Energy-Efficient Write Scheme for Nonvolatile Resistive Crossbar Arrays With Selectors

Albert Ciprut⬢, *Student Member, IEEE*, and Eby G. Friedman, *Fellow, IEEE*

Abstract—The write operation of a resistive memory based on a one-selector–one-resistor (1S1R) crossbar array consumes significant energy and is dependent on the device and circuit characteristics as well as the bias scheme. In this paper, the energy efficiency of a crossbar array of a 1S1R configuration during a write operation is explored for the $V/2$ and $V/3$ bias schemes. The characteristics that affect the most energy-efficient bias scheme are demystified. The write energy of a crossbar array is modeled in terms of the array size, number of selected cells, and nonlinearity factor. For a specific array size and selector technology, the number of selected cells during a write operation can affect the choice of bias scheme. The effect of leakage current due to partially biased unselected cells is explored. Furthermore, an energy-efficient write operation based on a hybrid bias scheme is proposed to reduce the write energy. This write operation adaptively sets the bias scheme based on the number of selected cells to enhance overall energy efficiency. Energy improvements of more than two times are demonstrated with this hybrid bias scheme.

Index Terms—Bias scheme, crossbar array, emerging memory technologies, nonlinearity factor, resistive RAM (RRAM), write energy.

I. INTRODUCTION

RESISTIVE memories are expected to replace charge-based conventional memories due to scalability limitations and energy benefits due to nonvolatility characteristics. Resistive memory devices, such as resistive RAM (RRAM), phase change memory, and magnetoresistive RAM (MRAM), have been explored for nonvolatile memories [1]–[4]. To achieve high density, these resistive devices are placed within a crossbar array structure. The area of a memory cell in an RRAM-based crossbar array utilizing a two-terminal one-selector–one-resistor (1S1R) configuration can be as low as $4F^2$, where F is the minimum feature size of a technology node [3]. These arrays can be placed within the metal layers, supporting cell placement above the CMOS logic, further reducing area. Moreover, a crossbar array can be configured as

a logic gate, providing a path to non-von Neumann in-memory computing [5]. To enable this capability, however, the energy consumption of a crossbar array should be within practical limits due to thermal design power envelope constraints in high-performance integrated circuits and the limited battery size of mobile devices. The energy consumption of 1S1R memories increases significantly as the size of the array grows. In particular, the write energy is a large portion of the total energy and is significantly greater than the read energy [6]. This difference is due to the long switching times of the selected devices, whereas the read latency primarily depends upon the sense amplifier, which improves with technology scaling. The write latency is typically on the order of a few hundred nanoseconds, whereas the read latency can be as low as 5 ns [7].

The energy consumption of a crossbar array during a write operation depends upon the bias scheme, typically a $V/2$ or $V/3$ bias scheme [8], [9] (see Section II). While most of the work described in the literature considers the $V/2$ bias scheme [6], [10], the advantages of one bias scheme over the other bias scheme are unclear in terms of energy efficiency. Furthermore, the $V/2$ bias scheme is often claimed to be more energy efficient than the $V/3$ bias scheme [1], [11]. The most energy-efficient bias scheme can, however, vary depending upon the circuit and device characteristics. In particular, the selector device has a profound effect on the energy consumption, since the leakage currents due to partially biased cells increase with array size. The selector device within a crossbar array suppresses the currents under a low voltage bias while supporting higher currents under a high voltage bias. Different three-terminal devices, such as an MOS transistor and a bipolar junction transistor, as well as two-terminal devices, such as a silicon-based diode and a metal–insulator–metal (MIM) tunneling barrier, have been considered [12], [13]. The three-terminal transistor based-selector devices provide greater isolation between the selected cells and unselected cells within a crossbar array. This solution, however, significantly increases cell area and inhibits scalability. Two-terminal selectors can, however, be vertically integrated within a nonvolatile resistive cell, preserving area. A wide range of two-terminal selectors exist that can be classified into two categories: unipolar and bipolar. In addition, depending upon the material and the nonvolatile resistive cell, the selector can be a silicon-based diode, a self-rectifying device, or MIM with different kinds of tunneling mechanisms depending upon the thickness of the insulator material [14]. In this paper, a 1S1R element is used to refer to a nonvolatile resistive cell integrated with a two-terminal selector device.
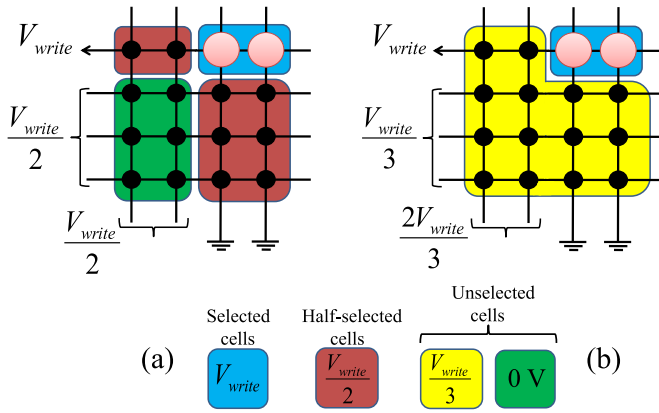
Fig. 1.   Bias schemes for a two-bit write operation. (a) $V/2$ bias scheme. (b) $V/3$ bias scheme.

To incorporate the effects of the selector within an array, the nonlinearity factor is used as the primary metric to quantify the isolation capability (see Section II).

In this paper, the write bias schemes are compared from an energy efficiency point of view for 1S1R crossbar arrays with two-terminal selectors. It is shown here that the bias scheme that provides the highest energy efficiency depends upon several parameters, such as the nonlinearity factor of the selectors, the size of the array, and the number of selected cells during a write operation. Simple closed-form expressions that model the write energy of a crossbar array in terms of these parameters, excluding the peripheral circuitry, are provided for the case where the interconnect resistance is negligible. The models are applicable to both unipolar and bipolar devices. Most of the existing work described in the literature does not consider the effects of writing multiple bits (i.e., multiple selected cells) on the energy consumption of an array. In [6] and [10], the power consumption when selecting multiple bits is considered, however, only for the $V/2$ bias scheme. In this paper, the effects of writing multiple bits on the energy efficiency of different bias schemes are explored for the first time. Moreover, the effects of leakage current on energy consumption are discussed. In addition, an energy-efficient write scheme is proposed that adaptively utilizes both the $V/2$ and $V/3$ bias schemes to lower the write energy, extending the preliminary work described in [15]. Based on the proposed write operation, the bias scheme is altered for maximum energy efficiency depending upon the number of selected bits, which can vary for different write operations. In Section II, the bias schemes during a write operation are reviewed. In Section III, models of the energy consumption are described. In Section IV, the proposed energy-efficient write scheme is described. The potential challenges and overhead are also discussed. In Section V, some conclusions are offered.

