A Novel Ferroelectric Superlattice Based Multi-Level Cell Non-Volatile Memory

Kai Ni[1], Jeffrey Smith[1], Huacheng Ye[1], Benjamin Grisafe[1], G. Bruce Rayner[3], Andrew Kummel[2], and Suman Datta[1]

[1]University of Notre Dame, Notre Dame, IN, USA; [2]University of California, San Diego, San Diego, CA, USA; [3]Kurt J. Lesker Co., PA, USA; email: kni@nd.edu

***Abstract*—**We demonstrate a novel ferroelectric (FE) superlattice based multi-level cell (MLC) memory, which outperforms previous multi-state FE memory implemented using partial polarization switching, from the standpoint of device-to-device variation. We show that the FE superlattice consisting of alternate FE and dielectric (DE) thin layers provides a scalable approach for MLC implementation because: 1) the superlattice constructs controlled layer-by-layer polarization switching in the constituent FE layers; 2) the number of FE layers equals the number of stored bits; and, finally, (3) the switching of all the domains in a given coercive field (E_C) distribution associated with one isolated peak suppresses the variation induced by partial polarization switching. Based on these, we experimentally demonstrate a 2-bit/cell FE superlattice memory and simulate a 3-bit/cell memory with excellent device-to-device variation.

I. Introduction

Doped HfO_2 based FeFET has emerged as one of the leading candidates for embedded nonvolatile memory (eNVM) due to its excellent CMOS compatibility and scalability. As a result, it has been integrated in advanced technology nodes and excellent performance has been demonstrated [1]. However, several challenges remain for it to become a competitive eNVM technology. The first challenge is the limited endurance caused by charge injection during the write operation and high write voltage to switch the polarization. Several device-level solutions have been proposed to overcome both the endurance and logic compatible voltage bottlenecks [2].

A remaining key bottleneck for FeFET is the exacerbation of device-to-device variation with scaling of device dimensions and eventual collapse of the memory window [3]. It is caused by partial polarization switching, where write pulses sample different portions of the E_C distribution for different devices. The reduced number of domains in scaled device, coupled with domain inhomogeneity and stochasticity of domain switching, make the variation induced by partial polarization switching hard to control even for binary memory. The variation only gets worse for MLC implemented through partial polarization switching, and the programmed states strongly overlap with one another and are practically indiscernible (Fig.2(a)).

Complete polarization switching with a single write pulse, where domains in an E_C distribution are all switched irrespective of their local environment, has been shown to be an effective method to suppress variation [3]. Thus, MLC FE memory with well-controlled variation can be realized if a multi-peak E_C distribution can be designed such that each distribution associated with an isolated peak represents precisely one bit of information and get fully switched during the write operation (Fig.2(b)). In this work, we design and experimentally demonstrate a novel FE superlattice memory (Fig.1), where FE layers are separated by non-FE layers, to realize a multi-peak E_C distribution. With theoretical modeling, we show that this multi-peak E_C landscape is achieved through controlled layer-by-layer switching of individual FE layers. This means that it is theoretically feasible to even realize a N-bit cell with N layers of FE. To that effect, we demonstrate a 3-bit/cell superlattice MLC with excellent control of variation using a kinetic Monte Carlo (KMC) modeling framework.

II. Device Fabrication Process

The key processing steps to fabricate metal-ferroelectric-metal (MFM) capacitors are shown in Fig.3(a). In additional to the HfO_2/ZrO_2 superlattice, two control devices, FE $Hf_{0.4}Zr_{0.6}O_2$ and anti-ferroelectric (AFE) ZrO_2 (all 10nm), are also fabricated for comparison. All the devices are fabricated on a highly doped (10^{19} cm^{-3}) p-Si substrate with tungsten (W) bottom and top electrodes. All samples are annealed at 600 ºC for crystallization and stabilization of the FE and AFE phases.

High resolution cross-section TEM images of the three stacks (Fig.3(b-d)) show that the superlattice is poly-crystalline but the composition remains intact for all the layers. There is no intermixing between the HfO_2 and the ZrO_2 layers. The Q_{FE}-V_{FE} hysteresis loops of the superlattice exhibit a two-step switching characteristic, different from the stand-alone FE and AFE controls. Unlike the $Hf_{0.4}Zr_{0.6}O_2$ FE control, the superlattice shows a weaker ferroelectricity. However, it is non-volatile and retains its polarization states, which is different from the AFE ZrO_2 control. The two-step switching corresponds to a double-peak distribution of E_C, as shown in the first-order reversal curve (FORC), which measures the Preisach density of polarization switching [4]. Two peaks of the same polarity appear for the same program voltage polarity, in sharp contrast to that in AFE, where two peaks are of opposite polarities for the same voltage polarity due to polarization relaxation.

