SoC Logic Compatible Multi-Bit FeMFET Weight Cell for Neuromorphic Applications   
 K. Ni1, J. A. Smith1, B. Grisafe1, T. Rakshit2, B. Obradovic2, J. A. Kittl2, M. Rodder2 and S. Datta1 1University of Notre Dame, Notre Dame, IN, USA; 2Samsung Advanced Logic Lab, Austin, TX, USA; email: [kni@nd.edu](mailto:kni@nd.edu)

***Abstract*—**We demonstrate an SoC logic compatible ferroelectric-metal field effect transistor (FeMFET) digital 2-bit weight cell by monolithic BEOL integration of a ferroelectric (FE) capacitor with the gate of a conventional Si HK/MG MOSFET. Through optimization of the area ratio between the FE capacitor and the MOSFET, we show: 1) program/erase write voltages can be scaled down to logic compatible level, ±1.8 V, simplifying write circuitry; 2) write speed of 100ns; 3) write endurance >1010 cycles without degradation due to elimination of charge trapping in FE; 4) 2 bits/cell achieving software levels of accuracy for inference on MNIST training database; 5) state retention approaching 104 s for a depolarization field of 0.3 MV/cm; 6) Multi-port (independent read and write) operations.

**INTRODUCTION**   
 Deep neural networks (DNNs) are ubiquitously applied in tasks such as image recognition, natural language processing and self-driving cars (Fig.1(a)). Various neuromorphic accelerators based on the cross-bar architecture with nonvolatile weight cells have been investigated (Fig.1(b)). Neuromorphic accelerator performance is improved via elimination of data movement between the processing unit and off-chip memory by storing the synaptic values locally (on-chip) in the weight cells and performing the multiply-accumulate (MAC) operation in the analog domain. The performance metrics of the on-chip weight cells vary depending on whether an in-situ training is supported or not. To support in-situ training, a weight with high bit precision is necessary to handle the incremental weight update during training, and complex peripheral circuitry is required to implement back-propagation. However, for inference only applications, low precision weight cell can achieve near-analog levels of accuracy. Therefore, by relegating the training to the cloud domain, we focus on a logic compatible embedded weight cell capable of excellent inference performance targeting edge devices [1].

The dynamics of multi-domain ferroelectric polarization switching in FeFET can be utilized for an analog weight cell [2]. However, all reported FeFETs suffer from high write voltages (±4.0 V) and limited endurance of ~105 cycles (Fig.2) [3]. This is because of the non-uniform field distribution across the gate stack, with the majority of the voltage drop across the interlayer (IL) and the semiconductor channel [4]. This increases the write voltage and degrades the reliability due to enhanced charge trapping in the FE [4]. This inefficiency can be overcome by integrating the FE capacitor in the back-end, electrically coupled to the gate of a conventional MOSFET (Fig.1(d)), rather than direct integration within the MOS gate stack. By independently optimizing the *A*FE/*A*MOS (AR: area ratio between MFM and MOSFET) of the metal-ferroelectric-metal (MFM)

capacitor and the MOSFET, the voltage drop across the FE can be maximized, which in turn reduces the write voltage, eliminates charge trapping and improves endurance(Fig.2). In this work, we experimentally demonstrate a 2bit FeMFET weight cell that has a logic compatible write voltage of 1.8V and endurance cycle >1010, and achieves the same inference accuracy as floating point weight.

**II.DEVICE FABRICATION PROCESS**   
 The key fabrication steps are shown in Fig.3(a). Boron doped (3x1017 cm-3) p-type silicon wafers are used to achieve enhancement mode MOSFET operation. After alignment mark etch, source/drain ion implantation and junction activation follow. SiO2 chemical oxide passivation and thermal ALD of 10nm thick HfO2 with 50nm thick tungsten (W) metal electrode completes the gate stack. CVD SiO2 is deposited for MFM capacitor isolation and a W via process is used to connect to the MOSFET gate. W bottom electrode, 10nm thick PEALD Hf0.5Zr0.5O2 deposited at 300°C and W top electrode forms the MFM stack before 600°C FE crystallization anneal. CVD SiO2 is used for MFM isolation. The FeMFET device fabrication is completed by via etch for connection to W top electrode and source/drain and Ti/Al contact metallization and anneal to form low-resistivity Ti-Silicide contacts.

