In-Memory Computing Primitive for Sensor Data Fusion in 28 nm HKMG FeFET Technology K. Ni1, B. Grisafe1, W. Chakraborty1, A. K. Saha2, S. Dutta1, M. Jerry1, S. Gupta2, and S. Datta1 1University of Notre Dame, Notre Dame, IN, USA;   
 2Purdue University, West Lafayette, IN, USA; email: kni@nd.edu

***Abstract*—**In this work, we exploit the spatio-temporal switching dynamics of ferroelectric polarization to realize an energy-efficient, and massively-parallel in-memory computational primitive for at-node sensor data fusion and analytics based on an industrial 28nm HKMG FeFET technology [1]. We demonstrate:(i) the spatio-temporal dynamics of polarization switching in HfO2-based ferroelectricsunder the stimuli of sub-coercive voltage pulses using experiments and phase-field modeling; (ii) an inherent rectifying conductance accumulation characteristic in FeFET with a large dynamic range of *G*max/*G*min> 100 in the case of 3.0V, 50ns gate pulses; (iii) transition to more abrupt accumulation characteristics due to single/few domain polarization switching in scaled FeFET (34nm LG); and (iv) successful detection of physiological anomalies from real-world multi-modal sensor data streams.

**I.INTRODUCTION**   
 In the era of Internet of Things (IoT), the ubiquity of sensors and the continuous flow of data streams require sensor data fusion and processing in real time to discover patterns for predictive analytics. Statistical correlation detection is one approach to extract the global information distributed among the sensor signals. A key application of statistical correlation is the detection of physiological anomalies for distributed health monitoring in the form of wearable devices (Fig.1). Anomalies can be detected based on the correlation among different physiological signals (e.g., blood pressure, heart rate, blood oxygen saturation, etc.). The correlation between the sensors change in accordance with a person’s health condition (highly correlated when anomaly occurs). In order to process large streams of data in real time, an energy-efficient hardware accelerator suitable for edge device is necessary.

In the conventional Von-Neumann approach, sensor data is loaded from the memory buffer into the computation unit which computes the distance norm between the data stream (correlation as one type of distance norm). Clustering algorithms (such as K-means clustering) are then applied to classify the data based on the calculated distance. This approach involves substantial data movement between the memory and the arithmetic logic unit, limiting the throughput and energy efficiency (Fig.2(a)), often called the von Neumann bottleneck. On the other hand, the in-memory computing approach exploits the physical dynamics of emerging memories to perform the computation within the memory with co-located computation and storage (Fig.2(b)) [2]. In the case of correlation detection with FeFET, each signal is associated with an exclusive memory cell. The signal data is fed to an encoder, which translates the correlation among signals to the rate of the pulse

applied to the FeFET cells. As a result, only cells associated with highly correlated signals receive frequent pulse update, whose ferroelectric polarization accumulates and eventually switches the total polarization along with the FeFET conductance. Thus, a conductance read can differentiate between the correlated and uncorrelated signals, thereby detecting the anomaly. In this work, we apply statistical correlation detection to detect physiological anomaly from physiological signal streams based on an industrial 28nm HKMG FeFET technology, using the spatial-temporal polarization switching dynamics of ferroelectrics (FE).

**II.RESULTS AND DISCUSSION**   
*A.Spatio-Temporal Dynamics of Ferroelectric Switching in*

*MFM Capacitor*   
 The channel conductance accumulation in FeFET results from the polarization accumulation properties of the ferroelectric when exposed to identical consecutive pulses with amplitudes below the coercive voltage threshold. It involves two physical processes, domain nucleation and domain growth (Fig.3). When sub-coercive field pulses are applied to the FE, a critical number of pulses are required for the reverse domain nuclei to be stabilized. The polarization remains unchanged until the domain reaches the critical nuclei size, beyond which the domain growth happens. When integrated into the gate stack of the FeFET, the polarization accumulation in FE manifests as a shift in the device threshold voltage (*V*TH) and conductance accumulation characteristics.

We first measure the polarization accumulation in an MFM capacitor. Fig.4 shows the measured *Q*FE-*V*FE hysteresis loops, which exhibit a saturation loop and non-saturated minor loops. Fig.5(a-c) show the polarization accumulation with the number of consecutive pulses for varying pulse amplitude, pulse width, and inter-pulse delay. A modified pulsed measurement (PUND) is used to measure the switched polarization by the accumulation pulses. The results show that the reduction of either pulse amplitude or pulse width increases the required pulse number for domain nucleation [3]. Moreover, the reduction in the amount of switched polarization with longer inter-pulse delay suggests polarization relaxation between the pulses. This is because the interaction from the unswitched domains switches back the flipped domain, as discussed later.

