

Hafnium-based ferroelectric devices for neuromorphic computing: From materials physics to systems

Zhenhua Song^{1,§}, Xiumin Xu^{1,§}, Jianxiu Liu⁴, Ruihong Ruan², Xingcan Guo², Wenhao Chen², Tianxing Wei², Jiajie Yu², Qingxuan Li¹, Yiqun Hu¹(✉), Zihan Liu¹, Qingqing Sun^{2,3}, David Wei Zhang^{2,3}, Zhenhai Li¹(✉), Lin Chen^{2,3}(✉)

¹ School of Integrated Circuits, Anhui University, Hefei 230601, China

² College of Integrated Circuits & Micro-Nano Electronics, Fudan University, Shanghai 200433, China

³ Nano Institute of Fudan University, Shanghai 201203, China

⁴ School of Electronic Information, Anhui Audit College, Hefei 230601, China

§ Zhenhua Song and Xiumin Xu contributed equally to this work.

Nano Res., **Just Accepted Manuscript** • <https://doi.org/10.26599/NR.2026.94908925>

<https://www.sciopen.com/journal/1998-0124> on Jun. 8, 2026

© The Authors(s)

Just Accepted

This is a “Just Accepted” manuscript, which has been examined by the peer-review process and has been accepted for publication. A “Just Accepted” manuscript is published online shortly after its acceptance, which is prior to technical editing and formatting and author proofing. Tsinghua University Press (TUP) provides “Just Accepted” as an optional and free service which allows authors to make their results available to the research community as soon as possible after acceptance. After a manuscript has been technically edited and formatted, and the page proofs have been corrected, it will be removed from the “Just Accepted” web site and published officially with volume and article number (e.g., *Nano Research*, **2025**, 18, 94906990). Please note that technical editing may introduce minor changes to the manuscript text and/or graphics which may affect the content, and all legal disclaimers that apply to the journal pertain. In no event shall TUP be held responsible for errors or consequences arising from the use of any information contained in these “Just Accepted” manuscripts. To cite this manuscript please use its Digital Object Identifier (DOI®), which is identical for all formats of publication.

Hafnium-based ferroelectric devices for neuromorphic computing: From materials physics to systems

Zhenhua Song^{1,§}, Xiumin Xu^{1,§}, Jianxiu Liu⁴, Ruihong Ruan², Xingcan Guo², Wenhao Chen², Tianxing Wei², Jiajie Yu², Qingxuan Li¹, Yiqun Hu^{1*}, Zihan Liu¹, Qingqing Sun^{2,3}, David Wei Zhang^{2,3}, Zhenhai Li^{1*}, and Lin Chen^{2,3*}

¹*School of Integrated Circuits, Anhui University, Hefei 230601, China*

²*College of Integrated Circuits & Micro-Nano Electronics, Fudan University, Shanghai 200433, China*

³*Nano Institute of Fudan University, Shanghai 201203, China*

⁴*School of Electronic Information, Anhui Audit College, Hefei 230601, China*

[§] Zhenhua Song and Xiumin Xu contributed equally to this work.

Received: 10 April 2026; **Revised:** 28 May 2026; **Accepted:** 8 June 2026

✉ Address correspondence to Yiqun Hu, huyiqun@ahu.edu.cn; Zhenhai Li, li.zhenhai@ahu.edu.cn; Lin Chen, linchen@fudan.edu.cn

Cite this article: *Nano Research*, 2026, 19, 94908925 <https://doi.org/10.26599/NR.2026.94908925>

Abstract

Technological progress is advancing at a rapid pace, and the rate of chip performance improvements and transistor density increases is gradually diverging from Moore's Law predictions. Simultaneously, the "memory wall" and "power wall" challenges inherent in the von Neumann architecture have become increasingly prominent. Today, neuromorphic computing—inspired by the human brain—is emerging as one of the most promising computational paradigms. To realize efficient brain-like hardware, a high-performance, scalable, non-volatile memory device is essential. Among these, ferroelectric field-effect transistors (FeFETs) based on hafnium oxide (HfO_2) stand out as the most representative. It effectively mimics biological synaptic plasticity while offering high compatibility with CMOS processes, excellent scalability, fast low-power switching characteristics, and inherent controllable three-terminal field-effect properties.

The paper systematically reviews the research progress of HfO_2 -based ferroelectric devices in the field of neuromorphic computing. Starting from the ferroelectric physical properties of HfO_2 , it first discusses engineering regulation of the ferroelectric phase

through methods such as doping and strain, and elaborates on the associated material-level challenges. Building upon this foundation, this review discusses representative HfO_2 -based ferroelectric device architectures, including FeFETs, ferroelectric tunnel junctions (FTJs), ferroelectric diodes (Fe-diodes), and advanced fin field-effect transistor (FinFET)/nanowire ferroelectric transistors. The primary focus is on their applications in simulating biological synapses and neurons, along with the impact on performance optimization. The study also extends from devices to system-level integration, addressing issues such as the non-idealities in crossbar arrays, power consumption in peripheral circuits, and challenges in hardware-software co-design. Finally, this research comprehensively summarizes the significant challenges currently facing science and technology from multiple perspectives and proposes several promising solutions. Through comprehensive research, this paper identifies contemporary development directions: integrated applications of two-dimensional ferroelectric materials, exploration of novel computational paradigms, and convergence of sensing, storage, and computation. It aims to provide a comprehensive reference framework for future research in this field.

1 Introduction

Over the past half century, computing systems based on the von Neumann architecture have driven the digital transformation of human society at an unprecedented pace. Today, Moore's Law is approaching its physical limits, while computationally intensive applications like artificial intelligence and big data continue to grow. Consequently, the "memory wall" and "power wall" bottlenecks in traditional architectures have become practical obstacles hindering technological progress[1, 2]. It is understood that data transfer time between processors and memory accounts for nearly 90% of the system's total time or energy consumption [3] . Therefore, a revolutionary breakthrough in computational paradigms is urgently needed. The biological brain stands as nature's most efficient computational system. Consequently, both academia and industry have been actively advancing the development of ferroelectric devices and two-dimensional material integration technologies. These innovations have significantly enhanced the energy efficiency and multi-bit storage capabilities of neuromorphic computing[4].

In fact, due to the significant area and power overhead introduced by on-chip static random-access memory (SRAM), coupled with the low efficiency of neural network model mapping, even the latest digital brain-inspired processors face the challenge of the "on-chip memory barrier." Memory has once again become an obstacle to performance breakthroughs. Ferroelectric memory computing technology, which achieves true deep integration of storage and computation, is currently the primary direction for developing analog computing architectures[5].

At present, neuromorphic computing inspired by the human brain stands as one of the most promising computational paradigms. Its core concept involves constructing computing systems that deeply integrate processing and storage, thereby overcoming the inherent data movement bottlenecks of the von Neumann architecture. To realize such efficient brain-like hardware, humanity must identify non-volatile physical devices capable of efficiently

emulating the plasticity of biological synapses[6]. As research progresses, technologies such as phase-change memory (PCM) and resistive random-access memory (RRAM) have demonstrated significant potential. However, challenges remain, including power consumption, stability, and analog accuracy.

Hafnium oxide (HfO_2)-based ferroelectric field-effect transistors (FeFETs), however, differ from these two-terminal resistive devices by possessing their own unique field-effect modulation pathway. This endows them with high compatibility with advanced complementary metal–oxide–semiconductor (CMOS) processes, excellent scalability, fast low-power switching characteristics, and three-terminal controllability, among other advantages. Consequently, FeFETs are currently regarded as one of the most promising devices for constructing high-performance electronic synapses[7]. The discovery of ferroelectricity in HfO_2 has brought ferroelectric memory back onto the path of Moore's Law development. Recent investigations further indicate that integrating two-dimensional semiconductors with ferroelectric HfO_2 can significantly enhance synaptic performance and energy efficiency[8].

The paper systematically reviews research progress on HfO_2 -based ferroelectric devices in the field of brain-inspired computing, focusing on aspects such as material physics fundamentals and system-level applications. Unlike traditional reviews that primarily concentrate on device structures, this study explores four dimensions: physical origins, device optimization strategies, system integration challenges, and technological directions. From these four perspectives, the paper first analyzes the unique metastable phase ferroelectric mechanism of HfO_2 , then explains how FeFET structures simulate synaptic and neuronal functions. It subsequently examines various non-ideal factors emerging in large-scale memory computing systems and corresponding hardware-software co-design solutions. Finally, it outlines future development directions, including two-dimensional material integration and novel computing paradigms. Figure 1 provides an intuitive overview of the paper's overall structure, serving as a clear and comprehensive technical roadmap for researchers and engineers in this rapidly evolving field.

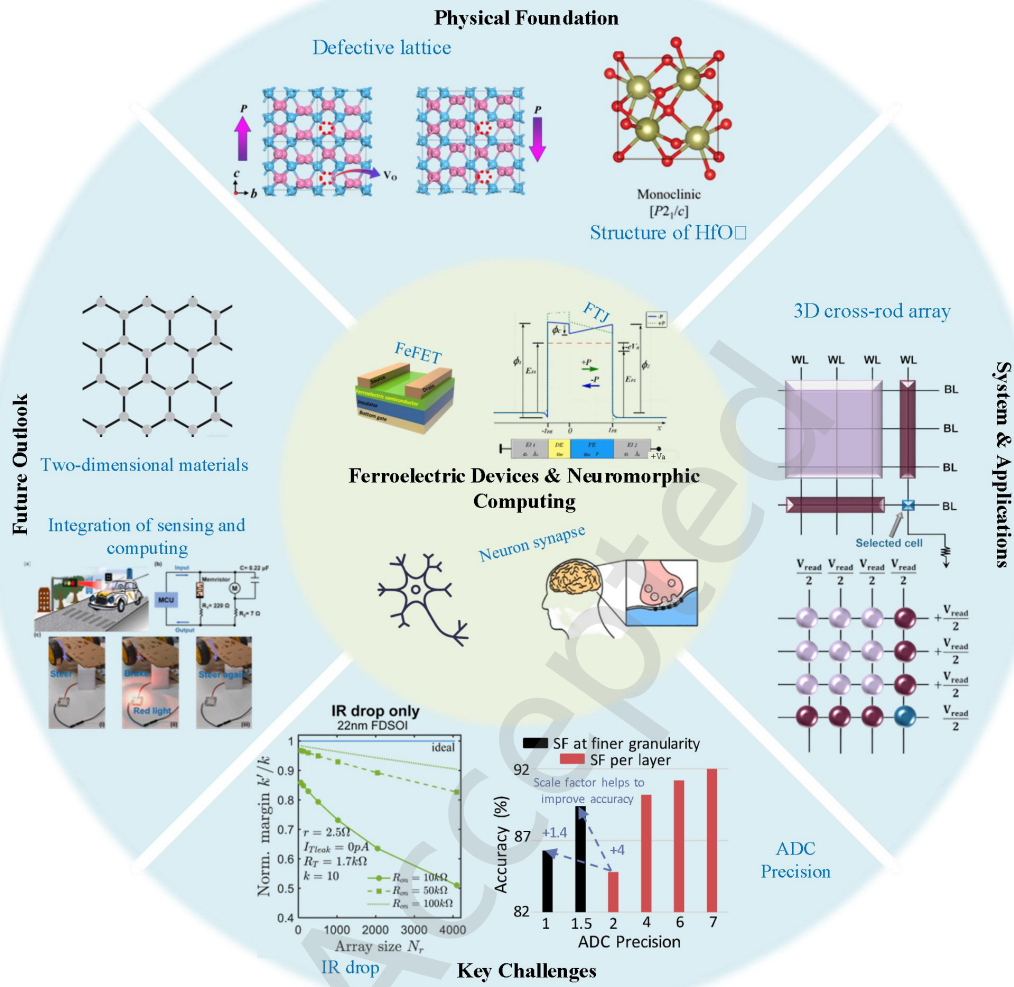


Figure 1 A panoramic overview of HfO_2 -based ferroelectric devices and their applications in neuromorphic computing[9-14]

Scope and contribution of this review

This review constructs a cross-scale analytical framework based on the materials physical properties and device operating mechanisms of HfO_2 -based ferroelectric materials, array-level non-ideal effects, circuit implementation methods, and neuromorphic system applications. On the other hand, in the content described in this review, material properties, device structures, and computing architectures are not independent existences from one another; instead, the connections among them are established to focus on clarifying phase stability, defect dynamic behavior, polarization switching processes, reliability issues, and their influences on various aspects. In addition, this review further introduces the working mechanisms and related application fields of several representative hafnium-based ferroelectric devices, including FeFETs, ferroelectric tunnel junctions (FTJs), ferroelectric diodes (Fe-diodes), and negative-capacitance-assisted ferroelectric circuits. Therefore, by discussing the above aspects in detail, this paper provides a systematic technical route for the development of ferroelectric hardware toward neuromorphic computing.

Literature scope and selection criteria.

The literature included in this review covers representative research achievements on HfO₂ -based ferroelectric materials published from the discovery of ferroelectricity in hafnium oxide in 2011 to the latest progress in 2025. The main covered contents include the intrinsic properties of the materials, various types of devices and their operating mechanisms, as well as related system-level applications. Meanwhile, priority is given to theoretical studies that have been fully validated.

2 Background and Current Research Status

2.1 Comparison of Emerging Non-Volatile Memory Technologies

Academic and industrial communities have been exploring various novel non-volatile memory technologies in pursuit of ideal electronic synapse devices. Currently, the primary candidate technologies include PCM, RRAM, and magnetoresistive random-access memory (MRAM). The content of Table 1 provides a qualitative comparison of several representative non-volatile memory technologies, including FeFETs, FTJs, Fe-diodes, RRAM, PCM, and MRAM, across multiple dimensions such as cell type, write speed, multilevel storage capability, linearity/symmetry, and CMOS process compatibility.

Table 1. Qualitative comparison of representative non-volatile memory technologies for neuromorphic and in-memory computing

Technology	Cell type	Write speed	Write energy / power	Endurance	Multi-level capability	Linearity / symmetry	CMOS compatibility	Array integration difficulty	Main limitation
FeFET	1T	Fast	Low	Moderate	Good	Requires pulse engineering	High	Moderate	Endurance, charge trapping, variability
FTJ	Two-terminal	Very fast	Very low read energy; low write energy	Moderate	Limited to moderate	Nonlinear	Moderate to high	Compact cell, but thickness-sensitive	Thickness sensitivity; low ON/OFF ratio
Fe-diode	Two-terminal self-rectifying	Fast	Low	To be optimized	Possible	Application-dependent	High	Favorable for cross-point arrays	Emerging technology; performance maturity
RRAM	1R / 1T1R	Fast	Low to moderate	Variable	Good	Often nonlinear/asymmetric	High	Selector often needed	Filament randomness; cycle-to-cycle variability
PCM	1R / 1T1R	ns-us	Relatively high write energy	Moderate	Good	Resistance drift	High	Mature, but energy-limited	Resistance drift, heating
MRAM	1T1MTJ	Very fast	Moderate	Excellent	Limited analog states	Limited analog tunability	High	Moderate	Small ON/OFF ratio; limited multi-level precision

Note. This table summarizes representative trends reported in reviews and key hafnia-based Fe-diode demonstrations [6, 15-18]. The assessments are intended for qualitative comparison rather than absolute ranking, because device performance depends strongly on material systems, device stacks, cell dimensions, programming schemes, array configurations, and measurement conditions.

It can be seen from this table that each type of technology has its own unique advantages and limitations. Among them, RRAM and PCM have compact two-terminal structures, but their disadvantage is that their analog precision is constrained by conductive-filament randomness and resistance drift. MRAM performs well in terms of endurance, but its biggest problem is the relatively low on/off ratio. Of course, FTJs and Fe-diodes, because they have more compact two-terminal structures, are indispensable in high-density integration. In contrast, HfO_2 -based FeFETs demonstrate the most promising characteristics. Featuring a 1T cell structure and CMOS compatibility, they are exceptionally well-suited for high-density integration. Concurrently, their transistor-level endurance and device variability remain active research areas. Furthermore, FeFETs exhibit sensitivity to stacked-layer design and operating conditions. Consequently, neuromorphic hardware based on FeFETs often requires joint optimization across materials, devices, and peripheral circuits—a topic explored in subsequent sections.

The first type to introduce is PCM. PCM essentially stores information by exploiting the resistance difference between crystalline and amorphous states of selenide compounds. While PCM technology is relatively mature today, it suffers from issues such as high write power consumption, resistance drift, and limited endurance. The second type to introduce is RRAM. RRAM primarily switches resistance by forming or breaking conductive filaments within a dielectric layer. Leon Chua theoretically predicted the existence of such a "memristor" in 1971[19]. Despite its relatively simple structure, its main issue is reduced device stability and consistency due to the random formation of conductive filaments, making it difficult to apply to neuromorphic computing requiring precise weight updates[20]. Finally, we introduce MRAM, which stores information through spin-dependent transport in magnetic tunnel junctions. Consequently, it offers ultra-fast read/write speeds and near-infinite endurance. However, it typically exhibits a small on/off ratio and struggles to achieve multiple analog states[21].

2.2 Evolution of Ferroelectric Devices and the Breakthrough of HfO_2

Unlike the resistive memory mentioned earlier, FeFETs offer a unique field-effect solution. Although the concept was first proposed by Moll and Tarui in 1963, practical implementation remained hindered for decades due to material limitations[22]. FeFET employs ferroelectric materials as the gate dielectric layer, where channel conductivity is controlled by the direction and strength of the remanent polarization (P_r). In two-dimensional FeFETs, its dual ferroelectric coupling effects have been applied to simultaneously enhance the efficiency of logic gates and synaptic functions, significantly advancing in-memory computing technology[23]. This three-terminal structure makes FeFETs particularly well-suited for simulating synaptic weight modulation. However, the incompatibility of traditional perovskite-based ferroelectric materials with standard CMOS processes and their scalability challenges have stifled significant advancements in FeFETs over the past several decades.

In 2011, researchers serendipitously discovered that hafnium oxide thin films exhibit

ferroelectric properties—a material long utilized as a high-dielectric-constant dielectric in CMOS technology[24]. This groundbreaking discovery immediately caused a sensation in the semiconductor industry, as it signaled that ferroelectric memory could for the first time be scaled according to Moore's Law[25]. Subsequent research confirmed the feasibility of FeFET devices using advanced 28-nanometer process nodes and analyzed their scaling potential [26]. Over the past decade, extensive research has been pushing the performance boundaries of ferroelectric devices.

Regarding scalability, traditional ferroelectric materials lose their ferroelectric properties at the nanoscale due to the presence of depolarization field effects. However, researchers discovered that HfO_2 exhibits the opposite behavior. In 2020, Cheema and colleagues achieved stable out-of-plane ferroelectricity in an HfO_2 film only 0.5 nm thick through precise stress control, pushing device scaling to its physical limit[27]. This research, published in Nature, dispelled longstanding industry doubts about the scalability of ferroelectric devices. In the field of devices and integration, the industry has also achieved remarkable accomplishments. In 2019, a research team at GlobalFoundries successfully fabricated embedded FeFET non-volatile memory using a 22-nm FD-SOI process. This memory demonstrated read/write speeds at the nanosecond level and exhibited outstanding endurance and retention characteristics[28]. Laleni and Kämpfe (2024) developed a 28 nm FeFET-based in-memory-computing macro, employing a novel readout circuit inspired by 3T pixel architectures, which significantly improved the energy efficiency of AI acceleration[29].

In terms of materials and physical mechanisms, existing technologies have surpassed traditional doped HfO_2 systems. Lee et al. (2020) experimentally demonstrated that inserting an ultrathin Al_2O_3 interfacial layer between $\text{Hf}_{0.5}\text{Zr}_{0.5}\text{O}_2$ and TiN electrodes significantly enhances remanent polarization and durability without doping. This indicates that ferroelectricity in HfO_2 -based multilayers can be induced or stabilized through pure interface engineering[30]. This discovery simplifies the associated fabrication process and offers new insights into the origin of HfO_2 ferroelectricity. HfO_2 -based ferroelectric devices have achieved major breakthroughs from fundamental physics research to commercial integration. This paper centers on HfO_2 -based ferroelectric devices. Its primary advantages are threefold: first, perfect compatibility with CMOS processes; second, enhanced ferroelectricity when HfO_2 is scaled below 10 nm. Consequently, it maintains stable ferroelectric behavior even at sub-nanometer thicknesses, demonstrating exceptional scalability. The third major point is environmental friendliness and material stability: HfO_2 is lead-free and chemically stable, making it ideal for sustainable electronics. In the following section, this paper will delve into the fundamental material physics underlying these phenomena.

3 Ferroelectric Physics of HfO_2 -Based Materials

HfO_2 -based ferroelectric materials represent a revolutionary technology, exhibiting unique physical and materials science properties in their crystal structure. Unlike traditional ferroelectrics such as perovskites, their greatest advantage lies in the fact that HfO_2 's

ferroelectricity originates from its metastable crystalline phase. However, this very characteristic also presents its own distinct challenges. This chapter will delve into the physical origins of HfO_2 ferroelectricity, engineering control methods, and the inherent reliability issues it faces at the material level.

3.1 Formation Mechanism of the Metastable Orthorhombic Phase

Unlike classical ferroelectrics with perovskite structures such as lead zirconate titanate (PZT), the ferroelectricity of hafnium oxide does not stem from its thermodynamically most stable state, but rather from a metastable crystalline phase. This unique physical origin is the key reason for HfO_2 's compatibility with CMOS processes, while also determining its pronounced sensitivity to processing conditions.

