Supporting information

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

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

[a] Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA, USA

[b] Material Science and Engineering, University of Pennsylvania, Philadelphia, PA, USA

[c] Laboratory for Research on the Structure of Matter, University of Pennsylvania, Philadelphia, PA, USA

[d] Microsystems Engineering, Science and Applications (MESA), Albuquerque, New Mexico, USA

[e] Brookhaven National Laboratory, Center for Functional Nanomaterials, Upton, NY, USA

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

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:

$$T(E) = \exp(-\frac{2}{\hbar}\int_{x_1}^{x_2}\sqrt{2m^*[(E_c + qV(x)) - E]}\mathrm{d}x) \tag{1.1}$$

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

$$N(E) = k_B T \ln\left(\frac{1+\exp\left(\frac{E-E_{f,1}}{k_B T}\right)}{1+\exp\left(\frac{E-E_{f,2}}{k_B T}\right)}\right) \tag{1.2}$$

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

$$J = \frac{4\pi m^* q}{\hbar^3}\int_{E_{min}}^{E_{max}} T(E)N(E)\mathrm{d}E \tag{1.3}$$

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.

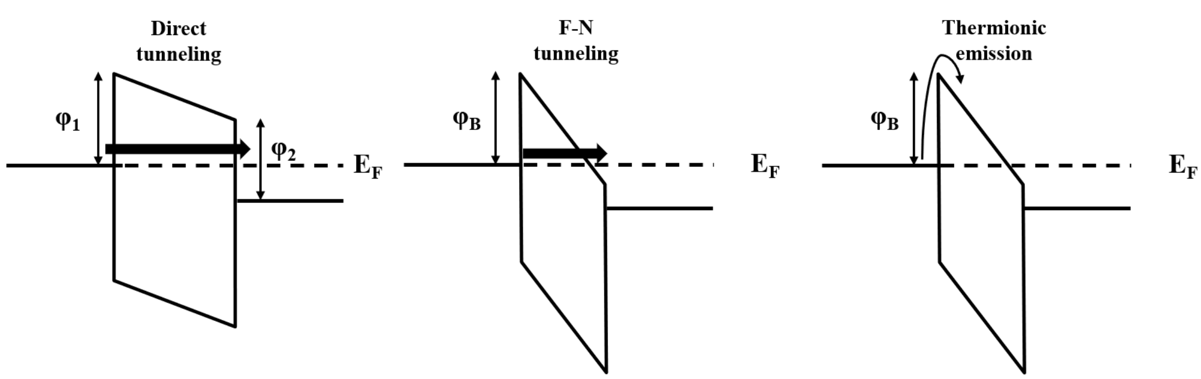

Fig. S1. Illustrations of the three methods of electron transport through a potential barrier. (a) direct tunneling (b) Fowler-Nordheim tunneling (c) thermionic emission.

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 ΔV as the diode transforms from LRS to HRS. In other words, it takes more voltage to offset the voltage of ΔV at HRS to generate the same current as the current at LRS. The shifted voltage ΔV can be derived as:

$$\Delta V = E_{dp} t \quad (2.1)$$

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:

$$E_{dp} = E_C \tanh\left(\frac{\beta P_r}{\epsilon_{fe} E_C}\right) \quad (2.2)$$

where E_C is the coercive field of the ferroelectric layer, P_r is the remanent polarization, ϵ_{fe} is the dielectric constant of the ferroelectric, and β is the parameter related to the oxide capacitance C_{ox} and the ferroelectric capacitance C_{fe}, as [2]:

$$\beta = \left[\left(\frac{C_{ox}}{C_{fe}} + 1\right)\right]^{-1} \quad (2.3)$$

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 $\mathrm{E_{dp}}$ would not exceed the coercive field $\mathrm{E_c}$ in the model.

