Design Implications of Memristor-Based RRAM Cross-Point Structures

Cong Xu†, Xiangyu Dong†‡, Norman P. Jouppi‡, and Yuan Xie†
†Department of Computer Science and Engineering, Pennsylvania State University
‡Intelligent Infrastructure Lab, Hewlett-Packard Labs

***Abstract*— Emerging non-volatile memory (NVM) technologies are getting mature in recent years. These emerging NVM technologies have demonstrated great potentials for the universal memory hierarchy design. Among all the technology candidates, resistive random- access memory (RRAM) is considered to be the most promising as it operates faster than phase-change memory (PCRAM), and it has simpler and smaller cell structure than magnetic memory (MRAM or STT-RAM). In contrast to a conventional MOS-accessed memory cell, memristor-based RRAM has the potential of forming a cross-point structure without using access devices, achieving ultra high density. The cross-point structure, however, brings extra challenges to the peripheral circuitry design. In this work, we study the memristor-based RRAM array design and focus on the choices of different peripherals to achieve the best trade-off among performance, energy, and area. In addition, a system-level model is built to estimate the performance, energy, and area values[1].**

I. Introduction

Memristor, a portmanteau of "memory resistor", is a generalized resistance that maintains a functional relationship between the time integrals of current and voltage. In 1971 memristor was defined by Chua [1] as the fourth fundamental circuit element from the completeness of relations between the four basic circuit variables, namely, current, voltage, charge, and flux-linkage. In 2008, HP Labs [2] presented the first experimental realization as well as a theoretical model of memristor. As rapid progress in nanotechnology has been achieved during the past few years, research has been widely explored in nanoscale bipolar RRAM built with different materials and working under different mechanisms [3], [4]. These devices have demonstrated many distinctive features, including non-volatility, non-linearity, fast access, high density, and good scalability [5].

Memristor-based RRAM is considered as the most promising universal memory technology since it has faster write latency compared to PCRAM and has smaller cell structure compared to MRAM (or STT-RAM). More importantly, Memristor-based RRAM has the potential to build cross-point memory array without access devices because the memristor cell itself has sufficient non-linearity to differentiate the accessing mode and non-accessing mode. However, the cross-point structure also brings extra challenges to the peripheral circuitry design. In this work, we implement a system-level performance, energy, and area model to estimate the impact of different peripheral circuitry choices on the RRAM array design.

Many modeling tools have been developed to enable system-level design exploration for SRAM- or DRAM-based cache and memory design. For example, CACTI [6] is a tool that has been widely used in the computer architecture community to estimate the speed, power, and area of SRAM and DRAM caches. Evans and Franzon [7] developed an energy model for SRAMs and used it to predict an optimum organization for caches. eCACTI [8] incorporated a leakage power model into CACTI. Muralimanohar *et al.* [9] modeled large capacity caches through the use of an interconnect-centric organization composed of mats and request/reply H-tree networks.

In addition, CACTI has also been extended to evaluate the performance, power, and area for STT-RAM [10], PCRAM [11], and NAND flash [12]. However, it is not straightforward to use the CACTI-based models to estimate the performance, energy, and area for ultra high density memristor designs. The conventional peripheral circuitry, which is usually optimized to reduce the memory access latency, would result in a memristor design with very low area efficiency and impair the benefits achieved by the minimum physical size of the memristor cells. As a result, several new array structures, such as cross-point access, non-H-tree organization, external sensing, and minimum-sized row decoder, are proposed in this work. By leveraging these new features, the unique property of the memristor's small cell size can be greatly exploited and an ultra high density non-volatile memory system can be implemented.

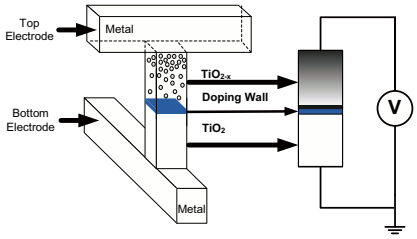

Fig. 1. The structure of memristor cells.

II. Memristor Cross-point Structure Modeling

Figure 1 shows a conceptual view of the memristor structure presented in [2]. The top electrode and bottom electrode are two metal nanowires on platinum, and the thin titanium dioxide film is sandwiched by the electrodes. The resistance of the memory cell can be changed in opposite directions when voltage with different polarities is applied to it. This resistive switching phenomenon results from the oxygen vacancy drift under electric field and has been observed in varieties of materials [13].

