

Computer Architecture

Lecture 16: Flash Memory and Solid-State Drives

Dr. Mohammad Sadrosadati

Prof. Onur Mutlu

ETH Zürich

Fall 2023

17 November 2023

Brief Self Introduction



■ Mohammad Sadrosadati

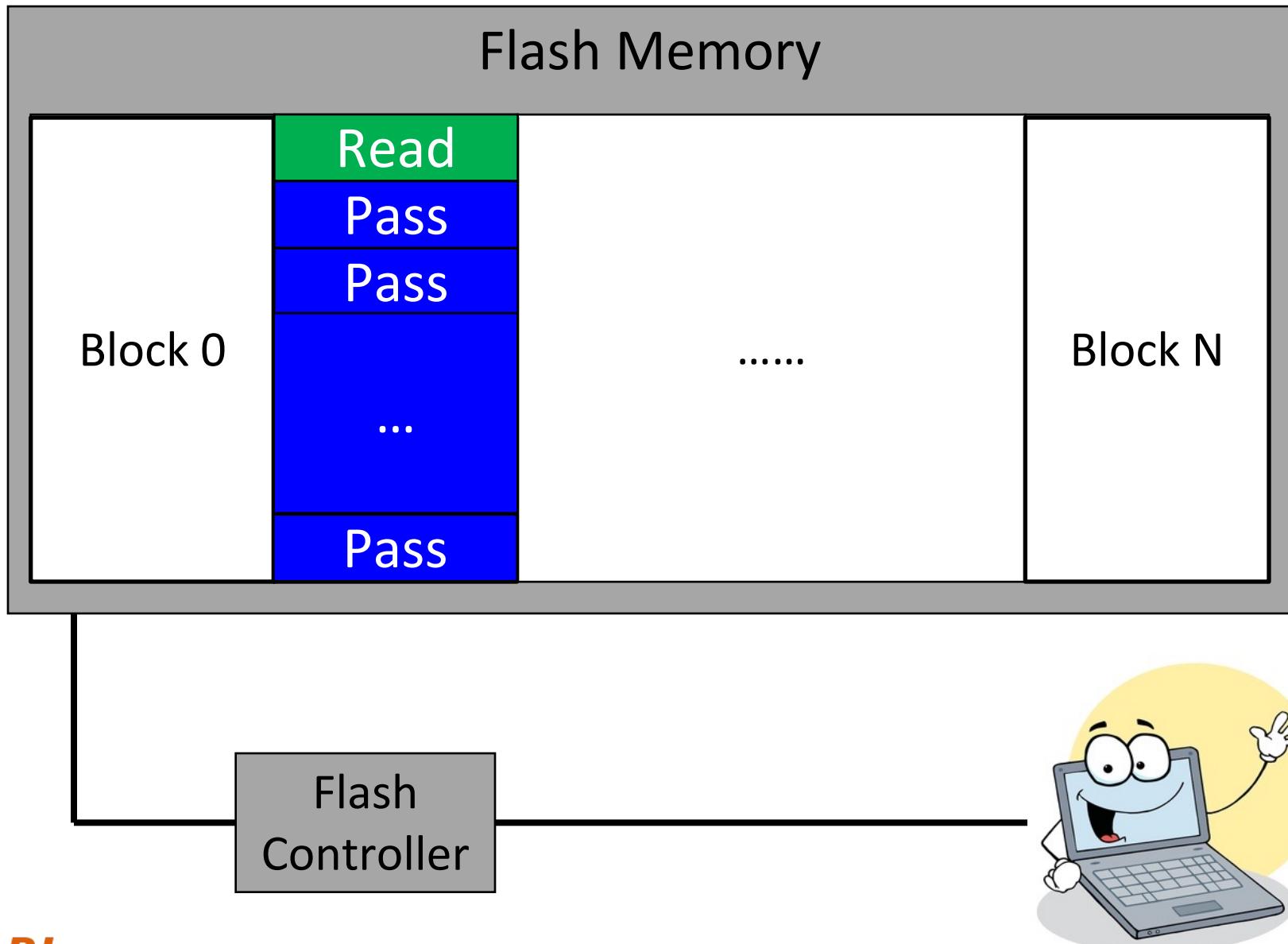
- Senior Researcher and Lecturer @ SAFARI Research Group, ETHZ
- Postdoc @ IPM 2019–2021
- PhD from Sharif University of Technology, 2014-2019
- mohammad.sadrosadati@safari.ethz.ch

■ Research Areas

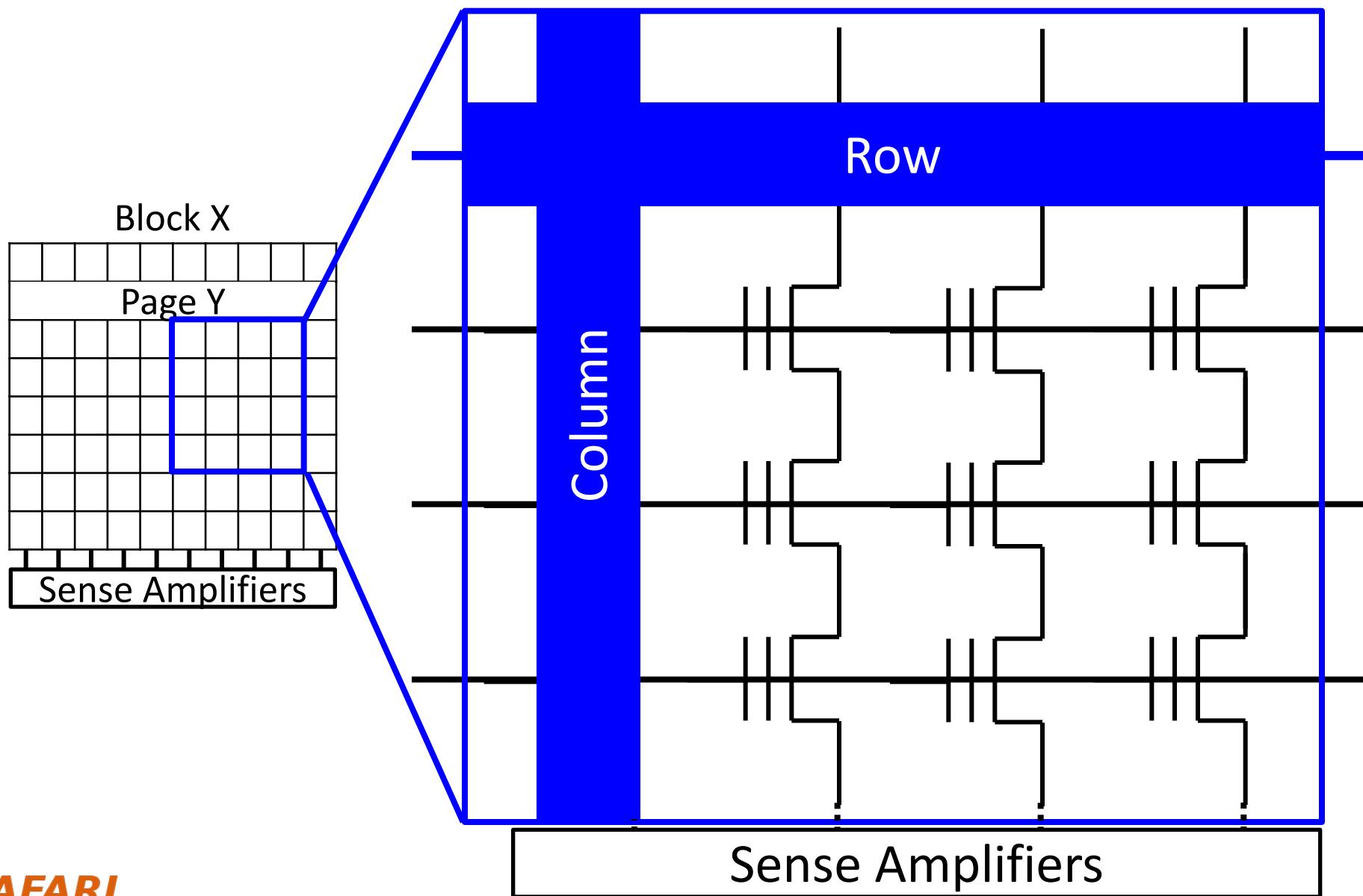
- Computer Architecture
- Memory & Storage Systems
- Near-Data Processing
- Heterogeneous System Architecture
- Bioinformatics
- Interconnection Networks

Short Background on NAND Flash Memory Operation

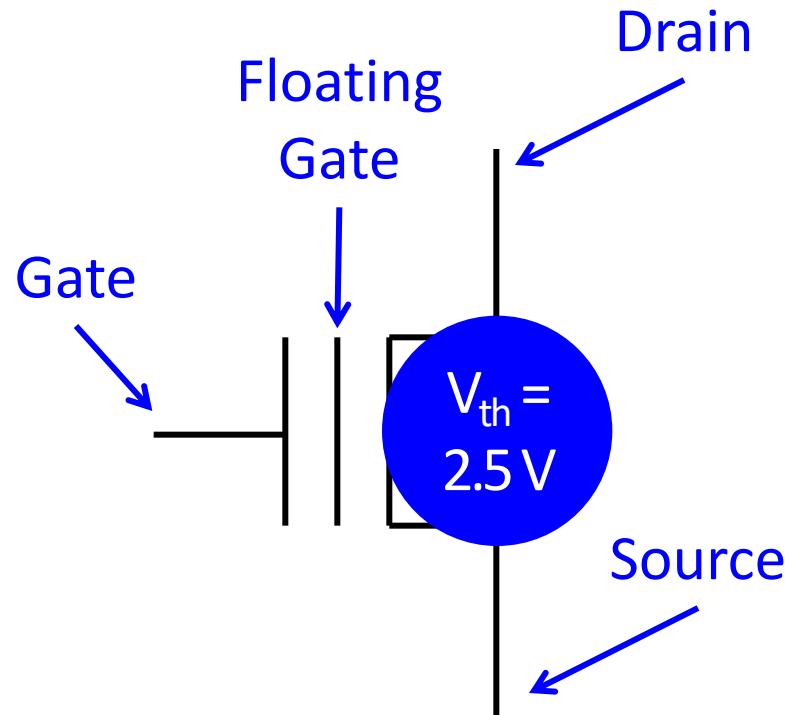
NAND Flash Memory Background



Flash Cell Array

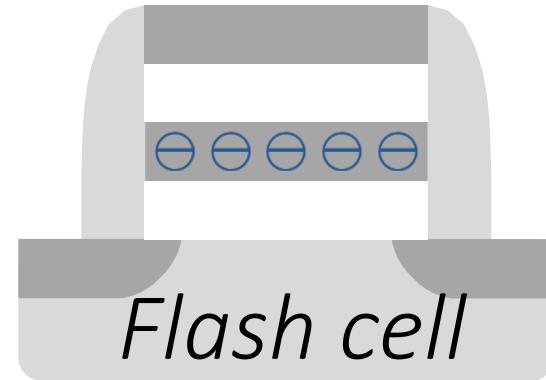
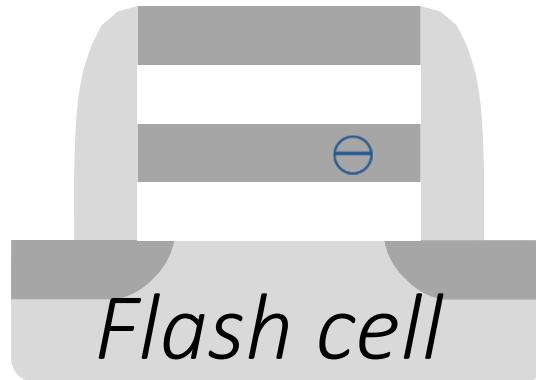


Flash Cell



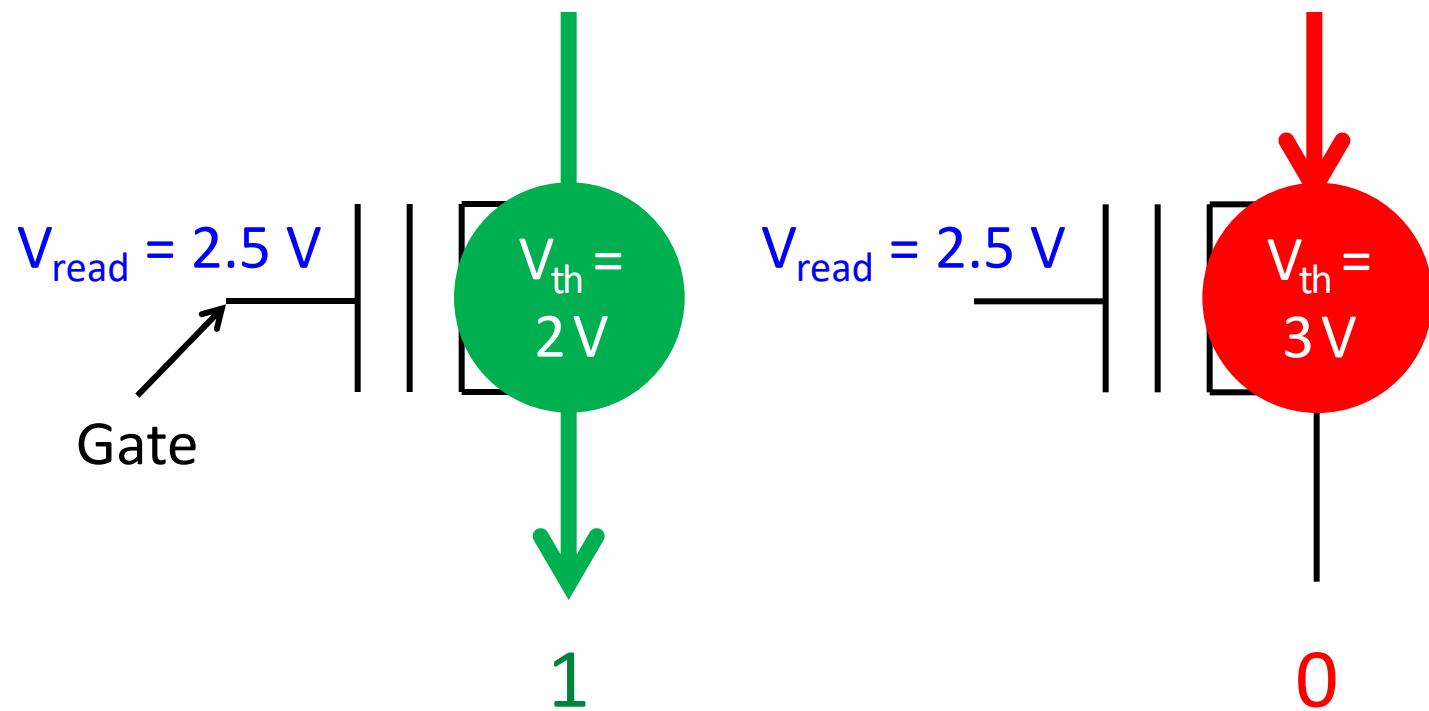
Floating Gate Transistor
(Flash Cell)

Threshold Voltage (V_{th})

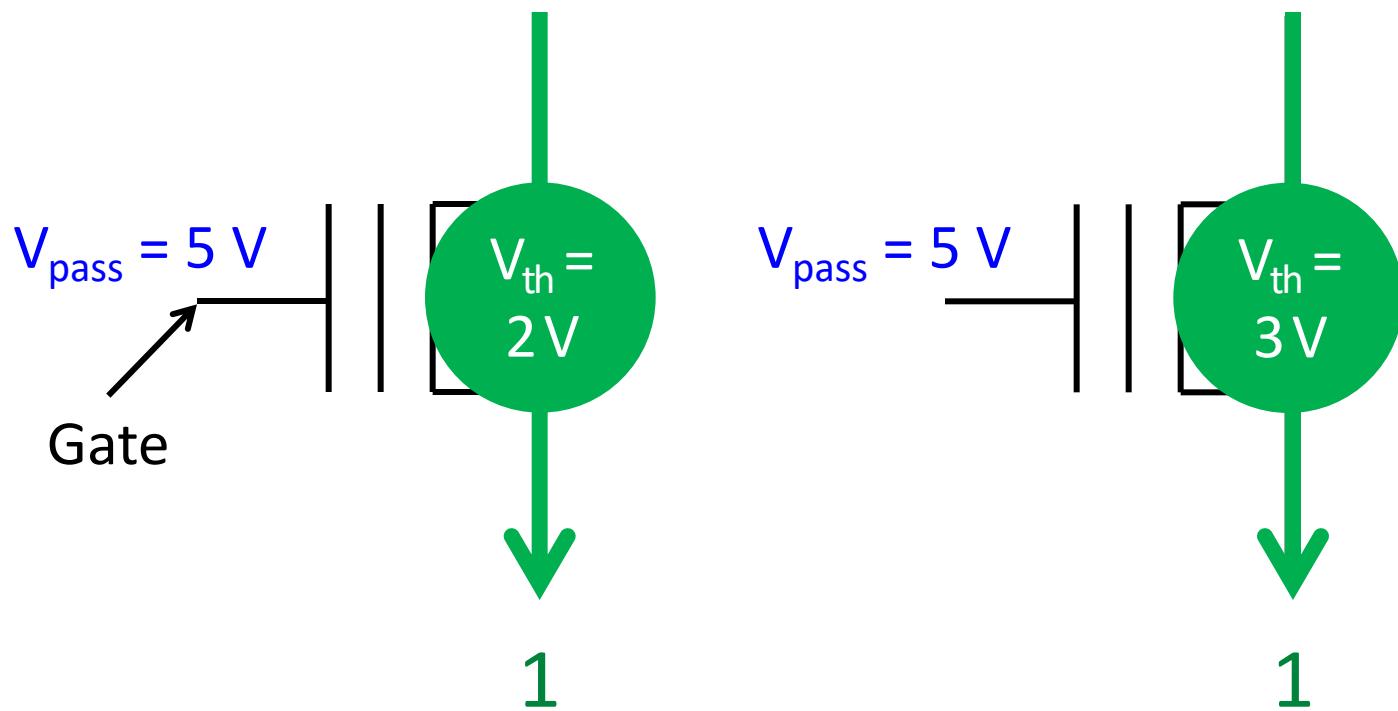


Normalized V_{th}

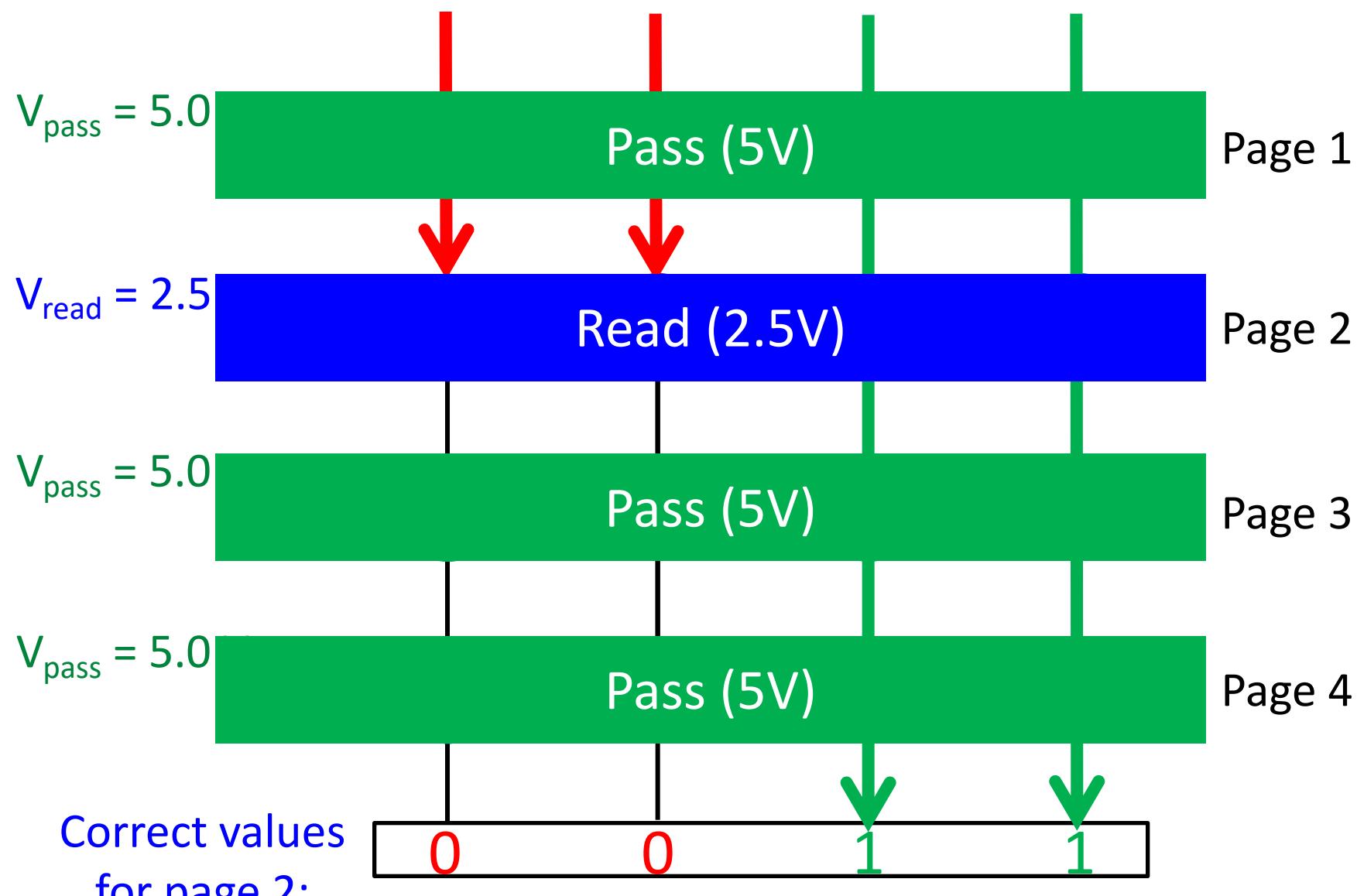
Flash Read



Flash Pass-Through

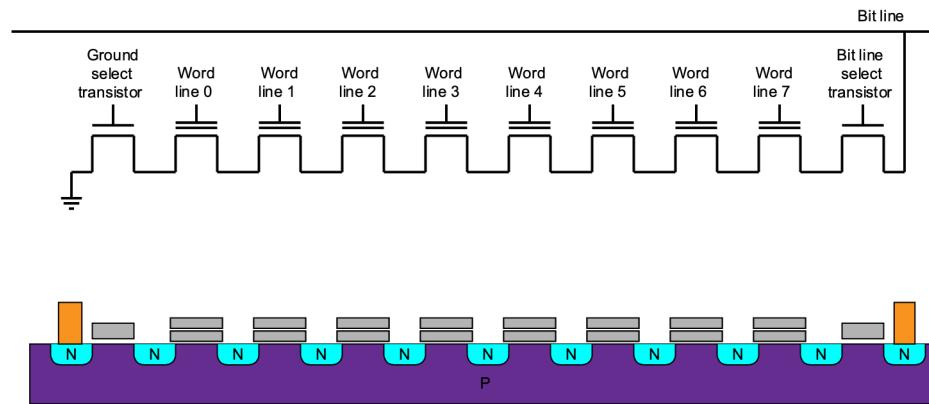


Read from Flash Cell Array

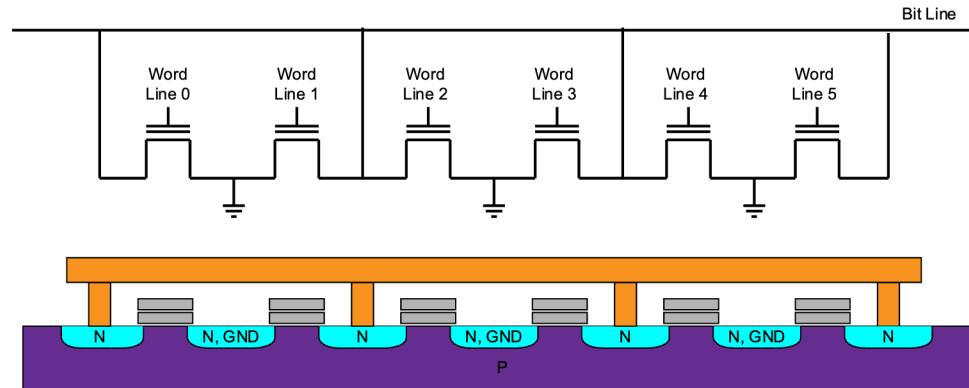


Aside: NAND vs. NOR Flash Memory

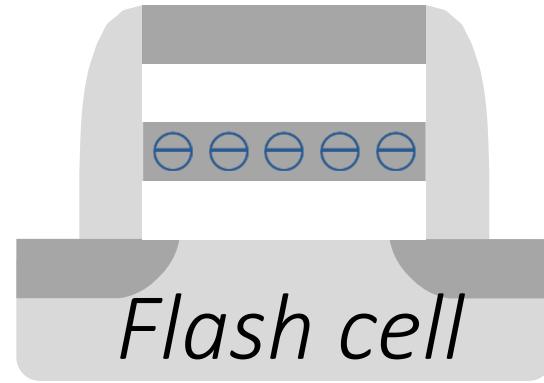
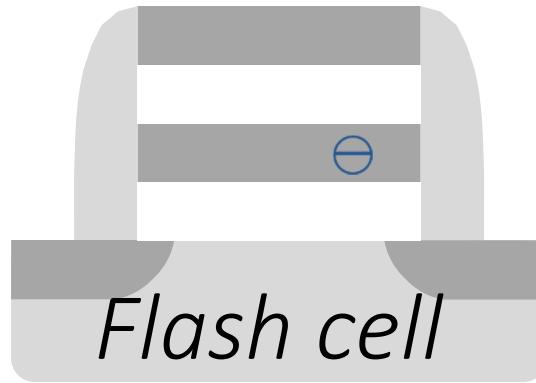
NAND



NOR



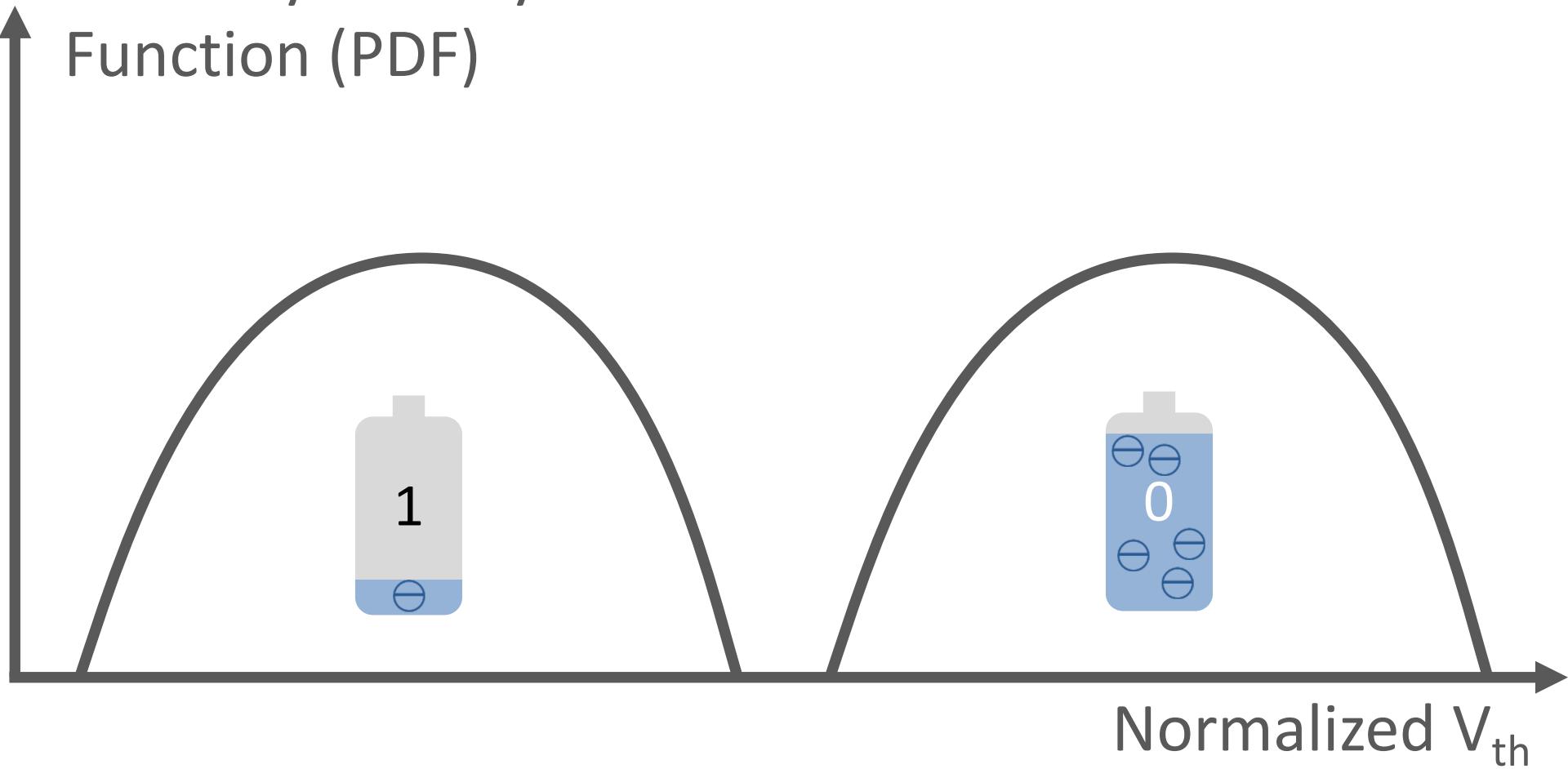
Threshold Voltage (V_{th})



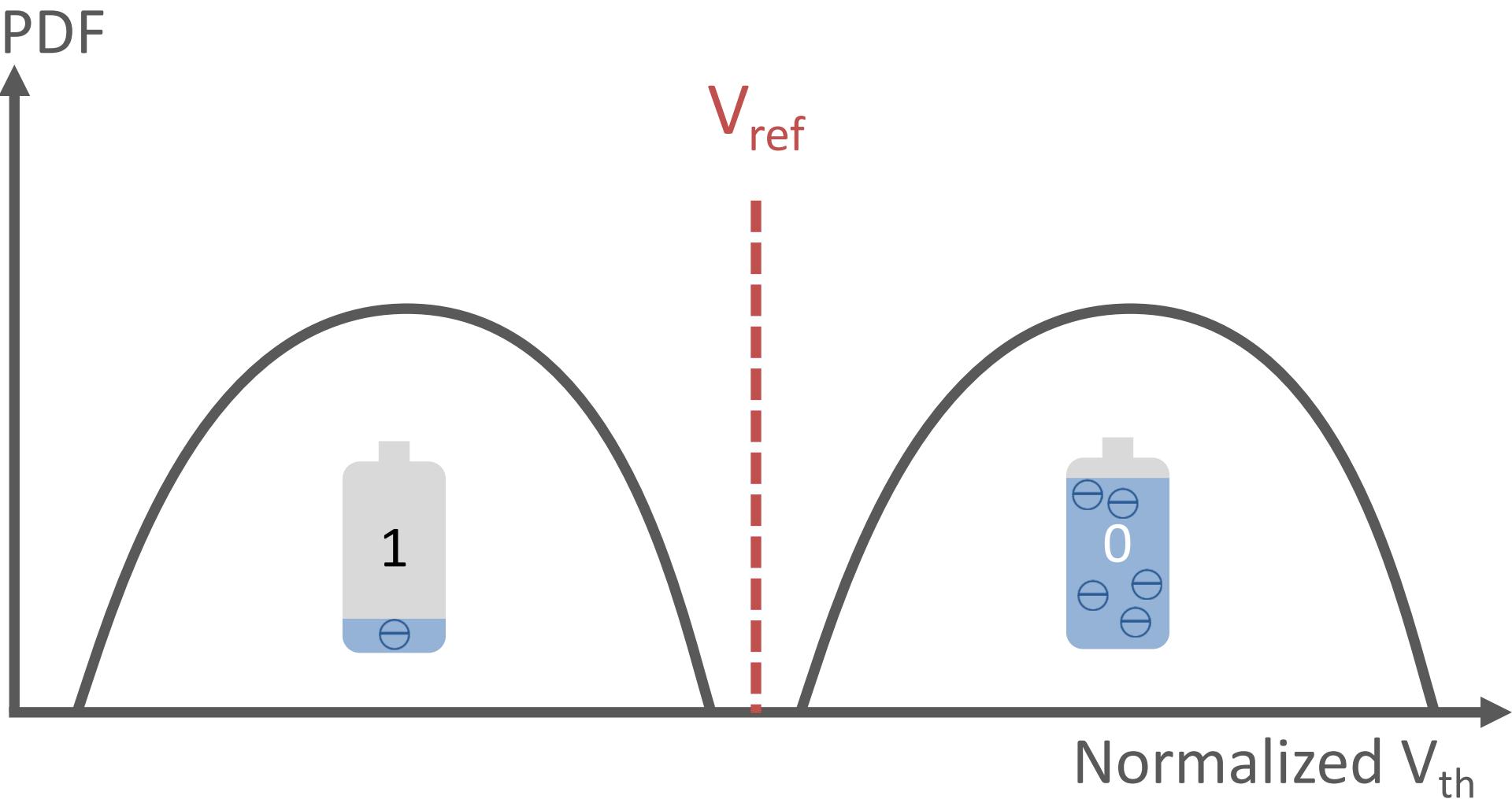
Normalized V_{th}

Threshold Voltage (V_{th}) Distribution

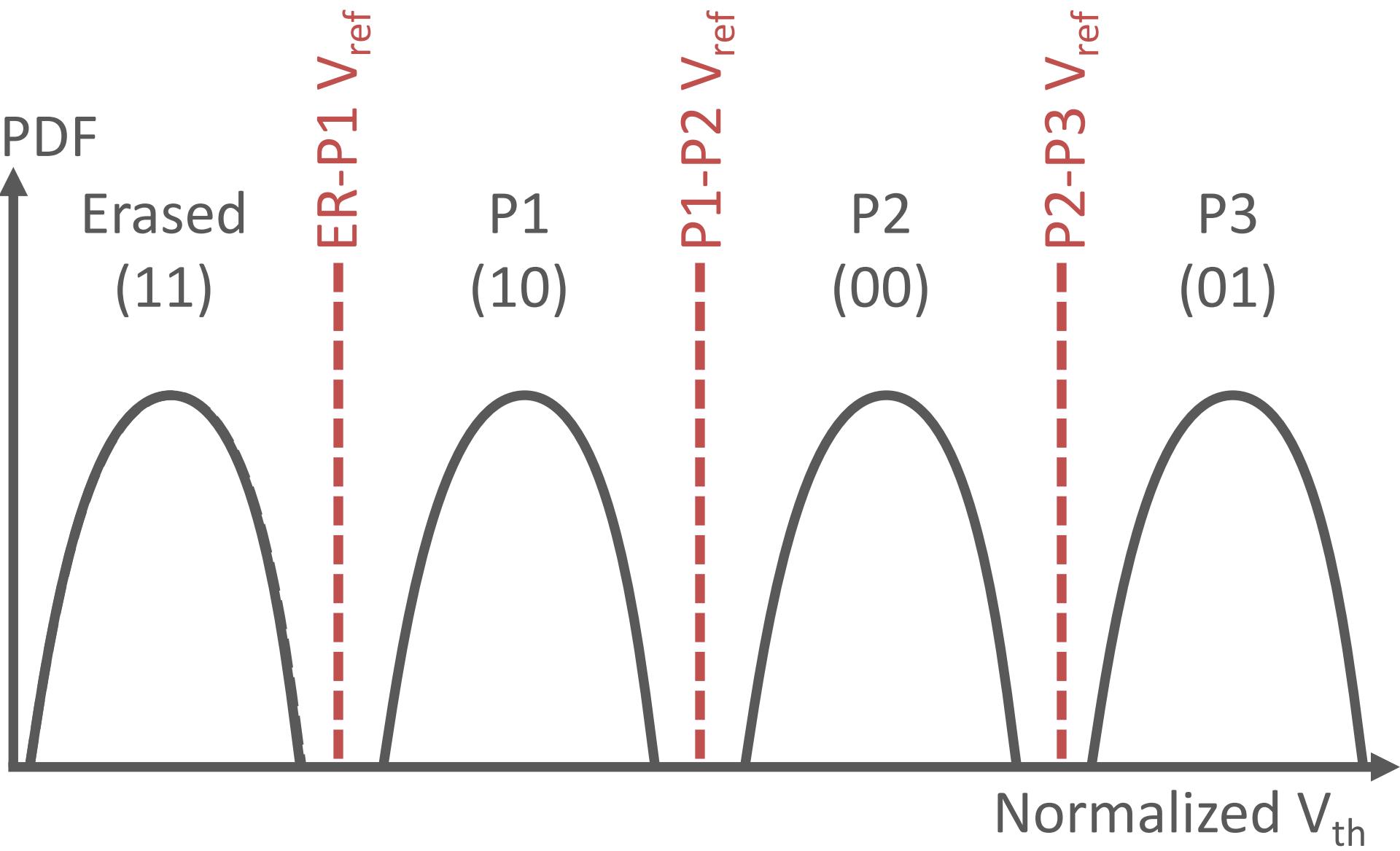
Probability Density
Function (PDF)



Read Reference Voltage (V_{ref})

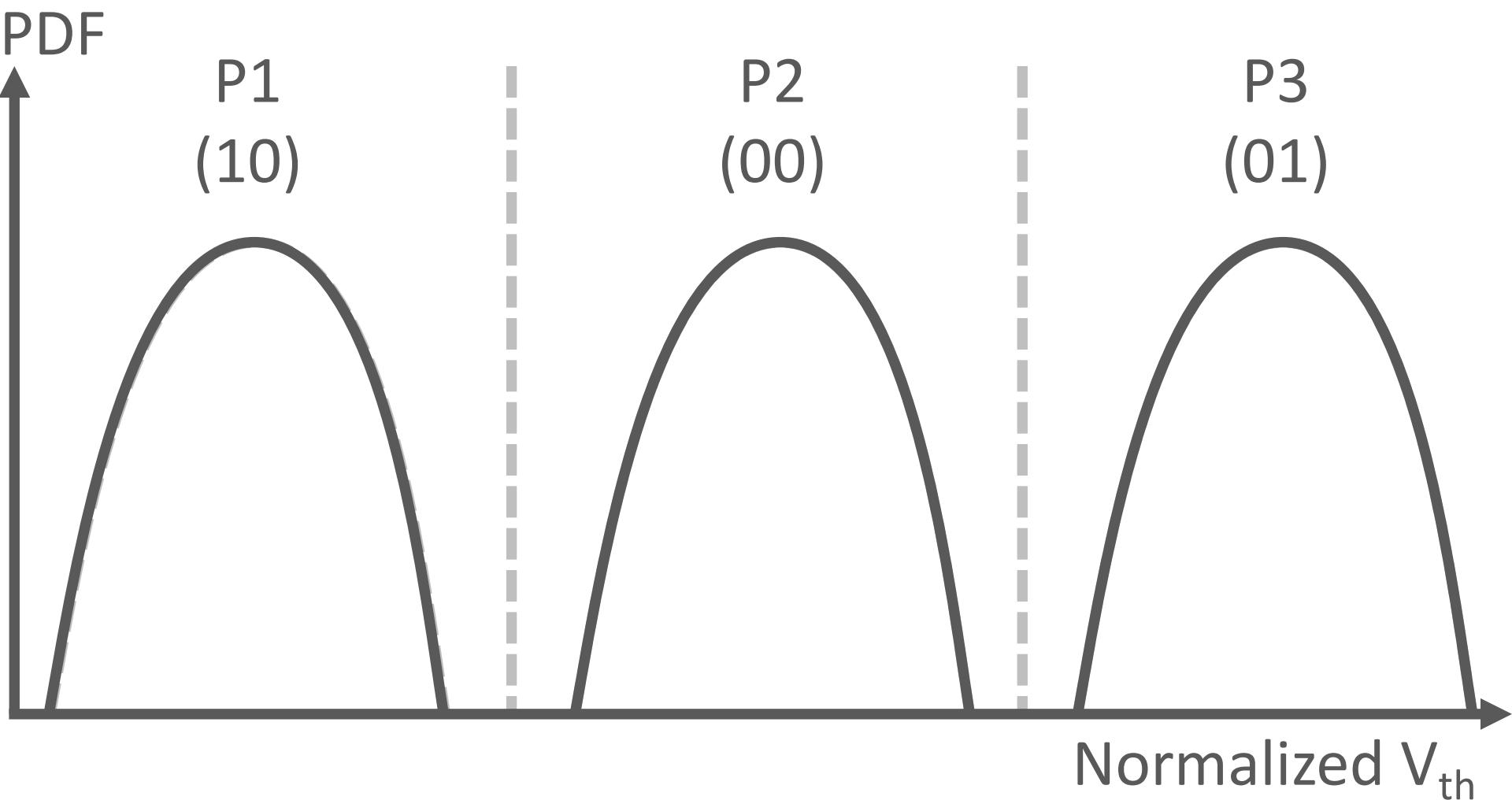


Multi-Level Cell (MLC)



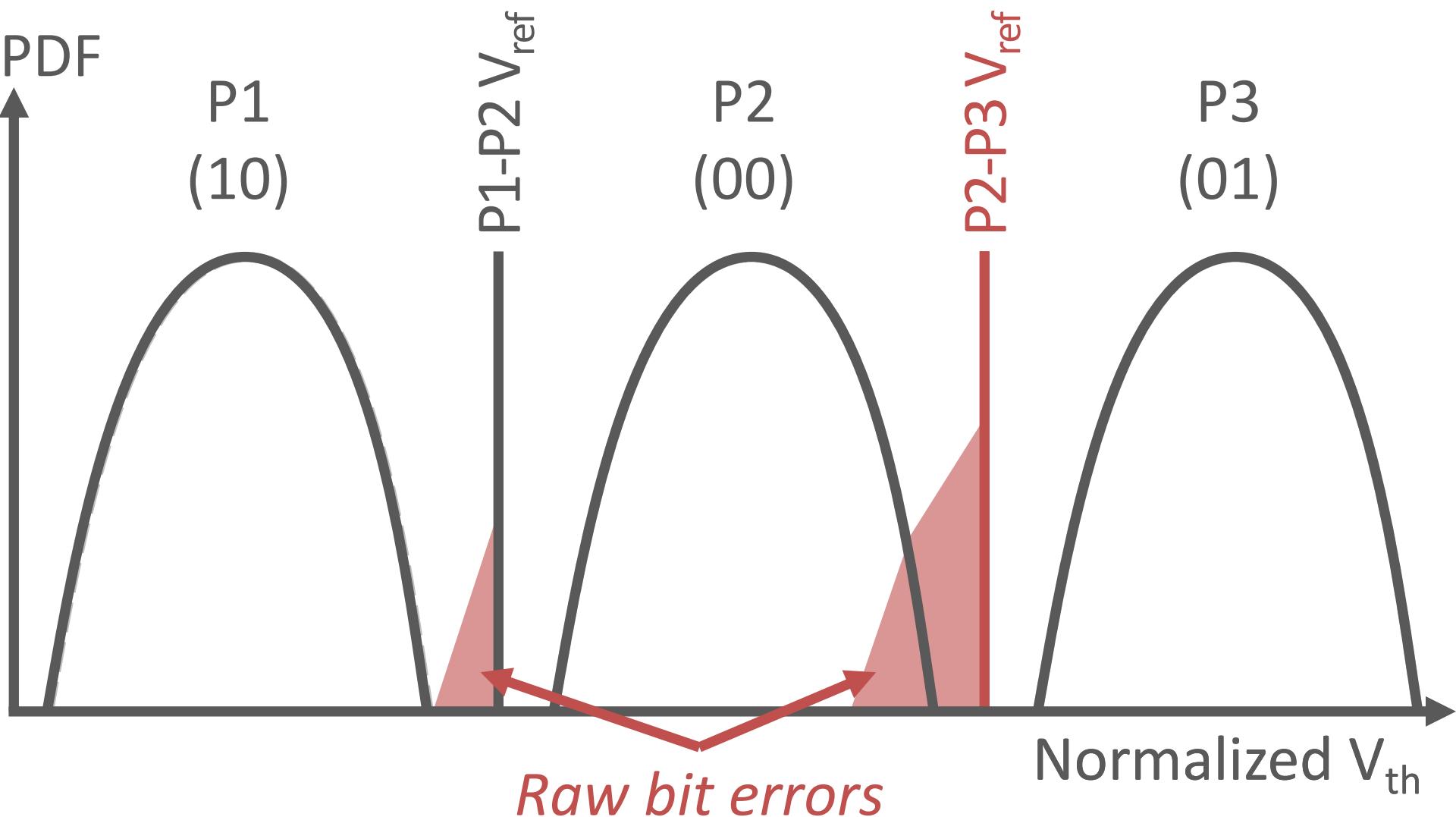
Threshold Voltage Reduces Over Time

After some retention loss:



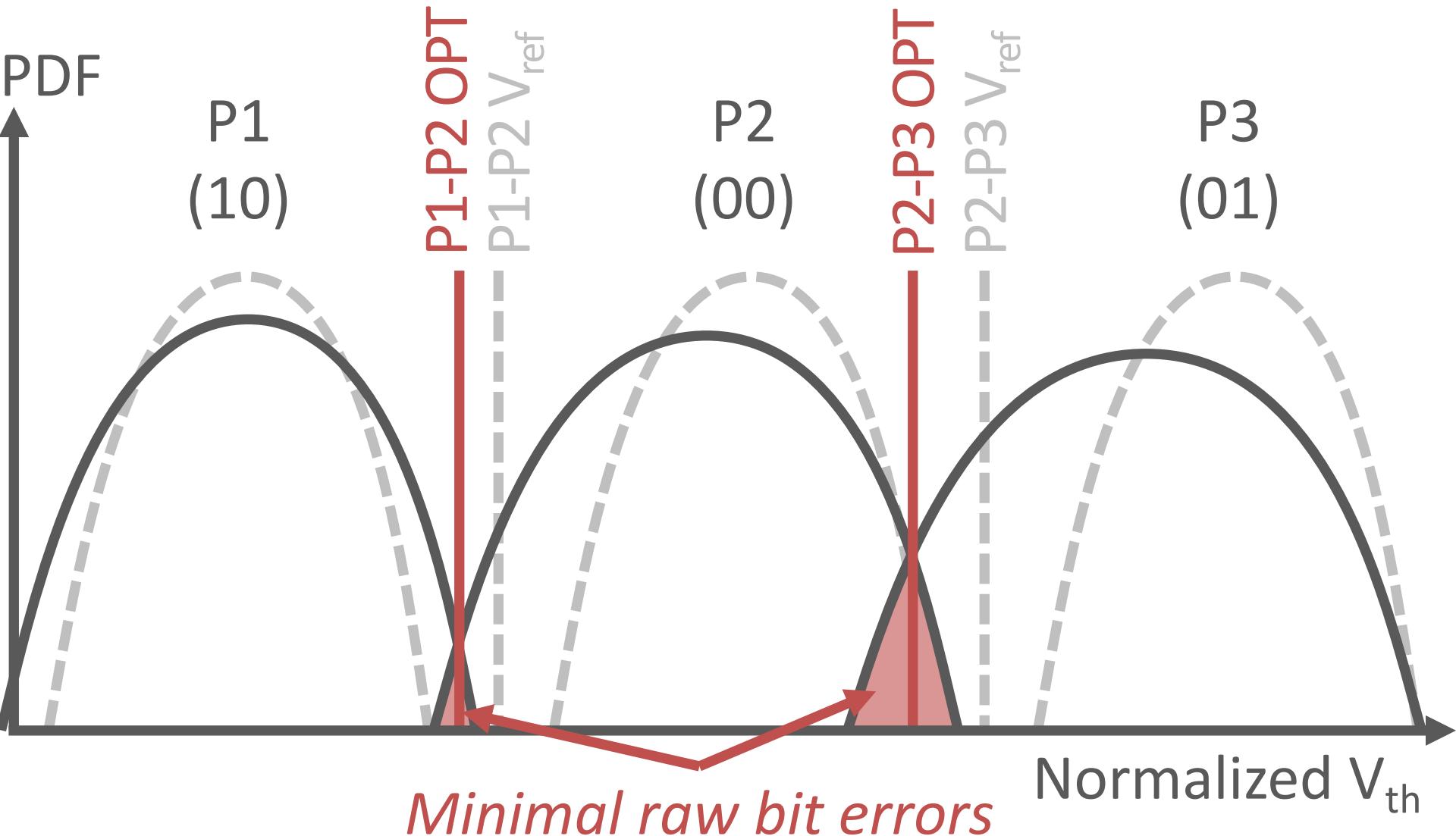
Fixed Read Reference Voltage Becomes Suboptimal

After some retention loss:



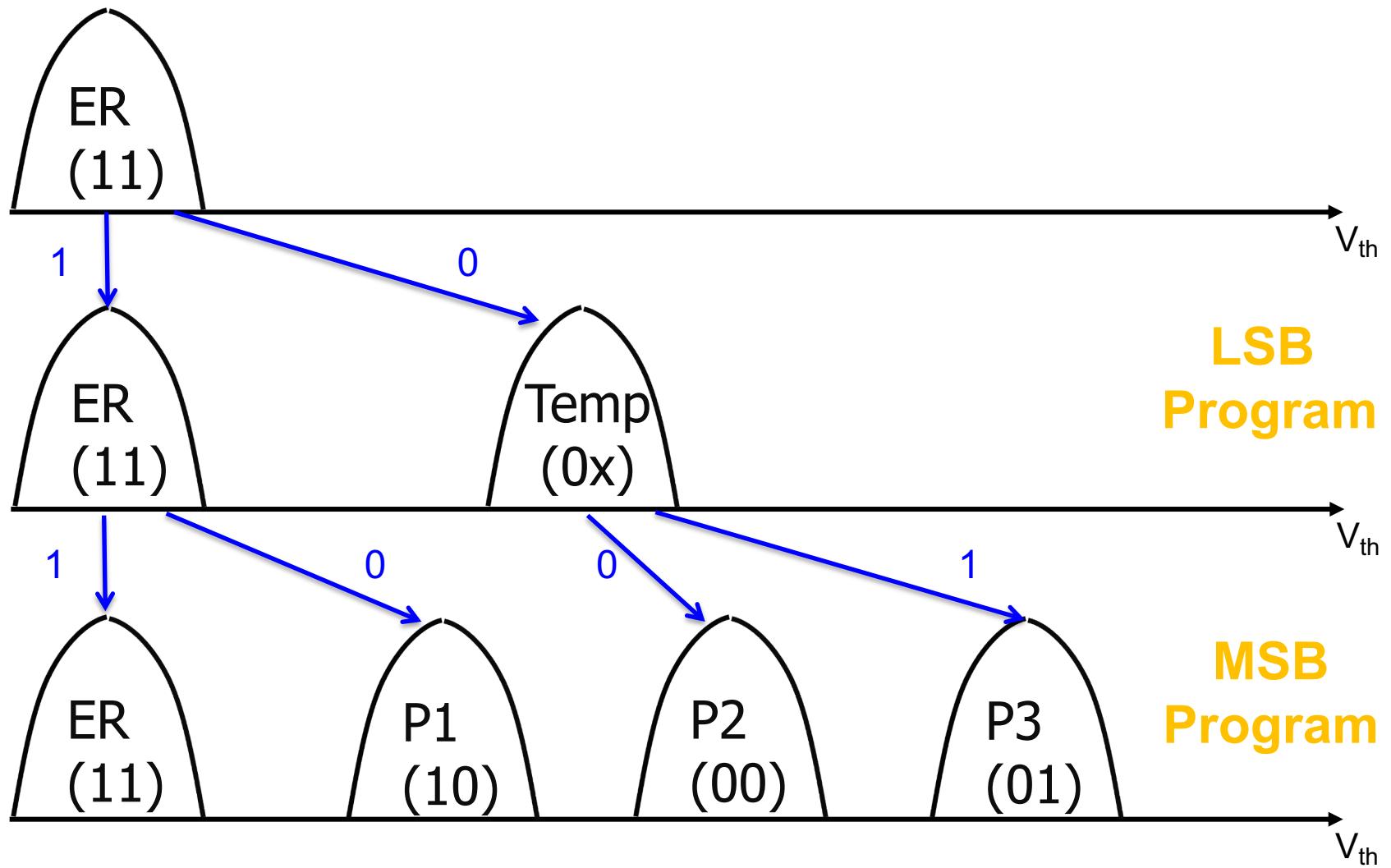
Optimal Read Reference Voltage (OPT)

After some retention loss:

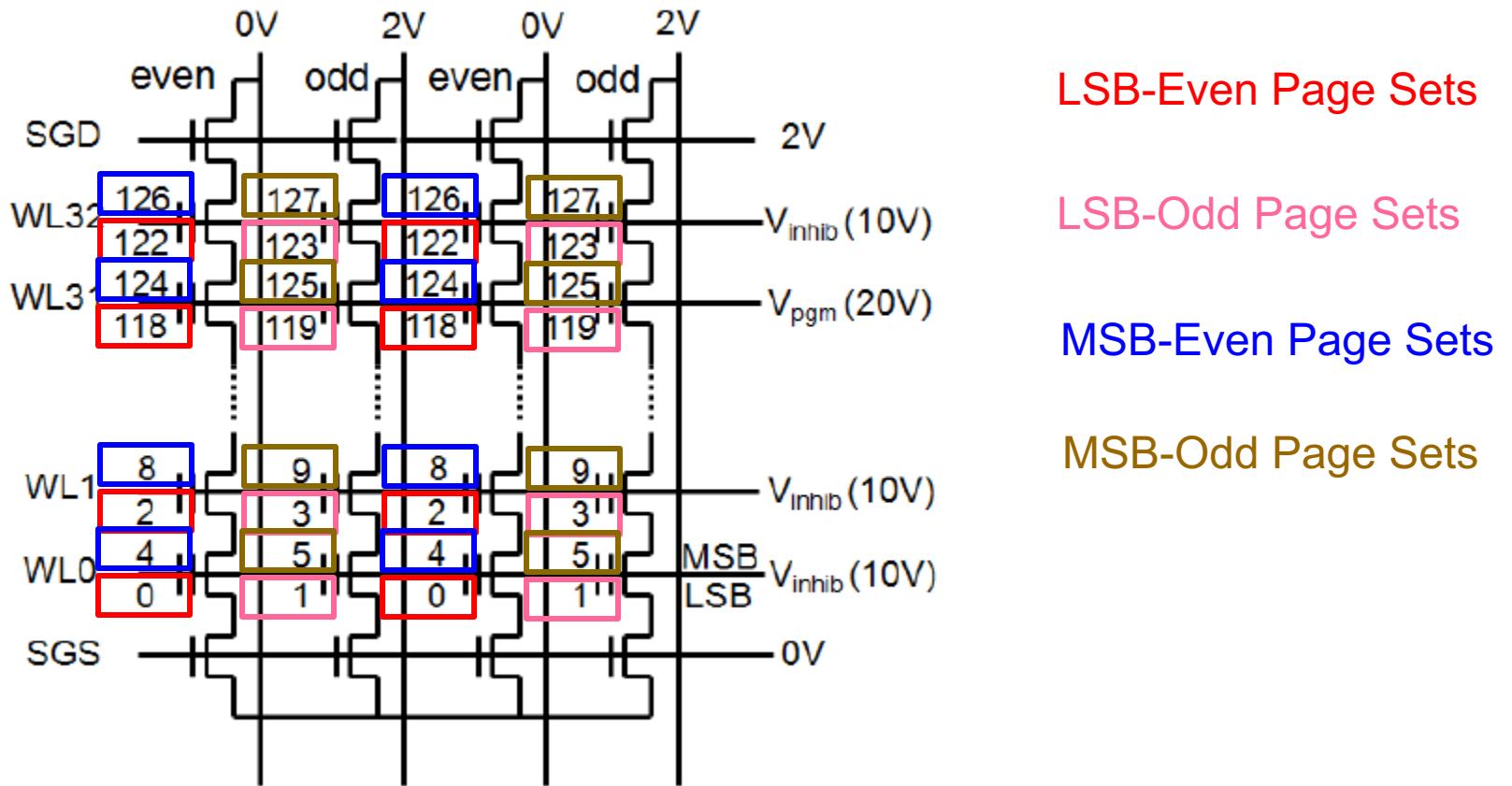


How Current Flash Cells are Programmed

- Programming 2-bit MLC NAND flash memory in two steps



MLC Architecture



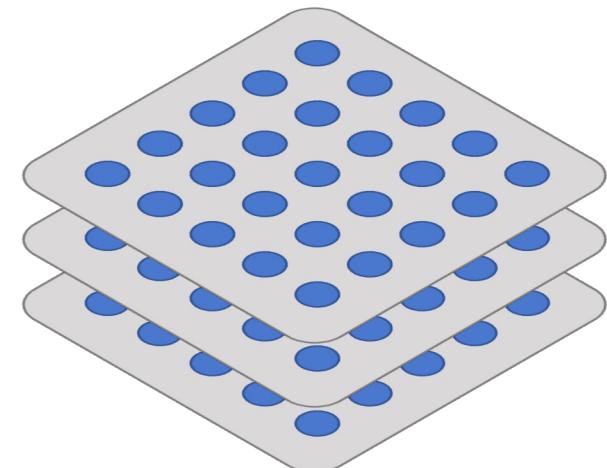
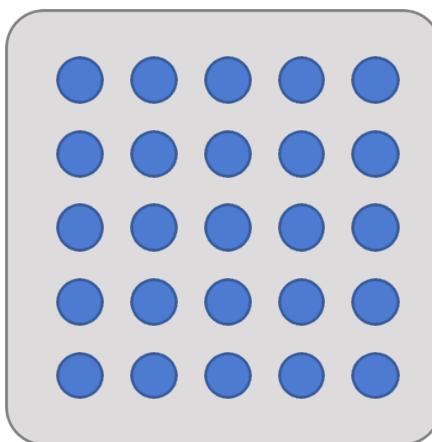
LSB-Even Page Sets

LSB-Odd Page Sets

MSB-Even Page Sets

MSB-Odd Page Sets

Planar vs. 3D NAND Flash Memory



**Planar NAND
Flash Memory**

Scaling

Reduce flash cell size,
Reduce distance b/w cells

Reliability

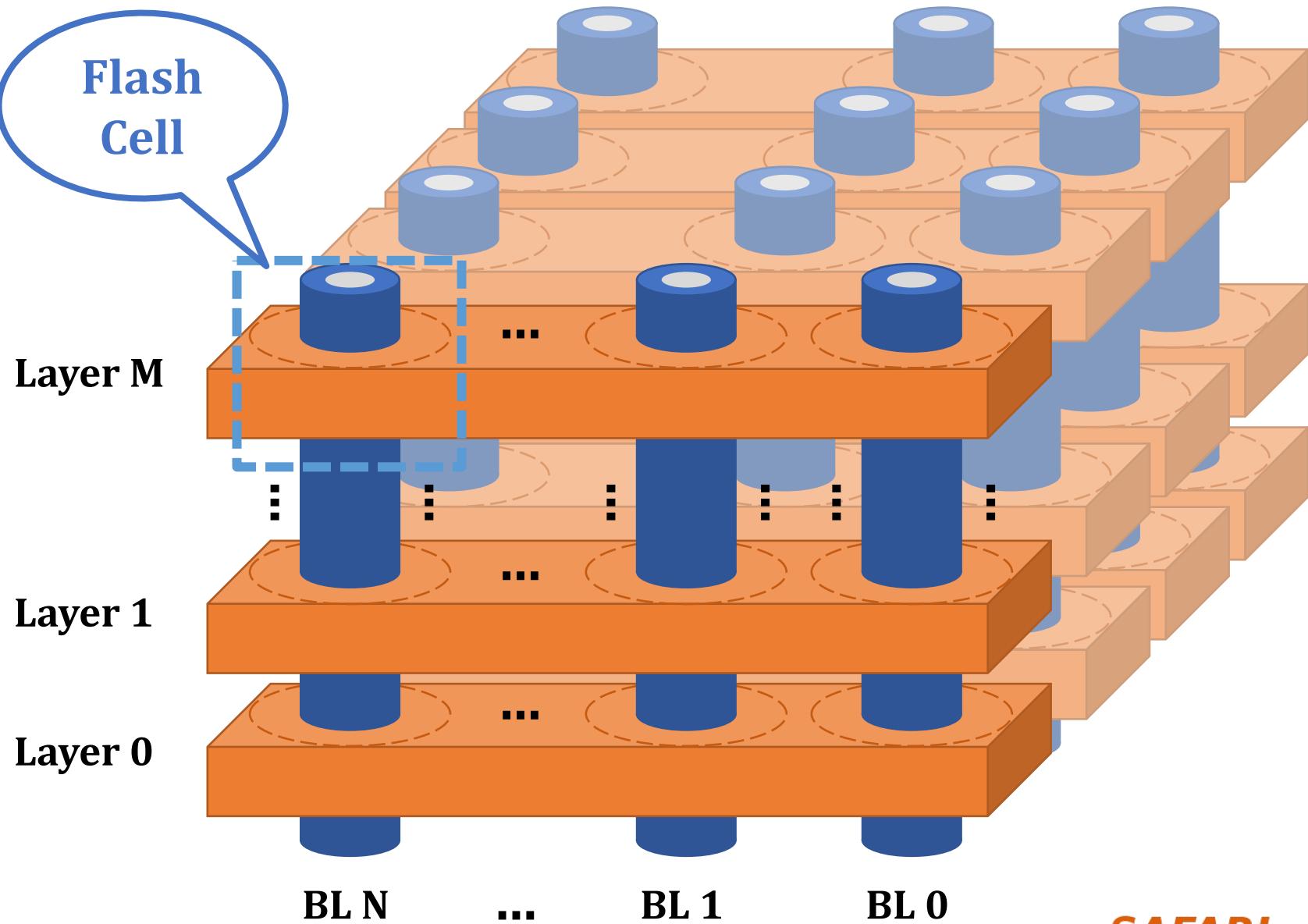
Scaling hurts reliability

**3D NAND
Flash Memory**

Increase # of layers

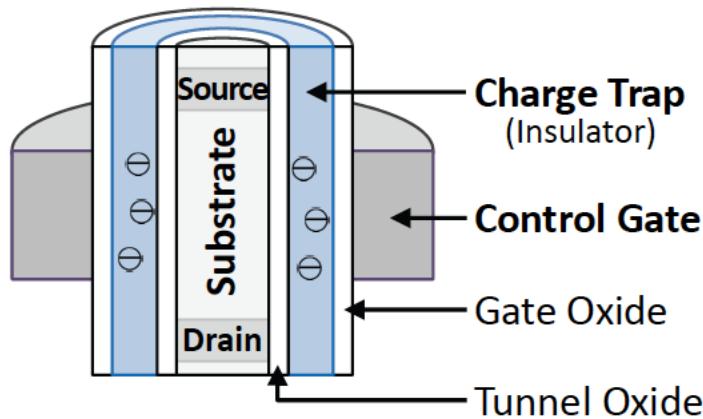
Not well studied!

3D NAND Flash Memory Structure



Charge Trap Based 3D Flash Cell

■ Cross-section of a charge trap transistor



3D NAND Flash Memory Organization

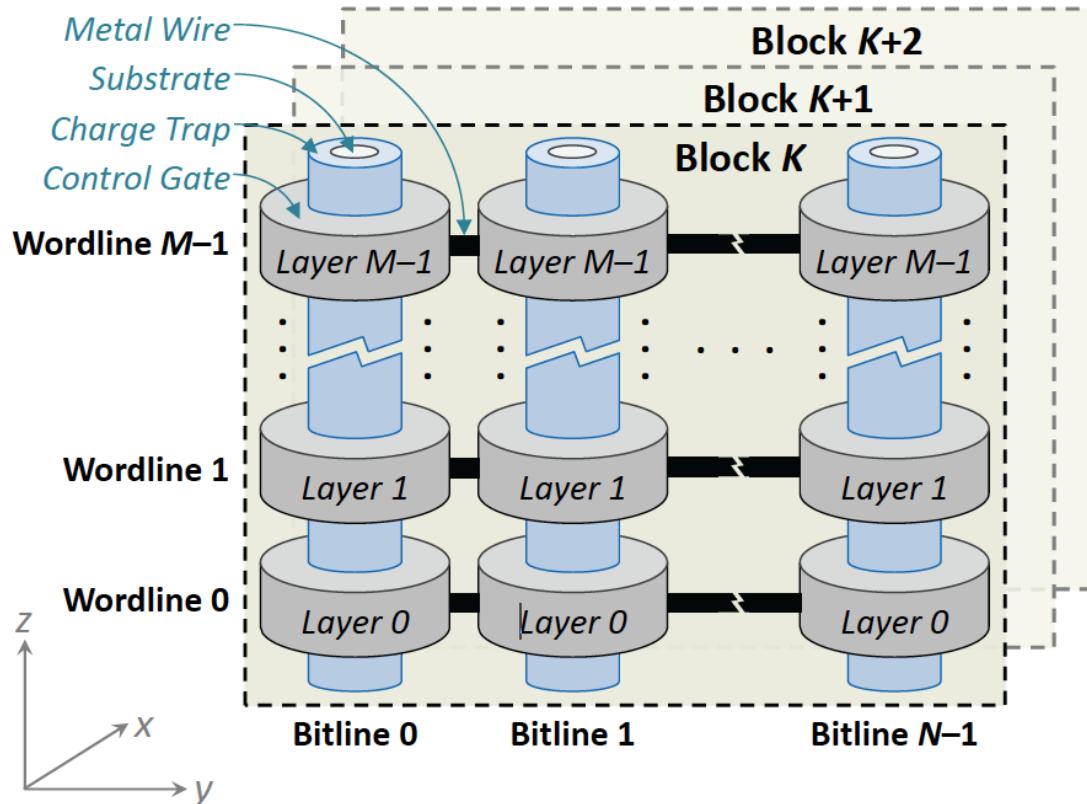


Fig. 43. Organization of flash cells in an M -layer 3D charge trap NAND flash memory chip, where each block consists of M wordlines and N bitlines.

More Background and State-of-the-Art



Proceedings of the IEEE, Sept. 2017

Error Characterization, Mitigation, and Recovery in Flash-Memory-Based Solid-State Drives

This paper reviews the most recent advances in solid-state drive (SSD) error characterization, mitigation, and data recovery techniques to improve both SSD's reliability and lifetime.

By YU CAI, SAUGATA GHOSE, ERICH F. HARATSCH, YIXIN LUO, AND ONUR MUTLU

<https://arxiv.org/pdf/1706.08642>



More Up-to-date Version

- Yu Cai, Saugata Ghose, Erich F. Haratsch, Yixin Luo, and Onur Mutlu,
"Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery"

Invited Book Chapter in Inside Solid State Drives, 2018.

