



Article

Optimal Energetic-Trap Distribution of Nano-Scaled Charge Trap Nitride for Wider V_{th} Window in 3D NAND Flash Using a Machine-Learning Method

Kihoon Nam ¹, Chanyang Park ¹, Jun-Sik Yoon ¹, Hyeok Yun ¹, Hyundong Jang ¹, Kyeongrae Cho ¹, Ho-Jung Kang ², Min-Sang Park ³, Jaesung Sim ², Hyun-Chul Choi ⁴ and Rock-Hyun Baek ^{1,*}

¹ Department of Electrical Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Korea; namgee4970@postech.ac.kr (K.N.); chypark@postech.ac.kr (C.P.); junsikyoon@postech.ac.kr (J.-S.Y.); myska315@postech.ac.kr (H.Y.); hdjang@postech.ac.kr (H.J.); krcho5252@postech.ac.kr (K.C.)

² SK hynix Inc., Cheongju-si 28429, Chungcheongbuk-do, Korea; hojung.kang@sk.com (H.-J.K.); jaesung1.sim@sk.com (J.S.)

³ SK hynix Inc., Icheon-si 17336, Gyeonggi-do, Korea; minsang.park@sk.com

⁴ Department of Electrical Engineering, Yeungnam University, Gyeongsan 38541, Korea; pogary@ynu.ac.kr

* Correspondence: rh.baek@postech.ac.kr

Abstract: A machine-learning (ML) technique was used to optimize the energetic-trap distributions of nano-scaled charge trap nitride (CTN) in 3D NAND Flash to widen the threshold voltage (V_{th}) window, which is crucial for NAND operation. The energetic-trap distribution is a critical material property of the CTN that affects the V_{th} window between the erase and program V_{th} . An artificial neural network (ANN) was used to model the relationship between the energetic-trap distributions as an input parameter and the V_{th} window as an output parameter. A well-trained ANN was used with the gradient-descent method to determine the specific inputs that maximize the outputs. The trap densities (N_{TD} and N_{TA}) and their standard deviations (σ_{TD} and σ_{TA}) were found to most strongly impact the V_{th} window. As they increased, the V_{th} window increased because of the availability of a larger number of trap sites. Finally, when the ML-optimized energetic-trap distributions were simulated, the V_{th} window increased by 49% compared with the experimental value under the same bias condition. Therefore, the developed ML technique can be applied to optimize cell transistor processes by determining the material properties of the CTN in 3D NAND Flash.

Keywords: 3D NAND Flash; charge trap nitride; gradient-descent method; machine learning; multi-level cell; trap distribution; threshold voltage window



Citation: Nam, K.; Park, C.; Yoon, J.-S.; Yun, H.; Jang, H.; Cho, K.; Kang, H.-J.; Park, M.-S.; Sim, J.; Choi, H.-C.; et al. Optimal Energetic-Trap Distribution of Nano-Scaled Charge Trap Nitride for Wider V_{th} Window in 3D NAND Flash Using a Machine-Learning Method. *Nanomaterials* **2022**, *12*, 1808. <https://doi.org/10.3390/nano12111808>

Academic Editors: Yung-Hsien Wu, Albert Chin and Mads Brandbyge

Received: 28 March 2022

Accepted: 23 May 2022

Published: 25 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, solid-state drives (SSDs) have been developed to satisfy the requirements for rapidly storing and using large amounts of data. NAND Flash is a key and widely used component in SSDs because of its suitability for mass production and reasonable bit cost [1]. NAND Flash has been developed with innovation in three-dimensional (3D) architecture and charge trap nitride (CTN) as a storage node [2–5]. Further, the bit cost of NAND Flash has been reduced using multi-level cell (MLC) technology, which stores multiple bits in each single cell [6].

An MLC creates more threshold-voltage (V_{th}) states in one cell by subdividing the amount of charge injected into the storage node. As more V_{th} states are created, the number of bits stored per cell increases, and therefore, the bit cost considerably reduces [7,8]. Unfortunately, the number of sections with overlaps between subdivided V_{th} states also increases, making it difficult to distinguish the actual V_{th} states and thereby degrading the device reliability [9–11].

Multiple V_{th} states are created within a limited V_{th} window between the erase (ERS) and the program (PGM) V_{th} [12,13]. Then, if the V_{th} window is widened, the errors in distinguishing V_{th} states can be reduced by obtaining the margin of each V_{th} state. In the previous works, wide V_{th} windows were achieved by adopting new materials and bias conditions [14,15]. However, qualitative analyses for improving the V_{th} window have not sufficiently focused on the material properties of the storage node.

The energetic-trap distribution is an important material property of the CTN, a nano-scaled thin film in 3D NAND Flash. Previously, it was discussed based on the analytical model for charges [16]. The energetic-trap distribution can be extracted using retention models [17] and trap spectroscopy by charge injection and sensing (TSCIS) [18]. Furthermore, the energetic-trap distributions are determined by the gas flow ratio deposited on the CTN [19,20]. These studies indicated that energetic-trap distributions are controllable and that their profiles determine the ERS V_{th} ($V_{th,ers}$) and PGM V_{th} ($V_{th,pgm}$) [17,21]. We tried to maximize the V_{th} window, that is, the sum of the absolute value of $V_{th,ers}$ ($|V_{th,ers}|$) and $V_{th,pgm}$, by optimizing the energetic-trap distributions.