Top and tilted view SEM images (Fig.3(b)) show the top MFM electrode and the internal gate, which are used to probe the C-V, I-V characteristics of the MFM and MOSFET individually. A zoomed-in view shows the W via that connects the MFM cap to the underlying MOSFET gate. Cross-section TEM images (Fig.3(c)) of FeMFET clearly show the connections between the MFM and MOSFET gate, highlighting the monolithic integration of the MFM with the MOSFET. The HRTEM images reveal poly-crystalline nature of the FE HZO in the MFM.

**III.RESULTS AND DISCUSSION**   
*A.FeMFET Operation: Program, Erase and Read*   
 The fundamental device operating principle is first demonstrated by electrically coupling a discrete MFM cap to the gate of a MOSFET (Fig.4(a)). The *Q*FE-*V*FE characteristics of MFM cap show the transition from saturated to non-saturated hysteresis loops as the *V*FE sweep range decreases (Fig.4(b)). When connected with a MOSFET, the DC *I*D-*V*G characteristics exhibit large hysteresis window (~ 2.0V) for a *V*G sweep voltage of -1.8V to 1.8V. The lowering of *V*G sweep range reduces the current ratio between the two FE polarization states, which corresponds to minor loop operation of the ferroelectric.

To quantitatively understand the program/erase operation of the FeMFET, a model is developed by introducing the AR design parameter to the Preisach FeFET model [5]. Fig.5(a) shows the bias conditions to operate the FeMFET weight cell.

The select transistor eliminates the disturb to half-select cells and pass the write voltage to the FeMFET gate. Fig.5(b) shows the waveform for the cell write and read operations. After erase, the internal gate voltage, *V*MOS, remains high, whereas it becomes low after program. *V*MOS after program/erase corresponds to the two intersection points between the MOSFET loadline and the FE subloop during program/erase cycle (Fig.6(a)). Their separation determines the current ratio between the two states and the available cell read window. By decreasing the AR, the MFM capacitance is reduced and the voltage across it is increased (Fig.6(b)). This mitigates the undesirable voltage division issue encountered in FeFET and reduces the write voltage from 4.0V to 1.8V. However, too small of an AR decreases the total polarization charge (Fig.6(c)), which in turn decreases the read window, leading to an optimal

approximately a linear behavior (Fig.11(b)), critical for neuromorphic inference application. Higher *V*DS and shorter channel length can further enhance the FeMFET conductance values and accelerate the read speed.

The deviation of the FeMFET channel conductance values or weights from ideal linearity and its impact on inference accuracy is illustrated in Fig.12(a). A nonlinear fit is performed on the measured weights and used for hardware aware quantization. Fig.12(b) shows the inference error incurred for the MNIST digit recognition task compared with a floating point weight, resulting from the weight quantization effect (1-bit, 2-bits, and 4-bits). It is clear that 2bits/cell is suffice to achieve analog-level accuracy, albeit with increase in number of the hidden layer neurons.

Fig.13(a) shows the FeMFET array architecture for

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| AR | range | maximizing | memory | window | (Fig.6(d)). | inference. The pseudo-crossbar architecture performs the MAC |

Furthermore, unlike other nonvolatile memories such as STT-MRAM or RRAM, FeMFET can be operated as a multi-port device due to the separation of read and write path.

*B. Integrated FeMFET Device*   
 The integrated FeMFET device characteristics are shown in Fig.7. The MFM capacitor *C*FE-*V*FE curve exhibits double peaks, indicating ferroelectric polarization switching. The MOSFET characteristics are shown in Fig.7(b). The *C*MOS-*V*MOS measurements indicate that the capacitance ratio *C*FE/*C*MOS is 1/8, which increases the voltage drop across the MFM cap, as shown in Fig.6. The FeMFET *I*D-*V*G characteristics are similar to the discrete MFM and MOSFET measurement. It exhibits ~1.5V hysteresis window for *V*G range of -1.8V to 1.8V. Fig.8 shows measured program/erase/read operation waveform of FeMFET with logic compatible write voltages.

The thick gate oxideand small voltage drop (due to small AR) in the MOSFET, ensures that there is no charge trapping in the FE or the internal floating node, thereby markedly enhancing the FeMFET endurance[4]. Fig.9(a) shows that the FeMFET endurance exceeding 1010 cycles without degradation. Fig.9(b) shows the read current, *I*D, corresponding to the programmed/erased state as a function of write pulse width for write voltages of 2V, 1.8V, and 1.5V. FeMFET achieves fast write operation (of 100ns for a program/erase voltage of 1.8V, while maintaining read current ratio of >10). The retention state of FeMFET is shown in Fig.10. During retention after erase, the internal node voltage *V*MOS is approximately the MOSFET *V*TH to have enough sensed current. This would result in voltage drop of –*V*TH across the MFM cap, which acts as depolarization field, decreasing the FE polarization(Fig.10(a)). Fig.10(b) shows the measured polarization retention under externally applied depolarization field. It suggests that, in order to reduce the polarization loss, it is necessary to design the underlying MOSFET *V*TH close to 0V.

