Novel Ultra-Low Voltage and High-Speed Programming/Erasing Schemes for SONOS Flash Memory with Excellent Data Retention

Steve S. Chung[1], Y. H. Tseng[2], C. S. Lai[2], Y. Y. Hsu[3], Eric Ho[3], Terry Chen[3], L. C. Peng[3], and C. H. Chu[3]
[1]*Department of Electronic Engineering, National Chiao Tung University, Taiwan*
[2]*Department of Electronic Engineering, Chang Gung University,Taiwan,* [3] *ProMOS Technology Inc., Hsinchu, Taiwan*

Abstract- A novel cell operation scheme featuring low voltage, high speed, and excellent data retention has been proposed for SONOS flash memory. First, in 1bit/cell operation, program is achieved by a back-bias assisted hot hole injection, while erase is achieved by forward-bias assisted electron injection. For a thick tunnel oxide (50A°), the ultra-low voltage (~5V) and ultra-fast speed (<1μsec) operation has been the ***record reported*** to date. On the other hand, a 2 bit/cell operation is also demonstrated, in which very good retention can be achieved in comparison to conventional operation schemes, e.g., CHE (channel hot electron) or BTB(Band-to-band) tunneling etc.

1. Introduction

In a SONOS cell, the program/erase operation is achieved via the tunneling of electrons/holes through the ***very thin*** tunnel oxide. As a result, the cell has poor retention due to this *thin* tunnel oxide. In order to improve the retention of the SONOS cells, a relatively thick bottom oxide is used to prevent charge loss such as NROM cell [1]. However, the thicker bottom oxide increases the difficulty for holes injection during erase, even for the BTBT HHI (Band to Band Tunneling induced Hot Hole Injection) [2] method. Although CHE exhibits a faster programming speed than the FN, the cell reliabilities, e.g., cycling endurance, and data retention are very poor. For the cell erase operation, the commonly used schemes, e.g., FN, BTBT HHI, a higher voltage is needed for erasing. Also, they exhibit poor endurance and retention after the cycling.

In this paper, a new combination of operation schemes for hole/electron injection which is **applicable for 1bit/cell or 2bit/cell** operation has been proposed. The tunnel oxide can be **kept thick enough** to overcome the data retention problem by using the proposed programming/erasing schemes.

2. Device Preparation

The n-channel flash SONOS cells with ONO thickness (50/40/50) (A°), W/L=0.2/0.2(μm), were fabricated using 0.13μm CMOS technology. Medium-doped n^+-region in the S/D extension was used to maintain a higher breakdown voltage.

3. Operation Schemes

A. Injection Scheme for Holes- BBHHI

Fig. 1 shows the schematic for hole injection, called back-bias assisted hot hole injection (**BB-HHI**). During T_1, source and drain are kept floating with a positive bias ($+V_{top}$) applied at the bulk, where source or drain-substrate junction capacitance is charged [3] (the junction is switched to forward bias mode). At the end of T_2, bulk bias is turned-off (the junction is switched to reverse bias mode) and changed to $+V_{base}$, the generated holes in the space charge region are then injected into the nitride storage through discharging, during the V_B falling pulse edge, T_3, as can be seen from Figs. 2(a) and 2(b).

Fig. 2(a) shows the waveform of the applied voltage with V_{top}- V_{base} (=5V-2V) and its decay due to the discharge measured from the drain or source, in which the difference of each case along the voltage axis is the built-in potential between the p-n+ junctions. Fig. 2(b) shows three kinds of discharging conditions for the same V_{top} but different V_{base}, in which the same charging height and initial discharging behavior for the three conditions due to the unique supplying pulse height V_{top}= 5V, are compared.

