第 23 卷第 10 期　　半 导 体 学 报　　Vol. 23, No. 10
2002 年 10 月　　CHINESE JOURNAL OF SEMICONDUCTORS　　Oct., 2002

Novel p-Channel Selected n-Channel Divided Bit-Line NOR Flash Memory Using Source Induced Band-to-Band Hot Electron Injection Programming *

Pan Liyang, Zhu Jun, Liu Kai, Liu Zhihong and Zeng Ying

(Institute of Microelectronics, Tsinghua University, Beijing 100084, China)

Abstract: A novel p-channel selected n-channel divided bit-line NOR (PNOR) flash memory, which features low programming current, low power, high access current, and slight bit-line disturbance, is proposed. By using the source induced band-to-band hot electron injection (SIBE) to perform programming and dividing the bit-line to the sub-bit-lines, the programming current and power can be reduced to 3.5μA and 16.5μW with the sub-bit-line width equaling to 128, and a read current of 60μA is obtained. Furthermore, the bit-line disturbance is also significantly alleviated.

Key words: flash memory; DINOR; band-to-band; SIBE; disturbance
EEACC: 2560A; 1265D
CLC number: TP333.5　　**Document code:** A　　**Article ID:** 0253-4177(2002)10-1031-06

1 Introduction

With the scaling down of the different flash cells, the flash array architecture is found new structures to increase the integrate density and to upgrade cell performance. Several memory array architectures, such as NAND[1], NOR[2], DINOR[3] and DuSNOR[4], have been proposed for single-power-supply flash memories. Among these arrays, DINOR (divided bit line NOR) realizes a high access speed, high reliability and a small chip size simultaneously. The unit of memory arry consists of a select gate and eight memory transistors connected by the third polysilicon layer serving as a sub bit line. The cell uses Fowler-Nordheim tunneling electron to perform programming and erasing operation, which requires relatively high voltage biasing on the terminals of the memory cell and has slow programming speed[1]. Ohnakado *et al*. proposed a p-channel DINOR flash memory using the band-to-band tunneling induced hot electron (BBHE)[5] to perform programming operation, which has the advantages of high efficiency and low voltage of programming[6]. However, the read current of the p-channel memory cell is smaller than one half of the one with n-channel. To overcome these problems, this paper proposed a novel p-channel selected n-channel divided bit-line NOR flash memory using source induced band-to-band hot electron injection to perform programming.

In this paper, we described the proposed PNOR flash memory array architecture and the principle of the source induced band-to-band hot electron injection programming, the programming characteristics, the read characteristics,

* Project supported by the 9th Five-Year Plan of China (No. 97-760-01-01)

Pan Liyang male, was born in 1975, PhD candidate. His research field is on the device, technology and reliability of flash memory.

Received 1 February 2002, revised manuscript received 10 May 2002

and the array performance in details.

2 Array architecture and its operation

A schematic illustration of the proposed p-channel selected n-channel divided bit-line NOR (PNOR) flash memory is shown in Fig. 1. One main bit-line (MBL) is divided into several sub-bit-lines (SBL), which are selected by the p-channel transistors. The memory cell is formed on the p-well, which is nearly the same as the traditional ETOX structure, except an accessorial p^+ region