[Preliminary arxiv.org version]

Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery

YU CAI, SAUGATA GHOSE

Carnegie Mellon University

ERICH F. HARATSCH

Seagate Technology

YIXIN LUO

Carnegie Mellon University

ONUR MUTLU

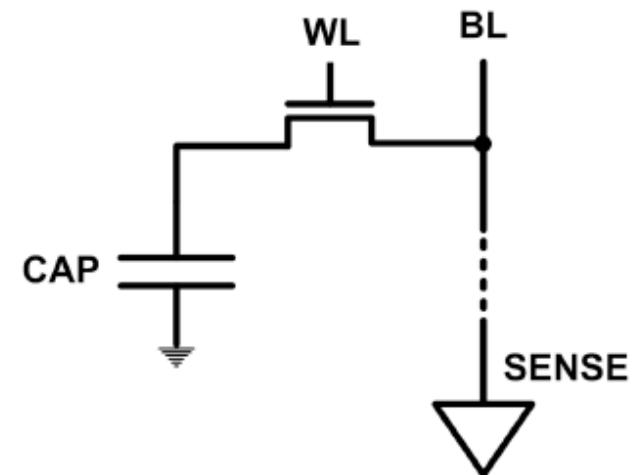
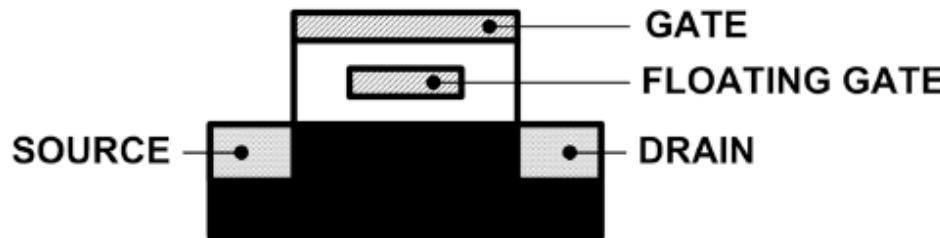
ETH Zürich and Carnegie Mellon University

Flash Memory Reliability and Security

Error Analysis and Management of NAND Flash Memory

Limits of Charge Memory

- Difficult charge placement and control
 - Flash: floating gate charge
 - DRAM: capacitor charge, transistor leakage
- Reliable sensing becomes difficult as charge storage unit size reduces



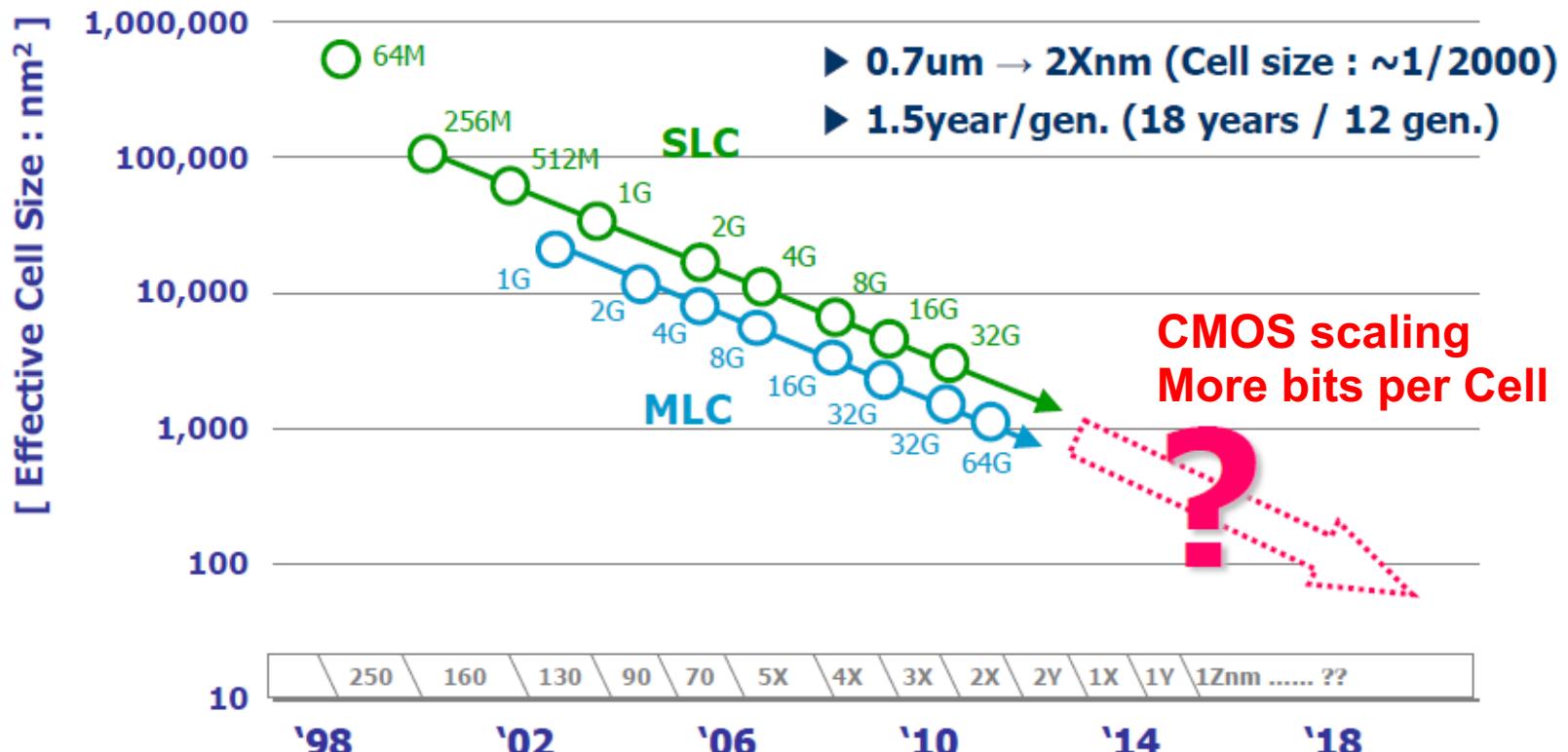
Executive Summary

- Problem: MLC NAND flash memory reliability/endurance is a key challenge for satisfying future storage systems' requirements
- Our Goals: (1) Build reliable error models for NAND flash memory via experimental characterization, (2) Develop efficient techniques to improve reliability and endurance
- This lecture provides a “flash” summary of our recent results published in the past 8 years:
 - Experimental error and threshold voltage characterization [DATE'12&13]
 - Retention-aware error management [ICCD'12]
 - Program interference analysis and read reference V prediction [ICCD'13]
 - Neighbor-assisted error correction [SIGMETRICS'14]
 - Read disturb error handling [DSN'15]
 - Data retention error handling [HPCA'15]

Agenda

- Background, Motivation and Approach
 - Experimental Characterization Methodology
 - Error Analysis and Management
 - Characterization Results
 - Retention-Aware Error Management
 - Threshold Voltage and Program Interference Analysis
 - Read Reference Voltage Prediction
 - Neighbor-Assisted Error Correction
 - Read Disturb Error Handling
 - Retention Error Handling
 - Large Scale Field Analysis
 - 3D NAND Flash Memory Reliability
 - Summary
-

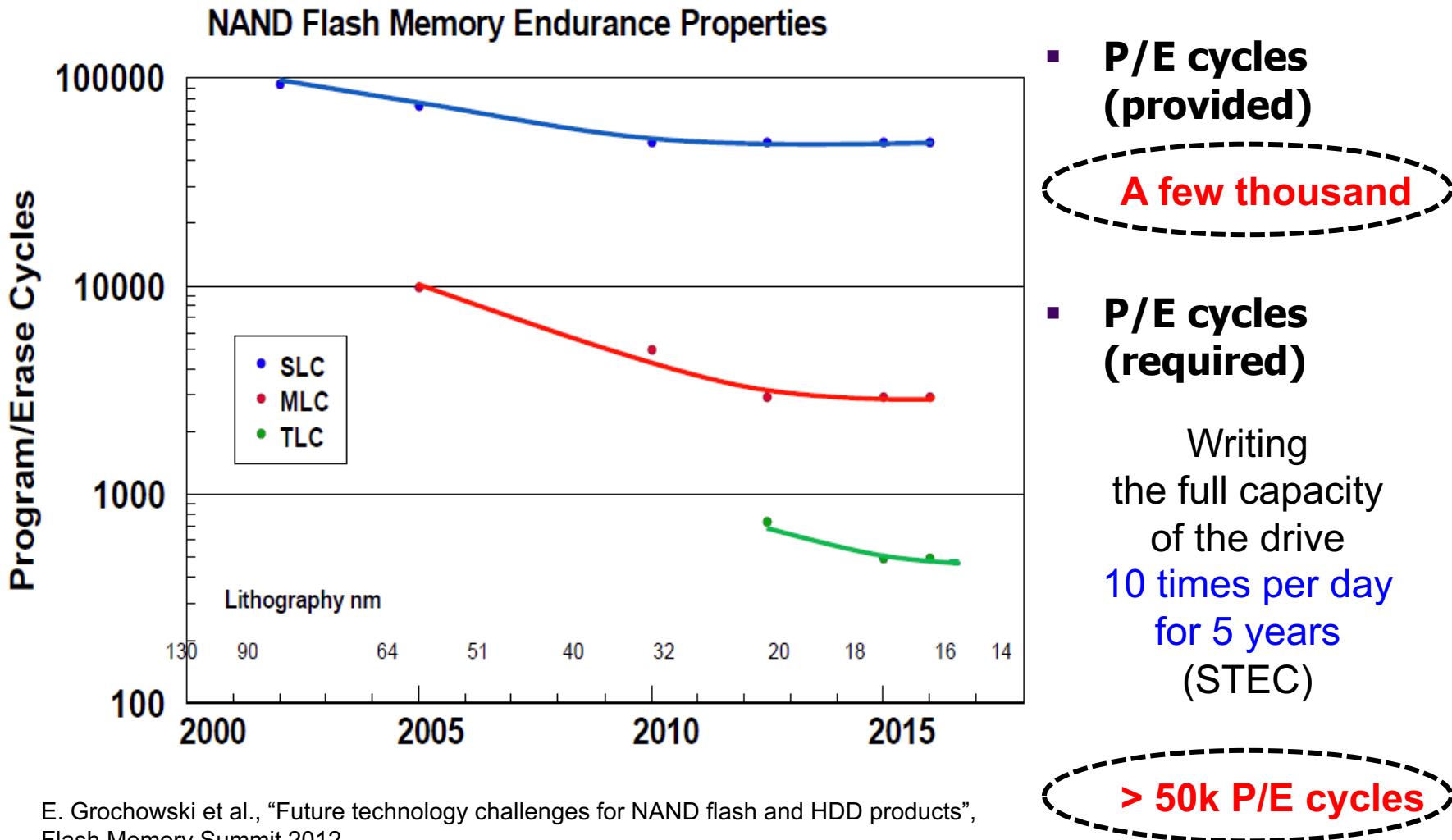
Evolution of NAND Flash Memory



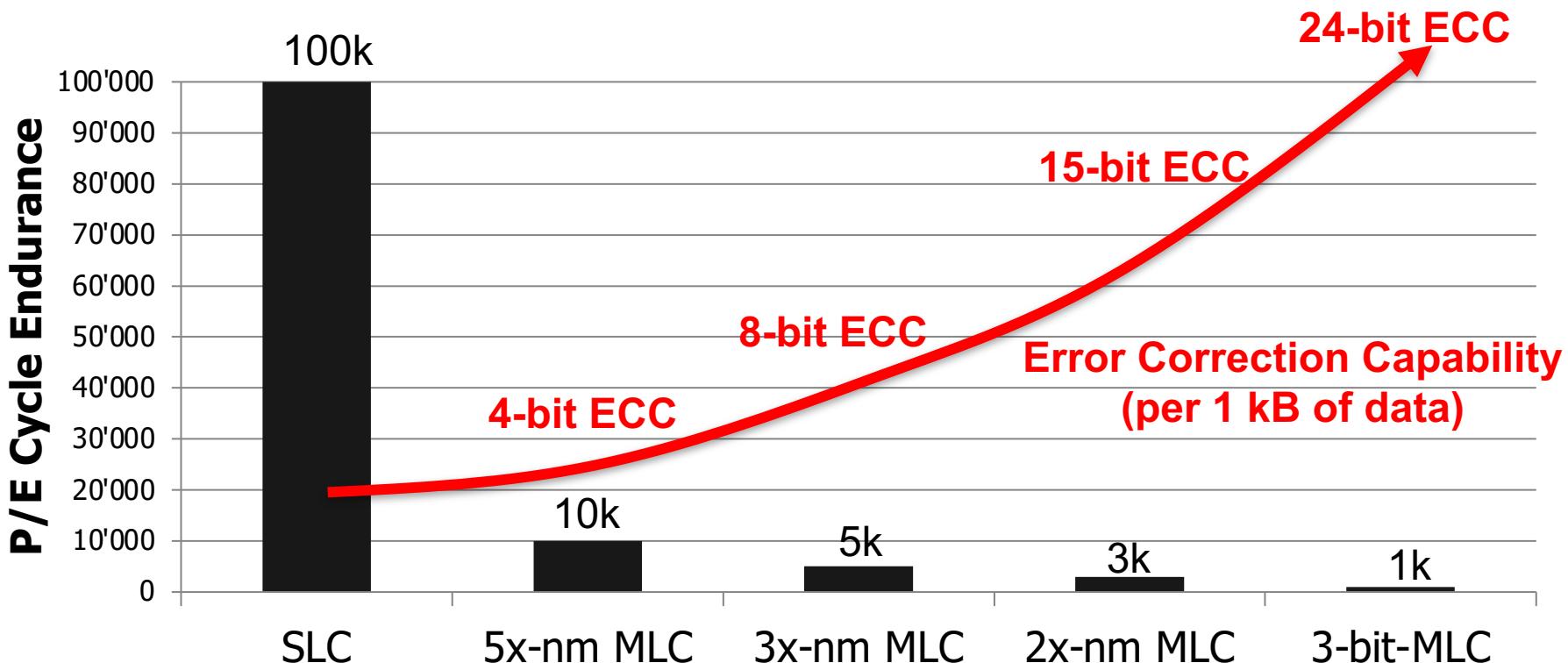
Seung Suk Lee, "Emerging Challenges in NAND Flash Technology", Flash Summit 2011 (Hynix)

- Flash memory is widening its range of applications
 - Portable consumer devices, laptop PCs and enterprise servers

Flash Challenges: Reliability and Endurance



Decreasing Endurance with Flash Scaling

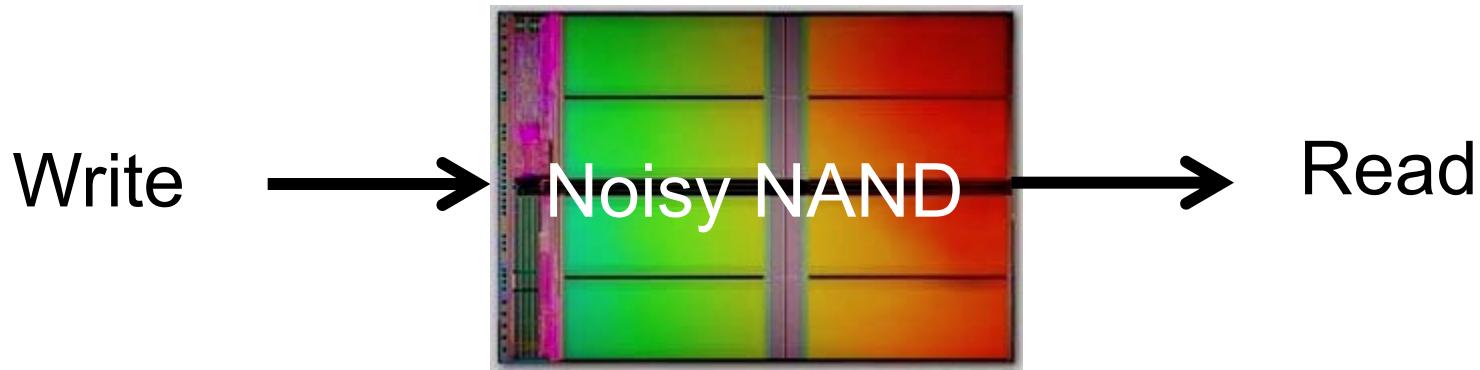


Ariel Maislos, "A New Era in Embedded Flash Memory", Flash Summit 2011 (Anobit)

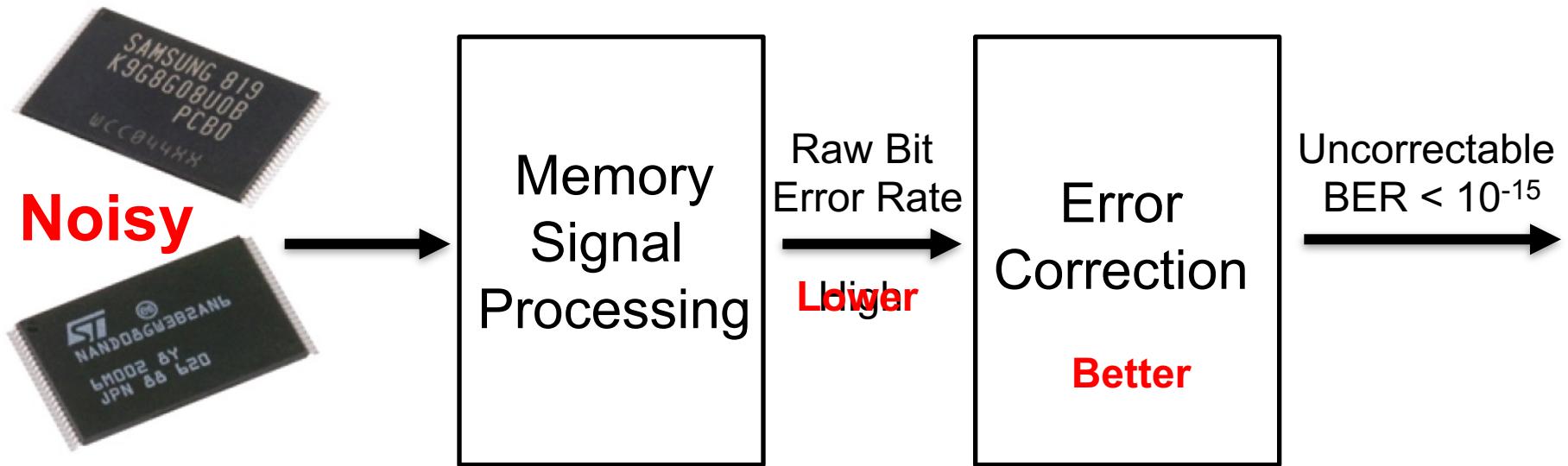
- Endurance of flash memory decreasing with scaling and multi-level cells
- Error correction capability required to guarantee storage-class reliability ($\text{UBER} < 10^{-15}$) is increasing exponentially to reach /less endurance

UBER: Uncorrectable bit error rate. Fraction of erroneous bits after error correction.

NAND Flash Memory is Increasingly Noisy



Future NAND Flash-based Storage Architecture

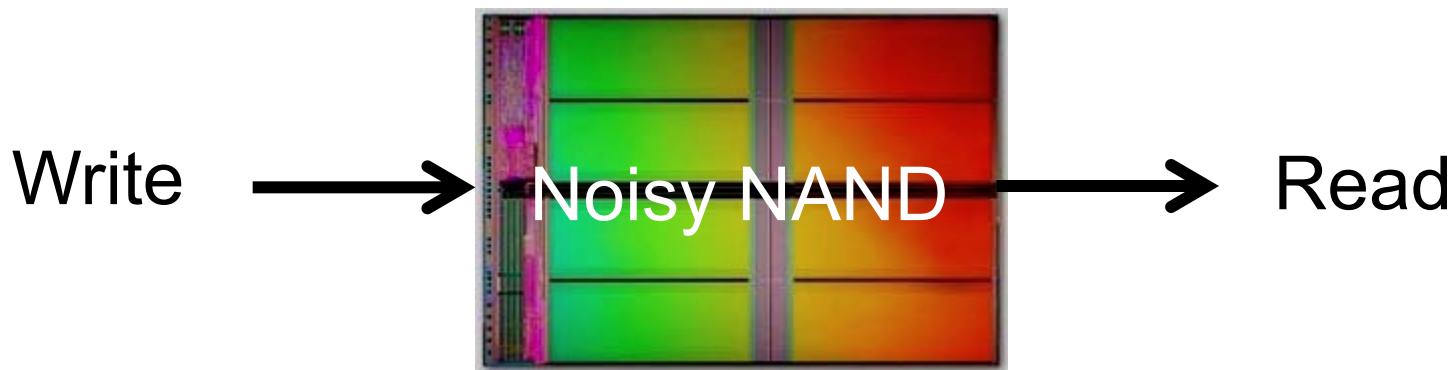


Our Goals:

Build reliable error models for NAND flash memory

Design efficient reliability mechanisms based on the model

NAND Flash Error Model



Experimentally characterize and model dominant errors

Cai et al., "Error Patterns in MLC NAND Flash Memory: Measurement, Characterization, and Analysis", **DATE 2012**
Luo et al., "Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory", **JSAC 2016**



Cai et al., "Threshold voltage distribution in MLC NAND Flash Memory: Characterization, Analysis, and Modeling", **DATE 2013**

Cai et al., "Vulnerabilities in MLC NAND Flash Memory Programming: Experimental Analysis, Exploits, and Mitigation Techniques", **HPCA 2017**

Cai et al., "Program Interference in MLC NAND Flash Memory: Characterization, Modeling, and Mitigation", **ICCD 2013**

Cai et al., "Neighbor-Cell Assisted Error Correction in MLC NAND Flash Memories", **SIGMETRICS 2014**

Cai et al., "Read Disturb Errors in MLC NAND Flash Memory: Characterization and Mitigation", **DSN 2015**

Cai et al., "Flash Correct-and-Refresh: Retention-aware error management for increased flash memory lifetime", **ICCD 2012**

Cai et al., "Error Analysis and Retention-Aware Error Management for NAND Flash Memory", **ITJ 2013**

Cai et al., "Data Retention in MLC NAND Flash Memory: Characterization, Optimization and Recovery", **HPCA 2015**

Our Goals and Approach

- Goals:
 - Understand error mechanisms and develop reliable predictive models for MLC NAND flash memory errors
 - Develop efficient error management techniques to mitigate errors and improve flash reliability and endurance

- Approach:
 - Solid experimental analyses of errors in real MLC NAND flash memory → drive the understanding and models
 - Understanding, models, and creativity → drive the new techniques

Many Errors and Their Mitigation [PIEEE'17]

Table 3 List of Different Types of Errors Mitigated by NAND Flash Error Mitigation Mechanisms

Mitigation Mechanism	Error Type				
	P/E Cycling [32,33,42] (§IV-A)	Program [40,42,53] (§IV-B)	Cell-to-Cell Interference [32,35,36,55] (§IV-C)	Data Retention [20,32,34,37,39] (§IV-D)	Read Disturb [20,32,38,62] (§IV-E)
Shadow Program Sequencing [35,40] (Section V-A)		X			
Neighbor-Cell Assisted Error Correction [36] (Section V-B)		X			
Refresh [34,39,67,68] (Section V-C)				X	X
Read-Retry [33,72,107] (Section V-D)	X			X	X
Voltage Optimization [37,38,74] (Section V-E)	X			X	X
Hot Data Management [41,63,70] (Section V-F)	X	X	X	X	X
Adaptive Error Mitigation [43,65,77,78,82] (Section V-G)	X	X	X	X	X

Cai+, "Error Characterization, Mitigation, and Recovery in Flash Memory Based Solid State Drives," Proc. IEEE 2017.

Many Errors and Their Mitigation [PIEEE'17]



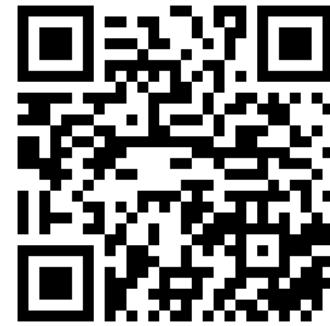
Proceedings of the IEEE, Sept. 2017

Error Characterization, Mitigation, and Recovery in Flash-Memory-Based Solid-State Drives

This paper reviews the most recent advances in solid-state drive (SSD) error characterization, mitigation, and data recovery techniques to improve both SSD's reliability and lifetime.

By YU CAI, SAUGATA GHOSE, ERICH F. HARATSCH, YIXIN LUO, AND ONUR MUTLU

<https://arxiv.org/pdf/1706.08642>



More Up-to-date Version

- Yu Cai, Saugata Ghose, Erich F. Haratsch, Yixin Luo, and Onur Mutlu,
"Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery"

Invited Book Chapter in Inside Solid State Drives, 2018.

[Preliminary arxiv.org version]

Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery

YU CAI, SAUGATA GHOSE

Carnegie Mellon University

ERICH F. HARATSCH

Seagate Technology

YIXIN LUO

Carnegie Mellon University

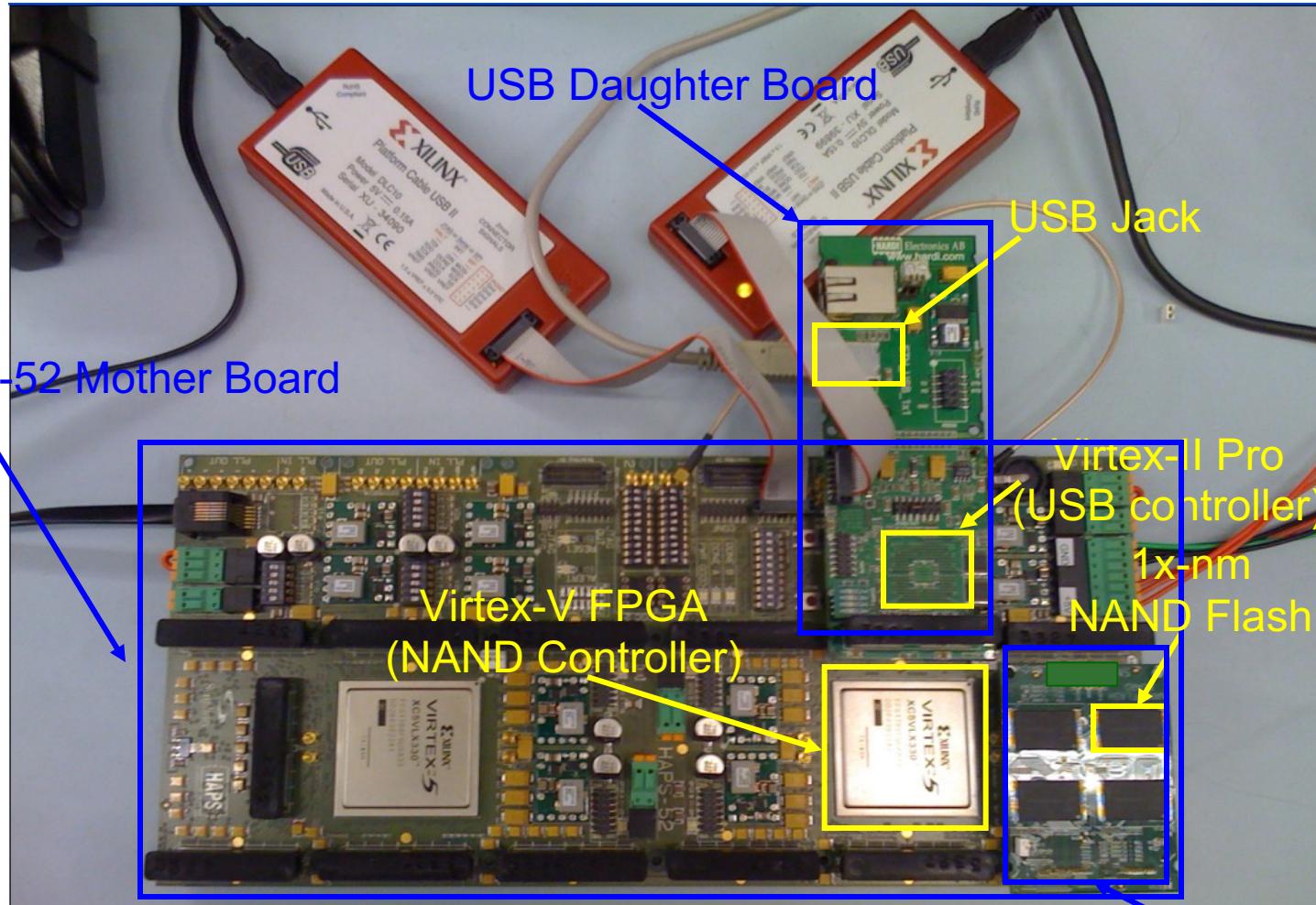
ONUR MUTLU

ETH Zürich and Carnegie Mellon University

Agenda

- Background, Motivation and Approach
 - Experimental Characterization Methodology
 - Error Analysis and Management
 - Main Characterization Results
 - Retention-Aware Error Management
 - Threshold Voltage and Program Interference Analysis
 - Read Reference Voltage Prediction
 - Neighbor-Assisted Error Correction
 - Read Disturb Error Handling
 - Retention Error Handling
 - Large Scale Field Analysis
 - 3D NAND Flash Memory Reliability
 - Summary
-

Experimental Testing Platform



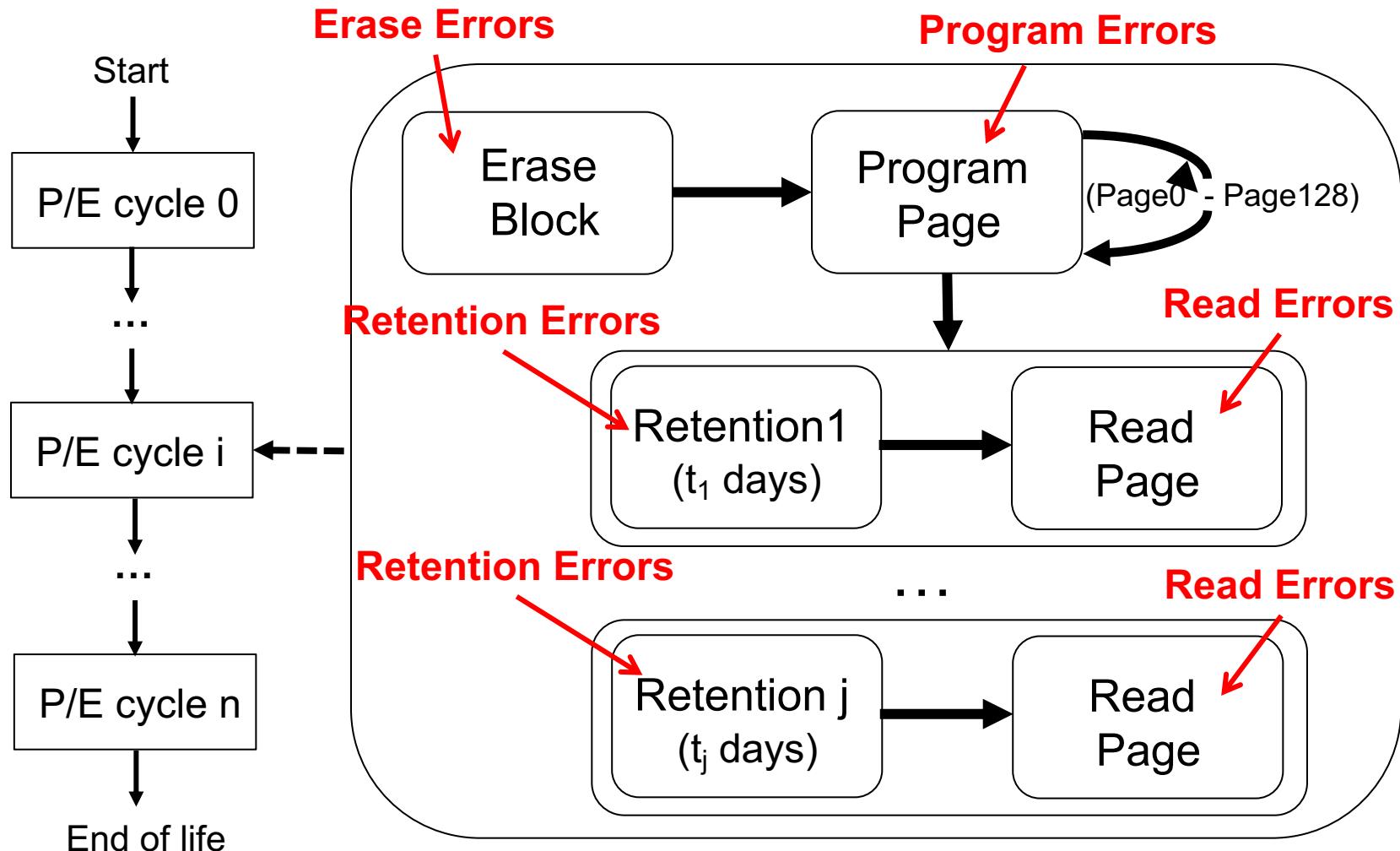
[DATE 2012, ICCD 2012, DATE 2013, ITJ 2013, ICCD 2013, SIGMETRICS 2014, HPCA 2015, DSN 2015, MSST 2015, JSAC 2016, HPCA 2017, DFRWS 2017, PIEEE 2017, HPCA 2018, SIGMETRICS 2018]

NAND Daughter Board

NAND Flash Error Types

- Four types of errors [Cai+, DATE 2012]
- Caused by common flash operations
 - Read errors
 - Erase errors
 - Program (interference) errors
- Caused by flash cell losing charge over time
 - Retention errors
 - Whether an error happens depends on required retention time
 - Especially problematic in MLC flash because threshold voltage window to determine stored value is smaller

NAND Flash Usage and Error Model



Methodology: Error and ECC Analysis

- Characterized errors and error rates of 3x and 2y-nm MLC NAND flash using an experimental FPGA-based platform
 - [Cai+, DATE'12, ICCD'12, DATE'13, ITJ'13, ICCD'13, SIGMETRICS'14]
- Quantified Raw Bit Error Rate (RBER) at a given P/E cycle
 - Raw Bit Error Rate: Fraction of erroneous bits without any correction
- Quantified error correction capability (and area and power consumption) of various BCH-code implementations
 - Identified how much RBER each code can tolerate
 - how many P/E cycles (flash lifetime) each code can sustain

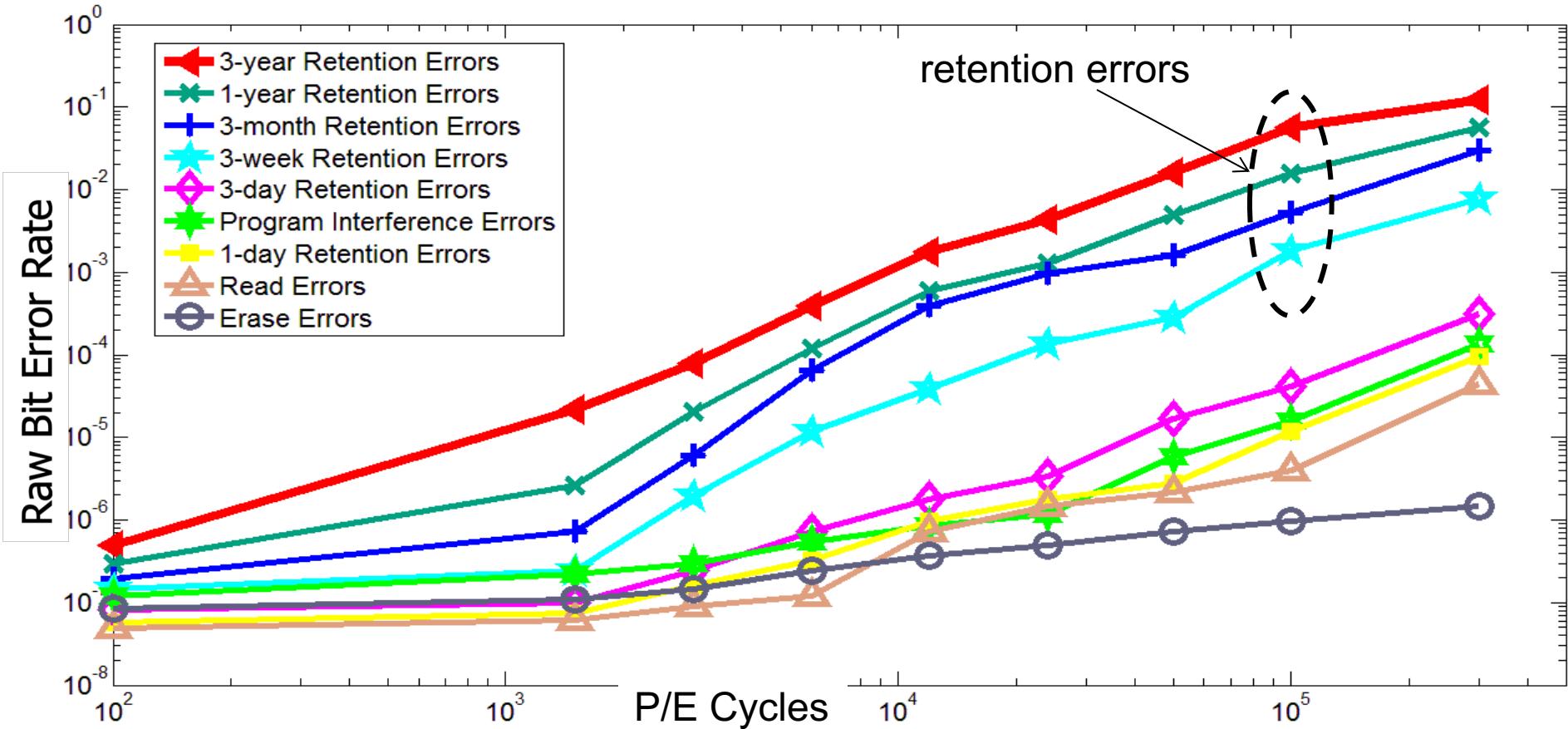
Agenda

- Background, Motivation and Approach
 - Experimental Characterization Methodology
 - Error Analysis and Management
 - Main Characterization Results
 - Retention-Aware Error Management
 - Threshold Voltage and Program Interference Analysis
 - Read Reference Voltage Prediction
 - Neighbor-Assisted Error Correction
 - Read Disturb Error Handling
 - Retention Error Handling
 - Large Scale Field Analysis
 - 3D NAND Flash Memory Reliability
 - Summary
-

Error Types and Testing Methodology

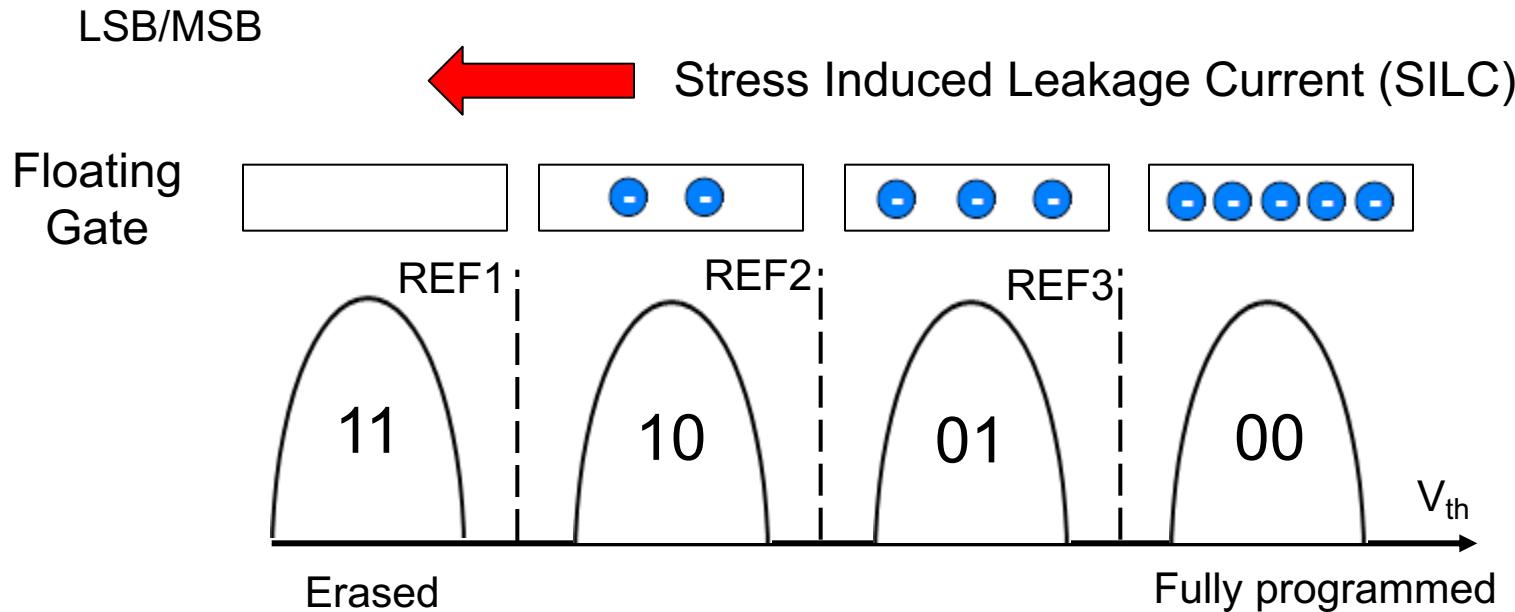
- Erase errors
 - Count the number of cells that fail to be erased to “11” state
- Program interference errors
 - Compare the data immediately after page programming and the data after the whole block being programmed
- Read errors
 - Continuously read a given block and compare the data between consecutive read sequences
- Retention errors
 - Compare the data read after an amount of time to data written
 - Characterize short term retention errors under room temperature
 - Characterize long term retention errors by baking in the oven under 125°C

Observations: Flash Error Analysis



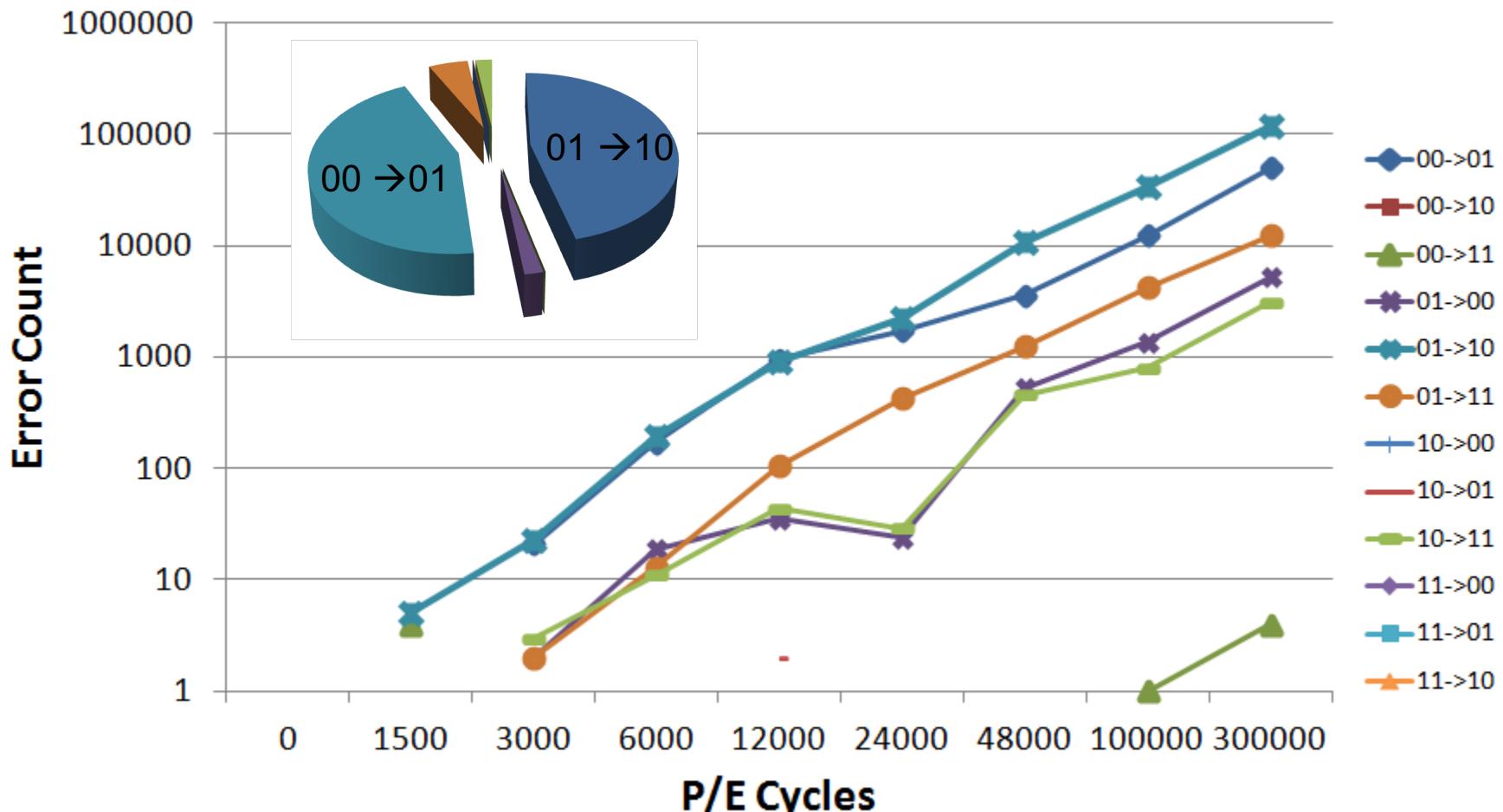
- Raw bit error rate increases exponentially with P/E cycles
- Retention errors are dominant (>99% for 1-year ret. time)
- Retention errors increase with retention time requirement

Retention Error Mechanism



- Electron loss from the floating gate causes retention errors
 - Cells with more programmed electrons suffer more from retention errors
 - Threshold voltage is more likely to shift by one window than by multiple

Retention Error Value Dependency



- Cells with more programmed electrons tend to suffer more from retention noise (i.e. 00 and 01)

More on Flash Error Analysis

- Yu Cai, Erich F. Haratsch, Onur Mutlu, and Ken Mai,
**"Error Patterns in MLC NAND Flash Memory:
Measurement, Characterization, and Analysis"**
*Proceedings of the Design, Automation, and Test in Europe
Conference (DATE)*, Dresden, Germany, March 2012. Slides
(ppt)

Error Patterns in MLC NAND Flash Memory: Measurement, Characterization, and Analysis

Yu Cai¹, Erich F. Haratsch², Onur Mutlu¹ and Ken Mai¹

¹Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA

²LSI Corporation, 1110 American Parkway NE, Allentown, PA

¹{yucai, onur, kenmai}@andrew.cmu.edu, ²erich.haratsch@lsi.com

Agenda

- Background, Motivation and Approach
 - Experimental Characterization Methodology
 - Error Analysis and Management
 - Main Characterization Results
 - Retention-Aware Error Management
 - Threshold Voltage and Program Interference Analysis
 - Read Reference Voltage Prediction
 - Neighbor-Assisted Error Correction
 - Read Disturb Error Handling
 - Retention Error Handling
 - Large Scale Field Analysis
 - 3D NAND Flash Memory Reliability
 - Summary
-

Solution to Retention Errors

- Refresh periodically
- Change the period based on P/E cycle wearout
 - Refresh more often at higher P/E cycles
- Use a combination of **in-place** and **remapping-based** refresh
- Cai et al. “**Flash Correct-and-Refresh: Retention-Aware Error Management for Increased Flash Memory Lifetime**”, ICCD 2012.

Flash Correct-and-Refresh (FCR)

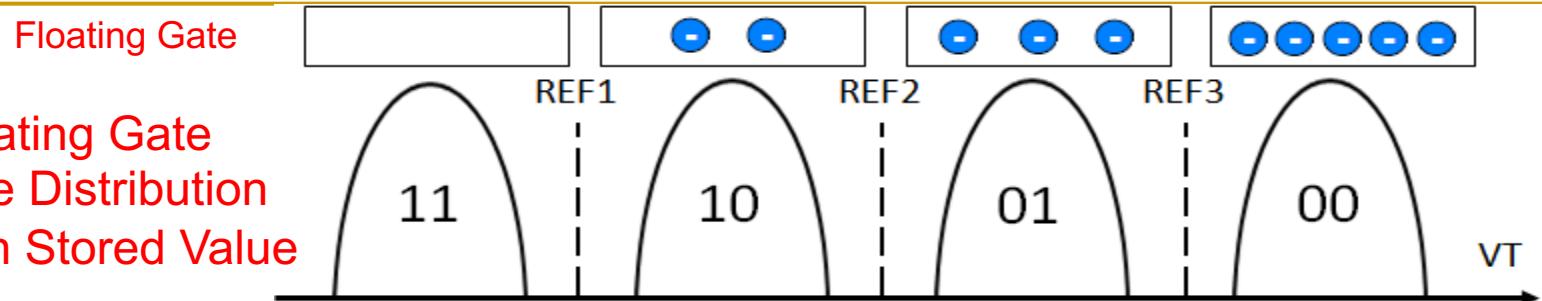
- Key Observations:
 - ❑ Retention errors are the dominant source of errors in flash memory [**Cai+ DATE 2012**][**Tanakamaru+ ISSCC 2011**]
→ limit flash lifetime as they increase over time
 - ❑ Retention errors can be corrected by “refreshing” each flash page periodically
- Key Idea:
 - ❑ Periodically read each flash page,
 - ❑ Correct its errors using “weak” ECC, and
 - ❑ Either remap it to a new physical page or reprogram it in-place,
 - ❑ Before the page accumulates more errors than ECC-correctable
 - ❑ Optimization: Adapt refresh rate to endured P/E cycles

FCR: Two Key Questions

- How to refresh?
 - Remap a page to another one
 - Reprogram a page (in-place)
 - Hybrid of remap and reprogram

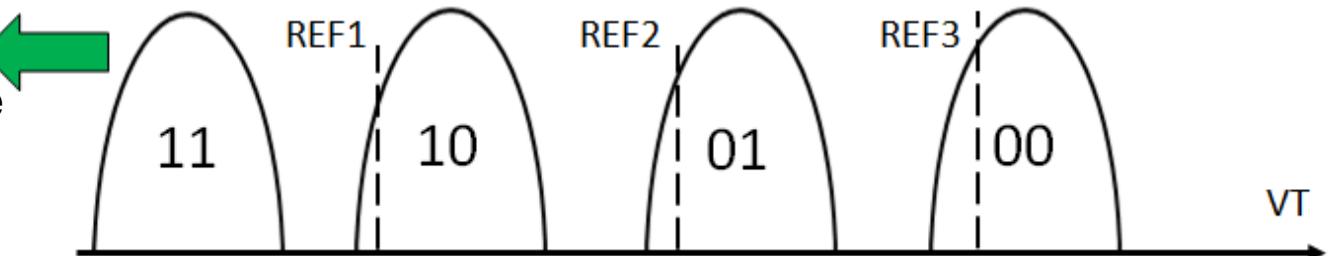
- When to refresh?
 - Fixed period
 - Adapt the period to retention error severity

In-Place Reprogramming of Flash Cells

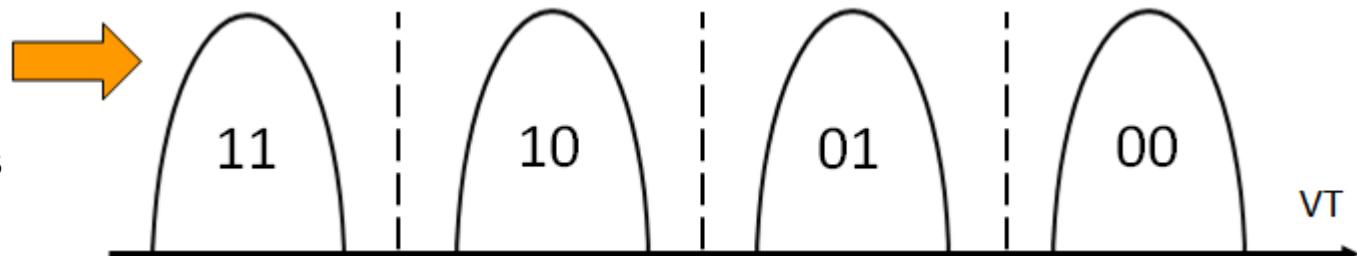


Floating Gate
Voltage Distribution
for each Stored Value

Retention errors are caused by cell voltage shifting to the left

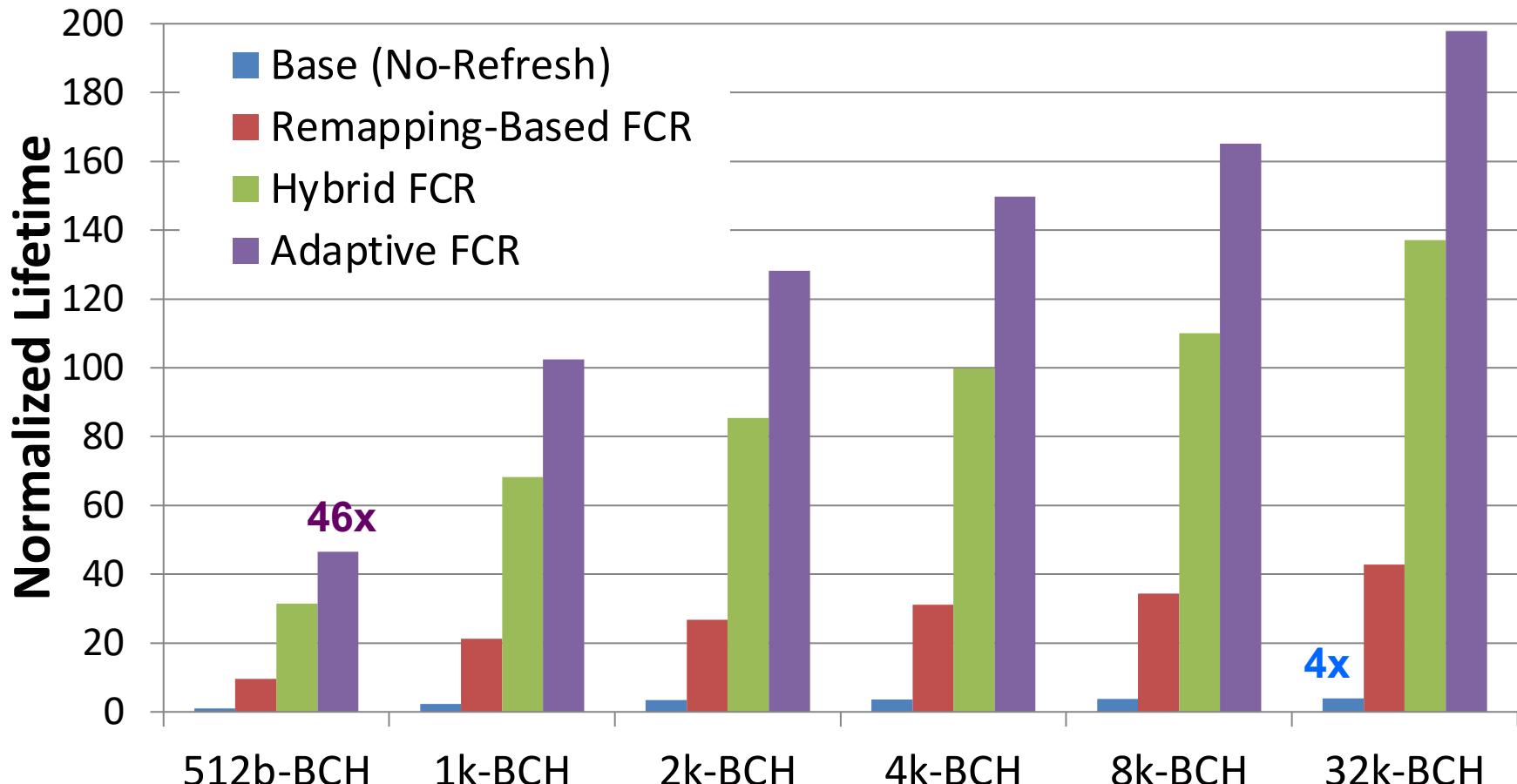


ISPP moves cell voltage to the right; fixes retention errors



- Pro: No remapping needed → no additional erase operations
- Con: Increases the occurrence of program errors

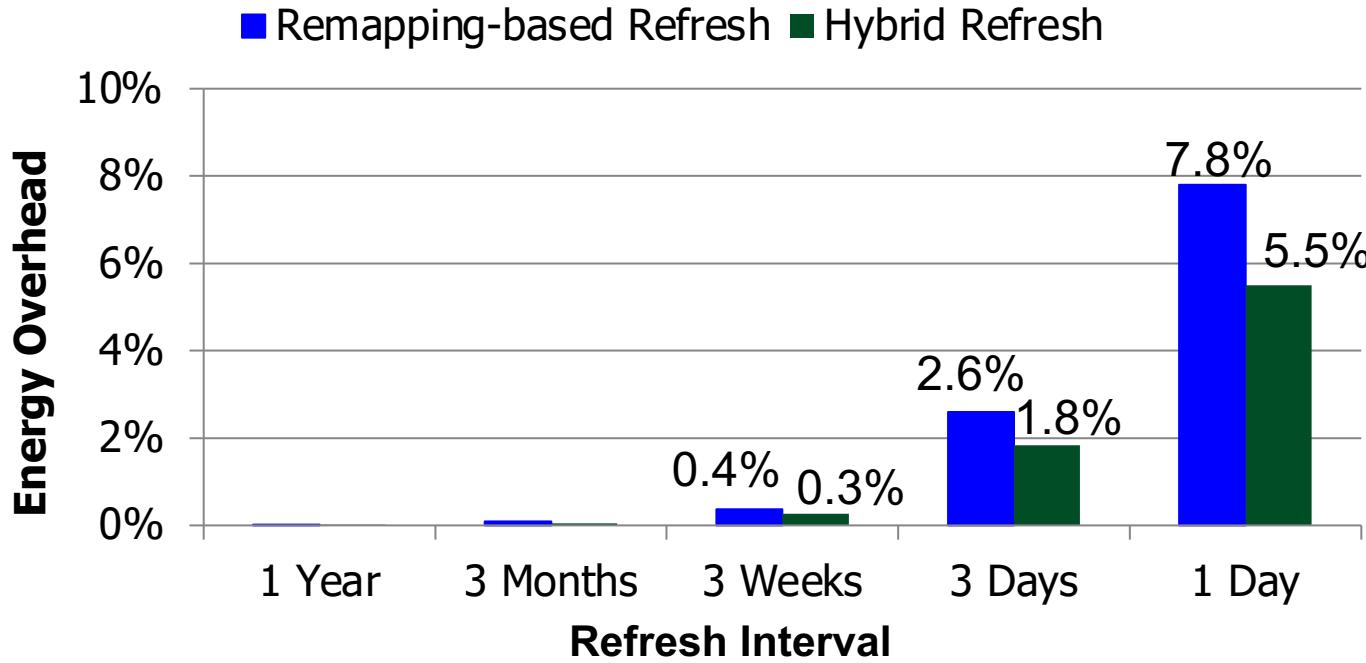
Normalized Flash Memory Lifetime



Adaptive-rate FCR provides the highest lifetime

Lifetime of FCR much higher than lifetime of stronger ECC

Energy Overhead



- Adaptive-rate refresh: <1.8% energy increase until daily refresh is triggered

Flash Correct-and-Refresh [ICCD'12]

- Yu Cai, Gulay Yalcin, Onur Mutlu, Erich F. Haratsch, Adrian Cristal, Osman Unsal, and Ken Mai,

"Flash Correct-and-Refresh: Retention-Aware Error Management for Increased Flash Memory Lifetime"

Proceedings of the 30th IEEE International Conference on Computer Design (ICCD), Montreal, Quebec, Canada, September 2012. Slides (ppt)(pdf)

Flash Correct-and-Refresh: Retention-Aware Error Management for Increased Flash Memory Lifetime

Yu Cai¹, Gulay Yalcin², Onur Mutlu¹, Erich F. Haratsch³, Adrian Cristal², Osman S. Unsal² and Ken Mai¹

¹DSSC, Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA

²Barcelona Supercomputing Center, C/Jordi Girona 29, Barcelona, Spain

³LSI Corporation, 1110 American Parkway NE, Allentown, PA

More Detail on Flash Error Analysis

- Yu Cai, Gulay Yalcin, Onur Mutlu, Erich F. Haratsch, Adrian Cristal, Osman Unsal, and Ken Mai,

"Error Analysis and Retention-Aware Error Management for NAND Flash Memory"

Intel Technology Journal (ITJ) Special Issue on Memory Resiliency, Vol. 17, No. 1, May 2013.

Intel® Technology Journal | Volume 17, Issue 1, 2013

ERROR ANALYSIS AND RETENTION-AWARE ERROR MANAGEMENT
FOR NAND FLASH MEMORY

Many Errors and Their Mitigation [PIEEE'17]

Table 3 List of Different Types of Errors Mitigated by NAND Flash Error Mitigation Mechanisms

Mitigation Mechanism	Error Type				
	P/E Cycling [32,33,42] (§IV-A)	Program [40,42,53] (§IV-B)	Cell-to-Cell Interference [32,35,36,55] (§IV-C)	Data Retention [20,32,34,37,39] (§IV-D)	Read Disturb [20,32,38,62] (§IV-E)
Shadow Program Sequencing [35,40] (Section V-A)		X			
Neighbor-Cell Assisted Error Correction [36] (Section V-B)		X			
Refresh [34,39,67,68] (Section V-C)				X	X
Read-Retry [33,72,107] (Section V-D)	X			X	X
Voltage Optimization [37,38,74] (Section V-E)	X			X	X
Hot Data Management [41,63,70] (Section V-F)	X	X	X	X	X
Adaptive Error Mitigation [43,65,77,78,82] (Section V-G)	X	X	X	X	X

Cai+, "Error Characterization, Mitigation, and Recovery in Flash Memory Based Solid State Drives," Proc. IEEE 2017.

Many Errors and Their Mitigation [PIEEE'17]



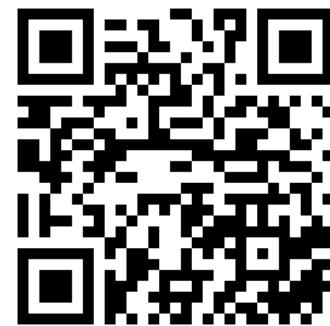
Proceedings of the IEEE, Sept. 2017

Error Characterization, Mitigation, and Recovery in Flash-Memory-Based Solid-State Drives

This paper reviews the most recent advances in solid-state drive (SSD) error characterization, mitigation, and data recovery techniques to improve both SSD's reliability and lifetime.

By YU CAI, SAUGATA GHOSE, ERICH F. HARATSCH, YIXIN LUO, AND ONUR MUTLU

<https://arxiv.org/pdf/1706.08642>



More Up-to-date Version

- Yu Cai, Saugata Ghose, Erich F. Haratsch, Yixin Luo, and Onur Mutlu,
"Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery"

Invited Book Chapter in Inside Solid State Drives, 2018.

[Preliminary arxiv.org version]

Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery

YU CAI, SAUGATA GHOSE

Carnegie Mellon University

ERICH F. HARATSCH

Seagate Technology

YIXIN LUO

Carnegie Mellon University

ONUR MUTLU

ETH Zürich and Carnegie Mellon University

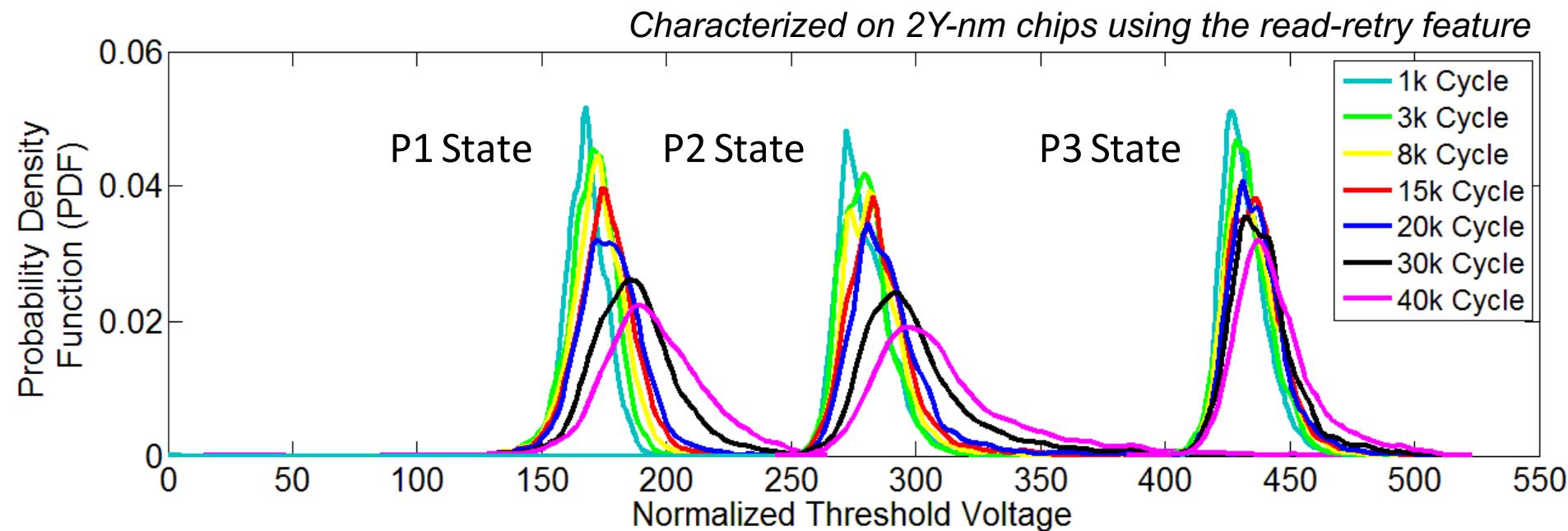
Agenda

- Background, Motivation and Approach
 - Experimental Characterization Methodology
 - Error Analysis and Management
 - Main Characterization Results
 - Retention-Aware Error Management
 - Threshold Voltage and Program Interference Analysis
 - Read Reference Voltage Prediction
 - Neighbor-Assisted Error Correction
 - Read Disturb Error Handling
 - Retention Error Handling
 - Large Scale Field Analysis
 - 3D NAND Flash Memory Reliability
 - Summary
-

Key Questions

- How does threshold voltage (V_{th}) distribution of different programmed states change over flash lifetime?
- Can we model it accurately and predict the V_{th} changes?
- Can we build mechanisms that can correct for V_{th} changes?
(thereby reducing read error rates)

Threshold Voltage Distribution Model



Gaussian distribution with additive white noise

As P/E cycles increase ...

