Supporting information

Reconfigurable Compute-In-Memory on Field-Programmable Ferroelectric Diodes

Xiwen Liu,\textsuperscript{a} John Ting,\textsuperscript{a} Yunfei He,\textsuperscript{a} Merrilyn Mercy Adzo Fiagbenu,\textsuperscript{a} Jeffrey Zheng,\textsuperscript{b} Dixiong Wang,\textsuperscript{a} Jonathan Frost,\textsuperscript{a} Pariasadat Musavigharavi,\textsuperscript{a,b} Giovanni Esteves,\textsuperscript{d} Kim Kisslinger,\textsuperscript{e} Surendra B. Anantharaman,\textsuperscript{a} Eric A. Stach,\textsuperscript{b,c} Roy H. Olsson III,\textsuperscript{a,*} Deep Jariwala\textsuperscript{a,*}

\textsuperscript{a} Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA, 19104, USA
\textsuperscript{b} Materials Science and Engineering, University of Pennsylvania, Philadelphia, PA, 19104, USA
\textsuperscript{c} Laboratory for Research on the Structure of Matter, University of Pennsylvania, Philadelphia, PA, 19104, USA
\textsuperscript{d} Microsystems Engineering, Science and Applications (MESA), Sandia National Laboratories, Albuquerque, New Mexico, 87185, USA
\textsuperscript{e} Brookhaven National Laboratory, Center for Functional Nanomaterials, Upton, NY, 11973, USA

*Corresponding author: rolsson@seas.upenn.edu, dmj@seas.upenn.edu
Experimental Methods:

Device fabrication

FeD consisted of a film stack of Al (80 nm)/Al$_{0.68}$Sc$_{0.32}$N (45 nm)/Al (30 nm) on top of a Si/Al$_{0.8}$Sc$_{0.2}$N (85 nm) substrate. To prepare this stack, we start by sputter depositing an 85-nm thick Al$_{0.8}$Sc$_{0.2}$N template on the top of a 6” Si <100> wafer. The Al$_{0.8}$Sc$_{0.2}$N was deposited using pulsed-DC reactive sputter deposition from a single-alloyed Al$_{0.8}$Sc$_{0.2}$ target using 5 kW of target power, a pressure of $7.47 \times 10^{-3}$ mbar, and a deposition temperature of 375°C. The first layer of 85 nm Al$_{0.8}$Sc$_{0.2}$N serves to orient the subsequent 80-nm thick Al layer into a {111}-orientation. This Al (80 nm thick) layer serves as the bottom electrode for the second layer of Al$_{0.68}$Sc$_{0.32}$N (45 nm thick), which is the ferroelectric layer used in this device. The 45-nm thick ferroelectric Al$_{0.68}$Sc$_{0.32}$N film was co-sputtered from separate 4-inch Al and Sc targets in an Evatec CLUSTERLINE® 200 II pulsed DC Physical Vapor Deposition System. The Al and Sc targets were operated at 1250 W and 695 W, respectively, at a chuck temperature of 350 °C with 10 sccm of Ar gas flow and 25 sccm of N$_2$ gas flow. The chamber pressure was maintained ~$1.45 \times 10^{-3}$ mbar. This sputter condition resulted in a deposition rate of 0.3 nm/sec. The highly oriented {111} Al layer promotes the growth of AlScN with its [0001] axis direction being perpendicular to the substrate, thus, yielding a highly textured FE film. Then, without breaking vacuum, a 30 nm Al layer was sputtered as the top electrode and as a capping layer to prevent the oxidation of ferroelectric Al$_{0.68}$Sc$_{0.32}$N surface.

