Extended Abstracts of the 1999 International Conference on Solid State Devices and Materials, Tokyo, 1999, pp. 528-529

A Novel High-Density and High-Speed NAND-Type EEPROM

Frank Ruei-Ling Lin, Mou-Lin Lee, Shih-Yun Lin, Chen-Hao Boe, Charles Ching-Hsiang Hsu, Ching-Pen Yeh*, Po-Hao Wu*, Wen-Sen Wang*, and James Ni*

Microelectronics Lab. Semiconductor Technology and Application Research Group (STAR), Dep. of Electrical Engineering, National Tsing Hua University, Hsinchu, 300 Taiwan, ROC. TEL/FAX/EMAIL: +886-3-515131 ext.4036/+886-3-5718300/chhsu@ee.nthu.edu.tw

**Analogy Technology Corporation (ATC), No. 7, Industry E. Rd. VII, Science-Based, Industrial-Park, Hsinchu, 300, Taiwan, ROC.*

1. Introduction

EEPROM's have been widely used in programmable IC's, such as PC BIOS, monitors, printers, network hubs, etc. The developing trend of EEPROM's directs toward low-cost and high-speed access. However, the cell size of EEPROM cell consisting of a select transistor and a stacked-gate transistor with tunnel oxide is too large for high-density applications. In addition, the write speed can not be improved too aggressively because the tunneling current is inherently small and large tunneling current density will result in premature threshold-voltage window closing between state 0 and state 1.

In this work, a new single-transistor NAND-type EEPROM, which uses (1) the novel **I**n-**C**ell **T**emporary **S**torage **(ICTS)** technique and (2) the novel **M**ultiple-**W**ordline **P**arallel **P**rogramming **(MWPP)** method, is proposed to shrink the unit cell size, reduce the input buffer size, and shorten the overall programming time.

2. Experimental Device Structure

The experimental device structure is shown in Fig. 1, where the EEPROM cells (Fig. 1a) in series compose the NAND array strings (Fig. 1b). It is noted that the averaged cell-size is reduced to half since the select transistor is removed from the unit cell and a long string of cells share two select transistors. The cell consists of 10nm tunnel oxide upon the buried drain diffusion, 35nm gate oxide, 300nm self-aligned thermal oxide upon source/drain buried diffusions, and effective 20nm ONO inter-poly dielectric.

3. Novel In-Cell Temporary Storage (ICTS) Technique and Multiple-Wordline Parallel Programming (MWPP) Method

Conventional NAND EEPROM operations stored information in the buffer latches and supplied constant bitline voltages, such as 0V or 5V, during programming according to the latched states. A string of NAND cells were programmed from the bottom cell to top cell sequentially. The novel ICTS technique, however, stores mass input-data directly inside the selected cells, instead of peripheral latches, by inverting different channel charge under each selected cells. To eliminate charge sharing (cross talk of adjacent cells), the NAND string is divided into two groups of even and odd cells, as shown in Fig. 1b. For example, when even cells are inverted with different charges, odd cells are shut off to block the charge sharing. In the following, the procedure of programming even cells is demonstrated. Programming odd cells is in a similar manner. The operation voltages used below are summarized in table1.

(a) In-Cell Temporary Storage: Referring to the timing diagram in Fig. 2 and cell operation in Fig. 3, data (BL voltages) from external system, such as CPU, are transferred directly to selected even cells (WL0, WL2, WL4, and WL6 in Fig. 1b), consecutively. Even WL voltages are raised to 8V (referred as 1st step gate voltage V_{G1} in Fig. 2) to invert different charges according to different BL voltages. Odd cells (WL1, WL3, WL5 and WL7 in Fig. 1b) are shut off sequentially, immediately after selected even cells invert sufficient charges, as shown in Fig. 2. It is noted that the even cell (WL0) has to wait for a period, defined as the temporary storage time T_{TS}, that other even cells (WL2, WL4, and WL6) need to finish storing inversion charge. During T_{TS}, the charge stored in the cell (WL0) has to remain unchanged for latter accurate programming.

(b) Multiple-Wordline Parallel Programming: After all data are transferred to even cells, all even wordlines are raised up to 20V (referred as the 2nd step gate voltage V_{G2} in Fig. 2) to transfer the inverted charges into the floating gate by FN tunneling. More wordlines (even cells in other NAND strings) incorporated into parallel programming will reduce the effective programming time per cell efficiently.

To assist the understanding of programming details, the cell operation is divided into three time frames in accord with the timing diagram in Fig. 2: **(1) data transfer step, (2) temporary storage step,** and **(3) program step,** as shown in Fig. 3. Immediately after data transferring in Fig. 3a, the adjacent unselected cells are shut off to float the source/drain of selected cells (Fig. 3b and Fig. 3c), e.g. odd cells are shut off to isolate even cells. The voltages of source (V_S), drain (V_D), and channel (V_{CH}), which adds a new floating node in addition to the floating gate, are floating at different potentials according to previous input voltages. It is noted that each cell, which behaves as a data latch, memorizes data in terms of inverted charge (Q_{CH}) in this work, not bitline voltages (V_{BL}) of peripheral latch output in conventional approaches. During programming step, the increased gate voltage (V_{G2}=20V) couples voltages not only to the floating gate but also to the channel. The V_{CH} is estimated to be 3V if large Q_{CH} exists in the channel and the considerable FN tunneling current transfers partial Q_{CH} into the floating gate, as shown in Fig. 3d. On the other hand, no charge transfer occurs if negligible Q_{CH} exists in the channel and the V_{CH} is estimated to be coupled to 8V.

