

# SRAM Leakage Suppression by Minimizing Standby Supply Voltage

Huifang Qin, Yu Cao, Dejan Markovic, Andrei Vladimirescu, and Jan Rabaey  
Department of EECS, University of California at Berkeley, Berkeley, CA 94720, USA

## Abstract

Suppressing the leakage current in memories is critical in low-power design. By reducing the standby supply voltage ( $V_{DD}$ ) to its limit, which is the Data Retention Voltage (DRV), leakage power can be substantially reduced. This paper explores how low DRV can be in a standard low leakage SRAM module and analyzes how DRV is affected by parameters such as process variations, chip temperature, and transistor sizing. An analytical model for DRV as a function of process and design parameters is presented, and forms the base for further design space explorations. This model is verified using simulations as well as measurements from a 4KB SRAM chip in a 0.13 $\mu\text{m}$  technology. It is demonstrated that an SRAM cell state can be preserved at sub-300mV standby  $V_{DD}$ , with more than 90% leakage power savings.

## 1. Introduction

One of the negative side effects of technology scaling is that leakage power of on-chip memory increases dramatically and forms one of the main challenges in future system-on-a-chip (SoC) design. In battery-supported applications with low duty-cycles, such as the Pico-Radio wireless sensor nodes [1], cellular phones, or PDAs, leakage power can dominate system power consumption and determine battery life. Therefore, an efficient memory leakage suppression scheme is critical for the success of ultra low-power design.

Various techniques have been proposed to reduce the SRAM sub-threshold leakage power. At the circuit level, dynamic control of transistor gate-source and substrate-source back bias were exploited to create low leakage paths during standby periods [2]. Yet these approaches require many modifications on the SRAM cell structure, resulting in a large design and area overhead. Furthermore, the application of reversed body bias exacerbates within-die variations in threshold voltage. At the architectural level, leakage reduction techniques include gating off the supply voltage ( $V_{DD}$ ) of idle memory sections, or putting less frequently used sections into drowsy standby mode. These approaches exploited the quadratic reduction of leakage power with  $V_{DD}$ , and achieved optimal power-performance tradeoffs with assistance of compiler-level cache activity analysis. The cache delay technique applied adaptive timing policies in cache line gating, achieving 70% leakage saving at modest performance penalty [3]. To further exploit leakage control on caches with large utilization ratio, the approach of drowsy

cache allocated inactive cache lines to a low-power mode, where  $V_{DD}$  was lowered but with memory data preserved [4].

While the drowsy caches approach can achieve leakage energy savings of over 70% in a data cache, the question remains on the lower bound of standby  $V_{DD}$  that still preserves data. Knowledge of the minimum low-power mode supply voltage allows a designer to exploit the maximum achievable leakage reduction for a given technology. Understandings of low voltage SRAM data preservation behavior further open the opportunity for aggressive memory supply voltage minimization, which has been the bottleneck in VLSI system voltage scaling. Driven by the requirements of ultra low-power applications, this paper presents the first work on exploring the limit of SRAM low voltage data preservation under realistic conditions.

In SRAM design, the Data Retention Voltage (DRV) defines the minimum  $V_{DD}$  under which the data in a memory unit is still preserved. An analytical model of DRV is developed to investigate the dependence of DRV on process and design parameters (Section 2). To verify the new model and further understand the limitations of DRV under realistic conditions, a 4KB SRAM test chip with dual-rail supply scheme was designed and fabricated in a 0.13 $\mu\text{m}$  technology, as introduced in Section 3. This scheme targets ultra low-power applications and uses a customized on-chip switch capacitor converter to generate standby  $V_{DD}$ . Section 4 presents measurement results of the SRAM data preservation and leakage suppression. Impact of various process and design factors on DRV is analyzed in Section 5. Finally, Section 6 concludes current work and proposes future directions.

## 2. Data retention voltage analysis

The circuit structure of a 6T SRAM cell is shown in Fig. 1. When  $V_{DD}$  is reduced to DRV, all six transistors are in the sub-threshold (sub- $V_{th}$ ) region, thus the capability of SRAM data retention strongly depends on the sub- $V_{th}$  current conduction behavior (i.e., leakage). In order to understand

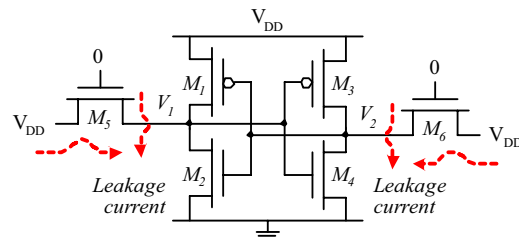


Figure 1. Standard 6T SRAM cell structure.

the low voltage data preservation behavior of SRAM and the potential for leakage saving through minimizing standby  $V_{DD}$ , analytical models of SRAM DRV and cell leakage current are developed in this section.

