Hybrid CMOS/Memristor Circuits

D. B. Strukov
Department of Electrical and Computer Engineering
UC Santa Barbara
Santa Barbara, CA 93106 USA
strukov@ece.ucsb.edu

D. R. Stewart
National Research Council of Canada
100 Sussex Dr, Ottawa, ON, K1A 0R6, Canada
duncan.stewart@nrc.gc.ca

*J. Borghetti, X. Li, M. Pickett, G. Medeiros Ribeiro, W. Robinett, G. Snider, J.P. Strachan, W. Wu, Q. Xia, J. Joshua Yang, R.S.Williams**
Information and Quantum Science Laboratory
Hewlett Packard Laboratories
Palo Alto, CA 94304 USA
*stan.williams@hp.com

***Abstract* — This is a brief review of recent work on the prospective hybrid CMOS/memristor circuits. Such hybrids combine the flexibility, reliability and high functionality of the CMOS subsystem with very high density of nanoscale thin film resistance switching devices operating on different physical principles. Simulation and initial experimental results demonstrate that performance of CMOS/memristor circuits for several important applications is well beyond scaling limits of conventional VLSI paradigm.**

I. Introduction

The prospects to continue the Moore Law with current VLSI paradigm, based on a combination of lithographic patterning, CMOS circuits, and Boolean logic, beyond the 10 nm frontier are at best uncertain [1, 2]. The main reason is that at gate length beyond 10 nm, the sensitivity of parameters (most importantly, the voltage threshold) of MOSFETs to inevitable fabrication spreads grows exponentially. As a result, the gate length should be controlled with a few-angstrom accuracy, far beyond even the long-term projections of the semiconductor industry [3]. This is why alternative electronic devices and circuits are being actively explored now. In this short review we present one very promising paradigm based on the combination of CMOS circuitry and thin film memristors which not only has a potential to prolong Moore's law for digital memories and reconfigurable Boolean logic circuits but might as well revive the field of artificial neuromorphic computing.

II. Memristors and Resistive Switching Devices

A memristor is a 2-terminal electrical circuit element that changes its resistance depending on the total amount of charge that flows through the device [4]. A memristance naturally arises in thin-film semiconductors for which electronic and dopant equations of motion are coupled in the presence of an applied electric field [5]. This property is actually common for nanoscale films and has been observed in a variety of material systems, e.g. transition metal oxides and perovskites, various superionic conductors composed of chalcogenides and metal electrodes, and organic polymer films [6]. Figure 1 shows typical switching I-V of the memristive device based on thin film of titanium dioxide where resistance change is due to electric field assisted modulation of oxygen vacancies profile [7]. Titanium dioxide by itself is wide band semiconductor, i.e. insulating, but even the slightest oxygen nonstochiometry turns it into very good conductor since oxygen vacancy acts as shallow donor. The device shown on the Figure 1 was intentionally fabricated with the excess of oxygen vacancies near the bottom electrode so that application of negative voltage to the top electrode forces positively charged vacancies to drift to the top electrode thus increasing the conductance of the device.

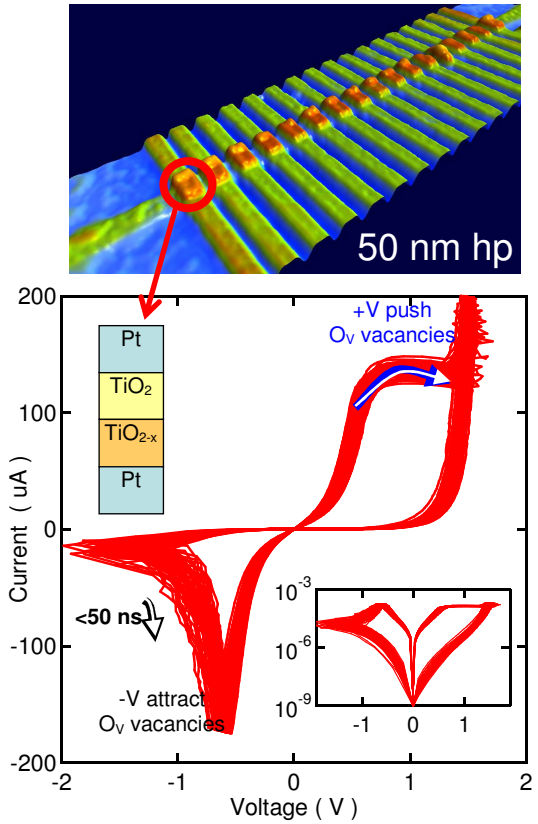

Figure 1. Micrograph of array of memristive devices and typical switching I-Vs [7].

