# Vertically Stacked Nanosheet FET: Charge Trapping Memory and Synapse with Linear Weight Adjustability for Neuromorphic Computing Applications



| | |
|---|---|
| Journal: | *Transactions on Electron Devices* |
| Manuscript ID | TED-2022-09-2185-R |
| Manuscript Type: | Regular |
| Date Submitted by the Author: | 06-Sep-2022 |
| Complete List of Authors: | Raza Ansari, Hasan; King Abdullah University of Science and Technology, CEMSE<br>Li, Hanrui; King Abdullah University of Science and Technology<br>El-Atab, Nazek; King Abdullah University of Science and Technology, Computer Electrical Mathematical Sciences and Engineering |
| Area of Expertise: | Memory devices, Neuromorphic computing, Nanosheet FET |
| | |




SCHOLARONE™
Manuscripts

# Vertically Stacked Nanosheet FET: Charge Trapping Memory and Synapse with Linear Weight Adjustability for Neuromorphic Computing Applications

Md. Hasan Raza Ansari, *Member, IEEE*, Hanrui Li, *Student Member, IEEE*, Nazek El-Atab, *Senior Member, IEEE*

***Abstract*— This work analyzes a vertically-stacked nanosheet field effect transistor (NSFET) for non-volatile charge trapping memory and artificial synaptic devices. The artificial synapse's operation, long-term potentiation (LTP), and depression (LTD) are analogous to Erase (ERS) and Program (PGM) of charge trapping memory, respectively. The NSFET device with a gate length of 50 nm achieves a wider memory window, long retention time for the Program, and infinite retention (>$10^8$ s) for Erase operation. The results also show linear synaptic features with non-linearity values of 2.5 and -0.42 for LTP and LTD, respectively. Furthermore, the device conductance values are utilized as synaptic weights for image recognition of MNIST datasets in neural networks and achieve 93.30 % accuracy. The results reveal that vertically stacked NSFET is suitable for next-generation charge trapping memory and neuromorphic computing due to its wider memory window, long retention, better accuracy, and high density.**

***Index Terms*—NSFET, Synaptic Transistor, LTP, LTD, Neural Network, neuromorphic computing.**

## I. Introduction

The bottleneck issue of data transfer between memory and processor of conventional von-Neumann computing can be resolved by neuromorphic computing through the integration of memory and computing functions in a single device [1], [2]. Artificial synapse is an essential building block for neuromorphic computing, which mimics human behaviors such as image recognition and detection and can be implemented in neural network systems [3], [4]. Synapses are used to transfer the information between pre-neurons to post-neurons and show the synaptic strength of the brain [3], [4]. The essential features of synapses are spike time-dependent plasticity (STDP), short-term potentiation (STP), long-term potentiation (LTP), and long-term depression (LTD). These features have been implemented through two terminal memristive (Resistive RAM [5], [6], Phase Change Memory [7], [8], and Magnetic RAM) [9], and as well as three-terminal devices (Charge Trapping Memory [10], [11] and Ferroelectric Transistors [12]).

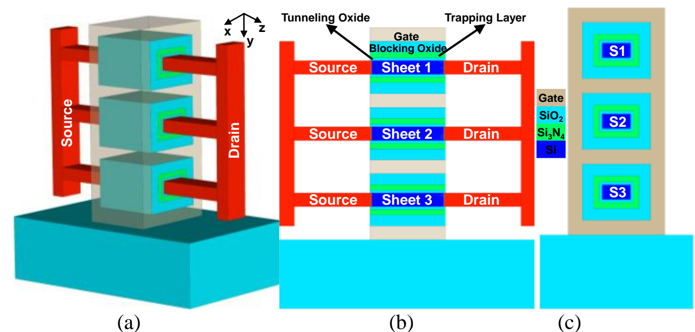



Fig. 1. Schematic representation of vertically Nanosheet FET for charge trapping memory and synaptic transistor in 3D, (b) vertical (y-z plane), and (c) horizontal (y-x plane) of 3D. S1, S2, and S3 indicate Sheet 1, Sheet 2, and Sheet 3 of vertical NSFET.

The scaling of the devices becomes necessary for large-scale integration of neural networks. The issues with these devices are scaling and other limitations, such as reliability. More recently, academia and industries have revealed the potential of vertically-stacked nanosheet field effect transistors for the next generation logic and memory applications [13]–[18]. In addition, the epitaxial process controls NSFET by stacking multiple nanosheets in a vertical direction and the spacing between sheets [19]. The increase in gate controllability of NSFET compared to finFet, Nanowire, and planer MOSFETs results in less short channel effect, which is beneficial not only for logic application but also for memory application [20], [21]. The advantages of nanosheet FETs include better gate controllability, larger effective channel width, thus high current density, and also compatibility with the crossbar array architecture and 3-D integration [13]–[15]. These advantages make the transistor feasible for synaptic devices to achieve high conductance values for LTP and LTD.

