Estimation of Trapped Charge Density in SONOS Flash Memory Using a Parallel Capacitor Model

Hua-Min Li, Gang Zhang, *Student Member, IEEE*, and Won Jong Yoo, *Senior Member, IEEE*

***Abstract*—This paper proposes a model established using the parallel connection of MOS and floating-gate MOS capacitors to examine the electric properties of polysilicon–oxide–nitride–oxide–silicon (SONOS) Flash memory in both the fresh and programmed states. A linear relationship between threshold voltage shift and effective trapped charge density was revealed via the threshold electric field. The simulation of the channel current was given based on our model, being in agreement with the experimental results for both forward and reverse read modes. The model and theory make an easy and time-saving approach to comprehensively analyze the trapped charge and its effect on other electric properties, e.g., electric field, oxide capacitance, and charge distribution, contributing to the control of the write/erase operation, the optimization of the device structure, and the characteristics of the retention properties in SONOS Flash memory.**

***Index Terms*—Polysilicon–oxide–nitride–oxide–silicon (SONOS), threshold electric field, threshold voltage shift, trapped charge density.**

I. Introduction

THE CHARGE distribution and energy level of the trapped charge from channel hot-electron injection (CHEI) are essential parameters affecting the cell endurance and scalability of polysilicon (poly-Si)–oxide–nitride–oxide–silicon (SONOS) Flash memory [1]–[3]. The spatial distribution of the trapped charge has been extensively examined using a variety of approaches, e.g., quasi-2-D analytical model [4], gate-induced-drain-leakage simulation, charge pumping [5], analysis of current transient, threshold voltage, and subthreshold slope (SS) [6]–[8]. Both the bulk and the interface trap depth in SONOS devices have been examined using capacitance–voltage (C–V) measurements and deep-level transient spectroscopy (DLTS) [9]. Electron retention in stacked dielectrics has been also modeled using a thermal emission mechanism [10]. However, the relations between the trapped charge density and the other electric properties, e.g., electric field, oxide capacitance, and charge distribution, have not been clearly investigated in the previous studies. These relations are essential for comprehensively understanding the influence of the trapped charge on the electrical performance, i.e., the control of the write/erase (W/E) operation, the optimization of the device structure, and the characterization of the retention properties in SONOS devices [11], [12].

In this paper, an equivalent model established by the parallel connection of MOS and floating-gate MOS capacitors was proposed for characterization of the trapped charge in SONOS memory devices. Threshold voltage shift $\Delta V_{\rm th}$ was obtained as a linear function of effective trapped charge density $\overline{\sigma}_t$ using threshold electric field $E_{\rm th}$, which was proposed as a signature for the formation of the inversion channel in order to link the threshold voltage in both fresh and programmed states. A comprehensive analysis of various electrical properties was carried out as a function of the gate voltage, providing the relations between the trapped charge density and the electric field, oxide capacitance, charge distribution, etc. A simulation of the channel current under both forward and reverse read modes was in agreement with experimental results. Compared to previous studies, our model has its own advantages. First, $E_{\rm th}$ was proposed newly as a key parameter in our model to connect $V_{\rm th}$ before and after programming, owing to its constant value in both fresh and programmed states. Second, our model provides comprehensive information of various electrical properties, e.g., charge distribution at the gate, trapping layer and substrate, electric field through the block and tunnel oxides, oxide capacitance in the fresh and programmed states, and drain-induced barrier lowering (DIBL) and its effect on the other properties, in which some pieces of information are rarely obtained in conventional approaches. Third, our model is very simple and time saving. According to the experimental $\Delta V_{\rm th}$ and the derived equations, various properties can be obtained via simple calculations in minutes. Finally, our model has the potential for wide applications, e.g., the control of the W/E operation, the optimization of the device structure, and the characterization of the retention properties. In addition, it can be modified and applied to the SONOS cells using various trapping materials or the modeling of the split-gate SONOS cells using separated charge-storage nodes. These indicate the parallel capacitor model to be an effective and convincing approach for a comprehensive analysis of the trapped charge in SONOS memory devices.

II. Device Fabrication and Electrical Analysis

The SONOS memory device was fabricated using a standard complementary metal–oxide–semiconductor process with a poly-Si/SiO_2/HfO_2/SiO_2/Si gate stack. After the thermal

Manuscript received August 13, 2010; revised April 1, 2011; accepted June 21, 2011. Date of publication August 18, 2011; date of current version September 21, 2011. This work was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology under Grant 2011-0006268, 2011-0010274. The review of this paper was arranged by Editor S. Deleonibus.

