



A compound memristive synapse model for statistical learning through STDP in spiking neural networks

Johannes Bill* and Robert Legenstein

Faculty of Computer Science and Biomedical Engineering, Institute for Theoretical Computer Science, University of Technology, Graz, Austria

Edited by:

Christian G. Mayr, University of Zurich and ETH Zurich, Switzerland

Reviewed by:

Nabil Imam, Cornell University, USA

Love Mattias Cederström, Technische Universität Dresden, Germany

***Correspondence:**

Johannes Bill, Institute for Theoretical Computer Science, University of Technology, Inffeldgasse 16b/1, A-8010 Graz, Austria
e-mail: bill@tugraz.at

Memristors have recently emerged as promising circuit elements to mimic the function of biological synapses in neuromorphic computing. The fabrication of reliable nanoscale memristive synapses, that feature continuous conductance changes based on the timing of pre- and postsynaptic spikes, has however turned out to be challenging. In this article, we propose an alternative approach, the compound memristive synapse, that circumvents this problem by the use of memristors with binary memristive states. A compound memristive synapse employs multiple bistable memristors in parallel to jointly form one synapse, thereby providing a spectrum of synaptic efficacies. We investigate the computational implications of synaptic plasticity in the compound synapse by integrating the recently observed phenomenon of stochastic filament formation into an abstract model of stochastic switching. Using this abstract model, we first show how standard pulsing schemes give rise to spike-timing dependent plasticity (STDP) with a stabilizing weight dependence in compound synapses. In a next step, we study unsupervised learning with compound synapses in networks of spiking neurons organized in a winner-take-all architecture. Our theoretical analysis reveals that compound-synapse STDP implements generalized Expectation-Maximization in the spiking network. Specifically, the emergent synapse configuration represents the most salient features of the input distribution in a Mixture-of-Gaussians generative model. Furthermore, the network's spike response to spiking input streams approximates a well-defined Bayesian posterior distribution. We show in computer simulations how such networks learn to represent high-dimensional distributions over images of handwritten digits with high fidelity even in presence of substantial device variations and under severe noise conditions. Therefore, the compound memristive synapse may provide a synaptic design principle for future neuromorphic architectures.

Keywords: neuromorphic, synapse, synaptic plasticity, STDP, memristor, WTA, Bayesian inference, unsupervised learning

1. INTRODUCTION

A characteristic property of massively parallel computation in biological and artificial neural circuits is the need for intensive communication between neuronal elements. As a consequence, area- and energy consumption of neuromorphic circuits is often dominated by those circuits that implement synaptic transmission between neural elements (Schemmel et al., 2008). Biological synapses are highly dynamic computational entities, exhibiting plasticity on various time scales (Malenka and Bear, 2004). In order to capture the most salient of these aspects, silicon synapses thus demand extensive circuitry if implemented in CMOS technology. On this account, novel nanoscale circuit elements have recently gained interest in the field of neuromorphic engineering as a promising alternative solution for the implementation of artificial synaptic connections. In particular, for mixed-signal neuromorphic CMOS architectures, which combine traditional digital circuits with analog components, memristors are considered a promising class of circuit elements due to high integration densities, synapse-like plasticity dynamics, and low power

consumption (Choi et al., 2009; Jo et al., 2010; Kuzum et al., 2011; Indiveri et al., 2013). One particularly important feature of memristors is that their electrical resistance (often termed “memristance”) can be altered in a persistent manner by applying a voltage to its terminals, leading to the eponymous perception of memristors as resistors with memory. As an important application of this property, it was shown that the memristance can be changed based on the spike timings of the pre- and post-synaptic neurons in a manner that approximates spike-timing dependent plasticity (STDP), a plasticity rule that is believed to represent a first approximation for the changes of synaptic efficacies in biological synapses (Markram et al., 1997; Caporale and Dan, 2008; Markram et al., 2012). On the level of single synapses, this important property has been confirmed experimentally in real memristors (Mayr et al., 2012) and has been included into computational models of memristive plasticity (Serrano-Gotarredona et al., 2013). On the network level, models of memristive STDP were employed in computer simulations to demonstrate the potential applicability of neuromorphic designs

with memristive synapses to pattern recognition tasks (Querlioz et al., 2011).

In practice however, the production of functional memristive synapses with nanoscale dimensions has proven difficult, mainly due to large device variations and their unreliable behavior. It turned out to be particularly challenging to fabricate reliable nanoscale memristive synapses that feature a continuous spectrum of conductance values. As an alternative solution, it was proposed to employ bistable memristors as neuromorphic synapses instead since they exhibit a high degree of uniformity (Fang et al., 2011) and high durability (Jo et al., 2009b). For switching between their two stable conductance states, bistable memristors can be operated in a deterministic as well as in a stochastic regime (Jo et al., 2009b).

Using machine learning theory, we show in this article that stochastically switching bistable memristors become computationally particularly powerful in mixed-signal neuromorphic architectures when multiple memristors are combined to jointly form one synapse. Such a joint operation of multiple memristors can be interpreted as the collective function of ion channels in biological synapses (Indiveri et al., 2013). Concretely, we propose the *compound memristive synapse model* which employs M bistable memristors operating in parallel to form a single synaptic weight between two neurons. To implement synaptic plasticity, we employ standard STDP pulsing schemes (Querlioz et al., 2011; Serrano-Gotarredona et al., 2013) and exploit the stochastic nature of memristive switching (Jo et al., 2009b; Gaba et al., 2013; Suri et al., 2013; Yu et al., 2013). For the analysis of the resulting plasticity dynamics, we perceive individual memristors as binary stochastic switches. This abstract description was previously utilized to capture the most salient features of experimentally observed memristive switching (Suri et al., 2013) and appears compatible with pivotal aspects of the experimental literature (Jo et al., 2009b; Gaba et al., 2013).

We show analytically and through computer simulations that the change of the synaptic efficacy for a given pairing of pre- and postsynaptic spikes follows an STDP-like plasticity rule such that the expected weight change depends on the momentary synaptic weight in a stabilizing manner. The resulting *compound-synapse STDP* enables a synapse to attain many memristive states depending on the history of pre- and postsynaptic activity. A stabilizing weight dependence of synaptic plasticity exists in biological synapses (Bi and Poo, 1998) and has been shown to facilitate learning and adaptation in neural systems (Van Rossum et al., 2000; Morrison et al., 2007). In particular, it has been shown in Nessler et al. (2013) that in stochastic winner-take-all (WTA) architectures, STDP with stabilizing weight dependence implements an online Expectation-Maximization algorithm. When exposed to input examples, neurons in the WTA network learn to represent the hidden causes of the observed input in a well-defined generative model. This adaptation proceeds in a purely unsupervised manner. We adopt a similar strategy here and apply the compound memristive synapse model in a network of stochastically spiking neurons arranged in a WTA architecture. We show analytically that compound-synapse STDP optimizes the synaptic efficacies such that the WTA network neurons in the hidden layer represent the most salient features of the input

distribution in a Mixture-of-Gaussians generative model. After training, the network performs Bayesian inference over the hidden causes for the given input pattern. We show in computer simulations that such networks are able to learn to represent high-dimensional distributions over images of handwritten digits. After unsupervised training, the network transforms noisy input spike-patterns into a sparse and reliable spike code that supports classification of images. It turns out that even small compound synapses, consisting of only four bistable constituents per synapse, are sufficient for reliable image classification in our simulations. We furthermore show that the proposed model is able to represent the input distribution with high fidelity even in the presence of substantial device variations and under severe noise conditions. These findings render the compound synapse model a promising design principle for novel high-density, low-power mixed-signal CMOS architectures.

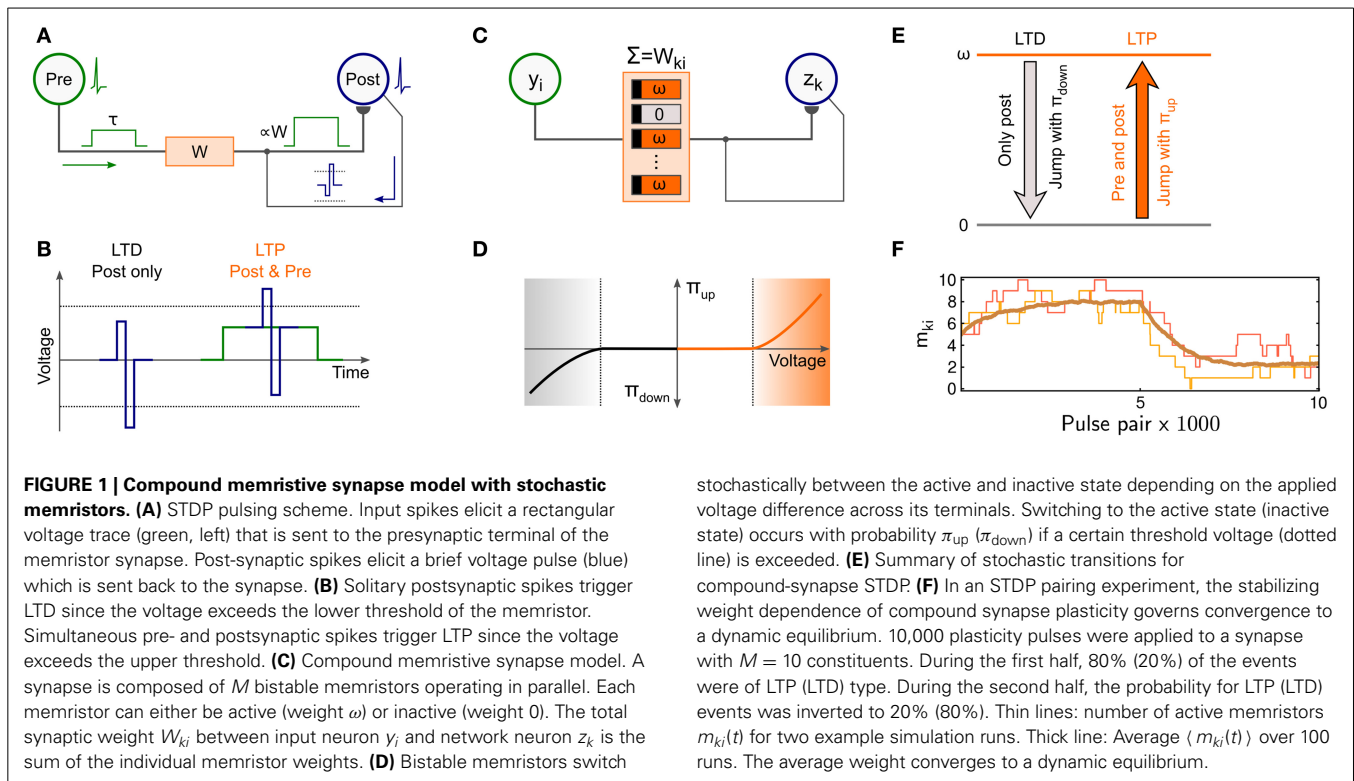
2. RESULTS

2.1. STOCHASTIC MEMRISTORS AS PLASTIC SYNAPSES

Memristors have gained increasing attention in neuromorphic engineering as possible substrates for plastic synapses (Jo et al., 2010) due to the possibility to change their electrical conductance without the requirement of extensive supporting circuitry. Recently, Zamarreño-Ramos et al. (2011) and Querlioz et al. (2011) have proposed a pulsing scheme to realize spike-timing dependent plasticity (STDP) with memristive synapses in response to pre- and postsynaptic activity as sketched in **Figure 1A**: Pre-synaptic spikes trigger a rectangular voltage pulse of duration τ (shown in green) that is sent to the memristor's presynaptic terminal. Similarly, postsynaptic neurons send back a copy of their spikes to the postsynaptic terminal as a brief voltage pulse (shown in blue). The combined effect of these pulses was shown to trigger STDP-type plasticity in the memristor as illustrated in **Figure 1B**, where we adopted the convention to measure the voltage drop at the memristor as "presynaptic minus postsynaptic potential." If only the postsynaptic neuron spikes (**Figure 1B**, left), the voltage pulse exceeds the lower threshold of the memristor, leading to long-term depression (LTD). Conversely, if a postsynaptic spike follows a presynaptic spike within duration τ (**Figure 1B**, right), the combined voltage trace exceeds the memristor's upper threshold, thus resulting in long-term potentiation (LTP). Pre-synaptic pulses alone do not trigger any plasticity. Overall, the memristor's conductance obeys an STDP-type plasticity rule.

The direct implementation of the above plasticity rule in mixed-signal neuromorphic architectures, however, faces practical challenges due to the continuous spectrum of the conductance values it relies on. Memristors that support a (quasi-) continuous spectrum of memristive states often suffer from instabilities to maintain their conductance value ("volatility") and typically show unreliable changes under repetitive application of identical pulses.

Here we explore an alternative approach that employs multiple bistable memristors, that support only two distinct conductance states per memristor, to jointly form one synapse. Such devices were reported to exhibit a high degree of uniformity (Fang et al., 2011). Concretely, we propose a synapse model which employs



M bistable memristors operating in parallel to form a single synaptic weight W_{ki} between the i -th input neuron y_i and the k -th network neuron z_k . The model which we refer to as *compound memristive synapse* is sketched in **Figure 1C**. Each memristor is assumed to provide two stable states: a high-conductive (active) state and a low-conductive (inactive) state. Since the dynamic range of memristors typically covers several orders of magnitude, the weight contribution of inactive memristors is almost negligible. In line with this notion, each inactive memristor contributes weight 0 in the synapse model and each active memristor contributes weight ω . Since parallel conductances sum up, the total weight of the compound memristive synapse reads

$$W_{ki} = \omega \cdot m_{ki} \quad (1)$$

with $m_{ki} \in \{0, 1, \dots, M\}$ denoting the number of active memristors. As a consequence, a compound memristive synapse supports $M + 1$ discrete weight levels, ranging from 0 to the maximum weight $W_{\max} = \omega \cdot M$.

