Modified Potential Well Formed by $Si/SiO_2/TiN/TiO_2/SiO_2/TaN$ for Flash Memory Application

Gang Zhang, *Student Member, IEEE*, Chang Ho Ra, Hua-Min Li, Tian-zi Shen, Byung-ki Cheong, and Won Jong Yoo, *Senior Member, IEEE*

***Abstract*—This paper proposes a modified engineered-potential-well (MW) for NAND flash memory application. The MW was formed by using a transitional SiO_2/SiO_xN_y-TiO_xN_y tunnel barrier, a trap-rich TiO_2 trapping layer, and an abrupt SiO_2 block barrier. The transitional tunnel barrier shrinks to enhance the tunneling of carriers during programming/erasing (P/E) and extends to suppress charge loss during data retention. Deep-level transient spectroscopy suggests that this tunnel barrier has few shallow traps after a $N_2 + O_2$ thermal treatment, and the TiO_2 trapping layer has deep electron traps. With the variable tunnel barrier and deep electron traps, the MW device showed promising performance in fast programming ($< \mu$s) at low-voltage operation (7–10 MV/cm), good P/E endurance ($> 10^6$ P/E cycles), large threshold voltage window ($\Delta V_{\rm th} = \sim$6 V), as well as improved data retention at 125 °C.**

***Index Terms*—Flash memory, modified engineered-potential-well (MW), TiO_2 trapping layer.**

I. Introduction

In THE development of a flash memory, operating speed, memory window, and data retention are the important parameters. Various designs, such as polysilicon-oxide-nitride-oxide-semiconductor (SONOS) devices using a high-dielectric-permittivity (k) trapping layer [1], a high-k block barrier and metal gate [2], a band-engineered (BE) tunnel barrier [3]–[5], and metal nanocrystal trapping layers [6]–[8], have been proposed to demonstrate fast speed and good data retention. Recently, an engineered-potential-well (EW) has also been proposed for the flash memory application [9]. Memory devices with an EW have transitional boundaries between the tunnel/block barriers and deep traps at the bottom of the EW. During programming/erasing (P/E), the EW is bent by a gate electric field $(E_{\rm ox})$, and the direct tunneling (DT) of carriers via the shrunk tunnel barrier is induced to enable fast P/E. However, under the retention mode, the EW is transformed due to charge trapping, i.e., the trapped electrons raise the bottom of the EW, and the enlarged tunnel barrier suppresses the reverse tunneling of the detrapped electrons, resulting in good data retention. The EW switches the retention and the P/E modes effectively, and it is insensitive to the P/E cycling-induced tunnel barrier degradation. Therefore, the EW accomplishes promising memory performance with fast speed at low-voltage operation with excellent endurance [9]. Nevertheless, the EW formed by the intermixing of TiN and the tunnel/block oxide (SiO_2) during rapid thermal process (RTP) suffers from a small $\Delta V_{\rm th}$ due to the finite trap density, a large $V_{\rm th}$ dispersion ($\sim$1 V), and the degradation of data retention due to the shallow traps at the transitional tunnel barrier.

In this paper, a modified-EW (MW) was proposed to improve the electrical performance. The MW was formed using a transitional SiO_2/SiO_xN_y-TiO_xN_y tunnel barrier, a trap-rich TiO_2 trapping layer, and an abrupt SiO_2 block barrier, as shown in Fig. 1(a). The transitional tunnel barrier was preformed by the intermixing of SiO_2 and TiN during 900 °C RTP, followed by a low-pressure thermal treatment at 550 °C in a $N_2 + O_2$ ambient. Deep-level transient spectroscopy (DLTS) results suggested that the thermal treatment effectively removes the shallow traps at the boundary of the tunnel barrier, and that the TiO_2 trapping layer has very deep electron traps. On the other hand, the TiO_2 trapping layer has proper conduction $(\Delta\varphi_c)$ and valence $(\Delta\varphi_v)$ band offsets to Si of 1.2 eV for both, to form the symmetric band structure of a MW [10]. The MW adopts the promising performance of EW in fast programming ($< \mu$s) at low-voltage operation (7–10 MV/cm) and good endurance ($> 10^6$ P/E cycles). Moreover, the MW shows an enlarged $\Delta V_{\rm th}$ ($\sim$6 V) and improved 10-year data retention at 125 °C. The experimental results suggested that the MW with a 3-nm-thick tunnel SiO_2 exhibits optimum performance, and the $\Delta V_{\rm th}$ of the MW can be further enlarged using a thicker TiO_2 trapping layer.

II. Device Design and Fabrication

The transitional tunnel barrier of a MW was formed using the same process as for an EW formation [9]. Memory devices

Manuscript received March 3, 2010; revised July 12, 2010; accepted July 16, 2010. Date of publication September 13, 2010; date of current version November 5, 2010. The 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 (2010-0015035). The review of this paper was arranged by Editor D. Esseni.

G. Zhang and C. H. Ra are with the Department of Nanoscience and Technology, SKKU Advanced Institute of Nano Technology, Sungkyunkwan University, Suwon 440-746, South Korea, and also with the Electronic Materials Research Center, Korea Institute of Science and Technology, Seoul 136-791, South Korea.

H.-M. Li, T.-z. Shen, and W. J. Yoo are with the Department of Nanoscience and Technology, SKKU Advanced Institute of Nano Technology, Sungkyunkwan University, Suwon 440-746, South Korea (e-mail: yoowj@skku.edu).

B.-k. Cheong is with the Electronic Materials Research Center, Korea Institute of Science and Technology, Seoul 136-791, South Korea.

