Transient program operation model considering distribution of electrons in 3D NAND flash memories

Dong Chan Lee¹, a) and Hyungcheol Shin¹

Abstract We developed a new compact model for the program operation of 3D NAND Flash memories. A modified 1-D Poisson equation was proposed that shows better accuracy than the existing model by reflecting the spatial distribution of electrons trapped by the program operation. Under various conditions of program voltage (V_{PGM}) and program time (t_{PGM}), the threshold voltage shift (ΔV_t) was extracted by TCAD (Technology Computer-Aided Design) simulation, and we used this data to validate our new model. It also provides validity of the model for program operation in 3D NAND flash memory along with various TCAD analysis data.

Keywords: 3D NAND flash memory, program operation, electron distribution, transient dynamics, poisson equation

1. Introduction

3D NAND Flash memory is considered as a promising memory technology by reducing bit cost and storing more data than NAND Flash memory with 2D planar structure and floating gate [1, 2, 3, 4]. Since it has a gate-all-around (GAA) structure, program dynamics are improved due to the high electric field at the substrate/tunneling oxide interface compared to planar devices [5, 6]. In NAND Flash memory, the logical value stored by a memory cell is determined by the cell’s threshold voltage (V_t) [7]. Thus, it is essential to predict the exact V_t value changed by the program operation. Previous studies have analyzed transient program operation through the 1-D Poisson equation with radial coordinates [8, 9, 10, 11, 12, 13]. In [8], the spatial distribution of trapped electrons in charge-trap nitride (CTN) was not considered, and it was assumed that the trapped electron density is uniform in all CTN regions. However, according to several papers and Technology Computer-Aided Design (TCAD) simulation, electrons are mainly trapped from the region close to the tunneling oxide in the CTN [14, 15]. Therefore, when using the existing model in [8], the distribution of trapped electrons is not properly considered, so the threshold voltage shift (ΔV_t) is not calculated accurately.

In this Letter, it is assumed that CTN is divided into two regions: region 1 where electron trapping occurs (close to tunneling oxide), and region 2 where electrons are not trapped, in order to reflect the distribution of trapped electrons effectively. A modified Poisson equation is presented accordingly, and the boundary of the regions in the CTN is determined at the point that showing best fits with the ΔV_t value. Various program voltage (V_{PGM}) conditions are used for the model, and several TCAD analysis data are presented.

2. Simulation details

The device parameters used in the TCAD simulation are specified in Table I [16]. To reflect only the effect of the distribution of trapped electrons, the substrate is assumed to be single crystal silicon without grain boundaries. The effects of interface trap and electron emission are not considered. The pass voltage of the unselected word lines is 6V and the voltage of the string-select line is 2.4V [17]. The threshold voltage (V_t) is defined as the voltage when a 10nA bit line current occurs by applying a constant current method [18, 19], and ΔV_t is defined as the difference between V_t in the initial state without trapped electrons and V_t in the programmed state with trapped electrons.

3. Electrostatic solution

Table II shows the number of trapped electrons and the results of ΔV_t calculations derived from the TCAD simulation and the previous model given in [8]. It was calculated based on our device parameters. When substrate effect (potential drop due to channel potential (V_{ch}) and flat band voltage

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word-line length (L_w)</td>
<td>28nm</td>
</tr>
<tr>
<td>Spacer length (L_s)</td>
<td>28nm</td>
</tr>
<tr>
<td>Substrate radius (r_0)</td>
<td>25nm</td>
</tr>
<tr>
<td>Tunneling oxide thickness (t_{ox})</td>
<td>5nm</td>
</tr>
<tr>
<td>CTN thickness (t_{ctn} = t_{ox})</td>
<td>5nm</td>
</tr>
<tr>
<td>Blocking oxide thickness (t_{b})</td>
<td>6nm</td>
</tr>
<tr>
<td>Tunneling/Blocking oxide dielectric constant (ε_{tox}/ε_{b})</td>
<td>3.5ε_0</td>
</tr>
<tr>
<td>CTN dielectric constant (ε_{ctn})</td>
<td>7ε_0</td>
</tr>
<tr>
<td>Program voltage (V_{PGM})</td>
<td>12V/14V/16V/18V</td>
</tr>
<tr>
<td>Program time (t_{PGM})</td>
<td>0 to 9μs</td>
</tr>
<tr>
<td>Rising time at (t_{rise})</td>
<td>1 μs</td>
</tr>
<tr>
<td>Total trap density (N_t)</td>
<td>4 × 10^{20}cm^{-3}</td>
</tr>
</tbody>
</table>

