A Configurable CMOS Memory Platform for 3D-Integrated Memristors

Melika Payvand, Advait Madhavan, Miguel Angel Lastras-Montaño, Amirali Ghofrani, Justin Rofeh, Kwang-Ting Cheng, Dmitri Strukov, Luke Theogarajan
Electrical and Computer Engineering
University of California, Santa Barbara
Santa Barbara, CA, U.S.A
{ltheogar,melika01}@ece.ucsb.edu

***Abstract*— Memristors are emerging as powerful nanoscale devices for diverse applications, such as high-density memories and neuromorphic applications. However, this nascent technology requires considerable advancement before this vision is realized. We present a highly configurable CMOS interface chip which enables the characterization of on-chip memristors, especially for memory applications. The chip was fabricated in On-Semi 3M2P 0.5 µm occupying 2×2 mm^2. The chip design allows for post-CMOS fabrication of memristors. The interface between the memristor and the CMOS circuitry was provided via a top metal contact. The chip was designed to support an area-distributed interface decoupling CMOS pitch and memristor pitch, enabling high-density memristor integration. Measurement results on post-CMOS fabricated Ag/SiO_2/Pt memristive devices are reported. Though we have shown the results from one memristive material stack, thorough chip characterization demonstrates the versatility of the chip enabling its use with a wide variety of materials stacks.**

Keywords*— *Memrsitor Read and Write Circuitry; Tunable Circuitry; 3D Integration; Hybrid circuits; CMOL; High Density Memory Array; Memristor Characterization; Crossbar Array

I. INTRODUCTION

Driven by the rapid growth of handheld devices, flash memory has become one of the fastest growing memory segments with state-of-the-art feature size being pushed into the 10nm regime. Lithography and leakage-limited physical scaling and material dependent read/write voltage of electrical scaling are some of the issues plaguing today's non-volatile memories (NVMs) [1]. In recent years, several emerging memory technologies such as ReRAM [2,17], phase change memories [3], and spin-transfer torque magneto-resistive memories [4] have been under study to overcome these issues and Memristor technology has shown a lot of potential as a candidate for the next generation of non-volatile memory. Rather than replacing CMOS with these nascent technologies, a symbiotic approach utilizing the favorable properties of memristive devices, such as high density, fast switching and high endurance [5], could extend nanoscale memories beyond their current limitations. More importantly, a key feature of memristive devices is their ability to be easily fabricated on top of CMOS foundry devices. Among the many proposed architectures for integration of CMOS and novel materials [6-11], CMOL [12] is a promising candidate. CMOL tackles fabrication issues such as interlayer alignment accuracy and integration of nanoscale devices over a CMOS sub-system with larger scale feature size. Moreover, it provides high-density memory with less parasitics by sharing select circuitry between multiple memristors (1T-1R vs 1T-NR).

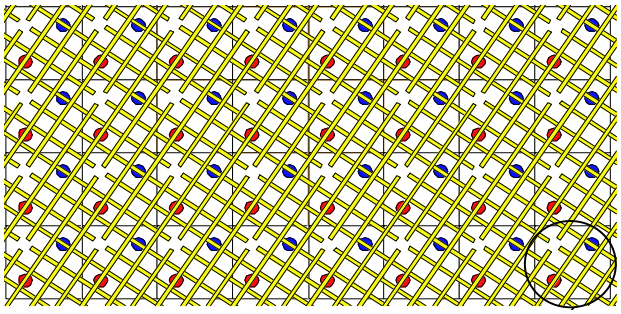

Fig. 1. CMOL concept. The circle highlights the CMOS cell. Figure is taken from [12]

With this vision of a monolithic, 3D-integrated CMOL memory platform in mind, we have designed and tested the first prototype of the CMOL architecture complete with integrated memristors. This paper focuses on the challenges involved from a circuit design perspective and the steps taken to support memristors with different ranges of resistance, threshold voltages, on/off ratio etc. More in-depth analysis of the architectural trade-offs can be found in [13] and details of the memristor integration is discussed in [14]. Section II provides some of the relevant background for the CMOL architecture. In section III the design of the chip, along with a brief discussion on the architecture, is discussed and in section IV we report the experimental results after memristor integration along with thorough characterization of the chip.

