Study of gate current in advanced MOS architectures

Ghulam Ali Gauhar[1], Abhishek Chenchety[1], Hashish Yenugula[1], Vihar Georgiev[2], Asen Asenov[2], Oves Badami[1,*]

Abstract

We have carried out a comprehensive study of the gate current (I_G) in advanced MOS architectures for different gate lengths and cross-section areas using an in-house simulation tool. We have considered only direct tunneling under the assumption that trap concentration and therefore the trap assisted current would be small in a matured technology. We have also studied the impact of the interfacial (IL) SiO_2 layer on the gate current in the high-κ gate stack. Our results suggest that IL leads to an increase in the gate current for equivalent EOT. They also highlight that reduction in the cross-section area leads to a significant increase in the I_G.

Keywords: Gate current, Tunneling, WKB Approximation

2010 MSC: 00-01, 99-00

1. Introduction

MOSFETs have been constantly scaled for more than five decades to improve the performance and increase the packing density of the microelectronic circuits. In this journey of scaling, the channel cross-section area and the gate oxide thickness have been scaled to reduce the short channel effects. To this end, several innovative device architectures like the double gate, FinFET, stacked nanosheets, have been implemented or are under investigation. These architec-

*Corresponding author

Email address: `oves.badami@ee.iith.ac.in` (Oves Badami)

[1]Dept. of Electrical Engg., Indian Institute of Technology Hyderabd, Telangana, India.

[2]James Watt School of Engineering, University of Glasgow, Scotland, UK.

tures have been extensively analyzed and optimized from the point of view of improving the electrostatic integrity and the on-state performance, and reducing the variability [1], [2]. [3], [4].

In continuing the scaling of MOS transistors, the channel length will be scaled in concert with the equivalent oxide thickness (EOT) and the cross-section area to maintain gate control over the channel region. The increase in the electron confinement (which is different in different architectures) with the reduction in the cross-section will lead to an increase in the subband energy levels, in turn, reducing the tunneling barrier height [5]. Thus, the reduction in the gate oxide thickness and cross-section area are expected to increase the gate current. However, a comprehensive comparison of the different transistor architectures from the gate leakage current perspective is missing in the literature. Hence, in this work, we compare the gate current (I_G) of transistors with different architectures and cross-sectional areas.

2. Methodology

To correctly capture the electrostatics in the modern-day MOS transistors, we have developed a 2D Finite Element method [6] based Schrodinger-Poisson solver along with equilibrium Fermi-Dirac statistics. Several different methodologies have been employed in the literature to compute the gate current. These methodologies range from full quantum transport formalism like Transfer Matrix method [7], Quantum Transmitting Boundary method [8] to semi-classical approaches [5], [9], [10]. In this work, we have used the semi-classical approach as it has been demonstrated in Ref. [10] that it gives correct life-times of the carriers in the semiconductor.

In 2D confined structures, the electrons in a subband (electron density per unit length) can leak out of the cross-section through multiple paths as shown in Fig. 1(a). Each tunneling path would have its own **escape rate**. Assuming that the electrons sequentially tunnel from the semiconductor into the gate terminal and tunneling events are independent, we consider an average value of the carrier

escape rate across the oxide-semiconductor interface length to account for multiple tunneling paths. Mathematically, I_G and the average carrier life ($\overline{\tau_i}$) time can be written as

$$I_G = -q\Sigma_i \frac{n_i}{\overline{\tau_i}} \ [A/m] \tag{1}$$

$$\frac{1}{\overline{\tau_i}} = \frac{\int_S \tau_i^{-1}(S)dS}{\int_S dS} = \frac{1}{W}\int_S \frac{T(S,\varepsilon_i)dS}{\int_{\eta_l}^{\eta_r} \sqrt{2m_\nu/(\varepsilon_i - E_C(\eta))}} \ [1/s] \tag{2}$$

where n_i, is the 1D of electron density in the ith subband with energy ε_i, $T(S,E)$ is the transmission coefficient calculated using WKB formulation at location S along the interface, $\overline{\tau_i}$ is escape rate averaged over the oxide-semiconductor interface length ($\int_S dS = W$). η_l and η_r are the classical turning points as illustrated in Fig. 1(b). For planar geometries, eq. 1 and eq. 2 reduces to the formulation discussed in [9], [10]. The gate to substrate current was calculated as discussed in [9]. In this work, I_G was evaluated after a self-consistent Schrodinger-Poisson solution. Hence, the impact of the electron tunneling out of the device on the device electrostatics was neglected.

