0.5-V sub-ns open-BL SRAM array with mid-point-sensing multi-power 5T cell

Kiyoo Itoh, Khaja Ahmad Shaik, Amara Amara

0.5-V Sub-ns Open-BL SRAM Array with Mid-Point-Sensing Multi-Power 5T Cell

Kiyoo Itoh, Khaja Ahmad Shaik, and Amara Amara
Institut supérieur d'électronique de Paris (ISEP),
10 Rue Vanves, Issy-les-Moulineaux, France.
kshaik@isep.fr

***Abstract*—To achieve 0.5-V high-speed SRAMs, two proposals are demonstrated. One is a multi-power-supply five-transistor cell (5T cell), combined with a boosted word-line voltage and a mid-point sensing enabled by precharging bit-lines to $V_{DD}/2$. The other is a partial activation of a multi-divided open-bit-line array without significant area penalty. Layout and post-layout simulation with a 28-nm fully-depleted planar-logic SOI MOSFET reveal that a 5T-cell 4-kb array in a 128-kb SRAM core is able to achieve x6 faster and x14 lower power than the counterpart 6T-cell array, suggesting a possibility of a 540-ps cycle time at 0.5 V.**

Keywords—0.5-V 5T-cell SRAM array, multi-divided open bit-lines, boosted word-line, and mid-point sensing.

I. INTRODUCTION

Traditional high-speed SRAMs have evolved through ceaseless developments towards low voltage and low power, playing a key role in achieving high-performance MPUs and SoCs. Consequently, the six-transistor memory cell (6T cell) becomes most popular due to its excellent device and voltage scalability. Along with advancing in process technologies into the nanoscale regime, however, even 6T-cell SRAMs are facing at least two serious challenges that mainly comes from the cell itself and from the array architecture. One is voltage (V_{DD}) scaling of the cell to 0.5 V or less, since the scaling has rapidly degraded the voltage margin under the highest necessary threshold voltage (V_t) for small leakage current and the ever-larger V_t-variation. The other is reduction of the ever-higher array power, which dominates SRAM power, with larger capacity. Major sources of the array power are the traditional V_{DD}-bit-line (BL) precharging scheme in which many BLs swing at a large voltage of V_{DD}, and the folded-BL arrangement with inherently large BL-capacitance C_B that not only increases power dissipation but also degrades speed and read stability of the cell. As for the V_{DD} scaling, many attempts to cope with resultant narrower voltage margin have been made so far [1, 2]. They include a column-based dynamic power supply [3, 4], in which the cell power supply during write is floating or lower than that during read, a dynamic forward-body-bias just after read and write operations for fast cell-node recovery [5], a two-step word pulse [6, 7] with boosting word voltage during write with the help of a sense amplifier, and a boosted WL-voltage FinFET cell [8]. In addition, a negative BL boosting [9] speeds up write operation despite instability or additional leakage current involved. Meanwhile, single-ended 5T cells suitable for an open-BL arrangement have been proposed as a potential replacement for the 6T cell. A typical example is to boost the word-voltage only during write [10, 11], as in the 6T cell. Another is to set the BL read-voltage at an intermediate level (e.g., 0.6 V at V_{DD} = 1.8 V) to ensure the read stability [12], although a higher BL voltage is necessary for write and stable reference-voltage generations for differential sensing remain unknown. Despite such intensive developments for the 6T and 5T cells, the cell margin remains small because the cell internal voltages are equal to V_{DD}, thus eventually preventing from V_{DD} scaling. The array power also remains high. Although multi-divided BL architecture is a key to small C_B, an area penalty by a large column circuit necessary at each division is intolerable.

In this paper, two proposals are presented to accomplish the challenges. One is a multi-power-supply 5T cell, assisted by a boosted word-line voltage and a mid-point sensing enabled by precharging BLs to $V_{DD}/2$. The other is a multi-divided open-bit-line 5T-cell array that enables a partial activation of the array without significant area penalty. The cell and array are then evaluated through layout and post-layout simulation with a 28-nm fully-depleted planar logic SOI MOSFET [13]. Finally, the high-speed potential is investigated.