*C.FeMFET Multi-Bit Weight Cell*   
 In this section, we demonstrate multi-bit programming of the weight cell using partial polarization switching in the MFM cap[2]. Fig.11(a) shows a well-tempered distribution of the cycle-to-cycle variation for the four distinct *I*D levels (2 bits/cell) for FeMFET. By tuning the erase voltage amplitude, four *I*D levels are achieved. The *I*D-*V*DS measurement after write indicates that the programmed conductance values follow

operation in the analog domain. Two FeMFETs are used to represent positive and negative weights, respectively, which can be achieved by current subtractor (e.g. current mirror). The inference accuracy of the FeMFET array matches that of the floating point weight. These results highlight the promising application of the multi-bit FeMFET weight cell for neuromorphic inference. Fig. 14 benchmarks the FeMFET with other types of embedded nonvolatile memory candidates [6][8],. FeMFET has matched or superior performance in almost all the aspects of write voltage, write energy, latency, and cycle endurance, etc. Thus, FeMFET is a promising weight cell candidate for low power, high performance neuromorphic inference applications for edge devices.

**IV.CONCLUSIONS**   
 In summary, we demonstrate an SOC-logic-compatible multi-bit FeMFET weight cell via monolithic BEOL integration of a MFM capacitor with the gate of front-end MOSFET. We show that, by optimizing the area ratio between MFM and MOSFET, we can reduce the program/erase voltage to logic-compatible level of <1.8V, improve write latency to <100ns, improve the endurance beyond 1010 cycles. Further, 2bits/cell FeMFET weight cell is demonstrated with high linearity and has no loss in inference accuracy. These characteristics render FeMFET a promising candidate for implementing low power, high performance inference task on edge devices.