Digital Object Identifier 10.1109/TED.2010.2066200

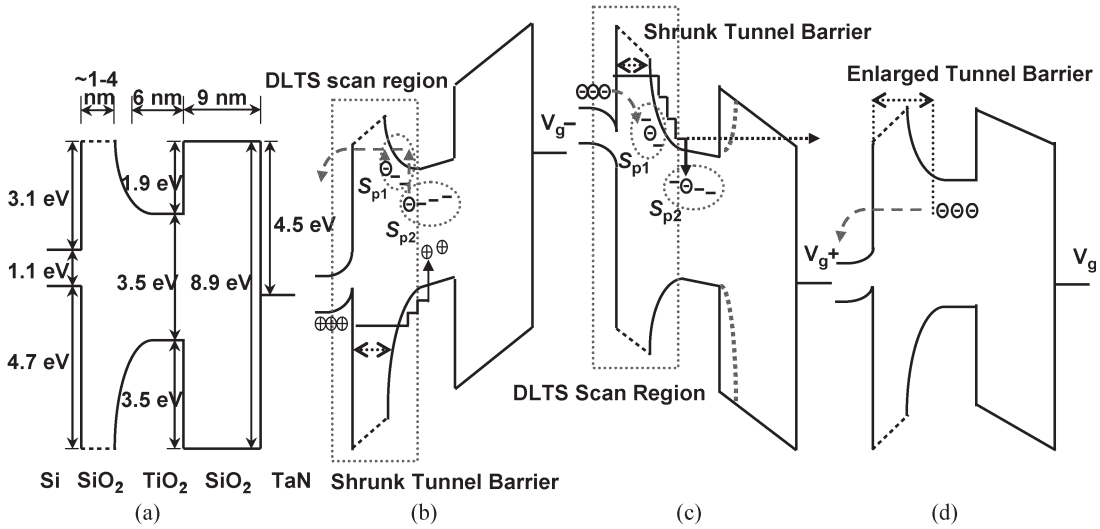


Fig. 1. Band diagrams of a MW (a) under flatband, (b) during erasing, (c) during programming, and (d) during data retention.

with a Si/SiO_2/TiN/TiO_2/SiO_2/TaN (SOTTOT: –/2–5/3/6/9/150-nm) gate stack were fabricated. An approximately 2–5-nm-thick $(D_{\rm ox})$ tunnel oxide was grown thermally at 875 °C on a p-type silicon wafer, and a ∼3-nm-thick TiN layer was then deposited on SiO_2 by electron-beam evaporation using a TiN target at 600 °C, followed by RTP at 900 °C under vacuum for 10 s. The thin TiN film deposited at low temperature is structurally irregular and thermodynamically not in equilibrium [11]. It interacts with SiO_2 [12] to form ∼3-nm-thick TiO_xN_y and ∼1-nm-thick SiO_xN_y transition layers during 900 °C RTP [9]. Accordingly, the remaining SiO_2 thicknesses would be ∼1–4 nm after the formation of the MW, as shown in Fig. 1(a). Note that this thin TiN film was only exposed to O_2 during RTP, and this eliminates the possibility of oxidation. A thermal treatment at 550 °C in an $85\%N_2 + 15\%O_2$ ambient at 130 mTorr was performed for 110 s. The preformed SiO_2/SiO_xN_y-TiO_xN_y tunnel barrier was then capped with a 6-nm-thick TiO_2 trapping layer formed by sputtering using a Ti target in an $Ar + O_2$ atmosphere, and a 9-nm-thick block oxide (SiO_2) was formed by low-pressure chemical vapor deposition. A ∼150-nm-thick TaN layer was deposited and patterned as the gate electrode. Postdeposition annealing was then performed at 900 °C under vacuum for 10 s, followed by forming gas annealing at 420 °C for 30 min. The reference devices with a SOTTOT: –/2/3/6/9/150-nm gate stack were fabricated using the same process, but on an n-type Si substrate, and with/without a 550 °C thermal treatment in the $N_2 + O_2$ ambient for an electron trap investigation. The equivalent oxide thickness (EOT) of the memory capacitor with a SOTTOT: –/3/3/6/9/150-nm gate stack was found to be about 13 nm by using the capacitance–voltage $(C$–$V)$ method, and the $\Delta V_{\rm th}$ was extracted by monitoring the C–V hysteresis during P/E.

It was expected that the low-pressure 550 °C thermal treatment in the $N_2 + O_2$ ambient could reduce the oxygen and nitrogen vacancies in the SiO_xN_y and TiO_xN_y intermixing layers [12] so as to remove the SiO_xN_y shallow traps and certain TiO_xN_y traps in the transitional SiO_2/SiO_xN_y-TiO_xN_y tunnel barrier [9]. The loosely bound TiN was converted to the oxidized states of Ti^{3+} and Ti^{4+}. The generated oxides carry a high density of oxygen vacancies, which act as deep traps [10]. The reaction of $TiN + O_2 \rightarrow TiO_2 + 1/2N_2$ is thermodynamically favorable with a Gibb's free energy of −138.7 kcal/mole [12]. Although the traps of TiN and TiO_xN_y in the EW are the desired trapping sites that result in good data retention [9], the ultrathin TiO_xN_y film formed by the intermixing of TiN and SiO_2 is unlikely to provide sufficient charge traps to produce to a large $\Delta V_{\rm th}$, i.e., a typical $\Delta V_{\rm th}$ of the EW saturates at ∼3 V [9]. In order to enhance the charge-trapping capacity of the MW, a trap-rich TiO_2 trapping layer was adopted. TiO_2 has a high-k value of 80–100, a good vertical scalability, a proper band structure with respect to Si, and a high density of traps [10]. The flatband diagram of a MW was estimated, as shown in Fig. 1(a) [10].

The electron traps of the MW were examined by DLTS. In the extraction of electron traps using DLTS, a pulse bias was applied to the reference samples to induce the capacitance-time $(C - t)$ transients. By analyzing the $C - t$ transients, the activation energy of the deep traps can be estimated [13]. First, a quiescent bias (−1.5 V) was applied for 180 s to retain the quiescent condition of the substrate and the DLTS scan region, as shown in (dashed arrow) Fig. 1(b). Deep traps in the DLTS scan region were discharged to eliminate the disturbances induced by the precharged traps. Subsequently, a majority carrier (electron) driven bias (1.7 V) was applied for 90 s to drive the n-type Si substrate into accumulation, and the deep traps in the DLTS scan region were filled with electrons, as shown in (dashed arrow) Fig. 1(c). An emptying bias $(-2.0- - 0.5$ V) was then applied to empty the deep traps. The discharging of the deep traps at $(T) = -200$ °C−350 °C induces the $C - t$ transients, as shown in (dashed arrow) Fig. 1(b) [13]. It shall be noted that the ultrathin SiO_2 (∼1-nm-thick) in the reference device is designed to be a buffer layer rather than a tunnel barrier. Thus, the principle of DLTS is applicable to the reference sample, despite the disturbance induced by the DT and the Fowler–Nordheim (F–N) tunneling of the detrapped electrons [14]. Fig. 2(a) and (b) show the DLTS signals $[S = (C_{t1} - C_{t2})/\Delta C_o]$ obtained for the reference devices before and after the 550 °C thermal

