Semiconductor Physics, Quantum Electronics & Optoelectronics, 2011. V. 14, N 2. P. 203-208.

PACS 85.30.Tv

Impact of sidewall spacer on gate leakage behavior of nano-scale MOSFETs

Ashwani K. Rana[1], Narottam Chand[2], Vinod Kapoor[1]
[1]*Department of Electronics and Communication, National Institute of Technology, Hamirpur Hamirpur (H.P)-177005, India*
E-mail: ashwani_paper@yahoo.com; kapoor@nitham.ac.in
[2]*Department of Computer Science and Engineering, National Institute of Technology, Hamirpur Hamirpur(H.P.)-177005, India*
E-mail: nar@nitham.ac.in

Abstract. Semiconductor devices with a low gate leakage current are preferred for low power application. As the devices are scaled down, sidewall spacer for CMOS transistor in nano-domain becomes increasingly critical and plays an important role in device performance evaluation. In this work, gate tunneling currents have been modeled for a nano-scale MOSFET having different high-k dielectric spacer such as SiO_2, Si_3N_4, Al_2O_3, HfO_2. The proposed model is compared and contrasted with Santaurus simulation results and reported experimental result to verify the accuracy of the model. The agreement found was good, thus validating the developed analytical model. It is observed in the results that gate leakage current decreases with the increase of dielectric constant of the device spacer. Further, it is also reported that the spacer materials impact the threshold voltage, on current, off current, drain induced barrier lowering and sub-threshold slope of the device.

Keywords: MOSFET, spacer, leakage current.

Manuscript received 30.09.10; accepted for publication 16.03.11; published online 30.06.11.

1. Introduction

To improve the performances of electronic devices, the size of their active components is scaled down according to the International Technology Roadmap for Semiconductors (ITRS) [1]. As we approach the nano-regime, a whole new set of problems regarding the device performance arises [2]. The control of leakage power is one of the most important issues for scaling MOSFET towards nano-regime [3]. For nano-scale MOSFET, gate leakage current is considered as a dominant leakage component as compared to sub-threshold leakage [4], as gate oxide thickness approaches its manufacturing and physically limiting value of less than 2 nm [2]. Hence, accurate estimation of the gate leakage current is essential to appreciate the total off-state power dissipation.

Numerous models have been developed numerically [5-7] in the past for calculating the tunneling current, but this approach is not always practical and is time consuming. Schuegraf *et al.* [8, 9] have derived a simple analytical formula to represent direct tunneling through a trapezoidal barrier. However, this model suffers from various limitations such as (i) gate current does not approach to zero as gate voltage goes to zero and does not fit experimental data at sub-1-V gate bias range, (ii) the assumption of constant effective mass for all energies is not accurate, (iii) non-consideration of quantum effects. Lee and Hu [10, 11] proposed a semi-empirical model by introducing the correction function to Schuegraf's analytical model to take care of above-mentioned secondary effect. However, this model has not considered the edge direct tunneling current (EDT). In [12], direct tunneling current expressions have been developed both for channel gate tunneling current and EDT including polydepletion effect and quantization effect with four adjustable parameters. This model does not include: (i) the non-uniform dopant profile in polygate in vertical direction resulted due to low energy ion implantation, (ii) additional depletion layer at the gate edges due to gate length scaling down, and (iii) gate oxide barrier

lowering due to image charges across the Si/SiO_2 interface. These nano-scale effects (NSE) are inevitable for nano-scale devices operating into deep sub-50-nm regime. Therefore, it is mandatory to include these NSE effects in nano-scale MOSFET to achieve an accurate estimation of the gate tunneling current.

In this work, an effective model has been developed for analyzing the gate tunneling current of nano-scale NMOSFET by considering the NSE effect that are difficult to ignore at nano-scale regime. This work mainly focuses on the impact of device spacer on gate leakage current and other device parameters.