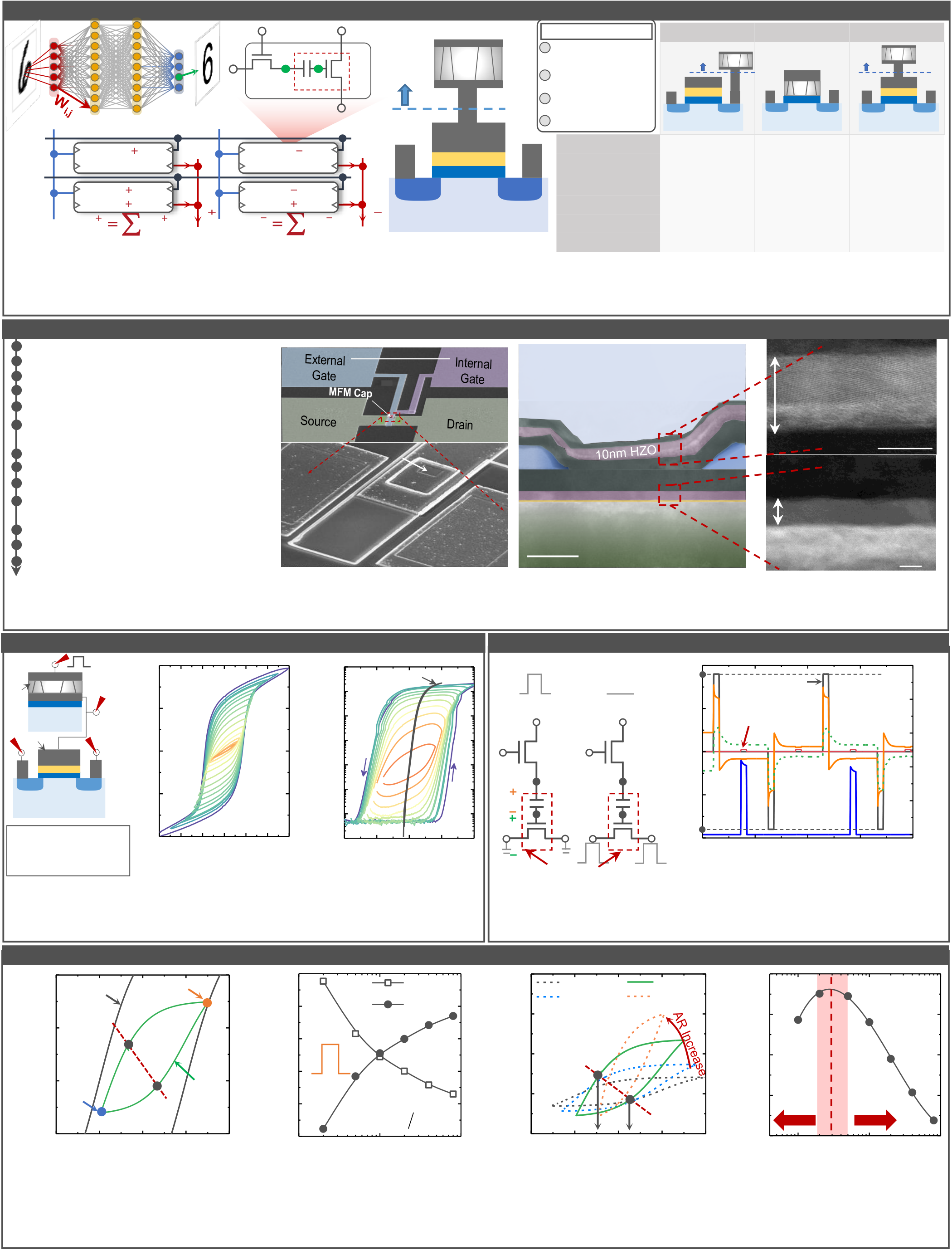
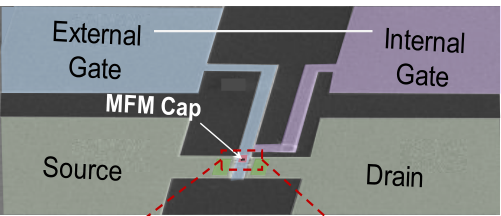
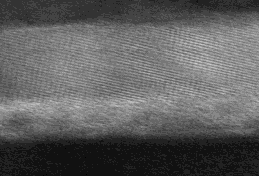
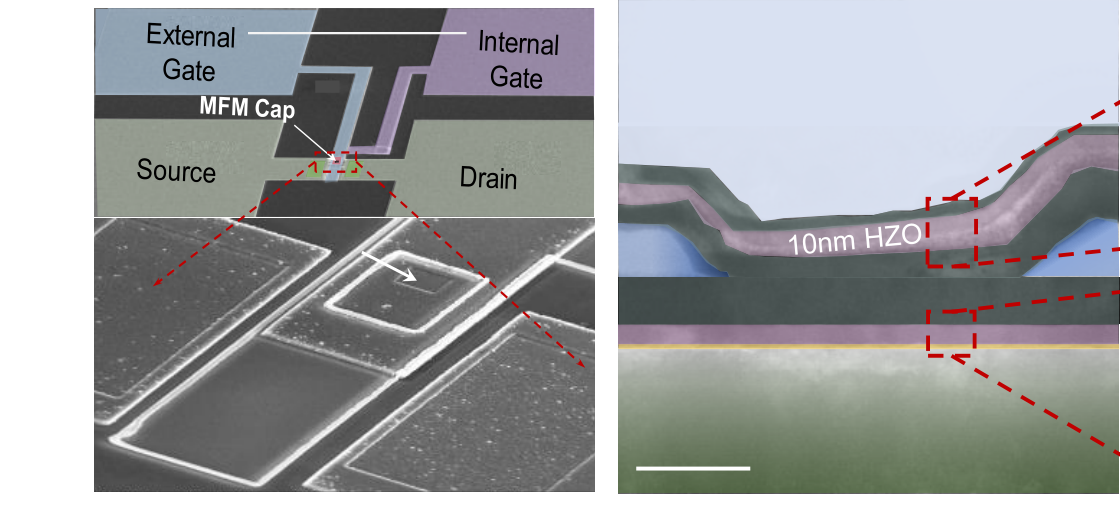
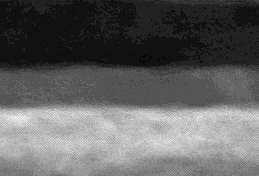
**ACKNOWLEDGEMENT**   
 This work supported in part by ASCENT, one of six centers in JUMP, Semiconductor Research Corporation (SRC) program sponsored by DARPA

REFERENCES   
[1]B. Obradovic, et al., “A multi-bit neuromorphic weight cell using ferroelectric FETs, suitable for SoC integration,” JEDS 2018.

[2]M. Jerry, et al., “Ferroelectric FET analog synapse for acceleration of deep neural network training,” IEDM 2017.

