Charge Loss Induced by Defects of Transition Layer in Charge-Trap 3D NAND Flash Memory

FEI WANG, (Member, IEEE), YUAN LI, (Member, IEEE), XIAOLEI MA, AND JIEZHI CHEN, (Senior Member, IEEE)

School of Information Science and Engineering (ISE), Shandong University, Qingdao 266237, China

Corresponding authors: Yuan Li (yuan.li@sdu.edu.cn) and Jiezhi Chen (chen.jiezhi@sdu.edu.cn)

This work was supported in part by the China Key Research and Development Program under Grant 2016YFA0201802; in part by the National Natural Science Foundation of China under Grant 62034006, Grant 61874068, and Grant 91964105; and in part by the Fundamental Research Funds of Shandong University under Grant 2018TB013.

ABSTRACT In charge-trap (CT) three-dimensional (3D) NAND flash memory, the transition layer between Si3N4 CT layer and SiO2 tunneling layer is inevitable, and the defects in the transition layer are expected to cause both lateral and vertical charge loss. Here, by first-principles calculations, we present a detailed study on the defects in the transition layer Si2N2O to comprehend their impacts on charge loss in CT 3D NAND flash memory. It is shown that shallow-trap centers, such as intrinsic nitrogen vacancy (V_N) and interstitial Ti (Ti_i), can couple with the conduction band of Si2N2O to lead to lateral charge loss. On the other hand, the N substituting Si atom (N_Si) and Ti substituting Si atom (Ti_Si) defects in the transition layer can couple through resonance with the trap centers in Si3N4, leading to vertical charge loss from the CT layer to the transition layer. Our results strongly suggest that appropriate treatment of the transition layer and hydrogen passivation are both important for avoiding charge loss in CT 3D NAND flash memory.

INDEX TERMS Silicon oxynitride (Si2N2O), Ti doping, charge loss, 3D NAND.

I. INTRODUCTION

Ultra-high density storages are urgently needed for data storage and management at the age of big data. Three-dimensional (3D) NAND flash memory represents a mainstream non-volatile memory aiming to break the inherent restrictions of 2D planar NAND in terms of storage density [1], [2]. While the transition from 2D to 3D structures has achieved a huge reduction in device size and bit cost, scale size is still highly required for further development of 3D NAND flash memory. For instance, 3D charge-trap (CT) NAND flash memory, which usually adopts a macaroni-like vertical channel structure, has been rapidly developed because of its ultra-high storage density, low cost, well reliability, and diversity in novel applications [3]–[8]. The cylindrical shape of 3D CT NAND with ONO stack (SiO2/Si3N4/SiO2) from inside out is beneficial to the field distribution in the CT layer (Si3N4), which allows thicker tunneling oxide (SiO2) to enlarge the memory window [9]. However, this well-established structure also brings issues that have yet to be solved. A typical example is the reliability issue generated as a result of the common CT layer in the same NAND string. In particular, lateral charge migration between adjacent cells can lead to read disturb (RD) and data retention (DR), [10], [11] which issues could be much more serious with size scaling. As such, several approaches, such as metallic doping [12] and process optimization [13], were proposed to suppress charge migration in 3D CT NAND flash memory. As one of the most frequently used methods, process optimization could greatly improve device reliability by reducing defects at the interface to guarantee the purity of Si3N4. However, previous experimental studies revealed that, at the interface between SiO2 and Si3N4, it is very likely to form silicon oxynitride of which the impact is poorly understood [14], [15].

It is found that silicon oxynitride represents a functional material with good mechanical stability, tunable refractive index, and cracking dielectric properties [16], [17], which are promising for applications in microelectronic devices [18]–[21]. As an essential procedure for the fabrication of 3D NAND, N2-annealing is expected to introduce a thin layer of Si2N2O on the surface of Si3N4 [22], [23]. This issue is also expected to be more serious with size scaling. As a unique compound in the SiO2-Si3N4...
quasi-binary system, the intrinsic \( \text{Si}_2\text{N}_2\text{O} \) has long been studied intensively. However, little attention was paid to the impacts of \( \text{Si}_2\text{N}_2\text{O} \) on the charge loss in 3D NAND, which essentially depends on the ability of trapping charges rather than the thickness of the CT layer itself. Therefore, it is important to reveal how \( \text{Si}_2\text{N}_2\text{O} \) as a transition layer impacts the reliability of a CT 3D NAND flash memory, which is instructive for size scaling of the devices.

In this work, by first-principles calculations, we present a detailed study on various defects in \( \text{Si}_2\text{N}_2\text{O} \) to comprehend their impacts on charge loss in CT 3D NAND flash memory, which is critical to the reliability and size scaling of the devices. As will be shown, our results reveal that appropriate treatment of the transition layer (\( \text{Si}_2\text{N}_2\text{O} \)) and hydrogen passivation are both important for improving the reliability in 3D NAND flash memory.

