Band-to-Band Tunnelling in SOI MOSFETs

EE724 Project Report
Abhineet Agarwal, Harsh Pujare, Kushagra Gehlot
Mentor - Jay Sonawane
Instructor - Prof. Udayan Ganguly
IIT Bombay, Powai, Mumbai, India - 400076

***Abstract*—We analyse Band-to-Band Tunneling in Partially and Fully-Depleted SOI MOSFETs, optimising design parameters to minimise and maximise BTBT current for applications such as BTBT Integrators and Inverters, respectively. Band-to-Band tunneling is used to achieve ultra-low power hardware neurons by having low current compared to other CMOS implementations, thus enabling large-scale Spiking Neural Networks(SNN) in hardware. The dependence of BTBT current on temperature and gate length scaling is also analysed.**

***Index Terms*—Neuromorphic computing, BTBT, SNN, FD/PD**

I. Introduction

Neuromorphic computing architectures consisting of biologically inspired neural networks are excellently manifested by the spiking neural networks (SNNs). The enormous number of neurons in the human brain makes it essential to use large SNNs to realize biological neural networks. Compact area and energy-efficient technology are essential for large-scale deployment of neuromorphic computing hardware for various applications.

Band-to-Band tunnelling is a quantum mechanical phenomena where electrons move from the valence band of the intrinsic region to the conduction band of the adjacent P-type doped region. This can occur when a large bias voltage is applied at the drain. BTBT can be used to create a "hole storage" at the floating body of a SOI MOSFET. This enables the usage of such devices for a variety of applications in neuromorphic computing.

In phase 1 of the project, we minimise BTBT current in a BTBT Partially Depleted(PD)-SOI MOSFET Integrator (used to trigger spikes in the neural circuit after the voltage crosses a threshold) to decrease the power consumption of the neuron to enable large-scale SNN implementations, by varying design parameters[9]. We identify three important parameters on whom the current depends. In phase 2, we maximise BTBT current in a Fully Depleted(FD)-SOI MOSFET BTBT-based Inverter to enable a fast build-up of holes in the floating MOSFET body. We do this by varying N_{CH}, N_{SD} and L_{OV}(overlap length, between the gate and source/drain). We also check for temperature dependence of the BTBT current, and the changes in current on scaling the gate length from 32nm to 22nm.

II. Motivation

To implement SNNs on a large scale, efficient neurons that are low area and energy efficient are necessary. With BTBT neurons being used as one such implementation, decreasing BTBT current is essential for increasing energy efficiency and decreasing spiking frequency for an enhanced biologically plausible neuron implementation. Determining device design parameters directly affecting tunnelling current is crucial for fine-tuning manufacturing processes for optimal performance. Using BTBT neurons for implementing an inverter offers the same benefits of energy efficiency. Maximising BTBT current is needed for high charging speeds to increase device bandwidth. For a future scaling of the device size to improve area efficiency, a study is necessary to see the effects of the same on BTBT current. Similarly, temperature variations can lead to varying tunnelling currents can vary device performance. We perform a detailed analysis on each of these performance factors.

III. Literature Survey

A. Band-to-Band Tunneling

In band-to-band tunneling charge carriers tunnel through the forbidden energy band gap from one band to another [10][3]. Several models have been proposed to describe band-to-band tunneling.

- **Hurkx Model**: It describes BTBT as a result of the coupling of electronic states across the tunnel barrier. The model considers both the direct and indirect tunneling processes. Direct tunneling involves the transfer of carriers through the energy barrier, while indirect tunneling occurs via trap-assisted tunneling mechanisms. It provides a description of tunneling in terms of carrier energies, tunneling distance, and the effective mass of carriers [5].
- **Schenk Model**: This model focuses on the impact of the electric field on band-to-band tunneling. It considers the electric field dependence of tunneling probabilities and incorporates the effects of band non-parabolicity. It provides insights into the role of electric fields in enhancing or suppressing tunneling currents [6].
- **Nonlocal Path Model**: This model takes into account the nonlocal nature of tunneling paths in the device structure. It considers the contributions from multiple tunneling paths, accounting for their varying lengths and associated probabilities. It offers a more accurate representation of tunneling behavior in realistic device geometries compared to simpler models. In our simulations, we use the NLP model. [4].

B. PD-SOI MOSFET BTBT Integrator

- The device is used as a leaky integrator in BTBT mode by applying a sufficiently high voltage between the body and source/drain junction. The body of the device is used as an integrator output terminal.
- The electrons in the body tunnel to the drain (Id,BTBT), leaving holes in the body to cause the body potential to rise. I(SB) is the hole leakage current between the source and body, which results in leaky integration. However, there is also an electron leakage current from the drain to the source through the conduction band, which results in power loss[8][2].