Conventionally, memory cells are connected together to form an array and isolated by using MOS access devices as illustrated in Figure 2. In the MOS-accessed array structure, the cell size is dominated by the large MOS device that

[1]C. Xu, X. Dong and Y. Xie were supported in part by NSF grants 0702617, 0903432, 0905365, and SRC grants.

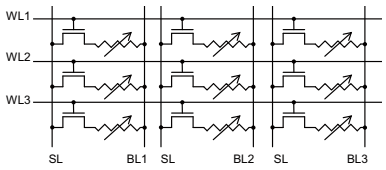

Fig. 2. The schematic view of MOS-accessed memristor arrays (WL=wordline, BL=bitline, SL=sourceline).

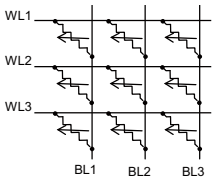

Fig. 3. The schematic view of cross-point memristor arrays without access devices (WL=wordline, BL=bitline).

TABLE I

ARRAY VOLTAGE AND CURRENT IN THE V/2 SCHEME

Item	Voltage	Current
Selected row	V	$N_{sc} \times I(V) + (N_c - N_{sc}) \times I(V/2)$
Selected column	0	$-[I(V) + (N_r - 1) \times I(V/2)]$
Unselected row	V/2	$N_{sc} \times I(V/2)$
Unselected column	V/2	$-I(V/2)$

is necessary to drive enough write current even though the memristor cell itself is much smaller. However, the memristor cell structure as shown in Figure 3 enables the cross-point memory array which dramatically reduces the memory cell size. But it also brings extra challenges in designing the memory peripheral circuitry. In this section, we first describe these challenges.

A. Design Constraint on Array Size

Write half-select problem is endemic in a cross-point array structure. When a memristor cell is selected during a write operation, other memory cells connected to the selected row and columns are subject to the voltage bias and defined as half-select cells. In this work, we analyze the write half-select problem in *V/2 biasing scheme* [14], which is the common accessing method in a cross-point structure. Table I lists the details of V/2 biasing[2].

In the *V/2 biasing scheme*, the write operation should guarantee that the states of these half selected memristor cells do not change over a specified number of write cycles or a specified time with V/2 applied. The half-select cells serve as current dividers in the selected row and columns, preventing the array size from growing unbounded since the available driving current is limited. The minimum current that a column write driver should provide is determined by the case when all the half-select cells are in their low resistance state (LRS),

$$I_{driver} = I_{reset} + (N_r - 1) \times I(V_{reset}/2) \quad (1)$$

where I_{reset} and V_{reset} are the reset current and reset voltage.

Non-linearity of memristor is reflected by the fact that the current through memristor is not directly proportional to the voltage applied on it, which means non-constant resistance of memristor. We here define a non-linearity coefficient to quantify the current divider effect of the half selected memory cells as follows,

$$K_r(p, V) = p \times \frac{R(V/p)}{R(V)} \quad (2)$$

where $R(V/p)$ and $R(V)$ are equivalent static resistance of memristor biased at V/p and V, respectively.

[2] N_r, N_c, and N_{sc} are the numbers of total rows, total columns, and selected columns per row

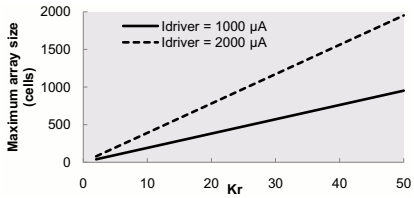

Fig. 4. Maximum array size versus non-linearity and driving current.

Then, we derive the upper limit in a cross-point memory array size when using the V/2 scheme by examining the worst case scenario, i.e., all the cells connected to the selected row/column are in LRS:

$$N_r = \left(\frac{I_{driver}}{I_{reset}} - 1\right) \times K_r\left(2, V_{reset}\right) + 1 \quad (3)$$

$$N_c = \left(\frac{I_{driver}}{I_{reset}} - N_{sc}\right) \times K_r\left(2, V_{reset}\right) + N_{sc} \quad (4)$$

where I_{driver} is the maximum driving current that the write driver attached to the selected row/column can provide. Thus, N_r and N_c are the maximum numbers of rows and columns in a cross-point array.

