This is the author-created version of the following work:


Access to this file is available from:
https://researchonline.jcu.edu.au/46904/

Please refer to the original source for the final version of this work:
http://dx.doi.org/10.1109/TBCAS.2016.2618351
A Hybrid CMOS-memristor Neuromorphic Synapse

Mostafa Rahimi Azghadi, Member, IEEE, Bernabe Linares-Barranco, Fellow, IEEE, Derek Abbott, Fellow, IEEE, Philip H.W. Leong, Senior Member, IEEE

Abstract—Although data processing technology continues to advance at an astonishing rate, computers with brain-like processing capabilities still elude us. It is envisioned that such computers may be achieved by the fusion of neuroscience and nano-electronics to realize a brain-inspired platform. This paper proposes a high-performance nano-scale Complementary Metal Oxide Semiconductor (CMOS)-memristive circuit, which mimics a number of essential learning properties of biological synapses. The proposed synaptic circuit that is composed of memristors and CMOS transistors, alters its memristance in response to timing differences among its pre- and post-synaptic action potentials, giving rise to a family of Spike Timing Dependent Plasticity (STDP). The presented design advances preceding memristive synapse designs with regards to the ability to replicate essential behaviours characterised in a number of electrophysiological experiments performed in the animal brain, which involve higher order spike interactions. Furthermore, the proposed hybrid device CMOS area is estimated as 600 $\mu m^2$ in a 0.35 $\mu m$ process—this represents a factor of ten reduction in area with respect to prior CMOS art. The new design is integrated with silicon neurons in a crossbar array structure amenable to large-scale neuromorphic architectures and may pave the way for future neuromorphic systems with spike timing-dependent learning features. These systems are emerging for deployment in various applications ranging from basic neuroscience research, to pattern recognition, to Brain-Machine-Interfaces.

Index Terms—Neuromorphic, Synaptic Plasticity, Learning, Memristor, Crossbar, STDP, Triplet, Quadruplet.

I. INTRODUCTION

MEMRISTORS, due to their special features including non-volatility, nanoscale dimensions, low power consumption, and the ability to be programmed while operating [1], have attracted attention for implementing an in-situ architecture [2], [3], [4]. These emerging nanoscale devices can implement and mimic the synaptic plasticity characteristics of well-known learning algorithms such as pair-based STDP and Spike Rate-Dependent Plasticity (SRDP) [5], [6], [7], [8]. Attempts have also been made to mimic experimental outcomes of higher order spike-based synaptic plasticity rules such as the suppressive STDP rule of Froemke and Dan [9] or Local Correlation Plasticity (LCP) rules to reproduce higher order synaptic plasticity in memristors [10], [11]. In order to advance our understanding of the fundamental properties of synapses and their role in large-scale learning, there is still a need to implement a versatile memristive synapse that is capable of faithfully reproducing a larger regime of experimental data that takes into account conventional STDP [12], frequency-dependent STDP [13], triplet [14], [15] and quadruplet [15], [16] plasticity experiments. In a recent study, Wei et al. replicated the outcome of a variety of synaptic plasticity experiments including STDP, frequency-dependent STDP, triplet, and quadruplet spike interactions, using a TiO$_2$ memristor [17].

This paper proposes a new hybrid CMOS-memristive circuit that aims to emulate all the aforementioned experimental data, with minimal errors close to those reported in a phenomenological model of Triplet STDP (TSTDP) rule presented in [15]. Similar to many previous studies that devised memristive synaptic devices/circuits with STDP, SRDP, or other synaptic properties, our aim is a circuit that implements the TSTDP learning algorithm of [15]. To the best of our knowledge, this has not been previously achieved using memristors. The proposed TSTDP memristive circuit advances the synaptic capabilities of previous designs to be more biologically realistic, and promotes our understanding of synaptic alteration mechanisms, believed to play a key role in learning and memory. Furthermore, the proposed design significantly decreases the silicon real estate required for implementing and utilizing a variety of learning rules.

Spiking neural networks with memristive synapses incorporating the proposed compact and biologically plausible triplet learning circuits, will be an important contribution to the neuroscience research, where a more faithful synaptic plasticity rule, compared to traditional STDP, can be implemented and simulated in a large-scale network. An interesting feature of memristive synapses that distinguishes them from their traditional pure Complementary Metal Oxide Semiconductor (CMOS) counterparts is the feasibility of arranging them in a dense crossbar structure [1] integrated with CMOS circuitry. We also show how the proposed CMOS-memristive circuit can be used in this fashion, to facilitate large-scale integration.

In order to promote reproducible research, Matlab and Cadence files to generate the experimental data and reproduce the results in this paper are made publicly available through Github.\(^1\)

II. MEMRISTIVE SYNAPSE WITH SPIKE TIMING DEPENDENT PLASTICITY (STDP)

Spike Timing Dependent Plasticity (STDP) is a well-known synaptic plasticity rule that modifies the synaptic weight according to the exact timing relationship of pre-
and post-synaptic spikes and brings about Long Term Potentiation (LTP) or Long Term Depression (LTD) [15]. In some electrophysiological experiments performed in cultured hippocampal neurons in 1998, the hypothesized dependence of the synaptic efficacy to the spike timing was experimentally confirmed [12]. Consequently, computational neuroscientists developed a model to approximate the findings of the experiment [18]. This model is today known as pair-based STDP (PSTDP) and is usually represented as

$$\Delta w = \begin{cases} \Delta w^+ = A^+ e^{(\frac{e}{\tau_+})} & \text{if } \Delta t > 0 \\ \Delta w^- = -A^- e^{(\frac{e}{\tau_-})} & \text{if } \Delta t \leq 0, \end{cases}$$

where $\Delta t = t_{post} - t_{pre}$ is the timing difference between a single pair of pre- and post-synaptic spikes. As demonstrated in Fig. 1, the amount of potentiation/depression will be determined as a function of the timing difference between pre- and post-synaptic spikes, their temporal order, and their relevant amplitude parameters ($A^+$ and $A^-$).

