A low-voltage sense amplifier for high-performance embedded flash memory*

Liu Jiang(柳江)[1], Wang Xueqiang(王雪强)[2], Wang Qin(王琴)[1], Wu Dong(伍冬)[2], Zhang Zhigang(张志刚)[2], Pan Liyang(潘立阳)[2], and Liu Ming(刘明)[1,†]

(1 Institute of Microelectronics, Chinese Academy of Sciences, Beijing 100029, China)
(2 Institute of Microelectronics, Tsinghua University, Beijing 100084, China)

Abstract: This paper presents a sense amplifier scheme for low-voltage embedded flash (eFlash) memory applications. The topology of the sense amplifier is based on current mode comparison. Moreover, an offset-voltage elimination technique is employed to improve the sensing performance under a small memory cell current. The proposed sense amplifier is designed based on a GSMC 130 nm eFlash process, and the sense time is 0.43 ns at 1.5 V, corresponding to a 46% improvement over the conventional technologies.

Key words: sense amplifier; current mode; embedded flash memory; low voltage
DOI: 10.1088/1674-4926/31/10/105001 **EEACC:** 2570

1. Introduction

The rapidly increasing portable electronic equipment market demands high-speed and low-power flash memories embedded in MCU or other microelectronic systems. To satisfy the requirement of very low power consumption in portable applications, the power supply voltage has been scaled down to sub-1.5 V. In embedded flash (eFlash) memory, the read access time is mainly determined by the read path, which is strongly affected by the sense amplifier's performance[1]. One of the main challenges for new generation eFlash memories is to develop a robust and high-speed read circuit with a supply voltage of sub-1.5 V. As a result, the design of a high-speed low-power sense amplifier becomes very critical as the power supply is lowered further.

Traditionally, the sense amplifier is based on a differential voltage amplifier[2]. The input differential voltage is obtained by the current–voltage conversion of the current coming from the selected memory cell and from a reference cell. The read operation is performed by comparing the input voltage, and data is output through the voltage sense amplifier. However, the voltage-mode approach has scaling limitations as the power supply voltage is reduced[3]. The voltage amplifier has the minimum allowed operating supply voltage, which may be larger than the system power supply voltage[4]. Though an improved single-ended sense amplifier using the voltage source follower and auto-zeroing technique is proposed in Ref. [5], it has no clamp circuit, which may result in worse data retention. Moreover, a long sense time is needed to reach the voltage variation in the source follower.

To overcome the above problems, the current-mode approach is proposed[3,4,6–8]. This approach implements a pure current comparison operation. The current of the selected memory cell can be directly obtained by a current mirror, and then, current comparison is performed between the mirrored current and the reference current. The current-mode sense amplifier is able to operate robustly with a nominal supply voltage lower than 1.5 V. However, as the power supply scales down, the threshold voltage of the memory cell cannot be scaled in the same manner since the cell channel doping is optimized for programming and erasing[2]. Hence, the memory cell current ($I_{\rm cell}$) during read operation decreases under lower $V_{\rm DD}$, which leads to a longer sense time and spoils the read performance.

In this work, a high-performance sense amplifier with the offset-voltage elimination technique is proposed to improve sense performance under low memory cell current. The offset-free sense amplifier employs a capacitor for DC isolation and makes the trip point of the sense amplifier irrespective of process variation. Therefore, the sense time can be reduced dramatically.

2. Conventional sense amplifier

A typical current-mode sense amplifier is illustrated in Fig. 1[7]. $I_{\rm cell}$ and $I_{\rm ref}$ are the cell current and the reference current, respectively, $C_{\rm bl}$ is the parasitic capacitance at the bitline, and node Out represents the output logical value. The read operation is divided into two stages: precharge and sense. When the signal Disc = '0' and Prec = '1', the sense amplifier starts the precharge operation which realizes fast precharging and maintains a stable voltage on the bitline. To avoid soft programming, the bitline is usually clamped to roughly 0.8 V. Inverter INV1 and transistor M5 form a unitary feedback loop that charges the bitline voltage to the clamped value. When the signal Disc = '0' and Prec = '0', the sense amplifier enters the sense stage, the current of the memory cell is directly mirrored by M1 into the drain side of M2, then current to voltage conversion takes place at node A. Inverter INV2 acts as a comparator that compares the voltage of node A with its switching threshold. If the cell current is larger than the reference current, the output node Out will be converted to logic '1', otherwise it will

