

Cite this: DOI: 00.0000/xxxxxxxxxx

Role of electrolyte layer in CMOS-compatible and oxide-based vertical three-terminal ECRAM[†]

Geonhui Han,^{‡a} Jongseon Seo,^{a‡} Hyejin Kim,^{a‡} and Deaseok Lee^a

Received Date

Accepted Date

DOI: 00.0000/xxxxxxxxxx

1 Supplementary information

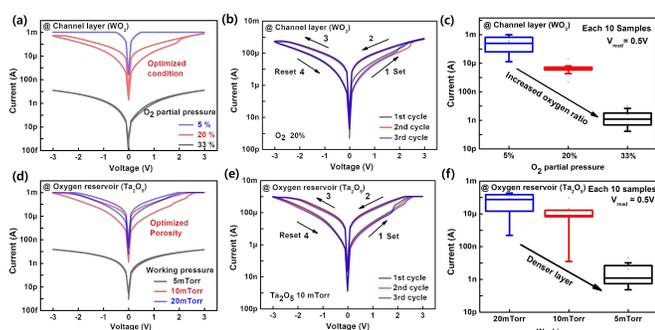


Fig. 1 (a) Logarithmic current versus voltage (log I-V) characteristics depending on the oxygen ion concentration of the channel layer. (b) Repeatable log I-V characteristics at the optimized oxygen concentration of the channel layer. (c) Comparison of initial current, which is dependent on the oxygen ion concentration of the channel layer. (d) Various working pressures during depositing oxygen reservoir layer are optimized based on the optimized condition of the channel layer. (e) The obtained log I-V characteristics at the optimized conditions of the channel and oxygen reservoir layers. (f) Comparison of initial current with varying working pressures of the oxygen reservoir layer.

To develop 2T- and 3T-ECRAMs, a TiN/WO₃/Ta₂O₅/W structure is fabricated on a 3D vertical wafer. First, the oxygen concentration of the channel layer (WO₃) is optimized (a–c). During the deposition of WO₃ (channel layer), the oxygen partial pressure is varied from 5 % to 20 % to 33 %. At 20 % oxygen partial pressure, gradual and sequential current changes are exhibited under an applied voltage bias (b). Moreover, current hysteresis is observed, which can be explained by the change in total resistance due to the movement of oxygen ions from the WO₃ layer to the Ta₂O₅ layer. When a positive voltage is applied to the W

electrode, mobile oxygen ions are attracted from the WO₃ layer to the Ta₂O₅ layer. Thus, oxygen vacancies are generated in the WO₃ layer, which decrease the total resistance. Subsequently, under a negative bias, the attracted oxygen ions return to the WO₃ layer, and the total resistance increases.

In contrast to the 20 % case, at 5 % and 33 % oxygen partial pressures, electric short and insulating characteristics are observed, respectively. In the 5 % case, a significantly high concentration of oxygen vacancies are formed in the WO₃ channel layer, whereas mobile oxygen ions are insufficient in the 33 % case. Thus, the statistical initial current level is higher in the 5 % case than in the 33 % case (c) ^{1–3}.

Next, for the oxygen reservoir (Ta₂O₅) layer, the working pressure is controlled from 5 mTorr to 10 mTorr to 20 mTorr (d–f). Because a higher working pressure can lead to a more porous film, the Ta₂O₅ layer is optimized to increase the effective interface area between the WO₃ and Ta₂O₅ layers ^{4,5}. Compared with the lowest working pressure, the higher working pressures (10 mTorr and 20 mTorr) exhibit higher current levels and hysteresis (d). This is because a more porous Ta₂O₅ layer can be formed at higher working pressures, and it can absorb more oxygen ions from the WO₃ layer. In contrast to the higher working pressure, mobile oxygen ions cannot be absorbed into the dense Ta₂O₅ layer, which is formed at the lowest working pressure. The statistical initial current levels for the working pressures are strongly correlated with the obtained results (f).

