$V_{\rm th}$ Control by Complementary Hot-Carrier Injection for SONOS Multi-Level Cell Flash Memory

Gang Zhang, *Student Member, IEEE*, and Won Jong Yoo, *Senior Member, IEEE*

***Abstract*—In this paper, a complementary hot-carrier injection programming/erasing (P/E) scheme is proposed for accurate threshold voltage ($V_{\rm th}$) control for polysilicon–oxide–nitride–oxide–silicon (SONOS) multi-level cell (MLC) Flash memory. Alternate hot electrons (HEs) and hot holes (HHs) are injected by pulse gate and drain biases, and complementary HE and HH injections spontaneously correct the $V_{\rm th}$ offsets via a feedback control enabled by the programming and the offset-correcting cycles. The HE and HH fluxes are balanced at a steady state induced by the tunnel oxide potential transformation. Accurate $V_{\rm th}$ control is accomplished, and endurance reliability is significantly improved for SONOS MLCs.**

***Index Terms*—Complementary hot-carrier injection (CHCI), multi-level cell (MLC), polysilicon–oxide–nitride–oxide–silicon (SONOS) Flash memory, $V_{\rm th}$ control.**

I. Introduction

RECENTLY, MLC as a promising technique to enhance the storage density of SONOS Flash memory [1] has been investigated extensively [2]–[5]. However, the MLC reliability is one of the major challenges for its application [2], [3]. For example, it is difficult to realize accurate $V_{\rm th}$ control for SONOS MLCs by using the self-convergent-programming (SCP) schemes based on channel hot electron injection (CHEI) and hot hole injection (HHI), due to the $V_{\rm th}$ offset [2]. Various band-engineered (BE) trapping layers have been proposed to improve the MLC reliability, although the BE gate stacks require complicated processes [6], [7]. The write/verify scheme [8] and the verified SCP scheme [9] may be other approaches to achieve the accurate MLC control. However, the write/verify scheme and the verified SCP scheme are only applicable to specially designed circuits [8], [9]. The circuit implementation will bring about considerable challenges to the application of these schemes.

In this paper, the tunneling/injection properties of HEs and HHs via the potential barrier formed by the tunnel oxide of a SONOS device are investigated. Simulation results suggest that the HE tunneling/injection property is a continuous function of φ_e and $\Delta\varphi$, where φ_e is the electron energy, and $\Delta\varphi$ is the shift in the trapping layer potential induced by charge trapping in the modified EOT model [10]. Note that φ_e obeys the Maxwell–Boltzmann distribution, and $\Delta\varphi$ affects the HE tunneling/injection probability [10]–[12]. It is understood that the $\Delta\varphi$ increase exponentially suppresses the tunneling probability for HEs, while it exponentially enhances the tunneling probability for HHs, as shown in Fig. 1(a) and (b).

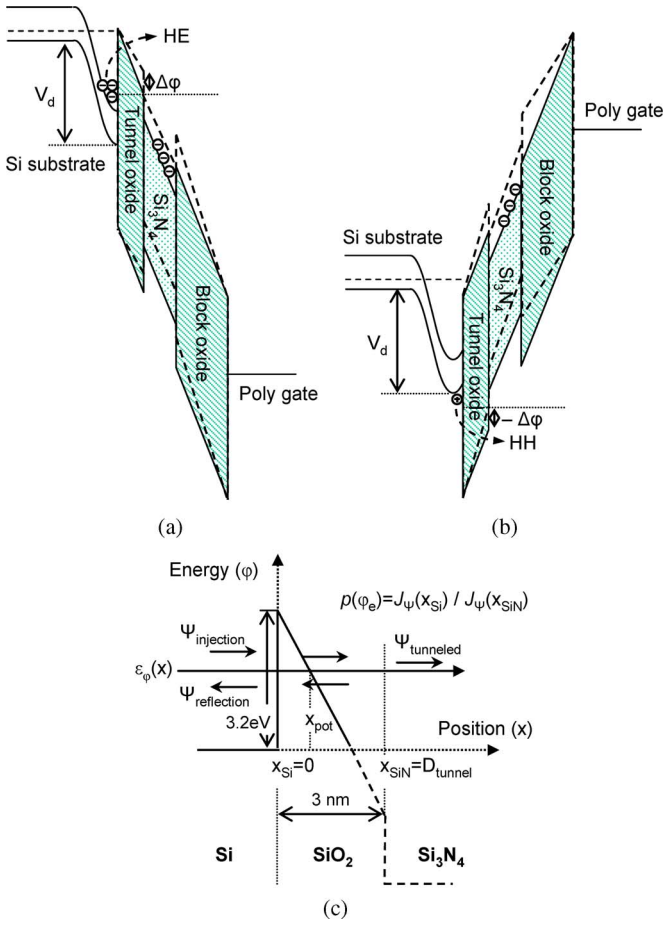

Fig. 1. Band diagram at the channel/drain junction of the SONOS device under (a) $+V_g$ and $+V_d$ and (b) $-V_g$ and $+V_d$. The trapping layer potential increase ($\Delta\varphi$) induced by the electron trapping suppresses the HE tunneling in (a) but enhances the HH tunneling in (b). (c) Potential barrier diagram to compute the electron tunneling/injection probability.

Based on the complementary tunneling/injection properties of HEs and HHs, a complementary hot-carrier injection (CHCI) P/E scheme is proposed for accurate $V_{\rm th}$ control for NOR SONOS MLCs. With grounded source and substrate, pulse gate ($\pm V_g$) and drain ($\pm V_d$) biases are applied to an n-channel device. Alternate pulse-agitated substrate HEI (PASHEI) and

Manuscript received May 8, 2009. First published September 18, 2009; current version published November 20, 2009. The review of this paper was arranged by Editor G.-T. Jeong.