In a standard SRAM cell (Fig. 1), when  $V_{DD}$  scales down to DRV, the voltage transfer curves (VTC) of the internal inverters degrade to such a level that noise margin of the SRAM cell degrades to zero, as illustrated in Fig. 2. Using the notations of Fig. 1, this condition is given by:

$$\left. \frac{\partial V_1}{\partial V_2} \right|_{\text{Left inverter}} = \left. \frac{\partial V_1}{\partial V_2} \right|_{\text{Right inverter}}, \text{ when } V_{DD} = \text{DRV}. \quad (1)$$

If  $V_{DD}$  is further reduced below DRV, the inverters flip to the biased state determined by the deteriorated VTC and lose the capability to preserve the stored data.

Based on this observation, the DRV of an SRAM cell can be determined by solving the sub- $V_{th}$  VTC equations of the two internal data holding inverters, since all the transistors conduct in weak inversion region when  $V_{DD}$  is around DRV. The derivation is presented below.

When an SRAM cell (Fig. 1) is in standby mode, the currents in each internal inverter are balanced:

$$\text{Node } V_1: I_1 + I_5 = I_2, \quad (2)$$

$$\text{Node } V_2: I_3 + I_6 = I_4. \quad (3)$$

Assuming that during standby

$$V_1 \approx 0 \text{ and } V_2 \approx V_{DD}, \quad (4)$$

and assuming that the bit-lines are set to  $V_{DD}$ ,  $I_6$  is negligible and Eq. (3) can be simplified to:

$$\text{Node } V_2: I_3 = I_4. \quad (5)$$

In Eqs. (2, 3, 5),  $I_i$  is the conduction current of the  $i^{\text{th}}$  transistor (Fig. 1) in the sub- $V_{th}$  region. Considering that  $I_i$  is dominated by the drain-source leakage in current technology (i.e., ignoring gate leakage and other leakage mechanisms which have minor effects compared to the sub- $V_{th}$  current),  $I_i$  is modeled as in [5]:

$$I_i = S_i I_0 \exp\left(\frac{-V_{th}}{n_i kT/q}\right) \cdot \exp\left(\frac{V_{GS}}{n_i kT/q}\right) \cdot (1 - e^{-V_{DS}/(kT/q)}), \quad (6)$$

where  $S_i$  is the transistor (W/L) ratio;  $I_0$  is a process-specific current at  $V_{GS}=V_{th}$  for a transistor with W/L=1;  $T$  is the chip temperature; and  $n_i$  is the sub- $V_{th}$  factor, (sub- $V_{th}$  swing /60mV at room temperature). If we further define:

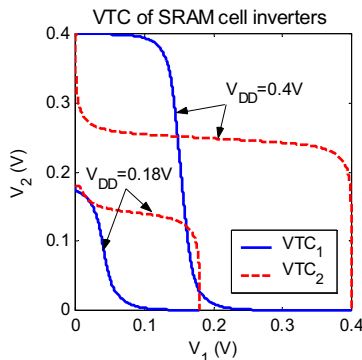


Figure 2. Deterioration of inverter VTC under low- $V_{DD}$ , with zero SRAM cell noise margins at DRV.

$$A_i = S_i I_0 \exp\left(\frac{-V_{th}}{n_i kT/q}\right), \quad (7)$$

$I_i$  can be expressed as:

$$I_i = A_i \cdot \exp\left(\frac{V_{GS}}{n_i kT/q}\right) \cdot (1 - e^{-V_{DS}/(kT/q)}). \quad (8)$$

Substituting these current models, which are functions of  $V_1$ ,  $V_2$ ,  $V_{DD}$ ,  $T$ , and other technology parameters, into Eqs. (2, 5), we obtain the VTCs of the inverters in the cell. Then, together with Eq. (1), the value of the DRV (and the corresponding  $V_1$  and  $V_2$ ) can be derived.