The authors are with the Department of Nanoscience and Technology, Samsung-SKKU Graphene Center, SKKU Advanced Institute of Nano Technology, Sungkyunkwan University, Suwon 440-746, Korea (e-mail: yoowj@skku.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2011.2162730

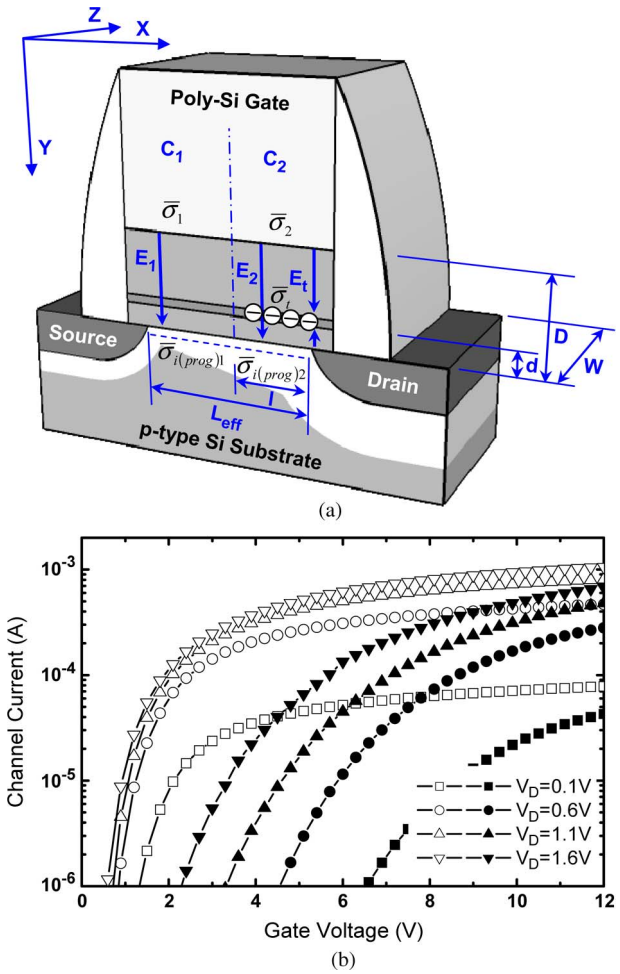


Fig. 1. (a) Illustration of the SONOS model formed by two capacitors, i.e., C_1 and C_2, connected in parallel. (b) Experimental result of I–V characteristics under the forward read mode for various V_D levels (0.1, 0.6, 1.1, and 1.6 V) before (hollow symbol) and after (solid symbol) programming.

growth of a 3-nm-thick tunnel oxide d on p-type Si substrate, a 6-nm-thick HfO_2 film with an effective oxide thickness (EOT) of $\sim$1 nm was deposited by sputtering. A 9-nm-thick block oxide was deposited, followed by the deposition of 180-nm-thick poly-Si. The device has a physical gate length $L_{\rm eff}$ of 200 nm, a channel width W of 2 μm, and an EOT of the $SiO_2/HfO_2/SiO_2$ stack D of 13 nm, as shown in Fig. 1(a). The channel current–gate voltage (I–V) characteristics under both forward and reverse read modes were carried out before and after CHEI programming at an applied gate voltage V_G of 5 V and a drain voltage V_D of 10 V for 1 ms with the source and substrate grounded, as shown in Fig. 1(b).

III. Modeling and Theory

The following assumptions are made to establish the parallel capacitor model.

1) The charge at the gate, trapping layer, and substrate are uniformly distributed over the surface, generating their corresponding uniform electric fields as

$$E = \frac{\overline{\sigma}}{\varepsilon_{\rm ox}} \tag{1}$$

where $\overline{\sigma}$ is the average or effective charge density, $\varepsilon_{\rm ox}$ is the permittivity of SiO_2, and the direction of E is perpendicular to the charge surface.

2) The charge density of the inversion channel σ_i is neglected when the gate voltage is less than the threshold voltage ($V_G < V_{\rm th}$) [13] because the strong inversion condition is not reached. The only uncompensated charge in the Si substrate is due to the doped ions in the depletion layer.

3) When the gate voltage is greater than the threshold voltage ($V_G > V_{\rm th}$), the surface potential φ_s and, consequently, the charge density in the depletion layer σ_d are constant, which can be expressed as $\varphi_s = 2\phi_F$ and $\sigma_d = 2\sqrt{\phi_F \varepsilon_{\rm Si} q N_A}$ [13], respectively, where ϕ_F is the Fermi level, $\varepsilon_{\rm Si}$ is the permittivity of Si, q is the unit of charge, and N_A is the substrate doping concentration. The gate voltage beyond the threshold voltage ($V_G - V_{\rm th}$) is used to form an inversion layer as $\sigma_i = (V_G - V_{\rm th})C_{\rm ox}$, where $C_{\rm ox}$ is the oxide capacitance.

When the gate voltage is equal to the threshold voltage ($V_G = V_{\rm th}$), the electric field at the substrate surface reaches $E_{\rm th}$, which is defined as a vertical electric field needed to form the inversion channel. The value of $E_{\rm th}$ is only determined by the substrate doping concentration, and thereby, it is the same before and after CHEI programming. This is a unique advantage compared to $V_{\rm th}$, and it is used to link $V_{\rm th}$ in both fresh and programmed states. In this model, the fresh SONOS cell is analyzed in an equivalent manner to the MOSFET device, where $E_{\rm th}$ is defined by the threshold voltage in the fresh state ($V_{\rm th(fresh)}$) as

$$E_{\rm th} = \frac{V_{\rm th(fresh)}}{D} \tag{2}$$

and the average charge density at the bottom of poly-Si is $\overline{\sigma}_1$, generating its electric field E_1 as $\overline{\sigma}_1/\varepsilon_{\rm ox}$. The programmed SONOS cell is considered a MOS capacitor C_1 ($0 < x < L - l$) and a floating-gate MOS capacitor C_2 ($L - l < x < L$) in a parallel connection [see Fig. 1(a)]. Since CHEI occurs near the drain end [7], C_1 is unaffected and its average gate charge density and electric field are equal to $\overline{\sigma}_1$ and E_1, respectively. In C_2, the injected charges are assumed to be trapped at the bottom of the trapping layer as $\overline{\sigma}_t$, establishing its electric field E_t of $\overline{\sigma}_t/\varepsilon_{\rm ox}$. The average gate charge density becomes $\overline{\sigma}_2$ with an established electric field E_2 of $\overline{\sigma}_2/\varepsilon_{\rm ox}$.

