Article

3D AND-Type Stacked Array for Neuromorphic Systems

Taejin Jang [1,2], Suhyeon Kim [1,2], Jeesoo Chang [1], Kyung Kyu Min [1,2], Sungmin Hwang [1,2], Kyungchul Park [1,2], Jong-Ho Lee [1,2] and Byung-Gook Park [1,*]

1 Department of Electrical and Computer Engineering, Seoul National University, Seoul 08826, Korea; star99@snu.ac.kr (T.J.); sh.kim@live.co.kr (S.K.); jess907@naver.com (J.C.); mkk2975@snu.ac.kr (K.K.M.); smh88@snu.ac.kr (S.H.); dremata104@gmail.com (K.P.); jhl@snu.ac.kr (J.-H.L.)

2 Inter-university Semiconductor Research Center (ISRC), Seoul 08826, Korea

* Correspondence: bgpark@snu.ac.kr; Tel.:+82-2-880-7282

Received: 20 July 2020; Accepted: 30 August 2020; Published: 31 August 2020

Abstract: NOR/AND flash memory was studied in neuromorphic systems to perform vector-by-matrix multiplication (VMM) by summing the current. Because the size of NOR/AND cells exceeds those of other memristor synaptic devices, we proposed a 3D AND-type stacked array to reduce the cell size. Through a tilted implantation method, the conformal sources and drains of each cell could be formed, with confirmation by a technology computer aided design (TCAD) simulation. In addition, the cell-to-cell variation due to the etch slope could be eliminated by controlling the deposition thickness of the cells. The suggested array can be beneficial in simple program/inhibit schemes given its use of Fowler–Nordheim (FN) tunneling because the drain lines and source lines are parallel. Therefore, the conductance of each synaptic device can be updated at low power level.

Keywords: neuromorphic; 3D AND-type synapse array; flash memory; Fowler–Nordheim tunneling

1. Introduction

Recently, artificial neural networks (ANNs) have shown tremendous performance capabilities in a variety of areas, particularly in the areas of recognition, face detection and voice recognition. However, software-based ANNs, which use GPUs for vector-by-matrix multiplication (VMM), consume a large amount of power. On the other hand, neuromorphic systems based on spiking neural networks (SNNs) have received much attention owing to their low power consumption through event-driven operation. Among the elements of a neuromorphic system, synapses can easily implement VMM by simply summing the current. Various memristors such as phase change memory (PCM) and resistive memory (RRAM) have advantages in terms of scaling and given their simple structures, but there are certain limitations, such as reliability and sneak current issues [1–10].

Another synaptic device candidate is flash memory, which is quite mature and has the advantage of stable operation.in particular, NOR/AND-type flash memory has been studied in an effort to implement a synaptic array [11–18]. However, the cell size in this case exceeds that of memristors, and there are challenges when attempting to scale a large number of synapse arrays. For this reason, we proposed a 3D stackable AND-type synapse array architecture. In the proposed structure, the sources and drains (S/Ds) are formed by tilted implantation and conformal doping is possible if buffer oxide is deposited onto the top layer. Based on the above structure, program/inhibit pulses were applied to verify the successful control of the threshold voltage by means of a TCAD simulation.

2. Characteristics of the AND-type Synapse Array

In SNNs, upon the input of voltage, the VMM is realized by summing the current through each synaptic device. Therefore, the input/output lines of synaptic devices must be perpendicular, and such a structure is termed a NOR/AND-type array (Figure 1). The advantage of the NOR flash structure is that when the same input voltage is applied to the gate and drain during the inference process, it can be operated in an event-driven mode to minimize the leakage current. However, a hot carrier injection should be used to control the conductance levels of individual cells, and this can cause a great deal of damage to the silicon oxide, adversely affecting the reliability of the device. In addition, the program/erase cycle through a hot carrier injection has a disadvantage in that it consumes much more power compared to the FN tunneling method [19]. Alternatively, FN tunneling can be used with an asymmetric dual gate, but the process can be somewhat complicated [20,21].

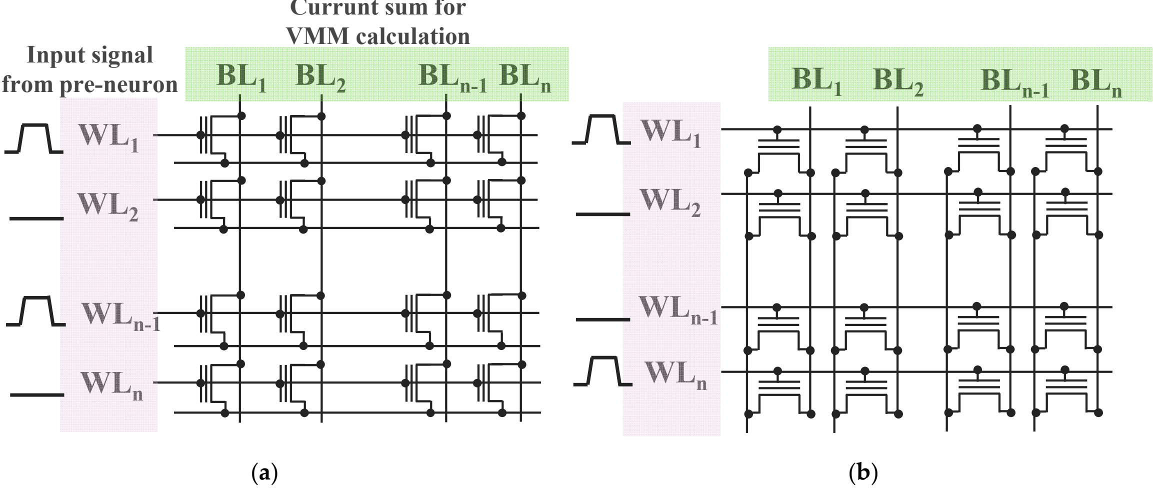

Figure 1. (**a**) NOR-type flash array structure. Because the directions of word lines and drain lines are identical, event-driven operation is simple if applying the same input pulses to those lines. (**b**) 3D AND-type stackable array with parallel source lines (SLs) and drain lines (DLs). The advantage of this structure is that the program and erase scheme is simple because the conductance can be adjusted through FN tunneling.

Therefore, the device consumes less energy for weight updates in less time as well. Additionally, because the cell size decreases as cells are stacked, another advantage is realized when integrating a large number of synapses.