Since the report of the first memristor, various attempts have been made to devise artificial memristive synapses with PSTDP characteristics [19], [6], [20], [21]. In almost all of these implementations, the programmable non-volatile memristance (resistance or conductance) of the device is considered to play the role of the synaptic weight, and voltage spikes are applied to the two terminals of the memristor to alter its memristance. However, the device physics, models and attributes, spike shape, and the method spike applied to elicit memristance changes differ [22], [23]. In this paper, we have utilized similar spikes and memristive device model to [20], to develop a new synaptic circuit for higher order timing- and rate-based synaptic plasticity. We discuss the utilized model and the approach we have taken to implement our new device in the following.

A. Memristor model

For this paper we have chosen a simple voltage/flux driven memristor model such as the one proposed in [21], and utilized in [20]. The I-V characteristics of this memristor, which is claimed to be physically implemented in [19] can be written as:

$$i_{MR} = g(w, v_{MR})v_{MR}$$

$$\frac{dw}{dt} = f(w, v_{MR}),$$

where $i_{MR}$ and $v_{MR}$ are the current passing through and the voltage across the device, $w$ denotes a memristor physical state variable, and $g$ represents the nonlinear conductance of the device. According to [21], this memristor is voltage/flux driven, because its structural parameter depends on $v_{MR}$. Considering this model for a memristive device, one should define the function $f$, so that account for the memristive behaviour observed in physically implemented devices. Here we utilize a simple function similar to the one employed in [20]. This function is written as

$$f(v_{MR}) = \begin{cases} I_0 \text{sign}(v_{MR})[e^{\frac{|v_{MR}|}{v_0}} - e^{\frac{-v_{th}}{v_0}}] & \text{if } |v_{MR}| > v_{th} \\ 0 & \text{otherwise}, \end{cases}$$

where $I_0$ and $v_0$ are some physical parameters of the device and $v_{th}$ is its threshold, beyond which the conductance of the device changes exponentially. This behavioural model of a memristive device can be illustrated as shown in Fig. 2(a). Note to the two thresholds and the exponential growth of the conductance. These are the features that we exploit to devise a memristive synapse with STDP.

Fig. 2(b-c) demonstrate the current-voltage and resistance characteristics of the utilized device, which is simulated using the memristor macromodel used in this paper and was presented in [20]. This macromodel depicts a thresholding behaviour similar to a commercially available ion-based physical memristor [24] with a current-voltage characteristic as shown in Fig. 2(d).

III. Memristive Synapse with triplet STDP

In 2002, Froemke and Dan presented a modified STDP rule, that takes into account a suppressive mechanism among spikes. This mechanism was hypothesized to account for non-linearities observed in triplet STDP experiments, where the PSTDP failed [9]. In 2006, triplet STDP rule was proposed by Pfister and Gerstner [15] to account for a larger set of higher order STDP experimental data. Recently Cai et al. [10], have developed a synaptic circuit, utilizing memristors with adaptive threshold, to implement the suppressive STDP rule of Froemke-Dan [9]. They have shown that using their proposed memristive synapse, the triplet STDP behaviour of the suppressive STDP model can be reproduced. However, they did not explore the strength of their developed synapse in reproducing other synaptic plasticity experimental data, such as quadruplet and pairing frequency experiments. In this paper, we focus on implementing the triplet rule of Pfister and Gerstner using memristors.

A. Triplet STDP

The triplet rule extends the conventional form of STDP rule (shown in Eq. 1) and introduces extra potentiation/depression contributions for both pre- and post-synaptic spikes. In the
triplet model, the post spike, in addition to its exponentially decaying pairing depression potential, \( r_1 \), shown in Fig. 3, triggers an extra potentiation potential, \( o_2 \), for interaction with upcoming post spike(s). Similarly, the pre spike also gives rise to an extra depression potential, \( r_2 \), to interact with next pre spikes, besides its usual pairing potentiation potential trace, \( o_1 \). These extra triplet potentials that are shown in Fig. 3, may differ in time constants and amplitudes, compared to conventional STDP potentials. As shown in Fig. 3, at the time of the first pre spike, \( t_{pre1} \), a depression happens due to the previous post spike that has left a depression trace, \( r_1 \). Next, at the time of the second post spike, \( t_{post2} \), two potentiations take place. The first is due to the pre-post pairing, and the potentiation trace, \( o_1 \), that the first pre spike left. The second potentiation though, is a result of a triplet interaction (pre-post-post) among the first and second post and the first pre spikes. This potentiation depends on the two potentiation traces, one left by the first pre spike, \( o_1 \), and the second one triggered by the second post spike, \( o_2 \). This second trace is the differential point to the standard STDP rule, as it introduces interactions among spikes of the same pre or post neuron, and may lead to extra potentiation/depression. This triplet STDP interaction can be represented as

\[
\Delta w(t) = \begin{cases} 
A_1^\tau o_1(t) + A_2^\tau o_1(t) o_2(t - \epsilon) & \text{if } t = t_{post} \\
-A_1^\tau r_1(t) - A_2^\tau r_1(t) r_2(t - \epsilon) & \text{if } t = t_{pre},
\end{cases}
\]

(5)

where \( o_1 \) and \( o_2 \) are potentiation potentials triggered by pre and post spikes, respectively. In addition, \( r_1 \) and \( r_2 \) are depression potentials elicited by the arrival of post and pre spikes, respectively. Parameters \( A_1^\tau \), \( A_2^\tau \), \( A_1^\tau \) and \( A_2^\tau \) are constant amplitude parameters that determine the contribution strength of each spike in potentiation/depression. Here, \( \epsilon \) is a small positive constant which ensures that the weight update uses the correct values occurring just before the pre- or post-synaptic spike of interest.

