A Multilevel Read and Verifying Scheme for Bi-NAND Flash Memories

Chiu-Chiao Chung, Hongchin Lin, Member, IEEE, and Yen-Tai Lin

Abstract—A multilevel sensing and read verifying circuit is proposed for Bi-NAND (Buried bit-line NAND) type flash memories. The Bi-NAND technology employs the negative programmed threshold voltage to facilitate the multilevel storage with lower program/erase bias and programming disturbance. The sensing circuit utilizes an advanced cross-coupled sense amplifier to achieve excellent immunity against mismatch effect and reduction of power consumption. As well, it acts as data latch during multilevel sensing and verifying operations. By comparing to the conventional and simultaneous verifying circuits, the proposed scheme with dichotomous architecture simplifies the verifying circuit and speeds up verification process for multilevel operation. By adding only one latch and a pair of switches, the circuit can be easily expanded for one more bit per cell.

Index Terms—Multilevel, Bi-NAND, flash memory, negative programmed threshold voltage, mismatch, read verifying, dichotomous.

I. INTRODUCTION

T HE mass storage applications such as video and image processing have accelerated the development of high-density and high-performance flash memories. For requirement of high density, the serial configured NAND flash memory is the most popular one for the lowest bit cost and high data throughput. However, since the sequential programming was commonly employed, the pass gate voltage disturbance during programming was inevitable. As to the cell technology, further enhancement of density has led to multilevel operation [1], [2]. It is important for multilevel cells that the threshold voltage deviation of the programmed memory cells has to be well controlled for operation of every threshold (V_{TH}) level. Therefore, the more efficient solution may be desired.

For flash memories, comparison of current difference between the flash cell and the reference cell is the direct and fast method to read the data. A new Buried bit-line NAND (Bi-NAND) [3] flash memory has been presented to lower the program/erase voltage and to alleviate the gate disturbance. For the Bi-NAND flash memory arrays, most of the sensing circuits developed for the conventional flash memory cells are

Y.-T. Lin is with eMemory Technology Inc., Hsinchu 300, Taiwan, R.O.C. Digital Object Identifier 10.1109/JSSC.2007.894822

not appropriate. Since those sense amplifiers were designed for draining cell current at the drain node of the flash cell, their bit-lines were usually pre-charged to high before sensing. That also increases power dissipation. On the contrary, the current of Bi-NAND cell flows to the sense amplifier at the source node, thus the bias at the bit-line source has to be low enough for the cell current flowing to the sense amplifier. To comply with these restrictions, a new sense amplifier (NSA) [4] that senses the current difference using the advanced cross-coupled structure by connecting the gates of the clamping MOS transistors to the cross-coupled nodes has been proposed to improve the mismatch characteristics and reduce the power consumption without sacrifice for sensing time.

The verifying circuit is usually required to limit the $V_{\rm TH}$ distribution of programmed cell while using tunneling mechanism for programming. For multilevel operation, V_{TH} distribution has to be restricted more precisely. A purely sequential program and program-verify approach [5] needs at least one program/ verify cycle in each V_{TH} level, so it is a time consuming process. A selective verifying circuit [6] switches the states to compare the data stored in the data buffers (DBs) with the programmed data by scanning the selected word-line voltages. Another verifying scheme employs the word-line sweeping read (WSR) [7] technique to reduce the circuit complexity. The read/verify control circuit utilizes various word-line voltages according to the control signals to verify the data during the verifying period for a multilevel flash cell. However, those circuits need more complicated control signals and procedures to do verification, especially for more bits per cell.

The first binary search sensing scheme (BSSS) [8] was introduced for multilevel storage by using two sense amplifiers connected in series to sense a single cell, and a reference select logic to enable one of three reference cells for comparison. However, the number of sense amplifiers is considerable as the bit number increases. The proposed multilevel sensing and read verifying scheme for Bi-NAND flash memories exploits dichotomous sequential architecture with the current-mode NSA for multilevel data sensing and read verification. The proposed scheme reduces the number of program/verify cycles to improve verification efficiency. Since the upper bit of the program data is checked before the lower bit by the dichotomy method, the verify/program sequence is shortened.

