Hardware Spiking Artificial Neurons, Their Response Function, and Noises

Doo Seok Jeong

Abstract In this chapter, overviewed are hardware-based spiking artificial neurons that code neuronal information by means of action potential, *viz*. spike, in hardware artificial neural networks (ANNs). Ongoing attempts to realize neuronal behaviours on Si 'to a limited extent' are addressed in comparison with biological neurons. Note that 'to a limited extent' in this context implicitly means 'sufficiently' for realizing key features of neurons as information processors. This ambiguous definition is perhaps open to a question as to what neuronal behaviours the key features encompass. The key features are delimited within the framework of neuromorphic engineering, and thus, they approximately are (i) integrate-and-fire; (ii) neuronal response function, i.e. spike-firing rate change upon synaptic current; and (iii) noise in neuronal response function. Hardware-based spiking artificial neurons are aimed to achieve these goals that are ambitious albeit challenging. Overviewing a number of attempts having made up to now illustrates approximately two seemingly different approaches to the goal: a mainstream approach with conventional active circuit elements, e.g. complementary metal-oxide-semiconductor (CMOS), and an emerging one with monostable resistive switching devices, i.e. threshold switches. This chapter will cover these approaches with particular emphasis on the latter. For instance, available types of threshold switches, which are classified upon underlying physics will be dealt with in detail.

1 Introduction

Neuromorphic systems based on full hardware artificial neural networks consist of a complex array of building blocks encompassing artificial neuron and synapse devices. The artificial neuron represents the information—input synaptic current—which it receives by relaying the corresponding response to neighbouring neurons via

D.S. Jeong (✉)
Center for Electronic Materials Research, Korea Institute of Science and Technology,
Hwarangno 14-gil 5, Seongbuk-gu, Seoul 136-791, Republic of Korea
e-mail: dsjeong@kist.re.kr

synapses. Namely, the neuron encodes the input information into a particular form of response and the response is subsequently processed by the synapses and further encoded by the bridged neurons. Encoded information at the outset, for instance, by sensory neurons flows throughout the entire network by repeating the aforementioned process. The neuron is therefore required to represent information clearly in order to prevent information loss in the network. Hereafter, a device in charge of information representation in a neuromorphic system is referred to as an artificial neuron—in short, neuron—for simplicity, at times its behaviour is barely biologically plausible though.

Choice of type of an artificial neuron depends upon the dimension of a neuromorphic system—including time frame, i.e. dynamic system, leads to an additional request for dynamic neuronal behaviour—and the information type in use such as binary, multinary, and analogue. The simplest case is when the neuromorphic system disregards time frame, i.e. static, and employs information in binary; time-independent binary neurons that merely represent 1 bit of information meet the requirements. For instance, a summing amplifier—summing inputs from adjacent neurons—in conjunction with a single transistor, representing binary states—channel on and off ($\log_2 n$ and $n = 2$, where n denotes the number of states)—upon the input in total, perhaps successfully works as an artificial neuron in this simplest system. Employing multinary or analogue information in the neuromorphic system requires the neuron to exhibit various states ($n > 2$) for $\log_2 n$ bit of information; for instance, a transistor, working in the subthreshold regime, and thus representing multiple states of channel conductance, may meet the need. Note that the use of multinary information gains the remarkable benefit that the same amount of information is represented by much less number of multinary neurons than binary neurons if preventing the response variability upon the same input and the consequent information loss [1, 2]. Later, this variability-induced information loss will be addressed in detail from the perspective of information decoding.

When it comes to time-dependent (dynamic) neuromorphic systems, the neuron should introduce a time-varying response to the input in total, which also varies in time, rendering the neuron complicated. If binary, the input elicits all—or—nothing output. The rest case, dynamic and multinary information-utilizing neuromorphic systems, is most biologically plausible regarding information representation by biological neurons. As for the dynamic binary system, the response of the neuron should be reliably distinguished in time. Besides, the response should vary upon the input in order for the response to be decodable. For instance, in case of a single-pulse output, the response may be parameterized by pulse height and/or width—varying with respect to the input—and used in information representation. A number of different types of neurons and their outputs can be used in such a system as far as they meet the aforementioned requirements and are compatible with the synapse as a whole in the system. Regarding the root of neuromorphic engineering, our intuition perhaps leads us to use the most biologically plausible one—satisfying fidelity to key features of biological neurons—among the possible candidates. Thus far, a great deal of effort on building biologically plausible artificial neurons—hereafter such a type of a neuron is referred to as *spiking* neuron—

has been made by capturing key features of biological neurons and then realizing them by means of electric circuit components [3].

The forthcoming Sects. 1.1–1.3 are dedicated to addressing essential characteristics of biological neurons, which artificial neurons are required to reflect. Sect. 1.4 is dedicated to introducing in silico—computational—neuron models that also form the basis of hardware artificial neurons. Section 2 addresses the state-of-the-art spiking artificial neurons being classified as silicon neuron (SiN), realized by very-large-scale integration (VLSI) circuits, and functional materials-based emerging technologies.