Similar to the pair-based STDP, the triplet rule can also be mathematically represented as

\[
\Delta w = \begin{cases} 
A_1^\tau e^{(-\Delta t_1 / \tau)} + A_2^\tau e^{(-\Delta t_1 / \tau)} e^{(-\Delta t_2 / \tau)} & \text{if } t = t_{post} \\
-A_1^\tau e^{(-\Delta t_1 / \tau)} - A_2^\tau e^{(-\Delta t_1 / \tau)} e^{(-\Delta t_2 / \tau)} & \text{if } t = t_{pre},
\end{cases}
\]

(6)

where \( \Delta t_1 = t_{post(n)} - t_{pre(m)} \), \( \Delta t_2 = t_{post(n)} - t_{post(n-1)} - \epsilon \) and \( \Delta t_3 = t_{pre(m)} - t_{pre(m-1)} - \epsilon \), are the time differences between combinations of pre- and post-synaptic spikes, and \( \tau_- \), \( \tau_+ \), \( \tau_x \) and \( \tau_y \) are time constants relating to the potentiation/depression potentials of \( r_1 \), \( o_1 \), and \( o_2 \), respectively.

Pfister and Gerstner [15] have shown that the full TSTDP rule of Eq. 6 can be simplified, without the performance of the model in reproducing the experiments being compromised, to a minimal rule that does not include a triplet depression term, \( r_2 \). Therefore, the triplet rule of Eq. 6 will be minimized to a minimal TSTDP rule as

\[
\Delta w = \begin{cases} 
A_1^\tau e^{(-\Delta t_1 / \tau)} + A_2^\tau e^{(-\Delta t_1 / \tau)} e^{(-\Delta t_2 / \tau)} & \text{if } t = t_{post} \\
-A_1^\tau e^{(-\Delta t_1 / \tau)} & \text{if } t = t_{pre},
\end{cases}
\]

(7)

Note that in all the results presented in this paper, the minimal triplet STDP is used.

B. Relating memristor model to triplet STDP

For implementing the triplet STDP rule, we used the combination of two memristors and by applying the superposition principle, we can sum the weight changes of all spike interactions applied to these two memristors, to obtain the final weight change. Hence, considering Eq. 7, let us assume that memristance (synaptic weight) changes as follows

\[
\frac{dw}{dt} = f(w, \Delta v_{pair}(t)) + f(w, \Delta v_{triplet}(t)),
\]

(8)
where
\[ \Delta v_{\text{pair}}(t) = v_{\text{post}} - v_{\text{pre}}, \quad (9) \]
in response to a pre-post or post-pre pair of spikes, is applied to the two terminals of the first memristor shown in Fig. 4(a), and
\[ \Delta v_{\text{triplet}}(t) = v_{\text{triplet}_{\text{post}}}(t) - v_{\text{triplet}_{\text{dep}}}(t), \quad (10) \]
where
\[ v_{\text{triplet}_{\text{post}}}(t) = [v_{\text{post}}(t) - v_{\text{pre}}(t)]^+, \quad (11) \]
\[ v_{\text{triplet}_{\text{dep}}}(t) = [v_{\text{pre}}(t) - v_{\text{post}}(t)]^+, \quad (12) \]
are respectively responses to post-pre-post and pre-post-pre spike combinations, applied to the two terminals of the second memristor. Here, \([x]^+\) is a rectifier function represented as
\[ [x]^+ = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases} \quad (13) \]
If the minimal TSTDP rule is considered, i.e. \(v_{\text{triplet}_{\text{dep}}} = 0\), Eq. 10 is simplified to
\[ \Delta v_{\text{triplet}}(t) = v_{\text{triplet}_{\text{post}}}(t). \quad (14) \]
If we integrate Eq. 8, we can find the weight changes across memristors for various set of pre- and post-synaptic voltages (spikes) applied to the memristors as follows
\[ \Delta w(\Delta t_1, \Delta t_2) = \int_{\Delta t_1} f(\Delta v_{\text{pair}}(t)) \, dt + \int_{\Delta t_2} f(\Delta v_{\text{triplet}}(t)) \, dt. \quad (15) \]
Considering Eq. 15, in a post-pre-post triplet case of spikes, the integration will be
\[ \Delta w(\Delta t_1, \Delta t_2) = \int_{t_{\text{pre}}}^{t_{\text{post}} + t_{\text{post2}}} f(\Delta v_{\text{pair}}(t)) \, dt \]
\[ + \int_{t_{\text{pre}}}^{t_{\text{post2} - \epsilon}} f(\Delta v_{\text{pair}}(t)) \, dt \]
\[ + \int_{t_{\text{post1}}}^{t_{\text{post2}}} f(\Delta v_{\text{triplet}}(t)) \, dt. \quad (16) \]

The three parts of Eq. 16 are demonstrated in Fig. 4(b), where the first integral is over a period of 5 ms between post1-pre, which resulted in a decrease in memristance of the first memristor in Fig. 4(a), i.e. a synaptic depression, which is demonstrated as a negative value in orange in the fourth graph. The second integration is over a period of 15 ms, i.e. between the pre and post2 spikes and resulted in a positive value, shown in green in the fourth graph, representing an increase in the memristance of the first memristor. These two negative and positive values are in relation to the pair-based STDP model. However, according to the minimal triplet STDP, formulated in Eq. 7, the post1-post2 spikes in the presence of a pre spike can result in potentiation. This potentiation shown in the bottom graph of Fig. 4(b), is demonstrated as an increase in the memristance of the second memristor in Fig. 4(a), and is in result of an integration over the post1-post2 spikes period as shown in Eq. 16. Note that, due to the lack of a second pre spike in this triplet, the second term of the second equation, in the equation array shown in Eq. 6 is zero, and therefore, no further depression will be elicited. This is also the case when a minimal TSTDP model is considered.

