Analog Integrated Circuits and Signal Processing
https://doi.org/10.1007/s10470-018-1276-4

A sub-nW neuromorphic receptors for wide-range temporal patterns of post-synaptic responses in 65 nm CMOS

Xuefei You[1] · Amir Zjajo[1] · Sumeet S. Kumar[1] · Rene van Leuken[1]

Received: 28 February 2018 / Revised: 19 July 2018 / Accepted: 25 July 2018

Abstract
Synaptic dynamics is of great importance in realizing biophysically accurate neural behaviors and efficient synaptic learning in neuromorphic integrated circuits. In this paper, we propose a current-based synapse structure with multi-compartment receptors AMPA, NMDA and GABAa and a weight-dependent learning algorithm. The designed circuit offers distinctive dynamic features of receptors as well as a joint synaptic function. A cross-correlation methodology is applied to a two-layer RNN built by multi-compartment receptors to demonstrate the proposed synapse structure. An increased computation efficiency is verified through temporal synchrony detection among the neural layers in a noisy environment. The design implemented in TSMC 65 nm CMOS technology consumes 1.92, 3.36, 1.11 and 35.22 pJ per spike event of energy for AMPA, NMDA, GABAa and the advanced learning circuit, respectively.

Keywords Neuromorphic design · Synapse · Receptor · Synchrony detection · Synaptic plasticity

1 Introduction

Neuromorphic electronic system is an innovative approach in brain-like system development [1], which utilizes the continuum physics of transistors to emulate biological elements, e.g. the exponential behavior of transistors in sub-threshold region is analogous to the relation between ionic conductance and corresponding membrane voltage [2]. Phenomenological models, a compromise between high fidelity and computational feasibility, allow deeper understanding of working principles underlying neuron networks. Synapse is the connecting structure between neurons with the presynaptic part located in axon terminals and the postsynaptic receptors on dendrites. It experiences synaptic state modifications based on incoming information and current activities in neural networks. Meanwhile, neurons convert resulting signals into spikes and pass them to next layer of synapses. When stimulated by incoming presynaptic spikes, synapses release vesicles, where neurotransmitters are stored, as a way of signal transmission. The binding of released transmitters and receptors at the dendrites of postsynaptic neurons activates receptor channels and induces electrical activities in postsynaptic neurons.

It should be noted that the synaptic state modification induced by a mechanism, commonly called *synaptic plasticity*, is an abstraction of synaptic learning. From a neurological perspective, synaptic plasticity is influenced by the quantity of transmitters, the synaptic weight carriers emitted from axon terminals, and the efficiency of the postsynaptic receptors [3]. Various types of receptors exist: NMDA, AMPA, GABAa, and GABAb, each exhibiting different temporal dynamics in response to neurotransmitters [4]. For example, the speed or response of NMDA receptor is considerably slower in comparison with AMPA. A more rigorous condition should be satisfied to open the receptor channel. Additionally, the unbinding of glutamate and receptors for NMDA is relatively slow [5]. A generic synapse structure does not capture diverse temporal dynamics of different types of receptors in biological synapses, which are essential for the realization of biophysically accurate neural behavior in spiking neural networks (SNN) [6]. In recent neural network studies [7], synapse circuit implementation varies from simple constant current sources activated by presynaptic spikes to more

✉ Xuefei You
x.you@student.tudelft.nl

[1] Circuits and Systems Group, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands

Published online: 11 August 2018

complex realizations of synaptic current dynamics. However, the receptor diversities have usually been ignored. Most research includes single type of synapse in whole neural system [8, 9] or they are partially mentioned in some cases but not described comprehensively [7, 10]. A conductance-based synapse configuration was proposed in [11] which discusses various types of receptors. Although robustness is highlighted, the switched-capacitor based architecture occupies considerable silicon area which limits the integration density. Ion-based model is employed in [12] to simulate receptor-supported synapse structure. Detailed ion dynamics are captured at the cost of high energy consumptions caused by extra active blocks.

In this paper, we propose a current-based phenomenological synapse model [13], consisting of efficient weight-dependent synaptic learning algorithms and multi-compartment synapses, namely AMPA, NMDA and GABAa receptors. The designed circuits [7, 9] offer distinctive receptor dynamics with a compact and power-efficient structure. We introduced a cross-correlation methodology to demonstrate the computation efficiency of synapse in small scale of recurrent neural networks (RNNs) which can also be expanded to other synapse models. The paper is organized as follows: In Sect. 1, the key biological features of receptors and classic synaptic models are described, and proposed multi-compartment circuit introduced. Section 2 discuss the cross-correlation methodology and test settings, and presents the simulation results. Finally, Sect. 3 provides a summery and the main conclusions.

2 Synaptic receptors

The binding transmitter and receptor pairs results in either positive current flow to postsynaptic neuron also termed as excitatory postsynaptic current (EPSC) or negative current flow to postsynaptic current also termed as inhibitory postsynaptic current (IPSC). Different types of receptors display different temporal dynamics due to their distinctive conducting mechanisms. However, we have limited the circuit implementation and simulation in this paper to two main glutamate receptors, AMPA and NMDA, and one GABAergic receptor, GABAa. The GABAb receptor is not covered due to its complex conducting mechanism.

2.1 Biological receptors

2.1.1 AMPA receptor

The AMPA receptor is one of the most common receptors in the nervous system. The AMPA receptor is mostly permeable to sodium (Na^+) via ion channels. The positively charged Na^+ enters the AMPA ion channels upon binding of transmitters on AMPA receptors, which depolarizes the cell thus inducing action potentials. AMPA receptor has a high conduction speed due to a straightforward mechanism of channel opening and closing, and are thus responsible for fast signal transmission [14].

2.1.2 NMDA receptor

The ion channel in NMDA receptor is voltage-dependent, which is distinctive compared with other glutamatergic receptors. This dependency initially arises from the non-selectivity of its ion channels. When ligand-binding occurs, the non-selective ion channels are open to extracellular magnesium (Mg^{2+}) and zinc (Zn^{2+}), which will bind to specific sites on the receptor and block the channels for any other ions. A certain level of depolarization of the cell is necessary to eliminate this blockage through the influx of Ca^{2+} [15]. Once cleared, the ion channels introduce both Ca^{2+} and Na^+ into the target cell. At the same time, in response to the increased level of depolarization, more AMPA receptors are inserted into the membrane, creating higher possibility that ion influx occurs. Thus, the conductance of NMDA receptor has a boost effect on the postsynaptic current. To activate NMDA receptors, the presynaptic activities introduce free transmitters to the dendrites, while the postsynaptic depolarization opens the receptor ion channels. This kind of dual function of presynapes and postsynapses implies the role of NMDA receptor in synchrony detection and biological emulation. On the temporal aspect, the NMDA receptors are typically three to six times slower than AMPA [4], which originates from a more complicated binding mechanism and small unchanneling speed.

