



Viability of Low Temperature Deep and Ultra Deep Submicron Scaled Bulk nMOSFETs on Ultra Low Power Applications

Subhra Dhar¹, Manisha Pattanaik¹, P. Rajaram²

¹Department of Information Technology, VLSI lab,
ABV-Indian Institute of Information Technology and Management Gwalior Gwalior-474010, M.P., India
sdharmos@gmail.com

²Department of Physics Jiwaji University, Gwalior-474011, M.P., India

Received 1/07/2011, online 17/10/2011

Abstract

Chip cooling is an attractive option for leakage control and power as well as thermal management of high performance ICs. Subthreshold leakage being the main leakage contributor in nanoscale CMOS, it rapidly increases with scaling due to continuous reduction in the supply voltage and is highly temperature sensitive. The authors in this work investigate Si bulk nMOSFETs using both constant voltage scaling as well as constant field scaling rules where temperature is employed as a design variable to improve subthreshold characteristics in the deep and ultradeep submicron regions of operation. Devices of gate lengths 90nm, 65nm in the deep submicron, and 45nm, 32nm in the ultradeep submicron region are designed using 2D Silvaco ATLAS device simulator. Encouraging subthreshold characteristics in the 150K-300K range are achieved. The optimum temperature at which these scaled devices would produce best subthreshold slope and maximum reduction in the OFF state leakage current is also suggested. Minimizing power consumption than pure subthreshold operation being a major concern, standby power is estimated at the optimized temperature to compare the scaling effects. This work recommends the 45nm device to be used for both ULP and HP applications at 200K.

Keywords: scaling, temperature, nMOSFET, deep submicron, subthreshold slope, leakage reduction, leakage power, ultra deep submicron region, cooling

I. INTRODUCTION

Much of the progress in semiconductor integrated circuit technology is attributed to the ability to shrink or scale the devices. As technology scales down into the ultra deep submicron region, the static power dissipation grows exponentially and become an increasing dominant component of the total power dissipated in CMOS circuits. The dominant mechanism in the OFF mode is the subthreshold leakage current for deep submicron technologies. The subthreshold region is particularly important when the MOSFET is used as a low voltage, low power device such as a switch in digital logic and memory applications. Two factors increase the subthreshold current as the temperature is increased, decrease in threshold voltage and the linear increase of subthreshold slope with temperature [1]. As suggested by various researchers in the last decade, effective static leakage reduction techniques at the device level includes halo doping and retrograde doping, minimizing gate oxide thickness and reduction in temperature. It has been shown that the halo and retrograde doping are not essential for subthreshold device design [2]-[3]. Minimizing oxide thickness to improve subthreshold slope does not provide minimum energy consumption in digital subthreshold operation [3]. Several authors have considered the potential advantages of operating semiconductor devices and integrated circuits at low temperatures. Low temperature operation of silicon MOSFETs is

considered a promising way to improve device as well as circuit performance. Temperature reduction allows a substantial increase of the carrier mobility and saturation velocity, better turn on capabilities, latch up immunity, reduction in activated degradation process, lower power consumption, decrease of leakage current, reduced thermal noise and increase thermal conductivity. Low temperature operation of MOSFETs at 77K by [4] suggests scaling of threshold voltage to get the maximum technological benefits in the submicron region. As power and thermal problems become more critical with technology scaling, device and circuit level power reduction techniques alone is not sufficient due to the tradeoffs between various design metrics that such techniques necessitate. Chip cooling [5] is an attractive option for leakage control and power as well as thermal management of high performance ICs. The impact of cooling should be analyzed at every level of an IC operation. At the device level, although threshold voltage increases at low operating temperatures, cooling provides a net improvement in performance. At the circuit level, higher device drive current capability at lower temperatures enhances circuit performance. Cooled operation benefits back-end performance and reliability. Lower operating temperatures lead to smaller wire resistance per unit length reducing delay in signal lines and static IR-drop in power/ground networks. Intra wire capacitance per unit length can be reduced significantly for smaller aspect