III. Results and Discussion

A. Ferroelectric Superlattice Working Principle

To understand the origin of multi-peak distribution of E_C in FE superlattice, we model the superlattice stack based on the Landau-Khalatnikov theory of FE (Fig.4) [5]. We combine the volume Gibbs free energy of the individual FE and DE layers and weight with their respective thicknesses to obtain the total surface Gibbs free energy. After considering the electrical boundary conditions, the surface Gibbs free energy is a function of the polarization in each FE layer (3 layers in our study). Therefore, for every bias point, the polarization states can be determined by minimizing the surface Gibbs free energy.

The simulated P_{FE}-V_{FE} hysteresis loop of the superlattice with three FE layers and two DE layers in between exhibits a three-step switching characteristic (Fig.5(a)), indicating a triple-peak E_C distribution. The polarization configuration in each layer is shown at several critical points on the loop. The intermediate states (point B, C and E, F) correspond to a single layer switching. When the voltage is swept from negative to positive, a controlled layer-by-layer switching is observed. This results from the electrostatic interaction between the polarizations in different layers. Switching in one layer modifies the electric field in the others, which in turn changes their switching characteristics. For a 3-bit MLC cell, only 6 states are shown in Fig.5(a), while the other 2 states are hidden. This is because those hidden states (states G and H, shown in Fig.5(b)) need a specific combination of pulses to access. For example, the state G can be accessed by reaching state C first, and then switch the first layer back again. In this way, we can see that FE superlattice provides an effective approach to realize MLC memory. The surface Gibbs free energy contour (Fig.5(c-j)) is projected onto the P_{FE1}-P_{FE2} plane (P_{FE3} value is fixed). With the change in electrical bias, the progressive switching of an individual FE layer is clearly shown.

B. Experimental Device Characteristics

We measured the MLC characteristics on two superlattice samples with 5 layers. In one case, the ZrO_2 layers are separated by 1nm thick HfO_2 spacer layer, while, in the other case, the HfO_2 spacer layer thickness is increased to 5nm (Fig.6(a-f)). A double-peak distribution is readily observed in both samples, suggesting that the ultra-thin ZrO_2 layer is ferroelectric [6]. The separation of the two peaks increases with the spacer layer HfO_2 thickness, which provides an insightful guideline towards the optimization of the E_C distributions for MLC operation, as it is critical to reduce the overlap between the peaks so that switching of each peak is independent of each other.

A modified positive-up-negative-down (PUND) pulse sequence is applied to verify memory write and read operations (Fig.7(a)). The device is initialized by the first two pulses and the last two pulses read out the residual polarization that left unswitched by the write pulse. In this double-peak distribution device, 2 bit/cell, corresponding to four write pulses are demonstrated (Fig.7(b-f)). The transient current waveform induced by the read pulses (Fig.7(b)) and I_{FE}-V_{FE} characteristics during the write pulses (Fig.7(c) and (e)) show that ±1V write pulses only switch the distribution associated with the first peak, while ±2V write pulses switch both the peaks. Therefore, through pulse engineering, we can individually switch all the domains in one distribution associated with an isolated peak. These results indicate that we can utilize the multi-peak E_C distribution to realize a MLC memory device. We also studied the endurance and retention of the superlattice, which shows more than 10^6 cycles and no polarization loss at 85°C for 1hr.

C. Advantages of Superlattice in Variation Control

Multi-state operation of FE memory can also be realized through partial polarization switching (Fig.2(a)), as shown before in ferroelectric synaptic weight cell [7]. To illustrate the advantages of full polarization switching in multi-peak superlattice over the partial polarization switching in FE in implementing MLC, we adopt the KMC framework to evaluate the device-to-device variation [3]. Figs.8 (a-c) and (f-h) show the calibration of the KMC model to the measured Q_{FE}-V_{FE}, I_{FE}-V_{FE} characteristics and the extracted E_C distribution for both the superlattice (Fig.6(d)) and stand-alone FE (Fig.3(b)), respectively. The fitted parameters are then used to simulate the device-to-device variation in 2-bit/cell MLC memory.

