A 125mV 2ns-Access-Time 16Kb SRAM Design based on a 6T Hybrid TFET-FinFET Cell

Hassan Afzali-Kusha, Alireza Shafaei, and Massoud Pedram
Department of Electrical Engineering, University of Southern California, USA
Email: {afzaliku, shafaeib, pedram}@usc.edu

***Abstract*—This paper proposes a robust and energy-efficient hybrid TFET-FinFET 6T SRAM cell which takes advantage of the higher ON/OFF current ratio of TFETs compared to that of FinFETs to reliably hold and access data at ultra-low supply voltages. More precisely, in the proposed hybrid cell, to achieve low static currents along with high noise margins, TFETs are used for cross-coupled inverters, and to speed up the access time, high-performance FinFETs are utilized for access transistors. The paper also presents a dual-V_t 6T SRAM, in which low-power (high-V_t) and high-performance (low-V_t) FinFETs are used for cross-coupled inverters and access transistors, respectively. For both SRAM cells, the V_{dd} boost read-assist technique is employed to improve the read stability. Characteristics of both SRAMs are analyzed using HSPICE simulations for technologies with the gate length of 20 nm for a 128×128 SRAM array. Simulation results reveal that the lowest operating V_{dd} for the dual-V_t cell is 225 mV, whereas that of the hybrid cell is 125 mV. Moreover, to further decrease the access delay of the hybrid cell for 125 mV $\leq V_{dd} <$ 225 mV, negative Gnd read-assist technique and a boosted voltage for the row decoder are used. Finally, the paper presents a 125mV 2ns-access-time 16Kb SRAM array based on the proposed hybrid TFET-FinFET SRAM cell.**

I. Introduction

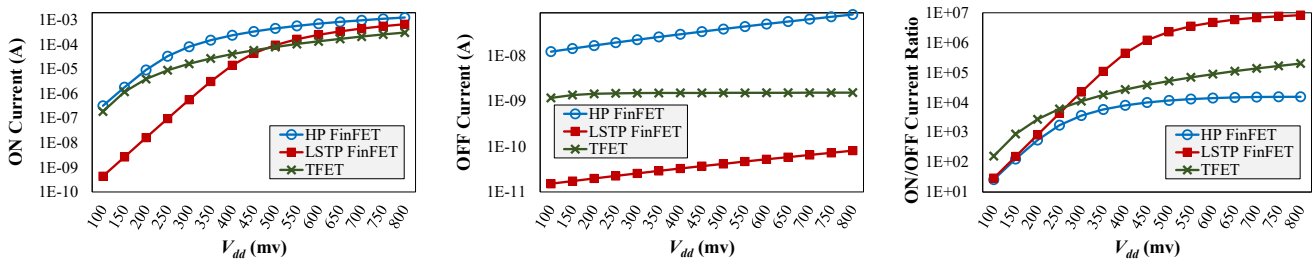
Reducing the power consumption has been one of the key objectives in the design of electronic systems. The demand for the power reduction has gained an ever-increasing importance due to the growth in the battery-powered mobile and wearable devices. In addition, decreasing the cost of these devices, or in general, system-on-chips (SoCs), which may be achieved by reducing the area of the fabricated circuit, is also an important design objective. A large fraction of area of such systems are consumed by large SRAM caches [1]. Therefore, keeping the density of SRAM cells as high as possible along with lowering their power consumption are highly desired. Furthermore, both static and dynamic power consumptions are greatly affected by the changes in the sizes of SRAM transistors.

An efficient approach to reduce the power consumption of digital circuits is supply voltage, V_{dd}, scaling. However, the minimum operating voltage of high-density SRAM bitcells implemented with minimum-sized transistors may be higher than that of logic circuits on the chip (due to the minimum required noise margin levels for SRAM cells). To be able to take advantage of the maximum power reduction for all circuits on the chip, a separate (minimum) operating voltage for on-die memory arrays [2] or larger-sized memory cells [3] may be utilized. These solutions come at the price of a higher system cost due to the use of an additional supply voltage or larger memory (die) area. On the other hand, using larger transistors or adopting a higher V_{dd} for SRAM cells lead to higher static (leakage) and dynamic (active) power consumptions, which are higher than when minimum-sized transistors under the minimum V_{dd} of logic could have been used for SRAMs [4].