4. Programming Characteristics

Programming occurs only sufficient electric field is established across the tunnel oxide. Based upon the previous analysis, programming or not is determined by the channel charge Q_{CH}, which is related to the gate voltage (V_{G1}) and drain voltage (V_D) during data transfer. Therefore, the programmed threshold voltage (V_{TH}) is indirectly related to V_{G1} and V_D, as shown in Fig. 4. Small V_D or large V_{G1} inverts larger Q_{CH} and results in larger threshold voltage (V_{TH}) shift. Adjusting V_D during data transfer can effectively modulate the programmed V_{TH}'s linearly, which is suitable for analog storage. During data transfer and temporary storage periods, the FN tunneling is inhibited. As shown in Fig. 5, tunneling current appears prematurely when V_{G1}>9V. Thus V_{G1}=8V is selected for accurate operation. The programming trend is illustrated in Fig. 6, where different 2nd step gate voltages (V_{G2}) are used. Larger V_{G2} will increase the V_{TH} shift for programmed state, as shown in Fig. 7. It is noted that 1V difference of V_{G2} only leads to 0.15V of V_{TH} shift because the gate-coupled channel potential reduces the electric field for tunneling.

Another concern is for how long the cells in the temporary storage state can wait. The temporary storage should maintain its data until all even or odd cells latch individual data. Undesired leakage may cause malfunctions. As an indirect evidence in Fig. 8, if the programmed V_{TH}'s for both program and erase state are not altered by longer waiting time, the temporary storage is believed to be reliable and the stored data is inferred to be unchanged. No V_{TH} change is observed up to 500ms at room temperature, which suggests that, during this period, a considerable amount of data can be transferred into even/odd cells before multiple-wordline parallel programming. Another preferable feature is the highly endurable characteristics, as shown in Fig. 9, that nearly no degradation is observed up to 10^6 cycles.

4. Conclusions

A high-density and high-speed NAND EEPROM using ***In-Cell Temporary Storage (ICTS) technique*** and ***Multiple-Wordline Parallel Programming (MWPP) method*** are, for the first time, demonstrated. Cell size is reduced to half by removing select transistors. The excellent programming characteristics and the reliable temporary storage give the promising features for future high-density and high-speed EEPROM applications.

References

[1] K.-D. Suh et. al., Journal of Solid State Circuits, p. 1149 (1995).
[2] R. Shirota et. al., IEDM Tech. Dig. p. 103 (1990)

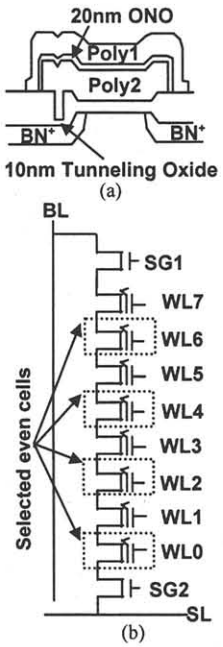

Fig. 1 (a) Cell cross section, (b) NAND array is divided into even/odd cells.

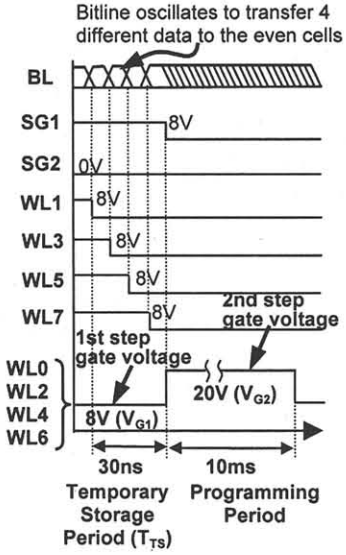

Fig. 2 Timing diagram of high-speed multiple-wordline programming. During temporary storage period, even cells function as data latches. After latching, even cells are programmed simultaneously.