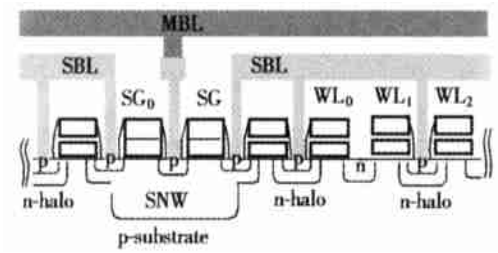

Fig. 1 Schematic illustration of the proposed PNOR flash memory

at the drain end. The impurity concentrations of the n-halo and p^+ region under the gate-to-drain overlap area are $2\times10^{19}cm^{-3}$ and $8\times10^{19}cm^{-3}$, respectively. The tunneling oxide is 10nm and the electric thickness of the oxide/nitride/oxide interpoly dielectric is 20nm. The W/L of the tested cell is 1μm/0.8μm. The select PMOS transistor is formed on the shallow n-well (SNW), and isolated from the n-channel memory cell by a dummy cell, whose source and drain is shorted. In order to increase the read current, the dimension of the select PMOS is enlarged to 2μm/0.6μm. The circuit diagram of the PNOR memory array is shown in Fig. 2. Each SBL connects to 128 memory cells, and each 2048 SBLs selected by the same select transistor is arranged as one sector. The additional size of the select transistor and the isolator is only about twice as the cell size, and the whole array area penalty is about 2%.

Figure 3 shows a schematic model of the memory and the select transistor. The memory cell consists an n-channel stacked-gate transistor and a parasitic p-channel

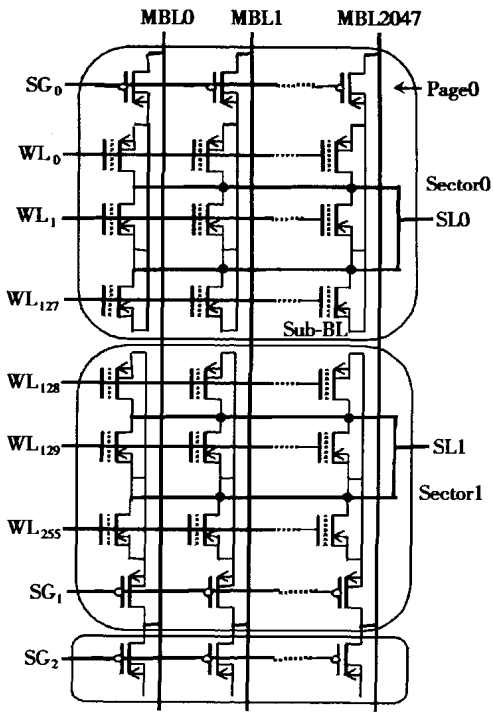

Fig. 2 Circuit diagram of the PNOR flash memory

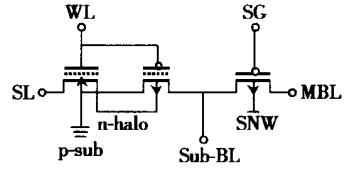

Fig. 3 Schematic model of the memory and the select transistor

stacked-gate transistor, which is made up of the p^+ drain, the p-substrate and the floating n-halo as the source, drain and substrate, respectively.

A novel source induced band-to-band hot electron injection (SIBE) method is proposed to perform the programming operation, where the MBL is set to -3V, the SG is set to -5V, the WL is set to 10V, the SL is set to 2V, the SNW and the p-sub are both set to ground. Under this bias condition, the select transistor and the memory cell are turned on and the parasitic p-channel transistor works as a reverse pn diode. The MBL voltage can be transferred to the SBL and the p^+ drain without loss of

voltage due to the zero turn-on current. Meanwhile, the source voltage can also be transferred to the floating n-halo region in the selected cell, and the drain p^+/n junction is thus negatively biased to nearly - 5V. Due to the high vertical electrical field induced by the high control gate voltage, electrons can be generated on the surface of the p^+ region by band-to-band tunneling[7]. These band-to-band tunneling-induced electrons may also flow into the n-halo region and generate hot electrons by gaining energy through the junction electric field and/or impact ionization. Some hottest electrons can overcome the barrier and be injected into the floating gate through the tunnel oxide with the aid of the positive bias on the control gate. In the unselected cells on the same bit-line, the control gate is set to ground, so the source voltage cannot be transferred to the n-halo region. As a result, the floating n-halo region is clamped to about - 0.5V, and the negative voltage on the drain p^+/n junction is only - 2.5V. Under this condition, the band-to-band tunneling mechanism is hardly induced, and the programming operation is prohibited.

The negative gate source side Fowler-Nordheim tunneling is used to perform the erasing operation, during which, the bit-lines are set to floating, the substrate is set to ground, and the word-lines and the source-line are set to - 8V, and 5V, respectively.