The rest of the paper is organized as follows. In Section 2, modeling of the gate tunneling current is developed. The device structure and design used for simulation of set up is presented in Section 3. The results obtained are discussed in Section 4. Finally, concluding remarks are offered in Section 5.

2. Theoretical model

In ultra short-channel MOSFETs, in addition to gate to channel direct tunneling current, the source/drain extensions (overlap regions) direct tunneling current known as edge direct tunneling current (EDT) has been identified as the principal source of off-state power dissipation in VLSI chips because source/drain extensions (overlap regions) under poly-silicon gate represent a significant fraction of the device, as they do not scale at the same rate as the gate length. Therefore, the evaluation of EDT is critical for state of the art MOSFETs.

Modeling of the direct tunneling current analytically has been largely based on the WKB approximation [8]. The discrepancies that were present in the original WKB approximation [8] have been rectified in [10, 11] by introducing few adjusting parameters but they neglected the nano-scale effects. In our work, we adopt this model to evaluate the direct tunneling current from channel and overlap region in the nano-scale regime by taking the nano-scale effect into account. In this scheme, the value of fitting parameter $\alpha_{(ch,ov)}$ for channel and overlap region has been used as 0.6 and 0.45, respectively, with oxide spacer to match the overall best fit with Santaurus simulation and also with the experimental results reported in [12]. The T_{ox} refers to the physical oxide thickness and effective mass of the carrier in the oxide has been used as 0.40 m_o through this work. The total gate leakage current is given by

$$I_g = I_{gc} + I_{gso} + I_{gdo}, \qquad (1)$$

where I_{gc} is the gate-to-channel tunneling current, I_{gso} is the gate-to-source overlap region gate tunneling current, and I_{gdo} is the gate-to-drain overlap region tunneling current. Since drain to source V_{ds} is taken to be zero for simplification, so I_g can be modified as below

$$I_g = I_{gc} + I_{ov}; \qquad I_{ov} = 2I_{gso}. \qquad (2)$$

The channel current I_{gc} and EDT current I_{ov} per micrometer can be written as: $I_{gc} = J_{ch} \times L_{eff}$; $I_{ov} = J_{ov} \times L_{ov}$. $L_{eff} = L_g - 2L_{ov}$, where L_g is the total gate length, L_{ov} is the overlap gate length, and L_{eff} is the effective gate length. The channel current density J_{ch} and overlap current density J_{ov} are modeled as follows:

$$J_{(ch,ov)} = AC_{F(ch,ov)}T_{WKB(ch,ov)}, \qquad (3)$$

where $A = q^3 / 8\pi\phi_{b\_eff}\varepsilon_{ox}$, $C_{F(ch,ov)}$ is the correction term incorporated in [10, 11] and $T_{(ch,ov)}$ is the modified WKB transmission probability and are modified for channel and overlap region. ε_{ox} is the permittivity of the gate oxide and ϕ_{b_eff} is the effective barrier height, calculated as below,

$$\phi_{b\_eff} = \phi_b - \Delta\phi, \qquad (4)$$

$$\Delta\phi = \sqrt{\frac{qE_{gi}}{4\pi\varepsilon_{gi}}} = \sqrt{\frac{qV_{gi}}{4\pi\varepsilon_{gi}T_{gi}}} = \left(\frac{2q^3N_{eff}\phi_{b\_eff}}{16\pi^2\varepsilon_{gi}^3}\right)^{1/4}. \qquad (5)$$

$\Delta\phi$ is the reduction in the barrier height at the high-k/Si interface from ϕ_b, so that the barrier height becomes ϕ_{b_eff}. This reduction in barrier height is due to image charges across the interface. This barrier reduction is of great interest, since it modulates the gate tunneling current.