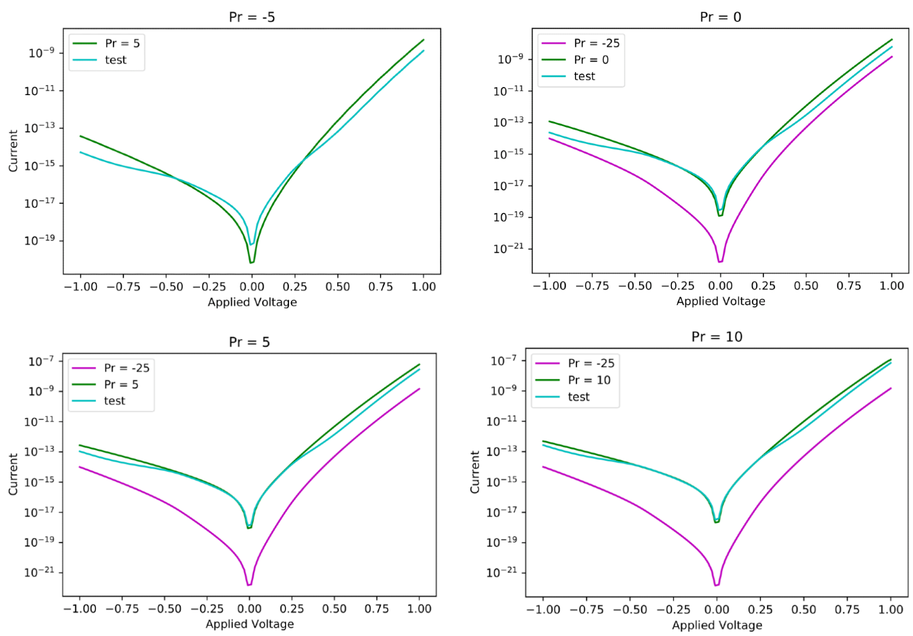

Fig. S2. Comparisons of the IV curves using the WKB approximation (green) vs. the new voltage shifting model for a ferroelectric diode (blue) for any arbitrary ferroelectric. The purple curve represents the base curve at high-resistance state modeled with the WKB approximation.

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 β using eq.2.3, and β 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 β is demonstrated in Fig. S3(b).

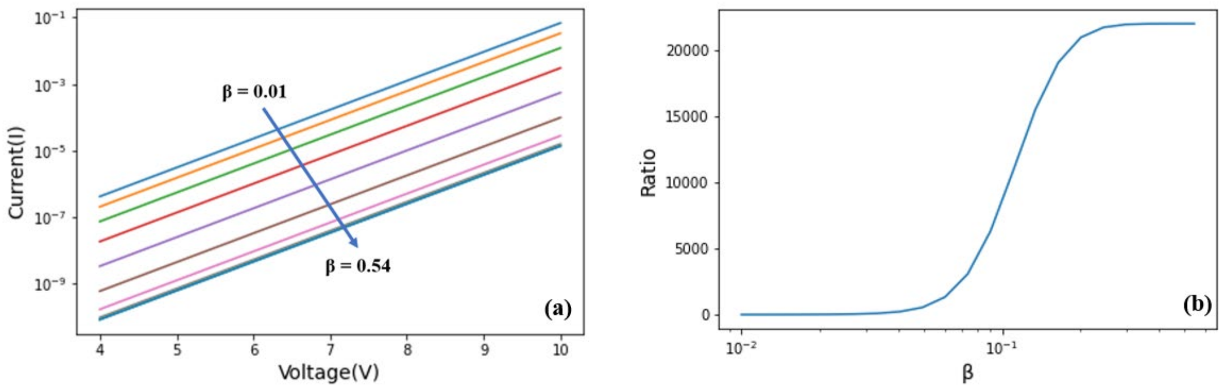

Fig. S3. The plot of HRS I-V curve shifting with the variation of the pre-coded oxide capacitance β is shown in (a), we vary β from 0.01 to 0.54 exponentially to get a better view of the I-V curve shifting trend. (b) is the ON/OFF ratio vs. β, and the ON/OFF ratio is the current ratio between LRS and 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 $\mathrm{P_r}$ from 1 to 135 $\mu\mathrm{C/cm^2}$ with a suitable insulator capacitance $\mathrm{C_{ox}}$ and coercive field $\mathrm{E_C}$, 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).