^aDepartment of Electronics Materials Engineering, 20 Kwangwoon, Seoul, 01897, Nowon-Gu, Republic of Korea

Email Address: ghhan@kw.ac.kr, seojs@kw.ac.kr, hyejinkim@kw.ac.kr and leeds@kw.ac.kr

[†] Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/cXsm00000x/

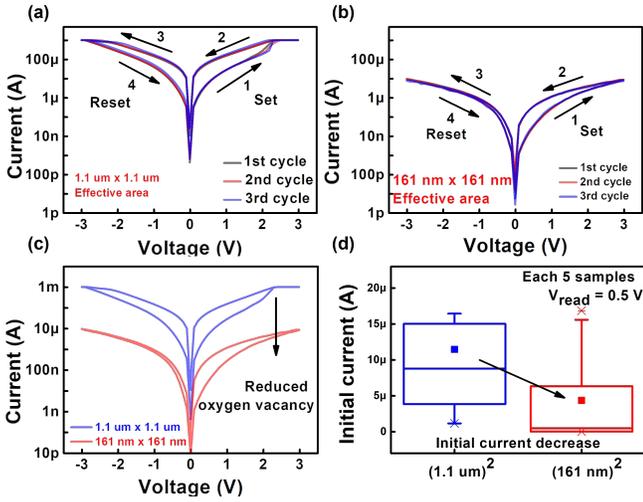


Fig. 2 (a) The log I-V characteristics of TiN/WO₃/Ta₂O₅/W structure formed under the optimized fabrication conditions. (b–c) Size dependence is evaluated to confirm bulk switching, which can exhibit gradual conductance changes. (d) Devices with larger areas exhibit higher initial current levels.

The area dependence of the TiN/WO₃/Ta₂O₅/W structure is determined based on the optimized fabrication conditions of the WO₃ and Ta₂O₅ layers. When the area is reduced, the current level of the log I-V characteristics also decreases. This implies that current hysteresis results from the movement of oxygen ions through the entire interface (bulk switching) and not through a localized region. In other words, in a larger area cell, more oxygen ions can be attracted to the interface between the WO₃ and Ta₂O₅ layers.

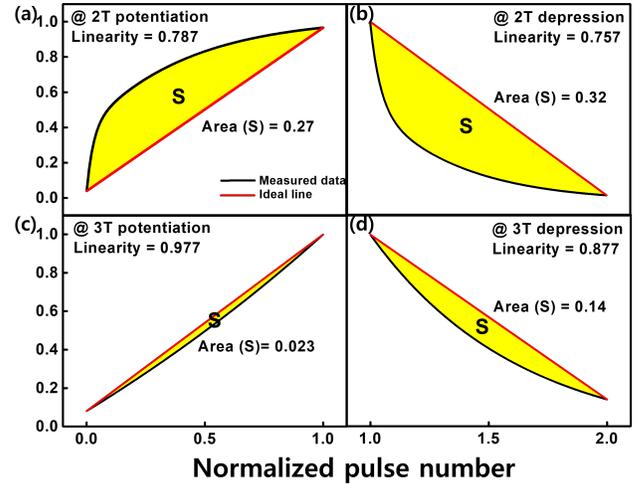


Fig. 3 Definition of linearity in potentiation and depression. The linearity of (a–b) 2T- and (c–d) 3T-ECRAMs is defined as $1/(1+S)$ based on the difference between the measured data and the ideal case.

The dynamic range and linearity of potentiation and depression are defined to quantitatively compare synaptic characteristics. First, dynamic range, which is the available conductance range, is defined as the ratio between the maximum and minimum conductance levels. Next, linearity is defined by S , which is an integrated value of the differences between the measured data and the ideal case for all conductance levels. Note that all differences at any conductance level must be considered because all conductance levels can be utilized during inference and training processes. Therefore, linearity is defined as follows:

$$Linearity = \frac{1}{(1+S)}$$

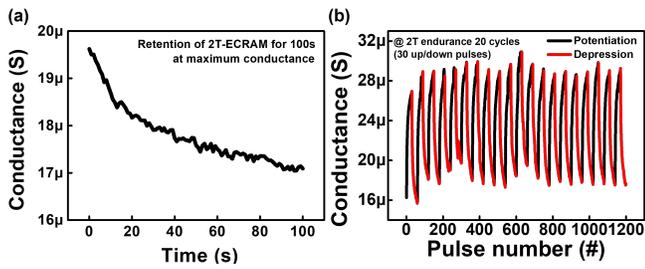


Fig. 4 2T-ECRAM exhibits relatively unstable synaptic characteristics, such as (a) degraded retention of the maximum conductance level (for 100 s) and (b) fluctuating P/D cycles under 1,200 applied pulses.