[3]S. Dunkel et al., “A FeFET based super-low-power ultra-fast embedded NVM technology for 22nm FDSOI and beyond,” IEDM 2017   
[4]K. Ni et al., “Critical role of interlayer in Hf0.5Zr0.5O2 ferroelectric FET nonvolatile memory performance,” TED 2018.

[5]K. Ni et al., “A circuit compatible accurate compact model for ferroelectric- FETs” VLSI 2018   
[6]D. Takashima et al., “A 100MHz ladder FeRAM design with capacitance- coupled-bitline cell,” JSSC 2011   
[7]C. Park et al., “Systematic optimization of 1Gbit perpendicular magnetic tunnel junction arrays for 28nm embedded STT-MRAM and beyond,” IEDM 2016 [8]S. R. Lee et al., “Multi-level switching of triple layered TaOx RRAM with excellent reliability for storage class memory,” VLSI 2012



Motivation: FeMFET as Weight Cell for Neural Network Inference   
(a) SEL (c) BL (d) FeMFET G **Decreased AFE/AMOS**  **FeRAM**  **FeFET**  **FeMFET**  WL

(GDS)i,j as Wi,j   
 G Gint

SL Back end FE area: AFE FE No charge trapping in FE Write Voltage decreases EFE increases, EIL decreases   
 S

n+

**Back end**

interlayer

high 𝜅

G

n+

FE

D S

n+

interlayer

FE

G

n+

D

**Back end**

S

n+

interlayer

high 𝜅

FE

Gint

G

n+

D

FeMFET

Vin,i+1

(b) Vin,i SELj+   
 SEL   
 SEL   
 WL   
 WL   
 I

j

W   
   
W 

i

,i j

1, j W V   
 SL   
 SL   
 BL   
 BL

i j in i

**jI**

SELj-  
 SEL   
 SEL   
 WL   
 WL   
 I

j

W   
   
W 

i

,i j

1, j W V   
 BL   
 SL   
 SL   
 BL

i j in i

**jI**

MOSFET area: AMOS   
S n+   
 Interlayer high 𝜅   
 Gint

n+   
 D   
 Improved endurance

**Charge trapping**   
**Read scheme**   
**Write voltage**   
**Endurance**   
**AFE/AMOS**

Destructive

Good   
None   
N/A

3 V   
 Non-Destructive

Significant   
 Bad   
 4 V   
 1 Non-Destructive Tunable

Good   
None   
1.8 V

Fig.1. (a) Inference in deep neural network requires dense weight cell; (b) i i **Multi-bits**  Bad Good Good Fig.2. Advantages of FeMFET compared with other FeMFET pseudo-crossbar array for inference accelerator; (c) FeMFET weight cell circuit diagram, including a select transistor and FeMFET weight cell; (d) types of ferroelectric memories, including write voltage FeMFET structure, including a MOSFET and back-end MFM capacitor. scaling and endurance improvement.

Integrated FeMFET Fabrication Process and Physical Structure   
P-Si (3x1017cm-3 B-doped) (a) (b) 100 μm (c) W(5nm) Alignment Mark Etch   
 HZO(10nm) ALD 10nm HfO2/W Gate Stack Source/Drain Formation Ti/Al via   
CVD SiO2 deposition/Gate Via Etch Sputtered W Bottom Electrode for connection to MFM   
 5nm **MFM-Gate Connection**  W W(50nm) Sputtered W Top Electrode for MFM PEALD 10nm Hf0.5Zr0.5O2 **MFM Cap**  SiO2 W Via (MFM to Gate) SiO2   
CVD SiO2 deposition/Via Etch for **MFM Top Electrode Connection** RTA for HZO Crystallization Process flow and device structure SiOx HfOx HfOx:9.1nm W(50nm)

Via Etch for Source/Drain Contacts SiOx:0.9nm Gate/Source/Drain Metallization (Ti/Al) Contact anneal (350°C, 15mins) 50 nm Si Substrate Si 5nm Fig.3. (a) Key processing steps in the fabrication of FeMFET; (b) SEM images of one FeMFET device (the zoomed in image shows the MFM capacitor); (c) TEM images of the whole stack (W/HZO/W MFM, W via, and W/HfO2/SiO2/Si MOSFET). High-resolution TEM images show poly-crystalline HZO in MFM capacitor.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Proof of Concept: Discrete MFM Cap + MOSFET  (a) **VG**  TE   (b) 40  MFM: 10mx10m (c) 10-4  MOSFET only  AFE BESiO2 FE  p-Si **VG,int**   20   10 KHz 10  10  -5  -6   2mx1m  2) QFE (C/cm  **VS**   AMOS  S interlayer  high 𝜅  G **VD**  D   0 10  10  -7  -8  ID (A)  n+ n+  -20 10-9  p-Si  10-10 | | | | | | | 2 |
| AFE/AMOS is not optimal due to the setup parasitic. | -40  -3 -2 -1 0 1  VFE (V) | 2 | 3 | 10-11  -2 | -1 | 0 1  VG (V) |

