A Geometrical Optimization Rule of the Synaptic Pass-Transistor for a Low Power Analog Accelerator

DANYOUNG CHA, **YEONSU KANG**, **SOOBIN LEE**, **AND SUNGSIK LEE**

Department of Electronics Engineering, Pusan National University, Pusan 46241, South Korea

Corresponding author: Sungsik Lee (sungsiklee@pusan.ac.kr)

ABSTRACT We present a simulation framework on a geometrical optimization rule of the synaptic pass-transistor (SPT) for a low power analog accelerator (AA). Here, the SPT is a synaptic transistor (Syn-Tr) in series with a load resistor to be scaled with respect to a geometrical ratio between a channel length and width of the Syn-Tr. When only the load resistance is increased for the reduction of the power consumption of the SPT, a synaptic characteristics (e.g. a synaptic dynamic ratio, DR_w) is hard to be maintained. To overcome this, the channel geometrical ratio and scaling factor of the load resistance are required to be increased equally, thus a geometrical optimization rule. Here, the load resistance is equivalent to a geometric mean where two extreme cases of the synaptic full facilitation and full depression are considered. To verify the proposed rule, we perform a semiconductor device simulation for a static and pulsed characteristics of the SPT. When the SPT is scaled based on the geometrical optimization rule, from the simulation results, it is found that the static-power consumption is decreased while maintaining the DR_w. As a trade-off relation, however, the noise power-spectral-density is found to be increased due to a higher thermal noise associated with a higher total resistance of the scaled SPT. Here, the increased noise power-spectral-density of a single SPT may affect the performance of the AA based on the SPT-array, so we also show a crossbar simulation, checking the classification accuracy against a standard dataset (e.g. MNIST).

INDEX TERMS Synaptic pass-transistor, geometrical optimization rule, analog accelerator, synaptic weight, synaptic dynamic ratio, full depression, full facilitation, maximum static-power consumption.
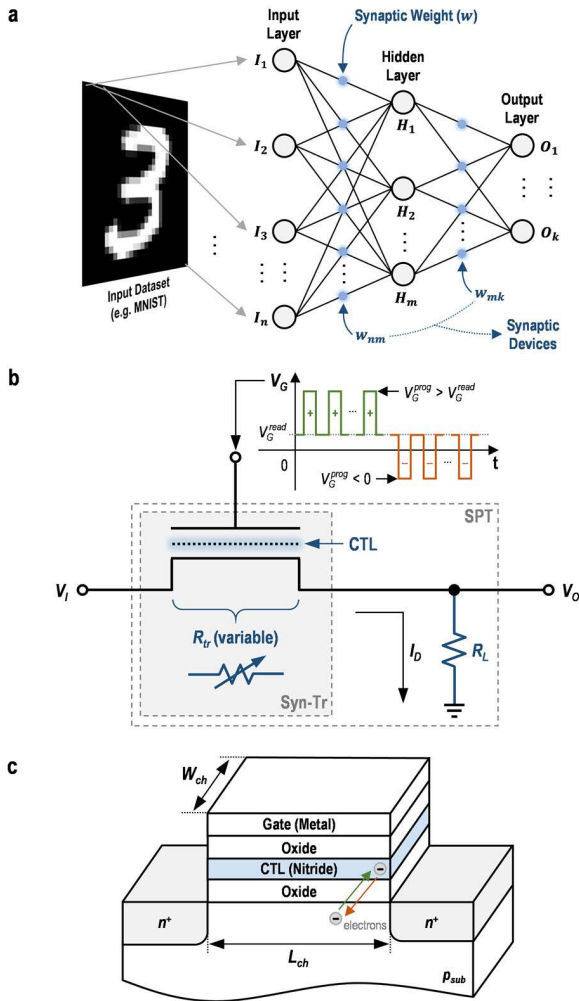
I. INTRODUCTION

A neuromorphic system, which mimics behaviors of a biological nervous system, has been considered as a promising candidate due to a more efficient computing capability compared to the conventional computer system based on the von Neumann architecture [1]–[5]. For example, the data can be efficiently processed with an artificial neural network (ANN) which mimics the connection of the biological neurons. Here, a multi-layer perceptron (MLP) structure can be employed for a simple ANN (see Fig. 1(a)) [6], [7]. Besides a higher data processing efficiency of the neuromorphic system, a low-power characteristics of a synaptic device, which is a fundamental building block, is essential for

The associate editor coordinating the review of this manuscript and approving it for publication was Marcelo Antonio Pavanello.

a high-level intelligence of the system [8]–[11]. Here, a level of the intelligence is proportional to the number of synaptic devices, resulting in the increase of a total power consumption proportionally. In this respect, the power consumption of a single synaptic device has to be minimized. In the case of the human brain having a high-level intelligence, it is known that its total power consumption is only 10 W in spite of a huge number of synapses (i.e. 10^{15}) in the brain [12], [13]. This is mainly because the power consumption per synapse is very small level of 0.01 pW. Note that a single central processing unit (CPU) for the conventional computer system shows a high-power consumption of 100 W approximately, which is 10 times higher compared to the total power consumption of the brain [14], [15].

As a typical two-terminal synaptic device, a memristor with a low-power consumption has been reported,

FIGURE 1. (a) Schematic of an artificial neural network (ANN) as a simple multi-layer perceptron (MLP) structure with three layers (e.g. the input layer, hidden layer, and output layer). Here, the I_n is n^{th} neuron of an input layer, H_m is m^{th} neuron of a hidden layer, O_k is k^{th} neuron of an output layer, w_{nm} is a synaptic weight between In and H_m, and w_{mk} is a synaptic weight between H_m and O_k. (b) Schematic of the synaptic pass-transistor (SPT) with a charge trapping layer (CTL) and (c) the three-dimensional view of the synaptic transistor (Syn-Tr). Here, the equivalent resistance (R_{tr}) of the Syn-Tr can be varied when electrons are trapped or de-trapped depending on the polarity of the programming-voltage (V_G^{prog}) and number of programming-pulses applied to the gate.

indicating its power consumption in the range of nano-watts [11], [16]–[19]. However, that is about six orders of magnitude larger compared to a biological synapse (e.g. 0.01 pW per synapse). It may be because a high voltage to generate memristive phenomena, such as an electro-chemical metallization and valency change, is still required [20], [21]. Therefore, it is expected that a higher-level intelligence of the neuromorphic system based on memristors with these memory phenomena is difficult to be achieved due to a limitation for reducing the power consumption of each device. As another type, three-terminal synaptic devices based on a metal-oxide-semiconductor field-effect transistor

(MOSFET) have been investigated for a low power operation [22]–[24]. For a memory function of this synaptic device, programming-pulses are required to be applied to a gate terminal for a memory function, such as a read/write operation, in these synaptic devices. However, a high programming-voltage for both short-term memory phenomena (e.g. an impact ionization) and long-term memory (e.g. a trapping and de-trapping of charges through a tunneling) phenomenon of the MOSFET-based synaptic device is needed giving rise to a high burst current. This is expected to lead to an increase of the static-power consumption, thus a limitation of a high-level intelligence in the neuromorphic system.

As a method to limit the burst current, synaptic devices can be connected with a load resistor, which leads to a low-power operation. For example, a synaptic pass-transistor (SPT), where a n-channel synaptic transistor (Syn-Tr) is connected in series with the load resistor, was introduced, as depicted in Fig. 1(b) [25]. Here, a load resistance (R_L) of the SPT limits the maximum current level as V_I / R_L where the V_I is a fixed input voltage. In addition, the power dissipation of the SPT can be reduced by the increase of R_L. However, the increase of R_L may cause side-effects, such as a degradation of a synaptic characteristics. Therefore, a selection of an appropriate R_L is important.

