Advanced Electronic Materials

ADVANCED ELECTRONIC MATERIALS Open Access

www.advelectronicmat.de

ARTICLE OPEN ACCESS

A Dual-Memory Ferroelectric Transistor Emulating Synaptic Metaplasticity for High-Speed Reservoir Computing

Yifan Wang[1] | Muhammad Sakib Shahriar[2] | Salma Soliman[3] | Noah Vaillancourt[4] | Lance Fernandes[3] | Andrea Padovani[5] | Asif Islam Khan[3] | Md Sakib Hasan[2] | Raisul Islam[1,4]

[1]Elmore Family School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana, USA | [2]Department of Electrical and Computer Engineering, University of Mississippi, Mississippi, USA | [3]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA | [4]School of Materials Engineering, Purdue University, West Lafayette, Indiana, USA | [5]Dipartimento di Scienze e Metodi dell'Ingegneria, Universitá di Modena e Reggio Emilia, Reggio Emilia, Italy

Correspondence: Md Sakib Hasan (mhasan5@olemiss.edu) | Raisul Islam (raisul@purdue.edu)

Received: 24 November 2025 | **Revised:** 24 March 2026 | **Accepted:** 1 May 2026

Keywords: ferroelectric field-effect transistor (FeFET) | neuromorphic computing | reservoir computing | short-term memory

ABSTRACT

The exponential growth of edge artificial intelligence demands material-focused solutions to overcome energy consumption and latency limitations when processing real-time temporal data. Physical reservoir computing (PRC) offers an energy-efficient paradigm but faces challenges due to limited device scalability and reconfigurability. Additionally, reservoir and readout layers require memory of different timescales, short-term and long-term respectively-a material challenge hindering CMOS-compatible implementations. This work demonstrates a CMOS-compatible ferroelectric transistor using hafnium-zirconium-oxide (HZO) and silicon, enabling dual-memory operation. This system exhibits non-volatile long-term memory (LTM) from ferroelectric HZO polarization and volatile short-term memory (STM) from engineered non-quasi-static (NQS) channel-charge relaxation driven by gate-source/drain overlap capacitance. Ferroelectric polarization acts as non-volatile programming of volatile dynamics: by modulating threshold voltage, the ferroelectric state deterministically switches the NQS time constant and computational behavior between paired-pulse facilitation (PPF) and depression (PPD). This establishes a generalizable material-design principle applicable to diverse ferroelectric-semiconductor heterostructures, extending beyond silicon to oxide semiconductors and heterogeneously-integrated systems. The device solves second-order nonlinear tasks with 3.69×10^{-3} normalized error using only 16 reservoir states-~5× reduction—achieving 20 μs response time (~1000× faster) and 1.5×10^{-7} J energy consumption, providing an immediately manufacturable pathway for neuromorphic hardware and energy-efficient edge intelligence.

1 | Introduction

Reservoir computing (RC) is a powerful and efficient computational paradigm for processing complex temporal and sequential data, offering a substantial reduction in training overhead compared to conventional recurrent neural networks (RNNs). Since the reservoir, functioning as the core nonlinear dynamic system, remains fixed and only the readout layer is trained, RC achieves much faster learning, resulting in a lower computational cost [1]. This streamlined training process, combined with the reservoirs inherent ability to project inputs into a high-dimensional feature space, makes RC exceptionally well-suited for resource-constrained and low-latency edge-computing applications, including speech recognition [2], image classification

[3, 4], and temporal sensor-signals analysis like electrocardiogram (ECG) [5, 6]. Physical reservoir computing (PRC) has emerged as a promising alternative to traditional RNNs, utilizing diverse nonlinear dynamic physical platforms, including electronic [7–9], photonic [10, 11], mechanical [12, 13], biological [14, 15], quantum [16], and spintronic [17] mechanisms, to achieve rapid information processing with minimal learning costs. The fundamental principle is that any nonlinear dynamic system can function as a reservoir if it satisfies four key criteria [1]: (i) high dimensionality, (ii) nonlinearity, (iii) fading (short-term) memory [18, 19], and (iv) the separation property, which ensures that distinct inputs are mapped to separable states in the high-dimensional space. Consequently, practical physical reservoir computing demands material platforms that simultaneously satisfy competing requirements: CMOS manufacturability, nonlinear dynamics with volatile memory, non-volatile memory, and reconfigurable behavior. Achieving these distinct functionalities typically requires the co-integration of disparate materials or complex heterostructures, which often introduces non-uniformity and processing incompatibilities. Therefore, state-of-the-art implementations of PRC face significant challenges, including device-to-device or cycle-to-cycle variations, ensuring stable and consistent readouts, limited reconfigurability, and overcoming inherent physical limitations on processing timescales [20, 21]. Solving all challenges in a CMOS-compatible heterogeneously-integrated system has not been demonstrated so far using PRC systems that utilize exotic materials for nonlinear dynamics.

In biological synapses, metaplasticity describes "the plasticity of synaptic plasticity": prior activity (or a persistent synaptic state) does not necessarily change baseline synaptic efficacy, but instead persistently regulates the ability, sign, or threshold for subsequent plastic changes. Functionally, metaplasticity can be viewed as a higher-order control variable that gates whether the same input history leads to facilitation vs. depression, or potentiation vs. depression, thereby stabilizing learning while preserving adaptability [22, 23]. Short-term synaptic plasticity (e.g., paired-pulse facilitation/depression, PPF/PPD) provides a canonical measurable manifestation of such history-dependent modulation on short timescales [24].

In recent years, there has been growing research interest in CMOS-compatible material platforms for emulating neural behavior [25]. Ferroelectric field-effect transistors (FeFETs) employing hafnium zirconium oxide (e.g. $Hf_{0.5}Zr_{0.5}O_2$/HZO) as the gate dielectric have emerged as one of the promising physical reservoir-computing platforms due to their fast, scalable memory capability and CMOS compatibility. A notable example is the silicon-channel FeFET incorporating HZO, which uses ferroelectric polarization dynamics and polarization—charge coupling to effectively demonstrate short-term memory and high-dimensional nonlinear transformations, successfully solving nonlinear time-series prediction tasks via simple regression models [26]. However, a key limitation of this approach is its reliance on binary or discrete input signals for complex tasks like the second-order nonlinear dynamic system (SONDS) prediction, wherein the utilization of discrete input signals defeats the purpose of employing PRC in real-time temporal signal processing. This constraint arises because the STM mechanism of this work relies on the gate-driven switching of ferroelectric domains, a process more naturally suited for distinct and non-volatile state changes rather than the volatility in storing the previous states required for continuous signals. In effect, this approach does not show a true STM but an effective STM through limited sampling using a temporal mask function. Additional demonstrations of physical reservoir computing based on ferroelectric polarization switching also encounter notable constraints. Liu et al. reported a PRC system employing a ferroelectric semiconductor (α-In_2Se_3) channel transistor, constructed by vertically stacking multiple α-In_2Se_3 layers [27]. While functional, this architecture lacks CMOS compatibility and scalability. Furthermore, it does not permit independent modulation of short-term and long-term memory dynamics and suffers from a slow temporal response (~100 ms). Similarly, Ju et al. demonstrated a PRC implementation using a HfAlO-based memristor [28], in which ferroelectric polarization switching modulates the tunnel barrier, with partial polarization switching utilized to emulate short-term memory. However, the resulting relaxation times (τ)—ranging from milliseconds to seconds—are prohibitively long, preventing the complete erasure of the conductive state at extended inter-pulse intervals.

Our research specifically exploits a FeFET structure integrating silicon channels with $Hf_{0.5}Zr_{0.5}O_2$ (HZO) ferroelectric material. The device uniquely incorporates dual memory modalities: long-term memory (LTM) derived from ferroelectric gate polarization switching and short-term memory (STM) arising from a non-quasi-static (NQS) channel-charge response that is engineered by the gate—source/drain overlap capacitance. Under appropriate drain-voltage pulsing, the overlap provides a localized capacitive coupling at the drain edge; because the inversion channel behaves as a distributed, bias-dependent RC network where the perturbed channel charge relaxes with a finite time constant, τ, equivalent to the relaxation time of the STM. (Details on the NQS model of transistors and its corresponding equivalent RC circuit model are detailed in Figure S13 and the Supporting Note that follows.) This yields microsecond-scale STM expressed as tunable paired-pulse facilitation/depression (PPF/PPD) in the I_d—V_{ds} transients. Because the STM mechanism is overlap-capacitance-driven NQS rather than ferroelectric switching or trapping, it is intrinsically CMOS-scalable: relaxation time, $\tau \sim R_{ch}C_{ov}$ can be tuned by geometry/materials (to set C_{ov}) and bias/length (which set R_{ch}), enabling high speed at low energy without sacrificing compatibility. As ferroelectric polarization modulates the effective threshold voltage, it selects the operating regime (strong versus weak inversion) and thereby programmatically tunes the sign and strength of the NQS STM under a fixed V_{gs}. This co-action, where a persistent, nonvolatile state configures the subsequent volatile plastic response, is directly analogous to the classical operational definition of synaptic metaplasticity as a persistent change in the propensity/sign of later plasticity [22, 23]. In our device, the ferroelectric polarization (LTM) acts as the enduring meta-state that deterministically selects/modulates the sign and strength of the NQS relaxation dynamics (STM), enabling programmable switching between PPF and PPD under otherwise identical pulsing conditions.

