Ultra-compact SRAM design using TFETs for low power low voltage applications

*Department of Electrical and Computer Engineering
†Northeastern University, Boston, MA
police.ab@northeastern.edu, ramgopal.r@northeastern.edu

***Abstract*—Tunnel Field-Effect Transistors (TFETs) are gaining significant attention as promising alternatives to conventional CMOS devices, especially for ultra-low power memory design. Unlike MOSFETs, TFETs utilize band-to-band tunneling, allowing them to operate at lower supply voltages while achieving subthreshold slopes below the thermionic limit of 60 mV/decade. This feature makes them highly suitable for energy-constrained applications such as IoT and edge computing devices. In this project, we investigate the design of an ultra-compact SRAM cell utilizing TFETs with the objective of reducing leakage power and improving overall efficiency under low voltage conditions. TFETs, due to their steep switching characteristics, enable robust static noise margins and compact cell designs without sacrificing stability. Furthermore, we explore material choices and structural modifications that enhance tunneling efficiency and device scalability. [1] This work synthesizes findings from recent research to evaluate the feasibility of integrating TFET-based SRAMs in next-generation systems. The analysis also considers practical constraints in fabrication, and evaluates the performance benefits compared to conventional SRAM cells. The insights gained here contribute to understanding the role of TFETs in the future of low-power VLSI design.**

I. Introduction and History

The continued scaling of CMOS technology has enabled exponential growth in computational power, but it also brings about challenges such as increased leakage power, short-channel effects, and limited voltage scalability. In traditional CMOS devices, the subthreshold swing (SS) cannot be reduced below 60 mV/decade at room temperature due to the Boltzmann limit, thereby restricting further reduction of supply voltage and dynamic power. To overcome these limitations, researchers have explored alternative transistor architectures. Tunnel Field-Effect Transistors (TFETs) present a compelling solution, leveraging quantum mechanical band-to-band tunneling instead of thermionic emission for carrier injection. This enables a steeper SS, often below the 60 mV/decade limit, and allows for ultra-low-voltage operation while maintaining a reasonable ON current. Originally conceptualized for low-power logic applications, TFETs have recently gained interest in memory design. Their intrinsic low leakage and steep switching characteristics make them highly suitable for Static Random Access Memory (SRAM), where power consumption, data stability, and area efficiency are critical. One of the pivotal studies in this direction is the 2016 ISCAS paper, which introduced a TFET-based SRAM cell design that achieved reduced static power consumption and enhanced cell compactness compared to CMOS-based designs.[2] This project builds on these developments by analyzing the integration of TFETs into SRAM cell design. We focus on the trade-offs between leakage power, voltage scalability, and static noise margin (SNM), aiming to optimize cell performance for future low-power VLSI systems such as IoT nodes and energy-harvesting devices. The evolution of TFET technology from a theoretical concept to a viable alternative in commercial-grade memory applications underscores its potential in shaping the next generation of ultra-low-power electronics.

II. Device Structure and Material Properties

TFETs differ from MOSFETs primarily in their carrier injection mechanism—tunneling rather than thermionic emission. A conventional n-type TFET consists of a p+-source, an intrinsic or lightly doped channel, and an n+-drain. Electrons tunnel from the valence band of the source to the conduction band of the channel when sufficient gate voltage is applied. In this project, we focus on silicon-based TFETs using vertical or double-gate structures to enhance electrostatic control and tunneling efficiency. Recent advancements in material engineering have also explored heterojunction TFETs (e.g., Si-Ge, InGaAs-GaAsSb) for better band alignment and tunneling probabilities. From a materials perspective, TFETs benefit from:

- High bandgap materials for the channel to suppress leakage
- Low bandgap materials in the source to enhance tunneling
- High-k dielectrics to improve gate control

