Asymmetric Sizing in a 45nm 5T SRAM to Improve Read Stability over 6T

Satyanand Nalam and Benton H. Calhoun
Dept. of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA, USA
E-mail: {svn2u, bcalhoun}@virginia.edu

***Abstract*—This paper describes a 5-transistor (5T) SRAM bitcell that uses a novel asymmetric sizing approach to achieve increased read stability. Measurements of a 32 kb 5T SRAM in a 45nm bulk CMOS technology validate the design, showing read functionality below 0.5V. The 5T bitcell has lower write margin than the 6T, but measurements of the 45nm 5T array confirm that a write assist method restores comparable writability with a 6T down to 0.7 V.**

I. Introduction and Related Work

Increased variations reduce SRAM noise margins and oppose scaling of the conventional 6T SRAM bitcell to lower V_{DD} and to new processes. Depending on the process and cell design, either read static noise margin (RSNM) or write noise margin (WNM) tends to limit the lowest operational V_{DD} (V_{DDmin}). In addition, embedded memories need to meet aggressive performance requirements. Since the 6T cannot meet stability and performance requirements simultaneously, alternative bitcells, such as the 8T have been proposed [1]. For example, the Nehalem processor [2] uses the 8T SRAM in the L1 and L2 caches to achieve high performance and adequate stability. In many designs, the area overhead of the 8T leads designers to explore other options for improving 6T stability such as read and write (e.g.[3]-[6]) assists. The requirement for dense SRAM that has lower V_{DDmin} while retaining stability and performance, especially in embedded memories for scaled technologies, suggests the need for an alternative bitcell to the 6T that provides a better tradeoff between area and these other critical metrics.

Several 5T bitcells have been proposed earlier as a potential replacement for the 6T. A 5T bitcell in [7] uses a bitline (BL) biased at mid-rail and careful sizing to balance read and write margins through the access device, which becomes infeasible due to variation in modern processes. More recently, a port-less 5T bitcell controlled by a single transistor between the storage nodes was proposed in [8], which trades-off performance for increased RSNM and leakage power savings. In this paper, we propose a 5T bitcell that closely mimics 6T access methods and that also uses a novel asymmetric sizing approach to increase RSNM and to provide an effective knob to tradeoff performance, area, and variation tolerance. We make the following key contributions:

- Present a 5T bitcell with a novel asymmetric sizing approach to increase RSNM over an iso-area 6T.

This work was supported by FCRP/DARPA C2S2. The authors thank Freescale Semiconductor for chip fabrication.

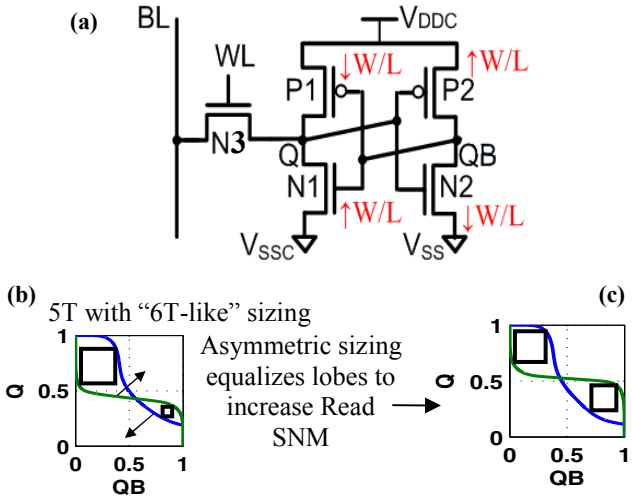

Fig. 1: (a) 5T bitcell schematic. (b) Read Static Noise Margin (RSNM) for 5T with 6T-based sizing. (c) RSNM for 5T with asymmetric sizing

- Show how asymmetric sizing can be used as a knob to achieve an efficient trade-off between read delay, variability, area and leakage.
- Demonstrate a functional 5T SRAM in a commercial 45nm bulk CMOS technology and analyze the pros and cons of a 5T relative to a 6T.
- Show measurements that confirm that the main problem with the 5T, writing a '1', can be overcome using write assist methods.
- Demonstrate the scalability of the 5T bitcell.

