



Spatial Properties of STDP in a Self-Learning Spiking Neural Network Enable Controlling a Mobile Robot

Sergey A. Lobov^{1,2*}, Alexey N. Mikhaylov¹, Maxim Shamshin¹, Valeri A. Makarov^{1,3} and Victor B. Kazantsev^{1,2}

¹ Neurotechnology Department, Lobachevsky State University of Nizhny Novgorod, Nizhny Novgorod, Russia,

² Neuroscience and Cognitive Technology Laboratory, Center for Technologies in Robotics and Mechatronics Components, Innopolis University, Innopolis, Russia, ³ Instituto de Matemática Interdisciplinar, Facultad de Ciencias Matemáticas, Universidad Complutense de Madrid, Madrid, Spain

OPEN ACCESS

Edited by:

Stefano Brivio,
Institute for Microelectronics and
Microsystems (CNR), Italy

Reviewed by:

Junxiu Liu,
Ulster University, United Kingdom
Valerio Milo,
Politecnico di Milano, Italy

*Correspondence:

Sergey A. Lobov
lobov@neuro.nnov.ru

Specialty section:

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

Received: 13 August 2019

Accepted: 22 January 2020

Published: 26 February 2020

Citation:

Lobov SA, Mikhaylov AN,
Shamshin M, Makarov VA and
Kazantsev VB (2020) Spatial
Properties of STDP in a Self-Learning
Spiking Neural Network Enable
Controlling a Mobile Robot.
Front. Neurosci. 14:88.
doi: 10.3389/fnins.2020.00088

Development of spiking neural networks (SNNs) controlling mobile robots is one of the modern challenges in computational neuroscience and artificial intelligence. Such networks, being replicas of biological ones, are expected to have a higher computational potential than traditional artificial neural networks (ANNs). The critical problem is in the design of robust learning algorithms aimed at building a “living computer” based on SNNs. Here, we propose a simple SNN equipped with a Hebbian rule in the form of spike-timing-dependent plasticity (STDP). The SNN implements associative learning by exploiting the spatial properties of STDP. We show that a LEGO robot controlled by the SNN can exhibit classical and operant conditioning. Competition of spike-conducting pathways in the SNN plays a fundamental role in establishing associations of neural connections. It replaces the irrelevant associations by new ones in response to a change in stimuli. Thus, the robot gets the ability to relearn when the environment changes. The proposed SNN and the stimulation protocol can be further enhanced and tested in developing neuronal cultures, and also admit the use of memristive devices for hardware implementation.

Keywords: spiking neural networks, spike-timing-dependent plasticity, learning, neurorobotics, neuroanimat, synaptic competition, neural competition, memristive devices

INTRODUCTION

The adoption of brain-inspired spiking neural networks (SNNs) constitutes a relatively novel paradigm in neural computations with high potential, yet not fully discovered. One of the most intriguing and promising experimental illustrations of SNNs was the development of robots controlled by biological neurons, the so-called neuroanimates, proposed at the end of the XX century and currently attracting much attention (Meyer and Wilson, 1991; Potter et al., 1997; Reger et al., 2000; Izhikevich, 2002; Pamiès et al., 2014; Dauth et al., 2016). In those experiments, neural networks self-organized in dissociated neuronal cultures, which was suggested to be used as a decision-making element in robotic systems. In the earlier 1990s, Meyer and Wilson introduced the term an animat, as a composition of words “animal” and “automat,” referring to a robot exhibiting the behavior of an animal (Meyer and Wilson, 1991). Later, several research groups developed

prototypes of hybrid systems composed of a robot controlled by a living neural network. The main idea was to achieve adaptive learning in biological SNNs with a real physical embodiment.

Learning is inevitably linked with the interaction of an agent with its environment. Therefore, to implement learning *in vitro*, a neural network should be equipped with a “body” interacting with the environment. The first neuroanimat was proposed by Mussa-Ivaldi’s group (Reger et al., 2000). To control a tiny wheeled robot Khepera, they used electric potentials recorded from brain slices of the sea lamprey fed by signals from light sensors. Almost in parallel with this study, Potter et al. (1997) suggested connecting a neuronal culture grown on a multielectrode array (MEA) to animate a roving robot (DeMarse et al., 2001). They succeeded in constructing a virtual neuroanimat capable of moving in the desired direction within 60° corridor after 2 h of “training” with a success rate of 80% (Bakkum et al., 2008). Shahaf et al. (2008) used ultrasonic sensors detecting the presence of an obstacle in the trajectory of a neuroanimat by stimulating a neuronal culture, which, in turn, controlled the movement. Obstacles located on the right or left side provoked population bursts with different spiking signatures. Then, a computer algorithm detected and classified the population bursts and moved the robot in the corresponding direction.

Despite extensive experimental studies conducted over the last decades, the high computational potential of SNNs has not been really achieved. The main problem faced by the researchers building “living computers” is the absence of robust learning algorithms. Unlike the backpropagation algorithm (Rumelhart et al., 1986) and deep learning approaches (Lecun et al., 1998), which revolutionized artificial neural networks (ANNs), SNNs still lack similar methodology. In a more general context, the learning principles of biological neural networks are not explored up to the level sufficient for designing engineering solutions (Gorban et al., 2019). Several attempts were made to adapt the backpropagation algorithm and its variations to SNNs (Hong et al., 2010; Xu et al., 2013). Within this approach, an ANN is subject to learning, and then the obtained weights are transferred with some limitations to a similar SNN (Esser et al., 2016). However, SNNs trained in such a way usually do not achieve a level of accuracy similar to their ANN counterparts. This can be explained both by the formulation of the recognition problem and by the nature of the tests (Tavanaei et al., 2019).

One of the intriguing brain features is the ability to associative learning. It is based on synaptic plasticity, most likely of a Hebbian type (Hebb, 1949). A classic example of associative learning is Pavlovian conditioning (Pavlov, 1927). Generally, it binds a conditional stimulus (CS) with an unconditional stimulus (US). The US always evokes a response in the nervous system, whereas the CS initially does not. After several presentations of the US and CS together, the nervous system starts responding to the CS alone. Hebbian associative learning can be extremely efficient, given that the neural input dimension is high enough (Gorban et al., 2019; Tyukin et al., 2019). Experimentally, associative learning is often achieved in the form of operant or instrumental conditioning, which is characterized by the

presentation of stimuli to an animal depending on its behavior (Pavlov, 1927; Hull, 1943; Dayan and Abbott, 2001).

There are several approaches to implement associative learning in mathematical models. One is to incorporate US and CS events as spiking waves or patches of activity propagating in neural tissue and associate them through a spatiotemporal interaction. Learning underlying such a “spatial computation” can be implemented by using spike-timing-dependent plasticity (STDP) (Gong and van Leeuwen, 2009; Palmer and Gong, 2014). The STDP implements the Hebbian rule. In this case, repeated arrival of presynaptic spikes a few milliseconds before the generation of postsynaptic action potentials leads to potentiation of the synapse, whereas the occurrence of presynaptic spikes after postsynaptic ones provokes synaptic depression (Markram et al., 1997; Bi and Poo, 1998; Sjöström et al., 2001). A different approach to the conditioning paradigm uses reinforcement learning, e.g. on the basis of an eligibility trace and dopamine modulated STDP (Houk et al., 1995; Izhikevich, 2007). Based on this type of plasticity, a robot interacting with humans capable of associating color and touch patterns was recently designed (Chou et al., 2015). However, this approach is quite complicated and was implemented only in model neural networks.

