ERROR CHARACTERIZATION AND CORRECTION TECHNIQUES FOR RELIABLE STT-RAM DESIGNS

by

Wujie Wen

B.S. in Electronic Engineering,

Beijing Jiaotong University, China, 2006

M.S. in Electronic Engineering,

Tsinghua University, China, 2010

Submitted to the Graduate Faculty of

the Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Wujie Wen

It was defended on

June 1, 2015

and approved by

Yiran Chen, Ph.D., Associate Professor, Department of Electrical and Computer Engineering

Rami Melhem, Ph.D., Professor, Department of Computer Science

Hai Li, Ph.D., Assistant Professor, Department of Electrical and Computer Engineering

Ervin Sejdic, Ph.D., Assistant Professor, Department of Electrical and Computer Engineering

Zhi-Hong Mao, Ph.D., Associate Professor, Department of Electrical and Computer Engineering

Dissertation Director: Yiran Chen, Ph.D., Associate Professor, Department of Electrical and Computer Engineering

Copyright © by Wujie Wen
2015

ERROR CHARACTERIZATION AND CORRECTION TECHNIQUES FOR RELIABLE STT-RAM DESIGNS

Wujie Wen, PhD

University of Pittsburgh, 2015

The concerns on the continuous scaling of mainstream memory technologies have motivated tremendous investment to emerging memories. Being a promising candidate, *spin-transfer torque random access memory* (STT-RAM) offers nanosecond access time comparable to SRAM, high integration density close to DRAM, non-volatility as Flash memory, and good scalability. It is well positioned as the replacement of SRAM and DRAM for on-chip cache and main memory applications. However, reliability issue continues being one of the major challenges in STT-RAM memory designs due to the process variations and unique thermal fluctuations, *i.e.*, the stochastic resistance switching property of magnetic devices.

In this dissertation, I decoupled the reliability issues as following three-folds: First, the characterization of STT-RAM operation errors often require expensive Monte-Carlo runs with hybrid magnetic-CMOS simulation steps, making it impracticable for architects and system designs; Second, the state of the art does not have sufficiently understanding on the unique reliability issue of STT-RAM, and conventional *error correction codes* (ECCs) cannot efficiently handle such errors; Third, while the information density of STT-RAM can be boosted by *multi-level cell* (MLC) design, the more prominent reliability concerns and the complicated access mechanism greatly limit its applications in memory subsystems.

Thus, I present a novel through solution set to both characterize and tackle the above reliability challenges in STT-RAM designs. In the first part of the dissertation, I introduce a new characterization method that can accurately and efficiently capture the multi-variable design metrics of STT-RAM cells; Second, a novel ECC scheme, namely, *content-dependent*

ECC (CD-ECC), is developed to combat the characterized asymmetric errors of STT-RAM at 0→1 and 1→0 bit flipping's; Third, I present a circuit-architecture design, namely *state-restricted multi-level cell* (SR-MLC) STT-RAM design, which simultaneously achieves high information density, good storage reliability and fast write speed, making MLC STT-RAM accessible for system designers under current technology node. Finally, I conclude that efficient robust (or ECC) designs for STT-RAM require a deep holistic understanding on three different levels–device, circuit and architecture. Innovative ECC schemes and their architectural applications, still deserve serious research and investigation in the near future.

TABLE OF CONTENTS

1.0 INTRODUCTION . . . 1

1.1 MOTIVATION . . . 1

1.1.1 Challenge 1: Error Characterization of STT-RAM . . . 2

1.1.2 Challenge 2: Asymmetric Error Correction of SLC STT-RAM . . . 3

1.1.3 Challenge 3: High-Reliable High-Performance MLC STT-RAM Design 4

1.2 Dissertation Contribution and Outline . . . 5

2.0 STATISTICAL METHODOLOGY–PS3-RAM . . . 8

2.1 Preliminary . . . 8

2.1.1 STT-RAM Basics . . . 8

2.1.2 Operation Errors of MTJ . . . 9

2.1.2.1 Persistent errors . . . 9

2.1.2.2 Non-persistent errors . . . 10

2.2 PS3-RAM Method . . . 11

2.2.1 Sensitivity Analysis on MTJ Switching . . . 11

2.2.1.1 Threshold voltage variation . . . 11

2.2.1.2 Sensitivity analysis on variations . . . 13

2.2.1.3 Variation contribution analysis . . . 15

2.2.1.4 Simulation results of sensitivity analysis . . . 16

2.2.2 Write Current Distribution Recovery . . . 18

2.2.3 Statistical Thermal Analysis . . . 19

2.3 Application 1: Write Reliability Analysis . . . 21

2.3.1 Reliability Analysis of STT-RAM Cells . . . 21

2.3.2 Array Level Analysis and Design Optimization 24
2.4 Application 2: Write Energy Analysis 26
2.4.1 Write Energy Without Variations 26
2.4.2 PS3-RAM for Statistical Write Energy 29
2.5 Computation Complexity Evaluation 31
2.6 Appendix 32
2.6.1 Sensitivity Analysis Model Deduction 32
2.6.2 Analytic Results Summary 34
2.6.3 Validation of Analytic Results 36
2.7 Chapter 2 Summary 39
3.0 CONTENT-DEPENDENT ECC DESIGNS 40
3.1 Research Motivations 40
3.1.1 Asymmetric STT-RAM Write Errors 40
3.1.2 Related Work 41
3.2 Asymmetric Write Channel 43
3.2.1 Asymmetric Write Channel (AWC) Model 43
3.2.1.1 Parametric Asymmetric Stages (PAS) 43
3.2.1.2 Random Asymmetric Stages (RAS) 44
3.2.1.3 Construction of AWC Model 45
3.2.2 Utilization of AWC model 47
3.3 Content-dependent ECC (CD-ECC) 49
3.3.1 Typical-Corner-ECC (TCE) 49
3.3.1.1 Static Differential Coding 50
3.3.1.2 Dynamic Differential Coding 51
3.3.1.3 Typical-Corner-ECC Design 52
3.3.2 Worst-Corner-ECC 53
3.3.2.1 The Codec of Worst-Corner-ECC 54
3.3.2.2 Efficacy of Worst-Corner-ECC 55
3.4 Evaluation of CD-ECC 57
3.4.1 Reliability 57

3.4.2 Performance Overhead . . . 59
3.5 Chapter 3 Summary . . . 61
4.0 STATE-RESTRICT MLC STT-RAM DESIGNS FOR HIGH-RELIABLE HIGH-PERFORMANCE MEMORY SYSTEM . . . 62
4.1 Background and Motivation . . . 63
4.1.1 MLC STT-RAM Basics . . . 63
4.1.2 Reliability of MLC STT-RAM Cells . . . 64
4.1.2.1 Write errors of MLC STT-RAM . . . 64
4.1.2.2 Read errors of MLC STT-RAM . . . 64
4.1.2.3 Practicability of ECC schemes . . . 65
4.2 SR-MLC STT-RAM Design . . . 66
4.2.1 State Restriction (StatRes) . . . 67
4.2.1.1 Basic concept of state restriction . . . 67
4.2.1.2 Optimization of StatRes . . . 67
4.2.2 Error-pattern Removal (ErrPR) . . . 70
4.2.2.1 Basic concept of ErrPR . . . 70
4.2.2.2 Reliability evaluation of SR-MLC with ErrPR . . . 72
4.2.3 Ternary Coding (TerCode) . . . 73
4.3 State Pre-recovery (PreREC) . . . 74
4.3.1 Motivation of PreREC . . . 75
4.3.2 Design of PreREC . . . 75
4.4 Evaluation of SR-MLC STT-RAM . . . 77
4.4.1 Experiment Setup . . . 77
4.4.2 Evaluation of PreREC . . . 79
4.4.3 Performance Comparison . . . 79
4.5 Chapter 4 Summary . . . 81
5.0 CONCLUSION AND FUTURE WORK . . . 82
5.1 Dissertation Conclusion . . . 82
5.1.1 Conclusion of Chapter 2 . . . 83
5.1.2 Conclusion of Chapter 3 . . . 83

5.1.3 Conclusion of Chapter 4 84
5.2 Future Work 85
5.2.1 Facts and Observations 85
5.2.2 Multi-bit ECC Design 87
5.2.3 Non-uniform ECC Design 88
5.2.4 Architecture Investigation 89
5.3 Research Summary and Insight 89
BIBLIOGRAPHY 90

LIST OF TABLES

1 Simulation parameters and environment setting 12
2 Parameter definition 32
3 Summary of variation contribution 35
4 The configuration of the microprocessor and baseline 58
5 Delay/overhead characterization of ECC schemes 59
6 Binary-to-Ternary storage mapping 74
7 System configuration 78
8 Different configurations of STT-RAM L2 cache 78
9 Reliability comparison of mixed-line, hard-line and soft-line 87

LIST OF FIGURES

1 STT-RAM basics. (a) Parallel (low resistance). (b) Anti-parallel (high resistance). (c) 1T1J cell structure. 9
2 Overview of PS3-RAM. 12
3 The normalized contributions under different W at '1'→'0' switching. 17
4 The normalized contributions under different W at '0'→'1' switching. 17
5 Basic flow for MTJ switching current recovery. 19
6 Relative Errors of the recovered I w.r.t. the results from sensitivity analysis. 19
7 Recovered I vs. Monte-Carlo result at '1'→'0'. 20
8 Recovered I vs. Monte-Carlo result at '0'→'1'. 20
9 Write failure rate at '0'→'1' when T=300K. 22
10 Write failure rate at '1'→'0' when T=300K. 22
11 P_{WF} under different temperatures at '0'→'1'. 23
12 STT-RAM design space exploration at '0'→'1'. 23
13 Write yield with ECC's at '0'→'1', T_w=15ns. 25
14 Design space exploration at '0'→'1'. 25
15 Average Write Energy under different write pulse width when T=300K. . . . 28
16 Average Write Energy vs write pulse width under different temperature. . . . 28
17 Statistical Write Energy vs write pulse width at '1'→'0'. 30
18 Statistical Write Energy vs write pulse width at '0'→'1'. 30
19 Contributions from W. 36
20 Contributions from L. 37
21 Contributions from R. 37

22 Square partial derivatives for V_{th}. 38

23 Contributions from V_{th}. 39

24 The relationship between block level reliability P_{block} and Hamming weight W for asymmetric errors. 42

25 Overview of the proposed asymmetric write channel (AWC) model. 44

26 Step breakdowns of AWC Model. 47

27 Asymmetric error rate ratio R at different T_w. 48

28 Normalized distribution of the Hamming weight of the cache data from benchmark *mcf* and *milc*. 49

29 Simulated Hamming weight distributions comparison before and after dynamic differential coding. 52

30 Overview of typical-corner-ECC. 53

31 The simulated block error rate $(1 - P_{block})$ w.r.t. the $P_{ER,0\to1}$ 56

32 The simulated block error rate $(1 - P_{block})$ for Worst-Corner-ECCs and Hammings at $P_{ER,0\to1} = 5 \times 10^{-3}$. 56

33 Cache line error rate under different schemes. 58

34 Normalized IPC of each benchmark under different schemes. 61

35 Illustrations of (a) MTJ. (b) MLC STT-RAM cell. (c) Two-step write scheme. (d) Two-step read scheme. 63

36 Comparison of different ECCs. 66

37 Overview and optimization of StatRes. 68

38 (a) 10 error patterns of C-MLC, (b) 6 error patterns of SR-MLC, (c) 2 error patterns of SR-MLC with ErrPR, (d) Overview of ErrPR. 71

39 Error rate comparison of SR-MLC vs C-MLC cells 72

40 (a) Error patterns of the state transitions of two SR-MLC cells, (b) Error patterns mapped to the 3-bit binary data. 74

41 Overview of PreREC. 76

42 The probability for a write performed in a PreRec-done L2 cache line. 79

43 Successful rate of pre-recovery operations and the average time intervals between two consecutive reads. 80

44 Normalized IPC of each benchmarks under three different cache designs. . . . 80
45 Illustration of ORIGINAL design vs. SPLIT design structure. 86

PREFACE

This dissertation is submitted in partial fulfillment of the requirements for Wujie Wen's degree of Doctor of Philosophy in Electrical and Computer Engineering. It contains the work done from September 2011 to May 2015. My advisor is Yiran Chen, University of Pittsburgh, 2010 – present.

The work is to the best of my knowledge original, except where acknowledgement and reference are made to the previous work. There is no similar dissertation that has been submitted for any other degree at any other university.

Part of the work has been published in the conference:

1. **DAC2014**: **W. Wen**, Y. Zhang, M. Mao and Y. Chen, "State-Restrict MLC STT-RAM Designs for High-Reliable High-Performance Memory System," Design Automation Conference (DAC), Jun. 2014, pp. 1-6 **(Best Paper Award Nomination, 1 out of 42 in track, 2.4%)**.
2. **ICCAD2013**: **W. Wen**, M. Mao, X. Zhu, S. Kang, D. Wang and Y. Chen, "CD-ECC: Content-Dependent Error Correction Codes for Combating Asymmetric Nonvolatile Memory Operation Errors," International Conference on Computer Aided Design (ICCAD), Nov. 2013, pp. 1-8. **(acceptance rate: 92/354 = 26%)**.
3. **DAC2012**: **W. Wen**, Y. Zhang, Y. Chen, Y. Wang and Y. Xie, "PS3-RAM: A Fast Portable and Scalable Statistical STT-RAM Reliability Analysis Method," Design Automation Conference (DAC), Jun. 2012, pp. 1191-1196. **(acceptance rate: 168/741 = 23%)**.

4. **ASP-DAC2013**: **W. Wen**, Y. Zhang, L. Zhang and Y. Chen, "Loadsa: A Yield-Driven Top-Down Design Method for STT-RAM Array," 18th Asia and South Pacific Design Automation Conference (ASP-DAC), Jan. 2013, pp. 291-296.

5. **ISCE2014**: **W. Wen**, Y. Zhang, M. Mao and Y. Chen, "STT-RAM Reliability Enhancement through ECC and Access Scheme Optimization", International Symposium on Consumer Electronics, Jun. 2014, pp. 1-2.

6. **DAC2014**: M. Mao, **W. Wen**, Y. Zhang, H. Li and Y. Chen, "Exploration of GPGPU Register File Architecture Using Domain-wall-shift-write based Racetrack Memory," Design Automation Conference (DAC), Jun. 2014, pp. 1-6. **(acceptance rate: 174/787 = 22.1%)**.

7. **DAC2014**: E. Eken, Y. Zhang, **W. Wen**, R. Joshi, H. Li and Y. Chen, "A New Field-Assisted Access Scheme of STT-RAM with Self-Reference Capability,", Design Automation Conference (DAC), Jun. 2014, pp. 1-6.

8. **ICCAD2012**: Y. Zhang, L. Zhang, **W. Wen**, G. Sun and Y. Chen, "Multi-level Cell STT-RAM: Is It Realistic or Just a Dream?" International Conference on Computer Aided Design (ICCAD), Nov. 2012, pp. 526-532. **(acceptance rate: 82/338 = 24.3%)**.

9. **DATE2013**: J. Guo, **W. Wen**, and Y. Chen, "DA-RAID-5: A Disturb Aware Data Protection Technique for NAND Flash Storage Systems," Design, Automation & Test in Europe (DATE), Mar. 2013, pp. 380-385.

10. **ISCAS2013**: Y. Zhang, X. Bi, **W. Wen**, and Y. Chen, "STT-RAM Design Considering Probabilistic and Asymmetric MTJ Switching," IEEE International Symposium on Circuits and Systems (ISCAS), May 2013, pp. 113-116.

11. **INTERMAG2012**: Y. Zhang, **W. Wen**, and Y. Chen, "The Prospect of STT-RAM Scaling from Read ability Perspective," IEEE International Magnetics Conference (INTERMag), May. 2012, BB-03.

Part of the work has been published in journal publications:

1. **TCAD2014**: **W. Wen**, Y. Zhang, Y. Chen, Y. Wang and Y. Xie, "PS3-RAM: A Fast Portable and Scalable Statistical STT-RAM Reliability/Energy Analysis Method," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD), Nov., 2014, vol. 33, no.11, pp.1644-1656.
2. **TMAG2014**: E. Eken, Y. Zhang, **W. Wen**, R. Joshi, H. Li, and Y. Chen, "A Novel Self-reference Technique for STT-RAM Read and Write Reliability Enhancement," IEEE Transaction on Magnetics (TMAG), Nov. 2014, vol. 50, no. 11, 3401404.
3. **TMAG2012**: Y. Zhang, **W. Wen**, and Y. Chen, "The Prospect of STT-RAM Scaling from Read ability Perspective," IEEE Transaction on Magnetics (TMAG), vol. 48, no.1, Nov. 2012, pp. 3035-3038.
4. **JETC2013**: Y. Chen, W. Wong, H. Li, C.-K. Koh, Y. Zhang, and **W. Wen**, "On-chip Caches built on Multi-Level Spin-Transfer Torque RAM Cells and Its Optimizations," ACM Journal on Emerging Technologies in Computing Systems (JETC), vol. 9, no 2, article 16, May 2013.
5. **SPIN2013**: Y. Zhang, **W. Wen**, and Y. Chen, "STT-RAM Cell Design Considering MTJ Asymmetric Switching," SPIN, vol. 2, no. 3, Nov. 2013, 1240007.

ACKNOWLEDGEMENTS

I would like to acknowledge the support of my advisor, Yiran Chen, whose support made this work possible, and to 49th Design Automation Conference (DAC 2012) A. Richard Newton Scholarship, Samsung Global MRAM Innovation (SGMI 2014) Program, and National Science Foundation Project (NSF CCF-1217947) for directly providing much of the financial support. I'd like to thank Professor Yiran Chen and Professor Hai (Helen) Li for their excellent guidance during the research. Professor Yiran Chen gives me guidance of emerging nonvolatile memory designs from device modeling, circuit implementation, CAD tool development to architecture simulations and validations. Special thanks go to Professor Rami Melhem, Professor Ervin Sejdic, Professor Zhi-Hong Mao, and Professor Hai (Helen) Li for being my committee members. I also would like to thank Professor Yuan Xie from University of California at Santa Barbara, for his guidance and encouragement during my Ph.D. study.

Besides, I'd like to express my gratitude to the members from Evolutional Intelligent (EI) lab at Swanson School of Engineering, especially Mengjie Mao, Yaojun Zhang, Xiang Chen and Jie Guo, for their consistent supports during my research. Finally, I'd like to thank my wife, Shuchun Yang, the MBA student in Arizona State University (ASU) and my parents in China for their great encouragement during the whole Ph.D. research.

1.0 INTRODUCTION

1.1 MOTIVATION

In modern computer systems, the demand on memory capacity grows sharply due to the exponentially increased data processing capability. However, the technology scaling of conventional memories, such as SRAM and DRAM, is facing severe challenges like the prominent leakage power consumption and the significant degradation in device reliability. The concerns on the continuous scaling of these mainstream technologies have motivated tremendous investment to emerging memories [1, 2, 3, 4, 5, 6], including *Phase Change RAM* (PCRAM), *Magnetic RAM* (MRAM), and *Resistive RAM* (RRAM) etc..

Being one promising candidate, *spin-transfer torque random access memory* (STT-RAM) has demonstrated great potentials in embedded memory and on-chip cache designs [7, 8, 9, 10, 11] through a good combination of the non-volatility of Flash, the comparable cell density to DRAM, and the nanosecond programming time like SRAM. In the past decade, many STT-RAM test chips ranging from 4Kb to 64Mb [4] have been successfully demonstrated by major semiconductor and data storage companies [2, 12, 13, 14, 15, 16, 17]. In November 2012, Everspin started shipping 64MB STT-RAM in DDR3 DIMM format [18], commencing the commercialization era of STT-RAM. Simultaneously, Crocus unveiled thermal-assisted STT-RAM chips to store transaction data on smartphones and smartcards [19].

In STT-RAM, the data is represented as the resistance state of a *magnetic tunneling junction* (MTJ) device. The MTJ resistance state can be programmed by applying a switching current with different polarizations. Compared to the charge-based storage mechanism of conventional memories, the magnetic storage mechanism of STT-RAM shows less dependency on the device volume and hence, better scalability.

Although STT-RAM demonstrates many attractive features, reliability issue remains as one of the main challenges in STT-RAM design and greatly hinders its wide applications. Process variations, for example, induce deviations of the electrical characteristics of MOS transistors and MTJs from their nominal values, leading to read and write errors of memory [20, 21, 22]. In addition, the resistance switching mechanism of MTJs suffers from a special source of randomness–thermal fluctuation, which generates the uncertainty of the MTJ switching time. As one major difference between STT-RAM and SRAM reliability concerns, the asymmetric structure of the popular one-transistor-one-MTJ (*a.k.a.* 1T1J) STT-RAM cell results in extremely unbalanced write error rates at the bit flipping's of 0→1 and 1→0. Finally, the emergence of some advanced technologies in STT-RAM development, such as *multi-level cell* (MLC) design [23, 24], further squeezes the safety margins of the read and write operations.

To summarize, in this dissertation, the complexity of reliability issue is further decoupled as following three-folds:

1. The difficulty of STT-RAM operation error characterization;
2. The inefficiency of the popular ECCs to repair the unique STT-RAM operation errors;
3. The infeasibility of system designers to leverage the advanced technologies for high reliable and high performance applications, *e.g.* *multi-level cell* (MLC), under current technology node.

1.1.1 Challenge 1: Error Characterization of STT-RAM

As pointed out by many prior arts [9, 21, 25, 26], the unreliable write operation and high write energy are to be the major issues in STT-RAM designs. And these design metrics are significantly impacted by the prominent statistical factors of STT-RAM, including CMOS/MTJ device process variations under scaled technology and the probabilistic MTJ switching behaviors. In particular, thermal fluctuations in the magnetization process introduce uncertainty to the MTJ switching time, leading to intermittent write failures if the actual MTJ switching time is longer than the applied write pulse width.

Many studies were performed to evaluate the impacts of process variations and thermal fluctuations on STT-RAM reliability [27, 28, 29]. The general error characterization flow is the follows: First, Monte-Carlo SPICE simulations are run extensively to characterize the distribution of the MTJ switching current I during the STT-RAM write operations, by considering the device variations of both MTJ and MOS transistor; Then I samples are sent into the macro-magnetic model to obtain the MTJ switching time (τ_{th}) distributions under thermal fluctuations; Finally, the τ_{th} distributions of all I samples are merged to generate the overall MTJ switching performance distribution. A write failure happens when the applied write pulse width is shorter than the needed τ_{th}. Nonetheless, there are two limitations here: 1) The costly Monte-Carlo runs and the dependency on the macro-magnetic and SPICE simulations incur huge computation complexity of such a method, limiting the application of such a simulation method at the early stage STT-RAM design and optimization; 2) The method is simply performed on the STT-RAM cells with fixed variation configurations, which means one variation configuration one simulation, and significantly reduces its scalability and portability. Meanwhile, the modeling of write energy in STT-RAM was also studied extensively [25]. However, many such works only assume that the write energy of STT-RAM is deterministic and cannot successfully take into account its statistical characteristic induced by process variations and thermal fluctuations.

1.1.2 Challenge 2: Asymmetric Error Correction of SLC STT-RAM

Error correction code (ECC) has been proven a "must-have" technology in STT-RAM designs [30, 31, 32, 33, 34, 35, 36]. However, the uniqueness of STT-RAM designs generates many new challenges in development of ECC scheme. We do not believe that the state of the art has sufficiently deep understanding on the reliability issue of STT-RAM operations, and conventional ECCs, can efficiently handle the highly asymmetric writing errors at different bit-flipping directions. The major limitations of conventional ECCs are: 1) Unable to differentiate the asymmetric bit error rate; 2) Extremely unbalanced block reliability after coding; and 3) High cost wasted on guaranteeing few worst corner blocks. Moreover, high operational error rate in STT-RAM designs (which indeed relies on the storage patterns) de-

mands for a very strong ECC scheme. However, such strong ECC usually implies long data encoding/decoding latency, which is usually against the requirement of the delay-sensitive on-chip cache applications.

1.1.3 Challenge 3: High-Reliable High-Performance MLC STT-RAM Design

Similar to other nonvolatile memory technologies, the information density of STT-RAM can be boosted by the advanced technology–*multi-level cell* (MLC) design, e.g., stacking two MTJ devices vertically [11]. However, the reliability concern [20] and the complicated access mechanism [37] greatly limit the application of MLC STT-RAM.

Compared to *single-level cell* (SLC) design, the reliability concerns of MLC STT-RAM are mainly from two perspectives: first, MLC STT-RAM cells often have narrower distinction between resistance states, resulting in a smaller sense margin of read operations; second, MLC STT-RAM cells have a higher write error rate because of more complex failure mechanisms, *i.e.*, incomplete write or overwrite (which is new for MLC STT-RAM cells [20]) and two-step write operations. Based on [20], the read and write error rates of conventional MLC STT-RAM can be as high as 10^{-2} and 10^{-4}, respectively, which are far beyond the error correcting capability of common simple *error correction code* (ECC) like *single-error-correction-double-error-detection* (SEC-DED) [31, 38, 39]. Applying stronger ECC like *Bose-Chaudhuri-Hocquenghem* (BCH) code, however, is usually impractical for on-chip applications due to the associated high area and performance overheads.

Two-step write scheme is required in conventional MLC STT-RAM to program each digit of the 2-bit data in sequence [37]. Hence, the write access time of an MLC STT-RAM cell can be at least 2× longer than that of an SLC STT-RAM cell, resulting in considerable performance penalty [40].

1.2 DISSERTATION CONTRIBUTION AND OUTLINE

According to above three challenges, our proposed work can be also decoupled as following three main research scopes: 1) Statistical simulation approaches to characterize the write reliability and write energy under both process variations and the intrinsic randomness in the physical mechanisms (*e.g.*, thermal fluctuations); 2) New design concept based ECCs to tolerate the highly asymmetric write errors of STT-RAM; 3) A holistic circuit-architecture solution set to promote the early adoption of MLC STT-RAM in high reliable and high performance applications under current technology node.

For research scope 1, we proposed "PS3-RAM" – a fast, **p**ortable and **s**calable **s**tatistical **S**TT-RAM reliability/energy analysis method, which includes three integrated steps: 1) characterizing the MTJ switching current distribution under both MTJ and CMOS device variations; 2) recovering MTJ switching current samples from the characterized distributions in MTJ switching performance evaluation; and 3) performing the simulation on the thermal-induced MTJ switching variations based on the recovered MTJ switching current samples. Our major technical contributions of PS3-RAM are:

- We developed a sensitivity analysis technique to capture the statistical characteristics of the MTJ switching at scaled technology nodes. It achieves multiple orders-of-magnitude ($> 10^5$) run time cost reduction with marginal accuracy degradation, compared to SPICE-based Monte-Carlo simulations;
- We proposed using dual-exponential model for the fast and accurate recovery of MTJ switching current samples in statistical STT-RAM thermal analysis;
- We released PS3-RAM from SPICE and macro-magnetic modeling and simulations, and extended its application into the array-level reliability analysis and the design space exploration of STT-RAM.
- We introduced the concept of statistical write energy of STT-RAM and performed the statistical analysis on write energy by leveraging our PS3-RAM.

For research scope 2, we developed an analytical *asymmetric write channel* (AWC) model to provide a detailed step-by-step analysis to answer the questions where and how such asym-

metric write errors of STT-RAM come from. Both cell-to-cell device variations and cycle-to-cycle stochastic MTJ switching variations are considered. To address such unique errors, we carefully demonstrated the inefficiency of the traditional worst-case view based ECC design and proposed the *content-dependent ECC* (CD-ECC) by leveraging the new probabilistic ECC design view, to balance the error correcting capability at both bit-flipping directions. Two CD-ECC schemes – *typical-corner-ECC* (TCE) and *worst-corner-ECC* (WCE), are designed for the codewords with different bit-flipping distributions. The main contributions of the research scope 2 are:

- We systematically decoupled the asymmetric factors into *"parametric asymmetric stages"* (PAS) and *"random asymmetric stages"* (RAS) in AWC model, both of which are described with mathematical modeling. The AWC model can provide a quick microscopic analysis for the step-by-step accumulated asymmetry phenomena;
- We proposed CD-ECC technique to improve and balance the block-level error rate for different data patterns. Two ECC schemes – typical-corner-ECC and worst-corner-ECC, are designed for the codewords with different bit-flipping distributions;
- We evaluated the efficacy of CD-ECC technique at circuit-design and architecture levels. Our simulation results show that CD-ECC can improve STT-RAM write reliability by $10 - 30\times$ with very marginal *instruction-per-cycle* (IPC) performance degradation and low hardware overhead.

For research scope 3, we proposed an circuit-architecture co-optimization solution to address the multi-objective optimization problem of MLC STT-RAM on reliability, performance and integration density. The major contributions can be summarized as:

- We proposed a novel MLC STT-RAM design, namely, *state-restrict MLC STT-RAM* (SR-MLC STT-RAM), which can dramatically reduce the read error rate by $\sim 10^4\times$.
- We developed *error-pattern removable* (ErrPR) technique that can significantly reduce both the number of write error patterns (from 6 to 2) and write error rate of an SR-MLC cell by $\sim 10\times$.
- We developed a fast and low cost *ternary coding* (TerCode) technique to make efficient transition between binary data and the tri-state SR-MLC based storage system.

- We proposed state *pre-recovery* (PreREC) technique to virtually eliminate the costly two-step programming of SR-MLC STT-RAM. Compared to *single-level cell* (SLC) STT-RAM, SR-MLC STT-RAM based cache design can boost the system performance by 6.2% on average by leveraging the increased cache capacity at the same area and the improved write and read latency.

For future work directions, we will further focus on the reliability, performance and power issues of the promising MLC STT-RAM, for example, the low-latency and cost multi-bit ECCs may need be seriously investigated due to the increased occurrence probability of the multi-bit errors in performance-driven MLC STT-RAM designs.

The outline of this dissertation is summarized as follows: Chapter 1 presents the overall picture of this dissertation, including the research motivations, research scopes and the research contributions; Chapter 2 gives the details of the proposed fast, portable, scalable and statistical method–"PS3-RAM", as well as its applications on reliability and write energy characterization; Chapter 3 describes the developed *asymmetric write channel* (AWC) to analyze the unique asymmetric operation errors of SLC STT-RAM, as well as the corresponding customized ECC design (CD-ECC) to tolerate such errors; Chapter 4 demonstrates the benefits of our proposed circuit architecture solution–SR-MLC, to provide intelligent balance between performance, reliability and density for MLC STT-RAM based storage system under current technology node. Chapter 5 finally summarizes the research work and presents the potential future research directions, as well as our insights for robust (or ECC) designs of emerging nonvolatile memories.

2.0 STATISTICAL METHODOLOGY–PS3-RAM

In this chapter, we will present the details of our error characterization methodology–PS3-RAM. The structure of this chapter is organized as the follows: Section 2.1 gives the preliminary of STT-RAM; Section 2.2 presents the details of PS3-RAM method; Section 2.3 presents the application of our PS3-RAM on cell and array level reliability analysis and design space exploration; Section 2.4 shows the deterministic/statistical write energy analysis based on our PS3-RAM; Section 2.5 discusses the computation complexity; Section 2.6 gives the detailed theatrical model deduction and its numerical validation for sensitivity analysis; Section 2.7 concludes this chapter.