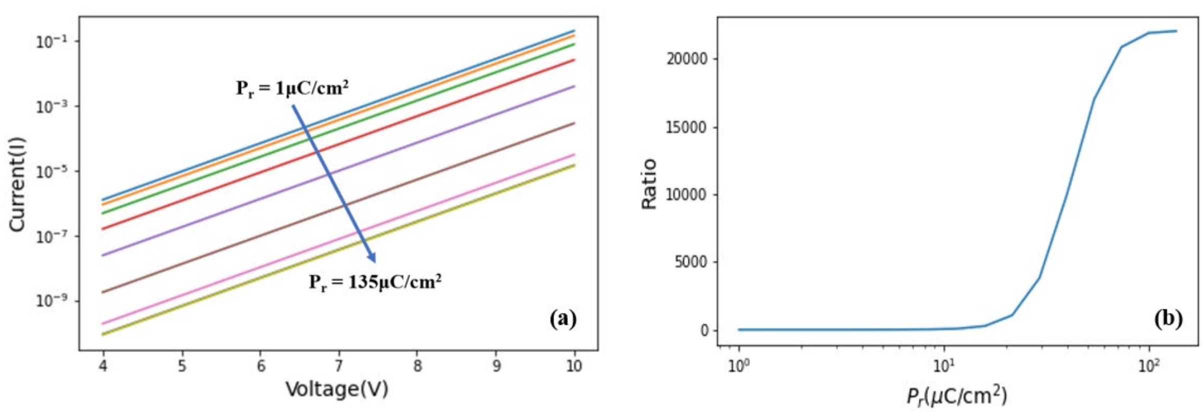

Fig. S4. The plot of HRS I-V curve shifting with the variation of the remanent polarization $\mathrm{P_r}$ is shown in (a), we vary $\mathrm{P_r}$ from 1 to 135 $\mu\mathrm{C/cm^2}$ exponentially to get a better view of the I-V curve shifting trend. (b) is the ON/OFF ratio vs. $\mathrm{P_r}$, and the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

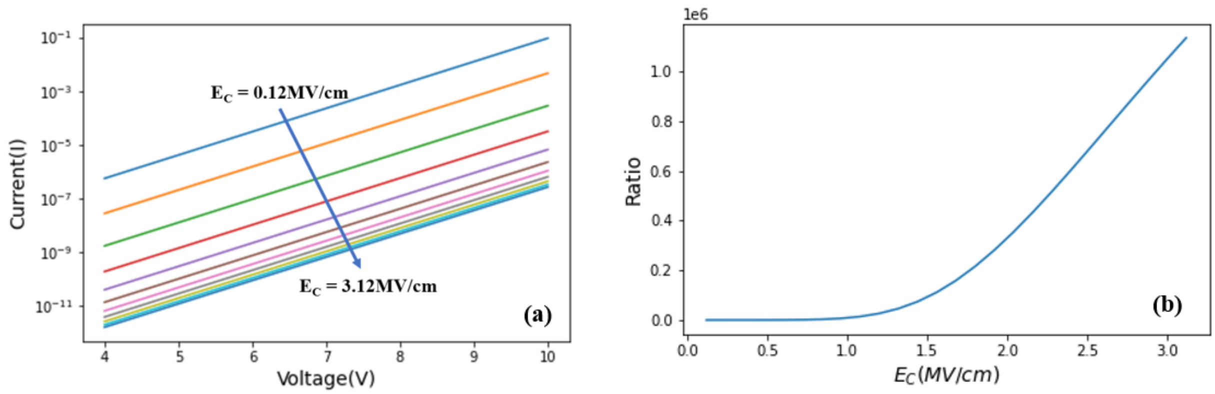

Fig. S5. The plot of HRS I-V curve shifting with the variation of the result of the coercive field E_C is shown in (a), we vary E_C from 0.12 to 3.12 MV/cm linearly to get a I-V curve shifting trend. (b) is the ON/OFF ratio vs. E_C, and the ON/OFF ratio is the current ratio between LRS and corresponding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

With the $\log I$ - V characteristics of the device shown in Fig. S6, the $\log I$ has excellent linearity with the voltage applied, and the slopes of the plots in 16 different states of the device are consistent in the fitting.

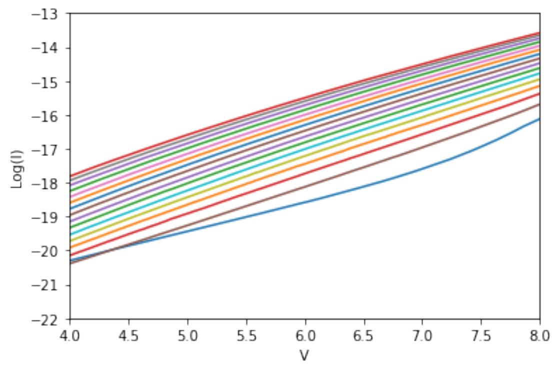

Fig. S6. Log(*I)* vs *V*. Current *I* measured by applying dc voltage sweeping to 16 different states.