II. CMOL ARCHITECTURE

The term "CMOL" comes from the combination of CMOS and Molecular scale devices and was conceived to mitigate the density scaling issues of CMOS. The main idea of CMOL is to construct 3D-integrated crossbars, i.e. mutually perpendicular layers of parallel nanowires (electrodes) forming two terminal, memristive devices at the cross-points. However, high- density memory requires long crossbars, causing excessive capacitances and resistances on the lines, limiting the speed and functionality of the system. By segmenting these long nanowires into shorter fragments, some of these issues are alleviated. However, an area-distributed interface becomes necessary to address individual devices.

Exploiting the intrinsic nanoscale dimensions of this memory requires decoupling the underlying CMOS feature size from the device. One method of decoupling is to rotate the nanowires [12]. Shown in Fig. 1 is the CMOL concept. The

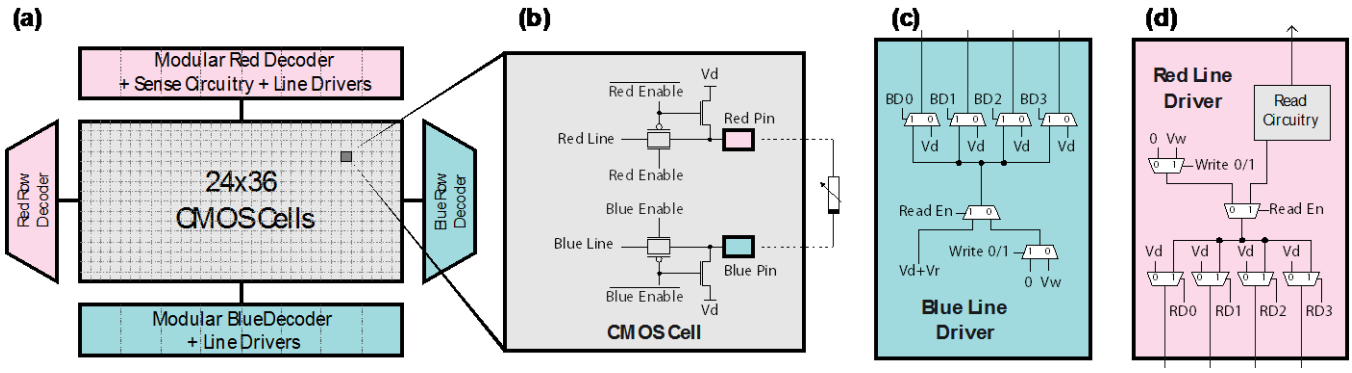

Fig. 2. a) Overall chip architecture. b) CMOS cell. When the transmission gates are selected by Red/Blue enable signals, they connect the Red/Blue lines to the Red/Blue pins which are the interface to the integrated memristors. c,d) Blue and Red line drivers which place the appropriate voltages on the Red/Blue lines.

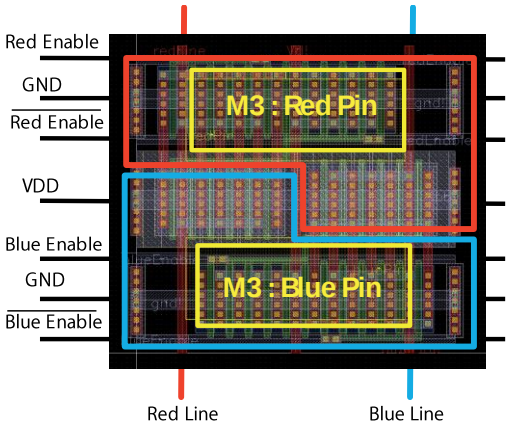

Fig. 3. CMOS cell layout. Metal 3 is used as the interface with integrated memristors. This cell occupies an area of 32×32 μm^2 in a 0.5µm process.

Red and Blue pins represent the area-distributed interface connecting the underlying CMOS to the integrated top and bottom crossbar nanowires, respectively. Each Red and Blue pair forms a so-called "CMOS cell," highlighted in Fig. 1. The cell houses transmission gates and drive circuitry supplying appropriate voltages to the electrodes required for correct operation. Unlike the standard 1T-1R architectures, CMOL is a 1T-NR architecture since every transmission gate is connected to N memristors. The next section addresses the design of the aforementioned CMOS subsystem.