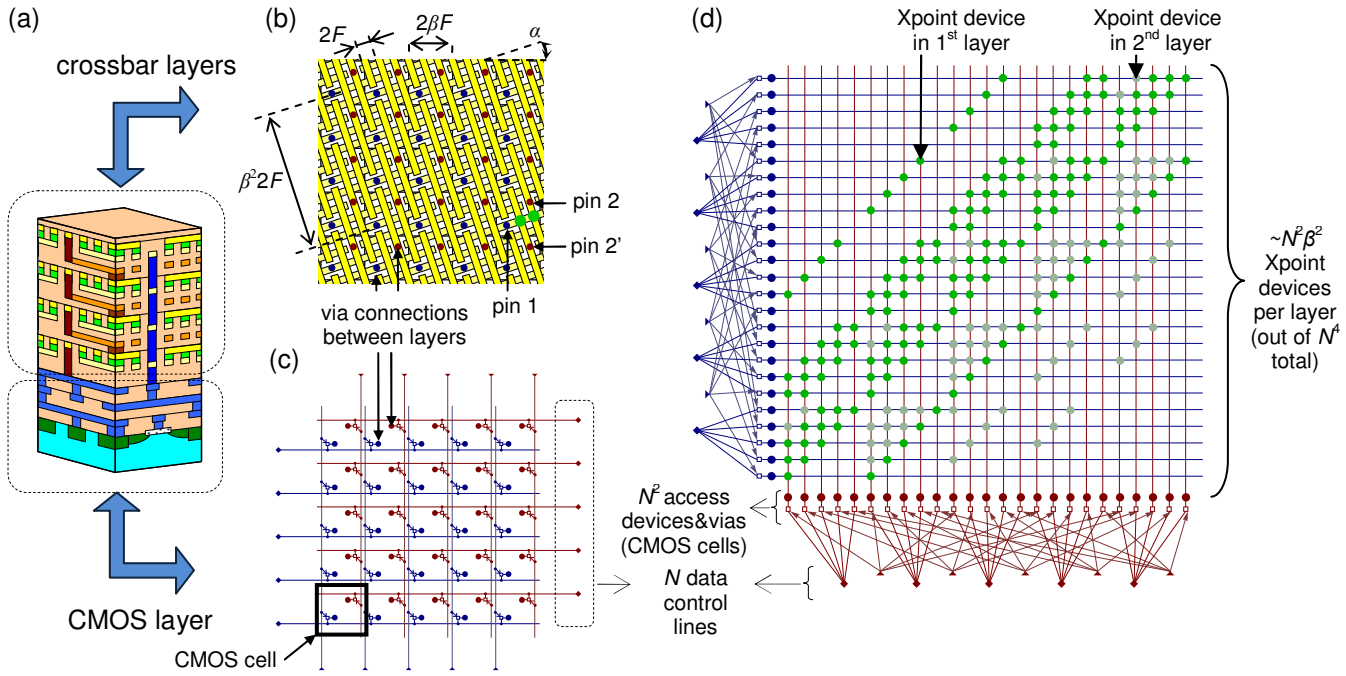

Figure 2. Three-dimensional hybrid CMOS/memristor circuit with area distributed interface. (a) Cut-away illustration of the circuit showing four crossbar layers ($M = 4$), (b) a zoom in of the crossbar structure with several adjecent pins to show that any nanodevice may be addressed via the appropriate pin pair (e.g., pins 1 and 2 for the right of the two shown devices, and pins 1 and 2' for the left device), (c) equivalent circuit diagram of CMOS layer for the $N = 5$ primitive cell array, and (d) equivalent circuit diagram of the virtual crossbar array for the case $N = 5$, $M = 2$, and $r = 3$.