In general, power reduction techniques, including V_{dd} scaling, adversely affect the speed of the circuit. More specifically, achieving a high-performance (fast) operation at low operating voltages is a challenging design task because of lower ON currents. In addition, lowering the supply voltage limits the read and write margins (stability parameters) of the standard 6T SRAM cell [5]. In fact, the conflicting requirements on access transistors for read and write operations in column-interleaved memories mandate a compromise between the two margins. Inevitably, the compromise limits lowering the minimum cell operating voltage or parametric cell yield [5]. To continue scaling the density and operating voltage of the cell, new materials, innovative transistor architecture enhancements, and/or memory assist circuit co-design have been considered (e.g., see [6] and [7]). These approaches help dealing with the aforesaid intrinsic limitations of the 6T bitcell structure.

The purpose of this paper is to present a robust memory cell for ultra-low voltage applications. To this end, we take advantage of Tunnel Field Effect Transistors (TFETs), known as steep switching devices, which are considered as one of the proper replacements of conventional MOSFETs for low power applications [8] [9]. This is because these transistors are not constrained by the subthreshold swing limit of 60 mV/decade for conventional MOSFETs. Further, TFET devices offer very high ON/OFF current ratios at ultra-low V_{dd} levels, which is an important feature for an SRAM cell to have high noise margins. Accordingly, we present a hybrid TFET-FinFET 6T SRAM cell, in which TFET devices are used for cross-coupled inverters to achieve low leakage power along with high noise margins, and high-performance FinFETs are utilized for access transistors to speed up the access time. It should be noted that the overall fabrication processes for TFETs is compatible with the MOSFET process [10]. Hence, it is possible to have hybrid circuits as stated in [11] and [12].

We also propose a a dual-V_t 6T SRAM cell, which adopts low-power (high-V_t) and high-performance (low-V_t) FinFETs for cross-coupled inverters and access transistors, respectively. For both SRAM cells, the V_{dd} boost read-assist technique is employed to improve the read stability. The comparative study of proposed SRAMs is performed using HSPICE simulations for technologies with the gate length of 20 nm using a 128×128 SRAM array. According to our simulation results, while the lowest operating V_{dd} for the dual-V_t cell is 225 mV, the hybrid



(a) (b) (c)

Fig. 1. (a) ON current, (b) OFF current, and (c) ON/OFF current ratio values of TFET, HP FinFET, and LSTP FinFET devices. Vertical axes are in logarithmic (base 10) scale. For supply voltage levels below 250 mV, TFET has the highest ON/OFF current ratio, which makes TFETs a suitable device for ultra-low voltage memory cells.

cell can reliably operate at V_{dd} levels as low as 125 mV. However, to further decrease the access delay of the hybrid cell for 125 mV $\leq V_{dd} <$ 225 mV, where the dual-V_t cell does not function, negative Gnd read-assist technique and a boosted voltage for the address decoder are used. Finally, we present a 125 mV 2ns-access-time 16Kb SRAM array based on the proposed hybrid TFET-FinFET SRAM cell.

The rest of this paper is organized as follows. In Section II, the adopted devices and two SRAM cell designs are described. Characteristics of SRAM cells (noise margins and leakage power) and SRAM arrays (access delay and energy consumption) are studied in Section III and Section IV, respectively. Section V discusses performance improvements for the proposed hybrid cell at ultra-low supply voltages. Finally, Section VI concludes the paper.

II. Devices and SRAM Cells

A. Devices

In recent years, several research work have concentrated on the operation and characteristics of inter-band TFETs [13]. Due to their potentials for sub-KT/q subthreshold-slope operation, TFETs can be considered for very low power digital circuits by aggressively scaling down the supply voltage. Conventional MOSFETs switch between the ON and OFF states via modulating an energy barrier which exists in the channel. Changing this barrier regulates the current flow (thermal diffusion) between the source and drain terminals. In the case of TFETs, the current flows through band to band tunneling and the switching between ON and OFF states is performed via aligning and misaligning the energy bands which participate in the tunneling.

In this work, for simulating TFET devices, the Universal TFET model with the gate length of 20 nm is used [14]. This device model contains the following four structures: (i) double-gate, (ii) single-gate, (iii) high-performance, and (iv) low-power. To make the conditions of devices similar, we choose the double-gate structure. Moreover, for FinFETs, we use the PTM multi-gate (PTM-MG) transistor model [15], which includes both high-performance (HP) and low-standby power (LSTP) FinFETs. Specifically, the 16 nm FinFET device whose gate length is equal to 20 nm is adopted for this work. The effective width of a FinFET device, W_{eff}, is obtained by $W_{eff} = 2 \times H_{fin} + W_{fin}$, where H_{fin} and W_{fin} denote the height and width of the fin, respectively [16]. For the adopted FinFET devices, the effective width is 64 nm. Based on this, we selected the width of the TFET device to be $3 \times l_g$, where l_g is the gate length.