As shown in Figure 4, the maximum cross-point array size increases with larger current driving capability or larger non-linearity coefficient. More driving current means larger driver size, which results in more area and energy penalty. On the other side, there are two directions in getting a larger K_r: to build a device with better intrinsic non-linearity which needs more engineering control and optimization in fabrication; or to use another biasing scheme during write operations, e.g., from the V/2 scheme to V/3 scheme since $K_r(3, V_{reset})$ should be much larger than $K_r(2, V_{reset})$ according to the non-linearity of memristors.

B. Design Challenges in Write Operations

Two-step write operations are required in the cross-point structure when multiple cells are selected in a single row. In that case no reasonable voltage on the unselected row can be applied: if the voltage applied on the unselected row is smaller than that on the selected row, the cell between the unselected row and the selected column that write a "0" will be undesirably RESET since the voltage drop on the cell is larger than $|V_{reset}|$; else the cell between the unselected row and the selected column that write a "1" will be SET undesirably since the voltage drop on the cell is larger than $|V_{set}|$. Therefore, SET and RESET operations can not be performed simultaneously in the cross-point structure.

We propose two write methods. The first one separates SET and RESET operations (as Figure 5 shows). The second one erases all the cells in the selected row before the selective RESET operation (as Figure 6 shows). For example, the 4-bit word to write is "0101". In SET-before-RESET method, we first write "x1x1" ("x" here means bias row and column of the corresponding cells at the same voltage to keep their original states) and then write "0x0x". While in ERASE-before-RESET method, we first SET all the four cells, that is, to write "1111" and then write "0x0x". The second method has smaller write latency since the erase operation can be performed before the arrival of the column selector signal

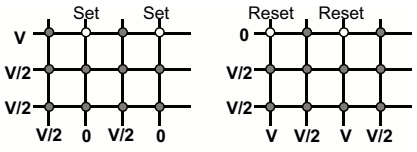

Fig. 5. Sequential write method: SET-before-RESET.

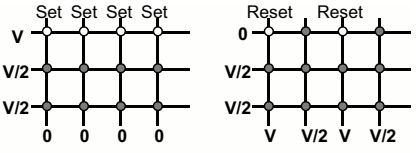

Fig. 6. Sequential write method: ERASE-before-RESET.

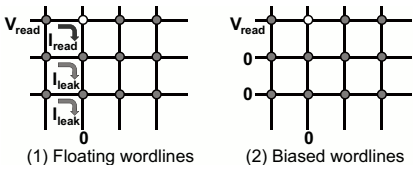

Fig. 7. The read sneak leakage path and the optimized biasing scheme for read operations.

but it needs more write energy due to the redundant SET on the cells that are RESET back in the second step. Here, ERASE-before-RESET is chosen rather than ERASE-before-SET because SET operation usually consumes less energy than RESET operation does.

C. Design Challenges in Read Operations

Read sneak leakage path is another issue in the cross-point structure as the total half-select cell current may lead to a read failure, especially when the selected memristor is in high resistance state (HRS) and the half-select cells are in low resistance state (LRS). Therefore, in order to alleviate read disturbance, as shown in Figure 7, one of the solutions is to bias all the unselected rows at the same voltage with the selected column. However, the IR drop and other variations may make the voltage drop on the unselected cells that are connected to the selected column slightly away from the ideal zero, thus limiting the maximum array size as well.

Another circuit-level solution is a two-step sampling method which could isolate the noise current from the parasitic half select cells. We first read the background current of the half-select cells and latch it. In the next step, the total current of the background current plus the current through the selected memristor cell is read. Finally, after removing of the background current, the state of the selected memory cell is identified by the sensing scheme.

III. Peripheral Circuitry Designs

Peripheral circuitry designs can vary across the specification spectrum from area-optimized to latency-optimized or energy-optimized. In this section, we discuss several peripheral circuitry design options that are necessary to build memristor-based RRAM based on different design optimization goals.