We also used a novel machine-learning (ML) method to improve the V_{th} window. Recently, ML has been used for predicting and optimizing nanoscale transistors [22–24]. Moreover, it can help quantitatively determine the complex material properties of the CTN with high speed and accuracy. Therefore, an artificial neural network (ANN) [25] was trained by modeling eight inputs that determine the energetic-trap distributions and the two outputs $|V_{th,ers}|$ and $V_{th,pgm}$. Next, the energetic-trap distributions were investigated to realize a large V_{th} window by the gradient-descent method, a widely used optimization algorithm.

In this study, the ML method helps determine the optimized energetic-trap distributions and thereby control the material properties of the CTN in 3D NAND Flash. We quantitatively found the energetic-trap distributions that resulted in the largest V_{th} window. The remainder of this paper is organized as follows. In Section 2, the simulation data and ML method used to train the ANN are introduced. In Section 3, we discuss the results of training and optimization of ML-based analysis. Finally, the conclusions are drawn in Section 4.

2. Simulation and Machine-Learning Method

We obtained data from TCAD simulations for training the ANN [26]. Shockley–Read–Hall recombination, drift-diffusion transport, mobility (high-field, doping and interface dependent) and Hurkx band-to-band tunneling models were adopted for the poly-Si channel. Furthermore, the nonlocal tunneling model was used to describe charge transport in ERS/PGM operations.

Figure 1 shows the schematic diagrams of 3D NAND Flash used in the simulation. Figure 1a shows a part of the 3D NAND Flash string. It has a cylindrical structure, and both ends are connected by a bit line (BL) and a source line (SL). V_{th} is extracted from the BL current (I_{BL}) of 1 μA vs. selected word line (WL_{sel}) voltage in the middle of the string. Figure 1b shows the half cross-sectional view of the cylindrical structure and a gate stack. Furthermore, ERS (PGM) state is the condition when holes (electrons) are filled in the CTN. The simulation structure was derived from the manufactured device [27].

Figure 2 shows two energetic-trap distributions following the Gaussian distribution of the bandgap in the CTN. Donor- and acceptor-like traps capture holes and electrons, respectively. Each distribution consisted of the trap densities (N_{TD} and N_{TA}), peak energy levels (E_{TD} and E_{TA}), capture cross sections (CCS_D and CCS_A), and standard deviations (σ_D and σ_A). These eight inputs determine the profile of the distributions, and $V_{th,ers}$ and $V_{th,pgm}$ change accordingly.

Figure 3 shows V_{th} with the ERS/PGM operating time. Each operation was performed by applying a single pulse; the PGM operation started from the ERS state, whereas the ERS operation started from the PGM state. Under the same bias condition ($V_{PGM} = 15\text{ V}$, $V_{ERS} = 20\text{ V}$), the simulation data (solid lines) were calibrated with the experimental data

(symbols) by adjusting the energetic-trap distributions of the CTN. The calibrated simulation data were then used as training data for the ML-based analysis. In this study, the V_{th} window is the sum of $|V_{th,ers}|$ and $V_{th,pgm}$ at time = 10^{-2} s. The corresponding V_{th} window contains the difference between the ERS and the highest PGM V_{th} states of an MLC.

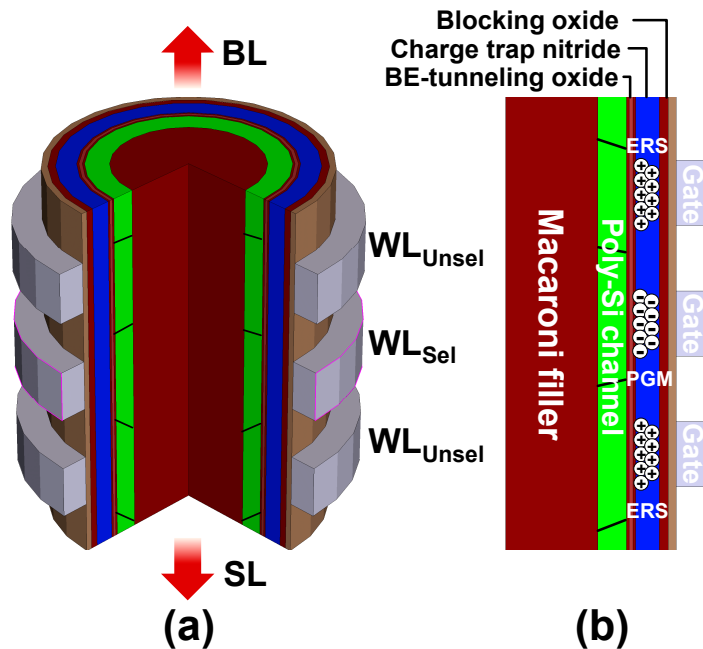


Figure 1. (a) Schematic diagram of the cylindrical structure of 3D NAND Flash. (b) Half cross-sectional view of string and cell stack. Depending on whether holes or electrons are present in the CTN, the state is ERS or PGM.