Our group has developed several models which support such mechanism [5, 7-9].

While memristance has been observed experimentally for at least fifty years before it was recognized as such, it has now become interesting for a variety of digital and analog applications, especially because a true memristor does not lose its state when the electrical power is turned off. Other unique properties of memristive devices include the fact that operation of memristive relies only on one critical dimension that is the film thickness and can be easily controlled to few angstroms without the use of expensive fabrication techniques. Additionally, fabrication for most of devices does not require high temperatures enabling back-end integration of multiple layers of memristors with CMOS substrate. Finally, the footprint of memristive devices can be very small (of the order of $4F^2$ where F is a minimum feature size) and essentially determined by the overlap area of two perpendicular wires contacting the device.

III. Hybrid Circuits

On the other hand, memristive devices are not active components, i.e. not the equivalent of the CMOS transistor, because they cannot supply energy to a circuit. The solution to that problem is to complement crossbar arrays of memristive devices with a conventional CMOS substrate that provides signal restoration and gain, but is much less dense. Figure 2 shows one example of such circuits, called CMOL (standing for Cmos + MOLecular scale devices hybrids), which employs area distributed interface [10]. More specifically, as Figures 2b, 2c show, two types of vias, one connecting to the lower (shown with red dots) and the other to the upper (blue dots) wire level in the crossbar, are arranged into a square array with side $2\beta F$ (which is also equal to the side length of the "cells" grouping two vias of each kind). Here β is a dimensionless number larger than 1 that depends on the cell size (i.e. complexity) in the CMOS subsystem. The crossbar is rotated by an angle $\alpha = \arcsin(1/\beta)$ relative to the via array such that vias naturally subdivide the wires into fragments of length $\beta^2 2F$. The factor β is not arbitrary, but is chosen from the spectrum of possible values $\beta = (r^2 + 1)^{1/2}$, where r is an integer so that the precise number of devices on the wire fragment is $r^2 - 1 \approx \beta^2$. The decoding scheme in CMOL is based on two separate address arrays (one for each level of wire in the crossbar so that there are a total of $4N$ edge channels to provide access to two different via controllers (one 'blue' and one 'red') in each of N^2 addressing cells in the CMOS plane. In contrast to standard memory arrays, in CMOL each control and data line pair electrically connects the peripheral input/outputs to a via instead of a single memory element. In turn, each via is connected to a wire fragment in the crossbar. The two perpendicular sets of wire fragments provide unique access to any crosspoint device even for large values of β. For example, selecting pins 1 and 2 provides access to the rightmost of the two shown devices on Fig. 2b, while pins 1 and 2' for the leftmost device. The total number of crosspoint devices that can be accessed by the $N \times N$ array of CMOS addressing cells is $\sim N^2\beta^2$, which can provide a significant multiplicative factor when comparing CMOS to crossbar implementation, especially if the lithographic feature size of the crossbar is smaller than that of the CMOS. An alternate way of viewing this is that one can use complex CMOS circuitry built with a significantly larger feature size to

address regular crossbars built at a much finer lithographic scale.

Figure 2d shows how original CMOL concept with only one layer of crossbar circuits [10] can be extended to three-dimensional system with multiple layers of crossbars [11]. To understand this idea, first, note that selecting only certain pairs of vias results in addressing of crosspoint memristive devices within a single layer since not all wire segments mutually overlap. By physically translating vias of one kind (red on Figure 2a) by certain distance in each new crossbar layer the devices in this layer are addressed with new unique pairs of vias (which are different from pairs of vias in any other crossbar layer). Effectively, each new layer adds $\beta^2 N^2$ crosspoints devices in $N^2 \times N^2$ virtual crossbar (Fig. 2d). The salient feature of such 3D architecture is that theoretical maximum number of layers with the same pitch which can be stacked and still uniquely access all of the crosspoint devices could be very large, of the order of $M = N^2/\beta^2$, thus allowing a significantly increased integration density of memristive devices beyond the scaling limits of lateral feature sizes. Moreover, the circuitry is compatible with state-of-the-art integrated circuit foundries and potentially cost efficient since the features in all crossbar layers are similar and can be patterned using the same set of masks, while adding new layer does not require changes in the CMOS substrate.