Many attempts to implement learning features in neuroanimats have been made in cultured neural networks grown *in vitro*. The use of synaptic plasticity as a mechanism of reinforcement or control of functional connections was demonstrated only in the case of relatively simple adaptive changes in the network. It has been suggested that the network homogeneity (e.g. unstructured connectivity) precludes the emergence of more complex forms of learning (Pimashkin et al., 2013, 2016). Earlier, we proposed an approach to explain the problems of learning in unstructured neural networks by the competition between different pathways conducting excitation to a neuron or set of neurons (Lobov S. A. et al., 2017; Lobov S. et al., 2017b). Recently, the possibility to structure the network geometry by directing axon growth was demonstrated experimentally (Malishev et al., 2015; Gladkov et al., 2017), which opens a new venue to build network architectures *in vitro*.

In this article, we study how spatial or topological properties of STDP can be used to implement associative learning in small SNNs. We show that the competition of spike-conducting pathways in a network plays an essential role in establishing the association of neural connections. In particular, on the network scale, STDP potentiates the shortest neural pathways and depresses alternative longer pathways. It permits replacing irrelevant associations by new ones in response to changes in the structure of external stimuli. We show that a roving robot controlled by an especially designed SNN can exhibit classical and operant conditioning. Application of the shortest-pathway rule allows the robot to relearn sensory-motor skills by rewiring the SNN on the fly when the environment changes. The developed SNN topology and the stimulation protocol can be adapted further for structured neural network cultured *in vitro* and for designing hardware SNNs based on, e.g. memristive plasticity.

MATERIALS AND METHODS

The SNN Model

To simulate the dynamics of a SNN, we adopt the approach described elsewhere (Lobov S. A. et al., 2017). Briefly, the dynamics of a single neuron is given by Izhikevich (2003):

$$\frac{dv}{dt} = 0.04v^2 + 5v + 140 - u + I(t), \quad (1)$$

$$\frac{du}{dt} = a(bv - u), \quad (2)$$

where v is the membrane potential, u is the recovery variable, and $I(t)$ is the external driving current. If $v \geq 30$, then $v \leftarrow c$, $u \leftarrow u + d$, which corresponds to generation of a spike. We set $a = 0.02$, $b = 0.2$, $c = -65$, and $d = 8$. Then, the neuron is silent in the absence of the external drive and generates regular spikes under a constant stimulus, which is a typical behavior of cortical neurons (Izhikevich, 2003, 2004). The driving current is given by:

$$I(t) = \zeta(t) + I_{syn}(t) + I_{stim}(t), \quad (3)$$

where $\zeta(t)$ is an uncorrelated zero-mean white Gaussian noise with variance D , $I_{syn}(t)$ is the synaptic current, and $I_{stim}(t)$ is the external stimulus. As a stimulus, we use a sequence of square electric pulses of the duration of 3 ms delivered at 10 Hz rate, with the amplitude sufficient to excite the neuron.

The synaptic current is the weighted sum of all synaptic inputs to the neuron:

$$I_{syn}(t) = \sum_j g_j w_j(t) y_j(t), \quad (4)$$

where the sum is taken over all presynaptic neurons, w_j is the strength of the synaptic coupling directed from neuron j , g_j is the scaling factor, in this paper we set them equal to 20 or -20 (Lobov S. A. et al., 2017) for excitatory and inhibitory neurons, respectively, and $y_j(t)$ describes the amount of neurotransmitters released by presynaptic neuron j .

To model the neurotransmitters, we use Tsodyks-Markram's model (Tsodyks et al., 1998) that accounts for short-term depression and facilitation. We use this model with the following parameters: the decay constant of postsynaptic currents $\tau_I = 10$ ms, the recovery time from synaptic depression $\tau_{rec} = 50$ ms, the time constant for facilitation $\tau_{facil} = 1$ s.

The dynamics of the synaptic weight w_{ij} of coupling from an excitatory presynaptic neurons j to a postsynaptic neuron i is governed by the STDP with two local variables (Song et al., 2000; Morrison et al., 2008). Assuming that τ_{ij} is the time delay of spike transmission between neurons j and i , a presynaptic spike fired at time t_j and arriving to neuron i at $t_j + \tau_{ij}$ induces a weight decrease proportional to the value of the postsynaptic trace s_i . Similarly, a postsynaptic spike at t_i induces a weight potentiation proportional to the value of the presynaptic trace s_j . The weighting

functions obey the multiplicative updating rule (Song et al., 2000; Morrison et al., 2008). Thus, the weight dynamics is given by:

$$\frac{ds_i}{dt} = -\frac{s_i}{\tau_S} + \sum_{t_i} \delta(t - t_i), \quad (5)$$

$$\frac{ds_j}{dt} = -\frac{s_j}{\tau_S} + \sum_{t_j} \delta(t - t_j - \tau_{ij}), \quad (6)$$

$$\frac{dw_{ij}}{dt} = \lambda [(1 - w_{ij})s_j\delta(t - t_i) - \alpha w_{ij}s_i\delta(t - t_j - \tau_{ij})], \quad (7)$$

where $\tau_S = 10$ ms is the time constant of spiking traces, $\lambda = 0.001$ is the learning rate, and $\alpha = 5$ is the asymmetry parameter.

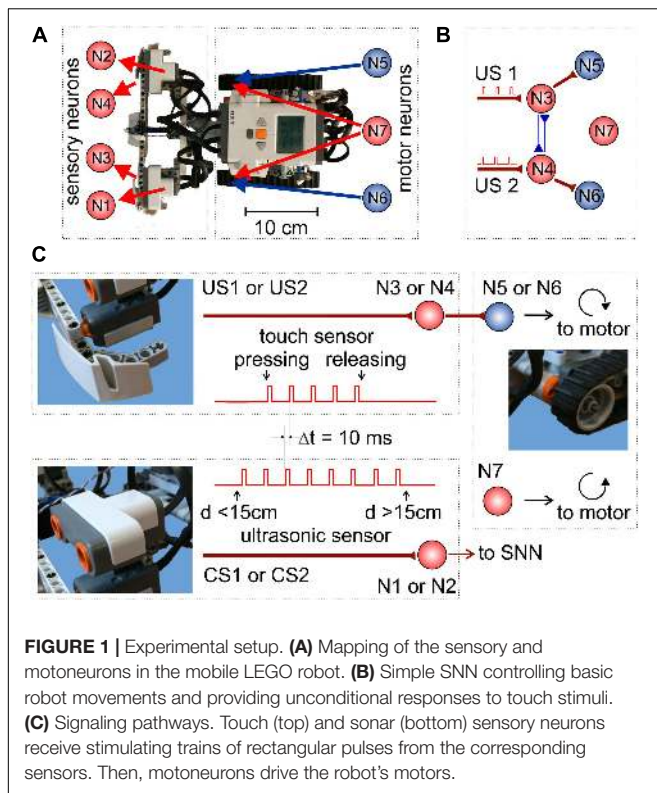
We implemented the SNN model (see below) as custom software NeuroNet developed in QT C++ environment. For the axonal delays, we used $\tau_{ij} = 3$ ms for parallel connections and $\tau_{ij} = 4.2$ ms for diagonal coupling. The selected delays are proportional to the interneuron distances and thus take into account the network topology. The app supports SNNs with up to 10^4 neurons. On an Intel® Core™ i3 processor, the simulation can be performed in real time for a SNN with tens of neurons.