3. Results

3.1. Process Flow of the 3D AND-type Synapse Array

Figure 2 summarizes the process flow of the overall synapse array, simulated here using the Sentaurus process tool. First, SiO_2 and in situ phosphorus-doped poly-Si are alternately deposited onto the stacked cells, and the active regions are patterned (Figure 2b). The basic structure is an AND-type array in which SLs and DLs are composed of doped poly-Si, with these two lines formed in parallel. Then, the doped poly-Si of the thin channel is removed through a chemical dry etching, after which poly-Si is deposited by chemical vapor deposition. This is done to minimize the leakage current by forming an intrinsic channel. Subsequently, dry etching is conducted to separate the cells of each layer, and a storage layer capable of storing the conductance of the synapse is deposited. Word lines are then formed in the vertical direction of each layer to transmit spiking signals to the devices on each layer. Double gates capable of controlling the conductance of cells and transferring the signals to post-neurons are formed at the side of each cell.

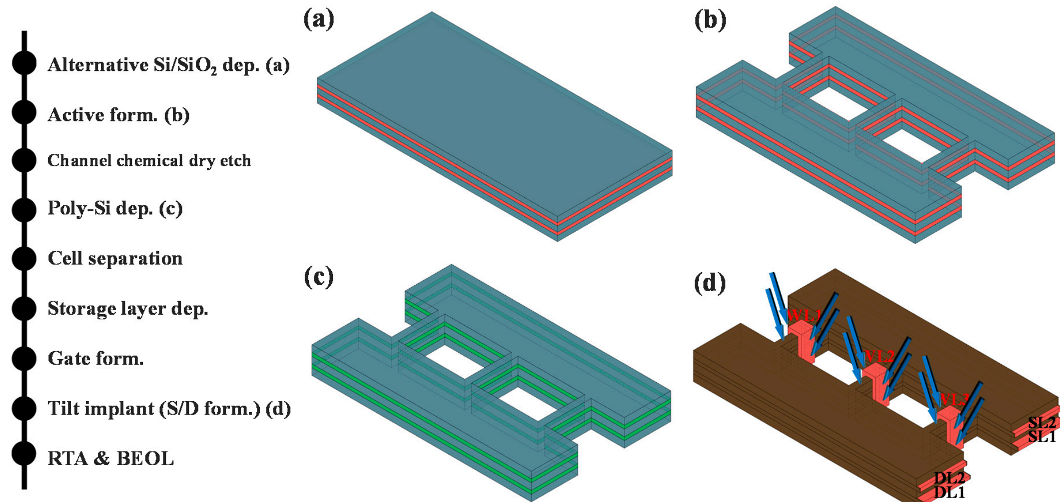

Figure 2. Process flow of the 3D AND-type synapse array (**a**) n+-doped poly-Si and SiO_2 are alternatively deposited; (**b**) Active regions are formed by dry etching. The SLs/DLs are positioned in parallel; (**c**) intrinsic poly-Si is deposited to make channels. These channels can reduce the leakage current caused by the drain bias; (**d**) Self-aligned sources and drains are formed by means of tilted ion implantation. Phosphorus at a dose of 1×10^{15} cm^{-2} is implanted at slight tilt angle.

Generally, when implementing a stacked array, it is difficult to form S/Ds. Therefore, junction-less metal-oxide-semiconductor field-effect transistors (MOSFETs) are used for stacked arrays. However, the leakage current of a junction-less MOSFET is large such that current sum errors can arise at the SLs. Accordingly, we propose tilted implantation as a means by which to form S/Ds (Figure 2d). Tilted implantation has already been extensively studied in relation to the formation of S/D in the FinFETs [22–33]. However, several problems arose during the commercialization process related to the tilted angle of implantation. When the tilted angle is small, the differences in the dopant concentration at the top, center and bottom of the fin become large. Therefore, conformal doping is difficult, and the performance is therefore degraded. On the other hand, at greater tilted angles, reducing the gate pitch is limited due to the shadow effect. In the proposed structure, because the buffer oxide is located above the top cell, the vertical components of the implanted dopants are removed. Moreover, the horizontal dopants are uniformly injected into the cells of each layer, as shown in Figure 3.

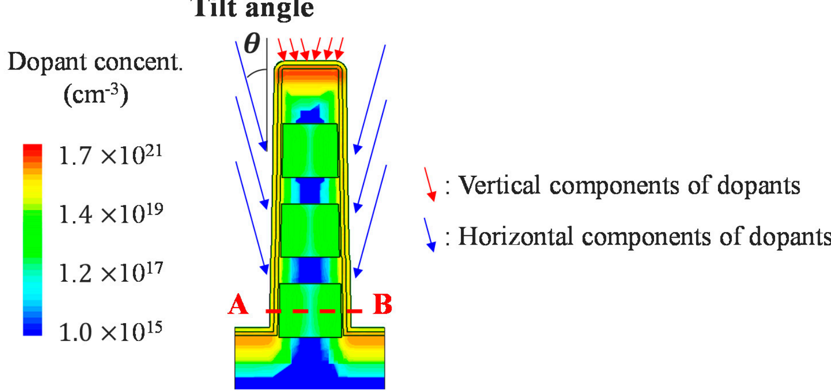

Figure 3. Dopant concentration after tilted implantation with a tilt angle of 15°. The vertical components of the dopants are implanted on the buffer oxide and are unable to affect the concentration of the cells. Because horizontal components are injected into each source and drain, this leads to similar doping concentrations among the cells.

Phosphorus at various tilted angles was simulated at a dose of 1×10^{15} cm^{-2} by a TCAD simulation (at an energy level of 20 keV). Figure 4 shows that the maximum dopant concentration is approximately 5×10^{18} cm^{-3} at 15° of tilt, which is a sufficient concentration for S/Ds. We analyzed the cell-to-cell variation in three layers at a tilt angle of 15° and confirmed that the dopant concentrations in the three layers did not differ significantly. Furthermore, the gate pitch can be reduced even further due to the low tilt angle.

