Int. J. Nanoelectronics and Materials **5**(2012) 39-45

Quantum Mechanical direct leakage currents in a sub 10nm mosfet: a rigorous

Amit Chaudhry[1,*] Jatindra Nath Roy[2]

[1]*Faculty of Engineering and Technology, University Institute of Engineering and Technology, Panjab University, Chandigarh, India.*

[2]*Solar Semiconductor Private Limited, Hyderabad, India*.

Received 22 Jan. 2010; Revised 13 Jan. 2011; Accepted 1 Feb. 2011

Abstract

In this paper, we have developed an analytical model for the quantum mechanical source to drain tunneling effects occurring in the sub 10nm nanometer scale MOSFETs. Inversion layer quantization, band-gap narrowing and drain induced barrier lowering effects have been included in the developed model. Results predict that the source to drain tunneling results in an increase of leakage currents in sub 10 nm MOSFETs and hence cannot be ignored. The results match closely with the numerical results already reported in literature proving the accuracy of the model.

Keywords: QME; WKB; Inversion Layer quantization; Tunneling.
PACS: 85.35.-p; 85.40.Bh; 73.40.-c; 73.63.Rt.

1. Introduction

The scaling of MOSFET into sub 10nm region is major challenge today due to severe leakage current occurring in MOSFETs. The International Technology Roadmap for Semiconductors (ITRS) [1] predicts that the scaling will continue, and that in 2016, the MOSFET gate length will reach 10nm region. Many research groups [2], [3] around the world are engaged in the development of sub 10nm-level MOSFETs. Due to the obstacles in down scaling, several groups in the world have shifted from the traditional MOSFET downscaling to an alternative technology or non classical structures such as dual gate MOSFET technology, carbon nanotubes, molecular electronics, resonant tunneling diodes etc.[4],[5]. But this too has some possible disadvantages such as, integrating these components with the existing silicon technology and other constraints like packaging issues still need to be seen. The use of the conventional CMOS technology for the microelectronic applications still dominates majority of the electronics industry. So, there is still an ample scope of continuing with the existing MOSFET scaling to sub 10nm region. The existing problems of scaling down the MOSFETs to nanoscale are inversion layer quantization effects which result in the degradation of gate capacitance, threshold voltage increase and decrease in drain current. These factors are the limiting factors in the scaling down of the MOSFET device. If we continue with the MOSFET technology and reach the sub 10nm level, another limiting factor in scaling down is the source to drain direct tunneling current density that contributes positively to the sub threshold currents. This increases the sub threshold leakage and hence the static power dissipation and failure to switch off the

[*]) For Correspondence; E-mail: amit_chaudhry01@yahoo.com.

MOSFET. Hence, it becomes inevitable that the downscaling to sub 10nm scale can only continue if suitable MOS technologies, design procedures and modeling tools are developed at that scale. Development of such a complex technology requires the accurate and analytical modeling of the MOSFET before they can be passed over to the fabrication phase.

In this paper, the direct source to drain tunneling current density of sub 10nm MOSFETs is explored by WKB technique. The physics of drain induced barrier lowering, band-gap narrowing and inversion layer quantization has been explained and suitably included in the tunneling equations. We use a transmission view of the device in which carriers are injected into the channel from the source, across a potential energy barrier whose height is modulated by the gate voltage, into the channel and then under drain bias into the drain. Here, we use a MOSFET with simple architecture and gate lengths in the range 6nm and 4 nm for study and finally the paper ends with conclusion and references.

2. Physics of tunneling

Electrons are injected from the source into the channel across a potential barrier the height of which is modulated by gate voltage. Carriers drift across the channel and are collected by the drain. The existence of the source to channel barrier is influenced/lowered by drain-induced barrier lowering. The physics of the transport describes the transmission of carriers from the source into the substrate and the second transmission across the substrate to the drain. A complete analysis is required to understand the source to drain tunneling by considering DIBL, inversion layer quantization process and band-gap narrowing. The research work in sub 10nm MOSFETs off leakage currents and quantum mechanical tunneling from source to drain tunneling in particular is largely centered around numerical models [6] and [7] .One-dimensional numerical calculations of the source to drain tunneling current were performed by the research group as reported in [8].