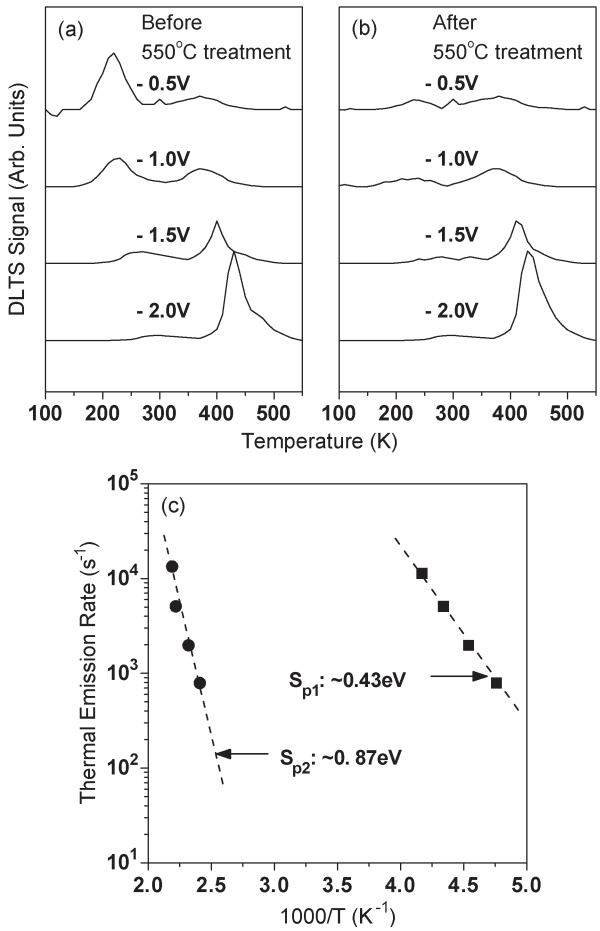


Fig. 2. DLTS signals obtained for the reference MW devices (a) before and (b) after the low-pressure thermal treatment and (c) Arrhenius plots of the electron thermal emission rates versus the reversed temperature.

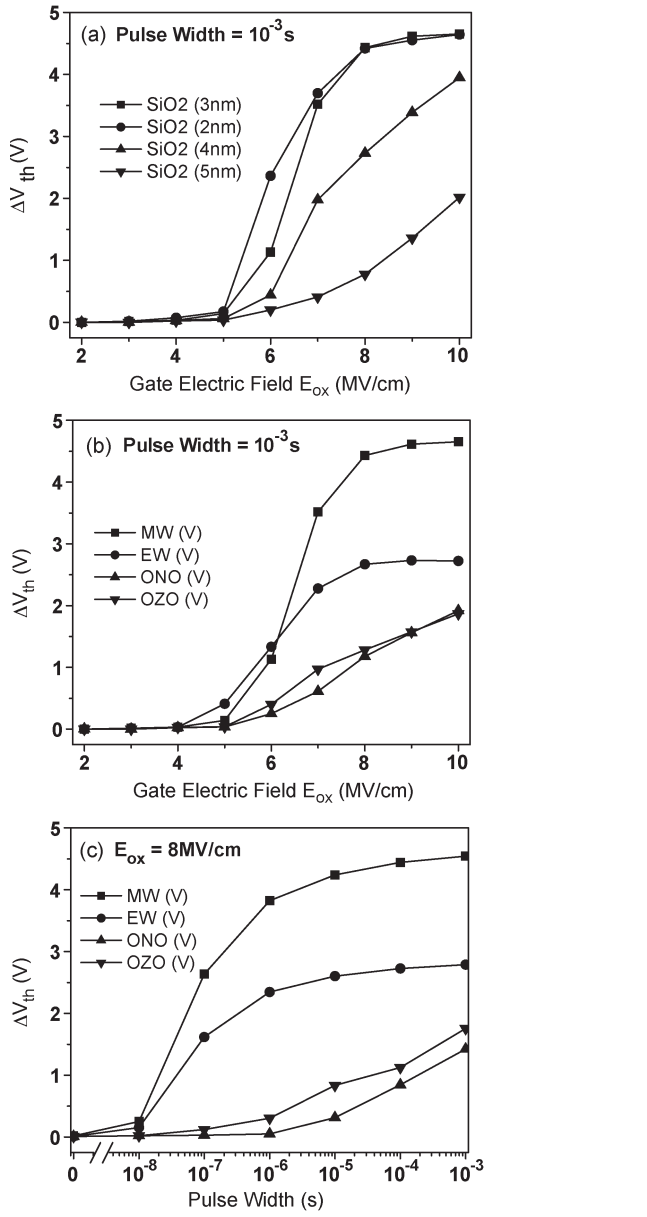


Fig. 3. (a) $\Delta V_{\rm th}$ transients of the MW devices with a 2–5-nm-thick tunnel SiO_2 programmed under a gate electric field of 4–10 MV/cm for 10^{-3} s, (b) $\Delta V_{\rm th}$ transients of memory devices with a MW, an EW, and a ONO/OZO gate stack programmed under a gate electric field of 4–10 MV/cm for 10^{-3} s, and (c) $\Delta V_{\rm th}$ transients of these memory devices during programming under a gate electric field of 8 MV/cm.

