A Novel Si-Tunnel FET based SRAM Design for Ultra Low-Power 0.3V V_{DD} Applications

J. Singh†, K. Ramakrishnan‡, S. Mookerjea‡, S. Datta‡, N. Vijaykrishnan‡, D. Pradhan†
†Department of Computer Science, University of Bristol, UK.
‡Department of Electrical Engineering, The Pennsylvania State University, University Park, USA - 16802.

Abstract— Steep sub-threshold transistors are promising candidates to replace the traditional MOSFETs for sub-threshold leakage reduction. In this paper, we explore the use of Inter-Band Tunnel Field Effect Transistors (TFETs) in SRAMs at ultra low supply voltages. The uni-directional current conducting TFETs limit the viability of 6T SRAM cells. To overcome this limitation, 7T SRAM designs were proposed earlier at the cost of extra silicon area. In this paper, we propose a novel 6T SRAM design using Si-TFETs for reliable operation with low leakage at ultra low voltages. We also demonstrate that a functional 6T TFET SRAM design with comparable stability margins and faster performances at low voltages can be realized using proposed design when compared with the 7T TFET SRAM cell. We achieve a leakage reduction improvement of 700X and 1600X over traditional CMOS SRAM designs at V_{DD} of 0.3V and 0.5V respectively which makes it suitable for use at ultra-low power applications.

I. INTRODUCTION

Continued miniaturization of the silicon CMOS transistor technology has resulted in an unprecedented increase in single-core and multi-core performance of modern-day microprocessors. However, the exponentially rising transistor count has also increased the overall power consumption making performance per watt of energy consumption the key figure-of-merit for today's high-performance microprocessors. Today, energy efficiency serves as the central tenet of high performance microprocessor technology at the system architecture level as well as the transistor level ushering in the era of energy efficient nanoelectronics. Aggressive supply voltage scaling while maintaining the transistor performance is a direct approach towards reducing the energy consumption since it reduces the dynamic power quadratically and the leakage power linearly. In MOSFETs, the OFF-state leakage current (I_{OFF}) increases exponentially with reduction of threshold voltage and hence there is a fundamental limit to the scaling of the MOSFET threshold voltage and hence the supply voltage. Scaling supply voltage limits the ON current (I_{ON}) and the $I_{ON}-I_{OFF}$ ratio. This fundamental limit to threshold voltage scaling arises from MOSFETs 60 mV/decade subthreshold swing at room temperature.

Leakage power consumption in SRAMs have been a major concern in caches since they occupy more than 50% of the processor chip area. Lower threshold voltages increase the sub-threshold current exponentially and ultra thin gate oxides cause a huge increase in gate current. Various methods such as multiple threshold voltages and increased gate oxide thicknesses have been explored to reduce leakage in SRAMs. Adaptive body biasing techniques have also been explored to reduce leakage.

Recently, leakage reduction using steep subthreshold transistors has gained great attention. A steep sub-threshold transistor allows us to operate at very low threshold voltages with ultra low leakage and low supply voltages (V_{DD}). Inter-band Tunnel Transistor also called as Tunnel Field Effect Transistor (TFETs) has shown to be a promising steep subthreshold transistor which works on the principle of inter-band tunneling [1]. TFETs have shown to be extremely power efficient in [2] for logic circuit applications. The authors in [2] also point out the problem of uni-directionality in TFETs and it's detrimental impact on 6T TFET SRAMs. To overcome this limitation, they have proposed a 7T SRAM with extra read port to achieve higher stability margins. In this paper, we propose a novel 6T SRAM cell to overcome the problem of uni-directionality and achieve tolerable stability margins and performance at the same area of a CMOS design.

The remainder of the paper is organized as follows. Section II explains the device physics behind the operation of a TFET and our models used for circuit simulation. Various existing SRAM designs using CMOS and TFETs are explained in section III. Section IV explains our proposed 6T TFET design. The comparison of metrics and the results obtained are presented in section V. Section VI concludes the paper.

II. TUNNEL FIELD EFFECT TRANSISTORS (TFETS)

In the recent times, inter-band Tunnel Field Effect Transistors (TFETs) have been extensively investigated [1][3][4][5] due to its potential for sub-KT/q subthreshold slope device operation and thus enabling supply voltage reduction for low power logic applications. Figure 1 shows our optimized double gate device structure of a Si based N-channel and P-channel TFET. A N-type TFET consists of a p+ source, intrinsic (i) channel and a n+ drain and the P-type TFET has n+ source, intrinsic (i) channel and p+ drain regions. The source and drain regions are heavily doped regions with the channel region being intrinsic. The gate work function of N-channel TFET is modified suitably to obtain a P-channel TFET.