* Project supported by the National High Technology Research and Development Program of China (No. 2008AA031403), the National Basic Research Program of China (No. 2010CB934204), and the National Science Fund for Distinguished Young Scholars of China (No. 60825403).
† Corresponding author. Email: liuming@ime.ac.cn
Received 18 March 2010, revised manuscript received 17 May 2010

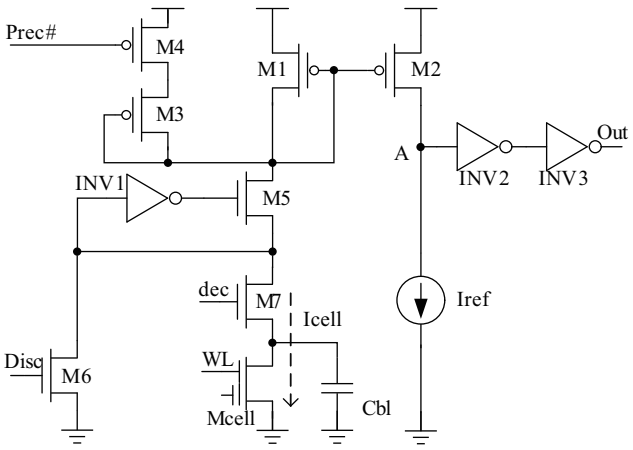

Fig. 1. Block scheme of the conventional single ended current mode sense amplifier[7].

be '0'.

There are two drawbacks of this circuit. One is that the current window between the memory cell current and the reference current becomes even smaller under lower power supply voltage, which results in a longer sense time. The other is that the switching threshold of inverter INV2 is highly dependent on the process corner, which may lead to a longer sense time.

3. Proposed high-performance sense amplifier

3.1. Description of the new sense amplifier

The circuit structure of the proposed sense amplifier is shown in Fig. 2(a). A current-mode sense amplifier with the offset-free technique is proposed to improve the sensing speed. The offset-free circuit structure is composed of the capacitance C_{AZ} used for AC coupling and the regenerative inverter INV2 controlled by the switch S2. For the sake of clarity, the timing diagram of SE, Prec, V_A, V_C, and Out during read operation is illustrated in Fig. 2(b). The detailed operation of the proposed sense amplifier is described in the following. In the precharge stage, transistor M6 is turned off and transistor M3 is turned on. Therefore, the precharge current flows to the selected bitline through transistor M5, and then inverter INV1 and transistor M5 clamp the bitline voltage to roughly 0.8 V. For the offset-free circuit, switch S2 is turned on; thereby the input voltage of inverter INV2 is equal to its output voltage. At the same time, switches S0 and S1 connect node a' and node b' to turn off transistor M2 and M8, and switch S3 is turned on to clamp node A to roughly half of the supply voltage. Under these conditions, the trip point is set at the point where INV2 has the maximum voltage gain. In addition, this trip point has no influence on the bias of node B because the capacitor C_{AZ} is used for DC isolation. Indeed, the voltage difference between the node A and the trip point is stored in the capacitor C_{AZ}. With this offset auto-zeroing technique, the trip point can be kept independent of the technology parameter, which is very critical for reliable sensing with a small memory cell current.

At the end of the precharge stage, Prec goes low, switch S3 is turned off, and switch S0 and S1 connect nodes a and b, respectively. Thereby, transistor M1 precharges the bitline merely. Subsequently, SE goes low, switch S2 is turned off, and then the sense amplifier switched to the sense stage. The difference between the memory cell current (I_{cell}) and the reference current I_{ref} determines whether the voltage of node A (V_A) increases or decreases. For the case of $I_{cell} > I_{ref}$, V_A rises, and vice versa. Due to the charge preservation of capacitor C_{AZ}, a small variation of node A is able to be coupled to node B, and then inverter INV2 senses the variation from the trip point and amplifies the variation with maximum voltage gain. Accordingly, when $I_{cell} > I_{ref}$, the state of the output node Out is changed to '1'. Otherwise, it remains at '0'. It is worth noting that in order to relieve the impact of channel charge injection of the MOS switch (S2 and S3), a transmission gate composed of a PMOS transistor in parallel with a NMOS transistor is employed. When sense operation is completed, the sense amplifier is reset for the next read operation. For the unselected bitline, transistor M6 will discharge the bitline voltage to improve the data retention.