Plasticity in this synapse model naturally emerges from transitions between the active and inactive state of the bistable constituents. However, deterministic transitions, which are, for instance, desirable in memristor-based memory cells, impair the performance in a neuromorphic online learning setup with compound memristive synapses: If all constituents of a synapse, that experience the same pre- and postsynaptic spikes, change their state simultaneously, the compound weight toggles between $W_{ki} = 0$ and $W_{ki} = W_{\max}$ depending on the latest pulse pair, not showing any gradual trace of memory formation as required for STDP.

A possible remedy to this issue can be found in memristors that exhibit stochastic rather than deterministic switching between their stable states. Yu et al. (2013), for instance, reported stochastic transitions in $\text{HfO}_x/\text{TiO}_x$ memristors (from the class of anion-based memristors) and explored in computer simulations how stochastic bistable memristors could be used in a neuromorphic learning architecture. Similar studies explored the usability of stochastically switching bistable cation-based memristive materials (Jo et al., 2009b; Gaba et al., 2013; Suri et al., 2013). In these nanoscale devices, changes in the memristance were shown to be dominated by the formation of a single conductive filament (Jo et al., 2009b). Stochastic switching was demonstrated for both directions (active \leftrightarrow inactive) (Suri et al., 2013) with switching probabilities being adjustable via the duration (Gaba et al., 2013; Suri et al., 2013) and amplitude (Jo et al., 2009b; Gaba et al., 2013) of the voltage applied across the terminals. These observations have led to the conclusion that “switching can be fully stochastic” with switching probabilities being almost unaffected by the rate at which consecutive plasticity pulses are delivered (Gaba et al., 2013). Furthermore, bistable memristors can be extremely durable, not showing any notable degradation over hundreds of thousands of programming cycles (Jo et al., 2009b). Owing to these pivotal properties, Suri et al. (2013) proposed an STDP-type plasticity rule that perceives bistable memristors as simple stochastic switches.

Here we generalize this idea to compound memristive synapses and investigate the computational function of the arising plasticity rules in spiking networks from a machine learning perspective. **Figure 1D** illustrates a simple model of stochastic switching of individual memristors that employs the STDP-pulsing scheme

from Zamarreño-Ramos et al. (2011) and Querlioz et al. (2011) discussed above: If the voltage between the pre- and postsynaptic terminal of a device exceeds a certain threshold (dotted lines) for the duration of the back-propagating spike signal, stochastic switching may occur. Inactive memristors jump to the active state with probability π_{up} provided a sufficiently strong positive voltage, resulting in stochastic LTP. Similarly, active memristors turn inactive with probability π_{down} given a sufficiently strong negative voltage, leading to stochastic LTD. No switching occurs if only a small (or zero) voltage is applied, or if the memristor is already in the respective target state. Since the applied pre- and postsynaptic pulse amplitudes are free parameters in the model, the jumping probabilities π_{up} and π_{down} can be controlled, to a certain extent, by the experimenter. This simple model captures the most salient aspects of stochastic switching in bistable memristive devices as discussed above, cp. also Suri et al. (2013) and the Discussion section. In order to distinguish the abstract memristor model from physical memristive devices, we will in the following refer to the model memristors as “stochastic switches,” “constituents,” or simply as “switches” for the sake of brevity. Furthermore, we refer to the resulting stochastic plastic behavior of compound memristive synapses in response to pre-post spike pairs, summarized in **Figure 1E**, as *compound-synapse STDP*.

From the transition probabilities of individual stochastic switches we can calculate the expected temporal weight change $\langle \frac{d}{dt} W_{ki} \rangle$ of the compound memristive synapse as a function of pre- and postsynaptic activity. Formally, we denote the presence of a rectangular input pulse (green in **Figure 1A**) of the i -th input by $y_i(t) = 1$ (and the absence by $y_i(t) = 0$). The brief pulses that are sent back from a postsynaptic neuron z_k to the synapse (blue in **Figure 1A**) are formally treated as point events at the spike times of the postsynaptic neuron. We denote the spike time of the f^{th} spike of neuron z_k by t_k^f . The spike train $s_k(t)$ of a neuron z_k is formally defined as the sum of Dirac delta pulses $\delta(\cdot)$ at the spike times: $s_k(t) = \sum_f \delta(t - t_k^f)$. When a synaptic efficacy W_{ki} is subject to a stochastic LTP update, there are $(M - m_{ki})$ constituents in the compound memristive synapse that are currently inactive and could undergo an LTP transition. Each constituent independently switches to its active state with probability π_{up} , thereby contributing ω to the compound weight W_{ki} . Hence, the expected weight change for the LTP condition reads $(M - m_{ki}) \omega \pi_{\text{up}}$. A similar argumentation applies to the LTD case. In summary, considering that plasticity always requires a postsynaptic spike, and that LTP is induced in the presence of a presynaptic pulse ($y_i(t) = 1$), while LTD is induced in the absence of a presynaptic pulse ($y_i(t) = 0$), the expected weight change of the compound memristive synapse reads:

$$\begin{aligned} \left\langle \frac{d}{dt} W_{ki} \right\rangle &= s_k(t) \cdot \left[\underbrace{(M - m_{ki}) \omega \pi_{\text{up}} y_i(t)}_{\text{LTP}} \right. \\ &\quad \left. - \underbrace{m_{ki} \omega \pi_{\text{down}} (1 - y_i(t))}_{\text{LTD}} \right] \\ &= s_k(t) \cdot [M \omega \pi_{\text{up}} y_i(t) - m_{ki} \omega \pi_{\text{up}} y_i(t)] \end{aligned} \quad (2)$$

$$\begin{aligned} &- m_{ki} \omega \pi_{\text{down}} + m_{ki} \omega \pi_{\text{down}} y_i(t)] \\ &= s_k(t) \cdot [W_{\text{max}} \pi_{\text{up}} y_i(t) - W_{ki} \pi_{\text{down}}] \\ &\quad + s_k(t) W_{ki} y_i(t) \cdot (\pi_{\text{down}} - \pi_{\text{up}}) \quad . \end{aligned}$$

In order to obtain a simple closed form solution of the weight changes, we set $\pi_{\text{up}} = \pi_{\text{down}}$, i.e., the probability of potentiation of a single switch under LTP equals its probability of depression under LTD. This choice will facilitate the theoretical analysis of learning in a spiking network, later on. In a hardware implementation, the switching probabilities could be adjusted via the pre- and postsynaptic amplitudes of the STDP pulses. We obtain the following closed form solution for the expected weight change:

$$\left\langle \frac{d}{dt} W_{ki} \right\rangle = \pi_{\text{up}} W_{\text{max}} \cdot s_k(t) \cdot [y_i(t) - W_{ki}/W_{\text{max}}] \quad (3)$$

Notably, the plasticity rule (3) differs from standard additive STDP rules in that it includes the weight dependent term W_{ki}/W_{max} . This weight dependence has its origin in the varying number of (in-)active stochastic switches m_{ki} that could actually undergo plastic changes and is in line with a prominent finding from neurobiology (Bi and Poo, 1998): Relative changes $\Delta W_{ki}/W_{ki}$ become weaker for strong weights under LTP, while under LTD the relative change is weight-independent. Studies in computational neuroscience, see e.g., Van Rossum et al. (2000), found that this type of weight dependence facilitates the formation of stable connections in spiking networks.

In **Figure 1F**, we illustrate the stochastic convergence of a compound synaptic weight to a stable, dynamic equilibrium in a simple STDP pairing experiment. The synapse consists of $M = 10$ bistable switches with switching probabilities set to $\pi_{\text{up}} = \pi_{\text{down}} = 0.001$. These synapse parameters will also be used in network level simulations, later on. In a small computer simulation, 5 of the 10 constituents are initially set active ($m_{ki}(t = 0) = 5$). Then, 5000 postsynaptic spikes are sent to the postsynaptic terminal of the synapse for triggering stochastic switching in the bistable constituents. 80% of these events are randomly paired with a presynaptic pulse, i.e., 80% of the events are of LTP type and 20% are of LTD type. After 5000 plasticity pulses, the statistics are inverted to 20% LTP and 80% LTD events for another 5000 plasticity pulses. The thin lines in **Figure 1F** show the evolution of $m_{ki}(t)$ in this STDP pairing experiment for two independent simulation runs. In the first half, the synaptic weight tends to settle around higher weight values, while in the second half, it stochastically declines toward lower weight values. The gradual convergence of the average weight, as suggested by Equation (3), becomes apparent by taking the mean over 100 simulation runs (thick line). The stabilizing weight dependence of the STDP rule leads to convergence to a dynamic equilibrium such that the mean value shows slow, continuous changes as expected from the theory. Individual synapses fluctuate stochastically around this mean.

Equation (3) is reminiscent of a theoretically derived synaptic plasticity rule for statistical model optimization in spiking neural networks proposed by Nessler et al. (2013). Building upon the theoretical approach developed by Nessler et al. (2013), we

will next turn to the question how to conceive stochastic learning with compound memristive synapses from a Bayesian perspective as unsupervised model optimization via a powerful optimization method that is known as Expectation-Maximization in the machine learning literature.

2.2. COMPOUND MEMRISTIVE SYNAPSES IN WINNER-TAKE-ALL NETWORKS

The winner-take-all (WTA) network structure is a ubiquitous circuit motif in neocortex (Douglas and Martin, 2004; Lansner, 2009) and is often utilized in neuromorphic engineering (Mead and Ismail, 1989; Indiveri, 2000). Recently, WTA networks attracted increasing attention in theoretical studies on statistical learning (Habenschuss et al., 2012; Keck et al., 2012; Habenschuss et al., 2013; Nessler et al., 2013; Kappel et al., 2014) because their comparatively simple network dynamics facilitate a comprehensive mathematical treatment. In this section, we introduce the WTA architecture that we will use to study the learning capabilities of spiking neural networks with compound memristive synapses subject to STDP.

The WTA network architecture is sketched in **Figure 2A**. The network consists of N spiking input neurons y_1, \dots, y_N and K spiking network neurons z_1, \dots, z_K with all-to-all connectivity in the forward synapses. Lateral inhibition introduces competition among the network neurons. Network neurons are stochastic spike response neurons (Gerstner and Kistler, 2002) with the membrane potential u_k of network neuron z_k being given by

$$u_k(t) = b_k + \sum_{i=1}^N W_{ki} \cdot y_i(t) . \quad (4)$$

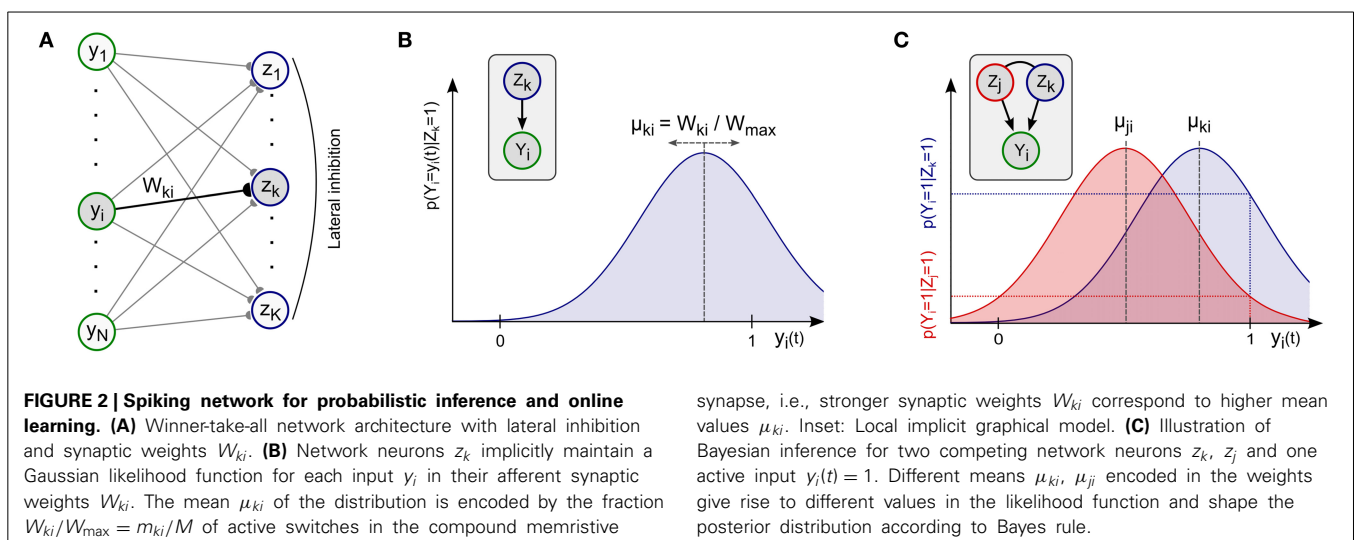
The membrane potential $u_k(t)$ integrates the inputs $y_i(t)$, i.e., the rectangular voltage pulses following each input spike, linearly through the synaptic weights W_{ki} . The parameter b_k denotes the intrinsic excitability of the neuron and controls its general disposition to fire. In Methods we outline how the linear membrane potential (4) can be realized with leaky integrators, a common neuron model in neuromorphic designs. For the spike response

of the network neurons z_k , a stochastic firing mechanism is employed. In the WTA network, neurons z_k spike in a Poissonian manner with instantaneous firing rate $\rho_k(t)$ that depends on the membrane potential $u_k(t)$ and on lateral inhibition $u_{\text{inh}}(t)$:

$$\rho_k(t) = r_{\text{net}} \cdot e^{u_k(t) - u_{\text{inh}}(t)} , \quad (5)$$

with a constant $r_{\text{net}} > 0$ that scales the overall firing rate of the network. In other words, the neuron spikes with probability $\rho_k(t) \cdot \delta t$ in a small time window $\delta t \rightarrow 0$. The inhibitory contribution $u_{\text{inh}}(t) := \log \sum_{j=1}^K \exp(u_j(t))$ summarizes the effect of lateral inhibition in the network and introduces WTA-competition between the network neurons to fire in response to a given stimulus $y_1(t), \dots, y_N(t)$. Notably, the exponential relationship (5) between an idealized membrane potential and neuronal firing is consistent with biological findings about the response properties of neocortical pyramidal neurons (Jolivet et al., 2006).