As a test case for this combined LTM+STM operation, we specifically demonstrate the capacity of our silicon—HZO FeFET reservoir to solve second-order nonlinear dynamic tasks with high accuracy and efficiency. We perform systematic device and bias optimization, characterize memory dynamics and nonlinearity,

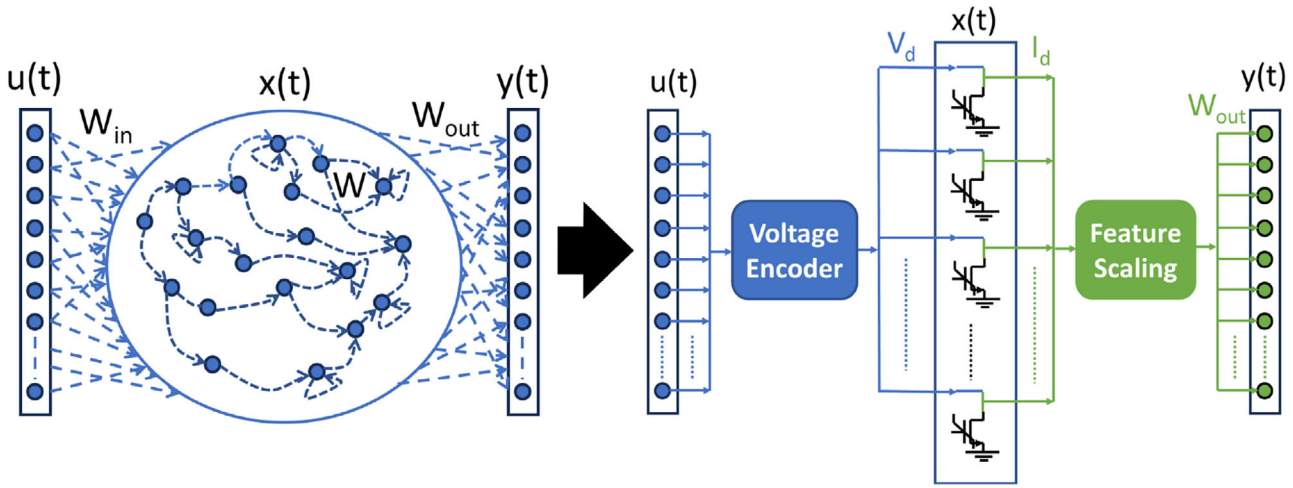

FIGURE 1 | A representation of a conventional reservoir and how it is converted into the FeFET reservoir. In the conventional approach (left), the reservoir is modeled as a recurrent neural network; the input weights (W_{in}) and the recurrent reservoir weights (W) remain fixed, while only the readout weights (W_{out}) are trained. In contrast, the FeFET reservoir (right) comprises heterogeneous FeFET devices exhibiting distinct behavior. The input sequence u(t) is first encoded into voltage signals, which drive the FeFET array. The resulting output currents are scaled according to device conductance to produce the reservoir states, which are then processed by the readout layer.

and benchmark performance. The results highlight the potential of co-integrated NQS STM and ferroelectric LTM in FeFET-based reservoir-computing architectures for next-generation, low-power edge computing. By co-engineering ferroelectric and NQS charge dynamics within a CMOS-compatible materials framework, this work establishes a scalable platform for physical reservoir computing with performance (1000× speed improvement, 5× efficiency gain) that previously required exotic, non-manufacturable material systems [29, 30]. Metaplasticity has also been implemented in two-terminal memristive devices, such as GQD-assisted FeO_x synapses that demonstrate history-dependent plasticity with pulse widths in the sub-millisecond range and synaptic decay timescales of tens of milliseconds or longer [31]. In contrast, our FeFET achieves deterministic LTM-programmed switching between PPF and PPD through overlap-capacitance-driven NQS dynamics operating on microsecond timescales (20 μs response time, 1.5×10^{-7} J), while maintaining CMOS material compatibility and nanoscale scalability in simulation. This combination addresses the speed, energy, and integration constraints essential for edge neuromorphic deployment. Figure 1 illustrates a conventional reservoir and its mapping to the proposed FeFET reservoir. Since amorphous oxide semiconductors (AOS) exhibit rich charge-trapping dynamics and bias-dependent percolation transport, the functionality of interconnected LTM and STM can also be demonstrated with FeFETs having AOS channel, potentially opening the path toward heterogeneously integrated CMOS+X type systems, where X can be overlap engineered FeFETs having AOS channel.

2 | Device Architecture and Characterization

The fabricated silicon-based FeFET (Figure 2a) is based on a conventional CMOS platform, incorporating an 8 nm HZO ferroelectric gate dielectric and a tungsten (W) gate electrode. Figure 2b gives the XRD analysis plot of the post-annealing HZO. A TEM image of the gate stack is shown in Figure 2c. The source and drain regions are heavily doped (n++) to ensure low-resistance electrical contacts. Two types of device geometries were implemented in this study: a planar FET and a ring FET (See the image of the fabricated devices in Figure S1). The ring FET design offers enhanced electrostatic control with significantly higher on/off ratio.

Devices were fabricated with channel lengths of 2, 3, and 5 μm. The planar FETs feature a 50 μm channel width, while the ring FETs achieve an effective channel width of approximately 400 μm. This geometric variation enables us to evaluate the FeFET's memory and computational behavior across multiple structural configurations. A key architectural feature is the deliberate gate-source/drain (G-S/D) overlap, which is critical for the device's dynamic (short-term) memory response and for the computational capabilities demonstrated here. While the overlap was intentionally introduced to enable NQS-driven STM, the overlap length and S/D doping are not yet optimized and will be explored systematically in future device generations.

Long-term memory (LTM) in FeFETs is controlled by the remanent polarization of the HZO ferroelectric layer. When specific gate voltages are applied, polarization states within HZO are set and remain stable after removing the external gate voltage, yielding robust non-volatile memory. As illustrated in Figure 2d, the drain-current versus gate-voltage (I_d–V_{gs}) characteristics exhibit pronounced anticlockwise hysteresis loops (e.g., −4 V to +4 V and −3 V to +3 V windows), reflecting the existence of distinct and multi-level switchable polarization states, thus justifying the LTM effect. Because the HZO thickness is 8 nm, we explicitly checked gate leakage during DC transfer measurements. At 50 mV drain voltage, the measured gate current I_g remains in the sub-nA per micrometer channel width range over most of the V_g sweep window and increases above 1 nA/μm only near the highest applied positive bias (+4 V), indicating that gate leakage is much smaller than the on-state drain current and does not dominate the reported device operation.