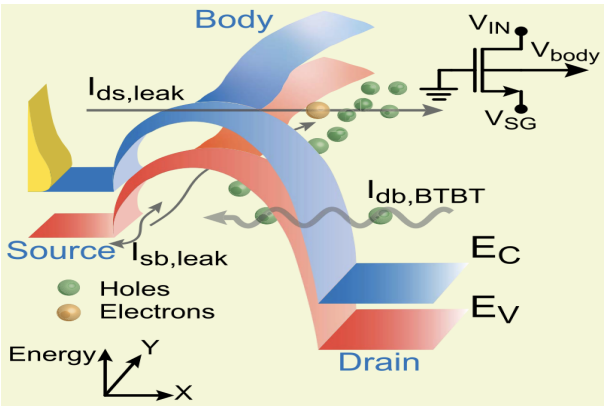

Fig. 1. Schematic of (PD)SOI-MOSFET

C. FD-SOI MOSFET BTBT Inverter

- The device is used as a leaky inverter in BTBT mode by applying a sufficiently high voltage between the body and the source/drain junction. The body of the device is used as the inverter output terminal.
- A mechanism similar to the BTBT Integrator is used. A high BTBT current is needed for quick rise time of V_{OUT} at low V_{IN} voltages. The SS regime transistor discharges the output terminal when the input goes high. This is the reset phase.

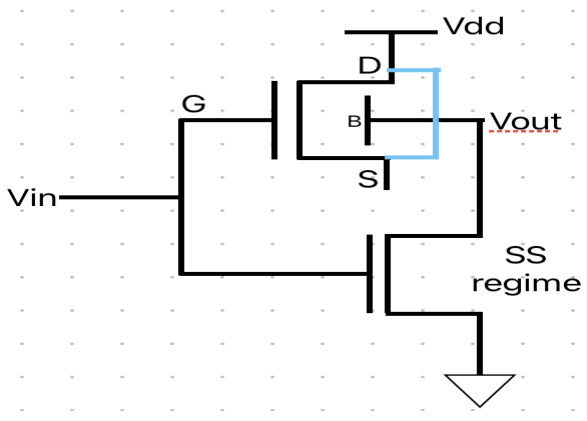

Fig. 2. Schematic of (FD)SOI-MOSFET Inverter

IV. Methodology

A. Phase 1: BTBT Integrator

The 32nm PD-SOI MOSFET fabricated at GF with gate length (LG)/width (W) of 40/450 nm was used. The electrical measurements were performed using an Agilent B1500 Semiconductor parameter analyzer. A well-calibrated Sentaurus TCAD[7] deck was used. For design space analysis, variation in channel height (H_{CH}), buried oxide thickness (H_{BOX}), substrate height (H_{SUB}), drain-gate (D-G) and source-gate (S-G) overlap length (L_{OV}), equivalent oxide thickness (T_{OX}), channel (N_{CH}) and source/drain doping (N_{SD}) were considered, and the simulation results were analysed to find ways to reduce the BTBT current.

B. Phase 2: BTBT Inverter

The 22nm FDXSOI MOSFET fabricated at GF with gate length (LG)/width (W) of 25/300 nm was used. The electrical measurements were performed using an Agilent B1500 Semiconductor parameter analyzer. A well-calibrated Sentaurus TCAD[7] deck was used. For increasing BTBT current, changes in drain-gate (D-G) and source-gate (S-G) overlap length (L_{OV}) and channel (N_{CH}) doping were considered to find the optimal device for maximum BTBT current.

We checked the I-V characteristics at three values of temperature for the base design (300 K, 358 K and 398 K) for both the PDSOI MOSFET and FDSOI MOSFET to study the variability of BTBT current with temperature.

The effects of L-G scaling from 32nm to 22nm were also studied again by comparison of I-V characteristics, before and after scaling, without changing the channel doping and with. Parameters in the base design:

Parameter	32nm PDSOI	22nm FDX
L_G	40 nm	25 nm
H_{CH}	40 nm	8 nm
H_{box}	100 nm	20 nm
H_{sub}	100 nm	10 nm
L_{OV}	2.5 nm	1 nm
L_{sp}	20 nm	7 nm
T_{ox}	1 nm	1 nm
N_{ch}	6.4e17	1e12cm^{-3}
N_{sd}	1e20	1e20cm^{-3}

V. Results

A. Phase 1: BTBT Integrator

- H_{BOX} and H_{CH} scaling has little effect on the BTBT current. This is primarily because BTBT predominantly occurs in the D-G overlap region.