The feed-forward synapses from inputs y_i to network neurons z_k are implemented as compound memristive synapses and their synaptic weights W_{ki} are adapted through stochastic STDP as described above. The intrinsic excitabilities b_k are adapted according to a homeostatic plasticity rule (Habenschuss et al., 2012) that ensures that all network neurons take part in the network response and thus facilitates the emergence of a rich neural representation that covers the entire input space. Besides its observed stabilizing effect (Querlioz et al., 2011), homeostatic intrinsic plasticity plays a distinct computational role when combined with synaptic learning: Network neurons that maintain many strong synapses W_{ki} gain an “unrightful advantage” during WTA competition over neurons that are specialized on low-activity input patterns (and therefore maintain weaker weights). A detailed analysis of the learning dynamics in the network shows that, in order to compensate for this advantage, the former must be burdened with a lower excitability b_k than the latter. A formal definition of the homeostatic plasticity mechanism and a discussion of its computational role from a theory



perspective are provided in Methods, see also Habenschuss et al. (2012).

2.3. MEMRISTIVE SYNAPSES SUPPORT INFERENCE AND ONLINE LEARNING

In this section, we thoroughly analyze the learning effects of STDP in compound memristive synapses in the stochastic WTA network model. For the mathematical analysis, we describe the inputs $y_i(t)$ and the stochastic neuron responses $s_k(t)$ with the help of probability theory. To this end, we perceive the spiking activity of the input neurons y_i and network neurons z_k as samples of random variables (RVs) Y_i and Z_k respectively. Consistent with the assumption that spikes from input neurons produce voltage pulses of duration τ in the circuit implementation (see **Figure 1A**), we set $Y_i = y_i(t)$, i.e., $Y_i = 1$ if input neuron y_i spiked within $[t - \tau, t]$ and $Y_i = 0$ otherwise. The assignment of output spikes $s_k(t)$ to RVs Z_k is different: The random variable Z_k labels the winner of the WTA network at spike times of network neurons. Hence, the value of Z_k is only defined at the moments when one of the K network neurons spikes, i.e., when the spike train $s_j(t) \neq 0$ for some neuron z_j . In this case, Z_k encodes which neuron spiked, and we set $Z_k = 1$ if $k = j$ and $Z_k = 0$ if $k \neq j$.

Using this interpretation of neural activity as realizations of RVs, the network's stochastic response $s_k(t)$ to an input configuration $\mathbf{y}(t) = (y_1(t), \dots, y_N(t))$ gives rise to a conditional probability distribution $p_{\text{net}}(\mathbf{Z} | \mathbf{Y})$ over the network RVs $\mathbf{Z} := (Z_1, \dots, Z_K)$ conditioned on the input RVs $\mathbf{Y} := (Y_1, \dots, Y_N)$. In line with the definition of the RVs Z_k , the distribution $p_{\text{net}}(\mathbf{Z} | \mathbf{Y})$ describes the network response only when one of the network neurons z_k fires. The response distribution $p_{\text{net}}(\mathbf{Z} | \mathbf{Y} = \mathbf{y}(t))$ for any fixed input configuration $\mathbf{y}(t)$ can directly be calculated from Equations (4) and (5) (note that the probability $p_{\text{net}}(Z_k = 1 | \mathbf{Y} = \mathbf{y}(t))$ for an individual RV Z_k to be active is proportional to the firing rate $\rho_k(t)$ of neuron z_k):

$$\begin{aligned} p_{\text{net}}(Z_k = 1 | \mathbf{Y} = \mathbf{y}(t)) &= \rho_k(t) / r_{\text{net}} = e^{u_k(t) - u_{\text{inh}}(t)} \\ &= \frac{e^{b_k + \sum_{i=1}^N W_{ki} \cdot y_i(t)}}{\sum_{j=1}^K e^{b_j + \sum_{i=1}^N W_{ji} \cdot y_i(t)}}. \end{aligned} \quad (6)$$

Equation (6) fully characterizes the network response distribution $p_{\text{net}}(\mathbf{Z} | \mathbf{Y})$ for any given input $\mathbf{Y} = \mathbf{y}(t)$. We next turn to the question how the response distribution $p_{\text{net}}(\mathbf{Z} | \mathbf{Y})$ can be understood as the result of a meaningful probabilistic computation. Specifically, we will show that the spike response of the WTA network approximates the Bayesian posterior distribution during inference in a well-defined probabilistic model. This probabilistic model is implicitly encoded in the synaptic weights W_{ki} , and synaptic plasticity can thus be perceived as an ongoing refinement of the involved probability distributions. Indeed, the STDP rule (3) of the compound memristive synapses turns out to be optimal in the sense that it instantiates *generalized Expectation-Maximization* in the WTA network, a powerful algorithm for unsupervised learning from machine learning theory. Our findings build upon theoretical work on synaptic learning in spiking neural networks from Nessler et al. (2009, 2013) and Habenschuss et al. (2012, 2013).

The key idea for identifying the response distribution $p_{\text{net}}(\mathbf{Z} | \mathbf{Y})$ as the result of a Bayesian computation, is to hypothetically reverse the network computation and view the spike response of a network neuron z_k as the *hidden cause* behind the observed input $\mathbf{y}(t)$. In this view, the network is treated as a *generative model* that implicitly defines a prior distribution $p(\mathbf{Z})$ over hidden causes Z_k and a set of likelihood distributions $p(\mathbf{Y} | Z_k = 1)$, one for each hidden cause Z_k . The shape of the distributions is defined by the parameters of the network, e.g., the synaptic weights W_{ki} . An important property of these implicitly encoded distributions—that also motivates the term “generative model”—is that they give rise to a (hypothetical) distribution over the inputs, $p(\mathbf{Y} = \mathbf{y}(t)) = \sum_{k=1}^K p(\mathbf{Y} = \mathbf{y}(t) | Z_k = 1) \cdot p(Z_k = 1)$. This equation also explains why a RV Z_k is called a hidden cause: If we observe an input vector $\mathbf{y}(t)$, that has high probability only in one of the likelihood distributions $p(\mathbf{Y} = \mathbf{y}(t) | Z_k = 1)$, then we can consider the RV Z_k as a likely (but unobservable) cause for the observation according to the generative model. A common objective in machine learning theory, known as *Maximum-Likelihood learning*, is to find parameters that bring the implicit distribution $p(\mathbf{Y})$ of the generative model as close as possible to the distribution of the actually observed input. Then the hidden causes of the generative model are expected to represent important features of the observed input (e.g., some typical input clusters). Leaving the hypothetical generative perspective again, in the network's real operation an input $\mathbf{y}(t)$ is presented to the network and the hidden causes Z_k need to be inferred (e.g., the cluster the input $\mathbf{y}(t)$ belongs to). The mathematically correct result of this inference is given by Bayes rule

$$p(Z_k = 1 | \mathbf{Y} = \mathbf{y}(t)) \propto p(Z_k = 1) \cdot p(\mathbf{Y} = \mathbf{y}(t) | Z_k = 1) \quad (7)$$

which combines the likelihood $p(\mathbf{Y} = \mathbf{y}(t) | Z_k = 1)$ with the prior $p(Z_k = 1)$.

Nessler et al. (2013) showed that in a WTA network architecture, that evolves according to Equations (4) and (5), the synaptic weights W_{ki} can be understood as an implicit neural encoding of likelihood distributions $p(\mathbf{Y} | Z_k = 1)$, and that the network response $p_{\text{net}}(\mathbf{Z} | \mathbf{Y})$ approximates the posterior distribution according to Equation (7). Hence, WTA networks can be regarded as implicit generative models. Furthermore—and even more importantly from a theoretical perspective—Nessler et al. (2013) showed that the implicit likelihood model, that is encoded in the weights W_{ki} , can be optimized in an unsupervised manner by a weight-dependent STDP rule. Indeed, there exists a tight link between the type of weight dependence in the STDP rule and the type of implicit likelihood model it optimizes. The exponential weight dependence in Nessler's rule, however, differs from the linear weight dependence we identified for the plasticity rule in Equation (3). This raises the question what type of implicit likelihood model is encoded and optimized by the compound memristive synapses.

An intuition about an appropriate probabilistic interpretation of compound-synapse STDP can be obtained from the equilibrium points of the plasticity rule (3). We first observe, that

plasticity is always triggered by a postsynaptic spike, i.e., $Z_k = 1$ for some neuron z_k . In the spirit of spike-triggered averaging (Simoncelli et al., 2004), we can then study the conditional distribution $p(Y_i = y_i(t) | Z_k = 1)$ of an input y_i at the moment of the network response since the coincidence of pre- and postsynaptic spiking activity is the driving force behind any weight change in the STDP rule. By assuming that plasticity has converged to a dynamic equilibrium, the average contributions of LTP and LTD cancel each other, i.e., $\langle \frac{d}{dt} W_{ki} \rangle_{p(Y_i | Z_k = 1)} = 0$. From this condition, we obtain the following relation between the synaptic weight W_{ki} and the conditional input distribution $p(Y_i | Z_k = 1)$:

$$\begin{aligned} 0 &\stackrel{!}{=} \left\langle \frac{d}{dt} W_{ki} \right\rangle_{p(Y_i | Z_k = 1)} \\ &= \langle \pi_{\text{up}} W_{\text{max}} \cdot [Y_i - W_{ki}/W_{\text{max}}] \rangle_{p(Y_i | Z_k = 1)} \\ \Rightarrow W_{ki} &= W_{\text{max}} \cdot \langle Y_i \rangle_{p(Y_i | Z_k = 1)}. \end{aligned} \quad (8)$$

According to this analysis, the synaptic weight W_{ki} represents the expected value of input neuron y_i at the moment of a postsynaptic spike in network neuron z_k in a linear manner. In the compound memristive synapse, this expectation is encoded in the $M + 1$ possible weight states $W_{ki} = 0, \omega, \dots, W_{\text{max}}$ of the synapse. The linear encoding (8) is compatible with the convergence points which we observed previously in the small STDP pairing experiment in **Figure 1F**.

The above analysis only serves as an intuition and is no substitute for a thorough mathematical treatment of the learning process. A rigorous formal derivation that, for instance, also takes into account the dynamically changing response properties of the network neurons due to recurrent interactions and plastic changes in the weights, is provided in Methods. It reveals that the above intuition holds. More precisely, the likelihood distributions $p(\mathbf{Y} | Z_k = 1)$ that are optimized in a WTA circuit with compound-synapse STDP are given by the product of the likelihoods of individual inputs

$$p(\mathbf{Y} = \mathbf{y}(t) | Z_k = 1) = \prod_{i=1}^N p(Y_i = y_i(t) | Z_k = 1), \quad (9)$$

and the likelihood for each individual input channel y_i is given by a Gaussian distribution

$$p(Y_i = y_i(t) | Z_k = 1) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(y_i(t) - \mu_{ki})^2}{2\sigma^2}}. \quad (10)$$

The mean values μ_{ki} and the standard deviation σ of the likelihood distributions (10) are identified as

$$\mu_{ki} = W_{ki}/W_{\text{max}} = m_{ki}/M \quad \text{and} \quad \sigma = 1/\sqrt{(W_{\text{max}})}. \quad (11)$$

Hence, the mean μ_{ki} of the distribution for input channel y_i is given by the fraction of active constituents in the compound memristive synapse. The width $\sigma = 1/\sqrt{(W_{\text{max}})}$ of the distribution is determined by the maximum weight $W_{\text{max}} = M \cdot \omega$ and could, for instance, be controlled by the weight contribution ω of an individual stochastic switch. The resulting probabilistic model

of the WTA network is a *Mixture-of-Gaussians* generative model (see Methods for a formal definition).

In order to illustrate the computational properties of this generative model, the likelihood distribution $p(Y_i | Z_k = 1)$ for a single input y_i and a single active hidden cause $Z_k = 1$ is sketched in **Figure 2B**. An active hidden cause $Z_k = 1$ assigns probabilities to all possible instantiations $y_i(t)$ of Y_i . In principle, the Gaussian likelihood distribution supports arbitrary real-valued input instantiations $y_i(t) \in \mathbb{R}$. We will come back to this observation in the Discussion section where we address possible extensions of the WTA network to support more complex input types. In this article, we consider only binary inputs that take on the value 0 (input pulse absent) or 1 (input pulse present), see the presynaptic pulses in **Figure 1A**. The corresponding likelihood values $p(Y_i = 0 | Z_k = 1)$ and $p(Y_i = 1 | Z_k = 1)$ are determined by the mean μ_{ki} and the variance σ of the likelihood distribution, see Equations (10) and (11). The task of the network when presented with an input $\mathbf{y}(t)$ is to infer the posterior distribution over hidden causes $p(Z_k = 1 | \mathbf{Y} = \mathbf{y}(t))$ and produce spikes according to this distribution. The optimal solution is given by Bayes rule (7) with the likelihood $p(\mathbf{Y} = \mathbf{y}(t) | Z_k = 1)$ given by Equations (9) and (10), and an (input independent) a priori probability $p(Z_k = 1)$. As we prove in Methods, the response distribution $p_{\text{net}}(\mathbf{Z} | \mathbf{Y} = \mathbf{y}(t))$ of the spiking WTA network implements a close and well-defined variational approximation of this posterior distribution. A minimal example of such Bayesian inference is sketched in **Figure 2C**, where we consider a small WTA network with only one input y_i and two network neurons z_k and z_j . For a given input instantiation $y_i(t)$ the values $p(Y_i = y_i(t) | Z_k = 1)$ and $p(Y_i = y_i(t) | Z_j = 1)$ measure the likelihoods of the two competing hypotheses that neuron z_k or neuron z_j is the hidden cause of the observed input $y_i(t)$. These likelihood values shape the Bayesian posterior distribution (7) by contributing one factor to the product in Equation (9).