treatment, respectively, where ΔC_o is the change in total capacitance induced by the majority carrier driven bias, and the thermal emission rate $e(T)$ is obtained by $\ln(t_1/t_2)/(t_1-t_2)$ at a temperature that yields the S peaks [13]. Two S peaks were obtained at $\sim$230 K (S_{p1}) and $\sim$430 K (S_{p2}) before the thermal treatment. In contrast, only S_{p2} was observed at $\sim$420–440 K after the thermal treatment. Fig. 2(c) presents the Arrhenius plots of $e(T)$ versus $1000/T$, and the slopes of the $\ln[e(T)/T^2]$ versus $1/T$ plots indicated an activation energy of $\sim$0.43 and $\sim$0.87 eV for S_{p1} and S_{p2}, respectively [13]. A comparison of Fig. 2(a) and (b) indicates that shallow traps have been removed by the thermal treatment. S_{p1} was obtained at emptying biases of $-0.5--1.0$ V, whereas S_{p2} was observed at emptying biases of $-1.0--2.0$ V. This suggested that the traps at S_{p1} were discharged without being negatively biased due to the rapid retention decay at 75 °C–150 °C [9]. In contrast, the traps at S_{p2} were discharged only when they were negatively biased. The traps at S_{p2} were attributed to steady data storage. However, the density and spatial distributions of the traps at S_{p1} and S_{p2} were difficult to be extracted using the DLTS method due to the low accuracy [14]. On the other hand, the hole traps were not investigated in this paper because holes play a minor role in the programming and data retention properties of the MW device.

III. Electrical Performance

A. P/E Property

Fig. 1(c) presents the band diagram of the MW during programming (solid arrow). The DT of electrons via the shrunk tunnel barrier occurs during programming. The enhanced tunneling of electrons enables rapid programming at low-voltage operation for memory devices with a MW compared to the SONOS device, which has F–N tunneling of electrons during programming [1]. Fig. 3(a) shows the $\Delta V_{\rm th}$ transients of the MW with various tunnel SiO_2 thicknesses ($D_{\rm ox}=2-5$ nm) under $E_{\rm ox}=4-10$ MV/cm at 10^{-3} s. Programming was enabled at $E_{\rm ox}=5-6$ MV/cm when $D_{\rm ox}=2$ nm,

TABLE I
COMPARISON OF THIS WORK TO OTHER WORKS

References	This work	[9] *IEDM* 2009 pp. 835-838	[4] *IEDM* 2007 pp. 75-78	[5] *IEDM* 2005 pp. 547-550	[2] *IEDM* 2005 pp. 327-330	[6] *IEDM* 2005 pp. 177-180
Device structure	SOTTOT	SOTOT	SOSONOS	SONONOS	TANOS	W nanocrystals
Layer Thicknesses (nm)	-/3/3/6/9/-	-/3/6/9/-	-/1/1.2/1/4/6.5/-	-/1.5/2/1.8/7/9/-	-/15/6.5/4/-	-
P/E electrical filed (MV/cm)	7-10	6-8	6.9-7.7	10-12	-	-
Prog. durations for ΔV_{th}=2V	~7×10^{-8}s	3×10^{-7}s	~2×10^{-4}s	~7×10^{-5}s	~2×10^{-6}s	~2×10^{-1}s
Erase durations for ΔV_{th}=2V	~1×10^{-4}s	~2×10^{-4}s	~2×10^{-4}s	~10^{-2}s	~1×10^{-4}s	~1 s
ΔV_{th}(V)	~6	~3	~1.8	~6	>6	~2
Endurance (P/E cycles)	$>10^6$ *	$>10^7$	$>10^5$	$>10^4$	$>10^4$	$>10^5$
10-year ΔV_{th} decay (%)	~35% at 125°C (post 10^6 cycles)	52% at 125°C (post 10^7 cycles)	50% at 125°C (before cycling)	~25% at 150°C (post 10^4 cycles)	(~0.6V after 1 hour @ 200°C)	~50% at 25°C (before cycling)

* Obtained from capacitors

under which condition the conduction band (φ_c) of the TiO_2 trapping layer was lowered by $E_{ox} \times D_{ox} \approx 1.2$ eV. On the other hand, $\Delta\varphi_c$ between the Si substrate and the TiO_2 trapping layer was 1.2 eV, as shown in Fig. 1(a). When φ_c of the TiO_2 trapping layer was lowered than φ_c of Si, the DT of the Si substrate electrons via a ~2-nm-thick tunnel barrier occurred to enable programming. Similarly, programming was enabled at $E_{ox} = 7-8$ MV/cm when $D_{ox} = 3$ nm. It is understood that modified F–N tunneling dominates the programming at $E_{ox} = 5-6$ MV/cm due to the enlarged tunnel barrier (> 3 nm) near the bottom of the MW [3], and the DT is due to the rapid programming at $E_{ox} = 7-8$ MV/cm when the MW is significantly bent, as shown in Fig. 1(c). In contrast, it is difficult to achieve rapid programming when $D_{ox} = 4-5$ nm, because DT via such a thick SiO_2 is disabled, while F–N tunneling occurs when the MW is significantly bent. Fig. 3(b) shows the ΔV_{th} transients of the memory devices with a MW ($D_{ox} = 3$ nm) and an EW (SOTOT: –/3/6/9/150 nm) [9] and the SONOS devices (ONO/OZO) with a 3-nm-thick tunnel SiO_2 and a Si_3N_4 or ZrO_2 trapping layer under $E_{ox} = 4-10$ MV/cm at 10^{-3} s. The ΔV_{th} of the MW, EW, ONO, and OZO devices were 4.65, 2.72, 1.92, and 1.87 V at 10^{-3} s under $E_{ox} =$ 10 MV/cm at 10^{-3} s, respectively. Programming was enabled by $E_{ox} = 7-10$ MV/cm for the MW devices, while it was enabled by $E_{ox} > 10$ MV/cm for the ONO and OZO devices. Therefore, the MW device shows promising capability for low-voltage operation. Fig. 3(c) presents the ΔV_{th} transients of the various devices under $E_{ox} = 8$ MV/cm for $10^{-9}-10^{-3}$ s, where $\Delta V_{th} = 3.82$, 2.35, 0.06, and 0.31 V at 10^{-6} s for the MW, EW, ONO, and OZO devices, respectively. The ΔV_{th} of the MW and EW devices saturates at 4.65 and 2.72 V under $E_{ox} = 8$ MV/cm at 10^{-4} s, respectively. In contrast, the ONO and OZO devices were underprogrammed under the same condition. It is understood that the enlarged ΔV_{th} for the MW device is attributed from the additional TiO_2 trapping layer and the abrupt block barrier that enhances the blocking effect of the MW compared to that of the EW device that has a transitional block barrier, as shown in (dotted arrow and line) Fig. 1(c). As a result, the MW device shows promising P/E properties, as shown in Fig. 3(b) and (c), and listed in Table I.