**II. METHODOLOGY**

We employ crystalline structures of \( \text{Si}_2\text{N}_2\text{O} \) to reveal the internal mechanisms of charge loss by introducing different defects. As shown in Fig. 1(a), a supercell of \( \text{Si}_2\text{N}_2\text{O} \) crystal (with side lengths of 17.7384 Å, 16.4793 Å, and 14.5473 Å; crystal system: Base-Centered Orthorhombic; H-M symbol: Cmcm [24]) containing 360 atoms was employed for calculating accurate trap energy levels. Due to the high symmetry of \( \text{Si}_2\text{N}_2\text{O} \) crystal, all the defect sites can be identified exclusively. Even though in the case of interstitial doping, the big ring of \( \text{Si}_2\text{N}_2\text{O} \) with the lowest energy configurations could be located easily.

![FIGURE 1. (a) Geometric structure and (b) projected density of states (PDOS) of a \( \text{Si}_2\text{N}_2\text{O} \) supercell (side lengths of 17.7384 Å, 16.4793 Å, and 14.5473 Å) with and without spin-orbit coupling, respectively. The inset in (b) displays schematically the location of the transition layer (\( \text{Si}_2\text{N}_2\text{O} \)).](image)

The geometric optimization of the supercell was calculated at the level of density functional theory (DFT) by using the generalized gradient approximation Perdew-Burke-Ernzerhof (GGA-PBE) exchange-correlation functional. A k-points grid of \( 1 \times 1 \times 1 \) was adopted in the calculations. In order to obtain the correct value of 5.97 eV for the band gap of \( \text{Si}_2\text{N}_2\text{O} \) [25], the electronic properties were calculated by using the screened hybrid functional of Heyd, Scuseria, and Ernzerhof (HSE) with Hartree-Fock mixing constant of \( \alpha = 0.14 \) and screen parameter of \( \omega = 0.2 \). The spin-orbital coupling (SOC) was omitted in our calculations because it has little impact on the electronic properties of \( \text{Si}_2\text{N}_2\text{O} \), in particular the band gap (see Fig. 1(b)). Similar calculations were performed for both \( \beta\text{-SiO}_2 \) and \( \beta\text{-Si}_3\text{N}_4 \). All the calculations were performed by using the GPU accelerated PWmat package based on the plane wave and the pseudopotential method [26]. In practice, total energy convergence of less than \( 10^{-4} \) eV, energy cut-off of 50 Ry, and residual force of 0.01 eV/Å on each atom were employed in the calculations.

To study the defect properties, we calculated the formation energy of various defects. The formation energy for a defect \( \alpha \) in the charged state \( q \) is defined by following formula [27]:

\[
\Delta H(\alpha, q) = E(\alpha, q) - E(\text{host}) + \sum n_i(E_i + u_i) + q[E_{\text{VBM}}(\text{host}) + E_F + \Delta V]
\]

Here, \( \alpha \) denotes the defect type; \( q \), the number of electrons transferred from the supercell to the reservoirs forming the defect cell; \( E(\alpha, q) \), the energy of the supercell with the defect; \( E(\text{host}) \), the energy of a perfect supercell without defect; \( E_i \), the atomic energy of element \( i \); \( E_F \), the Fermi level of the host relative to the valence band maximum \( E_{\text{VBM}} \); \( n_i \), the number of atoms for element \( i \); and \( \mu_i \), the chemical potential of element \( i \) added to (\( \mu_i < 0 \)) or removed (\( \mu_i > 0 \)) from the defect-free supercell, which depends on the experimental conditions. It is noted that, different from previous work [28], \( \mu_i \) is considered in our calculations such that more accurate formation energy can be obtained. The electrostatic potential correction term \( \Delta V \) was used to align the reference potential of the supercell with or without defects (\( \Delta V = V(\alpha, q, R) - V(\text{host}, R) \)), and \( R \) denotes a place far away from the defect. To correct for the spurious interaction of the charge with the neutralizing background charge as well as with its mirror images in periodic boundary condition, a charge-correction term was calculated by the equation (2/3)(\( e^2/4\pi\varepsilon_0 \))((\( q^2\alpha_M/2\varepsilon_rL \)) [29], where \( e \) denotes the elementary charge; \( \varepsilon_0 \), the vacuum permittivity; \( \alpha_M \), the Madelung constant (with value of 1.64) [30]; \( \varepsilon_r \), the relative dielectric constant; and \( L = V_{\text{supercell}}^{-1/3} \) [31]–[33]. Given the values of the parameters used in our calculations, a charge correction of only 76 meV is obtained. The impact of such small amount of charge correction, as will be shown, is negligible and will not be considered in our study here.