As a consequence, a network neuron z_k is particularly responsive to those input configurations $\mathbf{y}(t)$ that are associated with high likelihood values $p(\mathbf{Y} = \mathbf{y}(t) | Z_k = 1)$. The likelihood distributions are determined by the means μ_{ki} , i.e., by the number m_{ki} of active switches in the synapses that change according to the compound-synapse STDP rule. Through this mechanism, synaptic plasticity governs the emergence of prototypic patterns that the network neurons are most responsive to, and thereby turns each neuron z_k into a probabilistic expert for certain input configurations $\mathbf{y}(t)$. The aim of learning in the WTA network is to distribute the probabilistic experts z_k such that the likelihood of the presented input is (on average) as high as possible. This objective is an equivalent formulation of Maximum-Likelihood learning, and the *log-likelihood function*, which measures the average (logarithm of the) input likelihood, is a widely-used measure to determine how well a learning system is adapted to the presented input. Formally, the learning process can be described within the framework of generalized Expectation-Maximization (Dempster et al., 1977; Habenschuss et al., 2012; Nessler et al., 2013). In Methods we show that the generalized Expectation-Maximization algorithm is implemented in the WTA network via the interplay of the compound-synapse STDP rule (3), that adapts the synaptic weights W_{ki} , and the homeostatic intrinsic plasticity

rule, that regulates the intrinsic excitability b_k of the neurons such that each neuron maintains a long-term average firing rate. The interplay of synaptic and intrinsic plasticity achieves the aim of Maximum-Likelihood learning in the WTA network in the following sense:

The expected synaptic weight changes $\langle \frac{d}{dt} W_{ki} \rangle$ of the compound memristive synapses on average increase a lower bound of the log-likelihood function in a Mixture-of-Gaussians generative model during online learning until a (local) optimum is reached.

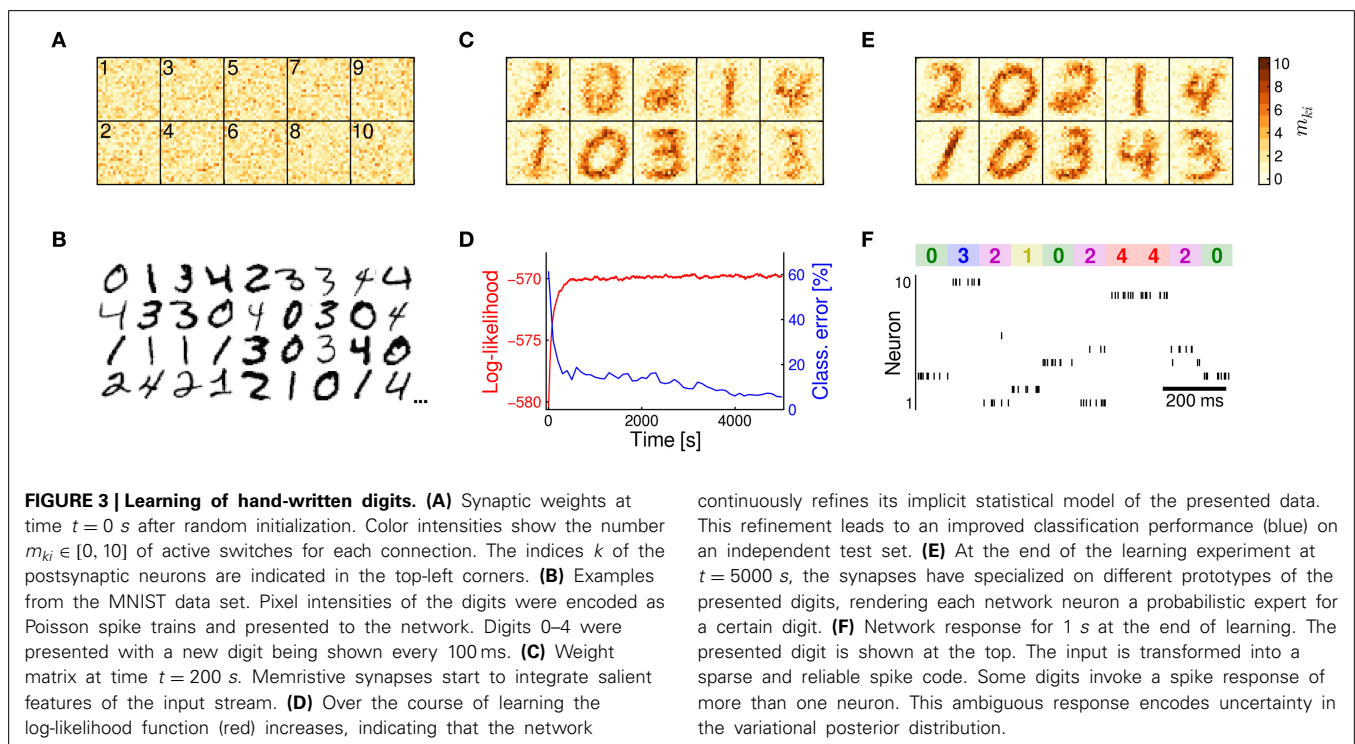
2.4. DEMONSTRATION OF UNSUPERVISED LEARNING

We tested the learning capabilities of the compound memristive synapse model in a standard machine learning task for hand-written digit recognition. In a computer simulation, we set up a WTA network with $N = 24 \times 24$ input neurons and $K = 10$ network neurons. Each synaptic weight W_{ki} was composed of $M = 10$ stochastic bistable switches, each contributing $\omega = 0.1$ in its active state. The switching probabilities $\pi_{\text{up}} = \pi_{\text{down}} = 0.001$ were set to a quite low value. This corresponds to a long integration time for gradual memory formation in order to assess the general ability of the synapse model in online learning tasks. Before training, the stochastic switches were initialized randomly as shown in **Figure 3A**. The network was then exposed to hand-written digits 0–4 from the MNIST training data set (LeCun et al., 1998). Examples from the data set are shown in **Figure 3B**. Each pixel was encoded by one input neuron y_i . Digits were presented as Poisson spike trains with firing rates depending on pixel intensities. Overall, the network was trained in an unsupervised setup for 5000 s with a new digit being presented every 100 ms. **Figure 3C** shows the weight matrix in an early stage of learning at $t = 200$ s. At this stage, the synapses begin to integrate

salient statistical features of the input, such as the generally low activity along the frame. Furthermore, a specialization to certain digit classes becomes apparent for some of the network neurons.

Over the course of learning, the synapses continuously improve the network's implicit generative model of the presented input. This refinement is reflected in the log-likelihood function shown in **Figure 3D** that measures how well the probabilistic model is adapted to the input distribution. The ongoing refinement also becomes apparent by a more intuitive—and practically more relevant—measure, namely the classification performance of the network on an independent test set of hand-written digits. The classification error (blue in **Figure 3D**; see Methods for details) continuously decreases as training progresses. The improved performance on an independent test set furthermore indicates that the network develops a generally well-suited representation of the input and evades the risk of over-fitting.

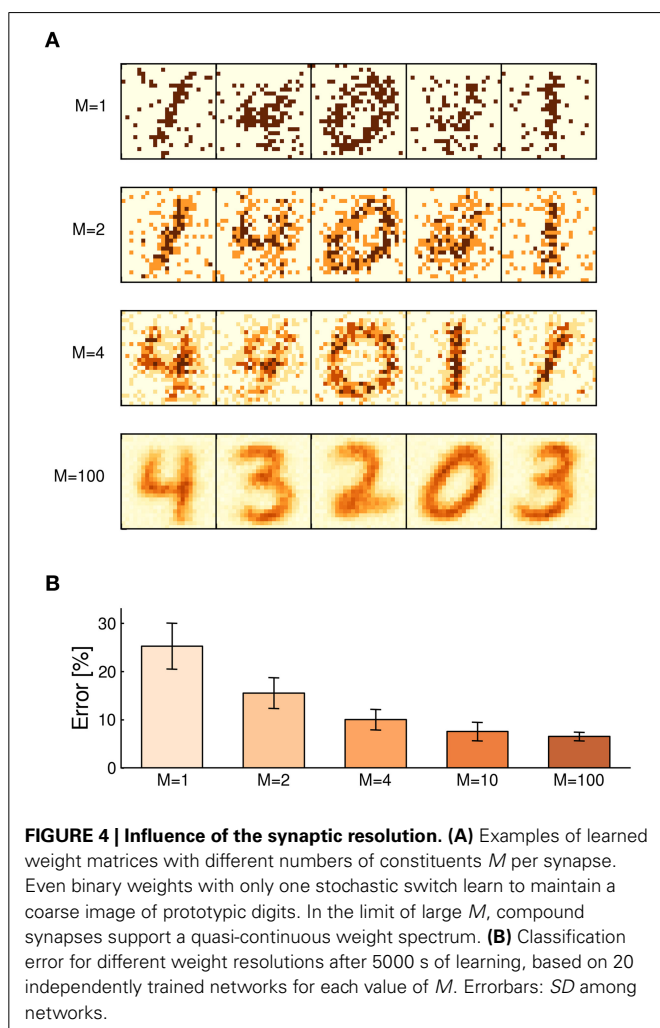
At the end of training, after $t = 5000$ s, a set of prototypic digits has emerged in the compound memristive synapses as shown in **Figure 3E**. The well-adapted synapse array turns each network neuron into a probabilistic expert for a certain digit class. As a consequence, the network has learned to transform the N -dimensional, noisy spike input into a sparse and reliable spike code, as shown in **Figure 3F**. Typically, exactly one network neuron z_k fires in response to the input. But also the seemingly unclear cases, when two neurons respond simultaneously, carry meaningful information in a Bayesian interpretation: Since network spikes approximate the posterior distribution $p(\mathbf{Z} | \mathbf{Y} = \mathbf{y}(t))$ through sampling, an ambiguous spike response encodes the level of uncertainty during probabilistic inference.



2.5. INFLUENCE OF SYNAPTIC RESOLUTION

In the previous section, we have demonstrated that the compound memristor synapse model is able to learn statistical regularities in the input stream and enables the spiking WTA network to perform probabilistic inference in a well-defined generative model. The demonstration employed compound synapses with $M = 10$ stochastic switches per synapse. Notably, the number of constituents M is a free parameter of the model and determines the weight resolution of the compound synapse. Increasing the synaptic resolution by recruiting more bistable switches per synapse is generally expected to improve the accuracy of the input representation, but comes at the cost of reduced integration density in a neuromorphic design. In the following, we therefore explore the opposite direction, i.e., unsupervised learning with a low weight resolution.

For estimating the influence of the weight resolution on the learning capabilities of the WTA network, we repeated the above computer simulation for different values of M , while holding the maximum weight $W_{\max} = \omega \cdot M$, and thus the variance $\sigma = 1/\sqrt{W_{\max}}$ of the implicit generative model, fixed. **Figure 4A** shows examples of the digits stored in the synapse array after 5000 s of learning for $M = 1, 2, 4$, and 100 stochastic



switches per synapse. Even binary synapses with $M = 1$ successfully identify noisy archetypes of the input digits. This observation is in line with previous studies on learning with binary weights (Fusi, 2002). The accuracy of the representation quickly increases with higher M -values. As an (academic) reference, we also included a simulation with $M = 100$ switches per synapse which support a quasi-continuous state spectrum.

The resulting ability of the WTA network to recognize handwritten digits is shown in **Figure 4B** in terms of the classification error on a test set. Each bar depicts the mean performance of 20 independently trained networks per M -value, errorbars show the standard deviation among networks. Taking the academic example with $M = 100$ as a reference for the performance achievable by the small WTA network, the computer simulations suggest that as few as $M = 4$ constituents per synapse may be sufficient for practical applications. While individual weights W_{ki} only store little information (ca. 2.3 bits in case of $M = 4$) about the expected input in channel y_i , the partial evidence received from each of the $N = 576$ input channels is integrated by the network in a statistically correct manner to form a sharply peaked posterior, most of the time.

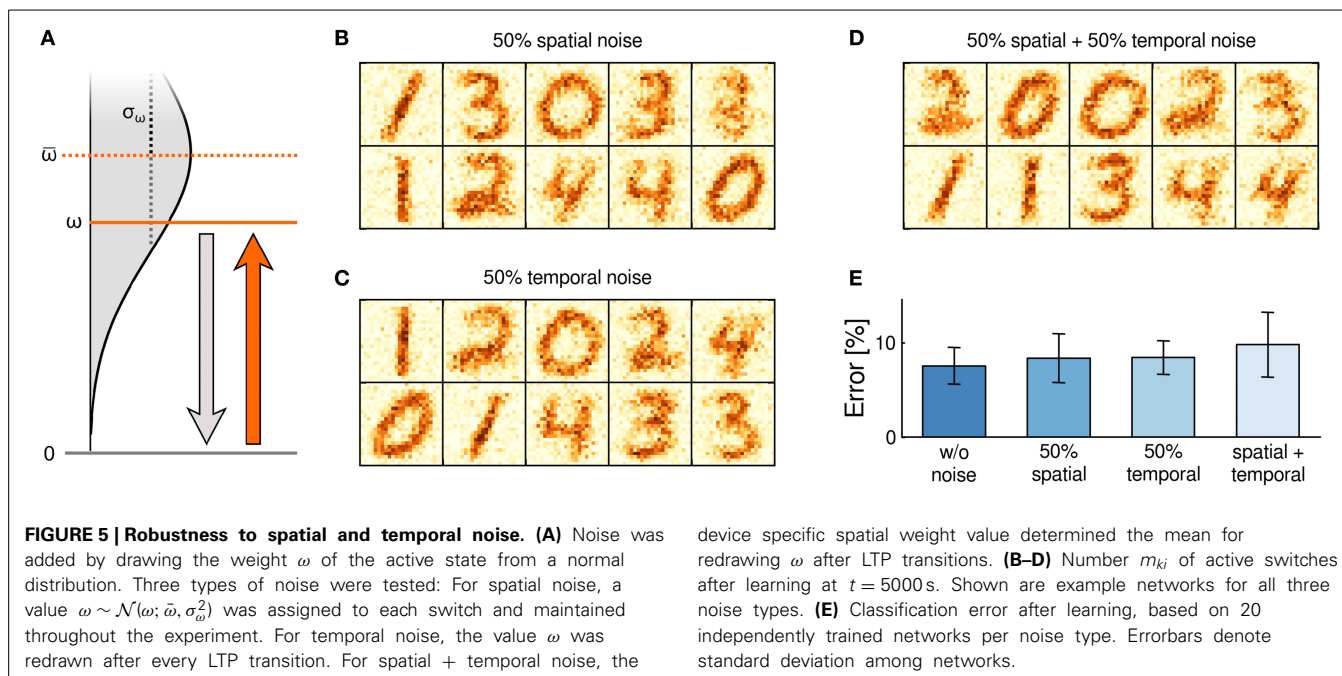
2.6. ROBUSTNESS TO DEVICE VARIATIONS

We have demonstrated so far that spiking networks with compound memristive synapses can learn a faithful representation of their input when synapses consist of idealized bistable constituents. Large-scale physical implementations, however, are likely to exhibit substantial device variabilities and imperfections. Plasticity in the compound synapse model depends on two device properties that are likely to be distorted in physical implementations: the conductance of each individual constituent ω and the switching probabilities $\pi_{\text{up}}/\pi_{\text{down}}$. In the following, we address the impact of distortions in these two properties on the WTA network learning capabilities, separately.

