Impact of Random Telegraph Signals on V_{min} in 45nm SRAM

Seng Oon Toh, Yasumasa Tsukamoto†, Zheng Guo, Lauren Jones, Tsu-Jae King Liu, and Borivoje Nikolić
Dept. of Electrical Engineering and Computer Sciences, University of California, Berkeley 94720, USA
†Renesas Technology Corp., 4-1, Mizuhara, Itami, Hyogo, 664-0005, Japan

Abstract

An alternating-bias random telegraph signal (RTS) characterization technique is presented, which shortens measurement time by 10x and also produces more accurate statistical distributions of RTS amplitudes. Measurements of RTS amplitudes in 45nm SRAM transistor I_{ds} and cell write margin are reported and used to demonstrate a complex dependence of write margin on RTS in multiple transistors. Fail bit rate of SRAM with RTS is estimated using a statistical model populated by I_{write} measurements. Statistical analysis indicates a V_{min} degradation of less than 50 mV due to RTS.

Introduction

Large temporal fluctuations of threshold voltage (V_{th}) in highly scaled CMOS transistors have been reported and attributed to random telegraph signals (RTS) [1, 2]. This paper presents a measurement technique for sampling worst-case V_{th} variation due to RTS. This technique is used to characterize RTS amplitude distributions in SRAM transistors of the 45nm technology. A padded-out SRAM cell array is utilized to extract RTS-induced fluctuations in SRAM write margin as well as to identify the transistors that contribute the most to write-margin fluctuation. Finally, a numerical method is developed to estimate V_{min} degradation for write operation in large SRAM arrays. This method is validated using V_{min} measurement data for a 64kb SRAM array.

Experimental Setup

A 6T SRAM testchip is fabricated in an industrial 45nm CMOS process with $0.252\mu m^2$ bit-cells. The testchip contains four 64kb SRAM arrays as well as a macro with internal nodes of 160 cells padded out through a switch network (Fig. 1) [3]. Source-meters are used to access these internal nodes using precision 4-terminal Kelvin sensing methods to accurately set voltages at all nodes. RTS measurements are conducted at a sampling rate of 60 Hz. The SRAM array is used for measuring V_{min} while the SRAM macro is used for measuring I_{ds} and I_{write}.

Alternating-Bias RTS Measurement Technique

RTS amplitude measurements are conventionally [1, 2] performed by measuring the drain current (I_{ds}) of the transistor under a constant gate bias. Long measurement periods are required to observe RTS-related fluctuations caused by deep traps with long time constants. This makes it prohibitive to analyze a large population of transistors, to obtain statistics necessary for estimation of the properties of large SRAM arrays. A measurement technique is therefore introduced to accelerate the oxide trapping and de-trapping processes by

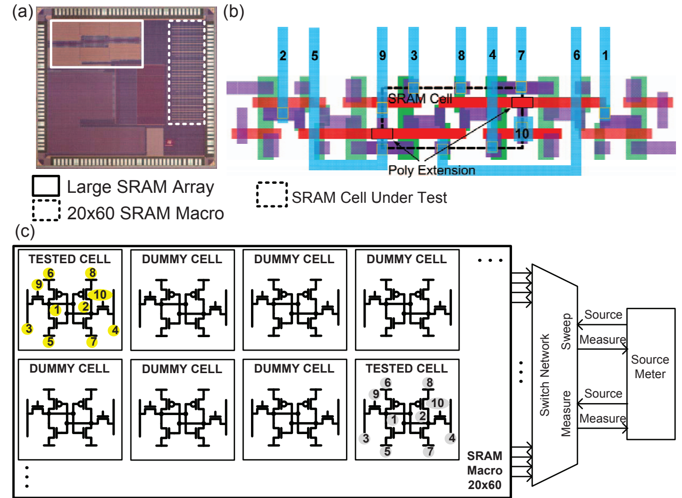

Fig. 1. (a) Die photo of the testchip. (b) Layout sketch of the padded out SRAM cell with poly connection for accessing internal nodes. (c) Experimental setup for the padded out cell RTS measurement.

pulsing the gate bias to a high stress voltage (V_{stress}) prior to sampling the current, and subsequently to a negative voltage ($V_{accumulation}$) prior to sampling the current (Fig. 2). Gate biases of V_{stress} and $V_{accumulation}$ force trapping and de-trapping of the oxide trap. The trap remains occupied/empty even though stress voltage is removed because instantaneous trap occupancy converges to a new steady-state value as a decaying exponential [4]. Fig. 3 plots the drain currents measured using this technique, demonstrating the 10x decrease in measurement time required to observe similar RTS fluctuation.