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

$$I = G_i \exp(\alpha V) \tag{3}$$

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

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

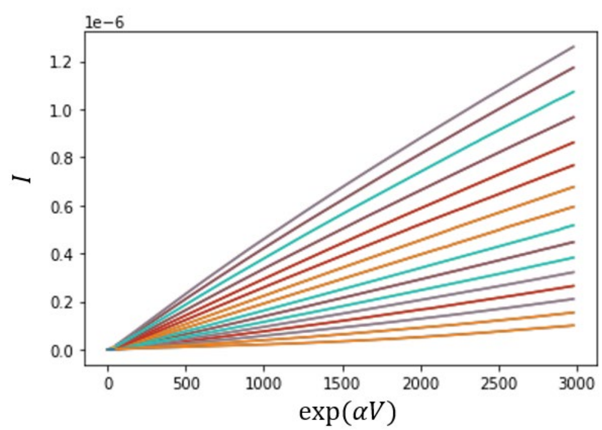

Fig. S7. Current *I* vs. $\exp(\alpha V)$ for 16 different states.

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.

With the linear characteristics between current I and the function of voltage $f(V) = \exp(\alpha V)$, we can linearly map the input to $f(V)$. 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 evenly between $\min(\exp(\alpha V))$ and $\max(\exp(\alpha V))$. Therefore, we can map each input to its corresponding $f(V)$, and decode $f(V)$ to find its corresponding Voltage V by:

$$V = \frac{1}{\alpha} \ln f(V) \tag{4}$$

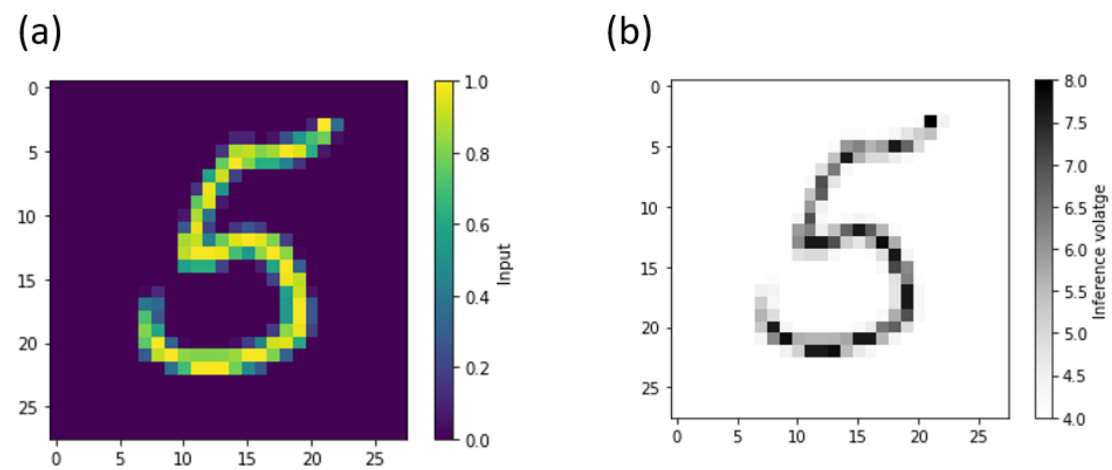

Fig. S8. (a) A picture of a number '5' as input from MNIST dataset. (b) Plotting after mapping and encoding the input signal into the realistic voltage V, the feature remains after converting the input signal to the voltage applied on the device.

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 V to the function $f(V)$, then we could have the current I measured as the output, shown below:

$$I = G_1 f(V_1) + G_2 f(V_2) + G_3 f(V_3) + \cdots \tag{5}$$

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 ferrolectric diode devices and meaure the current as the output which is responding to the original input.