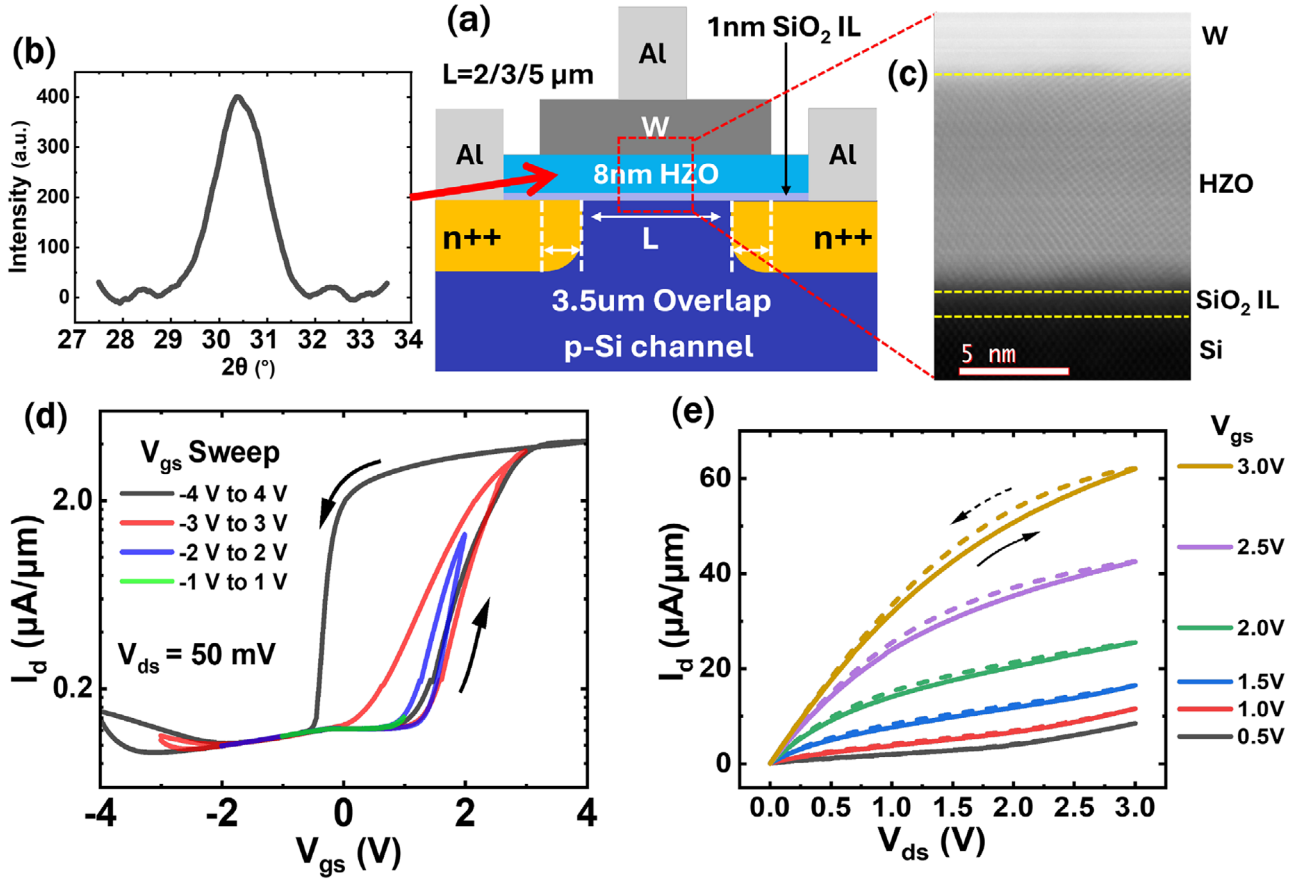

FIGURE 2 | (a) Cross-section of the planar FeFET device showing an overlap region of 3.5 μm × 50 μm between gate metal (W) and source/drain (n++) implant region forming a capacitor structure. (b) X-ray diffraction (XRD) spectrum of the 8 nm hafnium-zirconium-oxide (HZO) film after annealing. The prominent diffraction peak centered at approximately 30.4° corresponds to the (111) reflection of the ferroelectric orthorhombic phase. This confirms the successful crystallization of the HZO layer, which is essential for the ferroelectric behavior of the gate dielectric in the device. (c) A TEM image of the FeFET gate stack. The TEM clearly shows the crystalline nature of the HZO, with the repeating dot-like pattern arising from its polycrystalline structure. (d) I_d—V_{gs} plot for various gate voltage sweeping ranges showing anticlockwise hysteresis modulating the threshold voltage based on gate polarization. Drain voltage is kept at 50 mV. (e) I_d—V_{ds} plot for various gate voltages showing anticlockwise hysteresis that is only prominent for faster sweep speed (about 2 V/s) and higher gate voltage. See Figure S10 for sweep speed comparison result.

The overlap capacitance arises from the physical intersection of the gate electrode with the source and drain extensions beyond the channel, forming capacitors at both G–S and G–D interfaces. Its magnitude is governed by the overlap area, dielectric thickness, and the dielectric constant of interfacial layers. Note that in fast excitation regimes, the overlap is not merely parasitic; it supplies the input coupling that excites a non-quasi-static (NQS) channel charge response. In an NQS picture, the inversion layer behaves as a distributed, bias-dependent RC line; a rapid drain-voltage perturbation couples through the G–D overlap and perturbs the channel charge near the drain edge, which then relaxes with a finite time constant τ.

Accordingly, when voltage pulses (or sufficiently fast V_{ds} sweeps) are applied, the device exhibits transient, history-dependent changes in I_d that appear as apparent hysteresis in the drain-current versus drain-voltage (I_d—V_{ds}) characteristics (see Figure 2e). We attribute this behavior to NQS channel-charge relaxation rather than to quasi-static conduction changes: the short-term memory (STM) is volatile and decays as the NQS charge perturbation relaxes, yielding paired-pulse facilitation/depression (PPF/PPD) on microsecond timescales. Consistent with compact NQS theory [32], the dominant time constant scales approximately as

$$\tau \sim R_{\mathrm{ch}}(V_{gs}, V_{ds}, L, W)\, C_{\mathrm{ov}}(\text{overlap geometry}, t_{\mathrm{ox}})$$

so that shorter L and operation deeper in strong inversion (lower R_{ch}) reduce τ, whereas operation near/under threshold (higher effective R_{ch}) increases τ. The overlap geometry and dielectric stack set C_{ov}, providing a separate design knob for tuning τ. This engineered overlap—a designed material functionality rather than a parasitic effect—enables the microsecond-scale dynamics essential for real-time edge computing.

The interplay between overlap-driven NQS STM and ferroelectric LTM enables the coexistence of volatile and non-volatile memory within a single FeFET device. While we do not aim to reproduce specific biochemical pathways of metaplasticity, we adopt an engineering-operational definition, a persistent internal state variable that reconfigures subsequent plastic dynamics, and we experimentally verify this hallmark by showing

reversible, deterministic LTM-programmed transitions between facilitation- and depression-type STM responses. This dual-memory modality enhances the computational richness of the reservoir in three ways: (i) the overlap-driven NQS dynamics introduce nonlinear, history-dependent current responses, expanding the diversity of internal states; (ii) STM encodes and processes recent input sequences essential for temporal tasks; and (iii) LTM (via polarization) configures the operating regime (strong vs. weak inversion) and thereby programmatically tunes the sign and strength of STM (PPF vs. PPD) under a fixed V_{gs}. Together, fast, volatile STM and slow, persistent LTM project input signals into a high-dimensional feature space suitable for reservoir computing and low-power edge inference.

3 | Ferroelectric-Modulated Transient Charge Dynamics: Inter-dependence Between Short-Term and Long-Term Memory

To demonstrate the interaction between STM and LTM in our silicon-based FeFETs, we performed systematic electrical measurements using a series of drain voltage pulse sequences following controlled gate polarization. This approach reveals how the device's remanent polarization state—set by the gate voltage—modulates the transient current response, enabling tunable memory dynamics essential for reservoir computing.

After polarizing the FeFET gate to a specific voltage (either +4 V, +3 V, or −4 V), we applied a sequence of ten 10 μs rectangular drain voltage pulses (1 V amplitude), each separated by a 10 μs inter-pulse interval. The resulting drain current (I_d) was recorded in real time to capture the device's dynamic response to each pulse. This protocol was repeated on planar FETs with varying channel lengths, and the results are consistent, enabling a comprehensive analysis of memory effects with respect to the polarization states.

The drain current response shows that the device's gate polarization state directly controls the strength and nature of its short-term memory. When the device is polarized at +4 or +3 V, the drain current response to the pulse sequence exhibits a paired pulse facilitation (PPF) effect. Here, successive pulses generate incrementally larger current peaks, indicating that the device retains and amplifies the memory of recent inputs. This is observed in Figure 3a, where the current increases with each pulse during the sequence. In contrast, polarizing the gate at −4 V induces a paired pulse depression (PPD) effect. The drain current response diminishes with each subsequent pulse, reflecting a rapid decay of the STM trace. Figure 3b highlights this behavior, where the current response decreases over the pulse train. A summarized Figure 3c shows the calculated percentage increase in drain current against different polarization states. To further quantify and analyze this interplay between gate polarization and short-term memory, we developed a dedicated "read—write—read" testing protocol. The results, presented in Figures S2– S4, clearly demonstrate that the percentage change in drain current between read pulses systematically varies with the gate polarization state, confirming that long-term polarization modulates short-term memory strength and dynamics across all device geometries.

The physical origin of these tunable STM effects—PPF and PPD—is understood to arise from two interconnected mechanisms: a dominant non-quasi–static (NQS) channel-charge relaxation excited via the gate–source/drain overlap capacitance, and a secondary, minor contribution from charge trapping/detrapping. The device's long-term memory state, set by the ferroelectric gate polarization, directly modulates this NQS transient response.

The primary mechanism is the overlap-induced capacitive coupling that perturbs channel charge near the drain edge and relaxes with a finite NQS time constant. When the gate is positively polarized (+4 or +3 V at our fixed V_G), the local threshold is lowered, and the device operates in strong inversion; a drain pulse leaves a small residual pre-charge at the channel edge if the inter-pulse interval is shorter than the NQS time constant, leading to higher initial conductance on the next pulse (PPF). Conversely, with negative polarization (−4 V), the operating point shifts toward weak inversion/near-threshold; the relaxation is slower and diffusion-dominated, so a residual deficit persists at the second pulse, reducing the current (PPD). The proof that operating in weak or strong inversion dictates PPD or PPF is shown in Figure S5. In this case, the positive polarization demonstrates PPD when operated in weak inversion during drain pulsing, and negative polarization demonstrates PPF when operated in strong inversion. This signifies that the operating regime (strong vs. weak inversion) determines the type of short-term memory, which can be modulated using gate polarization during training, providing reconfigurability to the RC network for different training datasets.

A secondary, less dominant mechanism involves the trapping and detrapping of electrons into and out of defect states, either within the HZO gate oxide or at its interfaces. This process could explain a small portion of the hysteresis observed in slow-sweep DC I_d—V_{ds} characteristics (Figure 2e). However, our simulation results suggest that the trapping/detrapping is quite small for our pulsed experiments [33, 34]. Additionally, given that the device operates in a relatively low-voltage regime that limits carrier injection into the oxide, the contribution from trapping is considered minimal. Therefore, the fast pulsed response is dominated by the overlap-engineered NQS mechanism programmed by the non-volatile ferroelectric state.

