Distinguishing capture cross-section parameter between GIDL erase compact model and TCAD

Kul Lee* and Hyungcheol Shin

Inter-University Semiconductor Research Center (ISRC) and School of Electrical Engineering and Computer Science, Seoul National University, Seoul 151-747, Republic of Korea
*E-mail: yikul@snu.ac.kr

Received June 17, 2021; revised October 8, 2021; accepted October 13, 2021; published online November 30, 2021

A compact model of 3D NAND enables simulation at circuit- or system-level. Although a compact model for gate-induced-drain-leakage (GIDL)-assisted erase was proposed in a previous study, it is difficult to use practically because it has not been properly validated. In particular, a capture cross-section (CCS) value that is far from the real value is used. Furthermore, it does not consider the latest device structure and its operation. In this paper, a conventional GIDL-assisted erase compact model is validated using TCAD and has been improved more practically. It is confirmed that we should distinguish between CCS in TCAD and the compact model due to their differences. Based on their physical differences, an equation that can interconvert them is proposed and the model is successfully validated with proper CCS. Finally, the advanced GIDL-assisted erase compact model considering a tapered angle, single-side injection and word-line voltage is suggested.

© 2021 The Author(s). Published on behalf of The Japan Society of Applied Physics by IOP Publishing Ltd

1. Introduction

3D NAND flash memory has been developed with increasing bit storage density.1–5 CMOS-under-array (CUA) structure has contributed to its further increase. CUA structure enables minimum cell area and die size by placing a CMOS logic circuit under a NAND array.6–8 In this structure, gate-induced-drain-leakage (GIDL)-assisted erase is used instead of bulk erase.7–9 Therefore, predicting GIDL erase characteristics becomes more important in the latest 3D NAND.

Compact models for 3D NAND have been constantly developed because they can offer system- or circuit-level analysis, which cannot be rapidly accomplished by technology computer-aided design (TCAD).10–14 These compact models are generally validated using TCAD or experimental data.10,12,13 In particular, in the early stages of product development without any experimental data, TCAD is used to validate a compact model and to extract the required model parameters.

Figure 1(a) shows a conventional GIDL-assisted erase compact model (in short, GIDL erase compact model).12 However, the conventional model in Fig. 1(a) is difficult to use practically for two reasons. First, the model is not validated or calibrated against experimental data or TCAD. Second, it does not consider the latest device issues and its operation (tapered structure, double stacking and CUA structure).

In this study, to make the model more practical, we validate the conventional GIDL erase compact model using TCAD and improve it to include issues and operations of the latest device. In Sect. 2, model validation is attempted through TCAD before improving the conventional model. Through the validation process, it is confirmed that the compact model and TCAD results show severe disagreement in spite of the same tunneling current and capture cross-section (CCS). The cause of the disagreement and the way to obtain a proper CCS for the compact model are discussed using simple silicon-oxide-nitride-oxide-silicon (SONOS) structure. In Sect. 3, an advanced compact model considering a tapered angle, single-side injection and word-line (WL) voltage is suggested.

2. Validation of the compact model

2.1. Description of a conventional GIDL erase model

The conventional model in Fig. 1(a) is modeling erase operation using combination of a GIDL compact model (GIDL circuit)13 and a semi-analytical erase model (cell circuit).12 It models half of a symmetric 3D NAND string structure with $N_{WL}$ WLs. Therefore, it includes one bitline (BL), one select gate (SG) stack and $2N_{WL}$ WLs. $I_{GIDL}$ represents the band-to-band tunneling current between the BL and SG. $I_h$, $I_r$, and $I_t$ represent the hole tunneling current, hole recombination current and hole trapping current in a unit cell, respectively. $V_{n+}$ represents the built-in potential. Figure 2(b) shows capacitances used in the model on a cross-section of the string near the SG. The potential of the hole accumulated region (orange region) is considered to be constant as $V_B$ regardless of its position. $V_N$ represents the potential at the $n^+-n$ junction (boundary between the BL and channel). The hole accumulated region is coupled by $C_{NL}$ to the depletion layer of $n^+$ region. $L_x$ is the length of the depleted region under the SG. $C_{dep}$ represents the capacitance of the depletion layer of the $n^+$ region. $C_{out}$, $C_{fin}$ and $C_{fB}$ are fringing capacitances. $C_{ono}$ is the capacitance per unit length of the O/N/O layer. $C_{G1}$ represents capacitance between the channel and middle of the nitride layer in a unit cell, and $C_{G2}$ represents the capacitance between the WL and the middle of the nitride layer in a unit cell. Calculations for $I_{GIDL}$ and the capacitances are introduced in Ref. 13.

