2011 11th IEEE International Conference on Nanotechnology
Portland Marriott
August 15-18, 2011, Portland, Oregon, USA

Spike Timing-Dependent Synaptic Plasticity Using Memristors and Nano-Crystalline Silicon TFT Memories

Kurtis D. Cantley[a)], Anand Subramaniam, Harvey J. Stiegler, Richard A. Chapman, and Eric M. Vogel[b)]

University of Texas at Dallas, 800 W. Campbell Rd., Richardson, TX, USA 75080
[a)]kurtis.cantley@utdallas.edu; [b)]eric.vogel@utdallas.edu

***Abstract*–Neural circuits based on ambipolar nano-crystalline silicon TFTs and memristive synapses are investigated via SPICE simulations. The drive transistor for the memristive devices is an ambipolar TFT with memory that could be physically implemented using a metal nanoparticle layer within the gate dielectric. It is shown that using such a device adds spike-timing dependence to changes in the synaptic weight. In experiments with action potential pairs, the synaptic weight modification is similar to biological data. Further, asymmetric temporal integration of the weight change is demonstrated using pre-post-pre and post-pre-post spike triplets. Finally, the dependence of weight changes on frequency is presented. This is followed by a discussion of applications and issues which require further analysis.**

I. Introduction

The demonstration of TiO_2 resistive switches by HP Labs in 2008 [1] sparked significant interest in the possibility of using memristive devices as synapses in artificial neural circuits. Much of the resulting research has angled toward implementing spike-timing-dependent plasticity (STDP), a synaptic learning mechanism based on timing differences between action potentials [2,3]. However, due to the relatively long time scales of biological inter-spike intervals (on the order of tens of milliseconds) it is difficult to implement efficient schemes using (high-speed) electronics. Previously proposed solutions involve pulse width or height modulation or pulse shaping [4-6] and would require extensive auxiliary circuitry. In addition, there have been no reports of learning characteristics beyond pair-based trials in electronic systems. Experiments on biological samples indicate a much more complex reality, and the exact mechanisms of synaptic learning remain unclear [7,8]. Specifically, asymmetric temporal integration of the synaptic weight changes have been demonstrated in spike triplet and quadruplet, as well as frequency dependence experiments [9-11]. Progress is being made on developing models that explain more of the observed effects [12,13].

The circuits described in this work are the first to realize a compact physical implementation of the STDP learning mechanism with biologically realistic action potentials. Results of triplet and frequency dependence experiments are also shown. Complex synaptic behavior is a natural outcome of using the TFT memory devices based on nano-crystalline silicon (nc-Si) in conjunction with memristive (resistive switch) devices. The consequences of these results are twofold. First, they suggest the feasibility of fabricating a three-dimensional system with physical structure similar to the human neocortex, since nc-Si material can be deposited over large areas at low temperature. This feature also negates the need for any silicon CMOS circuits. Second, the results show that creative solutions utilizing nanotechnology can be applied to neural systems to provide important complex behaviors, such as STDP.

The following section discusses the specific device models and the circuit configurations used in the simulations. Various learning properties of the system are then presented in section III. Synaptic weight changes are measured during pair and triplet STDP simulation experiments, and the results are qualitatively similar to biological data. Then, the weight change dependence on frequency is briefly examined. Finally, section IV provides a discussion of more general STDP and neural circuit topics. These include applications of STDP learning rules, improvements to the circuits presented, and power consumption and scalability issues.

II. Device Models

The ambipolar nc-Si TFT SPICE models used in this work are based on characteristics reported in [14], but scaled to submicron dimensions to increase density. Fabrication of similar amorphous silicon and nc-Si nanowire TFTs have been reported [15,16]. The devices have been connected into the spiking neuron circuit configuration shown in Fig. 1a, which is a heavily modified version of that originally proposed by Mead [17]. The TFT and memristor models have been discussed extensively in previous work [18,19] in which each synapse consisted of one TFT and one memristor. Here, the synapse has been simplified to a single memristor whose state variable w/D (definition provided in [1]) represents the synaptic weight. At the output of each neuron, memory capability has been added to the drive transistor. Thus in a large network, a memory TFT would actually be connected to the output of each neuron circuit such that it would drive a large number of memristive synapses. In this work, the configuration has been simplified to include only the drive transistor of a pre-synaptic neuron and one memristor synapse. The drive TFT on the post-synaptic neuron has also been omitted.

