Extended memory lifetime in spiking neural networks employing memristive synapses with nonlinear conductance dynamics

Nanotechnology

PAPER

Extended memory lifetime in spiking neural networks employing memristive synapses with nonlinear conductance dynamics

IOP Publishing Nanotechnology
Nanotechnology 30 (2019) 015102 (12pp) https://doi.org/10.1088/1361-6528/aae81c

Extended memory lifetime in spiking neural networks employing memristive synapses with nonlinear conductance dynamics

S Brivio[1,4], D Conti[2], M V Nair[3], J Frascaroli[1], E Covi[1], C Ricciardi[2], G Indiveri[3] and S Spiga[1,4]

[1] CNR—IMM, Unit of Agrate Brianza, via C. Olivetti 2, I-20864 Agrate Brianza, Italy
[2] Politecnico di Torino, Dipartimento di Scienza Applicata e Tecnologia (DISAT), Corso Duca degli Abruzzi 24, I-10129 Torino, Italy
[3] Institute for Neuroinformatics, University of Zürich and ETH Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

E-mail: stefano.brivio@mdm.imm.cnr.it and sabina.spiga@mdm.imm.cnr.it

Received 9 October 2018
Accepted for publication 12 October 2018
Published 31 October 2018

Abstract

Spiking neural networks (SNNs) employing memristive synapses are capable of life-long online learning. Because of their ability to process and classify large amounts of data in real-time using compact and low-power electronic systems, they promise a substantial technology breakthrough. However, the critical issue that memristor-based SNNs have to face is the fundamental limitation in their memory capacity due to finite resolution of the synaptic elements, which leads to the replacement of old memories with new ones and to a finite memory lifetime. In this study we demonstrate that the nonlinear conductance dynamics of memristive devices can be exploited to improve the memory lifetime of a network. The network is simulated on the basis of a spiking neuron model of mixed-signal digital-analogue sub-threshold neuromorphic CMOS circuits, and on memristive synapse models derived from the experimental nonlinear conductance dynamics of resistive memory devices when stimulated by trains of identical pulses. The network learning circuits implement a spike-based plasticity rule compatible with both spike-timing and rate-based learning rules. In order to get an insight on the memory lifetime of the network, we analyse the learning dynamics in the context of a classical benchmark of neural network learning, that is hand-written digit classification. In the proposed architecture, the memory lifetime and the performance of the network are improved for memristive synapses with nonlinear dynamics with respect to linear synapses with similar resolution. These results demonstrate the importance of following holistic approaches that combine the study of theoretical learning models with the development of neuromorphic CMOS SNNs with memristive devices used to implement life-long on-chip learning.

Supplementary material for this article is available online

Keywords: RRAM, memristor, neuromorphic, spiking neural network, memory lifetime, ReRAM, HfO_2

(Some figures may appear in colour only in the online journal)

Introduction

Spiking neural networks (SNNs) that implement brain-inspired neural processing models can be endowed with life-

[4] Authors to whom any correspondence should be addressed.

long online learning features [1]. Hardware implementations of these networks can lead to the construction of efficient neuromorphic computing systems able to extract useful information from large amount of unstructured data in real-time [2]. These architectures represent a radical departure from the standard machine learning and information processing systems based on the von Neumann architecture [2, 3]. Indeed, the physical separation of storage and processing units typical of von Neumann scheme and their synchronous mode of operation are being questioned, in view of energy-efficient massively parallel brain-inspired in-memory computing operations. In the last decades, significant efforts have been devoted to design and produce computing systems based on standard very large scale integrated (VLSI) silicon technology that partially surpass von Neumann architectural limits [4–6]. The remaining hurdle of such innovative systems lies in the inefficiency of employing conventional memory elements as synapse units. Indeed, on the one hand, they cannot fulfil the *technical* requirements as nanometre size scaling, non-volatility, parallel programming and low voltage/power operation simultaneously. On the other hand, which is even more *fundamental*, the finite resolution of a generic hardware synaptic element still remains the critical issue that limits the memory capacity of SNNs [1, 7]. As a consequence, a finite lifetime exists for memories that are continuously replaced (forgotten) by new ones [8], thus deteriorating the learning and the overall performance of an SNN. Emerging resistive random access memories, classified into the category of memristive devices, i.e. two-terminal devices that undergo resistance change upon voltage stimulation, are expected to comply with all the aforementioned *technical* requirements for SNNs. On the contrary, the *fundamental* limitation of the memory lifetime in SNNs employing memristive devices has never been investigated, despite it can provide clues for the optimisation of SNNs as a whole.

In this study, we aim at filling this gap by investigating the role of the synaptic conductance dynamics on the memory lifetime of an SNN. An SNN constituted by memristive devices characterised by nonlinear conductance evolution that slowly approaches the boundary values is benchmarked to the same network architecture constituted by theoretical linear synapses with hard conductance boundaries. The model for the memristive synapses well reproduces the experimentally observed conductance evolution of HfO_2-based devices as a function of trains of identical pulses [9]. The constitutive equations for the neuronal units are derived from a fabricated asynchronous mixed digital-analogue sub-threshold neuromorphic CMOS processor [6]. Therefore, the synergy of hardware realisations of neuronal and synaptic elements is investigated and optimised in a holistic approach, as an additional novelty point with respect to pioneering works in which SNNs are simulated [10–20] or implemented in microcontrollers or field programmable gate arrays [21–23]. Indeed, the compatibility of such theoretical SNN with very large scale integration is not straightforward. The learning mechanism of the system relies on the update of specific synaptic weights and depends on two factors: The timing of the spikes generated by the pre- and post-neurons that the synapse connects, as in the case of spike timing-dependent plasticity (STDP) [24], and the occurrence rate of pre- and post-neuronal spikes [25–27]. These elements contribute to realise a spike timing and rate-dependent plasticity (STR-DP) learning rule. The vast majority of the investigations on STDP-based SNNs employing emerging memristive devices bases their timing evaluation on nearest-neighbour paired spikes [14, 15, 18, 20, 28]. Only few recent works propose implementations for rate-based computing [29, 30]. In general, most of the studies on memristor-based SNNs seldom considers other spike pairing schemes (triplets or all-to-all) or any kind of rate dependence despite these features are acknowledged both as a fundamental contribution in human brain cognition [31] and as a valuable tool for neuro-inspired computation, since they improve the sensitivity to spatio-temporal correlations [32–34].

From the simulations, we find that memristor soft-bound dynamics results in improved SNN memory lifetime and capacity and slower learning speed, thus ensuring slower forgetting and higher and more robust recognition rate in comparison to linear synapses with similar resolution. The findings mark a difference between requirements for synapses to be used in SNNs with online life-long feed-forward learning, as in the present study, and those to be employed in accelerators for deep learning, which need perfectly linear updates [35–37]. Furthermore, the results are in agreement with computational neuroscience mean-field simulations [1] and open the way for a joint optimisation of CMOS neurons and memristor dynamical features towards life-long online learning systems based on SNNs.

Methods

Memristive devices are two terminals metal/oxide/metal structures that undergo a voltage-controlled conductance change [9]. In this study, we use $TiN/HfO_2/Ti/TiN$ structures, whose operation relies on formation and dissolution of conductive filaments [17, 38–41] which short and disconnect the metal ends, respectively [42]. The device fabrication comprises sputtering deposition of the metal electrodes, atomic layer deposition of the oxide layer and patterning by photolithography and lift-off, as specified elsewhere [38, 39, 43–45].

Devices are tested in a standard probe station equipped with Keysight B1500A instrument. Pulses are sent through a B1525A semiconductor pulse generator unit and current is read through a B1511B source measuring unit, both interfaced with the device through a custom board [17]. Voltage is applied to the Ti/TiN top electrode and the bottom TiN contact is kept at ground voltage. Devices show resistance switching phenomenon after an electroforming process in which a current ramp is forced to flow through the device until the resistance drops to a low value [40, 41]. After the forming process, the resistance can be increased (in the following referred to as long term depression, LTD, of the conductance) and decreased (in the following referred to as long term potentiation, LTP, of the conductance) with negative and positive voltages, respectively. To characterise their dynamics,

the devices are stimulated by trains of identical pulses and their resistance is read after each pulse at 100 mV. Pulses are 10 μs-long with rise and fall times of 40 ns. The pulse voltage is 0.5 V for LTD and −0.45 V for LTP. The switching has been verified on 10 different devices for 10–150 LTD/LTP cycles. The pulsed operation can be carried out without any external element limiting the current [40, 46].