III. Memory Access Controller for Memristor Applications (MAMA)

The plethora of memristive device designs, each with their unique advantages, requires a flexible supporting circuit architecture. The circuit design is strongly influenced by the connectivity imposed by the area-distributed interface and the chip architecture, briefly discussed in Section A. We term this versatile chip the Memory Access controller for Memristor Applications (MAMA). A key circuit requirement for the MAMA chip is the ability to handle memristors with different R_{on}/R_{off} values, and provide the appropriate write and read voltages. These circuits are discussed in sections B and C.

A. Chip Level Architecture

Fig. 2a depicts the overall architecture of the chip. The chip consists of an array of CMOS cells, double decoders, programming drivers and sensing circuitry. Each CMOS cell houses select circuitry including Red and Blue pins required by the area-distributed interface (Fig. 2b). Selecting two of these Red and Blue pins accesses two of the segmented nanowires and hence a unique memristive device at the cross-point. A row-column decoder in turn accesses these pins. Thus it requires a double decoding scheme. The double decoders surround the CMOS cell array and have their function split among the Blue/Red row/column decoders. Depending on the desired operation (Read/Write) the Blue/Red line drivers place appropriate voltages on the Red and Blue lines (Fig. 2c-2d) which connects to the Red/Blue pins through the CMOS Cell select circuitry (Fig. 2b) [15]. Specifically, during the read operation, the sensing circuitry makes a binary decision regarding the memristor state and the data is shifted out serially.

B. Writing Circuitry (CMOS Cell Design)

To write on a particular memristor, the device is addressed and the appropriate write voltages are applied across it. This is done through CMOS cells shown in Fig. 2b. It includes two transmission gates controlled by Blue/Red enable signals routed from the double decoder. When the gates are asserted, they drive the Red and Blue pins with the appropriate voltages on the Blue/Red lines. De-assertion connects the pins to a default voltage, V_d, in order to avoid floating problems such as leakage or unpredictable state-changes due to unwanted noise sources.

The transmission gates together with the memristors comprise a voltage divider. To ensure the memristor's operation in the desired region (i.e. the write region), the voltage drop across the transmission gates must be negligible. Therefore, these pass gates need to be sized accordingly.

However, the size of the transmission gates imposes a limitation on the number of CMOS Cells which can fit in the chip and hence the size of the memory supported by the chip. As a result, there is a trade-off between the maximum current drive and the size of the integrated memory on the chip.

Given these constraints, a size of W/L=42µm/0.6µm in 0.5 µm process is chosen for the pass gate transistors. The maximum current supported by these transmission gates for a range of input voltages is reported in section IV. This maximum current can be considered as the compliance current limiting the current passing through the memristors and hence preventing device break down [5]. Depending on the required write voltage, the minimum resistance supported by the chip can be calculated.

Layout realization of a CMOS cell is shown in Fig. 3. Since this CMOS chip needs to be post-processed for 3D memristor integration, the last metal layer in the On-Semi 0.5µm process (Metal 3) is crucial to the area-distributed interface. General power and ground routing cannot be done on this metal layer as it risks exposing and damaging these lines, therefore they

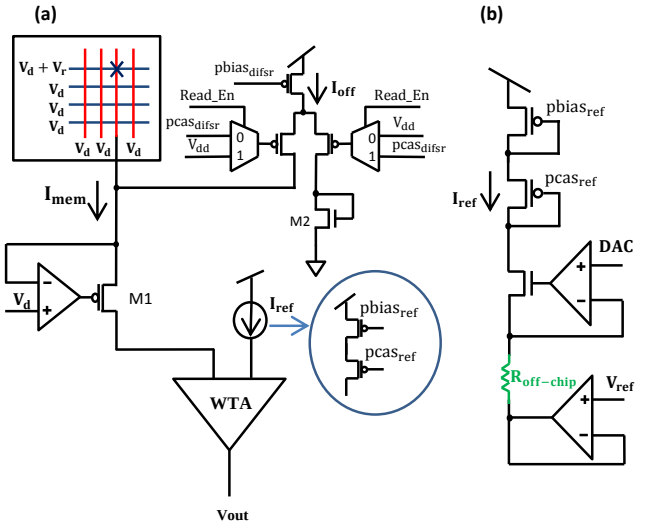

Fig 4. Sensing circuitry. a) The current-sensing scheme. The memristor's current from the crossbar is compared against a reference current by the winner-take-all (WTA) circuit. b) A tunable reference current. The current can be changed by tuning the $R_{off-chip}$ or the DAC voltage.

are routed in Metal 2. The size of the pins comprising this area-distributed interface has been intentionally made large (24×8 μm^2) to reduce the effect of cumulative alignment error.