To model the memory characteristics of the drive transistor, the gate voltage of the intrinsic ambipolar nc-si TFT is adjusted using a dependent voltage source. The adjustment accounts for trapped charge in the oxide. A schematic of the subcircuit is illustrated in Fig. 1b. The simulation testing procedure is as follows. The gate voltage is first ramped up and down to a maximum value that is less than the

978-1-4577-1515-0/11/$26.00 ©2011 IEEE

Authorized licensed use limited to: University of Groningen. Downloaded on July 12,2020 at 09:25:05 UTC from IEEE Xplore. Restrictions apply.

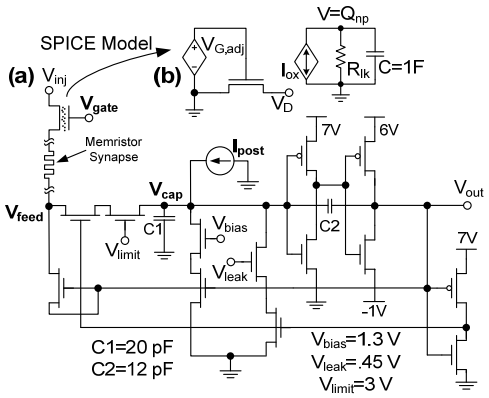

Fig. 1. (a) Diagram of the circuit used to investigate spike-timing-dependent learning. The memory transistor is part of the pre-synaptic neuron. (b) The SPICE model of the nanoparticle memory transistor consists of two parts: the intrinsic transistor with adjusted gate voltage, and an auxiliary circuit used for calculating the amount of trapped charge.

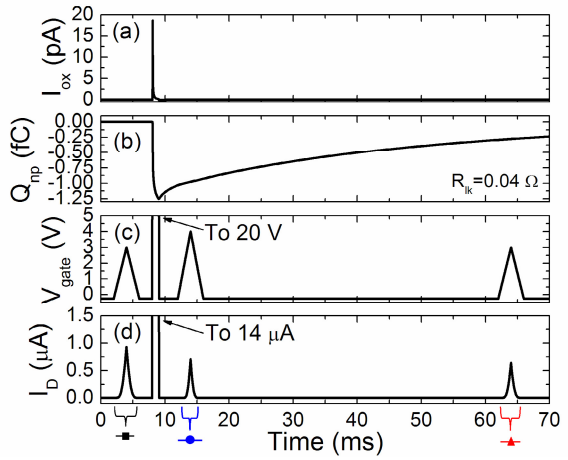

Fig. 2. Plots of the (a) tunneling current I_{ox}, (b) trapped charge Q_{np}, (c) gate voltage V_{gate}, and (d) drain current I_D versus time. I_D is probed with a low V_{gate} sweep immediately before and after programming, then again after 54ms, when much of the trapped charge has leaked off. The labeled probing regions correspond to the symbols in Fig. 3. Drain voltage V_D=1 V.