In this paper, a simulation framework on a geometrical optimization rule (GOR) of the SPT for a low power analog accelerator (AA). Here, the SPT is a Syn-Tr in series with the load resistor, is presented scaling R_L with respect to a geometrical ratio (L_{ch}/W_{ch}) between a channel length (L_{ch}) and width (W_{ch}) of the Syn-Tr for a low-power operation. To reduce the power consumption of the SPT, when only the R_L is increased, the synaptic characteristics, such as a synaptic dynamic ratio (DR_w) between the fully-facilitated and fully-depressed weights, are difficult to be maintained. To maintain this DR_w, the L_{ch}/W_{ch} and scaling factor of R_L are needed to be increased with the same rate, so a GOR is suggested. Here, the R_L is equivalent to a geometric mean where two extreme synaptic states of the full facilitation and full depression of the Syn-Tr are considered. To check the GOR, a semiconductor device simulation for the static and pulsed characteristics of the SPT is performed. When both the L_{ch}/W_{ch} and scaling factor of R_L are equally scaled up with applying the GOR, it is found that the static-power consumption is reduced while maintaining the DR_w wider compared to the case without applying the GOR. As a trade-off relation, however, it is found that the noise power-spectral-density (PSD) of the SPT is increased due to the increase of the thermal noise which is related to the increase of R_L and equivalent resistance (R_{tr}) of the Syn-Tr. Here, when the SPTs are arrayed as an AA, it is expected that the increased noise PSD of a single SPT can lead to the decrease of the classification accuracy of the AA based on the SPT-array. From the crossbar simulation for the Modified National Institute of Standards and Technology database (MNIST), it is also found that the classification accuracy is decreased due to the
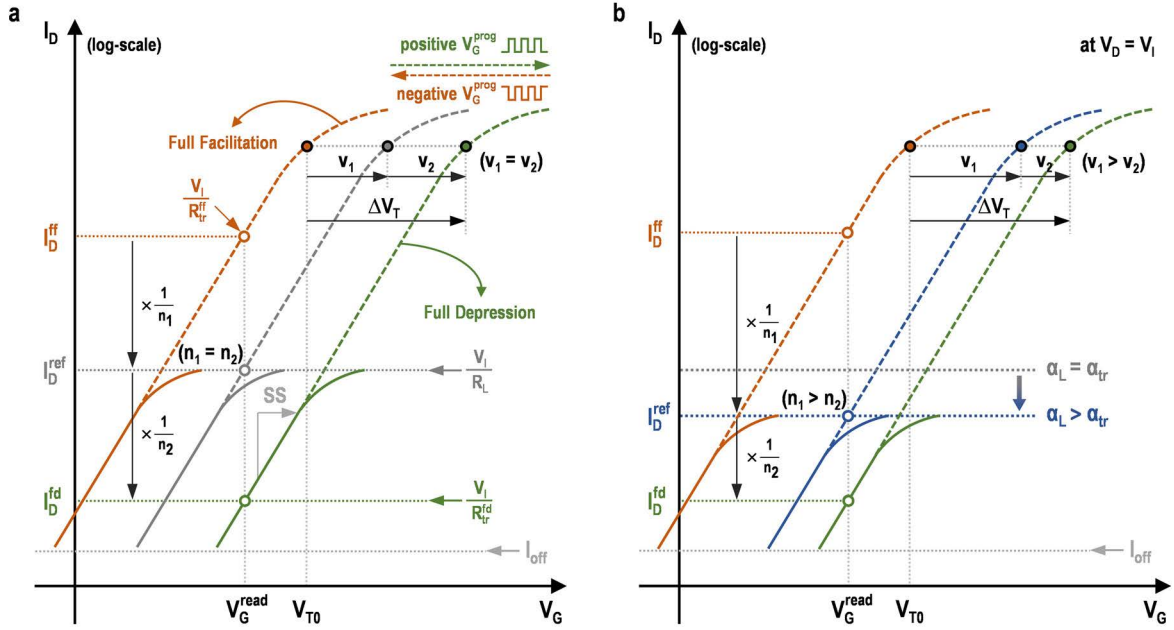
FIGURE 2. (a) Conceptual plot of the current-voltage characteristics to express the synaptic process through the threshold voltage shifts changed by programming-pulses. Here, the SS is a subthreshold slope. (b) Notional plot of the current-voltage characteristics when only α_L is increased without applying the geometrical optimization rule (GOR). Note that the v_1 and v_2 are an amount of the shifted V_T.

higher noise PSD, as increasing R_{tr} and R_L. Therefore, it is shown that the classification accuracy can be decreased due to the increased total resistance of each arrayed SPT although we can expect the reduction of the power consumption in the AA, suggesting a trade-off relation between the power consumption and classification accuracy.

II. SPT AND RELATED THEORIES

A. BASIC STRUCTURE AND OPERATING PRINCIPLE

As shown in Fig. 1(b), the SPT consists of the Syn-Tr and R_L. Here, the R_L is added for a voltage output (V_O), employing the pass-transistor concept. For the Syn-Tr, a memory function is implemented with a charge trapping layer (CTL) in the insulator system based on a semiconductor-oxide-nitride-oxide-semiconductor (SONOS), as seen in Fig. 1(c). With this CTL, charges (e.g. electrons) are trapped or de-trapped depending on the polarity of the programming-voltage (V_G^{prog}) and number of programming-pulses applied to the gate. Thus, a threshold voltage (V_T) is varied according to the number of electrons trapped in the CTL, which results in a change of the drain current (I_D) at a fixed read voltage (V_G^{read}). Since the R_{tr} can be expressed as V_I/I_D, the change of I_D leads to a variation of R_{tr}. Assuming that the SPT operates only in the subthreshold region, the R_{tr} can be represented as follows [25],

$$R_{tr} = \frac{L_{ch}}{W_{ch}} \frac{V_I}{K_{sub} exp\left\{q(V_G^{read} - V_T)/(n_t kT)\right\}}, \quad (1)$$

where K_{sub} is a constant related to the diffusion coefficient in the subthreshold regime, n_t is an ideality factor about the

interface state, kT is a thermal energy, q is an elementary charge, and $V_T = V_{T0} + \Delta V_T$. Here, ΔV_T is the amount of the threshold voltage shift from the initial value (V_{T0}). Also, ΔV_T can be represented as $-Q_e^{trap}/C_i$, where Q_e^{trap} is the magnitude of a negative charge density per area, corresponding to the number of electrons trapped in the CTL, and C_i is a gate-insulator capacitance per area. Note that the operation in the subthreshold regime can be a key aspect for an ultralow-power synaptic device due to a low operating current and high sensitivity of the current for a small variation of ΔV_T [26]. By defining $L_{ch}/W_{ch} \equiv \alpha_{tr}$ and $R_{tr}/\alpha_{tr} \equiv R_{tr0}$, Eq. (1) can be rewritten as,

$$R_{tr} = \alpha_{tr} R_{tr0}. \quad (2)$$

Here, the R_{tr0} at $V_G = V_G^{read}$ is the rest channel resistance normalized with α_{tr}, which is a function of V_T depending on the polarity and number of programming-pulses. Since the SPT can be considered as R_{tr} in series with R_L, the V_O is determined as,

$$V_O = \frac{R_L}{R_L + R_{tr}} V_I. \quad (3)$$

With the definition of w ($\equiv V_O/V_I$), the w can be expressed as shown below,

$$w = \frac{R_L}{R_L + R_{tr}}. \quad (4)$$

Thus, the w is varied depending on the change of R_{tr} which is a function of ΔV_T.

B. DETAILED PROCEDURE OF THE SYNAPTIC OPERATION

As the Syn-Tr is based on the n-channel MOSFET, the initial state of the SPT is a full facilitation (FF) with $V_T = V_{T0}(\Delta V_T = 0)$ (see Fig. 2(a)). The channel resistance at the FF (i.e. R_{tr}^{ff}) can be expressed as V_I/I_D^{ff}, where the I_D^{ff} is the drain current for the FF at $V_G = V_G^{read}$. From this FF condition, multiple positive programming-pulses are applied to get a positive ΔV_T of $v_1 > 0$, which leads to the decrease of I_D at $V_G = V_G^{read}$. As shown in Fig. 2(a), when the I_D has become n_1 times lower compared to the FF condition, it is defined as the reference drain current (I_D^{ref}). This state is to be called a reference state of the SPT with $V_T = V_{T0} + v_1$ at which the reference channel resistance (i.e. R_{tr}^{ref}) is denoted as V_I/I_D^{ref}. Since $I_D^{ref} = I_D^{ff}/n_1$, the R_{tr}^{ref} can be expressed as,

$$R_{tr}^{ref} = n_1 R_{tr}^{ff}. \qquad (5)$$

Here, the R_{tr}^{ref} is to be R_L for the SPT operation, so the R_L is given from Eq. (5),

$$R_L = n_1 R_{tr}^{ff}. \qquad (6)$$

Since the w at the FF (w^{ff}) is found by replacing the R_{tr} with R_{tr}^{ff} in Eq. (4), the w^{ff} can be represented with Eq. (6), as follows,

$$w^{ff} = \frac{R_L}{R_L + R_{tr}^{ff}} = \frac{n_1}{n_1 + 1} \approx 1, \qquad (7)$$

where w^{ff} is almost unity (i.e. the FF) with the condition of $n_1 \gg 1$. Note that the w^{ff} can be generalized as a maximum w (w_{max}) for any n_1, as follows,

