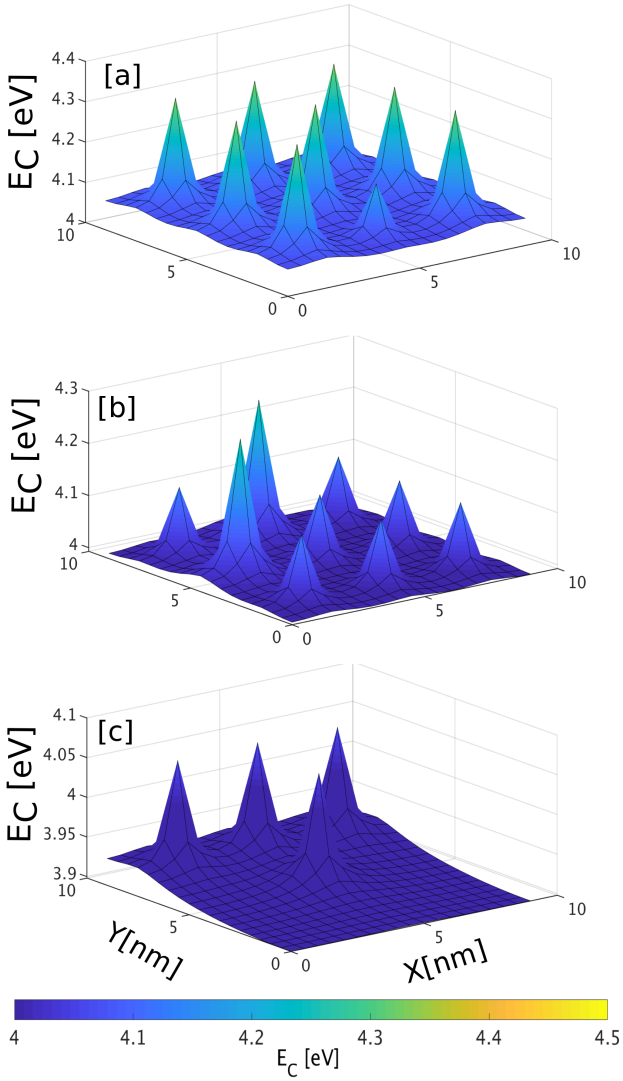

Fig. 6. The evolution of the conduction band edge in the plane of the POM molecules as the electrons tunnel out of the POM molecules with (a) 17 (b) 11 and (c) 4 electrons trapped in the oxide.

from the POM molecules tunnel out through the tunnel oxide. As it is evident, the landscape of the E_C changes dramatically at these time instants because of the different number of the electrons present in the oxide. The presence of the electrons in the POM molecule affects the local potential and hence the electric field. This has a significant impact on the electron tunneling rate out of the doubly charged POM molecules.

Fig. 7 shows the variation of V_T as a function of time where the tunneling rates are calculated self-consistently with the Poisson's equation solution and in the case where the self-consistency is not considered. For the purpose of simulations, $E_{REL} = 0.36eV$ was used [10]. In the absence of the self-consistency, the electrostatic potential in the oxide is not updated and hence all the tunneling rate calculations do not account for the correct potential. When self-consistency is applied, then during the time period $t < t_1$, all the POM molecules are significantly more likely to undergo transformation from doubly charged (2× reduced) to singly charged (1× reduced) state. This is followed by a significant reduction in the tunneling rates as the local electrostatic potential is modified (as shown in Fig. 6) out of the POM molecules which gives the distinct plateau in Fig. 7. Then, subsequently, POM molecules undergo a transition from 1× reduced state to their parent (neutral) state. Fig. 8 compares the impact of the E_{REL} on the retention behaviour of the flash memory cell shown in Fig. 4. It can seen that as the E_{REL} is increased the retention time increases because of the reduction in the multiphonon

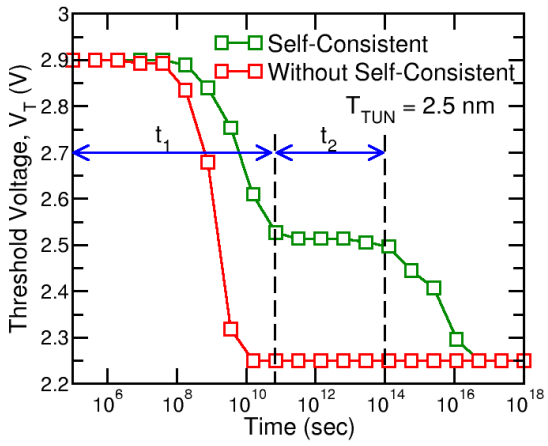

Fig. 7. The impact of the self-consistency between the tunneling rates and Poisson equation on the threshold voltage evolution. For the purpose of simulations, it was assumed that initially all the POM molecules contain two electrons.