We simulate an SNN trained to recognised the handwritten digits of the MNIST data-set [47]. Constitutive equations for synapses and neurons are derived from experimental data and from a fabricated asynchronous mixed digital-analogue and sub-threshold neuromorphic CMOS processor that comprises leaky integrate and fire (LIF) neurons with plasticity circuits implementing STR-DP [6]. The neuron contains also a differential pair integrator module (DPI) [48] that is a current-mode log-domain integrator operated in sub-threshold regime. The DPI implements a low-pass filter whose time constant is in the order of tens of ms and it shapes the pre-synaptic signal into an exponential current resembling the excitatory and inhibitory post-synaptic currents present in biological systems. The SNN is simulated through a Brian2 code [49, 50].

Results

State-dependent synaptic weight update

Long term storage of events stimulating an SNN is limited by hardware synapses that can only store a finite number of communication efficacy values or weights between neurons (w, normalised between [0, 1] in this paper). When the storage capacity is saturated, every new experience results in erasure of an old one, thus limiting the lifetime of the memories stored in the network. Fusi and Abbott [1] demonstrated that both the number of available weight values and dynamics of the weight update affect the memory lifetime. In particular, they proposed a weight-dependent update rule with soft boundary limits that results in improved SNN memory lifetime. The synaptic weight update, δw_{pulse}, produced by a single LTP or LTD pulse, is given by the following equations: [1]

$$\begin{aligned} \text{LTP: } \delta w_{pulse,P} &= \alpha_P \cdot (1 - w)^{\gamma_P} \\ \text{LTD: } \delta w_{pulse,D} &= -\alpha_D \cdot w^{\gamma_D}. \end{aligned} \quad (1)$$

These equations imply that the end values are strictly reached only after an infinite number of LTP (LTD) pulses. Therefore, the number of levels accessible by the synapses cannot be easily defined. However, the parameters $\alpha_{\{P,D\}}$ can be considered to be inversely proportional to the number of levels [1]. The parameters $\gamma_{\{P,D\}}$ affect the speed of approaching the asymptotic boundary values.

Figure 1 demonstrates the analogy between the conductance dynamics of $TiN/HfO_2/Ti/TiN$ memristors and the soft-bound law defined by equation (1). The device conductance, G, can be increased (figure 1(a)) and decreased (figure 1(b)) in an analogue and nonlinear fashion by trains of identical 10 μs-pulses with opposite polarities during LTP and LTD, respectively. The symbols in figures 1(a) and (b) show the average device dynamics and the grey areas indicate the 1σ dispersion in experimental data over 50 measurements as a function of the number of delivered pulses. The conductance evolution follows a fast or a slow variation when departing or approaching the boundary conductance values, in agreement with other observations [41, 51–53]. The rate of departure and approach can be tuned to some extent by changing pulse voltage and time-width [9, 15, 40, 41]. In this paper, we choose the pulse parameters that minimise the conductance change per pulse (i.e. minimise $\alpha_{\{P,D\}}$) and simultaneously preserve data reproducibility. Fitting equation (1) to the observed average memristor dynamics returns values of $\alpha_P = 0.0064$, $\alpha_D = 0.0053$, $\gamma_P = 3.2$ and $\gamma_D = 3.4$.

The maximum and minimum conductance values are reached at saturation as for a soft-bound model. The nonlinear dynamics implies a weight update (δw_{pulse}) that depends on the instantaneous value of the weight and nullifies when boundary values are approached, i.e. when $w \to 1$ for LTP and $w \to 0$ for LTD, as shown by figure 1(c) and as modelled by equation (1). It is worth specifying that, in the following, the weight will be considered to be the normalised version of the conductance between 0 and 1. The dashed lines in figures 1(a) and (b) correspond to the fitting lines of the average of the experimental curves, thus experimentally attesting the soft-bound behaviour first introduced by Fusi and Abbott [1]. The significant variability in the conductance dynamics visible in figures 1(a) and (b) is expected for filamentary devices [11, 54–56]. The memristive device simulation takes variability into account through an additive Gaussian stochastic term for each LTP and LTD pulse. Figure 1(d) compares 10 experimental and simulated LTP–LTD operations (small filled symbols and empty symbols, respectively). Simulations are performed according to equation (1) and include the additive Gaussian term with amplitude adapted to reproduce the experimental variation (refer to section 1 of the supplementary material available online at stacks.iop.org/NANO/30/015102/mmedia for the simulation parameters). The white line in figure 1(d) corresponds to the average of the simulated data.

As noticed above, the estimation of the effective number of levels covered by the memristive device is given by $1/\alpha_{\{P,D\}} \approx 150 - 200$. However, it is worth specifying that, strictly speaking, the device does not implement a multilevel operation, because of the large variability shown also in figures 1(a) and (b). In any case, also in agreement with Fusi and Abbott [1], in the following, we will use the quantity $1/\alpha_{\{P,D\}}$ as a measure of the effective device resolution and, for brevity, we will refer to it as *number of levels* and *resolution* of the memristive device.

Timing- and rate-dependent architecture

In the network, LIF neurons are simulated, according to equations derived from the circuit implementations [6]. Each neuron is characterised by two internal state variables, I_{Ca} and I_{mem}, as represented in figure 2(a). Let us consider input (INs) and output neurons (ONs) connected by a matrix of trainable synapses. The internal variables of the ONs control the

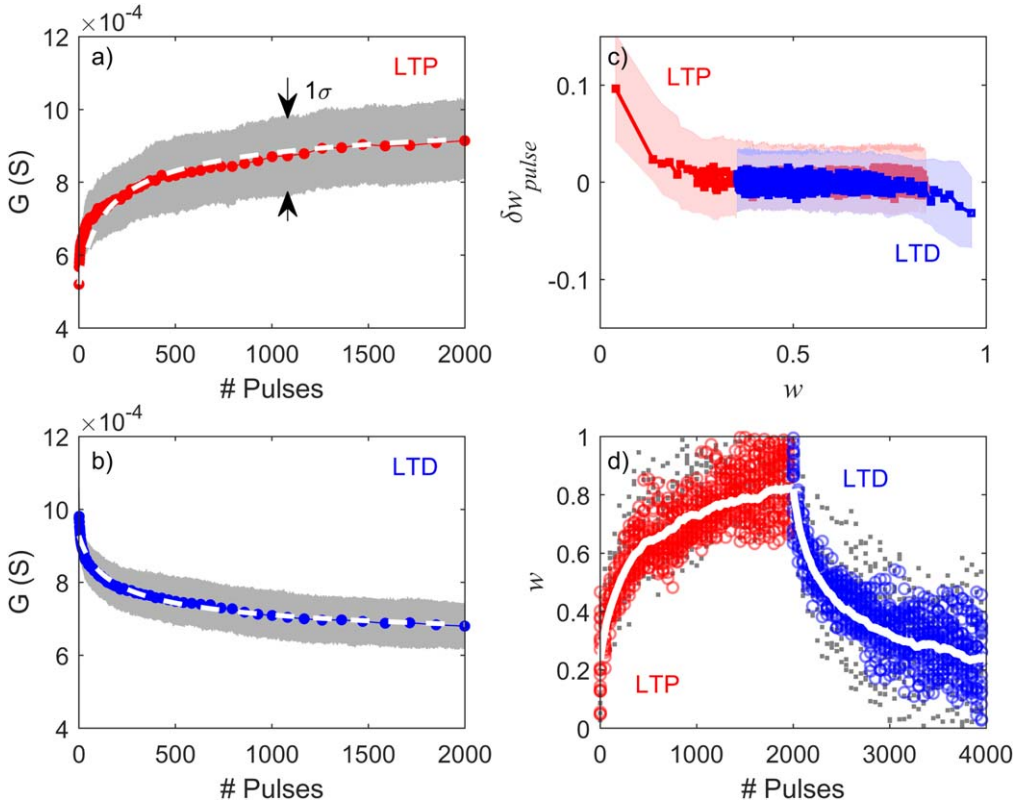


Figure 1. Representative conductance evolution as a function of the number of pulses from 1 to 2000 pulses for LTP (a) and LTD (b): symbols correspond to the average of 50 experimental curves; dashed lines are the fitting curves. Grey shaded regions correspond to 1σ variability. (c) Conductance change driven by one pulse as a function of the initial resistance for both LTP and LTD, in red and blue respectively. (d) Comparison of ten weight evolutions as obtained from measurements and simulations including pulse-to-pulse variability for LTP (from pulse 1 to 2000) and LTD (from pulse 2000 to 4000). Grey small filled symbols are experimental data and large empty symbols corresponds to the simulated data. Thick white lines are the average of the simulated data.