$$w_{max} = \frac{n_1}{n_1 + 1}. \qquad (8)$$

By applying multiple positive programming-pulses, the V_T can be further increased by $v_2 > 0$ from the reference condition. So, the I_D is further decreased by n_2 times, as described in Fig. 2(a). For this case, a full depression (FD) with $V_T = V_{T0} + v_1 + v_2(\Delta V_T = v_1 + v_2)$ is assumed to be made. The channel resistance at the FD (i.e. R_{tr}^{fd}) is represented as V_I/I_D^{fd}. Here, the I_D^{fd} is the drain current for the FD at $V_G = V_G^{read}$. Since $I_D^{fd} = I_D^{ref}/n_2$, the R_{tr}^{fd} is expressed as $n_2 R_{tr}^{ref}$. Along with Eq. (5), it can also be rewritten as,

$$R_{tr}^{fd} = n_1 n_2 R_{tr}^{ff}. \qquad (9)$$

In Eq. (4), the R_{tr} is R_{tr}^{fd} at the FD. Along with Eqs. (6) and (9), the w at the FD (w^{fd}) is derived as,

$$w^{fd} = \frac{R_L}{R_L + R_{tr}^{fd}} = \frac{1}{1 + n_2} \approx 0, \qquad (10)$$

where the w^{fd} approaches zero (i.e. the FD) with the condition of $n_2 \gg 1$. Similar to Eq. (8), the w^{fd} can also be generalized as a minimum w (w_{min}) for any n_2, as follows,

$$w_{min} = \frac{1}{1 + n_2}. \qquad (11)$$

In addition, with the w_{max} and w_{min} of the SPT, the DR_w is defined as,

$$DR_w \equiv \frac{w_{max}}{w_{min}} = \frac{n_1(1 + n_2)}{n_1 + 1}. \qquad (12)$$

Here, the DR_w becomes n_2 with the condition of both n_1 and $n_2 \gg 1$ while achieving both $w_{max} = w^{ff}$ and $w_{min} = w^{fd}$ of the SPT.

C. LOAD RESISTANCE AND TWO EXTREME SYNAPTIC STATES

Since the I_D^{fd} is limited to the off-current floor (I_{off}), the n_1 and n_2 are also limited, thus limited ΔV_T. In this respect, the optimum condition of n_1 and n_2 should be found with the limited number of programming-pulses. Assuming that the transistor is always operated in the subthreshold regime and subthreshold slope (SS) is constant, it is found that $n_1 n_2$ is a constant for the constant ΔV_T as $v_1 + v_2$. So, $n_1 = n_2$ is optimum to be both the n_1 and $n_2 \gg 1$, satisfying $v_1 = v_2$. Otherwise, either n_1 or n_2 is decreased, which results in the lack of the condition of $n_1 \gg 1$ or $n_2 \gg 1$. Especially, for the case of $n_1 > n_2$, because the FD is difficult to be made (see Eq. (11)), the DR_w could be reduced, as can be seen in Eq. (12). Now, in order to maintain this optimum after a scaling of the SPT, a scaling condition to keep $n_1 = n_2$ is discussed. Based on Eq. (2), the R_{tr}^{ff} can be defined as,

$$R_{tr}^{ff} \equiv \alpha_{tr} R_{tr0}^{ff}. \qquad (13)$$

Here, the R_{tr0}^{ff} is the rest channel resistance at the FF, which is R_{tr0} at $V_T = V_{T0}$. With this, Eqs. (6) and (9) can also be rewritten as, respectively,

$$R_L = n_1 \alpha_{tr} R_{tr0}^{ff}, \qquad (14)$$

$$R_{tr}^{fd} = n_1 n_2 \alpha_{tr} R_{tr0}^{ff}. \qquad (15)$$

Also, the R_L is individually defined as,

$$R_L \equiv \alpha_L R_{L0}, \qquad (16)$$

where the α_L is a scaling factor of R_L and R_{L0} is the rest load resistance. With Eqs. (14) and (16), we set $R_{L0} = n_1 R_{tr0}^{ff}$ and $\alpha_{tr} = \alpha_L$ to maintain $n_1 = n_2$. Under these conditions, the R_L is determined as a geometric mean associated with R_{tr} for two extreme synaptic states from Eqs. (13) to (16),

$$R_L = \sqrt{R_{tr}^{ff} R_{tr}^{fd}}. \qquad (17)$$

So, when the α_L and α_{tr} are changed to the same value, the R_L and R_{tr} are determined obtaining a wide DR_w from Eq. (12). And Eq. (17) is also satisfied. Accordingly, both the α_L and α_{tr} should be equally increased to maintain an optimum state in terms of the synaptic characteristics, thus the GOR.

D. GEOMETRICAL OPTIMIZATION RULE AND POWER CONSUMPTION

As mentioned earlier, since the SPT consists of the Syn-Tr and the load resistor, the increase of R_L can reduce the power
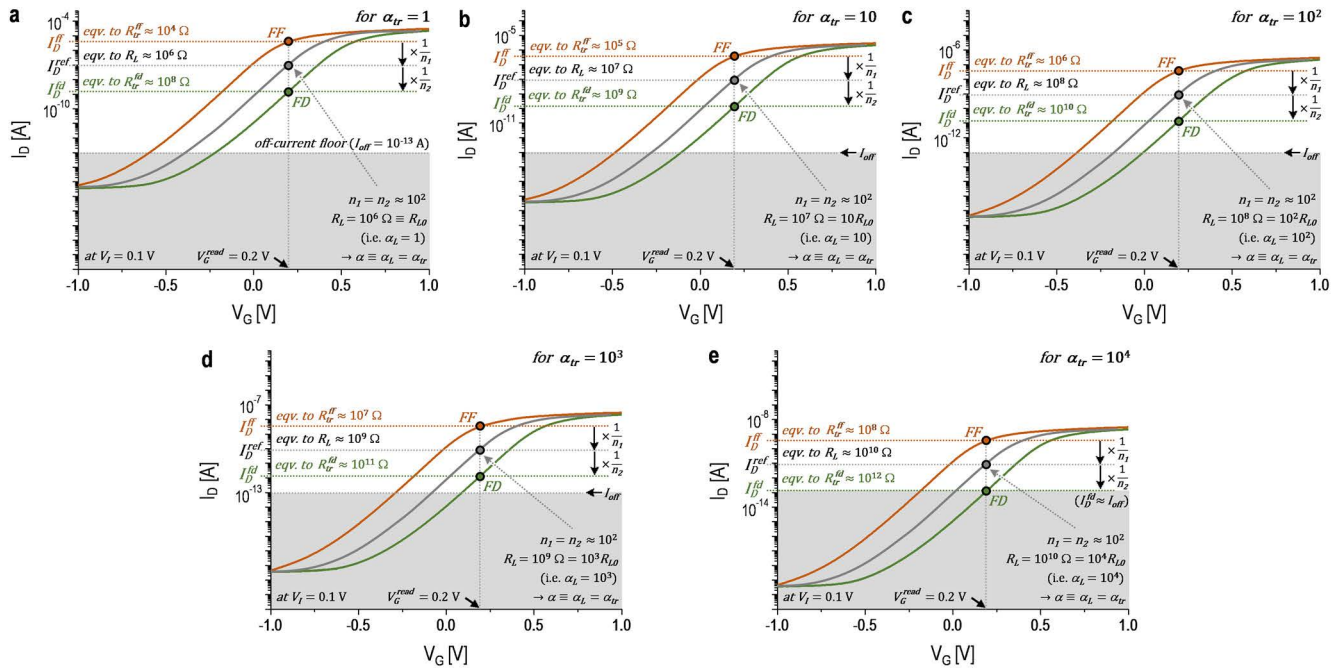
FIGURE 3. Current-voltage characteristics of the Syn-Tr at the full depression (FD) and the full facilitation (FF) for (a) $\alpha_{tr} = 1$, (b) $\alpha_{tr} = 10$, (c) $\alpha_{tr} = 10^2$, (d) $\alpha_{tr} = 10^3$, and (e) $\alpha_{tr} = 10^4$ ($\alpha_{tr} = \alpha_L \equiv \alpha$). The load resistance can be determined at each reference current level where $n_1 = n_2 \gg 1$, which satisfies the GOR. Here, the off-current floor (I_{off}) is assumed to be 10^{-13} A.