Mobile Robot and Unconditional Motor Response

We built a robotic platform from a LEGO® NXT Mindstorms® kit. **Figure 1A** shows the mapping of the robot sensors and motors to the sensory- and motoneurons, respectively. NeuroNet software was used to implement SNNs of different types controlling the robot behavior. **Figure 1B** illustrates the simplest SNN providing the robot with unconditional responses to touching events (see below). The software was run on a standalone PC connected to the robot controller through a Bluetooth interface.

The robot is equipped with two touch sensors and two ultrasonic sonars (**Figure 1C**). A sensitive bumper detects touch stimuli (collisions with obstacles) from the left and right side of the robot (**Figure 1B**). When a touch sensor is on, the corresponding sensory neuron (either N3 or N4) is stimulated by a train of pulses delivered at 10 Hz rate (**Figure 1C**, top-left panel). Such stimulation models signal processing in the sensory system of animals. The ultrasonic sonars are located above the bumper and are coupled to sensory neurons N1 and N2 (**Figure 1C**, bottom-left panel). A sonar sensor turns on if the distance to an obstacle is less than 15 cm. Then, the corresponding neuron is stimulated by a train of square pulses delivered at 10 Hz rate.

The SNN controls the robot movements through the activation of motoneurons. Motor neuron N7 produces tonic spiking with the mean frequency F , which is mapped simultaneously to the left and right motors. As a result, the robot moves straightforward with the velocity proportional to F . Neurons N5 and N6 are coupled to the right and left motors, respectively. The amount of



neurotransmitters released by these neurons modulates the rotation velocity of the corresponding motor. When N5 (N6) fires, the right (left) motor slows down (or even rotates backward if, e.g. $F = 0$), and the robot turns to the right (left).

The robot also has three LEDs facilitating its recognition in the arena by a zenithal video camera. Video frames, captured at 29 Hz rate, were analyzed offline. Trajectory tracking was performed by employing a computer vision algorithm implemented in the OpenCV library. Robot detection is based on the fact that the robot image is a high gradient area. The LEDs turn off when a touch sensor is activated, which allows such events to be detected by analyzing the overall glow of the robot image.

The touch sensors mediate US (**Figures 1B,C**, top). When one of them is activated due to a collision with an obstacle, the corresponding sensory neuron (N3 or N4) starts firing and directly excites a motoneuron (N5 or N6, **Figure 1B**). As a result, the corresponding motor starts rotating backward, and the robot turns away from the obstacle and thus avoids the negative stimulus (**Supplementary Video S1**).

The sonars are connected to sensory neurons N1 and N2 and mediate CS. At the beginning of learning, the CS in the form of an approaching obstacle does not evoke any robot's response. The goal of learning is to associate CS with US to avoid the obstacles in advance without touching them. To provide stimulation of "sensory neurons", according to the STDP protocol, the stimulating pulses from the touch sensors have a 10-ms delay relative to the sonar pulses (**Figure 1C**).

RESULTS

The Shortest Pathway Rule

Let us consider a pair of unidirectionally coupled neurons driven by periodic stimuli applied to one of them (**Figure 2A**). Stimuli excite the first neuron, and then the activation propagates along the "chain" to the second cell, which fires, given that the coupling strength w_{21} is strong enough. Then, the presynaptic spikes precede the postsynaptic ones, and, as a result, the weight increases following the STDP rule (the first term in the right-hand side of Eq. 7). Such a situation can be extended into a chain of three or even more neurons (**Figure 2B**). Thus, STDP increases the corresponding synaptic weights.

However, if we add a new connection from the first neuron to the third one (**Figure 2C**), the weight dynamics changes crucially. Although all synapses are excitatory, the coupling directed from the second to the third neuron is depressed, while the other two are potentiated. This occurs because the axonal delay via the direct way N1–N3 (τ_{31} , **Figure 2C**) is significantly shorter than the delay via the pathway N1–N2–N3 ($\tau_{321} = \tau_{21} + \tau_{32}$, **Figure 2C**). Thus, the first neuron makes fire directly the third one (which is also postsynaptic for w_{32}), and its spikes appear ahead of the spikes coming from the second neuron (presynaptic for w_{32}). Such an inverse sequence (**Figure 2D**) forces depression of the coupling w_{32} according to the STDP rule (the second term in the right-hand side of Eq. 7). We thus can formulate the shortest pathway rule:

- *On the network scale, STDP potentiates the shortest neural pathways and depresses alternative longer pathways.*