II. 5T CELL

A. Concept

Fig. 1 shows the proposed 5T cell featuring wide voltage margin and strong immunity to V_t variation (ΔV_t) due to increased gate-over-drive. It uses two different sets of power supplies, V_{DD} and V_{SS} (= 0 V) and boosted power supplies of V_{DH} (> V_{DD}) and V_{SL} (> 0 V). The M_{lb}-M_{db} inverter gives a bias to the M_l-M_d inverter. Either M_l or M_d is thus forward-biased for fast read operation with quickly driving the BL even at 0.5 V or less, while the other is reverse-biased for the small leakage current. For example, "0" read when N_1 node at 0 V speeds up with lower impedance of the read path composed of M_a and M_d. Write operation speeds up thanks to the boosted supplies enabling fast feedback of the flip flop. For example, for "1" write when flipping N_1-voltage from 0 V to V_{DD}, the speed improves with increasing

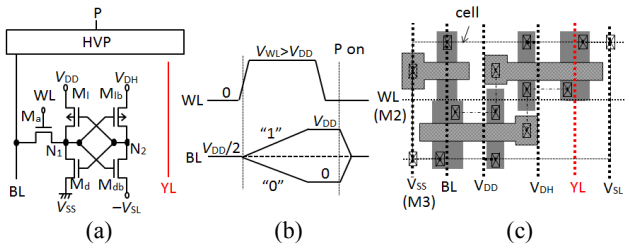


Fig. 1 Concept of proposed 5T cell. (a) Circuit, (b) timing, and (c) layout. P: precharge pulse.

the gate-over-drive of M_{db}, given as $V_r - V_{teff}(M_{db}) = V_r - V_t(M_{db}) + V_{SL}$, where V_r and $V_{teff}(M_{db})$ are the ratio voltage of M_a to M_d and the effective V_t of M_{db}, respectively. Hence, the boosted WL ($V_{WL} > V_{DD}$) enhances read and write speeds. The read instability involved in the well-known boosted WL and the conventional V_{DD}-BL-precharging is remedied by the mid-point sensing (MID-S) that discriminates "1" and "0" read signals, referring to $V_{DD}/2$, as explained later. MID-S also halves array power and provides a quiet array. In addition, the 5T cell that eliminates one set of MOSFET and BL from the 6T cell [14] reduces the cell size and affords to add a contactless column (YL) line with the same layer as BLs and power lines. The YL is a key to high-performance arrays, as explained later.

B. *Evaluation*

The cell is evaluated in the following condition. The regular V_t (RVT) defined by the saturation current is 218 mV for pMOSFET and 343 mV for nMOSFET, which are available at a silicon foundry. For the 5T cell (see Fig. 7), the RVT is used for M_a and M_l, while a V_t increased by 0.15 V from the RVT is for M_d for fast write, and another V_t increased by 0.2 V are for M_{lb} and M_{db} to suppress leakage currents caused by the boosted supplies. Here, the V_ts are changed through backplane bias although in practice, they should be fixed without the bias to minimize relevant overhead after cell-optimization. A channel width (W) of 120 nm is for M_l in the 5T cell and M_d and M_{db} in the 6T cell, while an 80-nm W is for other MOSFETs, under a fixed channel length of 30 nm. A 0.5-V V_{DD} and 1-V V_{WL} are fixed. Read speeds $t_R(1)$ for "1" and $t_R(0)$ for "0" are from WL activation to read-signal developing to 100 mV on BL with a BL-capacitance C_B of 10 fF. Write speeds $t_W(1)$ and $t_W(0)$ are from WL activation to stored-voltage flipping to $0.9V_{DD}$ in the cell.

