Analysis of Power Dissipation for Different SRAM Cells in Nanoscale Technologies

Tanvi Joshi, Aakanksha Pal and S. S. Rathod

Abstract-- **As per the 'International Technology of Roadmap for Semiconductors-2007', high leakage current in nanometer regime is becoming a significant portion of power dissipation in cmos circuits as threshold voltage, channel length and gate oxide thickness are scaled. This paper explores the possibility of reduction in the energy dissipation in different types of SRAM cells.**

Overall leakage in a stack of transistors reduces due to modification of gate to source voltage, threshold voltage and drain induced barrier lowering. This paper evaluates various SRAM cells with and without introducing stacking in nanometer regime. T-spice simulation results shows that compared to the conventional high performance SRAM cells, stacked cells offer significant reduction of power consumption. Comparisons from viewpoint of average power dissipation among 6T SRAM, passive load SRAM, dual ported double ended SRAM, content addressable SRAM, 5T single ended SRAM and dual port single ended SRAM are drawn.

Index Terms-- **MOSFET, Stack Effect, SRAM, Power Reduction, TCAD.**

I. Introduction

LOW power large scale integration of memory technology is an increasing important and growing area of electronics. Although rapid progress has been made in the reduction of power for SRAM subsystems there remains significant potential for future improvement. Low power SRAM technology is achieved by a combination of low power chip technology and low voltage chip to chip technology. SOI (Silicon on Insulator) technology is relatively new introduction in memory design and its substitution can improve performance through reduced junction capacitance and immunity to soft error rates. Ultra low voltage operation is crucial for modern SRAM designs especially for mobile and embedded applications. However, subthreshold current and threshold voltage mismatches developed in the memory cell are major concerns for low voltage operation. Short gate lengths in SOI technology can lead to short channel effects and mismatch between devices in the memory cell can cause instability and disrupt functionality.

State-of-the-art microprocessor designs devote a large fraction of the chip area to memory structures - e.g., multiple levels of instruction and data caches, translation look-aside buffers, and prediction tables [1]. Leakage power management is becoming indispensable for cost effective packaging and cooling solutions for high-end microprocessors. And it is critical for holding time design for battery supported low-end mobile System-On-Chips (SOC). Since leakage power is proportional to the number of on-chip transistors, on-chip LI and L2 caches, which comprise the vast majority of on-chip transistors and represent a sizable fraction of the total power consumption of microprocessors, should be paid much attention from the viewpoint of leakage. Recent power estimation for 100nm process indicates that leakage power accounts for 30% of L1 cache power and as much as 80% of L2 cache power [1]. With technology scaling, leakage power consumption in Static Random Access Memory (SRAM) will become more significant.

A. Stacking Effect

It has been shown that the stacking of two off transistors can significantly reduce leakage power than a single off transistor [7]. Increasing the source voltage of NMOS transistor reduces subthreshold leakage current exponentially due to negative Vgs, lowered signal rail, reduced DIBL and body effect. This effect is also called self-reverse biasing of transistor. The self-reverse bias effect can be achieved by turning off a stack of transistors. Turning off more than one transistor in a stack raises the internal voltage (source voltage) of the stack, which acts as reverse biasing the source. The voltages at the internal nodes depend on the input applied to the stack [4, 5, 6, 8].

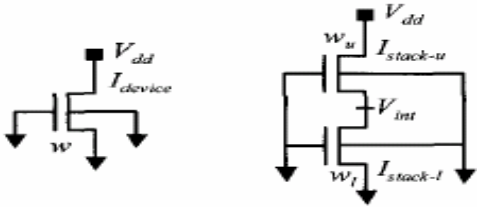

Fig. 1. Stacking of nMOS

For a two-device stack shown in Fig. 1, a steady state condition will be reached when the intermediate node voltage *Vint* approaches Vx such that the leakage currents in the upper and lower devices are equal. Under this condition, the leakage

Tanvi Joshi is with Department of Electronics Engineering at Sardar Patel Institute of Technology, Mumbai 400058 INDIA. (e-mail:sinv_by_cosv@yahoo.com).