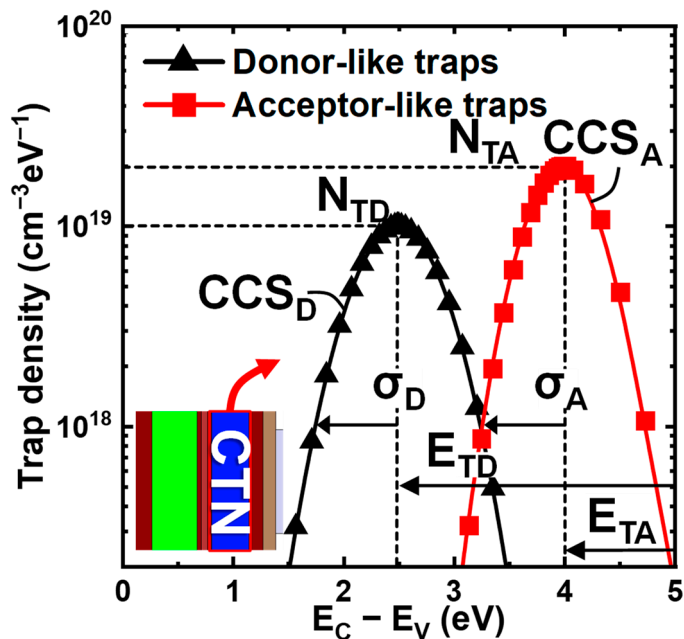


Figure 2. Two energetic-trap distributions comprising donor- and acceptor-like traps of the CTN. The peak energy levels (E_{TD} and E_{TA}) are from the conduction band (E_C). Their profiles determine the V_{th} window.

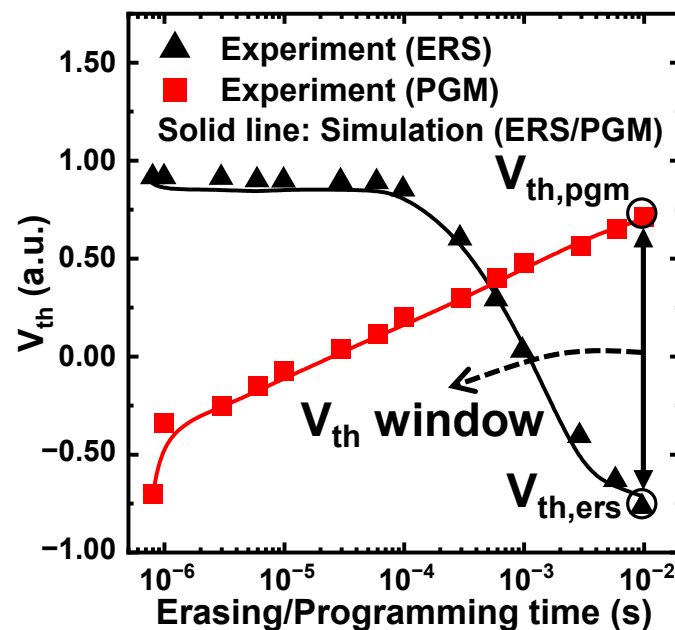


Figure 3. V_{th} vs. operating time started from different ERS/PGM states during a single pulse operation. The simulation (solid line) and the experimental (symbol) data are in good agreement. The V_{th} window is the sum of $|V_{th,ers}|$ and $V_{th,pgm}$ at time = 10^{-2} s.

Figure 4 shows a schematic of the ANN. First, datasets are produced from TCAD simulation for 3D NAND Flash cell. These extracted datasets are used to design the ANN. A multilayer perceptron (MLP), in which several layers of perceptrons are sequentially attached, was used for training [28]. MLP is suitable for solving nonlinear functions that cannot be solved using a single-layer perceptron, so it is useful for the training complex model. In this study, this feedforward network contains one hidden layer with 15 nodes and tanh as an activation function. In addition, it was trained using MATLAB to model the inputs and outputs [29]. The Levenberg–Marquardt method was used for the back-propagation of training, and the cost function was calculated by the mean squared error (MSE). The MSE indicates the accuracy of ML training. The smaller it is, the better is the training of the ANN. After training, the gradient-descent method was used in the backward direction to determine the optimal inputs that resulted in a large V_{th} window. This method is widely used for finding the minimum value of the cost function in ML. Finally, we set the cost function in the direction of making large outputs. In summary, we trained the ANN using well-calibrated simulation data and then used the well-trained ANN to determine the optimal inputs that resulted in the largest outputs.

Table 1 summarizes the calibrated values and ranges of the eight inputs for training and optimizing of the ANN. These inputs were randomized within each range uniformly, and different $|V_{th,ers}|$ and $V_{th,pgm}$ values were derived, accordingly. N_{TD} , N_{TA} , CCS_D , and CCS_A were logged, and then all parameters were standardized to improve the prediction accuracy of the ANN. In addition, there is no correlation between inputs. The calibrated values are generally the median of each range, and the entire range is reasonable [27,30]. Other material properties of the CTN were fixed only to verify the effect of energetic-trap distributions.