The existence and dynamic nature of short-term memory in the silicon-based FeFET devices are demonstrated in Figure 3d. This plot presents representative paired-pulse facilitation (PPF) data from three planar FeFET devices with channel lengths of 2, 3, and 5 μ m. In each case, two consecutive square voltage pulses are delivered to the device with varying inter-pulse intervals (IPI). The figure reveals a typical drain current response: upon applying two closely timed voltage pulses, the second pulse generates a noticeably higher current than the first, indicating the presence of short-term facilitation. The percentage PPF is calculated as:

$$\mathrm{PPF} = \frac{(B - A)}{A} \times 100\% \tag{1}$$

A key observation from Figure 3d is that the PPF effect diminishes gradually as the interval between pulses increases. This decay in PPF with increasing IPI is consistently observed across all three device geometries. Such an exponential-like decay with

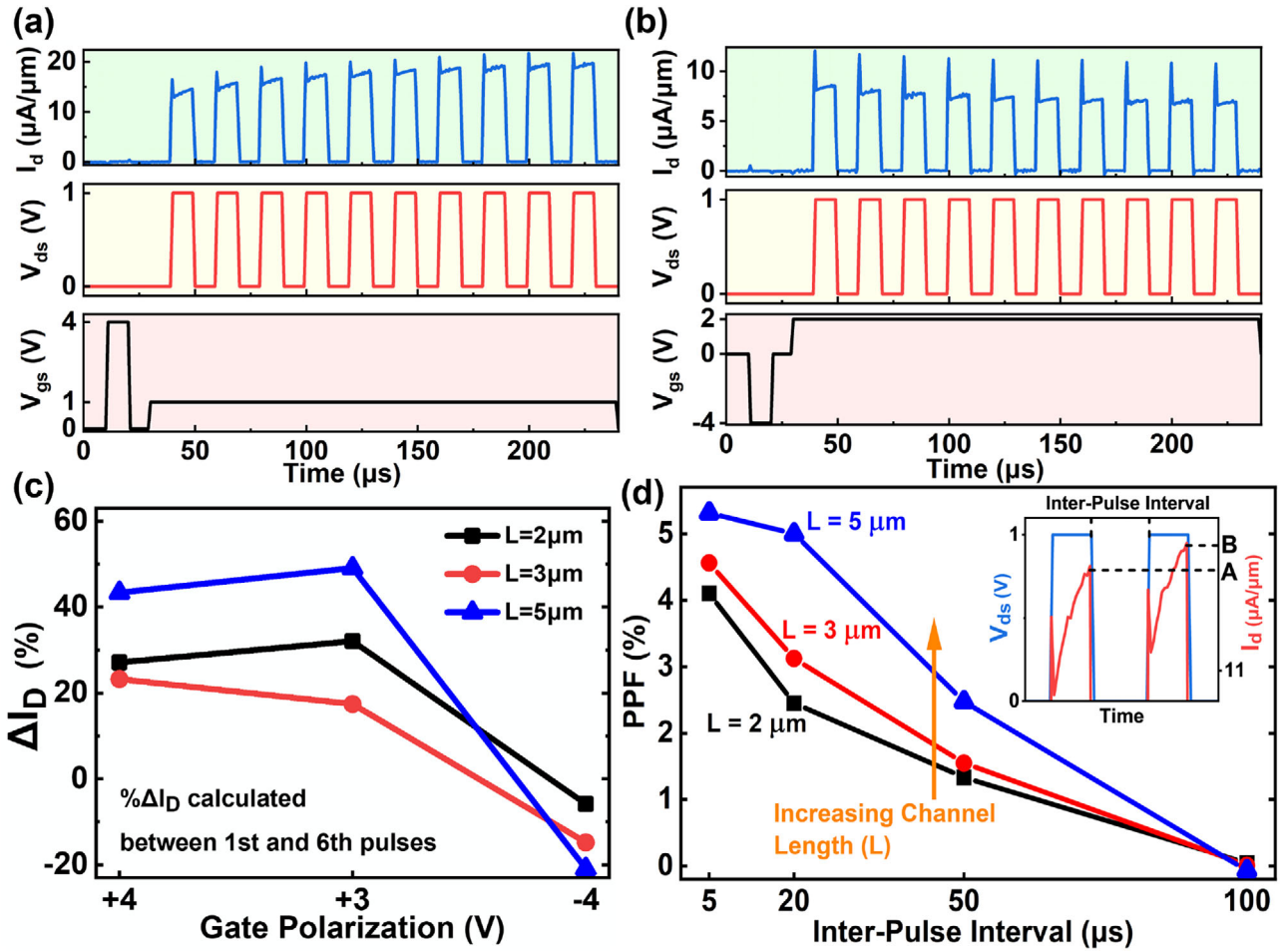

FIGURE 3 | (a) Current response to ten 10 μs rectangular voltage pulses of 1 V administered in close temporal proximity, separated by an inter-pulse interval of 10 μs. The pulses generate an amplified current response, illustrating the PPF of the FeFETs. Drain voltage pulse sequence showing PPF effect when the FeFET is polarized at +4 V on the gate side. (b) Drain voltage pulse sequence showing PPD effect when the FeFET is polarized at -4 V on the gate side. (c) Calculated percentage increase of drain current between the first and sixth pulses during the ten-pulse sequence. The percentage shows a positive increase in drain current for +4 and +3 V gate polarization, while a decrease in drain current occurs at –4 V gate polarization. (d) To demonstrate the unique temporal dynamics of each FeFET device, we display representative experimental data for PPF for three planar devices with various channel lengths of 2, 3, and 5 μm. The inset displays a typical drain current response to two square voltage pulses administered in close temporal proximity, separated by an inter-pulse interval (IPI). The pulses generate an amplified current response, illustrating the PPF of the FeFET. Quantitatively, PPF is defined in terms of the peak drain current of the first pulse (A) and second pulse (B) as follows: PPF = (B - A)/A × 100%. The PPF values for the three planar devices of our FeFETs are then plotted against IPI. The voltage pulses employed were 15 μs in duration and rectangular waveforms with an amplitude of 1 V.

IPI is characteristic of NQS relaxation governed by the channel time constant τ (see Supporting Note). The clear reduction in facilitation at longer inter-pulse intervals directly evidences the volatile, time-dependent nature of short-term memory encoded in the device response. These temporal dynamics are crucial for enabling temporal processing in reservoir computing, as they provide fading memory essential for sequence-dependent computation.

These results confirm that our FeFETs– long-term memory, encoded by the stable remanent polarization of the HZO layer, can control the STM strength and sign. By switching the gate polarization, we can reversibly toggle between facilitation and depression, effectively tuning the device's temporal response characteristics. These findings prove that our FeFETs possess robust LTM (via ferroelectric polarization) and highly tunable STM (via overlap-driven NQS dynamics coupled with the polarization controlled operating regime). The ability to modulate STM strength and sign through long-term polarization states significantly increases the heterogeneity of accessible reservoir states. This dual-memory mechanism enhances the computational richness and dynamic range of the reservoir, enabling more effective temporal information processing—a key requirement for advanced neuromorphic and reservoir computing applications. The CMOS-compatible materials foundation—exploiting established ferroelectric HZO and silicon processing-provides an immediately manufacturable pathway to scalable neuromorphic systems. In contrast, many reported memristive reservoir systems rely on exotic materials that are incompatible with CMOS processing, hindering practical deployment.

