Semiconductor Science and Technology

LETTER

Effect of charge trap layer thickness on the charge spreading behavior within a few seconds in 3D charge trap flash memory

Letters

IOP Publishing

Semiconductor Science and Technology

Semicond. Sci. Technol. **33** (2018) 10LT01 (4pp)

https://doi.org/10.1088/1361-6641/aade29

Letter

Effect of charge trap layer thickness on the charge spreading behavior within a few seconds in 3D charge trap flash memory

Bongsik Choi[1], Jiyong Lee[2], Jinsu Yoon[1], Minsu Jeon[1], Yongwoo Lee[1], Jungmin Han[1], Jieun Lee[1], Jinhee Park[1], Yeamin Kim[1], Dong Myong Kim[1], Dae Hwan Kim[1], Sungyong Chung[2], Chan Lim[2] and Sung-Jin Choi[1]

[1] School of Electrical Engineering, Kookmin University, Seoul 02707, Republic of Korea
[2] Research and Development Division, SK Hynix Semiconductor Inc., Icheon, Gyeongki 17336, Republic of Korea

E-mail: sjchoiee@kookmin.ac.kr

Received 23 April 2018, revised 20 August 2018
Accepted for publication 31 August 2018
Published 11 September 2018

Abstract

Charge spreading behavior within a few seconds, referred to as early retention, was comprehensively investigated in 24 word-line stacked tube-type 3D NAND flash memory. We thoroughly explored the charge spreading behavior from the perspectives of both electron and hole spreading in 3D NAND flash memory with different charge trap layer thicknesses at various programming and erasing levels for solid and checkerboard patterns to provide guidelines for minimizing and optimizing the charge spreading.

Keywords: 3D NAND, charge trap memory, charge loss, retention, SONOS, charge spreading

(Some figures may appear in colour only in the online journal)

1. Introduction

Charge trap flash (CTF) memory has been widely investigated as a possible replacement for floating-gate memory because it provides several advantages, including simpler process steps, superior vertical scalability, and reduced cell-to-cell interferences [1–5]. Recently, CTF memory has attracted more interest in 3D NAND flash architecture, where a higher bit density has been achieved by stacking multiple memory layers in the vertical dimension [6–8]. In addition, 3D NAND flash memory based on the CTF uses minimal processing steps to pattern the multilayer stacks only once, thus greatly reducing the bit cost [7]. However, it has been reported that the lateral spreading of trapped charges, i.e., electrons and holes, through the 'shared' charge trap layer (SiN) in 3D NAND flash memory produces a significant effect on charge retention [9–11]. In our previous study, it was observed that lateral charge spreading can occur, especially within a very short timescale, i.e., a few seconds, which we referred to as 'early retention' [12]. This phenomenon causes unwanted results, i.e., a negative shift of the threshold voltage (V_T) and an increase in the string current immediately after programming. Therefore, for higher density memory applications, a further in-depth investigation is required regarding early retention behavior, according to important physical parameters in 3D NAND flash memory, such as the SiN layer thickness, because it can directly affect the bit density in vertically stacked 3D NAND flash memory.

In this work, we evaluated the significant influences of the thickness of the SiN layer on the early retention behaviors in 24 word-line (WL) stacked tube-type 3D NAND flash memory to provide guidelines for minimizing and optimizing the charge spreading. Early retention can be caused by the lateral charge spreading of the electrons and holes trapped during the programming (P) and erasing (E) operations, respectively; thus, diverse bias conditions were employed for