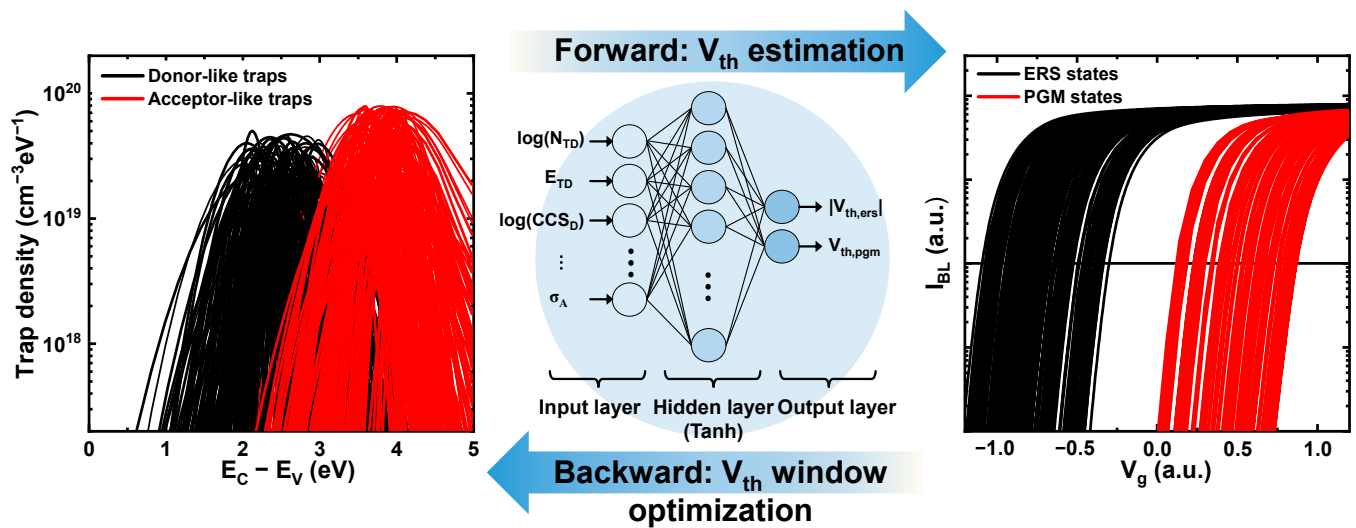


Figure 4. Scheme of eight inputs (Table 1) for two outputs ($|V_{th,ers}|$ and $V_{th,pgm}$) in ANN. An MLP was used as a learning algorithm to train the ANN. In the backward direction, the eight inputs can be found in the direction of making the two outputs large.

Table 1. Input parameters of energetic-trap distributions of the CTN.

Trap Parameter	Min	Max	Calibrated
Density of donor-like traps, N_{TD} ($\text{cm}^{-3}\cdot\text{eV}^{-1}$)	5.00×10^{18}	5.00×10^{19}	1.00×10^{19}
Peak energy level of donor-like traps, E_{TD} (eV)	2.00	3.00	2.50
Capture cross section of donor-like traps, CCS_D (cm^2)	1.00×10^{-15}	1.00×10^{-11}	1.00×10^{-13}
Standard deviation of donor-like traps, σ_D (eV)	0.10	0.50	0.35
Density of acceptor-like traps, N_{TA} ($\text{cm}^{-3}\cdot\text{eV}^{-1}$)	8.00×10^{18}	8.00×10^{19}	2.00×10^{19}
Peak energy level of acceptor-like traps, E_{TA} (eV)	0.80	1.50	1.00
Capture cross section of acceptor-like traps, CCS_A (cm^2)	1.00×10^{-15}	1.00×10^{-11}	1.00×10^{-13}
Standard deviation of acceptor-like traps, σ_A (eV)	0.10	0.50	0.30

3. Results and Discussion

Figure 5 shows the example of MSEs in the training and validation sets with epochs. Here, epoch refers to the number of times the entire data have passed through the neural network. A total data of 1980 samples were used for ML training. First, they were used for training and test sets with a weight of 80/20. Then, the training sets were divided five-fold. Each fold became a validation set once, and the mean value of the five evaluations was used to determine the performance of the corresponding model. We also repeated this entire process five times by splitting the nodes of the hidden layer to increase the reliability of the model. As a result, the 15 nodes of the hidden layer have the lowest MSE, and the results of applying the model to test sets are shown in Figure 6. Furthermore, the model was well-generalized with no overfitting and underfitting because of the five-fold cross validation.

Figure 6a compares the simulated and the estimated values of $|V_{th,ers}|$ and $V_{th,pgm}$ in the test sets. For one input set, two outputs appear simultaneously, and the estimation is expected to be good as the two symbols overlap. Figure 6b shows the errors in the individual values. Each error is extremely small and is within the acceptable range of $\pm 5\%$. Furthermore, each MSE is 0.7368×10^{-3} and 0.8472×10^{-3} ; therefore, the prediction accuracy is very high. These results confirmed that the ANN was trained well. Furthermore, the ANN model was superior in learning speed and accuracy in our datasets compared with other regression models, the random forest, etc.

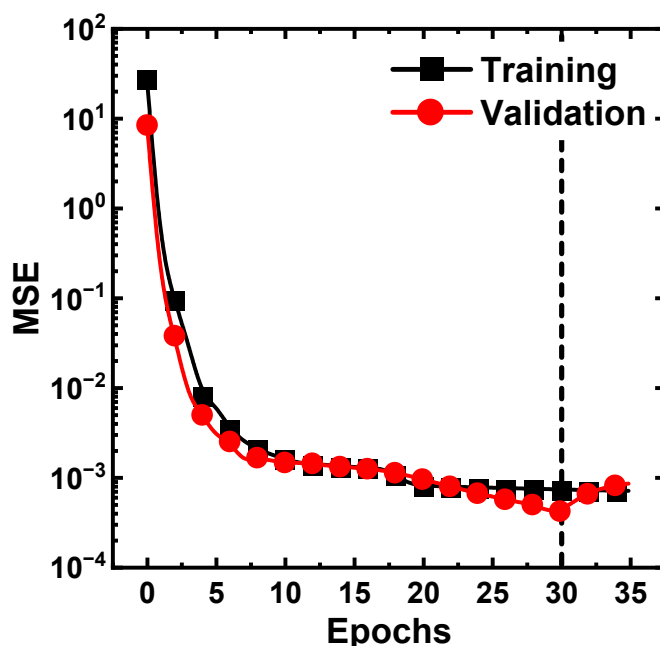


Figure 5. Example of MSE vs. epochs in the training and validation sets. MSEs were used as a measure of training. The model adopted five-fold cross validation later used to evaluate the test sets.

