Fundamental Understanding and Control of Device-to-Device Variation in Deeply Scaled Ferroelectric FETs

Kai Ni[1], Wriddhi Chakraborty[1], Jeffrey Smith[1], Benjamin Grisafe[1], and Suman Datta[1]
[1]University of Notre Dame, Notre Dame, IN; Email: kni@nd.edu

Abstract: In this work, we present a comprehensive Kinetic Monte Carlo (KMC) modeling based statistical framework to evaluate the device-to-device variation of thin-film HfO_2 ferroelectric FET (FeFET). We conclude that the closing of the memory window in a FeFET array with device scaling can be attributed to: 1) limited number of domains; 2) variation among domains; 3) intrinsic stochasticity of individual domain switching. To enable further scaling of FeFET, co-optimization approaches from material, process, and device operation to control variation are proposed: i) increase the number of domains through material/process optimization (e.g. decrease of deposition temperature, etc.); ii) improve the uniformity of domains (e.g. minimizing the domain size variation and defect distribution, etc.); iii) increase the pulse amplitude/width to ensure deterministic switching of individual domains.

Introduction: FeFETs based on ferroelectric HfO_2 are a promising candidate for embedded nonvolatile memory (NVM), due to its CMOS compatibility, scalability, and energy efficiency. Successful integration of HfO_2 FeFET within an industrial scale advanced CMOS technology platform has been demonstrated[1]. Promising as it is, several key challenges remain. The first challenge is to mitigate the electron and hole trapping during write to improve the endurance[2]. The second challenge is to reduce the write voltage to a logic-compatible level[3]. This work focuses on the third challenge, the device-to-device variation, which increases with the FeFET device size scaling, ultimately causing the memory window collapse for aggressively scaled FeFET arrays[1] (Fig.1). In this work, we develop a comprehensive statistical framework to model and evaluate the variation of FeFET with the scaling, supported by experimental data, which guides co-optimization of material /process conditions and prescribes operating bias conditions.

KMC Simulation Framework: The ferroelectric is assumed to contain N elementary domains (Fig.3) with each domain switching independently. The switching of an individual domain between two polarization states is modeled with a master equation[4] (Fig.2). The original framework treats the switching between two stable states as a double-well potential separated by an energy barrier[4]. This is not applicable for thin film HfO_2[5], where polarization switching is nucleation limited. Therefore, the framework is adapted to capture the nucleation-limited switching dynamics by adjusting the transition rates. As the single domain switching is stochastic[5], the KMC simulation is a natural means to study the variation of FeFET (Fig.3). A Gaussian distribution of the activation field (E_a) is assumed. The equations governing the voltage division and the charge balance between the ferroelectric and the MOSFET are solved to obtain the ferroelectric electric field (E_{FE}), which is then used to calculate the switching rates. Thereafter, the KMC algorithm is executed.

Variation in Metal-Ferroelectric-Metal (MFM) Capacitor: We validate this modeling approach using measured data from an MFM capacitor. Fig.4 shows the calibration of the Q_{FE}-V_{FE} saturation loop. The switching dynamics under various applied voltages are captured by the KMC model (Fig.5a). When the number of FE domains decreases from 1,000 to 20, the switching occurs in discrete steps and the variation increases significantly (Fig.5b). This is due to the intrinsic stochastic nature of individual domain switching. The effect of single domain switching (of the order of $1/N$) now becomes significant. Fig.6 quantifies the variation in the P_{FE} after program/erase operation. With the reduction in the number of domains with device area scaling, the program and erase states start to overlap, which leads to memory window collapse.

Variation in Ferroelectric FET: The switching dynamics is measured in a state-of-the-art 28nm high-κ metal gate FeFET technology[2] (Fig.7a). The device features an 8nm thick doped HfO_2 with 1nm SiO_2 interlayer (Fig.7c). The typical grain size in 10nm HfO_2 is around 10-15nm[6], which means that there are approximately 1,000 switching elements for a 500nm by 500nm FeFET. The KMC model accurately reproduces the measured switching dynamics (Fig.7b). The evolution of the polarization charge during the write operation (Fig.7d) explains how the device-to-device variation increases as the number of domains decreases. The deviation of P_{FE} from its expectation value originates from single domain switching events, which have a strong impact when limited number of domains participate in the switching process. The measured (Fig.7e) and modeled (Fig.7f) variation show the collapse of the memory window as a function of FeFET device dimensions.

Control of Variation in Ferroelectric FET: With the insights from the modeling, several approaches controlling the variation in FeFET are explored. The first is to increase the number of ferroelectric domains through careful material /process co-optimization (Fig.8a). Effective fabrication techniques, such as lower deposition temperature[7], enhanced electrode surface roughness, higher annealing temperature[8], can be applied to reduce the grain size. The second method is to increase the domain uniformity. The decrease of the variance of the activation field improves the memory window (Fig.8b), which may be achieved by minimizing the size variation, defect distribution, etc. among the domains[5]. Another set of techniques involve the optimization of the write pulses, where increasing the pulse width (Fig.8c) or pulse amplitude (Fig.8d), decreases the measured variation. Through modeling, we confirm this dependence (Fig.8e and Fig.8f), as the switching becomes less stochastic under high E_{FE} or large pulse width[5]. As the write in FeFET is voltage-driven with excellent energy efficiency, the increase in pulse amplitude or pulse width may be an acceptable compromise for improving variation.