In this work, we have shown the feasibility of a vertically stacked nanosheet field effect transistor as a charge trapping memory (CTM) and emulate synaptic properties such as long-term potentiation (LTP) and depression (LTD). The device is

This work was supported by the King Abdullah University of Science and Technology baseline fund.

The authors are with the SAMA Labs, Electrical and Computer Engineering, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia. (corresponding author: nazek.elatab@kaust.edu.sa

optimized to achieve high retention time and endurance characteristics for non-volatile memory. In addition, a synaptic device shows better device conductance linearity (non-linearity value of 2.5 for LTP and -0.42 for LTD) with a better dynamic range (maximum conductance/minimum conductance ($G_{max}/G_{min}$) > $10^3$). The estimated conductance values are utilized for artificial neural networks, which resulted in 93.36% accuracy of Modified National Institute of Standards and Technology (MNIST) image datasets [22].

TABLE. I
DEVICE PARAMETERS FOR NSFET AS CTM AND SYNAPSE

| Device Parameters | Values |
|---|---|
| Gate Length ($L_g$) | 50 – 20 nm |
| Silicon width ($W_{Si}$) | 20 nm |
| Silicon Height ($H_{Si}$) | 10 nm |
| Blocking Oxide ($SiO_2$) | 8 nm and 6 nm |
| Charge Trapping Layer ($Si_3N_4$) | 4 nm |
| Tunneling Oxide ($SiO_2$) | 2 nm |
| Number of sheets | 3 |
| Total height | 134 nm |
| Temperature | 300 K |
| Source/Drain doping ($N_{D(S/D)}$) | $10^{20}$ cm$^{-3}$ |
| Channel region doping (*n*-type) | $10^{17}$ cm$^{-3}$ |

TABLE. II
MATERIAL PARAMETERS FOR NANOSHEET FET BASED CHARGE TRAPPING MEMORY AND SYNAPTIC TRANSISTOR

| Parameters | Values |
|---|---|
| Trap density ($N_T$) | $10^{19}$ cm$^{-3}$ |
| CCS area | $10^{-13}$ cm$^{-2}$ |
| $m^*$ of Si | 0.36 $m_o$, 0.38 $m_o$ |
| $m^*$ of $SiO_2$ | 0.50 $m_o$, 0.25 $m_o$ |
| $m^*$ of $Si_3N_4$ | 0.36 $m_o$, 0.38 $m_o$ |
| $\varepsilon_{SiO2}$ | 3.9 |
| $\varepsilon_{Si3N4}$ | 7.5 |
| $\mu_{Si3N4}$ | 1.0 cm$^2$ / V s |
| Poole-Frankel trap | $10^{15}$ |

$m^*$, $m_o$, and CCS indicate effective mass, free electron mass, and capture cross section in the nitride layer. Si3N4 is considered a semiconductor in the simulation to trap electrons and holes [23].

## II. DEVICE DESCRIPTION AND METHODOLOGY

Fig. 1(a) represents the schematic representation of a *n*-type inversion mode vertically stacked nanosheet field effect transistor (NSFET) in 3D, Fig. 1(b) vertical (*y-z* plane), and Fig. 1(c) horizontal (*y-x* plane) planes of the 3D charge trapping memory (CTM) and artificial biological synapse. The device consists of three vertically stacked nanosheets. The device is simulated through a 3D Synopsys Sentaurus TCAD simulator with the optimized device parameters, as illustrated in Table. I [23]. Thanks to NSFET and the electrically coupled nanosheets, the impact of the external field is identical on every sheet. In order to perform the charge trapping and de-trapping mechanism, we have incorporated the charge trapping models, with Fowler–Nordheim (FN) tunneling and Poole-Frankel model. The basic drift-diffusion models include concentration-dependent mobility, carrier lifetime SRH models, and dynamic non-local band tunneling, models. The material parameters used for the simulations are illustrated in Table II and adopted from [24].

## III. NSFET BASED CHARGE TRAPPING MEMORY

The memory operations are performed through the F-N tunneling mechanism [25], [26]. The program (PGM) operation is performed by applying a positive (10 V) gate voltage due to thicker oxide, which allows the channel electrons to tunnel through tunneling oxide and traps them in the nitride layer. The contour plot of electron barrier tunneling during PGM operation is shown in Fig, 2(a). Erase (ERS) operation is performed by applying a negative gate voltage (-8 V) for 10 µs, which de-traps the electrons from the nitride layer. The contour plot of hole barrier tunneling during ERS operation is shown in Fig, 2(b). The trapped electrons and holes in the nitride layer contribute positive and negative potential on the channel of the transistor after PGM and ERS operation [25], [26]. Trapped electrons increase the barrier while holes lower the barrier for source electrons, which can be confirmed through the conduction band (CB) energy and contour plots of electron density (eDensity) during read operation as shown in Fig. 3(a). The energy band diagram is extracted 1 nm below of the tunneling oxide of top sheets. Fig. 3(b) shows the transfer characteristics of the device after PGM and ERS operations for different blocking oxide thicknesses (6 nm and 8 nm) with a gate length ($L_g$) of 50 nm at a drain voltage of 1.0 V. The memory window ($\Delta V_{th}$) is wider (1.86 V) for thicker blocking oxide due to the higher electric field at the blocking and nitride interface, which enhances the barrier for trapped carriers. Fig. 3(b) shows the retention time of NSFET charge trapping memory in terms of the threshold voltage ($V_{th}$) of PGM and ERS operation.

