Memristor networks learn by mistakes

Juan Pablo Carbajal
Institute for Energy Technology, University of Applied Sciences, Oberseestrasse 10, 8640 Rapperswil, Switzerland

Daniel A. Martin and Dante R. Chialvo*
Center for Complex Systems and Brain Sciences (CEMSC^3) & Instituto de Ciencias Físicas (ICIFI-Conicet), Escuela de Ciencia y Tecnología, Universidad Nacional de Gral. San Martín, Campus Miguelete, 25 de Mayo y Francia (1650), San Martín, Buenos Aires, Argentina
(Dated: April 12, 2022)

Recent results revived the interest in the implementation of analog devices able to perform brain-like operations. Here we introduce a training algorithm for a memristor network which is inspired in previous work on biological learning. Robust results are obtained from computer simulations of a network of voltage controlled memristive devices. Its implementation in hardware is straightforward, being scalable and requiring very little peripheral computation overhead.

I. INTRODUCTION

In the last decade we have witnessed an explosion in the interest on neuro-morphing, i.e., adaptive devices inspired in brain principles. Many of the current efforts focus on the replication of the dynamics of a single neuron, using a diversity of technologies including magnetic, optics, atomic switches, etc [1]. While the emulation of a single neuron seems achievable with present technology, we still lack learning algorithms to train large interconnected neuron-like elements, without resorting to *peripheral computation overheads.*

In thinking about this issue, it is soon realized that *in vivo* biological learning exhibits important features which are not presently considered in neuromorphing implementations. The most relevant one, in the context of these notes, is the fact that the only information a biological neuron has at its disposal to modify its synapses is either *global* or *local*. In other words, in real brains, there is no *peripheral computation overheads*, the strength of the synaptic weight between any two given neurons, is a function of the activity of its immediate neighbors and/or (through some so-called neuro-modulators) some global state of the brain (or partial region of), resulting from success or frustration in achieving some goal (or being happy, angry, excited, sleepy, etc). These observations leaded to the proposal [2–7] of a simple neural network model able to learn simple input/output associations.

The present article instantiates a translation of the work of Refs [2, 3] into the realm of memristive networks. The main objective is to design a device working on the principles described there, able to be fully implemented in hardware, requiring no access to the inner structure of the network and minimal (1 ammeter, one switch, and two batteries) external processing.

This article is organized as follows. First we will review previous work [2–7] describing a self-organized process by which biological learning may proceed. Such work is the inspiration for the algorithm proposed here to train a network of memristors [8, 9] which is introduced after that. Subsequently the main results describing the simulation results obtained from a three layer feedforward network are presented. The paper closes with a short list of expected hurdles to surpass and other possible similar implementations. Numerical details and model parameters are described on the Supplementary Material, together with some miscellaneous observations.

II. ALGORITHM, MODEL AND OBSERVABLES

A. A toy model of biological learning

A toy model of biological learning: Two decades ago, Chialvo & Bak [3] introduced an unconventional model of learning which emphasized self-organization. In that work, they re-examined the commonly held view that learning and memory necessarily require potentiation of synapses. Instead, they suggested that, for a naive neuronal network, the process of learning involves making more mistakes than successful choices, thus the process of adapting the synapses would have more opportunities to punish the mistakes than to positively reinforce the successes. Consequently, their learning strategy used two steps: the first involves extremal dynamics to determine the propagation of the activity through the nodes and the second using synaptic depression to decrease the weights involved in the undesired (i.e., mistaken) outputs. The first step implies to select only the strongest synapses for propagating the activity. The second step assumes that active synaptic connections are temporarily tagged and subsequently depressed if the resulting output turns out to be unsuccessful. Thus, all the synaptic adaptation leading to learning is driven only by the mistakes.

The toy model considered an arbitrary network of nodes connected by weights. Although almost any network topology can be used, for the sake of description let us discuss the simplest version of a three feedforward layered network (see Fig. 1-(a)). To describe the working principle, let us suppose that we wish to train the network to learn an arbitrary map. To be precise, a map is an association from each input neuron to an output neuron. For instance, the identity map is one where each input neuron is associated to the corresponding output neuron (the first input neuron is associated to the first output neuron, the second input neuron is associated to the second

* dchialvo@unsam.edu.ar

output neuron, and the same relation holds for all other input neurons), and a random map is one where one output neuron is chosen randomly for each input neuron.

The learning algorithm needs to modify the network's weights in such a way that a given input neuron connects to the prescribed output neuron.

The entire dynamical process goes as follows:

1. Activate an input neuron i chosen randomly from the set established by the task to learn.
2. Activate the neuron j_m in the middle layer connected with the input neuron i with the *largest* $w(j,i)$.
3. Activate the output neuron k_m with the *largest* $w(k,j_m)$.
4. If the output k happens to be the desired one, nothing is done.
5. Otherwise, that is if the output is not correct, $w(k_m,j_m)$ and $w(j_m,i)$ are both reduced (depressed) by an amount δ.
6. Go back to 1. Another input neuron from the task set is randomly chosen and the process is repeated.