programming of the synaptic matrix. For the sake of brevity, we define the neuron of figure 2(a) as an ON. An ON collects the spikes fired by N INs and weighted by N synapses. The ON integrates in time (with some leak) the input current, I, into the membrane current, I_{mem}, the first internal neuron variable:

$$\frac{dI_{mem}}{dt} = -\lambda + I(t), \tag{2}$$

where λ is the leak term. Every time I_{mem} becomes larger than the threshold θ_{fire}, the ON fires output spikes, whose rate is monitored by the second internal variable, the calcium current, I_{Ca}, defined as:

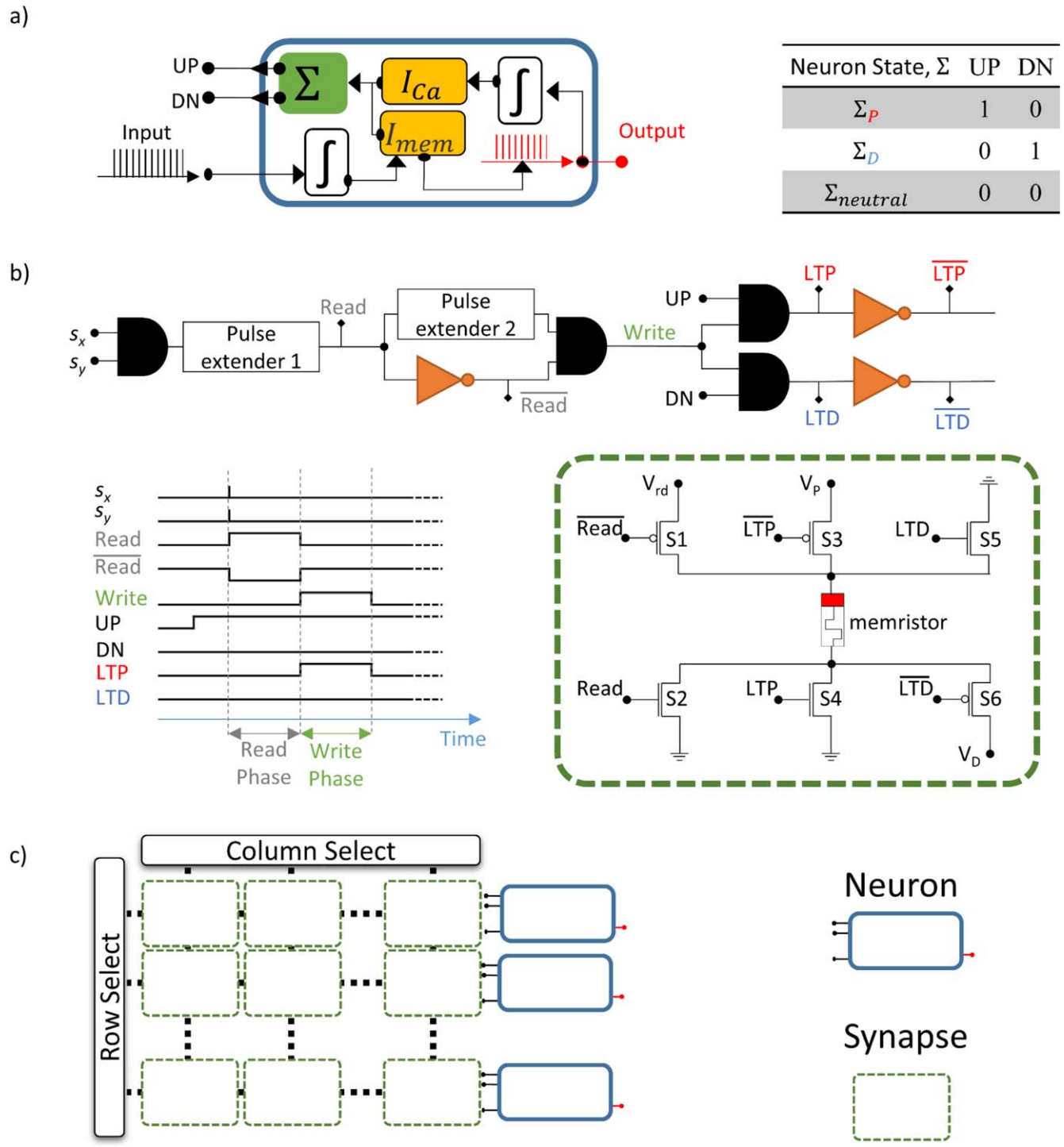
$$\frac{dI_{Ca}}{dt} = -\frac{1}{\tau_C} + J_C \sum_i \delta(t - t_i), \tag{3}$$

where τ_C is the calcium time constant and $\delta(t - t_i)$ is a spike occurring at time t_i at the output terminal of the ON, which increases I_{Ca} by J_C. The neuron can settle into potentiation (Σ_P), depression (Σ_D) or neutral ($\Sigma_{neutral}$) states depending on the values of its two internal variables according to equation (4): [25]

$$\begin{aligned} &\Sigma_P\text{: } I_{mem} > \theta_{mem} \quad and \quad I_{Ca} \in [\theta_{\mathrm{LTP},low}, \theta_{\mathrm{LTP},high}] \\ &\Sigma_D\text{: } I_{mem} < \theta_{mem} \quad and \quad I_{Ca} \in [\theta_{\mathrm{LTD},low}, \theta_{\mathrm{LTD},high}] \\ &\Sigma_{neutral}\text{: otherwise,} \end{aligned} \tag{4}$$

where $\theta_{mem} < \theta_{fire}$ is a threshold parameter; $\theta_{\{LTP,LTD\},high}$ and $\theta_{\{LTP,LTD\},low}$ are the low and high boundaries defining the intervals for specific Σ_P and Σ_D neuron states. Representative evolution of the neuron activity over time can be found in section 2 of the supplementary material.

The neuron state ($\Sigma_{\{P,D,neutral\}}$) is coded by two digital signals, UP and DN, according to the truth table in figure 2(a). When a spike arrives at the ON input, the current state of the ON is used to generate UP and DN signals that control the ON input terminal voltage leading to one of three outcomes—the synapse is only read, is read and potentiated or is read and depressed, depending on whether the neuron is in $\Sigma_{neutral}$, Σ_P or Σ_D states, respectively. Indeed, the asynchronous circuitry described in figure 2(b) is driven by UP and DN values to generate the signals *Read*, *LTP* and *LTD* and their inverse, $\overline{Read}$, $\overline{LTP}$, $\overline{LTD}$, that control the voltage drop across the memristive device using the circuitry composed of transistors S1–S6. The 6 transistor—1 memristor (6T-1 memristor) block is the elemental synaptic unit that is organised in a pseudo-cross-bar macro-structure, as shown in figure 2(c), which can perform a vector-matrix multiplication in a single step. With reference to the block scheme in figure 2(b), synapses in the pseudo-cross-bar are selected when both row and column signals, s_x and s_y, are high. The *Read* signal is extended for the required duration by the *Pulse extender 1*. The *Read* signal and its inverse turn on transistors S1 and S2 and allow



Neuron State, Σ	UP	DN
Σ_P	1	0
Σ_D	0	1
$\Sigma_{neutral}$	0	0

Figure 2. (a) Conceptual block diagram of the neuron module: input spikes are integrated into the I_{mem} variable (equation (2)), which generates output spikes each time the θ_{fire} is exceeded; the resulting output firing rate is integrated into the I_{Ca} variable (equation (3)); I_{mem} and I_{Ca} define the state of the neuron ($\Sigma_{\{P,D,neutral\}}$) and the UP and DN terminal voltages are adjusted depending on the neuron state Σ according to the truth table. (b) Logic blocks and timing diagram of the neuron-memristor interface and arrangement of the switches (S1–S6) driving the synaptic operation. The circuit comprising 6T and 1 memristor is the synaptic module. Conventional symbols are displayed for AND and NOT ports. (c) Overall system architecture comprising asynchronous row and column controllers, synaptic pseudo-cross-bar macro-structure and the ONs. Neuron and synapse modules are those reported in panel (a) and (b), respectively.

the read voltage V_{rd} to drop across the memristive device. A write phase is initiated when *Read* is low ($\overline{Read}$ is high) and lasts for the duration defined by the *Pulse extender 2*. The write phase manifests as an LTP or an LTD event only if either UP or DN signal is high. In that case, the signals *LTD* and $\overline{LTD}$ or *LTP* and $\overline{LTP}$ turn on transistors couples S5–S6 or S3–S4 that allow the LTD or LTP voltage (V_D or V_P) to drop across the memristive device, respectively. The block

diagram of figure 2(b) realises the truth table shown in figure 2(a) and applies the correct voltage drop on the memristive device in an asynchronous manner.