The rest of the paper is organized as follows. Section II describes the 5T bitcell with asymmetric sizing and discusses its advantages compared to an iso-area 6T bitcell. We also show how we can keep the same metrics as a 6T and save area instead. We then describe a write assist method that overcomes the reduced writability of the 5T cell while maintaining its other benefits. Section III presents the 45nm test chip architecture and measured results. Section IV concludes.

II. Asymmetrically Sized 5T Bitcell

A. Asymmetric sizing approach and its benefits

Fig. 1a shows the 5T bitcell schematic (a 6T bitcell missing one access transistor). Both read and write accesses occur identically to the 6T, except that they are single ended through the lone access device. Writing a '1' through the lone NMOS access transistor is difficult without using write

assist(s), which we discuss in Section II.. Fig. 1b shows the RSNM butterfly curve of a 5T bitcell obtained simply by dropping one access transistor from a conventional symmetrically sized 6T bitcell. One lobe of the curve is much smaller than the other due to the voltage-divider effect of N1 and N3. This lobe determines the RSNM of the "6T-like" 5T and is the same as the original 6T. Sizing the cross-coupled inverters in the bitcell asymmetrically skews the butterfly curve as indicated by the arrows in Fig. 1b. Fig. 1c shows the resulting increase in RSNM. This key insight provides the 5T bitcell with its beneficial features.

In order to skew the butterfly curve to achieve higher RSNM, we need to either strengthen N1 or P2, or weaken P1 or N2, or use a combination of these approaches. The missing access transistor allows us to size up one or more of these transistors until the bitcell area is the same as the original 6T. In particular, increasing the width of N1 improves RSNM, increases read current, and reduces read delay for the same bitcell area. Without loss of generality, we define the read delay as the time elapsed between WL activation and the BL voltage dropping below a certain threshold. We arbitrarily choose this as 900 mV for V_{DD}=1V. The delay will also be improved for other definitions of read time since the 5T drive transistor is larger than that of the iso-area 6T cell. Moreover, widening N1 reduces the standard deviation (σ) of its threshold voltage in the presence of local mismatch, which in turn reduces the variability in the read current.

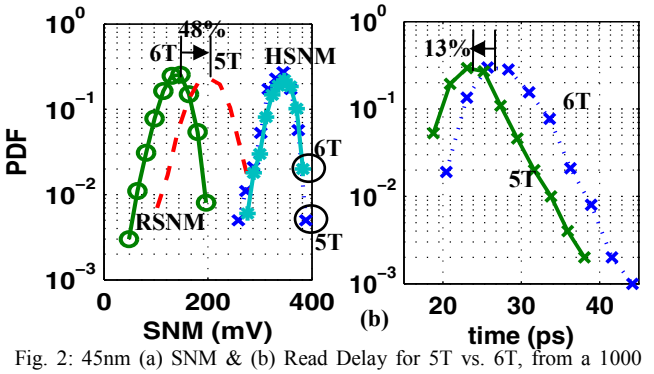

Fig. 2: 45nm (a) SNM & (b) Read Delay for 5T vs. 6T, from a 1000 point Monte Carlo Simulation at TT, 27 °C, 1 V.

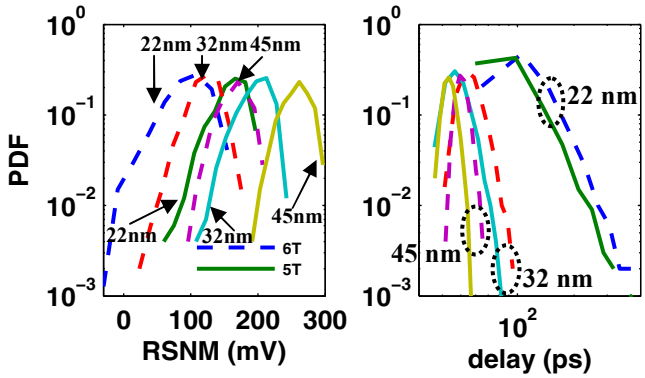

Fig. 3: Effect of scaling (using predictive models [10])on (a) RSNM and (b) read delay.

