micromachines

Article

Effect of Noncircular Channel on Distribution of Threshold Voltage in 3D NAND Flash Memory

Donghyun Go [1], Gilsang Yoon [1], Jounghun Park [1], Donghwi Kim [1], Jiwon Kim [1], Jungsik Kim [2,*] and Jeong-Soo Lee [1,*]

[1] Department of Electrical Engineering, Pohang University of Science and Technology, Pohang 37673, Republic of Korea; godhyun@postech.ac.kr (D.G.); ygs6233@postech.ac.kr (G.Y.); mancity@postech.ac.kr (J.P.); kdh3879@postech.ac.kr (D.K.); jiwon104@postech.ac.kr (J.K.)
[2] Department of Electrical Engineering, Gyeongsang National University, Jinju 52828, Republic of Korea
* Correspondence: jungsik@gnu.ac.kr (J.K.); ljs6951@postech.ac.kr (J.-S.L.); Tel.: +82-54-279-2380 (J.-S.L.)

Abstract: The instability in threshold voltage (V_{TH}) and charge distributions in noncircular cells of three-dimensional (3D) NAND flash memory are investigated. Using TCAD simulation, we aim to identify the main factors influencing the V_{TH} of noncircular cells. The key focus is on the nonuniform trapped electron density in the charge trapping layer (CTL) caused by the change in electric field between the circular region and the spike region. There are less-trapped electron (LT) regions within the CTL of programmed noncircular cells, which significantly enhances current flow. Remarkably, more than 50% of the total current flows through these LT regions when the spike size reaches 15 nm. We also performed a comprehensive analysis of the relationship between charge distribution and V_{TH} in two-spike cells with different heights (H_{Spike}) and angles between spikes (θ). The results of this study demonstrate the potential to improve the reliability of next-generation 3D NAND flash memory.

Keywords: 3D NAND flash memory; noncircular cell; spike; TCAD simulation; threshold voltage distribution; trapped charge

Citation: Go, D.; Yoon, G.; Park, J.; Kim, D.; Kim, J.; Kim, J.; Lee, J.-S. Effect of Noncircular Channel on Distribution of Threshold Voltage in 3D NAND Flash Memory. *Micromachines* **2023**, *14*, 2007. https://doi.org/10.3390/mi14112007

Academic Editor: Aiqun Liu

Received: 31 July 2023
Revised: 11 October 2023
Accepted: 27 October 2023
Published: 28 October 2023

1. Introduction

Two-dimensional (2D) NAND flash memory, a conventional data storage technology, is characterized by the arrangement of memory cells in a two-dimensional structure for the storage and retrieval of data. Nonetheless, as the demand for increased storage capacity persists, and the relentless pursuit of cost reduction per bit continues, 2D NAND technology has encountered notable challenges that have impeded its progress. These limitations encompass scaling issues, restricted durability, and ongoing challenges in cost reduction during manufacturing. Consequently, the emergence of 3D NAND technology has become an imperative solution to surmount these challenges and extend the horizons delineated by Moore's Law, primarily by achieving enhanced memory densities.

Three-dimensional (3D) NAND technology is primarily characterized by vertically stacked structures, forming a robust foundation for storing more data within a reduced physical footprint. A way to achieve improved capacity and cost efficiency in 3D NAND is to increase the word line (WL) stack [1–4]. The evolutionary changes in this architecture significantly transcend the limitations of traditional 2D NAND technology, effectively addressing the mounting demands for data storage and accessibility in modern computing devices.

Recently, 3D NAND technology has made remarkable progress with the adoption of the high-aspect-ratio (HAR) gate-all-around (GAA) structure, representing an advancement in NAND flash memory design. The GAA structure is designed to encompass the channel of a memory cell, thereby minimizing the gaps between memory cells and facilitating a higher cell packing density within the same physical space. This innovation translates to

enhanced storage capacity within a given area, addressing the ever-increasing demand for more data storage. Moreover, the use of GAA polysilicon channels offers improved gate controllability and superior management of electrical fluctuations [5–8]. These advancements in control mechanisms have paved the way for the successful commercialization of triple-level cell (TLC) and quadruple-level cell (QLC) technologies within the domain of 3D NAND. This achievement has further enhanced the appeal and versatility of this memory technology [9–12]. However, it is essential to acknowledge that the transition to multi-level cells presents a unique set of challenges, particularly concerning threshold voltage (V_{TH}) distribution. The narrower margins inherent in the programmed V_{TH} distribution of multilevel cells introduce novel complexities and potential issues, diverging from the relative simplicity of single-level cells (SLCs) [13]. Incorporating the HAR structure into 3D NAND introduces a notable rise in process complexity, which, in turn, influences electrical performance and reliability [14–17]. This technology relies on the utilization of multiple stacked control gates and insulators crafted from alternating thin films of metal and silicon dioxide (SiO_2). In this approach, a control gate is substituted for sacrificial silicon nitride (Si_3N_4), and holes within the stack are created through plasma etching. The procedure alternates between these thin layers of SiO_2 and Si_3N_4. The cumulative stacking of these layers and the increasing number of successive layers result in a rapid rise in aspect ratio. Consequently, the fabrication of HAR structures presents its own set of challenges. The presence of small feature sizes and high plasma density during the manufacturing process pose potential risks. While it has been possible to achieve etching of up to 128 layers, this technique encounters limitations when pushed beyond this threshold, primarily due to the constraints of available plasma etching methods. The management of the HAR etch process is a complex undertaking, and it introduces a spectrum of profile distortions, including tapering, center-line tilt, and warping [18–24]. Simultaneously, a microtrench structure frequently forms close to the bottom region of the tapered etch hole. The cross-section of the microtrench generates a noncircular channel with spike-wise deformation, which can further degrade the uniformity of the electrical characteristics [25,26]. Although the electrical behavior of noncircular channel shape has been reported in some studies [27], it is not clear how the spike affects electrical properties of the cell and how V_{TH} is determined by the spike.

Our study utilized Synopsys Sentaurus technology for computer-aided design (TCAD) to investigate the complexity of trap and channel inversion electron density (e-density) distributions within noncircular cells characterized by single or two spikes. Our objective is to establish a clear correlation between channel current density and V_{TH} with respect to spike height and the angle between spikes. Furthermore, we intend to provide valuable insights into the V_{TH} distribution of spiked noncircular cells and reveal how these structural complexities affect the electrical properties of memory cells.