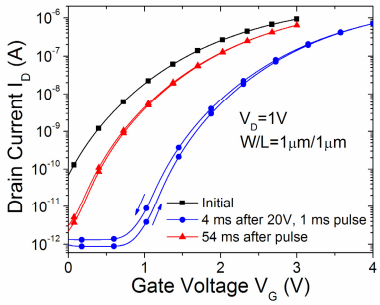

Fig. 3. I_D-V_G characteristics of the transistor immediately prior to, 4 ms after, and 54 ms after the programming pulse (pre-synaptic action potential V_{gate}). These curves correspond to the temporal V_{gate} sweep shown in Fig. 2c.

programming voltage (Fig. 2c). The drain current is measured during this time (Fig. 2d). A 20 V programming pulse is then applied to the device, causing electrons to tunnel to particles 30 nm from the semiconductor surface. The amount of trapped charge Q_{np} is calculated assuming Fowler-Nordheim tunneling to the particles (Figs. 2a and b). Two more gate voltage ramps are applied, one shortly after the programming pulse, and another some time later (Fig. 2c). Fig. 3 shows the transfer characteristics during each of these three ramps, and how the overall flatband voltage shifts with time. The maximum shift is approximately 1.5 V, and charge constantly leaks off (via R_{lk} in Fig. 1b) such that the retention time is in the 50-100 ms range, as shown in Fig. 3 (or Fig. 2b). These characteristics would ideally be obtained using a 1 ms programming pulse of the same magnitude as the action potential (<10 V). However, the TFT model being used is only valid for t_{ox}=300 nm, and thus a 20 V programming pulse is required. The consequences of using this relatively high voltage are discussed later. Although no measured data is available, this TFT memory would generally be considered low performance and fabrication of a comparable device is not inconceivable. Thinner oxides should be capable of achieving these metrics. Also, organic [20] and amorphous silicon [21] TFTs with similar characteristics have been demonstrated.

III. Synaptic Learning Mechanism

A. Action Potential Pairs

Changes in synaptic weight as a function of spike timing differences were measured using a 100 ms transient simulation of the circuit in Fig. 1a. The 20 V pre-synaptic action potential (V_{gate}) was always applied at t_{pre}=50 ms (Fig. 4b), and the post-synaptic action potential t_{post} was varied to achieve different $\Delta t=t_{post}-t_{pre}$. The voltage V_{inj} was also applied at the same time as V_{gate}, but with magnitude 1.8 V (and ground otherwise). Timing of the post-synaptic action potential at V_{out} was varied using a 12 nA current pulse I_{post} injected directly onto the node V_{cap}. This pulse was applied starting at a specific time, and turned off at the instant V_{feed} reached 1.8 V, as shown in Fig. 4a. The time of this turn off corresponds to t_{post}. The specific reason for both the location and behavior of I_{post} is to eliminate artificially induced errors in the synaptic weight measurement. If the current source is placed at V_{feed}, feedback of the injected current through the memristor when the input transistor turns off significantly increases potentiation. And although other input synapses will be injecting current directly to that node, it is preferable in these simulations to inject directly onto the "membrane" capacitor C1. However, in such a configuration, if the injection current is large while the neuron is firing, the charge leakage off C1 through the reset transistors (controlled by V_{bias}) is slower. The action potential width is thus larger, causing slightly more total current to flow back through the synapse. Careful evaluation of the circuit operation is therefore necessary when evaluating the performance of circuits and systems such as these.

In addition to the current injected to initiate specific post-synaptic spikes, a constant DC current $I_{post,DC}$ was also added. Its purpose is to imitate the excitation of the neuron by many other afferents. It causes the voltage V_{feed} to be greater than zero, which is a realistic operating condition in larger networks. During the application of the 20 V pre-synaptic action potential to the memory TFT, current flows through the memristor. The amount its weight is decreased during this

Authorized licensed use limited to: University of Groningen. Downloaded on July 12,2020 at 09:25:05 UTC from IEEE Xplore. Restrictions apply.

