3D Hybrid CMOS/Memristor Circuits: Basic Principle and Prospective Applications

Dmitri B. Strukov
Department of Electrical and Computer Engineering,
University of California at Santa Barbara, Santa Barbara, CA 93106, USA
Email: strukov@ece.ucsb.edu

Abstract: This paper is a brief overview of hybrid CMOS/memristor circuits and their applications for digital memories, programmable logic, and artificial neuromorphic networks.

1 Hybrid Circuits: CMOS/memristor hybrid circuits (also known earlier as CMOL circuits [1]) combine a CMOS subsystem with monolithically integrated several layers of nanowire crossbars, consisting of arrays of two-terminal memristors, all connected by an area-distributed interface between the CMOS subsystem and the crossbars (Fig. 1). This approach combines the advantages of CMOS technology, including its high flexibility, functionality and yield, with the extremely high density of nanowires, nanodevices and interface vias. As a result, the 3D hybrids can overcome limitations pertinent to other 3D integration techniques and enable 3D circuits with unprecedented memory density and aggregate interlayer communication bandwidth at manageable power dissipation.

The key element in hybrid circuits - "memristive" device [2] (Fig. 1c), consists of an active thin film layer of some switching material sandwiched between two metal electrodes. Application of relatively large electrical stress across electrodes changes the resistivity ("memory state") of the thin film material. For many devices the resistance of thin film can be changed continuously so that the memory state is effectively analog. In addition, the memory state of properly engineered devices is nonvolatile and could be read without disturbing it with a relatively small electrical stress [3].

2 Prospective Applications: Applications that can directly benefit from hybrid circuit include those demanding significant amounts of memory access and/or relying of low-precision data (such as imaging, networking, and neuromorphic computing), as well as those whose performance is dominated by interconnects (such as programmable circuits).

2.1 Digital Memories: Hybrid circuits could potentially combine all the desired properties of "universal memory" – high speed, low energy and high endurance of static random access memories, and high density, low cost and nonvolatility of flash memories [1, 2] (Figs. 2a, b). Though most of the research and development efforts are still largely focused on single memristive devices, demonstrations of passive crossbar memories are not uncommon.

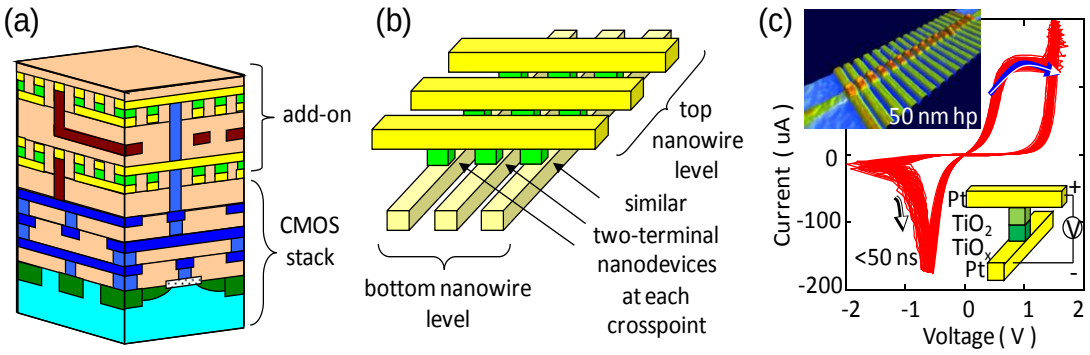

Fig. 1. The basic idea of CMOS/memristor circuits: (a) hybrid circuit cartoon, (b) crossbar topology, and (c) micrograph of array of metal oxide memristive devices, and typical switching *I-V* curves [4].

2.2 Programmable Logic Circuits: Hybrid circuits might eliminate the main inefficiency of field programmable gate arrays (FPGAs) — large overhead associated with memory storing the circuit configuration information [1, 2]. In hybrid FPGA circuits configuration information is stored in memristors which allows to improve density by over 10× and 100× for the conservative (Fig. 2c) and aggressive architectures, respectively, as compared to conventional circuits [1, 2]. The conservative version of hybrid FPGA has been recently demonstrated with nanoimprinted 100-nm scale TiO_{2-x} memristor devices integrated with 0.5-μm CMOS technology [4].