Fig. 3 shows the ΔV_T versus the pulse height, in which a higher level V_{top} gives a larger ΔV_T. Moreover, it also shows the ΔV_T as a function of ΔV_B (=V_{top}- V_{base}), and the ΔV_T reaches its maximum at ΔV_T = 4V. On the left hand side of the peak (ΔV_B< 4V), the V_t shift is smaller since the variation of the depletion region at the S/D p-n+ junction during the transition from forward bias mode to reverse bias mode is not wide enough. Therefore, fewer holes are generated via the junction breakdown in the space charge region. On the right hand side of the peak (ΔV_B > 4V), the variation of the depletion region is wider during the transition from the forward to reverse bias mode but the voltage of V_{base} which is feasible for the hot hole to inject into the nitride layer is smaller. For example, comparing the conditions of V_{top}- V_{base} =7V- 2V and V_{top}- V_{base} =7V- 0V, although the condition of V_{top}- V_{base} =7V- 0V has larger space charge region variation during the transition from forward bias mode to reverse bias mode and the hot holes generated via discharging induced breakdown are much more than the condition of V_{top}- V_{base} =7V- 2V; the base level of zero voltage does not assist such hot holes to surmount the oxide barrier. Hence, if we slowly adjust the voltage of the base level, the best injection condition can be estimated to be ΔV_B = 4V which also indicates at least 4V difference from V_{top} to V_{base} which is needed for sufficient hole generation. Fig. 4 shows the V_T versus pulse width

1-4244-0439-X/07/$25.00 © 2007 IEEE

which reaches a stable V_T after 1usec or longer T_2.

B. Injection Scheme for Electrons- FBEI

For electron injection as shown in Fig.5, called forward-bias assisted electron injection (**FBEI**), similar to **PASHEI** in [4], is developed. Initially, the source, drain, and gate are grounded, during time T_1, a negative voltage is applied at the drain, where electrons are generated due to forward bias drain/substrate p-n+ diode. During time T_2, both gate and drain are switched to positive bias, then, electrons are accelerated toward the drain and reach the nitride layer via the vertical field. Fig. 6 shows the V_T with respect to the emitting voltage V_1 during T_1, in which V_1= -1.0V is sufficient to reach a stable V_T. Fig. 7 determines the collecting voltage V_2 during T_2, which reaches the maximum V_T when V_2= 4.5V.

Unlike the conventional hot carrier and FN operation which depends largely on the pulse width (e.g., T_2 in Fig. 1), however, FBEI is less dependent on the pulse width since the injection of electrons is taken place during the transition from forward bias substrate/drain p-n+ diode (T_1 in Fig. 5) to reverse bias substrate/drain p-n+ diode (T_2 in Fig. 5). In Fig. 8, the total erase time is set to 1ms while varying the pulse period (T_1+ T_2). It can be observed that higher V_T is achieved with increasing pulse counts. In addition in Fig. 9, the threshold voltage keeps the same with different pulse period while setting the same pulse count. From Figs. 8 and 9, the threshold voltage is dependent on the pulse count but not on the pulse period. Fig. 10 shows the transient characteristics by using the scheme in Fig. 5, in which higher gate voltage gives faster injection speed as expected.

3. Performance and Reliability of 1bit/cell

In the first experiment, to demonstrate a **negative logic operation** by using BB-HHI for program to the logic-1 state and FBEI for erase to the logic-0 state, Fig. 11 shows the endurance test up to 10^4 P/E cycles, in which no window closure can be seen and a larger operation window around 3.2V can be achieved. Fig. 12 shows the data retention at room temperature and a fairly large V_T window (~2.8V) can be maintained after 10 years. Fig. 13 shows the data retention at T= 125°C, in which V_T= 2V window is maintained after 10 years. These will all easily meet the industry requirement. For the **positive logic**, these schemes are also feasible by using FBEI for programming and BB-HHI for erase, with a sacrifice of the programming speed.

4. Performance and Reliability of 2bit/cell

In the second experiment, combination of FBEI for program and BB-HHI for erase is also feasible for 2-bit operation. Fig. 14 shows the I_D-V_G characteristics of forward and reverse read, comparing to PASHEI, FBEI and CHEI, in which FBEI has largest V_T window between programmed bit-1 and erased bit-2 for all three cases.