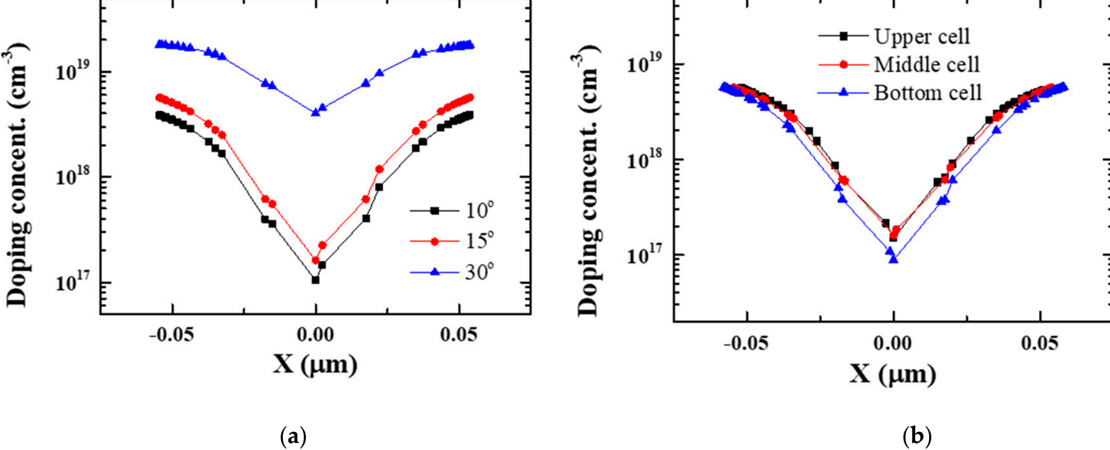

Figure 4. Doping concentration according to cut line shown in Figure 3: (**a**) Implantation with θ = 10°~15° is appropriate for the formation of SLs/DLs because the shadow effect is suppressed and a sufficient amount of dopant is implanted to SLs/DLs; (**b**) There are scant differences in the cell-to-cell dopant concentrations of the sources and drains.

Accordingly, it is possible to form a self-aligned S/D by tilted implantation, and rapid thermal annealing is then carried out at 950 °C for 10 s. after the metallization process, the 3D AND-type stacked array process is finally finished.

3.2. Cell-to-Cell Variation due to the Etch Slope

An etch slope of approximately 88° using an ICP poly-Si etcher with HBr gas can cause the cell-to-cell variation because it can affect the fin width and gate length. When the vertical cell pitch is 150 nm, the gate length and fin width increase by 5 nm as the cell goes down. Therefore, the on-current for the three layers was analyzed based on the upper cell, which had a gate length of 100-nm and a fin width of 100 nm, as shown in Figure 5a. Increasing the gate length and fin width has the effect of reducing the on-current, as shown in Figure 5b. In Figure 5c, we simulated the effects of both the gate length and the fin width on the current. As the gate length increases, the resistance of the channel also increases, which then reduces the current. Additionally, when the fin width increases, the spacing between the double gates increases, which reduces the gate controllability of the channel. Because both factors affect the current reduction, cell-to-cell variation occurs, with the current decreasing by approximately 4.7% each time the layer is lowered one-by-one.

There are two methods that can be used to resolve the issue of cell-to-cell variation caused by the etch slope as mentioned above. First, we can adjust the threshold voltage by applying more program pulses to the upper cells, as the current tends to decrease in the lower layer. However, in such a case, a crucial issue can arise because the threshold voltage of a large number of synapses must be set equally by applying a program pulse individually. The second method is to control the thickness of the lower layer to compensate for the low current. The thicker deposition of poly-Si in the lower cells is identical to increasing the effective width of the gate, implying that the current also increases. In addition, the cell-to-cell variation issue can be resolved simply, as the deposition thickness is easy to control.

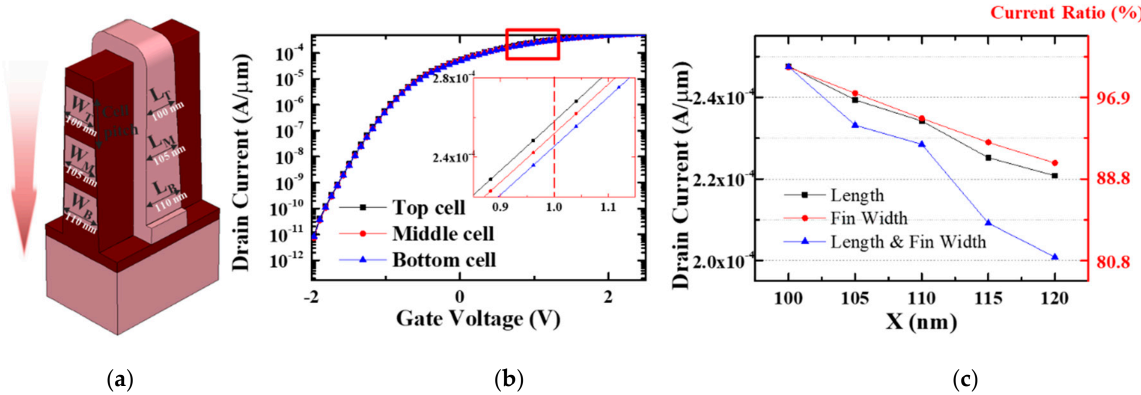

Figure 5. (**a**) Fin width and gate length of the lower cells increase due to the poly-Si etch slope; (**b**) The current of the lower cells tend to decrease when 1 V is applied to the gate; (**c**) The longer gate length increases the channel resistance and the wider fin width degrades the gate controllability. If the gate length and fin width become 120 nm, the current decreases by approximately 18.83% compared to when these values are both 100 nm.

3.3. Electrical Characteristics of Synapse Array

The proposed synapse array is an AND-type array structure; accordingly, individual conductance can be controlled through FN tunneling. The program and erase schemes are summarized in Table 1. When programming the selected cells, the FN tunneling condition is formed by applying 50 ns of a 13 V program pulse to the vertical word lines and 0 V to the SLs/DLs. Figure 6 shows the simulation results when applying ten program pulses to the selected cell, and it can be confirmed here that the current decreases as the threshold voltage shifts in the positive direction. For the unselected cells, the floating SLs/DLs inhibit FN tunneling by reducing the electric field at the tunneling oxide, which is a self-boost inhibit scheme (Figure 7) [15]. Because 0 V is applied to the word lines of the other inhibited cells, the threshold voltage of those cells does not change. When 100 inhibit pulses are applied to the unselected cells, the conductance does not change, and program disturb is well controlled. Using FN tunneling is highly cost-effective and convenient to control the conductance of numerous synapses, as the current is typically less than a few pA/μm.