The simulation parameters for the TFET-based SRAM cell[**?**] in this work are chosen based on these principles to optimize both static noise margin and area efficiency. Device sizing and doping concentrations are carefully tailored to balance tunneling current and short-channel effects. The 3-Transistor Tunnel Field-Effect Transistor (3T-TFET) SRAM cell that serves as the basis for the design makes use of the Negative Differential Resistance (NDR) region of operation, two TFETs (M0 and M1) in reverse bias configuration are used to create a latch.One more TFET (M3) functioning as a read-access transistor is used.TFETs are retained in reverse bias using the following biasing scheme: VDD - VSS 0.6V.To forward bias the TFETs during write, the supply rails are switched from VSS to VDD.The Important aspects of the device architecture is with just three TFETs in the memory bitcell, it is possible to obtain very low leakage (less than 0.1 fA/bit).Area of compact cell: 0.108 m²/bit.Scalability of voltage from 0.6V to 0.2V.For extremely low leakage, the SRAM array is entirely constructed

with TFETs. he peripheral circuits, which include sense amplifiers, decoders, and row/column drivers, are constructed with standard CMOS for high drive strength. [3]

By using band-to-band tunneling, the silicon-based Tunnel Field-Effect Transistors (TFETs) utilized in this design are able to achieve a subthreshold slope that is substantially lower than that of conventional MOSFETs. They are perfect for low-power applications because of this feature, which enables effective switching at significantly lower supply voltages. With a subthreshold slope as low as 30 mV/dec, aggressive voltage scaling is possible without sacrificing functionality.The TFET has three different operating regions when biased in reverse: a hump region where tunneling current predominates, a flat-current region where the current is negligible and mostly unaffected by the gate voltage, and a p-i-n turn-on region where thermionic emission becomes important. In the suggested SRAM architecture, these behaviors are essential for reliable data retention and switching.The Gate-to-Drain Capacitance (CGD) is the predominant component of the total gate capacitance, whereas the Gate-to-Source Capacitance (CGS) is intrinsically low and exhibits a weak dependence on the gate voltage. Due to the memory cell's compact layout and advantageous capacitance distribution, wiring capacitance is significantly reduced—up to 50 percent less than with conventional SRAM designs. This enables faster and more energy-efficient operation in addition to reducing dynamic power consumption.

III. Fabrication and Production

The fabrication of Tunnel Field-Effect Transistors (TFETs) builds upon conventional CMOS processes but incorporates key differences to enable band-to-band tunneling. The crucial step in TFET manufacturing is the formation of abrupt p-i-n junctions, essential for efficient tunneling, and precise doping to control carrier injection regions. These structures require careful process control to maintain device reliability and tunneling efficiency.

TFETs may be fabricated using silicon-on-insulator (SOI) substrates or compound semiconductors like InAs and GaSb to enhance tunneling rates. For this project, we consider a silicon-based vertical double-gate TFET using a high-k dielectric and work-function engineered gates. The proposed flow is compatible with standard VLSI infrastructure, which eases adoption into existing memory technologies.

Key fabrication steps:

1) Start with SOI wafer (or bulk silicon for planar TFETs)
2) Perform ion implantation to define p+ source and n+ drain regions
3) Use atomic layer deposition (ALD) for high-k gate dielectric (e.g., HfO_2)
4) Deposit metal gate (TiN, Mo) and pattern using photolithography
5) Annealing process to activate dopants and maintain sharp junction profiles

In TFET-based SRAM cells, symmetrical doping profiles and minimized parasitic capacitance are essential. Simulation results from recent works show the ability of fabricated TFETs to achieve sub-60 mV/dec subthreshold slopes and drive currents suitable for 6T and 8T SRAM configurations.

TABLE I
Comparison of CMOS and TFET Process Characteristics

Process Metric	CMOS	TFET
Junction Sharpness	Moderate	Requires steep doping profile
Gate Dielectric	SiO_2/HfO_2	High-k (HfO_2, Al_2O_3)
ON/OFF Ratio	High	Moderate to High
Fabrication Complexity	Mature	Under Research
Compatibility	High	Moderate

These fabrication strategies are critical to successfully implementing low-power SRAM using TFETs, paving the way for energy-efficient memory in next-generation VLSI systems[4].