Fig. 2 shows speeds and leakage currents of the cell. The 5T cell is rapidly faster with V_{DH} and V_{SL}, as expected, surpassing the 6T cell. Here, the 6T cell is slow, especially at V_{DD} = 0.5 V, due to non-boosted WL. In any event, for example, at V_{DH}/V_{SL}= 0.7/0.2 V, it improves $t_R(1)$ to 166 ps from 502 ps of the 6T cell, while $t_W(1)$ to 174 ps from 229 ps. The minimum necessary WL-pulse width T_{WLmin} is thus 174 ps, equal to the slowest of $t_W(1)$, $t_W(0)$, $t_R(1)$, and $t_R(0)$, for the 5T cell while 502 ps for the 6T cell. At V_{DH}/V_{SL}= 0.7/0.2 V, leakage currents per cell, $I_l(1)$ for "1" and $I_l(0)$ for "0", are about 80 pA while 54 pA for the 6T cell. The $I_l(1)$ and $I_l(0)$ are dominated by large leakage currents of the M_{lb}-M_{db} inverter caused by the boosted supplies. However, they become rapidly smaller with lowering the V_{DH}/V_{SL} due to reduced leakage current. At about V_{DH}/V_{SL} = 0.6/0.1 V, $I_l(0)$ tends to be saturated and then increases because the major component is changed from M_{db} to M_l and the leakage current of M_l, $i(M_l)$, rapidly increases due to shallower reverse-biasing. Another constant leakage-current component appearing at M_a, $i(M_a)$, is from the well-known drain-induced barrier lowering (DIBL) effect. The component is generated only when N_1 is at 0 V, and flows from the BL to N_1. It is as small as 6.5 pA for the 5T cell thanks to halved drain-source voltage (i.e., $V_{DD}/2$) of M_a while 15 pA for the 6T cell due to larger drain-source voltage (i.e., V_{DD}) of M_a. In any event, $I_l(0)$ increases at V_{DH}/V_{SL}= 0.5/0 V due to no reverse-biasing to M_l any more. On the other hand, $I_l(1)$ continues to decrease due to smaller component of $i(M_d)$ thanks to $V_t(M_d)$ larger than $V_t(M_l)$ (see Fig.7).

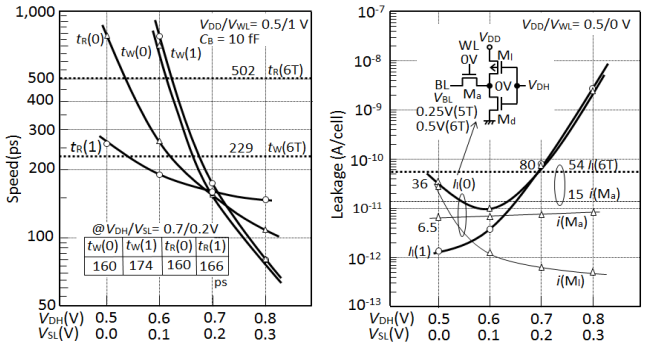
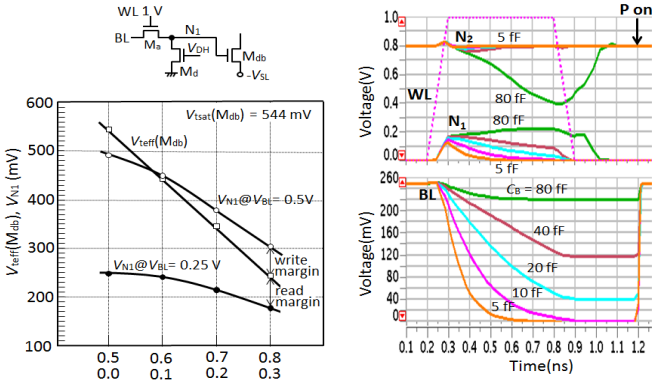


Fig. 2 Speeds and leakage current vs. V_{DH} and V_{SL}.



(a) (b)

Fig. 3 (a) DC characteristics of 5T cell and (b) waveforms at V_{DH}/V_{SL}= 0.8/0.3 V.

Fig. 3 depicts the read stability, which is closely related to successful write operation. For example, for "0" read, N_1-voltage $V_{N1}(0.25)$ when BL is at $V_{DD}/2$ must be lower than $V_{teff}(M_{db})(=V_t(M_{db})-V_{SL})$ to avoid activating the M_{db}-M_l

feedback loop. For "1" write, however, N_1-voltage $V_{N1}(0.5)$ when BL is at V_{DD} must be higher than $V_{teff}(M_{db})$ for successful write. Since the present V_{teff} is between $V_{N1}(0.25)$ and $V_{N1}(0.5)$, as seen in the figure, the cell ensures both operations over a wide range of V_{DH}/V_{SL}. In practice, $V_{N1}(0.25)$ is actually a floating voltage, which further widens the margin. In fact, even if it is initially raised by charge-sharing of BL and N_1 and/or by ratio of M_a to M_d, it never activates the feedback loop even at V_{DH}/V_{SL} = 0.8/03 V and C_B = 80 fF, as seen in Fig. 3(b). Note that the conventional V_{DD}-BL-precharging causes a read instability of "0" destruction due to the high $V_{N1}(0.5)$, calling for non-boosted WL to reduce the $V_{N1}(0.5)$.