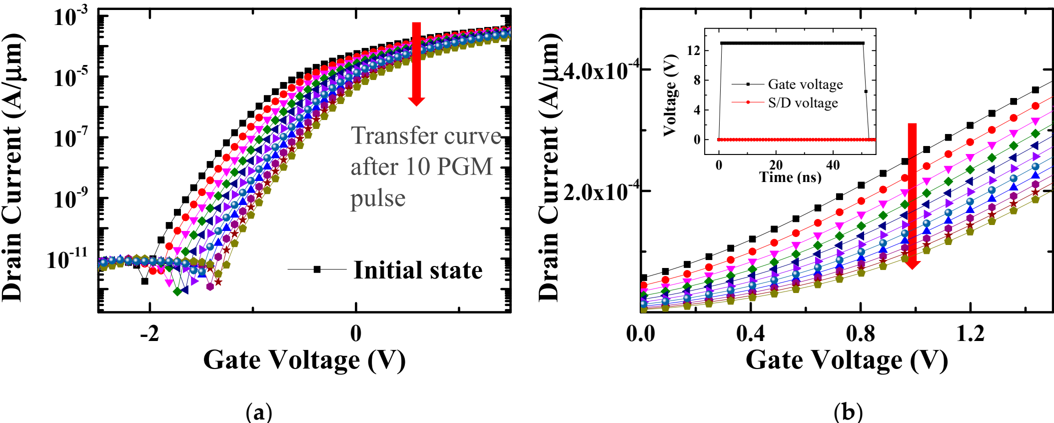

Figure 6. Transfer curves after applying several program/erase pulses: (**a**) The threshold voltages of selected cells increase as electrons become trapped in the trapping layer; (**b**) The transfer curves on a linear scale with different numbers of program pulses are shown.

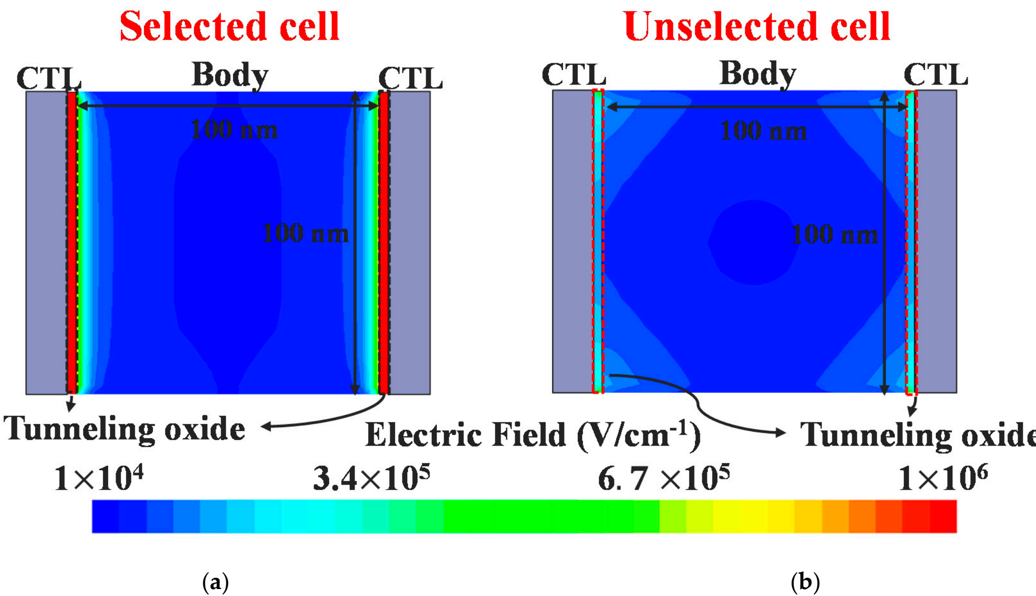

Figure 7. Contour plot of an electric field with the voltage condition shown in Table 1: (**a**) For the selected cells, the electrons are trapped in the storage layer by FN tunneling due to the high electric field of 1×10^6 V/cm^{-2}; (**b**) When the SLs/DLs are floating, the electric field of the tunneling oxide is too small to lead to FN tunneling owing to the self-boosted channel.

Table 1. Program/erase scheme for changing conductance of the synaptic device.

	Program	Erase
Selected WL	13 V	−19 V
Unselected WL	0 V	0 V
Selected SL/DL	0 V	0 V
Unselected SL/DL	Floating or $V_{pgm}/2$	Floating or $V_{er}/2$

Figure 8 shows a 4-bit operation of a synaptic device. Whenever a program pulse is applied to a target cell, the conductance gradually decreases, and it can be quantized to 16 levels. However, it is difficult to map the conductance linearly because with more program pulses applied to the cell, more charges fill the storage layer. In this case, the conductance can be more linearly quantized by means of incremental step pulse programming (ISPP) [34–38].

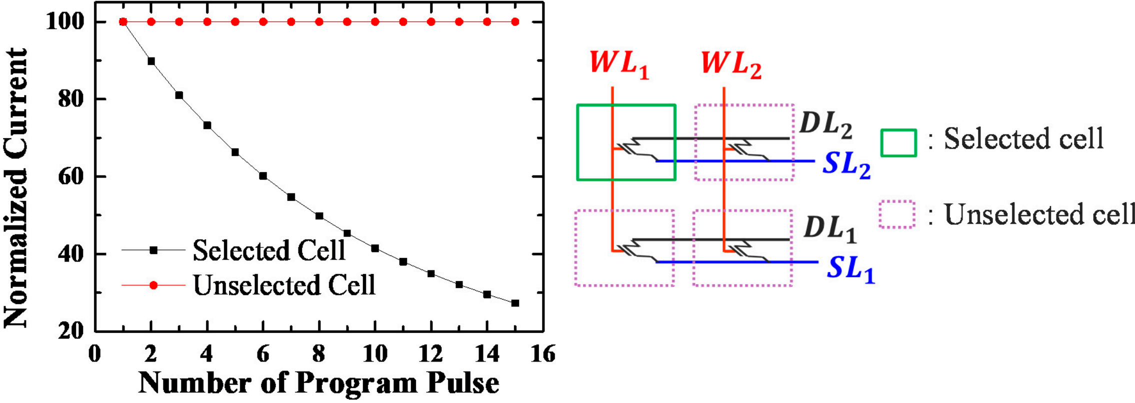

Figure 8. The normalized current of a selected cell decreases when applying 15 program pulses. On the other hand, the current levels of unselected cells in these three cases do not change. Each synapse can store 4-bit conductance.

Next, we simulate the erase operation to increase the conductance of the synaptic device. After applying ten program pulses, a few erase pulses are applied to a programmed cell and the change of conductance is then analyzed. In Figure 9, −19 V is applied to the gate and 0 V is applied to the S/D for 1 ms as an erase pulse. During the erase operation, trapped electrons are removed and holes are injected into the storage layer, resulting in an increase in the positive net charges. As the erase pulses are applied, the number of trapped electrons decreases, which degrades the erase efficiency. Through these program and erase operations, the conductance values of individual cells can be mapped through FN tunneling.