To understand the voltage and time dependent domain nucleation and growth, we solve the time-dependent Landau-Ginzburg-Devonshire (LGD) equation in a phase-field framework (Fig.6) [4]. We consider the domain-domain interactions across the cross-section (x-y plane) via the Laplacian of polarization (*P*), which can be interpreted as an interaction field (*E*INTR). In addition, we account for the random fluctuations in *P* switching due to thermal processes using an

effective thermal field *E*ACCU, which increases with pulse number*.* The total effective field (*E*EFF) in the ferroelectric is a sum of the applied external field (*E*APP), and the internal *E*INTR and *E*ACCU. The latter two fields are self-consistently updated in our simulations at each time step as polarization switching causes field redistribution (Fig.6). The interactions between these different terms leads to domain nucleation and growth.

The phase field model is calibrated with the measured *Q*FE-*V*FE hysteresis loop showing an acceptable match (Fig.7). Simulated polarization accumulation in a 200nmx200nm MFM capacitor (10000 domains) is shown in Fig.8. The simulation reproduces the measured accumulation characteristics, such as the delayed onset of polarization increase with the reduction in pulse amplitude and pulse width, and the increase in inter-pulse delay. Fig.9 shows the simulated spatial-temporal evolution of polarization distribution in the FE for different pulse numbers. It exhibits two stages for polarization accumulation. In the initial stage, when the reverse domain volume is smaller than the critical size (10nmx10nm in simulation), increasing number of pulses yields larger thermal fluctuations due to dissipative processes associated with polarization switching. This raises *EACCU* and *EEFF*, increasing the probability of reverse domain nucleation and stabilization. The second stage involves domain growth from the nucleation site, which propagates to the whole grain until a grain boundary is encountered. This process is mainly driven by *E*APP and *EINTR*. The polarization relaxation is attributed to *E*INTR from the surrounding unswitched domains as polarizations from different domains tend to align with each other. For scaled MFM cap (20nmx20nm, 100 domains), the polarization accumulation exhibits abrupt switching due to small number of domains in switching process (Fig.10).

*B.Conductance Accumulation Characteristics of FeFET*  Conductance accumulation and threshold behavior is characterized on an 28 nm HKMG FeFET platform [1]. The device cross-sectional TEM images show a poly-Si/TiN/Si:HfO2/SiON/p-Si gate stack (Fig.11(a)). Conductance response to input gate pulses are characterized for FeFETs with W/L=450/450nm (Fig.11(b-f)) and W/L=72/34nm (Fig.11(g-k)). Figs.11(b-c, g-h) show the erased state *I*D-*V*G characteristics with progressively increasing erase voltage amplitudes. Polarization switching decreases the device *V*TH, increasing the channel conductance. Compared with the large device, which shows continuous *V*TH shift, the scaled FeFET shows an abrupt *V*TH decrease. This is consistent with the phase field modeling and related with single/few domain switching in scaled FeFETs [3]. The number of domains in scaled device is small enough such that single/few domain switching is evident, whereas it is averaged out by the large number of domains for large device.

The conductance accumulation characteristics for different pulse amplitudes, pulse widths and pulse delays are shown in Fig.11(d-f, i-k). Similar to the erase amplitude dependence, the increase of the number of consecutive pulses cause a continuous shift in conductance for the large FeFET, while the scaled device displays an abrupt conductance increase. The conductance accumulation characteristics reflect the domain nuclei stabilization and growth processes shown in Fig.9. For scaled device, the threshold number of pulses required for domain nucleation is exponentially dependent on the applied

pulse amplitude and pulse width (Fig.11(i,j)), consistent with the nucleation physics [5]. Unlike MFM cap, the inter-pulse delay has a negligible effect on the accumulation characteristic of FeFET. This is likely due to the depolarization field in FeFET, which causes fast polarization relaxation at timescale less than 10s. Therefore, the increase in delay from 10s to 1ms does not have any effect. The pulse width can be decreased to as low as 50 ns for a pulse amplitude of 3.0V, indicating high speed operation of FeFET.