In summary, the update of the weight of a specific synapse depends on the occurrence of IN spikes that stimulate the neuron. Furthermore, the state of the ONs depends both on its input and on its output firing rate, realising, therefore, a STR-DP learning rule [25]. In many works in the literature, long-lasting overlapping pulses, possibly with complex pulse shapes are used to implement STDP [12, 15, 18, 20, 57] and SRDP [29, 30]. This makes the asynchronous control of large synaptic arrays difficult. In this work, the realisation of the proposed STR-DP learning rule relies on asynchronous adjustments of the voltage drop across the memristive devices. Silicon implementations of neuron units and control blocks implementing such schemes have been demonstrated in earlier works [6, 26, 27, 48].

Learning dynamics and performances

The circuitry building blocks described in the previous section can be used to build the SNN shown in figure 3(a). Neurons are grouped into four families. The INs convert the intensity of the 28 × 28 pixels of a hand-written digit, drawn from the MNIST data-set [47], into a firing rate that stimulates trainable synapses (empty dashed squares in figure 3(a)). The 10 ONs are responsible for the classification of the digits (1 ON for each class to be recognised): the one that fires most is the winner and decides the response of the SNN (see section 3 in the supplementary material for an alternative voting procedure). The teacher neurons send spikes through fixed synapses to the ON that is designated to recognise a specific digit. Finally, the inhibitory neurons (IhNs) are connected to INs and to ONs through fixed synapses. When stimulated by the INs, the IhNs feed the ONs with a negative current, reducing the overall signal at their input. In MNIST data-set, there are pixels with extremely high and low intensities in the same locations of the images belonging to different digit classes (e.g. the corners of an image correspond to the low intensity background for any digit class). These pixels cause a high and a low firing rate of the same INs even though they are not representative of some digit features and cause a high unbalance in the firing rates of the ONs. The reduction of the input signal into the ONs due to the inhibitory connection serves to level out ONs firing rate [58]. In the SNN, all the synapses are randomly initialised (see section 1 in the supplementary material for details) and the network parameters are optimised according to the guidelines reported in sections 4 and 5 in the supplementary material.

The SNN architecture realises a semi-supervised learning scheme that needs an initial labelling of the training set. We simulate the network with different models for the trainable synapses: nonlinear soft-bound weight update, as well as linear synapses with hard weight boundaries (figure 3(b)). As discussed above, synaptic units comprise 6T-1 memristor or a 6T-1 linear element, which will be referred to as memristive synapse and linear synapse, respectively, in the following. Linear synapses are simulated with resolutions of 10 levels ($\delta w_{pulse} \approx 0.1$), 100 levels ($\delta w_{pulse} \approx 0.01$) and 1000 levels ($\delta w_{pulse} \approx 0.001$) with $w \in [0, 1]$, and compared against the memristive device. Linear synapses with constant weight update δw_{pulse} are believed to be the elements of choice of hardware deep neural networks [35, 51, 59] but no compact non-volatile electronic device is able to easily reproduce the linear conductance evolution. For instance, the realisation of a linear conductance update by means of memristor technology still appears to be challenging [9, 35, 52, 60]. As discussed above, the memristive device is roughly estimated to cover a number of levels of 150–200 ($\approx 1/\alpha_{\{P,D\}}$) and, therefore, it compares well to the linear synapse with 100 levels ($\delta w_{pulse} \approx 0.01$) in terms of resolution. By comparing results obtained with memristive and linear synapses, we will thus be able to disentangle the effects of the weight update and the weight resolution on the network dynamics and performances. The SNN is trained and tested with two distinct MNIST sets to probe the SNN ability to generalise from the digits learnt during training and classify new ones during tests.

The digit recognition rate in our simulations as a function of the training epochs is reported in figure 3(c). The horizontal band represents the recognition rate ($10\% \pm 1\sigma$) of a randomly initialised untrained synaptic matrix. Linear synapses with $\delta w_{pulse} \approx 0.1$ (squares) give a recognition rate that is only slightly better than that of an untrained synaptic matrix. A recognition rate of ≈52% is reached after a few hundred training epochs by using linear synapses with $\delta w_{pulse} \approx 0.01$ (triangles). The SNN employing synapses with $\delta w_{pulse} \approx 0.001$ (rhombi) displays a slow increase of the recognition rate up to ≈55% after the presentation of 2000 training images. Note that, in this case, the training does not saturate with 2000 image samples. In summary, the increase of the synaptic resolution of linear synapses improves the SNN recognition rate and slows down the learning process, which is an indication of the increase of the memory capacity of the SNN and ensures a correspondingly slow memory forgetting. The weight maps after 2000 training images are reported in figure 3(d). For the $\delta w_{pulse} \approx 0.1$ case, the map shows not well resolved digits and further training is not expected to bring any improvement, because the recognition rate does not show any increasing trend in figure 3(c). The digits are clearly distinguishable for the $\delta w_{pulse} \approx 0.01$ case, while they are still quite unclear for the $\delta w_{pulse} \approx 0.001$ case because the training has not saturated [20, 57].

The SNN employing memristive devices shows a peculiar behaviour featuring a recognition rate of ≈30% in combination with a slow learning dynamics. Specifically, in comparison to the SNN with $\delta w_{pulse} \approx 0.01$ synapses, the memristor-based SNN exhibits a lower recognition rate and a slower learning rate. Interestingly, the digit patterns in the weight maps at the end of the training are well defined even though the weight values of the synapses storing the information of the digit features do not extend over the entire [0, 1] range, as shown in the colour-bar on the right side of figure 3(d). On the contrary, the weights of the linear synapses cover the full [0, 1] range, as discussed in more detail in section 6 of the supplementary material. In summary,

