

Novel Level-Identifying Circuit for Flash Multi-Level Memories

D. Montanari, J. Van Houdt, G. Groeseneken and H. E. Maes

IMEC, Kapeldreef 75, B-3001 Leuven, Belgium
Tel: + 32 16 281 561; Fax: + 32 16 281 501
E-mail: montanar@imec.be

This paper presents a high-speed, small-area circuit specifically designed to identify the levels in the read out operation of a Flash Multi-Level Memory. The analog computation of the Euclidean Distance between the current read current of a cell and the reference currents that represent the different logic levels. A new version of the circuit has been integrated in a standard double-metal process with a die area of only 140 x 100 μm^2 . Operating under a 5V power supply, it identifies the read-out current of a memory cell and associates it with a logic level in 9 nsec.

1. Introduction

In order for Flash technology to access the profitable market for portable applications, the cost per bit must be reduced. Device manufacturers are expected to take the bit cost down to \$1/Mbyte by the year 2000. However, extensive additional circuitry is required to write and read multiple levels. Charge Storage (MLCS) could further reduce this cost by at least a factor of 2 per technology generation.

Reading multiple charge levels does not only mean that the cell current is compared with reference currents (in the binary case), but also that it is compared with reference currents for each of the different logical levels it represents. This latter operation requires a dedicated circuit which consumes silicon area, enlarges the peripheral area, and consequently, decreases the area reduction produced by using MLC. This circuit introduces a delay which increases the access time of the memory. It is of uttermost importance that the design of this 'level identifying' circuit minimizes silicon area and propagation time.

The architectures presented so far in the literature, show a trade-off between area on one side and propagation time on the other. In fact, parallel architectures obtain high speeds through a single comparison step, independent from the cell content but they need $n-1$ comparators ($n =$ number of levels) to be implemented on silicon; sequential architectures [2] are much

time depends on the cell content, but only one comparator is needed. This architecture's utilise a single comparator, but extra logic circuitry state machine [3] and the propagation delay is larger than in the parallel architecture.

This paper presents a new fully parallel 'level identifying' circuit which utilises comparators and is designed specifically to minimise die area.

2. Working Principle

The circuit working principle and block diagram are shown in Fig. 1 and Fig. 2 respectively. In order to realise the multi-level signal comparator between the read-out current and each of the reference currents (L1, L2, L3, L4), four 'Distance Measuring' sub circuits of Fig. 2. Each of these sub circuits produces a MS voltage that is proportional to the similarity of the measured current (MS) to the reference current (L). The better the match, the higher the MS voltage becomes. In the second stage of the circuit of Fig. 2, the Winner-Take-All circuit decides which of the four MS's is the highest and brings the corresponding digital state, while the other three outputs remain low. Only one output is high for any given cell current value.

3. Circuit Implementation

This block receives two input signals (I_{dc} and I_{dx}) and delivers a MS voltage (V_{MS}) that is proportional to the similarity between the two input currents. The Euclidean distance between the currents, the higher is V_{MS} . The Euclidean Distance can be expressed as:

$$Dist = \sqrt{(R - L)^2} \quad (1)$$

where R and L are two generic monodimensional numbers. V_{MS} is a monotonically decreasing function of Dist. To implement expression (1) on silicon, a long channel MOSFET in saturation is exploited:

$$I_{DS} \cong K(V_G - V_S - V_T)^2 \quad (2)$$

Applying a voltage signal (R) to the gate of a MOSFET is straightforward. To apply a voltage (L) to its source, a second MOSFET in the source must be connected to the source of the driving n-MOS device. In this configuration, $(L)^2$ is implemented. The source is then fed to a p-MOS device in the same configuration which implements the square root function.

The considerations mentioned above lead to the DM block circuit which includes two driving n-MOS device (M2, M3) with the respective gates (M4, M1), a diode connected p-MOS device (M5) and two current-to-voltage MOS device (D1, D2). The dimensions of the transistors are reported in Table 1. They have an aspect ratio which is ten times that of M2 and M3 since the transistors utilised to bias the source of the driving transistors. The biasing is done with respect to the driving transistor, the more accurately the biasing is done. Simulations show that a factor of 10 is a good compromise between

consumption. D1 and D2 convert I_{cell} into V_D and V_C respectively. If I_{cell} is than I_x , then M2 is cut off since the voltage at its source, V_C , is h at its gate. However, M3 is on and the larger the difference between V_C and V_D , the larger is drain current and the smaller is V_{MS} . On the contrary, if I_{cell} is sn cut off and M2 is on.