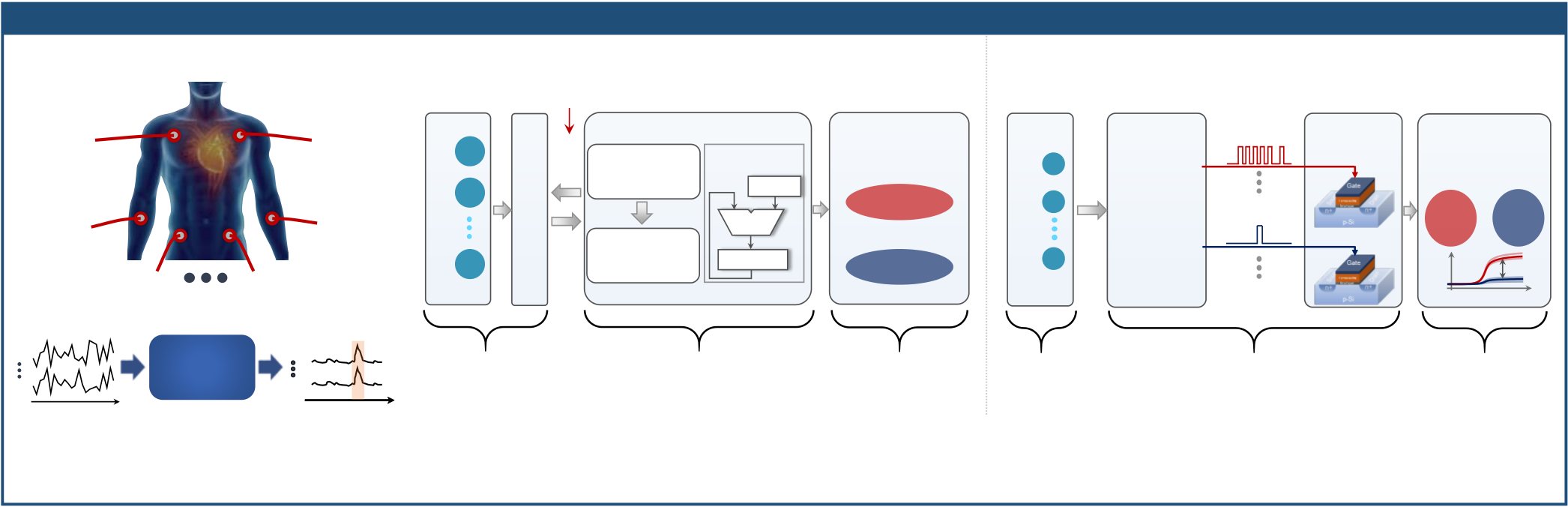
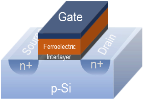
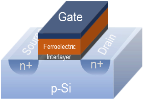
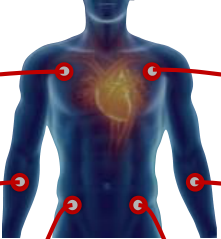
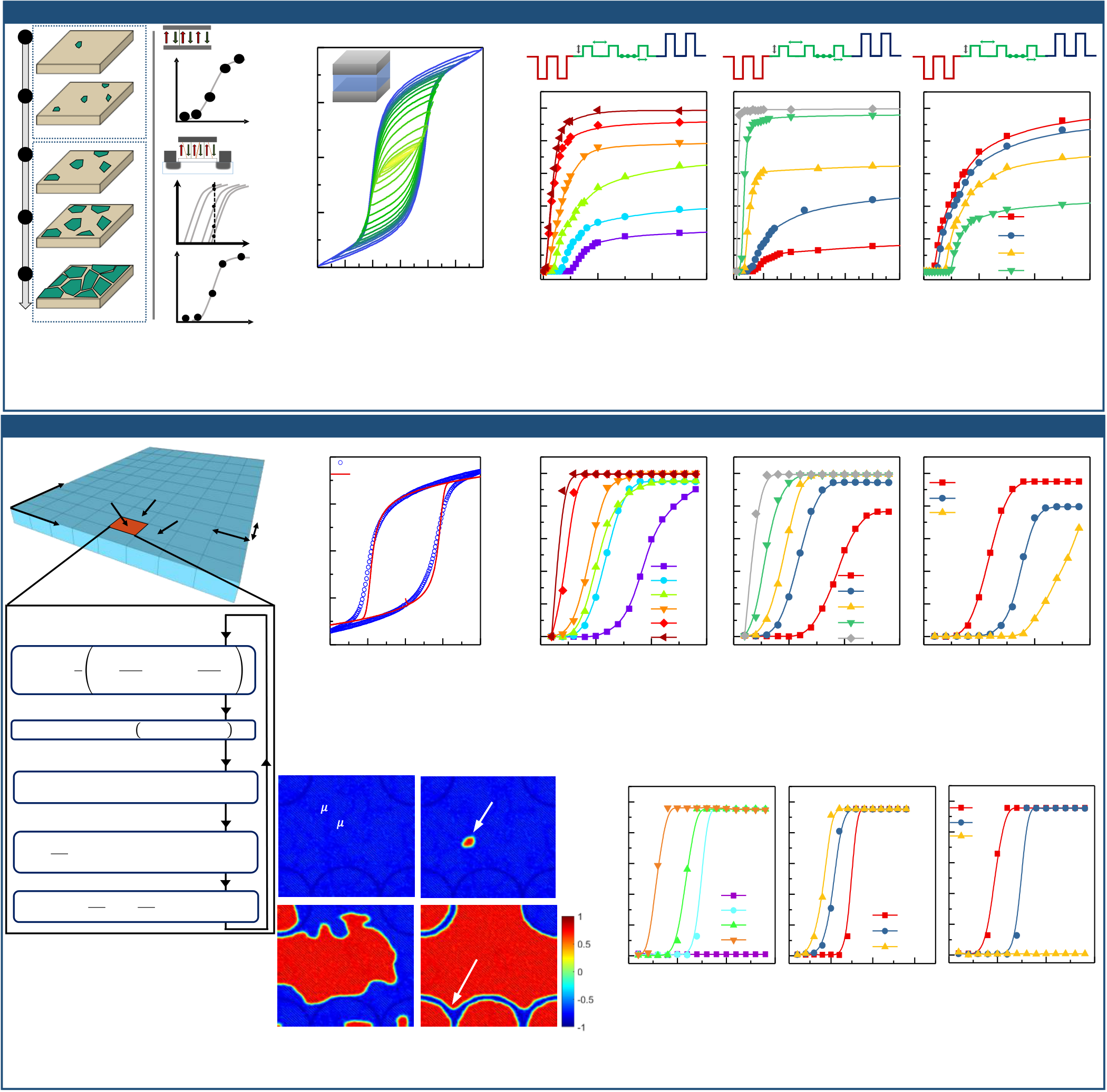
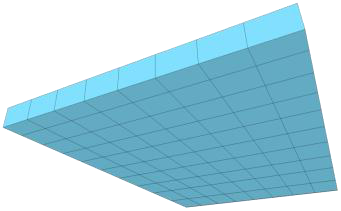
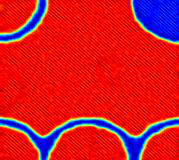
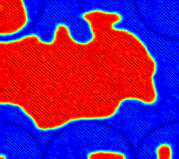
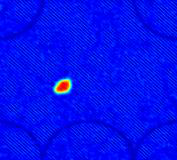
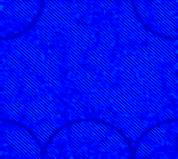
*C.Statistical Correlation Detection*   
 The polarization accumulation properties of the FeFET are applied to detect anomalies in physiological signals (for scaled geometry, multiple FeFETs are combined as a single device to have continuous accumulation characteristics). The gate pulse encoder translates the correlation among signals to the pulse rate(Fig.12). A stochastic translator (STR) is used to translate real value signals into random binary bit streams so that the multiplication operation can be greatly simplified [6]. At each time instance, the collected momentum is calculated (sum of all the signals) and a MUX is used to achieve multiplication operation necessary for correlation calculation.