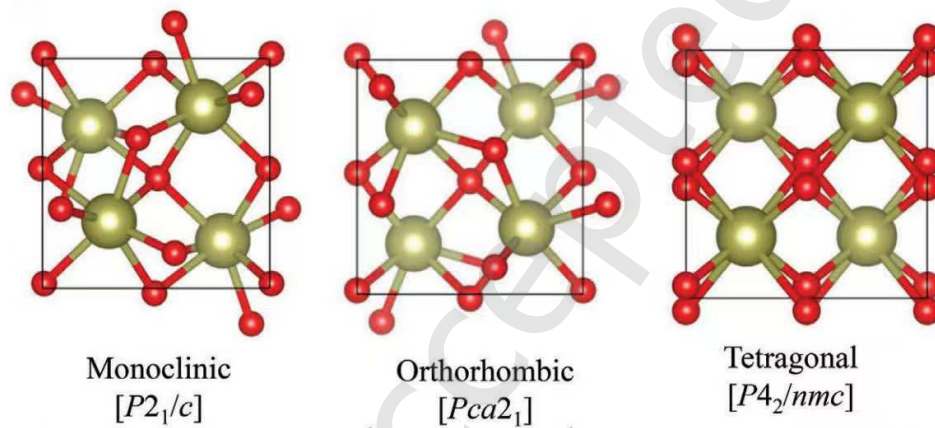


Figure 2 Schematic crystal structures of different HfO_2 phases[14]

HfO_2 can be classified as a polymorphic material whose crystal structure changes with variations in temperature and pressure. Under room temperature and ambient pressure, bulk HfO_2 has a single thermodynamically stable phase—the monoclinic phase—with space group $P2_1/c$. As shown in Figure 2, the monoclinic unit cell is centrosymmetric, meaning that the centers of positive and negative charges coincide. Consequently, it cannot generate spontaneous polarization and therefore does not exhibit ferroelectricity[31]. With increasing temperature, HfO_2 undergoes a series of phase transitions. At about 1720 °C, it transforms into a tetragonal phase; at even higher temperatures (around 2600 °C), it further transforms into a cubic phase. Although these high-temperature phases possess higher symmetry, they remain centrosymmetric and thus still lack ferroelectric characteristics.

Under conventional conditions, we generally consider the origin of hafnium-based ferroelectricity to be associated with the non-centrosymmetric orthorhombic $Pca2_1$ phase[32]. However, the ferroelectricity of hafnium-based materials is not limited to the orthorhombic polymorph. Recent studies have discovered a stable rhombohedral phase in ferroelectric $\text{Hf}(\text{Zr})_{1-x}\text{O}_2$ capacitors, which is a phase capable of exhibiting an ultralow coercive field; therefore, phase engineering can also be regulated based on this phase[33]. As shown in Figure 2, oxygen ions undergo cooperative displacement relative to hafnium ions within this phase, causing the positive and negative charge centers to no longer coincide within the unit cell. Consequently, it develops spontaneous polarization along the c-axis,

which can be reversed under an applied electric field, thus exhibiting ferroelectricity. Further analysis of this atomic structure demonstrates that the $Pca2_1$ space group is the fundamental cause of HfO_2 's ferroelectricity[34]. The central issue, however, is that this ferroelectric orthorhombic (*o*) phase is metastable. From the perspective of Gibbs free energy, at room temperature the relative energy ordering of different crystal phases is: $E(m\text{-phase}) < E(o\text{-phase}) < E(t\text{-phase})$. This indicates that if left to evolve naturally, all HfO_2 will eventually transform into the monoclinic phase—a state of minimum energy with no ferroelectricity. The formation of the ferroelectric orthorhombic phase occurs specifically during the cooling of the system from the high-temperature tetragonal phase. Researchers believe it is a product of a kinetic trap that emerges during the thin-film preparation process. Because the *o*-phase and *t*-phase have highly similar lattice structures, the *o*-phase can be regarded as a distorted variant of the *t* unit cell, separated by a very low energy barrier. Therefore, during cooling, this system can relatively easily transition from the *t* phase to the *o* phase. Unlike the process of transforming into the *o* phase, when the system further transforms from the *o* phase to the monoclinic (*m*) phase, it must overcome a higher energy barrier, and this barrier increases with growing volume. In bulk materials, the system possesses sufficient energy and space to complete this transformation, ultimately reaching a stable monoclinic phase.

However, in nanoscale films, multiple factors influence this process. This chapter will discuss how external factors such as doping, mechanical stress, and size effects affect the relative energy levels of different crystalline phases and the transformation energy barriers. These factors synergistically suppress the transformation from the metastable ferroelectric *o*-phase to the non-ferroelectric *m*-phase, enabling it to maintain a relatively stable state at room temperature. Thus, the ferroelectricity of HfO_2 is fundamentally a kinetically stabilized metastable phenomenon. To achieve high-performance, reliable HfO_2 -based ferroelectric devices, precise control over this mechanism is essential. Recent studies indicate that defect engineering and interface manipulation represent key approaches for stabilizing the metastable orthorhombic phase and mitigating fatigue[35].

3.2 Physical and Theoretical Understanding of the Polarization of Orthorhombic HfO_2

3.2.1 Fundamental Origins of Ferroelectricity, Dipoles, and Piezoelectric Coupling

As mentioned earlier, the ferroelectricity of HfO_2 actually originates from the orthorhombic $Pca2_1$ phase, which lacks center symmetry, as shown in Figure 3(b); in contrast, the cubic phase is a highly symmetric, nonpolar paraelectric phase [7], as shown in Figure 3(a). When the oxygen and hafnium sublattices undergo relative displacement within the unit cell, the centers of positive and negative charges no longer coincide, leading to the formation of a spontaneous electric dipole moment in the system. From a microscopic perspective, these dipoles arrange themselves in an ordered manner within a certain region, resulting in polarization.

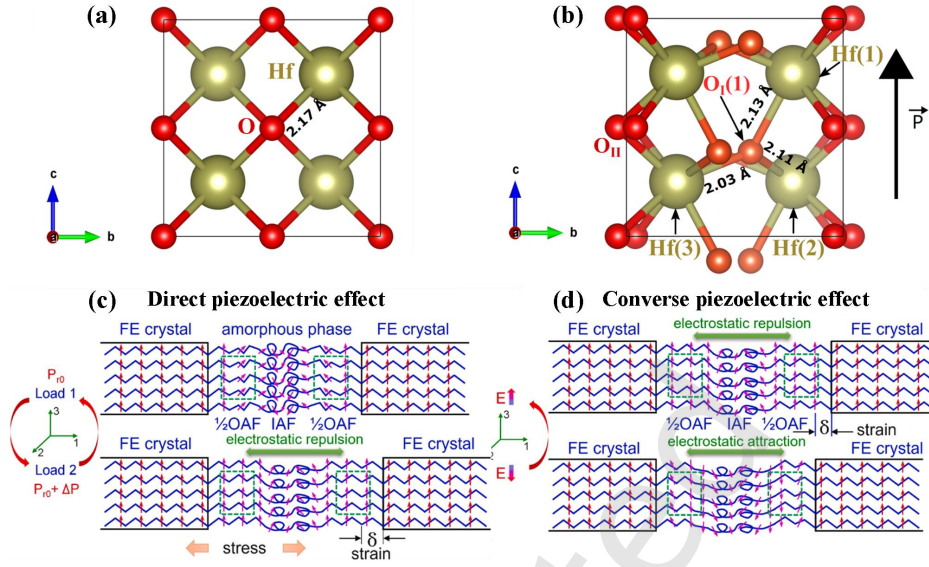


Figure 3 (a) Cubic paraelectric phase of HfO_2 [36]; (b) Orthorhombic ferroelectric phase of HfO_2 [36]; (c) Direct piezoelectric effect [37]; (d) Converse piezoelectric effect [37]

Generally speaking, ferroelectricity can be regarded as a special state of polar order. In this state, the dipole-based structure can remain stable in the absence of an external electric field. However, when a sufficiently strong external electric field is applied, the structure may undergo a reversal, meaning that the direction of polarization also flips. Of course, ordinary polar materials also exist in nature, but there is a fundamental difference between their polarization and that of ferroelectric materials: the polarization of the former cannot be reversed, whereas that of the latter can be switched. Upon further investigation into the origin of polarization reversal, the fundamental reason for the switchability of the orthorhombic phase HfO_2 lies in the presence of polarization states associated with symmetry in its free energy surface, and the energy barriers between these states are surmountable. Under an applied electric field, the polarization can be reversibly switched between symmetry-related polar states [38].

Similar to ferroelectricity, the orthorhombic phase of HfO_2 lacks inversion symmetry, resulting in a coupling relationship between mechanical deformation and electrical polarization [36]. When the orthorhombic phase of HfO_2 is subjected to external mechanical stimulation, it exhibits an electrical polarization response, as shown in Figure 3(c). Similarly, when an external electric field is applied, it induces lattice deformation, as shown in Figure 3(d). These correspond to the direct piezoelectric effect and the converse piezoelectric effect respectively, which are collectively referred to as the piezoelectric coupling effect. This demonstrates that the piezoelectricity and ferroelectricity observed in the orthorhombic HfO_2 phase are two interrelated manifestations of the same non-centrosymmetric crystal structure under different external fields.

In HfO_2 thin films, the piezoelectric response differs from that of conventional bulk piezoelectric ceramics and is far more complex. In actual thin films, the measured response is not entirely determined by the ideal crystal lattice itself, but is simultaneously influenced by various factors, such as domain wall motion, incomplete polarization inversion, interfacial constraints, residual stress, and the proportion of competing crystal phases. Lübben et al. conducted tests and analyses on the piezoelectric properties of HfO_2 - CeO_2 thin films. Their study demonstrated that both substrate stress and ferroelastic domain wall stability influence the piezoelectric characteristics [39]. This further confirms that the piezoelectric response is, in fact, a process involving the synergistic interaction of polarization, structure, and strain.

In actual devices, the piezoelectric response of thin films is not a fixed constant determined by a single crystal phase. It is generally believed to be closely related to the orthorhombic-phase fraction, polarization strength, stress state, thickness, and defect distribution. Tan et al. used a PFM system to investigate the effects of La and Zr doping, substrate stress, and thickness on the piezoelectric response of epitaxial doped HfO_2 films. The results indicate that the piezoelectric response of the film is primarily related to the polarization strength, which is mainly determined by the orthorhombic phase content; additionally, doping, stress, and thickness further modulate the intensity of the piezoelectric response [40]. This serves as a crucial foundation for the neuromorphic devices discussed in this paper. By altering the polarization state and regulating the internal electrostatic potential, external mechanical stimuli can further modify the storage state and conductive properties of FeFETs or related devices. For neuromorphic computing, the electromechanical coupling effect offers a novel possibility—an opportunity to directly integrate sensing, storage, and computation within the same material.

3.2.2 First Principles, Defects, and Dynamic Evolution

The previous section provided a detailed description of the nature of spontaneous polarization, offering an intuitive picture for understanding ferroelectricity. This section will adopt a modern computational perspective, particularly first-principles calculations based on density functional theory (DFT), to gain a deeper understanding of HfO_2 -based ferroelectrics. The theoretical foundation of DFT was established by Hohenberg and Kohn in 1964, and subsequently, Kohn and Sham proposed a self-consistent equation form that could be practically solved in 1965 [41]. Here, DFT is primarily used to calculate the relative stability of different crystal phases of HfO_2 , the origin of polarization, the electronic structure, and the piezoelectric response. Since the ferroelectric orthorhombic phase is not the thermodynamic ground state of bulk HfO_2 , its emergence must be understood through the

subtle energy competition among multiple competing crystal phases. DFT provides the theoretical tool for this.

In early calculations, the monoclinic phase of HfO_2 (a symmetric structure) was found to be the lowest-energy state, while the ferroelectric orthorhombic phase was a metastable phase. This conclusion is consistent with the structural picture shown in Figure 3 above. A metastable phase is an unstable structure; under most processing conditions, it is difficult to render it a globally thermodynamically stable phase by relying solely on conventional stress or strain. However, if HfO_2 is made exceptionally thin — reaching the nanoscale—the energy effects at its surface and interfaces become significant. Therefore, when the size of HfO_2 is reduced to a certain extent, it becomes possible to stabilize the orthorhombic ferroelectric phase. This also confirms the observation discussed in the previous chapter that the ferroelectricity of HfO_2 typically manifests in thin films.

In addition, Dutta et al. investigated the piezoelectric properties of ferroelectric HfO_2 using first-principles calculations and piezoresponse force microscopy (PFM) experiments, confirming that it can exhibit a negative longitudinal piezoelectric effect [36]. The negative longitudinal piezoelectric effect is an anomalous manifestation of the piezoelectric effect. As shown in Figure 4(a)–(c), the relationship between polarization and strain in HfO_2 exhibits different states under different epitaxial strain conditions. When the film is under compressive strain, it exhibits a negative longitudinal piezoelectric response; when the film is close to zero strain, the piezoelectric stress constant is also nearly zero; however, when the film is under tensile strain, the piezoelectric stress constant is positive, resulting in a positive piezoelectric response. Further studies have revealed that this anomalous piezoelectric response is related to the chemical coordination environment of active oxygen atoms [36].

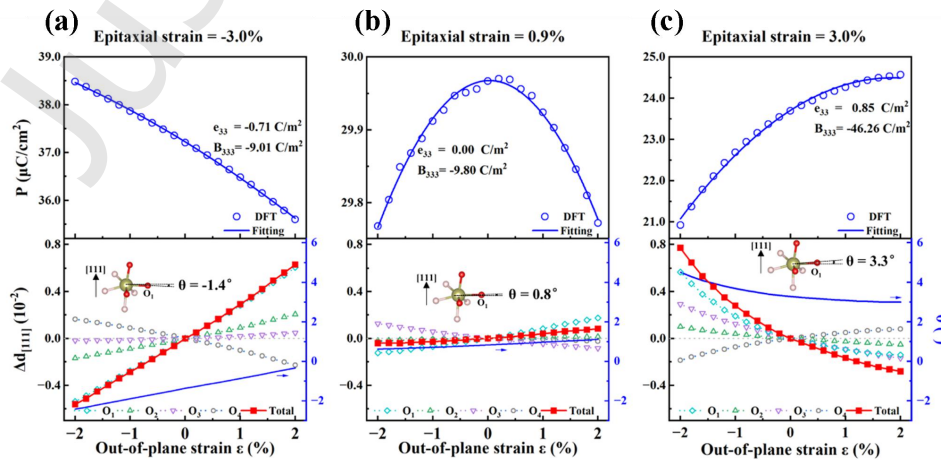


Figure 4 Epitaxial strain imposed on the film by the substrate [42]: (a) compressive strain, (b) near-zero strain, (c) tensile strain

First-principles calculations can also provide a detailed analysis of the specific roles played by various defects in HfO₂ ferroelectrics, particularly oxygen vacancy defects. In HfO₂-based ferroelectric materials, these defects are not merely parasitic with no positive effects; rather, they exert a bidirectional influence on the performance of both the material and the device. Regarding positive effects, oxygen vacancy defects alter the local chemical bonding patterns and charge balance in HfO₂, and induce slight lattice distortions (as shown in Figure 5(b)). These changes affect the relative stability of the orthorhombic, tetragonal, and monoclinic phases of HfO₂. As shown in Figure 5(a), the ferroelectric phase gradually stabilizes as the oxygen vacancy concentration increases. Regarding the negative effects, oxygen vacancy defects lead to various reliability issues in HfO₂-based ferroelectric devices [43], including the wake-up, fatigue, and imprint effects, which will be discussed later in this chapter. Consequently, while oxygen vacancy defects may to some extent promote the formation and stabilization of the ferroelectric phase, they simultaneously undermine the long-term stability of the device through migration and redistribution during electrical cycling.

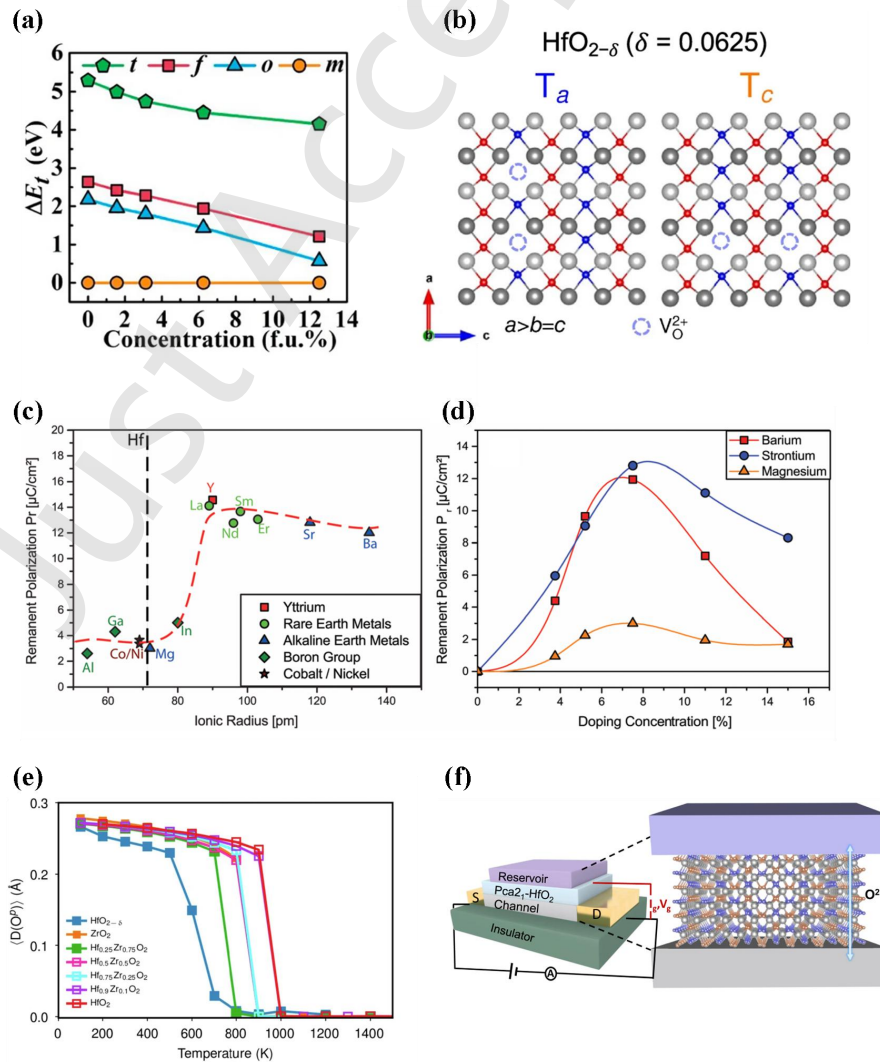


Figure 5: (a) Relative energies of different phases of HfO_2 as a function of oxygen vacancy concentration [44] (b) Representative defect structures in oxygen-deficient HfO_2 [44] (c) Remanent polarization as a function of dopant ion radius [45] (d) Remanent polarization of the dopant as a function of doping concentration [45] (e) Polar order parameter derived from molecular dynamics simulations as a function of temperature [44] (f) Electrochemical ionic synapse device with a Pca2_1 - HfO_2 electrolyte layer [46]

In addition to point defects, modern computational studies have also focused on various doping effects. Due to differences in ionic radius, valence state, and local chemical environment, different dopant elements alter the structure of HfO_2 in distinct ways. Clear doping patterns have also been observed experimentally. As shown in Figure 5(c), dopants with larger ionic radii, such as alkaline earth metals and rare earth metals, can significantly enhance the remanent polarization of HfO_2 films. As shown in Figure 5(d), the polarization intensity reaches its maximum when the doping concentration is between 5% and 8%, and it first increases and then decreases with concentration. Therefore, the influence of doping effects on the polarization intensity cannot be ignored.

It should be noted that traditional DFT remains, in essence, a quasi-static formalism. It is not well-suited for describing the dynamic structural evolution of defect migration over long time scales. In light of this, molecular dynamics (MD) simulations and related multiscale methods have become increasingly important in recent years. Unlike first-principles calculations, MD simulations can capture dynamic information about materials over time, such as atomic thermal motion induced by temperature, the diffusion and movement of defects such as oxygen vacancy defects, and the local atomic arrangement and distribution within the material. As shown in Figure 5(e), based on such molecular dynamics simulations, starting from the Pca2_1 polar phase, the system can transform into different nonpolar phases as the temperature increases, clearly demonstrating the structural evolution of HfO_2 's ferroelectric behavior over time.

Therefore, from the perspective of modern theory, it is essential to combine computational methods with different focuses, leveraging the strengths of each and conducting a comprehensive analysis, in order to establish a complete theoretical understanding of the HfO_2 ferroelectric. Of the three methods described above, first-principles calculations (such as DFT) are used to study the intrinsic properties of ideal, static materials at zero temperature, including various crystal phases with different energies and the origin of ferroelectric polarization; calculations related to doping and oxygen vacancy defects focus on local lattice distortions and regulation; while molecular dynamics and multiscale simulations are used to study non-equilibrium dynamic processes with time evolution. Only by integrating these methods can we gain a clearer and more comprehensive understanding of HfO_2 -based ferroelectrics.

As shown in Figure 5(f), based on the coupling between ferroelectric polarization in orthorhombic HfO_2 and oxygen-ion transport, researchers have successfully developed a silicon-compatible neuromorphic device. Building on this foundation, the next section will further discuss how to modulate the ferroelectric properties of HfO_2 through doping, stress, and size effects.

