Compact Graphene-Based Spiking Neural Network with Unsupervised Learning Capabilities

Wang, He ; Cucu Laurenciu, NicoletA; Jiang, Yande; Cotofana, Sorin Dan

DOI
10.1109/OJNANO.2020.3041198

Publication date
2020

Document Version
Final published version

Published in
IEEE Open Journal of Nanotechnology

Citation (APA)
Wang, H., Cucu Laurenciu, N., Jiang, Y., & Cotofana, S. D. (2020). Compact Graphene-Based Spiking Neural Network with Unsupervised Learning Capabilities. *IEEE Open Journal of Nanotechnology*, *1*, 135-144. Article 9272847. https://doi.org/10.1109/OJNANO.2020.3041198

IEEE Open Journal of
Nanotechnology

Received 5 November 2020; accepted 24 November 2020. Date of publication 27 November 2020;
date of current version 17 December 2020. The review of this article was arranged by Editor F. Lombardi.

Digital Object Identifier 10.1109/OJNANO.2020.3041198

Compact Graphene-Based Spiking Neural Network With Unsupervised Learning Capabilities

HE WANG (Student Member, IEEE), NICOLETA CUCU LAURENCIU (Member, IEEE), YANDE JIANG (Student Member, IEEE), AND SORIN DAN COTOFANA (Fellow, IEEE)

Department of Quantum and Computer Engineering, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, 2628, CD Delft, The Netherlands

CORRESPONDING AUTHOR: HE WANG (e-mail: H.Wang-13@tudelft.nlStudent)

ABSTRACT To fully unleash the potential of graphene-based devices for neuromorphic computing, we propose a graphene synapse and a graphene neuron that form together a basic Spiking Neural Network (SNN) unit, which can potentially be utilized to implement complex SNNs. Specifically, the proposed synapse enables two fundamental synaptic functionalities, i.e., Spike-Timing-Dependent Plasticity (STDP) and Long-Term Plasticity, and both Long-Term Potentiation (LTP) and Long-Term Depression (LTD) can be emulated with the same structure by properly adjusting its bias. The proposed neuron captures the essential Leaky Integrate and Fire spiking neuron behavior with post firing refractory interval. We demonstrate the proper operation of the graphene SNN unit by relying on a mixed simulation approach that embeds the high accuracy of atomistic level simulation of graphene structures conductance within the SPICE framework. Subsequently, we analyze the way graphene synaptic plasticity affects the behavior of a 2-layer SNN example consisting of 6 neurons and demonstrate that LTP significantly increases the number of firing events while LTD is diminishing them, as expected. To assess the plausibility of the graphene SNN reaction to input stimuli we simulate its behavior by means of both SPICE and NEST, a well established SNN simulation framework, and demonstrate that the obtained reactions, characterized in terms of total number of firing events and mean Inter-Spike Interval (ISI) length, are in close agreement, which clearly suggests that the proposed design exhibits a proper behavior. Further, we prove the unsupervised learning capabilities of the proposed design by considering a 2-layer SNN consisting of 30 neurons meant to recognize the characters "A," "E," "I," "O," and "U," represented with a 5 by 5 black and white pixel matrix. The SPICE simulation results indicate that the graphene SNN is able to perform unsupervised character recognition associated learning and that its recognition ability is robust to input character variations. Finally, we note that our proposal results in a small real-estate footprint (max. 30 nm^2 are required by one graphene-based device) and operates at 200 mV supply voltage, which suggest its suitability for the design of large-scale energy-efficient computing systems.

INDEX TERMS Character recognition Character recognition, graphene, spiking neural network, spiking neural network, synaptic plasticity, synaptic plasticity, unsupervised learning, unsupervised learning.

I. INTRODUCTION

Human brain is a natural high performance computing system that exhibits excellent properties, e.g., ultra-low energy consumption, highly parallel information processing, suitability for complex tasks solving, and robustness. As such, to obtain artificial bio-inspired systems with brain akin computation abilities that can help understand the complex functionality of human brain, numerous attempts have been done to design and implement neuromorphic systems [1]–[3].

However, the fact that human brain comprises billions of neurons, which are the fundamental information processing units, and trillions of synapses that interconnect them makes the design and implementation of large-scale brain-inspired computing systems quite a challenging task. In most of

state-of-the-art neuromorphic systems, neurons and synapses are implemented with complex CMOS circuitry [4]–[6], which have high energy consumption and limited scalability and integration density. Recently, resistive switching memory devices [7] have been utilized for artificial neuron and synapse implementations [8]–[11] due to their simple structure, good scalability, and state preservation ability. However, they suffer from variability induced undesired stochastic behavior that may result in the instability of the neuromorphic system. Artificial neurons and synapses based on phase-change devices were also proposed [12]–[14] as they exhibit small footprint and the intrinsic device properties provide natural support for capturing spiking neuron and synapse dynamics. However, such designs require additional CMOS circuitry and external signals to enable their basic functionality and operate at relatively high voltage, which impede their applicability in large-scale energy-efficient neuromorphic systems.

Graphene has emerged as a promising material for nanoelectronics, as it exhibits outstanding properties, e.g., ballistic transport, flexibility, ultimate thinness, and biocompatibility [15], [16]. Due to these attractive properties, graphene-based Boolean logic gate [17], [18] and spiking neuron and synapse [19]–[21] implementations have been reported. However, previous work concentrated on individual synapse and neuron designs while disregarding input-output compatibility aspects, which preclude their direct utilization for the implementation of graphene-based Spiking Neural Networks.

In this paper, we propose a graphene-based synapse (comprising 2 graphene devices) and a spiking neuron (comprising 6 graphene devices), which form together a basic Spiking Neural Network (SNN) unit and can be utilized for the implementation of complex graphene-based SNNs. Specifically, the proposed artificial synapse emulates two basic synaptic functionalities, i.e., Spike-Timing-Dependent Plasticity (STDP) and Long-Term Plasticity, while the same synapse can exhibit Long-term Potentiation (LTP) or Long-term Depression (LTD) by properly adjusting the back-gate bias voltage of one of its composing graphene device. The proposed artificial neuron exhibits the essential Leaky Integrate and Fire (LIF) spiking neuron behavior with post firing refractory interval and provides the feedback signal required for the SDTP associated synaptic transmission efficiency modulation.

We first demonstrate the proper operation of the graphene SNN unit by relying on a mixed simulation approach that embeds the high accuracy of atomistic level simulation of graphene structures conductance within the SPICE framework. Subsequently, we analyze the way the synaptic plasticity affects the graphene SNN behavior by making use of a 2-layer SNN example consisting of 6 neurons and the obtained results indicate that LTP significantly increases the number of SNN firing events while LTD is diminishing them, as expected. To get some inside on the 2-layer graphene SNN reaction to input stimuli plausibility we also simulate its behavior by means of NEST [22], a well established SNN simulation framework. Our experiments indicate that the SPICE obtained reaction, characterized in terms of total number of firing events and mean Inter-Spike Interval (ISI) length, is in close agreement with the one reported by means of NEST based simulation, which clearly suggests that the proposed design exhibit a proper behavior. Further, we demonstrate the unsupervised learning capabilities of the proposed design by considering a two layer SNN consisting of 30 neurons meant to recognize the characters (and variations of them) "A," "E," "I," "O," and "U," represented with a 5 by 5 black and white pixel matrix. The simulation results indicate that the graphene SNN is able to perform unsupervised learning and that the enabled recognition ability is robust to input character variations. Finally, we note that our proposal results in a small real-estate footprint (max. 30 nm^2 are required by one graphene-based device) and operates at 200 mV supply voltage, which suggest its suitability for the design of large-scale energy-efficient computing systems.