III. OPEN-BL ARRAY

A. Concept

Fig. 4 depicts an open-BL (O-BL) 4-kb array in a 128-kb SRAM core. The YL line is utilized as column selection, so the array can be divided into four pairs (PMA_0-PMA_3) of sub-arrays in the BL direction and a large column decoder and driver block (YDEC&DRV) can be shared with the PMAs. Hence, there is no substantial area increase even if an array is multi-divided for smaller C_B. To activate only one PMA of the four, each PMA adopts a hierarchical I/O of local and global pairs, LIOs and GIOs, through M_y-M_{yb} switches YSWs controlled by YLs, and sub-array switches (SWs). Hence, the array architecture enables a partial activation of the array while preserving a small array, thus reducing active power of the array to one-fourth. Moreover, it speeds up read operation with reduced capacitance of read path (BL-GIO), as explained later. Here, it is assumed that V_{DH}, V_{SL}, and $V_{DD}/2$ for cells, and V_{PP} (= V_{WL}= V_{YL}) for periphery are from on-chip voltage converters with V_{DD} and V_{SS} (= 0 V) [15], which are well established in DRAMs and the area penalty is not significant for 128-kb SRAMs. Even if peripheral circuits operate at 1 V, they don't contribute significantly to the total SRAM power due to a limited number of simultaneously activated circuits. In the figure, a $V_{DD}/2$ precharger (HVP) and a sense amplifier (SA), composed of four MOSFETs of small size, are placed on each pair of BLs, while V_{DD}-prechargers are on LIO and GIO pairs for faster signal transfer without involving slow pMOSFETs. When the cell read signal reaches point S_1 (i.e., SA offset voltage of 100 mV), amplification starts, as seen in Fig. 5. To do so, a dummy cell (DMC) generates a timing pulse, SA-enable SE, for driving SA driver (SAD) and then SPL and SNL from $V_{DD}/2$ to V_{DD} and to 0 V, respectively. When the amplified signal reaches point S_2 (≈ 400 mV), the selected YL is activated with SE, so the signal is outputted to LIOs and then to GIOs by turning on the relevant switches. When the signal on GIOs reaches 100 mV, it is amplified by a read write circuit RWC after isolating RWC from GIOs. At the beginning of precharge period, each pair of BLs is precharged and equalized, so the reference voltage V_{ref} (= $V_{DD}/2$) is automatically generated by the charge sharing. No need for any current from the external supply, allows the on-chip $V_{DD}/2$-generator (HVG [15]) design to be easier.

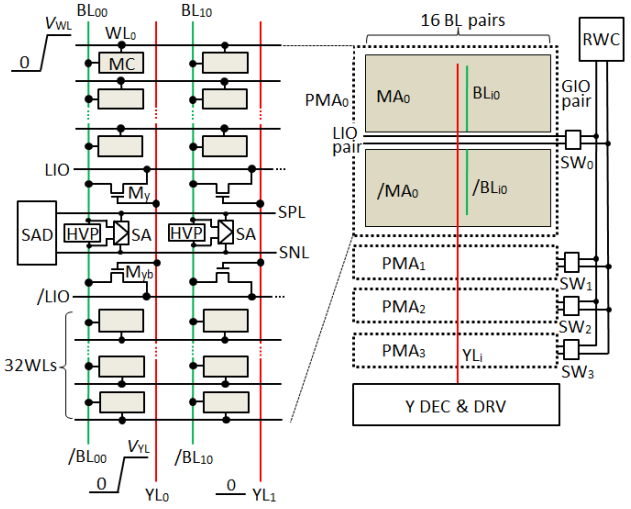


Fig. 4 Concept of the proposed open-BL 5T-cell array.