C. *Sensing Circuitry Design*

In order to make a binary decision regarding the memristor state, a current-sensing scheme is chosen over a voltage-sensing counterpart. In a conventional voltage-sensing scheme, a transimpedence amplifier (TIA) is utilized to convert the signal into a voltage which is then compared against a threshold voltage. However, the TIA needs at least a two stage op-amp with an appropriate output stage in order to drive the resistive load. Moreover, for a high resistive gain of the TIA, a large feedback resistor is needed, which takes up a large silicon area. Therefore, for a more compact design, the current-sensing scheme is utilized. Also, a current-sensing scheme has the advantage of a much larger dynamic range, which is required for the configurability.

Fig. 4a shows the schematic of the current-sensing circuitry. The current drawn by the device in response to a small read voltage, V_r, is compared against a reference current. The read voltage should be picked in a region where the memristor's state does not change. This read voltage is applied by pinning one terminal of the memristor of interest to the default voltage, V_d, by an op-amp, while the other terminal is driven by the blue line driver to V_d+V_r. This read current is then compared against a reference current using a winner-take-all (WTA) [16] circuit.

As the sensing circuitry is only connected to the memristors when the *Read En* signal is asserted, the pinning loop is not always closed. In order to avoid the settling time of the loop when the *Read En* signal asserts, a very small I_{off} current (50 pA, through $pbias_{difsr}$ and $pcas_{difsr}$ generated from a current diffuser) is passing through M1 while *Read En* is not active. As soon as *Read En* is activated, the I_{off} current is steered to an alternate path and is drained by M2.

In order to make the sensing circuitry compatible with different memristor types (e.g. different R_{on} and R_{off} values), a tunable reference current is designed. As is shown in Fig. 4b, this current reference can be tuned by two knobs: the off-chip resistor and the DAC output voltage across that resistor.

A flexible platform for generating the read and write voltages is designed on the PCB test board by using DAC-controlled voltage sources.

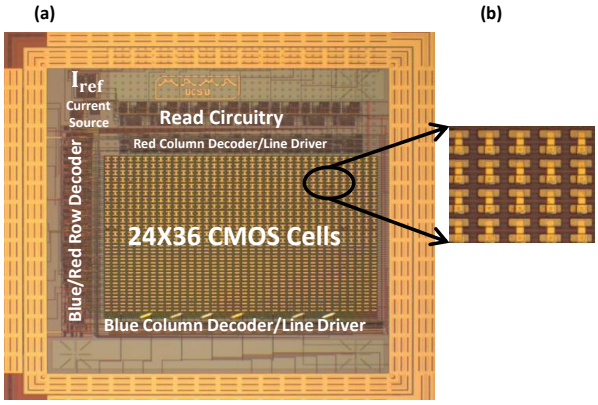

Fig. 5. a) Chip micrograph. Different parts of the chip are shown. b) Individual devices integrated on the chip.

IV. MEASUREMENT RESULTS

The chip micrograph is shown in Fig. 5a. It occupies an area of 2×2 mm^2 and was fabricated in On-Semi 3M2P 0.5µm technology through the MOSIS service. This area can potentially support 1kb of memory. Using an advanced CMOS technology node will allow for a larger memory size and smaller CMOS cell size. The range of voltages required for different memristor types coupled with the fabrication cost make 0.5µm technology ideal for this multipurpose chip.

In order to test the functionality of the chip independent of successful memristor integration, the last row of the CMOS cell arrays are connected to peripheral bond pads. We used a potentiometer connected to two of these pads to verify the functionality of the chip over a large range of resistances, since a memristor is essentially a resistor in steady state. The experimental characterization of the chip along with the memristor integration results are reported in this section.