G. Zhang is with the Department of Nano Science and Technology, SKKU Advanced Institute of Nanotechnology (SAINT), Sungkyunkwan University, Suwon 440-746, Korea, and also with the Thin Film Materials Research Center, Korea Institute of Science and Technology, Seoul 136-791, Korea.

W. J. Yoo is with the Department of Nano Science and Technology, SKKU Advanced Institute of Nanotechnology (SAINT), 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.2009.2030427

HHI are triggered [13], and the complementary HE and HH injections spontaneously correct the $V_{\rm th}$ offset via the $\Delta\varphi$-dependent feedback control. V_d is adjusted to control the programmed $V_{\rm th}$ levels, while the durations of HE and HH injections are modulated to switch the P/E operations. Different from the other SCP schemes, in which $V_{\rm th}$ are controlled by only the HE or HH injection [2]–[5], the CHCI P/E scheme enables the saturation of $V_{\rm th}$ levels by both the HE and HH injections. For the first time, a self-initialized feedback control, which is enabled by the programming cycles and the offset-correcting cycles, is accomplished for SONOS MLCs by CHCI. Accurate $V_{\rm th}$ control is achieved, and the MLC endurance reliability is significantly improved.

II. Mechanism of CHCI

It has been reported that φ_e during CHEI can be expressed by the Maxwell–Boltzmann distribution, in which the distribution of φ_e is a continuous function of the electron temperature $T_{\rm eo}$ at the steady state induced by V_g and V_d [10]–[12]. As a result, the coexistence of the HE tunneling/injection is inevitable. For example, it has been reported that $> 99.3\%$ of the electrons have φ_e in the range of 0.03–2.41 eV for $V_g = 6$ V, $V_d = 5$ V [14]. As φ_e, e.g., < 2.41 eV, rarely exceeds the potential barrier height of 3.2 eV, HE tunneling may play an important role during various HEI programming operations [10]. According to the Wentzel–Kramers–Brillouin approximation, the HE tunneling probability is greatly affected by $\Delta\varphi$ [10]. For example, when the tunneled HEs are captured in the trapping layer, the transformation of the tunnel oxide potential induced by the $\Delta\varphi$ increase suppresses the subsequent HE tunneling in an exponential manner, as shown in Fig. 1(a). The φ_e- and $\Delta\varphi$-dependent electron tunneling probability $p(\varphi_e)$ is estimated with the help of numerical simulation. Decoupling the 3-D Schrödinger equation, which describes the wave function of hot carriers (Ψ), in two independent equations for momentum and energy $(\varepsilon_P, \varepsilon_\varphi)$, one reports that ε_P and ε_φ are parallel and perpendicular to the Si-substrate–barrier interface, as shown in Fig. 1(c) [15]. Solving the 1-D equation of ε_φ in the band structure shown in Fig. 1(c), analytical solutions for Ψ are found in terms of Bessel functions of order $\pm 1/3$ in the potential barrier (SiO_2) (Airy functions), and of plane waves in Si-substrate and in the conduction band of the trapping layer (Si_3N_4) [15]. The $p(\varphi_e)$ is calculated from the ratio of the wave function probability flux (J_Ψ) at the Si/SiO_2 $(x = x_{\rm Si})$ and SiO_2/Si_3N_4 $(x = x_{\rm SiN})$ interfaces for a triangular barrier, as $p(\varphi_e) = (J_\Psi(x_{\rm SiN})/J_\Psi(x_{\rm Si})) = \exp\left(-2\int_0^{x_{\rm pot}} (\sqrt{2m^*}/\hbar)\sqrt{q(\varphi_{\rm pot} - \varphi_e)\,(1 - (x/x_{\rm pot}))}\,dx\right)$, where $\hbar$ is the reduced Planck constant of 6.58×10^{-16} eV $\cdot$ s, q is electron charge of 1.6×10^{-19} C, m^* is the electron effective mass of $1.08m_e \approx 9.83 \times 10^{-31}$ kg (m_e is the electron rest mass of 9.1×10^{-31} kg), and $\varphi_{\rm pot}$ and $x_{\rm pot}$ are the barrier height and the effective width of the potential barrier, respectively. Note that in thin SiO_2, i.e., ≤ 3 nm, the carrier transport is essentially ballistic, and $p(\varphi_e)$ can be modeled using the potential barrier profile, as shown in Fig. 1(c), in which the final states of the carrier tunneling are referred to the conduction band of Si_3N_4. It is understood that the deep states in SiO_2 and Si_3N_4 can significantly assist the tunneling when φ_e is very small. However, the trap-assistant tunneling induced by the deep states in SiO_2 is excluded in the simulation, since the effects of the trap-assistant tunneling may be overwhelmed by ballistic transport of electrons in the ultrathin ($\sim$3 nm) SiO_2 barrier. In addition, the deep states in Si_3N_4 are excluded in the simulation, since the trap states, i.e., their potential depth and spatial distribution, in Si_3N_4 are still unclear [10].

