Variation-Tolerant Ultra Low-Power Heterojunction Tunnel FET SRAM Design

Vinay Saripalli[†], Suman Datta[§] and Vijaykrishnan Narayanan[†]
[†]Department of Computer Science and Engineering and
[§]Department of Electrical Engineering
Pennsylvania State University, University Park, PA 16802
Email: {vxs924,vijay}@cse.psu.edu, sdatta@engr.psu.edu

Jaydeep P. Kulkarni
Circuit Research Lab, Intel Corporation
Hillsboro, OR 97124-5961
Email: jaydeep.p.kulkarni@intel.com

***Abstract*—Steep sub-threshold Interband Tunnel FETs (TFETs) are promising candidates for low supply voltage applications with higher switching performance than traditional CMOS. Unlike CMOS, TFETs exhibit uni-directional conduction due to their asymmetric source-drain architecture, and delayed output saturation characteristics. These unconventional characteristics of TFETs pose a challenge for providing good read/write noise margin characteristics in TFET SRAMs. We provide an analysis of 8T and 10T TFET SRAM cells, including Schmitt-Trigger (ST) based cells, to address these shortcomings. By benchmarking a variety of TFET-based SRAM cells, we show the utility of the Schmitt-Trigger feedback mechanism in improving the read/write noise margins, thus enabling ultra low-V$_{CC}$ operation for TFET SRAMs. We also propose a variation model for studying the impact of device-level variation on TFET SRAM cells. We show that the TFET ST SRAM cell has sufficient variation tolerance to operate at low-V$_{CC}$, and is a very promising cell to achieve a V$_{CC}$-min of 124mV. The TFET ST cell operating at its V$_{CC}$-min provides a 1.2x reduction in dynamic energy and 13x reduction in leakage power compared to the best CMOS-based SRAM implementation operating at it's V$_{CC}$-min, while giving better performance at the same time.**

I. Introduction

Voltage scaling is fundamental to achieving energy efficient operation in digital circuits due to the quadratic reduction in dynamic energy with V$_{CC}$ scaling. Numerous design techniques have been proposed both at the circuit-level and the architectural-level [1], [2] to enable low-V$_{CC}$ operation using CMOS digital circuits. Further requirement for energy reduction drives the operation of CMOS digital circuits into sub-threshold operation, thus increasing the sensitivity of the circuit parameters to device-level variation. It also causes exponential increases in delay, causing the circuit operation to be leakage-energy dominated. SRAM bit-cells employing minimum-sized transistors can be particularly vulnerable to device-level variation occurring due to the process flow (intradie as well as interdie) [3]. Due to the added sensitivity of the minimum sized transistors to variation at the device-level, SRAM bit-cells are most prone to access failures in reduced V$_{CC}$ operation. Thus, there is a need for robust variation-tolerant SRAM design, capable of sub-300mV operation. Numerous designs have been proposed to address the challenge of sub-threshold operation of CMOS SRAMs [4], [5], [6]. CMOS SRAM operation at 160 mV has also been shown [7].

Interband Tunnel FETs with a promise of sub-60 mV/decade sub-threshold slope have garnered tremendous interest in recent years. The idea is to enable low-V$_{CC}$ operation with strong On-current by taking advantage of the steep-slope. The only way to allow strong drive current operation at low-V$_{CC}$ ($<$ 300mV) in CMOS is to reduce the V$_T$, which in turn causes an unacceptable increase in the Off-Current - This is a fundamental limitation in CMOS due to the 60 mV/dec sub-threshold slope limitation. A number of TFETs have been experimentally demonstrated in recent literature, showcasing the progress in fabrication and experimental demonstration of novel tunneling devices. A vertically-oriented, gate-all-around silicon nanowire was demonstrated recently showing 50 mV/decade over 3 decades of drain current [8]. A horizontally-oriented Ultra-Thin-Body (UTB) InAs-on-Silion TFET was also demonstrated recently showing the utility of a III-V semiconductor layer-transfer-technique in TFET fabrication [9]. Further, a process flow for the creation of a side-gated vertical-mesa TFET which can be scaled down to achieve an UTB double-gated structure has also demonstrated [10]. Thus, the efforts being undertaken in the fabrication of vertical and horizontal UTB tunneling structures show great promise in experimental demonstration of UTB TFETs with a steep sub-threshold slope. However, TFETs are unconventional devices with unique properties which pose challenges to robust SRAM design, which we address in this paper.

In this paper, we study TFET SRAMs from a combined technology and architecture perspective. The paper is organized as follows - (1) characterization of 8T and 10T TFET SRAM cells to address the challenges of TFET SRAM design, in section II, (2) proposal of a model for studying variation in UTB TFETs and a small-signal variation model suitable for circuit simulations, sections III-A & III-B, and (3) characterization of read-failure probability vs. V$_{CC}$ in the presence of variation for TFET SRAMs using circuit-level Monte-Carlo simulations, in section III-C. Using the read-failure probabilities, we extract the V$_{CC}$-min for different CMOS and TFET SRAM cells and show that it is possible for a TFET Schmitt-Trigger based SRAM cell to achieve a lower V$_{CC}$-min compared to the CMOS SRAM cells, thus allowing ultra-low V$_{CC}$ operation.

II. SRAM Cell Design and Characterization

A. TFET and CMOS Device Models

In order to compare the characteristics of Si CMOS and TFET-based SRAMs, it is important to choose accurate models for the underlying transistors. The device models which are used to compare Si CMOS and TFET SRAM cell characteristics are described here. Fig. 1(B) shows the I$_d$-V$_g$

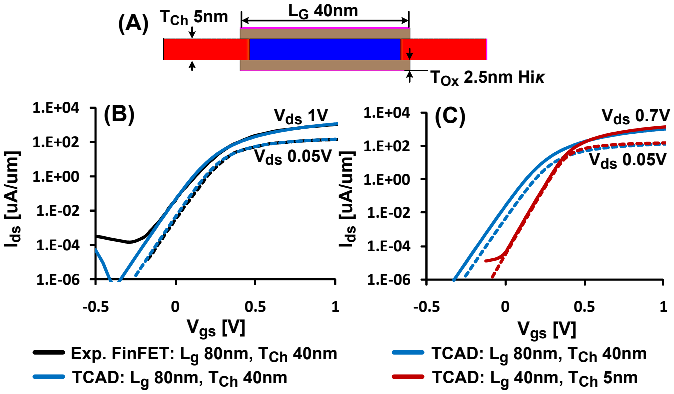


Fig. 1. (A) Structure of UTB Si FinFET (B) Simulation of I_d-V_g characteristics of an experimentally demonstrated FinFET [11] and (C) Simulation of I_d-V_g characteristics of a scaled FinFET

characteristics of an experimentally demonstrated n-channel Si FinFET device [11], and it's simulation using a TCAD double-gated structure. For CMOS SRAMs, we assume a highly scaled UTB Si FinFET, with a nearly ideal 60-mV/dec sub-threshold slope. Fig. 1(A) shows the TCAD structure and Fig. 1(C) shows the simulation of such a highly scaled Si FinFET, which is obtained by scaling the T_{Ch} and T_{Ox} of the experimental Si FinFET.