Fig. 2a shows that the mean RSNM for the asymmetric 5T with a wider N1 increases by 48% over an iso-area 6T at the 45nm node, with minimal degradation of hold SNM (HSNM). In addition, Fig. 2b shows that the asymmetric 5T bitcell reduces the mean read delay by 13.4% and the σ by 13.6%, making the 5T less sensitive to variation. Asymmetric sizing gives a similar improvement in read delay and variability for a 6T in [9]. More importantly, Fig. 3 shows that the relative improvements of the 5T over the 6T in terms of RSNM and read delay increase with scaling, with RSNM recovering nearly two process nodes at 22nm.

In general, since a transistor can be strengthened or weakened by changing either its width or length, it gives us a knob to achieve a trade-off between cell area and other metrics of interest in addition to a guaranteed increase in RSNM. The following example illustrates this idea. We start with a reference 6T bitcell in 45 nm with device sizes as shown in TABLE I. Dropping one access FET gives us additional area equivalent to roughly 100 nm of device width or 50 nm of device length, so that the resulting asymmetric 5T has the same area as the 6T. In addition, we weaken N2 by reducing its width to 100 nm, which gives us roughly an additional 50 nm of device width. TABLE I. shows five 5T bitcells that use different asymmetric sizing approaches, but have the same area as the reference 6T bitcell based on these sizing assumptions.

TABLE I. DEVICE SIZING FOR ISO-AREA 6T AND 5T BITCELLS

Bitcell	W (nm) / L (nm)				
	Pull up		*Pull Down*		*Pass Gate*
6T	100/40		150/40		100/60
	P1	*P2*	*N1*	*N2*	*N3*
5T1	100/40	100/40	300/40	100/40	100/60
5T2	100/90	100/40	200/40	100/40	100/60
5T3	70/40	100/40	330/40	100/40	100/60
5T4	100/40	250/40	150/40	100/40	100/60
5T5	100/40	100/40	200/40	100/90	100/60

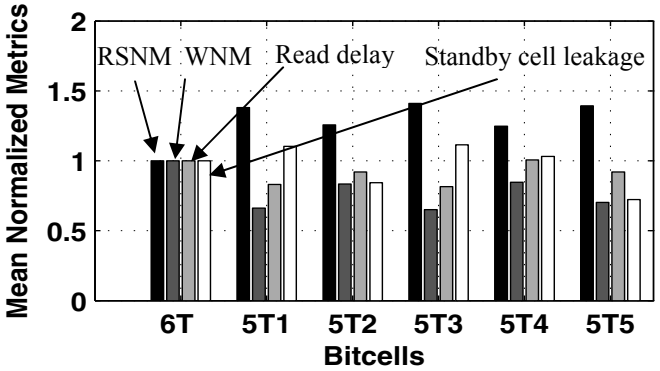

Fig. 4: Asymmetric sizing used as a knob to trade off cell area with RSNM, WNM (with write assist for 5T), read delay, and leakage. All bitcells have the same area. For 5T standby leakage, the average of the leakage for the cell storing '0' and '1' is shown.