A general solution to these equations requires numerical iterations. To avoid the iterations, we first estimate the initial value of DRV ( $DRV_1$ ), using the approximations in Eq. (4):

$$DRV_1 = \frac{kT/q}{n_2^{-1} + n_3^{-1}} \cdot \log\left[\left(n_3^{-1} + n_4^{-1}\right) \frac{A_4}{A_2 A_3} \left(\frac{A_5}{n_2} + \frac{A_1}{(n_1^{-1} + n_2^{-1})^{-1}}\right)\right], \quad (9)$$

where  $kT/q$  equals 26mV when  $T=27^\circ\text{C}$ . Then, using  $DRV_1$ , the approximations in Eq. (4) are refined as:

$$V_1 = \frac{kT}{q} \cdot \frac{A_1 + A_5}{A_2} \cdot \exp\left(\frac{-DRV_1}{n_2 kT/q}\right), \quad (10)$$

$$V_2 = DRV_1 - \frac{kT}{q} \cdot \frac{A_4}{A_3} \cdot \exp\left(\frac{-DRV_1}{n_3 kT/q}\right). \quad (11)$$

With Eqs. (10-11) available, we can refine the calculation of DRV and a final expression is obtained:

$$DRV = DRV_1 + \left[ \frac{V_1}{2} + \frac{(DRV_1 - V_2) \cdot n_2}{2} \right]. \quad (12)$$

The above DRV formula only relies on the values of  $A_i$  and  $n_i$ , which can be easily extracted from transistor characterizations (either by simulation or measurement). In addition, it captures the dependence of DRV on process variations in transistors, sizing  $S_i$ , and chip temperature  $T$ . Based on Eqs. (7-12), the impact of these process and design factors on DRV can be formulated as:

$$\Delta DRV = DRV_0 + \sum_i a_i \frac{\Delta S_i}{S_i} + \sum_i b_i \Delta V_{thi} + c \Delta T, \quad (13)$$

where  $DRV_0$  is the nominal value at room temperature;  $a_i$ ,  $b_i$ , and  $c$  are fitting coefficients. With the 0.13 $\mu\text{m}$  technology used and considering an industrial SRAM cell sized for optimal performance, the DRV formula predicts  $DRV_0 = 77\text{mV}$  at perfect matching and  $DRV_0 = 169\text{mV}$  with  $3\sigma$  variations in  $V_{th}$  and channel length. These analytical results match well with SPICE simulated values of 78mV and 170mV, respectively. The model coefficients  $a_i$ 's are extracted from simulations:  $a_1 = 10\text{mV}$ ,  $a_3 = -41\text{mV}$ ,  $a_4 = 11\text{mV}$  ( $a_2$  is negligible);

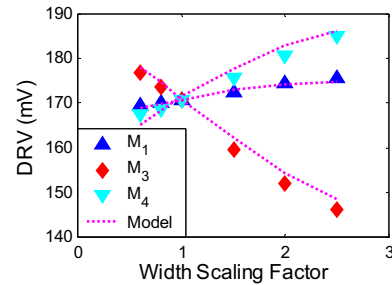


Figure 3. Modeled and simulated DRV as a function of transistor width scaling.

$b_i$  is about 10mV, depending on bias condition. As demonstrated in Fig. 3, Eq. (13) fully captures the impact of transistor sizing on DRV. Due to the competing behavior of PMOS devices  $M_1$  and  $M_3$ , they have different roles in state preservation. Fig. 3 infers that transistor sizing is an effective means in tuning the DRV of an SRAM cell, which will be further explored in Section 5. Temperature coefficient  $c$  is extracted as 0.169mV/°C, which predicts an increase of 12.3mV in DRV when  $T$  rises from 27°C to 100°C.

With DRV obtained, the total leakage of an SRAM cell in the sub- $V_{th}$  region can be calculated as:

$$I_{leak} = I_2 + I_4 \quad (14)$$

$$= S_2 i_2 \exp\left(\frac{V_2}{n_2 kT/q}\right) \cdot \left[1 - \exp\left(-\frac{V_1}{kT/q}\right)\right]$$

$$+ S_4 i_4 \exp\left(\frac{V_1}{n_1 kT/q}\right) \cdot \left[1 - \exp\left(-\frac{DRV - V_2}{kT/q}\right)\right]$$

where  $i = I_0 \cdot \exp[-V_{th}/(n \cdot kT/q)]$ ;  $i_2 = 2.95\mu\text{A}$  and  $i_4 = 6.78\mu\text{A}$ ; DRV,  $V_1$ , and  $V_2$  are calculated from Eqs. (10-13). Thus, the leakage power  $P_{leak}$  under DRV is:

$$P_{leak} = DRV \cdot I_{leak} \quad (15)$$

Leakage current  $I_{leak}$  as provided in Eq. (14) represents the minimum achievable leakage while preserving state.