The multilevel cells and circuit operations for Bi-NAND are described in the next section. Section III explains the new read/ verifying circuit in comparison with the selective verifying and word-line sweeping read verifying schemes. Section IV demonstrates the simulation and measurement results. The final part is the conclusion.

Manuscript received December 13, 2005; revised November 15, 2006. This work was supported by National Science Council of Taiwan (NSC 94-2215-E-005-013 and NSC-94-2220-E-005-005).

C.-C. Chung is with the Department of Electrical Engineering, National Chung-Hsing University, Taichung 402, Taiwan, R.O.C.

H. Lin is with the Department of Electrical Engineering and Graduate Institute of Communication Engineering, National Chung-Hsing University, Taichung 402, Taiwan, R.O.C. (e-mail: hclin@dragon.nchu.edu.tw).



Fig. 1. (a) Cross-sectional view of the Bi-NAND memory cell at the cut lines of BL and WL. (b) Schematic of Bi-NAND type cell array and the top view of layout.

TABLE I TYPICAL OPERATING CONDITION FOR THE BI-NAND FLASH CELL (UNIT IN VOLTS)

Operation	SGB	Bit line		Word line			SGS	Source
		select	unselect	select	unselect	pass	200	line
Program	9	5, 6, 7	0	- 9	0	0	Float	Float
Erase	- 6	Float	Float	9	Float	- 6	- 6	- 6
Read	3.3	0	1	0	0	5	3.3	1

II. MULTILEVEL CELLS AND CIRCUIT OPERATIONS

A. Bi-NAND Flash Memories

The Cross-sectional views of the Bi-NAND memory cell at the cut lines of BL and WL are shown in Fig. 1(a), and cell array with the top view of the Bi-NAND memory array is illustrated in Fig. 1(b) [3]. It employs buried bit-line with isolated "shallow P-well" structures to realize the lower program/erase bias for Fowler–Nordheim (FN) tunneling. Due to the divided voltage between word-line (WL) and buried bit-line (BL), the disturbance is also minimized.

The typical operation conditions for the multilevel Bi-NAND cell are listed in Table I [3]. Programming (PGM) and erasure (ERS) for Bi-NAND cell can be achieved in 1 ms and 10 ms, respectively. The cell can be programmed to different threshold

voltages by varying the bit-line voltage and complying with program time, as shown in Fig. 2(a) [3]. Then the cell currents refer to different threshold voltages are driven to the sense amplifier by the source-line voltage. The V_{TH} window for the multilevel programmed and erased state of a Bi-NAND cell is illustrated in Fig. 2(b). The pass gate disturbance is also alleviated due to the negative programmed V_{TH} . Moreover, since the negative programmed V_{TH} helps the conductivity of the cell string during read operation, the V_{TH} windows for different levels are much less influenced by the pass gate voltage as the conventional NAND cells suffered. Therefore, Bi-NAND cells are more promising for multilevel storage with bit-by-bit programming verification.

B. Sense Amplifier

The conventional sense amplifiers were usually designed for draining cell current at the drain node of a flash cell, thus the bit-lines were usually pre-charged to high before sensing. Furthermore, the sense amplifier drains the current of the Bi-NAND flash cell at the source node, so the bias at the source node (bit-line) has to be low enough for the cell current flowing into the sense amplifier.

Since the flash cells are connected in series and biased at 1 V at the source line (SL), the current flows into the sense amplifier at the bit line of the flash cell. Therefore, the bias at the bit line must be close to zero to comply with the requirement. This



Fig. 2. (a) Programming and erasure characteristics [3]. (b) $\rm V_{TH}$ window for programmed and erased states.

new operation makes most of the sense amplifiers designed for the conventional flash cell arrays inappropriate for the new cell array.

The sense amplifier utilizes the advanced cross-coupled structure [4] by connecting the gates of the clamping MOS transistors to the cross-coupled nodes. Since the currents from the selected cell and the reference cell slightly charges the drains before sensing, this improves the mismatch characteristics and reduces the power consumption without sacrificing the sensing time.