The proposed sense amplifier is designed based on the GSMC 130 nm eFlash process. The column decoding structure is included in the circuit diagram and the bitline capacitance is set to 0.5 pF, as shown in Fig. 2(a). Moreover, the current distribution of erased and programmed cells due to the read-cycle endurance is taken into account. In the worst case, the memory cell current I_{cell} is set to 16 μA for erased cells and 0.5 μA for programmed cells. Hence, the reference current I_{ref} is trimmed to 8 μA.

3.2. First order modeling

Since transistor M2 mirrors the memory cell current and converts the current difference to the voltage difference at node A, a further analysis is performed for the transient behavior of node A, which can be expressed as

$$\mathrm{d}V_A = \frac{I_{ref}\mathrm{d}t}{C_{AZ}+C_p} - \frac{I_{cell}\mathrm{d}t}{C_{AZ}+C_p}, \tag{1}$$

where I_{cell} and I_{ref} represent the selected memory cell current and the reference current, respectively. C_{AZ} is the depletion region capacitance of a PMOS transistor, and C_P is the parasitic capacitance at node A. The first term in Eq. (1) represents the voltage decrease due to the discharge of C_{AZ} by I_{ref}, while the second term represents the voltage increase due to the charge of C_{AZ} by I_{cell}. Assuming that the parasitic capacitance at the inverter input node B is equal to C_P, the transient behavior of node B can be formulated as

$$\mathrm{d}V_B = \frac{C_{AZ}}{C_{AZ}+C_p}\mathrm{d}V_A = \frac{C_{AZ}}{(C_{AZ}+C_p)^2}(I_{ref}-I_{cell})\mathrm{d}t. \tag{2}$$

Then, the transient behavior of the inverter output node C is given as

$$\frac{\mathrm{d}V_C}{\mathrm{d}t} = \frac{g_m}{C_{load}}\frac{\mathrm{d}V_B}{\mathrm{d}t}, \tag{3}$$

where g_m is the transconductance of the inverter and C_{load} is the parasitic capacitance at the invert output node C. Hence, we can get the transient behavior of node C from Eqs. (1)–(3):

$$V_C(t) = V_{trip} + \frac{g_m}{2C_{load}(C_{AZ}+C_p)}\frac{C_{AZ}}{C_{AZ}+C_p}(I_{ref}-I_{cell})t^2, \tag{4}$$