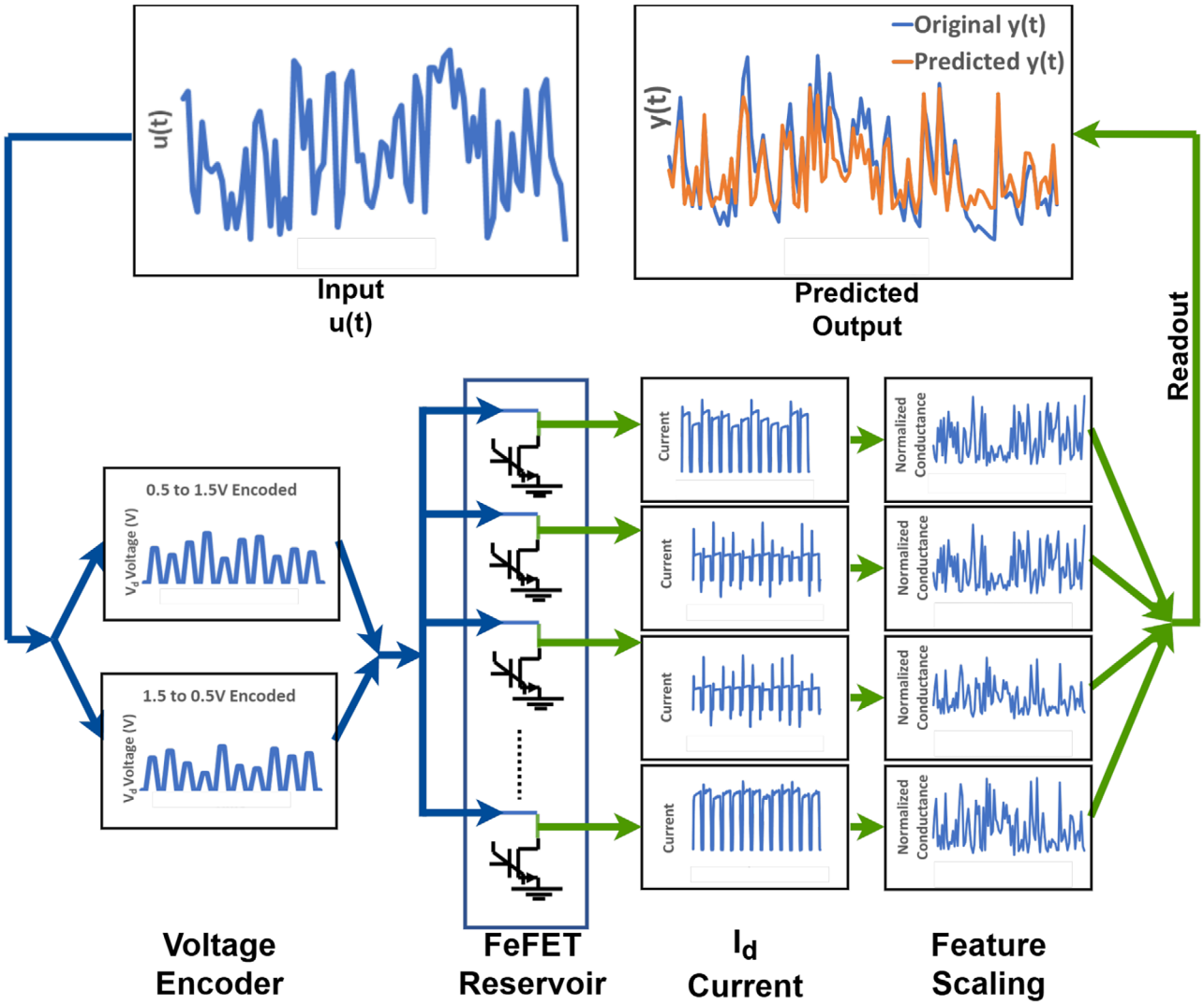

FIGURE 4 | Complete workflow solving the second-order dynamic nonlinear transfer function with our Si-FeFET devices.

4 | FeFET-Based Physical Reservoir Computing Implementation

To implement physical reservoir computing with our silicon-based FeFET devices, we used the FeFET reservoir to predict a second-order dynamic nonlinear transfer function. The transfer function is described as follows:

$$y(t) = 0.4y(t-1) + 0.4y(t-1)y(t-2) + 0.6u^3(t) + 0.1 \quad (2)$$

The output signal, $y(t)$, depends on the present input, $u(t)$, as well as the previous two inputs, $y(t-1)$ and $y(t-2)$ (that is, a time delay of two time steps), as shown in Equation (2). In this study, we trained the FeFET-based RC system to map a random input onto a higher-dimensional space. This mapping enables the generation of an accurate second-order dynamic nonlinear transfer function output from the input after training, without prior knowledge of the underlying mathematical relationship between input and output. The process flow of how the whole system works is demonstrated in Figure 4.

We encode both training and testing sequences as a series of voltage input pulses delivered to the drain terminal. Specifically, the input data, ranging from 0 to 0.5, is mapped onto a sequence of rectangular drain voltage pulses linearly ranging from 0.5 to 1.5 V. Each pulse has a period of 20 μs and a 75% duty cycle. As visualized in the attached experimental setup (Figure 5a), each measurement cycle begins with a designed pre-writing sequence to account for device initialization effects. This pre-writing (Figure 5b) consists of gate ferroelectric polarization set pulses followed by a drain 'read, write, read' voltage sequence. This protocol stabilizes the ferroelectric polarization state and ensures consistent device behavior across repeated measurements. Figure S4 provides more details on the pre-writing scheme.

Once initialized, the device is fed by a stream of 500 encoded drain voltage pulses, as shown partially in Figure 5c. Each pulse in this sequence represents a time point in the input signal, allowing the reservoir to process dynamic, time-dependent patterns. This pulse train is configured to maximize the device's nonlinearity and memory properties, which are critical for effective reservoir computing. During every voltage pulse, the corresponding drain

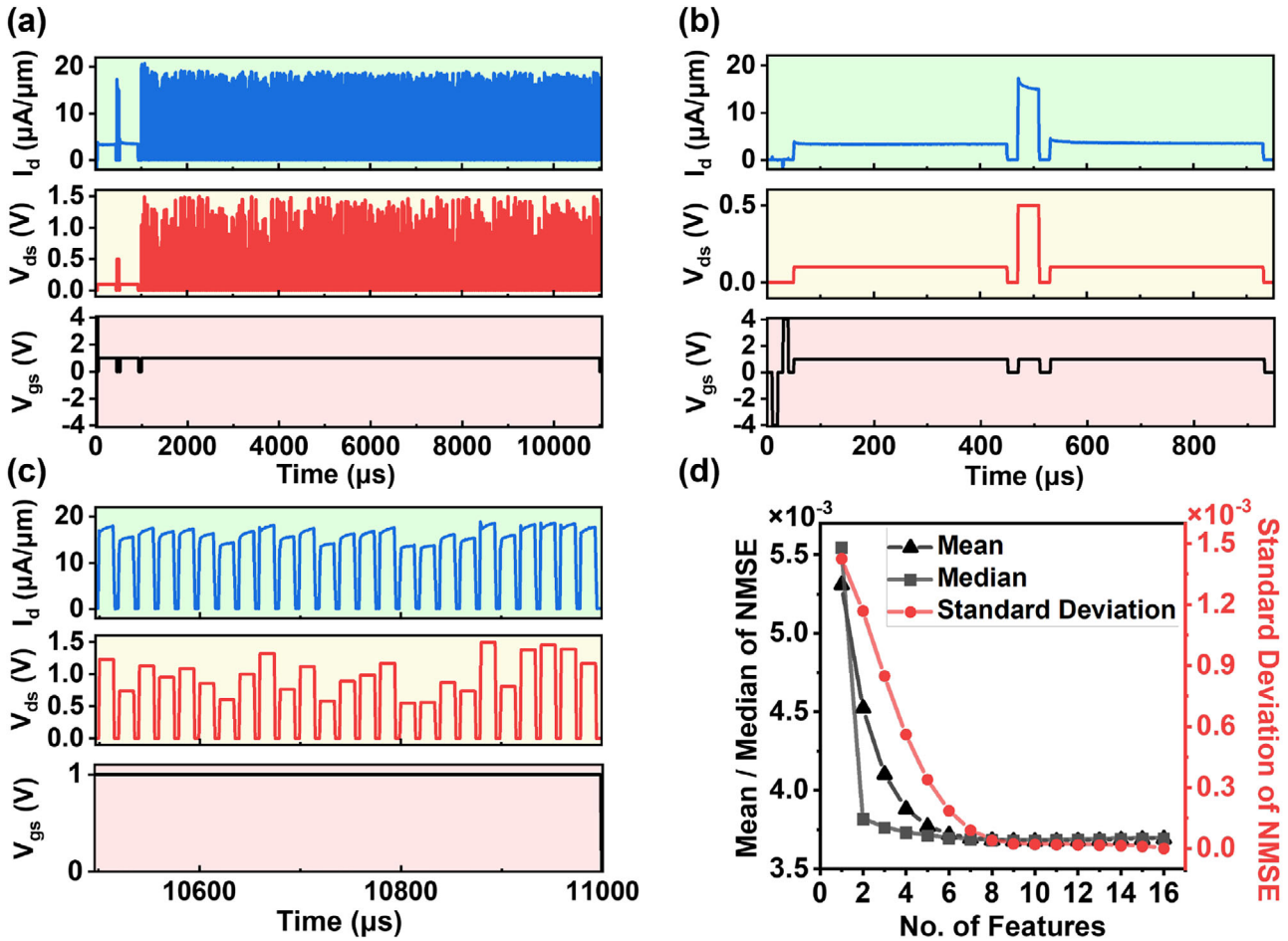

FIGURE 5 | (a) Complete timeline of the 500-point training set measurements. (b) The pre-writing scheme of the 500-point training set measurements. The pre-writing scheme is utilized to account for the initialization of the device. (c) Section of the encoded voltage pulses and drain current response toward the end of the 500-point training set measurements. (d) Due to device heterogeneity, each FeFET feature in the reservoir exhibits different polarization states, and different feature combinations result in varying NMSE values. The plot illustrates how the statistical measures of NMSE change with the number of features, showing a clear decrease in NMSE as the number of features increases.

current response is recorded and utilized as the instantaneous output of the physical reservoir. This time series of drain current values reflects the evolving internal state of the FeFET reservoir in response to the input sequence, capturing both short-term memory (capacitive) and long-term memory (ferroelectric polarization) contributions. These experimentally acquired current responses for each pulse form the feature set that defines the reservoir's output layer.

After pulse stimulation, this data is passed to the reservoir computing readout layer, the only trained portion of the system, where machine learning algorithms extract and map relevant patterns for the target computational task. The readout layer consists of the same number of neurons as the reservoir, with one extra for bias. Before going through the readout layer, the reservoir output current is first converted into conductance and then normalized with the first conductance sample of the set. The normalized conductance is then scaled between 0 and 1, depending on the minimum and maximum normalized conductance for each feature. For the training and testing sets, both conductance normalization and feature scaling are performed with reference to the training set. In the readout layer, Ridge Regression [35] is used for training.

5 | Results and Discussion

When evaluating performance in comparison with other published results, several key metrics highlight the advantages of our proposed FeFET-based reservoir computing systems. In terms of prediction accuracy, our FeFET approach demonstrates a competitive normalized mean square error (NMSE) of 3.31×10^{-3} for training and 3.69×10^{-3} for testing tasks. Most importantly, our system achieves this performance having only 16 reservoir states, >3× reduction from Hossain et al. and >5× reduction from Du et al., indicating a highly compact and efficient design. Also, this reduction in the number of reservoir states results in a similar reduction in the size of the readout layer, providing significant scalability to our system. Furthermore, our FeFET-based reservoir consumes only 1.50×10^{-7} J of energy while offering three orders of magnitude ($\approx 10^3\times$) faster response time of 20 μs compared to Hossain et al. and Du et al., which is essential

TABLE 1 | Comparison of key metrics between some of the notable reported results.