II. Write Operations

The two types of write bias schemes, $V/2$ and $V/3$, are shown in Fig. 1. For the $V/2$ bias scheme, the selected wordline is connected to the write voltage, while the selected bitlines are grounded. The unselected wordlines as well as bitlines are biased to half of the write voltage. Similarly, for the $V/3$ bias scheme, the selected wordline is connected to the write voltage, while the selected bitlines are grounded. The unselected wordlines are biased at one third of the write voltage, whereas the unselected bitlines are biased at two thirds of the write voltage. The voltage drop across the unselected cells along the selected wordline and selected bitlines, also called the half-selected cells, is, therefore, biased at one half of the write voltage for the $V/2$ bias scheme. For the $V/3$ bias scheme, this voltage decreases to one third of the write voltage. More importantly, the cells on the unselected wordlines and bitlines are at 0 V for the $V/2$ bias scheme and at one third of the write voltage for the $V/3$ bias scheme, resulting in a large number of cells leaking current when the $V/3$ bias scheme is applied.

The leakage current of the unselected cells depends upon the nonlinearity factor of the selector. The two-terminal selector is placed above a resistive cell to form a nonlinear $I–V$ characteristic. A selector with a higher nonlinearity factor further decreases the current of the cell when biased below the threshold voltage of the selector [16]. The leakage current due to the partially biased unselected cells is, therefore, suppressed, decreasing IR voltage drops and supporting larger array sizes [17]. The nonlinearity factor of a selector is the ratio of the current passing through a selected cell to the current passing through a half-selected cell. The nonlinearity factor of the $V/2$ and $V/3$ bias schemes is, respectively,

$$K_{V/2} = \frac{I_{\text{cell}}(V_{\text{write}})}{I_{\text{cell}}(V_{\text{write}}/2)} = 2 \times \frac{R_{\text{ON}@V_{\text{write}}/2}}{R_{\text{ON}}} \qquad (1)$$

$$K_{V/3} = \frac{I_{\text{cell}}(V_{\text{write}})}{I_{\text{cell}}(V_{\text{write}}/3)} = 3 \times \frac{R_{\text{ON}@V_{\text{write}}/3}}{R_{\text{ON}}} \qquad (2)$$

where $I_{\text{cell}}(V_{\text{write}})$, $I_{\text{cell}}(V_{\text{write}}/2)$, and $I_{\text{cell}}(V_{\text{write}}/3)$ are, respectively, the current passing through the cell when the cell voltage is equal to the write voltage, one half of the write voltage, and one third of the write voltage. R_{ON}, $R_{\text{ON}@V_{\text{write}}/2}$, and $R_{\text{ON}@V_{\text{write}}/3}$ are, respectively, the cell resistance during an ON state when the cell voltage is equal to the write voltage, one half of the write voltage, and one third of the write voltage. The leakage current, therefore, depends upon the bias scheme, which is related to the nonlinearity factor.

The nonlinearity factor $K_{V/2}$ of a 1S1R device is typically on the order of 10^1 to 10^2, whereas $K_{V/3}$ is on the order of 10^3 to 10^4 [16], [18]–[24]. A selector device with an ON/OFF ratio as high as 10^8 has recently been demonstrated [25]. The choice of bias scheme can, therefore, greatly affect the energy consumption.

III. Energy Models

In this section, a model of the energy consumption of the $V/2$ and $V/3$ bias schemes is provided. A design guideline for choosing the proper bias scheme is explained in Section III-A. Moreover, the effect of the nonlinearity factor on the choice of bias scheme is explored in Section III-B. The effect of leakage current on the total energy consumption is discussed in Section III-C.

To provide an intuitive closed-form expression that models the energy consumption of a crossbar array, the interconnect resistance is assumed to be negligibly small. Although this
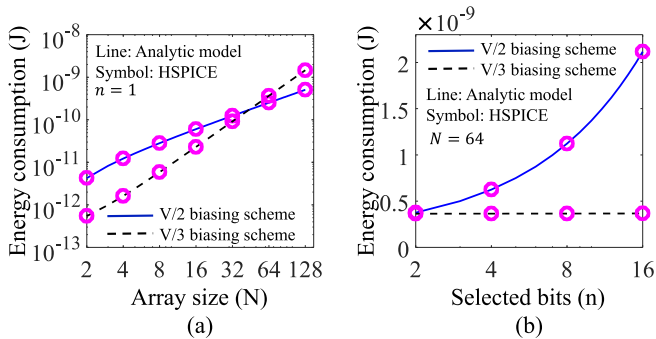
Fig. 2. Energy consumption of a crossbar array with respect to (a) array size and (b) number of selected cells, assuming $R_{ON} = 10^4$ Ω, $R_{OFF} = 10^7$ Ω, $K_{V/2} = 20$, $K_{V/3} = 1000$, and $V_{write} = 2$ V.
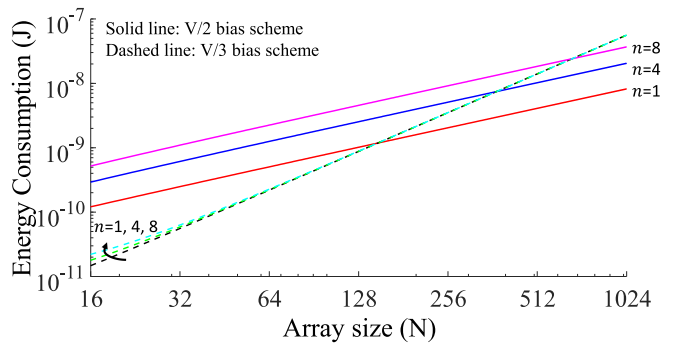

Fig. 3. Effect of the number of selected cells on the energy consumption of a crossbar array for the $V/2$ and $V/3$ bias schemes, assuming $K_{V/2} = 20$ and $K_{V/3} = 1000$.

assumption is not always practical in large arrays, it permits the effects of the critical parameters on the energy consumption, such as the nonlinearity factor, the size of the array, the number of selected cells, and a bias scheme, to be captured while retaining simplicity and providing intuitive expressions. An array with an equal number of rows and columns biased according to the $V/2$ and $V/3$ bias schemes, as shown in Fig. 1, is considered. The selected devices are modeled based on the VTEAM model [26] considering linear switching, and the remaining devices are modeled as resistors. The switching devices are considered to be symmetric with equal ON/OFF threshold voltages and equal set/reset times. The switching energy during the set and reset operations is, therefore, the same (see the Appendix). Based on these considerations and assumptions, the energy consumption of a crossbar array for the $V/2$ and $V/3$ bias schemes is, respectively,

$$E_{V/2} = V_{write} \frac{I_{ON}}{K_{V/2}} \frac{(Nn + N - 2n)}{2} t_{sw} + n E_{sw} \quad (3)$$

$$E_{V/3} = V_{write} \frac{I_{ON}}{K_{V/3}} \frac{(N^2 - n)}{3} t_{sw} + n E_{sw} \quad (4)$$

where V_{write} is the write voltage, I_{ON} is the cell current when biased at the write voltage during the ON state, N is the number of rows and columns, n is the number of selected cells, t_{sw} is the switching time, and E_{sw} is the switching energy consumption of the selected device

$$E_{sw} = \frac{V_{write}^2}{R_{OFF} - R_{ON}} \ln\left(\frac{R_{OFF}}{R_{ON}}\right) t_{sw}. \quad (5)$$

R_{ON} and R_{OFF} are, respectively, the 1S1R cell resistance during the ON and OFF states (for more details see the Appendix). The resistance of the selector device due to the limited current density is assumed to be considered in R_{ON} and R_{OFF}. Note that the second term in (3) and (4) is the dynamic portion of the total energy due to switching the selected cells, while the first term is due to the leakage current of the half-selected and unselected cells.

