

Ultra Low Power Circuit Design using Tunnel FETs

Ravindhiran Mukundrajan*, Matthew Cotter*, Vinay Saripalli*, Mary Jane Irwin*, Suman Datta[†] and Vijaykrishnan Narayanan*

*Department of Computer Science & Engineering

[†]Department of Electrical Engineering

The Pennsylvania State University, University Park, PA 16802

Email: {ravi, mjc324, vxs924, mji, vijay}@cse.psu.edu & sdatta@engr.psu.edu

Abstract—The proliferation of ubiquitous and mobile computing systems has created a new segment in the design space where energy efficiency is the most critical design parameter. With the end user expecting more functionality from these types of systems, there is a pressing need to evaluate emerging technologies that can overcome the limitations of CMOS. This work evaluates the potential of one such prospective MOSFET replacement device - the Tunnel FET (TFET). Novel circuit designs are presented to overcome unique design challenges posed by TFETs. The impacts of the proposed design techniques are characterized and a sparse prefix tree adder employing the proposed designs is presented. A considerable improvement in delay and significant reduction in energy is observed due to the combined impact of circuit and technology co-exploration.

Keywords-Tunnel FET; Low-Power; Energy-efficiency;

I. INTRODUCTION

Ubiquitous computing systems have seen an upward growth trajectory recently and market indicators predict that this trend will continue for the foreseeable future. Typically, these systems are battery powered and hence the primary design focus in this space is maximizing energy efficiency. Many techniques have been presented over the past two decades to improve the energy efficiency [1] at the circuit, architecture and system levels of abstraction. These techniques are predominantly improved variations and combinations of slower, simpler, dedicated, parallel and adaptive systems. Despite all these advances, overcoming the power wall still remains a major design challenge.

CMOS technology has been an ideal framework to realize digital designs over the past four decades due to its desirable performance, power, cost and reliability characteristics. However, as these devices are scaled down to feature sizes in the order of atomic dimensions, fundamental limits are approached causing transistors and wires to behave in a manner that is far from ideal [2]. The continued scaling of the MOSFET device leads to an increased leakage (OFF state) current due to short channel effects such as Drain Induced Barrier Lowering (DIBL). Further, the supply voltage cannot be scaled with the feature size in these nanometer devices without severely impacting the performance or energy consumption, as the sub-threshold slope of MOSFETs is limited to 60 mV/decade at room temperature. These challenges

have brought the future of CMOS into question [3] and researchers have commenced their quest for the next digital switch [4].

One such prospective MOSFET replacement device is the Tunnel FET (TFET) [5], which has been billed as “THE GREEN TRANSISTOR” and demonstrated to possess more attractive operating characteristics when compared to CMOS at future technology nodes [6]. Previous research efforts in this area have focused on uses of TFETs as replacements for MOS transistors in SRAM cells [7] [8] [9] for cache architectures and explored the feasibility of hybrid CMOS - TFET [10] [11] cores. This research attempts to provide an inspection of practical limitations in the design of TFET-based systems, propose novel circuit design techniques to overcome them and definitively demonstrate the usefulness and feasibility of such techniques as a path forward by characterizing their impact on a functional unit design.

The remainder of this paper is organized as follows. Section II provides details and descriptions of the Tunnel FET device used in this work along with some of its idiosyncrasies. Section III elaborates on the simulation environment used and the details of novel circuit design techniques that can be used to overcome some of the limitations of the device. The structural and circuit design of a novel sparse-tree adder is presented in Section IV along with detailed evaluation of the energy-delay characteristics of the circuit designs presented. Section V concludes the paper.

II. TUNNEL FET AND ITS IDIOSYNCRASIES

Scaling the supply voltage provides a quadratic reduction in switching energy. However, supply voltage scaling in MOSFET designs reaches a plateau due to the concerns of increased static energy consumption. This is because the threshold voltage (V_t) of the MOSFET must also be scaled along with the supply voltage in order to maintain a sufficiently high on-state drive current (I_{ON}) and thereby avoid performance degradation. This reduction of the threshold voltage (V_t), however, results in an exponential increase of the off-state leakage current (I_{OFF}) which in turn increases static energy consumption. Thus, there is a fundamental limit to the scaling of MOSFET threshold voltages and, consequently, the supply voltage. This limit is determined

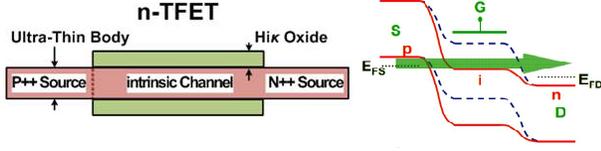


Figure 1: Structure and Band diagram of a generic ultra-thin body nTFET

by the sub-threshold slope of MOSFETs and overcoming this limit provides the primary motivation for research into alternative technologies.