2.3 Artificial Neural Networks (ANN): The motivation behind the development of ANN comes from the fact that the mammalian brain still remains much more efficient for a number of computational tasks, such as pattern recognition and classification, as compared to conventional computers. The structure of the ANN maps naturally on hybrid CMOS/ memristor circuit (Figs. 2d,e,f) and crude estimates have shown that circuits with ultimately scaled CMOS technology and sub-10 nm scale memristive devices could challenge the complexity and connectivity of the human brain [5].

Acknowledgements

The author would like to acknowledge useful discussions with J. Yang, D. Stewart, and the members of the HyNano team.

References

[1] 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.

[2] D.B. Strukov and H. Kohlstedt, "Resistive switching phenomena in thin films: Materials, devices and applications", MRS Bulletin, vol. 37(02), pp. 108-114, February 2012.

[3] F. Alibart, L. Gao, B. Hoskins, and D.B. Strukov, "High-precision tuning of state for memristive devices by adaptable variation-tolerant algorithm", Nanotechnology, vol. 23, art. 075201, 2012.

[4] D.B. Strukov et al., "Hybrid CMOS/memristor circuits", in: Proc. ISCAS'10, pp. 1967-170, Paris, France, May-Jun. 2010.

[5] K. K. Likharev, "CrossNets: Neuromorphic hybrid CMOS/nanoelectronic networks", Science of Advanced Materials, vol. 3, pp. 322-331, Jun. 2011.

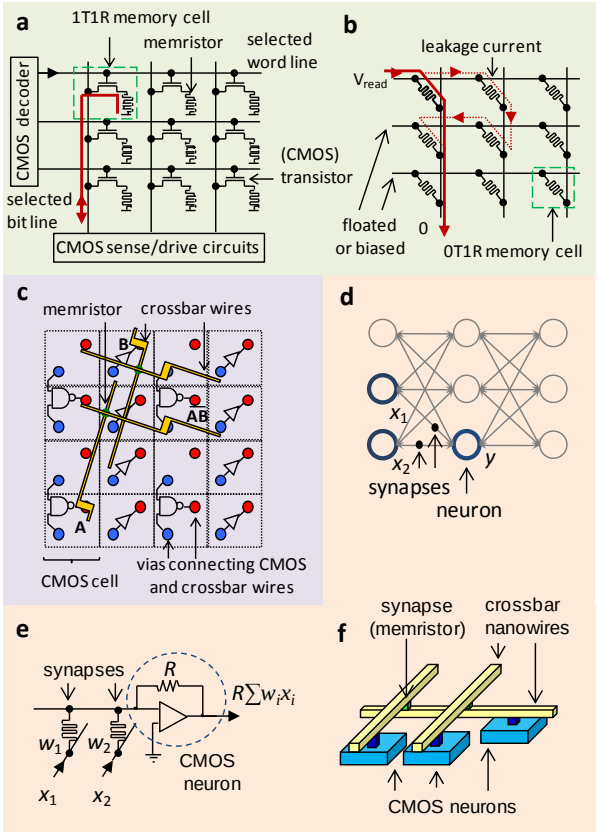

Fig. 2. (a, b) In memories a resistance state represents a bit(s) of information which are read by sensing current flowing through the device at non-perturbative bias. Memory cells contain memristor and select device, and are organized in roughly square arrays, with multiple cells sharing row (select) and column (bit) lines. The select device, which is implemented with transistor in (a) 1T1R cell architecture, enables unique access to a row of cells at a time via bit lines by asserting the corresponding select line. A denser alternative to 1T1R architecture is (b) passive crossbar memory in which select functionality is implemented with (Fig. 1c) diode-like I-V of the crosspoint device. (c) In FPGAs memristor controls whether two wires (e.g., input of some Boolean gates and output of another) are electrically connected. The operation of the FPGA consists of two stages: first, a particular computation task is mapped to the FPGA structure by setting connectivity of the gates via programming to the ON state corresponding memristors. During the second stage, connectivity pattern is fixed (i.e. memristors do not change their state), and FPGA runs a specific task with a programmed circuit.(d, e, f) Crudely, ANN can be represented by (d) a graph with nodes corresponding to neurons and graph edges that correspond to synapses. (e) Operation of the node involves summation of input signals, with each scaled by the corresponding edge weights, followed by a specific threshold function of the node. Memristors implement density-critical configurable analog weights, crossbar wires serve as axons

and dendrides, while CMOS is used for the summing amplifier and thresholding.