The remaining of this paper is organized as follows: Section II presents the fundamental structure and functionality of a Spiking Neural Network, describes the basic graphene device, and give some inside over the utilized simulation framework. Section III introduces the graphene-based SNN design and its basic operation. Section IV presents the simulation results and Section V concludes the paper.

II. BACKGROUND

In this section we introduce the basic structure and functionality of a Spiking Neural Network (SNN), present the generic graphene device, which constitutes the fundamental building block for the construction of graphene SNNs, and conclude by providing some inside on the utilized graphene circuit SPICE simulation framework.

A. SPIKING NEURAL NETWORKS

Synapses and neurons are basic SNN components, which serve as junctions connecting different neurons and as basic information processing units, respectively. Fig. 1 depicts a small SNN comprising three neurons connected via two synapses. Neuron N_i collects signals (input spikes S_j and S_k) from neuron N_j and N_k, and it fires (generates the output spike S_i) when the cumulated input signals effect reaches the neuron firing threshold. A spiking neuron comprises three components: (i) a soma, which is the neuron cell body and supports the main neuronal functionalities, (ii) dendrites, which collect signals from other neurons and generate input to the soma, and (iii) the axon, which transmits neuron output spike to other neurons. There are various spiking neuron models to describe its functionality [23]–[25], among which the Integrate-and-Fire model is of particular interest, as it captures the essential behavior of a spiking neuron while having a low complexity [26]. The model for a standard nonlinear Leaky Integrate-and-Fire (LIF) neuron is as follows:

$$du/dt = F(u) + G(u) \cdot I, \tag{1}$$

where u represents the membrane potential, which is an intrinsic neuron parameter related to its membrane electrical charge, $F(u)$ is a voltage-dependent leak term, and $G(u)$

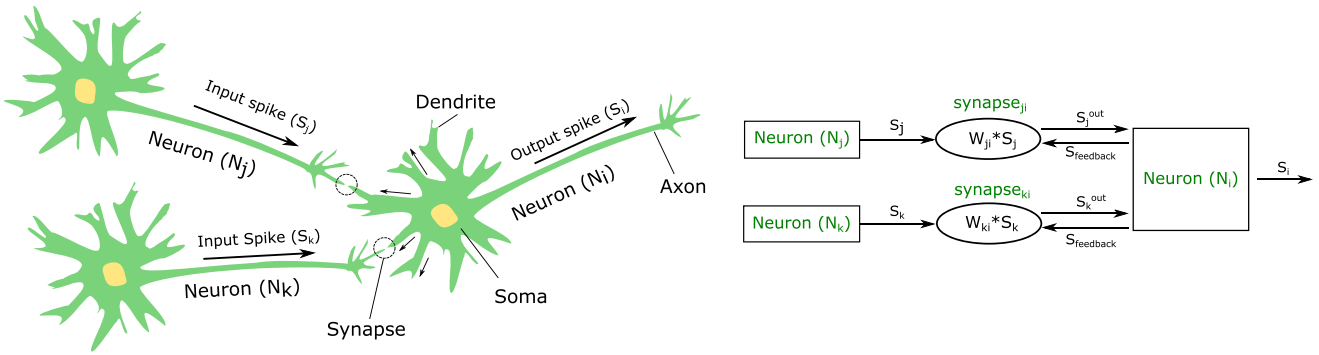


FIGURE 1. **Spiking neural network illustration.**

denotes the voltage-dependent input resistance, which contributes to the accumulation of the membrane potential due to the input current I. The dynamics of a nonlinear LIF neuron includes in chronological order: (i) an integration process, when the membrane potential increases due to input spikes arrival, (ii) an output firing event that generates an output spike when the membrane potential exceeds the neuron firing threshold, and (iii) a refractory period during which the neuron is not reacting to incoming spikes.

While synapses are essentially connecting neurons they are more than simple signal transmission lines, as their transmission efficiency (denoted as W) governed by the so-called synaptic plasticity process can either enhance or inhibit the transmitted signals. As such, a synapse, e.g., the one connecting neuron N_j and neuron N_i in Fig. 1, is actually processing two input signals (an input spike S_j from the pre-synaptic neuron N_j and a feedback signal from the post-synaptic neuron N_i) and produces one output signal S_j^{out}, which is transmitted to neuron N_i. It is believed that synaptic plasticity has an important impact on learning and memory of human brain [27], [28] and there are two basic plasticity types, i.e., Spike-Timing-Dependent Plasticity (STDP) and Long-Term Plasticity [26]. STDP follows the celebrated Hebbian learning principle [29] and adjusts the synaptic weight (transmission efficiency) according to the timing difference between the pre-synaptic spike (input spike from pre-synaptic neuron) and the post-synaptic spike (feedback signal from the post-synaptic neuron). Specifically, when the pre-synaptic spike arrives before the post-synaptic spike, the synapse transmission efficiency increases; otherwise the synapse transmission efficiency decreases. Long-term plasticity is a persistent synaptic weight change, which relies on the history of synaptic activities, and materialize in Long-Term Potentiation (LTP) and Long-Term Depression (LTD).

B. GENERIC GRAPHENE-BASED DEVICE

To enable the aforementioned synapse and neuron functionalities, we rely on instances of the generic graphene-based device depicted in Fig. 2, which comprises a monolayer Graphene Nanoribbon (GNR) placed on an insulating layer and a doped substrate that serves as a back-gate. The GNR sheet sustains

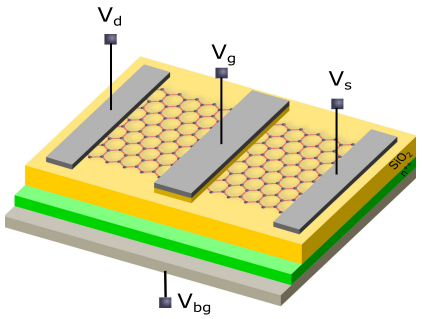


FIGURE 2. **Generic graphene-based device.**

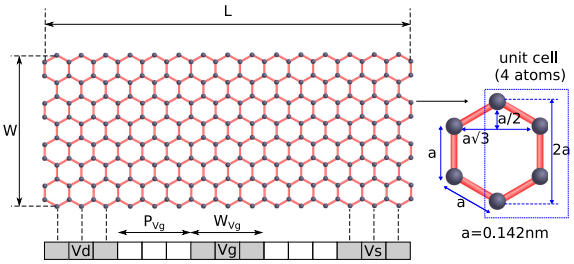


FIGURE 3. **GNR geometry and contact topology.**

a conduction channel under a drain-to-source bias voltage $V_d - V_s$, which conductance profile is determined by the GNR sheet geometry and contacts topology, while the actual channel conduction value is modulated by means of external voltages applied on the top/back gates [30]. Fig. 3 illustrates the parameters related to GNR geometry and contacts topology. Specifically, W and L denote the width and length of the graphene sheet, respectively, P_{V_g} the distance between the top-gate and the drain contact, and W_{V_g} the top-gate width. The distance between two neighbor carbon atoms is denoted as $a = 0.142$ nm. We note that the capability of such graphene-based device to provide a rich set of complex functionalities has been demonstrated by the utilization of specially tailored (in terms of topology and dimensions) versions of it for the implementation of Boolean gates and individual synapses and neurons [17], [19], [21].