1.1 Biological Neurons

Biological neurons (also known as nerve cells) are a *living* electrochemical system being made of a lipid membrane containing ion channels and pumps. A schematic of a biological neuron is illustrated in Fig. 1. The membrane separates intracellular and extracellular media in which there exist significant differences in concentration of important ions, e.g. Na^+, K^+, Ca^{2+}, and Cl^-, between the media. Na^+, Cl^-, and Ca^{2+} are rich in the extracellular medium whereas K^+ and fixed anions—in the sense that they cannot diffuse out through ion channels—such as organic acids and proteins, are rich on the other side. Namely, for each ion, chemical potential through the membrane is not equal. In the resting state, the extracellular charge in total is positive, whereas the total charge in the intracellular medium is negative as a whole; that is, the resting state indicates electrical polarization, leading to electric field, e.g. electrostatic potential gradient, evolution through the lipid membrane. This potential difference caused by the difference in chemical potential between the two sides of the membrane is referred to as the Donnan potential [4], which tends to recover the equilibrium distribution—no chemical potential gradient—of each ion. This ionic distribution is not spontaneous configuration as being far from energy-minimizing conditions; work should be done by a third party to maintain the polarized state, raising the internal energy. Ion pumps embedded in the membrane

Fig. 1 Schematic of neurons and chemical synapses in-between. The *grey* and the *orange* neurons denote a presynaptic and a postsynaptic neurons, respectively. The *blue arrows* indicate the spike propagating direction—from the pre- to the postsynaptic neuron
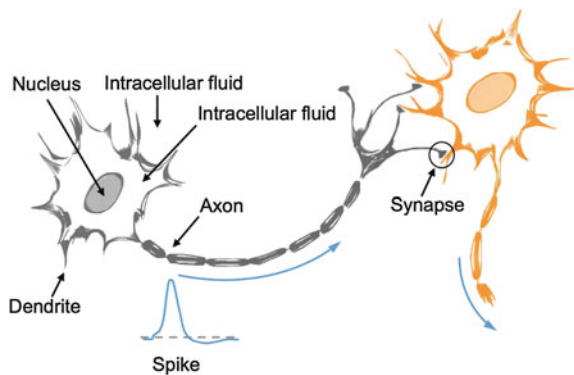
are in charge of the work by pumping out ions in order for the resting state to be maintained [5]. Ion pumps consume chemical energy that is subsequently converted into electrical energy as for batteries [6]. They are known as sodium–potassium adenosine triphosphatase (Na^+/K^-−ATPase), which drive Na^+ and K^+ ions—two important ions in membrane potential maintenance—into and out of the intracellular fluid, respectively.

Temporary breakdown of the resting state occurs upon significant fluctuation in ion concentration in the intracellular fluid. In particular, Upon release of chemical messengers, i.e. neurotransmitters, at a chemical synapse, N-methyl-D-aspartate receptor (NMDAR) and α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic receptor (AMPAR) open their ion channels, so that both monovalent ions, e.g. Na^+ and K^+, and divalent ions, e.g. Ca^{2+}, are able to diffuse into the intracellular fluid, leading to breakdown of the resting state ion concentration [7]. The inward diffusion corresponds to *synaptic current* that is of significant importance in spiking. Such concentration fluctuation causes a change in membrane potential regarding the aforementioned origin of the potential. Given that Na^+ and K^+ ion channels are voltage-gated [5], the change in the membrane potential can determine the channel conductance. However, the consequent change in the membrane potential should be larger than a *threshold* for the channel opening; otherwise, the change lasts for a short time without leading to any significant phenomena. Thus, introduction of synaptic current—sufficing for large potential evolution—opens voltage-gated Na^+ and K^+ channels, and thus, the resting state ion distribution is locally destroyed by inward Na^+ ions and outward K^+ ions through their characteristic ion channels. As a consequence, the membrane undergoes depolarization—down to almost zero volt or sometimes polarity reversal occurs.

Once such local depolarization occurs, ion pumps do not let the membrane maintain the depolarization and recovers the resting state ion distribution, and thus membrane potential. Meanwhile, this local depolarization in turn affects adjacent Na^+ and K^+ ion channels and the same depolarization procedure keeps being repeated throughout the entire axon down to the end of the axon that is termed as axon terminal—similar to dominoes [6]. Therefore, so far explained procedure in both time and spatial frames produces an action potential also known as spike and it propagates along the axon in due course. Neurons in a neural network communicate by firing spikes and receiving them. To be precise, a train of spikes, rather than a single spike, appears used in the communication, which enables representation of analogue information in terms of spike frequency also known as activity.

1.2   Neuronal Response Function

The biological neuron locally fires a spike if and only if the membrane potential goes above the threshold, which is attributed to synaptic current injection through channels such as NMDAR and AMPAR. For the time being, the question as to how these purely electrochemical phenomena are in association with information repre-

sentation arises. To answer this question, we need to elucidate the electrochemical phenomena in regard to a relationship between input and output and define types of both analogue input and output. Given that synaptic current injection initiates spiking at the outset, synaptic current works as input, varying its quantity upon synaptic weight. Synapse in this chapter implies chemical synapse unless otherwise stated. Synapses are spatially discrete given that each neuron has an immense number of synapses. Also, the synaptic current through them varies in time upon arrival of a train(s) of spikes from the presynaptic neuron. Fortunately, the membrane of the neuron resembles a capacitor that integrates the synaptic current varying in metric as well as temporal space. The integrated synaptic current at given time is therefore typically taken as input value. Activity—the number of spikes per unit time or at times per particular time period—caused by the input is typically regarded as output; that is, a train of spikes, rather than a single spike, is of concern in determining output value. Note that evaluating time-varying activity is fairly complicated, in particular, when the synaptic current varies in time—this situation is very often encountered in in vivo experiments. There exist several methods in the linear-filtering framework with different linear filters (see Ref. [8]).

The activity tends to increase with synaptic current and this tendency is the substrate of the neuronal response function that enables representation of analogue information. Given the role of the membrane as a capacitor, the larger the synaptic current, the sooner the membrane potential hits the threshold for spiking and the more there exist spikes in a unit time. This tendency holds for stereotypical neurons, the detailed behaviour is rather complicated in the light of presence of refractory time though. A relation between activity and synaptic current is referred to as a gain function, providing the key features of the tuning function [9]. From a perspective of information coding, the gain function provides single neuron's ability to encode a large number of different inputs—in comparison with digital components such as a transistor representing merely 1 bit of information—and make them distinguishable by the output, i.e. very clearly decodable in one-to-one correspondence. However, biological neurons are noisy [10–12], in particular, in vivo neurons, so that a great deal of information is lost. Neuronal noises and related information loss will be addressed in the following subsection in detail.