2.1.3 GABAa receptor

The GABAa receptor is a primary inhibitory channel carrier in the nervous system. The GABAa receptor is permeable to chloride (Cl^-). When activated, the GABAa receptor conducts Cl^- through the ion channels, causing the hyperpolarization of the cell and a lower possibility of neural firing. This inhibition function of the GABAa receptor is reported to be a prerequisite for balancing excitation and inhibition, thus stabilizing neural network [16]. The GABAa receptors have similar temporal dynamics as AMPA, i.e. both the rise and the fall time of EPSCs are comparable. A brief summary of the dynamic features of three receptors are listed in Table 1.

Table 1 Biological dynamics for three receptors [4, 16]

	Rise and fall times (ms)	Conduction remarks
AMPA (+)	0.4–0.8, 5	1-step, fast EPSC
NMDA (+)	20, 100	2-step, voltage dependency, slow EPSC
GABAa (−)	3.9, 20	1-step, fast IPSC

2.2 Model extraction

2.2.1 Learning rules

Spike timing dependent plasticity (STDP), a temporally symmetrical form of Hebbian's theory [17], is a learning process that can adapt the synaptic weight according to temporal correlations between prespikes and postspikes of a target synapse. These correlations should be within milliseconds time range in accordance with biological temporal features. If the prespike precedes the postspike, a potentiation of the synaptic weight occurs. In contrast, a depression is induced if a reversed sequence happens. A repetition of one of the two patterns described above evokes long-term memory effects which could be long-term potentiation (LTP) and long-term depression (LTD). A diagram of the STDP learning function is illustrated in Fig. 1. Two factors of concern in this learning window are time constants (τ) and amplitudes (A). The time constant indicates the temporal range where the correlation happens while the amplitude controls the adaptation level. The STDP rule is expressed as:

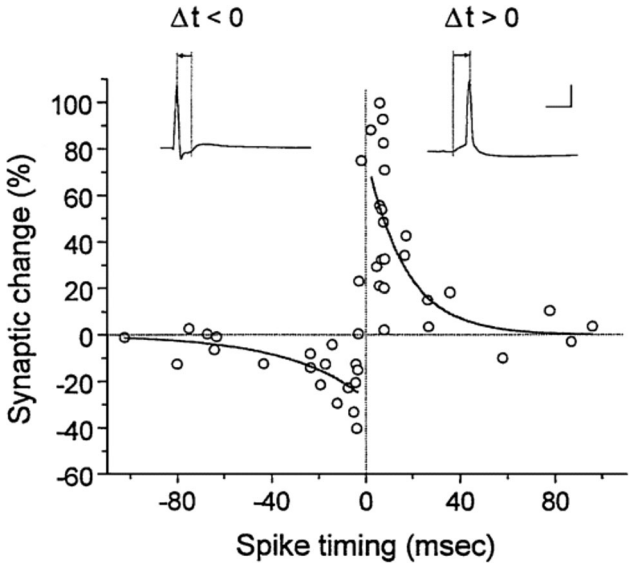


Fig. 1 The learning window of STDP learning rule. The hollow circles are experimental data of EPSC amplitude percentage change at 20–30 min after repetitive stimuli of pre and post spikes at a frequency of 1 Hz [18]. The spike timing is defined as the temporal interval between post and pre spikes. An exponential fit of those data points is outlined with two smooth curve (LTP and LTD). For LTP and LTD respectively, A = 0.777 and 0.273; $\tau = 16.8$ and 33.7 ms

$$\Delta w^{+} = A^{+} \cdot e^{-\Delta t/\tau_{+}} \quad \Delta t > 0 \tag{1}$$

$$\Delta w^{-} = -A^{-} \cdot e^{\Delta t/\tau_{-}} \quad \Delta t < 0 \tag{2}$$

where Δt is the temporal difference between a single pair of postspike and prespike. A^{+} and A^{-} are the maximum amplitude while τ_{+} and τ_{-} are time constants of the potentiation and the depression phase, respectively. These parameters impact the area of the weight update curves during potentiation and depression. It is observed that stable learning is realized when the aggregate area of depression exceeds that of potentiation in the weight update function. However, weaker depression results in the extreme potentiation of synaptic weights and the eventual shorting of outputs to inputs. This behavior prevents the realization of any practical network transfer function.

2.2.2 Synaptic dynamics

One of the basic and direct models describing the synaptic conductance properties is the exponential decay model where the rising phase of the synaptic conductance is assumed to be instant [19]. This means the release of transmitters, its corresponding diffusion across the cleft, the receptor binding, and channel opening all happen very fast. The conductance of the synapse at time t is then:

$$g_{syn}(t) = \bar{g}_{syn} \cdot e^{-(t-t_0)/\tau} \cdot \Theta(\mathrm{t} - t_0) \tag{3}$$

where $\bar{g}_{syn}$ is the maximum conductance of the synapse, t_0 is the onset time of the presynaptic spike while τ is the decaying time constant, and $\Theta(x)$ is the Heaviside step function.

The exponential profile is a match with the relationship between the ionic conductance of a neuron and its membrane potential [2]. For some IPSCs, the exponential decay model is validated to outline their activities because the rising phase of these currents is much shorter compared with the decaying phase, like the GABAa-induced currents (see Table 1). The exponential decay model is also satisfactory for AMPA receptors which are fast excitation contributors. The model fails to simulate EPSCs induced by NMDA receptors which have comparable temporal dynamics in both rising and decaying phases. A more detailed model with two separate exponential components is introduced:

$$g_{syn}(t) = \bar{g}_{syn} \cdot f \cdot (e^{-(t-t_0)/\tau_{decay}} - e^{-(t-t_0)/\tau_{rise}} \cdot \Theta(t - t_0) \tag{4}$$

The factor f is used to normalize the total amplitude of the sum to $\bar{g}_{syn}$. τ_{decay} and τ_{rise} are the time constant for decaying and rising phase, respectively. The neuromorphic design for this model is more complex due to one extra rising phase.

A transistor operating in the sub-threshold region not only provides low power consumption, but also its exponential relationship between the drain currents and the gate voltages is analogous to the biological model. The drain current of a transistor is:

$$I_{ds} = I_0 \cdot e^{\kappa_n V_g / U_T} \cdot (e^{-V_s/U_T} - e^{-V_d/U_T}) \tag{5}$$

where I_0 is the current scaling factor and is the value of current when V_{gs} equals the threshold voltage of the transistor. κ_n represents the sub-threshold slope for n-type MOSFET. V_g, V_d, V_s are the gate, drain, source voltages relative to the bulk and U_T is the threshold voltage of the transistor. The exponential relationship between V_{gs} and I_{ds} is as follows when the transistors enters the saturation region during sub- threshold conduction and satisfies the condition $V_{ds} \geq 4U_T \approx 100\,\text{mV}$:

$$I_{ds} = I_0 \cdot e^{\kappa_n V_{gs}/U_T} \tag{6}$$

2.3 Circuit implementation

The top-level of multiple receptor synapse architecture is shown in Fig. 2. The synaptic weight, generated by STDP learning block and transmitted through receptors, results in a wide range of temporal dynamics of EPSCs or IPSCs. Those overlapping responses are then integrated in integrate and fire (I&F) neuron [16] generating further firing spikes. The voltage dependent NMDA receptor receives membrane voltages as a feedback from the I&F neuron. The three receptors integrated with an advanced STDP learning circuit [9] form a multi-compartment cluster structure. The transistor level implementation is shown in Fig. 3.