The devices are first initialized and write pulses similar to Fig.7(a) are applied to access all the four levels. The variation of the P_{FE} during the write operation are shown in Figs.8(d) and (i) for both the superlattice and the FE. Clearly, with 100 domains, the superlattice exhibits a much better controlled variation compared with FE. This is because each peak in the double-peak distribution corresponds to one level and gets fully switched by the write pulses. Therefore, the variation is tight irrespective of the distribution. However, for the partial polarization switching in FE, the distribution makes a huge impact on the variation. Fig.8(e), (j) show the P_{FE} distribution after write pulses. States overlap in the FE case, whereas they are well separated in the superlattice.

The effect of scaling (here the domain number is reduced to 20) on the variation is shown in Figs.9(a) and (b) for the superlattice and FE, respectively. The reduced domain number greatly increases the variation, which remains a big challenge for the FE memory [3]. However, because of the full polarization switching in the superlattice, the variation is still tight even with limited number of domains. Therefore, the superlattice can potentially provide a scaling path for FE MLC memory. We further studied the possibility of implementing a 3bit/cell MLC using the superlattice. We add one more peak to the double-peak E_C distribution in Fig.8(c) to form a triple-peak distribution and simulate the variation of the 8 levels (Fig.10(a)). The states strongly overlap in FE; while they remain separated in the superlattice. From these comparisons between the superlattice and FE, and also considering the current power-hungry, high-latency, and limited endurance embedded flash memory (Fig.11), superlattice based ferroelectric memory emerges as a promising candidate for future eNVM with MLC capability.

IV. CONCLUSIONS

In summary, we demonstrate a novel ferroelectric superlattice based MLC non-volatile memory with excellent variation. It relies on the full switching of individual peaks in the multi-peak E_C distribution created by the superlattice. We experimentally demonstrate a 2bit/cell MLC and simulated a 3bit/cell MLC with well-controlled distribution. Considering its excellent performance (write power, latency, endurance, retention etc.), the FE superlattice provides a scalable approach for MLC non-volatile memory implementation.

ACKNOWLEDGEMENT

This work supported in part by ASCENT, one of six centers in JUMP, sponsored by DARPA and Semiconductor Research Corporation (SRC).

REFERENCES

[1] S. Dunkel et al., "A FeFET based super-low-power ultra-fast embedded NVM technology for 22nm FDSOI and beyond," IEDM 2017
[2] K. Ni et al., "SoC logic compatible multi-bit FeMFET weight cell for neuromorphic applications," IEDM 2018
[3] K. Ni et al., "Fundamental understanding and control of device-to-device variation in deeply scaled ferroelectric FETs," VLSI 2019.
[4] T. Schenk et al., "Complex internal bias fields in ferroelectric hafnium oxide," ACS Appl. Mater. Interfaces 2015.
[5] G. A. Boni, et al., "Multiple polarization states in symmetric ferroelectric heterostructures for multi-bit non-volatile memories," Nanoscale 2017.
[6] M. Dogan, et al., "Single atomic layer ferroelectric on silicon," Nano Lett. 2018
[7] M. Jerry et al., "Ferroelectric FET analog synapse for acceleration of deep neural network training," IEDM 2017.

Motivation: Ferroelectric Superlattice for Multi-Level Cell (MLC)

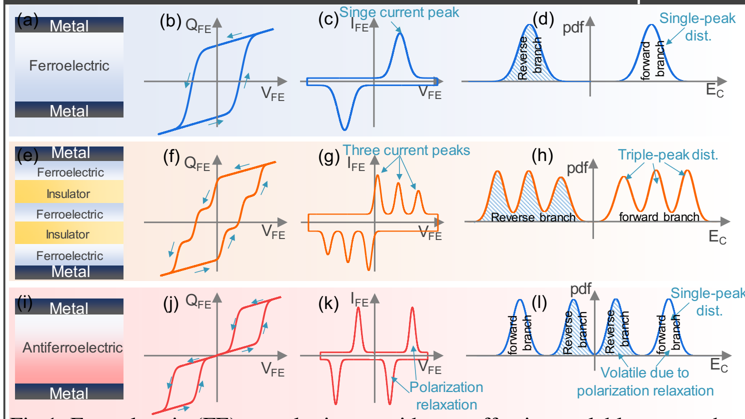


Fig.1. Ferroelectric (FE) superlattice provides an effective scalable approach to create a multi-peak E_C distribution, which could be harnessed to demonstrate non-volatile MLC memory device (N ferroelectric layers can store N bits).