The process involves a principle of extremal dynamics (here simplified by choosing the strongest weights) followed –in case of incorrect output– by negative feedback (i.e., the step 5 of adaptation). The only parameter of the model is δ, but, at least in the numerical simulations, it is not crucial at all, because its only role is to prevent the same path from the input to the undesired output to be selected more than once. Numerical explorations with this simple model showed that step 5 above can be modified in many different ways (including choosing random values) without serious consequences, as long as it makes less probable the persistence of "wrong paths". Initial values of $w(i,j)$ and $w(j,k)$ are not relevant either [3].

Notice that the toy model omits to consider the neuronal dynamics: it is not necessary to introduce spikes whose only role would be to propagate activity across the network. Since propagation occurs most often (statistically speaking) across the strongest synapses, the toy model omits including spikes and directly selects the strongest paths, as done in the steps 2 & 3 of the algorithm.

In Fig. 1-(b) and (c) we reproduce the results of a typical simulation in which few simple maps (indicated on Fig. 1-(d)) are successively learned by the model (see details in [3]). The error in learning map "a", computed as the squared distance between the actual output and the desired one, is seen to fluctuate until eventually vanishing at time $\simeq 600$ (see Fig. 1-(b)). After that, the network is given the task to learn map "b" (which is achieved at time $\simeq 800$), map "c" and so on. Interference between maps is expected for relatively small system size, since the same path can be chosen by chance for two different input-output maps. As was discussed earlier in Ref. 3 and 4, a system trained under these premises is robust with respect to noise, in the sense that depression of synaptic weights will self-adjust proceeding to correct the errors, until eventually achieving the desired outputs. Another interesting property of this set-up is that the learning time goes down with the size of the middle layer, a fact that is easily understood since the learning process implies to find an keep the strongest paths between the input and the desired nodes in the output layer. This and other scaling relations can be found in Refs. [3–7].

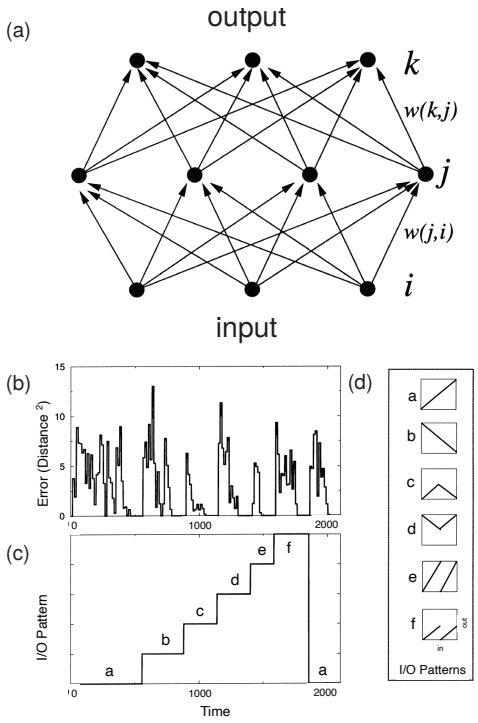


FIG. 1. An example of a three-layer network with 3 input nodes, 4 intermediary nodes and 3 output nodes is shown in (a). Each input node has a "synaptic" connection to every intermediary node, and each intermediary node has a similar connection to every output node. A typical run to learn the six simple input-output patterns (i.e., maps) for a network of 6 input neurons, 300 middle neurons and 6 output neurons, is shown in (b)-(d). As noted in (b), the error eventually reaches zero, after which learning of a new pattern (c) is attempted. The list of maps is shown in (d). Re-drawn from [3].

B. Memristor model

Now we turn to discuss how to implement the toy model just described on a network of memristors. Memristive devices are a family of two-terminal devices whose resistance evolves according to the bias and currents they experience [10]. In analogy to long term potentiation and depression taking place in neuronal synapses, memristor resistances can be increased or decreased through the application of relatively high voltage differences or currents.

In this article, we will consider Voltage Controlled Memristors with threshold [11], whose resistance R, can take any value between $R_{\rm min}$ and $R_{\rm max}$. When a voltage difference V is applied, a current I passes through the memristor, and the value of R may change, depending on V and R values.