For a triangular potential barrier with a bottom thickness $(D_{\rm tunnel})$ of 3 nm, $\varphi_{\rm pot} = 3.2$ eV, and biased by an electrical filed $(E_{\rm ox})$ of 5.5 MV/cm, as shown in Fig. 1(c), the estimated $p(\varphi_e)$ transients are shown in Fig. 2(a). It is found that $p(\varphi_e)$ decays exponentially for $\varphi_e < 2.5$ eV when $\Delta\varphi$ increases, i.e., $p(\varphi_e = 1.18$ eV) are on the order of 10^{-8}, 10^{-9}, and 10^{-10} for $\Delta\varphi = 0.0$, 1.0, and 2.0 eV, respectively. However, $p(\varphi_e)$ becomes insensitive to $\Delta\varphi$ for $\varphi_e \geq 2.5$ eV, i.e., $p(\varphi_e)$ are $\sim$2.6%, $\sim$31.8%, and $\sim$60.6% for $\varphi_e = 2.75$, 3.0, and 3.25 eV, despite $\Delta\varphi = 0.0$, 1.0, and 2.0 eV, respectively. It shall be noticed that the $\Delta\varphi$-dependent $p(\varphi_e)$ transients overlap for HEs with 2.5 eV $\leq \varphi_e \leq$ 3.5 eV due to scale of the plot. However, the effect of $\Delta\varphi$ is still remarkable. For example, $p(\varphi_e = 2.75$ eV) are $\sim$4.1%, $\sim$3.3%, and $\sim$2.6% for $\Delta\varphi =$ 0.0, 1.0, and 2.0 eV, respectively.

On the other hand, the $\Delta\varphi$ increase enhances HH tunneling exponentially, as shown in Fig. 1(b). The φ_h- and $\Delta\varphi$-dependent hole tunneling probability $p(\varphi_h)$ is estimated by using the same method as that used for electrons, but having $D_{\rm tunnel} = 3$ nm, $\varphi_{\rm pot} = 4.4$ eV, $E_{\rm ox} = -6.0$ MV/cm, and effective mass of $0.56m_e$ for holes in Si and SiO_2 [15]. The estimated $p(\varphi_h)$ transients are shown in Fig. 2(b). It is found that $p(\varphi_h)$ increases exponentially for $\varphi_h < 3.0$ eV when $|\Delta\varphi|$ increases, i.e., $p(\varphi_h = 2.13$ eV) are on the order of 10^{-9}, 10^{-8}, and 10^{-7} for $|\Delta\varphi|$ of 0.0, 1.0, and 2.0 eV, respectively. However, $p(\varphi_h)$ becomes insensitive to $|\Delta\varphi|$ for $\varphi_h \geq 3.5$ eV, i.e., $p(\varphi_h)$ are $\sim$14.6%, $\sim$43.8%, and $\sim$63.2% for $\varphi_h = 3.75$, 4.0, and 4.25 eV, despite $|\Delta\varphi| = 0.0$, 1.0, and 2.0 eV, respectively. The estimated tunneling/injection probabilities for electrons and holes are consistent with the reported results in [15], [16].

As both the HE and HH tunnelings are affected by the $\Delta\varphi$ shift, while their effects on the programmed $V_{\rm th}$ level are complementary, it is possible to eliminate the $V_{\rm th}$ offset by CHCI. Monte Carlo simulation suggests that the maximum HE flux $(\Phi_{\rm HE})$ is 5.41×10^{18} cm$^{-2} \cdot$ s^{-1} at the drain junction [5] by assuming the HE generation lifetime of 40 ns [17], while $\Phi_{\rm HE}$ is negligible beyond $\sim$60 nm from the drain junction [5]. By integrating the HE tunneling/injection amplitudes, the average $\Phi_{\rm HE}$ over the HE tunneling/injection region $(W_{\rm injection})$ can be obtained. The average HH flux $(\Phi_{\rm HH})$ can be obtained by the same method. For example, the average $\Phi_{\rm HE}$ and $\Phi_{\rm HH}$ are estimated to be 1.76×10^{18} and 2.75×10^{17} cm$^{-2} \cdot$ s^{-1} over $W_{\rm injection} = 50$ nm at $\Delta\varphi = 0$ eV for CHCI by $\pm E_{\rm ox} = \pm 5.5$ MV/cm, $V_d =$ 5–5.5 V [5], as shown in Fig. 2(c). The details of the device simulation on the hot-carrier profiles have been reported in [5], [18]. It has been reported that both the HE and HH profiles are localized within $W_{\rm injection} = 50$ nm for PASHEI and HHI [5], [13]. We thus assume that the overlapped HE and HH injections properly prevent the lateral