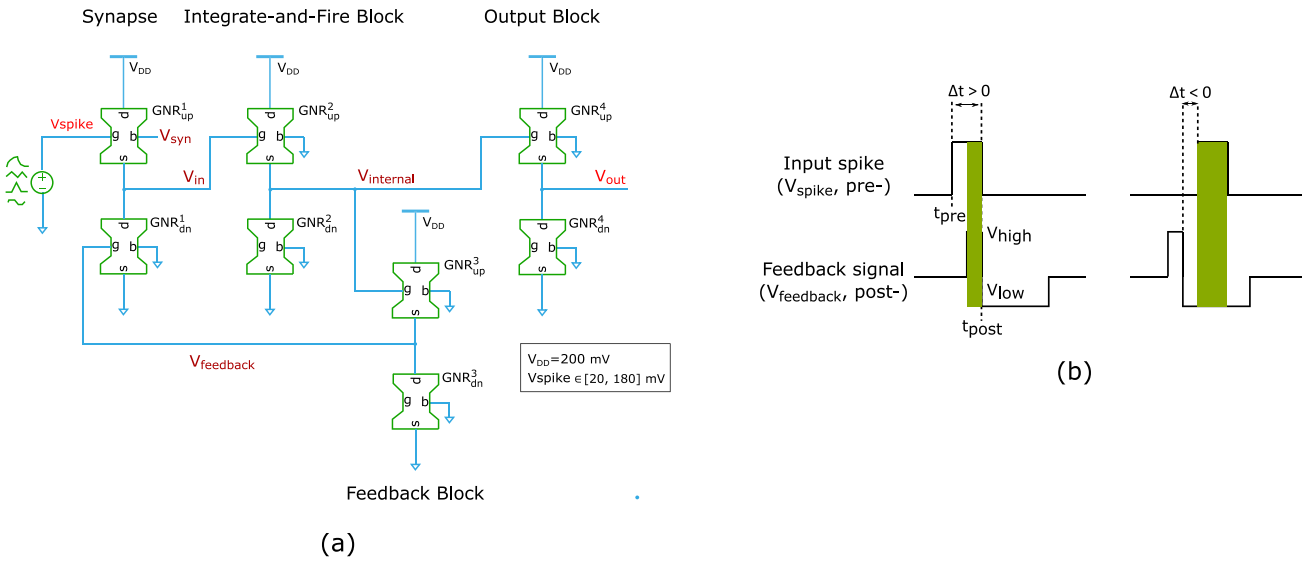


FIGURE 4. **Graphene-based spiking neural network: (a) SNN circuit, (b) Pre- vs post-synaptic spikes timing.**

Apart of the conduction channel creation it has been experimentally observed that graphene-based devices inherently exhibit interface charges trapping/detrapping phenomena [31], which are usually induced by top-gate oxide defects. Depending on the applied top-gate voltage V_g value charges are trapped or released by the graphene oxide interface, which causes an equivalent top-gate voltage shift and affects the top-gate conductance modulation ability. This phenomenon makes the GNR conductance dependent on the cumulated device history activities and as such makes GNR devices suitable for the emulation of synaptic plasticity [21] and neuron membrane potential dynamics [19].

C. SIMULATION FRAMEWORK

In order to properly validate and evaluate the graphene-based SNN circuits, we rely on a mixed simulation approach incorporating atomistic level graphene-based device modelling and SPICE simulation in Cadence [32].

For the graphene-based device electronic transport properties calculation, we utilize the atomistic level Tight-Binding Hamiltonian to model the carbon atom interactions and external potentials, the Non-Equilibrium Green Function (NEGF) to solve the Schrödinger equation, and the Landauer-Büttiker formula to calculate the GNR channel current and conductance [33]. The potential distribution on graphene sheet is obtained by solving a 3D Poisson equation self-consistently, and the effect of trapping/detrapping phenomenon on the device operation is accounted for by calculating the equivalent voltage shift caused by interface trapped charges [34].

To enable high accuracy circuit simulation, we make use of a Verilog-A graphene device generic model [35], which in order to enable time effective SPICE simulation of graphene circuit relies on GNR topology specific precomputed look-up tables containing graphene conduction simulation data obtained with the aforementioned atomistic level simulation methodology.

III. GRAPHENE-BASED SPIKING NEURAL NETWORKS

In this section we present the proposed graphene-based Spiking Neural Network design and describe its basic operation principle.

The schematic illustration of the graphene-based SNN unit (consisting of one synapse and one neuron) is depicted in Fig. 4(a) and comprises four blocks: (i) synapse, (ii) integrate-and-fire, (iii) feedback, and (iv) output. Each block consists of two GNR-based devices and its output voltage (V_{in}, V_{internal}, V_{feedback}, and V_{out}) is governed by the $V_{\text{DD}} \cdot G^i_{\text{up}}/(G^i_{\text{up}} + G^i_{\text{dn}})$ relation, where V_{DD} is the supply voltage (200 mV), and G^i_{up} and G^i_{dn} denote the conductance of the i^{th} GNR_{up} and GNR_{dn}, respectively.

The synapse receives input spikes V_{spike} from another neuron, potentiates or suppresses them according to its transmission efficiency (weight), and generates V_{in} to be utilized as neuron block input. The initial synaptic weight value is determined by the V_{syn} potential applied on the back-gate of GNR^1_{up}. The synapse exhibits two types of plasticity: Spike-Timing-Dependent Plasticity (STDP) and Long-Term Plasticity. STDP modulates the synaptic weight by accounting for the time difference $\Delta t = t_{post} - t_{pre}$ between the input spike V_{spike} occurrence and V_{feedback} transition from V_{high} to V_{low}, as depicted in Fig. 4(b). When $\Delta t > 0$, i.e., V_{feedback} is asserted before the end of the input spike, the synaptic transmission efficiency is increased and V_{spike} contribution to V_{in} is strengthened. When $\Delta t < 0$, i.e., the input spike occurrence is not generating a V_{feedback} transition to V_{high}, the synaptic transmission efficiency is decreased and V_{spike} contribution to V_{in} is weakened. The input spike potentiation/depression is controlled by the V_{feedback} signal, which by being connected

TABLE 1. GNR Topologies

GNR components	GNR^1_{up}	GNR^1_{dn}	GNR^2_{up}	GNR^2_{dn}	GNR^3_{up}	GNR^3_{dn}	GNR^4_{up}	GNR^4_{dn}
W [a]	23	29	29	29	23	29	29	29
L [a]	$25\sqrt{3}$	$25\sqrt{3}$	$30\sqrt{3}$	$30\sqrt{3}$	$30\sqrt{3}$	$25\sqrt{3}$	$25\sqrt{3}$	$30\sqrt{3}$
PV_g [a]	$6\sqrt{3}$	$6\sqrt{3}$	$6\sqrt{3}$	$6\sqrt{3}$	$6\sqrt{3}$	$4\sqrt{3}$	$6\sqrt{3}$	$6\sqrt{3}$
WV_g [a]	$8\sqrt{3}$	$6\sqrt{3}$	$8\sqrt{3}$	$6\sqrt{3}$	$6\sqrt{3}$	$8\sqrt{3}$	$6\sqrt{3}$	$4\sqrt{3}$

to GNR^1_{dn} top-gate modulates its conductance. The long-term plasticity emulation relies on the fact that when applying input spikes on GNR^1_{up} top-gate trapped charges are accumulated and as such modulate its conductance persistently, which depending on the V_{syn} value results in Long-Term Potentiation (LTP) or Long-Term Depression (LTD) of the synaptic weight, e.g., 0 mV for LTP and −100 mV for LTD.

