Leakage Minimization of SRAM Cells in a Dual-Vt and Dual-Tox Technology

Behnam Amelifard, Farzan Fallah, *Member*, Massoud Pedram, *Fellow, IEEE*

Abstract **— Aggressive CMOS scaling results in low threshold voltage and thin oxide thickness for transistors manufactured in deep submicron regime. As a result, reducing the subthreshold and tunneling gate leakage currents has become one of the most important criteria in the design of VLSI circuits. This paper presents a method based on dual-V_t and dual-T_{ox} assignment to reduce the total leakage power dissipation of SRAMs while maintaining their performance. The proposed method is based on the observation that read and write delays of a memory cell in an SRAM block depend on the physical distance of the cell from the sense amplifier and the decoder. Thus, the idea is to deploy different configurations of six-transistor SRAM cells corresponding to different threshold voltage and oxide thickness assignments for the transistors. Unlike other techniques for low-leakage SRAM design, the proposed technique incurs neither area nor delay overhead. In addition, it results in a minor change in the SRAM design flow. The leakage saving achieved by using this technique is a function of the values of the high threshold voltage and the oxide thickness, as well as the number of rows and columns in the cell array. Simulation results with a 65nm process demonstrate that this technique can reduce the total leakage power dissipation of a 64×512 SRAM array by 33% and that of a 32×512 SRAM array by 40%.**

Index Terms**—Low-power design, static random access memory (SRAM), subthreshold leakage, tunneling gate leakage, multiple V_t, multiple T_{ox}**

I. Introduction

CMOS scaling beyond the 90nm technology node requires not only very low threshold voltages (V_t) to retain the device switching speeds, but also ultra-thin gate oxides (T_{ox}) to maintain the current drive and keep threshold voltage variations under control when dealing with short-channel effects [1]. Low threshold voltage results in an exponential increase in the subthreshold leakage current, whereas ultra-thin oxide causes an exponential increase in the tunneling gate leakage current. The leakage power dissipation is roughly proportional to the area of a circuit. Since in many processors caches occupy about 50% of the chip area [2], the leakage power of caches is one of the major sources of power consumption in high performance microprocessors.

While one way of reducing the subthreshold leakage is to use higher threshold voltages in some parts of a design, to suppress tunneling gate leakage, high-*k* dielectrics or multiple gate oxides may be used. In [3, 4] a comparative study of using high-*k* dielectric and dual oxide thickness on the leakage power consumption has been presented and an algorithm for simultaneous high-*k* and high-T_{ox} assignment has been proposed. Although some investigation has been done on Zirconium- and Hafnium-based high-*k* dielectrics [5], there are unresolved manufacturing process challenges in way of introducing high-*k* dielectric material under the gate (e.g., related to the compatibility of these materials with Silicon [6] and the need to switch to metal gates); hence, high-*k* dielectrics are not expected to be used before 45nm technology node [5, 7], leaving multiple gate oxide thicknesses as the one promising solution to reduce tunneling gate leakage current at the present time.

There are different ways to achieve a higher threshold voltage [8], chief among them are adjusting the channel doping concentration and applying a body bias. To achieve multiple oxide thicknesses, on the other hand, Arsenic can be implanted into the Silicon substrate before thermal oxidation is done [9].

In the past, much research has been conducted to address the problem of leakage in SRAMs. In [10], for example, the authors used a dynamically controlled sleep transistor to reduce the leakage power dissipation of a large on-chip SRAM. In [11], a dynamic threshold voltage method to reduce the leakage power in SRAMs has been utilized. In that technique, the threshold voltage of the transistors of each cache line is controlled separately by using forward body biasing. In [12], on the other hand, by observing the fact that in ordinary programs most of the bits in data-cache and instruction-cache are zero, the authors proposed using asymmetric SRAM cells to reduce the subthreshold leakage. Dynamic resizable instruction caches [13], leakage biased bit-lines [14], and dynamic power gating [13, 15, 16] are other effective techniques for reducing the leakage power in SRAMs.

Although many techniques have been proposed to address the problem of low-leakage SRAM design, most of them address only the standby leakage power consumption, while it is known that in sub-100nm designs, active leakage comprises more than 20% of the total active power dissipation in

Manuscript received January 15, 2007; revised May 21, 2007. This research was funded in part by a grant from the National Science Foundation.

B. Amelifard and M. Pedram are with the Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, CA 90089 USA (e-mail: amelifar@usc.edu; pedram@ceng.usc.edu).

Farzan Fallah is with Fujitsu Laboratories of America, Sunnyvale CA 94085 USA (e-mail: farzan.fallah@us.fujitsu.com).

memories [17]. On the other hand, many of these techniques result in hardware overhead and hence increase chip's area and reduce the manufacturing yield. Furthermore, many of them try to reduce the subthreshold leakage current only, whereas for sub-100nm technology node, the tunneling gate leakage is comparable to the subthreshold leakage. In this paper we present a method for reducing both subthreshold and tunneling gate leakage current of an SRAM by using different threshold voltages and oxide thicknesses for transistors in an SRAM cell. The idea is to deploy different configurations of six-transistor SRAM cells corresponding to different threshold voltage and oxide thickness assignments for the transistors [18, 19]. We show that our heterogeneous cell SRAM (HCS) technique has several main advantages over previous techniques in that it:

- reduces both active and standby leakage current including subthreshold and tunneling gate leakage components,
- has no hardware or delay overheads,
- requires only a minor change in the SRAM design flow, and
- has the ability to improve the static noise margin under process variations.

The remainder of this paper is organized as follows. In Section II the SRAM design and operation is discussed and leakage components are briefly described. Our idea for reducing the leakage power dissipation is presented in Section III. Section IV is dedicated to the experimental results, while Section V concludes the paper.

II. Preliminaries

A. SRAM Architecture

A typical SRAM block consists of cell arrays, address decoders, column multiplexers, sense amplifiers, I/O, and a control unit. In the following, the functionality and design of each component is briefly discussed.

1) SRAM Cell

Fig. 1 shows a 6-transistor (6T) SRAM cell. In an SRAM cell, the pull-down NMOS transistors and the pass-transistors reside in the read path. The pull-up PMOS transistors and the pass-transistors, on the other hand, are in the write path. Traditionally all cells used in an SRAM block are identical (i.e., corresponding transistors have the same width, threshold voltage, and oxide thickness) which results in identical leakage characteristic for all cells. However, as we will show in this paper, by using non-identical cells, which have the same layout footprint, one can achieve more power efficient designs.

2) Cell Array

The size of the cell array depends on both performance and density requirements. Generally speaking, as technology shrinks, cell arrays are moving from tall to wide structures [10] [20]. However, since wider arrays need more circuitry for column multiplexers and sense amplifiers, if a small area overhead is desirable (e.g., large L3 caches), the number of rows is kept high [21] [22].