2.1 PRELIMINARY

2.1.1 STT-RAM Basics

Fig. 1(c) shows the popular "*one-transistor-one-MTJ* (1T1J)" STT-RAM cell structure, which includes a MTJ and a NMOS transistor connected in series. In the MTJ, an oxide barrier layer (*e.g.*, MgO) is sandwiched between two ferromagnetic layers. '0' and '1' are stored as the different resistances of the MTJ, respectively. When the magnetization directions of two ferromagnetic layers are parallel (anti-parallel), the MTJ is in its low (high) resistance state. Fig. 1(a) and (b) shows the low and the high MTJ resistance states, which are denoted by R_L and R_H, respectively. The MTJ switches from '0' to '1' when the switching current drives from reference layer to free layer, or from '1' to '0' when the switching current drives in the opposite direction.

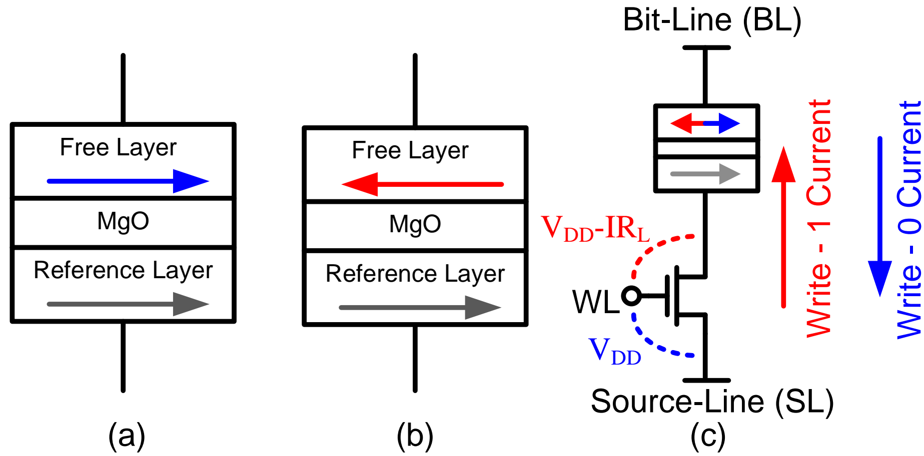

Figure 1: STT-RAM basics. (a) Parallel (low resistance). (b) Anti-parallel (high resistance). (c) 1T1J cell structure.

2.1.2 Operation Errors of MTJ

In general, the MTJ switching time decreases when the switching current increases. A write failure happens when the MTJ switching does not complete before the switching current is removed. There are two reasons can cause this failure:

2.1.2.1 Persistent errors The current through the MTJ is affected by the process variations of both transistor and MTJ. For example, the driving ability of the NMOS transistor is subject to the variations of transistor channel length (L), width (W), and threshold voltage (V_{th}). The MTJ resistance variation also affects the NMOS transistor driving ability by changing its bias condition. The degraded MTJ switching current leads to a longer MTJ switching time and consequently, results in an incomplete MTJ switching before the write pulse ends. This kind of errors is referred to as "persistent" errors, which are mainly incurred by only device parametric variations. Persistent errors can be measured and repeated after the chip is fabricated.

2.1.2.2 Non-persistent errors Another kind of errors is called "non-persistent" errors, which happen intermittently and may not be repeated. The non-persistent errors of STT-RAM are mostly caused by the intrinsic thermal fluctuations during MTJ switching [41]. In general, the impact of thermal fluctuations can be modeled by the thermal induced random field h_{fluc} in stochastic *Landau-Lifshitz-Gilbert* (LLG) equation (Eq. 2.1) [42, 43, 44] as

$$\frac{d\overrightarrow{m}}{dt} = -\overrightarrow{m} \times (\overrightarrow{h}_{eff} + \overrightarrow{h}_{fluc}) + \alpha\overrightarrow{m} \times (\overrightarrow{m} \times (\overrightarrow{h}_{eff} + \overrightarrow{h}_{fluc})) + \frac{\overrightarrow{T}_{norm}}{M_s} \tag{2.1}$$

Where $\overrightarrow{m}$ is the normalized magnetization vector. Time t is normalized by γM_s; γ is the gyro-magnetic ratio and M_s is the magnetization saturation. $\overrightarrow{h}_{eff} = \frac{\overrightarrow{H_{eff}}}{M_s}$ is the normalized effective magnetic field. $\overrightarrow{h}_{fluc}$ is the normalized thermal agitation fluctuating field at finite temperature which represent the thermal fluctuation. α is the LLG damping parameter. $\overrightarrow{T}_{norm} = \frac{\overrightarrow{T}}{M_s V}$ is the spin torque term with units of magnetic field. And the net spin torque $\overrightarrow{T}$ can be obtained through microscopic quantum electronic spin transport model. Due to thermal fluctuations, the MTJ switching time will not be a constant value but rather a distribution even under a constant switching current.

2.2 PS3-RAM METHOD

Fig. 2 depicts the overview of our proposed PS3-RAM method, mainly including the sensitivity analysis for MTJ switching current (I) characterization, the I sample recovery, and the statistical thermal analysis of STT-RAM. The first step is to configure the variation-aware cell library by inputting both the nominal design parameters and their corresponding variations, like the channel length/width/threshold voltage of NMOS transistor, as well as the thickness/area of MTJ device. Then a multi-dimension sensitivity analysis will be conducted to characterize the statistical properties of I, followed by an advanced filtering technology – smooth filter, to improve its accuracy. After that, the write current samples can be recovered based on the above characterized statistics and current distribution model. The write pulse distribution will be generated after mapping the switching current samples to the write pulse samples by considering the thermal fluctuations. Finally, the statistical write energy analysis and the STT-RAM cell write error rate can be performed based on the samples of the write current once the write pulse is determined. Array-level analysis and design optimizations can be also conducted by using PS3-RAM.

2.2.1 Sensitivity Analysis on MTJ Switching

In this section, we present our sensitivity model used for the characterization of the MTJ switching current distribution. We then analyze the contributions of different variation sources to the distribution of the MTJ switching current in details. The definitions of the variables used in our analysis are summarized in TABLE 1.

2.2.1.1 Threshold voltage variation The variations of channel length, width and threshold voltage are three major factors causing the variations of transistor driving ability. V_{th} variation mainly comes from *random dopant fluctuation* (RDF) and *line-edge roughness* (LER), the latter of which is also the source of some geometry variations (i.e., L and W) [45, 46]. It is known that the V_{th} variation is also correlated with L and W and its variance decreases when the transistor size increases.

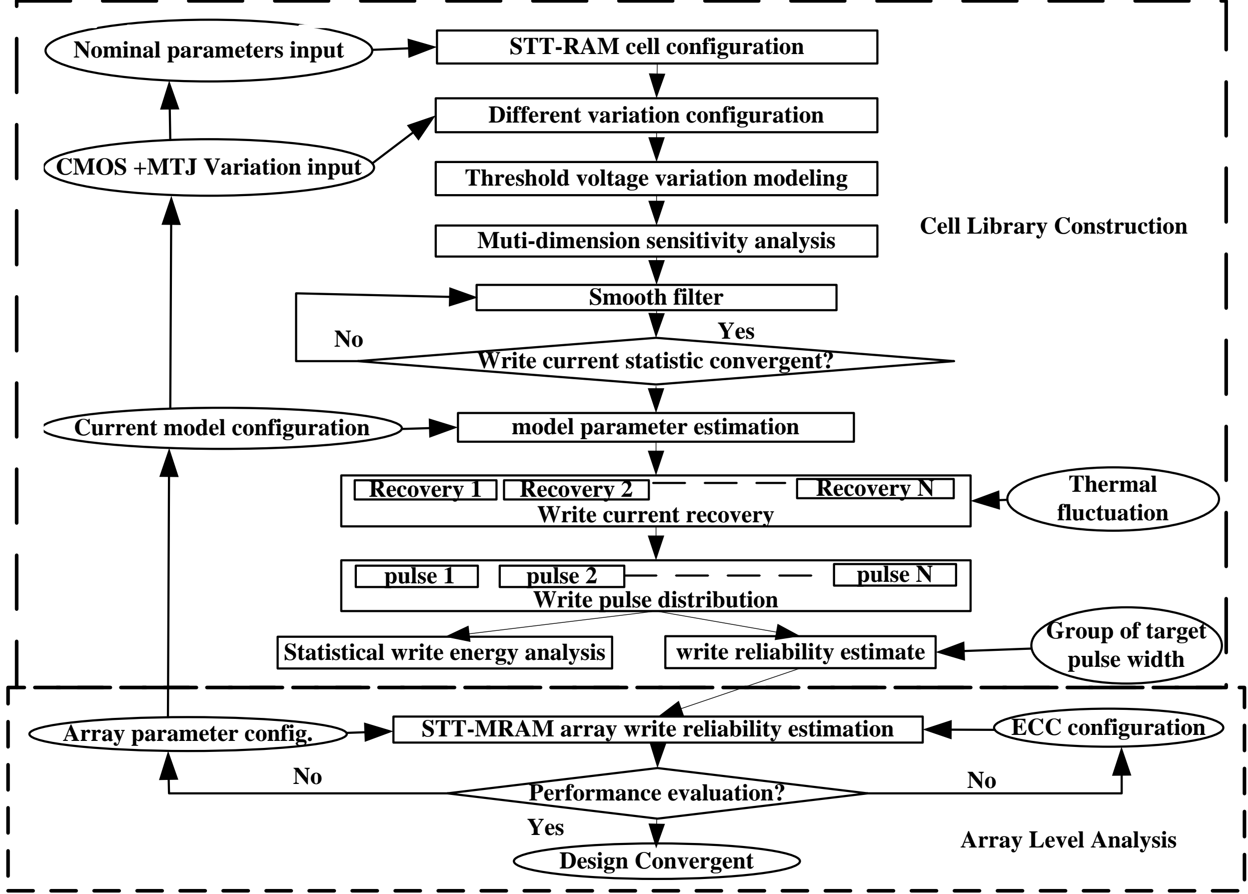

Figure 2: Overview of PS3-RAM.

Table 1: Simulation parameters and environment setting

Parameters	Mean	Standard Deviation
Channel length	$\overline{L} = 45$nm	$\sigma_L = 0.05\overline{L}$
Channel width	$\overline{W} = 90 \sim 1800$nm	$\sigma_W = 0.05\overline{L}$
Threshold voltage	$\overline{V}_{th} = 0.466$V	by calucaltion
Mgo thickness	$\overline{\tau} = 2.2$nm	$\sigma_\tau = 0.02\overline{\tau}$
MTJ surface area	$\overline{A} = 45 \times 90\text{nm}^2$	by calculation
Resistance low	$R_L = 1000\Omega$	by calculation
Resistance high	$R_H = 2000\Omega$	by calculation

The deviation of the V_{th} from the nominal value following the change of L (ΔL) can be modeled by [46]:

$$\Delta V_{th} = \Delta V_{th0} + V_{ds} exp(-\frac{L}{l'}) \cdot \frac{\Delta L}{l'}. \tag{2.2}$$

Then the standard deviation of V_{th} can be calculated as:

$$\sigma^2_{V_{th}} = \frac{C_1}{WL} + \frac{C_2}{exp\left(L/l'\right)} \cdot \frac{W_c}{W} \cdot \sigma^2_L. \tag{2.3}$$

Here W_c is the correlation length of *non-rectangular gate* (NRG) effect, which is caused by the randomness in sub-wavelength lithography. C_1, C_2 and l' are technology dependent coefficients. The first term in Eq. (2.3) describes the RDF's contribution to $\sigma_{V_{th}}$. The second term in Eq. (2.3) represents the contribution from NRG, which is heavily dependent on L and W. Following technology scaling, the contribution of this term becomes prominent due to the reduction of L and W.

2.2.1.2 Sensitivity analysis on variations Although the contributions of MTJ and MOS transistor parametric variabilities to the MTJ switching current distribution cannot be explicitly expressed, it is still possible for us to conduct a sensitivity analysis to obtain the critical characteristics of the distribution. Without loss of generality, the MTJ switching current I can be modeled by a function of W, L, V_{th}, A, and τ. A and τ are the MTJ surface area and MgO layer thickness, respectively. The 1st-order Taylor expansion of I around the mean values of every parameter is:

$$\begin{aligned} I\left(W, L, v_{th}, A, \tau\right) &\approx I\left(\overline{W}, \bar{L}, \bar{V}_{th}, \bar{A}, \bar{\tau}\right) + \frac{\partial I}{\partial W}\left(W - \overline{W}\right) \\ &+ \frac{\partial I}{\partial L}\left(L - \bar{L}\right) + \frac{\partial I}{\partial V_{th}}\left(V_{th} - \bar{V}_{th}\right) \\ &+ \frac{\partial I}{\partial A}\left(A - \bar{A}\right) + \frac{\partial I}{\partial \tau}\left(\tau - \bar{\tau}\right). \end{aligned} \tag{2.4}$$

Here W, L and τ generally follow Gaussian distribution [27], A is the product of two independent Gaussian distributions, V_{th} is correlated with W, L, as shown in Eq. (2.2) and (2.3). Because the MTJ resistance $R \propto \frac{e^{\tau}}{A}$ [27], we have:

$$\begin{aligned} \frac{\partial I}{\partial A}\Delta A + \frac{\partial I}{\partial \tau}\Delta \tau &= \frac{\partial I}{\partial R}\left(\frac{\partial R}{\partial A}\Delta A + \frac{\partial R}{\partial \tau}\Delta \tau\right) \\ &= \frac{\partial I}{\partial R}\Delta R. \end{aligned} \tag{2.5}$$

Eq. (2.5) indicates that the combined contribution of A and τ is the same as the impact of MTJ resistance. The difference between the actual I and its mathematical expectation μ_I can be calculated by:

$$I\left(W, L, V_{th}, R\right) - E\left(I\left(\overline{W}, \bar{L}, \bar{V}_{th}, \overline{R}\right)\right) \approx \frac{\partial I}{\partial W}\Delta W + \frac{\partial I}{\partial L}\Delta L + \frac{\partial I}{\partial V_{th}}\Delta V_{th} + \frac{\partial I}{\partial R}\Delta R. \tag{2.6}$$

Here we assume $\mu_I \approx E\left(I\left(\overline{W}, \bar{L}, \bar{V}_{th}, \overline{R}\right)\right) = I\left(\overline{W}, \bar{L}, \bar{V}_{th}, \overline{R}\right)$ and the mean of MTJ resistance $\overline{R} \approx R\left(\bar{A}, \bar{\tau}\right)$. Combining Eq. (2.2), (2.3), and (2.6), the standard deviation of I (σ_I) can be calculated as:

$$\begin{aligned}
\sigma_I^2 &= \left(\frac{\partial I}{\partial W}\right)^2 \sigma_W^2 + \left(\frac{\partial I}{\partial L}\right)^2 \sigma_L^2 + \left(\frac{\partial I}{\partial R}\right)^2 \sigma_R^2 \\
&+ \left(\frac{\partial I}{\partial V_{th}}\right)^2 \left(\frac{C_1}{WL} + \frac{C_2}{exp\left(L/l'\right)} \cdot \frac{W_c}{W} \cdot \sigma_L^2\right) \\
&+ 2\frac{\partial I}{\partial L}\frac{\partial I}{\partial V_{th}}\rho_1\sqrt{\frac{C_1}{WL}}\sigma_L + 2\frac{\partial I}{\partial W}\frac{\partial I}{\partial V_{th}}\rho_2\sqrt{\frac{C_1}{WL}}\sigma_W \\
&+ 2\frac{\partial I}{\partial L}\frac{\partial I}{\partial V_{th}}V_{ds}exp(-\frac{L}{l'})\frac{\sigma_L^2}{l'}.
\end{aligned} \tag{2.7}$$

Here $\rho_1 = \frac{\text{cov}(V_{th0}, L)}{\sqrt{\sigma_{v_{th0}}^2 \sigma_L^2}}$ and $\rho_2 = \frac{\text{cov}(V_{th0}, W)}{\sqrt{\sigma_{V_{th0}}^2 \sigma_W^2}}$ are the correlation coefficients between V_{th0} and L or W, respectively [46]. $\sigma_{V_{th0}}^2 = \frac{C_1}{WL}$. Our further analysis shows that the last three terms at the right side of Eq. (2.7) are significantly smaller than other terms and can be safely ignored in the simulations of STT-RAM normal operations.

The accuracy of the coefficient in front of the variances of every parameter at the right side of Eq. (2.7) can be improved by applying window based smooth filtering. Take W as an example, we have:

$$\left(\frac{\partial I}{\partial W}\right)_i = \frac{I\left(\overline{W} + i\Delta W, L, V_{th}, R\right) - I\left(\overline{W} - i\Delta W, L, V_{th}, R\right)}{2i\Delta W}, \tag{2.8}$$

where $i = 1, 2, ...K$. Different $\frac{\partial I}{\partial W}$ can be obtained at the different step i. K samples can be filtered out by a windows based smooth filter to balance the accuracy and the computation complexity as:

$$\overline{\frac{\partial I}{\partial W}} = \sum_{i=1}^{K} \omega_i \left(\frac{\partial I}{\partial W}\right)_i. \tag{2.9}$$

Here ω_i is the weight of sample i, which is determined by the window type, i.e., Hamming window or Rectangular window [47].

2.2.1.3 Variation contribution analysis The variations' contributions to I are mainly represented by the first four terms at the right side of Eq (2.7) as:

$$S_1 = \left(\frac{\partial I}{\partial W}\right)^2 \sigma_W^2, S_2 = \left(\frac{\partial I}{\partial L}\right)^2 \sigma_L^2, S_3 = \left(\frac{\partial I}{\partial R}\right)^2 \sigma_R^2$$
$$S_4 = \left(\frac{\partial I}{\partial V_{th}}\right)^2 \left(\frac{C_1}{WL} + \frac{C_2}{exp\left(L/l'\right)} \cdot \frac{W_c}{W} \cdot \sigma_L^2\right). \quad (2.10)$$

As pointed out by many prior-arts [36, 48, 49], an asymmetry exists in STT-RAM write operations: the switching time of '0'→'1' is longer than that of '1'→'0' and suffers from a larger variance. Also, the switching time variance of '0'→'1' is more sensitive to the transistor size changes than '1'→'0'. As we shall show later, this phenomena can be well explained by using our sensitivity analysis. To the best of our knowledge, this is the first time the asymmetric variations of STT-RAM write performance and their dependencies on the transistor size are explained and quantitatively analyzed.

As shown in Fig. 1, when writing '0', the *word-line* (WL) and *bit-line* (BL) are connected to V_{dd} while the *source-line* (SL) is connected to ground. $V_{gs} = V_{dd}$ and $V_{ds} = V_{dd} - IR$. The NMOS transistor is mainly working in triode region. Based on short-channel BSIM model, the MTJ switching current supplied by a NMOS transistor can be calculated by:

$$I = \frac{\beta \cdot \left[(V_{dd} - V_{th})(V_{dd} - IR) - \frac{a}{2}(V_{dd} - IR)^2\right]}{1 + \frac{1}{v_{sat}L}(V_{dd} - IR)}. \quad (2.11)$$

Here $\beta = \frac{\mu_0 C_{ox}}{1+U_0(V_{dd}-V_{th})}\frac{W}{L}$. U_0 is the vertical field mobility reduction coefficient, μ_0 is electron mobility, C_{ox} is gate oxide capacitance per unit area, a is body-effect coefficient and v_{sat} is carrier velocity saturation. Based on short-channel PTM model [50] and BSIM model [51, 52], we derive $\left(\frac{\partial I}{\partial W}\right)^2$, $\left(\frac{\partial I}{\partial L}\right)^2$, $\left(\frac{\partial I}{\partial R}\right)^2$, and $\left(\frac{\partial I}{\partial V_{th}}\right)^2$ as:

$$\left(\frac{\partial I}{\partial W}\right)_0^2 \approx \frac{1}{(A_1 W + B_1)^4}, \quad \left(\frac{\partial I}{\partial L}\right)_0^2 \approx \frac{1}{\left(\frac{A_2}{W} + B_2 W + C\right)^2}$$
$$\left(\frac{\partial I}{\partial R}\right)_0^2 \approx \frac{1}{\left(\frac{A_3}{W} + B_3\right)^4}, \quad \left(\frac{\partial I}{\partial V_{th}}\right)_0^2 \approx \frac{1}{\left(\frac{A_4}{\sqrt{W}} + B_4\sqrt{W}\right)^4}.$$

Our analytical deduction shows that the coefficients A_{1-4}, B_{1-4} and C are solely determined by W, L, V_{th}, and R. The detailed expressions of coefficients A_{1-4}, B_{1-4} and C can be found in the appendix. Here R is the high resistance state of the MTJ, or R_H. For a NMOS transistor at '0'→'1' switching, the MTJ switching current is:

$$I = \frac{\beta}{2a}\left[(V_{dd} - IR - V_{th}) - \frac{I}{WC_{ox}v_{sat}^2}\right]^2. \tag{2.12}$$

Here R is the low resistance state of the MTJ, or R_L. We have:

$$\left(\frac{\partial I}{\partial W}\right)_1^2 \approx \frac{1}{(A_5W + B_5)^4}, \quad \left(\frac{\partial I}{\partial L}\right)_1^2 \approx \frac{1}{\left(\frac{A_6}{W} + B_6\right)^2}$$
$$\left(\frac{\partial I}{\partial R}\right)_1^2 \approx \frac{1}{\left(\frac{A_7}{W} + B_7\right)^4}, \quad \left(\frac{\partial I}{\partial V_{th}}\right)_1^2 \approx \frac{1}{\left(\frac{A_8}{W} + B_8\right)^2}$$

Again, A_{5-8} and B_{5-8} can be expressed as the function of W, L, V_{th}, and R and the detailed expressions of those parameters can be found in the appendix.

In general, a large S_i corresponds to a large contribution to I variation. When W is approaching infinity, only S_3 is nonzero at '1'→'0' switching while both S_2 and S_3 are nonzero at '0'→'1' switching. It indicates that the residual values of S_1–S_4 at '0'→'1' switching is larger than that at '1'→'0' switching when $W \to \infty$. In other words, '0'→'1' switching suffers from a larger MTJ switching current variation than '1'→'0' switching when NMOS transistor size is large.

2.2.1.4 Simulation results of sensitivity analysis Sensitivity analysis [53] can be used to obtain the statistical parameters of MTJ switching current, i.e., the mean and the standard deviation, without running the costly SPICE and Monte-Carlo simulations. It can be also used to analyze the contributions of different variation sources to I variation in details. The normalized contributions (P_i) of variation resources, i.e., W, L, V_{th}, and R, are defined as:

$$P_i = \frac{S_i}{\sum_{i=1}^{4} S_i}, i = 1, 2, 3, 4 \tag{2.13}$$

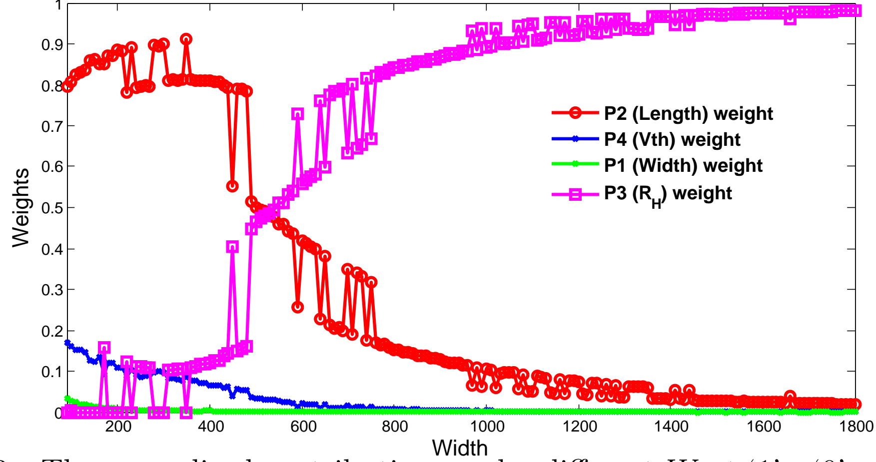

Figure 3: The normalized contributions under different W at '1'→'0' switching.

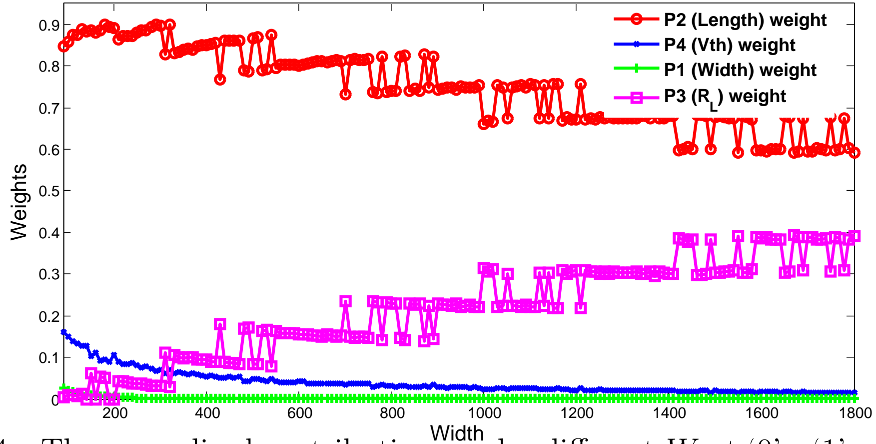

Figure 4: The normalized contributions under different W at '0'→'1' switching.

Fig. 3 and Fig. 4 show the normalized contributions of every variation source at '0'→'1' and '1'→'0' switching's, respectively, at different transistor sizes. We can see that L and V_{th} are the first two major contributors to I variation at both switching directions when W is small. At '1'→'0' switching, the contribution of L raises until reaching its maximum value when W increases, and then quickly decreases when W further increases. At '0'→'1' switching, however, the contribution of L monotonically decreases, but keeps being the dominant factor over the simulated W range. At both switching directions, the contributions of R ramps up when W increases. At '1'→'0' switching, the normalized contribution of R becomes almost 100% when W is really large.

2.2.2 Write Current Distribution Recovery

After the I distribution is characterized by the sensitivity analysis, the next question becomes how to recover the distribution of I from the characterized information in the statistical analysis of STT-RAM reliability. We investigate the typical distributions of I in various STT-RAM cell designs and found that dual-exponential function can provide the excellent accuracy in modeling and recovering these distributions. The dual-exponential function we used to recover the I distributions can be illustrated as:

$$f(I) = \begin{cases} a_1 e^{b_1(I-u)} & I \le u, \\ a_2 e^{b_2(u-I)} & I > u. \end{cases} \tag{2.14}$$

Here a_1, b_1, a_2, b_2 and u are the fitting parameters, which can be calculated by matching the first and the second order momentums of the actual I distribution and the dual-exponential function as:

$$\begin{aligned} &\int f(I)dI = 1, \\ &\int I f(I)dI = E(I), \\ &\int I^2 f(I)dI = E(I)^2 + \sigma_I^2. \end{aligned} \tag{2.15}$$

Here $E(I)$ and σ_I^2 are obtained from the sensitivity analysis.

The recovered I distribution can be used to generate the MTJ switching current samples, as shown in Fig. 5. At the beginning of the sample generation flow, the confidence interval for STT-RAM design is determined, *e.g.*, $[\mu_I - 6\sigma_I, \mu_I + 6\sigma_I]$ for a six-sigma confidence interval. Assuming we need to generate N samples within the confidence interval, say, at the point of $I = I_i$, a switching current sequence of $[N\mathrm{Pr}_i]$ samples must be generated. Here $\mathrm{Pr}_i \approx f(I_i)\Delta$. Δ equals $\frac{12\sigma_I}{N}$, or the step of sampling generation. $f(I_i)$ is the dual-exponential function.

Fig. 6 shows the relative errors of the mean and the standard deviation of the recovered I distribution *w.r.t.* the results directly from the sensitivity analysis (as Eq. (2.6) and (2.7) show). The maximum relative error $< 10^{-2}$, which proves the accuracy of our dual-exponential model.

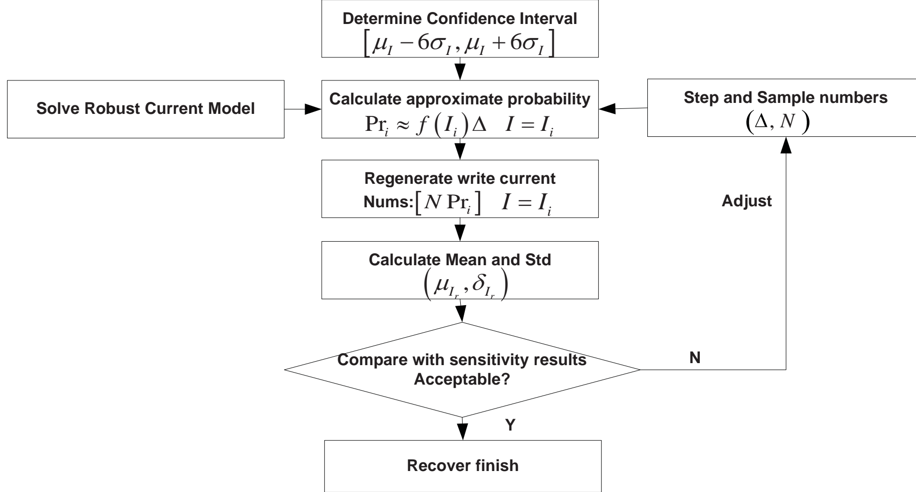

Figure 5: Basic flow for MTJ switching current recovery.

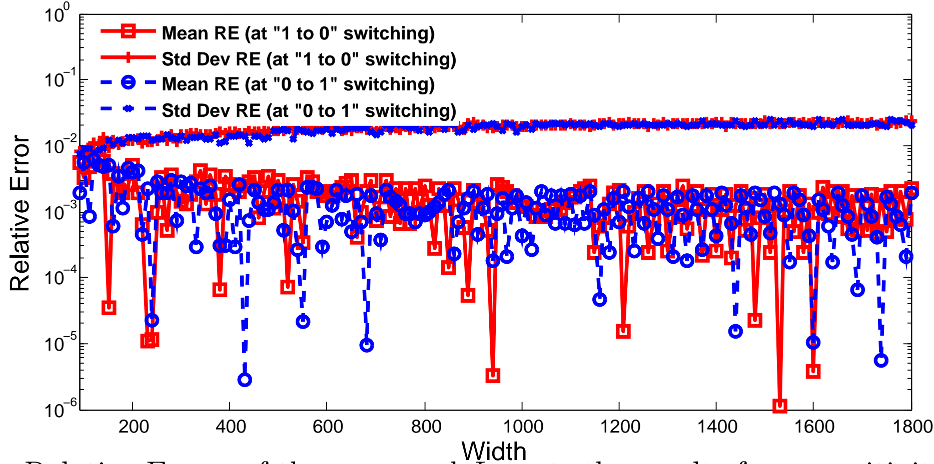

Figure 6: Relative Errors of the recovered I w.r.t. the results from sensitivity analysis.

Fig. 7 and Fig. 8 compare the probability distribution functions (PDF's) of I from the SPICE Monte-Carlo simulations and from the recovery process based on our sensitivity analysis at two switching directions. Our method achieves good accuracy at both representative transistor channel widths (W = 90nm or W = 720nm).