The closed-form expressions are in good agreement with SPICE, exhibiting an average error of 0.04% and a maximum error of 0.74%, as shown in Fig. 2. The energy consumption scales differently with respect to array size for different bias schemes. The $V/2$ bias scheme follows a linear trend,

whereas the $V/3$ bias scheme scales superlinearly with array size ($\sim N^2$). Moreover, while the energy consumption for the $V/2$ bias scheme is strongly dependent on the number of selected cells, the $V/3$ bias scheme is constant for large arrays ($N \gg n$). Note that $E_{V/3}$ quadratically scales with N, exhibiting a near constant profile with respect to n if $N \gg n$. Under this condition, while the switching energy continues to grow with large n since the leakage energy dominates for large array sizes, E_{sw} remains insignificant. The effect of n on the energy consumption for different array sizes is shown in Fig. 3. A summary of the parameters used in the following simulations is provided in Table I (unless otherwise noted).

The increasing number of selected bits per write operation significantly adds to the energy consumption of the $V/2$ bias scheme. The $V/3$ bias scheme remains relatively constant for large array sizes. This behavior is due to the increasing number of half-selected cells for the $V/2$ bias scheme with increasing n. In contrast, for the $V/3$ bias scheme, the variation in the number of unselected cells becomes negligible as n increases if the size of the array N is much larger than n.

One method to decrease the energy consumption is by using selectors with a higher nonlinearity factor. A higher nonlinearity factor decreases the leakage current of the unselected cells, improving the ability of the selector to isolate the switching cell from the rest of the unselected array. The effect of the nonlinearity factor on the energy consumption is shown in Fig. 4. Note that with increasing nonlinearity factor, the energy consumed during both bias schemes decreases, since (3) and (4) are, respectively, inversely proportional to $K_{V/2}$ and $K_{V/3}$.

A. Energy-Efficient Bias Scheme

Depending upon the array size, one bias scheme is more efficient than the other bias scheme. The number of selected
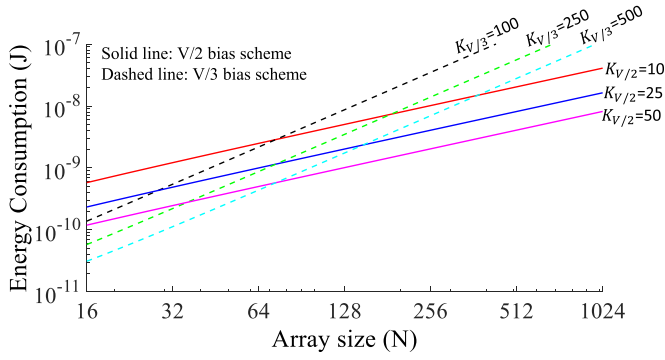
TABLE I
SUMMARY OF PARAMETERS FOR WRITE OPERATION

Parameters	Values
R_{on}	10^4 Ω
R_{off}	10^7 Ω
t_{sw}	100 ns
V_{write}	4 V

Fig. 4. Effect of the nonlinearity factor on the energy consumption of a crossbar array for the $V/2$ and $V/3$ bias schemes, assuming $n = 4$.

Fig. 6. Energy savings of the $V/3$ bias scheme as compared with the $V/2$ bias scheme assuming the same parameters listed in Fig. 5. Solid line: contour where the energy consumption between the two bias schemes is equal.
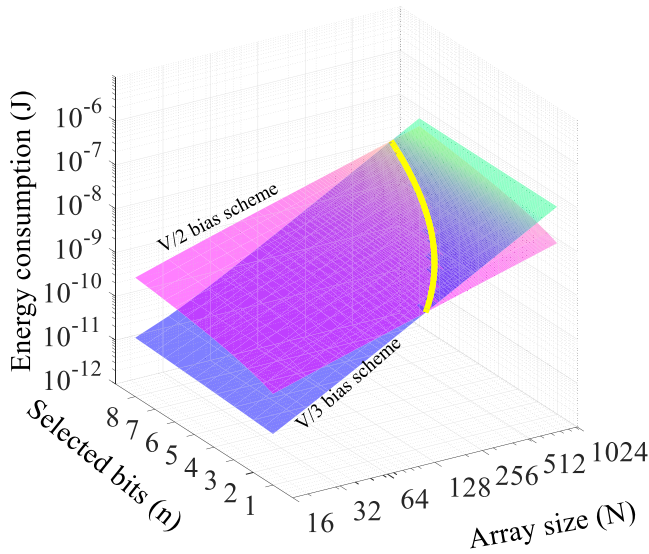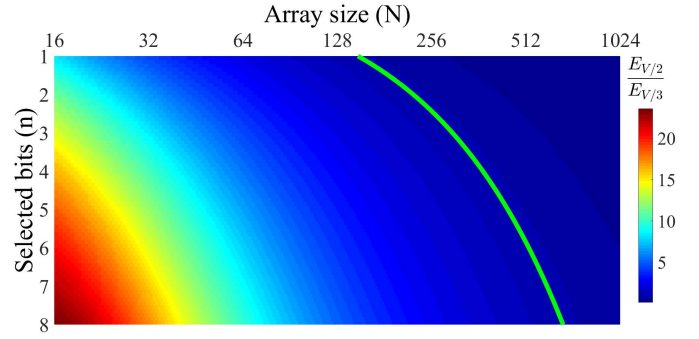

Fig. 5. Comparison of energy consumption in terms of the array size and the number of selected cells for the $V/2$ and $V/3$ bias schemes, assuming $K_{V/2} = 20$ and $K_{V/3} = 1000$.

energy efficient over a wider range of array sizes. The write energy can be as much as five times lower for a 128×128 array and ten times lower for a 64×64 array using the $V/3$ bias scheme with eight selected bits. For large arrays, however, since the number of cells leaking current during the $V/3$ bias scheme scales with N^2, the $V/2$ bias scheme can consume as much as seven times lower energy for an array size of 1024×1024 with single bit operation.

The interconnect resistance changes the location of the contour (see Fig. 6), where the energy for both bias schemes is equal. Since the leakage current due to the half-selected cells for the $V/2$ bias scheme is significantly greater than the leakage current of the cells biased at one third of the write voltage, the IR voltage drops are greater for the $V/2$ bias scheme [17]. Thus, the voltage drop across the selected cells for the $V/2$ bias scheme is smaller than the $V/3$ bias scheme. The switching time of the selected cells for the $V/2$ bias scheme is, therefore, longer, increasing the energy consumption [26] and resulting in the $V/3$ bias scheme being more energy efficient. This effect is more pronounced with larger IR voltage drops, resulting in slower switching times.

