Ferroelectric materials for neuromorphic computing

Cite as: APL Mater. 7, 091109 (2019); https://doi.org/10.1063/1.5108562
Submitted: 30 April 2019 . Accepted: 05 September 2019 . Published Online: 19 September 2019

S. Oh, H. Hwang, and I. K. Yoo

COLLECTIONS

Paper published as part of the special topic on Emerging Materials in Neuromorphic Computing
Note: This paper is part of the Special Topic on Emerging Materials in Neuromorphic Computing.

This paper was selected as an Editor’s Pick

ARTICLES YOU MAY BE INTERESTED IN

Ferroelectricity in hafnium oxide thin films
Applied Physics Letters 99, 102903 (2011); https://doi.org/10.1063/1.3634052

Perspective: A review on memristive hardware for neuromorphic computation
Journal of Applied Physics 124, 151903 (2018); https://doi.org/10.1063/1.5037835

A critical review of recent progress on negative capacitance field-effect transistors
Applied Physics Letters 114, 090401 (2019); https://doi.org/10.1063/1.5092684
Ferroelectric materials for neuromorphic computing

ABSTRACT
Ferroelectric materials are promising candidates for synaptic weight elements in neural network hardware because of their nonvolatile multilevel memory effect. This feature is crucial for their use in mobile applications such as inference when vector matrix multiplication is performed during portable artificial intelligence service. In addition, the adaptive learning effect in ferroelectric polarization has gained considerable research attention for reducing the CMOS circuit overhead of an integrator and amplifier with an activation function. In spite of their potential for a weight and a neuron, material issues have been pointed out for commercialization in conjunction with CMOS processing and device structures. Herein, we review ferroelectric synaptic weights and neurons from the viewpoint of materials in relation to device operation, along with discussions and suggestions for improvement. Moreover, we discuss the reliability of HfO$_2$ as an emerging material and suggest methods to overcome the scaling issue of ferroelectrics.

INTRODUCTION
The use of ferroelectric thin films was recommended for the connection of elements according to the need of the development of integrated circuit devices that directly facilitate the use of artificial neural network (ANN) as analog circuits in 1989. The ferroelectric synaptic weight for a neural network was demonstrated by a single bit memory effect or a multilevel memory effect in the form of a ferroelectric capacitor, ferroelectric field effect transistor (FeFET), and ferroelectric tunnel junction (FTJ). Moreover, the ferroelectric weight was reported to be controlled by spike-timing-dependent plasticity (STDP). The ferroelectric neuron, which generates a pulse that activates a new input, was demonstrated on the basis of an adaptive learning function. Neurons (pre- and post-) connecting via a three-terminal FeFET made of a ferroelectric thin-film transistor (TFT) was studied. In addition to ceramic ferroelectrics, organic ferroelectric weights were also investigated. Even the ferroelectric liquid crystal was studied for the optical neural network in 1989. Ferroelectric weights are applied to both ANN and STDP with different operation. Typical wave forms to operate ferroelectric weight are well summarized in Ref. 16, where amplitude and pulse width modes are shown for ANN and prespike and postspike wave forms are shown for STDP.

The FeFET type is mainly used for STDP and various device types, for ANN. The taxonomy of ferroelectric synaptic weight for ANN, organized on the basis of the above reports, is listed in Table I. It is similar to that of ferroelectric memories because the memory can be directly applied to synaptic weight. The three-terminal structure, namely, a single Tr node for STDP, is well demonstrated in Refs. 5, 14–16 and 22. This article will mainly focus on ANN when it comes to ferroelectric weight.

In this article, we present several issues and limitations of ferroelectrics, mainly conventional ferroelectrics and HfO$_2$-based ferroelectrics, and clarify how ferroelectrics can be used in neuromorphic applications such as synaptic weights and neurons.
TABLE I. Taxonomy of ferroelectric synaptic weight for ANN.

| Weight cell structures | Weight cell nodes | References |
|------------------------|-------------------|------------|
| Two-terminal           |                   |            |
| 1R FTJ                 |                   | 13 and 17  |
| 1T-1R FTJ              |                   | 11         |
| 1T-1C Capacitor        |                   | 1 and 6    |
| Three-terminal         |                   |            |
| 1Tr NOR                |                   | 28         |
| 1Tr AND                |                   | 29 and 30  |
| 1Tr Inverter           |                   | 31         |
| 1Tr Single Tr          |                   | 32 and 33  |

FERROELECTRIC SWITCHING

The size of ferroelectric capacitors should be carefully considered when studying synaptic weight and neuron of ferroelectrics because ferroelectric switching behavior changes with decreasing ferroelectric size. Figure 1 describes the ferroelectric switching behavior depending on ferroelectric size and the input signal. Ishiwara investigated partial switching of PbZrTiO$_x$ (PZT) in detail and termed it adaptive learning. Ferroelectric switching is caused by ferroelectric domain switching. Ishibashi and Takagi interpreted the switching transients of various ferroelectric crystals based on Kolmogorov’s method. They considered the domain nucleation rate and developed a model based on the time-dependent switching current. According to their model, a pulse width less than that of the full-domain switching time gives rise to partial domain switching. Ishiwara reported on polarization vs pulse width and pulse counts with a specific pulse width through partial switching modeling.