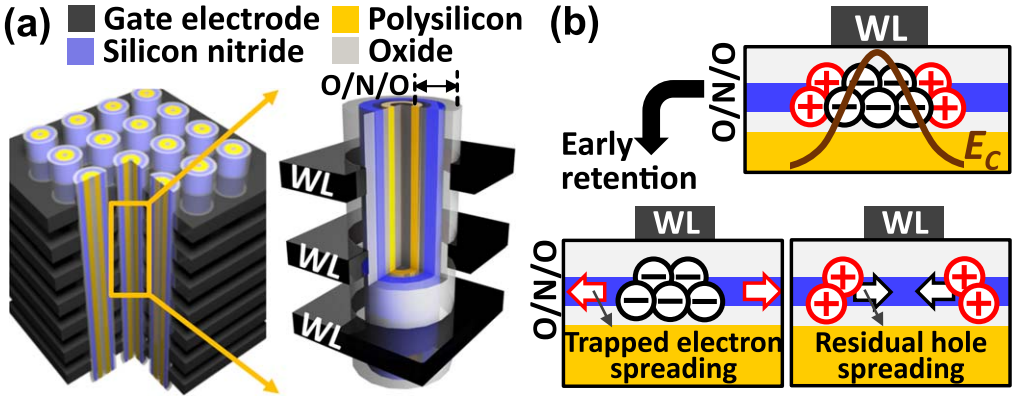

Figure 1. (a) Schematic of the 24 WL stacked tube-type 3D NAND flash memory. (b) Schematic explaining the early retention caused by the spreading of trapped electrons and residual holes in the SiN layer. The profile of the conduction band edge (E_C) in the SiN layer is also shown.

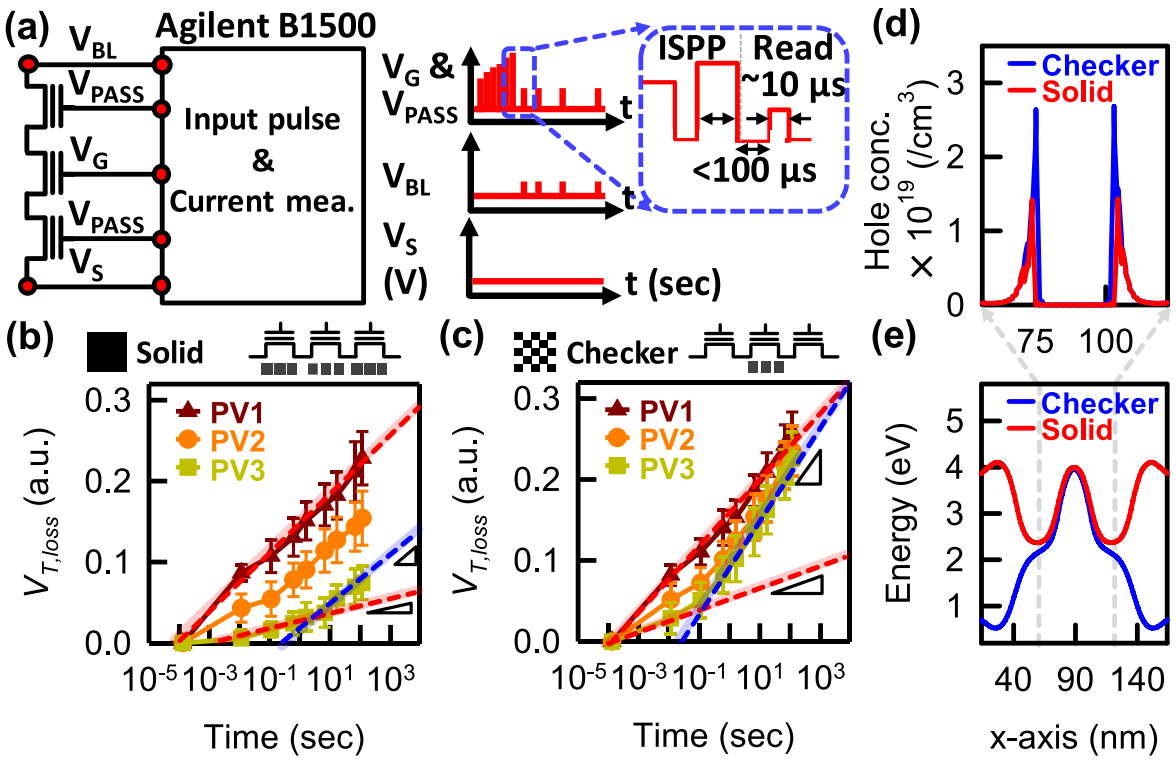

Figure 2. (a) Pulse I-V measurement setup using Agilent B1500 with the waveform generator/fast measurement unit (WGFMU) module. The waveforms designed for the measurement of early retention consist of the ISPP and read pulses. The measured string read current was transformed into the early retention defined as $V_{T,loss}(t) = V_T(t) - V_T(t = 0)$. $V_{T,loss}$ vs. retention time (t) as a parameter of the PV levels at a high EV level for (b) S/P and (c) C/P. Comparisons of (d) simulated residual hole concentration and (e) simulated E_C profiles in SiN layer for S/P and C/P. The number of residual holes for S/P was smaller than for C/P because the holes trapped during the erasing operation were more effectively compensated for by the higher fringing field for S/P compared to that for C/P. The E_C values for C/P have a larger gradient than those for S/P because of the erased neighbor cell states and the larger number of residual holes for C/P.

the comprehensive evaluation of early retention. Moreover, we evaluated the early retention behavior for different patterns, such as solid and checkerboard patterns (S/P and C/P, respectively). It was found that the early retention was caused by a combination of residual hole spreading for short time scales and trapped electron spreading for long time scales. Importantly, we also observed that a thinner SiN layer worsens the early retention under all conditions.