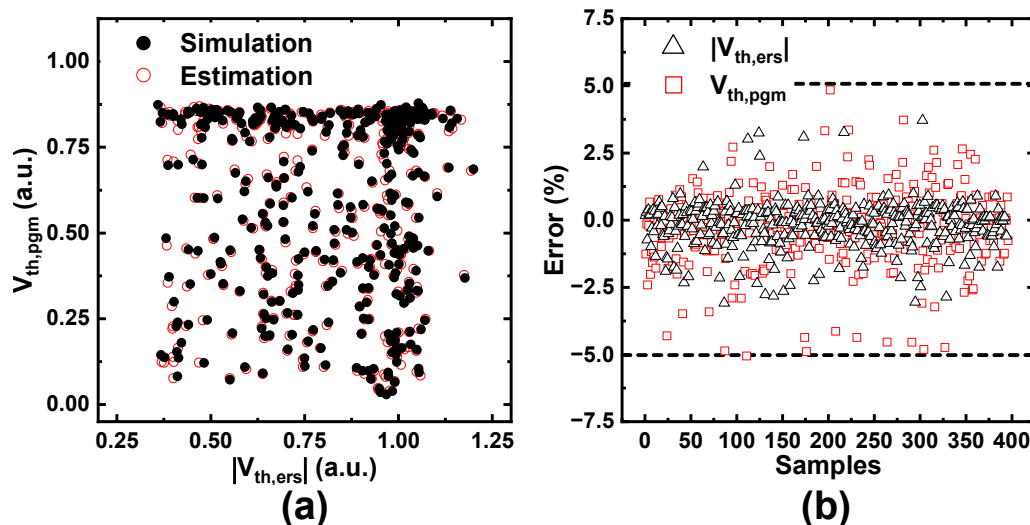


Figure 6. (a) Comparisons of $|V_{th,ers}|$ and $V_{th,pgm}$ between simulated (closed, black) and estimated (open, red) data, and (b) errors are within $\pm 5\%$ in the test sets. MSEs are 0.7368×10^{-3} and 0.8472×10^{-3} , respectively.

Figure 7 shows the raw simulated and optimized values of the V_{th} window; the latter are all larger than the former. We set the cost function in the direction of maximizing the V_{th} window. The gradient-descent method was used to determine the slope of the cost function. After setting the random inputs within the range listed in Table 1, the estimated V_{th} window was derived using the well-trained ANN. Then, the random inputs were set again until the slope reached the extreme value. Finally, the inputs that resulted in a large V_{th} window were found. Table 2 lists the optimized inputs that resulted in the largest V_{th} window after 2000 iterations.

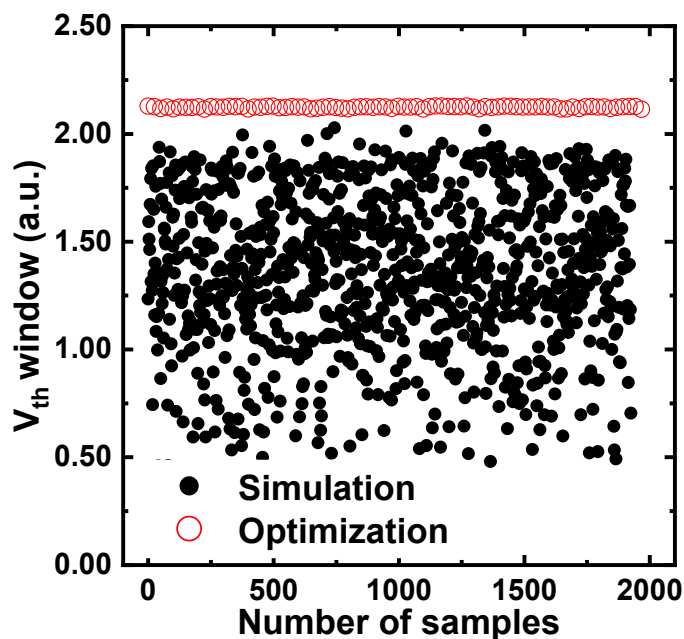


Figure 7. Comparison of optimized values (open, red) and simulated values (closed, black) of V_{th} window. A large value means that the V_{th} window is widened.

Table 2. Optimized input parameters of energetic-trap distributions of the CTN.

Trap Parameter	Value
N_{TD} ($\text{cm}^{-3} \cdot \text{eV}^{-1}$)	5.00×10^{19}
E_{TD} (eV)	2.00
CCS_D (cm^2)	1.00×10^{-15}
σ_D (eV)	0.50
N_{TA} ($\text{cm}^{-3} \cdot \text{eV}^{-1}$)	8.00×10^{19}
E_{TA} (eV)	1.45
CCS_A (cm^2)	1.00×10^{-15}
σ_A (eV)	0.50

Figure 8 compares the calibrated and the optimized energetic-trap distributions. The ML results indicated that N_{TD} , N_{TA} , σ_{TD} , and σ_{TA} strongly influenced the V_{th} window. When they increased, both $|V_{th,ers}|$ and $V_{th,pgm}$ increased significantly. This is because many available trap sites could capture more holes or electrons. When E_{TD} decreased and E_{TA} increased, both $|V_{th,ers}|$ and $V_{th,pgm}$ increased slightly. This is because the deeper energy of each distribution reduced the attempt-to-escape factor [31]. However, E_{TD} and E_{TA} mainly determined the retention characteristics; therefore, the correlation with the V_{th} window was weak. Similarly, CCS_D and CCS_A had small effects on the V_{th} window. In summary, large N_{TD} , N_{TA} , σ_{TD} , and σ_{TA} resulted in a large V_{th} window owing to the availability of a larger number of trap sites. However, E_{TD} , E_{TA} , CCS_D , and CCS_A had a small correlation with the V_{th} window because we did not consider the retention characteristics.

