A FeFET with a novel MFMFIS gate stack: towards energy-efficient and ultrafast NVMs for neuromorphic computing

Tarek Ali¹, Konstantin Mertens¹, Kati Kühnel¹, Matthias Rudolph¹, Sebastian Oehler¹, David Lehninger¹, Franz Müller¹, Ricardo Revello¹, Raik Hoffmann¹, Katrin Zimmermann¹, Thomas Kämpfe¹, Malte Czernohorsky¹, Konrad Seidel¹, Jan Van Houdt²,³ and Lukas M Eng⁴,⁵

¹ Fraunhofer IPMS - Center Nanoelectronics Technologies An der Bartlake 5, D-01109 Dresden, Germany
² Imec, Leuven B-3001, Belgium
³ Department of Physics and Astronomy Katholieke Universiteit Leuven, Leuven B-3000, Belgium
⁴ Institut für Angewandte Physik, Technische Universität Dresden, Nöthnitzer Str. 61, D-01187 Dresden, Germany
⁵ Center of Excellence - Complexity and Topology in Quantum Matter ct.qmat, TU Dresden, D-01062 Dresden, Germany

E-mail: tarek.ali@ipms.fraunhofer.de

Received 14 March 2021, revised 12 June 2021
Accepted for publication 14 July 2021
Published 29 July 2021

Abstract
The discovery of ferroelectricity in the fluorite structure based hafnium oxide (HfO₂) material sparked major efforts for reviving the ferroelectric field effect transistor (FeFET) memory concept. A novel metal-ferroelectric-metal-ferroelectric-insulator-semiconductor (MFMFIS) FeFET memory is reported based on dual ferroelectric integration as an MFM and MFIS in a single gate stack using Si-doped Hafnium oxide (HSO) ferroelectric (FE) material. The MFMFIS top and bottom electrode contacts, dual HSO based ferroelectric layers, and tailored MFM to MFIS area ratio (AR-TB) provide a flexible stack structure tuning for improving the FeFET performance. The AR-TB tuning shows a tradeoff between the MFM voltage increase and the weaker FET Si channel inversion, particularly notable in the drain saturation current \( I_{DS} \) when the AR-TB ratio decreases. Dual HSO ferroelectric layer integration enables a maximized memory window (MW) and dynamic control of its size by tuning the MFM to MFIS switching contribution through the AR-TB change. The stack structure control via the AR-TB tuning shows further merits in terms of a low voltage switching for a saturated MW size, an extremely linear at wide dynamic range of the current update, as well as high symmetry in the long term synaptic potentiation and depression. The MFMFIS stack reliability is reported in terms of the switching variability, temperature dependence, endurance, and retention. The MFMFIS concept is thoroughly discussed revealing profound insights on the optimal MFMFIS stack structure control for enhancing the FeFET memory performance.

Supplementary material for this article is available online

Keywords: ferroelectric, hafnium oxide, MFMFIS, FeFET, synaptic device, neuromorphic

(Some figures may appear in colour only in the online journal)
1. Introduction

The emerging hafnium oxide based ferroelectric FET (FeFET) has brought increasing interest for a CMOS compatible embedded memory co-integration and utilization in neuromorphic based computing architectures [1]. The massive parallelism and efficiency of the human brain both in computing and power consumption inspired for non-sequential, probabilistic, and efficient computing paradigms [2]. The FeFET technology development is accelerated by the strong merits such as low voltage operation compared to the embedded Flash technology and CMOS compatibility of the ferroelectric (FE) materials based on the fluorite structure hafnium oxide [3, 4].

At the stack structure level, the FeFET is available in different types; each can define unique underlying physics and operation principles. The simplest and de-facto structure of the FeFET realization, even to the smallest nodes [5, 6], is known as the metal-ferroelectric-insulator-semiconductor (MFIS) stack [7]. The interfacial layer growth at the Si interface has a critical dependence on the thermal budget [8] and leaves a limited scope for interface engineering, ultimately via the realization of a high permittivity interfacial layer growth [9]. Thereby, the MFIS based gate stack structure is limited in reliability when compared to the integration of a more symmetric MFM capacitor structure [10].

An alternative stack structure is the MFMIS, this concept combines an MFM capacitor integrated on top of a conventional MIS FE/FET for enhanced reliability and improved FeFET performance as recently reported [11]. In particular, the tuning knob of ferroelectric to insulator capacitance ratio enables a controlled voltage distribution of the two gate stack capacitors (MFM and MIS) by careful of engineering the ferroelectric-to-insulator area ratio. In turn, the lower field across a large MIS area and the large MIS capacitance, at first glance provide a low power operation at improved endurance and retention reliability [12]. Initial conception of the MFMIS structure for conventional perovskite based ferroelectric materials aimed at the pressing issue of improved memory retention [13]. At a limited scalability of the perovskite based ferroelectric film thickness, a single ferroelectric layer integration as MFMIS remained a de-facto practical choice for this concept. Today, the pronounced ferroelectric properties for ultra-thin film thicknesses (∼10 nm) of the hafnium oxide material enables versatility beyond standard MFMIS structures. Hereby, an accelerated performance and more opportunities can be achieved for the FeFET through the stack structure control where an MIS based FET structure is replaced by an MFIS one. The bottom stack uses an MFIS area larger than the MFM one, this has the merit of providing a possible knob for compensating potential threshold voltage variability at scaled MFM. For saturated switching of the full MFMFIS stack, higher MW and lower variability are expected compared to an MFMIS stack.