2.2.3 Statistical Thermal Analysis

The variation of the MTJ switching time (τ_{th}) incurred by the thermal fluctuations follows Gaussian distribution when τ_{th} is below 10∼20ns [48]. In this range, the distribution of

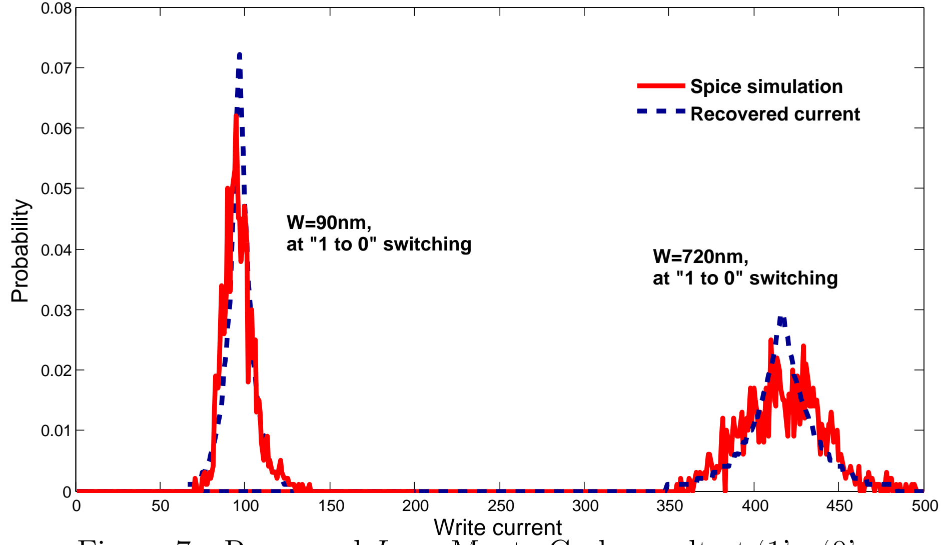

Figure 7: Recovered I vs. Monte-Carlo result at '1'→'0'.

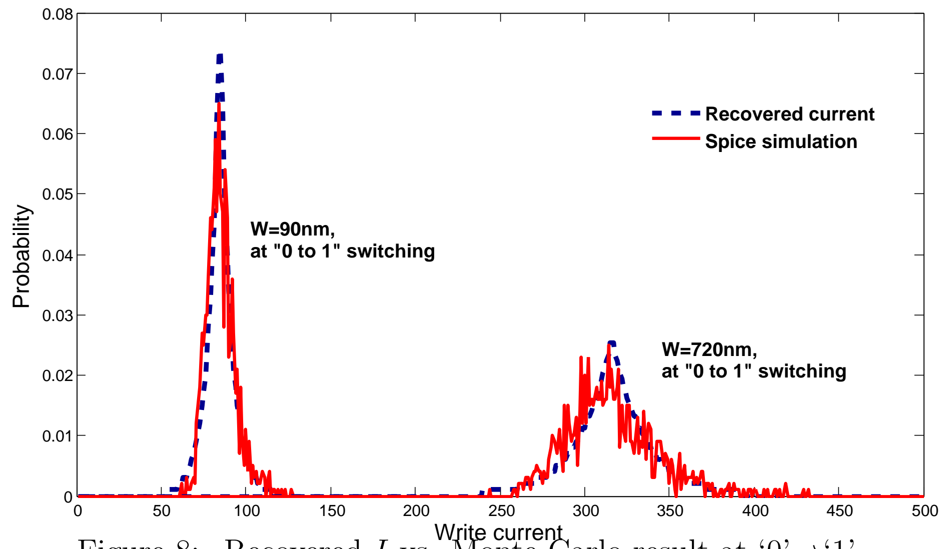

Figure 8: Recovered I vs. Monte-Carlo result at '0'→'1'.

τ_{th} can be easily constructed after the I is determined. The distribution of MTJ switching performance can be obtained by combining the τ_{th} distributions of all I samples.

2.3 APPLICATION 1: WRITE RELIABILITY ANALYSIS

In this section, we conduct the statistical analysis on the write reliability of STT-RAM cells by leveraging our PS3-RAM method. Both device variations and thermal fluctuations are considered in the analysis. We also extend our method into array-level evaluation and demonstrate its effectiveness in STT-RAM design optimizations.

2.3.1 Reliability Analysis of STT-RAM Cells

The write failure rate P_{WF} of a STT-RAM cell can be defined as the probability that the actual MTJ switching time τ_{th} is longer than the write pulse width T_w, or $P_{WF} = P\left(\tau_{th} > T_w\right)$. τ_{th} is affected by the MTJ switching current magnitude, the MTJ and MOS device variations, the MTJ switching direction, and the thermal fluctuations. The conventional simulation of P_{WF} requires costly Monte-Carlo runs with hybrid SPICE and macro-magnetic modeling steps. Instead, we can use PS3-RAM to analyze the statistical STT-RAM write performance. The corresponding simulation environment is also summarized in TABLE 1.

Fig. 9 and 10 depict the P_{WF}'s simulated by PS3-RAM for both switching directions at 300K. For comparison purpose, the Monte-Carlo simulation results are also presented. Different T_w's are selected at either switching directions due to the asymmetric MTJ switching performances [48], i.e., $T_w = 10, 15, 20ns$ at '0'→'1' and $T_w = 6, 8, 10, 12ns$ at '1'→'0'. Our PS3-RAM results are in excellent agreement with the ones from Monte-Carlo simulations.

Since '0'→'1' is the limiting switching direction for STT-RAM reliability, we also compare the P_{WF}'s of different STT-RAM cell designs under different temperatures at this switching direction in Fig. 11. The results show that PS3-RAM can provide very close but pessimistic results compared to those of the conventional simulations. PS3-RAM is also capable to precisely capture the small error rate change incurred by a moderate temperature shift (from T=300K to T=325K).

It is known that prolonging the write pulse width and increasing the MTJ switching current (by sizing up the NMOS transistor) can reduce the P_{WF}. In Fig. 12, we demonstrate an example of using PS3-RAM to explore the STT-RAM design space: the tradeoff curves

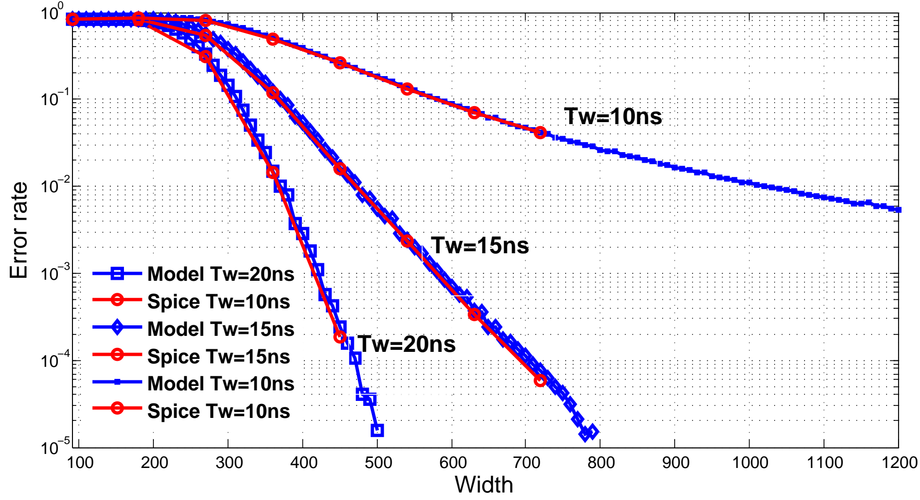

Figure 9: Write failure rate at '0'→'1' when T=300K.

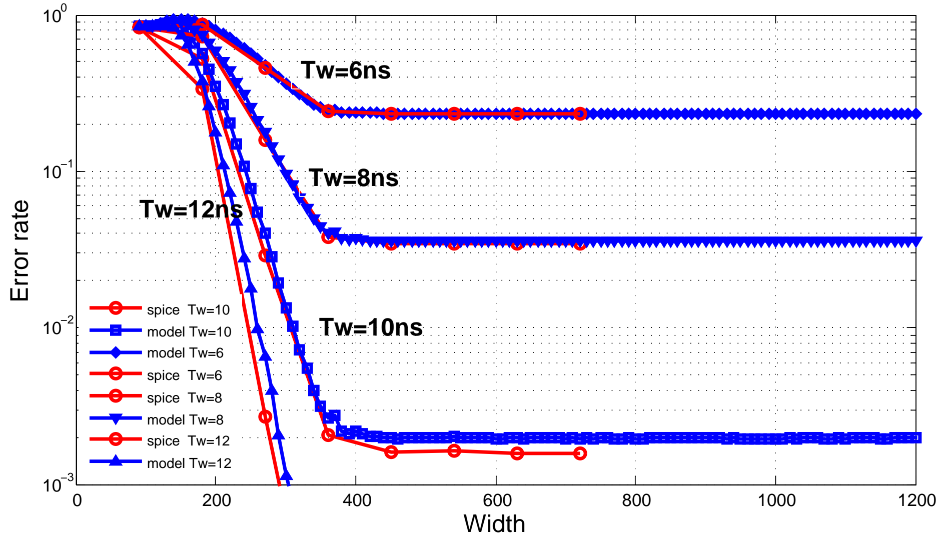

Figure 10: Write failure rate at '1'→'0' when T=300K.

between P_{WF} and T_w are simulated at different W's. For a given P_{WF}, for example, the corresponding tradeoff between W and T_w can be easily identified on Fig. 12.

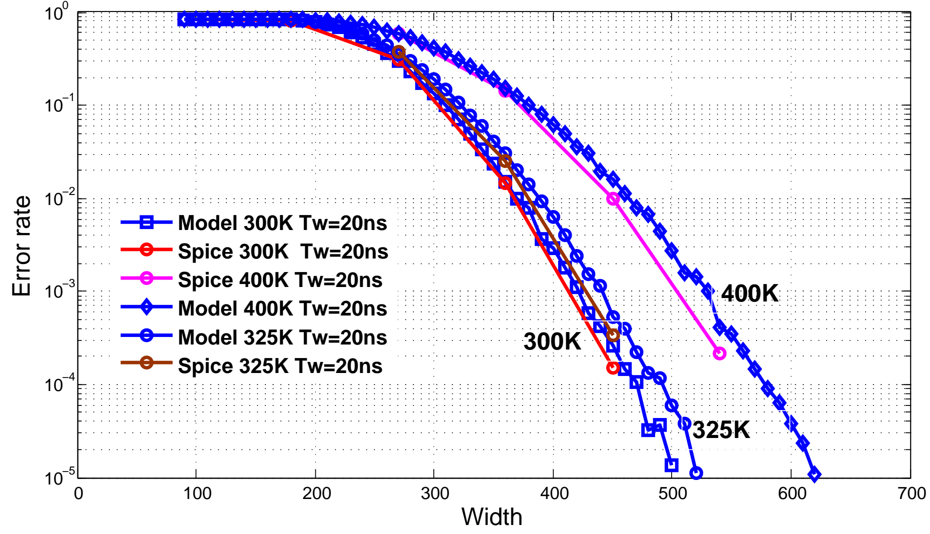

Figure 11: P_{WF} under different temperatures at '0'→'1'.

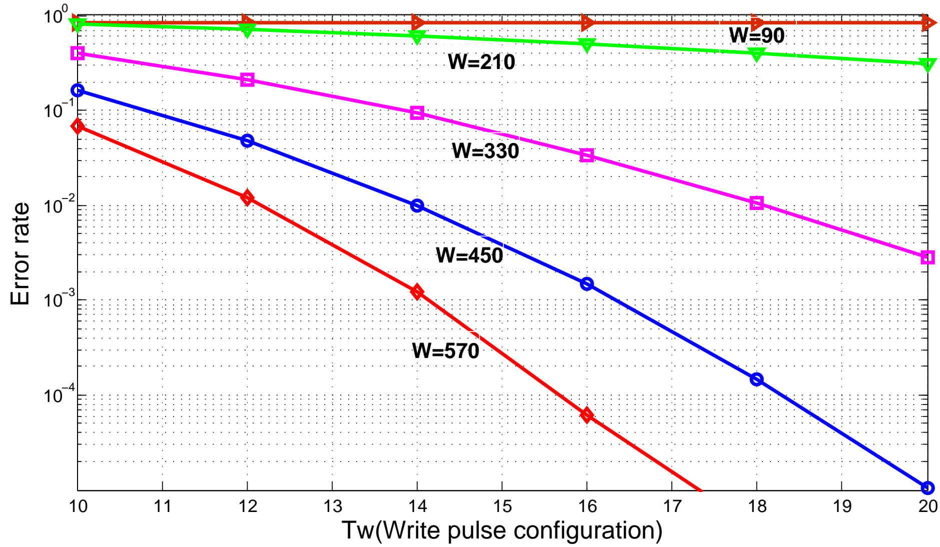

Figure 12: STT-RAM design space exploration at '0'→'1'.

2.3.2 Array Level Analysis and Design Optimization

We use a 45nm 256Mb STT-RAM design [39] as the example to demonstrate how to extend our PS3-RAM into array-level analysis and design optimizations. The number of bits per memory block $N_{bit} = 256$ and the number of memory blocks $N_{word} = 1\text{M}$. ECC (error correction code) is applied to correct the random write failures of memory cells. Two types of ECC's with different implementation costs are being considered, i.e., single-bit-correcting Hamming code and a set of multi-bits-correcting BCH codes. We use (n, k, t) to denote an ECC with n codeword length, k bit user bits being protected (256 bit here) and t bits being corrected. The ECC's corresponding to the error correction capability t from 1 to 5 are Hamming code $(265, 256, 1)$ and four BCH codes – BCH1 $(274, 256, 2)$, BCH2 $(283, 256, 3)$, BCH3 $(292, 256, 4)$ and BCH4 $(301, 256, 5)$, respectively. The write yield of the memory array Y_{wr} can be defined as:

$$Y_{wr} = P\left(n_e \leq t\right) = \sum_{i=0}^{t} C_n^i P_{WF}^i (1 - P_{WF})^{n-i}. \tag{2.16}$$

Here, n_e denotes the total number of error bits in a write access. Y_{wr} indeed denotes the probability that the number of error bits in a write access is smaller than that of the error correction code can fix.

Fig. 13 depicts the Y_{wr}'s under different combinations of ECC scheme and W when $T_w = 15\text{ns}$ at '0'→'1' switching. The ECC schmes required to satisfy $\sim 100\%$ Y_{wr} for different W are: (1) Hamming code for $W = 630\text{nm}$; (2) BCH2 for $W = 540\text{nm}$; and (3) BCH4 for $W = 480\text{nm}$. The total memory array area can be estimated by using the STT-RAM cell size equation $\text{Area}_{\text{cell}} = 3\left(W/L + 1\right)\left(F^2\right)$ [54]. Calculation shows that combination (3) offers us the smallest STT-RAM array area, which is only 88% and 95% of the ones of (1) and (2), respectively. We note that PS3-RAM can be seamlessly embedded into the existing deterministic memory macro models [54] for the extended capability on the statistical reliability analysis and the multi-dimensional design optimizations on area, yield, performance and energy.

Fig. 14 illustrates the STT-RAM design space in terms of the combinations of Y_{wr}, W, T_{sw} and ECC scheme. After the pair of (Y_{wr}, T_w) is determined, the tradeoff between W

and ECC can be found in the corresponding region on the figure. The result shows that PS3-RAM provides a fast and efficient method to perform the device/circuit/architecture co-optimization for STT-RAM designs.

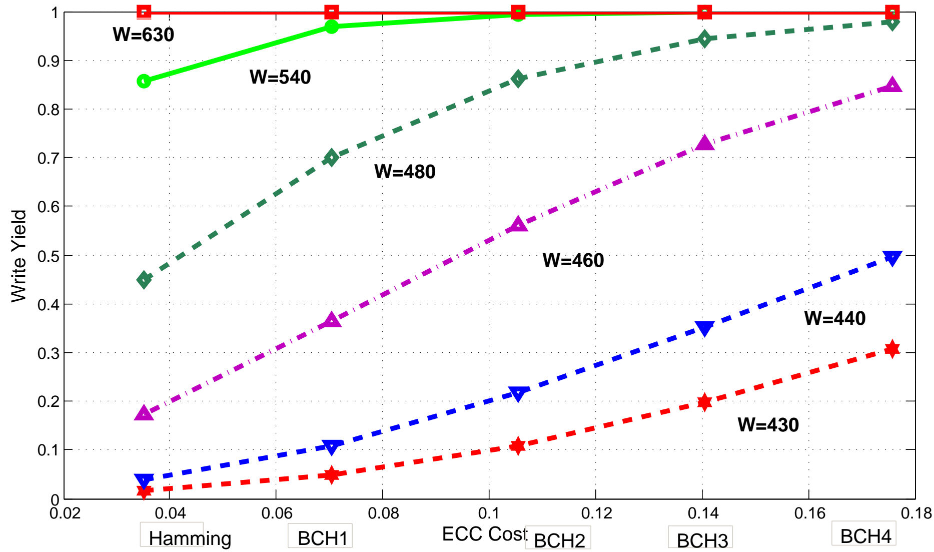

Figure 13: Write yield with ECC's at '0'→'1', T_w=15ns.

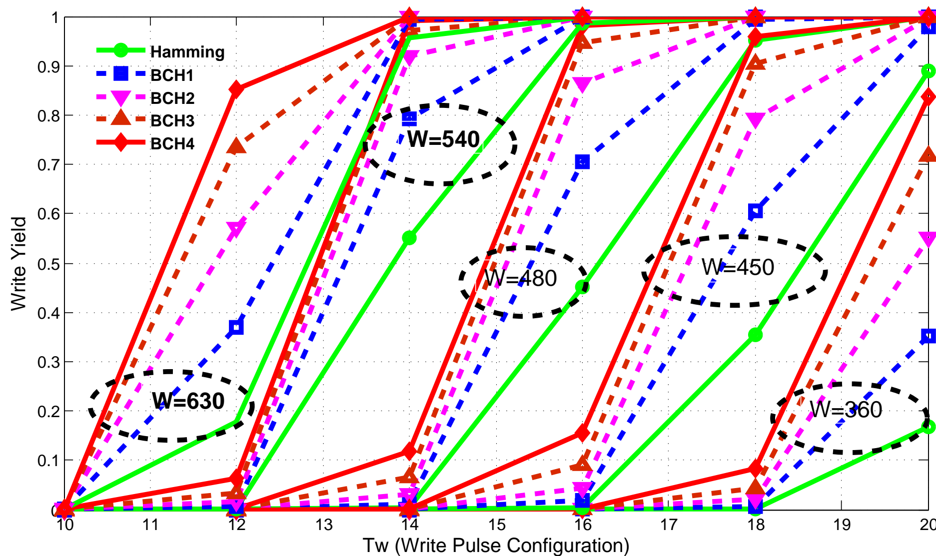

Figure 14: Design space exploration at '0'→'1'.

2.4 APPLICATION 2: WRITE ENERGY ANALYSIS

In addition to write reliability analysis, our PS3-RAM method can also precisely capture the write energy distributions influenced by the variations of device and working environment. In this section, we first prove that there is a sweet point of write pulse width for the minimum write energy without considering any variations. Then we introduce the concept of statistical write energy of STT-RAM cells considering both process variations and thermal fluctuations, and perform the statistical analysis on write energy using our PS3-RAM method.

2.4.1 Write Energy Without Variations

The write energy of a STT-RAM cell during each programming cycle without considering process and thermal variations is deterministic and can be modeled by Eq. (2.17) as:

$$E_{av} = I^2 R \tau_{th}. \tag{2.17}$$

Here I denotes the switching current at either '0'→'1' or '1'→'0' switching, τ_{th} is the corresponding MTJ switching time and R is the MTJ resistance value, i.e., R_L (R_H) for '0'→'1'('1'→'0') switching. As discussed in prior art [48], the switching process of an STT-RAM cell can be divided into three working regions:

$$I = \begin{cases} I_{C_0}\left(1 - \frac{\ln(\tau_{th}/\tau_0)}{\Delta}\right), & \tau_{th} > 10\text{ns} \\ I_{C_0} + C\ln\left(\frac{\pi}{2\theta}\right)/\tau_{th}, & \tau_{th} < 3\text{ns} \\ \frac{P}{\tau_{th}} + Q. \quad 3 \leq \tau_{th} \leq 10\text{ns} \end{cases} \tag{2.18}$$

Here I_{C_0} is the critical switching current, Δ is thermal stability, $\tau_0 = 1$ns is the relax time, θ is the initial angle between the magnetization vector and the easy axis, and C, P, Q are fitting parameters.

For a relatively long switching time range ($\tau_{th} \approx 10 \sim 300$ns), the undistorted write energy P_{av} can be calculated as:

$$\begin{aligned} E_{av} &= I_{C_0}^2\left(1 - \frac{\ln \tau_{th}}{\Delta}\right)^2 R\tau_{th} \\ &= \frac{I_{C_0}^2 R}{\Delta^2}(\Delta - \ln \tau_{th})^2 \tau_{th}. \end{aligned} \tag{2.19}$$

In the long switching time range, we have $\ln \tau_{th} < 0$. Thus, $(\Delta - \ln \tau_{th})^2 \tau_{th}$ or E_{av} monotonically raises as the write pulse τ_{th} increases and the minimized write energy E_{av} occurs at $\tau_{th} = 10$ns.

In the ultra-short switching time range ($\tau_{th} < 3$ns), E_{av} can be obtained as:

$$\begin{aligned} E_{av} &= \left[I_{C_0} + C \ln\left(\frac{\pi}{2\theta}\right) \Big/ \tau_{th}\right]^2 R\tau_{th} \\ &= 2I_{C_0} RC \ln\left(\frac{\pi}{2\theta}\right) + I_{C_0}^2 R\tau_{th} + \frac{C^2 \ln^2(\pi/2\theta) R}{\tau_{th}} \\ &\geq 2I_{C_0} RC \ln\left(\frac{\pi}{2\theta}\right) + 2\sqrt{I_{C_0}^2 R^2 C^2 \ln^2(\pi/2\theta)} \\ &\geq 4I_{C_0} RC \ln\left(\frac{\pi}{2\theta}.\right) \end{aligned} \tag{2.20}$$

As Eq. (2.20) shows, the minimum of E_{av} can be achieved when $\tau_{th} = \frac{C\ln(\pi/2\theta)}{I_{C_0}}$. However, for the ultra-short switching time range (usually $\frac{C\ln(\pi/2\theta)}{I_{C_0}} > 3$ns), E_{av} monotonically decreases as τ_{th} increases.

Similarly, in the middle switching time range ($3 \leq \tau_{th} \leq 10$ns), E_{av} can be expressed as:

$$\begin{aligned} E_{av} &= \left(\frac{P}{\tau_{th}} + Q\right)^2 R\tau_{th} \\ &= \left(\frac{P}{\sqrt{\tau_{th}}} + Q\sqrt{\tau_{th}}\right)^2 R. \\ &\geq 4PQR \end{aligned} \tag{2.21}$$

Again, the minimized E_{av} occurs at $\tau_{th} = \frac{P}{Q}$. Here $\frac{P}{Q} \geq 10$ns based on our device parameters characterization [48]. Thus, the write energy P_{av} in this range monotonically decreases as τ_{th} grows.

According to the monotonicity of E_{av} in the three regions, the most energy-efficient switching point of E_{av} should be at $\tau_{th} = 10$ns. To validate above theoretical deduction for the sweet point of E_{av}, we also conduct the SPICE simulations. Here the STT-RAM device model without considering process and thermal variations is also adopted from [48].

Fig. 15 shows the simulated write energy E_{av} over different write pulse at '0'→'1' switching. As Fig. 15 shows, E_{av} monotonically decreases in the ultra-short switching range and

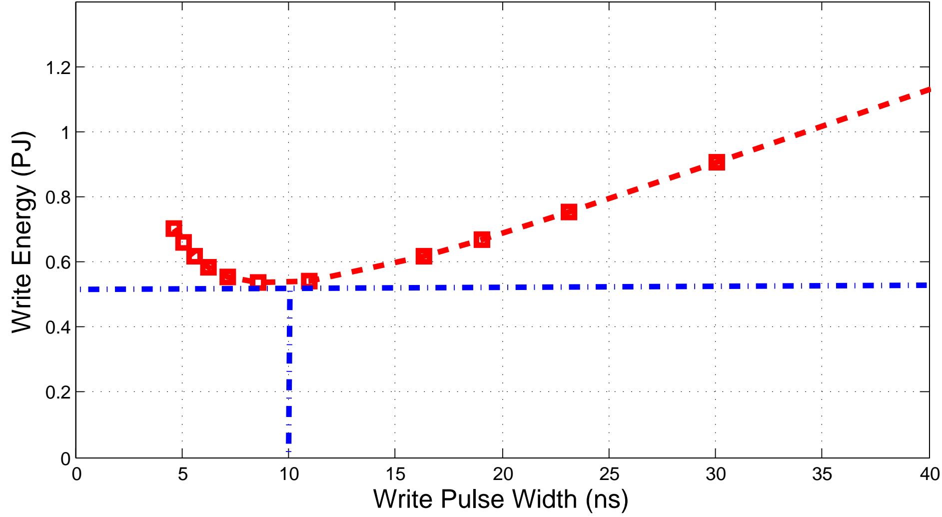

Figure 15: Average Write Energy under different write pulse width when T=300K.

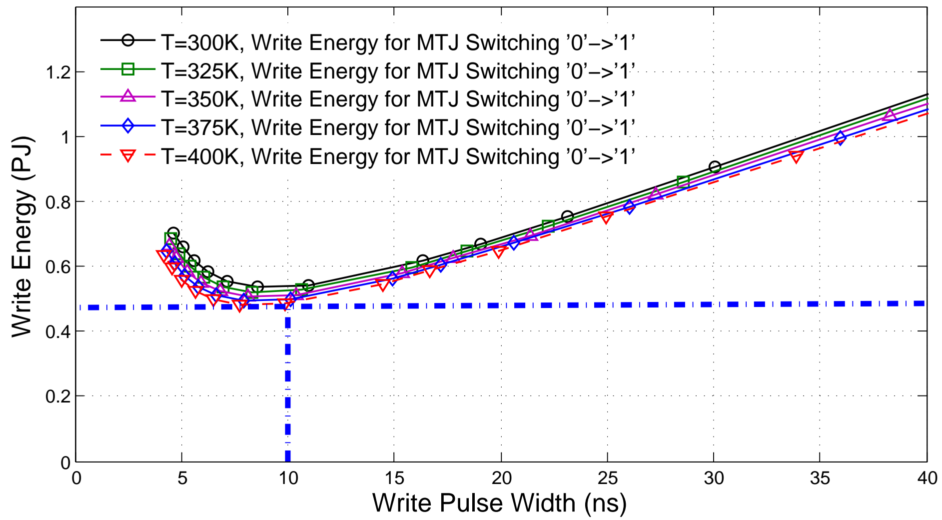

Figure 16: Average Write Energy vs write pulse width under different temperature.

continues decreasing in the middle range, but becomes monotonically increasing after entering the long switching time range. The sweet point of E_{av} occurs around $\tau_{th} = 10$ns, which validates our theoretical analysis for the write energy without considering any variations.

We also present the simulated E_{av}–τ_{th} curve under different temperatures in Fig. 16. The trend and sweet point of E_{av}–τ_{th} curves remain almost the same when the temperature increases from T=300K to T=400K. In fact, the write energy E_{av} decreases a little bit as the temperature increases. The reason is that the driving ability loss of the NMOS transistor (I) dominates E_{av} though the MTJ switching time (τ_{th}) slightly increases when the working temperature raises.

2.4.2 PS3-RAM for Statistical Write Energy

As discussed in Section 2.4.1, the write energy of a STT-RAM cell can be deterministically optimized when all the variations are ignored. However, since the switching current I, the resistance R, and the switching time τ_{th} in Eq. (2.17) may be distorted by CMOS/MTJ process variations and thermal fluctuations, the deterministic value will not longer be able to represent the statistic nature of the write energy of a STT-RAM cell. Accordingly, the optimized write energy at sweet point ($\tau_{th} = 10$ns) shown in Fig. 15 should be expanded as a distribution.

Similar to the write failure analysis in Section 2.3, we conduct the statistical write energy analysis using our PS3-RAM method. We choose the mean of NMOS transistor width $W = 540$nm. The remained device parameters and variation configurations keep the same as TABLE 1.

Fig. 17 and 18 show the simulated statistical write energy by PS3-RAM for both switching directions at 300K. For comparison, the SPICE simulation results are also presented. As shown in those two figures, the distribution of write energy captured by our PS3-RAM method are in excellent agreement with the results from SPICE simulations at both '1'→'0' and '0'→'1' switching's.

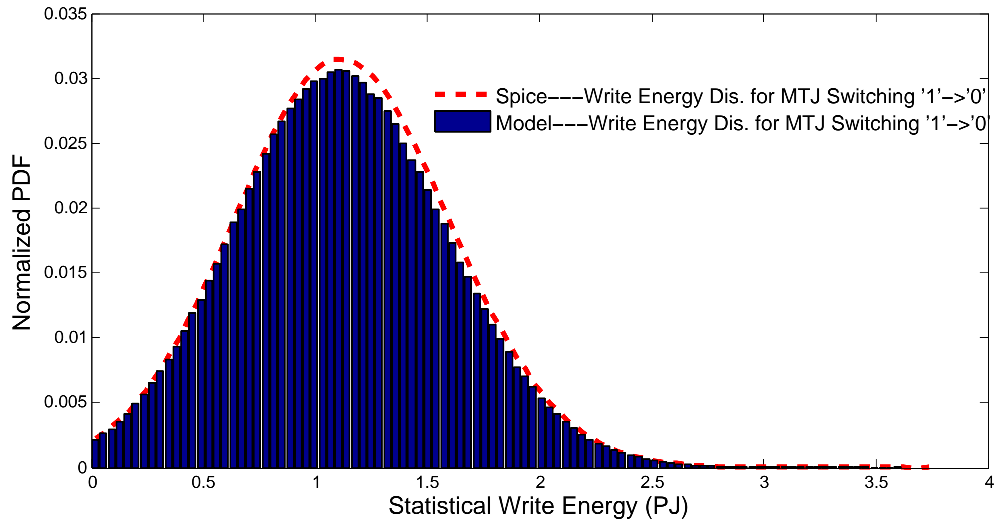

Figure 17: Statistical Write Energy vs write pulse width at ‘1’→‘0’.

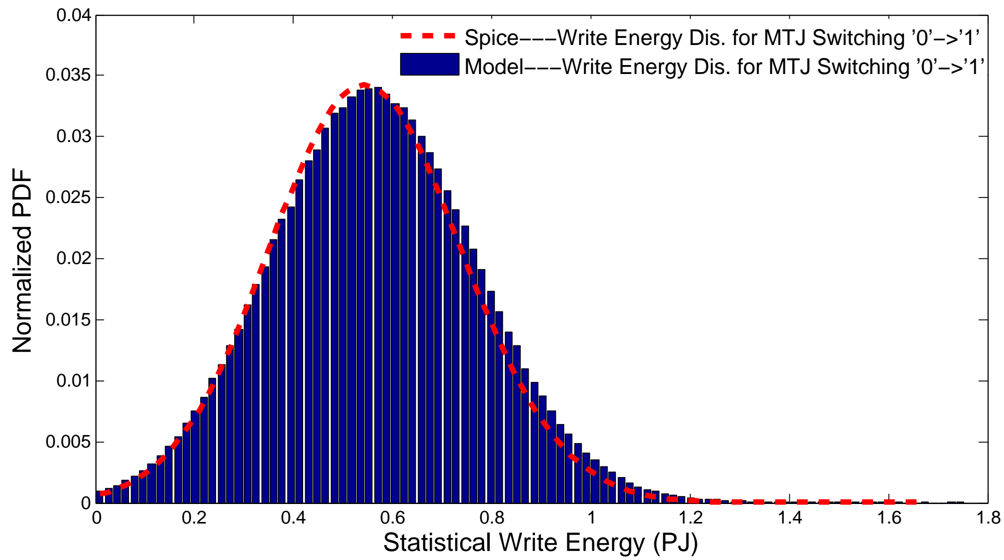

Figure 18: Statistical Write Energy vs write pulse width at ‘0’→‘1’.