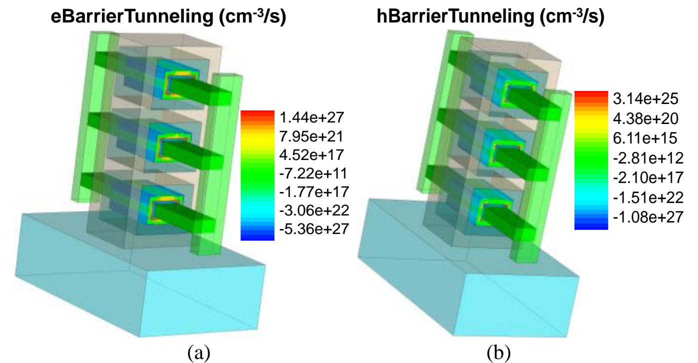



Fig. 2. Contour plots for (a) electron barrier tunneling (eBarrier Tunneling) during the Program and (b) hole barrier tunneling (eBarrier Tunnelin) during Erase operation. The contour plots are extracted at the end of the operation.

The threshold voltage is estimated through the maximum transconductance technique [26]–[28]. The retention time of trapped charges in the nitride layer is estimated by applying zero bias at all the electrodes and monitored for $3.15\times10^8$ seconds (10 years) at 300 K. The threshold voltage of PGM operation with thicker blocking oxide maintains up to 78.35 % of the initial $V_{th}$. The threshold voltage of ERS operation is maintained for an infinite time (longer than 10 years). The retention time of thicker blocking oxide is higher than the thinner blocking oxide. Thinner blocking oxide maintains 73.45% of initial voltage after years, while thicker blocking oxide maintains 78.85 %. The endurance of the NSFET based

charge trapping memory is shown in Fig 3(c) and estimated through program/erase (P/E) cycles. The device achieves the same $V_{th}$ after $10^3$ program/erase cycles, which confirms that the device has better endurance. Fig. 3(e) shows the gate length ($L_g$) scaling of NSFET device as CTM in terms of variation in $V_{th}$ and memory window after PGM and ERS operation. The device can be scaled down to 15 nm with achieving optimum memory window of 1.3 V. These results revealed that the device could be an alternate for next generation charge trapping memory applications. The NSFET based charge trapping memory achieves longer retention for PGM operation and infinite retention for ERS operation with endurance > $10^3$ and is feasible for the analog synaptic device.