The integrate-and-fire block is the kernel of the graphene spiking neuron and emulates the main neuronal functionalities, including membrane potential integration and the generation of the output firing events. The integrate and fire behavior builds upon the interface trapping phenomenon, which results in charge accumulation when V_{in} spikes are applied on GNR^2_{up} top-gate. The trapped charges cause an equivalent shift ΔV_g of the top-gate voltage V_g and when $V_g + \Delta V_g$ reaches a certain level, i.e., the neuron firing threshold, GNR^2_{up} conductance increases abruptly, which triggers a firing event, i.e., generates a spike on the $V_{internal}$ signal.

While this is enough to emulate spiking neuron functionality $V_{internal}$ requires some extra processing in order to be compatible in terms of voltage levels and duration with the input spike applied on V_{spike}, which assumes values between 20 mV and 180 mV and has a time duration of 2 ms. As such the output block further processes $V_{internal}$ and produces V_{out} that is level and duration compatible with synapse input spikes, which enables the direct cascading of SNN basic units. Moreover, as the neuron output is playing a crucial role in the STDP process the $V_{internal}$ spike occurrence has to be signalled to the synapse block. Again $V_{internal}$ cannot be directly utilized and the feedback block is responsible for the generation of $V_{feedback}$ that is connected to GNR^1_{dn} top-gate to internally signal the firing event occurrence. Apart of contributing to the synaptic weight adaptation $V_{feedback}$ is also placing the neuron into the refractory state, which has to occur after any output firing event. This is enabled by the $V_{feedback}$ transition from V_{high} to V_{low}, which is increasing GNR^1_{dn} conductance resulting in a significant V_{in} magnitude reduction that inhibits the trap accumulation and as such incoming input spikes cannot trigger a firing event while $V_{feedback} = V_{low}$.

The basic SNN unit behavior is actually dependent on the conductance variation exhibited by each of the GNRs it comprises. Thus to guaranty proper SNN functionality 4 GNR geometry pairs, which conductance maps fit the variation profile required to achieve the desired behavior of the V_{in}, $V_{internal}$, $V_{feedback}$, and V_{out} signals, respectively, should be find. Fig. 5 depicts the GNR topologies we identified for the proposed SNN circuit, by means of an atomistic model based Design Space Exploration (DSE) process and Table 1 summarizes their dimensions expressed in terms of the distance between

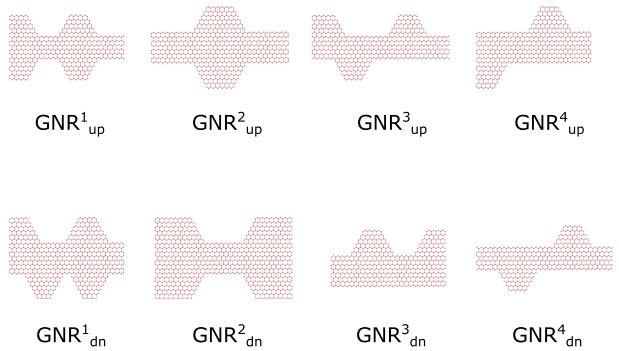


FIGURE 5. Basic SNN unit GNR shapes.

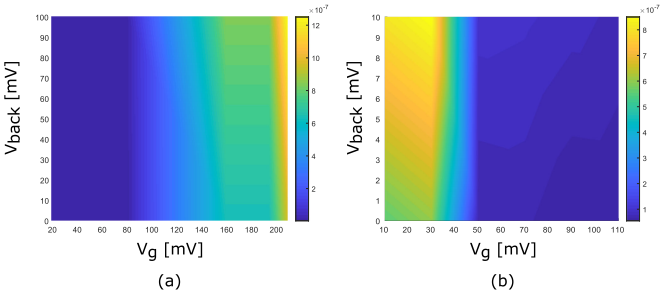


FIGURE 6. GNR conduction maps: (a) GNR^1_{up}, (b) GNR^1_{dn}.

adjacent carbon atoms in the graphene sheet $a = 0.142$ nm. Concerning the interface trap profile for the atomistic-level graphene-based device modelling in DSE, we assume an interface trap density of $2.363 \cdot 10^{13}$ $cm^{-2}(eV)^{-1}$ and a trapping/detrapping time constant of 1.6 ms [36], [37].

To provide inside on the relation between the chosen GNR topologies and SNN circuit behavior we present in Fig. 6 the conduction maps of GNR^1_{up} and GNR^1_{dn} that form the synapse block. As one can observe in Fig. 6 GNR^1_{up} conduction is high under large top-gate voltages and varies with back-gate voltage value thus can provide different initial synaptic weights. GNR^1_{dn} conductance is high under low top-gate voltages and small under high top-gate voltages, which allows $V_{feedback}$ to induce synaptic transmission potentiation and depression when being V_{high} and V_{low}, respectively. A similar analysis can be carried on for the other GNR pairs in the circuit but we omit it in view of page limit.

SPICE simulation results concerning the SNN unit basic operation (with $V_{syn} = 0$ mV) are illustrated in Fig. 7. As seen from the point of view of $V_{feedback}$ value the basic operation follows three phases. In Phase I $V_{feedback}$ has an initial after circuit reset value and the neuron input V_{in} follows

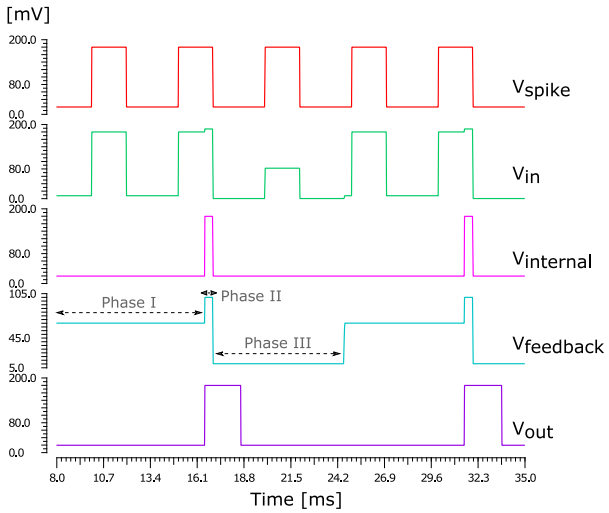


FIGURE 7. Graphene-based SNN unit basic operation.

the synapse input V_{spike}. When $V_g + \Delta V_g$ of GNR^2_{up} reaches the firing threshold a spike is induced on V_{internal} that makes V_{feedback} to enter into Phase II when V_{in} magnitude increases to $V_{\text{spike}} \times 1.1$ for a short time period as result of the pre-spike before post-spike $\Delta t > 0$ induced STDP potentiation. Immediately after the firing event V_{internal} returns to its initial value and as the V_{internal} induced trapped charges are still present V_{feedback} becomes V_{low} and the SNN unit enters Phase III. In this period V_{in} magnitude decreases to $V_{spike}/2.2$ as a result of pre-spike after post-spike $\Delta t < 0$ induced STDP depression. As no firing events can be triggered during Phase III, it accounts for the spiking neuron refractory interval. When the feedback block trapped charges decay to the initial level, V_{feedback} returns to its after reset value, Phase III finishes and the circuit switches back to Phase I. Related to the refractory interval influence on the neuron behavior one can observe in Fig. 7 that the first output spike is triggered by 2 input spikes while the second one occurs after 3 input spikes.