First, the electric fields through the block oxide ($0 < y < D - d$) and the tunnel oxide ($D - d < y < D$) in C_2 are estimated using the principle of superposition as $E_{\rm (block)} = E_2 + E_t = (\overline{\sigma}_2 + \overline{\sigma}_t)/\varepsilon_{\rm ox}$ and $E_{\rm (tunnel)} = E_2 - E_t = (\overline{\sigma}_2 - \overline{\sigma}_t)/\varepsilon_{\rm ox}$, respectively. Second, due to the parallel connection, V_G in C_2 as $\int_0^{D-d} E_{\rm (block)} dy + \int_{D-d}^{D} E_{\rm (tunnel)} dy$ is equal to that in C_1 as $\int_0^D E_1 dy$. Therefore, the relationships of $\overline{\sigma}_1$, $\overline{\sigma}_2$, and $\overline{\sigma}_t$ are obtained as

$$\overline{\sigma}_1 = \frac{\varepsilon_{\rm ox}}{D} V_G \tag{3}$$

$$\overline{\sigma}_2 = \overline{\sigma}_1 - \frac{D - 2d}{D}\overline{\sigma}_t. \tag{4}$$

Assuming that $\overline{\sigma}_t$ is infinitely small, $\overline{\sigma}_2$ would be the same as $\overline{\sigma}_1$, which is consistent with the experimental results. Third, according to (4), the electric fields through the block and tunnel oxides are rewritten as follows:

$$E_{(\mathrm{block})} = \frac{V_G}{D} + \frac{2d}{D} \cdot \frac{\overline{\sigma}_t}{\varepsilon_{\mathrm{ox}}} \tag{5}$$

$$E_{(\mathrm{tunnel})} = \frac{V_G}{D} - \frac{2(D-d)}{D} \cdot \frac{\overline{\sigma}_t}{\varepsilon_{\mathrm{ox}}}. \tag{6}$$

Since the formation of the conducting channel at the interface is induced only by $E_{(\mathrm{tunnel})}$, it is reasonable to assume that $E_{(\mathrm{tunnel})} = E_{\mathrm{th}}$ when V_G reaches the threshold voltage in the programmed state $V_{\mathrm{th(prog)}}$. Therefore, the threshold voltage shift, i.e., $\Delta V_{\mathrm{th}} = V_{\mathrm{th(prog)}} - V_{\mathrm{th(fresh)}}$, can be obtained as follows:

$$\Delta V_{\mathrm{th}} = 2(D-d)\frac{\overline{\sigma}_t}{\varepsilon_{\mathrm{ox}}}. \tag{7}$$

Equation (7) (which can be also represented with the capacitance of the block oxide C_{block} as $\Delta V_{\mathrm{th}} = 2\overline{\sigma}_t / C_{\mathrm{block}}$) shows that ΔV_{th} is a linear function of $\overline{\sigma}_t$, as well as the distance between the trapped charge and the gate electrode. Compared with the conventional expression as $\Delta V_{\mathrm{th}} = \overline{\sigma}_t / C_{\mathrm{block}}$ [11] for the bulk trapping layer in which the charge on the top surface is considered to interact with the gate, there is a coefficient "2" in our formula due to the assumption of the zero-thickness charge sheet that provides the double interaction with both gate and substrate. Assuming a more complicated model, where the thickness of the trapping layer is included and the charge density on the top and bottom surfaces is individually considered, the same expression as $\Delta V_{\mathrm{th}} = \overline{\sigma}_t / C_{\mathrm{block}}$ can be still derived in our approach.

According to the experimental value of ΔV_{th} [see Fig. 1(b)], the electric properties, e.g., electric field, oxide capacitance, charge distribution at the gate, inversion channel, and depletion region, are obtained by combining (7) with their corresponding equations. For example, the expressions of $E_{(\mathrm{block})}$ and $E_{(\mathrm{tunnel})}$ [see (5) and (6)] are rewritten as $[V_G + d\,\Delta V_{\mathrm{th}}/(D-d)]/D$ and $(V_G - \Delta V_{\mathrm{th}})/D$, respectively, which are shown in Fig. 2. E_{th} is a constant of 1.06×10^8 V/m, determined by (2). E_1, $E_{(\mathrm{block})}$, and $E_{(\mathrm{tunnel})}$ are expressed as linear functions of V_G (0–12 V) for various V_D (0.1, 0.6, 1.1, and 1.6 V) levels. In the fresh state, the location of the intersection between E_{th} and E_1 (point A) indicates its corresponding $V_{\mathrm{th(fresh)}}$. As a comparison, $E_{(\mathrm{block})}$ becomes higher and $E_{(\mathrm{tunnel})}$ becomes lower than E_1 in the programmed state for the same V_G condition due to the influence of the negative trapped charges. Therefore, $V_{\mathrm{th(prog)}}$ requires a higher value for compensation to make $E_{(\mathrm{tunnel})}$ equal to E_{th}, as indicated by the locations of the intersection between E_{th} and $E_{(\mathrm{tunnel})}$ (points B–E). In addition, the distribution of the electric field along the y-axis is also shown in the inset, suggesting that E_1 is equal to E_{th} for $V_G = V_{\mathrm{th(fresh)}}$, and $E_{(\mathrm{tunnel})}$ is equal to E_{th} for $V_G = V_{\mathrm{th(prog)}}$.