The polarization level based on partial switching in ferroelectrics is determined by both pulse width and voltage according to the following polarization equation:

![Figure 1](image-url)
\[
P(t) = P_r \left[ 1 - 2 \exp \left( -\frac{t}{t_s} \exp \left( -\frac{E_a}{V/d_{FE}} \right)^n \right) \right],
\]
where \(P_r\) represents spontaneous polarization, \(E_a\) is the activation electric field, \(d_{FE}\) is the ferroelectric film thickness, \(n\) is a constant used to describe the nucleation mechanism, and \(t_s\) is the switching time. Ishiwara observed partial polarization when applied pulse width was shorter than threshold switching time \(t_{TH}\) [Fig. 1(a)] where the maximum switching current was measured. He applied a short pulse train to the PZT capacitor and obtained the curve of polarization and pulse count showing pulse-voltage dependency. Figure 1(c) shows the schematic curve of this relationship. To explain pulse-count dependence on partial switching, Ishiwara assumed that the pulse count is directly proportional to \(t\) in Eq. (1). Recently, Oh et al. and Kinder et al. demonstrated the multilevel states of HfO\(_2\) with respect to an incremental voltage mode and a pulse-width effect on polarization reversal in both HfZrO\(_x\) (HZO) and PZT, respectively. Equation (1) explains the incremental voltage dependency of polarization in Refs. 32 and 33 in conjunction with the schematic model shown in Fig. 1(e).

However, Eq. (1) is valid for relatively large sizes of a ferroelectric capacitor with many grains and domains. In the neuromorphic synapse application, highly scaled devices are required eventually. In a nanoscale ferroelectric capacitor, multilevel switching is difficult. Ishiwara observed partial polarization when applied pulse width was shorter than threshold switching time \(t_{TH}\) [Fig. 1(a)] where the maximum switching current was measured. He applied a short pulse train to the PZT capacitor and obtained the curve of polarization and pulse count showing pulse-voltage dependency. Figure 1(c) shows the schematic curve of this relationship. To explain pulse-count dependence on partial switching, Ishiwara assumed that the pulse count is directly proportional to \(t\) in Eq. (1). Recently, Oh et al. and Kinder et al. demonstrated the multilevel states of HfO\(_2\) with respect to an incremental voltage mode and a pulse-width effect on polarization reversal in both HfZrO\(_x\) (HZO) and PZT, respectively. Equation (1) explains the incremental voltage dependency of polarization in Refs. 32 and 33 in conjunction with the schematic model shown in Fig. 1(e).

However, Eq. (1) is valid for relatively large sizes of a ferroelectric capacitor with many grains and domains. In the neuromorphic synapse application, highly scaled devices are required eventually. Therefore, a nano-sized FeFET will have limitation for use in multilevel states because it has either single grain or single domain in which multilevel switching is difficult. Therefore, different physics should be considered in a nanoscale ferroelectric capacitor. Mulaosmanovic et al. studied accumulative polarization reversal and single-domain switching in nanoscale FeFETs with a small number of ferroelectric grains. They applied a nucleation-dominated switching model to the nanoscale HfO\(_2\) FeFET with a purely nucleation-limited model. Moreover, they studied switching time \(t_s\) in detail in conjunction with ferroelectric nucleation time \(t_0\) and time \(\Delta T_i\) between critical nucleus generation,

\[
t_s = \sum_{i=1}^{m} \Delta T_i, \tag{2}
\]

Therefore, \(m \geq 1\). They assumed the nucleation to be the Poisson process and proposed the probability density function of \(\Delta T_i\) as

\[
p_{\Delta T_i} = \lambda e^{-\lambda \Delta T_i}, \tag{3}
\]

where \(\lambda\) is the nucleation generation rate. Furthermore, they observed electrical excitation in a FeFET, followed by a sudden and complete switching, as schematically shown in Fig. 1(d). As shown, binary switching (“off” and “on” switching) appears in fewer pulse counts with an increase in the pulse voltage in the nanoscale ferroelectric capacitor. Mulaosmanovic et al. also demonstrated multilevels in a nanoscale FeFET by controlling the program and erasing signals by applying incremental voltage pulses with reset pulses for every step, as schematically shown in Fig. 1(f). This implies that the use of multilevels is possible by using multidomains with different coercive voltages, even in a nanoscale ferroelectric device.

The above-mentioned switching properties provide us with an integration guideline for ferroelectric weight. The multilevel weight of a certain sized ferroelectric capacitor can be realized by integrating 1T-1C and 1Tr cells. However, intermediate levels disappear in a nanoscale ferroelectric capacitor with binary weight. Therefore, a stacked structure, such as a cross-point structure, should be considered for developing multibit weights, as ferroelectrics are scaled down. A typical stacked cross-point structure for 1Tr node may be similar to the structure of a vertical NAND. The ferroelectric TFT is also another option for stack integration, in which an oxide semiconductor replaces Si. This is discussed in the emerging material section in more detail.

FERROELECTRIC WEIGHT

Ferroelectrics have been considered for synaptic weights because of their nonvolatile multilevel memory properties. Although a floating gate transistor also exhibits a multilevel storage effect, its operating voltage is higher than that of a ferroelectric transistor. Moreover, a ferroelectric transistor exhibits endurance for more than \(10^9\) switching cycles when a fatigue free electrode or a buffer layer for a specific ferroelectric material is integrated as an electrode.