Since parameters of the SRAM cell are evaluated at different supply voltages in this work, we initially compare the I-V characteristics of the TFET and FinFET devices. Characteristics of the n-type and p-type transistors for both devices are in the same order, and hence, results are only presented for the n-type transistor. Characteristics include the ON current, OFF current, and ON/OFF current ratio which are plotted for a wide range of supply voltages in Fig. 1. These results, which have been obtained using HSPICE simulations, indicate that HP FinFET (LSTP FinFET) has the highest ON current (lowest OFF current) over all shown V_{dd} levels. However, the interesting result in terms of the stability of the SRAM cell is the ON/OFF current ratio. As shown in Fig. 1(c), for supply voltage levels below 250 mV, TFET has the highest ON/OFF current ratio. Because of this feature, cross-coupled inverters made of TFET devices will have a higher noise margin at such low voltages, which subsequently makes TFETs a suitable device for ultra-low voltage memory cells.

B. SRAM Cells

To have a high-density memory, the conventional 6T SRAM cell with minimum-size transistors, whose structure is shown in Fig. 2(a), has been used for all our designs. This cell which has a symmetric structure consists of two access transistors to read and write data and two cross-coupled inverters for storing data. Considering three device options of TFET (T), LSTP FinFET (F^L), and HP FinFET (F^H), and allowing independent selections of transistors for cross-coupled inverters and access transistors, we have nine different SRAM designs. However, assigning TFETs or LSTP FinFETs to access transistors significantly degrades the access latency of the SRAM array. Accordingly, we only consider cases where a HP FinFET is used for access transistors. On the other hand, using HP FinFETs for cross-coupled inverters increases the leakage power of the SRAM cell. Hence, we end up with the following two SRAM cell designs (cf. Fig. 2):

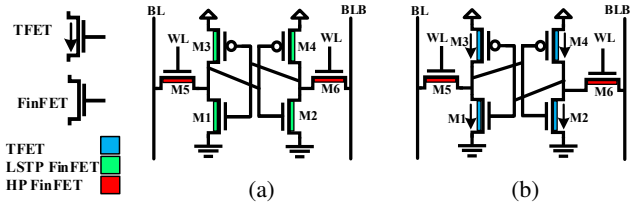


Fig. 2. 6T SRAM cell designs used in this paper. (a) Baseline cell: A dual-V_t FinFET SRAM denoted by F^LF^H. (b) Proposed cell: A hybrid TFET-FinFET SRAM denoted by TF^H.

1) A dual-V_t FinFET SRAM cell, denoted by F^LF^H, where LSTP FinFETs and HP FinFETs are used for cross-coupled inverters and access transistors, respectively.
2) A hybrid TFET-FinFET SRAM structure, denoted by TF^H, where TFETs and HP FinFETs are used for cross-coupled inverters and access transistors, respectively.

III. SRAM Cell Characteristics

In this section, we compare the cell-level characteristics, including hold, read, and write stability parameters and leakage power consumption, of TF^H and F^LF^H SRAMs at different V_{dd} levels. This study is especially important in order to assess the effectiveness of each cell for ultra-low voltage applications.

A. Hold Static Noise Margin

In large SRAM arrays, most of the cells are in the idle (hold) state, and hence, it is critical for SRAM cells to reliably hold their data while not being accessed. A parameter which is widely-used to measure the stability of the cell during the hold state is *hold static noise margin* (HSNM), which is measured using the butterfly curves [17]. Furthermore, *data retention voltage* (DRV) is the minimum V_{dd} level for which the SRAM cell has an acceptable HSNM.

HSNM values for the TF^H and F^LF^H SRAM cells for supply voltages ranging from 100 mV to 800 mV are shown in Fig. 3. In this paper, the minimum acceptable noise margin level is considered to be 20% of V_{dd}. Hence, according to Fig. 3, DRV values for F^LF^H and TF^H SRAM cells are 225 mV and 125 mV, respectively. The lower DRV value of TF^H cell is because of the higher ON/OFF current ratio of TFETs compared with LSTP FinFETs, which in turn enables cross-coupled inverters to maintain their data at ultra-low V_{dd} levels in the presence of external noises. In addition, below 300 mV, TF^H has a higher HSNM than F^LF^H, which makes the proposed TF^H design a suitable SRAM for extremely low voltage applications.

B. Static Power Consumption

For large SRAM arrays, static (leakage) power dissipation is the main component of the cache power consumption [18]. Hence, the static power of SRAM cells should be minimized to decrease the total power consumption of the cache memory. For this purpose, we use low-leakage devices (TFETs or LSTP FinFETs) for cross-coupled inverters. Keeping the SRAM cell at its DRV point during the hold state also helps reducing the static power consumption.