2T-ECRAM ($\text{TiN}/\text{WO}_3/\text{Ta}_2\text{O}_5/\text{W}$) exhibits relatively unstable retention and P/D cycles. The maximum conductance level gradually decreases over 100 s. In addition, P/D cycles fluctuate under 1,200 identical pulses. Considering these results, degraded retention can be explained by the self-diffusion of attracted oxygen ions. Additionally, fluctuating P/D cycles can result from the influence of interfacial oxide layer. The interfacial layer can be formed unstably due to self-diffusion behavior. Thus, considering that conductance is derived from the current flowing through all layers, including the interfacial layer, conductance can fluctuate during P/D cycles.

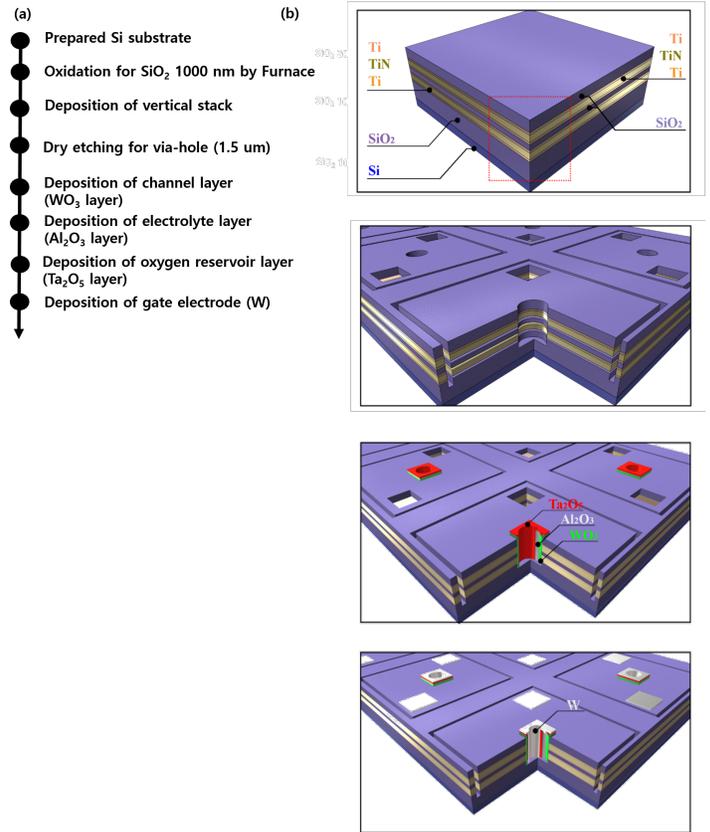


Fig. 5 (a) Detailed fabrication flow of 3T-eECRAM. (b) Simple illustration of the fabricated 3T-eECRAM.

To develop 3T-eECRAM, $\text{Ti}/\text{TiN}/\text{Ti}/\text{SiO}_2/\text{Ti}/\text{TiN}/\text{Ti}/\text{SiO}_2$ vertical stacks are sequentially deposited on a thermally oxidized 8-inch p-Si wafer. Ti (10 nm) and TiN (140 nm) layers are used as the adhesion layer and electrode, respectively. Between the two TiN electrodes (employed as the source and drain), a 100-nm-thick SiO_2 layer is deposited, which implies that the channel length is deposited as an inter-metal layer. Next, a 1.5 μm diameter via hole is formed by a typical dry etching process. A square-shaped metal contact via is also formed to contact the electrodes (source and drain). Considering that 3T-eECRAM is fabricated vertically, the effective cell area is approximately $680 \text{ nm} \times 680 \text{ nm}$ ($L \times W = 100 \text{ nm} \times 4.7 \text{ }\mu\text{m}$, where L is the channel length and W is the channel width). In the via hole, $\text{WO}_3/\text{Ta}_2\text{O}_5/\text{W}$ layers are deposited sequentially using a typical sputtering system. Finally, the fabrication parameters of each layer are optimized to improve synaptic characteristics, as shown in Fig. 2 and supplementary figures S1–S2.