The oxide capacitance for C_2 in the programmed state is expressed as $C_{\mathrm{ox(prog)2}} = \overline{\sigma}_2 / V_G = \varepsilon_{\mathrm{ox}} / [D + (D-2d)\overline{\sigma}_t / \overline{\sigma}_2]$, which is smaller than that for the fresh state or C_1 in the programmed state obtained as $C_{\mathrm{ox(fresh)}} = C_{\mathrm{ox(prog)1}} = \overline{\sigma}_1 / V_G = \varepsilon_{\mathrm{ox}} / D$, as shown in Fig. 3. $C_{\mathrm{ox(prog)2}}$ increases close to $C_{\mathrm{ox(prog)1}}$ with increasing V_G, owing to the increase in V_G-dependent $\overline{\sigma}_2$. The distribution of the oxide capacitance along the x-axis is also shown in the inset, suggesting a reduction of the oxide capacitance in the region where the trapped charge occupied (from $L_{\mathrm{eff}} - l$ to L_{eff}).

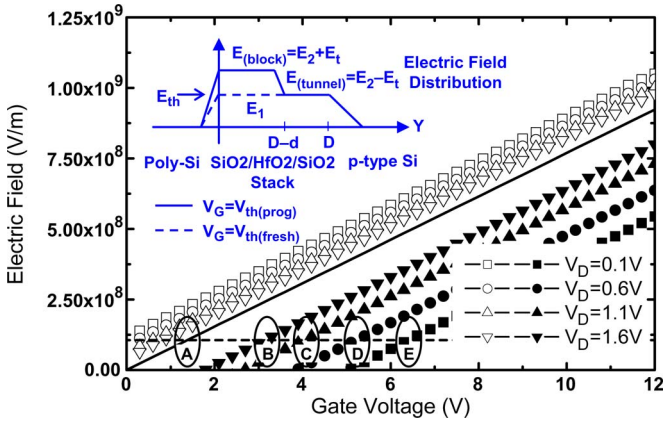


Fig. 2. Estimation of the electric field $E_{(\mathrm{block})}$ (hollow symbol) and $E_{(\mathrm{tunnel})}$ (solid symbol) as a function of V_G, as compared to E_1 (solid line) and E_{th} (dashed line). The location of the intersection between E_{th} and E_1 (point A) indicates its corresponding $V_{\mathrm{th(fresh)}}$. The location of the intersection between E_{th} and $E_{(\mathrm{tunnel})}$ (points B–E) indicates its corresponding $V_{\mathrm{th(prog)}}$ for various V_D levels, being consistent with the experimental results. The electric field distribution along the y-axis when $V_G = V_{\mathrm{th(fresh)}}$ and $V_G = V_{\mathrm{th(prog)}}$ is shown in the inset.

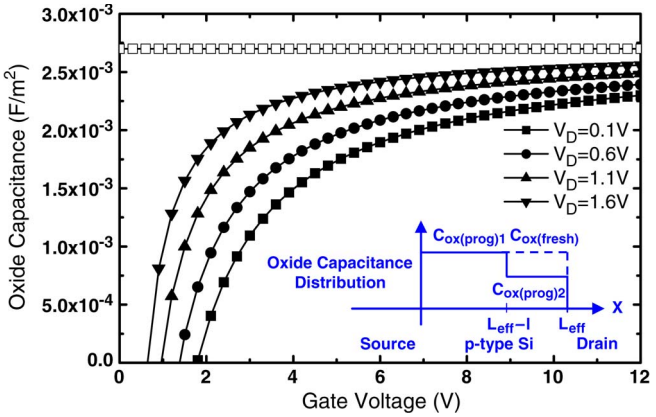


Fig. 3. Estimation of the oxide capacitance $C_{\mathrm{ox(fresh)}}$ or $C_{\mathrm{ox(prog)1}}$ (hollow symbol) and $C_{\mathrm{ox(prog)2}}$ (solid symbol) as a function of V_G with their distribution along the x-axis under the same V_G, as shown in the inset.

The charge distributions at the trapping layer, gate, and inversion channel are shown in Fig. 4. $\overline{\sigma}_t$ is obtained by (7), showing independence of V_G. Due to its effect, $\overline{\sigma}_2$ that is obtained as a function of V_G [see (4) and (7)] becomes smaller than $\overline{\sigma}_1$ [see (3)] for the same V_G condition, as shown in Fig. 4(a). Based on the oxide capacitance (see Fig. 3), the charge densities at the inversion channel in the fresh state $\overline{\sigma}_{i(\mathrm{fresh})}$ (or for C_1 in the programmed state $\overline{\sigma}_{i(\mathrm{prog})1}$) and for C_2 in the programmed state $\overline{\sigma}_{i(\mathrm{prog})2}$ are expressed as $(V_G - V_{\mathrm{th(fresh)}})C_{\mathrm{ox(fresh)}}$ and $(V_G - V_{\mathrm{th(prog)2}})C_{\mathrm{ox(prog)2}}$, respectively, as shown in Fig. 4(b). $\overline{\sigma}_{i(\mathrm{prog})2}$ becomes smaller than $\overline{\sigma}_{i(\mathrm{prog})1}$ for the same V_G condition due to the reduction of $C_{\mathrm{ox(prog)2}}$ compared with $C_{\mathrm{ox(prog)1}}$, and therefore, it requires