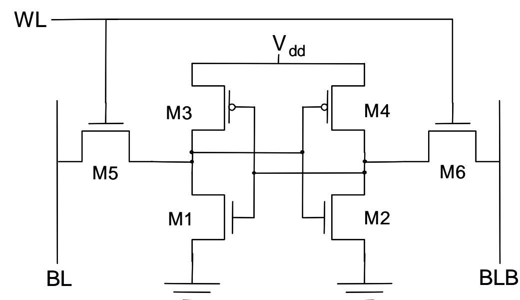

Fig. 1. A 6T SRAM cell

3) Address Decoder

Although the logical function of an address decoder is very simple, in practice designing it is complicated because the address decoder needs to interface with the core array cells and pitch matching with the core array can be difficult [23]. To overcome the pitch-matching problem and reduce the effect of wire's capacitance on the delay of the decoder, the address decoder is often broken into two pieces. The first piece, called pre-decoder, is placed before the long decoder wires and the second part, row decoder, which usually consists of a single NAND gate and buffers for driving the word-line capacitance, is pitch-matched and placed next to each row as shown in Fig. 2.

4) Column Multiplexers and Sense Amplifiers

Column multiplexing is inevitable in most SRAM designs because it reduces the number of rows in the cell array and as a result increases the speed. Since during a read operation one of the bit or bit-line is partially discharged, a sense amplifier is used to sense this voltage difference between bit and bitbar lines to create a digital voltage. To make the circuit more robust to noise, the sense amplifier is typically switched when the voltage difference between bit and bit-bar lines becomes 100-200mV.

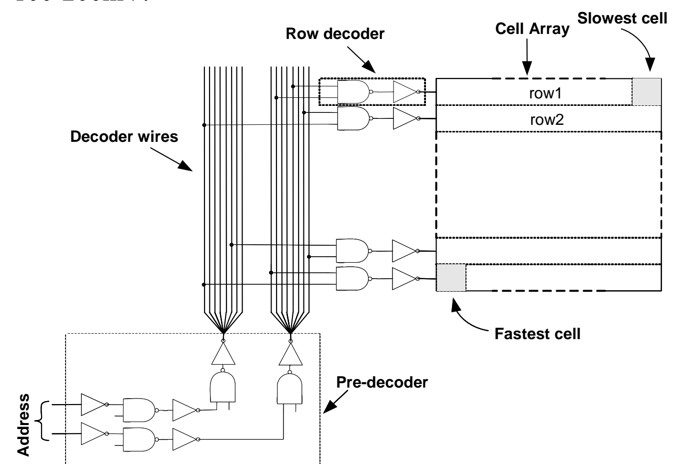

Fig. 2. An SRAM block with its decoder

5) Control Unit

The control unit generates internal signals of the SRAM, including the write and read enable signals, the pre-charge signal, and the sense amplifier enabler.

B. Leakage Current Components

The leakage current of a deep submicron CMOS transistor consists of three major components: junction tunneling current, subthreshold current, and tunneling gate current [24]. In this section, each of these three components is briefly described. Junction Tunneling Leakage

1) Tunneling Junction Leakage Current

The reversed biased P-N junction leakage has two main components: one corresponds to the minority carriers' diffusion near the edge of the depletion region and the other is due to electron-hole pair generation in the depletion region of the reverse biased junction [24]. The tunneling junction leakage current is an exponential function of the junction doping and reverse bias voltage across the junction. Since tunneling junction leakage current is quite small compared to other sources of leakage in state-of-the-art CMOS devices [24], we do not consider this component of leakage in the 6T SRAM cell.

2) Subthreshold Leakage Current

Subthreshold leakage is the drain-source current of a transistor when the gate-source voltage is lower than the threshold voltage. The subthreshold leakage is modeled as [24],

$$I_{sub} = A_{sub} \exp\left(\frac{q}{n'kT}\left(V_{GS} - V_{t0} - \gamma' V_{SB} + \eta V_{DS}\right)\right) \times \left(1 - \exp\left(-\frac{q}{kT} V_{DS}\right)\right) \quad 1$$

where $A_{sub}=\mu_0 C_{ox}\, W/L_{eff}(kT/q)^2 e^{1.8}$, μ_0 is the zero bias mobility, C_{ox} is the gate oxide capacitance per unit area, W and L_{eff} denote the width and effective length of the transistor, k is the Boltzmann constant, T is the absolute temperature, and q is the electrical charge of an electron. In addition, V_{t0} is the zero biased threshold voltage, γ' is the linearized body-effect coefficient, η denotes the drain-induced barrier lowering (DIBL) coefficient, and n' is the subthreshold swing coefficient of the transistor.

There are two dominant subthreshold leakage paths in a 6T SRAM cell: 1) V_{dd} to ground paths inside the SRAM cell and 2) the bit-line (or bit-bar line) to ground path through the pass transistor. To reduce the first type of leakage, the threshold voltages of the pull-down NMOS transistors and/or pull-up PMOS transistors can be increased, whereas to lower the second type of leakage, the threshold voltages of the pull-down NMOS transistors and/or pass transistors can be increased. If the threshold voltage of the pull up PMOS transistors is increased, the write delay increases while the effect on the read delay would be negligible. On the other hand, if the threshold voltage of the pull down NMOS transistors is increased, the read delay increases while the effect on the write delay would be marginal. By increasing the threshold voltage of the pass transistors, both read and write delays increase.

3) Tunneling Gate Leakage Current

Electron tunneling from the conduction band, which is only significant in the accumulation region, results in direct tunneling gate leakage current in NMOS transistors. In PMOS transistors, on the other hand, hole tunneling from the valence band results in the tunneling gate leakage current.

The tunneling gate current is composed of three main components: (1) gate-to-source and gate-to-drain overlap current, (2) gate-to-channel current, part of which goes to the source while the remainder goes to the drain, and (3) gate-to-substrate current. In CMOS technology, the gate-to-substrate leakage current is several orders of magnitude lower than the overlap tunneling and gate-to-channel current [6]. On the other hand, while the overlap tunneling current dominates the gate leakage in the OFF state, gate-to-channel tunneling dictates the gate current in the ON state of the transistor. Since the gate-to-source and gate-to-drain overlap regions are much smaller than the channel region, the tunneling gate current in the OFF state is much smaller than that in the ON state [6].