IV. APPLICATIONS

A. Digital Memories

The most straightforward application of described hybrid circuits is embedded and stand-alone memories with their simple matrix structure. In such circuits, each memristive device would play the role of a single (or multiple) bit memory cell, while the CMOS subsystem may be used for coding, decoding, line driving, sensing, and input/output functions [10, 12]. The top level architecture of CMOL memories is similar to the traditional ones, while at the block level two pairs of decoders is needed (instead of usual two) to facilitate accessing of a pair of pins independently. The CMOS cell in this case has the simplest structure and consists of just two pass transistors. So far only relatively small crossbar memories have been experimentally demonstrated without integration with CMOS – see, e.g. extensive review of experimental efforts in Refs. 6, 10-13. The major challenge is the requirement of high yield of memristive devices. For example, the simulation results showed that in order to obtain an order-of-magnitude density advantage from the transfer to hybrid memories (such a goal seems natural for the introduction of a novel technology), the single bit device fraction has to be below ~ 1%, even when using relatively powerful BCH error correcting codes [13]. Additionally, strong nonlinearity in I-V characteristics and low write currents in memristive devices are desired for competitive performance (as compared to conventional memories) which are yet to be demonstrated. Nevertheless, it is natural to expect quick progress in CMOS/memristor memories due to very attractive density scaling prospects which may approach hundred terabit per square centimeter over long term [11].

B. Reconfigurable Boolean Logic Circuits

Another application of CMOS/memristor circuits is in reconfigurable Boolean logic circuits such as field programmable gate arrays (FPGA). FPGAs can be thought of a sea of logic gates which can be selectively connected (programmed) after fabrication by uploading the configuration information into chip's memory. In the contemporary FPGAs typically more than 90% of the area is consumed by SRAM-based configuration bits and only a small fraction of chip real estate is used for the actual computation. The basic idea of hybrid CMOL FPGA circuits is to lift all configuration bits, specifically those used for programmable interconnect, to the crossbar layer while keeping the logic gate implementation in CMOS, so that memristive devices act as a routing switching connecting logic gates [14, 15] (Fig. 3). In particular, the structure of the FPGA is similar to that of depicted in Figure 2 with additional logic gates introduced in each cell with gate's inputs typically connected to pins of one (blue) kind and its outputs to the pins of the other (red) kind (Fig 3b). During the configuration stage, logic gates are disabled so that programming memristive devices is similar to the operation of CMOL memory. Note that only a small fraction of nanodevices is typically programmed to the ON state and once programmed, the state of memristive crosspoint device does not change during operation. These circumstances greatly relax the requirement for memristive devices and additionally make the circuits highly defect tolerable. Simulation results have shown that, in addition to high defect tolerance, CMOL FPGA circuits may have extremely high density (more than two orders of magnitude higher than that of usual CMOS FPGA with the same CMOS design rules) while operating at higher speed at acceptable power consumption [14]. Recently, hybrid reconfigurable-logic circuits were fabricated by integrating memristor-based crossbars onto a foundry-built CMOS (complementary metal-oxide-semiconductor) platform using nanoimprint lithography (NIL), as well as materials and processes that were compatible with the CMOS [16].

C. Artificial Neuromorphic Networks

Finally, hybrid circuit technology is uniquely fitted to provide very large (few thousands) connectivity between small (100-gate-scale) CMOS circuitry cells via nanowire-memristor-nanowire links. This enables building cost efficiently large-scale neuromorphic networks. In a several families of neuromorphic network architectures that map uniquely on the CMOL topology [10, 17, 18] neural cell bodies are relatively sparse and implemented in the CMOS subsystem. "Axons" and "dendrites" are implemented as mutually perpendicular nanowires of the CMOL crossbar, while "synapses" that control coupling between the axons and dendrites (and hence between neural cells) are based on memristive devices. Preliminary results indicate that CMOL-based neural nets can be efficiently used for advanced information processing including pattern recognition and classification [18], efficiently perform human visual cortex tasks such as self-organized spatial filtering [17] and, as a matter of principle, enable human brain complexity with a wafer-scale CMOL circuitry [10].