Fig. 4 shows the proposed WTA, which consists of 4 current i (MS0...MS3) and 4 current comparators. The WTA receives the four V_{MS} blocks as input signals at the gates of the current conveyors and de signals. The current conveyors have the source node (X) in common a same biasing current. When the V_{MS} 's are applied to the gates of tl follows the largest input voltage, turning off the other three conv will flow only through the conveyor with the largest V_{MS} . Only the co V_{MS} will drive current and result in a high output voltage.

4. Experimental Results

An experimental version of the circuit has been integrated in a double-metal 0.7- μ m process with a silicon area of $2f$, as shown in the picture of Fig. 5, and has been extensively evaluated.

Fig. 6 shows the experimentally evaluated function of a sweeping current in the case of a reference current of 25 μ A. As expected, the the higher the V_{MS} . Fig. 7 shows the experimental output voltages of circuit as a function of a sweeping cell current in the case of refer and 160 μ A. Only one output voltage is high for any given cell curren

5. Conclusions

The sensing and the correct identification of multi-level signa issue in the design of MLCS memories. An adequate fully parallel arch employs a system of four analog computational blocks and a winner-tal propagation delay and the die area.

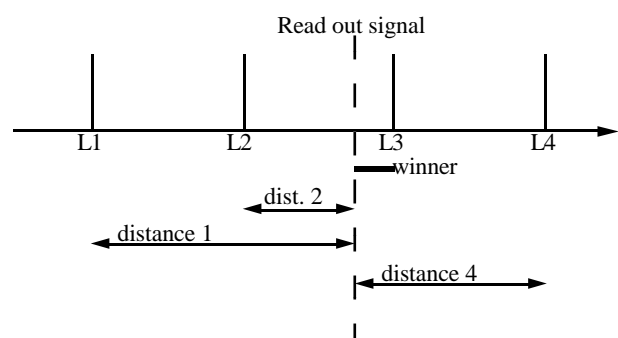
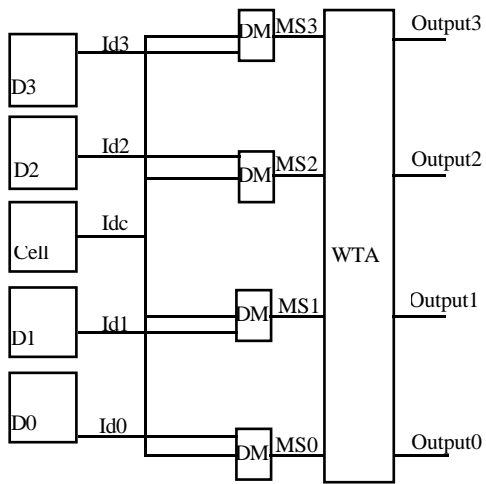


Fig.1: Level Identifying circuit: working principle.

	W (μ m)	L (μ m)
M1	70	0.7
M2	7	0.7
M3	7	0.7
M4	70	0.7
M5	1.5	0.7
MS0	70	0.7
MS1	70	0.7
MS2	70	0.7
MS3	70	0.7

Table 1: Transistor dimensions.



D0, D1, D2, D3: Dummy cells.
 Cell: Memory cell.
 Idx: Read-out current.

DM: Distance Measuring Block.
 WTA: Winner-take-all.

Fig. 2: Block diagram of the level identifying circuit.

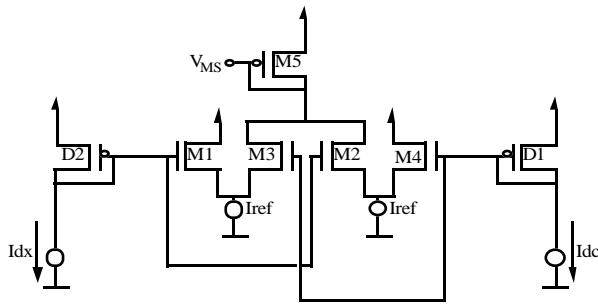


Fig. 3: Distance Measuring Block.