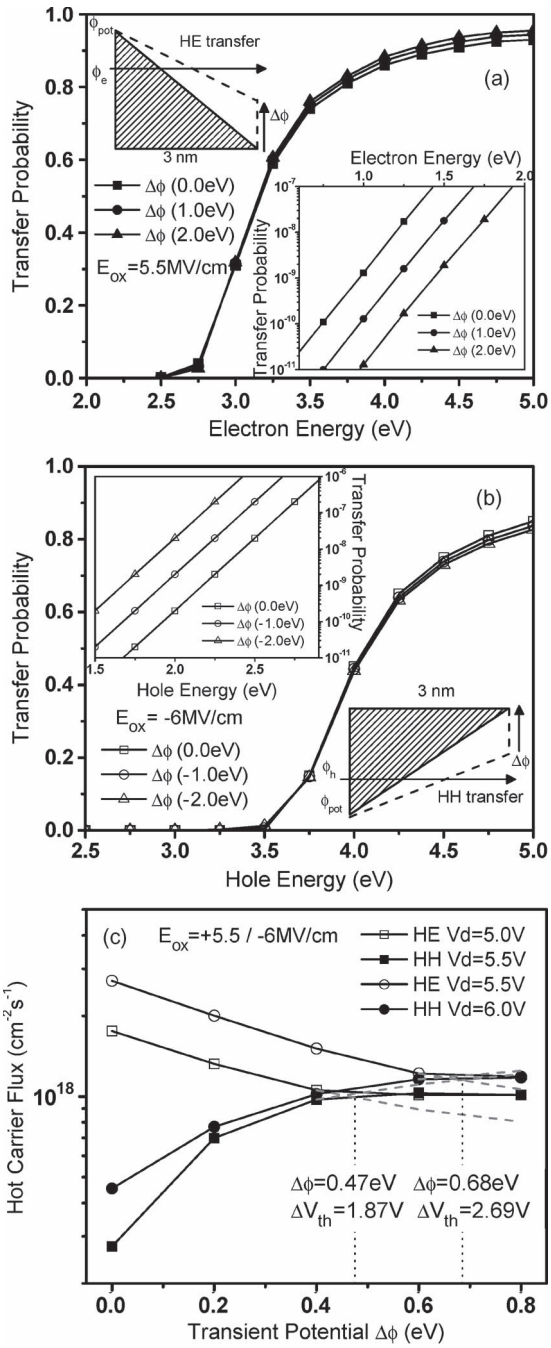

Fig. 2. (a) Estimated HE tunneling probability $p(\varphi_e)$ for $\varphi_e = 0.5$–5 eV, and $\Delta\varphi = 0$–2 eV. (b) Estimated HH tunneling probability $p(\varphi_h)$ for $\varphi_h = 1.5$–5 eV, and $\Delta\varphi = -2$–0 eV. (c) Estimated HE and HH fluxes when $\Delta\varphi = 0$–0.8 eV (solid lines) with and (dashed lines) without CHCI.

charge accumulation induced by the mismatch of HE and HH profiles, giving rise to accurate charge injection control and improved endurance reliability [5], [19].

By multiplying the HE and HH flux densities at $\Delta\varphi = 0$ eV and the $p(\varphi_e)$ and $p(\varphi_h)$ transients for $\Delta\varphi = 0$–0.8 eV, the $\Delta\varphi$-dependent $\Phi_{\rm HE}$ and $\Phi_{\rm HH}$ transients can be estimated, as shown in Fig. 2(c). It is noticed that the amplitude of $\Phi_{\rm HE}$ is around seven times of that of $\Phi_{\rm HH}$ at $\Delta\varphi = 0$ eV, thus $\Phi_{\rm HE}$ dominates CHCI, enabling programming. When some of the electrons are captured by the trapping layer, the $\Delta\varphi$ increase suppresses $\Phi_{\rm HE}$ but enhances $\Phi_{\rm HH}$, until a steady state for the HE and HH tunneling/injection is reached at $\Delta\varphi = 0.47$ eV, where $\Phi_{\rm HE}$ and $\Phi_{\rm HH}$ are balanced at 1.02×10^{18} cm$^{-2} \cdot$ s^{-1}, as shown (solid lines) in Fig. 2(c). We define the $\Delta\varphi$ at the steady state as $\Delta\varphi_{\rm SS}$. When $\Delta\varphi$ is assumed to exceed $\Delta\varphi_{\rm SS}$, the amplitude of $\Phi_{\rm HH}$ will exceed that of $\Phi_{\rm HE}$, i.e., the $\Phi_{\rm HE}$ and $\Phi_{\rm HH}$ are 8.95×10^{17} and 1.11×10^{18} cm$^{-2} \cdot$ s^{-1} at $\Delta\varphi = 0.60$ eV when the offset-correcting cycles (t_3) are excluded in CHCI, as shown (dashed lines) in Fig. 2(c). However, the continuous increase of $\Delta\varphi$ is prevented when the offset-correcting cycles are included in CHCI. For example, when $\Delta\varphi$ exceeds $\Delta\varphi_{\rm SS}$, the $\Phi_{\rm HH}$ starts to dominate CHCI and enables erasing. Recall the programming cycles (t_2), in which $\Delta\varphi$ is less than $\Delta\varphi_{\rm SS}$. The programming and the offset-correcting cycles then accomplish a feedback control, which eventually restores the $\Delta\varphi_{\rm SS}$ and enables saturated $\Delta V_{\rm th}$. For example, the programmed $\Delta V_{\rm th}$ may saturate at 1.87 V $(\Delta V_{\rm th} = \Delta\varphi \times ({\rm EOT}/D_{\rm tunnel})$, where EOT (11.9 nm) is the equivalent oxide thickness of the SONOS gate stack [10]), when CHCI is applied by $+E_{\rm ox} = +5.5$ MV/cm, $V_d = 5.0$ V for $t_2 = 2$ μs, and $-E_{\rm ox} = -6.0$ MV/cm, $V_d = 5.5$ V for $t_3 = 3$ μs. On the other hand, by adjusting the $\Phi_{\rm HE}$ and $\Phi_{\rm HH}$ with different sets of V_g and V_d [5], the steady states can be reached at various $\Delta\varphi_{\rm SS}$. For example, saturated $\Delta V_{\rm th} = 2.69$ V can be achieved at $\Delta\varphi_{\rm SS} = 0.68$ eV when CHCI is applied by $+E_{\rm ox} = +5.5$ MV/cm, $V_d = 5.5$ V for $t_2 = 2$ μs, and $-E_{\rm ox} = -6.0$ MV/cm, $V_d = 6.0$ V for $t_3 = 3$ μs, as shown in Fig. 2(c). The saturation of $V_{\rm th}$ is unachievable for the other SCP schemes without CHCI [2]–[7].

It must be mentioned that the programming speeds of CHCI may be lowered by the offset-correcting cycles. Nevertheless, the programming speeds can be largely enhanced by the modified CHCI scheme: 1) PASHEI (t_1 and t_2) or CHEI (t_2) are applied to achieve a fast-speed programming at the first stage and 2) CHCI (t_1, t_2, and t_3) is applied to correct the $V_{\rm th}$ offset at the second stage. The duration ratio of the second stage to the first stage requires intensive investigation. Theoretically, the modified CHCI may have programming speeds comparable with the other SCP schemes in [2]–[7]. Compared with the write/verify scheme [8] and the verified SCP scheme [9], the modified CHCI is applicable to current NOR circuits, being practical.