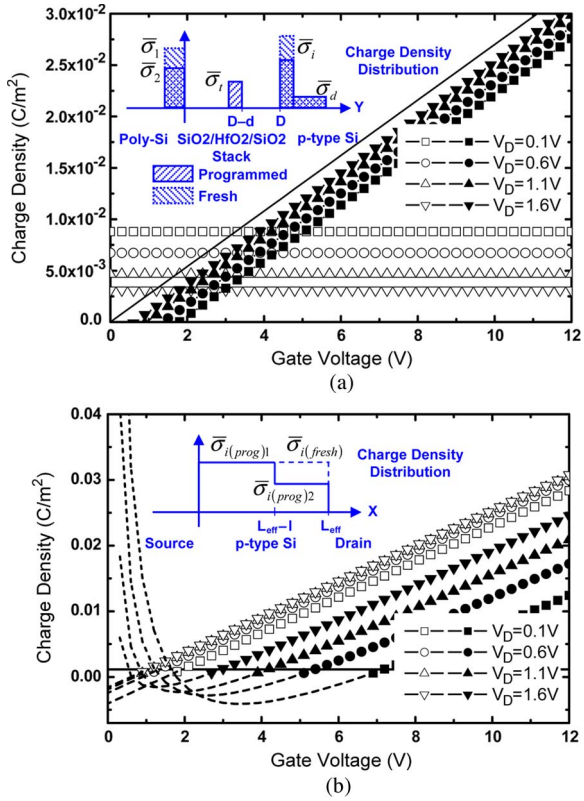


Fig. 4. (a) Estimation of the charge density $\overline{\sigma}_1$ (solid line), $\overline{\sigma}_2$ (solid symbol), and $\overline{\sigma}_t$ (hollow symbol) as a function of V_G. The charge distribution along the y-axis under the same V_G is shown in the inset, where $\overline{\sigma}_2$ and $\overline{\sigma}_i$ after programming are reduced due to the injected $\overline{\sigma}_t$. (b) Estimation of the charge density $\overline{\sigma}_{i(\mathrm{fresh})}$ or $\overline{\sigma}_{i(\mathrm{prog})1}$ (hollow symbol) and $\overline{\sigma}_{i(\mathrm{prog})2}$ (solid symbol) as a function of V_G with their distribution along the x-axis under the same V_G, as shown in the inset. The charge density in the depletion region $\overline{\sigma}_d$ (solid line) is also shown.

a higher V_G for compensation to form the inversion channel. The charge distributions along the y- and x-axes are shown in the insets in Fig. 4(a) and (b), respectively, suggesting a reduction of the charge density at the gate and the inversion layer where the trapped charge occupied (from $L_{\mathrm{eff}} - l$ to L_{eff}).

Moreover, DIBL is clearly observed via the reduction of $\overline{\sigma}_t$ from 8.80×10^{-3} to 3.11×10^{-3} C/m^2 by increasing V_D from 0.1 to 1.6 V [see Fig. 4(a)]. In the meantime, the effect of the DIBL on the other electric properties, e.g., the trapped charge-induced differences between the electric fields $E_{(\mathrm{block})}$ and $E_{(\mathrm{tunnel})}$, oxide capacitance values $C_{\mathrm{ox(prog)}1}$ and $C_{\mathrm{ox(prog)}2}$, charge distributions $\overline{\sigma}_1$ and $\overline{\sigma}_2$, and $\overline{\sigma}_{i(\mathrm{prog})1}$ and $\overline{\sigma}_{i(\mathrm{prog})2}$, is gradually and clearly reduced (see Figs. 2–4).

IV. Channel Current Simulation

Considering the effect of V_D that causes a nonuniform distribution of the inversion layer charge density along the channel, the varying threshold voltage in the fresh SONOS device $(v_{\mathrm{th(fresh)}})$ can be expressed in terms of the surface potential variation as $V_{\mathrm{FB}} + \varphi_s + (2\varepsilon_{\mathrm{Si}} q N_A \varphi_s)^{1/2}/C_{\mathrm{ox(fresh)}}$, where V_{FB} is the flatband voltage, and N_A is the acceptor concentration of the Si substrate. The drain current is obtained by the simulation of a MOSFET device via Simulation Program with Integrated Circuit Emphasis (SPICE) Level 2 model [13].