1.3 Neuronal Noises

Neuronal behaviour, particularly in vivo behaviour, is noisy including a number of unpredictable features. Neuronal noises are loosely classified as (i) background noise, e.g. white noise, in the membrane potential in conjunction with the neuronal response of focus [8] and (ii) irregularity in spiking, e.g. irregular periodicity of spikes in a single spike train [1, 2, 13, 14]. The former may be attributed to white-noise-like synaptic current injection that keeps on perturbing the membrane potential. [13, 15] These two types of noises may be correlated insomuch as the random synaptic current injection—causing the former—enables random spike firing at times [15]. Addition-

ally, the background noise is thought to enable gain modulation, altering the gain function [16]. For the time being, questions as to what roles of the noise in a neural network are and if it is good or bad have barely been answered clearly. Regarding information representation, of course, the noise is the root cause of information loss, and thus, it needs to be avoided [17]. The brain may, however, perform stochastic calculations making use of probability, rather than single bit, coding schemes, which can be robust to errors due to noises in the noisy neural network [18–20]. However, this hypothesis is open to objection due to lack of understanding of probability estimation [14].

Regarding the irregularity in spiking, the irregularity is well parameterized by variability in inter-spike interval (ISI) in an ensemble of spike trains [8]. In the biological neuron, the ISI distribution in the same spike train is dispersive. The distribution varies upon experimental conditions such as in vitro and in vivo experiments. Despite such variations, it is generally agreed that the biological neuron represents Poisson or Poisson-like noises, which is possibly justified by the relation between the mean neuronal response and the corresponding variance [8, 13]. The Poisson noise is a result of random spike generation, so that it is also required to be identified, for instance, by simply evaluating correlation between the spikes in the same train [21]. A similar type of noise may be present in hardware artificial neurons due to operational variability in the core device. We will revisit this later in Sect. 2.

1.4 Artificial Neuron Models

Thus far, biological neurons are featured by (i) integrate-and-fire—the membrane integrates synaptic current and fires a spike when the resulting membrane potential goes above the threshold; (ii) gain function—the output varies upon the input and they are in one-to-one correspondence in the ideal case; and (iii) noise leading to variability in ISI, and thus activity. There exist several artificial neuron models, being mostly for computational purpose but serving as the basis of hardware-based artificial neurons, which mainly reflect the first two features. The integrate-and-fire (IF) model is simplest; its equivalent circuit is shown in Fig. 2a. The capacitor in Fig. 2a corresponds to the membrane of a biological neuron, and it integrates the injected synaptic current until the electrostatic potential difference across the capacitor reaches the threshold. Crossing the threshold generates a spike by the lumped switch in parallel with the capacitor, and the refractory period follows spiking hindering another spiking in close succession. The lumped switch and its functions in spiking can simply be programmed, rendering the model very easily implementable. However, deficient fidelity of the model to the biological neuron is noticeable at a glance; particularly, the membrane cannot be lossless perfect dielectric as assumed in the model in the light of leakage of ions through the membrane. That is, underpinning the model and thus making it more realistic require a parallel resistor to the capacitor, which results in charge loss in due course when the model system is subject to no or smaller synaptic current than it used to be. This charge loss and

Fig. 2 Equivalent circuits of (**a**) IF and (**b**) LIF neurons. For each circuit, the component lumped in the box outlined with a *grey dashed line* generates a spike when the membrane potential crosses V_{th}, i.e. a threshold membrane potential for spiking
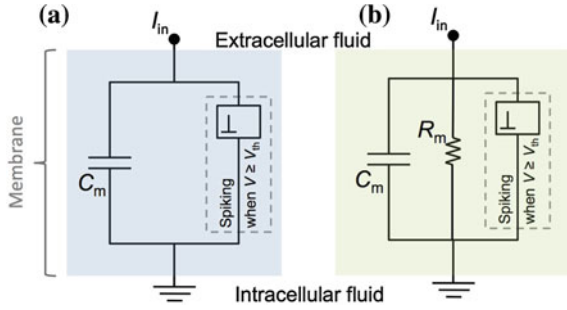
the corresponding potential decay away with time cannot be implemented in the IF model. This modified model is referred to as the leaky integrate-and-fire (LIF) model [8, 17, 22–25]. The equivalent circuit is illustrated in Fig. 2b. The membrane potential V_m with time in the subthreshold regime is described by the following equation: $V_m = I_{in} R_m \left(1 - e^{-t/\tau} \right)$, where I_{in}, R_m, and τ mean the input current, i.e. synaptic current, the membrane resistance, and the time constant for membrane charging and discharging—product of R_m and C_m (membrane capacitance)—respectively. Including the parallel resistor barely adds up further complexity so that the LIF model is still easy to implement in in silico neural networks. In fact, the LIF model is thought to reflect the most essential feature in a very simple manner, offering remarkable computational efficiency regarding computational time. In addition, this model likely bridges a computational neuron model to a hardware-achievable model in which realizing the lumped switch by means of scalable electrical components is of significant concern. This issue will be revisited in Sect. 2.