III. CHCI P/E Properties

The waveforms of the CHCI biases are shown in Fig. 3: At t_1, grounded V_g and $-V_d$ are applied to trigger the excess electrons emission over the forward-biased p-n junction between the p-Si substrate and the n$^+$ drain [13]; at t_2, $+V_g$ and $+V_d$ accelerate the excess electrons over the reverse-biased p-n junction and enable the programming cycles [13]; at t_3, HHI is triggered by $-V_g$ and $+V_d$ in the channel/drain junction region to enable the offset-correcting cycles. The feedback control enabled by the programming and offset-correcting cycling gives rise to saturated $V_{\rm th}$ levels for the CHCI P/E operations. The P/E biases are listed in Table I.

SONOS devices used in this paper have ONO thicknesses of 3/6/10 nm and channel length $L_g = 70$–300 nm [11]. Fig. 4(a) shows the MLC $V_{\rm th}$ control without CHCI. $V_g = 0$ V,

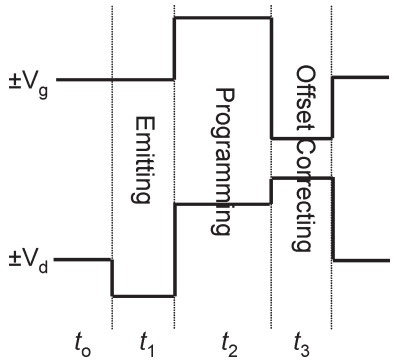

Fig. 3. Schematic of the CHCI waveforms of gate ($\pm V_g$) and drain ($\pm V_d$) pulse biases.

TABLE I
P/E CONDITIONS OF CHCI

P/E	Time (μs)		V_g (V)	V_d (V)
Pro.	t_1	2	0	−1.5
	t_2	2	10	5-6
	t_3	3	−13	5.5-6.5
Era.	t_1	2	0	−1.5
	t_2	1-2	10	5-6
	t_3	9	−13	5.5-6.5

$-V_d = -1.5$ V are applied during $t_1 = 2$ μs, and $V_g = 10$ V, $V_d = 5.0$, 5.5, and 6.0 V are applied during $t_2 = 2$ μs. Distinct $V_{\rm th}$ with the trends of saturation at $\Delta V_{\rm th} = 1.48$, 2.37, and 3.39 V are observed over 10^3 pulses for $V_d = 5.0$, 5.5, and 6.0 V, respectively. However, $V_{\rm th}$ saturation is unachievable for each level [2]. Fig. 4(b) shows the $V_{\rm th}$ control with CHCI. $-V_d = -1.5$ V is applied during $t_1 = 2$ μs, $V_g = 10$ V, $V_d =$ 5.5 V are applied during $t_2 = 2$ μs, and $-V_g = -13$ V, $V_d =$ 5.5, 6.0, and 6.5 V are applied during $t_3 = 3$ μs. The programmed $V_{\rm th}$ saturates at ∼3.49 V for all $V_d = 5.5$, 6.0, and 6.5 V at $t_3 = 3$ μs. This suggests that V_d during t_3 is ineffective to the $V_{\rm th}$ control, as the $\Phi_{\rm HE}$ during t_2 dominates the net P/E effect, and the $\Phi_{\rm HH}$ during t_3 only corrects the $V_{\rm th}$ offset under proper (t_3/t_2) ratio. However, by adjusting the pulsewidths of $t_3 = 3$–9 μs, the effect of CHCI is switched to erasing. When $-V_g = -13$ V, $V_d = 6.0$ V are applied for $t_3 = 3$, 6, and 9 μs, the programmed or erased $V_{\rm th}$ levels saturate at 3.52, 2.57, and 0.46 V, respectively, as shown in the inset of Fig. 4(b). It is understood that the t_1 and t_2 can be excluded to achieve a faster erasing, once overerasure is properly prevented.

Fig. 5(a) shows the CHCI MLC P/E properties of a 130-nm device. As listed in Table I, $V_g = 0$ V, $-V_d = -1.5$ V during $t_1 = 2$ μs, $V_g = 10$ V, $V_d = 5.0$, 5.5, and 6.0 V during $t_2 =$ 2 μs, and $-V_g = -13$ V, $V_d = 5.5$, 6.0, and 6.5 V during $t_3 = 3$ μs are applied to give rise to saturated $V_{\rm th}$ levels 10/01/00 at 2.71, 3.53, and 4.34 V, as shown (solid lines) in Fig. 5(a). In contrast, MLC $V_{\rm th}$ level saturation is unachievable without CHCI. For example, $V_{\rm th}$ levels 10/01/00 keep increasing to 2.93, 3.75, and 4.71 V at 10^3 P/E cycles, as shown (dashed lines) in Fig. 5(a). This makes the MLC $V_{\rm th}$ control difficult. The same pulse voltages are applied during $t_2 =$ 1–2 μs and $t_3 = 9$ μs for CHCI erasing, and the erased levels 10/01/00 saturates at ∼1.41 V. CHCI spontaneously corrects