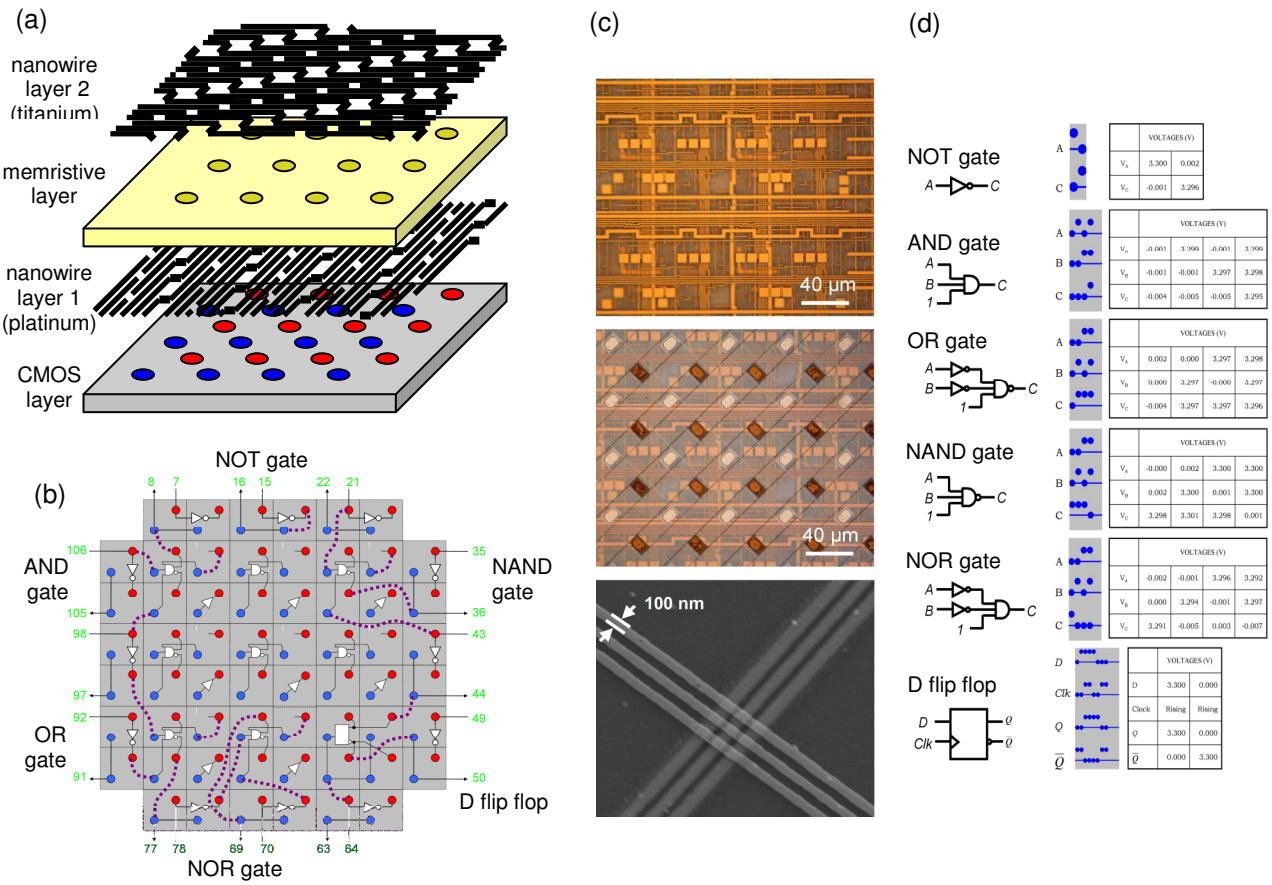

Figure 3. Hybrid CMOS/memristor programmable logic chip: (a) layering; (b) chip floorplan showing experimentally connected CMOS gates; (c) micrographs of the chip; and (d) equivalent circuits and digital logic results from the visualization system of the chip tester for the hybrid circuits with measured truth tables. For the digital signal visualization, the lower blue dots are logic 0 and the upper dots are logic 1.