A. Design Choices of Sense Amplifiers

Different sensing schemes have their impacts on the trade-off among performance, energy, and area. In this work, we propose three sensing schemes: current sensing, current-in voltage sensing, and voltage-divider sensing. In the current sensing scheme as shown in Figure 8, the memristor status is read out by measuring the resulting current through the selected memristor cell when a read voltage is applied: the current on the bit-line is compared to the reference current generated by reference cells, the current difference is amplified by current-mode sense amplifiers, and they are eventually converted to voltage signals. Figure 9 demonstrates an alternative sensing method by applying a current source on the selected memristor cell and sensing the voltage via the voltage-mode sense amplifier. The voltage-divider sensing scheme is presented by introducing a resistor (R_x) in series with the memristor cell as illustrated in Figure 10. The resistance value is selected to achieve the maximum read sensing margin, and it is calculated as follows,

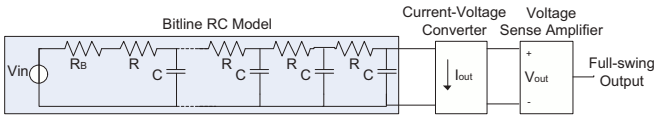

Fig. 8. Analysis model for memristor current sensing scheme.

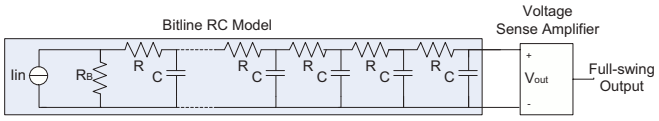

Fig. 9. Analysis model for memristor current-in voltage sensing scheme.

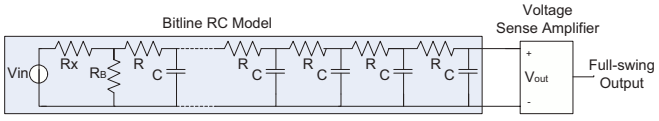

Fig. 10. Analysis model for memristor voltage-divider sensing scheme.

$$R_x = \sqrt{R_{on} \times R_{off}} \tag{5}$$

where R_{on} and R_{off} are the memristor cell resistance values in LRS and HRS, respectively.

We model the bit-line RC delay and power consumption analytically for each sensing scheme. The most significant difference between the current-mode sensing and voltage-mode sensing is that the input resistance of ideal current-mode sensing is zero while that of ideal voltage-mode sensing is infinite. And, the most significant difference between current-in voltage sensing and voltage-divider sensing is that the internal resistance of a ideal current source is infinite while the resistor R_x serving as a voltage divider can be treated as the internal resistance of a voltage source. Delays of current-in voltage sensing, voltage-divider sensing and current sensing are given by the follows equations using Seevinck's delay expression [15]:

$$\delta t_v = \frac{R_T C_T}{2} \times \left(1 + \frac{2R_B}{R_T}\right) \tag{6}$$

$$\delta t_{vd} = \frac{R_T C_T}{2} \times \left(1 + \frac{2(R_B || R_x)}{R_T}\right) \tag{7}$$

$$\delta t_i = \frac{R_T C_T}{2} \times \left(\frac{R_B + \frac{R_T}{3}}{R_B + R_T}\right) \tag{8}$$

where R_T and C_T are the total line resistance and capacitance, R_B is the pull-down resistance of the memristor cell, and R_x

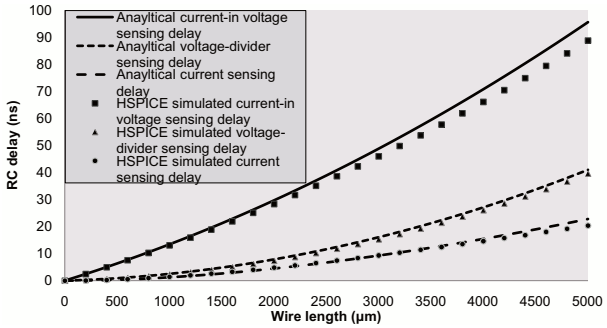

Fig. 11. Delay model verification of three sensing schemes comparing to HSPICE simulations.