For the TFET SRAMs, we assume a highly scaled UTB n-channel GaSb/InAs heterojunction TFET (HTFET), whose structure is shown in Fig. 2(A). The TCAD simulation of the HTFET compares well with an full-band atomistic simulation of the same structure [12], as shown in Fig. 2(B-C). The structure studied in the atomistic simulation [12] does not include a gate-source overlap, which is unavoidable when fabricating a UTB side-gated interband tunneling structure, similar to that shown in Fig. 2(A). Since the source is P+ doped, a positive V_g applied during the operation of the n-channel HTFET creates a depletion region under the gate-source overlap region, as shown in Fig. 3. In our study of TFET-based SRAMs, we assume a gate-source overlap of 2nm. This leads to a reduced On-current compared to the case without a gate-source overlap. An On-current reduction of 1.35**x** is shown for the HTFET @ V_{CC} 0.5V (Fig. 4). Further, a change in the position of the gate-edge over the source causes a fluctuation of the depletion region under the

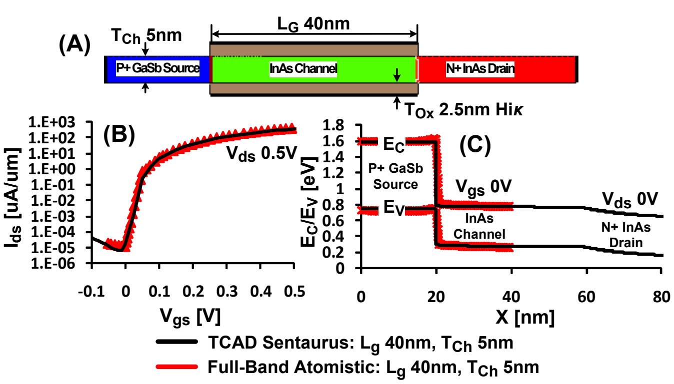


Fig. 2. (A) Structure of UTB GaSb/InAs nearly broken-gap TFET (B) Comparison of simulated I_d-V_g characteristics using TCAD and OMEN [12] and (C) Comparison of simulated band-structure using TCAD and Omen

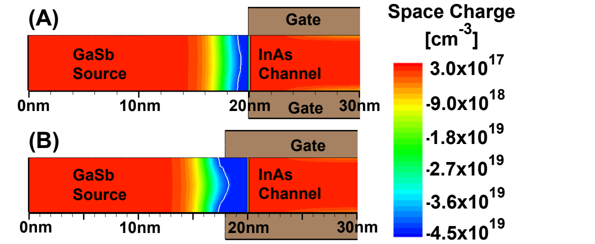


Fig. 3. Space-Charge region in P+ GaSb source (V_{ds} 0.5V, V_{gs} 0.5V) (A) without gate-source overlap and (B) with 2nm gate-source overlap

gate. We show in section III-B that this can be a major cause of On-current variation in side-gated tunneling structures.

The drive current of the simulated p-channel Si FinFET is $^1/_2$ of the drive current of the n-channel Si Finfet discussed here. The p-channel HTFET drive is $^1/_2$ of the n-channel counterpart. This reduction in the inter-band tunneling current is due to the reduced doping concentration of the n+ source region, which is needed to maintain the steep switching slope by reducing the amount of Fermi level degeneracy of the source [13].

B. TFET Saturation Characteristics and Impact on SRAM

Fig. 5(A-B) compares the I_{On} vs I_{On}/I_{Off} characteristics of a GaSb/InAs HTFET and a Si NMOS at V_{CC} 0.7V and V_{CC} 0.3V. It is clear that the HTFET is a superior device compared to Si NMOS in its sub-threshold region, showing both higher I_{On} as well as higher I_{On}/I_{Off} ratio. However, it is also important to consider the I_d-V_d characteristics of the HTFET, since the saturation voltage (V_{d-sat}) plays an important role in the noise-margin characteristics of digital circuits. Fig. 5(C-D) compares the saturation characteristics of a HTFET and a Si NMOS. The HTFET behaves like a device with a very low V_T (close to 0V), and hence shows delayed output saturation characteristics. Apart from delayed saturation, the HTFET also shows uni-directional conduction due to the asymmetric *p-i-n* structure.

To perform circuit simulations, we capture the I_d-V_g transfer characteristics of the CMOS and the HTFET obtained using the models discussed in section II-A, in a Verilog-A lookup table [14], [15], [16]. Because of delayed onset of saturation in the HTFET, the Voltage Transfer Characteristics (VTC) of a HTFET inverter are considerably degraded compared to that of a CMOS inverter, as shown in Fig. 6. Further, we consider a 6T TFET SRAM (Fig. 7(A)) with inward-facing access transistors, (*inward* is defined as being able to conduct current from the bit-line into the storage-node of the

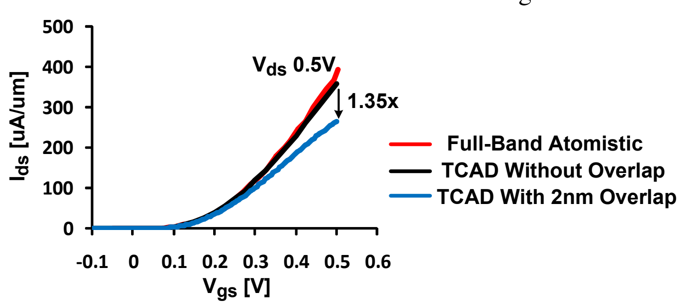


Fig. 4. On-current reduction due to source-depletion @ V_{cc} 0.5V

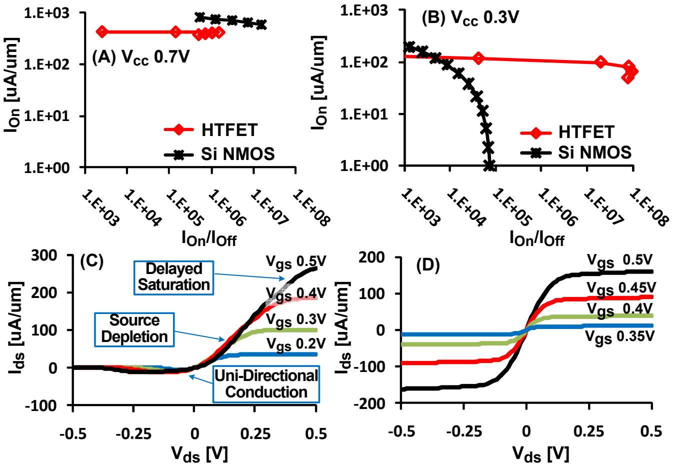


Fig. 5. I_{On} vs. I_{On}/I_{Off} comparison for CMOS and HTFETs at (A) V_{CC} 0.7V and (B) V_{CC} 0.3V. (C) I_d-V_g characteristic for HTFET and (D) I_d-V_g characteristic for Si NMOS

cell, i.e. the direction in which the arrow points), and a 6T CMOS SRAM (Fig. 7(B)) with similar sized transistors and compare their read-SNMs. We find that read-SNM for the 6T TFET SRAM is considerably degraded compared to the 6T CMOS SRAM(Fig. 7)(C-D). Further, due to uni-directional conduction, the Write-SNM for the TFET SRAM cell shown in Fig. 7(A) is zero [14].