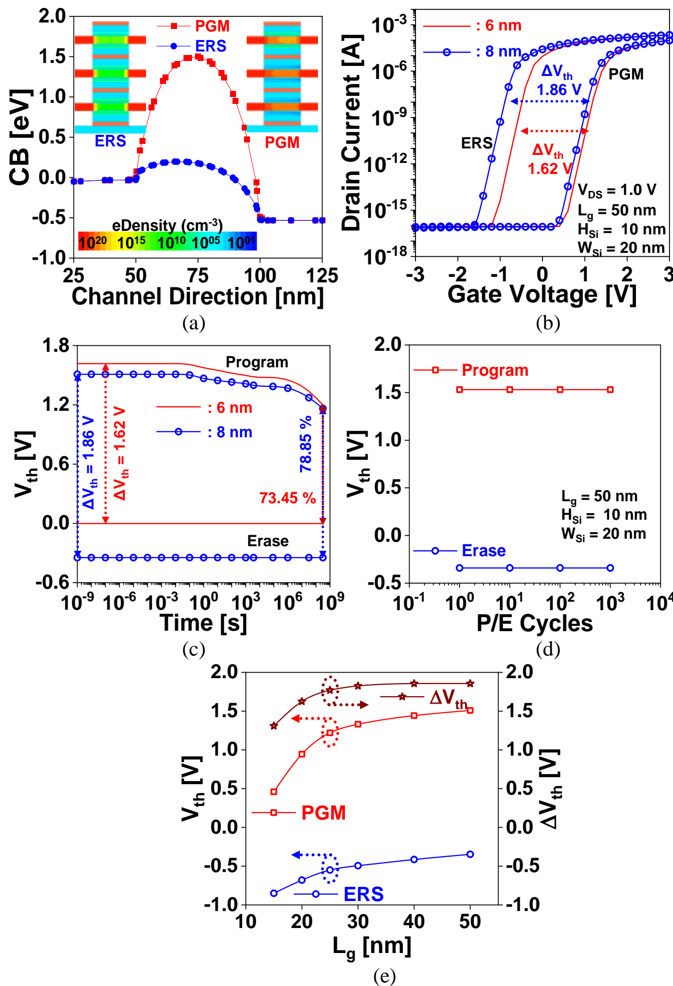



Fig. 3. (a) Conduction band energy during read operation after Program (PGM) and Erase (ERS) operation. Variation in (b) Transfer characteristic and (c) threshold voltage ($V_{th}$) of the device after PGM and ERS operation of the charge trapping memory for different blocking oxide at gate length of 50 nm. (d) Variation in $V_{th}$ after PGM and ERS for program/erase cycles. (e) Dependence of $V_{th}$ and memory window ($\Delta V_{th}$) with gate length ($L_g$).

## IV. NSFET for Neuromorphic Computing

Fig. 4 shows the schematic representation of an artificial synapse with vertically stacked nanosheet field effect transistor. In order to mimic biological behaviors (long term potentiation (LTP) and depression (LTD)) with vertically stacked NSFET, a repetitive pulse for potentiation and depression is applied. We have used 40 pulses for potentiation and depression. In the NSFET device, the gate and drain electrodes behave as pre- and post- synapses, respectively. Potentiation of the synapse in the brain is analogous to the Erase operation of charge trapping memory [10]. The potentiation operation is performed by applying identical pulses with a positive drain voltage ($V_{DS}$ = 3V) and negative gate voltage ($V_{GS}$ = -4V) with a pulse and interval time of 500 ns as shown in the inset of Fig. 5(a).

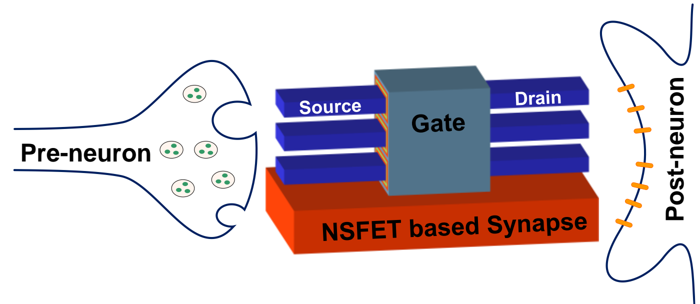



Fig. 4. Schematic representation of NSFET based artificial synapse for neuromorphic applications.

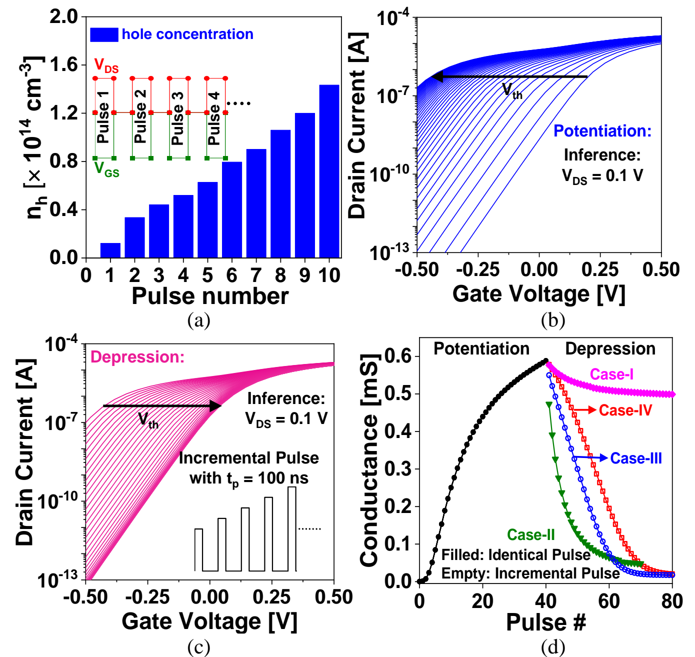



Fig. 5. (a) Variation of trapped hole concentration ($n_h$) in the nitride layer with pulse number. Transfer characteristics during inference operation after each pulse of (b) potentiation and (c) depression at a drain voltage of 0.1 V (d) Variation of conductance values of LTP and LTD with identical and incremental pulse time with pulse number.

The applied bias generates electron-hole pairs in the semiconductor. Due to a sufficient electric field at the oxide interface, some holes get trapped in the nitride layer [26]. Fig. 5(a) shows the trapped holes ($n_h$) in the nitride layer with pulse number during the potentiation. An increase in the pulse number increases the $n_h$ in the nitride and thus increases the channel conductivity, similar to strengthening the synaptic plasticity [29]. Fig. 5(b) shows the transfer characteristic of NSFET CTM device after each potentiation pulse. The trapped charges in the nitride layer reduce the threshold voltage ($V_{th}$) at each pulse of the potentiation and thus increase the conductance with increasing pulses. The main focus of this