For appropriate comparison between different architectures, the threshold voltage (V_{TH}) was set to 0.3 V in all simulations, and the cross-section area was assumed to be the same. The voltage at which the inversion charge density is equal to the doping concentration was set as the V_{TH}. The metal gate workfunction was modified to adjust the V_{TH}. The supply voltage was assumed to be constant ($V_{DD} = 0.6\ V$) in all simulations. We assume that in matured technology, the gate oxide would be of good quality (negligible number of traps) and this will lead to negligible trap assisted tunneling current.

3. Results and Discussion

Fig. 2 shows the comparison between the simulations and the experimental data [11]. The simulated I_G-V_G shows a good match with the experimental data. For comparison, we considered a large square cross-section nanowire with the side of L = 60 nm to mimic the bulk nature of the experimental devices. Tab. 1 lists the gate oxide (SiO_2) parameters. These parameters agree well

with those employed in [11]. The table also lists important HfO_2 parameters used in the simulations. Fig. 3 shows the schematics of the cross-sections employed in this work. Fig. 4 compares the I_G for different architectures with HfO_2 as the gate oxide. The natural length (Λ) is 7.5 times the L_G to calculate the gate oxide thickness (t_{ox}) [4]. This was done to minimize the short channel effects and keep them comparable across different architectures. The channel was p-doped with $N_A = 10^{16}$ cm^{-3}). The cross-section area is taken to be 25 nm^2 as any further reduction in the cross-section area leads to a significant degradation in the mobility [1]. The results highlight that scaling the channel length (and consequently scaling the gate oxide to maintain the electrostatic integrity) leads to a significant increase in the I_G. This increase in I_G is enhanced in circular GAA transistors because of the exponential relationship between the oxide thickness and Λ^2 [4]. Fig.5 shows the gate current comparison in a square cross-section device with and without SiO_2 IL (with the same EOT, **the physical oxide thickness are different in the two cases**). **The gate current is directly proportional to the the transmission coefficient ($T(S,E) = exp\left(-\frac{2}{\hbar}\int_0^{t_{ox}}\sqrt{m_{ox}(\eta)(E_C(\eta,S)-E)}d\eta\right)$) from eq. 1. Due to the reduction in the effective tunneling barrier ($-\frac{2}{\hbar}\int_0^{t_{ox}}\sqrt{m_{ox}(\eta)(E_C(\eta,S)-E)}d\eta$), an increase in I_G is seen in devices with IL. As IL has lower permittivity and hence leads to reduction in the physical thickness. The impact of reduction in the physical thickness more than that of the increased barrier height of the IL. This leads to reduction in the gate current**. Fig.6 shows a significant increase in the I_G for circular GAA cross-section with the reduction in the cross-section area. With the reduction in the cross-section area, the electron density (n_{1D}) reduces, and the tunneling rate increases because of the reduction in the tunneling barrier which leads to an increase in the average escape rate. The increase in the average escape rate ($1/\overline{\tau_i}$) is much faster than the reduction in the n_{1D} and hence the I_G increases.

4. Conclusion

We have extended the existing I_G simulation formalism and benchmarked the methodology with the experimental data. We have compared different transistor architectures from I_G point of view. Our results indicate that for $L_G < 14$ nm circular GAA has the lower I_G while $L_G > 14$ nm square GAA has the edge at a fixed cross-section area and overdrive voltage. The simulations suggest that removing or reducing the IL SiO_2 would help in reducing the gate current (due to direct tunneling). Our analysis also highlights that reduction in the cross-section area can be detrimental from the I_G perspective even though it will lead to better electrostatic integrity.