Work	Train [NMSE]	Test [NMSE]	Reservoir states	Reservoir energy (J)	Response time
This work	3.31×10^{-3}	3.69×10^{-3}	16	1.5×10^{-7}	20 μs
Hossain et al. [29] Work	5.75×10^{-4}	7.81×10^{-4}	50	2.72×10^{-8}	600 ms
Du et al. [30] Work	3.61×10^{-3}	3.43×10^{-3}	90	3.34×10^{-4}	20 ms

TABLE 2 | Comparison of key metrics between some of the notable reported results (cont…).

Work	CMOS compatibility	Memory type	Readout layer size	Device level heterogeneity	scalability
This work	Yes	Both Long & Short-Term	17	Yes[a]	Yes
Hossain et al. [29] Work	No	Short-Term	51	No	No
Du et al. [30] Work	Yes	Short-Term	91	No	Yes

aGate Polarization, Gate Voltage, Device Geometry.

for CMOS integration and large-scale deployment. Additionally, the FeFET-based system benefits from device-level heterogeneity through tunable parameters such as gate polarization, gate voltage, and device geometry; an important feature missing in prior works that enriches its computational dynamics without compromising scalability. Figure 5e shows the statistical change in NMSE with different combinations of FeFET heterogeneity. Our results suggest that at least 8 features are sufficient to reach close to the lowest error. Further increases in the number of features improve the stability of the result by reducing the standard deviation of the errors for different possible feature combinations. Figure S6 provides more details on FeFET heterogeneity in the context of error stability. Collectively, these attributes establish FeFET-based reservoir computing as a scalable, fast, energy-efficient, and CMOS-compatible platform capable of supporting complex temporal processing with a compact footprint. Tables 1 and 2 summarize the comparison with recently published work in the literature. Beyond performance metrics, the CMOS-compatible material platform enables heterogeneously integrated CMOS+X type device arrays — a scalability pathway absent in material systems requiring high-temperature deposition or unconventional substrates.

Takagi et al. [2] reported similar work that addresses the second-order nonlinear dynamic task using a FeFET-based reservoir computing architecture. However, the formulation of the problem reported in that paper diverges significantly from our approach. Specifically, instead of sampling the input values in Equation (2) from a continuous uniform distribution over the interval $[0, 0.5]$, the input is restricted to discrete values, namely $\{0, 0.5\}$. Due to this fundamental difference in input encoding and task formulation, the results published in Takagi et al. are not directly comparable and were therefore excluded from the quantitative evaluation presented in this work.

To evaluate cycling endurance, we compared DC and pulsed characteristics before and after 10 000 read/write cycles. The device shows a reduction in absolute I_d while preserving ferroelectric hysteresis (LTM) and the NQS-driven fading transient response (STM) (Figures S11 and S12). We attribute the post-cycling current reduction to cycle-induced changes that primarily affect the effective channel/contact resistance and/or interfacial scattering/trap population, which can lower the absolute drive current without eliminating the ferroelectric switching behavior or the overlap/NQS-driven transient response mechanism. Because reservoir computing relies on the reproducible nonlinear transient dynamics rather than absolute current magnitude, the preservation of hysteresis and STM fading indicates that the core dual-memory functionality remains operational after endurance stress.

Although functional memory dynamics are preserved after 10 000 cycles, the absolute drain current decreases. This degradation may primarily arise from the generation of interface traps and oxide charge trapping during repeated voltage stress, which increases Coulomb scattering and reduces channel carrier mobility [36]. However, this baseline current drift does not impair our system's long-term reliability. Because the reservoir states rely on relative transient dynamics (PPF/PPD) and the outputs are normalized before passing to the readout layer, the computation is highly resilient to absolute current reductions. Future optimizations, such as interfacial layer engineering, can further suppress trap generation and mitigate this degradation.

Material-level scalability, determined by whether ferroelectric-semiconductor coupling maintains functionality at advanced nodes, is critical for practical deployment. To investigate the scalability of our device, we modeled a silicon (Si) FeFET using the Ginestra EDA simulation software with the same device structure but scaled to 30 nm gate length. The simulated electrical characteristics show good agreement with the experimental data, with minor deviations arising from geometrical differences between the fabricated and modeled devices; these differences were introduced to improve simulation convergence and speed. The simulated device structure is shown in Figure S7 and its parameters are detailed in the Methods section.

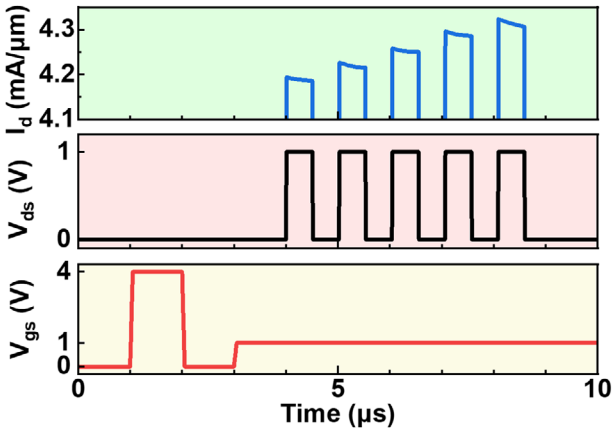

FIGURE 6 | Simulated drain current (I_d) as a function of time is shown with respect to the corresponding drain voltage (V_{ds}) and gate voltage (V_{gs}) pulsing schemes. An initial gate voltage pulse from 0 V to +4 V and back to 0 V precedes the final set at +1 V. This sets the positive polarization at the gate. Five drain pulses are encoded between 0 V and +1 V with a duty cycle of 50% and a pulse period of 1 μs. The resulting I_d shows clear PPF.

The simulated drain current (I_d) in response to the pulsed drain voltage (V_{ds}) is presented in Figure 6. Before the drain pulse, positive polarization at the gate is set by applying a gate voltage (V_{gs}) of +4 V. During the simulations, the source and bulk electrodes were grounded, V_{gs} was set to +1 V to turn the transistor on, and the temperature was held at 300 K. The results clearly exhibit paired-pulse facilitation (PPF). In our interpretation, these transients reflect a non-quasi-static (NQS) channel-charge response: drain-voltage steps couple through the gate–source/drain overlap capacitance, perturbing the inversion charge, which then relaxes with a finite time constant rather than instantaneously as assumed in a quasi-static picture. In order to demonstrate the decaying nature of this memory, simulations were run with a longer inter-pulse interval (IPI) between subsequent V_{ds} pulses. As seen in Figure S8, after the first pulse, the current amplitude does not change with successive pulses, confirming the presence of STM in the modeled device. This fading behavior with IPI is consistent with an NQS relaxation process of the channel charge. Additionally, there is an initialization effect (increase in I_d) only after the first pulse to set the baseline; an effect that was also observed in the experimental device (Figure S4). This strong agreement between simulation and experimental results validates the employed physics models, affirming their suitability for further scalability studies.

Simulation result analysis identifies the overlap capacitance between the gate and the source/drain regions as the primary driver of this NQS STM. Mechanistically, the overlap capacitance provides the input coupling while the inversion channel behaves as a distributed, bias-dependent RC; together they generate a non-quasi-static (NQS) "charge-deficit" relaxation with a characteristic time constant governed by the channel input resistance and oxide capacitance, [32, 37–39] yielding microsecond-scale short-term memory. Under our fixed gate-bias condition, the ferroelectric polarization selects the local operating regime at the drain edge: positive polarization lowers the effective threshold and places the device in strong inversion (shorter NQS time constant), whereas negative polarization shifts the operating point toward weak inversion (longer NQS time constant). Consistent with this NQS bias dependence, [37, 38] we observe paired-pulse facilitation in the strong-inversion case and paired-pulse depression in the weak-inversion case. This overlap-engineered NQS mechanism is highly scalable, as reducing the transistor area allows us to operate at a higher frequency regime, which may not always be possible for memristor based short-term memory. As detailed in Figure S9, after an initial +4 V gate pulse sets the device's polarization, the ferroelectric state remains largely stable during subsequent drain pulses, making its contribution to PPF/PPD minimal. Similarly, the simulation shows negligible changes in the trapped charge-carrier density, effectively ruling it out as a significant contributor to STM. With these other effects discounted, the analysis confirms that the overlap-capacitance-driven NQS channel response is the dominant mechanism behind the observed STM. Therefore, engineering the gate overlap remains the source of the STM as the device is scaled down, proving the scalability of our device structure.

It is important to note that scaling down the transistor also scales the overlap area, which reduces the overlap capacitance. In order to keep the capacitance the same, the thickness of the dielectric must also scale. Equivalently, from an NQS standpoint the characteristic time constant scales approximately with the product of channel resistance and overlap capacitance (schematically, $\tau \sim R_{\mathrm{ch}} C_{\mathrm{ov}}$), [37] so τ can be maintained or shortened by co-optimizing geometry/materials (to set C_{ov}) and bias/length (which set R_{ch}). Additionally, the pulse width can be scaled, thereby proportionally scaling the system-level response time, provided that the edge (or pulse) duration remains on the order of, or shorter than, τ, so that the NQS fading-memory behavior is preserved.