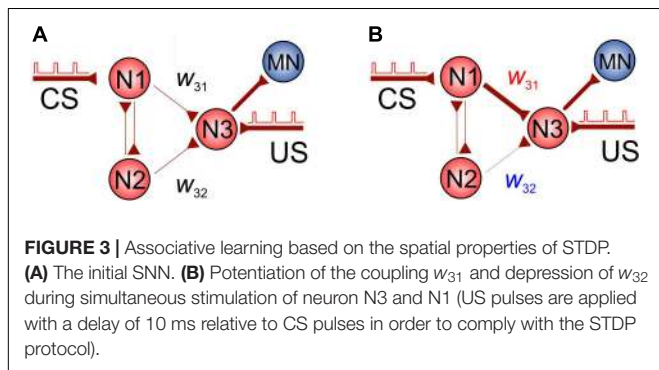
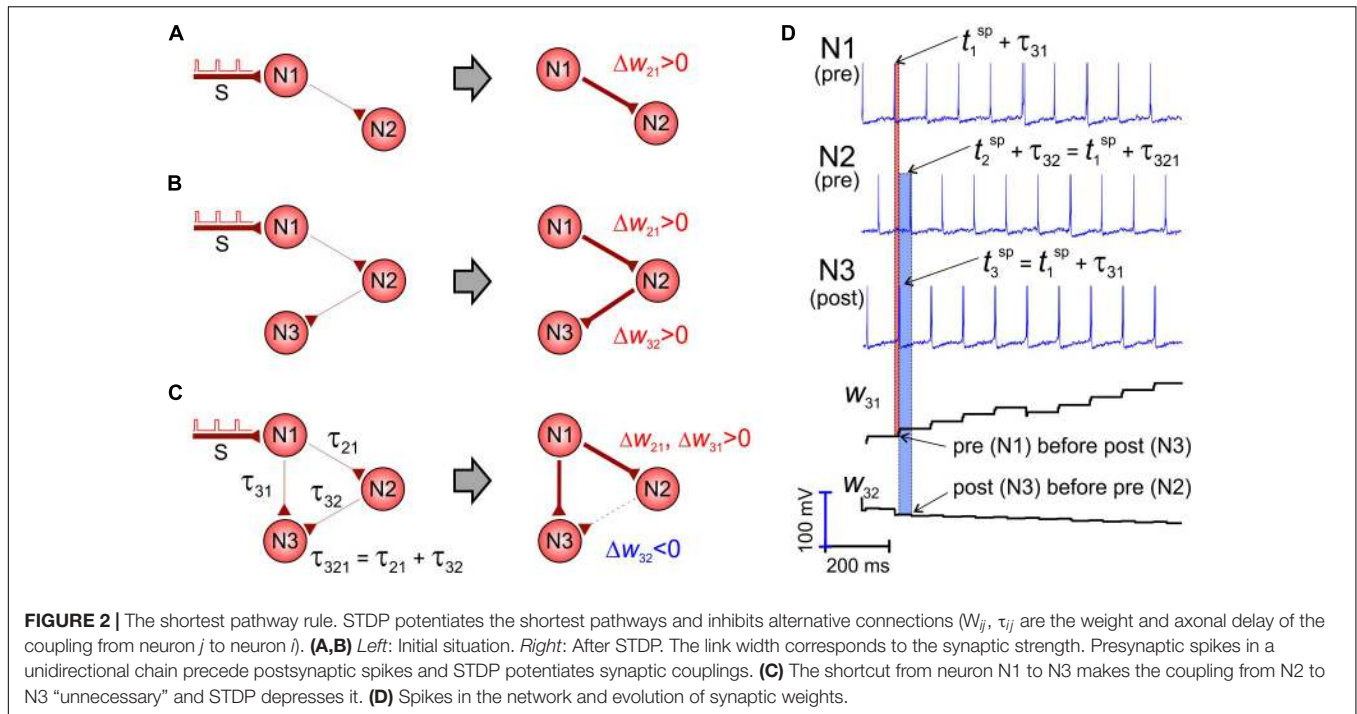
SNN Exhibiting Non-trivial Associative Learning

Let us now employ the shortest-pathway rule to implement conditional learning in an SNN. **Figure 3A** shows a simple SNN consisting of four neurons, which can exhibit associative learning. The SNN receives two types of inputs: CS and US applied to neurons N1 and N3, respectively. To comply with the STDP protocol of paired stimulation, we assume that the US pulses arrive with a delay of 10 ms relative to CS pulses (see also **Figure 1C**).

At the beginning, the coupling between N1 and N3, w_{31} , is not sufficient to excite N3 through the CS pathway. However, under stimulation, it is potentiated due to the appropriate delay between US and CS. At the same time, the coupling between N2 and N3, w_{32} , is depressed due to the shortest pathway rule. Thus, after learning, we get the network shown in **Figure 3B** and the CS alone can activate neuron N3 and then the motoneuron. We also note that, similarly, if the CS is applied to N2 instead of N1, then w_{32} will be potentiated, while w_{31} depressed, and we get the same effect of associative learning.

SNN Driving Robot

The above-discussed SNN (**Figure 3**) has one motoneuron and hence can drive one motor channel. To process events coming from the right and left sensors of the robot, we need to extend the SNN to account for two motor channels. Thus, we duplicate



the SNN shown in **Figure 3** but, at the same time, share some of the neurons between two copies of the SNN (**Figure 4A**). The resulting SNN contains four sensory neurons (N1, N2 for CS and N3, N4 for US, **Figure 4A**) and two motoneurons N5, N6 modulating the rotation velocities of the left and right motors, respectively (see also **Figure 1**). Neurons N3 and N4 are mutually inhibitory coupled with fixed synaptic weights ($w_{34} = w_{43} = 1$).

The pair of neurons receiving CS (N1, N2) can be connected to the pair of sonars in an arbitrary order (left-right or right-left). Depending on the connection, there can be two types of associations between the stimuli and motors: either with strong “parallel” (PA) or strong “diagonal” (DA) pathways (**Figure 4B**). Such freedom ensures that there is no *a priori* chosen structure in the complete SNN. Instead, the SNN adapts to the stimuli coming from the environment. Thus, the mutual exchange of the CS sources can simulate a situation with a change in the environment, which should induce relearning in the SNN and adaptation

to novel conditions. Note that the bidirectional coupling between neurons N1 and N2 plays a fundamental role by providing synaptic competition while training couplings to neurons N3 and N4.

Classical or Pavlovian Conditioning

To implement Pavlovian (classical) conditioning, let us, for a moment, deactivate neuron N7 responsible for forward movement. If an object approaches the robot from one side, the corresponding touch sensor is activated, and we get an unconditional response (**Figure 4C** and **Supplementary Video S1**). At the same time, the corresponding sonar is also triggered on, and paired trains of stimuli innervate sensory neurons with a time delay of 10 ms.

We repeated such a stimulation alternately on the left and right sides of the robot. This protocol led to the potentiation of two associations for the left and right sides. Five stimulating cycles applied to the right and left sides were sufficient to achieve robust learning. After switching the connections of the sonars between sensory neurons N1 and N2, the SNN was able to relearn the associations (i.e. to switch between PA and DA, **Figure 4B**) after about 10–15 stimulus cycles.

In practice, to avoid obstacles successfully, the robot should gain high selectivity of the right and left channels. Then, in the presence of an obstacle on the left side, neuron N5 fires while neuron N6 is silent, which occurs in part due to inhibitory connections between neurons N3 and N4. Experimentally, the channel selectivity can be monitored by measuring the ratio of synaptic weights of “parallel” and “diagonal” connections:

$$w_P = (w_{31} + w_{42})/2, w_D = (w_{41} + w_{32})/2. \tag{8}$$

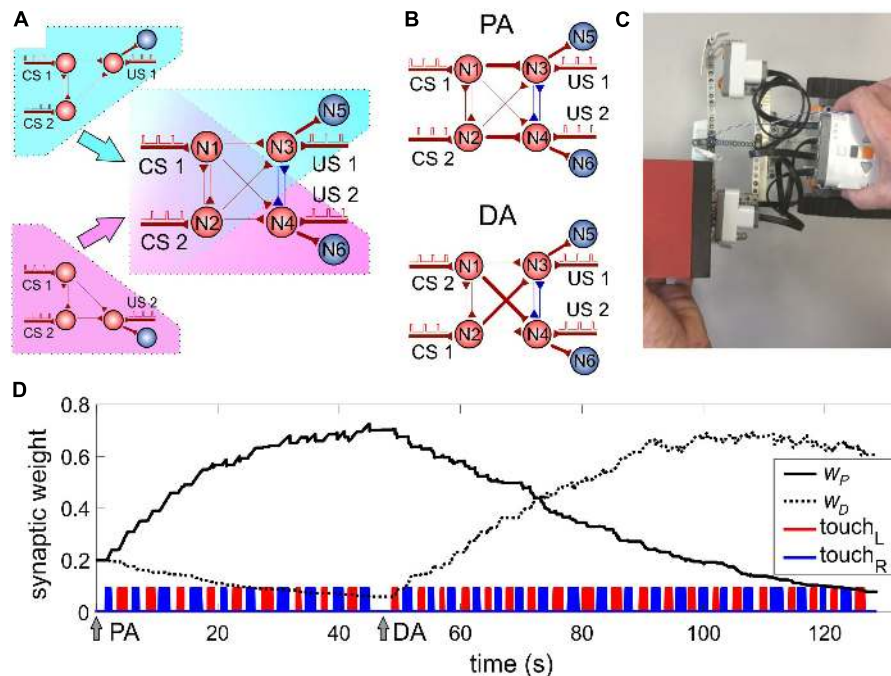


FIGURE 4 | Model of classical conditioning. **(A)** The design of a two-channel SNN by duplicating the single-channel SNN (Figure 3) sharing some neurons. The neural circuit includes neurons N1–N4 involved in learning. Motoneurons N5 and N6 provide turning the robot away from an obstacle. **(B)** The SNN after learning. PA, parallel association: N1 (N2) is associated with N3 (N4), couplings w_{31} and w_{42} are potentiated. DA, diagonal association: N2 (N1) is associated with N3 (N4), couplings w_{32} and w_{41} are potentiated. **(C)** Application of a stimulus to the touch and sonar sensors. **(D)** Evolution of the average weights of parallel (w_P) and diagonal (w_D) couplings under classical conditioning. Arrows PA and DA denote the time instants of the beginning of learning with correspondent scheme of the US mapping; touch_L (touch_R) is the time course of triggering the left (right) touch sensor.

Figure 4D shows the dynamics of these connections when simulating classical conditioning. Note that in the case of PA, the parallel connection w_P is potentiated, while the diagonal connection w_D is depressed. This happens due to simultaneous potentiation/depression of the pairs (w_{31} , w_{42}) and (w_{41} , w_{32}), according to the shortest pathway rule. After switching the CS inputs (Figure 4D, DA arrow), the opposite effect is observed, which leads to relearning in the SNN.

To achieve a high learning rate, our experiments show that the SNN should satisfy the following conditions:

1. Intermediate noise variance ($D = 5.5$ in experiments).
2. Bidirectional coupling between CS neurons (N1 and N2, Figure 4A).
3. Couplings between CS and US neurons are STDP-driven.
4. Inhibitory connections between US neurons (N3 and N4, Figure 4A).

Condition (1) agrees with our previous findings showing that the network rearrangement under stimulation takes place in a certain interval of the noise intensity (Lobov S. A. et al., 2017). At low noise intensity, the neuronal activation may not reach the level necessary for STDP-ordered pre and post-synaptic spiking. At high noise intensity, random STDP events dominate and break learning (see Supplementary Figure S1). Condition (2) expresses competition between the synapses involved in the associations increasing the SNN selectivity. Thus, competition plays a positive

role in learning, unlike the case study reported previously (Lobov S. et al., 2017b). Condition (3) implies a reduction of the SNN selectivity due to a negative effect that STDP can have on the synaptic couplings between CS neurons (w_{21} and w_{12}). Condition (4) leads to competition between neurons “for the right” to be activated and, as a result, to an increase in the selectivity of the connections of the right and left channels.

Operant or Instrumental Conditioning

Animals learn behaviors through active interaction with the environment. To model such natural learning, we use operant (or instrumental) conditioning. To implement it, we activated motoneuron N7 (Figures 1B,C) responsible for forward movement and introduced the robot in an arena with several obstacles (Figure 5A).

In the beginning, the robot could avoid obstacles only after touching them due to US (Figure 5A). Then, learning progressively established associations between approaching obstacles (sonars, CS) and touching events (US). Thus, the robot learned to avoid obstacles in advance, without touching them (Figure 5B and Supplementary Videos S2, S3). We then switched sonars. Similarly to classical conditioning, the robot was able to relearn the associations (Figure 5C, PA arrow).

The learning rate depends on the total time of activation of the touch sensors. In turn, this time depends on the configuration of the arena, i.e. the arena size and the number of obstacles.

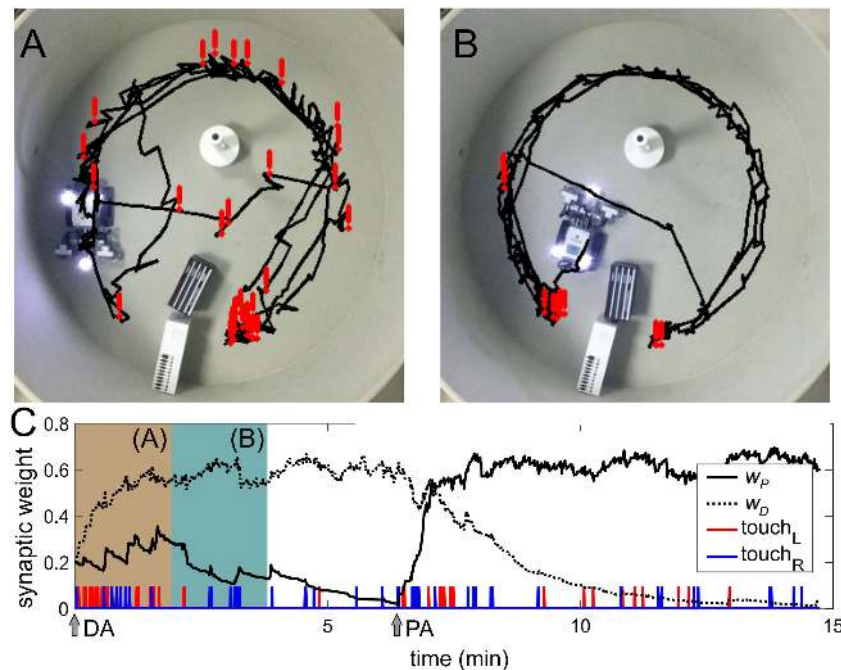


FIGURE 5 | Operant conditioning. **(A)** Trajectory of the robot in the first 2 min of the experiment. Exclamation marks indicate the positions of collisions with obstacles. **(B)** Same as in **(A)** but after learning. **(C)** Evolution of the weights of parallel (w_P) and diagonal (w_D) couplings (compare to **Figure 4D**). Beige and green-blue bars correspond to periods **(A,B)**, respectively.

In the Morris water maze (**Figure 5A**, 1 m²), learning takes about 2 min. In a larger room (50 m²) with a few obstacles, the learning time increases to 10–20 min. Relearning takes about twice a longer time.