C. Design of 8T and 10T TFET SRAM

As discussed in section II-B, the 6T TFET SRAM with inward-facing access transistors cannot perform a write successfully. It has been shown that a 6T TFET SRAM with either inward or outward-facing access transistors, cannot simultaneously do both read and write [14]. In order to circumvent this limitation, a 6T TFET SRAM with one-inward and one-outward facing access transistor has been proposed [15]. However, a virtual-ground write-assist is required to perform a write successfully in this design. Another proposed approach [17] is to use a 6T TFET cell with a cell-ratio (β) of 0.6 to provide a robust write. This cell has a read-SNM close to zero (because of the low β value), and is fundamentally unstable during read. Instead, it relies on the application of a short read-pulse width, relying on the read-dynamic noise margin (DNM) characteristic of the 6T TFET cell, along with a ground-lowering read-assist, to avoid an upset during read operation. Thus, only 6T TFET cells which require a read/write-assist, or cells which are fundamentally unstable during read/write-access have been studied. In this work, we consider the design and characterization of TFET SRAM cells with higher (8T and 10T) transistor counts, and compare them with 6T and 10T CMOS SRAM cells. We do not consider cells

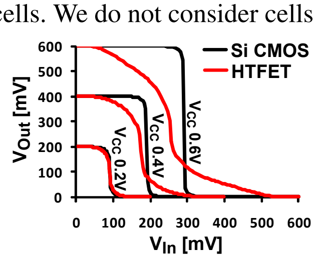


Fig. 6. VTC comparison for HTFET and Si CMOS (pull-up/pull-down is 1:1, p-channel drive current is $^1/_2$ the n-channel drive current)

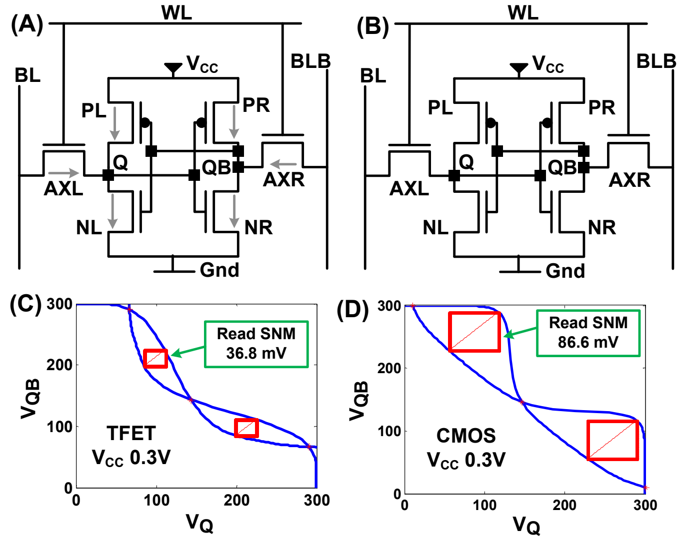


Fig. 7. (A) 6T TFET SRAM with inward-facing access transistors (arrows indicate direction of On-current), (PL/AXL/NL - 1/1/2) (B) 6T CMOS SRAM, (PL/AXL/NL - 1/1/2) (C) Static-RNM of 6T TFET SRAM @ V_{CC} 300 mV and (D) Static-RNM of 6T CMOS SRAM @ V_{CC} 300 mV

which are fundamentally unstable during read/write-access, or cells which require ground-assist to perform a read or write operation ([17], [15]).

Fig. 8 shows the read operation for various 8T and 10T TFET SRAM cell configurations. The TFET 8T Transmission-Gate SRAM cell in Fig. 8(A) has both inward and outward-facing TFET access transistors to overcome the problem of uni-directional conduction. Read is performed by enabling the inward-facing TFET access transistors AXL_{rd}(AXR_{rd})