The memristor equations can be written as:

$$I = V/R \tag{1}$$

$$\frac{\partial R}{\partial t} = -F(R,V), \tag{2}$$

where the function F describes the behavior of the memristor. Individual memristor dynamics is fully described by previous equations plus a definition of the function F, which describes how the characteristics of the memristor change upon applied voltage differences. For that, we have used the bipolar memristive system with threshold (BMS), which is described in detail in Section 3.2 of [11] (a different memristor model, known as Boundary Condition Memristor, BCM, [12, 13], is considered in the Supplementary Material [14]). According to that reference, we write:

$$F(R,V) = \begin{cases} \beta(V + V_{\updownarrow}) & \text{if } V < -V_{\updownarrow}, R > R_{\min} \\ 0 & \text{if } |V| < V_{\updownarrow} \\ \beta(V - V_{\updownarrow}) & \text{if } V > V_{\updownarrow}, R < R_{\max} \end{cases} \tag{3}$$

where $\beta > 0$ is the rate at which the resistance increases or decreases when a large enough voltage difference is applied. The function $F(R,V)$ is illustrated in Figure 2. From Eqs. 1-3, we find that a relatively large positive voltage difference tends to decrease resistance, while a negative voltage difference tends to increase resistance on the memristor (R is not modified if the absolute value of the voltage does not exceed the threshold $V_{\updownarrow}$).

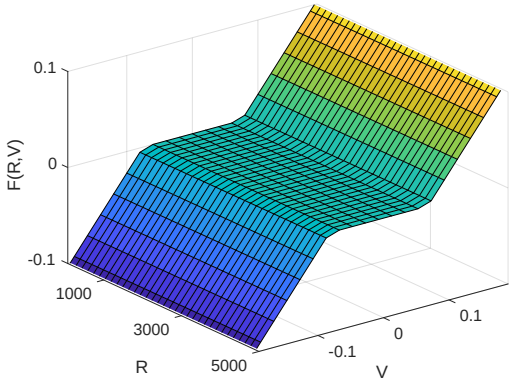


FIG. 2. Memristor behavior as a function of resistance and applied voltage. The resistance of the memristor does not change unless the absolute value of the applied voltage is greater than $V_{\updownarrow}$. Here we used $V_{\updownarrow} = 0.075$, $R_{\min} = 75$, $R_{\max} = 5000$, $\beta = 0.9$.

In Figure 3, we show an example of the typical changes exhibited by voltage controlled memristor when subjected to voltage sources of different amplitudes. The circuit and the sign convention are depicted in Fig. 3-(a), while the voltage, the current across the device and the memristor resistance are shown in Figs. 3 (b)-(d). The voltage source applies three triangular shaped low *positive* voltage pulses, which do not change memristor's resistance, followed by a high, *negative*, voltage excursion, which results on an increase of the memristor's resistance. The final triangular low voltage pulse shows

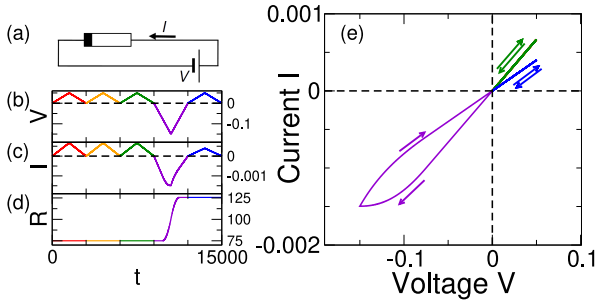


FIG. 3. Behavior of a voltage controlled memristor connected to the circuit shown in (a). Typical changes in the properties of the memristor as a function of a time dependent voltage V (b), the resulting current I and the instantaneous resistance R are shown in (c) and (d), respectively. The same data is presented in panel as an I-V curve (e). Notice that relatively small voltage excursions (i.e., upwards triangular sweeps) do not change the device resistance, while relatively large voltage excursions does it, resulting in the typical hysteresis loop. Results in panels (b)-(e) have been colored to ease the interpretation.

that the resulting resistance increase is permanent. This property will be used here to modify the network input/output paths, as explained in the following paragraphs [15].

C. Memristor Network

The results presented here correspond to numerical simulations of a three layer network of memristors [16], with N_{in} input nodes, N_{bulk} bulk nodes, and N_{out} output nodes. Pairs of nodes from successive layers are connected through bipolar memristive system with threshold (BMS) BMS memristors of Ref. [11] (see also Ref. [17]) . Notice that this three-layer network is equivalent to a memristor crossbar array [18] with N_{in} input nodes and N_{bulk} output nodes, connected to a second memristor crossbar array, of N_{bulk} inputs and N_{out} outputs. The training algorithm uses an ammeter and a voltage source with two possible values: V_{read} (read voltage) and V_{write} (punishment or correction voltage). Memristor polarity is set in such way that a negative V_{write} tends to increase their resistance (see Fig. 3-(a) and Fig. 4-(b)). The network and control resources are as sketched in Fig. 4.

