Medina Bailon, C., Sadi, T., Sampedro, C., Padilla, J. L., Donetti, L., Georgiev, V. , Gamiz, F. and Asenov, A. (2019) Impact of the trap attributes on the gate leakage mechanisms in a 2D MS-EMC nanodevice simulator. *Lecture Notes in Computer Science*, 11189, pp. 273-280.(doi:10.1007/978-3-030-10692-8_30)

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

http://eprints.gla.ac.uk/178378/

Deposited on: 22 January 2019

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk

# Impact of the Trap Attributes on the Gate Leakage Mechanisms in a 2D MS-EMC Nanodevice Simulator*

C. Medina-Bailon[1,2], T. Sadi[3], C. Sampedro[2], J.L. Padilla[2], L. Donetti[2], V. Georgiev[1], F. Gamiz[2], and A. Asenov[1]

[1] School of Engineering, University of Glasgow, Glasgow G12 8LT, Scotland, UK
Cristina.MedinaBailon@glasgow.ac.uk
[2] Nanoelectronics Research Group, Universidad de Granada, 18071 Granada, Spain
[3] Department of Neuroscience and Biomedical Engineering, Aalto University, P.O. Box 12200, FI-00076 Aalto, Finland

**Abstract.** From a modeling point of view, the inclusion of adequate physical phenomena is mandatory when analyzing the behavior of new transistor architectures. In particular, the high electric field across the ultra-thin insulator in aggressively scaled transistors leads to the possibility for the charge carriers in the channel to tunnel through the gate oxide via various gate leakage mechanisms (GLMs). In this work, we study the impact of trap number on gate leakage using the GLM model, which is included in a Multi-Subband Ensemble Monte Carlo (MS-EMC) simulator for Fully-Depleted Silicon-On-Insulator (FDSOI) field effect transistors (FETs). The GLM code described herein considers both direct and trap-assisted tunneling. This work shows that trap attributes and dynamics can modify the device electrostatic characteristics and even play a significant role in determining the extent of GLMs.

**Keywords:** gate leakage mechanism · direct tunneling · trap assisted tunneling · MS-EMC · FDSOI.

## 1 Introduction

Reducing the gate oxide thickness implies an increase in the field across the oxide. The high electric field coupled with thin oxides leads to the possibility of charge carriers traversing the barrier for transport set up by the dielectric layer, resulting in tunneling processes from (to) substrate to (from) gate through the gate oxide [1, 2], and thus giving rise to a certain gate current. This effect is known as the gate leakage mechanism (GLM) and includes both direct and trap-assisted tunneling.

The direct tunneling (DT) processes are always present, even if a dielectric film of perfect quality is assumed. In the case where a very thin oxide layer (less

* The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688101 SUPERAID7.

than $3-4$nm) is considered, electrons forming the inversion layer can tunnel to the gate through the energetically forbidden band gap of the dielectric material. Similarly, the trap assistant tunneling (TAT) processes are related to the existence of defect states which cause elastic or inelastic tunneling of electrons into and out of defects. In general, direct tunneling is the dominant phenomenon due to the small oxide thickness [3]. Nevertheless, trap attributes and dynamics can modify the device electrostatic properties.

Apart from the analysis of these tunneling mechanisms, the investigation of these processes on new technological nanodevices is mandatory. Currently, Fully-Depleted Silicon-On-Insulator (FDSOI) devices have been recognized as an alternative to bulk devices. However, the impact of GLM mainly depends on the electron confinement near the interface [3]. Accordingly, the number of electrons tunneling through the oxide is higher in the single gate FDSOI in contrast with other double gate devices, such as the vertical FinFET. Therefore, the study of this mechanism in FDSOI devices is of special interest.

The aim of this work is to perform a study of the impact of the trap attributes on the GLM and thus on the FDSOI performance. For this purpose, a detailed discussion of this transport mechanism is given together with the details of the stochastic simulation process in Section 2. The main findings are reported in Section 3 including a meticulous analysis of how the trap density modifies the GLM and the performance of FDSOI nano-transistors. Finally, conclusions are given in Section 4.

## 2 Methodology

The starting point of the simulation framework is a Multi-Subband Ensemble Monte Carlo (MS-EMC) code, which is based on the space-mode approach for quantum transport. The simulator solves the Schrödinger equation in the confinement direction and the Boltzmann Transport Equation (BTE) in the transport plane. The system is coupled by solving Poisson's equation in the 2D simulation domain. This tool has been widely used in different scenarios [4, 5] including the study of other tunneling mechanisms such as source-to-drain tunneling (S/D tunneling)[6] or band-to-band tunneling (BTBT) [7]. The main advantage of our MS-EMC code is that the additional modules needed for taking into account the tunneling processes are included as separate transport mechanisms without increasing the computational time in comparison to purely quantum simulators. Apart from that, they can be activated or deactivated depending on the simulation scenario, giving us the possibility of independently studying GLM.

The noisy nature of the GLM, due to the random number of electrons affected by leakage, is included by implementing it as a stochastic mechanism evaluated for each particle at the end of the Monte Carlo cycle. However, it is necessary to define the input of both physical and simulation parameters before starting the Monte Carlo iterations.

Firstly, the number of traps is deterministically calculated according to the oxide dimensions and the trap density, which in turn depends on the material and

the wafer orientation. The trap density is a particular input parameter in this approximation giving the possibility to the user to vary it in order to consider the fluctuating behavior of this quantity between samples. Secondly, the traps energy below the conduction band and its location along the oxide are chosen considering their random nature by reckoning a uniformly distributed random sequence of numbers. Their energy level is usually between 2.9eV and 3.9eV below the conduction band for the $SiO_2$ oxide [8, 9]. Accordingly, the shift of the conduction band is calculated with the initial conditions and fixed during the whole simulation. Thirdly, it is indispensable to keep in mind that the MS-EMC code makes use of a 2D description, whereas an electron can be trapped only when it is located near a trap location in the oxide with 3D coordinates. Therefore, the dimension of the trap is defined as a cube where the assigned charge is estimated according to the trap density. This percentage $n_{pery}$ will be compared to a random number in the MC iterations, so that it is possible to determine the probability of finding an electron located near the trap.

