# INTEGRATED CIRCUIT OUTLIER IDENTIFICATION

# BY MULTIPLE PARAMETER CORRELATION

A Dissertation

by

SAGAR SURESH SABADE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2004

Major Subject: Computer Engineering

**INTEGRATED CIRCUIT OUTLIER IDENTIFICATION**

**BY MULTIPLE PARAMETER CORRELATION**

A Dissertation

by

SAGAR SURESH SABADE

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

| | |
|---|---|
| Duncan Walker<br>(Chair of Committee) | Donald Friesen<br>(Member) |
| Vivek Sarin<br>(Member) | Henry Taylor<br>(Member) |
| | Valerie Taylor<br>(Head of Department) |

May 2004

Major Subject: Computer Engineering

## ABSTRACT

Integrated Circuit Outlier Identification by Multiple Parameter Correlation.

(May 2004)

Sagar Suresh Sabade, B.E., Pune University, India;

M.Tech., Indian Institute of Science, Bangalore, India

Chair of Advisory Committee: Dr. Duncan Walker

Semiconductor manufacturers must ensure that chips conform to their specifications before they are shipped to customers. This is achieved by testing various parameters of a chip to determine whether it is defective or not. Separating defective chips from fault-free ones is relatively straightforward for functional or other Boolean tests that produce a go/no-go type of result. However, making this distinction is extremely challenging for parametric tests. Owing to continuous distributions of parameters, any pass/fail threshold results in yield loss and/or test escapes. The continuous advances in process technology, increased process variations and inaccurate fault models all make this even worse. The pass/fail thresholds for such tests are usually set using prior experience or by a combination of visual inspection and engineering judgment.

Many chips have parameters that exceed certain thresholds but pass Boolean tests. Owing to the imperfect nature of tests, to determine whether these chips (called “outliers”) are indeed defective is nontrivial. To avoid wasted investment in packaging or further testing it is important to screen defective chips early in a test flow. Moreover, if seemingly strange behavior of outlier chips can be explained with the help of certain process parameters or by correlating additional test data, such chips can be retained in the test flow before they are proved to be fatally flawed.

In this research, we investigate several methods to identify *true* outliers (defective chips, or chips that lead to functional failure) from *apparent* outliers (seemingly defective, but fault-free chips). The outlier identification methods in this research primarily rely on wafer-level spatial correlation, but also use additional test parameters. These methods are evaluated and validated using industrial test data. The potential of these methods to reduce burn-in is discussed.

# DEDICATION

To my grandmother and my parents, whom I owe everything.

Without their unflagging support this wouldn't have been possible.

---

The following poem that I composed during one of my study breaks, expresses the general theme of my research.

VLSI Testing – A Poetic Perspective

There's an open or there is a short
These are defects that cause a fault
Current may rise or current may halt
Depends on where and what caused the fault

Technology today is called CMOS
Takes a tiny current, has small power loss
Current would increase for a defect gross
But if it's subtle, you are at a loss

Is chip good or faulty: who can say?
One has to find a better way
Throwing good chips means loss of yield
But shipping a bad one, can fail on field

Fail on field means customer return
That can hamper your reputation
"Poor quality" customers will say
Reducing defect level is the only way

## ACKNOWLEDGMENTS

I arrived on the campus of Texas A&M University in August 1999 with rather a vague notion of doing a Ph.D. As I reflect upon this academic journey so far, I realize I owe a lot to many people who supported me throughout this process. I am deeply indebted to Dr. Hank Walker for his financial, academic and moral support throughout my years at TAMU. He encouraged me, like his other students, to explore a research direction on my own in order to hone my research skills and independent thinking. He shared his insights to help me understand where my research fits in the long-term. I would like to thank my committee members Dr. Donald Friesen, Dr. Vivek Sarin and Dr. Henry Taylor for agreeing to serve on my committee and for their helpful comments and constructive criticism. I would like to express my thanks to my Graduate Committee Representative (GCR) Dr. Sharon Braunagel for completing the paperwork promptly. Thanks to Wanda Allen, Elena Catalena, Dr. Bart Childs, Margaret Dunaway, Glenda Elledge, Leah Lewis, and Sandy Morse for cooperating in the administrative tasks.

I gratefully acknowledge the financial support I received from the National Science Foundation (NSF), Texas Advanced Research Program/Advanced Technology Program (ARP/ATP) and Semiconductor Research Corporation (SRC).

Several people from industry and academia were helpful in shaping my views about the VLSI testing area in general and defect-based testing in specific. In particular, I would like to thank Dr. Anne Gattiker (IBM), my SRC Industrial Liaison, for sharing her views on some topics and for answering my queries. I would like to thank Dr. Phil Nigh (IBM) for providing SEMATECH data, a copy of his dissertation and for answering my numerous questions.

I was a summer intern at Texas Instruments, Inc. (Dallas) in 2000 and 2001. Working in industry while being in a graduate school was indeed a valuable experience as I got an opportunity to see some of the “real-world” issues. I would like to thank Hari Balachandran, Dr. Ken Butler, David Colby, Craig Force, Shri Jandhyala, Dr. Theo Powell, Ron Richmond, Dr. Jayashree Saxena, and Neil Simpson for discussions that gave me useful information. I am thankful to Bob Madge and Manu Rehani of LSI Logic for providing valuable data and Kevin Cota for useful discussions. Thanks to Dr. Peter Maxwell of Agilent Technologies for sharing his knowledge and allowing reproduction of figures from his papers. Thanks to David Bergman (NIST), Dr. Ali Keshavarzi (Intel), Thomas Vogel (CMU) for allowing reproduction of figures from their publications.

I would also like to thank Dr. Jim Plusquellic of University of Maryland Baltimore County for

his insightful comments. My friend Dr. Subashish Mitra of Intel extended all the help he could and I am thankful to him for the same. I wish to thank Wayne Needham who answered my questions in spite of having officially retired from the testing field!

I would like to thank my undergraduate advisors Dr. C. P. Gadgil (College of Engineering, Pune) and Dr. K. Gopakumar (Indian Institute of Science, Bangalore) who motivated me for further studies. Thanks to Prof. N. J. Rao and Dr. L. Umanand for encouraging me for further studies and for writing recommendation letters.

I would like to express my sincere thanks to my friends and colleagues Hoki Kim, Chul Lee, Abhijit Prasad, Ajay Thadhlani, Zoran Stanojevic, Wangqi Qiu, Ms. Jing Wang and Bin Xue. Several discussions with Wangqi Qiu during study breaks kept me enthused about my research.

Several friends helped me during graduate school directly or indirectly. Thanks to Prabha Acharya, Rajmohan Bharadwaj, Rajendra and Deepali Bhamare, Yogen Deshpande, Anagha Phadke, Supriya Phadke, Ramesh Setigam and Shailesh Waikar. I am especially grateful to my friends Jayesh Kulkarni and Sachin Upadhye for their encouragement and unfading friendship.

My cousin Sonalee and brother-in-law Himanshu were helpful in every possible way. Sonalee’s phone calls provided some sort of refreshment in my studies. I am thankful to both of them. Mrunal, my wife, made the last phase of my Ph.D. sustainable and more manageable by taking care of several minute details. I am thankful to her.

My decision to join the Ph.D. program was supported by my grandmother, parents and other family members. I certainly believe that I have come this far only due to their guidance and best wishes. I am deeply indebted to them for their love and moral support throughout my life. I earnestly hope this moderate achievement brings some happiness to their lives. As I finish my graduate studies and know there are still many challenges to solve the problem I embarked upon, I realize that true wisdom comes not from knowing all the answers but by knowing how many questions remain. I have often thought that a Ph.D. is more than a degree – it is an attitude to solve problems. Having gone through the process myself, now I know, that it is really so and it imposes more responsibility than privilege. I am glad to have had the opportunity to experience it.