If SiO_2 is used for the gate oxide, PMOS transistors will have about one order of magnitude smaller gate leakage than NMOS transistors [6] [25]. Therefore, one may conclude that the major source of tunneling gate leakage in CMOS circuits is the gate-to-channel tunneling current of the ON NMOS transistors which can be modeled as [26],

$$J_{tunnel} = \frac{4\pi m^* q}{h^3}(kT)^2\left(1 + \frac{\gamma kT}{2\sqrt{E_B}}\right) \times \exp\left(\frac{E_F}{kT} - \gamma\sqrt{E_B}\right) \quad 2$$

where m^*(=0.19M_0) is the electron transfer mass and M_0 is the electron rest mass. Moreover, h is Planck's constant, E_F is the Fermi level at the Si/SiO_2 interface, E_B is the height of barrier, and γ is defined as,

$$\gamma = \frac{4\pi T_{ox}\sqrt{2m_{ox}}}{h} \quad 3$$

where m_{ox}(=0.32M_0) is the effective electron mass in the oxide.

The major contributor to the tunneling gate leakage current in a 6T SRAM cell is the gate-to-channel leakage of the ON pull-down transistor. To weaken this leakage path, one needs to increase the gate-oxide thickness of the pull-down transistors. To reduce other (minor) tunneling gate leakage currents in the SRAM cell, one only needs to increase the gate oxide thickness of the pass transistors, because from the above discussion, it can be concluded that the gate leakage saving achieved by increasing the oxide thickness of the PMOS transistors would be quite small. Increasing the oxide thickness of a transistor not only increases the threshold voltage, but also reduces the drive current of the transistor. So, the effect of applying this technique to an SRAM cell is an increase in the read/write delay of the cell.

Based on the above discussion, the leakage currents of an SRAM cell storing “0” are the ones shown in Fig. 3.

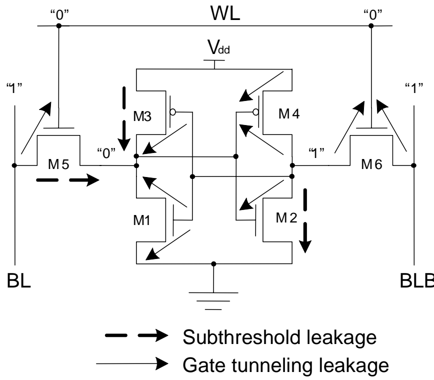

Fig. 3. Subthreshold and tunneling gate leakage of an SRAM cell storing “0”

III. MIXED CELL SRAM

Due to the non-zero delay of the interconnects of the address decoder, word-lines, bit-lines, and the column multiplexer, read and write delays of different cells in an SRAM block are different. Simulations show that for typical SRAM blocks, depending on the number of rows and columns, the read time of the closest cell to the address decoder and the column multiplexer may be 5-15% less than that of the furthest cell from the address decoder and the column multiplexer. This provides an opportunity to reduce the leakage power consumption of an SRAM by increasing the threshold voltage or oxide thickness of some of the transistors in the SRAM cells. The resulting SRAM is called “heterogeneous cell SRAM” (HCS). In this section, it is shown how to design an HCS without degrading the performance or robustness.

A. Technology

All results presented in this section are obtained by HSPICE [27] simulations using a predictive 65nm technology model [28] with 1.1V for the supply voltage, 0.18V for the threshold voltage, and 12°A as the gate oxide thickness. Moreover, unless otherwise stated, it is assumed that the value of the high threshold voltage is 0.28V and the value of the thicker gate oxide is 14°A. The SiO_2 layer in the gate stack is assumed to be 2A^{o} thicker than the thin oxide so as to achieve one order of magnitude reduction in tunneling gate leakage. All simulations are performed at a die temperature of 100 ^{o}C.

The SRAM module used in these simulations is a pre-designed 64Kb SRAM with a 64-bit word and comprised of two cell arrays, each of which containing 64 rows and 512 columns. All local and global interconnects, including bit and bit-bar lines, word line, and decoder wires have been modeled as distributed RC circuits. In this SRAM, the read delay difference between the slowest cell and the fastest one is about 9%.

Although the simulation results we present in this section are specific to the aforesaid technology and design parameters, the general methodology is applicable to any SRAM block designed in any technology. In Section IV we show how the results change with the change of the values of high-T_{ox} and high-V_t, and also as a function of the SRAM cell array size.

B. Library Generation

It is known that each additional threshold voltage or oxide thickness requires one additional mask layer in the fabrication process, which increases the manufacturing cost and reduces the yield [9, 29]. As a result, in many cases, only two threshold voltages and/or two oxide thicknesses are utilized in circuits. That is also why we shall concentrate on the problem of low-leakage SRAM design in a dual-V_t and dual-T_{ox} technology in this paper. Clearly, it is possible to extend the results to handle more than two threshold voltages and two oxide thicknesses. In the next section it is shown how the results are changed if only the option of dual-V_t is available in the technology. We show that in this case, although the efficacy of our technique is reduced, the leakage reduction still remains significant.

The maximum reduction in the subthreshold leakage currents in a SRAM cell is achieved by increasing the threshold voltage of all transistors in the cell. Unfortunately, this scenario also results in the largest read delay penalty for the cell. Therefore, we also consider other configurations which result in lower subthreshold leakage reductions, but also smaller delay penalties. On the other hand, as mentioned in Section II.B.3), to reduce the tunneling gate leakage of an SRAM cell, only the oxide thickness of the pull-down NMOS transistors and the pass-transistors must be increased. Although this is seemingly desirable from a low power point of view, it is not applicable for all cells in the cell array, i.e., thinner oxide thicknesses needs to be used in the cells that are far from the address decoder and the sense amplifiers. It is worth mentioning that due to roll-off effect, increasing the oxide thickness also raises the threshold voltage, resulting in a decrease in the subthreshold leakage. In the following, high-V_t transistors refer to the devices whose threshold voltages have been modified by increasing the channel doping only. Furthermore, our simulations show that when the gate oxide thickness of the PMOS transistors is increased, the reduction in subthreshold leakage due to roll-off effect is very small. That is, the overall leakage reduction achieved by using a thicker gate oxide for the PMOS transistor is negligible.

