A Sub-V_{T} 2T Gain-Cell Memory for Biomedical Applications

Pascal Meinerzhagen*, Adam Teman†, Anatoli Mordakhay†, Andreas Burg*, and Alexander Fish†
*Institute of Electrical Engineering, EPFL, Lausanne, VD, 1015 Switzerland
Email: pascal.meinerzhagen@epfl.ch, andreas.burg@epfl.ch
†VLSI Systems Center, Ben-Gurion University of the Negev, Be'er Sheva, Israel
Email: aditeman@gmail.com, anatoli16@gmail.com, alexander.fish@gmail.com

***Abstract*—Biomedical systems often require several kb of embedded memory and are typically operated in the subthreshold (sub-V_{T}) domain for good energy-efficiency. Embedded memories and their leakage current can easily dominate the overall silicon area and the total power consumption, respectively. Gain-cell based embedded DRAM arrays provide a high-density, low-leakage alternative to SRAM for such systems; however, they are typically designed for operation at nominal or only slightly scaled supply voltages. For the first time, this paper presents a gain-cell array which is fully functional in the sub-V_{T} regime and achieves a data retention time that is more than 10^4 times higher than the access time. Monte Carlos simulations show that the 2 kb gain-cell array, implemented in a mature 0.18 µm CMOS node and supplied with a sub-V_{T} voltage of 400 mV, exhibits robust write and read operations at 500 kHz under parametric variations and has over 99% availibilty for read and write access.**

I. Introduction

Biomedical sensor nodes and implants are expected to run on a single cubic-millimeter battery charge for days or even for years, and therefore are required to operate with extremely low power budgets. Aggressive supply voltage scaling, leading to subthreshold (sub-V_{T}) circuit operation, is widely used in this context to lower both active energy dissipation and leakage power consumption; albeit, at the price of severely degraded on/off current ratios ($I_{\mathrm{on}}/I_{\mathrm{off}}$) and increased sensitivity to process variations [1]. The majority of these biomedical systems require a considerable amount of embedded memory for data and instruction storage, often amounting to a dominant share of the overall silicon area and power. Typical storage capacity requirements range from several kb for low-complexity systems [2] to several tens of kb for more sophisticated systems [3]. Over the last decade, robust, low-leakage, low-power sub-V_{T} memories have been heavily researched [4–6]. In order to guarantee reliable operation in the sub-V_{T} domain, many new SRAM bitcells consisting of 8–14 transistors have been proposed. All these state-of-the-art sub-V_{T} memories are based on static bitcells, while the advantages and drawbacks of dynamic bitcells operated in the sub-V_{T} regime have not yet been studied.

Gain-cells are a promising alternative to SRAM and to conventional 1-transistor-1-capacitor eDRAM (incompatible with standard digital CMOS technologies), as they are both smaller than any SRAM bitcell, as well as fully logic-compatible. Much of the previous work on gain-cell eDRAMs focuses on high-speed operation, targeting on-die caches in processors [7,8], while only a few publications deal with the design of low-power near-V_{T} gain-cell arrays [9–11]. However, the possibility of operating gain-cell arrays in the sub-V_{T} regime for high-density, low-leakage, and voltage-compatible data storage in sub-V_{T} biomedical systems has not been exploited yet. Reasonably, one of the main objections to sub-V_{T} gain-cells are the degraded $I_{\mathrm{on}}/I_{\mathrm{off}}$ current ratios, leading to rather short data retention times compared to the achievable data access times. However, the present study shows that these current ratios are still high enough in the sub-V_{T} regime to achieve short access and refresh cycles and high memory availability. While gain-cells are considerably smaller than robust sub-V_{T} SRAM bitcells, they also exhibit lower leakage currents, especially in mature CMOS nodes where sub-V_{T} conduction is the dominant leakage mechanism. Recent studies show that gain-cell arrays can even have lower retention power (leakage power plus refresh power) than SRAM (leakage power only) [12]. Moreover, compared to SRAM, gain-cells are naturally suitable for two-port memory implementation, which gives an advantage in terms of memory bandwidth and allows for simultaneous and independent optimization of the write-ability and read-ability.

