Supervised Learning with Organic Memristor Devices and Prospects for Neural Crossbar Arrays

Christopher H. Bennett*, Djaafar Chabi*, Theo Cabaret†, Bruno Jousselme†, Vincent Derycke†,
Damien Querlioz*, Jacques-Olivier Klein*
*Institut d'Electronique Fondamentale, Univ. Paris-Sud, CNRS, 91405 Orsay, France
Email: christopher.bennett@u-psud.fr
†CEA Saclay, IRAMIS, NIMBE (UMR 3685), LICSEN, 91191 Gif-sur-Yvette, France.

***Abstract*—The integration of memristive nanodevices within transistor-based electronic systems offers the potential for computing structures smaller, lower power and cheaper than traditional high-performance systems. Among emerging memristive technologies, a novel device based on organic materials distinguishes itself, in that it can feature several threshold voltages on the same die, and possesses unipolar behavior. In this work, we highlight that these two features can be beneficial for neural network-inspired learning systems. An on-chip supervised learning method for hybrid memristors / CMOS systems – an analogue synaptic array paired with a hybrid learning cell – is extended to the case of this novel organic memristor device. The organic device can be trained with only one pulse per row (two for the entire array) per presentation of input- as compared to four for a bipolar memristor array. The device also works universally- in both the synaptic grid as well as learning cell- paving the way to single die integration. The proposed scheme learns successfully, even while incorporating non-ideal circuit phenomena such as a wide range of parasitic wire resistances and associated sneak paths. These encouraging first results suggest that these multi-threshold, unipolar organic memristive devices are a useful species for inclusion in adaptive next generation electronic systems.**

***Index Terms*—memristor, organic memristor, neural network, nanoscale crossbar, on-chip learning, supervised learning**

I. Introduction

Memristor arrays may enable new power-saving frontiers in computing, as they require zero power in standby mode. Moreover, since memristor elements are reminiscent of biological synapses, a natural opportunity exists for them to store the synaptic weights of neural network structures at ultrahigh density [1], [2]. The transfer of data between physically separate memory and computing cells could be obviated in such a unified, reconfigurable structure. Yet wiring schemes at the nano-scale come with trade-offs, in particular stateful access issues and sneak paths [3]. These issues highlight the importance of hybrid approaches that combine memristor arrays with partner CMOS access device cells or layers [4].

These arrays have the capacity to be trained rather than directly programmed, using learning rules. Several technological options have been considered in this context. The present paper focuses on a recently developed organic memristor device. Previously considered as nonvolatile memories, organic switching media have recently come under consideration as candidates in adaptive circuits [5], [6]. In this context, their strengths include rich analog behavior with multi-level read/ write operation, and highly tunable R_{on} / R_{off} ratios and retention times. From the fabrication perspective, the combination of nano-imprint lithography and electro-grafted active polymer layers scales readily below 25 nm feature size at high throughput, low cost, and already achieved 3D integration; when roll-to-roll printed on a flexible substrate, this may offer a path towards adaptive embedded electronics and sensors [7], [8]. We show that our particular organic memristor device- which follows a unipolar, multi-threshold conductance evolution with no required compliance current- can elegantly implement such an adaptive learning system.

Several learning approaches have been considered with memristive devices, including the bioinspired spike-timing dependent plasticity (STDP) [2], [9], [10] and the algorithm of back-propagation from artifical neural networks [11]. Here, we focus on the single-layer perceptron, a canonical supervised neural network, capable of learning linearly-separable logic functions [12]. This allows us to assess the lower level questions relating to using organic memristors as synapses. Additionally, small perceptron blocks trained with the Widrow-Hoff's Delta rule [13] might act as "Neural Logic Blocks" (NLBs) the equivalent of configurable logic blocks (CLBs) in reinvented Field Programmable Gate Arrays (FPGAs) capable of harnessing the benefits of the memristive synaptic array approach [14]. Recent work has compacted the learning cell further with the use of memristive elements in addition to CMOS [15], [16]. It has been shown already on other technologies that successful on-chip learning is possible with the perceptron [17]–[19]. Yet our device- a unipolar, two-threshold model- trains differently than the supervised learning scheme required for past devices. Notably, we find that only one pulse per cycle per row is required; in the case of 2 input logic functions, 4 for an epoch. This compares favorably to the bipolar scheme of 4 (16). A novel scheme using analog instead of binary latches that allows all-organic devices is also introduced; this could improve integration of this system onto a single die. Nevertheless, both cases are considered and compared for completeness. Transient electrical simulations confirm that both organic memristor on-chip schemes successfully learn several linearly separable Boolean functions simultaneously, and are remarkably resistant to sneak paths at a wide range of parasitic wire resistances.