# TABLE OF CONTENTS

Page


1. INTRODUCTION ........ 1
    - 1.1 Introduction ........ 1
    - 1.2 Motivation ........ 2
    - 1.3 Objectives of the Research ........ 7
    - 1.4 Contributions of the Dissertation ........ 8
    - 1.5 Structure of the Dissertation ........ 8
2. BACKGROUND ........ 10
    - 2.1 Introduction ........ 10
    - 2.2 IC Manufacturing Flow ........ 10
    - 2.3 Need for Testing ........ 10
    - 2.4 Contamination, Defects, Faults and Degradation ........ 13
    - 2.5 Test Quality ........ 16
    - 2.6 Test Method Classification ........ 17
    - 2.7 Principle of $I_{DDQ}$ Testing ........ 18
    - 2.8 $I_{DDQ}$ Testing Benefits ........ 19
    - 2.9 $I_{DDQ}$ Test in Practice ........ 21
    - 2.10 Components of $I_{DDQ}$ ........ 21
    - 2.11 Impact of Technology on $I_{DDQ}$ Test ........ 22
3. REVIEW OF $I_{DDX}$ –BASED TEST METHODS ........ 27
    - 3.1 Introduction ........ 27
    - 3.2 $I_{DDQ}$ Reduction Techniques ........ 28
    - 3.3 Reverse Body Biasing (RBB) Technique ........ 28
    - 3.4 Stack Effect ........ 29
    - 3.5 Multiple-$V_{TH}$ CMOS Transistors ........ 30
    - 3.6 Use of Alternative Technology ........ 31
    - 3.7 Partitioning of Power Supply Network ........ 32
    - 3.8 $I_{DDQ}$ Variance Reduction/Data Analysis Techniques ........ 32
    - 3.9 Current Signature ........ 32
    - 3.10 Delta $I_{DDQ}$ ........ 34
    - 3.11 Statistical Clustering ........ 35
    - 3.12 Current Ratio ........ 36

Page

3.13 Eigen Signatures ... 39
3.14 $I_{DDQ}$ versus Temperature ... 41
3.15 $I_{DDQ}$ versus $F_{max}$ ... 43
3.16 $I_{DDQ}$ versus Voltage ... 45
3.17 Principal Component Analysis-Based Linear Prediction ... 46
3.18 Transient Current ($I_{DDT}$)-Based Test Methods ... 47
3.19 $I_{DD}$ Pulse Response Testing (PRT) ... 47
3.20 Dynamic Power Consumption Current-Based Testing ... 48
3.21 Transient Signal Analysis (TSA) ... 49
3.22 Frequency Spectrum Analysis of Dynamic Current ... 49
3.23 On the Fly Depowering ... 49
3.24 Energy Consumption Ratio (ECR) ... 50
3.25 Fault-Free Parameter Value ... 51
4. OUTLIER IDENTIFICATION AND PARAMETRIC TEST DATA ... 55
4.1 Introduction ... 55
4.2 What Are Outliers? ... 55
4.3 Outlier Identification, Detection and Rejection ... 55
4.4 Why Do Outliers Exist? ... 57
4.5 Challenges in Outlier Rejection ... 58
4.6 Outlier Detection Methods ... 64
4.7 Resistant Outlier Detection ... 65
4.8 Philosophy Behind Outlier Rejection ... 65
4.9 Chauvenet's Criterion ... 65
4.10 Tukey Method ... 66
4.11 Median of Absolute Deviations (MAD) About Medians ... 67
4.12 Applicability of Outlier Rejection Methods to VLSI Testing ... 68
4.13 Measuring Outlier Screening Effectiveness ... 70
5. MULTIPLE PARAMETER CORRELATION ... 72
5.1 What Is Parameter Correlation? ... 72
5.2 Why Correlate Multiple Parameters? ... 72
5.3 Correlating Multiple Factors ... 73
5.4 Multiple Test Parameter Correlation ... 87

Page


5.5 Use of Multiple Test Metrics ... 92

5.6 CR-NCR Combination ... 92

5.7 Combination of CR, NCR and Delay Data ... 93

5.8 Conclusion ... 97

6. ANALYSIS OF INDUSTRIAL TEST DATA ... 98

6.1 Introduction ... 98

6.2 Analysis Using Spatial Fit Method ... 99

6.3 Evaluation of Statistical Outlier Rejection Methods ... 107

6.4 Correlating Two Test Parameters ... 110

6.5 Exploiting Wafer Spatial Correlation Using NCR ... 115

6.6 Immediate Neighbor Difference $I_{DDQ}$ Test (INDIT) ... 120

6.7 Combination of CR and NCR ... 124

6.8 Combination of CR, NCR and Delay Data ... 136

6.9 Conclusion ... 141

7. APPLICATION TO BURN-IN REDUCTION ... 144

7.1 Introduction ... 144

7.2 Methodology ... 144

7.3 Burn-in Pass/Fail Decision ... 147

7.4 Experimental Results ... 148

7.5 Discussion of Results ... 149

7.6 Conclusion ... 149

8. CONCLUSIONS AND FUTURE WORK ... 150

8.1 Summary ... 150

8.2 Future Directions for Research ... 151

8.3 Conclusion ... 157

REFERENCES ... 158

APPENDIX A SEMATECH EXPERIMENT S-121 ... 176

APPENDIX B LSI DATA ... 180

APPENDIX C TI DATA ... 181

APPENDIX D 3-D LINEAR REGRESSION ... 182

APPENDIX E CHAUVENET'S CRITERION ... 184

APPENDIX F MAD SCORE COMPUTATION ... 185

Page

APPENDIX G COST MODEL FOR OUTLIER REJECTION BENEFIT EVALUATION..186
VITA……………………………………………………………………………………………188

## LIST OF FIGURES

Page

Fig. 1. Basic manufacturing test flow for semiconductor chips. 2
Fig. 2. Illustration of IC parameter variation at various levels. 4
Fig. 3. Reliability bathtub curve and reliability screen using burn-in. 5
Fig. 4. IC parameters show a continuous distribution. 6
Fig. 5. Conflicting goals of yield loss and defect level. 7
Fig. 6. Typical CMOS IC fabrication flow. 11
Fig. 7. (a) Original mask pattern and (b) actual silicon. 12
Fig. 8. Process variation induces inter-layer variation in metal thickness. 13
Fig. 9. Global and local defects and associated yield loss. 14
Fig. 10. Yield loss due to parametric, systematic and random defects. 15
Fig. 11. Subtle and gross defects. 16
Fig. 12. Shmoo plot showing relationship between supply voltage and chip speed. 18
Fig. 13. (a) CMOS inverter circuit and (b) faulty and fault-free waveforms. 19
Fig. 14. Leakage current mechanisms for DSM transistors. 22
Fig. 15. Reduction in threshold voltage increases leakage current exponentially. 23
Fig. 16. Defect detection using $I_{DDQ}$ test depends on defect severity. 24
Fig. 17. (a) Single threshold test for older and (b) for DSM technologies. 25
Fig. 18. Chip-to-chip $I_{DDQ}$ variation for a vector for five wafers. 26
Fig. 19. Leakage current reduction by stacking of transistors. 29
Fig. 20. Basic principle of MTCMOS circuits showing sleep control transistors. 30
Fig. 21. Cross section of SOI transistor and leakage components. 31
Fig. 22. Current signatures for fault-free and faulty chips. 33
Fig. 23. Histograms of delta-$I_{DDQ}$ for fault-free and faulty chips. 35
Fig. 24. Normalized histogram for four clusters of $I_{DDQ}$ test data. 36
Fig. 25. Two chips having different $I_{DDQ}$ magnitude but similar CR. 37
Fig. 26. Current ratios of fault-free and faulty chips from a wafer. 38
Fig. 27. Reduction in CR with increasing background leakage. 39
Fig. 28. Fault-free chips show regular variation in $I_{DDQ}$. 40
Fig. 29. Scatter plot of mean and standard deviation of $I_{DDQ}$ for SEMATECH data. 41
Fig. 30. $I_{DDQ}$ values at two different temperatures for outlier screening. 42
Fig. 31. Temperature-based outlier screening. 43
Fig. 32. Relationship between $I_{DDQ}$ and Fmax. 44