Modeling of source to drain direct tunneling

The source to drain tunneling current is mainly dependent on the width and height of the potential barrier between source and drain. Using the WKB method [9],[10] for source to drain tunneling process, we get the transmission probability. The main issue in the current density calculation is transmission probability T(E). The general expression for the transmission probability is given by [10]:

$$T(E) = \exp\left[-2\int_{x_1}^{x_2} |k(x)|\, dx\right] \tag{1}$$

$$k(x) = \left[2m_1 \frac{V(x) - Ex}{\mathrm{h}^2}\right]^{1/2}$$

The direct tunneling current density is given by [9]

$$J_T = \left(\frac{4\pi m_t q}{h^3}\right)\int_0^V \left\{\int_0^\infty [f_S(E) - f_D(E)]\, dE_y\right\} T(E\ )dE_x \tag{2}$$

Electron distribution at the source/substrate interface = $f_S(E) = \left[1 + \frac{\exp(E - E_{fs})}{kT}\right]^{-1}$,

Electron distribution at the drain/substrate interface = $f_D(E) = \left[1 + \frac{\exp(E - E_{fd})}{kT}\right]^{-1}$,

E_{fs} is fermi energy at the source, E_{fd} is fermi energy at the drain, E is the total energy of electrons, E_x is the energy in the direction of tunneling, i.e., from source to drain. E_y is the energy in the direction of tunneling, i.e., from source to drain .Using (1) and (2), the tunneling current density from source to drain can be evaluated.

$$J_T = \left(4\pi m_l q(kT)^2 / h^3\right)\left(1 + \gamma_t kT / 2\sqrt{V}\right) exp\left(E_{fs}\right) exp\ \left(-\gamma_t \sqrt{V}\right) \quad (3)$$

Ex=Energy of electron from source to drain, ħ = Planck's constant, ml = longitudinal electron mass= 0.916mo , Barrier Height (V)= 0.5(2Vbi +Vds) -φsqm- Vm ,Efs =0.5 Eg-qφf, Eg = Silicon band-gap, φf = Fermi potential, Vbi = Built in potential, Vds= Drain voltage, Vm = Surface potential increase due to drain induced barrier lowering effect=, γt= 4πL(2ml)1/2/h , φsqm =Quantum surface potential from (14). L is the channel length.

Calculation of V_m

Solving the Poisson's equation at the source/substrate interface and applying the boundary conditions, the potential variation as:

$$V_s(x) = V_{bi}\,(1\text{-}x/W_1)^2 \quad (4)$$

Similarly, solving the Poisson's equation at the drain/substrate interface and applying the boundary conditions, the potential variation in the substrate as:

$$V_d\,(y) = (V_{bi}+V_{ds})\,(1\text{-}y/W_2)^2 \quad (5)$$

y=L-x, W1= (2εoεsiVbi/qNb)1/2 = depletion width at the source, W2 ={2εoεsi (Vbi+Vds)/qNb}1/2 = depletion width at the drain. Nb is the substrate concentration

Equating (4) and (5) and solving for x, we get,

$$x= (1\text{-}R+RL/W_2)/(R/W_2+1/W_1);\quad R= \{(V_{bi}+V_{ds})/V_{bi}\}^{1/2} \quad (6)$$

Therefore, the minimum potential in the substrate is given by:

$$V_m=V_{bi}\,\{1\text{-}(x/W_1)\}^2 \quad (7)$$

3. Inversion layer quantization

The nanometer-scale MOSFETs use highly-doped substrate and ultra-thin gate oxides to control short-channel effects such as drain induced barrier lowering (DIBL) and the punch through effect. All these methods used to control short channel effects result in a high electric field in the direction vertical to the silicon/silicon oxide interface. Although the high electric field in the vertical direction can keep the charges in the channel under gate control against the influence of drain potential, it confines the movement of carriers in a narrow potential well existing between the surface potential distribution and the infinite oxide potential. According to Heisenberg principle, the energy of the channel carriers can only take discrete values and not a continuous energy distribution as described by classical device physics.