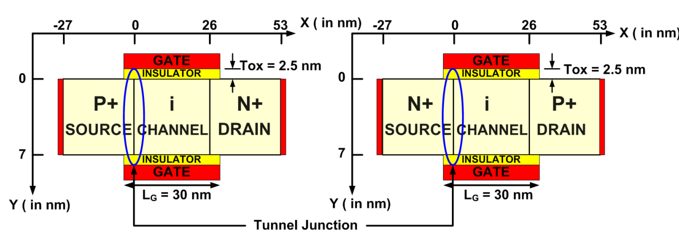

Fig. 1. Double gate N-channel and P-channel Si-TFET

Figure 2 shows the band-diagram of a N-type TFET during the ON and OFF state. In the OFF state ($V_{GS}=0V, V_{DS}=1V$), the conduction in MOSFET is limited by the source side p-n junction barrier which prevents the thermionic emission of carriers. In the ON state ($V_{GS}=1V, V_{DS}=1V$), the source barrier is negligible enabling over the barrier thermionic emission. In contrast, TFETs operate by tunneling of carriers from the valence band in the source to the conduction band in the channel. In the OFF state ($V_{GS}=0V, V_{DS}=1V$), the transmission probability is low due to the

This research was funded in part by Intel Corporation, NRI and NSF grant no. 0829926.

TABLE I
NOMINAL SI-TFET PARAMETERS

Gate Length, L_G	30 nm
Oxide thickness, T_{OX}	2.5 nm
Gate di-electric constant, ξ	21 (HfO_2)
Body thickness, T_{Si}	7 nm
Gate overlap	2 nm
Source/Drain Doping, $N_{S/D}$	10^{20} cm^{-3}
Channel Doping, N_{Ch}	10^{15} cm^{-3}
Gate work-function	3.9 eV (N-type) 5.2 eV (P-Type)

thick depletion region associated with the source to channel tunnel junction resulting in very low OFF currents. With the application of the gate voltage ($V_{GS} = 1V, V_{DS} = 1V$), the depletion region shrinks and the carriers tunnel through the barrier. Since the TFET ON current is limited by the inter-band quantum mechanical tunneling compared to thermionic emission over the barrier the ON current in silicon TFETs is much lower than MOSFETs. The reverse biased leakage current under the condition of OFF state ($V_{GS} = 0V, V_{DS} = 1V$) yields extremely low OFF current in the order of pico-femto amperes.

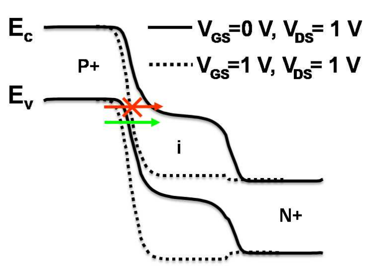

Fig. 2. Band diagram of a Si-NTFET under ON and OFF conditions

Table I shows the nominal parameters of our device structure. A non-local tunneling model [6] is used for the simulation of tunnel current which accounts for the actual spatial charge transfer across the tunnel barrier by considering the actual potential profile along the entire path connected by tunneling. The inter-band tunneling current in the TFET depends on the potential profile along the entire path between two points connected by tunneling. In contrast to the local tunneling models commonly used [7][8], we use a non-local tunneling model [9] which reflects the real space carrier transport through the barrier taking into account the potential profile along the entire tunneling path. Band edge tunneling masses of m_c=0.5*m_0 and m_v=0.65*m_0 (where m_0 is electron rest mass) for silicon are used to calculate the local imaginary wave numbers within the forbidden gap. Kane's two band model is then used to calculate the tunneling probability. The results presented here are obtained through drift-diffusion simulation where the Poisson and carrier continuity equations are solved self consistently. The inter-band tunneling component is added to the carrier continuity equation as a generation-recombination (G-R) term. The G-R term contains adjustable scaling factors g_c and g_v kept at value equal to 0.1 and 0.4 respectively for Si which set the effective Richardson constant. We also obtained a excellent fit of our nonlocal tunneling model with the experimental data from Fair & Wivell [10] for a reverse biased Si zener diode.

Figure 3 shows the $I_D V_G$ characteristics of a Si NTFET and PTFET for V_{DS} = 1V. We obtain a I_{DSAT} = 120 μA/μm and the PTFET characteristics are matched to it. The reverse biased leakage can be set to the order of pico-femto amperes by modifying the gate work function. We assume that the gate leakage is negligible due to the use of high-k dielectrics. We have also denoted the symbols for NTFET and PTFET in figure 3. The source side tunneling barriers are represented by a bracket symbol and the current directions are also shown.