Figure 2 schematically shows the semi-analytical erase model used in Fig. 1(a). Holes are injected into nitride by Fowler–Nordheim (FN) tunneling ($I_h$). Then, they are recombined ($I_{r}$) with electrons or trapped ($I_{t}$). Electrons are emitted by Poole–Frenkel (PF) emission ($I_e$). The following differential equations represent this mechanism:12,15

$$\frac{d}{dt} [\text{electron density}] = I_e - I_r - I_t$$

$$\frac{d}{dt} [\text{hole density}] = I_h - I_r - I_t.$$
\[ \frac{dn_{t,e}}{dt} = -e_n n_{t,e} - \frac{J_h \sigma_e n_{t,e}}{q}, \quad (1) \]

\[ \frac{dn_{t,h}}{dt} = J_h \sigma_h (n_{\text{total,h}} - n_{t,h}) \quad (2) \]

\( n_{t,e}, n_{t,h} \) and \( n_{\text{total,h}} \) represent concentrations of trapped electrons, trapped holes and total hole traps, respectively.

\( e_n \) is the emission rate calculated using a PF equation. \( J_h \) is hole FN tunneling current density. \( \sigma_e \) and \( \sigma_h \) are hole CCSs for trapped electrons and empty hole traps. The first term and the second term in the right-hand side of Eq. (1) are related to \( I_e \) and \( I_{h,t} \), respectively. The term in the right-hand side of Eq. (2) is related to \( I_{h,t} \). Theoretically calculated parameters based on tunneling barrier and tunneling mass, \( A = 3.85 \times 10^{-7} \) A V\(^{-2}\) and \( B = 386 \) MV cm\(^{-1}\), are used for hole FN.
tunneling in the compact model.\textsuperscript{17,18} The electron trap is set at 1.5 eV below the conduction band, and its concentration is $5 \times 10^{19}$ cm$^{-3}$.

2.2. TCAD simulation setup for validation

Figure 3 shows the assumed simple SONOS structures to observe the capture mechanism in TCAD. Heavily p-doped silicon is commonly used for the gate, drain and source. This is done to make analysis simple by eliminating the initial potential differences between them. Monocrystalline channel is assumed. Shockley–Read–Hall recombination and a nonlocal band-to-band tunneling model is assumed in the silicon channel. The nonlocal tunneling model in TCAD is used to simulate charge exchange at the nitride layer. The electron tunneling from the gate is ignored, as assumed in the semi-analytical erase model.\textsuperscript{12} The tunneling mass of the hole in oxide ($m_{h,ox}^*$) is assumed to be $0.5m_0$ ($m_0$ is free electron mass) and the tunneling barrier is assumed to be 4 eV.\textsuperscript{19,20} The energy level of the acceptor trap and donor trap are set at 1.5 and 2.5 eV below the conduction band of nitride, and their concentrations are $5 \times 10^{19}$ cm$^{-3}$ and $3 \times 10^{19}$ cm$^{-3}$, respectively.\textsuperscript{12,21} Constant CCS values are assumed for holes and electrons in TCAD ($p_0 = 0$ in J-model).\textsuperscript{22}

2.3. Different definition of CCS between TCAD and the compact model

As shown in Fig. 4, the nonlocal tunneling model in TCAD and the FN tunneling model in the compact model show similar tunneling current for the same electric field. However, the results show that there are significant differences between the compact model and TCAD, as shown in Fig. 5.
compact model overestimates both drain current and $\Delta V_T$ ($\Delta V_T$ is calculated in the same way as in Ref. 13.)

Disagreement is due to the different definition of CCS. Figure 6 compares capture mechanisms of TCAD and the compact model. In TCAD, CCS is defined for holes that have already been injected and remained on the valance band of nitride, whereas the compact model defines it for holes that are tunneling into nitride. Therefore, CCS in TCAD and compact models should be distinguished. From this section onwards, CCS in TCAD will be expressed as $\sigma_T$, and CCS in the compact model will be expressed as $\sigma_C$. Since $\sigma_h$ and $\sigma_r$ in Sect. 2.2 are defined in the compact model, they should be denoted as $\sigma_{C,h}$ and $\sigma_{C,r}$. The corresponding values in TCAD are denoted as $\sigma_{T,h}$ and $\sigma_{T,r}$.