We first turned to the conductance value ω of individual switches and investigated the robustness of learning to two fundamentally different types of noise in ω , namely spatial noise and temporal noise:

- *Spatial noise* describes device-to-device variations and addresses peculiarities of individual memristors that remain stable over time.
- *Temporal noise* refers to trial-to-trial variations and covers device instabilities over the course of learning.

Both types of noise can be suspected to induce serious disturbances during learning: In case of spatial noise, although device-to-device variations could average out if many memristors are employed, any remaining deviations give rise to sustained systematic errors that may build up over the course of learning. In case of temporal noise, while trial-to-trial variations could average out over time, any synaptic update rests upon a disturbed instantiation of the weight matrix, i.e., on a noisy (and false) assumption. In computer simulations, we accounted for these types of noise separately by disturbing the active-state weight value ω of the switches as sketched in **Figure 5A**. To model spatial noise, the weight ω was randomly drawn prior to training for



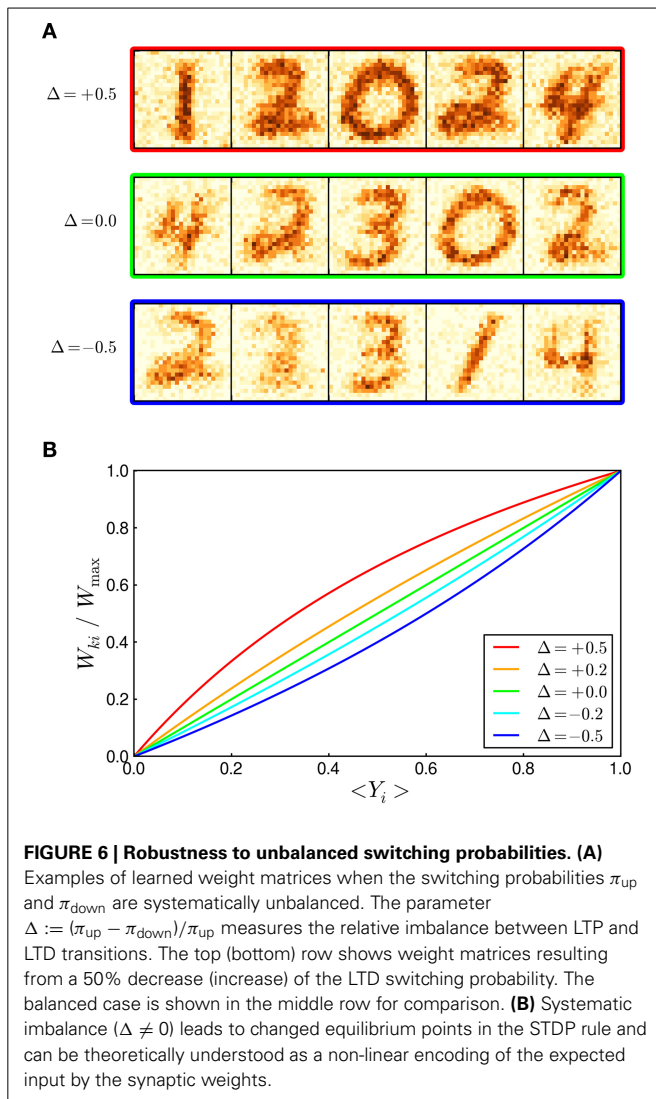
each stochastic switch from a normal distribution $\mathcal{N}(\omega; \bar{\omega}, \sigma_{\omega}^2)$ with mean $\bar{\omega}$ and standard deviation σ_{ω} . To capture the effect of temporal noise, in contrast, the weight ω was redrawn from $\mathcal{N}(\omega; \bar{\omega}, \sigma_{\omega}^2)$ whenever the constituent switched to its active state in an LTP transition. Furthermore, we examined the combined effect of both noise types being present simultaneously. In this combined case, the mean value for temporal noise was determined by the device-specific spatially perturbed weight value of each constituent. In any case, the range of perturbed weights was truncated to $\omega \geq 0$ to rule out negative conductances.

We repeated the experiment of **Figure 3** under each of these noise conditions. The mean $\bar{\omega} = 0.1$ was set to the undisturbed weight value of the previous, idealized experiment. The noise level was set to $\sigma_{\omega} = 0.05$, i.e., to 50% of the mean. Example cases of weight matrices after learning (shown are the m_{ki} 's from individual simulation runs) are presented in **Figures 5B–D** for the three noise conditions “spatial,” “temporal” and “spatial+temporal,” respectively. Surprisingly, hardly any difference to the idealized setup is observable. Nevertheless, under 20 repetitions of the learning simulation the detrimental influence of noise becomes visible in the classification performance (see **Figure 5E**) as noise appears to slightly increase the mean of the classification error. In summary, these results reveal a remarkable robustness of learning with compound memristive synapses to substantial device variability and severe temporal instability.

We next turned to the question how distorted switching probabilities π_{up} and π_{down} influence the learning dynamics in the WTA network. In the theory section, we had assumed that $\pi_{\text{up}} = \pi_{\text{down}}$, i.e., that the switching probabilities underlying LTP and LTD are balanced. This assumption, which could to some extent be achieved in a calibration step, yielded the elegant learning rule (3) and thereby facilitated the theoretical analysis. A physical implementation, however, will likely exhibit

unbalanced switching probabilities in the majority of memristors. We examined the influence of unbalanced switching, in two ways. First, we applied spatial noise to the switching probabilities of individual constituents by drawing π_{up} and π_{down} (separately) from normal distributions with 50% noise level (truncated to $0 \leq \pi_{\text{up}}, \pi_{\text{down}} \leq 1$). Thus, about half of the stochastic switches were more responsive to LTP pulses, the other half more to LTD pulses; even more, some of the constituents only showed switching in one direction, or were completely unresponsive. Nevertheless, synapses developed a faithful representation of prototypic digits in a repetition of the experiment in **Figure 3** (data not shown). Also the classification performance was only mildly impaired (*classification error*: $8.7 \pm 2.7\%$ based on 20 networks) compared to ideal, noise-free synapses (*classification error*: $7.5 \pm 1.9\%$).

In a second step, we pursued a slightly different—more principled—approach that permits a theoretical interpretation of how the altered synapse dynamics give rise to a different encoding of the expected input by the synaptic weights. Instead of drawing random parameters for each constituent, we systematically chose the LTD switching probability π_{down} larger (or smaller) than the LTP probability π_{up} throughout the synapse array. This systematic imbalance displays a worst-case scenario during learning since all synapses either favor (or suppress) LTD over LTP. We denote the relative imbalance between π_{up} and π_{down} by $\Delta := (\pi_{\text{up}} - \pi_{\text{down}})/\pi_{\text{up}}$. For instance, $\Delta = -0.5$ means that the probability for LTD transitions is 50% higher than for LTP transitions. **Figure 6A** shows examples of weight matrices after 5000 s of learning in a repetition of the experiment in **Figure 3**. In the top row, $\Delta = +0.5$, LTP transitions are favored over LTD transitions, resulting in generally stronger weights in comparison with balanced STDP ($\Delta = 0.0$, middle row). Conversely, in the bottom row, $\Delta = -0.5$, the systematic strengthening of LTD



leads to weaker weight patterns. Nevertheless, synaptic weight values converged to a dynamic equilibrium in either case since the STDP rule preserves its general stabilizing weight dependence. As can be expected from the prototypic digits that emerged in the weight matrices, the classification performance of the WTA networks was not considerably impaired by the unbalanced switching (classification errors estimated from 20 networks: $6.7 \pm 0.8\%$ for $\Delta = +0.5$; $7.5 \pm 1.9\%$ for $\Delta = 0$; $11.8 \pm 4.0\%$ for $\Delta = -0.5$). Indeed, positive Δ -values even performed slightly (but not significantly) better than balanced switching. A conceptual understanding of the altered learning dynamics can be obtained from the equilibrium points of the unbalanced STDP rule. The short calculation, that had led to Equation (8) for the balanced case, can be repeated for unbalanced switching probabilities. **Figure 6B** shows the resulting encoding of the expected input value $\langle Y_i \rangle_{p(Y_i|Z_k=1)}$ by the synaptic weight W_{ki} for different Δ -values. The example digits shown in panel A correspond to the red, green and blue graph in panel B, respectively. This analysis illustrates how unbalanced switching probabilities give rise to a non-linear encoding of the input in the WTA network.

In particular, it can be seen how the same expected input value $\langle Y_i \rangle_{p(Y_i|Z_k=1)}$ leads to stronger weights for $\Delta > 0$, and weaker weights for $\Delta < 0$.

3. DISCUSSION

We have proposed the compound memristive synapse model for neuromorphic architectures that employs multiple memristors in parallel to form a plastic synapse. A fundamental property of the synapse model is that individual memristors exhibit stochastic switching between two stable memristive states rather than obeying a deterministic update rule. Yet, the expected weight change of the compound memristive synapse, as it arises from the stochastic switching of its constituents, yielded an STDP-type plasticity rule with a stabilizing, linear weight dependence. We examined the computational capabilities of the compound-synapse STDP rule in WTA networks, a common circuit motif in cortical and neuromorphic architectures, by analyzing the network and synapse dynamics from the perspective of probability theory and machine learning. The comprehensive mathematical treatment revealed that compound memristive synapses enable a spiking network to perform Bayesian inference in and autonomous statistical optimization of a Mixture-of-Gaussians generative model via generalized Expectation-Maximization. Accompanying computer simulations demonstrated the practical capability of the synapse model to perform unsupervised classification tasks and, furthermore, revealed a remarkable robustness of the compound synapses to substantial device variations and imperfections.

3.1. COMPREHENSIVE LEARNING THEORY OF MEMRISTIVE PLASTICITY

Our work contributes a theoretical foundation for memristive learning in neural networks to the endeavor to employ memristors as plastic synapses in self-calibrating systems. Snider (2008); Querlioz et al. (2011), and Serrano-Gotarredona et al. (2013) have investigated how different pre- and postsynaptic waveforms can shape a memristive STDP learning window. Jo et al. (2010) and Mayr et al. (2012) have demonstrated STDP-type plasticity in Ag/Si and BiFeO₃ memristors. Yu et al. (2013) reported stochastic switching between stable states in oxide-based memristive synapses. Gaba et al. (2013) studied the parameter dependence of switching probabilities in metal filament based memristors, indicating a renewal process that is independent of the overall network firing rate. A strategy for integrating nanoscale memristive synapses into a hybrid memristor-CMOS network architecture was proposed by Indiveri et al. (2013). The beneficial contribution of stochasticity to learning with CMOS synapse circuits was explored by Chicca et al. (2014). Here we have established a firm link between the emergent synapse configurations observed in such architectures (see e.g., Querlioz et al., 2011) and a rigorous mathematical description of memristive learning on the system level using machine learning theory. Our findings on memristive learning from a Bayesian perspective build upon a series of theoretical contributions on synaptic learning in spiking neural networks: Nessler et al. (2009, 2013) identified a general link between STDP-type synaptic plasticity and statistical model optimization for probabilistic inference in WTA networks. Habenschuss et al. (2012) extended this work to incorporate also

homeostatic intrinsic plasticity, thereby overcoming several limiting assumptions on the input presentation. Finally, Habenschuss et al. (2013) investigated how the learning framework can be generalized to support a broad class of probability distributions. We expect that utilizing machine learning theory for describing the effects of specific memristor synapse models can significantly promote our understanding of memristive learning and its computational prospects.

3.2. HEADING FOR A FULL HARDWARE INTEGRATION

Plasticity in the compound memristor synapse model relies on stochastic transitions between two stable states. Such bistable devices—or more generally, devices with a clearly discrete state spectrum—were reported to exhibit a high degree of uniformity (Lee et al., 2006a; Fang et al., 2011) and temporal stability (Indiveri et al., 2013). Notably, the theoretical approach we have pursued in this work could likely be extended to cover memristors with more than two stable states and to support more complex input and plasticity mechanisms. For instance, the Gaussian likelihood distributions $p(Y|Z)$ identified in the present study, in principle support inference over arbitrary real-valued input states $y(t)$. Such states could arise if the input is presented in the form of exponentially decaying or additive postsynaptic potentials. Such more complex input types could afford more versatile STDP pulsing schemes, and the resulting memristor plasticity rules could likely be incorporated in an adapted model of statistical learning. The reason we restricted the input to binary values $y_i(t)$ is found in the STDP pulsing scheme that employs binary presynaptic waveforms. In this case, the theoretically derived generative model reveals how active and inactive inputs contribute to the network's spike response by means of a Gaussian likelihood distribution $p(Y_i = y_i(t) | Z_k = 1)$ that is sampled only at $y_i(t) = 0$ and $y_i(t) = 1$.