To make the memory cells more manufacturable, unlike [12], we use a symmetric cell configuration, which means that symmetrically located transistors within an SRAM cell will have the same threshold voltages and oxide thicknesses. Thus, there are 32 different possibilities for assigning high and low threshold voltages and oxide thicknesses to the transistors within a cell. Since increasing the oxide thickness increases the threshold voltage of a transistor as well, we do not increase both the oxide thickness and threshold voltage for a transistor because the delay penalty will be too high. Therefore, the number of different configurations is reduced to eighteen (there are two choices for the pair of PMOS

transistors, three choices for the pull-down NMOS pair, and three choices for the pass-transistor pair). Each configuration is shown by a triplet (*x*,*y*,*z*) where the first entry *x* in the triplet corresponds to the pair of pull-down transistors M1 and M2, the second entry *y* corresponds to the pair of pull-up transistors M3 and M4, and the third entry *z* corresponds to the pass-transistors M5 and M6 as shown in Fig. 1. Each entry is zero, one, or two, if the corresponding transistors are respectively normal, high-V_t, or high-T_{ox}. For example, (0,0,0) corresponds to the original configuration where all transistors in the cell assume default (low) V_t and (low) T_{ox} values whereas (0,1,2) corresponds to a configuration with nominal pull-down transistors, high-V_t pull-up transistors, and high-T_{ox} pass-transistors.

It should be emphasized that our technique does not require all configurations to be used in the optimization process. If a configuration cannot be manufactured due to process restriction or if it has a high manufacturing cost, it can be excluded from the library. Since using eighteen configurations in the optimization process is too expensive, we next show how to eliminate some 'inferior' configurations.

Each configuration has a specific delay and leakage characteristics. We denote the leakage power of the configuration *C* with $P(C)$ and its read and write delays with $D_R(C)$ and $D_W(C)$, respectively. More specifically, $D_R(C)$ is the difference between the time the address bit's voltage reaches $1/2V_{dd}$ and the time the output of the read buffer reaches 90% of its final value. On the other hand, $D_W(C)$ is the write delay, defined as the difference between the time the address bit's voltage reaches $1/2V_{dd}$ and the voltage of bitbar inside the cell reaches 90% of their final values.

Due to the delay of sense amplifiers and output buffers in a read path, the read delay of a cell is higher than its write delay. Therefore, the read delay specifies the performance of an SRAM. Considering the fact that the PMOS transistors in a 6T SRAM cell have a marginal impact on the read delay, it can be seen that increasing the threshold voltage of these transistors increases the write delay without having much effect on its read delay; so one may reduce the leakage power by increasing the threshold voltage of the PMOS transistors as long as the write time is below a target value.

Definition 1: Assume when only the original configuration (0,0,0) is used, the read-delay of the closest and furthest cells to the address decoder and the column multiplexer are $T_{\min}$ and $T_{\max}$, respectively (cf. Fig. 2). Configuration *C* is called *feasible*, if its read and write delays are less than T_{max}. The set of all feasible configurations is called the *Feasible Configuration Set* (*FCS*).

Definition 2: Configuration $C_1 \in$FCS is *inferior* if there exists a configuration $C_2 \in$FCS, whose leakage power and read-delay are no larger than those of C_1, i.e., $P(C_2) \le P(C_1)$ and $D_R(C_2) \le D_R(C_1)$.

It should be noted that the inferiority of a cell depends on different parameters, including the size of the transistors in the cell, the size of the array, and the technology library being used. Changing any of these parameters may change the dominancy relation between two cells.

Definition 3: The maximum subset of *FCS* which does not contain any inferior configuration is called the *Non-Inferior Configuration Set* (*NICS*).

NICS may be obtained by simulating all configurations and removing the inferior ones. When designing a heterogeneous cell SRAM, instead of using the complete set of configurations, *NICS* can be used without degrading the results. Table I shows the set of *NICS* along with their leakage power reduction and read delay increase for the technology described in Section III.A. From this table one can see the delay penalty of some configurations is very small while their leakage saving is significant, e.g., (1,0,0). These configurations are ideal candidates for HCS.

TABLE I
NON-INFERIOR FEASIBLE CONFIGURATION SET (NICS)

Cell	% Leakage Reduction over (0,0,0) Cell	% Read Delay Increase over (0,0,0) Cell
(0,0,0)	-	-
(1,0,0)	43.39	3.02
(0,1,0)	7.60	0.00
(1,1,0)	50.96	2.98
(2,0,0)	16.35	0.43
(2,1,0)	23.93	0.41
(1,1,2)	55.57	6.63

C. Stability

The Static Noise Margin (SNM) of a CMOS SRAM cell is defined as the minimum DC noise voltage necessary to flip the state of a cell [30]. SRAM cells are especially sensitive to noise during a read operation because the "0" storage node rises to a voltage higher than ground due to a resistive voltage divider comprised of the pull-down NMOS transistor and the pass transistor. If this voltage is high enough, it can change the cell's value.

To design an HCS as robust as the conventional SRAM, only configurations that do not degrade the SNM should be used during design.

Definition 4: Configuration *C* is *robust,* if its static noise margin is not any smaller than that of the original cell (0,0,0).

Definition 5: The maximum subset of *FCS* which contains only robust configurations is called *Robust Configuration Set* (*RCS*). The maximum subset of *RCS* which does not contain any inferior configuration is called *Non-Inferior RCS* (*NIRCS*).

To obtain the robust configurations, we consider three separate criteria for SNM: SNM under nominal conditions, worst-case corner-based SNM, and statistical SNM.

1) Stability under Nominal Conditions

Table II lists the set of *NIRCS* when the criterion for robustness is the SNM under nominal condition ($NIRCS_{NC}$). Also shown are the nominal SNM of each configuration in

this set along with the percentage of its improvement over the original cell.

2) Worst Case Stability

Since small transistors are typically used in SRAM cells to achieve a compact design, the most significant source of random intra-die variations in SRAM cells is the threshold voltage variation due to the Random Dopant Fluctuation (RDF) and the line width variation [31]. On the other hand, it is known that gate oxides are very well controlled compared to other dimensions such as the effective channel length [6]. Hence, in this section, we only consider threshold voltage variation for transistors in the 6T SRAM cell.

TABLE II
NOMINAL SNM OF CONFIGURATIONS IN $NIRCS_{NC}$

Cell	Nominal SNM (mV)	% Increase over (0,0,0) Cell
(0,0,0)	185	-
(1,0,0)	208	12.43
(1,1,0)	201	8.65
(1,1,2)	208	12.43

TABLE III
SET OF $NIRCS_{WC}$

Cell	Worst-Case SNM (mV)	% Increase over (0,0,0) cell
(0,0,0)	25	-
(1,0,0)	44	76.00
(1,1,0)	40	60.00
(1,1,2)	47	88.00

In the presence of RDF, the threshold voltage of the SRAM cell transistors can be considered as independent Gaussian random variables [31] where the standard deviation of each transistor depends on its length and width parameter values, i.e.,

$$\sigma = \sigma_{\min}\sqrt{\frac{W_{\min}L_{\min}}{WL}} \qquad 4$$

where σ is the standard deviation of the threshold voltage of a transistor with the channel length and width of L and W, and σ_{min} is the standard deviation of the threshold voltage for the minimum sized transistor in a given manufacturing process [32].