### 3. Ultra low voltage SRAM standby: Design and implementation

To obtain silicon verification of the presented DRV model and explore the potential of SRAM leakage suppression with ultra low standby  $V_{DD}$ , a 4KB SRAM test chip with dual rail standby control was implemented in a 0.13 $\mu\text{m}$  technology. Designed for ultra low-power applications, this scheme puts the entire SRAM into deep sleep during the system standby period. As shown in Fig. 4, the SRAM supply rail are connected to the standard  $V_{DD}$  and the standby  $V_{DD}$  through two big power switches. A customized on-chip SC converter generates the standby  $V_{DD}$  with high conversion efficiency. Compared to existing SRAM leakage control techniques, the simplicity of this scheme leads to minimized design effort and maximum leakage saving during standby mode, which is desired for battery-supported systems.

#### 3.1. Design considerations

**3.1.1. Standby stability and noise margin analysis.** In an actual implementation, reducing  $V_{DD}$  all the way down to the DRV is not really an option, as other mechanisms may

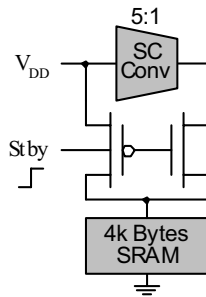


Figure 4. Standby leakage suppression scheme.

disrupt the state of the memory cell. More particularly, these are the noise on the supply rail (mostly caused by the output voltage ripple of the switched-capacitor converter), as well as radiation particles. To offset these effects, an appropriate noise margin has to be provided. Simulation shows that assigning a guard band of 100mV above DRV for standby  $V_{DD}$  gives 60mV in SRAM cell Static Noise Margin (SNM), and that the SNM degrades linearly with the  $V_{DD}$  guard band (becoming zero at the DRV). Here, the SNM is defined as the edge of the maximum square that can fit into the cross section of the VTC diagram of the cross-coupled inverters [6]. A guard band of 100mV over the DRV is sufficient to overcome the 20mV peak-to-peak ripple on the standby  $V_{DD}$  (with worst case SNM of 55mV).

The radiation particle events pose a more serious hazard. With data storage node parasitic capacitance around 1fF in a 0.13 $\mu\text{m}$  technology SRAM cell, the critical charge ( $Q_{critical}$ ) for a 1V  $V_{DD}$  is approximately 3fC. For a reduced  $V_{DD}$  at 100 mV above the DRV,  $Q_{critical}$  is reduced to 0.5fC. To ensure reliable state preservation, future memories may have to add additional storage capacitance [7]. Another option to combat the effect of soft errors is to apply error-correction schemes.

**3.1.2. Dual voltage scheme design concerns.** Main considerations in a dual supply scheme include the operation delay overhead due to the power switch resistance, memory wake up delay and the power penalty during mode transition. Targeted for ultra low-power applications, the system requirements of this design are much more stringent on power than performance [1]. In such a situation the concern of the operation delay overhead is not that crucial. A 200 $\mu\text{m}$  wide PMOS power switch with 30 $\Omega$  conducting resistance is used to connect the memory module to a 1V operation supply voltage. With the same switch the memory wake up time is simulated to be within 10ns, which is typically a small fraction of the target application system cycle time.

The wake up power penalty incurred during switching from the standby mode to the full- $V_{DD}$  mode determines the minimum standby time for the scheme if net power saving over one standby period are to be achieved. This break-even time is an important system-design parameter, as it helps the power control algorithm to decide when a power-down would be beneficial. With the parasitic capacitance information attained from process model, the minimum standby time in this design is estimated to be around tens of microseconds, which is much shorter than the typical system idle time in a battery-supported system.