The physiological signals of patient 221 from the MIMIC database show three anomalies (Fig.13(a)) [7]. By translating the real signals to gate pulses to FeFET, a conductance read can differentiate the anomalies from normal signals reliably as the anomalies are highly correlated among different signals. A majority voting on the number of FeFET conductance over the threshold identifies the anomaly faithfully. Fig.13(b) shows that detection based on 3 signals cause information loss while in the case of 7 signals the detection is successful as more global information is extracted. An array level implementation with the FeFET cells is shown in Fig. 14. Compared with a CMOS ASIC based counter [8] or PCM cell [2], FeFET exhibits high speed, large *G*max/*G*min ratio and low write energy (Fig.15).

**CONCLUSIONS**   
 In summary, we demonstrate an in-memory computational primitive for statistical correlation detection based on an industrial 28nm HKMG FeFET technology. The spatio-temporal polarization switching dynamics, involving the domain nucleation and growth, are responsible for polarization accumulation, which causes continuous conductance accumulation in large FeFET. Extremely scaled FeFET exhibits abrupt switching but can be compensated by multiple devices grouping. Fast operation with 3.0V, 50 ns is also demonstrated. The conductance accumulation characteristic is successfully applied to detect real-world physiological signal anomalies. These results make FeFET based correlation detection an ultra-dense, highly energy-efficient, and massively parallel system for real-time signal processing for sensor analytics.

**ACKNOWLEDGEMENT**   
 We would like to thank M. Trentzsch, S. Dunkel, S. Beyer, and W. Taylor at Globalfoundries Dresden, Germany for providing 28nm HKMG FeFET test devices. This work was supported in part by the Semiconductor Research Corporation (SRC) and DARPA.