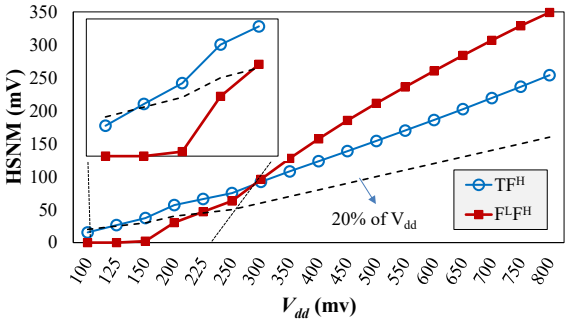


Fig. 3. Hold SNM (HSNM) values of TF^H and F^LF^H SRAM cells for different supply voltage levels. The dashed line shows the minimum acceptable value for HSNM at each V_{dd}.

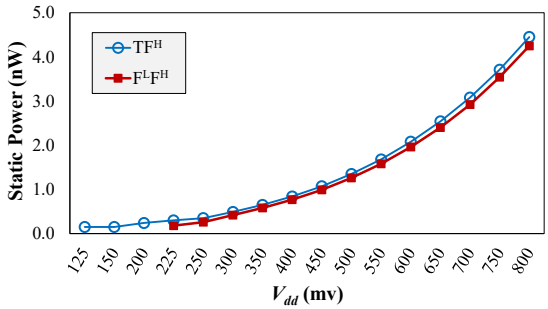


Fig. 4. Static power consumption of TF^H and F^LF^H SRAM cells for different supply voltage levels.

As shown in Fig. 4, the static power consumption of TF^H and F^LF^H SRAM cells are very close to each other. This is because there are two main sources of leakage path in the standard 6T SRAM cell. (i) V_{dd} to ground paths inside the cross-coupled inverters (cell leakage), and (ii) the bitline to ground path through the access transistor (bitline leakage). As mentioned earlier, by adopting low-leakage devices in our designs, we are able to significantly reduce the cell leakage paths. However, using HP FinFETs for access transistors to achieve fast SRAM cells is the main cause of static power dissipation in both cell designs. It should be noted that the OFF current of the adopted TFET is higher than that of the LSTP FinFET, which subsequently makes the TF^H design to slightly consume more power than the F^LF^H counterpart.

C. Write Margin

There are different definitions for the write stability of the SRAM cell. We chose the one which is based on the difference between V_{dd} and the minimum wordline (WL) voltage that can cause a successful write. This is called the combined wordline margin, which is denoted by WM in this paper [19]. A high write-ability may be achieved by strengthening access transistors and/or weakening pull-up transistors. In our SRAM designs, access transistors (which are made of HP FinFETs) have a significantly higher ON current than pull-up transistors (which are made of TFETs or LSTP FinFETs). As a result, high WMs for both TF^H and F^LF^H designs are expected. This can be confirmed in Fig. 5, which shows that WM valuess of

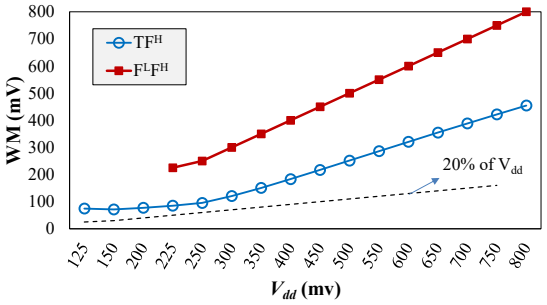


Fig. 5. Write margin (WM) values of TF^H and F^LF^H SRAM cells for different supply voltage levels. The dashed line shows the minimum acceptable value for WM at each V_{dd}.

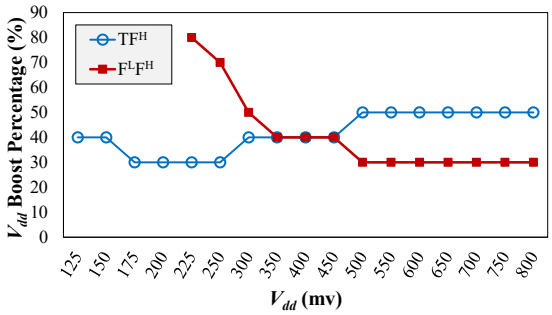


Fig. 6. The percentage of cell V_{dd} increase from the nominal supply voltage to achieve an acceptable read SNM (RSNM) for TF^H and F^LF^H SRAM cells.

both TF^H and F^LF^H designs are above the minimum acceptable level in all V_{dd} values. Hence, there is no need for a write-assist technique for both SRAM cells.