IV. Device Operation Principle and Physics

Tunnel Field-Effect Transistors (TFETs) operate on a fundamentally different principle compared to MOSFETs. While MOSFETs rely on thermionic emission over a potential barrier, TFETs utilize band-to-band tunneling (BTBT), a quantum mechanical process where electrons tunnel directly from the valence band of the source to the conduction band of the channel..

In an n-type TFET, the source is heavily p-doped and the drain is heavily n-doped, forming a reverse-biased p-i-n junction. When a positive gate voltage is applied, the conduction band of the channel aligns with the valence band of the source, enabling tunneling and resulting in a steep subthreshold slope. When the gate is off, the bands are misaligned, and no significant current flows[6].

The unique tunneling mechanism allows TFETs to achieve subthreshold slopes below the 60 mV/decade limit of MOSFETs, enabling operation at lower supply voltages. This drastically reduces static power consumption, making TFETs ideal for ultra-low power applications such as SRAM and IoT.

Analytical Expressions

The BTBT current in TFETs can be approximated using the Kane's model:

$$I_{BTBT} \propto A \cdot E_{field}^2 \cdot \exp\left(-\frac{B}{E_{field}}\right)$$

Where: - E_{field}: Electric field across tunneling junction - A, B: Material-dependent constants

This exponential relationship shows why TFETs require high electric fields and abrupt junctions for efficient tunneling.

Tunneling Efficiency Factors

Several key parameters affect TFET performance:

- **Tunneling barrier width**: Reduced by high gate voltage and sharp doping gradients
- **Material band alignment**: Heterojunctions with staggered bandgap profiles increase tunneling probability
- **Effective mass**: Lower effective mass materials (e.g., InAs, GaSb) improve current drive

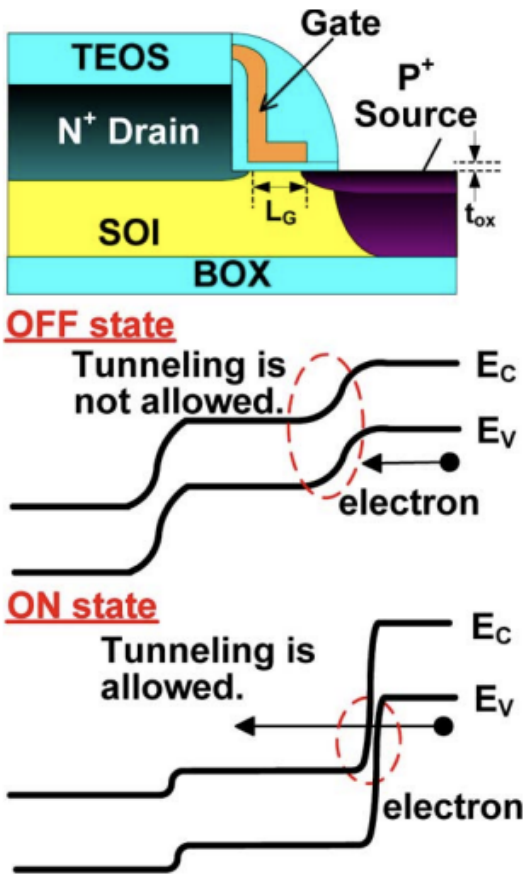

Fig. 1. Device schematic and band diagram
[5]

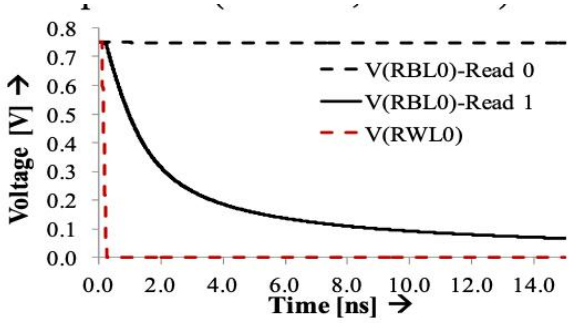

Fig. 2. Read Operation
[7]