B. Impact of Nonlinearity Factor

The bias scheme affects the total leakage current due to the difference between the nonlinearity factors and the number of leaking cells. While the size of the array as well as the number of selected bits affect the choice of energy-efficient bias scheme, the difference between the nonlinearity factors ($K_{V/2}$ and $K_{V/3}$) determines the range of N and n at which the two energy consumptions, $E_{V/2}$ and $E_{V/3}$, are equal. For instance, if one nonlinearity factor is much greater than the other nonlinearity factor, the bias scheme that provides the higher nonlinearity factor will be the most energy-efficient bias scheme for a wide range of N and n. The ratio of the two nonlinearity factors, $K_{V/2}$ and $K_{V/3}$, is a function of the array size and the number of selected cells. Based on this ratio, for the $V/3$ bias scheme to be more energy efficient than the $V/2$ bias scheme, the following condition must be satisfied:

$$\frac{K_{V/3}}{K_{V/2}} \geqslant \frac{2}{3} \frac{N^2 - n}{Nn + N - 2n}. \tag{6}$$

cells n during a write operation may, however, alter the most energy-efficient bias scheme, as shown in Fig. 5. Note that the line of intersection between the two bias schemes (where $E_{V/2} = E_{V/3}$) spans a range of array sizes ($N = 128, 256,$ and 512) depending upon the number of selected bits. Since the $V/2$ bias scheme scales with the number of selected cells as opposed to the $V/3$ bias scheme which remains relatively constant, the line of intersection bends for different values of n.

Extra energy is expended due to an incorrect choice of bias scheme, wasting significant power during a write operation. The ratio of the energy consumption between the two bias schemes is shown in Fig. 6. The right hand side of the contour is the region where the $V/2$ bias scheme is more efficient than the $V/3$ bias scheme, and the left hand side is where the $V/3$ bias scheme is more efficient than the $V/2$ bias scheme. Since increasing the number of selected cells consumes more energy for the $V/2$ bias scheme for low n, the $V/2$ bias scheme remains more energy efficient over a wider range of array sizes. In contrast, for high n, the $V/3$ bias scheme is more
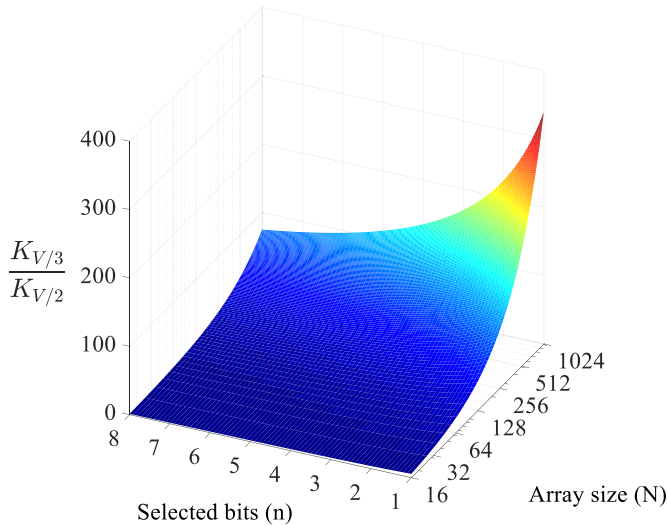
Fig. 7. Ratio of the nonlinearity factors $K_{V/3}$ to $K_{V/2}$ to maintain equal energy consumption for the $V/2$ and $V/3$ bias schemes in terms of the array size and the number of selected cells.
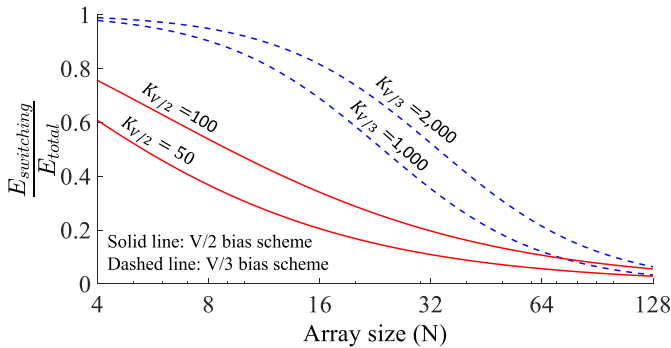

Fig. 8. Ratio of the switching energy to the total energy in terms of the array size, $R_{ON} = 10^4 \, \Omega$, $R_{OFF} = 10^6 \, \Omega$, and $n = 4$.

Note that for negligible parasitic interconnect resistance, (6) is a function of the size of the array and the number of selected cells. The variation of $K_{V/3}$ to satisfy (6) is shown in Fig. 7. The $V/3$ bias scheme is more energy efficient if $K_{V/3}$ is at least two orders of magnitude greater than $K_{V/2}$ for array sizes up to 1024×1024 with six selected bits or an array size up to 256×256 with a single selected bit.

C. Write Pulsewidth

The pulsewidth to successfully program the selected cells depends upon the switching time of the cells. While shorter pulses may produce write failures, extended pulse widths may consume excessive power, degrading the energy efficiency. Due to the significance of the leakage current of the unselected cells, it is crucial to accurately set the pulsewidth with high precision. For large arrays, the leakage current portion of the total energy dominates, making the switching energy E_{sw} negligible, as shown in Fig. 8.

Note that the switching energy for the $V/3$ bias scheme is a larger portion of the total energy as compared with the $V/2$ bias scheme. This difference is due to the smaller leakage
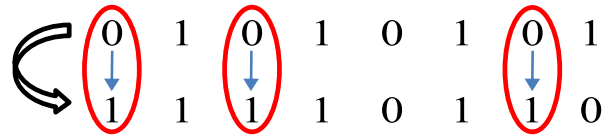

Fig. 9. Writing an eight-bit word. Four bits of the new string are the same as the old string; however, only three bits are selected, since one bit requires a reset, whereas the other three bits require a set operation.

current for the $V/3$ bias scheme due to the larger nonlinearity factor, $K_{V/3}$. Similarly, a higher nonlinearity factor reduces the leakage energy, resulting in the switching energy being more pronounced and exhibiting greater energy efficiency. The switching energy is less than 10% of the total energy for array sizes exceeding $N = 128$.

To lower the energy due to leakage currents, the pulsewidth is set as precisely as possible, sufficient to switch the selected cells. This excess energy due to leakage currents requires write termination circuitry to isolate the write voltage from the array once successful switching is achieved. While write termination techniques have been adopted for resistive cells based on STT-MRAM due to the stochastic nature of the switching process [27], a similar approach in RRAM-based 1S1R crossbar arrays can be useful to save energy, since an overextended write pulse can significantly reduce the energy efficiency due to the large leakage currents. The write termination circuitry exhibits a negligible energy overhead of, on average, less than 100 fJ [27].