The multilevel polarization states of ferroelectrics can be achieved by applying either an incremental or constant voltage with a pulse width that results in partial polarization reversal \({34,35}\) for a ferroelectric capacitor with a certain size and thickness. A ferroelectric capacitor has nonlinear multilevel polarization under both incremental voltage pulses and partial switching short pulses with a constant voltage. Furthermore, it is desirable for a ferroelectric capacitor to apply incremental and decremental pulses for multilevel synaptic weight in full range of polarization.

A ferroelectric synaptic weight comprises either three-terminal (1Tr) or two-terminal (FTI) weight nodes. A very thin ferroelectric capacitor may create an FTJ that produces a binary polarization, in general. The FTI cell forms a 1T-1R structure and comprises a two-terminal weight. Both 1Tr and 1T-1R nodes are operated in the conducting mode during weighting and multiplication. Figure 2 shows the operational scheme for both three- and two-terminal structures of a ferroelectric synaptic weight for ANN and STDP. For ANN, the weight is controlled by the gate that is independent of the input line in a three-terminal cell structure, whereas the weight is controlled by the input line in a two-terminal weight cell structure. The conductivity of the weight node increases (potentiation) and decreases (depression) based on pulse amplitude modulation or pulse width modulation. Furthermore, the relative timing between the prespike and postspike is converted to voltage height modulation across the synapse in STDP.

In a neural network, the input signal and weight are independently controlled. The input signal of the neural network typically comprises voltage pulses, and both input voltage and input pulse counts may vary depending on the input information. In addition, the weight is determined during learning and should be accessible randomly, directly, and independently. The dynamic range of the weight is determined by \(P_t\), of ferroelectrics, i.e., positive and negative \(P_t\). The weight resolution and multilevel of polarization are controlled in this dynamic range, and this limits the weight bit-width in the ferroelectric synapse. Therefore, ferroelectrics with higher \(P_t\) are desired. Nonlinear polarization against weighting (programming) voltage is also an issue. Thus, it is recommended that the information about weighting voltage and the corresponding polarization value should be stored separately to be used when weighting and unweighting. Figure 3 demonstrates how to control weight values.
randomly and directly. Although a narrow pulse width with incremental pulse voltage may be considered for controlling multilevel polarization, nonlinearity would still be observed, and pulse count information should be stored.

**1Tr for multilevel ferroelectric weight**

1Tr-type multilevel weight is useful for vector matrix multiplication (VMM) in ANN. VMM in ANN can be classified as a conductance- and capacitance-based multiplication. The relation $I = GV$ corresponds to a typical conductance-based multiplication, where $I$ is the output current, $G$ is the conductance, and $V$ is the input voltage. The capacitance-based multiplication adopts the relation $Q = CV$ for a capacitive weight in a neural network, where $Q$ is the output charge, $C$ is the capacitance, and $V$ is the input voltage. A ferroelectric gate transistor with a three-terminal structure is operated in the capacitive and conductive modes during weighting and matrix multiplication, respectively. This study mainly focused on conductance-based multiplication with respect to ferroelectric weight. Figure 4(a) shows a NOR-type ferroelectric weight in ANN that can perform VMM. The input lines are perpendicular to the output lines, and each cell conductance can be read by applying voltage pulses to the input lines simultaneously, by which output currents are measured simultaneously. Multilevel FeFETs can be recognized by reading channel conductance of each NOR cell. Therefore, a NOR structure with multilevel ferroelectric weight is an ideal hardware architecture for VMM.

**1Tr for binary ferroelectric weight**

The binary weight system is indispensable when a nanoscale ferroelectric device is integrated. However, binary weight may be applied even to a larger ferroelectric capacitor depending on an application such as the binary neural network. As ferroelectrics have a bistable state, bipolar voltage pulses at voltages higher than the coercive voltage are applied to achieve the binary weight. The positive and negative remanent polarizations ($P_r$) are designated as “1” and “0,” respectively.
FIG. 4. NOR-type (a) and AND-type (b) ferroelectric synaptic weights. The NOR weight has nonlinear multilevel weights. In contrast, the AND weight with a group of single bits gives linear multibit weights.

and “0,” respectively. Then, a single weight bit or binary weight per cell can be realized. A group of 1T1R binary-weight ferroelectric cells is required for achieving a multibit weight. A typical 1T1R group weight structure is similar to the AND-type ferroelectric weight, as shown in Fig. 4(b),\textsuperscript{2,29,31,42} which originates from the AND-type ferroelectric storage.\textsuperscript{43} Each FeFET represents a single-bit weight (binary weights “1” and “0”), and multilevels are recognized by reading as many as “on state” ferroelectric transistors.

The AND-type ferroelectric weight may guarantee a linear wide dynamic range of weights, but the wiring circuit becomes a critical issue in this case. The current level will also become limited to guarantee a practical dynamic range of weights. For example, when the weight bit is 16 bit and the current resolution is 1 nA, the system should be able to read from 1 nA through 66 μA/weight cell. Then, the number of weight cells will be limited by available power even for the input voltage pulses. Therefore, the channel current of a ferroelectric transistor should be minimized for AND-type binary synaptic weights.