978-1-4673-7849-9/15/$31.00 ©2015 IEEE

II. Nanodevice Characterization and Modeling

A. Experimental Details

The organic memristor studied herein has as its active layer a polymeric film electro-grafted in between metal contacts; an applied current implements memristive behavior as conductive filaments are created or destroyed between the two electrodes above or below given critical thresholds. The electro-chemistry of the polymer used as well as the metal type and work function of the electrodes determine the respective thresholds and conductance evolution graph. Chemically, the switching media is a thin-film polymer of tris-bipyridine iron complexes (TBFe), where memory effects emerge as a dynamic redox system [20], [21]. Each molecular complex $Fe(bpy)_3{}^{2+}$ contains three bi-pyridine ligands surrounding a central iron core and three diazonium functions that allow for the covalent binding between molecular complexes and between the complexes and the electrodes. As described in greater detail in [21], planar junctions were fabricated by depositing robust thin films (height $\approx 10nm$) of these complexes by electro-grafting into the gap between gold electrodes (gap size $\approx 30nm$) on a SiO2 insulating substrate. Altogether, this robust thin film consists of about 6-7 molecules in the vertical direction and 20 molecules along the gap direction. TBFe cations are counterbalanced by PF_{6^-} anions. Upon bias application, the electrochemical characteristics of the grafted molecular layer allows for a dynamic filamentary behavior as conductive paths are formed or destroyed. A schematic of the individual TBFe complexes surrounded by negative anions and their connection through both "azo" bridges and carbon bonds is visible in Figure 1 panel B, while in panel C a schematic evolution of conductive filaments is presented. We hypothesize that the change in conductivity within filaments arises from reversible charge separation between the iron core, the ligands and the azo-bridges. At the bottom most image roughly corresponding to G_{off}, some TBFe complexes connect but none provide a clear path between the electrodes; some parts of the films form conductive domains but none provide a effective path between the electrodes; in the intermediate states, the domains develop; finally, in the top-most state corresponding to G_{on}, many conductive filaments exist between the two electrodes. A high voltage reversibly disrupts these conductive filaments. Note that the precise elucidation of the electrochemical switching mechanism is outside the scope of this study and will be presented separately.

Electrical characterization reveals two distinct voltage thresholds and three operating modes. In the range of $V = 0V$ to $V = 3V$, the device's conductance is stable and does not change (READ mode). Above $V = 3V$ (V_{th1}) and below $V = 5V$ (V_{th2}), the conductance increases as a product of the applied bias (SET/WRITE mode). Finally, above V_{th2}, the device's conductance decreases (RESET/ERASE mode). These thresholds are visible in the panel A of Figure 1; note that this depicts a forming cycle (first application of current, hence the change in conductance within READ. Conductance evolution is symmetrical (the memristor is unipolar), so that

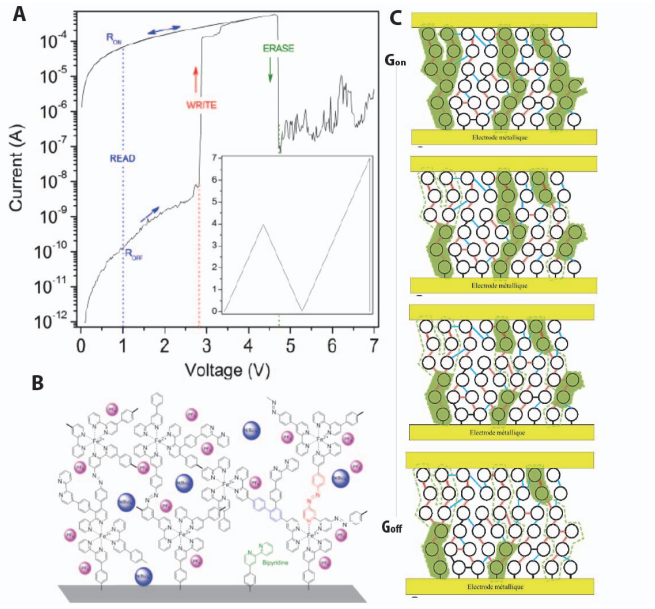


Fig. 1. A) Electrical characterization results for a single TBFe nanodevice in the positive voltage domain. Behavior in the negative voltage domain is symmetrical (flipped across the y-axis). B) Schematic of molecules that comprise the thin-film C) Filamentary development within the thin-film as current is applied between the two electrodes; conductive areas shaded in green. Source: [20]