**REFERENCES**   
[1] M. Trentzsch et al., IEDM 2016 [2] A. Sebastian, Nature Comm. 2017 [3] H. Mulaosmanovic et al., Appl. Mater. Interfaces 2017 [4] Y. H. Shin, et al., Nature 2007 [5] H. Mulaosmanovic et al., Appl. Mater. Interfaces 2018 [6] B. R. Gaines et al., 1969 [7] G. B. Moody et al., Computers in Cardiology 1996 [8] S. Mathew, ISSCC, 2004



**In-Memory Computing for Sensor Analytics**

Physiological Signal **(a)**  **Conventional Correlation Detection**  **(b)**  **In-Memory Correlation Detection**

**Sensor 1**  **Sensor 4**  Von Neumann Bottleneck

(Heart rate)

**X1**  **X4**   
 (ECG)

X1 **Distance Metrics**   
 **CPU/GPU**

**MAC Processing**   
 **Clustered**

**Output**   
 **Gate Pulse**

**Encoder**   
 **Correlated**

**Signal**  **Processing**

**in FeFET**   
 **Conductance**

**Read**

**Multiple Sensors**   
 **Multiple Sensors**  Store Load   
(Blood pressure)   
 **Sensor 2**

**X2**  **X5**   
 **Sensor 5**

(SpO2)   
 X2

**(e.g. distance norm)**

**Among Sensors**

Adder   
**Unit**

Multiplier

**Uncorrelated Processes**   
**Correlated Processes**

X1, X5, X11,… X2   
 X1

Pulse rate



j   
N

1   
 R ij  
 i

**Uncorrelated**

**Signal**   
 **Correlated**

**Processes**

X1,

X5,

X11,…  
 **Uncorrelated**

**Processes**

X3,

X4,

X7,…**Memory**

Load

**Clustering Algorithm**

(Respiration) **Sensor 3 X3**  **XN**  **Sensor N**  Xi **(e.g. K-means)**  Accumulator X3, X4, X7,… Xi Ri,j: correlation

between Xi and Xj

Time   
*G*DS

**Sensed Signal**  **Anomaly Detected**

***X1***

Correlation Statistical ***Y1***  **Sensed**  **Correlation**  **Anomaly**  **Sensed**  **Correlation**  **Anomaly**

***XN***  Detection ***YN***  **Signal**  **Detection**  **Detection**  **Signal**  **Detection**  **Detection**

time time

**Fig. 1:** Physiological anomaly detection **Fig. 2:** (a) Conventional approach for correlation detection involves data transfer between the memory and

based on the change of correlation computing unit, limiting the throughput and energy efficiency. (b) In-memory computing approach

among signals. utilizing conductance accumulation of FeFET maps correlation among signals to FeFET gate pulse rate.

**Physics of Polarization Switching Dynamics (MFM): Experiment**

**1**  **MFM**

**Capacitor**  40 **(a)**  **VPulse**  **10 *μ*s**  **(b)**  **0.8 V**  **10 *μ*s**  **(c) 0.8 V**  **TDelay**

**ΔPr**  **4 5**  **W**  **1 *μ*s**  **Set**  **PW**  **Set**  **1 *μ*s**  **Set**

**3**  **10nm Hf0.6Zr0.4O2**  **Reset**  **Reset**  **Reset**

**2**  **1 2**  **Pulse**

**Number**  20 **W**

1.0 VPulse=1.2V PW=100s

QFE (C/cm2)   
**Pulse Number**   
 PW=10s

**G G**  0.8 VPulse=1.0V

**3**

 **ID**   
 **n+**

**S**

**p-Si**

**4**

**5**   
**n+**

**D**

**FeFET**  0

0.6   
 VPulse=0.9V

VPulse=0.8V PW=5s P/2PsP/2Ps   
**4**   
 **3**

**2**  
 -20 0.4

VPulse=0.7V PW=1s TDelay=10s

**5**   
 **ID**

**1 2**   
 **3**   
 **1**

**4**

**Pulse**

**Number**   
**5**   
**Vg**

-40

-3 -2 -1

VFE (V)   
 0 1 2 3   
 0.2

0.0

0

Pulse Number   
 20 40   
 VPulse=0.6V

60 0 10 20 30 40 50 60

Pulse Number   
 PW=400ns

0

Pulse Number   
 20  
 TDelay=100s

TDelay=1ms

TDelay=1s

40 60

**Fig. 3:** Polarization accumulation **Fig. 4:** Measured *Q*FE-*V*FE loop for **Fig. 5:** Measured polarization accumulation as a function of pulse number

due to domain nuclei stabilization W/HZO/W MFM capacitor showing on W/HZO/W MFM capacitor for different (a) pulse amplitude, (b) pulse

and growth causes conductance non-saturated minor loops due to width and (c) inter-pulse delay. Decrease of amplitude and pulse width

accumulation in FeFET. partial polarization switching delays the onset of polarization accumulation.