A. TFET device and structure

A promising alternative to the MOSFET, which does not suffer from these limitations, is the Tunnel FET (TFET) [6]. The structure of a basic TFET is simply a $p-i-n$ diode with a gate over the intrinsic region as shown in Figure 1. Tunnel FETs work on the principle of inter-band tunneling of electrons through a barrier instead of flowing over one as in MOSFETs [5]. The gate on the intrinsic region is used to induce a strong band-bending at the source-channel interface, as shown in Figure 1, such that the length of the tunneling path decreases allowing more electrons to tunnel through the barrier.

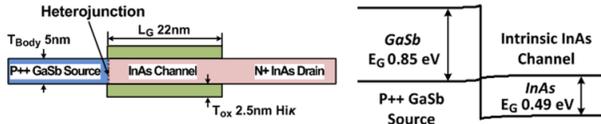


Figure 2: GaSb-InAs Heterojunction nTFET with its Band Diagram [10]

Tunnel FETs are capable of achieving a sub-60mV/decade subthreshold slope and are resilient to short channel effects [12]. These characteristics provide an opportunity to scale the supply voltage without significantly impacting the circuit delay or leakage component of energy consumption. Further, the leakage current (I_{OFF}) of TFETs is far lower than MOSFETs, increasing its attractiveness for energy-constrained designs.

In this study, we utilize a TFET employing a GaSb-InAs heterojunction in the source-channel interface [10] as presented in Figure 2. A higher I_{ON} is observed with heterojunction TFETs compared to homojunction TFETs because the staggered P-N heterojunction, at the source-channel interface, provides a higher critical-field strength for efficient inter-band tunneling. Further, the heterojunction used in this study employs InAs, a lower bandgap material, which further enhances the ON state current.

B. Operating Characteristics of TFETs

The transfer ($I_d - V_g$) and output ($I_d - V_{ds}$) characteristics of a 20nm GaSb-InAs heterojunction nTFET are shown in Figures 3 and 4 respectively. The steep subthreshold

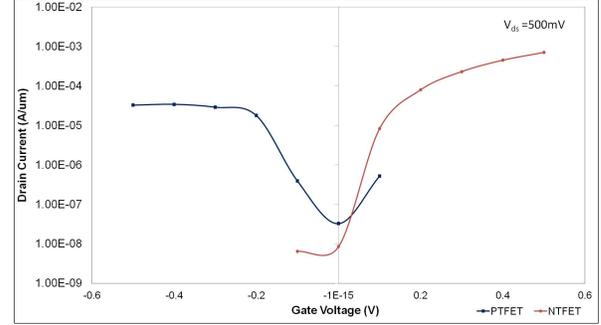


Figure 3: $I_d - V_{gs}$ Characteristics of nTFET & pTFET

slope, mentioned earlier, is clearly observed in the transfer characteristics. The output characteristics also provide some interesting insights about the device. Unlike MOSFETs, we observe asymmetric current conduction as TFET conduction currents are present only in the reverse-bias region. Thus, the device acts like a unidirectional switch with minimal conduction currents under moderate forward bias. Under high forward bias, there is significant I_{DS} regardless of the applied gate voltage.

From the operating characteristics of TFETs, we can infer that TFETs are not direct replacements for MOSFETs in all digital designs. Functionally, neither static nor dynamic design styles are directly impacted by transitioning to TFETs. Other design styles such as Pass Transistor Logic (PTL) must be reassessed as they become dysfunctional due to the asymmetric behavior of TFETs. Certain unique circuit design challenges must be overcome in order to incorporate TFETs into such mainstream designs. Table I summarizes the idiosyncrasies associated with TFETs along with the challenges and opportunities they present to the designer. The challenges tackled and the opportunities exploited in this work are highlighted.