2. Simulation Structure and Methods

Figure 1a presents a schematic diagram illustrating the structure of 3D NAND flash memory, featuring three word lines (WLs). Each WL is separated by the spacer, and the central WL is designated as the target WL. Both the WL length and spacer length are set to 30 nm. The radial structure of a cell is composed of a metal gate, blocking oxide (BOX), charge trap layer (CTL), and band-engineered tunneling layer (BE-TOX) consisting of O1/N1/O2, a polysilicon (poly-Si) channel, and macaroni oxide. At both ends of the poly-Si channel, there are the source line (SL) and bit line (BL). Additionally, Figure 1b–d show cross-sectional schematic diagrams of the circular cell (C-cell), single-spike cell (SSC), and two-spike cell (TWSC), respectively. To investigate the influence of the gap between spikes in TWSC cells, we introduced the parameter θ, as shown in Figure 1d. In this analysis, we varied the height of the spike (H_{Spike}) and the angle between the two spikes (θ) over the ranges of 5 to 15 nm and 20° to 135°, respectively. Cell configurations featuring three or more spikes were not considered in this study due to the complexity associated with the etching process, which would require careful recalibration.

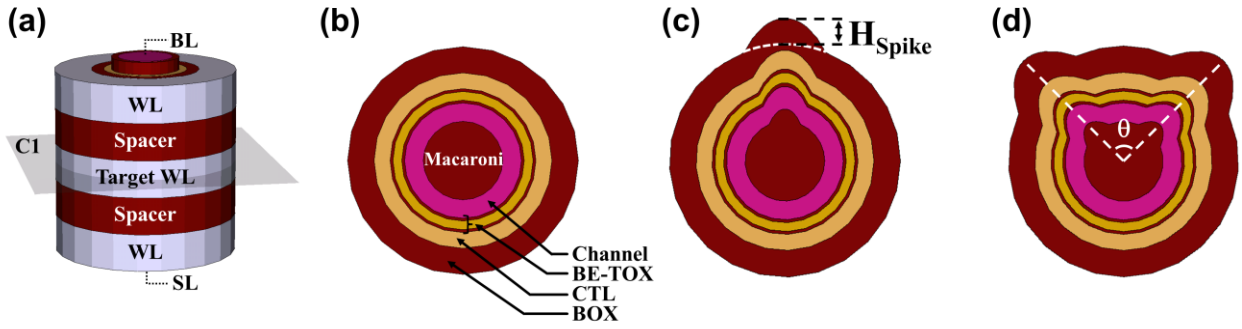
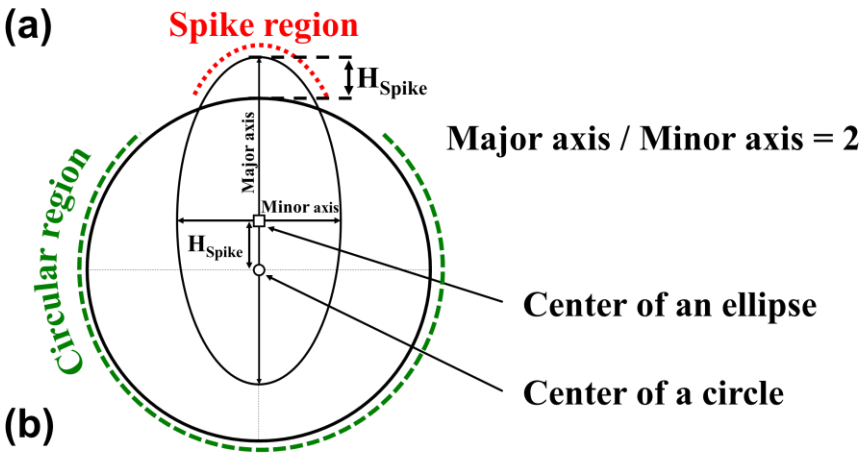


Figure 1. Schematic diagrams of simulated devices using gate-all-around (GAA) and vertical polysilicon (poly-Si) channel with macaroni channel structure. (**a**) Schematic diagrams of a 3D NAND flash memory string with three word lines (WLs), two spacers, a source line (SL), and a bit line (BL). Among the WLs, the middle WL is set as the target WL. (**b**) Cross-sectional schematics of the circular cell (C-cell), (**c**) single-spike cell (SSC), and (**d**) two-spike cell (TWSC).

Figure 2 shows a detailed cross-sectional view of the spike cell used in the simulation. In our simulation, a portion of the ellipse's structure is dedicated to shaping the spike shape. As shown in Figure 2a, the boundaries of the circular and elliptical shapes intersect, seamlessly creating the spike structure. Notably, the center of the ellipse is intentionally situated at a distance of H_{Spike} from the center of the circular shape. The outer perimeters of both the circle and the ellipse serve as the boundary between the metal gate and the BOX.



Region \ Layer		BOX (nm)	CTL (nm)	BE-TOX O2/N1/O1 (nm)	Channel (nm)	Macaroni (nm)
Circular region		10	6	1.1 / 3.3 / 1.1	7	30
Spike region	H_{Spike} = 5	12	7.2	1.3 / 4 / 1.3	7.7	30
	H_{Spike} = 10	12	7.2	1.3 / 4 / 1.3	9.1	34.3
	H_{Spike} = 15	12	7.2	1.3 / 4 / 1.3	9.1	39.3

Figure 2. Description of noncircular cell: (**a**) Outline of a noncircular cell created by overlapping circle and ellipse. (**b**) Detailed values for the thickness of each layer (BOX, CTL, BE-TOX, channel, and macaroni) in circular and spike regions for H_{spike} = 5, 10, and 15 nm.

In accordance with the thickness of each layer, as shown in Figure 2b, the layers are assembled in the following order, from the outermost to the innermost: BOX, CTL, BE-TOX, channel, and macaroni. The thickness of each layer in the spike region was intentionally configured to be greater than that of the circular region. Consequently, the cell is separated into a spike region and a circular region.