**Phase Field Modeling of Polarization Switching Dynamics (MFM)**

**y**

**x**   
 **(i, j)**  **(i, j+1)**

**(i+1, j)**  2 nm   
 40

20  
 Experiment

Phase Field Simulation   
 **(a)**

1.0

0.8   
 **(b)**  **(c)**

TDelay

10s

100s

1ms

QFE (C/cm2)   
2 nm 0 0.6

VPulse

0.4 0.6V

0.7V   
 PW

400ns

**Phase-Field Simulation Setup** -20 0.2 0.8V

0.9V  
 1s

5s

**t = t + 1** 1.0V 10s

Domain-domain interaction -40 0.0 1.2V 100s

-2 -1 0 1 2 0 5 10 15 0 5 10 15 0 5 10 15

VFE (V) Pulse Number Pulse Number Pulse Number

Accumulation **Fig.**  **7:**  *Q*FE-*V*FE loop **Fig. 8:** Simulated polarization accumulation in 200nmx200nm MFM cap as

calculated with phase a function of pulse number using calibrated phase field model for different

field model is well (a) pulse amplitude, (b) pulse width and (c) inter-pulse delay. Simulations

Total effective electric field calibrated to experiment. are in good agreement with experimental results.

**Vpulse=0.7V**

**PW=1**𝜇**s**  **Critical Nuclei Size**  1.0 **(a)**  **(b)**  TDelay

10s   
 **(c)**

Solve time-dependent LGD **Tdelay=10**𝜇**s**  0.8 100s

1ms

P/2Ps Neumann boundary condition   
 **Pulse #2**  **Pulse #3**  0.6

VPulse

**ΔP/2Ps**   
 0.4 0.6V

0.7V   
 PW

400ns

0.2 0.8V 1s

P : polarization; : Viscosity coefficient

: Landau coefficients **Grain Boundary**  0.0 0.9V 5s

KP : domain wall coupling parameter 0 5 10 15 0 5 10 15 0 5 10 15

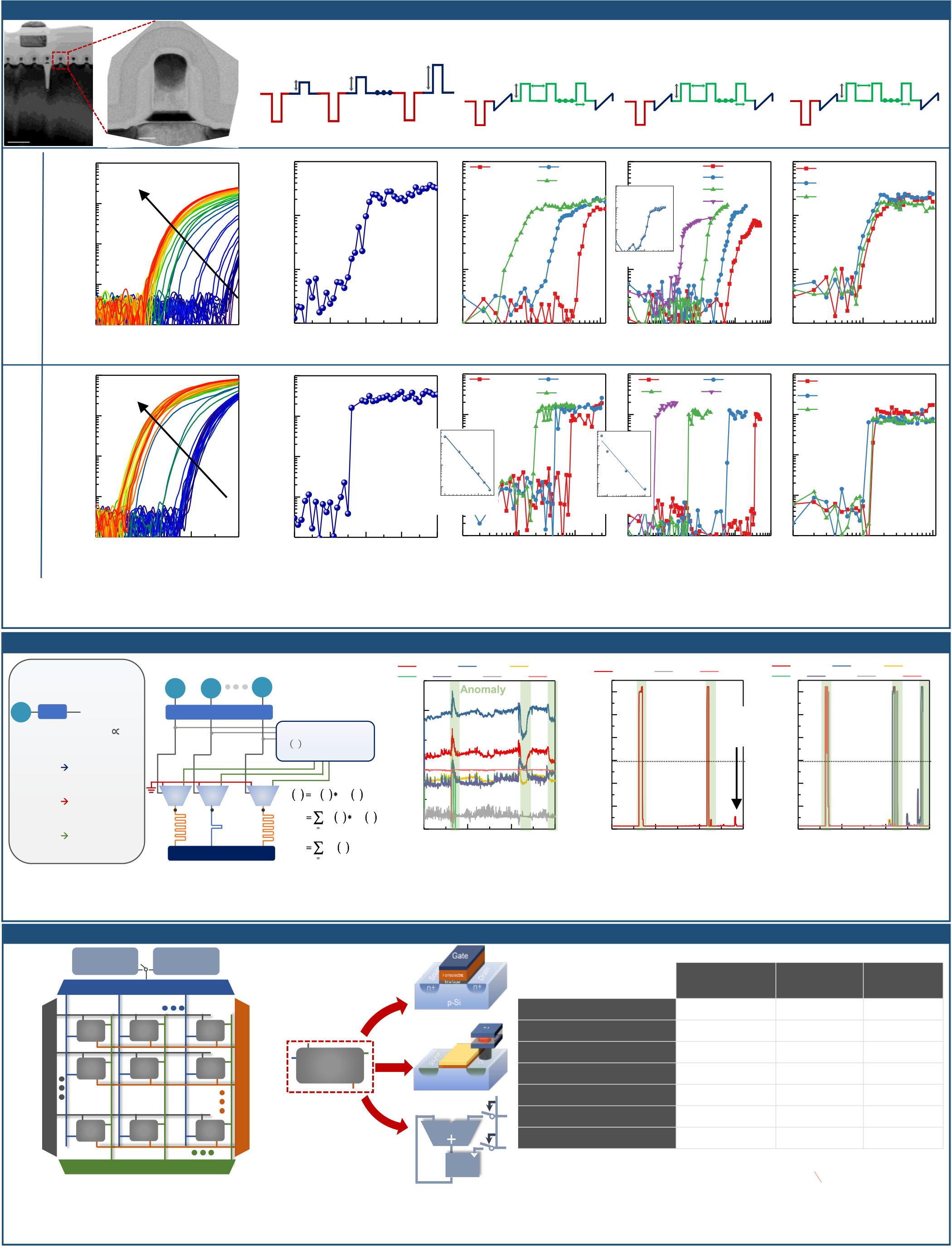
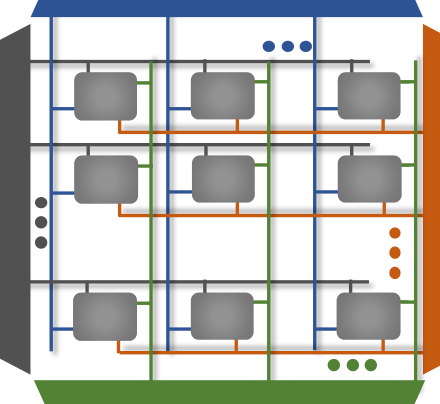
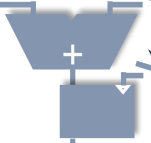
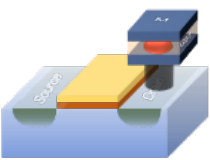
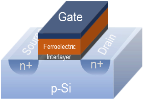
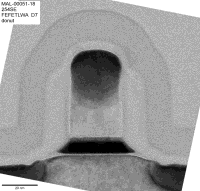