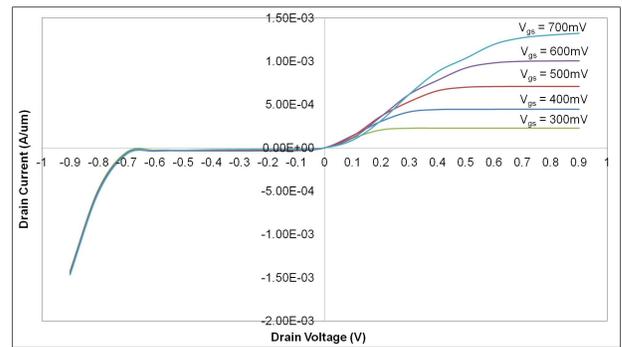


Figure 4: $I_d - V_{ds}$ Characteristics of nTFET

An interesting observation (from the TFET characteristics shown in Figure. 3) is that the pTFETs exhibit a noticeably weaker sub-threshold slope. Further, pTFETs also exhibit a higher I_{OFF} and a lower I_{ON} . The p-type heterojunction TFET has a structure similar to that of the n-type heterojunction TFET, except that the source and channel materials are

reversed. The source region in a pTFET is heavily n-doped InAs, the channel region is intrinsic GaSb, and the drain region is heavily p-doped GaSb. The weaker sub-threshold in pTFET is the result of enhanced source-side degeneracy caused by the heavily doped n+ InAs source. The relatively weaker sub-threshold slope, higher I_{OFF} and lower I_{ON} all contribute negatively to the overall energy consumption and delay performance of a design. As nTFETs exhibit more desirable characteristics of these devices, design styles that incorporate solely nTFETs such as PTL can see further improvement over complementary designs when compared with MOSFETs.

III. SIMULATION FRAMEWORK & CIRCUIT EXPLORATIONS

A compact SPICE model for TFETs has yet to be developed and so, in our experiments, we have used the TFET Verilog-A models obtained from Penn State’s NDCL [13]. Verilog-A models are based on look-up tables and provide an efficient and accurate way of modeling emerging devices which do not yet have compact or SPICE models [7]. A similar model was developed for the 20nm FinFETs in our experiments for comparison.

The following sub-sections elaborates on how the challenges described in section II can be overcome using the opportunities provided at the circuit-level of abstraction to create systems with improved performance and energy-efficiency.

A. Dynamic Circuits utilizing only nTFETs

In traditional MOSFET based designs, only p-channel devices are employed as pull-up transistors. Typically, this is because using a n-channel device results in a degradation of the output voltage levels equivalent to the threshold voltage of the device (V_t drop). However, nTFETs do not exhibit a significant V_t drop and hence, can be used in the pull-up network without affecting the robustness of the circuit. This unique property of TFETs provides designers with the opportunity to attempt novel circuit designs utilizing only nTFETs. Pseudo-NMOS designs which were popular before the advent of CMOS are one such design choice. However, the constant static current problem that necessitated the switch to CMOS will re-surface.

In order to fully exploit this property without paying a significant energy penalty, we propose the use of a dual-clocked dynamic design as shown in Figure. 5. Using only nTFETs necessitates that both the pre-charge and evaluation transistors in the circuit are driven by a logic-high gate voltage bias and hence, the use of a dual-clocked scheme is advocated. However, as with any pre-charge and evaluate design, there exists the possibility of output voltage degradation due to charge-sharing among internal nodes. Traditionally, this drawback is overcome by using a pull-up transistor with an inverter to form a level-restorer circuit.

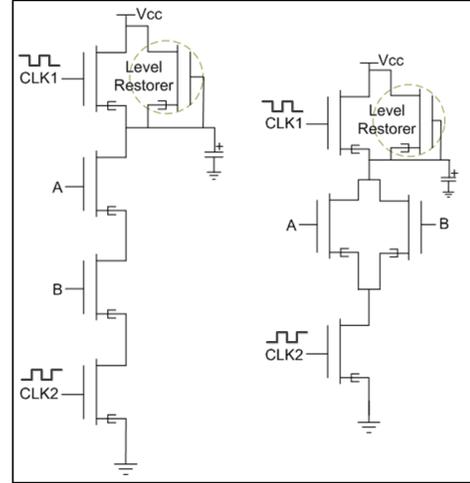


Figure 5: Dual-Clocked Dynamic NAND and NOR circuits employing only nTFETs

However, in this design, a single nTFET is sufficient to implement level restoration as shown in Figure 5. This also improves the response of the level-restore circuitry and improves the energy-delay characteristics by eliminating the impact of the inverter.