To measure the worst-case SNM, each configuration is tested under all corners of V_t variation. To limit the yield loss, we consider a large range of parametric variation, i.e., 5σ, for the transistors in each configuration; so, each configuration is tested in all corners of $\{-5\sigma, 0, +5\sigma\}$. The number of these corners for each configuration is 3^6=729. In these simulations, the standard deviation of each transistor is obtained from (4) by assuming σ_{min}=30mV which is a typical standard deviation of the threshold voltage in the 65nm technology node [7]. By simulating all configurations, $NIRCS_{WC}$, which denotes *NIRCS* with the worst-case SNM robustness condition, is obtained (see Table III.)

3) Statistical Stability

To measure the statistical stability of each configuration, we used a Monte Carlo simulation of 500 samples to obtain the statistical mean and variance of the SNM for each configuration.

The threshold voltage of each transistor has been modeled as an independent Gaussian random variable whose standard deviation is obtained from (4) by assuming σ_{min}=30mV [7]. By simulating all configurations, $NIRCS_{MC}$, which denotes *NIRCS* with the statistical SNM robustness condition, is obtained (see Table IV.) Here the measure of robustness has been assumed to be μ-5σ. Interestingly, from Table II-Table IV, one can see that for the technology we are using, the three different criteria for robustness result in the same set of configurations. This result may not hold for other technologies or technology nodes.

TABLE IV
SET OF $NIRCS_{MC}$

Cell	μ_{SNM} (mV)	σ_{SNM} (mV)	μ_{SNM}-5σ_{SNM} (mV)	% (μ_{SNM}-5σ_{SNM}) Increase over (0,0,0) Cell
(0,0,0)	186	24	66	-
(1,0,0)	210	26	80	21.21
(1,1,0)	202	25	77	16.67
(1,1,2)	209	25	84	27.27

TABLE V
READ STABILITY FOR NICS CELLS

Cell	I_{trip}/I_{read}	% Decrease over (0,0,0) cell
(0,0,0)	1.69	-
(1,0,0)	1.63	3.5
(0,1,0)	1.64	3.0
(1,1,0)	1.60	5.3
(2,0,0)	1.62	4.1
(2,1,0)	1.57	7.1
(1,1,2)	1.68	0.6

D. Read Stability

The read stability is a transient stability metric which specifies the likelihood of inverting an SRAM cell's stored value during a read operation [12]. It is typically computed as the ratio of I_{trip}/I_{read}, where I_{trip} is the current through the pull-down NMOS transistor on the "0" side of the cell when the state of the cell is inverted by an external current I_{test} injected at the node storing the "0" value. Notice that I_{read} is the maximum current through the pass-transistor during the read operation [20]. The larger the I_{trip}/I_{read} ratio, the higher the read stability of a cell is.

The read stability simulation results on *NICS* configurations are reported in Table V. From this table, it is seen that for different configurations in *NICS*, the maximum reduction in I_{trip}/I_{read} is 7.1%.

E. Writability

The write-trip voltage is a measure of the writability of an SRAM cell [33]. The write-trip voltage is the highest voltage on the bit-line, which can still flip the SRAM cell content. The write-trip voltage is mainly determined by the pull-ups' ratio of the cell [34]. A higher value for the write-trip voltage

represents ease of writability, but the write-trip voltage should be sufficiently lower than the supply voltage so noise cannot cause a write failure or a write during a read operation [33].

Table VI shows the write-trip voltage of different configurations in *NICS*. From this table, one can see that the configurations in *NICS* become slightly easier to write, but at the same time write-trip voltage is far enough from the supply voltage to guarantee safe read/write operations.

TABLE VI
WRITE-TRIP VOLTAGE FOR NICS CELLS

Cell	Write Trip Voltage(mV)	% Increase over (0,0,0) cell
(0,0,0)	424	-
(1,0,0)	438	3.3
(0,1,0)	452	6.6
(1,1,0)	466	9.9
(2,0,0)	428	0.9
(2,1,0)	458	8.0
(1,1,2)	443	4.5

F. Soft Error

Commensurate with down-scaling of the minimum feature size and the critical dimension in the bulk CMOS process technology, soft errors in SRAM memories have become a critical issue [35-37]. In this section we evaluate the effect of our technique on the soft error rate (SER) of the SRAM cells.

A high-energy alpha particle or an atmospheric Neutron striking a capacitive node of a circuit deposits charge which leads to a time-varying voltage pulse at the node. In the case of atmospheric Neutrons, the current flow created by the charge deposited into the node is modeled as (similar models exist for alpha-particle related soft errors):

$$I(Q,t) = \frac{2Q}{\sqrt{\pi}T_S}\sqrt{\frac{t}{T_S}}\exp\left(\frac{-t}{T_S}\right) \quad 5$$

where Q is the collected charge and T_S is the technology-dependent collection waveform time constant [37]. If the collected charge Q exceeds the critical charge Q_{CRIT} in an SRAM cell, it will upset the bit value and cause a soft error. In [37] a methodology for estimating the Neutron-induced soft error rate (SER) in SRAM has been proposed, according to which the dependence of SER on circuit and environmental parameters is expressed as:

$$SER \propto N_{flux} A_S \exp\left(-\frac{Q_{CRIT}}{Q_S}\right) \quad 6$$

where N_{flux} is the intensity of the Neutron flux and A_S is the area of the cross section of the node (i.e., the area of the drain or source region). Moreover, Q_S is the collection slope, which depends strongly on the doping concentration of the drain and source and also the supply voltage level.

In this section we concentrate on Q_{CRIT} when investigating the effect of increasing the threshold voltage and gate-oxide thickness on SER, since the other parameters in (6) are not affected by our proposed technique.

We have used SPICE simulation to measure Q_{CRIT} of each SRAM cell configuration. In these simulations, equation (5) is used to model the collection waveform, and T_S is assumed to be 20ps [37].

Table VII reports Q_{CRIT} for configurations of *NICS*. From this table one can see that Q_{CRIT} of an SRAM cell is only marginally affected by increasing the threshold voltage or oxide thickness.

TABLE VII
QCRIT FOR NICS CELLS

Cell	Q_{CRIT} (fC)	% Decrease over (0,0,0) cell
(0,0,0)	7.87	-
(1,0,0)	7.40	5.9
(0,1,0)	7.83	0.6
(1,1,0)	7.44	5.4
(2,0,0)	7.56	3.9
(2,1,0)	7.56	3.9
(1,1,2)	7.44	5.4

G. Cell Type Assignment

To design a HCS, we need to find out the slowest read and write delay starting with all low-V_t SRAM cells (configuration C_0=(0,0,0)). Next, all remaining configurations are sorted in decreasing order of their leakage reduction. Starting from the configuration that results in the highest leakage reduction among all configurations, say (x,y,z), we replace as many (0,0,0) cells as possible with cell (x,y,z) subject to the condition that the access delay of the replaced cells does not exceed the slowest access delay of the SRAM array. Next we try to replace the remaining (0,0,0) cells with the remaining configurations according to the aforesaid order. As long as design rules are met modifying V_t and T_{ox} (i.e., assigning a cell type) does not change the footprint of a cell. Therefore, the cell type assignment does not change the layout of the SRAM cell array.