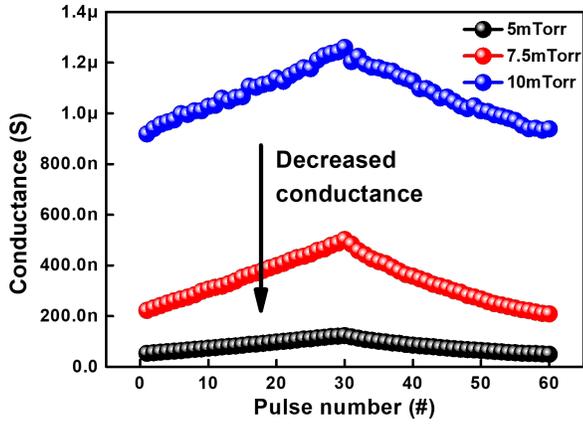


Fig. 6 Potentiation and depression characteristics, which are dependent on working pressures during the deposition of the electrolyte layer (Al_2O_3). A lower conductance level is obtained at a lower working pressure, which leads to a denser electrolyte layer^{4,5}.

The vertical leakage currents and conductance levels of potentiation and depression can be optimized by controlling the working pressure during deposition, as shown in Fig. 2c and Fig. S6. Considering that more porous electrolyte layers form at higher working pressures, the effective interface area between the electrolyte and channel layers can be increased^{4,5}. Thus, a relatively higher oxygen absorption from the channel layer to the electrolyte layer can occur in the porous electrolyte layer. Consequently, more oxygen vacancies can be generated, which leads to increased channel conductance.

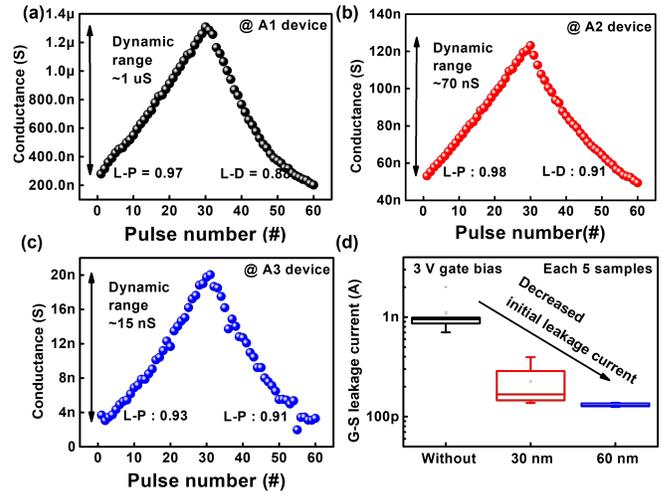


Fig. 7 (a–c) Potentiation and depression characteristics, which are dependent on the thickness of electrolyte layer with (d) the vertical leakage current. Not only the conductance levels of potentiation and depression but also the vertical leakage current are decreased when the thickness of electrolyte layer increases.

With respect to device fabrication, the vertical leakage current must be prevented because it can cause disturbances during the measurement of channel conductance. Thus, each layer of 3T-eECRAM is optimized; in particular, vertical leakage current is successfully prevented by optimizing the electrolyte layer.

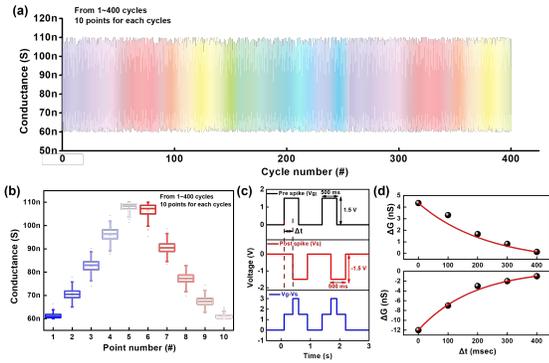


Fig. 8 (a) More stable P/D cycles of 3T-eECRAM can be achieved by limiting the number of conductance states as 10 (4000 pulses = 10 conductance states \times 400 P/D cycles). (b) Statistical plot also exhibits a reliable distribution of the 400 P/D cycles. In addition, spike-timing-dependent plasticity is also evaluated for applying a spike neural network (SNN). (c) Conductance changes can be modulated by applying spike pulses from pre- and post-synapse with different spike timings (Δt). (d) Higher conductance changes are observed as Δt is reduced.