The Hurkx Band-to-Band Tunneling model was employed for analyzing gate-induced drain leakage, along with the high field saturation mobility model applied to the channel region. To ensure accurate analysis of channel current, the quantum confinement effect was integrated into the study using the density-gradient model. In our assumptions, we considered that there were no defects present in the oxide layers within the BE-TOX and BOX layers. To replicate the program/erase (P/E) operation, transient simulations were carried out, employing a nonlocal tunneling model. This nonlocal tunneling model was applied to both electrons and holes at the interface between the channel and BE-TOX as well as at the junction of the metal gate and BOX during P/E operations. In the oxide layer, the effective tunneling mass for both electrons and holes was defined as 0.4. The carriers within the charge trap layer (CTL) were conveyed according to the drift-diffusion model, and the capture or emission of carriers into traps was described using the Shockley–Read–Hall (SRH) model. The material properties of the charge trap layer (CTL) employed in this study can be found in Table 1. The electron and hole traps in Si_3N_4 both exhibit gaussian energy distributions. The spatial distribution of the traps was maintained as constant since our primary focus was on evaluating alterations in electrical properties attributable to the noncircular channel shape.

Table 1. Material properties of silicon nitride used in CTL.

Parameter	Value
Bandgap	5.0 eV
Peak Energy Level of Electron Trap	1.0 eV
Standard Deviation of Electron Trap	0.1 eV
Total Density of Electron Trap (N_T)	5×10^{19} cm^{-3}
Peak Energy Level of Hole Trap	2.5 eV
Standard Deviation of Hole Trap	0.1 eV
Total Density of Hole Trap	5×10^{18} cm^{-3}

For the program (PGM), erase (ERS), and read operations, the conditions were as follows: V_{PGM} = 16 V with t_{PGM} = 100 µs, V_{ERS} = −16 V with t_{ERS} = 1 ms, and V_{PASS} = 5 V, respectively. The voltage levels for the bit line (BL) and source line (SL) were maintained at 0.05 V and 0 V, respectively.

These simulations were conducted in an environment with a temperature of 300 K, which is equivalent to room temperature. Figure 3 illustrates the bit-line current (I_{BL}) plotted against the gate voltage (V_G) for the C-cell and cells with spikes at various states, including the initial, ERS, and PGM states. These results were derived from the 3D TCAD simulations. The current curve is graphed with both logarithmic and linear scales. To extract the V_{TH}, we employed the constant current method, identifying V_{TH} as the voltage at which 2 µA flows through the bit line (BL). In the initial, ERS, and PGM states, the V_{TH} values for the C-cell were measured at −0.5 V, −2 V, and 3.67 V, respectively. To explore the variation in V_{TH} as influenced by the number of spikes, a comparison was made between the C-cell, SSC (with H_{Spike} = 10 nm), and TWSC (with H_{Spike} = 10 nm and θ = 90°). In the case of the SSC and TWSC cells, it was observed that the V_{TH} values in the initial and ERS states exhibited minimal differences when compared to the V_{TH} of the C-cell, as depicted in Figure 3a. Nevertheless, during the PGM state, a noticeable negative shift in V_{TH} was distinctly observed for noncircular cells due to structural deformations, which is evident in Figure 3b. It is worth noting that TWSC cells with θ = 90° exhibited a lower V_{TH} compared to SSC. To gain a deeper understanding of their electrical behaviors, SSC and TWSC cells were subjected to simulations with varying H_{Spike} and θ values.

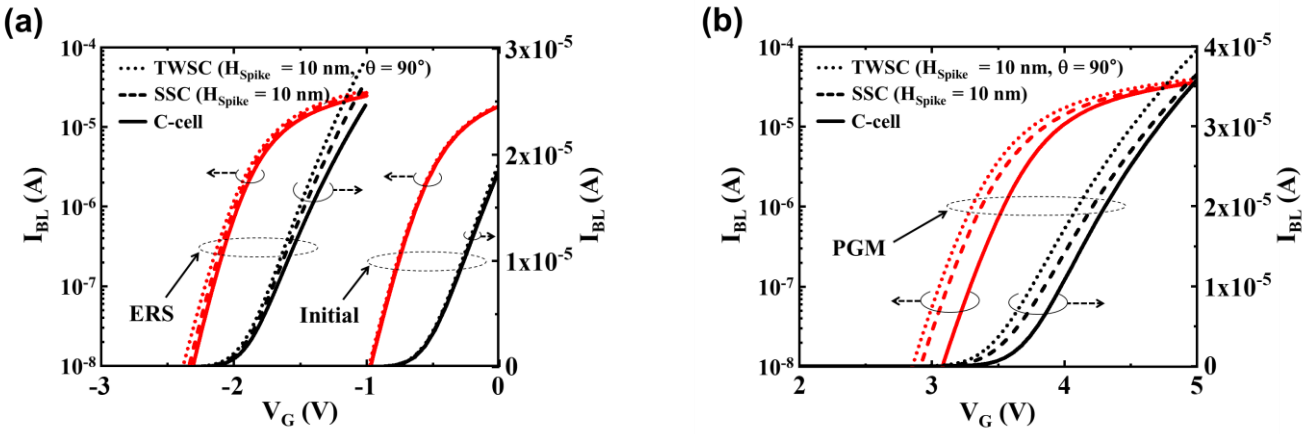


Figure 3. The bit-line current vs. gate voltage (I_{BL}–V_G) curves of the C-cell and noncircular cells plotted on the log and linear scales: (**a**) Initial, ERS, and (**b**) PGM states for the C-cell, SSC (H_{Spike} = 10 nm), and TWSC (H_{Spike} = 10 nm, θ = 90°) using 3D TCAD simulation.

3. Results and Discussion

3.1. SSC Characteristics

In Figure 4a, the distribution of trapped e-density within the charge trap layer (CTL) during the PGM state of the SSC is depicted. Notably, the trapped e-density in the CTL within the spike region surpasses that within the circular region. There is an observable gradual decline in trapped e-density within the space situated between the spike and circular regions. This region, where the trapped e-density experiences a reduction of 10% compared to the circular region, is now defined as the less-trapped electron (LT) region. It is essential to highlight that in the SSC, two distinct LT regions are formed.