- Distribution shifts to the right
- Distribution becomes wider

Threshold Voltage Distribution Model

- V_{th} distribution can be modeled with ~95% accuracy as a Gaussian distribution with additive white noise
- Distortion in V_{th} over P/E cycles can be modeled and predicted as an exponential function of P/E cycles
 - With more than 95% accuracy

More Detail on Threshold Voltage Model

- Yu Cai, Erich F. Haratsch, Onur Mutlu, and Ken Mai,
"Threshold Voltage Distribution in MLC NAND Flash Memory: Characterization, Analysis and Modeling"
Proceedings of the Design, Automation, and Test in Europe Conference (DATE), Grenoble, France, March 2013. [Slides \(ppt\)](#)

Threshold Voltage Distribution in MLC NAND Flash Memory: Characterization, Analysis, and Modeling

Yu Cai¹, Erich F. Haratsch², Onur Mutlu¹ and Ken Mai¹

¹DSSC, Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA

²LSI Corporation, 1110 American Parkway NE, Allentown, PA

¹{yucai, onur, kenmai}@andrew.cmu.edu, ²erich.haratsch@lsi.com

More Accurate and Online Channel Modeling

- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu,
**"Enabling Accurate and Practical Online Flash Channel Modeling
for Modern MLC NAND Flash Memory"**
to appear in IEEE Journal on Selected Areas in Communications (JSAC),
2016.

Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory

Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, Onur Mutlu

Non-Gaussian V_{th} Distributions (1X-nm)

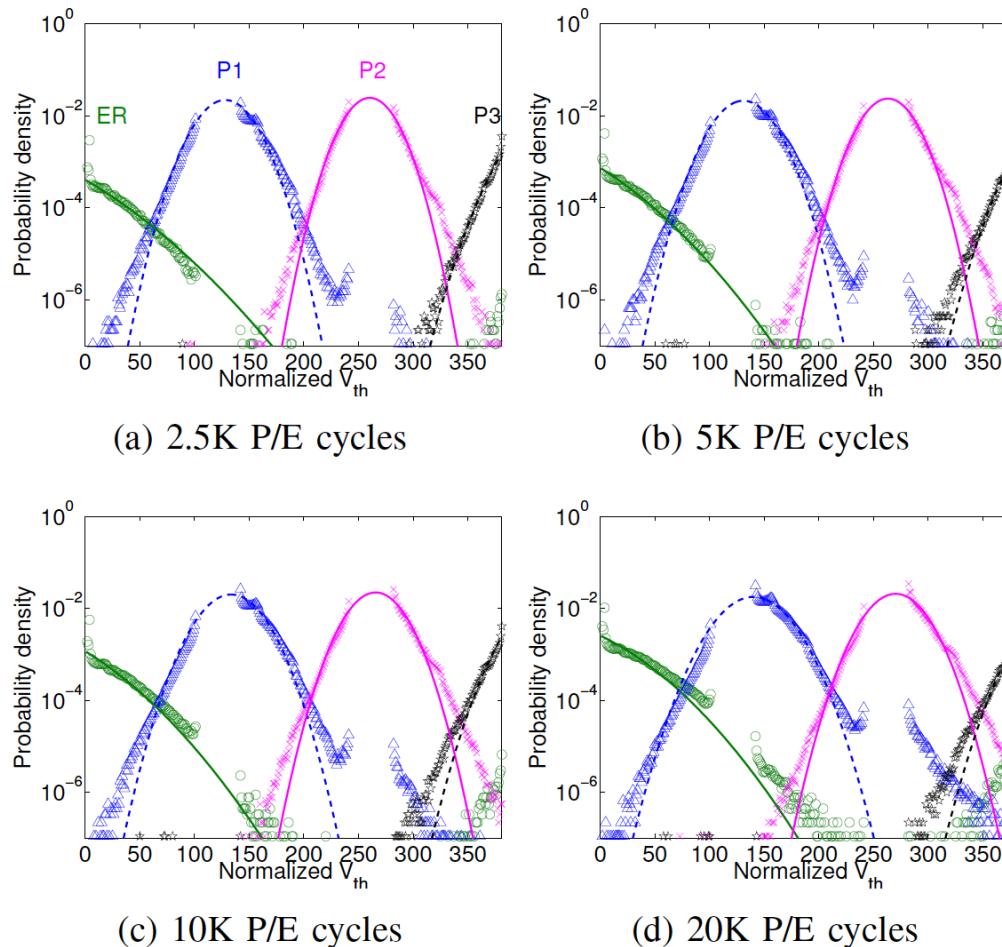


Fig. 4: Gaussian-based model (solid/dashed lines) vs. data measured from real NAND flash chips (markers) under different P/E cycle counts.

Better Modeling of V_{th} Distributions (I)

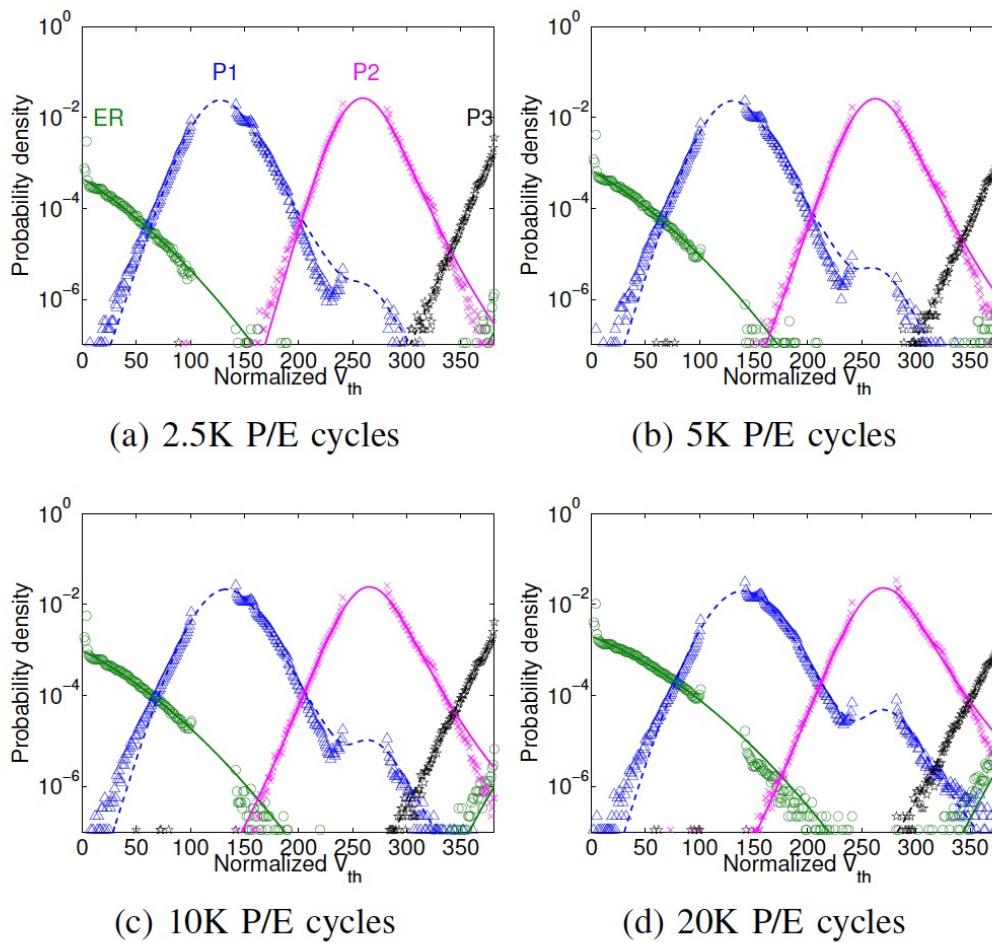


Fig. 6: Our new Student's t-based model (solid/dashed lines) vs. data measured from real NAND flash chips (markers) under different P/E cycle counts.

Better Modeling of V_{th} Distributions (II)

P/E Cycles	0	2.5K	5K	7.5K	10K	12K	14K	16K	18K	20K	AVG
Gaussian	.99%	1.8%	1.6%	1.8%	1.9%	2.4%	3.1%	8.7%	2.1%	2.3%	2.6%
Normal-Laplace	.34%	.46%	.55%	.61%	.63%	.67%	.68%	.70%	.67%	.67%	.61%
Student's t	.37%	.51%	.61%	.68%	.70%	.76%	.76%	.78%	.76%	.78%	.68%

TABLE 1: Modeling error of the evaluated threshold voltage distribution models, at various P/E cycle counts.

Prediction vs. Reality with Better Modeling

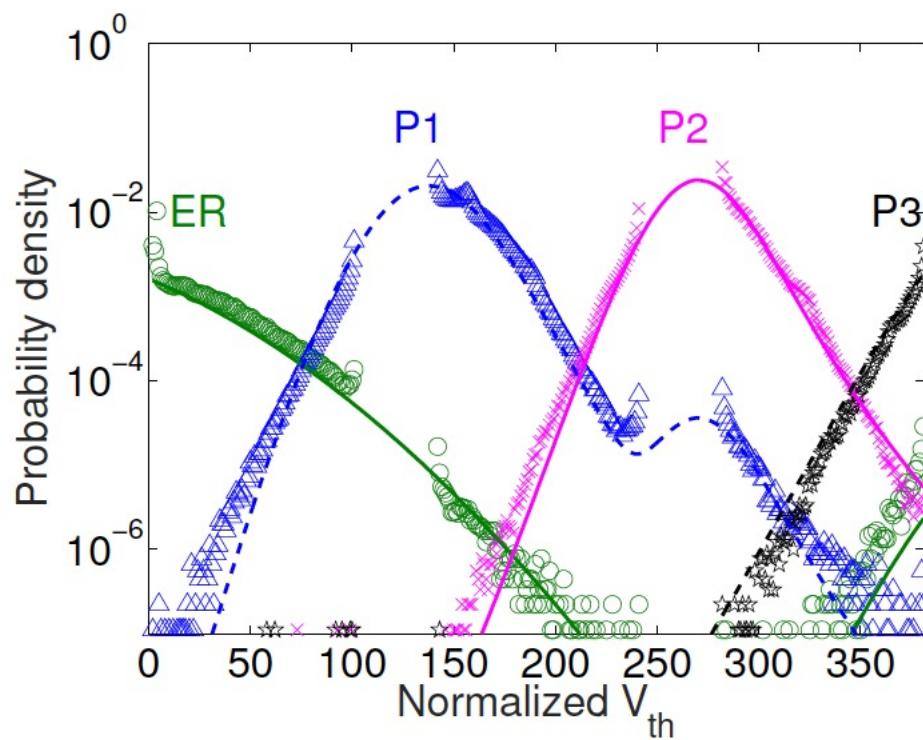


Fig. 13: Threshold voltage distribution as predicted by our dynamic model for 20K P/E cycles, using characterization data from 2.5K, 5K, 7.5K, and 10K P/E cycles, shown as solid/dashed lines. Markers represent data measured from real NAND flash chips at 20K P/E cycles.

More Accurate and Online Channel Modeling

- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu,
**"Enabling Accurate and Practical Online Flash Channel Modeling
for Modern MLC NAND Flash Memory"**
to appear in IEEE Journal on Selected Areas in Communications (JSAC),
2016.

Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory

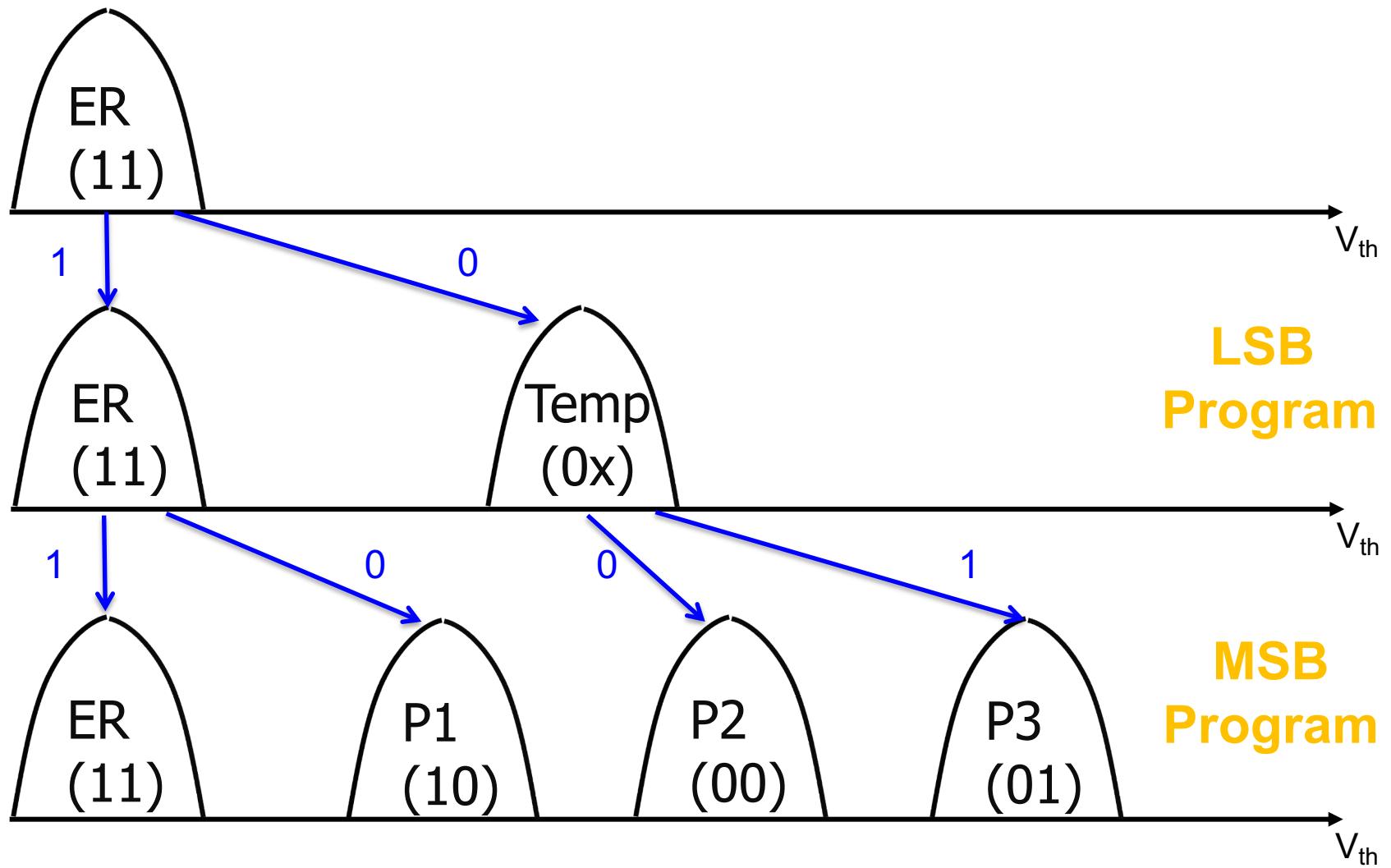
Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, Onur Mutlu

Program Interference Errors

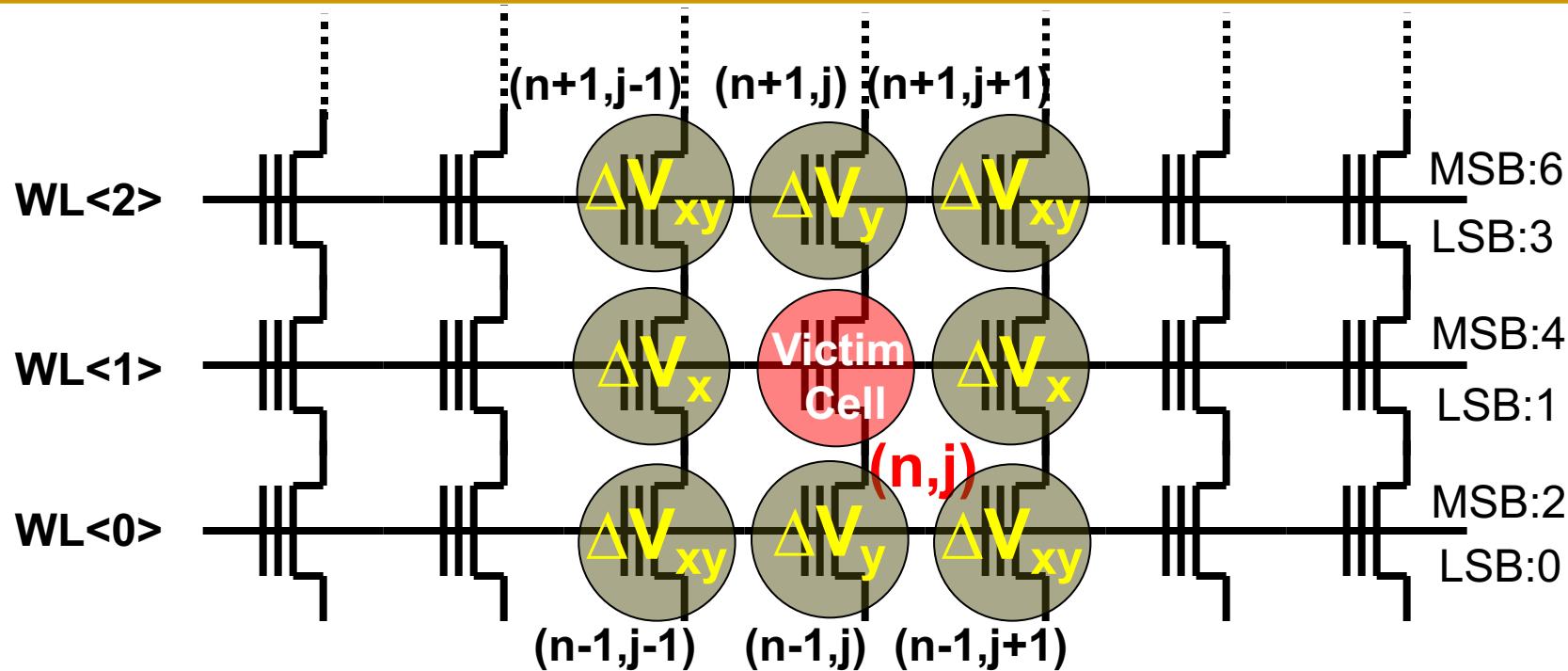
- When a cell is being programmed, **voltage level of a neighboring cell changes** (unintentionally) due to parasitic capacitance coupling
→ can change the data value stored
- Also called program interference error
- Causes neighboring cell voltage to increase (shift right)
- Once retention errors are minimized, these errors can become dominant

How Current Flash Cells are Programmed

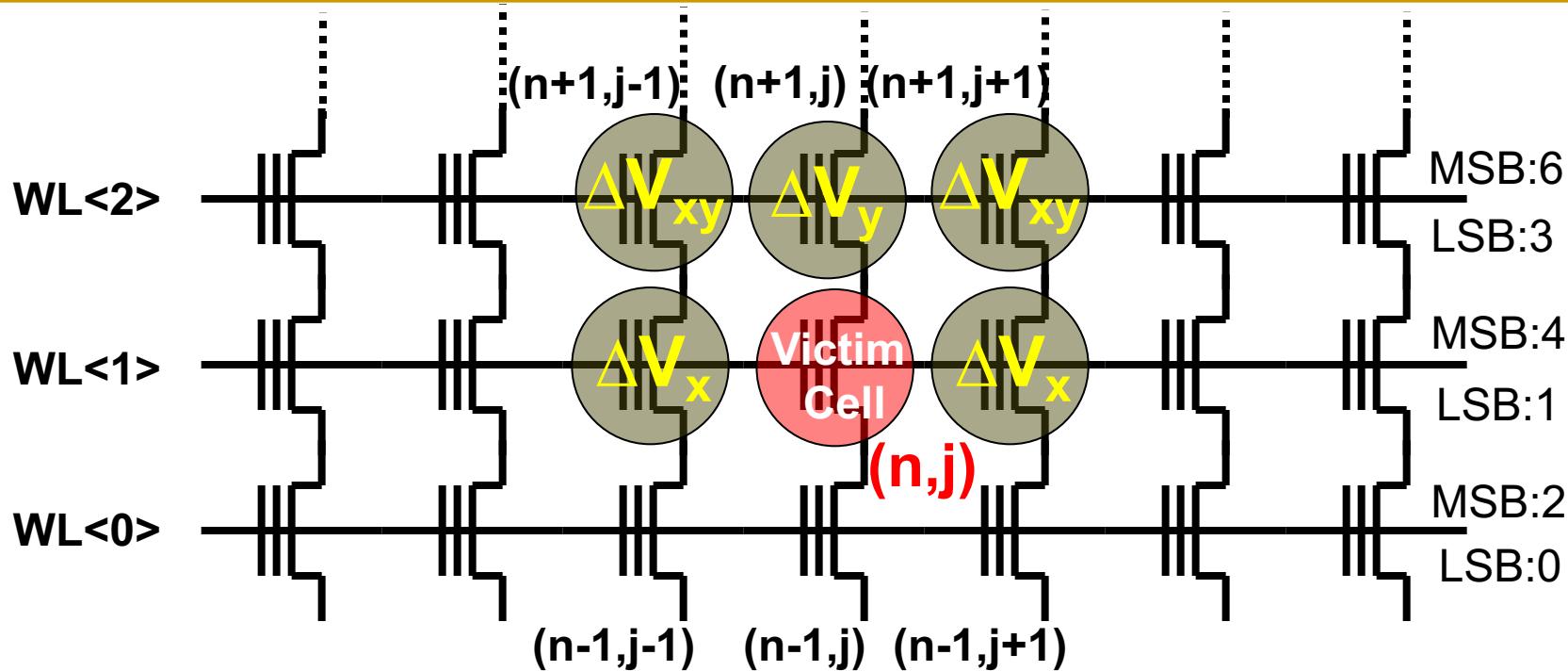
- Programming 2-bit MLC NAND flash memory in two steps



Basics of Program Interference



Traditional Model for V_{th} Change



- Traditional model for victim cell threshold voltage change

$$\Delta V_{victim} \square (2C_x \Delta V_x \square C_y \Delta V_y \square 2C_{xy} \Delta V_{xy}) / C_{total}$$

Not accurate and requires knowledge of coupling caps!

Our Goal and Idea

- Develop a new, more accurate and easier to implement model for program interference
- Idea:
 - Empirically characterize and model the effect of neighbor cell Vth changes on the Vth of the victim cell
 - Fit neighbor Vth change to a linear regression model and find the coefficients of the model via empirical measurement

$$\Delta V_{victim}(n, j) \square \sum_{y \square j-K}^{j \square K} \sum_{x \square n \square 1}^{n \square M} \alpha(x, y) \Delta V_{neighbor}(x, y) \square \alpha V_{victim}^{before}(n, j)$$

Can be measured

Developing a New Model via Empirical Measurement

- Feature extraction for V_{th} changes based on characterization
 - Threshold voltage changes on aggressor cell
 - Original state of victim cell
- Enhanced linear regression model

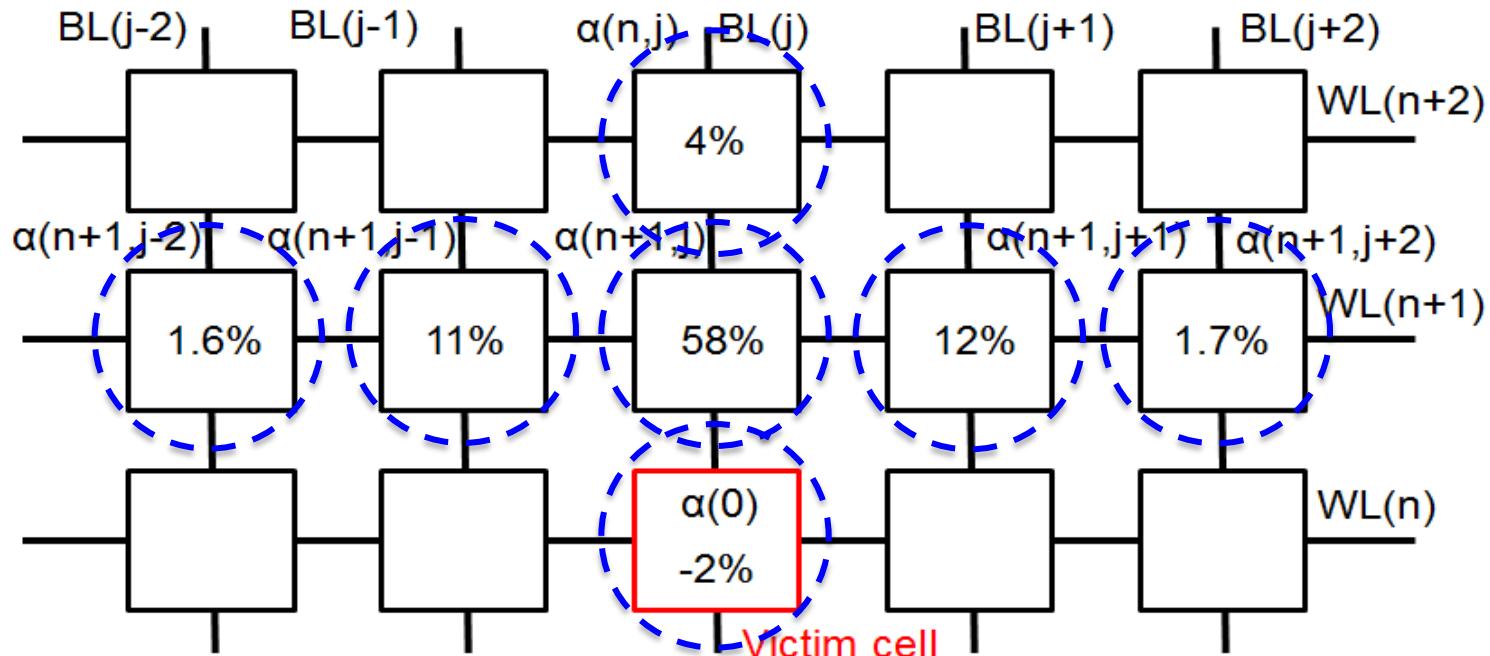
$$\Delta V_{victim}(n, j) \square \sum_{y \square j-K}^{j \square K} \sum_{x \square n \square 1}^{n \square M} \alpha(x, y) \Delta V_{neighbor}(x, y) \square \alpha_0 V_{victim}^{before}(n, j)$$

$$Y \square X\alpha \square \varepsilon \quad (\text{vector expression})$$

- Maximum likelihood estimation of the model coefficients

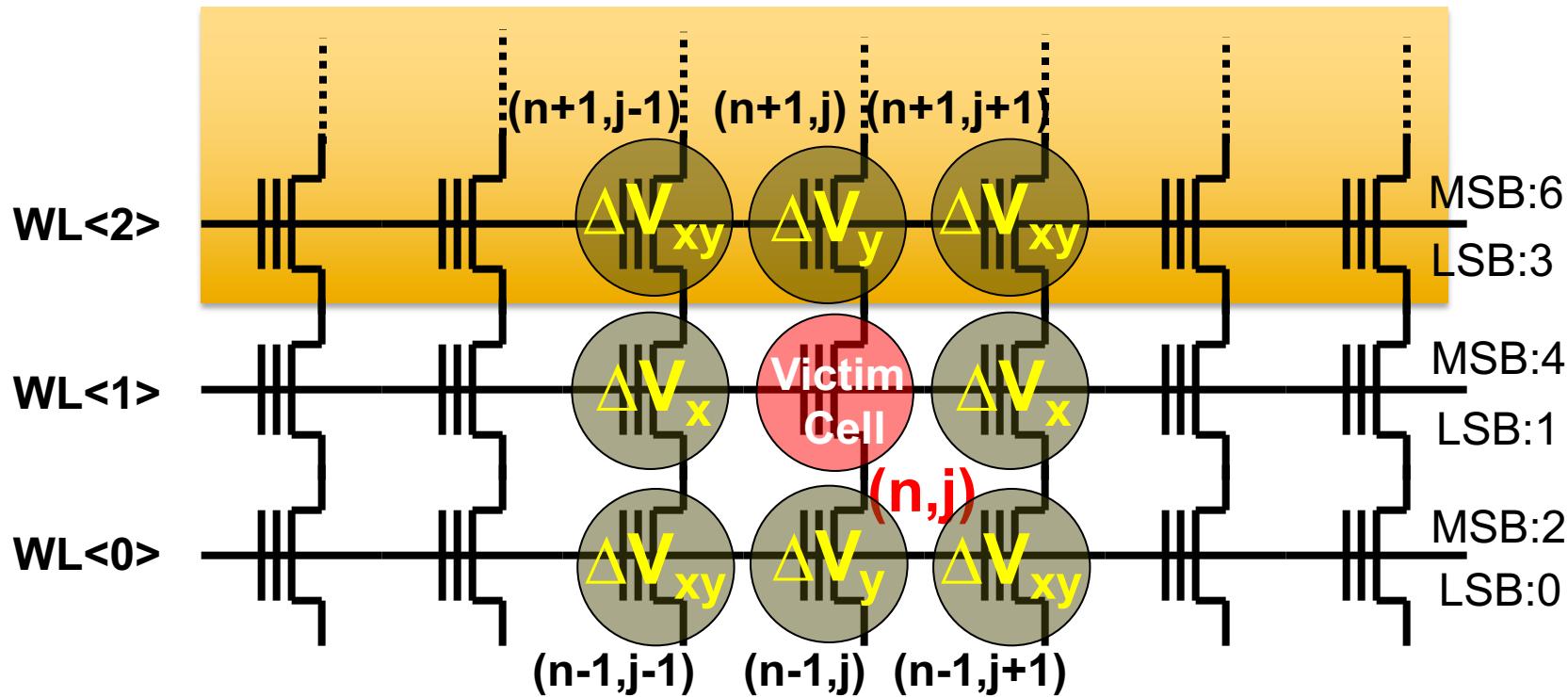
$$\arg \min_{\alpha} (\|X \times \alpha - Y\|_2^2 \square \lambda \|\alpha\|_1)$$

Effect of Neighbor Voltages on the Victim



- Immediately-above cell interference is dominant
- Immediately-diagonal neighbor is the second dominant
- Far neighbor cell interference exists
- Victim cell's V_{th} has negative effect on interference

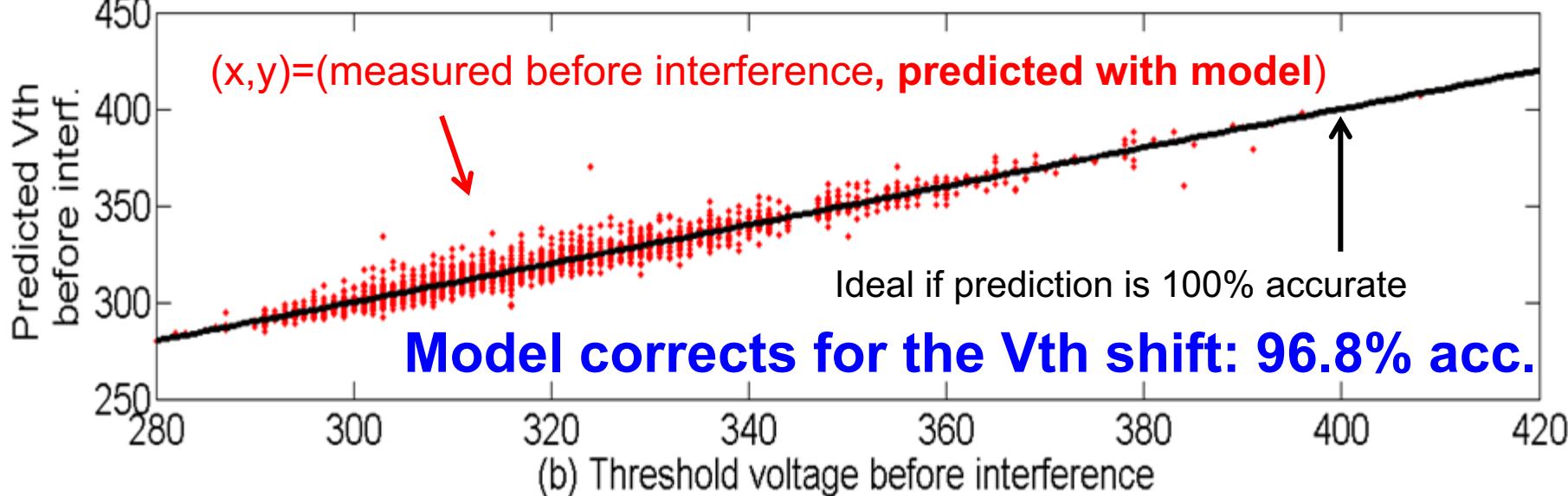
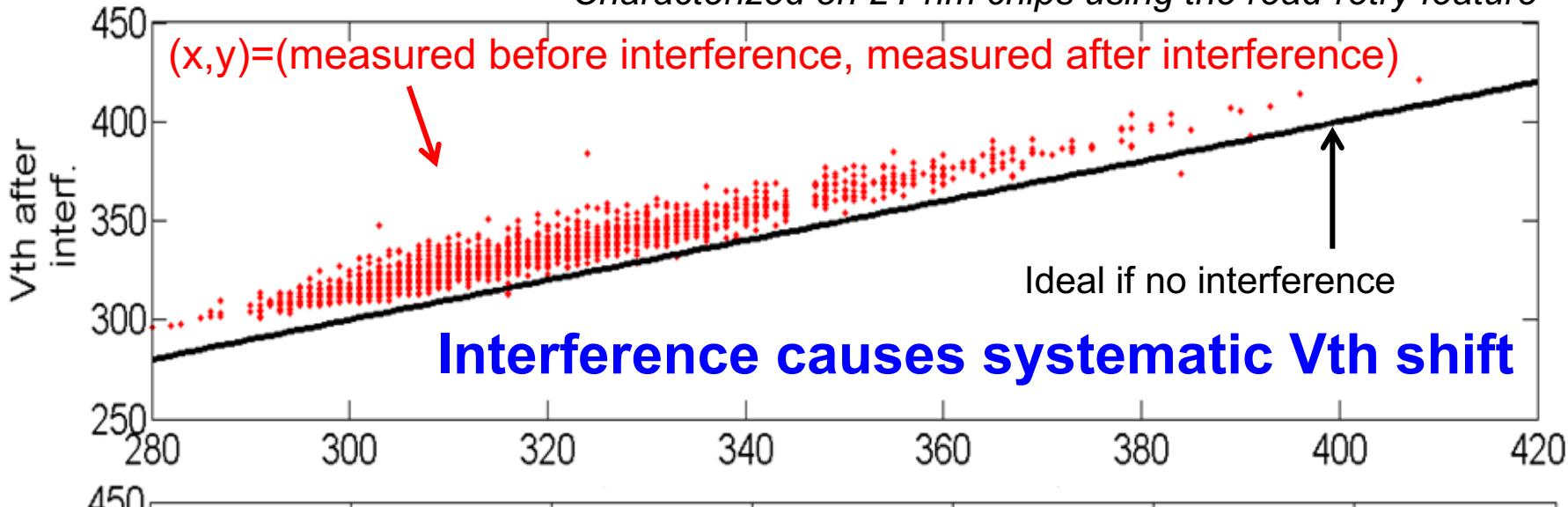
New Model for Program Interference



$$\Delta V_{victim}(n, j) \square \sum_{y \square j-K}^{j \square K} \sum_{x \square n \square 1}^{n \square M} \alpha(x, y) \Delta V_{neighbor}(x, y) \square \alpha_0 V_{victim}^{before}(n, j)$$

Model Accuracy

Characterized on 2Y-nm chips using the read-retry feature



Many Other Results in the Paper

- Yu Cai, Onur Mutlu, Erich F. Haratsch, and Ken Mai,
**"Program Interference in MLC NAND Flash Memory:
Characterization, Modeling, and Mitigation"**

*Proceedings of the 31st IEEE International Conference on
Computer Design (ICCD)*, Asheville, NC, October 2013. [Slides](#)
(pptx) (pdf) [Lightning Session Slides](#) (pdf)

Program Interference in MLC NAND Flash Memory: Characterization, Modeling, and Mitigation

Yu Cai¹, Onur Mutlu¹, Erich F. Haratsch² and Ken Mai¹

1. Data Storage Systems Center, Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA

2. LSI Corporation, San Jose, CA

yucaicai@gmail.com, {omutlu, kenmai}@andrew.cmu.edu

Agenda

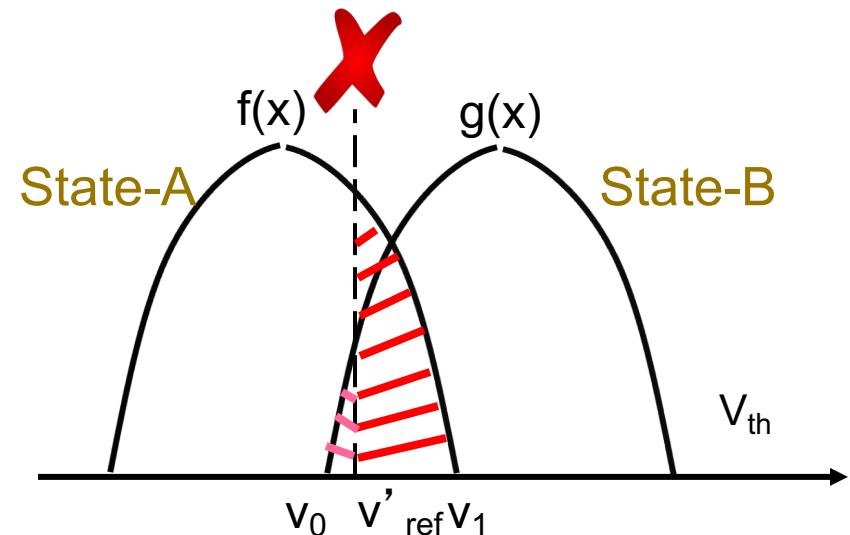
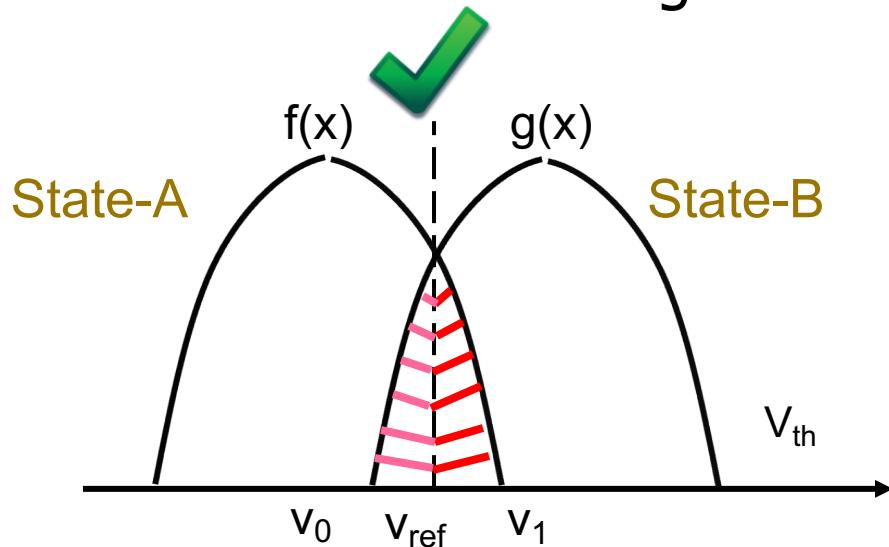
- Background, Motivation and Approach
 - Experimental Characterization Methodology
 - Error Analysis and Management
 - Main Characterization Results
 - Retention-Aware Error Management
 - Threshold Voltage and Program Interference Analysis
 - Read Reference Voltage Prediction
 - Neighbor-Assisted Error Correction
 - Read Disturb Error Handling
 - Retention Error Handling
 - Large Scale Field Analysis
 - 3D NAND Flash Memory Reliability
 - Summary
-

Mitigation: Applying the Model

- So, what can we do with the model?
- Goal: Mitigate the effects of program interference caused voltage shifts

Optimum Read Reference for Flash Memory

- Read reference voltage affects the raw bit error rate

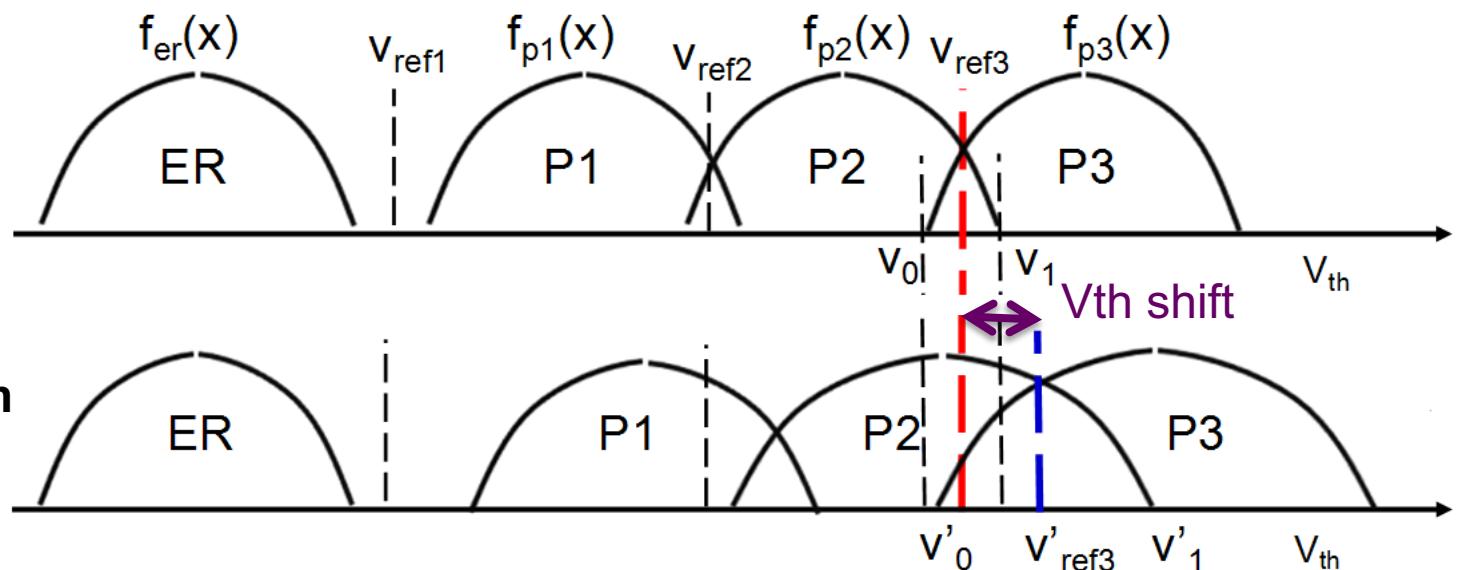


$$BER1 \square \int_{v_{ref}}^{\infty} f(x)dx \square \int_{-\infty}^{v_{ref}} g(x)dx$$

$$BER2 \square \int_{v'^{ref}}^{\infty} f(x)dx \square \int_{-\infty}^{v'^{ref}} g(x)dx$$

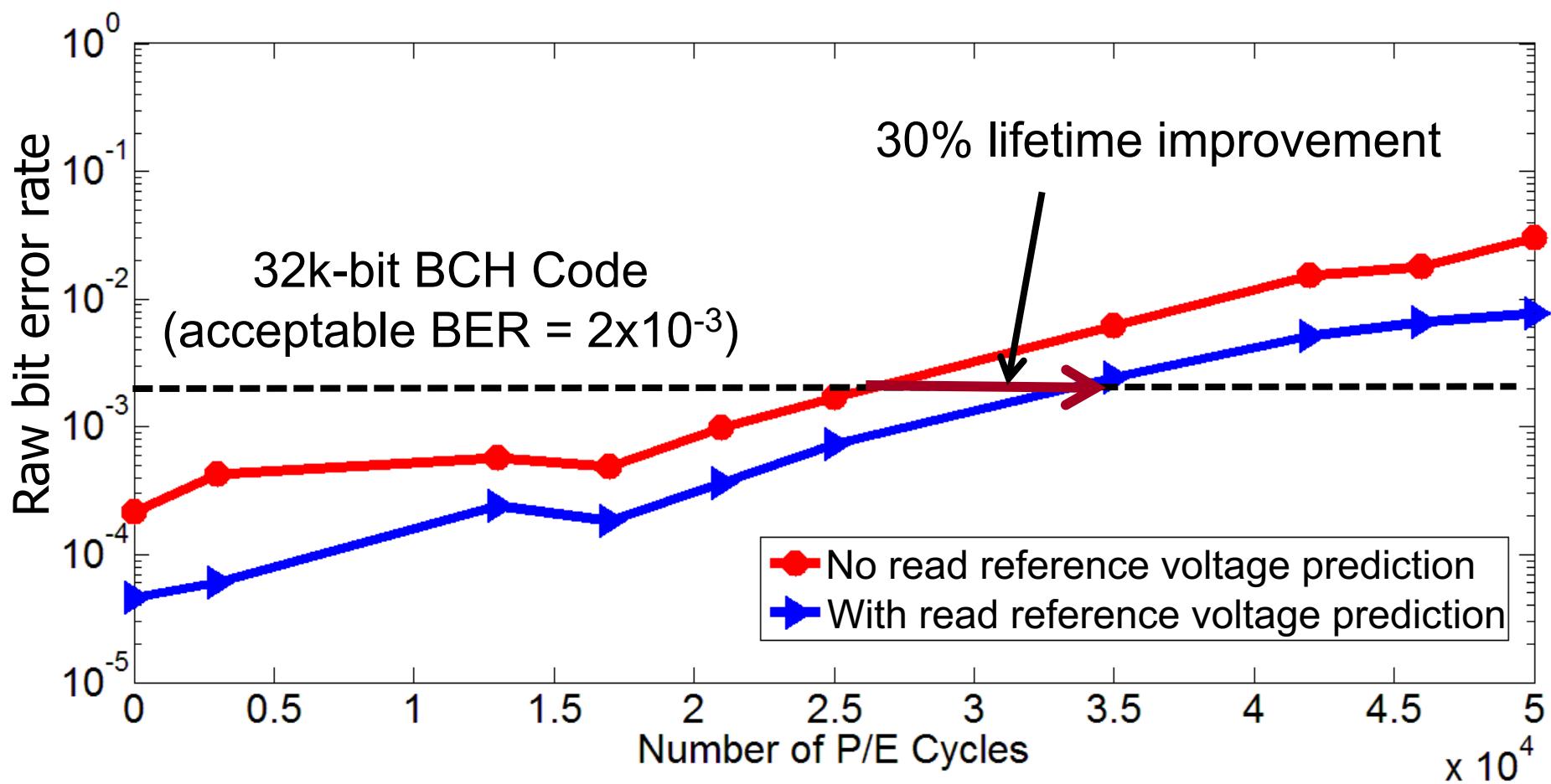
- There exists an optimal read reference voltage
 - Predictable if the statistics (i.e. mean, variance) of threshold voltage distributions are characterized and modeled

Optimum Read Reference Voltage Prediction



- Vth shift learning (done every $\sim 1k$ P/E cycles)
 - Program sample cells with known data pattern and test Vth
 - Program aggressor neighbor cells and test victim Vth after interference
 - Characterize the mean shift in Vth (i.e., program interference noise)
- Optimum read reference voltage prediction
 - Default read reference voltage + Predicted mean Vth shift by model

Effect of Read Reference Voltage Prediction



- Read reference voltage prediction reduces raw BER (by 64%) and increases the P/E cycle lifetime (by 30%)

More on Read Reference Voltage Prediction

- Yu Cai, Onur Mutlu, Erich F. Haratsch, and Ken Mai,
**"Program Interference in MLC NAND Flash Memory:
Characterization, Modeling, and Mitigation"**
*Proceedings of the 31st IEEE International Conference on
Computer Design (ICCD)*, Asheville, NC, October 2013.
Slides (pptx) (pdf) Lightning Session Slides (pdf)

Program Interference in MLC NAND Flash Memory: Characterization, Modeling, and Mitigation

Yu Cai¹, Onur Mutlu¹, Erich F. Haratsch² and Ken Mai¹

1. Data Storage Systems Center, Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA

2. LSI Corporation, San Jose, CA

yucaicai@gmail.com, {omutlu, kenmai}@andrew.cmu.edu

More Accurate and Online Channel Modeling

- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu,
**"Enabling Accurate and Practical Online Flash Channel Modeling
for Modern MLC NAND Flash Memory"**
to appear in IEEE Journal on Selected Areas in Communications (JSAC),
2016.

Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory

Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, Onur Mutlu

Non-Gaussian V_{th} Distributions (1X-nm)

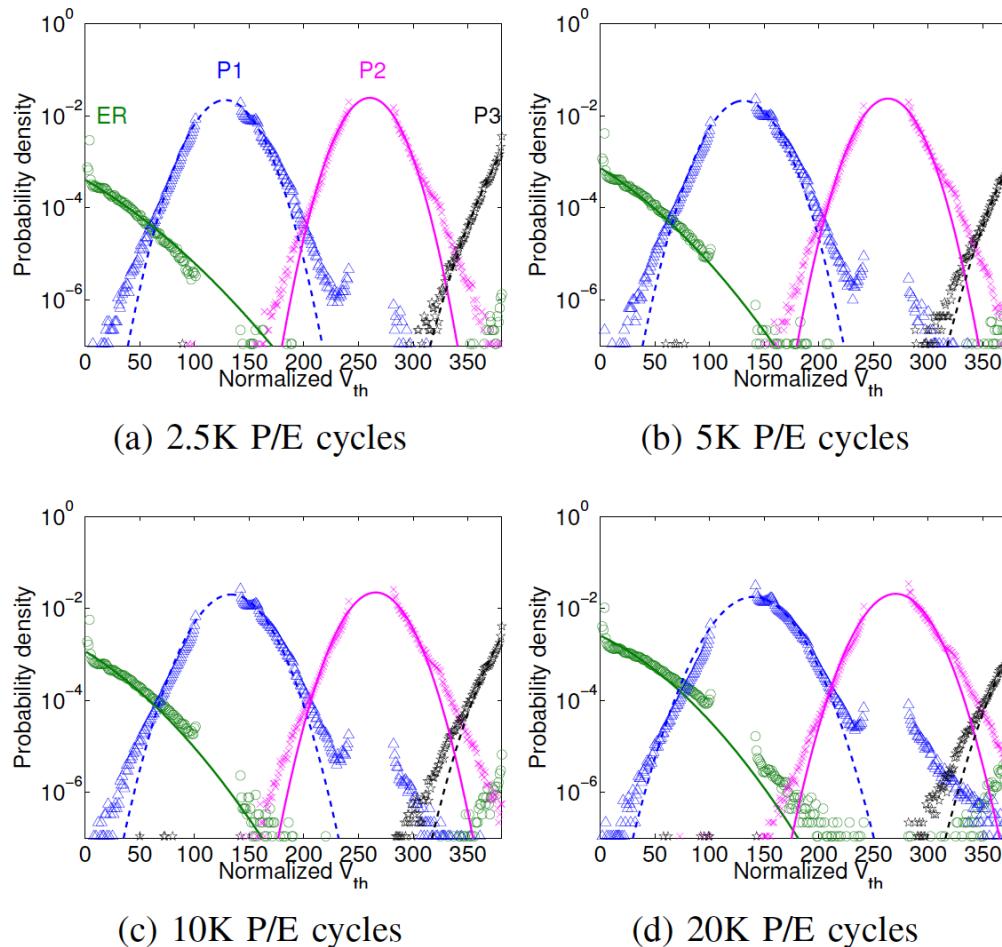


Fig. 4: Gaussian-based model (solid/dashed lines) vs. data measured from real NAND flash chips (markers) under different P/E cycle counts.

Better Modeling of V_{th} Distributions (I)

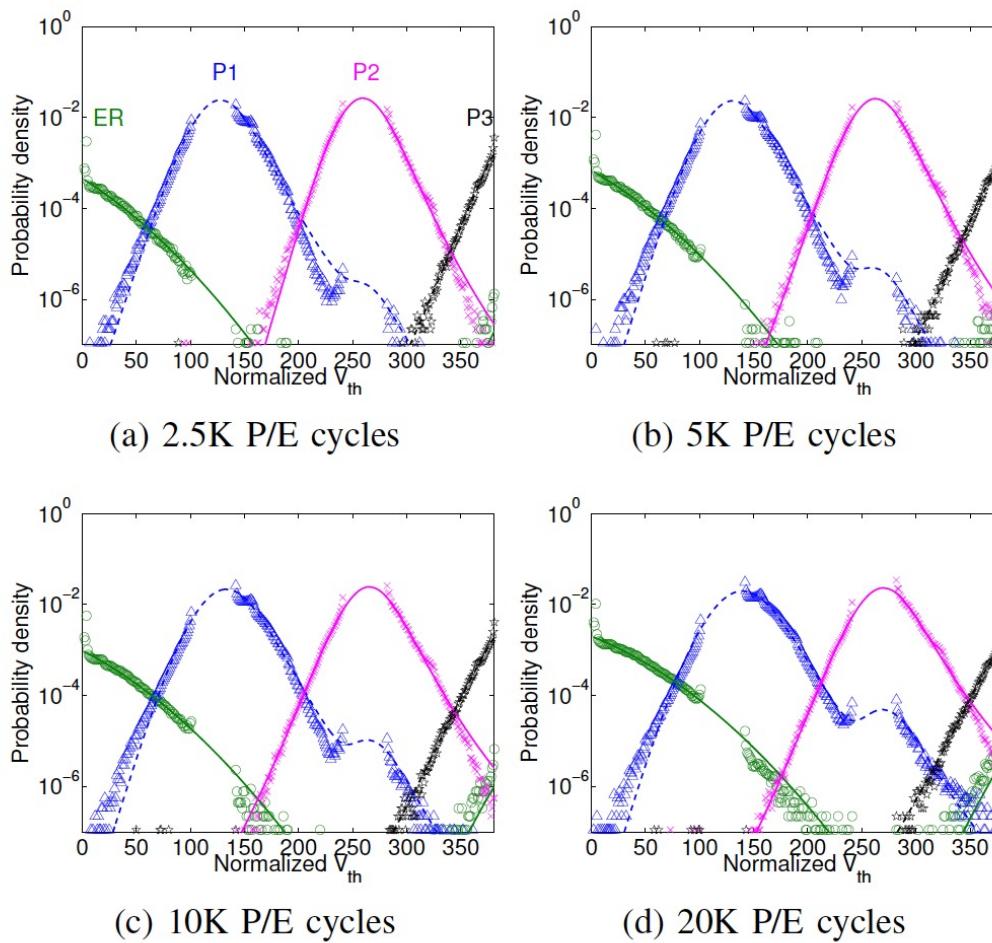


Fig. 6: Our new Student's t-based model (solid/dashed lines) vs. data measured from real NAND flash chips (markers) under different P/E cycle counts.

Better Modeling of V_{th} Distributions (II)

P/E Cycles	0	2.5K	5K	7.5K	10K	12K	14K	16K	18K	20K	AVG
Gaussian	.99%	1.8%	1.6%	1.8%	1.9%	2.4%	3.1%	8.7%	2.1%	2.3%	2.6%
Normal-Laplace	.34%	.46%	.55%	.61%	.63%	.67%	.68%	.70%	.67%	.67%	.61%
Student's t	.37%	.51%	.61%	.68%	.70%	.76%	.76%	.78%	.76%	.78%	.68%

TABLE 1: Modeling error of the evaluated threshold voltage distribution models, at various P/E cycle counts.

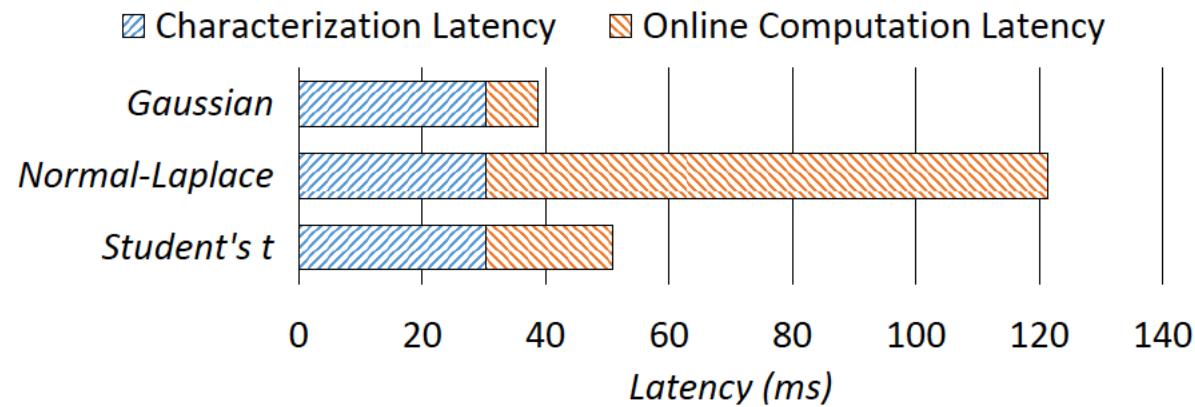


Fig. 8: Overall latency breakdown of the three evaluated threshold voltage distribution models for static modeling.

V_{th} Prediction vs. Reality with Better Modeling

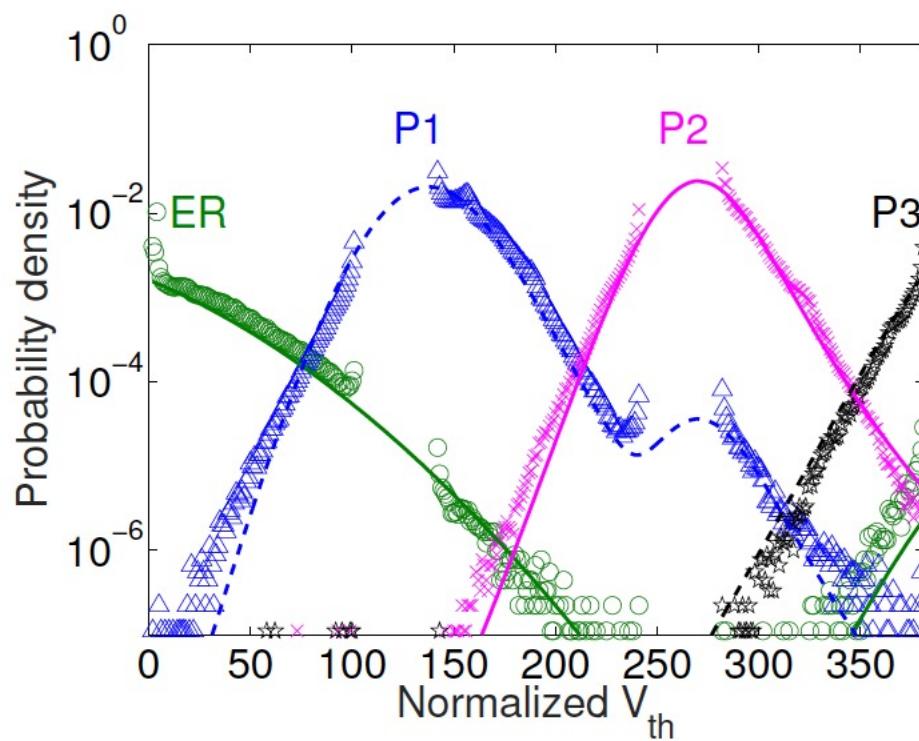


Fig. 13: Threshold voltage distribution as predicted by our dynamic model for 20K P/E cycles, using characterization data from 2.5K, 5K, 7.5K, and 10K P/E cycles, shown as solid/dashed lines. Markers represent data measured from real NAND flash chips at 20K P/E cycles.

Online Read Reference Voltage Prediction

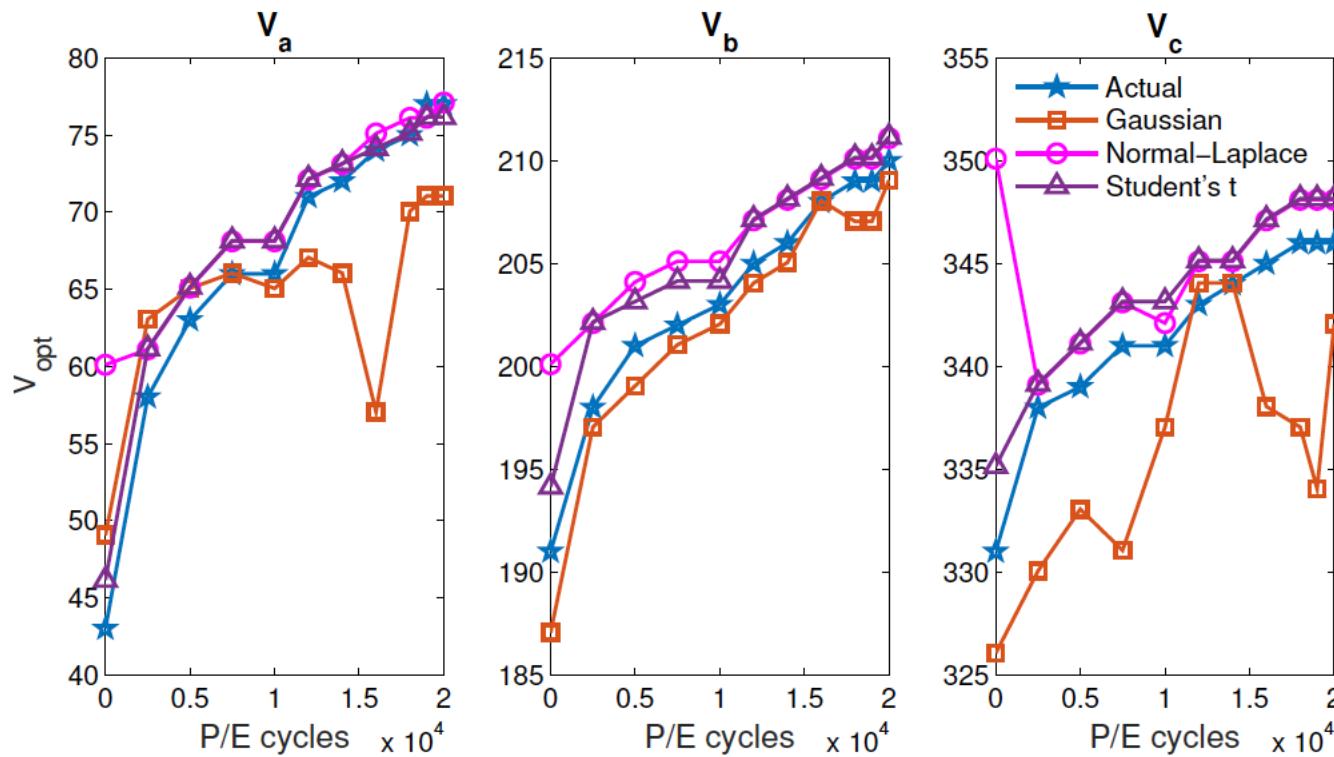


Fig. 16: Actual and modeled *optimal* read reference voltages (V_{opt}) using the three evaluated threshold voltage distribution models at different P/E cycle counts.

Effect on RBER of Read Ref V Prediction

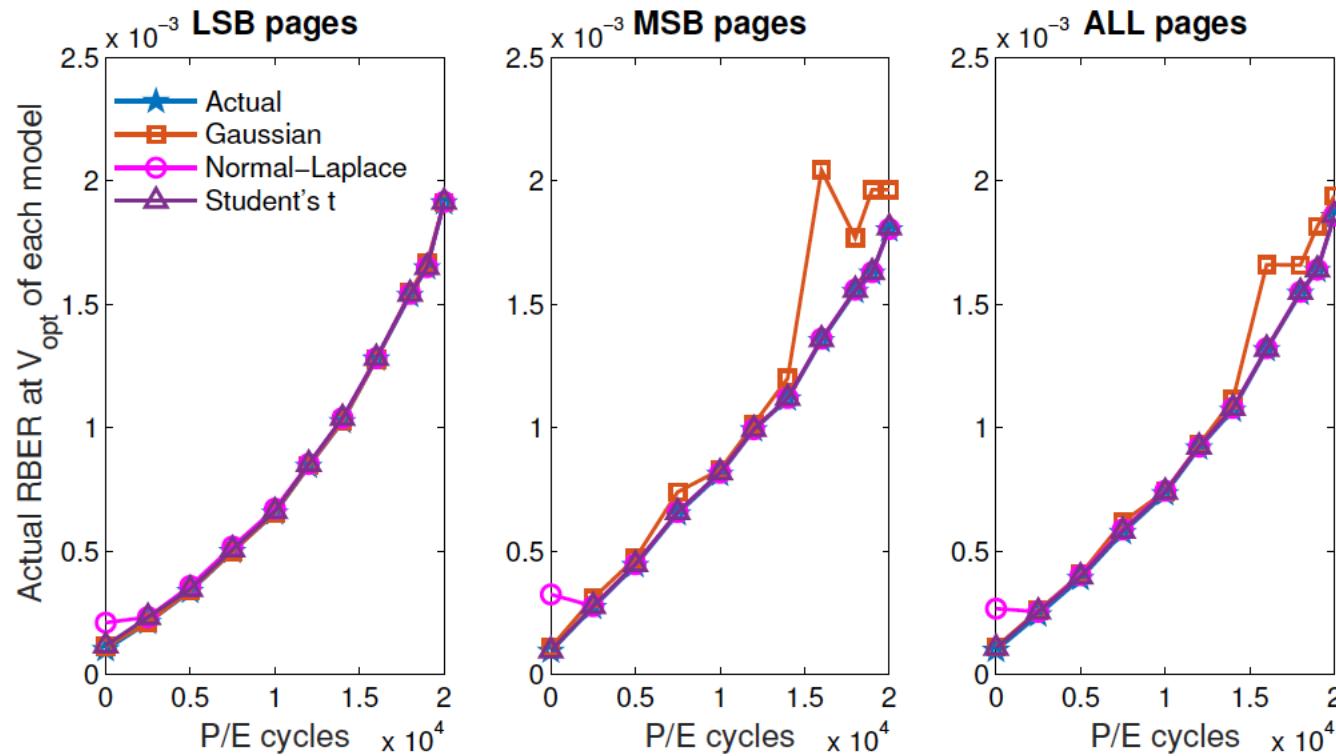


Fig. 17: RBER achieved by actual and modeled *optimal* read reference voltages (V_{opt}) using the three evaluated threshold voltage distribution models at different P/E cycle counts.

More Accurate and Online Channel Modeling

- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu,
**"Enabling Accurate and Practical Online Flash Channel Modeling
for Modern MLC NAND Flash Memory"**
to appear in IEEE Journal on Selected Areas in Communications (JSAC),
2016.

Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory

Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, Onur Mutlu

Agenda

- Background, Motivation and Approach
 - Experimental Characterization Methodology
 - Error Analysis and Management
 - Main Characterization Results
 - Retention-Aware Error Management
 - Threshold Voltage and Program Interference Analysis
 - Read Reference Voltage Prediction
 - Neighbor-Assisted Error Correction
 - Read Disturb Error Handling
 - Retention Error Handling
 - Large Scale Field Analysis
 - 3D NAND Flash Memory Reliability
 - Summary
-

Goal

- Develop a better error correction mechanism for cases where ECC fails to correct a page

Observations So Far

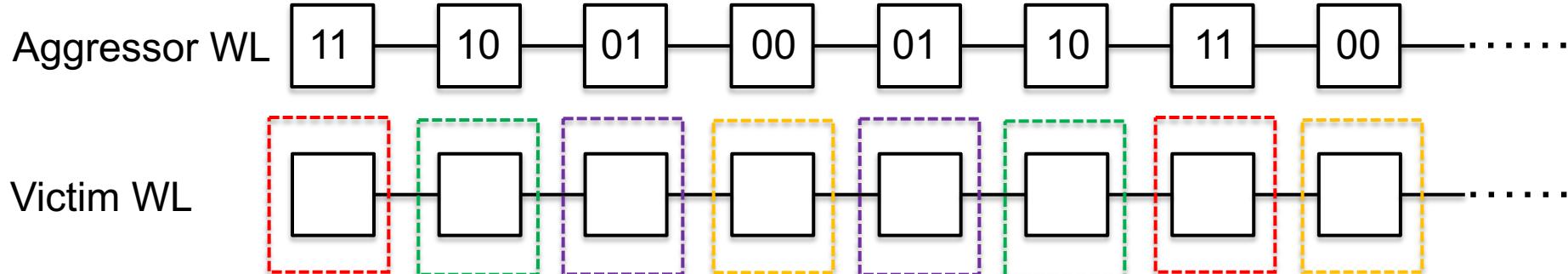
- Immediate neighbor cell has the most effect on the victim cell when programmed
- A single set of read reference voltages is used to determine the value of the (victim) cell
- The set of read reference voltages is determined based on the ***overall threshold voltage distribution of all cells*** in flash memory

New Observations [Cai+ SIGMETRICS'14]

- Vth distributions of **cells with different-valued immediate-neighbor cells** are significantly different
 - Because neighbor value affects the amount of Vth shift
- **Corollary:** If we know the value of the immediate-neighbor, we can find a more accurate set of read reference voltages based on the “conditional” threshold voltage distribution

Cai et al., Neighbor-Cell Assisted Error Correction for MLC NAND Flash Memories, SIGMETRICS 2014.

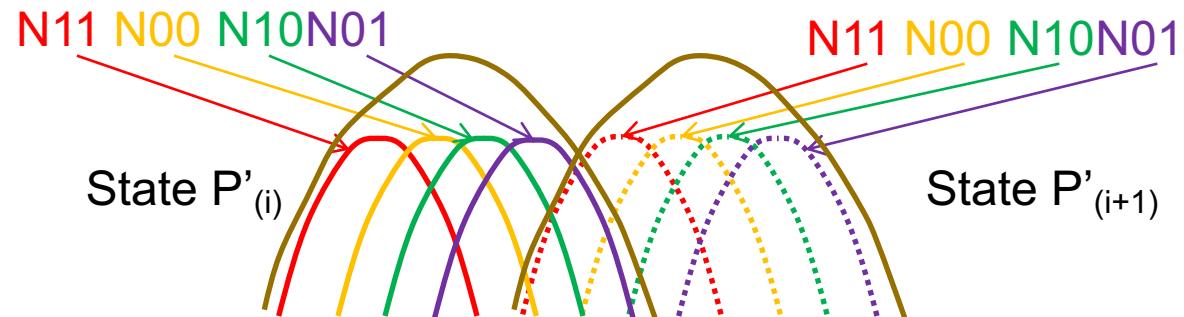
Secrets of Threshold Voltage Distributions



Victim WL **before** MSB
page of aggressor WL
are programmed



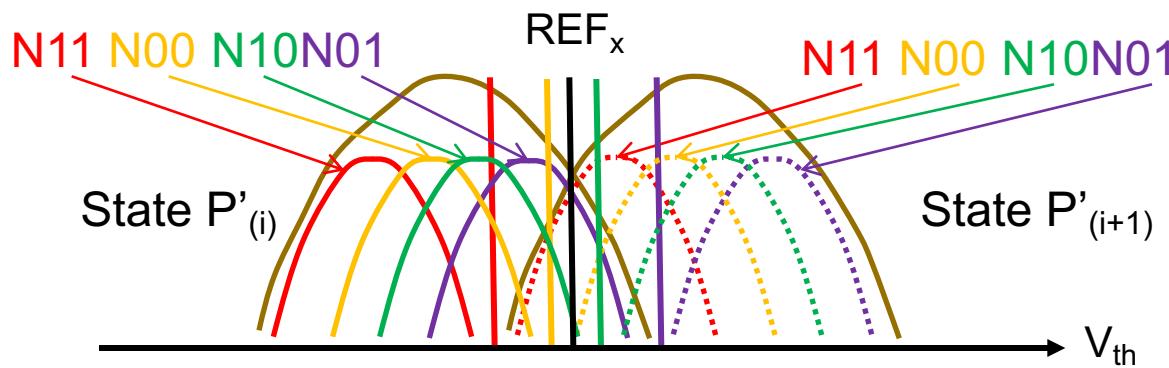
Victim WL **after** MSB
page of aggressor WL
are programmed



If We Knew the Immediate Neighbor ...

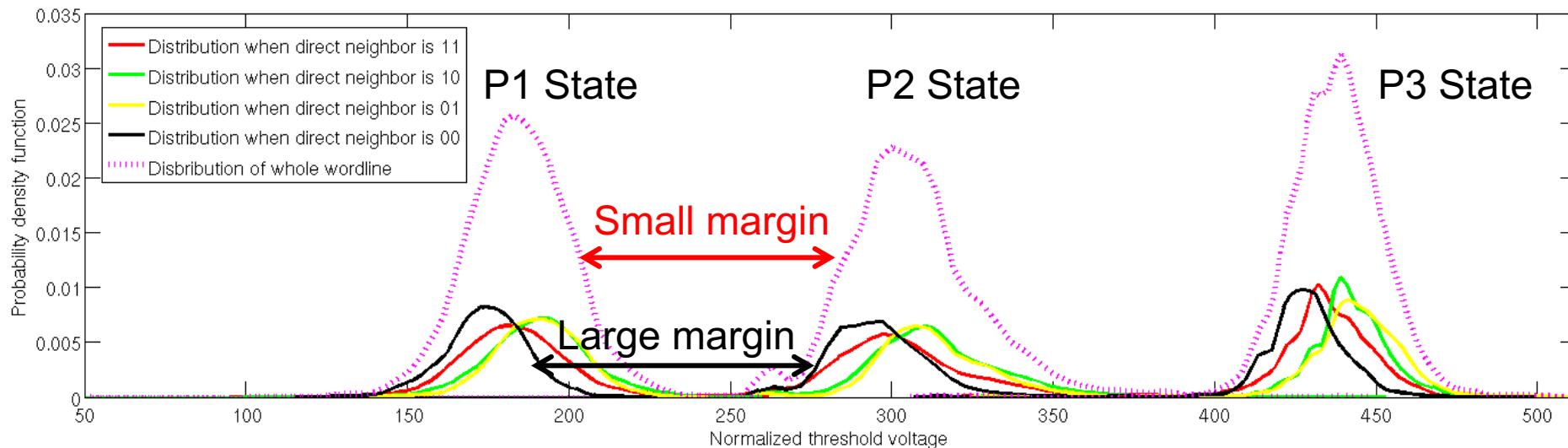
- Then, we could choose a different read reference voltage to more accurately read the “victim” cell

Overall vs Conditional Reading



- Using the optimum read reference voltage based on the overall distribution leads to more errors
- Better to use the optimum read reference voltage based on the conditional distribution (i.e., value of the neighbor)
 - Conditional distributions of two states are farther apart from each other

Real NAND Flash Chip Measurement Results



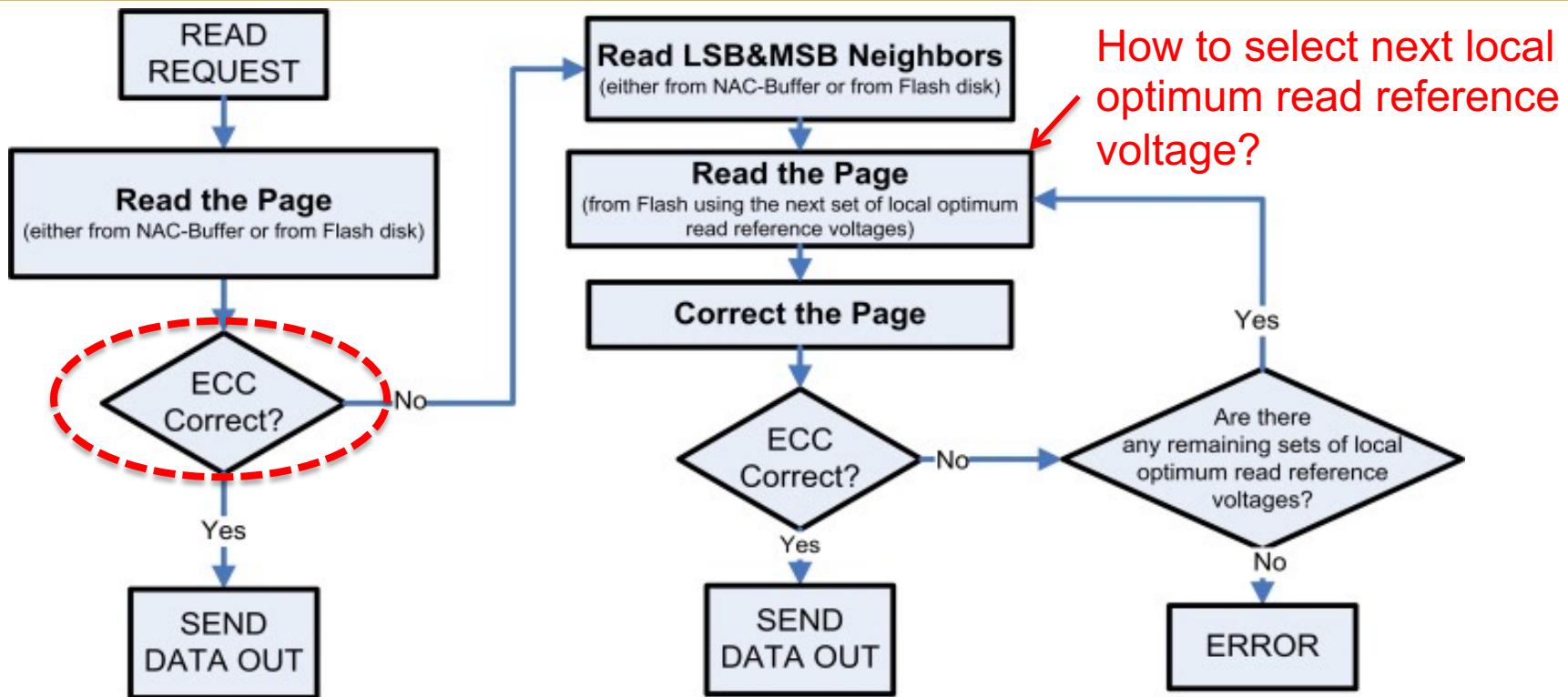
	Overall	x_{11} (ER)	x_{10} (P1)	x_{00} (P2)	X_{01} (P3)
Distance	65.4	65.4	64.7	66.4	65.8
Variance	385.9	286.2	256.7	242.8	252.1
SNR	3.4	3.8	3.9	4.2	4.1
BER	3×10^{-4}	7×10^{-5}	5×10^{-5}	2×10^{-5}	3×10^{-5}

Raw BER of conditional reading is much smaller than overall reading

Idea: Neighbor Assisted Correction (NAC)

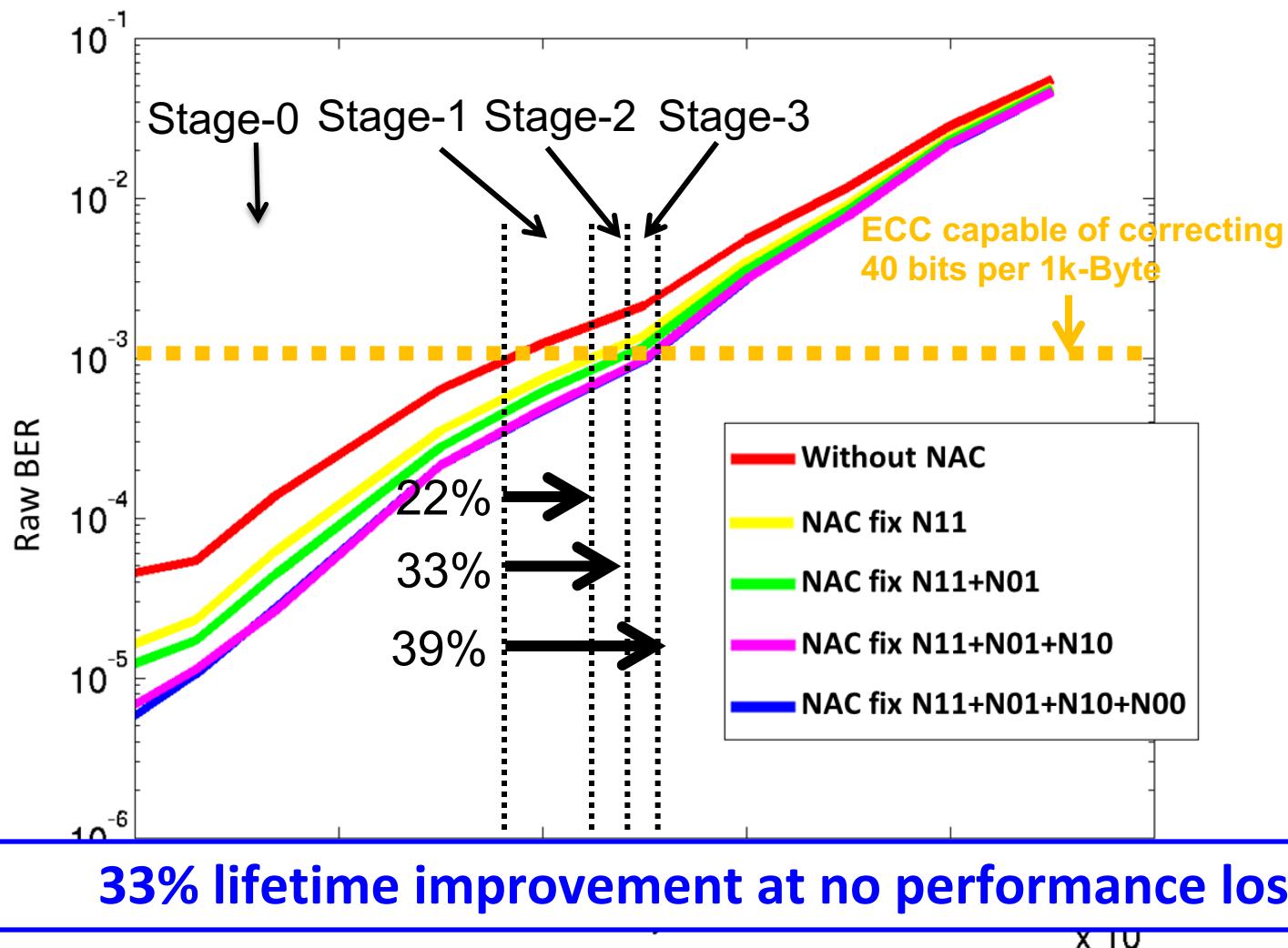
- Read a page with the read reference voltages based on overall V_{th} distribution (same as today) and buffer it
- If ECC fails:
 - Read the immediate-neighbor page
 - Re-read the page using the read reference voltages corresponding to the voltage distribution assuming a particular immediate-neighbor value
 - Replace the buffered values of the cells with that particular immediate-neighbor cell value
 - Apply ECC again

Neighbor Assisted Correction Flow

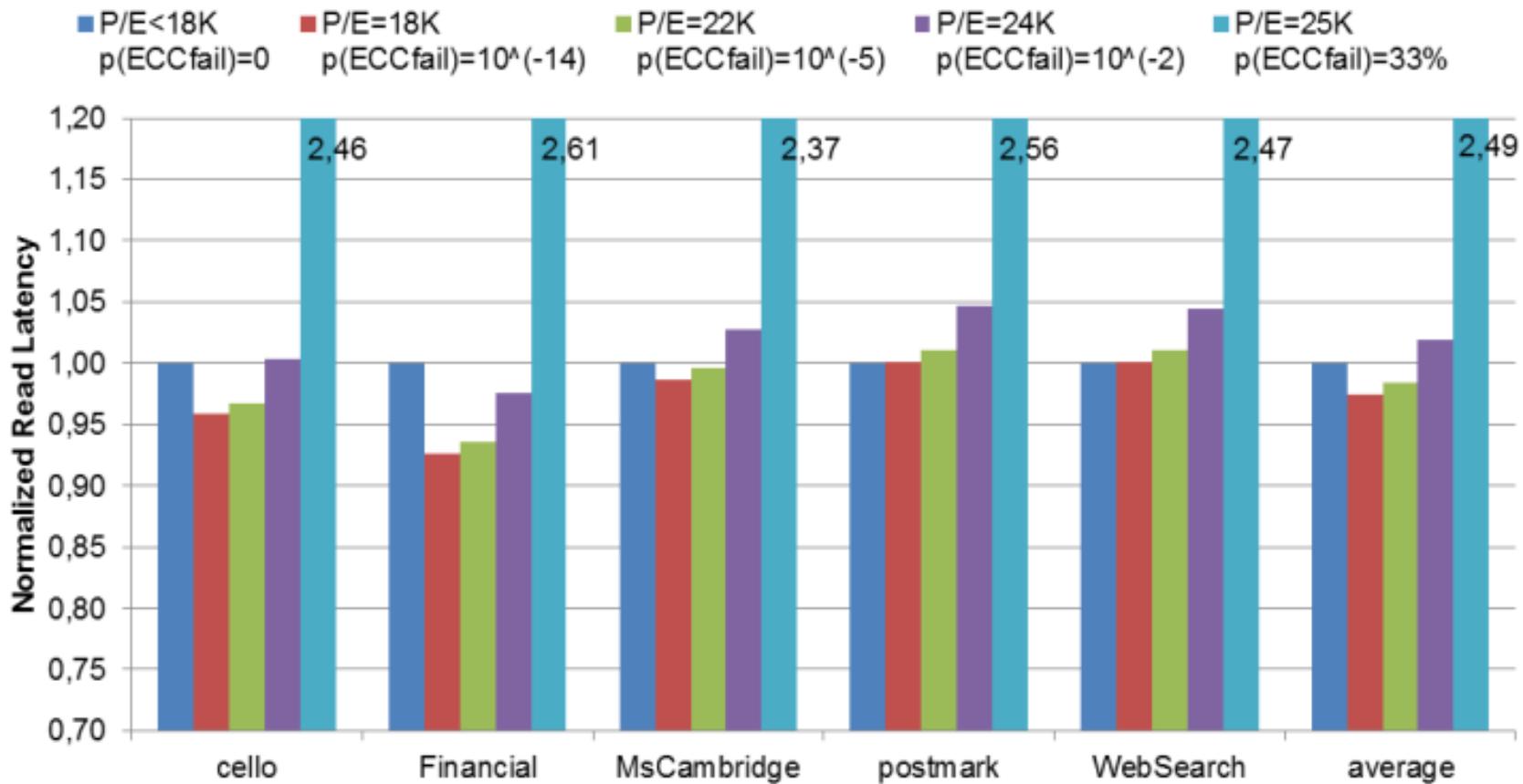


- Trigger neighbor-assisted reading only when ECC fails
- Read neighbor values and use corresponding read reference voltages in a prioritized order until ECC passes

Lifetime Extension with NAC



Performance Analysis of NAC



No performance loss within nominal lifetime
and with reasonable (1%) ECC fail rates

More on Neighbor-Assisted Correction

- Yu Cai, Gulay Yalcin, Onur Mutlu, Eric Haratsch, Osman Unsal, Adrian Cristal, and Ken Mai,
"Neighbor-Cell Assisted Error Correction for MLC NAND Flash Memories"

Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS), Austin, TX, June 2014. [Slides \(ppt\)](#) [\(pdf\)](#)

Neighbor-Cell Assisted Error Correction for MLC NAND Flash Memories

Yu Cai¹, Gulay Yalcin², Onur Mutlu¹, Erich F. Haratsch⁴,
Osman Unsal², Adrian Cristal^{2,3}, and Ken Mai¹

¹Electrical and Computer Engineering Department, Carnegie Mellon University

²Barcelona Supercomputing Center, Spain ³IIIA – CSIC – Spain National Research Council ⁴LSI Corporation
yucaicai@gmail.com, {omutlu, kenmai}@ece.cmu.edu, {gulay.yalcin, adrian.cristal, osman.unsal}@bsc.es

Agenda

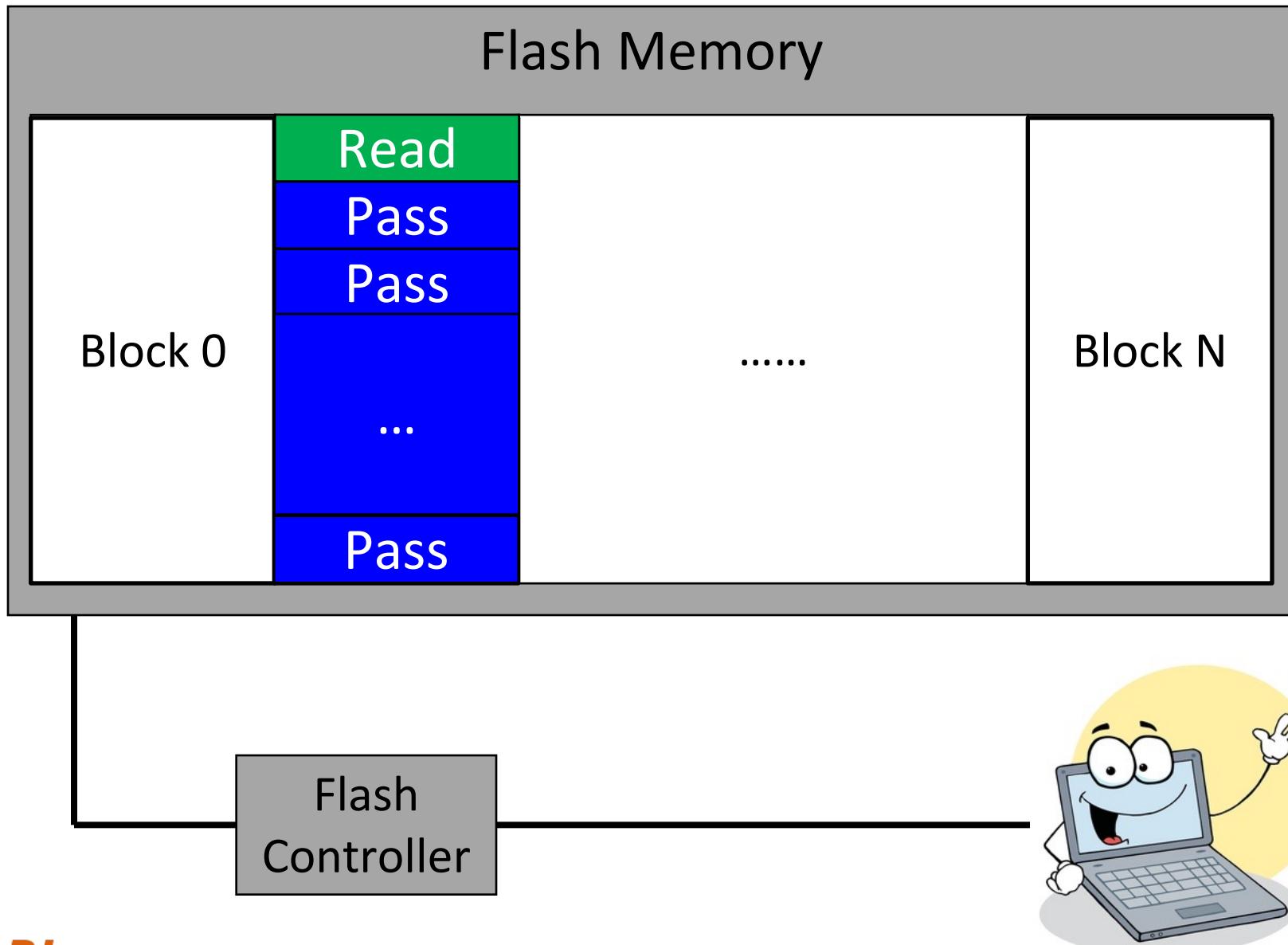
- Background, Motivation and Approach
 - Experimental Characterization Methodology
 - **Error Analysis and Management**
 - Main Characterization Results
 - Retention-Aware Error Management
 - Threshold Voltage and Program Interference Analysis
 - Read Reference Voltage Prediction
 - Neighbor-Assisted Error Correction
 - **Read Disturb Error Handling**
 - Retention Error Handling
 - Large Scale Field Analysis
 - 3D NAND Flash Memory Reliability
 - Summary
-

Read Disturb Errors in Flash Memory

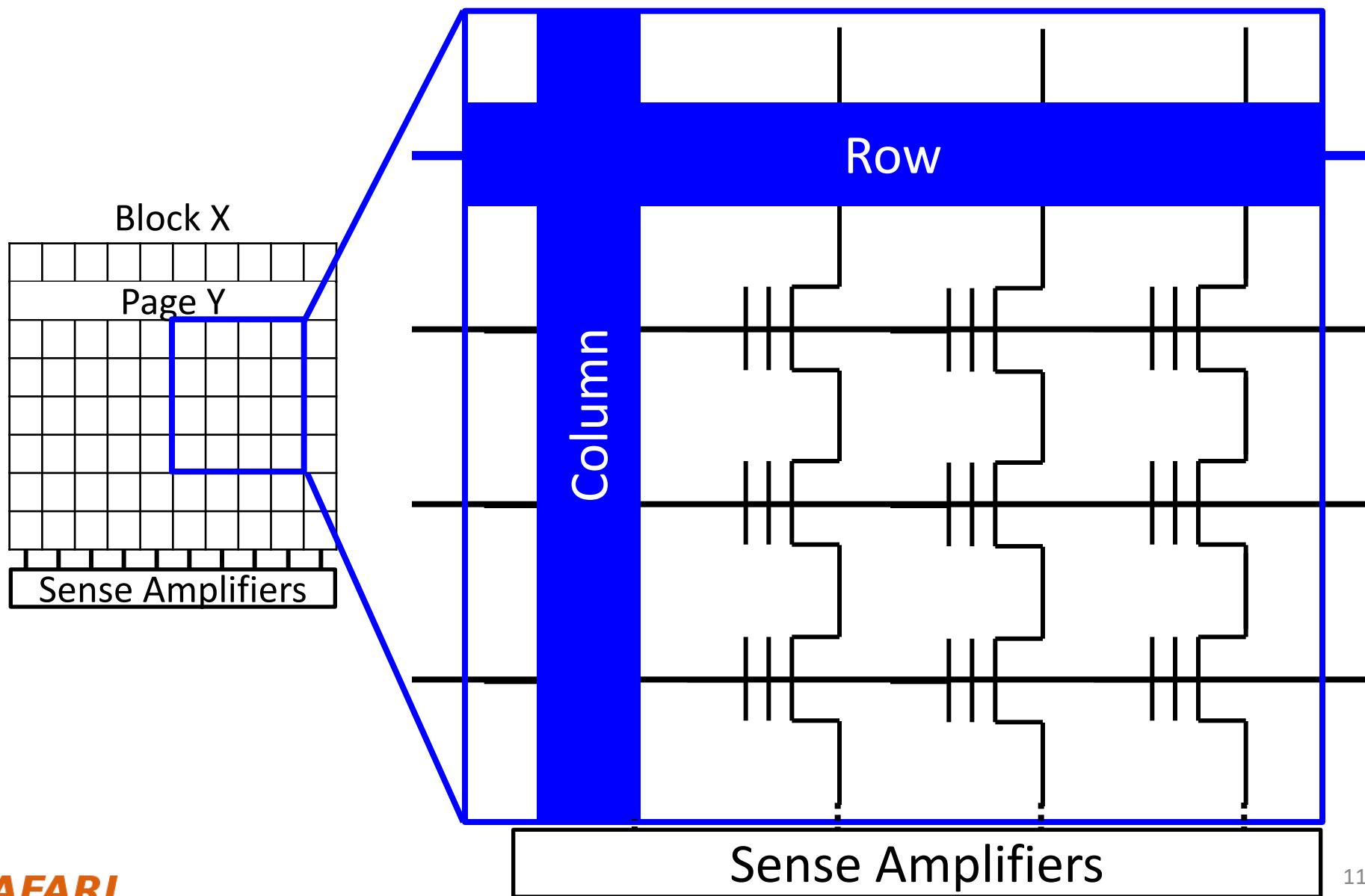
One Issue: Read Disturb in Flash Memory

- All scaled memories are prone to read disturb errors
- DRAM
- SRAM
- Hard Disks: Adjacent Track Interference
- NAND Flash

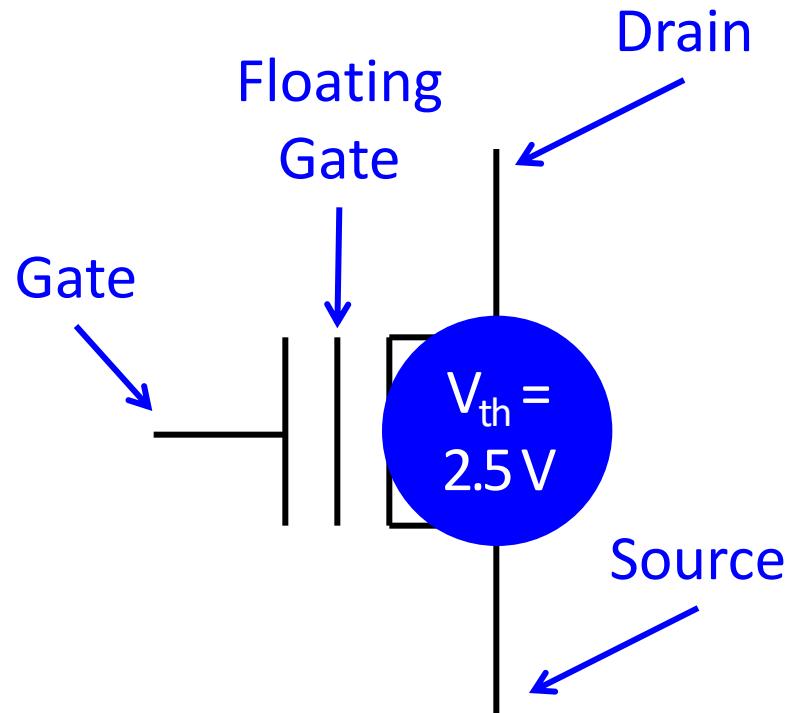
NAND Flash Memory Background



Flash Cell Array

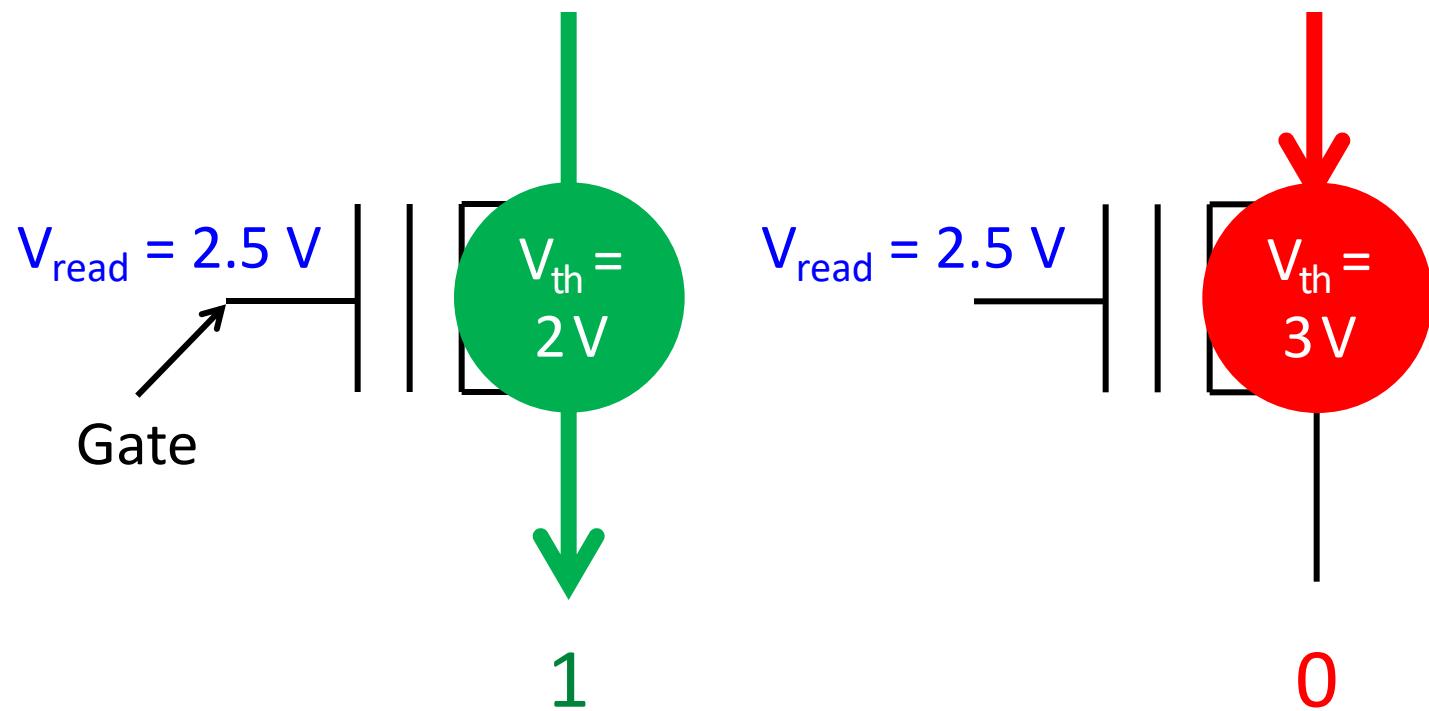


Flash Cell

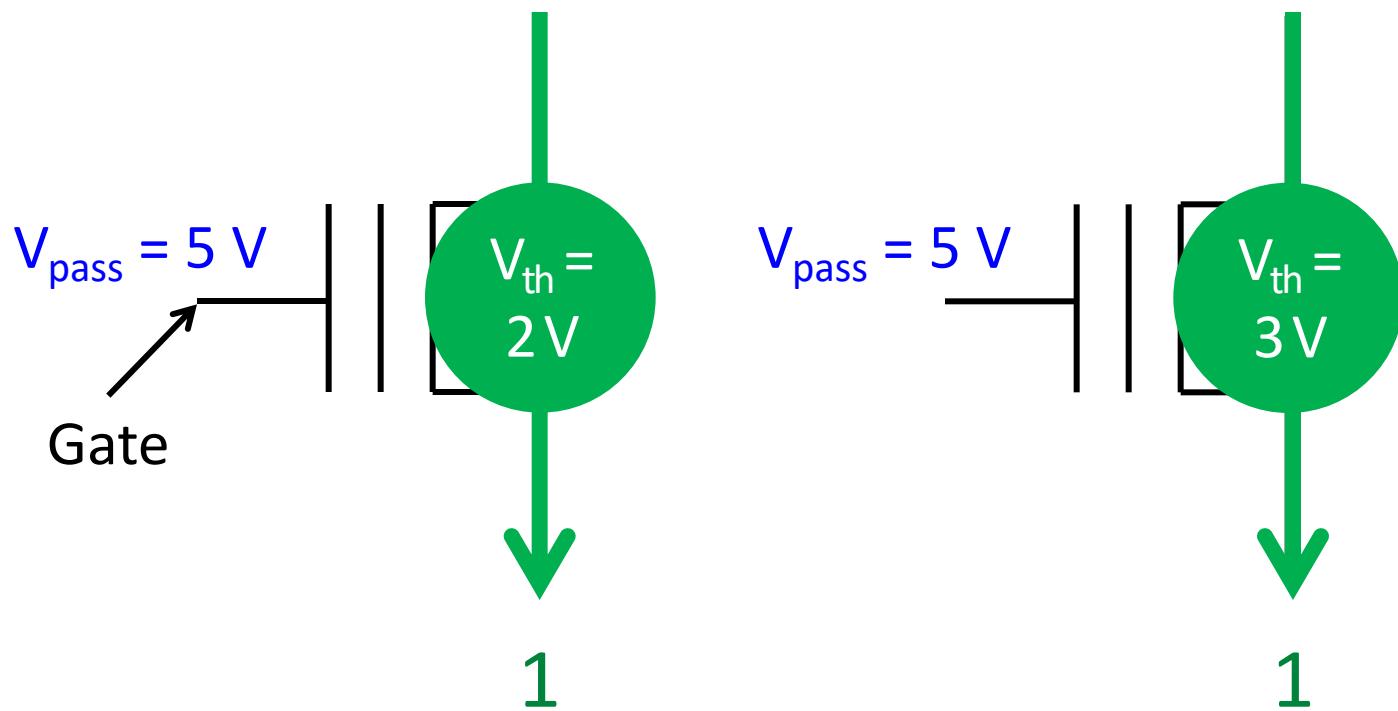


Floating Gate Transistor
(Flash Cell)

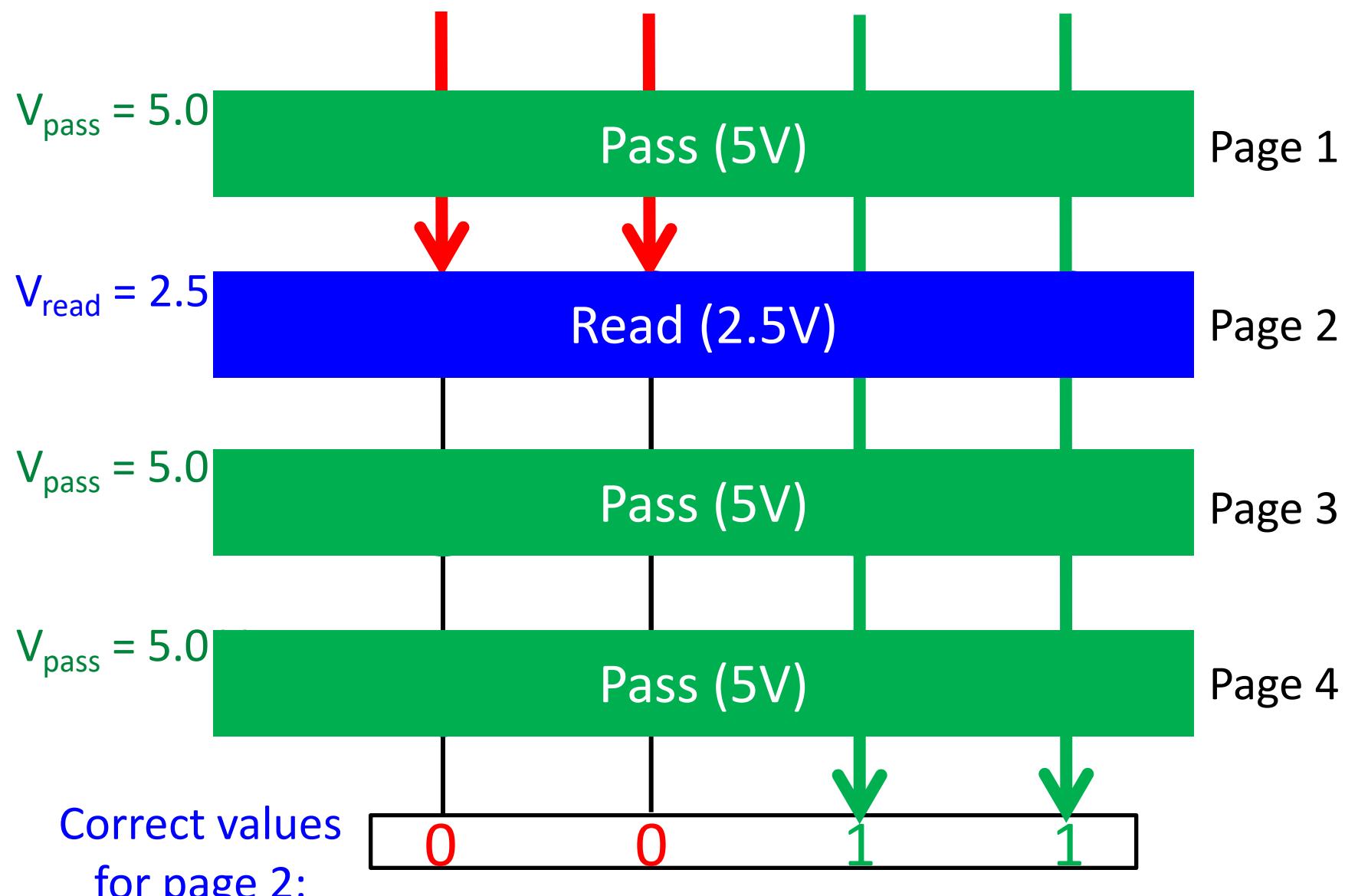
Flash Read



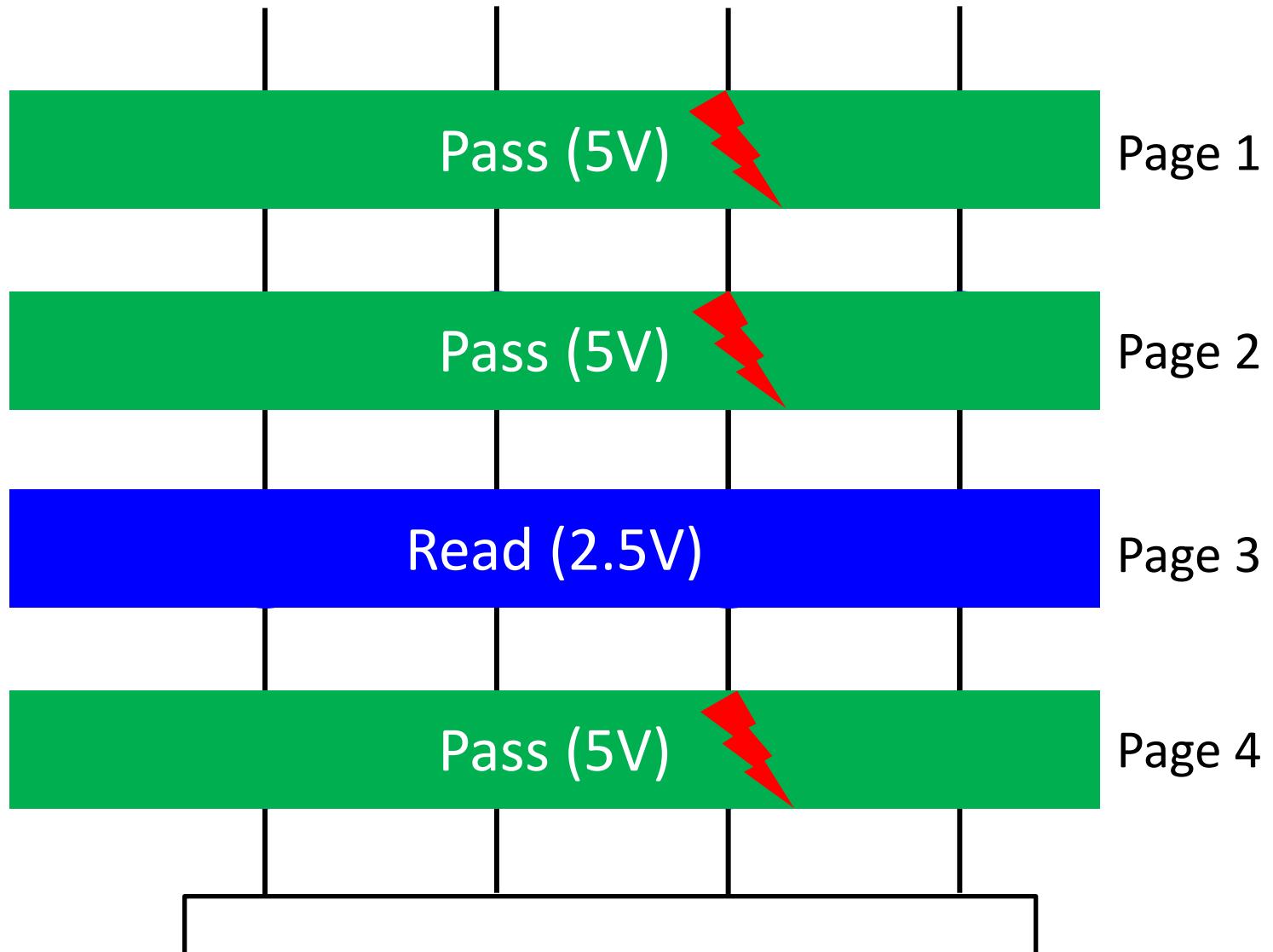
Flash Pass-Through



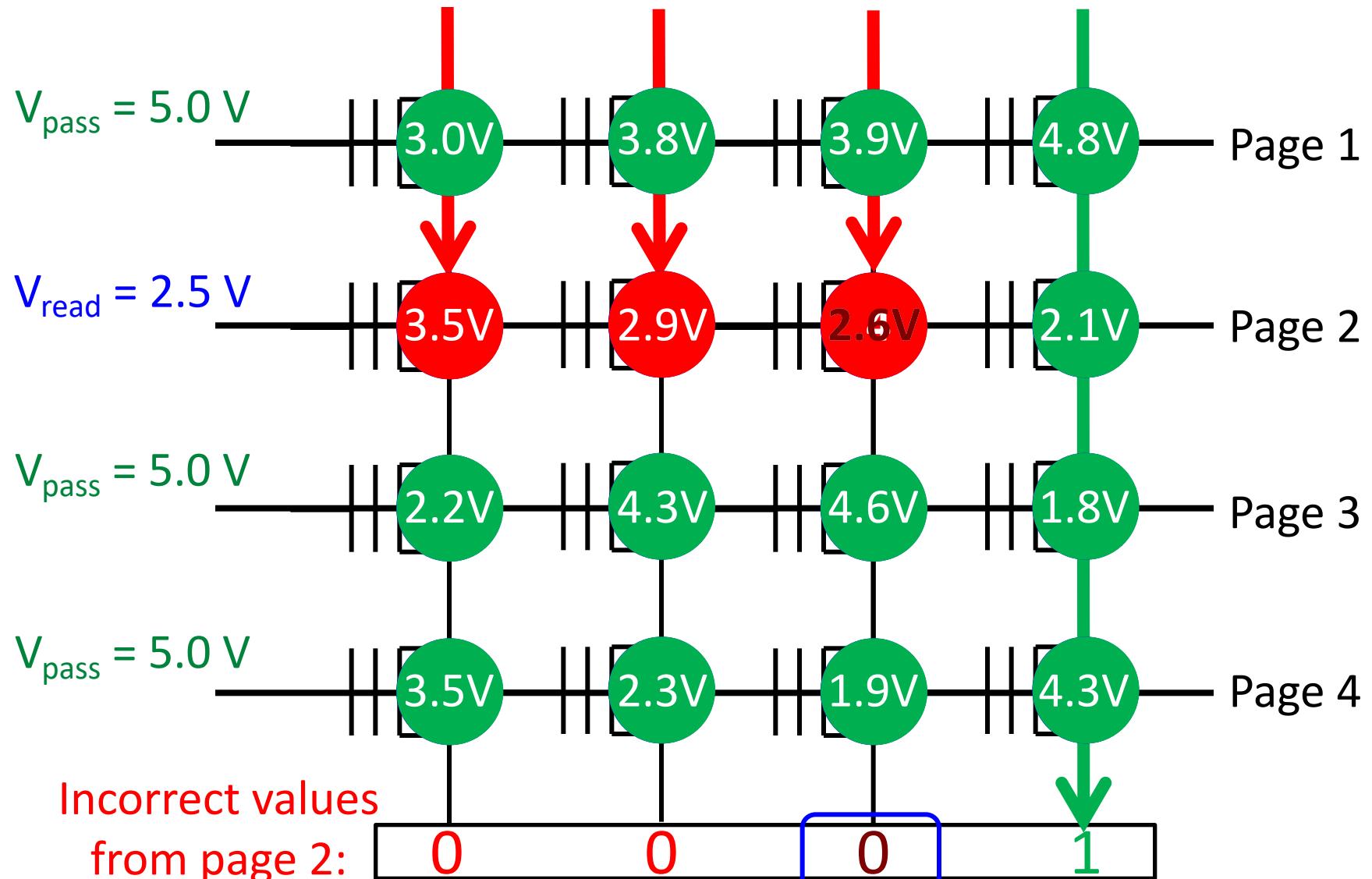
Read from Flash Cell Array



Read Disturb Problem: “Weak Programming” Effect



Read Disturb Problem: “Weak Programming” Effect



Executive Summary [DSN'15]

- ***Read disturb errors*** limit flash memory lifetime today
 - Apply a *high pass-through voltage (V_{pass})* to multiple pages on a read
 - Repeated application of V_{pass} can alter stored values in unread pages
- We **characterize read disturb** on real NAND flash chips
 - Slightly lowering V_{pass} greatly reduces read disturb errors
 - Some flash cells are more prone to read disturb
- **Technique 1: Mitigate read disturb errors online**
 - V_{pass} **Tuning** dynamically finds and applies a lowered V_{pass} per block
 - Flash memory **lifetime improves by 21%**
- **Technique 2: Recover** after failure to prevent data loss
 - ***Read Disturb Oriented Error Recovery*** (RDR) selectively corrects cells more susceptible to read disturb errors
 - Reduces raw bit error rate (RBER) by up to 36%

Key Observation 1: Slightly lowering V_{pass} greatly reduces read disturb errors

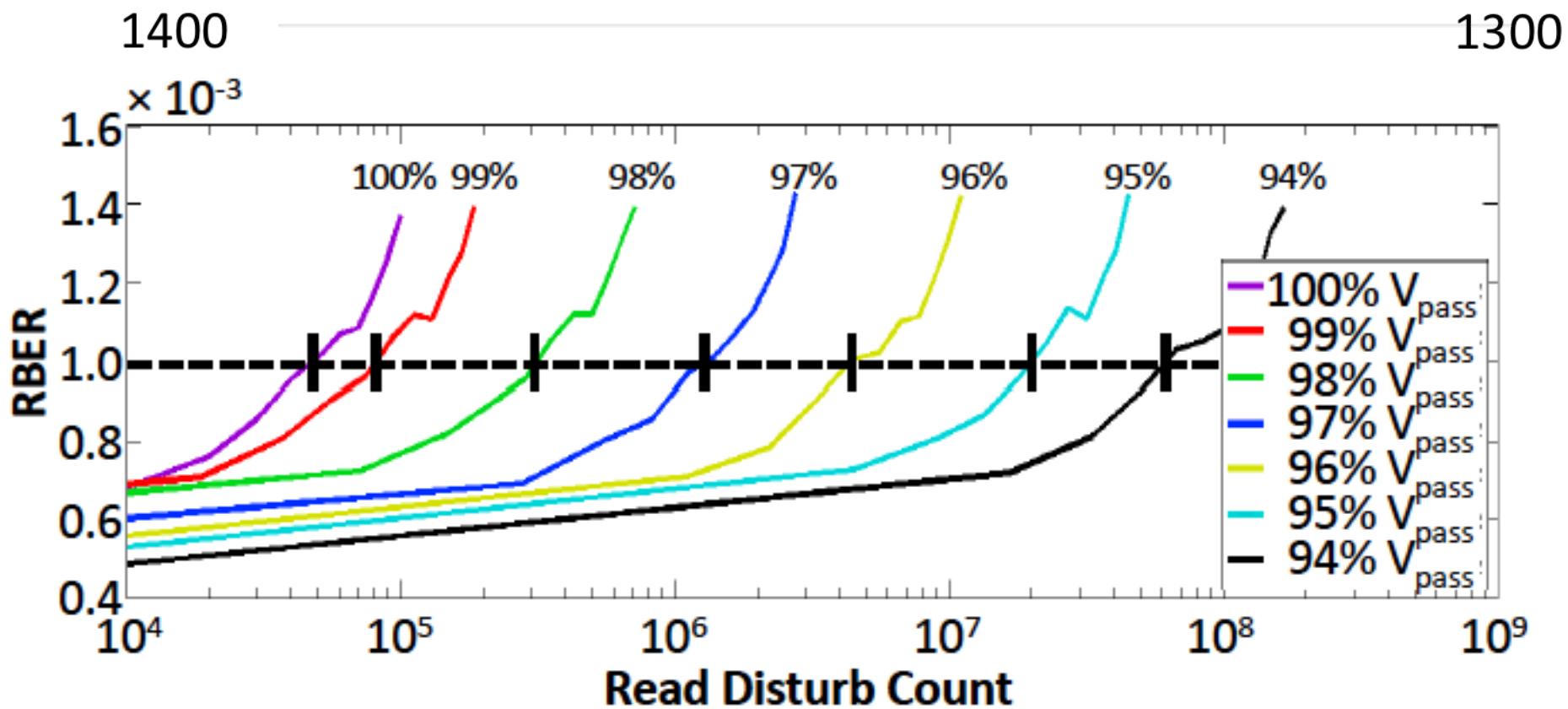


Fig. 11. Raw bit error rate vs. read disturb count for different V_{pass} values, for flash memory under 8K P/E cycles of wear.

Percentage of Vpass Reduction

Outline

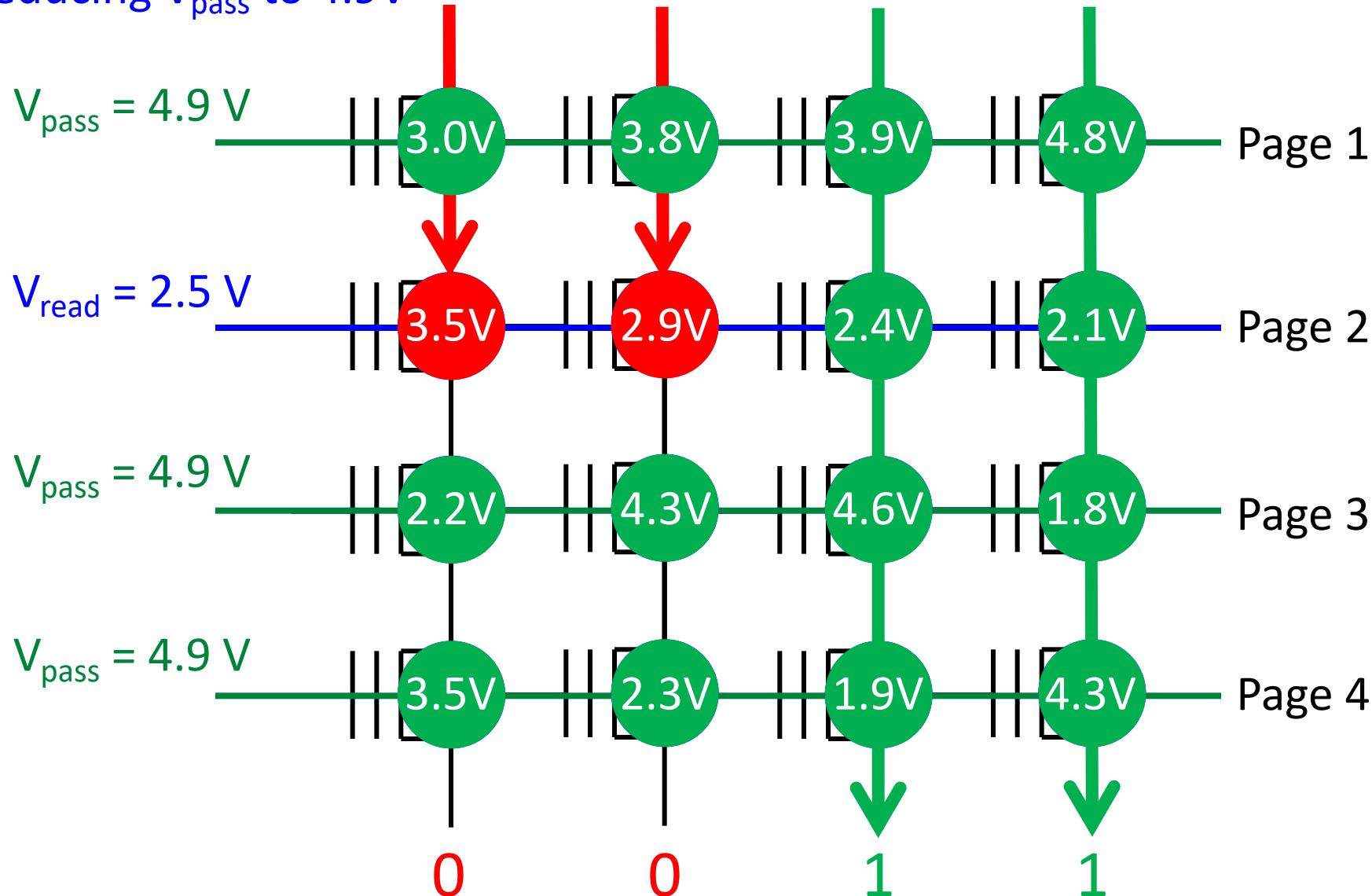
- Background (Problem and Goal)
- Key Experimental Observations
- Mitigation: V_{pass} Tuning
- Recovery: Read Disturb Oriented Error Recovery
- Conclusion

Read Disturb Mitigation: V_{pass} Tuning

- Key Idea: Dynamically find and apply a lowered V_{pass}
- Trade-off for lowering V_{pass}
 - + Allows more read disturbs
 - Induces more read errors

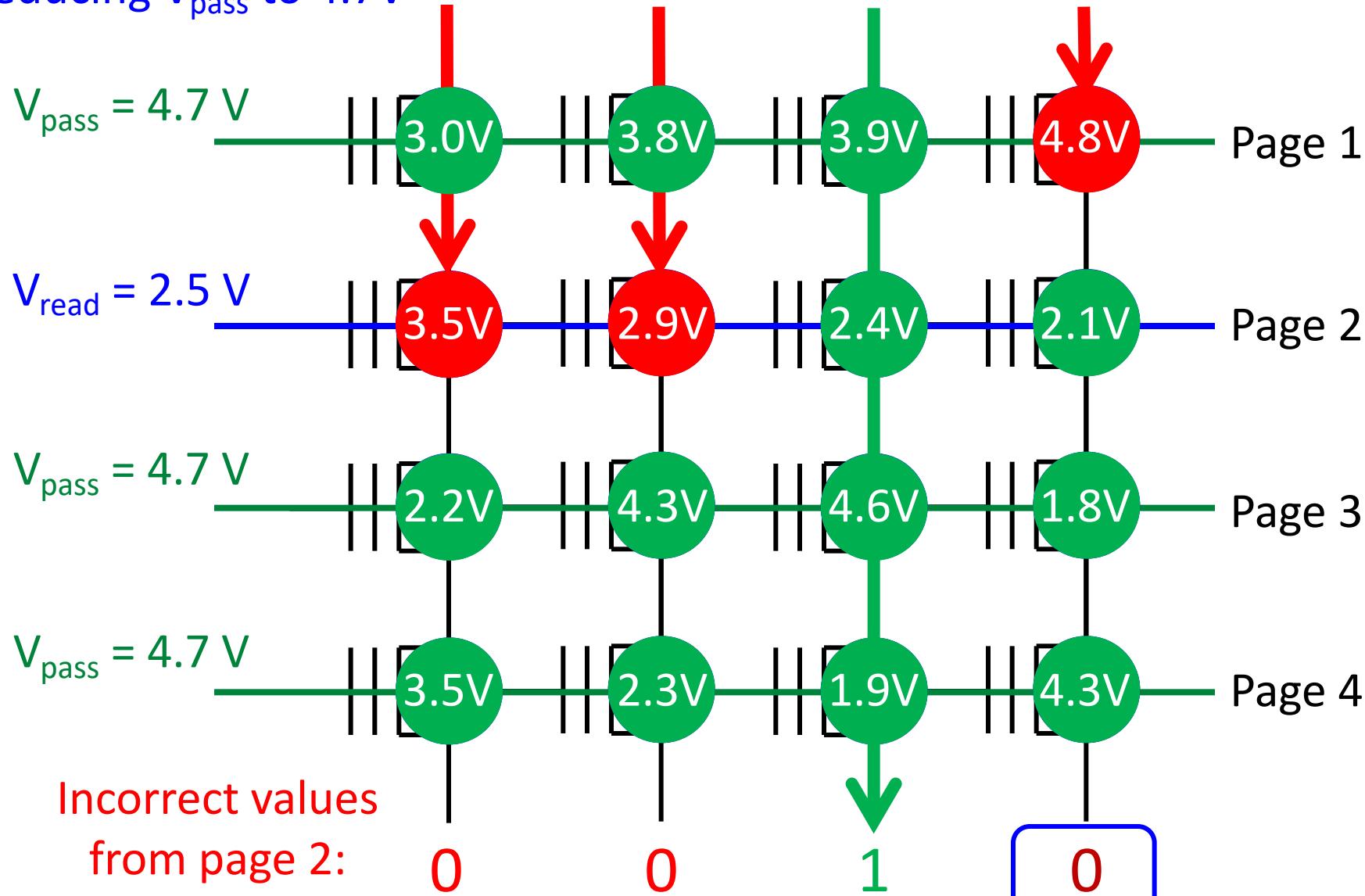
Read Errors Induced by V_{pass} Reduction

Reducing V_{pass} to 4.9V

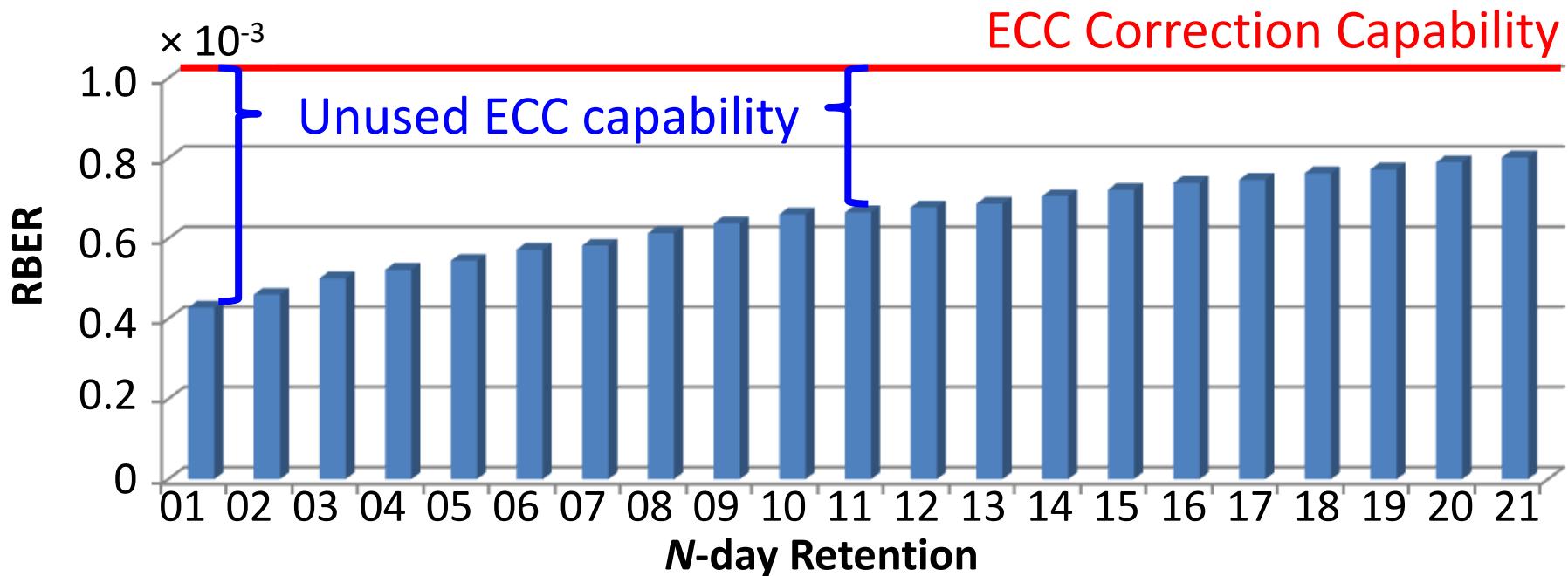


Read Errors Induced by V_{pass} Reduction

Reducing V_{pass} to 4.7V



Utilizing the Unused ECC Capability



1. ECC provisioned for high retention “age”
 2. Unused ECC capability can be used to fix read errors
 3. Unused ECC capability decreases over retention age
- Dynamically adjust V_{pass} so that read errors fully utilize the unused ECC capability

V_{pass} Reduction Trade-Off Summary

- Today: Conservatively set V_{pass} to a high voltage
 - Accumulates more read disturb errors at the end of each refresh interval
 - + No read errors
- Idea: Dynamically adjust V_{pass} to unused ECC capability
 - + Minimize read disturb errors
 - Control read errors to be tolerable by ECC
 - If read errors exceed ECC capability, read again with a higher V_{pass} to correct read errors