IV. ENERGY-EFFICIENT HYBRID WRITE SCHEME

In this section, a write scheme is proposed to improve the energy efficiency of a crossbar array during write operations. The optimal choice of the energy-efficient bias scheme is explained in Section IV-A. The overhead and challenges of the proposed system are discussed in Section IV-B.

The number of selected cells affects the energy of an array and can be used to determine the most energy-efficient bias scheme. The proposed write scheme improves the energy efficiency by adaptively switching between the $V/2$ and $V/3$ bias schemes depending upon the number of selected cells during a write operation. The number of selected bits during a write operation depends upon the difference between the patterns of the old data and the new data, as shown in Fig. 9. Consider a word size of eight bits. If the new data are the same as the old data, the number of selected cells is equal to zero. If, however, the new data are different than the previous data, the number of selected cells separately depends upon the number of sets and resets, since in resistive memories, writing a 1 or a 0 requires two different write operations. To determine the number of bits, a read-before-write technique is typically used [28]. This approach detects those cells that require switching, reducing excessive energy consumption during a write operation. By adopting a similar approach to monitor the number of selected cells during each write operation, the optimal energy-efficient bias scheme can be determined.
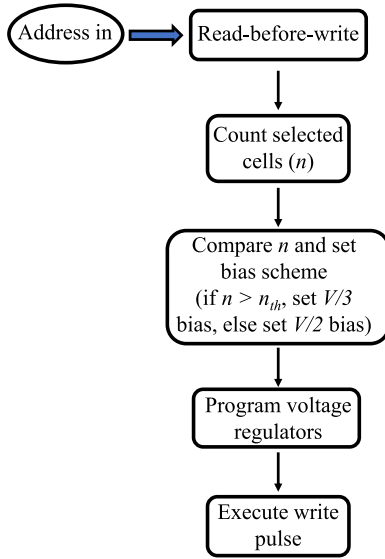
Fig. 10.   Steps during the proposed energy-efficient write scheme.

Fig. 11.   Energy improvement in terms of the number of selected cells, assuming $N = 128$, $K_{V/2} = 20$, and $K_{V/3} = 345$. The proposed write operation chooses the most energy-efficient bias scheme based on the number of selected cells n with respect to n_{th}.

The steps summarizing the write process using the energy-efficient write scheme are shown in Fig. 10. The initial step is a read-before-write operation followed by counting the number of cells that will switch for the new string of data. Once the number of selected cells n is known, n is compared with n_{th} (see Section IV-A) for a specific array. Following this step, the power delivery system is configured to support either the $V/2$ or $V/3$ bias scheme to lower the energy. During this step, the crossbar array remains idle, and no energy is, therefore, consumed. Finally, once the regulator voltage converges to the appropriate bias scheme, the write pulse is executed to write the new data and complete the write process.

A. Optimal Choice of Bias Scheme

The bias scheme of a crossbar array is altered when the number of selected cells n crosses a threshold, n_{th}. At this threshold, the write energy of the $V/2$ and $V/3$ bias schemes is equal. Since the energy for the $V/2$ bias scheme grows with increasing n, if $n < n_{\text{th}}$, the power delivery system switches to the $V/2$ bias scheme. If $n > n_{\text{th}}$, the power delivery system switches to the $V/3$ bias scheme. The energy savings in terms of the number of selected cells n are shown in Fig. 11. Note that if n_{th} is four, the $V/2$ bias scheme provides as much as a $2.5\times$ energy improvement for a 128×128 array when a single bit is selected. The $V/3$ bias scheme provides up to a $1.8\times$ savings in energy when eight bits are selected. Note, however, the size of the array N as well as the ratio of the nonlinearity factor K_r can affect the most energy-efficient bias scheme. Depending upon N and K_r, n_{th} may reside outside the range of the allowed values of n.

The effect of N and K_r on the energy savings is shown in Fig. 12. Note that the hybrid bias scheme only benefits specific array sizes for a fixed value of K_r. For instance, according to Fig. 12(a)–(c), the hybrid bias scheme can be used for an array size of, respectively, 512×512 or 256×256, 128×128 (the same as shown in Fig. 11), and 64×64. The
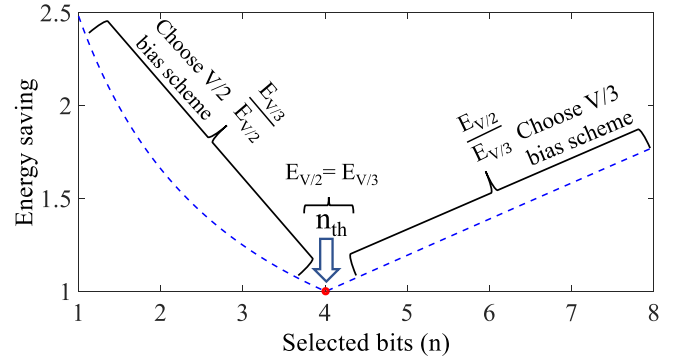
curve along the N and n axes spans the regions where no energy savings exist (i.e., unity). If the array size is above this curve, the bias scheme is set to $V/2$. If below this curve, the bias scheme is set to $V/3$. If the array size is neither above nor below this curve, the hybrid bias scheme can be used to improve the energy efficiency.

By setting the energy for both bias schemes, (3) and (4), equal, the number of bits in which both bias schemes consume the same energy n_{th} can be determined. Based on this equality, n_{th} is