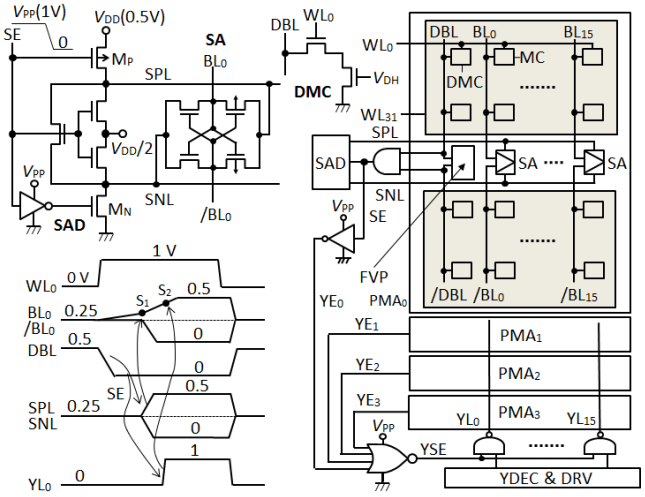
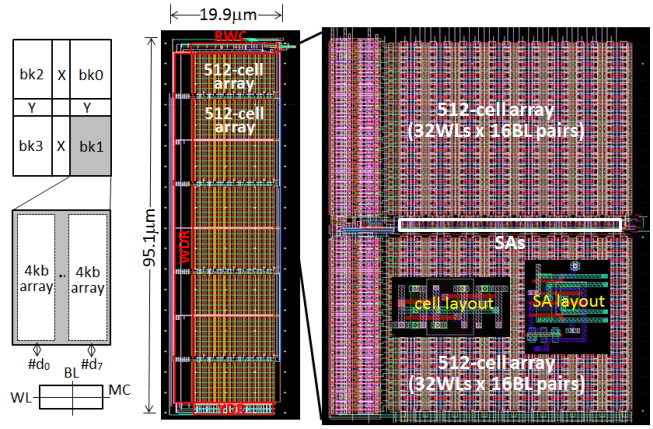


Fig. 5 Peripheral circuits and timing diagram of the array.

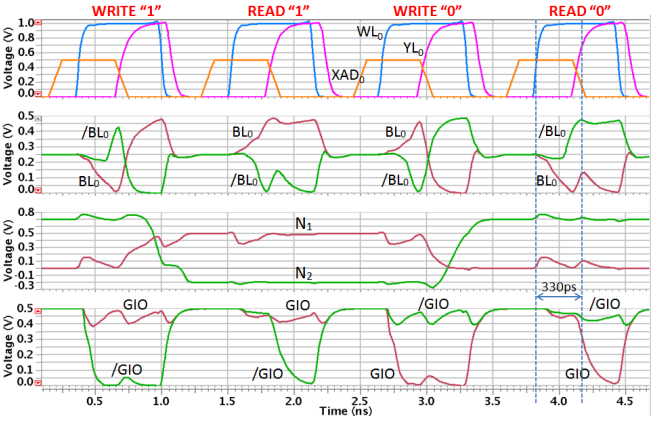
B. Evaluation

Fig. 6 illustrates layout and waveforms by post-layout simulation. The 5T cell measures 306 by 861 nm^2 (logic design rule) and the SA increases the height of the cell by about 5%. The total capacitance $C_T(R)$ of read path (BL-GIO) is as small as 12 fF (see Fig. 7) due to no cell-connection to GIOs whereas 25 fF for the 6T-cell array. Read waveform shows an array speed (WL to GIOs) of 330 ps with a delay from row addresses to WL of 180 ps, as seen in Fig. 6. The array reduces the active power of the 6T-cell array to 1/14, as seen in Fig. 7, if all BLs of half-selected cells along the selected WL in the 6T-cell array are discharged to 0 V in the worst-case V_t-variation. The power reduction is due to reducing the total capacitance $C_T(P)$ contributing to array active power to one-seventh thanks to

the partial activation of the array and halving the voltage swing on BLs, ΔV(BL).



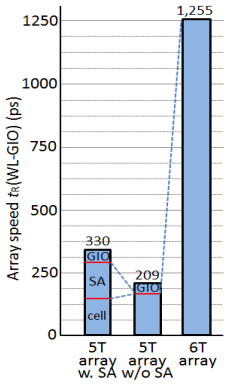
(a)



(b)

Fig. 6 (a) Layout and (b) simulated waveforms of 4-kb array in 128-kb SRAM core with 4 banks (bk0-bk3).