Figure 9 shows the experimental and simulated $|V_{th}|$ values optimized by ML. Here, the simulated V_{th} window was calculated from the best inputs in Table 2. In this case, the $|V_{th}|$ error between the simulation and the ANN was within 4.34%. Therefore, the simulated V_{th} windows were reliable. The V_{th} window increased by 49%. In other words, the V_{th} window can be sharply increased by optimizing the energetic-trap distributions. Therefore, we can provide a guideline for maximizing the V_{th} window, although the precise process remains difficult.

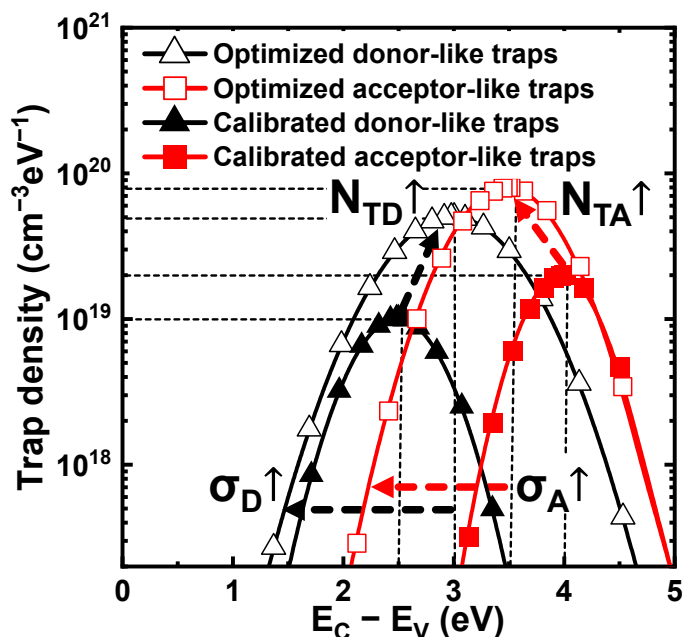


Figure 8. Comparison of profiles between optimized (open) and calibrated (closed) energetic-trap distributions. N_{TD} , N_{TA} , σ_D , and σ_A strongly influence the V_{th} window.

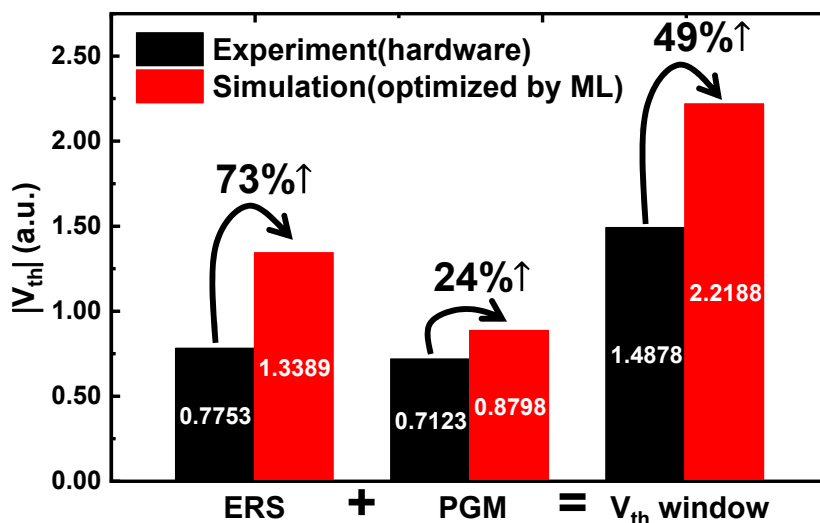


Figure 9. $|V_{th}|$ values in ERS/PGM operations and V_{th} window. Values estimated using ML outperform experimental values overall.

4. Conclusions

ML-based analysis was used to obtain optimized energetic-trap distributions for the CTN in 3D NAND Flash to improve the V_{th} window. The ANN enables modeling the relationship between eight inputs that determine the energetic-trap distributions and the two outputs, $|V_{th,ers}|$ and $V_{th,pgm}$. The ANN was trained using well-calibrated simulation data with experiments, and the MSEs were found to be small. Then, we used the gradient-descent method to determine the best inputs that resulted in the largest V_{th} window. N_{TD} , N_{TA} , σ_{TD} , and σ_{TA} significantly influenced the V_{th} window. As they increased, the V_{th} window grew because of the large number of trap sites. In particular, when the best inputs obtained using ML were employed, the V_{th} window increased by 49% compared with the experimental value. This study should enable the determination of the V_{th} window from the material properties of the CTN in 3D NAND Flash. More generally, this work implies

that ML can help to solve the complex problem of nanomaterials accurately and optimize it rapidly.

Author Contributions: Conceptualization, K.N.; methodology, K.N. and H.Y.; software, K.N. and J.-S.Y.; validation, K.N., C.P., H.J., J.-S.Y. and R.-H.B.; formal analysis, K.N. and H.Y.; investigation, K.N., C.P. and K.C.; resources, K.N., C.P., H.Y. and H.J.; data curation, K.N., C.P. and J.-S.Y.; writing—original draft preparation, K.N.; writing—review and editing, C.P., J.-S.Y. and R.-H.B.; visualization, K.N. and K.C.; supervision, H.-C.C. and R.-H.B.; project administration, H.-J.K. and M.-S.P.; funding acquisition, J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Semiconductor Industry Collaborative Project between POSTECH and SK hynix Inc., in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2020R1A4A4079777), and in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2020M3F3A2A02082436).