3.3 Ferroelectric Phase Engineering: Synergistic Regulation via Doping, Stress, and Size Effects

Since the ferroelectricity of HfO_2 originates from its metastable orthorhombic phase, relying solely on dynamic trapping during device fabrication is insufficient. This approach also compromises reliability and narrows the process window. To actively and controllably adjust the energy levels between different crystal phases, ensuring stability at room temperature and maximizing the proportion of the ferroelectric orthorhombic phase, we must investigate how to regulate the ferroelectric phase. Consequently, the academic community has developed a relatively mature methodology primarily based on synergistically utilizing three key factors: doping, stress, and size effects. Recently, researchers discovered that pseudo-binary ferroelectrics and hafnium-doped systems can effectively control CMOS-compatible polarization[47]. The primary approach is doping, which regulates phase transitions through lattice distortion. The most direct and effective experimental method involves introducing heteroatoms into the HfO_2 lattice. This is primarily because the dopant atoms, differing in size or valence from Hf^{4+} ions, induce localized lattice distortion, thereby altering the relative stability of different crystalline phases.

Zirconium (Zr) doping is relatively well-established in the industry, as it forms $\text{Hf}_{1-x}\text{Zr}_x\text{O}_2$ solid solutions. Within this solid solution, the ionic radius of Zr^{4+} is slightly smaller than that of Hf^{4+} . Consequently, Zr doping induces lattice contraction, effectively lowering the formation energy of the highly symmetric tetragonal phase. This allows the tetragonal phase to remain stable at lower temperatures. Because the ferroelectric orthorhombic phase is structurally closer to the tetragonal phase than to the monoclinic ground state, Zr incorporation stabilizes a tetragonal-like precursor and suppresses transformation to the non-ferroelectric monoclinic phase, thereby broadening the process window for ferroelectricity, thereby broadening the process window for achieving ferroelectricity[7]. Other dopants, such as silicon (Si), aluminum (Al), and lanthanide elements (La, Gd, etc.), are also highly effective. Silicon, being a relatively small tetravalent ion, significantly enhances the remanent polarization when doped. However, its process window is relatively narrow; excessive Si doping readily forms amorphous silicate phases at grain boundaries, thereby degrading ferroelectricity[7]. Aluminum, being a trivalent ion, induces size-dependent effects upon Al^{3+} doping while introducing oxygen vacancies to maintain charge balance. Moderate aluminum doping can stabilize the ferroelectric phase and enhance device durability[25]. Regarding lanthanide elements, these are large trivalent ions that, upon doping, tend to stabilize highly symmetric cubic or tetragonal phases, macroscopically preventing their transformation into monoclinic phases. This approach, often termed kinetic stabilization, primarily relies on stabilizing the high-temperature phase to

suppress the formation of non-ferroelectric ground states[48].The second major factor is stress, which suppresses phase transitions through mechanical boundary conditions. Its mechanism lies in the volume change accompanying phase transitions. For instance, during the transformation from the t-phase or o-phase to the m-phase, the volume expands by approximately 4%.

In a standard metal–ferroelectric–metal (MFM) structure, a HfO_2 film lies between two TiN electrodes. During cooling from high-temperature annealing to room temperature, the TiN electrodes undergo significant contraction due to their much higher thermal expansion coefficient compared to HfO_2 . This contraction induces substantial biaxial tensile stress in the HfO_2 layer [25]. Park et al. directly measured this capping stress via X-ray diffraction and confirmed its strong correlation with ferroelectric phase stability [49]. This tensile effect constrains the film's in-plane deformation, hindering volumetric expansion and thereby influencing the monoclinic phase transition process.

The third major point is the size/thickness effect, whose primary mechanism involves surface energy-driven phase stability reversal. Conventional perovskite ferroelectrics lose their ferroelectricity at ultra-thin thicknesses due to depolarization effects. However, HfO_2 exhibits enhanced ferroelectricity when its thickness is reduced below 10 nanometers. This unique "anomalous size effect" in HfO_2 is precisely what current advanced process node scaling requires. Currently, researchers have even observed ferroelectricity in laboratory-grown films thinner than 3 nanometers[27].

This mechanism stems from the competition between surface/interface energy and bulk free energy. While the monoclinic phase possesses the lowest bulk free energy, its symmetry is lower, resulting in relatively higher surface and interface energies. Conversely, the highly symmetric orthorhombic (o) and tetragonal (t) phases exhibit lower surface/interface energies. When film thickness reaches the nanoscale, the surface-to-volume ratio increases dramatically, causing surface energy to dominate. Below a specific critical thickness, the surface/interface-energy contribution can compensate for the bulk free-energy disadvantage, making the metastable ferroelectric orthorhombic phase energetically more competitive than the monoclinic phase. This phenomenon stands in sharp contrast to the Kay–Dunn scaling law. The Kay–Dunn relationship indicates that the coercive electric field follows $E_c \propto d^{-2/3}$ and predicts that ultrathin ferroelectric films will lose ferroelectricity[50]. However, the presence of HfO_2 completely overturns this view.

In the actual device fabrication process, these three effects work synergistically to ensure the formation of the ferroelectric phase. A typical process flow involves first depositing an ultra-thin (less than 10 nm) film with a specific composition using atomic layer deposition. Next, top and bottom TiN electrodes are deposited to encapsulate the film. The final step is rapid thermal annealing to crystallize the film. This process flow clearly demonstrates the synergistic effect. During the cooling process, the favorable phase transition pathway is determined by doping, while stress suppresses undesirable transformations. At the nanoscale, the ferroelectric orthorhombic phase can maintain an effective free-energy advantage over the monoclinic phase. It is precisely the synergistic interaction of these three mechanisms that sustains a stable and reproducible process window, enabling HfO_2 -based ferroelectric

devices to transition from laboratory research to industrial production.

3.4 Materials-Level Challenges: Physical Origins of Wake-Up, Fatigue, and Imprint Effects

Although tuning this ferroelectric phase enables the fabrication of high-performance HfO_2 films, certain undesirable effects persist during device cycling and long-term data retention, such as wake-up effects, fatigue effects, and imprint effects. The presence of these effects significantly compromises device reliability. Therefore, investigating their physical origins can facilitate the development of highly reliable ferroelectric devices. Both academia and industry widely recognize that these effects stem from intrinsic defects within HfO_2 , particularly the migration, aggregation, and interactions of oxygen vacancies at interfaces and domain walls.

The primary effect is the wake-up effect. In HfO_2 -based ferroelectric devices, the initial polarization-electric field (P-E) hysteresis loop typically exhibits low remanent polarization (P_r) and high coercive field (E_c). However, after a certain number of electrical cycling operations, the P-E curve gradually changes: P_r increases while E_c decreases, ultimately approaching a saturated, nearly rectangular shape. This performance optimization process is the wake-up effect. To mitigate this undesirable phenomenon, Samanta et al. (2021) demonstrated that multi-step low-temperature annealing under oxygen atmosphere can reduce oxygen-vacancy-related defects and realize wake-up-free ferroelectricity in Y: HfO_2 films, confirming that defect redistribution is the dominant origin of the wake-up effect[51].

The wake-up effect is generally attributed to electric-field-driven redistribution of oxygen vacancies and domain depinning. In some HfO_2 -based films, it may also be accompanied by a transition from an antiferroelectric-like or non-ferroelectric state to a ferroelectric state[52]. Chouprik et al. (2019) employed PFM to investigate the domain structure evolution in $\text{Hf}_{0.5}\text{Zr}_{0.5}\text{O}_2$ films during the wake-up process. Their experiments demonstrated that the remanent polarization accumulates and strengthens with cycling cycles, providing direct evidence for domain depinning[53]. During thin-film deposition and annealing, devices inevitably form numerous point defects, with oxygen vacancies being the most abundant. In the early stages of preparation, these positively charged oxygen vacancies are readily captured by electrically neutral domain walls and grain boundaries, generating local electric fields. Under applied bias, this enables suppressed domains to switch. Phase field simulations similarly explain the wake-up effect: during electrical cycling, strong electric fields provide sufficient energy for defects to redistribute and migrate to more stable positions, ultimately releasing pinned domain walls. Although the wake-up effect can optimize device performance, it inherently stems from initial instability in the device, increasing design complexity.

The second major effect is fatigue. After undergoing numerous switching cycles, the device's performance gradually degrades: the remanent polarization (P_r) weakens, the P-E hysteresis loop narrows or distorts, ultimately closing the memory window and causing device failure. This degradation phenomenon is termed the fatigue effect. The fundamental mechanisms of fatigue primarily involve charge injection and defect pinning[30]. During

each polarization reversal, high switching currents and strong electric fields cause charge injection from the electrodes into the ferroelectric layer. This injected charge is captured by defects, forming charged defect nuclei. As cycling continues, captured charge eventually accumulates near domain walls or interfaces. This generates a strong built-in bias field, stabilizing the ferroelectric domains. When the number of pinned domains reaches a critical threshold, the macroscopic remanent polarization value significantly decreases, causing fatigue effect failure. Additionally, charge injection may induce new defects at interfaces, leading to the formation of dead zones that progressively degrade performance. Simultaneously, the quality of the interface SiO_2 layer significantly influences the suppression of charge injection and the enhancement of fatigue endurance[54].

The third major reliability issue is the imprint effect. The imprint effect primarily impacts data retention capability. This refers to the phenomenon where the P-E hysteresis loop shifts along the direction of the electric field-induced polarization state. This shift causes an asymmetric distribution of the coercive field, making it increasingly difficult for the device to switch to the opposite state. When the built-in bias field is strong enough, it can trigger spontaneous polarization reversal back to the initial state even after the external electric field is removed, leading to data loss. To mitigate this effect, structural design modifications and material optimization are required. The industry employs a metal-ferroelectric-metal-insulator-semiconductor (MFMIS) structure. By rationally designing the area ratio between the ferroelectric capacitance and the MOS capacitance, the depolarization field can be effectively suppressed, thereby enhancing data retention capability[55].

The imprint effect and fatigue effect share similar characteristics, both driven by the establishment of an internal bias field through defect migration. However, the driving force behind the imprint effect originates not from an alternating electric field, but from the intrinsic depolarization field and polarized charge within the ferroelectric material. When the ferroelectric layer becomes polarized, a significant amount of bound polarized charge accumulates at its surface, generating an internal depolarization field opposite to the polarization direction. Under this field, mobile charge defects gradually drift and redistribute to shield the depolarization field. Their ordered distribution forms an internal bias field that effectively stabilizes the device's current state[56]. Temperature-accelerated testing indicates that the imprint effect is the primary factor limiting the long-term data retention capability of ferroelectric field-effect transistors[57]. As summarized in Figure 6, defect redistribution during electrical cycling can lead to wake-up, fatigue and imprint.

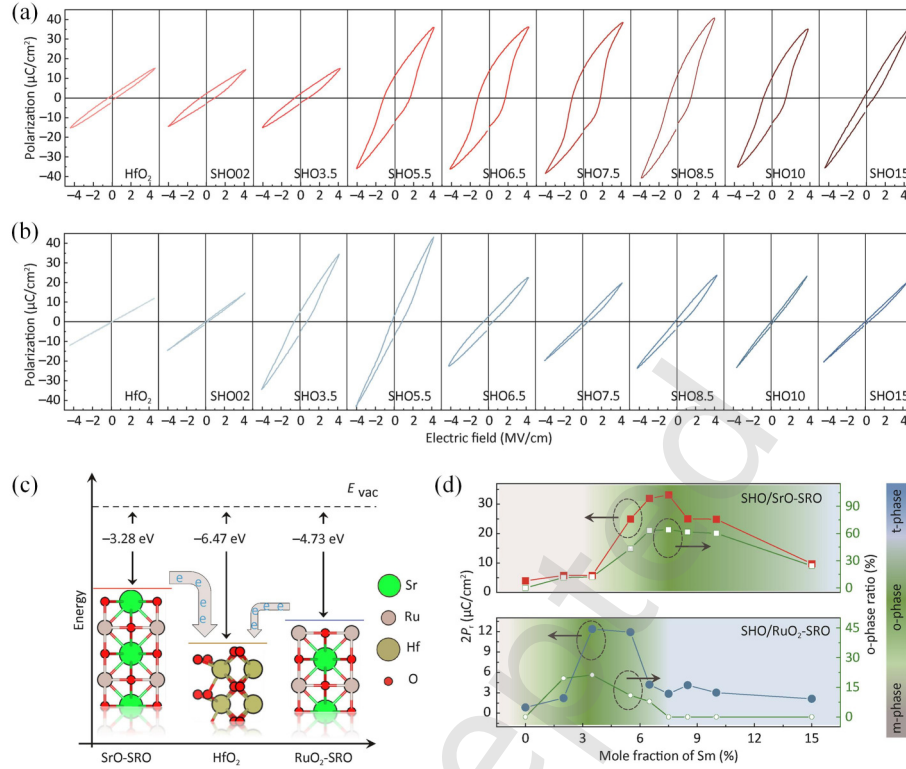


Figure 6 Illustration of wake-up, fatigue, and imprint effects in ferroelectric materials[58]. (a) Initial weak polarization state before cycling. (b) Wake-up process showing polarization enhancement. (c) Fatigue effect with polarization degradation after many cycle, (d) Imprint effect represented by a shifted hysteresis loop due to internal bias.

In the HfO₂ -based materials, point defects such as oxygen vacancies not only synergize with dopants to stabilize the ferroelectric phase, but also migrate and redistribute under electric fields. This migration is the fundamental cause of the wake-up effect, fatigue effect, and imprint effect. To enhance the stability and reliability of HfO₂ -based ferroelectric devices, the industry must optimize fabrication processes to control both the initial defect concentration and its migration.

Therefore, understanding and further controlling these defect-related phenomena determines the stability, programmability, and retention capability of ferroelectric devices. Building upon this material foundation, we will next explore how to transform HfO₂ -based ferroelectric materials into functional synaptic units and neuronal units.

4 Ferroelectric Devices: From Units to Systems

After exploring the ferroelectric physics underlying HfO₂ -based materials, this chapter investigates how these materials can be utilized to construct functional devices and introduces their applications in neuromorphic computing. Furthermore, starting with the most fundamental FeFET, this chapter analyzes the operating principles and structure of the transistor, along with its advantages and limitations in simulating biological synapses. Finally, this chapter will explain similar devices such as FTJs, analyze how these devices can implement fundamental neuronal functions, and construct a complete neuromorphic system.

4.1 Ferroelectric Field-Effect Transistor

The most extensively studied HfO_2 -based neuromorphic device is the FeFET. Additionally, it features various structures, such as ultra-thin ITO-channel flexible FeFETs, which exhibit excellent durability and transparency, making them suitable for wearable edge AI systems[59]. Essentially, it remains a standard MOSFET where the gate insulator is wholly or partially replaced by ferroelectric material. By integrating the nonvolatile polarization characteristics of ferroelectrics with the field-effect switching properties of transistors, a three-terminal controllable unit is formed, integrating nonvolatile storage and computational functions.

4.1.1 Operating Principle and Gate Stack Structures

The core operating principle of the FeFET lies in the ability of the ferroelectric gate dielectric's remanent polarization (P_r) to induce a persistent, non-volatile electrostatic field in the semiconductor channel, thereby modulating the transistor's threshold voltage (V_{th}). As shown in Figure 7, when a positive write pulse is applied to the gate of an n-channel FeFET, the ferroelectric layer becomes polarized, generating a remanent polarization directed toward the semiconductor channel. This polarization field attracts a large number of electrons into the channel, ultimately enabling the transistor to conduct at a lower gate voltage. This low threshold voltage state can be defined as logic "1" or a strong synaptic weight.

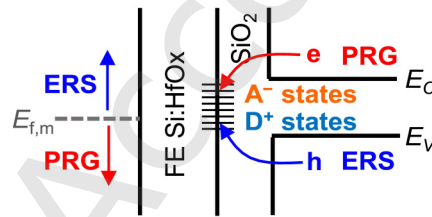


Figure 7 Energy band diagram and operating states of a FeFET[60]

When a negative write pulse is applied to the gate, the polarization direction reverses, generating a remanent polarization away from the channel. This polarization repels electrons from the channel, raising the transistor's conduction voltage. This high-threshold voltage state can be defined as logic "0" or a weak synaptic weight. Due to the non-volatile nature of ferroelectric polarization, this threshold voltage difference persists after the write pulse is removed, enabling non-volatile storage. The difference in threshold voltage between the two states is called the memory window, which is regarded as a key performance metric for the FeFET. In practical, functional devices, two mainstream gate stack structures are predominantly used: the metal–ferroelectric–insulator–semiconductor (MFIS) gate stack and the metal–ferroelectric–metal–insulator–semiconductor (MFMIS) gate stack, each with its own advantages and considerations. This is the simplest structure. To build it, one must first form an ultra-thin insulating interface layer on the semiconductor channel surface, followed by deposition of the ferroelectric layer and metal gate. The insulating intermediate layer prevents interface reactions and charge injection between the ferroelectric material and the semiconductor, thereby ensuring this switching characteristic. The primary advantage of this structure lies in its compact design, which closely resembles standard high-k metal gate processes, enabling high integration density. However, uncertainties during ferroelectric layer

deposition and subsequent high-temperature annealing may damage the underlying semiconductor channel or introduce defects, potentially compromising device performance and reliability.

In order to overcome the process barriers of MFIS, the MFMIS architecture was developed, and in this structure, the device is divided into two independent parts: a standard high-performance MOSFET and a separate MFM capacitor, where the MOSFET can first be fabricated using conventional processes, and then the independently prepared MFM capacitor is connected to the MOSFET gate through back-end metallization. The primary advantage of this approach lies in process decoupling: the high-temperature annealing required for magnetic field gate capacitors does not interfere with the normal operation of the already-formed MOSFETs, thereby ensuring optimal performance for both components. This structure has been validated in HfO_2 -based device research and represents a viable pathway for achieving high-density back-end integrated memory. With complementary engineering implemented via optimized interfacial layers and read-voltage control, the retention characteristics of FeFETs for neuromorphic computing are further improved[61], and its drawback is that extra metal layers and interconnections are needed, making the structure more complex and possibly causing a reduction in integration density, and this configuration is also referred to as a series FeFET, and optimizing such a series structure is a key research direction, and by adjusting the thickness and area of the series ferroelectric capacitor, the capacitance matching of the device can be actively corrected so as to maximize the memory window[62].

4.1.2 Synaptic Plasticity: Optimization of Linearity and Symmetry

The core advantage of FeFETs acting as electronic synapses is their multi-level storage capability, unlike digital memories that only have binary "0/1" states, the macroscopic polarization of ferroelectric materials is in fact the accumulated result of reversals of numerous microscopic domains, and by finely tuning the amplitude, width, or number of gate-voltage pulses, the proportion of switched domains can be delicately adjusted, thereby obtaining multiple—usually quasi-continuous—intermediate polarization states between fully positive polarization (P_r^+) and fully negative polarization (P_r^-), and these intermediate states correspond to different threshold voltages, enabling the channel conductance (G) of the FeFET to emulate the multi-level plasticity of biological synaptic weights[63].

However, synaptic weight updates must also satisfy the requirements of linearity and symmetry to enable efficient brain-inspired hardware based on FeFETs. In terms of linearity, from an ideal perspective, the change in synaptic weights should remain constant when a series of identical enhancement pulses are applied. However, in actual devices, the conductivity variation of FeFETs is nonlinear, exhibiting significant changes during the initial phase and diminishing as saturation approaches. This nonlinearity impairs the training convergence of hardware neural networks and reduces their accuracy. To evaluate this non-ideal effect, researchers developed a framework embedding device physics models into neural network simulations[64].

Regarding symmetry, under ideal conditions, the enhanced conductivity induced by the excitation pulse would be counterbalanced by the suppression pulse with opposite parameters.

However, in actual devices, the complex charge capture and release dynamics within the FeFET result in asymmetric pathways for both enhancement and suppression. The nonlinearity and asymmetry under identical pulse trains are evident in Figure 8(a)-(c), while incremental pulse schemes can improve update linearity and symmetry (Figure 8(d)-(f)).

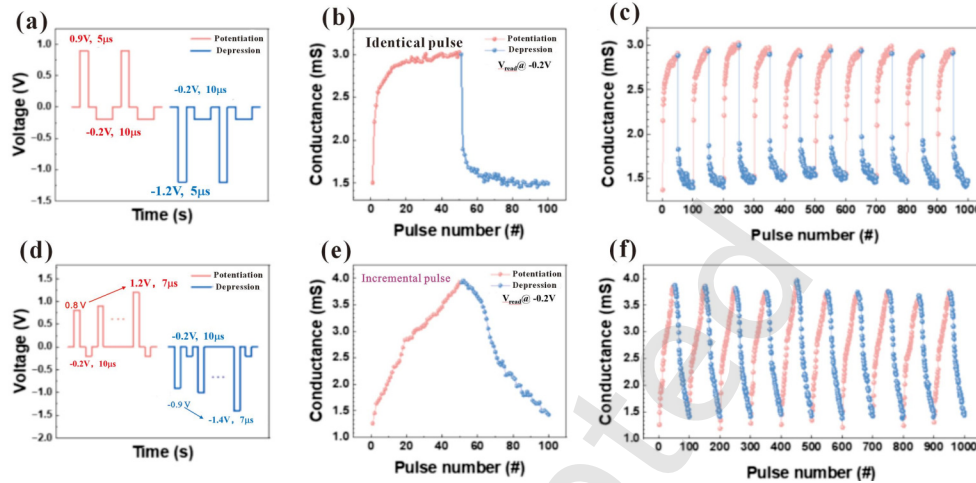


Figure 8 Pulse-scheme engineering for FeFET synaptic weight update[65]. (a) Identical potentiation/depression pulse trains. (b) Conductance update under identical pulses (nonlinearity/asymmetry). (c) Endurance/repeatability of conductance modulation under identical pulses. (d) Incremental (staircase) potentiation/depression pulse trains. (e) Conductance update under incremental pulses (improved linearity/symmetry). (f) Endurance/repeatability of conductance modulation under incremental pulses.