Fig. 4 shows the pseudo-code of the heterogeneous cell assignment (HCA). In this figure, *ROW* and *COL* denote the number of rows and columns of the cell array, respectively. If *robustness=1*, only robust configurations are used in the optimization process of HCA. The fastest cell is denoted by index [0,0], while the slowest one is denoted by index [*COL*–1, *ROW*–1]. Subroutines *ReadDelay*(*col,row,C*) and *WriteDelay*(*col,row,C*) return the read and write delays of cell with index of [*col,row*] when configuration *C* is used. If configuration *C* fails for cell [*col,row*], then it will fail for all cells [*i*, *j*], where $i \geq col$ and $j \geq row$. Therefore, a large number of cells can be pruned as soon as a configuration fails for a given cell. In the pseudo-code, *flag*[*col*,*row,C*] is a flag that specifies if *cell*[*col*,*row*] can work with configuration *C*. Initially all flags are set to 1.

To speed up the process, instead of checking for possible replacements of each SRAM cell, one can select $2^n \times 2^m$ cell blocks and do the checking for the slowest cell in the block. If the slowest cell passes the delay test, the whole block will be uniformly optimized based on the current configuration; otherwise, the next configuration for the block is examined (in the case that the block fails the delay test for all

configurations, it will remain unchanged and the next block will be taken up). Evidently, choosing a larger value for n or m decreases the design time, but may degrade the quality of the final result.

```
HCA (ROW, COL, robustness)
Begin
1. T_max=ReadDelay (COL–1, ROW–1, C_0)
2. If (robustness == 1) ConfigSet=NIRCS
3. Else ConfigSet=NICS
4. Sort ConfigSet in decreasing order of leakage
5. For each C in ConfigSet do
6.     For (0≤col<COL, 0≤row<ROW) do
7.         flag [col,row,C] =1;
8. For col=0 to COL–1 do
9.     For row=0 to ROW–1 do
10.        For each C in ConfigSet do
11.            If (flag[col,row,C] == 1)
12.                If  (ReadDelay(col,row,C)<T_max
                     && WriteDelay(col,row,C)<T_max)
13.                    Replace cell[col][row] with C;
14.                    Break;
15.                Else
16.                    For (i≥col, j≥row)
17.                        flag[i,j,C] =0;
End
```

Fig. 4. Pseudo-code for the heterogeneous cell assignment.

It is noteworthy that using the configurations where the pass transistors have thick gate oxides decreases the word-line capacitance, and thereby, reduces the delay of the word-line. To avoid short-circuit power consumption in the SRAM cell-array (which could occur due to simultaneous activation of the pre-charge and WL drivers), one may have to redesign the timing of these two signals for the cell array. The required modification will, however, be minor.

IV. Simulation Results

To study the efficiency of the proposed technique, we performed extensive simulations. To reduce the simulation time, all simulations were done on a simplified version of the memory circuit comprising only of elements in the read/write path of a cell; this included the critical path of the decoder, all cells in corresponding row and column of the SRAM array, the corresponding pre-charge devices, column multiplexers, sense amplifiers, write drivers, and the output buffer.

In the first set of experiments, we applied the proposed technique on the SRAM block described in Section III.A. Table VIII shows the leakage power reduction achieved and the percentage utilization of each configuration by the HCA algorithm for two cases *NICS* and *NIRCS* (i.e., non- robust and robust cases) denoted by HCS and RHCS, respectively. As mentioned in Section III.C, for the technology parameters described earlier, the three different criteria that we defined for the robustness resulted in the same set of configurations for RHCS, as shown in Table II-Table IV. From Table VIII it is seen that the power reduction in HCS and RHCS are 32.6% and 21.2%, respectively.

TABLE VIII
THE LEAKAGE REDUCTION AND THE UTILIZATION OF EACH CONFIGURATION IN THE HETEROGENEOUS CELL SRAM

	% Leakage Reduction	% Utilization of Each Configuration				
		(0,0,0)	(0,1,0)	(1,1,0)	(2,1,0)	(1,1,2)
HCS	32.6	5.0	10.9	21.5	44.3	18.3
RHCS	21.2	60.2	-	21.5	-	18.3

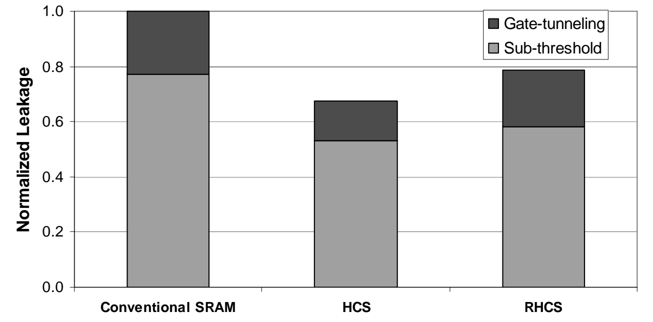

Fig. 5: Subthreshold and tunneling gate leakage in the conventional and heterogeneous cell SRAMs.

Fig. 5 shows the share of subthreshold and tunneling gate currents in the total leakage power dissipation of the conventional SRAM, HCS and RHCS.

A. Effect of high-Vt and high-Tox Selection

To study the effect of specific values of high-V_t and high-T_{ox} on the efficacy of heterogeneous cell SRAM technique, we invoked the HCA algorithm with different values of high-V_t and high-T_{ox}. In these experiments, whose results are reported in Table IX, we considered three values for high-V_t (i.e., 0.23V, 0.28V, and 0.33V) and three values for high-T_{ox} (i.e., 13A^o, 14A^o, and 15A^o) parameters. For each pair of high-V_t and high-T_{ox}, we ran the HCA algorithm with and without the robustness option. From this table one can see that up to 33% leakage power reduction is achieved by using the HCA algorithm. Furthermore, the power reduction is a weak function of the value of high-T_{ox}. On the other hand, for very high values of high-V_t, power reduction drops. The reason is that in this case the delay overhead of high-V_t configurations becomes too high and these configurations are used less frequently in the SRAM block, which in turn results in less power reduction.