$$C_{(ch,ov)} = \exp\left[\frac{20}{\phi_{b\_eff}}\left(\frac{|V_{ox(ch,ov)}| - \phi_{b\_eff}}{\phi_{b\_eff}} + 1\right)^{\alpha(ch,ov)}\left(1 - \frac{|V_{ox(ch,ov)}|}{\phi_{b\_eff}}\right)\right]\left(\frac{V_g}{T_{ox}}\right)N_{DTC(ch,ov)},$$

$$T_{WKB(ch,ov)} = \exp\left[\frac{-8\pi\sqrt{2m_{ox}}\phi_{b\_eff}^{3/2}\left[1 - \left(1 - \frac{|V_{ox(ch,ov)}|}{\phi_{b\_eff}}\right)\right]^{3/2}}{3hq\,|E_{ox(ch,ov)}|}\right],$$

$$N_{DTC(ch)} = \begin{cases} \dfrac{\varepsilon_{ox}}{t_{gi}}\left\{n_{acc}v_t \ln\left[1 + \exp\left(-\dfrac{(V_g - V_{FB})}{n_{acc}v_t}\right)\right]\right. \\ \text{for } V_g < 0 \\ \dfrac{\varepsilon_{ox}}{t_{gi}}\left\{n_{inv}v_t \ln\left[1 + \exp\left(-\dfrac{(V_g - V_{th})}{n_{inv}v_t}\right)\right]\right. \\ \text{for } V_g > 0 \end{cases}$$

$$N_{DTC(ov)} = \begin{cases} \dfrac{\varepsilon_{ox}}{t_{gi}}\left\{n_{acc}v_t \ln\left[1 + \exp\left(-\dfrac{(V_g)}{n_{acc}v_t}\right)\right]\right. \\ \text{for } V_g < 0 \\ \dfrac{\varepsilon_{ox}}{t_{gi}}\left\{n_{inv}v_t \ln\left[1 + \exp\left(-\dfrac{(V_{ge})}{n_{inv}v_t}\right)\right]\right. \\ \text{for } V_g > 0 \end{cases}$$

Where $\alpha_{(ch,ov)}$ is the fitting parameter depending upon channel or source/drain overlap tunneling, n_{inv} and n_{acc} are the swing parameters, V_{FB} represents the flat band voltage, $N_{DTC(ch,ov)}$ denotes the density of carrier in channel/overlap region depending upon MOSFET biasing condition, and V_{ge} is the effective gate voltage excluding polygate non-uniformity and gate length effect and is equal to $V_g - V_{poly}$, where V_{poly} is the voltage drop due to polydepletion in the polygate.

The default values of n_{inv} and n_{acc} are S/v_t (S is the sub-threshold swing and v_t is the thermal voltage) and 1, respectively. The correction factor $C_{F(ch,ov)}$ and transmission probability $T_{WKB(ch,ov)}$ are different for channel and source/drain overlap region, because both channel and overlap components have different values of $V_{ox(ch,ov)}$ and $N_{DTC(ch,ov)}$. It is because of the fact that overlap region has almost zero flat band voltage, as both SDE region and overlying polygate Si are heavily doped n^+ regions. $N_{DTC(ch,ov)}$ has been given differently for both region as in (12) and (13). The gate oxide voltage V_{ox} for the channel and SDE overlap are calculated as follows.

Case (*i*): when $V_g > 0$.

In this biasing condition for MOSFET device, there is a depletion layer in the polygate thereby causing an additional potential drop across the gate. The SDE region enters into accumulation and substrate region enters into the week inversion below V_{th} and strong inversion beyond V_{th}. Therefore, both the channel and EDT component are present and are comparable.

Case (*ii*): when $V_{FB} < V_g < 0$.

Here, gate tunneling current is dominated by the EDT where electric field is such that electron are directed from the accumulated polygate into the overlap region. On the other hand, substrate is in depletion/weak inversion and constitutes a negligible tunneling current. This region of biasing is primarily responsible for off-state power dissipation. Thus, EDT plays an important role in the evaluation of off-state power dissipation.

Case (*iii*): when $V_g < V_{FB}$.