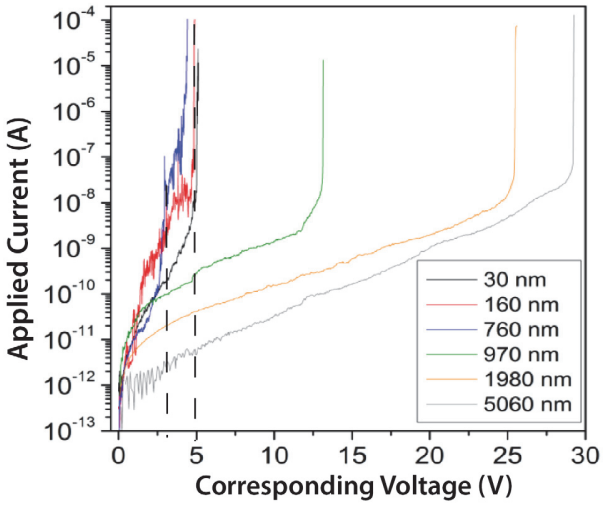


Fig. 2. Current as a function of voltage at the forming stage, measured for different device size (inter-electrode distance L). Source: [20]

negative and positive biases applied in the given ranges have the same effect. This behavior is inverse of most other unipolar devices, where the Erase mode is at lower voltages than the Set. Importantly, our device does not require a compliance current in the 'Set' mode, which is a decisive asset from a circuit point of view. Note that the numerical values of the thresholds just referenced are not singular; tunability exists as a function of the selected molecular compounds, grafted parameters and device geometry as shown for example

for the initial (forming) step in Figure 2. On-chip learning tasks benefit from such a modularity in thresholds. Values of V_{th1}=1.2V and V_{th2}=2.5V were chosen for the analog array memristor; a larger device is used in the learning cell, and its threshold values $V_{\text{th1}} = 3.3V$, $V_{\text{th2}} = 4.8V$, closely correspond to the experimental values from Figure 1 A.

B. Mapping Conductance Evolution to a Compact Model

In a simple bipolar memristor model, conductance does not vary when the device is exposed to a given voltage value in between the thresholds (below the positive threshold V_{th} and yet above $-V_{th}$); else, it increments at a voltage above the positive threshold or decrements at a voltage below the negative threshold. A unipolar two-threshold device such as the organic TBFe species evolves its conductance by a notably different scheme. When a given voltage value applied at time t: $V(t) = V_{\text{M}}$ has an absolute value greater than first threshold V_{th1} and smaller than the second threshold V_{th2} an increment to conductance occurs (WRITE mode); when its absolute value is greater than V_{th2} a decrement occurs (ERASE); and when its absolute value is less than V_{th1}, no conductance change occurs (READ). Mathematically:

$$\frac{dG}{dt} = \begin{cases} V_{th2} > |V_M| > V_{th1} & \alpha(|V_M - V_{th1}|) \\ |V_M| > V_{th2} & -\beta(|V_M - V_{th2}) \\ |V_M| < V_{th1} & 0. \end{cases} \quad (1)$$

Constants $|\alpha|$ and $|\beta|$ are adjusted to match experimental results with the organic memristor; integrating over a small time window, $10^{-8}s$ for instance, is equivalent to a positive conductance change of $4{\times}10^{-8}S$ in the write mode. Following a linear approximation, absolute voltage values at the top of each range produce maximal effect. While a maximal conductance increment voltage value is given at the absolute voltage value just below $|V_{th2}|$, the maximum value for decrement is given at an absolute voltage value just about 0.5V above $|V_{th2}|$, and saturates at this maximum rather than linearly scaling at even higher voltages. Based on these analytical insights and with additional benchmarking for experimental values (G_{off}=150nS, G_{on}= $100\mu S$), a simple model for the unipolar organic memristor was implemented in verilog-A for electrical simulations. This model captures the essential features observed in experiments without requiring intense computations. It is used for all simulations of our unipolar organic memristors in the following sections.