Fig. 4 shows the ΔV_{th} transients of a fresh MW device ($D_{ox} = 3$ nm) P/E by $E_{ox} = (\pm)6-8$ MV/cm for $10^{-9}-10^{-3}$ s, where $\Delta V_{th} = 1.14$, 3.33, and 4.54 V at 10^{-3} s when $E_{ox} = 6$, 7, and 8 MV/cm were applied for programming, respectively, and $\Delta V_{th} = -0.32, -0.70,$ and -1.18 V at 10^{-3} s when $E_{ox} = -6, -7,$ and -8 MV/cm were applied for erasing, respectively. The DLTS results suggest that the discharging of the TiO_2 traps occurs at ~125 °C. Thus, the detrapping of electrons is unlikely to occur at room temperature during erasing. On the other hand, the tunneling of holes via the shrunk tunnel barrier shall be responsible for erasing [8], as shown in (solid arrow) Fig. 1(b). Note that the MW has a symmetric band structure, where the $\Delta\varphi_v$ between the Si substrate and the TiO_2 trapping layer is also 1.2 eV. Erasing was performed at the same $|E_{ox}|$ as programming. However, the erasing speed under $|E_{ox}| = 6-8$ MV/cm was less than the programming speed. For example, $\Delta V_{th} = 3.82$ and -0.84 V at 10^{-6} s when $E_{ox} = \pm 8$ MV/cm were applied for P/E. This may be due to larger $\Delta\varphi_v$ (4.7 eV) than $\Delta\varphi_c$ (3.1 eV) between the Si substrate and the tunnel SiO_2. The programmed ΔV_{th} tends to saturate at about 4.5 V for $10^{-4}-10^{-1}$ s under $E_{ox} = 6-8$ MV/cm. This indicates that the trap density in TiO_2 may be on the order of $10^{17}-10^{18}$ cm^{-3} when extracted by using the C–V method [1]. This trap density is consistent with the results ($10^{17}-10^{19}$ cm^{-3}) reported for memory application [14]. In contrast, overerasure occurred over $10^{-4}-10^{-1}$ s, resulting in the erased ΔV_{th} of -6.36 V larger than the programmed ΔV_{th} of 4.58 V at 10^{-1} s, as shown in Fig. 4. It is understood that hole traps also exist in the TiO_2 trapping layer. Hole trapping occurred during erasing so as to induce overerasure. In NAND architecture, overerasure may be employed to enlarge the total memory window. Fig. 5 shows the ΔV_{th} transients of

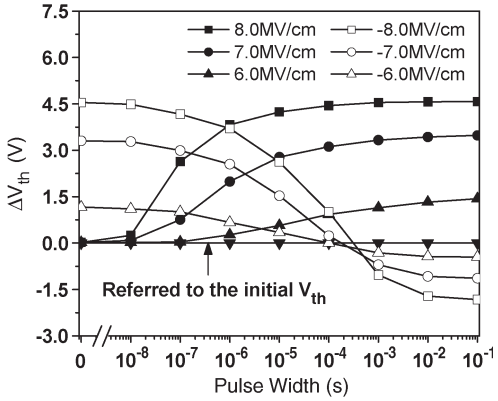


Fig. 4. ΔV_{th} transients of the MW device with a 3-nm-thick tunnel SiO_2 P/E by $E_{ox} = (\pm)6-8$ MV/cm for $10^{-9}-10^{-3}$ s.

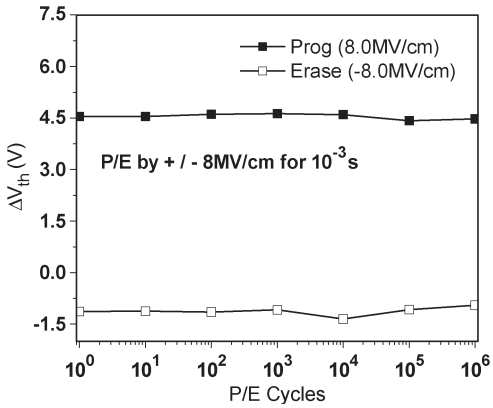


Fig. 5. $\Delta V_{\rm th}$ transients of the MW device with a 3-nm-thick tunnel SiO_2 P/E by $E_{\rm ox} = (\pm)8$ MV/cm for 10^{-3} s for 10^6 P/E cycles.

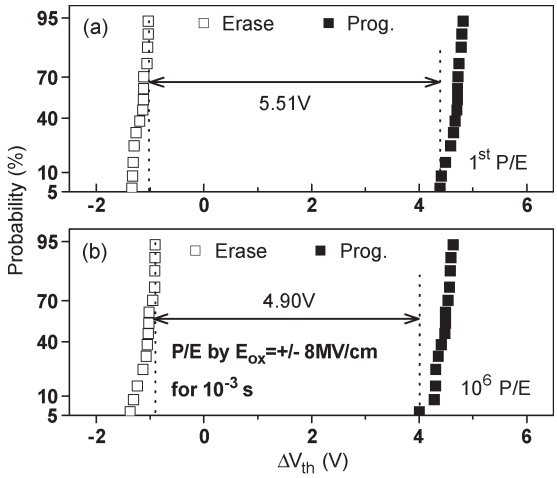


Fig. 6. Cumulative distribution of $\Delta V_{\rm th}$ of the MW devices with a 3-nm-thick tunnel SiO_2 before/after 10^6 P/E cycling.

the same MW device P/E by $E_{\rm ox} = (\pm)8$ MV/cm for 10^{-3} s over the 10^6 P/E cycles. The programmed $\Delta V_{\rm th}$ at the first and 10^6th P/E cycle were 4.54 and 4.47 V, respectively, and the erased $\Delta V_{\rm th}$ at the first and 10^6th P/E cycle were -1.13 and -0.94 V, respectively. The endurance reliability of the MW device was well maintained throughout the 10^6 P/E cycles.