$$n_{\text{th}} = \frac{2N^2 - 3K_r N}{3K_r N - 6K_r + 2} \qquad (7)$$

where K_r is the ratio of the nonlinearity factors

$$K_r = \frac{K_{V/3}}{K_{V/2}}. \qquad (8)$$

Note that n_{th} is a function of K_r and the array size N when the interconnect resistance is negligible. The change of n_{th} as a function of K_r and the array size N is shown in Fig. 13. For large arrays with low K_r, n_{th} increases significantly, reaching 16. This effect is due to the diminishing savings in energy of the $V/3$ bias scheme with increasing array size, resulting in a large number of unselected cells leaking current, which scales with N^2. Furthermore, a lower K_r means the difference in leakage current between the half-selected cells for both bias schemes decreases. Thus, the $V/2$ bias scheme is more energy efficient for a wider number of selected cells. Since the leakage current of the unselected cells for the $V/3$ bias scheme decreases relative to the leakage current of the $V/2$ bias scheme, as K_r increases, the $V/3$ bias scheme becomes more energy efficient for a wider number of selected cells, hence decreasing n_{th}. In addition, if the interconnect resistance incurs significant IR voltage loss, n_{th} decreases, since the switching time for the $V/2$ bias scheme is larger than that for the $V/3$ bias scheme due to the voltage degradation across the selected cells [17]. Increasing K_r from a few tens to 100 can reduce n_{th} from 16 to six. If n_{th} is larger or equal to the maximum number of selected cells (i.e., word size), the array is biased with only the $V/2$ bias scheme rather than the hybrid bias scheme (see Fig. 11). The nonlinearity factor of a 1S1R
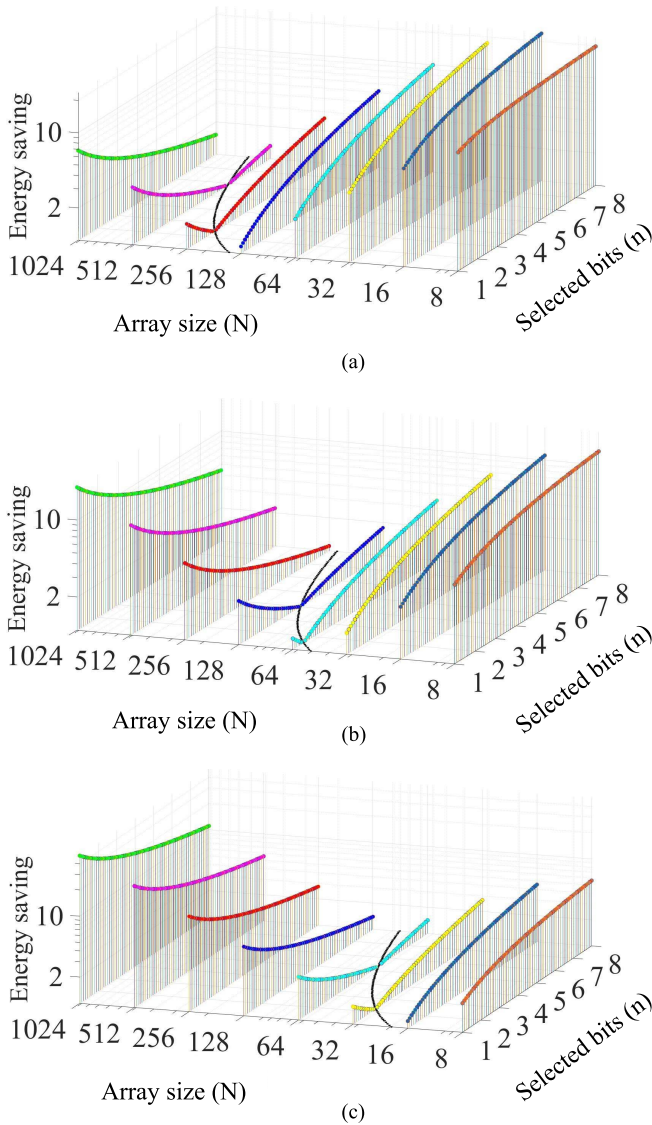
Fig. 12. Energy savings for different array sizes and the number of selected cells considering (a) $K_r = 1000/20$, (b) $K_r = 345/20$, and (c) $K_r = 345/50$.

Fig. 13. Number of selected cell in which the energy for both bias schemes is equal with respect to K_r and the array size N.

cell for the $V/2$ bias scheme is typically less than 100, whereas the nonlinearity factor for the $V/3$ bias scheme reaches a few thousands. K_r is typically in the range of a few tens to several hundreds.

B. Overhead and Challenges

While the $V/2$ bias scheme requires two voltages, V_{write} and $V_{write}/2$, the $V_{write}/3$ bias scheme requires three voltages, namely, V_{write}, $V_{write}/3$, and $2V_{write}/3$. A hybrid solution using both bias schemes requires four voltage levels. Providing a large number of heterogeneous on-chip voltages is challenging due to the limited board area for the off-chip power supplies and the limited number of power I/Os. In [29], a boost converter with a charge pump is used to bias the array. This approach is, however, not feasible for a hybrid bias scheme with multiple voltage levels, since the switching converter requires large off-chip inductors as well as large capacitors, greatly increasing the area and, therefore, the cost [30].
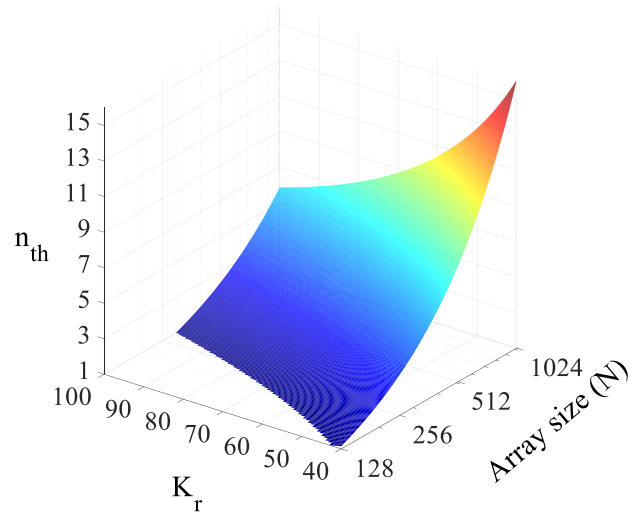
Linear regulators, alternatively, are less power efficient as opposed to switching converters; however, linear converters are much smaller, since bulky capacitors or inductors are not required [31]. Heterogeneous power delivery systems with a large number of voltages using on-chip linear regulators have been proposed [32], [33]. These on-chip voltage regulators can be placed close to the load, further reducing the response time while providing fast local power management to control the bias scheme (as opposed to an off-chip power management solution which exhibits higher latency) [34]. By programming the reference voltage of the on-chip regulators, the bias scheme can be altered between $V/2$ and $V/3$ [34]–[36].

The proposed energy-efficient write scheme provides energy savings as high as $2.5\times$ as compared with a conventional system with a single bias scheme. The write process, however, incurs additional steps as compared with a conventional write operation with a constant bias scheme (see Fig. 10), increasing the write latency. The write latency is typically the switching time of the 1S1R cell. In the proposed hybrid write scheme, however, the read-before-write operation necessitates a read operation for every write operation. The time required to compute and compare n with respect to n_{th} has to be considered in addition to the switching time of the 1S1R cell.

In memory systems, the read operation is typically a primary performance bottleneck. If, however, the write latency increases significantly, it can inhibit memory performance. Thus, a fast power delivery system is required for time constrained memory applications, such as DRAM and cache memory. For slower memory systems, such as flash, the stringent timing requirements can be relaxed. While the read latency is significantly smaller than the write latency [6] and can be as low as 5 ns [7], the time required to program the voltage regulators has to be within a few nanoseconds to prevent write dependent performance limitations. Hence, the need for an on-chip voltage regulator (as opposed to an off-chip regulator) becomes necessary, since, unlike on-chip local regulation, off-chip power management and regulation cannot provide sub-μs bandwidth [34].

The energy overhead of the energy-efficient write scheme is insignificant. The write operation for a 1S1R crossbar array is typically on the order of hundreds of nanojoules [6]. The read operation during the read-before-write requires negligible energy, typically less than 1 nJ, since the read latency is significantly less than the write latency. The programmable CMOS reference voltage consumes a few picojoules [36], assuming a switching time on the order of hundreds of nanoseconds. The primary challenge for the proposed write scheme is lowering the overhead of the write latency in time constrained memory applications.