Data Availability Statement: Not applicable.

Acknowledgments: The EDA tool was supported by the IC Design Education Center (IDEC).

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

References

1. Ishimaru, K. Future of non-volatile memory from storage to computing. In Proceedings of the IEEE International Electron Device Meeting (IEDM), San Francisco, CA, USA, 5–8 December 2019; pp. 1.3.1–1.3.6. [\[CrossRef\]](#)
2. Choi, E.-S.; Park, S.-K. Device considerations for high density and highly reliable 3D NAND Flash cell in near future. In Proceedings of the IEEE International Electron Device Meeting (IEDM), San Francisco, CA, USA, 5–7 December 2011; pp. 211–214. [\[CrossRef\]](#)
3. Katsumata, R.; Kito, M.; Fukuzumi, Y.; Kido, M.; Tanaka, H.; Komori, Y.; Ishiduki, M.; Matsunami, J.; Fujiwara, T.; Nagata, Y.; et al. Pipe-shaped BiCS Flash memory with 16 stacked layers and multi-level-cell operation for ultra high density storage devices. In Proceedings of the Symposium on VLSI Technology, Kyoto, Japan, 15–17 June 2009; pp. 136–137.
4. Jang, J.; Kim, H.-S.; Cho, W.; Cho, H.; Kim, J.; Shim, S.I.; Jang, Y.; Jeong, J.-H.; Son, B.-K.; Kim, D.W.; et al. Vertical cell array using TCAT(Terabit Cell Array Transistor) technology for ultra high density NAND Flash memory. In Proceedings of the Symposium on VLSI Technology, Kyoto, Japan, 15–17 June 2009; pp. 192–193.
5. Tanaka, H.; Kido, M.; Yahashi, K.; Oomura, M.; Katsumata, R.; Kito, M.; Fukuzumi, Y.; Sato, M.; Nagata, Y.; Matsuoka, Y.; et al. Bit cost scalable technology with punch and plug process for ultra high density Flash memory. In Proceedings of the Symposium on VLSI Technology, Kyoto, Japan, 12–14 June 2007; pp. 14–15. [\[CrossRef\]](#)
6. Lee, S.; Lee, Y.-T.; Han, W.-K.; Kim, D.-H.; Kim, M.-S.; Moon, S.-H.; Cho, H.C.; Lee, J.-W.; Byeon, D.-S.; Lim, Y.-H.; et al. A 3.3 V 4 Gb four-level NAND Flash memory with 90 nm CMOS technology. In Proceedings of the IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 15–19 February 2004; pp. 52–53. [\[CrossRef\]](#)
7. Lee, S.; Kim, C.; Kim, M.; Joe, S.-M.; Jang, J.; Kim, S.; Lee, K.; Kim, J.; Park, J.; Lee, H.-J.; et al. A 1Tb 4b/Cell 64-Stacked-W 3D NAND Flash memory with 12MB/s program throughput. In Proceedings of the IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 11–15 February 2018; pp. 340–342. [\[CrossRef\]](#)
8. Shibata, N.; Kanda, K.; Shimizu, T.; Nakai, J.; Nagao, O.; Kobayashi, N.; Miakashi, M.; Nagadomi, Y.; Nakano, T.; Kawabe, T.; et al. A 1.33-Tb 4-bit/cell 3-D Flash memory on a 96-word-line-layer technology. *IEEE J. Solid-State Circuits* **2020**, *55*, 178–188. [\[CrossRef\]](#)
9. Mizoguchi, K.; Takahashi, T.; Aritome, S.; Takeuchi, K. Data retention characteristics comparison of 2D and 3D TLC NAND Flash memories. In Proceedings of the IEEE Memory Workshop (IMW), Monterey, CA, USA, 14–17 May 2017; pp. 1–4. [\[CrossRef\]](#)
10. Shibata, N.; Maejima, H.; Isobe, K.; Iwasa, K.; Nakagawa, M.; Fujiu, M.; Shimizu, T.; Honma, M.; Hoshi, S.; Kawaai, T.; et al. A 70 nm 16 Gb 16-level-cell NAND Flash memory. *IEEE J. Solid-State Circuits* **2008**, *43*, 929–937. [\[CrossRef\]](#)
11. Cai, Y.; Luo, Y.; Ghose, S.; Mutlu, O. Read disturb errors in MLC NAND Flash memory: Characterization, mitigation, and recovery. In Proceedings of the 45th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, Rio de Janeiro, Brazil, 22–25 June 2015; pp. 438–449. [\[CrossRef\]](#)
12. Parat, K.; Goda, A. Scaling trends in NAND Flash. In Proceedings of the IEEE International Electron Device Meeting (IEDM), San Francisco, CA, USA, 1–5 December 2018; pp. 2.1.1–2.1.4. [\[CrossRef\]](#)
13. Yoon, C.-W.; Kim, H.-G.; Lee, S.-K.; Lee, J.; Song, J.H. Issues and key technologies for next generation 3D NAND. In Proceedings of the International Conference on Electronics, Information, and Communication (ICEIC), Jeju, Korea, 31 January–3 February 2021; pp. 1–4. [\[CrossRef\]](#)
14. Shin, Y.; Choi, J.; Kang, C.; Lee, C.; Park, K.-T.; Lee, J.-S.; Sel, J.; Kim, V.; Choi, B.; Sim, J.; et al. A novel NAND-type MONOS memory using 63nm process technology for multi-gigabit flash EEPROMs. In Proceedings of the IEEE International Electron Device Meeting (IEDM), Washington, DC, USA, 5–7 December 2005; pp. 327–330. [\[CrossRef\]](#)