1T1R for ferroelectric bi-weight

The most powerful property of FeFET is bistability, by which both excitatory and inhibitory weights can be achieved in a single cell. Figure 5 illustrates ferroelectric bi-weight cell operation, in which the cell structure consists of NMOS and PMOS. The weight cell with a ferroelectric inverter for both positive and negative weights was patented by Hynix.\textsuperscript{29} Nishitani et al. fabricated PZT TFT by depositing an oxide semiconductor on PZT.\textsuperscript{15} The biweight effect is enhanced when p-type and n-type materials are deposited on ferroelectrics. Figure 5(c) shows PZT TFT with transition metal dichalcogenides deposited on it where p-type WSe\textsubscript{2} and n-type MoS\textsubscript{2} were used as channel materials. The ferroelectric TFT is useful for stacked ferroelectric weight.

1T-1R or 1R for FTJ weight

Tsymbal and Kohlstedt reported about tunneling across a ferroelectric,\textsuperscript{44} although the basic principle was reported by Esaki et al. in IBM.\textsuperscript{45} Esaki et al. studied and formulated polar switch in 1971, following which Tsymbal and Kohlstedt established the FTJ concept.

FIG. 5. Biweight of ferroelectrics for ANN. Multilevel biweight in ferroelectric synaptic weight with a pair of (a) NMOS and (b) PMOS ferroelectric transistors. Positive output current and negative output current are designated by the current direction. (c) Ferroelectric TFT for biweight n-type and p-type weights is determined by semiconducting materials on ferroelectrics for TFT. Partially adapted from Patent KR 10-2018-0035251.
FTJ is a nondestructive readout node as that of the resistive change memory. Figure 6 represents a simple description of the binary FTJ weight principle. A ferroelectric thin film should be thin enough to allow tunneling and thick enough to guarantee ferroelectric polarization switch. This contradictory requirement makes it difficult for FTJ to have multilevel weight. In addition, the polycrystalline ferroelectric thin film in FTJ induces high leakage current, and its practical device applications are thus difficult. Therefore, some insulating layer is integrated with FTJ to improve its performance.\textsuperscript{46,47}

The tunneling model of the multilayer FTJ is illustrated in Fig. 6(b). Chanthbouala et al.\textsuperscript{48} suggested new applications of FTJ as the hardware basis of neuromorphic computational architectures. The electrode effect on tunneling resistance in BaTiO\textsubscript{3} (BTO)\textsuperscript{49} and ion migration in PbZr\textsubscript{0.52}Ti\textsubscript{0.48}O\textsubscript{3}\textsuperscript{50} was studied. Neuromorphic modeling and simulation were reported for BTO FTJ in the STDP circuit.\textsuperscript{11,13–17,22,23} Unsupervised STDP learning was simulated for multilevel BiFeO\textsubscript{3} (BFO) FTJ.\textsuperscript{17} Yamaguchi \textit{et al.} showed that the HfO\textsubscript{2} FTJ can achieve $10^3$ cycles by testing its reliability on the basis of time-dependent dielectric breakdown (TDDB) and stress-induced leakage current (SILC).\textsuperscript{52} IR FT was used for pattern recognition with off-chip learning by applying conventional voltage pulses of potentiation and depression\textsuperscript{-1} and spikes for spiking neural networks (SNNs).\textsuperscript{17}

**1T-1C for ferroelectric weight**

The 1T-1C ferroelectric synaptic weight in ANN was suggested by Clark, Grondin, and Dey in 1989.\textsuperscript{2} They reported on both binary and multilevel weights, as shown in Fig. 7, and proposed incremental voltage pulses and partial switching with a short pulse width for achieving multilevel weights. The nondestructive readout (NDRO) technique has been applied to the multilevel weight by defining conductance as a weight.\textsuperscript{2} The NDRO mode is the same operational mode as that in the 1Tr ferroelectric weight. Obradovic \textit{et al.} fabricated a weight cell by applying this scheme to demonstrate multilevel FeFET, in which a ferroelectric HZO capacitor is connected to the FET gate.\textsuperscript{5} Clark \textit{et al.} demonstrated partial switching in a PZT capacitor by varying the pulse width and pulse voltage to demonstrate the analog programming behavior of the synapse circuit.\textsuperscript{2}

**Ferroelectric weight in STDP**

As mentioned earlier, the STDP operation differs from that of ANN, in which relative timing between the prespikes and postspikes is converted to voltage height. Therefore, many studies have been conducted on ferroelectric synaptic weights in STDP.\textsuperscript{11,13–17,22,23} Nishitani \textit{et al.} studied the ferroelectric weight by applying STDP to a ferroelectric TFT with a structure contradictory to that of a ferroelectric transistor.\textsuperscript{14} Boynt \textit{et al.} studied the ferroelectric multilevel weight with respect to STDP,\textsuperscript{17} and Mulaosmanovic \textit{et al.} obtained STDP-like curves from ferroelectric HfO\textsubscript{2}.\textsuperscript{16} Nishitani, Kaneko, and Ueda studied supervised learning by using STDP of ferroelectric TFT,\textsuperscript{15} and Wang \textit{et al.} demonstrated ferroelectric spiking neurons for unsupervised clustering.\textsuperscript{2} Furthermore, Chen \textit{et al.} studied STDP and neural networks by using HZO FTJ,\textsuperscript{13} and Wang \textit{et al.} designed STDP for supervised learning by using 1T-1R BaTiO\textsubscript{3} FTJ.\textsuperscript{11} The neural network with the ferroelectric weight controlled by STDP is represented in Fig. 2.
Both 1Tr with a ferroelectric capacitor and 1T-1R with a FTJ are conductance-based synaptic weights. It is important to verify whether they can be used in on-chip learning for mobile application. Figure 8 describes schematics of two- and three-terminal synapses for the STDP function. The 1T-1R weight is simple for neural networks; however, signal processing and learning cannot be simultaneously performed as reported by Nishitani et al. As prespikes and postspikes are applied to the learning process, a feedback signal flow is required during learning. Such a signal can be fed to the gate of the FeFET (three-terminal) during learning but cannot be fed to a two-terminal synaptic weight. Thus, the 1Tr ferroelectric weight is indispensable when on-chip learning is required.