The energy-delay characteristics of 2-input NAND and NOR logic gates designed using this technique are shown in Figure 6 along with those of a standard pTFET-inclusive implementation. We observe significant improvements in both energy consumption and delay by employing the proposed technique. Both of these improvements are attributable to the stronger sub-threshold swing and reduced capacitances associated with nTFETs.

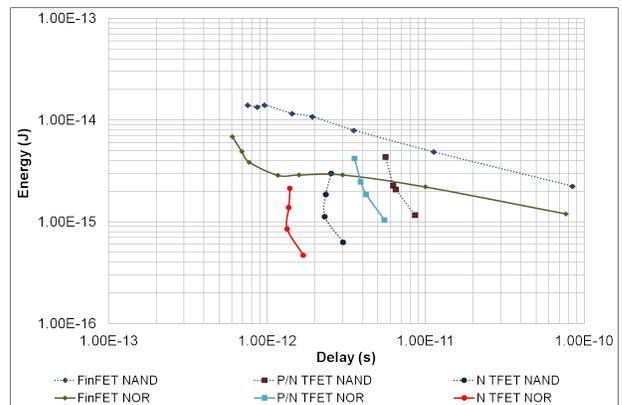


Figure 6: Energy-Delay Plot of 2-input NAND & NOR

B. Pass Transistor Logic for TFETs

Pass Transistor Logic (PTL) [14] is widely used to implement many important logic functions and circuits such as XOR, MUX etc. This is because PTL can implement a given logic circuit with fewer transistors as compared to its static counterpart. Using PTL, logic operations are performed

Characteristics	Opportunities	Challenges
Steep Sub-threshold slope	<ul style="list-style-type: none"> Voltage scaling without performance degradation Mimics “zero” V_t device: <ol style="list-style-type: none"> Better Pass Transistor Logic? Design Styles negating pTFETs? 	Performance concerns due to comparatively weaker pTFET sub-threshold slope
Unidirectional Conduction		<ul style="list-style-type: none"> SRAM Cells Pass Transistor Logic

Table I: Design Opportunities & Challenges for TFET based systems. Those addressed here are highlighted in **Bold**

by connecting and disconnecting the input signal(s) to the output and the same pass transistor stack is used to perform both pull-up and pull-down operations. This, in turn, reduces the latency and switching energy consumed by the circuit due to reduced capacitance in the network.

Thus, the fundamental requirement for PTL is that every device in the PTL stack should be able to source and sink current when needed. As noted in Section II, TFETs are predominantly unidirectional devices that exhibit asymmetric current conduction. This inherent property of tunneling devices cannot be eliminated by structural or material changes. Therefore, the onus lies on the designer to work around this limitation when using TFETs for PTL implementations.

1) *Bi-directional Switch based PTL*: A simple and effective way to construct a bi-directional switch using TFETs is to use two nTFETs, with their drains oriented in opposite directions, creating a bi-directional switch as shown in the inset of Figure 7. The bi-directional switch operates just like a NMOS pass transistor allowing the complete re-use of existing PTL designs, synthesis methods, and tools. The obvious drawback of this implementation is that circuit area is doubled. Additionally, the range of operating voltages must be limited to ensure that no nTFETs in the PTL stack become significantly forward-biased, resulting in unwanted large conduction currents without regard to the gate voltage. A 4:1 multiplexer designed using these bi-directional switches is shown in Figure 7.

2) *Pre-charge Dynamic PTL*: An alternative design, that does not double the area, is the dynamic pre-charge design. In this case, the TFETs in the pass transistor stack are oriented to only discharge the output node, which is pre-charged to V_{cc} every cycle. The inputs of the PTL stack must be isolated from the output node while pre-charging to prevent the possibility of a direct V_{cc} to GND short-circuit. Figure 8 shows an area efficient implementation of 4:1 MUX using pre-charge based PTL. The transistor shown with dotted lines is used to isolate the inputs during the pre-charge cycle. A nTFET based level-restorer circuit, as described above, is also added to negate the impact of charge sharing. The biggest advantage of this design is that only the internal capacitances on the path connecting the input to output in the stack are charged. This is unlike the static design, presented above, where some charged nodes may