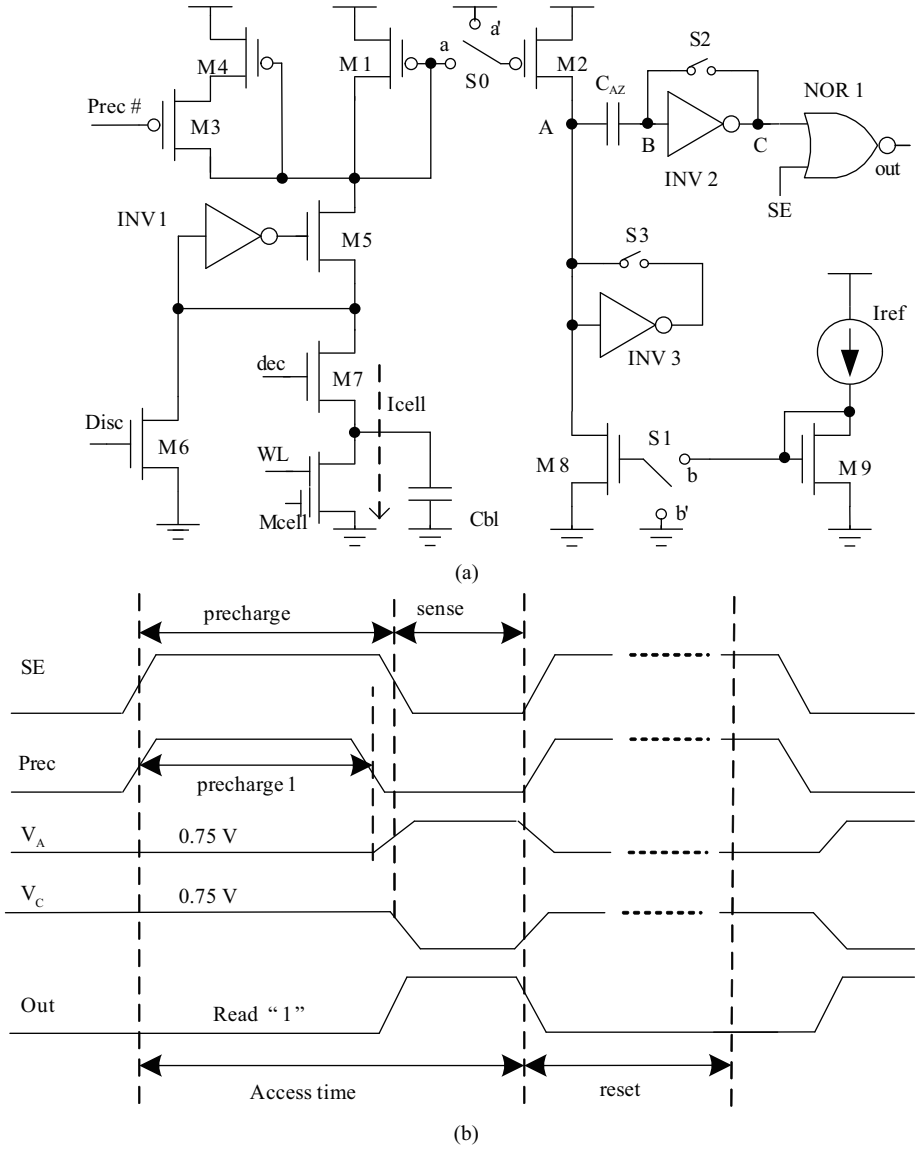

Fig. 2. (a) Circuit diagram of the proposed sense amplifier. (b) Timing diagram for the operation of the proposed sense amplifier.

where V_{trip} is the inverter trip point voltage.

From Eq. (4), it can be seen that the output voltage of the inverter is independent of the bitline capacitance. For comparison, the transient behavior of the output inverter in the conventional sense amplifier is given in Ref. [5]:

$$V(t) = V_{\text{trip}} + \frac{g_{\text{m}}}{2C_{\text{load}}(C_{\text{AZ}} + C_{\text{p}})}(I_{\text{ref}} - \beta I_{\text{cell}})t^2, \quad \beta = \frac{C_{\text{AZ}}}{C_{\text{BL}}}, \tag{5}$$

where C_{BL} is the bitline capacitance, and $\beta < 1$.

In flash memories, the erased cell draws a current $I_{\text{on}} > 0$ and the programmed cell draws a current $I_{\text{off}} \approx 0$. To achieve the maximum slope during transient time, the reference current I_{ref} is chosen as $I_{\text{ref}} = 0.5(I_{\text{on}} + I_{\text{off}})$ in Eq. (4) and $I_{\text{ref}} = 0.5\beta(I_{\text{on}} + I_{\text{off}})$ in Eq. (5).

4. Simulation results and analysis

With the supply voltage of 1.5 V, simulated waveforms of the erased cell with the proposed sense amplifiers under typical conditions are shown in Fig. 3. It can be seen that in the proposed sense amplifier, at the beginning of the sense stage, V_{B} samples the voltage variation and rises from the trip point. With the offset-free technique, inverter INV2 amplifies the variation with maximum voltage gain and the output can quickly be set to a certain value.

As mentioned above, the memory cell current becomes smaller as the read voltage decreases. Figure 4 gives the simulated sense time with respect to the memory cell current. As can be observed, the sense time of the proposed scheme is much less than the conventional scheme in Ref. [5] at a wide memory cell current range of 6 to 22 μA. It can be explained that in the proposed design, the slope in Eq. (4) is larger than that in Eq. (5) by a factor of β. It is worth noting that β is much less than 1 since C_{AZ} is much smaller than C_{BL}. Therefore, the proposed design achieves much better performance. In addition, in the conventional scheme[5], it is difficult to bias the reference current I_{ref} for a very small memory current.