**Summary of ferroelectric weight**

In summary, in ANN application, ferroelectrics are competitive as a synaptic weight, especially the 1T-1C type, because of their nonvolatile multilevel property, as mentioned earlier. Even though they may not show perfect linearity in both potentiation and depression, their polarization level is still controllable because the curve shape of potentiation and depression is symmetric. Multilevel weights may reduce the circuit overhead, but the precise control of the polarization level is challenging. As data scattering, level resolution, and dynamic range of weight bits should be simultaneously considered, multilevels are generally limited to 16–32 levels. This implies that the dynamic range or weight bit-width of the multilevel is limited to 5 bits (32 levels) at most. The synaptic weight in ANN may require a wider bit-width. In addition, when considering level-to-level resolution, the ferroelectric weight may not be able to guarantee multilevels higher than 16 levels (4 bits) with decreasing device size.

A 1Tr-type synapse (FeFET or TFT-type FeFET) that is used in both ANN and STDP has the unique property of bistability, which is useful for excitatory and inhibitory functions. However, it is noted that a pair of n- and p-type channel materials are required for the biweight function, as demonstrated by Nishitani for a TFT-type FeFET combined with an oxide semiconductor such as ZnO. This function is emphasized for STDP applications, where both positive and negative charge differences are defined. In addition, weighting (programming) is performed between the gate and the body of FeFET and reading is performed between the source and the drain, making the network circuit more complicated than that of a two-terminal weight. Above all, reliability, such as endurance, retention, and imprint, of ferroelectrics should be improved in advance for commercial applications.

**FERROELECTRIC NEURON**

A ferroelectric neuron is defined as the CMOS neuron that contains a ferroelectric capacitor(s) or ferroelectric transistor(s). Ferroelectrics can be used in a neuron because they exhibit an integration effect in the course of partial switching. That is, the polarization state increases without discharging, because of nonvolatility, under consecutive input pulses. Moreover, a ferroelectric neuron has the same structure as that of a 1Tr weight, and it can be used with either a leaky integrator or signal generator. The activation function in association with a ferroelectric transistor can be achieved by devices such as unijunction transistors and oscillators. Ishiwara et al. proposed a ferroelectric neuron by utilizing ferroelectrics with pulse oscillation circuits. Wang et al. demonstrated ferroelectric spiking neurons, as shown in Fig. 9(a), by connecting a ferroelectric transistor to a CMOS transistor. The ferroelectric transistor consists of a ferroelectric capacitor on the gate of the CMOS transistor (FeFET). CMOS leaky integrators are connected to the gate of the FeFET and CMOS transistor. This combination results in an activation function. Nishitani et al. connected a ferroelectric TFT, which generates prepulses, to a CMOS integrator to generate postpulses. This combination generates leaky integrate-and-fire, as shown in Fig. 9(b). A recent study confirmed that a single ferroelectric transistor [fully depleted silicon-on-insulator (FDSOI) FeFET] exhibits an integrate-and-fire behavior; the use of the integration capabilities (accumulative switching) of a nanoscale FeFET was proposed to replace the bulky integration capacitor that is normally used in neuron circuits. The activation function can be achieved by forming a circuit with a bipolar resistive switch, as shown in Fig. 9(c). It switches to the “off” state when the integrate-and-fire current flows through its “on” state resulting in voltage out. A CMOS neuron consists of an integrator and an amplifier and often comprises a large capacitor to emulate the integration of the current.
FIG. 9. Ferroelectric neurons: (a) using the negative weight effect of the ferroelectric transistor and (b) using the STDP effect of the ferroelectric transistor. A ferroelectric capacitor with (c) integration-and-fire behavior was observed in an ultrascaled fully depleted silicon-on-insulator (FDSOI) FeFET. This effect can be used for the activation function by utilizing a current driven bipolar resistive switch that gives rise to activation voltage.

sum from each output of neural networks. Although a ferroelectric can reduce the size of a capacitor to some extent, the scaling limitation still exists.

ISSUES

Ferroelectrics such as PbZrO$_x$ (PZT), SrBi$_2$Ta$_2$O$_9$ (SBT), BaTiO$_3$ (BTO), and BiFeO$_3$ (BFO) have been investigated for the neuromorphic application. Their main issues are scalability, thickness, CMOS compatibility, and reliability including leakage current. It is relatively easier to achieve the multilevel of polarization for a large-size ferroelectric capacitor than control the polarization for a small ferroelectric capacitor because shorter pulse width and smaller incremental voltage are required for the latter.