More reliable P/D cycles can be achieved by limiting the number of conductance levels. When the number of conductance levels is limited to 10, 3T-eECRAM exhibits more stable P/D cycles. In addition, 3T-eECRAM can be operated by spike pulses for SNN applications, providing highly time- and energy-efficient computing based on spatiotemporal data⁶.

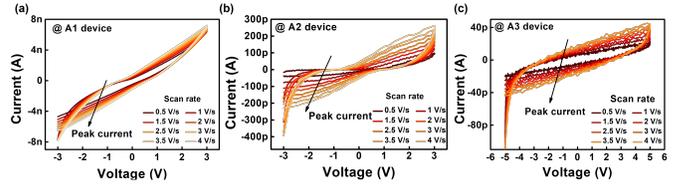


Fig. 9 Cyclic voltammetry (CV) of 3T-eECRAM with different thicknesses of electrolyte layer ((a) 0 nm, (b) 30 nm, and (c) 60 nm). CV is measured for all cases between the gate and source; note that CV indicates ionic motions.

Ionic motion can be investigated based on CV. The peak current of CV curve is fitted using the Randles–Sevcik equation:

$$I_{peak} = 0.4958(Fn)^{3/2}(RT)^{-1/2}Ac_o(\alpha D_o v)^{1/2}$$

where F is the Faraday constant, A is the device area, c_o is the ion concentration, α is the transfer coefficient, R is the gas constant, T is the temperature, D_o is the diffusion coefficient, and v is the scan rate^{7–9}.

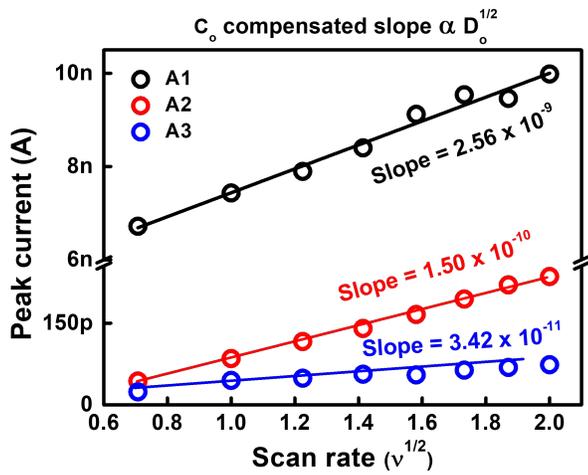


Fig. 10 Peak current fitted using the Randles–Sevcik equation. It is compensated by the ion concentration values (c_o).

The peak currents of CV can be fitted based on the Randles–Sevcik equation; the slope of the fitted peak current is proportional to both c_o and the diffusion coefficient (D_o), as shown in Fig. 4d. Considering that c_o is correlated to concentration of oxygen in the channel layer, among A1, A2, and A3, the ratio of c_o can be derived from the results of XPS analysis in Fig. 5. Consequently, the slope of fitted peak current can be compensated by the ratio of c_o to extract an influence of only D_o . In other words, c_o of the samples: A1, A2, and A3 is compensated as the same value, and the slope compensated by c_o is proportional to only D_o .

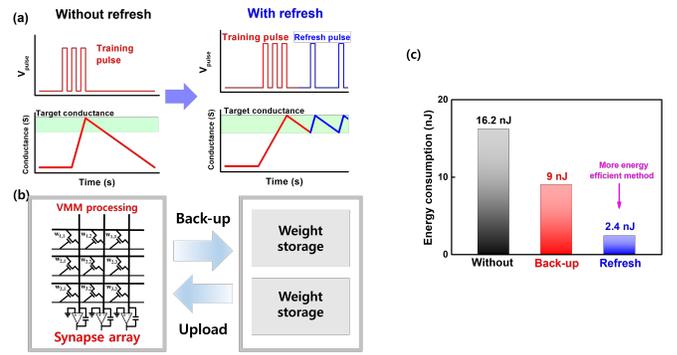


Fig. 11 (a) Data refreshing method to recover degraded retention. (b) Data back-up method as another recovery method. (c) Comparison of energy consumption among all approaches (without data recovery, data refreshing, and data back-up methods).