The presented sub-V_{T} gain-cell eDRAM is designed for a mature 0.18 µm CMOS node which is typically used to 1) easily fulfill the high reliability requirements of biomedical sensor nodes and implants; 2) reach the highest energy-efficiency of such biomedical systems typically requiring low frequencies and duty cycles [13]; and 3) achieve low manufacturing cost.

II. 2T sub-V_{T} Gain-Cell Design

Previously reported gain-cell cell topologies include either two or three transistors and an optional MOSCAP or diode [14]. While the basic two-transistor (2T) bitcell has the smallest area cost, it limits the number of cells which can connect to the same read bitline (RBL) due to leakage currents from unselected cells masking the sense current [14]. However, as typical biomedical sensor nodes require only small memory arrays with relatively few cells per RBL, we consider the implementation of a sub-V_{T} 2T bitcell as a viable option. Both the write transistor (MW) and the combind storage and read transistor (MR) of the 2T gain-cell can be implemented with either an NMOS or a PMOS device, as shown in Fig. 1(a)-(d). Moreover, both MW and MR can be implemented with standard-V_{T} core or high-V_{T} I/O devices in the considered CMOS technology. Due to V_{T} drop across MW, a boosted write wordline (WWL) voltage is required during write access;

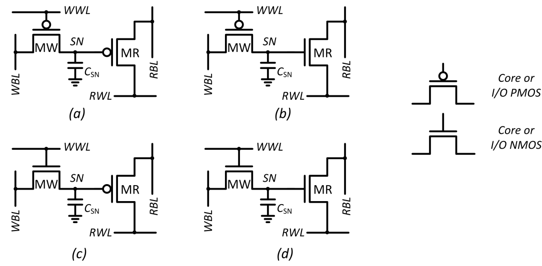


Fig. 1. Overview of two-transistor gain cell implementations.

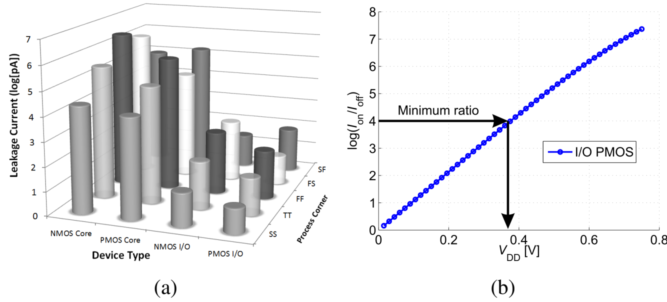


Fig. 2. (a) Leakage current of various transistor types, and (b) I/O PMOS $I_{\rm on}/I_{\rm off}$ current ratio as a function of $V_{\rm DD}$.

above $V_{\rm DD}$ for the NMOS option and below $V_{\rm SS}$ for the PMOS option. For a read operation, a PMOS MR requires a pre-discharge of the parasitic RBL capacitance followed by raising the read wordline (RWL). If the selected bitcell's storage node (SN) holds a '0', MR is conducting and charges RBL past a detectable sensing threshold. If SN holds a '1', MR is cut off, such that RBL remains discharged below the sensing threshold. For the NMOS implementation of MR, the operation is exactly opposite, *i.e.*, RBL is precharged and RWL is lowered to initiate a read.