In this work, scaled FeFETs are reported that are based on a combination of MFM with MFIS structures, as integrated within one MFMFIS stack. The stack structure in the form of MFM to MFIS area ratio tuning and dual ferroelectric layer operation in a single stack, are explored for optimal FeFET operation. The analog modulation of the FeFET current via partial ferroelectric polarization make these devices promising candidates for emulating the biological synapse function [14]. Therefore, the powerful MFMFIS stack-structure tuning is explored for optimized synaptic device to achieve controlled linearity and symmetry of the synaptic weight update. The MFMFIS concept was recently reported in [15], this paper aims to highlight the merits of this novel concept with detailed insight explanation.

2. Experimental details

2.1. The MFMFIS-based FeFET integration

The MFMFIS-based FeFET concept as schematically depicted in figure 1(a) is fabricated on 300 nm Si wafers using a gate first integration scheme. The process scheme in figure 1(c) conceptually shows a two-step of selective stack patterning to define the MFM to MFIS capacitor area ratio. Initially, an optimized SiON interface layer with a thickness of ∼1.8 nm is thermally grown by a rapid thermal nitridation process using ammonia (NH₃) ambient followed by a rapid thermal oxidation process in a N₂ and O₂ gas mixture as reported earlier [12]. As a next step, a 10 nm Si:H₂O₂ (HSO) layer is ALD deposited as the bottom ferroelectric layer using (SiCl₄:HCl₂) precursors and H₂O as the oxidizer. A bottom electrode (BE) of 35 nm TiN is deposited by TiCl₄/NH₃ in a chemical vapor deposition (CVD) process, thereby constructing the layer structure of the bottom MFIS stack.

An oxide layer at a thickness of 100 nm is deposited in a CVD process as an etch and self-aligned S/D implant hard mask (HM). A first step patterning is used to structure the MFIS stack where the TiN BE is selectively etched by the standard clean 1 (SC1) chemical etching followed by a reactive ion etching etch of the bottom HSO layer at elevated temperature using chlorine based etch chemistry. This patterning step defines the bottom MFIS area. A 10 nm SiO₂ implant scatter liner is ALD deposited and followed by a phosphorous doped ion implantation through the oxide HM. A selective etch of the HM by DHF is used for non-invasive expose of the TiN BE in the patterned stack. Hereby, a second 10 nm HSO layer is deposited as the top ferroelectric layer followed by a 10 nm CVD deposited TiN as a capping metal to form together with the BE an MFM configuration. The top and bottom ferroelectric layers are deposited in 10 nm film thickness at optimal ferroelectric properties (1:16) of (Si:H₂O₂) cycle ratio [16]. This is followed by the CVD deposition of a 100 nm of highly doped amorphous Si (aSi) as a top electrode (TE) gate contact.

A second patterning step with the top gate electrode as a HM is used to structure the top stack and define the MFM area. A SC1 etch of the TiN capping metal as well as the dry etch of the top HSO layer with a selective etch stop at the BE; both defines an MFM to MFIS area ratio (AR-TB) as in figure 1(b). The FeFETs were annealed at 1050 °C, 5 s for the source/drain dopant activation and ferroelectric HSO film.
crystallization in a single step. The inset of figure 1(b) shows the TEM cross section of the stack by TEM for an MFM/MFIS area ratio (AR-TB ∼ 0.5), the inset shows the stack layers of the MFMFIS device, and (c) simplified process flow of the integration concept showing the two-step patterning to realize the top/bottom MFM/MFIS stack. (b)–(c) Reproduced from [15] with permission. Copyright 2020, IEEE.

3. Results and discussion

3.1. The MFMFIS FeFET P–E and FET characteristics

The unique FeFET structure in figure 1(a) allows to measure the properties of the integrated ferroelectric capacitor via an optional monitor terminal of the bottom MFMFIS electrode. The polarization of the ferroelectric material as a normalized charge over the area is sensed by reversible dipole switching current. This is based on a measurement of the displacement current where the integral resembles the polarization values. Thus, accurate polarization extraction relies on a reduced intrinsic dielectric leakage current and pronounced ferroelectric dipole displacement currents [17]. In conventional MFIS FeFETs, the interfacial layer at the Si side has incomplete charge compensation and non-ideal dielectric properties compared to the more symmetric MFM case. Thereby, the MFIS configuration typically has non-ideal hysteresis shape, special technique such as the pulse up negative down measurement is used for accurate extraction of the ferroelectric polarization to reduce the dielectric leakage contribution. In contrast, the MFMFIS configuration provides a top stack with the ferroelectric material integrated between two metal electrodes. Hence, an accurate P–E hysteresis measurement and assessment for process optimization is possible based on effective ferroelectric thin film properties as integrated inside the FeFET gate stack.