Figs. 15 and 16 show the fixed top level charge pumping (CP) measurement with drain pulses [5], where we can investigate a "pump" after programming with V_D= 0 for CHEI and FBEI, respectively. Furthermore, the "pump" indicates a region where the charges are stored in the nitride layer, and a similar result can be seen after CHEI and FBEI by using this CP test. In the meantime, the drain voltage is raised in order to cover the "pump", where FBEI has smaller V_D (~1V) to cover the "pump" compared to CHEI (>1.2V). Consequently, it seems that FBEI has a much localized behavior near the drain side as illustrated from Figs. 14-16. Fig. 17 shows the extraction of V_T profile by the CP technique [6]-[8], which is consistent with the results of Figs. 14-16. Fig. 18 shows the charge density profile extracted from Fig. 17 by the expression:

$$Q_N = C_{ox}\left(\frac{V_{PGM} - V_{vir}}{q}\right)$$

where the charge distribution of FBEI exhibits more localized behavior than CHEI. This is crucial since the more localized charge is, the better is the cell retention. In Fig. 19, two memory cells are programmed by CHEI and FBEI once and the data retention is measured at 80 °C, in which FBEI shows much better retention ability than CHEI after 10 years, since it covers a smaller portion of the charge stored inside the channel.

Fig. 20 shows V_T as a function of V_{read}, where a smaller V_{read} is needed for FBEI to reach 1V window between programmed bit-1 and erased bit-2. The programming speed for bit-1 and bit-2 is shown in Fig. 21. First, bit-1 is programmed near the drain side, after that bit-2 is programmed near the source side without erasing bit-1. Although CHEI has faster bit-1 programming speed, longer bit-2 programming time is needed. In Fig. 22, in order to find out the reason in Fig. 21, bit-1 is programmed by CHEI first and bit-2 is performed by FBEI later in the same memory cell. It was found that 1ms programming time is still unchanged, since the charged bit-1 will affect the acceleration of the electrons under the channel by using CHEI to program bit-2.

The cycling characteristic is shown in Fig. 23 with the repeated cycling by programming bit-1→ programming bit-2 → erase, and good endurance can be achieved. The data retention for cells before and after the cycling is shown in Fig. 24, where an acceptable value of the programmed state is achieved after 10 years. In short, for the 2bit/cell operation, the new scheme exhibits much better retention and endurance characteristics.

In summary, reliable ultra-low voltage and high speed operation schemes have been demonstrated for NOR-type SONOS flash memory applications. Two different schemes have been developed for 1bit/cell and 2bit/cell operations respectively. For **1bit/cell operation**, a lowest 5V operation voltage **reported-to-date** and **very fast speed** (<1μsec) has been achieved for a thick tunnel oxide (50A°) such that excellent data retention can be preserved. In other words, tunnel oxide thickness is not an issue for the future scaling using this low voltage operation scheme. **For 2bit/cell operation**, the charge storage in FBEI is more localized, as verified from the charge pumping analysis, such that this new scheme shows much better retention and endurance comparing to the conventional CHEI ones.

Acknowledgments- This work was supported in part by the National Science Council, Taiwan, under contract *NSC95 -2221-E009-280* and *MoE ATU program, National Chiao Tung University.*

References

[1] B. Eitan et al., in ***IEEE EDL,*** Vol. 21, p. 543, 2000.
[2] T. Y. Chan et al., ***IEEE EDL,*** V0. 18 , p. 93, 1987.
[3] M. K. Cho et al., ***IEEE EDL***, Vol. 24, p. 260, 2003.
[4] Z. Liu et al., in ***Symp. on VLSI Tech.,*** p. 17, 1999.
[5] S.-H. Gu et al., in ***IEEE ED***, Vol. 53, p. 103, 2006.
[6] H. Pang et al., ***Proceedings of ESSDERC,*** *Grenoble, France,* p. 2.C.4, 2005.
[7] M. Rosmeulen et al., ***Solid-State Electronics,*** p. 1525, 2004.
[8] C. Chan et al., in ***IEEE ED***, Vol.45, p.1524, 1998.

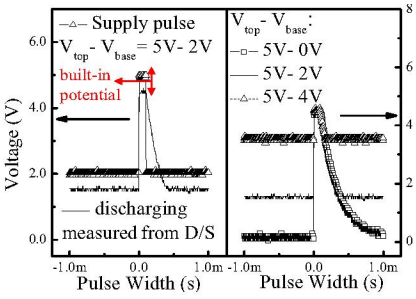

Fig. 2 (a) The oscilloscope trace of the amplitude of the voltage and its decay due to discharge at S/D for the case of V_{top}- V_{base}= 5V- 2V. (b) The discharging condition for three different V_{top}- V_{base} cases.