2.5 COMPUTATION COMPLEXITY EVALUATION

We compared the computation complexity of our proposed PS3-RAM method with the conventional simulation method. Suppose the number of variation sources is M, for a statistical analysis of a STT-RAM cell design, the numbers of SPICE simulations required by conventional flow and PS3-RAM are $N_{std} = {N_s}^M$ and $N_{PS3-RAM} = 2KM + 1$, respectively. Here K denotes the sample numbers for window based smooth filter in sensitivity analysis, N_s is average sample number of every variation in the Monte-Carlo simulations in conventional method, $K \ll N_s$. The speedup $X_{speedup} \approx \frac{N_s^M}{2KM}$ can be up to multiple orders of magnitude: for example, if we set $N_s = 100$, $M = 4$, (note: V_{th} is not an independent variable) and $K = 50$, the speed up is around 2.5×10^5.

2.6 APPENDIX

In this appendix, we give the details on the model deduction in sensitivity analysis and the summary of the analytic results involved in the PS3-RAM development. We also present the validation of our analytic results based on Monte-Carlo simulations. TABLE 2 [51] summarizes some additional parameters used in this section.

2.6.1 Sensitivity Analysis Model Deduction

The sensitivity analysis model is developed based on the electrical MTJ model and the simplified BSIM model [52, 51]. At '1'→'0' switching, the MTJ switching current supplied by an NMOS transistor working in the triode region is:

$$I = \frac{\beta \cdot \left[(V_{dd} - V_{th}) (V_{dd} - IR) - \frac{a}{2} (V_{dd} - IR)^2 \right]}{1 + \frac{1}{v_{sat} L} (V_{dd} - IR)}. \tag{2.22}$$

Here $\beta = \frac{\mu_0 C_{ox}}{1 + U_0 (V_{dd} - V_{th})} \frac{W}{L}$. As summarized in Table 2, U_0 is the vertical field mobility reduction coefficient, μ_0 is electron mobility, C_{ox} is gate oxide capacitance per unit area, a is body-effect coefficient and v_{sat} is carrier velocity saturation. The MTJ is in its high resistance state, or $R = R_H$.

Table 2: Parameter definition

Variable	Definition
U_0	Vertical field mobility reduction coefficient
μ_0	Electron mobility
C_{ox}	Gate oxide capacitance per unit area
a	Body-effect coefficient
v_{sat}	Carrier velocity saturation

Based on PTM [50] and BSIM [51], the partial derivatives in Eq. (2.6) can be calculated by ignoring the minor terms in the expansion of Eq. (2.22) as:

$$\left(\frac{\partial I}{\partial W}\right)_0^2 \approx \frac{1}{(A_1 W + B_1)^4}, \quad \left(\frac{\partial I}{\partial L}\right)_0^2 \approx \frac{1}{\left(\frac{A_2}{W} + B_2 W + C\right)^2},$$
$$\left(\frac{\partial I}{\partial R}\right)_0^2 \approx \frac{1}{\left(\frac{A_3}{W} + B_3\right)^4}, \quad \left(\frac{\partial I}{\partial V_{th}}\right)_0^2 \approx \frac{1}{\left(\frac{A_4}{\sqrt{W}} + B_4\sqrt{W}\right)^4}.$$

Here,

$$\begin{aligned}
A_1 &= \sqrt{\frac{\mu_0 C_{ox} V_{dd}\left(V_{dd} - V_{th}\right)}{L}} R, \\
B_1 &= \sqrt{\frac{L}{\mu_0 C_{ox} V_{dd}\left(V_{dd} - V_{th}\right)}}, \\
A_2 &= \frac{L^2}{\mu_0 C_{ox} V_{dd}\left(V_{dd} - V_{th}\right)}, \\
B_2 &= R^2 \mu_0 C_{ox} \frac{V_{dd} - V_{th}}{V_{dd}}, \\
A_3 &= \frac{L}{\mu_0 C_{ox} \sqrt{V_{dd}}\left(V_{dd} - V_{th}\right)}, \\
B_3 &= \frac{R}{\sqrt{V_{dd}}}, C = \frac{2LR}{V_{dd}}, \\
A_4 &= \sqrt{\frac{L}{\mu_0 C_{ox} V_{dd}}}, \\
B_4 &= \sqrt{\frac{\mu_0 C_{ox}}{L V_{dd}}} R\left(V_{dd} - V_{th}\right).
\end{aligned}$$

At '0'→'1' switching, the NMOS transistor is working in the saturation region. The current through the MTJ is:

$$I = \frac{\beta}{2a}\left[\left(V_{dd} - IR - V_{th}\right) - \frac{I}{W C_{ox} v_{sat}^2}\right]^2. \tag{2.23}$$

The MTJ is in its low resistance state, or $R = R_L$. the derivatives can be also calculated as:

$$\left(\frac{\partial I}{\partial W}\right)_1^2 \approx \frac{1}{(A_5 W + B_5)^4}, \quad \left(\frac{\partial I}{\partial L}\right)_1^2 \approx \frac{1}{\left(\frac{A_6}{W} + B_6\right)^2},$$
$$\left(\frac{\partial I}{\partial R}\right)_1^2 \approx \frac{1}{\left(\frac{A_7}{W} + B_7\right)^4}, \quad \left(\frac{\partial I}{\partial V_{th}}\right)_1^2 \approx \frac{1}{\left(\frac{A_8}{W} + B_8\right)^2}.$$

by ignoring the minor terms in the expansion of Eq. (2.23). Here, all the parameters, including A_5, B_5, A_6, B_6, A_7, B_7 and A_8, are shown as below:

$$
\begin{aligned}
A_5 &= \sqrt{\frac{2C_{ox}v_{sat}\mu_0}{La+\mu_0\left(V_{dd}-V_{th}\right)}}R,\\
B_5 &= \frac{\mu_0}{2C_{ox}v_{sat}\left[La+\mu_0\left(V_{dd}-V_{th}\right)\right]},\\
A_6 &= \frac{\mu_0}{2aC_{ox}v_{sat}^2},\\
B_6 &= \frac{R\mu_0}{av_{sat}},\\
A_7 &= \frac{1}{2C_{ox}v_{sat}}\sqrt{\frac{\mu_0}{Lav_{sat}+\mu_0\left(V_{dd}-V_{th}\right)}},\\
B_7 &= \sqrt{\frac{\mu_0}{Lav_{sat}+\mu_0\left(V_{dd}-V_{th}\right)}}R,\\
A_8 &= \frac{1}{2C_{ox}v_{sat}}, B_8 = R.
\end{aligned}
$$

The contributions of different variation sources to I are represented by:

$$
\begin{aligned}
S_1 = \left(\frac{\partial I}{\partial W}\right)^2\sigma_W^2, S_2 = \left(\frac{\partial I}{\partial L}\right)^2\sigma_L^2, S_3 = \left(\frac{\partial I}{\partial R}\right)^2\sigma_R^2,\\
S_4 = \left(\frac{\partial I}{\partial V_{th}}\right)^2\left(\frac{C_1}{WL}+\frac{C_2}{exp\left({L}/{l'}\right)}\cdot\frac{W_c}{W}\cdot\sigma_L^2\right).
\end{aligned}
\tag{2.24}
$$

Here S_1, S_2, S_3 and S_4 denote the variations induced by W, L, R (R_H or R_L) and V_{th}, respectively.

2.6.2 Analytic Results Summary

TABLE 3 shows the monotonicity and the upper or lower bounds of the variation contributions $S_1 - S_4$ as the transistor channel width W increases. Here, "↑" , "↓" and "↗↘" denotes monotonic increasing, monotonic decreasing and changing as a convex function. $K_1 = \frac{C_1}{L}+\frac{C_2W_c\sigma_L^2}{\exp\left({L}/{l'}\right)}$. TABLE 3 also gives the maximum and minimum values of $S_i\,(i=1\cdots4)$ and their corresponding W's.

Table 3: Summary of variation contribution

	Variation	Monoto	bounds	$W \to \infty$
'0'	S_1	$\downarrow$	$\min S_1 = 0$ $W = \infty$	$S_1 \to 0$
	S_2	$\nearrow\searrow$	$\max S_2 = \left(\frac{V_{dd}}{4LR_H}\sigma_L\right)^2$ $W = \frac{L}{\mu_0 C_{ox}(V_{dd}-V_{th})R_H}$	$S_2 \to 0$
	S_3	$\uparrow$	$\max S_3 = \left(\frac{V_{dd}}{R_H^2}\sigma_{R_H}\right)^2$ $W = \infty$	$\max S_3$
	S_4	$\nearrow\searrow$	$\max S_4 = \frac{K_1 \mu_0 C_{ox} V_{dd}^2}{16LR_H(V_{dd}-V_{th})}$ $W = \frac{L}{\mu_0 C_{ox} R_H (V_{dd}-V_{th})}$	$S_4 \to 0$
'1'	S_1	$\downarrow$	$\min S_1 = 0$ $W = \infty$	$S_1 \to 0$
	S_2	$\uparrow$	$\max S_2 = \left(\frac{a v_{sat}}{R_L \mu_0}\sigma_L\right)^2$ $W = \infty$	$\max S_2$
	S_3	$\uparrow$	$\max S_3 \approx \left(\frac{V_{dd}-V_{th}}{R_L^2}\sigma_{R_L}\right)^2$ $W = \infty$	$\max S_3$
	S_4	$\nearrow\searrow$	$\max S_4 = \frac{C_{ox} v_{sat}}{2R_L} K_1$ $W = \frac{1}{2C_{ox} v_{sat} R_L}$	$S_4 \to 0$

2.6.3 Validation of Analytic Results

As Eq. (2.24) shows, $\left(\frac{\partial I}{\partial W}\right)^2$, $\left(\frac{\partial I}{\partial L}\right)^2$, and $\left(\frac{\partial I}{\partial R}\right)^2$ solely determine the trends of S_1, S_2, S_3, respectively, when W increases at both switching directions. The corresponding Monte-Carlo simulation results of S_1, S_2, S_3 are shown in Fig. 19, 20, and 21, respectively.

Fig. 19 shows S_1 monotonically decreases to zero as W increases to infinity at both switching directions. Its value at '1'→'0' switching is always greater than that at '0'→'1' switching because $A_1 < A_5$.

Fig. 20 shows that the variation contribution of L at '0'→'1' switching is always larger than that at '1'→'0' switching. The gap between them reaches the maximum when $W \to \infty$.

Fig. 21 shows that the contribution from MTJ resistance R becomes dominant in the MTJ switching current distribution when W is approaching infinity. Because $\left(\frac{V_{dd}-V_{th}}{R_L^2}\sigma_{R_L}\right)^2 < \left(\frac{V_{dd}}{R_H^2}\sigma_{R_H}\right)^2$, the normalized contribution of R is always larger at '1'→'0' switching than that at '0'→'1' switching.

We note that the additional coefficient $\left(\frac{C_1}{WL} + \frac{C_2}{exp\left(L/l'\right)}\frac{W_c}{W}\sigma_L^2\right)$ at the right side of Eq. (2.24) after $\left(\frac{\partial I}{\partial V_{th}}\right)^2$ results in different features of $\left(\frac{\partial I}{\partial V_{th}}\right)^2$ from S_4 in our simulations.

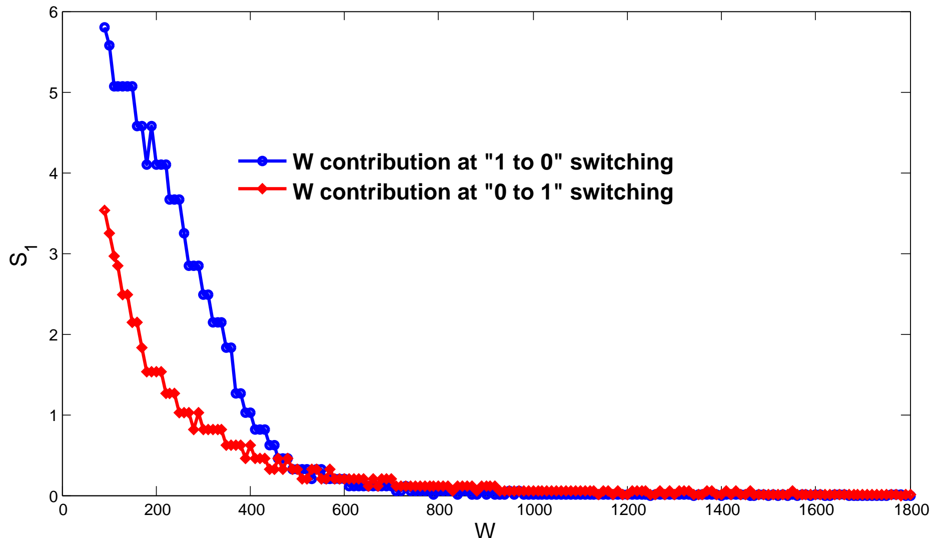

Figure 19: Contributions from W.

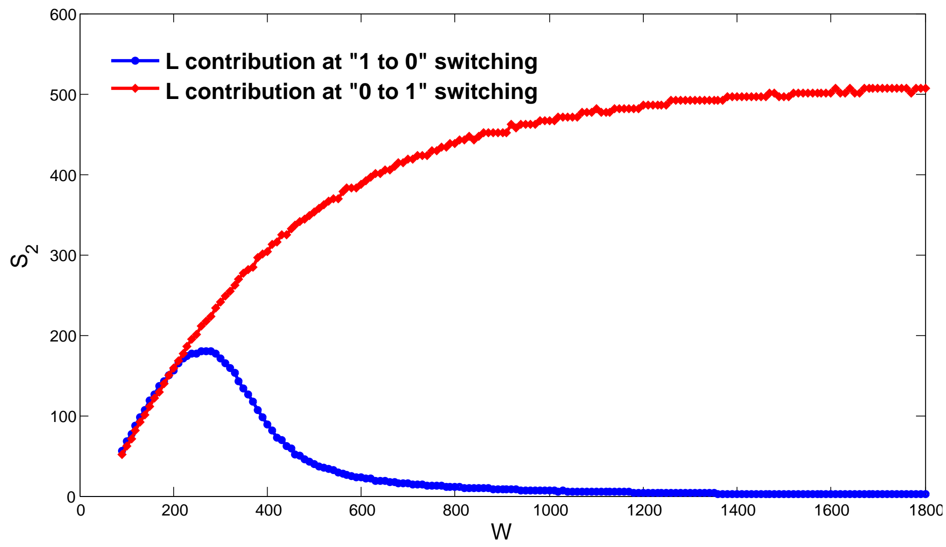

Figure 20: Contributions from L.

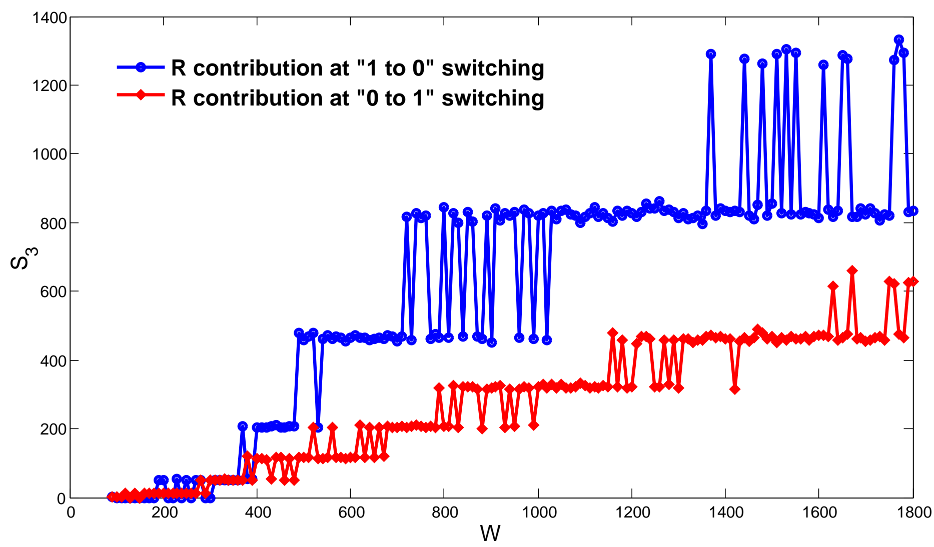

Figure 21: Contributions from R.

Fig. 22 shows the values of $\left(\frac{\partial I}{\partial V_{th}}\right)^2$ at both switching directions. At '0'→'1' switching, $\left(\frac{\partial I}{\partial V_{th}}\right)^2$ increases monotonically when W grows. At '1'→'0' switching, $\left(\frac{\partial I}{\partial V_{th}}\right)^2$ increases first, then quickly decays to zero after reaching its maximum. These trends follow the expressions of $\left(\frac{\partial I}{\partial V_{th}}\right)^2$ at either switching directions very well.

However, because of the additional coefficient on the top of $\left(\frac{\partial I}{\partial V_{th}}\right)^2$, S_4 does not follow the same trend of $\left(\frac{\partial I}{\partial V_{th}}\right)^2$ at either switching directions. Fig. 23 shows that at '0'→'1' switching, S_4 increases first and then slowly decreases when W rises. At this switching direction, S_4 will become zero when $W \to \infty$ due to the existence of the additional coefficient $\left(\frac{C_1}{WL} + \frac{C_2}{exp\left(L/l'\right)}\frac{W_c}{W}\sigma_L^2\right)$.

All these above results are well consistent with our analytic analysis in TABLE 3.

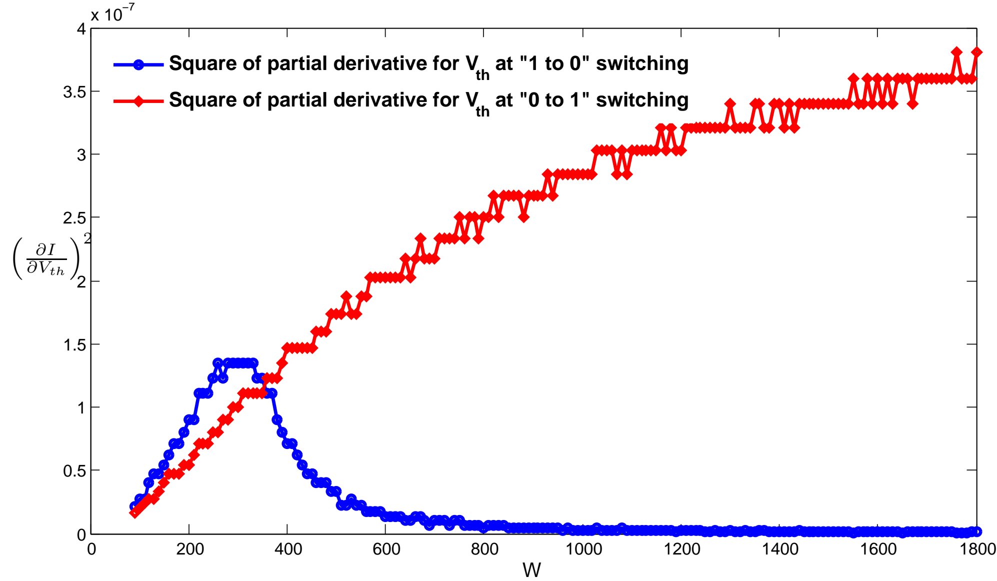

Figure 22: Square partial derivatives for V_{th}.

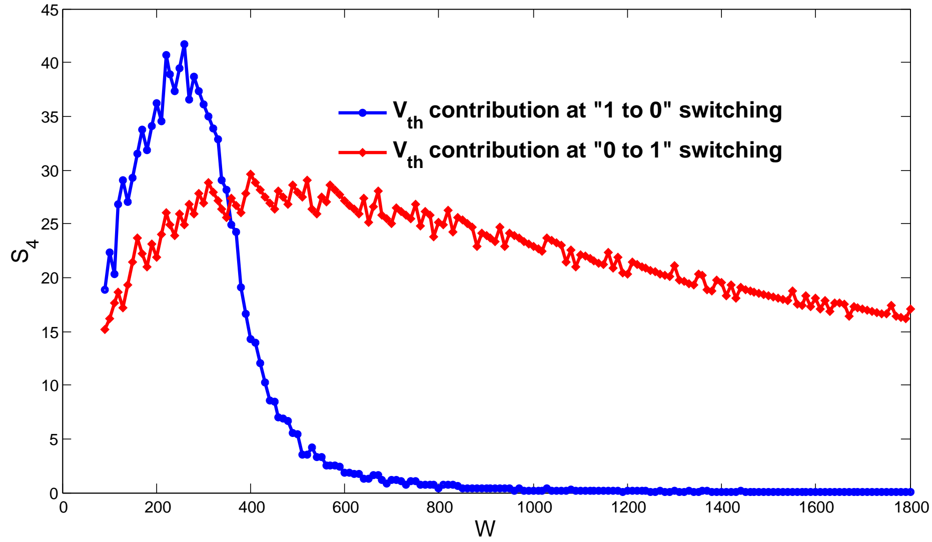

Figure 23: Contributions from V_{th}.

2.7 CHAPTER 2 SUMMARY

In this chapter, we developed a fast and scalable statistical STT-RAM reliability/energy analysis method called PS3-RAM. PS3-RAM can simulate the impact of process variations and thermal fluctuations on the statistical STT-RAM write performance or write energy distributions, without running costly Monte- Carlo simulations on SPICE and macro-magnetic models. Simulation results show that PS3-RAM can achieve very high accuracy compared to the conventional simulation method, while achieving a speedup of multiple orders of magnitude. The great potentials of PS3-RAM in the application of the device/circuit/achitecture co-optimization of STT-RAM designs are also demonstrated.

3.0 CONTENT-DEPENDENT ECC DESIGNS

In Chapter 2, PS3-RAM shows that the *bit error rate* (BER) and/or the required switching time of writing "1" is significantly larger or longer than that of writing "0", indicating the asymmetric write error rate of STT-RAM cells at two bit-flipping directions. In this chapter, we will design high-efficiency ECCs by leveraging the understanding on the inherent stochastic MTJ switching process and the new ECC design concept–content-dependent. The rest of this chapter is organized as follows: Section 3.1 gives the motivation of our research; Section 3.2 describes proposed *asymmetric write channel* (AWC) model to analyze the accumulated asymmetry of write errors step-by-step; Section 3.3 illustrates the CD-ECC technique details, including both typical-corner-ECC and worst-corner-ECC; Section 3.4 presents the efficacy evaluations of CD-ECC technique at both memory design and architecture level. Section 3.5 summarizes this chapter.

3.1 RESEARCH MOTIVATIONS

3.1.1 Asymmetric STT-RAM Write Errors

The MTJ switching time can be reduced by increasing the magnitude of switching current. A write failure happens if the switching current is removed before the MTJ switching completes. Some factors may cause the uncertainty of MTJ switching time, such as 1) the driving ability variation of the NMOS transistor, which is caused by the the parametric variations of NMOS transistor and MTJ; and 2) the stochastic MTJ magnetization switching process induced by random thermal fluctuations [41].

In the write operations of STT-RAM cells, the MTJ switching from low-resistance state to high-resistance state ($0 \rightarrow 1$) is considered as "unfavorable" switching direction compared to the MTJ switching at the opposite direction: $0 \rightarrow 1$ flipping requires larger switching current than $1 \rightarrow 0$ flipping due to the lower spin-transfer efficiency [41, 49]. Also, the variation of MTJ switching time at $0 \rightarrow 1$ flipping is more prominent, leading to a higher write error rate [36, 48, 55].

3.1.2 Related Work

ECC has been widely used to repair the errors of memory subsystems. Popular ECC schemes, such as SEC-DED [56], BCH [57, 39, 38], etc., are designed by assuming the error rates of the stored data with different values are always identical. These ECC schemes, however, are not suitable for STT-RAM designs because they are generally designed for the worst-case that rarely happens and cannot address the asymmetric bit error rates at different flipping directions efficiently. Such a limitation of conventional ECCs can be explained by using the following example: We assume the length of a pre-coding codeword is 256 and all bits of the codeword flip at each write. The bit error rate of $0 \rightarrow 1$ flipping is $P_{ER,0\rightarrow 1} = 0.01$ while the one of $1 \rightarrow 0$ flipping is $P_{ER,1\rightarrow 0} = 0$ (the extreme case for asymmetric errors). A Hamming code with the pre-coding codeword dimension $k = 256$, the post-coding codeword length $n = 265$ and the error correction capability $t = 1$ is applied. Fig. 24 shows that the block-level reliability P_{block}, which denotes the probability that all bits successfully flip, decreases as the Hamming weight W (i.e., the number of '1's, $0 \leq W \leq n$) of the destination codeword increases. Applying a strong ECC to cover the worst case, i.e., $W = n$, is very inefficient because such an extremely asymmetric pattern rarely happens in reality. Also, the content-independent design, i.e., applying the same error correcting capability ($t = 1$) to all the data patterns in the conventional Hamming ECCs, results in extremely unbalanced P_{block} of different codewords (*e.g.* $W = n$ and $W = \frac{n}{2}$).

Also, in information theory, many studies have been conducted to design the ECC schemes tailored for the asymmetric errors [58, 59, 60, 61, 62, 63, 64, 57, 65]. However, these theoretical studies mainly emphasized on the approaches to estimate the upper bound

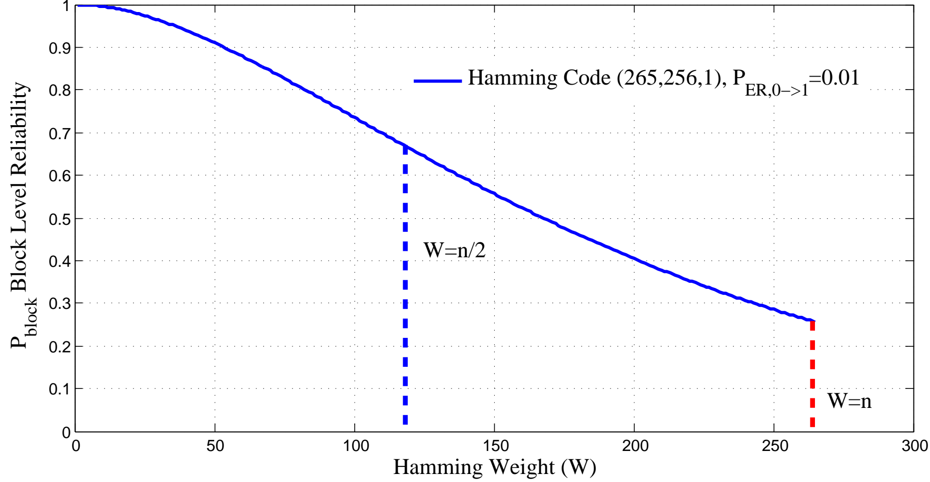

Figure 24: The relationship between block level reliability P_{block} and Hamming weight W for asymmetric errors.

of codeword size or the asymptotic code rate based on the improper pre-code or post-code length assumptions, rather than coding/decoding algorithms themselves and hardware designs. The long coding/decoding latency and high hardware cost prevent these schemes from being integrated in the memory hierarchy on computer systems.

Finally, many proposals about using ECCs to improve the lifetime, energy, and reliability of memory subsystems have been proposed [30, 31, 32, 33, 34, 35, 66]. In [34], J. Kim *et al.* developed a 2-D coding scheme to fix multi-bit errors. The long decoding latency (100's∼1000's cycles) makes it unsuitable to latency-sensitive applications like on-chip cache. In [30], the *error correction pointer* (ECP) was proposed to fix the hard bit errors of phase change memory. However, ECP is basically a redundant design and cannot fix the random or soft errors in STT-RAM. None of these solutions can efficiently handle the asymmetric errors in STT-RAM.

Considering the fact that the number of bit-flipping's of the cache codewords is approximately proportional to the hamming weight of the destination codeword [67] during the normal operations of the memory hierarchy in computer systems, we are motivated to develop new ECCs, namely, *content-dependent ECCs* (CD-ECCs) to enhance and balance

the reliability of the STT-RAM with asymmetric write errors by minimizing the Hamming weights of the codewords.Here the "content" means the hamming weight or reliability degree for different data patterns.

3.2 ASYMMETRIC WRITE CHANNEL

In multi-level cell (MLC) memory design, the optimization of reference signals in read circuitry is critical for improving the read reliability [11, 20]. Read channel model for MLC NAND Flash memory is also created to enable fast simulation and analysis on the performance and robustness of the read circuitry [68]. Due to the considerable complexity of STT-RAM write mechanism, we borrow the concept of read channel model and construct a write channel model to efficiently analyze the asymmetric write errors of STT-RAM.

3.2.1 Asymmetric Write Channel (AWC) Model

Fig. 25 illustrates the overview of our proposed *asymmetric write channel* (AWC) model. Based on the modeled contributors to the MTJ switching time variation, the AWC model can be divided into five stages. The asymmetry of STT-RAM write operations is mainly generated in the first four steps. Based on the sources of asymmetry generation, we further categorize the stages from *ideal switching current programming* to *switching-current-to-switching-time mapping* as "**parametric asymmetric stages (PAS)**" and the steps from *switching current distorted by PVs (process variations)* to *switching time distorted by thermal fluctuations* as "**random asymmetric stages (RAS)**".

3.2.1.1 Parametric Asymmetric Stages (PAS) In PAS, the asymmetry is mainly generated from the unbalanced driving ability of the NMOS transistor and the asymmetric switching-current-to-switching-time mapping at different bit-flipping directions. In Fig. 1(c), when writing '0', both the *word-line* (WL) and the *bit-line* (BL) are connected to V_{dd} while the *source-line* (SL) is connected to ground. The V_{gs} of the NMOS transistor is V_{dd}; when

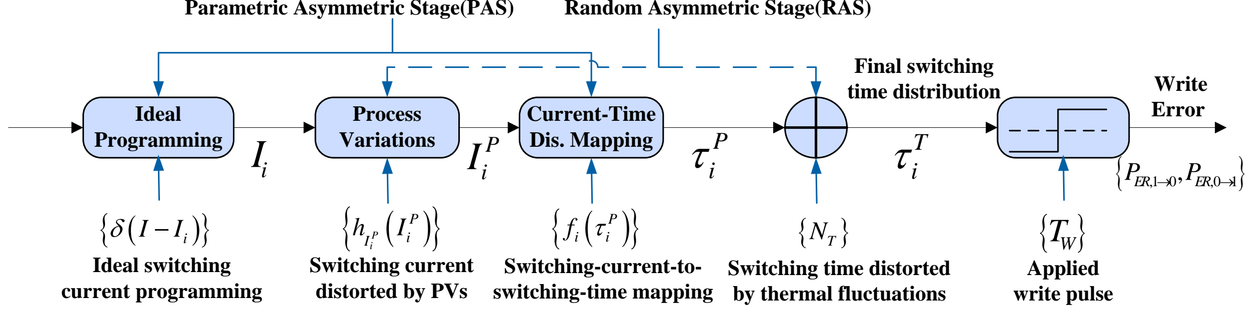

Figure 25: Overview of the proposed asymmetric write channel (AWC) model.

writing '1', the WL and the SL are connected to V_{dd} while the BL is connected to ground. The V_{gs} becomes lower than V_{dd} due to the potential loss on the MTJ. In such a case, the switching current supplied by the NMOS transistor to the MTJ (I_1) is lower than that of writing '0' (I_0).

The asymmetry introduced in the switching-current-to-switching-time mapping stage is from the different switching time of the MTJ at two bit-flipping directions even driven by the same switching current. This effect further deteriorates the asymmetry of MTJ switching induced by the unbalanced NMOS transistor driving ability at two bit-flipping directions.

