Sense Amplifier Offset Characterisation and Test Implications for Low-Voltage SRAMs in 65 nm

Dhruv Patel*†, Derek Wright*, Manoj Sachdev*
Email: {dr3patel, derek.wright, msachdev}@uwaterloo.ca
*Department of Electrical and Computer Engineering, University of Waterloo, ON N2L 3G1, Canada;
†Department of Electrical and Computer Engineering, University of Toronto, ON M5S 3G4, Canada;

***Abstract*—Variability in offset voltage, bitcell transistor conductance, and leakage currents can lead to marginal and intermittent failures in low-voltage SRAMs. In this paper, we develop a model of these marginal faults that includes such sense amplifier and bitcell variability. Using simulations and measurement data from a 65 nm test chip, we investigate the likelihood of these failures and propose how to stimulate their occurrence during testing.**

***Keywords*— SRAM, Sense Amplifier, Offset Voltage, Test Algorithms, Threshold Voltage Mismatch, Weak Cell Fault**

I. Introduction

Static random access memories (SRAMs) often occupy a significant area of integrated circuits (ICs) and therefore significantly determine their energy consumption, yield, and reliability [1]. SRAM test complexity emerges from (i) the integration of a large number of SRAM bits; (ii) bitcell and peripheral circuit sensitivity to process, voltage, and temperature (PVT) variations and manufacturing defects; and (iii) the incomplete observability and controllability of critical internal nodes. Design for Test (DfT) circuits, such as Built-In Self-Test (BIST), can be added to increase the test fault coverage or to speed its execution, but this requires additional die area. So, it is preferable to achieve these goals exclusively via test algorithms and operating conditions when possible [2-4].

When a fault is observed during SRAM test algorithm execution, it is not easy to determine its root cause [5-7]. For example, Sense Amplifier (SA) offset, V_{OS}, related failures may appear as intermittent, weak cell failures, or column failures. We postulate that a weak bitcell coupled with a high offset voltage SA will fail under certain conditions. The objective of this research is to bring this to the attention of the test community, to investigate the conditions necessary for their testing.

The rest of the paper is organized as follows: Section II provides a background on SA operation and non-idealities. Section III develops a statistical yield model for the failure mode of a marginal SA coupled with a marginal bitcell. Section IV presents simulation and measurement results that serve as inputs to the statistical model. Section V applies these results to the model and proposes a test algorithm to detect such faults. Finally, Section VI concludes the paper.

II. SRAM Sense Amplifier Operation and Non-Idealities

An SA interprets and amplifies the Boolean values stored in a bitcell in a time and energy efficient manner. An idealized SA is shown in Fig. 1(a), where the inputs are tied to the bitlines (BL

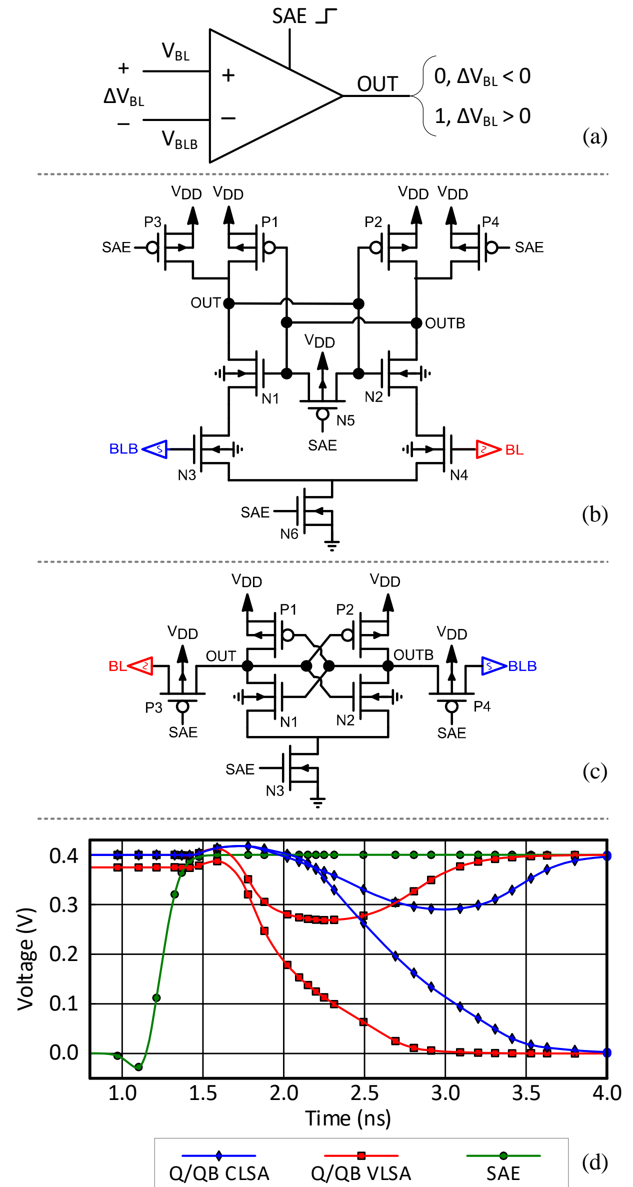


Fig. 1. (a) Idealized sense amplifier with typical connections. The output of the SA is latched when Sense Amp Enable (SAE) goes high and is determined by the differential input voltage, called ΔV_{BL}; (b) CLSA, and (c) VLSA conventional sense amplifier topologies; (d) Simulated transient response showing that the VLSA converges faster than the CLSA under particular operating conditions (V_{DD} = 0.4 V, ΔV_{BL} = -25 mV, T = 20 °C, TT process corner).