To handle these non-ideal phenomena, researchers have put forward a variety of pulse-optimization schemes, Wu et al. proposed and verified at the array level a double-pulse (write + verify) programming strategy for FeFET arrays, and this strategy can improve the linearity of weight updates and suppress write disturbance[16], and the incremental staircase-pulse method—namely gradually increasing the pulse amplitude or width—has been proven to improve update linearity to a certain extent[66]. Research also claims that applying small-amplitude reverse pulses after the main programming pulse can release trapped charges at the interface, and thus enhance update symmetry[67], and developing hardware-compatible programming schemes that can effectively cope with device non-idealities is a key step in pushing FeFETs from single-device demonstrations toward large-scale array applications, and recently, results on Si-doped HfO₂ FeFETs showing up to 64 distinguishable conductance states have further validated their potential to achieve analog-like synaptic behavior in brain-inspired computing[68].

4.2 Diversity of Ferroelectric Devices

Although the FeFET is the most extensively studied HfO₂-based neuromorphic device, it is not the only option. To meet diverse requirements for density, speed, and power consumption across different application scenarios, researchers have developed multiple device architectures based on ferroelectric physics. Many representative device development directions have currently been formed, including two-terminal ferroelectric devices

represented by FTJs and Fe-diodes, as well as FeFETs integrated with advanced transistor architectures. These devices differ in their operating mechanisms, readout methods, and neuromorphic applications. HZO-based FTJ structures, optimized for ultra-fast response and stable multistate operation, are well-suited for synapse emulation[69]. These technologies complement standard FeFETs, collectively forming a comprehensive HfO_2 -based ferroelectric device technology system.

4.2.1 FTJs, Fe-Diodes, and High-Density Integration

An FTJ is an extremely simple two-terminal device, typically consisting of an ultra-thin ferroelectric film sandwiched between two different metal electrodes. It primarily exploits the direct tunneling effect described by quantum mechanics. Figure 9 provides an overview of representative ferroelectric devices, including ferroelectric random-access memory (FeRAM), FeFETs, and FTJs. For FTJs, the polarization-dependent barrier modulation shown in Figure 9(c) directly gives rise to the tunneling electronic resistance effect. When the polarization points toward one electrode, the effective barrier height decreases, electrons can tunnel more easily, and the device is in a low-resistance state. When the polarization reverses and points toward the opposite electrode, the effective barrier height rises, tunneling becomes more difficult, and the device is in a high-resistance state. The resistance-modulation phenomenon induced by polarization switching is called the tunneling electroresistance effect, and since read operations are generally carried out at very low voltages, FTJs exhibit extremely low read power consumption, and Ma et al. demonstrated, under room-temperature conditions, sub-nanosecond (about 0.6 ns) resistance-state switching in FTJ memristors[70]. The key advantage of FTJs lies in their unparalleled integration density[71], as two-terminal devices, they can be embedded at metal-line cross-points into three-dimensional crossbar array structures, and in theory the unit-cell size can be scaled down to $4F^2$, far smaller than transistor-based structures that require isolation regions, and FTJs are therefore highly suitable for building ultra-high-density in-memory-computing arrays and next-generation storage-class memory, and through engineering optimization of electrode work functions, the on-current and energy efficiency of FTJs are further improved, thereby enhancing the neuromorphic computing capability of this device[72], and in addition, a parallel-FTJ architecture has been proposed to optimize linearity and reduce variability, thus improving the functionality of edge neuromorphic devices[73].

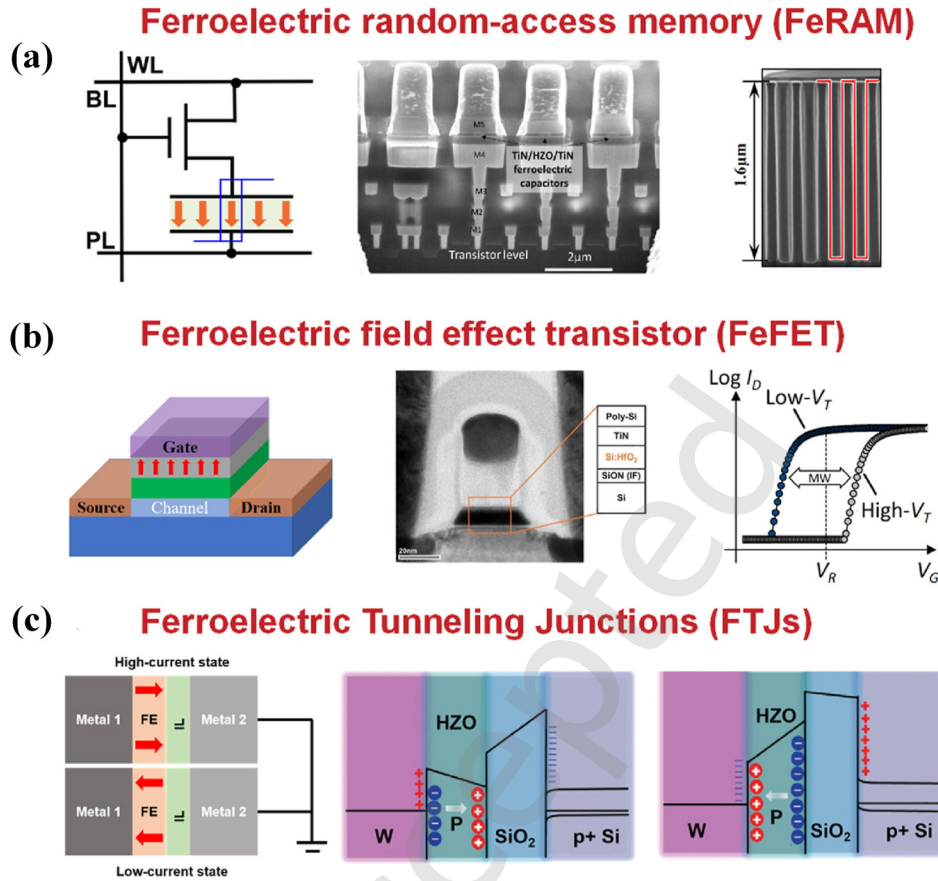


Figure 9 Representative ferroelectric devices enabled by HfO_2 ferroelectricity[74]. Top: dielectric-to-ferroelectric transition in HfO_2 (phase change and P - E hysteresis emergence enabled by annealing/doping/stress). (a) Ferroelectric random-access memory (FeRAM): 1T - 1C cell concept and example integration/structure. (b) Ferroelectric field-effect transistor (FeFET): device concept/stack cross-section and memory-window (threshold-shift) behavior. (c) Ferroelectric tunneling junctions (FTJs): polarization-controlled tunneling barrier and high-/low-current states.

However, FTJs have certain limitations: the tunnel current is extremely sensitive to the thickness and quality of the ferroelectric layer. Minor thickness variations or defects can cause significant performance changes, making it difficult to achieve consistent mass production. HfO_2 -based FTJs typically exhibit a low on/off current ratio, limiting their precision in certain analog neuromorphic applications. For instance, they struggle to meet the requirements of applications demanding a wide dynamic range. To address this issue, researchers adopted a layer-by-layer structural design approach to optimize device performance.

In addition to FTJs, in recent years, hafnium-based Fe-diodes have also been considered as an option for high-density information storage and neuromorphic computing. Different from FTJs, Fe-diodes combine switchable ferroelectric polarization with rectifying transport characteristics. Due to this self-rectifying behavior, under the selected bias, they can generate a relatively large read current, while under reverse or unselected bias, they maintain a low current. Therefore, the sneak current in cross-point arrays can be greatly suppressed. Luo et al. investigated a hafnium-based Fe-diode with high CMOS compatibility, demonstrating the

feasibility of integrating Fe-diode cells into low-power, high-density memory hardware[17]. In addition, field-programmable Fe-diodes can be applied to Bayesian neural network applications to construct complementary memory cells[75]. These research results indicate that the intrinsic rectifying characteristics of Fe-diodes and their array-level advantages can effectively complement FTJs, further enabling hardware for highly energy-efficient neuromorphic computing.

4.2.2 Advanced FET Structures: FinFETs, Nanowires, and Junctionless Devices

Current CMOS process nodes have shrunk below 10 nanometers, where traditional planar MOSFET structures suffer from severe short-channel effects, leading to their gradual obsolescence. In contrast, three-dimensional transistor architectures such as fin field-effect transistors (FinFETs) and gate-all-around (GAA) nanowires have gained traction. Consequently, integrating ferroelectric gate stacks with these advanced structures represents an inevitable trend, promising significant performance enhancements.

In FinFET-based components or nanowire-device configurations, the gate wraps the channel from three sides or four sides, thereby markedly enhancing electrostatic control capability, and this stronger and more powerful gate-control effect can effectively suppress short-channel effects such as barrier lowering and subthreshold-swing degradation[76], and when a ferroelectric layer is incorporated into these three-dimensional structures, its polarization electric field can more effectively modulate the potential of the entire channel, so that a stable and wide memory window can be achieved even in ultra-short-channel devices, and the excellent electrostatic characteristics can also reduce leakage current and improve subthreshold behavior, thereby lowering both static and dynamic energy consumption.

Traditional MOSFETs form PN junctions by using heavily doped source/drain regions, and at the nanoscale it is highly challenging to create steep doping gradients, and the junctionless transistor first experimentally demonstrated by Colinge et al. in 2010[77], solves this problem by adopting a uniformly doped nanowire channel that is fully depleted by the gate in the off state[76], and when a ferroelectric gate is integrated, its polarization can assist the depletion or accumulation of the channel, thereby achieving nonvolatile switching, and this structure simplifies the device fabrication flow and shows extraordinary potential at sub-5-nm technology nodes. Researchers have already realized wafer-scale graphene-HfO₂ hybrid ferroelectric transistors, and these devices exhibit three-terminal memristive behavior suitable for neuromorphic computing [78], and extending this concept to junctionless ferroelectric FET architectures based on two-dimensional materials may push device scaling toward the physical limit[79].

Currently, ferroelectric devices come in numerous types, offering brain-inspired hardware designers a wide array of choices and combinations. However, the industry's mainstream choice remains FeFETs, which leverage mature processes and excellent analog characteristics. Additionally, FTJs demonstrate strong performance in ultra-high-density and memory-computing integration applications. Ferroelectric devices based on advanced structures like FinFETs and nanowires represent the future direction for cutting-edge process nodes. Process selection must balance these devices according to specific application requirements to ultimately achieve optimal system-level performance.

4.3 Hardware Realization of Neuronal Functions

Considering only the impact of devices makes it difficult to form a complete neuromorphic computing system. Therefore, the system also requires circuits with neuronal functionality. In biological neural networks, neurons serve as computational units capable of receiving and integrating signals from numerous synapses. When their internal state exceeds a certain threshold, neurons fire spikes as output signals while simultaneously resetting their state. This transition from integration to triggering constitutes the fundamental operating principle of spiking neural networks.

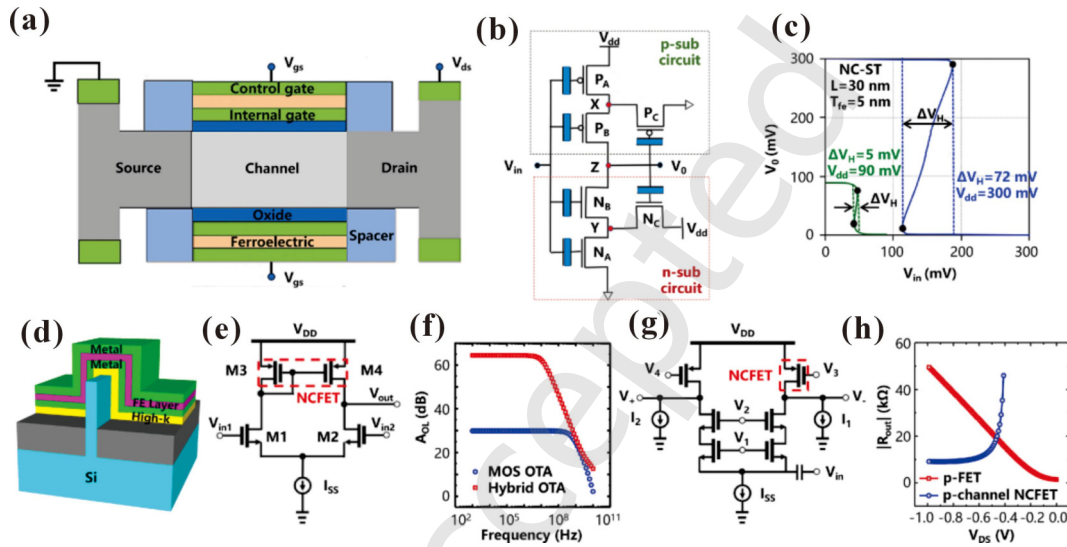


Figure 10 Voltage transfer and circuit-level demonstrations enabled by ferroelectric/negative-capacitance FETs[80]. (a) Cross-section schematic of FeFET structure with ferroelectric gate; (b) Complementary p-sub/n-sub Schmitt-trigger subcircuits; (c) Simulated voltage transfer curve (VTC) and hysteresis window; (d) 3D schematic of an integrated negative-capacitance field-effect transistor (NC-FET) device; (e) Circuit of hybrid NC-FET-based OTA; (f) Frequency response comparison between MOS and hybrid OTA; (g) NC-FET-based analog enhancement circuit; (h) Output resistance comparison of p-FET and p-channel NCFET versus V_{DS} .

Even though traditional CMOS circuits can construct fully functional integrate-and-fire (I&F) neurons, they generally require dozens of transistors, resulting in relatively large area and power overhead, and by leveraging the unique physical properties of ferroelectric devices, researchers have explored new routes for building ultra-compact, low-power neuron circuits so as to achieve higher-density integration with synaptic arrays, and constructing coupled-oscillator networks using the oscillation characteristics of ferroelectric devices has become a frontier academic exploration direction for realizing neuronal functions[81].

When designing neurons based on negative capacitance, the complex interactions within the ferroelectric layer must also be considered. Research utilizing segmented SPICE models for simulation has revealed that these interactions can alter the device's internal conduction current and saturation characteristics. In more severe cases, negative differential resistance (NDR) may even occur. These are all physical effects that must be accounted for in circuit design[82]. One characteristic of FeFETs is the hysteresis loop present in their input-output characteristics, enabling them to function in conjunction with Schmitt triggers. The Schmitt

trigger is a voltage comparator circuit with hysteresis characteristics that effectively suppresses noise in the input signal and provides a stable output signal. Research indicates that ultra-compact Schmitt triggers can be realized through simple circuit designs, such as combining a FeFET with a pull-up resistor or a CMOS inverter[83]. As shown in Figure 10(a)–(c), the FeFET-based Schmitt-trigger circuit exhibits hysteretic switching with distinct high/low thresholds, while NC-FET-enabled analog circuit demonstrations and metrics are summarized in Figure 10(d)–(h).

In neuromorphic applications, FeFET-based Schmitt trigger circuits can serve as discharge modules for I&F neurons. Here, the input voltage represents the accumulated membrane potential. When this voltage exceeds the threshold, the circuit generates a pulse to simulate the discharge behavior of biological neurons. Due to its extremely simple structure, this neural circuit offers lower area and static power consumption compared to traditional, relatively complex CMOS approaches.

In addition to hysteretic FeFET operation, ferroelectric gate stacks can also be engineered to operate in a negative-capacitance regime, which is typically discussed in negative-capacitance field-effect transistors (NC-FETs) or negative-capacitance-assisted circuits rather than in conventional memory-type FeFETs. Negative capacitance refers to a differential regime with $dQ/dV_{FE} < 0$ on the ferroelectric free-energy landscape, which can appear transiently during polarization switching or be stabilized by a series dielectric/capacitor in NC-FETs. Recent analyses indicate this effect can enable ultra-steep switching and energy-efficient inference, thereby realizing low-power neural circuits[84]. When connected in series with positive capacitance, this effect amplifies voltage. In 2008, Salahuddin and Datta first proposed leveraging ferroelectric negative capacitance to overcome thermodynamic limitations in transistors[85]. This technology breaks through traditional physical constraints on subthreshold swing. NC-FETs exhibit steep switching characteristics, making them suitable as highly sensitive threshold detectors. When the accumulated membrane potential reaches a certain threshold, the neuron fires an output pulse, thereby enhancing transmission speed and temporal precision. This unique switching behavior enables circuits to operate at lower voltages, significantly reducing dynamic power consumption.

Researchers have designed and validated multiple neuromorphic circuits based on NC-FETs, such as leaky-integrating-discharging neurons[86]. In these circuits, the NC-FET serves as the primary switching device. Its unique characteristics render the neuron highly sensitive to minute input variations while maintaining extremely low energy consumption. The multifunctional characteristics of HfO_2 -based ferroelectric devices not only make them suitable candidates for emulating analog synapses, but also provide innovative building blocks for composing compact and energy-efficient neuron circuits, and recently demonstrated nanoscale HfO_2 -based FeFET neuron units have successfully mimicked biological integrate-and-fire behavior, achieving a bridge between the device level and neuron models[87], and whether by using their inherent hysteresis characteristics to implement Schmitt triggers, or by using NC-FETs and negative-capacitance-assisted circuits to build steep-switching neurons, HfO_2 -based ferroelectric devices can, in future monolithic

neuromorphic chips, simultaneously assume the dual roles of "data storage units" and "data computing units" — thereby realizing in-memory computing in a true sense.

However, achieving practical neuromorphic computing cannot be accomplished with individual synapses and neurons alone. Therefore, we must organize them into large-scale arrays, integrated with peripheral circuits and system architectures, to truly realize neuromorphic computing. This is precisely what the next chapter will explore.

5 Hardware Systems and Emerging Applications

In the previous chapters, this paper has carried out a comprehensive discussion on the material physics and device-unit characteristics of HfO_2 -based ferroelectric devices, and isolated synaptic or neuronal devices by themselves cannot form a computing system, and to perform complex cognitive tasks, it is necessary to scientifically organize tens of thousands of such devices into large-scale hardware arrays, together with a system framework in which algorithms can run on them, and this chapter focuses on the transition from devices to systems: it first explains the mainstream in-memory-computing architecture—the crossbar array—and analyzes the physical principles behind its core operations; it then comprehensively examines the key challenges faced in system integration, and finally introduces the most cutting-edge applications and the corresponding solution strategies.

5.1 In-Memory Computing Architectures

In neural networks—particularly deep neural networks—the most fundamental and frequent computational task is vector–matrix multiplication (VMM), which essentially consists of a series of multiply-accumulate (MAC) operations. In traditional von Neumann architectures, executing VMM operations incurs significant energy and time costs. First, it requires reading weight matrices and input vectors from memory, then performing computations on the CPU/GPU, and finally writing results back to memory. The greatest breakthrough of in-memory computing lies in eliminating data movement bottlenecks. It primarily leverages physical laws to execute VMM operations directly within memory arrays using massively parallel processing.

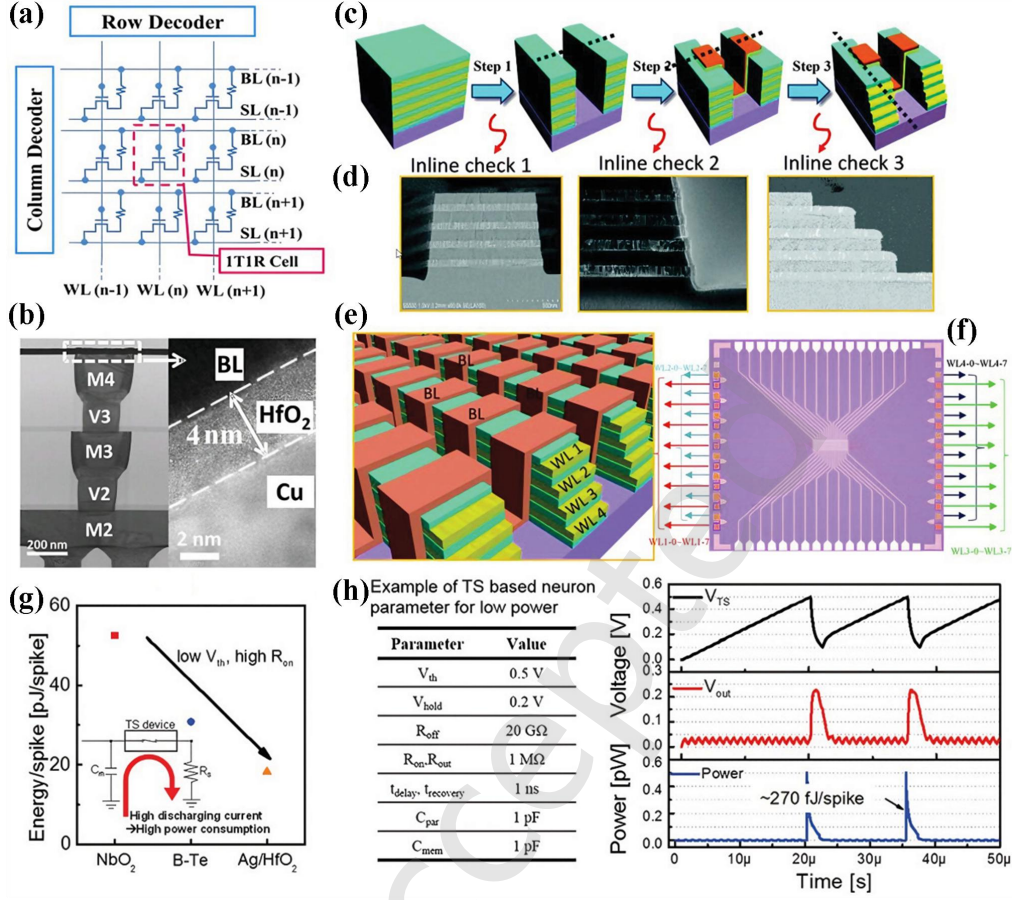


Figure 11 Array/cell integration examples related to ferroelectric memory and computing hardware[14]. (a) Circuit-level schematic of a 1T1R array with row/column decoding. (b) Cross-sectional microscopy/TEM view highlighting device stack and nanoscale HfO₂ thickness. (c) Step-by-step 3D integration/process illustration for array formation. (d) Inline inspection images corresponding to key fabrication steps. (e) 3D schematic of a dense crossbar/array architecture. (f) Optical/SEM image of a fabricated array and its routing/line assignment. (g) Energy-per-spike benchmark comparison for different threshold-switching device materials (illustrative neuron power trade-off). (h) Example neuron parameters and corresponding voltage/power waveforms for low-power operation.