Aakanksha Pal is with Department of Electronics Engineering at Sardar Patel Institute of Technology, Mumbai 400058 INDIA. (e-mail:aakankshapal@yahoo.com).

S. S. Rathod is with Department of Electronics Engineering at Sardar Patel Institute of Technology, Mumbai 400058 INDIA. (e-mail: rathod_spce@yahoo.com).

currents in the upper and lower devices can be expressed as [7],

$$I_{stack-u} = w_u I_1 \quad 10^{\frac{-(1+\lambda_d+k_\gamma)V_x}{S}}$$

$$I_{stack-l} = w_l I_1 \quad 10^{\frac{-\lambda_d(V_{dd}-V_x)}{S}}$$

and the intermediate node voltage will be,

$$V_x = \frac{\lambda_d V_{dd} + S\log\frac{w_u}{w_l}}{1+k_\gamma+2\lambda_d}$$

Where S is the sub-threshold swing, λ_d is the drain-induced barrier lowering (DIBL) factor, and k_γ is the body effect coefficient. The above equation assumes that the resulting

$$Vds > 3kT/q$$

For short channel devices the body terminal's control on the channel is negligible compared to gate and drain terminals, implying $k\gamma << 1 + 2\lambda_d$. Hence the steady state value, Vx of the intermediate node voltage can be approximated as,

$$V_x \approx \frac{\lambda_d V_{dd} + S\log\frac{w_u}{w_l}}{1+2\lambda_d}$$

Substituting Vx in either $I_{stack-u}$ or $I_{stack-l}$ will yield the leakage current in a two-stack given by,

$$I_{stack} = w_u^{\alpha} w_l^{1-\alpha} I_1 \quad 10^{\frac{-\lambda_d V_{dd}(1-\alpha)}{S}}$$

$$\text{where} \quad \alpha \approx \frac{\lambda_d}{1+2\lambda_d}$$

Positive potential at the intermediate node has three effects:

1) gate-to-source voltage of upper transistor becomes negative. As the subthreshold current is exponentially proportional to V_{GS}, it is also reduced.

2) negative body-to-source potential of upper transistor causes more body effect and V_T is increased. Since the subthreshold current is exponentially proportional to V_T also, it is reduced.

3) drain-to-source potential of upper transistor decreases, resulting in less drain-induced barrier lowering (DIBL). As a result the subthreshold leakage is further reduced.

This phenomenon is the "***stacking effect***."

From BSIM2 MOS transistor model [9], subthreshold current flowing through "off" transistor is given by

$$I_{sub} = A\, e^{\frac{q}{nkT}(V_{gs}-V_{th0}+\gamma V_{bs}+\eta V_{ds})}\left(1-e^{\frac{-qVds}{kT}}\right)$$

It is observed that a negative Vgs, an increase in the body effect (negative Vbs), and a reduction in Vds (less DIBL) reduce the subthreshold current exponentially.

It is possible to facilitate delay-leakage trade-off by increasing the channel length of devices that are in non-critical paths. To maintain iso-input load the channel width will have to be reduced along with increase in the channel length [2]. Since we assume that the forced stack technique breaks each existing transistor into two half sized transistors the resistance of each transistor of the forced stack technique is doubled, compared to the standard inverter; furthermore, in this way, we can maintain input capacitance equal [3].

II. DIFFERENT TYPES OF PROPOSED STATIC RAM

This section describes six different types of static RAM memories and there proposed stack circuit. These are constructed using T-Spice [11].

A. 6T SRAM

The CMOS SRAM cell [10] consists of two load devices and two storage transistors together with two access transistors as shown in fig. 2. The access and storage transistors are enhancement type nMOS. The load device offsets the charge leakage at the drains of the storage and select transistors. When the load transistor is pMOS, the resulting six transistors CMOS cell as shown in fig. 1 has only leakage current through the cell except during switching. Either nMOS or pMOS is always off. To reduce the power dissipation, 6T SRAM cell with stack effect is proposed as shown in fig.3.