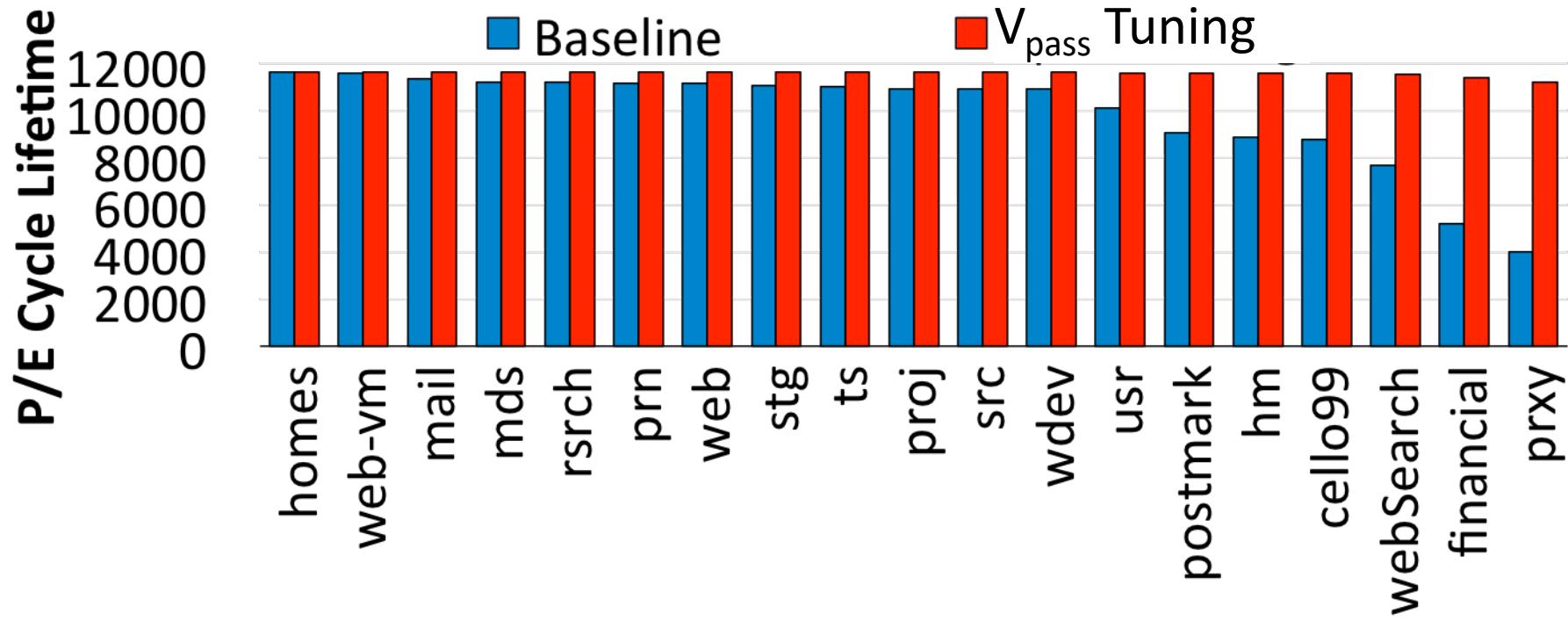
V_{pass} Tuning Steps

- Perform once for each block every day:
 1. Estimate *unused ECC capability (using retention age)*
 2. Aggressively reduce V_{pass} until *read errors exceeds ECC capability*
 3. Gradually increase V_{pass} until read error becomes just less than ECC capability

Evaluation of V_{pass} Tuning

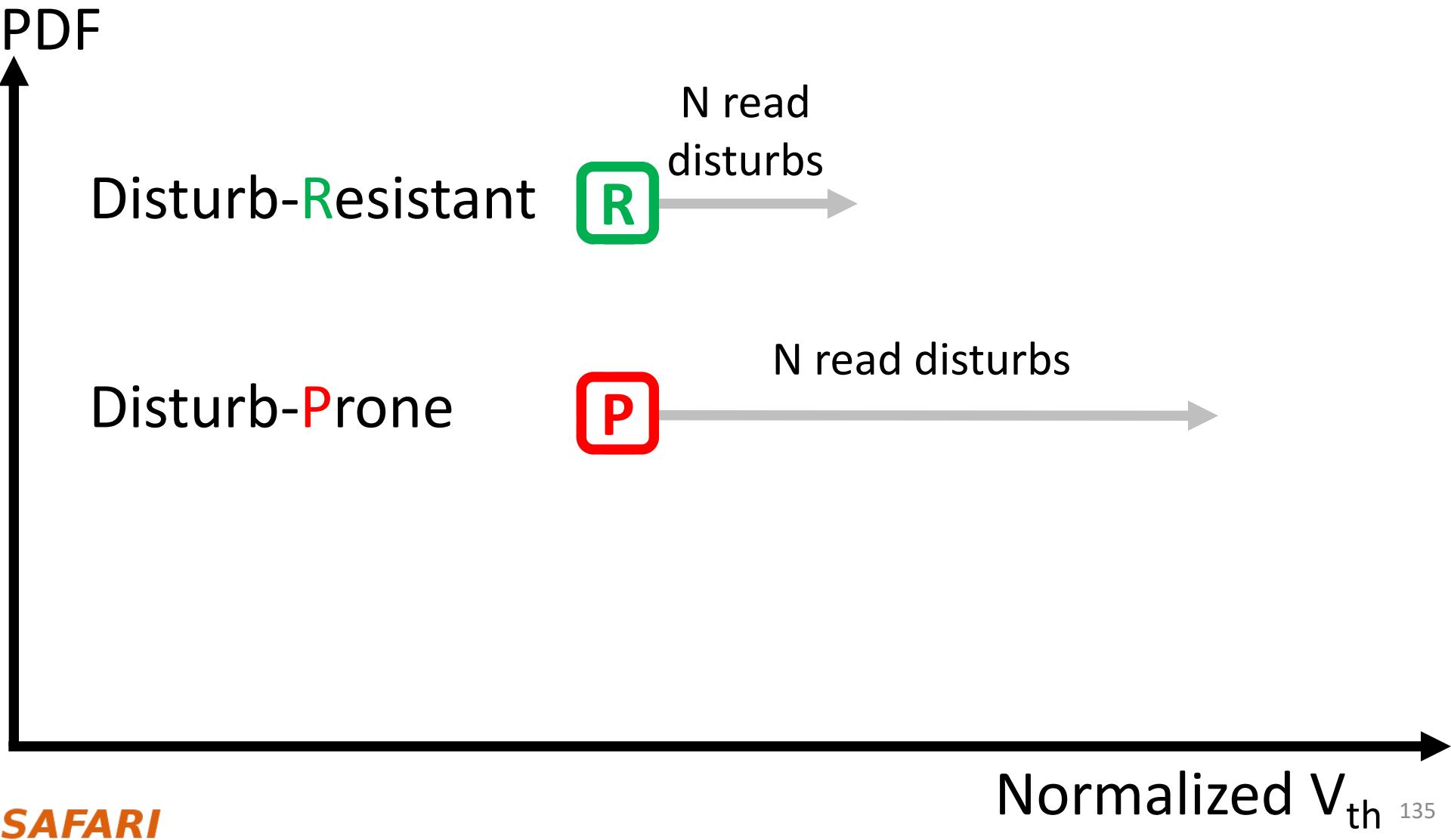
- 19 real workload I/O traces
- Assume 7-day refresh period
- Similar methodology as before to determine acceptable V_{pass} reduction
- Overhead for a 512 GB flash drive:
 - 128 KB storage overhead for per-block V_{pass} setting and worst-case page
 - 24.34 sec/day average V_{pass} Tuning overhead

V_{pass} Tuning Lifetime Improvements



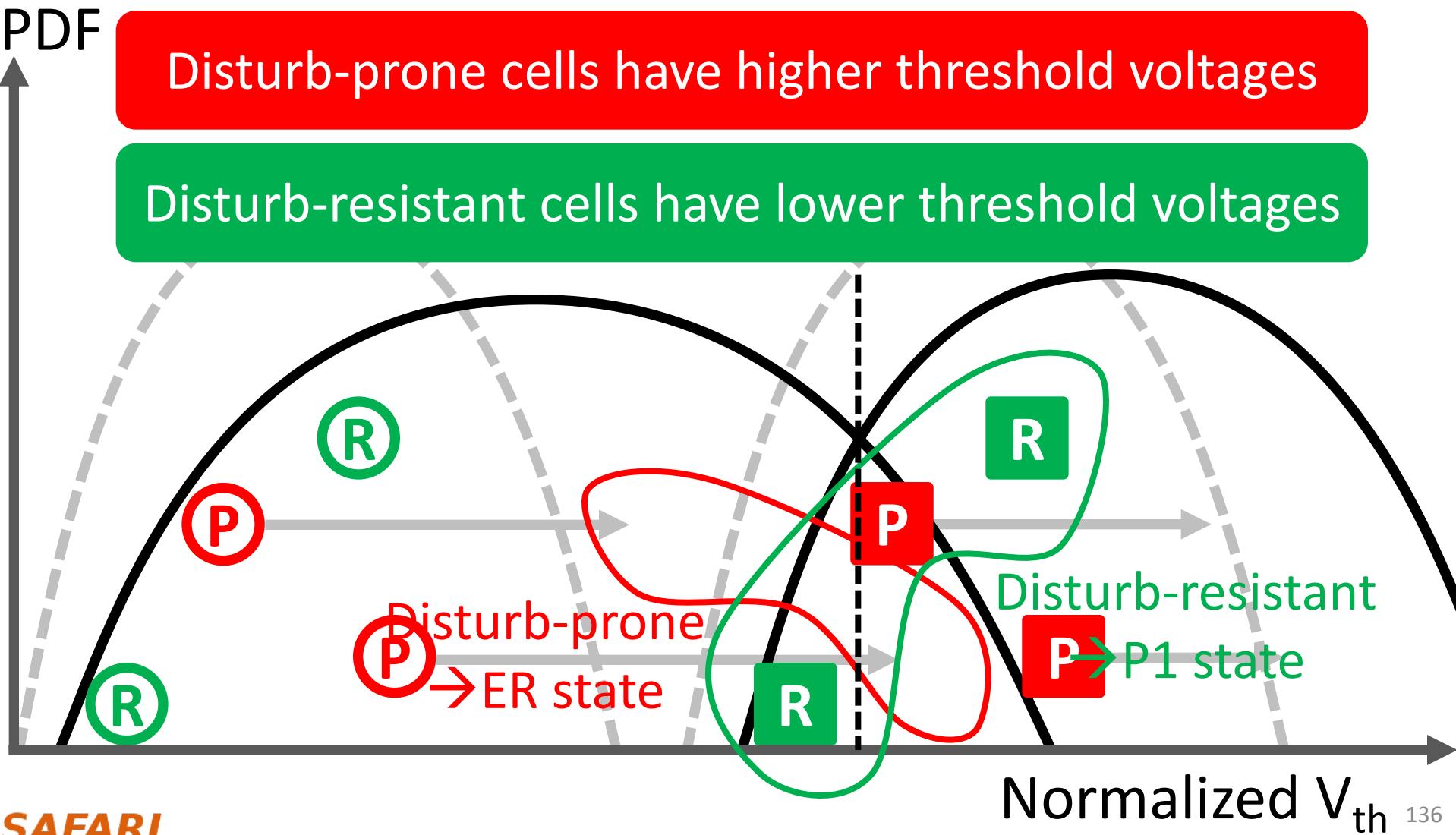
Average lifetime improvement: 21.0%

Read Disturb Prone vs. Resistant Cells



Observation 2: Some Flash Cells Are More Prone to Read Disturb

After 250K read disturbs:



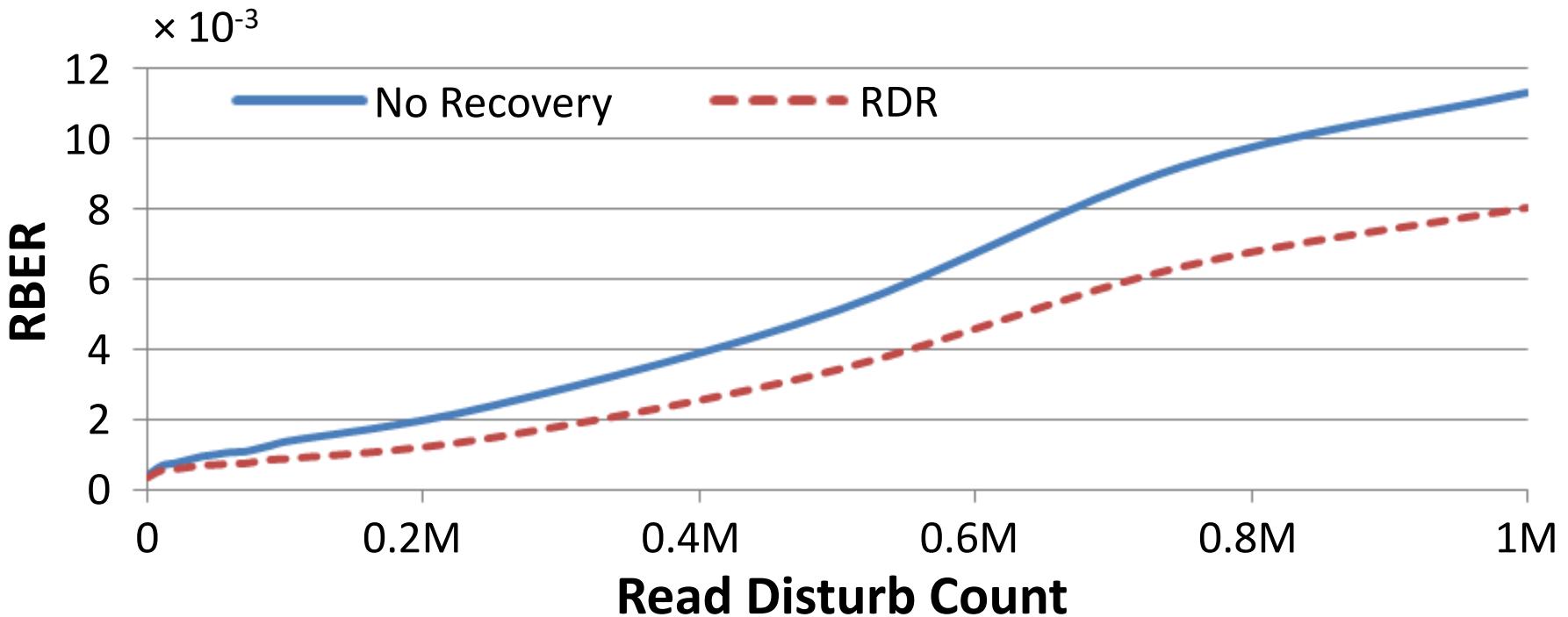
Read Disturb Oriented Error Recovery (RDR)

- Triggered by an uncorrectable flash error
 - Back up all valid data in the faulty block
 - Disturb the faulty page 100K times (more)
 - Compare V_{th} 's before and after read disturb
 - Select cells susceptible to flash errors ($V_{ref}-\sigma < V_{th} < V_{ref}+\sigma$)
 - Predict among these susceptible cells
 - Cells with more V_{th} shifts are disturb-prone → Lower V_{th} state
 - Cells with less V_{th} shifts are disturb-resistant → Higher V_{th} state

Reduces total error count by up to 36% @ 1M read disturbs

ECC can be used to correct the remaining errors

RDR Evaluation



Reduces total error counts by up to 36% @ 1M read disturbances
ECC can be used to correct the remaining errors

More on Flash Read Disturb Errors [DSN'15]

- Yu Cai, Yixin Luo, Saugata Ghose, Erich F. Haratsch, Ken Mai, and Onur Mutlu,

"Read Disturb Errors in MLC NAND Flash Memory: Characterization and Mitigation"

Proceedings of the 45th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), Rio de Janeiro, Brazil, June 2015.

Read Disturb Errors in MLC NAND Flash Memory: Characterization, Mitigation, and Recovery

Yu Cai, Yixin Luo, Saugata Ghose, Erich F. Haratsch*, Ken Mai, Onur Mutlu
Carnegie Mellon University, *Seagate Technology
yucaicai@gmail.com, {yixinluo, ghose, kenmai, onur}@cmu.edu

Agenda

- Background, Motivation and Approach
 - Experimental Characterization Methodology
 - **Error Analysis and Management**
 - Main Characterization Results
 - Retention-Aware Error Management
 - Threshold Voltage and Program Interference Analysis
 - Read Reference Voltage Prediction
 - Neighbor-Assisted Error Correction
 - Read Disturb Error Handling
 - **Retention Error Handling**
 - Large Scale Field Analysis
 - 3D NAND Flash Memory Reliability
 - Summary
-

Data Retention in Flash Memory

Characterize

retention loss in real NAND chip

Optimize

read performance for old data

Recover

old data after failure

An unfortunate tale about Samsung's SSD 840 read performance degradation

An avalanche of reports emerged last September, when owners of the usually speedy Samsung SSD 840 and SSD 840 EVO detected the drives were no longer performing as they used to.

The issue has to do with older blocks of data: reading old files consistently slower than normal as slow as 30MB/s whereas newly-written files ones used in benchmarks, perform as fast as new – are 500 MB/s for the well regarded SSD 840 EVO. The reason no one had noticed (we reviewed the drive back in September 2013) is that data has to be several weeks old to show the problem. Samsung promptly admitted the issue and proposed a fix.

Reference: (May 5, 2015) Per Hansson, "When SSD Performance Goes Awry"

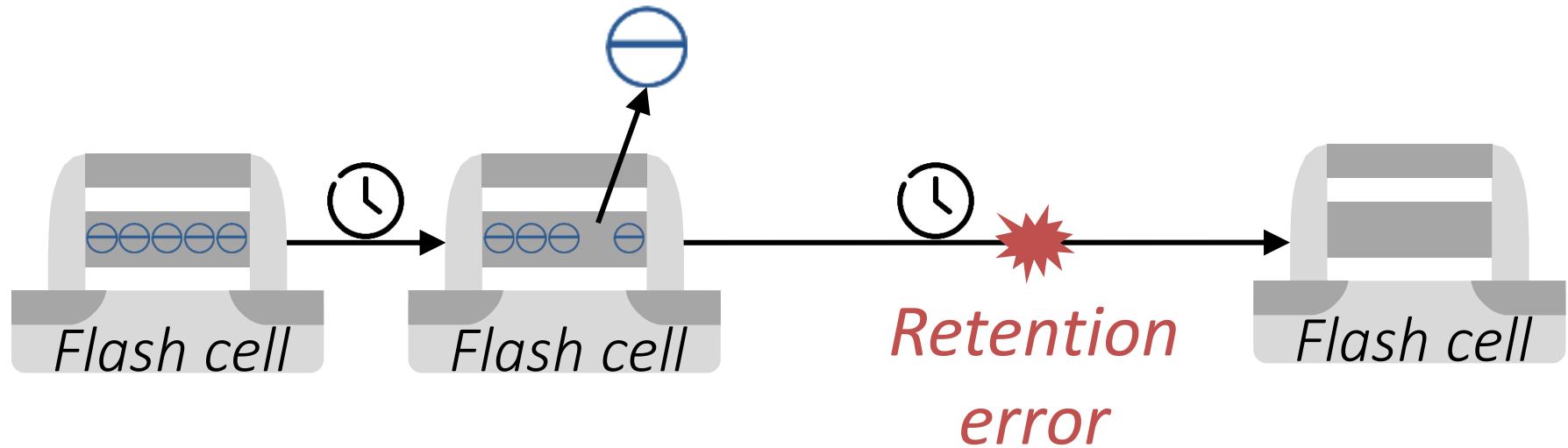
<http://www.techspot.com/article/997-samsung-ssd-read-performance-degradation/>

Why is old data slower?



Retention loss

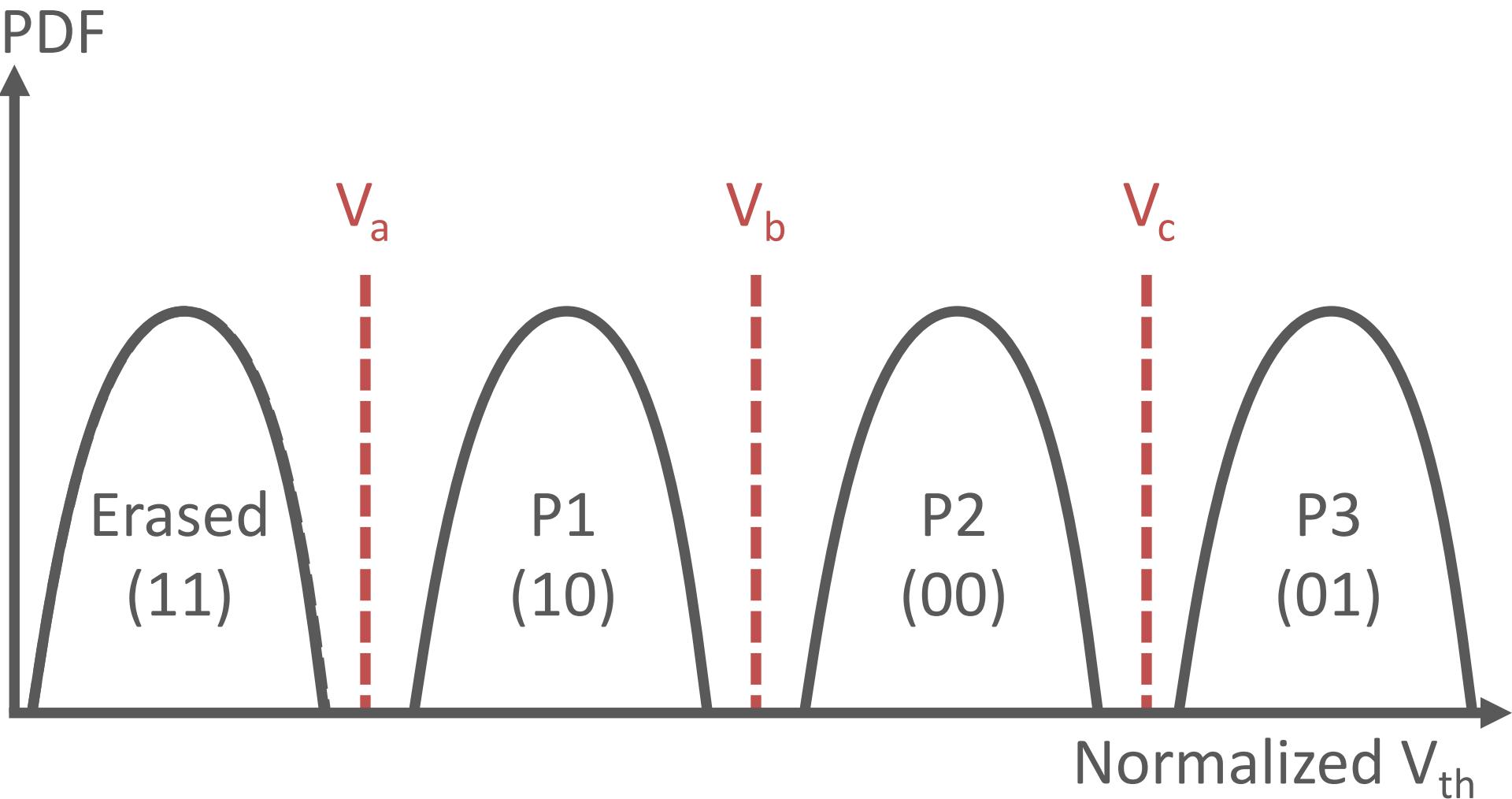
Charge leakage over time



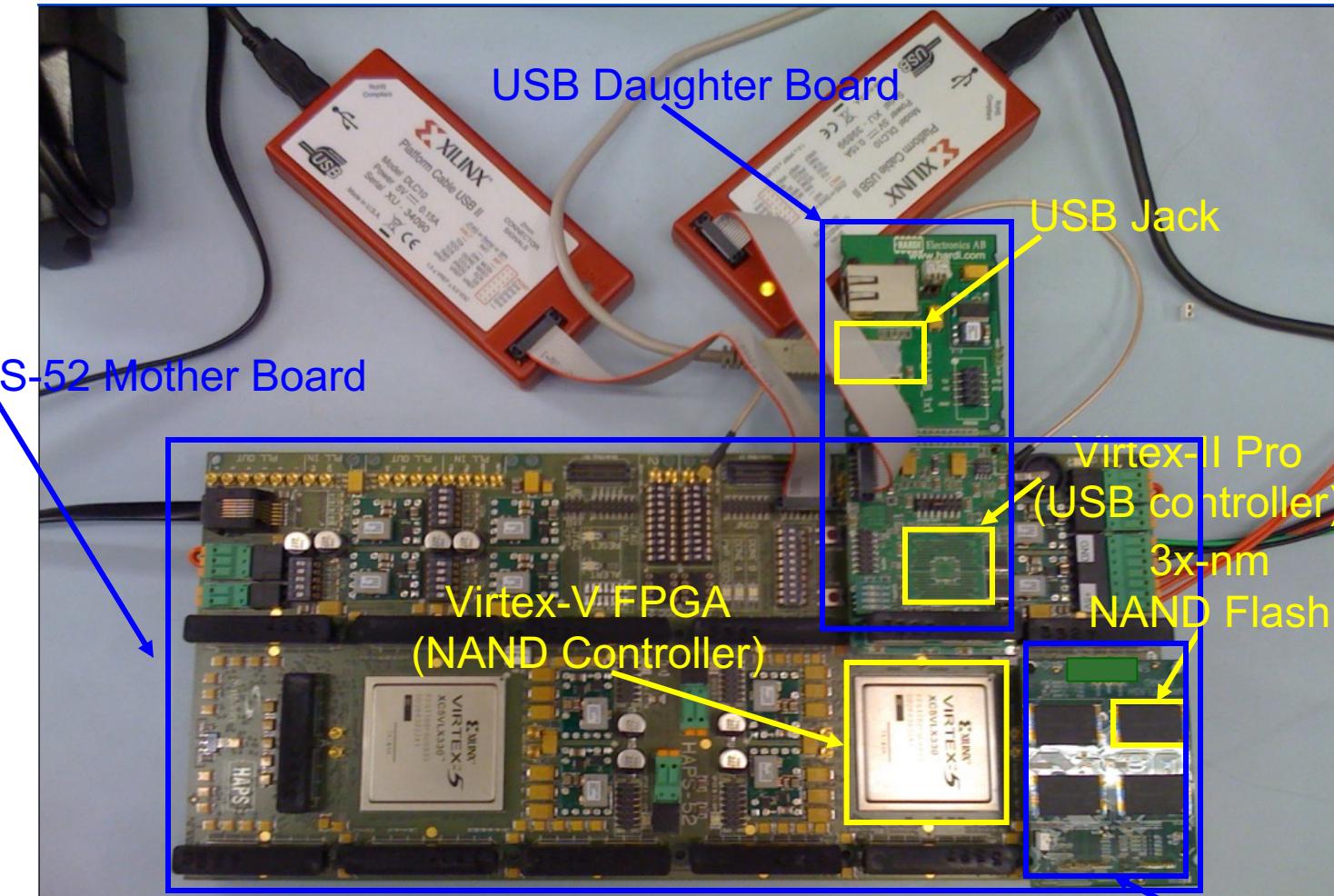
*One dominant source of flash
memory errors [DATE '12, ICCD '12]*

Side effect: Longer read latency

Multi-Level Cell (MLC) threshold voltage distribution



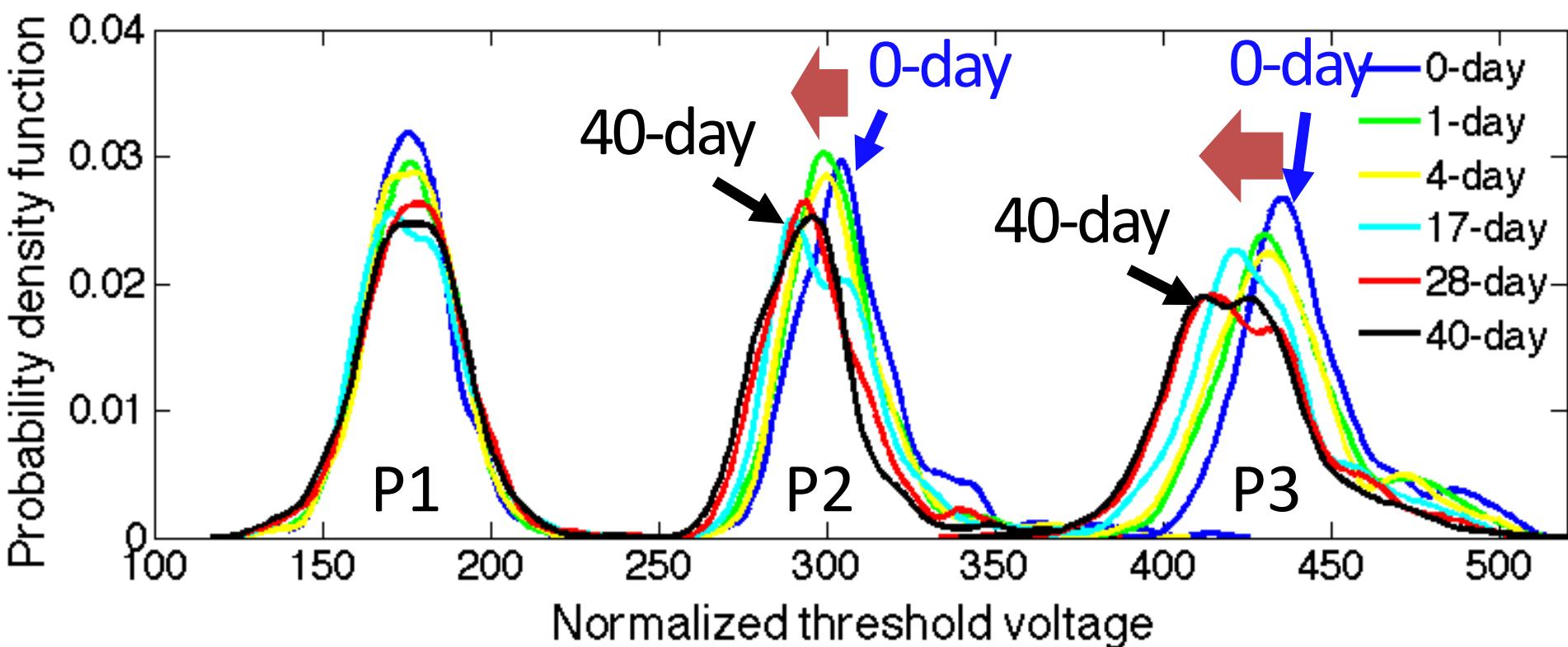
Experimental Testing Platform



[Cai+, FCCM 2011, DATE 2012, ICCD 2012, DATE 2013, ITJ 2013, ICCD 2013, SIGMETRICS 2014, DSN 2015, HPCA 2015]

NAND Daughter Board

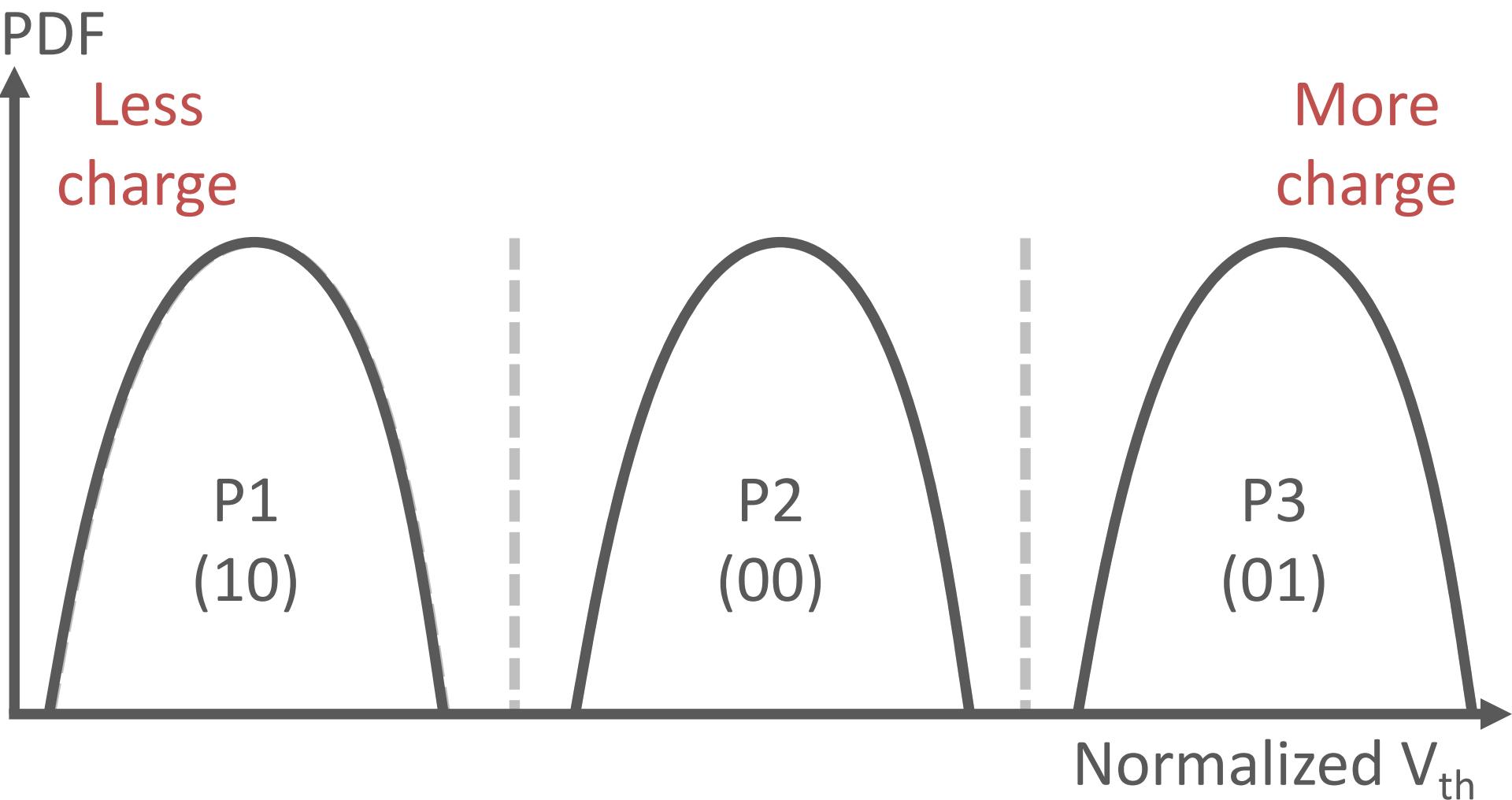
Characterized threshold voltage distribution



Finding: Cell's threshold voltage decreases over time

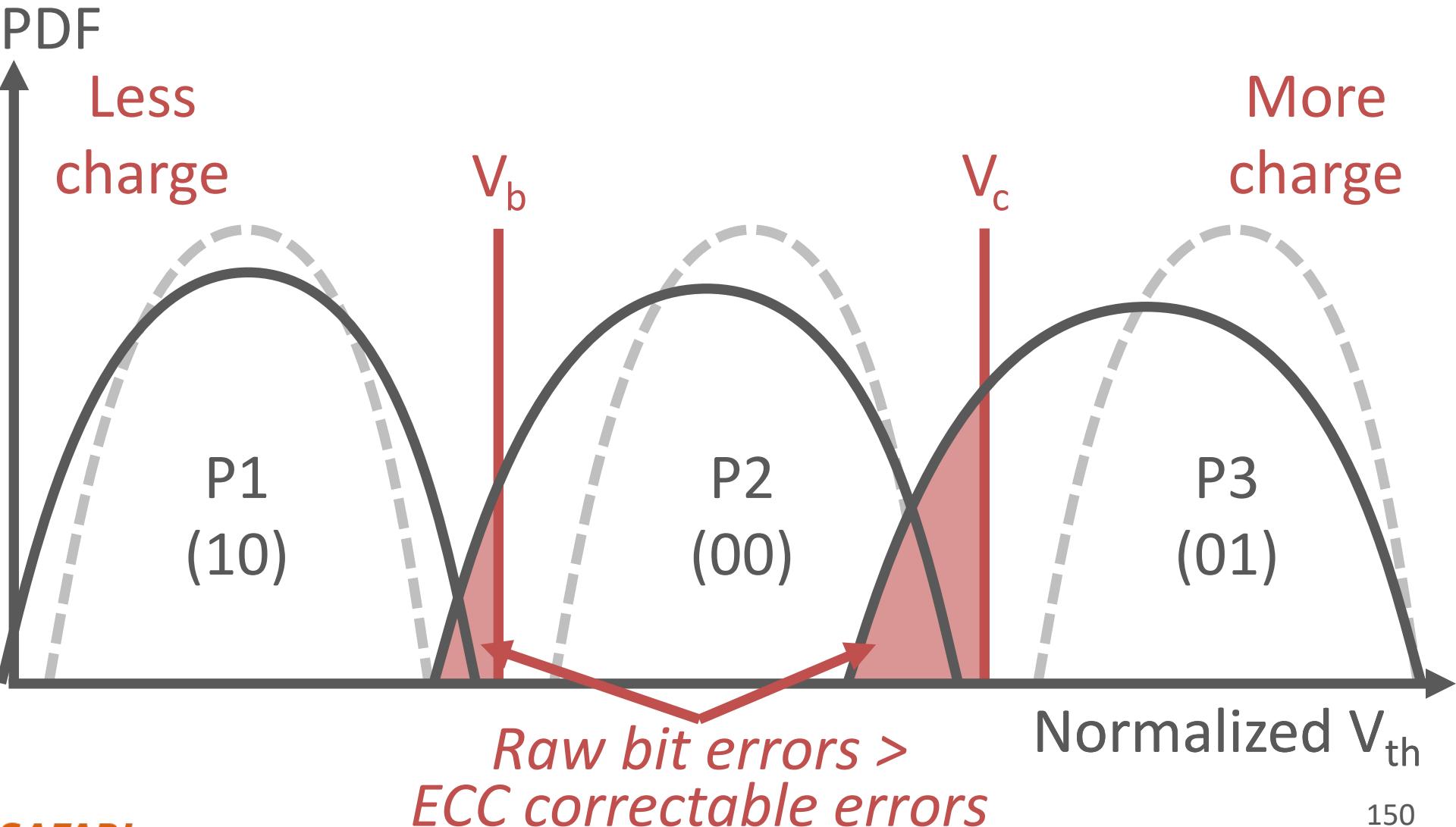
Threshold voltage reduces over time

Old data



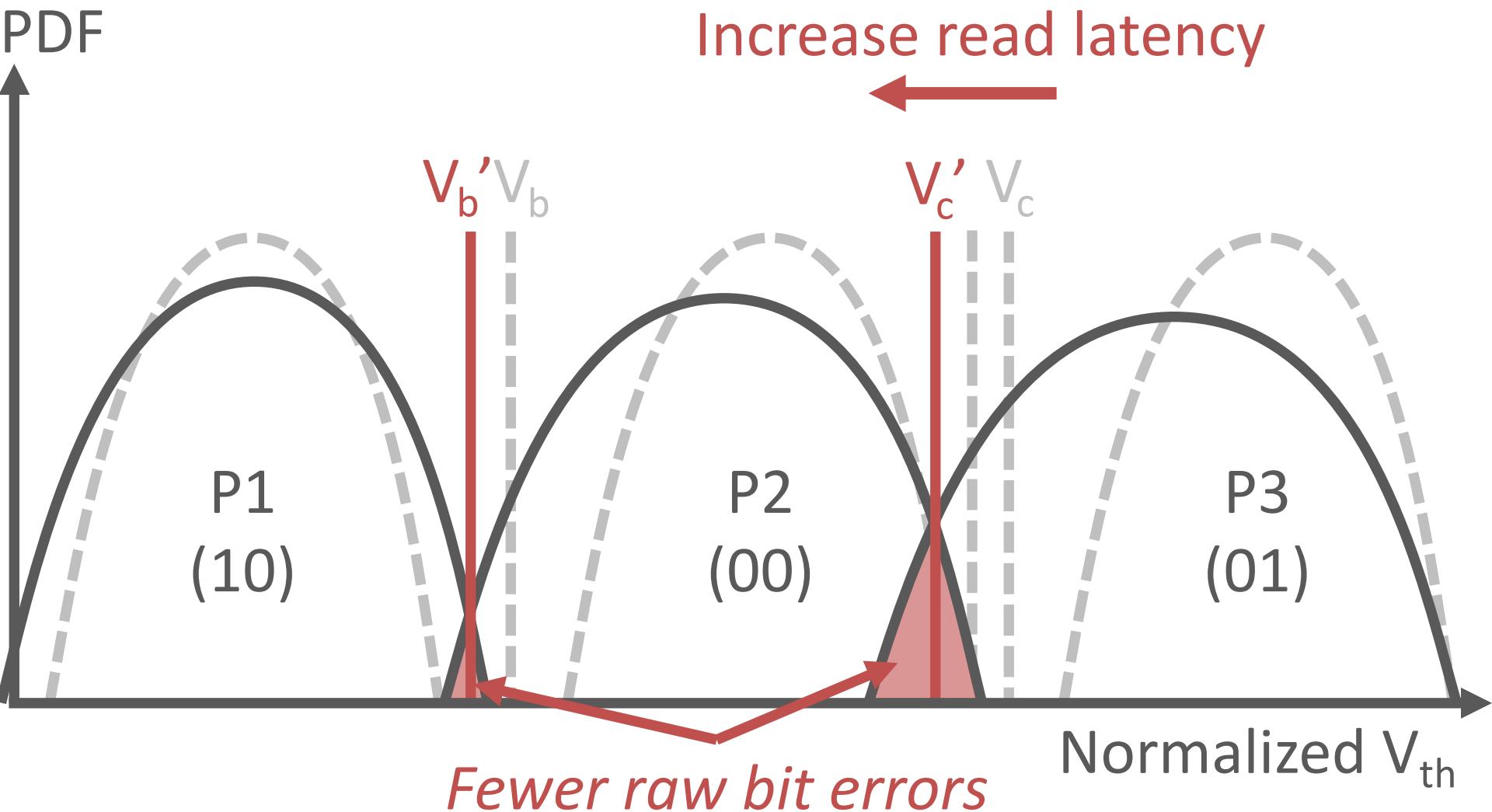
First read attempt fails

Old data



Read-retry

Old data



Why is old data slower?

Retention loss

- *Leak charge over time*
- *Generate retention errors*
- *Require read-retry*
- *Longer read latency*

Characterize

retention loss in real NAND chip

Optimize

read performance for old data

Recover

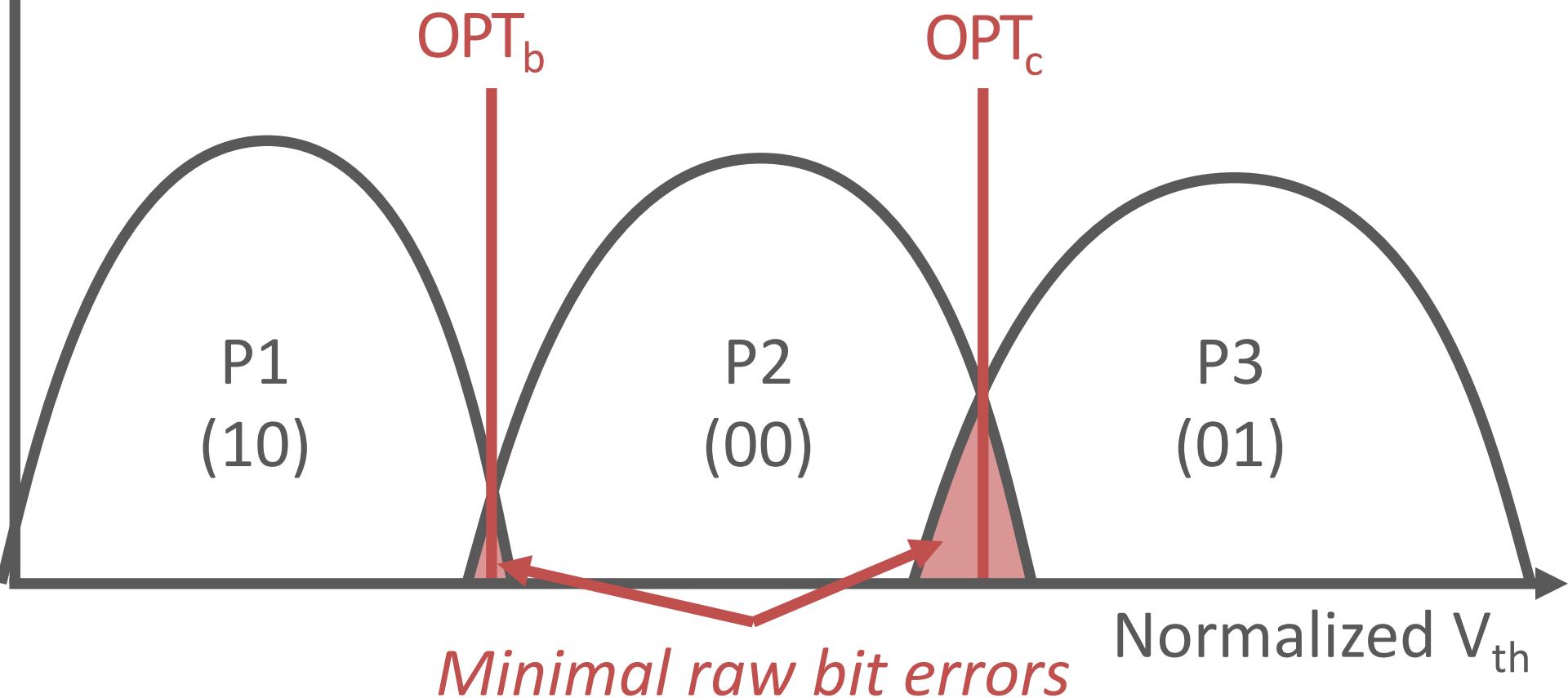
old data after failure

The ideal read voltage

Old data

PDF

*OPT: Optimal read reference voltage
→ minimal read latency*



In reality

- *OPT changes over time due to retention loss*
- *Luckily, OPT change is:*
 - Gradual
 - Uni-directional (decreases over time)

Retention Optimized Reading (ROR)

Components:

1. Online pre-optimization algorithm

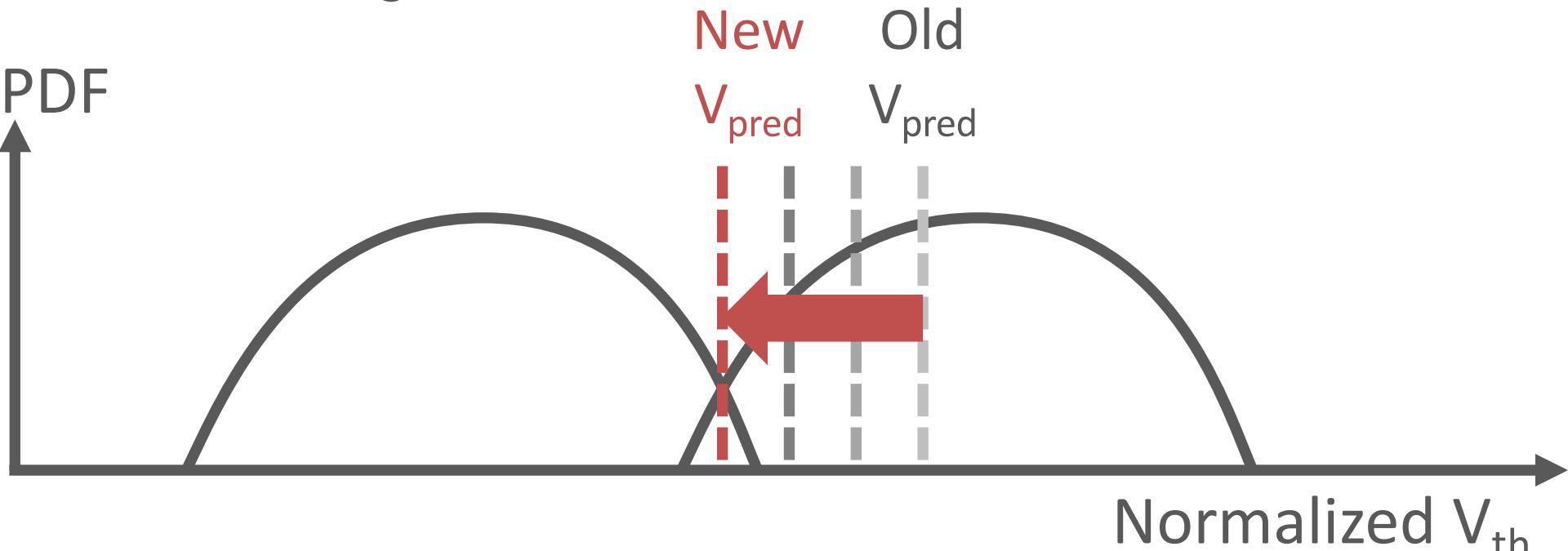
- Learns and records OPT
- Performs in the background once every day

2. Simpler read-retry technique

- If recorded OPT is out-of-date, read-retry with *lower voltage*

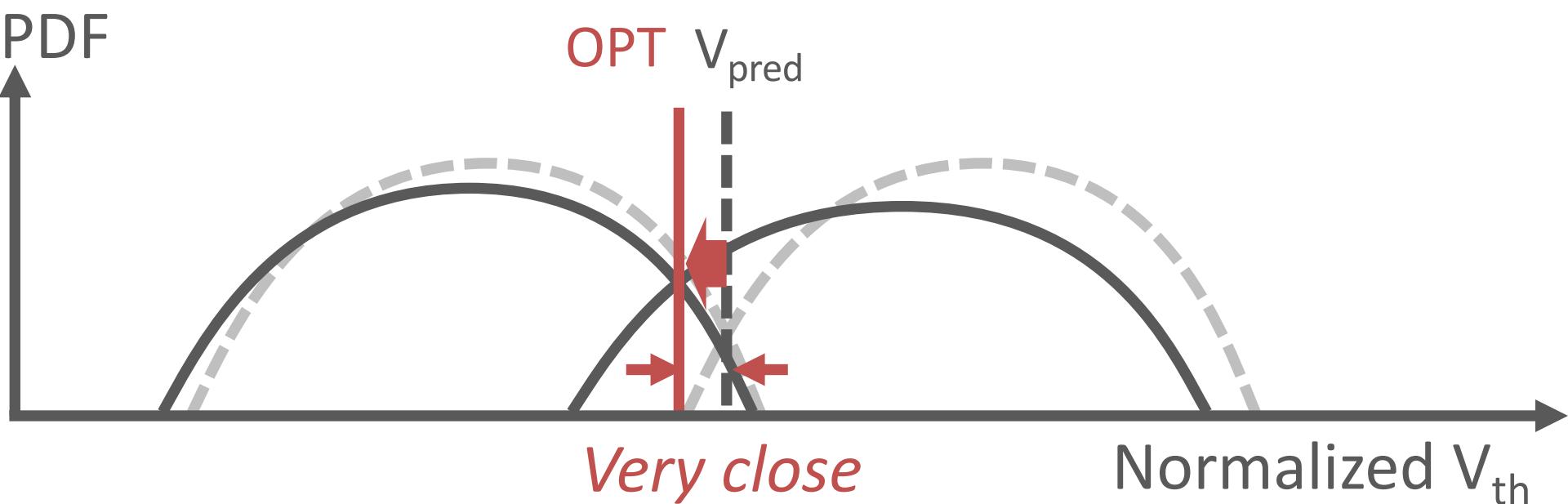
1. Online Pre-Optimization Algorithm

- Triggered periodically (e.g., per day)
- Find and record an OPT as per-block V_{pred}
- Performed in background
- Small storage overhead

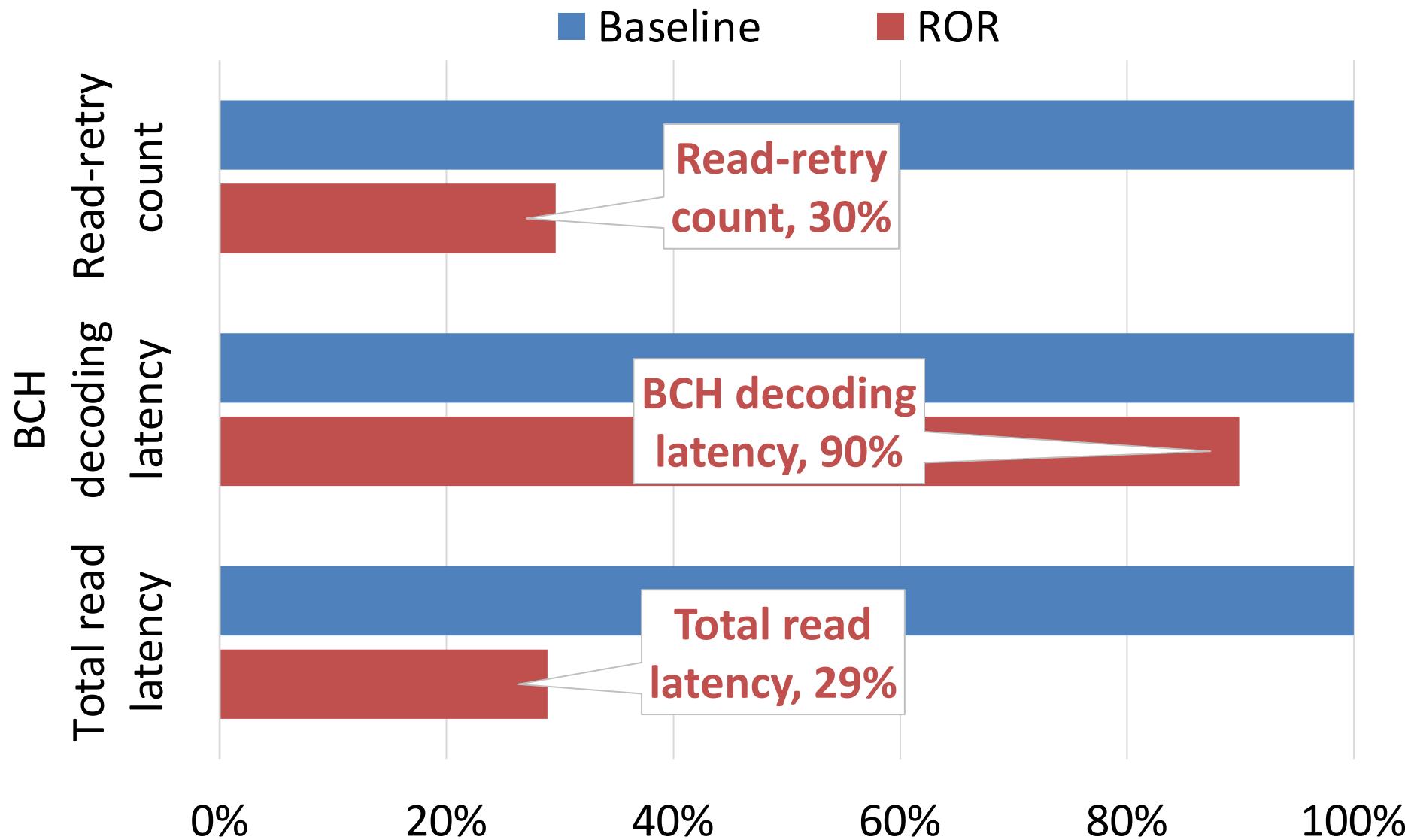


2. Improved Read-Retry Technique

- Performed as normal read
- V_{pred} already close to actual OPT
- Decrease V_{ref} if V_{pred} fails, and retry



ROR result



Retention optimized reading

Retention loss → longer read latency

Optimal read reference voltage (OPT)

→ Shortest read latency

→ Decreases gradually over time (retention)

→ Learn OPT periodically

→ Minimize read-retry & RBER

→ Shorter read latency

Characterize

retention loss in real NAND chip

Optimize

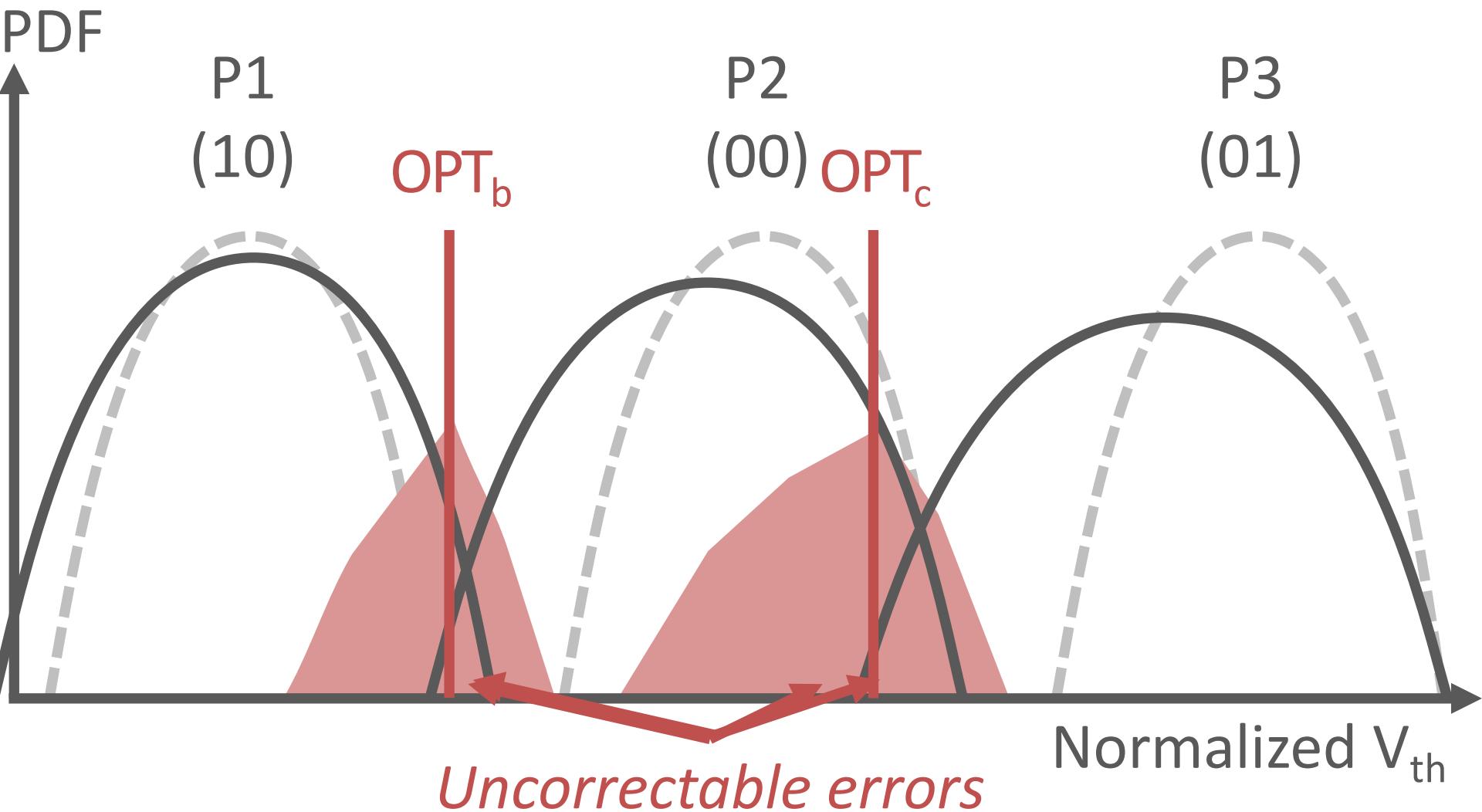
read performance for old data

Recover

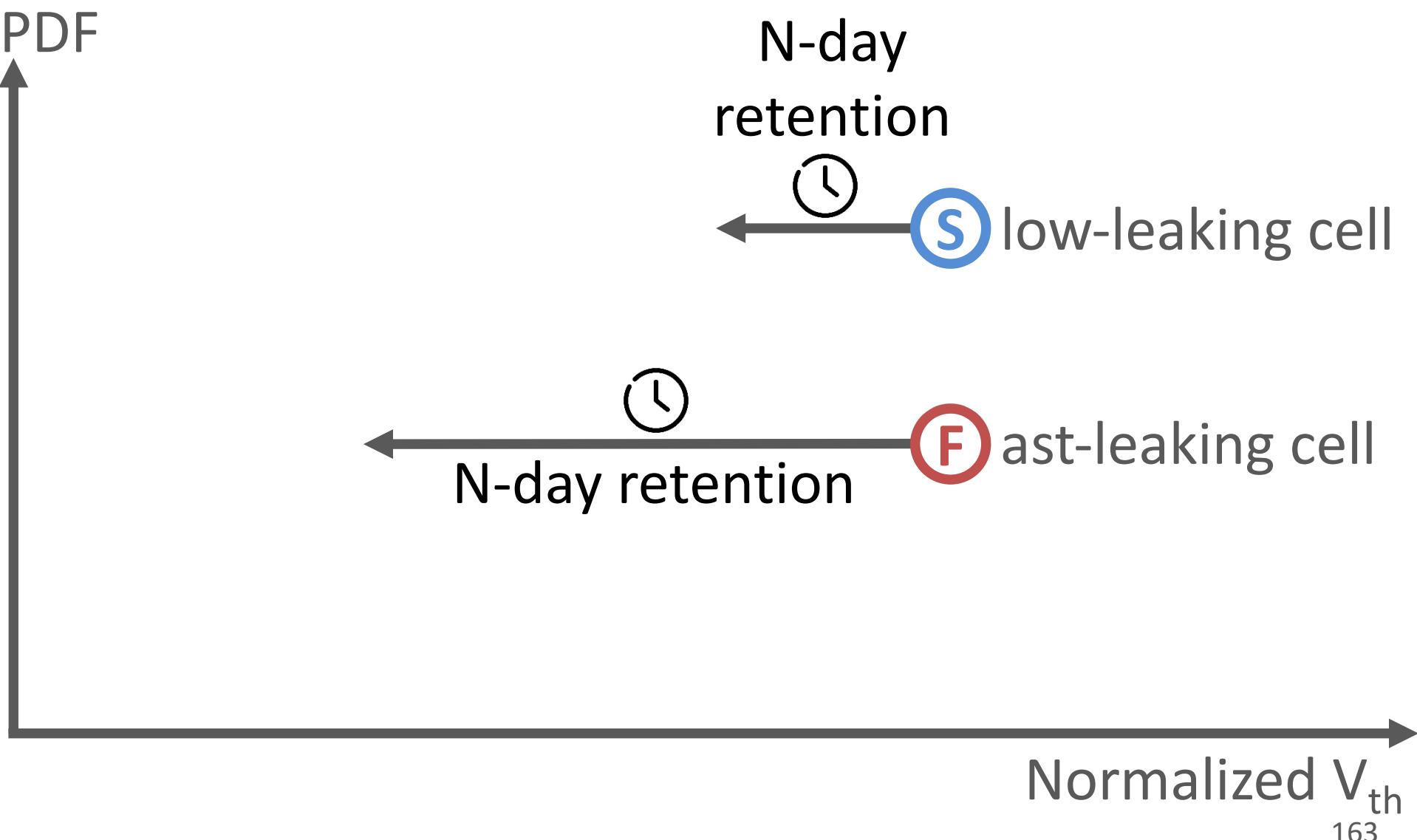
old data after failure

Retention failure

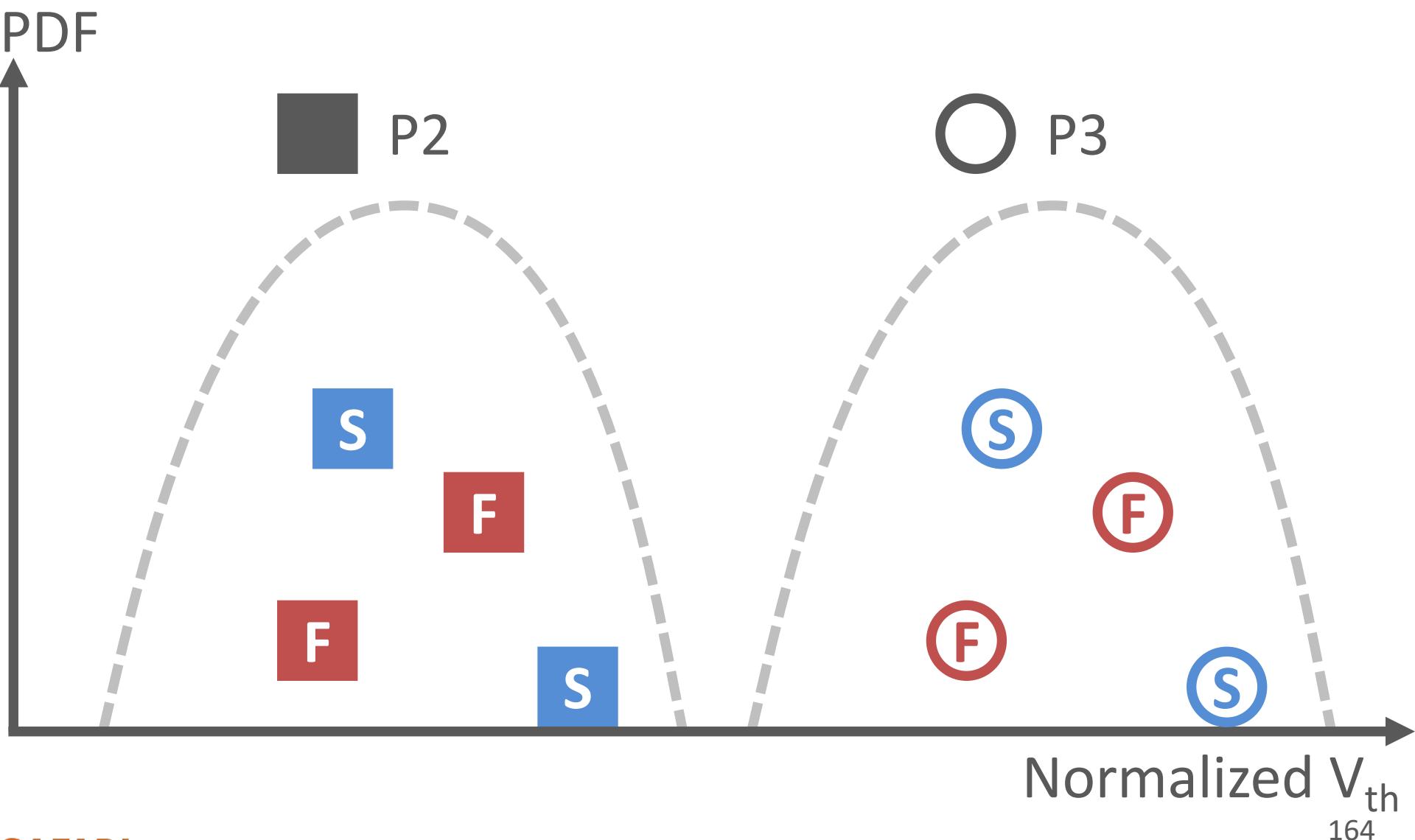
Very old data



Leakage speed variation



A simplified example

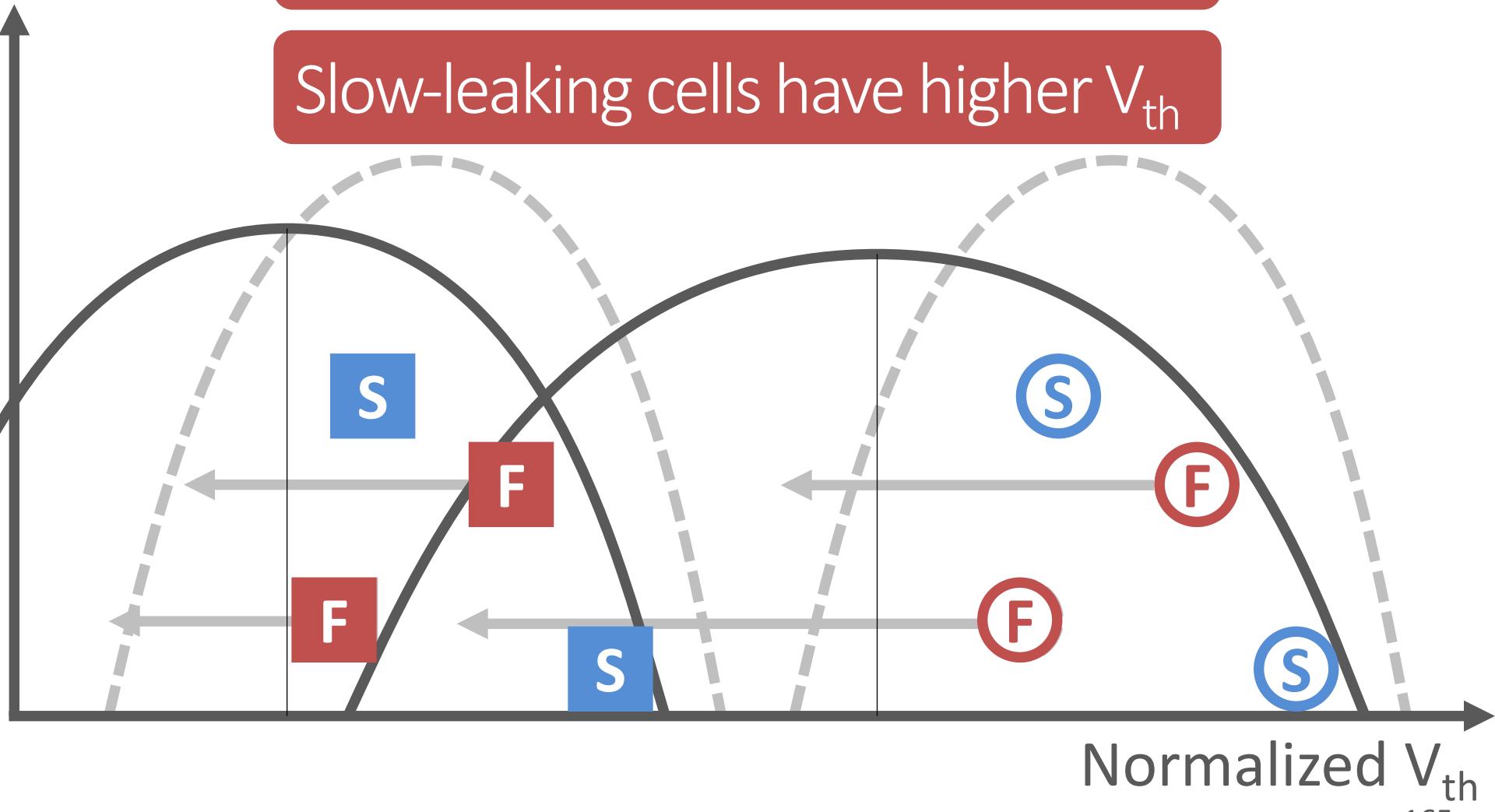


Reading very old data

Very old
PDF

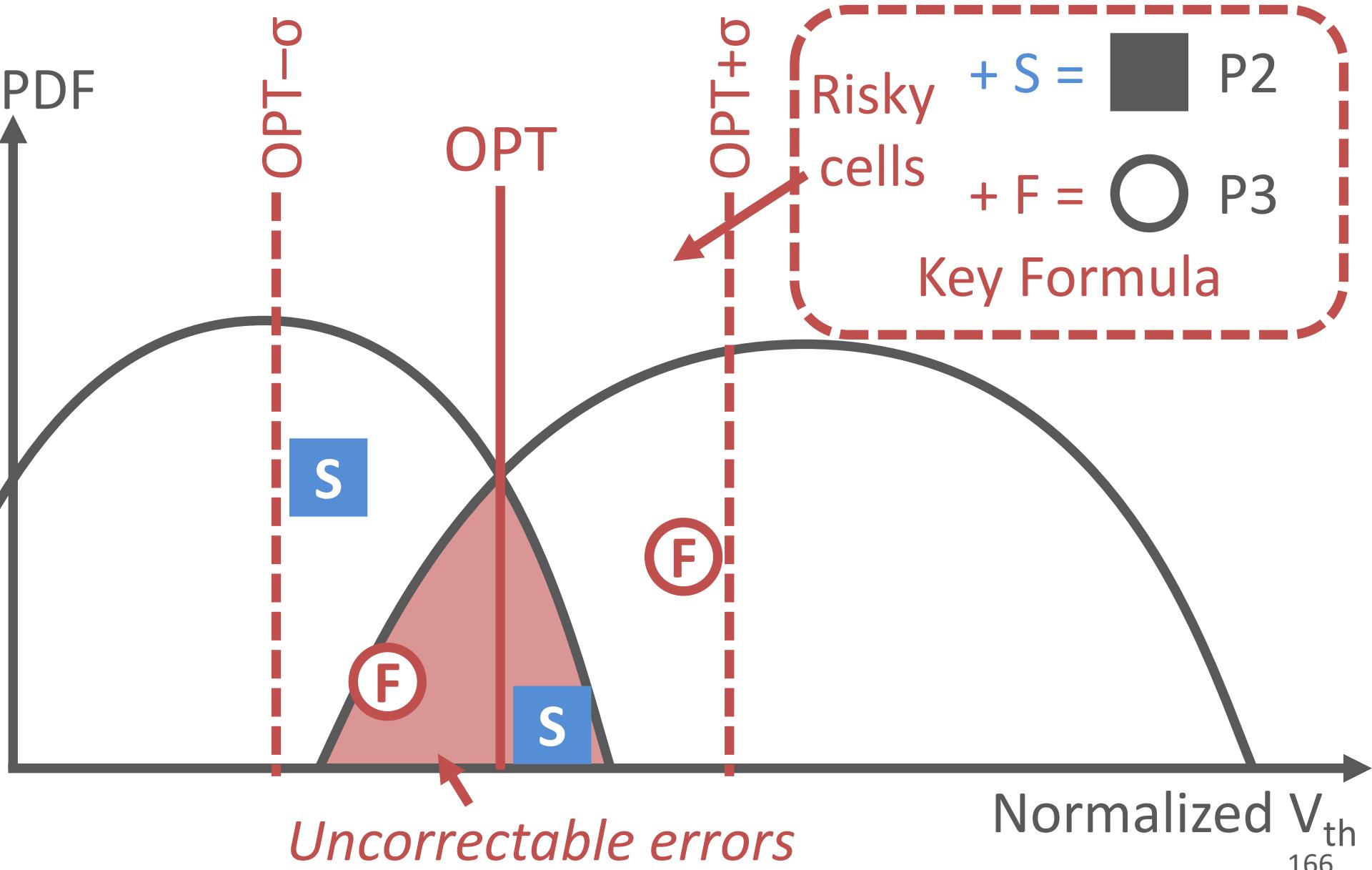
Fast-leaking cells have lower V_{th}

Slow-leaking cells have higher V_{th}



Normalized V_{th}

“Risky” cells

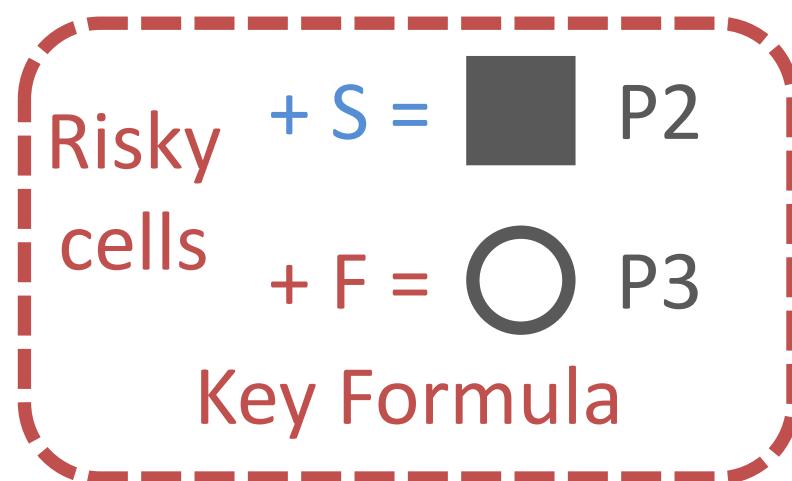


Retention Failure Recovery (RFR)