III. Implementation in Neural Crossbar Scheme

A. Learning Principles

Taking the perceptron as a starting point, we wire a crossbar such that each row (neuron) can learn a logically separable function given an arbitrary number of logic inputs (bits). This can be achieved in a neural crossbar scheme, which operates as follows: for n logic inputs, $2n + 2$ physical wires are required, as each input requires a negative and positive wire to separate states along with separate negative and positive bias lines to configure the entire row's weights. A given input

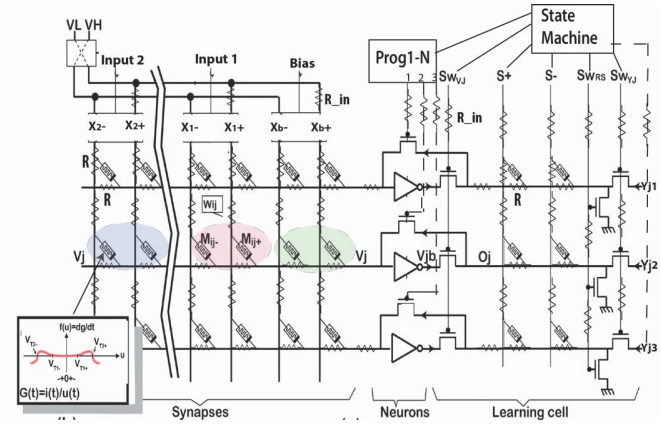


Fig. 3. Schematic for an all-unipolar neural crossbar setup (all constituent memristors follow conductance evolution graph in inset). Each row takes a separate programming pulse per cycle so SW_{prg} is now SW_{PrgJ} for each row. No input polarity I_p. Parasitic resistances are materialized by the R and R_{in} variables. Adapted from [22]

X_i, then, will have two nanowires (X_{i+}, X_{i-}) at whose intersection with row j, two memristors (M_{ij+}, M_{ij-}) will encode a unique synaptic weight pairs as a difference function ($G_{ij+} - G_{ij-}$) for all input/neuron combinations in the entire array, as follows:

$$W_{ij} = K_j(G_{ij+} - G_{ij-}) \quad (2)$$

where K_j is a normalizing factor for that row. This is visualized in Figure 3 as blue, red and green pairs corresponding to Input 2, Input 1, and Bias respectively. Summing the weights for all pairs along the row produces a post-synaptic potential V_j; once inverted at a simple neuron inverter with ground as the threshold, we obtain the final output state O_j which is high (+1) or low (-1):

$$V_{jb} = O_j = sign(V_j) \quad (3)$$

When the output state O_j is the same as the target function Y_j that we wish to converge to, there is no error; but when the sign is different, configuration is required to increase the weights along the row appropriately. The subset of the truth table for all combinations of sign on the input X_i, sign of initial output state V_{jb} = O_j, and sign of target function Y_j in which $Y_j \neq O_j$, produces the four active configuration steps given in the following table:

X_i	Y_j	O_j	ΔW_{ij}	Step
-1	-1	1	1	S1
-1	1	-1	-1	S2
1	-1	1	-1	S3
1	1	-1	1	S4

TABLE I
ACTIVE PROGRAMMING STEPS THAT SOLVE THE DELTA RULE.

As shown above, these four active configurations can implement a Boolean version of Widrow-Hoff's 'Delta' rule:

$$\Delta W_{ij} = \alpha X_i(Y_j - V_j) \quad (4)$$

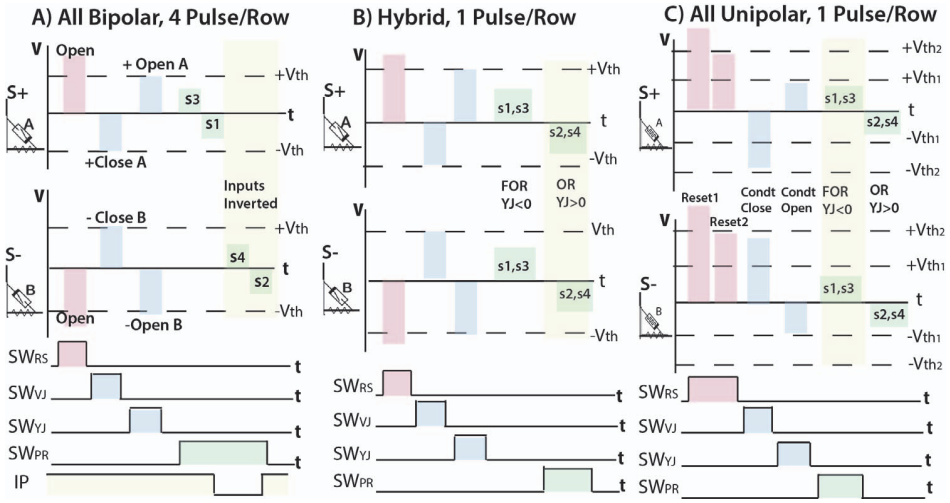


Fig. 4. A): reset, configuration, and programming scheme for bipolar devices. Altogether, 3 gate pulses, one very long gate programming pulse during which 4 individual pulses sent, and a pulse to invert inputs is required. B): hybrid approach (unipolar synapse, bipolar neuron latch) with only 4 smaller gate voltage pulses and one programming pulse required. C): all-organic scheme. It requires only three gate pulses, one programming pulse, but a slightly longer gate voltage and larger magnitude reset pulse.