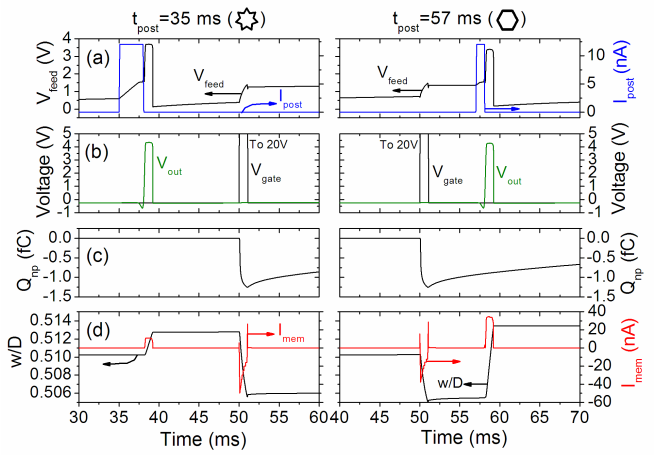

Fig. 4. Spike pair testing procedure. (a) The input voltage V_{feed} and injection current I_{post} of the post-synaptic neuron. (b) The output (V_{out}) voltage of the post-synaptic neuron along with the gate voltage applied to the TFT memory (V_{gate}). This voltage induces the trapped charge Q_{np} shown in (c). (d) The current through the memristive synapse I_{mem}, which relates to the change in the memristor state variable w/D. The left column is a post-pre pair resulting in depression and the right side is a potentiating pre-post pair (corresponding points on the STDP curve are highlighted by star and hexagon in Fig. 5b).

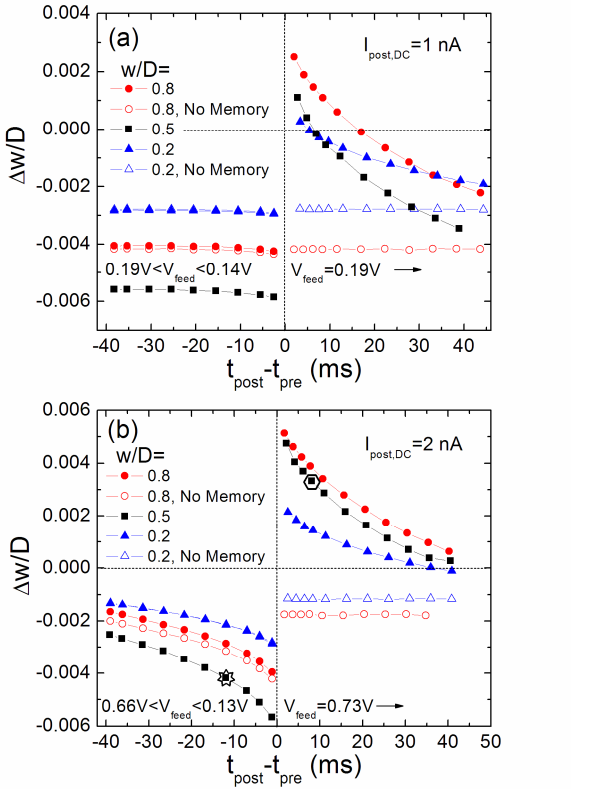

Fig. 5. STDP curves for various synaptic weight values (w/D) and with and without the memory device. (a) Direct current injection into the post-synaptic neuron is 1 nA, and (b) 2 nA, with the corresponding variation in V_{feed} at the time of the first spike labeled. Each data point is obtained from individual simulations like those shown in Fig. 4.

time depends on the magnitude of V_{inj} (1.8 V) and the instantaneous value of V_{feed}. If V_{feed} is nearly 1.8 V, there is almost no voltage drop across the device and very little depression. Upon the termination of the pre-synaptic spike, the memory transistor channel is on for an extended period while the trapped charge leaks into the channel (Fig. 4c). The length of time after the pre-synaptic pulse that the post-synaptic neuron fires changes the amount of reverse current through the memristor (Fig. 4d) as the channel conductance of the memory TFT changes. That is why the magnitude of the weight increase varies with the amount of time following the pre-synaptic action potential (Fig. 4d).