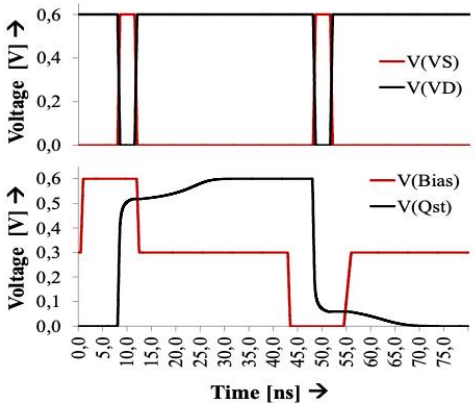

Fig. 3. Write Operation
[7]

These principles directly impact the SRAM cell behavior by determining the ON-current, leakage characteristics, and write-access speed. By carefully choosing material systems and gate work functions, TFETs can be optimized for specific memory performance goals. Compared to a traditional MOSFET, the Tunnel Field-Effect Transistor (TFET) operates on a fundamentally different principle. Band-to-band tunneling, a quantum mechanical process where charge carriers (electrons or holes) tunnel through a narrow energy barrier when sufficient electric field is applied across the gate and source junction, is the basis for TFET operation rather than thermionic emission over a potential barrier. Because of this tunneling effect, TFETs can function well at lower supply voltages and achieve incredibly steep subthreshold slopes.A typical n-type TFET has a p-i-n structure, with the drain being n-doped and the source being heavily p-doped. The energy bands bend when a positive gate voltage is applied, causing the p-type source's valence band to line up with the intrinsic channel's conduction band. Current flows as a result of electrons tunneling from the source into the channel due to this alignment. TFETs can achieve subthreshold slopes below the 60 mV/decade limit of MOSFETs because this process does not require thermal energy, which makes them ideal for ultra-low-power applications. The TFET is highly directional and likewise displays asymmetric behavior. Although this unidirectional current flow reduces leakage, it also presents design difficulties for bidirectional logic circuits. TFETs exhibit three distinct regions of current behavior under reverse-biased conditions: (1) the hump region, where tunneling current is dominant; (2) the flat-current region, where the current is low and mostly independent of the gate voltage; and (3) the p-i-n turn-on region, where thermionic emission causes a sudden increase in current.In this work, the design of the 3T-SRAM cell depends on these areas. In particular, a latch-like behavior between two TFETs is implemented using the NDR (Negative Differential Resistance) behavior observed in the hump region, enabling the SRAM cell to store data with high stability and little leakage. The leakage current stays very low—typically in the attoampere range—because the storage TFETs function in reverse bias and only one of them is active at a time in retention mode. The fundamental requirements of low-power electronics, such as embedded systems and Internet of Things nodes, are met by this physical operation model, which not only lowers standby power but also guarantees dependable operation under aggressive voltage scaling.

V. Scaling Laws for the device

Because Tunnel Field-Effect Transistors (TFETs) rely on quantum mechanical band-to-band tunneling, their scaling behavior is governed by principles that are very different from those of conventional CMOS devices. TFETs allow current conduction through tunneling at the source-channel junction, enabling steeper subthreshold swings and better performance at low supply voltages than MOSFETs, which rely on thermionic emission across a potential barrier. In advanced technology nodes, this fundamental difference results in special scaling laws that promote low-power operation and improved energy efficiency.TFETs outperform MOSFETs in terms of electrostatic control as the device's dimensions decrease, especially the channel length. Nonetheless, the alignment and sharpness of the source-to-channel junction have a significant impact on the tunneling behavior. Because the current path depends on the tunneling junction rather than the channel length, a shorter channel length does not substantially increase leakage as in MOSFETs. However, excessively aggressive scaling can result in increased ambipolar conduction, or unwanted current flow from the drain, which needs to be controlled with careful device engineering.Another important factor in TFET operation is gate dielectric scaling. Enhancing the electric field and raising the tunneling probability, gate control over the tunneling region is improved by reducing the gate oxide thickness or by using high- materials. Lower switching voltages and improved subthreshold swing are the outcomes of this. Reduced gate-to-source capacitance (CGS), which stays low and largely independent of gate voltage, also helps TFETs by lowering dynamic power consumption. But as devices get bigger, the gate-to-drain capacitance (CGD) becomes more important. This can impact switching speed, so layout and material selection must be optimized.In scaled TFETs, doping profile engineering becomes more crucial. The abruptness of the source-channel junction has a significant impact on the tunneling current. ON-current performance is improved and tunneling efficiency is increased with sharper doping gradients. Advanced fabrication techniques, like epitaxial growth or laser annealing, are necessary to achieve such abrupt junctions at nanoscale dimensions in order to precisely control dopant placement and minimize diffusion.The ability of TFETs to function at ultra-low supply voltages (VDD), even below 0.3V, is one of their most noteworthy advantages when it comes to scaling. Their steep subthreshold slopes enable fast switching without the need for high gate overdrive, enabling this. This feature allows for aggressive voltage scaling as technology nodes continue to shrink, which significantly lowers both active and standby power consumption—a crucial feature for applications like wearable electronics and the Internet of Things.