Fig.4. (a) Discretely connected MFM cap and MOSFET for proof of

FeMFET Program/Erase Operation

Erase (a) (b) 2 5

Program VG-VSL Simulation

1.8 V   
SEL:   
WL: 1.8 V

0.7 V   
SEL: WL: 0 V 1 VBL-VSL

VFE   
 4

3 IBL\_to\_SL (A) Voltage (V)   
 0

VG: VG: VMOS 2

~1.8 V ~0 V

VFE -1 1

SL: BL: SL: IBL\_to\_SL

VMOS 0 V 1.8 V 1.8 V -2 0

0 10 20 30 40

BL:

0 V FeMFET Time (s)

Fig.5. (a) Bias conditions during program/erase operation of a

concept; (b) MFM *Q*FE-*V*FE loops for different *V*FE ranges; (c) *I*D-*V*G FeMFET weight cell; (b) Simulated waveform of write and sensing

characteristics showing hysteresis window ~2V for VG (-1.8V to 1.8V). operation on the FeMFET cell.

FeMFET Operation Principle and Modeling

(a) 20 Saturation Erase peak (b) 1.4

VFE (c) 150 AR=0.01 AR=0.05 (d) 0.4 Optimum

MOSFET loop 1.2 VMOS 100 AR=0.02 AR=0.2

0.3   
 AR ratio

(VFE)pgm-(VFE)ers 2) PFE (C/cm   
 Voltage (V)   
 10

0   
 @VG=0 V

Program   
loadline

peak Subloop   
 1.0

0.8

0.6   
 1.8 V

50

0   
@VG=0 V   
MOSFET

QMOS 0.2

0.1 decrease charge   
 FE

hysteresis

decrease   
 Subloop   
PFE (fC)