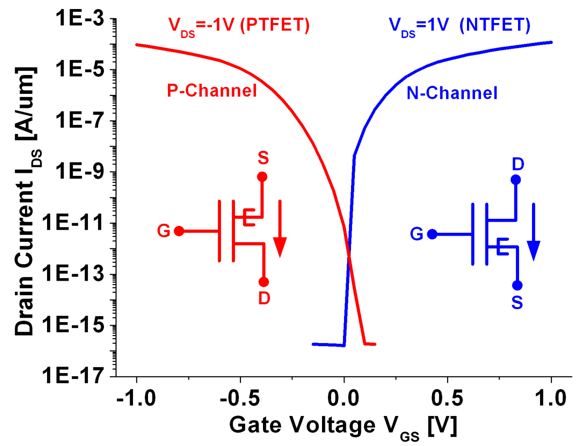

Fig. 3. $I_D V_G$ characteristics of a Si NTFET and PTFET

Figure 4 shows the $I_D V_D$ characteristics of the same device. The device exhibits expected characteristics due to tunneling during positive V_{DS} (reverse bias conditions) while I_{DS} increases significantly for two conditions when V_{DS} is negative (forward bias). When V_{DS} is $\sim$-1V, there is a significant I_{DS} irrespective of the value of V_{GS}. Significant current conduction is also observed when V_{DS} is slightly negative and V_{GS} is positive. This is due to electrons tunneling from the conduction band of intrinsic 'i' region to the valence band of p+ source region.

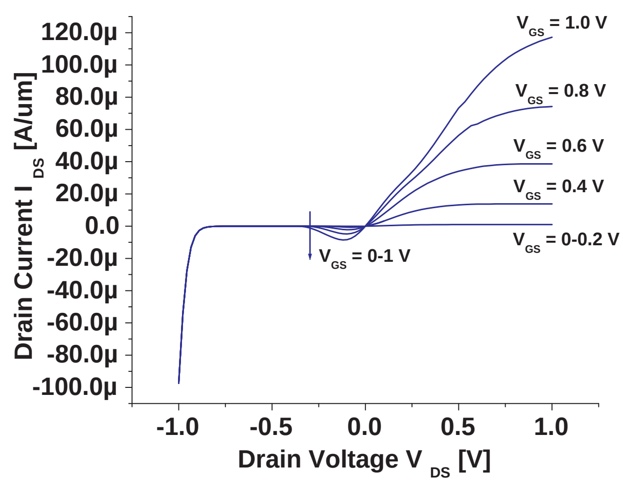

Fig. 4. $I_D V_D$ characteristics of a Si-NTFET

Since analytical models for TFETs are not available, we have built a look-up table based model using Verilog-A for circuit simulation. The Verilog-A module is then used as instances for circuit simulation in Cadence Spectre. This efficient and accurate way of modeling is well suited for the emerging devices for which compact or SPICE models are not available [11]. In this model, I-V and C-V characteristics of the TFET devices extracted using Sentaurus [6] TCAD simulations and stored as a two dimension look-up tables. We also observed enhanced miller capacitance (High C_{GD}) values for our devices and their effect was observed to be negligible for circuits with high electrical effort as explained in [2] and [12].

III. SRAM DESIGNS

Figure 5 shows various SRAM designs. Figure 5 (a) is the standard 6T SRAM cell and (b) and (c) show the 6T TFET SRAM