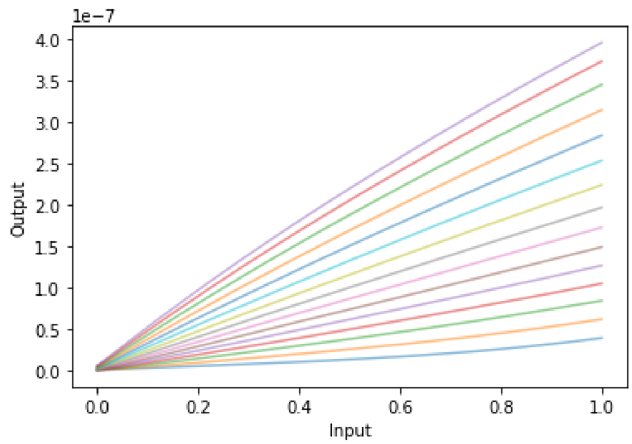

Fig. S9. The meaured current at the output with applying the encoded voltages on the ferrolectric diode devices which are responding to the original input.

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 where the conductance typically changes rapidly at the beginning then gradually saturates. This is one of the main reasons inhibiting highly accurate hardware matrix mulplication. Those non-linear weight updates of real devices can be evaluated by a factor A:

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

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 equaltions that as A decreases, the devices perform worse non-linear weight updates.

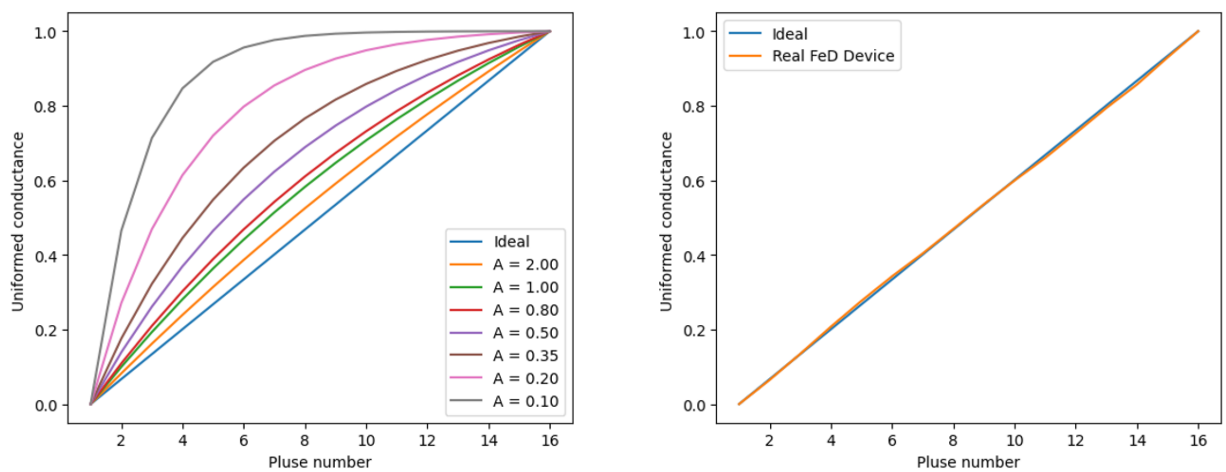

Fig. S10. (a) The normalized conductances by non-linear weight updates with different factor A. (b) The normalized conductances measured in the realistic FeD devices, which shows a near-ideal weight updates.

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	$\overline{SL}$	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.