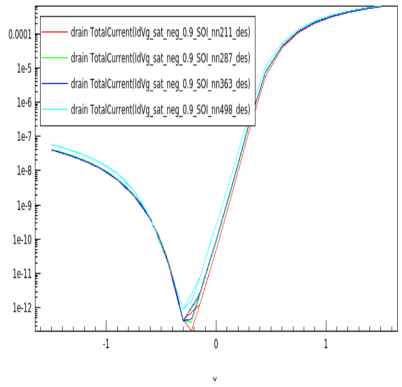

Fig. 3. Variation with H_{BOX}
I-V characteristics plotted at H_{BOX} of 10 nm, 25 nm, 40 nm and 100 nm.

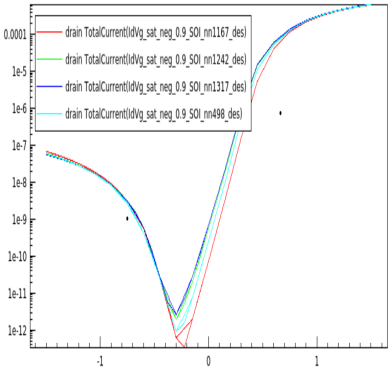

Fig. 4. Variation with H_{CH}
I-V characteristics plotted at H_{CH} of 10 nm, 25 nm, 40 nm and 100 nm.

- A slight increase in the BTBT current is observed as the D-G overlap (L_{OV}) increases. This is maybe because of enhanced electric field at the junction.

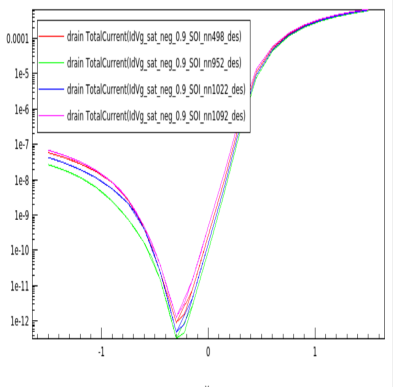

Fig. 5. Variation with L_{OV}
I-V characteristics plotted at L_{OV} of 1 nm, 1.5, 2.5 nm and 3.5 nm.

- BTBT current increases with an increase in the channel doping (N_{CH}).

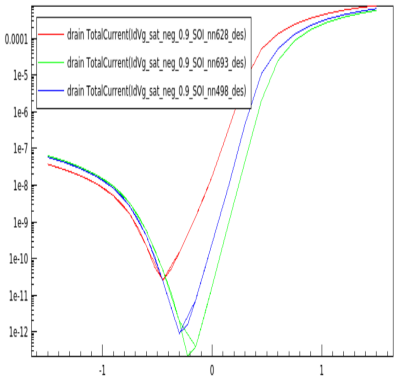

Fig. 6. Variation with N_{CH}
I-V characteristics plotted at N_{CH} of 6.4e17, 1e17 and 5e16

- Increase in the S-D doping (N_{SD}) leads to a large increase in the BTBT current.

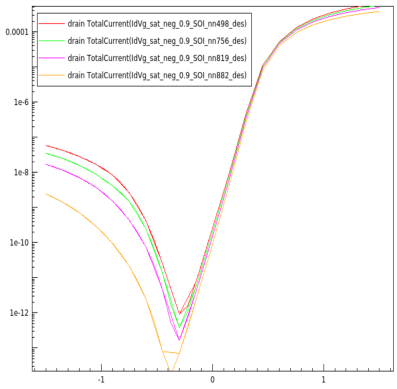

Fig. 7. Variation with N_{SD}
I-V characteristics plotted at N_{SD} of 1e20, 7e19, 5e19 and 3e19

Higher N_{CH} and N_{SD} values lead to a reduction in the depletion width at the drain-channel junction. This results in a steeper energy band profile, shorter tunnel lengths, and a higher electric field in the overlap region.

B. Phase 2: BTBT Inverter

- Increasing L_{OV} increases the BTBT current slightly, but increasing N_{CH} increases it by a large amount. N_{SD} already being 10^{20} in the base design cannot be increased any further. Optimal design for maximum BTBT current is found at L_{OV} of 1.4nm and N_{CH} of 10^{17}. Increasing L_{OV} beyond this at an N_{CH} of 10^{17} leads to convergence issues.