Charge loss affecting the performance of 3D CT memory mainly takes place by means of electron emission and drift-diffusion. Here, we focus on the emission rate from the trap to the conduction band. The Poole-Frenkel (PF) model was used to obtain the electron emission rate \( R_E(\Delta E_t) \) at temperature \( T \) by the following equation:

\[
R_E(\Delta E_t) = v_{PF} \cdot \exp\left(\frac{\beta \sqrt{F - E_t}}{k_B T}\right), \quad \beta = \sqrt{\frac{e^3}{\pi \varepsilon_0 \varepsilon_r}}
\]

where \( v_{PF} \) denotes the attempt-to-escape frequency for the emission [34]; \( F \), the strength of electric field at the transition layer; \( E_t \), the trap energy with respect to the nitride.
conduction band; and \( \varepsilon_r \), the relative dielectric constant of silicon oxynitride tuned by chemical composition. In practice, \( T = 85 \, ^\circ \text{C}, \nu_{\text{PF}} = 5 \times 10^{12}, \) and \( \varepsilon_r = 6.2 \) [35], [36] were utilized in the calculations.

The SiO\(_2\), Si\(_3\)N\(_2\)O, and Si\(_3\)N\(_4\) were combined to construct the interface layers of SiO\(_2\)/Si\(_3\)N\(_4\) or SiO\(_2\)/Si\(_3\)N\(_2\)O/Si\(_3\)N\(_4\) without interface defects. The dangling bonds at the sidewall of interface were saturated by hydrogen to eliminate the gap defect states. The electrostatic interaction between the interface and its mirror was eliminated in the calculations by employing a vacuum space of 2 nm. The total density of states (DOS) was calculated based on the constructed interface layers. The band alignment was obtained from the partial density of states (PDOS) by projecting the DOS onto the atoms at the interface. Similar methods were employed for estimating interfacial band alignment in previous studies [37], [38]. However, different from other method [39] in which the interfacial materials were calculated separately, our calculations were performed by treating the interface as a whole such that the quantum confinement effect is also included.

### III. RESULTS AND DISCUSSION

We first consider the geometric and band structures of Si\(_3\)N\(_2\)O. The DFT-calculated results have been shown in Fig. 1. The band gap of Si\(_3\)N\(_2\)O obtained in our calculation is about 5.97 eV both with and without spin-orbit coupling. This result agrees with the value range of 5.2 to 5.97 eV reported previously for the band gap of silicon oxynitrides [25].

#### A. DEFECTS-INDUCED LATERAL CHARGE LOSS

Due to the N\(_2\)-annealing processes in 3D NAND flash memory, Si\(_3\)N\(_2\)O could be formed at the Si\(_3\)N\(_4\)/SiO\(_2\) interface, whose defects could also trap charges and thereby cause reliability issues similar to that in the CT layer. Therefore, based on the intrinsic Si\(_3\)N\(_2\)O supercell, we consider the formation energy of different types of defects, such as vacancy, substitution, and interstice. A schematic diagram of the formation energy has been shown in Fig. 2(a), in which the electron trap level (\( E_t \)) is marked as charge loss. Electrons trapped in the shallow-trap centers could lead to reliability issues such as charge loss. Therefore, the shallow-trap centers could be easily identified.

It is found that the defects V\(_N\), interstitial O (O\(_I\)), and O\(_N\) have lower formation energy in the cases of vacancy, substitution, and interstice, labeled by filled dashed ellipses. For estimating interfacial band alignment in previous studies [37], [38], different from other method [39] in which the interfacial materials were calculated separately, our calculations were performed by treating the interface as a whole such that the quantum confinement effect is also included.

![Diagram of formation energy](image)

**FIGURE 2.** (a) Schematic diagram of transition level (vertical dashed line) and electron trap level (solid lines). (b) Geometric structures of typical defects of intrinsic nitrogen vacancy (V\(_N\)), interstitial Si (Si\(_I\)), Si substituting N atom (Si\(_N\)), and O substituting N atom (O\(_N\)). (c)-(e) Calculated formation energy of the defects of vacancy, interstice, and substitution, respectively. In (c)-(e), the electron trap levels are marked by filled circles and a region of \( E_t < 1.4 \) eV is labelled to highlight the energy levels close to or shallower than electron trap level.
instance, three $E_t$ (0.63 eV, 1.98 eV, and 2.67 eV) of Si$_i$ defect could be observed. It is noted that emission from traps can take place in terms of two different mechanisms, i.e., quantum tunneling and thermionic emission. The quantum tunneling can occur at any temperature between all allowed states, such as from trap states to band states or from one trap to another. The thermionic emission causes the release of carriers from traps to band at finite temperature, which can be described by the Poole-Frenkel model. By focusing on the thermionic emission from the traps to the conduction band of Si$_2$N$_2$O, both shallow and deep trap centers (labelled in red and blue, respectively) in Fig. 3(a) were selected to calculate the emission rate by using the Poole-Frenkel equation. The results have been shown in Fig. 3(b), in which the emission rates are distinct for different types of defects. Large emission rates are obtained in the case of shallow-trap centers, such as V$_N$, V$_O$, and Si$_i$. In this situation, charge loss occurs during the early stage of data retention. Meanwhile, low emission rates are obtained in the case of deep-trap centers, such as the defects of O$_N$ ($E_t$ $\sim$ 1.58 eV), Si$_i$ ($E_t$ $\sim$ 1.74 eV), and Si$_O$ ($E_t$ $\sim$ 1.98 eV), which hardly contribute to fast electron diffusion. Thus, according to the emission rates, the shallow- and deep-trap centers could be distinguished in intrinsic Si$_2$N$_2$O.