It should be noted that the resting state of both IF and LIF models indicates depolarization and, upon incidence of synaptic current, a build-up of membrane potential, i.e. polarization, evolves in due course unlike the biological neuron—membrane polarization in the resting state and depolarization evolving upon synaptic current incidence. It is, however, believed that this reversal is trivial. The most significant discrepancy between the simple LIF model and the biological neuron consists in the fact that the LIF deals with a single type of charge, whereas the biological neuron involves two different ions, i.e. Na^+ and K^+, and their transport through different ion channels in the membrane. In this regard, the LIF model is often regarded as being oversimplified to gain computational efficiency.

Another class of artificial neuron models is conductance-based LIF model in which the conductance of the ion channel varies upon arrival of a spike from the postsynaptic neuron [26–29]. This class is more biologically plausible in the sense that the assumed channel conductance increase upon spike transmission is the feature of the biological neuron albeit lumped. Furthermore, unlike the integrate-and-fire-based models, the resting state denotes polarization and the polarization is perturbed by the increase in the channel conductance, leading to depolarization. Various conductance-based models are available, ranging from a simple one offering similar computational efficiency to the LIF model [29] to a one, offering a great

deal of fidelity to the biological neuron, such as the Hodgkin–Huxley (HH) model [26]. Later, several attempts to simplify the HH model have been made by reducing the number of auxiliary variables; the FitzHugh–Nagumo model is an example [27, 28, 30].

2 Hardware Spiking Neurons

Several prototypical hardware spiking neurons have been achieved in search of models fulfilling the aforementioned requirements for neuronal information representation. Basically, they should produce spike trains in association with input synaptic current and the activity of the neuronal response is required to vary upon the input synaptic current. When it comes to hardware neurons, integration, for which scalability is a premise, should be taken into account, and thus scalability is an additional requirement.

2.1 Silicon Neurons

The neuromorphic engineering stemmed from VLSI technologies, dating back in 1980s [31]. In the past few decades, vigorous attempts to realize neuromorphic circuits, encompassing basic building blocks, i.e. neuron and synapse, and related physiological phenomena, have been made by utilizing VLSI circuit elements such as field-effect transistor (FET), capacitor, and resistor [3, 32, 33]. As a consequence, they have led the mainstream trend of neuromorphic engineering. An evident advantage of the SiN is such that designed circuits can readily be fabricated using the conventional complementary metal-oxide-semiconductor (CMOS) technology. Also, the SiN is likely scalable unless the case of real-time operation, i.e. the operational timescale is comparable to that of biological neurons, which perhaps requires a large capacitor(s) working as the membrane. Note that the same difficulty in scalability due to a large capacitor also holds for the emerging neurons. A number of SiNs mimicking different in silico neuron models and the progress in the related technologies are very well overviewed by Indiveri et al. in their review paper [3].

2.2 Emerging Spiking Neurons

As mentioned in Sect. 1.4, the LIF model includes the lumped switch and one should be in search of the switch—achievable by means of electrical components—so as to realize the LIF model. In 2012, Pickett et al. at the Hewlett Packard Labs proposed a neuron model in the framework of the LIF neuron [34]. Later, Lim et al. named the model the neuristor-based LIF (NLIF) model [21]. In the NLIF circuit, a neuristor—
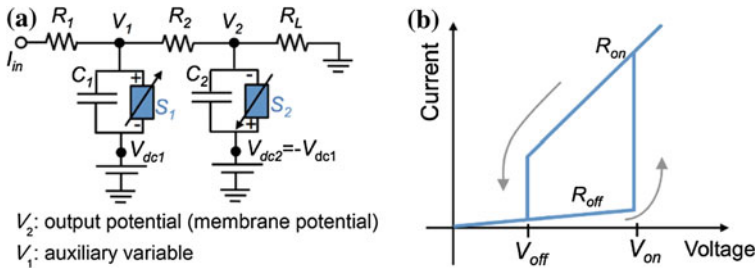
Fig. 3 a Equivalent circuit of the NLIF neuron model. The *blue boxes* stand for the threshold switch whose current–voltage hysteresis loop is sketched in **b**. Reprinted with permission from Ref. [21]. Copyright 2015, Nature Publishing Group

its first proposal dates back to 1962 [35]—is employed as the lumped switching; the neuristor consists of a pair of Pearson–Anson oscillators in parallel [34]. The oscillators are made of a pair of a capacitor and a *threshold switch* in parallel. The circuit of the proposed NLIF neuron circuit is illustrated in Fig. 3a. The key role in the oscillation is played by the S-shape negative differential resistance (NDR) effect that is not followed by memory effect. This class of resistive switching is termed as 'monostable' resistive switching in order to distinguish it from 'bistable' switching typically referring to memory-type resistive switching [36]. Given the lack of memory effect, this class of switching is also referred to as volatile switching. The threshold switching effect is a typical example of volatile switching. The effect has been seen in various systems such as amorphous higher chalcogenides [37–40], Mott insulators [34, 36, 41, 42], Si n+/p/n+ junctions [43], and particular transition metal oxides such as NbO_x [44, 45].

A typical current–voltage (I-V) hysteretic loop of a threshold switch is shown in Fig. 3b. Two thresholds for critical resistance changes are defined in Fig. 3b: V_{on} for high-to-low resistance transition and V_{off} for the reversed one. The low resistance state (LRS; R_{on}) emerges at V_{on}, and a drastic transition from the high resistance state (HRS; R_{off}) to the LRS takes place. The LRS can be maintained unless the applied voltage falls below V_{off} ($V_{off} < V_{on}$), i.e. the HRS is of the only stability under no voltage application.