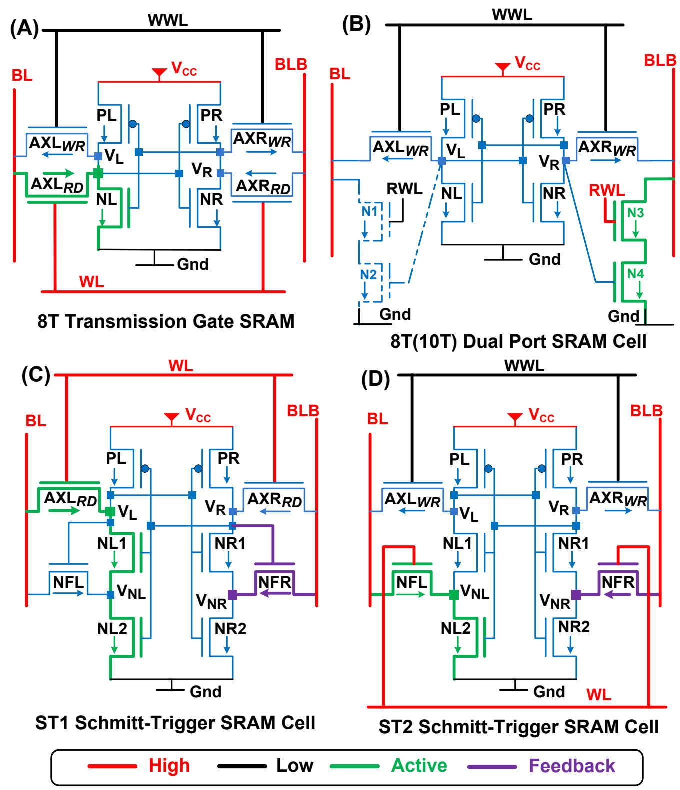


Fig. 8. Read-operation in various TFET SRAM cell configurations

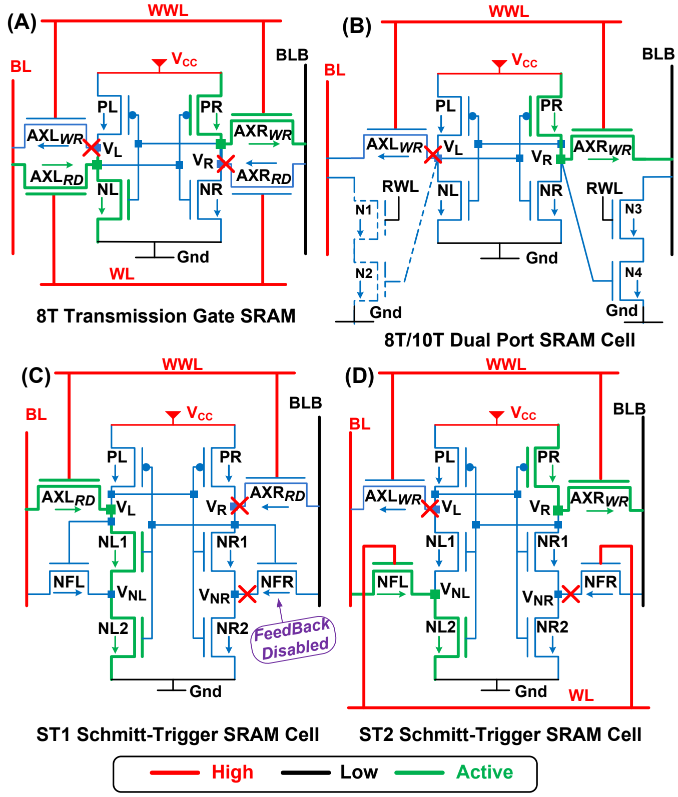


Fig. 9. Write-operation in various TFET SRAM cell configurations

using the read word-line (WL). The 8T(10T) dual-port SRAM cell ([18], [19], [5]), shown in Fig. 8(B), has separate read and write access ports. Thus, the read-SNM of the dual-port SRAM is same as the hold-SNM. CMOS Schmitt-Trigger (ST)-based SRAM cells have been proposed in [7] and [20]. The TFET ST-based SRAM cells (ST-1 and ST-2) differ from the CMOS counterparts only in the orientation of the access transistors. The TFET ST-1 SRAM cell (Fig. 8(C)) has inward-facing access transistors (AXL_{rd}/AXR_{rd}), as well as inward-facing feedback transistors (NFL/NFR). Read operation is performed by enabling the inward-facing TFET access transistors AXL_{rd}(AXR_{rd}) using the word-line (WL), while feedback is provided by the inward-facing NFR(NFL) transistors. The TFET ST-2 SRAM cell (Fig. 8(D)) has outward-facing access transistors (AXL_{wr}/AXR_{wr}) and inward-facing feedback transistors (NFL/NFR). Read operation is performed by enabling the inward-facing TFET feedback transistor NFL(NFR) using the read word-line (WL), while feedback is provided by the other inward-facing NFR(NFL) transistor.

Fig. 9 shows the write operation for different TFET SRAM cell configurations. For the TFET 8T Transmission-Gate SRAM cell (Fig. 9(A)), both word-lines (WL/WWL) are enabled during a write. The outward facing access transistor AXR_{wr}(AXL_{wr}) drives the cell-node voltage V_R(V_L) to **0**, and the inward facing access transistor AXL_{rd}(AXR_{rd}) assists in the write by raising the voltage of the complementary cell-node V_L(V_R). For 8T(10T) dual-port SRAM (Fig. 9(B)), the write operation is uni-axial due to the uni-directional conduction property of the outward facing access transistors. During a write, the cell-node voltage V_R(V_L) is driven to **0** only by the AXR_{wr}(AXL_{wr}) transistor, while the other transistor AXL_{wr}(AXR_{wr}) does not assist. The TFET ST-1 SRAM (Fig. 9(C)) also suffers from a uni-axial write operation because all the access transistors face inwards in this cell. In the TFET ST-2 SRAM, both word-lines (WL/WWL) are enabled during a write. The outward facing access transistor AXR_{wr}(AXL_{wr}) drives the cell-node voltage V_R(V_L) to **0**, and the inward facing feedback transistor NFL(NFR) assists by raising the voltage of the complementary node.

Cell-sizing has to be studied carefully for TFET SRAM cells in which the write operation is uni-axial (i.e. only one transistor participates in operation). For the 8T and 10T Dual Port SRAM cells, the write-access transistors face outwards and need to have sufficient width for a write operation to be completed unassisted. The cell-sizing in Table I suffices for the dual-port SRAMs. For the TFET ST-1 SRAM cell, inward-facing access transistors are used for both read, as well as write operation. Fig. 10(A) shows the dependence of SNMs on the Pull-up Ratio (PR) @ V_{CC} 300mV, assuming a fixed cell-ratio (β) of 1. The hold-SNM as well as the read-SNM are sufficiently large even for a very low Pull-up Ratio (PR **0.1**) because of the Schmitt-Feedback action of the NFL and NFR transistors (Fig. 8(C)). In fact, Fig. 10(A) also shows that, without feedback and for low PR values, the read-SNM is very low making the cell unstable during reads (this is consistent with the observation made previously [14]). Write operation can be performed in the ST-1 cell by setting one of the bit-lines to **0**, and by enabling the word-line (WL). Since, the feedback transistor is powered by the bit-line supply voltage, the feedback is disabled when the bit-line is set to **0** (Fig. 9(C)). When the pull-up is sufficiently weak (PR $<$ **0.2**), write can be performed successfully with a good write-SNM. When the pull-up is strong (PR $\geq$ **1**), the cell retains its data even when the feedback is disabled, causing the uni-axial write to fail. Further, disabling the feedback during write has the unwanted side-effect of disabling Schmitt-Feedback for all the cells which are column-neighbors of the row being written. Fig. 10(A) also shows the hold-SNM, with and without feedback, showing that there is sufficient hold-SNM ($>$ 95mV @ V_{CC} 300mV) even when the feedback is disabled, suggesting that temporary disabling of feedback in

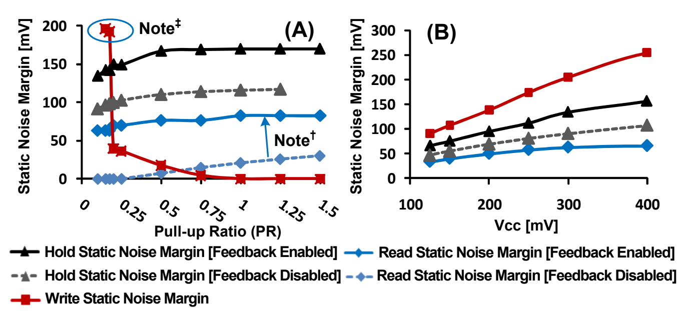


Fig. 10. (A) Static Noise Margins vs. Pull-up Ratio (PR) for TFET ST-1 Cell. Note†: Improvement in read-SNM is due to Schmitt-Feedback. Note‡: The weak p-channel is overpowered by the Fwd. biased current of the n-channel access transistor resulting in a strong write. (B) Static Noise Margins as a function of V_{CC} for a Pull-up Ratio of 0.15

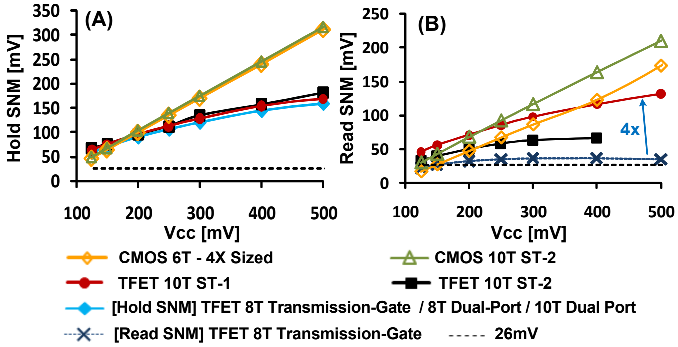


Fig. 11. Comparison of (A) Hold-Static Noise Margin and (B) Read-Static Noise Margin, of various SRAM cells

the hold-state is not a serious hindrance. Fig. 10(B) shows the read, write and hold-SNM (with and without feedback), as a function of V_{CC} for a Pull-up Ratio of **0.1**, illustrating that this sizing strategy is valid at all V_{CC}. Thus, using a sizing study for the TFET ST-1 cell, we show that it is possible to take advantage of the Schmitt-Feedback principle to circumvent the problem of weak read-SNMs in TFET SRAMs, and design an ST-1 SRAM cell capable of unassisted read and write operation. The sensitivity of the read-SNM of the TFET ST-1 cell to device-level variation is explored in section III-C. Section II-D illustrates a benchmarking study using various SRAM cells discussed here.