We turn to consider the impact of metallic doping, such as Ti-doping, which was proved to be efficient in Si$_3$N$_4$ for suppressing the shallow-trap centers in 3D NAND [12]. During this process, Ti-doping could easily take place in Si$_2$N$_2$O. The calculated formation energy of defects in Ti-doped Si$_2$N$_2$O has been shown in Fig. 4(a). The inset displays the geometric structures of the Si defects. The defect Ti substituting Si atom (Ti$_i$) has the lowest formation energy and a shallow-trap energy ($E_t$ $\sim$ 0.46 eV). Because the Si$_2$N$_2$O is Si-rich, there are fewer Si atomic defects such that Ti substituting N atom (Ti$_N$) could represent the dominant defect. Importantly, since Ti$_N$ has deep trap energy larger than 0.88 eV, it will not contribute to shallow traps. In terms of defect interstitial Ti (Ti$_i$), ultra-shallow trap energy ($E_t$ $\sim$ 0.14 eV) is obtained, which can be reduced by doping optimization. As a result, Ti-doping in Si$_2$N$_2$O will result in additional shallow-trap centers. These results suggest that, to achieve robust reliability of 3D NAND flash memory, it is necessary to propose an effective method for deepening the trap energy in both intrinsic and Ti-doped Si$_2$N$_2$O.

We also consider H passivation because it represents an alternative effective measure to eliminate shallow-trap centers [41]. Previous experimental evidence demonstrated that hydrogen annealing could be an effective approach for introducing H-passivation in the devices [42]. The properties of shallow-trap centers, such as V$_N$, V$_O$, Si$_N$, Ti$_i$, and Ti$_S$, with H passivation were calculated, and the results have been shown in Fig. 4(b). The inset displays the geometric structures of these defects with H passivation. It is noted that the trap energy can be obviously deepened by H passivation. For instance, in the case of V$_N$ defect, the H atom will fill the N vacancy, forming V$_N$-H defect, in which the H-passivation can change the defect type by passivating one Si dangling-bond, forming Si-H bond. As a result, intrinsic trap energy of V$_N$ defect is deepened from 0.37 eV to 1.33 eV, which is deep enough to trap electrons. Similarly, by H passivation, the shallow-trap V$_O$ and Si$_i$ defects turn to be V$_O$-H (2.24 eV) and Si$_i$-H (1.41 eV) in intrinsic Si$_2$N$_2$O, which both become deep-trap centers. In the case of defects in Ti-doped Si$_2$N$_2$O, H passivation can deepen the shallow-trap energy by 0.3 eV from Ti$_N$-H and 0.41 eV from Ti$_i$ to Ti$_i$-H defect, respectively. Indeed, these results demonstrate that hydrogen passivation represents an effective method to deepen trap energy for suppressing lateral charge loss in both intrinsic and Ti-doped Si$_2$N$_2$O. Our results also reveal that it is important to optimize the transition layer by hydrogen passivation, which is efficient to avoid the generation of shallow traps for suppressing lateral charge loss.

**B. DEFECTS-INDUCED VERTICAL CHARGE LOSS**

In addition to lateral charge loss, vertical charge loss, which takes place between different layers, also represents an
important factor responsible for the reliability of CT NAND flash memory [43]. The defects in the transition layer Si$_{3}$N$_{4}$O are expected to serve as links connecting different layers of the device, leading to vertical charge loss. However, the current understanding of such effects still remains very limited.

We now turn to consider the vertical charge loss induced by defects in Si$_{3}$N$_{4}$, which can occur theoretically between all allowed resonant states in different layers. Due to the stacked structure of SiO$_{2}$/Si$_{3}$N$_{4}$/SiO$_{2}$, the vertical charge loss could take place between layers. The formation energy of V$_{0}$ and V$_{O}$-H, which are typical defects in the tunneling layer (SiO$_{2}$), were calculated and the results have been shown in Fig. 5(a). It is noted that defect V$_{0}$ in SiO$_{2}$ serves as a hole trap instead of an electron trap and an extremely deep trap energy level ($E_t \sim 6.5$ eV) is obtained. Moreover, single V$_{0}$ defect in bulk system could be formed more easily than V$_{O}$-H because of its low formation energy. In previous work [44], [45], the virtual interactions between vacancy clusters and their images result in higher V$_{0}$ defect formation energy, which error is eliminated in our calculations here for single V$_{0}$ defect. In contrast, as compared with V$_{0}$, defect V$_{O}$-H has shallower trap energy ($E_t \sim 2$ eV), which is expected to be a key defect for vertical charge loss from Si$_{3}$N$_{4}$ to SiO$_{2}$.