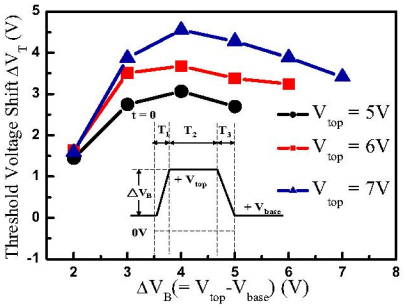

Fig. 3 Relationship between threshold voltage and the difference of bulk bias (V_{top}-V_{base}).

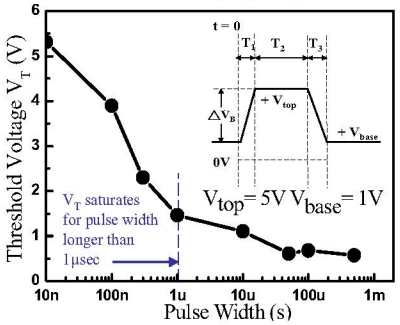

Fig. 4 Threshold voltage window saturates with increasing pulse width after 1μsec.

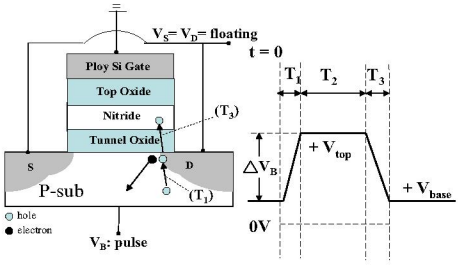

Fig. 1 Experimental set up and timing diagram for BB-HHI scheme.

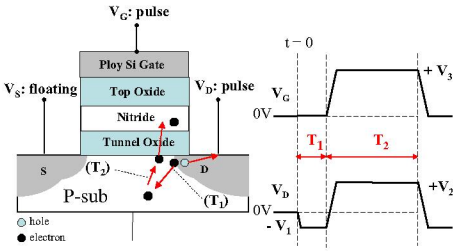

Fig. 5 Experimental set up and timing diagram for FBEI (forward-bias electron injection).

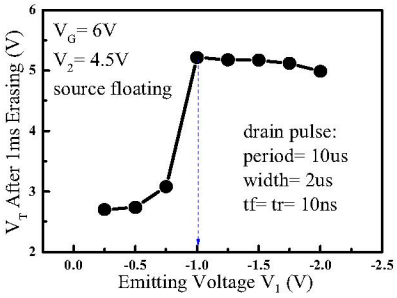

Fig. 6 Changing the emitting voltage at T_1 to find the best bias condition for injection in Fig. 5.

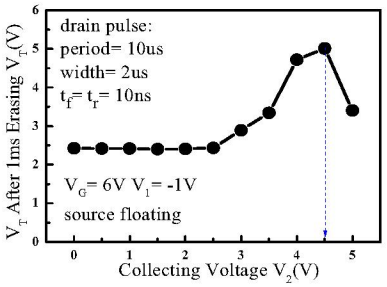

Fig. 7 Changing the collecting voltage to find the best condition for V_2 in Fig. 5. After 4.5V, V_T is dropping since the lateral field dominate the mechanism of motion of electrons.

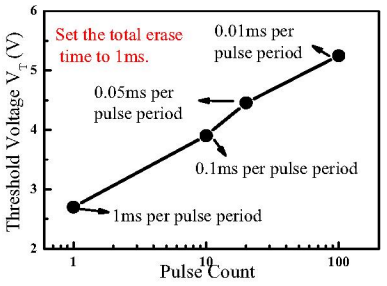

Fig. 8 The independence of V_T on the pulse period for FBEI injection in Fig. 5.

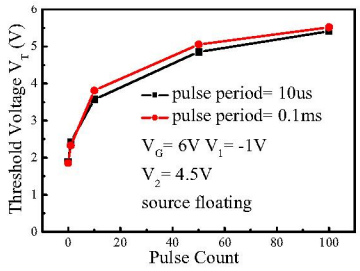

Fig. 9 Different pulse period but with the same pulse counts shows the dependence of V_T on the pulse count for FBEI injection in Fig. 5.

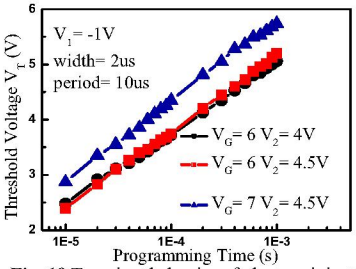

Fig. 10 Transient behavior of electron injection for various gate and drain biases in Fig. 5.