In the operant conditioning, the SNN selectivity did not reach the value achieved in classical conditioning (compare **Figures 4D**, **5C**). It occurs due to the fact that in the arena, the robot can approach objects in front. In this case, both sonars detect them, which leads to a simultaneous generation of stimuli on the left and right sides and competition between two connections from the same sensory neuron. Technical constraints, such as a narrow sensing angle of the sonars, also affect the correct implementation of the obstacle-avoidance task negatively. All these factors diminish the learning quality. Therefore, the robot sometimes collides with obstacles. Thus, in a real environment, learning does not reach 100% collision avoidance.

DISCUSSION

Competition is a universal paradigm well-extended both in neurophysiology, e.g. in the form of lateral inhibition (Kandel et al., 2000) and the ANN studies, e.g. in the form of competitive learning in Kohonen networks (Kohonen, 1982) or imitation learning (Calvo Tapia et al., 2018). In this work, we have proposed an SNN model implementing associative learning through an STDP protocol and temporal coding of sensory stimuli. To achieve successful learning, the SNN makes use of two mechanisms of competition. The first type is neuronal

competition, i.e. different neurons compete to be the first to get excited. In our case, this mechanism was provided by inhibitory connections between US neurons.

The second type of mechanism is synaptic competition; i.e. different synaptic inputs to a single neuron compete to be the one exciting the neuron. This mechanism has been less addressed in the literature on learning. Earlier, it was shown that in unstructured networks, synaptic competition leads to negative consequences for learning (Lobov S. A. et al., 2017; Lobov S. et al., 2017b). We have shown that the proposed structured architecture of the SNN, together with synaptic competition implementing the STDP-mediated rule of the shortest pathway, can ensure learning. We also note that the proposed mechanism of synaptic competition works well in the case of temporal coding of stimuli. Stimulus coding by the firing rate may require the development of a different approach. For example, in our recent study (Lobov et al., 2020), we implemented synaptic competition using synaptic forgetting, depending on the activity of the postsynaptic neuron. This allowed performing a mixed type of coding (temporal and rate) in the problem of recognition of electromyographic signals.

To test the SNN, we used it for controlling a mobile robot. We have shown that indeed, the robot exhibits successful learning at the behavioral level in the form of classical and operant conditioning. During navigation in an arena, the SNN self-organizes in such a way that after learning, the robot avoids obstacles without collisions, relying on CS only. Moreover, it can also relearn if the connection of CS sensors is switched between the corresponding sensory neurons, and a network rewiring, widely observed in biological neural networks, is

required (Calvo Tapia et al., 2020). The mechanism of relearning can be considered as a model of the animals' ability to adapt to changes in the environment. In the SNN, it is possible due to synaptic competition. Our experiments have also shown that learning is robust. The robot can operate in environments of different sizes and with varying densities of obstacles.

The proposed SNN implements a model with two associations: left and right sensors "coupled" to the right and left turns. In general, such associative learning can be extended to multiple inputs and outputs. Thus, the proposed architecture can be considered as a perceptron composed of spiking neurons with two inputs and two outputs, where logical 1 or 0 at an input corresponds to the presence or absence of a CS, respectively. Then, the US provides a learning mechanism on how to excite the target neuron in the output layer, i.e. how to obtain the desired output. Thus, we get a simple mechanism for supervised learning, i.e. a replacement of the backpropagation algorithm for SNNs. However, the question of how many neurons such a spiking perceptron can contain and, hence, how many classes can be discriminated in this way requires additional studies.

We note that the parameters of sensory stimuli play a crucial role in the learning of behaviors. For example, longer delays between stimuli or their inverse order (CS after US) can impair learning. In this sense, the temporal coding in SNNs requires fine-tuning of the neuronal circuits and maybe not robust. The rate coding using, e.g. the triplet-based STDP rule (Pfister and Gerstner, 2006), voltage-based STDP with homeostasis (Clopath et al., 2010), or STDP together with BCM rule (Wade et al., 2008; Liu et al., 2019) is likely to increase the reliability of robot control. However, in this case, we may end up with a mixed type of coding (temporal and rate).

Due to structural simplicity, the proposed SNN and the learning algorithm admit a hardware implementation by, e.g. using memristors, which are adaptive circuit elements with memory. Memristors change their resistance depending on the history of electrical stimulation (Wang et al., 2019). Since the first experiments and simulations (Linares-Barranco et al., 2011), significant progress has been achieved in the implementation of excitatory and inhibitory STDP by using resistive-switching devices (RRAM), which are a particular class of memristors with two-terminal metal-insulator-metal structure. Although most of STDP demonstrations still rely on a time overlap of pre- and postsynaptic spikes (Yu et al., 2011; Kuzum et al., 2013; Emelyanov et al., 2019), the rich internal dynamics of higher-order memristive devices related to multi-time-scale microscopic transport phenomena provides timing- and frequency-dependent plasticity in response to non-overlapping input signals in a biorealistic fashion (Du et al., 2015; Kim et al., 2015). Memristive plasticity can be realized at different time scales, in particular with STDP windows of the order of microseconds (Kim et al., 2015), which is essential for the development of fast spike encoding systems.

Upon reaching the technology maturity, arrays of memristive synapses offer unique scalability being integrated with CMOS layers and showing spatiotemporal functions (Wang W. et al., 2018), as well as combined with artificial memristive neurons (Wang Z. et al., 2018) within a single network. Simple spiking

architectures of Pavlov's dog association have been proposed on memristors (Ziegler et al., 2012; Milo et al., 2017; Tan et al., 2017; Minnekhanov et al., 2019). However, more sophisticated architectures are required to reproduce different types of associative learning to be adopted in advanced robotic systems. We anticipate that, soon, artificial neurons can be realized on the CMOS architecture, whereas the STDP can be implemented by incorporating memristors (Emelyanov et al., 2019). It seems convenient to have paired micro-scaled memristive devices to reproduce bipolar synaptic weights. They can be mounted in a standard package for easier integration into the SNN circuits.

Finally, we also foresee that the provided architecture can be implemented in biological neural networks grown in neuronal cultures *in vitro*. Modern technology of microfluidic channels permits building different network architectures (Gladkov et al., 2017). On the one hand, such a living SNN could verify if our understanding of the learning mechanism at the cell level is correct. From the other side, biological neurons have a much higher level of flexibility mediated by different molecular mechanisms that may shed light on how learning and sensory-motor control are organized in nature.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

SL, VM, and VK conceived and designed the research. SL developed the NeuroNet software, designed the robot configuration, and implemented the control of the robot by SNN. SL and MS carried out experiments with the robot and video tracking the movements of the robot. AM suggested the approach for emulating synaptic plasticity by memristive microdevices. All authors participated in the interpretation of the results and wrote the manuscript.