Key idea: Guess original state of the cell from its leakage speed property

Three steps

1. Identify risky cells
2. Identify fast-/slow-leaking cells
3. Guess original states



RFR Evaluation

*Program with
random data*



28 days

*Detect failure,
backup data*



*12 add'l.
days*

Recover data



- *Expect to eliminate 50% of raw bit errors*
- *ECC can correct remaining errors*

Characterize

retention loss in real NAND chip

Optimize

read performance for old data

Recover

old data after failure

Conclusion

Retention loss → Longer read latency

Retention optimized reading (ROR)

→ Learns OPT periodically

→ 71% shorter read latency

Retention failure recovery (RFR)

→ Use leakage property to guess correct state

→ 50% error reduction before ECC correction

→ Recover data after failure

More on Flash Read Disturb Errors

- Yu Cai, Yixin Luo, Erich F. Haratsch, Ken Mai, and Onur Mutlu,
"Data Retention in MLC NAND Flash Memory: Characterization, Optimization and Recovery"
Proceedings of the 21st International Symposium on High-Performance Computer Architecture (HPCA), Bay Area, CA, February 2015.
[Slides (pptx) (pdf)]

Data Retention in MLC NAND Flash Memory: Characterization, Optimization, and Recovery

Yu Cai, Yixin Luo, Erich F. Haratsch*, Ken Mai, Onur Mutlu
Carnegie Mellon University, *LSI Corporation

yucaicai@gmail.com, yixinluo@cs.cmu.edu, erich.haratsch@lsi.com, {kenmai, omutlu}@ece.cmu.edu

Agenda

- Background, Motivation and Approach
 - Experimental Characterization Methodology
 - **Error Analysis and Management**
 - Main Characterization Results
 - Retention-Aware Error Management
 - Threshold Voltage and Program Interference Analysis
 - Read Reference Voltage Prediction
 - Neighbor-Assisted Error Correction
 - Read Disturb Error Handling
 - Retention Error Handling
 - **Large Scale Field Analysis**
 - 3D NAND Flash Memory Reliability
 - Summary
-

Large Scale Field Analysis of Flash Memory Errors

SSD Error Analysis of Facebook Systems

- First large-scale field study of flash memory errors
- Justin Meza, Qiang Wu, Sanjeev Kumar, and Onur Mutlu,
"A Large-Scale Study of Flash Memory Errors in the Field"
Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS), Portland, OR, June 2015.
[[Slides \(pptx\)](#)] [[Coverage at ZDNet](#)] [[Coverage on The Register](#)] [[Coverage on TechSpot](#)] [[Coverage on The Tech Report](#)]

A Large-Scale Study of Flash Memory Failures in the Field

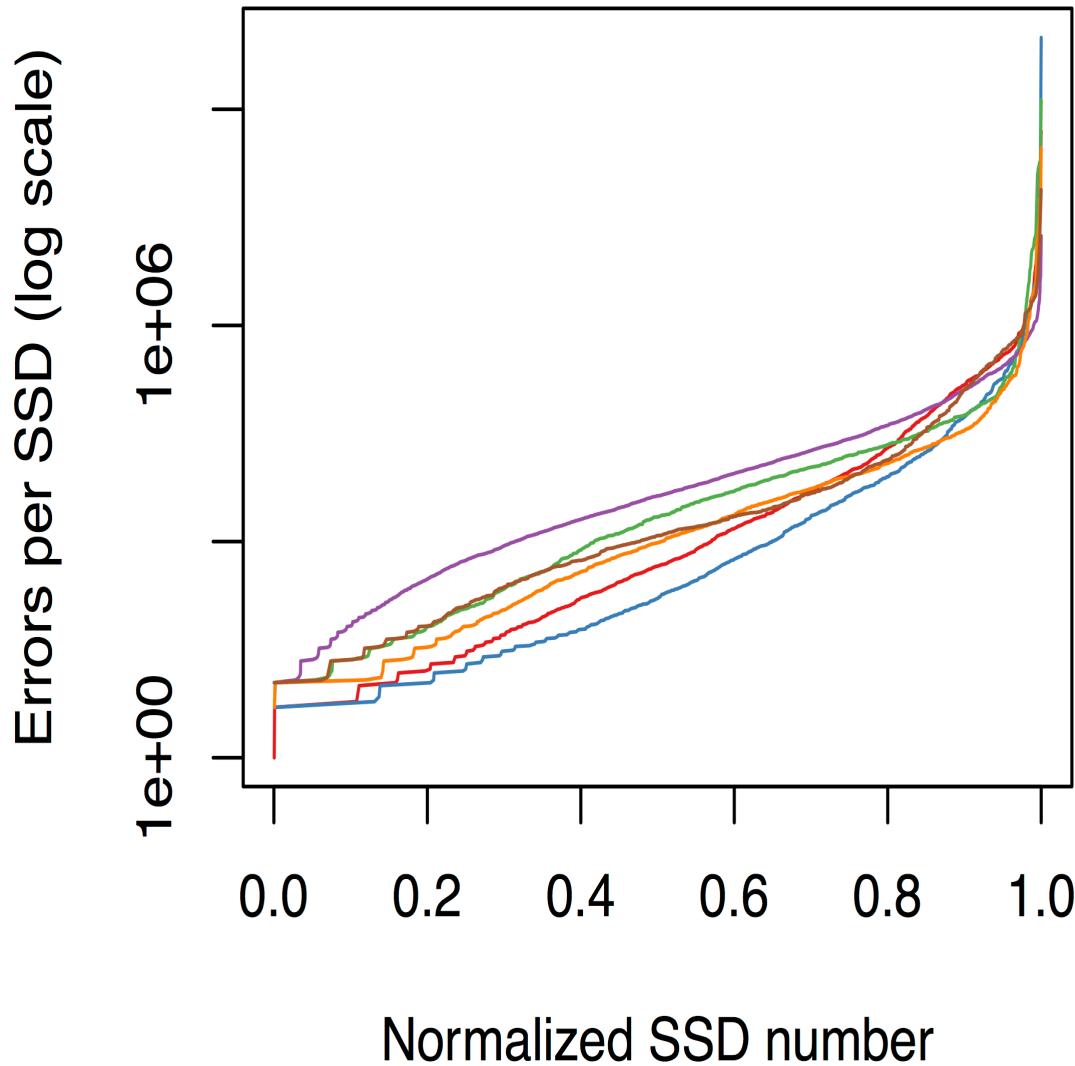
Justin Meza
Carnegie Mellon University
meza@cmu.edu

Qiang Wu
Facebook, Inc.
qwu@fb.com

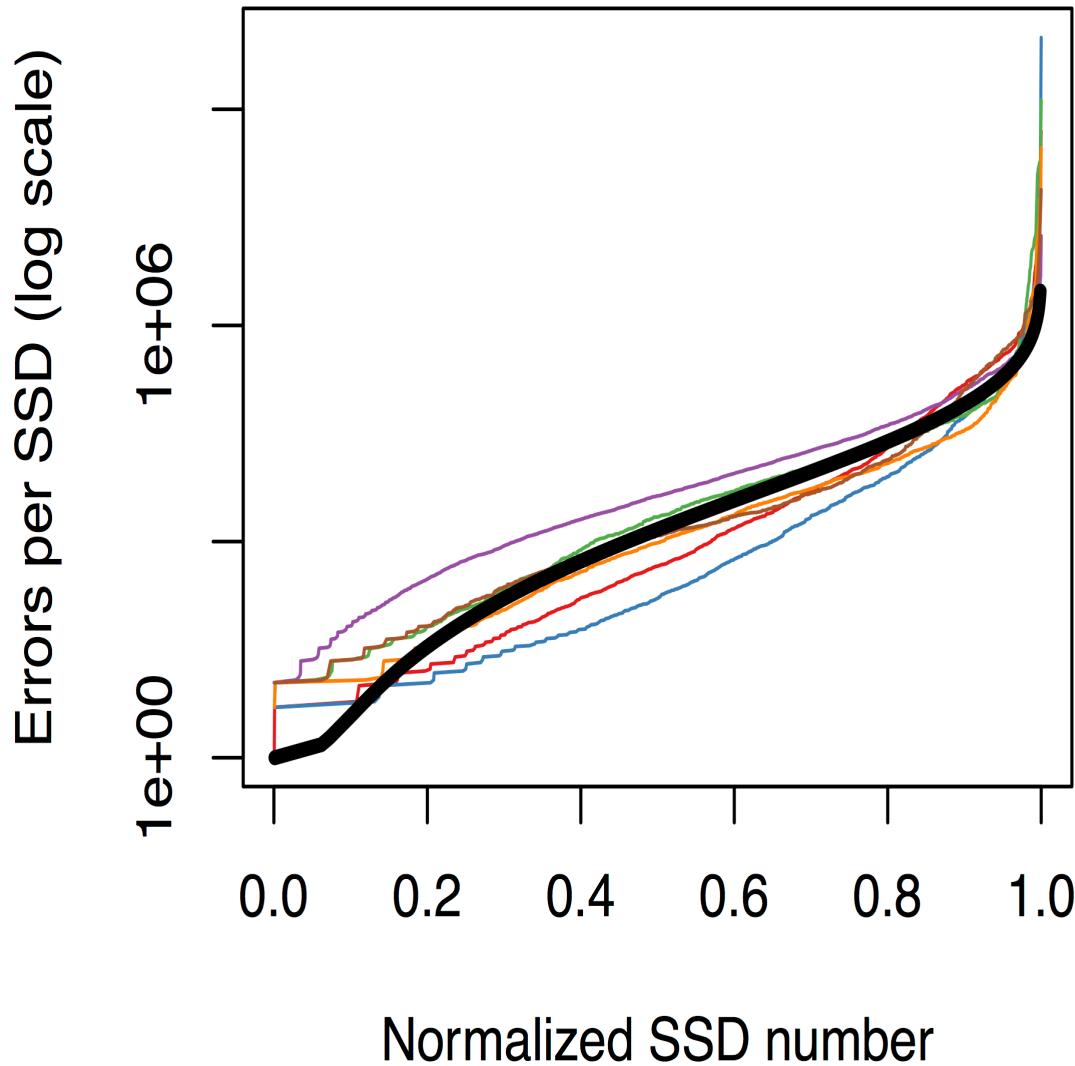
Sanjeev Kumar
Facebook, Inc.
skumar@fb.com

Onur Mutlu
Carnegie Mellon University
onur@cmu.edu

A few SSDs cause most errors



A few SSDs cause most errors



Summary

SSD lifecycle

*Access pattern
dependence*



*Read
disturbance*

Temperature

Summary

SSD lifecycle



Access ***Early detection*** lifecycle period
distinct from hard disk drive
lifecycle.

Temperature

SSD lifecycle

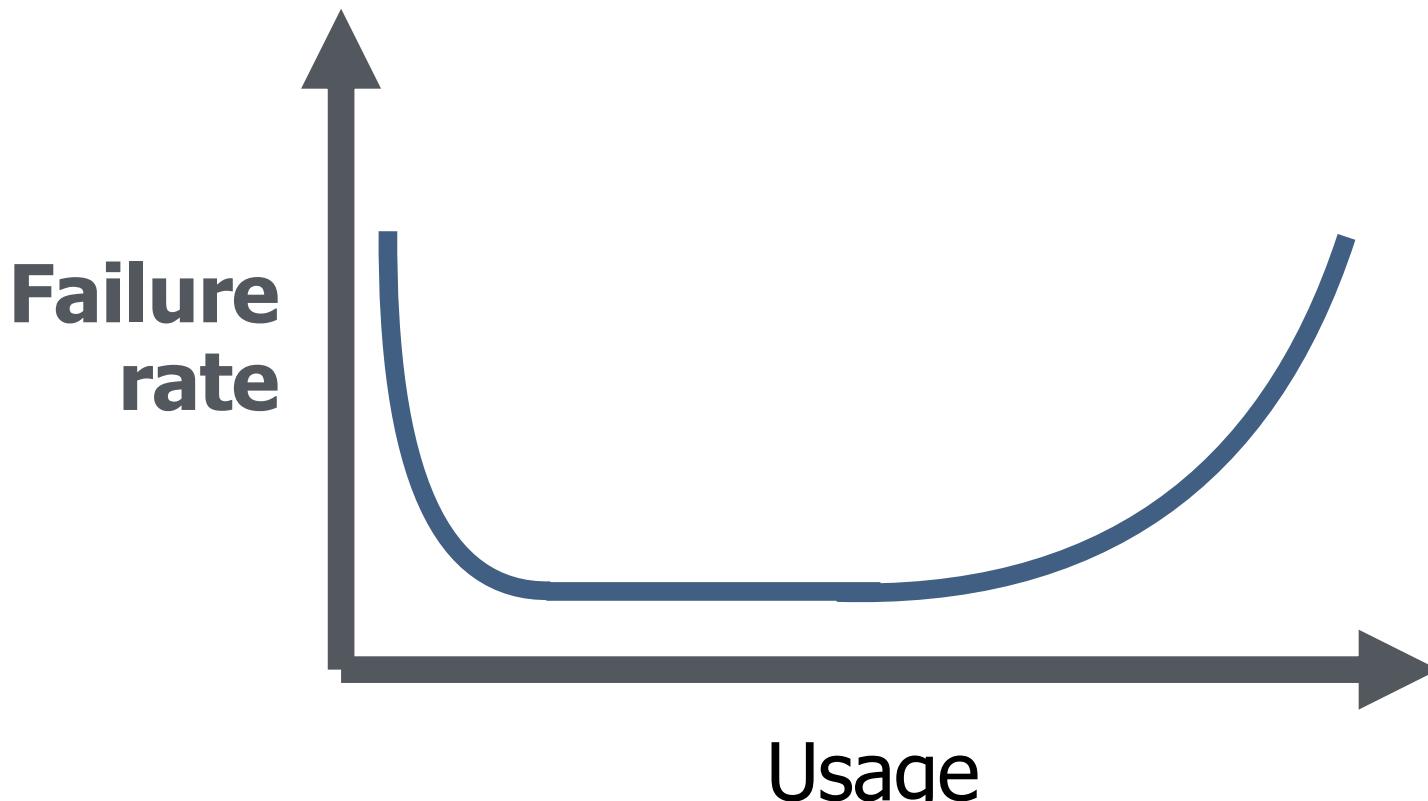
*Access pattern
dependence*



*Read
disturbance*

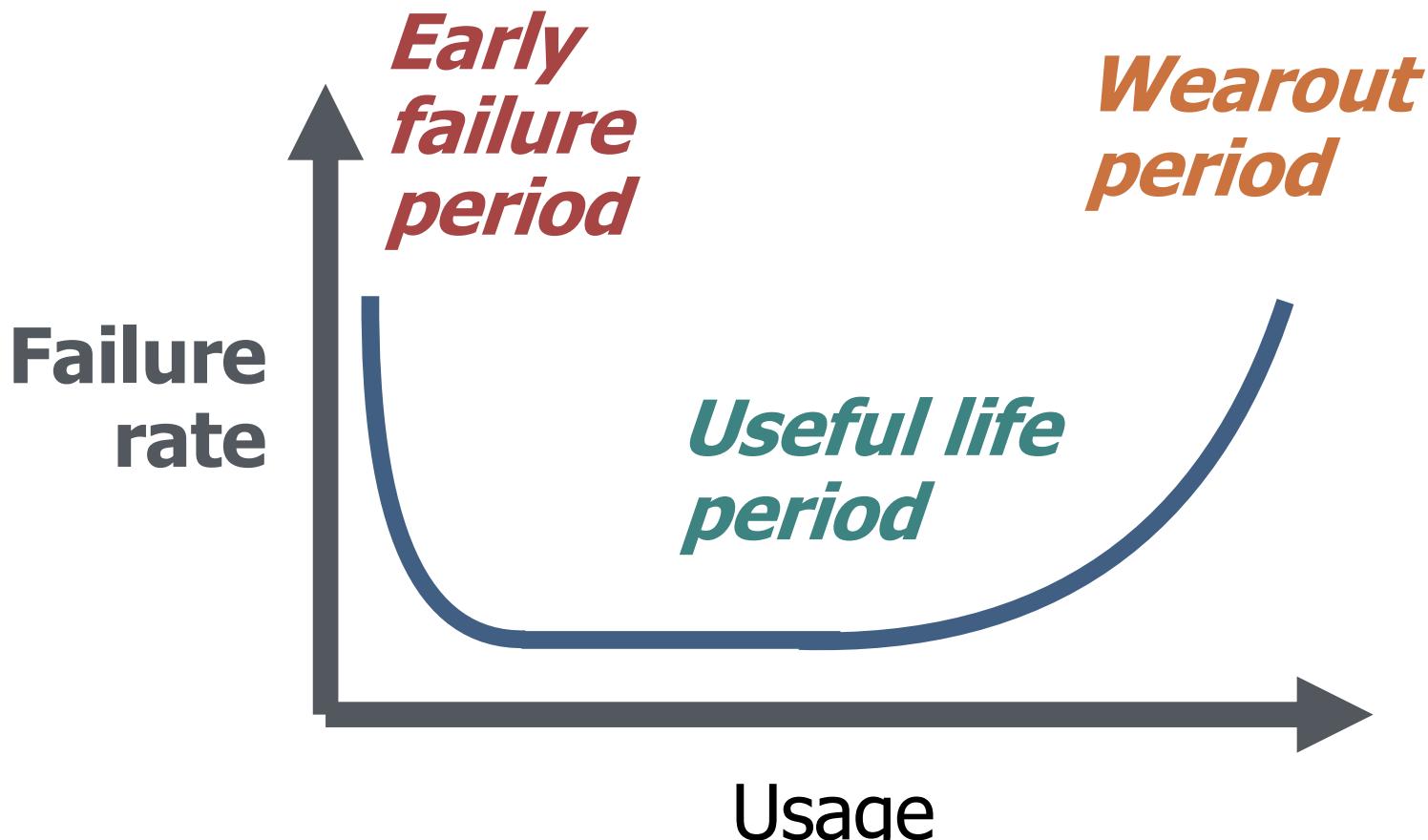
Temperature

Storage lifecycle background: the bathtub curve for disk drives



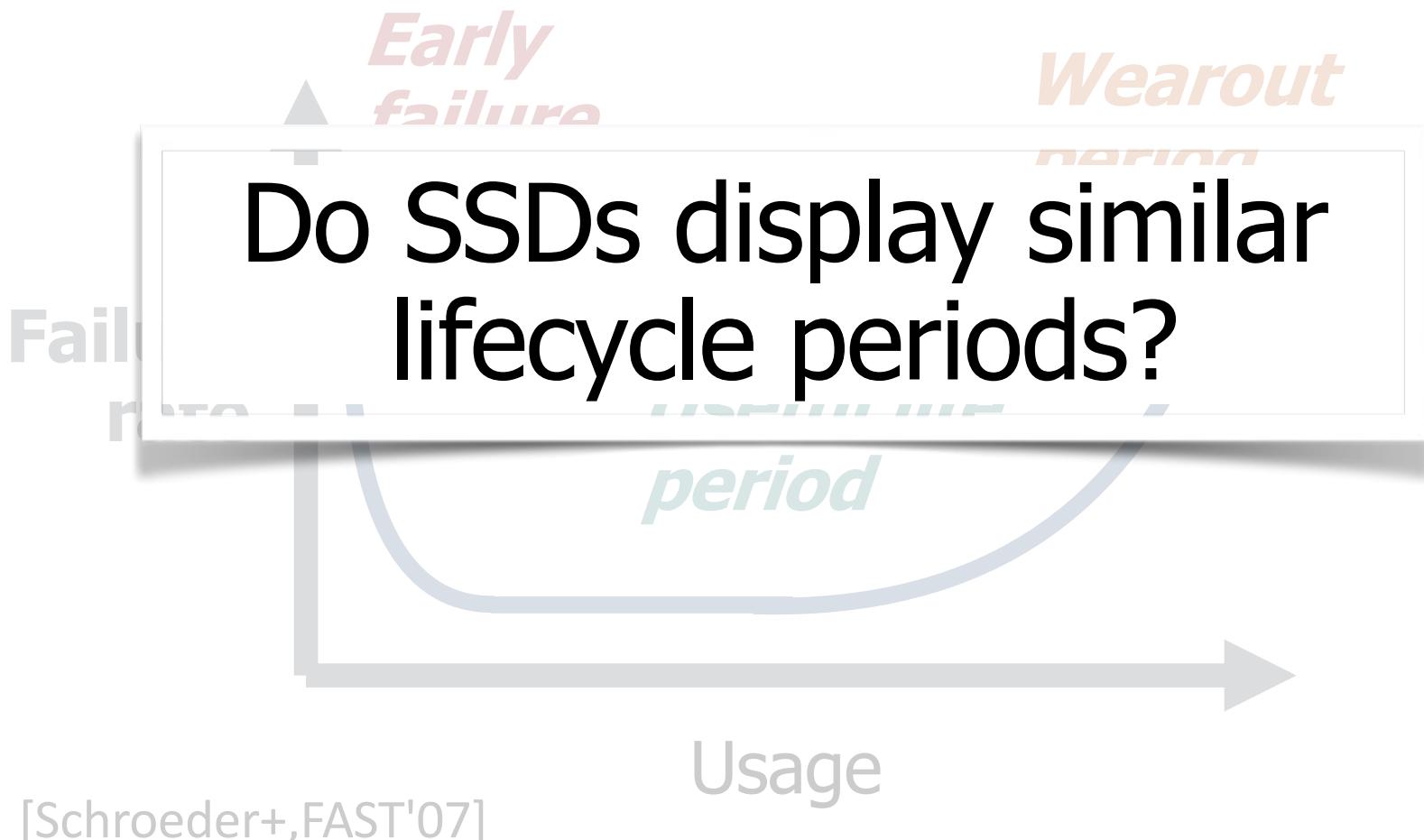
[Schroeder+, FAST'07]

Storage lifecycle background: the bathtub curve for disk drives



[Schroeder+, FAST'07]

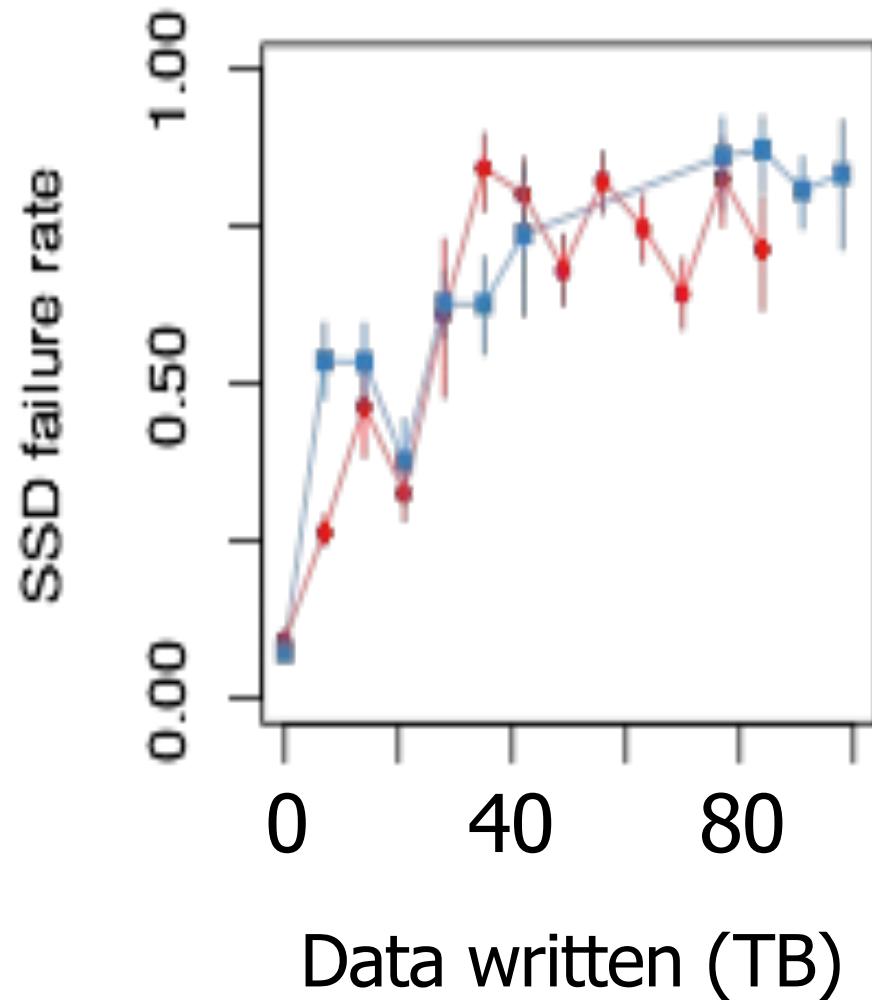
Storage lifecycle background: the bathtub curve for disk drives



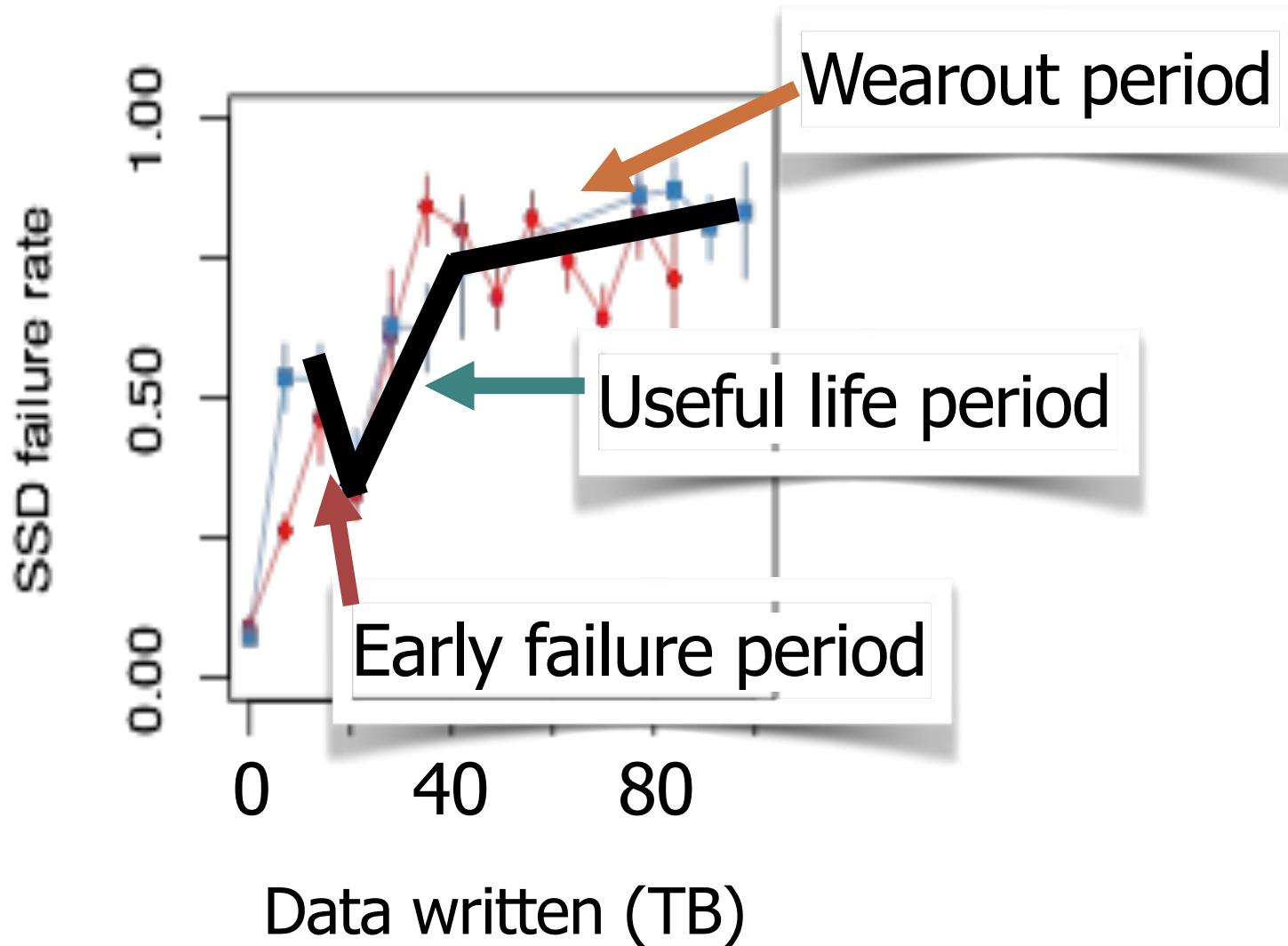
Use data written to flash
to examine SSD lifecycle

(time-independent utilization metric)

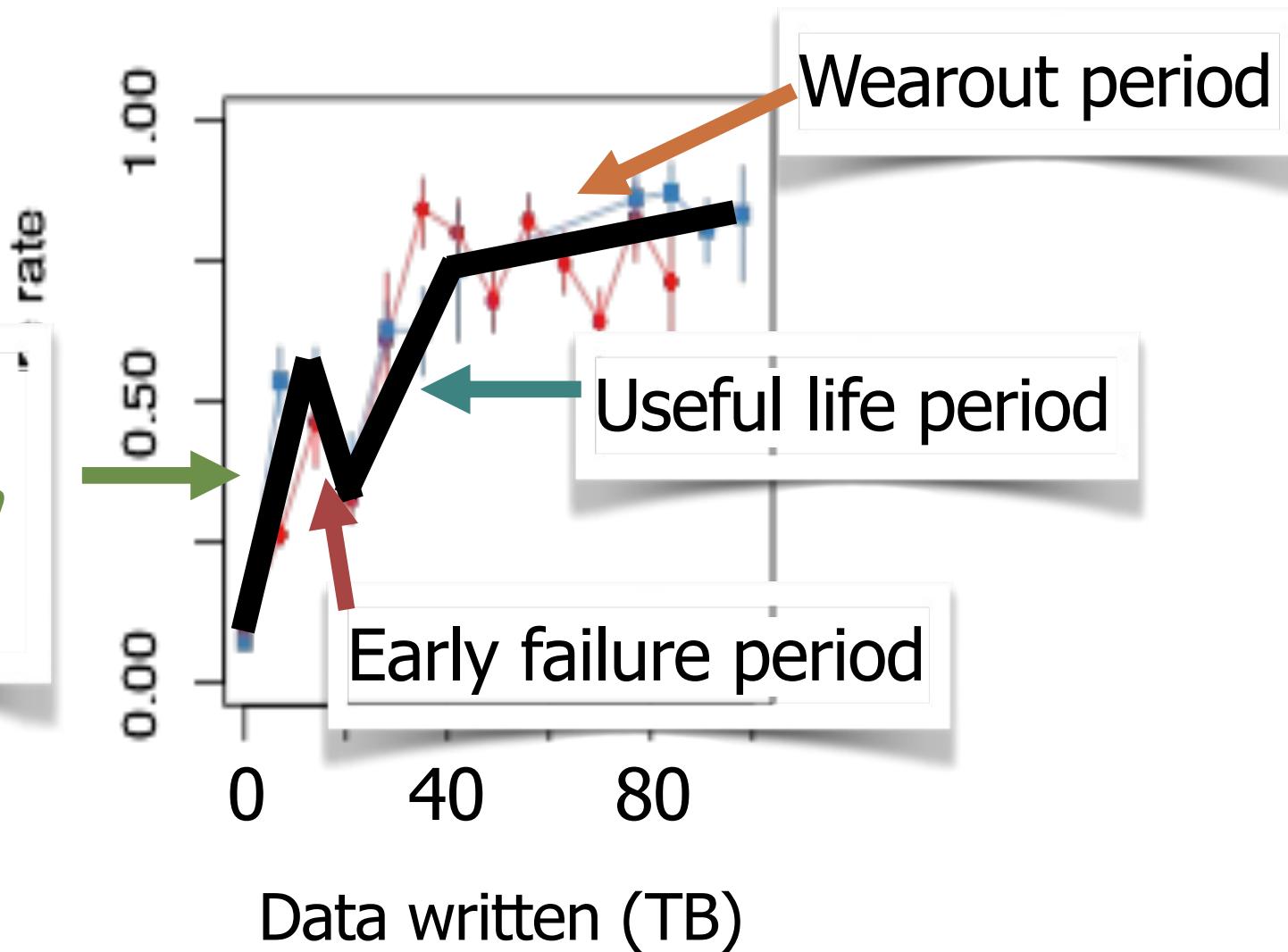
720GB, 1 SSD 720GB, 2 SSDs



720GB, 1 SSD 720GB, 2 SSDs



720GB, 1 SSD 720GB, 2 SSDs



SSD lifecycle



Early detection lifecycle period
distinct from hard disk drive
lifecycle.

Temperature

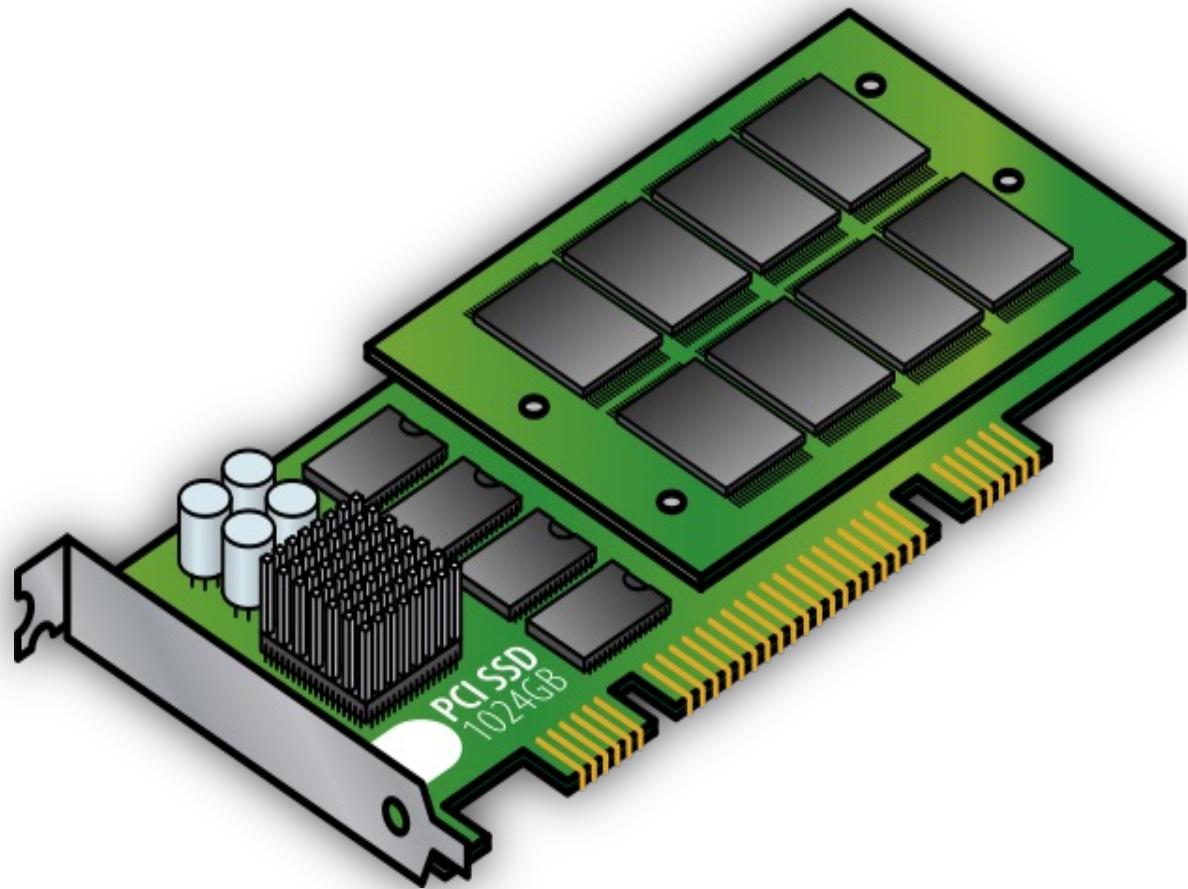
SSD lifecycle

*Access pattern
dependence*

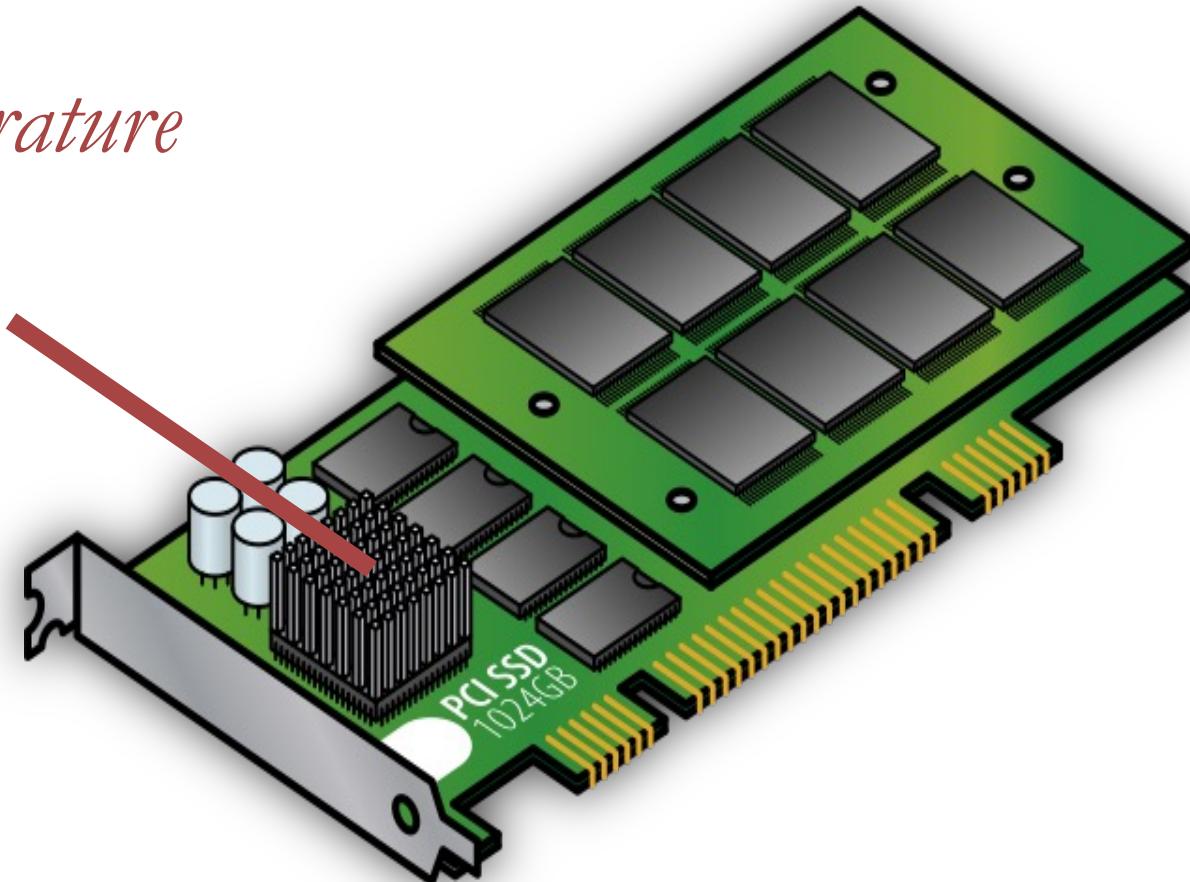


*Read
disturbance*

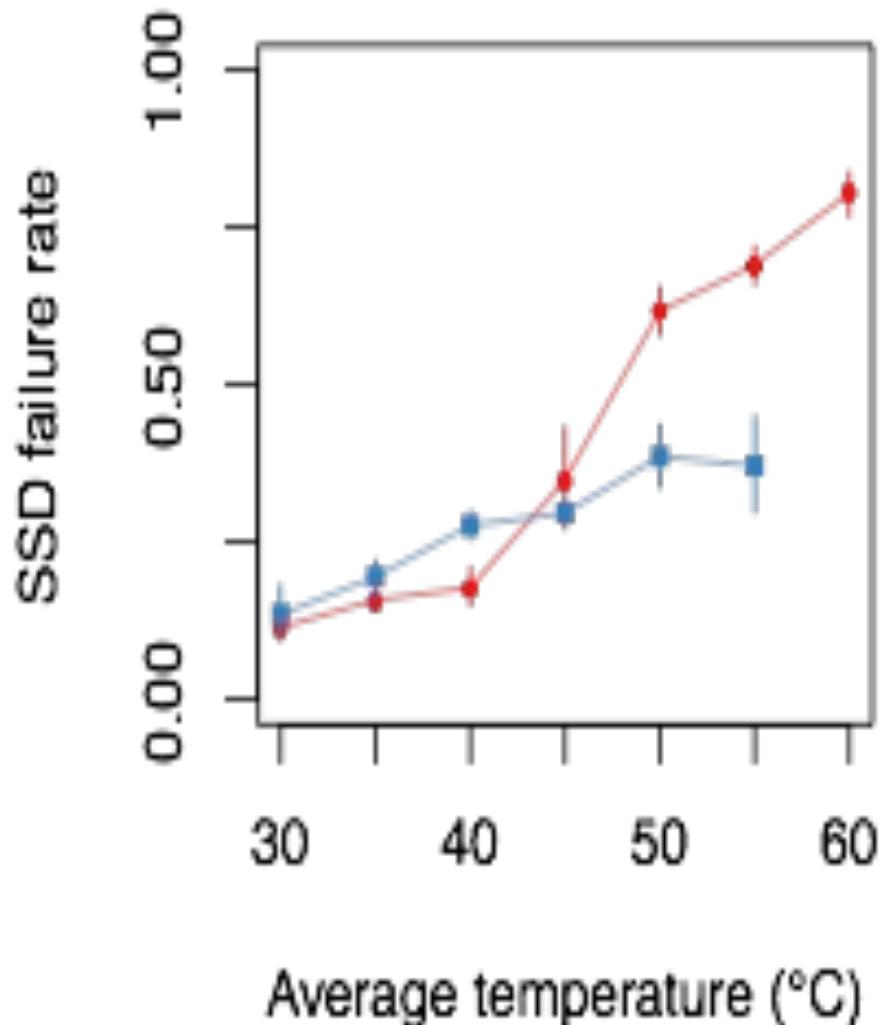
Temperature



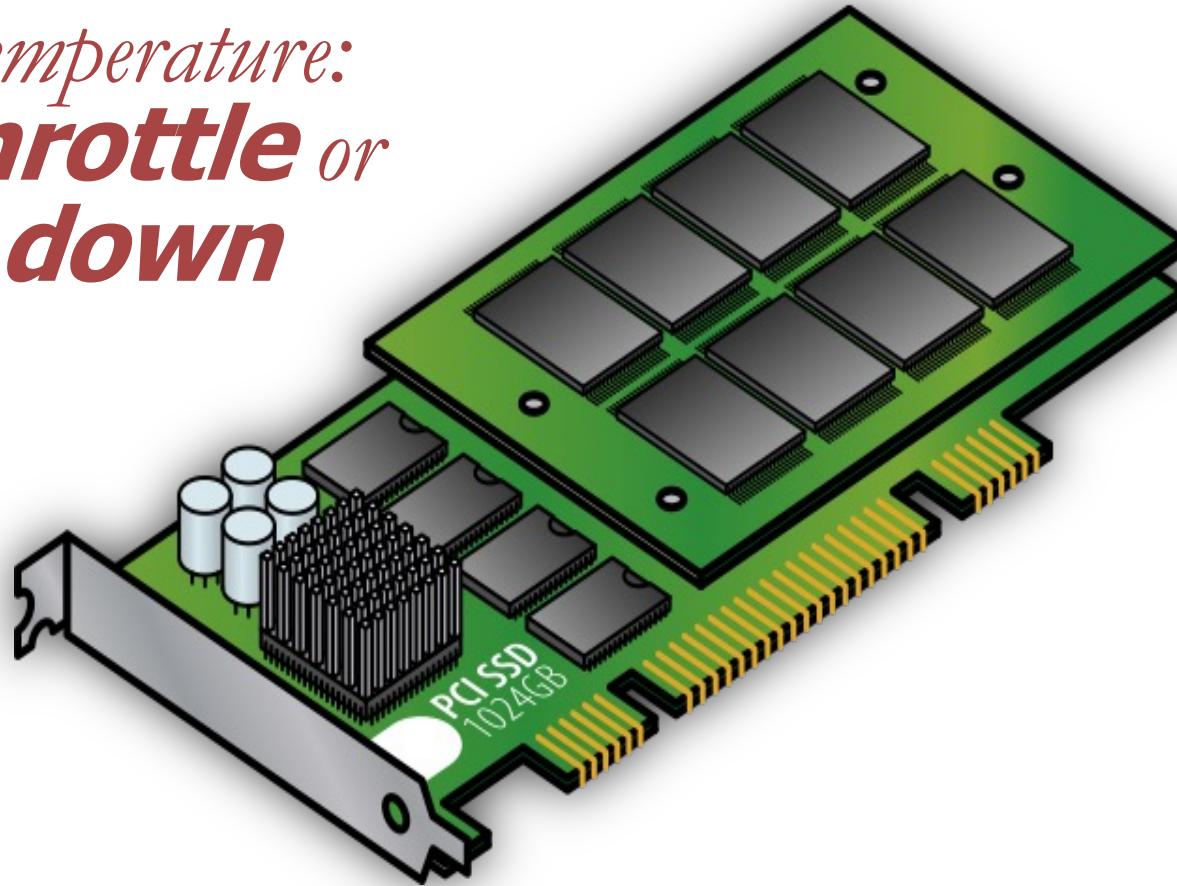
*Temperature
sensor*



720GB, 1 SSD 720GB, 2 SSDs

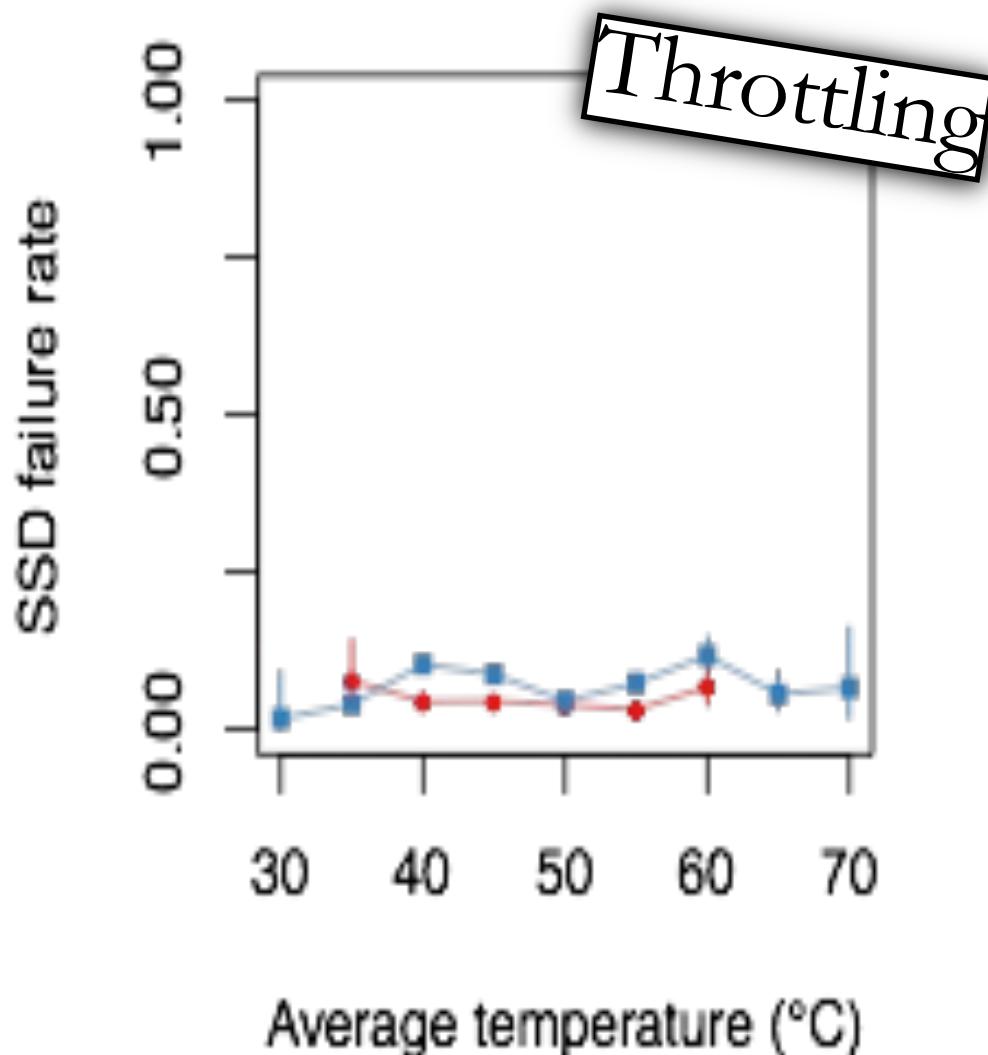


*High temperature:
may **throttle** or
shut down*

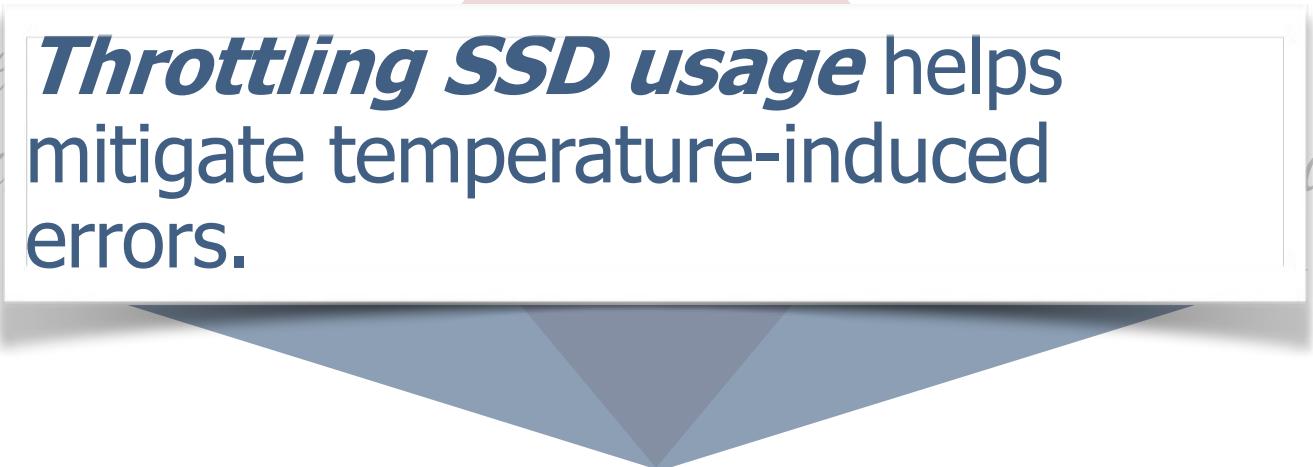


1.2TB, 1 SSD

3.2TB, 1 SSD



SSD lifecycle



Access
and
Throttling SSD usage helps
mitigate temperature-induced
errors.
ce

Temperature

Summary

SSD lifecycle

We ***do not*** observe the effects of ***read disturbance*** errors in the field.

ility

*Read
disturbance*

Temperature

Summary

SSD lifecycle



Access
ce

Throttling SSD usage helps
mitigate temperature-induced
errors.

ce

Temperature

Summary

*Access pattern
dependence*

SSD lifecycle



Temperature

Large-Scale SSD Error Analysis [SIGMETRICS'15]

- First large-scale field study of flash memory errors
- Justin Meza, Qiang Wu, Sanjeev Kumar, and Onur Mutlu,
"A Large-Scale Study of Flash Memory Errors in the Field"

Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS), Portland, OR, June 2015.

[[Slides \(pptx\)](#)] [[Coverage at ZDNet](#)] [[Coverage on The Register](#)]
[[Coverage on TechSpot](#)] [[Coverage on The Tech Report](#)]

A Large-Scale Study of Flash Memory Failures in the Field

Justin Meza
Carnegie Mellon University
meza@cmu.edu

Qiang Wu
Facebook, Inc.
qwu@fb.com

Sanjeev Kumar
Facebook, Inc.
skumar@fb.com

Onur Mutlu
Carnegie Mellon University
onur@cmu.edu

Other Works on NAND Flash Memory Modeling & Issues

Flash Memory Programming Vulnerabilities

- Yu Cai, Saugata Ghose, Yixin Luo, Ken Mai, Onur Mutlu, and Erich F. Haratsch,

"Vulnerabilities in MLC NAND Flash Memory Programming: Experimental Analysis, Exploits, and Mitigation Techniques"

Proceedings of the 23rd International Symposium on High-Performance Computer Architecture (HPCA) Industrial Session, Austin, TX, USA, February 2017.

[[Slides \(pptx\)](#) ([pdf](#))] [[Lightning Session Slides \(pptx\)](#) ([pdf](#))]

Vulnerabilities in MLC NAND Flash Memory Programming: Experimental Analysis, Exploits, and Mitigation Techniques

Yu Cai[†] Saugata Ghose[†] Yixin Luo^{‡†}
[†]*Carnegie Mellon University* [‡]*Seagate Technology*

Ken Mai[†] Onur Mutlu^{§†} Erich F. Haratsch[‡]
[§]*ETH Zürich*

Accurate and Online Channel Modeling

- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu,
**"Enabling Accurate and Practical Online Flash Channel Modeling
for Modern MLC NAND Flash Memory"**
to appear in IEEE Journal on Selected Areas in Communications (JSAC),
2016.

Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory

Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, Onur Mutlu

Agenda

- Background, Motivation and Approach
 - Experimental Characterization Methodology
 - Error Analysis and Management
 - Main Characterization Results
 - Retention-Aware Error Management
 - Threshold Voltage and Program Interference Analysis
 - Read Reference Voltage Prediction
 - Neighbor-Assisted Error Correction
 - Read Disturb Error Handling
 - Retention Error Handling
 - Large Scale Field Analysis
 - 3D NAND Flash Memory Reliability
 - Summary
-

3D NAND Flash Memory

3D NAND Flash Reliability I [HPCA'18]

- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu,
"HeatWatch: Improving 3D NAND Flash Memory Device Reliability by Exploiting Self-Recovery and Temperature-Awareness"

Proceedings of the 24th International Symposium on High-Performance Computer Architecture (HPCA), Vienna, Austria, February 2018.

[Lightning Talk Video]

[Slides (pptx) (pdf)] [Lightning Session Slides (pptx) (pdf)]

HeatWatch: Improving 3D NAND Flash Memory Device Reliability by Exploiting Self-Recovery and Temperature Awareness

Yixin Luo[†] Saugata Ghose[†] Yu Cai[‡] Erich F. Haratsch[‡] Onur Mutlu^{§†}
 [†]*Carnegie Mellon University* [‡]*Seagate Technology* [§]*ETH Zürich*

3D NAND Flash Reliability II [SIGMETRICS'18]

- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu,
"Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation"

Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS), Irvine, CA, USA, June 2018.