Fig. 6 presents the cumulative distribution of $\Delta V_{\rm th}$ before and after a 10^6 P/E cycling for the MW devices with $D_{\rm ox} =$ 3 nm. The $\Delta V_{\rm th}$ dispersions were $\sim$0.42 and $\sim$0.30 V at the programmed/erased states, and the sensing window was $\sim$5.51 V at the first P/E cycle, as shown in Fig. 6(a). After the 10^6 P/E cycles, the $\Delta V_{\rm th}$ dispersions were $\sim$0.45 and $\sim$0.40 V at the programmed/erased states, respectively, and the sensing window was $\sim$4.90 V, as shown in Fig. 6(b). The $\Delta V_{\rm th}$ dispersions for the MW devices were suppressed compared to those of the EW devices before ($\sim$0.94 V) and after ($\sim$1.13 V) 10^7 P/E cycles [9]. It is believed that the 550 °C thermal treatment improves the quality of the tunnel barrier by removing the oxygen and nitrogen vacancies (shallow traps), giving rise to uniform performance of the MW devices. In contrast, the EW devices suffer from the large $V_{\rm th}$ dispersions due to its trap-rich tunnel barrier. The $\Delta V_{\rm th}$ dispersions were well controlled throughout the 10^6 P/E cycles. This suggests that the

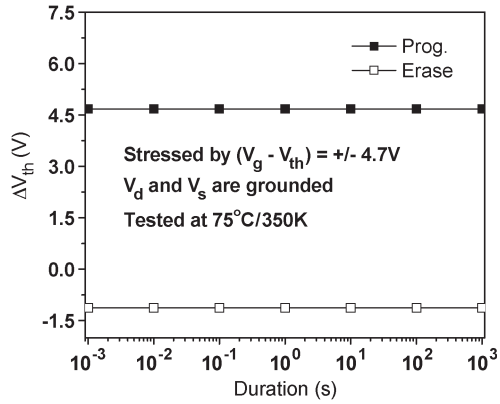


Fig. 7. $\Delta V_{\rm th}$ transients of the MW devices with a 3-nm-thick tunnel SiO_2 at the programmed/erased states disturbed by $E_{\rm ox} < \pm 5$ MV/cm over the 10^6 P/E cycles at 75 °C.

tunnel barrier is degraded insignificantly during 10^6 P/E cycling by $E_{\rm ox} = \pm 8$ MV/cm for 10^{-3} s. On the other hand, the read disturbance was negligible for the MW devices with $D_{\rm ox} =$ 3 nm at the erased/programmed states when they were stressed by $V_g - V_{\rm th} = \pm 4.8$ V ($E_{\rm ox} < \pm 5$ MV/cm) at 75 °C for 10^3 s, with the drain and source ($V_d = V_s$) grounded, as shown in Fig. 7. The $\Delta V_{\rm th}$ degradation between the programmed/erased states was < 0.02 V during the electric stress. The immunity to a read disturbance ensures the reliable circuit implementation of the MW device.

B. Data Retention

Fig. 1(d) shows the enlarged tunnel barrier of the MW under the retention mode, which was interpreted using the F–N model [9]. The trapped electrons stored in the deep electron traps of TiO_2 raise the bottom of the MW, enlarging the tunnel barrier. The experimental and calculation results showed that the tunnel barrier thickness of MW could be enlarged by $\sim$0.7–1.3 nm [9]. The enlarged tunnel barrier and deep electron traps of TiO_2 suppress the electron loss. Meanwhile, the internal $E_{\rm ox}$ across the tunnel barrier which is induced by electron trapping is less than $\Delta V_{\rm th}/{\rm EOT} \approx 4.6$ MV/cm[1], which is too small to induce F–N tunneling of holes from the Si substrate to the TiO_2 trapping layer. Thus, the MW enables good data retention. Fig. 8(a) presents the $\Delta V_{\rm th}$ transients of the MW devices with various $D_{\rm ox}$ of 2–5 nm at 75 °C. The programmed $\Delta V_{\rm th}$ decays by 0.41, 0.10, 0.06, and 0.05 V, and the erased $\Delta V_{\rm th}$ decays by 0.28, 0.20, 0.16, and 0.09 V when $D_{\rm ox} = 2-5$ nm after retention for 10^5 s. The extrapolated 10-year data retention suggests that $D_{\rm ox} = 3-5$ nm ensures good data retention at 75 °C. In contrast, $D_{\rm ox} = 2$ nm is too thin to retain good data retention, and the MW devices with $D_{\rm ox} = 2$ nm suffer from $\sim$50% $\Delta V_{\rm th}$ decay at 75 °C. Recall that rapid programming at low voltage is achievable when $D_{\rm ox} \leq 3$ nm. It is thus believed that the optimum performance of MW can be obtained when $D_{\rm ox} = 3$ nm when a single tunnel SiO_2 is employed. Fig. 8(b) shows the $\Delta V_{\rm th}$ transients of the MW devices with $D_{\rm ox} = 3$ nm at 25 °C–150 °C. The programmed $\Delta V_{\rm th}$ decays by 0.04, 0.10, 0.34, and 0.42 V, and the erased $\Delta V_{\rm th}$ decays by 0.15, 0.20, 0.27, and 0.28 V at 25 °C–150 °C after retention for