Differing from the BSSS, the dichotomous sensing approach for 4-level storage utilizes a sense amplifier and a latch for current sensing instead of two sense amplifiers for voltage sensing. According to the data to be programmed and the threshold voltages of the memory cell shown in Fig. 2(b), the upper bits with threshold voltages of V_{T1} and V_{T2} are "1" and those of V_{T3} and V_{T4} are "0". Therefore, the reference level corresponding to the middle of V_{T2} and V_{T3} is used to determine whether the upper bit is "1" or "0". Once the upper bit is known, the reference level for the lower data bit either the middle of V_{T1} and V_{T2} or the middle of V_{T3} and V_{T4} can be determined.

The reference generator produces three reference currents, I_1 , I_2 , and I_3 , which are the lower, middle, and higher reference levels, respectively. The reference generator is initially set to I_2 level. At first, the sense amplifier compares the cell current with I_2 level. If the upper data bit is "1" (or "0"), the reference level is switched to I_3 (or I_1). The sensing operation is accomplished by comparing the current differences between the cell current and the reference current.

C. Read Verifying Circuit

The key component of verifying circuit combining with a sense amplifier is illustrated in Fig. 3(a). Transistors M7 and M8 acted as switches are controlled by the outputs of sense amplifier. Transistors M9 and M10 are controlled by the control signals CV2 and CV1, respectively. Two series switches (M7/M9 and M8/M10) are connected in parallel for data checking and verifying. Before checking/verifying operation, "Ver" is precharged to high. "Ver" is then discharged to "zero", either the data to be programmed is zero or the cell is programmed to the target level.

First, the control signal "CV1" is activated for checking if the upper bit of programmed data is "0" or not. If the upper bit to be programmed is "0", "Ver" is discharged to zero immediately. If "Ver" signal remains high, the programming operation of the corresponding bit is performed. As to the verifying operation, after the control signal (CV2) is activated, the voltage at node "Ver" is discharged to low if the cell is programmed correctly. Otherwise, the node "Ver" keeps high if the cell is programmed incorrectly.

After Signal "Ver" becomes "0" for the upper bit verification, the lower data bit can be programmed, and "Ver" is pre-charged again. The data to be programmed stored in the latch is used to determined whether "Ver" should be discharged or not. Then, the upper data bit is transferred to the latch. After the lower bit is programmed, CV2 is activated again for verifying if the programming procedure is completed. It will be repeated until "Ver" switching to "0".

The proposed circuit consists of three parts: a sense amplifier, the verifying circuit, and a data latch. Through Switches T0, the cell bit-line is connected to the sense amplifier at the node of "cin", while the reference circuit is connected at the node of "rin". The sensing outputs or the data in the latch are transferred to the verifying circuit through Switches T1 and T2, or Switches T2 and T3, respectively. The data on Doi can be accessed to the IO bus with appropriate timing controls. Fig. 3(b) illustrates the timing diagram for stages of bit check, equalization, and bit sensing/verifying. The operations of lower bit are similar to those of upper bit except the additional data exchanging.

III. SIMULATION AND MEASUREMENT

A. Simulation Results

The proposed circuit was simulated by using 0.35- μ m technology according to the timing sequence in Fig. 3(b). The device sizes of Wp/Wn are $4\mu/1\mu$ in the sense amplifier, $4\mu/1\mu$ in the verifying circuit, and $2\mu/1\mu$ in the latches, respectively.

The typical bit-line charging time for the upper bit is 50 ns. The simulated results are demonstrated in Fig. 4. With the proposed control timings and generated reference currents ($I_1 = 5 \ \mu A, I_2 = 15 \ \mu A$, and $I_3 = 25 \ \mu A$ as refer to the cell programmed for about 2 ms), the read out data on the node "Vo" for "11", "10", "01", and "00" are shown. The node "Ver" for verifying generates the corresponding results for those data. It is discharged as CV1 activates for data "0" before programming, or discharged as CV2 activates for data "1" after being programmed completely.



Fig. 3. (a) Sensing and verifying circuit diagram. (b) Timing diagram for 2-bit/cell operation.

The upper bit data is read out after equalization at about 100 ns. The verification is completed (discharging to low) after 40 ns. The lower bit data is processed continuously by using the similar procedures to get the result.

B. Measurement Results

The multilevel read and verifying circuit was fabricated by using 0.35- μ m technology. The microphotograph of the sensing



Fig. 4. Simulation results for the cases of 11, 10, 01, and 00.