In this region of operation, substrate goes into accumulation. As a result, both current components become comparable. The voltage across the gate oxide for different region of operation is as follows:

$$V_{ox} = \begin{cases} \left(V_g - \phi_s - V_{FB}\right) & \text{for } V_g < 0 \\ \left(V_{ge} - \phi_s - V_{FB}\right) & \text{for } V_g > 0 \end{cases} \qquad (6)$$

Where ϕ_s is the surface band bending of the substrate and are calculated for channel and overlap region depending upon the biasing condition of the MOSFET device including the poly-non-uniformity, gate length effects and image force barrier lowering.

The gate effective voltage in the gate is derived as follows:

$$\therefore V_{ge} = (V_{FB} + \phi_{so} - \Delta V_{p1} - \Delta V_{p2}) + \frac{\left(q\varepsilon_{si} N_{poly} T_{ox}^2\right)}{\varepsilon_{ox}^2}\left[\sqrt{1 + \frac{2\varepsilon_{ox}^2 (V_g - V_{FB} - \phi_{so})}{q\varepsilon_{si} N_{poly} T_{ox}^2}} - 1\right]. \qquad (7)$$

This equation includes the non-uniformity in the gate dopant profile through the term ΔV_{p1} and fringing field effect, *i.e.* gate length effect through a term ΔV_{p2}. ϕ_{so} is surface band bending of the substrate by taking the quantization effect into account. The potential drop ΔV_{p1} due to non-uniform dopant profile in poly-Si gate, caused by low energy implantation, is calculated as below

$$\Delta V_{p1} = \left(\frac{kT}{q}\right) \ln\left(\frac{N_{poly-top}}{N_{poly-bottom}}\right) \qquad .(8)$$

$N_{poly-top}$ and $N_{poly-bottom}$ are the doping concentration at the top and bottom of the poly-silicon gate. The potential drop ΔV_{p2} due to the gate length effect, caused by very short gate lengths, is given below:

$$\Delta V_{p2} \approx \frac{\Delta Q}{C_d} = \frac{2qAN_d}{L_g C_d} \left(\text{V}/\text{cm}\right), \qquad (9)$$

$$C_d = \delta \frac{\varepsilon_{spacer}}{\pi} \ln\left[\frac{3 - \cos\left\{\pi\left(\frac{T_F - T_{gi}}{T_F}\right)\right\}}{1 + \left\{\pi\left(\frac{T_F - T_{gi}}{T_F}\right)\right\}}\right], \qquad (10)$$

where A denotes the triangular area of the additional charge, L_g is the gate length, C_d is the depletion capacitance in the sidewalls [13], ε_{gi} is the permittivity of the device spacer, T_F is the thickness of the device spacer, T_{gi} is the thickness of the gate insulator, and δ is the fitting parameter normally equal to 0.95.

3. Simulation set up

Fig. 1 shows the device structure used for simulation in Santaurus simulator. The deep S/D region is composed of a heavily doped silicon and a silicide contact. Doping the silicon S/D region is assumed to be very high, 1×10^{20} cm^{-3}, which is close to the solid solubility limit and introduces negligible silicon resistance. The dimension of the silicon S/D region is taken as 50 nm long and 20 nm high. This gives a large contact area resulting in a small contact resistance.

The heavily doped silicon called deep S/D region extends into the silicon film at both ends and constitutes the extended S/D for the device (labelled by ''S_1''and ''D_1'' in Fig. 1). The doping concentration of the acceptors in the silicon channel region is assumed to be graded due to diffusion of dopant ions from heavily doped S/D region with a peak value of 1×10^{18} cm^{-3} and

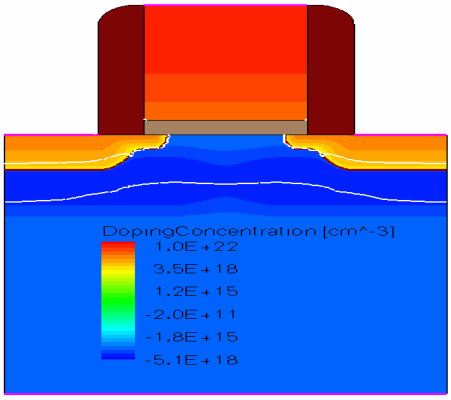

Fig. 1. NMOSFET device structure used in simulation.