The calculated bandgap of the separate SiO$_{2}$ and Si$_{3}$N$_{4}$ is consistent with that of previous work [46], which is labelled as color filled rectangles in Figs. 5(b). The band alignment extracted by PDOS at the Si$_{3}$N$_{4}$/SiO$_{2}$ interface shows that band bending effect and quantum confinement take place, which impact, however, is ignorable and does not affect the rationality of our results. Because of the consistency between the band diagrams of isolated bulks and those of SiO$_{2}$/Si$_{3}$N$_{4}$ and SiO$_{2}$/Si$_{2}$N$_{2}$O/Si$_{3}$N$_{4}$ heterojunctions, the defect levels computed in isolated bulk systems also make much sense to the heterojunctions.

Based on the band alignment, the trap energy levels were extracted and the results have been displayed in Fig. 5(c). Here, Si$_{3}$N$_{4}$ serves as a CT layer in 3D NAND flash memory, in which the electron trap centers are V$_{N}$-H and Si$_{3}$-H defects [40]. While electrons could be trapped in the shallow-trap centers (V$_{N}$ and V$_{N}$-O) of Si$_{3}$N$_{4}$, these trapped electrons are able to escape from the CT layer due to the PF emission. The escaped electrons could be trapped again by V$_{O}$-H in SiO$_{2}$, leading to vertical charge loss. Meanwhile, charge-transfer from electron trap centers (V$_{N}$-H and Si$_{3}$-H) in Si$_{3}$N$_{4}$ to V$_{O}$-H in SiO$_{2}$ could also take place directly because of resonance. Our results indicate that the V$_{O}$-H defects in SiO$_{2}$ are responsible for charge loss, which is consistent with previous work [47]. From our results, it is very important to minimize the V$_{O}$-H defects in SiO$_{2}$ in order to suppress the vertical charge loss. However, while the transition layer may affect the vertical charge loss, the impacts of Si$_{2}$N$_{2}$O defects on the tunneling from Si$_{3}$N$_{4}$ to SiO$_{2}$ still remains open for further study in CT 3D NAND flash memory.

Due to the stacked structure of SiO$_{2}$/Si$_{3}$N$_{4}$/SiO$_{2}$ and N$_{2}$-annealing processes, Si$_{2}$N$_{2}$O can be easily formed at the interface between the stacked layers. In this context, a schematic diagram including Si$_{3}$N$_{4}$ heterojunctions, the defect levels of SiO$_{2}$, SiO$_{2}$/Si$_{3}$N$_{4}$/SiO$_{2}$ interfaces have been shown in Fig. 6(b). While similar band bending effects are observed, our results still make difference. To understand how the defects impact vertical charge loss, we have summarized the results of trap energy levels of the defects based on the band alignment, as shown in Figs. 6(c) and 6(d) in the cases of intrinsic and Ti-doped Si$_{2}$N$_{2}$O, respectively. As discussed above, the defects with shallow-trap energy in Si$_{2}$N$_{2}$O is critical for lateral charge loss. However, the coupling between the defects in Si$_{3}$N$_{4}$, Si$_{2}$N$_{2}$O, and SiO$_{2}$ plays a more important role in causing vertical charge loss. Thus, we focus on the defects in Si$_{2}$N$_{2}$O that can easily couple with their counterparts in Si$_{3}$N$_{4}$ and SiO$_{2}$.

In the case of intrinsic systems, as shown in Fig. 6(c), there exist multiple electron trap centers for data storage in Si$_{3}$N$_{4}$, such as V$_{N}$-H and Si$_{3}$-H, for which the electron trap level is around 1.4 eV. Other typical defects with similar magnitude of trap energy in SiO$_{2}$, such as V$_{O}$ and V$_{O}$-H, and in Si$_{2}$N$_{2}$O, such as O substituting Si atom (O$_{Si}$), N substituting Si atom (N$_{Si}$), Si, and Si substituting O atom (Si$_{O}$), are also summarized for comparison, as shown in Fig. 6(c). Because these intrinsic defects have energy levels close to each other and part of them has shallow-trap levels, they could induce not only lateral charge loss in Si$_{2}$N$_{2}$O but also vertical charge loss between layers. As an effective measure restraining lateral charge loss, H-passivation could also deepen these...
trapped charge energy levels to break resonance. As a result, reduced vertical charge loss could be achieved. Thus, it is important to reduce defects in intrinsic Si$_2$N$_2$O by H-passivation, which is expected to improve the reliability in 3D NAND flash memory.