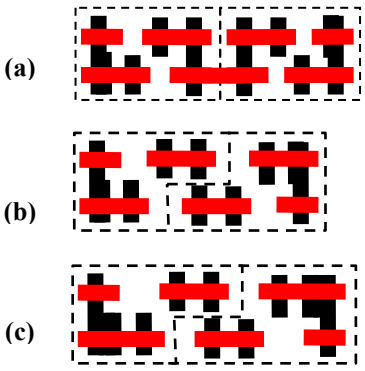

Fig. 5: Layout Options

Based on 45nm simulation results, Fig. 4 shows how the mean RSNM, read delay, Write Noise Margin (WNM), and the standby leakage of the 5T bitcell change with different use of the additional area available. As discussed earlier, sizing schemes that strengthen N1 significantly improve RSNM and lower read delay, but they also hurt the WNM (e.g. 5T1 and 5T3). In addition, the wider transistors also increase cell leakage. This is partially offset since bitline leakage is eliminated when the cell stores a '1'. Secondly, schemes that weaken P1 *and* do not strengthen N1 much reduce the WNM less, but they also reduce the improvement in delay and RSNM (e.g. 5T2). Thirdly, sizing schemes that increase transistor lengths result in significant leakage power reduction (e.g. 16% reduction in 5T2 and 27% in 5T5), which adds to the reduction in bitline leakage. Finally, all sizing schemes provide at least a 25% improvement in RSNM. An alternative approach is to keep devices sizes similar the 6T and save area by using a smaller bitcell with roughly the same stability as the 6T.

Fig. 5a-c shows layout options to exercise the tradeoffs in the 5T. At the same area as 6T (Fig. 5c), the 5T has better metrics, *or*, for the same RSNM and read delay as the 6T, the 5T can save area (Fig. 5b). Keeping logic design rules, an 11.2% bitcell area reduction relative to the 6T is possible in 45nm for the same RSNM. This number can increase if "pushed" design rules are followed.

B. Solving the write problem

The main limitation of the 5T is degraded WNM compared with a 6T of iso-size, due to difficulty writing '1' through N3. We show that by collapsing V_{DDC} ([4], [5]), we can solve this problem. As the timing waveforms in Fig. 6 show, collapsing V_{DDC} weakens the cell feedback, enabling it to flip despite the weak '1' passed by N3. As Fig. 7 shows, sufficient reduction of V_{DDC} restores the 5T WNM to near that of a 6T. Collapsing V_{DDC} reduces the HSNM of the half-selected cells (e.g. same V_{DDC}, but WL=0), but Fig. 7 shows that the HSNM remains sufficiently high even when V_{DDC} reduces enough to provide the 5T with WNM equal to the original 6T. Other write assist techniques can be used in combination with collapsing V_{DDC}. For example, as shown in Fig. 7, boosting WL ([6]) during write allows for a lower V_{DDC} drop to achieve similar WNM. This also reduces the possibility of upsetting a half-selected cell. The measurement

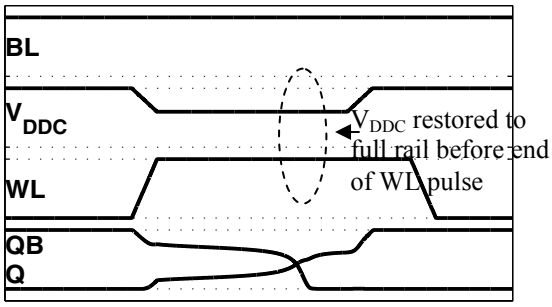

Fig. 6: Timing for V_{DDC} collapse write-assist.

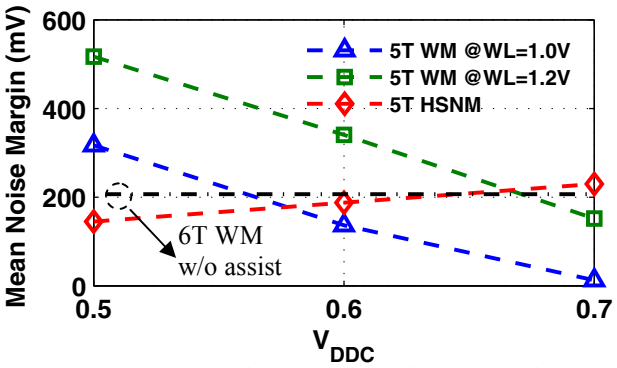

Fig. 7: Retention margin of half-selected cells limits V_{DDC} collapse. Boosting WL fixes this problem by allowing for a smaller drop in V_{DDC} for the same WNM.

results in Section III further confirm the effectiveness of these write assists.

V_{DDC} can be routed either column-wise or row-wise in the array. Routing V_{DDC} column-wise only works for processes with decent WNM since V_{DDC} cannot drop below the retention voltage of half-selected cells, although pulsing V_{DDC} maintains a dynamic margin in half-selected cells higher than static NM [11]. For row-wise routed V_{DDC}, we can either write the entire row at once or use a read-modify-write approach for the row.