Regarding the working principle of the NLIF model, the membrane potential evolution and the resulting spiking can be described by membrane potential V_2 and auxiliary variable V_1 as follows [21]:

$$C_1 \frac{dV_1}{dt} = I_{in} - \frac{1}{R_{S1}}(V_1 - V_{dc1}) - \frac{1}{R_2}(V_1 - V_2) \tag{1}$$

and

$$C_2 \frac{dV_2}{dt} = \frac{1}{R_2}(V_1 - V_2) - \frac{1}{R_{S2}}(V_2 - V_{dc2}) - \frac{1}{R_L}V_2, \tag{2}$$

where R_{S1} and R_{S2} denote the resistance of threshold switches S1 and S2, respectively. The spiking dynamics of the NLIF model can be mapped onto a $V_2 - V_1$ phase plane; this phase-plane analysis clearly shows the dynamics in a compact manner. In fact, such a phase-plane analysis is often employed to account for membrane potential change in time for the FitzHugh–Nagumo model [27, 28]. Towards this end, V_1- and V_2-nullclines need to be defined, which denote the lines on which every (V_2, V_1) point satisfies dV_1/dt and dV_2/dt, respectively. The V_1- and V_2-nullclines are readily obtained by equating the left-hand side of Eqs. (1) and (2) to zero as follows:

$$V_1 - \frac{R_{S1}}{R_2 + R_{S1}}V_2 - \frac{R_2 V_{dc1}}{R_2 + R_{S1}} - \frac{R_2 R_{S1}}{R_2 + R_{S1}}I_{in} = 0, \tag{3}$$

and

$$V_1 - R_2\left(\frac{1}{R_2} + \frac{1}{R_{S2}} + \frac{1}{R_L}\right) + \frac{R_2 V_{dc2}}{R_{S2}} = 0, \tag{4}$$

respectively.

A crossing point between these nullclines is termed as fixed point at which both dV_1/dt and dV_2/dt are zero—no gradient in either axis exists so that the (V_2, V_1) configuration becomes stuck at this stable point unless perturbation is applied to the configuration, for instance, by changing I_{in}. It should be noted that, as explained, R_{S1} and R_{S2} vary upon the threshold switching defined by two thresholds, V_{on} and V_{off}, so that the nullclines also vary in association with R_{S1} and R_{S2} as shown in Eqs. (3) and (4). This change in the nullclines features the spiking dynamics of the NLIF model, rendering this model distinguishable from other dynamic models such as the HH and the FHN models [21]. The results of the phase-plane analysis under a constant synaptic current are plotted in Fig. 4, sorted in due course. Following this first cycle of (V_2, V_1) at the outset, the trajectory repeats the grey cycle plotted in Fig. 4f, which is referred to as limit cycle.

The limit cycle shows important information as to the spiking dynamics at a glance: V_2 maximum (spike height) and V_2 minimum (spike undershoot). In addition, such phase-plane analysis intuitively provides a current threshold for spiking in comparison between Figs. 4a, b—the fixed point without synaptic current application (Fig. 4a) is required to be driven out of the subthreshold region in order for spiking to be initiated as for the case shown in Fig. 4b. The fixed point for the nullclines in Fig. 4b is definitely placed in the above-threshold region, so that the trajectory meets the threshold switching condition. Otherwise, a fixed point stays in the subthreshold region and thus no threshold switching takes place throughout the entire current application.

A following question is if the NLIF model successfully represents a gain function. Recent calculations on the NLIF model have properly answered the question by exhibiting an *ideal* gain function shown in Fig. 5a. The activity tends to increase with synaptic current above the threshold for spiking. In particular, for case α, the change in activity with respect to synaptic current is very large compared to the other cases— suitable for discriminating the applied synaptic current without uncertainty [21].
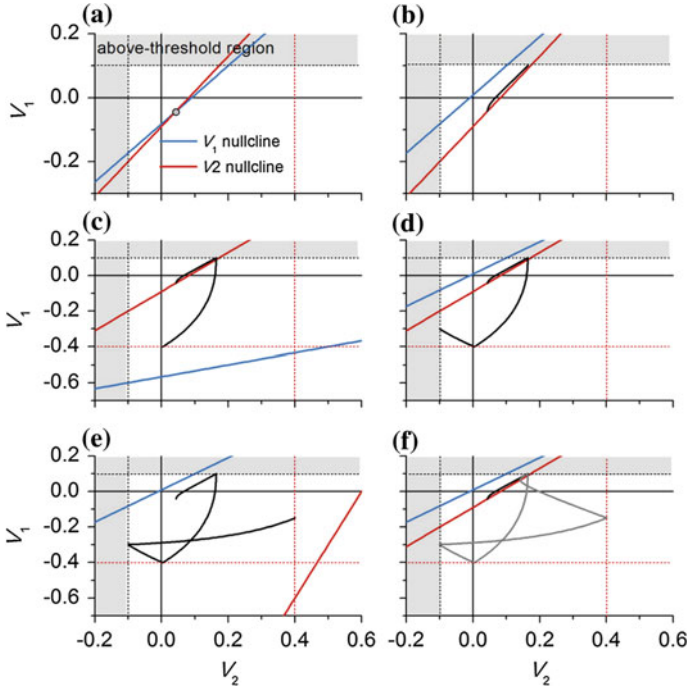
Fig. 4 Phase-plane analysis on the NLIF neuronal dynamics. The V_1- and V_2-nullclines Eqs. (3) and (4) cross each other at a particular point—fixed point—on the plane. The resting state is described by the nullclines and fixed point in **a**. The state of the neuron stays at the fixed point unless perturbation. **b** Upon application of a current, the V_1-nullclile shifts upwards and the fixed point leaves the subthreshold region; the trajectory moves towards the fixed point in due course. **c** The trajectory confronts the threshold switching condition, so that V_1-nullclile shifts down according to Eq. (3) and accordingly a new fixed point appears. **d**–**f** Threshold switching of both switches subsequently takes place and the dynamic procedure finishes a single complete cycle, i.e. limit cycle,—a *grey* cycle in **f**. The cycle in succession follows this limit cycle and the same dynamics is repeated. Reproduced with permission from Ref. [21]. Copyright 2015, Nature Publishing Group