In the case of Ti-doped systems, as shown in Fig. 6(d), the intrinsic trap centers in Si$_3$N$_4$ are replaced by defects with Ti doping, such as the 2.12 eV T$_i$ ($E_t \sim 0.88$ eV) and 1.57 eV T$_{Si}$ ($E_t \sim 1.43$ eV) defects serving as electron trap centers [12]. As discussed above, the electrons stored in the T$_i$ and T$_{Si}$ defects in Si$_3$N$_4$ could tunnel to the Ti-related defects in Si$_2$N$_2$O because of the resonance of the trap energy. Obviously, there also exist multiple Ti-related defects in Si$_3$N$_2$O for vertical charge loss, such as T$_i$ (1.6 eV) and T$_{Si}$ (1.14 eV). Importantly, these defects are close in trap energy to the V$_{O}$-H defects in SiO$_2$. As a result, vertical charge loss can take place from Si$_3$N$_4$ to Si$_2$N$_2$O and then to SiO$_2$. Thus, when Ti doping strategy in Si$_3$N$_4$ is adopted, it is necessary to appropriately treat the interface to suppress the transition layer. In particular, the Ti defects in Si$_2$N$_2$O could promote vertical charge loss because all Ti-related defects have trap energy levels close to that in both Si$_3$N$_4$ and SiO$_2$. Similar impacts of H-passivation in Ti-doped Si$_2$N$_2$O on vertical charge loss could be obtained. With Ti-doping in Si$_3$N$_4$, it is particularly important to reduce the transition layer Si$_2$N$_2$O to improve the reliability of 3D NAND flash memory.

It should be noted that the stack layers in 3D NAND flash memory are usually amorphous materials, for which the morphology is too diverse and complicated to model accurately. As such, our modeling here was carried out based on crystalline materials instead of amorphous structures to avoid the morphology issue. While there are differences in the V$_{O}$ and V$_{O}$-H defects between our results with that of the studies in literature because of the materials diversity and methodology difference, the major findings of our study here are adequate and reasonable for the crystalline materials of Si$_3$N$_4$, Si$_2$N$_2$O, and SiO$_2$. It can be expected that, in an amorphous Si$_3$N$_2$O, more localized states near the band edges would be generated, bringing huge difficulties for identifying the defect states of interest. Although the approximation made in our modeling fails to account for the diversity of the amorphous structures, the use of crystalline Si$_2$N$_2$O could reveal the internal mechanism without loss of generality. Moreover, as compared with its 2D counterpart, 3D NAND flash memory has the same ONO stacking such that the interfacial geometry at the atomic scale is expected to be very similar. As such, we believe that the major findings obtained in our study here are also applicable to CT type 2D flash memory.

**IV. CONCLUSION**

In summary, by first-principles calculations, we have systematically studied the properties of defects in Si$_2$N$_2$O, which serves as a transition layer between Si$_3$N$_4$ and SiO$_2$ in charge-trap 3D NAND flash memory. It is found that V$_N$, V$_O$, and Si$_i$ defects in intrinsic Si$_2$N$_2$O and T$_{Si}$ and T$_i$ defects in Ti-doped Si$_2$N$_2$O are shallow-trap centers, which could result in lateral charge loss in the device. On the other hand, because the O$_{Si}$, N$_{Si}$, Si$_i$, and all Ti-related defects in Si$_2$N$_2$O are close in trap energy, they can couple with the defects with electron traps in Si$_3$N$_4$ and significant defects (V$_{O}$=H) in SiO$_2$, leading to vertical charge loss from Si$_3$N$_4$ to SiO$_2$. Hydrogen passivation is found to be an effective way in deepening shallow-trap centers and reducing resonant trap levels for charge loss. It suggests that the transition layer should be appropriately treated to suppress charge loss, especially with Ti doping. These findings are instructive for improving the reliability and performance of 3D NAND flash memory from a perspective at the atomic scale.
REFERENCES

[1] C. Monzio Compagnoni, A. Goda, A. S. Spinelli, P. Feeley, A. L. Lacaita, and A. Visconti, “Reviewing the evolution of the NAND flash technology,” Proc. IEEE, vol. 105, no. 9, pp. 1690–1633, Sep. 2017.

[2] A. Goda and K. Parat, “Scaling directions for 2D and 3D NAND cells,” in IEDM Tech. Dig., Dec. 2012, pp. 2.1–2.4.

[3] Y. J. Song, J. H. Lee, H. C. Shin, K. H. Lee, K. Suh, J. R. Kang, S. S. Pyo, H. T. Jung, S. H. Hwang, G. H. Koh, S. C. Oh, O. Park, J. K. Kim, J. C. Park, J. Kim, K. H. Hwang, G. T. Jeong, K. P. Lee, and E. S. Jung, “Highly functional and reliable 8 Mbit STT-MRAM embedded in 28 nm logic,” in IEDM Tech. Dig., Dec. 2016, pp. 27.2.1–27.2.4.

[4] K. Parat and A. Goda, “Scaling trends in NAND flash,” in IEDM Tech. Dig., Dec. 2018, pp. 2.1–2.1.4.

[5] Y.-H. Liu, C.-M. Jiang, W.-C. Chen, T. Wang, W.-J. Tsai, T.-C. Lu, K.-C. Chen, and C.-Y. Lu, “Electric field induced nitride trapped charge lateral migration in a SONOS flash memory,” IEEE Electron Device Lett., vol. 38, no. 1, pp. 48–51, Jan. 2017.