Even though the inserted electrolyte layer can significantly improve synaptic characteristics, retention is still degraded. Therefore, recovery methods, such as data refreshing or data back-up, can be implemented to improve retention. After the synaptic weight updating process, when synaptic weight is degraded, a refresh pulse bias, which can update the synaptic weight with its original weight value, can be applied. In addition, updated synaptic weights can be saved in other digital memory devices. Subsequently, before the inference process, the stored synaptic weight can be uploaded to 3T-eECRAM. The energy consumption for weight updating is compared based on these recovery methods, and the data refreshing method consumes the least energy among the three cases.

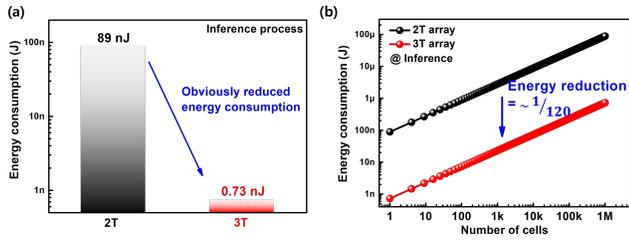


Fig. 12 (a) Comparison of energy consumption between 2T-ECRAM and 3T-eECRAM during the inference process. (b) Energy consumption can be demonstrably reduced by approximately a factor of 1/120.

3T-eECRAM consumes significantly less energy than 2T-ECRAM. The energy consumption of each device is approximated as¹⁰:

$$E = \text{pulse amplitude} \times \text{current} \times \text{pulse width}$$

Compared with 2T-ECRAM, 3T-eECRAM consumes approximately 800,000 and 120 times lesser energy during the synaptic weight update (Fig. 6c) and inference processes, respectively.

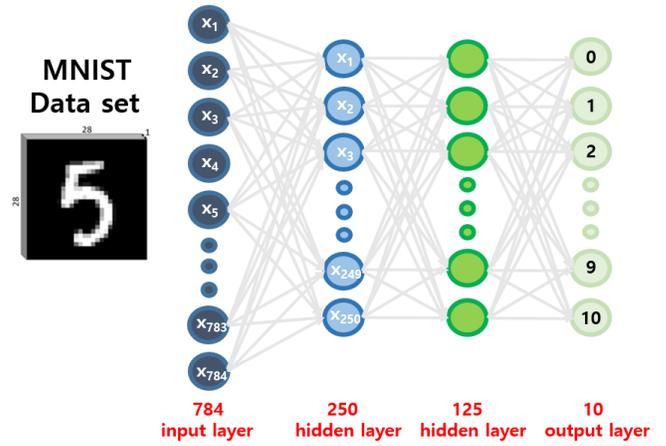


Fig. 13 (a) Schematic of the utilized ANN for recognizing the MNIST data set. ANN has 784, 250, 125, and 10 nodes at input, first hidden, second hidden, and output layers, respectively. Compared with 3T-ECRAM, a higher recognition accuracy (92%) is achieved based on linear P/D cycles of 3T-eECRAM.

To evaluate the influence of the achieved synaptic characteristics on the system, an artificial neural network (ANN) is constructed that can recognize the MNIST dataset^{11,12}. Compared with 2T-ECRAM, 3T-eECRAM with linear potentiation and depression demonstrably achieves an increase in recognition accuracy from 42.3% to 92% (Fig. 6d).