Then, when the traps are totally defined, the number of particles near the dielectric is required, given that the distance between their location and the interface modifies the tunnel probability. It is of note that the 2D MS-EMC code characterizes the semiclassical motion of the particles in the transport direction ($x$) even though its location in the confinement direction ($z$) is unknown. The simulated particles are distributed along the whole device and hence the percentage of the ones near the interface ($n_{intf}$) with respect to the total number of particles ($n(x, z)$) is estimated:

$$n_{intf} = \frac{\sum_{intf} \sum_x n(x, z)}{\sum_z \sum_x n(x, z)}, \tag{1}$$

where $intf$ represents the region near the interface in the $z$ direction. In this study, $intf$ is taken as 10% of the $T_{Si}$.

The last step required before starting the Monte Carlo iterations is the calculation of the initial tunneling probabilities for each mechanism [10, 11, 9]. In general, the probability for the tunneling processes is calculated using the Wentzel Kramers Brillouin (WKB) approximation [12]. This transmission coefficient depends on the barrier thickness and height (which, in our case, is set up by the band gap of the dielectric material):

$$T_{WKB}(E) = \exp\left\{-\frac{2}{\hbar}\int_a^b \sqrt{2m_z^*(E_{CB}(x, z) - E)}\,\mathrm{d}x\right\}, \tag{2}$$

where $a$ and $b$ are the starting and ending points, $E$ and $m_z^*$ are the energy and the confinement effective mass of the electron, respectively, and $E_{CB}(x, z)$ corresponds to the energy of the conduction band at the point $(x, z)$. Five tunneling processes have been implemented for the GLM in this simulation tool as illustrated in Fig. 1.

Direct tunneling probability is given directly by the WKB approximation. In general, when the ending point is the gate electrode, a Fermi-Dirac distribution of the electrons and available states at any given energy in the gate electrodes

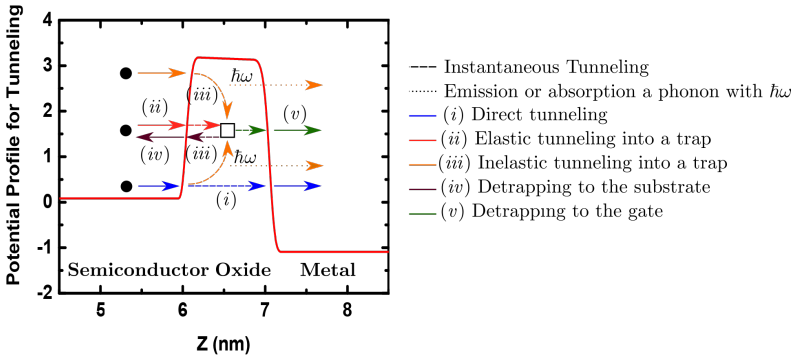



**Fig. 1.** Schematic band diagram of a MOS structure with metal gate and silicon substrate where the transport mechanisms implemented in the MS-EMC simulator are described: (*i*) direct tunneling, (*ii*) elastic tunneling and (*iii*) inelastic tunneling into a trap emitting or capturing a phonon with energy $\hbar\omega$, (*iv*) detrapping to the substrate, and *v*) tunneling from the trap to the gate.

is assumed considering that, after tunneling, the electrons thermalize. For trap assisted tunneling, the probability depends on both the WKB approximation and some specific factors related to each mechanism. In the first place, the trap occupation must obey the Pauli exclusion principle so that no more than two particles as a maximum can be located in a trap. Secondly, if the tunneling is inelastic into a trap, it must emit or absorb a phonon. When the particle energy is higher (resp. lower) than the trap energy, a phonon is emitted (resp. absorbed). Finally, as a result of the doping level of the substrate and the large electric field at the oxide surface, the energy states within the semiconductor substrate are quantized. This leads to less occupied energy states from which electrons can tunnel and so this effect is forbidden if the energy trap state is lower than the first subband. Furthermore, the excess energy is transferred to a phonon via inelastic collisions. When a trapped electron tunnels again to the substrate, a new energy level must be chosen. As the carriers tend to be at the subband with lower kinetic energy, the approximation used in this mechanism calculates the subband in which the electron has lower kinetic energy. More details about how to calculate tunneling probabilities can be found in [10, 11, 9].

When all the initial parameters of the system are introduced, the Monte Carlo iterations begin and so the positions of each electron in the transport direction after a random flight time are calculated. There are two different scenarios regarding the GLM, as determined by the particle location:

– Particle in the channel: The first step is to determine if the particle is located near the substrate-dielectric interface using a uniformly distributed random number $r_{ch1}$. If $r_{ch1} > n_{intf}$, the particle continues with its normal motion, whereas if $r_{ch1} < n_{intf}$ the particle can undergo both DT and TAT or only DT. This choice is made again using another uniformly distributed random number $r_{ch2}$. If $r_{ch2} \leq n_{pery}$, the particle can undergo both DT and

TAT, choosing the selected one from the comparison between the tunneling probabilities and another uniformly distributed random number. Otherwise, the particle undergoes DT through the insulator.
- Particle in a trap: In this scenario, the particle can experience three subscenarios: (i) going back to the substrate (only if the trap energy is higher than the lower subband energy), (ii) leaving the device going to the gate contact, or (iii) remaining in the trap. Due to its random nature, this choice is made again by comparing the tunneling probability with a uniformly distributed random number.