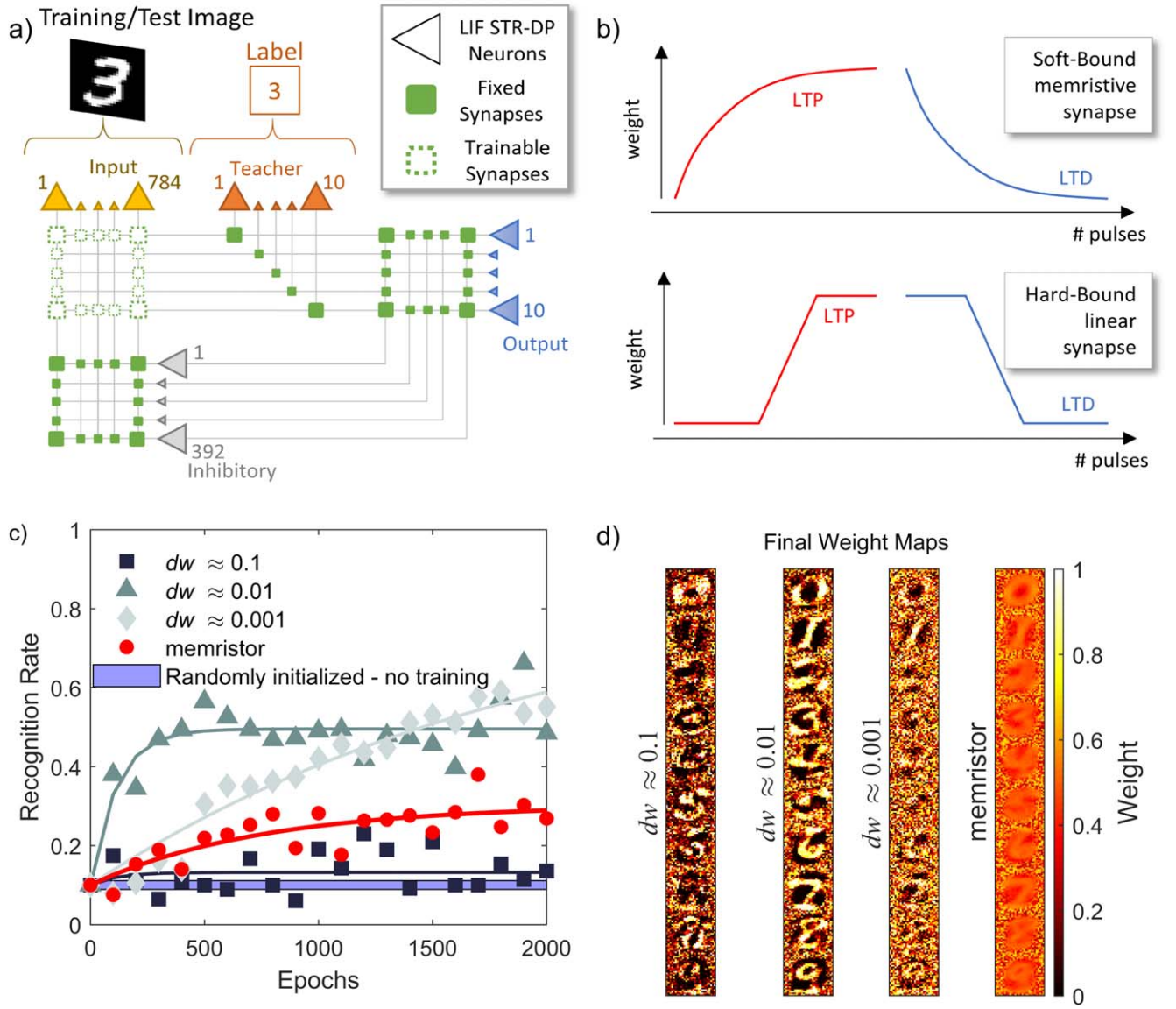


Figure 3. (a) Functional architecture of the simulated SNN, including 28 × 28 input, 10 teacher, 392 inhibitory and 10 output LIF neurons; adjustable synaptic matrix (empty dashed squares) and fixed synapses with random weights (filled squares). (b) Sketch of the simulated synapses: with hard bounds and linear update and with soft bounds (memristive device). (c) Evolution of the simulated recognition rate as a function of the training epochs for the investigated synaptic models: squares refer to linear synapses with $dw_{pulse} \approx 0.1$, triangles refer to linear synapses with $dw_{pulse} \approx 0.01$, rhombi to linear synapses with $dw_{pulse} \approx 0.001$, circles refer to memristive synapses and the continuous band corresponds to the recognition rate of an SNN with randomly initialised and untrained synaptic matrix (i.e. 10% ± 1σ). Lines correspond to exponential fitting of the simulated results. (d) Final weight maps for all the simulated synaptic models.

it appears that the learning rate is slowed down when using memristive synapses because of their slow update dynamics and their soft-bound behaviour. On the other hand, the permanence the synaptic weights in the middle of the [0, 1] range limits the sharpness of the learnt digit features and, consequently, the SNN recognition rate is low.

According to literature [20, 57], increasing the number of ONs for each class and the corresponding number of synapses in an all-to-all connection scheme is a simple way to improve the network performance without complicating its structure. Simulations carried out with 10 neurons for each digit class (100 in total) results in a significant increase in performance. Figure 4 shows the recognition rate evolution for the SNNs employing linear synapses with $\delta w_{pulse} \approx 0.01$, with $\delta w_{pulse} \approx 0.001$ (a) and memristive devices with and without variability (b). In this case, the training stage and the recognition rate is evaluated up to 20 000 image presentations to assess the stability of learning. The maximum value of the recognition rate is achieved with linear synapses with 1000 levels. Interestingly, the recognition rate of the SNN employing memristive synapses is slightly larger than that obtained with linear synapses with 100 levels, i.e. the results is inverted with respect to the case of the simulations up to 2000 training images and only one ON per class (10 in total). It is worth noticing that during long training sessions in which new images are continuously presented and the network learns their significant features, the effect of saturation of the storage capacity starts to play a role and the effect of memory

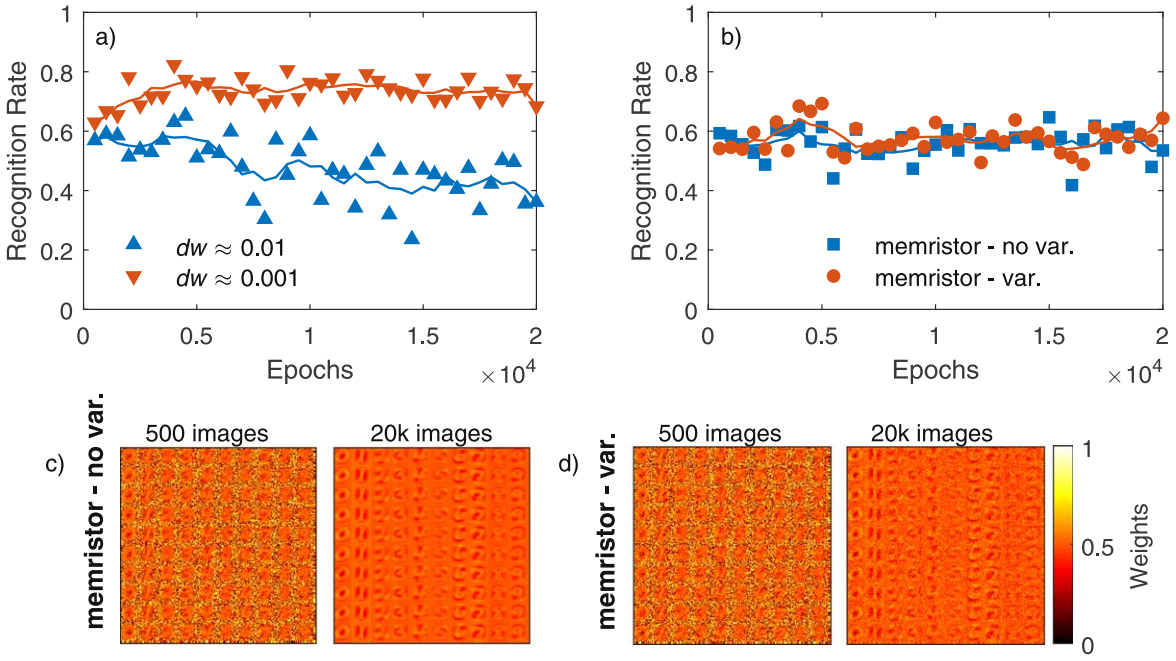


Figure 4. Recognition rate as a function of the training epochs for the SNN employing 10 ONs per class for linear synaptic models (a) and memristive device including and not including variability (b). Lines correspond to smoothed data and serve as a guide for the eyes. Memristive synaptic maps after 500 image presentations and at the end of the learning after 20 000 image presentations both neglecting (c) and including (d) variability.

loss affects the recognition rate. In the SNN with 100 evenly spaced synaptic levels ($\delta w_{pulse} \approx 0.01$), the recognition rate decreases as a function of the training epochs and, after about 5000 epochs, is subjected to large variability as compared to the other recognition rate trends reported in figures 4(a) and (b). Both these findings can be interpreted as consequences of the saturation of the SNN storage capacity. On the one hand, the continuous replacement of memories, occurring after storage capacity saturation, makes the recognition rate very sensitive to how much a testing image is similar to the training images shown recently or a long time ago. On the other side, memory replacement produces a progressive degradation of the stored digit patterns leading to recognition rate decrease. Conversely, memristive synapses with comparable resolution result in a higher and more stable recognition rate (figure 4(b)), which can be ascribed to an increased memory capacity, in agreement with theoretical predictions [1].