consumption of the SPT. There are two kinds of the power consumption of the SPT: a static-power consumption (P_{static}) and a dynamic-power consumption (P_{dyn}). Here, the P_{dyn} is negligible because it can be sufficiently smaller than P_{static} by extending the rising and falling time of the programming-pulse. So, it is important to reduce P_{static}, which is dominant in the power consumption of the SPT. Here, since $P_{static} = V_I^2/(R_{tr} + R_L)$, the maximum of P_{static} (P_{static}^{max}) in the SPT is expressed for $R_{tr} \ll R_L$, as follows [27],

$$P_{static}^{max} \approx \frac{V_I^2}{R_L}. \tag{18}$$

Thus, the increase of α_L makes R_L bigger, which results in the decrease of P_{static}^{max} as seen in Eq. (18). Here, according to the proposed GOR, the α_{tr} should be increased together to keep the condition of $\alpha_{tr} = \alpha_L$ for $n_1 = n_2$. This can also keep an optimally large DR_w. However, when only the R_L is increased, the α_L is greater than α_{tr}, thus $n_1 > n_2$ (see Fig. 2(b)). For this case, the FD is difficult to be made because the w_{min} is hard to approach zero, which can be explained with the decreased n_2 (see Eq. (11)). So, based on Eq. (12), the lack of the synaptic depression happens while the DR_w is smaller. To examine this and a low-power consumption of the SPT scaled with the GOR, it is suggested to firstly find a proper R_L with the static characteristics of the Syn-Tr. After finding a proper R_L based on the GOR, the power consumption and DR_w of the scaled SPT can be analyzed with the pulsed characteristics for the synaptic depression and facilitation.

III. RESULTS AND DISCUSSION

A. METHODOLOGY

Based on the theoretical analysis on the GOR in the previous section, a low-power consumption and wide DR_w of the SPT are expected to be achieved, as applying the GOR. To check this, a static and pulsed characteristics of the SPT (i.e. a combination of the Syn-Tr and R_L) are shown in the following Sections III-B and C. For this, the semiconductor device simulation is performed with the SILVACO™ ATLAS™. Here, based on the GOR, the α_{tr} and α_L are set to be the same value, which is defined as a single scaling factor $\alpha(\equiv \alpha_{tr} = \alpha_L)$ for a convenience. As increasing α, the R_{tr} is scaled up with the same rate as the R_L. Besides α_{tr} and α_L, other device properties are given in Ref. 26. As increasing the α, it is also expected that the noise characteristics (e.g. a noise PSD) of the SPT is varied. To verify this, the noise characteristics is monitored.

With the extracted data for both the synaptic weight-update and noise characteristics in the device level, when a SPT-array composed of these SPTs is used as the AA, it is expected that the performances (e.g. classification accuracy (CA)) of the AA are affected due to the variation of the noise PSD. To check this, the Δw-w plot and classification accuracy are described in the following Section III-D. For this, the AA simulation is performed using a resistive memory simulator, called the CrossSim™ of which the detailed description is shown in the Supplementary Material (see Section S1) [28]–[30]. Also, the descriptions with respect to the validation of the proposed models (i.e. the SPT

and AA) are shown in the Supplementary Material (see Section S2).

B. STATIC CHARACTERISTICS OF THE SYN-TR

As seen with the graphical analysis in Fig. 2(a), to find the value of R_L which satisfies the condition of $n_1 = n_2$, the static characteristics of the Syn-Tr needs to be analyzed. In addition, to check whether this condition of $n_1 = n_2$ can be maintained with applying the GOR, the current-voltage characteristics is required to be monitored as scaling the Syn-Tr. Hence, we show the transfer characteristics of each Syn-Tr scaled up with α_{tr}. Fig. 3 shows the I-V curves at the FF and FD of the Syn-Tr, respectively, for five different values of α_{tr}. Here, when the R_L as the geometric mean of R_{tr}^{ff} and R_{tr}^{fd} is determined with Eq. (17), the α_{tr} and α_L are to be the same value, thus the GOR. As depicted in Fig. 3(a), for $\alpha_{tr} = 1$ as a reference case, it is found that $I_D^{ff} \approx 10^{-5}$ A and $I_D^{fd} \approx 10^{-9}$ A at $V_G^{read} = 0.2$ V. With them, the level of I_D^{ref} is chosen as 10^{-7} A where $n_1 = n_2 \approx 10^2$. With this, the equivalent R_L is calculated as $10^6 \Omega$ for $V_I = 0.1$ V. This is because $R_L \equiv V_I / I_D^{ref}$, which is also corresponding to the geometric mean for $R_{tr}^{ff} = 10^4 \Omega$ and $R_{tr}^{fd} = 10^8 \Omega$, using Eq. (17). When the value of this load resistor is set to the R_{L0}, the α_L becomes unity, thus $\alpha_L = \alpha_{tr} = 1$. Based on this reference case, as seen in Figs. 3(b)-(e), when the α_{tr} is gradually scaled up by a factor of 10, it is found that the level of I_D^{ff} and I_D^{fd} is equally decreased tenfold. Meanwhile, the I_D^{ref} is selected as a level decreased tenfold, so the R_L is increased with the rate of α_{tr}, which is found to satisfy $\alpha_L = \alpha_{tr}$ with Eq. (16). It is because the n_1 and n_2 are maintained as the same value of 10^2, approximately. This is also consistent with the geometric mean of R_{tr}^{ff} and R_{tr}^{fd} for each case. These results indicate that the α_{tr} and α_L are required to be increased equally to maintain the synaptic characteristics, such as a DR$_w$ with the optimum condition of n_1 and n_2.

C. PULSED CHARACTERISTICS OF THE SYN-TR

From the analysis on the static characteristics of the Syn-Tr, it is expected that the DR$_w$ of the SPT is maintained, which satisfies $n_1 = n_2 \approx 10^2$ with the GOR. In addition, the increase of both R_{tr} and R_L based on the GOR can result in the decrease of I_D. So, it is expected that the P_{static} of the SPT is reduced with the decrease of I_D. To verify these DR$_w$ and P_{static}, the pulsed characteristics for V_O and I_D of the SPT needs to be monitored for the synaptic depression and facilitation processes, respectively, depending on the polarity and number of the programming-pulses.

For the synaptic depression, the positive programming-pulses are applied with $V_G^{prog} = 8$ V, $V_G^{read} = 0.2$ V, and pulse-width $= 1~\mu$s (see Fig. 4(a)). After 20 cycles of the positive programming-pulses, it starts to apply the negative programming-pulses for the synaptic facilitation, where $V_G^{prog} = -4$ V, $V_G^{read} = 0.2$ V, and pulse-width $= 2~\mu$s
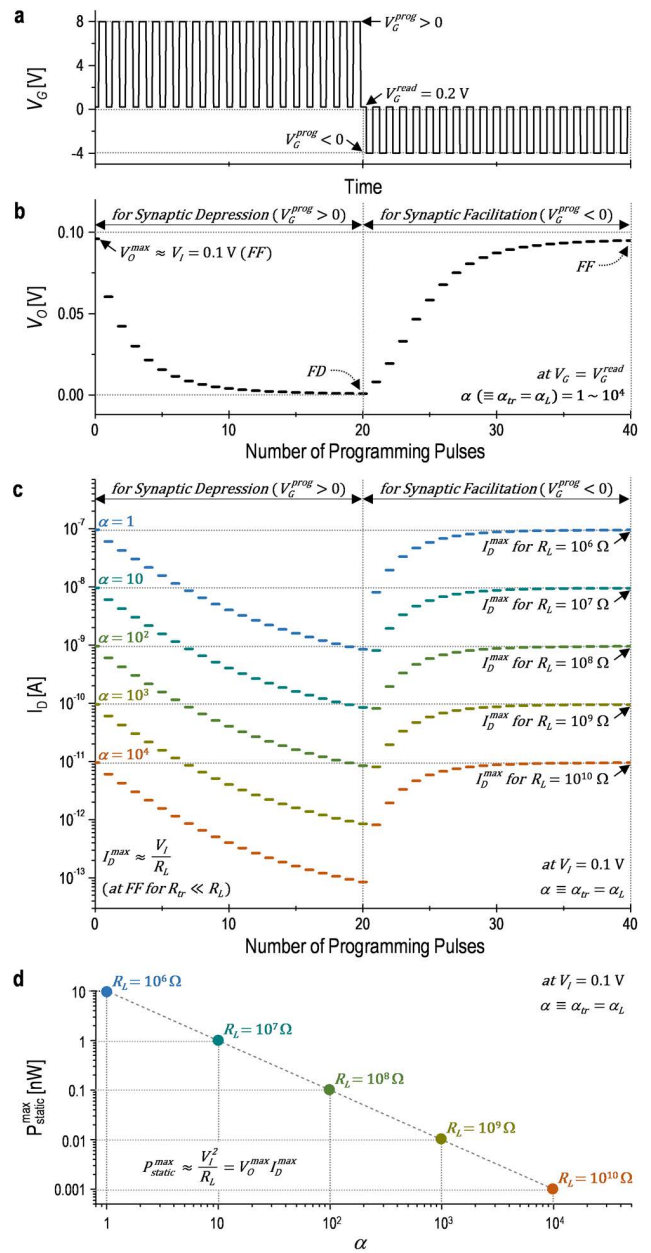

FIGURE 4. (a) Waveform of the positive programming-pulses (with $V_G^{prog} = 8$ V and $V_G^{read} = 0.2$ V) and negative programming-pulses (with $V_G^{prog} = -4$ V and $V_G^{read} = 0.2$ V) for the synaptic depression and facilitation, respectively. (b) Voltage output which appears on the load resistor and (c) drain current of the SPT over 40 cycles of programming-pulses for five different values of α. (d) The maximum static power consumption as a function of α which is calculated with the maximum output voltage (V_O^{max}) and maximum drain current (I_D^{max}) indicated in (b) and (c), respectively.