and BLB) of a column of bitcells. A differential voltage, ΔV_{BL}, is developed on the bitlines. Then the SA Enable signal (SAE) causes the SA to converge to a digital output. The output is '1' if $V_{BL} > V_{BLB}$ and '0' if $V_{BLB} > V_{BL}$.

Two popular SA topologies are the voltage latch SA (VLSA) and the current latch SA (CLSA), which are both shown in Fig. 1(b) and (c), respectively. VLSAs use the ΔV_{BL} present on the bitlines to pre-bias the SA before enabling it, and the resulting imbalance causes the SA to converge to the correct value. CLSAs equalize the internal nodes before enabling the SA, then rely on ΔV_{BL} to modulate the inrush of current. The asymmetric current flow causes the SA to converge to the correct value. The choice of particular architecture is dependent on the technology node used [8,9]. The simulated transient operation of a VLSA and CLSA designed in 65 nm CMOS (iso-area) is shown in Fig. 1(d).

How ΔV_{BL} is generated deserves greater explanation. During a read operation, the bitlines are both precharged to V_{DD}. Then, after activating the wordline (WL) of interest, the voltage of bitline coupled to the node storing '0' drops sufficiently to create ΔV_{BL}. The SAE signal causes the SA to latch the digital output corresponding to which bitline voltage dropped. However, the resultant $|\Delta V_{BL}|$ must be large enough to exceed the non-idealities of the SA reliably, called $\Delta V_{BL\text{-}min}$.

SA V_{OS} arises from mismatches in the gain factors, drain currents, threshold voltages, V_t, and layouts of the otherwise matched transistors that constitute the SA [10-12]. Among these factors, V_t mismatch of the sensing and input transistors is dominant [13,14]. Unfortunately, aggressive device scaling has resulted in increased device variations, which leads to increased threshold voltage mismatches [15,16]. A wider spread in the V_{OS} distribution (higher σ_{VOS}) has a significant negative impact on SRAM speed and read access yield, Y_{read}. The analysis in [17] shows that for a 16 Mb SRAM with Y_{read} of 97% and implemented in 28 nm technology, every 1 mV increase in the input offset voltage of the SA requires a 10 mV increase in $\Delta V_{BL\text{-}min}$.

At the minimum, $\Delta V_{BL\text{-}min}$ must be greater than $|V_{OS}|$. Other factors like noise voltage, temperature, changes in V_{DD}, and so on, only serve to increase $\Delta V_{BL\text{-}min}$ above this minimum. For an entire SRAM array, one must determine the worst-case (largest) $\Delta V_{BL\text{-}min}$ of every SA and ensure that the worst-case (smallest) ΔV_{BL} developed on the bitlines under given timing and voltage constraints exceeds this threshold.

During testing, violating $\Delta V_{BL\text{-}min}$ would appear as a read or read stability fault, and in the worst case, it may appear as an SA stuck-at fault. It is impractical to measure the V_{OS} of every SA to determine the worst-case. Often, Monte Carlo or corner simulations can provide statistics on the distribution of these values, from which an appropriate $\Delta V_{BL\text{-}min}$ can be estimated. However, simulation accuracy in capturing real-life situations depends on its underlying models, and such estimations should be interpreted with an abundance of caution.

III. Coupled SA and Bitcell Marginal Faults

Consider the case when an SA has a ΔV_{OS} very close to, but just under the ΔV_{BL} developed on the bitlines. This marginal case is dangerous because the SRAM may pass standard testing procedures nearly 100% of the time. However, it may experience intermittent in-field faults because transient and operating conditions cause ΔV_{BL} to dip below ΔV_{OS}.

To develop a model of such an occurrence, we begin by letting V_{OS} and the bitcell discharge current, I_{ON}, be Gaussian random variables. Consider the ΔV_{BL} generated by discharging the bitline through the bitcell NMOS transistors (both the access and driver). During bitline discharge, the bitcell is approximately a constant current source with discharge current I_{ON}. If the bitlines are precharged to V_{DD}, have capacitance C_{BL}, and neglecting leakage currents from half-selected cells, then after time Δt, the voltage difference will be

$$\Delta V_{BL} = \frac{I_{ON}\Delta t}{C_{BL}} \tag{1}$$

Since I_{ON} and C_{BL} will be subject to process variation, $\Delta V_{BL}(\Delta t)$ will be a random variable. Fig. 2 illustrates this point by showing an exaggerated distribution of generated ΔV_{BL} for both stored '0's and '1's (ΔV_{BL0} and ΔV_{BL1}, respectively) along with a distribution of V_{OS}. If a bitcell has a low $|\Delta V_{BL}|$ but its SA has a high $|\Delta V_{OS}|$ as highlighted in Fig. 2, then it is possible for the SA to latch the wrong value. In this marginal case, the bitcell is storing the correct value and the SA works, but together they fail because they are both marginal in their respective performance. One may shift the bitcell curves away from the origin (increase $|\Delta V_{BL}|$) by increasing the discharge time, Δt.