Conclusion: We have developed a KMC statistical modeling framework capable of quantifying the device-to-device variation of FeFET and explaining the collapse of the memory window with the scaling of device dimensions. Several approaches to control device variation have been proposed, thereby providing guidelines for the future effective scaling of FeFET for nonvolatile embedded memory applications.

References: [1] S.Dünkel et al., *IEDM* 2017; [2] M.Trentzsch et al., *IEDM* 2016; [3] K.Ni et al., *IEDM* 2018; [4] M.Vopsaroiu et al., *PRB* 2010; [5] H. Mulaosmanovic et al., *Appl. Mater. Interfaces* 2018; [6] M.H.Park et al., *APL* 2014; [7] K.D.Kim et al., *J. Mater.Chem.C.* 2016; [8] J.Y.Lee et al. *CAP* 2017

Acknowledgment: We would like to thank Globalfoundries Dresden, Germany for providing 28nm FeFET test devices. This work was supported in part by the Semiconductor Research Corporation (SRC) and DARPA.

Motivation: Understanding the FeFET Variation Using Kinetic Monte Carlo (KMC) Simulation

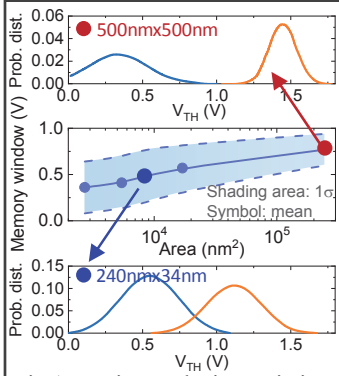

Fig.1. Device-to-device variation increases with scaling of FeFET[1].

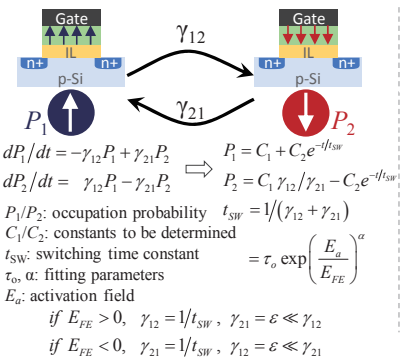

Fig.2. Rate equation modeling of polarization switching in individual domain.

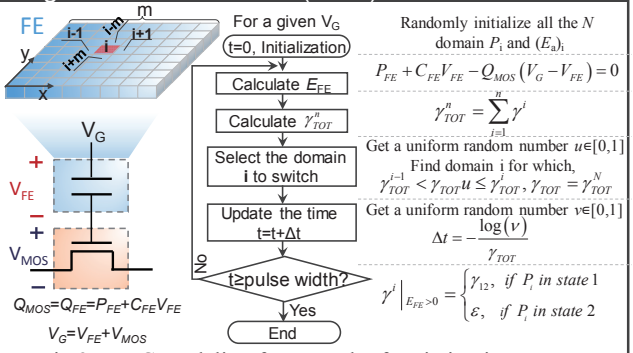

Fig.3. KMC modeling framework of variation in FeFET. FE contains N domains, whose switching dynamics is modeled with rate equation shown in Fig.2.

Variation in Metal-Ferroelectric-Metal Capacitor

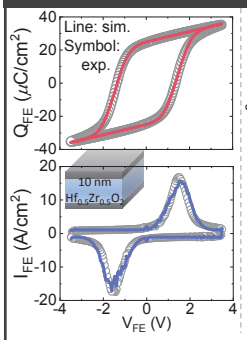

Fig.4. Calibration of Q_{FE}-V_{FE} loop in a $Hf_{0.5}Zr_{0.5}O_2$ (10nm) MFM capacitor.

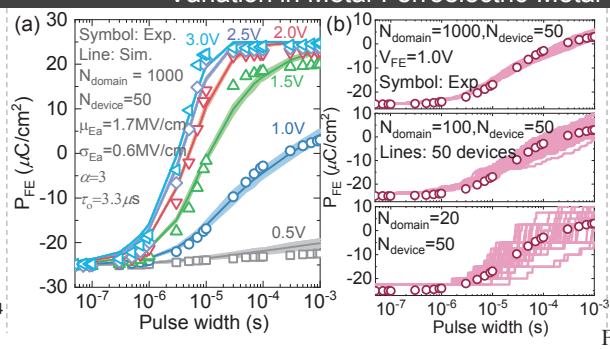

Fig.5. (a) Calibration of switching dynamics in a MFM capacitor; (b) switching dynamics variation increases when the domain number decreases from 1,000 to 20.