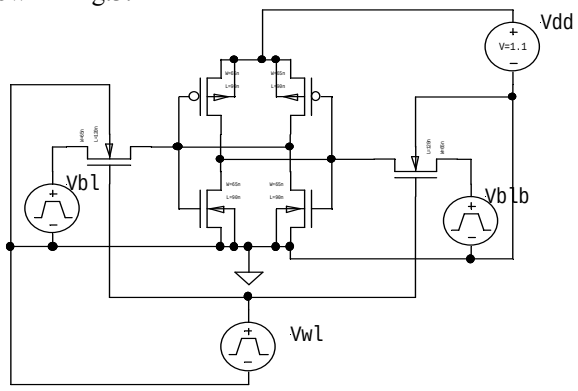

Fig. 2. Conventional 6T SRAM

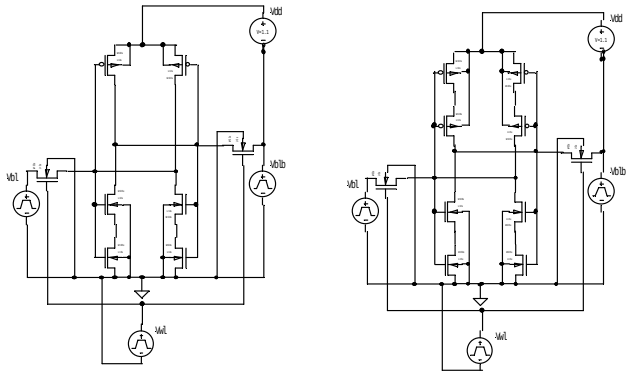

Fig. 3. Proposed Half and Full Stack 6T SRAM

B. Passive Load SRAM

The standby power dissipation is higher for a resistive load cell than a full CMOS six transistor cell [10]. One of the disadvantages using a pMOS load in a SRAM cell is, difficulty to write into the cell. If a double polysilicon process is used, the load resistor cell as shown in fig. 3 offers approximately 30% reduction in cell size over that of the pMOS load cell with double polysilicon technology for the resistor. The cell has lower performance, but the reduced die area of this cell provides a major cost benefit. To reduce the power passive load SRAM with stack effect is proposed as shown in fig.5.

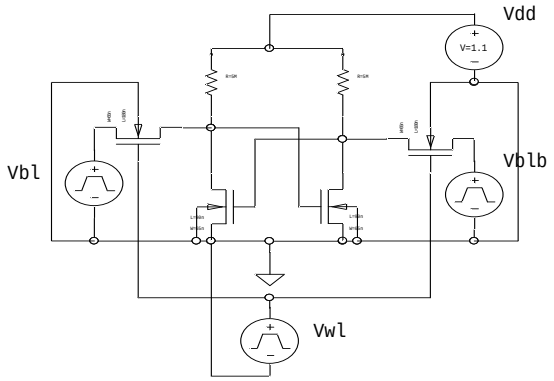

Fig. 4 Passive load SRAM

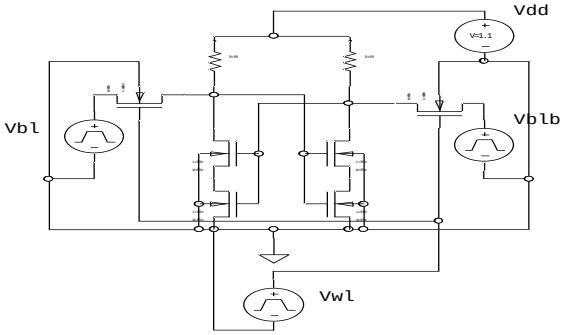

Fig. 5. Proposed Passive load SRAM with stack

C. Dual Ended Double Port SRAM

The dual port cell is useful as an embedded cache memory in microprocessors since it can be accessed simultaneously through both ports [10]. Simultaneous access eliminates wait states for the microprocessor unless one port wants to read while the other is writing, or both ports intend to write simultaneously. The proposed dual ended double port SRAM with stack is as shown in fig. 7.