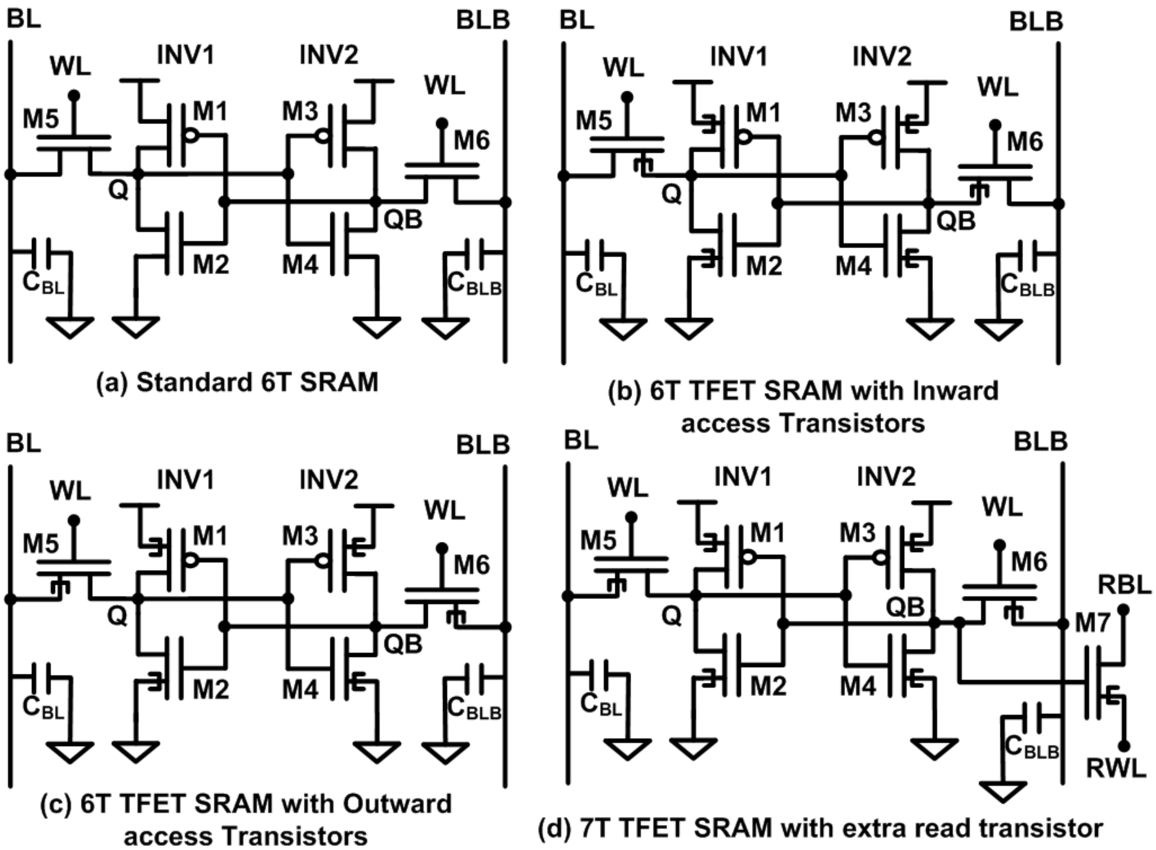

Fig. 5. SRAM designs

design configuration with inward and outward access transistors.

The read noise margin (RNM) of a SRAM design is estimated graphically as the length of a side of the largest square that can be embedded inside the lobes of a butterfly curve. The Write Noise Margin (WNM) is measured through the write trip point defined as the difference between V_{DD} and the minimum bit-line voltage required to flip the data storage nodes Q or QB. Figure 6 and figure 7 show an example of RNM measurement, read failure and WNM for a 6T TFET inward access transistors configuration shown in figure 5 (b). The 6T TFET SRAM design suffers from severe noise margin deficiencies due to the uni-directionality issues as shown in figure 8 and figure 9. Figure 8 and 9 show the read and write noise margins (RNM and WNM) for 6T TFET SRAM with inward and outward access transistor configurations. As mentioned above, we observe in figure 8 that the WNM reduces to 0 for cell ratio (β=$W_{Pull-Down}/W_{Access}$) > 0.3 while RNM is 0 for $\beta < 0.3$. Similarly, figure 9 shows that RNM starts to increase only for Pull-up ratios ($W_{Pull-Up}/W_{Access}$) greater than 2 while WNM reduces to 0 for the same. Thus, a 6T TFET SRAM with acceptable stability margins is not possible. In order to have enough read and write margins, a 7T TFET SRAM configuration with outward access transistors was proposed in [2] as shown in figure 5 (d). In this design, outward access transistor configuration is used to obtain the adequate write margin while the read margin is improved by providing a read-buffer with an extra transistor and separate read bit-line and word-line.

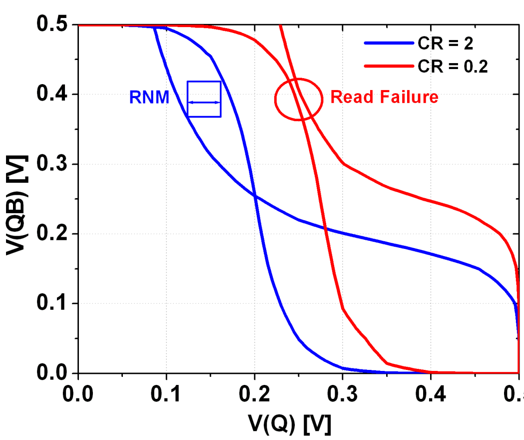

Fig. 6. Measurement of RNM and Read failure

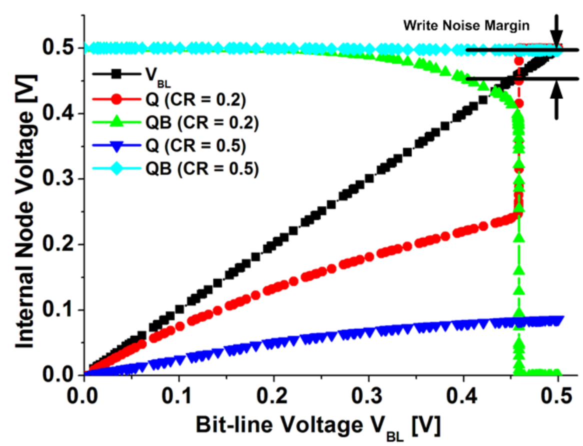

Fig. 7. Measurement of Write Trip Point (WTP) and WNM

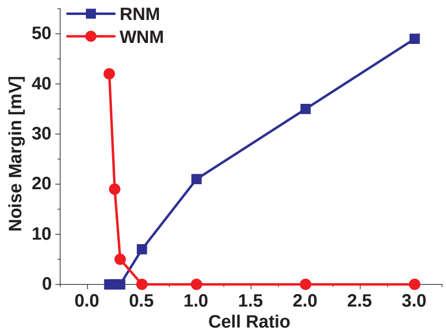

Fig. 8. Noise margins for 6T TFET SRAM with inward access transistors at V_{DD}=0.5V