Explicitly, all successful steps implement negative or positive increments (where $\alpha = 1$) towards the correct configuration. The steps lower conductances on a differential memristor pair if they are too high, or raise them if they are too low. In practice, these configurations are achieved through the application of programming pulses V_{p+} and V_{p-} which are passed from the learning cell back along post-synaptic line.

B. Programming Pulse Schemes

Two memristor latches in combination with a state machine can provide all the stateful logic necessary to implement all active states by sending the appropriate pulses through the signal lines, S_+ and S_-. In the bipolar scheme, these lines control two anti-parallel memristor latches, while the other four lines, SW_{Vj}, SW_{PR}, SW_{RS} and SW_{Yj} control the gate voltages at each of the four control transistors needed to connect or disconnect certain signal lines from each other at various moments in the learning cycle. S_+ and S_- feed into a line of memristor 'latches' $A_1..A_j$, $B_1..B_j$ respectively, where each device is connected to the appropriate output line O_j. By connecting this line to $V_j b$ to get the output state V_j and then target Y_j sequentially, each latch is passively programmed so a programming pulse sent through line S_+ will reach all rows that are low while Y_j is high (S2,S4) and a pulse sent through S_+ will reach the opposite (S1,S3) [16]. This scheme's temporal implementation is visible in Figure 4 Pane A, along with the corresponding operation of the gate voltages. To achieve a proper implementation of the delta rule, this sequence is repeated in cycles of $1\mu s$; depending on the size of the logic function learned, this is 2^n cycles per epoch, or full presentation of the truth table; here, 1 epoch is $4\mu s$. Epochs repeat until the function is learned.

However, the bipolar programming method can not be readily extended to a two-threshold unipolar device. Now, the proper voltage level for programming pulses (V_{p+}, V_{p-}) must sit at either second threshold (V_{th2} or $-V_{th2}$), since slight changes in either direction can induce an increment or decrement. The polarity of the pulse must follow the sign of the expected output for the row or function, Y_j , with the condition that $V_{p+} = -V_{th2}$ and $V_{p-} = +V_{th2}$ (since conductance drops along the second threshold) [22]. Given this, a single pulse fed simultaneously to signal lines S_+ and S_- can then implement learning with only one programming pulse per cycle.

Explicitly, if $Y_j > 0$, V_{p-} satisfies both S2 decrementing conductance and S4 incrementing it depending on whether the input line is low or high respectively; if $Y_j < 0$, then V_{p+} implements S3 decrementing or S1 incrementing conductance when input is high or low respectively. This scheme is visible in Figure 4 panes B and C, where the yellow bordered rows show the appropriate case (only one pulse is sent of the two options, (V_{p+}, V_{p-}). For many rows learning simultaneously where both types of errors exist, a negative and positive pulse can be applied sequentially to all rows; in this case, the gate voltage level for each of the programming transistors SW_{PrgJ} ensures that each row only gets the pulse it needs. Thus, at scale this approach requires two separate pulses per cycle (8 per epoch) yet only one is still sufficient to program each row. This is more economical than the bipolar approach, which not only requires four pulses but maintains a separate line for the momentary inversion of all inputs to program successfully.

A novel scheme for unipolar device integration- in which the same organic device class serves as analog array and latch- is introduced in Pane C. As pictured in Figure 3, in this case anti-parallelism is no longer required- the devices are arranged exactly like in the normal analog array. As in the other cases a large pulse is needed to unconditionally open the latch; here it is the value just below $V_{th2} = 4.8V$. Moreover, due to the

analog nature of the latches, it was found that a large negative pulse $V = 5.5V$ was needed just before this to properly reset states for each cycle due to gradually climbing conductance levels. While conditionally closing and conditionally opening the unipolar device latches is more or less synonymous in polarity with the earlier approach, it requires the use of pulses at both thresholds in order to function properly. Since the programming pulse for the row devices must also sit at their own second threshold voltage, the first threshold $|V_{\text{th1-Latch}}|$ of the latch must be greater than $|V_{\text{th2-Array}}|$ so as to keep the analog latch safely in its configured state while these pulses are applied at the end of each cycle. This requirement has motivated the use of two differently sized organic devices, a possibility offered by organic memristors.