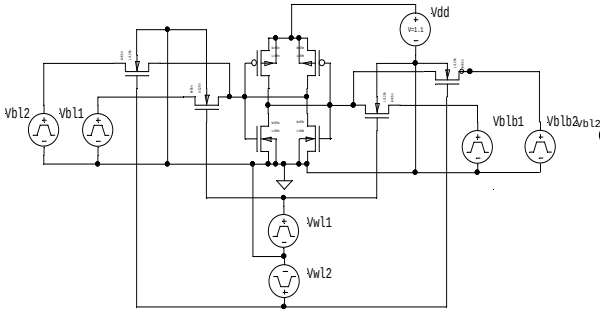

Fig. 6. Dual ended double port SRAM

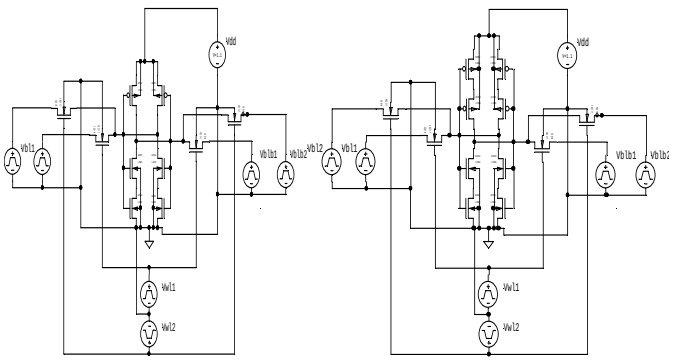

Fig. 7. Proposed Dual ended double port SRAM with stack

D. Content Addressable SRAM

The content addressable or associative memory cell is used in applications where both the contents and the location of the cell are required [10]. During normal operation, read and write is performed like a normal cell. During comparison operation, DATA is placed on BIT bar and DATA bar is place on the BIT lines. If the data match those in the cell, then the match transistor will stay OFF. If any cell has data that does not match, the match transistor pulls a previously pre-charged 'match' low. The proposed content addressable SRAM with stack is as shown in fig. 9.

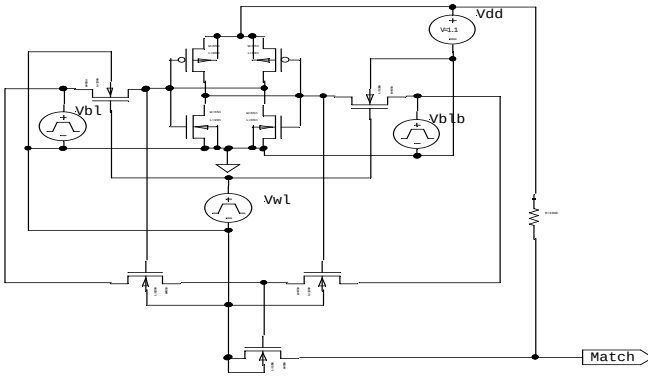

Fig. 8. Content Addressable SRAM

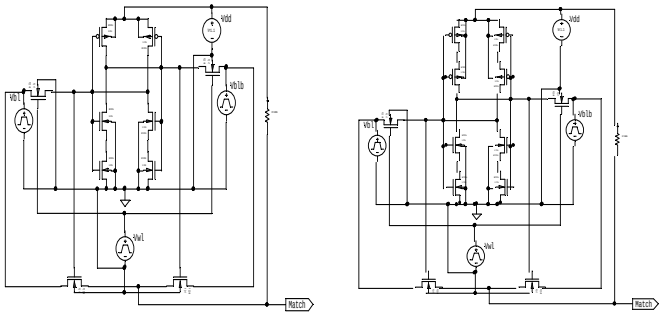

Fig. 9. Proposed Content Addressable SRAM with stack

E. Single Access Transistor 5T-SRAM

Fig. 9 shows a single ended five transistor full static CMOS cell, which can replace the basic six transistor cell where a smaller array is needed [10]. It is not widely used because of lower operating margins and difficulty in write operations. The proposed single access transistor 5T-SRAM with stack is as shown in fig. 11.