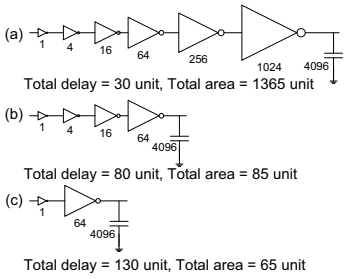

Fig. 12. Buffer design optimization: (a) latency-optimized; (b) balanced; (c) area-optimized.

is the resistance of voltage divider. In these equations, t_v, t_{vd}, and t_i are the RC delays of current-in voltage sensing, voltage-divider sensing, and current sensing schemes, respectively. $R_x||R_B$, instead of R_B, is used as the new effective pull-down resistance in Equation 7 according to the transformation from a Thevenin Equivalent to a Norton Equivalent.

Equation 6 and 7 show that voltage-divider sensing is faster than current-in voltage sensing with the extra cost of fabricating a large resistor. Comparing Equation 8 with Equation 6 and 7, we can see the current sensing is much faster than current-in voltage sensing and voltage-divider sensing since the former delay is less than the intrinsic line delay $R_T C_T/2$ while the latter delays are larger than $R_T C_T/2$. The bit-line delay analytical models are verified by comparing them with the HSPICE simulation results. As shown in Figure 11, the RC delays derived by our analytical RC models are consistent with the HSPICE simulation results.

B. Design Choices of Output Buffers

The conventional way of designing a output buffer is to calculate the inverter chain stage and the size of each inverter using logical effort for a given load capacitance. This is always optimized from latency perspective. However, for the cross-point structure of memristor array, a word-line row driver with large current driving capability is required so that sufficient amounts of current can pass through the memristor cells during write operations. This gives a lower limit of driver size and nullifies the area advantage of memristor cross-point memory array.

Therefore, we offer three buffer design choices in our model: one optimizing latency, one optimizing area, while another balancing latency and area. An example is illustrated in Figure 12 demonstrating the different sizing methods when an output buffer with 4096 times the capacitance of a minimum-sized inverter is to be designed. In a latency-optimized buffer design, the number of stages and all of the inverter sizing in the inverter chain is calculated by logical effort to achieve minimum delay (30 units) while paying a huge area penalty (1365 units). In an area-optimized buffer design, there are only two stages of inverters, and the size of the last stage is determined by the minimum driving current requirement. This type of buffer has the minimum area (65 units), but is much slower than the latency-optimized buffer. The balanced option determines the size of last stage inverter by its driving current requirement and calculates the size of the other inverters by logical effort. This results in a balanced delay and area metric.

TABLE II
MEMRISTOR TECHNOLOGY ASSUMPTIONS

	MOS-accessed	Cross-point
Cell size	$20F^2$	$4F^2$
Maximum NMOS driver size	$100F$	
RESET voltage and pulse duration	$2.0V$, $100ns$	
SET voltage and pulse duration	$-2.0V$, $100ns$	
READ input	$0.4V$ voltage source, or $2\mu A$ current source	
LRS resistance	$10k\Omega$	
HRS resistance	$500k\Omega$	
Half-select resistance	-	$100k\Omega$

C. Interval vs. External Sensing

Desirable memristor cross-point structures have relatively small cell sizes and also the design constraints of building large cell arrays. However, including sense amplifiers in each array can occupy a dominant portion of the total array area. As a result, in order to achieve a high-density memristor array, it is necessary to move the sense amplifiers out of the array and thus to use external sensing.

Figure 13 shows a common H-tree organization that connects all the sense amplifier-included arrays together. In contrast, a new external sensing organization is proposed in this work and is illustrated in Figure 14. In this external sensing scheme, all the sense amplifiers are located at the bank level and the output signals from each sense amplifier-free array are partial-swing. It is obvious that the external sensing scheme has much higher area efficiency compared to its internal sensing counterpart. However, as a penalty, sophisticated global interconnect technologies, such as repeater inserting, cannot be used in the external sensing scheme since all the global signals are partial-swing before passing through the sense amplifiers. Later in Section IV, the area-performance trade-off between internal and external sensing is evaluated.

IV. MEMRISTOR DESIGN OPTIMIZATION

In this section, we use the proposed model to explore the memristor design space and find optimized solutions for area, latency, and energy. In our evaluation, we use the memristor technology assumptions as listed in Table II.