Figs. 4 and 5 compare the statistical distributions of I_{ds}, as measured with constant time using the conventional and proposed techniques. A significant difference between the statistical distributions obtained from these two techniques is observed in the bulk of the distributions. The conventional technique underestimates the magnitude of RTS fluctuations in the bulk of the distribution due to insufficient measurement period. Alternately, measurement period of the conventional technique can be increased while sacrificing the sample population that can be collected within a reasonable amount of time.

RTS in 45nm SRAM Transistors

Fig. 6 plots statistical distributions of RTS amplitudes that are measured from transistors in SRAM cells. The $\geq$ 4x difference in RTS amplitudes between PMOS and NMOS devices can be explained by the difference in the capture cross-section between electrons and holes [2]. The pull-down

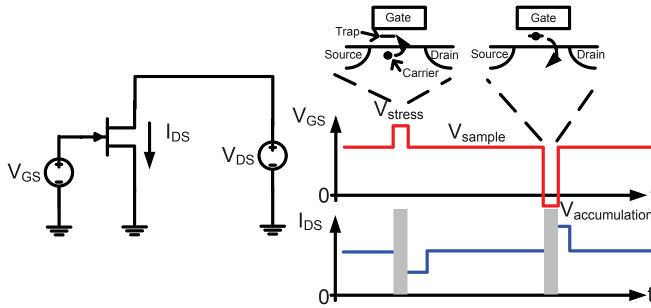

Fig. 2. Waveforms of the alternating-bias technique.

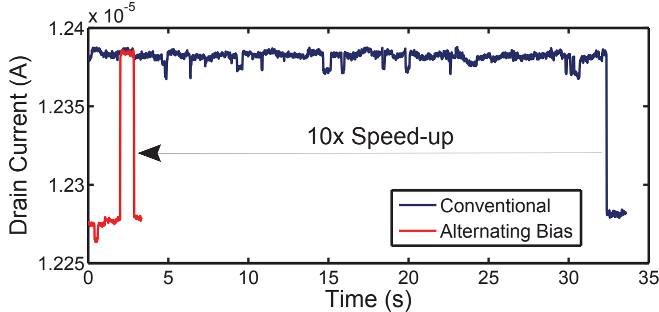

Fig. 3. Drain current of the same transistor measured using conventional and alternating-bias techniques. It takes 32s for a trap to be occupied in the conventional technique. By applying an initial 0.5ms gate stress before measurement, the trap is occupied from the start of the experiment.

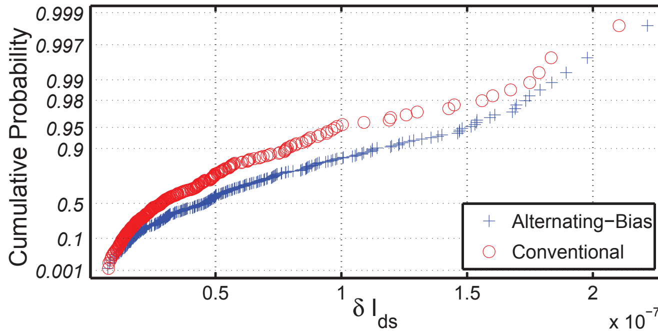

Fig. 4. Gumbel plots of RTS drain current fluctuations measured using conventional and alternating-bias techniques with constant time.

transistors exhibit slightly larger RTS amplitudes compared to the pass-gate transistors even though the larger effective area transistor should have a smaller fluctuation (Eqn. 1) [5]. A similar trend is also reported in [1]. This discrepancy is partially due to a 6% higher mobility in the pass-gate devices compared to the pull-down devices.

$$\frac{\delta I_d}{I_d} = \frac{1}{W_{eff} \times L_{eff}} \left(\frac{1}{N} \frac{\delta N}{\delta N_T} \pm \frac{1}{\mu} \frac{\delta \mu}{\delta N_T} \right) \quad (1)$$