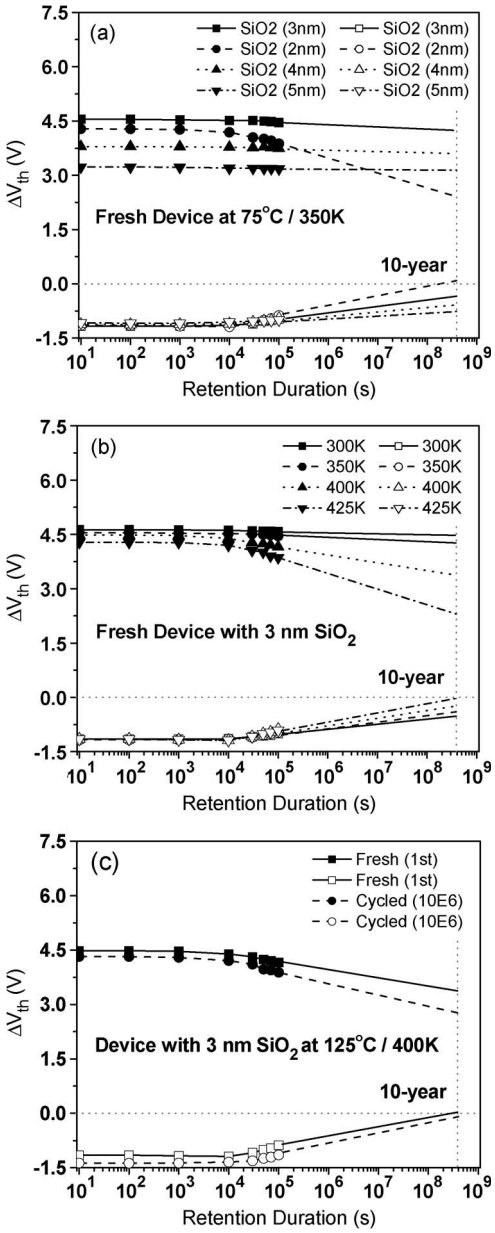


Fig. 8. (a) Data retention of the MW devices with a 2–5-nm-thick tunnel SiO_2 at 75 °C, (b) data retention of the MW device with a 3-nm-thick tunnel SiO_2 at 25 °C–150 °C, and (c) data retention of this device at 125 °C before and after 10^6 P/E cycling.

10^5 s. The extrapolated 10-year data retention suggests that the MW devices suffer from $<10\%$, $\sim 19\%$, and $\sim 50\%$ $V_{\rm th}$ decay at ≤ 75, 125 °C, and 150 °C, respectively. Significant electron detrapping occurs at $\sim$ 150 °C, which is consistent with the DLTS results. Therefore, the deep traps of TiO_2 shall be responsible for the good data retention. The good data retention was maintained throughout the 10^6 P/E cycles at 125 °C, as shown in Fig. 8(c). The programmed and erased $\Delta V_{\rm th}$ of the cycled device decayed by 0.44 and 0.23 V, respectively, after 10^5 s, and the extrapolated 10-year data retention retained $\sim 65\%$ of the $\Delta V_{\rm th}$ window. The insignificant electric stress on the tunnel barrier of the MW devices does not result in the development of defects. Thus, the MW devices show good endurance and data retention [9].

It should be noted that the activation energy of the electron traps in TiO_2 ($\sim$0.87 eV) is larger than that of the electron traps in bulk Si_3N_4 (0.25–0.45 eV) [14] when the TiO_2 and Si_3N_4 devices are investigated by using the same DLTS method. It is understood that the larger activation energy suppresses the electron emission more effectively, especially during the high-temperature baking. Thus, it is reasonable to assume that the larger activation energy is associated with deeper electron traps, although the correlation between the activation energy and the trap depth is complicated. In this regard, we think that the TiO_2 trapping layer in the MW device has deeper electron traps than the Si_3N_4 trapping layer in the SONOS devices. As a consequence, the data retention of the MW devices can be superior to that of the SONOS devices, as shown in Table I. Similar results have been reported by the other group [15]. Recall the diminishing DLTS spectra S_{p1} (which indicate the activation energy of $\sim$0.43 eV) during the thermal treatment in Fig. 2. We think that the shallow traps have been removed from the MW stack, and only the deep traps determine the data retention of the MW devices. Moreover, it is believed that BE technology can be used to engineer the SiO_2 in the SiO_2/SiO_xN_y-TiO_xN_y tunnel barrier of the MW to further enhance the vertical scalability of the MW device.

IV. Conclusion

A MW has been proposed for the NAND Flash memory. The MW has a transitional tunnel barrier, a trap-rich TiO_2 trapping layer with deep electron traps, and an abrupt block barrier. The MW devices exhibited promising performance with rapid programming at low-voltage operation, a large $\Delta V_{\rm th}$ window, good endurance, and improved data retention at 125 °C.

References

[1] G. Zhang, X.-P. Wang, W. J. Yoo, and M.-F. Li, "Spatial distribution of charge traps in a SONOS-type Flash memory using a high-k trapping layer," *IEEE Trans. Electron Devices*, vol. 54, no. 12, pp. 3317–3324, Dec. 2007.

[2] Y. C. Shin, J. D. Choi, C. S. Kang, C. H. Lee, K.-T. Park, J.-S. Lee, J. S. Sel, V. Kim, B. I. Choi, J. S. Sim, D. C. Kim, H.-J. Cho, and K. N. Kim, "A novel NAND-type MONOS memory using 63 nm process technology for multi-gigabit Flash EEPROMs," in *IEDM Tech. Dig.*, 2005, pp. 327–330.

[3] Y. Q. Wang, W. S. Hwang, G. Zhang, G. Samudra, Y.-C. Yeo, and W. J. Yoo, "Electrical characteristics of memory devices with a high-k HfO_2 trapping layer and dual SiO_2/Si_3N_4 tunneling layer," *IEEE Trans. Electron Devices*, vol. 54, no. 10, pp. 2699–2705, Oct. 2007.

[4] R. Ohba, Y. Mitani, N. Sugiyama, and S. Fujita, "15 nm planar bulk SONOS-type memory with double junction tunnel layers using sub-threshold slope control," in *IEDM Tech. Dig.*, 2007, pp. 75–78.

[5] H. T. Lue, S.-Y. Wang, E.-K. Lai, Y.-H. Shih, S.-C. Lai, L.-W. Yang, K.-C. Chen, J. Ku, K.-Y. Hsieh, R. Liu, and C.-Y. Lu, "BE-SONOS: A bandgap engineered SONOS with excellent performance and reliability," in *IEDM Tech. Dig.*, 2005, pp. 547–550.

[6] S. K. Samanta, P. K. Singh, W. J. Yoo, G. Samudra, Y.-C. Yeo, L. K. Bera, and N. Balasubramanian, "Enhancement of memory window in short channel non-volatile memory devices using double layer tungsten nanocrystals," in *IEDM Tech. Dig.*, 2005, pp. 177–180.