A differential-pair integrator [7] structure is applied to simulate the fast rising and decaying dynamics of AMPA receptor. The receptor generates biologically analogous synaptic currents, which are modeled as a time-dependent alpha function with finite duration in rising phase [20]. The extra scaling factor V_{thr} and leakage rate adjustment V_{tau1} offer flexibility to implement amplitude and time constant control of AMPA EPSCs. Additionally, linear filtering properties makes it possible to sum multiple currents from identical synapses, yielding significant area savings.

Unlike the single exponential dynamics used for AMPA receptor, the charging phase of NMDA receptor cannot be ignored due to its relatively large portion in the whole temporal range. Consequently, a double exponential function should be displayed in NMDA receptor design as well as its distinctive weight dependence. The circuit for the falling phase of NMDA receptor is derived from AMPA receptor while the circuit for the rising phase is extended via M26–M28. The presynaptic spike enables an instantaneous current influx into C_{nmda1} in the rising phase, the amplitude of which is mediated by V_w. The bias V_{tau2} determines the discharge speed of C_{nmda1}. Capacitor C_{nmda2} is charged when transistor M30 is conducting. After that, capacitor C_{nmda2} begins to discharge through M32 biased by V_{tau3}, adjusting the falling time constant. In this way, controllable double exponential dynamics are generated. To incorporate the distinctive voltage dependence of NMDA receptors, a differential pair is added to the circuit, forming a comparison between the membrane voltage V_{mem} and the threshold V_{mth}. When the postsynaptic neuron is depolarized, V_{mem} surpasses V_{mth}, introducing valid current flux into C_{nmda2}.

The GABAa receptor has analogous dynamics to AMPA receptor except for the polarity. However, the inhibitory synapses do not exhibit learning properties. Since the inhibitory level is independent of the synaptic weight, it is not necessary to have two control voltages over the inhibitory level i.e., V_{thr} and V_w. A log-domain integrator [7] is chosen for the implementation of GABAa receptor due to its simplicity and linear dynamics.

The advanced STDP learning circuit [9] incorporates presynaptic and postsynaptic spike trains, and conducts

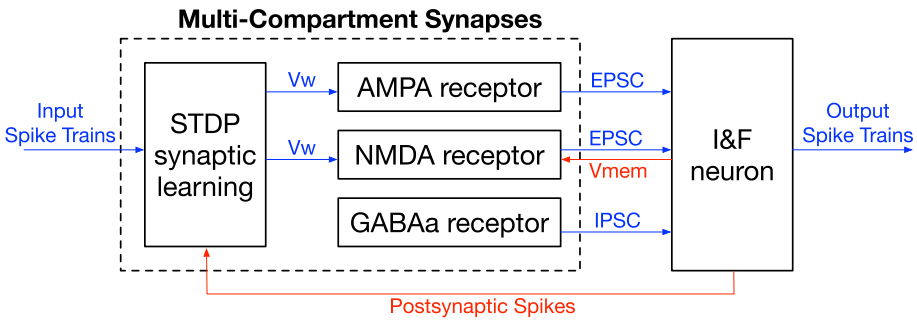


Fig. 2 Top-level diagram of multi-receptor mediated synapse architecture. The blue arrowed lines represent signal transmissions while red ones represent feedback signals. The dashed block denotes the multi-compartment synapse. Created by OmniGraffle

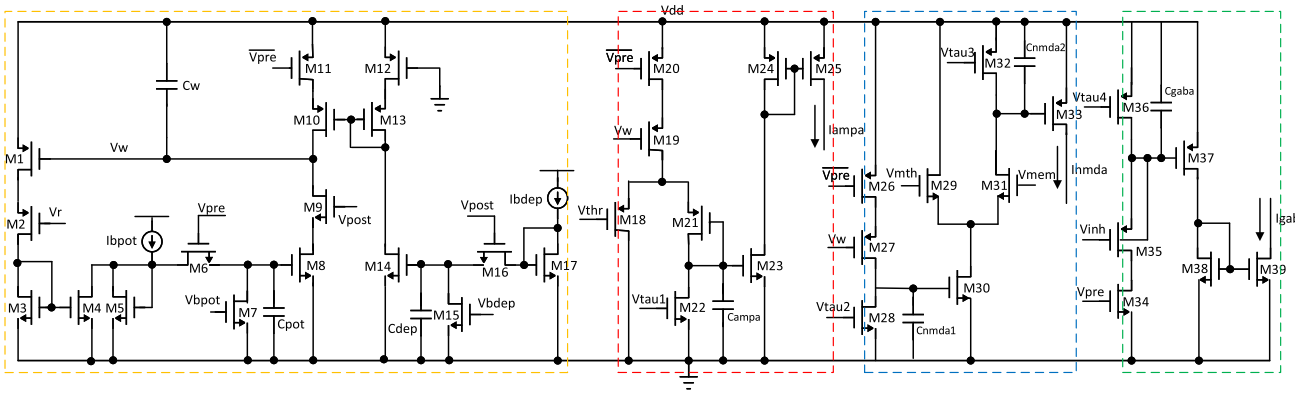


Fig. 3 The cluster structure including an advanced STDP learning circuit with three receptors. Different color dashed blocks denote different functional unit: yellow—advanced STDP learning, red—AMPA receptor, blue—NMDA receptor, green—GABAa receptor. The synaptic weight value is transmitted through node V_w between component circuits. Created by Microsoft Visio (Color figure online)

weight adaptation according to the STDP learning rule. The potentiation and depression phases exhibit almost symmetrical structures, inducing charge influx and efflux from the weight capacitor C_w depending on the relative timing of prespike and postspike pairs. Note that the circuit in this paper follows a complementary design, i.e. larger V_w represents smaller weight value. The time constants and amplitudes of the learning window are tuned through V_{bpot}, V_{bdep} and I_{bpot}, I_{bdep}, respectively. A weight dependence feature is added to the potentiation domain via M1–M3 to match with experimental observations obtained in [21]. V_r is connected to a controllable voltage, inducing adjustable weight dependence level. When synaptic weight increases which corresponds to a decrease in V_w, a larger current is subtracted from I_{bpot} through M5, resulting in lower current influx to C_{pot}. The above circuits have low power consumptions.

The transistors in above circuits are set in sub-threshold region except the switch transistors M6, M9, M11, M16, M20, M26 and M34. The initial value of V_w is set to VDD/2. The current fluxes from C_w are adjusted that the weight control transistors M19 and M27 always operate in sub-threshold region. The circuit conducts only in the presence of the presynaptic or postsynaptic spikes, which lasts for no more than 2 ms.