To further study the effect of the specific values of high-V_t and high-T_{ox}, we repeated the simulations for the case that only the dual threshold option is available in the technology. Table X shows the power reduction achieved by using the HCA algorithm for three different values of high-V_t. From this

table it is seen that the power reduction in this case is still significant and is as high as 24%.

TABLE IX
THE LEAKAGE REDUCTION IN HETEROGENEOUS CELL SRAM FOR DIFFERENT VALUES OF HIGH-VT AND HIGH-TOX

(high-V_t, high-T_{ox})	% Leakage Reduction	
	HCS	RHCS
(0.23V, 13A^o)	30.3	26.1
(0.23V, 14A^o)	31.3	26.1
(0.23V, 15A^o)	30.8	25.7
(0.28V, 13A^o)	30.8	23.7
(0.28V, 14A^o)	32.1	21.2
(0.28V, 15A^o)	33.4	20.7
(0.33V, 13A^o)	19.1	13.1
(0.33V, 14A^o)	19.1	13.1
(0.33V, 15A^o)	19.1	13.1

TABLE X
THE LEAKAGE REDUCTION IN HETEROGENEOUS CELL SRAM FOR A DUAL-VT TECHNOLOGY

High-V_t	% Leakage Reduction	
	HCS	RHCS
0.23V	22.7	21.0
0.28V	24.5	20.3
0.33V	19.1	13.1

B. *Effect of the Number of Configurations*

Table XI reports the power reduction of the SRAM block for different values of the high-V_t and high-T_{ox} when the number of configurations allowed to be used in the optimized SRAM, including the original configuration, is limited to two or three. As one can see the power reduction is substantial even when only a small number of configurations are used. More precisely, when only two configurations are allowed in the design, 20% power reduction can be achieved; if three configurations can be used in the optimization process, the quality of the results is comparable with the case that all configurations are used in the cell assignment.

C. *Effect of the Array Size*

To further study the efficacy of the HCA algorithm, we conducted another set of experiments for different sizes of the SRAM cell array whose results are reported in Table XII. As discussed in Section II, as technology scales, cell arrays are moving from tall to wide structures; so, here we have considered cell array sizes of 32×256, 32×512, 64×256, and 64×512. In all these simulations the values of high threshold voltage and thick oxide are set to 0.28V and 14°A, respectively.

TABLE XI
THE LEAKAGE REDUCTION IN HETEROGENEOUS CELL SRAM FOR DIFFERENT VALUES OF HIGH-VT AND HIGH-TOX

(high-V_t, high-T_{ox})	% Leakage Reduction			
	HCS		RHCS	
	Two Configs	Three Configs	Two Configs	Three Configs
(0.23V, 13A^o)	23.9	27.9	23.9	24.8
(0.23V, 14A^o)	23.9	28.3	23.9	24.8
(0.23V, 15A^o)	23.9	24.5	23.9	24.5
(0.28V, 13A^o)	17.8	22.8	15.0	19.0
(0.28V, 14A^o)	20.3	26.0	20.3	21.2
(0.28V, 15A^o)	22.0	29.3	20.3	20.8
(0.33V, 13A^o)	13.1	19.1	13.1	13.1
(0.33V, 14A^o)	13.1	19.1	13.1	13.1
(0.33V, 15A^o)	13.1	19.1	13.1	13.1

TABLE XII
SUMMARY RESULTS FOR LEAKAGE REDUCTION AND PERCENTAGE OF REPLACED CELLS IN HCS FOR DIFFERENT ARRAY SIZES

Cell Array Size	% Leakage Reduction	% Replaced Cells
64×256	20.6	90.3
64×512	32.6	95.1
32×256	25.8	94.3
32×512	40.7	96.4

From Table XII one can see that based on the size of cell array, the leakage power reduction resulted from HCA algorithm ranges from 20% to 40%. Moreover, it is seen that the leakage power reduction for a 64×256 cell array is less than that for the 32×256 array. This counter-intuitive result may be explained by noting that when 32 cells are connected to the bit-line, the bit-line becomes less capacitive compared to a 64-cell bit-line. As a result, in a 32-cell bit-line, the delay overheads of some configurations will be less than the delay overheads of them in the 64-cell bit-line (if we use a simple RC model for the delay, changing the threshold voltage of transistors of a cell, changes the R. Now for a 64-cell bit-line the value of C is higher, therefore, the change in the delay is larger. On the other hand, increasing the length of the bit line due to doubling the number of cells connected to it, has a small effect on the delay difference between the fastest cell and the slowest one. This is because of the fact that SRAM arrays are wide structures and the length of the word line has a higher impact on the delay difference) and hence these configurations will be used more frequently, which in turn results in more power reduction.

V. CONCLUSION

In this paper we have presented a novel technique for low-leakage SRAM design. Our technique is based on the fact that due to the non-zero delay of interconnects of the address decoder, word-line, bit-line and the column multiplexers, cells of an SRAM have different access delays. Thus, the threshold voltage or gate oxide thickness of some transistors of cells can

be increased without degrading the performance. We showed by using this technique significant power saving can be achieved without scarifying performance or area. We have showed that this leakage saving is a function of the value of high threshold voltage and oxide thickness, as well as the number of rows and columns in the cell array. By applying the proposed technique to a 64Kb SRAM in 65nm technology node, the total leakage power dissipation of the SRAM has been reduced by up to 40%.

REFERENCES

[1] Y. Taur, "CMOS scaling and issues in sub-0.25 µm systems," in *Design of High-Performance Microprocessor Circuits*, Chandrakasan, W. J. Bowhill, and F. Fox, Eds. Piscataway, NJ: IEEE, 2001, pp. 27–45.

[2] C. Molina, C. Aliagas, M. Garcia*, et al.*, "Non inferior data cache," in *Proc. of International Symposium on Low Power Electronics and Design*, 2003, pp. 274-277.

[3] S. P. Mohanty, R. Velagapudi, and E. Kougianos, "Dual-k versus dual-T technique for gate leakage reduction: a comparative perspective," in *Proc. of International Symposium on Quality Electronic Design*, 2006, pp. 564-569.

[4] V. Mukherjee, S. P. Mohanty, and E. Kougianos, "A dual dielectric approach for performance aware gate tunneling reduction in combinational circuits," in *Proc. of International Conference on Computer Design*, 2005, pp. 431-437.

[5] R. Chau, S. Datta, M. Doczy*, et al.*, "Gate dielectric scaling for high-performance CMOS: From SiO to high-k," in *Proc. of International Workshop on Gate Insulator*, 2003, pp. 124–126.

[6] D. Lee, D. Blaauw, and D. Sylvester, "Gate oxide leakage current analysis and reduction for VLSI circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 12, no. 2, Feb. 2004, pp. 155-166.