At this point, it is imperative to emphasize some global concepts. All probabilities must be recalculated when the conduction band changes, and when an electron is trapped or detrapped in each Monte Carlo iteration. Apart from that, the charge trapped is dynamically included in the 2D Poisson solution in order to preserve the self-consistency during the simulation time. Moreover, as the GLM has a very low frequency, the particles can only undergo this type of tunneling according to a certain period of occurrence and not after each integration step. Due to the negligible tunneling time through the thin oxide and the low frequency of these tunneling events, it is reasonable to assume that the electron goes directly from the starting point to the ending point at the same time step.

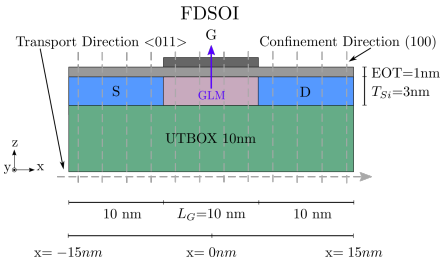
## 3 Results

### 3.1 Description of simulated devices and processes

Device parameters, effective masses and orientation for the FDSOI structure herein analyzed are outlined in Fig. 2. The gate length ranges from $L_G$=7.5nm to $L_G$=20nm, whereas the rest of technological parameters have been fixed: the channel thickness $T_{Si}$ is 3nm, the gate oxide has an Equivalent Oxide Thickness $EOT$=1nm, and the gate work function is 4.385eV. A Back-Plane with a $UTBOX$=10nm, a Back-Bias polarization ($V_{BB}$) of 0V, and a Back-Plane work function of 5.17eV have been chosen. The number of traps is estimated considering the typical trap density for a good quality gate oxide when the dielectric is $SiO_2$ and the wafer orientation is (100). In this particular work, the trap density ($N_{Trap}$) ranges from $10^{11}$cm$^{-2}$ and $10^{13}$cm$^{-2}$.

### 3.2 Results and discussion

An increase in the number of traps or their close location to the interface directly changes the tunneling probability from (to) the substrate to (from) the traps. Fig. 3 shows the number of particles that can experience any GLM for different $N_{Trap}$ values. Let us make several remarks. First, the number of particles that suffers trap assisted tunneling is higher as the trap density increase. Second, the direct tunneling through the oxide is the dominant phenomenon due to the ultra-thin oxide even for the highest $N_{Trap}$. Third, the probability of a trapped




| Valley | $m_x$ | $m_y$ | $m_z$ |
|---|---|---|---|
| $\Delta_2$ | $m_t = 0.198m_0$ | $m_t = 0.198m_0$ | $m_l = 0.916m_0$ |
| $\Delta_4$ | $\frac{2m_l m_t}{m_l+m_t} = 0.326m_0$ | $\frac{m_l+m_t}{2} = 0.557m_0$ | $m_t = 0.198m_0$ |

**Fig. 2.** (left) FDSOI structure analyzed in this work with $L_G$=10nm. 1D Schrödinger equation is solved for each grid point in the transport direction and BTE is solved by the MC method in the transport plane. (right) Effective masses in silicon for the device studied in this work: $m_x$ is the transport mass, $m_z$ is the confinement mass, $m_0$ is the electron free-mass, and the subindex in $\Delta$ represents the corresponding degeneracy factor, where $\Delta_2$ is the most populated valley.

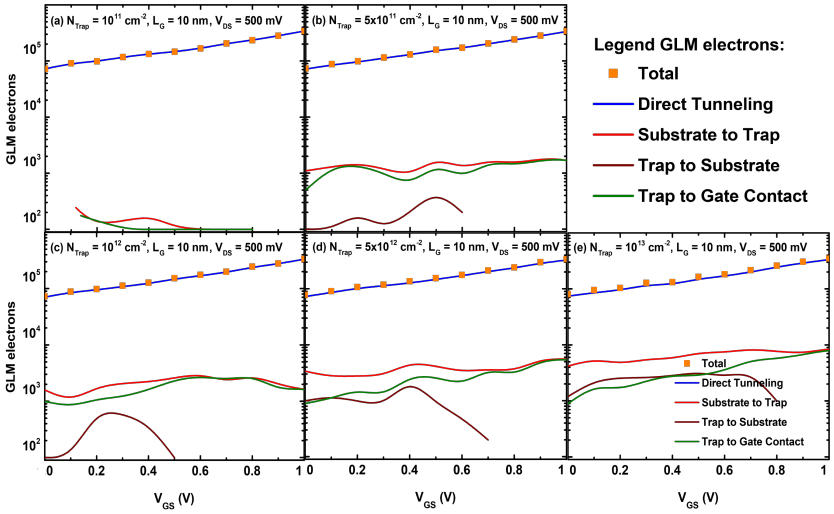



**Fig. 3.** Average number of electrons in arbitrary units affected by the total GLM, and by each individual mechanism as a function of $V_{GS}$ in the 10nm device, where $T_{Si}$=3nm and $V_{DS}$=500mV for different trap densities ($N_{Trap}$): $10^{11}$cm$^{-2}$ (a), 5x$10^{11}$cm$^{-2}$ (b), $10^{12}$cm$^{-2}$ (c), 5x$10^{12}$cm$^{-2}$ (d), and $10^{13}$cm$^{-2}$ (e).

electron to return to the substrate directly depends on the available energy states and so this type of TAT becomes forbidden as the gate bias increases.

In general, and as direct tunneling is the dominant mechanism, the particles that leave the device through the oxide reduce the drain current as depicted in Fig. 4.a. This effect is almost negligible for this particular device because the total number of particles that undergoes any GLM is very reduced. However, the increase of the trap density can modulate the thermionic current as shown in the inset of Fig. 4.a. The charge trapped in the oxide is dynamically included

in the 2D Poisson solution in order to preserve self-consistency during the simulation time. It can be appreciated in Fig. 4.b where the electron distribution is shown along the transport and confinement directions. Accordingly, an increase of the trapped charge reduces the subband levels (Fig. 4.c) causing an enhanced thermionic current.