$\sigma_T$ and $\sigma_C$ can be interconverted considering their physical meaning. In TCAD, the capture rate is calculated as follows:

$$c = \sigma_T v_{th}^p, \tag{3}$$

where $\sigma_T$ is the CCS for TCAD, $v_{th}^p$ is thermal velocity of holes ($1.56 \times 10^7$ cm s$^{-1}$ is used in this study), and $p$ is the concentration of holes on the nitride valence band. As shown in Eq. (3), $\sigma_T$ is defined for holes in the valance band of nitride.

On the other hand, considering Eqs. (1) and (2), the capture rate can be calculated as follows:

$$c = \frac{J_h \sigma_C}{q}, \tag{4}$$

where $\sigma_C$ is the CCS for the compact model. $\sigma_C$ is being defined for tunneling current density in the compact model.

Therefore, properly converted CCS should be used to compare the results of TCAD and the compact model. If the same capture rate is assumed considering the same structure, we can obtain their relation as follows (i.e. supposing Eqs. (3) and (4) are equal):

$$Conversion ratio (CR) = \frac{\sigma_C}{\sigma_T} = k \frac{q v_{th}^p p}{J_h} \tag{5}$$

$k$ is the fitting parameter. Considering that $q$ and $v_{th}^p$ are constants, the CR depends on the ratio between the hole concentration in the valance band and the tunneling hole concentration.
current density. Therefore, $\sigma_{C,h}$ and $\sigma_{C,r}$ in the compact model should be set $CR$ times larger than $\sigma_{T,h}$ and $\sigma_{T,r}$ in TCAD.

### 2.4. Validation in simple SONOS structure

Figure 7 shows the result of Eq. (5) using extracted $p$ and $J_h$ from TCAD for two cases. This shows that $\sigma_{C}$ is approximately $2.1 \times 10^5$ times larger than $\sigma_{T}$ in the given condition. Figure 8 shows the results of TCAD and the compact model with a converted CCS, as calculated in Fig. 7. The same CCS value is used for both recombining and hole trapping. With converted CCS, compact model results show good agreement with TCAD results for $I_D$ and $\Delta V_T$. However, when the CCS exceeds a critical value, erase characteristics are not affected by the CCS anymore, neither in TCAD nor in the compact model.

### 2.5. Validation in string structure including GIDL

Figure 9 shows the cross-section of half of the assumed 3D NAND string structure for GIDL erase. Ten WLs, two selectlines (DSL and SSL), heavily n-doped BL and source-line (SL) are assumed for the full string. The WL length is reduced to 50 nm, and a lightly n-doped channel with macaroni structure is used. The other physical conditions are the same as in the previous simple SONOS structure.

Although the actual 3D NAND channel is made of a polycrystalline silicon channel, a monocristalline silicon channel is assumed in the conventional model and in this paper. This is not only for simplicity of the compact model, but for consideration of the significantly reduced number of polysilicon grains along the string.\(^2\) \(^2\) \(^2\) \(^2\) \(^2\) \(^2\)

Figure 10 compares the results of TCAD and those of the GIDL erase compact model (Fig. 1). It is confirmed that the transient results of $I_{BL}$ and $\Delta V_T$ of the compact model are in good agreement with the results of TCAD. Unlike in a simple SONOS structure, two current peaks appear, as shown in Fig. 10(a). This is because GIDL current occurs at DSL before FN tunneling current occurs at WL.\(^1\)\(^2\)

### 3. Advanced erase compact model

The compact model validated in Sect. 2 assumed a constant channel radius (no tapered angle). In addition, it assumed symmetric structure and the same GIDL injection current from DSL and SSL. However, these conditions are far from the actual 3D NAND string. Figure 11 schematically shows a...
more realistic 3D NAND string structure. A string consists of two stacks, an upper stack and a lower stack. It has a tapered angle due to the limitation of the etch process. Unlike the model in Fig. 1(a), single-side injection (GIDL injection only from DSL) is assumed in this structure. This is the result of considering the recent operation of 3D NAND with CUA structure.