III. 45NM 5T TEST CHIP MEASUREMENT RESULTS

We implemented a 45nm bulk CMOS test chip (die photo in Fig. 8) with two 16 kb 5T arrays, divided into 4 kb banks, each with a different asymmetric sizing. The banks have 128 cells per BL. The chip also had a 16 kb 6T array. Fig. 9 shows the schematic of a 4 kb 5T block on the chip. V_{DDL} and V_{DDWL} are supplied externally to simplify testing. The single-ended read uses an inverter to "sense" the BL, but other single ended sensing mechanisms could improve read speed (e.g. [12]).

Both the 5T and 6T read correctly to below 0.5 V, where pad ring issues limit further measurement. Nevertheless, this measurement confirms a robust read operation for the 5T to very low V_{DDmin}. Time constraints limited initial testing to a 4 kb bank. Measurements also verify that write assists provide good writability. Collapsing V_{DDC} as shown in Fig. 6 provides full write functionality at a nominal voltage of 1V. We measured the impact of the sizing approach on the write

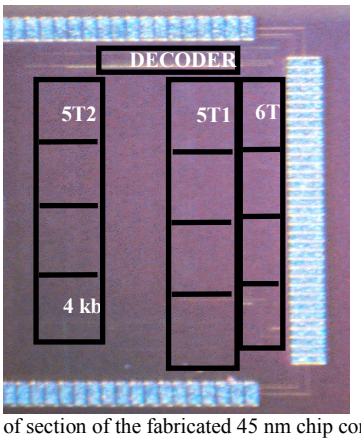

Fig. 8: Die photo of section of the fabricated 45 nm chip containing 32kb of 5T and 16kb of 6T SRAM.

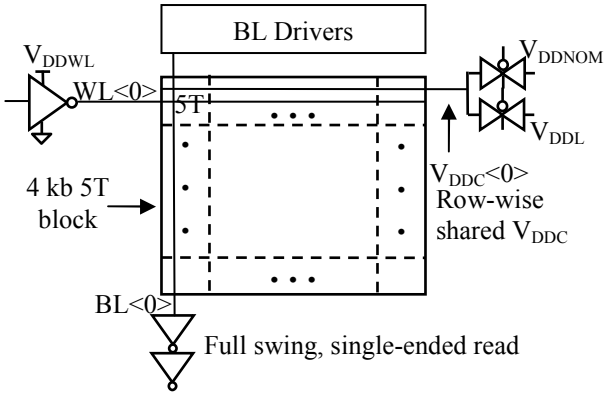

Fig. 9: Schematic of Write Assist implementations on chip

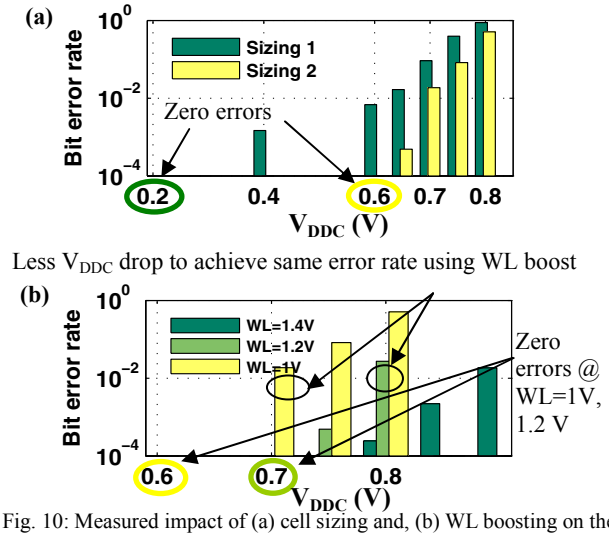

Fig. 10: Measured impact of (a) cell sizing and, (b) WL boosting on the effectiveness of the V_{DDC} reduction write-assist, on a 4 kb bank.

assist (Fig. 10a). Longer N2 and P1 (sizing 2) result in better writability (e.g. lower number of erroneous bits) than using wider N1 and P2 (sizing 1), which agrees with the conclusions drawn regarding WNM in Fig. 4. Also, using additional write assists along with V_{DDC} collapse significantly improves writability, as predicted in Fig. 7. Boosting the WL during write allows a smaller drop in V_{DDC} to ensure the same writability (Fig. 10b). This also enables the 5T to have a comparable writability to the 6T down to 0.7 V (Fig. 11).