3 Simulation results

3.1 Synaptic receptors

3.1.1 Single receptor

The EPSC amplitude is determined by the width of prespike signal and the synaptic weight value. In our experiments, pulse width of spike trains is set to 100 μs. The time constants of receptors are regulated by transistor control voltages V_{tau1}–V_{tau4} to cover wide temporal range. The time constant ranges from several to tens of milliseconds for AMPA receptors. Similarly, the time constants for NMDA receptor in both rising and falling phases are adjustable via two separate transistors M28 and M32.

The weight dependence of NMDA receptor is demonstrated through a comparison of V_{mem} and a reference voltage V_{mth}. A sequence of presynaptic spikes are introduced to synapse. The V_{mem} is a step signal from 0 to 500 mV at the onset time of 40 ms larger than V_{mth}. The membrane voltages are in spike forms in reality. We specifically examine the threshold function of NMDA receptor in this setting. It can be observed in Fig. 4(a) that a growing current output appears from the onset of V_{mem}. A linear increase of EPSC amplitudes can be found at each stimuli. When stimuli are densely distributed, single NMDA EPSC fails to return to resting line before the next stimuli comes due to large decaying time constants, which results in summation of previous activities.

In Fig. 4(b), AMPA currents are introduced at the onset time of 5 ms, and NMDA currents are injected at different times. Two cases need to be discussed individually. In the first case where AMPA stimuli precedes NMDA, the excitatory function of NMDA receptors can be demonstrated by the increase of V_{mem} ($\Delta t = 2, 5, 8$ ms). As NMDA stimuli approaches AMPA stimuli, larger V_{mem} is detected by NMDA synapse, which gives a greater voltage amplification. However, if delivered in reversed sequence ($\Delta t = -1$ ms), no modification is observed. This can principally be explained by the cooperation mechanism of those two receptors; AMPA receptors usually act as preliminary depolarization with post neurons by inducing small amount of ions (Na^+) into cells. When depolarization threshold is surpassed, NMDA receptors are activated, which allows substantial incursion of Na^+ and Ca^{2+} ions.

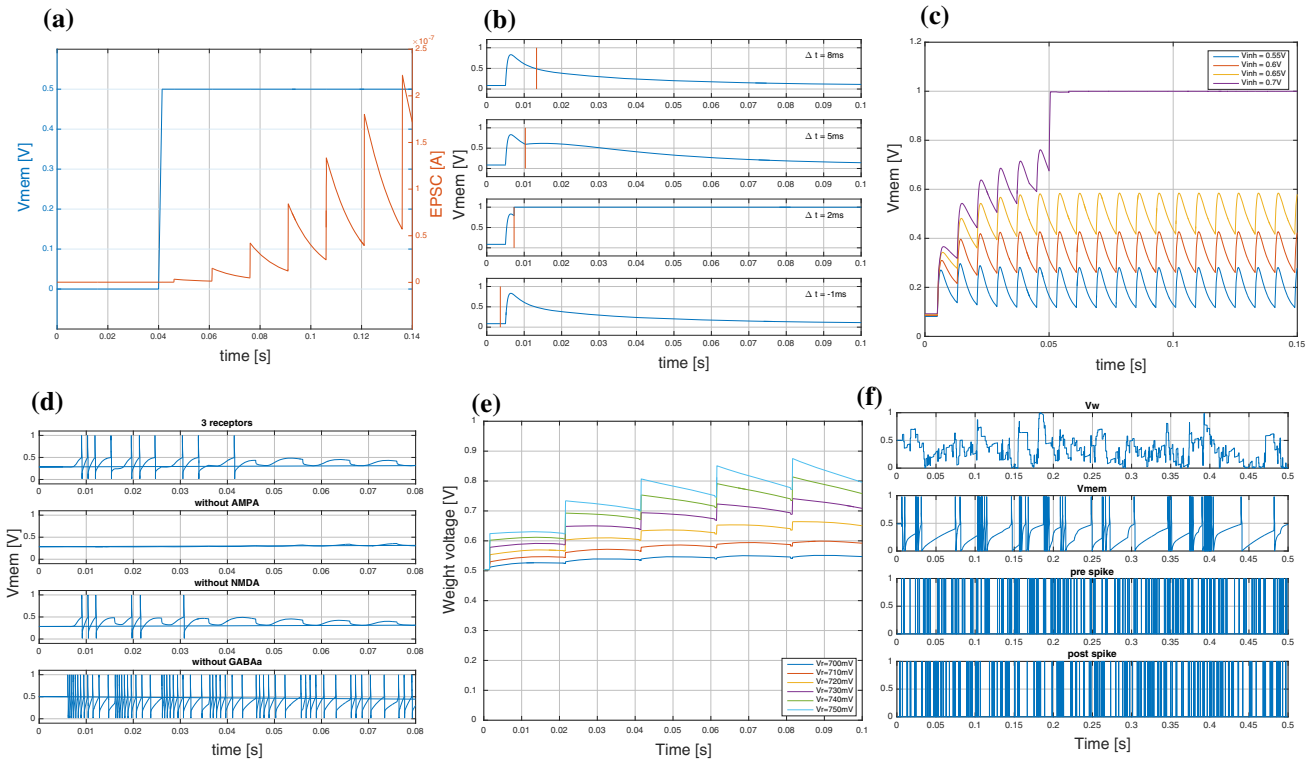


Fig. 4 **a** Weight dependence of NMDA receptors. **b** NMDA receptor mediated potentiation function on membrane voltage. AMPA currents are induced at an onset time of 5 ms while the time for NMDA receptor varies (labeled with red vertical lines). Δt represents the interval between NMDA and AMPA activations, ranging from -1 to 8 ms. **c** The role of GABAa receptor in synaptic integration. Due to a complementary design, larger inhibitory weight voltage means a smaller function of GABAa receptor. **d** The membrane voltage dynamics with different receptor configurations. From top to bottom, the strips represent configurations with three receptors, without AMPA, without NMDA and without GABAa, respectively. **e** Weight dependence of the advanced STDP circuit. In this example, the presynaptic and postsynaptic signals are of 50 Hz frequency with 1 ms delay. V_r determines the weight dependence level. **f** Sample synaptic weight evolvement and the membrane voltage distribution in the advanced STDP circuit. From top to bottom, the strips represent synaptic weight, membrane voltage, presynaptic spikes and postsynaptic spikes, respectively. Created by Matlab (Color figure online)

Bigger electrical stimuli are produced. Thus, it is implied that NMDA receptors are not self-initiated. However, the NMDA receptors act as major contributions to electrical signal transmission once they are activated in neuron system.

In Fig. 4(c), the contribution of inhibitory synapses to synapse integration is identified. Various levels of inhibition are applied to the system while the setting of excitatory synapses are maintained. When inhibition behavior is larger than certain level ($V_{inh} \leq 0.65$ V), neuron system operates normally. Conversely, if the inhibition level decreases, excitation prevails, which drives membrane state to the upper boundary and consequently loss of information may occur during this process. This result suggests that inhibitory synapses are of great importance in balancing membrane activities, especially in the case of NMDA receptors where long-term summation of multiple receptors may exist.