The silicon energy band is composed of six equal energy lobes orienting towards six directions. Every energy lobe has two directions also. One is longitudinal and the other is the transverse direction. So, the electrons present in these two directions have masses 0.916mo and 0.19mo respectively. Let the Si/SiO2 interface is towards (100) direction. So, the electrons in two lobes along the interface have mass 0.916mo and in the other four lobes have transverse mass 0.19mo along the Si/SiO2 interface. So, combining these four lobes of transverse mass 0.19mo are grouped together and the other two lobes are grouped together as shown in figure 1. When inversion layer quantization occurs, the electrons reside in lower energy valleys i.e. 0.916mo mass. So, 90% of the electron population is in lower valley having longitudinal mass 0.916mo and transverse mass 0.19mo. Also the lower valley is slightly above the conduction band edge of the silicon conduction band as also given by Heisenberg principle. This causes a significant decrease in the inversion carrier density at a Si/SiO2 interface in MOSFETs as compared to that of the classical case. Thus, it is important to model accurately the inversion layer quantization effect in a nanoscale MOSFET and understand the relationship between the inversion charge density and the surface potential. All the calculations done in this paper are based on the lower energy valley having longitudinal mass 0.916mo and transverse mass 0.19mo.

The research in the area of inversion layer quantization started in the early 1950s. The research[11]-[15] mainly focused on only calculating the inversion charge density in the presence of inversion layer quantization effects using variation approach and triangular well approach in the MOSFET. The use of such techniques required the calculation of surface potentials at the interface of silicon and its oxide. The lack of availability or slow development of surface potential models six decades ago, never allowed the growth of research in the area of modeling QME in MOSFETs. But as the MOSFETs are being scaled down to the nm scale, there is a need to analytically model the inversion layer quantization in nanoscale MOSFETs. Now we discuss the model, simulation and analysis of the inversion layer quantization process in the nanoscale MOSFETs. Solving the Poisson equation in the inverted channel, we get the total charge density, Qs.

$$Q_s = -(2qN_a\varepsilon_{si}\varepsilon_0)^{1/2}\left[\varphi_s + V_t e^{-2\varphi_f/V_t}\left(e^{\varphi_s/V_t} - 1\right)\right]^{1/2} \tag{8}$$

q is electron charge, ε_{si} is silicon relative permittivity, ε_0 is permittivity of free space, φ_s is surface potential, φ_f is fermi potential, Na is substrate concentration, and $V_t = kT/q$ is thermal voltage .

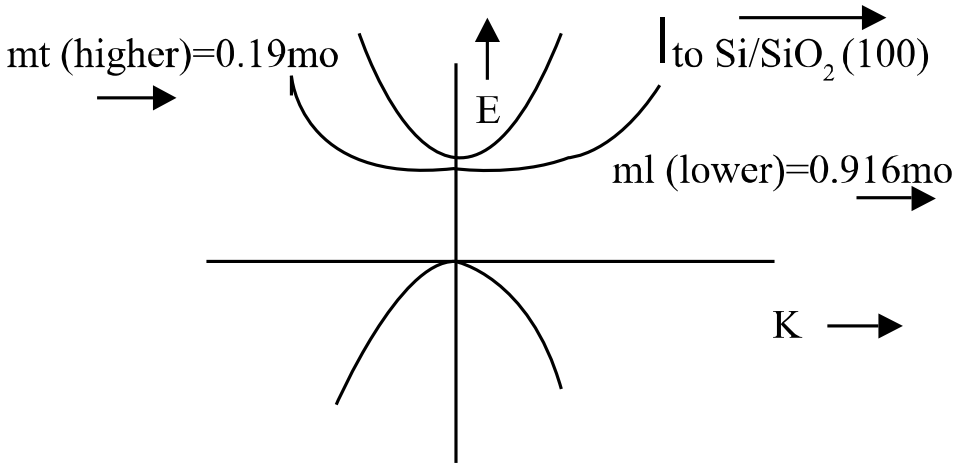

Fig. 1: E-k diagram showing inversion layer lower energy and upper energy and masses in the conduction band valleys.

Similarly, the depletion charge Q_b is approximated as

$$Q_b = -(2\varepsilon_{si}\varepsilon_0 qN_a\varphi_s)^{1/2} \tag{9}$$

Therefore, the inversion charge density Q_{inv} is given by (8) and (9):

$$Q_{\text{inv}} = -\gamma C_{\text{ox}}\left\{\left[\varphi_s + \frac{kT}{q}\exp\left(\frac{q(\varphi_s - 2\varphi_f)}{kT}\right)\right]^{1/2} - (\varphi_s)^{1/2}\right\} \tag{10}$$