-10

-1.0 -0.5 0.0 0.5 1.0 0.4

0.01   
 AR

0.1  
A FE A MOS

1  
 -50

-1.0   
 (VFE)ers

-0.5 0.0   
 (VFE)pgm

0.5 1.0 0.0

0.01 0.1 1

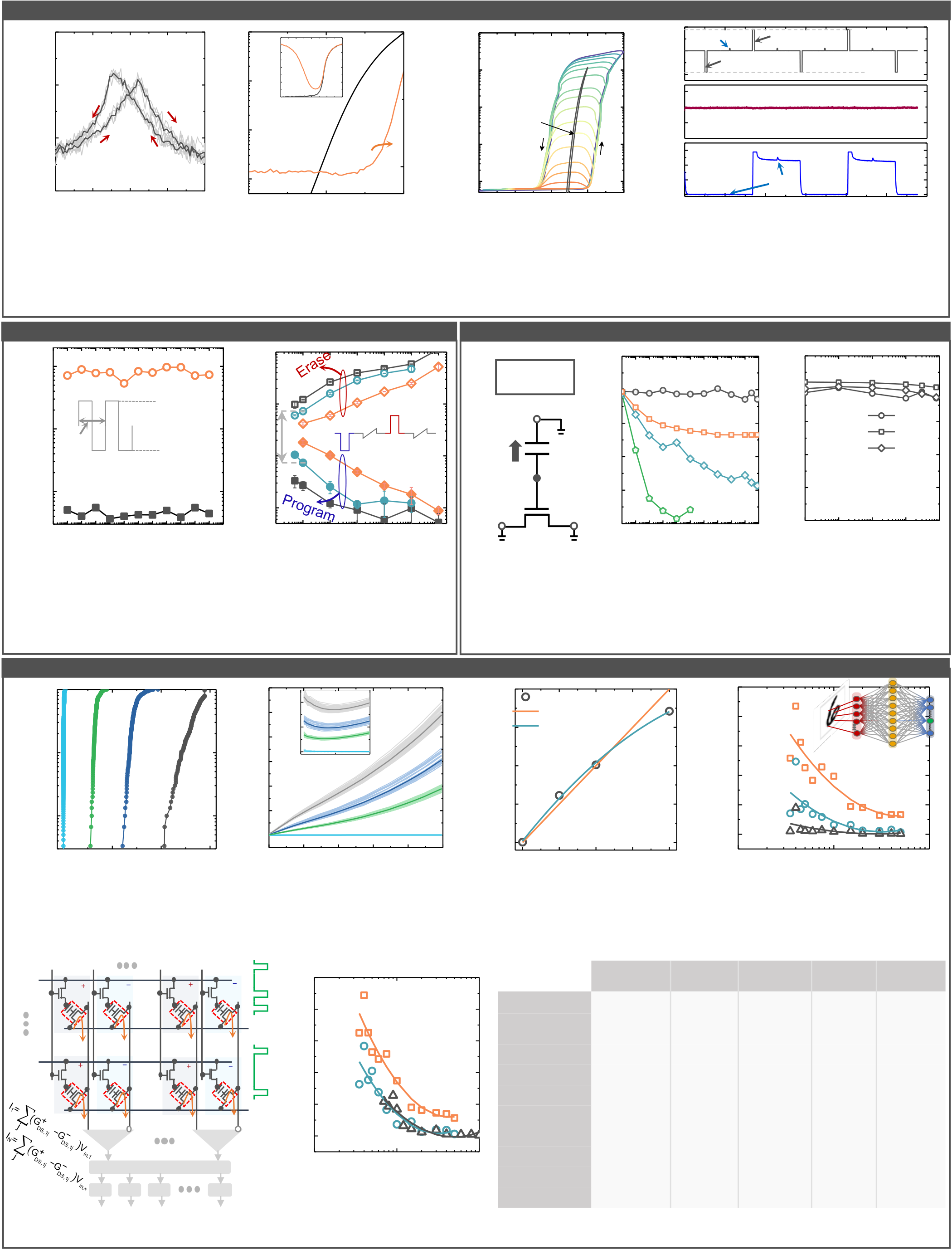
VFE (V) AR VFE (V) AR

Fig.6. (a) *P*FE-*V*FE hysteresis loops for several program/erase cycles. The saturation loop is shown in black line. The MOSFET loadline at

*V*G=0 V is shown in red; (b) *V*FE increases (*V*MOS decreases) with the decrease of *A*FE/*A*MOS ratio; (c) *P*FE-*V*FE subloop modulation as a function

of the area ratio (intersection points between the *P*FE-*V*FE loop and *Q*MOS-*V*FE loadline, red dashed line, sets the memory window); (d) Memory

window vs. area ratio characteristic shows a peak (small ratio reduces the total charge and large ratio reduces FE voltage)



FeMFET Device Characteristics

(a) 240

220

200   
 30 kHz MFM area: 1m x1m (b) 10

10

10  
 -5

-6

-7 MOSFET:   
 15mx0.8m   
 2.0   
 1.5   
 1.0   
 0.5   
 0.0-2 -1 0 Split C-V

VMOS (V)  
 30 kHz

1 2   
 30 (c) 10

10

10

10  
 -4

-5

-6

-7   
FeMFET:   
MFM: 1mx1m   
MOSFET:   
15mx0.8m

MOSFET   
only  
 -2.0   
 2.0

0.0

0.1

0.0

1.2   
 Read

Program   
 Erase Experiment VG (V) VD (V) ID (A)   
 Capacitance (pF)   
CFE (fF)   
ID (A) IG (pA)   
 ID (A)   
 ID (A)

180  
 -2 -1 0 1 2   
 10-8

0.0   
 CFE/CMOS

0.5 1.0   
 20   
 10-8

-2   
VDS=50 mV

-1 0 1 2 0.8   
 0.4 0.0 0 200 400   
 Read current

600 800 1000 1200 VFE (V) VMOS (V) VG (V) Time (s)   
Fig.7. (a) Measured MFM cap *C*FE-*V*FE characteristics shows good ferroelectric behavior; (b) Fig.8. Measured waveform of program/erase

MOSFET *I*-*V* and *C-V* characteristics. *C*FE/*C*MOS=1/8 implying efficient voltage dividing; (c) operation and sensing of the memory state on a

*I*D-*V*G characteristics of FeMFET for different *V*G sweep ranges, showing ~1.5V hysteresis FeMFET device. ±1.8 V write voltage is

window. The reduction of *I*D with *V*G range corresponds to FE operating on minor loops. applied. Successful memory operation is shown.