The basic SNN unit in Fig. 4(a) assumes that the neuron process input spikes coming from one previous neuron only, i.e., has a fan-in of 1, which is certainly not the case in any relevant SNN. To accommodate for a fan-in of n we extend the synapse block by replacing GNR^1_{up} with n GNRs as illustrated in Fig. 8. In this case the in-between voltage V_{in} is calculated as:

$$V_{in} = V_{DD} \cdot \frac{G^{11}_{up} + G^{12}_{up} + \cdots + G^{1n}_{up}}{G^{11}_{up} + G^{12}_{up} + \cdots + G^{1n}_{up} + G^{1}_{dn}}, \qquad (2)$$

where $\text{G}^{1n}_{\text{up}}$ denotes the conductance of the n^{th} up GNR.

IV. SIMULATION RESULTS

To get inside into the actual capabilities of the proposed SNN unit we consider and evaluate by means of SPICE simulation two graphene-based SNN examples. We first study the effect of the graphene enabled synaptic plasticity on a 2-layer

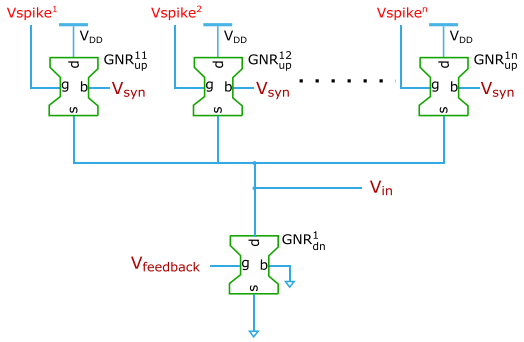


FIGURE 8. Multi-input synapse block.

6-neuron SNN and compare its SPICE derived behavior with the one obtained by means of NEST based simulations [22]. Subsequently, we demonstrate the capability of our proposal to perform unsupervised character recognition. In all simulations, the input spikes are 2 ms long pulses varying between 20 mV and 180 mV, and $V_{\text{DD}} = 200\,\text{mV}$. We note however that our proposal is general and can be adapted to operate on different power supply values and input spike formats.

A. GRAPHENE-BASED SNN BEHAVIOR EVALUATION

To evaluate how the long-term plasticity exhibited by the graphene devices modulates the neuron input signal V_{in}, we consider a single synapse block comprising GNR^1_{up} and GNR^1_{dn}, as depicted in Fig. 4(a), and set the feedback signal V_{feedback} to the Phase I value. In such a setup the synapse output magnitude follows the synapse input and if V_{spike} receives a train of spikes Long-Term Plasticity should be observed. To capture this phenomenon we apply a 200 Hz periodic input spike train with 180 mV peak amplitude on the GNR^1_{up} top-gate and simulate the circuit evolution for 300 ms. The obtained dynamics of the synapse output signal V_{in} is depicted in Fig. 9(a) and (b) for Long-Term Potentiation (LTP) and Long-Term Depression (LTD), respectively. Note that both LTP and LTD are acquired with the same synapse by properly changing the back-gate voltage of GNR^1_{up}, i.e., 0 mV for LTP and -100 mV for LTD. As expected, we observe a continuous V_{in} magnitude increase and decrease for LTP and LTD, respectively. After 300 ms the amplitude potentiation and depression are around 3.3% and 4.5%, for LTP and LTD. respectively, and exhibit an obvious saturation trend.

To explore the implication of the obtained long-term plasticity on SNN's firing events profile, we make use of a 2-layer SNN consisting of 6 neurons as illustrated in Fig. 10. The neuron in layer 2 is fully connected with all the neurons in layer 1 via identical synapses. In the simulations we considered three synapse types: (i) without long-term plasticity (assuming that the trapped charges do not affect the graphene device conductance), (ii) with long-term potentiation, and (iii) with long-term depression.

To evaluate the SNN behavior in the previously mentioned conditions we perform SPICE simulations assuming that all

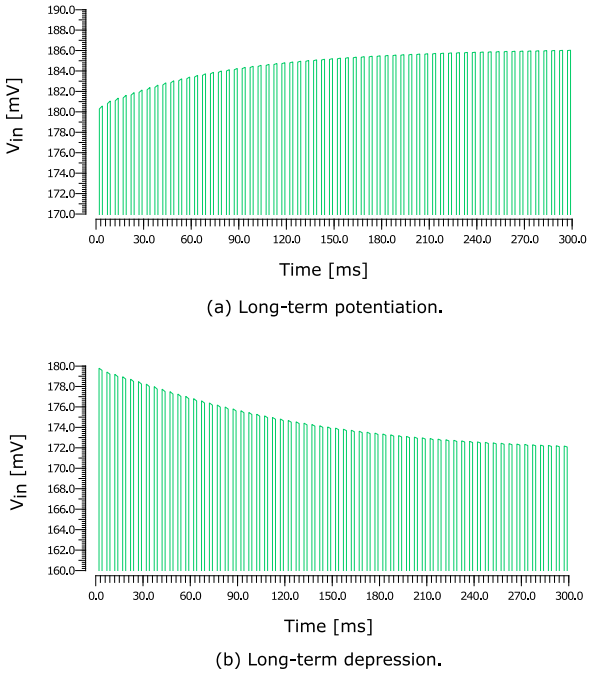


FIGURE 9. Synapse long-term plasticity.

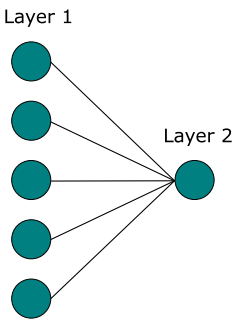


FIGURE 10. Two layers 6-neuron SNN.