It is worth noticing that the training is sampled every 500 epochs, which does not allow for capturing the initial increasing trend of the for a network with 10 ONs per class, except for the case of linear synapses with $\delta w_{pulse} \approx 0.001$. Furthermore, we cannot expect a precise correspondence of the initial training dynamics of the SNN with 1 and 10 ONs per class, respectively in figures 3(c) and 4. The overall performances are in line with the literature results dealing with unsupervised or semi-supervised learning and employing networks with the same number of adjustable synaptic elements based on memristive devices [20, 28, 57].

The recognition rate as a function of the training epochs, as shown in figure 4(b), settles on the same value of ≈60% for both the cases of memristive devices without (squares) and with the additive variability contribution (circles) as described with reference to figure 1. This result demonstrates the robustness of the network against synaptic variability, in agreement with literature reports [15, 17, 53, 57, 59–61]. The evolution of the synaptic matrix is captured after 500 and 20 000 training epochs in figures 4(c) and (d), for memristive synapses both without and with variability, respectively. The maps after 500 epochs are equivalent, while after 20 000 epochs, memristive device variability produces a noisy synaptic map, that produces the same recognition rate as the map corresponding to memristive device without variability. Moreover, while after 500 epochs the digits start to appear over a noisy background, after 20 000 epochs the ten digits appear well defined over a uniform background lying around a value of 0.5. The concentration of the background weight values around the median value of the entire range distinguishes the operation of the present network from those presented in the literature, whose background is pushed to zero [18, 20, 28, 57]. The source of this difference will become clear in the following.

Discussion

In the present paper, we have described the effect of the synaptic update dynamics on the SNN learning performances. More specifically, we have found that the soft-bound synaptic weight update featured by the memristive device guarantees slow learning (and consequently slow forgetting), which results in improved recognition performance with respect to linear synaptic update with comparable weight resolution.

The investigated SNN utilises a semi-supervised learning mechanism with stop-learning conditions that endow the network with online life-long learning features. One novel aspect of this work is represented by the simulation of neural and learning circuits that have already been fabricated in silicon but never tested on high level simulations in combination with memristive devices. Indeed, most of the recent studies deals with proof-of-principle investigations of theoretical networks without clearly addressing the issue of feasibility in VLSI silicon technology [14, 15, 17, 20, 28, 57]. The present SNN is based on an event-driven STR-DP rule as discussed with reference to figure 2. The STR-DP rule is a generalisation of the class of plasticity rules based only on timing and ensures an improved sensitivity to spatio-temporal firing correlations [32–34].

Memristive devices exhibit gradual conductance evolution in both LTP and LTD operations when stimulated by sequences of identical spikes as reported in figure 1. Memristor conductance dynamics displays unavoidable soft approach to extreme conductance levels, which contributes to a weight-dependent weight update and to a multiplicative network dynamics. Indeed, in case of weight-independent (additive) plasticity, as in the case of linear synapses, an initial synapse reinforcement provokes the post neuron to fire more and more, which usually causes further synaptic potentiation up to the boundary weight value. Such mechanism establishes a positive feedback which gives rise to a strong synapse specialisation and finally results in a bi-modal distribution of the synaptic weights [20, 61–63], as confirmed also by our simulations reported in section 6 of the supplementary material. Contrarily, weight-dependent (multiplicative) plasticity, according to which, e.g., strong synapses are weakly potentiated, leads to stable dynamics which causes the concentration of the weight values in the middle of the available range according to a uni-modal distribution [20, 61–63]. This kind of weight distribution is also in agreement with biological observations [63]. In order to exploit the strong tendency to specialisation, linear conductance update is usually desired and obtained by artificially keeping the memristor operation within its linear regime [13], which goes at the expenses of the width of the available conductance window (G_{max}/G_{min}) [51, 52]. Conversely, weight-dependent plasticity is characterised by a reduced tendency to specialisation that can be switched on either by lateral inhibition of the ONs or by activity-dependent plasticity [63]. The first route has already been run by some proof-of-principle SNNs where spike timing is given by overlapping pulses and in which Winner-Takes-All and, possibly, homeostatic mechanisms are included [15–17, 20, 57]. On the contrary, the present approach exploits activity-dependent plasticity of the ONs through the STR-DP rule to develop a specialisation in the synaptic matrix [63], as detailed in the following.

The dynamics of synaptic specialisation of the memristive SNN constituted by 10 ONs per class is analysed with reference to figure 5. The synapses are divided into three categories: one includes the synapses associated mainly to the background; one, the synapses associated mainly to the pattern and the third, the synapses not undoubtedly associated to either of the previous categories (namely, mixed category). To group the synapses, we consider the normalised average intensity of each pixel over the entire MNIST data-set (the result is shown in figure 5(a)). Then, we arbitrarily define two thresholds at 0.10 and at 0.25. The synapses connected to the pixel (i.e. INs) whose intensity is above (below) the higher (lower) threshold belong to the *pattern* (*background*) category, whereas the remaining synapses are classified as belonging to the mixed category. INs associated to high (low) intensity pixels stimulate the afferent synapses with a high (low) firing rate. Therefore, for brevity, *pattern* and *background* synapses are also named high and low firing rate (HFR and LFR) synapses, respectively, as reported in figures 5(a) and (b). The different stimulation rates of HFR and LFR synapses result in different weight dynamics, as shown in figures 5(c) and (d). Indeed, HFR synapses almost reach their asymptotic values after 5000 epochs, as visible in figure 5(c). In comparison, LFR synapses undergo a slower dynamics that needs almost the entire training process to reach the asymptotic values (figure 5(d)). Also the asymptotic configuration of HFR and LFR synapses values is quite different (figure 5(e)): the final LFR synaptic weights display a purely uni-modal distribution with 0.5 average value, whereas the HFR synaptic weight distribution extends to lower and higher values and roughly develops two modes.

The weights of the LFR synapses develop a uni-modal distribution as a consequence of the weight-dependent plasticity of the memristive device and of the scarce correlation between the low-rate firing of the INs and the ONs. Conversely, HFR (i.e. *pattern*) synapses are subjected to stronger specialisation because the sensitivity of the STR-DP to the spatio-temporal correlation leads to the development of a bi-modal distribution, useful for the recognition task.

As noticed above, the character of plasticity, being it multiplicative or additive, drives opposite weight dynamics during the network training phase. Indeed, the results of the simulation employing memristive synapses are compared to those that involve linear synapses. In particular, memristive synapses can be thought to accommodate a resolution of roughly one hundred levels. From figure 3(c), which refers to simulations with only 1 ON per class, it is evident that the memristive matrix guarantees a lower recognition rate than the one obtained using linear synapses with almost comparable resolution ($\delta w_{pulse} \approx 0.01$, 100 levels). Conversely, memristive synapses require a longer learning time than linear synapses with comparable resolution. Indeed, recognition rate and learning time are influenced by both synaptic resolution and weight dynamics, though in a slightly different way. Obviously, increasing the number of levels reduces the learning speed and improves the recognition rate, in principle. The transition from constant to weight-dependent conductance change per pulse, i.e. from additive to multiplicative plasticity, has the positive consequence of slowing the learning down, but it might have a negative impact on the recognition rate. As shown in figure 3(c), linear synapses with 100 levels result in a recognition rate higher than memristive synapses on a relatively short training stage of 2000 epochs,