2. Measurement results and discussion

The device structure of the 24 WL stacked tube-type 3D NAND flash memory used in this work, which is referred to as 'SMArT (Stacked Memory Array Transistor)' [6], is shown in figure 1(a). In this scheme, the stack height was minimized by inserting an oxide-nitride-oxide (ONO) layer into the plug. In addition, the gate-last process in the SMArT scheme

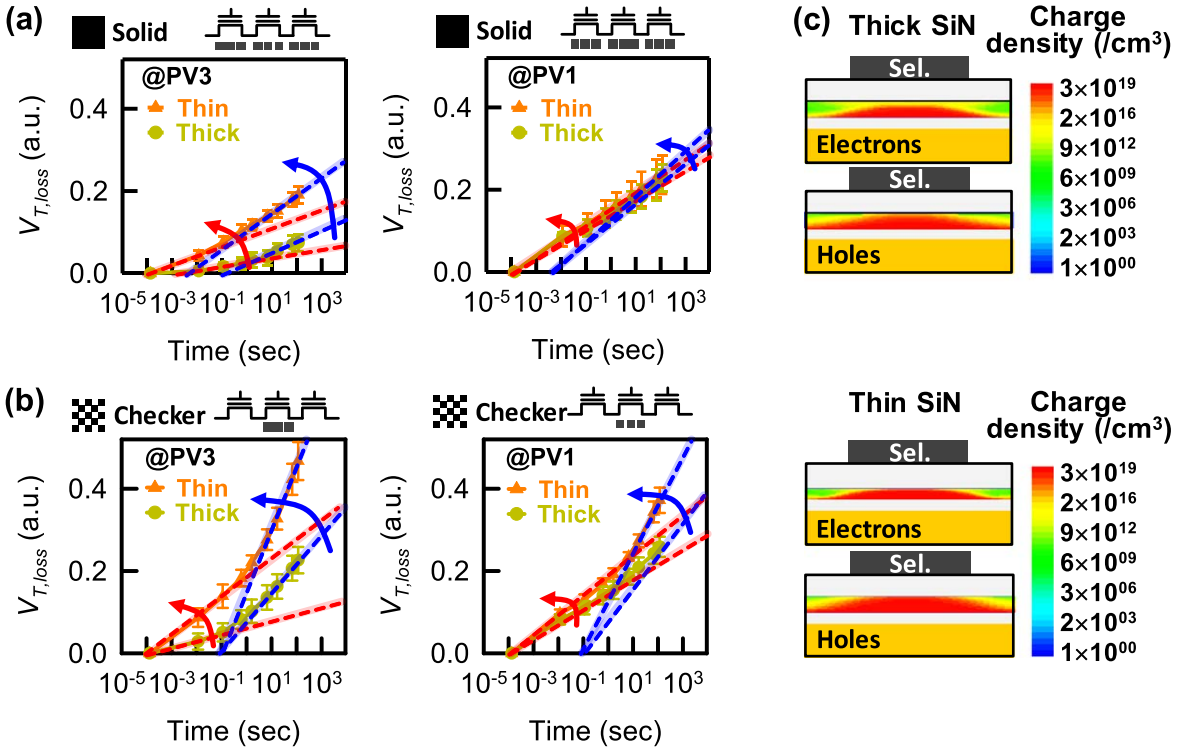

Figure 3. The SiN thickness dependence of early retention at a high PV level (PV3) and low PV level (PV1) for (a) S/P and (b) C/P. The red dotted line and blue dotted line indicate the hole spreading and electron spreading, respectively. More hole spreading (red arrow) and more electron spreading (blue arrow) occur for a thinner SiN layer for both patterns. (c) Simulated contours of the trapped electrons and holes during the P/E operations for thick and thin SiN layers. The difference in thickness between thick and thin SiN layers was 0.4 nm.