The figure 2(a) shows the MFIS based FeFET with the integrated MFM characteristics for (a) polarization versus electric field hysteresis loops measured at electric field sweep (2.0–3.9 MV cm⁻¹), (b) the equivalent switching current versus electric field, (c) an endurance cycling for 10⁶ cycles measured at different switching field (2.0, 2.25, 2.5 MV cm⁻¹), (a)–(c) shows characteristics measured between the top and the bottom metal electrode contacts. (a)–(c) Reproduced from [15] with permission. Copyright 2020, IEEE.
Figure 3. (a)–(c) The MFMFIS FeFET drain current versus gate voltage $I_{D_{\text{sat}}}$–$V_{G}$ characteristics measured for the different MFM to MFIS area ratio $AR-TB$, inset shows the corresponding device structure schematic highlighting the area factor tuning and shows a weaker inversion of the Si channel as the AR-TB decreases, further illustrated by (d) the drain saturation current ($I_{D_{\text{sat}}}$) at a readout voltage $V_G = 3$ V showing a decreasing trend of $I_{D_{\text{sat}}}$ as the area ratio AR-TB decreases. (a)–(d) Reproduced from [15] with permission. Copyright 2020, IEEE.

$P$–$E$ hysteresis is measured at a frequency of 1.0 KHz using triangular sweep pulses. The increase in the pulse amplitude leads to a higher $P$, and features the transition from sub-loop to saturated switching when the switching peak amplitude increases. The $P$–$E$ measurement in figure 2(a) shows a tilted hysteresis shape indicating a broader $E_c$ distribution of grains, a distinctive feature for lower device switching variability. Likewise, the corresponding $I$–$E$ current in figure 2(b) shows a broader peak at the switching instant. This indicates a strong difference in the film crystallization that in turn appears in the grain size and orientation distribution, hence the device variability, particularly when compared to the conventional MFIS configuration. Similarly, the top MFM endurance is measured at 10 KHz cycling frequency and shows in figure 2(c) the polarization $P$, over cycling for sub-loop operation measured at $10^6$ cycles versus the earlier breakdown at higher fields. The electric field for sub-loop operation (2.0–2.5 MV cm$^{-1}$) is defined by applying 2.0–2.5 V between the top and BE of the MFM in the stack. Thus, the BE contact allows to characterize the integrated ferroelectric layers as well as a dynamic and independent modulation of the ferroelectric layers switching during memory operation.

The floating metal insertion allows for a separate control of the MFM to MFIS capacitor area ratio. This role appears to have strong influence on the $I_{D_{\text{sat}}}–V_{G}$ characteristics [figures 3(a)–(c)]. The inset of figures 3(a)–(c) shows the depicted device structure in reference to the cell design at different AR-TB ratios. The $I_{D_{\text{sat}}}–V_{G}$ characteristics is measured for the pristine state of the devices. Due to the equal charge between two capacitors in series, the voltage division can be controlled directly by the area ratio control. For an $AR-TB \gg 1.0$, the top gate is wrapped around the bottom MFIS stack such that a maximized electrostatic coupling is present at the bottom stack. In essence, this configuration aims for enhanced interfacial oxide fields as commonly applied in the floating gate memory concept [18]. As AR-TB decreases to unity, the MFMFIS is patterned at once to produce equal electrostatic coupling between the top and bottom stack structures. Ultimately, when $AR-TB \ll 1.0$, the top MFM capacitor area becomes smaller compared to the MFIS one and that leads to maximized electrostatic coupling across the top MFM capacitor. Thus, in turn a weaker channel inversion occurs as AR-TB decreases. For the opposite case where $AR-TB (\gg 1.0)$, the higher AR-TB means the TE is wrapped around the BE hence a maximized voltage distribution at the Si surface is achieved. This in turn permits a stronger Si channel inversion.

Hence, a general trend of the $I_{D_{\text{sat}}}$ can be extracted in figure 3(d) versus the AR-TB change. This dependence as tailored stack voltage distribution via AR-TB tuning shows in figure 3(d) that $I_{D_{\text{sat}}}$ decreases when AR-TB area is smaller. The $I_{D_{\text{sat}}}–V_{G}$ variation is a function of the AR-TB where the smallest AR-TB (small MFM area) reflect an increased variation, particularly due to the area dependence of the ferroelectric film variability. At first glance, the trend in figure 3(d) with the sole $I_{D_{\text{sat}}}–V_{G}$ characteristics, independent of memory switching, indicates the possible tailoring of the stack voltage distribution. This may resemble a challenge in terms of readout of the stored state. Likewise, in a practical array implementation, this may increase the potential for higher disturbance on the cells as in the case of AR-TB $\ll 1$. The AR-TB can be viewed as a design knob where a tradeoff between the low voltage operation versus the reliable sensing of the stored state may hinder a strong reduction of the AR-TB factor [19].