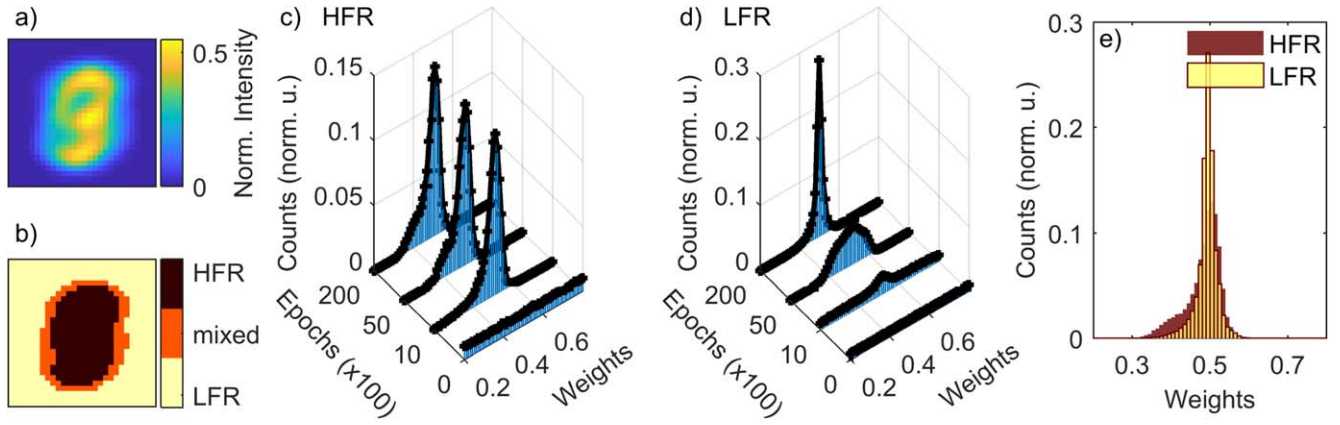


Figure 5. (a) Pixel-by-pixel average of the MNIST digits. (b) Identification of the pixel contributing to low (LFR), high (HFR) and mixed firing rates of the INs. (c) Evolution of the histogram of the synaptic weights in the HFR region. (d) Evolution of the histogram of the synaptic weights in the LFR region. (e) Histograms of the synaptic weights at the end of the training for the HFR and LFR regions.

because they naturally develop a bi-modal weight distribution after training [64], which produces a high signal to noise ratio of ONs firing rate in response to learned (or unlearned) patterns. In comparison, the final weight distribution of memristive synapses displaying a uni-modal character results in a lower signal to noise ratio. However, in the long run, slow learning of soft-bound synapses is more robust than fast learning with linear synapses. Indeed, the memristor-based SNN ensures improved performance with respect to one based on linear synapses with roughly the same resolution of 100 levels after a 20 000 epochs training, as shown in figure 4 for the SNN with 10 ON per class. This result confirms the enhanced memory capacity of soft-bound synapses with respect to linear hard bound ones with comparable number of levels [1]. The increase of synaptic resolution, even in case of linear synapses, further improves the network performance, as attested by the results shown in figure 4(a), where 1000-level ($\delta_{pulse} \approx 0.001$) linear synapses are used. However, state of the art solutions for memristive devices still appear far from those high resolution values. For instance, [51, 52, 65] report material engineering routes and conductance evolution throughout only tens to few hundreds of pulses. In particular, Park *et al* [65] demonstrate 64 conductance levels measured upon 30-cycle endurance tests. Stathopoulos *et al* [66] exploit the extreme low noise operation of their bi-layer device to programme through a write-and-verify scheme almost 100 states with stable resistance values up to 8 h at room temperature. Therefore, synaptic resolutions running through thousands of levels are still long from being realised with only one memristive synaptic element. However, the conductance-dependent memristor dynamics helps in increasing the network performance to values that are not achievable with linear synapses at similar resolutions.

In this paper, we propose a 6T-1 memristor elemental synapse architecture that drives the correct voltage drop on the memristive element for programming and reading operations. The adjustment of the voltage occurs in response to spikes as short as few μs [26] and in agreement with a learning rule that is sensitive to real-time events, i.e. to events occurring with an average period in the ms range (a rate of tens-to-hundreds of Hz) [25]. In the literature, in order to implement a learning rule in SNNs, the spikes controlling the memristive devices are extended above the ms range so that their temporal overlap ensures the sensitivity to real-time events [11, 12, 15, 18, 20, 57]. However, in this latter case, the occurrence of many nearly-simultaneous input events can raise the voltage of some lines of a synaptic array thus worsening the possible sneak path issue, even in presence of selector devices. In this work, the use of a relatively large 6T-1 memristor elemental synapse-block and of short spikes ensures an improved control of a large synaptic array. The sensitivity to correlation over time-scales in the ms range is obtained through the integration of the spike rates into the neuron state variables through capacitors. The sub-threshold operative regime of the CMOS circuits implementing the described neuron functionality limits the charging and discharging currents within the range of the femto- to nano-Ampéres, which allows the use of relatively small capacitors compatible with a VLSI approach [67]. Therefore, the proposed SNNs is devised to allow a relatively easy implementation in the current available technology.

In summary, from this work, it becomes clear that the joint optimisation of the memristor nonlinear conductance dynamics and the network parameters is crucial to reach the best possible trade-off between slow learning and high signal to noise ratio and to maximise the recognition rate. In particular, memristive devices should be developed in order to reduce the value of the parameters $\alpha_{P,D}$, while maintaining the parameters $\gamma_{P,D}$ to low values [1]. The network parameters, conversely, should be adjusted to increase the competition among the output neurons that promote the specialisation of the network and the opening of a bi-modal distribution of the synapses in the HFR region.

Conclusions

We simulate an SNN based on memristive synapses, mixed-signal analogue-digital neurons implementing a STR-DP rule, arranged in an asynchronous SNN with feasible

implementation in silicon technology. We compare the training dynamics of an SNN comprising soft-bound nonlinear memristive synapses with one employing ideal linear synapses. We demonstrate that the former SNN performs better than the one employing linear synapses with comparable resolution as a consequence of a slower learning dynamics and an improved memory capacity. The work, thus, provides a pathway for a holistic optimisation of future hardware neural SNNs based on memristive devices.

Acknowledgments

The work is partially supported by the European project H2020-ICT-2015 NEUral computing aRchitectures in Advanced Monolithic 3D-VLSI nanotechnologies (NEURAM^3, grant agreement No. 687299).

Author Contributions

All authors discussed the results and implications and contributed to the writing of the manuscript. SB, EC, GI and SS conceived the idea and designed the experiments. JF acquired electrical data with EC's contribution. JF analysed memristor data with SB's contribution. EC developed the experimental setup for memristor characterisation. DC and GI developed the code for the SNN simulations. DC performed the SNN simulations. DC and SB analysed the SNN data. MN and GI conceived the neuron's building blocks and neuron/synapse interface. GI and SS supervised the research.

ORCID iDs

S Brivio https://orcid.org/0000-0003-2386-7953
M V Nair https://orcid.org/0000-0002-0182-8358
J Frascaroli https://orcid.org/0000-0002-3204-0811
E Covi https://orcid.org/0000-0003-0479-6897
C Ricciardi https://orcid.org/0000-0002-4703-7949
G Indiveri https://orcid.org/0000-0002-7109-1689
S Spiga https://orcid.org/0000-0001-7293-7503