work is to adjust the non-linearity conductance issue during long term depression. The depression operation is based on F-N tunneling. It is performed by applying a positive gate voltage with identical and incremental pulse voltage and different pulse times (100 ns and 500 ns). Case-I → identical pulse voltage ($V_{GS}$ = 2) with a pulse time of 500 ns. Case-II → identical pulse voltage ($V_{GS}$ = 3) with a pulse time of 500 ns. Case-III → incremental voltage pulse $V_{GS}$ starts from 2 V and step voltage of 50 mV with a pulse time of 500 ns. Case-IV → incremental voltage pulse $V_{GS}$ starts from 2 V and step voltage of 50 mV with a pulse time of 100 ns. Applying a positive voltage de-traps the holes from the nitride layer at each pulse and thus increases the threshold voltage as shown in Fig. 5(c). The conductance values (synaptic weight) at each pulse for neural networks are estimated from Fig. 5(b) and (c). Fig. 5(d) shows the conductance values during potentiation (LTP) and depression (LTD) with pulse number. As shown in Fig. 4(d), for Case-IV, the saturation in conductance values occurs after the 75$^{th}$ pulse, which conveys that lower pulse time with incremental pulse enables a higher available conductance state with the same memory window and controlling of synaptic weights.

In synaptic weight updates for LTP, a large dynamic range ($G_{max}/G_{min}$) ~= 2×10$^3$ is achieved where $G_{max}$ = 0.56 mS and $G_{min}$ = 216 nS are the maximum and minimum conductance values). With a large dynamic range > 10, linear, and symmetrical conductance updates are essential criteria for neural network applications [30]. Fig. 6 shows the non-linearity in conductance values (synaptic weights) are estimated through MATLAB simulation with derived equations as in [30], [31].

$$G_{LTP} = B\left[1 - e^{\left(\frac{-P}{A}\right)}\right] + G_{min} \quad (1)$$

$$G_{LTD} = -B\left[1 - e^{\left(\frac{P-P_{max}}{A}\right)}\right] + G_{max} \quad (2)$$

$$B = \frac{(G_{max} - G_{min})}{\left(1 - e^{\left(\frac{-P_{max}}{A}\right)}\right)} \quad (3)$$

where $G_{LTP}$, $G_{LTD}$, $G_{min}$ and $G_{max}$ indicate the potentiation, depression, minimum and maximum synaptic device conductance value. $P_{max}$ represents the maximum number of pulses applied for the device to transition from the lowest to highest conductance states. B is a fitting factor that is a function of A.

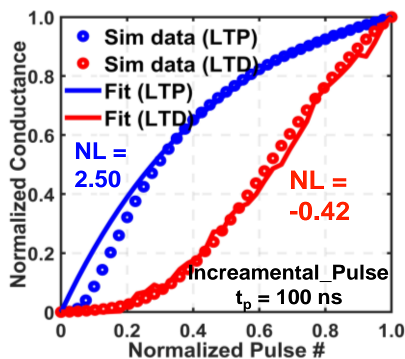


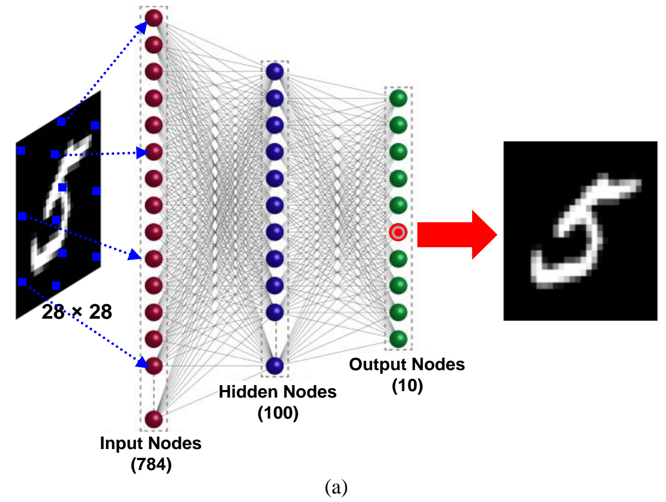

Fig. 6. Non-linearity calculation of NSFET synaptic device conductance values for an incremental pulse with a pulse time of 100 ns.

In order to achieve better accuracy for online training, linearity and symmetry The results show that the incremental pulse with lower pulse time shows better linearity compared to other cases as demonstrated in Table III. The NSFET device achieves non-linearity value of 2.5 for LTP and -0.42 for LTD.

