

# Multi-gate FeFET Discriminates Spatiotemporal Pulse Sequences for Dendrocentric Learning

Hugo J.-Y. Chen<sup>1†</sup>, Matthew Beauchamp<sup>1†</sup>, Kasidit Toprasertpong<sup>1</sup>, Fei Huang<sup>1</sup>, Louis Le Coeur<sup>1</sup>, Thorgund Nemec<sup>4</sup>, H.-S. Philip Wong<sup>1</sup>, and Kwabena Boahen<sup>1,2,3</sup>

<sup>1</sup>Department of Electrical Engineering, <sup>2</sup>Department of Bioengineering, <sup>3</sup>Department of Computer Science, Stanford University, Stanford, CA, USA, <sup>4</sup>GlobalFoundries, Dresden, Germany,

email: {hugochen, bchampm, topas, lecoeur, hspwong, boahen}@stanford.edu, thorgund.nemec@globalfoundries.com

<sup>†</sup> These authors contributed equally

**Abstract**—This paper presents a dendrite-like device that discriminates spatiotemporal patterns of pulses for parallel processing in 3D neuromorphic architectures. The device utilizes the ferroelectric layer in a segmented multi-gate FeFET design to detect a consecutive sequence of input pulses. Experimental results demonstrate successful emulation of highly selective sequence discrimination in dendrites of neurons in the cortex and highlight up to 100× signal-margin (output current differences). This nanodendrite design offers a neuromorphic solution to thermally scalable parallel processing in 3D systems.

## I. INTRODUCTION

With 2D device scaling reaching an apparent wall, innovative techniques to increase device density and computation capacity are imperative. A common theme is designing 3D devices to increase density and 3D architectures to increase capacity. Computation based on binary logic is not entirely suited to this shift from 2D to 3D integration. Binary computation activates one of every two wires per cycle. Thus, it generates heat at a rate that scales as volume ( $n^3$ ), but this heat dissipates at a rate that as surface ( $n^2$ ). Consequently, heat dissipation becomes incommensurate with heat generation. To restore the balance, activity must scale as  $1/n$ . For instance, in a stack of  $n$  layers, the fraction of its wires that each layer activates must decrease as  $1/n$ . This  $1/n$  activity scaling in 3D systems is fundamentally incompatible with binary signaling [1].

This paper introduces a device that performs logic not on binary-valued signals, such as a high or low voltages, but rather on unary-valued signals, such as voltage pulses. When voltage pulses are applied to  $m$  out of  $n$  wires, keeping  $m$  constant as  $n$  increases scales activity as  $1/n$ . This  $m$ -of- $n$  signaling encodes information most efficiently as a spatiotemporal pattern of  $m$  pulses on  $n$  wires. To discriminate such spatiotemporal patterns, we designed, fabricated, and experimentally demonstrated a *nanodendrite* that emulates nonlinear spatiotemporal processes in dendrites of neurons in the cortex. The conductance between the nanodendrite's source and drain terminals switches from a low value to a high value only when its gate terminals are pulsed sequentially, from the source to the drain (**Fig. 1**). It is the first nanoscale device to emulate sequence detection in a stretch of dendrite and paves the way to exploiting recent progress in understanding the brain to solve the unique thermal constraint faced by 3D systems [1].

Our nanodendrite design utilizes charge dipoles in a ferroelectric layer of a segmented multi-gate Field-Effect Transistor (FET). It was fabricated in GlobalFoundries conventional 28nm HKMG CMOS embedded FeFET technology without any modification [3] (**Fig. 2**). Here, we present results from a 3-gate version of our multi-gate FeFET design (**Fig. 3**) and demonstrate that its threshold voltage transitions from a high- $V_T$  state to a low high- $V_T$  state only when its gates are pulsed in the correct order.

## II. OPERATION PRINCIPLE

The nanodendrite relies on the two fundamental principles of FeFET operation. First, the ferroelectric layer's polarization will only flip if the electric field ( $\mathcal{E}_F$ ) within it exceeds a critical threshold known as the coercive field ( $\mathcal{E}_C$ ). Second, for  $\mathcal{E}_F$  to exceed  $\mathcal{E}_C$ , the channel must be in inversion rather than deep depletion (**Fig. 4**) [4]. The initial polarization direction puts the device in the high- $V_T$  state and the channel in deep depletion. To change the channel's state from depletion to inversion, inversion carriers must be supplied either by the source or the substrate, but the latter relies on thermal generation, a slow process.