ratio wires while maintaining the same resistance per unit length. At system level, similar to device and circuit level, system performance always improves under cooled operation. Leakage power dissipation decreases significantly as more cooling power is applied and the reduction of leakage power becomes greater as technology scales. Also, lowering the operating temperature in leakage dominant nanometer scale CMOS technologies can reduce overall cost, since the power needed for cooling may be regained from the lower leakage of the cooled devices. Therefore the benefit that can be derived from cooling increases as technology scales. The primary motivation for employing cooling has been the increase in performance due to the improvement of carrier mobility. Mobility increases as temperature decreases mainly because of the reduction of carrier scattering caused by thermal vibrations of the semiconductor crystal lattice. Transistor carrier mobility is a function of electric field, doping concentration and temperature. At less than 77K carrier mobility may be limited by impurity scattering. While MOSFETs can operate down to liquid helium temperature, there is added complexity due to increase in freeze out, cost and inconvenience of refrigeration below 100K. Thus, this paper does not consider operating temperatures below 100K. Impurity freeze out becomes significant for temperatures lower than 100k. This paper considers both the scaling methods at the deep and ultra deep submicron region while the temperature is varied between

150K and 300K to rule out the impurity freeze out, kink phenomena and transient effects that may occur at much lower temperatures [6]. The devices are operated in the subthreshold region to bring out commendable subthreshold characteristics in terms of S values and I_{OFF} . The optimum temperature at which these scaled devices would produce best subthreshold slope and maximum reduction in the OFF state leakage current is also suggested. Static power of the devices thus obtained due to scaling at optimized temperatures is presented at the end of the paper to enable the best possible subthreshold design at these regions.

II. DEVICE STRUCTURE

This work highlights the combined impact of technology scaling along with temperature reduction at the deep and ultra deep submicron region. Few important device design parameters used for the 65nm and 45nm nMOSFETs are as shown in **Table 1**.

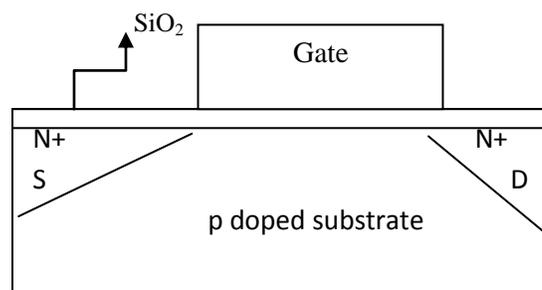


Fig.1 Schematic of planar bulk Si nMOSFET used for simulation

A planar bulk Si nMOSFET device structure (**Fig.1**) is designed such that both constant voltage scaling and constant field scaling rules are applied on all important device design

parameters. The scaled parameters include the supply voltage (V_{dd}), gate oxide thickness (T_{ox}), junction depth (X_j), and the doping concentrations (N_a , N_d). In our study, the physical gate lengths investigated in the deep submicron region are the 90nm and 65nm, whereas it is the 45nm and 32nm in the ultra deep submicron region. $N_{S/D}$ is kept at $1 \times 10^{20} \text{ cm}^{-3}$. Polysilicon gate, SiO_2 insulator is used and no type of halo doping or implants is present in

Table 1

L_g (nm)	T_{ox} , (nm) CVS	X_j (nm) CVS	N_a (cm^{-3}) CVS	V_{dd} (V) CVS	T_{ox} , (nm) CFS	X_j (nm) CFS	N_a (cm^{-3}) CFS	V_{dd} (V) CFS
65	2.17	43.47,	1.9×10^{17}	1.2	2.17	43.47	1.38×10^{17}	0.86
45	1.59	31.5	3.61×10^{17}	1.2	1.59	31.5	2.62×10^{17}	0.62

these devices. Furthermore, in order to minimize short channel effects, it is decided to use a substrate potential of zero volts. DIBL effect remains a serious issue for deep submicron MOSFETs even under cryogenic conditions since it cannot be minimized by temperature reduction. As the channel length reduces, all results are obtained at a low drain bias of 0.1V, to avoid the DIBL effect. CVT Lombardi mobility model is used for simulation and Newton-Gummel method is adopted to solve the equations in this model. The DEVEDIT and DECKBUILD modules of Silvaco simulator is used to obtain all the results.