FUNDING

This work was supported by the Russian Science Foundation (project 19-12-00394, the SNN development and implementation), by the Russian Foundation for Basic Research (grant 18-29-23001, the emulation of synaptic plasticity by using memristive microdevices), and by the Spanish Ministry of Science, Innovation, and Universities (grant FIS2017-82900-P, the analysis of the robot dynamics).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2020.00088/full#supplementary-material>

REFERENCES

- Bakkum, D. J., Chao, Z. C., and Potter, S. M. (2008). Spatio-temporal electrical stimuli shape behavior of an embodied cortical network in a goal-directed learning task. *J. Neural Eng.* 5, 310–323. doi: 10.1088/1741-2560/5/3/004
- Bi, G., and Poo, M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* 18, 10464–10472. doi: 10.1523/jneurosci.18-24-10464.1998
- Calvo Tapia, C., Makarov, V. A., and van Leeuwen, C. (2020). Basic principles drive self-organization of brain-like connectivity structure. *Commun. Nonlinear Sci. Numer. Simul.* 82:105065. doi: 10.1016/j.cnsns.2019.105065
- Calvo Tapia, C., Tyukin, I. Y., and Makarov, V. A. (2018). Fast social-like learning of complex behaviors based on motor motifs. *Phys. Rev.* 97:052308. doi: 10.1103/PhysRevE.97.052308
- Chou, T.-S., Bucci, L., and Krichmar, J. (2015). Learning touch preferences with a tactile robot using dopamine modulated STDP in a model of insular cortex. *Front. Neurobot.* 9:6. doi: 10.3389/fnbot.2015.00006
- Clopath, C., Büsing, L., Vasilaki, E., and Gerstner, W. (2010). Connectivity reflects coding: a model of voltage-based STDP with homeostasis. *Nat. Neurosci.* 13, 344–352. doi: 10.1038/nn.2479
- Dauth, S., Maoz, B. M., Sheehy, S. P., Hemphill, M. A., Murty, T., Macedonia, M. K., et al. (2016). Neurons derived from different brain regions are inherently different in vitro: a novel multiregional brain-on-a-chip. *J. Neurophysiol.* 117, 1320–1341. doi: 10.1152/jn.00575.2016
- Dayan, P., and Abbott, L. F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: The MIT Press.
- DeMarse, T. B., Wagenaar, D. A., Blau, A. W., and Potter, S. M. (2001). The neurally controlled animat: biological brains acting with simulated bodies. *Auton. Robots* 11, 305–310.
- Du, C., Ma, W., Chang, T., Sheridan, P., and Lu, W. D. (2015). Biorealistic implementation of synaptic functions with oxide memristors through internal ionic dynamics. *Adv. Funct. Mater.* 25, 4290–4299. doi: 10.1002/adfm.201501427
- Emelyanov, A. V., Nikiruy, K. E., Demin, V. A., Rylkov, V. V., Belov, A. I., Korolev, D. S., et al. (2019). Ytria-stabilized zirconia cross-point memristive devices for neuromorphic applications. *Microelectron. Eng.* 215:110988. doi: 10.1016/j.mee.2019.110988
- Esser, S. K., Merolla, P. A., Arthur, J. V., Cassidy, A. S., Appuswamy, R., Andreopoulos, A., et al. (2016). Convolutional networks for fast, energy-efficient neuromorphic computing. *Proc. Natl. Acad. Sci. U.S.A.* 113, 11441–11446. doi: 10.1073/pnas.1604850113
- Gladkov, A., Pigareva, Y., Kutyna, D., Kolpakov, V., Bukatin, A., Mukhina, I., et al. (2017). Design of cultured neuron networks in vitro with predefined connectivity using asymmetric microfluidic channels. *Sci. Rep.* 7:15625. doi: 10.1038/s41598-017-15506-2
- Gong, P., and van Leeuwen, C. (2009). Distributed dynamical computation in neural circuits with propagating coherent activity patterns. *PLoS Comput. Biol.* 5:e1000611. doi: 10.1371/journal.pcbi.1000611
- Gorban, A. N., Makarov, V. A., and Tyukin, I. Y. (2019). The unreasonable effectiveness of small neural ensembles in high-dimensional brain. *Phys. Life Rev.* 29, 55–88. doi: 10.1016/j.plrev.2018.09.005
- Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. New York, NY: Wiley.
- Hong, S., Ning, L., Xiaoping, L., and Qian, W. (2010). “A cooperative method for supervised learning in Spiking neural networks,” in *Proceedings of the 14th International Conference on Computer Supported Cooperative Work in Design, Shanghai*.
- Houk, J., Adams, J., and Barto, A. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. *Model. Inf. Process. Basal Ganglia* 13.
- Hull, C. L. (1943). *Principles of Behavior: An Introduction to Behavior Theory*. New York, NY: Appleton-Century-Crofts.
- Izhikevich, E. M. (2002). Resonance and selective communication via bursts in neurons having subthreshold oscillations. *Biosystems* 67, 95–102. doi: 10.1016/S0303-2647(02)00067-9
- Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Trans. Neural Netw.* 14, 1569–1572. doi: 10.1109/TNN.2003.820440
- Izhikevich, E. M. (2004). Which model to use for cortical spiking neurons? *IEEE Trans. Neural Netw.* 15, 1063–1070. doi: 10.1109/TNN.2004.832719
- Izhikevich, E. M. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex* 17, 2443–2452. doi: 10.1093/cercor/bhl152
- Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (2000). *Principles of Neural Science*, 4th Edn, New York, NY: McGraw-Hill.
- Kim, S., Du, C., Sheridan, P. J., Ma, W., Choi, S., and Lu, W. D. (2015). Experimental demonstration of a second-order memristor and its ability to biorealistically implement synaptic plasticity. *Nano Lett.* 15, 2203–2211. doi: 10.1021/acs.nanolett.5b00697
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biol. Cybern.* 43, 59–69. doi: 10.1007/bf00337288
- Kuzum, D., Yu, S., and Philip Wong, H.-S. (2013). Synaptic electronics: materials, devices and applications. *Nanotechnology* 24:382001. doi: 10.1088/0957-4484/24/38/382001
- Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2324. doi: 10.1109/5.726791
- Linares-Barranco, B., Serrano-Gotarredona, T., Camuñas-Mesa, L., Perez-Carrasco, J., Zamarreño-Ramos, C., and Masquelier, T. (2011). On spike-timing-dependent-plasticity, memristive devices, and building a self-learning visual cortex. *Front. Neurosci.* 5:26. doi: 10.3389/fnins.2011.00026
- Liu, J., Mcdaid, L. J., Harkin, J., Karim, S., Johnson, A. P., Millard, A. G., et al. (2019). Exploring self-repair in a coupled spiking astrocyte neural network. *IEEE Trans. Neural Netw. Learn. Syst.* 30, 865–875. doi: 10.1109/TNNLS.2018.2854291
- Lobov, S. A., Chernyshov, A. V., Krilova, N. P., Shamshin, M. O., and Kazantsev, V. B. (2020). Competitive learning in a spiking neural network: towards an intelligent pattern classifier. *Sensors* 20:E500. doi: 10.3390/s20020500
- Lobov, S. A., Zhuravlev, M. O., Makarov, V. A., and Kazantsev, V. B. (2017). Noise enhanced signaling in STDP driven spiking-neuron network. *Math. Model. Nat. Phenom.* 12, 109–124. doi: 10.1051/mmnp/2017121409
- Lobov, S., Balashova, K., Makarov, V. A., and Kazantsev, V. (2017b). “Competition of spike-conducting pathways in stdp driven neural networks,” in *Proceedings of the 5th International Congress on Neurotechnology, Electronics and Informatics, Setúbal*.
- Malishev, E., Pimashkin, A., Arseniy, G., Pigareva, Y., Bukatin, A., Kazantsev, V., et al. (2015). Microfluidic device for unidirectional axon growth. *J. Phys.* 643:12025. doi: 10.1088/1742-6596/643/1/012025
- Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275, 213–215. doi: 10.1126/science.275.5297.213
- Meyer, J. A., and Wilson, S. W. (1991). *From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior*. Cambridge: MIT Press.
- Milo, V., Ielmini, D., and Chicca, E. (2017). “Attractor networks and associative memories with STDP learning in RRAM synapses,” in *Proceedings of the 2017 IEEE International Electron Devices Meeting (IEDM)*, San Francisco, CA.
- Minnekhanov, A. A., Emelyanov, A. V., Lapkin, D. A., Nikiruy, K. E., Shvetsov, B. S., Nesmelov, A., et al. (2019). Parylene based memristive devices with multilevel resistive switching for neuromorphic applications. *Sci. Rep.* 9:10800. doi: 10.1038/s41598-019-47263-9
- Morrison, A., Diesmann, M., and Gerstner, W. (2008). Phenomenological models of synaptic plasticity based on spike timing. *Biol. Cybern.* 98, 459–478. doi: 10.1007/s00422-008-0233-1
- Palmer, J. H. C., and Gong, P. (2014). Associative learning of classical conditioning as an emergent property of spatially extended spiking neural circuits with synaptic plasticity. *Front. Comput. Neurosci.* 8:79. doi: 10.3389/fncom.2014.00079
- Pamies, D., Hartung, T., and Hogberg, H. T. (2014). Biological and medical applications of a brain-on-a-chip. *Exp. Biol. Med.* 239, 1096–1107. doi: 10.1177/1535370214537738
- Pavlov, I. P. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. Oxford: Oxford University Press.
- Pfister, J.-P., and Gerstner, W. (2006). Triplets of spikes in a model of spike timing-dependent plasticity. *J. Neurosci.* 26, 9673–9682. doi: 10.1523/JNEUROSCI.1425-06.2006