3.2.1.2 Random Asymmetric Stages (RAS) In RAS, the asymmetry is mainly generated by the difference between the impacts of the CMOS/MTJ process variations and the thermal fluctuations on the MTJ switching process at two bit-flipping directions. The bias condition difference of the NMOS transistor results in a much larger variation of I_1 compared to I_0 even though $I_1 < I_0$ [48]. For a certain nominal value of MTJ switching time, the ratio between the standard deviation and mean of the MTJ switching time at $0 \to 1$ flipping is also higher than that at $1 \to 0$ flipping. When the MTJ switching time is shorter than $3ns$, it roughly follows Gaussian distribution; Otherwise, it follows Poisson distribution or mixed Gaussian-Poisson distribution [48, 28].

3.2.1.3 Construction of AWC Model

The construction of AWC model starts with the ideal case that no any variations exist in the MTJ switching process. Consequently, the PDF (probability density function) of the corresponding MTJ switching current ($P_I(I)$) can be expressed as:

$$P_I\left(I\right) = \delta\left(I - I_i\right), i = 0, 1 \tag{3.1}$$

$$\delta\left(I - I_i\right) = 0, I \neq I_i \tag{3.2}$$

$$\int_0^\infty \delta\left(I - I_i\right) dI = 1. \tag{3.3}$$

Here, δ is Dirac Delta function, which is widely utilized in wireless communication channel model development [69]. I_0 and I_1 are the ideal switching current at the bit-flipping's of $1 \rightarrow 0$ and $0 \rightarrow 1$, respectively. Normally $I_0 < I_1$.

We propose to use dual-exponential function to model the impacts of CMOS/MTJ process variations (PVs) on the NMOS transistor driving ability. The corresponding statistical transfer function is defined as:

$$h_{I_i^P}\left(I_i^P\right) = \frac{1}{k_i \lambda_i} e^{-\lambda_i \left|I_i^P\right|}, i = 0, 1. \tag{3.4}$$

Here I_i^P, $i = 0, 1$ is the MTJ switching current distorted by PVs at each bit-flipping direction. k_i is chosen to ensure the integral of $h_{I_i^P}\left(I_i^P\right)$ equals 1. λ_i describes the severeness of the impact of PVs on I_i^P. The PDF of I_i^P can be obtained from the convolution between Eq. (3.1) and (3.4) as:

$$\begin{aligned} P_{I_i^P}\left(I_i^P\right) &= P_I\left(I_i^P\right) \otimes h_{I_i^P}\left(I_i^P\right) \\ &= h_{I_i^P}\left(I_i^P - I_i\right) \\ &= \frac{1}{k_i \lambda_i} e^{-\lambda_i \left|I_i^P - I_i\right|}, i = 0, 1. \end{aligned} \tag{3.5}$$

Here $\otimes$ denotes "convolution" operation.

The distribution of the MTJ switching time τ_i^P, $i = 0, 1$ at each bit-flipping direction can be derived from the I_i^P based on the characterized relationship between the ideal MTJ

switching current I_i and switching time τ_i. By applying the corresponding $I_i - \tau_i$ mapping curve, the PDF of τ_i^P ($P_{\tau_i^P}\left(\tau_i^P\right)$) can be expressed as:

$$P_{\tau_i^P}\left(\tau_i^P\right) = P_{I_i^P}\left(f_i\left(\tau_i^P\right)\right)\left|\frac{df_i\left(\tau_i^P\right)}{d\tau_i^P}\right|. \tag{3.6}$$

Here $I_i^P = f_i\left(\tau_i^P\right), i = 0, 1$ denotes the statistical $I_i - \tau_i$ mapping curve at each bit-flipping directions [28, 48].

The impact of thermal fluctuations on the MTJ switching time variation can be modeled as a thermal noise N_T with mean μ_{N_T} and standard deviation σ_{N_T}, or $P_{N_T,j}\left(N_T\right) = P_{N_T,j}\left(\tau_i^T|\tau_i^P\right), i = 0, 1, j = 1, 2, 3$ at three working regions. At the different working regions, N_T follows Gaussian distribution $P_{N_T,1}\left(N_T\right)$ ($\tau_i^p < 3ns$), Mixed Gaussian and Poisson distribution $P_{N_T,2}\left(N_T\right)$ ($3 \leq \tau_i^p \leq 10ns$), and Poisson distribution $P_{N_T,3}\left(N_T\right)$ ($\tau_i^p > 10ns$), respectively. μ_{N_T} and σ_{N_T} can be calculated by $\mu_{N_T} = \tau_i^p$ and $\sigma_{N_T} = g_i\left(\mu_{N_T}\right), i = 0, 1$, where g_i is the transfer function between the mean and the standard deviation of MTJ switching time under the impact of thermal noise at each switching direction [48]. Thus, the PDF of MTJ switching time τ_i^T under both process variations and thermal fluctuations can be derived by:

$$\begin{aligned} P_{\tau_i^T}\left(\tau_i^T\right) = P_{N_T,1}\left(N_T\right)\int_0^3 P_{\tau_i^P}\left(\tau_i^P\right)d\tau_i^P + P_{N_T,2}\left(N_T\right) \\ \int_3^{10} P_{\tau_i^P}\left(\tau_i^P\right)d\tau_i^P + P_{N_T,3}\left(N_T\right)\int_{10}^{\infty} P_{\tau_i^P}\left(\tau_i^P\right)d\tau_i^P. \end{aligned} \tag{3.7}$$

Finally, the write error rate at each bit-flipping direction under a write pulse width of T_w can be calculated by:

$$P = P\left(\tau_i^T > T_w\right) = \int_{T_w}^{\infty} P_{\tau_i^T}\left(\tau_i^T\right)d\tau_i^T. \tag{3.8}$$

3.2.2 Utilization of AWC model

We use our AWC model to simulate an in-plane MTJ with an elliptical shape of 45nm×90nm under 45nm PTM model [50]. The device variation assumption and simulation setup are adopted from [48]. The NMOS transistor width $W = 540$nm and the $V_{dd} = 1.0$V. Fig. 26(a) shows the ideal driving ability of the NMOS transistor when writing '1' and '0', which are 273 μA and 379 μA, respectively. The corresponding parameters for PVs are $k_0 = 116.062, \lambda_0 = 0.131$ and $k_1 = 657.566, \lambda_1 = 0.055$, respectively. After considering the CMOS/MTJ process variation, the MTJ switching current becomes a distribution, as shown in Fig. 26(b). Fig. 26(c) and (d) show the distributions of the MTJ switching time that is directly mapped from the MTJ switching current distributions and further distorted by the thermal fluctuations, respectively. The impacts of different variation sources are accumulated at each stage of the AWC model. As expected, the distributions of the MTJ switching current and switching time at $0 \rightarrow 1$ bit-flipping are much broader than those at $1 \rightarrow 0$ bit-flipping. In Figure 26(d), the right tail of the distribution of the MTJ switching time at $0 \rightarrow 1$ bit-flipping extends much farther than that at $1 \rightarrow 0$ bit-flipping. It indicates a higher error rate under a fixed write pulse width T_w (black solid line), or $P_{ER,1\rightarrow 0} \ll P_{ER,0\rightarrow 1}$. Increasing the T_w can reduce the write error rates at both bit-flipping directions.

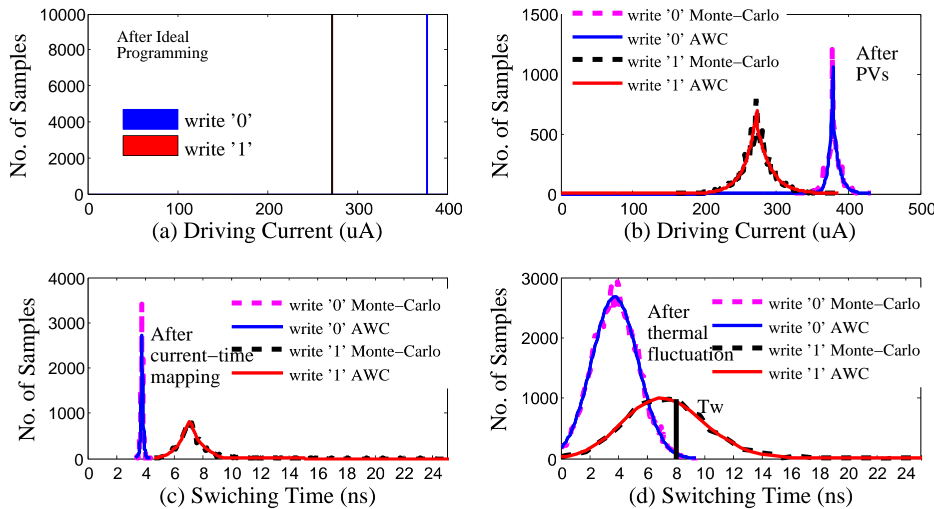

Figure 26: Step breakdowns of AWC Model.

We also conduct Monte-Carlo simulations to obtain the distributions of the MTJ switching time based on 10000 samples. The results of our AWC model match the Monte-Carlo simulation results very well in both PAS (Fig. 26(c)) and RAS (Fig. 26(b)(d)). Fig. 26(b) further shows that our proposed double-exponential function can precisely characterize the switching current distorted by CMOS/MTJ process variations.

Fig. 27 compares values of the asymmetric error rate ratio $R = \frac{P_{ER,0\to1}}{P_{ER,1\to0}}$ respectively extracted from Monte-Carlo simulations and AWC model at different T_w's. AWC model achieves good accuracy over the whole MTJ working regions. Following the increase of T_w, the MTJ switching asymmetry keeps deteriorating as R climbs up. Around the typical STT-RAM working region, e.g., $T_w \sim 10ns$, the R is between $10^3 \sim 10^4$ while the $P_{ER,0\to1} = 5 \times 10^{-3}$. It clearly shows that the $0 \to 1$ bit-flipping is the bottleneck of STT-RAM write reliability. In the rest of this chapter, we choose $T_w = 10ns$ as the working condition of our STT-RAM design.

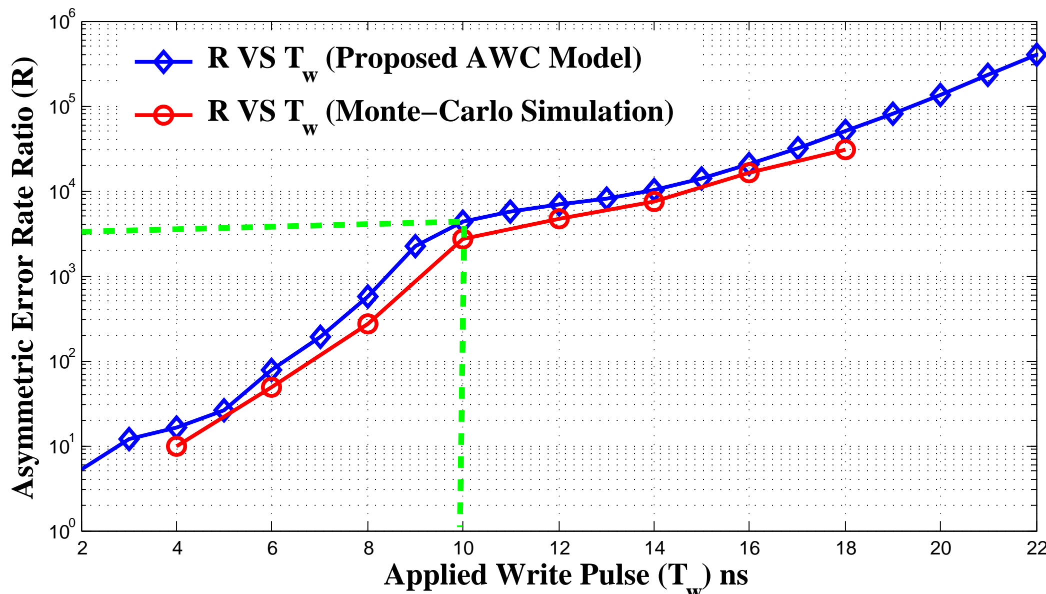

Figure 27: Asymmetric error rate ratio R at different T_w.

3.3 CONTENT-DEPENDENT ECC (CD-ECC)

In this section, we will discuss the details on *content-dependent ECC* (CD-ECC) technology. Two ECC schemes – *typical-corner-ECC* (TCE) and *worst-corner-ECC* (WCE), are developed to handle the codewords with different Hamming weight distributions. These two ECCs are also evaluated at both circuit-design and architecture levels.

3.3.1 Typical-Corner-ECC (TCE)

Fig. 28 depicts the typical Hamming weight distributions of the cache data for SPEC CPU2006 benchmarks [70] *mcf* and *milc*. The majority of the data's Hamming weights locates in the range $0 \leq W \leq \frac{n}{2}, n = 64$, or the typical corner. If the ECC is still designed for the worst case of above cache data with asymmetric error rates, the number of errors will rarely reach the maximum error correcting capability of such a costly ECC. However, we noticed that the cache line data are usually highly correlated at block-level: the adjacent blocks (*e.g.*, each block includes 64 bits) often contain the same or similar data. Based on

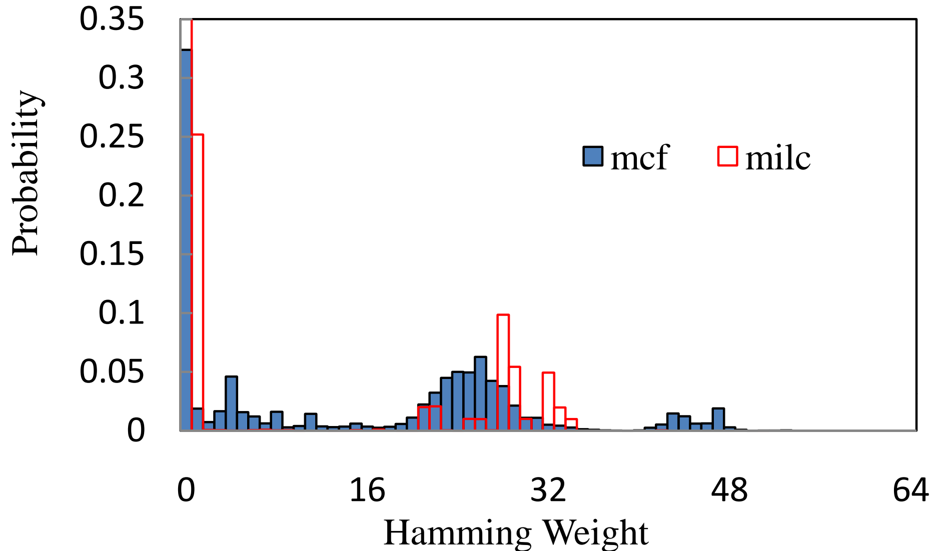

Figure 28: Normalized distribution of the Hamming weight of the cache data from benchmark *mcf* and *milc*.

this observation, we are able to design *typical-corner-ECC* (TCE) to migrate the Hamming weight of the codeword from $0 \leq W \leq \frac{n}{2}$ to the left, such as $0 \leq W \leq \frac{n}{4}$, $0 \leq W \leq \frac{n}{8}$, etc., by leveraging the data difference between the correlated blocks. Accordingly, the required error correcting capability of the ECC will be reduced.

In *typical-Corner-ECC* (TCE) scheme, we first use differential coding to de-correlate the cache data and reduce its Hamming weight. We then select appropriate ECC to protect the de-correlated data. The whole development process can be summarized as follows:

3.3.1.1 Static Differential Coding We first introduce a static differential coding scheme to de-correlate two data blocks as:

$$B_i^{'} = \begin{cases} B_{i-1} \oplus B_i, i = 1, 2 \cdots n \\ B_0, i = 0. \end{cases} \tag{3.9}$$

Here B_i and $B_i^{'}$ are the values of a data block i before and after coding, respectively. n is the number of data blocks in a cache line. '$\oplus$' denotes the XOR operation. Similarly, the decoding algorithm can be expressed as:

$$B_i = \begin{cases} B_i^{'} \oplus B_{i-1}, i = 1, 2 \cdots n \\ B_0, i = 0. \end{cases} \tag{3.10}$$

As shown in Eq. (3.9) and (3.10), the decoding of B_i always refers to the previously decoded block B_{i-1} while the initial reference starts with B_0. As a result, the critical path includes $n-1$ XOR gates. Due to performance concern, n cannot be too large in the implementation of static differential coding, say, less than 7 (or 8 data blocks in a cache line).

3.3.1.2 Dynamic Differential Coding

Static differential coding is based on a very strong assumption: Correlations always exist among the data blocks following a certain sequence. In the real cache data, however, such correlations may be interrupted if a block with low Hamming weight (e.g., all zeros) appears between two correlated blocks with high Hamming weight. In such a case, data blocks with very high Hamming weight may be generated in static differential coding. To resolve this issue, we further propose an enhanced differential coding scheme – dynamic differential coding to selectively perform de-correlation between only the correlated blocks.

A flag bit is introduced to mark the correlated blocks based on the Hamming weight of a block. For example, we divide a 64Byte cache line into 8 blocks {B0,B1,...,B7}. The flag bit of each block $\{Ii, i = 0, 1, ..., 7\}$ can be calculated as:

$$Ii = \begin{cases} 0, W_i \leq W_{th} & i = 0, 1 \cdots, 7 \\ 1, W_i > W_{th}. \end{cases} \tag{3.11}$$

Here W_i is the Hamming weight of block Bi. W_{th} is the threshold Hamming weight, which can be set on-the-fly based on the Hamming weight distributions of the cache data trace. Only the block with a flag bit $Ii = 1$ will be applied with the differential coding. The calculation of the flag bits can be executed only when the block is written into the cache to avoid impacting the read access performance of the cache.

We simulated the Hamming weight distributions of the cache data for 4 representative benchmarks from SPEC CPU2006 benchmark suite [70] before and after applying our dynamic differential coding. The length of a cache line is 64Byte, which is divided into 8 blocks. Threshold Hamming weight W_{th} is set to 8. As shown in Fig. 29, the Hamming weight distributions of all four simulated benchmarks are successfully shifted to the left (the region with lower Hamming weight), proving the effectiveness of dynamic differential coding. The average Hamming weight reductions in the benchmarks *mcf*, *milc*, *hmmer* and *lbm* are 35%, 28%, 75% and 81%, respectively. High Hamming weight reduction is achieved particularly in *hmmer* and *lbm*, of which the data traces have a large number of similar or duplicated blocks. However, some data blocks with high Hamming weights ($W > \frac{n}{2}$ or 32) are still left after dynamic differential coding is applied. We will handle them by the proposed worst-corner-ECC scheme.

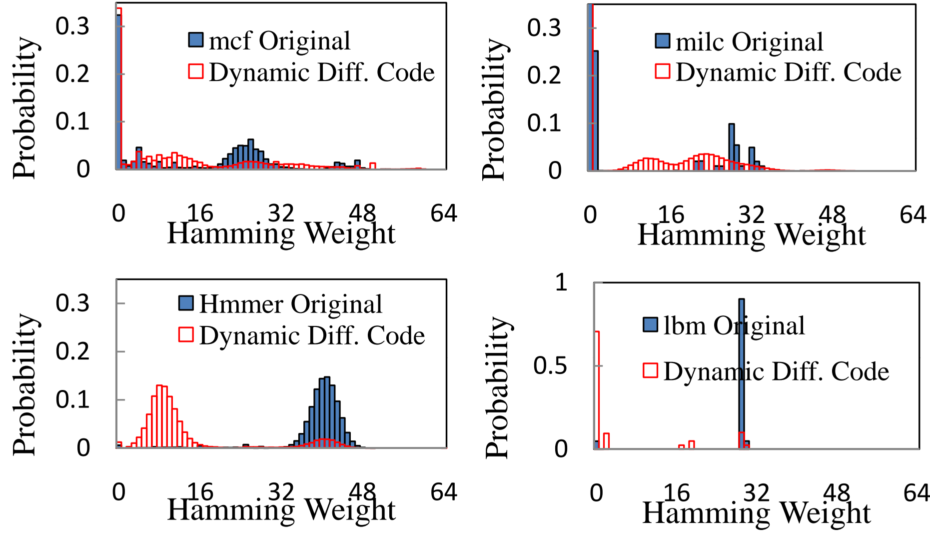

Figure 29: Simulated Hamming weight distributions comparison before and after dynamic differential coding.

3.3.1.3 Typical-Corner-ECC Design Fig. 30 illustrates the overview of our proposed typical-corner-ECC scheme using a 64Byte cache line as an example. During the initialization, the flag bit of each 64-bit data block is generated when the data is written into the cache. In the presented example, the flag bits (I1,I5,I7) are '1' while the others are '0'. Hence, Blocks $\{B1, B5, B7\}$ and $\{B0, B2, B3, B4, B6\}$ are marked as 'touch' and 'dont-touch', respectively. At step 1, dynamic differential coding on $\{B1, B5, B7\}$ and the ECC protection on the 8 flag bits are conducted simultaneously. $\{B0, B2, B3, B4, B6\}$ remain unchanged since they are marked as 'dont-touch'. The 8 flag bits can be protected by ECC1, i.e., the extended hamming code (12,8,1); At step 2, the 'dont-touch' data and the differentially coded 'touch' data will be further protected by ECC2 as a whole. Since the Hamming weight of the whole cache line has been significantly minimized, the requirement on the error correcting capability of the ECC2 is reduced.

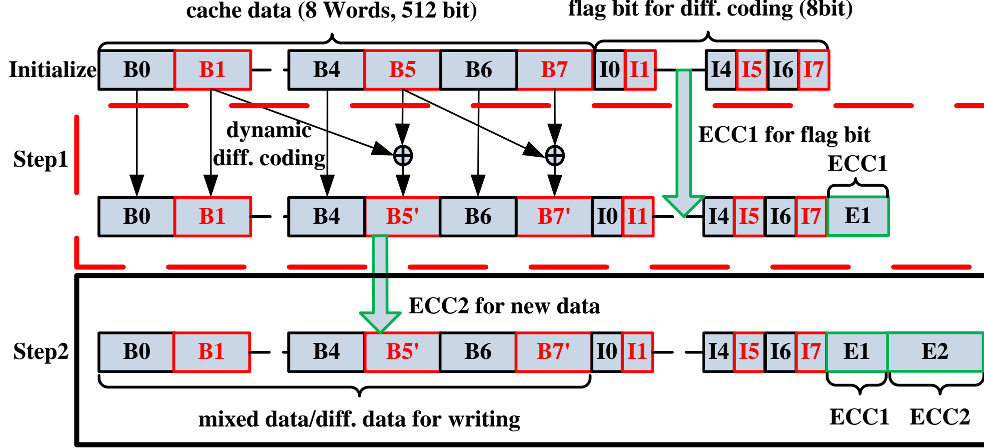

Figure 30: Overview of typical-corner-ECC.

3.3.2 Worst-Corner-ECC

As shown in Fig. 29, some data blocks with a high Hamming weight ($W > \frac{n}{2}$) or the worst corner still exist after the differential coding is applied. The ECC2 in TCE (see Fig. 30) has to cover these data blocks, leading to high hardware cost and performance overhead. *Worst-corner-ECC* (WCE) is designed to protect the data blocks with Hamming weight located in $\frac{n}{2} \leq W \leq n$. Since $P_{ER,1\to0} \ll P_{ER,0\to1}$, the correctness possibility of a codeword or a data block P_{block} can be approximately calculated as:

$$P_{block}\left(W_j, t\right) \approx \sum_{i=0}^{t} C_{W_j}^{i} P_{ER,0\to1}^{i} (1 - P_{ER,0\to1})^{W_j - i}. \tag{3.12}$$

Here n and k are the length of the post-coding codeword and the contained data, respectively. t is the error correcting capability. W_j is the Hamming weight of the codeword j, $0 \leq W_j \leq n$. As illustrated in Fig. 24, P_{block} decreases dramatically as W_j increases. In WCE, we will map the codewords with the Hamming weight between $\frac{n}{2} \leq W_j \leq n$ to the ones with the Hamming weight between $0 \leq W_j \leq \frac{n}{2}$ or even $0 \leq W_j \leq \frac{n}{4}$ in the codebook. As a consequence, the overall reliability of the codebook will be improved, followed by the reduction of the required ECC correcting capability.

3.3.2.1 The Codec of Worst-Corner-ECC The coding process of WCE can be described as follows: Assume $\mathbf{C}(n, k, t)$ is a traditional linear code set with the codeword length n, the length of information bits k and the error correcting capability t. The generator matrix $\mathbf{G}$ and the parity-check matrix $\mathbf{H}$ can be expressed as [38]:

$$G_{k\times n} = \left[Q_{k\times(n-k)}, I_{k\times k}\right].$$
$$H_{(n-k)\times n} = \left[I_{(n-k)\times(n-k)}, P_{(n-k)\times k}\right], P = Q^T. \tag{3.13}$$

Here I is the identity matrix and T is the matrix transpose. We add an extra bit '0' at the end of the information bits x to construct the input data $(x, 0)$ whose length is k. The corresponding codeword is $y_{1\times n} = (x, 0)_{1\times k} G_{k\times n}$. We use β to denote the code of which the last bit is $\beta_n = 1$ and reaches the maximum Hamming weight among the codebook $\mathbf{C}$ as:

$$\beta = \arg\max_{y\in\mathbf{C}, y_n=1} W(y). \tag{3.14}$$

Normally the Hamming weight $W(\beta)$ of β is close to n. We define the "relative weight" between a codeword y and β as:

$$W(y|\beta) = \left|\{1 \le i \le n | y_i = 1, \beta_i = 1\}\right|. \tag{3.15}$$

Then the following mapping mechanism can be applied to map a codeword with a relative Hamming weight higher than $\frac{W(\beta)}{2}$ to the one with a relative lower Hamming weight (below $\frac{W(\beta)}{2}$) as:

$$z = \begin{cases} y \oplus \beta, & if \quad W(y|\beta) > \frac{W(\beta)}{2} \\ y, & otherwise. \end{cases} \tag{3.16}$$

Note that the result of the XOR operation between any two codewords still belongs to the linear codebook $\mathbf{C}$. Hence, the maximum Hamming weight of the subset codeword z is reduced to $\left(n - \frac{W(\beta)}{2}\right)$.

z can be distorted by soft or hard errors to a different codeword z'. During the decoding, we first obtain the temporary results y by applying the decoding algorithm of a conventional ECC on the distorted codeword z' if the number of errors in z' is smaller than t. The original

data $x = \{x_1 x_2 \cdots x_{k-1}\}$ can be extracted by checking the last bit of y: If the last bit of y is '1', $x = y + \beta$; Otherwise, $x = y$.

During applications of WCE, we may encounter the problem that the length of $(x, 0)$ does not match the dimension requirement of the generator matrix $G_{k\times n}$. For example, assume we have 512 bits in a cache line, we can easily encode every 64 bits by a conventional (72,64,1) Hamming code with 8-bit overhead. In WCE, however, we need to encode 65 bits because 1 bit needs to be added to the original data x. Practically, we can remove 55 rows and 55 columns from the generator matrix of the (127,120,1) Hamming code to generate a truncated (72,65,1) Hamming code and then extend it to (73,65,1) by adding one more parity check column. By doing so, the generated (73,65,1) WCE can still correct one bit error and detect two bit errors as the (72,64,1) Hamming code does. Here the check parity matrix can be directly derived from Eq. (3.13). Note that WCE protects not only the information bits but also the postfix bits, e.g., the bit '0' appended to x. Also, WCE maintains almost the same dimensions of the generator/parity check matrix as those of SEC-DED's, incurring very low additional hardware overhead.

3.3.2.2 Efficacy of Worst-Corner-ECC To evaluate the efficacy of WCE w.r.t. conventional ECCs, we simulated the block error rate (i.e., $1 - P_{block}$, P_{block} is block reliability) based on the trace *hmmer* where the Hamming weight of the cache block data primarily locates in the worst corner. The cache block size is 64 bits.

Fig. 31 compares the simulated block error rate of the worst-corner-ECC WCE1 (73,65,1) and H64 (72,64,1) at different bit error rate $P_{ER,0\to1}$. The corresponding asymmetric ratio R is obtained from Fig. 27. WCE1 always has lower block error rate than that of H64 in the simulated $P_{ER,0\to1}$ range. Compared to H64, the block error rate improvement introduced by WCE1 is $\sim 7\times$ over the range $10^{-3} \leq P_{ER,0\to1} \leq 10^{-2}$.

Fig. 32 shows the simulated block error rate for four different WCEs and Hamming ECCs at $P_{ER,0\to1} = 5\times10^{-3}$, respectively, including WCE1 (73,65,1), WCE2 (138,129,1), WCE3 (267,257,1), WCE4 (524,513,1), H64 (72,64,1), H128 (137,128,1), H256 (266,256,1) and H512 (523,512,1). When the data block size increases, the ECC overhead decreases as only one bit error can be corrected; The error rate of Hamming ECCs becomes quickly unaffordable

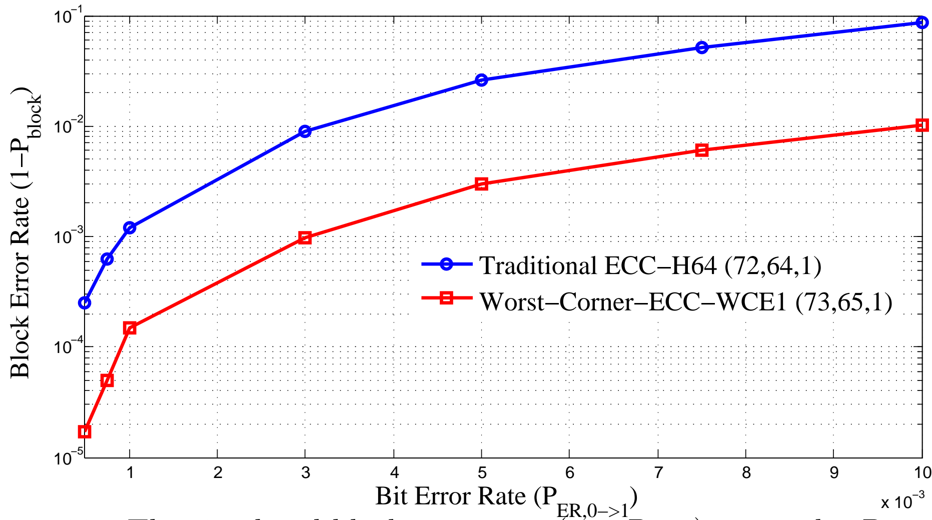

Figure 31: The simulated block error rate $(1 - P_{block})$ w.r.t. the $P_{ER,0\to1}$.

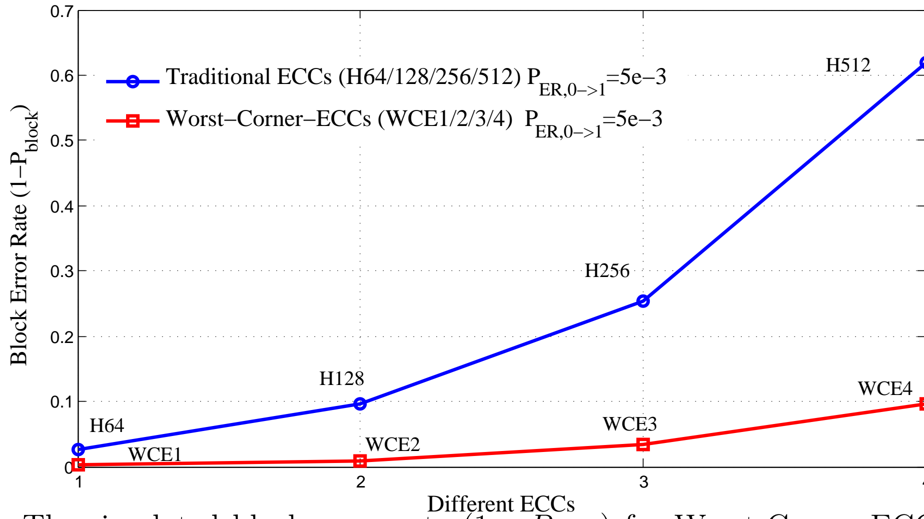

Figure 32: The simulated block error rate $(1 - P_{block})$ for Worst-Corner-ECCs and Hammings at $P_{ER,0\to1} = 5 \times 10^{-3}$.

while that of WCEs still maintains a relatively low level. As mentioned in Section 3.3.2.1, WCEs cost almost the same hardware resources compared to its corresponding hamming code, e.g., WCE1 vs. H64, WCE2 vs. H128, etc.