Analytical Expressions for Scaling Behavior

The subthreshold swing (SS) is a critical metric that influences the voltage scalability of TFETs. It is defined as:

Parameter	TFET Scaling Behavior
Subthreshold Slope	Improves (steeper than CMOS)
Leakage Current	Decreases (better than CMOS)
$_{DD}$ Scalability	Excellent (operates at 0.3V)
ON-current	Limited, depends on tunneling efficiency
Capacitance	Reduces dynamic power with C_{GS} scaling
Junction Sensitivity	High – requires abrupt doping control

TABLE II
Scaling behavior of TFET parameters.

$$SS = \left(\frac{dV_G}{d(\log I_D)} \right)$$

For ideal TFETs, the SS can be lower than the 60 mV/decade limit seen in MOSFETs, enabling aggressive scaling to lower supply voltages[8].

Additionally, the tunneling current sensitivity to doping gradient and electric field can be described by:

$$I_{BTBT} \propto \exp\left(-\frac{E_g^{3/2}}{E_{field}} \right)$$

where:

- E_g is the material's bandgap energy,
- E_{field} is the electric field across the tunneling junction.

VI. Noise and Limitations

TFETs face a number of noise-related issues that can affect their scalability and reliability, despite their benefits in ultra-low power operation. One of the main issues is thermal noise, which is more noticeable in TFETs running at very low voltages even though it is typically lower due to lower current levels. Because TFETs have a much lower ON-state current than MOSFETs, the signal-to-noise ratio may deteriorate, making it more challenging to discern between logical states in noisy environments. In high-density SRAM arrays, where even minor noise disturbances can cause data corruption, this is especially crucial.Random Telegraph Noise (RTN), which is brought on by charge trapping and detrapping at the interface or inside the gate dielectric, is another important noise mechanism in TFETs. Because TFETs rely on abrupt junctions and small active areas, they are more vulnerable to RTN. Small changes in tunneling current can cause timing errors and bit flips in memory applications, making these fluctuations particularly troublesome in subthreshold operation. Furthermore, any fluctuation or variability has a proportionally larger effect on performance because TFETs operate at low drive currents.In addition to noise, TFETs have a number of basic drawbacks. One of the most notable is that, in comparison to MOSFETs, their ON-current is comparatively low, which limits their performance in high-speed logic applications. TFETs are less appropriate for applications needing high drive strength because of the tunneling mechanism's inherent limitation on the amount of current that can flow through the device. Moreover, the performance of TFETs is highly sensitive to the abruptness of the source-to-channel junction

and the material properties at that interface. One of the biggest fabrication challenges is still achieving the sharp doping profiles needed for effective tunneling.The unidirectional nature of TFETs presents another drawback. Because of their p-i-n structure, TFETs are inherently asymmetric, in contrast to MOSFETs, which have bidirectional conductivity. This limits their application in specific logic styles and necessitates extra design considerations in order to implement symmetric or bidirectional functionality. Furthermore, ambipolar conduction, in which undesired tunneling from the drain side increases leakage, can affect TFETs. Complex design trade-offs or material engineering are frequently needed to suppress this effect.