It is desirable for the thickness of ferroelectrics in FeFET to be minimized because the thickness of the gate stack should be minimized in a vertical NAND structure. This is the reason why the ferroelectric HfO$_2$ has attracted attention because it demonstrates ferroelectricity at the thickness of 7–10 nm and gate length of 20 nm.²⁴

The ferroelectric transistor may be integrated in the form of metal-ferroelectric-metal-insulator-silicon (MFMIS), metal-ferroelectric-insulator-silicon (MFIIS), and metal-ferroelectric-silicon (MFS). The MFS structure is desired when considering a voltage drop at the oxide gate to minimize device operational voltage. However, the direct deposition of the ferroelectric thin films on Si substrates is difficult because of the chemical reaction between Si and ferroelectric materials.¹⁵ The direct deposition of HfO$_2$ on Si
also shows an unstable interface property between HfO$_2$ and Si, although HfO$_2$ is more stable on Si than other ferroelectric materials. In addition, a ferroelectric HfO$_2$ needs a large operation voltage for ferroelectric switching, thus inducing deep charge traps, although HfO$_2$ has been used as a gate dielectric in MOS devices. Thus, the use of a semiconductor-ferroelectric-metal (SFM) structure has been attempted, which is based on the model proposed by Evans.

Inferences, such as pattern recognition, can be performed in real time when ferroelectric weight is fabricated in a neural network because vector matrix multiplication becomes possible. However, if artificial-intelligence service requires high weight bit-width, stacked weights should be integrated to reduce the neuromorphic chip size. For this purpose, the SFM integration is feasible for this purpose. Therefore, a combination of an oxide semiconductor, thin ferroelectric, and fatigue-free electrode should be intensively investigated, and a low-loss ferroelectric film with no leakage path should be selected.

**HAFNIUM OXIDE AS AN EMERGING MATERIAL**

Ferroelectric HfO$_2$ was reported in 2011 by Boscke et al. and has attracted attention because it shows the ferroelectric property at a thickness less than 10 nm, which is compatible with the vertical NAND flash process. We reported the possibility of the neuromorphic synapse application using ferroelectric HZO. The TiN/10-nm-thick HZO/TiN capacitor was evaluated under various pulse schemes, as shown in Fig. 10. It was confirmed that scheme C with incremental voltage is acceptable for achieving HZO with multilevels.

The drain current was simulated to convert the multilevel $P_r$ into the channel current as a function of gate voltage by assuming a MFISFET. Figure 11(c) shows that the multilevel currents corresponding to each polarization state exhibit a symmetric potentiation and depression curve. Figure 11(d) shows that nonlinearity parameter $\alpha$ for potentiation and depression curves, in which scheme C gives a symmetric linear region.

The ferroelectric HfO$_2$ family transistors were investigated to confirm the presence of multilevel weights in them, as shown in Table II. Some authors focused on multilevel memory.

Ferroelectrics lose their multilevel weight effect as their size reduces as discussed earlier. When the signal control technique in Fig. 10(f) is not so practical in nanoscale ferroelectric capacitors for multilevel weights, the AND-type ferroelectric weight using the binary ferroelectric switch can be considered. But, a group of single ferroelectric cells will occupy a large area. Therefore, the stacked AND weights were proposed to reduce chip size. Figures 12(a) and 12(b) illustrate vertically stacked AND-type and horizontally stacked ferroelectrics, respectively. In the stacked AND weight group, oxide semiconducting materials replace Si in order to reduce processing cost and avoid the reaction between Si and ferroelectric materials.

Figure 12 does not confine ferroelectric material to HfO$_2$. PZT combined with fatigue proof electrode SRO may be fabricated in the form of either vertical AND or horizontal AND as a matter of course. 3D integration of HfO$_2$ has been demonstrated by IMEC for vertical processing that may be applied to a vertical NAND device. Ferroelectric HfO$_2$ based FeFET has several critical issues due to the interfacial layer (IL), although it is very attractive for emerging memory. The large voltage drop at the IL, which has a low dielectric constant increases operation voltage for switching. In addition, such a large voltage drop at the IL causes not only partial switching in the ferroelectric layer, not full switching, but also increases charge trapping, resulting in the reduction of memory window. Ali et al. reduced the operation voltage by adopting a SiON IL, which has a higher dielectric constant than SiO$_2$. In addition, it showed a large memory window because full switching is possible at low operation voltage and less charge trapping. Ni et al. also reported that
FIG. 11. (a) Simulation results of $I_d$-$V_g$ of each polarization state for scheme C and (b) the simulation parameters. (c) 32 level potentiation and depression curve at $V_{read} = 0.5$ V by simulation. (d) Change of nonlinearity parameter $\alpha$ for each pulse scheme for potentiation and depression. Reproduced with permission from Oh et al., IEEE Electron Device Lett. 38(6), 732–735 (2017). Copyright 2017 IEEE.