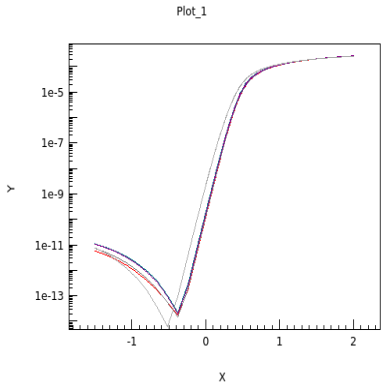

Fig. 8. L_{OV} variation
I-V characteristics plotted at L_{OV} of 1 nm, 1.05 nm, 1.1 nm. 1.15 nm, 1.2 nm and 1.25 nm.

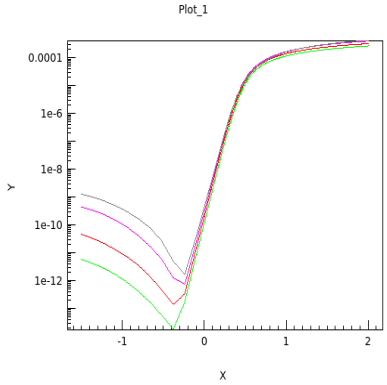

Fig. 9. Variation with N_{ch} at L_{OV} of 1nm
I-V characteristics plotted at N_{ch} of 1e12, 1e14, 1e16 and 1e17.

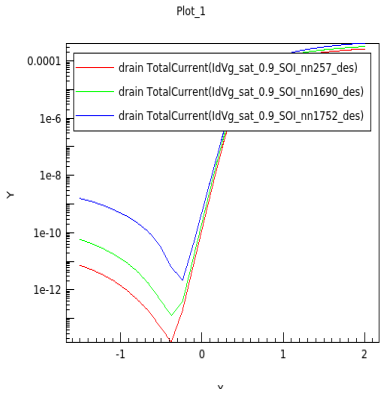

Fig. 10. Variation with N_{ch} at $L_{OV} of 1.25nm$
I-V characteristics plotted at N_{ch} of 1e12, 1e14, 1e16 and 1e17.

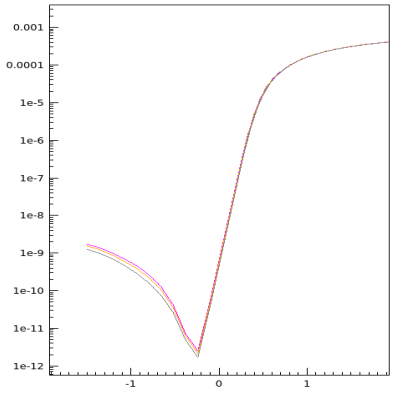

Fig. 11. Finding the optimal design

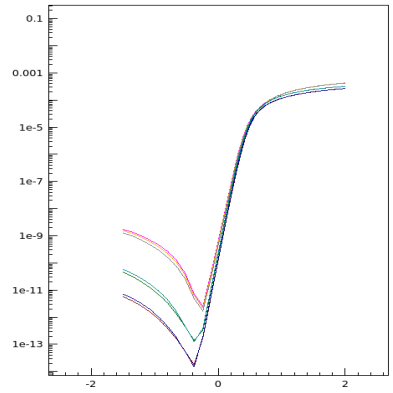

Fig. 12. Finding the optimal design

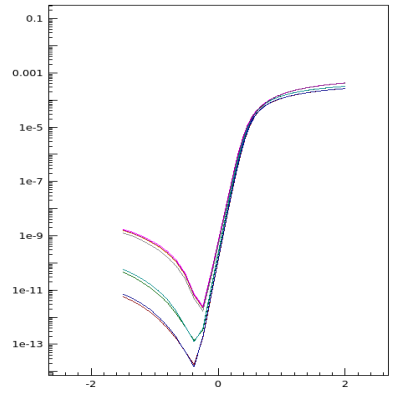

Fig. 13. Finding the optimal design; the maximum current is observed at L_{OV} of 1.4nm and N_{ch} of 1e17.

- A regular and fairly large increase in BTBT current is observed at increasing temperature in both the cases. This is because if temperature increases then the bandgap will decrease and hence tunnelling probability increases leading to increase in tunnelling current[1].

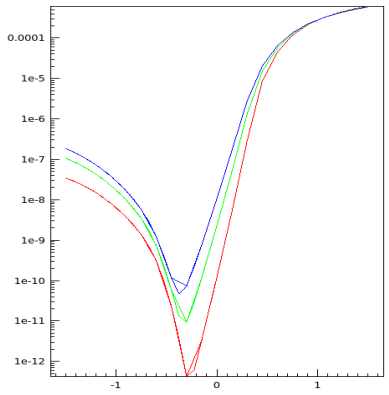

Fig. 14. Temperature variability for 32nm PDSOI Plots at temperatures of 300 K, 358 K and 398 K.