[7] Semiconductor Industry Association, International Technology Roadmap for Semiconductors, 2003 edition, http://public.itrs.net/.

[8] N. Sirisantana, L. Wei, and K. Roy, "High performance low power CMOS circuits using multiple channel length and multiple oxide thickness," in *Proc. of International Conference on Computer Design*, 2000, pp. 227-232.

[9] M. Togo, K. Noda, and T. Tanigawa, "Multiple-thickness gate oxide and dual-gate technologies for high-performance logic embedded DRAMs," in *Proc. of IEDM Technical Digest*, 1998, pp. 347–350.

[10] K. Zhang, U. Bhattacharya, Z. Chen*, et al.*, "SRAM design on 65-nm CMOS technology with dynamic sleep transistor for leakage reduction," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 4, Apr. 2005, pp. 895-901.

[11] C. Kim and K. Roy, "Dynamic Vt SRAM: a leakage tolerant cache memory for low voltage microprocessor," in *Proc. of International Symposium on Low Power Electronics and Design*, 2002, pp. 251-254.

[12] N. Azizi, F. Najm, and A. Moshovos, "Low-leakage asymmetric-cell SRAM," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 11, no. 4, Aug. 2003, pp. 701-715.

[13] M. D. Powell, S. Yang., B. Falsafi*, et al.*, "Gated-Vdd: a circuit technique to reduce leakage in cache memories," in *Proc. of International Symposium on Low Power Electronics Design*, 2000, pp. 90–95.

[14] S.Heo, K. Barr, M. Hampton*, et al.*, "Dynamic fine-grain leakage reduction using leakage-biased bitlines," in *Proc. of International Symposium on Computer Architecture*, 2002, pp. 137-147.

[15] A. Agarwal, H. Li, and K. Roy, "DRG-cache: A data retention gated-ground cache for low power," in *Proc. of Design Automation Conference*, 2002, pp. 473- 478.

[16] K. Flautner, N. Kim, S. Martin*, et al.*, "Drowsy caches: simple techniques for reducing leakage power," in *Proc. of International Symposium on Computer Architecture*, 2002, pp. 148-157.

[17] Y. Tsai, D. Duarte, N. Vijaykrishnan*, et al.*, "Characterization and modeling of run-time techniques for leakage power reduction," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 12, no. 11, Nov. 2004, pp. 1221-1233.

[18] B. Amelifard, F. Fallah, and M. Pedram, "Reducing the sub-threshold and tunneling gate leakage of SRAM cells using Dual-Vt and Dual-Tox assignment," in *Proc. of Design, Automation and Test in Europe*, 2006, pp. 995 -1000.

[19] B. Amelifard, F. Fallah, and M. Pedram, "Low-leakage SRAM design with dual Vt transistors", in *Proc. of International Symposium on Quality of Electronic Design*, 2006, pp. 729 -734.

[20] F. Hamzaoglu, Y. Te, A. Keshavarzi*, et al.*, "Dual Vt-SRAM cells with full-swing single-ended bit line sensing for high-performance on-chip cache in 0.13µm technology generation," in *Proc. of International Symposium on Low Power Electronics and Design*, 2000, pp. 15–19.

[21] K. Z. Zhang, U. Bhattacharya, Z. Chen*, et al.*, "A 3-GHz 70Mb SRAM in 65nm CMOS technology with integrated column-based dynamic power supply," in *Proc. of International Solid-State Circuits Conference*, 2005, pp. 474–475.

[22] D. Weiss, J. Wuu, and V. Chin, "The on-chip 3MB subarray based 3rd level cache on an Itanium microprocessor," in *Proc. of International Solid-State Circuits Conference*, 2002, pp. 112–113.

[23] R. Preston, "Register files and caches," in *Design of High-Performance Microprocessor Circuits*, A. Chandrakasan, W. J. Bowhill, and F. Fox, Eds. Piscataway, NJ: IEEE, 2001, pp. 285-308.

[24] V. De, A. Keshavarzi, S. Narendra*, et al.*, "Techniques for leakage power reduction," in *Design of High-Performance Microprocessor Circuits*, A. Chandrakasan, W. J. Bowhill, and F. Fox, Eds. Piscataway, NJ: IEEE, 2001.

[25] F. Hamzaoglu and M. Stan, "Circuit-level techniques to control gate leakage for sub-100nm CMOS," in *Proc. of International Symposium on Low Power Electronics and Design*, 2002, pp. 60-63.

[26] K. Bowman, L. Wang, X. Tang*, et al.*, "A circuit-level perspective of the optimum gate oxide thickness," *IEEE Transactions on Electron Devices*, vol. 48, no. 8, Aug. 2001, pp. 1800-1810.

[27] http://www.synopsys.com/products/mixedsignal/hspice/hspice.html

[28] http://www.eas.asu.edu/~ptm/

[29] A. Sirvastava, "Simultaneous Vt selection and assignment for leakage optimization," in *Proc. of International Symposium on Low Power Electronics and Design*, 2003, pp. 146-151.

[30] A. J. Bhavnagarwala, X. Tang, and J. Meindl, "The impact of intrinsic device fluctuations on CMOS SRAM cell stability," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 4, Apr. 2001, pp. 658-665.

[31] S. Mukhopadhyay, H. Mahmoodi-Meimand, and K. Roy, "Modeling of failure probability and statistical design of SRAM array for yield enhancement in nanoscaled CMOS," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 12, Dec. 2005, pp. 1859-1880.

[32] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices*. New York, NY: Cambridge Univ. Press, 1998.

[33] R. Heald and P. Wang, "Variability in sub-100nm SRAM designs," in *Proc. of International Conference on Computer-Aided Design*, 2004, pp. 347-352.

[34] E. Grossar, M. Stucchi, K. Maex*, et al.*, "Read stability and write-ability analysis of SRAM cells for nanometer technologies," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 11, Nov. 2006, pp. 2577-2588.

[35] N. Seifert, D. Moyer, N. Leland*, et al.*, "Historical trend in alpha-particle induced soft error rates of the alpha microprocessor," in *Proc. of International Reliability Physics Symposium*, 2001, pp. 259-265.

[36] T. Kamik, B. Bloechel, K. Soumyanath*, et al.*, "Scaling trends of cosmic rays induced soft errors in static latches beyond 0.18um," in *Proc. of Symposium on VLSI Circuits*, 2001, pp. 61-62.

[37] P. Hazucha and C. Svensson, "Impact of CMOS technology scaling on the atmospheric neutron soft error rate," *IEEE Transactions on Nuclear Science*, vol. 47, no. 6, Dec. 2000, pp. 2586–2594.