3.2. The MFMFIS-based memory characteristics

The memory switching characteristic of the MFIS and MFMFIS part of the FeFET stack are measured in figures 4(a)–(b) using an applied program/erase (PG/ER) pulse followed by readout as depicted in the figure 4(d) sequence. The separate contacts of the BE and TE allow for an independent write or readout of the state for the MFIS and MFM stacks, respectively. In the first case as in figure 4(a), the BE is used to PG/ER and readout the bottom MFIS stack using [5.0 V], 300 ns pulses. In contrast, figure 4(b) shows the memory switching characteristics of the same device via PG/ER pulse [7.0 V], 300 ns applied on the full MFMFIS stack structure. The difference in applied write amplitude of the MFIS versus MFMFIS stack represents the MFM capacitor voltage. As the saturated switching of the MFIS stack is obtained around [5 V], a higher voltage [7.0 V] is applied for the full MFMFIS stack switching. The device bottom MFIS structure is $W \times L$ (20 $\mu$m $\times$ 15 $\mu$m) while the top MFM
structure is \( W \times L \) (20 \( \mu \text{m} \times 5 \mu \text{m} \)) corresponding to an overall stack AR-TB ratio of \( \sim 0.3 \).

Thereby, a dual ferroelectric layer integration in the MFMFIS stack allows for tuning the memory window (MW) size. Thereby, a higher MW can be obtained for the MFMFIS case compared to the MFIS one is obtained. This is due to the superposition sum of the remnant polarization contribution of the two ferroelectric layers as depicted schematically in figure 4(c) of the PG state. Besides the thickness and coercive field dependence of the MW, a low values of the remnant polarization impacts the maximum MW as recently examined \[20\]. The gate current versus voltage (\( I_G-V_G \)) measured on the MFIS or the full MFMFIS stacks is plotted in figures 4(a)–(b) indicates a low gate stack leakage even with the MFMFIS floating metal insertion.

3.3. The impact of MFMFIS stack ratios tuning

The impact of tuning AR-TB is considered by a gradual increase of an ER pulse amplitude while preceded by PG reference pulse as shown in the figure 5(c) test sequence. The measured MW is defined as the difference between the \( V_T \) shift of each ER amplitude, and the initial \( V_T \) of PG reference. The independent measurement of the MW evolution versus pulse amplitude applied on the MFIS and MFMFIS part of the stack is shown in figures 5(a)–(b) at different area ratios. The coupled area ratio parameter AR-TB is defined as the area of the MFM capacitor divided by the MFIS one whereas the AR-BT is the reciprocal ratio of AR-TB [figure 5(d)]. Practically, the MFM capacitor area depicted in figure 4(c) is constant for a certain \( W \times L \) dimension while the MFIS area is a design variable.

In figure 5(a), the increase in AR-BT i.e. the MFIS area increase, shows as expected indifference in the trend of MW over pulse amplitude sweep. Below a critical amplitude threshold e.g. \(-3.0 \text{ V}\), the field across the ferroelectric layer is insufficient for switching. As the amplitude increases, a transition toward a maximum MW (\( \sim 0.6 \text{ V} \)) that saturates at higher amplitudes (ER >\( 5.0 \text{ V} \)) indicates a full polarization switching of the bottom MFIS ferroelectric layer. In contrast,
the full MFMFIS stack at MFM $W \times L$ (5.0 $\mu$m $\times$ 5.0 $\mu$m) shows as in figure 5(b) a strong dependence for the MW on the AR-TB as measured for different ER amplitude. For the same pulse amplitude, a lower AR-TB produces higher MW. As depicted in figure 5(d), the AR-TB decreases hence a higher field across the MFM capacitor is accompanied with a $P_r$ increase and as a result the MW increases. Due to the MFM switching, the MW saturation occurs at a higher pulse amplitude compared to the MFIS case. At higher erase write amplitudes (e.g. $-8.5$, $-9.0$ V), the stack reaches breakdown conditions in figure 5(b). This role of the MFM switching can be further highlighted in figure 6(a) by considering the maximum MW of the MFIS and MFMFIS stacks.