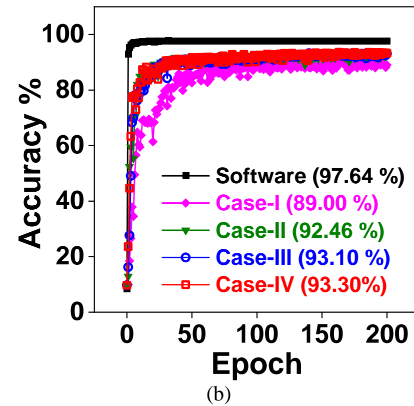
TABLE. III
NON-LINEARITY CALCULATION OF LTP AND LTD SYNAPTIC DEVICE WEIGHT VALUES.

| Cases | LTP | LTD |
|---|---|---|
| Case I | 2.50 | -6.20 |
| Case II | 2.50 | -5.10 |
| Case III | 2.50 | -2.23 |
| Case IV | 2.50 | -0.42 |

Case I→ Identical with $V_{GS}$ = 2 V and $t_p$ of 500 ns. Case II→ Identical with $V_{GS}$ = 3 V and $t_p$ of 500 ns. Case III→ Incremental starts with $V_{GS}$ = 2 V and $t_p$ of 500 ns. Case IV→ Incremental starts with $V_{GS}$ = 2 V and $t_p$ of 100 ns.




(a)




(b)

Fig. 7. (a) Schematic of the three-layer neural networks with the vertically NSFET as the synapse for MNIST digit recognition. (b) Variation of digit recognition accuracy (%) with the number of training epochs for different cases of LTP.

In order to verify the learning capabilities of the NSFET based synaptic device for hardware-based neural networks, a three-layer artificial neural network with input, one hidden, and output layers is shown in Fig. 7(a). The designed neural network is used to classify image data from a Modified National Institute of Standards and Technology (MNIST) dataset, which consists of 60,000 training images and 10,000 test images of handwritten digits from “0” to “9” [22]. It is simulated with 28 × 28 pixels of MNIST datasets. The

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

conductance values extracted from the device for all the cases are utilized as synaptic weights for image recognition in the algorithm. The image is normalized in pixel intensities in the interval of 0 and 1 as shown in Fig. 7(a). The normalized pixel is transformed into a column matrix with 784 elements, which are then fed into the neural network. Rectified Linear Unit (ReLU) activation is used in network simulation. The results indicate that the better linearity of LTD has better recognition accuracy. The recognition accuracies for Case-IV after 200 training epochs reach 93.30 %, which is only 3 % lower than the ideal software-based neural network accuracy as shown in Fig 7(b). The higher accuracy in image recognition and higher density of the device show that the proposed device is highly suitable for next generation neuromorphic applications.

## V. Conclusion

The works show the feasibility of novel vertically stacked nanosheet FET for charge trapping and artificial synaptic properties such as long-term potentiation (LTP) and depression (LTD). The non-linearity of depression scheme is minimized to -0.42 with incremental pulse voltage and pulse time of 100 ns. An optimized depression scheme with incremental pulse achieves better linearity and accuracy (93.30 %) than conventional depression operation (89.00 %). This NSFET device achieves winder memory window and infinite retention time for Erase (potentiation) operation. The device is scaled down to 20 nm with memory window of 1.77 V. These results convey that nanosheet FET is a promising candidate for next-generation memory and neuromorphic computing for longer retention, endurance, optimal memory window, higher density with linear weight, and better accuracy.

## References

[1] C. D. Schuman, S. R. Kulkarni, M. Parsa, J. P. Mitchell, P. Date, and B. Kay, “Opportunities for neuromorphic computing algorithms and applications,” *Nat. Comput. Sci.*, vol. 2, no. 1, pp. 10–19, Jan. 2022, doi: 10.1038/s43588-021-00184-y.

[2] G. Zhou, Z. Wang, B. Sun, F. Zhou, L. Sun, H. Zhao, X. Hu, X. Peng, J. Yan, H. Wang, W. Wang, J. Li, B. Yan, D. Kuang, Y. Wang, L. Wang, and S. Duan, “Volatile and Non-volatile Memristive Devices for Neuromorphic Computing,” *Adv. Electron. Mater.*, vol. 2101127, p. 2101127, Feb. 2022, DOI: 10.1002/aelm.202101127.

[3] S. Yu, “Neuro-Inspired Computing With Emerging Nonvolatile Memorys,” *Proc. IEEE*, vol. 106, no. 2, pp. 260–285, Feb. 2018, doi: 10.1109/JPROC.2018.2790840.

[4] W. Banerjee, R. D. Nikam, and H. Hwang, “Prospect and challenges of analog switching for neuromorphic hardware,” *Appl. Phys. Lett.*, vol. 120, no. 6, p. 060501, Feb. 2022, doi: 10.1063/5.0073528.

[5] Z. Yu, Z. Wang, J. Kang, Y. Fang, Y. Chen, Y. Cai, and R. Huang, “Early-Stage Fluctuation in Low-Power Analog Resistive Memory: Impacts on Neural Network and Mitigation Approach,” *IEEE Electron Device Lett.*, vol. 41, no. 6, pp. 940–943, Jun. 2020, doi: 10.1109/LED.2020.2986889.

[6] K. Moon, S. Lim, J. Park, C. Sung, S. Oh, J. Woo, J. Lee, and H. Hwang, “RRAM-based synapse devices for neuromorphic systems,” *Faraday Discuss.*, vol. 213, pp. 421–451, 2019, doi: 10.1039/C8FD00127H.