For the programmed SONOS device, the threshold voltage in C_1 $(v_{\mathrm{th(prog)}1})$ can be still expressed as $v_{\mathrm{th(fresh)}}$. However, it is affected by the trapped negative charges in C_2 and is expressed as follows:

$$\begin{aligned} v_{\mathrm{th(prog)}2} &= v_{\mathrm{th(fresh)}} + \Delta V_{\mathrm{th}} \\ &= V_{\mathrm{FB}} + \varphi_s + \frac{\sqrt{2\varepsilon_{\mathrm{Si}} q N_A}}{C_{\mathrm{ox(fresh)}}}\sqrt{\varphi_s} + \Delta V_{\mathrm{th}}. \end{aligned} \tag{8}$$

Therefore, the average charge density of the inversion layer is obtained by combining capacitors C_1 and C_2 with their corresponding threshold voltages when φ_s was varied from $2\phi_F$ at the source end to $2\phi_F + V_D$ at the drain end [see Fig. 1(a)] as follows:

$$\begin{aligned} \overline{\sigma}_i &= \frac{1}{V_D}\int_{2\phi_F}^{2\phi_F+V_D} \left[V_G - v_{\mathrm{th}}(\varphi_s)\right] C_{\mathrm{ox}} d\varphi_s \\ &= \frac{1}{V_D}\left\{ \int_{2\phi_F}^{2\phi_F+(1-l/L_{\mathrm{eff}})V_D} \left[V_G - v_{\mathrm{th(fresh)}}\right] C_{\mathrm{ox(fresh)}} d\varphi_s \right. \\ &\quad \left. + \int_{2\phi_F+(1-l/L_{\mathrm{eff}})V_D}^{2\phi_F+V_D} \left[V_G - v_{\mathrm{th(prog)}2}\right] C_{\mathrm{ox(prog)}2} d\varphi_s \right\}. \end{aligned} \tag{9}$$

The expression of $\overline{\sigma}_i$ can be obtained by solving the above integral with (8). The drain current after programming, therefore, is obtained as follows [13]:

$$I_{D(\mathrm{prog})} = \frac{\mu_0 W}{L_{\mathrm{eff}}} \overline{\sigma}_i V_D \tag{10}$$

where μ_0 is the electron mobility. According to (7)–(10), the drain current after programming is affected by the density $\overline{\sigma}_t$ and the occupied length l of the injected charge in the trapping layer, indicating the height and the width of the potential barrier in the inversion channel, respectively. The simulation results of I_D–V_G for various V_D levels in the programmed state are compared with the experimental data, as shown in Fig. 5(a).

In addition, all previous analyses are based on the forward read mode, in which V_D is applied as a positive voltage and substrate voltage V_S is grounded. The parallel capacitor model for the reverse read mode is still valid. Assuming that a positive voltage applied to V_S and V_D is grounded, (9) can be revised as follows:

$$\begin{aligned} \overline{\sigma}_i &= \frac{1}{V_S}\int_{2\varphi_F+V_S}^{2\varphi_F} \left[V_G - v_{\mathrm{th}}(\varphi_s)\right] C_{\mathrm{ox}} d\varphi_s \\ &= \frac{1}{V_S}\left\{ \int_{2\varphi_F+V_S}^{2\varphi_F+(l/L_{\mathrm{eff}})V_S} \left[V_G - v_{\mathrm{th(fresh)}}\right] C_{\mathrm{ox(fresh)}} d\varphi_s \right. \\ &\quad \left. + \int_{2\varphi_F+(l/L_{\mathrm{eff}})V_S}^{2\varphi_F} \left[V_G - v_{\mathrm{th(prog)}2}\right] C_{\mathrm{ox(prog)}2} d\varphi_s \right\}. \end{aligned} \tag{11}$$

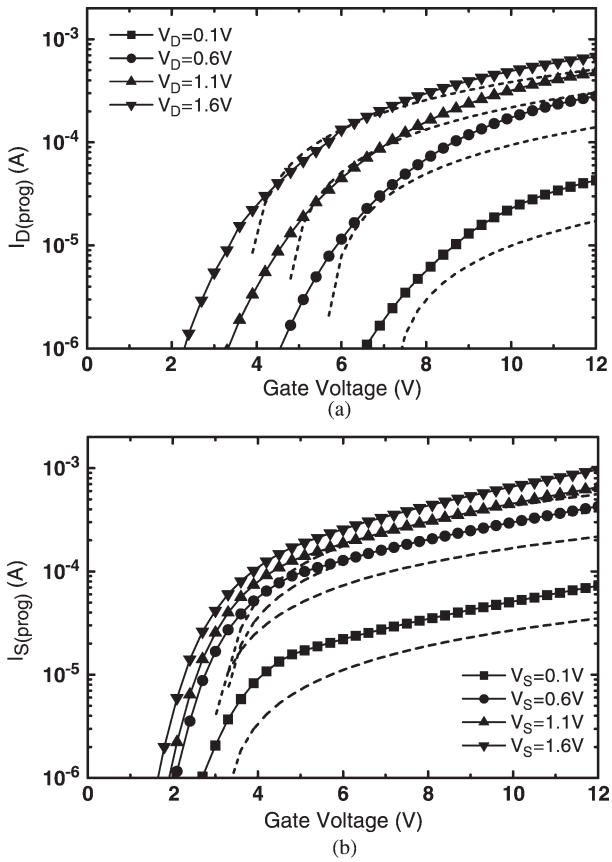


Fig. 5. Comparison of the experimental (solid symbol) and simulated (dashed line) channel current in the programmed state under (a) the forward read mode for various V_D levels and (b) the reverse read mode for various V_S levels.