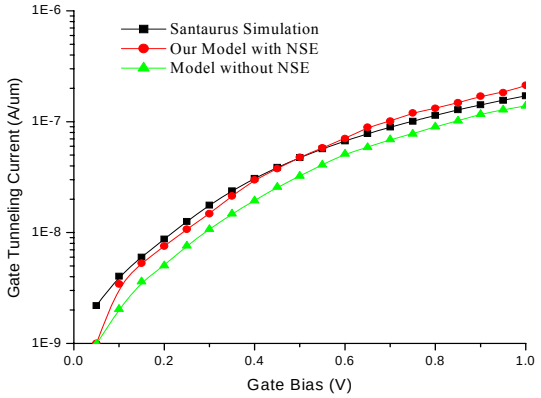

Fig 2. Comparison of model with simulated data for gate oxide thickness of T_{ox} = 1.0 nm with oxide spacer, metallurgical gate length of L_{met} = 25 nm and S/D overlap length of L_{ov} = 10 nm in nano-scale regime.

1×10^{17} cm^{-3}near the channel. The doping concentration in the poly-silicon gate is 1×10^{22} cm^{-3} at the top and 1×10^{20} cm^{-3} at the bottom of poly-silicon gate, i.e. interface of oxide and silicon. The halo implantation made around S/D also reduces short-channel effects, such as the punch-through current, drain induced barrier lowering (DIBL), and threshold voltage roll-off for different non-overlap lengths.

The MOSFET has a 50-nm-thick n^+ poly-Si gate with the metallurgical gate length of 25 nm and a 1-nm gate oxide. The MOSFET with L_{met} of 25 nm was designed to have a V_T of 0.23 V with SiO_2 as spacer. We determined V_T by using a linear extrapolation of the linear portion of the $I_{DS} - V_{GS}$ curve at low drain voltages. The operating voltage for the devices is 1 V. The simulation study has been conducted in two dimensions, hence all the results are in the units of per unit channel width.

The simulation of the device is performed by using Santaurus design suite [14, 15] with drift-diffusion, density gradient quantum correction and advanced physical model being turned on.

4. Results and discussion

In this section, computation of gate tunneling currents for a n-channel nano-scale MOSFET having different sidewall spacer such as SiO_2 (k = 3.9), Si_3N_4 (k = 7.5), Al_2O_3 (k = 9.0) and HfO_2 (k = 22) have been carried out. This model is computationally efficient and easy-to-realize. This model calculates the gate tunneling current by using $\alpha_{(ch/ov)}$ as fitting parameters. Thus, this model is applicable to many alternate high-k nano-MOSFET simply by adjusting the fitting parameter. Variation of the total gate tunneling current with a gate bias for a given values of gate insulator thickness has been presented for possible alternative sidewall spacer such as SiO_2, Si_3N_4, Al_2O_3 and HfO_2. The impact of sidewall spacer on the device threshold voltage, off current, on current, DIBL and sub-threshold slope (SS) is also reported in results.

The comparison between the simulated data and the model value for the gate tunneling current is presented in Fig. 2 for the value L_{met} = 25 nm, L_{ov} = 10 nm, T_{ox} = 1.0 nm. The model value while considering the nano-scale effect shows good agreement with the simulated value over the entire positive gate bias range, certifying the high accuracy of the proposed analytical modelling. Model also shows good agreement with simulated data for various sidewall spacers but with different values of the fitting parameter as listed in Table 1.