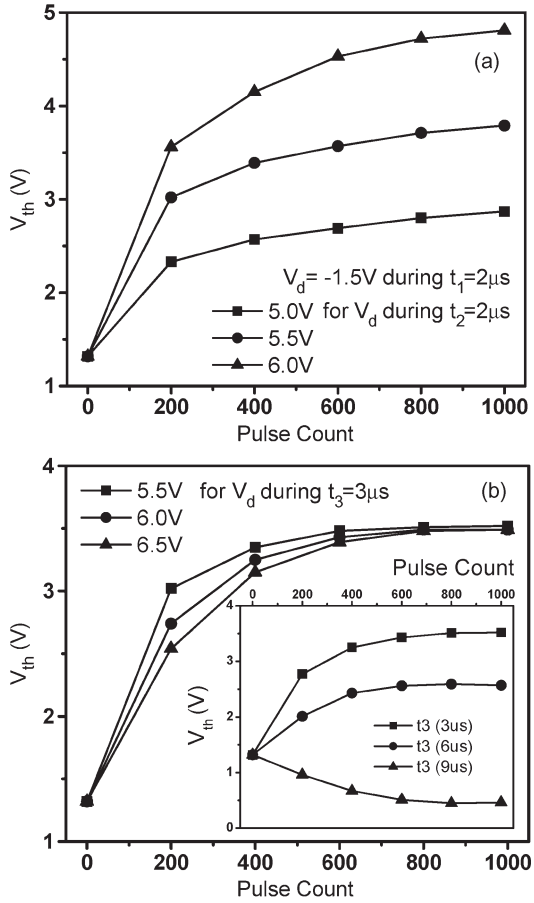

Fig. 4. (a) $V_{\rm th}$ transient for $V_g = 0$ V, $-V_d = -1.5$ V during $t_1 =$ 2 μs and $V_g = 10$ V, $V_d =$ 5–6 V during $t_2 = 2$ μs. (b) $V_{\rm th}$ transient for $V_g = 0$ V, $-V_d = -1.5$ V during $t_1 = 2$ μs, $V_g = 10$ V, $V_d = 5.5$ V during $t_2 = 2$ μs, and $-V_g = -13$ V, $V_d = 5.5$–6.5 V during $t_3 = 3$ μs. (Inset) Switching of the CHCI P/E by adjusting the duration of t_3/t_2, with $V_g = 0$ V, $-V_d = -1.5$ V during $t_1 = 2$ μs, $V_g = 10$ V, $V_d = 5.5$ V during t_2, and $-V_g = -13$ V, $V_d = 6.0$ V during t_3.

overprogramming and overerasure, giving rise to accurate MLC $V_{\rm th}$ control. Note that the devices used in this paper have a high-quality Si_3N_4 layer formed by plasma-enhanced chemical vapor deposition [11], thus its trapping efficiency is lower compared to the devices used in the other works [3]–[7]. Even so, the programming speeds of CHCI (∼2 ms) in this paper are still comparable with that of the write/verify scheme (∼1 ms) [8]. However, the P/E speeds of CHCI can be enhanced by using the modified CHCI scheme or by engineering the trapping layer properties [5], [11]. Therefore, we think that CHCI has greater potential of application compared to the writing/verify scheme.

Cumulative distributions of the programmed $V_{\rm th}$ levels for SONOS MLCs at 1 and 10^4 P/E cycles are shown in Fig. 5(b). The average sensing windows between the adjacent $V_{\rm th}$ levels, e.g., for 63% of the devices, and the minimum sensing windows between the head and the tail $V_{\rm th}$ of the adjacent levels are well maintained by the $\Delta\varphi$ $(\Delta V_{\rm th})$ control after 10^4 P/E cycles. For fresh cells, the average sensing windows are 0.91, 1.03, and 1.13 V, and the minimum sensing windows are 0.83, 0.91, and 0.95 V between levels 11/10/01/00, respectively. Throughout the 10^4 P/E cycles, the average sensing windows are 0.88, 1.08, and 1.25 V, and the minimum sensing windows are 0.76, 0.97,

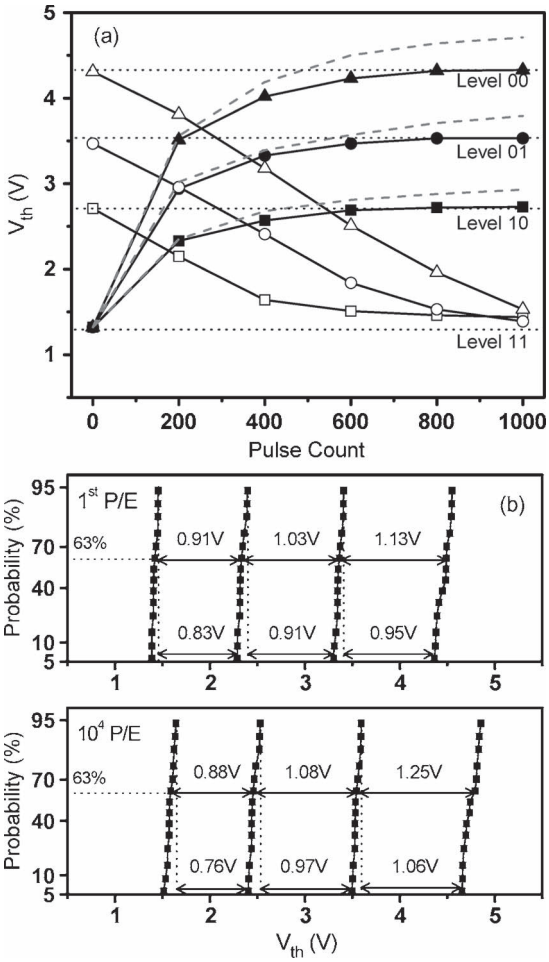

Fig. 5. (a) MLC $V_{\rm th}$ transient for CHCI P/E (solid lines). $V_g = 10$ V, $V_d =$ 5, 5.5, and 6 V give rise to saturated $V_{\rm th}$ levels 10/01/00 after programming, while $-V_g = -13$ V, $V_d = 5.5$, 6, and 6.5 V give rise to saturated level 11 after erase. The saturation of MLC $V_{\rm th}$ is unachievable without using CHCI (dashed lines). (b) MLC $V_{\rm th}$ cumulative distribution for SONOS MLCs at 1 and 10^4 P/E cycles. $V_{\rm th}$ clearances are maintained throughout the 10^4 P/E cycles between levels 11/10/01/00.

and 1.06 V, respectively. In contrast, the MLC $V_{\rm th}$ diverges significantly without CHCI. For example, the minimum sensing windows are 0.93, 0.53, and 0.17 V throughout the 10^4 P/E cycles, and level merge becomes inevitable between levels 10/01 and levels 01/00 at the 3×10^4 P/E cycles when CHEI/HHI is applied for the SONOS MLC operation [2]. However, stress-induced tunnel oxide degradation is still unavoidable, resulting in consistent $V_{\rm th}$ increments for all levels 11/10/01/00. The performances of CHCI are compared with the other works in Table II.