#### 3.2. Test chip implementation

Layout of the 0.13 $\mu\text{m}$  SRAM test chip is shown in Fig. 5. The two main components are a 4KB SRAM module and a Switch Capacitor (SC) converter. This memory is an IP module with no modifications from its original design. As shown in Fig. 6, a representative five-stage step down dc-dc switch capacitor topology is selected to implement the on-chip standby  $V_{DD}$  generator [8]. Compared to magnetic-based voltage regulators, SC converter provides higher efficiency, smaller output current ripple, and easier on-chip integration for small loads in the microwatt range. The design challenge

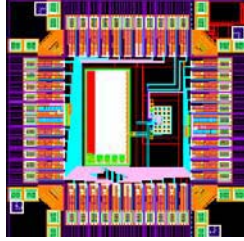


Figure 5. A 0.13 $\mu$ m SRAM leakage-control test chip.

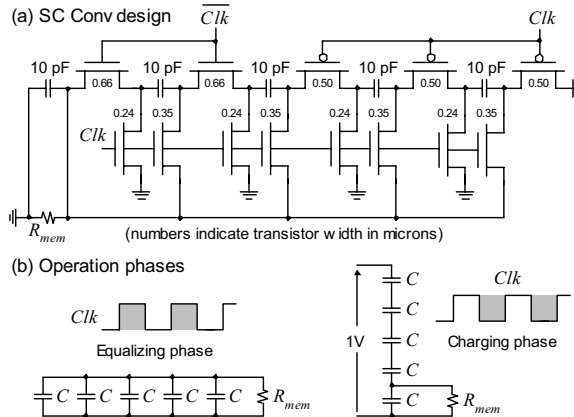


Figure 6. (a) Schematic of switch capacitor converter, (b) Operation phases.

here resides in handling small output load in the range of 10–20 $\mu$ W. With such low power operation, power loss incurred by short-circuit currents during phase switching becomes comparable to output power and forms a significant portion of total power loss. To maximize power efficiency, it is desirable to minimize both the switching voltage drop and short-circuit current, which have opposite dependence on device sizes. Hence the size and type selection of the switch devices (i.e., NMOS vs. PMOS) are carefully tuned to balance these two requirements. Fig. 6 shows the optimized design, in which an 85% conversion efficiency is achieved with a 1V input and an output load of estimated SRAM module leakage at standby mode.

## 4. Measurement results

### 4.1. DRV measurement

The DRV is measured by monitoring the data retention capability of an SRAM cell with different values of standby  $V_{DD}$ , as demonstrated in Fig. 7. With  $V_{DD}$  switching between active and standby modes, a specific state is written into the SRAM cell under test at the end of each active period ( $t_2$ ), and then read out at the beginning of the next active period ( $t_1$ ). Preservation of the assigned logic state is observed when standby  $V_{DD}$  is higher than DRV, while the state is lost when standby  $V_{DD}$  is below DRV.

The DRV was measured for 50 SRAM cells. Figure 8 shows the distribution of the total measurement results. The DRV values range from 80mV to 250mV with the mean value of 170mV at an approximately normal distribution. Such a wide range of DRV uncertainty reflects the existence

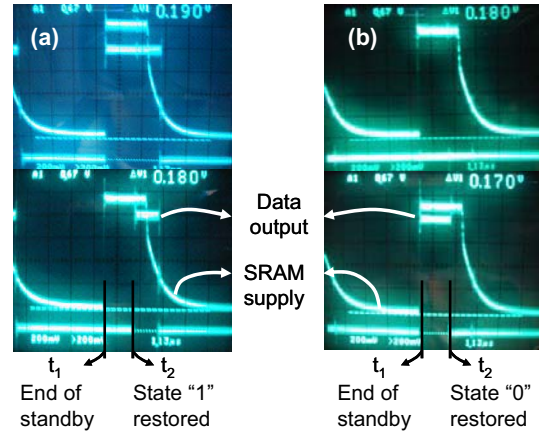


Figure 7. Waveform of DRV measurement. (a) DRV = 190mV in SRAM cell 1 with state “1”, (b) DRV = 180mV in SRAM cell 2 with state “0”.