We considered a relatively small (quenched) variability in the parameters of the device: the parameters for each memristor were randomly chosen from a uniform distribution with $0.8 < \beta < 1$, $0.05 < V_{\updownarrow} < 0.1$, $50 < R_{\min} < 100$, and $R_{\max} = 5000$. Initial condition was set to $R = R_{\min}$. The reading step lasted 1 time step using a voltage value $V_{\text{read}} = 0.0001$. The correction step lasted 5 time steps using a voltage $V_{\text{write}} = -0.2$.

Voltage values were chosen such that $|V_{\text{read}}| \ll V_{\updownarrow} < |V_{\text{write}}/2|$. In this way the memristor properties do not change on the reading step, and only few memristors change its resistance during the correction step of the algorithm.

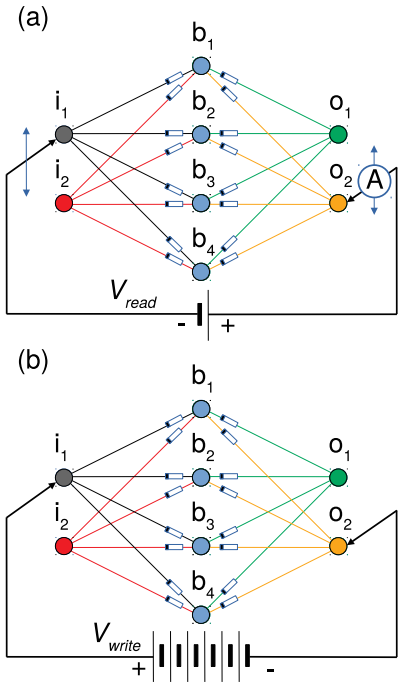


FIG. 4. Sketch of the learning algorithm for a network with $N_{\text{in}} = 2$, $N_{\text{bulk}} = 4$ and $N_{\text{out}} = 2$. In the Reading step (a) a relatively small V_{read} voltage is applied and the current at each output node is measured by the ammeter. The output node with the largest current is defined as the output. If that output is not the desired one, in the Correction step (b) a relatively large V_{write} voltage is applied to alter the resistance of the memristor path. The cycle is repeated until the desired map is learned.

D. Learning Algorithm

Here we discuss the implementation defining a simple input/output association task, similar to the one already discussed in the previous section for the case of the toy model of Fig. 1: for each input node i_n, we ask the memristor network to learn a randomly chosen output node.

The proposed training involves the following sequence at each training step:

1. Randomly choose an input node (i_n).
2. Read the current flowing through all the output nodes. Herein this is done by setting i_n voltage to V_{read}, and moving the ammeter tip (which closes the circuit) to each output node sequentially (o_1, o_2, etc), while measuring the current (see Fig. 4-(a)).
3. Determine the output node with the maximum current.
   (a) If the node with maximum current is the desired one, do nothing.
   (b) Otherwise, that is if the output maximum current is not at the desired node, apply V_{write} (see Fig. 4-(b)), and go back to point 2.
4. Go back to point 1.

The value of V_{read} needs to be small (such that it does not change the values of the resistances in the network). V_{write} is large and with inverted polarity, hence inducing an increase of the resistances of the network. This is the only crucial factor, to ensure that the reading (in step 2) is not modifying the network conductances and conversely that the correction (in step 3b) decreases the likelihood of having large currents in the undesired paths. It is evident that the memristor learning algorithm preserves the same spirit of the earlier work: to punish wrong paths by increasing the resistance of the involved memristors. We applied step 3b a maximum of $n_{\max} = 80$ times on each training step [19].

To collect the statistics presented here, at the end of each training step we calculate the learning error as follows. For each input node, we find the largest output (we apply V_{read} among that input node and all output nodes sequentially, and take the node through which current flow is largest). We define the error as the Hamming distance from the vector of largest outputs and the desired map. If the error is null, the network has learned. Otherwise, a new training step is performed. In some cases, after the network has learned, we will consider to train it with a different map. The pseudocode for this algorithm is shown in the Supplemental Material [14].

III. RESULTS

Now we proceed to describe the parametric behavior of the algorithm just explained in the previous paragraphs. First we explored the dependence of the learning time on the size of the middle layer N_{bulk} for random maps. As discussed, larger values of N_{bulk} in the neuron network model of Ref. [3, 4] provided more paths to the correct output, which leads to shorter learning time. We found very similar performance for the memristive network as shown in the results of Fig. 5 where success (the fraction of networks that have learned a map) is shown as a function of the training step and N_{bulk} for a three layer networks with $N_{\text{in}} = N_{\text{out}} = 3$, and $N_{\text{in}} = N_{\text{out}} = 4$, and several values of N_{bulk}. In panels (a) and (b), success as a function of the step number improves with larger N_{bulk}. This is also apparent when we plot the fraction of networks that have learned at (or before) correction step 1000 (see panels (c) and (d)) showing that performance is an increasing function of the middle layer size, N_{bulk}. It can also be noticed, from Fig. 5-(c) and (d), that Success in equal to 1 for long enough N_{bulk}, which means that any map can be learned on those networks. Similar numerical simulation results, for a different memristor model [12, 13], with parameters fixed to represent the behavior of the first reported memristor [10, 20], are shown in the Supplementary Material [14]. This similarity suggests that the performance of the learning algorithm presented here does not depend strongly on the details of the memristors used.