TABLE II. HfO$_2$ family FET and their multilevels. P: potentiation, D: depression.

| Structure | Channel materials | Multilevels | Pulse scheme | $G_{\text{max}}/G_{\text{min}}$ | Operation voltage (P: D: | Pulse width | Nonlinearity | References |
|-----------|-------------------|-------------|--------------|-------------------------------|--------------------------|-------------|-------------|------------|
| FET       | Si:HfO$_2$        | ~10 to 12 levels | Incremental | N/A | P: 3–5 V D: –2 to –4 V | 1 $\mu$s | N/A | 16 |
| FET       | HZO               | 32 levels | Incremental | 45 | P: 2.85–4.45 V D: –2.1 to –3.8 V | 75 ns | 1.22/–1.75 | 7 |
| FinFET    | HZO               | >32 levels | Identical | 4.98 | P: 3.7 V D: –3.2 V | 100 $\mu$s | 1.58/–7.57 | 9 |
| Nanowire  | HZO               | >256 levels | Identical | ~200 | ±5 V | 50 ns | 1.75/1.46 | 63 |
| TFT       | HZO IGZO          | >32 levels | Incremental | 14.4 | P: 2.7–4.3 V D: –2 to –3.6 V | 10 ms | –0.8028/–0.6979 | 10 |

FIG. 12. Stacked AND-type ferroelectric multibit weight. (a) Vertical cross-point AND stack that is compatible with vertical NAND flash processing. Partially adapted from Ref. 42. (b) Horizontal cross-point AND stack that is compatible with cross-point storage processing. Partially adapted from Ref. 59.
TABLE III. Strategies for endurance in HfO$_2$ based ferroelectric devices.

| Approach                        | Structure             | Investigation                      | References |
|---------------------------------|-----------------------|------------------------------------|------------|
| Modulation of polarization axis | MFIS FET              | Trade-off between the spontaneous polarization and switching voltage | 67         |
| Subloop operation               | MFIS FET              | Reduction of the memory window     | 67         |
| Capacitive divider              | MFMIS FET             | Complexity of the gate stack       | 67         |
| Channel modulation (SiGe)       | MFIS FET              | SiGe channel, not Si               | 77         |
| Heating pulse scheme            | MFIS FET              | Complexity of the pulse scheme     | 78         |

charge trapping can be minimized by modulating the pulse scheme (e.g., time delay, pulse width), resulting in an increase in memory window.

The retention and endurance of ferroelectric HfO$_2$ based FeFETs have been critical issues, although those of the MFM structure have been researched vigorously$^{69-71}$ and improved by NH$_3$ interface treatment. In addition, the MFS structure is desired in order to reduce operational voltage. Therefore, retention loss and fatigue were studied for ferroelectric HfO$_2$ on Si. In the case of retention, Gong and Ma reported that ferroelectric HfO$_2$ is more suitable than conventional ferroelectric materials because of its high coercive field. In addition, Ali et al. extrapolated that a 0.5-V memory window can be maintained after 10 years at 150°C by adopting a SiON interface, which reduces the depolarization field. In the case of endurance, IL degradation by charge trapping and trap generation induced by repetitive bipolar pulses cause endurance degradation. In order to improve the endurance property, many studies have been researched. Muller et al. suggested several strategies which can reduce voltage drop at IL during cycling. First, they suggested the modulation of the polarization axis to reduce the effect of spontaneous polarization. Ferroelectric forms large spontaneous polarization when applied voltage, resulting in a large voltage drop at the IL. Therefore, the voltage drop in the IL can be reduced by tilting the polarization axis with respect to the surface normal of the channel. However, it has a trade-off between the spontaneous polarization effect and switching voltage because the tilted polarization axis needs large switching voltage, which leads to the increase in operation voltage. Second, they suggested the use of subloop operation. The endurance property can be enhanced by using the ferroelectric subloop which needs low operation voltage, although the memory window is reduced. In addition, they suggested that operation voltage can be reduced by adopting an IL with high permittivity. Third, they suggested the capacitive divider by modulating the area ratio of the ferroelectric layer and dielectric ILs. However,
to control the area of ferroelectric and dielectric ILs separately, the MFMIS FET gate stack should be adopted.\textsuperscript{37} Chen et al. reported 10\textsuperscript{9} cycles in the MFIS stack by adopting a SiGe channel which forms a thinner suboxide IL with better quality.\textsuperscript{77} Furthermore, Mulaomanivic et al. reported recovery of the ferroelectric memory of FeFET after endurance failure by applying forward bias.\textsuperscript{78} The local Joule heating is generated by the forward bias, which heals out the damage at the IL induced by cycling.\textsuperscript{79} Therefore, they suggested optimizing the pulse scheme by adopting an alternative heating pulse.\textsuperscript{78}

Although they are excellent strategies for endurance, they have some investigations which have to be further researched and enhanced (see Table III). In addition, the research for improving the IL by direct treatment is rarely reported, although the IL is very important for endurance in the MFIS structure. We studied the effects of high pressure hydrogen annealing (HPHA), which improves the IL quality\textsuperscript{79} on the MFIS structure.\textsuperscript{68}

A 15-nm-thick poly crystalline Al doped HfO\textsubscript{2} thin film was deposited on heavily p-doped Si.\textsuperscript{80} It was found that Si reacts with HfO\textsubscript{2} to achieve the SiO\textsubscript{2} layer, which gives rise to domain imprinting (locking or pinning) and high coercive voltage, as shown in Fig. 13(b). The increase in coercive voltage is due to the voltage drop at the SiO\textsubscript{2} insulating layer. In addition, the resultant MFIS (W-HfO\textsubscript{2}-SiO\textsubscript{2}-p\textsuperscript{+}Si) capacitor shows poor endurance as compared to that of the MFM (W-HfO\textsubscript{2}-W) capacitor, as illustrated in Fig. 13(c).

High pressure hydrogen annealing (HPHA) was performed on the MFIS capacitor at various temperatures in order to improve the capacitor properties. Figure 14 shows that endurance was improved by increasing HPHA temperature, particularly at 100 kHz. A low leakage current was maintained up to 10\textsuperscript{7} cycles at 300°C, which led to endurance up to 10\textsuperscript{9} cycles.