In this article, we have employed a simple model for stochastic switching in memristive devices where switching occurs with probabilities π_{up} , π_{down} which depend on the applied voltage difference across the memristor terminals. This phenomenological model captures the most salient aspects of switching in real memristive materials (Jo et al., 2009b; Gaba et al., 2013) and was used as an abstraction of memristive switching in a recent experimental study (Suri et al., 2013). In future research, it will be important to evaluate the effectiveness of this model either with physical memristors or in simulations based on detailed memristor models. The authors of Suri et al. (2013) raised the concern that the precise switching probabilities of individual devices are potentially hard to control in large-scale systems. In this regard, our simulation results indicate that learning with compound memristive synapses tolerates significant noise levels in the switching probabilities. We expect the origin for the observed robustness to be twofold: Firstly, imbalances in π_{up} , π_{down} between different constituents are expected to partly average out in compound synapses according to the central limit theorem; secondly, the stabilizing weight dependence of compound-synapse STDP ensures that even unbalanced switching leads to stable weight configurations, albeit with slightly shifted convergence points.

Another potential issue for learning with compound memristive synapses is the absolute value of the switching probability.

The product $\pi_{\text{up}} \cdot W_{\text{max}}$ can be linked to a learning rate in the theory domain (see **Table 1**) which controls how many samples from the input history are integrated into the implicit generative model during online learning. A slow and gradual memory formation, which is desirable for developing a representation of large and complex input data sets, relies on small learning rates, i.e., on small switching probabilities. It has to be seen if memristive materials that exhibit stochastic switching provide sufficiently small switching probabilities. A possible remedy in a hardware integration could be to multiplex the back-propagating signals from network neurons such that only a random subset of the memristors is notified of a network spike at a time (Fusi, 2002).

Regarding the physical model neurons of a hybrid memristor-CMOS architecture, two types of currents occur in the WTA network. The input integration via forward-synapses is spike based and could be realized with standard leaky integrators (see Methods). Lateral inhibition, in contrast, depends on the neuronal membrane potentials, and the involved inhibitory circuits should ideally transmit potentials instead of spikes. Alternatively, the effect of lateral inhibition could be approximated in a spike-based manner by populations of inhibitory neurons. Independent of the specific implementation of lateral inhibition, the resulting potential $u_k - u_{\text{inh}}$ controls the stochastic response of the WTA neurons that could either be implemented genuinely with a stochastic firing mechanism or be emulated with integrate-and-fire neurons (Petrovici et al., 2013).

3.3. INHERENTLY STOCHASTIC NATURE OF COMPOUND-SYNAPSE STDP

The spiking WTA network architecture with compound memristive synapses exploits stochasticity in various ways, in that the stochastic firing of network neurons in response to a transient input trajectory triggers stochastic STDP updates in the synaptic weights. From a learning perspective, the high degree of stochasticity contributes to the network's ongoing exploration for potential improvements in the parameter space. While the learning theory only guarantees convergence to a local optimum of the weight configuration, the stochastic nature of the ongoing exploration enables the network to evade small local optima in the parameter landscape, and thereby improves the robustness of learning (compared to traditional batch Expectation-Maximization).

For the derivation of the learning algorithm, we have focused on the weight-dependent STDP rule (3) which describes the expected temporal weight change $\langle \frac{d}{dt} W_{ki} \rangle$ of the compound synapse. The stochasticity of memristive switching, however, gives rise to a probability distribution over the weights, as well.

Table 1 | Correspondence of synapse parameters between the hardware and theory domain.

Parameter name	Hardware	Theory
Learning rate	$\pi_{\text{up}} \cdot W_{\text{max}}$	η_W
Max. weight W_{max}	$\omega \cdot M$	$1/\sigma^2$
Likelihood mean μ_{ki}	m_{ki}/M	$\sigma^2 \cdot W_{ki}$
Synaptic resolution	ω	$1/(M \cdot \sigma^2)$

Indeed, in equilibrium we expect that the number of active constituents m_{ki} follows a binomial-type weight distribution. This points to a potential knob for adjusting the amount of stochasticity used during online learning: when many memristors are recruited per synapse, i.e., for large M , we expect a reduced variance in the weight distribution.

Besides the level of stochasticity, the parameter M also controls the weight resolution of the compound synapse. In **Figure 4**, we have investigated the impact of the weight resolution on the learning capabilities of the WTA network. Notably, we observed that even with $M = 4$, i.e., with synapses that feature only 5 weight levels, the network performed reasonably well in the handwritten digit recognition task. The observation that even a low synaptic weight resolution can yield a satisfactory performance has important practical implications for nanoscale circuit designs, where integration density and power consumption impose crucial constraints, since the area allocated by the synapse array grows linearly in the synaptic size M . For instance, SRAM cells can be fabricated with a cell size of $0.127 \mu\text{m}^2$ (Wu et al., 2009), corresponding to a memory density of $\lesssim 1 \text{ Gb}/\text{cm}^2$. Importantly, this estimate does not include any additional plasticity circuits for implementing STDP or similar plasticity mechanisms. Functional Ag/Si memristive crossbars with $2 \text{ Gb}/\text{cm}^2$ memory density were demonstrated by Jo et al. (2009a), with densities up to $10 \text{ Gb}/\text{cm}^2$ being envisioned (Jo et al., 2009a,b). In the long term, memristive crossbars are expected to combine the advantages of SRAM and Flash memory regarding energy efficiency, non-volatility and integration density (Yang et al., 2013).

3.4. GENERALIZATION TO OTHER MATERIALS AND FUTURE RESEARCH

In recent years, a plethora of (in a broader sense) memristive materials has been discovered, and the characterization and refinement of their switching dynamics is evolving rapidly. At least four types of stochastically switching memristive devices can be distinguished: Switching in (1) anion-based (e.g., $\text{HfO}_x/\text{TiO}_x$ Yu et al., 2013) and (2) cation-based (e.g., Ag/GeS_2 Suri et al., 2013) devices mainly originates from conductive filament formation (Yang et al., 2013). In contrast, (3) single-electron latching switches [e.g., CMOS/MOLecular (CMOL) CrossNets Lee et al., 2006b] rely on electronic tunneling effects and, thus, their stochastic switching dynamics arise directly from the underlying physical process. Similarly, (4) magnetoresistive devices (e.g., spin-transfer torque magnetic memory (STT-MRAM) Vincent et al., 2014) can inherit stochastic switching dynamics from fundamental physical properties. Some manufacturing processes related to these ideas (like conductive-bridging RAM and STT-MRAM) reached already an early industrial stage, others are still primarily subject of academic research. While the microscopic origins of plasticity in these memristor types are fundamentally different, they all share stochastic, persistent switching between bistable memory states on a phenomenological level. We therefore believe that the compound memristive synapse model displays a promising concept for future work in diverse research fields.

Independent of the underlying switching mechanism, any nanoscale synaptic crossbar will likely exhibit imperfections and imbalances due to process variations. Here, we have investigated spatial and temporal noise in the weight values as well as

deviations in the switching probabilities under unbiased, uniform conditions. However, physical implementations can be expected to also suffer from more systematic imperfections, such as structural imbalances (e.g., one corner of the array being more reactive) or crosstalk between neighboring devices. While our computer simulations indicate a remarkable general robustness against device variations of various types, additional research is required to estimate the influence of such systematic, and potentially coupled, deviations.

3.5. CONCLUSIONS

In this article, we have introduced the compound memristive synapse model together with the compound-synapse STDP rule for weight adaptation. Compound-synapse STDP, a stabilizing weight-dependent plasticity rule, naturally emerges under a standard STDP pulsing scheme. In addition, by employing memristors with bistable memristive states, compound memristive synapses may circumvent practical challenges in the design of reliable nanoscale memristive materials. Both, our theoretical analysis and our computer simulations confirmed that compound-synapse STDP endows networks of spiking neurons with powerful learning capabilities. Hence, the compound memristive synapse model may provide a synaptic design principle for future neuromorphic hardware architectures.

4. METHODS

4.1. PROBABILISTIC MODEL DEFINITION

The probabilistic model that corresponds to the spiking network is a mixture model with K mixture components and Gaussian likelihood function. Formally, we define a joint distribution $p(\mathbf{Y} = \mathbf{y}(t), \mathbf{Z} = \mathbf{z}(t) | \boldsymbol{\theta})$ over K hidden binary random variables $\mathbf{Z} = (Z_1, \dots, Z_K)^T$ with values $z_k(t) \in \{0, 1\}$, and N real-valued visible random variables $\mathbf{Y} = (Y_1, \dots, Y_N)^T$ with values $y_i(t) \in \mathbb{R}$. The parameter set $\boldsymbol{\theta} = \{\hat{\mathbf{b}}, \mathbf{W}\}$ consists of a real-valued bias vector $\hat{\mathbf{b}} = (\hat{b}_1, \dots, \hat{b}_K)^T$ and a real-valued $K \times N$ weight matrix \mathbf{W} . The hidden RVs Z_k display an unrolled representation of a multinomial RV $\tilde{Z} \in \{1, \dots, K\}$ that enumerates the mixture components, and we identify $\tilde{Z} = k \Leftrightarrow Z_k = 1$, i.e., exactly one binary RV Z_k is active in the random vector \mathbf{Z} . In the following, we stick to the unrolled vector notation \mathbf{Z} , and, for readability, omit the time-dependent notation and further shorten the notation by identifying the RVs with their values, e.g., we write $p(\mathbf{z} | \boldsymbol{\theta})$ for $p(\mathbf{Z} = \mathbf{z}(t) | \boldsymbol{\theta})$.

The likelihood distribution fulfills the naïve Bayes property, i.e., all statistical dependencies between visible RVs y_i, y_j are explained by the hidden state \mathbf{z} . More precisely, the generative model has the following structure:

$$p(\mathbf{y}, \mathbf{z} | \boldsymbol{\theta}) = p(\mathbf{z} | \boldsymbol{\theta}) \cdot \prod_{k=1}^K \prod_{i=1}^N p(y_i | z_k = 1, \boldsymbol{\theta})^{z_k}, \quad (12)$$

with the prior

$$p(\mathbf{z} | \boldsymbol{\theta}) = \frac{e^{\mathbf{z}^T \hat{\mathbf{b}}}}{\sum_{j=1}^K e^{\hat{b}_j}} \quad (13)$$

and the likelihood

$$p(y_i | z_k = 1, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(y_i - \mu_{ki})^2}{2\sigma^2}}, \quad (14)$$

with σ^2 denoting an arbitrary (but fixed) variance which displays a constant in the model. Equation (14) defines a Gaussian likelihood model for each input RV y_i with mean μ_{ki} which is selected by the active hidden cause $z_k = 1$ in Equation (12). For theoretical considerations, it is convenient to reorganize Equation (14) according to its dependency structure:

$$p(y_i | z_k = 1, \theta) = \frac{e^{-y_i^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}} \cdot e^{\frac{\mu_{ki}}{\sigma^2} \cdot y_i} \cdot e^{-\mu_{ki}^2/(2\sigma^2)} \quad (15)$$

$$=: h(y_i) \cdot e^{W_{ki} \cdot y_i} \cdot e^{-A_{ki}}, \quad (16)$$

where we set $W_{ki} = \mu_{ki}/\sigma^2$ and $A_{ki} = \mu_{ki}^2/(2\sigma^2) = \sigma^2 W_{ki}^2/2$. The first factor does neither depend on the hidden causes z_k nor on the weight W_{ki} and will play no role during inference and learning. The second factor describes the coupling between the visible RV y_i and the active latent RV z_k through the mean $W_{ki} = \mu_{ki}/\sigma^2$. Finally, the third factor solely depends on the weight W_{ki} (and not on y_i) and ensures correct normalization of the distribution.

4.2. INFERENCE

The posterior distribution given an observation \mathbf{y} follows directly from Bayes rule:

$$p(\mathbf{z} | \mathbf{y}, \theta) = p(\mathbf{z} | \theta) \cdot p(\mathbf{y} | \mathbf{z}, \theta) / \text{Norm}. \quad (17)$$

$$= e^{\mathbf{z}^T \hat{\mathbf{b}}} \prod_{k=1}^K \prod_{i=1}^N h(y_i)^{z_k} \cdot e^{z_k \cdot W_{ki} \cdot y_i} \cdot e^{-z_k \cdot A_{ki}} / \text{Norm}. \quad (18)$$

$$= e^{\mathbf{z}^T \cdot [\hat{\mathbf{b}} - \mathbf{A} + \mathbf{W} \cdot \mathbf{y}]} \cdot \prod_{i=1}^N h(y_i) / \text{Norm}. \quad (19)$$

with $\mathbf{A} := (A_1, \dots, A_k, \dots, A_K)^T$ and $A_k := \sum_{i=1}^N A_{ki}$. We evaluate the posterior for a specific hidden RV z_k to be active and provide the normalization constant explicitly:

$$p(z_k = 1 | \mathbf{y}, \theta) = \frac{e^{\hat{b}_k - A_k + \sum_{i=1}^N W_{ki} \cdot y_i} \cdot \prod_{i=1}^N h(y_i)}{\sum_{j=1}^K e^{\hat{b}_j - A_j + \sum_{i=1}^N W_{ji} \cdot y_i} \cdot \prod_{i=1}^N h(y_i)} \quad (20)$$

$$= \frac{e^{\hat{u}_k}}{\sum_{j=1}^K e^{\hat{u}_j}} \quad (21)$$

where we defined $\hat{u}_k = \hat{b}_k - A_k + \sum_{i=1}^N W_{ki} \cdot y_i$. The quantities \hat{u}_k are reminiscent of neuronal membrane potentials which consist of bias terms $\hat{b}_k - A_k$ and synaptic input $\sum_{i=1}^N W_{ki} y_i$. However, implicitly the bias terms depend on all afferent synaptic weights since $\hat{b}_k - A_k = \hat{b}_k - \frac{\sigma^2}{2} \sum_i W_{ki}^2$ and, thus, rely on information not locally available to the neurons. This issue will be

resolved in the context of learning: We will identify update rules for both biases and synapses which only use information available locally and thereby make a neural network implementation feasible.

4.3. LEARNING VIA GENERALIZED EXPECTATION-MAXIMIZATION

We investigate unsupervised learning of the probabilistic model based on generalized online Expectation-Maximization (EM) (Dempster et al., 1977), an optimization algorithm from machine learning theory. To this end, we impose additional constraints on the posterior distribution (Graça et al., 2007) which will enable a neural network implementation via homeostatic intrinsic plasticity (Habenschuss et al., 2012) and STDP-type synaptic plasticity (Habenschuss et al., 2013; Nessler et al., 2013). Since the derivation is almost identical to Habenschuss et al. (2012), we only outline the key steps and main results in the following and refer to Habenschuss et al. (2012) for the details.