In **Fig. 2**, the schematic illustration of the scaling of a Si MOSFET by a factor alpha is shown.

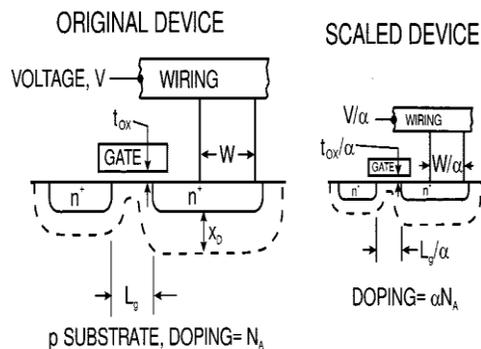


Fig. 2 Schematic illustration of the scaling of a Si MOSFET by a factor alpha [7].

III. RESULTS & DISCUSSION

As CMOS technology advanced, the performance gains due to the use of cryogenic temperatures decreased in magnitude and concerns arose about inferior hot carrier reliability. This dimming of the technical promise of

cryogenic CMOS combined with the need for refrigeration systems in any practical usage scenario, caused this option to be unrealistic and not worthy the trouble. But over the years, CMOS technology has evolved and is more friendlier for cryogenic usage. Several semiconductor devices like the superconducting structures and Josephson devices can operate only at low temperatures. Cryogenic power electronics is of interest to deep space exploration where low temperature is the norm than the exception. Due to the better electronic, electrical, and thermal properties of certain semiconductor materials at low temperatures cryogenic power electronics is expected to have

better efficiency, higher speed, reduced leakage current, and reduced latch-up susceptibility [8].

The conflict between the need to reduce V_{dd} while the basic physics prevents a proportional reduction in V_{TH} can only be addressed by

L_g (nm)	V_{TH} (mv) Const.	S (mv/dec) Voltage	I_{OFF} (A/um) Scaling	V_{TH} (mv) Const.	S(mv/dec) Field	I_{OFF} (A/um) Scaling
90	0.491	70.7	6.94×10^{-7}	0.463	69.4	2.34×10^{-6}
65	0.477	99.5	5.32×10^{-9}	0.457	69.2	4.63×10^{-7}
45	0.419	135.7	2.34×10^{-11}	0.400	78.1	1.94×10^{-12}
32	0.415	115.9	9.44×10^{-14}	0.320	109.9	1.204×10^{-10}

decreasing the device operating temperature. Lower operating temperature provide pure scaling and offer the only straight forward way of coping with the catastrophic headroom loss that threatens to cause the collapse of analog CMOS design techniques in sub 100nm technologies[4]. Digital VLSI circuits often operate at elevated temperatures. Heat generation along with degradation of subthreshold slope with temperature causes the leakage current at $V_g=0$ to increase considerably over its room temperature value. In this work, the device parameters are systematically reduced through constant field scaling and constant voltage scaling. Though the V_{TH} obtained in this study conforms to the lower limit of the OFF current specifications given in [9], the I_{OFF} and S values obtained at 300K (**Table 2**) doesn't appear to be encouraging enough due to various reasons including short channel effects, hot

carrier effects, punchthrough and various leakage mechanisms encountered in this region.