(see Fig. 4(a)). Note that common specifications of the programming-pulses are as follows, the duty cycle, rising time, and falling time are 50 %, 10^{-7}s, and 10^{-7}s, respectively. Fig. 4(b) shows the V_O sampled at $V_G = V_G^{read}$ for each α, where the positive and negative programming-pulses are repeated 20 times for the synaptic depression and facilitation processes, respectively.

As shown in Fig. 4(b), for $\alpha = 1$ (reference case), it is found that the V_O is gradually decreased with applying the positive programming-pulses for the synaptic depression, approaching $V_O \approx 0$. It is because the R_{tr} is gradually increased, which can be explained with Eqs. (1) and (3). And the V_O finally arrives at around zero, corresponding to the FD as a synaptic state, because the R_{tr} is approximately 100 times larger compared to R_L, suggesting $n_2 \approx 10^2$ (see Eq. (3)). Afterward, for the synaptic facilitation, it is also found that the V_O is gradually increased with a series of negative programming-pulses. This is because the R_{tr} is gradually decreased, which can also be explained with Eqs. (1) and (3). After 20 cycles of negative programming-pulses for the synaptic facilitation, the V_O approaches the maximum value (V_O^{max}) of V_O (i.e. $V_O^{max} \approx V_I = 0.1$ V), which corresponds to the synaptic state of the FF. This is due to a relative dominance between R_{tr} and R_L (i.e. $R_{tr} \ll R_L$), satisfying $n_1 \approx 10^2$ (see Eq. (3)). As this reference case, although the α is increased up to 10^4, it is shown that the trends of V_O remain the same in all cases of α. This is because both the R_{tr} and R_L are equally scaled up with α. It is consistent with Eq. (3) where the α is canceled out.

Since it is expected that these increased R_{tr} and R_L can lead to the decrease of I_D, the I_D as another factor of P_{static} needs to be monitored. Fig. 4(c) describes the I_D also sampled at $V_G = V_G^{read}$ for the synaptic depression and facilitation processes, respectively, for five different values of α. As seen in Fig. 4(c), for the reference case (i.e. $\alpha = 1$), it is found that the I_D decays with a series of the positive programming-pulses for the synaptic depression. This is also because the R_{tr} is gradually increased with the positively shifted V_T (i.e. $\Delta V_T > 0$), which is explained with Eq. (1) and $I_D = V_I/(R_{tr} + R_L)$ for $V_I = 0.1$ V. After the positive programming-pulses continued for 20 cycles, the I_D reaches the current level of 10^{-9} A, which is two orders of magnitude smaller compared to its maximum level (I_D^{max}). This is because the R_{tr} is 100 times larger compared to R_L (i.e. $n_2 \approx 10^2$). Afterward, for the synaptic facilitation, the I_D is gradually increased with applying the negative programming-pulses, recovering to I_D^{max}. It is because of the R_{tr} decreased gradually with a negative shift (i.e. $\Delta V_T < 0$) of V_T, which is also explained with Eq. (1) and $I_D = V_I / \alpha(R_{tr} + R_L)$ for $V_I = 0.1$ V. And the I_D finally arrives at I_D^{max} due to $R_{tr} \ll R_L$, satisfying $n_1 \approx 10^2$. When the α is increased by a factor of 10, it appears that the I_D is scaled down tenfold while maintaining the same trend as the reference case. It is because the R_{tr} and R_L are equally scaled up with α, satisfying $I_D = V_I / (R_{tr} + R_L)$ for $V_I = 0.1$ V (see Eqs. (2) and (16)). Accordingly, as seen in Figs. 4(b) and (c), since the V_O is equally maintained for all cases of α, it implies that the P_{static} can be reduced with the decrease of I_D. Here, as a maximum value of P_{static}, the P_{static}^{max} can be calculated as a product of V_O^{max} and I_D^{max} at the FF for each α, as shown in Fig. 4(d). For the reference case (i.e. $\alpha = 1$), as can be seen, it is found that the P_{static}^{max} is about 10 nW, which is consistent with the calculation for $R_L = 10^6 \Omega$ and $V_I = 0.1$ V (see Eq. (18)). When the α is increased
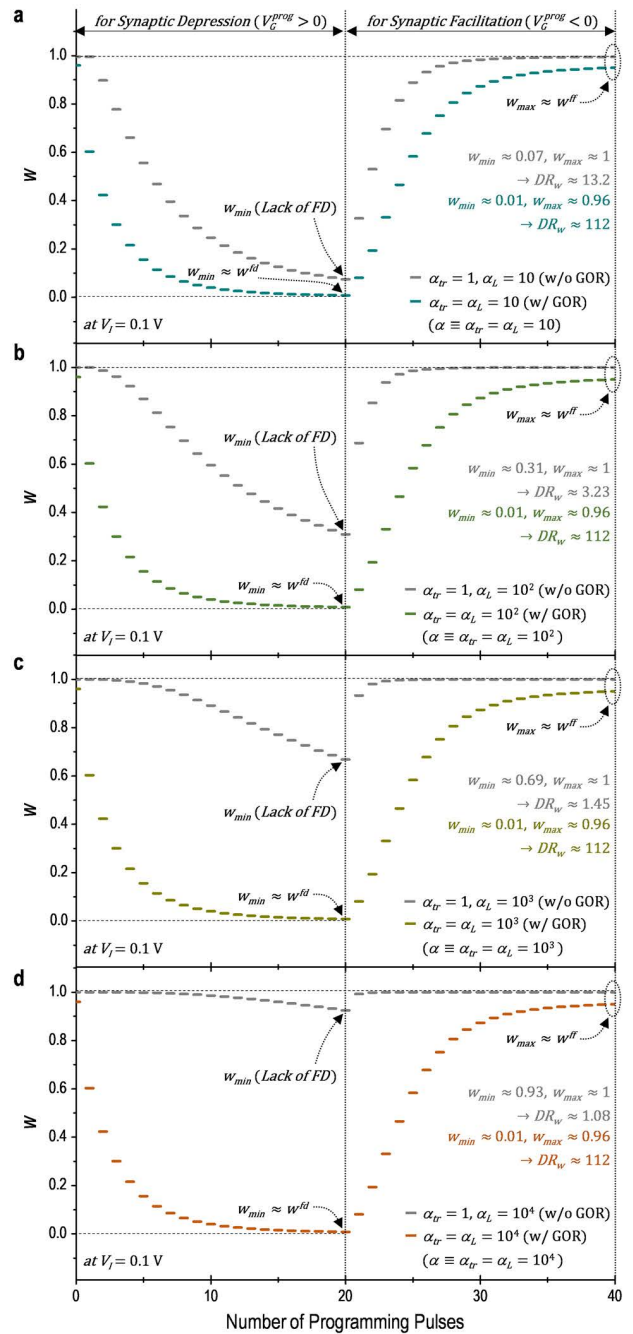

FIGURE 5. Synaptic weight versus the number of programming-pulses for two cases of $\alpha_{tr} = \alpha_L$ (with the GOR) and $\alpha_{tr} = 1 \ll \alpha_L$ (without applying the GOR) under four different conditions of (a) $\alpha_L = 10$, (b) $\alpha_L = 10^2$, (c) $\alpha_L = 10^3$, and (d) $\alpha_L = 10^4$. We apply the same number of pulses for the synaptic depression and facilitation, respectively, considering symmetrical aspect of the weight-update.