	[13]	[14]	[2]	[3]	This work
CMOS compatibility	X	O	O	O	O
Device size (dimension)	> 10 F ²	> 10 F ²	> 10 F ²	4 F ²	4 F ²
Structure	Planar	Planar	Planar	Vertical	Vertical
Channel length	210 um	4 um	~ 10 um	< 0.1 um	< 0.1 um
Speed	~ 5 s	~ 100 ns	~ 1 ms	~ 10 us	~ 100 us
Linearity	-	-	$\alpha_p = 0.8$ $\alpha_d = 0.7$	$\alpha_p = 1.43$ $\alpha_d = -1.36$	$\alpha_p = 0.93$ $\alpha_d = -0.51$
Array size	5 x 5	2 x 2	2 x 2	2 x 2	4 x 4

Fig. 14 Performance comparison between previously reported O-ECRAM. As compared with previously reported ECRAM, our 3T-eECRAM exhibits small dimension, high-speed switching, and feasibility of mass production.

Table summarizes the performance comparison between previously reported O-ECRAM and our devices. Our device has small dimension compared with planar structure (4 F²), high-speed switching characteristics (100 us), and feasibility of mass production (4 x 4) compared with the other reported papers^{2,3,13,14}.

Author Contributions

D.S.Lee conceived and directed the research. G.H. Han and D.S. Lee designed and planned the experiments. G.H. Han, J.S. Seo, and H.J. Kim performed the experiments and acquired the data. G.H. Han and D.S. Lee wrote the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2023- RS-2022-00156225) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation).

Notes and references

- 1 Y. Jeong, H. Lee, D. G. Ryu, S. H. Cho, G. Lee, S. Kim, S. Kim and Y. S. Lee, *Advanced Electronic Materials*, 2021, **7**, 2100185.
- 2 J. Lee, R. D. Nikam, M. Kwak, H. Kwak, S. Kim and H. Hwang, *Advanced Electronic Materials*, 2021, **7**, 2100219.
- 3 C. Lee, W. Choi, M. Kwak, S. Kim and H. Hwang, 2021 Symposium on VLSI Technology, 2021, pp. 1–2.
- 4 J. A. Thornton, *Journal of Vacuum Science and Technology*, 1974, **11**, 666–670.
- 5 Y. Shen, *Materials Science and Engineering: A*, 2003, **359**, 158–167.
- 6 Y. Li, J. Lu, D. Shang, Q. Liu, S. Wu, Z. Wu, X. Zhang, J. Yang, Z. Wang, H. Lv *et al.*, *Advanced Materials*, 2020, **32**, 2003018.
- 7 P. Zanello, C. Nervi and F. F. De Biani, *Inorganic electrochemistry: theory, practice and application*, Royal Society of Chemistry, 2019.
- 8 X. Tao, J. Wang, C. Liu, H. Wang, H. Yao, G. Zheng, Z. W. Seh, Q. Cai, W. Li, G. Zhou *et al.*, *Nature communications*, 2016, **7**, 1–9.
- 9 M. Strømme, G. A. Niklasson and C. G. Granqvist, *Solid state communications*, 1995, **96**, 151–154.
- 10 Y. Li, Z. Xuan, J. Lu, Z. Wang, X. Zhang, Z. Wu, Y. Wang, H. Xu, C. Dou, Y. Kang *et al.*, *Advanced Functional Materials*, 2021, **31**, 2100042.
- 11 J. Lee, R. D. Nikam, S. Lim, M. Kwak and H. Hwang, *Nanotechnology*, 2020, **31**, 235203.
- 12 G. W. Burr, R. M. Shelby, S. Sidler, C. Di Nolfo, J. Jang, I. Boybat, R. S. Shenoy, P. Narayanan, K. Virwani, E. U. Giacometti *et al.*, *IEEE Transactions on Electron Devices*, 2015, **62**, 3498–3507.
- 13 H.-Y. Huang, C. Ge, Q.-H. Zhang, C.-X. Liu, J.-Y. Du, J.-K. Li, C. Wang, L. Gu, G.-Z. Yang and K.-J. Jin, *Advanced Functional Materials*, 2019, **29**, 1902702.
- 14 S. Kim, T. Todorov, M. Onen, T. Gokmen, D. Bishop, P. Solomon, K.-T. Lee, M. Copel, D. B. Farmer, J. A. Ott *et al.*, 2019 IEEE International Electron Devices Meeting (IEDM), 2019, pp. 35–7.