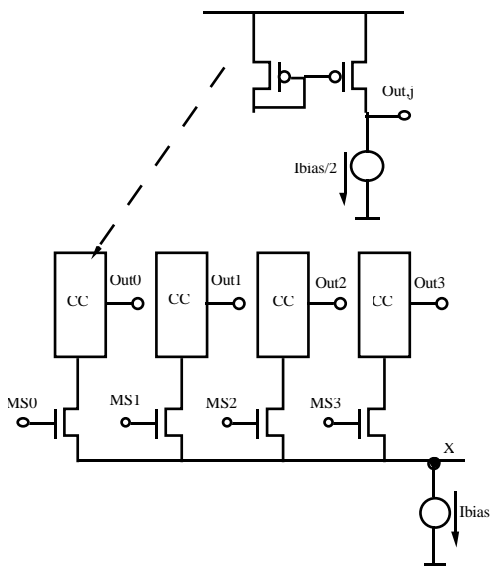


Fig. 4: Winner-Take-All

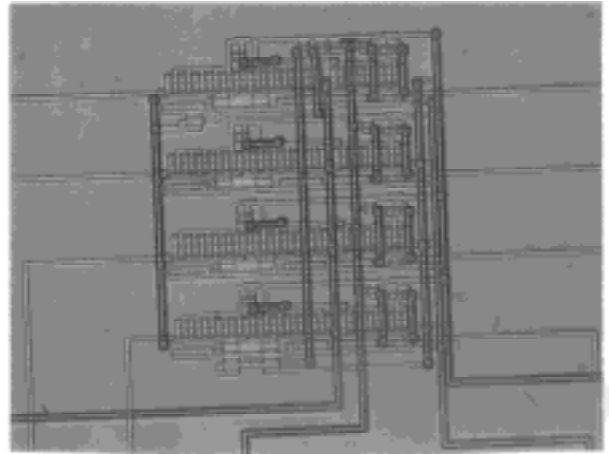


Fig. 6: Micrograph of the level identifying circuit.

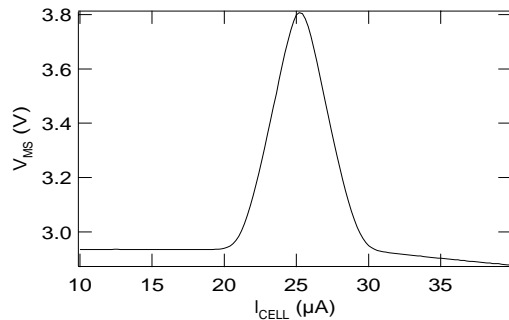


Fig. 6: 'Distance Measuring' transfer characteristic.

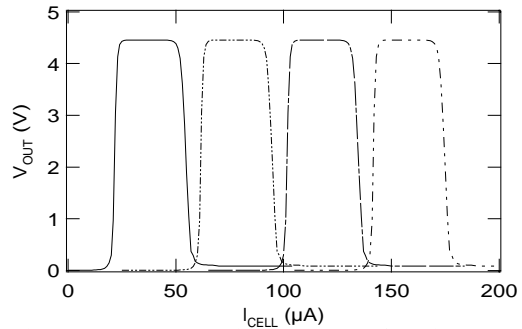


Fig. 7: 'Level Identifying' transfer characteristic.

References

- [1] S. Aritome et al., "A novel side-wall transfer-transistor cell (SWATT cell) for multi-level NAND EEPROMs", IEDM Tech. Dig., p. 275, 1995.
- [2] T. Jung et al., "A 117-mm 3.3-V Only 128-Mb Multilevel NAND Flash Memory for Mass Storage Applications", IEEE Journal of Solid State Circuits, VOL. 31, No 11, p. 1575-1583.
- [3] C. Calligaro et al., "Mixed Sensing Architecture for 64Mbit 16-Level-Cell Non-Volatile Memories", Proc. IEEE Int. Conf. Innovative Systems on Silicon, Austin, TX, Oct. 1996 p. 133-140.
- [4] Patent pending.