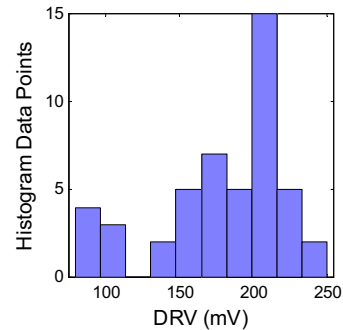


Figure 8. Distribution of measured DRV data.

of considerable process variations during fabrication. The 78 mV ideal DRV, assuming perfect process matching as obtained by simulation and analytical modeling, shows up as the lower bound of the measured DRV distribution. With mismatches included, the measured mean value of 170mV matches the analytical model with  $3\sigma$  variations in  $V_{th}$  and channel length. Moreover, temperature dependency of DRV was investigated with measurement. When the test chip was heated up to 100 $^{\circ}$ C, DRV was obtained as 183mV, which verifies the calculation of 181mV by Eq. (13). As evaluated in Sec. 2, our analytical DRV model not only predicts the ideal DRV values, but also fully captures the impact of process and temperature variations. Thus, it can serve as a convenient base for further design optimizations.

### 4.2. SRAM leakage measurement

Leakage measurement result of the 4KB SRAM is shown in Fig. 9. It increases exponentially when  $V_{DD}$  is high. Similar as observed in the DRV measurement, this phenomenon reflects the impact of process variations on SRAM leakage, more specifically the fluctuations in channel length and  $V_{th}$ . For short channel transistors, drain-induced-barrier-lowering (DIBL) effect further causes severe  $V_{th}$  degradation, resulting in even higher leakage in high- $V_{DD}$  conditions. Furthermore, the shaded area in Fig. 9 indicates the range of measured DRV (80-250mV). While the memory states can be preserved at sub-300mV  $V_{DD}$ , adding an extra guard band of 100mV to the standby  $V_{DD}$  enhances the noise robustness of

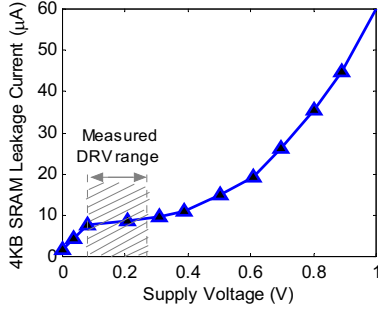


Figure 9. Measured SRAM leakage current.

state preservation as analyzed in section 3.1.1. With the resulting 350mV standby  $V_{DD}$ , SRAM leakage current can still be reduced by over 80%. Subsequently the leakage power, as the product of  $V_{DD}$  and leakage current, is efficiently reduced by more than 90%.

### 4.3. Dual rail standby scheme measurement

The dual rail scheme is shown to be fully functional through the DRV measurements. With 10MHz switch control signal, the SC converter generates the standby  $V_{DD}$  with less than 20mV peak-to-peak ripple. Wake up time of 10ns is observed during mode transition, while the sleep time spans around 10 $\mu$ s. Delay overhead during active operation is measured to be about 2X, which is reasonable for an ultra low-power application where the system clock period is typically 10 times the operation cycle of a low leakage SRAM.

## 5. SRAM DRV minimization: Analysis

While SRAM designs have been well optimized for speed and power metrics, improving DRV for future ultra low-voltage applications poses a new challenge for low-power SRAM designers. Results from previous sections have illustrated the importance of process variations, chip temperature, and transistor sizing on DRV. A thorough understanding of their impacts on DRV and other performance metrics provides insight into SRAM cell design for ultra low-voltage and ultra low-power designs.

### 5.1. Process variations and temperature impacts

Process variation and temperature fluctuation are the main imperfections in a real environment that cause degradations in circuit performance. Figure 10 shows the simulated effect of these two parameters on SRAM DRV. Data shows that process variations play a critical role in determining DRV, which increases by 100mV in the presence of  $3\sigma$  mismatches in  $V_{th}$  and channel length ( $L$ ). On the other hand, temperature affects DRV in a less severe manner. When  $T$  changes from 27°C to 100°C, DRV rises about 13mV, which was verified by test chip measurement. The difference in process variation and  $T$  effects on DRV is because fluctuation in  $T$  affects all transistors in an SRAM cell uniformly, while local process variations may change the drive strength ratio between transistors, resulting in an exacerbated sub- $V_{th}$  VTC mismatch and thus, a substantial DRV increment. Both effects are well captured by the analytical models in Sec. 2.