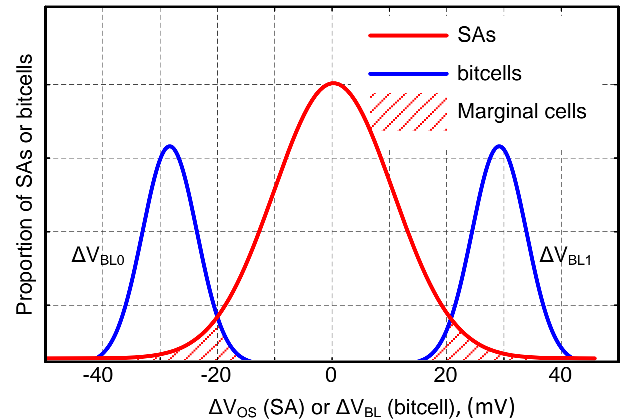


Fig. 2. An illustration of the risk imposed by an accidental pairing of marginal bitcells with marginal SAs.

The probability of a marginal bitcell being paired with a marginal SA is determined as follows, similar to the development in [18]. Let X_{VOS} be a Gaussian random variable representing V_{OS} of the SAs with a mean of 0 V. Its PDF is:

$$f_{X_{VOS}}(v, \sigma_{VOS}) = \frac{1}{\sigma_{VOS}\sqrt{2\pi}} \exp\left(-\frac{v^2}{2\sigma_{VOS}^2}\right), \tag{2}$$

where σ_{VOS} is the standard deviation of V_{OS} and v is the voltage of interest. Let $Y_{\Delta VBL}$ be a Gaussian random variable representing ΔV_{BL} generated by the bitcells. Its PDF is:

$$f_{Y_{\Delta VBLj}}(v, \mu_{\Delta VBLj}, \sigma_{\Delta VBLj}) = \frac{1}{\sqrt{2\pi}\cdot\sigma_{\Delta VBLj}} \exp\left[-\frac{(v-\mu_{\Delta VBLj})^2}{2\sigma_{\Delta VBLj}^2}\right], \tag{3}$$

where $\mu_{\Delta VBLj}$ and $\sigma_{\Delta VBLj}$ are the mean and standard deviation of ΔV_{BLj}, respectively, and j refers to the '0' or '1' case. Since there are l bits per column, one must consider the worst case of l bitcells. For ΔV_{BL1}, the probability of choosing at least one of l bitcells with $\Delta V_{BL1} < v$ is equivalent to

$$\begin{aligned} P_{BLF1}(v) &= (\exists i) P(\Delta V_{BL1,i} < v) \\ &= 1 - (\forall i) P(\Delta V_{BL1,i} \geq v), \end{aligned} \tag{4}$$

where BLF1 means bitline failure in the '1' case. That is, it is easier to find the probability of the inverse case of all bitcell ΔV_{BL1} exceeding v, equivalent to

$$\begin{aligned} (\forall i \in (0, l-1)) P(\Delta V_{BL1,i} \geq v) &= [P(\Delta V_{BL1} \geq v)]^l \\ &= \left[\int_v^\infty f_{Y_{\Delta VBL}}(v_Y, \mu_{\Delta VBL}, \sigma_{\Delta VBL}) \partial v_Y\right]^l \\ &= \left[\frac{1}{2}\operatorname{erf}\left(\frac{\mu_{\Delta VBL} - v}{\sqrt{2}\cdot\sigma_{\Delta VBL}}\right) + \frac{1}{2}\right]^l \end{aligned} \tag{5}$$

V_{OS} and ΔV_{BL1} are assumed to be independent, so a failure is expected when $\Delta V_{BL1} < \Delta V_{OS}$, which, using (4) and (5), will occur with probability

$$\begin{aligned} P(\Delta V_{BL1} < V_{OS}) &= \iint_{0\ 0}^{V_{DD}\ v_X} f_{X_{VOS}}(v_X) \cdot P_{BLF}(v_y) \partial v_Y \partial v_X \\ &= \int_0^{V_{DD}} f_{X_{VOS}}(v_X, \sigma_{VOS}) \cdot \left(1 - \left[\frac{1}{2}\operatorname{erf}\left(\frac{\mu_{\Delta VBL1} - v_X}{\sqrt{2}\cdot\sigma_{\Delta VBL1}}\right) + \frac{1}{2}\right]^l\right) \partial v_X \\ &= \frac{\int_0^{V_{DD}} \exp\left(-\frac{v_X^2}{2\sigma_{VOS}^2}\right)\cdot\left(1 - \left[\frac{1}{2}\operatorname{erf}\left(\frac{\mu_{\Delta VBL1} - v_X}{\sqrt{2}\cdot\sigma_{\Delta VBL1}}\right) + \frac{1}{2}\right]^l\right)\partial v_X}{\sqrt{2\pi}\sigma_{VOS}} \end{aligned} \tag{6}$$

If, as assumed, $\mu_{VOS} = 0$, and the ΔV_{BLj} distributions are symmetric and not strongly correlated, then $P(\Delta V_{BL1} < V_{OS}) = P(\Delta V_{BL0} > V_{OS})$, and the overall probability of failure would be double that of the result in (6). Equations (1) and (6) can be solved numerically to investigate the impact of various Δt to ensure a low enough probability of weak bitcells being paired with high V_{OS} SAs.