Crossbar arrays represent one of the most efficient hardware architectures for in-memory computing. Figure 11(a)–(d) provide representative examples of array/cell integration and key fabrication/inspection views for ferroelectric memory technologies, while Figure 11(e)–(f) illustrate the dense crossbar wiring topology used for in-memory computing. As shown in Figure 11(e)–(f), a crossbar array consists of two sets of vertical wiring, and at each intersection point, corresponding non-volatile memory devices are uniformly deployed. Additionally, Figure 11(g) primarily illustrates the impact of device selection on energy efficiency at the neuron level, while Figure 11(h) presents the key parameters and corresponding waveforms for a low-power neuron example.

Although the cross-array structure appears relatively simple, it ingeniously leverages two fundamental physical laws to perform vector-matrix multiplication (VMM) operations:

$$\text{Ohm's Law: } V = i \times R \text{ or equivalently } i = G \times V$$

Kirchhoff's Current Law (KCL): $\Sigma I_{in} = \Sigma I_{out} \Rightarrow \sum I_{in} = I_{out}$

Operation process (VMM/MAC in a crossbar):

Consider an $M \times N$ crossbar array. The memory cell at row i and column j is programmed to a conductance G_{ij} that encodes the synaptic weight W_{ij} , where $i=1, \dots, M$ and $j=1, \dots, N$. The input vector is applied as row voltages $\{V_{ij}\}$.

Parallel multiplication (Ohm's law): when a voltage V_i is applied on row i , the current through the cell (i,j) is

$$I_{ij} = G_{ij}V_i, \quad i=1 \dots M, j=1 \dots N \quad (1)$$

Analog accumulation (KCL): the total column current collected at the end of column j is the sum of all row currents in that column,

$$I_j^{col} = \sum_{i=1}^M I_{ij} = \sum_{i=1}^M G_{ij}V_i, \quad j=1 \dots N \quad (2)$$

Therefore, the crossbar performs a multiply-accumulate (MAC) operation in the analog domain, and the vector-matrix multiplication (VMM) can be written as

$$\mathbf{I}^{col} = \mathbf{G}^T \mathbf{V} \quad (3)$$

(where $\mathbf{V} \in M \times 1$, $\mathbf{G} \in M \times N$, and $\mathbf{I}^{col} \in N \times 1$; the transpose form follows the above row/column convention).

Finally, the output of the array is a current vector $\mathbf{I}^{col}$. This output current vector $\mathbf{I}^{col}$ physically corresponds to the analog MAC result of the input voltage vector $\mathbf{V}$ and the conductance matrix $\mathbf{G}$ stored in the array. In other words, the entire vector-matrix multiplication operation can be completed within a single clock cycle. Subsequently, peripheral circuits convert these analog output currents into digital signals for processing by the next computational layer. Figure 12 illustrates a typical in-memory computing system architecture.

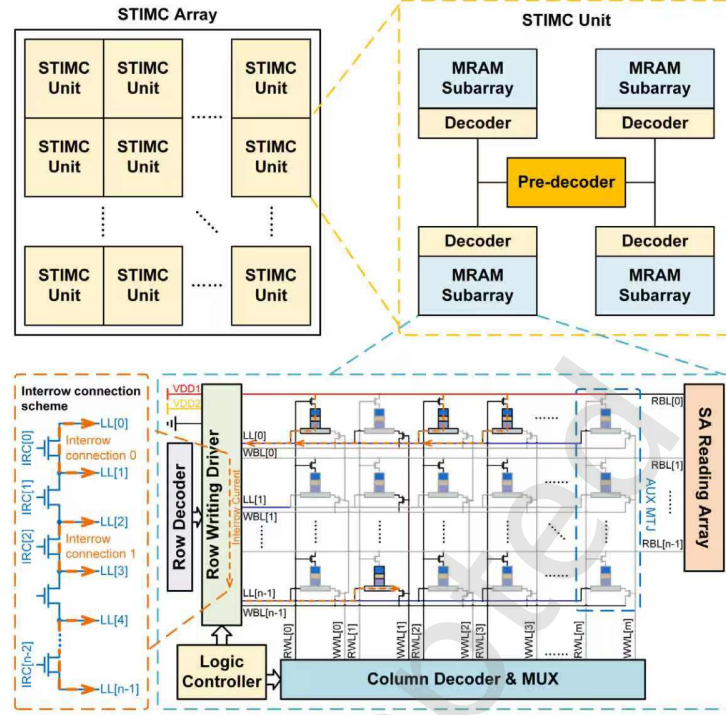


Figure 12 Schematic system architecture of in-memory computing based on crossbar arrays[88]

The core advantage of crossbar arrays in in-memory computing lies in their massive parallelism. For an $N \times N$ array, a single operation step can execute N^2 multiplications and $N(N - 1)$ additions, offering a computational throughput far beyond that of traditional digital architectures. To better leverage the speed and density advantages of analog computing, devices such as multi-conductivity-state FeFETs can be employed. Soliman et al. achieved the first memory-computing crossbar switch chip using multi-level-cell FeFETs fabricated with a 28-nm HKMG process, attaining an energy efficiency of approximately 885 TOPS/W at the chip level[89]. This breakthrough lays a solid architectural foundation for energy-efficient edge AI and large-scale neural network hardware acceleration[90].

5.2 Challenges and Countermeasures in System Integration

Although crossbar-based in-memory computing architectures offer considerable potential, researchers still face challenges in scaling individual devices to large-scale hardware systems. These include array non-idealities, peripheral circuit power consumption, and inherent device defects. Importantly, these limitations are strongly dependent on the selected device architecture, because two-terminal ferroelectric cells and three-terminal FeFET cells differ significantly in their selection mechanisms and readout characteristics. Therefore, this section first discusses architecture-dependent array non-idealities and then addresses general system-level bottlenecks associated with peripheral circuits and hardware-software co-design.

5.2.1 Array Non-Idealities: Sneak Paths, Line Resistance, and Parasitic Effects

It should be emphasized that array-level non-ideal effects are closely related to device architectures. In passive two-terminal crossbar arrays, sneak-path current is an unavoidable problem that may even disrupt circuit operation, that is, unselected cells in the array form unintended leakage paths. In contrast, three-terminal FeFET arrays possess gate-controlled selection capability, and therefore can intrinsically suppress sneak paths. However, each type of device also has its own shortcomings, and three-terminal FeFET arrays are no exception. In three-terminal FeFET arrays, parasitic capacitance, transistor leakage current, and peripheral circuit overhead are the issues that need to be improved.

When millions or even billions of devices are integrated into high-density arrays, ideal circuit models become inadequate. Various internal effects significantly impact computational accuracy and reliability. The bypass effect is the most critical factor in passive crossbar arrays. When reading or writing a single cell, current does not necessarily flow along the intended path but may also flow to adjacent low-resistance cells, forming multiple parallel paths. Figure 13 illustrates the sneak-path issue in passive crossbar arrays (Figure 13 (a)–(c)) and its impact on read-voltage margin scaling (Figure 13 (d)–(e)). Unintended current flow can severely compromise read accuracy and may even affect the state of neighboring cells. Figure 13(d) illustrates a typical half-biased ($V/2$) read scheme. This scheme primarily ensures that the selected cell can withstand the full read voltage, while unselected or partially selected cells remain at a lower bias level. For example, when the selected word line and bit line are biased at $+V_{\text{read}}/2$ and $-V_{\text{read}}/2$ respectively, the selected junction experiences the full voltage V_{read} , while adjacent cells experience at most $V_{\text{read}}/2$.

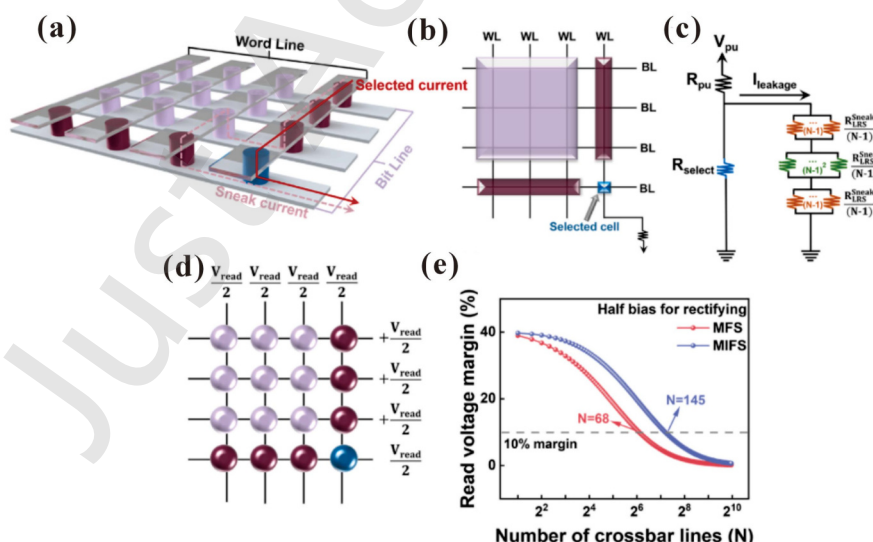


Figure 13 Sneak-path issue in crossbar arrays and read-margin scaling[12]. (a) 3D illustration of selected current versus sneak current paths. (b) Crossbar selection schematic indicating the addressed cell. (c) Equivalent leakage-path network illustrating accumulated sneak currents. (d) Half-bias ($V/2$) read scheme example across the array. (e) Read-voltage margin versus array size, comparing different stack configurations.

The most direct approach involves connecting a selector device in series with each storage cell to form a 1S1R unit. This selector is a dual-terminal device exhibiting highly

nonlinear I-V characteristics, maintaining a high-resistance state at low voltages and only conducting when the voltage exceeds a threshold. By designing different voltage bias schemes, only the target cell can be activated while keeping the selectors of other cells closed, thereby suppressing leakage current paths[66]. As shown in Figure 14(a)–(c), the effective sensing margin gradually decreases as the array size increases. Moreover, in large-scale arrays, the combined effects of interconnect IR voltage drop and transistor leakage current become more pronounced. Such selector-based mitigation is important for passive FTJ-type arrays, whereas self-rectifying Fe-diodes can suppress leakage through their intrinsic nonlinear transport characteristics. For three-terminal devices such as FeFETs, the transistor inherently functions as a selector. The word line precisely controls the on/off state of each row memory cell, significantly reducing leakage current issues. Compared to two-terminal memristors, this key feature highlights FeFETs advantage in constructing large-scale arrays[91]. In large-scale arrays, the resistance of thin metal wires can no longer be ignored. When current flows through these metal interconnects, an IR voltage drop occurs. Consequently, the actual voltage applied to a cell and the current read from it vary depending on its position along the metal wire. It is evident that cells closer to the voltage source receive higher voltages, while those farther away receive lower voltages. This voltage non-uniformity introduces systematic errors into Vector Matrix Multiplication (VMM), significantly degrading computational accuracy[66]. Regarding countermeasures, architecturally, large arrays can be partitioned into multiple smaller subarrays to reduce wiring lengths. Material-wise, metals with superior conductivity can be employed as interconnect materials. Algorithmically, hardware non-idealities can be incorporated into the neural network training process. This approach enables the algorithm to learn and compensate for system errors, with experiments demonstrating improved inference accuracy in deployed hardware[92].

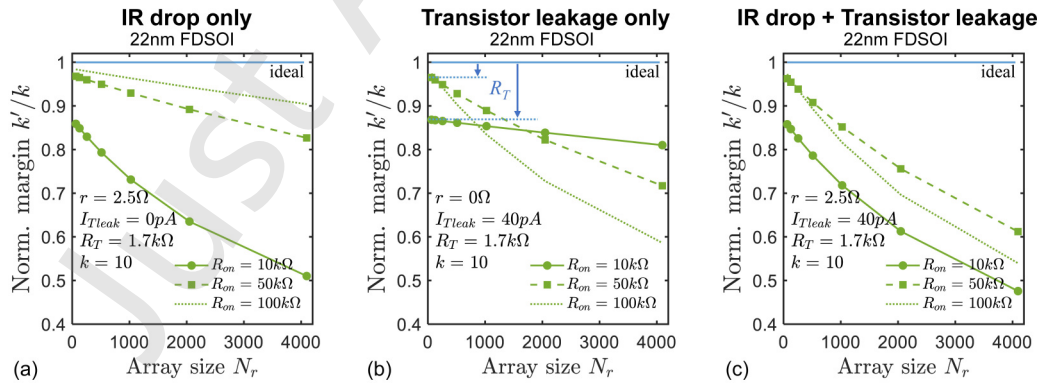


Figure 14 Representative impact of array size on normalized sensing margin under IR drop and transistor leakage.[10] (a) IR drop only (b) Transistor leakage only (c) IR drop + Transistor leakage

5.2.2 Peripheral Circuit Overhead and Hardware–Software Co-Design

The memory array itself serves as the core of computation. Its complex peripheral circuits determine its efficient operation, making the power consumption and area overhead of these circuits the true system bottlenecks. The core of the memory architecture is analog computation, yet input/output typically remains digital. This necessitates corresponding analog-to-digital/digital-to-analog converters. For instance, high-precision digital-to-analog

converters (DACs) must be integrated at the input to convert digital vectors into multi-level analog voltage pulses, while high-precision analog-to-digital converters (ADCs) are required at the output to reconvert analog currents into digital signals. These converters not only occupy substantial chip area but may even dominate the system's total power consumption[93]. Recent research indicates that circuit-level ADC/DAC co-design can minimize the computational overhead of analog memory operations[94]. A key route is to relax the precision requirements on ADCs/DACs, and one way is to develop neural networks that are robust to quantization, endowing the network with the ability to tolerate computational noise so that, even when low-precision converters are used, it can still maintain an acceptable performance level, and another potential route is to explore alternative encoding schemes, such as time-domain encoding—relying on the duration or frequency of pulses to encode information rather than on more energy-hungry analog voltages.

Beyond material-level challenges, the devices themselves exhibit inherent characteristics such as randomness, nonlinearity, and asymmetry. These non-ideal conditions reduce the accuracy of neural network inference. This situation has driven researchers to adopt hardware–software co-design approaches, and the Yu group has developed a training framework that incorporates hardware non-idealities, greatly improving the inference accuracy of real RRAM arrays, and achieving idealized behavior solely through device fabrication technologies is extremely difficult and also very costly, and a more practical approach is to accept device imperfections at the hardware level while performing adaptive correction via software and algorithms. This includes: designing fault-tolerant network models with noise resistance capabilities and employing network learning mechanisms to efficiently execute tasks on non-ideal Si-based FeFET hardware. Additionally, advanced materials engineering methods are required to reduce coupling between polarization and charge trapping[95]. Periodic online calibration combined with weight refresh can effectively mitigate weight drift caused by fatigue effects and imprint effects[96]. By comprehensively considering materials, components, circuits, architectures, and algorithmic implementations to achieve high-performance memory-centric computing systems, and then through multi-level collaborative optimization, we can tangibly enhance computational efficiency. Table2 summarizes the major non-idealities discussed in Section5.2 and their representative quantitative impacts.

Table2. Semi-quantitative summary of major array-level non-idealities

Non-ideality	Main impact	Representative quantitative evidence	Mitigation	Refs.
Sneak path	Read-margin loss; crosstalk	~20% lower sneak current; 7.5–22× lower normalized leakage with bilayer/self-rectifying design	1T1R / 1S1R; selector; V/2 or V/3 bias	[97]
IR drop	Position-dependent VMM error; accuracy loss	DNN accuracy drop reduced from 5.6% to 2.9% / 2.1% by remapping	Sub-arrays; remapping; IR-aware	[98]

ADC/DAC overhead	Quantization noise; energy/efficiency bottleneck	Optimal ADC $\approx$ 6 b (MRAM), 7 b (ReRAM/FeFET); ADC-less ternary scheme: energy -28% vs 7-bit and -12% vs 4-bit baseline	training Mixed precision; QAT; ADC-less partial sums; peripheral co-design	[99]
------------------	--	---	--	------

In summary, hafnium dioxide-based memory systems require not only attention to device performance but also suppression of array non-idealities. Building upon this foundation, recent research has further extended ferroelectric memory-computing architectures to emerging application scenarios.

5.3 Emerging Applications: From Edge AI to In-Sensor Computing

Regarding system integration challenges, memory computing technology continues to advance, giving rise to numerous cutting-edge applications. These applications demand exceptional performance in computational energy efficiency, latency, and physical footprint. Ferroelectric memory architectures based on hafnium dioxide demonstrate outstanding capabilities across these dimensions.

Most current AI applications rely on uploading data to cloud data centers, where it undergoes cluster processing by GPUs before results are returned to edge devices. This model suffers from three major flaws: First, it introduces high latency, with data round-trips taking too long to meet the real-time demands of applications like autonomous driving or industrial robotics. Second, it poses significant privacy risks, as sensitive user data must be processed outside the device. Finally, it relies heavily on high bandwidth, rendering the system inoperable in unstable or disconnected network environments. Edge AI involves moving AI model inference from the cloud to edge devices. This imposes stringent hardware requirements, demanding sufficient computational power within milliwatt-level or even microwatt-level power budgets. FeFET-based memory-computing chips, featuring non-volatility and highly efficient analog MAC operations, serve as the hardware foundation for edge AI. Their non-volatility ensures weight data persists after power loss, significantly reducing standby power consumption and enabling instant wake-up. Efficient analog MAC operations execute core neural network computations at power levels far below digital processors. A typical application is wearable health monitoring devices. Utilizing FeFET-based memory cores, these devices continuously analyze real-time sensor data streams—such as heart rate and blood oxygen levels—at ultra-low power. Upon detecting abnormal patterns like arrhythmia, it immediately alerts the patient without requiring raw data upload to the cloud. This approach safeguards user privacy while extending battery life. Additionally, novel low-power image processing hardware based on FeFET arrays has validated its energy-saving effectiveness for real-time edge detection tasks in edge AI vision systems[100]. Current memory computing research primarily focuses on AI inference—executing pre-trained fixed-weight models. However, truly intelligent systems

should possess online learning capabilities, enabling real-time model updates and optimization based on new data. The above-mentioned networks are all deterministic neural networks. However, when the input data of the system are incomplete or disturbed by noise, deterministic neural networks can hardly function effectively. In this case, Bayesian neural networks can quantify prediction uncertainty and make more reliable decisions. In recent years, Huang et al. experimentally realized a Bayesian neural network through a three-dimensional array composed of 16-layer Fe-diodes. In this system, the entropy source and synaptic weights were uniformly integrated in the same ferroelectric diode hardware platform[18]. Therefore, in terms of edge intelligence, the role of Bayesian neural networks also cannot be ignored. This demands hardware that not only efficiently supports inference but also accommodates training requirements. In the stage of implementing in-memory training, achieving efficient forward propagation is far less critical than achieving efficient backward propagation, and the latter is equally crucial for gradient accumulation required for weight updates as well as for the computation of transpose vector–matrix multiplications, and compared with inference, the difficulties encountered in in-memory training are much greater, and device characteristics (such as the linearity and symmetry of FeFETs) play a decisive role in the accuracy of gradient accumulation and weight updates.

Even in the face of these obstacles, in-memory training based on ferroelectric devices still shows enormous potential[6] , and one core application is personalized intelligent devices, and if a voice assistant adopts an in-memory computing chip, it can securely learn a user's accent and speaking habits locally, continuously improving the speech-recognition model without sending sensitive data to external servers, and this capability is extremely crucial for achieving truly personalized and adaptive edge artificial intelligence.

6 Challenges and Opportunities of Ferroelectric Technology

After more than a decade of rapid development, HfO_2 -based ferroelectric devices and their applications in neuromorphic computing have evolved from an initial scientific discovery into a vibrant and promising frontier research field. This paper has systematically reviewed this progress—from the unique ferroelectric origins of HfO_2 materials, to diverse device implementations such as FeFETs and FTJs, and further to system architectures for in-memory computing and their emerging applications. Research indicates that hafnium dioxide-based ferroelectric devices exhibit high compatibility with CMOS processes, excellent scalability, and energy-efficient analog computing capabilities. Consequently, they have emerged as one of the optimal choices for constructing brain-inspired hardware in the post-Moore's Law era. However, significant challenges remain for large-scale commercial deployment. This chapter first summarizes the core difficulties currently facing the field, then outlines technical pathways and future development directions. It aims to provide guidance for subsequent research.