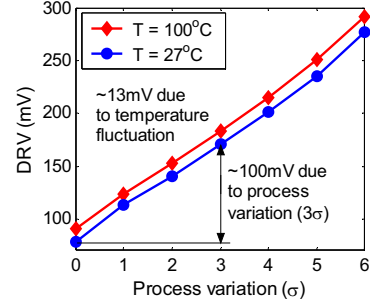


Figure 10. DRV under Process and T variations.

As observed from measurement and simulation results in previous sections, minimizing process variation is the most effective method to reduce DRV. While process variation control becomes more difficult in future technology nodes, it will define an effective lower bound on SRAM  $V_{DD}$ . Another way to improve data preservation under low  $V_{DD}$  is to ensure low temperature operation (e.g., at room temperature), which is easier for ultra low-power applications.

### 5.2. Sizing optimization

While sizing has long been an effective tool in conventional power and speed optimization, taking DRV into account further improves future ultra low-power SRAM designs. This section presents DRV-aware SRAM cell optimization analysis based on transistor sizing. Here, leakage power is considered as the power metric since this analysis targets ultra low-power design.

**5.2.1. Read delay and area cost models.** To facilitate analytical SRAM cell optimization, delay and area models as a function of transistor sizing are developed. In this model, the SRAM access delay is considered as the time required to pull down the bit-line voltage by 15% of  $V_{DD}$ . Starting from a simple hand-calculation model, a curve-fitting approach is used to develop a more accurate and scalable delay model as given by analytical derivation:

$$t_{read} = \frac{K \cdot S_W + (C_{BL0} + K_{BL} \cdot S_W)}{K_{read} \cdot \left( A \cdot S_W + B \cdot S_W^3 / S_N^2 - C \cdot S_W^2 / S_N - \frac{1}{2} \right)}, \quad (16)$$

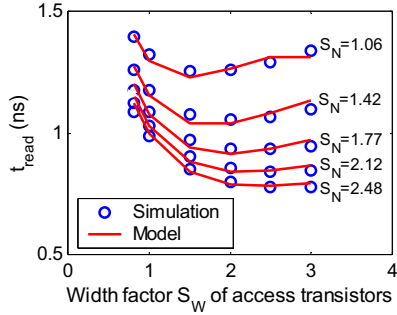
In this model  $S_W$ ,  $S_N$  represent the (W/L) ratio of the write access transistor and the NMOS pull down transistor, respectively, while parameters  $K$ ,  $K_{read}$ ,  $A$ ,  $B$ , and  $C$  are fitting parameters. They are related, but not equal to the following physical parameters:

$$K_{read} = k_n' \cdot V_{DSAT}^2 \cdot \frac{2}{V_{DD}}, \quad A = \frac{V_{DD} - V_{thW}}{V_{DSAT}}, \quad (17.a)$$

$$B = \frac{V_{DSAT}}{2(V_{DD} - V_{thN})}, \quad C = \frac{V_{DD} - V_{thW} - V_{DSAT}}{V_{DD} - V_{DSAT}}. \quad (17.b)$$

Fitting parameters  $C_{BL0}$  and  $K_{BL}$  define a relationship between the bit-line capacitance  $C_{BL}$  and the W/L ratio  $S_W$  of access transistors. The delay model fits simulated results as shown in Fig. 11. Due to a high layout density of an SRAM, the SRAM cell area is simply modeled as a linear function of the total transistor area.

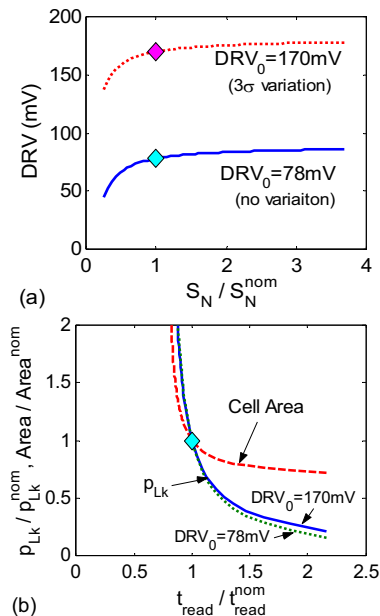
**5.2.2. DRV minimization.** In conventional performance-optimized SRAM cell design, the pull-down NMOS devices



**Figure 11. Read delay model fitting at  $V_{DD} = 1V$ .**

are sized substantially larger than the PMOS devices. This imbalance in the pull-up and pull-down paths leads to exacerbated VTC deterioration at low  $V_{DD}$ , and degrades DRV. It can be observed from Fig. 3 that by either decreasing the NMOS size or increasing the PMOS size DRV can be effectively reduced due to better balance in VTC. Further taking delay and area costs into account, NMOS tuning proves to be more desirable. This is because smaller NMOS improves both DRV and area at certain delay penalty, while larger PMOS only improves DRV and pays large area cost, which is expensive in memory design.