Page

Fig. 33. Adjustable limit setting for leakier and faster chips. ........ 45
Fig. 34. I-V curve for some uniquely detected defective chips. ........ 46
Fig. 35. Test vector to fabrication process correlation. ........ 47
Fig. 36. Comparison of good circuit $I_{DDT}$ response to faulty circuit $I_{DDT}$ response. ........ 48
Fig. 37. Overview of $I_{DDF}$ testing method. ........ 50
Fig. 38. Distribution of $I_{DDQ}$ values from different chips shows a long tail. ........ 51
Fig. 39. Test strategy can be modified depending on “outlier-ness”. ........ 57
Fig. 40. Outlier behavior can stem from different reasons. ........ 58
Fig. 41. (a) A symmetrical distribution gets skewed (b) due to outliers. ........ 59
Fig. 42. $I_{DDQ}$ variation across a wafer. ........ 60
Fig. 43. $I_{DDQ}$ variation across a lot. ........ 61
Fig. 44. Successive outlier rejection can nibble away data. ........ 62
Fig. 45. Transformation of Lognormal data to Normal distribution. ........ 64
Fig. 46. Parameter definition for Tukey method. ........ 67
Fig. 47. Outliers can be subtle or gross. Gross outliers hide the subtle outliers. ........ 69
Fig. 48. Illustrations for positive (a), negative (b) and no (c) correlation. ........ 72
Fig. 49. Two dimensional grayscale map for $I_{DDQ}$ across a wafer. ........ 74
Fig. 50. Neighborhood die definition. ........ 75
Fig. 51. Wafer surface plot for a single $I_{DDQ}$ vector. ........ 76
Fig. 52. Wafer surface plot for estimated $I_{DDQ}$ values. ........ 76
Fig. 53. Wafer surface plot for residual values for a vector. ........ 77
Fig. 54. Variation in minimum and maximum $I_{DDQ}$ for SEMATECH chips. ........ 78
Fig. 55. Variation in minimum and maximum $I_{DDQ}$ for LSI chips. ........ 79
Fig. 56. Variation in CRs across a wafer. ........ 81
Fig. 57. Variation in NCRs across a wafer. ........ 82
Fig. 58. INDIT algorithm. ........ 83
Fig. 59. Wafer surface plot of maximum self-deltas for a wafer. ........ 83
Fig. 60. Wafer surface plot of maximum neighbor-deltas for a wafer. ........ 84
Fig. 61. Wafer signatures for two wafers for two vectors. ........ 85
Fig. 62. Wafer signature for a wafer from LSI data. ........ 86
Fig. 63. Flush delay and $I_{DDQ}$ correlation for SEMATECH data ($R^2 = 0.40$). ........ 88
Fig. 64. Correlation between ring oscillator frequency and $I_{DDQ}$ for LSI data ($R^2 = 0.48$). ........ 88
Fig. 65. Leakage current variation for 20 vectors for two wafers. ........ 89

Page

Fig. 66. Systematic shift in $I_{DDQ}$ across lots. 90

Fig. 67. Radial variation in $I_{DDQ}$. 91

Fig. 68. IC leakage with and without RBB of 0.5 V at room temperature. 92

Fig. 69. Categorization of chips based on CR and NCR. 93

Fig. 70. Typical analysis flow for multiple test metric screening. 94

Fig. 71. CR-NCR scatter plot for chips that passed all wafer tests or failed only $I_{DDQ}$ test. 96

Fig. 72. CR, NCR and flush delay scatter plot for chips having CR≤10 and NCR≤10. 96

Fig. 73. $I_{DDQ}$ projections on wafer XY plane. 100

Fig. 74. $I_{DDQ}$ projections on wafer XY plane after outlier removal. 101

Fig. 75. Execution flow for outlier screening in spatial fit method. 102

Fig. 76. Distribution of residuals for spatial fit approach for a vector across wafer. 104

Fig. 77. A three-dimensional view of residual variation for a vector across a wafer. 104

Fig. 78. Wafer-level minimum $I_{DDQ}$ distribution for SEMATECH chips. 111

Fig. 79. Correlating actual and estimated $I_{DDQ}$ values. 112

Fig. 80. Actual and estimated $I_{DDQ}$ values for a chip. 113

Fig. 81. Residual distribution for all chips. 114

Fig. 82. Current Ratios for IBM/SEMATECH chips in the dataset. 117

Fig. 83. Distribution of post-BI results of chips for different maximum NCR values 120

Fig. 84. CDFs for self and neighbor deltas. 122

Fig. 85. CDFs for chips with different wafer probe test results. 122

Fig. 86. Distribution of post-BI results of chips for different maximum neighbor-delta values. 124

Fig. 87. Scatter plot of CR/NCR values for all pass or $I_{DDQ}$-only fail SEMATECH chips. 125

Fig. 88. CR/NCR scatter plot for LSI data showing number of chips in each region. 126

Fig. 89. Distribution of CR values for LSI data. 127

Fig. 90. Distribution of NCR values for LSI data. 127

Fig. 91. Distribution of CRs in the range 1-1.02 for LSI data. 129

Fig. 92. Distribution of NCRs for LSI chips having CR less than 1.02. 130

Fig. 93. Distributions of CR and NCR values less than 20. 131

Fig. 94. Current signatures for die from regions shown in Fig. 69. 132

Fig. 95. CR/NCR scatter plot for SEMATECH chips that passed all tests. 134

Fig. 96. CR/NCR scatter plot for chips that failed only $I_{DDQ}$ test before and after BI. 135

Fig. 97. CR/NCR scatter plot for chips that failed Boolean test after BI. 136

Fig. 98. Flush delay distribution for the SEMATECH sample. 137

Page

Fig. 99. Cumulative distribution for CR values for the SEMATECH sample............................ 138

Fig. 100. Cumulative distribution of NCR values for the SEMATECH sample. ........................ 138

Fig. 101. CR-NCR scatter plot for CROWNE chips. ................................................................ 140

Fig. 102. Distribution of flush delays for delay failed chips...................................................... 142

Fig. 103. CR-NCR scatter plot for delay fail and/or $I_{DDQ}$ fail chips. .......................................... 142

Fig. 104. NCR and flush delay scatter plot for delay fail and/or $I_{DDQ}$ fail chips.......................... 143

Fig. 105. $I_{DDQ}$ before and after burn-in for a sample of chips that passed all tests at wafer probe; chips for which post burn-in $I_{DDQ}$ < 8 μA are shown..................................... 146

Fig. 106. $I_{DDQ}$ before and after burn-in for a sample of chips that failed only $I_{DDQ}$ test at wafer probe; chips for which post burn-in $I_{DDQ}$ is less than 50 μA are shown. .......... 146

Fig. 107. Comparison of effectiveness of various methods......................................................... 148

Fig. 108. Vector-to-vector correlation between different chips.................................................. 152

Fig. 109. Histograms of vectors resulting in minimum $I_{DDQ}$. ..................................................... 153

Fig. 110. Histogram of vectors resulting in maximum $I_{DDQ}$. ...................................................... 154

Fig. 111. Histogram of vectors resulting in max NCR. .............................................................. 154

Fig. 112. $I_{DDQ}$ surface plot for a wafer showing outlier behavior of wafer edge dice.................. 155

Fig. 113. Test flow for SEMATECH experiment S-121. ........................................................... 176