[7] A. I. Khan, H. Kwon, M. E. Chen, M. Asheghi, H.-S. Philip Wong, K. E. Goodson, and E. Pop, “Electro-Thermal Confinement Enables Improved Superlattice Phase Change Memory,” *IEEE Electron Device Lett.*, vol. 43, no. 2, pp. 204–207, Feb. 2022, doi: 10.1109/LED.2021.3133906.

[8] M. Suri, O. Bichler, D. Querlioz, O. Cueto, L. Perniola, V. Sousa, D.Vuillaume, C. Gamrat, and B. DeSalvo, “Phase change memory as synapse for ultra-dense neuromorphic systems: Application to complex visual pattern extraction,” in *2011 International Electron Devices Meeting*, Dec. 2011, pp. 4.4.1-4.4.4. doi: 10.1109/IEDM.2011.6131488

[9] G. Srinivasan, A. Sengupta, and K. Roy, “Magnetic Tunnel Junction Based Long-Term Short-Term Stochastic Synapse for a Spiking Neural Network with On-Chip STDP Learning,” *Sci. Rep.*, vol. 6, no. 1, p. 29545, Sep. 2016, doi: 10.1038/srep29545.

[10] M. H. R. Ansari, S. Cho, J.-H. Lee, and B.-G. Park, “Core-Shell Dual-Gate Nanowire Memory as a Synaptic Device for Neuromorphic Application,” *IEEE J. Electron Devices Soc.*, vol. 9, no. June, pp. 1282–1289, 2021, doi: 10.1109/JEDS.2021.3111343.

[11] M.-S. Lee, J.-W. Lee, C.-H. Kim, B.-G. Park, and J.-H. Lee, “Implementation of Short-Term Plasticity and Long-Term Potentiation in a Synapse Using Si-Based Type of Charge-Trap Memory,” *IEEE Trans. Electron Devices*, vol. 62, no. 2, pp. 569–573, Feb. 2015, doi: 10.1109/TED.2014.2378758.

[12] M.-K. Kim and J.-S. Lee, “Ferroelectric Analog Synaptic Transistors,” *Nano Lett.*, vol. 19, no. 3, pp. 2044–2050, Mar. 2019, doi: 10.1021/acs.nanolett.9b00180.

[13] N. Loubet, T. Hook, P. Montanini , C.-W. Yeung , S. Kanakasabapathy , M. Guillorn , T. Yamashita , J. Zhang , X. Miao , J. Wang , A. Young , R. Chao , M. Kang, Z. Liu , S. Fan , B. Hamieh , S. Sieg , Y. Mignot , W. Xu , S.-C. Seo , J. Yoo, S. Mochizuki , M. Sankarapandian , O. Kwon, A. Carr , A. Greene , Y. Park, J. Frougier, R. Galatage, R. Bao , J. Shearer , R. Conti , H. Song, D. Lee, D. Kong , Y. Xu , A. Arceo , Z. Bi , P. Xu , R. Muthinti , J. Li , R. Wong , D. Brown, P. Oldiges , R. Robison , J. Arnold , N. Felix , S. Skordas , J. Gaudiello , T. Standaert , H. Jagannathan , D. Corliss , M.-H. Na , A. Knorr, T. Wu , D. Gupta , S. Lian, R. Divakaruni , T. Gow , C. Labelle, S. Lee, V. Paruchuri , H. Bu , and M. Khare, “Stacked nanosheet gate-all-around transistor to enable scaling beyond FinFET,” in *2017 Symposium on VLSI Technology*, Jun. 2017, vol. 5, no. 1, pp. T230–T231, doi: 10.23919/VLSIT.2017.7998183.

[14] S. Barraud, B. Previtali, C. Vizioz, J.-M. Hartmann, J. Sturm, J. Lassarre, C. Perrot, Ph. Rodriguez, V. Loup, A. Magalhaes-Lucas, R. Kies, G. Romano, M. Cassé, N. Bernier, A. Jannaud, A. Grenier, F. Andrieu., “7-Levels-Stacked Nanosheet GAA Transistors for High Performance Computing,” *Dig. Tech. Pap. - Symp. VLSI Technol.*, vol. 2020-June, pp. 22–23, 2020, doi: 10.1109/VLSITechnology18217.2020.9265025.

[15] A. Agrawal, S. Chouksey, W. Rachmady, S. Vishwanath, S. Ghose, M. Mehta, J. Torres, A.A. Oni, X. Weng, H. Li, D. Merrill, M. Metz, A. Murthy and J. Kavalieros, “Gate-All-Around Strained Si 0.4 Ge 0.6 Nanosheet PMOS on Strain Relaxed Buffer for High Performance Low Power Logic Application,” in *2020 IEEE International Electron Devices Meeting (IEDM)*, Dec. 2020, pp. 2.2.1-2.2.4, doi: 10.1109/IEDM13553.2020.9371933.