Fig. 5. Chip microphotograph of the sensing and verifying core circuit.

and verifying core circuit is shown in Fig. 5. The test chip was designed by using the currents generated from the selected cell

and reference cell. Each cell has a resistor of 320 Ω and two parallel capacitors of 2 pF in between to mimic the parasitic effects in the memory arrays equivalent to 2K bit lines with 16 cells per string (16 word lines). The cell currents are obtained by applying 1.5 V to the drains of the selected cell and reference cell with different word-line voltages to the gates of the cells, since the threshold voltage differences are usually measured instead of the current differences for flash memories.

Fig. 6(a)–(c) demonstrates the on-chip measured waveforms at the output pad of node "Vo" for reading data "11", "10", and "01" for 2-bit/cell operations, respectively. The verification results corresponding to the read out data are also demonstrated. Since it is trivial for data of "00", the measurement is not shown.

IV. CIRCUIT EVALUATIONS

A selective verify circuit [6], with the key circuit diagram shown in Fig. 7, utilizes different switch states to compare the data stored in the data buffers (DBs) with the programmed data by scanning the selected word-line voltage from low to high



Fig. 6. Measurement of reading data for 2-bit/cell operations: (a) 11, (b) 10, and (c) 01.



Fig. 7. Core of the selective verify circuit [6].

according to the verifying timing sequence. The gates of the M1 and M2, which act as switches, are controlled by the corresponding outputs of the DL1 and DL2 which hold the upper and lower bits of data to be programmed, respectively. There are two latches in DBs for data storage, and four control signals (S11, S12, S21, and S22) for verification in the case of 2-bit per cell. For *n*-bit/cell operation, it needs *n* data buffers (i.e., *n* latches), 2n control signals, and *n* switches connected in series



Fig. 8. Word-line sweeping read/verify control circuit [7].

for the selective verify circuit. Besides, the circuits for data to read out through the sense amplifier were not given in the paper [6].

The eight-level/3-bit per cell NAND flash memory employing the word-line sweeping read (WSR) technology [7] is shown in Fig. 8. During the read operation, the selected word line sweeps from high to low voltages, and the control signals (Lat0 ~ Lat9) are generated in the corresponding reading periods. For the verifying operation, the selected word line also sweeps from high to low voltages and generates the corresponding control signals. After verification, page check was performed by the wired-OR logic circuit. However, this read/verify control circuit may be too large to be used. When n bits data are stored in a cell, it needs $2^n - 1$ control signals and n-bit latches.

For the proposed multilevel verifying scheme, if one more bit is stored in the cell, it only requires an additional data latch (DL) and switches for data transferring. It is thanks to the sense amplifier, which acts like a latch circuit to hold data after the last sensing period, two control signals (CV1 and CV2), and (n-1)data latches along with (n-1) pairs of parallel switches are needed for *n*-bit/cell operation, as shown in Fig. 9. The new verifying circuit efficiently reduces the numbers of control signals and latches for a multilevel cell (*n*-bit data per cell). The comparison of those verifying circuits is given in Table II.

To evaluate the speed of circuits, the time required for a program/verify cycle of 1-bit cell can be estimated as [9]

$$Tw = T_{\rm SET} + (T_{\rm pulse} + T_{\rm vfy}) \times N_p \tag{1}$$

where T_{SET} is the sum of the precharge time before programming and the discharge time after programming, T_{pulse} is the

Reset -Reset BLi Reference **T**1 T1 Sense Amplifier SGB T2 Т2 Verify Circuit CV2 Ø. SGS Data Latch 1 Ø Memory Array Data Latch 2 Data Latch n-1 Doi Doi

Fig. 9. New verification circuit extended to n bits per cell.

 TABLE II

 COMPARISONS IN CONTROL SIGNALS AND LATCHES FOR n-BIT DATA PER CELL

No. of signal	Selective		Word-line		New		
& latch	Verifying circuit		Sweeping read		Verifying circuit		
bit per cell	Control signal	latch	Control signal	latch	Control signal	latch	
2-bit	4	2	-	-	3	1	
3-bit	6	3	8	3	5	2	
n-bit	2n	n	$2^{n} - 1$	n	2n – 1	n – 1	

program pulse width, and the $T_{\rm vfy}$ is the verification time. The parameter N_p is the number of program pulses.