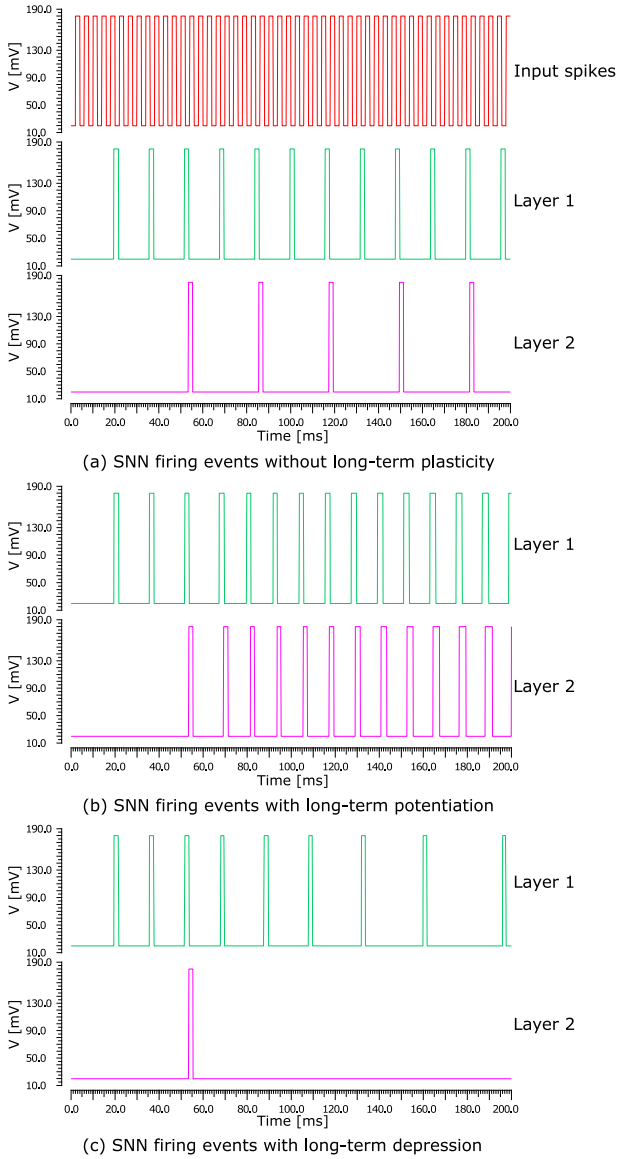
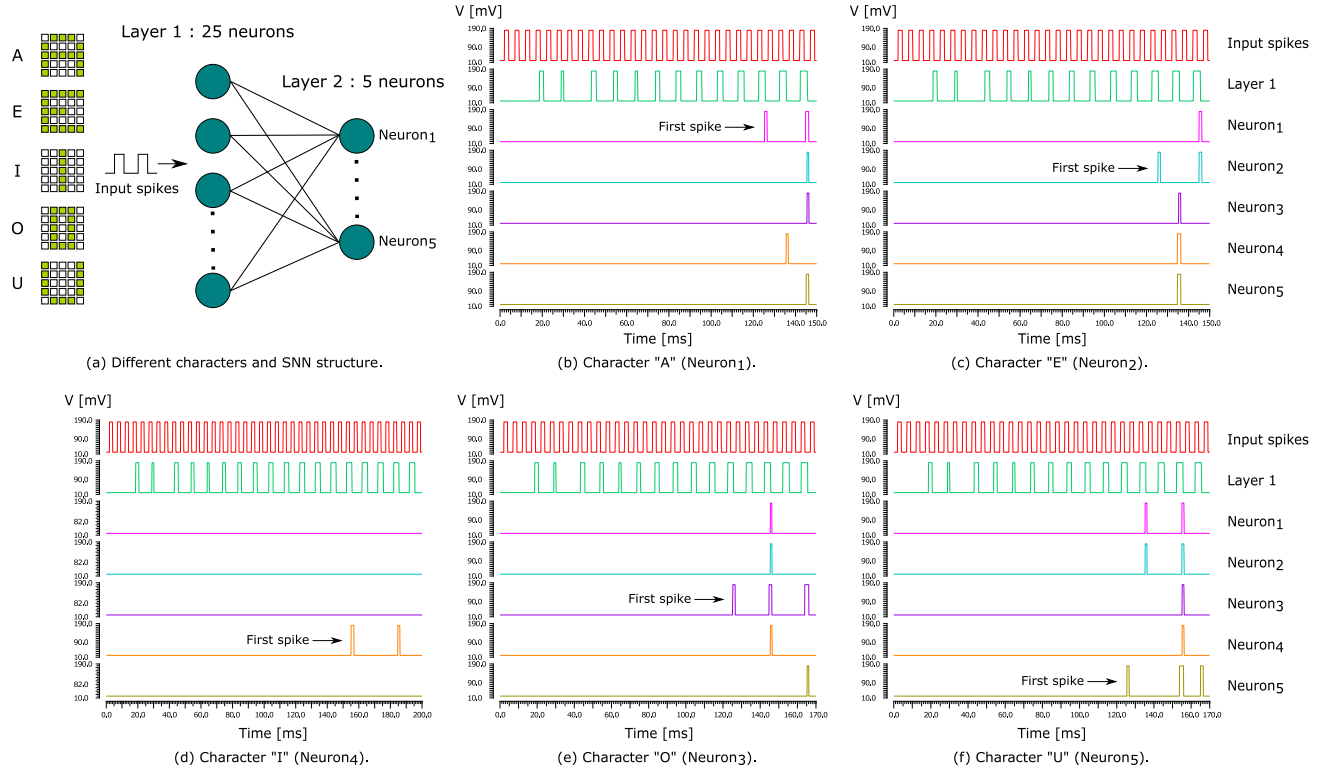


FIGURE 11. Two layers 6-neuron SNN output firing events.

layer 1 neurons receive identical 200 Hz periodic input signals (V_{spike}) on their synapse block for 200 ms, thus all layer 1 neurons generate identical firing events. Fig. 11(a) depicts the SNN output reaction with the synapses do not exhibit long-term plasticity. We observe that periodic output spike trains are generated by all neurons while layer 2 neuron firing rate of the neuron is lower than that of the layer 1 neurons. During the simulation there are in total 12 output spikes for every neuron in layer 1 and 5 output spikes for neuron in layer 2. Fig. 11(b) depicts the SNN output firing events with Long-Term Potentiation. As expected LTP induces an increase of the number of firing events in both layers, which now raise to 15 and 13 for neurons in layer 1 and layer 2, respectively. Thus LTP induces a 25% firing event increase in layer 1 and 160% in layer 2. On the contrary, in the case of SNN with Long-Term Depression, the simulation result is depicted in Fig. 11(c), we observe a significant decrease tendency of the number of firing events in both layers. Specifically, the layer 2 neuron stops generating any fire event after 60 ms, which is related to the fact that due to LTD layer 1 neurons are less active and as such cannot trigger a firing event of the neuron in layer 2. The total number of firing events for neurons in layer 1 and layer 2 are 9 and 1, which is equivalent with a 25% and 80% decrease, respectively.

To get some inside of the plausibility of the LTP and LTD influence on the considered SNN example we implement it in NEST with standard leaky Integrate-and-Fire neurons connected via synapses with Long-Term Potentiation, apply 200 Hz and 250 Hz periodic input spike trains, and record its reaction a time period of 200 ms. The number of layer 2 neuron firing events as well as the mean Inter-Spike Interval (ISI) between output spikes obtained by the SPICE simulation of the graphene-based SNN with LTD and the ones reported by means of NEST simulation are summarized in Table 2.



(a) Different characters and SNN structure. (b) Character "A" (Neuron$_1$). (c) Character "E" (Neuron$_2$).

(d) Character "I" (Neuron$_4$). (e) Character "O" (Neuron$_3$). (f) Character "U" (Neuron$_5$).

FIGURE 12. "A," "E," "I," "O," and "U" recognition associated unsupervised learning.

TABLE 2. SNN Activity Overview

	Input firing rate: 200 Hz		Input firing rate: 250 Hz	
	Total Number	Mean ISI	Total Number	Mean ISI
Graphene-based SNN	16	30	22	22.6
NEST simulator	17	28.5	22	22.3

In terms of the total number of firing events, the graphene SNN produces an almost identical response with the NEST based simulation one, i.e., 1 spike difference at 200 Hz input and the same number at 250 Hz input. As for the mean ISI, which represents the average time interval between adjacent output spikes, the values are quite close with a maximum difference of 5% (1.5 ms) between the SPICE and NEST predicted results. The obtained results clearly suggest that the proposed graphene SNN exhibits similar behavior with the one predicted by the well established NEST simulation framework.

B. UNSUPERVISED CHARACTER RECOGNITION

To demonstrate the learning abilities of our proposal, we consider a 2-layer SNN consisting of 30 neurons as depicted in Fig. 12(a), which is meant to recognize the characters (and variations of them) "A," "E," "I," "O," and "U," represented with a 5 by 5 black and white pixel matrix. Layer 1 comprises 25 neurons, which receive input spikes if the pixel in their position is black and no spikes if the pixel is white, and layer 2 consists of 5 neurons meant to indicate the recognition result. We assume that: (i) LTP synapses with identical initial synaptic weight are utilized for every neuron in layer 1 and (ii) Every neuron in layer 2 is connected with all the layer 1 neurons via LTP synapses with randomly initialized synaptic weights. This is achieved by biasing V_{syn} (the back-gate voltage of GNR^1_{up}) with fixed values between 0 mV and 100 mV, such that layer 2 neurons exhibit different firing profile. For a given input character, we stimulate the layer 1 neurons corresponding to black pixels with identical 200 Hz periodic spike trains as illustrated in Fig. 12(a). Each layer 2 neuron is meant to signal the recognition of one character in the vowel set and to indicate that we employ the "time-to-first-spike" scheme [26], i.e., the layer 2 neuron that first fires is the one that recognized the input character.