In order to verify the functionality of the proposed triplet memristive device, it was used to replicate an experimental data set generated using post-pre-post triplet spikes in [16]. Fig. 4(c) demonstrates a very close match between the memristance changes obtained using the proposed device, and those
measured in the triplet experiments [15], [16]. This match was obtained by optimizing the STDP and memristive parameters to reach the least error.

IV. EXPERIMENTAL RESULTS

So far we only considered synaptic weight changes for one pair or triplet of spikes using the proposed CMOS-memristive synapse. However, an extensive set of simulations should be carried out for reproducing the outcomes of a variety of essential experimental data, which has been the subject of research in both neuromorphic [25], [26] and computational neuroscience research [27]. In this section, we report our findings and demonstrate the limitations of the PSTDP memristive synapse of [20] in reproducing a number of experiments. We then show how our proposed synapse can closely replicate the outcomes of a number of previous experiments, using a single set of STDP parameters.

A. Experimental protocols

Certain standard experimental protocols are predominantly followed in the area of electrophysiological experiments to study synaptic plasticity [9], [12], [13], [16]. The same protocols should therefore be employed while verifying the performance of the devised computational models [9], [14], [15] or neuromorphic devices [10], [25], [28] in approximating/replicating the experimental data observed in biological synapses. Here we have followed similar protocols to those deployed in synaptic plasticity experiments to examine the functionality and performance of our proposed CMOS-memristive synapse, and to compare it with its conventional STDP counterparts. The utilized protocols are Pair-based STDP, frequency-dependent STDP, triplet-based STDP, extra triplet STDP, and quadruplet, which are defined in our previous studies [25] and in the TSTDP modelling paper [15].

B. Data fitting approach and experimental scenarios

In order to test the efficacy of a synaptic model/device, one can define an error function that represents the difference among the weight changes predicted by a candidate model/device, and those measured in electrophysiological experiments. A suitable error function, is the Normalised Mean Square Error (NMSE) function proposed and utilised in [15],

\[
\text{NMSE} = \frac{1}{P} \sum_{i=1}^{P} \left( \frac{\Delta w_{i}^{\text{exp}} - \Delta w_{i}^{\text{model}}}{\sigma_{i}} \right)^2 ,
\]

where \(\Delta w_{i}^{\text{exp}}, \Delta w_{i}^{\text{model}}\) and \(\sigma_{i}\) are the mean weight change obtained from biological experiments, the weight change obtained from the model or circuit under consideration, and the standard error mean of \(\Delta w_{i}^{\text{exp}}\) for a given data point \(i\), respectively. Here, \(P\) represents the number of data points in the data set under consideration.

In all experiments performed in this paper, we utilized the MATLAB built-in \texttt{fminsearch}, an unconstrained non-linear minimization function, to minimize the NMSE for the synaptic device under consideration. For instance, for the proposed bi-memristor hybrid synapse, in Fig. 4, five parameters including four of the triplet STDP rule embedded in the spike shapes \((A_1^+, A_1^-, A_2^+, \tau_y)\) and one relating to the memristor \(f\) function, \(v_0\), were optimized. We present results obtained from various experiments, in which these parameters along with some other parameters are optimized to reach the best NMSE in different scenarios.

Two different scenarios can be considered to verify the functionality and performance of pair-based and triplet-based memristive circuits in reproducing the outcomes of experiments using the aforementioned protocols. Under first scenario, similar to the experiments in [15], PSTDP time constants, i.e. \(\tau_+\) and \(\tau_-\), are kept fixed and equal to 16.8 ms and 33.7 ms respectively, while other parameters are optimized. Under scenario two, these parameters are optimized along with other parameters to study the effect of higher parameter flexibility on synaptic plasticity.

C. Frequency-dependent pairing (visual cortex) experiments

Under the first scenario, pair-based STDP fails to mimic experimental data, where synaptic weight changes are examined against the frequency of pairs of spikes, \(\rho\). The optimum NMSE is achieved using the pair-based memristive synapse presented in [20] is 8.19. The resulting weight prediction is shown in Fig. 5(a), which interestingly is similar to the weight changes predicted by the PSTDP computational model shown in Eq. 1, as reported in [15]. Scenario two results in a lower NMSE of 1.69 and an improved match to the experimental data (see Fig. 5(b)). However, the optimization results in a very long potentiation time constant, \(\tau_{+} = 110\) ms, and a very short depression constant of almost 1 ms. These time constants lead to only potentiation when the repetition frequency is high enough, i.e. \(\rho > 10\) Hz, for the spikes to overlap. Hence, we can conclude that the pair-based memristive STDP synapse is not capable of reproducing the outcome of frequency-dependent pairing experiments. This is in agreement with presented results in [15]. On the other hand, further simulations suggest that regardless of the optimization of time constants, the proposed hybrid circuit can closely approximate the behaviour observed in the experiments. These behaviours, which resulted in NMSE = 0.45 for the first scenario, and NMSE = 0.34 for the second one, are shown in Fig. 5(c) and (d), respectively.