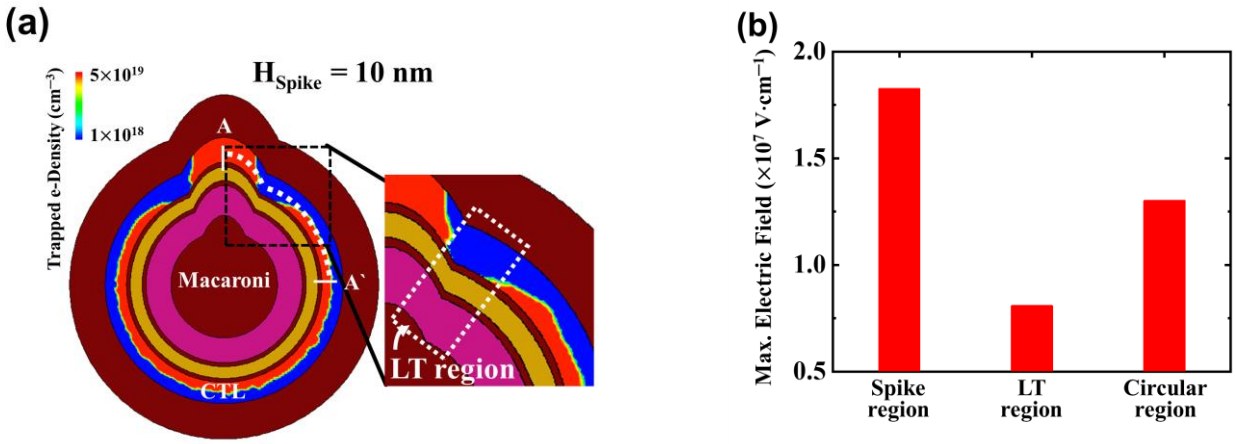


Figure 4. (**a**) Cross-sectional distribution of the trapped electron density (e-density) in the CTL in the PGM state of the SSC with 10 nm of H_{spike} and (**b**) maximum vertical electric field (E_{max}) extracted in the O1 layer of the BE-TOX layer by different regions (spike, LT, circular) during the PGM operation.

In Figure 4b, the diagram illustrates the maximum electric field (E_{max}) on the O1 layer of BE-TOX at V_{PGM} = 16 V in the SSC with H_{Spike} = 10 nm. In accordance with Gauss's law, the relationship between the electric field of BE-TOX and BOX, denoted as $E_{BE\text{-}TOX}$ and E_{BOX} respectively, is expressed as $E_{BE\text{-}TOX} = E_{BOX}\,(R + EOT)/R$, where R is the radius of curvature [28]. Given that the spike region exhibits a smaller R in comparison to the circular region, $E_{BE\text{-}TOX}$ in the spike region is notably higher. $E_{BE\text{-}TOX}$ directly impacts electron tunneling probability and trapped e-density during the PGM operation.

The concave shape observed in the LT region has the potential to reduce E_{max} and induce a lower quantity of trapped charges within the LT region.

In Figure 5a, the diagram shows the distribution of the average trapped e-density during the PGM state along the A-A′ perimeter, taking into account varying H_{Spike} values. The average trapped e-density was determined by radially integrating the cumulative charge at each position within the charge trap layer (CTL). In the case of the circular cell (C-cell), a consistent and uniform trapped e-density of 1.6×10^{19} cm^{-3} was established. As a result, the LT region in spike cells is defined as an area with a trapped e-density lower than 1.6×10^{18} cm^{-3}. For the SSC, a fully trapped e-density of 5×10^{19} cm^{-3} was identified in the spike region, while the circular region of SSC shared the same trapped e-density value as the C-cell. With an increase in H_{Spike}, the minimum trapped e-density decreased, and the LT region expanded from 5.6 to 10 nm. The distribution of trapped e-density is inherently dependent on the H_{Spike} values, which in turn affect the current flow within the channel and the corresponding V_{TH} values. As illustrated in Figure 5b, SSC exhibits a lower V_{TH} compared to the C-cell, and this value further decreases as H_{Spike} increases. In general, C-cells with lower trapped e-density within the CTL tend to have smaller V_{TH} values. However, as the H_{Spike} of SSC increases, the overall trapped e-density within the CTL increases, thus contradicting the trend observed for V_{TH} in C-cells.

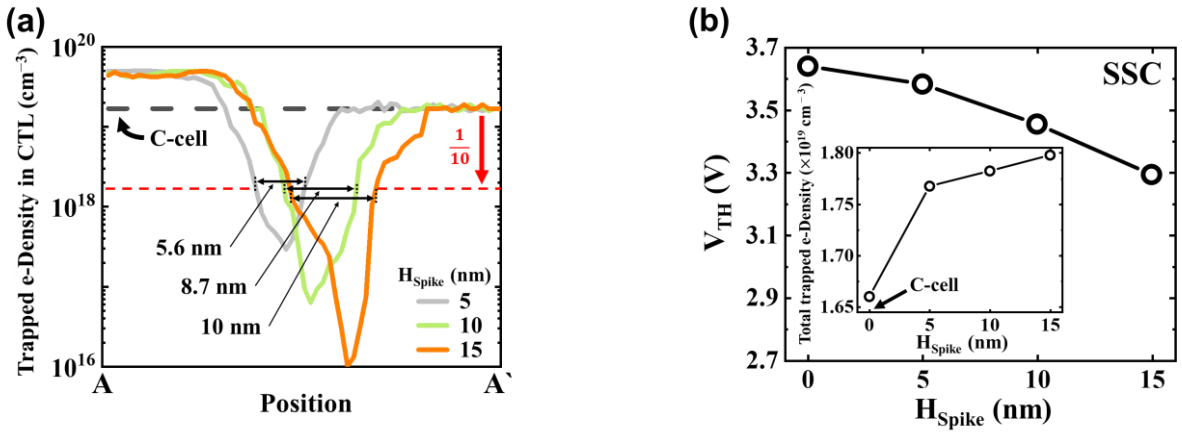


Figure 5. (**a**) Concentration of the trapped electron density (e-density) in CTL with different H_{Spike} values (5, 10, and 15 nm) along the A-A′ perimeter. The less-trapped electron (LT) region, where the e-density decreases by 10% of the circular region, is indicated with the red dashed line. (**b**) V_{TH} of the C-cell (H_{Spike} = 0 nm) and the SSC in the PGM state with different H_{Spike} values (0, 5, 10, and 15 nm). Inset: total trapped e-density vs. H_{Spike} value.