3.1.2 Joint function

Input spikes are introduced to synaptic learning circuit at a rate of 100 Hz with prespikes preceding postspikes for 1 ms which induce consecutive depression to synaptic weight. In the presence of three receptors, ten membrane spikes are generated as shown in Fig. 4(d). A gradually sparser distribution of the spikes is observed along with the decline of synaptic weight. When synaptic weight reaches lower bound, the network fails to produce any spike trains. In absence of AMPA receptor, no postspike trains are observed. This result is analogous to biological experiments observed in hippocampal region [5]. Synapse with only NMDA receptors, also called *silent synapse*, will only transmit information when the postsynaptic neuron is depolarized due to synchrony pairing of other synapses with AMPA receptors. When NMDA receptor is inhibited, the temporal intervals to generate equal number of spikes are larger, while the amount of spike clusters is lower due to a lack of long-term dynamics. If GABAa receptor is inhibited, a burst of postspikes is produced even though the

spike dynamics should decline with a decreasing synaptic weight. Consequently, the network fails to transmit learning information carried by synapses. Hence, GABAa receptor is essential to create stable signal transmission in SNNs. The average energy consumptions in a typical corner are 1.92, 3.36, 1.11 and 35.22 pJ per spike event for AMPA, NMDA, GABAa and the advanced learning circuit, respectively.

3.2 Advanced learning circuit

Figure 4(e) demonstrates the functionality of the weight dependence block in the advanced STDP circuit. The input signal pairs induce a stable increment of the synaptic weight value. The increment in V_r becomes less effective designating a unimodal weight distribution [6] as the weight adjustment level decreases. An example weight evolvement and the corresponding membrane voltage distribution with Poisson distributed presynaptic and postsynaptic input signals of 200 Hz are displayed in Fig. 4(f).

3.3 Synchrony detection

The method used in synchrony detection experiments is called *cross-correlograms*. It is a visualization of cross-correlation between two spike trains, i.e. the similarity of two series as a function of the temporal displacement of one relative to the other. In cross-correlograms, the temporal differences between every single pair of spikes are summed for certain temporal bin. A peak present in the cross-correlogram indicates a correlated relation at this certain temporal bin between target spike groups. For discrete signals, the cross-correlation is defined as [22]:

$$(f \bigstar g)[n] \underline{\underline{\text{def}}} \sum_{m=-\infty}^{\infty} f^*[m] g[m+n] \tag{7}$$

where f^* denotes the complex conjugate of f and n represents the displacement bins, which correspond to the temporal difference between two target spikes.

The magnitude of the correlation in the cross-correlogram indicates the causality level between sequential neural units. If two spike trains are strongly correlated, a large correlation amplitude displays in the cross-correlogram. It is an efficient way to characterize the computation efficiency of artificial neural network, and especially the SNN where the weight update information is encoded into the temporal difference between spikes. When valid learning updates happen, the spikes generated are more closely distributed regardless of the transmission delay, which leads to a higher cross-correlation between two spike trains.

A two-layer RNN consisting of the cluster structures is utilized to explore the parallel and hierarchical synchrony detection and amplification function of the multi-compartment synapse layers as described in [13] shown in Fig. 5. The system detects the spike-timing synchrony between spike trains embedded in a noisy environment and amplify this correlation from layers.

The same simulation environment (network configuration, input and noise patterns) is applied to three receptor settings: multi-receptor, AMPA receptor and NMDA receptor as shown in the dashed blocks in Fig. 5. Each setting consists of three neural clusters, which includes one classic I&F neuron and four functional synapses with either multiple or single receptor implementations. A correlated Poisson distributed spike train C1 is introduced to the first two synapses out of four in cluster N1, N4 and N7 while another correlation C2 is added to the first two synapses in cluster N2, N5 and N8, which adds an additional correlation between N1 and N2, N4 and N5, N7 and N8. The rest of the synapses obtain Poisson distributed spike trains as noise. The Poisson distributed spike trains are generated from Matlab. The correlated spike trains are more likely to coincide in the defined learning window of STDP, which will increase valid weight update events. The cross-correlogram is used to demonstrate the temporal synchrony between two output spike patterns in whether the same or different layers of neurons.

Noise has a great impact on the response dynamics of neural system. The synaptic noise is the main contributor among many other sources of noise existing in neurons [23]. The chemical synapses release packets of neurotransmitter at the axon terminals depending on the probabilistic firing history of both the presynaptic and postsynaptic neurons. However, the learning induces long-term effect on the postsynaptic neurons, which transmit the resultant spike trains to every spatial location they can reach and further form a recurrent network. This non-unidirectional transmission may pass on those spikes to a cell as noise. The summation of the thousands of synaptic inputs of one neuron forms irregular fluctuations on the neural response, ranging from completely random Poisson inputs to periodic inputs [23]. The probability of k events in an interval is given by:

$$P(k \text{ events in interval}) = \frac{e^{-\lambda}\lambda^k}{k!} \tag{8}$$

where λ is the average number of events per interval, k ranges from 0 to the total event amount. The Poisson distribution with five different λ values are illustrated in Fig. 6.

The learning width is chosen as 100 ms [21], which corresponds to a frequency of 10 Hz. The average input signal frequency should range from 10 to 1000 Hz (the