# LIST OF TABLES

Page

Table I. ITRS projections for $I_{DDQ}$ of high-performance ICs. ....... 24
Table II. Summary $I_{DDQ}$-based test methods. ....... 52
Table III. Summary of $I_{DDT}$-based test methods. ....... 54
Table IV. Properties of distributions shown in Fig. 44. ....... 63
Table V. Various possibilities as implied by NCR. ....... 80
Table VI. Test strategy when test metrics disagree. ....... 94
Table VII. Combination of CR, NCR and delay data. ....... 95
Table VIII. Comparison of predicted and actual $I_{DDQ}$ ....... 103
Table IX. Comparison of different limit setting schemes Considering post-BI $I_{DDQ}$ fails ....... 105
Table X. Comparison of different limit setting schemes ignoring post-BI $I_{DDQ}$ fails ....... 106
Table XI. Results for Chauvenet's criterion ....... 109
Table XII. Results for Tukey test ....... 109
Table XIII. Results for static threshold method ....... 109
Table XIV. Distribution of chips for two methods ....... 115
Table XV: Distribution of dice in the original dataset ....... 116
Table XVI. Distribution of chips for different test methods ....... 118
Table XVII. Overkill and DL for different test methods ....... 119
Table XVIII. Distribution of wafer test and post BI results of chips in the dataset ....... 121
Table XIX. Distribution of chips according to SEMATECH test results for self-delta and neighbor-delta test methods ....... 123
Table XX. DL and OveRkill comparison of two methods ....... 124
Table XXI. Distribution of SEMATECH chips. ....... 132
Table XXII . DL and YL for various categories. ....... 134
Table XXIII. Distribution of sample according to chip test results (12521 chips) ....... 139
Table XXIV. Comparison of DL and YL for various methods. ....... 148
Table XXV. Description of SEMATECH test methods. ....... 177
Table XXVI. Distribution of chips at different test levels. ....... 178
Table XXVII. Distribution of BI sample based on test results. ....... 179
Table XXVIII. MAD-based rejection example. ....... 185
Table XXIX. List of variables used in cost model. ....... 186

# 1. INTRODUCTION

## 1.1 Introduction

Continuous advances in semiconductor manufacturing technology permit reduction in transistor geometries. This allows higher levels of integration of transistors on a chip. The number of transistors per unit area doubles approximately every 18 months [1]. The International Technology Roadmap for Semiconductors (ITRS) outlines the projections for the future improvements in Very Large Scale Integration (VLSI) semiconductor technology [2]. The recent projections indicate that this progress will continue at least for a decade or more. Of course, several challenges exist for the development to occur. One of these is testing of these chips, an area this dissertation will elaborate on.

Semiconductor manufacturers must ensure that chips shipped to customers conform to their specifications. This is achieved by testing them for various parameters at different stages of chip manufacturing as shown in Fig. 1. These tests can be classified into two broad categories: Boolean tests and parametric tests. We consider Boolean tests as those tests whose failure clearly means device malfunction. These tests have a pass/fail type decision. Since failure of this type of test means the circuit under test (CUT) has a *hard fault* and cannot perform its desired function, it is rejected. Examples of this type of test include functional [3] and stuck-at tests [4]. On the other hand, parametric tests measure a certain parameter of the CUT. If the parameter falls outside the acceptable range, the CUT has a *soft fault* (parametric fault). Examples of parametric test include quiescent leakage current ($I_{DDQ}$) test and speed ($F_{max}$) test. The failure of a parametric test does not necessarily imply that the device cannot perform its intended function. However, it does not meet its specifications completely [5]. For example, a chip may function correctly but consume more power or operate at a slower speed.

Generally, parametric tests are not part of specification tests. The threshold for a parametric test is always a subjective decision. Too stringent a threshold rejects many fault-free chips; resulting in lost revenue and too loose a threshold accepts many faulty chips, resulting in

---

This dissertation follows the style and format of *IEEE Transactions on VLSI Systems*.

shipment of defective chips. Because parametric failed chips are functional, the distinction between "fault-free" and "faulty" chips is not straightforward, which makes deciding the pass/fail threshold for a parametric test challenging.

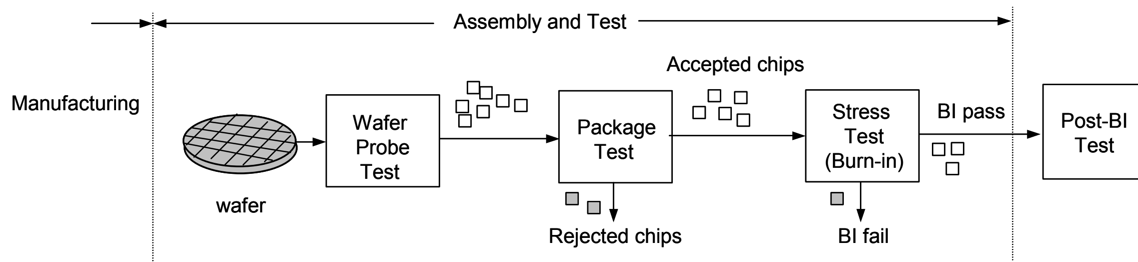


**Fig. 1. Basic manufacturing test flow for semiconductor chips.**

Apart from screening faulty chips, testing also provides a way to monitor the semiconductor manufacturing process. As semiconductor wafers are processed through the production line in batches called 'lots' (each containing 25-50 wafers) they undergo hundreds of processing steps. Detailed data is collected following many of the process steps and monitored using Statistical Process Control (SPC) [6]. The purpose of SPC is to maintain and insure or improve process quality [7]. Test data is analyzed to understand process variation and determine process corners. It is crucial for semiconductor manufacturers to set reasonable thresholds for SPC as it has direct impact on the outgoing quality and parametric yield.

## 1.2 Motivation

Testing is one of the final stages in the production cycle and is of fundamental importance to assure that customers will get working units. No product can be manufactured without an efficient test strategy that guarantees the percentage of defective products passing undetected (measured as *defective parts per million* or *DPM)* is sufficiently low to be acceptable. The test strategy that simplifies the production test results in lower test costs and products that are more reliable. Thus, in the long run it influences customer satisfaction. The value of a test can be defined as the ratio of improvement in quality (quantified as reduction in DPM) to the cost of the test. It is possible to reduce DPM by adding more tests but that increases test cost and reduces its value. Screening all defects is unnecessarily expensive. Manufacturers wish to maximize the value of test by balancing outgoing quality and the associated test costs.

Testing is a complex task. The growing complexity of semiconductor chips makes it even more challenging. The cost of testing does not fall as fast as the cost of manufacturing transistors. This results in increased overall percentage of manufacturing cost attributed to testing. To remain competitive in the market semiconductor manufacturers need to be able to keep test cost as low as possible. Tests themselves are not perfect and therefore cannot catch all defects [8]. Although testing does give some confidence to manufacturers about the quality of the shipped lot, it is always an overhead cost. As illustrated in Fig. 1, testing is carried out at different processing steps. Typically the first tests applied to a chip after continuity test check the integrity of the chip structure [9]. Chips with gross defects usually fail this initial screening. As a rule of thumb, test cost increases by an order of magnitude with each integration level (wafer, package, board and system level) [10]. If defective chips are identified earlier in the test cycle, investment in the later steps can be saved. By detecting defects at the wafer level, the cost of manufacturing is kept the lowest as investment in packaging and further testing a defective part is saved. The packaging costs vary greatly depending on the package type. For example, a 16-pin ceramic DIP (dual-in-line package) is about 2 cents while a 400-pin ceramic PGA can cost as much as $50 [11]. Therefore, identifying defective chips earlier in the manufacturing cycle becomes even more important as the relative percentage of packaging cost increases [12]. This forms the basic motivation for semiconductor manufacturers to screen defective chips at the wafer level test called *wafer sort*.