A. Area Optimization

The non-linearity of memristor cells is a promising property that enables the capability of building cross-point structure with minimum achievable cell size of only $4F^2$ compared to the traditional MOS-accessed structure that requires cell

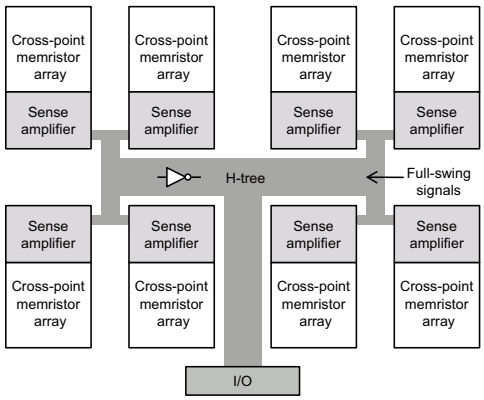

Fig. 13. An example of internal sensing using the H-tree routing organization.

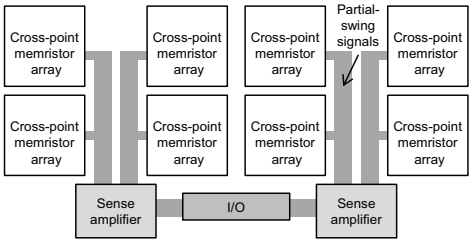

Fig. 14. An example of external sensing using the non-H-tree routing organization.

TABLE III

COMPARISON BETWEEN MOS-ACCESSED AND CROSS-POINT DESIGNS

	MOS-accessed	Cross-point (baseline)	Cross-point (area-optimized)
Process Node	32nm		
Capacity	8MB		
Cell Size	$20F^2$	$4F^2$	$4F^2$
Array Size	2048×8192	512×512	512×512
Total Area	$1.63mm^2$	$1.79mm^2$	$0.72mm^2$
Area Efficiency	84.3%	15.3%	38.4%

sizes of $20F^2$ or more. However, while cross-point structures have high-density cells, without proper area optimization, the peripheral circuitry becomes the dominant part of the chip area and nullifies the area saving from the memristor cell size.

As Table III shows, 515×512 is the maximum array size for cross-point memory structure calculated by Equation 3 and 4 under the assumptions in Table II. Therefore, a conventional peripheral circuitry design for latency optimization (using internal current-sensing, multiple-level output buffer, and H-tree organization) can lead to an unacceptable low area efficiency. Table III shows such a baseline cross-point design has an area efficiency of only 15.3%. As a result, the actual cross-point memristor chip area is larger than that of its MOS-accessed counterpart even if the MOS-accessed memristor cell size is five times larger.

Figure 15 illustrates the ratio of each component that contributes to the total memristor chip area. As shown by the results, using a cross-point structure without building MOS access devices for each cell dramatically reduces the area ratio of memristor cells. As the peripheral circuitry starts to dominate the total chip area, using external voltage-sensing and area-aware buffer design become the key techniques for achieving an area-optimized memristor design.

To have a more insightful area evaluation, Figure 16 and

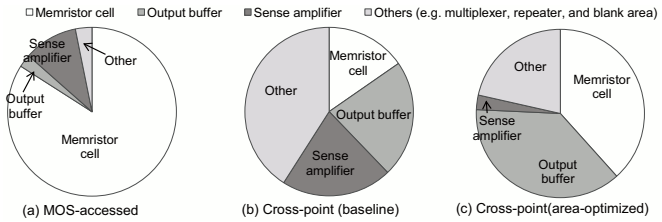

Fig. 15. Area ratio comparison between MOS-accessed and cross-point designs.

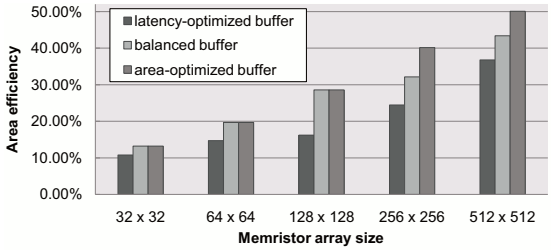

Fig. 16. Memristor array with buffers of different optimization targets (current-in voltage sensing is used for a fair comparison).