3.4 EVALUATION OF CD-ECC

We compare the following four different ECC schemes to evaluate the efficacy of our proposed CD-ECC technique:

1. Baseline: *H64* – pure H64 (72,64,1) Hamming code;
2. CD-ECC1: *DIFF+H64* – dynamic differential coding followed by H64 (72,64,1);
3. CD-ECC2: *DIFF+WCE1* – dynamic differential coding followed by WCE1 (73,65,1);
4. CD-ECC3: *DIFF+WCE2* – dynamic differential coding followed by WCE2 (138,129,1).

Note that here WCE1 (73,65,1) and WCE2 (138,129,1) can correct one bit error and detect two bits errors out of 73 bits and 138 bits, respectively. For dynamic differential coding, (12,8,1) Hamming code is applied to protect the flag bits.

3.4.1 Reliability

We analyze the reliability of the cache line data of 7 benchmarks from SPEC CPU2006 in our simulations. We assume the L2 cache is implemented with STT-RAM and the cache line length is 64Byte. $P_{ER,0\to1} = 5 \times 10^{-3}$ and asymmetric ratio $R = 6 \times 10^{3}$ at 10ns write pulse width. The length of each data block is 64 bits and dynamic differential coding is applied at data block level.

Fig. 33 compares the average cache line error rates of the STT-RAM based L2 cache for each benchmark under different error correction schemes. The worst error rate occurs at ‘H64’ due to the poor error correcting capability of the conventional Hamming code for asymmetric write errors. ‘DIFF+H64’ substantially improves the error rate up to 10× as dynamic differential coding significantly reduces the Hamming weight of the cache data. The only exception is *xalancbmk* which has poor data correlations among the blocks. Compared to ‘DIFF+H64’, ‘DIFF+WCE1’ further reduces the error rate by normally $2 - 5\times$ through minimizing the Hamming weight of the cache data at the worst corner. However, very marginal improvements are observed at *xalancbmk* and *omnetpp* because very few cache data is at the worst corner after dynamic differential coding is applied. Except for *xalancbmk*,

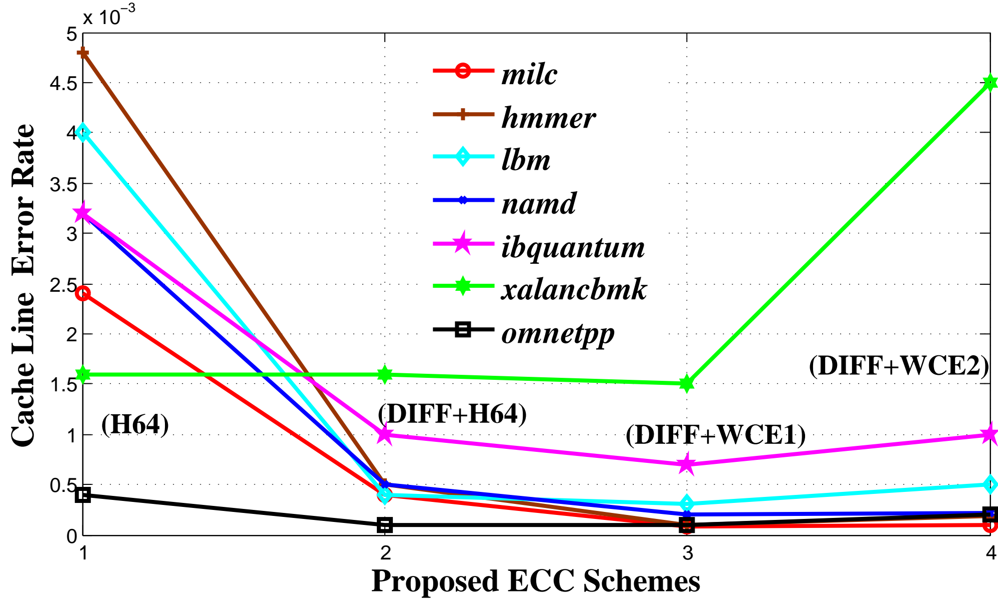

Figure 33: Cache line error rate under different schemes.

Table 4: The configuration of the microprocessor and baseline

Processor	4GHz 4 issues OOO, ROB size 256
SRAM L1 cache	32+32KB I/D, 64B line, 4-way write-back, 2-cycle read/write, 1 read +1 write ports
STT-RAM L2 cache	8 MB, 16-way, 64B line, 16 banks write-back, 16-cycle read, 48-cycle write
Main Memory	4GB, 8 channels, 16 banks, 400-cycle latency bank conflict, port contention, queuing modeled

the error rates of other benchmarks under 'DIFF+WCE2' are just slightly worse than that under 'DIFF+WCE1' due to the lower error correcting capability of WCE2.

Table 5: Delay/overhead characterization of ECC schemes

ECC Schemes	Encoding Delay	Decoding Delay	Check Bit Overhead
H64	187ps	391ps	64 (12.5%)
DIFF+H64	203ps	489ps	12+64 (14.8%)
DIFF+WCE1	417ps	547ps	12+64 (14.8%)
DIFF+WCE2	539ps	616ps	12+36 (9.4%)

3.4.2 Performance Overhead

We conduct architecture-level simulation to evaluate the impact of the write reliability of the STT-RAM based L2 cache on system performance under different ECCs. Our simulation configuration is shown in Table 4. *Macsim* [71] is used in performance evaluations. The baseline error correction scheme of our STT-RAM based L2 cache is 'H64'. The paralleled Odd-weight-column SEC-DED structure [56] is used to enhance the throughput of the 4 included ECC schemes. The synthesis results of encoding/decoding delay of the RTL implementation for all ECC schemes with TSMC 28nm process are summarized in Table 5. Table 5 also shows the storage overhead of all the ECCs. We use NVSim [72] to simulate the design parameters of an 8MB STT-RAM with 16 banks and 64Byte cache line under the 4 ECC schemes, including the latencies of codec and peripheral circuits. Table 5 shows that compared to the baseline design 'H64', the additional read latencies incurred by 'DIFF+H64', 'DIFF+WCE1' and 'DIFF+WCE2' are only (489-391)=98ps, (547-391)=156ps and (616-391)=225ps, respectively, at 28nm technology node. To be conservative, we assume 'DIFF+H64', 'DIFF+WCE1' and 'DIFF+WCE2' increase the read latency of the L2 cache by one clock cycle. The additional write latencies incurred by WCEs include the calculation of flag bit and ECC encoding, which contributes a small portion of the entire STT-RAM write latency (usually longer than 10ns). Therefore, we assume it is absorbed by the write queue operation of the L2 cache in our simulations.

The possible cache access errors can be categorized into the following four cases:

- CASE 1: The correctable errors, i.e., zero or one error, occur on a read hit of the L2 cache line;
- CASE 2: The uncorrectable errors, i.e., two errors, are detected in a 'clean' L2 cache line;
- CASE 3: The uncorrectable errors, i.e., two errors, are detected in a 'dirty' L2 cache line;
- CASE 4: The unrecoverable errors, i.e., three or more errors, happen in an L2 cache line.

In CASE 1, the corresponding L2 cache line is directly corrected by the ECCs and sent back to L1 cache; In CASE 2, the 'clean' L2 cache line can be recovered by fetching a correct copy from the main memory; In CASE 3, the detected 'dirty' L2 cache line is unrecoverable, thus causing a failure of system simulation. We do not focus on CASE 4 since it is very rare and beyond the capability of all compared ECCs. We repeat 200 simulations on each benchmark and find that the average ratio of 'clean' and 'dirty' line is roughly 1:1. In each simulation, we fast-forward to the ROI (Region of Interest), warm up the cache with 200 million instructions, and then run 500 million instructions. Not a single event of CASE 4 is captured in all simulations. However, 'DIFF+H64', 'DIFF+WCE1' and 'DIFF+WCE2' can reduce the occurrence of CASE 3 by up to 10x, 30x and 21x on average w.r.t. that of 'H64' (on average 14 failures per simulation), respectively.

Fig. 34 shows the normalized *Instruction Per Cycle* (IPC) of all the benchmarks with different ECC schemes, w.r.t. the results of our baseline 'H64'. The average IPC degradation of all the three CD-ECCs are merely less than 0.5% compared to that of 'H64'. 'DIFF+WCE1' achieves the lowest average IPC performance degradation (0.265% w.r.t. 'H64') by minimizing the number of CASE 2 and the accesses to the main memory during the execution. As shown in Table 5, the check bit overhead of 'DIFF+WCE1' is the same as that of 'DIFF+H64'(76 bits). 'DIFF+WCE2' offers the similar IPC performance to 'DIFF+H64', however, with much less check bit overhead (48 bits vs. 76 bits).

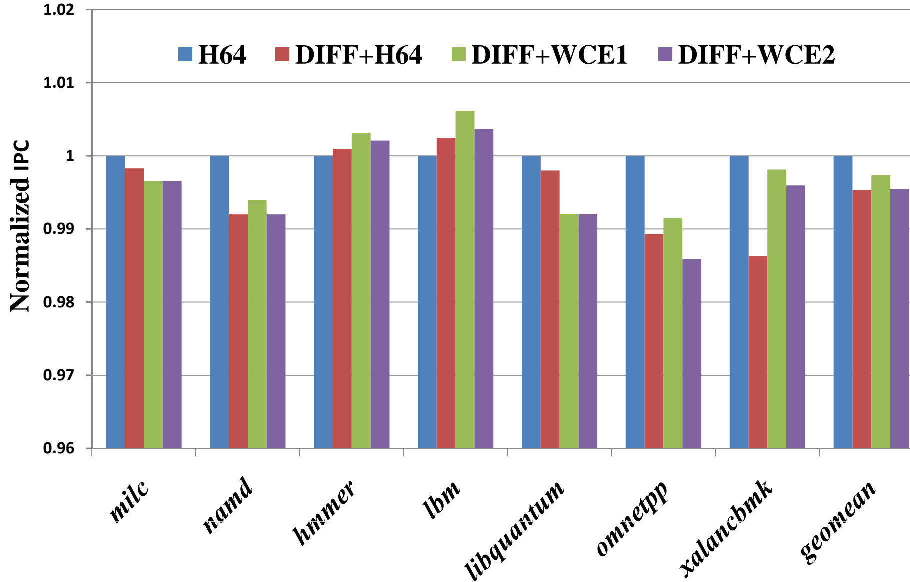

Figure 34: Normalized IPC of each benchmark under different schemes.

3.5 CHAPTER 3 SUMMARY

In this chapter, we proposed an *analytic write channel* (AWC) model to systematically analyze the STT-RAM asymmetric operation errors. We then developed *content-dependent ECC* (CD-ECC) technique, including two ECC schemes, namely, *typical-corner-ECC* (TCE) and *worst-corner-ECC* (WCE) to fix the asymmetric STT-RAM write errors based on the different bit-flipping distributions of the data. Simulations show that compared to conventional ECCs, our techniques can improve the reliability of the cache data in the simulated computer systems by $10-30\times$ with low hardware cost and very marginal ($< 0.5\%$) *Instruction Per Cycle* (IPC) degradation.

4.0 STATE-RESTRICT MLC STT-RAM DESIGNS FOR HIGH-RELIABLE HIGH-PERFORMANCE MEMORY SYSTEM

As discussed in chapter 1, *multi-level cell Spin-Transfer Torque Random Access Memory* (MLC STT-RAM) is a promising nonvolatile memory technology for high-capacity and high-performance applications. However, the reliability concerns and the complicated access mechanism greatly hinder the MLC STT-RAM application under current technology node. In this chapter, we will develop a holistic solution set, namely, *state-restrict MLC STT-RAM* (SR-MLC STT-RAM) to improve the data integrity and performance of MLC STT-RAM with the minimized information density degradation.

The structure of this chapter is organized as the follows: Section 4.1 presents the MLC STT-RAM basics and research motivation; Section 4.2 gives the details on the three circuit-level techniques: *state restriction* (StatRes), *error pattern removal* (ErrPR), and *ternary coding* (TerCode) for read and write reliability enhancement; Section 4.3 describes the architecture technique–*state pre-recovery* (PreREC) for write latency improvement; Section 4.4 illustrates our experimental results; Section 4.5 gives the conclusion of this chapter. As this chapter only focuses on STT-RAM designs, the term "STT-RAM" may be ignored in the following statement (*e.g.*, MLC STT-RAM cells vs. MLC cells).

4.1 BACKGROUND AND MOTIVATION

4.1.1 MLC STT-RAM Basics

As presented in Chapter 2, Fig. 35(a) shows the basic structure a SLC STT-RAM, which can store only one bit per cell. Raising the amplitude of the programming current will speedup the switching of the MTJ [73].

Fig. 35(b) shows a proposal of 2-bit MLC cell which is adopted in this work. Two MTJs with different sizes are stacked vertically atop an NMOS transistor. The four resistance states are defined by the four combinations of relative MDs of the two MTJs. Since the resistance-area product (RA) remains the same in both MTJs, the resistance of the small MTJ is higher than that of the large one. Also, the small MTJ will experience a higher current density than the large one when the programming current is applied, leading to a faster switching speed.

Based on the difference between the switching speed of the two MTJs, a two-step write scheme is designed for MLC cells, as depicted in Fig. 35(c). The resistance state transitions of an MLC cell are classified into three types: 1) *soft transition* (ST), which switches only the

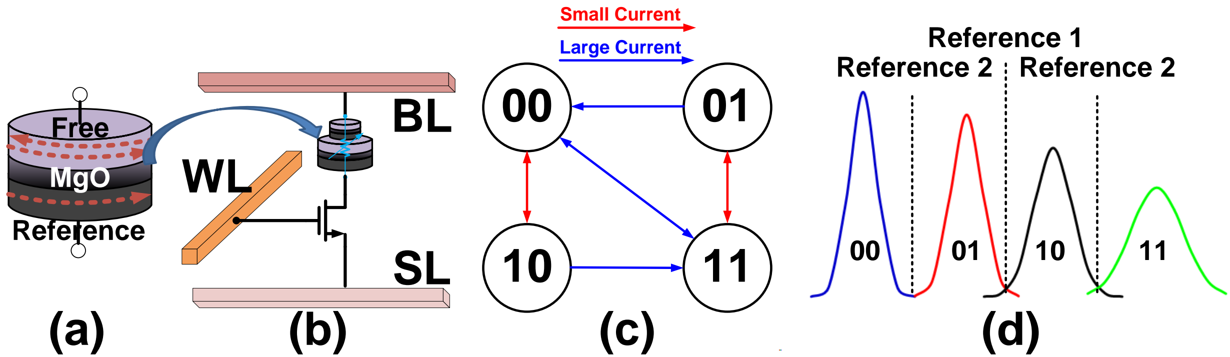

Figure 35: Illustrations of (a) MTJ. (b) MLC STT-RAM cell. (c) Two-step write scheme. (d) Two-step read scheme.

small MTJ with a low programming current (*e.g.*, '11'→'01'); 2) *hard transition* (HT), which switches both MTJs with a large current (*e.g.*, '00'→'11'); and 3) *two-step transition* (TT), which flips the resistance state of only the large MTJ in two steps, i.e., one HT followed by one ST ('00'→'11'→'01'). Based on the programming difficulty of the two data bits in an MLC cell, we name the first bit that is decided by the resistance state of the small MTJ as "soft-bit" while the second bit that is decided by the resistance state of the large MTJ as "hard-bit". As shown in Fig. 35(d), reading an MLC cell requires three references to detect both the hard-bit and soft-bit.

4.1.2 Reliability of MLC STT-RAM Cells

The performance and operation reliability of MLC cells are heavily influenced by the process variations and the thermal fluctuations in MTJ switching process.

4.1.2.1 Write errors of MLC STT-RAM In an SLC cell, write errors happen only when the programming current is removed before the MTJ switching process completes. Raising the amplitude of programming current can effectively reduce the MTJ switching time and improve the write reliability of the SLC cell. In an MLC cell, however, raising the amplitude of programming current during a soft transition (e.g., switching the resistance state of the small MTJ only) may also increase the probability of overwriting the cell, *e.g.*, accidently flipping the resistance state of the big MTJ. Thus, the write operation of MLC cells generally has a higher write error rate than that of SLC cells.

4.1.2.2 Read errors of MLC STT-RAM The read errors of an STT-RAM cell can be categorized into two types: sensing error and read disturbance.

Sensing error denotes the scenario that the MTJ resistance state cannot be read out correctly before the sensing period ends. It is usually because of the small or even false sense margin incurred by the process variations of the MTJ and NMOS transistor. Hence, maintaining a sufficient distinction between two adjacent resistance states becomes essential in STT-RAM cell designs. MLC cells, however, generally have smaller sense margin than

SLC cells: in Fig. 35(d), the distinction between resistance states R_{00} and R_{10} equals the difference between the high and the low resistance states of the (small) MTJ in an SLC cell. In an MLC cell, however, such a distinction is further partitioned into two smaller parts by an intermediate state R_{01}, resulting in smaller sense margins between the adjacent states.

Read disturbance denotes the scenario that the read current accidently flips the resistance state of the MTJ under the impact of thermal fluctuations. In MLC cell designs, the read error is mainly dominated by the sensing errors as the read current amplitude is always selected as low as possible to suppress the disturbance of both MTJs. In this work, we ignore the read disturbance in our analysis.

4.1.2.3 Practicability of ECC schemes A common approach to combat the operation errors of conventional MLC STT-RAM (C-MLC STT-RAM) is using error correction code (ECC). To evaluate the efficacy of different ECCs, we use "cell-level error rate" (P_{CER}) to denote the error probability of a memory cell. The designed ECC can be represented as (n, k, t). Here n and k are the post-coding and pre-coding length, respectively. t is the error correcting capability of the ECC, i.e., how many erroneous bits can be repaired among the n bits. We also assume the n bits are stored in m memory cells, e.g., $m = n/2$ for a 2-bit MLC STT-RAM. The probability that the n bits can not be corrected by the ECC can be calculated as:

$$P_{fail}\left(nbit\right) = 1 - \sum_{i=0}^{t} C_m^i P_{CER}^i \left(1 - P_{CER}^i\right)^{m-i}. \tag{4.1}$$

In a 4MB C-MLC STT-RAM based cache with 2^{16} 64B cache lines, if we assume the target yield is 99.9%, the maximum allowable failure rate of each cache line is $\sim 1.5 \times 10^{-8}$. However, as technology continues scaling down, the read errors induced by the increased process variations and degraded sensing margin will emerge and dominate the reliability of MLC STT-RAM systems. The two types of read errorssensing error and read disturbance are heavily correlated. On the one hand, increasing sensing accuracy and hence reducing sensing errors require prolonging the sensing period. On the other hand, the possibility of read disturbance increases quickly as the sensing time increases. Our recent analysis showed that the lowest read error rate of an MLC cell that can be achieved under the current manufacturing condition is merely $P_{CER} = 1.57 \times 10^{-2}$ [20]. As shown in Figure 36, if

four different ECCs [74], i.e., H64-1 (72,64,1), BCH-8 (592,512,8), BCH-16 (672,512,16), and BCH-24 (752,512,24), are applied, the error rates of the 64B cache line become 6.06×10^{-1} (H64-1), 4.61×10^{-2} (BCH-8), 3.14×10^{-5} (BCH-16) and 2.66×10^{-9} (BCH-24), respectively. The results indicate that only an extremely strong ECC like BCH-24 is sufficient to deliver the required yield. However, BCH-24 requires 240 redundant bits on top of 512 data bits (or 47% spare cells) and a decoding latency of nearly ($7 - 10$ns) [74]. Such high design and performance overheads are apparently unaffordable in conventional embedded and on-chip applications. In other words, the error rate of an MLC cell must be significantly reduced (i.e., down to $\sim 10^{-6}$) so that the error rate of a 64B cache line can be improved to $\sim 1.5 \times 10^{-8}$. However, the authors in [20] have shown that such a low error rate cannot be solely achieved through the circuit design optimizations of MLC cells.

4.2 SR-MLC STT-RAM DESIGN

Our proposed *state-restrict MLC* STT-RAM (SR-MLC) design includes three circuit-level techniques – *state restriction* (StatRes), *error pattern removal* (ErrPR), and *ternary coding*

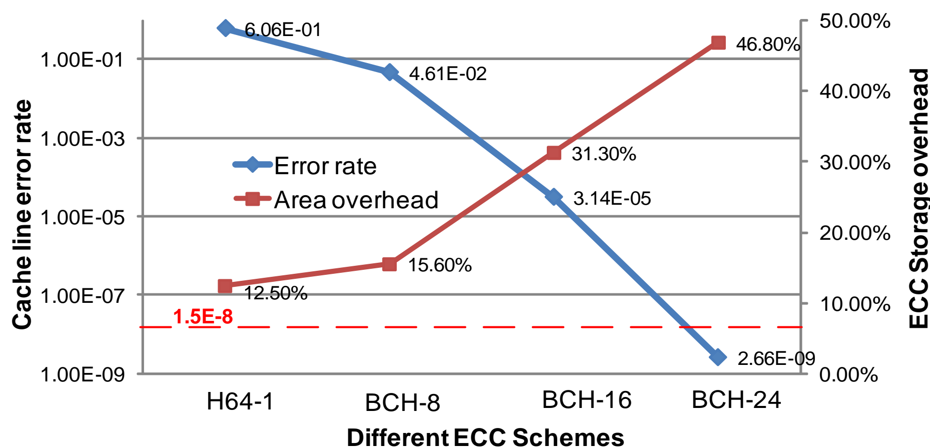

Figure 36: Comparison of different ECCs.

(TerCode). StatRes focuses on the reduction of read errors while ErrPR and Tercode aim at the suppression of the write errors of SR-MLC cells and the fast conversion between binary data and ternary storage system, respectively.

4.2.1 State Restriction (StatRes)

4.2.1.1 Basic concept of state restriction Fig. 37(a) shows the conceptual overview of the distributions of the four resistance states of a C-MLC cell under the influences of process variations. The references utilized in the read operations are usually selected between two adjacent states. If the distributions of two resistance states overlap with each other, read errors will be generated. The overlap area is mainly determined by the distinction between the adjacent resistance states and the distributions of the states. In an MLC cell, the read reliability is limited by the largest resistance state overlap. Hence, if we can eliminate one resistance state (ideally, the most error-prone one) out of the four states and use only the rest three states to store the data, the read reliability of the MLC cell will be improved. We name this technique as *state restriction (StatRes)* .

Fig. 37(b) depicts a state restriction example by eliminating resistance state 'R_{10}'. Before removing 'R_{10}', the largest overlaps occur at the two sides of 'R_{10}' distribution. The removal of 'R_{10}' does not only eliminate the two overlaps of 'R_{10}' vs. 'R_{01}' and 'R_{10}' vs. 'R_{11}' but also introduces a wide distinction between 'R_{01}' and 'R_{11}'. Apparently, removing the boundary resistance states – 'R_{00}' or 'R_{11}' gives us less benefit than removing intermediate resistance states – 'R_{01}' or 'R_{10}' as it does not increase the distinction between the states. As we will prove in next Section, removing resistance state 'R_{10}' offers the best enhancement on the read reliability of MLC cells among all state restriction options. Note that the data storage capacity of the MLC STT-RAM cell is reduced by 1/4 due to the removal of the state.

4.2.1.2 Optimization of StatRes In a 2-bit MLC cell, the low (R_{1L}) and the high resistance state (R_{1H}) of the small MTJ generally follow Gaussian distribution. If we define $\beta = R_{1H}/R_{1L} = 1 + TMR$, then:

$$R_{1L} \sim N(\mu, \sigma), R_{1H} \sim N(\beta\mu, \beta\sigma). \quad (4.2)$$

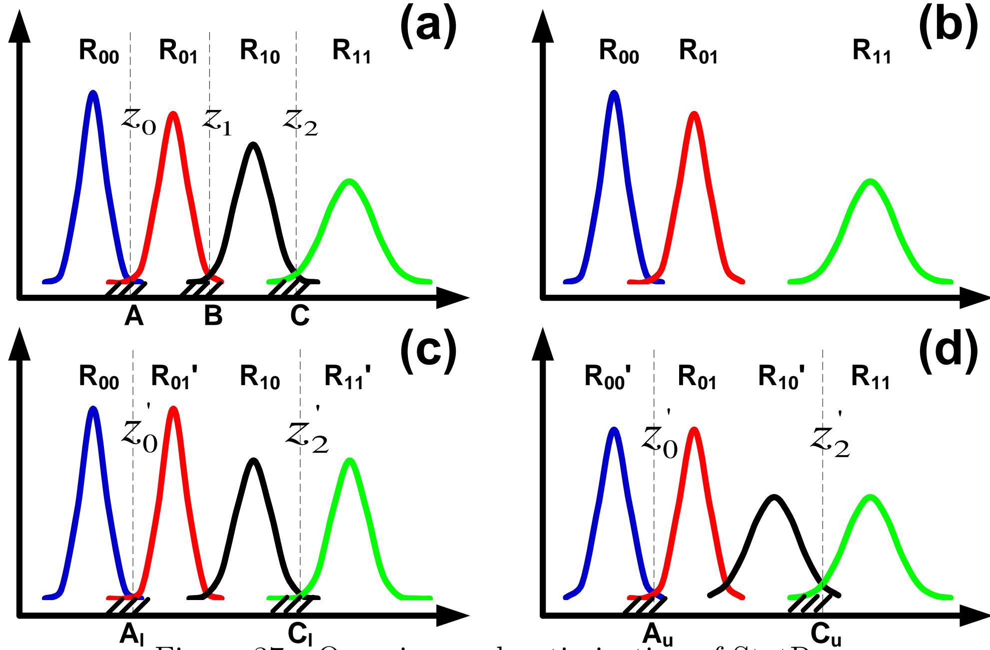

Figure 37: Overview and optimization of StatRes.

Here TMR is tunnel magnetoresistance ratio, which is the ratio between the difference of the two resistance states and the low resistance state. Normally $TMR = 1 \sim 1.2$ or $\beta = 2 \sim 2.2$.

As the large MTJ in the MLC cell shares the same manufacturing process and geometry ratio with the small MTJ, the distributions of the resistance states of the large MTJ can be expressed as:

$$R_{2L} \sim N\left(\frac{1}{k}\mu, \frac{1}{k}\sigma\right), R_{2H} \sim N\left(\frac{\beta}{k}\mu, \frac{\beta}{k}\sigma\right). \tag{4.3}$$

Here k is the ratio between the cell areas of the large and the small MTJs. We can derive the resistance state distributions of the two MTJs in a C-MLC cell as:

$$\begin{aligned} R_{00} &\sim N\left((1+1/k)\,\mu, \sqrt{1+1/k^2}\sigma\right), \\ R_{01} &\sim N\left((1+\beta/k)\,\mu, \sqrt{1+\beta^2/k^2}\sigma\right), \\ R_{10} &\sim N\left((\beta+1/k)\,\mu, \sqrt{\beta^2+1/k^2}\sigma\right), \\ R_{11} &\sim N\left((1+1/k)\,\beta\mu, \sqrt{1+1/k^2}\beta\sigma\right). \end{aligned} \tag{4.4}$$

Since $\beta = 2 \sim 2.2$ and $k > 1$, we can easily obtain the following inequalities for their means and standard deviations, respectively:

$$\mu_{R_{00}} < \mu_{R_{01}} < \mu_{R_{10}} < \mu_{R_{11}},$$
$$\sigma_{R_{00}} < \sigma_{R_{01}} < \sigma_{R_{10}} < \sigma_{R_{11}}. \tag{4.5}$$

As shown in Fig. 37(a), the maximum read error rate of a C-MLC cell, which decides the required ECC strength, is determined by the largest overlap among the overlaps A, B and C. Removing resistance state 'R_{01}' or 'R_{10}' will eliminate the overlaps $A+B$ or $B+C$, respectively. Hence, the read error rate of the corresponding SR-MLC cell is limited by the area of overlap C or A, respectively.

The area of overlap A can be calculated by:

$$A = \min\left(\int_{z}^{\infty} f_{R_{00}}(R)dR + \int_{0}^{z} f_{R_{01}}(R)dR\right). \tag{4.6}$$

Here $f_{R_{00}}(R)$ and $f_{R_{01}}(R)$ are the probability distribution functions of resistance states R_{00} and R_{01}, respectively. The optimum reference z_0 for the minimized value of A can be derived from $\frac{dA}{dz} = 0$.

Directly calculating and comparing the minimum areas of overlaps A and C are generally impractical. Instead, we construct the lower and upper bounds of the areas of A and C – $[A_l,\ A_u]$, and $[C_l,\ C_u]$ by introducing four assisting resistance states $R'_{00} \sim N(\mu_{R_{00}}, \sigma_{R_{01}})$, $R'_{01} \sim N(\mu_{R_{01}}, \sigma_{R_{00}})$, $R'_{10} \sim N(\mu_{R_{10}}, \sigma_{R_{11}})$, and $R'_{11} \sim N(\mu_{R_{11}}, \sigma_{R_{10}})$, as shown in Fig. 37(c) and (d). As an example, the lower bound of A – A_l is defined as the overlap between R_{00} and R'_{01}, both of which have the same standard deviation $\sigma_{R_{00}}$. Therefore, the corresponding optimum reference $z'_0 = (\mu_{R_{00}} + \mu_{R_{01}})/2 = (1 + (\beta+1)/2k)\,\mu$. The area of A_l can be calculated as:

$$A_l = \frac{1}{4} erfc\left(\frac{\mu_{R_{01}} - \mu_{R_{00}}}{2\sqrt{2}\sigma_{R_{00}}}\right) = \frac{1}{4} erfc\left(\frac{(\beta-1)\,\mu/2k}{2\sqrt{2}\sigma_{R_{00}}}\right) \tag{4.7}$$

Here $erfc\,(x) = (2/\sqrt{\pi}) \int_x^\infty e^{-t^2} dt$. Eq. (4.7) implies that A_l decreases as the ratio between $(\mu_{R_{01}} - \mu_{R_{00}})$ and $\sigma_{R_{00}}$ grows. Following the same routine, we can derive the expressions of A_u, C_l, and C_u, which satisify $A_l < A_u < C_l < C_u$. Considering $A_l < A < A_u$ and $C_l < C < C_u$, we conclude that the area of A is always smaller than that of C.

Hence, removing resistance state 'R_{10}' is more helpful than removing 'R_{01}' to minimize the read error rate of an SR-MLC cell. In our design, we restrict resistance state 'R_{10}' for the aforementioned optimizations.

4.2.2 Error-pattern Removal (ErrPR)

Besides proposing StatRes technique for read reliability enhancement, we also develop *error-pattern removal* (ErrPR) technique to improve the write reliability of SR-MLC cells. In the following analysis, we use the notation of $[X, Y]$ = [soft-bit, hard-bit] to denote the binary data stored in an SR-MLC cell.