From a device-engineering perspective, the NQS relaxation time provides several practical design knobs. First, reducing channel resistance through shorter channel length or operating deeper in strong inversion shortens the relaxation time and enables faster STM. Second, increasing the gate-source/drain overlap capacitance through a larger overlap area or a high-k dielectric stack increases the relaxation time and strengthens fading memory behavior. Third, operating near the threshold increases the effective channel resistance and therefore extends the memory window, while strong inversion favors shorter time constant operation. Finally, under device scaling, the reduction in overlap area must be co-optimized with dielectric thickness and pulse width to preserve the desired NQS regime.

6 | Conclusion

In this work, we have successfully demonstrated a CMOS-compatible ferroelectric-semiconductor platform featuring an engineered gate-source/drain (G-S/D) overlap capacitance that can be programmed to exhibit non—quasi—static (NQS) short-term memory (STM) tunable via gate polarization switching. With this engineered material architecture, we have demonstrated a highly efficient and tunable physical reservoir computing platform. By systematically characterizing and leveraging the device's unique dual-memory modalities, we have demonstrated its capacity to solve a second-order nonlinear dynamic

task with high accuracy, achieving a competitive test NMSE of 3.69×10^{-3} using only 16 reservoir states. This performance highlights the computational capability of this highly compact and efficient design.

The core of our approach lies in the synergistic interplay between ferroelectric long-term memory (LTM) and overlap-capacitance-engineered NQS short-term memory (STM). We established that LTM, governed by the remanent polarization of the HZO gate dielectric, acts as a non-volatile control for the device's transient dynamics. By setting the gate polarization, we could deterministically switch the device's response between paired-pulse facilitation (PPF) and paired-pulse depression (PPD), effectively tuning the STM strength. This tunable STM, which originates from the engineered overlap capacitance between the gate and source/drain regions, provides the rich, nonlinear, and fading memory essential for reservoir computing.

Beyond demonstrating reservoir computing performance, this work establishes material-engineering principles for neuromorphic hardware: (i) dual-timescale functionality through orthogonal mechanisms (ferroelectric switching + capacitive dynamics), (ii) material-parameter tunability (overlap capacitance, polarization state, operating regime) enabling application-specific optimization, and (iii) CMOS-compatible material choices (HZO, silicon, standard processing) ensuring manufacturability. These principles suggest broader applicability: the ferroelectric-modulated charge dynamics demonstrated here may make it possible to design nanosecond to millisecond tunable volatile memory using oxide semiconductors. This can ultimately enable adaptive analog circuits, reconfigurable sensors, and materials-based signal processing beyond neuromorphic computing.

The CMOS-compatible fabrication, low energy consumption (1.5×10^{-7} J), fast response time (20 μ s), and potential scalability position this ferroelectric-semiconductor platform as an excellent candidate for next-generation edge computing applications. Although our current devices have a fixed overlap geometry, our simulation results indicate that this structural parameter is a critical design knob for optimizing NQS-driven STM. This suggests a clear path for future work in co-designing device physics and reservoir performance by further exploring the gate-S/D overlap and enabling multi-modal transport using oxide semiconductors. Ultimately, this research establishes that incorporating both LTM and overlap-driven NQS STM within a single CMOS-compatible material system provides a powerful, scalable, and energy-efficient pathway for realizing advanced neuromorphic systems for complex temporal data processing.

7 | Methods

7.1 | Device fabrication

The fabrication of the silicon-based FeFET begins with the etching of alignment marks on a silicon wafer. A screen oxide is thermally grown at 950°C for 30 min to prepare for source and drain implantation. Ion implantation follows, involving photolithographic patterning, n++ implantation, and photoresist stripping. The screen oxide is then removed using a buffered oxide etch (BOE). Subsequently, the implanted dopants in the source and drain regions are activated through annealing at 900 °C for 20 min.

A standard CMOS RCA clean is performed before gate oxide formation. During the SC1 step, approximately 1 nm interfacial SiO_2 layer is chemically grown on the silicon surface. The gate stack, consisting of an 8 nm ferroelectric $Hf_{0.5}Zr_{0.5}O_2$ (HZO) layer, is deposited using thermal atomic layer deposition (ALD) at 250°C. A 40 nm tungsten (W) gate electrode is then deposited, followed by rapid thermal annealing at 500°C for 30 s to crystallize the ferroelectric film. The tungsten is patterned and etched to form the gate electrode, which is extended over the doped source and drain regions to create the desired overlap of 3.5 μm.

Next, the gate oxide is selectively etched to remove the HZO from the source and drain areas. A 400 nm SiO_2 interlayer dielectric is deposited and patterned to define contact vias using dry etching. Finally, aluminum pads (containing 1% silicon) are deposited via a lift-off process, and the device undergoes a forming gas anneal at 400°C for 30 min to improve contact integrity and overall reliability.

7.2 | Electrical Measurements

Device electrical measurements were performed at room temperature under standard laboratory lighting. A Micromanipulator P200L probe station was used to establish contact with the devices. The station was connected to a Keysight B1500A Semiconductor Device Analyzer mainframe, which served as the primary instrument for characterization. For dynamic characterization, a Keysight B1530A Waveform Generator/Fast Measurement Unit (WGFMU) module was utilized. This specialized module generated the necessary voltage pulses applied to the gate and drain terminals for the pulsed current–voltage (I–V) measurements while concurrently measuring the resulting drain current.

7.3 | Computational Modeling

Device simulations were performed using the Ginestra software package, a trap-centric modeling platform that self-consistently describes all the relevant physical mechanisms occurring in semiconductors, dielectric layers, and novel materials. The simulated device was based on a conventional n-type MOSFET structure with a physical gate length of 30 nm and an effective channel length of approximately 6 nm (calculated by excluding the overlap regions).

7.3.1 | Device Structure and Parameters:

The simulated structure consisted of a 200 nm thick silicon substrate with a width of 200 nm and a total length of 100 nm. The substrate was uniformly doped with a p-type doping concentration of $2 \times 10^{15}\ cm^{-3}$. The source and drain regions were defined with a length of 30 nm each and doped with an n-type concentration of $1 \times 10^{19}\ cm^{-3}$, which extends under the gate stack. A 5 nm spacer separated the gate from the source and drain

contacts. In order to maintain consistency with the experiments, the gate stack was composed of an 8 nm thick ferroelectric $Hf_{0.5}Zr_{0.5}O_2$ (HZO) layer on top of a 1 nm SiO_2 interfacial layer, capped with a tungsten (W) metal gate. An ideal bulk electrode at the bottom of the substrate and ideal source/drain electrodes on the top surface were used, with the bulk contact held at 0 V bias.

7.3.2 | Physics and Material Models:

The ferroelectric behavior of the HZO layer was modeled using the Landau-Khalatnikov formalism combined with the Ginzburg domain coupling term, available within Ginestra [40]. The accurate modeling of the interaction between atomic defects and FE properties was ensured by the coupled solution with Poisson's equation, while considering a uniform oxygen vacancy distribution with a concentration of $5 \times 10^{17}\ cm^{-3}$ in the HZO layer. Electron and hole transport models considering intrinsic (direct/Fowler-Nordheim tunneling, thermionic emission, drift) and defect-assisted mechanisms were implemented in the framework of the multi-phonon trap-assisted tunneling (TAT) theory.

Author Contributions

Y.W., M.S.H and R.I. conceived the idea. Y.W. did all the electrical characterization, and M.S.S. performed the reservoir computing training and tests based on the data provided by Y.W. S.S. and L.F. participated in the device fabrication, N.V., A.P. contributed to the simulation of the fabricated device using Ginestra^{TM}. A.I.K., M.S.H., and R.I. provided the overall supervision of the work. All the authors participated in manuscript writing and proofreading.

Acknowledgments

Simulation was performed using Ginestra, an Applied Materials proprietary simulation software platform. More information can be accessed at [41]. The authors acknowledge the suggestion and guidance of Valerio Lunardelli, Luca Larcher, and Gaurav Thareja from Applied Materials, Inc. throughout the project. R.I. also acknowledges the help and support of Purdue University College of Engineering.

Funding

Y.W. and M.S.S. acknowledge the support by the U.S. National Science Foundation under Standard Grant No. 2521468 and 2521469, respectively. N.V. acknowledges the support by the U.S. Nuclear Regulatory Commission Faculty Development Grant. A. P. acknowledges the FAR 2024 project of the "Enzo Ferrari" Engineering department of the University of Modena and Reggio Emilia, Italy, for the financial support.

Conflicts of Interest

The authors declare no conflicts of interest.

Consent

All authors certify that the work has been reviewed by each author and all authors are consenting for the publication if accepted.

Data Availability Statement

All data and code are fully available upon reasonable request to the corresponding authors.

References

1. G. Tanaka, T. Yamane, J. B. Héroux, et al., "Recent Advances in Physical Reservoir Computing: A Review," *Neural Networks* 115 (2019): 100–123.

2. S. Takagi, K. Toprasertpong, E. Nako, et al., "Physical Reservoir Computing Using HZO-Based FeFETs for Edge-AI Applications," in *2023 International Electron Devices Meeting (IEDM)* (2023): 1–4.