Device characterization

Current-voltage measurements were performed in air at ambient temperature using a Keithley 4200A semiconductor characterization system. P-E hysteresis loops and PUND measurements of ferroelectric AlScN were conducted using Keithley 4200A semiconductor characterization system and a Radiant Precision Premier II testing platform. The TEM cross-sectional sample was prepared in a FEI Helios Nanolab 600 focused ion beam (FIB) system using the in-situ lift-out technique. The sample was coated with a thin carbonaceous protection layer by writing a line on the surface with a Sharpie® marker. Subsequent electron beam and ion beam deposition of Pt protection layers were used to prevent charging and heating effects during FIB milling. At the final cleaning stage, a low-energy Ga+ ion beam (5 keV) was used to reduce FIB-induced damage. TEM characterization and image acquisition were carried out on a JEOL F200 operated at 200 kV accelerating voltage. The sample was orientated to the [001] zone axis for imaging. All of the captured TEM images were collected using Digital Micrograph software.
Supplementary Note 1  Compact model for evaluation of ON/OFF ratios of ferroelectric diode

1. General method of ON/OFF ratio evaluation

To evaluate the ON/OFF ratio and capture the IV characteristics of the ferroelectric diodes (FeDs), generally, we need to resolve the electron transport in the ferroelectric diode. Within the ferroelectric, there are three main methods of electron transport: direct tunneling, Fowler-Nordheim tunneling, and thermionic emission, the band diagrams illustrating each effect are shown in Fig.1. We can use the Wentzel-Kramers-Brillouin (WKB) approximation to encompass all three methods with one formula. In the approximation, the tunneling probability is given by:

\[ T(E) = \exp\left(-\frac{2}{\hbar} \int_{x_1}^{x_2} \sqrt{2m^* \left[ (E_c + qV(x)) - E \right]} \, dx \right) \]  

(1.1)

where \( m^* \) is the effective mass of the electron, \( E_c \) is the coercive field of the ferroelectric, \( V(x) \) is the voltage across the ferroelectric, and \( E \) is the applied field. We can give the integrand for the density of states as

\[ N(E) = k_B T \ln \left( \frac{1 + \exp \left( \frac{E - E_{f,1}}{k_B T} \right)}{1 + \exp \left( \frac{E - E_{f,2}}{k_B T} \right)} \right) \]  

(1.2)

where \( K_B T \) is the Boltzmann constant multiplied by temperature, \( E_{f,1} \) and \( E_{f,2} \) are the Fermi-levels at the left and right of the ferroelectric. With these formulas, we can define the current density \( J \) as

\[ J = \frac{4\pi m^* q}{h^3} \int_{E_{min}}^{E_{max}} T(E) N(E) \, dE \]  

(1.3)

This current density multiplied by the area of the ferroelectric film gives the tunneling current through the device. While this model does effectively capture the I-V characteristics of a ferroelectric diode, it lacks efficiency.

Fig. S1. Illustrations of the three methods of electron transport through a potential barrier. (a) direct tunneling (b) Fowler-Nordheim tunneling (c) thermionic emission.
2. Voltage shifting model and its verification.

2.1. Voltage shifting model.

In the analysis, we proposed a new compact shift model to describe the I-V characteristics of the FeDs. In the FeD, we can treat the changes in the I-V curve of the diode from the low resistance state (LRS) to the high resistance state (HRS) as the I-V curve shifting from left to right by the amount of $\Delta V$ as the diode transforms from LRS to HRS. In other words, it takes more voltage to offset the voltage of $\Delta V$ at HRS to generate the same current as the current at LRS. The shifted voltage $\Delta V$ can be derived as:

$$\Delta V = E_{dp} t$$  \hspace{1cm} (2.1)

where $t$ is the thickness of the ferroelectric layer, and $E_{dp}$ is the depolarization field of the ferroelectric, and we found that we can express the depolarization field as:

$$E_{dp} = E_C \tanh\left(\frac{P_r}{\epsilon_{fe} E_C}\right)$$  \hspace{1cm} (2.2)

where $E_C$ is the coercive field of the ferroelectric layer, $P_r$ is the remanent polarization, $\epsilon_{fe}$ is the dielectric constant of the ferroelectric, and $\beta$ is the parameter related to the oxide capacitance $C_{ox}$ and the ferroelectric capacitance $C_{fe}$, as \[3\]:

$$\beta = \left[\frac{C_{ox}}{C_{fe}} + 1\right]^{-1}$$  \hspace{1cm} (2.3)
2.2. Comparison between the voltage shifting model and WKB approximation.