D. Read Static Noise Margin

In the conventional 6T SRAM cell, when SRAM is storing 0 (i.e., Q = "0"), the voltage of node Q rises during the read operation. If this voltage becomes larger than the trip voltage of the right inverter (i.e., transistors M2 and M4 in Fig. 2), the stored data in the cell will flip, resulting in a *destructive* read operation. The stability of the cell during the read operation is typically measured by the read SNM (RSNM), which, similar to HSNM, is calculated using the butterfly curves with access transistors turned on [17]. For a stable read operation, weaker access transistors compared with pull-down transistors are preferred. In addition, a steep subthreshold slope for the powered-off transistors in the cross-coupled inverters increases the stability of the cell during the read operation. The reason is because, under this scenario, butterfly curves will have a steeper transition, and thus, a larger rectangle can be fitted in the butterfly curves.

TF^H and F^LF^H SRAM cells have destructive read operations, since access transistors in both cells are significantly stronger than pull-down transistors. To overcome this issue and increase the RSNM to be at least 20% of V_{dd}, we will apply a read-assist technique. Common and effective read-assist techniques include: (i) V_{dd} boost, (ii) negative Gnd, and (iii) wordline underdrive (WLUD). WLUD technique improves the RSNM by weakening the access transistor, which on the other hand, decreases the read current and hence increases the read time (resulting in a slower SRAM). Negative Gnd technique also has a small impact on the RSNM. Therefore, to improve the read stability, we adopt the V_{dd} boost read-assist technique, which is an effective method for bringing RSNM values above the acceptable levels [20] [21]. Moreover, V_{dd} boost, because of strengthening pull-downs, can improve the read access time.

After applying the V_{dd} boost technique, the new cell supply voltage will be $(1+p)\cdot V_{dd}$, where p is the percentage of increase in cell supply voltage from the nominal V_{dd}. A different value of p (V_{dd} boost percentage) is needed for each supply voltage level to achieve the acceptable RSNM value. Hence, for each V_{dd} between 125 mV and 800 mV, we increase the value of p with steps of 0.1 until the cell has an acceptable RSNM value. In Fig. 6, values of p for different V_{dd} levels in TF^H and F^LF^H SRAM cells are reported. As can be seen, in low supply voltages (below 300 mV), due to higher ON/OFF current ratio of TFETs compared with FinFETs, TF^H SRAM cell requires a lower p value to achieve the acceptable RSNM level.

IV. SRAM Array Characteristics

Access delays and energy consumptions of SRAM arrays made of TF^H and F^LF^H cells at different V_{dd} levels are presented in this section.

A. Array Structure

A 128×128 (16Kbit) SRAM array has been implemented in this paper. All peripheral circuits are made of HP FinFETs because of performance considerations and the small contribution of such circuits to the overall leakage power of the array. The resistance and capacitance of interconnects have also been considered for the SRAM array. To do this, we model each SRAM cell as shown in Fig. 7. Resistance and capacitance values of interconnects are from [22] for the 16-nm technology node. More specifically, the interconnect resistance is 36.5 Ω/μm and the interconnect capacitance is 0.145 fF/μm. The

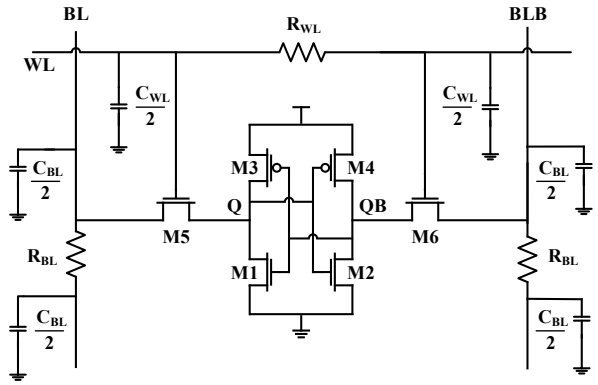


Fig. 7. SRAM cell model considering interconnect resistances and capacitances. R_{WL} (R_{BL}) and C_{WL} (C_{BL}) denote the wire resistance and capacitance along the width (height) of the SRAM cell, respectively.

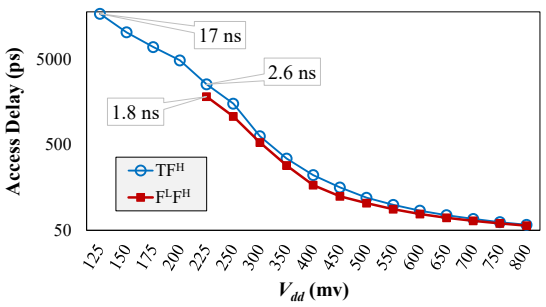


Fig. 8. Access time of a 128×128 array made of TF^H and F^LF^H SRAM cells for different supply voltage levels.