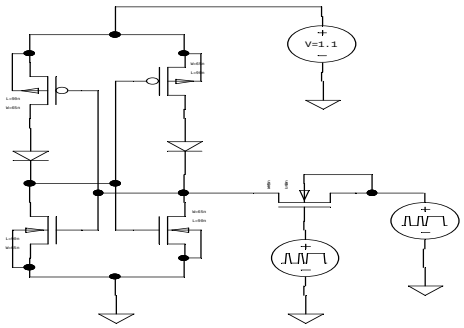

Fig. 10. Single Access Transistor 5T-SRAM

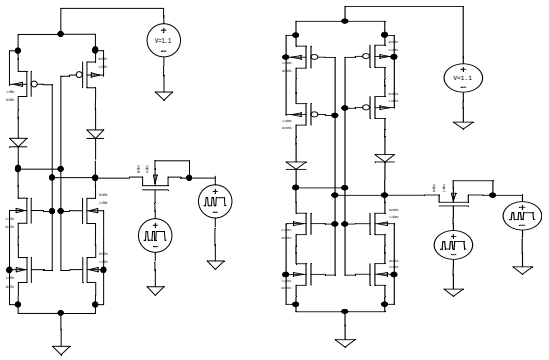

Fig. 11. Proposed Single Access Transistor 5T-SRAM with stack effect

F. Dual Port Single Ended SRAM

Fig. 11 shows the dual port single ended SRAM [10] and fig. 12 shows the proposed dual ended double port SRAM with stack.

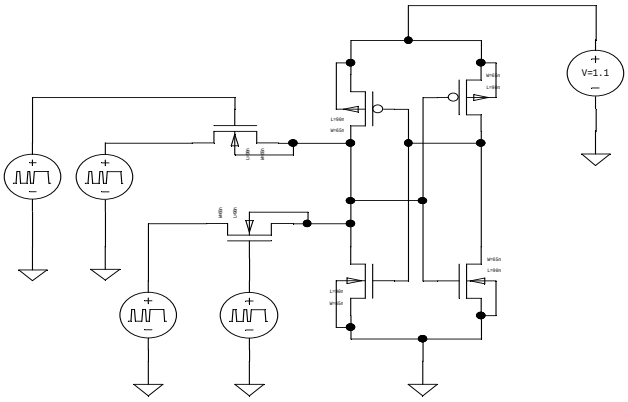

Fig. 12. Dual Port Single Ended SRAM

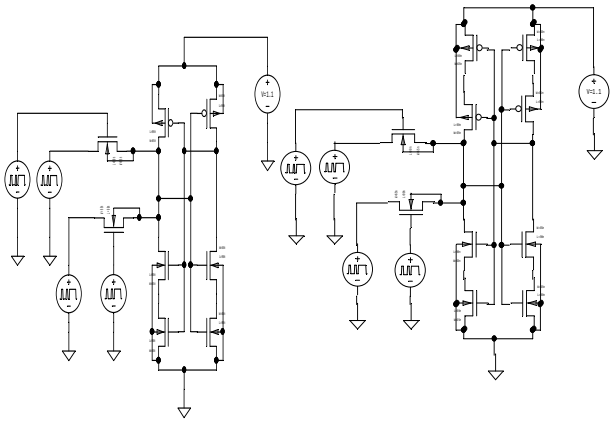

Fig. 13. Proposed Dual Port Single Ended SRAM with stack effect

III. RESULTS AND DISCUSSION

All types of static SRAM discussed in section II are constructed using T-Spice [11] and their performance is noted for power and current. Tables I to VI shows the comparison for power and current for various types of static SRAM under discussion. Tables also show the variation in device parameters of basic cell and access transistor. From these results it is clear that the average power is reduced by the introduction of the stack.