Table 2: Subthreshold characteristics attained due to scaling in terms of S and I_{OFF} at 300K

Weak inversion or subthreshold current between source and drain in a MOS transistor occurs when gate voltage is slightly smaller than V_{TH} . Weak inversion typically dominates in modern device OFF state leakage due to the low V_{TH} that is used in lower geometries. As the temperature is reduced, the V_{TH} increases resulting in exponential decrease in weak inversion current. As the subthreshold leakage current is exponentially dependent on temperature and supply voltage V_{dd} , temperature reduction is considered an effective method to enhance the OFF state characteristics in deep submicron devices. At lower temperatures, the density of thermal acoustic phonons is greatly reduced. These are the dominant scattering mechanism faced by the conducting MOS inversion layer, a reduction in the thermal phonon density leads to significant leakage reduction. In [6] a generalized mobility law has been proposed for

a temperature range 4.2K-300K valid for strong inversion above threshold whereas the room temperature law is applicable for N channel devices down to 100K-150K. It studied in detail the mobility and velocity modeling effects from 300K-4.2K along with parasitic leakage currents and short channel effects. To gain the most performance out of low-temperature CMOS the threshold voltage should be tuned to lower values while maintaining the same OFF-current as the temperature decreases suggests [10]. In [11], the authors have analyzed the impact of technology scaling on the subthreshold current as a function of temperature, but restricted it till 0.13 μm . node and not any further.

In this study, **Fig.3** shows the variation of I_{OFF} with the gate lengths L_g at the operating temperatures of 250K-200K-150K. The plot clearly shows almost steady decreasing trend of I_{OFF} due to constant field scaling as the technology scales down to the ultra deep submicron region. Comparing this to the constant voltage scaling performance in these devices, reduction in temperature cannot be appreciated much as the built in potential doesn't not scale with voltage, also higher fields cause hot electron and oxide reliability problems adding to increased leakage. The reliability of MOS devices is strong function of operating voltage and temperature. Reliability issues include gate oxide integrity, electro migration and hot carriers. Scaling of gate oxide improves the reliability at low temperatures.

Electromigration decreases with decreasing temperature. As the temperature is lowered the carrier mean free path increases due to reduction in thermally generated lattice vibrations. This results in a larger fraction reaching the gate and a higher susceptibility to hot carrier degradation at low temperature [12]. The amount of reduction achieved can be attributed to the reduced reliability issues due to reduction in temperature. Also, the amount of reduction achieved is attributed to lowered (scaled) gate oxide thickness of the devices. In energy constrained design, optimizing T_{ox} helps minimize overlap and fringe capacitances [3]. Moreover, reducing/scaling T_{ox} has other subthreshold benefits like improved subthreshold swing S [**Fig.4**].

Low temperature operation of a MOSFET for sharper subthreshold characteristics for switching off of the transistor being an attractive option, the subthreshold slope indicates how effectively the transistor can be turned OFF when V_{gs} is decreased below V_{TH} . **Fig. 4** shows subthreshold slopes due to impact of reduced temperature on the scaled devices. As the gate length enters the ultra deep submicron region, the subthreshold slope S deteriorates due to the loss of electrostatic control of the gate over the potential barrier in the channel. The static power obtained due to impact of temperature reduction on the leakage current in the scaled deep and ultra deep submicron nMOSFETs is shown in **Fig.5**. Other leakage mechanisms

activated at this region include the punchthrough current. A strong punch through current also degrades the S and V_{th} . The punchthrough condition which arises due to large drain biases in small geometry devices is restricted by applying low drain biases at deep submicron region. As the temperature is lowered, the punchthrough current is reduced.

For very low power applications, reducing supply voltage is a way of reducing power dissipation. At 90nm and 65nm const. voltage scaling proves better, at 45nm constant field scaling proves better due to reduced hot electron effects. The operating temperature of 200K holds good for the 45nm nMOSFET for both the scaling conditions. This device can be applicable for ULP applications if it is field scaled, and can be useful for LSTP and HP applications if the supply voltage is kept constant. Deviation of the 32nm device may be due to impact ionization. The increase of the electric field near the drain region and the increase of current density going through the high field region are key issues for deep submicron devices. The generation hole current due to impact ionization leads to significant increase in substrate current. Because the channel current is significantly increased under technology scaling the substrate leakage component due to impact ionization is increased as well. This device has to be operated below

150K.