Fig. 5 shows the results of performing many trials of post-pre (left column of Fig. 4) and pre-post (right column of Fig. 4) action potential pairs. Different DC injection currents of 1 nA and 2 nA are displayed in Figs. 5a and b, respectively, for various initial values of the synaptic weight w/D. The similarity to biological measurements [2] is striking when the effects of the memory transistor are included. It should be noted again that the weight change depends not only on the initial weight, but also on the voltage at the input to the post-synaptic neuron V_{feed} at t_{pre} (=50 ms). This value is labeled in the figures, and shows that more depression is initiated when the post-synaptic neuron is further from the firing threshold of ~1.6 V. With greater numbers of afferent neurons, the value of V_{feed} is actually more likely to fluctuate above approximately 1 V, increasing the overall amount of potentiation. Reducing the voltage required to "program" the memory TFT by decreasing the oxide thickness should also help to limit the amount of depression.

B. Spike Triplets

Synaptic weight changes resulting from spike triplets were obtained in a manner similar to the previous section. Temporal plots of the main circuit variables are shown in Fig. 6. These include the post-synaptic injection current I_{post} and the input voltage V_{feed} in Fig. 6a, and the output neuron and pre-synaptic voltage V_{out} and V_{gate} in Fig. 6b. The charge in the memory TFT and the memristor state variable w/D are then shown in Figs. 6c and d, respectively. A 2 nA DC injection was used in all cases.

Data for the full set of different triplets is shown in Fig. 7, with direct comparison to measured data [10] and the model of [12]. There is clearly a difference in the amount of potentiation induced, but generally a similar trend is observed. Again, since the amount that a pre-synaptic action potential depresses the synapse depends on V_{feed}, even stronger potentiation is likely in a larger network. In the same manner, reducing the value of V_{inj} will decrease the depression current through the memristor as well. Then, the drain of the feedback transistor which is currently diode-connected could be connected to a voltage source. The magnitude of that source would subsequently help control the amount of potentiation during a post-synaptic action potential. More discussion about reducing V_{inj} will be presented in section IV.

C. Frequency Dependence

A final simulation experiment tests the changes in synaptic weight as a function of the frequency of applied action potentials. With Δt set to specific positive values, the pairs are repeated ten times with different frequencies. The periods are kept larger than Δt so that the spike train is always pre-post alternating. However, the post spike may occur at different times during the duty cycle of t_{pre} depending on the

Authorized licensed use limited to: University of Groningen. Downloaded on July 12,2020 at 09:25:05 UTC from IEEE Xplore. Restrictions apply.

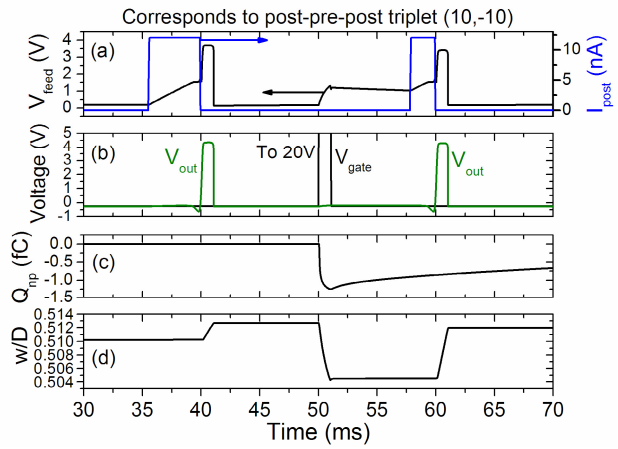

Fig. 6. Plots of circuit variables as a function of time during a (10,-10) post-pre-post triplet test. (a) The injection current I_{post} increases V_{feed} to the threshold, causing output action potentials at 40 and 60 ms. These are shown along with the pre-synaptic spike V_{gate} in (b). (c) The memory transistor particles are charged (programmed) by V_{gate} at 50 ms, and slowly decay back to the normal value. (d) Change in synaptic weight w/D versus time.