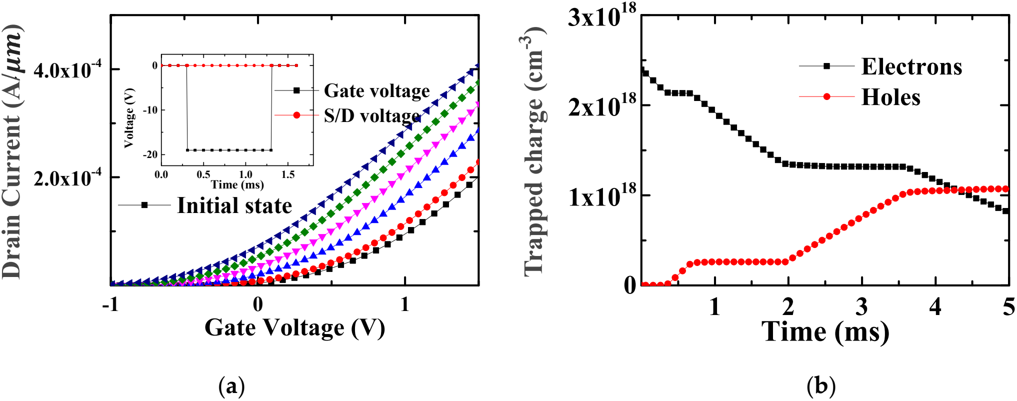

Figure 9. (**a**) When applying five erase pulses, the trans-characteristic curves are analyzed; (**b**) During the erase operation, positive net charges increase.

Read disturb as well as program disturb can change the conductance of a synaptic device, reducing its accuracy. When implementing a synapse array with a NAND-type array, a pass voltage must be applied to de-selected cells of the same string during the inference operation, causing a read disturb (Figure 10a) [39–49]. However, in the proposed structure, there is little risk of a read disturb because there is no need to apply pass voltage to the word lines of de-selected cells (Figure 10b). Moreover, the read voltage is very low at 1 V, and the conductance does not change even if read pulses at a rate of more than 10^5 times are applied. This makes the device more resistant to read disturb than a NAND-type array, and it has retention and endurance advantages compared to a NOR-type array, which uses hot carrier injection [50–53].

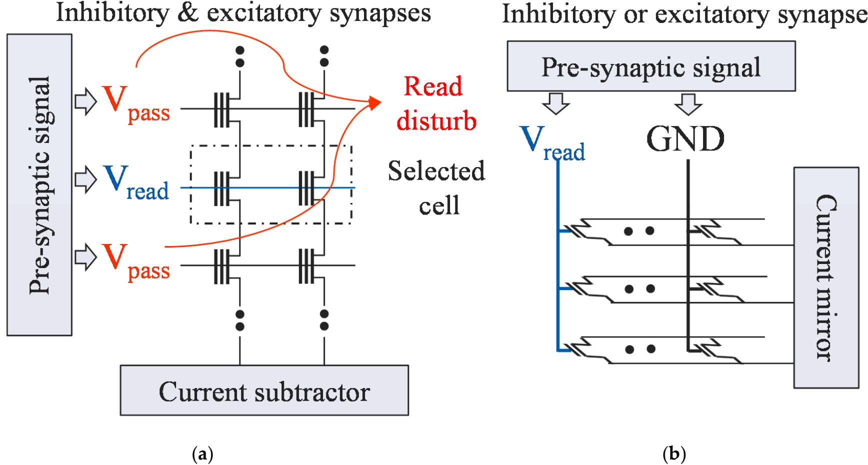

Figure 10. (**a**) Schematic of neural networks using NAND flash memory [40]; (**b**) Schematic of the proposed structure: pass voltage is not used on the de-selected cells and read disturb error can be suppressed.

The retention characteristics are analyzed at room temperature after applying ten program pulses with different program voltages. When 13 V is applied to the selected cell, the electric field caused by trapped charges is too small to degrade the retention. The conductance changes by approximately 2.1% over ten years, but if the program voltage increases, the retention characteristics worsen, as shown in Figure 11. Therefore, we selected 13 V for the program voltage and thus obtain better retention characteristics.

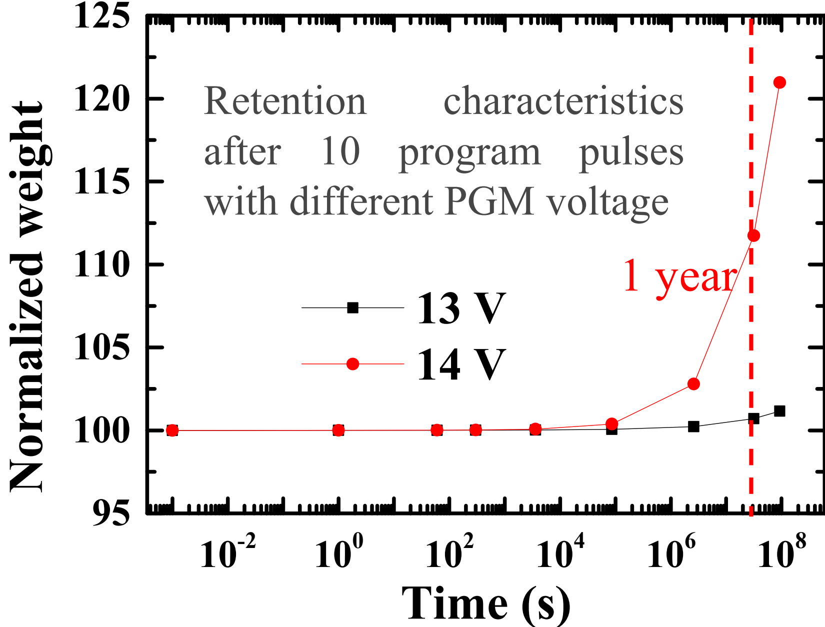

Figure 11. Retention characteristics according to the program voltages. After one year, the conductance programmed at 14 V changes by approximately 9.7%, whereas the change of the conductance programmed at 13 V is within 1%.

4. Conclusions

We proposed a 3D AND-type stacked array to implement a synapse array. Tilted implantation can be used to form conformal S/Ds of individual cells by depositing a buffer oxide onto the top of the cell. Additionally, the cell-to-cell variation caused by the etch slope can be reduced by thickly depositing the cells in the lower layers. It was verified through a simulation that the conductance levels of individual cells could be adjusted gradually at a low power using FN tunneling. In addition, the 10-year retention characteristic can be guaranteed through a low program voltage.