Owing to the gain function, the tuning function of the NLIF model is consequently acquired by presuming the synaptic current in association with a stimulus and the preferred stimulus of the synapse [21]. Invoking the Bienenstock–Cooper–Munro rule, stimulus selectivity is spontaneously given to synapses so that each synapse is supposed to possess strong preference for a unique stimulus [46]. In case of one-dimensional stimulation, such as edge orientation and light intensity in visual stimulation, the synaptic current can be assumed to be of a bell-shaped, i.e. Gaussian, function centred at the preferred stimulus (zero degree; see Fig. 5b), which eventually gives a tuning function (see Fig. 5c). The tuning function is also bell shaped and represents the maximum activity at the preferred stimulus. One of the most significant implications of a tuning function is to bridge external sensory information, i.e. stimulus, to the corresponding neuronal response. In this way, the sensory
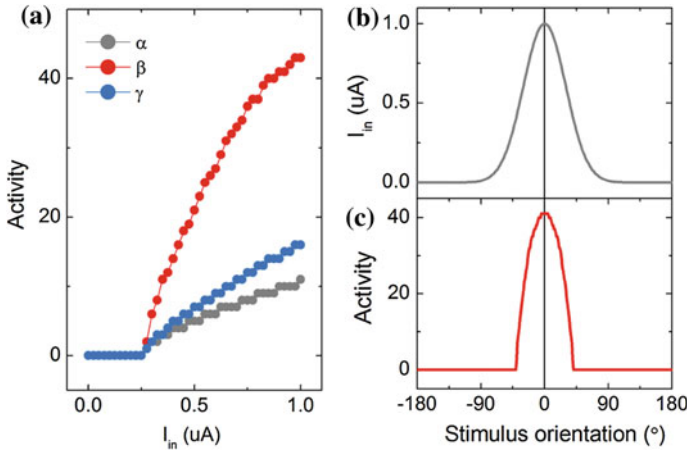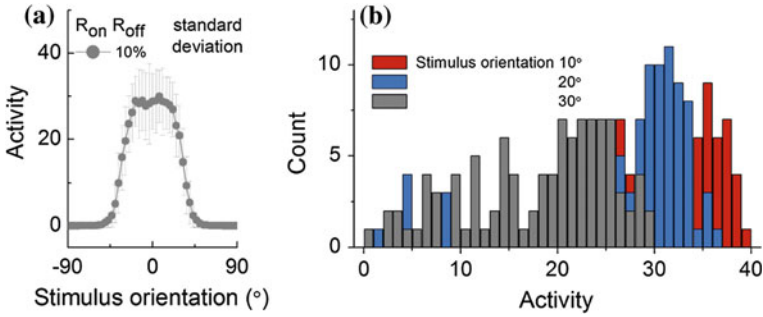
Fig. 5 a Gain functions of the NLIF neuron with three different sets of circuit parameters. **b** Assumed synaptic current profile with respect to one-dimensional stimulus (here, bar orientation). **c** The resulting tuning function of the given NLIF neuron clearly represents good selectivity to input stimulus. Reproduced with permission from Ref. [21]. Copyright 2015, Nature Publishing Group

information is encoded by the neuron into the value for activity and later the activity is most likely decoded in communication between adjacent neurons.

It is well known that resistive switching encompasses variability in switching parameters such as V_{on}, V_{off}, R_{on}, and R_{off}. The same most likely holds for the threshold switches in the NLIF model. The variation of such switching parameters is generally a switching event-driven, rather than real-time-varying, phenomenon, i.e. upon termination of one switching cycle—depicted in Fig. 3b—the parameters are updated, and this random update lasts throughout the entire spiking period [21]. Given this inevitable variability in switching parameters, the consequent spiking dynamics undergoes variation, mainly alternating ISI in a spike train. Taking into account variation of switching parameters, following a normal distribution, ISI alternation in a single spike train could be predicted theoretically; interestingly, the results identified that the noise is featured by a Poisson-like noise—often seen in biological neurons [13]—and the uncorrelated spikes in the train [21].

The noise in the NLIF neuron inevitably elicits a noise in the gain function, which consequently leads to a noisy tuning function (e.g. Fig. 6a). Despite the presence of the noise, the average behaviour of the tuning function is comparable to the ideal one albeit noisy. This noisy tuning function brings about serious problems in information decoding. For instance, in Fig. 6b, the distributions of activities at three different stimuli (10, 20, and 30°) exhibit a remarkable overlap, so that discriminating two different stimuli from the neuronal activity is of difficult—especially if the observed activity is placed in the overlap. That is, a difficulty lies in decoding the neuronal response, implying information loss [1]. Several hypotheses on noise cancelling at the

Fig. 6 a Tuning curve of the NLIF neuron with 10% R_{on} and R_{off} variability-tolerant threshold switches. The variability elicits large error bars compared with the ideal one in Fig. 5c. The distributions of activities at three different stimuli (10, 20, and 30°) represent large overlaps that lead neurons to difficulty in decoding the encoded information, i.e. orientation in this case

network scale have been proposed and theoretically identified their feasibility albeit not experimentally evidenced [47]. Probability of noise correlation and its positive effect on information decoding have also been proposed to reduce the uncertainty in stimulus discrimination to some extent [1, 2]. In addition, a recent study has theoretically proven that uncertainty in information decoding can be largely reduced by representation of a population of neurons [21]. All these hypotheses presume that information decoding can be conducted in a statistical manner; however, a biological mechanism for probabilistic coding has barely been understood for the time being.