Consider a FeFET with three segments, S1, S2, and S3, controlled by three gates, G1, G2, and G3 (see **Fig. 3**). These gates receive pulses P1, P2, and P3, respectively. When P1 is applied to G1, the source supplies inversion carriers to S1's channel and thus its state changes from depletion to inversion (**Fig. 5a, left**). That allows  $\mathcal{E}_F$  to exceed  $\mathcal{E}_C$  and flip S1's polarization. Hence, S1 transitions to the low- $V_T$  state. Now when P2 is applied to G2, S1's channel supplies inversion carriers to S2's channel and thus its state also changes from depletion to inversion (**Fig. 5a, middle**). That allows  $\mathcal{E}_F$  to exceed  $\mathcal{E}_C$  and flip S2's polarization. Hence, S2 also transitions to the low- $V_T$  state. Now when P3 is applied to G3, S2's channel supplies inversion carriers to S3's channel and thus its state also changes from depletion to inversion (**Fig. 5a, right**). That allows  $\mathcal{E}_F$  to exceed  $\mathcal{E}_C$  and flip S3's polarization. Hence, S3 also transitions to the low- $V_T$  state. With all 3 segments in the low- $V_T$  state, the drain-source conductance drops to a low value.

Now consider a different pulse sequence, such as P2, P1, P3 (**Fig. 5b**). When P2 is applied G2, S2's channel does not receive inversion carriers from S1's channel. Hence, it cannot change state from deep depletion to inversion. Thus,  $\mathcal{E}_F$  does

not exceed  $\mathcal{E}_c$  and S2's polarization does not flip. Hence, S3 remains in the high- $V_T$  state. Thus, the drain-source conductance will remain low, even if P1 and P3 transition S1 and S3 to the low- $V_T$  state, leading to a low drain current.

**Fig. 6** presents the outcomes of Sentaurus TCAD simulations, illustrating the characteristics of multi-gate FeFET including electron density, polarization, and electric field. The highly-doped regions between the gates would ensure the connectivity of gaps between the channel regions. If this  $n^{++}$  implant is blocked, then fringing electric fields between gate gaps couples the inversion layer of the preceding gate to the channel of the pulsed gate and helps establish a continuous conducting channel from the source to the drain that leads to a high drain current.

### III. SEQUENCE DETECTION

#### A. Single-gate FeFET and deep depletion condition

The transfer characteristics of a single-gate FeFET were examined (**Fig. 7**). The device exhibited a hysteresis of 0.5 V when sweeping from  $V_G = -2$  V to 2 V and 2 V to -2 V, with high- $V_T$  and low- $V_T$  states at 1 V and 0.5 V, respectively.

We first examine the conditions under which the deep depletion state would prevent dipole flipping. The effect of pulse width on the generation of the inversion layer was investigated with the source/drain grounded or floating. We investigate various pulse widths applied to the single-gate FeFET. **Fig. 8 and 9** illustrate the single-gate FeFET in both the high- $V_T$  and low- $V_T$  states, which were pulsed by applying a 3.5 V pulse for various pulse width with the source/drain either grounded or floating. The FeFET is initialized to a high- $V_T$  state by a pulse of -3.5 V for a duration of 5  $\mu$ s, while the source and drain are grounded. When the source/drain were grounded (**Fig. 8**), an inversion layer can be formed under the gate and the coercive field of the FE-HfO<sub>2</sub> is exceeded, causing a transition from a high- $V_T$  to low- $V_T$  state for the FeFET. Such behavior is observed even with pulses down to 0.1  $\mu$ s.

In the case of floating source/drain (**Fig. 9**) The minority carriers are not supplied from the source/drain but rather has to be thermally generated from the p-type substrate. State transitions occurred only for pulse widths exceeding 1  $\mu$ s (**Fig. 10**). Thus, pulse widths shorter than 1  $\mu$ s are necessary to maintain the deep depletion conditions.