PN: pulse number; : Fitting parameter Pulse Number Pulse Number Pulse Number

**Pulse #6**  **Pulse #10**

**Fig.**  **6:**  Phase field simulation **Fig. 9:** Simulated polarization snapshot **Fig. 10:** Simulated polarization accumulation in 20nmx20nm single

framework. at different pulse numbers showing the grain MFM cap as a function of pulse number for different (a) pulse

domain nuclei stabilization and growth. amplitude, (b) pulse width and (c) inter-pulse delay.



**FeFET Conductance Accumulation Characteristics**

**(a)**

**Erase Amplitude**  **Voltage Amplitude**  **Pulse-Width**  **Time-Delay**

**VERS**  **VERS**  **VERS**  **VPulse10 *μ*s**  **Read**  **2.3V 10 *μ*s**  **Read**  **2.3V TDelay**  **Read**

**Erase**

**1 *μ*s**  **PW**  **1 *μ*s**

**200nm**  **20nm**  **Program -4V**  **Program**  **Program**  **Program**

**(b)** 10-5

**VERS=2 to 4V**  **(c)** 10-6

**(d)** VPulse=2.4V

VPulse=2.6V  
 VPulse=2.5V **(e)**

PW=50 ns  
 PW=50ns

PW=100ns

PW=1s   
 **(f)** TDelay=10s

TDelay=100s

10-6

10-7   
 VAMP=3.0 V PW=10s TDelay=1ms

**L,W=450nm L=34nm, W=74nm**   
 ID (A) ID (A)   
 ID (A) ID (A)   
 10-7

10 0 10 1 10 2

10-8 Pulse Number

10-8

10-9

0.0 0.5 1.0 1.5   
 10-9

2.0 2.5 3.0 3.5 4.0 10 0 10 1 10 2 10 0 10 1 10 2 10 3 10 4 10 0 10 1 10 2

VGS (V) VERS (V) Pulse Number Pulse Number Pulse Number

**(g)** 10-5

**VERS=2 to 4V**  **(h)** 10-5

**(i)** VPulse=2.4V

VPulse=2.6V  
 VPulse=2.5V **(j)** PW=50ns

PW=1s PW=10s  
 PW=100ns **(k)** TDelay=10s

TDelay=100s

10-6 10-6 TDelay=1ms

10 3

10 3

Pulse Number   
 Pulse Number 10-7 10-7

10 2 10 2

10-8 10-8

10 1

2

Voltage (V)