D. Pair, triplet, and quadruplet (hippocampal) experiments

In the triplet-based STDP study by Pfister and Gerstner [15], one set of parameters and the minimal version of the triplet STDP model were utilized, to minimize the NMSE for a set of experimental data composed of 13 data points, including pairs (2 data points), triplets (8 data points), and quadruplets (3 data points). These data points and their respective error bars, which are shown in black in our figures, represent experimental data obtained from hippocampal culture as reported in [16].

Our performed experiments using the PSTDP memristive synapse of [20] show that this circuit fails to account for the hippocampal culture data set under both scenarios. The results for scenario two, are shown in Fig. 6. As expected the PSTDP memristive synapse with optimized parameters can
Fig. 5. (a) Pair-based STDP memristive synapse of [20] fails to mimic experimental data of [13], under scenario 1. (b) Similar to (a), only for scenario 2. (c-d) The proposed triplet-based STDP memristive synapse successfully mimics experimental data, under scenario 1 shown in (c) and scenario 2 presented in (d). Here \( \rho \) represents the frequency of spike pairs with time differences of 10 ms and -10 ms.

Successfully replicate the STDP learning window. However, in the case of quadruplet experiment (Fig. 6(c)), the PSTDP memristive synapse shows similar behaviour to the failure of PSTDP model as presented in [15]. Beside these, the PSTDP memristive synapse clearly lacks the ability to distinguish between the pre-post-pre and post-pre-post experiments as shown in Fig. 6(c)-(d). This is simply due to the accumulative nature of the PSTDP rule and its memristive synapse, which sum the effect of post-pre and pre-post spike pairs in a post-pre-post triplet, and similarly aggregate the effect of pre-post and post-pre pairs in a pre-post-pre triplet. Therefore, no difference between the two triplets is expected as both of them consist of a pre-post along with a post-pre spike pair. However, the experimental data, demonstrated in black, suggests significant difference between the two triplets. Furthermore, our simulations also demonstrate that the pair-based memristive STDP synapse, using scenario 1, i.e. with fixed \( \tau_+ \) and \( \tau_- \) and while only the other four parameters are optimized, cannot reach an NMSE smaller than 12.25, and clearly lacks the ability to mimic the experimental data.

In contrast to the PSTDP device, our minimal TSTDP memristive synapse shows a very close match to the experimental data, and achieves a very low NMSE of 0.87 under scenario 2. This is much lower than the PSTDP memristive synapse, where we reached an NMSE of 7.42 under same scenario. Experimental results using the triplet synapse demonstrated in Fig. 7(c-d) show how well this synaptic circuit distinguishes between two different cases of triplet data. The memristive synapse, results in strong potentiation in case of post-pre-post triplet (Fig. 7(d)) as expected. This is due to the triplet potentiation interaction, which is absent in case of pre-post-pre triplet (Fig. 7(c)). Fig. 7(a) also demonstrates the conventional STDP learning window generated by our TSTDP circuit, which closely matches the two targeted PSTDP experimental data. Finally, in case of quadruplet experiments (Fig. 7(b)), the proposed synapse closely fits the data points and follows similar behaviour to the data obtained using the minimal triplet computational model of Pfister and Gerstner [15]. Additionally, in case of scenario 1, i.e. when the pair time constants \( \tau_+ \) and \( \tau_- \) are kept fixed, an NMSE of 3.61 was reached using our proposed triplet circuit, which is much lower than the NMSE = 12.25, and NMSE = 7.42, obtained using the pair-based device.

Table I summarises NMSEs and their respective optimized parameters for all the experiments and scenarios mentioned in previous subsections. In all cases, the triplet memristive circuit results in a better NMSE compared to its pair-based counterpart of [20].

### E. Triplet experiments of Froemke-Dan [9]

Apart from the triplet experiments performed in previous subsection, where only pre-post-pre and post-pre-post triplets were considered, one may investigate the effect of other combinations of three spikes. This is the experiment, reported in [9], where six different combinations of spike triplets are studied (see Fig. 8(a)). Interestingly, our proposed triplet device, is capable of reproducing a close approximation of the data from the suppressive STDP model of Froemke-Dan [9], for these extra triplet experiments. Fig. 8(b) demonstrates the outcome of extra triplet experiments using the proposed triplet memristive circuit, under the first (Fig. 8(b1-b2)), and second (Fig. 8(b3-b4)) scenarios. In the first scenario, the optimized parameters for the triplet case of hippocampal culture experiments [15] were used, where the best NMSE achieved was 3.61 (see Table I). Part (b1) in Fig. 8 depicts the variety of combinations of 2 post and 1 pre spikes, their time differences, \( \Delta t_1 = t_{post1} - t_{pre} \), \( \Delta t_2 = t_{post2} - t_{pre} \), and their resulting weight modification, \( \Delta w \), shown as a colorbar. Part (b2) demonstrates the weight changes achieved using the same set of parameters in the triplet circuit, when 2 pre and 1 post spikes are combined, as shown in the figure.
In order to further test the performance of the proposed triplet memristive circuit, we utilized the optimized parameters in case of scenario 2, where the achieved minimal NMSE was 0.87. Figure 8(b3-b4) presents the strength of the triplet circuit in closely approximating the outcome of the triplet experiments as shown in [9].

Although using different scenarios and optimized parameter sets, both parts (b1-b2) and (b3-b4) in Fig. 8 present similar potentiation/depression characteristics for triplet spike combinations. Both these parts correctly mimic the weight changes as observed in the triplet experiments reported in [9], except for the post-pre-post triplet case. The reason for this difference has been explained in [15].