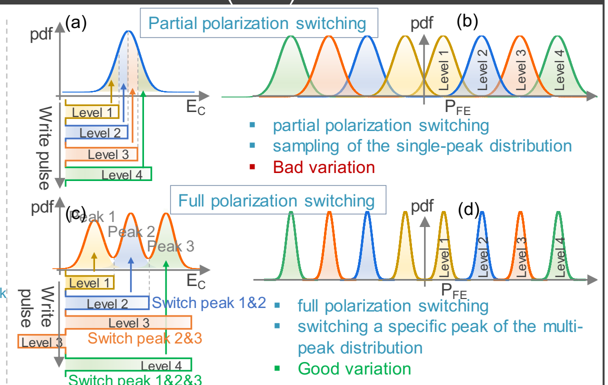


Fig.2. Selective switching of all domains in an isolated peak in multi-peak E_C distribution suppresses the device-to-device variation in MLC (c-d) compared with partial polarization switching in single-peak ferroelectric (a-b).

Ferroelectric Superlattice Fabrication Process

(a)
- W bottom electrode sputtering on 10^{19}cm-3 p-Si substrate
- Thermal ALD deposition (250°C)
  - $Hf_{0.4}Zr_{0.6}O_2$
  - ZrO_2
  - HfO_2/ZrO_2 superlattice
- Lithography for top electrode
- W top electrode sputtering
- RTA anneal at 600°C for 30s in N_2

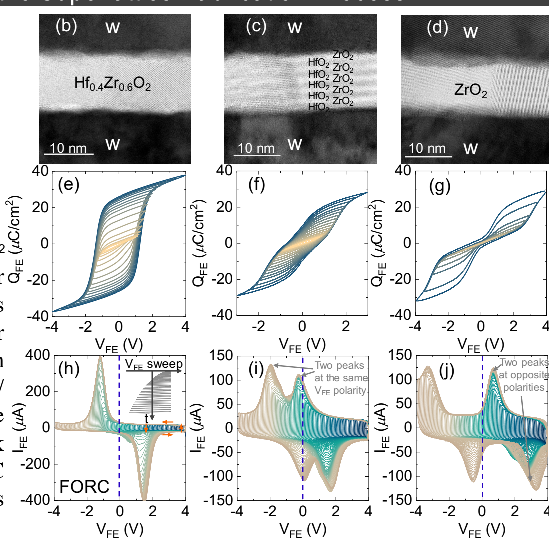
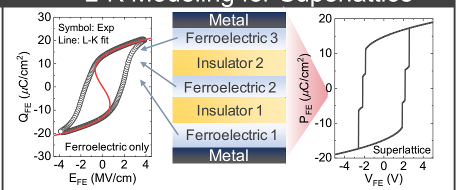


Fig.3. (a) Process flow for superlattice. (b-d)/(e-g)/(h-j) shows the TEM/Q_{FE}-V_{FE}/first-order reversal curve (FORC) for 10nm $Hf_{0.4}Zr_{0.6}O_2$, superlattice (HfO_2/ZrO_2), and ZrO_2 samples. The superlattice exhibits double-peak distribution, critical to realize MLC memory. ZrO_2 is volatile and loses its state without bias.

L-K Modeling for Superlattice



Gibbs Energy

$$\begin{cases} g_{DEi} = -\frac{1}{2}\varepsilon_{DEi}E_{DEi}^2 \\ g_{FEi} = -\frac{1}{2}\varepsilon_{FEi}E_{FEi}^2 - P_{FEi}E_{FEi} + \frac{1}{2}\alpha_{FEi}P_{FEi}^2 + \frac{1}{4}\beta_{FEi}P_{FEi}^4 \end{cases}$$

Boundary cond.