Author Contributions: Conceptualization, T.J., S.K., J.C., K.K.M., and S.H.; data curation, K.P.; writing—original draft preparation, T.J.; writing—review and editing, T.J. and B.-G.P.; supervision, B.-G.P. and J.-H.L.; project administration, B.-G.P. and J.-H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Brain Korea 21 Plus Project in 2020 and in part supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (2020–0–01294) and in part by Nano Material Technology Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (2016M3A7B4910348).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yu, S.; Wu, Y.; Jeyasingh, R.; Kuzum, D.; Philip Wong, H.-S. An electronic synapse device based on metal oxide resistive switching memory for neuromorphic computation. *IEEE Trans. Electron. Devices* **2011**, *58*, 2729–2737. [CrossRef]
2. Milo, V.; Pedretti, G.; Carboni, R.; Calderoni, A.; Ramaswamy, N.; Ambrogio, S.; Ielmini, D. Demonstration of hybrid CMOS/RRAM neural networks with spike time/rate-dependent plasticity. In Proceedings of the International Electron Device Meeting (IEDM), San Francisco, CA, USA, 3–7 December 2016; pp. 1–4. [CrossRef]
3. Ambrogio, S.; Balatti, S.; Milo, V.; Carboni, R.; Wang, Z.-Q.; Caldernoi, A.; Ramaswamy, N.; Ielmini, D. Neuromorphic Learning and Recognition with One-Transistor-One-Resistor Synapses and Bistable Metal Oxide RRAM. *IEEE Trans. Electron. Devices* **2016**, *63*, 1508–1515. [CrossRef]
4. Tosson, A.M.S.; Yu, S.; Anis, M.H.; Wei, L. A Study of the Effect of RRAM Reliability Soft Errors on the Performance of RRAM-Based Neuromorphic Systems. *IEEE Trans. Very Large Scale Intergr. Vlsi Syst.* **2017**, *25*, 3125–3137. [CrossRef]
5. Kuzum, D.; Jeyasingh, R.G.D.; Lee, B.; Philip Wong, H.-S. Nanoelectronic Programmable Synapses Based on Phase Change Materials for Brain-Inspired Computing. *Nano Lett.* **2011**, *12*, 2179–2186. [CrossRef] [PubMed]
6. Tan, S.H.; Lin, P.; Yeon, S.; Choi, S.; Park, Y.; Kim, J. Perspective: Uniform switching of artificial synapses for large-scale neuromorphic arrays. *Appl. Phys. Lett.* **2018**, *6*, 120901. [CrossRef]
7. Hoffer, B.; Rana, V.; Menzel, S.; Waser, R.; Kvatinsky, S. Experimental Demonstration of Memristor-Aided Logic (MAGIC) Using Valence Change Memory (VCM. *IEEE Trans. Electron. Devices* **2020**, *67*, 3115–3122. [CrossRef]
8. Liao, Y.; Wu, H.; Wan, W.; Zhang, W.; Gao, B.; Philip Wong, H.-S.; Qian, H. Novel In-Memory Matrix-Matrix Multiplication with Resistive Cross-Point Arrays. In Proceedings of the 2018 Symposium on VLSI Technology (VLSIT), Honolulu, HI, USA, 18–22 June 2018; pp. 31–32. [CrossRef]
9. Huang, P.; Zhou, Y.; Xiang, Y.; Han, R.; Liu, L.; Liu, X.; Kang, J. Hardware implementation of RRAM based binarized neural networks. *APL Mater.* **2019**, *7*, 081105. [CrossRef]
10. Kang, S.; Lee, J.; Kang, M.; Song, Y. Achievement of Gradual Conductance Characteristics Based on Interfacial Phase-Change Memory for Artificial Synapse Applications. *Electronics* **2020**, *9*, 1268. [CrossRef]
11. Malavena, G.; Filippi, M.; Spinelli, A.S.; Compagnoni, C.M. Unsupervised Learning by Spike-Timing-Dependent Plasticity in a Mainstream NOR Flash Memory Array—Part I: Cell Operation. *IEEE Trans. Electron. Devices* **2019**, *66*, 4727–4732. [CrossRef]
12. Malavena, G.; Filippi, M.; Spinelli, A.S.; Compagnoni, C.M. Unsupervised Learning by Spike-Timing-Dependent Plasticity in a Mainstream NOR Flash Memory Array—Part II: Array Learning. *IEEE Trans. Electron. Devices* **2019**, *66*, 4733–4738. [CrossRef]
13. Guo, X.; Merrikh Bayat, F.; Prezioso, M.; Chen, Y.; Nguyen, B.; Do, N.; Strukov, D.B. Temperature-insensitive analog vector-by-matrix multiplier based on 55 nm NOR flash memory cells. In Proceedings of the IEEE Custom Integrated Circuits Conference (CICC), Austin, TX, USA, 30 April–3 May 2017; pp. 1–4. [CrossRef]
14. Malavena, G.; Spinelli, A.S.; Monzio Compagnoni, C. Implementing Spike-Timing-Dependent Plasticity and Unsupervised Learning in a Mainstream NOR Flash Memory Array. In Proceedings of the International Electron Device Meeting (IEDM), San Francisco, CA, USA, 1–5 December 2018; pp. 1–4. [CrossRef]
15. Kim, C.-H.; Lee, S.; Woo, S.Y.; Kang, W.-M.; Lim, S.; Bae, J.-H.; Kim, J.; Lee, J.-H. Demonstration of Unsupervised Learning with Spike-Timing-Dependent Plasticity Using a TFT-Type NOR Flash Memory Array. *IEEE Trans. Electron. Devices* **2018**, *65*, 1774–1780. [CrossRef]
16. Guo, X.; Merrikh Bayat, F.; Bavandpour, M.; Klachko, M.; Mahmoodi, M.R.; Prezioso, M.; Likharev, K.K.; Strukov, D.B. Fast, Energy-Efficient, Robust, and Reproducible Mixed-Signal Neuromorphic Classifier Based on Embedded NOR Flash Memory Technology. In Proceedings of the International Electron Device Meeting (IEDM), San Francisco, CA, USA, 2–6 December 2017; pp. 1–4. [CrossRef]
17. Noh, Y.; Seo, Y.; Park, B.-G.; Lee, J.-H. Synaptic Devices Based on 3-D AND Flash Memory Architecture for Neuromorphic Computing. In Proceedings of the International Memory Workshop (IMW), Monterey, CA, USA, 12–15 May 2019; pp. 1–4. [CrossRef]