D. Characterization of 8T and 10T TFET SRAM

In order to compare the SRAM figures-of-merit, the transistor sizing has to be such that an *iso-area* condition is met. This requirement means that the memory sub-arrays realized using the candidate SRAM cells, while accounting for single-ended or differential read peripheral circuitry, should have the same area footprint. A sizing strategy has been proposed in [21] in order to study the figures-of-merit for various CMOS SRAM cells. In this paper, we assume that the relative cell sizes of SRAM cells realized using CMOS and HTFETs are comparable. Hence, we adopt a sizing strategy similar to [21], in order to compare the figures-of-merit of various CMOS and TFET SRAM cells. The cell-sizing used for various SRAM cells is shown in Table I. Only the TFET ST-1 SRAM cell has a greatly downsized pull-up transistor because an extremely weak pull-up is *necessary* for unassisted write operation, as discussed in section II-C. This changes the iso-area condition for the TFET ST-1 SRAM by a negligible amount ($< 5\%$) compared to the TFET ST-2 SRAM.

Fig. 11(A-B) compare the hold and read-SNMs of various TFET and CMOS SRAM cells. While CMOS SRAMs exhibit a better SNM at higher V_{CC}, the TFET SRAM cells provide better read-SNM characteristics at the desired low V_{CC} regime, due to their better drive currents at low V_{CC}. Among the TFET SRAM cells (Fig. 11(A)), ST-1 and ST-2 cells have marginally better hold-SNM than the 8T Transmission-Gate SRAM cell because of the feedback. The ST-1 cell is capable of giving a better hold-SNM than the ST-2 cell, but it is only marginally better in this comparison because the pull-up transistor has been downsized to enable write operations.

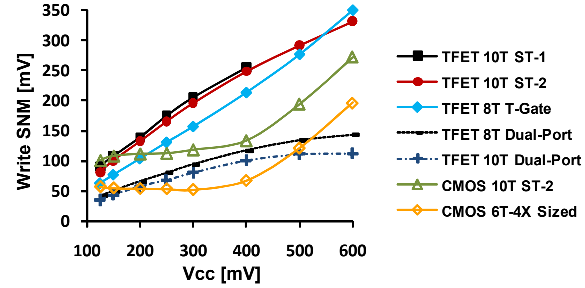


Fig. 12. Write-Noise Margin for various SRAM cells

Fig. 11(B) shows that the read-SNM of the 8T Transmission-Gate SRAM cell is considerably degraded because of the delayed saturation in TFETs, as explained in section II-B. The Schmitt-Feedback in the ST-2 cell improves the read-SNM by 4**x**. The read-SNM in the ST-2 cell is better than that of the ST-1 cell because the read-access occurs at a secondary node, V_{NL} (Fig. 8(D)). The ST-1 cell has a downsized pull-up transistor, and the read-access occurs directly at the cell storage-node V_L, causing the read-SNM to become lower than that of the ST-2 cell.

Fig. 12 shows the WNM characteristics of all the cells being considered. All the TFET SRAM cells, including those with uniaxial write (i.e. driven by only one access transistor), have a write-SNM of atleast 35mV, showing that TFET cells with higher transistor counts can perform unassisted writes unlike the 6T TFET SRAM cells. The TFET ST-2 cell and the 8T Transmission-Gate SRAM cell have the best write-SNM because write is performed using two access-transistors, one facing inwards and one facing outwards. We also observe that the 8T(10T) dual-port SRAM cells have the weakest write-SNM due to the uniaxial write operation. The ST-1 cell has a greatly improved write-SNM due to the use of a very weak pull-up, which can be afforded because its read and hold-SNM are protected by the Schmitt-Feedback. Downsizing the pull-up in the 8T(10T) dual-port SRAM cells would degrade the hold-SNM further, which is already the weakest among all TFET SRAM cells (Fig. 11(A)). This shows the utility of the Schmitt-Feedback in achieving significant noise-margin benefits in TFET SRAM cells.

We use a 256×256 SRAM array with 50fF bit-line capacitance (estimated using the cache estimation tool-CACTI [22]), to estimate dynamic energy consumption and read-access delay for different SRAM configurations. The word-line drives the access-transistors of 256 bit-cells in a row, and

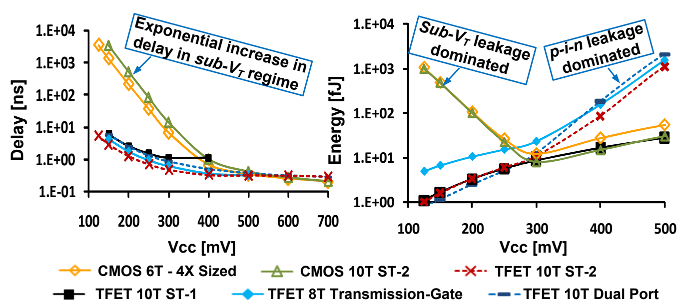


Fig. 13. (A) Access-delay comparison and (B) Dynamic Energy comparison for different SRAM cells

TABLE I
SUMMARY OF SRAM CELL SIZING FOR ISO-AREA COMPARISON (W IS THE NOMINAL WIDTH OF A TRANSISTOR).

	NL1/NR1	NL2/NR2	PL/PR	AXL_{Wr}/AXR_{Wr}	AXL_{Rd}/AXR_{Rd}	NFL/NFR	N1/N2	N3/N4
6T (CMOS Only) - 4X Sized	8W	-	4W	4W	-	-	-	-
8T Transmission-Gate (TFET Only)	2W	-	W	3W	W	-	-	-
8T Dual Port	2W	-	W	3W	-	-	W	W
10T Dual Port	2W	-	W	2W	-	-	W	W
Schmitt-Trigger (ST-1)	2W	2W	0.1W	-	W	2W	-	-
Schmitt-Trigger (ST-2)	2W	2W	W	2W	-	2W	-	-

is enabled using an appropriately sized driver circuit. The time taken to develop a 50mV differential bit-line voltage is used to estimate the read-access delay. The energy consumed by the driver in turning on the access transistors, together with the leakage energy consumption of the 256×256 bit-cell array is used to estimate the dynamic energy for the read-access. Fig. 13(A) shows that the TFET based cells have a lower delay compared to sub-threshold CMOS (V_{CC} < 500mV), whereas the CMOS based cells outperform the TFET cells at higher V_{CC}. This is consistent with the observation made in Fig. 5(A-B). The TFET ST-2 cell has the least delay out of all the cells at low V_{CC} due to its wider read-access transistors.