3. X. Song, A. Gaurav, P. B. Pillai, et al., "Reservoir Computing Based on a Solid Electrolyte ZnO TFT: An Attractive Platform for Flexible Edge Computing," in *2023 IEEE International Flexible Electronics Technology Conference (IFETC)* (2023): 1–3.

4. F. Nowshin, Y. Huang, M. R. Sarkar, Q. Xia, and Y. Yi, "MERRC: A Memristor-Enabled Reconfigurable Low-Power Reservoir Computing Architecture at the Edge," *IEEE Transactions on Circuits and Systems I: Regular Papers* 71, no. 1 (2024): 174–186.

5. D. W. Liedji, G. Kenné, and J. H. Talla Mbé, "Delay-Based Reservoir Computing Using Mackey–Glass Oscillator and Arduino Board for Edge Intelligence Applications," *AIP Advances* 13, no. 12 (2023): 125217.

6. R. Gaurav, T. C. Stewart, and Y. C. Yi, "Spiking Reservoir Computing for Temporal Edge Intelligence on Loihi," in *2022 IEEE/ACM 7th Symposium on Edge Computing (SEC)* (2022): 526–530.

7. M. C. Soriano, S. Ortín, L. Keuninckx, et al., "Delay-Based Reservoir Computing: Noise Effects in a Combined Analog and Digital Implementation," *IEEE Transactions on Neural Networks and Learning Systems* 26, no. 2 (2015): 388–393.

8. B. Schrauwen, M. D'Haene, D. Verstraeten, and J. V. Campenhout, "Compact Hardware Liquid State Machines on FPGA for Real-Time Speech Recognition," *Advances in Neural Networks Research: IJCNN '07* 21, no. 2 (2008): 511–523.

9. N. Soures, L. Hays, and D. Kudithipudi, "Robustness of a Memristor-Based Liquid State Machine," in *2017 International Joint Conference on Neural Networks (IJCNN)* (2017): 2414–2420.

10. G. Van der Sande, D. Brunner, and M. C. Soriano, "Advances in Photonic Reservoir Computing," *Nanophotonics* 6, no. 3 (2017): 561–576.

11. D. Brunner, B. Penkovsky, B. A. Marquez, M. Jacquot, I. Fischer, and L. Larger, "Tutorial: Photonic Neural Networks in Delay Systems," *Journal of Applied Physics* 124, no. 15 (2018): 152004.

12. H. Hauser, A. J. Ijspeert, R. M. Füchslin, R. Pfeifer, and W. Maass, "Towards a Theoretical Foundation for Morphological Computation with Compliant Bodies," *Biological Cybernetics* 105, no. 5 (2011): 355–370.

13. K. Caluwaerts, J. Despraz, A. Işçen, et al., "Design and Control of Compliant Tensegrity Robots Through Simulation and Hardware Validation," *Journal of the Royal Society Interface* 11, no. 98 (2014): 20140520.

14. M. R. Dranias, H. Ju, E. Rajaram, and A. M. J. VanDongen, "Short-Term Memory in Networks of Dissociated Cortical Neurons," *Journal of Neuroscience* 33, no. 5 (2013): 1940.

15. D. Sussillo and L. F. Abbott, "Generating Coherent Patterns of Activity from Chaotic Neural Networks," *Neuron* 63, no. 4 (2009): 544–557.

16. K. Fujii and K. Nakajima, "Harnessing Disordered-Ensemble Quantum Dynamics for Machine Learning," *Physical Review Applied* 8, no. 2 (2017): 024030.

17. J. Torrejon, M. Riou, F. A. Araujo, et al., "Neuromorphic Computing with Nanoscale Spintronic Oscillators," *Nature* 547, no. 7664 (2017): 428–431.

18. W. Maass, T. Natschläger, and H. Markram, "Fading Memory and Kernel Properties of Generic Cortical Microcircuit Models," *Decoding and Interfacing the Brain: From Neuronal Assemblies to Cyborgs* 98, no. 4 (2004): 315–330.

19. I. B. Yildiz, H. Jaeger, and S. J. Kiebel, "Re-Visiting the Echo State Property," *Neural Networks* 35 (2012): 1–9.

20. J. Liu, G. Feng, W. Li, et al., "Physical Reservoir Computing for Edge AI Applications," *The Innovation Materials* 3, no. 2 (2025): 100127.

21. Z. Ma, J. Ge, and S. Pan, "Spiking Reservoir Computing Based on Stochastic Diffusive Memristors," *Advanced Electronic Materials* 11, no. 3 (2025): 2400469.

22. W. C. Abraham and M. F. Bear, "Metaplasticity: The Plasticity of Synaptic Plasticity," *Trends in Neurosciences* 19, no. 4 (1996): 126–130.

23. W. Abraham and B. Philpot, "Metaplasticity," *Scholarpedia* 4, no. 5 (2009): 4894.

24. R. S. Zucker and W. G. Regehr, "Short-Term Synaptic Plasticity," *Annual Review of Physiology* 64, no. 1 (2002): 355–405.

25. S. Pazos, K. Zhu, M. A. Villena, et al., "Synaptic and Neural Behaviours in a Standard Silicon Transistor," *Nature* 640, no. 8057 (2025): 69–76.

26. K. Toprasertpong, E. Nako, Z. Wang, R. Nakane, M. Takenaka, and S. Takagi, "Reservoir Computing on a Silicon Platform with a Ferroelectric Field-Effect Transistor," *Communications Engineering* 1, no. 1 (2022): 21.

27. K. Liu, B. Dang, T. Zhang, et al., "Multilayer Reservoir Computing Based on Ferroelectric α-In_2Se_3 for Hierarchical Information Processing," *Advanced Materials* 34, no. 48 (2022): 2108826.

28. D. Ju, M. Noh, G. Kim, Y. Park, S. Lee, and S. Kim, "Reservoir Computing System with Diverse Input Patterns in HfAlO-Based Ferroelectric Memristor," *ACS Applied Materials & Interfaces* 16, no. 48 (2024): 66250–66261.

29. M. R. Hossain, A. S. Mohamed, N. X. Armendarez, J. S. Najem, and M. S. Hasan, "Biomembrane-Based Memcapacitive Reservoir Computing System for Energy-Efficient Temporal Data Processing," in *Advanced Intelligent Systems* (2023): 1–12.

30. C. Du, F. Cai, M. A. Zidan, W. Ma, S. H. Lee, and W. D. Lu, "Reservoir Computing Using Dynamic Memristors for Temporal Information Processing," *Nature Communications* 8, no. 1 (2017): 2204.

31. X. Fan, A. Chen, Z. Li, et al., "Metaplasticity-Enabled Graphene Quantum Dot Devices for Mitigating Catastrophic Forgetting in Artificial Neural Networks," *Advanced Materials* 37, no. 6 (2025): 2411237.

32. S. Y. Oh, D. E. Ward, and R. W. Dutton, "Transient Analysis of MOS Transistors," *IEEE Transactions on Electron Devices* ED-27, no. 8 (1980): 1571–1578.

33. T. Kim, S. Park, and S. Jeon, "Fast and Slow Transient Charging of Oxide Semiconductor Transistors," *Scientific Reports* 7, no. 1 (2017): 11850.

34. X. Ma, Y. Y. Liu, L. Zeng, et al., "Defects-Induced Charge Trapping/Detrapping and Hysteresis Phenomenon in MoS_2 Field-Effect Transistors: Mechanism Revealed by Anharmonic Marcus Charge Transfer Theory," *ACS Applied Materials & Interfaces* 14, no. 1 (2022): 2185–2193.

35. T. Hastie, "Ridge Regularization: An Essential Concept in Data Science," *Technometrics* 62 (2020): 426–433.

36. J. Duan, H. Xu, S. Zhao, et al., "Impact of Mobility Degradation on Endurance Fatigue of FeFET with $TiN/Hf_{0.5}Zr_{0.5}O_2/SiO_x/Si$ (MFIS) Gate Structure," *Journal of Applied Physics* 131, no. 13 (2022): 134102.

37. X. J. Xi, M. Dunga, J. He, et al., *BSIM4.3.0 MOSFET Model User's Manual* (Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA, 2003), Chapter 8 describes the charge-deficit NQS model (transient and AC), https://pages.hmc.edu/harris/cmosvlsi/4e/BSIM4_manual.pdf.

38. Y. Cheng, "Non-Quasi-Static (NQS) Model," in *MOSFET Modeling & BSIM3 User's Guide* (Springer, Boston, MA, 2002), 281–302.

39. J. J. Paulos and D. A. Antoniadis, "Limitations of Quasi-Static Capacitance Models for the MOS Transistor," *IEEE Electron Device Letters* EDL-4, no. 7 (1983): 221–224.

40. S. Sharma, S. Kundu, B. H. Lee, S. K. Gupta, K. Roy, and S. K. Gupta, "Reliability of Non-Volatile Memory Devices for Neuromorphic Applications: A Modeling Perspective (Invited)," in *2022 IEEE International Reliability Physics Symposium (IRPS)* (Dallas, TX, USA, 2022): 1–10.

41. Applied Materials: Ginestra Simulation Platform, accessed: 2025-09-12 (2025), https://www.appliedmaterials.com/us/en/semiconductor/solutions-and-software/software-solutions/ginestra-simulation-platform.html.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting File: aelm70422-sup-0001-SuppMat.pdf.