3.4. The maximum FeFET MW

The maximum MW is measured with optimal PG/ER pulses for saturated ferroelectric switching as in the case of the MFIS (5.0 V), 300 ns) or the MFMFIS (7.0 V), 300 ns) pulse condition. The cross wafer MW statistics in figure 6(a) is measured on device with MFIS area $W \times L$ (10 $\mu$m $\times$ 10 $\mu$m) while the MFIS area increased from $W \times L$ (10 $\mu$m $\times$ 11 $\mu$m) to $W \times L$ (10 $\mu$m $\times$ 28 $\mu$m) to cover a broad range of the AR-TB change. In figure 6(a), the change in MFIS area i.e. the AR-BT increases for constant MFM area $W \times L$ (5.0 $\mu$m $\times$ 10 $\mu$m), shows minor impact on the maximum MW. In contrast, the MFMFIS MW becomes higher as the MFM to MFIS area ratio AR-TB decreases. In essence, unlike the MW trend of the MFIS stack, the MFMFIS ones shows that a small change in AR-TB has a strong impact on the MW size, particularly at lower AR-TB. This is mainly due to the smaller voltage needed for the case of an MFM switching. At constant write amplitude, the transition from a sub-loop MFM switching to a full saturated loop is selectively set via the controlled sweep of AR-TB. The additional MFM polarization switching contributes to a higher MW size of the overall MFMFIS stack.

3.5. The role of device scalability and variability

In terms of the device scalability, the top MFM area resembles the critical feature size for scaling the MFMFIS concept. At a constant AR-TB, the MFIS area is in turn very dependent on the ultimate top MFM capacitor scaling. Thus, the impact of the top MFM scaling on the device performance is an important optimization aspect. In figure 6(b), the transition between a large MFM $W \times L$ (15 $\mu$m $\times$ 15 $\mu$m) to a small $W \times L$ (0.5 $\mu$m $\times$ 0.5 $\mu$m) indicates a higher variability in the measured maximum MW. This variability is intrinsic to the polycrystalline ferroelectric film properties of the top MFM capacitor superimposed on the bottom MFIS stack where the area is relatively larger as $W \times L$ (0.5 $\mu$m $\times$ 18.5 $\mu$m). In essence, the selected critical dimension of both the MFIS and MFM covers a broad picture due to the strong implications and tradeoffs between scalability, operating voltage, and variability. The AR-TB transition to unity is optimal for scaling the concept although it rises the operating voltage. Besides the AR-TB ratio tuning, the thickness parameter of both ferroelectric layers and permittivity or thickness of the interfacial layer with Si becomes another tuning knobs for stack optimization. Thereby, these tuning knobs give opportunities for design, process, and technology optimization for a target device performance. In terms of full wafer i.e. global variability of the maximum MW, the trend in figure 6(c) is measured using MFM area $W \times L$ (5.0 $\mu$m $\times$ 5.0 $\mu$m) at 7.0 V, 300 ns pulse condition. The trend shows the distribution over different AR-TB (0.8–0.16) where, as outlined earlier, the MW increases upon smaller AR-TB. Although, a constant PG/ER condition 7.0 V, 300 ns is applied for saturated ferroelectric stack switching, the merit of area ratio tuning is a lower memory operating voltage.

3.6. The MFMFIS low voltage operation

An MFMFIS stack switching in dependence of the AR-TB tuning is studied in figure 7 by the symmetric PG/ER sweep for pulse amplitude ([5.0 V]–[8.0 V]) and pulse width.
(50–700 ns). Initially, for each pulse amplitude, the width of PG/ER is symmetrically varied while measuring the corresponding $V_T$ shift and extracting the MW. The measured device has a top MFM area as $W \times L$ (10 $\mu m \times 2.0 \mu m$). At area ratio AR-TB $\sim 0.7$, the MW in figure 7(a) is defined mainly by the MFIS contribution of switching due to the negligible voltage across the MFM structure. Thus, the MW becomes relatively constant over the pulse width. As AR-TB decreases toward an AR-TB $\sim 0.3$, the dependence on selected amplitude is more pronounced as in figure 7(b) such that a higher MW occurs at higher amplitudes. This indicates an increased MFM contribution to the full stack ferroelectric switching. Accordingly, a higher amplitude results in an increased fraction of the voltage across the MFM part of the stack. In turn, the MFM switching contribution becomes ultimate in figure 7(c) at the smallest AR-TB $\sim 0.2$. For the smallest AR-TB, the MFM voltage is sufficiently high for saturated switching even at the lowest [5.0 V] amplitude. The low voltage switching is a strong merit of the AR-TB tuning, this is further outlined in figure 7(d). For a selected amplitude, the trend in figure 7(d) shows a higher MW as the AR-TB decreases. Likewise, the same MW magnitude can be realized at lower voltage upon smaller AR-TB ratio.

3.7. The MFMFIS-based FeFET reliability

In terms of device reliability, the effect of operating temperature, endurance at extended cycles of PG/ER pulses, and the retention of the stored state are studied. The temperature influence is outlined in figure 8(a) by the MW evolution over the temperature sweep (−40 $^\circ$C to 40 $^\circ$C). Initially, the pristine MW is measured at 25 $^\circ$C followed by a low temperature measurement at −40 $^\circ$C and an incremental 10 $^\circ$C temperature increase with measurement up to 40 $^\circ$C and as a final 25 $^\circ$C measurement. The MW versus temperature is measured for different AR-TB ratios (0.83, 0.18, 0.016). The trend in figure 8(a) shows a maximized MW size at lower temperatures, while the MW decreases as the operating temperature increases independent of the AR-TB ratio. This effect is recently reported for the FeFET memory and attributed to the role of $E_c$ and pyroelectric $P_e$ change with temperature [20–22]. The recovery of the MW size at the final 25 $^\circ$C indicates a reversible temperature effect. In particular, the reversible $P_e$ change with temperature, becoming larger at lower temperatures and vice versa causes a reciprocal MW dependence on the temperature [22].