In the conventional scheme, as the number of bits increases, the number of program/verify cycles is proportional to $(2^n - 1)$. Thus, the program time increases in the same way. The sequential and simultaneous program algorithms for 4-level cells are illustrated in Fig. 10(a) and (b), respectively. For the sequential scheme, the program time required for an *n*-bit cell can be written as

$$Tw(n)_{\text{Sequential}} = T_{\text{SET}} + (2^n - 1) \cdot (T_{\text{pulse}} + T_{\text{vfy}}) \times N_p$$
(2)

For the simultaneous scheme with very complicated control signals, the program cycle time required for n bits/cell can be expressed as

$$Tw(n)_{\text{Simultaneous}} = T_{\text{SET}} + [T_{\text{pulse}} + (2^n - 1) \cdot T_{\text{vfy}}] \times N_p$$
(3)

The simultaneous programming scheme [7] also reduces the verification time $T_{\rm vfy}$ using the precharge capacitive decoupling sensing scheme (PCDSS) [10]. However, to alleviate the array ground line (AGL) noise, metal bypass lines were added every 16 bit-lines. That results in the variations of source voltage.

The proposed dichotomous verifying sequence illustrated in Fig. 10(c) uses the bit checking sequence, in which the first

reference level corresponds to the middle of V_{T2} and V_{T3} , in order to determine whether the upper bit is "1" or "0" before the second bit programming. Then, the second bit verification is followed after the second bit is programmed. Therefore, the proposed verifying sequence for *n*-bit cells compresses the program time as shown below:

$$Tw(n)_{\text{dichotomous}} = T_{\text{SET}} + n \cdot (T_{\text{pulse}} + T_{\text{vfy}}) \times N_p$$
 (4)

V. CONCLUSION

A new multilevel sensing and verifying circuit for Bi-NAND flash memory technology was designed and measured. The advanced cross-coupled sense amplifier complies with low bit-line bias for Bi-NAND flash to lessen the power consumption. For the checking and verifying scheme, the pre-charged verifying circuit is discharged as soon as the bit to be programmed is zero. Otherwise, it will be discharged as long as the bit is correctly programmed. The proposed scheme shows good agreements between simulation and measurement for sensing and verifying operations at supply voltage of 3.3 V. Using the simple read/verifying circuit with the dichotomous approach reduces the cycle time of read or verifying. Since the current mode sense amplifier for each memory cell also acts as the data latch, the read/verifying circuit is even more compact. The circuit can be easily expanded by increasing the number of latches for more bits per cell without increment of sensing amplifiers and verifying circuits.





Fig. 10. Program and verify algorithms for (a) sequential, (b) simultaneous, and (c) dichotomous architectures.

ACKNOWLEDGMENT

The authors would like to acknowledge Chip Implementation Center (CIC) of National Applied Research Laboratories (NARL) of Taiwan for the support in chip fabrication and measurement.

REFERENCES

- [1] H. S. Kim, J. D. Choi, J. Kim, W. C. Shin, D. J. Kim, K. M. Mang, and S. T. Ahn, "Fast parallel programming of multi-level NAND flash memory cells using the booster-line technology," in *Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 1997, pp. 65–66.
- [2] A. Nozoe, H. Kotani, T. Tsujikawa, K. Yoshida, K. Furusawa, M. Kato, T. Nishimoto, H. Kume, H. Kurata, N. Miyamoto, S. Kubono, M. Kanamitsu, K. Koda, T. Nakayama, Y. Kouro, A. Hosogane, N. Ajika, and K. Kobayashi, "A 256 Mb multilevel flash memory with 2 MB/s program rate for mass storage applications," *J. Solid-State Circuits*, vol. 34, no. 11, pp. 1544–1550, Nov. 1999.
- [3] S. Chang, E. Yang, T. Chen, L. Huang, B. Hsu, D. Sung, J.-C. Duh, C.-W. Hung, V. Huang, Y.-C. King, C.-H. Chu, and C. C.-H. Hsu, "New buried bit-line NAND (Bi-NAND) flash memory for data storage," in *Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2003, pp. 95–96.
- [4] C.-C. Chung, H. Lin, and Y.-T. Lin, "A novel sense amplifier for Bi-NOR flash memory," *IEEE J. Solid-State Circuits*, vol. 40, no. 2, pp. 515–522, Feb. 2005.
- [5] T. Tanaka, Y. Tanaka, H. Nakamura, K. Sakui, H. Oodaira, R. Shirota, K. Ohuchi, F. Masuoka, and H. Hara, "A quick intelligent pageprogramming architecture and a shielded bitline sensing method for 3V-only NAND flash memory," *IEEE J. Solid-State Circuits*, vol. 29, no. 11, pp. 1366–1373, Nov. 1994.