To validate the learning ability of the proposed design, we apply the 5 characters "A," "E," "I," "O," and "U," to the graphene-based SNN one at a time and the learning process for each of them is depicted in Fig. 12(b), (c), (d), (e), and (f), respectively.

In each case, we observe that initially there are no firing events on any layer 2 neurons. However, during the learning process, the connections corresponding to the layer 1 stimulated neurons (the one driven by black pixels) are strengthened because of long-term potentiation. Thus, after some time one neuron in layer 2 fires (indicating the recognition result) and eventually other neurons in layer 2 may fire afterwards. Fig. 12(b), (c), (d), (e), and (f), clearly indicate that characters "A," "E," "I," "O," and "U" are recognized by Neuron$_1$, Neuron$_2$, Neuron$_4$, Neuron$_3$, and Neuron$_5$, respectively. The

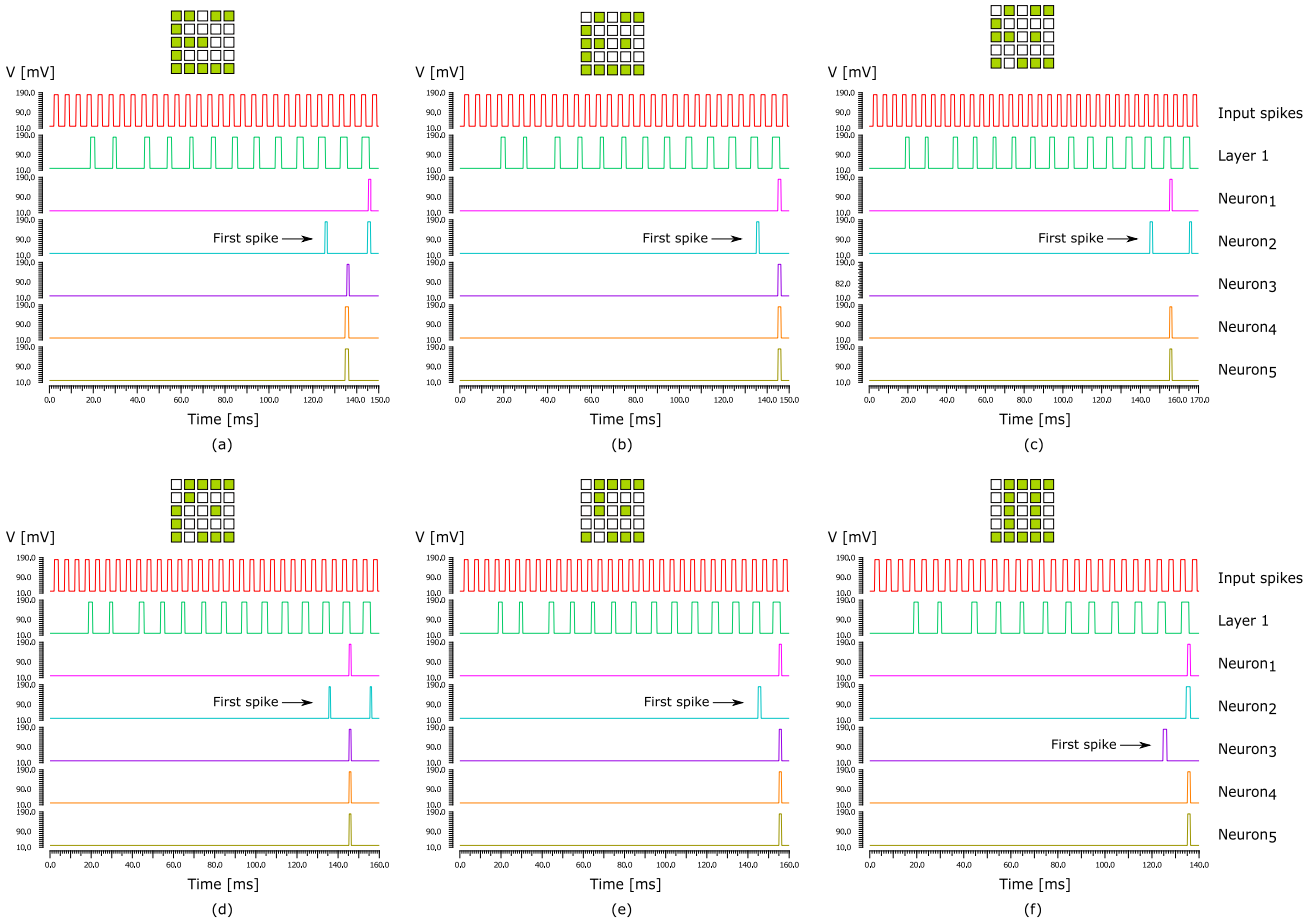


FIGURE 13. **Character "E" recognition.**

learning time for characters "A," "E," "O," and "U" is around 125 ms while for "I" is around 165 ms as it stimulates less layer 1 neurons than the other characters. As the result of this unsupervised learning process each layer 2 neuron is labeled with the character which presence in the input it recognizes and based on this labelling one can tell if a new unknown character is one of the 5.

To test the recognition ability of the graphene SNN we make use of different variations of the original characters as inputs. As an example, we present the recognition processes for six character variations that gradually degrade from "E" to "O," as illustrated in Fig. 13. When applying inputs that maintain the "E" character profile as depicted in Fig. 13(a), (b), (c), (d), and (e), one can observe that $Neuron_2$ first fires, which indicates that the graphene SNN correctly recognizes those inputs as character "E". The time needed for the SNN to recognize the inputs in each case are 125 ms, 135 ms, 155 ms, 135 ms, and 145 ms, respectively, which is in line with the observation that when an input character stimulates less input neurons in layer 1, the SNN recognition takes more time. When applying an input that fundamentally deviate from "E" as depicted in Fig. 13(f), $Neuron_3$ first fires after around 125 ms, which indicates that the SNN recognizes the input character as an "O" and not as an "E". The fact that the degraded character is closer to an "O" than to an "E" is also obvious by visual inspection and as such the SNN made the correct decision. The aforementioned results demonstrate the applicability of the proposed graphene SNN for provide support for unsupervised character recognition, and that the learning ability is robust.

V. CONCLUSION

In this paper we proposed a basic graphene-based Spiking Neural Network (SNN) unit consisting of a synapse and a spiking neuron that can be utilized to implement complex SNNs. The proposed design enables Spike-Timing-Dependent Plasticity (STDP) and Long-Term Plasticity, and both Long-Term Potentiation (LTP) and Long-Term Depression (LTD) can be induced in the same synapse by properly bias adjustments. By means of SPICE simulation, we validated the basic operation of the proposed design and analyzed how the enabled synaptic plasticity affects the SNN behavior. To this end we assumed a 2-layer SNN, derived its reaction to the same input stimuli by means of SPICE and NEST simulations, and demonstrated the close agreement between the obtained results in terms of total number of firing events and mean

Inter-Spike Interval (ISI) length. Further, we demonstrated the unsupervised learning capabilities of the proposed design by considering a two layer SNN consisting of 30 neurons meant to recognize the characters (and variations of them) "A," "E," "I," "O," and "U," represented with a 5 by 5 black and white pixel matrix. The simulation results indicated that the graphene SNN is able to perform unsupervised character recognition and that its recognition ability is robust to input character variations.