exhibited good reliability characteristics. Nevertheless, the early retention behavior was unavoidable and still needs to be improved for the implementation of ultrahigh density memory. As schematically depicted in figure 1(b), lateral spreading of the electrons trapped during programming proceeded outside the cells because of the electric field serving as the driving force of redistribution, which was accelerated at higher temperatures. Moreover, lateral spreading of the residual holes at the edges of cells that were not fully compensated by the electrons trapped during programming, because of the relatively long erasing time and high diffusivity of holes [13], proceeded into the cells from the electric field distribution and then were also redistributed [12]. Therefore, an in-depth study of the early retention is essential in terms of both electron and hole spreading. We only focused on the lateral spreading of the trapped electrons and residual holes in this work, because the SiN thickness is thin enough that the vertical spreading of the charges is expected to be negligible. However, the separation and extraction of two spreading components is important; hence, the quantitative analysis with a sophisticated simulation should be required in near future.

To accurately measure the early retention, we utilized the fast-response pulse I-V measurement using incremental step pulse programming (ISPP), as shown in figure 2(a), because it was difficult to detect early retention using the conventional DC I-V system [14]. The reading phase was enabled within 100 μs after ISPP, which allowed for accurately measuring the early retention. First, all cells were erased by a high erase verify (EV) level; then, the early retention under various program verify (PV) levels (PV1, PV2, and PV3) using ISPP was measured for S/P and C/P at 90 °C. Here, V_T is increased from PV1 to PV3. Note that, as was globally observed, the early retention was clearly aggravated when the PV levels for both S/P and C/P were decreased, as shown in figures 2(b) and (c). This trend was more noticeable for the shorter retention time. This is because a larger number of uncompensated residual holes were generated at a lower PV level (PV1) for both patterns [12]. Additionally, it should be noted that for a higher PV level (PV3), the time dependency of early retention for both S/P and C/P clearly differs from the lower PV level (PV1). That is, the early retention at the PV3 level for both patterns showed a transition in the slope of $V_{T,loss}$ vs. retention time that was not shown at the PV1 level. This was most likely caused not only by hole spreading but also by electron spreading at the high PV3 level. The diffusivity of the holes is surely larger than that of the electrons in the SiN layer; thus, for short time scales, the hole spreading (red dotted line) was dominantly observed [13]. However, for long time scales at the high PV3 level, significant electron spreading also occurred due to the large number of trapped electrons; hence, the slope was changed accordingly (blue dotted line). Since the amount of residual holes was not large at the PV3 level, the slope of the electron spreading was observed to be larger than the slope of the hole spreading. In addition, we observed that for C/P, the charge spreading of both electrons and holes was more evident, especially at PV3,

than for S/P. This is attributed to the larger potential gradient and larger amount of residual holes due to erased neighbor cells for C/P, compared to that for S/P, as shown in figures 2(d) and (e) [10–12].

Based on the aforementioned results of the early retention, we further investigated the early retention behavior for different SiN thicknesses to provide guidelines for minimizing and optimizing the early retention (figure 3). The P/E characteristics were almost similar regardless of the different SiN thicknesses (data are not shown), because the difference in thickness between thick and thin SiN layers was only 0.4 nm. But, we observed that the cells with thinner SiN thickness more rapidly reached to the program saturation at a lower V_T level, compared to those with thicker SiN thickness, which is due to the thickness difference. Notably, the early retention between them was clearly different: in particular, the difference can be clearly observed at a high PV level (PV3) because at this level, both electron and hole spreading were obviously distinguished. A thinner SiN thickness resulted in worse early retention of several tens of mV in both S/P and C/P, as shown in figures 3(a) and (b), respectively. It was observed that the two slopes indicating the charge spreading of both trapped electrons and residual holes increased as the SiN thickness decreased. These results would simply be caused by a wider distribution of trapped electrons and holes during programming and erasing operations in the thinner SiN layer, respectively, as simulated in figure 3(c). The thinner SiN layer did not have enough charge trap sites in the vertical direction compared to the thicker SiN layer [15]; hence, more lateral trapping of both electrons and holes was likely to occur during the P/E operations.