The implementation of the training algorithm shows also that the memristor network learns a series of maps in a way similar to that exhibited by the earlier neuronal model [3, 4]. This can be seen in the example of Fig. 6, which shows the evolution of the network with $N_{\text{in}} = 4$, $N_{\text{bulk}} = 200$, $N_{\text{out}} = 4$. The network was trained in one of the labeled maps until even-

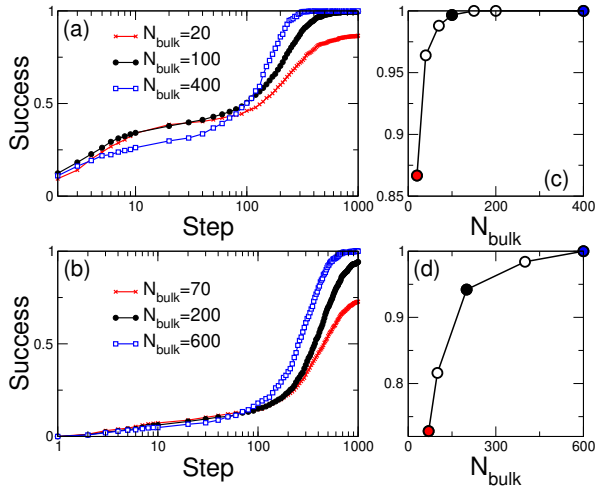


FIG. 5. Learning performance as a function of the middle layer size. Results show the success as the fraction of networks that learn a random input/output association map after a given number of correction steps for $N_{\rm in} = N_{\rm out} = 3$, $N_{\rm bulk} = 20$, 100 and 400 (a), and for $N_{\rm in} = N_{\rm out} = 4$, $N_{\rm bulk} = 70$, 200 and 600 (b). Success at 1000 steps, as a function of $N_{\rm bulk}$ for $N_{\rm in} = N_{\rm out} = 3$ (c) and for $N_{\rm in} = N_{\rm out} = 4$ (d) are also shown. In (c) and (d) the coloured symbols correspond to the results in (a) and (b) obtained with the respective $N_{\rm bulk}$ values. All results are averages over at least 500 network realisations, yielding values of standard errors smaller than the symbols size ($SEM \sim 0.025$).

tually the error is zero, at this point it starts being trained on a different map and so on. Notice the resemblance with the results in Fig. 1, which suggests that the training strategy proposed here is capturing the essence of the learning algorithm of Ref. [3, 4].

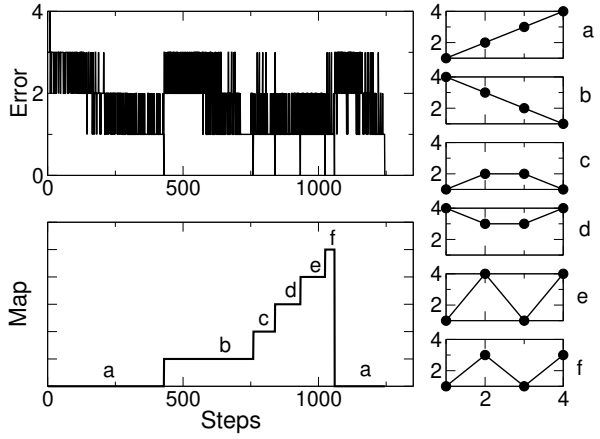


FIG. 6. Typical evolution of the training of a network learning seven successive maps. Upper main panel shows the error (Hamming distance from the desired to the current output) as a function of training steps. The input/output maps, labeled **a** to **f**, are depicted in the right column and presented sequentially as indicated in the lower main panel. Network with $N_{\rm in} = N_{\rm out} = 4$, $N_{\rm bulk} = 200$, other parameters same as in Fig. 5.

Another distinctive property of the proposed training strategy is the fact that the memristive networks are robust to perturbations of the device properties. This alterations can be seen, for instance, as changes in resistance, which in real networks can be due to volatility, defects, etc. As an example, we plot in Fig. 7 the evolution of a network ($N_{\rm in} = N_{\rm out} = 4$, $N_{\rm bulk} = 200$), which after learning the identity map, is periodically perturbed. It can be seen that after each perturbation, the network recovers to null error learning the map in a few additional steps. This ability is not surprising given the fact that the network proceeds with the perturbation in the same way than during the usual learning process.