A. Best-Practice Write Transistor Implementation

For the considered sub-$V_{\rm T}$ target applications, long retention times that minimize the number of power-consuming refresh cycles are of much higher precedence than fast write access. In the chosen 0.18 µm CMOS process, sub-$V_{\rm T}$ conduction of MW is the dominant leakage mechanism that causes the destruction of stored data levels. Fig. 2(a) shows that the I/O PMOS device has the lowest leakgage current $I_{\rm off}$ ($V_{\rm GS} = 0$ V, $V_{\rm SD} = V_{\rm DD}$) among all device options and across all standard process corners, leading to the longest retention time. At a sub-$V_{\rm T}$ $V_{\rm DD}$, as low as 400 mV, the on-current $I_{\rm on}$ ($V_{\rm GS} = -V_{\rm DD}$, $V_{\rm SD} = V_{\rm DD}$) of this preferred I/O PMOS device is still four orders of magnitude larger than $I_{\rm off}$, as shown in Fig. 2(b), which results in sufficiently fast write and refresh operations compared to the achievable retention time. As the silicon area of the 2T bitcell is dominated by contacts, the area penalty due to an I/O transistor is small.

With the chosen PMOS I/O write transistor, the worst-case retention time, corresponding to a write bitline (WBL) voltage that is constantly opposite to the stored data level during idle, is estimated at 40 ms, as illustrated in Fig. 3(a). Moreover, a logic '0' level decays much faster than a logic '1' level, corresponding with previous reports for the above-

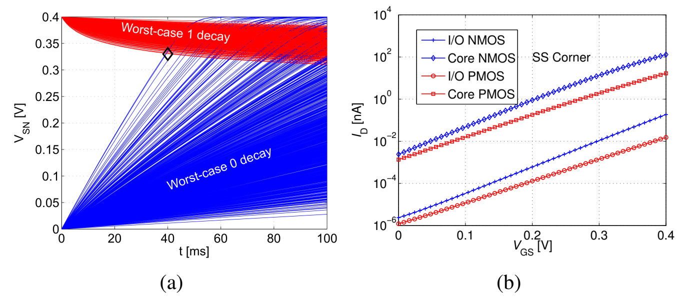


Fig. 3. (a) Retention time estimation through worst-case decay of '1' and '0' data levels. (b) Drain current ($I_{\rm D}$) of available devices as a function of the gate-to-source voltage ($V_{\rm GS}$).

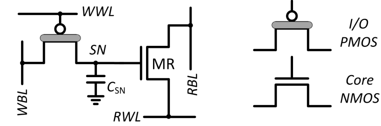


Fig. 4. The most convenient gain cell for sub-$V_{\rm T}$ operation consists of a I/O PMOS write transistor and a core NMOS read transistor.

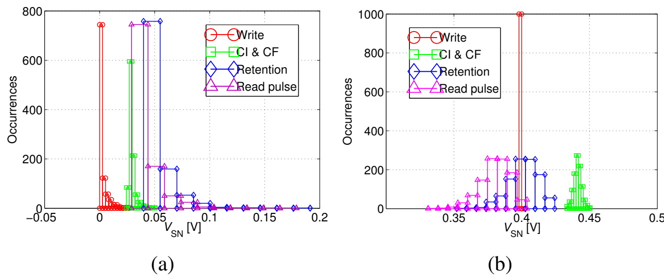
$V_{\rm T}$ domain [7,9]. In fact, the decay of a '1' level is self-limited due to the steady increase of the reverse gate overdrive and body effect of MW with progressing decay. Both of these effects suppress the device's leakage. Furthermore, the charge injection and clock feedthrough that occur at the end of a write access (when MW is turned off), cause the SN voltage level to rise, strengthening a '1' and weakening a '0' level.

B. Best-Practice Read Transistor Implementation

At the onset of a read operation, capacitive coupling from RWL to SN causes an additional voltage step on SN. Therefore, it is preferable to implement MR with an NMOS transistor that employs a negative RWL transition for read assertion. The resulting decrease in voltage on SN counteracts the previous effects, thus improving the '0' state during a read operation. Fig. 3(b) shows that NMOS devices are significantly stronger than their PMOS counterparts and that core transistors are over two orders of magnitude stronger than I/O transistors. This reinforces the choice of a core NMOS device as MR to also achieve fast read access even with a minimum-sized device. The resulting NMOS/PMOS gain-cell shown in Fig. 4 shares the n-well on three sides between neighboring cells [14] to keep the area cost low. The storage node capacitance is extended from 0.5 fF (primarily diffusion and gate capacitance) to 2.5 fF by metal stacking. Finally, the WWL underdrive voltage is carefully selected to be −650 mV for proper level transfer at minimum storage-node voltage disturb during WWL de-assertion.