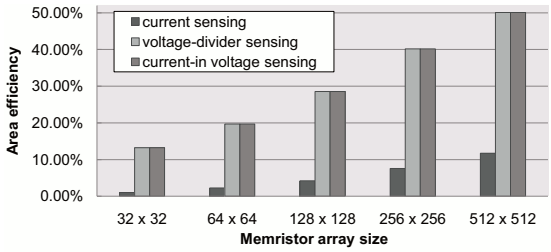

Fig. 17. Memristor array area efficiency with different sensing schemes (area-optimized buffer design is used for a fair comparison).

Figure 17 show the effects of different design choices of buffers and sense amplifiers. Figure 16 shows that switching from latency-optimized buffer to balanced and area-optimized buffer increases area efficiency. This is because the load capacitance of row decoder buffer increases with the number of cells driven by it, which results in more stages in the inverter chain by logical effort but area-optimized design can avoid such area overhead. Figure 17 shows that voltage-sensing always achieves higher area efficiency compared to the current-sensing. This is due to the fact that the chip I/Os are voltage signals and the current-sensing scheme needs to include an I-V converter for the final output stage, which occupies large chip area. However, current-sensing is necessary for latency-optimized design targets as later shown in the next discussion.

B. Latency and Energy Optimization

In addition to area optimization, our model can also be used to explore different design options and find optimized memristor-based RRAM chip design styles depending on the optimization goals (such as read/write latency, read/write dynamic energy, and leakage power). Table IV tabulates the design space exploration of a 32nm 8MB memristor-based random access memory. As the table shows, different peripheral designs are used for different optimization targets and each design is valuable in certain circumstances.

TABLE IV

MEMRISTOR CHIP DESIGN OPTIMIZATIONS WITH DIFFERENT GOALS

	Area opt.	Read latency opt.	Write latency opt.	Read energy opt.	Write energy opt.	Leakage opt.
Area (mm^2)	**0.664**	5.508	8.071	2.971	3.133	1.399
Read latency (ns)	107.1	**1.773**	1.917	5.711	6.182	426.8
Write latency (ns)	204.3	200.7	**100.6**	202.8	203.1	518.2
Read energy (nJ)	1.884	0.195	0.234	**0.012**	0.014	4.624
Write energy (nJ)	13.72	25.81	13.06	12.82	**12.81**	12.99
Leakage (mW)	1372	3872	7081	6819	7841	**26.64**
Array structure	Cross-point	Cross-point	MOS-accessed	Cross-point	Cross-point	MOS-accessed
Inter-array routing	Non-H-tree	H-tree	H-tree	H-tree	H-tree	Non-H-tree
Sense amp placement	External	Internal	Internal	Internal	Internal	External
Sense amp type	Current-in voltage	Current	Current	Voltage-divider	Voltage-divider	Voltage-divider
Write method	SET-before-RESET	Erase-before-RESET	Normal	SET-before-RESET	SET-before-RESET	Normal
Interconnect wire	Normal	Repeated	Repeated	Low-swing	Low-swing	Normal
Output buffer type	Area opt.	Latency opt.	Latency opt.	Area opt.	Area opt.	Area opt.

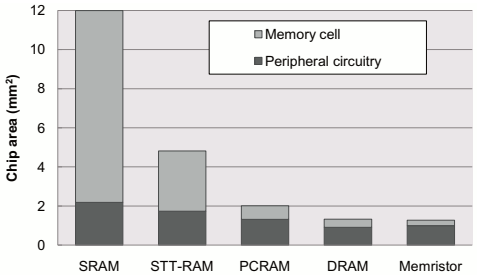

Fig. 18. Chip area of 32nm 8MB memory of using different technologies.

V. MEMORY TECHNOLOGY COMPARISON

Five 8MB memory macros of different memory types - SRAM, STT-RAM (or MRAM), PCRAM, DRAM and memristor have been simulated. The areas of memory cell array and peripheral circuity for each macro are illustrated in Figure 18. The memory cell size factors chosen in the simulation are $146F^2$ for SRAM [6], $6F^2$ for DRAM [6], $45F^2$ for STT-RAM [16], $10F^2$ for PCRAM [16], and $4F^2$ for cross-point memristor. Voltage sensing scheme for all memory macros was assumed in the simulations for fair comparison. It can be seen that although the memristor memory cell size is only 2% of the SRAM cell size, the area of cross-point memristor memory macro is more than 10% of the SRAM macro due to poor peripheral circuitry scaling. This comparison result reiterates the importance of proper peripheral circuitry design as its area occupation ratio increases with the reduced memory cell size.