Transistor geometries are shrunk with each technology node to obtain higher performance and density. However, controlling transistor geometries precisely becomes harder as transistors are scaled down. Thus, the impact of process variations on integrated circuit (IC) performance increases for deep sub-micron (DSM) technologies [13]. This is further exacerbated by the increasing number of process steps, complex transistor geometries and introduction of new materials. The process variations occur due to variations in physical and environmental conditions such as impurity concentrations, oxide thickness, diffusion of impurities and the imperfect spatial uniformity of the processing steps such as gas flows, etching and deposition of

materials, ion implantation, and chemical mechanical polishing (CMP). Variations in the dimensions occur due to limited resolution of the photolithographic process. This gives rise to variation in the chip performances between lots, wafers, across a wafer (*inter-die*[1]) and within a chip (*intra-die*) as shown in Fig. 2. The impact of these variations on chip performance can be as high as 30-35% [14]. Parametric tests or guard bands are used to compensate for discrepancies due to possible test equipment variations within their specified limits. The purpose of guard bands is to minimize the number of devices that fall out due to marginality [3]. Parametric tests must determine which variations are acceptable and which are not. Chips having unacceptable variation in parameter(s) are *outliers*. Empirical evidence suggests that outlier chips have low reliability [15]. Thus to improve or insure quality it is important to screen outliers.

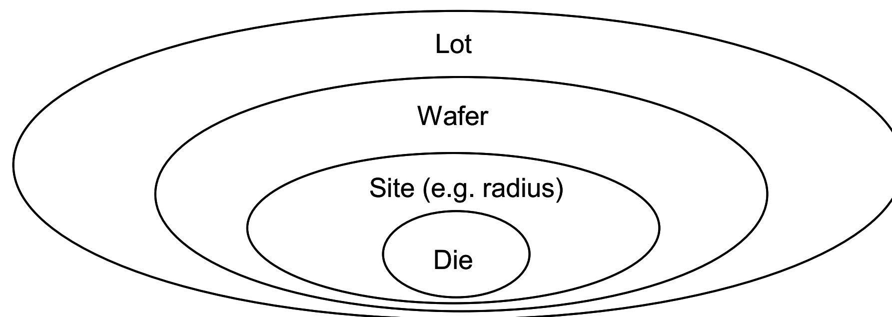


**Fig. 2. Illustration of IC parameter variation at various levels.**

The traditional test method to screen low-reliability chips is burn-in (BI) test. The basis for the concept of BI test is that a chip will most likely fail in the early hours of its life, if it is going to fail at all [16]. BI involves subjecting chips to high temperature and voltage stress to accelerate defects. Chip failure rate is most severe during the infancy of a device as shown in the bathtub curve in Fig. 3. Chips that fail BI test earlier are called *infant mortality.* BI accelerates

---

[1] We use the words ‘die’ and ‘chip’ synonymously in this dissertation.

the aging of the device or compresses the time scale so that infant mortality can be screened. Since chips are aged beyond their infant-mortality life stage, the chips that survive BI are expected to be more reliable and have very low and steady failure rates.

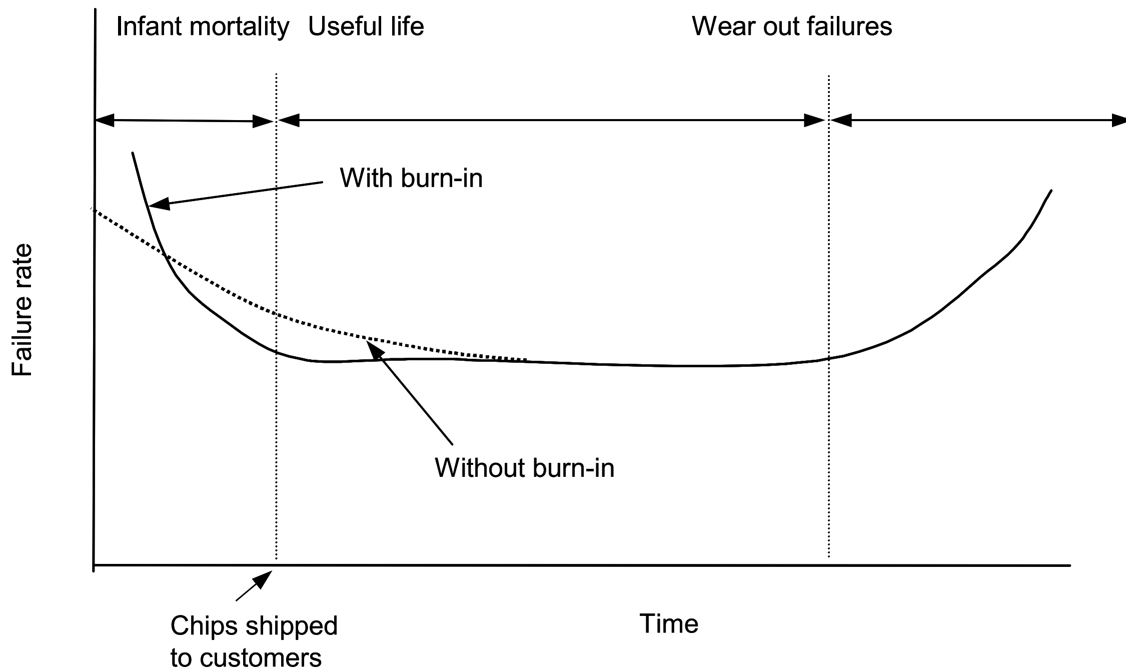


**Fig. 3. Reliability bathtub curve and reliability screen using burn-in.**

The effectiveness of BI depends on the stress exerted on a chip, which is a function of the difference between nominal voltage (temperature) and elevated voltage (temperature). As supply voltages are reduced with each technology node and operating temperatures are increased, the effective stress is reduced, thus making BI less effective. Secondly, BI is getting prohibitively expensive. This has fueled research to seek alternatives to BI. Chips having higher leakage current are observed to have low reliability and, therefore, the use of quiescent leakage current ($I_{DDQ}$) test has been investigated as an alternative to BI [17]. However, increasing fault-free leakage current and current variation due to technology advances makes distinction between faulty and fault-free leakage difficult and hence this option less viable for state-of-the-art technologies. One line of research is to screen chips based on the comparison of their behavior with the population. If a chip is an outlier, semiconductor manufacturers would rather accept the

yield loss incurred by rejecting it than risk shipping it. Due to higher cost of a customer return and long-term likely impact on customer-perceived quality, test economics is complicated. Statistical post-processing (SPP) of wafer-level test data has been shown to be a powerful alternative to BI in the production flow [15].