6.1 Key Challenges: Material Stability, Device Reliability, and System Integration

Although HfO_2 -based ferroelectric technology holds promising prospects, challenges

remain to be overcome at the material, device, and system levels. At the material level, there exists a contradiction between intrinsic defects and the process window. As discussed in Chapter 2, the ferroelectric properties of HfO_2 originate from its metastable crystalline phase. However, the stability of this phase depends on the synergistic interaction between factors such as doping, stress, and size, potentially resulting in a narrow process window. Additionally, the thin film exhibits extreme sensitivity to parameters including composition, thickness uniformity, annealing temperature, and annealing time. Minor deviations can trigger significant fluctuations in device performance.

The more critical issue lies in point defects such as oxygen vacancies, which can stabilize the ferroelectric phase while simultaneously serving as sources of reliability problems (e.g., wake-up effects, fatigue effects, and imprint effects). Thus, the fundamental challenge in materials science is: How can we ensure the stability of the ferroelectric phase while suppressing the formation and migration of harmful defects. This requires a comprehensive understanding of defect dynamics at the atomic scale and the development of more advanced thin-film growth and interface engineering techniques.

Dual Challenges of Reliability and Non-Idealism. The performance and reliability of individual devices form the foundation for the entire system. For ferroelectric field-effect transistors (FeFETs) serving as artificial synapses, addressing material-level durability and retention issues is crucial, while FeFETs must also contend with the inherent non-idealities of analog devices. The primary obstacles preventing FeFETs from achieving high-precision memory training are the nonlinearity and asymmetry of weight updates. While advanced pulse programming techniques can partially compensate for these shortcomings, they increase peripheral circuit complexity and power consumption, potentially undermining the inherent advantages of the memory-computing architecture itself. In large-scale arrays, minor process variations cause discrepancies in critical FeFET parameters (threshold voltage, storage window, switching speed). Worse still, frequent operations induce random fluctuations within individual devices. This randomness manifests as computational noise, severely degrading neural network inference accuracy. Therefore, optimizing manufacturing processes to enhance device uniformity is vital for achieving high-precision computation. As mentioned in Chapter 4, even when device arrays approach ideal conditions, system integration remains a significant challenge, with peripheral circuit overhead becoming particularly prominent. High-precision analog-to-digital converters (ADCs)/digital-to-analog converters (DACs), complex pulse generators, and voltage controllers can occupy more area and consume more power than the memory computing array itself, becoming the system bottleneck. Furthermore, in large arrays, physical phenomena such as wire resistance and parasitic capacitance can introduce systematic computational errors.

Optimizing devices or circuits alone is insufficient; a holistic co-design paradigm spanning algorithms to hardware must be established. A significant gap currently exists between algorithm researchers and device physicists. Neural network designers typically assume ideal hardware conditions, while device engineers may not fully grasp the specific hardware requirements dictated by the algorithms. Eliminating this divide—developing intelligent algorithms that can recognize hardware characteristics and adaptively correct

non-idealities, while simultaneously designing hardware structures that are appropriate rather than over-engineered—is the key to transitioning this technology from vision to practical implementation[101].

6.2 Future Technological Pathways

To address these challenges, researchers are actively developing next-generation multiscale ferroelectric technologies aimed at fundamentally overcoming existing technical bottlenecks and pioneering new application domains. These cutting-edge directions not only chart a roadmap for HfO_2 -based materials but also sketch a grand vision where new materials, computational models, and application scenarios intertwine.

Two-dimensional (2D) materials—represented by graphene, transition metal dichalcogenides (TMDs), and intrinsic 2D ferroelectrics—offer a breakthrough opportunity to resolve various challenges currently facing ferroelectric devices. Recent studies reveal that in HfO_2 -based FeFETs, chemically vapor-deposited monolayer MoS_2 channels enable atomic-level low-voltage control in neuromorphic circuits[102]. The combination of black phosphorus with the ferroelectric polymer P(VDF-TrFE) exhibits high-mobility, high-efficiency artificial synapse behavior, achieving over 93% recognition accuracy in brain-inspired networks[103]. The convergence of these two fields is primarily explored through three avenues. Two-dimensional semiconductor channels, utilizing two-dimensional transition metal dichalcogenides (2D TMDs) as FeFET channel materials instead of silicon or indium gallium arsenide, offer numerous advantages. The atomically flat surfaces of these two-dimensional materials lack dangling bonds, forming near-ideal interfaces with HfO_2 ferroelectric layers. This effectively eliminates mobility degradation, threshold voltage fluctuations, and reliability issues caused by interfacial trap charges[104]. Their ultra-thin morphology provides excellent resistance to short-channel effects, making them ideal channel material candidates for process nodes below 3 nm. The development of intrinsic two-dimensional ferroelectrics has emerged as another research direction. $\alpha\text{-In}_2\text{Se}_3$ exhibits both in-plane and out-of-plane ferroelectricity at room temperature while inherently behaving as a semiconductor. Ding et al. theoretically predicted this property in 2017, which was subsequently experimentally confirmed[105]. This breakthrough opened the door to creating "ferroelectric semiconductors," materials capable of performing both storage and logic functions within a single medium, thereby simplifying device architecture[106]. Van der Waals heterostructures demonstrated the feasibility of this concept. By integrating ferroelectric polymers with black phosphorus, neuromorphic transistors with high recognition accuracy and ultra-low power consumption were developed[103]. Using n-type ferroelectric semiconductor $\alpha\text{-In}_2\text{Se}_3$ as the gate and p-type MoTe_2 as the channel, a junction field-effect transistor with inherent non-volatile memory characteristics was fabricated. This demonstrates a feasible approach for integrating logic and memory functions within a fully two-dimensional material system. The layered structure of two-dimensional ferroelectric materials enables the construction of novel heterojunction devices via van der Waals stacking. Research on the electric field control of quantum effects such as exciton states and Moire superlattices has opened new avenues for scientific exploration[107].

The application of 2D materials as functional interface layers significantly enhances performance even when core materials are not fully replaced. Incorporating 2D layers into traditional MFIS stack structures markedly improves functionality. Monolayer hexagonal boron nitride (h-BN) serves as an excellent diffusion and tunneling barrier layer, effectively suppressing oxygen vacancy migration and charge injection. This substantially enhances device endurance and data retention capabilities[106]. Tan et al. achieved write endurance exceeding 10^{10} cycles in FeFETs by inserting a high-dielectric-constant interface layer between the gate dielectric and ferroelectric stack[108].

Current research predominantly focuses on emulating deep neural networks and spiking neural networks. However, the unique physical properties of ferroelectric devices make them an exceptional platform for exploring post-deep learning computing paradigms. Leveraging the inherent switching randomness of ferroelectric devices enables efficient hardware-based random number generators. Random computing encodes numerical values as long bit streams, simplifying complex arithmetic operations into basic logic operations. This approach offers high fault tolerance and low hardware overhead. The analog and multilevel characteristics of ferroelectric field-effect transistors (FeFETs) make them highly suitable for executing random computations using bit streams[109].

Bio-inspired models for hyperdimensional computing encode information as high-dimensional sparse binary vectors. Its core operations—vector binding, bundling, and similarity search—can be efficiently mapped onto nonvolatile, massively parallel crossbar arrays. FeFET arrays leverage nonvolatility and parallelism, offering a highly promising path for low-power, robust hyperdimensional computing hardware[110].

Oscillator Networks and Ising Machines: Compactly coupled nano-oscillator networks can be constructed using the negative resistive or capacitive effects of ferroelectric devices. Csaba and Porod pioneered brain-inspired computing using nano-oscillator networks[111]. Such networks leverage synchronization and phase dynamics to effectively tackle complex combinatorial optimization problems. Sutton et al. demonstrated that this system can be mapped to the Ising model, thereby solving NP-hard problems like the maximum cut problem[112]. Ising machines based on memristor hardware have already demonstrated the ability to accelerate NP-hard problem solving, with applications in drug discovery, financial modeling, and logistics planning[113].

Toward the Unification of Perception, Storage, and Computation. Future intelligent systems will not consist of isolated modules for perception, storage, and computation. Instead, they will evolve toward the deep integration of these three functions—a concept termed perception-storage-computing convergence. Devices fabricated on flexible substrates will play a pivotal role in this evolutionary process. Ferroelectric HfO_2 inherently exhibits piezoelectric and thermoelectric effects, enabling it to not only store data but also directly sense changes in pressure, vibration, or temperature. Khan and collaborators systematically demonstrated how multi-physics coupling in ferroelectric devices achieves the tight integration of sensing, storage, and computation[86].

A highly promising concept involves fabricating "smart skin" based on FeFETs on flexible substrates [113]. Beyond leveraging the inherent piezoelectric or thermoelectric

effects of the material for sensing, another innovative approach utilizes the gate structure of FeFETs for chemical sensing. By employing a catalytic metal sensitive to a specific gas as the gate electrode, gas molecules interacting with this metal alter its work function. This change modifies the threshold voltage of the ferroelectric field-effect transistor, enabling highly sensitive detection of gas type and concentration[114]. This embodies a design philosophy that deeply integrates sensing functionality with the storage and switching capabilities of transistors at the device level. Within this system, each FeFET gate simultaneously functions as both a pressure sensor and a storage unit. When an external object contacts the skin, the resulting pressure modulates the ferroelectric polarization state of the transistor, enabling instantaneous writing and preprocessing of perceived events. The entire FeFET array can leverage its memory-computing capabilities to recognize and analyze tactile patterns in real time. This integration of sensing, storage, and computation at the device level eliminates the bottlenecks of conversion and transmission inherent in traditional systems, promising revolutionary applications in wearable devices, flexible robotics, and human-machine interaction. Against this backdrop, flexible ferroelectric polymers have also been employed for tactile sensing[115]. Figure 15 summarizes sensing–memory–computing co-integration concepts, spanning from the algorithmic DNN/CNN pipeline mapped to hardware acceleration needs (Figure 15(a)) to an application-level neuromorphic sensory system example (Figure 15(d)), while highlighting synaptic-device-based sensory perception (Figure 15(b)) and in-sensor convolution for feature extraction (Figure 15(c)).

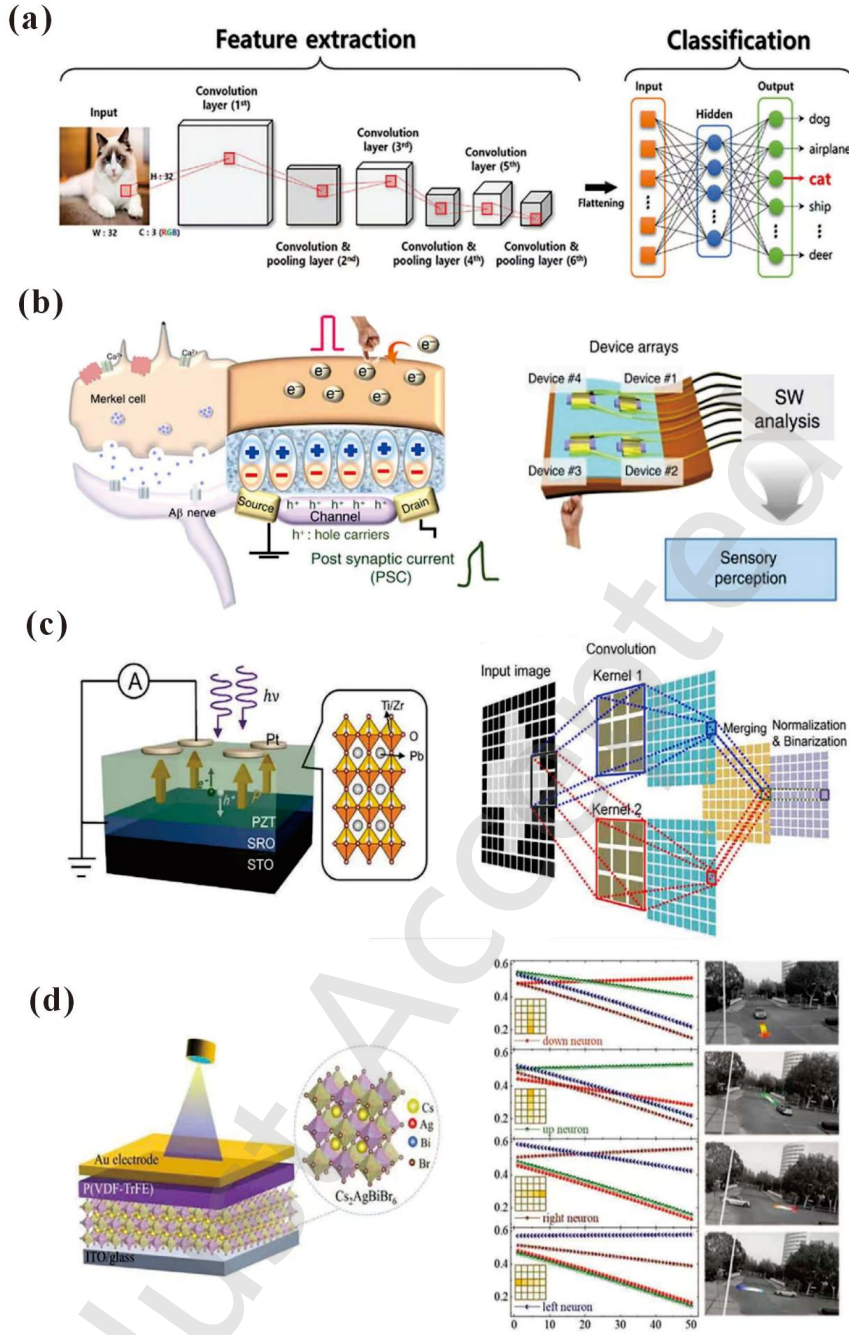


Figure 15 Representative concepts of sensing–memory–computing co-integration[116]. (a) Example DNN/CNN pipeline (feature extraction + classification) mapped to hardware acceleration needs. (b) Sensory-perception concept using synaptic devices/arrays to process stimuli into signals. (c) In-sensor convolution concept: device-level sensing front-end combined with convolution kernels/feature extraction. (d) Neuromorphic sensory system example illustrating application-level outputs enabled by integrated devices.

Two-dimensional (2D) material deep integration can overcome current technological bottlenecks and drive process nodes toward more advanced directions. This paradigm not only enables novel computing architectures but also provides innovative hardware solutions for tackling complex problems beyond artificial intelligence. CAM content-addressable memory based on ferroelectric devices serves as an efficient alternative for search and

associative storage in AI applications[4]. Such multifunctional integration technologies will ultimately drive the emergence of next-generation intelligent systems—systems capable of genuinely mimicking biological intelligence and freely interacting with the physical world. Achieving such closed-loop, adaptive intelligent systems is one of the ultimate goals of neuromorphic engineering[117].

6.3 Circuit-Level Innovations

Although breakthroughs in materials science and device structure optimization have laid a solid physical foundation for ferroelectric brain-inspired computing, innovation at the circuit level remains the critical link for translating the exceptional performance of individual devices into high energy efficiency and reliability in macroscopic computing systems[118]. This challenge parallels the developmental stages of traditional flash memory technology, whose success relies not only on the floating-gate transistor itself but more critically on a mature circuit ecosystem encompassing read/write control, error correction, and wear leveling[119]. For the core of analog computing—FeFETs—circuit design faces uniquely formidable challenges, demanding profound innovation across four dimensions: sensing, programming, reliability, and system integration[116]. Representative circuit and array architectures enabled by ferroelectric devices are summarized in Figure 16.

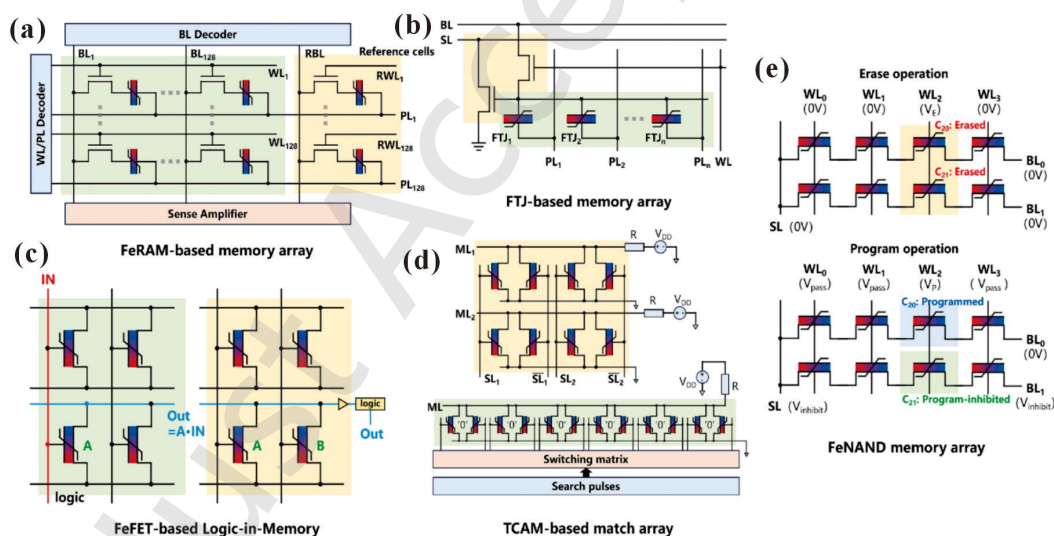


Figure 16 Circuit/array architectures built from ferroelectric devices[80]. (a) FeRAM-based memory array with decoders, sense amplifier, and reference cells. (b) FTJ-based memory array schematic. (c) FeFET-based logic-in-memory example implementing logic/dataflow in-array. (d) TCAM-style match array architecture with search pulses and switching matrix. (e) FeNAND array operation schemes showing erase and program biasing (including inhibited/selected states).

6.3.1 High-Fidelity, Low-Overhead Analog Sensing and Readout Circuits

In memory computing architectures, the channel conductance or threshold voltage of ferroelectric field-effect transistors (FeFETs) directly encodes the analog weights of neural networks. Therefore, rapid, accurate, and low-power sensing of these subtle multi-level analog states is a prerequisite for ensuring computational precision[120].

The core challenge lies in the fact that traditional digital memory sense amplifiers can

only distinguish between high and low logic states, failing to meet the requirements for reading multi-bit analog weights. While integrating high-precision analog-to-digital converters (8-bit SAR-ADCs) per column is feasible, it incurs significant area and power consumption costs[121], potentially exceeding those of the computational array itself, thereby undermining the fundamental purpose of memory-based computing. This is an ingenious method for converting analog signals to the digital domain. Its core idea is to use the integrated current on the bitline to charge/discharge a known capacitor. According to the formula $\Delta t = C \cdot \Delta V / I_{\text{column}}$, the larger the current, the shorter the time Δt required to reach a preset voltage threshold ΔV [122]. Measuring the time interval Δt with a simple digital counter enables high-precision quantization of current values. This current-to-time-to-digital conversion scheme replaces complex analog comparators with low-power digital logic, making it particularly suitable for large-scale parallel readout. The primary trade-off is a relatively long conversion time in low-conductance states, introducing data-dependent delays.

This approach applies a linear ramp voltage to the gate of a ferroelectric field-effect transistor (FeFET). Devices with higher conductivity turn on earlier, and their turn-on times are recorded by a latch. Analyzing the turn-on time sequences of different devices yields weight levels[123]. This technique maps conductance values to a time-series dataset, similarly avoiding direct measurement of weak analog currents. Inspired by the common 3T pixel cell structure in CMOS image sensors, a source follower and selection transistor are added adjacent to each FeFET cell. This design converts the local conductance signal into a more reliable output voltage signal, effectively eliminating bitline parasitic capacitance and noise interference, significantly improving signal-to-noise resolution[124]. The trade-off of this approach is reduced storage density, reflecting the inherent compromise between integration density and readout accuracy.

6.3.2 Intelligent Programming Control Circuits for Achieving Linearity and Symmetry

The nonlinearity and asymmetry of FeFET weight updates pose the greatest obstacle to achieving high-precision on-chip training [125]. Device improvements alone cannot fundamentally resolve this issue; closed-loop or open-loop compensation at the circuit level must be introduced.

The core challenge lies in the need for accurate, linear accumulation of weights when employing gradient descent algorithms in neural networks. However, the physical characteristics of ferroelectric field-effect transistors (FeFETs) cause their step size to decrease as they approach saturation, with distinct amplification and suppression response curves. A high-precision closed-loop control strategy. The controller first sets a target conductance value, then immediately performs a rapid, low-precision "verification" read after applying a short programming pulse to compare the current value with the target[126]. If the target is not met, a correction pulse is applied based on the deviation, iterating cyclically until the error falls below a threshold. This approach trades time for precision, achieving exceptionally high programming accuracy and currently stands as the most reliable method for deploying high-precision weights. This represents a more efficient open-loop control method. It requires detailed prior characterization of device properties to establish a lookup table (LUT)[95]. The LUT stores the optimal pulse parameters required to transition from any

initial state G_{start} to the target state G_{end} . During programming, the controller queries the LUT to generate "customized" pulses. It compensates for nonlinearity by automatically increasing pulse amplitude as weights approach saturation; it compensates for asymmetry by employing entirely different pulse waveforms for enhancement and suppression operations.

6.3.3 Active Reliability and Endurance Management Circuits

Reliability issues at the material level—such as wake-up, fatigue, and imprinting—ultimately require proactive management and compensation at the circuit system level to ensure stable operation of computing systems throughout their entire lifecycle[119]. This approach draws inspiration from the mature bad-block management and wear-leveling concepts employed in NAND flash memory.