TABLE I
RESULTS FOR 6T-SRAM

SRAM 6 transistor	conventional	Half stack	Full stack
Power	Avg=3.73843uW	2.025503uW	0.1011452uW
	Max=5.27937uW	4.690279uW	0.374269uW
	Min=3.118220uW	1.856209uW	0.382669uW
Current	Avg=1.5353uA	1.0933uA	97.635nA
	Max=3.1030uA	2.5309uA	8.3221nA
Basic cell: L=90n W=65n W/L=.722 Access transistor: L=120n W=65n W/L=0.5416			

TABLE II
RESULTS FOR PASSIVE LOAD SRAM

SRAM passive load	conventional	Half stack
Power	Avg 33.88556uW	25.28322uW
	Max=39.51317uW	35.26691uW
	Min=2.529451uW	2.454577uW
Current	Max=15.567uA	19.022uA
	Avg=3.1295uA	0.75122uA
Basic cell: L=90n W=65n W/L=.722 Access transistor: L=180n W=65n W/L=20769		

TABLE III
RESULTS FOR DUAL ENDED DOUBLE PORT SRAM

Dual ended double port	conventional	Half Stack	Full Stack
Power	Avg=7.427069uW	5.590653uW	3.723603uW
	Max=12.21446uW	11.20991uW	7.170344uW
	Min=6.831008uW	5.145751uW	3.346694uW
Current	Max=2.4458uA	1.0311uA	0.72427uA
	Avg=1.7599uA	1.0311uA	0.10910uA
Basic cell: L=90n W=65n W/L=.722 Access transistor: L=120n W=65n W/L=0.5416			

TABLE IV
RESULTS FOR CONTENT ADDRESSABLE SRAM

Content addressable	conventional	Half Stack	Full Stack
Power	Avg=3.162810uW	0.1164642uW	4.805e-008
	Max= 3.162810uW	3.908278uW	2.877992uW
	Min=2.205103uW	2.706362e-015	7.5939e-014
Current	Max=2.4772uA	1.4868uA	1.1404uW
	Avg=1.4855uA	6.6633e-008	2.4782e-008
Basic cell: L=90n W=65n W/L=.722 Access transistor: L=120n W=65n W/L=0.5416			

TABLE V
RESULTS FOR SINGLE ACCESS TRANSISTOR 5T-SRAM

5Tsingle access	Conventional	Half Stack	Full Stack
Power	Avg=3.162810e-006	1.164642e-007	4.805006e-008
	Max=3.162810e-006	3.908278e-006	2.877992e-006
	Min=2.205103e-006	2.706362e-015	7.593990e-014
Current	Max=2.4772e-006	1.4868e-006	1.1404e-006
	Avg=1.4855uA	6.6633e-008	2.4782e-008
Basic cell: L=90n W=65n W/L=.722 Access transistor: L=120n W=65n W/L=0.5416			

TABLE VI
RESULTS FOR DUAL PORT SINGLE ENDED SRAM

Dual port single ended	conventional	Half Stack	Full Stack
Power	Avg=3.943209e-006	2.814843e-006	2.631593e-006
	Max=5.896564e-006	5.306037e-006	4.857101e-006
	Min=3.872188e-006	2.628264e-006	1.695409e-006
Current	Max=3.4189e-006	2.6758e-006	2.6401e-006
	Avg=1.6880e-006	8.4879e-007	7.9563e-007
Basic cell: L=90n W=65n W/L=.722 Access transistor: L=120n W=65n W/L=0.5416			

IV. Conclusions

From various simulations it is found that stack technique helps in the reduction of power consumption and hence it can be used for the design of low power SRAM circuits. This reduction of power consumption is at the cost of increased silicon area and propagation delay. A calculation of static noise margins of all types of stacked SRAM is yet to be done and is considered for the future course of action.