#### B. Sequence detection demonstrated by 3-gate FeFET

We demonstrate the time-dependent sequence detection function and keep monitoring the drain current as shown in **Fig. 11** (the time scale is not plotted according to actual time). Initially, we establish the high- $V_T$  state by applying a -3.5 V pulse for 5  $\mu$ s with the three gates tied together.

The cycle of the FeFET device starts with the sequential application of pulses to G1, then G2 followed by G3. In the readout period between the pulses, these three gates are interconnected and biased at 0.5 V to measure the conductance of the FeFET channel. The drain voltage is maintained at 0.1 V to monitor any changes in the current. The cycle concludes after G3 receives the pulse. Subsequently, the output current

increased to  $10^{-7}$  A is observed, attributable to the formation of a continuous conducting channel from the source to the drain.

In the next cycle, the high- $V_T$  state is first established by applying a -3.5 V pulse for 5  $\mu$ s while reconnecting the three gates together. The incorrect pulse sequence is initiated by pulsing G2 then G1 followed by G3. In this case, the output current slightly increases and reaches a value of  $5 \times 10^{-9}$  A after the cycle ends with the pulse on G3. A difference of  $10^{1.5}$  in measured current is obtained for the correct sequence compared to the incorrect sequence. This disparity emphasizes the nanodendrite's capability for sequence discrimination.

**Fig. 12 and 13** present the transfer characteristics of the 3-gate FeFET, demonstrating the impact of different pulse widths in both the correct and incorrect sequences. Consistent with the single-gate FeFET analysis, it is crucial to complete the pulse procedure prior to the thermal generation of minority carriers to avoid the formation of an inversion channel when presented with the incorrect sequence. **Fig. 14** indicates that the pulse duration should be limited to less than 1  $\mu$ s to maintain high sequential discrimination capability.

To enhance the sequential discrimination capability, an alternative approach is to omit the highly-doped regions between the gates. This modification leads to a higher output current difference, reaching  $10^2$  (**Fig. 15**) and higher range of  $V_T$  state discrimination (**Fig. 16**).

### IV. CONCLUSIONS

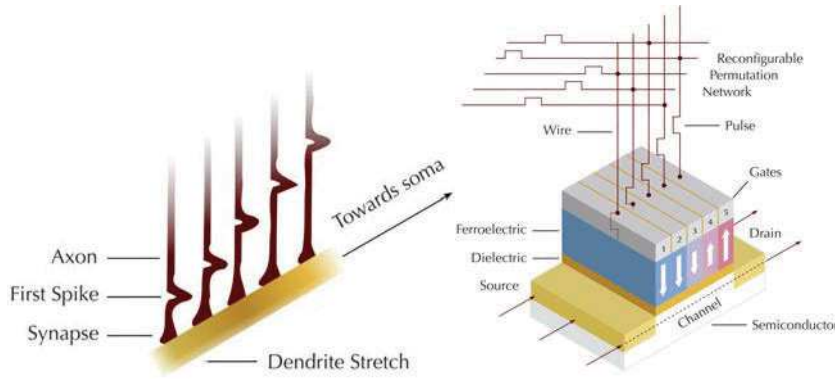
We demonstrate successful detection of spatiotemporal sequences using a multi-gate FeFET nanodendrite fabricated in a conventional GF 28-nm process. The multi-gate FeFET marks a significant advancement in parallel processing within 3D neuromorphic architectures, addressing the thermal challenges associated with traditional computation methods. We show a significant sense signal margin, highlighting the highly selective sequence discrimination capabilities of the nanodendrite device, reaching up to 100 times output current difference between the correct and incorrect pulse sequences. This research provides a stepping stone for further exploration and optimization of the nanodendrite device.