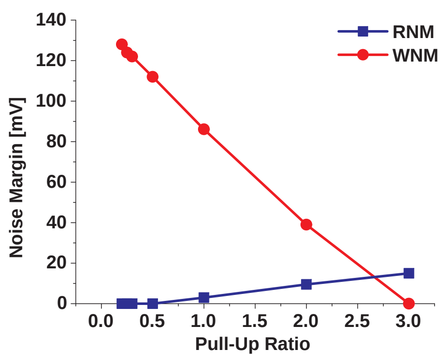

Fig. 9. Noise margins for 6T TFET SRAM with outward access transistors at V_{DD}=0.5V

IV. Proposed 6T TFET SRAM Design

As shown in the previous section, a practical 6T TFET SRAM design is not feasible. We have proposed a novel 6T TFET SRAM, keeping minimum number of devices and preserving the adequate RNM and WNM as shown in figure 10. Our proposed design consists of cross coupled inverters (INV1 and INV2) with the bit-lines BL and BLB connected to node Q through the access transistors M5 and M6 (Note that both the access transistors are connected to the same node Q). It is a design strategy to provide a virtual ground to INV1 while writing either '1' or '0' to node Q. This virtual grounding helps in improving the WNM, by decoupling (or weakening of the re-generative action) of the cross-coupled inverters.

A. Read Operation

We use differential read operation in our proposed design. Both the bit-lines (BL and BLB) are pre-charged to V_{DD} and then the WL is asserted to '1'. If the bit stored at node Q is a '0', then BL discharges from V_{DD} and the sense amplifier is triggered. Otherwise, the bit-line BL remains pre-charged at V_{DD} unperturbed. Figure 11 (a) shows the current path during a read operation in our proposed design. We chose inward access transistor for read operation in our design since this configuration allows us to have a higher RNM than outward access transistor configuration as shown in figure 8 and figure 9 while our design strategy significantly improves the WNM as explained in the later sections.

B. Write Operation

The write operation in our design is done through one of our access transistors depending on the bit to be written onto the SRAM cell. To write a '1' onto Q, we charge the bit-line BL to V_{DD} and

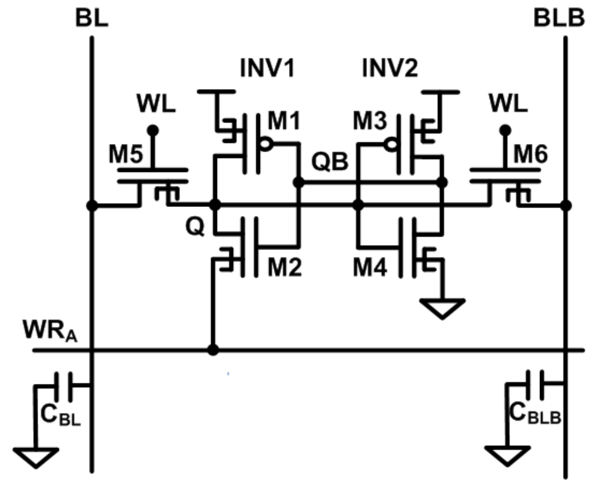

Fig. 10. Proposed 6T TFET SRAM

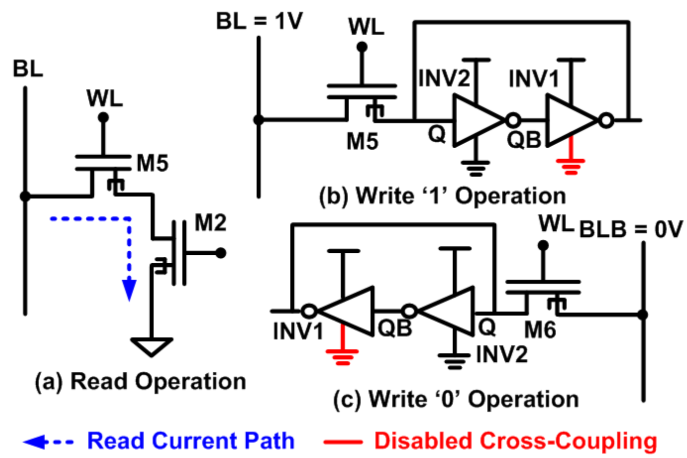

Fig. 11. Read and write operation of the proposed design

then enable WL = '1' for access transistor M5. The write enable line WR_A is also raised simultaneously to weaken the inverter INV1 and disable the cross-coupling between the two inverters. Once Q settles to a '1' and QB reaches '0', the WR_A line is connected to ground and the cross coupling is enabled. Figure 11(b) shows the write '1' operation.