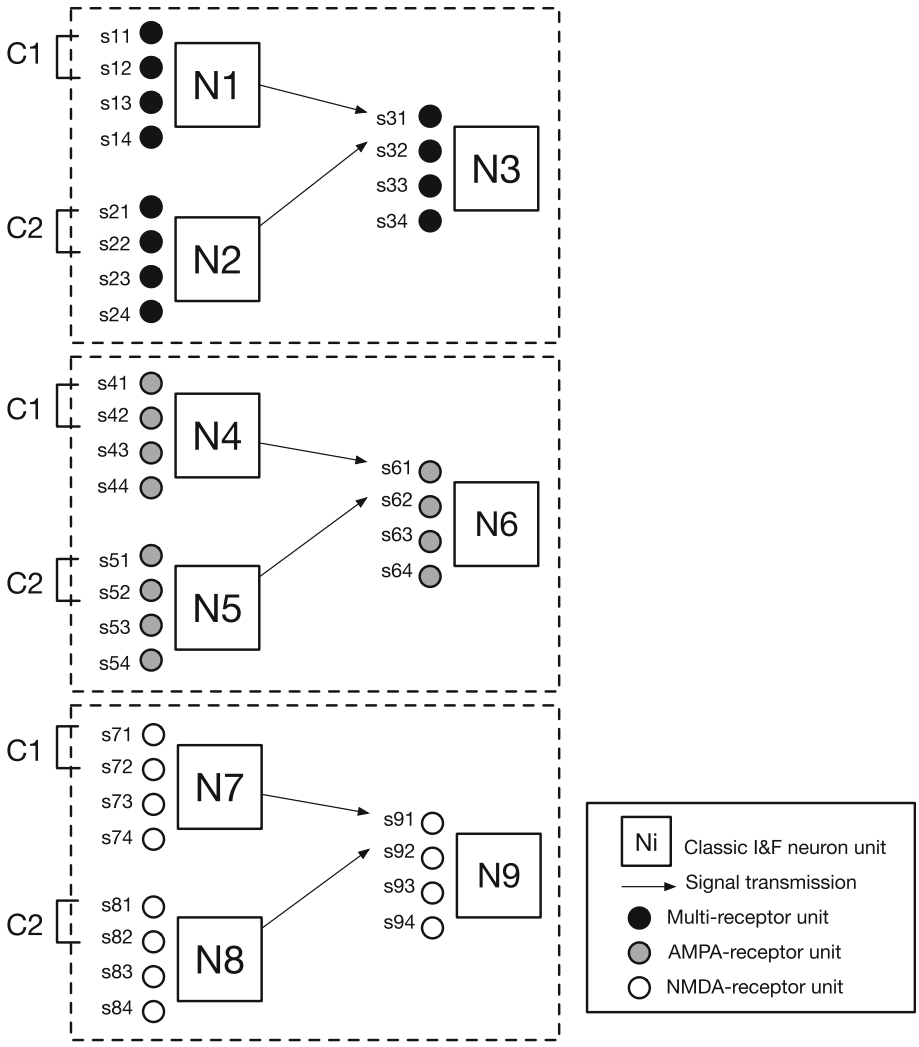


Fig. 5 Top-level diagram of the two-layer recurrent testing network. The symbol interpretations are listed in the box on the right. Each neuron *Ni* is connected with four synapses *si*1–*si*4, forming a cluster unit. Clusters belonging to different dashed block includes different types of synaptic receptor configurations. Input C1 is a Poisson distributed spike train of 40 Hz. Input C2 is correlated with C1, and this correlation can be in any correlation form. Here a delay of 2 ms is used. The rest of the synapses receive Poisson distributed spike trains of 15 Hz. Created by OmniGraffle

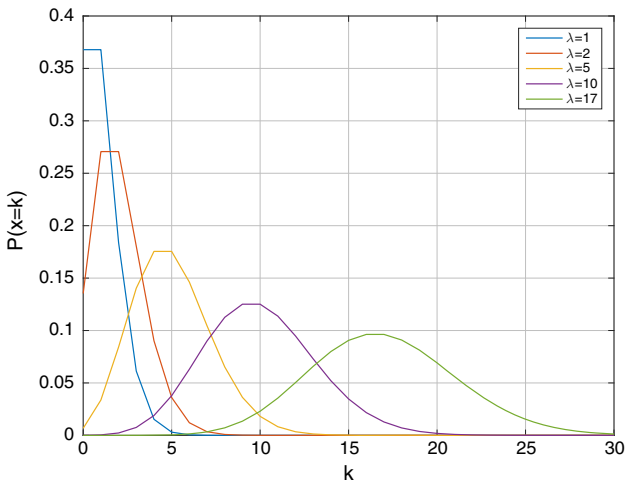


Fig. 6 The Poisson distribution with five different λ values. Created by Matlab

refractory period is 1 ms) in order to induce valid learning process. We applied 15 and 40 Hz Poisson distributed signals to demonstrate the functionality of the system.

The correlated input spikes induced by C1 and C2 correlation can be in several forms. It can be perfectly simultaneous or one precedes the other or one leads the other. The spike trains within the same cluster are perfectly simultaneous for our case while a delay is added to C2 to form correlation with C1 between the clusters. An example of cross-correlation is shown in the form of graphs with different delay values in Fig. 7. Larger delay time decreases the possibility of coincidence between two spike trains within target time range, thus reducing the peak correlation amplitude. In addition, a temporal shift is generated, and is proportional to the delay introduced to C2. As we want to compare the synchrony detection and amplification of different receptor configurations, the exact magnitudes of the correlation levels are of no interest. Therefore, any time delay is possible for simulation as long

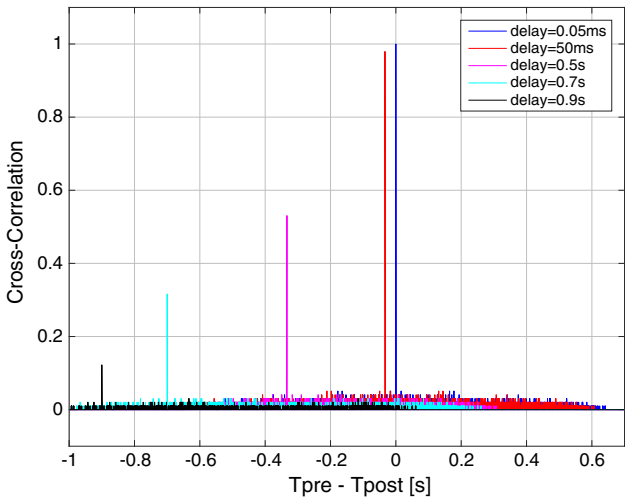


Fig. 7 The effect of delay of spike trains on cross-correlation. Created by Matlab

as the magnitude of correlation peak is above the noise horizon. In our experiments, the decay time is set as 2 ms.

Figure 8 shows the normalized cross-correlogram results from the two-layer recurrent testing network described above. It compares the synchrony level of three different receptor configurations: multi-receptor, AMPA receptor and NMDA receptor. In every plot, the temporal range of interest is 0.1 s, which covers the range of the learning window reported in [18]. The shift of correlation spikes as observed in Fig. 7 implies the temporal difference between two spike trains. Similarly, the amplitude of the correlation peak denotes the correlation level, i.e. the total amount of the coincidence occurring at this target lag point.

The histograms in Fig. 8(a, d) evaluate the cross-correlations between parallel and hierarchical clusters. In both cases, a large level of correlation is observed at a point close to zero time. This indicates a strong synchrony between parallel and hierarchical spike trains after synaptic learning process with multi-receptor settings. The synchrony level is decreased almost by half in Fig. 8(b, d), the. Similarly, the background noise as well as sub-peaks occurring near the origin is observed in the hierarchical relations, which implies a relatively poor stability performance. A peak shift occurs along with the amplitude decay. The delay between the inputs to the clusters is passed through layers while that of the multi-receptor synapses is mitigated. Finally, Fig. 8(c, f) characterize the synchrony detection function of NMDA-receptor network. Both parallel and hierarchical pairs have similar correlation plots as multiple receptor network but with reduced amplitudes