#### ACKNOWLEDGMENT

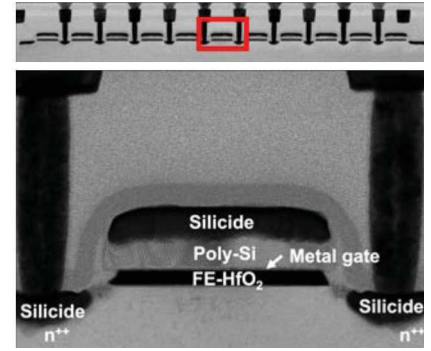
The authors gratefully acknowledge the contributions of Tarek Ali, Halid Mulaosmanovic, Steve Soss, Stefan Duenkel, Sven Beyer to provide access and guidance to GlobalFoundries 28nm FeFET technology. H.C. appreciates Koustav Jana (Stanford U.) for the invaluable and insightful discussions. This work is supported in part by the NSF BRAID (2223827), Stanford HAI Hoffman-Yee Research Grants and AFOSR ExPlor Center (FA9550-22-1-0532).

#### REFERENCES

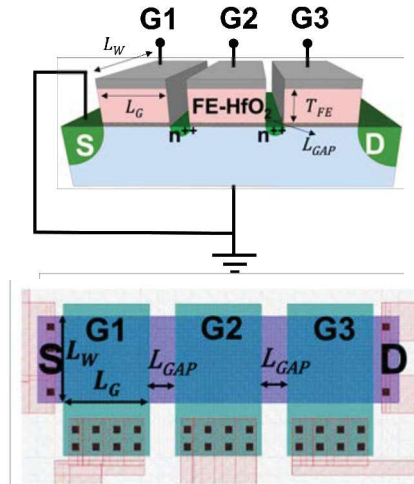
- [1] Boahen, K., *Nature*, vol. 612, pp. 43–50, (2022).
- [2] L. Le Coeur, et al., in *Proc. ESANN* (2023).
- [3] S. Beyer et al., in *IEEE International Memory Workshop* (2020).
- [4] K. Toprasertpong, et al., *Appl. Phys. Lett.*, vol. 116, 242903, (2020).



**Fig. 1:** *Left* – Detecting a spike sequence in the brain. Axons route the first spike in the sequence to the left of a stretch of dendrite, followed by the second spike, etc. [1]. These spikes consecutively propel a plateau potential along the stretch towards the soma. *Right* – Detecting a pulse sequence with the nanodendrite. A reconfigurable permutation network routes the first spike to the first gate, the second spike to the second gate, etc. [2]. These pulses consecutively flip dipoles from up to down, and thus create a conductive channel from the source to the drain [2].

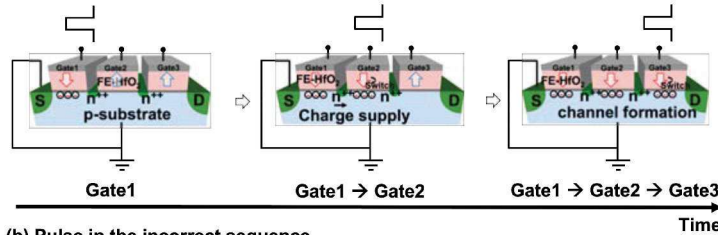


**Fig. 2:** *Top* – TEM cross-section of a multi-gate FeFET. This example has 11 gates with  $n^{++}$  source-drain implants in the gaps between that receive metal contacts. *Bottom* – Gate-stack highlighted above (red box) with layers of silicide, polysilicon (Poly-Si), metal gate, and ferroelectric-doped hafnium oxide (FE-HfO<sub>2</sub>).

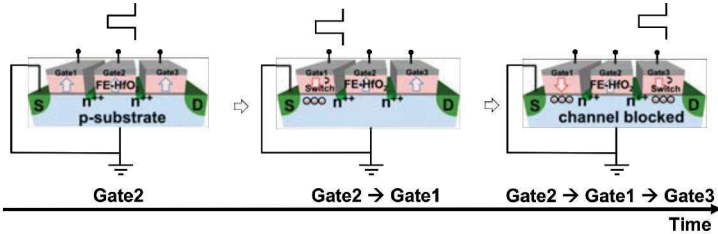


**Fig. 3:** Cross section (*top*) and layout (*bottom*) of 3-gate FeFET studied in this paper. Note that the  $n^{++}$  regions between gates are not contacted.