Table. 1 Proposed operation voltages

	Selected WL	Unselected WL	BL	SL	P-Sub
Data Transfer	8V	8V	0V/5V	0V	0V
Temporary Storage	8V	0V	Next Data	0V	0V
Program	20V	0V	Float	0V	0V
Erase	0V	15V	15V	0V	0V
Read	2V	5V	1V	0V	0V

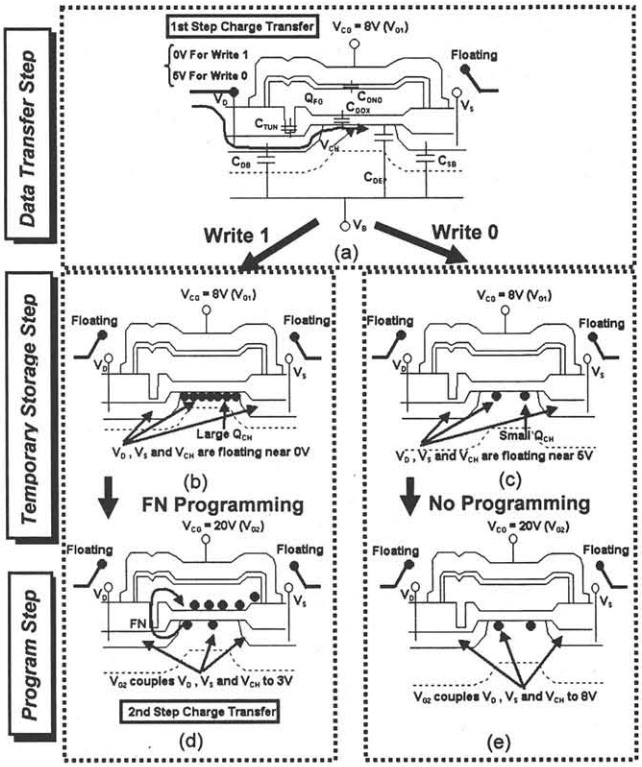

Fig. 3 Three-step programming procedure: (a) ***the 1st step of data transfer***: Different bitline voltages invert different channel charges. (b)(c) ***the 2nd step of temporary storage***: Cells which have stored data in the channel have to wait for data-transfer of other cells. FN tunneling is inhibited. (d)(e) ***the 3rd step of programming***: Charges are tunneled into the floating gate if abundant charges existing in the channel inversion. Otherwise, no tunneling occurs.

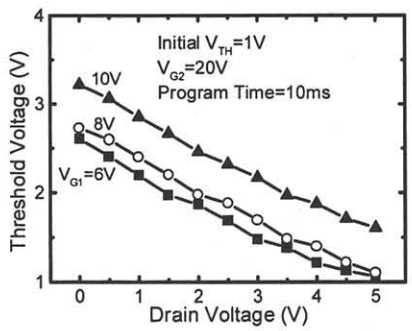

Fig. 4 Threshold voltage vs. drain voltage for different 1st step gate voltage (V_{G1}).

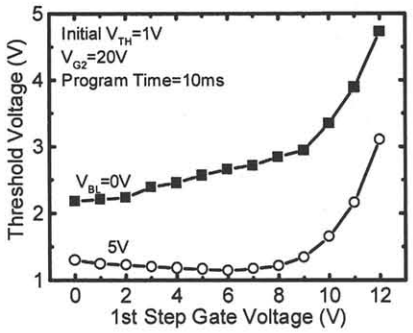

Fig. 5 Threshold voltage vs. 1st step gate voltage (V_{G1}) for programming and non-programming

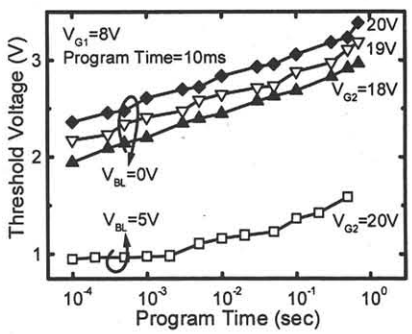

Fig. 6 Programming trend for different programming and non-programming.

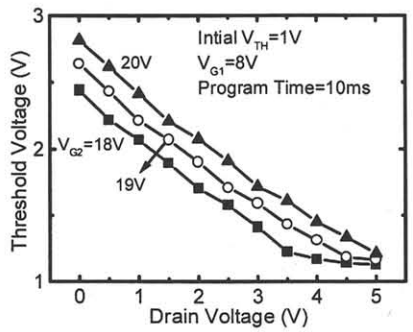

Fig. 7 Threshold voltage vs. drain voltage for different 2nd step gate voltage (V_{G2}).

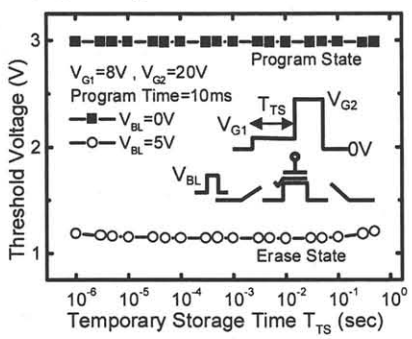

Fig. 8 Programmed threshold voltage vs. temporary storage time (T_{TS}).

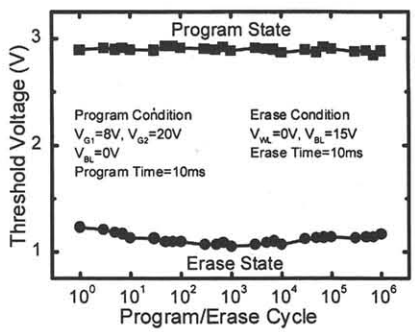

Fig. 9 Endurance characteristics. No window closing is observed.