The endurance is obtained in figure 8(b) using [5.0 V], [7.0 V] at 300 ns cycling conditions and intermediate PG/ER $V_T$ readout of the MFIS and the MFMFIS stack, respectively. A maximum endurance of $10^5$ cycles for the MFIS or MFMFIS stack is limited by the ER $V_T$ degradation after extended cycling. The retention is measured in figure 8(c) at room temperature (RT) by initially writing the PG/ER state while measuring the $V_T$ evolution for 10 d. The stable PG/ER $V_T$ readout versus retention time is extrapolated for 10 year and indicates a stable retention of the MFMFIS stack.

3.8. The MFMFIS FeFET based synapse

The biological brain resembles an exemplary system for efficient computing particularly in cognitive tasks at extremely low energy. In brain inspired computing, the constituting biological elements of the brain as synapses and neurons as well as the neuro-synaptic dynamics, are emulated by artificial non-volatile technologies [23]. The unsupervised learning resembles a localized learning algorithm in the sense that a synaptic weight update is dependent solely on the activities of connected neurons [24]. One type of unsupervised learning is the spike timing dependent plasticity (STDP) that relates to the controlled synaptic efficacy by the relative timing of pre and post neuron’s output spike as originated from the Hebbian theory [25]. In particular, the long term potentiation/depression (LTP/LTD) signifies the synaptic weight increase and decrease, respectively. An accurate bio-fidelity relies on the intrinsic physics of the artificial device with an operation that can map closely to the biological STDP learning.

Hereby, the conversion of an MFIS stack toward an MFMFIS one expands the paradigm for broad tuning of the underlying FeFET operation physics for accurate LTP and LTD. In particular, the MF capacitor integration in the gate stack gives a fine control of the device switching due to the small voltage required to induce polarization change. The LTP and LTD is emulated in figures 9(a), (d) with a fine control of the switching current (synaptic weight) by varying the pulse width and amplitude as depicted in figures 9(b)–(c) sequence [26]. After an initial reference pulse, a pulse width scheme [figure 9(b)] uses the sweep range (50–300 ns) in
steps of 10 ns at constant positive or negative amplitude of LTP or LTD, respectively while measuring the $I_D$–$V_G$ after each pulse. Likewise, the amplitude scheme [figure 9(c)] uses the range ($|4.0\ V|$–$|6.0\ V|$) at 50 mV step and a rather constant pulse width. As the ferroelectric switching in fluorite-structure ferroelectric materials follows a nucleation limited switching, a tradeoff between the pulse width and amplitude enables a dominant time-driven or field-driven saturated switching [27, 28]. Thereby, the synaptic plasticity in figure 9(a) shows saturated LTP/LTD as the pulse width increases indicating full polarization switching. In contrast, the amplitude dependence in figure 9(d) shows a very linear weight update that originates in part from the fine change in the field across the MFM capacitor.

This can be further outlined by exploring the AR-TB tuning in the MFMFIS stack. As in figures 10(a)–(b), an increased linearity and symmetry of the synaptic LTP/LTD at wider current dynamic range occurs as the AR-TB decreases. The decrease in AR-TB marks the transition between a dominant MFIS switching where a narrow window is produced hence a smaller dynamic range of current change to a wider but more fine controlled window when the MFM begins switching.

Additionally, a constant PG/ER pulse sequence [figure 11(a)] is explored for achieving the synaptic LTP/LTD operation. A symmetric PG/ER pulse amplitude sweep ($|4.5\ V|$–$|6.0\ V|$) shows in figure 11(b) a higher current dynamic range as the pulse amplitude increases. The synaptic current weight update shows a decrease in the number of intermediate levels compared to the pulse width or pulse amplitude based LTP/LTD discussed earlier in figure 9. For a constant $|5.5\ V|$ amplitude, the impact of the MFM to MFIS...
Area ratio control is shown in Figure 11(c). The decrease in AR-TB shows a faster convergence toward synaptic LTD current saturation whereas the LTP one shows full synaptic switching independent of the AR-TB. The dynamic range dependence on the pulse amplitude at a limited number of weight update current levels indicates the pulse width or pulse amplitude based sequences more suitable for emulating the MFMFIS synaptic LTP/LTD. The dynamic control of each ferroelectric layer switching in the MFMFIS device via tailored pulse conditions as well as via the area ratio AR-TB tuning renders such an MFMFIS stack concept as a promising electronic synapse.