It is also shown that the model value without any nano-scale effect does not show good agreement with the simulated value, emphasizing the need to include nano-scale effect.

Similarly, the model is also verified in Fig. 3 with experimental data published in [12] for value L_g = 0.17 μm, L_{ov} = 10 nm, T_{ox} = 1.85 nm and W_g = 10 μm. The substrate doping and poly-silicon gate doping have been taken to be 4.1×10^{17} cm^{-3} and 5×10^{20} cm^{-3}, respectively. The model value also shows good agreement with the experimental data over the entire gate bias range, certifying the high accuracy of the proposed analytical modelling.

Fig. 4 shows the variation of the gate tunneling current and off current of the NMOS device with S/D overlap length for optimization of the latter at a given gate bias of 0.6 V and gate oxide thickness of 1.0 nm. It is observed that off current for a device under consideration is slightly less at S/D overlap length of 5 nm. Therefore, S/D overlap length of 5.0 nm is considered in further results.

Table 1. Fitting parameter for calculation of the gate tunneling current through different sidewall spacers.

Parameter	SiO_2	Si_3N_4	Al_2O_3	HfO_2
α_{ch}	0.6	0.62	0.71	0.78
α_{ov}	0.45	0.49	0.53	0.55

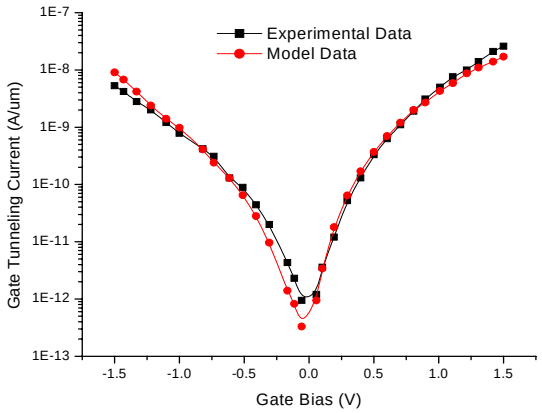

Fig. 3. Comparison of the model with experimental data for N_{sub} = 4.1×10^{17} cm^{-3} and N_{poly} = 5×10^{19} cm^{-3} with oxide spacer.

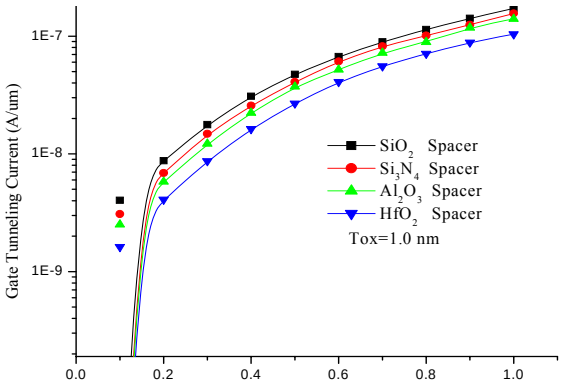

Fig. 5. Gate tunneling current vs gate bias for different sidewall spacer in nano-scale regime at L_{met} = 25 nm and L_{ov} = 5.0 nm.

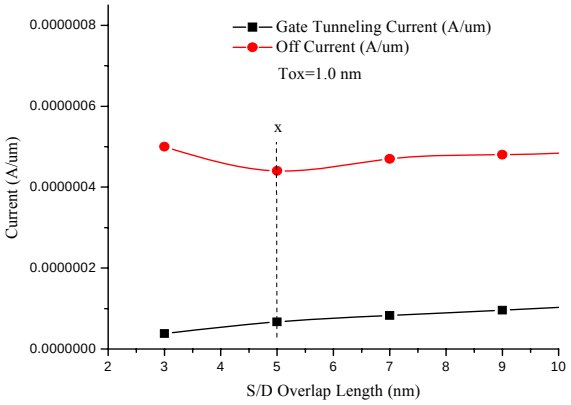

Fig. 4. Gate tunneling current and off current vs S/D overlap length for T_{ox} = 1.0 nm, L_{met} = 25 nm.