4.2.2.1 Basic concept of ErrPR In general, the occurrences of the write errors of a C-MLC cell can be summarized as:

1. In a soft transition $[X, Y] \rightarrow [\bar{X}, Y]$, the erroneous state after the write will be either $[X, Y]$ (incomplete write error) or $[\bar{X}, \bar{Y}]$ (overwrite error);
2. In a hard transition, we need to consider two scenarios: 1) if $X = Y$, the correct transition will be $[X, Y] \rightarrow [\bar{X}, \bar{Y}]$. The erroneous state after the write will be either $[X, Y]$ (incomplete write errors at both bits) or $[\bar{X}, Y]$ (incomplete write error at hard bit); 2) if $X \neq Y$, the correct transition will be $[X, Y] \rightarrow [X, \bar{Y}]$. The erroneous state after the write will be $[X, Y]$ (incomplete write error at hard bit).
3. In a two-step transition composed of a hard transition and a soft transition, the erroneous states after the write can be derived from the above two cases.

Fig. 38(a) summarizes the *ten* possible erroneous transitions in the writes of a C-MLC cell. StatRes technique reduces the number of possible erroneous transitions from *ten* to *six* by removing state '10', as illustrated in Fig. 38(b). However, we can further reduce the number of possible erroneous transitions down to *two* by introducing error-pattern removal (ErrPR) technique to optimize the writes of SR-MLC cells as follows (see Fig. 38(d)):

1. CASE 1: In a state transition from any arbitrary states to '00' or '11', i.e., 'XX'→'00'/'11', a large programming current $I_{00}^{EPC}/I_{11}^{EPC}$ is applied;

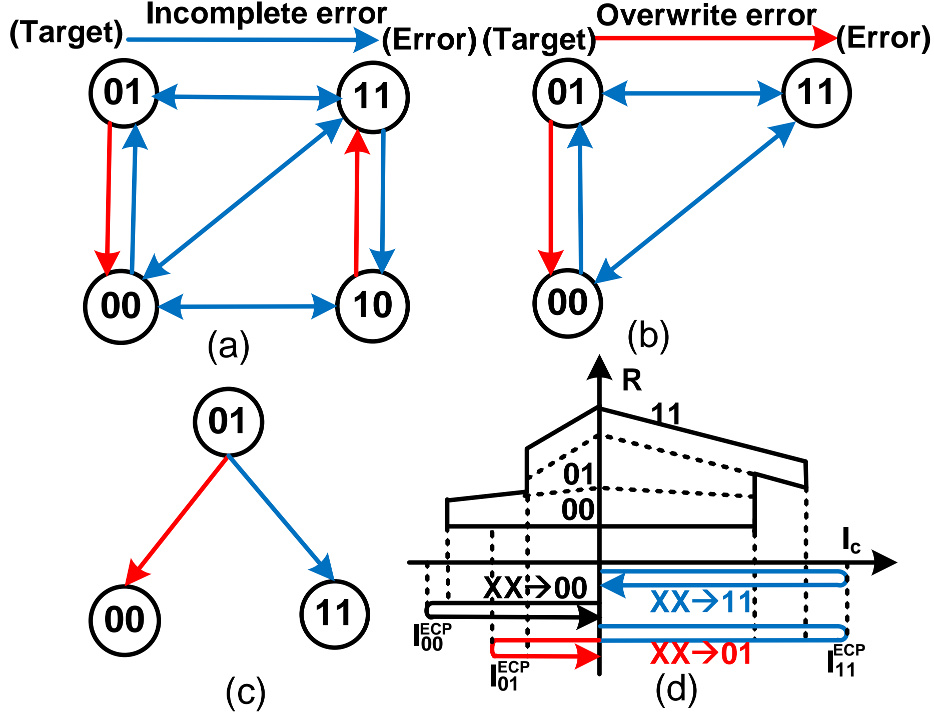

Figure 38: (a) 10 error patterns of C-MLC, (b) 6 error patterns of SR-MLC, (c) 2 error patterns of SR-MLC with ErrPR, (d) Overview of ErrPR.

2. CASE 2: In a state transition from any arbitrary states to ‘01’, i.e., ‘XX’→‘01’, the state of the SR-MLC cell is first switched to ‘11’ under the current I_{11}^{EPC} and then switched to ‘01’ under a current I_{01}^{EPC}. The whole transition can be expressed as ‘XX’→‘11’→‘01’.

As shown in Fig. 38(c), there are no overwrite errors existing in CASE 1. Thus, the amplitude of the programming current I_{00}^{EPC} or I_{11}^{EPC} can be raised to a sufficiently large level to suppress the incomplete write errors. In CASE 2, the write errors are dominated by the ones occurring in the ‘11’→‘01’ transition if the errors in the ‘XX’→‘11’ transition are well suppressed by carefully choosing I_{11}^{EPC}. The optimum amplitude of I_{01}^{EPC} can be found to minimize both the incomplete write and overwrite errors in the transition of ‘11’→‘01’.

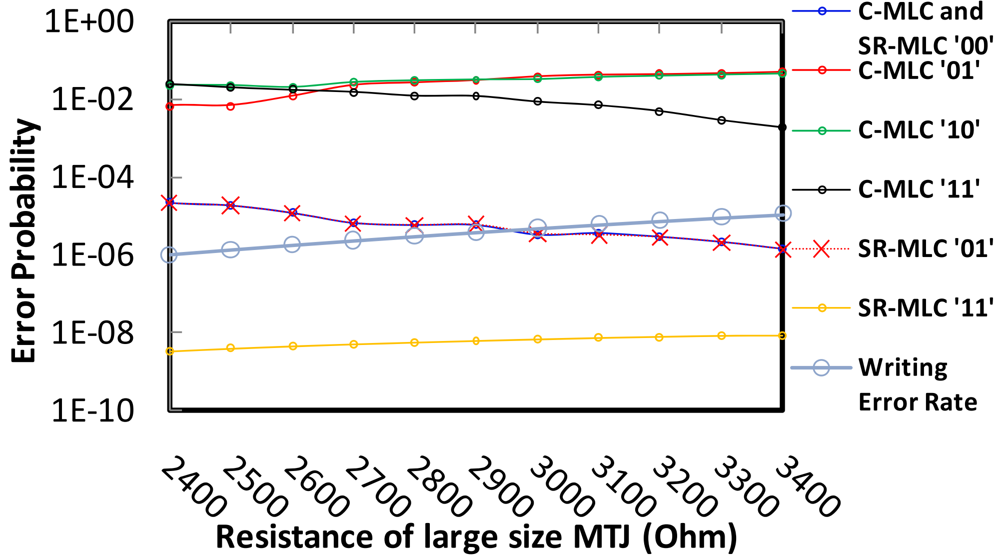

Figure 39: Error rate comparison of SR-MLC vs C-MLC cells

4.2.2.2 Reliability evaluation of SR-MLC with ErrPR To evaluate the effectiveness of SR-MLC cell design with ErrPR technique, we designed the SR-MLC and C-MLC STT-RAM designs based on 45nm PTM model [50]. The MTJ pillar has a 45nm×90nm elliptical shape while SHR-MLC structure [40] is adopted to offer a 2× capacity of SLC STT-RAM at the same area. The simulation setup and the assumptions on process variations and thermal fluctuations are adopted from [37, 20]. Fig. 39 compares the read error rates of the different states of a C-MLC cell and an SR-MLC cell when the size (represented as the low resistance state R_{2L}) of the large MTJ changes. Here we have $\beta = 2$ and $k = 1.5$.

Compared to C-MLC, SR-MLC improves the cell-level read error rate (the worst case among all states) by about 10^4 (i.e., 1.5×10^{-2} of state R_{11} at $R_{2L} = 2600\Omega$ in C-MLC vs. 1×10^{-6} of state R_{01} at $R_{2L} = 3000\Omega$ in SR-MLC). Removing state R_{01}, however, does not really improve the read reliability of MLC STT-RAM cells as the read error rate is still constrained by the state R_{11}. It indeed agrees with our theoretical analysis on the impact of removing R_{01} or R_{10} in Section 4.2.1.2.

The write error rate of an SR-MLC cell with ErrPR, which is constrained by state R_{01}, is also shown in Fig. 39. Here the write pulse width is set to 10ns. We note that raising

R_{2L} beyond 3000Ω may further decrease the read error rate of the SR-MLC cell but also increase its write error rate. Hence, we select $R_{2L} = 3000\Omega$ as the optimized point for both read and write error rates. The corresponding three optimized programming currents are $\left|I_{00}^{EPC}\right| = 185.6\mu A$, $|I_{11}^{EPC}| = 183.2\mu A$, and $\left|I_{01}^{EPC}\right| = 118.3\mu A$, respectively. The cell-level write error rate is successfully minimized under 5×10^{-6}. As a comparison, the original write error rate of C-MLC cells, which is constrained by either R_{01} or R_{10}, is about 4.3×10^{-5} [20].

4.2.3 Ternary Coding (TerCode)

The design of SR-MLC STT-RAM needs to combine two tri-state SR-MLC cells to represent 3 binary bits. Hence, we propose an efficient coding/decoding technique – *ternary coding* (TerCode), to convert binary data to ternary states of SR-MLC cells and vice verse.

Among all 9 combined states of the two SR-MLC cells, '0101' suffers from the highest programming errors as state '01' is associated with two possible erroneous transitions (see Fig. 38(c)). Hence, our TerCode technique excludes '0101' and only adopts the other 8 combined states to represent the 3 binary bits. Table 6 depicts the optimized mapping relationship between binary data and ternary states in TerCode scheme. Here the 3-bit binary data ($b_2b_1b_0$) and the combined ternary state of two SR-MLC cells ($s_3s_2s_1s_0$) are all coded with Gray code. TerCode technique guarantees that there is at most one erroneous state transition will be generated during any writes due to the removal of the combined state '0101'. As illustrated in Fig. 40, one erroneous state transition of the two SR-MLC cells produces only one bit error in the corresponding 3-bit binary data, resulting in a minimized write error rate.

The mathematic expression of our proposed TerCode technique is summarized in Eq. (4.8). As the critical path of the Codec is only about two gates, the impact of the additional latency of TerCode technique on system performance is negligible.

$$
\begin{aligned}
b_0 &= \widetilde{s_3}\widetilde{s_1}\widetilde{s_0} + s_2\widetilde{s_1}\widetilde{s_0} \\
b_1 &= \widetilde{s_3}\widetilde{s_1}\widetilde{s_0} + s_2 s_1 s_0, \quad b_2 = s_3 + s_2 s_1 \\
s_3 &= b_2\widetilde{b_1} + b_2\widetilde{b_0}, s_2 = b_2 + b_1\widetilde{b_0} \\
s_1 &= \widetilde{b_2}\widetilde{b_0} + b_1\widetilde{b_0}, s_0 = \widetilde{b_2}\widetilde{b_1} + \widetilde{b_0}.
\end{aligned}
\tag{4.8}
$$

Table 6: Binary-to-Ternary storage mapping

3-bit data $b_2b_1b_0$	$cell_1$ s_3s_2	$cell_0$ s_1s_0	3-bit data $b_2b_1b_0$	$cell_1$ s_3s_2	$cell_0$ s_1s_0
011	00	00	001	00	01
000	00	11	010	01	11
110	11	11	100	11	01
101	11	00	111	01	00

4.3 STATE PRE-RECOVERY (PREREC)

Although the SR-MLC cell design with ErrPR technique presented in Section 4.2 successfully suppresses the access error rates of MLC STT-RAM, a costly two-step programming, i.e., 'XX'→'11'→'01', is still required. In this section, we propose *state pre-recovery* (PreREC) technique – an architectural solution that can virtually remove the two-step programming of SR-MLC STT-RAM during system executions, achieving an access performance similar

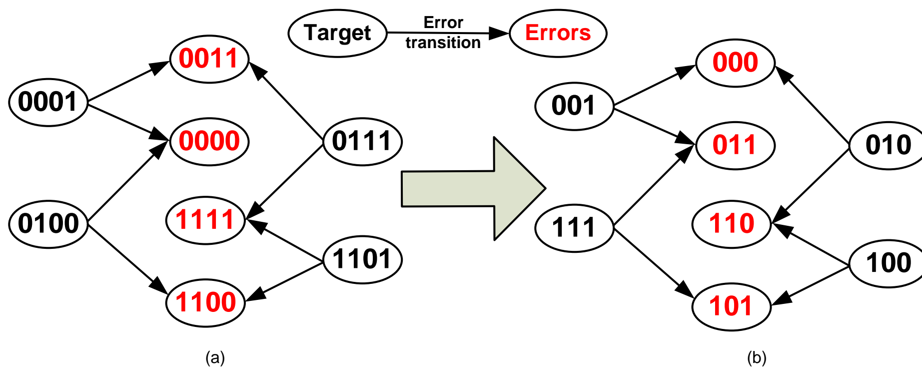

Figure 40: (a) Error patterns of the state transitions of two SR-MLC cells, (b) Error patterns mapped to the 3-bit binary data.

to SLC STT-RAM. For illustration purpose, we assume a cache hierarchy same as [75] in our analysis. It includes an SRAM based L1 cache and an SR-MLC based L2 cache, both of which are writeback.

4.3.1 Motivation of PreREC

PreREC is motivated by the key observation that if the initial state of an SR-MLC cell is '11', then programming the cell to any state just takes at most one step, as discussed in Section 4.2.2.1. Hence, if we can pre-recover the states of all SR-MLC cells in a cache line to '11' before a write access, the write of the whole cache line can complete with one-step programming. Note that PreREC is NOT applicable to C-MLC cells because in a C-MLC cell, the programming's of states '01' and '10' have different intermediate states, i.e., 'XX'→'11'→'01' and 'XX'→'00'→'10'.

4.3.2 Design of PreREC

Obviously, pre-recovery needs to be performed on the cache line before the write starts. To avoid destroying the useful cache line data, pre-recovery must be also conducted only when the data in the L2 cache line is recognized as stale and will not be needed. Hence, we proposed to perform pre-recovery in only the following three cases:

1. CASE 1: A write hit to a clean L1 cache line;
2. CASE 2: A write miss to L1 cache followed by a read hit to the corresponding L2 cache line;
3. CASE 3: A write miss to L1 cache followed by a read miss to L2 cache.

In CASE 1, the L2 cache line corresponding to the hit L1 cache line will not be accessed before the L1 cache line is evicted. Hence, pre-recovery can be performed on the L2 cache line as soon as the L1 cache line becomes dirty but no later than the L1 cache line is written back to the L2 cache. We note that pre-recovery needs to be performed only once on the L2 cache line though the L1 cache line may be written several times before being evicted.

In CASE 2, the L2 cache line will be loaded to the L1 cache. Pre-recovery can start as soon as the hit L2 cache block is written to the L1 cache. Compared to CASE 1, the

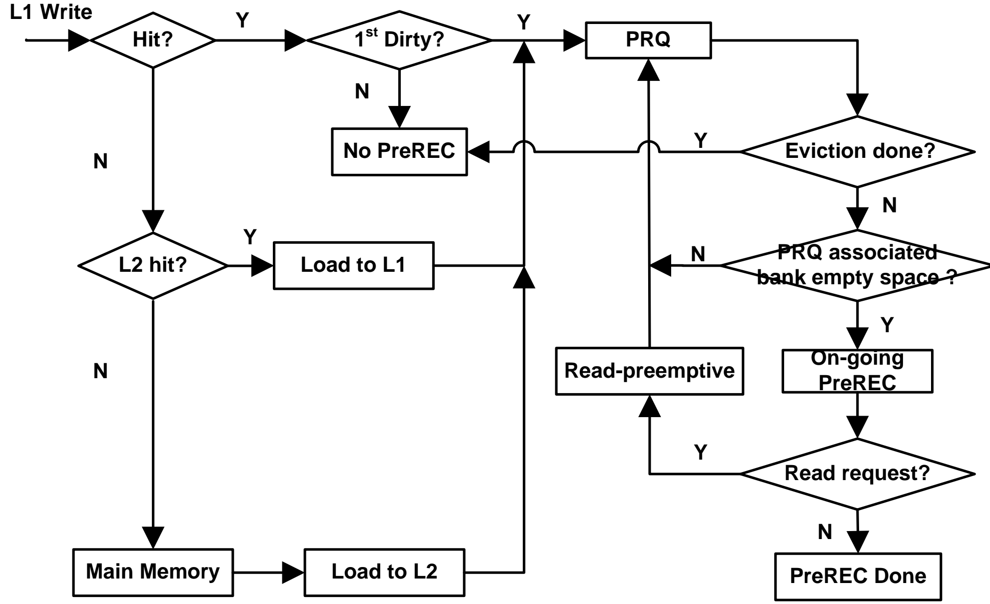

Figure 41: Overview of PreREC.

performing time window of pre-recovery is wider in CASE 2 as the updated L1 cache line is unlikely to be evicted in near future;

In CASE 3, the cache line will be loaded from the main memory. Pre-recovery can be performed on the L2 cache line anytime after the read miss occurs with an even wider performing window.

Fig. 41 illustrates the operation flow of PreREC. A PreREC buffer queue (PRQ) is introduced to store the address of the cache lines that are performed with pre-recovery. Besides the address, only a 5-bit counter is included in PRQ to track the status of the pre-recovery operation, which lasts 23 cycles at 1.8GHz clock frequency (or 12.7ns including peripheral circuit delay and MTJ programming time). The hardware design cost of the PRQ is very small. Two flag bits (F_1F_2) are also added to the tag entry of each cache block to guide pre-recovery operations. In particular, F_1F_2 are set based on the status of the pre-recovery operation as: 1) F_1F_2 = '01' for an on-going pre-recovery (counter=00001); 2) F_1F_2 = '11' for a complete pre-recovery (counter=11000); and 3) F_1F_2 = '00' for no or pending pre-recovery (counter=00000). When a dirty cache line is evicted, its flag bits will be also moved to the WRQ together with the tag information.

When a write operation is scheduled, the flag bits determine if the write will be performed as normal two-step mode (F_1F_2 = '00'), waiting mode (F_1F_2 = '01'), or fast one-step mode (F_1F_2 = '11'). In our scheme, an on-going pre-recovery will not be preempted by a write request because write is generally not on the critical path of cache access. However, the read-preemptive rule similar to [76] and [7] is applied in our scheme so that pre-recovery operations can be preempted by the read access to the same cache bank. Unlike storing the whole stalled write request (including both address info and data) in an additional WRQ [7], here PRQ only stores the stalled pre-recovery request (i.e., address info).

4.4 EVALUATION OF SR-MLC STT-RAM

4.4.1 Experiment Setup

Our architectural simualtions are conducted on cycle-accurate simulator MacSim [77], the cache model of which has been modified based on the proposed PreREC technique. The baseline architecture is selected as an embedded processor with a two-level cache hierarchy similar to Intel Atom [78]. The configurations of the CPU core (except for L2 cache) are summarized in Table 7. A set of benchmarks selected from SPEC CPU2006 benchmark suite [70] is adopted in our simulations. Three STT-RAM based L2 cache designs are evaluated in our experiments, including:

- **SLC**: SLC STT-RAM cache (baseline).
- **C-MLC**: Conventional four-level MLC STT-RAM cache.
- **SR-MLC**: SR-MLC STT-RAM cache with PreREC technique.

Table 8 summarizes the design details of the evaluated L2 cache designs. All designs have the same area though the capacity of the SLC cache is only half of that of the C-MLC and SR-MLC caches [40]. SEC-DED (72,64,1) or BCH-20 (712,512,10) is needed by different cache designs to offer a low error rate of each 64B cache line at 1.8×10^{-7} (or a 98% cache yield). The cell-level error rates of C-MLC and SR-MLC are obtained from Section 4.2.2.2. Information density, which is 0.88, 1.44 and 1.33 for SLC, C-MLC, and SR-MLC designs,

Table 7: System configuration

Processor	1.8GHz, in-order, single-issue
SRAM L1 cache	32+32KB I/D, 64B line, 4-way write-back, 1 cycles R/W, 1R +1W ports
Main Memory	2GB, 100 cycles for the critical block

Table 8: Different configurations of STT-RAM L2 cache

	SLC	**C-MLC**	**SR-MLC**
Capacity (Byte)	4M	8M	8M
STT-RAM L2 cache	16-way, 64B line, 8 banks write-back, 1 R/W port per bank		
Block-level error rate	$\sim 1.8 \times 10^{-7}$		
Cell-level error rate	3×10^{-6} [67]	1.5×10^{-2}	5×10^{-6}
ECC	SEC-DED	BCH-20 [35]	SEC-DED
Information density	0.88bit/cell	1.44bit/cell	1.33bit/cell
Write Lat. (Cycles)	23	54	PreREC: 23 Normal: 46
Read Lat. (Cycles)	7	18	7

respectively, is defined as the effective number of information bits stored in an STT-RAM cell by considering ECC overhead. Thus, the effective capacity of the 4MB SLC, 8MB C-MLC and 8MB SR-MLC is 3.5MB, 5.7MB and 5.3MB, respectively. The logic area overhead of ECC is very marginal in all designs, e.g., less than 1% of the total cache area [35]. The latencies of SEC-DED and BCH-20 are characterized as 500ps and 6ns, respectively, at 45nm technology node [74]. NVSim [79] is used to obtain the performance parameters of different cache designs with 16 banks and 64B cache line, including the contributions from Codec and peripheral circuits. We also assume a parallel sensing scheme in C-MLC and SR-MLC designs to achieve the read latency similar to SLC design. A 20-entry PRQ is used in SR-MLC design [7].

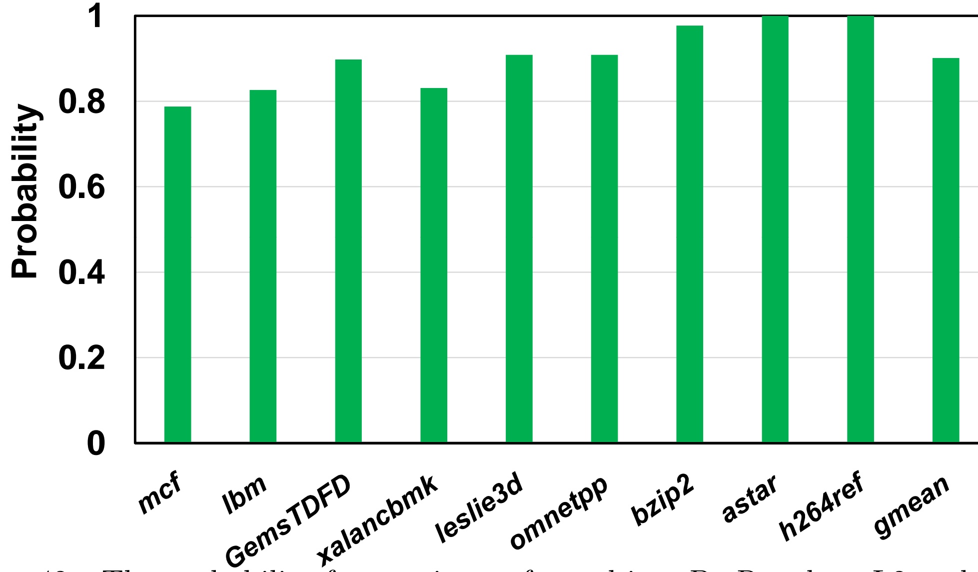

Figure 42: The probability for a write performed in a PreRec-done L2 cache line.

4.4.2 Evaluation of PreREC

Fig. 42 shows the probabilities that writes are performed on the L2 cache lines where pre-recovery is completed (PreREC-done) under different benchmarks. On average, 90% of the writes of L2 cache lines can be conducted with one-step. Also, the pre-recovery cancelation rate is very low even the read-preemptive rule is applied, as shown in Fig. 43: more than 90% PreREC operations can be finished without being interrupted by a read request. In fact, Fig. 43 also shows that the average time duration between two consecutive read accesses to the same bank of the L2 SR-MLC STT-RAM is $66 - 877$ clock cycles, which is much longer than a pre-recovery operation lasts (i.e., 23 cycles).

4.4.3 Performance Comparison

Fig. 44 shows the normalized *Instruction-Per-Cycle* (IPC) of the different L2 cache designs w.r.t. our 'SLC' baseline as well as the geometric mean over all benchmarks. Simply applying 'C-MLC' degrades the system performance by 12% on average mainly due to the long read access latency introduced by the costly ECC decoding. 'SR-MLC', however, improves the system performance by averagely 6.2%, which is mainly contributed by the increase in cache

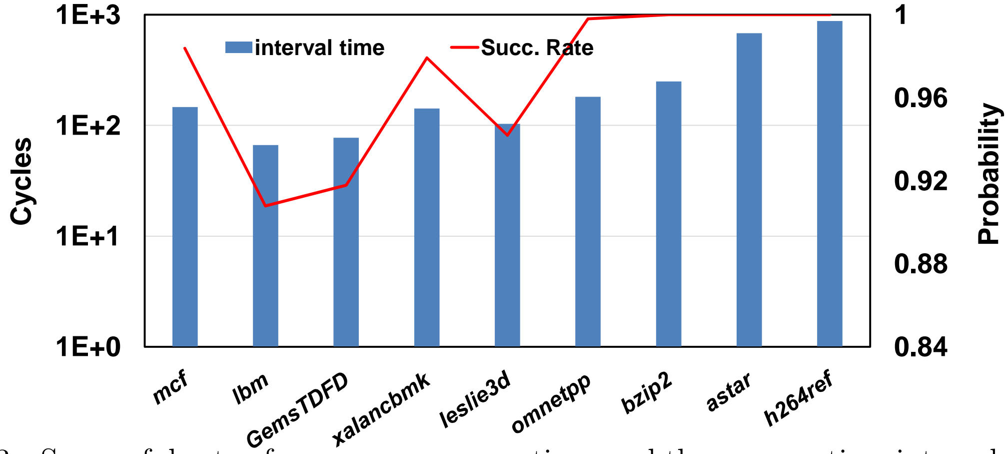

Figure 43: Successful rate of pre-recovery operations and the average time intervals between two consecutive reads.

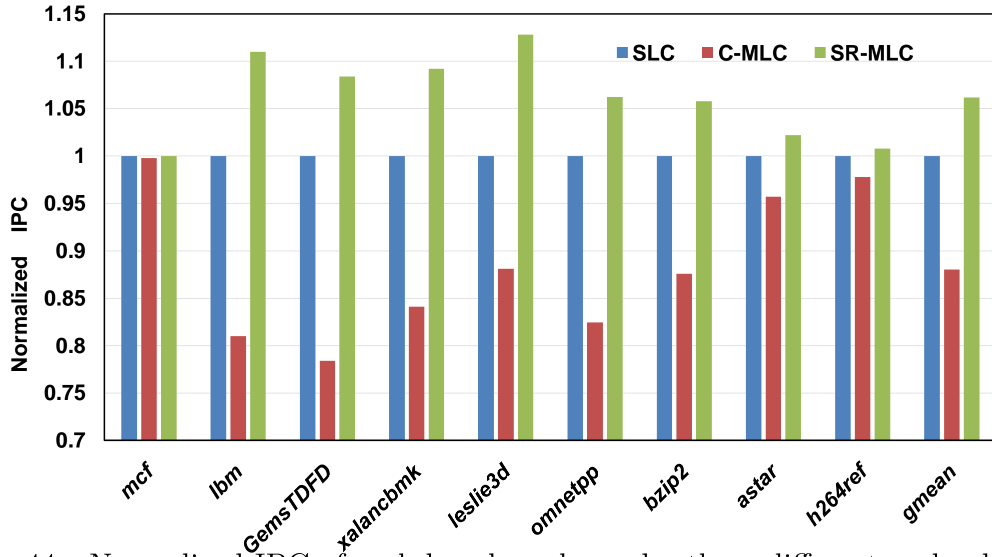

Figure 44: Normalized IPC of each benchmarks under three different cache designs.

capacity and the reduction in read and write latencies after the simple ECC and PreREC are applied, respectively. As expected, the benchmarks with heavy L2 cache accesses greatly benefit from SR-MLC design, such as *leslie3d* (12.8%), *lbm* (11.1%), *xalancbmk* (9.2%) and *GemsTDFD* (8.4%). The little improvement under benchmark *mcf* is because the increased cache capacity still cannot meet the demand of large working set of the application.

4.5 CHAPTER 4 SUMMARY

Similar to many other MLC memory technologies, MLC STT-RAM greatly suffers from the significantly degraded operation reliability as well as the high programming cost. In this chapter, three integrated techniques – StatRes, ErrPR, and TerCode, were proposed to construct a novel MLC STT-RAM cell design, namely, SR-MLC. The cell-level read and write reliability are enhanced by pruning the most error-prone resistance states and transitions. PreREC technique is also developed to minimize the expensive two-step write of the MLC STT-RAM during system executions. Experimental results show that under the same area budget, SR-MLC STT-RAM based L2 cache outperforms conventional SLC and MLC STT-RAM L2 cache while still offering high information density and significantly enhanced operation reliability.

5.0 CONCLUSION AND FUTURE WORK

5.1 DISSERTATION CONCLUSION

As billions of transistors and complex nano-systems are continuing to be integrated onto a single chip and the era of the big data application is approaching, the demand on memory and massive data storage capacity grows sharply due to the exponentially increased data processing capabilities. The large area of the chip memories are especially vulnerable to one-bit or multi-bit soft errors caused by single energetic particles like high-energy neutrons and alpha particles as the technology continues to shrink [80].

To cover those errors, *Error correction code* (ECC) has been proven a "must-have" technology in modern memory subsystem designs. The traditional memory technologies, such as SRAM and DRAM, can usually be equipped with the common ECCs, such as *single-error-correction double error detection code* (SEC-DED), BCH codes *etc.*, to better tradeoff between reliability, performance or energy due to the inherent reliability and fast programming in storage devices. Also, the extremely strong but slow *Low-Density-Parity-Check* (LDPC) codes are widely utilized in the high-density NAND flash memories because of the extremely degraded device reliability but slow programming speed requirement [81]. In recent years, the concerns on the continuous scaling of these technologies have motivated the tremendous investment in *emerging memory technologies* (EMTs), including *Spin-Transfer Torque Random Access Memory* (STT-RAM), *Resistive Random Access Memory* (RRAM) and *Phase Change Memory* (PCM). However, taking STT-RAM as an example, before benefiting from its attractive features-fast access speed, low leakage power and non-volatility, reliability issue becomes more and more prominent due to its unique storage stability from the aggravated process variations, stochastic device behaviors and environment fluctuations, and the ever-

increasing reliability requirement in massive data systems. As such, the complicated device or cell level reliability characterization becomes extremely expensive; Also, popular ECCs, such as SEC-DED, BCH or LPDC etc. may not be sufficient or suitable for STT-RAM based memory systems, and the demanding for stronger *error correction codes* (ECCs) or other solutions with minimized performance and hardware overhead for delay-sensitive on-chip/off chip applications are becoming essential.

This dissertation has looked at many facets of reliability issues of STT-RAM in designing memory systems, including the statistical *computer-aided design* (CAD) tool, the novel ECC design for asymmetric errors of SLC STT-RAM and the holistic circuit-architecture solution set for advanced MLC STT-RAM.