REFERENCES

[1] K.K. Likharev, "Electronics below 10 nm," in: *Nano and Giga Challenges in Microelectronics*, J. Greer et al., Eds., Amsterdam: Elsevier), pp. 27-68, 2003.

[2] D. J. Frank, D. J. Frank, R. H. Dennard, E. Nowak, P. M. Solomon, Y. Taur, and H. S. P. Wong, " Device scaling limits of Si MOSFETs and their application dependencies," *Proc. IEEE*, vol. 89(3), pp. 259–288, 2001.

[3] International Technology Roadmap for Semiconductors, 2007 Edition.

[4] L.O. Chua, "Memristor – The missing circuit element," *IEEE Transactions on Circuits Theory*, vol. 18, pp. 507-519, 1971.

[5] D. B. Strukov, G. Snider, D. Stewart, R. S. Williams, "The missing memristor found," *Nature*, vol. 453, pp. 80-83, 2008.

[6] D. B. Strukov, "Hybrid integrated circuits for digital electronics," in: *Nanoelectronics and Photonics*, A. Korkin and R. Federico, Eds., Berlin: Springer, pp. 15-57, 2008.

[7] J. J. Yang, M. D. Pickett, X. Li, D. A. A. Ohlberg, D. R. Stewart, and R.S. Williams, "Memristive switching mechanism for metal/oxide/metal nanodevices", *Nature Nanotechnology*, vol. 3, pp. 429-433, 2009.

[8] D. B. Strukov, J. L. Borghetti, and R. S. Williams, "Coupled ionic and electronic transport model of thin-film semiconductor memristive behavior," *Small*, vol. 5(9), pp. 1058-1063, 2009.

[9] M. Pickett, D. B. Strukov, J. Borghetti, J. Yang, G. Snider, D. Stewart, and R. S. Williams, "Switching dynamics in a titanium dioxide memristive device," *J. Applied Physics*, vol. 106, 074508, 2009.

[10] K. K. Likharev and D. B. Strukov, "CMOL: Devices, circuits, and architectures," in *Introducing Molecular Electronics*, G. Cuniberti, G. Fagas, and K. Richter, Eds. Berlin: Springer, pp. 447-478, 2005.

[11] D. B. Strukov and R. S. Williams, "Four-dimensional address topology for circuits with stacked multilayer crossbar arrays," *PNAS*, 2009, in print.

[12] D. B. Strukov and K. K. Likharev, "Defect-tolerant architectures for nanoelectronic crossbar memories," *Journal of Nanoscience and Nanotechnology*, vol. 7, pp. 151-167, 2007.

[13] Y. Chen, G.-Y. Jung, D. A. A. Ohlberg, X. Li, D. R. Stewart, J. O. Jeppesen, K. A. Nielsen, J. F. Stoddart, and R. S. Williams, "Nanoscale molecular-switch crossbar circuits," *Nanotechnology*, vol. 14, pp. 462-468, 2003.

[14] D. B. Strukov and K. K. Likharev, "CMOL FPGA: A cell-based, reconfigurable architecture for hybrid digital circuits using two-terminal nanodevices," *Nanotechnology*, vol. 16, pp. 888-900, 2005.

[15] G.S. Snider and R.S. Williams, " Nano/CMOS architectures using a field-programmable nanowire interconnect" *Nanotechnology*, vol. 18, 035204, 2007.

[16] Q. Xia, W. Robinett, M. W. Cumbie, N. Banerjee, T. J. Cardinali, J. J. Yang, W. Wu, X. Li, W. M. Tong, D. B. Strukov, G. S. Snider, G. Medeiros-Ribeiro and R. S. Williams, " Memristor-CMOS Hybrid Integrated Circuits for Reconfigurable Logic," *Nano Letters*, 2009, vol. 9 (10), pp.3640-3645, 2009.

[17] G. Snider, "Self-organized computation with unreliable, memristive nanodevices," Nanotechnology, vol. 18, pp. 365202, 2007.

[18] K.K. Likharev, A. Mayr, I. Muckra, and Ö. Türel, "CrossNets: High-Performance Neuromorphic Architectures for CMOL Circuits", *Ann. New York Acad. Sci.*, vol 1006, pp. 146-163, 2009