It is worth noting that, for obtaining the results demonstrated in Fig. 8, the same parameters that were utilized for reproducing the hippocampal experiments are applied. For instance, the results shown in Fig. 8(b) are obtained using the same parameters, using which Fig. 7 weight changes were attained. This feature further testifies to the strength of the proposed triplet device, which can reproduce the outcome of quadruplet, pairing and various triplet experiments including those that have not been explored in [15].

V. THE HYBRID SYNAPSE IN CROSSBAR ARRAY

In the structures presented in Fig. 4, only one instance of the proposed bi-memristor hybrid synapse is demonstrated, without considering the inclusion of such a synapse in a crossbar array structure. In order to utilize the proposed circuit in crossbar arrays and employ it for simultaneous learning and computation, the structure must be slightly modified.

Figure 9(a) shows a pre-synaptic neuron connected through a bi-memristor hybrid synapse to a post-synaptic neuron. Here, a modified CMOS neuron circuit compared to that of [20] has been utilized. The difference between this new CMOS neuron and those utilized in the implementation of PSTDP learning is that, not only does this neuron generate post spikes to interact with pre spikes, it also produces other post spikes, i.e. post1 spk, required for triplet learning. These spikes as shown in Eq. 11 should be multiplied by the spikes coming from afferent pre-synaptic neuron, and then the result should be rectified and applied to the triplet memristor ($R_{trip}$), as shown in Fig. 9(a).

This figure also shows two switches across the triplet memristor. These switches, controlled by a signal from the post-synaptic neuron, ensure the correct integration and learning in the memristive synapses. When integrating spikes from the pre-synaptic neuron, the switches are closed as shown in the figure, and therefore apply the pre spikes to the left sides of both memristors at the junction. On the post-synaptic (right) side of the memristors, a reference voltage is generated by the neuron, that shows no post-synaptic activity and at the same time, ensures no change in the memristive weights. This operation can be interpreted as a weight read phase, where the pre spikes are integrated into post neurons and excite it proportional to the strength of their respective memristive synapses.

On the other hand, when the post-synaptic neuron, in result of the integrations, fires a spike, the switches turn to the second position and the learning phase starts. The weights of memristive synapses are then changed due to the existence of overlapping pre- and post-synaptic spikes and the timing differences between them. In this case, due to a possible overlap of a pre-synaptic spike and a triplet post-synaptic spike, the weight of the second memristor can be modified. Note that, in this figure, the minimal version of the TSTDP rule is implemented, i.e. the triplet depression interactions from pre-synaptic neuron are neglected. This results in having the left side of the memristor connected to ground during learning, as shown in the figure.
The circuit structure shown in Fig. 9(a), which includes a multiplier/rectifier circuit and a number of switches, was simulated in Cadence Spectre. The multiplier is a CMOS Gilbert cell that along with a comparator and two pass gates perform the required multiplication/rectification. In addition, four other pass gates were used to properly control the weight read and learning (weight change) phases. The utilized memristor is the one employed in the experiments performed in [20]. Results using the implemented hybrid circuit are demonstrated in Fig. 9(b).

The proposed structure demonstrated in Fig. 9(a) is scalable. Figure 10 depicts a 3-by-3 crossbar array that incorporates CMOS neurons and CMOS-memristor synapses. This figure suggests that the proposed TSTDP bi-memristor synapses, analogous to their former single memristor PSTDP counterparts [20], can be integrated with CMOS neurons and implement large scale neural arrays.

In the neural array presented in Fig. 10, each post-synaptic neuron is driven by three pre-synaptic neurons connected to it through three bi-memristor hybrid synapses. Here, the second pre-synaptic neuron is assumed silent, hence it is replaced by a voltage source of the value of the post-synaptic neuron spiking threshold, \( V_{\text{REF}} \). The other two pre-synaptic neurons generate regular spike trains with various inter spike intervals. In addition, memristors in the three rows of the array are given various initial weights, which results in difference in spiking activities of their respective post-synaptic neurons.

Cadence simulation results that demonstrate changes in synaptic weights across all memristors in the 3-by-3 crossbar array of Fig. 10 are shown in Fig. 11. In this figure, the first row depicts weight changes corresponding to the synapses in the first column of the array, i.e. related to the first pre-synaptic neuron. The second row shows that no change takes place on the memristive synapses in the second column. This is due to the fact that the pre-synaptic neuron associated with this column is silent and produces no spike. The third row manifests weight changes occurring across the synapses driven by the third pre-synaptic neuron. Here, various weight change profiles are due to different initial weights set on the synapses, and because of the different timings among postsynaptic spikes generated by the post neuron in each row. As expected, only potentiation occurs for the triplet memristors, while both potentiation and depression are observed on pair memristors of each hybrid synapse, due to various pre-post or post-pre spike combinations.

VI. DISCUSSION AND CONCLUSION

Implementing area efficient, low-power, and large-scale neural-inspired learning architectures can be facilitated using memristors [23], [29]. Nanoscale dimensions, intrinsic non-volatility, and ultra low power consumption [30] combine to make memristors perfect candidates to implement synapses in neuromorphic architectures. The learning performance of these architectures is strongly governed by the plasticity mechanisms their synapses implement [31], [32]. Therefore, careful consideration must be taken when synaptic plasticity mechanism of the targeted neural platform is being chosen.
In terms of synaptic mechanisms, many studies have explored the implementation of the simple yet naive pair-based STDP rule using memristive devices [19], [5], [6], [20], [21], [22]. Only a few studies report implementations of other more powerful synaptic plasticity mechanisms such as suppressive STDP [10]. These mechanisms that have advanced synaptic plasticity (learning) abilities compared to the PSTDP rule, can improve the performance of the developed neuromorphic architectures in learning and computation. In order to reach higher learning capabilities in future neural architectures, this paper proposes a novel CMOS-memristive design for a higher order STDP rule, namely triplet STDP, which has advantages over its previous CMOS [25], [33], [34] as well as memristive [5], [6], [10], [20], [22] counterparts and significantly improves learning capabilities of neuromorphic synapses. The proposed synaptic circuit is composed of two memristors along with several CMOS transistors to account for the non-linearities of the triplet rule proposed by Pfister and Gerstner [15]. Although this hybrid CMOS-memristive circuit, compared to its memristive PSTDP counterparts [6], [20], has higher complexity in terms of implementation, it offers significantly improved learning performance. This higher performance is achieved by adding a second (triplet) memristor, as well as a CMOS multiplier/rectifier circuit.