We verified that the voltage shifting model showed a relatively consistent result as the I-V characteristics of the ferroelectric diode modeled by the WKB approximation, shown in Fig. S2. In addition, the voltage shifting model is much more efficient since it utilizes analytical equations as opposed to numerical integration as the WKB approximation did. Furthermore, the new model's hyperbolic function helps us focus on the readout voltage region of the I-V curve since the depolarization field $E_{dp}$ would not exceed the coercive field $E_c$ in the model.

![Fig. S2. Comparisons of the IV curves using the WKB approximation (green) vs. the new voltage shifting model for a ferroelectric diode (blue) for any arbitrary ferroelectric. The purple curve represents the base curve at high-resistance state modeled with the WKB approximation.](image)

3. I-V Curves and ON/OFF ratio related to the voltage shifting model.

After the verification of this compact model, we used this model to find the overall trend of the relationships between the ON/OFF ratios of the ferroelectric diodes and parameters such as oxide capacitance, remanent polarization, and coercive field.

3.1. I-V curves and ON/OFF ratio changing with oxide capacitance.

First, we plot the I-V curve of the FeD shifting under different insulator capacitance. We pre-coded the oxide capacitance into $\beta$ using eq.2.3, and $\beta$ should vary between 0 and 1. By choosing the proper value of remanent polarization and coercive field, the simulated I-V curve of HRS
shifting is shown in Fig. S3(a). Correspondingly, the ON/OFF ratio under different $\beta$ is demonstrated in Fig. S3(b).

**Fig. S3.** The plot of HRS I-V curve shifting with the variation of the pre-coded oxide capacitance $\beta$ is shown in (a), we vary $\beta$ from 0.01 to 0.54 exponentially to get a better view of the I-V curve shifting trend. (b) is the ON/OFF ratio vs. $\beta$, and the ON/OFF ratio is the current ratio between LRS and corresponding shifted curve at the voltage of 7 V in (a).

3.2. I-V curves and ON/OFF ratio changing with remanent polarization.

Second, we vary the remanent polarization $P_r$ from 1 to 135 $\mu$C/cm$^2$ with a suitable insulator capacitance $C_{ox}$ and coercive field $E_C$, and the resulting I-V characteristics are shown in Fig. S4(a). The corresponding ON/OFF ratio under different remanent polarization is demonstrated in Fig. S4(b).

**Fig. S4.** The plot of HRS I-V curve shifting with the variation of the remanent polarization $P_r$ is shown in (a), we vary $P_r$ from 1 to 135 $\mu$C/cm$^2$ exponentially to get a better view of the I-V curve shifting trend. (b) is the ON/OFF ratio vs. $P_r$, and the ON/OFF ratio is the current.
3.3. I-V curves and ON/OFF ratio changing with coercive field.

Last, the I-V curve shifting plot for coercive fields $E_c$ linearly ranging from 0.12 to 3.12 MV/cm is shown in Fig. S5(a). The corresponding plot of the ON/OFF ratio of the FeD under different coercive fields is shown in Fig. S5(b).

![Graphs showing I-V curves and ON/OFF ratio vs. coercive field](image)

**Fig. S5.** The plot of HRS I-V curve shifting with the variation of the result of the coercive field $E_c$ is shown in (a), we vary $E_c$ from 0.12 to 3.12 MV/cm linearly to get a I-V curve shifting trend. (b) is the ON/OFF ratio vs. $E_c$, and the ON/OFF ratio is the current ratio between LRS and corresponding shifted curve at the voltage of 7V in (a).