Fig. 13(B) shows the dynamic-energy SRAM cell access. For sub-threshold CMOS, the dynamic-energy is dominated by the leakage-component due to exponential increase in the access delay. The 8T Transmission Gate TFET cell, the TFET ST-2 cell and the TFET 8T(10T) dual-port cell, all have outward-facing access-transistors (to enable write) which have forward-biased p-i-n junctions in the hold-state. These forward-biased access transistors consume a significant amount of p-i-n leakage energy for V_{CC} > 300 mV. Only the TFET ST-1 cell does not have this forward-biased leakage because all its access transistors face inwards and are reverse-biased. As a result, for V_{CC} > 300mV, the TFET SRAM cells consume more dynamic-energy than their CMOS counterparts, mainly due to p-i-n leakage-energy domination. However, at V_{CC} < 300mV, all the TFET SRAM cells show sufficiently low forward-bias leakage to allow significantly energy-efficient operation compared to CMOS.

The conclusion of this benchmarking exercise is that ST-based TFET SRAM cells are the best choice for low-V_{CC} (< 300mV) operation because they consume significantly *lower energy* as well as deliver *improved performance* compared to Si CMOS, which is a direct consequence of the steep sub-threshold characteristic.

III. VARIATION STUDY

A. Device-Level Variation Model

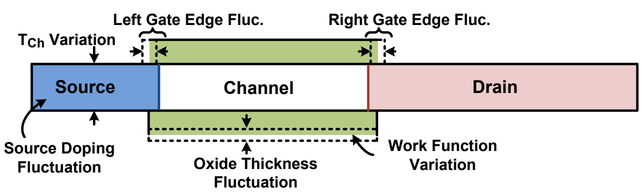


Fig. 14. Illustration of UTB HTFET variation model

It is important to consider process variations in TFETs because tunneling-based devices have an exponential dependence of the On-current on the tunneling-barrier. Any source of variation which can affect the effective tunneling-barrier width of the TFET can cause a significant On-current variation. Fig. 14 shows the variation model that is used to study the impact of variations in the structure and doping of the TFET. In this variation model, we also take the impact of quantum confinement of the UTB channel into consideration. Fig. 16 shows how the effective band-gap at the source-channel heterojunction interface changes with T_{Ch} due to quantization. These effective band-gaps were computed using a self-consistent Schrodinger-Poisson solver assuming that the channel is placed in a potential well (i.e. the oxide) [23].

In order to simplify our analysis, we assume only small fluctuations in various sources of variation considered. This allows us to express the variation in I_{On} as:

$$\delta I_{On} = \frac{\partial I_{On}}{\partial T_{Ch}} \times \delta T_{Ch} + \frac{\partial I_{On}}{\partial T_{Ox}} \times \delta T_{Ox} + \frac{\partial I_{On}}{\partial \phi_M} \times \delta \phi_M + ... \quad (1)$$

Further, we also assume that the sources of variation (δT_{Ch}, δT_{Ox}, $\delta \phi_M$, etc.) are independent, which allows us to calculate the variance of I_{On} as:

$$\sigma I_{On}^2 = (\frac{\partial I_{On}}{\partial T_{Ch}})^2 \times \sigma T_{Ch}^2 + (\frac{\partial I_{On}}{\partial T_{Ox}})^2 \times \sigma T_{Ox}^2 + ... \quad (2)$$

We simulated two-thousand Monte-Carlo samples of NMOS and HTFET devices in TCAD Sentaurus [24], assuming independent Gaussian distributions for various sources of variation. The Gaussian distributions used for the variation sources are listed in Table II. The statistical distribution of I_{On} @ V_{CC} 0.5V obtained through TCAD Monte-Carlo simulation compares well with the shape of the distribution predicted by the small-signal variation model (eq. 2), both for Si NMOS (Fig. 15(F)) and for HTFET (Fig. 15(L)), showing the validity of this approximation technique.

Our basic assumption for studying variation is that the deviations from a nominal device are always small. Based on this assumption, we can calculate the variation-coefficients ($\frac{\partial I_{On}}{\partial T_{Ch}}, \frac{\partial I_{On}}{\partial T_{Ox}}$, etc.), as shown in Fig. 15(A-E) and Fig. 15(G-K), which can then be used to study the impact of variation at the circuit-level. By using these variation coefficients, we are able to extend the Verilog-A table look-up model to study the impact of device-level variation on circuit characteristics.

TABLE II
SOURCES OF VARIATION FOR ULTRA-THIN-BODY DEVICE.

Variation N(μ,σ)	μ	**3σ**
Channel Thickness, T_{Ch} $[nm]$	5	0.5
Oxide Thickness, T_{Ox} $[nm]$	2.5	0.3
Source Doping (HTFET) cm^{-3}	4.5×10^{19}	5×10^{18}
Gate Work Func., ϕ_M (HTFET) [eV]	4.85	0.005
Gate Work Func., ϕ_M (NMOS) [eV]	4.48	0.005
Left Gate Edge [nm] (w.r.t channel center)	-20	2
Right Gate Edge [nm] (w.r.t channel center)	+20	2

B. Summary of Device-Level Variations

Variation in TFETs has been studied previously considering only two parameters - T_{Ch} and T_{Ox} [25]. However, there are also other prominent sources of variation that occur in a side-gated TFET, which are taken into consideration in our model (Fig. 14). Table II shows a summary of various small-signal fluctuations that are used to study variation-impact using our model. Further, an ultra-thin-body device can also be very sensitive to quantum effects due to structural quantization of the semiconductor channel. Fig. 16 shows the effects of quantization in TFETs, as a change in the effective band-gap of the channel material due to T_{Ch} fluctuation.

Fig. 17 (based on eq. 2) shows a break-down of the contribution of various sources of variation in TFETs, to the total Variance. It shows that the gate-source overlap can be a significant source of variation in an UTB TFET, for $V_{CC} \geq$ 500mV. The gate-source overlap in a side-gated TFET structure results in the formation of a depletion-region in the source, underneath the gate (Fig. 3). Fluctuations of the gate-edge can cause this depletion-region width to fluctuate, thus significantly increasing or decreasing the effective width over which tunneling takes place. Our model also shows that T_{Ch} fluctuation, *quantum effects included*, can also be a major sources of variation, for $V_{CC} \leq$ 300mV. Fluctuations in T_{Ch} can cause the effective tunneling-width at the source-channel heterojunction interface to change due to fluctuations in the effective band-gap (Fig. 16).