To elucidate the reason behind this contrasting V_{TH} trend in noncircular cells, a more comprehensive examination of the connection between trapped e-density in spike cells and channel current density, which directly influences V_{TH}, is imperative.

Figure 6a provides a cross-sectional view of channel electron distribution within the SSC with H_{Spike} = 10 nm, observed at $V_{read} = V_{TH}$. Despite the uniform application of V_{read} across all regions, the spike regions with high trapped e-density result in locally smaller channel e-density in comparison to other channel regions. However, the LT region, characterized by lower trapped e-density in the CTL, exhibits an increased channel electron count and subsequently a higher local current density. Figure 6b presents the average channel e-density along the A-A′ perimeter, considering various H_{Spike} values. The channel e-density for the C-cell is calculated to be 4.6×10^{16} cm^{-3}. Notably, the spike region displays the lowest local current, and this low current further decreases as H_{Spike} values increase. The channel e-density within the circular region aligns with that of the C-cell. A local e-density peak is evident in the LT region, and this peak coincides with the position of the minimum trapped e-density in the CTL (as shown in Figure 5a). With the increase in H_{Spike}, the local e-density peak value rises from 1.6×10^{17} to 2.1×10^{17} cm^{-3}. Figure 6c presents the proportion of currents flowing within the three regions concerning the total channel current (at $V_{read} = V_{TH}$) in programmed SSCs. The total channel current amounts

to 2 μA. Remarkably, less than 1% of the current is directed through the spike region. In the case of a circular cell (C-cell) with H_{Spike} = 0 nm, the total current belongs to the circular region. However, as H_{Spike} values increase, there is a notable rise in local current within the LT region. The normalized perimeter length, determined at H_{Spike} = 15 nm, indicates that 70% of the current is attributed to the circular region, 18% to the spike region, and 12% to the LT region. Despite the LT region having the smallest spatial area among the three regions, in SSCs with H_{Spike} = 15 nm, approximately 57% of the total current flows through the LT region.

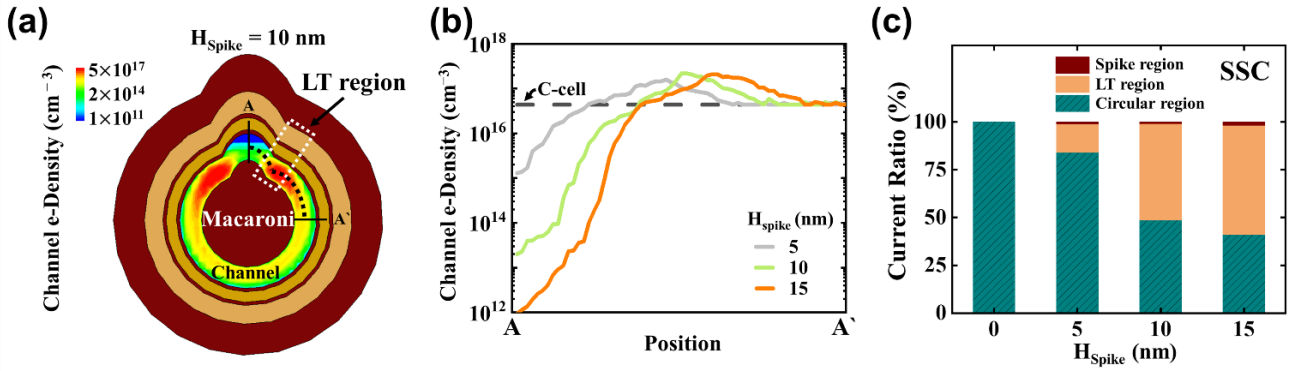


Figure 6. (**a**) Cross-sectional distribution of the channel e-density with H_{Spike} = 10 nm at $V_{read} = V_{TH}$; (**b**) channel e-density with different H_{Spike} along A-A′ perimeter; (**c**) ratio of current flowing through circular, LT, and spike regions to total channel current (I_{BL} = 2 μA) with different H_{Spike} values (0, 5, 10, and 15 nm).

In conclusion, it can be inferred that a relatively high current is directed towards the LT region in the programmed spike cell, ultimately resulting in a low VTH for the spike cell.

3.2. TWSC Characteristics

Figure 7 provides an overview of the trapped e-density within the CTL and the channel e-density at $V_{read} = V_{TH}$ in the vicinity of two spikes within the programmed TWSCs. In this simulation, H_{Spike} is kept constant at 10 nm, while the parameter θ is adjusted within the range of 20° to 135°. The insets display the cross-sectional distribution of trapped electrons in the CTL. Observations reveal that at θ = 135°, four distinct LT regions are identified, with two of these LT regions positioned within the B-B′ region (as shown in Figure 7d). Each LT region spans a length of 8.7 nm. Within all four LT regions, the location of minimum trapped e-density coincides with the peak location of local channel e-density. The two LT regions positioned between the spikes exhibit a minimum trapped e-density of 2.1 × 10^{16} cm^{-3}, and the peak channel e-density is 1.3 × 10^{17} cm^{-3}. As θ decreases to 90°, the two LT regions positioned between the spikes draw closer together, with no alteration in their length. As the circular region between the LT regions narrows, the relative proportion of LT regions within the B-B′ region increases significantly. Consequently, while the minimum trapped e-density remains unchanged, the peak channel e-density increases to 5 × 10^{17} cm^{-3} (as shown in Figure 7c). Further reducing the angle to θ = 45° in the TWSCs results in the merging of the two LT regions positioned between the spikes. At θ = 45°, there is a single LT region with a length of 8.1 nm between the spikes. The region situated between the two spikes exhibits a structure characterized by a low electric field, resulting in reduced levels of both trapped e-density in the CTL and channel e-density (as shown in Figure 7b). Upon reaching θ = 20°, the analysis reveals the presence of only two LT regions, each with a length of 8.7 nm. At both θ = 45° and θ = 20° within the TWSC, there is a notable absence of e-density between the two spikes in comparison to the circular region. This observation suggests that the majority of the channel current flows through the circular region and the two LT regions (as shown in Figure 7a). Furthermore, the spike region at θ = 20° is narrower in comparison to θ = 45°, indicating that the area experiencing current flow is broader.