18. Malavena, G.; Petro, S.; Spinelli, A.S.; Monzio Compagnoni, C. Impact of Program Accuracy and Random Telegraph Noise on the Performance of a NOR Flash-based Neuromorphic Classifier. In Proceedings of the ESSDERC 2019—49th European Solid-State Device Research Conference (ESSDERC), Cracow, Poland, 23–26 September 2019; pp. 122–125. [CrossRef]
19. Wang, Y.; Zhao, Y.; Khan, B.M.; Doherty, C.L.; Krayer, J.D.; White, M.H. A novel SONOS nonvolatile flash memory device using substrate hot-hole injection for write and gate tunneling for erase. *Solid State Electron.* **2004**, *48*, 2031–2034. [CrossRef]
20. Kim, H.; Park, J.; Kwon, M.-W.; Lee, J.-H.; Park, B.-G. Silicon-Based floating-body synaptic transistor with frequency-dependent short- and long-term memories. *IEEE Electron. Device Lett.* **2016**, *37*, 249–252. [CrossRef]
21. Baek, M.-H.; Jang, T.; Kim, H.; Park, J.; Kwon, M.-W.; Hwang, S.; Kim, S.; Lee, J.-J.; Park, B.-G. Grain boundary induced short-term memory effect in fully depleted thin-polysilicon devices. *Jpn. J. Appl. Phys.* **2019**, *58*, 1–5. [CrossRef]
22. Mody, J.; Duffy, R.; Eyben, P.; Goossens, J.; Moussa, A.; Polspoel, W.; Berghmans, B.; Dal, M.J.H.V.; Pawlak, B.J.; Kaiser, M.; et al. Experimental studies of dose retention and activation in fin field-effect-transistor-based structures. *J. Vac. Sci. Technol. B* **2010**, *28*, C1H5. [CrossRef]
23. Wang, L.; Brown, A.R.; Cheng, B.; Asenov, A. Simulation of 3D FinFET Doping Profiles by Ion Implantation. In Proceedings of the American Institute of Physics (AIP) Conference Proceedings, Valladolid, Spain, 25–29 June 2012; pp. 217–220. [CrossRef]
24. Duffy, R.; Curatola, G.; Pawlak, B.J.; Doornbos, G.; Tak, K.V.D.; Breimer, P.; van Berkum, J.G.M.; Roozeboom, F. Doping fin field-effect transistor sidewalls: Impurity dose retention in silicon due to high angle incident ion implants and the impact on device performance. *J. Vac. Sci. Technol. B* **2008**, *26*, 402–407. [CrossRef]
25. Cho, S.; Li, D.H.; Kim, D.-H.; Cho, I.H.; Park, B.-G. Channel Doping Concentration and Fin Width Effects on Self-Boosting in NAND-Type SONOS Flash Memory Array Based on Bulk-FinFETs. In Proceedings of the Nanotechnology Materials and Devices Conference (NMDC), Traverse City, Michigan, USA, 2–5 June 2009; pp. 251–254. [CrossRef]
26. Liu, Y.; Matsukawa, T.; Endo, K.; O'uchi, S.; Sakamoto, K.; Tsukada, J.; Ishikawa, Y.; Yamauchi, H.; Masahara, M. Investigation of Low-Energy Tilted Ion Implantation for Fin-Type Double-Gate Metal–Oxide–Semiconductor Field-Effect Transistor Extension Doping. *Jpn. J. Appl. Phys.* **2010**, *49*, 04DC18. [CrossRef]
27. Lenoble, D.; Doornbos, G.; de Keersgieter, A.; Pawlak, B.; Vandervorst, W.; Skotnicki, T. The junction challenges in the FinFETs device. In Proceedings of the International Workshop on Junction Technology, Shanghai, China, 15–16 May 2006; pp. 78–83. [CrossRef]
28. Chang, R.-D.; Lin, P.-H. Simulation Study of Implantation Angle Variation and Its Impact on Device Performance. In Proceedings of the International Conference on Ion Implantation Technology (IIT), Tainan, Taiwan, 26–30 September 2016; pp. 1–4. [CrossRef]
29. Witters, L.; Son, N.J.; San, T.; Singanamalla, R.; Kerner, C.; Collaert, N.; De Meyer, K.; Jurczak, M. Treshold voltage modulation in FinFET devices through Arsenic Ion Implantation into TiN/HfSiON gate stack. In Proceedings of the IEEE International SOI Conference, Indian Wells, CA, USA, 1–4 October 2007; p. 32. [CrossRef]
30. Mody, J.; Kambham, A.K.; Zschatzsch, G.; Schatzer, P.; Chiarella, T.; Collaert, N.; Witters, L.; Jurczak, M.; Horiguchi, N.; Gilbert, M.; et al. Dopant and carrier profiling in FinFET-based devices with sub-nanometer resolution. In Proceedings of the 2010 Symposium on VLSI Technology (VLSIT), Honolulu, HI, USA, 15–17 June 2010; pp. 195–196. [CrossRef]
31. Gossmann, H.-J.L.; Agarwal, A.; Parrill, T.; Rubin, L.M.; Poate, J.M. On the FinFET extension implant energy. *IEEE Trans. Nanotechnol.* **2003**, *2*, 285–290. [CrossRef]
32. Kedzierski, J.; Ieong, M.; Nowak, E.; Kanarsky, T.S.; Zhang, Y.; Roy, R.; Boyd, D.; Fired, D.; Wong, H.-S.P. Extension and source/drain design for high-performance FinFET devices. *IEEE Trans. Electron. Devices* **2003**, *50*, 952–958. [CrossRef]
33. Zschatzsch, G.; Sasaki, Y.; Hayashi, S.; Togo, M.; Chiarella, T.; Kambham, A.K.; Mody, J.; Douhard, B.; Horiguchi, N.; Mizuno, B.; et al. High performance n-MOS finFET by damage-free, conformal extension doping. In Proceedings of the International Electron Device Meeting (IEDM), Wasington, DC, USA, 5–7 December 2011; pp. 1–4. [CrossRef]