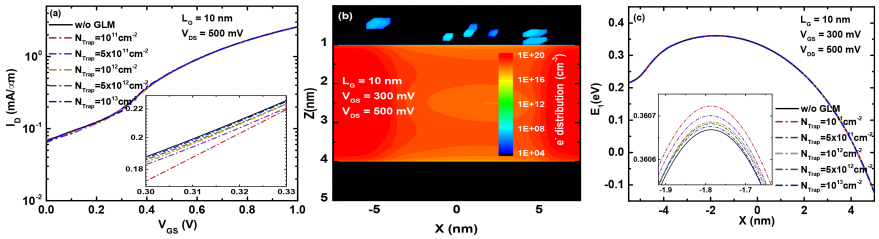



**Fig. 4.** (a) $I_D$ vs. $V_{GS}$ in the 10nm FDSOI device at $V_{DS}$=500mV considering a simulation without GLM and others ones with GLM for different $N_{it}$ values. (b) Electron distribution in $cm^{-3}$ along the transport ($X$) and confinement ($Z$) directions in the same device as in (a) with $V_{GS}$=0.3V and $N_{Trap} = 10^{12}cm^{-2}$. Recall that $X$=0nm corresponds to the center of the device.(c) Energy profiles of the lowest energy subband in the same device as in (a) at $V_{GS}$=0.3V, considering the case without GLM and others ones with GLM for different $N_{Trap}$ values.

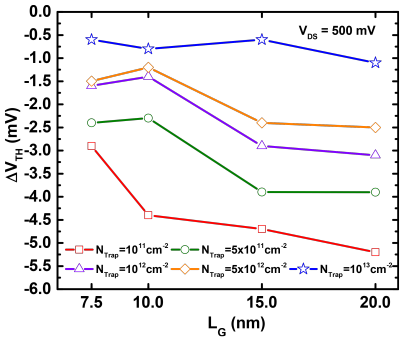



**Fig. 5.** Threshold voltage variation ($\Delta V_{th}$) as a function of $L_G$ calculated as the difference between simulations without and with GLM for different $N_{Trap}$ values, for the FDSOI device at $V_{DS}$=500mV.

The impact of GLM on the threshold voltage variation ($\Delta V_{th}$), as a function of the channel length, is shown in Fig. 5. It has been calculated as the difference between simulations without and with GLM for different $N_{Trap}$ values. This mechanism is more important as the channel length increases because the area in which the particle can undergo GLM is higher and the number of traps increases.

## 4 Conclusions

This work presents the implementation of the gate leakage mechanism (GLM) including direct and trap assisted tunneling in a MS-EMC tool for the study

of how the trap attributes can modify the device electrostatic properties in ultrascaled FDSOI devices. Our calculations show that direct tunneling is the dominant mechanism due to the ultra-thin oxide, resulting in the reduction of the drain current. However, the increase of the trap density slightly decreases the subband levels. Accordingly, this enhances thermionic current in comparison to the case where we only consider direct tunneling.

## References


1. K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage Current Mechanisms and Leakage Reduction Techniques in Deep-Submicrometer CMOS Circuits," *Proceedings of the IEEE*, vol. 91, pp. 305–327, February 2003.
2. Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI devices*. New York: Cambridge University Press, 2009.
3. C. Medina-Bailon, T. Sadi, C. Sampedro, A. Godoy, L. Donetti, V. Georgiev, F. Gámiz, and A. Asenov, "Assessment of Gate Leakage Mechanism utilizing Multi-Subband Ensemble Monte Carlo." 2017 Joint International EUROSOI and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS), 2017.
4. C. Sampedro, F. Gámiz, A. Godoy, R. Valín, A. García-Loureiro, and F. G. Ruiz, "Multi-Subband Monte Carlo study of device orientation effects in ultra-short channel DGSOI," *Solid-State Electronics*, vol. 54, no. 2, pp. 131–136, 2010.
5. C. Sampedro, F. Gámiz, and A. Godoy, "On the extension of ET-FDSOI roadmap for 22 nm node and beyond," *Solid-State Electronics*, vol. 90, pp. 23–27, 2013.
6. C. Medina-Bailon, C. Sampedro, F. Gámiz, A. Godoy, and L. Donetti, "Confinement orientation effects in S / D tunneling," *Solid-State Electronics*, vol. 128, pp. 48 – 53, 2017.
7. C. Medina-Bailon, J. Padilla, C. Sampedro, C. Alper, F. Gámiz, and A. Ionescu, "Implementation of Band–to–Band Tunneling Phenomena in a Multi–Subband Ensemble Monte Carlo Simulator: Application to Silicon TFETs," *IEEE Transactions on Electron Devices*, vol. 64, no. 8, pp. 3084 –3091, 2017.
8. L. Vandelli, S. S. Member, A. Padovani, L. Larcher, R. G. Southwick, W. B. Knowlton, and G. Bersuker, "A Physical Model of the Temperature Dependence of the Current Through SiO 2 / HfO 2 Stacks," *IEEE Transactions on Electron Devices*, vol. 58, no. 9, pp. 2878–2887, 2011.
9. T. Sadi, A. Mehonic, L. Montesi, M. Buckwell, A. Kenyon, and A. Asenov, "Investigation of resistance switching in SiOx RRAM cells using a 3D multi-scale kinetic Monte Carlo simulator," *Journal of Physics: Condensed Matter*, vol. 30, no. 8, pp. 084 005 –, 2018.
10. G. C. Jegert, "Modeling of leakage currents in high–$\kappa$ dielectrics," Ph.D. dissertation, Technische Universität München, München, March 2012.
11. T. Sadi, L. Wang, D. Gao, A. Mehonic, L. Montesi, M. Buckwell, A. Kenyon, A. Shluger, and A. Asenov, "Advanced physical modeling of SiOx resistive random access memories." 2016 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), 2016, pp. 149–152.
12. D. J. Griffiths, "The WKB approximation," in *Introduction to Quantum Mechanics*. New Jersey: Prentice Hall, 1995, ch. 8, pp. 274–297.