[6] M. Zhang, F. Wu, X. Chen, Y. Du, W. Liu, Y. Zhao, J. Wan, and C. Xie, “RBER aware multi-sensing for improving read performance of 3D MLC NAND flash memory,” IEEE Access, vol. 6, pp. 61934–61947, 2018.

[7] R. Ma, F. Wu, M. Zhang, Z. Lu, J. Wan, and C. Xie, “RBER-aware lifetime prediction scheme for 3D TLC NAND flash memory,” IEEE Access, vol. 7, pp. 44696–44708, 2019.

[8] S.-T. Lee, D. Kwon, H. Kim, H. Yoo, and J.-H. Lee, “NAND flash based novel synaptic architecture for highly robust and high-density quantized neural networks with binary neuron activation of (1,0),” IEEE Access, vol. 8, pp. 141330–141339, 2020.

[9] A. Arregghini, G. Van Den Bosch, G. S. Kar, and J. Van Houdt, “Ultimate scaling projection of cylindrical 3D SONOS devices,” in Proc. 4th IEEE Int. Memory Workshop, May 2012, pp. 1–4.

[10] K. Mizoguchi, T. Takahashi, S. Arimoto, and K. Takeuchi, “Data-retention characteristics comparison of 2D and 3D TLC NAND flash memories,” in Proc. IEEE Int. Memory Workshop (IMW), May 2017, pp. 1–4.

[11] B. Choi, S. Hyun Jung, J. Yoon, J. Lee, M. Yeon, Y. Lee, J. Han, J. Lee, D. M. Kim, D. H. Kim, C. Lim, S. Park, and S.-J. Choi, “Comprehensive evaluation of earliness for fast charge loss within a few seconds” characteristics in tube-type 3-D NAND flash memory,” in Proc. IEEE Symp. VLSI Technol., Jun. 2016, pp. 1–2.

[12] F. Wang, J. Wu, Y. Li, R. Cao, and J. J. A. P. E. Chen, “Metallic doping: A new strategy for suppressing shallow-trap centers in vertically stacked charge-trapping flash memories,” Appl. Phys. Express, vol. 13, no. 4, pp. 1–4, 2020.

[13] K. Mizoguchi, S. Kotaki, Y. Deguchi, and K. Takeuchi, “Lateral charge migration suppression of 3D-NAND flash by vth nerving for near data computing,” in IEDM Tech. Dig., Dec. 2017, pp. 19.2.1–19.2.4.

[14] Y. Iwase, Y. Horie, Y. Daiko, S. Honda, and Y. J. M. Iwamoto, “Synthesis and oxidation behavior of silicon oxynitride ceramics,” J. Amer. Ceram. Soc., vol. 74, no. 1, pp. 109–114, Jan. 1991.

[15] S. Lin, F. Ye, J. Ding, C. Yang, J. Ma, S. Dong, and Q. Liu, “Effects of pore diameters on phase axis, oxidation resistance, and thermal shock resistance of the porous Si3N4 ceramic,” J. Amer. Ceram. Soc., vol. 90, no. 5, pp. 2190–2198, May 2017.

[16] T. R. Durrant, S. T. Murphy, M. B. Watkins, and A. L. Slushler, “Relation between image charge and potential alignment corrections for charged defects in periodic boundary conditions,” J. Chem. Phys., vol. 149, no. 2, pp. 1–16, 2018.

[17] H.-P. Komsa, T. T. Rantala, and A. Pasquarello, “Finite-size supercell correction schemes for charged defect calculations,” Phys. Rev. B, Condens. Matter, vol. 86, no. 4, pp. 1–16, Jul. 2012.

[18] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, “Fully ab initio finite-size corrections for charged-defect supercell calculations,” Phys. Rev. Lett., vol. 102, no. 1, pp. 1–4, Jan. 2009.

[19] Z. Liu, F. Hayashi, S. Fujieda, M. Wilde, and K. Fukutani, “Reliability driven guideline for BEOL optimization: Protecting MOS stacks from hydrogen-related impurity penetration,” in Proc. IEEE Int. Conf. IC Design Technol., May 2012, pp. 1–4.

[20] Z. Liu, S. Ito, M. Wilde, K. Fukutani, I. Hirozawa, and T. Koganezawa, “A hydrogen storage layer on the surface of silicon nitride films,” Appl. Phys. Lett., vol. 92, no. 19, pp. 1–3, 2008.

[21] P. Kroll and M. Milko, “Theoretical investigation of the solid state reaction of silicon nitride and silicon dioxide forming silicon oxynitride (Si3N4O) under pressure,” Zeitschrift für anorganische und allgemeine Chem., vol. 629, no. 10, pp. 1737–1750, Sep. 2003.

[22] W. Ching and S.-Y. Ren, “Electronic structures of Si3N4O and Ge2N2O crystals,” Phys. Rev. B, Condens. Matter, vol. 24, no. 10, pp. 5788–5795, 1981.