The algorithmic approach rests upon the generalized EM decomposition:

$$\mathcal{F}(\mathbf{W}, q(\mathbf{z}|\mathbf{y})) = \mathcal{L}(\theta) - \langle D_{\text{KL}}(q(\mathbf{z}|\mathbf{y}) \| p(\mathbf{z}|\mathbf{y}, \theta)) \rangle_{p^*(\mathbf{y})} \rightarrow \text{E-step} \quad (22)$$

$$= \langle \log p(\mathbf{y}, \mathbf{z}|\theta) \rangle_{p^*(\mathbf{y})q(\mathbf{z}|\mathbf{y})} + \langle H(q(\mathbf{z}|\mathbf{y})) \rangle_{p^*(\mathbf{y})} \rightarrow \text{M-step} \quad (23)$$

with the log-likelihood $\mathcal{L}(\theta) = \langle \log p(\mathbf{y} | \theta) \rangle_{p^*(\mathbf{y})}$, the Kullback-Leibler divergence $D_{\text{KL}}(q(\mathbf{z}) \| p(\mathbf{z})) = \sum_{\mathbf{z}} q(\mathbf{z}) \cdot \log(q(\mathbf{z})/p(\mathbf{z}))$ and the entropy $H(q(\mathbf{z})) = -\sum_{\mathbf{z}} q(\mathbf{z}) \cdot \log q(\mathbf{z})$. The distribution $p^*(\mathbf{y})$ denotes the input distribution actually presented to the system. The distribution $p(\cdot | \theta)$ is the probabilistic model defined above. The distribution $q(\mathbf{z}|\mathbf{y})$ is called variational posterior and will ultimately be implemented by the spiking network. The short hand notation $\langle \cdot \rangle_{p^*(\mathbf{y})q(\mathbf{z}|\mathbf{y})}$ denotes the concatenated average $\langle \langle \cdot \rangle_{q(\mathbf{z}|\mathbf{y})} \rangle_{p^*(\mathbf{y})}$ with respect to the input distribution and the resulting variational posterior. In principle, the above decomposition holds for any choice of q , and since the Kullback-Leibler divergence in Equation (22) is strictly non-negative, the objective function \mathcal{F} is a lower bound of the log-likelihood \mathcal{L} . During optimization the algorithm will pursue two goals: to increase \mathcal{L} , i.e., to better adapt the probabilistic model to the data, and to keep $\langle D_{\text{KL}}(q \| p) \rangle$ small, i.e., to maintain a reliable approximation $q(\mathbf{z}|\mathbf{y})$ of the exact posterior $p(\mathbf{z} | \mathbf{y}, \theta)$.

We first impose a homeostatic constraint on the variational posterior $q(\mathbf{z}|\mathbf{y})$, namely that the long term average activation of any hidden RV z_k matches a predefined target value c_k (with $\sum_k c_k = 1$). Formally, we define a set of constrained distributions $\mathcal{Q} = \{q : \langle z_k \rangle_{p^*(\mathbf{y})q(\mathbf{z}|\mathbf{y})} = c_k \ \forall 1 \leq k \leq K\}$ and demand $q(\mathbf{z}|\mathbf{y}) \in \mathcal{Q}$. The optimization algorithm then relies on the joint application of an E(expectation)-step and an M(aximization)-step: During the E-step, we aim to minimize the Kullback-Leibler divergence with respect to $q \in \mathcal{Q}$ in Equation (22); during the M-step, we perform gradient ascent on $\langle \log p(\mathbf{y}, \mathbf{z} | \theta) \rangle$ with respect to the weights W_{ki} in Equation (23). The E- and M-step will be discussed separately.

The E-step is a constrained optimization problem, namely the minimization of $\langle D_{\text{KL}}(q \| p) \rangle$ such that $q \in \mathcal{Q}$, that can be

solved through Lagrange multipliers. Since we imposed K constraints (one per RV z_k), we need K Lagrange multipliers β_k . It turns out that the solution to this optimization problem simply adds the multipliers β_k to the biases $\hat{b}_k - A_k$ in Equation (20). This convenient result gives rise to the definition of intrinsic excitabilities $b_k := \hat{b}_k - A_k + \beta_k$ which unify biases and multipliers in a single quantity. Furthermore, it turns out that the optimal values of the β_k 's (and thus the b_k 's) can be determined via iterative update rules that solely rely on the hidden RVs z_k under the variational response $q(\mathbf{z}|\mathbf{y})$ and overwrite the non-local terms A_k . In summary, we obtain the variational posterior distribution q that solves the E-step:

$$q(\mathbf{z}|\mathbf{y}) = \frac{e^{u_k}}{\sum_{j=1}^K e^{u_j}} \quad \text{with} \quad u_k = b_k + \sum_{i=1}^N W_{ki} \cdot y_i \quad (24)$$

$$\Delta b_k \propto \langle c_k - z_k \rangle_{p^*(\mathbf{y})q(\mathbf{z}|\mathbf{y})} \quad (25)$$

The variational posterior in Equation (24) is described in terms of membrane potentials u_k which consist of synaptic input $\sum_i W_{ki} y_i$ and intrinsic excitabilities b_k . Equation (25) regulates the intrinsic excitabilities b_k in a homeostatic fashion: When the average response exceeds the target c_k , the excitability is reduced, and vice versa.

The M-step can be solved via gradient ascent on \mathcal{F} with respect to the weights W_{ki} . The variational posterior q is a constant during the M-step in EM, and thus the log-joint distribution $\log p(\mathbf{y}, \mathbf{z} | \boldsymbol{\theta})$ remains as the only W_{ki} -dependent term in Equation (23). By taking the derivative of the log-joint defined by Equation (12), (13), and (16) with respect to W_{ki} , we obtain:

$$\frac{d\mathcal{F}}{dW_{ki}} = \langle \partial_{W_{ki}} \log p(\mathbf{y}, \mathbf{z} | \boldsymbol{\theta}) \rangle_{p^*(\mathbf{y})q(\mathbf{z}|\mathbf{y})} \quad (26)$$

$$= \langle \partial_{W_{ki} z_k} \cdot \log p(y_i | z_k = 1, \boldsymbol{\theta}) \rangle_{p^*(\mathbf{y})q(\mathbf{z}|\mathbf{y})} \quad (27)$$

$$= \langle z_k \cdot (y_i - \sigma^2 \cdot W_{ki}) \rangle_{p^*(\mathbf{y})q(\mathbf{z}|\mathbf{y})} \quad (28)$$

The gradient with respect to the weights W_{ki} yields Hebbian-type update rules that use pre- (y_i) and post- (z_k) synaptic activity and the current weight W_{ki} given the input $p^*(\mathbf{y})$ and the variational response $q(\mathbf{z}|\mathbf{y})$. Importantly, only local information is required during the E- and M-step.

4.4. SPIKING NETWORK IMPLEMENTATION

The spiking neural network model instantiates Equation (24), (25), and (28), i.e., it represents the variational posterior $q(\mathbf{z}|\mathbf{y})$ for probabilistic inference through its spike response and implements the derived update rules for generalized online EM learning through intrinsic and synaptic plasticity.

Each of the RVs z_k is represented by one of the K network neurons, and each spike in the network is a sample from the variational posterior $q(\mathbf{z}|\mathbf{y})$ by identifying $z_k = 1$ for a spike of the k -th network neuron. By setting the instantaneous firing rate ρ_k to be

$$\rho_k = \lim_{\delta t \rightarrow 0} p(\text{spike in } [t, t + \delta t]) / \delta t = r_{\text{net}} \cdot e^{u_k - u_{\text{inh}}} \quad (29)$$

with $u_{\text{inh}} := \log \sum_{j=1}^K \exp(u_j)$ the network thus implements Equation (24) for any choice of r_{net} , i.e., $p_{\text{net}} = q$.

The learning rules (25) and (28) rely on expected values $\langle \cdot \rangle_{p^*(\mathbf{y})q(\mathbf{z}|\mathbf{y})}$. The expectations can be approximated from input samples $\mathbf{y} \sim p^*(\mathbf{y})$ and posterior samples $\mathbf{z} \sim q(\mathbf{z}|\mathbf{y})$ in response to this input. The input vector \mathbf{y} is defined at any time t in the network as it measures the instantaneous presence or absence of rectangular input pulses. Samples of the latent variable \mathbf{z} , in contrast, are only defined at the spike times of the network. Hence integrating expected values $\langle z_k \rangle$ from the spike response can be expressed most conveniently in terms of the spike train function $s_k(t) = \sum_f \delta(t - t_k^f)$ of the network neurons. We obtain the following plasticity rules:

$$\frac{db_k}{dt} = \eta_b \cdot [r_{\text{net}} \cdot c_k - s_k(t)] \quad (30)$$

$$\frac{dW_{ki}}{dt} = \eta_W \cdot s_k(t) \cdot [y_i - \sigma^2 \cdot W_{ki}] \quad (31)$$

with small learning rates η_b and η_W . The homeostatic rule (30) regulates the intrinsic excitabilities b_k such that the average target activations $\langle z_k(t) \rangle \approx c_k$ are maintained over the presentation of many different input patterns $\mathbf{y}(t) \sim p^*(\mathbf{y})$ in accordance with Equation (25), and thereby implements the E-step. Building on the network response shaped by the E-step, the synaptic rule (31) on average increases the objective function \mathcal{F} since synaptic changes $\frac{d}{dt} W_{ki}$ on average point in the direction of the \mathbf{W} -gradient of \mathcal{F} given by Equation (28), thereby implementing the M-step. Since synaptic updates rely on a (sufficiently) precise E-step which, in turn, needs to integrate any changes in the network response due to synaptic plasticity, homeostatic intrinsic plasticity is required to act on faster time scales than synaptic plasticity. As a consequence, the learning rate η_b will typically exceed the learning rate η_W in the spiking network implementation.

The homeostatic intrinsic plasticity rule (30) can readily be implemented by the spiking neurons: The intrinsic excitability b_k of each neuron increases linearly in time with a slow drift $\eta_b \cdot r_{\text{net}} \cdot c_k$ and is lowered abruptly by η_b at the spike times of neuron z_k . Similarly, mapping the synaptic plasticity rule (31) to the compound-synapse STDP rule (3) is straight-forward due to the structural equivalence of Equations (3) and (31): The learning rate η_W in the theory domain corresponds to $\pi_{\text{up}} \cdot W_{\text{max}}$ in the hardware domain, e.g., high jumping probabilities π_{up} and large weight contributions $\omega = W_{\text{max}}/M$ of individual stochastic switches lead to high learning rates η_W . Furthermore, the maximum weight W_{max} can directly be identified with the precision $1/\sigma^2$ of the likelihood distribution (14). In the theory domain, we know that $\mu_{ki} = \sigma^2 \cdot W_{ki}$, and hence, $\mu_{ki} = W_{ki}/W_{\text{max}} = m_{ki}/M$ for the compound synapses. Finally, due to the structural equivalence of Equations (3) and (31) we find that the compound memristor plasticity rule (3) inherits the convergence properties from the theoretically derived plasticity rule (31) during online learning. The resulting translation of memristor synapse parameters to the abstract model is summarized in **Table 1**.

4.5. COMPUTER SIMULATIONS

All computer simulations were performed with customized Python scripts. For the computer simulations, we employed a network architecture with $K = 10$ network neurons and $N = 24 \cdot 24 = 576$ inputs. The simulation time step was $\delta t = 1$ ms and the PSP time constant $\tau = 10$ ms. The overall network firing rate was set to $r_{\text{net}} = 100$ Hz, the homeostatic target activation uniformly to $c_k = 1/K$. Synapses were composed of $M = 10$ constituents with weight $\omega = 0.1$ each. Switching probabilities were set to $\pi_{\text{up}} = \pi_{\text{down}} = 10^{-3}$. This corresponds to a Gaussian likelihood model with variance $\sigma^2 = 1$ and learning rate $\eta_W = 10^{-3}$. The learning rate for homeostatic intrinsic plasticity was set to $\eta_b = 20 \cdot \eta_W$. For the simulations in **Figures 4–6**, certain parameters deviated from the above, depending on the simulation setup. For **Figure 6**, the switching probability $\pi_{\text{up}} = 10^{-3}$ was kept fixed, and π_{down} was adapted for different Δ -values. All other changes are described directly in the Results section.

For learning experiments, digits 0, 1, 2, 3, 4 were extracted in equal proportion from the MNIST training data set (LeCun et al., 1998). A frame of two pixels width was removed, leaving images of size 24×24 . The images (indexed by s) were scaled linearly to activity patterns $x_i^s \in [0.05, 0.9]$, with $i = 1, \dots, N$, which were presented to the network as follows. For given activity pattern $\mathbf{x}^s = (x_1^s, \dots, x_N^s)$, each input i spiked with probability $p_i^{\text{spike}} = 1 - (1 - x_i^s)^{\delta t/\tau}$ per time step δt . During training, a new activity pattern \mathbf{x}^s was randomly drawn from the training set every 100 ms. Each network was trained for 5000 s. To obtain the unweighted PSP values y_i , the resulting spike patterns were convolved with a box kernel of duration τ and amplitude 1, and then clipped to values $[0, 1]$. This defined the input $y(t)$, and thus (implicitly) the data distribution $y(t) \sim p^*(y)$. Notably, the spiking probability p_i^{spike} is chosen such that $\langle y_i \rangle = x_i^s$.

The estimate of the log-likelihood in **Figure 3D** was based on 5000 input samples $y(t)$, which were randomly drawn from the training data, and assumed a uniform prior $p(z_k = 1 | \theta) = 1/K$ in accordance with the homeostatic target activation. The classification performance in **Figures 3D, 4B, 5E** was determined as follows. For given configuration of the synapses and intrinsic excitabilities, 100 versions of each digit from the training data set were presented to the network for 1 s each. Each neuron was labeled to be tuned to the digit class it was most responsive to. Then 500 versions of each digit from the MNIST test data set were presented to the network for 1 s each. The network neuron that spiked most during the 1 s period determined the network's classification of the input digit. The classification error is the fraction of wrongly classified digits.