4. Conclusion.

The I-V curves in Fig. S3-S5 show that the HRS curve shifts further when encoded insulator capacitance $\beta$, remanent polarization $P_r$, or coercive field $E_c$ increases. And from both eq.2.2 and Fig. S3-5, we can find out the hyperbolic function in eq.2.2 limits the influence on the I-V curve shifting by the remanent polarization $P_r$ and the oxide capacitance $\beta$. On the other hand, we can find that the ON/OFF ratio increases drastically with the coercive field $E_c$ increasing, and the hyperbolic function does not limit it.
**Supplementary Note 2**: Linearization on the current-voltage relationship of ferroelectric diode

While Radu Berdan *et al.* proposed a logarithmic amplifier and a transimpedance amplifier by a well-fitted model in Ref. 3, we propose a simple min-max normalization and linear mapping procedure to encode the input voltage on software. With the log $I$ - $V$ characteristics of the device shown in Fig. S6, the log $I$ has excellent linearity with the voltage applied based on the nature of the conduction in a diode device [4], and the slopes of the plots in 16 different states of the device are consistent in the fitting.

![Fig. S6. Log(I) vs V. Current I measured by applying dc voltage sweeping to 16 different states.](image1)

Therefore, to linearize the $I$ - $V$ characteristics of a ferroelectric diode, we can have:

$$I = G_i \exp(\alpha V)$$

where the $G_i$ is a parameter that is related to the $i$’th conductance of the diode.

The constant slope $\alpha$ can be estimated through a linear regression method over the 16 distinct log $I$ - $V$ characteristics of the ferroelectric diode device.

![Fig. S7. Current I vs. $\exp(\alpha V)$ for 16 different states.](image2)
From Fig. S7, we can see that a ferroelectric diode's $I - \exp(\alpha V)$ characteristics in 16 different states show superior linearity and intersection at the origin, which is the same as the $I - V$ characteristics of an Ohmic resistor. The different slopes of all 16 states are proportional to the conductance of the 16 states, which also behave in a linear manner.

Given a input $X$ from software in range $[0, 1]$ and a read voltage window $[V_{\text{min}}, V_{\text{max}}]$ on the FeD device, the min-max normalization and linear mapping procedure to encode the input voltage $f(X)$ are described below:

\[
f(x) = \log(C)/\alpha + V_{\text{min}}
\]

With the linear characteristics between current $I$ and the function of voltage $f(X)$, we can linearly map the input to $f(X)$. For example, we have an input ranges from 0 to 1, we can map the input 1 to the max($\exp(\alpha V)$), input 0 to the min($\exp(\alpha V)$), and input between 0 and 1 to be distributed linearly between min($\exp(\alpha V)$) and max($\exp(\alpha V)$). Therefore, we can map each input to its corresponding $f(X)$, and decode $f(X)$ to find its corresponding Voltage $V$. 
Fig. S8. (a) A picture of a number ‘5’ as input from MNIST dataset. (b) Plotting after mapping and encoding the input signal into the realistic voltage \( V \), the feature remains after converting the input signal to the voltage applied on the device.

Fig. S8(b) shows a transformation from the given input from the MNIST shown in Fig. S8(a), with the intensity as the input signal, to the encoded voltage amplitudes applied on the device. By this method, we could use the AlScN ferroelectric diode as a simple resistor in neural network computation by simply encoding the input \( X \) to the function \( V = f(X) \), then we could have the current \( I \) measured as the output, shown below:

\[
I = G_1 f(X_1) + G_2 f(X_2) + G_3 f(X_3) + \cdots
\]  

(4)

To verify the encoding scheme on the realistic device, we encode a series of inputs in the range of [0, 1] to our read voltage range [4V, 8V]. Then, we directly apply the encoded voltages on the ferroelectric diode devices and measure the current as the output which is responding to the original input.

Fig. S9. The measured current at the output with applying the encoded voltages on the ferroelectric diode devices which are responding to the original input.

As shown in Fig. S9, the output current presents a superior linearity on the input showing a \( R^2 \) score of 0.9998 for a linear fit, which verifies that we could use the AlScN ferroelectric diode as a simple resistor in neural network computation by simply using this linearization encoding method.
Supplementary Note 3  Nonlinear weight update of ferroelectric diode

As we are typically using a linear quantization scheme when mapping the pre-trained weights to memristors, ideally, the amount of weight increase and weight decrease should be linearly proportional to the number of write pulses. However, the realistic devices reported in the literature do not follow such ideal trajectory. This is one of the main reasons inhibiting highly accurate hardware matrix multiplication. Those non-linear weight updates of real devices can be evaluated by a factor A [5]:

\[ G(N) = G_{\min} + (G_{\max} - G_{\min}) \frac{1 - e^{-N/A}}{1 - e^{-N_0/A}} \]

where \( G_{\min} \) and \( G_{\max} \) are the minimum and maximum conductance measured in the device, \( N \) is the pulse number we applied and \( N_0 \) is the maximum number we will apply. We could conclude from the above equations that as \( A \) decreases, the devices perform worse non-linear weight updates.