References

[1] Fusi S and Abbott L F 2007 *Nat. Neurosci.* **10** 485
[2] Indiveri G and Liu S C 2015 *Proc. IEEE* **103** 1379–97
[3] Jeong D S, Kim K M, Kim S, Choi B J and Hwang C S 2016 *Adv. Electron. Mater.* **2** 1600090
[4] Merolla P A *et al* 2014 *Science* **345** 668–73
[5] Furber S B, Galluppi F, Temple S and Plana L A 2014 *Proc. IEEE* **102** 652–65
[6] Qiao N, Mostafa H, Corradi F, Osswald M, Stefanini F, Sumislawska D and Indiveri G 2015 *Frontiers Neurosci.* **9** 141
[7] Parisi G 1986 *J. Phys. A: Math. Gen.* **19** L617
[8] Kirkpatrick J *et al* 2017 *Proc. Natl Acad. Sci.* **114** 3521–6
[9] Frascaroli J, Brivio S, Covi E and Spiga S 2018 *Sci. Rep.* **8** 7178
[10] Suri M, Querlioz D, Bichler O, Palma G, Vianello E, Vuillaume D, Gamrat C and DeSalvo B 2013 *IEEE Trans. Electron Devices* **60** 2402–9
[11] Garbin D, Vianello E, Bichler O, Rafhay Q, Gamrat C, Ghibaudo G, DeSalvo B and Perniola L 2015 *IEEE Trans. Electron Devices* **62** 2494–501
[12] Ambrogio S, Balatti S, Milo V, Carboni R, Wang Z Q, Calderoni A, Ramaswamy N and Ielmini D 2016 *IEEE Trans. Electron Devices* **63** 1508–15
[13] Yao P, Wu H, Gao B, S B Eryilmaz X H, Zhang W, Zhang Q, Deng N, Shi L, Wong H S P and Quian H 2017 *Nat. Commun.* **8** 15199
[14] Serb A, Bill J, Khiat A, Berdan R, Legenstein R and Prodromakis T 2016 *Nat. Commun.* **7** 1–9
[15] Covi E, Brivio S, Serb A, Prodromakis T, Fanciulli M and Spiga S 2016 *Frontiers Neurosci.* **10** 482
[16] Covi E, Brivio S, Serb A, Prodromakis T, Fanciulli M and Spiga S 2016 *2016 IEEE Int. Symp. on Circuits and Systems (ISCAS)* pp 393–6
[17] Covi E, Brivio S, Frascaroli J, Fanciulli M and Spiga S 2017 *ECS Trans.* **75** 85
[18] Yu S, Gao B, Fang Z, Yu H, Kang J and Wong H S P 2013 *Frontiers Neurosci.* **7** 186
[19] Eryilmaz S B, Kuzum D, Jeyasingh R, Kim S, BrightSky M, Lam C and Wong H S P 2014 *Frontiers Neurosci.* **8** 1–11
[20] Bill J and Legenstein R 2014 *Frontiers Neurosci.* **8** 412
[21] Pedretti G, Milo V, Ambrogio S, Carboni R, Bianchi S, Calderoni A, Ramaswamy N, Spinelli A S and Ielmini D 2017 *Sci. Rep.* **7** 1–10
[22] Park S, Chu M, Kim J, Noh J, Jeon M, Hun Lee B, Hwang H, Lee B and Lee B G 2015 *Sci. Rep.* **5** 10123
[23] Choi S, Shin J H, Lee J, Sheridan P and Lu W D 2017 *Nano Lett.* **17** 3113–8
[24] Bi G Q and Poo M M 1998 *J. Neurosci.* **18** 10464–72
[25] Brader M, Senn W and Fusi S 2007 *Neural Comput.* **19** 2881–912
[26] Mostafa H, Mayr C and Indiveri G 2016 *2016 IEEE Int. Symp. on Circuits and Systems (ISCAS)* pp 926–9
[27] Covi E, George R, Frascaroli J, Brivio S, Mayr C, Mostafa H, Indiveri G and Spiga S 2018 *J. Phys. D: Appl. Phys.* **51** 344003
[28] Hansen M, Zahari F, Ziegler M and Kohlstedt H 2017 *Frontiers Neurosci.* **11** 91
[29] He W, Huang K, Ning N, Ramanathan K, Li G, Jiang Y, Sze J, Shi L, Zhao R and Pei J 2015 *Sci. Rep.* **4** 4755
[30] Milo V, Pedretti G, Carboni R, Calderoni A, Ramaswamy N, Ambrogio S and Ielmini D 2016 *2016 IEEE Int. Electron Devices Meeting (IEDM)* pp 16.8.1–4
[31] Sjöström P J, Turrigiano G G and Nelson S B 2001 *Neuron* **32** 1149–64
[32] Gjorgjieva J, Clopath C, Audet J and Pfister J P 2011 *Proc. Natl Acad. Sci.* **108** 19383–8
[33] Pfister J P and Gerstner W 2006 *Advances in Neural Information Processing Systems* vol 18 (Cambridge: MIT Press) pp 1083–90
[34] Morrison A, Aertsen A and Diesmann M 2007 *Neural Comput.* **19** 1437
[35] Ambrogio S *et al* 2018 *Nature* **558** 60–7
[36] Gokmen T and Vlasov Y 2016 *Frontiers Neurosci.* **33** 1–13
[37] Kadetotad D, Xu Z, Mohanty A, Chen P Y, Lin B, Ye J, Vrudhula S, Yu S, Cao Y and Seo s J 2015 *IEEE J. Emerg. Sel. Top. Circuits Syst.* **5** 194–204
[38] Brivio S, Tallarida G, Cianci E and Spiga S 2014 *Nanotechnology* **25** 385705
[39] Brivio S, Frascaroli J and Spiga S 2015 *Appl. Phys. Lett.* **107** 023504

[40] Brivio S, Covi E, Serb A, Prodromakis T, Fanciulli M and Spiga S 2016 *Appl. Phys. Lett.* **109** 133504
[41] Covi E, Brivio S, Fanciulli M and Spiga S 2015 *Microelectron. Eng.* **147** 41–4
[42] Chen H Y *et al* 2017 *J. Electroceram.* **39** 21–38
[43] Frascaroli J, Volpe F G, Brivio S and Spiga S 2015 *Microelectron. Eng.* **147** 104–7
[44] Frascaroli J, Brivio S, Ferrarese Lupi F, Seguini G, Boarino L, Perego M and Spiga S 2015 *ACS Nano* **9** 2518–29
[45] Brivio S, Frascaroli J and Spiga S 2017 *Nanotechnology* **28** 395202
[46] Brivio S, Covi E, Serb A, Prodromakis T, Fanciulli M and Spiga S 2015 *Proc. IEEE Int. Conf. on Memristive Systems (MEMRISYS)* pp 1–2
[47] Lecun Y, Bottou L, Bengio Y and Haffner P 1998 *Proc. IEEE* **86** 2278–324
[48] Bartolozzi C and Indiveri G 2007 *Neural Comput.* **19** 2581–603
[49] Goodman D and Brette R 2009 *Frontiers Neurosci.* **3** 26
[50] Stimberg M, Goodman D, Benichoux V and Brette R 2014 *Frontiers Neuroinf.* **8** 6
[51] Jang J W, Park S, Burr G W, Hwang H and Jeong Y H 2015 *IEEE Electron Device Lett.* **36** 457–9
[52] Wang Z, Yin M, Zhang T, Cai Y, Wang Y, Yang Y and Huang R 2016 *Nanoscale* **8** 14015–22
[53] Kim S, Lim M, Kim Y, Kim H D and Choi S J 2018 *Sci. Rep.* **8** 2638
[54] Brivio S and Spiga S 2017 *J. Comput. Electronic.* **16** 1154–66
[55] Ambrogio S, Balatti S, Cubeta A, Calderoni A, Ramaswarmy N and Ielmini D 2013 *2013 IEEE Int. Electron Devices Meeting (IEDM)* 31.5 (https://doi.org/10.1109/IEDM.2013.6724732)
[56] Gaba S, Sheridan P, Zhou J, Choi S and Lu W 2013 *Nanoscale* **5** 5872–8
[57] Querlioz D, Bichler O, Vincent A F and Gamrat C 2015 *Proc. IEEE* **103** 1398–416
[58] Chicca E, Stefanini F, Bartolozzi C and Indiveri G 2014 *Proc. IEEE* **102** 1367–88
[59] Chen P Y, Lin B, Wang I T, Hou T H, Ye J, Vrudhula S, Seo s J, Cao Y and Yu S 2015 *2015 IEEE/ACM Int. Conf. on Computer-Aided Design (ICCAD)* pp 194–9
[60] Yu S, Chen P Y, Cao Y, Xia L, Wang Y and Wu H 2015 *2015 IEEE Int. Electron Devices Meeting (IEDM)* pp 17.3.1–4
[61] Agarwal S, Plimpton S J, Hughart D R, Hsia A H, Richter I, Cox J A, James C D and Marinella M J 2016 *2016 Int. Joint Conf. on Neural Networks (IJCNN)* pp 929–38
[62] Rubin J, Lee D D and Sompolinsky H 2001 *Phys. Rev. Lett.* **86** 364–7
[63] van Rossum M C W, Bi G Q and Turrigiano G G 2000 *J. Neurosci.* **20** 8812–21
[64] Billings G and van Rossum M C W 2009 *J. Neurophysiol.* **101** 2775–88
[65] Park J, Kwak M, Moon K, Woo J, Lee D and Hwang H 2016 *IEEE Electron Device Lett.* **37** 1559–62
[66] Stathopoulos S, Khiat A, Trapatseli M, Cortese S, Serb A, Valov I and Prodromakis T 2017 *Sci. Rep.* **7** 17532
[67] Indiveri G, Linares-Barranco B, Legenstein R, Deligeorgis G and Prodromakis T 2013 *Nanotechnology* **24** 384010