If the node Q stores a '1' and we intend to write a '0', the bit-line BLB is pulled low to 0V. Simultaneously, the write enable line WR_A is also raised simultaneously to virtual ground and the word-line WL is asserted. This breaks the cross-coupling and Q is drained to ground through the access transistor M6. Once the node QB settles to a '1', the cross-coupling is enabled. Figure 11(c) shows the write '0' operation.

In order to demonstrate a successful read and write operation, we have simulated the RNM and WNM of the proposed 6T TFET SRAM for different cell ratios (β) at V_{DD}=0.5V when the pull up ratio is kept at minimum. In figure 12, the RNM at half pre-charged bitline is much better than fully pre-charged bit-line. For $\beta > 2$, there is no significant improvement in the RNM while a slight degradation in the WNM is observed, also using the higher cell ratio will increase the cell area. Hence, in all the simulations we have used cell ratio (β) of 2 unless specified. Due to the asymmetric nature of the proposed design, writing a '1' is more difficult than writing '0', hence, we have only measured the WNM for writing a '1'.

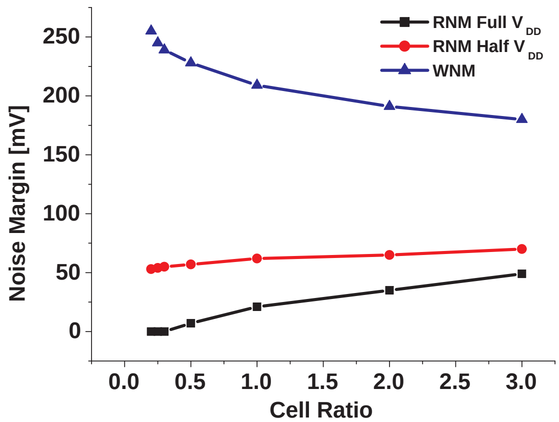

Fig. 12. Noise margins for modified 6T TFET SRAM with inward access transistors at V_{DD}=0.5V

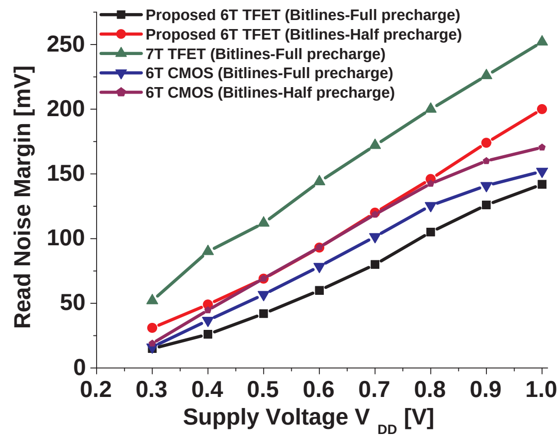

Fig. 13. Read Noise Margins for different designs at various supply voltages V_{DD}

V. Results

Stability, performance and power of a SRAM design are the three key design metrics in the nanometer regime. For comparison, we have used the existing 6T CMOS SRAM and 7T TFET SRAM design. We use 32nm Predictive Technology Models (PTM) [13] for 6T CMOS, while the 6T and 7T TFET SRAMs are simulated with the same device as the explained in table I. In this section, we have compared these designs for all the design metrics.

A. Stability

An adequate read and write stability is highly desirable for a successful realization of a SRAM cell. The RNM and WNM are the widely used metrics for stability analysis of a SRAM cell. Figure 13 shows the RNM of different designs. The proposed 6T TFET SRAM and 6T CMOS have bit-lines BL and BLB pre-charged to full V_{DD} and half V_{DD}. The 7T TFET SRAM cell shows the highest RNM, because of the isolated read-buffer which yields the RNM equivalent to Hold Static Noise Margin (SNM). The isolated read buffer concept has been widely explored in CMOS SRAM designs to improve RNMs. However, the proposed 6T TFET with fully pre-charged bit-line has the lowest RNM. This is because of the single access transistor which conducts during the read operation and rises the internal node (Q) voltage to a higher value than a 6T CMOS SRAM (while the other access transistor does not assist because of its uni-directionality).

The RNM of the proposed 6T TFET with half-swing is much better than the 6T CMOS with half and full pre-charged bit-lines. In 6T CMOS SRAM, half pre-charged bit-lines are not as effective as 6T TFET SRAM. This is due to the symmetric nature of SRAM where one of the bit-lines connected to a node (Q or QB) via access devices storing a V_{DD} is also pre-charged to half V_{DD}. This scenario is not effective in holding that node at V_{DD} as compared to pre-charging to full V_{DD} due to conduction from the node to bit-line in the former case. However in our proposed 6T TFET design, M6 in figure 10 does not conduct in the reverse direction and this contributes to higher RNM at half pre-charged V_{DD}. At V_{DD}=0.3V, we observe a 63% improvement in RNM over a 6T CMOS while it is 59% lesser than a 7T TFET. The advantage of 7T TFET purely comes from the extra transistor used as a read port.