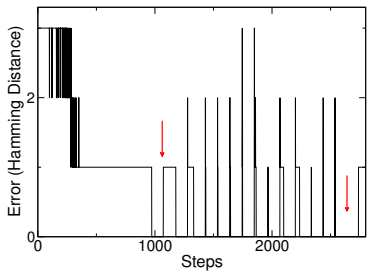


FIG. 7. Example of the network recovery after a single perturbation. We plot the error (i.e., the Hamming distance from the desired to the current output) as a function of the training steps for a network which is learning the identity map. 100 steps after the map is learnt, 10% randomly chosen memristors are perturbed by increasing their resistances by 5%. The first downward arrow indicates the first perturbation, while the second downward arrow denotes the 14th perturbation (which did not increase the network error). Network with $N_{\rm in} = N_{\rm out} = 4$, $N_{\rm bulk} = 200$, other parameters same as in Fig. 5

To gain insight on the dynamics of the memristor network resistances during learning of successive maps a $N_{\rm in} = N_{\rm out} = 3$, $N_{\rm bulk} = 400$ network was trained to learn all the possible $N_{\rm out}^{N_{\rm in}} = 27$ maps. After all maps were learned, the network location of each memristors were randomly shuffled (preserving their resistance values). After that, the network was re-trained to learn the same $N_{\rm out}^{N_{\rm in}}$ maps.

In Fig. 8-(a), we show the histogram of resistances, at the beginning of the simulation, after learning all possible maps once (labeled 27) after shuffling and relearning all maps four times (labeled 108), and after shuffling and re-learning all maps ten times (labeled 270). In panel B, we show the evolution of the average resistance, as a function of the number of learned maps. As expected from the nature of the training algorithm, resistances can only grow when performing correction steps. Moreover, from Fig. 8-(a), the range of resistance values increases with the number of steps. We remove the effect of growing resistance by normalizing with the average resistance of the ith network, R_i. In Panel (c) we plot a histogram of the normalized resistances, where each resistance value was divided by the average resistance of the network. Finally, in Panel (d), we show the coefficient of variation, $< CV >$, of a single network, computed as the standard deviation divided by the mean value of all resistances in the network. After repeated learning the distribution tends to a Gaussian distribution, approaching a $< CV > \sim 1/3$.

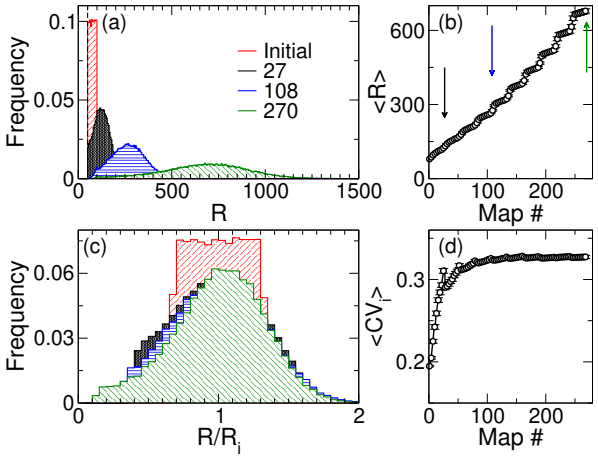


FIG. 8. Evolution of the distribution of resistance values after repeated re-learning of the same set of maps. Resistance histogram at the beginning of the simulation, after learning all possible maps once (labeled 27) after shuffling and relearning all maps four times (labeled 108) and after shuffling and re-learning all maps ten times (labeled 270) are shown in (a). The dashed vertical line shows the minimum possible value $R_{\rm min} = 50$. Average resistance as a function of number of learned maps is shown in (b). Arrows indicate the maps used for the histograms in (a). The same data as in (a) after normalising each memristor's resistance by the average value of the whole network is shown in (c). Coefficient of variation of memristor's resistance, as a function of the number of learned map is shown in (d). Results are calculated for a single network and then averaged over 25 network realisations. Bin-size $\Delta R = 5$ for histogram in (a) and $\Delta R = 0.05 R_i$ in (b). Network with $N_{\rm in} = N_{\rm out} = 3$, $N_{\rm bulk} = 400$, other parameters values as in Fig. 5.

IV. DISCUSSION

Several learning strategies for memristor networks have been recently proposed [21–24].

For problems with time dependent inputs, one strategy consists on using *reservoir computing* [25], where the inputs are connected to a reservoir network composed of interconnected time dependent elements (whose properties continue evolving *after* a stimulus is applied), such as *volatile* memristors. The response of the network is then classified through an output layer, that needs to be trained. Reservoir computing has also been applied to time independent inputs that are re-encoded as time-dependent ones [23]. While this kind of networks differ from those studied here, an error penalization learning mechanism, such as the one studied here, may be useful for training the output layer of this scheme.