$$\begin{cases} \varepsilon_{FE1}E_{FE1} + P_{FE1} = \varepsilon_{DE1}E_{DE1} = \varepsilon_{FE2}E_{FE2} + P_{FE2} \\ = \varepsilon_{DE2}E_{DE2} = \varepsilon_{FE3}E_{FE3} + P_{FE3} \\ E_{FE1}l_{FE1} + E_{DE1}l_{DE1} + E_{FE2}l_{FE2} + E_{DE2}l_{DE2} + E_{FE3}l_{FE3} = V_a \end{cases}$$

Minimize surface Gibbs energy density with respect to $P_{FE1}, P_{FE2}, P_{FE3}$ at a given bias V_a

Surface Gibbs Energy

$$G(V_a, P_{FE1}, P_{FE2}, P_{FE3}) = \sum_i l_{FEi}g_{FEi} + \sum_j l_{DEj}g_{DEj}$$

Fig.4. Landau-Khalatnikov model of ferroelectric superlattice. The surface Gibbs free energy is minimized to obtain polarization in all ferroelectric layers at a given bias.

Device Working Principle

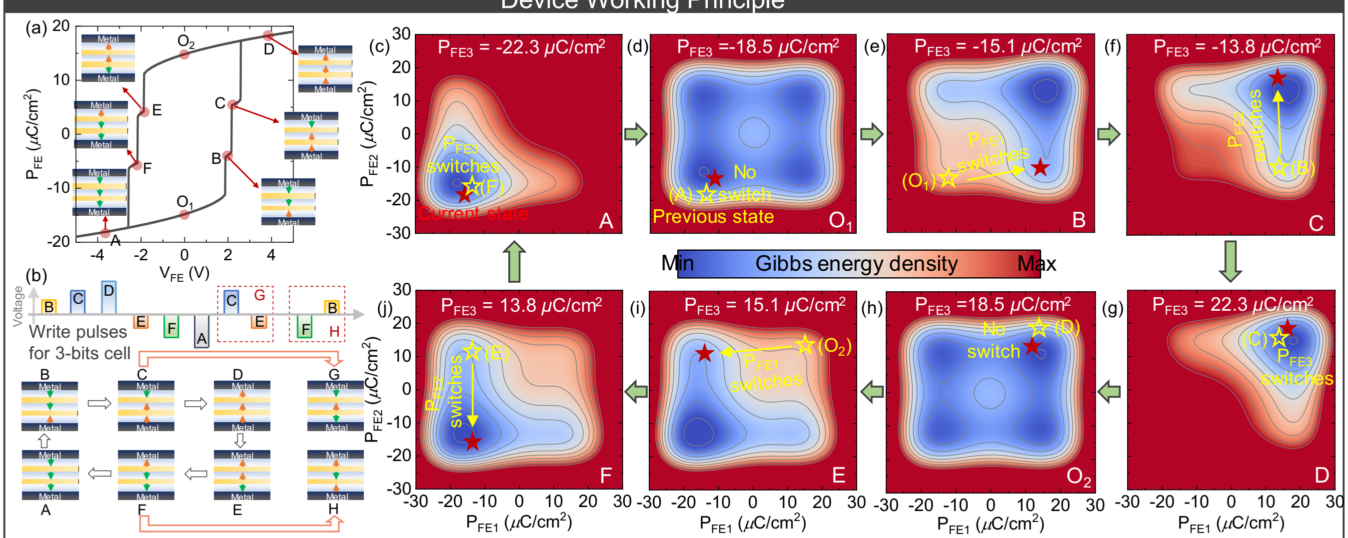


Fig.5. (a) Simulated P_{FE}-V_{FE} characteristics show three step switching, corresponding to individual layer switching. This effectively stores 3 bit per cell. (b) The write pulses to access individual states in a 3 bit cell. Only 6 states are accessed in (a). Two hidden states (G and H) can be accessed with combinations of pulses. (c-h) The surface Gibbs free energy contour projected onto the P_{FE1}, P_{FE2} plane (fixed P_{FE3}) at points highlighted in (b). It clearly shows that controlled layer-by-layer switching happens in ferroelectric superlattice, thereby enabling MLC.