We can extend the model in [18] by considering the role of leakage, which we may use to our advantage in developing a strenuous test case. With reference to Fig. 3, assume that one bitcell in a column of l bitcells stores '1' while all other $l - 1$ store '0'. Then, we attempt to read the cell containing '1'. Without leakage, BL will remain at V_{DD} while BLB will discharge to $V_{DD} - \Delta V_{BL}$ through the QB node of the cell storing '1'. However, with leakage, every cell storing '0' will experience $V_{DS} = V_{DD}$ across its access transistors from BL to Q and will experience a leakage current of I_{OFF} due to subthreshold conduction. Therefore, BL will also drop by ΔV_{leak} during the discharge of BLB, and ΔV_{BL} will be degraded by ΔV_{leak}. Furthermore, I_{OFF} is also a random variable due to process variation.

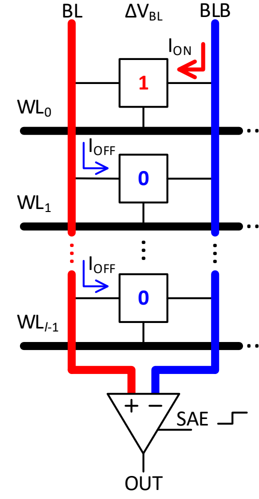


Fig. 3. Exploiting bitcell leakage to minimize ΔV_{BL} during testing.

ΔV_{leak} can be determined from I_{OFF} as in (7). Assuming that C_{BL} is approximately proportional to l, such that $C_{BL} = l\cdot C_{BL0}$, where C_{BL0} is the additional bitline capacitance per bitcell, then (8) expresses ΔV_{BL} including leakage.

$$\Delta V_{leak} = \frac{(l-1)I_{OFF}\Delta t}{C_{BL}} \tag{7}$$

$$\therefore \Delta V_{BL} = \frac{\Delta t[I_{ON} - (l-1)I_{OFF}]}{lC_{BL0}} \tag{8}$$

Jitter, layout asymmetries, and process variation also affect C_{BL} and Δt, but we restrict our analysis to V_{OS}, I_{ON}, and I_{OFF} since leakage current is often a dominant contributor in modern technology nodes. If I_{ON} and I_{OFF} are normal with μ_{ION}, σ_{ION}, μ_{IOFF}, and σ_{IOFF}, respectively, then

$$\mu_{\Delta VBL} = \frac{\Delta t[\mu_{ION} - (l-1)\mu_{IOFF}]}{lC_{BL0}}, \text{ and} \tag{9}$$

$$\sigma_{\Delta VBL} = \frac{\Delta t\sqrt{\sigma_{ION}^2 + (l-1)^2\sigma_{IOFF}^2}}{lC_{BL0}}. \tag{10}$$

Using (6), (9), and (10), one may choose an l and Δt that places $P(\Delta V_{BL1} < V_{OS})$ below a threshold to achieve a particular yield.

IV. Sense Amplifier Characterization

We now examine SAs in more detail to determine the V_{OS} statistics of two popular SA architectures for use in the model developed in Section III. We also characterize the I_{ON} and I_{OFF} statistics via Monte Carlo simulation.

A. Simulation Results

Since the model in Section III considers the impact of leakage, it is useful to know what test conditions minimize I_{ON}/I_{OFF}. The role of temperature, V_{DD}, and column size (l) were investigated using a fixed-width access transistor with $V_{DS} = V_{DD}$. As expected, Fig. 4(a) and (b) show that this ratio drops as the temperature increases, as V_{DD} decreases, and as l increases. These are sensible results since leakage increases with temperature, the single NMOS conducting I_{ON} is pushed into the subthreshold region as V_{DD} decreases, and there are more and more bitcells contributing to I_{OFF} as l increases.