Fig. S10. (a) The normalized conductances by non-linear weight updates with different factor A. (b) The normalized conductances measured in the realistic FeD devices, which shows a near-ideal weight updates.

As shown in Fig. S10(b), the normalized conductances measured in our demonstrated FeD devices show near-ideal weight updates showing an \( A \) over 10 and \( R^2 \) score of 0.9997 to the ideal values.
| D1             | D2             | Store       | SL | SL | Search | Results |
|----------------|----------------|-------------|----|----|--------|---------|
| negative-forward | negative-forward | Don’t care  | Vr | 0  | 1      | Match   |
| negative-forward | negative-forward | Don’t care  | 0  | Vr | 0      | Match   |
| positive-forward | negative-forward | 1           | Vr | 0  | 1      | Match   |
| positive-forward | negative-forward | 1           | 0  | Vr | 0      | Mis     |
| negative-forward | positive-forward | 0           | Vr | 0  | 1      | Mis     |
| negative-forward | positive-forward | 0           | 0  | Vr | 0      | Match   |

**Table. S1.** Voltage modes and encoding table for various values of stored states, search voltages, and search results for the demonstrated TCAM.
**Fig. S11.** a, Schematics of the signal sequences for the PUND measurements to differentiate the ferroelectric and non-ferroelectric contributions to the polarization. b, PUND current densities showing ferroelectric switching within 400 ns of the onset of the voltage switching pulse.
Fig. S12. Dynamic current response in 45 nm AlScN. a, Schematics of the signal sequences for the dynamic current response measurements to observe the ferroelectric switching induced current response. b, The current-voltage hysteresis loops (extracted from I-t curve) of a 45 nm thick AlScN corresponding to the signal sequences shown in a. The above plot shows a positive coercive field of +4.62 MV/cm and a negative coercive field of -3.79 MV/cm. Leakage optimization of the ferroelectric films is the subject of ongoing work.
**Fig. S13.** The TCAM cell structure makes it natural to utilize the FeD crossbar memory array, in which the signal lines connecting to the anode and to the cathode are parallel in a bit search for the TCAM demonstration.
Fig. S14. The equivalent resistor-capacitor (RC) circuits have been shown in (a) for 2-FeD and in (b) for 2T-2R based TCAMs, which could explain why the search delay in our FeD-based TCAM can be reduced in comparison to prior TCAM architectures based on 2T-2R architecture.
Fig. S15. I-V curve of a single sweep over the demonstrated FeD device. After being programmed by a positive voltage (sweep 1-2), the resistance changes from high-to-low, and the polarity of the device changes from a negative-forward diode to a positive-forward diode. Similarly, it can be observed in a negative voltage sweep that the polarity of the device changes from that of a positive forward diode to that of a negative-forward diode.
**Fig. S16. P-E loop measurements at 100 kHz in 45 nm AlScN.** A saturation in polarization is clearly observed at negative applied voltages while the leakage convolutes observation of saturation for positive voltages. The coercive field is observed to be slightly larger than the values observed in our DC measurements and the values reported in Ref. [6-8] at 300 K. We think that this is because the coercive field is reported to significantly increase as frequency increases [6, 9-10].
**Fig. S17.** Fitting of experimental current-voltage data to the Poole-Frenkel tunneling model showing a good fit.
Fig. S18. Plot showing scaling down of the operating voltages of the FeD devices with reducing AlScN thickness.
Fig. S19. (a) The transient SPICE simulation on match line voltage vs search line voltage timing. (b) Benchmark comparison chart of lateral footprint of various TCAM cells vs. search delay [11-15]: resistive random-access memories (RRAMs) [11], magnetic tunnel junction (MTJ) RAMs [12], floating gate transistor memory (FLASH) [13], phase change memories (PCMs) [14], and ferroelectric field-effect-transistor (FeFET) [15].
References

[1] Guo, R., Lin, W., Yan, X., Venkatesan, T., & Chen, J. (2020). Ferroic tunnel junctions and their application in neuromorphic networks. Applied physics reviews, 7, 011304.