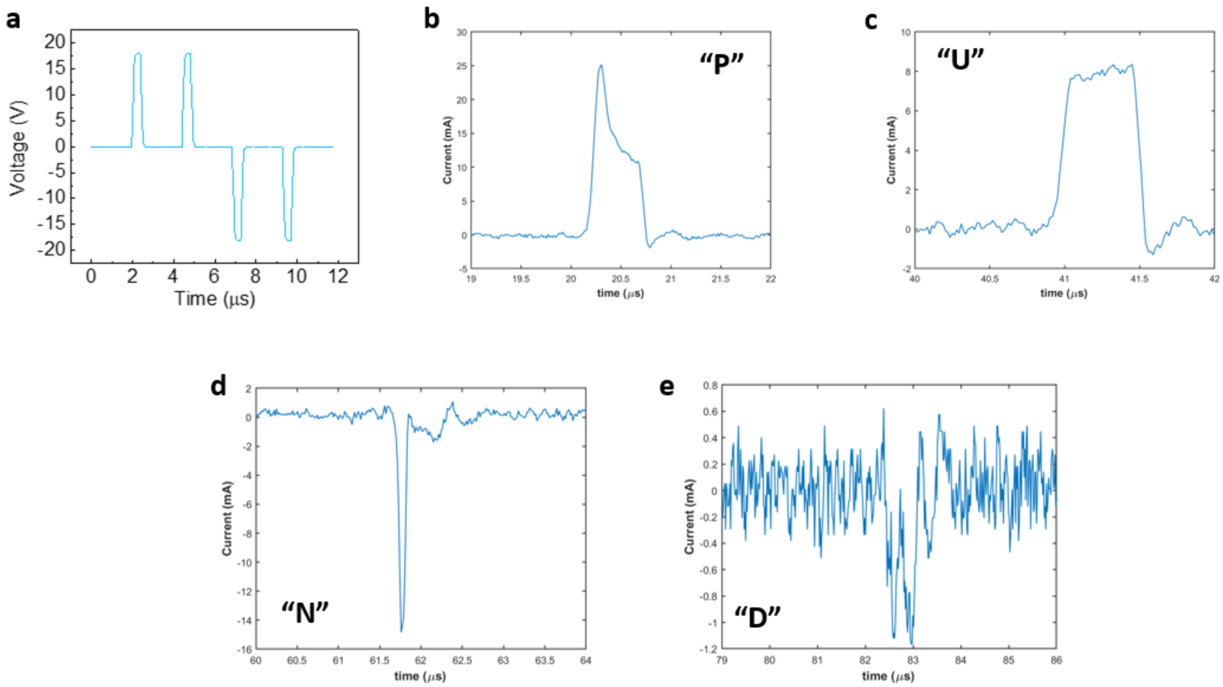

Fig. S11. a, Schematics of the signal sequences for the PUND measurements to differentiate the ferroelectric and non-ferroelectric contributions to the polarization. b, PUND current densities showing ferroelectric switching within 400 ns of the onset of the voltage switching pulse.

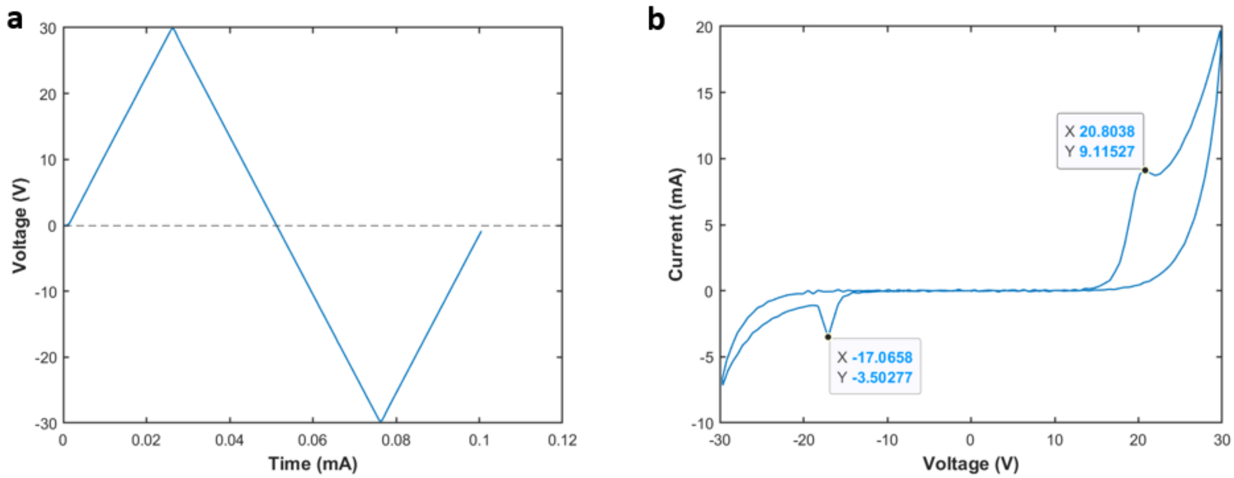

Fig. S12. Dynamic current response in 45 nm AlScN. a, Schematics of the signal sequences for the dynamic current response measurements to observe the ferroelectric switching induced current response. b, The current-voltage hysteresis loops of a 45 nm thick AlScN corresponding to the signal sequences 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 subject of ongoing work.