Figure 14 shows the WNM of SRAM cell designs for different V_{DD}. The WNM of the proposed 6T TFET SRAM design is higher than its counterpart designs due to weakening of the inverter which enables a faster write. At V_{DD}=0.3V, we observe a 46% and 32%improvement in WNM over 6T CMOS and 7T TFET respectively.

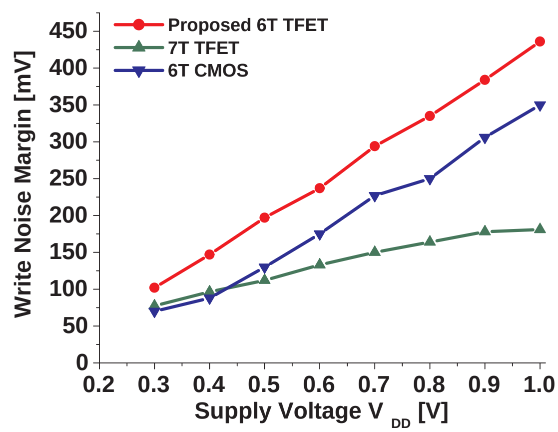

Fig. 14. Write Noise Margins for different designs at various supply voltages V_{DD}

B. Performance

Read and write delays are the metrics used to compare the performance of different SRAM designs. In 6T CMOS and 6T TFET read delay is defined as the time delay between 50% of word line (WL) activation to 10% of pre-charged voltage difference between the bit lines. In 7T and 8T SRAM designs, bit-line sensing is done using CMOS logic gates and not by using differential sense amps [14] [15]. So, for the 7T TFET design, read delay is measured between 50% of word line (WL) activation to 50% of pre-charged bit line voltage. Figure 15 shows the read delay of different SRAM designs. We observe that CMOS performs better than TFETs in the entire voltage range due to it's high drive current. At V_{DD}=0.3V, 6T CMOS design has a better read delay than 6T TFET and 7T TFET by 40% and 58% respectively. However, this problem can be solved in TFETs by moving to lower band-gap and low effective mass materials such as Indium Arsenide (InAs) which have a higher tunneling rate through the barrier and higher drive current (I_{ON}) of $\sim$85 μA/μm for V_{DD}=0.25V [5].

The write delay is defined as the time between the 50% activation of the word line (WL) to when the internal Q is flipped to 90% of its

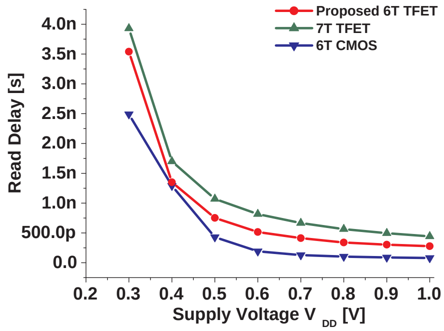

Fig. 15. Read delay for various supply voltages V_{DD}

full swing. At lower voltages, write delay of the proposed 6T TFET SRAM design is significantly less than the 6T CMOS and 7T TFET SRAM designs as shown in figure 16. This is due to the simple fact of breaking the cross coupling which enables a faster write speed than other designs. The write delays for 6T CMOS and 7T TFETs are 8.1X and 4.7X times higher than the proposed 6T TFET design at V_{DD}=0.3V.

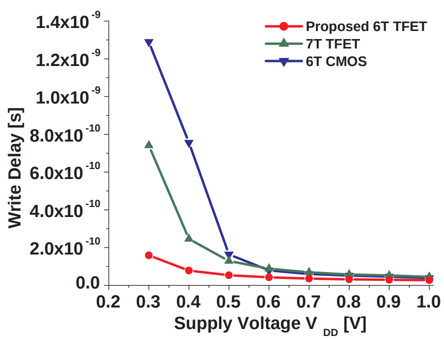

Fig. 16. Write delay for various supply voltages V_{DD}

C. Leakage Power

Due to the inherent nature of TFETs, the OFF state leakage current of a TFET is orders of magnitude lower than CMOS. Thus, we see a huge improvement in terms of leakage reduction. Figure 17 shows the standby leakage/cell of various SRAM designs. Both 6T and 7T TFET has equal leakage power due to the presence of the same leakage paths. We obtain a 700X and 1600X improvement in leakage reduction over CMOS designs at 0.3V and 0.5V V_{DD}. This shows that TFETs are a potential replacement candidate for CMOS transistors at low voltage and low power applications.