**SUMMARY**

The driving force of ferroelectric materials for neuromorphic applications is their bistable memory effect, but their main drawback is the reliability of the Si process. The SFM structure, instead of MFS, may be one of the solutions to improve the reliability of ferroelectric HfO\textsubscript{2} family transistors for neuromorphic applications. SFM structure is also useful for vertical and horizontal stacked ferroelectric weights. The use of semiconducting materials such as transition metal dichalcogenide (TMD) materials and oxide semiconductors can be used for SFM structure with a potential for improving reliability such as imprinting. However, the combination of the HfO\textsubscript{2} family and a fatigue-free electrode needs to be discovered.

**ACKNOWLEDGMENTS**

The authors would like to thank reviewers for their sincere advice and comments on this article. This work was supported by the National Research Foundation of Korea funded by the Korea government (MSIT), Grant No. NRF-2018R1A3B1052693.

**REFERENCES**

1. T. L. Clark, R. O. Grondin, and S. K. Dey, in First IEEE International Conference on Artificial Neural Networks (IEEE, 1989), pp. 47–51.
2. T. L. Clark, R. O. Grondin, and S. K. Dey, in Proceedings of 11th Annual International Phoenix Conference on Computers and Communications, 1992.
3. H. Ishiwara, Jpn. J. Appl. Phys. 32, 442–446 (1993).
4. K. Nakano, K. Nishitani, M. Ueda, A. Tsujimura, in Symposium on VLSI Technology (IEEE, 2013), pp. T238–T239.
5. B. Obradovic, T. Rakshit, R. Hatcher, J. Kittil, R. Sengupta, J. G. Hong, and M. S. Rodder, IEEE J. Electron Devices Soc. 6, 438–448 (2018).
6. M. Jerry, Y. Chen, J. Zhang, P. Sharma, K. Ni, S. Yu, and S. Datta, in International Electron Devices Meeting (IEDM) (IEEE, 2017), pp. 6.2.1–6.2.4.
7. M. Jerry, S. Dutta, A. Kazemi, K. Ni, I. Zhang, P. Y. Chen, P. Sharma, S. Yu, K. S. Hu, M. Niwmiere, and S. Datta, J. Phys. D: Appl. Phys. 51, 434001 (2018).
8. M. Seo, M. H. Kang, S. B. Jeon, H. Bae, J. Hur, B. C. Jang, S. Yun, S. Cho, W. K. Kim, M. S. Kim, K. M. Hwang, S. Hong, S. Y. Choi, and J. K. Choi, IEEE Electron Device Lett. 39(9), 1445–1448 (2018).
9. M. J. Kim and S. Lee, Nano Lett. 19, 2044–2050 (2019).
10. T. Wang, W. Zhao, W. Kang, Y. Zhang, J. O. Klein, and C. Chappert, in International Joint Conference on Neural Networks (IEEE, 2014), pp. 29–34.
11. R. Guo, Y. Zhou, L. Wu, Z. Wang, Z. Lin, X. Yan, W. Lin, H. Wang, H. Y. Yoong, S. Chen, Ariando, T. Venkatesan, J. Wang, G. M. Chow, A. Gruevman, X. Miao, Y. Zhu, and J. Chen, ACS Appl. Mater. Interfaces 10, 12862–12869 (2018).
12. L. Chen, T. Y. Wang, Y. W. Dai, M. Y. Cha, H. Zhu, Q. Q. Sun, J. S. Ding, P. Zhou, L. Chua, and D. W. Zhang, Nanoscale 10, 15826–15833 (2018).
13. Y. Nishitani, Y. Kaneko, M. Ueda, E. Fujii, and A. Tsujimura, Jpn. J. Appl. Phys. 52, 04CE06 (2013).
14. Y. Nishitani, Y. Kaneko, and M. Ueda, IEEE Trans. Neural Networks Learn. Syst. 26(12), 2999–3008 (2015).
15. H. Mulaomanivic, J. Ocker, S. Müller, M. Noack, J. Müller, P. Polakowski, T. Mikolajick, and S. Sleszacek, in Symposium on VLSI Technology (IEEE, 2017), pp. T176–T177.
75. N. Gong and T. P. Ma, IEEE Electron Device Lett. 39(1), 15–18 (2018).
76. E. Yurchuk, J. Muller, S. Muller, J. Paul, M. Pasic, R. Bentum, U. Schroeder, and T. Mikolajick, IEEE Trans. Electron Devices 63(9), 3501–3507 (2016).
77. K. Y. Chen, Y. H. Huang, R. W. Kao, Y. X. Lin, and Y. H. Wu, in Symposium on VLSI Technology (IEEE, 2018), pp. 119–120.

78. H. Mulaosmanovic, E. T. Breyer, T. Mikolajick, and S. Slesazeck, IEEE Electron Device Lett. 40(2), 216 (2019).
79. H. Park, M. S. Rahman, M. Chang, B. H. Lee, R. Choi, C. D. Young, and H. Hwang, IEEE Electron Device Lett. 26(10), 725–727 (2005).
80. S. Oh, J. Song, I. K. Yoo, and H. Hwang, IEEE Electron Device Lett. 40(7), 1092–1095 (2019).