Finally, Pane B represents a hybrid case where the same anti-parallel binary latches are used as in all past on-chip learning papers. This middle case was considered as a contrast to the all unipolar case to see if an analog latch adversely affects performance; considering the high reset pulse values needed in the all unipolar scheme (two pulses $V = 5.5V$, $V_1 = 4.7V$ required each cycle, compared to one of $V = 4.0V$ for the bipolar latch with $V_{th} = 3.7V$), it also would require the least energy expenditure per cycle of the three.

IV. Learning Performance

In order to verify the validity of these neural crossbar architectures, extensive transient simulations were performed on the Cadence platform. CMOS elements were simulated using a commercial 45nm low power design kit. Transient simulations reveal that both unipolar programming pulse schemes B and C introduced do learn successfully. Learning takes several epochs (presentations of the target) to complete; it is finished around $19\mu s$, or just before the end of the entire learning period in both cases. For both hybrid and all-organic models it was seen that final conductance values stay low overall, evolving from 150ns to somewhere in between $1\mu S - 10\mu S$, one to two orders of magnitude change but no where near the G_{max} value.

Next we evaluated the ability of hybrid and all-organic neural crossbars alike to properly learn 8 functions simultaneously over a wide variety of parasitic line resistances. Since metallic nano-wire resistances can vary within orders of magnitude depending on diameter, length, grain boundaries, scattering, and contact resistance issues, resilience to non-negligible wire resistance is an interesting topic for further exploration in nano-device array design. The 8 functions chosen (in the left of Panel A, Figure 6) are the 8 non-trivial and logically separable options amongst 2-bit functions.

At lower wire resistances (for $R < 5k\Omega$), all 8 functions are learned simultaneously in both systems without any incident. As depicted in Figure 6 B, above $5k\Omega$, the first breakdown values are obtained with the NOR function in both systems. Between $5k\Omega$ to $20k\Omega$, functions drop off sequentially, with 2 or 3 still persisting in correct output far beyond the ceiling value of $20k\Omega$. Overall, neither hybrid nor all organic perform significantly better than the other- the lines closely resemble each other despite some deviations. This indicates that the single die scheme (moving from the previous binary latch design to an analog latch design) entails no major trade-off in learning performance.

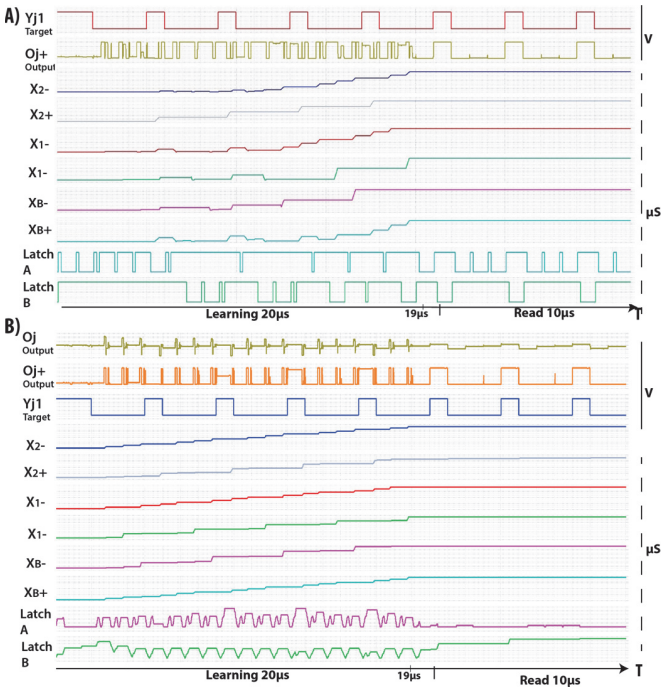


Fig. 5. A) Waveforms depicting successful learning of the AND function in the Hybrid (organic unipolar memristors, binary latch) case; target, output, 6 analog memristors being trained, and latches are depicted from top to bottom. B) Equivalent waveforms for the All-unipolar organic memristor learning case. As visible Latch behavior is now also analog.

Overall, both systems are very resilient to wire resistances. This is due to the fact that in all simulations initial conductance values are at G_{off}. As the final values exist within the range $1\mu S - 10\mu S$, this implies that even the most open memristors are still relatively resistive as compared to the wires. Yet, within the same order of magnitude (eg $R_{\text{mem}} = 100k\Omega$ while $R = 15k\Omega$), the voltage pulse needed to switch certain critical memristor devices may be attenuated enough that a threshold is no longer reached and the increments needed to implement supervised learning do not occur [14].