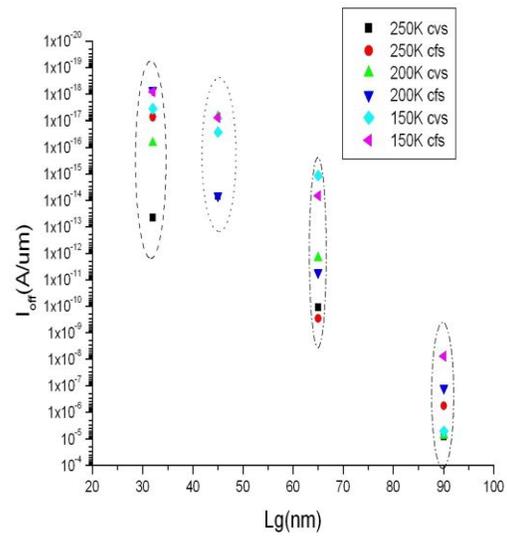


Fig.3. Impact of technology scaling on weak inversion current as a function of temperature

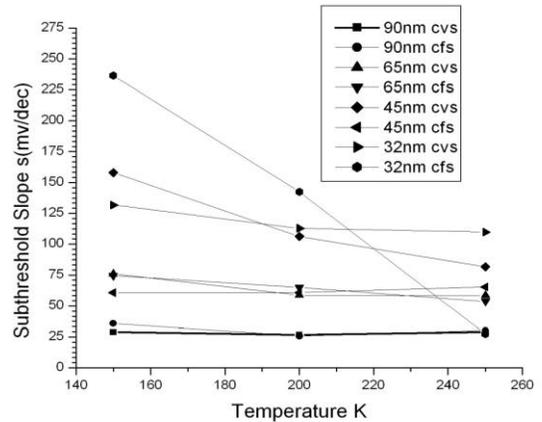


Fig.4 Impact of temperature on subthreshold slope in the deep submicron and ultra deep submicron region

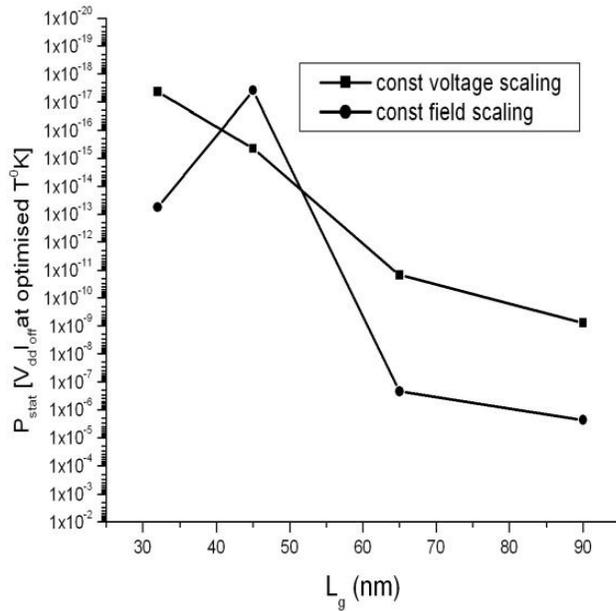


Fig. 5 Effect of scaling on static power in the deep and ultra deep submicron region.

arrives at optimized temperatures at which the investigated devices could give excellent S and I_{OFF} values due to the combined effect of scaling and temperature reduction. When the voltage scaling part is considered, the temperature had to be reduced further to achieve better S values at the 90nm node, whereas further reduction helped achieve better I_{OFF} values at the 32nm node (Table 3). Depending upon the application to be used for, designers may use table 3 to suit the need of reduced leakage current and/or better switching off capacity.