Many previous CMOS STDP synapse circuits occupy a large silicon area, even if the synaptic weight storage is not considered. This could be improved by using a memristive design such as the proposed circuit in this paper. For instance, the presented PSTDP synapse in [33] occupies an area of $145 \times 31 \, \mu m^2$ in a 0.8 $\mu m$ CMOS process, and the PSTDP weight update design proposed in [34] takes up $131.3 \times 139.7 \, \mu m^2$ in a 0.6 $\mu m$ CMOS process. In addition, a previous TSTDP circuit implemented by our group has an area of $165 \times 60 \, \mu m^2$ in a 0.35 $\mu m$ CMOS process, from which over 75 percent is occupied by five large capacitors [35]. Note that these areas are only related to the weight update circuitries.
and do not include the permanent weight storage devices such as memory cells and required data converters, that are not needed in a memristive synapse design. The proposed design that includes a simple CMOS Gilbert multiplier, a comparator (for rectification) and six pass gates occupies $\approx 600 \ \mu$m$^2$ in a 0.35 $\mu$m CMOS process, improving area by a factor of 10, while implementing a more powerful synaptic plasticity algorithm.

Note that, presented results in Table I are obtained using ideal behavioural models of multiplier/rectifier and switches. However, the results shown in Figures 9 and 11 are generated using non-ideal CMOS circuitry, which have limitations in the range of inputs/outputs and cannot be deployed for all the optimized values of pre- and post-synaptic spikes shown in Table I. To address this issue, one can either scale down the optimized amplitude values to the power rails of the multiplier/rectifier circuit, or trade off with a higher error.

To further verify the proposed circuit, it was simulated in the presence of noise with the bandwidth in the range of 100 Hz to 1 KHz. Results demonstrate good stability in the synaptic plasticity of the circuit in the presence of noise. In addition, simulations were performed to measure power consumption of the proposed hybrid synaptic circuit when responding to trains of pre- and post-synaptic spikes for 5 seconds. The average power consumption for the CMOS circuit (shown in 9(a)) is 310 $\mu$W. This could be further optimised as power consumption was not a design goal of this work.

As part of our experiments, we also investigated the effect of spike fusion mechanism presented in [10] on the synaptic plasticity ability of our proposed circuit, as well as the previous PSTDP circuit of [20]. These investigations demonstrate positive effect, i.e. reducing NMSEs, of the spike fusion in the performance of the PSTDP device of [20], while confirming proper functionality and low NMSEs of the proposed TSTDP device using fused spikes for almost all experiments. For instance, note that in the circuit simulations demonstrated in Figures 9 and 11, the post1 spikes are fused, i.e. start over once the next post1 spikes arrive.

When operating a memistor one must ensure to a maximum current (compliance) is not exceeded, because a high current can destroy the device. To address this issue in the proposed hybrid device, each synaptic memristor should have an NMOS in series. The gate of the NMOS should be set to maximum (so the NMOS is a closed switch) when reading the device memristance, i.e. when only pre-synaptic spikes are available. On the other hand, the gate should be set to a given voltage when the synaptic weight change, i.e. write phase, is taking place. This voltage is chosen to limit the current to the desired compliance. The voltage can be conveniently provided row-wise together with the post-synaptic pulse, similar to the approach adopted in [36].

In addition to studies that merely propose devices to mimic the plasticity of biological synapses [19], [6], [7], [17], some previous studies have utilized memristive synaptic circuits along with CMOS neurons in engineering applications such as extracting visual features [2] and pattern recognition [37]. The majority of these applications employ memristive devices with pair-based STDP learning capability. However, the use of other more complex, yet more powerful learning rules such as TSTDP, are yet to be explored. The TSTDP rule that is the subject of our study is shown to not only inherit all the properties of PSTDP, but also applicable to more complex tasks such as direction and speed selectivity [27], which may improve the image and pattern classification abilities of the previously developed STDP circuits. This is an open question for future research.

All the above mentioned experiments and verifications confirm that our bi-memristor hybrid synapse surpasses available CMOS and memristive synaptic devices and circuits in terms of area and synaptic plasticity strength, and hence can improve the learning capabilities of large-scale neuromorphic systems for learning and computation.

ACKNOWLEDGEMENT

The authors would like to thank The 2015 CapoCaccia Cognitive Neuromorphic Engineering Workshop, where collaboration and discussions on this work were initiated. M. Rahimi Azghadi and P. Leong also acknowledge the support of an Australian Institute for Nanoscale Science and Technology (AIST) accelerator grant from the University of Sydney. M. Rahimi Azghadi is also thankful to Prof Robert Froemke for useful discussion and for providing some experimental data.

REFERENCES

Fig. 11. Cadence simulation results of the 3 x 3 memristor cross-bar array shown in Fig. 10.