The channel current in the reverse read mode can be expressed as

$$I_{S(\mathrm{prog})} = \frac{\mu_0 W}{L_{\mathrm{eff}}} \overline{\sigma}_i V_S \tag{12}$$

and the simulation results of I_S–V_G for various V_S levels in the programmed state are compared with the experimental data, as shown in Fig. 5(b). Compared to the variation of $V_{\mathrm{th(prog)}}$ as a function of V_D in the forward read mode, its variation as a function of V_S in the reverse read mode is not obvious since the charge potential barrier is located at the drain end where the depletion region of the source cannot reach and reduce it when V_S is not high enough [5].

In addition, our model follows the conventional theory of MOSFET devices, and therefore, a variety of techniques can be included to improve the simulation accuracy. For example, the estimated current density can be improved close to the experimental value by considering channel-length modulation, mobility reduction with gate/drain voltage, subthreshold current, fitting parameter, etc. [4], [13], which give rise to the improvement of the accuracy in future applications.

V. Application and Discussion

Our model gives a comprehensive analysis of SONOS devices, particularly on the relations between the trapped charge density and the electric field, oxide capacitance, charge distribution, etc. For example, the effective trapped charge density is estimated based on the threshold voltage shift [see (7)], and this provides the essential information for the control of the W/E operation. Moreover, compared to the fresh state, the oxide capacitance is reduced in the region where the trapped charge occupied (from $L_{\mathrm{eff}} - l$ to L_{eff}), resulting in an insufficient charge density at the substrate surface to form the conducting channel and requiring additional ΔV_{th} for compensation. These relations have been clearly provided in our model [see (3)–(7)], and their applications are discussed as follows.

First, since these relations are dependent on the geometry of the oxide stack and the magnitude of ΔV_{th}, the information of the electric field, oxide capacitance, and charge distribution can be used for optimizing the device structure and the W/E condition. For example, $E_{(\mathrm{tunnel})}$ that directly relates to E_{th} is determined by the geometry of the oxide stack [see (6)], and therefore, the required trapped charge density to achieve E_{th} can be estimated for a specific device structure.

Second, the retention characteristics can be performed based on our model. For example, the influence from the reduction of the trapped charge density $(\overline{\sigma}_t)$ can be observed in terms of the variation of the electric field, oxide capacitance, and charge distribution [see (3)–(7)]. In addition, the influence from the shrinkage of the occupied region l can be estimated by the variation of the channel current [see (9)–(12)].

Third, all analyses in our model so far are based on the assumption that the trapped charge is located at the bottom of the trapping layer. The vertical location of $\overline{\sigma}_t$ is negligible because the trapping layer is quite thin ($\sim$1 nm of EOT). Assuming an opposite condition that $\overline{\sigma}_t$ is located at the top interface, all the equations still work by replacing d with $d + 1$ in nanometer units, and their corresponding simulation results, which are not shown here, are quite similar to the results shown in Fig. 5(a) and (b). Therefore, the parallel capacitor model could be modified and applied to other SONOS devices using different trapping materials, e.g., Si_3N_4, where the trapped charge has been reported to be located at the top SiO_2/Si_3N_4 interface [5].

Finally, the parallel capacitor model can be also used to model the split-gate SONOS device [14]. Because the trapped charges are separately located at the drain and source ends, the cell is equivalently modeled as a parallel connection of three capacitors, i.e., two floating-gate MOS capacitors at the channel ends and one MOS capacitor in the middle. Their electric properties can be individually obtained in the corresponding capacitor, and their combinations are used to model the electrical performance of the entire device, such as the channel current in both read modes.

VI. Conclusion

In this paper, the parallel capacitor model has been proposed as a quite simple and time-saving approach to comprehensively analyze the electric properties of the SONOS memory device, including the charge density at the gate, trapping layer and substrate, electric field through the block and tunnel oxides, oxide capacitance in both fresh and programmed states, as well as channel current in both forward and reverse read modes. Our

model can be applied for the control of the W/E operation, the optimization of the device structure, and the characterization of the retention properties in SONOS Flash memory.

REFERENCES

[1] E. Lusky, Y. Shacham-Diamand, G. Mitenberg, A. Shappir, I. Bloom, and B. Eitan, "Investigation of channel hot electron injection by localized charge-trapping nonvolatile memory devices," *IEEE Trans. Electron Devices*, vol. 51, no. 3, pp. 444–451, Mar. 2004.

[2] T. Y. Chan, K. K. Young, and C. Hu, "A true single-transistor oxide–nitride–oxide EEPROM device," *IEEE Electron Device Lett.*, vol. EDL-8, no. 3, pp. 93–95, Mar. 1987.

[3] G. Zhang, W. J. Yoo, and C. H. Ling, "Hot-electron capture for CHEI programming in SONOS-type flash memory using high- *k* trapping layer," *IEEE Trans. Electron Devices*, vol. 55, no. 6, pp. 1502–1510, Jun. 2008.

[4] L. Perniola, S. Bernardini, G. Iannaccone, P. Masson, B. D. Salvo, G. Ghibaudo, and C. Gerardi, "Analytical model of the effects of a nonuniform distribution of stored charge on the electrical characteristics of discrete-trap nonvolatile memories," *IEEE Trans. Nanotechnol.*, vol. 4, no. 3, pp. 360–368, May 2005.

[5] P. B. Kumar, P. R. Nair, R. Sharma, S. Kamohara, and S. Mahapatra, "Lateral profiling of trapped charge in SONOS Flash EEPROMs programmed using CHE injection," *IEEE Trans. Electron Devices*, vol. 53, no. 4, pp. 698–705, Apr. 2006.