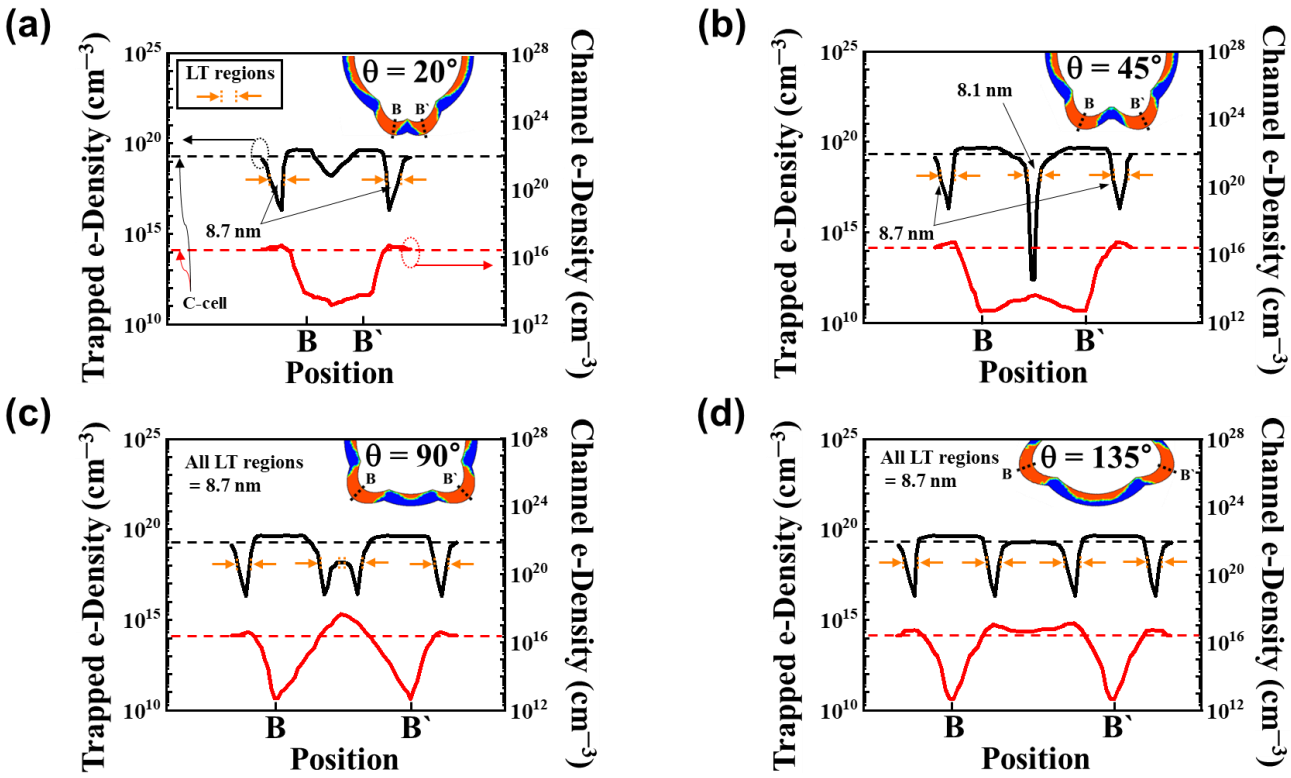


Figure 7. TWSC with H_{Spike} = 10 nm and θ for (**a**) 20°, (**b**) 45°, (**c**) 90°, and (**d**) 135° in PGM state. The trapped e-density in the CTL and channel e-density with different θ values along the B-B′ perimeter are shown. Each inset indicates the corresponding cross-sectional distribution of the trapped e-density in the CTL of programmed TWSCs.

Figure 8a illustrates the current ratio that flows through the LT regions relative to the total current as a function of θ. Notably, the current in the spike region remains less than 1%, emphasizing that the majority of the channel current is directed through the LT and circular regions. In the TWSC, the current ratio reaches its maximum at θ = 90°, irrespective of the spike's size. This ratio increases from 40% for H_{Spike} = 5 nm to 80% for H_{Spike} = 15 nm. At θ = 180°, where four LT regions are present, the current ratio is higher compared to θ = 0°, which has only two LT regions. Figure 8b presents the V_{TH} distribution for TWSCs with different H_{Spike} and θ values. The dashed line represents the V_{TH} value of the C-cell. All SSCs and TWSCs exhibit lower V_{TH} values than the C-cell, with a larger spike inducing a more substantial decrease in V_{TH}. Particularly, around θ = 90° in the TWSC cell, the noticeable increase in channel e-density between the spikes contributes to the reduced V_{TH} (as shown in Figure 7c). As the size of the spike increases, the V_{TH} distribution expands within the noncircular cell, leading to increased V_{TH} instability. Mitigation of the significant V_{TH} instability caused by the spike can be achieved through the implementation of a rigorous Incremental Step Pulse Program (ISPP) method. However, it is essential to strike a balance between V_{TH} distribution and PGM time when optimizing the ISPP conditions, as longer PGM times may lead to reduced operational speed. To enhance V_{TH} consistency, reducing the curvature in the spike region to achieve a uniform E-field is recommended. It is important to adjust the curvature when the layer thickness in the spike region changes. Controlling layer thickness to minimize V_{TH} variability in spike cells is an area of focus for future research and development.