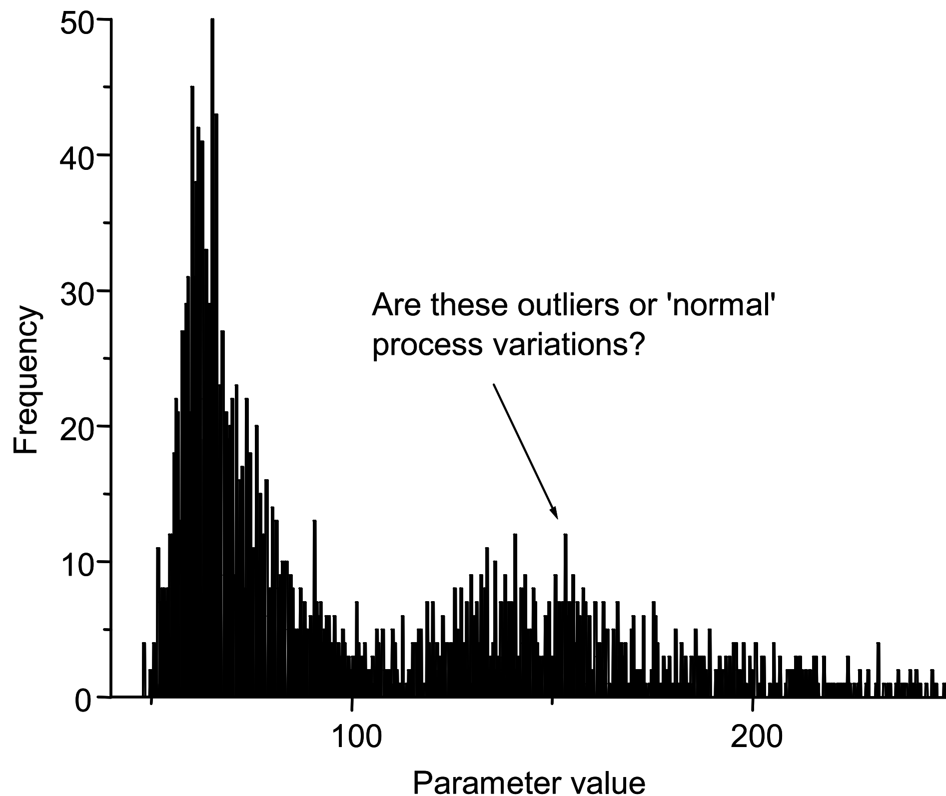


**Fig. 4. IC parameters show a continuous distribution.**

To determine whether a chip is an outlier it is necessary to define the “normal” behavior of a chip. Fig. 4 shows the distribution of a parameter for all chips from a wafer. Note that due to continuous parameter distribution many chips appear to be outliers. However, their seemingly strange behavior may be explained by understanding underlying process variations or by comparison with some parameters of other fault-free chips. Hence, outlier identification is a difficult challenge. This becomes even more challenging as subtle defects become crucially important for advanced technologies [18]. Differentiating between “true outliers” (defective chips) and “apparent outliers” is the primary goal of this research. Here further we use the word “outlier” to mean a defective chip.

1.3 Objectives of the Research

The main goal of this research is to evaluate outlier-screening capability of different data analysis methods for measured IC parameters. Ideally, outlier screening should result in a reduced number of defective parts shipped to customers (i.e. lower DPM) as well as reduced number of good parts getting rejected (i.e. yield loss). Due to continuous parameter distributions, these are conflicting goals as shown in Fig. 5. Therefore, semiconductor manufacturers need to balance these goals.

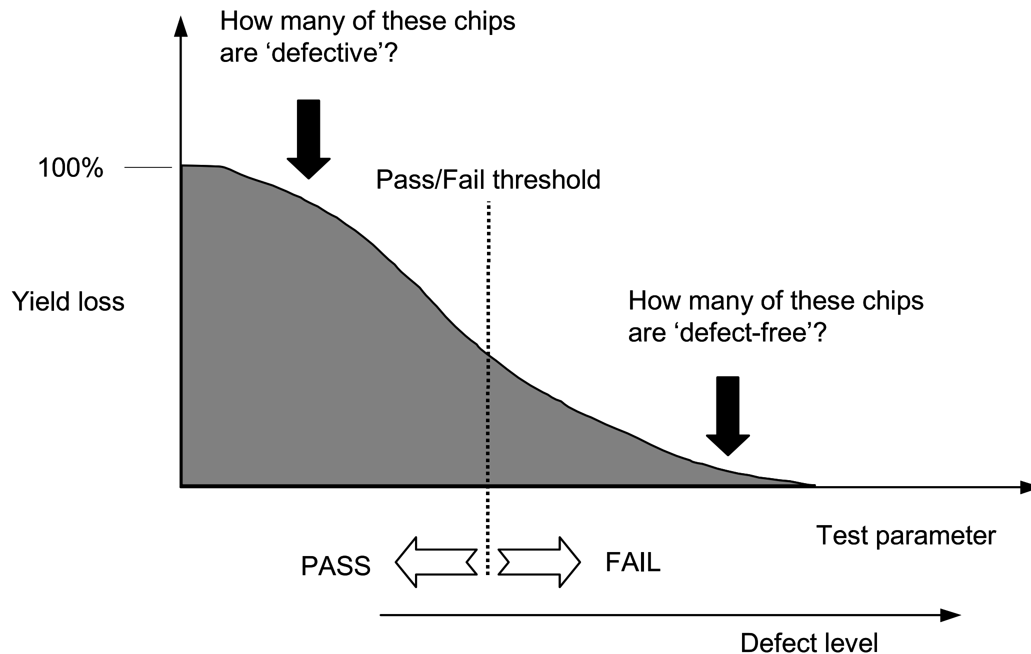


**Fig. 5. Conflicting goals of yield loss and defect level.**

In this dissertation, we evaluate some methods to determine their effectiveness in identifying outlier chips. These methods are evaluated using industrial test data from IBM (SEMATECH experiment), LSI Logic and Texas Instruments (TI). The information about the data is included in Appendices A, B and C.

The quiescent leakage current ($I_{DDQ}$) forms our basic parameter for outlier identification. However, as will be shown later, $I_{DDQ}$ data alone may not be sufficient to screen outliers. We therefore define additional test metrics and use a combination of multiple test parameters (e.g. delay and leakage current) for screening outliers with higher resolution. We primarily use wafer-

level test data for outlier identification and supplement it with wafer-level spatial data. If additional tests are conducted after packaging, we use them to evaluate the effectiveness of the outlier detection methods. While extensive BI and other test data is available for the SEMATECH research experiment, our evaluation is limited for the LSI Logic and TI data since further testing is not carried out when a chip fails in the test flow ("*stop on first fail*" test flow) and no BI stress test is performed.

## 1.4 Contributions of the Dissertation

The focus of this dissertation is to evaluate different techniques to screen defective chips using parametric test data and statistical outlier rejection methods. To achieve it, this dissertation exploits wafer-level spatial correlation. The methods evaluated, metrics proposed or examined primarily revolve around this basic theme.

The contributions of this dissertation can be summarized as follows:

- Evaluation of applicability of statistical outlier rejection methods to VLSI testing using parametric data
- Exploitation of wafer-level spatial correlation for outlier screening
- Evaluation of a novel self-scaling test metric called Neighbor Current Ratio (NCR)
- Evaluation of two new methods, Immediate Neighbor Difference $I_{DDQ}$ Test (INDIT) and Wafer Signature, for outlier identification
- Combination of multiple test metrics for obtaining some insights into defect type
- Evaluation of applicability of outlier rejection using Median of Absolute Deviation (MAD) resistant estimator for BI reduction

## 1.5 Structure of the Dissertation

The dissertation discusses the background in Section 2. We define the VLSI testing problem in general and describe different test methods used. This section should prepare the reader with the background necessary to understand the material covered in the later sections.

The focus of this research work is outlier identification. The outlier identification methods are applicable to any parametric test data. We use $I_{DDQ}$ test data extensively to illustrate the basic concepts in outlier analysis. The prior work reported in the literature for $I_{DDQ}$ data analysis is reviewed in Section 3. This section also discusses how these methods lose their effectiveness in screening defective chips (outliers) for new semiconductor technologies. This motivates the use of multiple test metrics.

Outliers are defined by knowing the distribution of nominal measured parameters. However, the distribution is affected by the presence of outliers. This is equivalent to the chicken and egg problem. Section 4 describes this difficulty in screening outliers. Some outlier identification criteria and different metrics are proposed in this section. Since outlier detection depends on resistant estimators, they are also discussed. Outlier detection itself is a separate topic in statistics and a wealth of literature is available [19]. Our goal here is just to provide a glimpse of this area and evaluate its applicability to VLSI testing.

Section 5 discusses multiple-parameter correlation and its effectiveness in screening outlier chips. Some discussion on why two parameters should exhibit correlation is provided. We also show how a combination of multiple test metrics is useful for increasing confidence in outlier detection.

We use the test data from IBM/SEMATECH, LSI Logic and TI for the experimental validation of our methods. Results of the analyses are presented in Section 6. The conclusions drawn are our own and do not necessarily represent views of these companies.

Section 7 presents our preliminary analysis of using an outlier screening method to reduce burn-in. The results of analysis of IBM/SEMATECH data are presented in this section.