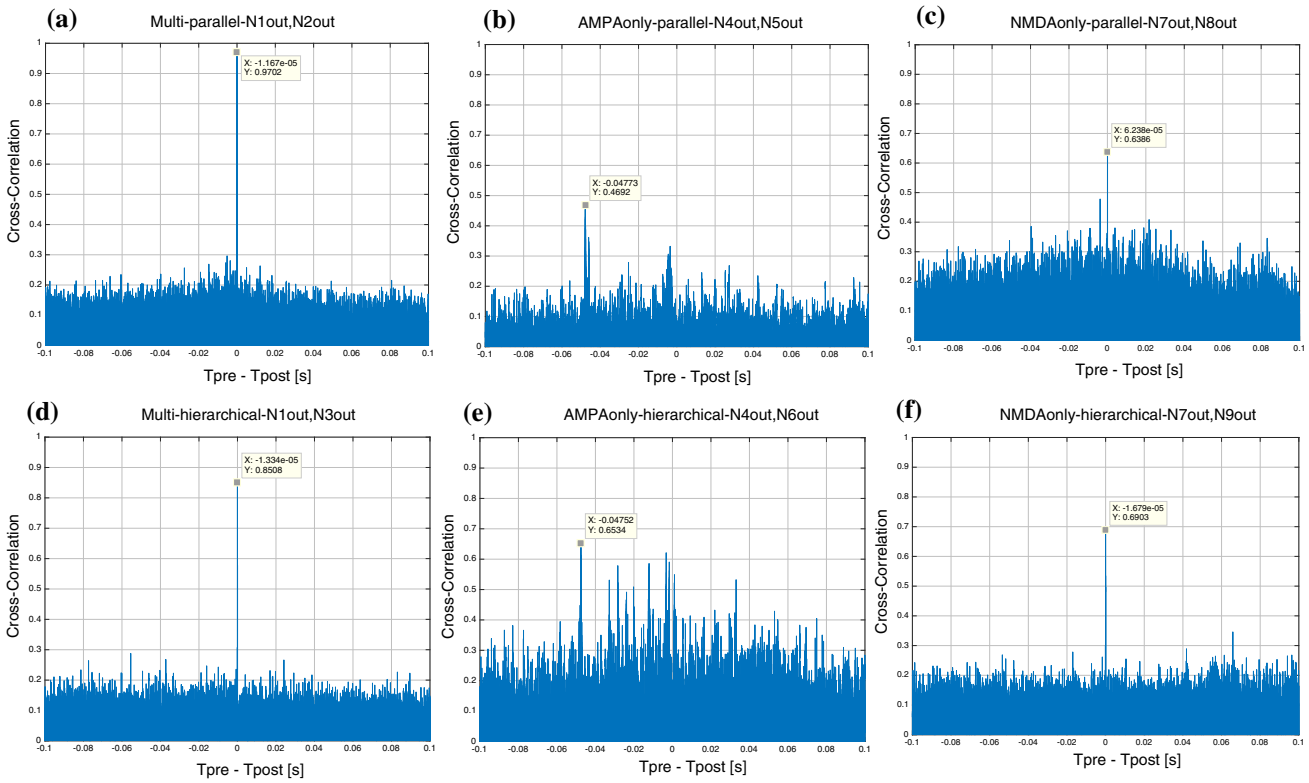


Fig. 8 Normalized cross-correlogram results from the two-layer recurrent testing network. **a–f** are the parallel and hierarchical cross-correlation plots for multi-receptor, AMPA receptor and NMDA receptor configurations, respectively. The annotation above each figure tells "receptor configuration and correlation source type". For example, "Multi-parallel" means the parallel correlation of multi-receptor settings. Created by Matlab

Table 2 Normalized cross-correlation comparison

	AMPA receptor		NMDA receptor		Multiple receptors	
	Temporal shift (s)	Amplitude	Temporal shift (s)	Amplitude	Temporal shift (s)	Amplitude
Parallel	− 0.04773	0.4692	− 6.238e−5	0.6386	− 1.167e−5	0.9702
Hierarchical	− 0.04752	0.6534	− 1.679e−5	0.6903	− 1.333e−5	0.8508

Table 3 Performance comparison between current works

	[11]	[12]	[24]	This work
Model type	Conductance-based	Current-based	Current-based	Current-based
Receptors included	AMPA, NMDA, GABA	AMPA, NMDA, GABA	AMPA, NMDA, GABA	AMPA, NMDA, GABAa
Technology	180 nm	1.5 μm	0.35 μm	65 nm
Supply voltage	1.8 V	5 V	–	1 V
Power consumption	45 μW for 1 neuron and 1 receptor	10–100 nW per receptor	–	111–336 pW per receptor
Remarks	Switched-capacitor structure, large area	Ion-based, biologically realistic	Phenomenological, dendritic dynamics	Phenomenological, achieving high fidelity

which are approximately 60 to 70% of multiple receptor configuration.

The detection and amplification level of the cross-correlation function of various spike train pairs using both single and multiple receptors is shown in Table 2. The maximum amplification level of multiple receptor configuration is about 2 times as compared to single receptor ones.

The circuit shows an efficient learning ability as the consecutive neural clusters generate almost synchronized output spike patterns in the presence of delay in the inputs signals. It takes shorter time for system with multiple-receptor to achieve synchrony. Analysis indicate that this ability originates from the NMDA receptor as NMDA receptor displays similar correlations except for a decrement in the amplitude of correlation level. AMPA and NMDA receptors have a collaborate relation in inducing efficient synchrony detection and amplification for synapse structures. A performance comparison of the existing research work with previous references is summarized below in Table 3.

4 Conclusion

In this paper, we have proposed a synaptic circuit with biologically-accurate temporal dynamics. The receptors within the synapse structure operate in a collective way: the AMPA receptor works as the preliminary trigger to signal transmission, the NMDA receptor enhances the learning efficiency, while the GABAa receptor balances the membrane voltage. The collaboration of receptors offers high efficiency in synchrony detection and amplification. The synaptic design, built in TSMC 65 nm CMOS technology, consumes extremely low energy in picojoule level per spike event.

References

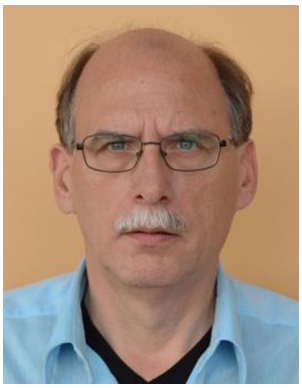
1. Mead, C. (1990). Neuromorphic electronic systems. *Proceedings of the IEEE, 78*(10), 1629–1636.
2. Azghadi, M. R., Iannella, N., Al-Sarawi, S. F., Indiveri, G., & Abbott, D. (2014). Spike-based synaptic plasticity in silicon: Design, implementation, application, and challenges. *Proceedings of the IEEE, 102*(5), 717–737.
3. Gaiarsa, J. L., Caillard, O., & Ben-Ari, Y. (2002). Long-term plasticity at GABAergic and glycinergic synapses: Mechanisms and functional significance. *Trends in Neurosciences, 25*(11), 564–570.
4. Destexhe, A., Mainen, Z. F., & Sejnowski, T. J. (1998). Kinetic models of synaptic transmission. *Methods in Neuronal Modeling, 2,* 1–25.
5. Liao, D., Hessler, N. A., & Malinow, R. (1995). Activation of postsynaptically silent synapses during pairing-induced LTP in CA1 region of hippocampal slice. *Nature, 375*(6530), 400.
6. Morrison, A., Diesmann, M., & Gerstner, W. (2008). Phenomenological models of synaptic plasticity based on spike timing. *Biological Cybernetics, 98*(6), 459–478.
7. Bartolozzi, C., & Indiveri, G. (2007). Synaptic dynamics in analog VLSI. *Neural Computation, 19*(10), 2581–2603.
8. Indiveri, G., Chicca, E., & Douglas, R. (2006). A VLSI array of low-power spiking neurons and bistable synapses with spike-timing dependent plasticity. *IEEE Transactions on Neural Networks, 17*(1), 211–221.