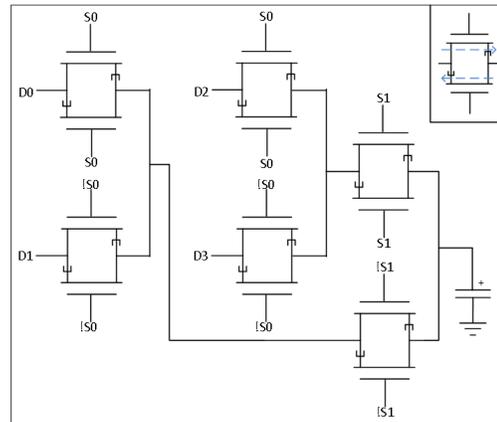


Figure 7: A 4:1 Multiplexer implemented using bi-directional switches. Direction of current flow through nTFETs in bi-directional switch shown in inset

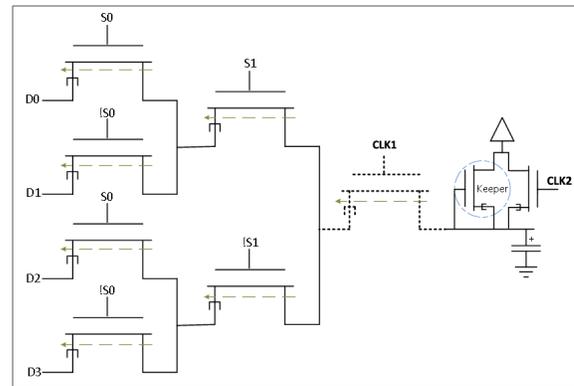


Figure 8: A 4:1 MUX implemented using Pre-Charge PTL. Direction of current flow is indicated by the dotted arrow

be charged needlessly. Despite these advantages, a subtle drawback inherent to this type of design is a limited range of operating voltages. As in the Bi-directional switch PTL, this is to prevent any nTFETs in the PTL stack from becoming significantly forward-biased.

Figure 9 plots the E-D characteristics for 16:1 MUX implemented using the circuit styles discussed above. A FinFET dynamic MUX implementation is also evaluated to assess the contribution of circuit and technology changes holistically. The dynamic energy consumed by the FinFET designs, both traditional PTL and the dynamic PTL imple-

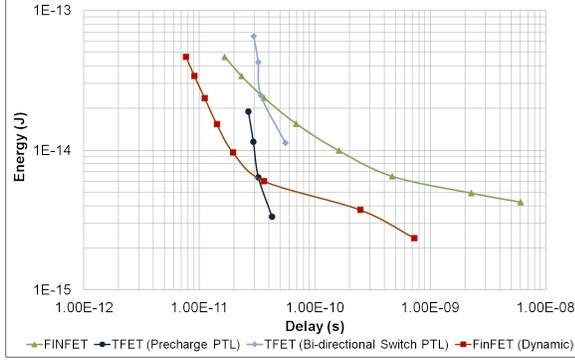


Figure 9: Energy-Delay Plot of FinFET and TFET MUX

mentation, is the same across the design space. However, it must be noted that in traditional PTL design, dynamic energy is determined only by the number of 0-1 transitions of the output node whereas in the case of dynamic PTL both 0-1 and 0-0 transitions are contributors.

The leakage power dissipated by both FinFET circuits is the same; however, the leakage energy varies due to differences in delay. With regard to TFET based designs, the energy-delay characteristic of the pre-charge design is more impressive. Better delay characteristics can be attributed to the isolated charging of solely the output capacitance during the pre-charge phase, resulting in an enhanced rise time and thus lower overall delay. With regard to energy consumption, the dynamic energy consumption of the pre-charge based designs is lower as no internal node capacitances are needlessly charged. Furthermore, the orientation of devices in the PTL stack of the pre-charge design results in reduced leakage energy consumption as no sneak leakage paths may be created by the inputs. The bi-directional switch based designs suffer from the presence of sneak leakage paths, which result in slightly higher leakage energy consumption.

IV. SPARSE MC CHAIN ADDER

For evaluating the impact of these circuit designs at a higher level of abstraction, we chose to model them together in the context of a larger functional circuit. To this end, we designed a sparse prefix tree adder, inspired by [15], which utilizes smaller Manchester Carry Chain (MCC) units in addition to the prefix tree.