γ is body effect parameter and Cox is oxide capacitance (Fcm-2). The main problem with (3) is that the surface-potential has to be evaluated explicitly in all the regions of inversion and then only, (3) can be solved. An explicit solution has been evaluated in [16]. The wave function solution of the Schrödinger's equation is given by using variation approach [11]:

$$\psi(x) = \frac{b^{3/2}x}{\sqrt{2}}\exp\left(\frac{-bx}{2}\right) \tag{11}$$

$$b = \left[\frac{48\pi^2 m_t q}{\varepsilon_{si}\varepsilon_0 h^2}\left((11/32)Q_{\text{inv}} + Q_{\text{dep}}\right)\right]^{1/3} \tag{12}$$

(12) is then included in the explicit surface potential expression given by [13]:

$$\begin{aligned}
&\varphi_s = f + a \\
&f = \varphi_f + 0.5\varphi_{\text{swi}} - 0.5\left[(\varphi_{\text{swi}} - 2\varphi_f)^2 + 0.0016\right]^{1/2} \\
&a = 0.025\ln\left\{\left[x - y\left(1+100y^2\right)^{-1/2}\right]^2\left(0.16\gamma\right)^{-2} - 40f + 1\right\} \\
&\varphi_{\text{swi}} = \left[\left(V_{\text{gs}} - V_{\text{fb}} + 0.25\gamma^2\right)^{1/2} - 0.5\gamma\right]^2
\end{aligned} \tag{13}$$

And φ_{swi} is the weak inversion surface potential, $x = V_{gs} - V_{fb} - f$, and $y = \varphi_{swi} - f$. The quantum surface potential is given by

$$\varphi_{sqm} = 2\varphi_f + \delta\varphi \tag{14}$$

Using the surface potential model (14) in (9) and (10), we can calculate explicitly inversion charge density and depletion charge density.

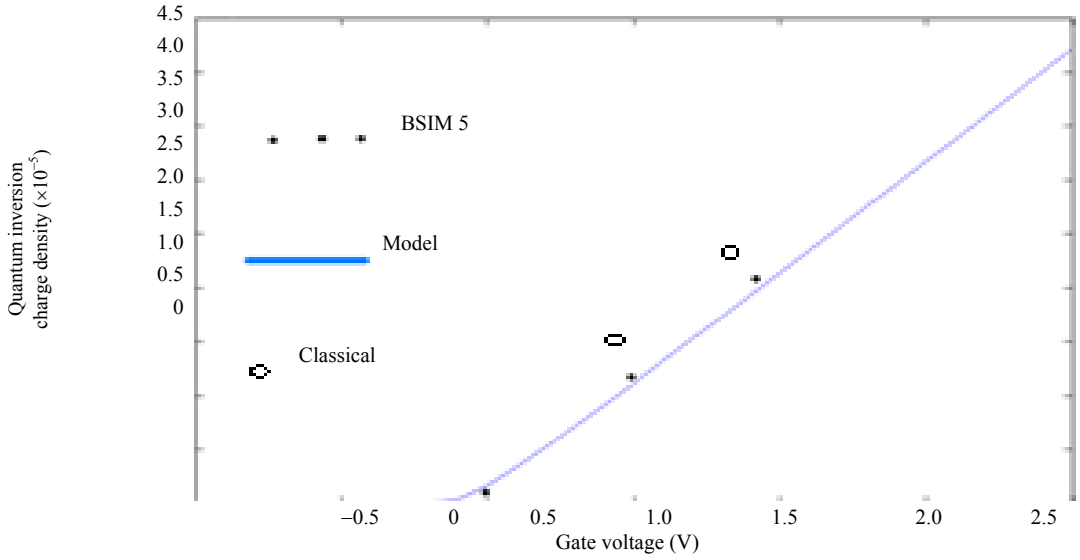

Fig. 2: Simulated results of quantum mechanical inversion charge density with gate voltage under the model parameters: substrate doping $1x10^{18}cm^{-3}$ and oxide thickness 1.5 nm. Black dots show the reported BSIM 5 results [17], round circles show the classical results and blue line show the quantum mechanical model developed.

The results in figure 2 match quite closely with the BSIM 5 results [17]. The results have been achieved by accurately modeling the shift in the surface potential. The results show that the inversion layer quantization leads to reduced inversion charge density.