REFERENCES

[1] C. Mead, "Neuromorphic electronic systems," *Proc. IEEE*, vol. 78, no. 10, pp. 1629–1636, 1990.

[2] G. Indiveri *et al.*, "Neuromorphic silicon neuron circuits," *Front. Neurosci.*, vol. 5, p. 73, 2011.

[3] Y. van De Burgt, A. Melianas, S. T. Keene, G. Malliaras, and A. Salleo, "Organic electronics for neuromorphic computing," *Nat. Electron.*, vol. 1, no. 7, pp. 386–397, 2018.

[4] S. B. Furber *et al.*, "Overview of the spinnaker system architecture," *IEEE Trans. Comput.*, vol. 62, no. 12, pp. 2454–2467, Dec. 2013.

[5] P. A. Merolla *et al.*, "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, no. 6197, pp. 668–673, 2014.

[6] J. H. Wijekoon and P. Dudek, "A CMOS circuit implementation of a spiking neuron with bursting and adaptation on a biological timescale," in *Proc. IEEE Biomed. Circuits Syst. Conf.*, 2009, pp. 193–196.

[7] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," *Nature*, vol. 453, no. 7191, pp. 80–83, 2008.

[8] M. Al-Shedivat, R. Naous, E. Neftci, G. Cauwenberghs, and K. N. Salama, "Inherently stochastic spiking neurons for probabilistic neural computation," in *Proc. 7th Int. IEEE/EMBS Conf. Neural Eng.*, 2015, pp. 356–359.

[9] Y. Babacan, F. Kaçar, and K. Gürkan, "A spiking and bursting neuron circuit based on memristor," *Neurocomputing*, vol. 203, pp. 86–91, 2016.

[10] B. Linares-Barranco and T. Serrano-Gotarredona, "Memristance can explain spike-time-dependent-plasticity in neural synapses," *Nat. Proc.*, 2009. [Online]. Available: https://doi.org/10.1038/npre.2009.3010.1

[11] S. Saïghi *et al.*, "Plasticity in memristive devices for spiking neural networks," *Front. Neurosci.*, vol. 9, p. 51, 2015.

[12] T. Tuma, A. Pantazi, M. Le Gallo, A. Sebastian, and E. Eleftheriou, "Stochastic phase-change neurons," *Nat. Nanotechnol.*, vol. 11, no. 8, p. 693, 2016.

[13] C. D. Wright, P. Hosseini, and J. A. V. Diosdado, "Beyond von-neumann computing with nanoscale phase-change memory devices," *Adv. Funct. Mater.*, vol. 23, no. 18, pp. 2248–2254, 2013.

[14] S. Ambrogio *et al.*, "Unsupervised learning by spike timing dependent plasticity in phase change memory (PCM) synapses," *Front. Neurosci.*, vol. 10, p. 56, 2016.

[15] P. Avouris and C. Dimitrakopoulos, "Graphene: Synthesis and applications," *Mater. Today*, vol. 15, no. 3, pp. 86–97, 2012.

[16] M. J. Allen, V. C. Tung, and R. B. Kaner, "Honeycomb carbon: A review of graphene," *Chem. Rev.*, vol. 110, no. 1, pp. 132–145, 2010.

[17] Y. Jiang, N. C. Laurenciu, H. Wang, and S. D. Cotofana, "Graphene nanoribbon based complementary logic gates and circuits," *IEEE Trans. Nanotechnol.*, vol. 18, pp. 287–298, 2019.

[18] G. Liu, S. Ahsan, A. G. Khitun, R. K. Lake, and A. A. Balandin, "Graphene-based non-boolean logic circuits," *J. Appl. Phys.*, vol. 114, no. 15, 2013, Art. no. 154310.

[19] H. Wang, N. C. Laurenciu, Y. Jiang, and S. D. Cotofana, "Ultra-compact, entirely graphene-based nonlinear leaky integrate-and-fire spiking neuron," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2020, pp. 1–5.

[20] H. Tian *et al.*, "Graphene dynamic synapse with modulatable plasticity," *Nano Lett.*, vol. 15, no. 12, pp. 8013–8019, 2015.

[21] H. Wang, N. C. Laurenciu, Y. Jiang, and S. D. Cotofana, "Graphene nanoribbon-based synapses with versatile plasticity," in *Proc. IEEE/ACM Int. Symp. Nanoscale Architectures (NANOARCH)*, 2019, pp. 1–6.

[22] T. Fardet *et al.*, "Nest 2.20.0," Jan. 2020. [Online]. Available: https://doi.org/10.5281/zenodo.3605514

[23] W. Gerstner and W. M. Kistler, *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge, U.K.: Cambridge Univ. Press, 2002.

[24] E. M. Izhikevich, "Which model to use for cortical spiking neurons?," *IEEE Trans. Neural Netw.*, vol. 15, no. 5, pp. 1063–1070, Sep. 2004.

[25] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Trans. Neural Netw.*, vol. 14, no. 6, pp. 1569–1572, Nov. 2003.

[26] W. Gerstner, W. M. Kistler, R. Naud, and L. Paninski, *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge, U.K.: Cambridge Univ. Press, 2014.

[27] L. F. Abbott and S. B. Nelson, "Synaptic plasticity: Taming the beast," *Nat. Neurosci.*, vol. 3, no. 11, pp. 1178–1183, 2000.

[28] S. J. Martin, P. D. Grimwood, and R. G. Morris, "Synaptic plasticity and memory: An evaluation of the hypothesis," *Annu. Rev. Neurosci.*, vol. 23, no. 1, pp. 649–711, 2000.

[29] N. Caporale and Y. Dan, "Spike timing–dependent plasticity: A Hebbian learning rule," *Annu. Rev. Neurosci.*, vol. 31, no. 1, pp. 25–46, 2008.

[30] S. Cotofana, P. Dimitrakis, M. Enachescu, I. Karafyllidis, A. Rubio, and G. C. Sirakoulis, "On graphene nanoribbon-based nanoelectronic circuits viability," in *42nd Workshop Compound Semicond. Devices Integr. Circuits Held Eur.*, pp. 35–36, 2018.

[31] A. C. Ferrari *et al.*, "Science and technology roadmap for graphene, related two-dimensional crystals, and hybrid systems," *Nanoscale*, vol. 7, no. 11, pp. 4598–4810, 2015.

[32] Cadence. [Online]. Available: https://www.cadence.com/

[33] S. Datta, *Quantum Transport: Atom to Transistor*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[34] H. Wang, N. C. Laurenciu, Y. Jiang, and S. D. Cotofana, "Atomistic-level hysteresis-aware graphene structures electron transport model," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2019, pp. 1–5.

[35] Y. Jiang, N. C. Laurenciu, and S. Cotofana, "Non-equilibrium green function-based verilog-a graphene nanoribbon model," in *Proc. 18th IEEE Int. Conf. Nanotechnol.*, 2018, pp. 1–4.

[36] U. Jung, Y. J. Kim, Y. Kim, Y. G. Lee, and B. H. Lee, "Extraction of the interface state density of top-gate graphene field-effect transistors," *IEEE Electron Device Lett.*, vol. 36, no. 4, pp. 408–410, Apr. 2015.

[37] Y. G. Lee *et al.*, "Fast transient charging at the Graphene/sio 2 interface causing hysteretic device characteristics," *Appl. Phys. Lett.*, vol. 98, no. 18, p. 183508, 2011.