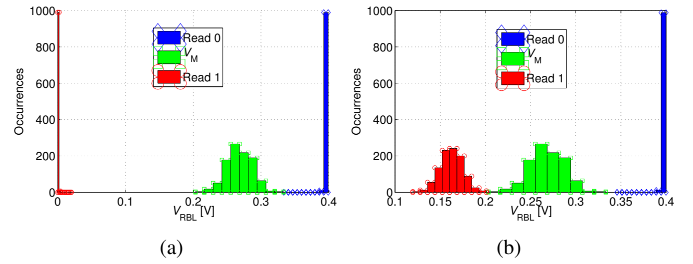
III. Macrocell Implementation Results

This section presents a 64×32 bit (2 kb) memory macro based on the previously elaborated 2T gain-cell configuration (Fig. 4), implemented in a bulk CMOS 0.18 µm technology. The considered $V_{\rm DD}$ of 400 mV is clearly in the sub-$V_{\rm T}$ regime, as $V_{\rm T}$ of MW and MR are −720 and 430 mV, respectively. Special emphasis is put on the analysis of the reliability of sub-$V_{\rm T}$ operation under parametric variations.



(a) (b)

Fig. 5. Distribution of the SN voltage of a logic ‘0’ (a) and a logic ‘1’ (b) at critical time points: 1) [circles] directly after a 1 µs write access (before turning off MW); 2) [squares] after turning off MW; 3) [diamonds] after a 40 ms retention period under worst-case WBL conditions; and 4) [triangles] during a read operation.



(a) (b)

Fig. 6. Distribution of RBL voltage (V_{RBL}) after read ‘1’ [circles] and read ‘0’ [diamonds] operations and distribution of the trip-point V_{M} of the read buffer [squares], for favorable (a) and unfavorable (b) read ‘1’ conditions.

While the address decoders and the sense buffers are built from combinational CMOS gates and operate reliably in the sub-V_{T} domain [15], the analysis focuses on the write-ability, data retention, and read-ability of the gain-cell. All simulations assume a 1 µs pulse width for both WWL and RWL; a target retention time of 40 ms; a temperature of 37 °C typically found in biomedical implants; and account for global and local parametric variations (1k-point Monte Carlo sampling).

Fig. 5(a) and (b) plot the distribution of the bitcell's SN voltage at critical time points for the ‘0’ and the ‘1’ states, respectively. As expected, nominal 0 V and 400 mV levels are passed to SN just before the positive edge of the write pulse. Charge injection and clock feedthrough cause the internal levels to rise by 20–50 mV, resulting in a slightly degraded ‘0’ level and an enhanced ‘1’ level, while the distributions remain sharp. After a 40 ms retention period with worst-case opposite WBL voltage, the distributions are spread out, but the ‘1’ levels are still strong, while the extreme cases of the ‘0’ levels have severely depleted, approaching 200 mV. However, the ‘0’ levels are improved again following the falling RWL transition, resulting in a 10–20 mV decrease.

To verify the read-ability of the bitcell, Fig. 6 shows the distribution of the RBL voltage (V_{RBL}) following read ‘0’ and read ‘1’ operations after the 40 ms retention period. In addition, the figure plots the distribution of the trip-point V_{M} of the sense buffer. While read ‘0’ is robust in any case (RBL stays precharged), read ‘1’ is most robust if all unselected cells on the same RBL as the selected cell store ‘0’ (see Fig. 6(a)), while it becomes more critical if all unselected cells store ‘1’ (see Fig. 6(b)), thereby inhibiting the discharge of RBL through the selected cell. However, the V_{RBL} distributions for read ‘0’ and read ‘1’ are still clearly separated, and the distribution of V_{M} is shown to comfortably fit between them.