Analytical Expressions for Noise Behavior

The thermal noise current (i_n) in a TFET can be estimated using:

$$i_n^2 = 4kTg_m$$

where:

- i_n^2 is the noise current spectral density,
- k is Boltzmann's constant,
- T is the absolute temperature,
- g_m is the transconductance of the device.[9]

Since TFETs have lower g_m compared to MOSFETs, thermal noise is generally reduced, but their low current levels can still pose challenges for signal integrity.

Random Telegraph Noise (RTN), another significant noise source in TFETs, can be modeled as:

$$\Delta I = q \times \Delta n$$

where:

- ΔI is the current fluctuation,
- q is the electronic charge,
- Δn is the change in trapped carrier number[10].

RTN becomes more critical in TFETs due to their small device dimensions and low drive currents.

VII. CURRENT APPLICATIONS

The Internet of Things (IoT) devices are the most direct and significant application for the suggested 3T-TFET SRAM design. The TFET-based memory's ultra-low leakage and compact design greatly benefit these systems, which function under strict power and space constraints. Embedded SRAM can take up over 50 percent of the chip area and control standby power consumption in contemporary SoCs for the Internet of Things. IoT devices can operate more sustainably and with longer battery life in energy-constrained environments by incorporating TFET-based SRAMs, which have leakage currents as low as 0.1 fA/bit.[11]

Wearable electronics, like fitness trackers, smartwatches, and health monitoring gadgets, are another crucial application area. These gadgets need to be always on and have extended standby times, where data must be stored in memory blocks without consuming a lot of power. The 3T-TFET SRAM is well suited for these use cases due to its strong data retention capability, low leakage, and ability to function at supply voltages as low as 0.2V. It enables designers to create memory systems that support user responsiveness and data integrity while remaining active and using little energy.

Memory efficiency is crucial in edge computing scenarios and sensor nodes, which process data locally before communicating. These systems frequently use energy harvesting and subthreshold logic, which calls for ultra-low-power memory components. The TFET-based SRAM design is ideal for memory blocks in edge processors and remote sensing units because it allows aggressive voltage scaling without compromising stability or access speed. Its suitability for these decentralized computing platforms is further enhanced by its low energy per read/write access and compatibility with dynamic power gating techniques.

Additionally, the SRAM design is well suited to the requirements of ultra-low-power microcontrollers utilized in low-duty-cycle embedded applications and wireless sensor networks (WSNs). Memory is crucial to data retention in these systems, which spend the majority of their operating time in sleep or idle modes. By reducing this overhead, the suggested TFET solution prolongs battery life and permits dependable operation under constrained energy budgets. Traditional SRAMs greatly contribute to leakage during these times.

Lastly, always-on memory blocks are frequently needed in complex system-on-chip (SoC) designs in order to preserve important data while the system is in deep sleep or standby. For such memory partitions, the hybrid TFET/CMOS SRAM architecture presented in this work is perfect. The TFET-based array guarantees low leakage, while the use of CMOS in the periphery guarantees compatibility with common design flows. Because of this, the suggested SRAM architecture is very good at lowering both dynamic and static power, which makes it appropriate for next-generation low-power SoCs in autonomous embedded systems, consumer electronics, and medical devices.

VIII. DISCUSSIONS

Working on the 3T-TFET-based SRAM design gave us valuable insights into the potential of TFET technology for ultra-low-power and compact memory applications. Throughout this project, we realized how TFETs, with their band-to-band tunneling mechanism, enable operation at much lower supply voltages compared to conventional CMOS devices. This property made it possible to achieve extremely low leakage currents — less than 0.1 fA per bit — which is a significant improvement over traditional 6T CMOS SRAM cells.