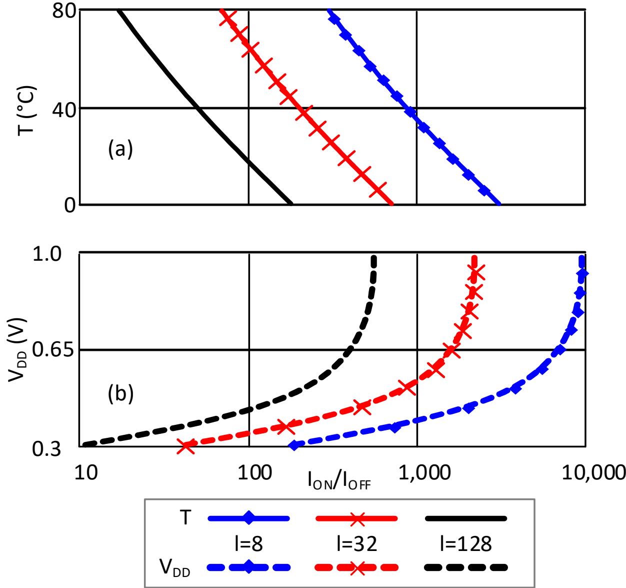


Fig. 4. Simulated results of conduction to leakage current ratio (I_{ON}/I_{OFF}) versus (a) temperature and (b) V_{DD} for various column sizes, l. These results validate the well-known strategy that a lower current ratio, and hence a more strenuous test environment, is achieved at high temperature and low voltage.

The I_{ON}, I_{OFF}, and C_{BL0} values were determined using the 65 nm technology node simulation model parameters in Table I. The I_{ON} and I_{OFF} values were determined via Monte Carlo simulations of 2000 bitcells. C_{BL0} was determined both from parasitic capacitance extraction from schematics and by modeling the metal parallel plate and coupling capacitances of the M2 bitlines assuming a 0.25 μm pitch cell.

B. Test Chip Implementation and Measurement Setup

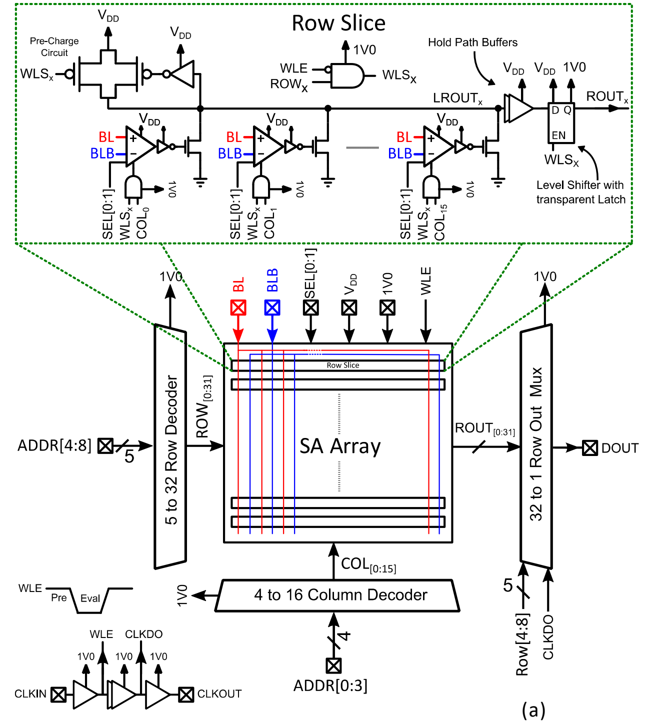
An array of various SA architectures was fabricated in a 65 nm general-purpose CMOS process. Fig. 5(a) shows the architecture of the test chip and a schematic of a row slice. There are 32 rows each of CLSAs and VLSAs, and each row has 16 cells arranged in 16 columns. Thus, the array contains 512 CLSAs and 512 VLSAs that are individually addressable with a row and a column decoder. The BL and BLB signals are routed vertically throughout the array. The ΔV_{BL} signal is supplied by setting appropriate voltages at the BL and BLB analog pads.

The test bench setup is shown in Fig. 5(b). Externally provided parameters and signals, such as V_{DD}, CLK, BL and BLB (ΔV_{BL}) are set via power supplies and signal generators, and a logic analyzer captures the output (DOUT). The

TABLE I. MODEL PARAMETERS

Parameter	Value	Method
Corner	TT	Chosen
W_{access}	200 nm	Chosen
V_{DD}	0.4 V	Chosen
T	80 °C	Chosen
C_{BL0}	0.209 fF/bitcell	Extracted from Schematic
μ_{ION}, σ_{ION}	6.78, 1.78 μA	Monte Carlo
μ_{IOFF}, σ_{IOFF}	3.52, 1.56 nA	Monte Carlo

environment is controlled by placing the test chip and PCB in a temperature chamber. Automation software (LabVIEW™) controls these devices to apply various test scenarios and to capture the raw DOUT data for a given cell address. The captured raw data is then processed and plotted.



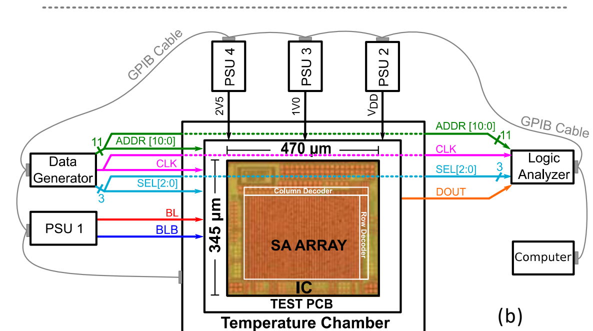


Fig. 5. (a) Test chip architecture, including pads and a row slice; (b) Test bench setup.