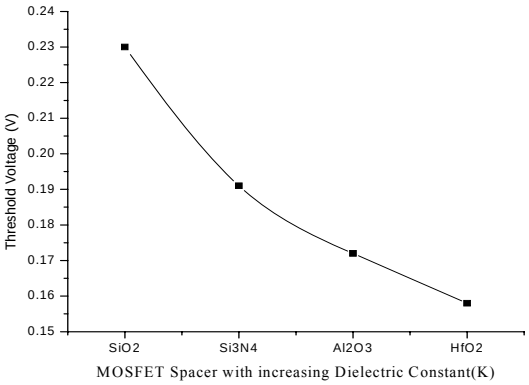

Fig. 6. Device threshold voltage vs sidewall spacer with L_{met} = 25 nm and L_{ov} = 5.0 nm.

Fig. 5 shows variation of the gate tunneling current with the gate bias for various sidewall spacers at gate oxide thickness of 1.0 nm. It is observed that gate leakage current improves with the introduction of sidewall spacer of increasing dielectric constant *K*. The application of high-k spacer enhances the fringing electric field thereby reducing the effective gate voltage. This reduction lowers the transverse electric field responsible for carrier tunneling through gate oxide. Consequently, the gate leakage current reduces as dielectric constant of the sidewall spacer increases.

In Fig. 6, variation of the device threshold voltage with sidewall spacer is presented. As the dielectric constant of the sidewall spacer increases, the fringing field increases. These field lines finally induce an electric field from the source-to-channel thereby reducing the source-to-channel barrier height. Since the threshold voltage of the device is controlled by the injection of electrons over this potential barrier, it decreases with increasing dielectric constant of the sidewall spacer. Thus, sidewall spacers with a larger dielectric permittivity reduce the threshold voltage owing to the enhanced value of fringing electric field.

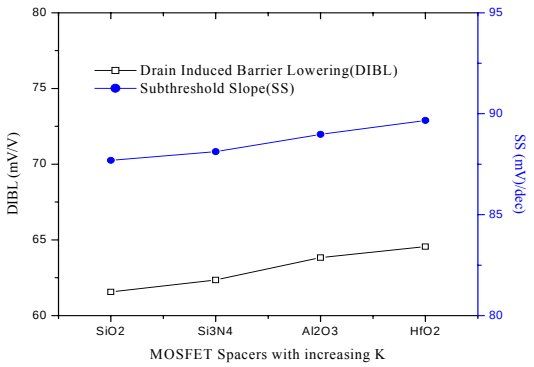

Fig. 7. Drain induced barrier lowering (DIBL) and sub-threshold slope (SS) vs sidewall spacer with L_{met} = 25 nm and L_{ov} = 5.0 nm.

Fig. 7 shows that DIBL increases with increase in dielectric constant of the sidewall spacer. It is due to the fact that the increased effect of fringing field on channel by the application of high-k sidewall spacer weakens the gate control over the channel region of a MOSFET. Due to this decrease in gate control, the drain electrode is tightly coupled to the channel, and the lateral electric

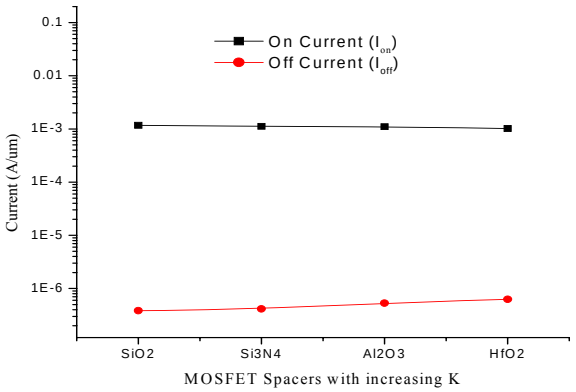

Fig. 8. On and off currents vs sidewall spacer with L_{met} = 25 nm and L_{ov} = 5.0 nm.

field from the drain reaches a larger distance into the channel. Consequently, this electrically closer proximity of drain to source gives rise to higher drain-induced barrier lowering in MOS transistors. It is also shown in Fig. 7 that sub-threshold characteristics degrade due to decrease in the threshold voltage.