One major takeaway from this project was the importance of energy efficiency, especially in systems like IoT devices, wearables, and edge computing platforms, where minimizing standby power is critical. The ability of the 3T-TFET SRAM cell to maintain stable data at ultra-low voltages without

compromising performance is particularly impressive. Additionally, I learned how the compact cell area (0.108 μm^2) not only helps reduce the overall memory size but also enhances speed and reduces dynamic power consumption by minimizing parasitic capacitance.

However, the project also highlighted several practical challenges. Even though TFETs offer clear advantages in leakage reduction and voltage scalability, issues like limited ON-current, difficulties in material integration, and manufacturing complexities still remain barriers for widespread adoption. Another important observation was that, while the SRAM array uses TFETs, the peripheral circuits still rely on CMOS technology for sufficient drive strength, which could complicate the overall integration.

Moreover, although the simulations showed promising performance, real-world factors like variability, temperature changes, and noise could affect stability, emphasizing the need for further robust modeling and design optimizations.

In conclusion, this project showed me that 3T-TFET-based SRAM architectures have strong potential for enabling future ultra-low-power memory systems. The combination of extremely low leakage, wide voltage scalability, and compact design makes TFET SRAMs a strong candidate for next-generation applications. However, continued research in fabrication techniques and hybrid integration will be necessary to fully unlock their practical potential.

REFERENCES

1 Gupta, N., Makosiej, A., Vladimirescu, A., Amara, A., and Anghel, C., "3t-tfet bitcell based tfet-cmos hybrid sram design for ultra-low power applications," in *2016 Design, Automation Test in Europe Conference Exhibition (DATE)*, 2016, pp. 361–366.

2 Bernstein, K., Cavin, R. K., Porod, W., Seabaugh, A., and Welser, J., "Device and architecture outlook for beyond cmos switches," *Proceedings of the IEEE*, vol. 98, no. 12, pp. 2169–2184, 2010.

3 Jafari, B., Tavakoli, J., and Sheikhaei, S., "Bish–qvco: A low–power, low–phase noise bulk–injected super–harmonic coupling qvco," *Microelectronics Journal*, vol. 111, p. 105001, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0026269221000124

4 Saripalli, V., Datta, S., Narayanan, V., and Kulkarni, J. P., "Variation-tolerant ultra low-power heterojunction tunnel fet sram design," in *2011 IEEE/ACM International Symposium on Nanoscale Architectures*, 2011, pp. 45–52.

5 Choi, W. Y., Park, B.-G., Lee, J. D., and Liu, T.-J. K., "Tunneling field-effect transistors (tfets) with subthreshold swing (ss) less than 60 mv/dec," *IEEE Electron Device Letters*, vol. 28, no. 8, pp. 743–745, 2007.

6 Wei, R., Wang, K., Wang, Z., Yang, L., Zhang, F., Wu, Y., Ren, Y., Ye, L., Zhang, L., Bu, W., Huang, R., and Huang, Q., "A novel tfet-mosfet hybrid sram for ultra-low-power applications," in *2023 IEEE 15th International Conference on ASIC (ASICON)*, 2023, pp. 1–4.

7 Gupta, N., Makosiej, A., Vladimirescu, A., Amara, A., and Anghel, C., "Ultra-compact sram design using tfets for low power low voltage applications," in *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2016, pp. 594–597.

8 Seabaugh, A. C. and Zhang, Q., "Low-voltage tunnel transistors for beyond cmos logic," *Proceedings of the IEEE*, vol. 98, no. 12, pp. 2095–2110, 2010.

9 Schroder, D. K., *Semiconductor Material and Device Characterization*, 3rd ed. Wiley-IEEE Press, 2006.

10 Ghibaudo, G., "Low frequency noise and rtn in advanced cmos devices," *Microelectronics Reliability*, vol. 53, no. 9-11, pp. 1299–1302, 2013.

11 Ahmad, S., Alam, N., and Hasan, M., "Robust tfet sram cell for ultra-low power iot applications," *AEU - International Journal of Electronics and Communications*, vol. 89, pp. 70–76, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1434841117325049