The tapered structure causes variations in the dimension parameters of cells in the same string. Each cell should be modeled to have different $C_{G1}, C_{G2}, C_{FB}, I_c, L_e$ and $I_h$. Therefore, to model the string that has $N_{WL}$ WLs, tapered string structure and single-side injection voltage scheme, one GIDL circuit and $N_{WL}$ cell circuits are required, as shown in Fig. 12. However, this model causes high computing cost because differential equations [Eqs. (1) and (2)] should be solved $N_{WL}$ times to simulate $N_{WL}$-layer structure. Therefore, we propose the model in Fig. 13 to increase the simulation efficiency by reducing the number of differential equations to be solved.

Figure 13 shows the advanced compact model considering a tapered angle, single-side injection and WL voltage. The advanced model assumes that the cells, except for the top and bottom cells in each stack, have the same size as the cell located at the center of the stack (mid cell). Therefore, the model consists of a total of four circuits. Each circuit represents GIDL, two top cells (one from the upper stack and the other from the lower stack), two bottom cells and $(N_{WL}-4)$ mid cells. The GIDL circuit is the same as the conventional model. $C_{G1,\text{top}}, C_{G2,\text{top}}, C_{BL,\text{top}}, I_{c,\text{top}}, L_{e,\text{top}}$ and $I_{h,\text{top}}$ represent the changed $C_{G1}, C_{G2}, C_{FB}, I_c, L_e$ and $I_h$ of the top cell due to the tapered angle. Components in other cell circuits also represent changed $C_{G1}, C_{G2}, C_{FB}, I_c, L_e$ and $I_h$ according to their position (mid or bottom). This simple modeling is based on the fact that the distance from the GIDL injection point does not affect the erase performance of the cell.

This also means that users of this model can further subdivide cells and add them to the channel node for more precise analysis. Although the simulation time can vary depending on the number of cell circuits, the compact model shows overwhelmingly fast analysis compared to TCAD, which takes tens or hundreds of hours or more. In addition, the voltage source ($V_{WL}$) is added to the cell circuit to consider the WL voltage. In recent NAND operation, different WL voltages have been applied depending on the positions of the cells. This is to suppress the cell-to-cell variation caused by the tapered angle. This operation can be analyzed by using our model.

4. Conclusion

TCAD defines CCS for holes below the valance band of nitride. On the other hand, the conventional compact model defines it for tunneling hole current density. Therefore, we should distinguish between the CCS of TCAD and that of the compact model. Proper CCS for the compact model can be
Fig. 12. Conceptual diagram of the compact model for the structure in Fig. 11.

Fig. 13. Advanced GIDL erase compact model for the structure in Fig. 11. It also includes the effect of tapered, single-side injection and WL voltage. $N_{WL}$ represents the number of WLs.
obtained using the proposed equation. For the structure used in this study, it is necessary to use $2.1 \times 10^5$ times larger CCS for the compact model than that of TCAD. The conventional GIDL erase compact model is successfully validated with this properly converted CCS.

An advanced compact model considering tapered angle, single-side injection and WL voltage is suggested. The advanced model can provide an analysis of the erase operation of recent 3D NAND.

We expect that the relationship between CCSs revealed in this paper will contribute to the validation and development of compact models. We also expect that the proposed advanced compact model will contribute to more realistic analysis of erase operation in recently developed 3D NAND flash memory.

1) Y. Fukuzumi et al., 2007 IEEE Int. Electron Devices Meeting (IEDM), 2007, p. 449.
2) R. Katsumata et al., 2009 IEEE Symp. on VLSI Symp Technology, 2009, p. 136.
3) I. Kim, A. J. Hong, S. M. Kim, E. B. Song, J. H. Park, J. Han, S. Choi, D. Jang, J.-T. Moon, and K. L. Wang, 2009 IEEE Symp. on VLSI Symp Technology, 2009, p. 186.
4) W. Kim, S. Choi, J. Sung, T. Lee, C. Park, H. Ko, J. Jung, I. Yoo, and Y. Park, 2009 IEEE Symp. on VLSI Symp Technology, 2009, p. 188.
5) A. Hubert et al., 2009 IEEE Int. Electron Devices Meeting (IEDM), 2009, p. 637.
6) P. Kalavade, 2020 IEEE Int. Memory Workshop (IMW) 2020, p. 1.
7) K. Parat and C. Dennison, 2015 IEEE Int. Electron Devices Meeting (IEDM), 2015, p. 33.1.