The sparse prefix tree provides a quick and efficient design for combining large clusters of internal carry signals into a single carry-out signal. By using a sparse tree rather than a full tree, we trade-off availability of all internal carry signals to obtain better area, energy and performance characteristics. These missing signals are instead “reconstructed” and propagated to the appropriate bit-adders through the use of the MCC units. The MCC units are simple and efficient in achieving quick propagation of carry signals through a short chain to bit-adders.

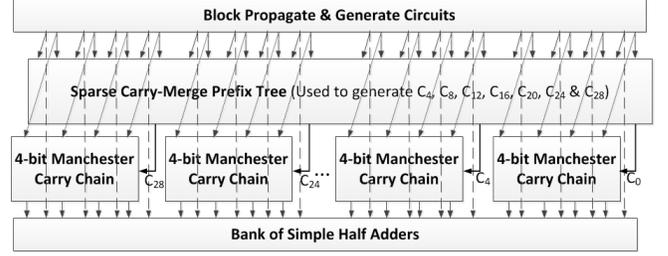


Figure 10: Structural design of the Sparse-prefix tree adder

A. Structural & Circuit Design

The structural implementation of the proposed 32-bit adder is presented in Figure 10. This adder architecture consists of 4 distinct types of circuits, which are composed to form the complete adder. At the first level, the PG unit generates propagate (P) and generate (G) signals for each pair of input bits. A sparse tree is then used to selectively generate specific carry signals that are fed to MCC units. The 32-bit adder is partitioned into eight 4-bit blocks and each contains a 4-bit MCC. The MCC accepts the appropriate P, G, and carry signals, propagating them as needed to the bank of single bit adders which compute the final sum bits. Finally, the carries produced by the MCC and P signals are combined in the bank of single bit adders to produce the final sum.

The proposed adder design was implemented with FinFETs for a baseline reference. Two TFET explorations, one employing both pTFETs and nTFETs (P/N TFET), and another using only nTFETs are implemented. With regards to circuit design, the PG block and the prefix tree are typically implemented using dynamic circuits styles in most adders as they are performance critical. A similar approach is utilized in this work. The circuit implementation of the MCC chain unit with nTFETs is shown in Figure.11. The MCC circuit for the stand alone nTFET implementation is a combination of both techniques described in Section III.

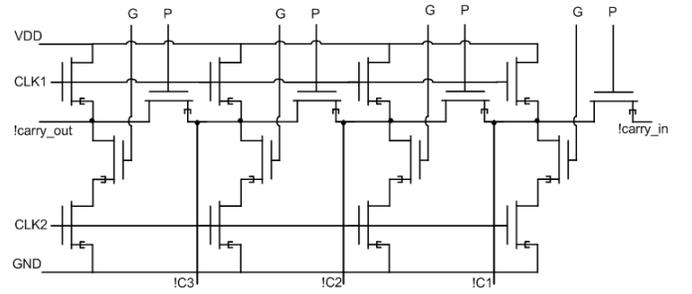


Figure 11: Dynamic nTFET Manchester Carry Chain Circuit

B. Results & Discussion

In order to evaluate the energy and delay characteristics of these circuits, we simulated their timing behavior and energy consumption over a range of voltages. For the FinFET

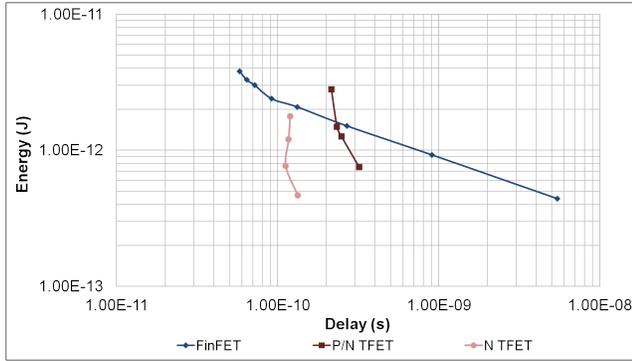


Figure 12: Energy-Delay Plot of 32-bit Sparse Adder incorporating MC Chains

designs, a voltage sweep was done over the range of 1.0V to 0.3V. For the TFET designs, in order to avoid over-biasing the devices incorrectly, voltages above 0.7V were deemed outside of the useful operating range. Therefore, the voltage sweeps for the TFET designs were done only over the range of 0.6V to 0.3V.