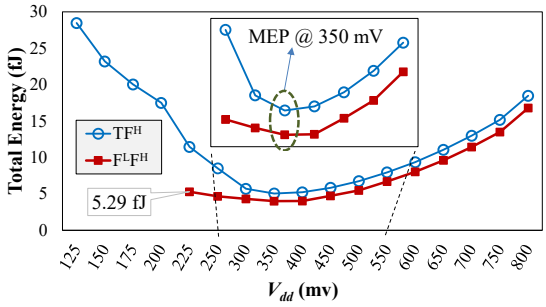


Fig. 9. Total energy consumption of a 128×128 array made of TF^H and F^LF^H SRAM cells for different supply voltage levels. For both cases, the minimum energy point (MEP) is 350 mV.

width and height of TF^H and F^LF^H SRAM cells are determined based on their layout, which are then used to derive resistances and capacitances per width and height of the cell.

Read and write access delays for the cell at the top-right corner of the array are measured as follows. Read access delay is calculated from the time the address decoder is activated until the difference between BL and BLB becomes 100 mV [23]. The write access delay is defined from the time when address decoder is activated until voltage levels of nodes Q and QB cross each other. Read and write access energy consumptions are calculated during their corresponding operation. The energy overhead of boosted cell V_{dd} and negative cell Gnd which is used only in Section V for read operation have also been considered during HSPICE simulations.

B. Access Delay and Energy Consumption

The access delay and total energy consumption of the array, denoted by D_{Array} and E_{Array}, respectively, are obtained as follows:

$$D_{Array} = \max(D_{Read}, D_{Write}) \quad (1)$$

$$E_{Access} = \beta \cdot E_{Read} + (1-\beta) \cdot E_{Write} \quad (2)$$

$$E_{Standby} = P_{Static} \cdot D_{Array} \quad (3)$$

$$E_{Array} = \alpha \cdot E_{Access} + (1-\alpha) \cdot E_{Standby} \quad (4)$$

where D_{Read} (D_{Write}) and E_{Read} (E_{Write}) denote the read (write) access delay and read (write) access energy consumption, respectively, β is the ratio of read accesses to the total accesses, α denotes the probability of accessing the array in each cycle, P_{Static} is the leakage power of the SRAM array, and E_{Access} ($E_{Standby}$) denotes the energy consumption of the array during the access (standby) mode.

Using above equations, D_{Array} and E_{Array} of SRAM arrays made of TF^H and F^LF^H cells at different supply voltages assuming α = 0.15 and β = 0.5 have been calculated and shown in Fig. 8 and Fig. 9, respectively. These results show that for 225 mV $\leq V_{dd} \leq$ 800 mV, the F^LF^H array achieves the lowest access delay (on average 14% lower than TF^H array) and the lowest energy consumption (on average 21% lower than TF^H array). This is mainly because of larger Miller capacitances in TFETs compared to FinFETs [12].

The minimum energy point (MEP) for both designs occur at 350 mV. Under the MEP, the total energy consumption of the F^LF^H design is 4.01 fJ per access, which is 21% lower than that of the TF^H counterpart. On the other hand, for 125 mV $\leq V_{dd} <$ 225 mV, where only the TF^H array can operate, significant increase in the D_{Array} of TF^H-based design is observed, which subsequently increases its E_{Array}. Since this is the region of interest in this paper, in the next section we will apply two read-assist techniques to decrease D_{Array}, and thus, obtain a relatively fast SRAM array at such ultra-low voltages.

V. Improving the Performance of the TF^H Structure in Sub-225mV Supply Voltages

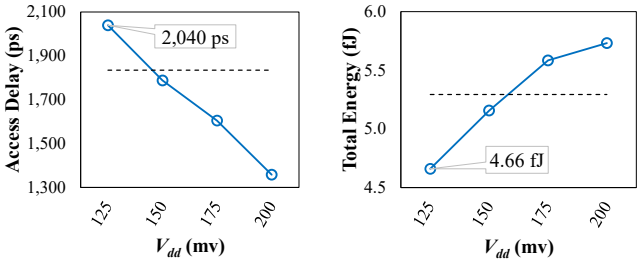
In this section, we focus on ultra-low supply voltages, covering V_{dd} levels between 125 mV and 200 mV. In this region, only the TF^H design can reliably hold data. However, as mentioned in the previous section, TF^H design in this region suffers from high access delays which also increases the total energy consumption of the array. Accordingly, techniques for performance improvement are described next. Our target is to make access delays of TF^H array close to that of the F^LF^H array at 225 mV.