[7] J. D. Blauwe, "Nanocrystal nonvolatile memory devices," *IEEE Trans. Nanotechnol.*, vol. 1, no. 1, pp. 72–77, Mar. 2002.

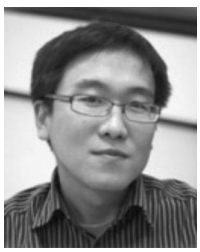
[8] C. M. Compagnoni, D. Ielmini, A. S. Spinelli, A. L. Lacaita, C. Gerardi, L. Perniola, B. De Salvo, and S. Lombardo, "Program/erase dynamics and channel conduction in nanocrystal memories," in *IEDM Tech. Dig.*, 2003, pp. 549–552.

[9] G. Zhang, C. H. Ra, H.-M. Li, C. Yang, and W. J. Yoo, "Potential well engineering by partial oxidation of TiN for high-speed and low-voltage Flash memory with good 125 °C data retention and excellent endurance," in *IEDM Tech. Dig.*, 2009, pp. 835–838.
[10] G. D. Wilk, R. M. Wallace, and J. M. Anthony, "High-k gate dielectrics: Current status and material properties considerations," *J. Appl. Phys.*, vol. 89, no. 10, pp. 5243–5275, 2001.
[11] A. R. Chourasia and D. R. Chopra, "X-ray photoelectron study of TiN/SiO_2 and TiN/Si interfaces," *Thin Solid Films*, vol. 266, no. 2, pp. 298–301, Oct. 1995.
[12] M. Wittmer, J. Noser, and H. Melchior, "Oxidation kinetics of TiN thin films," *J. Appl. Phys.*, vol. 52, no. 11, pp. 6659–6664, Nov. 1981.
[13] D. V. Lang, "Deep-level transient spectroscopy: A new method to characterize traps in semiconductors," *J. Appl. Phys.*, vol. 45, no. 7, pp. 3023–3032, Jul. 1974.
[14] 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, pp. 132 104-1–132 104-3, Mar. 2008.
[15] H. Sunamura, T. Ikarashi, A. Morioka, S. Kotsuji, M. Oshida, N. Ikarashi, S. Fujieda, and H. Watanabe, "Suppression of lateral charge migration using advanced impurity trap memory for improving high temperature retention," in *IEDM Tech. Dig.*, 2006, pp. 975–978.

Gang Zhang (S'07) received the B.Eng. (Hons.) degree in electrical and computer engineering from the National University of Singapore, Singapore, in 2005 and the M.Eng. degree in electronics from the Department of Nanoscience and Technology, SKKU Advanced Institute of Nanotechnology, Sungkyunkwan University, Suwon, South Korea, in 2008, where he is currently working toward the Ph.D. degree.

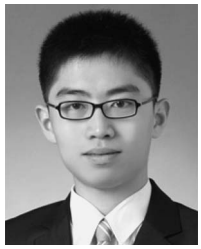
Meanwhile, he is a Research Student with the Electronic Materials Research Center, Korea Institute of Science and Technology, Seoul, South Korea. His research interests include high-k dielectric materials and charge-trapping/phase-change nonvolatile memory.

Mr. 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, in 2006, the 2008 Gold Paper Award at the IEEE Student Paper Contest, Seoul, in 2008, and the 2009 Chinese Government Award for Outstanding Self-Financed Students Abroad.



Chang Ho Ra received the B.S. and M.S. degrees from Sungkyunkwan University, Suwon, South Korea, in 2007 and 2009, respectively, where he is currently working toward the Ph.D. degree at the Department of Nanoscience and Technology, SKKU Advanced Institute of Nanotechnology. Meanwhile, he is a Research Student with the Electronic Materials Research Center, Korea Institute of Science and Technology, Seoul, South Korea.

His current research interests are simulation of semiconductor devices and application of new materials to thin-film transistors.



Hua-Min Li received the B.S. degree from the College of Physics and Electronics, Shandong Normal University, Jinan, China, in 2007 and the M.Eng. degree from the Department of Mechanical Engineering, Sungkyunkwan University, Suwon, South Korea, in 2010, where he is currently working toward the Ph.D. degree at the Department of Nanoscience and Technology, SKKU Advanced Institute of Nanotechnology.

His research interests are nonvolatile memory and solar cells.

Tian-zi Shen received the B.S. degree from the Department of Information and Communication Engineering, Kangwon National University, Chuncheon, South Korea, in 2008. He is currently working toward the Master degree at the Department of Nanoscience and Technology, SKKU Advanced Institute of Nanotechnology, Sungkyunkwan University, Suwon, South Korea.

His research interests are the organic/oxide semiconductor-based thin-film transistors.

Byung-ki Cheong received the B.E. and M.E. degrees in metallurgy from Seoul National University, Seoul, South Korea, and the Ph.D. degree in materials science and engineering with a thesis on phase transformation in solids from Carnegie Mellon University, Pittsburgh, PA.

He has been a Senior/Principal Researcher with the Korea Institute of Science and Technology (KIST), Seoul, since 1994. His major area of research has been chalcogenide thin-film materials for phase-change optical recording, superresolution optical storage, and nonvolatile phase-change electrical memory. He spent two years as a Postdoctoral Research Associate with the Data Storage Systems Center of Carnegie Mellon University, doing research on thin-film materials for magnetic and magnetooptical information storage until joining KIST, in 1994.

Won Jong Yoo (M'00–SM'02) received the B.S. and M.S. degrees from Seoul National University, Seoul, South 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 is currently a Professor with the Department of Nanoscience and Technology, SKKU Advanced Institute of Nanotechnology, Sungkyunkwan University (SKKU), Suwon, South Korea. Before joining SKKU in 2006, 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. His main industrial experiences were research and development in the areas of semiconductor material/device processes at Samsung Semiconductor Research Center, South Korea, and IBM Research Center in Yorktown Heights, NY. Areas of his current research interests are the investigation of the electrical property and reliability of nonvolatile memory devices using high-dielectric-constant materials, the investigation of plasma etching processes for future nanoelectronic devices, and the investigation of electrical properties of graphene devices. He has authored and coauthored about 140 peer-reviewed journals and conference papers.