by a factor of 10, as expected, the P_{static}^{max} is reduced due to I_D^{max} decreased with α, as shown in Fig. 4(c). This can also be explained with Eq. (18). The increase of α leads to the increased R_L (see Eq. (16)), thus a reduction of P_{static}^{max} for a fixed V_I. Note that the P_{dyn} is expected to be observed for the synaptic depression and facilitation processes, respectively,
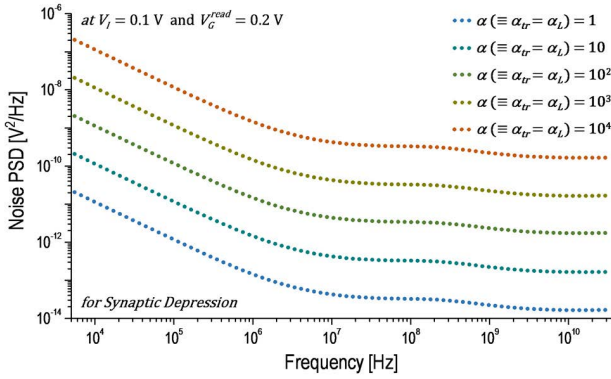
FIGURE 6. PSD of the voltage noise extracted for the SPT scaled with five different values of α.
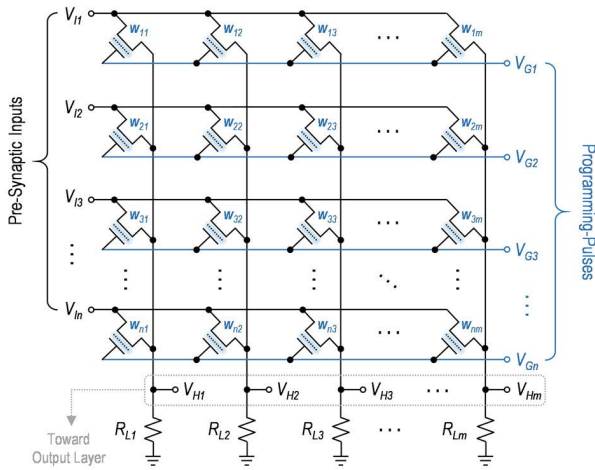
FIGURE 7. Conceptual diagram of the $n \times m$ SPT-array (i.e. SPT-based AA) constructed for the crossbar simulation. For every column of the SPT-array, SPTs are connected with a single load resistor. Here, the synaptic weight of each SPT (w_{nm}) is varied with the nth input voltage (V_{In}), n^{th} gate voltage (V_{Gn}), and m^{th} load resistance (R_{Lm}). The voltage output for the mth column, which appears on each load resistor, is the signal for the hidden layer (V_{Hm}) to be transferred toward the output layer of the ANN.

due to the existence of the displacement current associated with the voltage difference changed for a rising time and falling time of the programming-pulse.

According to the trend of V_O observed in Fig. 4(b), the w, which is defined as V_O/V_I, is found to be fully depressed (i.e. $w_{min} \approx w^{fd}$) and fully facilitated (i.e. $w_{max} \approx w^{ff}$) regardless of the increased α to reduce the power consumption of the SPT. In other words, the DR_w is expected to be wider than the DR_w for the case without applying the GOR. To clarify this, the weight dynamics needs to be monitored for the comparison between SPTs scaled with the GOR and without applying the GOR. Fig. 5 depicts the synaptic weight dynamics for two cases of $\alpha_{tr} = \alpha_L$ (i.e. with the GOR) and $\alpha_{tr} = 1 \ll \alpha_L$ (i.e. without applying the GOR), which is plotted for each α_L. Here, for the scaling based on the GOR, the w is calculated with its definition (i.e. $w \equiv V_O / V_I$) and the given input

voltage (i.e. $V_I = 0.1$ V), as seen in Fig. 4(b). Similarly, for the case without applying the GOR, the w is also extracted from the V_O over 40 cycles of programming-pulses.

As depicted in Fig. 5, the trends of w remain identical in all the GOR-based cases, corresponding to the trends of V_O, as seen in Fig. 4(b). This appears due to the R_{tr} and R_L scaled up with the same rate of α, which can also be explained with Eq. (4) where all scaling factors are canceled out. Further, it is found that the DR_w is maintained as 112, which is calculated with Eq. (12). This is because the optimum condition of $n_1 = n_2 \approx 10^2$ is always satisfied with the GOR, as shown in Fig. 3. On the other hand, for the cases of $\alpha_{tr} = 1 \ll \alpha_L$, it is also found that the DR_w is reduced with the increase of α_L. As mentioned earlier, since the product of n_1 and n_2 is constant, the n_1 increases and n_2 decreases when only α_L is increased for a fixed α_{tr} (i.e. $\alpha_{tr} = 1$). Then, the w_{min} is very difficult to be zero due to the decreased n_2, which is explained with Eq. (11). In this respect, the DR_w is reduced (see Eq. (12)), approaching $1+n_2$ with the increase of n_1 and decrease of n_2. Note that, for the case of $\alpha_{tr} = 1 \ll \alpha_L$, the w_{max} is easy to be almost unity because of n_1 which is larger compared to n_2. And this can also be explained with Eq. (8).

Consequently, it is verified that the DR_w can be maintained as a relatively large value of 112 while achieving a low-power consumption of the SPT when both the R_{tr} and R_L are scaled up with applying the GOR.

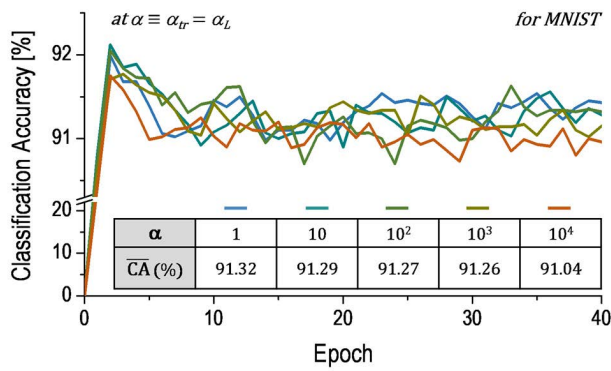### D. CLASSIFICATION ACCURACY OF THE SPT-ARRAY AS THE ANALOG ACCELERATOR

When the α is increased, the P_{static} is reduced due to the increase of R_{tr} and R_L while maintaining a wide DR_w of 112 for $w_{min} \approx w^{fd}$ and $w_{max} \approx w^{ff}$. However, it is expected that the increased R_{tr} and R_L can lead to a higher PSD of the voltage noise (e.g. a thermal noise and flicker noise) for the SPT scaled with the GOR. Indeed, as shown in Fig. 6, when the α increases, it is found that the PSD of the voltage noise is gradually increased with the rate of α. This is because the thermal noise, which is a dominant noise source in the SPT, is proportional to the R_{tr} and R_L [31].

When a SPT-array composed of these SPTs (see Fig. 7) is used as the AA based on the MLP structure seen in Fig. 1(a), it is expected that the increased noise PSD of each SPT can result in the degradation of performances, such as a classification accuracy (CA), in the AA. To verify this, it is required to firstly monitor the weight-update of the SPT-array based on the noise characteristics shown in Fig. 6. Fig. 8 shows the weight-update of the SPT-array with color maps of the cumulative distribution function (CDF) for five different values of α. The CDF distribution illustrated in Figs. 8(a) and (b) represents the probabilistic change of w with respect to a given synaptic state for each synaptic process (e.g. a synaptic depression and facilitation). Here, Δw is defined as a difference between w_n and w_{n-1} where the w_n is w after applying the nth programming-pulse and w_{n-1} is w before applying the nth programming-pulse. These are extracted based on the weight dynamics
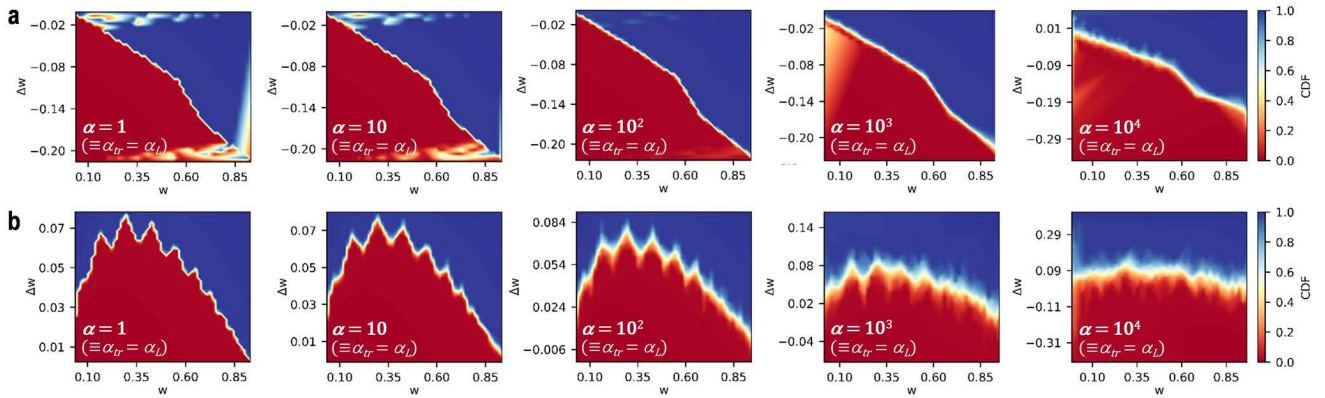
FIGURE 8. Δw vs w plots for (a) the synaptic depression and (b) synaptic facilitation for each α shown with the cumulative distribution function (CDF) of the probabilistic change in w. Here, 20 cycles of the programming-pulses are applied to arrayed SPTs for the synaptic depression and facilitation, respectively.