9. Bofill-i-Petit, A., & Murray, A. F. (2004). Synchrony detection and amplification by silicon neurons with STDP synapses. *IEEE Transactions on Neural Networks, 15*(5), 1296–1304.
10. Qiao, N., Mostafa, H., Corradi, F., Osswald, M., Stefanini, F., Sumislawska, D., et al. (2015). A reconfigurable on-line learning spiking neuromorphic processor comprising 256 neurons and 128K synapses. *Frontiers in Neuroscience, 9,* 141.
11. Noack, M., Krause, M., Mayr, C., Partzsch, J., & Schuffny, R. (2014). VLSI implementation of a conductance-based multi-synapse using switched-capacitor circuits. In *2014 IEEE international symposium on circuits and systems (ISCAS)* (pp. 850–853). IEEE.
12. Rachmuth, G., & Poon, C. S. (2008). Transistor analogs of emergent iono-neuronal dynamics. *HFSP Journal, 2*(3), 156–166.
13. You, X., Zjajo, A., Kumar, S.S., & van Leuken, R. (2017). Energy-efficient neuromorphic receptors for wide-range temporal patterns of post-synaptic responses. In *Nordic circuits and systems conference (NORCAS)* (pp. 1–6).
14. Platt, S. R. (2007). The role of glutamate in central nervous system health and disease—A review. *The Veterinary Journal, 173*(2), 278–286.
15. VanDongen, A. M. (Ed.). (2008). *Biology of the NMDA receptor*. Boca Raton: CRC Press.
16. Wu, S. H., Ma, C. L., & Kelly, J. B. (2004). Contribution of AMPA, NMDA, and GABAA receptors to temporal pattern of postsynaptic responses in the inferior colliculus of the rat. *Journal of Neuroscience, 24*(19), 4625–4634.
17. Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.
18. Bi, G. Q., & Poo, M. M. (2001). Synaptic modification by correlated activity: Hebb's postulate revisited. *Annual Review of Neuroscience, 24*(1), 139–166.
19. Schutter, E. D. (2009). *Computational modeling methods for neuroscientists*. Cambridge: The MIT Press.
20. van Vreeswijk, C., Abbott, L. F., & Bard Ermentrout, G. (1994). When inhibition not excitation synchronizes neural ring. *Journal of Computational Neuroscience, 1,* 313–321.
21. Bi, G. Q., & Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of Neuroscience, 18*(24), 10464–10472.
22. Amit, D. J., & Fusi, S. (1994). Learning in neural networks with material synapses. *Neural Computation, 6*(5), 957–982.
23. Brunel, N., Chance, F. S., Fourcaud, N., & Abbott, L. F. (2001). Effects of synaptic noise and filtering on the frequency response of spiking neurons. *Physical Review Letters, 86*(10), 2186.
24. Wang, Y., & Liu, S. C. (2011). A two-dimensional configurable active silicon dendritic neuron array. *IEEE Transactions on Circuits and Systems I: Regular Papers, 58*(9), 2159–2171.

Xuefei You received the M.S. degree from Delft University of Technology, Netherlands, and the B.E. degree from University of Electronic Science and Technology of China, China, both in electrical engineering, in 2017 and 2015, respectively. Her master work focused on analog neuromorphic synapse, investigating the role of multiple synaptic receptors for wide-range temporal patterns in recurrent neural network.

Amir Zjajo received the M.Sc. and DIC degrees from the Imperial College London, London, U.K., in 2000 and the PhD. degree from Eindhoven University of Technology, Eindhoven, The Netherlands in 2010, all in electrical engineering. In 2000, he joined Philips Research Laboratories as a member of the research staff in the Mixed-Signal Circuits and Systems Group. From 2006 until 2009, he was with Corporate Research of NXP Semiconductors as a Senior Research Scientist. He joined Delft University of Technology in 2009 as a Faculty Member within Circuits and Systems group. Dr. Zjajo has published more than 80 papers in referenced journals and conference proceedings, and holds more than 10 US patents or patent pending. He is the author of the books Brain-Machine Interface: Circuits and Systems (Springer, 2016), Low-Voltage High-Resolution A/D Converters: Design, Test and Calibration (Springer, 2011, Chinese translation, China Machine Press, 2015) and Stochastic Process Variations in Deep-Submicron CMOS: Circuits and Algorithms (Springer, 2014). He served as a member of Technical Program Committee of IEEE International Symposium on Quality Electronic Design, IEEE Design, Automation and Test in Europe Conference, IEEE International Symposium on Circuits and Systems, IEEE International Symposium on VLSI, IEEE International Symposium on Nanoelectronic and Information Systems, and IEEE International Conference on Embedded Computer Systems. His research interests include energy-efficient digital/mixed-signal circuit and system design for biomedical and mobile applications, on-chip machine learning and inference, sensor fusion, and bionic electronic circuits for autonomous cognitive systems.

Sumeet S. Kumar received the Bachelor of Engineering degree in Electronics and Communications from the CMR Institute of Technology, India, in 2008, the Master of Science and PhD degrees in Microelectronics from the Delft University of Technology, The Netherlands, in 2010 and 2015 respectively. In 2015, he joined Intel's Imaging and Camera Technologies Group in Eindhoven, The Netherlands to design domain-specific tools for the development of complex media processor architectures. Since 2016, Dr. Kumar serves as a project manager at the Delft University of Technology, where he is responsible for leading industrial consortia in collaborative R&D projects aiming to develop compute hardware for highly automated vehicles. Dr. Kumar is a member of the IEEE, and serves on the TPC of several leading IEEE/ACM conferences and journals. His research primarily focuses on the development of efficient compute architectures for cognitive applications.



Rene van Leuken received an MSc. and a Ph.D. degree in electrical engineering from the Delft University of Technology in 1983 and 1988 respectively. At the moment he is a professor at the Circuit and Systems group at Faculty of Electrical Engineering, Mathematics and Computer Science of the Delft University of Technology (TU Delft), The Netherlands. His current research interests include high level digital system design, system design optimization, VLSI design, and high performance compute (DSP) engines. Currently his major research activity is neuromorphic computing. He has been involved over the years in the inception, creation, and execution of many major research and development projects: JESSI, MEDEA, ENIAC/CATRENE, ARTEMIS and recently in ECSEL and PENTA projects. He is a member of the PATMOS steering committee and the DATE Technical Program Committee. He has published papers in all major journals, conferences and workshops proceedings and has received several best paper awards over the years.