Experimental Device Characteristics: Multi-Level Cell Write and Read

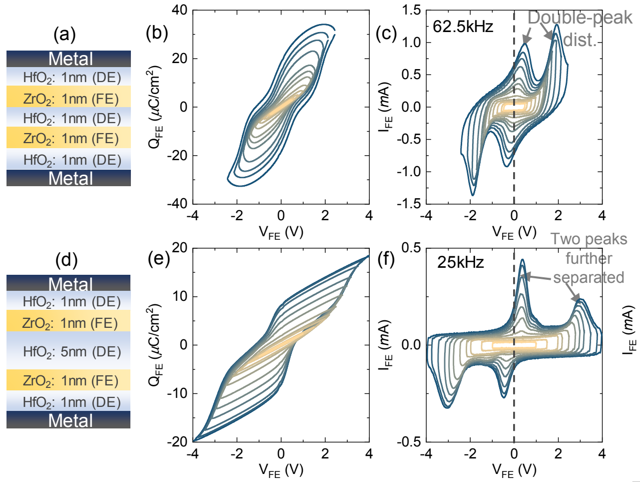


Fig.6. Superlattice stacks (a) and (d) to achieve multi-level cell (MLC) and their corresponding Q_{FE}-V_{FE} ((b) and (e)) and I_{FE}-V_{FE} ((c) and (f)) characteristics. The increase of middle HfO_2 layer thickness provides a way to increase the separation of the peaks in the double-peak distribution. Selective switching of a distribution peak, while not disturbing domains in other peaks can achieve MLC with full polarization switching.

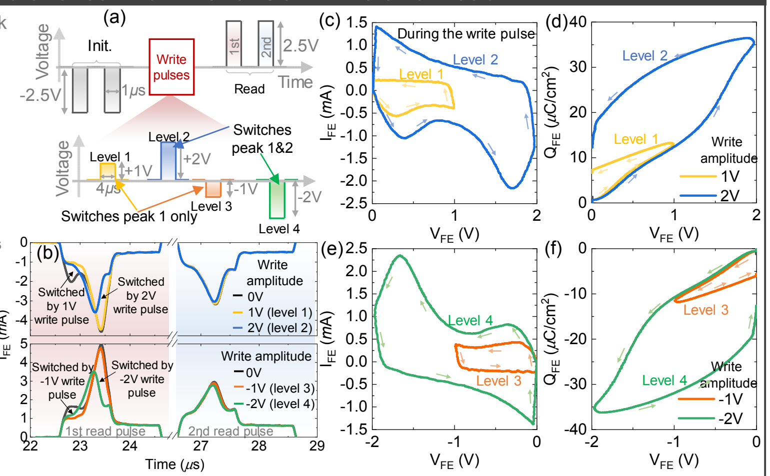


Fig.7. (a) Pulse schemes for MLC write and read operation. The write pulse amplitude to achieve one level is selected to ensure that the distribution associated with a specific peak is fully switched. (b) Measured I_{FE} for the read pulses, which measures the polarization left unswitched by the write pulses. ±1V switches the first peak while ±2V switches both peaks. (c)/(e) I_{FE}-V_{FE} and (d)/(f) Q_{FE}-V_{FE} during the write pulses. Distribution associated with different peaks can be selectively switched with different pulse amplitudes.

Advantages of Full Polarization Switching over Partial Polarization Switching: Well-Controlled Variation