# Impact of the Trap Attributes on the Gate Leakage Mechanisms in a 2D MS-EMC Nanodevice Simulator*

C. Medina-Bailon[1,2], T. Sadi[3], C. Sampedro[2], J.L. Padilla[2], L. Donetti[2], V. Georgiev[1], F. Gamiz[2], and A. Asenov[1]

[1] School of Engineering, University of Glasgow, Glasgow G12 8LT, Scotland, UK
Cristina.MedinaBailon@glasgow.ac.uk
[2] Nanoelectronics Research Group, Universidad de Granada, 18071 Granada, Spain
[3] Department of Neuroscience and Biomedical Engineering, Aalto University, P.O. Box 12200, FI-00076 Aalto, Finland

**Abstract.** From a modeling point of view, the inclusion of adequate physical phenomena is mandatory when analyzing the behavior of new transistor architectures. In particular, the high electric field across the ultra-thin insulator in aggressively scaled transistors leads to the possibility for the charge carriers in the channel to tunnel through the gate oxide via various gate leakage mechanisms (GLMs). In this work, we study the impact of trap number on gate leakage using the GLM model, which is included in a Multi-Subband Ensemble Monte Carlo (MS-EMC) simulator for Fully-Depleted Silicon-On-Insulator (FDSOI) field effect transistors (FETs). The GLM code described herein considers both direct and trap-assisted tunneling. This work shows that trap attributes and dynamics can modify the device electrostatic characteristics and even play a significant role in determining the extent of GLMs.

**Keywords:** gate leakage mechanism · direct tunneling · trap assisted tunneling · MS-EMC · FDSOI.

## 1 Introduction

Reducing the gate oxide thickness implies an increase in the field across the oxide. The high electric field coupled with thin oxides leads to the possibility of charge carriers traversing the barrier for transport set up by the dielectric layer, resulting in tunneling processes from (to) substrate to (from) gate through the gate oxide [1, 2], and thus giving rise to a certain gate current. This effect is known as the gate leakage mechanism (GLM) and includes both direct and trap-assisted tunneling.

The direct tunneling (DT) processes are always present, even if a dielectric film of perfect quality is assumed. In the case where a very thin oxide layer (less

* The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688101 SUPERAID7.

than $3-4$nm) is considered, electrons forming the inversion layer can tunnel to the gate through the energetically forbidden band gap of the dielectric material. Similarly, the trap assistant tunneling (TAT) processes are related to the existence of defect states which cause elastic or inelastic tunneling of electrons into and out of defects. In general, direct tunneling is the dominant phenomenon due to the small oxide thickness [3]. Nevertheless, trap attributes and dynamics can modify the device electrostatic properties.

Apart from the analysis of these tunneling mechanisms, the investigation of these processes on new technological nanodevices is mandatory. Currently, Fully-Depleted Silicon-On-Insulator (FDSOI) devices have been recognized as an alternative to bulk devices. However, the impact of GLM mainly depends on the electron confinement near the interface [3]. Accordingly, the number of electrons tunneling through the oxide is higher in the single gate FDSOI in contrast with other double gate devices, such as the vertical FinFET. Therefore, the study of this mechanism in FDSOI devices is of special interest.

The aim of this work is to perform a study of the impact of the trap attributes on the GLM and thus on the FDSOI performance. For this purpose, a detailed discussion of this transport mechanism is given together with the details of the stochastic simulation process in Section 2. The main findings are reported in Section 3 including a meticulous analysis of how the trap density modifies the GLM and the performance of FDSOI nano-transistors. Finally, conclusions are given in Section 4.

## 2 Methodology

The starting point of the simulation framework is a Multi-Subband Ensemble Monte Carlo (MS-EMC) code, which is based on the space-mode approach for quantum transport. The simulator solves the Schrödinger equation in the confinement direction and the Boltzmann Transport Equation (BTE) in the transport plane. The system is coupled by solving Poisson's equation in the 2D simulation domain. This tool has been widely used in different scenarios [4, 5] including the study of other tunneling mechanisms such as source-to-drain tunneling (S/D tunneling)[6] or band-to-band tunneling (BTBT) [7]. The main advantage of our MS-EMC code is that the additional modules needed for taking into account the tunneling processes are included as separate transport mechanisms without increasing the computational time in comparison to purely quantum simulators. Apart from that, they can be activated or deactivated depending on the simulation scenario, giving us the possibility of independently studying GLM.

The noisy nature of the GLM, due to the random number of electrons affected by leakage, is included by implementing it as a stochastic mechanism evaluated for each particle at the end of the Monte Carlo cycle. However, it is necessary to define the input of both physical and simulation parameters before starting the Monte Carlo iterations.

Firstly, the number of traps is deterministically calculated according to the oxide dimensions and the trap density, which in turn depends on the material and

the wafer orientation. The trap density is a particular input parameter in this approximation giving the possibility to the user to vary it in order to consider the fluctuating behavior of this quantity between samples. Secondly, the traps energy below the conduction band and its location along the oxide are chosen considering their random nature by reckoning a uniformly distributed random sequence of numbers. Their energy level is usually between 2.9eV and 3.9eV below the conduction band for the $SiO_2$ oxide [8, 9]. Accordingly, the shift of the conduction band is calculated with the initial conditions and fixed during the whole simulation. Thirdly, it is indispensable to keep in mind that the MS-EMC code makes use of a 2D description, whereas an electron can be trapped only when it is located near a trap location in the oxide with 3D coordinates. Therefore, the dimension of the trap is defined as a cube where the assigned charge is estimated according to the trap density. This percentage $n_{pery}$ will be compared to a random number in the MC iterations, so that it is possible to determine the probability of finding an electron located near the trap.