15. Lee, C.-H.; Choi, J.; Kang, C.; Shin, Y.; Lee, J.-S.; Sel, J.; Sim, J.; Jeon, S.; Choe, B.-I.; Bae, D.; et al. Multi-level NAND flash memory with 63 nm-node TANOS (Si-Oxide-SiN-Al₂O₃-TaN) cell structure. In Proceedings of the Symposium on VLSI Technology, Honolulu, HI, USA, 13–15 June 2006; pp. 21–22. [[CrossRef](#)]
16. Yang, Y.; White, M.H. Charge retention of scaled SONOS non-volatile memory devices at elevated temperatures. *Solid-State Electron.* **2000**, *44*, 949–958. [[CrossRef](#)]
17. Kang, H.-J.; Choi, N.; Joe, S.-M.; Seo, J.-H.; Choi, E.; Park, S.-K.; Park, B.-G.; Lee, J.-H. Comprehensive analysis of retention characteristics in 3-D NAND Flash memory cells with tube-type poly-Si channel structure. In Proceedings of the Symposium on VLSI Technology, Kyoto, Japan, 27 August 2015; pp. T182–T183. [[CrossRef](#)]
18. Degraeve, R.; Cho, M.; Govoreanu, B.; Kaczer, B.; Zahid, M.B.; Van Houdt, J.; Jurczak, M.; Groeseneken, G. Trap spectroscopy by injection and sensing (TSCIS): A quantitative electrical technique for studying defects in dielectric stacks. In Proceedings of the IEEE International Electron Device Meeting (IEDM), San Francisco, CA, USA, 15–17 December 2008; pp. 1–4. [[CrossRef](#)]
19. Kim, T.H.; Park, I.H.; Lee, J.D.; Shin, H.C.; Park, B.-G. Electron trap density distribution of Si-rich silicon nitride extracted using the modified negative charge decay model of silicon-oxide-nitride-oxide-silicon structure at elevated temperatures. *Appl. Phys. Lett.* **2006**, *89*, 063508. [[CrossRef](#)]
20. Suhane, A.; Arreghini, A.; Degraeve, R.; Van den bosch, G.; Breuil, L.; Zahid, M.B.; Jurczak, M.; De Meyer, K.; Van Houdt, J. Validation of retention modeling as a trap-profiling technique for SiN-Based charge-trapping memories. *IEEE Electron Device Lett.* **2010**, *31*, 77–79. [[CrossRef](#)]
21. Choi, E.-S.; Yoo, H.-S.; Park, K.-H.; Kim, S.-J.; Ahn, J.-R.; Lee, M.-S.; Hong, Y.-O.; Kim, S.-G.; Om, J.-C.; Joo, M.-S.; et al. Modeling and characterization of program/erase speed and retention of TiN-gate MANOS (Si-Oxide-SiN_x-Al₂O₃-Metal gate) cells for NAND Flash memory. In Proceedings of the IEEE Non-Volatile Semiconductor Memory Workshop (NVSMW), Monterey, CA, USA, 26–30 August 2007; pp. 83–84. [[CrossRef](#)]
22. Choi, H.-C.; Yun, H.; Yoon, J.-S.; Baek, R.-H. Neural approach for modeling and optimizing Si-MOSFET manufacturing. *IEEE Access* **2020**, *8*, 159351–159370. [[CrossRef](#)]
23. Yun, H.; Yoon, J.-S.; Jeong, J.; Lee, S.; Choi, H.-C.; Baek, R.-H. Neural network based design optimization of 14-nm node fully-depleted SOI FET for SoC and 3DIC applications. *IEEE J. Electron Devices Soc.* **2020**, *8*, 1272–1280. [[CrossRef](#)]
24. Yoon, J.-S.; Lee, S.; Yun, H.; Baek, R.-H. Digital/Analog performance optimization of vertical nanowire FETs using machine learning. *IEEE Access* **2021**, *9*, 29071–29077. [[CrossRef](#)]
25. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [[CrossRef](#)] [[PubMed](#)]
26. *Sentaurus Device User Guide*; Version: O-2018.06-SP2; Synopsys, Inc.: Mountain View, CA, USA, 2018.
27. Nam, K.; Park, C.; Yoon, J.-S.; Jang, H.; Park, M.S.; Sim, J.; Baek, R.-H. Origin of incremental step pulse programming (ISPP) slope degradation in charge trap nitride base multi-layer 3D NAND Flash. *Solid-State Electron.* **2021**, *175*, 1–6. [[CrossRef](#)]
28. Funahashi, K.-I. On the approximate realization of continuous mappings by neural networks. *Neural Netw.* **1989**, *2*, 183–192. [[CrossRef](#)]
29. *MATLAB User Manual*; Version: R2019b; MathWorks, Inc.: Natick, MA, USA, 2019.
30. Gritsenko, V.A.; Nekrashevich, S.S.; Vasilev, V.V.; Shaposhnikov, A.V. Electronic structure of memory traps in silicon nitride. *Microelectron. Eng.* **2009**, *86*, 1866–1869. [[CrossRef](#)]
31. Vishnyakov, A.V.; Novikov, Y.N.; Gritsenko, V.A.; Nasyrov, K.A. The charge transport mechanism in silicon nitride: Multi-phonon trap ionization. *Solid-State Electron.* **2009**, *53*, 251–255. [[CrossRef](#)]