As shown in Fig. 8, on current (I_{on}) and off current (I_{off}) degrade slightly due to a decrease in the threshold voltage as well as due to the degraded sub-threshold characteristics.

5. Conclusion

The impact of sidewall spacer on gate leakage current and other device parameters is studied using the gate tunnel model and extensive device simulations. A high-k sidewall spacer lowers the gate leakage current while increases the sub-threshold slope with drain induced barrier lowering. Sidewall spacers with a larger dielectric permittivity reduce the threshold voltage owing to the enhanced value of fringing electric field. It is found that the use of high-k sidewall spacers also degrades the on and off current marginally.

References

1. Online. Available: http://www.itrs.net/reports.html
2. C.H. Choi, K.Y. Nam, Z. Yu, and R.W. Dutton, Impact of gate direct tunneling current on circuit performance: A simulation study // *IEEE Trans. Electron Devices*, **48**, p. 2823-2829 (Dec. 2001).
3. F. Rana, S. Tiwari, and D. Buchanan, Self-consistent modeling of accumulation layers and tunneling currents through very thin oxides // *Appl. Phys. Lett.*, **69**, p. 1104-1106 (1996).
4. S.-H. Lo, D.A. Buchanan, Y. Taur, and W. Wang, Quantum-mechanical modeling of electron tunneling current from the inversion layer of ultra-thin-oxide nMOSFETs // *IEEE Electron Device Lett.*, **18**, p. 209-211 (May 1997).
5. W.K. Shih *et al.*, Modeling gate leakage current in nMOS structures due to tunneling through an ultra-thin oxide // *Solid State Electron.*, **42**, p. 997-1006 (June 1998).
6. K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits // *Proc. IEEE*, **91**(2), p. 305-327 (Feb. 2003).
7. Y. Ando and T. Itoh, Calculation of transmission tunneling current across arbitrary potential barriers // *J. Appl. Phys.*, **61**(4), p. 1497 (1987).
8. K.F. Schuegraf, C.C. King, and C. Hu, Ultra-thin silicon dioxide leakage current and scaling limit // *Symposium on VLSI Technology Digest of Technical Papers*, p. 18-19 (1992).
9. K.F. Schuegraf and C. Hu, Hole injection SiO_2 breakdown model for very low voltage lifetime extrapolation // *IEEE Trans. on Electron Devices*, **41**(5), 1994.
10. W.C. Lee and C. Hu, Modeling gate and substrate currents due to conduction- and valence-band electron and hole tunneling // *Symposium on VLSI Technology Digest of Technical Papers*, p. 198 (2000).
11. W.C. Lee and C. Hu, Modeling CMOS tunneling currents through ultrathin gate oxide due to conduction- and valence-band electron and hole tunneling // *IEEE Trans. Electron Devices*, **48**(7), (2001).
12. Xin (Ben) Gu, Ten-Lon Chen, Gennady Gildenblat,, Glenn O. Workman, Surya Veeraraghavan, Shye Shapira, and Kevin Stiles, A surface potential-based compact model of n-MOSFET gate-tunneling current // *IEEE Trans. Electron Devices*, **51**(1), (2004).
13. Steve Shao-Shiun Chung and Tung-Chi Li // *IEEE Trans. Electron Devices*, **39**(3), (1992).
14. ISE TCAD: Synopsys Santaurus Device User Manual, 1995-2005, Synopsys, Mountain View, CA.
15. ISE TCAD: Synopsys Santaurus Device simulator.