Then, when the traps are totally defined, the number of particles near the dielectric is required, given that the distance between their location and the interface modifies the tunnel probability. It is of note that the 2D MS-EMC code characterizes the semiclassical motion of the particles in the transport direction ($x$) even though its location in the confinement direction ($z$) is unknown. The simulated particles are distributed along the whole device and hence the percentage of the ones near the interface ($n_{intf}$) with respect to the total number of particles ($n(x, z)$) is estimated:

$$n_{intf} = \frac{\sum_{intf} \sum_x n(x, z)}{\sum_z \sum_x n(x, z)}, \tag{1}$$

where $intf$ represents the region near the interface in the $z$ direction. In this study, $intf$ is taken as 10% of the $T_{Si}$.

The last step required before starting the Monte Carlo iterations is the calculation of the initial tunneling probabilities for each mechanism [10, 11, 9]. In general, the probability for the tunneling processes is calculated using the Wentzel Kramers Brillouin (WKB) approximation [12]. This transmission coefficient depends on the barrier thickness and height (which, in our case, is set up by the band gap of the dielectric material):

$$T_{WKB}(E) = \exp\left\{-\frac{2}{\hbar}\int_a^b \sqrt{2m_z^*(E_{CB}(x, z) - E)}\,\mathrm{d}x\right\}, \tag{2}$$

where $a$ and $b$ are the starting and ending points, $E$ and $m_z^*$ are the energy and the confinement effective mass of the electron, respectively, and $E_{CB}(x, z)$ corresponds to the energy of the conduction band at the point $(x, z)$. Five tunneling processes have been implemented for the GLM in this simulation tool as illustrated in Fig. 1.

Direct tunneling probability is given directly by the WKB approximation. In general, when the ending point is the gate electrode, a Fermi-Dirac distribution of the electrons and available states at any given energy in the gate electrodes

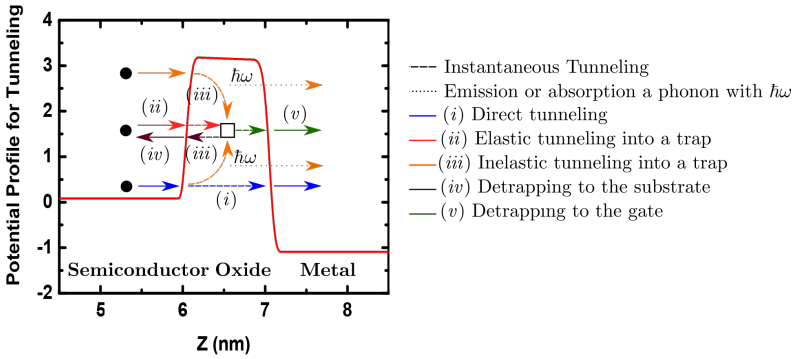



**Fig. 1.** Schematic band diagram of a MOS structure with metal gate and silicon substrate where the transport mechanisms implemented in the MS-EMC simulator are described: (*i*) direct tunneling, (*ii*) elastic tunneling and (*iii*) inelastic tunneling into a trap emitting or capturing a phonon with energy $\hbar\omega$, (*iv*) detrapping to the substrate, and *v*) tunneling from the trap to the gate.

is assumed considering that, after tunneling, the electrons thermalize. For trap assisted tunneling, the probability depends on both the WKB approximation and some specific factors related to each mechanism. In the first place, the trap occupation must obey the Pauli exclusion principle so that no more than two particles as a maximum can be located in a trap. Secondly, if the tunneling is inelastic into a trap, it must emit or absorb a phonon. When the particle energy is higher (resp. lower) than the trap energy, a phonon is emitted (resp. absorbed). Finally, as a result of the doping level of the substrate and the large electric field at the oxide surface, the energy states within the semiconductor substrate are quantized. This leads to less occupied energy states from which electrons can tunnel and so this effect is forbidden if the energy trap state is lower than the first subband. Furthermore, the excess energy is transferred to a phonon via inelastic collisions. When a trapped electron tunnels again to the substrate, a new energy level must be chosen. As the carriers tend to be at the subband with lower kinetic energy, the approximation used in this mechanism calculates the subband in which the electron has lower kinetic energy. More details about how to calculate tunneling probabilities can be found in [10, 11, 9].

When all the initial parameters of the system are introduced, the Monte Carlo iterations begin and so the positions of each electron in the transport direction after a random flight time are calculated. There are two different scenarios regarding the GLM, as determined by the particle location:

– Particle in the channel: The first step is to determine if the particle is located near the substrate-dielectric interface using a uniformly distributed random number $r_{ch1}$. If $r_{ch1} > n_{intf}$, the particle continues with its normal motion, whereas if $r_{ch1} < n_{intf}$ the particle can undergo both DT and TAT or only DT. This choice is made again using another uniformly distributed random number $r_{ch2}$. If $r_{ch2} \leq n_{pery}$, the particle can undergo both DT and

TAT, choosing the selected one from the comparison between the tunneling probabilities and another uniformly distributed random number. Otherwise, the particle undergoes DT through the insulator.
- Particle in a trap: In this scenario, the particle can experience three subscenarios: (i) going back to the substrate (only if the trap energy is higher than the lower subband energy), (ii) leaving the device going to the gate contact, or (iii) remaining in the trap. Due to its random nature, this choice is made again by comparing the tunneling probability with a uniformly distributed random number.

At this point, it is imperative to emphasize some global concepts. All probabilities must be recalculated when the conduction band changes, and when an electron is trapped or detrapped in each Monte Carlo iteration. Apart from that, the charge trapped is dynamically included in the 2D Poisson solution in order to preserve the self-consistency during the simulation time. Moreover, as the GLM has a very low frequency, the particles can only undergo this type of tunneling according to a certain period of occurrence and not after each integration step. Due to the negligible tunneling time through the thin oxide and the low frequency of these tunneling events, it is reasonable to assume that the electron goes directly from the starting point to the ending point at the same time step.