As shown before, 1 µs write and read pulses are sufficiently long for array access. Assuming an additional 1 µs latency of peripherals, a full refresh cycle of 64 rows takes approximately 256 µs. With a worst-case 40 ms retention time, the resulting availability for write and read is over 99%.

IV. Conclusions

This paper proposes a two-transistor sub-V_{T} gain-cell memory for use in ultra-low-power biomedical systems. The main design goals of the bitcell are long retention time and high data integrity. A low-leakage I/O PMOS write transistor and an extended storage node capacitance ensure a retention time of at least 40 ms. At low voltages, data integrity is severely threatened by charge injection and capacitive coupling from read and write wordlines. Therefore, the positive storage-node voltage disturb at the culmination of a write operation is counteracted by a negative disturb at the onset of a read operation, which is only possible with an NMOS read transistor. Monte Carlo simulations of an entire 2 kb memory array operated at 500 kHz with a 400 mV sub-V_{T} supply voltage confirm robust write and read operations under global and local variations, as well as a minimum retention time of 40 ms leading to over 99% availability for read and write.

Acknowledgment

This work was kindly supported by the Swiss National Science Foundation under the project number PP002-119057.

References

[1] M. Sinangil *et al.*, “A reconfigurable 65nm SRAM achieving voltage scalability from 0.25-1.2V and performance scalability from 20kHz-200MHz,” in *Proc. IEEE ESSCIRC*, 2008.

[2] S. Hanson *et al.*, “A low-voltage processor for sensing applications with picowatt standby mode,” *IEEE JSSC*, April 2009.

[3] J. Constantin *et al.*, “An ultra-low-power application-specific processor for compressed sensing,” in *Proc. IFIP/IEEE VLSI-SoC*, 2012.

[4] B. H. Calhoun and A. P. Chandrakasan, “A 256-kb 65-nm sub-threshold SRAM design for ultra-low-voltage operation,” *IEEE JSSC*, 2007.

[5] A. Teman *et al.*, “A 250mV 8kb 40nm ultra-low power 9T supply feedback SRAM (SF-SRAM),” *IEEE JSSC*, 2011.

[6] P. Meinerzhagen *et al.*, “A 500fW/bit 14fJ/bit-access 4kb standard-cell based sub-Vt memory in 65nm CMOS,” in *Proc. IEEE ESSCIRC*, 2012.

[7] K. C. Chun *et al.*, “A 3T gain cell embedded DRAM utilizing preferential boosting for high density and low power on-die caches,” *IEEE JSSC*, 2011.

[8] D. Somasekhar *et al.*, “2GHz 2Mb 2T gain-cell memory macro with 128GB/s bandwidth in a 65nm logic process,” in *Proc. IEEE ISSCC*, 2008.

[9] Y. Lee *et al.*, “A 5.4nW/kB retention power logic-compatible embedded DRAM with 2T dual-Vt gain cell for low power sensing applicaions,” in *Proc. IEEE A-SSCC*, 2010.

[10] K. C. Chun *et al.*, “Logic-compatible embedded DRAM design for memory intensive low power systems,” in *Proc. IEEE ISCAS*, 2010.

[11] R. Iqbal *et al.*, “Two-port low-power gain-cell storage array: voltage scaling and retention time,” in *Proc. IEEE ISCAS*, 2012.

[12] K. C. Chun *et al.*, “A 667 MHz logic-compatible embedded DRAM featuring an asymmetric 2T gain cell for high speed on-die caches,” *IEEE JSSC*, 2012.

[13] M. Seok *et al.*, “Optimal technology selection for minimizing energy and variability in low voltage applications,” in *Proc. ACM/IEEE ISLPED*, 2008.

[14] P. Meinerzhagen *et al.*, “Design and failure analysis of logic-compatible multilevel gain-cell-based DRAM for fault-tolerant VLSI systems,” in *Proc. IEEE GLSVLSI*, 2011.

[15] B. H. Calhoun *et al.*, “Modeling and sizing for minimum energy operation in subthreshold circuits,” *IEEE JSSC*, 2005.