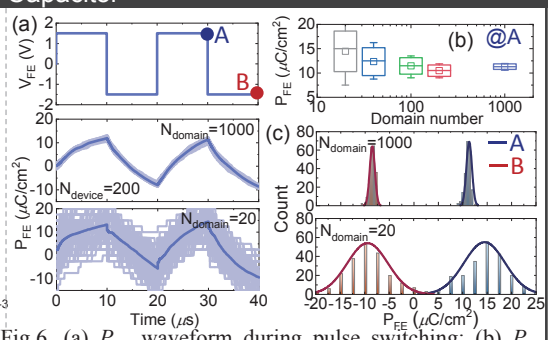

Fig.6. (a) P_{FE} waveform during pulse switching; (b) P_{FE} shows a wider distribution for reduced number of domains; (c) P_{FE} distribution after program/erase overlaps when the domain number reduces from 1,000 to 20.

Variation in Ferroelectric FET

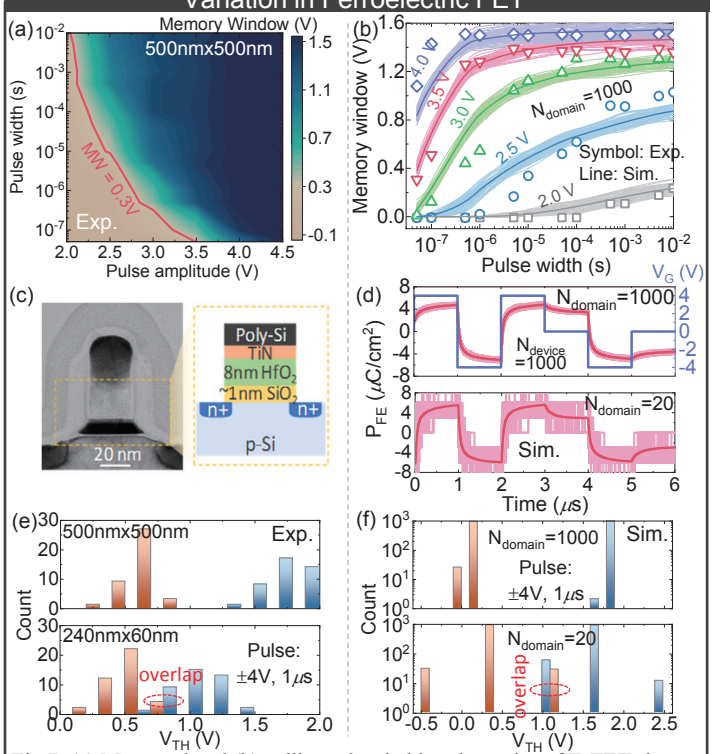

Fig.7. (a) Measured and (b) calibrated switching dynamics of FeFET shown in (c). (μ_{Ea}/σ_{Ea}=2.2/0.4 MV/cm, α=3.6, τ_o=100ns) (d) P_{FE} waveform variation in 1000 devices increases with reduced number of domains; (e) Measured V_{TH} distributions overlap when FeFET scales down; (f) Simulated V_{TH} distributions reproduce the observed phenomena.

Approaches: Variation Control in Ferroelectric FET

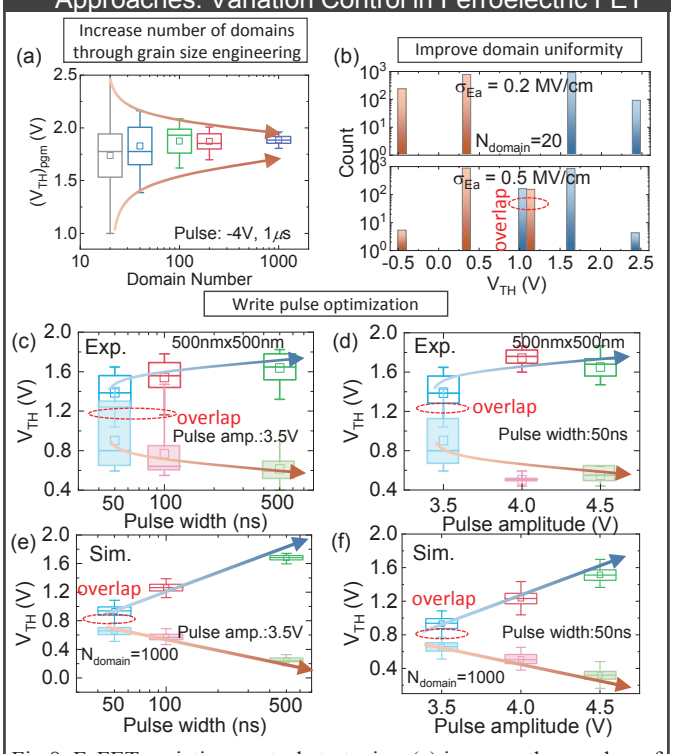

Fig.8. FeFET variation control strategies: (a) increase the number of domains through grain size engineering; (b) increase the uniformity among the domains (reduce variance of E_a); (c),(d) measured V_{TH} variation for different write pulse width & amplitude; (e),(f) simulated V_{TH} variation under various pulse width & amplitude.