- Pimashkin, A., Gladkov, A., Agrba, E., Mukhina, I., and Kazantsev, V. (2016). Selectivity of stimulus induced responses in cultured hippocampal networks on microelectrode arrays. *Cogn. Neurodyn.* 10, 287–299. doi: 10.1007/s11571-016-9380-6
- Pimashkin, A., Gladkov, A., Mukhina, I., and Kazantsev, V. (2013). Adaptive enhancement of learning protocol in hippocampal cultured networks grown on multielectrode arrays. *Front. Neural Circuits* 7:87. doi: 10.3389/fncir.2013.00087
- Potter, S. M., Fraser, S. E., and Pine, J. (1997). “Animat in a petri dish: cultured neural networks for studying neural computation,” in *Proceedings of the 4th Joint Symposium on Neural Computation, UCSD, San Diego, CA*.
- Reger, B. D., Fleming, K. M., Sanguineti, V., Alford, S., and Mussa-Ivaldi, F. A. (2000). Connecting brains to robots: an artificial body for studying the computational properties of neural tissues. *Artif. Life* 6, 307–324. doi: 10.1162/106454600300103656
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature* 323, 533–536. doi: 10.1038/323533a0
- Shahaf, G., Eytan, D., Gal, A., Kermany, E., Lyakhov, V., Zrenner, C., et al. (2008). Order-based representation in random networks of cortical neurons. *PLoS Comput. Biol.* 4:e1000228. doi: 10.1371/journal.pcbi.1000228
- Sjöström, P. J., Turrigiano, G. G., and Nelson, S. B. (2001). Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron* 32, 1149–1164. doi: 10.1016/S0896-6273(01)00542-6
- Song, S., Miller, K. D., and Abbott, L. F. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nat. Neurosci.* 3:919. doi: 10.1038/78829
- Tan, Z.-H., Yin, X.-B., Yang, R., Mi, S.-B., Jia, C.-L., and Guo, X. (2017). Pavlovian conditioning demonstrated with neuromorphic memristive devices. *Sci. Rep.* 7:713. doi: 10.1038/s41598-017-00849-7
- Tavanaei, A., Ghodrati, M., Kheradpisheh, S. R., Masquelier, T., and Maida, A. (2019). Deep learning in spiking neural networks. *Neural Netw.* 111, 47–63. doi: 10.1016/j.neunet.2018.12.002
- Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput.* 10, 821–835. doi: 10.1162/089976698300017502
- Tyukin, I., Gorban, A. N., Calvo, C., Makarova, J., and Makarov, V. A. (2019). High-dimensional brain: a tool for encoding and rapid learning of memories by single neurons. *Bull. Math. Biol.* 81, 4856–4888. doi: 10.1007/s11538-018-0415-5
- Wade, J. J., McDavid, L. J., Santos, J. A., and Sayers, H. M. (2008). “SWAT: an unsupervised SNN training algorithm for classification problems,” in *Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, Hong Kong.
- Wang, F. Z., Li, L., Shi, L., Wu, H., and Chua, L. O. (2019). Φ memristor: real memristor found. *J. Appl. Phys.* 125:54504. doi: 10.1063/1.5042281
- Wang, W., Pedretti, G., Milo, V., Carboni, R., Calderoni, A., Ramaswamy, N., et al. (2018). Learning of spatiotemporal patterns in a spiking neural network with resistive switching synapses. *Sci. Adv.* 4:eaat4752. doi: 10.1126/sciadv.aat4752
- Wang, Z., Joshi, S., Savel'ev, S., Song, W., Midya, R., Li, Y., et al. (2018). Fully memristive neural networks for pattern classification with unsupervised learning. *Nat. Electron.* 1, 137–145. doi: 10.1038/s41928-018-0023-2
- Xu, Y., Zeng, X., Han, L., and Yang, J. (2013). A supervised multi-spike learning algorithm based on gradient descent for spiking neural networks. *Neural Netw.* 43, 99–113. doi: 10.1016/j.neunet.2013.02.003
- Yu, S., Wu, Y., Jeyasingh, R., Kuzum, D., and Wong, H. S. P. (2011). An electronic synapse device based on metal oxide resistive switching memory for neuromorphic computation. *IEEE Trans. Electron Devices* 58, 2729–2737. doi: 10.1109/TED.2011.2147791
- Ziegler, M., Soni, R., Patelczyk, T., Ignatov, M., Bartsch, T., Meuffels, P., et al. (2012). An electronic version of pavlov's dog. *Adv. Funct. Mater.* 22, 2744–2749. doi: 10.1002/adfm.201200244

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

Copyright © 2020 Lobov, Mikhaylov, Shamshin, Makarov and Kazantsev. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.