3. Conclusions

We comprehensively investigated the early retention, i.e., fast charge loss within very short time scales, in 3D NAND flash memory. It was found that the early retention was closely correlated with the combination of residual holes at short times scales and trapped electrons at long time scales. Moreover, the early retention behavior was shown to be strongly dependent on the program and erase levels; hence, they should be balanced and optimized to reduce the early retention. The early retention became worse when the thickness of the SiN layer was decreased because of the widened distribution of trapped charges, which indicates that the thickness of the SiN layer is also an important factor to consider to reduce the early retention. We believe that our work provides the optimal design guideline for minimizing and optimizing early retention in 3D NAND flash memory.

Acknowledgments

This work was supported by the SK Hynix Incorporation and the National Research Foundation of Korean Government (Nos. 2016R1A2B4011366 and 2016R1A5A1012966) and partially supported by the Future Semiconductor Device Technology Development Program (Grant 10067739) funded by MOTIE (Ministry of Trade, Industry & Energy) and KSRC (Korea Semiconductor Research Consortium).

ORCID iDs

Sung-Jin Choi https://orcid.org/0000-0003-1301-2847

References

[1] Cho M K and Kim D M 2002 High performance SONOS memory cells free of drain turn-on and over-erase: compatibility issue with current flash technology *IEEE Electron Device Lett.* **21** 399–401
[2] Kim K 2005 Technology for sub-50nm DRAM and NAND flash manufacturing *IEDM Tech. Dig.* pp 323–6
[3] Kim K and Jeong G 2007 Memory technologies for sub-40 nm node *IEDM Tech. Dig.* pp 27–30
[4] Lee C-H, Choi J, Park Y, Kang C, Choi B-I, Kim H, Oh H and Lee W-S 2008 Highly scalable NAND flash memory with robust immunity to program disturbance using symmetric inversion-type source and drain structure *Proc. VLSI Symp. Technol.* pp 118–9
[5] Sim J S *et al* 2007 Self aligned trap-shallow trench isolation scheme for the reliability of TANOS (TaN/AlO/SiN/Oxide/Si) NAND flash memory *Proc. Non-Volatile Semicond. Memory Workshop* pp 110–1
[6] Choi E-S and Park S-K 2012 Device considerations for high density and highly reliable 3D NAND flash cell in near future *IEDM Tech. Dig.* pp 211–4
[7] Chen S-H *et al* 2012 A highly scalable 8-layer vertical gate 3D NAND with split-page bit line layout and efficient binary-sum MiLC (Minimal Incremental Layer Cost) staircase contacts *IEDM Tech. Dig.* pp 21–4
[8] Tanaka H *et al* 2007 Bit cost scalable technology with punch and plug process for ultra high density flash memory *Proc. VLSI Symp. Technol.* pp 14–5
[9] Whang S *et al* 2010 Novel 3-dimensional dual control-gate with surrounding floating-gate (DC-SF) NAND flash cell for 1Tb file storage application *IEDM Tech. Dig.* pp 668–71
[10] Kang C, Choi J, Sim J, Lee C, Shin Y, Park J, Sel J, Jeon S, Park Y and Kim K 2007 Effects of lateral charge spreading on the reliability of TANOS (TaN/AlO/SiN/Oxide/Si) NAND flash memory *Proc. IEEE 45th Annu. Int. Reliab. Phys. Symp.* pp 167–70
[11] Oh D, Lee B, Kwon E, Kim S, Cho G, Park S, Lee S and Hong S 2015 TCAD simulation of data retention characteristics of charge trap device for 3D NAND flash memory *Proc. Int. Memory Workshop* pp 1–4
[12] Choi B *et al* 2016 Comprehensive evaluation of early retention (fast charge loss within a few seconds) characteristics in tube-type 3D NAND flash memory *Proc. VLSI Symp. Technol.* pp 14–6
[13] Baik S J, Lim K S, Choi W, Yoo H, Lee J-S and Shin H 2011 Lateral redistribution of trapped charges in nitride/oxide/Si (NOS) investigated by electrostatic force microscopy *Nanoscale* **3** 2560–5
[14] Chen C-P, Lue H-T, Hsieh C-C, Chang K-P, Hsieh K-Y and Lu C-Y 2010 Study of fast initial charge loss and it's impact on the programmed states Vt distribution of charge-trapping NAND flash *IEDM Tech. Dig.* pp 5–6
[15] You H-W and Cho W-J 2010 Charge trapping properties of the HfO_2 layer with various thicknesses for charge trap flash memory applications *Appl. Phys. Lett.* **96** 093506