[[Abstract](#)]

[[POMACS Journal Version \(same content, different format\)](#)]

[[Slides \(pptx\)](#) ([pdf](#))]

Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation

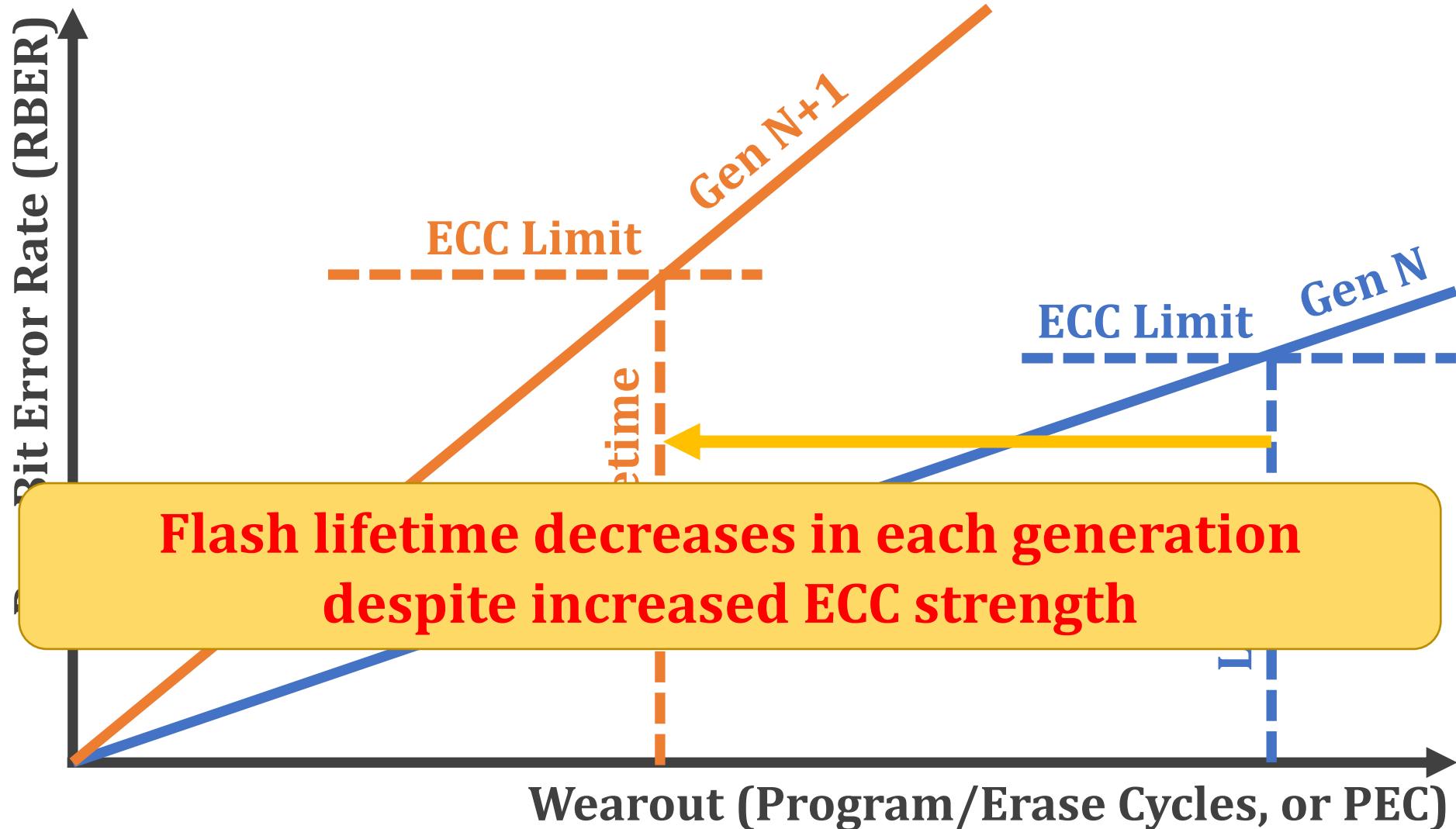
Yixin Luo[†] Saugata Ghose[†] Yu Cai[†] Erich F. Haratsch[‡] Onur Mutlu^{§†}

[†]Carnegie Mellon University

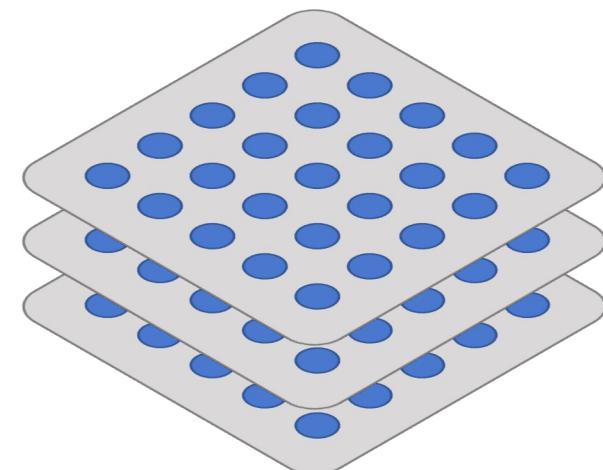
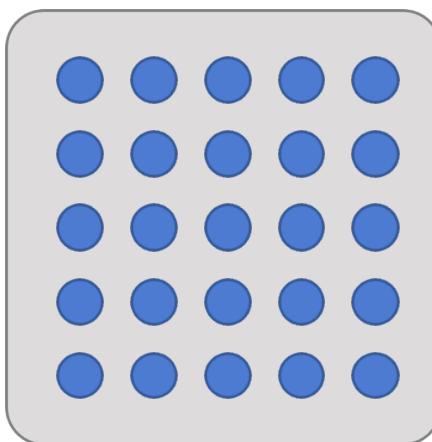
[‡]Seagate Technology

[§]ETH Zürich

NAND Flash Memory Lifetime Problem



Planar vs. 3D NAND Flash Memory



**Planar NAND
Flash Memory**

Scaling

Reduce flash cell size,
Reduce distance b/w cells

**3D NAND
Flash Memory**

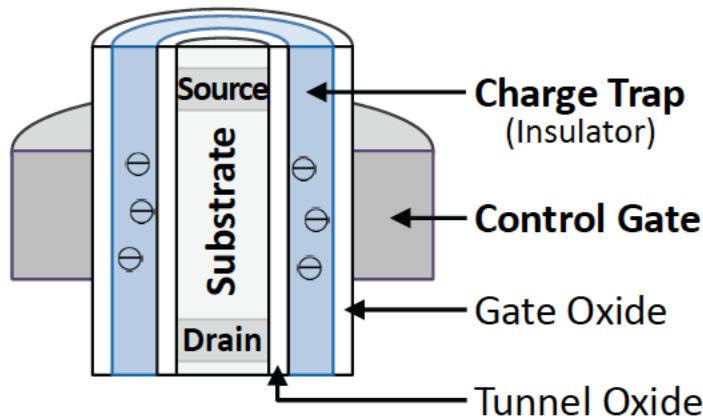
Reliability

Scaling hurts reliability

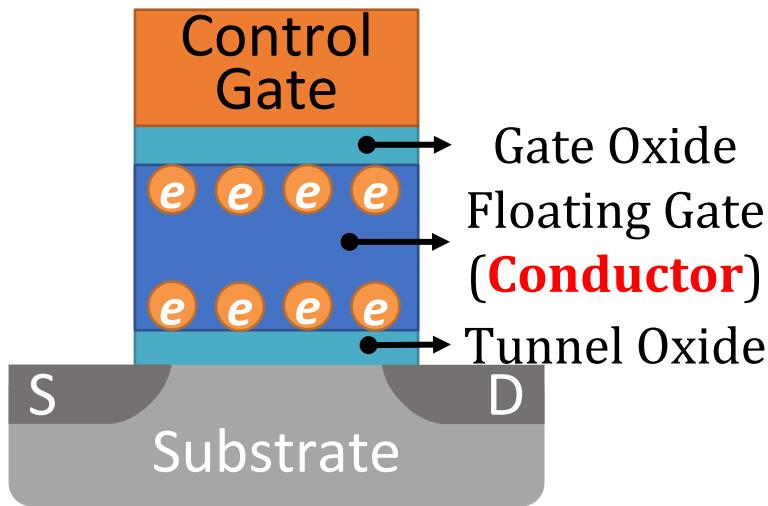
Not well studied!

Charge Trap Based 3D Flash Cell

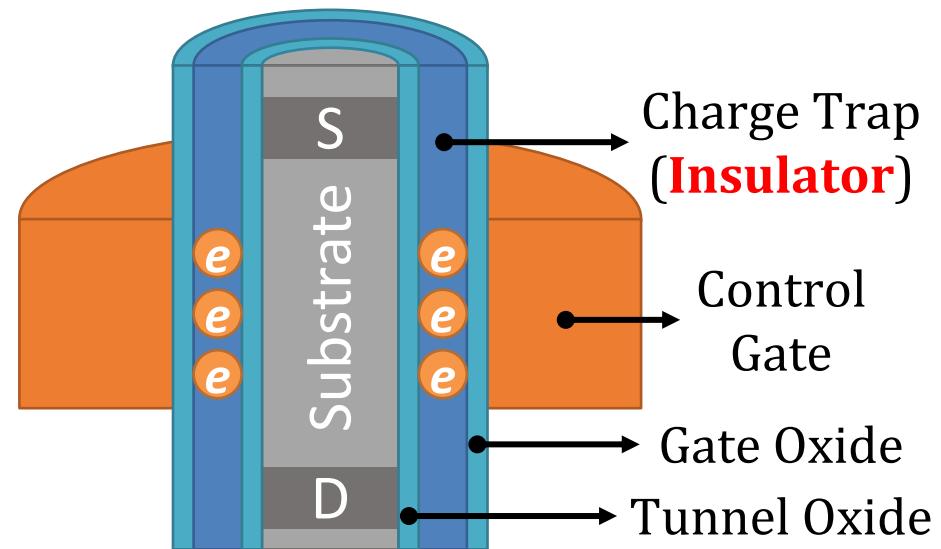
■ Cross-section of a charge trap transistor



2D vs. 3D Flash Cell Design



2D Floating-Gate Cell



3D Charge-Trap Cell

3D NAND Flash Memory Organization

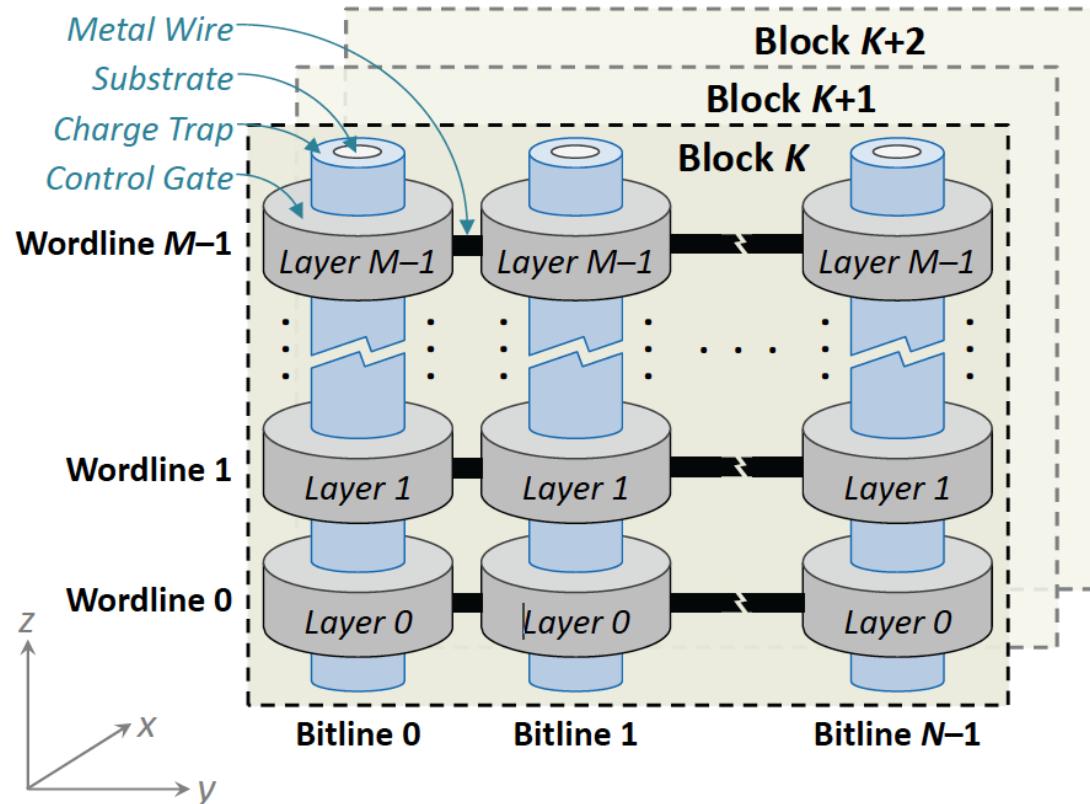


Fig. 43. Organization of flash cells in an M -layer 3D charge trap NAND flash memory chip, where each block consists of M wordlines and N bitlines.

More Background and State-of-the-Art

- Yu Cai, Saugata Ghose, Erich F. Haratsch, Yixin Luo, and Onur Mutlu,
"[Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery](#)"

Invited Book Chapter in [Inside Solid State Drives](#), 2018.

[[Preliminary arxiv.org version](#)]

Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery

YU CAI, SAUGATA GHOSE

Carnegie Mellon University

ERICH F. HARATSCH

Seagate Technology

YIXIN LUO

Carnegie Mellon University

ONUR MUTLU

ETH Zürich and Carnegie Mellon University

3D vs. Planar NAND Errors: Comparison

Table 4. Changes in behavior of different types of errors in 3D NAND flash memory, compared to planar (i.e., two-dimensional) NAND flash memory. See Section 6.2 for a detailed discussion.

Error Type	Change in 3D vs. Planar
P/E Cycling (Section 3.1)	3D is <i>less susceptible</i> , due to current use of charge trap transistors for flash cells
Program (Section 3.2)	3D is <i>less susceptible for now</i> , due to use of one-shot programming (see Section 2.4)
Cell-to-Cell Interference (Section 3.3)	3D is <i>less susceptible for now</i> , due to larger manufacturing process technology
Data Retention (Section 3.4)	3D is <i>more susceptible</i> , due to early retention loss
Read Disturb (Section 3.5)	3D is <i>less susceptible for now</i> , due to larger manufacturing process technology

Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation

Yixin Luo Saugata Ghose Yu Cai Erich F. Haratsch Onur Mutlu

Carnegie Mellon

SAFARI



ETH Zürich



Executive Summary

- Problem: 3D NAND error characteristics are **not well studied**
- Goal: *Understand & mitigate* 3D NAND errors to improve lifetime
- **Contribution 1: Characterize** real 3D NAND flash chips
 - *Process variation:* $21\times$ error rate difference across layers
 - *Early retention loss:* Error rate increases by $10\times$ after 3 hours
 - *Retention interference:* Not observed before in planar NAND
- **Contribution 2: Model** RBER and threshold voltage
 - *RBER (raw bit error rate) variation model*
 - *Retention loss model*
- **Contribution 3: Mitigate** 3D NAND flash errors
 - *LaVAR: Layer Variation Aware Reading*
 - *LI-RAID: Layer-Interleaved RAID*
 - *ReMAR: Retention Model Aware Reading*
 - *Improve flash lifetime by $1.85\times$ or reduce ECC overhead by 78.9%*

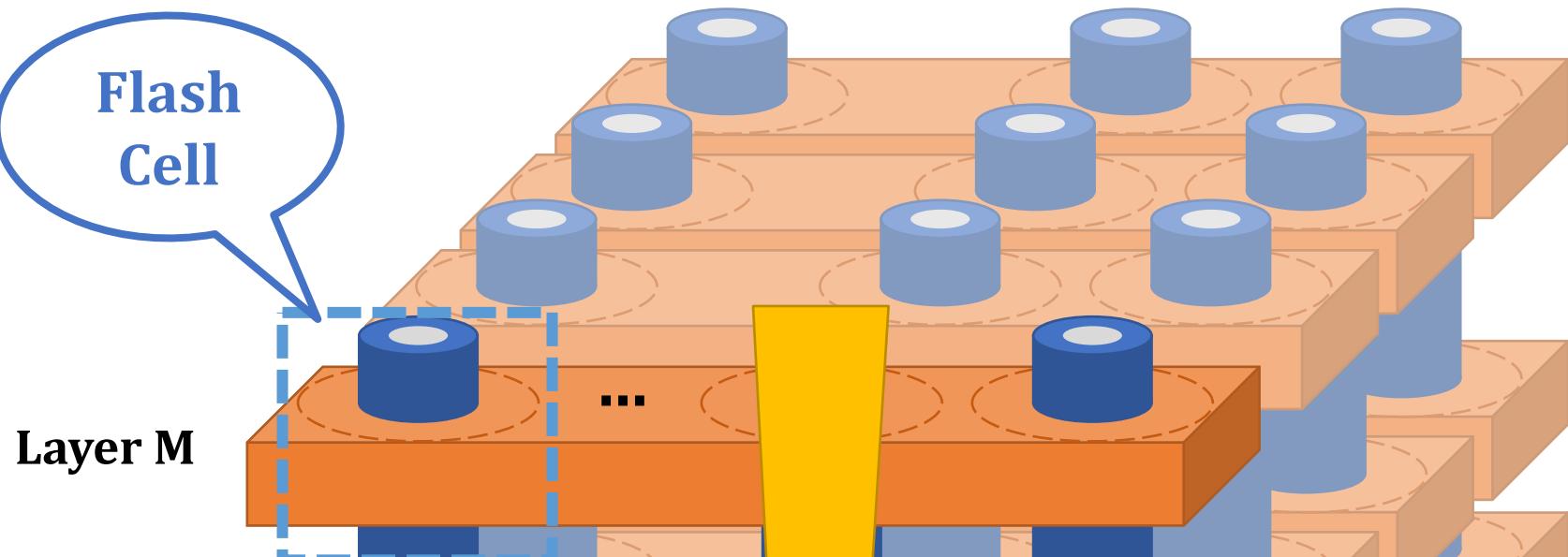
Agenda

- Background & Introduction
- Contribution 1: Characterize real 3D NAND flash chips
- Contribution 2: Model RBER and threshold voltage
- Contribution 3: Mitigate 3D NAND flash errors
- Conclusion

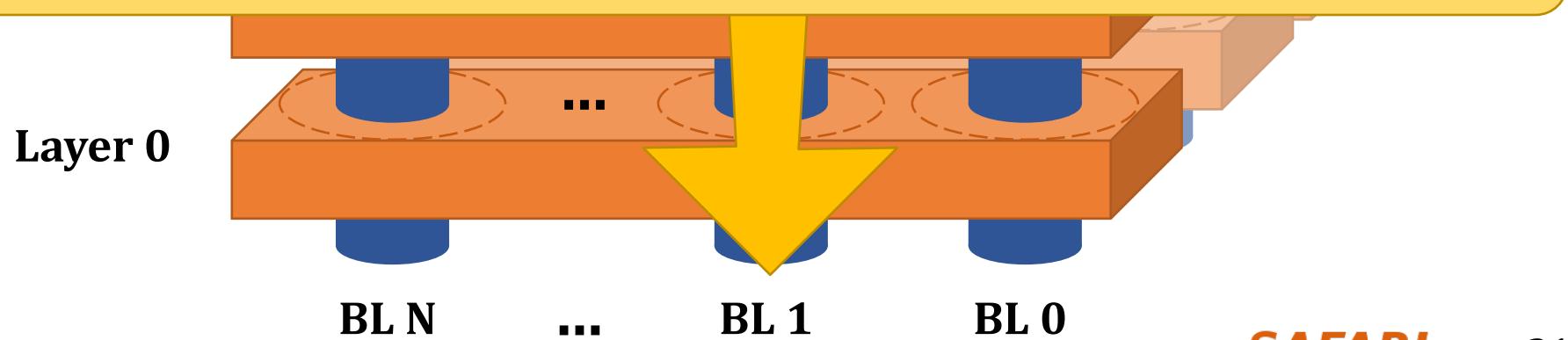
Agenda

- Background & Introduction
- Contribution 1: Characterize real 3D NAND flash chips
 - Process variation
 - Early retention loss
 - Retention interference
- Contribution 2: Model RBER and threshold voltage
- Contribution 3: Mitigate 3D NAND flash errors
- Conclusion

Process Variation Across Layers



Flash cells on different layers may have different error characteristics

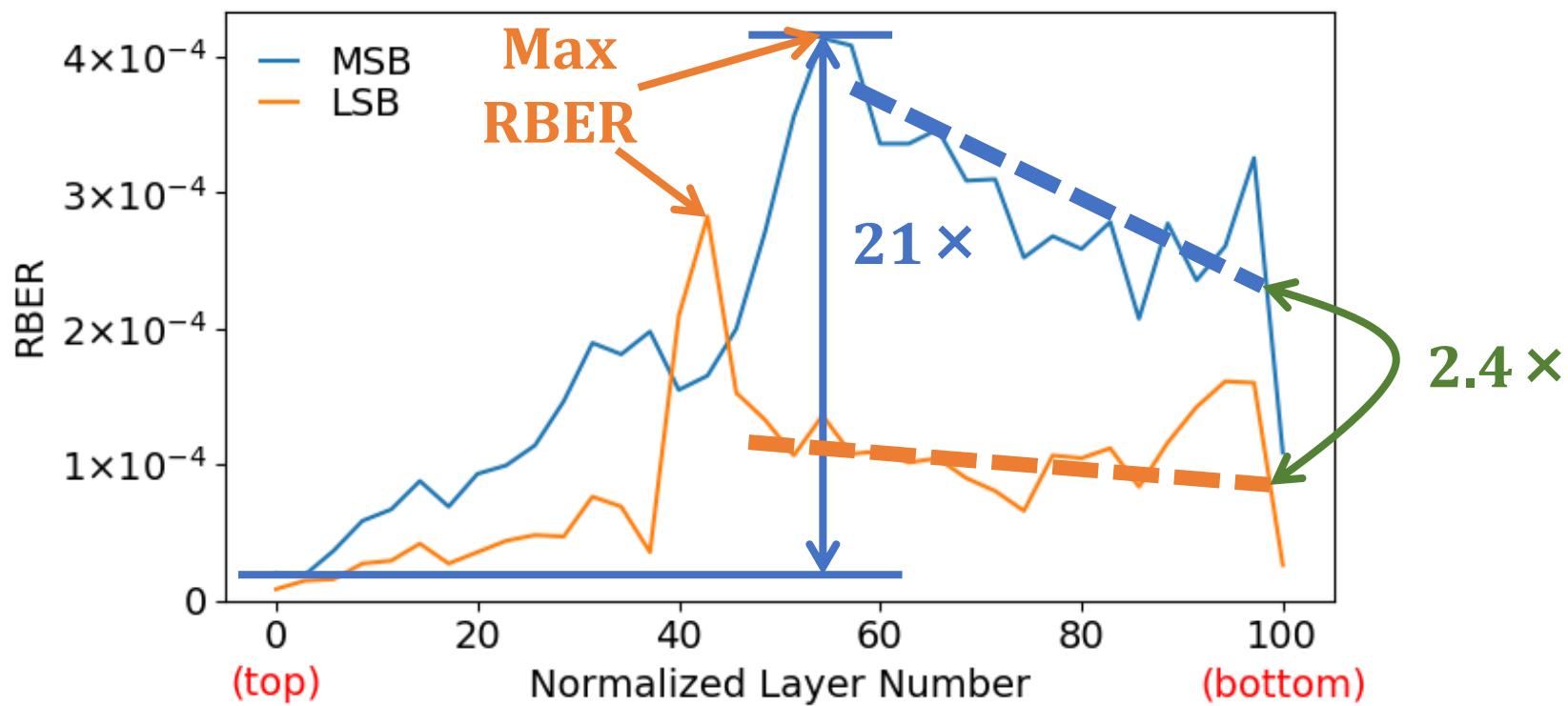


Characterization Methodology

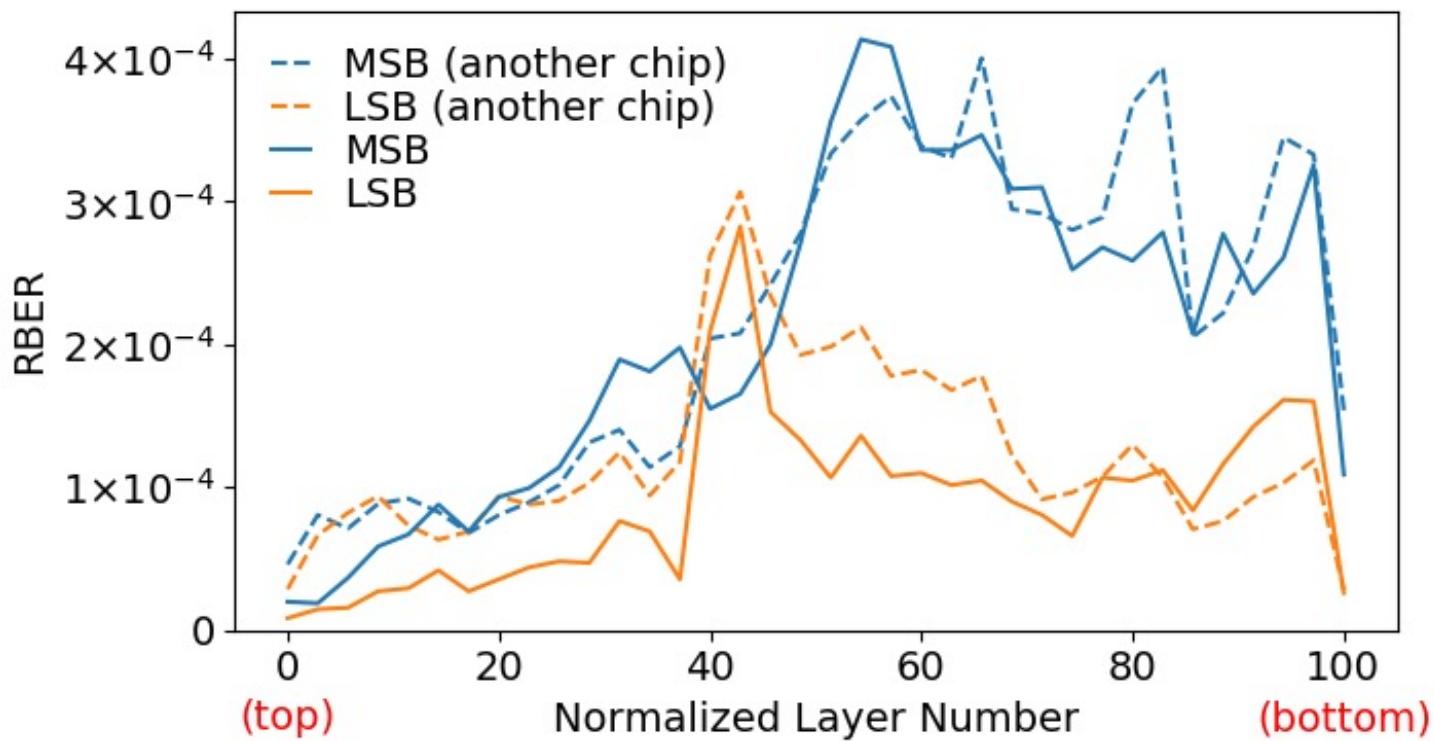
- Modified firmware version in the flash controller
 - Controls the read reference voltage of the flash chip
 - Bypasses ECC to get raw data (with raw bit errors)
- Analysis and post-processing of the data on the server



Layer-to-Layer Process Variation



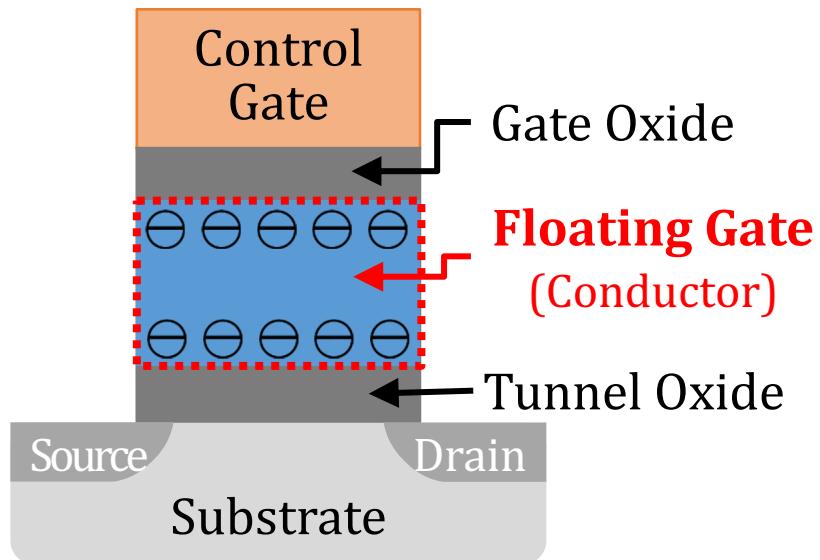
Layer-to-Layer Process Variation



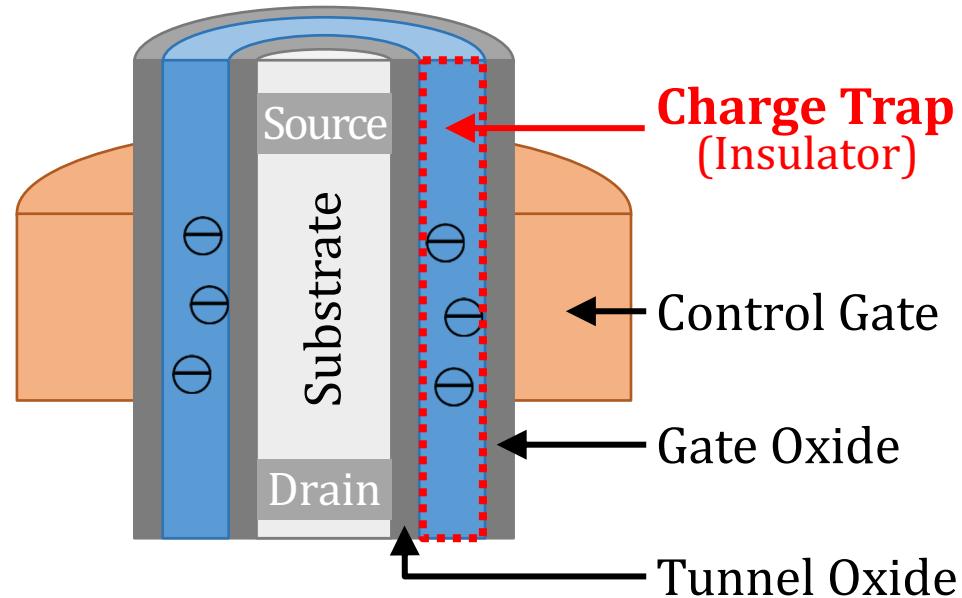
**Large RBER variation
across layers and LSB-MSB pages**

Retention Loss Phenomenon

Planar NAND Cell

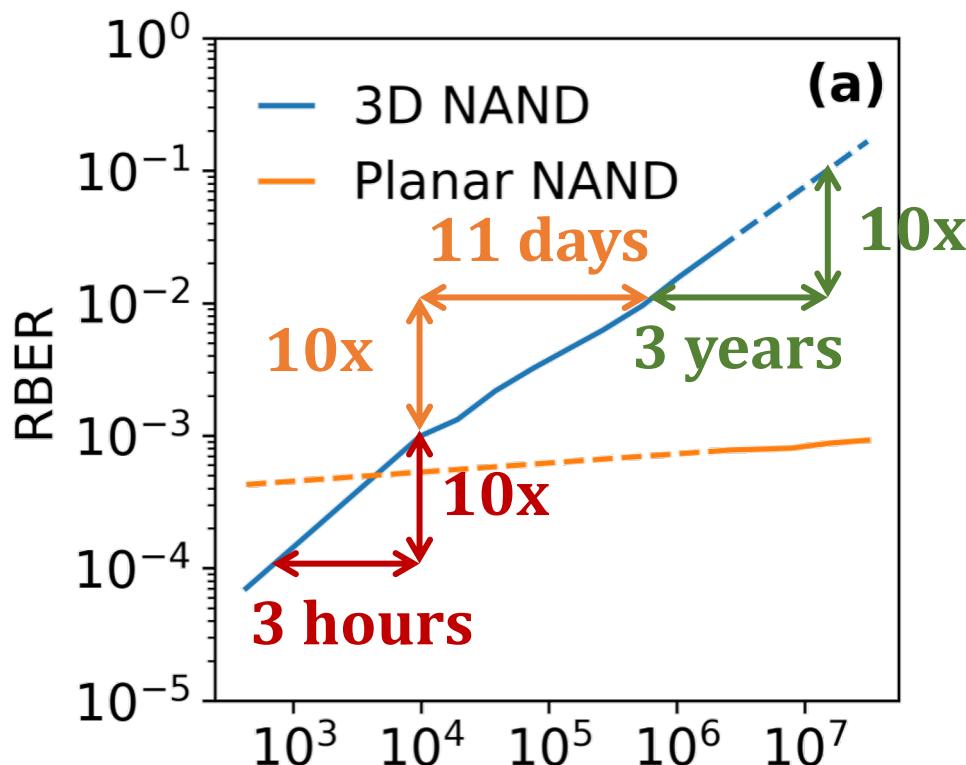


3D NAND Cell



Most dominant type of error in planar NAND.
Is this true for 3D NAND as well?

Early Retention Loss



Retention errors increase quickly immediately after programming

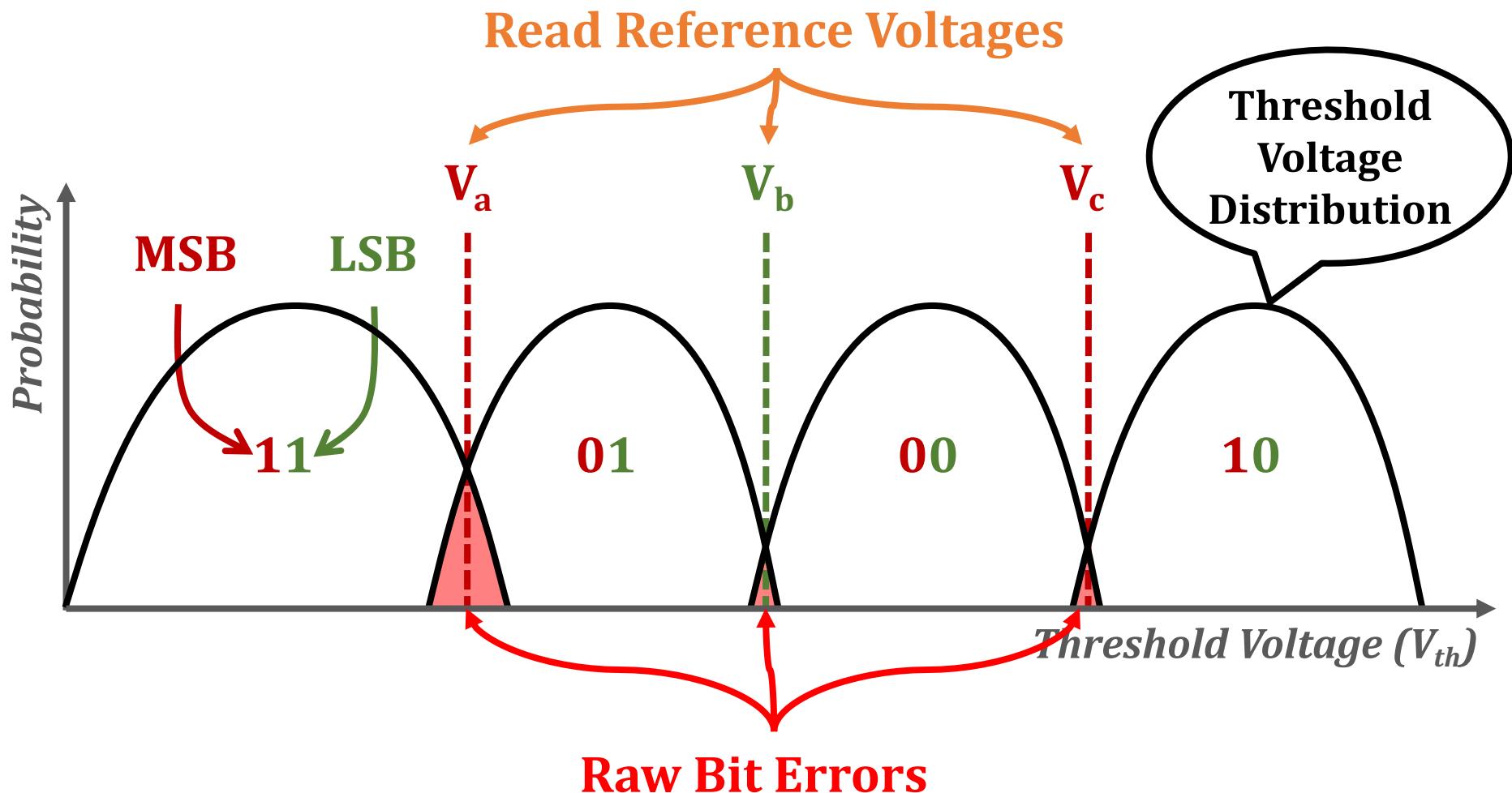
Characterization Summary

- **Layer-to-layer process variation**
 - Large RBER variation across layers and LSB-MSB pages
 - → Need new mechanisms to tolerate RBER variation!
- **Early retention loss**
 - RBER increases quickly after programming
 - → Need new mechanisms to tolerate retention errors!
- **Retention interference**
 - Amount of retention loss correlated with neighbor cells' states
 - → Need new mechanisms to tolerate retention interference!
- **More *threshold voltage* and *RBER* results in the paper:**
3D NAND P/E cycling, program interference, read disturb, read variation, bitline-to-bitline process variation
- **Our approach** based on insights developed via our experimental characterization: Develop **error models**, and build online **error mitigation mechanisms** using the models

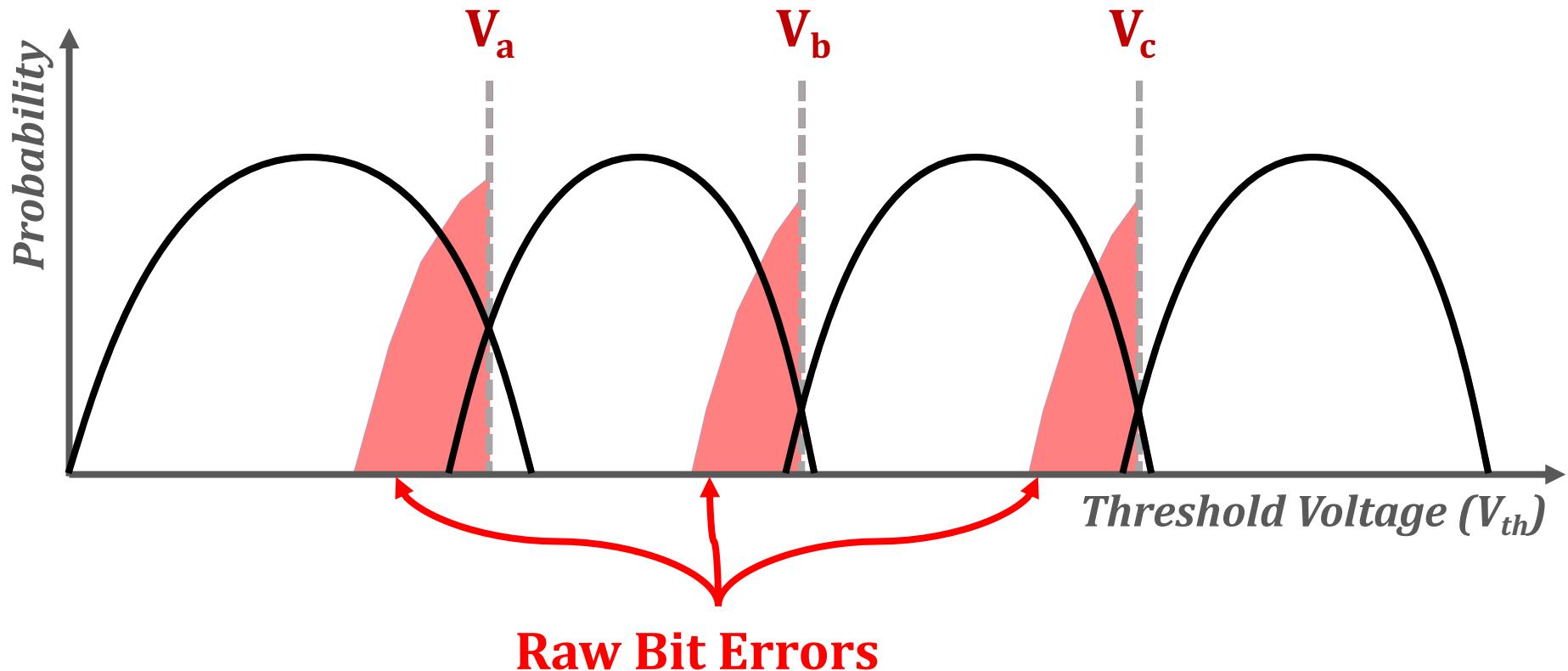
Agenda

- Background & Introduction
- Contribution 1: Characterize real 3D NAND flash chips
- Contribution 2: Model RBER and threshold voltage
 - Retention loss model
 - RBER variation model
- Contribution 3: Mitigate 3D NAND flash errors
- Conclusion

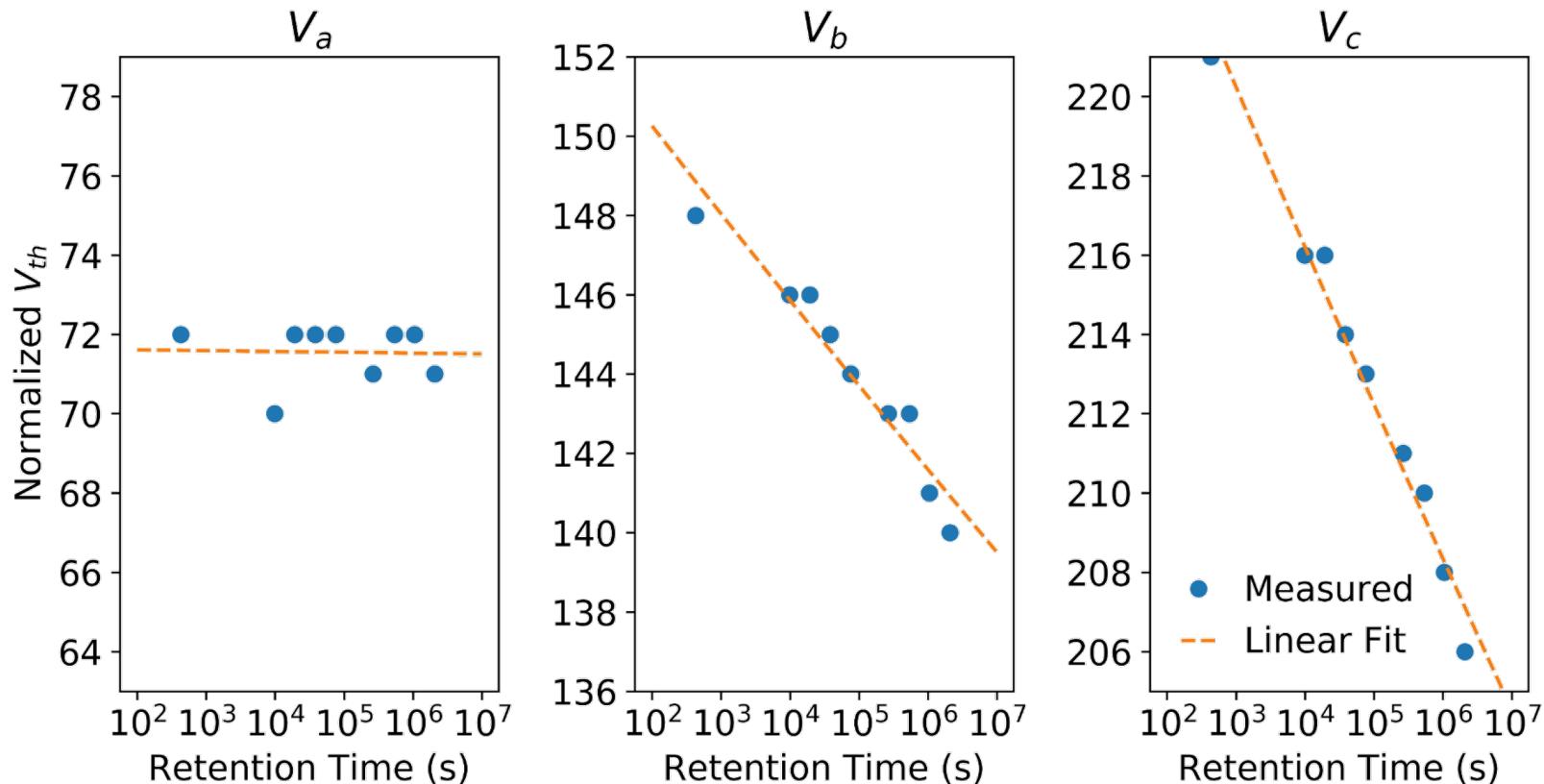
What Do We Model?



Optimal Read Reference Voltage



Retention Loss Model

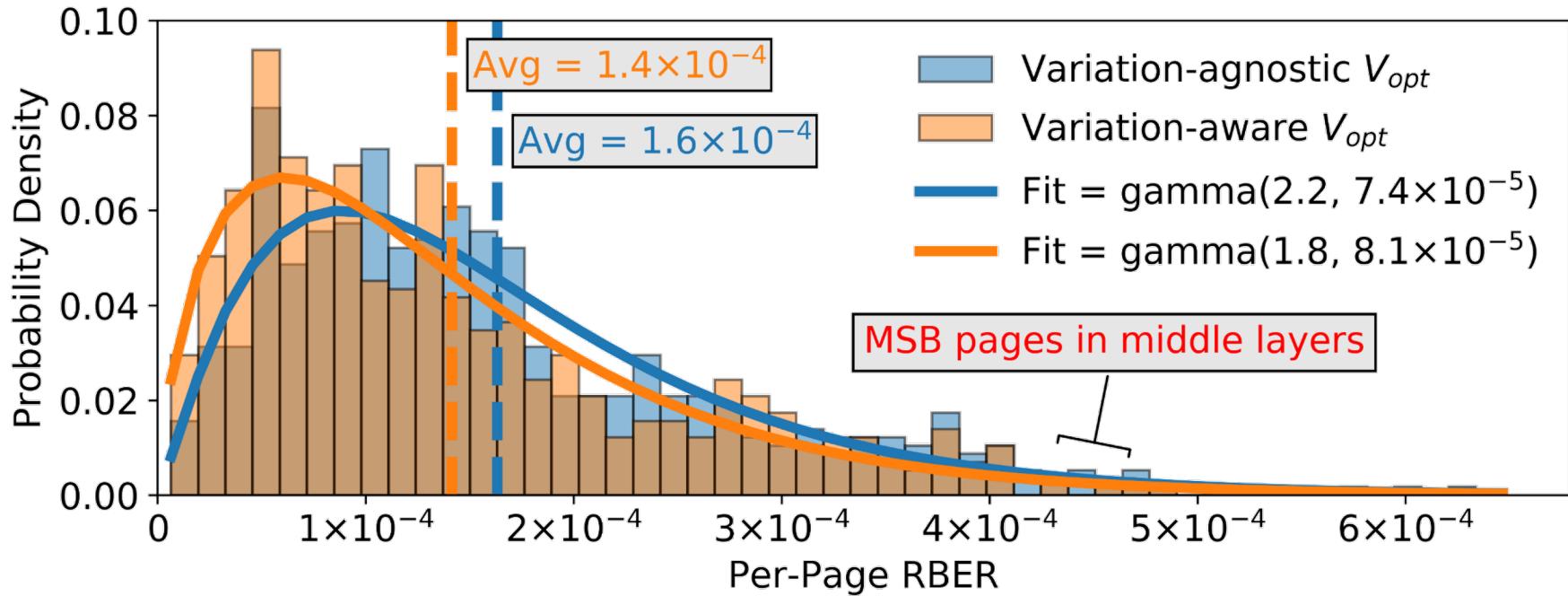


Early retention loss can be modeled as a simple linear function of $\log(\text{retention time})$

Retention Loss Model

- Goal: Develop a simple linear model that can be used online
- Models
 - Optimal read reference voltage (V_b and V_c)
 - Raw bit error rate ($\log(RBER)$)
 - Mean and standard deviation of threshold voltage distribution (μ and σ)
- As a function of
 - Retention time ($\log(t)$)
 - P/E cycle count (PEC)
- e.g., $V_{opt} = (\alpha \times PEC + \beta) \times \log(t) + \gamma \times PEC + \delta$
- Model error <1 step for V_b and V_c
- Adjusted R² > 89%

RBER Variation Model



Variation-agnostic V_{opt}

- Same V_{ref} for all layers optimized for the entire block

RBER distribution follows gamma distribution

KL-divergence error = 0.09

Agenda

- Background & Introduction
- Contribution 1: Characterize real 3D NAND flash chips
- Contribution 2: Model RBER and threshold voltage
- Contribution 3: Mitigate 3D NAND flash errors
 - LaVAR: Layer Variation Aware Reading
 - LI-RAID: Layer-Interleaved RAID
 - ReMAR: Retention Model Aware Reading
- Conclusion

LaVAR: Layer Variation Aware Reading

- **Layer-to-layer process variation**
 - Error characteristics are different in each layer
- **Goal:** Adjust read reference voltage **for each layer**
- **Key Idea:** Learn a **voltage offset (Offset)** for each layer
 - $V_{opt}^{Layer\ aware} = V_{opt}^{Layer\ agnostic} + Offset$
- **Mechanism**
 - **Offset:** Learned once for each chip & stored in a table
 - *Uses (2×Layers) Bytes memory per chip*
 - $V_{opt}^{Layer\ agnostic}$: Predicted by any existing V_{opt} model
 - *E.g., ReMAR [Luo+Sigmetrics'18], HeatWatch [Luo+HPCA'18], OFCM [Luo+JSAC'16], ARVT [Papandreou+GLSVLSI'14]*
- Reduces RBER on average by **43%**
(based on our characterization data)

LI-RAID: Layer-Interleaved RAID

- **Layer-to-layer process variation**
 - Worst-case RBER much higher than average RBER
- **Goal:** Significantly reduce worst-case RBER
- **Key Idea**
 - Group flash pages on *less reliable layers* with pages on *more reliable layers*
 - Group *MSB pages* with *LSB pages*
- **Mechanism**
 - Reorganize RAID layout to eliminate worst-case RBER
 - <0.8% storage overhead

Conventional RAID

<i>Wordline #</i>	<i>Layer #</i>	<i>Page</i>	Chip 0	Chip 1	Chip 2	Chip 3
<i>0</i>	<i>0</i>	<i>MSB</i>	Group 0	Group 0	Group 0	Group 0
<i>0</i>	<i>0</i>	<i>LSB</i>	Group 1	Group 1	Group 1	Group 1
<i>1</i>	<i>1</i>	<i>MSB</i>	Group 2	Group 2	Group 2	Group 2
<i>1</i>	<i>1</i>	<i>LSB</i>	Group 3	Group 3	Group 3	Group 3
<i>2</i>	<i>2</i>	<i>MSB</i>	Group 4	Group 4	Group 4	Group 4
<i>2</i>	<i>2</i>	<i>LSB</i>	Group 5	Group 5	Group 5	Group 5
<i>3</i>	<i>3</i>	<i>MSB</i>	Group 6	Group 6	Group 6	Group 6
<i>3</i>	<i>3</i>	<i>LSB</i>	Group 7	Group 7	Group 7	Group 7

**Worst-case RBER in any layer
limits the lifetime of conventional RAID**

LI-RAID: Layer-Interleaved RAID

<i>Wordline #</i>	<i>Layer #</i>	<i>Page</i>	Chip 0	Chip 1	Chip 2	Chip 3
0	0	MSB	Group 0	Blank	Group 4	Group 3
0	0	LSB	Group 1	Blank	Group 5	Group 2
1	1	MSB	Group 2	Group 1	Blank	Group 5
1	1	LSB	Group 3	Group 0	Blank	Group 4
2	2	MSB	Group 4	Group 3	Group 0	Blank
2	2	LSB	Group 5	Group 2	Group 1	Blank
3	3	MSB	Blank	Group 5	Group 2	Group 1
3	3	LSB	Blank	Group 4	Group 3	Group 0

Any page with worst-case RBER can be corrected by other reliable pages in the RAID group

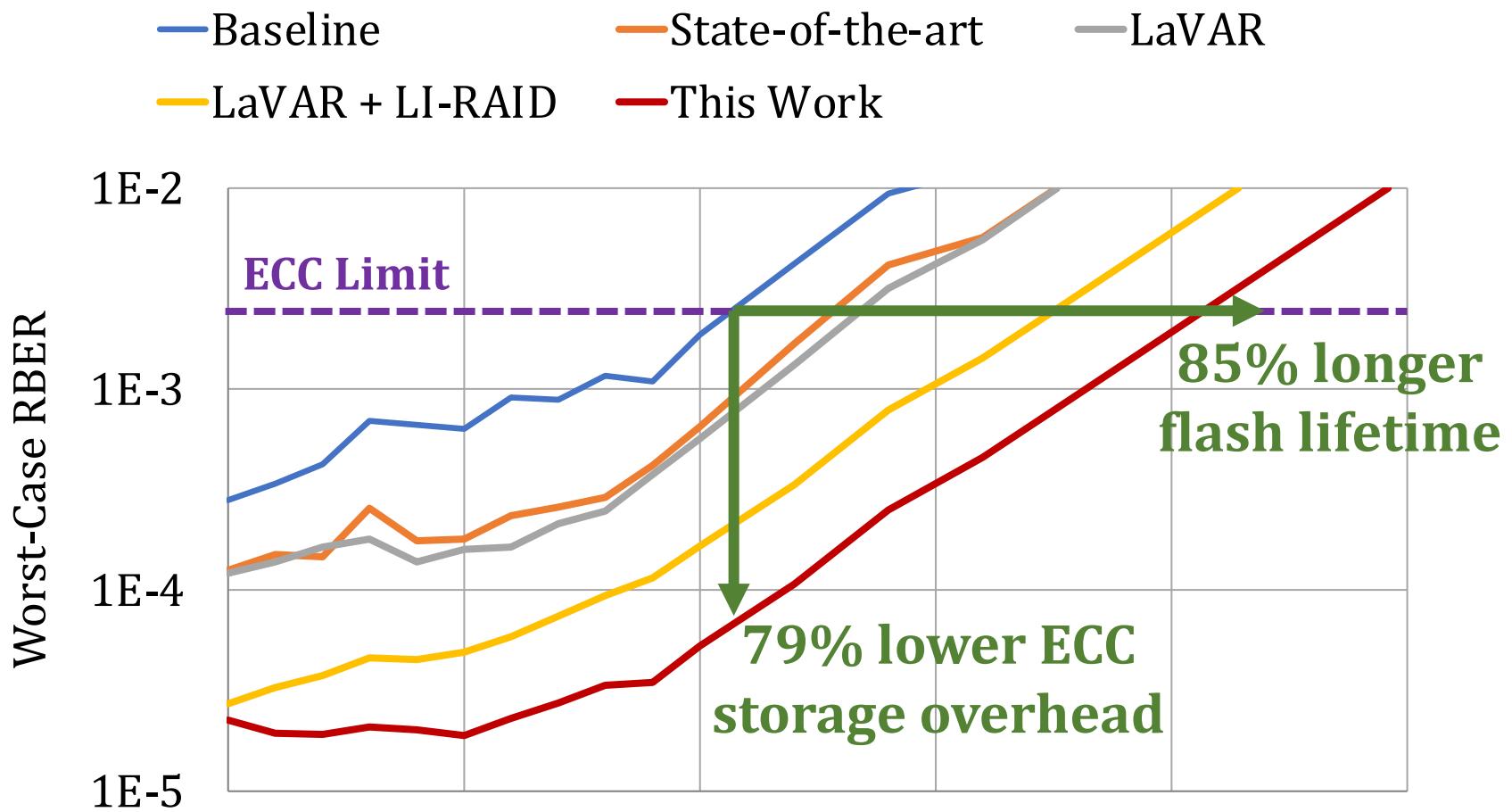
LI-RAID: Layer-Interleaved RAID

- **Layer-to-layer process variation**
 - Worst-case RBER much higher than average RBER
- **Goal:** Significantly reduce worst-case RBER
- **Key Idea**
 - Group flash pages on *less reliable layers* with pages on *more reliable layers*
 - Group *MSB pages* with *LSB pages*
- **Mechanism**
 - Reorganize RAID layout to eliminate worst-case RBER
 - <0.8% storage overhead
- Reduces worst-case RBER by **66.9%**
(based on our characterization data)

ReMAR: Retention Model Aware Reading

- **Early retention loss**
 - Threshold voltage shifts quickly after programming
- **Goal: Adjust read reference voltages based on retention loss**
- **Key Idea:** Learn and use a retention loss model online
- **Mechanism**
 - Periodically characterize and learn retention loss model online
 - Retention time = Read timestamp - Write timestamp
 - *Uses 800 KB memory to store program time of each block*
 - Predict retention-aware V_{opt} using the model
- Reduces RBER on average by **51.9%**
(based on our characterization data)

Impact on System Reliability



LaVAR, LI-RAID, and ReMAR improve flash lifetime
or reduce ECC overhead significantly

Error Mitigation Techniques Summary

- **LaVAR: Layer Variation Aware Reading**
 - Learn a V_{opt} offset for each layer and apply *layer-aware V_{opt}*
- **LI-RAID: Layer-Interleaved RAID**
 - Group flash pages on *less reliable layers* with pages on *more reliable layers*
 - Group *MSB pages* with *LSB pages*
- **ReMAR: Retention Model Aware Reading**
 - Learn retention loss model and apply *retention-aware V_{opt}*
- **Benefits:**
 - Improve flash lifetime by **1.85×** or reduce ECC overhead by **78.9%**
- **ReNAC (in paper):** Reread a failed page using V_{opt} based on the *retention interference* induced by neighbor cell

Agenda

- Background & Introduction
- Contribution 1: Characterize real 3D NAND flash chips
- Contribution 2: Model RBER and threshold voltage
- Contribution 3: Mitigate 3D NAND flash errors
- Conclusion

Conclusion

- Problem: 3D NAND error characteristics are **not well studied**
- Goal: *Understand & mitigate* 3D NAND errors to improve lifetime
- **Contribution 1: Characterize** real 3D NAND flash chips
 - *Process variation:* $21\times$ error rate difference across layers
 - *Early retention loss:* Error rate increases by $10\times$ after 3 hours
 - *Retention interference:* Not observed before in planar NAND
- **Contribution 2: Model** RBER and threshold voltage
 - *RBER (raw bit error rate) variation model*
 - *Retention loss model*
- **Contribution 3: Mitigate** 3D NAND flash errors
 - *LaVAR: Layer Variation Aware Reading*
 - *LI-RAID: Layer-Interleaved RAID*
 - *ReMAR: Retention Model Aware Reading*
 - *Improve flash lifetime by $1.85\times$ or reduce ECC overhead by 78.9%*

Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation

Yixin Luo Saugata Ghose Yu Cai Erich F. Haratsch Onur Mutlu

Carnegie Mellon

SAFARI



ETH Zürich



3D NAND Flash Reliability II [SIGMETRICS'18]

- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu,
"Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation"

Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS), Irvine, CA, USA, June 2018.

[[Abstract](#)]

[[POMACS Journal Version \(same content, different format\)](#)]

[[Slides \(pptx\)](#) ([pdf](#))]

Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation

Yixin Luo[†] Saugata Ghose[†] Yu Cai[†] Erich F. Haratsch[‡] Onur Mutlu^{§†}

[†]Carnegie Mellon University

[‡]Seagate Technology

[§]ETH Zürich

One More Idea

WARM

Improving NAND Flash Memory Lifetime with Write-hotness Aware Retention Management

Yixin Luo, Yu Cai, Saugata Ghose, Jongmoo Choi, Onur Mutlu*

*Carnegie Mellon University, *Dankook University*

SAFARI

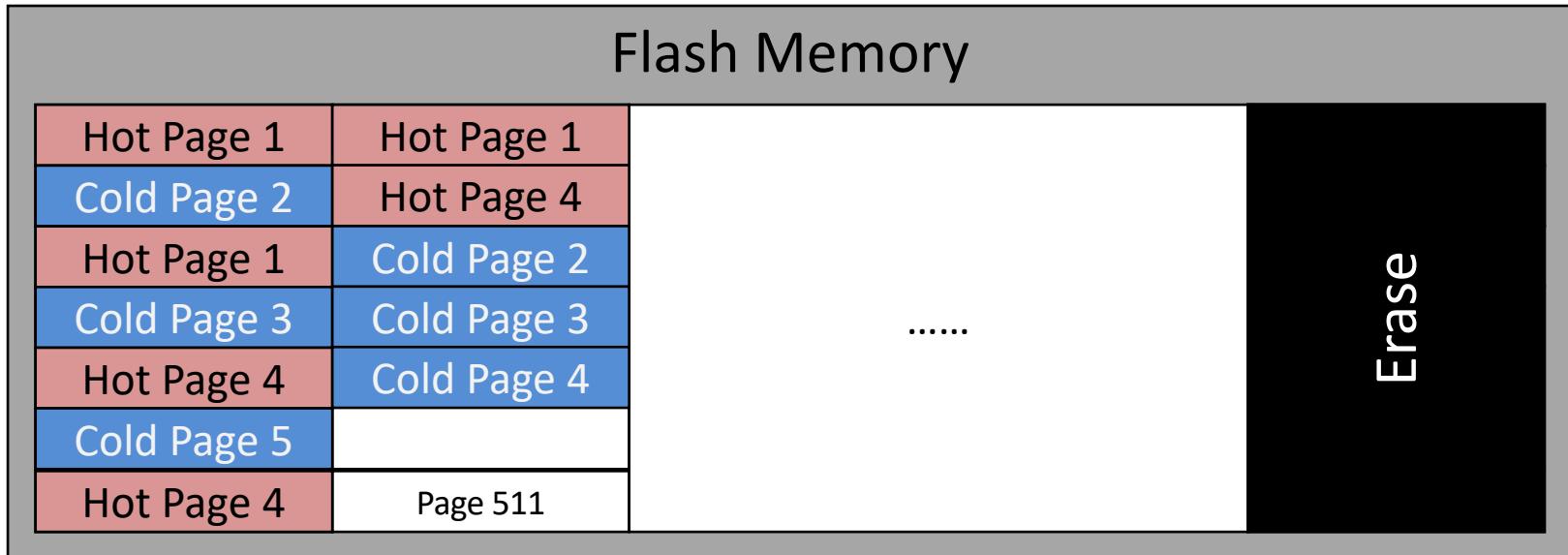
Carnegie Mellon



Executive Summary

- Flash memory can achieve **50x endurance improvement by relaxing retention time using refresh** [Cai+ ICCD '12]
- *Problem: Frequent refresh consumes the majority of endurance improvement*
- *Goal: Reduce refresh overhead to increase flash memory lifetime*
- *Key Observation: Refresh is unnecessary for write-hot data*
- *Key Ideas of Write-hotness Aware Retention Management (WARM)*
 - **Physically partition write-hot pages and write-cold pages** within the flash drive
 - **Apply different policies** (garbage collection, wear-leveling, refresh) to each group
- *Key Results*
 - WARM w/o refresh **improves lifetime by 3.24x**
 - WARM w/ adaptive refresh **improves lifetime by 12.9x** (1.21x over refresh only)

Conventional Write-Hotness Oblivious Management



Unable to relax retention time for blocks with write-hot and cold pages



Key Idea: Write-Hotness Aware Management

Flash Memory				
Hot Page 1	Cold Page 2	Hot Page 4		Page M
Hot Page 1	Cold Page 3	Hot Page 1		Page M+1
Hot Page 4	Cold Page 5			Page M+2
Hot Page 4			
Hot Page 1
Hot Page 4				
Hot Page 1	Page 511			Page M+255

Can relax retention time for blocks with write-hot pages only



Write-Hotness Aware Retention Management

- Yixin Luo, Yu Cai, Saugata Ghose, Jongmoo Choi, and Onur Mutlu,
"WARM: Improving NAND Flash Memory Lifetime with Write-hotness Aware Retention Management"

Proceedings of the 31st International Conference on Massive Storage Systems and Technologies (MSST), Santa Clara, CA, June 2015.
[Slides (pptx) (pdf)] [Poster (pdf)]

WARM: Improving NAND Flash Memory Lifetime with Write-hotness Aware Retention Management

Yixin Luo
yixinluo@cs.cmu.edu

Yu Cai
yucaicai@gmail.com

Saugata Ghose
ghose@cmu.edu

Jongmoo Choi[†]
choijm@dankook.ac.kr

Onur Mutlu
onur@cmu.edu

Agenda

- Background, Motivation and Approach
 - Experimental Characterization Methodology
 - Error Analysis and Management
 - Main Characterization Results
 - Retention-Aware Error Management
 - Threshold Voltage and Program Interference Analysis
 - Read Reference Voltage Prediction
 - Neighbor-Assisted Error Correction
 - Read Disturb Error Handling
 - Retention Error Handling
 - Large Scale Field Analysis
 - 3D NAND Flash Memory Reliability
 - Summary
-

Summary of Key Works

- Yu Cai, Saugata Ghose, Erich F. Haratsch, Yixin Luo, and Onur Mutlu,
"Error Characterization, Mitigation, and Recovery in Flash Memory Based Solid State Drives"

Proceedings of the IEEE, September 2017.

- Cai+, "Error Patterns in MLC NAND Flash Memory: Measurement, Characterization, and Analysis," DATE 2012.
- Cai+, "Flash Correct-and-Refresh: Retention-Aware Error Management for Increased Flash Memory Lifetime," ICCD 2012.
- Cai+, "Threshold Voltage Distribution in MLC NAND Flash Memory: Characterization, Analysis and Modeling," DATE 2013.
- Cai+, "Error Analysis and Retention-Aware Error Management for NAND Flash Memory," Intel Technology Journal 2013.
- Cai+, "Program Interference in MLC NAND Flash Memory: Characterization, Modeling, and Mitigation," ICCD 2013.
- Cai+, "Neighbor-Cell Assisted Error Correction for MLC NAND Flash Memories," SIGMETRICS 2014.
- Cai+, "Data Retention in MLC NAND Flash Memory: Characterization, Optimization and Recovery," HPCA 2015.
- Cai+, "Read Disturb Errors in MLC NAND Flash Memory: Characterization and Mitigation," DSN 2015.
- Luo+, "WARM: Improving NAND Flash Memory Lifetime with Write-hotness Aware Retention Management," MSST 2015.
- Meza+, "A Large-Scale Study of Flash Memory Errors in the Field," SIGMETRICS 2015.
- Luo+, "Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory," IEEE JSAC 2016.
- Cai+, "Vulnerabilities in MLC NAND Flash Memory Programming: Experimental Analysis, Exploits, and Mitigation Techniques," HPCA 2017.
- Fukami+, "Improving the Reliability of Chip-Off Forensic Analysis of NAND Flash Memory Devices," DFRWS EU 2017.
- Luo+, "HeatWatch: Improving 3D NAND Flash Memory Device Reliability by Exploiting Self-Recovery and Temperature-Awareness," HPCA 2018.
- Luo+, "Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation," SIGMETRICS 2018.
- Cai+, "Error Characterization, Mitigation, and Recovery in Flash Memory Based Solid State Drives," Proc. IEEE 2017.

NAND Flash Vulnerabilities [HPCA'17]

HPCA, Feb. 2017

Vulnerabilities in MLC NAND Flash Memory Programming: Experimental Analysis, Exploits, and Mitigation Techniques

Yu Cai[†] Saugata Ghose[†] Yixin Luo^{‡‡} Ken Mai[†] Onur Mutlu^{§†} Erich F. Haratsch[‡]
[†]Carnegie Mellon University [‡]Seagate Technology [§]ETH Zürich

Modern NAND flash memory chips provide high density by storing two bits of data in each flash cell, called a multi-level cell (MLC). An MLC partitions the threshold voltage range of a flash cell into four voltage states. When a flash cell is programmed, a high voltage is applied to the cell. Due to parasitic capacitance coupling between flash cells that are physically close to each other, flash cell programming can lead to cell-to-cell program interference, which introduces errors into neighboring flash cells. In order to reduce the impact of cell-to-cell interference on the reliability of MLC NAND flash memory, flash manufacturers adopt a two-step programming method, which programs the MLC in two separate steps. First, the flash memory partially programs the least significant bit of the MLC to some intermediate threshold voltage. Second, it programs the most significant bit to bring the MLC up to its full voltage state.

In this paper, we demonstrate that two-step programming exposes new reliability and security vulnerabilities. We expe-

belongs to a different flash memory *page* (the unit of data programmed and read at the same time), which we refer to, respectively, as the least significant bit (LSB) page and the most significant bit (MSB) page [5].

A flash cell is programmed by applying a large voltage on the control gate of the transistor, which triggers charge transfer into the floating gate, thereby increasing the threshold voltage. To precisely control the threshold voltage of the cell, the flash memory uses *incremental step pulse programming* (ISPP) [12, 21, 25, 41]. ISPP applies multiple short pulses of the programming voltage to the control gate, in order to increase the cell threshold voltage by some small voltage amount (V_{step}) after each step. Initial MLC designs programmed the threshold voltage in *one shot*, issuing all of the pulses back-to-back to program *both* bits of data at the same time. However, as flash memory scales down, the distance between neighboring flash cells decreases, which

https://people.inf.ethz.ch/omutlu/pub/flash-memory-programming-vulnerabilities_hPCA17.pdf

NAND Flash Errors: A Modern Survey



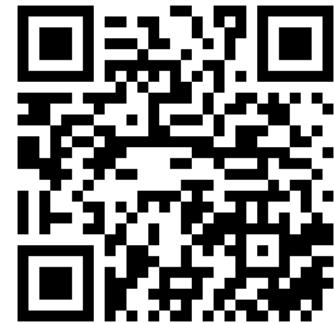
Proceedings of the IEEE, Sept. 2017

Error Characterization, Mitigation, and Recovery in Flash-Memory-Based Solid-State Drives

This paper reviews the most recent advances in solid-state drive (SSD) error characterization, mitigation, and data recovery techniques to improve both SSD's reliability and lifetime.

By YU CAI, SAUGATA GHOSE, ERICH F. HARATSCH, YIXIN LUO, AND ONUR MUTLU

<https://arxiv.org/pdf/1706.08642>



More Up-to-date Version

- Yu Cai, Saugata Ghose, Erich F. Haratsch, Yixin Luo, and Onur Mutlu,
"Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery"

Invited Book Chapter in Inside Solid State Drives, 2018.