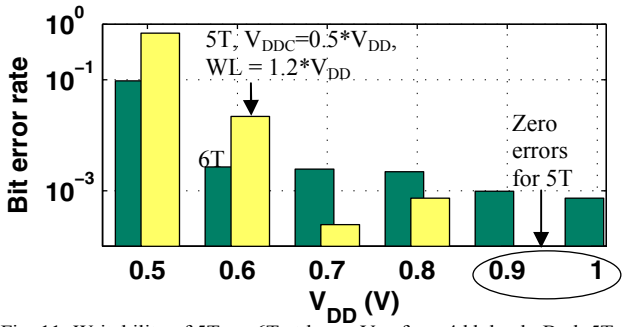

Fig. 11: Writability of 5T vs. 6T at lower V_{DD} for a 4 kb bank. Both 5T and 6T are stable for read to below 0.5 V for these banks.

IV. Conclusions

The 5T SRAM presented in this work uses asymmetric sizing to significantly improve the RSNM versus 6T. In addition, this approach provides a means to trade-off area and WNM for performance and reduced variability more efficiently than the 6T. Measurements from a 45nm chip confirm the successful functionality and read stability of the 5T SRAM, and show how writability can be improved through write assist techniques. Although the 5T has an inferior writability when compared to the 6T, its read stability and the benefits of asymmetric sizing make it an attractive choice for applications where writability is less of a concern. Finally, the benefits of the 5T relative to the 6T improve with process scaling.

References

[1] L. Chang, et al., "A 5.3GHz 8T-SRAM with Operation Down to 0.41V in 65nm CMOS," *Symposium on VLSI Circuits,* 2007.

[2] R. Kumar and G. Hinton, "A Family of IA Processors," *ISSCC,*2009

[3] A. Bhavnagarwala, et al., "A Sub-600mV, Fluctuation tolerant 65nm CMOS SRAM Array with Dynamic Cell Biasing", *Symp. VLSI Circuits,* pp.78-79, 2007.

[4] S. Ohbayashi, et al., "A 65-nm SoC Embedded 6T-SRAM Designed for Manufacturability with Read and Write Operation Stabilizing Circuits", *JSSC*, vol 42, No. 4, pp.820-829, April 2007.

[5] M. Yamaoka, et al., "Low-power embedded SRAM modules with expanded margins for writing," *ISSCC. 2005*

[6] Y. Morita, et al., "A Vth-Variation-Tolerant SRAM with 0.3-V Minimum Operation Voltage for Memory-Rich SoC Under DVS Environment," *Symposium on VLSI Circuits,* 2006

[7] I. Carlson, S. Andersson, S. Natarajan and A. Alvandpour, "A high density, low leakage, 5T SRAM for embedded caches," *ESSCIRC* , 2004

[8] Wieckowski, M.; Margala, M., "A portless SRAM Cell using stunted wordline drivers," *ISCAS*, 2008

[9] A. Kawasumi, et al., "A Single-Power-Supply 0.7V 1GHz 45nm SRAM with An Asymmetrical Unit-β-ratio Memory Cell," *ISSCC* , 2008

[10] W. Zhao, and Y. Cao, "New generation of predictive technology model for sub-45nm design exploration," *ISQED,* 2006

[11] M. Khellah, et al., "PVT-variations and supply-noise tolerant 45nm dense cache arrays with Diffusion-Notch-Free (DNF) 6T SRAM cells and dynamic multi-Vcc circuits," *Symposium on VLSI Circuits,* 2008

[12] N.Verma, A. Chandrakasan, "A High-Density 45nm SRAM using small-signal non-strobed regenerative signal", *ISSCC*, 2008.