SRAM cell optimization tradeoffs are shown in Fig. 12, where the marked nominal case is a commercial performance-optimized cell design. Fig. 12a plots the lower bound on DRV for both  $3\sigma$  process variation ( $DRV_0 = 170mV$ ) and perfect matching ( $DRV_0 = 78mV$ ) versus NMOS size. In both cases DRV can be reduced by 20~30mV at some delay penalty. Tradeoffs between delay, leakage power and area are illustrated in Fig. 12b. It can be observed that around the nominal design point, area is best traded off for performance. However, leakage power is then best traded for performance at delay of about 20% higher than the nominal point, with



**Figure 12. SRAM cell design tradeoffs:**  
**(a) DRV vs. NMOS transistor size,**  
**(b) Leakage power and area vs. read delay.**

area reduced at the same time. Therefore, reducing size of the cell both lowers the area and the leakage power at the expense of increased delay. At a 30% delay increase, a 50% leakage power reduction can be achieved. Furthermore, since  $I_{leak}$  is relatively insensitive to  $V_{DD}$  in the range around DRV as compared to high- $V_{DD}$  condition (see Fig. 10), there is only slight difference in leakage power for the cases with no process variation and  $3\sigma$  variation.

## 6. Conclusions and future work

This paper explores the limit of SRAM data preservation under ultra-low standby  $V_{DD}$ . An analytical model of the SRAM DRV is developed and verified with measurement results. A commercial SRAM module with high- $V_{th}$  process is shown to be capable of sub-300mV standby data preservation. Under this low standby  $V_{DD}$ , leakage power saving of more than 90% can be achieved with a dual-rail standby scheme designed for ultra low-power applications. The DRV is observed to be a strong function of process variation and SRAM cell sizing, while the design tradeoffs between read delay, area and leakage power can be optimized.

While existing approaches generally involve significant tradeoffs with other performance metrics or technology limits, more opportunities exist on architectural level innovations. As an example, more SRAM leakage savings can be achieved with assistance from error tolerant schemes when the standby supply voltage is scaled down beyond the limit of DRV. Future work will be focused on exploiting design techniques to achieve even lower power and higher reliability in memory design.

## Acknowledgements

The sponsorship of the GSRC MARCO center and fabrication support from ST Microelectronics are greatly acknowledged. The authors would also like to thank to Professor Seth Sanders and Dr. Bhusan Gupta for their enlightening technical advice.

## References

- [1] J. Rabaey et al, "PicoRadios for Wireless Sensor Networks: The Next Challenge in Ultra-Low-Power Design," *Proc. of ISSCC*, pp. 200-201, Feb 2002.
- [2] K. Itoh, "Low Voltage Memories for Power-Aware Systems," *Proc. ISLPED*, pp. 1-6, Aug. 2002.
- [3] S. Kaxiras, Z. Hu, and M. Martonosi, "Cache decay: Exploiting generational behavior to reduce cache leakage power," *Proc. of ISCA*, pp. 240-251, Jun-Jul 2001.
- [4] K. Flautner et al, "Drowsy caches: simple techniques for reducing leakage power," *Proc. of ISCA*, pp. 148-157, May 2002.
- [5] J. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital Integrated Circuits: A Design Perspective*, 2<sup>nd</sup> edition, Prentice-Hall 2002.
- [6] E. Seevinck, F. J. List, and J. Lohstroh, "Static-noise margin analysis of MOS SRAM cells", *IEEE J. Solid-State Circuits*, vol. SC-22, No. 5, pp. 748-754, Oct. 1987.
- [7] C. Lage et al., "Soft error rate and stored charge requirements in advanced high-density SRAMs," *Proc. of IEDM*, pp. 821-824, Dec.1993.
- [8] K.D.T. Ngo and R. Webster, "Steady-state analysis and design of a switched-capacitor DC-DC converter," *Proc. of PESC*, pp. 378-385, Jun-Jul 1992.