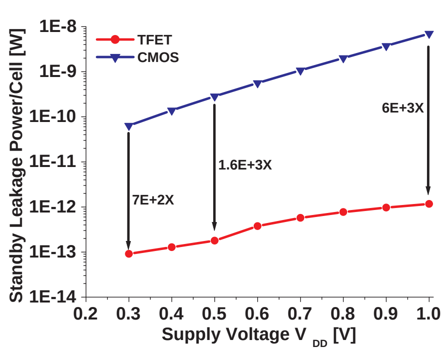

Fig. 17. Standby leakage/Cell for CMOS and TFET SRAM designs

D. Area

The proposed 6T TFET SRAM cell is not expected to have an area increase while a 7T TFET SRAM is bound to have an increase of around 15% [2]. Thus, a design with comparable margins and better performances can be obtained using a 6T instead of 7T.

VI. Conclusions

In this paper, we have proposed a novel 6T Si-TFET based SRAM design to enable ultra-low voltage and low power applications. We show that our proposed 6T Si-TFET SRAM cell has comparable margins and better performances than the 7T TFET SRAM design. We also obtain a significant improvement in leakage reduction over the entire voltage range and find TFETs a suitable candidate for replacement for CMOS in SRAM designs at ultra low voltages such as 0.3V. Our design has superior margins and performance except for read delay than CMOS due to the low drive current. Our future work will be to explore the use of lower band gap materials such as Indium Arsenide (InAs) for better performance.

References

[1] W. M. Reddick and G. A. J. Amaratunga, "Silicon surface tunnel transistor," *Applied Physics Letters*, vol. 67, no. 4, pp. 494–496, 1995. [Online]. Available: http://link.aip.org/link/?APL/67/494/1

[2] 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 *ISLPED '09: Proceedings of the 14th ACM/IEEE international symposium on Low power electronics and design*. New York, NY, USA: ACM, 2009, pp. 219–224.

[3] K. Bhuwalka, S. Sedlmaier, A. Ludsteck, C. Tolksdorf, J. Schulze, and I. Eisele, "Vertical tunnel field-effect transistor," *Electron Devices, IEEE Transactions on*, vol. 51, no. 2, pp. 279–282, Feb. 2004.

[4] P. F. Wang, "Complementary tunneling fets (ctfet) in cmos technology," Ph.D. dissertation, TU Munchen, Munich, Germany, May 2003.

[5] S. Mookerjea and S. Datta, "Comparative study of si, ge and inas based steep subthreshold slope tunnel transistors for 0.25v supply voltage logic applications," in *Device Research Conference, 2008*, June 2008, pp. 47–48.

[6] *TCAD Sentaurus Device Manual, Release: Z-2007.03*, Synopsys, 2003.

[7] G. Hurkx, D. Klaassen, and M. Knuvers, "A new recombination model for device simulation including tunneling," *Electron Devices, IEEE Transactions on*, vol. 39, no. 2, pp. 331–338, Feb 1992.

[8] "Rigorous theory and simplified model of the band-to-band tunneling in silicon," *Solid-State Electronics*, vol. 36, no. 1, pp. 19 – 34, 1993.

[9] M. Ieong, P. Solomon, S. Laux, H.-S. Wong, and D. Chidambarrao, "Comparison of raised and schottky source/drain mosfets using a novel tunneling contact model," in *Electron Devices Meeting, 1998. IEDM '98 Technical Digest., International*, Dec 1998, pp. 733–736.

[10] R. Fair and H. Wivell, "Zener and avalanche breakdown in as-implanted low-voltage si n-p junctions," *Electron Devices, IEEE Transactions on*, vol. 23, no. 5, pp. 512–518, May 1976.

[11] J. Lin, E. Toh, C. Shen, D. Sylvester, C. Heng, G. Samudra, and Y. Yeo, "Compact hspice model for imos device," *Electronics Letters*, vol. 44, no. 2, pp. 91–92, 17 2008.

[12] S. Mookerjea, R. Krishnan, S. Datta, and V. Narayanan, "On enhanced miller capacitance effect in inter-band tunnel transistors," *Electron Device Letters (In Press), IEEE*.

[13] W. Zhao and Y. Cao, "New generation of predictive technology model for sub-45nm design exploration," in *ISQED '06: Proceedings of the 7th International Symposium on Quality Electronic Design*. Washington, DC, USA: IEEE Computer Society, 2006, pp. 585–590.

[14] G. K. Chen, D. Blaauw, T. Mudge, D. Sylvester, and N. S. Kim, "Yield-driven near-threshold sram design," in *ICCAD '07: Proceedings of the 2007 IEEE/ACM international conference on Computer-aided design*. Piscataway, NJ, USA: IEEE Press, 2007, pp. 660–666.

[15] M. Meterelliyoz, J. P. Kulkarni, and K. Roy, "Thermal analysis of 8-t sram for nano-scaled technologies," in *ISLPED '08: Proceeding of the thirteenth international symposium on Low power electronics and design*. New York, NY, USA: ACM, 2008, pp. 123–128.