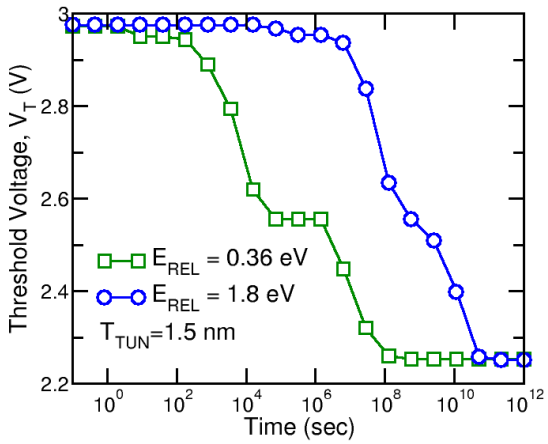

Fig. 8. Impact of the $E_{REL} = S\hbar\omega$ on the retention behaviour. For the purpose of simulation, the phonon energy was assumed to be 0.06 eV while the Huang-Rhys factor (S) was varied.

transition probability. In addition to this, with the increase in the E_{REL} the distinct plateau like feature is also softened. Comparison of Fig.7 and Fig.8 highlights an important fact that with reduction in the T_{TUN} the retention time decreases significantly.

IV. Conclusion

We have developed a kMC module within the in-house simulation framework NESS. This module enables us to perform a complete 3D analysis of carrier transport in oxide layers using a comprehensive set of physics-based models. We have bench-marked the module by comparing it with the experimental results for gate tunneling currents. The module allows the user to specify the individual trap location in the 3D space, its type (positive or neutral), the energy depth and cross-section area. Furthermore, by employing a very simple user interface, we can switch on or off different mechanisms, such as electrode-to-defect, defect-to-electrode and defect-to-defect tunneling, enabling the study of their individual impact.

References

[1] S. Tiwari *et al.*, "Volatile and non-volatile memories in silicon with nano-crystal storage," Proceedings of International Electron Devices Meeting, Washington, DC, USA, 1995, pp. 521-524.

[2] L. Larcher *et al.*, "Multiscale modeling of neuromorphic computing: From materials to device operations," 2017 IEEE International Electron Devices Meeting (IEDM), pp. 11.7.1-11.7.4, 2017.

[3] M. A. Zidan *et al.*, "A General Memristor-Based Partial Differential Equation Solver," Nature Electronics, 1 (7), 411420, 2018.

[4] S. Berrada *et al.*, "NESS: new flexible Nano-Electronic Simulation Software" SISPAD 2018.

[5] G. C. Gunther, "Modeling of leakage currents in high-κ dielectrics," PhD Thesis, Technical University Munich, Germany 2015.

[6] T. Sadi *et al.*, "Investigation of resistance switching in SiOx RRAM cells using a 3D multiscale kinetic Monte Carlo simulator," Jour. Phys.: Condens. Matter, 30, pp. 084005, 2018.

[7] M. Herrmann *et al.*, "Field and high-temperature dependence of the long term charge loss in erasable programmable read only memories: Measurements and modeling," Jour. of Appl. Phys. vol., 77 no. 9, 1995.

[8] B. Ricco *et al.*, " Modeling and simulation of stress-induced leakage current in ultrathin SiO_2 films," Trans. on Electron Devices, vol. 47, no. 7, 1998.

[9] F. Jimenez-Molinos *et al.*, "Physical model for trap-assisted inelastic tunneling in metal-oxide-semiconductor structures," Jour. of Appl. Phys., vol. 90 no. 7, 2001.

[10] L. Vandelli *et al.*, " Physical Model of the Temperature Dependence of the Current Through SiO_2/HfO_2 Stacks," Trans. on Electron Devices, vol. 58, no. 9, 2011.

[11] V. P. Georgiev *et al.*, "Comparison Between Bulk and FDSOI POM Flash Cell: A Multiscale Simulation Study," Trans. on Electron Devices, vol. 62, no. 2, Feb. 2015.