[23] W. Jia, Z. Cao, L. Wang, J. Fu, X. Chi, W. Gao, and L.-W. Wang, “The analysis of a plane wave pseudopotential density functional theory code on a GPU machine,” Comput. Phys. Commun., vol. 184, no. 1, pp. 9–18, Jan. 2013.

[24] C. G. Van de Walle and J. Neugebauer, “First-principles calculations for defects and impurities: Applications to III-nitrides,” J. Appl. Phys., vol. 95, no. 8, pp. 3851–3879, Apr. 2004.

[25] A.-M. El-Sayed, Y. Wimmer, W. Goes, T. Grasser, V. A. Afanas’ev, and A. L. Slushler, “Theoretical models of hydrogen-induced defects in amorphous silicon dioxide,” Phys. Rev. B, Condens. Matter, vol. 92, no. 1, pp. 014107–1–014107–11, Jul. 2015.

[26] W. Liu, M. Wang, L. Yin, J. Shi, Y. Tian, and J. Wang, “The analysis of a plane wave pseudopotential density functional theory code on a GPU machine,” Comput. Phys. Commun., vol. 184, no. 1, pp. 9–18, Jan. 2013.
[43] C. Woo, M. Lee, S. Kim, J. Park, G.-B. Choi, M.-S. Seo, K. H. Noh, M. Kang, and H. Shin, “Modeling of charge loss mechanisms during the short term retention operation in 3-D NAND flash memories,” in Proc. Symp. VLSI Technol., Jun. 2019, pp. T214–T215.

[44] M. S. Munde, D. Z. Gao, and A. L. Shluger, “Diffusion and aggregation of oxygen vacancies in amorphous silica,” J. Phys. Condens. Matter, vol. 29, no. 24, pp. 1–10, 2017.

[45] S. Mukhopadhyay, P. V. Sushko, A. M. Stoneham, and A. L. Shluger, “Correlation between the atomic structure, formation energies, and optical absorption of neutral oxygen vacancies in amorphous silica,” Phys. Rev. B, vol. 71, no. 23, pp. 235204-1–235204-9, Jun. 2005.

[46] K. A. Nasyrov, S. S. Saimiev, V. A. Gritsenko, J. H. Han, C. W. Kim, and J.-W. Lee, “Electron and hole injection in metal-oxide-nitride-oxide-silicon structures,” J. Exp. Theor. Phys., vol. 102, no. 5, pp. 810–820, May 2006.

[47] X. Ma, R. Cao, F. Wang, X. Zhan, and J. Chen, “Charge-assisted recovery and degradation in charge-trapping 3D NAND flash memory, experimental evidences and theoretical perspectives,” in Proc. IEEE Silicon Nanoelectron. Workshop (SNW), Jun. 2020, pp. 41–42.

FEI WANG (Member, IEEE) received the bachelor’s degree in physics from Shandong Normal University, in 2016. She is currently pursuing the Ph.D. degree with the School of Physics and the School of Information Science and Engineering, Shandong University. Her research interests include the reliability of 3D NAND flash memory and nano-device fabricating.

YUAN LI (Member, IEEE) received the Ph.D. degree in condensed matter physics from Shandong University, in 2009. From 2010 to 2016, he was a Postdoctoral Fellow and a Research Scientist with the Georgia Institute of Technology and the King Abdullah University of Science and Technology (KAUST). In 2017, he joined the Faculty of Shandong University, as an Associate Professor with the School of Information Science and Engineering. He has coauthored more than 60 peer-reviewed articles published in international journals, including one invited review paper and one book chapter. His research interests include the computational design and characterization of advanced materials for information applications, in particular the flexible electronic materials, such as organic semiconductors and conjugated polymers. He also serves as an independent Reviewer for the National Natural Science Foundation of China and several international journals.

XIAOLEI MA received the bachelor’s degree in physics from Qilu Normal University, in 2015, and the Ph.D. degree from the School of Information Science and Engineering, Shandong University, in 2020. He currently holds a postdoctoral position with Peking University.

JIEZHI CHEN (Senior Member, IEEE) received the Ph.D. degree from the Department of Informatics and Electronics, The University of Tokyo, in 2009. In 2010, he joined the Research and Development Center, Toshiba Corporation. He is currently a Professor with the School of Information Science and Engineering, Shandong University, China. He has published many articles in journals and conference proceedings. His research interests include the characterization and process engineering of nano-scale transistors and non-volatile memories, with main focus on reliability physics. He acted as a reviewer of several international journals. Since 2008, he has been the leading author or the corresponding author, working on nanoscale devices and device reliabilities, for more than ten times in the VLSI Symposium and the IEEE International Electron Device Meeting (IEEE IEDM). He also served as a Technical Program Committee (TPC) Member for IEDM, in 2016 and 2017. He has been serving as a TPC Member for the Silicon Nanoelectronics Workshop since 2018, the IEEE International Reliability Physics Symposium since 2019, and the IEEE International Memory Workshop since 2020.

* * *