4. Conclusion

In this paper, a novel dual ferroelectric layer FeFET concept is reported using an MFMFIS stack structure with Si-doped hafnium oxide (HSO) integrated ferroelectric material. The bottom metal electrode insertion is used to tailor the voltage distribution inside the stack via the MFM to MIS area ratio (AR-TB) control. This enables a dynamic and independent control of the ferroelectric layers switching as well as a low voltage FeFET operation for smaller AR-TB. However, a tradeoff between the low voltage operation and the reliable sensing of the stored state occurs due to the weaker gate voltage coupling to the Si channel for the smallest AR-TB.

In relation to the pulse conditions, the MW size can be modulated by the superposition of the dual ferroelectric layer polarization switching effect on the Si channel, additionally controlled by the AR-TB tuning. The scalability of the MFM capacitor area indicates higher variability as monitored on the cross wafer statistics of maximum MW. Thereby, the optimal stack definition relies on the essential tradeoffs between operating voltage, device MFIS critical area, and variability. The temperature dependence of polarization and coercive field reflects a trend of decreased MW at higher temperatures. An endurance of $10^5$ cycles becomes limited due to an erase threshold voltage degradation of both the MFIS and MFMFIS stacks. A stable retention is measured for 10 h at RT and further extrapolated to 10 years indicates a stable MFMFIS stack construct.

The stack structure tuning and dual ferroelectric integration turns out as a pivotal merits for FeFET implementation as a synaptic device. The synaptic plasticity as potentiation or depression exhibits a broad transition from saturated response toward a fine weight update at improved linearity and symmetry via the controlled AR-TB and MFM switching. The MFMFIS FeFET offers a novel low voltage and high-speed multi-ferroelectric layer integration concept for optimized storage and neuromorphic computing.

Acknowledgments

The authors are very thankful to Patrick Polakowski for his support during device processing. The authors are also thankful to Uwe Mühle for his support with the TEM stack analysis.

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

Tarek Ali @ https://orcid.org/0000-0002-9840-3531
Sebastian Oehler @ https://orcid.org/0000-0002-5394-2857
Franz Müller @ https://orcid.org/0000-0002-6564-9121

References

[1] Jerry M, Chen P, Zhang J, Sharma P, Ni K, Yu S and Datta S 2017 Ferroelectric FET analog synapse for acceleration of deep neural network training 2017 IEEE Int. Electron Devices Meeting (IEDM) (https://doi.org/10.1109/iedm.2017.8268338)
[2] Nawrocki R A, Voyles R M and Shaheen S E 2016 A mini review of neuromorphic architectures and implementations IEEE Trans. Electron Devices 63 3819–29
Hu V P, Lin H, Zheng Z, Lin Z, Lu Y, Ho L and Su C 2019
Ali T, Polakowski P, Riedel S, Buttner T, Kampfe T, Oh S, Hwang H and Yoo I K 2019 Ferroelectric materials for
Ali T, Kuhnel K, Czernohorsky M, Mart C, Rudolph M, Patzold B and Eng L M 2020 A study on the temperature-dependent operation of fluorite-structure-based ferroelectric HfO2 memory FeFET: pyroelectricity and reliability IEEE Trans. Electron Devices 67 2981–7
Gupta A, Ni K, Prakash O, Hu X S and Amrouch H 2020 Temperature dependence and temperature-aware sensing in ferroelectric FET 2020 IEEE Int. Reliability Physics Symp. (IRPS). (https://doi.org/10.1109/irps45951.2020.9129226)
Ali T, Kuhnel K, Czernohorsky M, Mart C, Rudolph M, Patzold B and Eng L M 2020 A study on the temperature-dependent operation of fluorite-structure-based ferroelectric HfO2 memory FeFET: a temperature-modulated operation IEEE Trans. Electron Devices 67 2793–9
Kendall J D and Kumar S 2020 The building blocks of a brain-inspired computer Appl. Phys. Rev. 7 011305
Chakraborty I, Jaiswal A, Saha A K, Gupta S K and Roy K 2020 Pathways to efficient neuromorphic computing with non-volatile memory technologies Appl. Phys. Rev. 7 021308
Langille J J and Brown R E 2018 The synaptic theory of memory: a historical survey and reconciliation of recent opposition Frontiers Syst. Neurosci. 12 52
Jerry M, Dutta S, Kazemi A, Ni K, Zhang J, Chen P and Datta S 2018 A ferroelectric field effect transistor based synaptic weight cell J. Phys. D: Appl. Phys. 51 434001
Gong N, Sun X, Jiang H, Chang-Liao K S, Xia Q and Ma T P 2018 Nucleation limited switching (NLS) model for HfO2-based metal-ferroelectric-metal (MFM) capacitors: Switching kinetics and retention characteristics Appl. Phys. Lett. 112 262903
Mulaosmanovic H, Muller F, Lederer M, Ali T, Hoffmann R, Seidel K and Slesazeck S 2020 Interplay between switching and retention in HfO2-based ferroelectric FETs IEEE Trans. Electron Devices 67 3466–71