2.5   
 10 1

Pulse Width (s)

10-7 10-6 10-5

10-9

0.0 0.5 1.0 1.5   
 10-9

2.0 2.5 3.0 3.5 4.0 10 0 10 1 10 2 10 0 10 1 10 2 10 3 10 4 10 0 10 1 10 2

VGS (V) VERS (V) Pulse Number Pulse Number Pulse Number

**Fig. 11:** (a) TEM image of 28nm FeFET; Conductance accumulation in (b-f) large device (L,W=450nm) and (g-k) scaled (L=34nm,W=74nm)

device. Continuous *V*TH shift and abrupt *V*TH shift are observed for large and small device, respectively. The abrupt change is a signature of single-

domain switching. 50ns operation is demonstrated.

**Correlation Detection using FeFET**

**STR: Stochastic**

**Translator**  X1   
 **Gate Pulse Encoder**

X2 XN   
 **(a)** ABPmean

HR

**1**  
 PULSE   
 ABPsys

RESP

**2**  
 ABPdias

SpO2

**3**   
 **(b)**

6  
 APBMean

**3 Signals**  
 RESP SpO2 **(c)**

6  
 ABPmean

HR PULSE

**7 Signals**  
 ABPsys

RESP   
 ABPdias

SpO2

Xi **STR**  streams(0/1) random bit **STR**  200 5 **Anomaly**  5 Conductance (S) Conductance (S)   
 Intensity (a.u)   
 p(1) ∝Xi O1 O2 ON **Momentum Calc.**

4 **Detected**   
 **# 3 Not**

4

**Bit Stream**  **1 1 1 1 1 1 1 1 1 1**   
**Xi=1**

100   
 3 **Threshold**  3 **Threshold**

**Xi=1/2**  2 2

**Bit Stream**  **0 1 0 0 1 1 1 0 1 0**  **MUX**  **MUX**  **MUX** Y N M k

N  
 O N 

1 1

**Xi=1/5**  Y1 Y2 YN O k j O N 

**Bit Stream**  **0 0 0 0 0 1 0 0 0 1**

**FeFET Gates**   
 j

j   
N



  
1

1   
R jN   
 0

0 5000

Time (s)   
 10000 15000   
 0

0 5000

Time (s)   
 10000 15000   
 0

0 5000

Time (s)   
 10000 15000

**Fig. 12:** Gate pulse encoder translates the input signal **Fig. 13: :** (a) Physiological signals from MIMIC database; Evolution of FeFET conductance

to random bit streams so that the multiply operation corresponding to each signal for (b) 3 signals input and (c) 7 signals input. Detection based

can be simplified to MUX gate. on 3 signals lose information and miss one anomaly, which is avoided in 7 signals detection.

**System & Benchmarking**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **(a)** | WL Decoder | Read  circuitry | Write  circuitry | | **Single Cell** | **(b)** | **FeFET Single Cell** | | Gate | n+ | **Cond. accumulation type** | **90 nm CMOS ASIC [8]** | **PCM [2]** | **FeFET**  **(this work)** |
| BL Decoder | | |
| **Super-linear** |
| **Linear** | **Rectifying** |
| C1,1 | C1,2 | C1,j | Ci,j | **PCM Single Cell**  n+ | **# of Transistors per cell** | **~ 500** | **1** | **1** |
| C2,1 | C2,2 | C2,j | **Nonvolatile cell** | **No** | **Yes** | **Yes** |
| **Pulse width** | **1 ns** | **100 ns** | **50 ns** |
| Ci,1 | Ci,2 | Ci,j | **CMOS ASIC**  **Single Cell** | | p-Si |
| ***Gmax*/*Gmin* ratio** | **N/A** | **20** | **100** |
| **Energy per reset pulse** | **5 pJ** | **580 pJ** | **0.06 pJ** |
| **Energy per set pulse** | **5 pJ** | **1.5 pJ** | **0.06 pJ** |
| Multiplexer | | | REG |

**Fig. 14:** (a) Array level implementation of the correlation detection. The accumulation pulse update is performed row-wise. (b) The individual cell is either FeFET, phase-change memory (PCM) or CMOS counter.

**Fig.**  **15:**  Unit cell for statistical correlation detection

benchmarking. 28 nm FeFET exhibits good speed, Gmax/Gmin ratio

and write energy efficiency compared with CMOS ASIC or PCM.