5.1.1 Conclusion of Chapter 2

Process variations and thermal fluctuations significantly affect the write reliability and write energy of STT-RAM, traditionally, modeling the impacts of these variations on STT-RAM designs requires expensive Monte-Carlo runs with hybrid magnetic-CMOS simulation steps. Also, those solutions are usually performed on the STT-RAM cells with fixed variation configurations, and significantly reduce their scalability and portability. Thus, in Chapter 2, we proposed PS3-RAM–a fast, portable and scalable statistical STT-RAM reliability/energy analysis method. By introducing the sensitivity analysis technique to capture the statistical characteristics of the MTJ switching, and dual-exponential model to efficiently and accurately recover the MTJ switching current samples for statistical STT-RAM thermal analysis, PS3- RAM can achieve multiple orders-of-magnitude run time cost reduction with marginal accuracy degradation under any variation configurations when compared to SPICE-based Monte-Carlo simulations.

5.1.2 Conclusion of Chapter 3

In chapter 3, we proposed the first analytical *asymmetric write channel* (AWC) to deeply understand the unique operation errors of STT-RAM write mechanism–its write failure rate is extremely asymmetric (the writing ‘1’ error rate can be even several orders higher than

that of writing ‘0’). By carefully investigating the common ECC solutions to tolerate such errors, we discovered interesting observations neglected before in memory systems: Generic ECCs like SEC-DED code, etc. are all designed under the assumption that the symmetric error rate always exists at $0 \rightarrow 1$ and $1 \rightarrow 0$ flipping and such ECCs cannot efficiently handle the highly asymmetric writing errors at different bit-flipping directions. Thus, to efficiently address such challenges, we introduced the new design concept based ECCs-the content dependent view instead of the worst-corner design view. The original data is intentionally partitioned into two different corners based on their reliability degree, and can be further processed through the proposed low cost circuit-level solutions–*typical-corner-ECC* (TCE) scheme or the *worst-corner-ECC* (WCE) scheme, respectively. By proposing the *content-dependent ECC* (CD-ECC) technique to balance and enhance the reliability of the STT-RAM with asymmetric write errors, our CD-ECC improves the reliability of the STT-RAM based cache system significantly with marginal performance degradation.

5.1.3 Conclusion of Chapter 4

The invention of *multi-level cell* (MLC) technology doubles the storage density by integrating two MTJs with different dimensions in one memory cell to represent multiple logic bits. However, MLC STT-RAM design further aggravates the reliability and write latency *w.r.t.* the single-level cell (SLC) version. In chapter 4, we demonstrated the infeasibility of applying extremely strong ECCs on MLC STT-RAM based memory systems for high-reliable and high-performance applications due to the associated decoding latency and storage overhead. Thus, we proposed a cross-layer solution, named *State-Restrict MLC* STT-RAM (SR-MLC), to address the reliability, performance and information density simultaneously. Three techniques: state restriction, error pattern removal, and ternary coding are proposed at circuit level to reduce the read and write errors of MLC STT-RAM cells. State pre-recovery technique is further developed at architecture level to improve the access performance of SR-MLC STT-RAM by eliminating unnecessary two-step write operations. Simulation results show that our SR-MLC design can enhance the write/read error rate by $10 - 10000\times$ over traditional MLC designs, while simultaneously boosting the system performance by averagely

6.2% over SLC designs. In summary, our solution delivers similar information density as traditional MLC design, comparable reliability and programming speed as SLC design, but significantly improved IPC performance.

5.2 FUTURE WORK

For future research work, the research of high-reliable, high-performance and energy-efficient MLC STT-RAM based memory systems still requires serious investigation since we believe that the multi-level design for STT-RAM, like the MLC NAND Flash memory, may eventually become available for commercialization.

5.2.1 Facts and Observations

In MLC STT-RAM designs, soft-bit and hard-bit show different write radiabilities. As mentioned in chapter 4, there are two kinds of write failures in MLC STT-RAM: incomplete write and overwrite. Note that incomplete write may occur at soft-bit and hard-bit, resulting in an error at the bit being programmed. Considering that the hard-bit requires a larger critical current I_C than the soft-bit, the incomplete write failure at hard-bits is severer. On the contrary, overwrite leads to unexpected error of only hard-bit when writing to the corresponding soft-bit. Thus, the average error rate of hard-bits is much larger than that of soft-bits, leading to the reliability asymmetry in MLC bits. For instance, if we assume the area ratio of the two MTJs is 2 [15] and they are fabricated at 32nm technology node [11], the bit error rates of soft-bit and hard-bits are roughly $P_{f_s} = 1.5 \times 10^{-8}$ and $P_{f_h} = 3.5 \times 10^{-3}$, respectively. That is, the asymmetry ratio of the error rates of hard-bits and soft-bits can be as high as five orders of magnitude.

Different logic-to-physical data mapping schemes have been well studied in the applications of MLC STT-RAM. Figure 45(a) shows conventional 2-bit MLC deign which stores a data block of N bits in $N/2$ MLC cells. Half of the data (i.e., $A_0 A_2 \cdots A_{N-2}$ of Data Block 0 in Figure 45 are stored in soft-bits and the other half $(A_1 A_3 \cdots A_{N-1})$ are saved in hard-

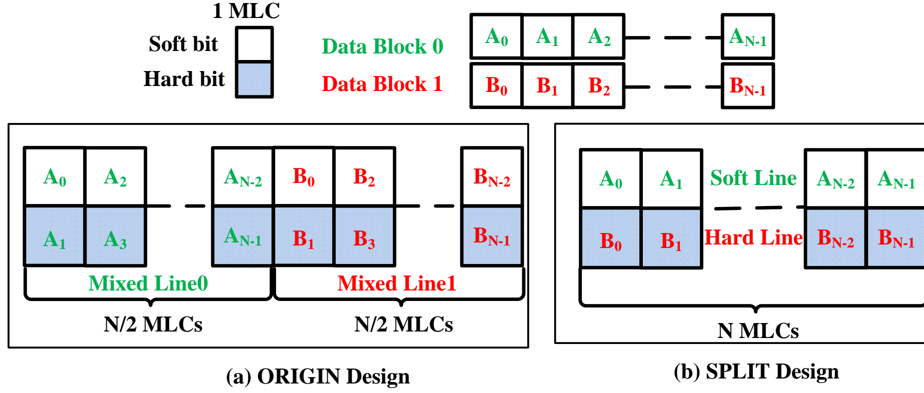

Figure 45: Illustration of ORIGINAL design vs. SPLIT design structure.

bits. We name such storage structure as the ORIGINAL design. When accessing a data block containing both soft- and hard-bits (mixed-line), the costly two-step read or write in Figure 35(c,d) are always needed. To solve the slow access issue of the ORIGINAL design, many prior works were performed [40, 75]. Figure 45(b) illustrates a SPLIT design that maps two data blocks ($A_0A_1 \cdots A_{N-1}$ and $B_0B_1 \cdots B_{N-1}$) into the entire soft-bits (soft-line) and hard-bits (hard-line) of N MLC cells, respectively. As such, only one-step read and write are needed by accessing the soft-line. If more and critical data blocks can be placed to the soft-lines, the overall system performance can be substantially improved. Note that two-step read and write are still needed when accessing the hard-lines.

The SPLIT design well leverages the access asymmetry, yet neglects the reliability asymmetry of the soft- and hard-bits. More specifically, the reliability of hard-lines degrades severely from the level that the mixed-line in ORIGINAL design can offer. We analyzed the write error rates for the mixed-lines of the ORIGINAL design as well as the soft-lines and hard-lines of the SPLIT design. We assume the length of data block is 64 bits and adopt the above P_{f_s} and P_{f_h}. Table 9 summarizes the calculated probabilities of 0-bit (failure free), 1-bit and 2-bit errors for the three types of lines. The preliminary results show that the hard-line is the reliability bottleneck of the SPLIT design. Compared to the mixed-line in the ORIGINAL design, the hard-line in the SPLIT design increases the probabilities of

Table 9: Reliability comparison of mixed-line, hard-line and soft-line

Design	ORIGINAL Design	SPLIT Design	
Line mode	Mixed Line	Hard Line	Soft Line
0 bit error rate	0.8939	0.7990	$9.9e-5$
1 bit error rate	0.1005	0.1796	$9.6e-7$
2 bit error rate	0.0055	0.0199	$4.54e-13$
sum 0,1bit error rate	0.9944	0.9786	1.0000
sum 0,1,2 bit error rate	0.9999	0.9985	1.0000

1-bit and 2-bit errors by 79% and 262%, respectively. If a target yield of 99.4% is required, the required ECC strength for a mixed-line and a hard-line will become 1 bit and 2 bits, respectively. Single-bit ECCs (e.g., SEC-DED) that are sufficient for a mixed-line become no longer enough for a hard-line. *Bose-Chaudhuri-Hocquenghem* (BCH) code with double-error correction capability could be a choice. However, their classical decoding schemes based on Berkeamo Massey algorithm and Chien search demand a very long decoding latency [82, 83]. After including these design factors, the performance advantage offered by the SPLIT design might be gone.

Considering the discovered reliability issue of the SPLIT design, for the future work, the cross-layer approach to address the performance and reliability challenges of MLC STT-RAM simultaneously shall be explored. For example, a fast multi-bit ECC design, namely, *hierarchical ECC* (HECC), can be developed. The solutions in memory hierarchy, i.e., the *non-uniform strength ECC* (NUS-ECC), should then be proposed by leveraging HECC for performance and/or energy efficiency enhancement.

5.2.2 Multi-bit ECC Design

The decoding latency of standard BCH is determined by the number of errors in the data [84, 85]. The decoding latency of BCH for an error-free data is relatively fast. However, when

there are 1 or 2 errors, the decoding latency will significantly increase. Note that the decoding latencies of standard BCH for 1 or 2 errors are the same as it always goes through the same procedure, including syndrome computation, error locator polynomial generation and error locator polynomial. As shown in Table 9, more than 19% of hard-line accesses potentially have at least 1 erroneous bit and therefore, may suffer from a long decoding latency of BCH. This fact may cause considerable performance loss in applications. Since the occurrence probability of 2-bit error in a hard-line is much lower than that of the other situations (e.g., 1.99% vs. 97.8% in Table 9), a *hierarchical ECC* (HECC) design might be proposed to particularly speedup the access latency of 0- or 1-bit error rather than the 2-bit error. Two ECC modes can be introduced: mode 1 is a high-speed mode associated with 1-bit error correction capability. Its access latency and cost are as low as SEC-DED; mode 2 is dedicated to correct 2-bit error by paying a longer latency (but still shorter than the BCH) [86]. Apparently, most accesses fit into mode 1. In the rare case that a 2-bit error happens, mode 2 will be adopted immediately. The generator matrix and the parity check matrix of HECC can be studied and developed.

5.2.3 Non-uniform ECC Design

In cache design, if the capacities of different types of cache way (or mixed-line/soft-line/hard-line) are fixed, the overall memory reliability is mainly determined by the ECC schemes applied on each cache way. Thus, a *non-uniform strength ECC* (NUS-ECC) scheme for different cache lines can also be studied to achieve the balanced reliability across different cache ways with the minimum storage cost and decoding latency: 1) the soft ways have the lowest error rate. So a simple ECC scheme with 1-bit error correction capability might be sufficient to maintain a reasonably low error rate. The incurred storage cost is also low, i.e., 11 extra bits for a 512-bit cache line; 2) the increase of error rate in the mixed ways may be corrected by some medium ECC schemes, e.g., partitioning the whole cache line into smaller segments, each of which is applied with a simple ECC scheme; 3) The HECC developed above can be leveraged to protect the most error-prone hard-ways.

5.2.4 Architecture Investigation

Carefully balancing the ratio between the capacities of different cache ways is an essential to achieve a combined enhancement on performance, energy, and hardware cost compared to the sole SPLIT or ORIGINAL design. Besides the "static" design metrics of different cache line designs such as error rate, ECC error correction capability etc., the "dynamic" characteristics of the data stored into the cache lines, such as the data flipping and access patterns *etc.*, also greatly influence the design optimization. Hence, architectural innovations become a "must-have" in the potential future research. Some techniques, *e.g.*, data migration and placement, can be introduced to minimize the performance cost incurred in the tri-way cache, which may include mixed-way, soft-way and hard-way.

5.3 RESEARCH SUMMARY AND INSIGHT

STT-RAM has demonstrated great potentials in next-generation storage and computing systems. However, reliability continues to be one of the most critical design challenges before adopting STT-RAM in future memory/storage subsystems. In this dissertation, we demonstrate our understanding on the error characterization and correction techniques for reliable STT-RAM designs. We believe that fast, portable and scalable statistical tools to calibrate STT-RAM reliability issues are essential for architects to conduct reliability-driven optimizations. Also, the efficient robust (or ECC) designs for STT-RAM require a deep holistic understanding on the three different levels-unique device features, circuit implementation and architecture applications, as chapter 3 and chapter 4 show. Innovative ECC schemes and their architectural applications shall be seriously investigated to accelerate the deployment of the modern microprocessors using the emerging nonvolatile memories in the near future.

BIBLIOGRAPHY

[1] Samsung, “Samsung Global MRAM Innovation Program,” 2013, http://www.samsung.com/global/busi-ness/semiconductor/news-events/mram.

[2] M. Hosomi, H. Yamagishi, T. Yamamoto, K. Bessho, Y. Higo, K. Yamane, H. Yamada, M. Shoji, H. Hachino, C. Fukumoto, H. Nagao, and H. Kano, “A Novel Nonvolatile Memory with Spin Torque Transfer Magnetization Switching: Spin-RAM,” in *IEEE International Conference on Electron Devices Meeting Technical Digest*, Dec. 2005, pp. 459–462.

[3] F. Bedeschi *et al.*, “A Bipolar-Selected Phase Change Memory Featuring Multi-Level Cell Storage,” *JSSC*, vol. 44, no. 1, pp. 217 – 227, Jan 2009.

[4] Everspin, “Second Generation MRAM: Spin Torque Technology,” 2012, http://www.everspin.com/products/second-generation-st-mram.html.

[5] H. Lv *et al.*, “Resistive memory switching of cuxo films for a nonvolatile memory application,” *Electron Device Letters, IEEE*, vol. 29, no. 4, pp. 309–311, April 2008.

[6] N. Mojumder, C. Augustine, D. Nikonov, and K. Roy, “Effect of Quantum Confinement on Spin Transport and Magnetization Dynamics in Dual Barrier Spin Transfer Torque Magnetic Tunnel Junctions,” *Journal of Applied Physics*, vol. 108, no. 10, pp. 104 306–104 306–12, Nov. 2010.

[7] G. Sun, X. Dong, Y. Xie, J. Li, and Y. Chen, “A Novel Architecture of the 3D Stacked MRAM L2 Cache for CMPs,” in *the 15th International Symposium on High-Performance Computer Architecture*. IEEE, 2009, pp. 239–249.

[8] W. Xu, H. Sun, Y. Chen, and T. Zhang, “Design of Last-Level On-Chip Cache Using Spin-Torque Transfer RAM (STT RAM),” in *IEEE Trans. on VLSI System*. IEEE, 2011, pp. 483–493.

[9] P. Zhou, B. Zhao, J. Yang, and Y. Zhang, “Energy Reduction for STT-RAM Using Early Write Termination,” in *The 2009 International Conference on Computer-Aided Design*. ACM, 2009, pp. 264–268.

[10] C. Smullen, V. Mohan, A. Nigam, S. Gurumurthi, and M. Stan, "Relaxing non-volatility for fast and energy-efficient stt-ram caches," in *High Performance Computer Architecture (HPCA), 2011 IEEE 17th International Symposium on*, Feb 2011, pp. 50–61.

[11] Y. Chen, W.-F. Wong, H. Li, C.-K. Koh, Y. Zhang, and W. Wen, "On-chip caches built on multilevel spin-transfer torque RAM cells and its optimizations," *J. Emerg. Technol. Comput. Syst.*, vol. 9, no. 2, pp. 16:1–16:22, May 2013.

[12] T. Kawahara *et al.*, "2mb spin-transfer torque ram (spram) with bit-by-bit bidirectional current write and parallelizing-direction current read," in *Solid-State Circuits Conference, 2007. ISSCC 2007. Digest of Technical Papers. IEEE International*, Feb 2007, pp. 480–617.

[13] K. Tsuchida *et al.*, "A 64mb mram with clamped-reference and adequate-reference schemes," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2010 IEEE International*, Feb 2010, pp. 258–259.

[14] D. Halupka *et al.*, "Negative-resistance read and write schemes for stt-mram in 0.13 um cmos," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2010 IEEE International*, Feb 2010, pp. 256–257.

[15] R. Nebashi *et al.*, "A 90nm 12ns 32mb 2t1mtj mram," in *Solid-State Circuits Conference - Digest of Technical Papers, 2009. ISSCC 2009. IEEE International*, Feb 2009, pp. 462–463,463a.

[16] C. Lin, S. Kang, Y. Wang, K. Lee, X. Zhu, W. Chen, X. Li, W. Hsu, Y. Kao, M. Liu, W. Chen, Y. Lin, M. Nowak, N. Yu, and L. Tran, "45nm low power cmos logic compatible embedded stt mram utilizing a reverse-connection 1t/1mtj cell," in *Electron Devices Meeting (IEDM), 2009 IEEE International*, Dec 2009, pp. 1–4.

[17] Y. Chen *et al.*, "Combined magnetic- and circuit-level enhancements for the nondestructive self-reference scheme of stt-ram," in *Low-Power Electronics and Design (ISLPED), 2010 ACM/IEEE International Symposium on*, Aug 2010, pp. 1–6.

[18] S. Dent, "Everspin throws first ST-MRAM chips down, launches commercial spintorque memory era," 2012, http://www.engadget.com/2012/11/14/everspin-throws-first-st-mram-chips-down/.

[19] B. Browdie, "Crocus Unveils Chip for Storing Transaction Data on Smartphones and Smart Cards," 2012, http://www.americanbanker.com/issues/177_214/crocus-chip-storing-transaction-data-on-smartphones-1054128-1.html.

[20] Y. Zhang *et al.*, "Multi-level Cell STT-RAM: Is It Realistic or Just a Dream?" in *ICCAD*, Nov 2012, pp. 526–532.

[21] Y. Zhang, W. Wen, and Y. Chen, "The prospect of stt-ram scaling from readability perspective," *Magnetics, IEEE Transactions on*, vol. 48, no. 11, pp. 3035–3038, Nov 2012.

[22] Z. Sun, H. Li, Y. Chen, and X. Wang, "Variation Tolerant Sensing Scheme of Spin-Transfer Torque Memory for Yield Improvement," in *IEEE/ACM International Conference on Computer-Aided Design*, Nov. 2010, pp. 432 –437.

[23] R. Sbiaa, R. Law, S. Y. H. Lua, E. L. Tan, T. Tahmasebi, C. C. Wang, and S. N. Piramanayagam, "Spin Transfer Torque Switching for Multi-bit Per Cell Magnetic Memory with Perpendicular Anisotropy," *Applied Physics Letters*, vol. 99, no. 9, pp. 092 506 –092 506–3, Aug. 2011.

[24] Y. Chen, X. Wang, W. Zhu, H. Li, Z. Sun, G. Sun, and Y. Xie, "Access Scheme of Multi-Level Cell Spin-Transfer Torque Random Access Memory and Its Optimization," in *53rd IEEE International Midwest Symposium on Circuits and Systems*, Aug. 2010, pp. 1109 –1112.

[25] S. Chatterjee, M. Rasquinha, S. Yalamanchili, and S. Mukhopadhyay, "A Scalable Design Methodology for Energy Minimization of STTRAM: A Circuit and Architecture Perspective," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 19, no. 5, pp. 809–817, 2011.

[26] Y. Huai, "Spin-Transfer Torque MRAM (STT-MRAM): Challenges and Prospects," *AAPPS Bulletin*, vol. 18, no. 6, pp. 33–40, 2008.

[27] J. Li, H. Liu, S. Salahuddin, and K. Roy, "Variation-Tolerant Spin-Torque Transfer (STT) MRAM Array for Yield Enhancement," in *CICC*, Sep. 2008, pp. 193 –196.

[28] C. W. Smullen, A. Nigam, S. Gurumurthi, and M. R. Stan, "The STeTSiMS STT-RAM Simulation and Modeling System," in *ICCAD*, Nov 2011, pp. 318–325.

[29] Y. Chen, X. Wang, H. Li, H. Xi, Y. Yan, and W. Zhu, "Design margin exploration of spin-transfer torque ram (stt-ram) in scaled technologies," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 18, no. 12, pp. 1724–1734, Dec 2010.

[30] S. Schechter, G. H. Loh, K. Straus, and D. Burger, "Use ecp, not ecc, for hard failures in resistive memories," *SIGARCH Comput. Archit. News*, vol. 38, no. 3, pp. 141–152, June 2010.

[31] D. H. Yoon and M. Erez, "Memory mapped ECC: low-cost error protection for last level caches," in *ISCA*, 2009, pp. 116–127.

[32] N. H. Seong, D. H. Woo, V. Srinivasan, J. Rivers, and H.-H. Lee, "Safer: Stuck-at-fault error recovery for memories," in *Microarchitecture (MICRO), 2010 43rd Annual IEEE/ACM International Symposium on*, Dec 2010, pp. 115–124.

[33] D. H. Yoon, N. Muralimanohar, J. Chang, P. Ranganathan, N. Jouppi, and M. Erez, "Free-p: Protecting non-volatile memory against both hard and soft errors," in *High Performance Computer Architecture (HPCA), 2011 IEEE 17th International Symposium on*, Feb 2011, pp. 466–477.

[34] J. Kim, N. Hardavellas, K. Mai, B. Falsafi, and J. Hoe, "Multi-bit error tolerant caches using two-dimensional error coding," in *Proceedings of the 40th Annual IEEE/ACM International Symposium on Microarchitecture*, 2007, pp. 197–209.

[35] A. R. Alameldeen *et al.*, "Energy-efficient Cache Design Using Variable-strength Error-correcting Codes," in *ISCA*, 2011, pp. 461–472.

[36] W. Wen, Y. Zhang, Y. Chen, Y. Wang, and Y. Xie, "PS3-RAM: A fast portable and scalable statistical STT-RAM reliability analysis method," in *49th DAC*, June 2012, pp. 1187–1192.

[37] T. Ishigaki *et al.*, "A multi-level-cell spin-transfer torque memory with series-stacked magnetotunnel junctions," in *VLSIT*, 2010, pp. 47–48.

[38] T. K. Moon, *Error Correction Coding: Mathematical Methods and Algorithms.* John Wiley and Sons, 2005.

[39] W. Xu, Y. Chen, X. Wang, and T. Zhang, "Improving STT MRAM storage density through smaller-than-worst-case transistor sizing," in *46th DAC*, July 2009, pp. 87 –90.

[40] X. Bi, M. Mao, D. Wang, and H. Li, "Unleashing the Potential of MLC STT-RAM Caches," in *ICCAD*, Nov 2013, pp. 429–436.

[41] X. Wang, Y. Zheng, H. Xi, and D. Dimitrov, "Thermal Fluctuation Effects on Spin Torque Induced Switching: Mean and Variations," *Journal of Applied Physics*, vol. 103, no. 3, pp. 034 507–034 507–4, Feb. 2008.

[42] Z. Diao, Z. Li, S. Wang, Y. Ding, A. Panchula, E. Chen, L. Wang, and Y. Huai, "Spin-transfer Torque Switching in Magnetic Tunnel Junctions and Spin-transfer Torque Random Access Memory," *Journal of Physics: Condensed Matter*, vol. 19, p. 165209, 2007.

[43] L. Berger, "Emission of Spin Waves by a Magnetic Multilayer Traversed by a Current," *Phys. Rev. B*, vol. 54, pp. 9353 –9358, Oct 1996.

[44] T. Gilbert, "A Lagrangian Formulation of the Gyromagnetic Equation of the Magnetization Field," *Phys.Tev.*, vol. 100, no. 1243, 1955.

[45] R. Singha, A. Balijepalli, A. Subramaniam, F. Liu, and S. Nassif, "Modeling and Analysis of Non-Rectangular Gate for Post-Lithography Circuit Simulation," in *44th DAC*, June 2007, pp. 823 –828.

[46] Y. Ye, F. Liu, S. Nassif, and Y. Cao, "Statistical Modeling and Simulation of Threshold Variation under Dopant Fluctuations and Line-Edge Roughness," in *45th DAC*, June 2008, pp. 900 –905.

[47] F. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51 – 83, Jan. 1978.

[48] Y. Zhang, X. Wang, and Y. Chen, "STT-RAM Cell Design Optimization for Persistent and Non-Persistent Error rate Reduction: A statistcal Design View," in *ICCAD*, Nov. 2011, pp. 471–477.

[49] Y. Zhang, W. Wen, and Y. Chen, "STT-RAM Cell Design Considering MTJ Asymmetric Switching," *SPIN*, vol. 02, no. 03, p. 1240007, 2012.

[50] ASU, "Predictive Technology Model (PTM)," http://www.eas.asu.edu/~ptm/.

[51] BSIM, "http://www-device.eecs.berkeley.edu/bsim3/," *UC Berkeley*.

[52] B. Sheu, D. Scharfetter, P.-K. Ko, and M.-C. Jeng, "BSIM: Berkeley short-channel IGFET model for MOS transistors," *JSSC*, vol. 22, no. 4, pp. 558 – 566, Aug 1987.

[53] P. Doubilet, C. Begg, M. Weinstein, P. Braun, and B. McNeil, "Probabilistic Sensitivity Analysis Using Monte Carlo Simulation. A Practical Approach," 1985.

[54] X. Cong, N. Dimin, Z. Xiaochun, K. H. Seung, N. Matt, and Y. Xie, "Device Architecture Co-Optimization of STT-RAM Based Memory for Low Power Embedded Systems," in *ICCAD*, Nov 2011, pp. 463–470.

[55] Y. Zhang, Y. Li, A.K.Jones, X. Wang, and Y. Chen, "Asymmetry of MTJ Switching and Its Implication to the STT-RAM Designs," *Design Automation and Test in Europe*, Mar. 2012.

[56] M. Hsiao, "A Class of Optimal Minimum Odd-weight-column SEC-DED Codes," *IBM Journal of Research and Development*, vol. 14, no. 4, pp. 395–401, 1970.

[57] C. Chen and M. Hsiao, "Error-Correcting Codes for Semiconductor Memory Applications: A State-of-the-Art Review," *IBM Journal of Research and Development*, vol. 28, no. 2, pp. 124–134, 1984.

[58] B. Bose and S. Al-Bassam, "On systematic single asymmetric error-correcting codes," *Information Theory, IEEE Transactions on*, vol. 46, no. 2, pp. 669–672, Mar 2000.

[59] T. Etzion, "New lower bounds for asymmetric and unidirectional codes," *Information Theory, IEEE Transactions on*, vol. 37, no. 6, pp. 1696–1704, Nov 1991.

[60] F.-W. Fu, S. Ling, and C. Xing, "New lower bounds and constructions for binary codes correcting asymmetric errors," *Information Theory, IEEE Transactions on*, vol. 49, no. 12, pp. 3294–3299, Dec 2003.

[61] T. Klove, “Upper bounds on codes correcting asymmetric errors (corresp.),” *IEEE Trans. Inf. Theor.*, vol. 27, no. 1, pp. 128–131, Sept. 2006.

[62] S. D. Constantin and T. R. N. Rao, “On the theory of binary asymmetric error-correcting codes,” *IEEE Information and Control*, vol. 40, pp. 20–36, 1979.

[63] P. Delsarte and P. Piret, “Bounds and constructions for binary asymmetric error-correcting codes (corresp.),” *Information Theory, IEEE Transactions on*, vol. 27, no. 1, pp. 125–128, Jan 1981.

[64] J. Weber, C. de Vroedt, and D. Boekee, “Bounds and constructions for binary codes of length less than 24 and asymmetric distance less than 6,” *Information Theory, IEEE Transactions on*, vol. 34, no. 5, pp. 1321–1331, Sep 1988.

[65] H. Zhou, A. Jiang, and J. Bruck, “Nonuniform codes for correcting asymmetric errors,” in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, July 2011, pp. 1046–1050.

[66] C. Wilkerson, A. R. Alameldeen, Z. Chishti, W. Wu, D. Somasekhar, and S.-L. Lu, “Reducing cache power with low-cost, multi-bit error-correcting codes,” in *ISCA*, 2010, pp. 83–93.

[67] X. Bi, Z. Sun, H. Li, and W. Wu, “Probabilistic Design Methodology to Improve Run-time Stability and Performance of STT-RAM Caches,” in *ICCAD*, Nov 2012, pp. 88–94.

[68] Y. Pan, G. Dong, and T. Zhang, “Exploiting memory device wear-out dynamics to improve NAND flash memory system performance,” in *Proceedings of the 9th USENIX conference on File and stroage technologies (FAST’11)*, 2011, pp. 18–18.

[69] A. A. M. Saleh and R. Valenzuela, “A Statistical Model for Indoor Multipath Propagation,” *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 128–137, 1987.

[70] S. CPU2006, http://www.spec.org/cpu2006/.

[71] H. Kim *et al.*, “MacSim Simulator,” http://code.google.com/p/macsim/.

[72] X. Dong, C. Xu, Y. Xie, and N. Jouppi, “NVSim: A Circuit-Level Performance, Energy, and Area Model for Emerging Nonvolatile Memory,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 31, no. 7, pp. 994–1007, 2012.

[73] M. Hosomi *et al.*, “A novel nonvolatile memory with Spin torque transfer magnetization switching: Spin-RAM,” in *IEDM*, 2005, pp. 459–462.

[74] N. H. Seong *et al.*, “Tri-level-cell Phase Change Memory: Toward an Efficient and Reliable Memory System,” in *ISCA*, 2013, pp. 440–451.

[75] L. Jiang *et al.*, "Constructing Large and Fast Multi-level Cell STT-MRAM Based Cache for Embedded Processors," in *49th DAC*, 2012, pp. 907–912.

[76] M. Qureshi *et al.*, "PreSET: Improving performance of phase change memories by exploiting asymmetry in write times," in *ISCA*, 2012, pp. 380–391.

[77] H. Kim *et al.*, "MacSim Simulator," http://code.google.com/p/macsim/.

[78] Atom, http://ark.intel.com/products/family/29035.

[79] NVSim, http://www.rioshering.com/nvsimwiki/index.php.

[80] E. Ibe, H. Taniguchi, Y. Yahagi, K. i. Shimbo, and T. Toba, "Impact of scaling on neutron-induced soft error in srams from a 250 nm to a 22 nm design rule," *Electron Devices, IEEE Transactions on*, vol. 57, no. 7, pp. 1527–1538, Jul 2010.

[81] K. Zhao, W. Zhao, H. Sun, T. Zhang, X. Zhang, and N. Zheng, "Ldpc-in-ssd: Making advanced error correction codes work effectively in solid state drives," in *Proceedings of the 11th USENIX Conference on File and Storage Technologies*, ser. FAST'13, 2013, pp. 243–256.

[82] P. Reviriego, C. Argyrides, and J. A. Maestro, "Efficient error detection in double error correction bch codes for memory applications," *Microelectronics Reliability*, vol. 52, no. 7, pp. 1528–1530, 2012.

[83] H. Burton, "Inversionless decoding of binary bch codes," *Information Theory, IEEE Transactions on*, vol. 17, no. 4, pp. 464–466, Jul 1971.

[84] S. Rizwan, "Retimed decomposed serial berlekamp-massey (bm) architecture for high-speed reed-solomon decoding," in *VLSI Design, 21st International Conference on*, Jan 2008, pp. 53–58.

[85] J. Cho and W. Sung, "Strength-reduced parallel chien search architecture for strong bch codes," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 55, no. 5, pp. 427–431, May 2008.

[86] Z. Wang, "Hierarchical decoding of double error correcting codes for high speed reliable memories," in *Design Automation Conference (DAC), 2013 50th ACM / EDAC / IEEE*, May 2013, pp. 1–7.