The imprint effect causes weight drift over time, compromising long-term inference accuracy; the fatigue effect limits the total number of weight updates, which is a bottleneck for on-chip training. Periodic calibration and refresh: Design a background management unit to read partial or full weight values during system idle periods. By comparing these values against reference values stored in digital logic, drifting units are identified and corrected with minute adjustment pulses to restore them to accurate values. This mechanism, akin to DRAM refresh, ensures long-term reliability[127]. During training, certain weights update far more frequently than others. Wear leveling algorithms leverage dynamic logical-to-physical address mapping to sequentially write frequently updated logical weights to different physical array locations. This evenly distributes write pressure across the entire chip, preventing premature failure of localized units due to excessive fatigue. Consequently, it significantly extends the chip's training endurance[128]. While traditional error-correcting codes are designed for digital systems, lightweight error-correction schemes applicable to analog computation can be explored. Treating a set of weights as a data block, an analog checksum is computed and stored. When a cell experiences a significant error, the checksum detects the anomaly. An algorithmic redundancy mechanism then restores or masks the affected cell[129].

6.3.4 Peripheral Circuit Co-optimization in In-Memory Computing Architectures

The overall system-level energy efficiency of a complete memory-computing chip is not solely determined by its core computing array but is also constrained by its peripheral circuitry [130]. Therefore, a cross-layer co-design approach must be adopted.

The area and power consumption of peripheral circuits can account for over 50% of the total, becoming the true system bottleneck. Without optimization, the theoretical advantages of in-memory computing will be consumed by the overhead of these peripheral circuits. Introducing hardware noise models during neural network training—specifically "quantization/noise-aware training"[128]—enhances the robustness of trained models against hardware non-idealities. This enables circuit designers to significantly reduce precision requirements for ADCs/DACs, achieving exponential reductions in area and power consumption. Architecturally partitioning large computational arrays into independent subarrays. Modular design shortens wordlines and bitlines, mitigating voltage drop penalties. Consequently, the drive and readout circuits surrounding each subarray become smaller, faster, and more energy-efficient[123], ultimately achieving dual optimization of

system-level performance and efficiency.

Implementing these systemic innovations at the circuit level fundamentally bridges the gap between ideal device physics and practical system applications. This approach transforms the unique advantages of HfO₂-based ferroelectric devices into powerful computational capabilities for next-generation AI hardware.

6.4 Summary

Hafnium-based ferroelectric devices have emerged as a key technology for overcoming traditional computing bottlenecks in the post-Moore's Law era. They offer perfect compatibility with standard CMOS processes, excellent scalability, and ultra-low power consumption, among other advantages. Furthermore, FeFETs and FTJs serve as ideal hardware platforms for neuromorphic computing. They efficiently emulate the weight plasticity of biological synapses and the nonlinear firing behavior of neurons, while eliminating energy barriers associated with data movement through compute-in-memory architectures. Currently, this technology is transitioning from laboratory exploration to system-level applications. It demonstrates energy efficiency ratios far surpassing traditional digital architectures in edge AI, low-power IoT, and real-time image processing, laying the physical foundation for brain-inspired intelligent systems.

Looking ahead, ferroelectric technology still faces challenges in large-scale array integration, including material stability, device consistency, and peripheral circuit overhead. However, with advancements in science and technology, the introduction of two-dimensional material heterointegration and novel computational paradigms, and deepening hardware-software co-design, these challenges are gradually being overcome. This progress accelerates the deep integration of ferroelectric neuromorphic computing across sensing, storage, and computation functions. It is believed that ferroelectric technology will inevitably drive the next generation of intelligent hardware toward higher density and lower power consumption. It will also serve as a core driving force in realizing general artificial intelligence and a sustainable digital future.

Acknowledgements

This work was supported by the NSFC (92573202 and 62574061), Science and Technology Commission of Shanghai Municipality (25LN3201100), and the “Shuguang Program” supported by the Shanghai Education Development Foundation, and the Shanghai Municipal Education Commission. Fudan University Nano Information Science Innovation Support Center Project, Ministry of Education of China.

Conflict of Interest

The authors declare no conflict of interest.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- [1] McKee, S. Reflections on the memory wall. **2004**, 10.1145/977091.977115, 162.
- [2] Kaxiras, S. Architecture at the End of Moore. **2013**, 10.1007/978-3-642-33137-4_1, 1-10.
- [3] Wulf, W. A.; McKee, S. A. Hitting the memory wall: implications of the obvious. *SIGARCH Comput. Archit. News* **1995**, 23, 20–24.
- [4] Tarkov, M.;Tikhonenko, F.;Popov, V.;Antonov, V.;Miakonkikh, A.; Rudenko, K. Ferroelectric Devices for Content-Addressable Memory. *Nanomaterials* **2022**, 12.
- [5] Yousefzadeh, A.;Sohail, S.; Varbanescu, A. L. Memory Wall is not gone: A Critical Outlook on Memory Architecture in Digital Neuromorphic Computing. In *2025 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, 2025; pp 1-4.
- [6] Ielmini, D.; Wong, H. S. P. In-memory computing with resistive switching devices. *Nat. Electron.* **2018**, 1, 333-343.
- [7] Böске, T. S.;Müller, J.;Bräuhäus, D.;Schröder, U.; Böttger, U. Ferroelectricity in hafnium oxide thin films. *Appl. Phys. Lett.* **2011**, 99.
- [8] Xiang, H.;Chien, Y.-C.;Li, L.;Zheng, H.;Li, S.;Duong, N. T.;Shi, Y.; Ang, K. Enhancing Memory Window Efficiency of Ferroelectric Transistor for Neuromorphic Computing via Two - Dimensional Materials Integration. *Adv. Funct. Mater.* **2023**, 33.
- [9] Tibeica, C.;Sandu, T.;Nedelcu, O.;Plugaru, R.; Plugaru, N. Parametric simulations of composite barrier FTJs under external bias at room temperature. *Results Phys.* **2023**, 47, 106350.
- [10] Chen, J.; Indiveri, G. Scaling Effects of Transistor Leakage Current and IR Drop on 1T1R Memory Arrays.
- [11] Negi, S.;Saxena, U.;Sharma, D.; Roy, K. HClM: ADC-Less Hybrid Analog-Digital Compute in Memory Accelerator for Deep Learning Workloads. In *Proceedings Proceedings of the 30th Asia and South Pacific Design Automation Conference*, Tokyo, Japan, pp 648–655.
- [12] Lee, Y.; Kim, S. Enhanced synaptic performance in hafnia-based ferroelectric memristors with MIFS structure for neuromorphic computing. *Ceram. Int.* **2025**, 51, 44919-44929.
- [13] Zhou, Y.;Zhang, Y. K.;Yang, Q.;Jiang, J.;Fan, P.;Liao, M.; Zhou, Y. C. The effects of oxygen vacancies on ferroelectric phase transition of HfO₂-based thin film from first-principle. *Comput. Mater. Sci* **2019**, 167, 143-150.
- [14] Banerjee, W.;Kashir, A.; Kamba, S. Hafnium Oxide (HfO₂) – A Multifunctional Oxide: A Review on the Prospect and Challenges of Hafnium Oxide in Resistive Switching and Ferroelectric Memories. *Small* **2022**, 18.
- [15] Chen, A. A review of emerging non-volatile memory (NVM) technologies and applications. *Solid-State Electron.* **2016**, 125, 25-38.
- [16] Wu, S.;Chen, B.;Wang, C.;Mei, J.;Bai, M.;Zhan, X.;Wu, J.;Chai, J.;Xu, H.;Wang, X.; Chen, J. Dual-pulse disturb-free programming scheme for FeFET based neuromorphic computing. *Microelectron. J.* **2023**, 137.

- [17] Luo, Q.; Cheng, Y.; Yang, J.; Cao, R.; Ma, H.; Yang, Y.; Huang, R.; Wei, W.; Zheng, Y.; Gong, T.; Yu, J.; Xu, X.; Yuan, P.; Li, X.; Tai, L.; Yu, H.; Shang, D.; Liu, Q.; Yu, B.; Ren, Q.; Lv, H.; Liu, M. A highly CMOS compatible hafnia-based ferroelectric diode. *Nat. Commun.* **2020**, *11*, 1391.
- [18] Huang, Y.; Wu, Q.; Gong, T.; Yang, J.; Luo, Q.; Liu, M. Bayesian neural network with unified entropy source and synapse weights using 3D 16-layer Fe-diode array. *Nat. Commun.* **2025**, *16*, 8063.
- [19] Chua, L. How we predicted the memristor. *Nat. Electron.* **2018**, *1*, 322.
- [20] Hu, H.; Peng, J.; Liu, G.; Yu, S.; Zhao, Z.; Zhang, Y.; Zhang, C.; Li, Z.; Liu, H.; Xu, H.; Wang, Y. An efficient numerical simulation method based on practical 1T1R devices measurement for compute in memory chip design. *Integration* **2025**, *103*.
- [21] Kent, A. D.; Worledge, D. C. A new spin on magnetic memories. *Nat. Nanotechnol.* **2015**, *10*, 187-191.
- [22] Moll, J. L.; Tarui, Y. A new solid state memory resistor. *IEEE Trans. Electron Devices* **1963**, *10*, 338-338.
- [23] Luo, Z. D.; Zhang, S.; Liu, Y.; Zhang, D.; Gan, X.; Seidel, J.; Liu, Y.; Han, G.; Alexe, M.; Hao, Y. Dual-Ferroelectric-Coupling-Engineered Two-Dimensional Transistors for Multifunctional In-Memory Computing. *ACS Nano* **2022**, 10.1021/acsnano.2c00079.
- [24] Böске, T.; Müller, J.; Braeuhaus, D.; Schroeder, U.; Bottger, U. Ferroelectricity in Hafnium Oxide Thin Films. *Appl. Phys. Lett.* **2011**, *99*, 102903-102903.
- [25] Mikolajick, T.; Slesazek, S.; Park, M. H.; Schroeder, U. Ferroelectric hafnium oxide for ferroelectric random-access memories and ferroelectric field-effect transistors. *MRS Bull.* **2018**, *43*, 340-346.
- [26] Yurchuk, E.; Muller, J.; Paul, J.; Schlosser, T.; Martin, D.; Hoffmann, R.; Mueller, S.; Slesazek, S.; Schroeder, U.; Boschke, R.; van Bentum, R.; Mikolajick, T. Impact of Scaling on the Performance of HfO₂-Based Ferroelectric Field Effect Transistors. *IEEE Trans. Electron Devices* **2014**, *61*, 3699-3706.
- [27] Cheema, S. S.; Kwon, D.; Shanker, N.; dos Reis, R.; Hsu, S.-L.; Xiao, J.; Zhang, H.; Wagner, R.; Datar, A.; McCarter, M. R.; Serrao, C. R.; Yadav, A. K.; Karbasian, G.; Hsu, C.-H.; Tan, A. J.; Wang, L.-C.; Thakare, V.; Zhang, X.; Mehta, A.; Karapetrova, E.; Chopdekar, R. V.; Shafer, P.; Arenholz, E.; Hu, C.; Proksch, R.; Ramesh, R.; Ciston, J.; Salahuddin, S. Enhanced ferroelectricity in ultrathin films grown directly on silicon. *Nature* **2020**, *580*, 478-482.
- [28] Lederer, M.; Kämpfe, T.; Vogel, N.; Utess, D.; Volkmann, B.; Ali, T.; Olivo, R.; Müller, J.; Beyer, S.; Trentzsch, M.; Seidel, K.; Eng, A. Structural and Electrical Comparison of Si and Zr Doped Hafnium Oxide Thin Films and Integrated FeFETs Utilizing Transmission Kikuchi Diffraction. *Nanomaterials* **2020**, *10*.
- [29] Lalenı, N.; Kämpfe, T. An Efficient FeFET-Based Compute-In-Memory Macro Readout Inspired by the 3T Pixel Array Cell. *2024 20th International*

Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD) **2024**, 10.1109/smacd61181.2024.10745435, 1-4.

[30] Hyuk Park, M.;Joon Kim, H.;Jin Kim, Y.;Lee, W.;Moon, T.; Seong Hwang, C. Evolution of phases and ferroelectric properties of thin $\text{Hf}_{0.5}\text{Zr}_{0.5}\text{O}_2$ films according to the thickness and annealing temperature. *Appl. Phys. Lett.* **2013**, 102.

[31] Park, M. H.;Lee, Y. H.;Kim, H. J.;Kim, Y. J.;Moon, T.;Kim, K. D.;Müller, J.;Kersch, A.;Schroeder, U.;Mikolajick, T.; Hwang, C. S. Ferroelectricity and Antiferroelectricity of Doped Thin HfO_2 - Based Films. *Adv. Mater.* **2015**, 27, 1811-1831.

[32] Materlik, R.;Künneth, C.; Kersch, A. The origin of ferroelectricity in $\text{Hf}_{1-x}\text{Zr}_x\text{O}_2$: A computational investigation and a surface energy model. *J. Appl. Phys.* **2015**, 117.

[33] Wang, Y.;Tao, L.;Guzman, R.;Luo, Q.;Zhou, W.;Yang, Y.;Wei, Y.;Liu, Y.;Jiang, P.;Chen, Y.;Lv, S.;Ding, Y.;Wei, W.;Gong, T.;Wang, Y.;Liu, Q.;Du, S.; Liu, M. A stable rhombohedral phase in ferroelectric $\text{Hf}(\text{Zr})_{1-x}\text{O}_2$ capacitor with ultralow coercive field. *Science* **2023**, 381, 558-563.

[34] Sang, X.;Grimley, E. D.;Schenk, T.;Schroeder, U.; LeBeau, J. M. On the structural origins of ferroelectricity in HfO_2 thin films. *Appl. Phys. Lett.* **2015**, 106.

[35] Chen, H.;Zhou, X.;Tang, L.;Chen, Y.;Luo, H.;Yuan, X.;Bowen, C.; Zhang, D. HfO_2 -based ferroelectrics: From enhancing performance, material design, to applications. *Appl. Phys. Rev.* **2022**, 10.1063/5.0066607.

[36] Dutta, S.;Buragohain, P.;Glinsek, S.;Richter, C.;Aramberri, H.;Lu, H.;Schroeder, U.;Defay, E.;Gruverman, A.; Íñiguez, J. Piezoelectricity in hafnia. *Nat. Commun.* **2021**, 12, 7301.

[37] Zhu, Z.;Rui, G.;Allahyarov, E.;Zhang, H.;Li, R.;Taylor, P. L.; Zhu, L. Mechanisms of direct and converse piezoelectricity in ferroelectric polymers. *Polymer* **2025**, 325.

[38] Raeliarijaona, A.; Cohen, R. E. Hafnia HfO_2 is a proper ferroelectric. *Physical Review B* **2023**, 108, 094109.

[39] Lübben, J.;Berg, F.; Böttger, U. Investigation of the large-signal electromechanical behavior of ferroelectric HfO_2 - CeO_2 thin films prepared by chemical solution deposition. *J. Appl. Phys.* **2024**, 135.

[40] Tan, H.;Estandía, S.;Sánchez, F.; Fina, I. Effects of Doping, Stress, and Thickness on the Piezoelectric Response and Its Relation with Polarization in Ferroelectric HfO_2 . *ACS Appl. Electron. Mater.* **2023**, 5, 6630-6639.

[41] Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review* **1965**, 140, A1133-A1138.

[42] Cheng, H.;Jiao, P.;Wang, J.;Qing, M.;Deng, Y.;Liu, J.-M.;Bellaiche, L.;Wu, D.; Yang, Y. Tunable and parabolic piezoelectricity in hafnia under epitaxial strain. *Nat. Commun.* **2024**, 15, 394.

[43] Chen, L.;Liang, Z.;Shao, S.;Huang, Q.;Tang, K.; Huang, R. First direct observation of the built-in electric field and oxygen vacancy migration in ferroelectric Hf_{0.5}Zr_{0.5}O₂ film during electrical cycling. *Nanoscale* **2023**, *15*, 7014-7022.

[44] Zhu, T.;Ma, L.;Deng, S.; Liu, S. Progress in computational understanding of ferroelectric mechanisms in HfO₂. *npj Comput. Mater.* **2024**, *10*.

[45] Batra, R.;Huan, T. D.;Rossetti, G. A., Jr.; Ramprasad, R. Dopants Promoting Ferroelectricity in Hafnia: Insights from a comprehensive Chemical Space Exploration. *Chem. Mater.* **2017**, *29*, 9102-9109.

[46] Ma, L.;Wu, J.;Zhu, T.;Huang, Y.;Lu, Q.; Liu, S. Ultrahigh Oxygen Ion Mobility in Ferroelectric Hafnia. *Phys. Rev. Lett.* **2023**, *131*.

[47] Jayakrishnan, A.;Kim, J.;Hellenbrand, M.;Marques, L.;MacManus-Driscoll, J.; Silva, J. Growth of emergent simple pseudo-binary ferroelectrics and their potential in neuromorphic computing devices. *Mater. Horiz.* **2024**, *11*, 2355-2371.

[48] Lomenzo, P. D.;Zhao, P.;Takmeel, Q.;Moghaddam, S.;Nishida, T.;Nelson, M.;Fancher, C. M.;Grimley, E. D.;Sang, X.;LeBeau, J. M.; Jones, J. L. Ferroelectric phenomena in Si-doped HfO₂ thin films with TiN and Ir electrodes. *Journal of Vacuum Science & Technology B* **2014**, *32*.

[49] Park, M.;Chung, C.;Schenk, T.;Richter, C.;Opsomer, K.;Detavernier, C.;Adelmann, C.;Jones, J.;Mikolajick, T.; Schroeder, U. Effect of Annealing Ferroelectric HfO₂ Thin Films: In Situ, High Temperature X - Ray Diffraction. *Adv. Electron. Mater.* **2018**, *4*.

[50] HU, Z.-N.; LO, V. C. THICKNESS DEPENDENCE OF THE COERCIVE FIELD IN FERROELECTRIC THIN FILMS. *International Journal of Modern Physics B* **2006**, *20*, 3223-3231.

[51] Samanta, S.;Anoop, G.;Joh, H.;Seol, W.;Park, S. M.;Unithrattil, S.;Lee, J.;Kim, T. Y.;Kim, H.;Yeom, J.;Hong, S.; Jo, J. Multi - Step Chemical Solution Deposition - Annealing Process Toward Wake - Up Free Ferroelectricity in Y:HfO₂ Films. *Adv. Mater. Interfaces* **2021**, *8*.

[52] Tan, Y.;Wu, Y.;Duan, T.;Xiong, R.;Jiang, J.;Liao, M.; Yang, Q. Antiferroelectric-ferroelectric phase transition of HfO₂ and its influencing factors. *Comput. Mater. Sci* **2025**, *253*.

[53] Chouprik, A.;Spiridonov, M.;Zarubin, S.;Kirtaev, R.;Mikheev, V.;Lebedinskii, Y.;Zakharchenko, S.; Negrov, D. Wake-Up in a Hf_{0.5}Zr_{0.5}O₂ Film: A Cycle-by-Cycle Emergence of the Remnant Polarization via the Domain Depinning and the Vanishing of the Anomalous Polarization Switching. *ACS Appl. Electron. Mater.* **2019**, *10*.1021/acsaelm.8b00046.

[54] Scott, J. F. Charge Injection and Fatigue. In *Ferroelectric Memories*. Springer Berlin Heidelberg; Berlin, Heidelberg, 2000; pp 133-143.

[55] Yoon, S.-J.;Min, D.-H.;Moon, S.-E.;Park, K. S.;Won, J. I.; Yoon, S.-M. 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* **2020**, *67*, 499-504.

[56] Brennan, C. J.; Parrella, R. D.; Larsen, D. E. Temperature dependent fatigue rates in thin-film ferroelectric capacitors. *Ferroelectrics* **1994**, *151*, 33-38.

[57] Li, X.; Tai, L.; Sang, P.; Dou, X.; Zhan, X.; Xu, H.; Wang, X.; Wu, J.; Chen, J. Imprint-Correlated Retention Loss in Hf_{0.5}Zr_{0.5}O₂ Ferroelectric Thin Film Through Wide-Temperature Characterizations. *IEEE Trans. Electron Devices* **2024**, *71*, 5361-5366.

[58] Yue, Y.; Jing, X.; Zheng, W.; Li, J.; Zheng, C.; Fan, X.; Song, K.; Cao, T.; Wen, Z. Engineering phase transition and ferroelectric properties in acceptor-doped HfO₂/SrRuO₃ thin-film heterostructures by a charge-balance synergistic strategy. *J. Materiomics* **2025**, *11*.

[59] Li, Q.; Wang, S.; Li, Z.; Hu, X.; Liu, Y.; Yu, J.; Yang, Y.; Wang, T.; Meng, J.; Sun, Q.; Zhang, D. W.; Chen, L. High-performance ferroelectric field-effect transistors with ultra-thin indium tin oxide channels for flexible and transparent electronics. *Nat. Commun.* **2024**, *15*.

[60] Dahan, M. M.; Mulaosmanovic, H.; Levit, O.; Dunkel, S.; Müller, J.; Beyer, S.; Yalon, E. Origin of charges in bulk Si:HfO₂ FeFET probed by nanosecond polarization measurements. *Microelectron. Eng.* **2025**, 296.