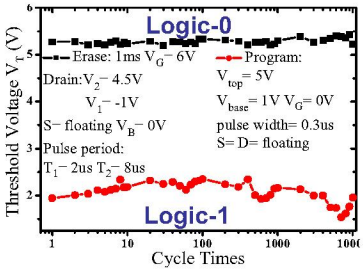

Fig. 11 The endurance characteristics after long term P/E cycling. A large window 3.2V can still be maintained after the cycling for 1bit/cell operation.

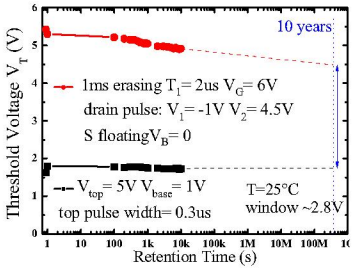

Fig. 12 Data retention characteristic after 10^4 cycling at room-temperature, where 2.8V window can be maintained after 10 years.

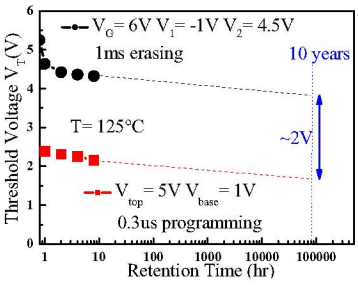

Fig. 13 Data retention characteristic after 10^4 P/E cycling at 125℃. A 2V window can be maintained after 10 years.

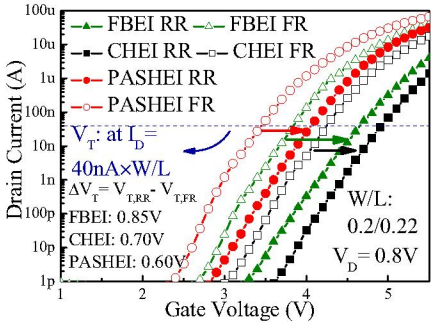

Fig. 14 I_D-V_G measurement investigated from forward read (FR) and reverse read (RR) for FBEI, CHEI, and PASHEI.

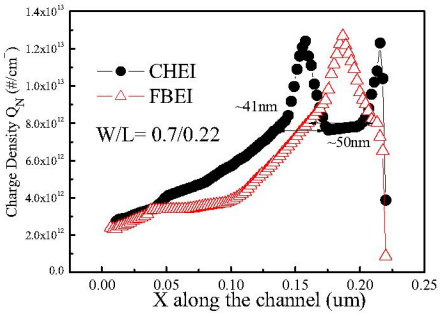

Fig. 18 The charge density distribution along the channel. Note that FBEI is more localized toward the drain.

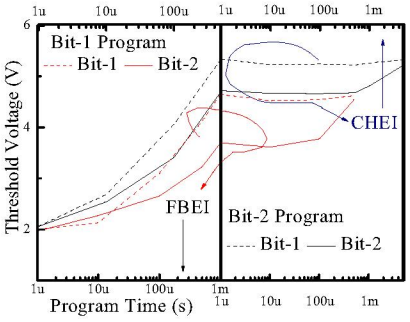

Fig. 21 The programming speed for Bit-1 and Bit-2, and CHEI shows much longer Bit-2 programming time.

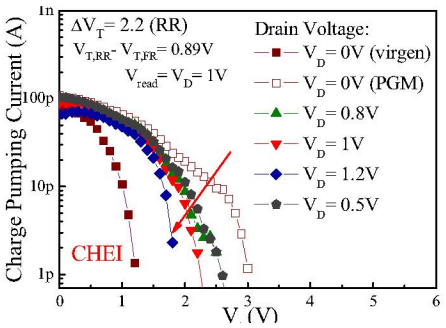

Fig. 15 The fix top level CP (V_D= 0V) and fixed top level CP with drain pulses technologies for CHEI writing.

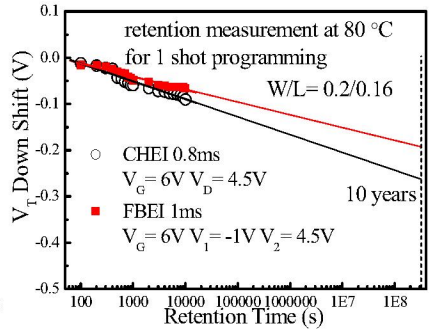

Fig. 19 V_T downshift during 10000s retention time at 80°C, where CHEI shows worse retention ability.