Impact of RTS on SRAM Write Margin

The I_{write} metric, derived from write N-curves [6] is used to characterize write margin fluctuation of the SRAM cells caused by RTS (Fig. 7-8). An SRAM cell requires a positive I_{write} to be writeable. This current-based metric is favored

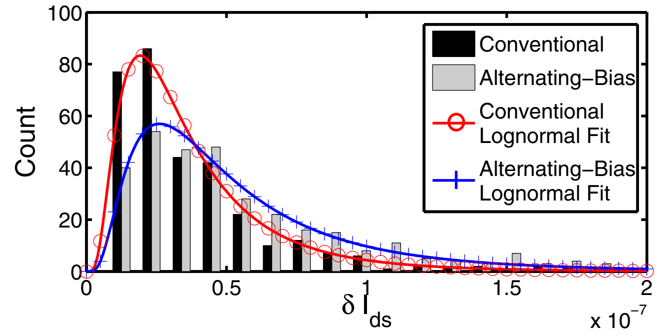

Fig. 5. Histogram and lognormal distribution fits of RTS I_{ds} fluctuations measured using conventional and alternating-bias techniques with constant time. The presented technique provides a more accurate sampling of the worst-case RTS amplitude, especially in the bulk of the distribution.

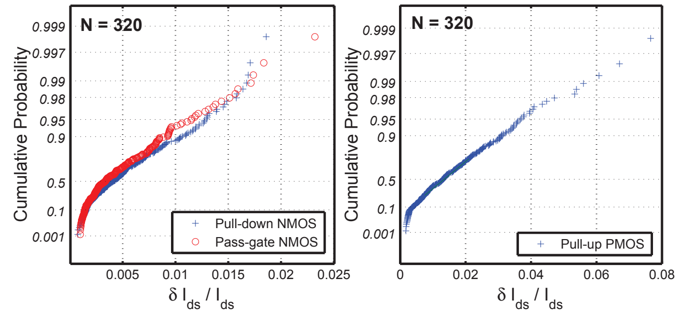

Fig. 6. Gumbel plots of normalized RTS fluctuations in drain currents, measured from transistors in the padded-out SRAM cells.

over voltage-based metrics, such as wordline and bitline write metrics [3], because the measured margin fluctuations can be easily correlated with drain current fluctuations observed in the transistors. Fig. 9 plots the RTS in I_{write} and its constituents, measured from one cell. Although PU5 contributes a small amount to the N-curve current, its RTS is reflected in I_{write} because RTS amplitudes are much larger for the PMOS transistors than for the NMOS transistors (Fig. 6). Fig. 10 plots the statistical distributions of nominal I_{write} and RTS at two operating voltages. Worst-case fluctuations in I_{write} are extracted by applying stress voltages to the pass-gate and pull-up transistors before measurement. These results indicate that RTS in SRAM write margin is dependent on both bias and RTS in multiple transistors, and requires a more accurate model than fixed shifts in V_{th} [2] or single transistor RTS [7].

Fail Bit Rate (FBR) Estimation

A statistical model of SRAM write failure is developed in order to estimate the impact of RTS on write margin in large arrays. Measured nominal I_{write} data is fitted to a normal distribution while RTS fluctuation data is fitted to a hybrid distribution (Fig. 11). This hybrid distribution allows accurate modeling of both the bulk and tail of the distribution. The joint probability density function (PDF) of these two distributions is obtained through numerical analysis (Fig. 12-13). The most probable failure point (MPFP) corresponds to the point in the design space where the first failure is likely to occur. The FBR

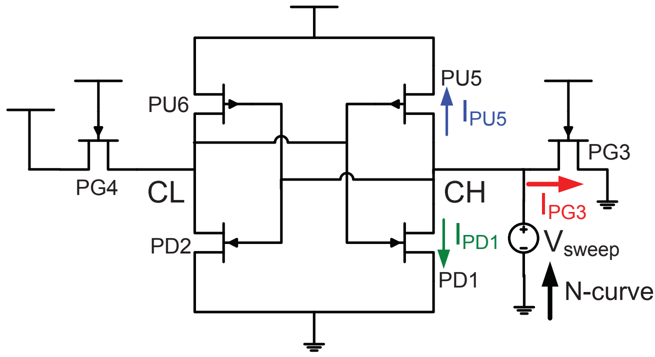

Fig. 7. Schematic of I_{write} measurement. The N-curve is the sum of the currents flowing out of the internal node.