¹ Inter-University Semiconductor Research Center (ISRC) and School of Electrical Engineering and Computer Science, Seoul National University, Seoul 151-747, Korea

a) ldc1207@snu.ac.kr

DOI: 10.1587/elex.17.20200335
Received October 9, 2020
Accepted October 26, 2020
Publicized November 6, 2020
Copyedited December 10, 2020
Table II  The number of trapped electrons and $\Delta V_t$ calculated by applying transient dynamics at $V_{PGM} = 14V, t_{PGM} = 9\mu s$.

<table>
<thead>
<tr>
<th></th>
<th># of eTrappedCharge</th>
<th>$\Delta V_t$ [V]</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCAD</td>
<td>59</td>
<td>0.301</td>
</tr>
<tr>
<td>[8] without substrate</td>
<td>145</td>
<td>0.751</td>
</tr>
<tr>
<td>effect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[8] with substrate</td>
<td>68</td>
<td>0.35</td>
</tr>
<tr>
<td>effect</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1  TCAD simulation results of 3D NAND Flash memory. (a) Cross-sectional view of unit cell structure at the programmed state. (b) Trapped electron density according to the program voltage ($V_{PGM}$) along the direction of A-A’ (middle position of the target cell).

$V_f$ is considered, only the voltage applied to the tunneling oxide/CTN/blocking oxide (ONO) layer is reflected in the model [20, 21, 22, 23]. Thus, the $\Delta V_t$ calculation becomes similar to TCAD simulation data. However, there is still an error and it becomes larger in proportion to $V_{PGM}$ and $t_{PGM}$. The reason for this error is that the previous study [8] modeled assuming uniform trapped electron density ($n_t$; units:cm$^{-3}$) for all regions in CTN without considering the spatial distribution and number of electrons. Fig. 1(a) is a result of the transient simulation of program operation in TCAD, showing that the electrons are mainly trapped in the CTN region close to the tunneling oxide. Also, Fig. 1(b) shows that the trapped electron distribution changes as $V_{PGM}$ increases, but still shows the same tendency. Therefore, electron distribution must be reflected in the model. To solve this problem, we divided CTN into two regions as shown in Fig. 2: region 1 where electron trapping occurs, and region 2 that does not. This assumption allows the spatial distribution of trapped electrons to be reflected effectively in the model so that the number of electrons and $\Delta V_t$ can be calculated accurately. The boundary-value of region 1 and region 2 is set to a new parameter $r_x$ as follows:

$V_t(r) = C_1 \ln \frac{r}{r_0} \left( r_0 < r < r_{to} \right) + C_2 \ln \frac{r}{r_t} + C_3 \ln \frac{r}{r_n} \left( r_{to} < r < r_x \right) + C_4 \ln \frac{r}{r_{bo}} \left( r_x < r < r_{bo} \right)$

where $C_1 \sim C_4$ is a constant, $r_x = r_{to} + \eta f_n$.

Note that the setting of the boundary condition is the same as [8], except for the substrate effect and region 1 are reflected in the model. The expression for constants ($C_1 \sim C_4$) is as follows:

$C_1 = \frac{V_{PGM} - V_{fch} - V_{fbo}}{\alpha} + \frac{qn_t}{2\alpha} \frac{1}{r_{bo}} \ln \frac{r_x}{r_{to}}$

$C_2 = \frac{\varepsilon_0}{\varepsilon_n} \left( C_1 - \frac{qn_t^2}{2\varepsilon_n} \frac{1}{r_{to}^2} \right) \left( r_x^2 - r_{to}^2 \right)$

Fig. 2  Conceptual image of the trapped electron distribution in CTN assumed in [8] (dash) and in this work (line).