[Preliminary arxiv.org version]

Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery

YU CAI, SAUGATA GHOSE

Carnegie Mellon University

ERICH F. HARATSCH

Seagate Technology

YIXIN LUO

Carnegie Mellon University

ONUR MUTLU

ETH Zürich and Carnegie Mellon University

Computer Architecture

Lecture 16: Flash Memory and Solid-State Drives

Dr. Mohammad Sadrosadati

Prof. Onur Mutlu

ETH Zürich

Fall 2023

17 November 2023

Other Works on Flash Memory

HeatWatch

Improving 3D NAND Flash Memory Device Reliability by
Exploiting Self-Recovery and Temperature Awareness

Yixin Luo Saugata Ghose Yu Cai Erich F. Haratsch Onur Mutlu

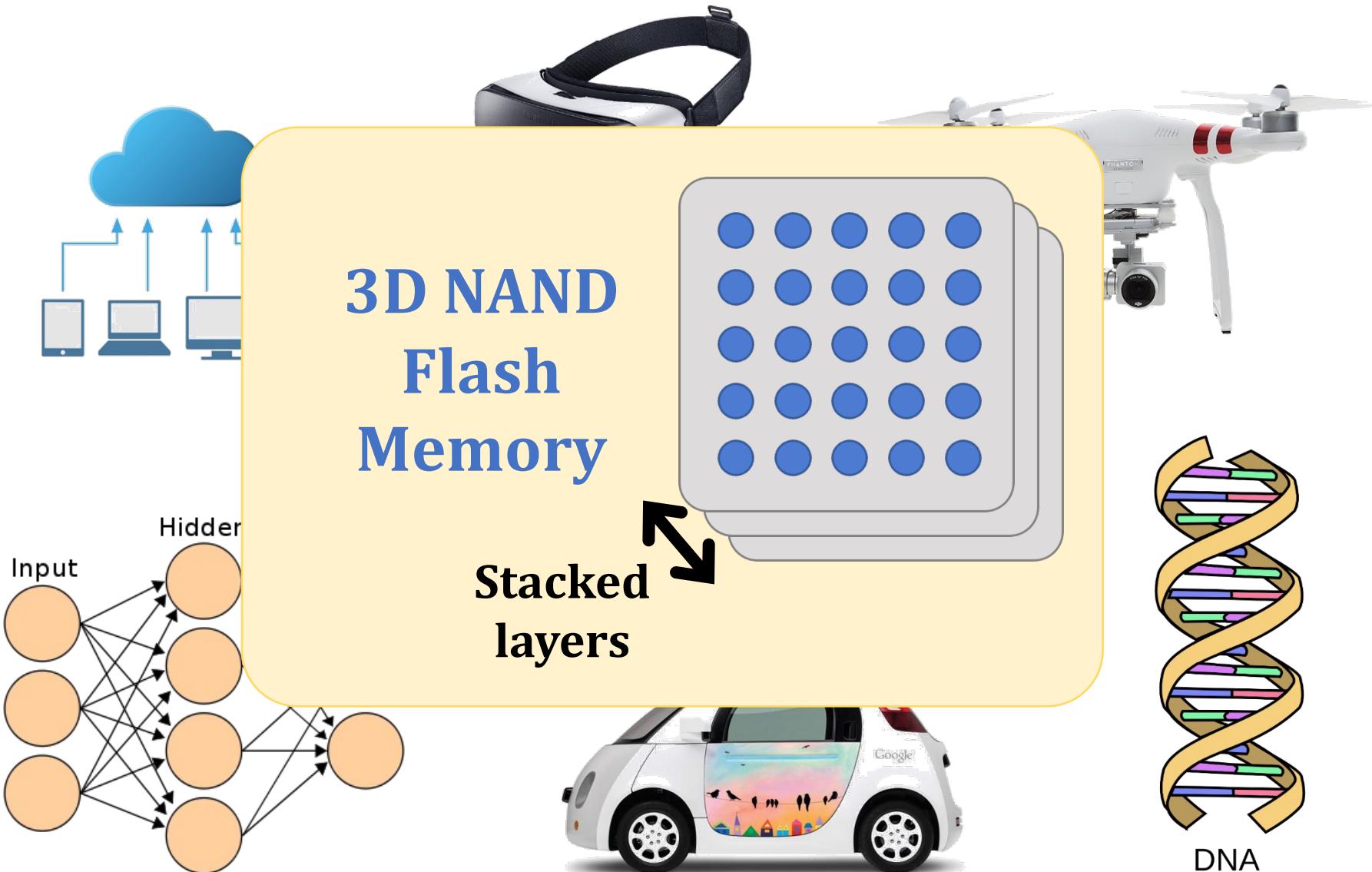
Carnegie Mellon

SAFARI



ETH Zürich

Storage Technology Drivers - 2018



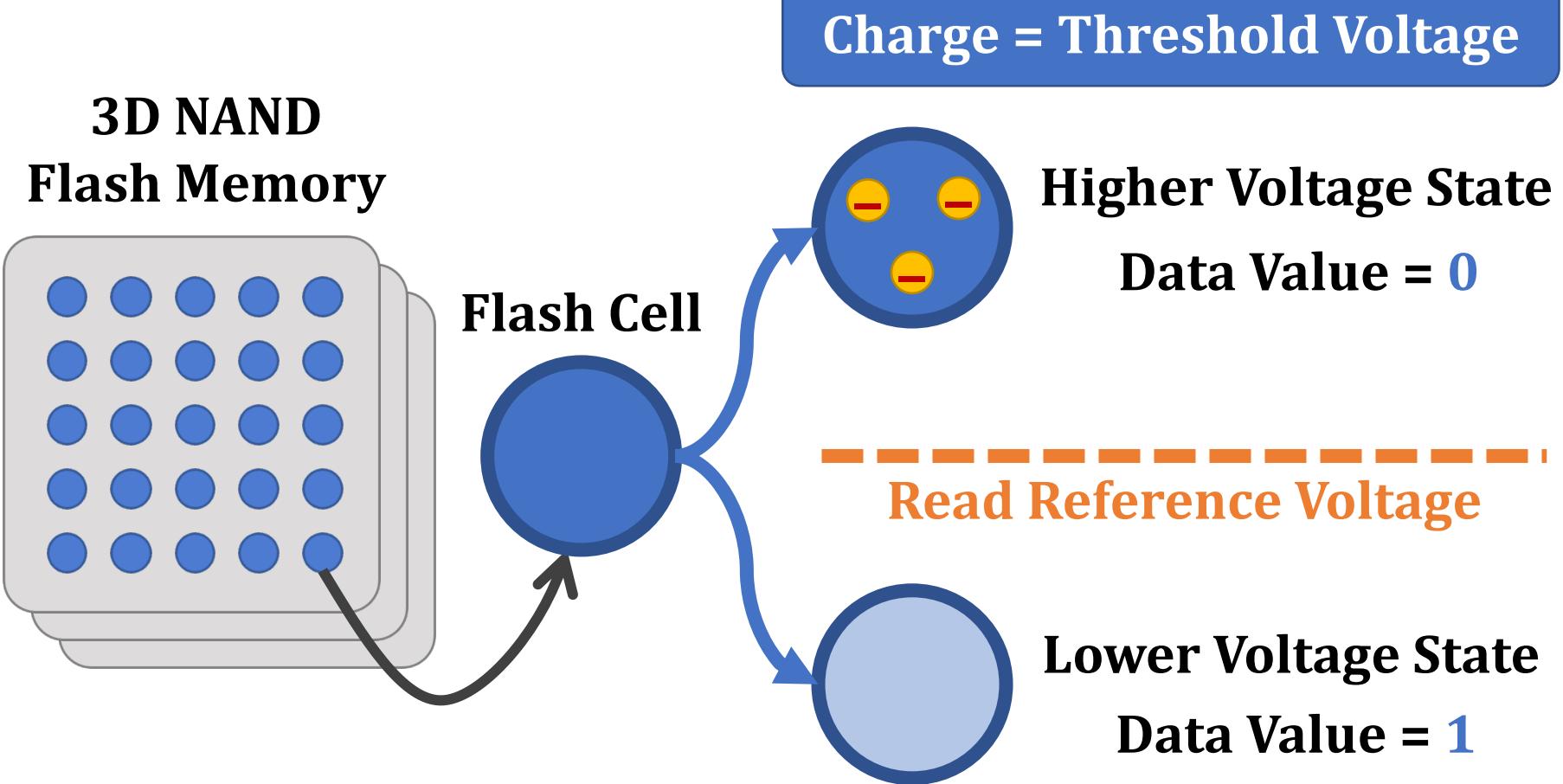
Executive Summary

- 3D NAND flash memory susceptible to **retention errors**
 - Charge leaks out of flash cell
 - Two unreported factors: *self-recovery* and *temperature*
- We study *self-recovery* and *temperature* effects
 - **Experimental characterization** of *real* 3D NAND chips
- **Unified Self-Recovery and Temperature (URT) Model**
 - Predicts impact of retention loss, wearout, self-recovery, temperature on **flash cell voltage**
 - **Low prediction error rate: 4.9%**
- We develop a new technique to improve flash reliability
 - **HeatWatch**
 - Uses URT model to find optimal read voltages for 3D NAND flash
 - **Improves flash lifetime by 3.85x**

Outline

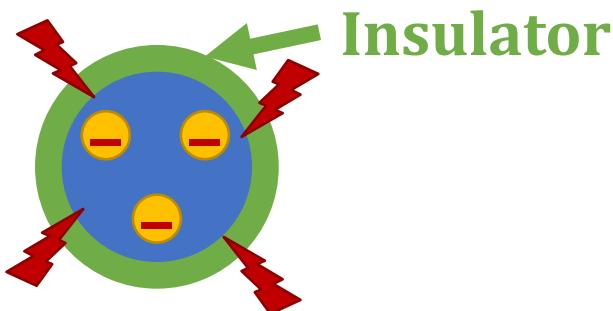
- Executive Summary
- **Background on NAND Flash Reliability**
- Characterization of Self-Recovery and Temperature Effect on Real 3D NAND Flash Memory Chips
- URT: Unified Self-Recovery and Temperature Model
- HeatWatch Mechanism
- Conclusion

3D NAND Flash Memory Background



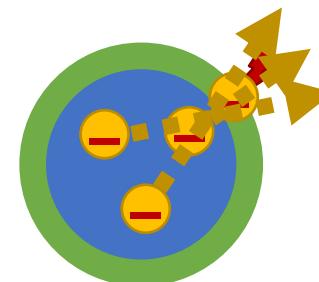
Flash Wearout

Program/Erase (P/E) → Wearout



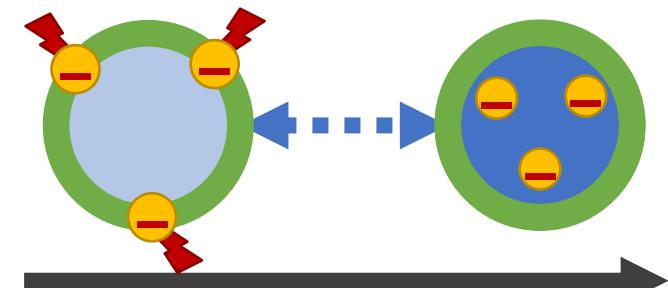
Wearout Effects:

1. **Retention Loss**
(voltage shift over time)



2. **Program Variation**
(init. voltage difference b/w states)

Wearout Introduces Errors



Voltage

Improving Flash Lifetime

**Errors introduced by wearout
limit flash lifetime**
(measured in P/E cycles)

**Two Ways to Improve
Flash Lifetime**

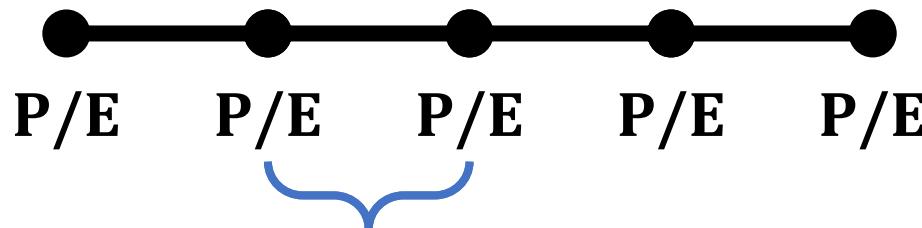


**Exploiting the
Self-Recovery Effect**

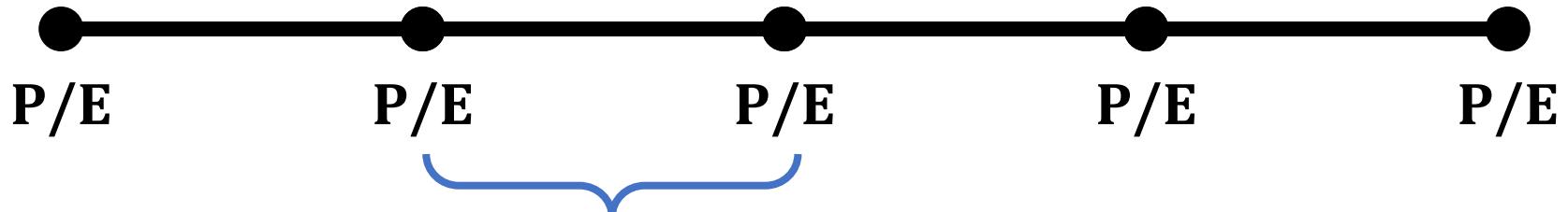
**Exploiting the
Temperature Effect**

Exploiting the Self-Recovery Effect

Partially repairs damage due to wearout



Dwell Time: Idle Time Between P/E Cycles

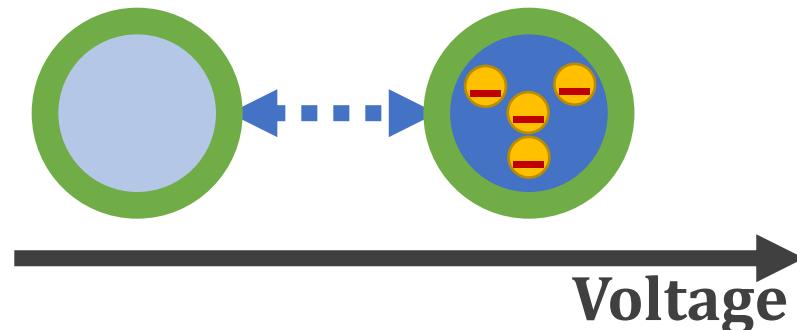


Longer Dwell Time: More Self-Recovery

Reduces Retention Loss

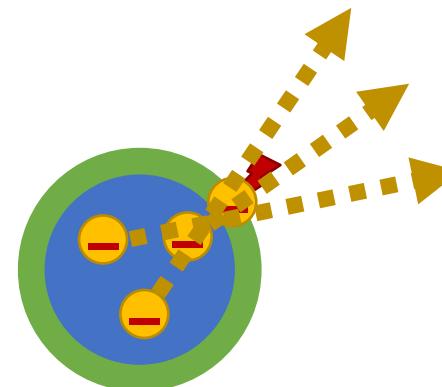
Exploiting the Temperature Effect

**High Program
Temperature**



Increases Program Variation

**High Storage
Temperature**



Accelerates Retention Loss

Prior Studies of Self-Recovery/Temperature

**Self-Recovery
Effect**

Planar (2D) NAND



3D NAND



**Temperature
Effect**



Mielke 2006
JEDEC 2010
(no characterization)



Outline

- Executive Summary
- Background on NAND Flash Reliability
- **Characterization of Self-Recovery and Temperature Effect on Real 3D NAND Flash Memory Chips**
- URT: Unified Self-Recovery and Temperature Model
- HeatWatch Mechanism
- Conclusion

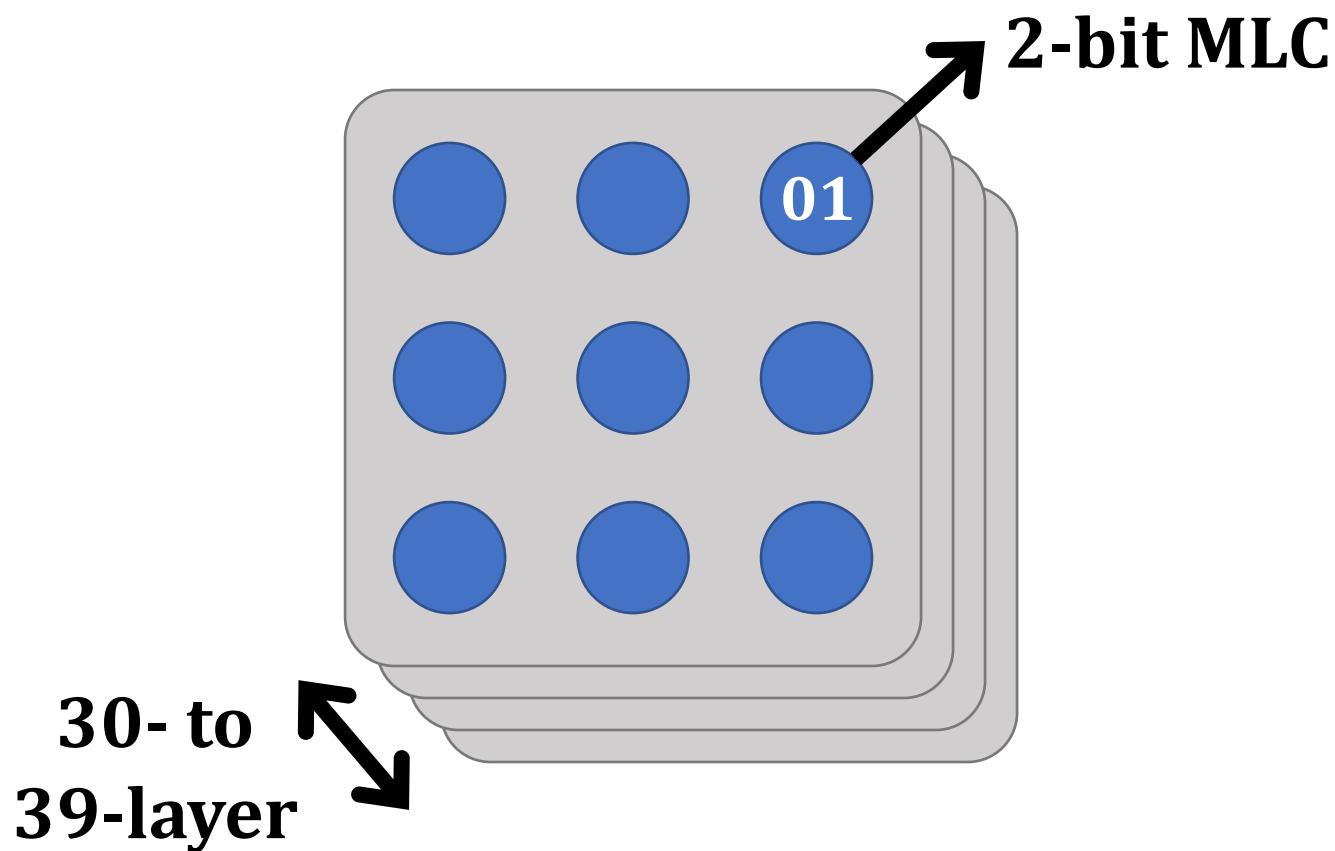
Characterization Methodology

- Modified firmware version in the flash controller
 - Control the read reference voltage of the flash chip
 - Bypass ECC to get raw NAND data (with raw bit errors)
- Control temperature with a heat chamber

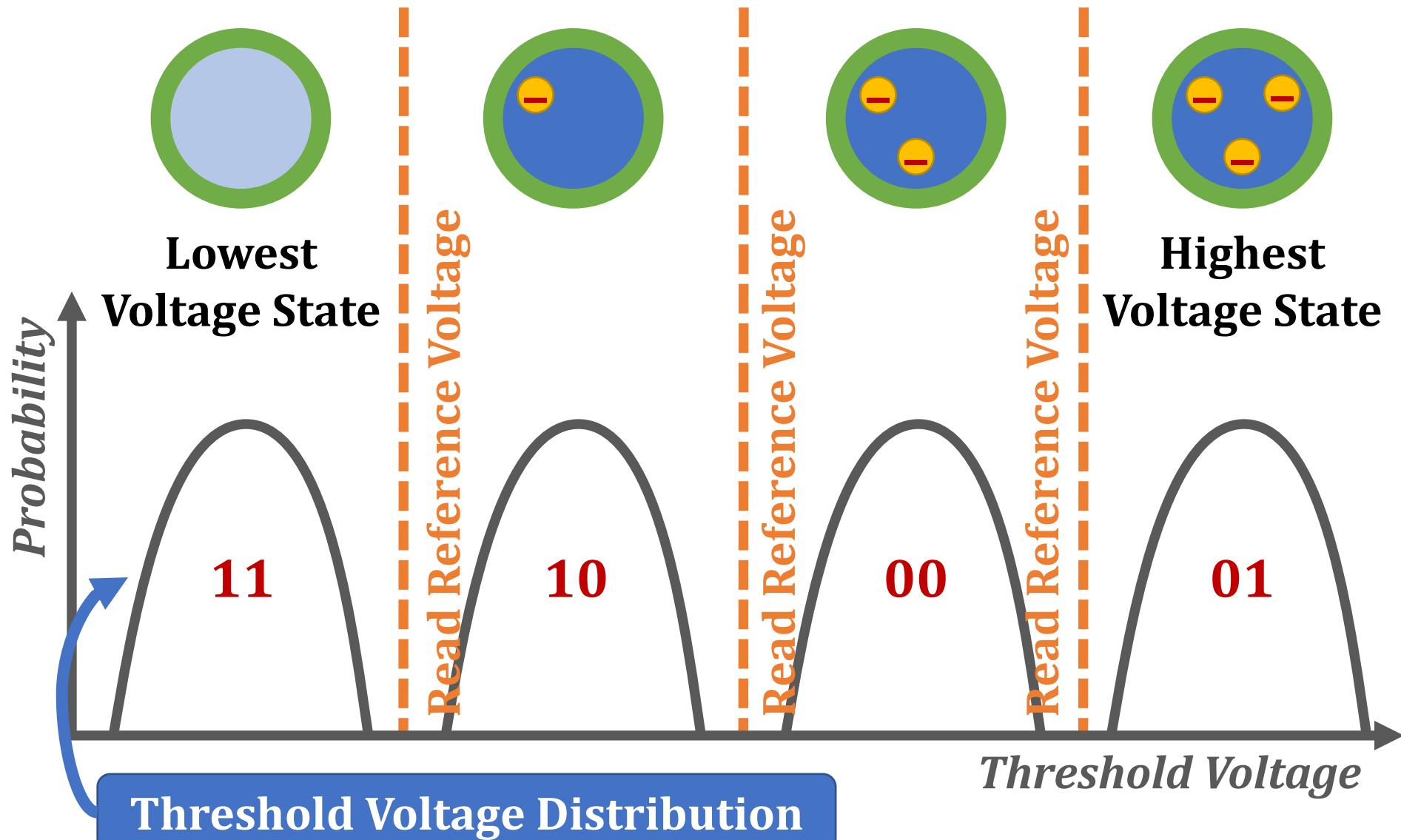


Characterized Devices

Real 30-39 Layer 3D MLC NAND Flash Chips



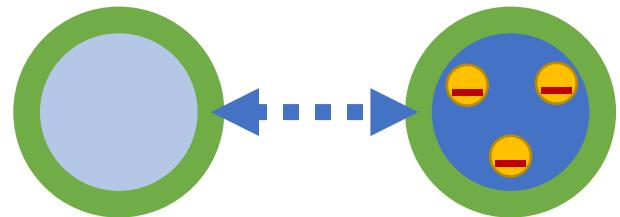
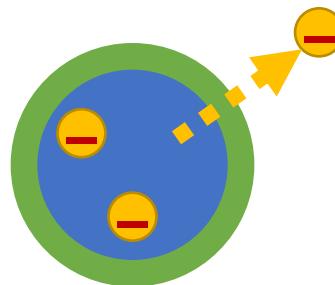
MLC Threshold Voltage Distribution Background



Characterization Goal

Characterized
Metrics

Characterized
Phenomena



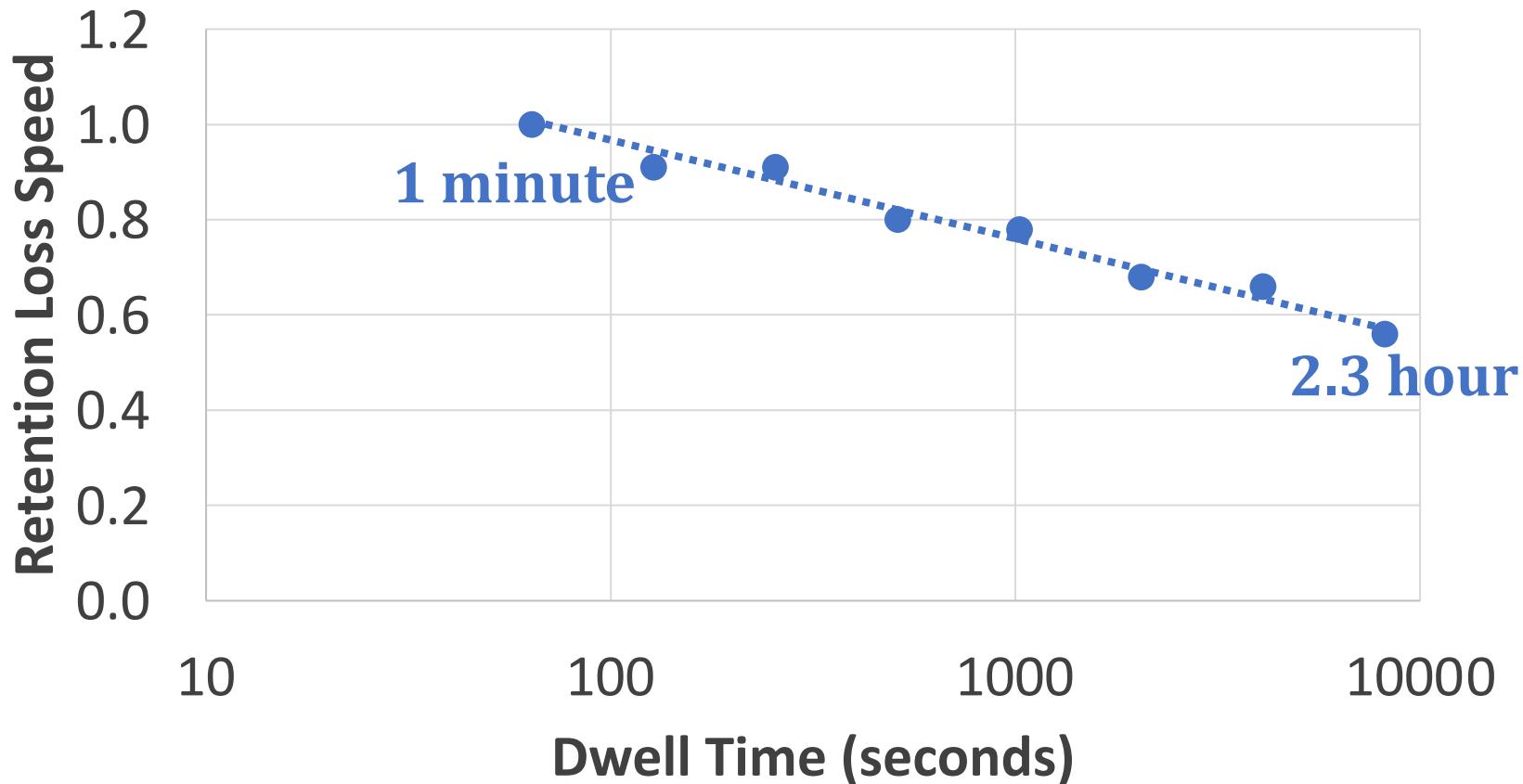
Retention Loss Speed
(how fast voltage shifts
over time)

Program Variation
(initial voltage difference
between states)

Self-Recovery
Effect

Temperature
Effect

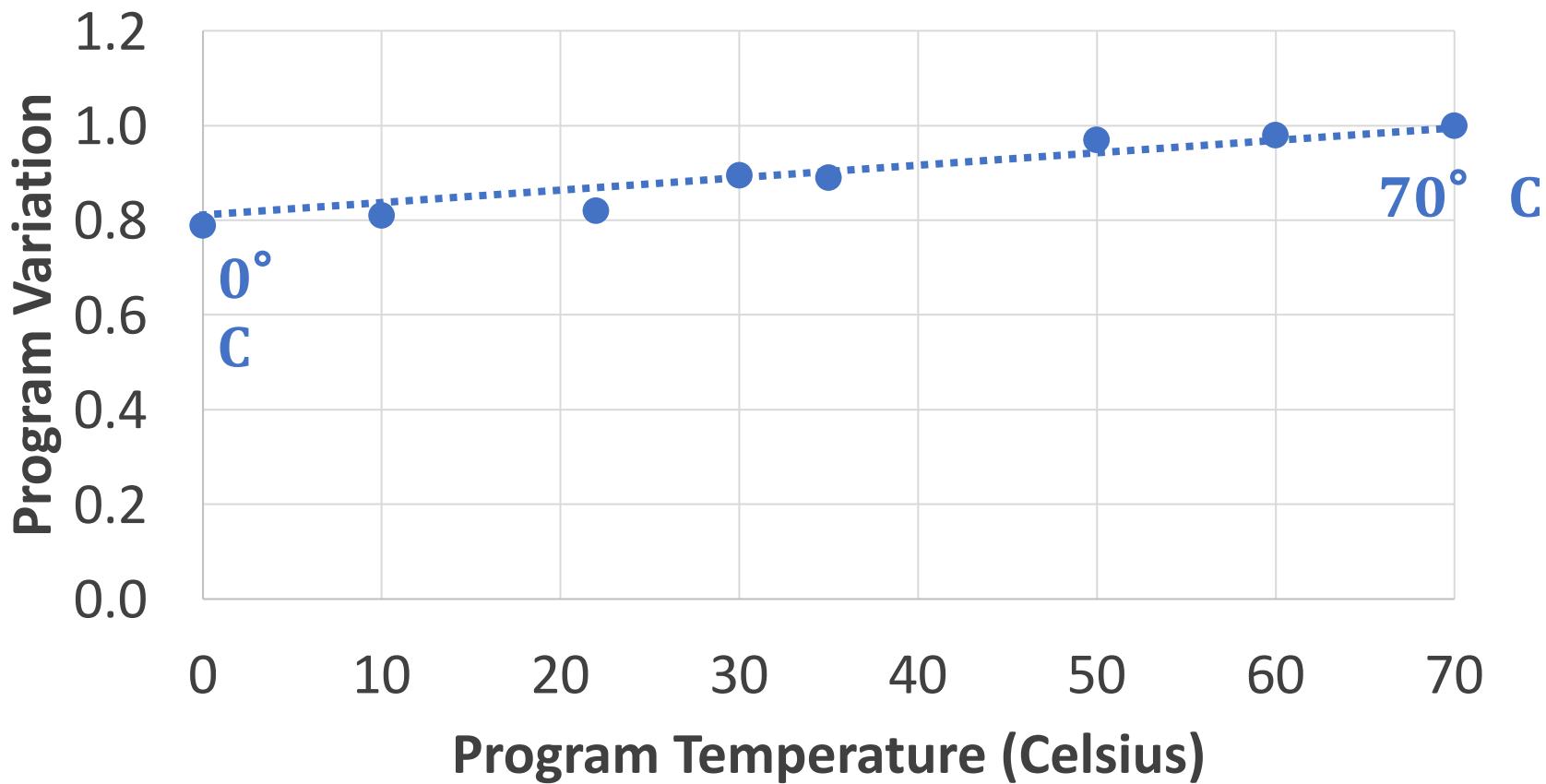
Self-Recovery Effect Characterization Results



Dwell time: Idle time between P/E cycles

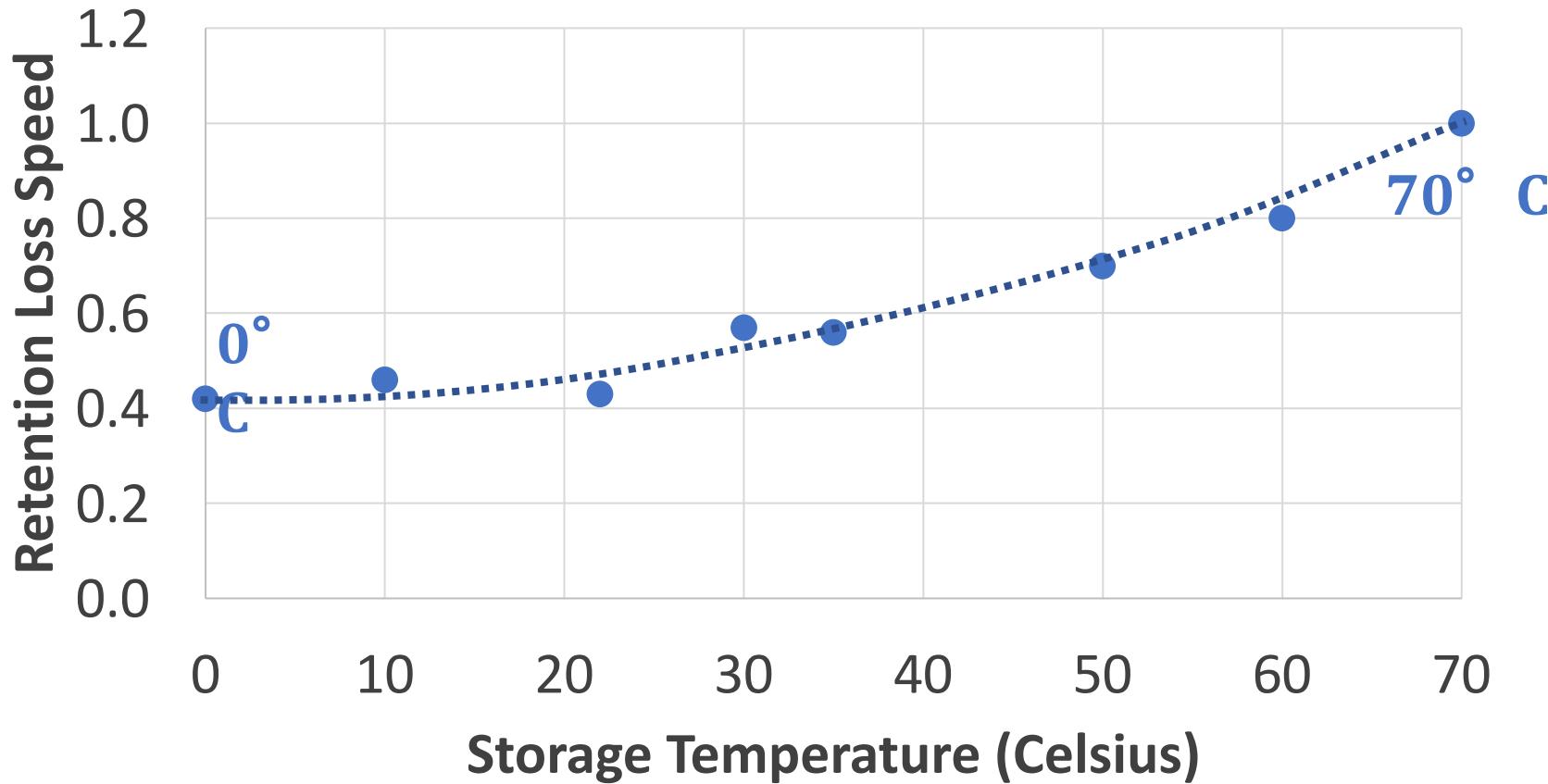
Increasing dwell time from 1 minute to 2.3 hours
slows down retention loss speed by 40%

Program Temperature Effect Characterization Results



Increasing program temperature from 0°C to 70°C
improves program variation by 21%

Storage Temperature Effect Characterization Results



Lowering storage temperature from 70°C to 0°C
slows down retention loss speed by 58%

Characterization Summary

Major Results:

- *Self-recovery* affects retention loss speed
- Program *temperature* affects program variation
- *Storage temperature* affects retention loss speed

Unified Model

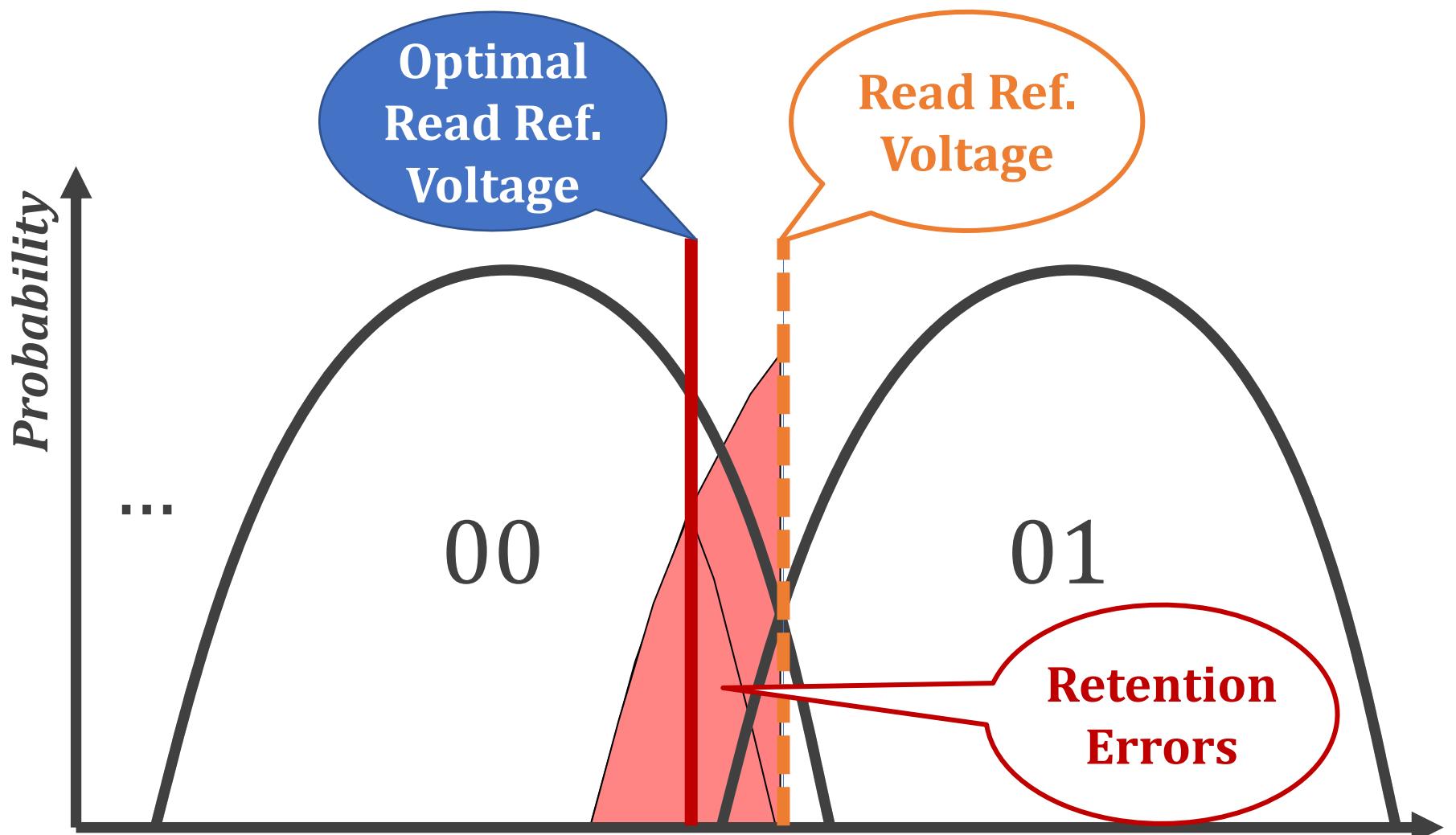
Other Characterizations Methods in the Paper:

- More detailed results on self-recovery and temperature
 - Effects on error rate
 - Effects on threshold voltage distribution
- Effects of recovery cycle (P/E cycles with long dwell time) on retention loss speed

Outline

- Executive Summary
- Background on NAND Flash Reliability
- Characterization of Self-Recovery and Temperature Effect on Real 3D NAND Flash Memory Chips
- **URT: Unified Self-Recovery and Temperature Model**
- HeatWatch Mechanism
- Conclusion

Minimizing 3D NAND Errors



Optimal read reference voltage
minimizes 3D NAND errors

Predicting the Mean Threshold Voltage

Our URT Model:

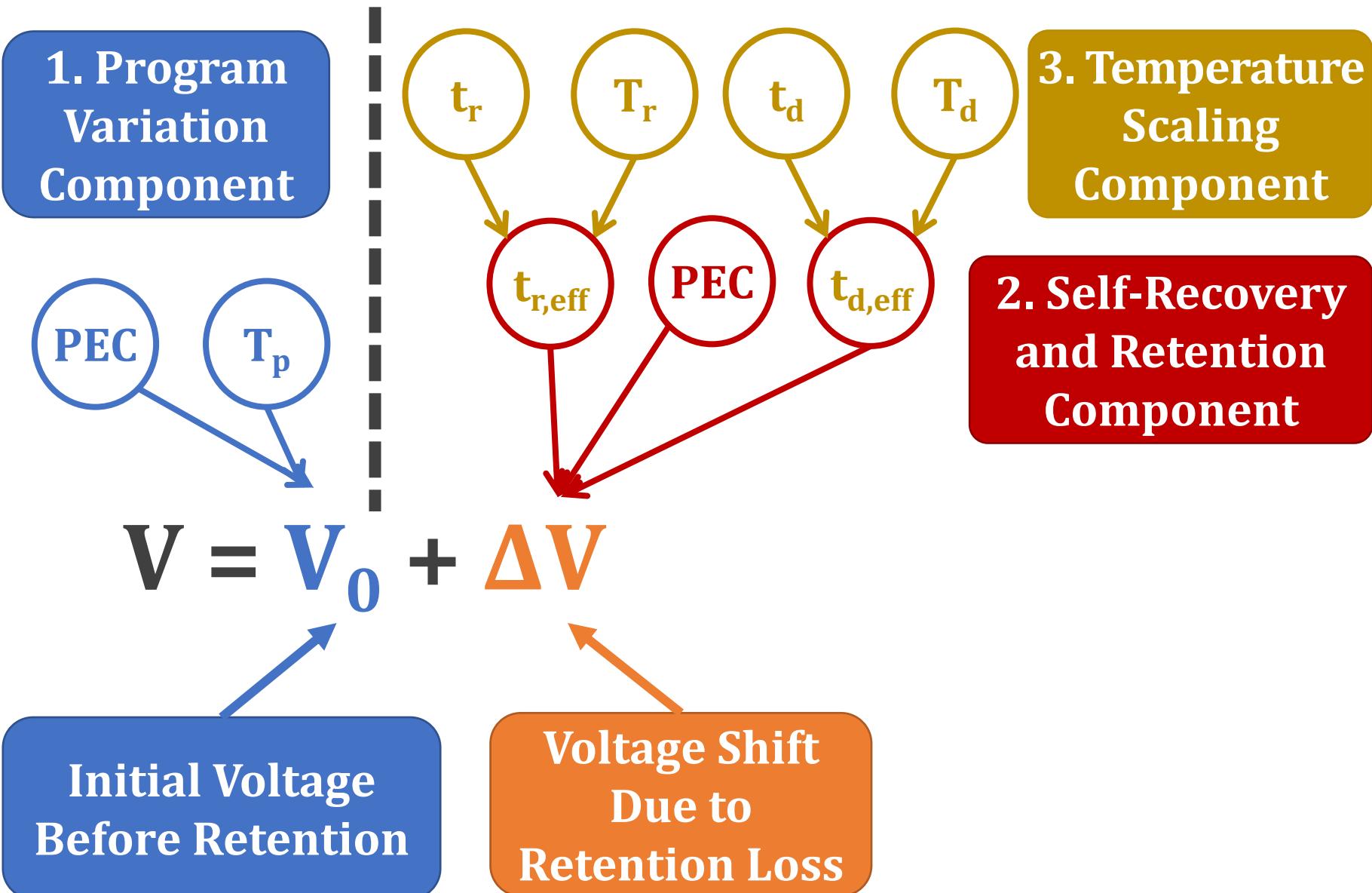
$$V = V_0 + \Delta V$$

Mean
Threshold
Voltage

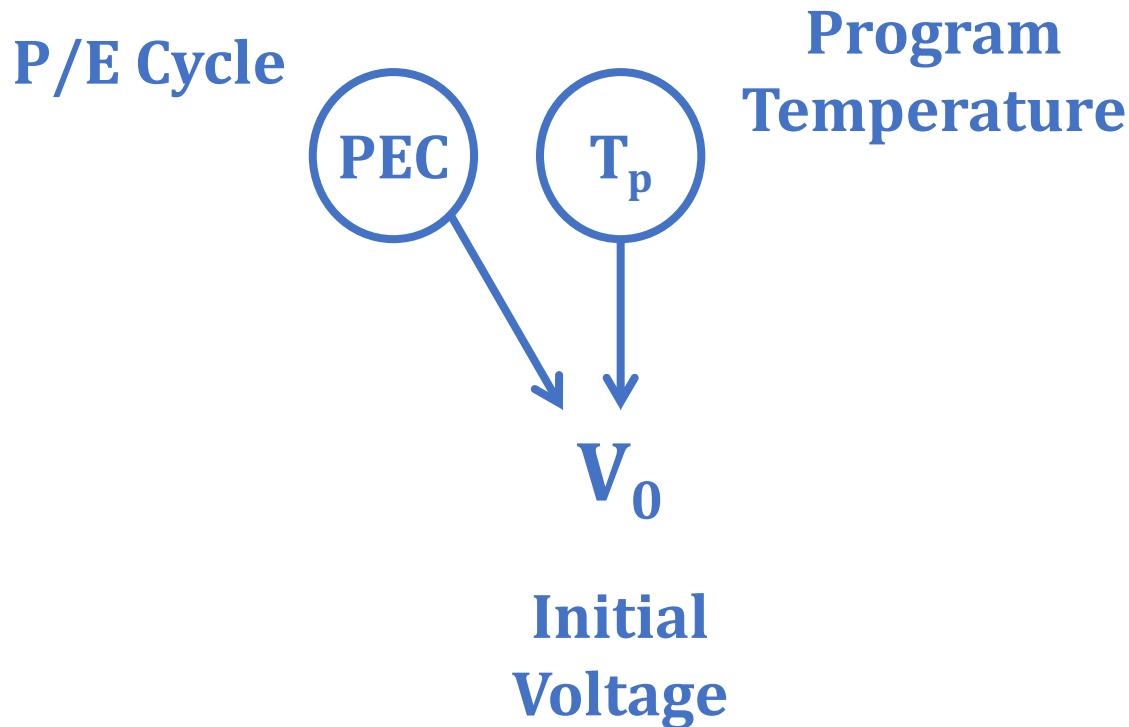
Initial Voltage
Before Retention
(Program Variation)

Voltage Shift
Due to
Retention Loss

URT Model Overview



1. Program Variation Component

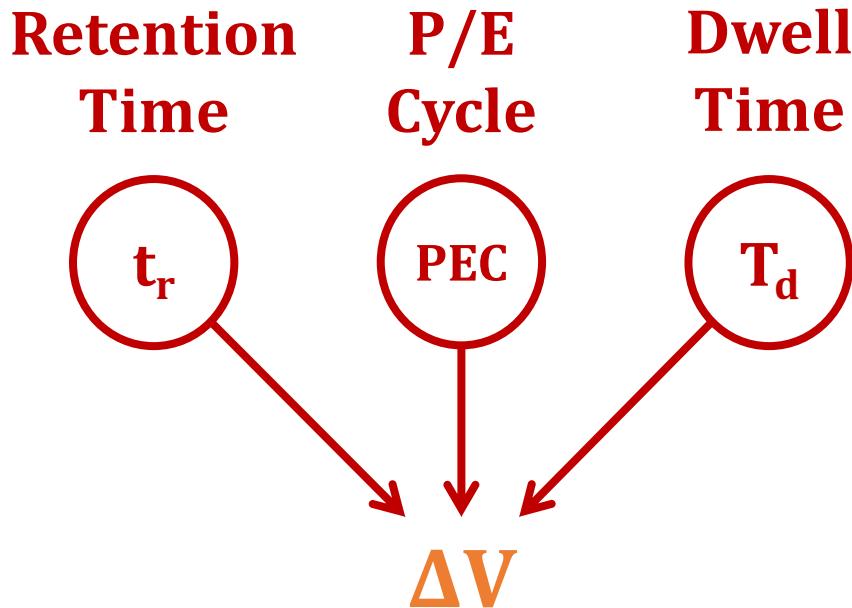


$$V_0 = A \cdot T_p \cdot PEC + B \cdot T_p + C \cdot PEC + D$$



Validation: $R^2 = 91.7\%$

2. Self-Recovery and Retention Component



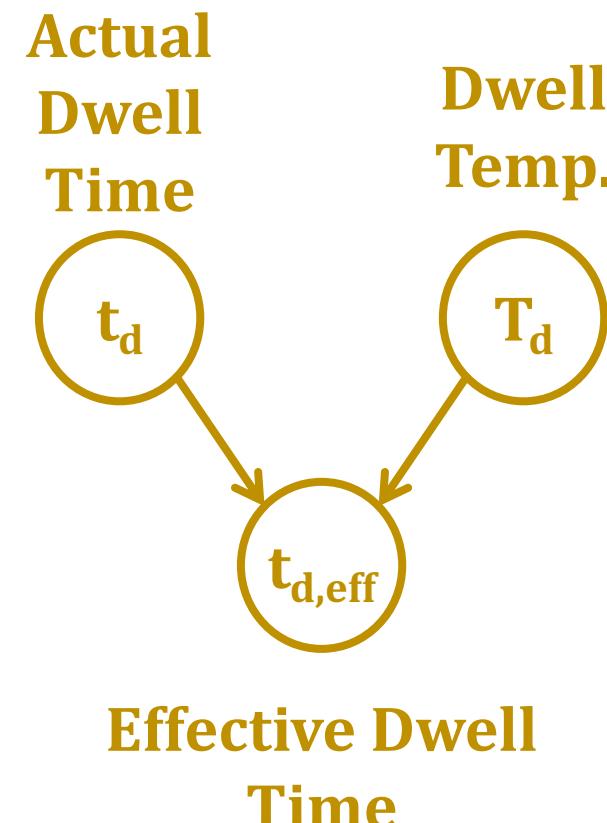
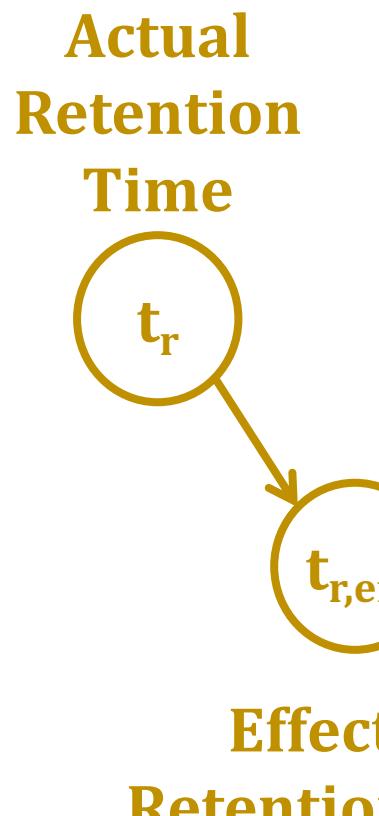
Retention Shift

$$\Delta V(t_{er}, t_{ed}, PEC) = b \cdot (PEC + c) \cdot \ln \left(1 + \frac{t_{er}}{t_0 + a \cdot t_{ed}} \right)$$



Validation: 3x more accurate
than state-of-the-art model

3. Temperature Scaling Component



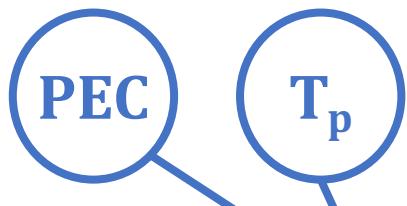
Arrhenius Equation: $AF = \frac{t_{real}}{t_{room}} = \exp\left(\frac{E_a}{k_B} \cdot \left(\frac{1}{T_{real}} - \frac{1}{T_{room}}\right)\right)$



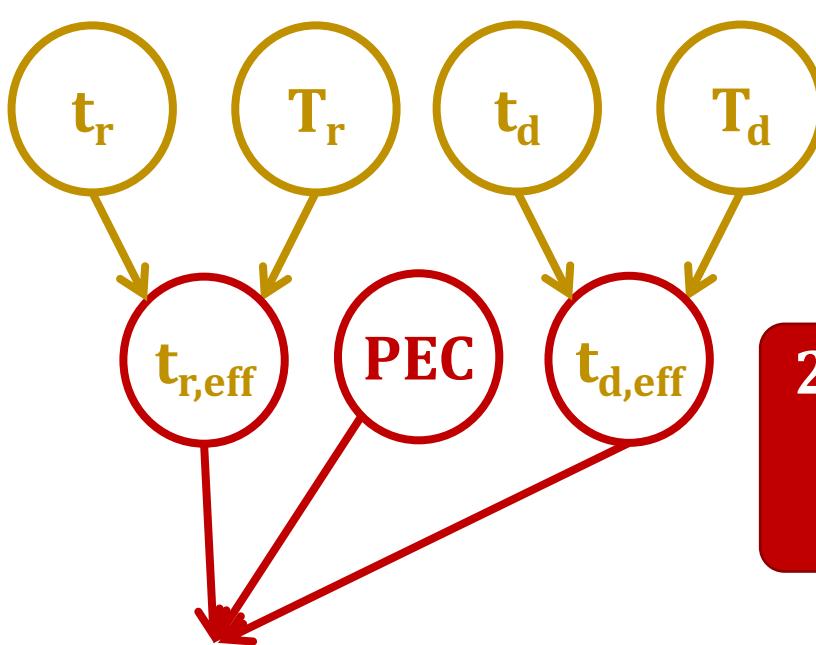
Validation: Adjust an important parameter, E_a , from 1.1 eV to 1.04 eV

URT Model Summary

1. Program Variation Component



$$V = V_0 + \Delta V$$



3. Temperature Scaling Component

2. Self-Recovery and Retention Component

Validation:

Prediction Error Rate = 4.9%

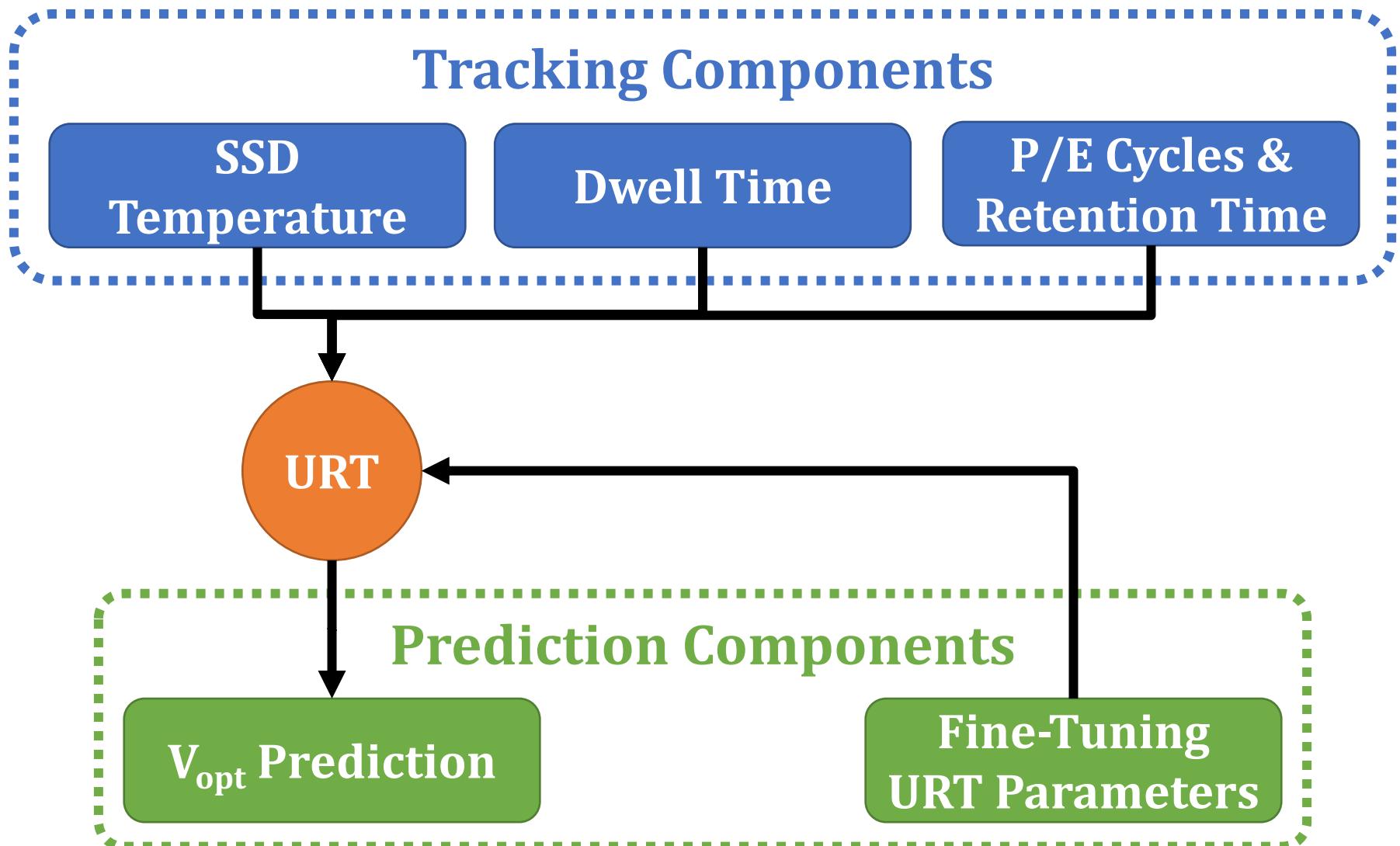
Outline

- Executive Summary
- Background on NAND Flash Reliability
- Characterization of Self-Recovery and Temperature Effect on Real 3D NAND Flash Memory Chips
- URT: Unified Self-Recovery and Temperature Model
- **HeatWatch Mechanism**
- Conclusion

HeatWatch Mechanism

- Key Idea
- Predict change in threshold voltage distribution by using the URT model
- Adapt read reference voltage to near-optimal (V_{opt}) based on predicted change in voltage distribution

HeatWatch Mechanism Overview



Tracking SSD Temperature

Tracking Components

SSD
Temperature

Dwell Time

P/E Cycles &
Retention Time

- Use existing sensors in the SSD
- **Precompute** temperature scaling factor at **logarithmic time intervals**

Prediction Components

V_{opt} Prediction

Fine-Tuning
URT Parameters

Tracking Dwell Time

Tracking Components

SSD
Temperature

Dwell Time

P/E Cycles &
Retention Time

- Only need to log the timestamps of **last 20 full drive writes**
- Self-recovery effect diminishes after 20 P/E cycles

Prediction Components

V_{opt} Prediction

Fine-Tuning
URT Parameters

Tracking P/E Cycles and Retention Time

Tracking Components

SSD
Temperature

Dwell Time

P/E Cycles &
Retention Time

- P/E cycle count **already recorded** by SSD
- **Log write timestamp** for each block
- Retention time = read timestamp – write timestamp

Prediction Components

V_{opt} Prediction

Fine-Tuning
URT Parameters

Predicting Optimal Read Reference Voltage

Tracking Components

SSD
Temperature

Dwell Time

P/E Cycles &
Retention Time

- Calculate URT using tracked information
- Modeling error: 4.9%

Prediction Components

V_{opt} Prediction

Fine-Tuning
URT Parameters

Fine-Tuning URT Parameters Online

Tracking Components

SSD
Temperature

Dwell Time

P/E Cycles &
Retention Time

- Accommodates **chip-to-chip variation**
- Uses **periodic sampling**

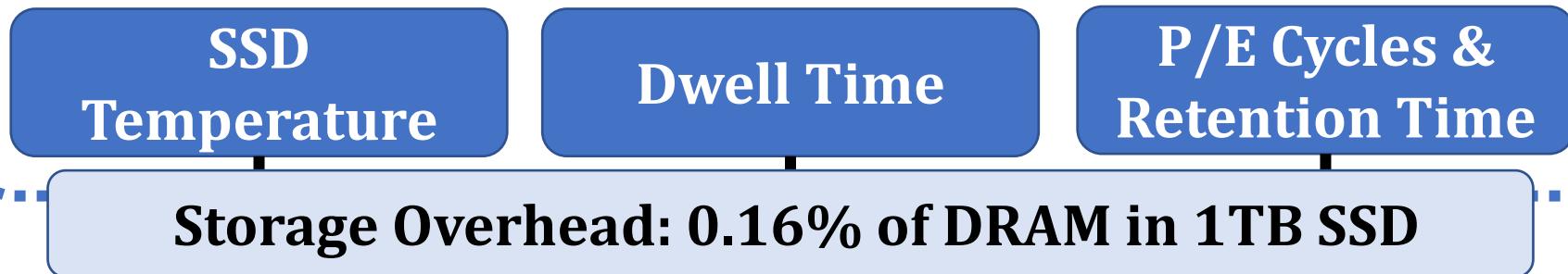
Prediction Components

V_{opt} Prediction

Fine-Tuning
URT Parameters

HeatWatch Mechanism Summary

Tracking Components



URT

Prediction Components

V_{opt} Prediction

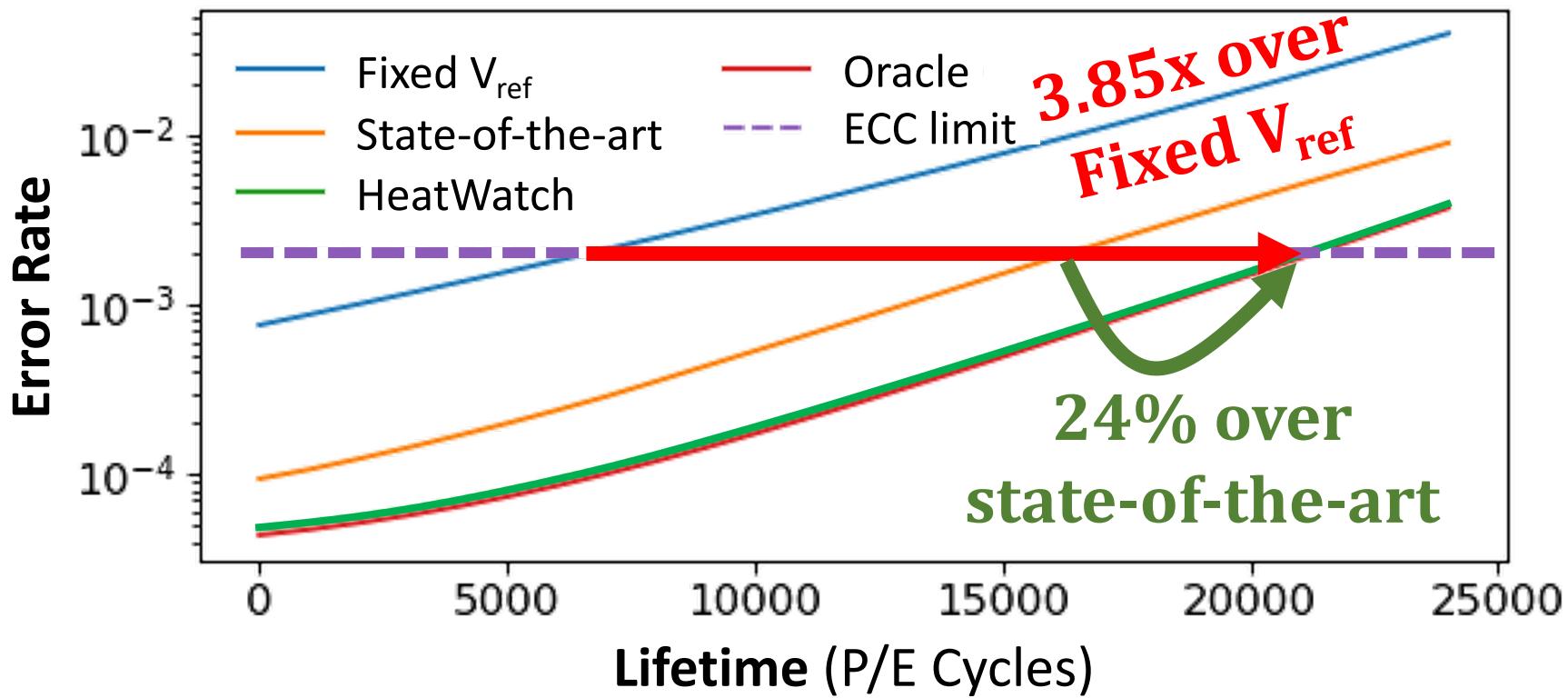
Fine-Tuning
URT Parameters

Latency Overhead: < 1% of flash read latency

HeatWatch Evaluation Methodology

- **28 real workload storage traces**
 - MSR-Cambridge
 - We use **real dwell time, retention time values** obtained from traces
- **Temperature Model:**
Trigonometric function + Gaussian noise
 - Represents **periodic temperature variation** in each day
 - Includes **small transient temperature variation**

HeatWatch Greatly Improves Flash Lifetime



HeatWatch improves lifetime by
capturing the effect of
retention, wearout, self-recovery, temperature

Outline

- Executive Summary
- Background on NAND Flash Reliability
- Characterization of Self-Recovery and Temperature Effect on Real 3D NAND Flash Memory Chips
- URT: Unified Self-Recovery and Temperature Model
- HeatWatch Mechanism
- **Conclusion**

Conclusion

- 3D NAND flash memory susceptible to **retention errors**
 - Charge leaks out of flash cell
 - Two unreported factors: *self-recovery* and *temperature*
- We study *self-recovery* and *temperature* effects
 - **Experimental characterization** of *real* 3D NAND chips
- **Unified Self-Recovery and Temperature (URT) Model**
 - Predicts impact of retention loss, wearout, self-recovery, temperature on **flash cell voltage**
 - **Low prediction error rate: 4.9%**
- We develop a new technique to improve flash reliability
 - **HeatWatch**
 - Uses URT model to find optimal read voltages for 3D NAND flash
 - **Improves flash lifetime by 3.85x**

HeatWatch

Improving 3D NAND Flash Memory Device Reliability by
Exploiting Self-Recovery and Temperature Awareness

Yixin Luo Saugata Ghose Yu Cai Erich F. Haratsch Onur Mutlu

Carnegie Mellon

SAFARI



ETH Zürich