The trends seen in our results, shown in figure 12, demonstrate that the P/N TFET circuit is competitive with the FinFET implementation. The P/N TFET design is initially outperformed by the FinFET design. However, as the supply voltage drops to 0.5V and below, the P/N TFET design quickly outperforms the FinFET, in terms of both energy and delay metrics.

Migration to a stand-alone N-TFET design only extends this marked E-D improvement even further. The nTFET design begins beating the FinFET design immediately at the 0.6V node. As the supply voltage continues to decrease, the margin of victory for the nTFETs continues to grow rapidly. The nTFET-only design easily dominates the P/N TFET at all supply voltage nodes.

V. CONCLUSION

While we acknowledge that TFETs pose unique challenges to circuit designers, from this work we see that overcoming these challenges is not only possible, but also allows for additional design exploration. We also show that these designs are not only capable of eliminating or mitigating many of the design challenges posed by TFETs, they also provide additional performance benefits in terms of both energy and delay for logic designs. These results clearly demonstrate that TFET devices are viable and attractive candidates for the future of digital logic designs, especially at ultra-low voltages.

VI. ACKNOWLEDGEMENTS

We thank the anonymous reviewers for their comments and suggestions. This work was supported in part by NSF Awards 1147388, 1160980, 0903432 and 1028807.

REFERENCES

- [1] J. Rabaey, *Low Power Design Essentials (Integrated Circuits and Systems)*. Springer, 2009.
- [2] Y. Taur and T. H. Ning, *Fundamentals of modern VLSI devices*, 2nd ed. Cambridge University Press, Aug. 2009.
- [3] M. Horowitz, E. Alon, D. Patil, S. Naffziger, R. Kumar, and K. Bernstein, "Scaling, power, and the future of cmos," in *Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International*, dec. 2005.
- [4] K. Bernstein, R. Cavin, W. Porod, A. Seabaugh, and J. Welser, "Device and architecture outlook for beyond cmos switches," *Proceedings of the IEEE*, vol. 98, no. 12, dec. 2010.
- [5] C. Hu, "Green transistor as a solution to the ic power crisis," in *Solid-State and Integrated-Circuit Technology, 2008. ICSICT 2008. 9th International Conference on*, oct. 2008.
- [6] A. Seabaugh and Q. Zhang, "Low-voltage tunnel transistors for beyond cmos logic," *Proceedings of the IEEE*, vol. 98, no. 12, dec. 2010.
- [7] 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 *Proceedings of the ACM/IEEE international symposium on Low power electronics and design*, 2009.
- [8] J. Singh, K. Ramakrishnan, S. Mookerjee, S. Datta, N. Vijaykrishnan, and D. Pradhan, "A novel si-tunnel fet based sram design for ultra low-power 0.3v vdd applications," in *Design Automation Conference (ASP-DAC), 2010 15th Asia and South Pacific*, jan. 2010.
- [9] V. Saripalli, S. Datta, V. Narayanan, and J. P. Kulkarni, "Variation-tolerant ultra low-power heterojunction tunnel fet sram design," *Nanoscale Architectures, IEEE International Symposium on*, vol. 0, 2011.
- [10] V. Saripalli, A. K. Mishra, S. Datta, and V. Narayanan, "An energy-efficient heterogeneous cmp based on hybrid tfet-cmos cores," in *DAC*, 2011.
- [11] K. Swaminathan, E. Kultursay, V. Saripalli, V. Narayanan, M. T. Kandemir, and S. Datta, "Improving energy efficiency of multi-threaded applications using heterogeneous cmos-tfet multicores," in *Proceedings of the ACM/IEEE international symposium on Low power electronics and design*, 2011.
- [12] A. S. Verhulst, W. G. Vandenberghe, D. Leonelli, R. Rooyackers, A. Vandooren, S. D. Gendt, M. M. Heyns, and G. Groeseneken, "Tunnel field-effect transistors for future low-power nano-electronics," *ECS Transactions*, vol. 25, no. 7, 2009.
- [13] "Verilog-a models for tunnel fets," <http://www.ndcl.ee.psu.edu/downloads.asp>.
- [14] J. M. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital integrated circuits- A design perspective*, 2nd ed. Prentice Hall, 2004.
- [15] S. Mathew, M. Anders, R. Krishnamurthy, and S. Borkar, "A 4-ghz 130-nm address generation unit with 32-bit sparse-tree adder core," *Solid-State Circuits, IEEE Journal of*, vol. 38, no. 5, may 2003.