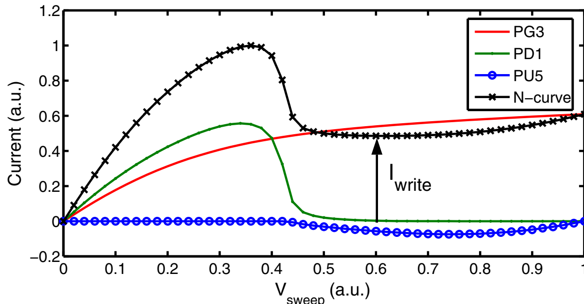

Fig. 8. Currents contributing to I_{write}. The I_{write} metric corresponds to the minimum point of the N-curve after the peak. It measures the relative strengths of PG3 and PD1 compared to PU5 during the critical phase of the write operation.

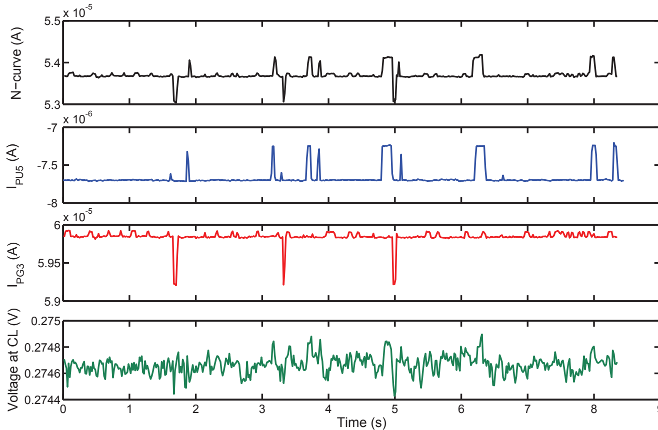

Fig. 9. Currents and voltages measured from padded-out SRAM cell with RTS. RTS characteristics of the N-curve are influenced by RTS in both PU5 and PG3. Voltage fluctuation at node CL is minimal due to the low impedance of this node.

of the SRAM is calculated by integrating the joint PDF up to the boundary defined by the probability corresponding to the MPFP.

Figs. 13 and 14 plot the joint PDFs, estimated based on measured I_{write} at nominal and low operating voltages. The MPFP of both operating conditions occurs within the same region, characterized by low nominal I_{write} and small RTS

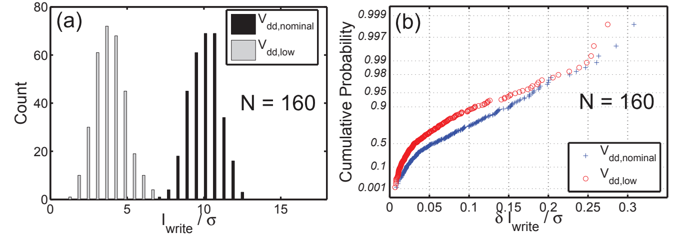

Fig. 10. (a) Histogram of nominal I_{write}. (b) Gumbel plots of I_{write} RTS fluctuation. The Maximum RTS amplitude normalized to σI_{write} at each operating voltage does not change significantly, although the shape of the distribution changes.

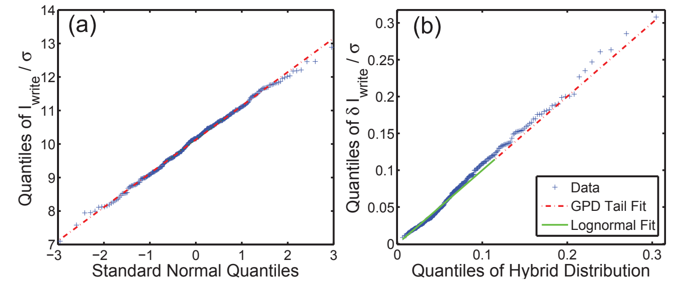

Fig. 11. (a) Quantile-quantile plot of I_{write} with normal distribution fit. (b) Quantile-quantile plot of RTS fluctuation in I_{write} with hybrid distribution fit. The bulk of the distribution is fitted using a log-normal distribution while data above a certain threshold is fitted to a generalized pareto distribution. Accurate fitting of the tail of the distribution is critical for estimating RTS-induced fluctuation in large arrays.