For time-independent inputs, several learning strategies considering *non-volatile* memristors (as done here), have also been proposed. Most of them are inspired in *Machine learning* algorithms, such as *Gradient descent* training (see e.g. [21, 24]), where the value of each resistance needs to be known and modified at each correction step. As an alternative, a *random weight change algorithm*, which does not require to know the precise values of all elements, has recently been proposed for training a network where each synapse is composed of four memristors, and several transistors [22].

Similarly, here we have described a training procedure allowing a simple network of memristors to learn any arbitrary input/output association map. This is achieved without any detailed information of the inner structure of the layers. The method is inspired by a learning algorithm proposed about 20 years ago [3, 4], however, there are some differences among them.

First, in the original model, the active neurons transmitted their activity to just one neuron on the following layer, while here, the current passing through each node is determined by Kirchoff's laws. Probably, the inclusion of diodes or transistors (such as in Refs. 22 and 23) may generate dynamics which are closer to the original proposal. The study of this possibilities is an interesting avenue for future research.

Second, in the correction step for the toy model, only the (two) intervening connections are punished. Here, instead, we apply V_{write} over the input and output nodes, in such way that the correction voltage difference (and consequently the correction) over each memristor in the network is proportional to the voltage involved in the wrong answer. In this way, the correction is performed without the need to measure or to have control over the middle layer: the learning method does not need to control individual memristors and requires only access to read and/or perturb two nodes from the (arbitrarily large) networks at any given time (consider, as an example, the network of Fig. 8, which has 2400 memristors, and only 3 input nodes plus 3 output nodes need to be considered in the present learning method).

Despite the extreme simplicity of the approach, it is demonstrated that an iterative reading of the current flowing between two points of the network, and its eventual perturbation, can modify the overall network connectivity until arriving to one of the possible solutions. By design, the learning is robust against different types of perturbations, including differences and fluctuations in the memristors parameters, noise and defects as well as distribution of polarities (see Supplementary Material [14]). Although the present results are limited to numerical simulations, in case of being implemented in hardware, the method is easily scalable to arbitrarily large network sizes.

Notice that the memristor network is able to learn despite lacking an important feature of its biological neural contrapart: the spikes. That is not a limitation, because as commented in the introduction, in this kind of networks, the only role for the spikes would be to propagate the information from the input node/s through the network to some output node/s. Instead, in the present implementation, this is also achieved, but by the current flow from an external battery. Thus, knowing which input node is connected the current flowing through the network will be reflected on the value of the output current at a given node. In this context, our approach is a simple solution which does not require at all of the implementation of any spiking mechanism. Obviously, the absence of additional electronic to implement neurons becomes very relevant when considering a hardware implementation of this concept. Concerning hardware, the proposed learning method benefits from variability in the network properties, therefore not requiring of

precise control on the memristor parameter during manufacturing.

Several analytical results have been provided for the original model [5, 6]. For instance, it has been shown that a geometric transition from no-learning to learning takes place at $N_{Bulk} = N_{in} \times N_{out}$. While our numerical simulation results show similarities with the original learning algorithm, it would be useful to verify whether and how these results are valid here. Also, extending other analytical approaches, such as the work by Caravelli *et. al.* [26, 27] to the system studied here, should give a deeper understanding on how learning in this memristor network takes place.

In these notes we limited ourselves to the presentation of the most fundamental aspects of the results. As a salient feature, we reported how a single network can learn many maps, and this causes an expansion of the distribution of their resistances, possibly due to the multiplicity of solutions to learn a single map. Nonetheless some few caveats must be mentioned. In the first place we consider adversarial situations for the algorithm, concerning some simple variations on the approach, where changes in the initial condition of memristor conductances, polarity as well as minor changes in the type of correction step are described. These studies are presented in the Supplementary Material [14]. Second, we did not expand here on discussing the type of problems that the present approach can solve. This issue would require of extensive numerical simulations and it seems to deserve being explored on a hardware implementation, since it will work tens of orders of magnitude faster than any of our current numerical simulations. Finally, we expect the approach to be useful on a variety of network topologies including less ordered systems such as a random network of nanowires [28] whose conductivity can be varied by applying a voltage difference among pairs of points in the network. The algorithm may also be useful on other structures [29] as long as a correction mechanism increasing resistance over not desired paths can be generated.

V. CONCLUSIONS

In summary, we have introduced an algorithm able to train a memristor network to learn arbitrary associations. Robust results for its performance were demonstrated using numerical simulations of a network of voltage controlled memristive devices. Given the design principles, the results suggest that its implementation in hardware would be straightforward, being scalable and requiring very little peripheral computation overhead.

ACKNOWLEDGMENTS

Work partially supported by 1U19NS107464-01 NIH BRAIN Initiative (USA) and CONICET (Argentina).