V. Conclusion

The energy consumption of a 1S1R crossbar array for two bias schemes, $V/2$ and $V/3$, for optimal energy efficiency is discussed. Closed-form expressions that intuitively model the energy consumption in terms of the nonlinearity factor, size of the array, and number of selected cells are presented. The most energy-efficient bias scheme depends upon the size of the array as well as the number of selected cells during a write operation. The energy consumed during both bias schemes scales differently. The $V/2$ bias scheme is more energy efficient for large arrays. As the number of selected cells increases, however, the $V/3$ bias scheme achieves greater energy efficiency. The $V/3$ bias scheme provides higher efficiency, decreasing the energy consumption by an order of magnitude for a 64×64 array with eight selected cells. As the array size increases and the number of selected cells decreases, the energy benefits of the $V/3$ bias scheme diminish.

For the $V/3$ bias scheme to be as energy efficient as the $V/2$ bias scheme for large arrays ($N > 128$), $K_{V/3}$ should be two orders of magnitude greater than $K_{V/2}$. The appropriate choice of bias scheme can save an order of magnitude of energy. A higher nonlinearity factor significantly decreases the energy consumption by suppressing leakage currents within the half-selected and unselected cells. The switching energy is a negligible portion of the total energy for large arrays ($N > 128$). To prevent excess energy consumption due to leakage currents, write termination circuitry can be used to prevent overextended write pulses.

To improve the energy efficiency during write operations, an energy-efficient write scheme is proposed. The write operation uses a hybrid bias scheme to exploit both the $V/2$ and $V/3$ bias schemes to enhance the energy efficiency based on the number of selected cells. The critical number of selected cells in which the bias scheme switches (n_{th}) is characterized. Energy improvements provided by the hybrid write scheme can be as high as $2.5\times$. To effectively exploit the energy efficient write scheme in time constrained memory systems, the program time of the voltage regulators and the time to compute n_{th} need to be on the order of a few nanoseconds. The proposed write scheme incurs negligible energy overhead. Future work will focus on integrating the interconnect resistance into the energy models to capture the effects of IR voltage drops on the switching time of the selected cells and the energy consumption of the crossbar array.

Appendix

To estimate the switching energy of a resistive cell, the resistance is modeled as a linear function during the switching interval. The resistance during a set operation is

$$R(t) = R_{\mathrm{OFF}} + \frac{t}{t_{\mathrm{set}}}(R_{\mathrm{ON}} - R_{\mathrm{OFF}}) \tag{9}$$

assuming the set operation is initiated between $t = 0$ and $t = t_{\mathrm{set}}$. If the interconnect resistance is negligible, the voltage across the cell is equal to the write voltage V_{write}. The power consumption is

$$P_{\mathrm{set}}(t) = \frac{V_{\mathrm{write}}^2}{R(t)}. \tag{10}$$

Integrating (10) over the set period, the energy consumption is

$$E_{\mathrm{set}} = \int_{t=0}^{t_{\mathrm{set}}} P_{\mathrm{set}}(t)dt = \frac{V_{\mathrm{write}}^2}{R_{\mathrm{ON}} - R_{\mathrm{OFF}}} \ln\left(\frac{R_{\mathrm{ON}}}{R_{\mathrm{OFF}}}\right) t_{\mathrm{set}}. \tag{11}$$

For a symmetric resistive cell where the set and reset voltages as well as switching times are equal, the set and reset energy consumption is also the same.

References

[1] S. Yu and P. Y. Chen, "Emerging memory technologies: Recent trends and prospects," *IEEE Solid State Circuits Mag.*, vol. 8, no. 2, pp. 43–56, Spring 2016.

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

[3] H.-S. P. Wong *et al.*, "Metal–oxide RRAM," *Proc. IEEE*, vol. 100, no. 6, pp. 1951–1970, Jun. 2012.

[4] H.-S. P. Wong *et al.*, "Phase change memory," *Proc. IEEE*, vol. 98, no. 12, pp. 2201–2227, Dec. 2010.

[5] S. Kvatinsky, G. Satat, N. Wald, E. G. Friedman, A. Kolodny, and U. C. Weiser, "Memristor-based material implication (IMPLY) logic: Design principles and methodologies," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 22, no. 10, pp. 2054–2066, Oct. 2014.

[6] X. Dong, C. Xu, Y. Xie, and N. P. Jouppi, "NVSim: A circuit-level performance, energy, and area model for emerging nonvolatile memory," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 31, no. 7, pp. 994–1007, Jul. 2012.

[7] T. Na, B. Song, J. P. Kim, S. H. Kang, and S.-O. Jung, "Offset-canceling current-sampling sense amplifier for resistive nonvolatile memory in 65 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 52, no. 2, pp. 496–504, Feb. 2017.

[8] Y.-C. Chen *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 *IEDM Tech. Dig.*, Dec. 2003, pp. 37.4.1–37.4.4.

[9] Y. Chen *et al.*, "Nanoscale molecular-switch crossbar circuits," *Nanotechnology*, vol. 14, no. 4, pp. 462–468, Mar. 2003.

[10] C. Xu *et al.*, "Overcoming the challenges of crossbar resistive memory architectures," in *Proc. IEEE Int. Symp. High Perform. Comput. Archit.*, Feb. 2015, pp. 476–488.

[11] S. Yu, *Resistive Random Access Memory (RRAM): From Devices to Array Architectures*. San Rafael, CA, USA, Morgan & Claypool, 2016.

[12] A. Chen, "A review of emerging non-volatile memory (NVM) technologies and applications," *Solid-State Electron.*, vol. 125, pp. 25–38, Nov. 2016.

[13] R. Aluguri and T.-Y. Tseng, "Overview of selector devices for 3-D stackable cross point RRAM arrays," *IEEE J. Electron Devices Soc.*, vol. 4, no. 5, pp. 294–306, Sep. 2016.

[14] B. Song, Q. Li, H. Liu, and H. Liu, "Exploration of selector characteristic based on electron tunneling for RRAM array application," *IEICE Electron. Exp.*, vol. 14, no. 17, pp. 1–8, Aug. 2017.

[15] A. Ciprut and E. G. Friedman, "On the write energy of non-volatile resistive crossbar arrays with selectors," in *Proc. IEEE Int. Symp. Quality Electron. Design*, Mar. 2018, in press.

[16] J.-J. Huang, Y.-M. Tseng, C.-W. Hsu, and T.-H. Hou, "Bipolar nonlinear Ni/TiO$_2$/Ni selector for 1S1R crossbar array applications," *IEEE Electron Device Lett.*, vol. 32, no. 10, pp. 1427–1429, Oct. 2011.

[17] A. Ciprut and E. G. Friedman, "Modeling size limitations of resistive crossbar array with cell selectors," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 1, pp. 286–293, Jan. 2017.

[18] W. Lee *et al.*, "High current density and nonlinearity combination of selection device based on TaO$_x$/TiO$_2$/TaO$_x$ structure for one selector–one resistor arrays," *ACS Nano*, vol. 6, no. 9, pp. 8166–8172, Aug. 2012.