Fig. 18 compares the % $\sigma I_{On}/I_{On}$ change in CMOS and HTFETs due to variations. It can be seen that TFETs are

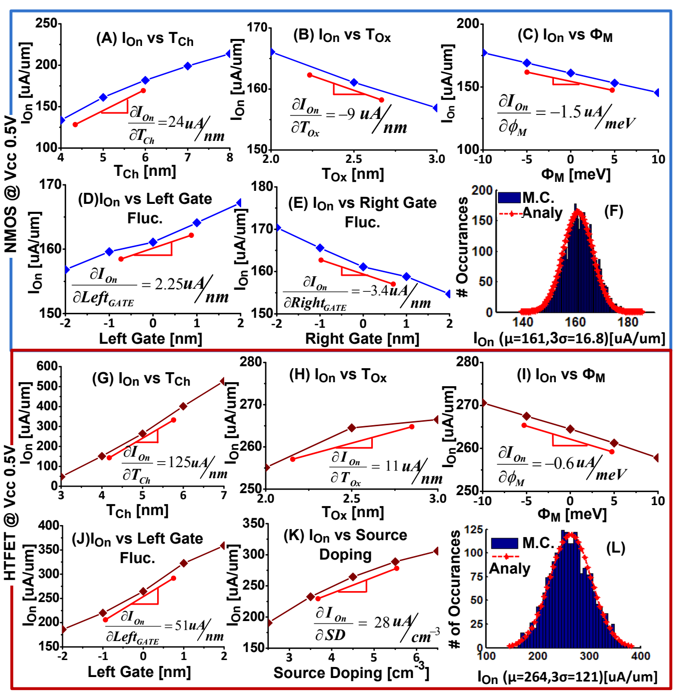


Fig. 15. (A-E) Small-signal variation coefficients for NMOS @ V_{CC} 0.5V (F) Comparison of statistical distribution of On-current obtained through Monte-Carlo vs. Analytical eq. 2 for NMOS (G-K) Small-signal variation coefficients for HTFET @ V_{CC} 0.5V (L) Comparison of statistical distribution of On-current obtained through Monte-Carlo vs. Analytical eq. 2 for HTFET

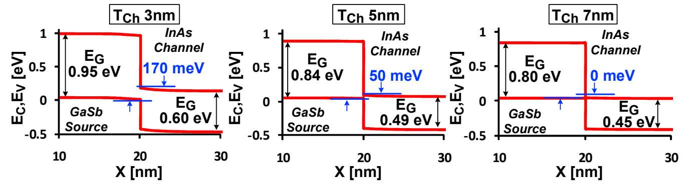


Fig. 16. Change in effective band-gap with T_{Ch} due to quantization

in general prone to variations, where-as CMOS is prone to variation only in the sub-threshold region. Thus, it is prudent to compare how variations impact the read-write SNMs in HTFET and sub-threshold CMOS SRAMs. The following sub-section summarizes the impact of variation on SRAM read-write noise margin characteristics using the model described in this section.

C. Monte-Carlo Simulation of Read-Write Noise Margins

Monte-Carlo simulation at the circuit-level for TFET-based SRAMs was shown in [17] assuming only one source of variation (T_{Ox}). We perform Monte-Carlo simulation at the SRAM circuit-level using our proposed small-signal variation model assuming all the possible sources of variation. The read-failure probability is defined as [21] :

$$P_{read-upset} = Pr\{read - SNM < kT\}$$

where, kT = 26mV at 300K. We generate one-thousand Monte-Carlo samples for various CMOS and TFET SRAM cells, and estimate the mean and sigma of the read-SNM at different voltage points. Using these estimates, we plot the the read-upset probability as a function of V_{CC} for different SRAM cells, as shown in Fig. 19. The V_{CC}-min is defined as the voltage for which $P_{read-upset}$ is $\leq 10^{-9}$. The CMOS ST-2 cell has the best read V_{CC}-min of 134mV among the CMOS SRAM cells because of its improved variation tolerance [21]. The 8T Transmission-Gate SRAM had degraded read-SNM due to the delayed saturation in TFETs (Fig. 11). In addition, TFETs are prone to variation as discussed in this section. As a result, the 8T Transmission-Gate SRAM shows a very

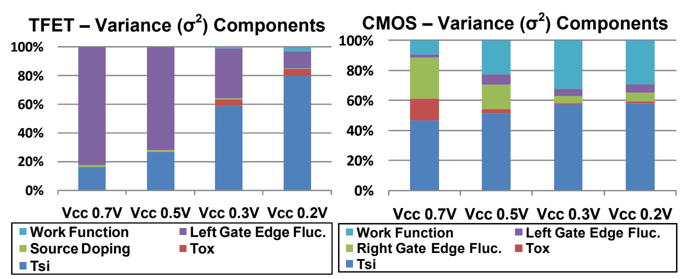
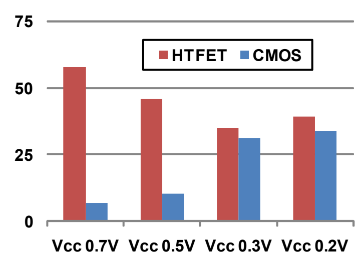


Fig. 17. Components of Variance in Si NMOS and HTFET due to various sources of variation @ different V_{CC}

Fig. 18. Comparison of % $\sigma I_{On}/I_{On}$ for Si NMOS and HT-FET @ different V_{CC}



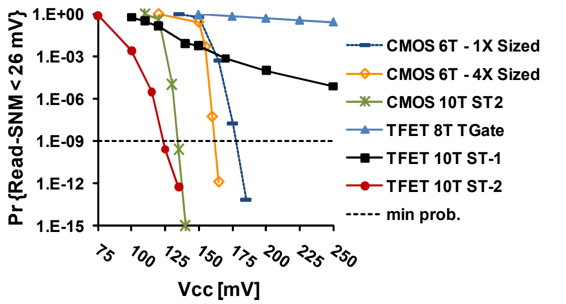


Fig. 19. Probability of read-upset for various SRAM configuration

high probability of upset even when the V_{CC} is increased, showing its unsuitability for low-V_{CC} applications. The TFET ST-1 SRAM shows improvement in the read-upset probability compared to the 8T TFET SRAM due to the use of the Schmitt-Feedback. However, due to its weak pull-up, the cell is still prone to variation-induced upsets, causing its V_{CC}-min to be large compared to the CMOS SRAM cells. In contrast, the TFET ST-2 cell, shows sufficient variation tolerance and also shows a V_{CC}-min of 124mV, showing the suitability of this cell for ultra low-V_{CC} operation.

Fig. 20(A) shows the V_{CC}-min for different cell configurations. A comparison of the dynamic energy ($C_{gg} \times V^2$) and the leakage power consumption of different SRAM cells operating at their respective V_{CC}-min is also shown in Fig. 20(B)&(C) (normalized to the TFET ST-2 cell). The TFET ST-2 cell at its V_{CC}-min provides 1.2x lower dynamic energy and 13x lower leakage power consumption compared to the CMOS ST-2 cell operating at its V_{CC}-min. At the same time, as explained in section II-D, the TFET ST-2 cell has far better performance than the sub-threshold CMOS ST-2 cell, due to the better drive currents of HTFETs in the low V_{CC} regime.

(A)

SRAM Cell	CMOS 6T - 1X Sized	CMOS 6T - 4X Sized	CMOS 10T ST2	TFET 10T ST1	TFET 10T ST2
Vcc-min [mV]	177	162	134	423	124

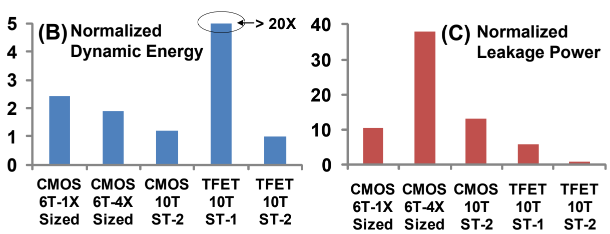


Fig. 20. (A) V_{CC}-min for different SRAM cell configurations. Comparison of (B) Dynamic Energy and (C) Leakage Power for different SRAM cells operating at their V_{CC}-min. The energy and leakage numbers are normalized to TFET ST-2 cell.