$r_x = r_{to} + \eta f_n$  \hspace{1cm} (1)

where $\eta$ is a proportional factor associated with the thickness of the nitride ($T_n$), with a value between 0 and 1. The value of $\eta$ will be determined when fitting the $\Delta V_t$ with TCAD data by applying transient dynamics in results section. The electrostatics of the 3D NAND flash memory can be calculated by solving a modified 1-D Poisson equation in radial coordinates that considering the distribution of trapped electrons is taken into account:

$\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial V}{\partial r} \right) = -\frac{qn_t}{\varepsilon_n} \left[ H(r - r_{to}) - H(r - r_x) \right] \hspace{1cm} (2)$

where $V(r)$ is the electrostatic potential, and $q$ is the electron charge. The Heaviside step function $H$ is used to assume that uniform $n_t$ exists only in region 1. By solving the modified Poisson equation at Eq. (2), expressions for $V(r)$ and electric field $F(r)$ are obtained in four regions:

$V(r) = \begin{cases} C_1 \ln \frac{r}{r_0} & \left( r_0 < r < r_{to} \right) \\ \frac{q}{4\varepsilon_n} \left( r^2 - r_{to}^2 \right) + C_1 \ln \frac{r}{r_{to}} + C_3 \ln \frac{r}{r_n} & \left( r_{to} < r < r_x \right) \\ C_1 \ln \frac{r}{r_{to}} + C_2 \ln \frac{r}{r_t} + C_3 \ln \frac{r}{r_n} + \frac{qn_t}{4\varepsilon_n} \left( r_x^2 - r_{to}^2 \right) & \left( r_x < r < r_{bo} \right) \end{cases}$

$F(r) = -\frac{\partial V}{\partial r} \begin{cases} -\frac{C_1}{r} & \left( r_0 < r < r_{to} \right) \\ -\frac{q}{4\varepsilon_n} r^2 - C_1 & \left( r_{to} < r < r_x \right) \\ -C_2 & \left( r_x < r < r_{bo} \right) \end{cases}$

![Image](image-url)
obtained:

\[ C_3 = \frac{\varepsilon_{t0}}{\varepsilon_n} C_1 + \frac{q n_t}{2\varepsilon_n} \left( r_x^2 - r_{t0}^2 \right) \]  

(7)

\[ C_4 = \frac{\varepsilon_{b0}}{\varepsilon_n} C_1 + \frac{q n_t}{2\varepsilon_b} \left( r_x^2 - r_{b0}^2 \right) \]  

(8)

\[ \alpha = \ln \frac{r_{t0}}{r_{t0}} + \frac{e_n}{e_n} \ln \frac{n_t}{r_{b0}} + \frac{e_n}{e_b} \ln \frac{n_t}{r_n} \]  

(9)

where \( \alpha \) is same formula as [8]. From Eq. (4) and Eq. (5), \( \Delta V_t \) is calculated as the amount of change in \( V_{PGM} \) required to recover the electric field of the substrate/tunneling oxide interface \( (F_i) \) when \( n_t = 0 \), and the following equation is obtained:

\[ \Delta V_t = -\frac{q n_t}{2} \left[ \frac{1}{\varepsilon_{t0}} \ln \frac{r_x}{r_{t0}} - \left( \frac{1}{\varepsilon_{b0}} \ln \frac{n_t}{r_n} + \frac{1}{\varepsilon_n} \ln \frac{n_t}{r_x} + \frac{1}{2\varepsilon_n} \left( r_x^2 - r_{t0}^2 \right) \right] \]  

(10)

4. Transient dynamics

The FN equation is used as a theoretical expression for tunneling current density \( (J_n) \) [24, 25, 26, 27, 28, 29].

\[ J_n = A F_i^2 \exp \left[ -\frac{B}{F_i} \right] \]  