- [6] H. Kurata, N. Kobayashi, K. Kimura, and S. Saeki, "A selective verify scheme for achieving a 5-MB/s program rate in 3-bit/cell flash memories," in *Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2000, pp. 166–167.
- [7] H. Nobukata, S. Takagi, K. Hiraga, T. Ohgishi, M. Miyashita, K. Kamimura, S. Hiramatsu, K. Sakai, T. Ishida, H. Arakawa, M. Itoh, I. Naiki, and M. Noda, "A 144-Mb, eight-level NAND flash memory with optimized pulsewidth programming," *IEEE J. Solid-State Circuits*, vol. 35, no. 5, pp. 682–690, May 2000.
- [8] M. Bauer, "A multilevel-cell 32 Mb flash memory," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 1995, vol. 38, pp. 132–133.
- [9] K. Takeuchi, T. Tanaka, and T. Tanzawa, "A Multi-page cell architecture for high-speed programming Multi-level NAND Flash Memories," in Symp. VLSI Circuits Dig. Tech. Papers, Jun. 1997, pp. 67–68.
- [10] J.-K. Kim, "A 120-mm² 64-Mb NAND memory achieving 180 ns/Byte effective program speed," *IEEE J. Solid-State Circuits*, vol. 32, no. 5, pp. 670–680, May 1997.



Chiu-Chiao Chung was born in Taiwan, R.O.C. She received the B.S. degree in electronic engineering from Tam-Kang University, Taipei County, Taiwan, R.O.C., in 1983, and the M.S. degree in electrical engineering from the University of Texas, El Paso, in 1989.

She joined the Department of Electrical Engineering, Nan-Kai College, Nan-Tao County, Taiwan, in August 1990. She transferred to the Department of Computer and Communication in 2002. She is currently pursuing the Ph.D. degree in the Depart-

ment of Electrical Engineering, National Chung-Hsing University, Taichung, Taiwan. Her research involves memory circuit design, and Flash memory technology and device design.



Hongchin Lin received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, R.O.C., in 1986, and the M.S. and Ph.D. degrees from University of Maryland, College Park, in 1989 and 1992, respectively.

From 1992 to 1995, he was a Senior Device Engineer with Integrated Technology Division, Advanced Micro Devices, Sunnyvale, CA. In 1995, he joined the Department of Electrical Engineering, National Chung-Hsing University, Taichung, Taiwan, and was promoted to a full Professor in 2003. Since

September 2006, he is also with the Graduate Institute of Communication Engineering at the same university. His current research interests include VLSI circuit design, semiconductor memory devices and circuits, VLSI implementation of wireless communication systems, and system-on-package technology.



Yuan-Tai Lin was born in Taiwan, R.O.C. He received the B.S. and M.S. degrees in electrical engineering from National Tsing Hua University (NTHU), Hsinchu, Taiwan, in 1981 and 1983.

He joined the ERSO (Electronic Research & Service Organization) of ITRI (Industrial Technology Research Institute) for SRAM/DRAM circuit design since he graduated from NTHU. From 1996 to 1998, he was with Vanguard International Semiconductor Cooperation in Taiwan as a SRAM/DRAM design manager. From 1998 to 2000, he was with Macronix

International Cooperation in Taiwan as a Flash design manager. He is presently with eMemory Technology Incorporation, HsinChu in Taiwan. He is currently responsible for the design of embedded nonvolatile memory including Flash, MTP and OTP products and IP.