[2] T. P. Ma and Jin-Ping Han, "Why is nonvolatile ferroelectric memory field-effect transistor still elusive?" in IEEE Electron Device Letters, vol. 23, no. 7, pp. 386-388, July 2002.

[3] R. Berdan et al., Low-power linear computation using nonlinear ferroelectric tunnel junction memristors. Nature Electronics 3, 259-266 (2020).

[4] C. Hu, “Modern Semiconductor Devices for Integrated Circuits,” Pearson/Prentice Hall, New Jersey, 351 pages, 2010.

[5] P.-Y. Chen, X. Peng, S. Yu, NeuroSim: A circuit-level macro model for benchmarking neuro-inspired architectures in online learning. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 37, 3067-3080 (2018).

[6] S. Fichtner, N. Wolff, F. Lofink, L. Kienle, and B. Wagner, J. Appl. Phys. 125, 114103 (2019).

[7] Dixiong Wang, Jeffrey Zheng, Zichen Tang, Michael D’Agati, Paria S. M. Gharavi, Xiwen Liu, Deep Jariwala, Eric A. Stach, Roy H. Olsson III, Volker Roebisch, Martin Kratzer, Bernd Heinz, Myung-Geun Han, Kim Kisslinger, Joint Conference of the IEEE International Frequency Control Symposium and International Symposium on Applications of Ferroelectrics 2020, (2020).

[8] Dixiong Wang, Jeffrey Zheng, Pariasadat Musavigharavi, Wanlin Zhu, Alexandre C. Foucher, Susan E. Trolier-McKinstry, Eric A. Stach, Roy H. Olsson III, IEEE Electron Device Lett. 41, 1774 (2020).

[9] S. Fichtner, D. Kaden, F. Lofink and B. Wagner, International Conference on Solid-State Sensors, Actuators and Microsystems & Eurosensors XXXIII (TRANSUDCERS & EUROSENSORS XXXIII) 2019, 289 (2019).

[10] Wei Li, Zhijun Chen, Orlando Auciello, J. Phys. D: Appl. Phys. 44, 105404 (2011).

[11] Nii, K.; Amano, T.; Watanabe, N.; Yamawaki, M.; Yoshinaga, K.; Wada, M.; Hayashi, I. In 13.6 A 28nm 400MHz 4-parallel 1.6 Gsearch/s 80Mb ternary CAM, 2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC), IEEE: 2014; pp 240-241.

[12] Song, B.; Na, T.; Kim, J. P.; Kang, S. H.; Jung, S.-O., A 10T-4MTJ nonvolatile ternary CAM cell for reliable search operation and a compact area. IEEE Transactions on Circuits and Systems II: Express Briefs 2016, 64 (6), 700-704.
[13] Fedorov, V. V.; Abusultan, M.; Khatri, S. P. In An area-efficient ternary cam design using floating gate transistors, 2014 IEEE 32nd International Conference on Computer Design (ICCD), IEEE: 2014; pp 55-60.

[14] Li, J.; Montoye, R. K.; Ishii, M.; Chang, L., 1 Mb 0.41 µm² 2T-2R cell nonvolatile TCAM with two-bit encoding and clocked self-referenced sensing. IEEE Journal of Solid-State Circuits 2013, 49 (4), 896-907.

[15] Ni, K.; Yin, X.; Laguna, A. F.; Joshi, S.; Dünkel, S.; Trentzsch, M.; Müller, J.; Beyer, S.; Niemier, M.; Hu, X. S., Ferroelectric ternary content-addressable memory for one-shot learning. Nature Electronics 2019, 2 (11), 521-529.