Endurance and Speed Retention   
(a) 10-6 (b) 10-6 (a) Retention (b) 25 MFM: 1 mx1 m (c) 25

Erase after erase 20 25 C EDep=0 20

10-7

1*μ*s   
 +1.8 V

-1.8 V   
 10

10  
 -7

-8   
 >10x **EDepolarize**  VG=0 V 15

10   
 15

10 25 C 80 C 150 C 2) Pr (C/cm   
 2) Pr (C/cm   
ID (A)

10-8 FeMFET Program 10-9 VMOS~VTH 5 EDep=-0.6 MV/cm 5 EDep=0

10-1 10 1 10 3 10 5 10 7 10 9 10 11 10-8 10-7 10-6 10-5 10-4 10-3 10-2 10 0-6 10-4 10-2 10 0 10 2 10 4 0 10 0 10 1 10 2 10 3 10 4

Cycle number Pulse width (s) Retention time (s) Retention time (s)

Fig.9.(a) Endurance of FeMFET device. No degradation is observed after 1010cycles due to small write voltage (1.8 V) and minimized charge trapping; (b) *I*D, as a function of write pulse width. *I*D ratio >10 between program/erase state under ±1.8V, 100ns is achieved.

Fig.10.(a) After erase, finite depolarization field is present in the MFM cap, which could cause retention loss; Polarization retention of MFM cap under (b)externally applied depolarization field at 25 °C and (c) different temperature. Properly designed *V*TH close to 0 V is necessary.

FeMFET Weight Cell and System

10 0

(a)   
 12

20 (b) (a) 20 Measured Data 50

(b) Test quantization error (%) Meas. conductance (S)   
Cumulative probablity   
 GDS (S) Test quantization error (%)   
 10-1   
 250 cycles 10

8

6

4 50 cycles   
10

0

0.0

VDS (V)   
 0.5 11

10   
 15

10  
 Ideal Linear Weight

Fitted Weight   
 40

30

20   
 **1-bit**

**2-bit**   
 Ideal weight

quantization 784 n 10 Program: -1.8 V 00   
 ID (A)   
 Erase: 1.7 V   
10 Erase: 1.6 V 01   
 Erase: 1.8 V

10-2 2

0 -1.8 V   
 01

00   
 5

0   
 10

0 **4-bit**

11

0.0 0.4 0.8 1.2 0.0 0.1 0.2 0.3 0.4 0.5 0 5 10 15 20 10 1 10 2 10 3

ID (A) VDS (V) Target conductance (S) Number of hidden neurons

Fig.11.(a) Cycle-to-cycle variation of four levels in FeMFET is well- Fig.12.(a) Comparison between the measured FeMFET weights and

behaved; (b) Measured *I*D-*V*D characteristics show small non- ideal linear weight; (b) Errors induced by quantization of the ideal

linearity. Large *I*D can be obtained with large *V*DS for fast read speed. linear weight (2 bits) is negligible compared with floating point

(a)

WL1   
 SEL1+

W   
 SEL1-

11 W 11   
 SELN+ SELN-

W 1N  W 1N  50   
 (b)

FeMFET   
 weight by increasing the number of hidden neurons.

**FeRAM**

**[6]**   
 **FeFET**

**[3]**  (this work)   
 **FeMFET**  **STT-MRAM**

**[7]**   
 **RRAM**

**[8]**

40 **1-bit**  weight **Structure**  1T-1C (drain) 1T 1T-1C (gate) 1T-1MTJ 1T-1R

BL1

30 quantization **Read scheme**  Destructive destructive Non-

destructive   
Non-

destructive   
Non-

destructive   
Non-

WLN

W N

1 W N 1 W NN W NN 

20   
 **2-bit**  **Write voltage**

**Write energy**  ~0.1 pJ   
 3.3 V

~0.1 pJ   
4.0 V

~0.1 pJ   
1.8 V

~5 pJ   
1.5 V

~ 10 pJ   
4 V

BLN 10 **4-bit**  **Multi-bit**  Bad Good Good Bad Good

SL1+ SL1- SLN+ SLN- **Read speed**  Medium Good Good Good Good

Current

Sub.

Current

Sub. 0 **Charge trapping**  None Significant None None None

I1

MUX   
 IN

10 1 10 2 10 3 **Endurance**  ~1014 ~105 >1010 ~1015 >108

ADC ADC ADC ADC Area (Arb. Unit)

**Multi-port**   
**Variation**  Good

No   
 Good

Yes   
 Good

Yes   
 Good

No   
 Bad

No

Fig.13.(a) FeMFET array architecture for inference; (b) Inference error due to FeMFET weight quantization compared with floating point weight.

Fig.14. Benchmarking FeMFET with other nonvolatile memories. FeMFET achieves excellent performance in all aspects.