[61] Raffel, Y.; De, S.; Lederer, M.; Olivo, R.; Hoffmann, R.; Thunder, S.; Pirro, L.; Beyer, S.; Chohan, T.; Kämpfe, T.; Seidel, K.; Heitmann, J. Synergistic Approach of Interfacial Layer Engineering and READ-Voltage Optimization in HfO₂-Based FeFETs for In-Memory-Computing Applications. *ACS Appl. Electron. Mater.* **2022**, *4*, 5292-5300.

[62] Liu, W.-D.; Huang, Z.-Y.; Ma, J.; Zheng, Z.-W.; Cheng, C.-H. Impact of Series-Connected Ferroelectric Capacitor in HfO₂-Based Ferroelectric Field-Effect Transistors for Memory Application. *IEEE J. Electron Devices Soc.* **2020**, *8*, 1076-1081.

[63] Jerry, M.; Chen, P. Y.; Zhang, J.; Sharma, P.; Ni, K.; Yu, S.; Datta, S. Ferroelectric FET analog synapse for acceleration of deep neural network training. In *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017; pp 6.2.1-6.2.4.

[64] Sheridan, P. M.; Cai, F.; Du, C.; Ma, W.; Zhang, Z.; Lu, W. D. Sparse coding with memristor networks. *Nat. Nanotechnol.* **2017**, *12*, 784-789.

[65] Ju, S.; Jang, H.; Park, W.; Jung, S.; Kim, S. Emulating nociceptor and synaptic functions in GaOx-based resistive random-access memory for bio-inspired computing. *Appl. Surf. Sci.* **2025**, 697.

[66] Burr, G. W.; Shelby, R. M.; Sebastian, A.; Kim, S.; Kim, S.; Sidler, S.; Virwani, K.; Ishii, M.; Narayanan, P.; Fumarola, A.; Sanches, L. L.; Boybat, I.; Le Gallo, M.; Moon, K.; Woo, J.; Hwang, H.; Leblebici, Y. Neuromorphic computing using non-volatile memory. *Adv. Phys.: X* **2016**, *2*, 89-124.

[67] Kim, G.; Choi, H.; Lee, S.; Shin, H.; Lee, S.; Nam, Y.; Kang, H.; Shin, S.; Kim, H.; Lim, Y.; Kim, K.; Oh, I.; Park, S. H. K.; Ahn, J.; Jeon, S. Experimental Analysis on the Interaction Between Interface Trap Charges and Polarization on the Memory

Window of Metal-Ferroelectric-Insulator-Si (MFIS) FeFET. *IEEE Trans. Electron Devices* **2024**, *71*, 6627-6632.

[68] Wang, Y.-Y.; Wang, K.-C.; Wu, C.-H.; Chang, T.-Y.; Ronchi, N.; Banerjee, K.; Van Den Bosch, G.; Van Houdt, J.; Wu, T.-L. Demonstration of 64 Conductance States and Large Dynamic Range in Si-doped HfO₂ FeFETs under Neuromorphic Computing Operations. *2022 International Symposium on VLSI Technology, Systems and Applications (VLSI-TSA)* **2022**, 10.1109/vlsi-tsa54299.2022.9770994, 1-2.

[69] Cao, Y.; Yang, Y.; Li, Q.; Zhang, T.; Ji, L.; Zhu, H.; Chen, L.; Sun, Q.; Zhang, D. Structural Engineering of H_{0.5}Z_{0.5}O₂ - Based Ferroelectric Tunneling Junction for Fast - Speed and Low - Power Artificial Synapses. *Adv. Electron. Mater.* **2023**, *9*.

[70] Ma, C.; Luo, Z.; Huang, W.; Zhao, L.; Chen, Q.; Lin, Y.; Liu, X.; Chen, Z.; Liu, C.; Sun, H.; Jin, X.; Yin, Y.; Li, X. Sub-nanosecond memristor based on ferroelectric tunnel junction. *Nat. Commun.* **2020**, *11*, 1439.

[71] Chanthbouala, A.; Garcia, V.; Cherifi, R. O.; Bouzehouane, K.; Fusil, S.; Moya, X.; Xavier, S.; Yamada, H.; Deranlot, C.; Mathur, N. D.; Bibes, M.; Barthélémy, A.; Grollier, J. A ferroelectric memristor. *Nat. Mater.* **2012**, *11*, 860-864.

[72] Mulaosmanovic, H.; Breyer, E.; Dünkler, S.; Beyer, S.; Mikolajick, T.; Slesazeck, S. Ferroelectric field-effect transistors based on HfO₂: a review. *Nanotechnology* **2021**, *32*.

[73] Moon, T.; Lee, H.-Y.; Nam, S.; Bae, H.; Choe, D.-H.; Jo, S.; Lee, Y.; Park, Y.-H.; Yang, J.; Heo, J. Parallel synaptic design of ferroelectric tunnel junctions for neuromorphic computing. *Neuromorph. Comput. Eng.* **2023**, *3*.

[74] Shao, M.-H.; Zhao, R.-T.; Liu, H.; Xu, W.-J.; Guo, Y.-D.; Huang, D.-P.; Yang, Y.-Z.; Li, X.-R.; Shao, W.; Shen, P.-H.; Liu, J.; Wang, K.; Zheng, J.; Yan, Z.-Y.; Yan, J.-L.; Lu, T.; Yang, Y.; Ren, T.-L. Challenges and recent advances in HfO₂-based ferroelectric films for non-volatile memory applications. *Chip* **2024**, *3*.

[75] Luo, Q.; Chen, B.; Cao, R.; Xue, X.; Zhou, K.; Yang, J.; Zheng, X.; Yu, H.; Yu, J.; Gong, T.; Xu, X.; Yuan, P.; Li, X.; Tai, L.; Liu, Q.; Lv, H.; Liu, M. Complementary Memory Cell Based on Field-Programmable Ferroelectric Diode for Ultra-Low Power Current-SA Free BNN Applications. In *2019 IEEE International Electron Devices Meeting (IEDM)*, 2019; pp 38.35.31-38.35.34.

[76] Lien, A. S.-Y.; Lai, C.-Y.; Wei, J.-D.; Yang, H.-M.; Yeh, J.-T.; Tai, H.-C. A Granulation Tissue Detection Model to Track Chronic Wound Healing in DM Foot Ulcers. *Electronics* **2022**, *11*.

[77] Colinge, J.-P.; Lee, C.-W.; Afzal, A.; Akhavan, N. D.; Yan, R.; Ferain, I.; Razavi, P.; O'Neill, B.; Blake, A.; White, M.; Kelleher, A.-M.; McCarthy, B.; Murphy, R. Nanowire transistors without junctions. *Nat. Nanotechnol.* **2010**, *5*, 225-229.

[78] Dragoman, M.; Dinescu, A.; Dragoman, D.; Palade, C.; Moldovan, A.; Dinescu, M.; Teodorescu, V.; Ciurea, M. Wafer-scale graphene-ferroelectric HfO₂/Ge-HfO₂/HfO₂ transistors acting as three-terminal memristors. *Nanotechnology* **2020**, *31*.

[79] Leblanc, C.; Song, S.; Jariwala, D. 2D Ferroelectrics and ferroelectrics with

2D: Materials and device prospects. *Curr. Opin. Solid State Mater. Sci.* **2024**, *32*.

[80] Xu, X.;Luo, Z.;Sun, H.;Xu, Y.;Gao, L.; Yu, Z. A review of hafnium-based ferroelectrics for advanced computing. *Solid-State Electron.* **2025**, *225*, 109053.

[81] Torrejon, J.;Riou, M.;Araujo, F. A.;Tsunegi, S.;Khalsa, G.;Querlioz, D.;Bortolotti, P.;Cros, V.;Yakushiji, K.;Fukushima, A.;Kubota, H.;Yuasa, S.;Stiles, M. D.; Grollier, J. Neuromorphic computing with nanoscale spintronic oscillators. *Nature* **2017**, *547*, 428-431.

[82] Lu, P. S.;Lin, C. C.; Su, P. Impact of Multi-Domain Interaction on ON-State Characteristics of MFIS-Type 2D Negative-Capacitance FETs. In *2019 International Symposium on VLSI Technology, Systems and Application (VLSI-TSA)*, 2019; pp 1-2.

[83] Liu, M.;Kang, J.;Zhang, J.;Liu, X.; Han, R. A MOS-based behavioral macro-model for ferroelectric capacitors. *Microelectron. Eng.* **2003**, *66*, 813-817.

[84] Dragoman, M.;Aldrigo, M.;Dragoman, D.;Iordanescu, S.;Dinescu, A.; Modreanu, M. HfO₂ - Based Ferroelectrics Applications in Nanoelectronics. *physica status solidi (RRL) – Rapid Research Letters* **2021**, *15*.

[85] Salahuddin, S.; Datta, S. Can the subthreshold swing in a classical FET be lowered below 60 mV/decade? In *2008 IEEE International Electron Devices Meeting*, 2008; pp 1-4.

[86] Khan, A. I.;Keshavarzi, A.; Datta, S. The future of ferroelectric field-effect transistor technology. *Nat. Electron.* **2020**, *3*, 588-597.

[87] Mulaosmanovic, H.;Chicca, E.;Bertele, M.;Mikolajick, T.; Slesazek, S. Mimicking biological neurons with a nanoscale ferroelectric transistor. *Nanoscale* **2018**, *10* 46, 21755-21763.

[88] Song, M.;Tang, Q.;Ouyang, X.;Duan, W.;Xu, Y.;Zhang, S.; You, L. SOT-MRAM-based true in-memory computing architecture for approximate multiplication. *Chip* **2025**, *4*.

[89] Soliman, T.;Chatterjee, S.;Laleni, N.;Müller, F.;Kirchner, T.;Wehn, N.;Kämpfe, T.;Chauhan, Y. S.; Amrouch, H. First demonstration of in-memory computing crossbar using multi-level Cell FeFET. *Nat. Commun.* **2023**, *14*, 6348.

[90] Yu, S. Neuro-Inspired Computing With Emerging Nonvolatile Memorys. *Proc. IEEE* **2018**, *106*, 260-285.

[91] Samanta, P.;Yadav, D. N.;Das, P. P.; Sengupta, I. An efficient read approach for memristive crossbar array. *Memories - Materials, Devices, Circuits and Systems* **2023**, *4*.

[92] Gokmen, T.;Onen, M.; Haensch, W. Training Deep Convolutional Neural Networks with Resistive Cross-Point Devices. *Front. Neurosci.* **2017**, *11*.

[93] Vignali, R.;Zurla, R.;Pasotti, M.;Rolandi, P. L.;Singh, A.;Gallo, M. L.;Sebastian, A.;Jang, T.;Antolini, A.;Scarselli, E. F.; Cabrini, A. Designing Circuits for AiMC Based on Non-Volatile Memories: A Tutorial Brief on Trade-Off and Strategies for ADCs and DACs Co-Design. *IEEE Trans. Circuits Syst. II Express Briefs* **2024**, *71*, 1650-1655.

[94] Majumdar, S. Back-End CMOS Compatible and Flexible Ferroelectric

Memories for Neuromorphic Computing and Adaptive Sensing. *Adv. Intell. Syst.* **2022**, *4*, 2100175.

[95] Toprasertpong, K.; Takenaka, M.; Takagi, S. On the strong coupling of polarization and charge trapping in HfO₂/Si-based ferroelectric field-effect transistors: overview of device operation and reliability. *Appl. Phys. A* **2022**, *128*.

[96] Sebastian, A.; Tuma, T.; Papandreou, N.; Le Gallo, M.; Kull, L.; Parnell, T.; Eleftheriou, E. Temporal correlation detection using computational phase-change memory. *Nat. Commun.* **2017**, *8*.

[97] Chen, Y.-C.; Lin, C.-C.; Chang, Y.-F. Post-Moore Memory Technology: Sneak Path Current (SPC) Phenomena on RRAM Crossbar Array and Solutions. *Micromachines* **2021**, *12*, 50.

[98] Agrawal, A.; Lee, C.; Roy, K. X-CHANGR: Changing Memristive Crossbar Mapping for Mitigating Line-Resistance Induced Accuracy Degradation in Deep Neural Networks. arXiv, Ed., 2019.

[99] Negi, S.; Saxena, U.; Sharma, D.; Roy, K. HCiM: ADC-Less Hybrid Analog-Digital Compute in Memory Accelerator for Deep Learning Workloads. arXiv, Ed., 2024.

[100] Chen, J.; Xu, J.; Gu, J.; Chen, B.; Zhang, H.; Qian, H.; Liu, H.; Shen, R.; Lin, G.; Yu, X.; Zhang, M.; Ding, Y.; Liu, Y.; Tang, J.; Wu, H.; Jin, C.; Han, G. Low-power edge detection based on ferroelectric field-effect transistor. *Nat. Commun.* **2025**, *16*.

[101] Wang, Z.; Li, C.; Song, W.; Rao, M.; Belkin, D.; Li, Y.; Yan, P.; Jiang, H.; Lin, P.; Hu, M.; Strachan, J. P.; Ge, N.; Barnell, M.; Wu, Q.; Barto, A. G.; Qiu, Q.; Williams, R. S.; Xia, Q.; Yang, J. J. Reinforcement learning with analogue memristor arrays. *Nat. Electron.* **2019**, *2*, 115-124.

[102] Shen, P.-C.; Lin, C.; Wang, H.; Teo, K.; Kong, J. Ferroelectric memory field-effect transistors using CVD monolayer MoS₂ as resistive switching channel. *Appl. Phys. Lett.* **2020**, *116*, 33501.

[103] Dang, Z.; Guo, F.; Duan, H.; Zhao, Q.; Fu, Y.; Jie, W.; Jin, K.; Hao, J. Black Phosphorus/Ferroelectric P(VDF-TrFE) Field-Effect Transistors with High Mobility for Energy-Efficient Artificial Synapse in High-Accuracy Neuromorphic Computing. *Nano Lett.* **2023**, 10.1021/acs.nanolett.3c01687.

[104] Akinwande, D.; Huyghebaert, C.; Wang, C.-H.; Serna, M. I.; Goossens, S.; Li, L.-J.; Wong, H. S. P.; Koppens, F. H. L. Graphene and two-dimensional materials for silicon technology. *Nature* **2019**, *573*, 507-518.

[105] Ding, W.; Zhu, J.; Wang, Z.; Gao, Y.; Xiao, D.; Gu, Y.; Zhang, Z.; Zhu, W. Prediction of intrinsic two-dimensional ferroelectrics in In₂Se₃ and other III₂-VI₃ van der Waals materials. *Nat. Commun.* **2017**, *8*.

[106] Du, J. y.; Ge, C.; Riahi, H.; Guo, E. j.; He, M.; Wang, C.; Yang, G. z.; Jin, K. j. Dual - Gated MoS₂ Transistors for Synaptic and Programmable Logic Functions. *Adv. Electron. Mater.* **2020**, *6*.

[107] Shi, B.; Qi, P.; Jiang, M.; Dai, Y.; Lin, F.; Zhang, H.; Fang, Z. Exotic physical properties of 2D materials modulated by moiré superlattices. *Mater. Adv.*

2021, 2, 5542-5559.

[108] Tan, A. J.; Liao, Y. H.; Wang, L. C.; Shanker, N.; Bae, J. H.; Hu, C. M.; Salahuddin, S. Ferroelectric HfO₂ Memory Transistors With High- κ Interfacial Layer and Write Endurance Exceeding 10¹⁰ Cycles. *IEEE Electron Device Lett.* **2021**, 42, 994-997.

[109] Mulaosmanovic, H.; Mikolajick, T.; Slesazek, S. Random Number Generation Based on Ferroelectric Switching. *IEEE Electron Device Lett.* **2018**, 39, 135-138.

[110] Kanerva, P. Hyperdimensional Computing: An Introduction to Computing in Distributed Representation with High-Dimensional Random Vectors. *Cognit. Comput.* **2009**, 1, 139-159.

[111] Nikonov, D. E.; Csaba, G.; Porod, W.; Shibata, T.; Voils, D.; Hammerstrom, D.; Young, I. A.; Bourianoff, G. I. Coupled-Oscillator Associative Memory Array Operation for Pattern Recognition. *IEEE J. Explor. Solid-State Comput. Devices Circuits* **2015**, 1, 85-93.

[112] Camsari, K. Y.; Sutton, B. M.; Datta, S. p-bits for probabilistic spin logic. *Appl. Phys. Rev.* **2019**, 6.

[113] Pyo, S.; Lee, J.; Bae, K.; Sim, S.; Kim, J. Recent Progress in Flexible Tactile Sensors for Human-Interactive Systems: From Sensors to Advanced Applications (Adv. Mater. 47/2021). *Adv. Mater.* **2021**, 33, 2170373.

[114] Garg, S.; Kaur, J.; Goel, A.; Haldar, S.; Gupta, R. S. High-K, Gate-Stack Engineered, Junctionless, Nanowire Ferroelectric Field Effect Transistor (J-NW-Fe-Fet) with various catalytic metals for PH₃ Sensing Applications. In *2025 Devices for Integrated Circuit (DevIC)*, 2025; pp 401-406.

[115] Mannsfeld, S. C. B.; Tee, B. C. K.; Stoltenberg, R. M.; Chen, C. V. H. H.; Barman, S.; Muir, B. V. O.; Sokolov, A. N.; Reese, C.; Bao, Z. Highly sensitive flexible pressure sensors with microstructured rubber dielectric layers. *Nat. Mater.* **2010**, 9, 859-864.

[116] Wang, D.; Hao, S.; Dkhil, B.; Tian, B.; Duan, C. Ferroelectric materials for neuroinspired computing applications. *Fundam. Res.* **2023**, 4, 1272-1291.

[117] Roy, K.; Jaiswal, A.; Panda, P. Towards spike-based machine intelligence with neuromorphic computing. *Nature* **2019**, 575, 607-617.

[118] Wang, Z. Ferroelectric Materials for Synaptic Transistors and Their Neuromorphic Applications. *ArXiv* **2024**, abs/2406.13946.

[119] Pešić, M.; Schroeder, U.; Slesazek, S.; Mikolajick, T. Comparative Study of Reliability of Ferroelectric and Anti-Ferroelectric Memories. *IEEE Trans. Device Mater. Reliab.* **2018**, 18, 154-162.

[120] Yu, E.; K, G. K.; Saxena, U.; Roy, K. Ferroelectric capacitors and field-effect transistors as in-memory computing elements for machine learning workloads. *Sci. Rep.* **2024**, 14.

[121] Oh, S.; Hwang, H.; Yoo, I. Ferroelectric materials for neuromorphic computing. *APL Mater.* **2019**, 10.1063/1.5108562.

- [122] Yin, X.;Li, C.;Huang, Q.;Zhang, L.;Niemier, M.;Hu, X.;Zhuo, C.; Ni, K. FeCAM: A Universal Compact Digital and Analog Content Addressable Memory Using Ferroelectric. *IEEE Trans. Electron Devices* **2020**, *67*, 2785-2792.
- [123] Ram, A.;Maity, K.;Marchand, C.;Mahmoudi, A.;Kshirsagar, A.;Soliman, M.;Taniguchi, T.;Watanabe, K.;Doudin, B.;Ouerghi, A.;Reichardt, S.;O'Connor, I.; Dayen, J. Reconfigurable Multifunctional van der Waals Ferroelectric Devices and Logic Circuits. *ACS Nano* **2023**, 10.1021/acsnano.3c07952.
- [124] Majumdar, S. Back - End CMOS Compatible and Flexible Ferroelectric Memories for Neuromorphic Computing and Adaptive Sensing. *Adv. Intell. Syst.* **2021**, *4*.
- [125] Dahan, M.;Breyer, E.;Slesazeck, S.;Mikolajick, T.; Kvatinsky, S. C-AND: Mixed Writing Scheme for Disturb Reduction in 1T Ferroelectric FET Memory. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2022**, *69*, 1595-1605.
- [126] Park, J. Y.;Choe, D.-H.;Lee, D. H.;Yu, G.;Yang, K.;Kim, S. H.;Park, G.;Nam, S.;Lee, H.-Y.;Jo, S.;Kuh, B.;Ha, D.;Kim, Y.;Heo, J.; Park, M. Revival of Ferroelectric Memories Based on Emerging Fluorite - Structured Ferroelectrics. *Adv. Mater.* **2022**, *35*.
- [127] Lehninger, D.;Müller, F.;Raffel, Y.;Yang, S.;Neuber, M.;Abdulazhanov, S.;Kämpfe, T.;Seidel, K.; Lederer, M. Ferroelectric Hafnium Oxide: A Potential Game - Changer for Nanoelectronic Devices and Systems. *Adv. Electron. Mater.* **2025**, *11*.
- [128] Chen, Z.;Li, W.;Fan, Z.;Dong, S.;Chen, Y.;Qin, M.;Zeng, M.;Lu, X.;Zhou, G.;Gao, X.; Liu, J.-M. All-ferroelectric implementation of reservoir computing. *Nat. Commun.* **2023**, *14*.
- [129] Feng, G.;Zhu, Q.;Liu, X.;Chen, L.;Zhao, X.;Liu, J.;Xiong, S.;Shan, K.;Yang, Z.;Bao, Q.;Yue, F.;Peng, H.;Huang, R.;Tang, X.;Jiang, J.;Tang, W.;Guo, X.;Wang, J.;Jiang, A.;Dkhil, B.;Tian, B.;Chu, J.; Duan, C. A ferroelectric fin diode for robust non-volatile memory. *Nat. Commun.* **2024**, *15*.
- [130] Yao, K.;Chen, S.;Lai, S.; Yousry, Y. M. Enabling Distributed Intelligence with Ferroelectric Multifunctionalities. *Adv. Sci.* **2021**, *9*.