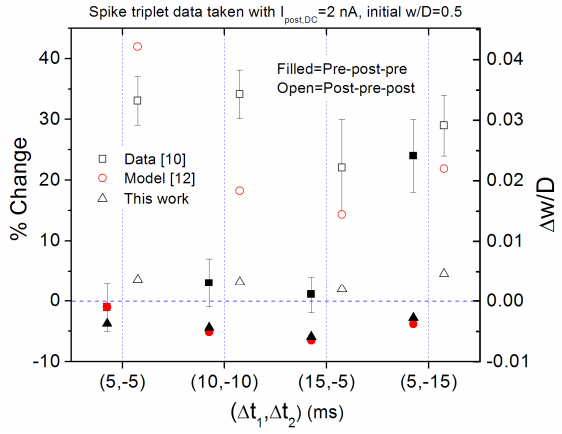

Fig. 7. Comparison of synaptic weight changes in triplet experiments between [10], [12], and this work. The biological data taken in hippocampus cells indicates very little synaptic depression.

frequency. For instance, two cases for Δt=10 ms are shown along with the corresponding synaptic weight changes in Fig. 8. In the high frequency case (top inset), most of the pairs appear as post-pre with the exception of the first (pre) and last (post) spikes. In the bottom inset, corresponding to lower frequency, all the pairs appear to be pre-post. Asymmetric temporal integration is again illustrated by this system since the weight change is not constant for a given Δt.

IV. Discussion

Although there are several drawbacks of simple pair-based STDP learning algorithms, it remains a fundamental first-order approximation of true learning rules. Specific applications of STDP include learning to recognize temporal patterns within large spike trains [22]. In addition, timing-dependent learning similar to STDP appears to be critical in the auditory system of many animals. It is used to filter out interaural time differences for sound localization [23]. Finally, there is evidence that STDP contributes to network stability by regulating the synaptic weights of loops

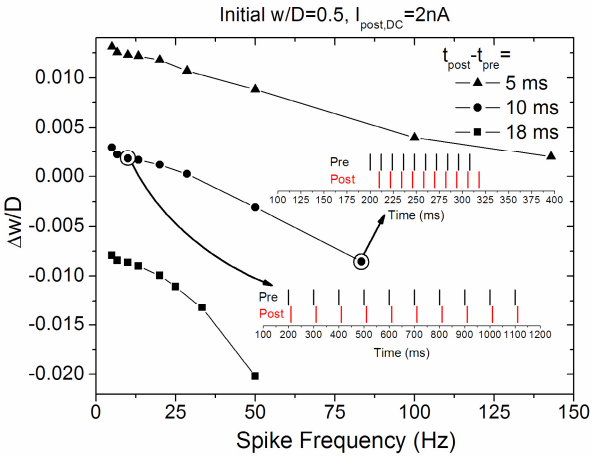

Fig. 8. Frequency dependence of synaptic weight change for three Δt values. Ten action potentials we applied in succession with the given offset and period (frequency) and the total weight change was measured.

within the network [24]. It is therefore not an inconsequential learning mechanism, and deserves significant attention and some form of implementation in artificial neural systems.

Aside from the questions surrounding the realities of STDP, further analysis is warranted regarding other related issues. One is the power dissipation of the circuits presented in this work. It can be seen in Fig. 4d that the current flow through the memristor during any action potential, given the voltages used, is on the order of 20 nA. Assuming the synaptic weights are all w/D=0.5, corresponding R_{mem}=25 MΩ, the average energy per synapse per spike is calculated to be 10 pJ. SPICE is then used to determine that the neuron circuits require approximately 20 nJ per neuron per spike. Now imagine a system on the scale of the human neocortex (10^{10} neurons [25] and 10^{14} synapses [26]) was to be built with overall average firing rate of 100 Hz. The total consumed power could approach 30 kW. This value is clearly far too large for any realistic electronic system, but it can easily be reduced by two orders of magnitude or more with some simple design alterations.