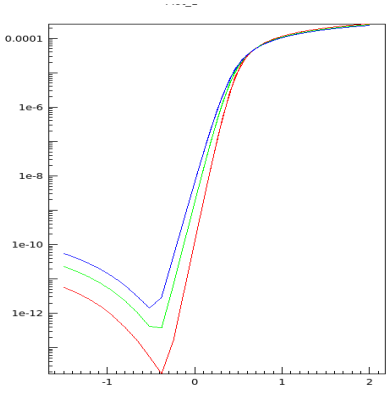

Fig. 15. Temperature variability for 22nm FDXSOI Plots at temperatures of 300 K, 358 K and 398 K.

- We observe that just scaling down the lengths from 32nm PDSOI to 22nm FDXSOI or scaling them up from 22nm to 32nm is not enough. In order to replicate the I-V characteristics, the channel doping (N_{CH}) also needs to be changed accordingly to ensure the channel is fully depleted in the 22nm case, and partially depleted in the 32nm case.

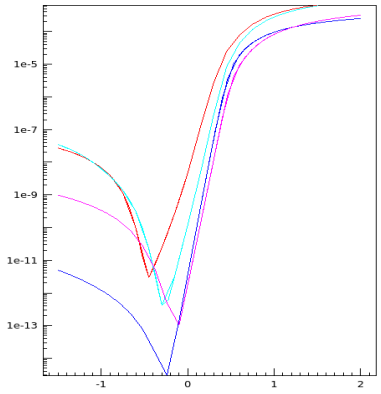

Fig. 16. L_G scaling

VI. Conclusion

In the first stage, we performed design space and variability analysis to achieve a lower BTBT current. We inferred from the simulations that a low LOV, low NSD, and low NCH are favorable parameters for reducing the BTBT current.

In the second stage, we investigated the effects of L_G scaling from a 32nm PDSOI to a 22nm FDX SOI MOSFET. We performed design space and variability analysis of the FDX22 MOSFET to achieve higher BTBT current, and found an optimal design at an L_{OV} of 1.4 nm and an Nch of 1e17 cm^{-3}. We also investigated the effect of temperature on BTBT current and observed that it increases with increase in temperature.

References

[1] Paula Agopian et al. “Temperature impact on the tunnel fet off-state current components”. In: *Solid-State Electronics* 78 (Dec. 2012), pp. 141–146. DOI: 10.1016/j.sse.2012.05.053.

[2] Tanmay Chavan et al. “Band-to-Band Tunneling Based Ultra-Energy-Efficient Silicon Neuron”. In: *IEEE Transactions on Electron Devices* 67.6 (2020), pp. 2614–2620. DOI: 10.1109/TED.2020.2985167.

[3] Robert Entner. “Modeling and Simulation of Negative bias temperature Instability”. PhD thesis. Vienna University of Technology, 2007.

[4] Koichi Fukuda et al. “On the nonlocal modeling of tunnel-FETs Device and Compact models”. In: 2013. URL: https://api.semanticscholar.org/CorpusID:221968337.

[5] G.A.M. Hurkx, D.B.M. Klaassen, and M.P.G. Knuvers. “A new recombination model for device simulation including tunneling”. In: *IEEE Transactions on Electron Devices* 39.2 (1992), pp. 331–338. DOI: 10.1109/16.121690.

[6] A. Schenk. “Rigorous theory and simplified model of the band-to-band tunneling in silicon”. In: *Solid-State Electronics* 36.1 (1993), pp. 19–34. ISSN: 0038-1101. DOI: https://doi.org/10.1016/0038-1101(93)90065-X.

[7] *Sentaurus Device Manual*. Synopsys. Santa Clara, 2010.

[8] Ajay Kumar Singh et al. “Quantum Tunneling Based Ultra-Compact and Energy Efficient Spiking Neuron Enables Hardware SNN”. In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 69.8 (2022), pp. 3212–3224. DOI: 10.1109/TCSI.2022.3172176.

[9] Jay Sonawane et al. *Design Space and Variability Analysis of SOI MOSFET for Ultra-Low Power Band-to-Band Tunneling Neurons*. 2023. arXiv: 2311.18577 [physics.app-ph].

[10] Yuan Taur and Tak H. Ning. *Fundamentals of modern VLSI devices*. Cambridge University Press, 1998.