A. *Writing Circuitry Characterization*

The most important factor of the MAMA chip writing circuitry is the maximum current drive for the various memristive loads. As is explained in section III B, decreasing the load resistance lowers the voltage drop across the memristor, which limits the current drive. Fig. 6a reports the voltage across a large range of load resistances for different writing voltages. The highlighted region (below 2kΩ) shows the memristor range not supported by this chip since the voltage drop across the transmission gates becomes dominant. Table 1 depicts the maximum current provided by the chip in the writing mode for a 10% drop of the writing voltages.

Table 1. Maximum current in the writing mode for a 10% drop of the writing voltage across the gates.

Writing Voltage (V)	1	2	3	4	5
Current (µA)	90	178	183	213	260

B. *Sensing Circuitry Characterization*

To characterize the sensing circuitry, a read voltage of 0.8V is applied across the varying load resistances. The generated current is then compared to the tunable reference current varying from 40nA to 4µA and the measurement results are

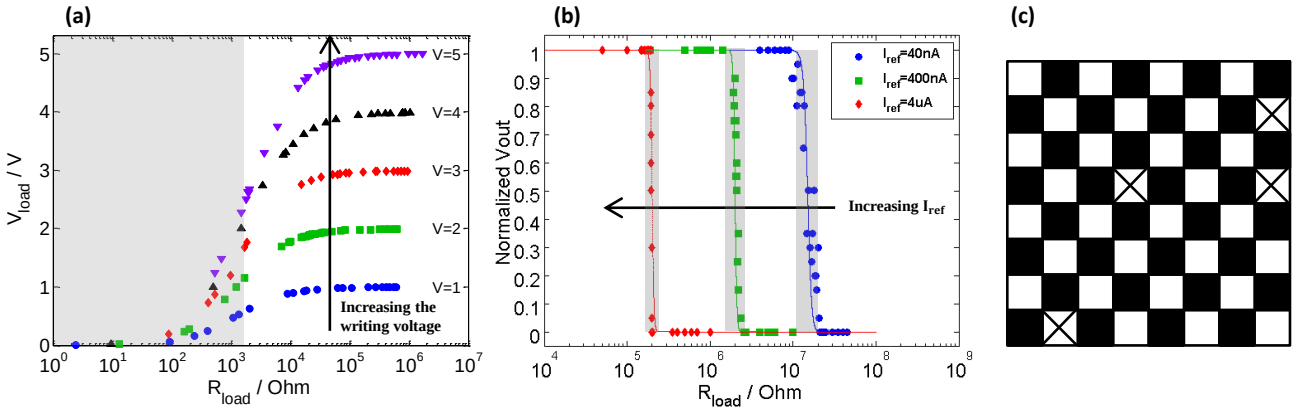

Fig. 6. a) Write circuitry characterization. As the resistive load decreases, the writing voltage drop across the load also decreases. b) Read circuitry characterization. The reference current to the WTA is tuned by two orders of magnitude and the response of the read circuitry is plotted. The highlighted region shows the forbidden zone. c) A checkerboard pattern is used to program an array of 8×8 devices. The devices with the X's are either shorted or failed to get programmed.

shown in Fig. 6b. The highlighted region illustrates the forbidden zone in which the winner-takes-all comparator is in the metastable state. As a result, fluctuations on the chip can affect the output state, giving it a probabilistic nature which can be seen in the forbidden zone. The forbidden zone can be defined as a region between a maximum low resistance (RIL) and a minimum high resistance (RIH). A curve is fit to the data and the relationship between RIL, RIH and I_{ref} in this chip is as follows:

$$R_{IH} = 1.82\, I_{ref}^{-0.9375} + \frac{1}{50\, I_{ref}} \qquad R_{IL} = 1.82\, I_{ref}^{-0.9375} - \frac{1}{50\, I_{ref}}$$

C. 3D Memristor-Integration Results

Ag/SiO_2/Pt memristors are 3D-integrated on the chip [14] and are shown in Fig. 5b. These integrated devices, once addressed, are programmed with 500ms, 5V pulses. To verify the programming, a 10ms pulse of 0.8V is applied to read the device's state. Fig. 6c shows the checker board pattern created from an array of 8×8 devices after a program and a read pulse. Black squares represent the devices in the low resistive state. Each row of the pattern (8 bit) is programmed in parallel and read out serially. The squares with an "X" mark show the memristors which are either shorted or do not get programmed because of the fabrication yield.