Table 3

Lg nm	Optimized Temp K	S	I_{OFF}	Optimized Temp K	S	I_{OFF}
		CVS	CVS		CFS	CFS
90	135	52.9	6.53×10^{-10}	280	64.5	1.93×10^{-6}
65	250	65.2	1.26×10^{-11}	275	62.9	2.55×10^{-7}
45	200	63	3.72×10^{-16}	200	57.5	6.21×10^{-18}
32	100	80.8	3.55×10^{-18}	225	53.2	1.26×10^{-13}

After taking all the above discussed aspects of subthreshold device design into consideration and rigorous laboratory simulation time, Table 3

IV CONCLUSION

This study embarks the benefits of constant voltage as well as constant field scaling in deep

and ultradeep submicron bulk nMOSFETs and reduces the operating temperature till 150K to achieve better subthreshold characteristics. The simulation results of the impact of scaling of important device parameters and temperature reduction on the subthreshold leakage current, subthreshold slope and static power is presented in this work. Many aspects of deep and ultra deep submicron devices including DIBL, reduction in V_{TH} , punchthrough problems and reliability issues are discussed in this work. The operating temperature of 200K holds good for the 45nm nMOSFET for both the scaling conditions. This device can be applicable for ULP applications if it is field scaled, and can be useful for LSTP and HP applications if the supply voltage is kept constant. In case of industrial limitations of cooling to 100K, constant field scaling is advocated over constant voltage scaling to achieve better subthreshold characteristics. The improved subthreshold characteristics and operation over a large temperature range can make these devices better choice for future low power, low voltage VLSI-ULSI cryogenic applications.

REFERENCES

- 1 Christian Piguet, "Low power CMOS Circuits, Technology, Logic Design and CAD Tools", Taylor and Francis, 3,3-1 (2005)
- 2 Gupta et al "Digital computation in subthreshold region for ultra low power operation", Proc of IEEE, **98**, 168 (2010)
- 3 Ramesh V, S. Dasgupta, and R.P. Agarwal, "Device and circuit design challenges in the digital subthreshold region for ultralow power applications", VLSI Design, Hindawi Publishing Corporation, 2009, **1** (2009)
- 4 Daniel Foty "An Evaluation of Deep Submicron CMOS Design Optimized For Operation At 77k", Analog Integrated Circuits Signal Process, **49**, 97 (2006)
- 5 Sheng-cheng Lin, Kaustav Banerjee, "COOL CHIPS: Opportunities and Implications for Power and Thermal Management "IEEE Trans. on Electron Devices, **55**, 245 (2008)
- 6 G.Ghiboudo, F.Balestra, "Low Temperature Characterization of Silicon CMOS Devices", Microelectronics Reliability, **37**, 1353 (1997)
- 7 David J. Frank, "Device Scaling Limits of Si MOSFETs and their Application Dependencies, Proceedings of the IEEE, **89**, 259 (2001)
- 8 Hue Ye, C.Lee, J.Raynolds, Pradeep Haldar, Michael J.Hennessy, E.K Mueller, "Silicon Power MOSFET At Low Temperatures: A Two Dimensional Computer Simulation Study", Cryogenics **47**, 243 (2007)
- 9 Y.Taur and T.H Ning, "Fundamentals of Modern VLSI Devices" Cambridge University Press ,**4**, 188, (1998)
- 10 Y.Taur, "CMOS Scaling Beyond 0.1um: How Far Can It Go?" VLSI-TSA, **6**, (1999)
- 11 O.Semenov, Arman Vassighi, Manoj Sachdev, "Impact of technology scaling on thermal behavior of leakage current in sub quarter micron MOSFETs: perspective of low temperature current testing", Microelectronics Journal, **33**, 985 (2002)
- 12 William.F.Clark, Badih El-Kareh, Renato G.Pires, Stephan L. Titcomb, Richard L.Anderson, "Low Temp CMOS-Brief Review", IEEE Trans. on components, Hybrids and manufacturing technology,**15**,397 (1992).