V. Acknowledgment

The authors gratefully acknowledge the facilities provided in VLSI Laboratory of Electronics Engineering Department, Sardar Patel Institute of Technology, Mumbai.

VI. References

[1] C. H. Kim, I. J. Kim, S. Mukhopadhyay, and K. Roy, " Forward Body-biascd-low-leakage SRAM Cache: Device and Architecture Considerations," *Proc. International Symposium Low Power Electronics and Design,* p.p. 69-70, 2003.

[2] D. Dobberpuhl, "The Design of a High Performance Low power Microprocessor", *Proc. International Symposium Low Power Electronics and Design*, p.p. 11-16, 1996

[3] Jun Cheol Park and Vincent J. Mooney, "Sleepy Stack Leakage Reduction", IEEE trans. On VLSI systems, Vol. 14, No. 11, Nov. 2006.

[4] N. Kim, T. Austin, D. Baauw, T. Mudge, K. Flautner, J. Hu, M. Irwin, M. Kandemir, and V. Narayanan, "Leakage current: Moore's Law meets static power", IEEE Comput., vol. 36, no. 12, 68–75, Dec. 2003.

[5] K. Usami and M. Horowitz, "Clustered voltage scaling techniques for low-power Design", ISLPED, 1995.

[6] Q. Wang and S. Vrudhula, "Algorithms for minimizing standby power in deep submicron, dual-Vt CMOS circuits", IEEE Transactions on CAD, vol.21, 306-318, 2002.

[7] S. Narendra, S.Borkar, V. De, D.Antoniadis and A.Chandrakasan, "Scaling of Stack Effect and its Application for Leakage Reduction", IEEE Proc. of Low Power Electronics and Design, 195-200, 2001.

[8] Amit Agarwal,Saibal Mukhopadhyay and Kaushik Roy, "Leakage in Nano-Scale Technologies: Mechanisms, Impact and Design Considerations", DAC 2004, June 7-11, 6-11, 2004

[9] B. J. Sheu, D. L. Scharfetter, P. Ko, and M. Jeng, "BSIM: Berkeley short-channel IGFET model for MOS transistors", IEEE J. Solid-State Circuits, vol. 22, 558–566, Aug. 1987.

[10] Andrew Marshall and Sreedhar Natarajan, "SOI Design: Analog, Memory and Digital Techniques", vol. I., 2nd ed, London: Kluwer, 2002, p. p. 181-229.

[11] "T-Spice User Manual", Tanner EDA Inc, USA, 2007

VII. Biographies

Tanvi Joshi (M'2007) was born in Mumbai in India on 1986. She is graduate student of Sardar Patel College of Engineering, Mumbai. She won the 4th award in Technical Paper Presentation organized by ESA,06. She was the chief editor of Electronics Students Association. She is also a member of IEEE. Her area of interest includes microelectronics, VLSI design and embedded systems.

Aakansha Pal (M'2007) was born in Calcutta in India on 1986. She is graduate student of Sardar Patel College of Engineering, Mumbai. She won the 4th award in Technical Paper Presentation organized by ESA,06. She is also a member of IEEE. Her area of interest includes microelectronics, VLSI design and embedded systems.

S. S. Rathod (M'2007) was born in Amravati in India on Feb 28, 1975. He graduated from the College of Engineering Badnera, and completed his post graduation in the field of Electronics Engineering from V.J.T.I., Mumbai-India. Currently he is pursuing his doctoral research at Indian Institute of Technology, Roorkee-India.

His employment experience includes eight years as an educationalist. He has published more that 20 papers in various national and international conferences. His special fields of interest include VLSI Design, process and device modeling. His name is listed in the science category of Marques Who's Who USA. He received outstanding achieving ward by the Energy Society, India. He is member of IEEE, ISTE and ISNT.