Figure 5 illustrates the sense time with respect to the bitline capacitance. The simulation result shows that the sense time of

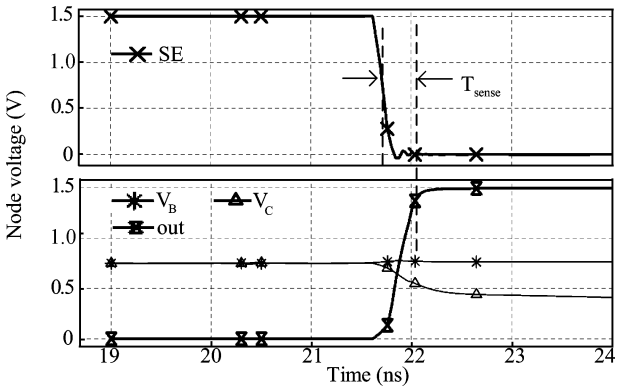

Fig. 3. Read operation simulated waveforms of the erased cell under typical conditions.

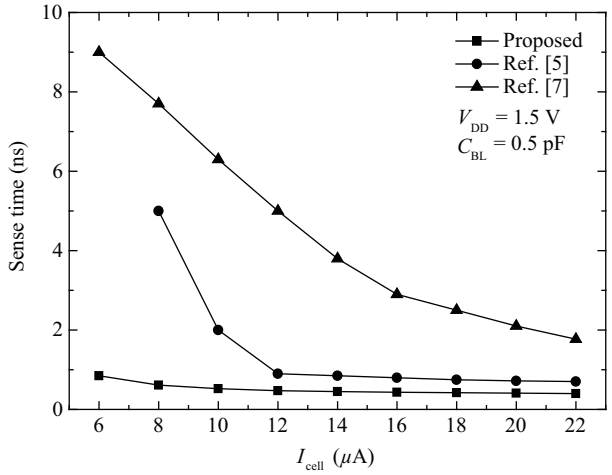

Fig. 4. Sense time sensitivity to the memory cell current.

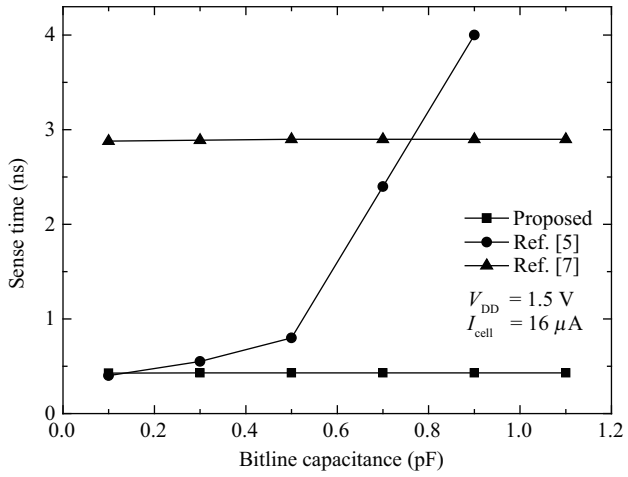

Fig. 5. Sense time sensitivity to the bitline capacitance.

the proposed scheme is independent of the bitline capacitance, which verifies the expression in Eq. (4). At the given time Δt, the voltage difference at node C between data '0' and data '1' can be deduced by substituting I_{on} and I_{off} into Eq. (4) and

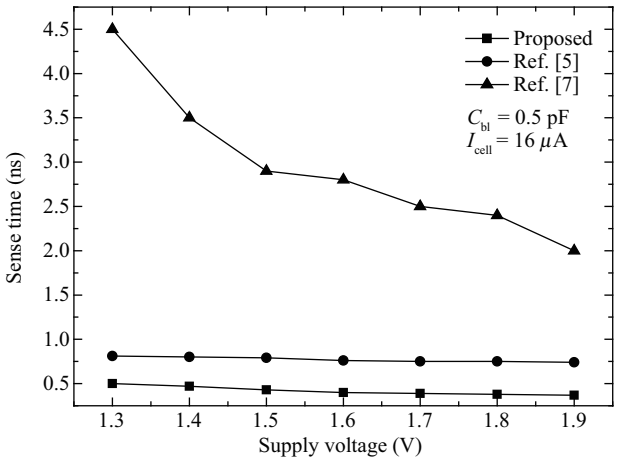

Fig. 6. Sense time sensitivity to the supply voltage.