As seen in in Figure 6 C, some functions (eg NOR) break down early while others (eg NAND and IMP) do not. Interestingly, functions that are inverse of each other (IMP, N-IMP; CONV, N-CONV) are the ones that survive to high values of R. There is an underlying pattern: in the all-organic system functions with 3 highs (1) and one low (0) survive longer- the opposite is true for the hybrid system (functions with 3 low and 1 high survive longer). Close inspection of latch values reveals an explanation: for the analog (organic) devices, conductance varies each time but is overall higher (mean, $G_{\text{on}} = 18\mu S$, $G_{\text{off}} = 1.5\mu S$), than the simple binary latch values, which deterministically switch between $G = 10\mu S$ and $G = 100nS$, as visible in Figure 5.

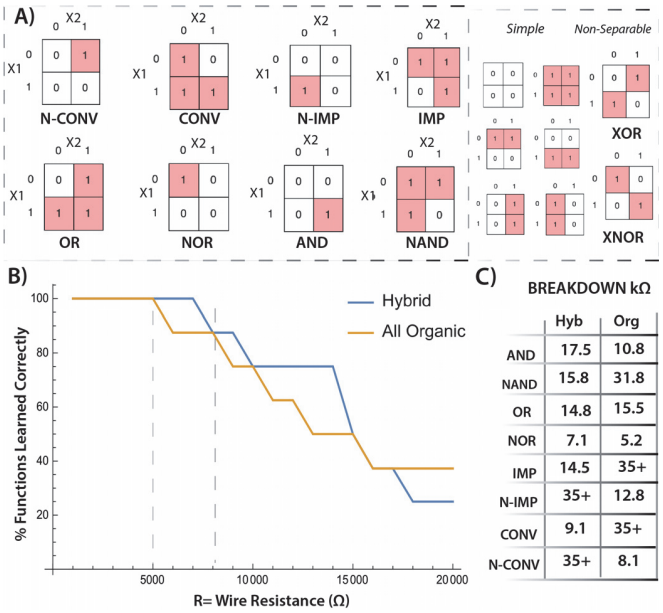


Fig. 6. A) All possible 2-bit functions. The 8 functions in the left box are the non-obvious ones on which the system is trained. B) Rate of success at learning these 8 functions for each variant system- hybrid and organic-, as a function of horizontal and vertical nanowire resistance R (40 parametric simulations), while R_{in} at the input wires is held constant at 500Ω. C) 'Breakdown' value in $R = k\Omega$: the given wire resistance at which the output for one of the four cases of the truth table is incorrect.

V. Conclusion

In this paper we showed that a unipolar organic memristor device with multiple thresholds offers advantages for on-chip supervised learning. These advantages stem from the symmetry of conductance evolution, and reduce the complexity of programming operations to one pulse per cycle for a single function learned, and 2 per cycle in array. Moreover, an all-unipolar, single die design has been introduced and confirmed to learn successfully in a non-ideal circuit with no significant tradeoffs relative to earlier architectures. While an on-chip learning scheme previously proposed for bipolar devices has been extended to new working frontiers, a head-to-head comparison between this scheme and other learning algorithms such as backpropagation and STDP using the same device model could be fruitful. Additionally, while our organic devices proved extremely resilient to parasitics and sneak paths, it remains unclear if this is an intrinsic advantage or an extrinsic effect due to smaller crossbar model systems and low conductances. Providing larger organic memristor crossbars with a richer learning task such as image recognition in the face of device variation and random initial conductances should clarify. If these simulations are similarly robust, then a hardware demonstrator will provide the ultimate test.

Acknowledgment

This work was supported by the French Agence Nationale de la recherche (MOOREA grant, ANR-12-BS03-004) and the Nanodesign Paris-Saclay lidex.

References

[1] L. Chua, V. Sbitnev, and H. Kim, "Hodgkin–Huxley Axon Is Made of Memristors," *Int. J. Bifurc. Chaos*, vol. 22, no. 3, p. 1230011, 2012.

[2] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, "Nanoscale memristor device as synapse in neuromorphic systems," *Nano letters*, vol. 10, no. 4, pp. 1297–1301, 2010.

[3] P. O. Vontobel, W. Robinett, P. J. Kuekes, D. R. Stewart, J. Straznicky, and R. S. Williams, "Writing to and reading from a nano-scale crossbar memory based on memristors," *Nanotechnol.*, vol. 20, no. 42, 2009.

[4] D. B. Strukov and K. K. Likharev, "Cmol fpga: a reconfigurable architecture for hybrid digital circuits with two-terminal nanodevices," *Nanotechnology*, vol. 16, no. 6, p. 888, 2005.