Parameter [units]	SiO_2	HfO_2
Tunneling mass, m_T [m_0]	0.45	0.18
ϕ_B [eV]	2.9	1.7
ε_r [ε_0]	6.6	21

Table 1: Gate oxide parameters that were employed in the simulation. SiO_2 parameters are the same as used in the benchmarking simulations of Fig.2. HfO_2 parameters are taken from [11].

References

[1] C. Medina-Bailon, T. Sadi, M. Nedjalkov, H. Carrillo-Nuñez, J. Lee, O. Badami, V. Georgiev, S. Selberherr, A. Asenov, Mobility of circular and elliptical Si nanowire transistors using a multi-subband 1D formalism, IEEE Electron Device Letters 40 (10) (2019) 1571–1574. doi:10.1109/LED.2019.2934349.

[2] O. Badami, D. Lizzit, F. Driussi, P. Palestri, D. Esseni, Benchmarking of 3-D MOSFET Architectures: Focus on the Impact of Surface Roughness and Self-Heating, IEEE Transactions on Electron Devices 65 (9) (2018) 3646–3653. doi:10.1109/TED.2018.2857509.

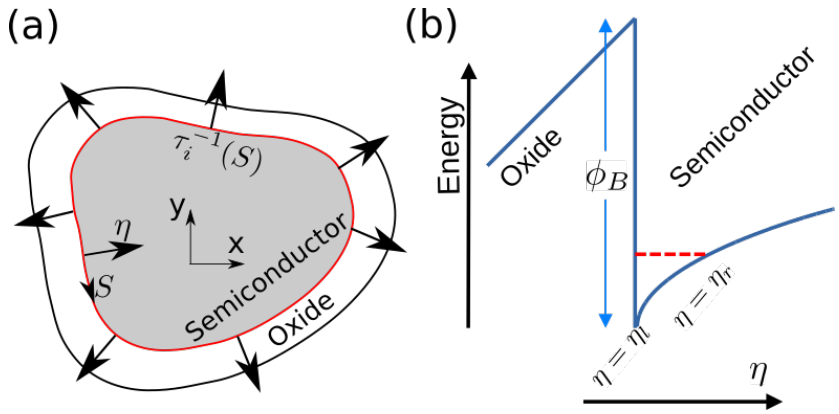

Figure 1: (a)Illustration of multiple paths through which 1D electron density in the 2D confined cross-section can leak from the channel to gate electrode. $\tau_i^{-1}(S)$ is the escape rate for the 1D electron density in the ith subband from location S along the oxide semiconductor interface. (b) Schematic of the conduction band edge normal to the oxide semiconductor interface showing the classical turning points η_l and η_r for a subband (dashed line)

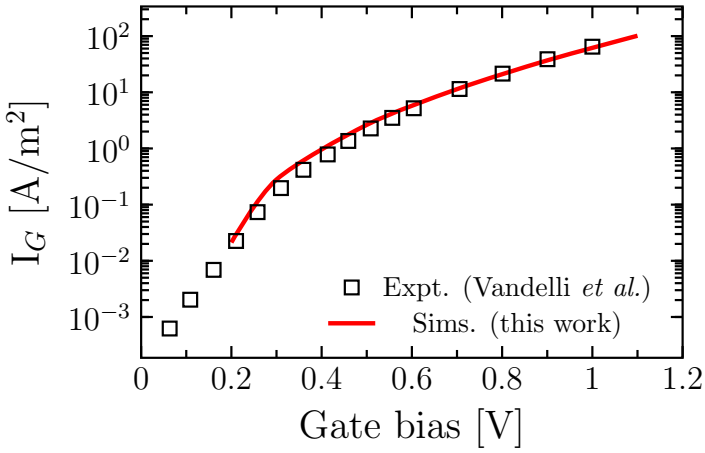

Figure 2: Comparison between the measured gate current and simulations. Experimental data are from Ref.[11]. I_G (A/m) was normalized by the perimeter of the oxide-semiconductor interface. The parameters for gate oxide (SiO_2) used in the simulations are listed in Table I.

[3] A. Sudarsanan, O. Badami, K. Nayak, Superior interface trap variability immunity of horizontally stacked si nanosheet fet in Sub-3-nm technology node, International Semiconductor Conference (2021) 161–164.