The selected cell is accessed through a select transistor in the PNOR array. During reading, the MBL and the WL are set to 1.8V and 3V, and the SG is set to - 5V to turn on the select transistor and transfer the MBL voltage to the SBL. Because the memory cell and the select transistor are turned on, there is a voltage lowing between the source and drain of the select transistor, which makes the SBL voltage biased to about 1.5V. Under this bias condition, the drain PN junction of the memory cell is positively biased and the voltage of the n-halo region is clamped at about 0.9V. With the access voltage (V_{CG} = 3V), the p-type channel region may be inverted (in the "0" cell) so that the electrons can flow from the source to the n-halo and an access current of about 60μA/μm is read out from the drain. The detailed operation bias of the proposed flash memory array is illustrated in Table 1.

Table 1　Operation bias of the PNOR flash memory

Operation		V_{MBL}	V_{WL}	V_{SL}	V_{SG}	V_{Sub}	V_{SNW}
Write	selected	- 3	10	2	- 5	0	0
	unselected	0	0	2	0	0	0
Erase	selected	Floating	- 8	5	0	0	0
Read	selected	1.8	3	0	- 3	0	1.8
	unselected	0	0	0	0	0	1.8

3　Result and discussion

Figure 4 shows the program drain and floating gate currents with different V_{FG}. By the proposed SIBE programming with drain voltage equaling to - 3V, the drain and the floating gate currents are 3μA/μm and 1nA/μm, respectively, and the programming efficiency can reach 3×10^{-4}. In the unselected cells lined to the selected bit-line (according to the floating gate voltages equaling 0V), the drain leakage current and the floating gate disturb current can be limited to less than 4nA and 5×10^{-14}A, respectively.

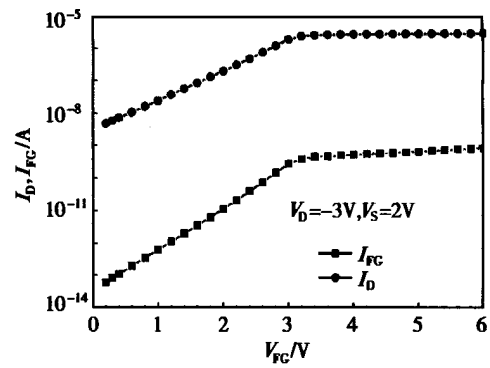

Fig. 4　Program drain and floating gate currents with drain and source equaling to - 3V and 2V, respectively

Considering a flash memory array, the bit-line current is made up of the drain programming current ($I_{d,w}$) of the selected cell and the drain leakage currents ($I_{d,leak}$) of the unselected cells. When the bit-line width is n, the bit-line current and the programming power consumption can be deduced by the following equations.

$$I_{\text{bit-line}} = I_{\text{d,w}} + (n-1)\, I_{\text{d,leak}}$$

$$P_{\text{w}} = (V_{\text{BL}} - V_{\text{SL}})\, I_{\text{d,w}} + (n-1)\, V_{\text{BL}}\, I_{\text{d,leak}}$$

where the $I_{\text{bit-line}}$ is the overall programming current on the bit-line; P_{w} is the programming power consumption; V_{BL} and V_{SL} are bit-line and source-line voltages during programming.

Figure 5 shows $I_{\text{bit-line}}$ and P_{w} with different sub-bit-line widths. Due to the leakage current of the unselected cells, the larger of the bit-line width, the higher the programming current and power. With the DINOR architecture, the main bit-line is divided into short sub-bit-lines, so the programming current and power are significantly reduced. The $I_{\text{bit-line}}$ and P_{w} are 3.5μA and 16.5μW with the SBL width equaling to 128, however, they will be as large as 11μA and 40μW with 2048bit SBL. It is clear that the bit-line leakage current is greatly reduced by using the DINOR architecture.