There are two main reasons for the delay increase in the region of interest. (i) As indicated in Fig. 1(a), for 125 mV $\leq V_{dd} \leq$ 200 mV, ON current of transistors significantly degrades. Consequently, read current decreases, resulting in longer read access delays. However, the read current of the SRAM cell can be increased by applying the negative Gnd read-assist technique. (ii) At ultra-low supply voltages, the delay of the address decoder of the SRAM array also significantly increases. For example, based on our HSPICE simulations on a 7-to-128-line decoder, the delay of the decoder at 125 mV is 11× larger than the delay at 225 mV. To reduce this delay, we will use the boosted cell V_{dd}, which is needed to maintain the RSNM above the minimum acceptable range, for the supply voltage of the address decoder as well. In case of 125 mV, as shown in Fig. 6, 40% boosted cell V_{dd} is needed. However, this V_{dd} boost percentage is not sufficient to achieve the desired access delay. As a result, a higher supply voltage (80% of V_{dd}) is used for supply voltages of SRAM cells and address decoder.

Percentages of V_{dd} boost and negative Gnd techniques applied to the TF^H array for performance enhancement are reported in Table I. Access delay and total energy consumption values are also shown in Fig. 10. As can be seen, the lowest

TABLE I. Percentages of V_{dd} boost and negative Gnd read-assist techniques applied to TF^H array at ultra-low voltages for performance enhancement. For TF^H array, V_{dd} boost has also been applied to the suplly voltage of the address decoder.

SRAM	V_{dd}	V_{dd} Boost	Cell V_{dd}	Negative Gnd	Cell V_{ss}
TF^H	125 mV	80%	225.0 mV	90%	-112.5 mV
TF^H	150 mV	40%	210.0 mV	80%	-120.0 mV
TF^H	175 mV	30%	227.5 mV	50%	-87.5 mV
TF^H	200 mV	30%	260.0 mV	30%	-60.0 mV
F^LF^H	225 mV	80%	405.0 mV	N/A	N/A



(a) (b)

Fig. 10. Access delay and total energy consumption of the TF^H array after applying negative Gnd read-assist technique and boosted V_{dd} for the address decoder. Dashed line on each figure shows the result of F^LF^H array at 225 mV (its lowest operating voltage) without the performance enhancement techniques.

energy consumption of the TF^H design is obtained at 125 mV, which has an access time of 2 ns. Techniques discussed in this section can also be applied to the F^LF^H design on its operating region to further decrease the access delay.

VI. Conclusion

In this paper, we proposed two 6T SRAM cell designs using minimum-sized transistors. In both designs, we used HP FinFETs to increase the performance of the SRAM array, and low leakage devices with high ON/OFF current ratios (TFETs or LSTP FinFETs) to increase noise margins and decrease the leakage power. This device selection for an SRAM cell results in a destructive read operation. Therefore, we used V_{dd} boost technique to not only achieve a non-destructive read operation, but also increase the RSNM. Based on our HSPICE simulations the DRV of F^LF^H design is 225 mV, whereas that of TF^H, due to the higher ON/OFF current ratio at ultra-low voltages, is 125 mV. We showed that for a 128×128 array, the F^LF^H design has a better access delay and total energy consumption when $V_{dd} \geq 225$ mV. However, for ultra-low supply voltages (125 mV $\leq V_{dd} <$ 225 mV), only the TF^H design operates, but it has a high access time. Accordingly, we used negative Gnd read-assist technique along with a boosted supply voltage for the address decoder to reduce the access time. After our optimizations, we presented a 125mV 2ns-access-time 16Kb SRAM array based on the proposed hybrid TFET-FinFET cell.

Acknowledgment

This research is supported by a grant from the National Science Foundation.

References

[1] International Technology Roadmap for Semiconductors (ITRS). [Online]. Available: http://www.itrs.net/Links/2012ITRS/Home2012.html

[2] O. Hirabayashi *et al.*, "A process-variation-tolerant dual-power-supply sram with 0.179 μm^2 cell in 40nm cmos using level-programmable wordline driver," in *IEEE International Solid-State Circuits Conference (ISSCC)*, Feb 2009, pp. 458–459,459a.

[3] L. Chang *et al.*, "Stable SRAM Cell Design for the 32 nm Node and Beyond," in *Symposium on VLSI Technology*, June 2005, pp. 128–129.