[4] J. P. Colinge, "FinFETs and Other Multi-Gate Transistors", 1st Edition,

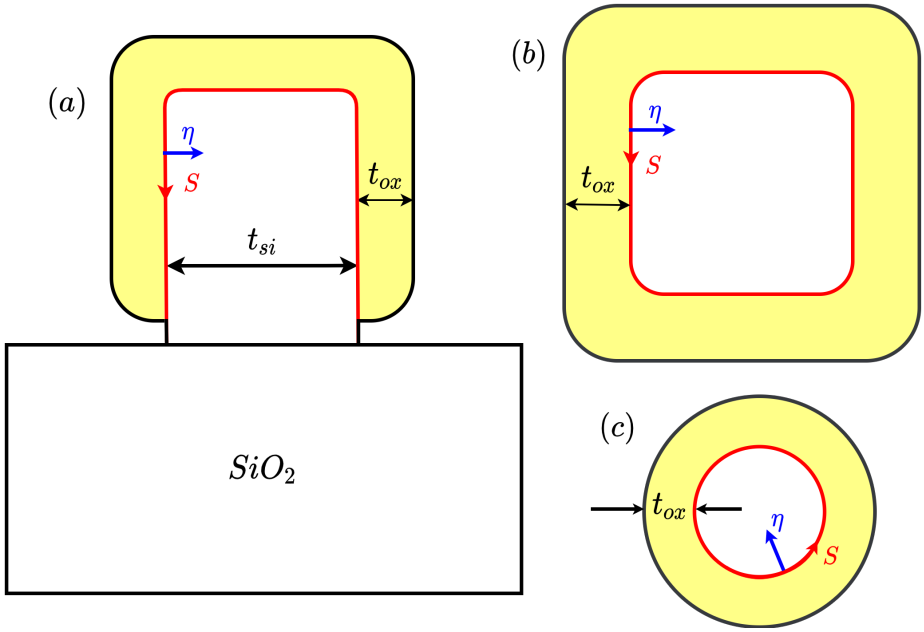

Figure 3: Schematics of the structures used in this work (a)SOI-trigated FinFET (b) Square GAA (c) Circular GAA. For trigated FinFET the aspect ratio was set to 1:1. The radius of curvature at the corners was taken to be 0.5 nm.

Springer, 2007.

[5] M. V. Fischetti, B. Fu, W. G. Vandenberghe, Theoretical study of the gate leakage current in Sub-10-nm field-effect transistors, IEEE Transactions on Electron Devices 60 (11) (2013) 3862–3869. doi:10.1109/TED.2013.2280844.

[6] M. G. Larson, F. Bengzon, "The Finite Element Method: Theory, Implementation, and Applications", 1st Edition, Springer, Berlin, Heidelberg, 2013. doi:https://doi.org/10.1007/978-3-642-33287-6.

[7] A. Gehring, T. Grasser, B.-H. Cheong, S. Selberherr, Design optimization of multi-barrier tunneling devices using the transfer-matrix method, Solid-State Electronics 46 (10) (2002) 1545–1551. doi:https://doi.org/10.1016/S0038-1101(02)00103-X.

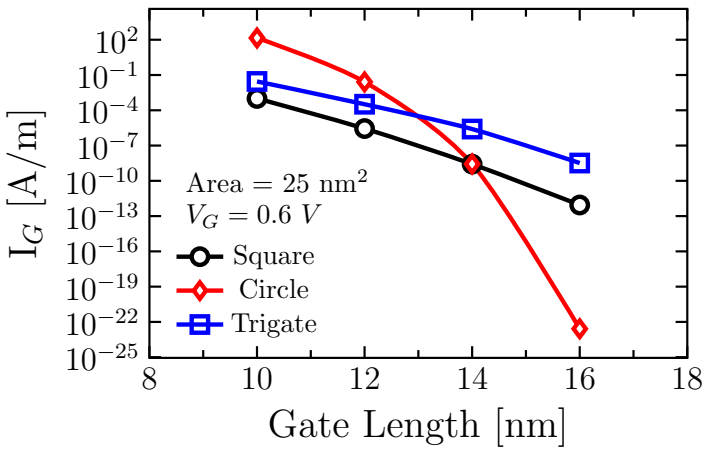

Figure 4: Comparison of the gate current in different architectures. The results indicate that the square GAA has an advantage over other architectures for $L_G > 14\ nm$ in terms of the gate current. t_{ox} for different channel lengths was calculated from the natural length for different architectures [4]. The cross-section area was assumed to be 25 nm^2.