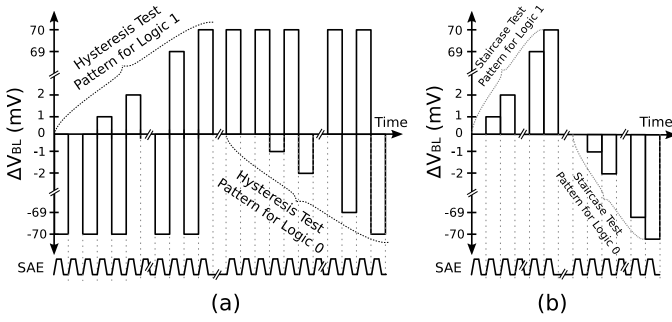


Fig. 6. (a) Hysteresis test pattern (b) Staircase test pattern. The worst-case test pattern characterizes a sense amplifier's minimum required ΔV_{BL} for making reliable decisions and corresponds to the pattern requiring the highest $|\Delta V_{BL}|$ to make a correct decision.

Fig. 6 summarizes two test strategies used to exercise the SA test chip. In both cases, the procedure starts with $\Delta V_{BL} = 0$ V and increases until the SA returns a correct result. It is then tested by applying decreasing values of ΔV_{BL} until the opposite value is returned. For example, $\Delta V_{BL} > 0$ corresponds to a correct value of logic '1', but if there is a negative skew to the cell of $V_{OS} = +10$ mV, then the cell will continue to return logic '0' until ΔV_{BL} exceeds 10 mV. This test is performed on all cells to determine the minimum $|\Delta V_{BL}|$ necessary for 100% of the SAs to return the correct value.

Since the SA cells may exhibit a memory effect, which could be due to many reasons such as parasitic influences of neighboring cells or charge retention from the previous read on the same cell, a test pattern was designed to reveal and mitigate these effects. Therefore, a 'Hysteresis' test pattern described by Fig. 6(a) was developed to test the sensitivity of the SAs to the previous read cycle. Before applying a test ΔV_{BL} to the SA, the inverse data was applied with a high enough ΔV_{BL} (±100 mV) to ensure that the SA evaluated the correct, but opposite logic. For example, applying $\Delta V_{BL} = -100$ mV forces the cell to return logic '0', then $\Delta V_{BL} > 0$ V is applied to test its ability to return logic '1'. This sequence was repeated for each test value of ΔV_{BL} and is expected to clear any confounding memory effect.

A second 'Staircase' test pattern, shown in Fig. 6(b), does not include the predischarge step of applying an opposite ΔV_{BL} before testing the cell. It begins at $\Delta V_{BL} = 0$ V and either increases or decreases stepwise until the cell consistently returns the correct value. Thus, any memory effect should be revealed by examining the difference between these two test cases.

C. *Measurement Results*

Ten test ICs were fabricated and tested, and were designated IC_1 through IC_10. The differences between the various $\Delta V_{BL\text{-}min}$ determined by the Hysteresis and Staircase test methods were negligible; we did not detect a significant memory effect in the SAs. Across 10 ICs and various V_{DD} voltages, the means of the differences (Hysteresis – Staircase) was less than 1 mV in all cases, which is small relative to the calculated $\Delta V_{BL\text{-}min}$ values. There may truly be a memory effect present that is not evident due to the nature of the test chip architecture. For example, it was difficult to adequately control the frequency and slew rate as the primary objective of the test chip design was to characterize V_{OS} of a variety of SA architectures. Nevertheless, the Hysteresis test pattern was used to derive the data presented in this section, and future research will further examine the memory effect of SAs.

All ten ICs were tested by sweeping ΔV_{BL} from –100 mV to +100 mV in 1 mV increments. Fig. 7 shows the PDF of SAs flipped to logic '1' as a function of ΔV_{BL} for all ten test chips, essentially a measure of V_{OS}. The tighter distribution of the VLSA means that it requires a lower $\Delta V_{BL\text{-}min}$. The ten ICs were

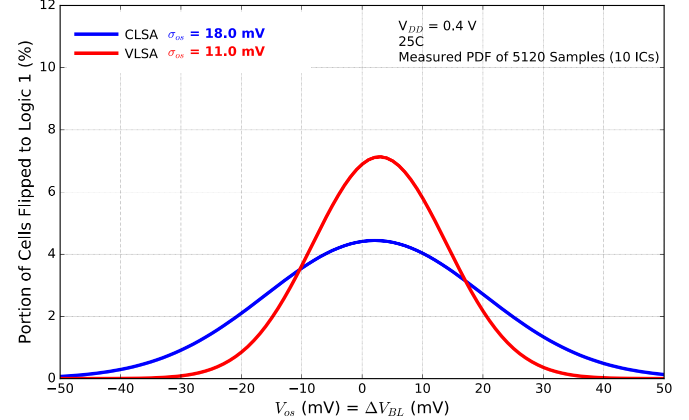


Fig. 7. A PDF of the measured percent of SA cells out of 5120 flipped to logic '1' by sweeping ΔV_{BL} from -50 mV to +50 mV at a resolution of 1 mV using the hysteresis test pattern from Fig. 6(a). and V_{DD} = 0.4 V.