Reduction of the injection voltage V_{inj} would significantly diminish the current flowing through the memristive devices. It is currently set at 1.8 V because it must be greater than the threshold of the input inverter (1.6 V) in order for the post-synaptic neuron to fire. Reducing the value of V_{inj} to 0.18 V or less would be ideal from the power reduction perspective, but also from a fan-in and density viewpoint. A smaller value for C1 could be used to achieve the same voltage at V_{cap} (Fig. 1a). Solutions may come in the form of reduced device threshold voltage, or simple front-end amplification to translate ~100 mV at V_{cap} into 1.6 V to trip the inverter. Along the same lines, reducing the device threshold and optimizing the inverter transfer curve would allow much lower power supply voltages and switching currents. We are currently working on fabricating full circuits with high-quality nc-Si devices, and progressing toward device and circuit designs that provide much higher efficiency. A final point to be made is that even after significant improvements, the power dissipation might still be too great for a small area chip, but could easily be dissipated

Authorized licensed use limited to: University of Groningen. Downloaded on July 12,2020 at 09:25:05 UTC from IEEE Xplore. Restrictions apply.

over larger areas. This work is meant to show the conceptual possibilities for neural circuit learning using nanoscale devices with low-temperature processing, and does not imply these designs and results are final.

V. Conclusions

Neural circuits comprised of ambipolar nc-Si TFTs are simulated using SPICE. The synapses are memristors driven by a TFT memory device connected to the pre-synaptic neuron output. In this configuration, these compact circuits exhibit timing-dependent synaptic learning with biologically realistic action potential shapes. This is demonstrated through pair and triplet-based simulation experiments, as well as the frequency dependence of the weight changes. The significance of these results is highlighted by the use of materials and devices which could conceivably be used to build dense, large-scale systems with 3-D physical structure similar to the neocortex. Along with the simulated results for the circuits, multiple comparisons were made with results from neuroscience. In addition to the investigation of these learning mechanisms, considerable discussion was devoted to other important issues. These include improving the circuit performance for use in large networks, and reducing power dissipation and reliability concerns. Evidence was provided for the conclusion that these types of devices and systems could be significant contributors to future neuromorphic computing applications.

References

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

[2] G. Q. Bi and M. M. Poo, "Synaptic modification by correlated activity: Hebb's postulate revisited," *Annual Review of Neuroscience,* vol. 24, pp. 139-166, 2001.

[3] S. Song, K. D. Miller, and L. F. Abbott, "Competitive Hebbian learning through spike-timing-dependent synaptic plasticity," *Nature Neuroscience,* vol. 3, pp. 919-926, 2000.

[4] B. Linares-Barranco and T. Serrano-Gotarredona, "Exploiting memristance in adaptive asynchronous spiking neuromorphic nanotechnology systems," in *9th IEEE Conference on Nanotechnology*, Genoa, Italy, pp. 601-604, 2009.

[5] G. S. Snider, "Spike-timing-dependent learning in memristive nanodevices," in *IEEE International Symposium on Nanoscale Architectures (NANOARCH),* pp. 85-92, 2008.

[6] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, "Nanoscale Memristor Device as Synapse in Neuromorphic Systems," *Nano Letters,* vol. 10, pp. 1297-1301, 2010.

[7] H. Z. Shouval, S. S. H. Wang, and G. M. Wittenberg, "Spike timing dependent plasticity: a consequence of more fundamental learning rules," *Frontiers in Computational Neuroscience,* vol. 4, no. 19, 2010.

[8] J. Lisman and N. Spruston, "Questions About STDP as a General Model of Synaptic Plasticity," *Frontiers in Synaptic Neuroscience,* vol. 2, no. 140, 2010.

[9] H. Markram, J. Lubke, M. Frotscher, and B. Sakmann, "Regulation of Synaptic Efficacy by Coincidence of Postsynaptic APs and EPSPs," *Science,* vol. 275, pp. 213-215, 1997.