Table 1. Sense time and power comparisons of the sense amplifier.

Parameter	Ref. [7]	Ref. [5]	Proposed
Sense timing (ns)	2.9	0.8	0.43
Average power (μW)	83.1	90	85

calculating the difference. The result is described as follows:

$$\Delta V_C = \frac{g_m \Delta t^2}{2C_{load}(C_{AZ}+C_p)} \frac{C_{AZ}}{C_{AZ}+C_p}(I_{on}-I_{off}). \quad (6)$$

Using a similar method, the voltage difference between data '0' and data '1' in Ref. [5] is given by

$$\Delta V = \frac{g_m \Delta t^2}{2C_{load}(C_{AZ}+C_p)} \frac{C_{AZ}}{C_{BL}}(I_{on}-I_{off}). \quad (7)$$

From Eq. (7), it can be seen that the voltage difference ΔV is inverse to the bitline capacitance C_{BL}. In high density memories, ΔV becomes relatively small due to large C_{BL}, thereby a longer sense time is required to distinguish logic '1' from logic '0'. However, for the proposed topology, the voltage difference ΔV_C is independent of the bitline capacitance, as shown in Eq. (6). Hence, the sense speed is able to be improved significantly.

As the power supply voltage scales down, the performance of the sense amplifier degrades. Figure 6 illustrates the sense time with respect to the supply voltage. It is clear that the proposed sense amplifier has a smaller sense time at a wide range of supply voltage. Therefore, the proposed sense amplifier is very suitable for low voltage applications.

5. Conclusions

In this work, a high-speed single-ended sense amplifier based on the current mode approach is designed for low-voltage embedded flash memories. An offset-voltage elimination technique is employed to improve sense performance. In GSMC 130 nm eFlash technology, the simulation results show that the worst sense time of the proposed sense amplifier is 0.43 ns with almost the same power consumption, as shown in Table 1. It achieves 46% improvement compared to conventional schemes[5]. Furthermore, the proposed sense amplifier

has better performance in lower memory cell current and larger storage capacity at a supply voltage of 1.5 V.

References

[1] Micheloni R, Crippa L, Sangalli M, et al. The flash memory read path: building blocks and critical aspects. Proc IEEE, 2003, 91(4): 537

[2] Otsuka N, Horowitz M A. Circuit techniques for 1.5-V power supply flash memory. IEEE J Solid-State Circuits, 1997, 32(8): 1217

[3] Conte A, Giudice G L, Palumbo G, et al. A high-performance very low-voltage current sense amplifier for nonvolatile memories. IEEE J Solid-State Circuits, 2005, 40(2): 507

[4] Tanzawa T, Takano Y, Taura T, et al. A 1.2 V sense amplifier for high-performance embeddable NOR flash memories. Proc ISCAS, 2005, 2: 1266

[5] Kajiyama S, Fujito M, Kasai H, et al. A 300 MHz embedded flash memory with pipeline architecture and offset-free sense amplifiers for dual-core automotive microcontrollers. Proc ASSCC, 2008: 257

[6] Papaix C, Daga J M. A new single ended sense amplifier for low voltage embedded EEPROM non volatile memories. Proc Int Workshop on MTDT, 2002: 149

[7] Seo M K, Sim S H, Oh M H, et al. A 130-nm 0.9-V 66-MHz 8-Mb (256K×32) local SONOS embedded flash EEPROM. IEEE J Solid-State Circuits, 2005, 40(4): 877

[8] Deml C, Jankowski M, Thalmaier C. A 0.13 μm 2.125 Mb 23.5 ns embedded flash with 2 Gb/s read throughput for automotive microcontrollers. Proc ISSCC Dig Tech Papers, 2007: 478