The O-BL 5T-cell array is potentially much faster than the above-simulated speed. This is because the high-speed potential is degraded by sophisticated timing requirements of the SA for properly activating the dummy cell, SE, SPL and SNL, and YL. If the SA is eliminated with precharging LIOs and GIOs to $V_{DD}/2$, WL, YL, and other switches are simultaneously activated. A read signal thus goes directly to the RWC. In this case, the 5T-cell array speed, simply calculated from the cell speed at C_B = 10 fF, is thus about 209 ps while the 6T-cell array speed is as slow as 1,255 ps (i.e., x6 slower), as seen in Fig. 7. Here, no BL-division is assumed for the 6T-cell array because the YL is not available. If divided with the same process, the array is intolerably large. Consequently, even a 540-ps cycle time with a row-address-WL delay of 180 ps and a precharge time of 150 ps is feasible, if the address-GIO path and the RWC succeeding path are pipelined with a switch at the input of RWC. The array active power is also reduced, as with SAs. Fortunately, even without SAs, the partial activation of the O-BL array allows the maximum V_{ref} variation after precharging to be $\pm V_{DD}/16$ when all the cells along the selected WL store "1" or "0". Such a small fluctuation can be compensated for with the conventional HVG.



(a)

4-kb Array Architecture		5T-cell array w. SA	5T-cell array w/o SA	6T-cell array
Cell	V_t(Ma/Mab)	343 mV	343 mV	343/343 mV
	V_t(Ml/Md)	218/491	218/491	218/343
	V_t(Mlb/Mdb)	420/544	420/544	218/343
	size	0.95	0.95	1.0
	t_R@C_B=10fF	174 ps	174 ps	502 ps
	leakage/cell	80 pA	80 pA	54 pA
Array	area incr.	+5%	0	0
	C_T(R)	12 fF	12 fF	25 fF
	t_R@C_T(R)	330 ps	209 ps	1,255 ps
	C_T(P)	57 fF	57 fF	400 fF
	ΔV(BL)	0.25 V	0.25 V	0.5 V(worst)
	P_{act} (ratio)	1/14	1/14	1(worst)

(b)

Fig. 7 (a) Expected speed improvement and (b) comparisons with a counterpart 6-T cell array. V_{DD} = 0.5 V, V_{WL} = 1.0 V, V_{DH} = 0.7V, and V_{SL}= 0.2 V. t_R; WL-GIO delay. ΔV(BL); voltage swing on BLs.

IV. CONCLUSION

To achieve 0.5-V high-performance SRAMs, two proposals were demonstrated. One is a multi-power-supply 5T cell, combined with boosted WL voltage and mid-point sensing. The other is a partial activation of a multi-divided open-BL array without significant area penalty. Layout and post-layout simulation with a 28-nm fully-depleted planar-logic-SOI MOSFET revealed that a 5T-cell 4-kb array in a 128-kb SRAM core is able to achieve x6 faster and x14 lower power than the counterpart 6T-cell array, suggesting a possibility of a 540-ps cycle time at 0.5 V.

ACKNOWLEDGEMENT

The authors thank Circuit Multi Projects (CMP), Grenoble, France for providing the FD-SOI 28nm design kit (DK), and STMicroelectronics, Crolles, France for the DK support.

REFERENCES

[1] K. Itoh et al., IEEE J. SSC, vol. 43, no. 4, pp.762-768, April 2008.
[2] K. Itoh, IEEE Solid-State Circuits Magazine, pp. 27-39, Winter 2013.
[3] K. Zhang et al., ISSCC Dig., pp. 474-475, Feb. 2005.
[4] M. Yamaoka et al., ISSCC Dig., pp. 480-481, Feb. 2005.
[5] F. Hamzaoglu et al., ISSCC Dig. pp. 376-377, Feb. 2008.
[6] K. Takeda et al., VLSI Circuits Symp. Dig., pp. 101-102, June 2010.
[7] M. E. Sinangil et al., ISSCC Dig., pp. 260-261, Feb. 2011.
[8] A. Kotabe et al., CICC Dig., 16-1, Sept. 2011.
[9] Y. Fujimura et al., ISSCC Dig., pp. 348-349, Feb. 2010.
[10] M. Ukita et al., ISSCC Dig., pp. 252-253, Feb. 1993.
[11] S. Nalam et al., CICC Dig., pp. 709-712, 2009.
[12] I. Carlson et al., ESSCIRC Dig., pp. 215-218, 2004.
[13] P. Flatresse et al., ISSCC Dig., pp. 424-425, Feb. 2013.
[14] K. Osada et al., IEEE J. SSC, 36(11), pp. 1738-1744, 2001.
[15] K. Itoh, M. Horiguchi, and H. Tanaka, *Ultra-low Voltage Nano-scale Memories*, Springer, pp. 277-279, 2007.