Mostafa Rahimi Azghadi (S’07–M’14) received the M.Sc. degree in Computer Architecture Engineering from Shahid Beheshti University, Tehran, Iran, in 2009. In 2010, he was awarded an International Postgraduate Research Scholarship (IPRS) to pursue his PhD in the School of Electrical & Electronic Engineering, The University of Adelaide, Australia. In 2014 he completed his PhD earning the Doctoral Research Medal, as well as the Adelaide University Alumni Medal. In 2012 and 2013, he was a visiting PhD student in the Neuromorphic Cognitive System group, Institute of Neuroinformatics, University and Swiss Federal Institute of Technology (ETH) Zurich, Switzerland. From 2014 to 2015, Dr Rahimi was a postdoctoral research fellow in Centre for Biomedical Engineering (CBME) of The University of Adelaide, working in the area of smart sensors for biomedical applications. From 2015 to 2016, Mostafa was a postdoctoral research associate in the School of Electrical and Information Engineering of the University of Sydney, Australia, working in the area of memristive neuromorphic systems, VLSI design, and VLSI interfaces for nanoscale sensors. Since 2016, he has been a lecturer in College of Science and Engineering, James Cook University, Townsville, Australia. Dr Rahimi was a recipient of several national and international awards and scholarships such as Endeavour Postdoctoral Fellowship (2015), South Australian Science Excellence Awards Finalist (2015), Australian Nanotechnology Network Young Science Ambassador Award (2015), Doreen McCarthy Research Bursary (2014), Simon Rockloff Scholarship (2014), and D. R. Stranks fellowship (2013). He has served as a review editor of *Frontiers in Neurorhic Engineering* and is a reviewer for several recognised international journals and conferences such as IEEE TBCAS, IEEE TCAS, IEEE TVLSI, IEEE TC, IEEE ISCAS, and IEEE BioCAS.

Bernabe Linares-Barranco (F’10) received the B.S. degree in Electronic Physics, the M.S. degree in Microelectronics, and a first Ph.D. degree in June 1990 from the University of Sevilla, Sevilla, Spain, in 1986, 1987, and 1990, respectively, and a second Ph.D. degree from Texas A&M University, College Station, USA, in 1991.

Since 1991, he has been with the Sevilla Microelectronics Institute (IMSE-CNM), from the Spanish Research Council (CSIC) of Spain, where he currently is Full Professor of Research. He has been visiting professor/fellow at the Johns Hopkins University, Baltimore (USA), Texas A&M University (USA), and The University of Manchester (UK). His recent interests are in Address-Event-Representation VLSI, real-time AER vision sensing and processing chips, memristor circuits, and extending AER to the nanoscale. He has received two IEEE TRANSACTIONS Best Paper Awards, and has been an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–PART II, the IEEE TRANSACTIONS ON NEURAL NETWORKS, and Frontiers in Neuromorphic Engineering. From 2011 to 2013, he was the Chair of the IEEE Circuits and Systems Society Spain Chapter.


He was a Research Engineer with the GEC Hirst Research Centre, London, from 1978 to 1986. From 1986 to 1987, he was a VLSI Design Engineer with Austek Microsystems, Australia. Since 1987, he has been with The University of Adelaide, where he is currently a Full Professor with the School of Electrical and Electronic Engineering. He co-edited Quantum Aspects of Life (London, U.K.: Imperial College Press, 2008), and co-authored the book entitled Stochastic Resonance (Cambridge, U.K.: Cambridge University Press, 2012) and Terahertz Imaging for Biomedical Applications (New York, NY, USA: Springer-Verlag, 2012). He holds over 800 publications/patents and has been an Invited Speaker at over 100 institutions. His interests are in the area of multidisciplinary physics and electronic engineering applied to complex systems. His research programs span a number of areas of stochastics, game theory, photonics, biomedical engineering, and computational neuroscience. Prof Abbott is a fellow of the Institute of Physics. He has won a number of awards, including the South Australian Tall Poppy Award for Science (2004), the Premier’s SA Great Award in Science and Technology for outstanding contributions to South Australia (2004), an Australian Research Council Future Fellowship (2012), and The David Dewhurst Medal (2015). He has served as an Editor and/or Guest Editor for a number of journals, including the IEEE JOURNAL OF SOLID-STATE CIRCUITS, Journal of Optics B, Microelectronics Journal, Chaos, Smart Structures and Materials, Fluctuation and Noise Letters, PROCEEDINGS OF THE IEEE, the IEEE PHOTONICS JOURNAL, and PLOSONE. He is on the Editorial Boards of Scientific Reports (Nature), the IEEE ACCESS, Royal Society Online Science, and Frontiers in Physics.
Philip Leong (SM’02) received the B.Sc., B.E. and Ph.D. degrees from the University of Sydney. In 1993 he was a consultant to ST Microelectronics in Milan, Italy working on advanced flash memory-based integrated circuit design. From 1997–2009 he was with the Chinese University of Hong Kong. He is currently Professor of Computer Systems in the School of Electrical and Information Engineering at the University of Sydney, Visiting Professor at Imperial College, Visiting Professor at Harbin Institute of Technology, and Chief Technology Advisor to Cluster Technology.

Prof. Leong was co-founder and program co-chair of the International Conference on Field Programmable Technology (FPT); program co-chair of the International Conference on Field Programmable Logic and Applications (FPL); Senior Associate Editor for ACM Transactions on Reconfigurable Technology and Systems (TRETS) and Associate Editor for IEEE TRANSACTIONS ON COMPUTER AIDED DESIGN (TCAD). The author of more than 150 technical papers and 5 patents, Dr. Leong was the recipient of the 2005 FPT conference Best Paper as well as the 2007 and 2008 FPL conference Stamatis Vassiliadis Outstanding Paper awards.