4.6. IMPLEMENTATION WITH LEAKY INTEGRATOR NEURONS

The idealized stochastic neurons in the WTA network model feature abstract membrane potentials u_k that integrate the input $y(t)$ through the weights W_{ki} linearly. In a hardware integration, however, synaptic weights arise from the conductance of memristors, and neurons are physical implementations based on capacitors and various other circuit elements. Here, we outline

one possible hardware integration and consider a leaky integrator with membrane potential U_k that obeys the following dynamics:

$$\tau_m \cdot \frac{dU_k}{dt} = -(U_k - B_k) + I_k/G_L, \quad (32)$$

with membrane time constant τ_m , leak conductance G_L , resting potential B_k and synaptic input current I_k . In the setup of **Figure 1A**, input spikes trigger a rectangular voltage pulse of duration τ and with amplitude U_{pre} . Denoting the conductance of the memristive synapse by G , this generates a synaptic current $I = U_{\text{pre}} \cdot G$. The equilibrium membrane potential (i.e., $\frac{d}{dt}U_k = 0$) under this current is $U_k = B_k + (U_{\text{pre}}/G_L) \cdot G$. For small membrane time constant $\tau_m \rightarrow 0$, e.g., for a small neuron capacitance, the fast membrane will closely resemble the rectangular presynaptic pulse shape, and the PSP amplitude ($U_k - B_k$) will be proportional to the weight G . This linear integration property of U_k also holds in case of multiple memristive synapses acting in parallel, and we find

$$U_k = B_k + \frac{U_{\text{pre}}}{G_L} \cdot \sum_i G_{ki} \cdot y_i. \quad (33)$$

Consequently, the membrane potential U_k of the leaky integrator matches the idealized membrane potential u_k employed in the Results section up to a linear function that serves to translate the voltage-based potential U_k to the unitless potential u_k . Using the membrane potential U_k , the exponential firing behavior (29) of the neurons could either be realized with an inherently stochastic firing mechanism. Alternatively, deterministic leaky integrate-and-fire neurons could be operated in a stochastic regime by adapting, for instance, the approach taken in Petrovici et al. (2013).

AUTHOR CONTRIBUTIONS

Johannes Bill and Robert Legenstein conceived the study, designed the experiments and wrote the manuscript. Johannes Bill developed the theory and performed the computer simulations.

FUNDING

Written under partial support of CHIST-ERA ERA-Net (Project FWF #1753-N23, PNEUMA) (Johannes Bill, Robert Legenstein) and the European Union project #604102 The Human Brain Project (HBP) (Robert Legenstein).

ACKNOWLEDGMENTS

We thank Radu Berdan, Alex Serb and Themis Prodromakis for valuable discussions on memristors, Giacomo Indiveri for initial discussions on the compound memristive synapse model, Georg Goeri for preparatory studies on robustness to device variability, and David Kappel as well as two anonymous reviewers for helpful comments on the manuscript. Johannes Bill thanks Bernhard Nessler and Stefan Habenschuss for all the fruitful discussions on the SEM theory and statistical learning in general.

REFERENCES

- Bi, G.-Q., and Poo, M.-M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* 18, 10464–10472.
- Caporale, N., and Dan, Y. (2008). Spike timing-dependent plasticity: a hebbian learning rule. *Annu. Rev. Neurosci.* 31, 25–46. doi: 10.1146/annurev.neuro.31.060407.125639
- Chicca, E., Stefanini, F., Bartolozzi, C., and Indiveri, G. (2014). Neuromorphic electronic circuits for building autonomous cognitive systems. *Proc. IEEE* 102, 1367–1388. doi: 10.1109/JPROC.2014.2313954
- Choi, H., Jung, H., Lee, J., Yoon, J., Park, J., Seong, D.-J., et al. (2009). An electrically modifiable synapse array of resistive switching memory. *Nanotechnology* 20:345201. doi: 10.1088/0957-4484/20/34/345201
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *J. R. Stat. Soc. B* 39, 1–38.
- Douglas, R. J., and Martin, K. A. (2004). Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* 27, 419–451. doi: 10.1146/annurev.neuro.27.070203.144152
- Fang, Z., Yu, H., Li, X., Singh, N., Lo, G., and Kwong, D. (2011). Hfox/tiox/hfox/tiox multilayer-based forming-free rram devices with excellent uniformity. *IEEE Electron Device Lett.* 32, 566–568. doi: 10.1109/LED.2011.2109033
- Fusi, S. (2002). Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates. *Biol. Cybern.* 87, 459–470. doi: 10.1007/s00422-002-0356-8
- Gaba, S., Sheridan, P., Zhou, J., Choi, S., and Lu, W. (2013). Stochastic memristive devices for computing and neuromorphic applications. *Nanoscale* 5, 5872–5878. doi: 10.1039/c3nr01176c
- Gerstner, W., and Kistler, W. M. (2002). *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511815706
- Graça, J. V., Ganchev, K., and Taskar, B. (2007). “Expectation maximization and posterior constraints,” in *Advances in Neural Information Processing Systems* (Vancouver).
- Habenschuss, S., Bill, J., and Nessler, B. (2012). “Homeostatic plasticity in bayesian spiking networks as expectation maximization with posterior constraints,” in *Advances in Neural Information Processing Systems* (Lake Tahoe), 773–781.
- Habenschuss, S., Puh, H., and Maass, W. (2013). Emergence of optimal decoding of population codes through stdp. *Neural Comput.* 25, 1371–1407. doi: 10.1162/NECO-a-00446
- Indiveri, G. (2000). Modeling selective attention using a neuromorphic analog vlsi device. *Neural Comput.* 12, 2857–2880. doi: 10.1162/089976600300014755
- Indiveri, G., Linares-Barranco, B., Legenstein, R., Deligeorgis, G., and Prodromakis, T. (2013). Integration of nanoscale memristor synapses in neuromorphic computing architectures. *Nanotechnology* 24:384010. doi: 10.1088/0957-4484/24/38/384010
- Jo, S. H., Chang, T., Ebong, I., Bhadviya, B. B., Mazumder, P., and Lu, W. (2010). Nanoscale memristor device as synapse in neuromorphic systems. *Nano Lett.* 10, 1297–1301. doi: 10.1021/nl904092h
- Jo, S. H., Kim, K.-H., and Lu, W. (2009a). High-density crossbar arrays based on a si memristive system. *Nano Lett.* 9, 870–874. doi: 10.1021/nl8037689
- Jo, S. H., Kim, K.-H., and Lu, W. (2009b). Programmable resistance switching in nanoscale two-terminal devices. *Nano Lett.* 9, 496–500. doi: 10.1021/nl803669s
- Jolivet, R., Rauch, A., Lüscher, H.-R., and Gerstner, W. (2006). Predicting spike timing of neocortical pyramidal neurons by simple threshold models. *J. Comput. Neurosci.* 21, 35–49. doi: 10.1007/s10827-006-7074-5
- Kappel, D., Nessler, B., and Maass, W. (2014). Stdp installs in winner-take-all circuits an online approximation to hidden markov model learning. *PLoS Comput. Biol.* 10:e1003511. doi: 10.1371/journal.pcbi.1003511
- Keck, C., Savin, C., and Lücke, J. (2012). Feedforward inhibition and synaptic scaling—two sides of the same coin? *PLoS Comput. Biol.* 8:e1002432. doi: 10.1371/journal.pcbi.1002432
- Kuzum, D., Jeyasingh, R. G., Lee, B., and Wong, H.-S. P. (2011). Nanoelectronic programmable synapses based on phase change materials for brain-inspired computing. *Nano Lett.* 12, 2179–2186. doi: 10.1371/journal.pcbi.1002432
- Lansner, A. (2009). Associative memory models: from the cell-assembly theory to biophysically detailed cortex simulations. *Trends Neurosci.* 32, 178–186. doi: 10.1016/j.tins.2008.12.002
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2324. doi: 10.1109/5.726791
- Lee, D., Seong, D.-J., Jung Choi, H., Jo, I., Dong, R., Xiang, W., et al. (2006a). “Excellent uniformity and reproducible resistance switching characteristics of doped binary metal oxides for non-volatile resistance memory applications,” in *Electron Devices Meeting, 2006. IEDM’06. International* (San Francisco: IEEE), 1–4.
- Lee, J. H., Ma, X., and Likharev, K. (2006b). Cmol crossnets: possible neuromorphic nanoelectronic circuits. *Adv. Neural Inf. Process. Syst.* 18:755.
- Malenka, R. C., and Bear, M. F. (2004). Ltp and ltd: an embarrassment of riches. *Neuron* 44, 5–21. doi: 10.1016/j.neuron.2004.09.012
- Markram, H., Gerstner, W., and Sjöström, P. J. (2012). Spike-timing-dependent plasticity: a comprehensive overview. *Front. Synaptic Neurosci.* 4:2. doi: 10.3389/fnsyn.2012.00002
- Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic aps and epsps. *Science* 275, 213–215. doi: 10.1126/science.275.5297.213
- Mayr, C., Stärke, P., Partzsch, J., Cederstroem, L., Schüffny, R., Shuai, Y., et al. (2012). “Waveform driven plasticity in bifeo3 memristive devices: model and implementation,” in *Advances in Neural Information Processing Systems* (Lake Tahoe), 1700–1708.
- Mead, C., and Ismail, M. (1989). *Analog VLSI Implementation of Neural Systems*. Norwell, MA: Springer. doi: 10.1007/978-1-4613-1639-8
- Morrison, A., Aertsen, A., and Diesmann, M. (2007). Spike-timing-dependent plasticity in balanced random networks. *Neural Comput.* 19, 1437–1467. doi: 10.1162/neco.2007.19.6.1437
- Nessler, B., Pfeiffer, M., Buesing, L., and Maass, W. (2013). Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent plasticity. *PLoS Comput. Biol.* 9:e1003037. doi: 10.1371/journal.pcbi.1003037
- Nessler, B., Pfeiffer, M., and Maass, W. (2009). “Stdp enables spiking neurons to detect hidden causes of their inputs,” in *Advances in Neural Information Processing Systems* (Vancouver), 1357–1365.
- Petrovici, M. A., Bill, J., Bytschok, I., Schemmel, J., and Meier, K. (2013). Stochastic inference with deterministic spiking neurons. arXiv preprint arXiv:1311.3211.
- Querlioz, D., Bichler, O., and Gamrat, C. (2011). “Simulation of a memristor-based spiking neural network immune to device variations,” in *Neural Networks (IJCNN), The 2011 International Joint Conference on* (IEEE), 1775–1781.
- Schemmel, J., Fieres, J., and Meier, K. (2008). “Wafer-scale integration of analog neural networks,” in *Neural Networks, 2008. IJCNN 2008 (IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on* (Hong Kong: IEEE), 431–438.
- Serrano-Gotarredona, T., Masquelier, T., Prodromakis, T., Indiveri, G., and Linares-Barranco, B. (2013). Stdp and stdp variations with memristors for spiking neuromorphic learning systems. *Front. Neurosci.* 7:2. doi: 10.3389/fnins.2013.00002
- Simoncelli, E. P., Paninski, L., Pillow, J., and Schwartz, O. (2004). Characterization of neural responses with stochastic stimuli. *Cogn. Neurosci.* 3, 327–338.
- Snider, G. S. (2008). “Spike-timing-dependent learning in memristive nanodevices,” in *Nanoscale Architectures, 2008. NANOARCH 2008. IEEE International Symposium* (Anaheim, CA: IEEE), 85–92.
- Suri, M., Querlioz, D., Bichler, O., Palma, G., Vianello, E., Vuillaume, D., et al. (2013). Bio-inspired stochastic computing using binary cbram synapses. *Elect. Devices IEEE Trans.* 60, 2402–2409. doi: 10.1109/TED.2013.2263000
- Van Rossum, M. C., Bi, G. Q., and Turrigiano, G. G. (2000). Stable hebbian learning from spike timing-dependent plasticity. *J. Neurosci.* 20, 8812–8821.
- Vincent, A., Larroque, J., Zhao, W., Romdhane, N. B., Bichler, O., Gamrat, C., et al. (2014). “Spin-transfer torque magnetic memory as a stochastic memristive synapse,” in *Circuits and Systems (ISCAS), 2014 IEEE International Symposium* (Melbourne: IEEE), 1074–1077.
- Wu, S.-Y., Liaw, J., Lin, C., Chiang, M., Yang, C., Cheng, J., et al. (2009). “A highly manufacturable 28nm cmos low power platform technology with fully functional 64mb sram using dual/tripe gate oxide process,” in *VLSI Technology, 2009 Symposium on* (IEEE), 210–211.
- Yang, J. J., Strukov, D. B., and Stewart, D. R. (2013). Memristive devices for computing. *Nat. Nanotechnol.* 8, 13–24. doi: 10.1038/nnano.2012.240
- Yu, S., Gao, B., Fang, Z., Yu, H., Kang, J., and Wong, H.-S. P. (2013). Stochastic learning in oxide binary synaptic device for neuromorphic computing. *Front. Neurosci.* 7:186. doi: 10.3389/fnins.2013.00186

Zamarreño-Ramos, C., Camuñas-Mesa, L. A., Pérez-Carrasco, J. A., Masquelier, T., Serrano-Gotarredona, T., and Linares-Barranco, B. (2011). On spike-timing-dependent-plasticity, memristive devices, and building a self-learning visual cortex. *Front. Neurosci.* 5:26. doi: 10.3389/fnins.2011.00026

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 September 2014; accepted: 24 November 2014; published online: 16 December 2014.

Citation: Bill J and Legenstein R (2014) A compound memristive synapse model for statistical learning through STDP in spiking neural networks. *Front. Neurosci.* 8:412. doi: 10.3389/fnins.2014.00412

This article was submitted to *Neuromorphic Engineering*, a section of the journal *Frontiers in Neuroscience*.

Copyright © 2014 Bill and Legenstein. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.