[6] H. T. Lue, Y. H. Shih, K. Y. Hsieh, R. Liu, and C. Y. Lu, "A transient analysis method to characterize the trap vertical location in nitride-trapping devices," *IEEE Electron Device Lett.*, vol. 25, no. 12, pp. 816–818, Dec. 2004.

[7] L. Larcher, G. Verzellesi, P. Pavan, E. Lusky, I. Bloom, and B. Eitan, "Impact of programming charge distribution on threshold voltage and subthreshold slope of NROM memory cells," *IEEE Trans. Electron Devices*, vol. 49, no. 11, pp. 1939–1946, Nov. 2002.

[8] E. Lusky, Y. Shacham-Diamand, I. Bloom, and B. Eitan, "Characterization of channel hot electron injection by the subthreshold slope of NROMTM device," *IEEE Electron Device Lett.*, vol. 22, no. 11, pp. 556–558, Nov. 2001.

[9] Y. J. Seo, K. C. Kim, T. G. Kim, Y. M. Sung, H. Y. Cho, M. S. Joo, and S. H. Pyi, "Analysis of electronic memory traps in the oxide-nitride-oxide structure of a polysilicon–oxide–nitride–oxide–semiconductor flash memory," *Appl. Phys. Lett.*, vol. 92, no. 13, p. 132 104, Mar. 2008.

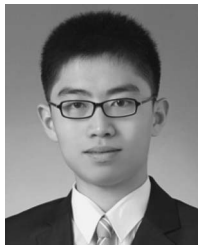
[10] E. Lusky, Y. Shacham-Diamand, I. Bloom, and B. Eitan, "Electrons retention model for localized charge in oxide–nitride–oxide (ONO) dielectric," *IEEE Electron Device Lett.*, vol. 23, no. 9, pp. 556–558, Sep. 2002.

[11] P. Pavan, R. Bez, P. Olivo, and E. Zanoni, "Flash memory cells—An overview," *Proc. IEEE*, vol. 85, no. 8, pp. 1248–1271, Aug. 1997.

[12] G. Zhang and W. J. Yoo, "$V_{\rm th}$ control by complementary hot-carrier injection for SONOS multi-level cell flash memory," *IEEE Trans. Electron Devices*, vol. 56, no. 12, pp. 3027–3032, Dec. 2009.

[13] S. Dimitrijev, *Understanding of Semiconductor Devices*. New York: Oxford Univ. Press, 2000.

[14] G. Zhang, S.-H. Lee, C. H. Ra, H.-M. Li, and W. J. Yoo, "Second-bit-effect-free multibit-cell flash memory using Si_3N_4/ZrO_2 split charge trapping layer," *IEEE Trans. Electron Devices*, vol. 56, no. 9, pp. 1966–1973, Sep. 2009.



Hua-Min Li received the B.S. degree from Shandong Normal University, Jinan, China, in 2007 and the M.S. degree from Sungkyunkwan University (SKKU), Suwon, Korea, in 2010. He is currently working toward the Ph.D. degree with the SKKU Advanced Institute of Nanotechnology, SKKU.

His research interests include the electrical characterization of traps in nonvolatile memory devices and the Si- and graphene-based high-performance optoelectronic devices.

Gang Zhang (S'07) received the B.Eng. (with honors) degree in electrical and computer engineering from the National University of Singapore, Singapore, in 2005 and the M.Eng. and Ph.D. degrees in nanoelectronics from Sungkyunkwan University (SKKU), Suwon, Korea, in 2008 and 2011, respectively.

From 2009 to 2011, He was a Research Assistant with the Electronic Materials Research Center, Korea Institute of Science and Technology, Seoul, Korea. He is currently with Samsung Electronics, Hwasung, Korea. His research interests include high-k dielectric materials, and charge-trapping/phase-change nonvolatile memory.

Dr. Zhang was the recipient of the 2006 Best Student Paper Award at the Eighth International Conference on Solid-State and Integrated-Circuit Technology, Shanghai, China; the 2008 Gold Paper Award at the IEEE Student Paper Contest, Seoul, Korea; and the 2009 Chinese Government Award for Outstanding Self-Financed Student Abroad.

Won Jong Yoo (M'00–SM'02) received the B.S. and M.S. degrees from Seoul National University, Seoul, Korea, and the Ph.D. degree in the area of the plasma etching properties of Si and SiO_2 from Rensselaer Polytechnic Institute, Troy, NY, in 1993.

He was an Associate Professor with the National University of Singapore, Singapore, where he conducted his research in the areas of silicon device and plasma processing. Since 2006, he has been a Professor with Sungkyunkwan University (SKKU) Advanced Institute of Nanotechnology, SKKU, Suwon, Korea. His main industrial experiences were research and development in the areas of semiconductor material/device processes with Samsung Semiconductor Research Center, Kiheung, Korea, and IBM Research Center, Yorktown Heights, NY. He is currently working as the Director of the Samsung–SKKU Graphene Center, SKKU. He is the author or coauthor of about 160 peer-reviewed journal and conference papers. His current research interests include the investigation of the materials and electrical properties of graphene devices and nonvolatile memory devices using high-dielectric-constant materials, and the investigation of plasma etching processes of these devices.