[10] H.-X. Wang, R. C. Gerkin, D. W. Nauen, and G.-Q. Bi, "Coactivation and timing-dependent integration of synaptic potentiation and depression," *Nature Neuroscience,* vol. 8, pp. 187-193, 2005.

[11] R. C. Froemke and Y. Dan, "Spike-timing-dependent synaptic modification induced by natural spike trains," *Nature,* vol. 416, pp. 433-438, 2002.

[12] C. Clopath and W. Gerstner, "Voltage and spike timing interact in STDP - a unified model," *Frontiers in Synaptic Neuroscience,* vol. 3, no. 25, 2010.

[13] C. G. Mayr and J. Partzsch, "Rate and pulse based plasticity governed by local synaptic state variables," *Frontiers in Synaptic Neuroscience,* vol. 2, no. 33, 2010.

[14] C.-H. Lee, A. Sazonov, and A. Nathan, "High hole and electron mobilities in nanocrystalline silicon thin-film transistors," *Journal of Non-Crystalline Solids,* vol. 352, pp. 1732-1736, 2006.

[15] K. D. Cantley et al., "Hydrogenated amorphous silicon nanowire transistors with Schottky barrier source/drain junctions," *Applied Physics Letters,* vol. 97, no. 143509, 2010.

[16] A. Subramaniam, K. D. Cantley, H. J. Stiegler, R. A. Chapman, and E. M. Vogel, "Ambipolar Nano-crystalline-silicon TFTs with Submicron Dimensions and Reduced Threshold Voltage Shift," in *69th Device Research Conference Digest*, Santa Barbara, CA, 2011.

[17] C. Mead, *Analog VLSI and Neural Systems*: Addison-Wesley, 1989.

[18] K. D. Cantley, A. Subramaniam, H. J. Stiegler, R. A. Chapman, and E. M. Vogel, "SPICE simulation of nanoscale non-crystalline silicon TFTs in spiking neuron circuits," in *53rd IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, Seattle, WA, pp. 1202-1205, 2010.

[19] K. D. Cantley, A. Subramaniam, H. Stiegler, R. Chapman, and E. Vogel, "Hebbian Learning in Spiking Neural Networks with Nano-Crystalline Silicon TFTs and Memristive Synapses," *IEEE Transactions on Nanotechnology,* DOI 10.1109/TNANO.2011.2105887, 2011.

[20] F. Alibart et al., "An Organic Nanoparticle Transistor Behaving as a Biological Spiking Synapse," *Advanced Functional Materials,* vol. 20, pp. 330-337, 2010.

[21] Y. Kuo and H. Nominanda, "Nonvolatile hydrogenated-amorphous-silicon thin-film-transistor memory devices," *Applied Physics Letters,* vol. 89, no. 173503, 2006.

[22] B. Linares-Barranco et al., "On Spike-Timing-Dependent-Plasticity, Memristive Devices, and building a Self-Learning Visual Cortex," *Frontiers in Neuroscience,* vol. 5, no. 26, 2011.

[23] W. Maass and C. M. Bishop, *Pulsed Neural Networks*. Cambridge, MA: The MIT Press, 1999.

[24] J. Kozloski and G. A. Cecchi, "A theory of loop formation and elimination by spike timing-dependent plasticity," *Frontiers in Neural Circuits,* vol. 4, no. 7, 2010.

[25] B. Pakkenberg and H. J. G. Gundersen, "Neocortical neuron number in humans: Effect of sex and age," *The Journal of Comparative Neurology,* vol. 384, pp. 312-320, 1997.

[26] Y. Tang, J. R. Nyengaard, D. M. G. DeGroot, and H. J. G. Gundersen, "Total regional and global number of synapses in the human brain neocortex," *Synapse,* vol. 41, pp. 258-273, 2001.

Authorized licensed use limited to: University of Groningen. Downloaded on July 12,2020 at 09:25:05 UTC from IEEE Xplore. Restrictions apply.