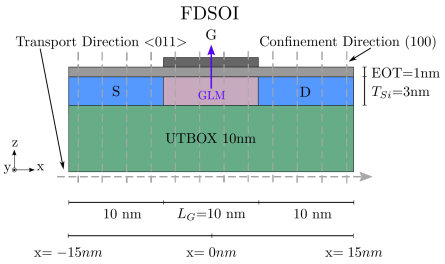
## 3 Results

### 3.1 Description of simulated devices and processes

Device parameters, effective masses and orientation for the FDSOI structure herein analyzed are outlined in Fig. 2. The gate length ranges from $L_G$=7.5nm to $L_G$=20nm, whereas the rest of technological parameters have been fixed: the channel thickness $T_{Si}$ is 3nm, the gate oxide has an Equivalent Oxide Thickness $EOT$=1nm, and the gate work function is 4.385eV. A Back-Plane with a $UTBOX$=10nm, a Back-Bias polarization ($V_{BB}$) of 0V, and a Back-Plane work function of 5.17eV have been chosen. The number of traps is estimated considering the typical trap density for a good quality gate oxide when the dielectric is $SiO_2$ and the wafer orientation is (100). In this particular work, the trap density ($N_{Trap}$) ranges from $10^{11}$cm$^{-2}$ and $10^{13}$cm$^{-2}$.

### 3.2 Results and discussion

An increase in the number of traps or their close location to the interface directly changes the tunneling probability from (to) the substrate to (from) the traps. Fig. 3 shows the number of particles that can experience any GLM for different $N_{Trap}$ values. Let us make several remarks. First, the number of particles that suffers trap assisted tunneling is higher as the trap density increase. Second, the direct tunneling through the oxide is the dominant phenomenon due to the ultra-thin oxide even for the highest $N_{Trap}$. Third, the probability of a trapped




| Valley | $m_x$ | $m_y$ | $m_z$ |
|---|---|---|---|
| $\Delta_2$ | $m_t = 0.198m_0$ | $m_t = 0.198m_0$ | $m_l = 0.916m_0$ |
| $\Delta_4$ | $\frac{2m_l m_t}{m_l+m_t} = 0.326m_0$ | $\frac{m_l+m_t}{2} = 0.557m_0$ | $m_t = 0.198m_0$ |

**Fig. 2.** (left) FDSOI structure analyzed in this work with $L_G$=10nm. 1D Schrödinger equation is solved for each grid point in the transport direction and BTE is solved by the MC method in the transport plane. (right) Effective masses in silicon for the device studied in this work: $m_x$ is the transport mass, $m_z$ is the confinement mass, $m_0$ is the electron free-mass, and the subindex in $\Delta$ represents the corresponding degeneracy factor, where $\Delta_2$ is the most populated valley.

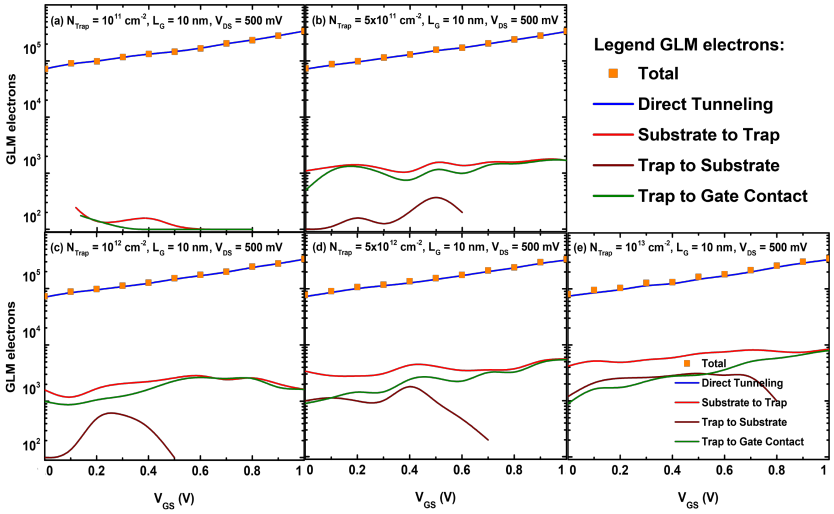



**Fig. 3.** Average number of electrons in arbitrary units affected by the total GLM, and by each individual mechanism as a function of $V_{GS}$ in the 10nm device, where $T_{Si}$=3nm and $V_{DS}$=500mV for different trap densities ($N_{Trap}$): $10^{11}$cm$^{-2}$ (a), 5x$10^{11}$cm$^{-2}$ (b), $10^{12}$cm$^{-2}$ (c), 5x$10^{12}$cm$^{-2}$ (d), and $10^{13}$cm$^{-2}$ (e).

electron to return to the substrate directly depends on the available energy states and so this type of TAT becomes forbidden as the gate bias increases.

In general, and as direct tunneling is the dominant mechanism, the particles that leave the device through the oxide reduce the drain current as depicted in Fig. 4.a. This effect is almost negligible for this particular device because the total number of particles that undergoes any GLM is very reduced. However, the increase of the trap density can modulate the thermionic current as shown in the inset of Fig. 4.a. The charge trapped in the oxide is dynamically included

in the 2D Poisson solution in order to preserve self-consistency during the simulation time. It can be appreciated in Fig. 4.b where the electron distribution is shown along the transport and confinement directions. Accordingly, an increase of the trapped charge reduces the subband levels (Fig. 4.c) causing an enhanced thermionic current.