[4] E. Karl *et al.*, "A 0.6 V, 1.5 GHz 84 Mb SRAM in 14 nm FinFET CMOS Technology With Capacitive Charge-Sharing Write Assist Circuitry," *IEEE Journal of Solid-State Circuits (JSSC)*, vol. 51, no. 1, pp. 222–229, Jan. 2016.

[5] E. Karl *et al.*, "A 4.6 GHz 162 Mb SRAM Design in 22 nm Tri-Gate CMOS Technology With Integrated Read and Write Assist Circuitry," *IEEE Journal of Solid-State Circuits (JSSC)*, vol. 48, no. 1, pp. 150–158, Jan. 2013.

[6] A. Shafaei, H. Afzali-Kusha, and M. Pedram, "Minimizing the Energy-Delay Product of SRAM Arrays using a Device-Circuit-Architecture Co-Optimization Framework," in *53rd Design Automation Conference (DAC)*, June 2016, pp. 1–6.

[7] D. H. Morris, U. E. Avci, and I. A. Young, "Variation-tolerant dense TFET memory with low VMIN matching low-voltage TFET logic," in *Symposium on VLSI Technology (VLSI Technology)*, June 2015, pp. T24–T25.

[8] A. C. Seabaugh and Q. Zhang, "Low-Voltage Tunnel Transistors for Beyond CMOS Logic," *Proceedings of the IEEE*, vol. 98, no. 12, pp. 2095–2110, Dec 2010.

[9] R. Mukundrajan *et al.*, "Ultra Low Power Circuit Design Using Tunnel FETs," in *IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, Aug 2012, pp. 153–158.

[10] F. Mayer *et al.*, "Impact of SOI, Si1-xGexOI and GeOI substrates on CMOS compatible Tunnel FET performance," in *IEEE International Electron Devices Meeting (IEDM)*, Dec 2008, pp. 1–5.

[11] Y. N. Chen *et al.*, "Evaluation of Stability, Performance of Ultra-Low Voltage MOSFET, TFET, and Mixed TFET-MOSFET SRAM Cell With Write-Assist Circuits," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems (JETCAS)*, vol. 4, no. 4, pp. 389–399, Dec 2014.

[12] T. C. Wu *et al.*, "Evaluation of energy-efficient latch circuits with hybrid tunneling FET and FinFET devices for ultra-low-voltage applications," in *28th IEEE International System-on-Chip Conference (SOCC)*, Sept 2015, pp. 339–344.

[13] K. K. Bhuwalka *et al.*, "Vertical tunnel field-effect transistor," *IEEE Transactions on Electron Devices*, vol. 51, no. 2, pp. 279–282, Feb 2004.

[14] H. Lu, D. Esseni, and A. Seabaugh, "Universal analytic model for tunnel fet circuit simulation," *Solid-State Electronics*, vol. 108, pp. 110 – 117, 2015.

[15] Predictive Technology Model (PTM). [Online]. Available: http://ptm.asu.edu

[16] G. Pei *et al.*, "FinFET Design Considerations based on 3-D Simulation and Analytical Modeling," *IEEE Transactions on Electron Devices*, vol. 49, no. 8, pp. 1411–1419, Aug 2002.

[17] E. Seevinck, F. List, and J. Lohstroh, "Static-Noise Margin Analysis of MOS SRAM Cells," *IEEE Journal of Solid-State Circuits*, vol. 22, no. 5, pp. 748–754, Oct 1987.

[18] M. F. Chang *et al.*, "A 130 mV SRAM With Expanded Write and Read Margins for Subthreshold Applications," *IEEE Journal of Solid-State Circuits (JSSC)*, vol. 46, no. 2, pp. 520–529, Feb 2011.

[19] M. Yabuuchi *et al.*, "A 45nm 0.6V cross-point 8T SRAM with negative biased read/write assist," in *Symposium on VLSI Circuits*, June 2009, pp. 158–159.

[20] B. Zimmer *et al.*, "SRAM Assist Techniques for Operation in a Wide Voltage Range in 28-nm CMOS," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 59, no. 12, pp. 853–857, 2012.

[21] R. W. Mann *et al.*, "Impact of circuit assist methods on margin and performance in 6t sram," *Solid-State Electronics*, vol. 54, no. 11, pp. 1398 – 1407, 2010.

[22] C. Pan and A. Naeemi, "A Paradigm Shift in Local Interconnect Technology Design in the Era of Nanoscale Multigate and Gate-All-Around Devices," *IEEE Electron Device Letters*, vol. 36, no. 3, pp. 274–276, March 2015.

[23] A. Islam and M. Hasan, "Leakage Characterization of 10T SRAM Cell," *IEEE Transactions on Electron Devices*, vol. 59, no. 3, pp. 631–638, March 2012.