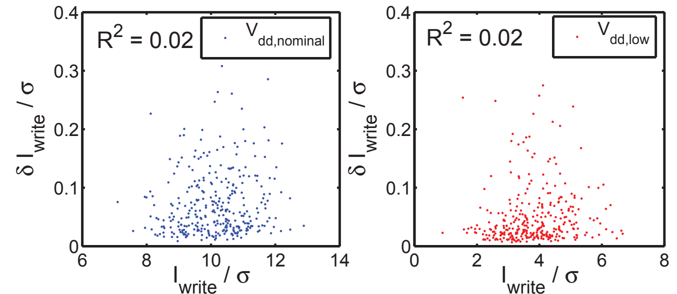

Fig. 12. Scatter plots of measured nominal I_{write} and RTS fluctuation at nominal and low operating voltages. The RTS amplitude appears to be uncorrelated with nominal I_{write} values.

fluctuation despite the different shapes of the joint PDFs. The proximity of the MPFP to small values of RTS fluctuation highlights the importance of modeling an accurate RTS fluctuation distribution in the bulk. The alternating-bias technique provides this accurate model.

FBR and V_{min} Degradation

The FBR of the SRAM cell is estimated at different voltages based on statistical distributions fitted to measured data (Fig. 15). The FBR degradation by RTS at the nominal voltage is minimal even in large SRAM arrays (10σ) because the MPFP is dominated by the bulk of the RTS fluctuation distributions rather than the tail. V_{min} degradation is estimated from Fig. 15 by observing the increase in V_{dd} required to maintain a

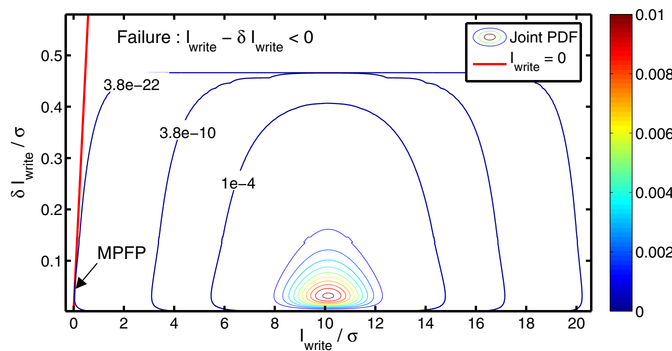

Fig. 13. Joint probability density function of nominal I_{write} and RTS fluctuation at nominal V_{dd}. The MPFP is found by tracing the write failure contour and finding the point with the highest probability.

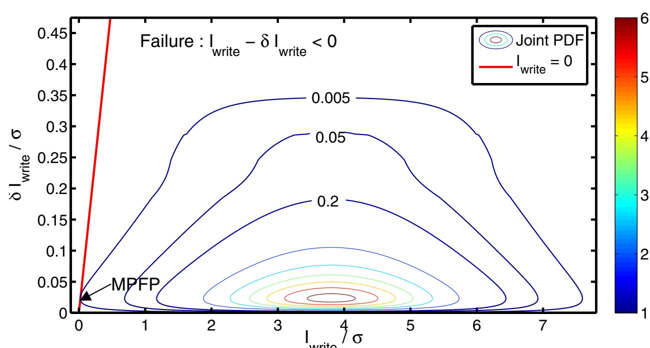

Fig. 14. Joint probability density function of nominal I_{write} and RTS fluctuation at low V_{dd}.