IV. Conclusions

In this paper, we consider the characteristics of emerging interband tunneling transistors and analyze the impact of the unique properties of these device on the stability characteristics of SRAM cells. We cover a wide design space of SRAM cells, and show that better read/write noise margin characteristics can be obtained by using higher transistor count (8T and 10T) SRAMs. Further, we show that Schmitt-Trigger-based (10T) TFET SRAM cells operating at low-V_{CC} are a very attractive alternative compared to sub-threshold CMOS, both from an energy as well as a performance perspective. We propose a small-signal variation model to analyze the impact of variation on the stability characteristics of SRAM cells, and show that Schmitt-Trigger-based TFET SRAM cells have sufficient variation tolerance to allow ultra-low V_{CC}-min operation.

V. Acknowledgments

This work was supported in part by NSF grants 1028807, 0903432, 0829926, the Semiconductor Research Corporations Nanoelectronic's Research Initiative and National Institute of Standards & Technology through the Midwest Institute for Nanoelectronics Discovery (MIND) and Intel Academic Research Office (ARO) program on Post-CMOS Circuits.

References

[1] A. Chandrakasan and R. Brodersen, *Low Power Digital CMOS Design*. Kluwer Academic Publishers, 1995.

[2] K. Roy and S. Prasad, *Low-Power CMOS VLSI Circuit Design*. John Wiley and sons, Inc., 2000.

[3] Y. Nakagome, M. Horiguchi, T. Kawahara, and K. Itoh, "Review and future prospects of low-voltage RAM circuits," *IBM Journal of Research and Development*, 2003.

[4] B. Zhai, D. Blaauw, D. Sylvester, and S. Hanson, "A Sub-200mV 6T SRAM in 0.13 um CMOS," in *IEEE ISSCC*, 2007.

[5] N. Verma and A. Chandrakasan, "A 65nm 8t sub-vt sram employing sense-amplifier redundancy," in *IEEE ISSCC*, 2007.

[6] B. Calhoun and A. Chandrakasan, "A 256kb Sub-threshold SRAM in 65nm CMOS," in *IEEE ISSCC*, 2006.

[7] J. P. Kulkarni, K. Kim, and K. Roy, "A 160 mv, fully differential, robust schmitt trigger based sub-threshold sram," in *Proc. of ISLPED*, 2007.

[8] R. Gandhi, Z. Chen, N. Singh, K. Banerjee, and S. Lee, "Vertical Si-Nanowire n -Type Tunneling FETs With Low Subthreshold Swing ($\leq$ 50mV/decade) at Room Temperature," *IEEE Elec. Dev. Letters*, 2011.

[9] A. C. Ford, C. W. Yeung, S. Chuang, H. S. Kim, E. Plis, S. Krishna, C. Hu, and A. Javey, "Ultrathin body InAs tunneling field-effect transistors on Si substrates ," *Applied Physics Letters*, 2011.

[10] D. Mohata, R. Bijesh, V. Saripalli, T. Mayer, and D. S., "Self-aligned Gate NanoPillar In0.53Ga0.47As Vertical Tunnel Transistor ," in *Device Research Conference (DRC)*, 2011.

[11] I. Ok, K. Akarvardar, S. Lin, M. Baykan, C. D. Young, P. Hung, M. P. Rodgers, S. Bennett, H. O. Stamper, D. L. Franca, C. Nadeau, J. P. Hobbs, P. Kirsch, M. P., and R. Jammy, "Strained sige and si finfets for high perf. logic with sige/si stack on soi," in *IEDM*, 2009.

[12] M. Luisier and G. Klimeck, "Performance comparisons of tunneling field-effect transistors made of insb, carbon, and gasb-inas broken gap heterostructures," in *IEEE IEDM*, 2009.

[13] V. Saripalli, D. Mohata, S. Mookerjea, S. Datta, and V. Narayanan, "Low power loadless 4t sram cell based on degenerately doped source (dds) in0.53ga0.47as tunnel fets," in *Device Research Conference*, 2010.

[14] D. Kim, Y. Lee, J. Cai, I. Lauer, L. Chang, S. J. Koester, D. Sylvester, and D. Blaauw, "Low Power Circuit Design Based on Heterojunction Tunneling Transistors (HETTs)," in *ACM/IEEE ISLPED*, 2009.

[15] J. Singh, K. Ramakrishnan, S. Mookerjea, S. Datta, N. Vijaykrishnan, and D. Pradhan, "A Novel Si-Tunnel FET Based SRAM Design for Ultra Low-Power 0.3V VDD Applications," in *Proc. 15th Asia and South Pacific Design Automation Conf. (ASP-DAC)*, 2010.

[16] V. Saripalli, A. K. Mishra, S. Datta, and V. Narayanan, "An Energy-Efficient Heterogeneous CMP based on Hybrid TFET-CMOS Cores," in *Proc. Design Automation Conference*, 2011.

[17] X. Yang and K. Mohanram, "Robust 6T Si Tunneling Transistor SRAM Design," in *Proc. Design Automation and Test in Europe*, 2011.

[18] H. Noguchi, S. Okumura, Y. Iguchi, H. Fujiwara, Y. Morita, K. Nii, H. Kawaguchi, and M. Yoshimoto, "Which is the best dual-port sram in 45-nm process technology ; 8t, 10t single end, and 10t differential," in *IEEE Intl. Conf. on Integrated Circuit Design and Technology*, 2008.

[19] L. Chang, D. Fried, J. Hergenrother, J. Sleight, R. Dennard, R. Montoye, L. Sekaric, S. McNab, A. Topol, C. Adams, K. Guarini, and W. Haensch, "Stable sram cell design for the 32 nm node and beyond," in *Symposium on VLSI Technology.*, 2005.

[20] J. Kulkarni, K. Kim, S. Park, and K. Roy, "Process variation tolerant sram array for ultra low voltage applications," in *ACM/IEEE DAC*, 2008.

[21] J. P. Kulkarni and K. Roy, "Ultralow-voltage process-variation-tolerant schmitt-trigger-based sram design," *IEEE Tran. VLSI*, 2011.

[22] S. Thoziyoor, N. Muralimanohar, J. H. Ahn, and N. P. Jouppi, "CACTI 5.1," HP Labs, Tech. Rep., 2008.

[23] *nextnano^3 - next generation 3D nanodevice simulator (http://www.nextnano.de/nextnano3/)*, nextnano, 2008.

[24] *TCAD Sentaurus Device Manual, Release: C-2009.06*, Synopsys, 2009.

[25] S. Saurabh and M. Kumar, "Estimation and Compensation of Process-Induced Variations in Nanoscale Tunnel Field-Effect Transistors for Improved Reliability," *IEEE TDMR*, 2010.