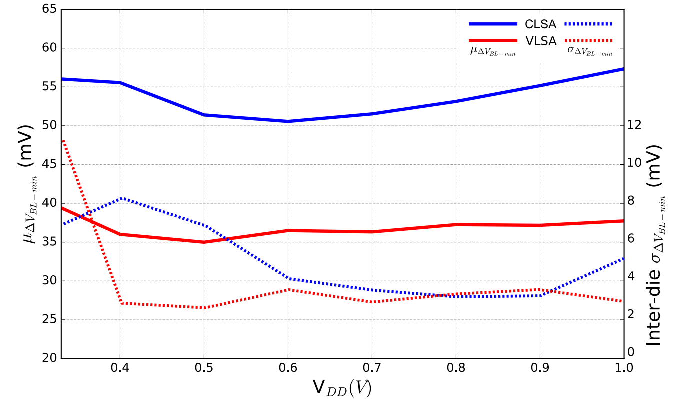


Fig. 8. Measured mean, μ, and standard deviation, σ, of $\Delta V_{BL\text{-}min}$ across ten test chips vs. V_{DD}.

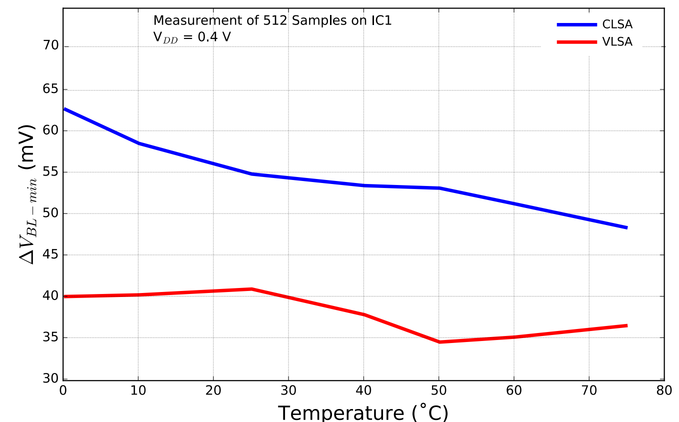


Fig. 9. Measured $\Delta V_{BL\text{-}min}$ of 512 cells of IC_1 vs. temperature.

then tested with V_{DD} ranging from 0.4 V through 1.0 V at 0.1 V increments to investigate inter-die behavior. Fig. 8 shows the mean and standard deviation of $\Delta V_{BL\text{-}min}$ for all ten ICs ($\mu_{\Delta VBL\text{-}min}$ and $\sigma_{\Delta VBL\text{-}min}$) versus V_{DD}. Fig. 9 shows that the VLSA continued to outperform the CLSA across a wide range of operating temperatures for IC_1.

V. Yield Calculations and Test Algorithm

Using the results summarized in Table I, the VLSA measurement results in Fig. 7, (6), (9), and (10), we calculate the failure rate for various l and Δ_t, both including and excluding leakage current. Fig. 10 shows numerical results that illustrate how the predicted failure rate increases when leakage is taken into account (solid lines) and neglecting leakage current (dashed lines). The grey dashed lines indicate 1% and 3σ failure rates. As a representative example, for $l = 256$ at 1% and 3σ failure rates, the model predicts that leakage effects can account for a sizeable increase in the number of detected failures. Since these leakage effects are data dependent, the real-world failure rate would be somewhere between these two extremes. Therefore, it is sensible to apply a test pattern as illustrated in Fig. 3 to check for marginal cases. Otherwise, a weak bitcell coupled with an unfavorably high V_{OS} SA may escape detection.

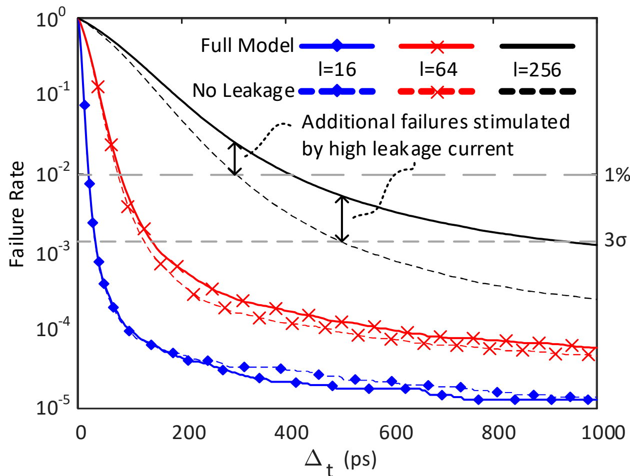


Fig. 10. Predicted failure rates both using the full statistical model and neglecting leakage currents. The dashed lines indicate 1% and 3σ failure rates. ΔV_{BL} increases with Δt, lowering the failure rate. C_{BL} increases with l, slowing the development of ΔV_{BL} and increasing the failure rate for a given Δt. The increase in failure rates due to leakage current for $l = 256$ at 1% and 3σ is indicated in the figure.