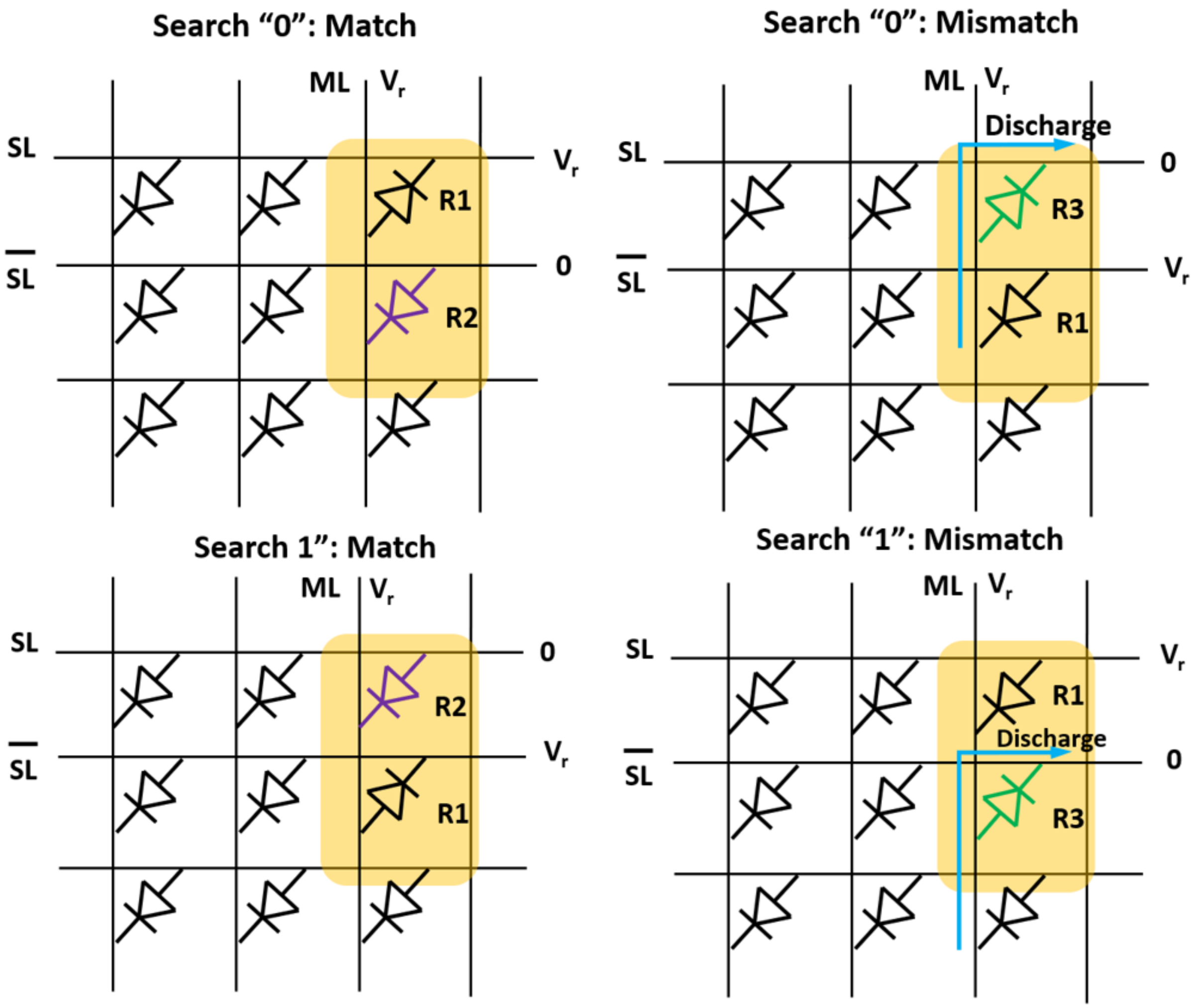

Fig. S13. The TCAM cell structure makes it natural to utilize the FeD crossbar memory array, in which the signal lines connecting to the anode and to the cathode are parallel in a bit search for the TCAM demonstration

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, doi: 10.1109/LED.2002.1015207.