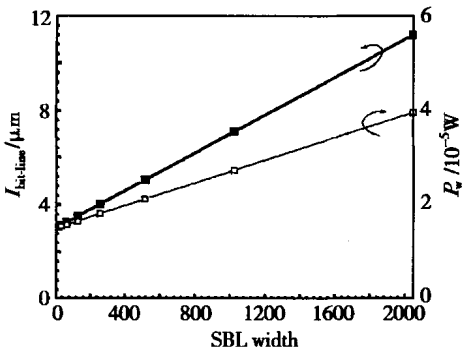

Fig. 5 Calculated $I_{\text{bit-line}}$ and P_{w} as a function of bit-line width

Figure 6 shows the time dependent programming current and threshold voltage shift during programming, where the threshold voltage is tested with the 1μm/0.8μm memory cell, and the floating gate current is deduced from the V_{FG}-t curve and V_{FG}-I_{FG} curve in 1μm/0.8μm dummy cell with contacted floating gate. A programming time of high speed for less than 20μs was obtained with the threshold voltage shift for more than 4V.

Figure 7 shows the read characteristics of the PNOR flash memory, during which the MBL is set to 1.8V, and the SBL voltage is about 1.5V. When the control gate voltage is 3V, the read current can be more than 60μA/μm in the 0.8μm cell, which is twice the access currents in other introduced p-channel flash memory cells.

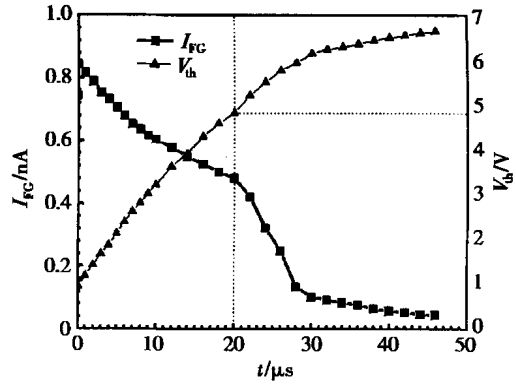

Fig. 6 Schematic evolutions of threshold voltage and floating gate current during programming

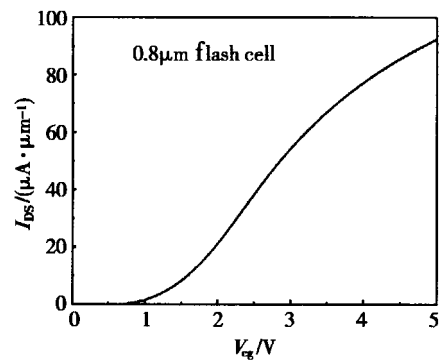

Fig. 7 Read characteristics of the flash cell Read operation is biased at $V_{\text{d}} = 1.5\text{V}$, $V_{\text{cg}} = 3\text{V}$, $V_{\text{s}} = 0$.

The most severe disturbance is the bit-line disturbance during BBHE programming, which can also be remarkably improved by using the SIBE programming method. With the help of the source voltage, the drain voltage can be reduced, and the lower drain voltage can alleviate the bit-line disturbance. Moreover, owing to the zero bias on the control gate for the disturbed cell, the source voltage can not be transferred to the drain p^+/n junction and the band-to-band hot electron's generation and injection are depressed. For further insight into the bit-line disturbance, the programming tolerant time with different programming drain voltages is shown in Fig. 8, in

which the tolerant time is defined as the disturbance time required for 500mV shift of threshold voltage. It can be seen that at least 100s (far longer than 1000 program cycles) of tolerant time is guaranteed by the proposed program under the condition of V_{drain} = - 3V, which illustrates that the proposed PNOR flash memory, combining with the SIBE programming, has a slight bit-line disturbance with the SBL width smaller than 1000.