VI. Discussion and Conclusions

We have investigated how one can test the pass/fail limits of SAs in light of how closely they are coupled with the bitcells in an SRAM. Simulations and measurements show that strenuous test conditions can be created by reducing voltage and increasing temperature, and by applying a data pattern that maximizes the detrimental role of leakage. Testing under these conditions may reveal weak cell faults that are not otherwise evident. Designing and fabricating SAs with bitcells that have known or controllable characteristics will allow a more in-depth investigation into the impact that coupled weak cell faults may play in SRAM reliability.

References

[1] K. Zhang, K. Hose, V. De, and B. Senyk, “The scaling of data sensing schemes for high speed cache design in sub-0.18 µm technologies,” VLSI Symp. Tech. Dig., 2000, pp. 226–227.

[2] J. Kinseher, M. Voelker, and I. Polia, “Improving Testability and Reliability of Advanced SRAM Architectures,” *IEEE Trans. Emerg. Topics Comput.*, 2017.

[3] C. Yeh, Y.-N. Liu, J.-S. Wang, and P.-Y. Chang, “Variation-resilient voltage generation for SRAM weak cell testing,” *2011 9th IEEE Int. Conf. on ASIC*, Xiamen, 2011, pp. 248-251.

[4] B. S. Mohammad, H. Saleh and M. Ismail, “Design Methodologies for Yield Enhancement and Power Efficiency in SRAM-Based SoCs,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst*, vol. 23, no. 10, pp. 2054-2064, Oct. 2015.

[5] A. J. van de Goor, S. Hamdioui and R. Wadsworth, “Detecting faults in the peripheral circuits and an evaluation of SRAM tests,” *Proc. IEEE Int. Test Conf.*, 2004, pp. 114-123.

[6] A. Pavlov and M. Sachdev, “SRAM Cell Stability: Definition, Modeling and Testing,” in *CMOS SRAM Circuit Design and Parametric Test in Nano-Scaled Technologies*, Springer, 2008.

[7] D. Khalil, M. Khellah, N. S. Kim, Y. Ismail, T. Karnik and V. De, “SRAM dynamic stability estimation using MPFP,” *2007 Int. Conf. on Microelectronics*, Cairo, 2007, pp. 167-170.

[8] M. Sinha, S. Hsu, A. Alvandpour, W. Burleson, R. Krishnamurthy, and S. Borkar, “Low voltage sensing techniques and secondary design issues for sub-90 nm caches,” *Proc. ESSCIRC*, Sep. 2003, pp. 413–416.

[9] Mohammad, Baker, et al. "Comparative study of current mode and voltage mode sense amplifier used for 28nm SRAM." *2012 24th International Conference on Microelectronics (ICM)*. IEEE, 2012.

[10] S. Lovett, G. Gibbs, and A. Pancholy, “Yield and matching implications for static RAM memory array sense-amplifier design,” *IEEE J. Solid-State Circuits*, vol. 35, no. 8, pp. 1200–1204, Aug. 2000.

[11] D. Laurent, “Sense amplifier signal margins and process sensitivities,” *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 3, pp. 269–275, Mar. 2002.

[12] J. S. Shah, D. Nairn, and M. Sachdev, “Am Energy-Efficient Offset-Cancelling Sense Amplifier.” *IEEE Trans. Circuits and Systems –II*, 2013.

[13] A. Bhavnagarwala, S. Kosonocky, C. Radens, K. Stawiasz, R. Mann, Y. Qiuyi, and K. Chen, “Fluctuation limits and scaling opportunities for CMOS SRAM cells,” *IEEE IEDM Tech. Dig.*, Dec. 2005, pp. 659–662.

[14] S. H. Woo, H. Kang, K. Park and S. O. Jung, “Offset voltage estimation model for latch-type sense amplifiers,” *IET Circuits, Devices & Systems*, vol. 4, no. 6, pp. 503-513, Nov. 2010.

[15] J. Pineda de Gyvez and H. P. Tuinhout, “Threshold voltage mismatch and intra-die leakage current in digital CMOS circuits,” *IEEE J. Solid-State Circuits*, vol. 39, no. 1, pp. 157–168, Jan. 2004.

[16] M. Steyaert *et al.*, “Threshold voltage mismatch in short-channel MOS transistors,” *Electronics Letters*, vol. 30, no. 18, pp. 1546–1548, Sept. 1994.

[17] M. Abu-Rahma, Y. Chen, W. Sy, W. L. Ong, L. Y. Ting, S. S. Yoon, M. Han, and E. Terzioglu, “Characterization of SRAM sense amplifier input offset for yield prediction in 28 nm CMOS,” *Proc. IEEE CICC*, Sep. 2011, pp. 1–4.

[18] P. Sharma, A. K. Gundu and M. S. Hashmi, “Modeling and yield estimation of SRAM sub-system for different capacities subjected to parametric variations,” *20th Int. Symp. VLSI Design and Test (VDAT)*, pp. 1-6, 2016