(11)

where \( A \) and \( B \) are determined by considering the potential barrier and effective tunneling mass of the electron, and same values are used for both TCAD simulation and the model \( (A = 3.07 \times 10^{-7} A/V^2, B = 2.41 \times 10^8 V/cm) \). The transient dynamics of program operation is expressed by the following equation [30]:

\[ \frac{d n_t}{d t} = \frac{J_n}{q} \left( \frac{r_{t0}}{r_{t0} + \gamma n_t} \right) \sigma_P (N_t - n_t) - e_a n_t \]  

(12)

Here, the new parameter \( \gamma \) is a charge centroid factor related to region 1 where electron trapping occurs in CTN and corresponds to \( \eta/2 \). In [8], the value of \( \gamma \) was a constant of 1/2 because electron trapping was assumed in all areas within the CTN. Since the effects caused by electron emission are not considered in this study, the emission rate \( (e_a) \) is set to 0. Fig. 3 shows the modeling methodology in this study. Since there is no analytical solution for Eq. (12), we calculated it numerically for discontinuous time. Starting from the initial condition of \( n_t = 0 \), \( F_i \) and \( J_n \) are calculated, and then the increase in trapped electron density \( (\Delta n_t) \) is calculated. The \( \Delta V_t \) is derived by the modified \( n_t \) value, and all of the processes are repeated with increasing time step \( (\Delta t) \).

5. Results

Fig. 4 shows the \( \Delta V_t \) calculated according to the electron density through the previous model in [8] and TCAD simulation. In TCAD simulation, a fixed electron density is set by assuming uniform electron distribution in all regions of the CTN in the target cell. Since the spatial distribution and the number of the electrons are set the same, a slight error is caused by physical phenomena that reflected only in TCAD simulation. However, it is negligible because it is not large enough to affect the modeling process. Fig. 5(a) shows the \( \Delta V_t \) calculated by applying transient dynamics. In the case of \( \eta = 0.77 \) that using our model, the result shows a good agreement with the TCAD data at all \( V_{PGM} \) and it is more accurate than the previous model in [8]. For reference, when \( \eta = 1 \) is applied to our model, the calculation results are the same as the results of [8] with considering the substrate effect. Fig. 5(b) shows the \( n_t \) according to the position in CTN, and it is calculated by TCAD and model when \( V_{PGM} = 9\mu \) in Fig. 5(a). In our proposed model, \( r_x \) and region 1 are set according to the value of \( \eta \) and the \( n_t \) is determined by the solution of the modified 1-D Poisson equation. The number of trapped electrons is 55, which is calculated based on \( n_t \) and the volume of the CTN layer in the proposed model \( (\eta = 0.77) \). It is closer to TCAD data than the results calculated through [8] in Table II. Fig. 6(a) shows the \( \Delta V_t \) fitted for the \( V_{PGM} \) in various \( V_{PGM} \). Fig. 6(b) shows the value of \( \eta \) with the best fitting result for each \( V_{PGM} \) in Fig. 6(a), where its values are proportional to \( V_{PGM} \). This indicates that the proposed model could reflects the tendency for the
expansion of the trapped electron distribution according to $V_{PGM}$ as mentioned in Fig. 1(b). Because the $F_i$ and $J_n$ are small in low $V_{PGM}$, fewer electron trapping occurs, resulting in less variation in the distribution of trapped electrons over $t_{PGM}$. Conversely, in high $V_{PGM}$, there is a large variation in trapped electrons distribution. In order to reflect this characteristic more accurately, at high $V_{PGM}$, $\eta$ should be set to increase in proportion to $t_{PGM}$ at $V_{PGM} = 18\,V$. It is more accurate when time dependency is considered. Here, the value of $\eta$ changed in proportion to the $t_{PGM}$ is shown in Fig. 7(b).

6. Conclusion

In this letter, a compact model for transient program dynamics in 3D NAND Flash memories was presented. The 1-D Poisson equation was modified so that our model reflects the trapped electron distribution in CTN layer effectively. The CTN was divided into two regions, and it was assumed that electron trapping occurred only in the region close to the tunneling oxide. The parameter for determining the boundary value of the two regions were used to fit the $\Delta V$ correctly. The proposed model shows a good agreement with the TCAD simulation results. Lastly, the program time dependency was considered, and the accuracy was improved.

References