**(a) Pulse in the correct sequence**

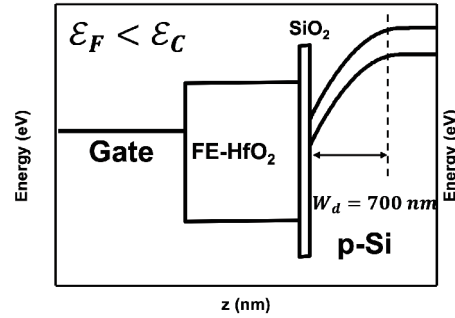


**(b) Pulse in the incorrect sequence**

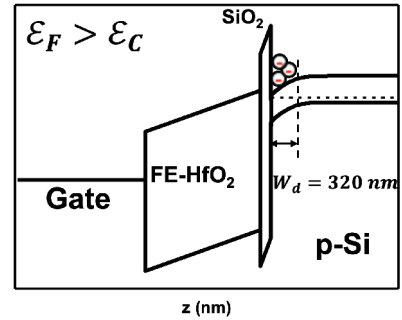


**Fig. 5:** Schematic diagram of (a) a correct pulse sequence and (b) an incorrect pulse sequence. In the correct sequence, inversion carriers are supplied from the source to G1, from the inversion layer of G1 to G2, and from the inversion layer of G2 to G3. Consequently, all the dipoles are flipped.

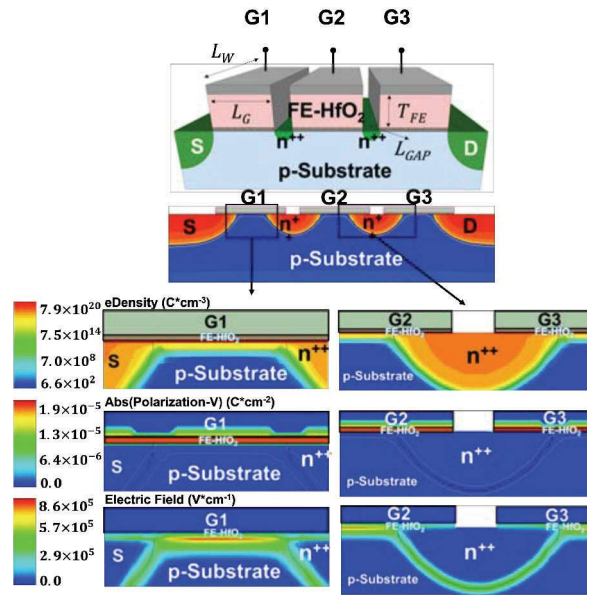
**(a) Deep depletion**



**(b) Inversion**



**Fig. 4:** Band diagrams across a gate-stack. (a) In deep depletion, the voltage applied to the gate drops primarily across the substrate, keeping the electric field ( $\mathcal{E}_F$ ) in the FE-HfO<sub>2</sub> layer from surpassing the coercive field ( $\mathcal{E}_C$ ). (b) In inversion, the voltage applied to the gate drops primarily across the ferroelectric layer, allowing  $\mathcal{E}_F$  to surpass  $\mathcal{E}_C$  and flip the dipoles.



**Fig. 6:** Sentaurus TCAD simulation of electron density, polarization, and electric field. Inversion channel formation after G1, G2, G3 were pulsed sequentially. The fringing field and the highly-doped regions ensure the connectivity of gaps between the channel regions.



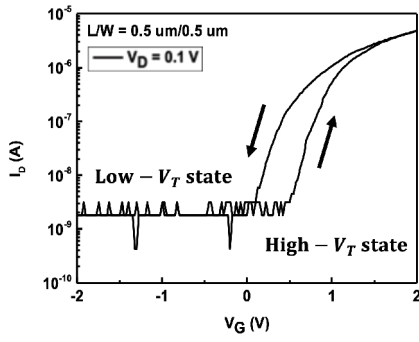


Fig. 7. The transfer characteristics of a single-gate FeFET. The device exhibited a hysteresis of 0.5 V when sweeping from  $V_G = -2$  V to 2 V and 2 V to -2 V, with high- $V_T$  and low- $V_T$ .