TABLE II
COMPARISON OF THIS WORK TO THE OTHER WORKS

References	This work	IEDM 2005, p.539 [3]	IEDM 2006, p.963 [4]	TED v54n12, p.3177, 2007[5]
Device structure	SONOS	NROM	SONOS	SOHOS (HfO_2)
Circuit	NOR	NOR	NAND	NOR
P/E mode	CHCI	CHEI/HHI	FN	CHEI/HHI
Sensing window	~0.9V	~1.2V	>1V	~1.2V
$V_{\rm th}$ offset of an cell at 10^4 cycles	<0.3V (Fig. 15)	-	~0.6V (Fig. 15 of [4])	~0.7V (Fig. 9(a) of [5])
$V_{\rm th}$ offset of an array at 10^4 cycles	<0.2V (Fig. 17)	~0.9V (Fig. 6 of [3])	-	-
P/E cycles	$>10^4$	10^4	10^4	10^5

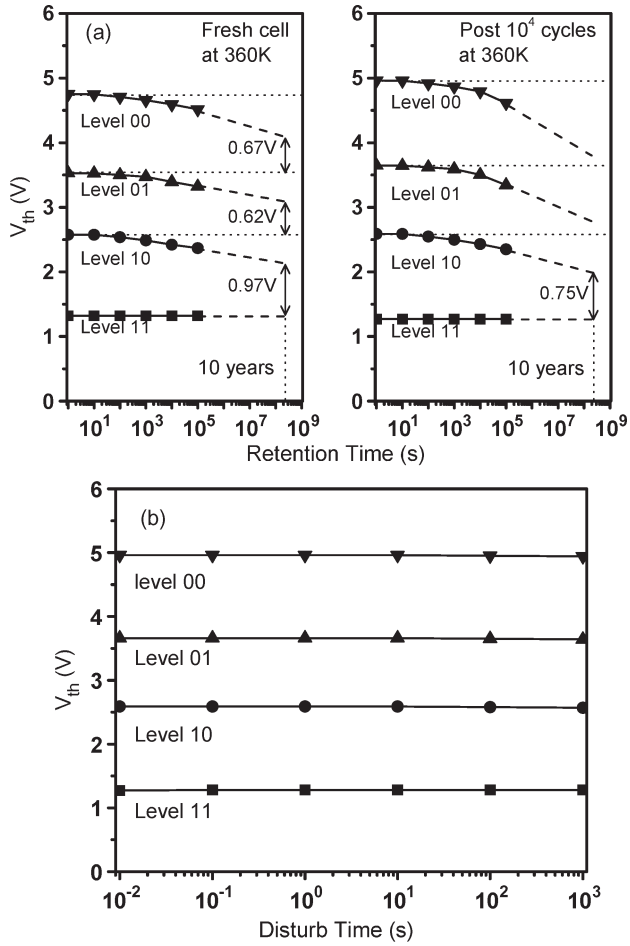

Fig. 6. (a) Retention property of SONOS MLCs at 360 K for CHCI programming before and after the 10^4 P/E cycling. (b) Effects of gate and drain disturbances at $V_d = 0$ V, $\pm V_g = 10$ V/ -13 V and $V_d = 6$ V, $V_g = 0$ V for up to 10^3 s.

IV. CHCI RELIABILITY

Fig. 6(a) shows the MLC retention properties at 360 K for fresh SONOS devices and those post-10^4 P/E cycles. The estimated ten-year $V_{\rm th}$ windows between levels 11/10/01/00 are 0.97, 0.62, and 0.67 V for fresh devices. After 10^4 P/E cycles, the retention property is degraded due to the tunnel oxide damage induced by repetitive hot-carrier injections. For example, level merge is observed for levels 00 and 01 after ten-year retention. The MLC retention properties of CHCI P/E are consistent with that of CHEI/HHI in our previous work [5]. We understand the $V_{\rm th}$ decay is attributed to the leakage via the ultrathin (~3 nm) tunnel oxide, particularly after the P/E cycling. The engineered tunnel barriers are reported to significantly improve the retention [6], [7]. Fig. 6(b) shows the immunity to the P/E disturbances. $V_d = 0$ V, $\pm V_g = 10$ V/ -13 V and $V_d = 6$ V, $V_g = 0$ V are applied for up to 10^3 s. Negligible disturbances to levels 11/10/01/00 are observed for the SONOS devices after 10^4 P/E cycles. We understand that the low-voltage operation of CHCI [13] suppresses the P/E disturbances for SONOS devices.

V. Conclusion

The mechanism of CHCI is investigated, and the accurate $V_{\rm th}$ control for SONOS MLCs by CHCI is accomplished. CHCI spontaneously corrects the $V_{\rm th}$ offset via the $\Delta\varphi$-dependent feedback control, improving the MLC endurance reliability significantly.

References

[1] ITRS 2007. [Online]. Available: http://www.itrs.net/Links/2007ITRS/2007_Chapters/2007_PIDS.pdf

[2] G. Zhang, W. S. Hwang, S.-H. Lee, B.-J. Cho, and W. J. Yoo, "Endurance reliability of multi-level cell Flash memory using ZrO_2/Si_3N_4 dual charge storage layer," *IEEE Trans. Electron Devices*, vol. 55, no. 9, pp. 2361–2369, Sep. 2008.