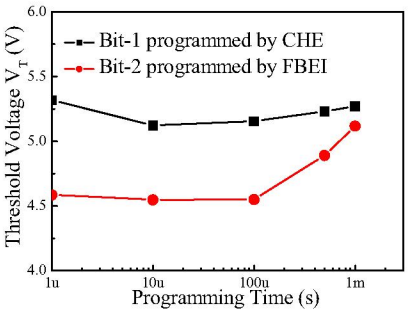

Fig. 22 The programming speed for Bit-2 by FBEI while Bit-1 has been performed by CHEI.

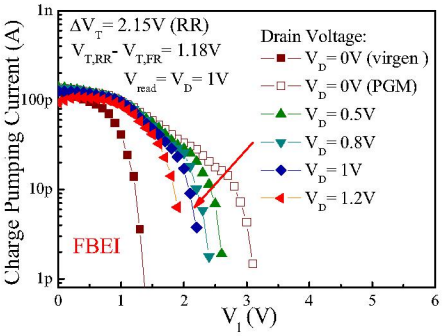

Fig. 16 The fix top level CP (V_D= 0V) and fixed top level CP with drain pulses technologies for FBEI writing.

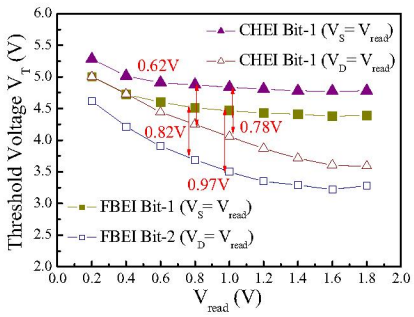

Fig. 20 V_T as a function of V_{read} for Bit-1 and Bit-2. A smaller V_{read} for FBEI to obtain a larger window between the two bits.

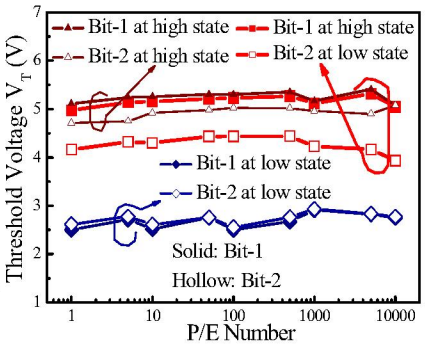

Fig. 23 Endurance characteristics of two-bit-per-cell application. The cycle is repeated by programming Bit-1=> programming Bit-2=> erase.

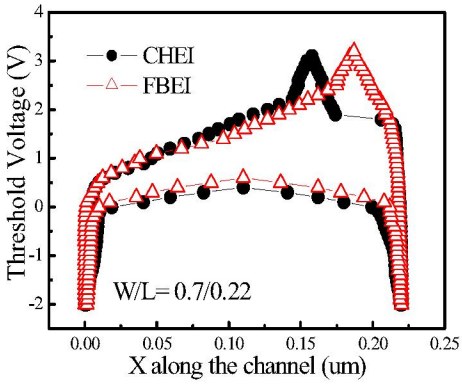

Fig. 17 The V_T distribution for both cases of FBEI and CHEI.

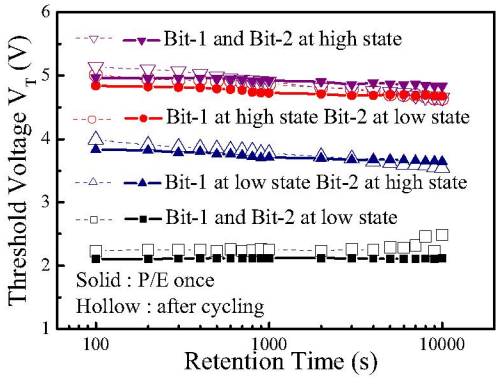

Fig. 24 The retention behavior of Bit-1 in one and 10^4 P/E cycled cells at room temperature.
Solid symbol is for one time P/E cycle and open symbol is for 10^4 P/E cycling.