1. Band gap narrowing effect

It is the reduction of the band-gap energy due to the excess doping concentration in the semiconductor. The band gap narrowing effect is applicable to both source and the drain. The reduction in energy band-gap is given by [18]

$$\delta E_g = 18.7 ln\left(N_D / 7x10^{17}\right) \tag{15}$$

N_D is doping concentration cm^{-3} in the source. The increase in the intrinsic carrier concentration due to excess doping is obtained by putting (15) in (16).

$$n = n_I exp\ \left(\delta E_g / 2kT\right) \tag{16}$$

n_1 is the intrinsic carrier concentration. The band-gap narrowing effect will reduce the built in potential existing at the source/drain and the substrate. This effect is also included in the sub 10nm MOSFET source to drain direct leakage model.

2. Drain Induced Barrier Lowering (DIBL)

DIBL is an influence of drain potential into substrate surface potential thus lowering the potential barrier in the substrate with increasing drain-source voltage and causing increasing sub threshold currents. DIBL results in decreased concentration in the substrate due to the depletion caused by the drain potential.

To account for the DIBL effect, the substrate concentration (Nb) is replaced with the effective substrate concentration [19] in all the equations i.e.

$$N_B=N_b-(2\varepsilon_o\varepsilon_{si}V_{ds}/qL^2) \quad (17)$$

So, using (7),(14),(16) and (17) in (3), source to drain tunneling current density can be calculated as shown in figure 3 and figure 4.

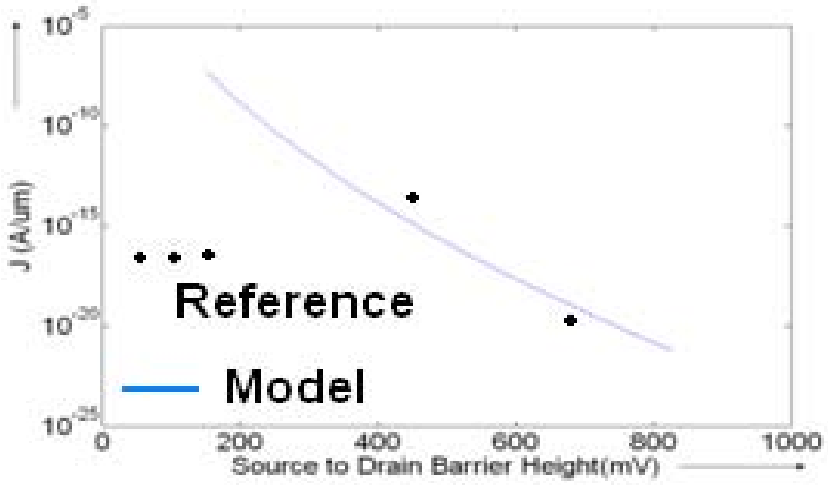

Fig. 3: Tunneling current variation with the barrier height at 6nm effective channel length. Black dots show the reported numerical results [20] blue line show the quantum mechanical model developed.

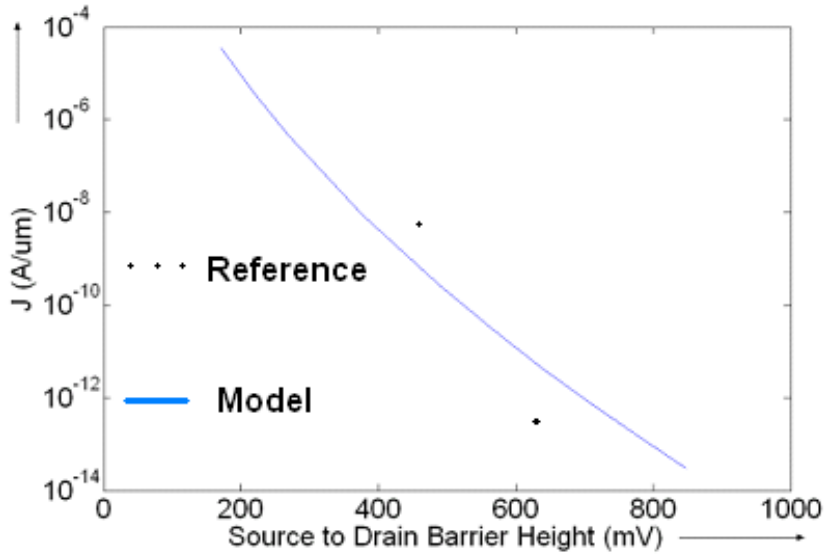

Fig. 3: Tunneling current variation with the barrier height at 4nm effective channel length. Black dots show the reported numerical results [20] blue line show the quantum mechanical model developed.