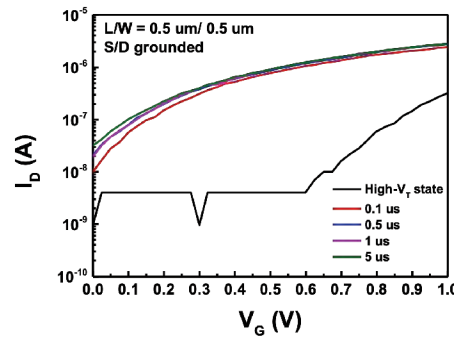


Fig. 8. Single-gate FeFET in High- $V_T$  state and low- $V_T$  state pulsed by 3.5 V pulse for various time with source/drain grounded. Even pulses down to 0.1  $\mu$ s transited the state and flipped the dipole.

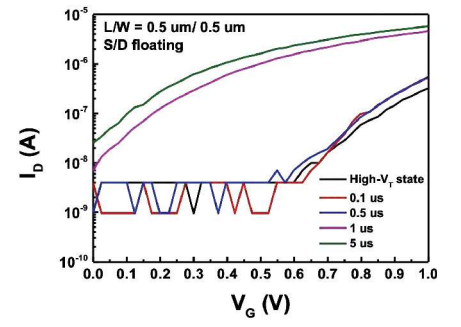


Fig. 9. Single-gate FeFET in High- $V_T$  state and low- $V_T$  state pulsed by 3.5 V pulse for various time with source/drain floating. State transitions occurred for pulse widths exceeding 1  $\mu$ s. The minority carriers came not supplied from the source/drain but rather thermally generated from the p-type substrate.

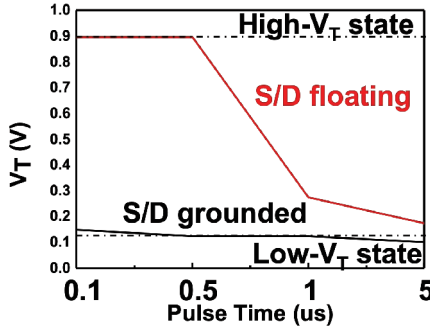


Fig. 10. The relationship between  $V_T$  and pulse time. Pulse widths shorter than 1  $\mu$ s are necessary to maintain high- $V_T$  state and the deep depletion conditions.  $V_T$  extracted at current equals to  $1 \times 10^{-7}$  A.

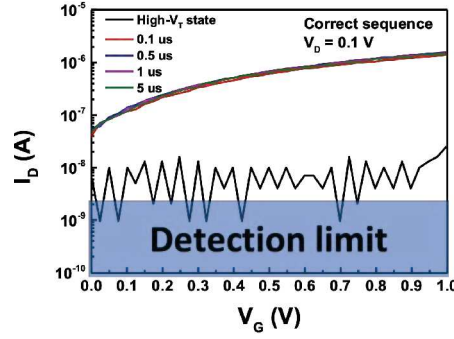


Fig. 12. Read of transfer characteristics of 3-gate FeFET with various pulse width in the correct sequence. Even the shortest pulse width 0.1  $\mu$ s can flip the dipole if the sequence is correct.

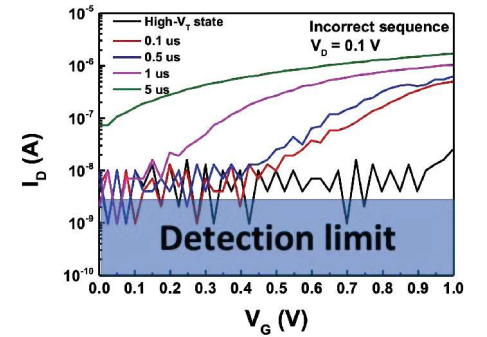


Fig. 13. Read of transfer characteristics of 3-gate FeFET with various pulse width in the incorrect sequence. The pulse process should be finished less than 1  $\mu$ s to ensure high sequential discrimination capability.