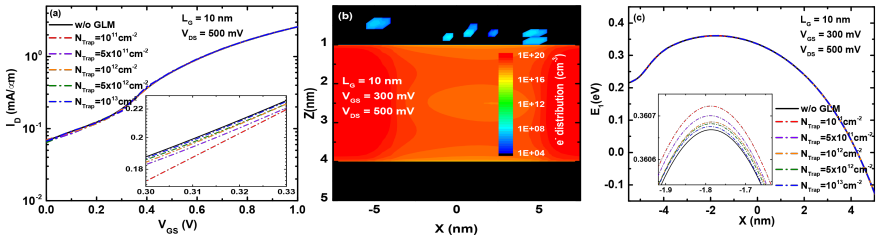



**Fig. 4.** (a) $I_D$ vs. $V_{GS}$ in the 10nm FDSOI device at $V_{DS}$=500mV considering a simulation without GLM and others ones with GLM for different $N_{it}$ values. (b) Electron distribution in $cm^{-3}$ along the transport ($X$) and confinement ($Z$) directions in the same device as in (a) with $V_{GS}$=0.3V and $N_{Trap} = 10^{12}cm^{-2}$. Recall that $X$=0nm corresponds to the center of the device.(c) Energy profiles of the lowest energy subband in the same device as in (a) at $V_{GS}$=0.3V, considering the case without GLM and others ones with GLM for different $N_{Trap}$ values.

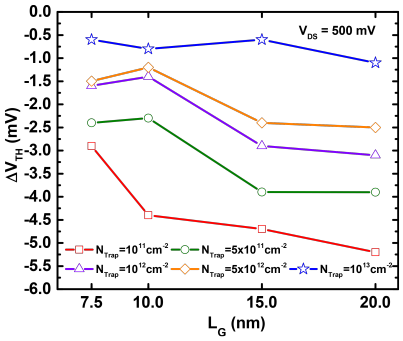



**Fig. 5.** Threshold voltage variation ($\Delta V_{th}$) as a function of $L_G$ calculated as the difference between simulations without and with GLM for different $N_{Trap}$ values, for the FDSOI device at $V_{DS}$=500mV.

The impact of GLM on the threshold voltage variation ($\Delta V_{th}$), as a function of the channel length, is shown in Fig. 5. It has been calculated as the difference between simulations without and with GLM for different $N_{Trap}$ values. This mechanism is more important as the channel length increases because the area in which the particle can undergo GLM is higher and the number of traps increases.

## 4 Conclusions

This work presents the implementation of the gate leakage mechanism (GLM) including direct and trap assisted tunneling in a MS-EMC tool for the study

of how the trap attributes can modify the device electrostatic properties in ultrascaled FDSOI devices. Our calculations show that direct tunneling is the dominant mechanism due to the ultra-thin oxide, resulting in the reduction of the drain current. However, the increase of the trap density slightly decreases the subband levels. Accordingly, this enhances thermionic current in comparison to the case where we only consider direct tunneling.

## References


1. K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage Current Mechanisms and Leakage Reduction Techniques in Deep-Submicrometer CMOS Circuits," *Proceedings of the IEEE*, vol. 91, pp. 305–327, February 2003.
2. Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI devices*. New York: Cambridge University Press, 2009.
3. C. Medina-Bailon, T. Sadi, C. Sampedro, A. Godoy, L. Donetti, V. Georgiev, F. Gámiz, and A. Asenov, "Assessment of Gate Leakage Mechanism utilizing Multi-Subband Ensemble Monte Carlo." 2017 Joint International EUROSOI and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS), 2017.
4. C. Sampedro, F. Gámiz, A. Godoy, R. Valín, A. García-Loureiro, and F. G. Ruiz, "Multi-Subband Monte Carlo study of device orientation effects in ultra-short channel DGSOI," *Solid-State Electronics*, vol. 54, no. 2, pp. 131–136, 2010.
5. C. Sampedro, F. Gámiz, and A. Godoy, "On the extension of ET-FDSOI roadmap for 22 nm node and beyond," *Solid-State Electronics*, vol. 90, pp. 23–27, 2013.
6. C. Medina-Bailon, C. Sampedro, F. Gámiz, A. Godoy, and L. Donetti, "Confinement orientation effects in S / D tunneling," *Solid-State Electronics*, vol. 128, pp. 48 – 53, 2017.
7. C. Medina-Bailon, J. Padilla, C. Sampedro, C. Alper, F. Gámiz, and A. Ionescu, "Implementation of Band–to–Band Tunneling Phenomena in a Multi–Subband Ensemble Monte Carlo Simulator: Application to Silicon TFETs," *IEEE Transactions on Electron Devices*, vol. 64, no. 8, pp. 3084 –3091, 2017.
8. L. Vandelli, S. S. Member, A. Padovani, L. Larcher, R. G. Southwick, W. B. Knowlton, and G. Bersuker, "A Physical Model of the Temperature Dependence of the Current Through SiO 2 / HfO 2 Stacks," *IEEE Transactions on Electron Devices*, vol. 58, no. 9, pp. 2878–2887, 2011.
9. T. Sadi, A. Mehonic, L. Montesi, M. Buckwell, A. Kenyon, and A. Asenov, "Investigation of resistance switching in SiOx RRAM cells using a 3D multi-scale kinetic Monte Carlo simulator," *Journal of Physics: Condensed Matter*, vol. 30, no. 8, pp. 084 005 –, 2018.
10. G. C. Jegert, "Modeling of leakage currents in high–$\kappa$ dielectrics," Ph.D. dissertation, Technische Universität München, München, March 2012.
11. T. Sadi, L. Wang, D. Gao, A. Mehonic, L. Montesi, M. Buckwell, A. Kenyon, A. Shluger, and A. Asenov, "Advanced physical modeling of SiOx resistive random access memories." 2016 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), 2016, pp. 149–152.
12. D. J. Griffiths, "The WKB approximation," in *Introduction to Quantum Mechanics*. New Jersey: Prentice Hall, 1995, ch. 8, pp. 274–297.