FIGURE 9. Classification accuracies of the SPT-array based AA versus training epochs for each α. In the inset, the comparison table of average classification accuracies for different values of α is shown. Here, the classification task is tested for the MNIST.

and noise PSD shown in Figs. 5 and 6, respectively, with a resistive memory simulator, called the CrossSim^{TM}. As seen in Fig. 8(a) and (b), the variability of Δw becomes greater with the increase of α for the synaptic depression and facilitation, respectively. It is because of the increased noise PSD of each scaled SPT, as depicted in Fig. 6. This implies that the weight-update can be less predictable for the SPT-array scaled with the increased α compared to the reference case. These results also suggest that the training for the classification task may be more disturbed with the increased noise PSD of scaled SPTs, leading to a decrease of CA in the AA.

To clearly check the relation between the CA of the SPT-array and noise PSD of a single SPT, the CA with respect to the weight-update shown in Fig. 8 also needs to be monitored with a standard dataset, the MNIST [32], [33]. Fig. 9 describes the CA for the MNIST classification task which is tested for the SPT-array composed of SPTs scaled with five different cases of α. When the α is gradually scaled up by powers of 10, as can be seen, an average value of the CA extracted for 40 epochs is slightly reduced from

91.32 % to 91.04 %, as increasing α. This is because the noise PSD increased with α can reduce the signal to noise ratio (SNR) which is inversely proportional to the noise PSD [31]. In addition, since it is known that the CA is proportional to the SNR, the decreased SNR can lead to a reduction of the CA [34]–[36].

Consequently, these results indicate that the power consumption of a single SPT is reduced while maintaining a relatively wide $\mathrm{DR_w}$ due to the increased R_{tr} and R_L. However, since the noise characteristics of the SPT is degraded by the increase of total resistance, the CA of the AA, where SPTs are arrayed, is decreased, thus a trade-off relation between the power consumption and CA in the AA. In addition, in terms of a scaling of the semiconductor device, the SPT is scaled down with the GOR where both W_{ch} and R_L are considered for maintaining $\mathrm{DR_w}$ and satisfying a low power operation. This is expected that the level of integration would be high, thus a high-level intelligence.

IV. CONCLUSION
In this article, the simulation framework on the GOR has been presented for a low-power SPT which is the Syn-Tr in series with the load resistor scaled with respect to the geometrical ratio between the channel length and width of the Syn-Tr. It has been shown that the synaptic characteristics, such as the $\mathrm{DR_w}$ between the fully-facilitated and fully-depressed weights, can be varied due to the simply-increased load resistance of the SPT for a low-power consumption. To optimize this, the GOR has been suggested, allowing the $\mathrm{DR_w}$ to remain constant by equally increasing the channel geometrical ratio and scaling factor of the load resistance. Here, it has been found that the load resistance can be determined as a geometric mean with respect to extreme states of the synaptic full facilitation and full depression. To substantiate the GOR, the results of the semiconductor device simulation have been analyzed for the static and pulsed characteristics of the scaled SPT. From the SPT scaled with the GOR, it has been found that the low-power operation of the SPT

can be achieved due to a reduced operating current at the read voltage while maintaining the maximum DR_w of 112. As a trade-off relation, however, it has been also found that a higher noise PSD appears due to the increase of the thermal noise which is proportional to the total resistance of the scaled SPT. Additionally, to check the influence of this scaled SPT on the performance of the AA based on the SPT-array, the weight-update and classification task against the MNIST have been simulated. From the crossbar simulation, it has been found that the classification accuracy of this AA is decreased due to the increase of the noise PSD in the scaled SPTs. Here, the power consumption over the SPT-array as the AA can be reduced due to the decreased static-power consumption of a single SPT, thus a trade-off relation between the power consumption and classification accuracy in the AA based on the SPT-array. Consequently, when SPTs are scaled down for a high-level integration associated with the level of intelligence which is proportional to the number of synaptic devices, it can be believed that the GOR would be essential for the semiconductor device scaling, maintaining DR_w and achieving a low power consumption.

REFERENCES

[1] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, "Nanoscale memristor device as synapse in neuromorphic systems," *Nano Lett.*, vol. 10, no. 4, pp. 1297–1301, Apr. 2010.

[2] K. Kim, C.-L. Chen, Q. Truong, A. M. Shen, and Y. Chen, "A carbon nanotube synapse with dynamic logic and learning," *Adv. Mater.*, vol. 25, no. 12, pp. 1693–1698, Mar. 2013.

[3] K. H. Kim, S. Gaba, D. Wheeler, J. M. Cruz-Albrecht, T. Hussain, N. Srinivasa, and W. Lu, "A functional hybrid memristor crossbar-array/CMOS system for data storage and neuromorphic applications," *Nano Lett.*, vol. 12, no. 1, pp. 389–395, Jan. 2012.

[4] J. Shi, S. D. Ha, Y. Zhou, F. Schoofs, and S. Ramanathan, "A correlated nickelate synaptic transistor," *Nature Commun.*, vol. 4, p. 2676, Jan. 2013.

[5] F. Alibart, S. Pleutin, D. Guérin, C. Novembre, S. Lenfant, K. Lmimouni, C. Gamrat, and D. Vuillaume, "An organic nanoparticle transistor behaving as a biological spiking synapse," *Adv. Funct. Mater.*, vol. 20, no. 2, pp. 330–337, Jan. 2010.

[6] E. J. Fuller, F. E. Gabaly, F. Léonard, S. Agarwal, S. J. Plimpton, R. B. Jacobs-Gedrim, C. D. James, M. J. Marinella, and A. A. Talin, "Li-ion synaptic transistor for low power analog computing," *Adv. Mater.*, vol. 29, no. 4, Jan. 2017, Art. no. 1604310.

[7] W. Wang, R. Wang, T. Shi, J. Wei, R. Cao, X. Zhao, Z. Wu, X. Zhang, J. Lu, H. Xu, Q. Li, Q. Liu, and M. Liu, "A self-rectification and quasi-linear analogue memristor for artificial neural networks," *IEEE Electron Device Lett.*, vol. 40, no. 9, pp. 1407–1410, Sep. 2019.

[8] N. K. Upadhyay, H. Jiang, Z. Wang, S. Asapu, Q. Xia, and J. Joshua Yang, "Emerging memory devices for neuromorphic computing," *Adv. Mater. Technol.*, vol. 4, no. 4, Apr. 2019, Art. no. 1800589.

[9] Z. Yu, A. M. Abdulghani, A. Zahid, H. Heidari, M. A. Imran, and Q. H. Abbasi, "An overview of neuromorphic computing for artificial intelligence enabled hardware-based Hopfield neural network," *IEEE Access*, vol. 8, pp. 67085–67099, 2020.

[10] X. Yan, Q. Zhao, A. P. Chen, J. Zhao, Z. Zhou, J. Wang, H. Wang, L. Zhang, X. Li, Z. Xiao, K. Wang, C. Qin, G. Wang, Y. Pei, H. Li, D. Ren, J. Chen, and Q. Liu, "Vacancy-induced synaptic behavior in 2D WS$_2$ nanosheet–based memristor for low-power neuromorphic computing," *Small*, vol. 15, no. 24, Jun. 2019, Art. no. 1901423.

[11] T.-Y. Wang, J.-L. Meng, M.-Y. Rao, Z.-Y. He, L. Chen, H. Zhu, Q.-Q. Sun, S.-J. Ding, W.-Z. Bao, P. Zhou, and D. W. Zhang, "Three-dimensional nanoscale flexible memristor networks with ultralow power for information transmission and processing application," *Nano Lett.*, vol. 20, no. 6, pp. 4111–4120, Jun. 2020.