[19] J.-J. Huang, Y.-M. Tseng, W.-C. Luo, C.-W. Hsu, and T.-H. Hou, "One selector-one resistor (1S1R) crossbar array for high-density flexible memory applications," in *IEDM Tech. Dig.*, Dec. 2011, pp. 31.7.1–31.7.4.

[20] Q. Luo *et al.*, "Demonstration of 3D vertical RRAM with ultra low-leakage, high-selectivity and self-compliance memory cells," in *IEDM Tech. Dig.*, Dec. 2015, pp. 10.2.1–10.2.4.

[21] M. Wang, J. Zhou, Y. Yang, S. Gaba, M. Liu, and W. D. Lu, "Conduction mechanism of a TaO$_x$-based selector and its application in crossbar memory arrays," *Nanoscale*, vol. 7, no. 11, pp. 4964–4970, Feb. 2015.

[22] B. Choi *et al.*, "Trilayer tunnel selectors for memristor memory cells," *Adv. Mater.*, vol. 28, no. 2, pp. 356–362, Jan. 2016.

[23] C.-Y. Lin *et al.*, "Attaining resistive switching characteristics and selector properties by varying forming polarities in a single HfO$_2$-based RRAM device with a vanadium electrode," *Nanoscale*, vol. 9, pp. 8586–8590, May 2017.

[24] Q. Luo *et al.*, "Super non-linear RRAM with ultra-low power for 3D vertical nano-crossbar arrays," *Nanoscale*, vol. 8, pp. 15629–15636, Aug. 2016.

[25] R. Midya *et al.*, "Anatomy of Ag/hafnia-based selectors with 1010 nonlinearity," *Adv. Mater.*, vol. 29, no. 12, p. 1604457, Jan. 2017.

[26] S. Kvatinsky, M. Ramadan, E. G. Friedman, and A. Kolodny, "VTEAM: A general model for voltage-controlled memristors," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 62, no. 8, pp. 786–790, Aug. 2015.

[27] H. Farkhani, M. Tohidi, A. Peiravi, J. K. Madsen, and F. Moradi, "STT-RAM energy reduction using self-referenced differential write termination technique," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 2, pp. 476–487, Feb. 2017.

[28] H. L. Lung, "Method, apparatus and computer program product for stepped reset programming process on programmable resistive memory cell," U.S. Patent 7 440 315 B2, Oct. 21, 2008.

[29] T. Ishii, S. Ning, M. Tanaka, K. Tsurumi, and K. Takeuchi, "Adaptive comparator bias-current control of 0.6 V input boost converter for ReRAM program voltages in low power embedded applications," *IEEE J. Solid-State Circuits*, vol. 51, no. 10, pp. 2389–2397, Oct. 2016.

[30] G. Villar-Piqué, H. J. Bergveld, and E. Alarcón, "Survey and benchmark of fully integrated switching power converters: Switched-Capacitor versus inductive approach," *IEEE Trans. Power Electron.*, vol. 28, no. 9, pp. 4156–4167, Sep. 2013.

[31] J. Torres *et al.*, "Low drop-out voltage regulators: Capacitor-less architecture comparison," *IEEE Circuits Syst. Mag.*, vol. 14, no. 2, pp. 6–26, May 2014.

[32] I. Vaisband and E. G. Friedman, "Heterogeneous methodology for energy efficient distribution of on-chip power supplies," *IEEE Trans. Power Electron.*, vol. 28, no. 9, pp. 4267–4280, Sep. 2013.

[33] S. Kose and E. G. Friedman, "Distributed on-chip power delivery," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 2, no. 4, pp. 704–713, Dec. 2012.

[34] Z. Toprak-Deniz *et al.*, "Distributed system of digitally controlled microregulators enabling per-core DVFS for the POWER8 microprocessor," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2014, pp. 98–99.

[35] Y.-C. Wu, C.-Y. Huang, and B.-D. Liu, "A low dropout voltage regulator with programmable output," in *Proc. IEEE Conf. Ind. Electron. Appl.*, May 2009, pp. 3357–3361.

[36] V. Srinivasan, G. Serrano, C. M. Twigg, and P. Hasler, "A floating-gate-based programmable CMOS reference," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 11, pp. 3448–3456, Dec. 2008.

Albert Ciprut (S'15) received the B.S. degree in electronics engineering from Sabanci University, Istanbul, Turkey, in 2013 and the M.S. degree in electrical and computer engineering from the University of Rochester, Rochester, NY, USA, in 2016, where he is currently working toward the Ph.D. degree in electrical engineering under the supervision of Prof. E. G. Friedman.

He was an Intern in the Power Team, Google Inc., Mountain View, CA, USA, in 2016. His current research interests include memory systems and integrated circuit design based on emerging memory technologies as well as on-chip power delivery systems.

Eby G. Friedman (F'00) received the B.S. degree from Lafayette College, Easton, PA, USA, in 1979 and the M.S. and Ph.D. degrees from the University of California at Irvine, Irvine, CA, USA, in 1981 and 1989, respectively, all in electrical engineering.

He was with Hughes Aircraft Company, Westchester, CA,USA, from 1979 to 1991, rising to Manager of the Signal Processing Design and Test Department, where he was responsible for the design and test of high performance digital and analog ICs. He has been with the Department of Electrical and Computer Engineering, University of Rochester, Rochester, NY, USA, since 1991, where he is currently a Distinguished Professor and the Director of the High Performance VLSI/IC Design and Analysis Laboratory. He is also a Visiting Professor with the Technion–Israel Institute of Technology, Haifa, Israel. He has authored over 500 papers and book chapters. He holds 15 patents. He is the author or editor of 18 books in the fields of high-speed and low-power CMOS design techniques, 3-D design methodologies, high-speed interconnect, and the theory and application of synchronous clock and power distribution networks. His current research and teaching interests include high-performance synchronous digital and mixed-signal microelectronic design and analysis with application to high-speed portable processors, low-power wireless communications, and server farms.

Dr. Friedman is a Senior Fulbright Fellow. He was a member of the Circuits and Systems Society Board of Governors, and is currently a member of the Technical Program Committee of numerous conferences. He was a recipient of the IEEE Circuits and Systems Charles A. Desoer Technical Achievement Award, a University of Rochester Graduate Teaching Award, and a College of Engineering Teaching Excellence Award. He was the program and technical chair of several IEEE conferences. He was the Editor-in-Chief and Chair of the Steering Committee of the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION SYSTEMS, the Regional Editor of the *Journal of Circuits, Systems and Computers*, a member of the Editorial Board of the PROCEEDINGS OF THE IEEE, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II: ANALOG AND DIGITAL SIGNAL PROCESSING, *Analog Integrated Circuits and Signal Processing*, the IEEE JOURNAL ON EMERGING AND SELECTED TOPICS IN CIRCUITS AND SYSTEMS, and *Journal of Signal Processing Systems*. He is the Editor-in-Chief of the *Microelectronics Journal* and a member of the Editorial Board of the *Journal of Low Power Electronics* and *Journal of Low Power Electronics and Applications*.