As stated earlier, scalability of emerging artificial neurons is an important issue. The threshold switch is perhaps scaled down without remarkable degradation of its functionality [48]. Both R_{on} and R_{off} likely increase as the cell size shrinks so that materials engineering—able to adjust the resistance accordingly—needs to follow. However, making use of capacitors is a challenge. The capacitance determines the spiking dynamics in the way that it determines a time constant of the capacitance charging—immediately following spiking—and thus ISI. Provided the shrinkage of the capacitor area and the resulting decline in the capacitance, the spiking activity is seen at high-frequency scale and it may cause technical problems regarding the compatibility with the artificial synapse devices.

3 Summary and Outlook

Neuromorphic engineering provides versatile platform technologies that are possibly applied to various functional devices that recognize patterns in general such as visual, auditory, and semantic patterns. For the time being, a few neuromorphic products are already in the pipeline and expected to appear in the market in the near future [49, 50]. They are only a few examples in sight, however, a number of precious gems are probably hidden for the moment and waiting to be discovered. Gaining a better understanding of brain functionalities, e.g. recognition, perception, long-term

memory, working memory, is an important premise for accelerating the progress in neuromorphic technologies. Towards this end, a significant portion of this chapter is devoted to the explanation as to biological neurons. Addressing artificial neuron models that are not merely delimited within the framework of hardware design is of importance given that the models are perhaps the roots of hardware artificial neurons.

Recalling the emphasis in Introduction, the target application determines the type of a neuron as well as a synapse in use; therefore, it is fruitful to diversify available artificial neuron models—ranging from a static and binary neuron to a dynamic and analogue neuron that is fairly biologically plausible. The latter may be most versatile in the light of its similarity to the biological neuron. This class of neurons follows the fidelity to the key features of the biological neuron—integrate-and-fire and variation in the response upon the input, i.e. gain function.

In designing a neuron, its detailed specifications need to be determined in combination with other significant components such as synapses and details of receiving data. Spiking neurons are required to allow synapses in connection to them to change their synaptic weights in the case that reweighting is supposed to occur. Thus, spiking dynamics, e.g. activity and spike height and width, should be compatible with the operational conditions of the synaptic weight change; otherwise, no weight change regardless of the activity of the adjacent neurons. For instance, use of an emerging memory element, e.g. phase-change memory and resistive switching memory, as an artificial synapse imposes remarkable constraints on the neuron design insomuch as the operational window of such an emerging memory is fairly narrow. In addition, when the target neuromorphic system deals with external dynamic stimulation—as our eyes do—the response rate should be comparable to the rate of the external stimulus. Otherwise, the cost of neuronal representation—approximately (the number of spikes) $\times$ (the energy consumption/spike)—outweighs the benefit, which is against one of the mottos of neuromorphic engineering: low power consumption. Thus, use of fairly large capacitors is inevitable regarding the reconciliation between low power consumption and high sensitivity to a time-varying stimulus. This serves as a significant obstacle to integrating the neuromorphic system, so that it is a challenge for the moment. A workaround solution may be to employ high dielectric constant materials representing much larger capacitance density at a given thickness compared to a conventional one such as SiO_x. Further technical issues in integration will appear with maturation of the technology, and multidisciplinary approaches to issues most likely provide shortcuts to the solutions.

References

1. Averbeck, B.B., Latham, P.E., Pouget, A.: Nat. Rev. Neurosci. **7**, 358–366 (2006)
2. Averbeck, B.B., Lee, D.: J. Neurophysiol. **95**, 3633–3644 (2006)