34. Lue, H.-T.; Hsu, T.-H.; Wang, S.-Y.; Lai, E.-K.; Hsieh, K.-Y.; Liu, R.; Lu, C.-Y. Study of incremental step pulse programming (ISPP) and STI edge effect of BE-SONOS NAND Flash. In Proceedings of the IEEE International Reliability Physics Symposium, Phoenix, AZ, USA, 27 April–1 May 2008; pp. 693–694. [CrossRef]
35. Kim, M.; Kim, S.; Shin, H. A Compact Model for ISPP of 3-D Charge-Trap NAND Flash Memories. *IEEE Trans. Electron. Devices* **2020**, *67*, 3095–3101. [CrossRef]
36. Chen, W.-C.; Lue, H.-T.; Hsiao, Y.-H.; Hsu, T.-H.; Lin, X.-W.; Lu, C.-Y. Charge storage efficiency (CSE) effect in modeling the incremental step pulse programming (ISPP) in charge-trapping 3D NAND flash devices. In Proceedings of the International Electron Device Meeting (IEDM), Washington, DC, USA, 7–9 December 2015; pp. 1–4. [CrossRef]
37. Maconi, A.; Monzio Compagnoni, C.; Amoroso, S.M.; Mascellino, E.; Ghidotti, M.; Padovini, G.; Spinelli, A.S.; Lacaita, A.L.; Mauri, A.; Ghidini, G.; et al. Investigation of the ISPP dynamics and of the programming efficiency of charge-trap memories. In Proceedings of the European Solid State Device Research Conference, Sevilla, Spain, 14–16 September 2010; pp. 444–447. [CrossRef]
38. Liu, S.; Zou, X. QLC NAND study and enhanced Gray coding methods for sixteen-level-based program algorithms. *Microelectron. J.* **2017**, *66*, 58–66. [CrossRef]
39. Kim, S.; Baek, M.-H.; Hwang, S.; Jang, T.; Park, K.; Park, B.-G. A Novel Vector-matrix Multiplication (VMM) Architecture based on NAND Memory Array. *J. Semicond. Tech. Sci.* **2020**, *20*, 242–248. [CrossRef]
40. Lee, S.-T.; Lim, S.; Choi, N.Y.; Bae, J.-H.; Kwon, D.; Park, B.-G.; Lee, J.-H. Operation Scheme of Multi-Layer Neural Networks Using NAND Flash Memory as High-Density Synaptic Devices. *IEEE J. Electron. Devices Soc.* **2019**, *7*, 1085–1093. [CrossRef]
41. Lee, S.-T.; Kim, H.; Bae, J.-H.; Yoo, H.; Choi, N.Y.; Kwon, D.; Lim, S.; Park, B.-G.; Lee, J.-H. High-Density and Highly-Reliable Binary Neural Networks Using NAND Flash Memory Cells as Synaptic Devices. In Proceedings of the International Electron Device Meeting (IEDM), San Francisco, CA, USA, 7–11 December 2019; pp. 1–4. [CrossRef]
42. Kim, M.; Song, J.; Kim, C.H. Reliability Characterization of Logic-Compatible NAND Flash Memory based Synapses with 3-bit per Cell Weights and 1µA Current Steps. In Proceedings of the International Reliability Physics Symposium (IRPS), Dallas, TX, USA, 28–30 April 2020; pp. 1–4. [CrossRef]
43. Wang, P.; Xu, F.; Wang, B.; Gao, B.; Wu, H.; Qian, H.; Yu, S. Three-Dimensional NAND Flash for Vector–Matrix Multiplication. *IEEE Trans. Very Large Scale Intergr. Vlsi Syst.* **2019**, *27*, 988–991. [CrossRef]
44. Minucci, U.; Santis, L.D.; Vali, T.; Irrera, F. A Neural Network implemented on NAND memory. In Proceedings of the 2020 IEEE International Memory Workshop (IMW), Dresden, Germany, 17–20 May 2020; pp. 1–4. [CrossRef]
45. Lue, H.-T.; Lai, S.-C.; Hsu, T.-H.; Du, P.-Y.; Wang, S.-Y.; Hsieh, K.-Y.; Liu, R.; Lu, C.-Y. Understanding barrier engineered charge-trapping NAND flash devices with and without high-K dielectric. In Proceedings of the IEEE International Reliability Physics Symposium, Montreal, QC, Canada, 26–30 April 2009; pp. 874–882. [CrossRef]
46. Kang, M.; Park, K.-T.; Song, Y.; Hwang, S.; Choi, B.Y.; Song, Y.; Lee, Y.-T.; Kim, C. Improving Read Disturb Characteristics by Self-Boosting Read Scheme for Multilevel NAND Flash Memories. *Jpn. J. Appl. Phys.* **2009**, *48*, 04C062. [CrossRef]
47. Lue, H.-T.; Hsu, P.-K.; Wang, K.-C.; Lu, C.-Y. Introduction of Non-Volatile Computing in Memory (nvCIM) by 3D NAND Flash for Inference Accelerator of Deep Neural Network (DNN) and the Read Disturb Reliability Evaluation. In Proceedings of the IEEE International Reliability Physics Symposium, Dallas, TX, USA, 28–30 May 2020; pp. 1–6. [CrossRef]
48. Kobayashi, A.; Tokutomi, T.; Takeuchi, K. Versatile TLC NAND flash memory control to reduce read disturb errors by 85% and extend read cycles by 6.7-times of Read-Hot and Cold data for cloud data centers. In Proceedings of the 2016 Symposium on VLSI Technology (VLSIT), Honolulu, Hawaii, 15–17 June 2016; pp. 1–2. [CrossRef]
49. Zhang, Y.; Jin, L.; Jiang, D.; Zou, X.; Liu, H.; Huo, Z. A Novel Read Scheme for Read Disturbance Suppression in 3D NAND Flash Memory. *IEEE Electron. Device Lett.* **2017**, *12*, 1669–1672. [CrossRef]
50. Della Marca, V.; Postel-Pellerin, J.; Just, G.; Canet, P.; Ogier, J.-L. Impact of endurance degradation on the programming efficiency and the energy consumption of NOR flash memories. *Microelectron. Reliab.* **2014**, *54*, 2262–2265. [CrossRef]

51. Driussi, F.; Esseni, D.; Selmi, L. Performance, degradation monitors, and reliability of the CHISEL injection regime. *IEEE Trans. Device Mater. Reliab.* **2004**, *4*, 327–334. [CrossRef]
52. Postel-Pellerin, J.; Chiquet, P.; Della Marca, V.; Wakrim, T.; Just, G.; Ogier, J.L. Improving Flash memory endurance and consumption with ultra-short channel-hot-electron programming pulses. In Proceedings of the International Semiconductor Conference (CAS), Sinaia, Romania, 13–15 October 2014; pp. 197–200. [CrossRef]
53. Park, M.; Park, S.; Lee, D.-K.; Jeong, Y.; Hong, C.; Lee, H.S.; Cho, M.K.; Ahn, K.-O.; Koh, Y. NAND Flash reliability degradation induced by HCI in boosted channel potential. In Proceedings of the IEEE International Reliability Physics Symposium, Anaheim, CA, USA, 2–6 May 2010; p. MY.7.1-2.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).