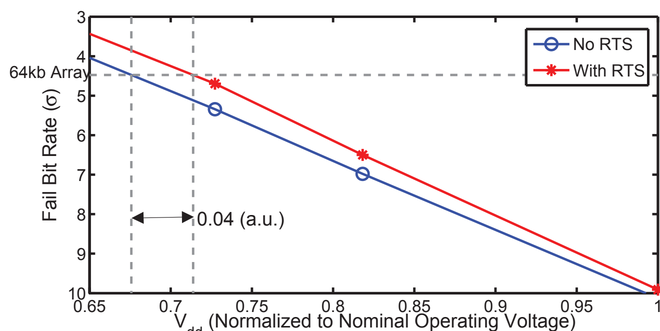

Fig. 15. FBR of SRAM at different voltages. FBR increases by 0.6σ at low voltages and 0.2σ at nominal operating voltage. V_{min} is degraded by less than 50mV for small arrays (4σ) and is not significant in large arrays.

similar FBR. The estimated V_{min} degradation corresponding to a 64kb SRAM array is experimentally verified by measuring V_{min} fluctuation in an actual array. The minimum (nominal) and maximum V_{min} values of each cell is measured and illustrated in Fig. 16 as fail bit count. Measured results (0.04 a.u.) matches estimated V_{min} degradation. Mismatch in the absolute V_{min} values between these two figures is caused by layout-induced differences between the 64kb SRAM arrays and the SRAM macros used for I_{write} measurements [3].

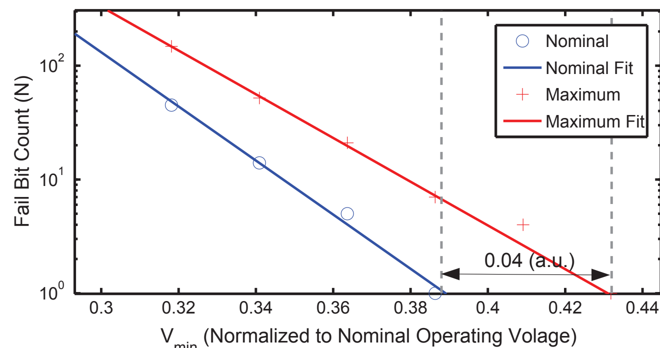

Fig. 16. Measured fail bit count of a 64kb SRAM array. Fail bit count measures the number of cells with V_{min} greater than a corresponding voltage. Extrapolation of fail bit count to 10^0 estimates the largest V_{min} in the array. Measured V_{min} degradation is less than 50 mV.

Conclusion

Statistical analysis confirmed by measurements indicates that V_{min} degradation due to RTS is less than 50 mV, even in the presence of large RTS fluctuations. Furthermore, degradation due to RTS becomes less significant in larger SRAM arrays. To demonstrate this, a measurement technique to accelerate RTS testing is presented. Measurements of worst-case RTS amplitudes in 45nm SRAM transistor I_{ds} and cell I_{write} show the complex dependence of write margin on RTS in multiple transistors. V_{min} degradation is numerically estimated based on the joint distribution of nominal I_{write} and RTS amplitude and shows good agreement with experimental data.

Acknowledgment

This work was supported by the Center for Circuit & System Solutions (C2S2) Focus Center, one of five research centers funded under the Focus Center Research Program, a Semiconductor Research Corporation program and the National Science Foundation Infrastructure Grant No. 040342. Chip fabrication was donated by STMicroelectronics. Andrew Carlson contributed to the chip design.

References

[1] N. Tega *et al.*, *VLSI Tech. Technical Digest*, pp. 50–51, 2009.
[2] N. Tega *et al.*, *IRPS Technical Digest*, pp. 541–546, 2008.
[3] Z. Guo, A. Carlson, L. T. Pang, K. Duong, T. J. King Liu, and B. Nikolic, *VLSI Circ. Technical Digest*, pp. 42–43, 2008.
[4] J. Kolhatkar *et al.*, *IEDM Technical Digest*, pp. 759–762, 2004.
[5] K. Hung, P. Ko, C. Hu, and Y. Cheng, *IEEE Electron Device Letters*, pp. 90–92, 1990.
[6] A. Carlson, Z. Guo, S. Balasubramanian, L. T. Pang, T. J. King Liu, and B. Nikolic, *IEEE SOI Conference Proceedings*, pp. 105–106, Oct. 2006.
[7] K. Fukuda, Y. Shimizu, K. Amemiya, M. Kamoshida, and C. Hu, *IEDM Technical Digest*, pp. 169–172, 2007.