3. Indiveri, G., Linares-Barranco, B., Hamilton, T.J., van Schaik, A., Etienne-Cummings, R., Delbruck, T., Liu, S.-C., Dudek, P., Häfliger, P., Renaud, S., Schemmel, J., Cauwenberghs, G., Arthur, J., Hynna, K., Folowosele, F., Saighi, S., Serrano-Gotarredona, T., Wijekoon, J., Wang, Y., Boahen, K.: Front. Neurosci. **5**, 1–23 (2011)
4. Hamann, C.H., Hamnett, A., Vielstich, W.: Electrochemistry, 2nd edn. WILEY-VCH Verlag GmbH & Co., Weinheim (2007)
5. Gadsby, D.C.: Nat. Rev. Mol. Cell Biol. **10**, 344–352 (2009)
6. Jeong, D.S., Kim, I., Ziegler, M., Kohlstedt, H.: RSC Adv. **3**, 3169–3183 (2013)
7. Malenka, R.C., Nicoll, R.A.: Science **285**, 1870–1874 (1999)
8. Dayan, P., Abbott, L.F.: Theoretical Neuroscience. The MIT Press, London (2001)
9. Butts, D.A., Goldman, M.S.: PLoS Biol. **4**, e92 (2006)
10. Deneve, S., Latham, P.E., Pouget, A.: Nat. Neurosci. **4**, 826–831 (2001)
11. Knill, D.C., Pouget, A.: Trends Neurosci. **27**, 712–719 (2004)
12. Ma, W.J., Beck, J.M., Latham, P.E., Pouget, A.: Nat. Neurosci. **9**, 1432–1438 (2006)
13. Calvin, W.H., Stevens, C.F.: Science **155**, 842–844 (1967)
14. Bialek, W., Rieke, F.: de Ruyter van Steveninck R. Warland D. Science **252**, 1854–1857 (1991)
15. Gerstner, W.: Neural Comput. **12**, 43–89 (2000)
16. Chance, F.S., Abbott, L.F., Reyes, A.D.: Neuron **35**, 773–782 (2002)
17. Eliasmith, C., Anderson, C.H.: Neural Engineering: Computation, Representation, and Dynamics in Neurobiological Systems. MIT Press, Cambridge, MA (2003)
18. Anderson, C.: Computational Intelligence Imitating Life. IEEE Press, New York (1994)
19. Pouget, A., Dayan, P., Zemel, R.S.: Annu. Rev. Neurosci. **26**, 381–410 (2003)
20. Weiss, Y., Fleet, D.J.: Velocity Likelihoods in Biological and Machine Vision. In Statistical Theories of the Cortex. MIT Press, Cambridge, MA (2002)
21. Lim, H., Kornijcuk, V., Seok, J.Y., Kim, S.K., Kim, I., Hwang, C.S., Jeong, D.S.: Sci. Rep. (2015)
22. Lapicque, L.: J. Physiol. Pathol. Gen. **9**, 620–635 (1907)
23. Abbott, L.F.: Brain Res. Bull. **50**, 303–304 (1999)
24. Gerstner, W., Kistler, W.M.: Spiking Neuron Models: Single Neurons, Populations. Cambridge University Press, Plasticity (2002)
25. Jolivet, R., Lewis, T.J., Gerstner, W.: Generalized Integrate-and-Fire Models of Neuronal Activity Approximate Spike Trains of a Detailed Model to a High Degree of Accuracy, vol 92. vol 2. doi:10.1152/jn.00190.2004 (2004)
26. Hodgkin, A.L., Huxley, A.F.: J. Physiol. **117**, 500–544 (1952)
27. FitzHugh, R.: Biophys. J. **1**, 445–466 (1961)
28. Nagumo, J., Arimoto, S., Yoshizawa, S.: Proc. IRE **50**, 2061–2070 (1962)
29. Destexhe, A.: Neural Comput. **9**, 503–514 (1997)
30. Kepler, T., Abbott, L.F., Marder, E.: Biol. Cybern. **66**, 381–387 (1992)
31. Mead, C.: Analog VLSI and Neural Systems. Adison-Wesley, Reading, MA (1989)
32. Mead, C.: Proc. IEEE **78**, 1629–1636 (1990)
33. Indiveri, G., Chicca, E., Douglas, R.: IEEE Trans. Neural Netw. **17**, 211–221 (2006)
34. Pickett, M.D., Medeiros-Ribeiro, G., Williams, R.S.: Nat. Mater. **12**, 114–117 (2012)
35. Crane, H.D.: Proc. IRE **50**, 2048–2060 (1962)
36. Jeong, D.S., Thomas, R., Katiyar, R.S., Scott, J.F., Kohlstedt, H., Petraru, A., Hwang, C.S.: Rep. Prog. Phys. **75**, 076502 (2012)
37. Jeong, D.S., Lim, H., Park, G.-H., Hwang, C.S., Lee, S.: Cheong B-k. J. Appl. Phys. **111**, 102807 (2012)
38. Ovshinsky, S.R.: Phys. Rev. Lett. **21**, 1450–1453 (1968)
39. Lee, M.-J., Lee, D., Cho, S.-H., Hur, J.-H., Lee, S.-M., Seo, D.H., Kim, D.-S., Yang, M.-S., Lee, S., Hwang, E., Uddin, M.R., Kim, H., Chung, U.I., Park, Y., Yoo, I.-K.: Nat. Commun. **4**, 2629 (2013)
40. Ahn, H.-W., Jeong, D.S., Cheong, B.-K., Kim, S.-D., Shin, S.-Y., Lim, H., Kim, D., Lee, S.: ECS Solid State Lett. **2**, N31–N33 (2013)
41. Mott, N.F.: Proc. Phys. Soc. Sect. A **62**, 416 (1949)

42. Cario, L., Vaju, C., Corraze, B., Guiot, V., Janod, E.: Adv. Mater. **22**, 5193–5197 (2010)
43. Han, J.-W.: Choi Y-K Bistable resistor (biristor)—gateless silicon nanowire memory. Symp. VLSI Technol. **15–17**(2010), 171–172 (2010). doi:10.1109/vlsit.2010.5556215
44. Pickett, M.D., Williams, R.S.: Nanotechnology **23**, 215202 (2012)
45. Liu, X., Sadaf, S.M., Son, M., Shin, J., Park, J., Lee, J., Park, S., Hwang, H.: Nanotechnology **22**, 475702 (2011)
46. Bienenstock, E., Cooper, L., Munro, P.: J. Neurosci. **2**, 32–48 (1982)
47. Deneve, S., Latham, P.E., Pouget, A.: Nat. Neurosci. **2**, 740–745 (1999)
48. DerChang, K., Tang, S., Karpov, I.V., Dodge, R., Klehn, B., Kalb, J.A., Strand, J., Diaz, A., Leung, N., Wu, J., Lee, S., Langtry, T., Kuo-wei, C., Papagianni, C., Jinwook, L., Hirst, J., Erra, S., Flores, E., Righos, N., Castro, H.: Spadini G A stackable cross point Phase Change Memory. In: IEEE International Electron Devices Meeting, vol. 7–9, pp. 1–4 (2009)
49. Merolla, P.A., Arthur, J.V., Alvarez-Icaza, R., Cassidy, A.S., Sawada, J., Akopyan, F., Jackson, B.L., Imam, N., Guo, C., Nakamura, Y., Brezzo, B., Vo, I., Esser, S.K., Appuswamy, R., Taba, B., Amir, A., Flickner, M.D., Risk, W.P., Manohar, R., Modha, D.S.: Science **345**, 668–673 (2014)
50. Gehlhaar, J.: Neuromorphic processing: a new frontier in scaling computer architecture. In: Architectural Support for Programming Languages and Operating Systems, Salt Lake City, UT, USA, pp. 317–318 (2014)