[3] B. Eitan, G. Cohen, A. Shappir, E. Lusky, A. Givant, M. Janai, I. Bloom, Y. Polansky, O. Dadashev, A. Lavan, R. Sahar, and E. Maayan, "4-bit per cell NROM reliability," in *IEDM Tech. Dig.*, 2005, pp. 539–542.

[4] C. Friederich, M. Specht, T. Lutz, F. Hofmann, L. Dreeskornfeld, W. Weber, J. Kretz, T. Melde, W. Rosner, E. Landgraf, J. Hartwich, M. Stadele, L. Risch, and D. Richter, "Multi-level p^+ tri-gate SONOS NAND string arrays," in *IEDM Tech. Dig.*, 2006, pp. 963–966.

[5] G. Zhang, S. K. Samanta, P. K. Singh, F.-J. Ma, M.-T. Yoo, Y. H. Roh, and W. J. Yoo, "Partial crystallization of HfO_2 for 2-bit/4-level SONOS-type Flash memory," *IEEE Trans. Electron Devices*, vol. 54, no. 12, pp. 3177–3185, Dec. 2007.

[6] Z. L. Huo, J. K. Yang, S. H. Lim, S. J. Baik, J. Y. Lee, J. H. Han, I.-S. Yeo, U.-I. Chung, J. T. Moon, and B.-I. Ryu, "Band engineered charge trapping layer for highly reliable MLC Flash memory," in *VLSI Symp. Tech. Dig.*, 2007, pp. 138–139.

[7] H.-T. Lue, E.-K. Lai, Y. H. Hsiao, S. P. Hong, M. T. Wu, F. H. Hsu, N. Z. Lien, S. Y. Wang, L. W. Yang, T. Yang, K. C. Chen, Y. K. Hsieh, R. Liu, and C.-Y. Lu, "A novel junction-free BE-SONOS NAND Flash," in *VLSI Symp. Tech. Dig.*, 2008, pp. 140–141.

[8] Y. Yamashita, T. Shibata, and T. Ohmi, "Write/verify free analog nonvolatile memory using a neuron—MOS comparator," in *Proc. IEEE Int. Symp. Circuits Syst.*, 1996, vol. 4, pp. 229–232.

[9] P. Candelier, F. Mondon, B. Guillaumot, G. Reimbold, H. Achard, F. Martin, and J. Hartmann, "Hot carrier self convergent programming method for multi-level Flash cell memory," in *Proc. IRPS*, 1997, pp. 104–109.

[10] 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.

[11] 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.

[12] A. Toriumi, M. Yoshimi, M. Iwase, Y. Akiyama, and K. Taniguchi, "A study of photon emission from n-channel MOSFETs," *IEEE Trans. Electron Devices*, vol. ED-34, no. 7, pp. 1501–1508, Jul. 1987.

[13] S. S. Chung, Y. H. Tseng, C. S. Lai, Y. Y. Hsu, E. Ho, T. Chen, L. C. Peng, and C. H. Chu, "Novel ultra-low voltage and high-speed programming/erasing scheme for SONOS Flash memory with excellent data retention," in *IEDM Tech. Dig.*, 2007, pp. 457–460.

[14] D. Cassi and B. Ricco, "An analytical model of the energy distribution of hot electrons," *IEEE Trans. Electron Devices*, vol. 37, no. 6, pp. 1514–1521, Jun. 1990.

[15] A. Ghetti, L. Selmi, E. Sangiorgi, A. Abramo, and F. Venturi, "A combined transport-injection model for hot-electron and hot-hole in the gate oxide of MOS structure," in *IEDM Tech. Dig.*, 1994, pp. 363–366.

[16] L. Selmi, F. Sangiorgi, R. Bez, and B. Ricco, "Measurement of the hot hole injection probability from Si into SiO_2 in p-MOSFETs," in *IEDM Tech. Dig.*, 1993, pp. 333–336.

[17] A. L. Fahrenbruch and R. H. Bube, *Fundamentals of Solar Cells*. New York: Academic, 1983.

[18] 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.

[19] D. Fuks, A. Kiv, Y. Roizin, M. Gutman, R. A. Bibi, and T. Maximova, "The nature of HT $V_{\rm th}$ shift in NROM memory transistors," *IEEE Trans. Electron Devices*, vol. 53, no. 2, pp. 304–313, Feb. 2006.

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

He is currently a Research Student with the Thin Film Materials Research Center, Korea Institute of Science and Technology, Seoul, Korea. His research interests include high-κ dielectric materials and charge-trapping/phase-change nonvolatile memories.

Mr. Zhang is the recipient of the 2006 Best Student Paper Award at the 8th International Conference on Solid-State and Integrated Circuit Technology, Shanghai, China, and the 2008 Gold Paper Award at the IEEE Student Paper Contest, Seoul.

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 from Rensselaer Polytechnic Institute, Troy, NY, in 1993, focusing on the plasma etching properties of Si and SiO_2.

He was an Associate Professor with National University of Singapore, Singapore, where he conducted his research in the areas of silicon device and plasma processing. Since 2006, he has been with Department of Nano Science and Technology, SKKU Advanced Institute of Nanotechnology (SAINT), Sungkyunkwan University, Suwon, Korea, where he is currently a Professor. His main industrial experiences were research and development in the areas of semiconductor material/device processes with the Samsung Semiconductor Research Center, Korea, and the IBM Research Center, Yorktown Heights, NY. He has authored and coauthored about 130 peer-reviewed journal and conference papers. 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 3-D interconnects, and the improvement of thin-film solar cell efficiency using nano-structures.