[3] Böscze T S, Müeller J, Bräuhaus D, Schröder U and Böttger U 2011 Ferroelectricity in hafnium oxide thin films Appl. Phys. Lett. 99 102903
[4] Müller J, Böscze T S, Schröder U, Mueller S, Bräuhaus D, Böttger U and Mikolajick T 2012 Ferroelectricity in simple binary ZrO2 and HfO2 Nano Lett. 12 4318–23
[5] Trentzsch M, Flachowsky S, Richter R, Paul J, Reimer B, Utes D and Rice B 2016 A 28 nm HKMG super low power embedded NVM technology based on ferroelectric FETs 2016 IEEE Int. Electron Devices Meeting (IEDM) (https://doi.org/10.1109/iedm.2016.7838397)
[6] Dunkel S, Trentzsch M, Richter R, Moll P, Fuchs C, Gehring O and Beyer S 2017 A FeFET based super-low-power ultra-fast embedded NVM technology for 22 nm FDSOI and beyond 2017 IEEE Int. Electron Devices Meeting (IEDM) (https://doi.org/10.1109/iedm.2017.8268425)
[7] Mueller S, Muller J, Hoffmann R, Yurchuk E, Schlosser T, Boschke R and Mikolajick T 2013 From MFM capacitors toward ferroelectric transistors: endurance and disturb characteristics of HfO2-based FeFET devices IEEE Trans. Electron Devices 60 4199–205
[8] Chen Y, Su C, Hu C and Wu T 2019 Effects of annealing on ferroelectric hafnium–zirconium–oxide–based transistor technology IEEE Electron Device Lett. 40 467–70
[9] Ali T, Polakowski P, Riedel S, Buttner T, Kampfe T, Rudolph M and Muller J 2018 High endurance ferroelectric hafnium oxide-based FeFET memory without retention penalty IEEE Trans. Electron Devices 65 3769–74
[10] Muller J, Polakowski P, Paul J, Riedel S, Hoffmann R, Drescher M and Kolodinski S 2015 (Invited) integration challenges of ferroelectric hafnium oxide based embedded memory ECS Trans. 69 85–95
[11] Hu V P, Lin H, Zheng Z, Lin Z, Lu Y, Ho L and Su C 2019 Split-Gate FeFET (SG-FeFET) with dynamic memory window modulation for non-volatile memory and neuromorphic applications 2019 Symp. on VLSI Technology (https://doi.org/10.23919/vlsit.2019.8776555)
[12] Yoon S, Min D, Moon S, Park K S, Won J I and Yoon S 2020 ‘Improvement in long-term and high-temperature retention stability of ferroelectric field-effect memory transistors with metal–ferroelectric–metal–insulator–semiconductor gate-stacks using al-doped HfO2 thin films IEEE Trans. Electron Devices 67 499–504
[13] Sakai S and Takahashi M 2010 Recent progress of ferroelectric-gate field-effect transistors and applications to nonvolatile logic and FeNAND flash memory Materials 3 4950–64
[14] Oh S, Hwang H and Yoo I K 2019 Ferroelectric materials for neuromorphic computing APL Mater. 7 091109
[15] Ali T et al 2020 A novel dual ferroelectric layer based MFMFIS FeFET with optimal stack tuning toward low power and high-speed NVM for neuromorphic applications 2020 IEEE Symp. on VLSI Technology

[16] Ali T et al 2018 Silicon doped hafnium oxide (HSO) and hafnium zirconium oxide (HZO) based FeFET: a material relation to device physics Appl. Phys. Lett. 112 222903
[17] Schenk T, Yurchuk E, Mueller S, Schroeder U, Starchisch S, Böttger U and Mikolajick T 2014 About the deformation of ferroelectric hystereses Appl. Phys. Rev. 1 041103
[18] Gan L, Wang Y, Chen L, Zhu H and Sun Q 2019 A floating gate memory with U-shape recessed channel for neuromorphic computing and MCU applications Micromachines 10 558
[19] Ali T, Polakowski P, Buttner T, Kampfe T, Rudolph M, Patzold B and Muller J 2019 ‘Theory and experiment of antiferroelectric (AFE) Si-doped hafnium oxide (HSO) enhanced floating-gate memory IEEE Trans. Electron Devices 66 3356–64
[20] Ali T, Kuhnel K, Czernohorsky M, Mart C, Rudolph M, Patzold B and Eng L M 2020 A study on the temperature-dependent operation of fluorite-structure-based ferroelectric HfO2 memory FeFET: pyroelectricity and reliability IEEE Trans. Electron Devices 67 2981–7
[21] Gupta A, Ni K, Prakash O, Hu X S and Amrouch H 2020 Temperature dependence and temperature-aware sensing in ferroelectric FET 2020 IEEE Int. Reliability Physics Symp. (IRPS). (https://doi.org/10.1109/irps45951.2020.9129226)