[12] S. Yu, Y. Wu, R. Jeyasingh, D. Kuzum, and H.-S.-P. Wong, "An electronic synapse device based on metal oxide resistive switching memory for neuromorphic computation," *IEEE Trans. Electron Devices*, vol. 58, no. 8, pp. 2729–2737, Aug. 2011.

[13] M. Suri, O. Bichler, D. Querlioz, B. Traoré, O. Cueto, L. Perniola, V. Sousa, D. Vuillaume, C. Gamrat, and B. DeSalvo, "Physical aspects of low power synapses based on phase change memory devices," *J. Appl. Phys.*, vol. 112, no. 5, Sep. 2012, Art. no. 054904.

[14] Y. S. Lee, K. M. Kim, J. H. Lee, J. H. Choi, and S. W. Chung, "A high-performance processing-in-memory accelerator for inline data deduplication," in *Proc. IEEE 37th Int. Conf. Comput. Design (ICCD)*, Nov. 2019, pp. 515–523.

[15] R. Meade, S. Ardalan, M. Davenport, J. Fini, C. Sun, M. Wade, A. Wright-Gladstein, and C. Zhang, "TeraPHY: A high-density electronic-photonic chiplet for optical I/O from a multi-chip module," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2019, pp. 1–3.

[16] Q. Chen, M. Lin, Z. Wang, X. Zhao, Y. Cai, Q. Liu, Y. Fang, Y. Yang, M. He, and R. Huang, "Low power Parylene-Based memristors with a graphene barrier layer for flexible electronics applications," *Adv. Electron. Mater.*, vol. 5, no. 9, Sep. 2019, Art. no. 1800852.

[17] X. Zhao, Z. Wang, W. Li, S. Sun, H. Xu, P. Zhou, J. Xu, Y. Lin, and Y. Liu, "Photoassisted electroforming method for reliable low-power organic–inorganic perovskite memristors," *Adv. Funct. Mater.*, vol. 30, no. 17, Apr. 2020, Art. no. 1910151.

[18] K. Wang, L. Li, R. Zhao, J. Zhao, Z. Zhou, J. Wang, H. Wang, B. Tang, C. Lu, J. Lou, J. Chen, and X. Yan, "A pure 2H-MoS$_2$ nanosheet-based memristor with low power consumption and linear multilevel storage for artificial synapse emulator," *Adv. Electron. Mater.*, vol. 6, no. 3, Mar. 2020, Art. no. 1901342.

[19] N. Himmel, M. Ziegler, H. Mähne, S. Thiem, H. Winterfeld, and H. Kohlstedt, "Memristive device based on a depletion-type SONOS field effect transistor," *Semicond. Sci. Technol.*, vol. 32, no. 6, May 2017, Art. no. 06LT01.

[20] Z. Q. Wang, H. Y. Xu, X. H. Li, H. Yu, Y. C. Liu, and X. J. Zhu, "Synaptic learning and memory functions achieved using oxygen ion migration/diffusion in an amorphous InGaZnO memristor," *Adv. Funct. Mater.*, vol. 22, no. 13, pp. 2759–2765, Jul. 2012.

[21] X. Zhang, S. Liu, X. Zhao, F. Wu, Q. Wu, W. Wang, R. Cao, Y. Fang, H. Lv, S. Long, Q. Liu, and M. Liu, "Emulating short-term and long-term plasticity of bio-synapse based on Cu/a-Si/Pt memristor," *IEEE Electron Device Lett.*, vol. 38, no. 9, pp. 1208–1211, Sep. 2017.

[22] X. Yang, J. Yu, J. Zhao, Y. Chen, G. Gao, Y. Wang, Q. Sun, and Z. L. Wang, "Mechanoplastic tribotronic floating-gate neuromorphic transistor," *Adv. Funct. Mater.*, vol. 30, no. 34, Aug. 2020, Art. no. 2002506.

[23] H. Kino, T. Fukushima, and T. Tanaka, "Tunnel field-effect transistor charge-trapping memory with steep subthreshold slope and large memory Window," *Jpn. J. Appl. Phys.*, vol. 57, no. 4S, Mar. 2018, Art. no. 04FE07.

[24] N. Himmel, M. Ziegler, H. Mähne, S. Thiem, H. Winterfeld, and H. Kohlstedt, "Memristive device based on a depletion-type SONOS field effect transistor," *Semicond. Sci. Technol.*, vol. 32, no. 6, May 2017, Art. no. 06LT01.

[25] Y. Kang, J. Jang, D. Cha, and S. Lee, "Synaptic weight evolution and charge trapping mechanisms in a synaptic pass-transistor operation with a direct potential output," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 10, pp. 4728–4741, Oct. 2021.

[26] S. Lee and A. Nathan, "Subthreshold Schottky-barrier thin-film transistors with ultralow power and high intrinsic gain," *Science*, vol. 354, no. 6310, pp. 302–304, Oct. 2016.

[27] W. H. Hayt, J. E. Kemmerly, and S. M. Durbin, *Engineering Circuit Analysis*. New York, NY, USA: McGraw-Hill, 1978.

[28] S. Agarwal, S. J. Plimpton, I. Richter, A. H. Hsia, and D. R. Hughart. *CrossSim*. Accessed: Jun. 2018. [Online]. Available: http://crosssim.sandia.gov

[29] S. Agarwal, S. J. Plimpton, D. R. Hughart, A. H. Hsia, I. Richter, J. A. Cox, C. D. James, and M. J. Marinella, "Resistive memory device requirements for a neural algorithm accelerator," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 929–938.

[30] J. A. Cox, C. D. James, and J. B. Aimone, "A signal processing approach for cyber data classification with deep neural networks," *Proc. Comput. Sci.*, vol. 61, pp. 349–354, Nov. 2015.

[31] B. Razavi, *Design of Analog CMOS Integrated Circuits*. New York, NY, USA: McGraw-Hill, 2002.

[32] Y. LeCun. (1998). *The MNIST Database of Handwritten Digits*. [Online]. Available: http://yann.lecun.com/exdb/mnist/

[33] L. Schott, J. Rauber, M. Bethge, and W. Brendel, "Towards the first adversarially robust neural network model on MNIST," 2018, *arXiv:1805.09190*.

[34] R. B. Jacobs-Gedrim, S. Agarwal, K. E. Knisely, J. E. Stevens, M. S. van Heukelom, D. R. Hughart, J. Niroula, C. D. James, and M. J. Marinella, "Impact of linearity and write noise of analog resistive memory devices in a neural algorithm accelerator," in *Proc. IEEE Int. Conf. Rebooting Comput. (ICRC)*, Nov. 2017, pp. 1–10.

[35] S. Liao, M. W. K. Law, and A. C. S. Chung, "Dominant local binary patterns for texture classification," *IEEE Trans. Image Process.*, vol. 18, no. 5, pp. 1107–1118, May 2009.

[36] R. Wang and K. Wang, "A comparison study on the different SNR levels to the accuracy of two deep learning techniques in fault diagnosis of planetary gearbox," in *Proc. Asia–Pacific Int. Symp. Adv. Rel. Maintenance Modeling (APARM)*, Aug. 2020, pp. 1–6.

YEONSU KANG received the bachelor's degree from the Department of Electronics, Pusan National University, South Korea, in 2019, where she is currently pursuing the graduate degree in the combined master's and Ph.D. program. Her research interests include synaptic devices and physics for neuromorphic circuits and systems.

SOOBIN LEE currently pursuing the undergraduate degree with the Department of Electronics, Pusan National University, South Korea. Her research interests include neuromorphic devices and physics for neuromorphic circuits and systems.

DANYOUNG CHA received the bachelor's degree from the Department of Electronics, Pusan National University, South Korea, in 2019, where he is currently pursuing the graduate degree in the combined master's and Ph.D. program. His research interests include synaptic devices and physics for low-power operation of neuromorphic circuits and systems.

SUNGSIK LEE received the Ph.D. degree from University College London (UCL), London, U.K., in 2013. From 2013 to 2017, he was as a Research Associate at the University of Cambridge, Cambridge, U.K. He has published over 70 articles in the field, including one journal *Science* on the "most-off transistor" as the first author. His research interests include semiconductor devices and physics for futuristic electronics, such as neuromorphic circuits and systems. He is a winner of the IEEE EDS Ph.D. Student Fellowship 2011. Recently, he was awarded the Best Teaching Prize 2017 from the Korean Society for Engineering Education (KSEE), South Korea.

• • •