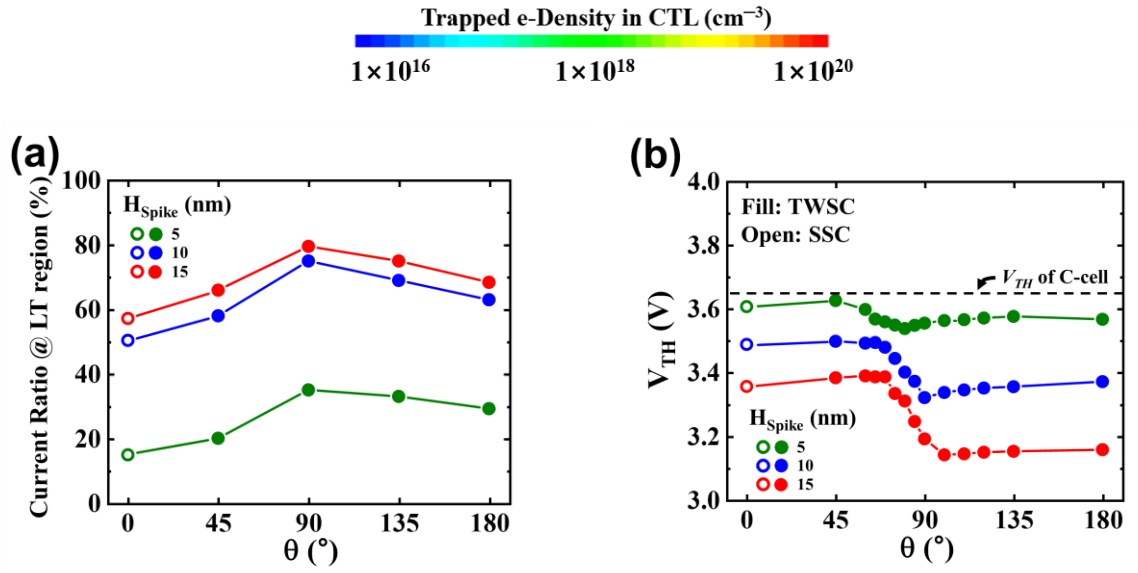


Figure 8. The electrical characteristics (current ratio and threshold voltage) as functions of H_{Spike} and θ for the TWSCs. (**a**) Ratio of current flowing in the LT regions to total channel current (I_{BL} = 2 μA) with the H_{Spike} and θ of TWSCs and (**b**) threshold voltage (V_{TH}) in programmed TWSCs depending on H_{Spike} and θ of spikes.

4. Conclusions

This study has delved into the instability and charge distribution of the threshold voltage (V_{TH}) in noncircular cells of three-dimensional (3D) NAND flash memory. Utilizing TCAD simulations with spike cells, we effectively pinpointed the crucial factors that exert significant influence on V_{TH}. A major focus is placed on the role of electric field fluctuations, which give rise to non-uniform trapped e-density within the charge trapping layer (CTL). Our investigation revealed the presence of regions with less-trapped (LT) regions within the CTL of programmed noncircular cells, resulting in increased current flow. Furthermore, we conducted a comprehensive analysis of the relationship between charge distribution and V_{TH} for two-spike cells with varying heights of spike (H_{Spike}) and angle between the spikes (θ). These findings clearly demonstrate that V_{TH} instability in noncircular cells intensifies as H_{Spike} values increase. The results of this study offer a comprehensive understanding of V_{TH} instability in noncircular cells, provide valuable insights for optimizing 3D NAND flash memory techniques, and furnish guidelines for enhancing both reliability and performance.

Author Contributions: Conceptualization, D.G. and J.K. (Jiwon Kim); methodology, J.P., D.K. and G.Y.; writing—original draft preparation, D.G., J.K. (Jungsik Kim) and J.-S.L.; writing—review and editing, D.G., J.K. (Jungsik Kim) and J.-S.L.; supervision, J.-S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by National R&D Program through the National Research Foundation of Korea (NRF) funded by Ministry of Science and ICT (2020M3H2A1078045) and by Samsung POSTECH Research Center (SPRC) funded by Samsung Electronics.

Acknowledgments: The EDA tool was supported by the IC Design Education Center. This research was supported by the "Leaders in INdustry-university Cooperation 3.0" Project funded by the Ministry of Education and National Research Foundation (NRF) of Korea.