[16] IEEE International Roadmap for Devices and Systems (IRDS) (https://irds.ieee.org/)

[17] U. K. Das, M. G. Bardon, D. Jang, G. Eneman, P. Schuddinck, D. Yakimets, P. Raghavan, and G. Groeseneken, “Limitations on Lateral Nanowire Scaling Beyond 7-nm Node,” *IEEE Electron Device Lett.*, vol. 38, no. 1, pp. 9–11, Jan. 2017, doi: 10.1109/LED.2016.2629420.

[18] U. K. Das and T. K. Bhattacharyya, “Opportunities in Device Scaling for 3-nm Node and Beyond: FinFET Versus GAA-FET Versus UFET,” *IEEE Trans. Electron Devices*, vol. 67, no. 6, pp. 2633–2638, Jun. 2020, doi: 10.1109/TED.2020.2987139.

[19] T. B. Hook, “Power and Technology Scaling into the 5 nm Node with Stacked Nanosheets,” *Joule*, vol. 2, no. 1, pp. 1–4, Jan. 2018, doi: 10.1016/j.joule.2017.10.014.

[20] T. Bang, B.-H. Lee, C.-K. Kim, D.-C. Ahn, S.-B. Jeon, M.-H. Kang, J.-S. Oh, and Y.-K. Choi, “Low-frequency noise characteristics in SONOS flash memory with vertically stacked nanowire FETs,” *IEEE Electron Device Lett.*, vol. 38, no. 1, pp. 40–43, 2017, doi: 10.1109/LED.2016.2632182.

[21] J. Lee, B. G. Park, and Y. Kim, “Implementation of Boolean Logic Functions in Charge Trap Flash for In-Memory Computing,” *IEEE Electron Device Lett.*, vol. 40, no. 9, pp. 1358–1361, 2019, doi: 10.1109/LED.2019.2928335.

[22] Li Deng, “The MNIST Database of Handwritten Digit Images for Machine Learning Research [Best of the Web],” *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 141–142, Nov. 2012, doi: 10.1109/MSP.2012.2211477.

[23] Synopsys, “Sentaurus 3D TCAD Manual,” 2022.

[24] E. Gnani *et al.*, "Modeling of gate-all-around charge trapping SONOS memory cells," *Solid. State. Electron.*, vol. 54, no. 9, pp. 997–1002, Sep. 2010, doi: 10.1016/j.sse.2010.04.026.

[25] N. El-Atab, F. Cimen, S. Alkis, B. Ortac¸ M. Alevli, N. Dietz, A. K. Okyay, and A. Nayfeh, "Enhanced memory effect via quantum confinement in 16 nm InN nanoparticles embedded in ZnO charge trapping layer," *Appl. Phys. Lett.*, vol. 104, no. 25, pp. 1–5, 2014, doi: 10.1063/1.4885397.

[26] A. Nayfeh and N. El-Atab, Nanomaterials-Based Charge Trapping Memory Devices (Elsevier, 2020).

[27] L. Dobrescu, M. Petrov, D. Dobrescu, and C. Ravariu, "Threshold voltage extraction methods for MOS transistors," in *2000 International Semiconductor Conference. 23rd Edition. CAS 2000 Proceedings (Cat. No.00TH8486)*, 2000, vol. 1, no. 2, pp. 371–374. doi: 10.1109/SMICND.2000.890257.

[28] N. El-Atab, A. Ozcan, S. Alkis, A. K. Okyay, and A. Nayfeh, "Low power zinc-oxide based charge trapping memory with embedded silicon nanoparticles via poole-frenkel hole emission," *Appl. Phys. Lett.*, vol. 104, no. 1, p. 013112, Jan. 2014, doi: 10.1063/1.4861590.

[29] G. Ding, B.Yang, K. Zhou, C. Zhang, Y. Wang, J.-Q. Yang, S.-T. Han, Y. Zhai, V. A. L. Roy, and Ye Zhou, "Synaptic Plasticity and Filtering Emulated in Metal–Organic Frameworks Nanosheets Based Transistors," *Adv. Electron. Mater.*, vol. 6, no. 1, p. 1900978, Jan. 2020, doi: 10.1002/aelm.201900978.

[30] P.-Y. Chen, X. Peng, and S. Yu, "NeuroSim+: An integrated device-to-algorithm framework for benchmarking synaptic devices and array architectures," in *2017 IEEE International Electron Devices Meeting (IEDM)*, Dec. 2017, pp. 6.1.1-6.1.4. doi: 10.1109/IEDM.2017.8268337.

[31] X. Sun and S. Yu, "Impact of Non-Ideal Characteristics of Resistive Synaptic Devices on Implementing Convolutional Neural Networks," *IEEE J. Emerg. Sel. Top. Circuits Syst.*, vol. 9, no. 3, pp. 570–579, Sep. 2019, doi: 10.1109/JETCAS.2019.2933148.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60