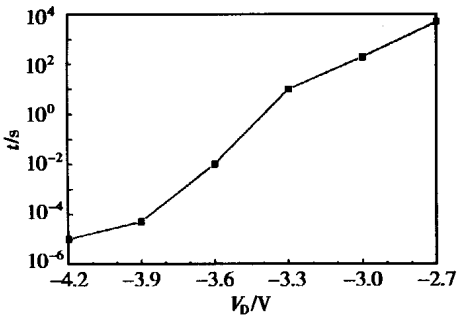

Fig. 8 Programming bit-line disturb tolerant time with different programming drain voltages

4 Conclusion

A novel p-channel selected n-channel divided bit-line NOR flash memory and a source induced band-to-band hot electron injection programming method are described in the paper. The whole array area penalty of the select transistor is only about 2 %. By using the SIBE programming method and the divided bit-line architecture, the programming current and power can be reduced to 3.5μA and 16.5μW with the sub-bit-line width equaling to 128, and a read current of 60μA is obtained. Furthermore, the tolerant time of bit-line disturbance can get up to 100s when the programming drain voltage is - 3V. The tested results show that the proposed PNOR flash memory features a low programming current, low power, high access current, and slight bit-line disturbance.

Acknowledgement The authors would like to express their thanks to Miss Zhang Wei, Fu Yuxiang, Xia Shu, and all other members in the process line of IMETU for their great supports on the technology development.

References

[1] Iwata Y, Momodomi M, Tanaka T, et al. A high-density NAND EEPROM with block-page programming for microcomputer applications. IEEE J Solid-State Circuits, 1990, SC-25(4): 417

[2] Ajika B, Ohi M, Arima H, et al. A 5V only 16M bit Flash EEPROM cell with a simple stacked gate structure. In: IEDM Tech Dig, 1990: 115

[3] Kobayashi S, Nakai H, Kunori Y, et al. Memory array architecture and decoding scheme for 3V-only sector erasable DINOR Flash memory. IEEE J Solid-State Circuits, 1994, SC-29(4): 454

[4] Kim K S, Kim J Y, Yoo J W, et al. A novel dual string NOR (DuSNOR) memory cell technology scalable to the 256M bit and 1G bit Flash memory. In: IEDM Tech Dig, 1995: 263

[5] Ohnakado T, Mitsunaga K, Nunoshita M, et al. Novel electron injection method using band-to-band tunneling induced hot electron (BBHE) for Flash memory with a p-channel cell. In: IEDM Tech Dig, 1995: 279

[6] Ohnakado T, Takada H, Hayashi K, et al. Novel self-limiting program scheme utilizing N-channel select transistors in P-channel DINOR flash memory. In: IEDM Tech Dig, 1996: 181

[7] Su Yu, Zhu Jun, Chen Yuchuan, et al. Charge pumping measurement for determining band-to-band tunneling induced interface damage during erasing operation of Flash. Chinese Journal of Semiconductors, 2001, 22(1): 69

采用源极增强带带隧穿热电子注入编程的新型 p 沟选择分裂位线 NOR 快闪存贮器 *

潘立阳 朱 钧 刘 楷 刘志宏 曾 莹

(清华大学微电子学研究所, 北京 100084)

摘要: 提出一种新型的 PMOS 选择分裂位线 NOR 结构快闪存贮器,具有高编程速度、低编程电压、低功耗、高访问速度和高可靠性等优点. 该结构采用源极增强带带隧穿热电子注入进行编程,当子位线宽度为 128 位时,位线漏电只有 3.5μA 左右,每位编程功耗为 16.5μW,注入系数为 4×10^{-4},编程速度可达 20μs,存贮管的读电流可达 60μA/μm 以上. 分裂位线结构和低编程电压使得该结构具有很好的抗位线串扰特性和可靠性.

关键词: 快闪存贮器; 带带隧穿; 分裂位线 NOR; 源极增强带带隧穿热电子注入; 位线串扰

EEACC: 2560A; 1265D

中图分类号: TP333.5 **文献标识码**: A **文章编号**: 0253-4177(2002)10-1031-06

* 国家九五计划资助项目(No. 97-760-01-01)

潘立阳 男,1975 年出生,博士研究生,主要从事快闪存贮器器件、工艺和可靠性研究.

2002-02-01 收到,2002-05-10 定稿