V. Conclusion and Future Work

We have demonstrated the first CMOS chip to support the CMOL architecture for various memristor types. Memristors are 3D-integrated as single devices on top of the CMOS chip. The analog part of the chip is characterized and the performance of the chip for different resistive loads is reported.

Although we have mainly focused on the functionality of the integrated CMOS chip with single devices, work is on-going with our collaborators to integrate segmented rotated crossbars on the MAMA chip.

Acknowledgment

This work is funded by the Air Force Office of Scientific Research (AFOSR) MURI grant FA9550-12-1-0038.

References

[1] G. Atwood, "Future directions and challenges for ETox Flash Memory Scaling", IEEE Trans. Dev. Mater., vol. 4, no.3, pp. 301-305, Sep. 2004.

[2] R. Waser and M.Aono, "Nanoionics-based resistive switching memories," Nature Materials, vol. 6, pp. 833-840, 2007.

[3] F. Bedeschi, R. Feckenthal, C. Resta, E.M. Donze, M. Jagasivamani, et al, " A multi-level-cell bipolar-selected phase change memory", *ISSCC*, p. 428, Feb. 2008.

[4] E. Chen, D. Apalkov, Z. Diao, A. Driskill-Smith, D. Druist, D. Lottis, et al, " Advances and Future Prospects of Spin-Transfer Torque Random Access Memory", IEEE. Tans. Magnetics, Vol. 46, No. 6, Jun 2010.

[5] J. J. Yang, D.B. Strukov and D. R. Stewart, " Memristive devices for computing", Nature Nanotech., vol. 8, pp. 13-24, Jan 2013.

[6] C. Kugeler, M. Meier, R. Rosezin, S. Gilles, R. Waser, " High density 3D memory architecture based on the resistive switching effect", Solid-State Electronics, vol. 53, pp. 1287-1292, Dec. 2009.

[7] W. Lian, S. Long, H. Lv, Q. Liu, Y. Li, Y. Wang, S. Zhang et al, " Multilevel storage characteristics in ZrO2-ReRAM brought about by ideal current limiter", ICSICT, pp. 1112-1114, Nov. 2010.

[8] M. Miyamura, S. Nakaya, M. Tada, T. Sakamoto, K. Okamoto, N. Banno et al, " Programmable cell array using rewritable solid-electrolyte switch integrated in 90nm CMOS", ISSCC, pp. 228-229, Feb. 2011

[9] D. Walczyk, Ch. Walczy, T. Schroeder, T. Bertaud, et al, " Resistive switching characteristics of CMOS embedded HfO2-based 1T1R cells", Microelectronic Engineering, vol. 88, pp. 1133-1135, Jul. 2011.

[10] J. Kim, A.J. Hong, S.M. Kim, K. Shin, E.B. Song, Y. Hwang et al, " A stacked memory device on logic 3D technology for ultra-high-density data storage", Nanotechnology, vol. 22, no. 25, May 2011.

[11] H. Kim, S. Gaba, D. Wheeler, J.M. Cruz-Albrecht, T. Hussain, N. Srinivasa and W. Lu, "A functional hybrid memristor crossbar-array/CMOS system for data storage and neuromorphic applications", Nano Letters, vol. 12, pp. 389-395, Dec. 2011.

[12] D.B. Strukov, K.K. Likharev, "CMOL FPGA: a reconfigurable architecture for hybrid digital circuits with two-terminal nanodevices", Nanotechnology, vol. 16, no. 6, Jun. 2005.

[13] M. Lastras, A. Ghofrani, K-T.T. Cheng "Architecting energy efficient crossbar-based resistive random access memories ", NANOARCH, Boston, July 2015.

[14] J. Rofeh, A. Sodhi, M. Payvand, M. Lastras-Montaño, A. Ghofrani, et al, "Vertical integration of memristors onto foundry CMOS dies using wafer-scale integration", ECTC, 2015 IEEE 65th , in press.

[15] M.S. Qureshi, M. Pickett, F. Miao and J.P. Strachan, "CMOS interface circuits for reading and writing memristor crossbar array", ISCAS, pp. 2954-2957, May 2011.

[16] C. Mead, Analog VLSI and neural systems, Addison-Wesley Pub. ,1989.

[17] A. Ghofrani, M. Lastras, K-T. Cheng, "Toward Large-Scale Access-Transistor-Free Memristive Crossbars," ASPDAC, Japan, Jan 2015.