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 2019 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 7–11 December 2019; pp. 1.3.1–1.3.6.
2. Heineck, L.; Liu, J. 3D NAND Flash Status and Trends. In Proceedings of the 2022 IEEE International Memory Workshop (IMW), Dresden, Germany, 15–18 May 2022; pp. 1–4.
3. Choe, J. Memory Technology 2021: Trends & Challenges. In Proceedings of the 2021 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Dallas, TX, USA, 27–29 September 2021; pp. 111–115.
4. Shirota, R. 3D-NAND flash memory and technology. In *Advances in Non-Volatile Memory and Storage Technology*; Woodhead Publishing: Sawston, UK, 2019; pp. 283–319.
5. Seo, J.Y.; Kim, Y.; Park, S.H.; Kim, W.; Kim, D.-B.; Lee, J.-H.; Shin, H.; Park, B.-G. Investigation into the effect of the variation of gate dimensions on program characteristics in 3D NAND flash array. In Proceedings of the 2012 IEEE Silicon Nanoelectronics Workshop (SNW), Honolulu, HI, USA, 10–11 June 2012; pp. 1–2.
6. Silvagni, A. 3D NAND Flash Based on Planar Cells. *Computers* **2017**, *6*, 28. [CrossRef]
7. Goda, A. 3-D NAND Technology Achievements and Future Scaling Perspectives. *IEEE Trans. Electron Devices* **2020**, *67*, 1373–1381.
8. Barraud, S.; Previtali, B.; Vizioz, C.; Hartmann, J.-M.; Sturm, J.; Lassarre, J.; Perrot, C.; Rodriguez, P.; Loup, V.; Magalhaes-Lucas, A.; et al. 7-Levels-Stacked Nanosheet GAA Transistors for High Performance Computing. In Proceedings of the 2020 IEEE Symposium on VLSI Technology, Honolulu, HI, USA, 16–19 June 2020; pp. 1–2.
9. Liang, S.; Qiao, Z.; Tang, S.; Hochstetler, J.; Fu, S.; Shi, W.; Chen, H.-B. An Empirical Study of Quad-Level Cell (QLC) NAND Flash SSDs for Big Data Applications. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; pp. 3676–3685.
10. Papandreou, N.; Ioannou, N.; Parnell, T.; Pletka, R.; Stanisavljevic, M.; Stoica, R.; Tomic, S.; Pozidis, H. Reliability of 3D NAND flash memory with a focus on read voltage calibration from a system aspect. In Proceedings of the 2019 19th Non-Volatile Memory Technology Symposium (NVMTS), Durham, NC, USA, 28–30 October 2019; pp. 1–4.
11. Papandreou, N.; Pozidis, H.; Parnell, T.; Ioannou, N.; Pletka, R.; Tomic, S.; Breen, P.; Tressler, G.; Fry, A.; Fisher, T. Characterization and Analysis of Bit Errors in 3D TLC NAND Flash Memory. In Proceedings of the 2019 IEEE International Reliability Physics Symposium (IRPS), Monterey, CA, USA, 31 March–4 April 2019; pp. 1–6.
12. Mizoguchi, K.; Takahashi, T.; Aritome, S.; Takeuchi, K. DataRetention Characteristics Comparison of 2D and 3D TLC NAND Flash Memories. In Proceedings of the 2017 IEEE International Memory Workshop (IMW), Monterey, CA, USA, 14–17 May 2017; pp. 1–4.
13. Fukuchi, M.; Suzuki, S.; Maeda, K.; Matsui, C.; Takeuchi, K. BER Evaluation System Considering Device Characteristics of TLC and QLC NAND Flash Memories in Hybrid SSDs with Real Storage Workloads. In Proceedings of the 2021 IEEE Int. Symp. Circuits and Systems (ISCAS), Daegu, Republic of Korea, 22–28 May 2021; pp. 1–4.
14. Reiter, T.; Klemenschits, X.; Filipovic, L. Impact of plasma induced damage on the fabrication of 3D NAND flash memory. *Solid-State-Electron* **2022**, *192*, 108261.
15. Kang, J.-K.; Lee, J.; Yim, Y.; Park, S.; Kim, H.; Cho, E.; Kim, T.; Lee, J.; Kim, J.; Lee, R.; et al. Highly Reliable Cell Characteristics with CSOB(Channel-hole Sidewall ONO Butting) Scheme for 7th Generation 3D-NAND. In Proceedings of the 2021 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 11–16 December 2021; pp. 10.1.1–10.1.4.
16. Chung, Y.-A.; Lung, C.-Y.; Chiu, Y.-C.; Lee, H.-J.; Lian, N.-T.; Yang, T.; Chen, K.-C.; Lu, C.-Y. Study of Plasma Arcing Mechanism in High Aspect Ratio Slit Trench Etching. In Proceedings of the 2019 30th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), Saratoga Springs, NY, USA, 6–9 May 2019; pp. 1–4.
17. Parat, K.; Goda, A. Scaling Trends in NAND Flash. In Proceedings of the 2018 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 1–5 December 2018; pp. 2.1.1–2.1.4.
18. Goda, A.; Parat, K. Scaling directions for 2D and 3D NAND cells. In Proceedings of the IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 10–13 December 2012; pp. 2.1.1–2.1.4.
19. Kim, H.; Ahn, S.-J.; Shin, Y.G.; Lee, K.; Jung, E. Evolution of NAND Flash Memory: From 2D to 3D as a Storage Market Leader. In Proceedings of the 2017 IEEE Int. Memory Workshop (IMW), Monterey, CA, USA, 14–17 May 2017; pp. 1–4.
20. Yanagihara, Y.; Miyaji, K.; Takeuchi, K. Control Gate Length, Spacing and Stacked Layer Number Design for 3D-Stackable NAND Flash Memory. In Proceedings of the 2012 4th IEEE International Memory Workshop (IMW), Milan, Italy, 20–23 May 2012; pp. 1–4.
21. Ko, K.; Lee, J.K.; Shin, H. Variability-Aware Machine Learning Strategy for 3-D NAND Flash Memories. *IEEE Trans. Electron Devices* **2020**, *67*, 1575–1580. [CrossRef]
22. Bhatt, U.M.; Manhas, S.K.; Kumar, A.; Pakala, M.; Yieh, E. Mitigating the Impact of Channel Tapering in Vertical Channel 3-D NAND. *IEEE Trans. Electron Devices* **2020**, *67*, 929–936. [CrossRef]
23. Lee, J.G.; Jung, W.J.; Park, J.H.; Yoo, K.-H.; Kim, T.W. Effect of the Blocking Oxide Layer With Asymmetric Taper Angles in 3-D NAND Flash Memories. *IEEE J. Electron Devices Soc.* **2021**, *9*, 774–777. [CrossRef]
24. Kim, K.T.; An, S.W.; Jung, H.S.; Yoo, K.; Kim, T.W. The Effects of Taper-Angle on the Electrical Characteristics of Vertical NAND Flash Memories. *IEEE Electron Device Lett.* **2017**, *38*, 1375–1378. [CrossRef]
25. Nowakowski, P.; Ray, M.; Fischione, P.; Sagar, J. Top-down delayering by low energy, broad-beam, argon ion milling—A solution for microelectronic device process control and failure analyses. In Proceedings of the 28th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), Saratoga Springs, NY, USA, 15–18 May 2017; pp. 95–101.

26. Han, C.; Wu, Z.; Yang, C.; Xie, L.; Xu, B.; Liu, L.; Yin, Z.; Jin, L.; Huo, Z. Influence of accumulated charges on deep trench etch process in 3D NAND memory. *Semicond. Sci. Technol.* **2020**, *35*, 045003. [CrossRef]
27. Fayrushin, A.; Liu, H.; Mauri, A.; Carnevale, G.; Cho, H.; Mao, D. Numerical Study of Non-Circular Pillar Effect in 3D-NAND Flash Memory Cells. In Proceedings of the 2019 IEEE Workshop on Microelectronics and Electron Devices (WMED), Boise, ID, USA, 26–26 April 2019; pp. 1–4.
28. Hsu, T.-H.; Lue, H.-T.; Lai, E.-K.; Hsieh, J.-Y.; Wang, S.-Y.; Yang, L.-W.; King, Y.-C.; Yang, T.; Chen, K.-C.; Hsieh, K.-Y.; et al. A High-Speed BE-SONOS NAND Flash Utilizing the Field-Enhancement Effect of FinFET. In Proceedings of the 2007 IEEE International Electron Devices Meeting (IEDM), Washington, DC, USA, 10–12 December 2007; pp. 913–916.