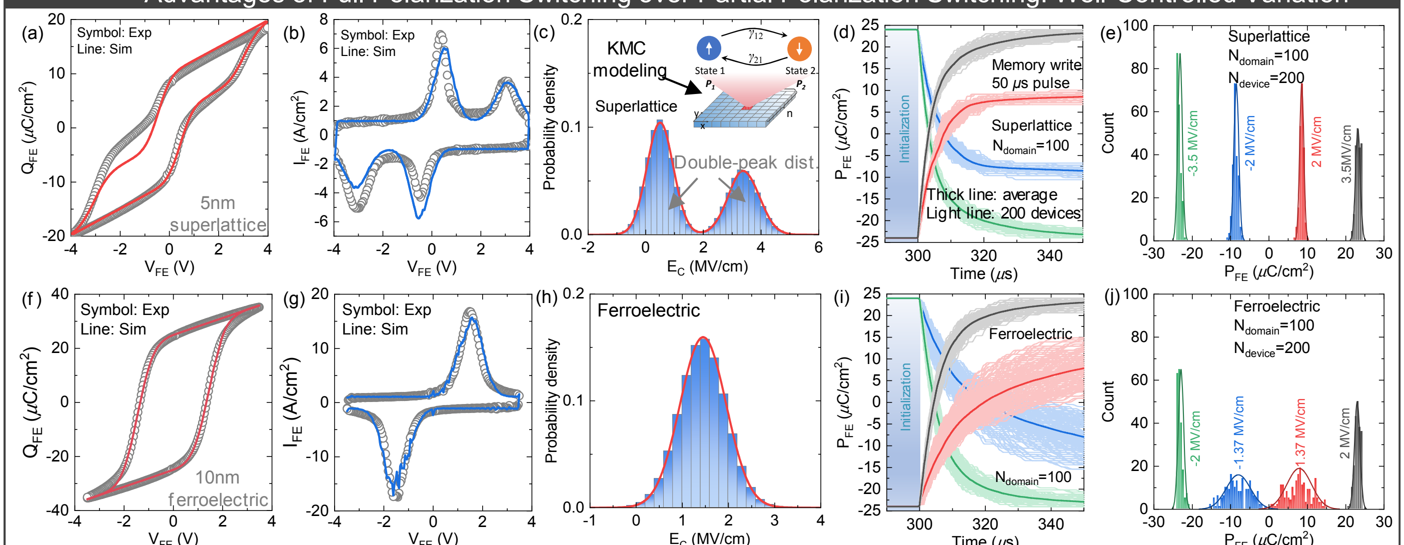


Fig.8. Comparison of device-to-device variation between ferroelectric devices with multi-peak distribution (a-e) and those with single-peak distribution (f-j). Coercive field (E_C) distribution (c)/(h) is extracted by calibration of the kinetic Monte Carlo (KMC) model (inset figure in (c)) with the measured Q_{FE}-V_{FE} (a)/(f) and I_{FE}-V_{FE} (b)/(g) characteristics. Variation is significantly degraded for MLC realized utilizing partial polarization switching (i) compared with full polarization switching (d). MLC states overlap for regular ferroelectric whereas tight distribution can be achieved in full polarization switching. Domain inhomogeneity, small number of domains, and stochasticity in switching cause variation degradation, which can be efficiently suppressed with full polarization switching.

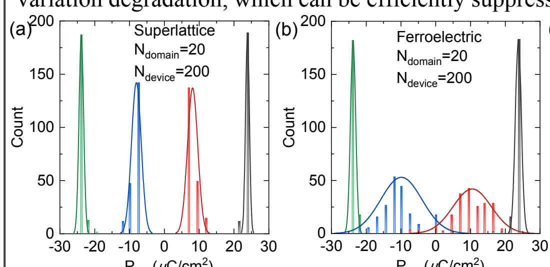


Fig.9. Device scaling (20 domains) causes significant states overlap for partial polarization switching (b) compared with full polarization switching (a). Superlattice provides an effective pathway for scaling.

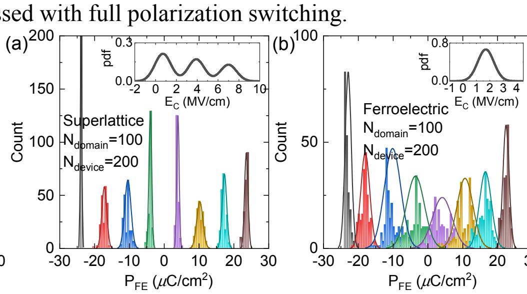


Fig.10. Eight MLC levels achieved in (a) triple-peak distribution devices using full polarization switching and (b) single-peak distribution devices. Superlattice is promising to achieve 3 bits or even 4 bits per cell.

MLC	Flash	Ferroelectric	Superlattice (this work)
Storage Mechanism	Charge storage	Partial polarization switching	Full polarization switching
Write Mechanism	Current driven	Field driven	Field driven
Write Energy	High (~100 pJ)	Low (~100 fJ)	Low (~100 fJ)
Write Speed	Slow (~1 ms)	Fast (~10 ns)	Fast (~10 ns)
Variation	Tight	Poor	Tight
Endurance	Poor (~10^3)	Good (>10^8)	Good (>10^6)
Retention	Good	Good	Good
Number of Bits	Good (3-4 bits/cell)	Poor (1 bit/cell)	Good (2 bits/cell: expt. 3 bits/cell: sim.)

Fig.11. Benchmarking of various MLC technologies. Ferroelectric superlattice emerges as one promising candidate for MLC eNVM application.