Finally, Section 8 presents conclusions and contributions of this work. It also outlines the direction of future research and comments on impact of technology and usefulness of this research.

# 2. BACKGROUND

## 2.1 Introduction

This section provides an overview of VLSI testing. The information presented here should help the reader understand the key terms used in the later sections and gain an appreciation for the complexity and challenges involved in testing semiconductor chips. We use quiescent leakage current measurement data extensively in this dissertation. Therefore, we discuss $I_{DDQ}$ test in detail in this section.

## 2.2 IC Manufacturing Flow

The Integrated Circuit (IC) manufacturing process consists of a series of photolithographic printing, etching, and doping (impurity addition) steps. A typical fabrication process for manufacturing a CMOS (Complementary Metal Oxide Semiconductor) transistor pair is illustrated in Fig 6. A layer of silicon dioxide ($SiO_2$) is grown on the surface of a $P^-$ wafer. Photoresist is laid down on top of $SiO_2$ layer. Using ultraviolet (UV) light a pattern is projected onto the photoresist through a photographic mask. The area of photoresist exposed to UV light gets hardened. Using an organic solvent the nonexposed portion of the photoresist is washed away. After baking the remaining photoresist, exposed areas of oxide are removed using an etching process. The exposed areas of silicon are doped to form an N-well using either diffusion or ion implantation. Through a series of hundreds of steps of printing, masking, etching, implanting, and chemical vapor deposition (CVD), a complex IC is fabricated [20].

## 2.3 Need for Testing

What is illustrated in Fig. 6 is an idealized approximation of actual fabricated circuit. In today's state of the art technology, the transistor geometries, metal pitch, and depositions are extremely small. For example, the state of the art microprocessor uses 130 nm technology. At such small geometries, even a tiny dust particle, metallic sliver, or even a piece of human hair can lead to shorting of two wires. The wire patterns are not exactly rectangular when printed. Fig. 7 (a) is the intended wiring pattern while Fig. 7(b) is the actual fabrication as observed through a Scanning Electron Microscope (SEM). The SEM photo shows that wires have non-uniform width as well as a particle that shorts two wires (called a *bridging defect*).

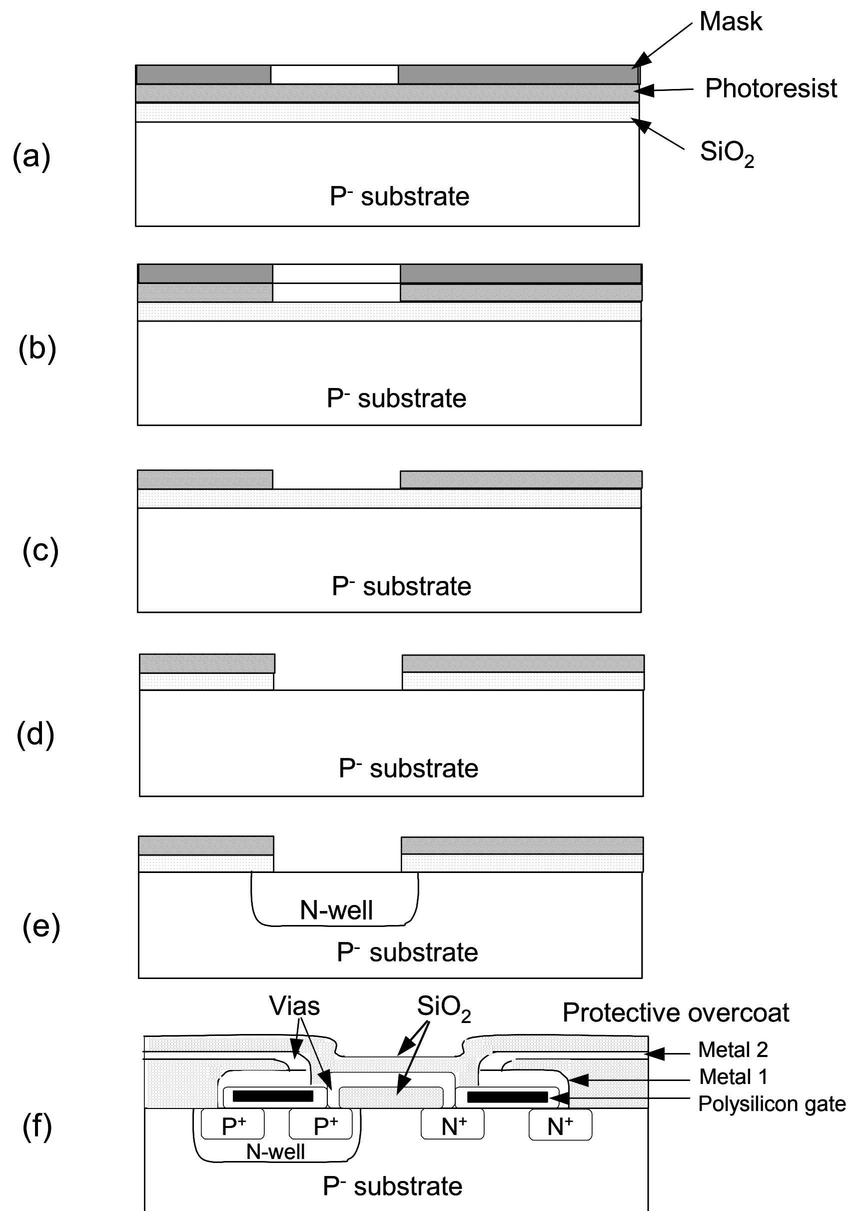


**Fig. 6. Typical CMOS IC fabrication flow.**

Variations in process conditions, physical deformations in the processing material and contaminations can cause a chip to malfunction [21]. Fig. 8 shows an example of interconnections on multiple metal layers. This clearly illustrates the nonuniformity caused in wires due to different processing conditions.