[5] A. Bandyopadhyay and A. J. Pal, "Large conductance switching and memory effects in organic molecules for data-storage applications," *Applied physics letters*, vol. 82, no. 8, pp. 1215–1217, 2003.

[6] V. Erokhin, T. Berzina, A. Smerieri, P. Camorani, S. Erokhina, and M. Fontana, "Bio-inspired adaptive networks based on organic memristors," *Nano Communication Networks*, vol. 1, no. 2, 2010.

[7] S. Song, B. Cho, T.-W. Kim, Y. Ji, M. Jo, G. Wang, M. Choe, Y. H. Kahng, H. Hwang, and T. Lee, "Three-dimensional integration of organic resistive memory devices," *Advanced Materials*, vol. 22, no. 44, pp. 5048–5052, 2010.

[8] N. Kooy, K. Mohamed, L. T. Pin, and O. S. Guan, "A review of roll-to-roll nanoimprint lithography," *Nanoscale research letters*, vol. 9, no. 1, pp. 1–13, 2014.

[9] D. Querlioz, O. Bichler, P. Dollfus, and C. Gamrat, "Immunity to Device Variations in a Spiking Neural Network with Memristive Nanodevices," *IEEE Trans. Nanotechnol.*, vol. 12, no. 3, pp. 288 – 295, 2013.

[10] M. Suri, D. Querlioz, O. Bichler, G. Palma, E. Vianello, D. Vuillaume, C. Gamrat, and B. DeSalvo, "Bio-Inspired Stochastic Computing Using Binary CBRAM Synapses," *IEEE Transactions on Electron Devices*, vol. 60, no. 7, pp. 2402–2409, 2013.

[11] D. Soudry, D. Di Castro, A. Gal, A. Kolodny, and S. Kvatinsky, "Memristor-based multilayer neural networks with online gradient descent training," *IEEE Trans. Neural Netw., in press*, 2015.

[12] M. L. Minsky and S. A. Papert, *Perceptrons - Expanded Edition*. MIT press Boston, MA:, 1987.

[13] D. Tank and J. J. Hopfield, "Simple'neural'optimization networks: An a/d converter, signal decision circuit, and a linear programming circuit," *IEEE Trans. Circ. Syst.*, vol. 33, no. 5, pp. 533–541, 1986.

[14] D. Chabi, D. Querlioz, W. Zhao, and J.-O. Klein, "Robust learning approach for neuro-inspired nanoscale crossbar architecture," *J. Emerg. Technol. Comput. Syst.*, vol. 10, no. 1, pp. 5:1–5:20, Jan. 2014.

[15] Z. Wang, W. Zhao, W. Kang, Y. Zhang, J.-O. Klein, D. Ravelosona, and C. Chappert, "Compact modelling of ferroelectric tunnel memristor and its use for neuromorphic simulation," *Appl. Phys. Lett.*, vol. 104, no. 5, p. 053505, 2014.

[16] D. Chabi, Z. Wang, W. Zhao, and J.-O. Klein, "On-chip supervised learning rule for ultra high density neural crossbar using memristor for synapse and neuron," in *IEEE/ACM Int. Symp. Nanoscale Architectures (NANOARCH)*, July 2014, pp. 7–12.

[17] S.-Y. Liao and et al, "Design and modeling of a neuro-inspired learning circuit using nanotube-based memory devices," *IEEE Trans. Circ. Syst.*, vol. 58, no. 9, pp. 2172–2181, Sept 2011.

[18] K. Gacem, J.-M. Retrouvey, D. Chabi, A. Filoramo, W. Zhao, J.-O. Klein, and V. Derycke, "Neuromorphic function learning with carbon nanotube based synapses," *Nanotechnol.*, vol. 24, p. 384013, 2013.

[19] F. Alibart, E. Zamanidoost, and D. B. Strukov, "Pattern classification by memristive crossbar circuits using ex situ and in situ training," *Nat. Comm.*, vol. 4, 2013.

[20] T. Cabaret, "Etude, réalisation et caractérisation de memristors organiques électro-greffés en tant que nanosynapses de circuits neuro-inspirés," Ph.D. dissertation, Paris 11, 2014.

[21] T. Cabaret, L. Fillaud, B. Jousselme, J. Klein, and V. Derycke, "Electrografted organic memristors: Properties and prospects for artificial neural networks based on stdp," in *Int. Conf. Nanotechnology (IEEE-NANO)*. IEEE, 2014, pp. 499–504.

[22] D. Chabi, W. Zhao, D. Querlioz, and J.-O. Klein, "On-chip universal supervised learning methods for neuro-inspired block of memristive nanodevices," *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, vol. 11, no. 4, p. 34, 2015.