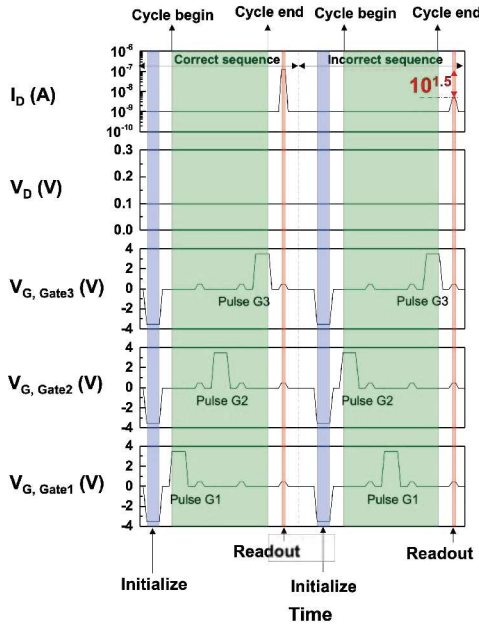


Fig. 11. Time-dependent  $V_G$ ,  $V_D$ , and  $I_D$  in standard 3-gate FeFET (the time scale is not plotted according to actual time). In the correct pulse sequence, the output current shows a substantial increase, with a difference of  $10^{1.5}$  compared to the incorrect sequence. This significant disparity underscores the nanodendrite's ability to discriminate between different pulse sequences.

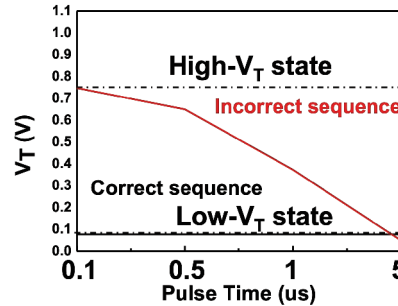


Fig. 14. The relationship between  $V_T$  and pulse time for 3-gate FeFET. Pulse widths shorter than 1  $\mu$ s are necessary to maintain high- $V_T$  state and the deep depletion conditions also in 3-gate FeFET.  $V_T$  extracted at current equals to  $1 \times 10^{-7}$  A.

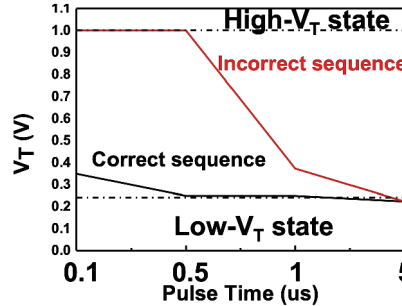


Fig. 16. The relationship between  $V_T$  and pulse time for 3-gate FeFET without highly-doped regions between gates. This device achieves a higher level of discrimination in the  $V_T$  states, allowing for improved differentiation between different voltage threshold levels.  $V_T$  extracted at current equals to  $1 \times 10^{-7}$  A.

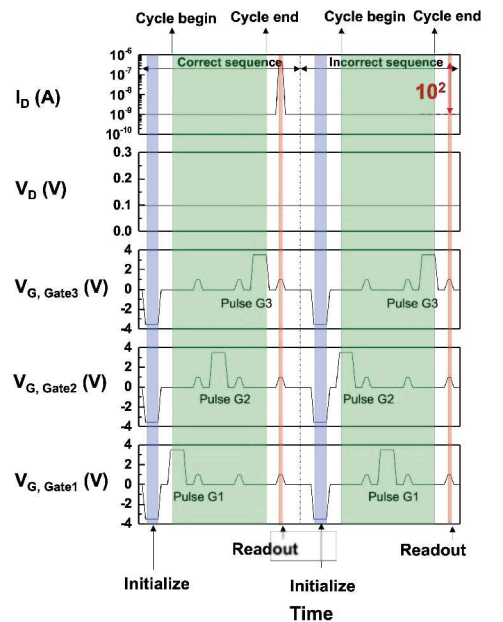


Fig. 15. Time-dependent  $V_G$ ,  $V_D$ , and  $I_D$  in 3-gate FeFET without highly-doped between gates (the time scale is not plotted according to actual time). The sequential discrimination capability can be further enhanced, resulting in an increased output current difference of  $10^2$ .