VI. CONCLUSION

Memristor-based RRAM is one of the most promising emerging memory technologies and has the potential of being a universal memory technology. Its attractive properties include fast access, zero standby leakage, and non-volatility. Moreover, the non-linearity of memristor cells makes it feasible to build cross-point structures with ultra high density. However, the cross-point memristor design brings extra challenges to the memory peripheral circuitry. In this work, we evaluated various data sensing schemes, buffer designs, and inter-array organizations to meet a large spectrum of the memristor optimization goals. In addition, a new performance, energy, and area model for memristor-based RRAM was developed. This model is expected to help and accelerate the design evaluation and exploration of the next-generation of computer memory/storage subsystems that rely on memristor technology.

REFERENCES

[1] L. Chua, "Memristor - The missing circuit element," *IEEE Transactions on Circuit Theory*, vol. 18, no. 5, pp. 507–519, 1971.

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

[3] S. Yu, B. Gao, H. B. Dai, B. Sun, *et al.*, "Improved uniformity of resistive switching behaviors in HfO_2 thin films with embedded Al layers," *Electrochemical and Solid-State Letters*, vol. 13, no. 4, pp. H36–H38, 2010.

[4] L. Chen, Y. Xu, Q.-Q. Sun, H. Liu, *et al.*, "Highly uniform bipolar resistive switching with Al_2O_3 buffer layer in robust $NbAlO$-based RRAM," *IEEE Electron Device Letters*, vol. 31, no. 4, pp. 356–358, 2010.

[5] R. Williams, "How we found the missing memristor," *IEEE Spectrum*, vol. 45, no. 12, pp. 28–35, 2008.

[6] S. Thoziyoor, N. Muralimanohar, J.-H. Ahn, and N. P. Jouppi, "CACTI 5.1 technical report," HP Labs, Tech. Rep. HPL-2008-20, 2008.

[7] R. J. Evans and P. D. Franzon, "Energy consumption modeling and optimization for SRAM's," *IEEE Journal of Solid-State Circuits*, vol. 30, no. 5, pp. 571–579, 1995.

[8] M. Mamidipaka and N. Dutt, "eCACTI: An enhanced power estimation model for on-chip caches," Center for Embedded Computer Systems, Tech. Rep. TR04-28, 2004.

[9] N. Muralimanohar, R. Balasubramonian, and N. P. Jouppi, "Architecting efficient interconnects for large caches with CACTI 6.0," *IEEE Micro*, vol. 28, no. 1, pp. 69–79, 2008.

[10] X. Dong, X. Wu, G. Sun, Y. Xie, H. Li, *et al.*, "Circuit and Microarchitecture Evaluation of 3D Stacking Magnetic RAM (MRAM) as a Universal Memory Replacement," in *Proceedings of the Design Automation Conference*, 2008, pp. 554–559.

[11] X. Dong, N. P. Jouppi, and Y. Xie, "PCRAMsim: System-level performance, energy, and area modeling for phase-change RAM," in *Proceedings of the International Conference on Computer-Aided Design*, 2009, pp. 269–275.

[12] V. Mohan, S. Gurumurthi, and M. R. Stan, "FlashPower: A detailed power model for NAND flash memory," in *Proceedings of Design, Automation and Test in Europe*, 2010, pp. 502–507.

[13] R. Waser and M. Aono, "Nanoionics-based resistive switching memories," *Nature Materials*, vol. 6, no. 11, pp. 833–840, 2007.

[14] Y.-C. Chen, C. Chen, C. Chen, J. Yu, *et al.*, "An access-transistor-free (0T/1R) non-volatile resistance random access memory (RRAM) using a novel threshold switching, self-rectifying chalcogenide device," in *Proceedings of the International Electron Devices Meeting*, 2003, pp. 37.4.1–37.4.4.

[15] E. Seevinck, P. van Beers, and H. Ontrop, "Current-mode techniques for high-speed VLSI circuits with application to current sense amplifier for CMOS SRAM's," *IEEE Journal of Solid-State Circuits*, vol. 26, no. 4, pp. 525–536, 1991.

[16] International Technology Roadmap for Semiconductors, "Process Integration, Devices, and Structures 2009 Edition," http://www.itrs.net/.