Table 1: Barrier height and source to drain quantum tunneling current variation

Barrier Height (mV) and the channel length (nm)	Tunneling Current A/um at t_{ox}= 1.5nm, $N_a=2x10^{20}cm^{-3}$ and $N_b=2x10^{18}cm^{-3}$
400 and 4	10^{-8}
400 and 6	10^{-12}

Results and discussion

The quantum mechanical direct tunneling source to drain currents is of the order of nano-amperes as shown in figures 3 and 4. The tunneling current density has been found in the presence of the gate and drain voltage including the energy quantization effects, drain induced barrier lowering and the band-gap narrowing effects. As shown in figure 3, at effective channel length of 6nm, the quantum mechanical tunneling off currents are 10-6nA/um at drain voltage of 0.5V. In figure 3, at 4nm, the quantum mechanical tunneled off currents are 0.1nA/um at a drain voltage of 0.5V. Simulation is done at source/drain doping of 2x1020cm-3, oxide thickness of 1.5nm and substrate doping of 2x1018cm-3. The calculations have been done by taking band-gap narrowing effect at the source/substrate and drain/substrate junctions. The tunneling current densities increase as the barrier heights and channel lengths are reduced. The results are well in accordance with the numerical calculations as given in the standard references [20].

Conclusion

In this paper, an analytical model has been developed for the direct source to drain tunneling in the sub 10nm MOSFETs. Three major effects such as energy quantization drain induced barrier lowering and band-gap narrowing has been included in the calculations. The dependence of source-to-drain tunneling with channel length and barrier height has been investigated. The major conclusion is that the tunneling current density cannot be ignored in such nano-scaled MOSFETs. The inversion layer quantization, drain induced barrier lowering and band-gap narrowing effects increase the leakage currents significantly. We have found that the tunneling currents have a dominating role for channel lengths below 4nm and further scaling down may prove fatal with the device.

Acknowledgment

The authors thank the Director, UIET, Panjab University, Chandigarh, India for allowing to carry out the work. The authors would like to thank to Panjab University, Chandigarh, India for providing excellent research environment to complete this work. The authors wish to thank all individuals who have contributed directly or indirectly in completing this research work.

References

[1] P. M. Zeitoff, J. E. Chung, IEEE Circ. Dev. **21** (2005) 4
[2] D. Foty, *MOSFET Modeling with SPICE, Principles and Practice*, Prentice Hall, New York, (1997)
[3] BSIM 4.3.0 MOSFET Model—User's Manual.
[4] W. C. Lee, C. M. Hu, IEEE Trans. Electr. Dev. **48** (2001) 1366
[5] X. Xi *et al*, Proceedings ICSSICT, China, pp. 920 (2004)
[6] H. Nakajima *et al*, IEEE Trans. Elect. Dev. **49** (2002) 1775
[7] M. Bescond *et al*, Proceedings of ESSDERC, Portugal, pp. 395 (2003)
[8] H. Kawaura *et al*, IEEE Trans. Elect. Dev. **47** (2000) 856
[9] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. New York: Wiley, pp 520 521 (1981)
[10] R. Holm, J. Appl. Phys. **1** (1951) 569
[11] F. Stern, Phys. Rev. B **5** (1972) 4891
[12] F. F. Fang, W. E. Howard, Phys. Rev. Lett. **16** (1966) 797
[13] F. Stern, W. E. Howard, Phys. Rev. **163** (1967) 816
[14] F. Stern, J. Vac. Sci. Techno. **9** (1972) 752
[15] T. Ando, A. B. Fowler, F. Stern, Rev. Mod. Phys. **54** (1982) 437
[16] R. Van Langevelde, F. M. Klaassen, Solid-State Electr. **44** (2000) 409
[17] J. He, *et al*, Solid-State Electr. **51** (2007) 433
[18] N. Dasgupta, A. Dasgupta, *Semiconductor devices: Modeling and Technology*, PHI, India (2007)
[19] H. Graff, F. Klassen, *Compact Transistor modeling for circuit design*, Springer (1990)
[20] S. Oda, D. Ferry, *The scaling limit of MOSFETs due to direct source –drain tunneling in silicon nanoelectronics*, Taylor and Francis, Boca Raton (2006)