[1] D. Marković, A. Mizrahi, A. Querlioz and J. Grollier, Nature Reviews Physics **2**, 499–510 (2020).

[2] P. Bak, *How nature works: The science of self-organized criticality* (Springer Science, New York, 1996).

[3] D.R. Chialvo and P. Bak Neuroscience **90**, 1137–1148 (1999).

[4] P. Bak and D.R. Chialvo Phys. Rev. E **63** 031912 (2001).

[5] J. Wakeling Physica A **325**, 561–569, (2003).

[6] J. Wakeling Physica A **340**, 766–773, (2004).

[7] M. Brigham M. *Self-Organised learning in the Chialvo-Bak model*. MSc Dissertation. Artificial Intelligence, School of Informatics. University of Edinburgh. UK (2009).

[8] L.O. Chua, IEEE Transactions on Circuit Theory **18**, 507–519 (1971).

[9] F. Caravelli and J.P. Carbajal, Technologies **6**, (4) 118 (2018).

[10] D.B. Strukov, G.S. Snider, D.R. Stewart, and R.S. Williams, Nature **453**, 80–83 (2008).

[11] D. Biolek, M. Di Ventra, and Y.V. Pershin. Radioengineering, **22** (4), 945–968 (2013).

[12] F. Corinto and A. Ascoli, *IEEE Trans. Circuits Syst. I*, **59** (11) 2713–2726 (2012).

[13] A. Ascoli, F. Corinto, V. Senger and R. Tetzlaff, *IEEE Circuits and Systems Magazine*, **13** (2), 89-105 (2013).

[14] See Supplemental Material at [URL will be inserted by publisher] for the mathematical description of the memristor model; the flow diagram of the learning algorithm; observations on the learning performance of a memristor network with random polarity or random correction voltage; and the pseudocode of the numerical simulations.

[15] In the following, it is customary to express all resistances in terms of an arbitrary unit resistance R_0, thus omitting units. Similarly, all voltages are written in terms of a unit voltage V_0, and time is also measured in terms of time unit t_0. Consequently, currents are measured in units of V_0/R_0, and parameter β (which describes the rate at which memristor's resistance change, see Supplemental material [14]) in units of $R_0/V_0/t_0$.

[16] The computer code to reproduce the following results can be downloaded from *https://github.com/DanielAlejandroMartin/Memristor121*(tag v1.0). A complementary library written in python is deposited at *https://gitlab.com/kakila/memnet*. Further considerations regarding potential memristive implementation can be consulted in *https://kakila.gitlab.io/mistake-learning/*

[17] D. Querlioz, P. Dollfus, O. Bichler, and C. Gamrat. IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH) 150-156 (2011).

[18] Xia, Q., and Yang, J.J. Nat. Mater. **18**, 309–323 (2019).

[19] The actual value of n_{max} does not determine whether the network learns. It only changes the number of required learning steps. Taking $n_{max} = 1$ would make our computational implementation spend too much time reading for each correction, while setting $n_{max} = \infty$ may make some numerical simulations get struck (i.e., in networks that do not learn).

[20] M. D. Pickett, D. B. Strukov, J. L. Borghetti, J. J. Yang, G. S. Snider, D. R. Stewart, and R. S. Williams, *J. Appl. Phys*,**106**, 074508, (2009).

[21] R. Hasan, T.M. Tarek, C. Yakopcic, *Microelectronics Journal* **66**, 31–40, (2017).

[22] C. Yang, S. P. Adhikari and H. Kim, *IEEE Transactions on Circuits and Systems I* **66** 10, 3906-3916 (2019).

[23] Midya, R., Wang, Z., Asapu, S., Zhang, X., Rao, M., Song, W., Zhuo, Y., Upadhyay, N., Xia, Q. and Yang, J.J. *Adv. Intell. Syst.* , **1**, 1900084, (2019).

[24] D. Soudry, D. Di Castro, A. Gal, A. Kolodny and S. Kvatinsky, *IEEE Transactions on Neural Networks and Learning Systems* , **26**, 10, 2408-2421 (2015).

[25] J. P. Carbajal, J. Dambre, M. Hermans and B. Schrauwen, *Neural Computation* ***27****, (3), 725-747, (2015)*

[26] F. Caravelli, *Phys. Rev. E* **96**, 052206 (2017).

[27] F. Caravelli, F. C. Sheldon, and F. L. Traversa, *Science Advances* **7** 52, eabh1542 (2021).

[28] H.G. Manning, F. Niosi, C.G. da Rocha, A.T. Bellew, C. O'Callaghan, S. Biswas, P.F. Flowers, B.J. Wiley, J.D. Holmes, M.S. Ferreira, and J.J. Boland, Nat. Commun., **9**, 3219, (2018).

[29] S. Fostner, R. Brown, J. Carr, and S. A. Brown, *Phys. Rev. B* **89**, 075402 (2014).