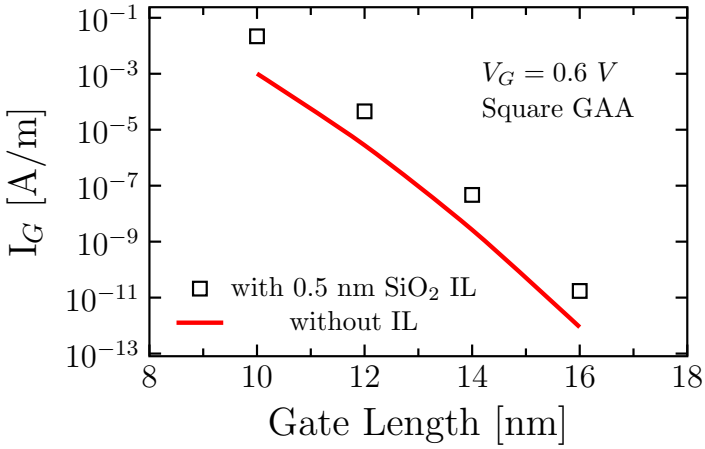

Figure 5: Comparison of the gate current with and without SiO_2 IL ($t_{IL} = 0.5\ nm$). The simulations suggest that using the IL leads to an increase in the gate current. For fixed L_G, the EOT in both the structures was kept the same.

[8] C. S. Lent, D. J. Kirkner, The quantum transmitting boundary method, Journal of Applied Physics 67 (10) (1990) 6353–6359. doi:10.1063/1.345156.

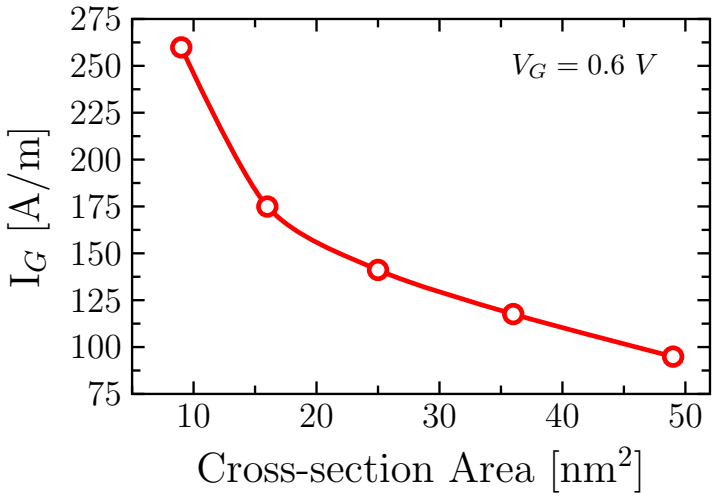

Figure 6: Variation in the gate current with change in the cross-section area of a circular GAA. The increase in I_G is due to increase in $1/\overline{\tau_i}$ of the electrons due to the increase in the ε_i.

[9] F. Rana, S. Tiwari, D. A. Buchanan, Self-consistent modeling of accumulation layers and tunneling currents through very thin oxides, Applied Physics Letters 69 (8) (1996) 1104–1106. doi:10.1063/1.117072.

[10] A. D. Serra, A. Abramo, P. Palestri, L. Selmi, F. Widdershoven, Closed- and open-boundary models for gate-current calculation in n-MOSFETs, IEEE Transactions on Electron Devices 48 (8) (2001) 1811–1815. doi:10.1109/16.936711.

[11] L. Vandelli, A. Padovani, L. Larcher, R. G. Southwick, W. B. Knowlton, G. Bersuker, A physical model of the temperature dependence of the current through SiO_2/HfO_2 stacks, IEEE Transactions on Electron Devices 58 (9) (2011) 2878–2887. doi:10.1109/TED.2011.2158825.