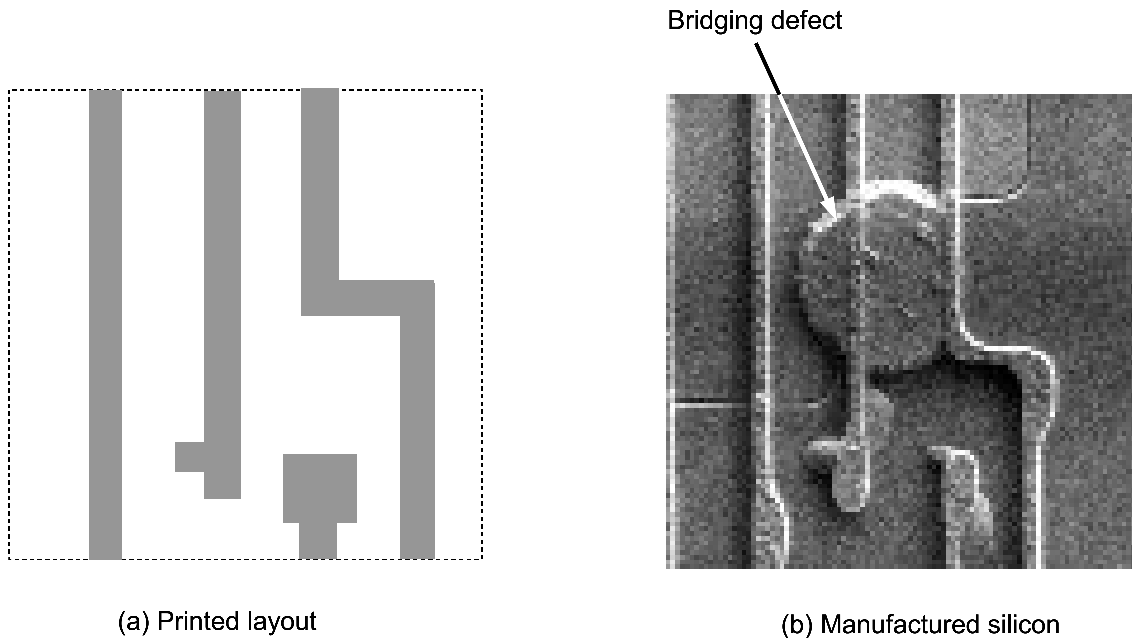


**Fig. 7. (a) Original mask pattern and (b) actual silicon.**

To ensure that all transistors and interconnect work in tandem to achieve the desired circuit functionality is a challenging task. To verify that a chip conforms to specifications, manufacturers need to test chips. However, from a manufacturer's perspective, the purpose of testing is not only to ensure conformance to specifications, but also to reduce the long-term cost involved in handling a customer return. Seen from this perspective, testing is an investment made in a "cost-avoidance" strategy [10]. In other words, testing is an investment made by the manufacturer to ensure customer satisfaction. Naturally, a semiconductor manufacturer wishes to maximize the return on investment (ROI), that is, minimize test cost while improving outgoing quality. Obviously, if test costs more than the manufacturing cost, it would be cheaper to replace the returned (defective) parts. In a nutshell, the role of testing is to screen chips that do not function and also chips that are likely to fail early in the customer system (reliability risk). Testing is also used for understanding, controlling and diagnosing process excursions to improve yield in the so-called *yield learning* phase.

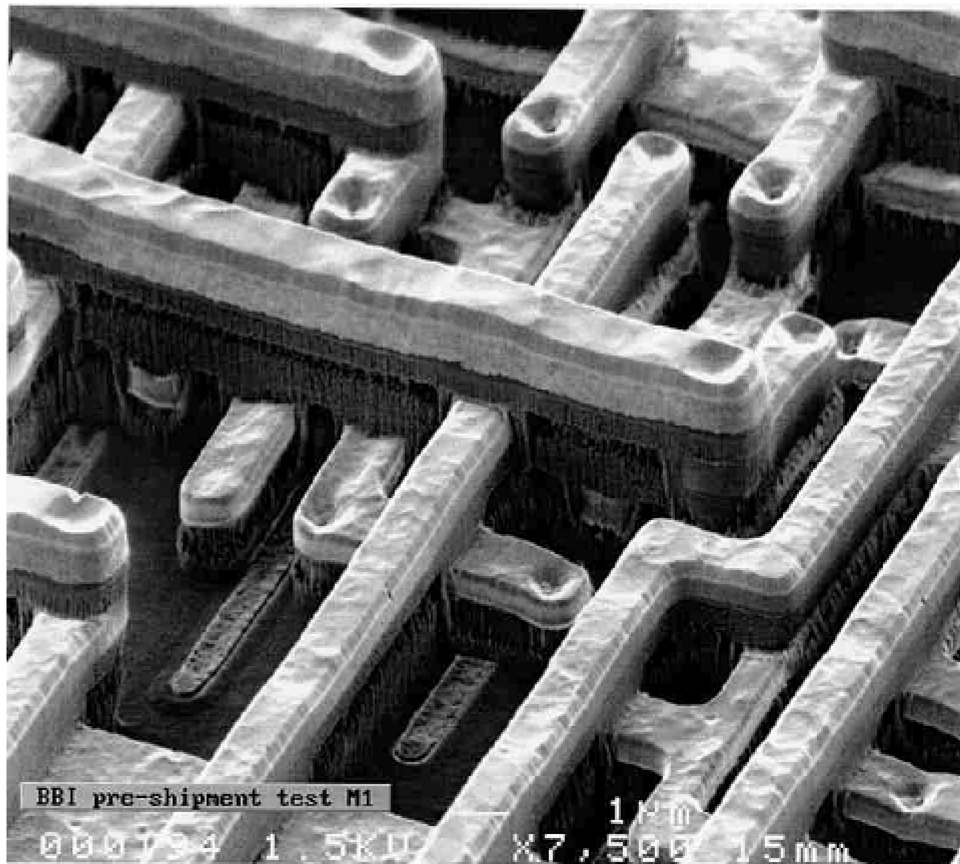


**Fig. 8. Process variation induces inter-layer variation in metal thickness.**

## 2.4 Contamination, Defects, Faults and Degradation

The purpose of test is detection of a malfunction in the operation of an IC. A distinction is made between defects, faults and degradation [22]. Contamination is a foreign material on a wafer surface. It can be from human skin, dirt, dust particle, residual chemicals, etc. A *defect* is a physical deformation that leads to device malfunction. A defect may be extra material (e.g. a short between two nodes called *bridge*) on a layer or between layers or missing material (e.g. discontinuity in metal or polysilicon line called *open defect*)

Different types of defects lead to similar IC abnormalities called *faults*. A fault is thus a higher-level abstraction of a defect. It is a hypothesis of how a logic gate would malfunction without assigning any specific attributes to the defect that can cause it [22]. A fault is said to occur only if a test to detect it exists (i. e. it is detectable) and it is detected. Obviously, if a fault

does not cause any change in the chip behavior then faulty and fault-free chips cannot be distinguished. These faults, called *redundant faults*, are of no consequence to the functional behavior of the circuit, although they may affect the reliability of the circuit.

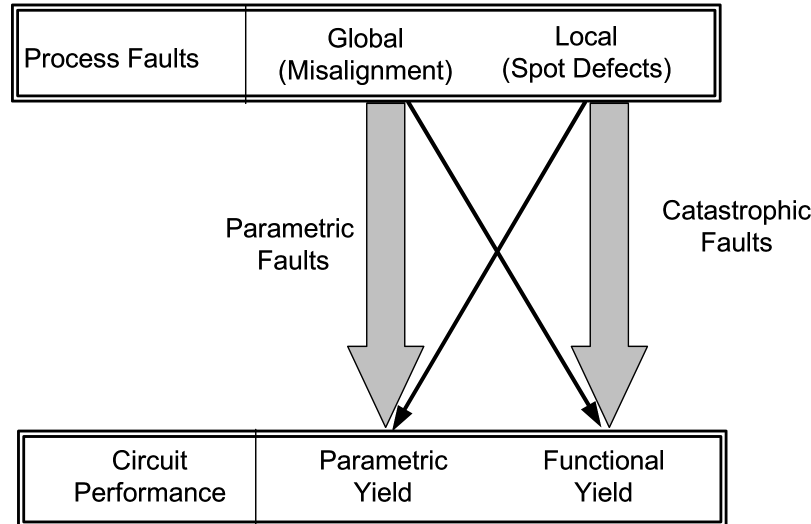


**Fig. 9. Global and local defects and associated yield loss.**

Many defects cause degradation (a weakness in the physical construction of the circuit) that does not lead to logical malfunction but only degrades system performance (e.g. propagation delay or noise margin). For example, a thinner metal line will result in higher resistance and increased delay. The circuit still functions but at a reduced speed. Such a chip, even if it passes all functional tests, is a reliability risk. CMOS chips are more prone to degradation faults [23]. In general, there are two types of faults than can occur: local process faults and global process faults [24]. Examples of global process faults include mask misalignment and line width variations. Local process faults include spot defects like oxide pinholes or bridges. As shown in Fig. 9, global process faults primarily affect parametric yield as they cause variation in speed or power consumption. This gives rise to *parametric faults*. Parametric faults are often the consequence of bad design, but they may also be an indication of a defect [25]. Local process faults primarily affect circuit topology, cause a chip to fail functionally, and affect functional yield. They are also called *catastrophic faults.* However, as thin arrows in Fig. 9 indicate, global faults can affect functional yield and local faults can affect parametric yield as well. Note that

whether a fault leads to structural or performance failure depends on the overall disturbance caused by the process deformation [26]. There is an overlap between yield loss caused by parametric, systematic and random defects as shown in Fig. 10. Recent studies indicate that systematic variation (e.g. via alignment, optical proximity correction (OPC) based defects, etc.) is a more dominant cause of failures for new technologies [27].

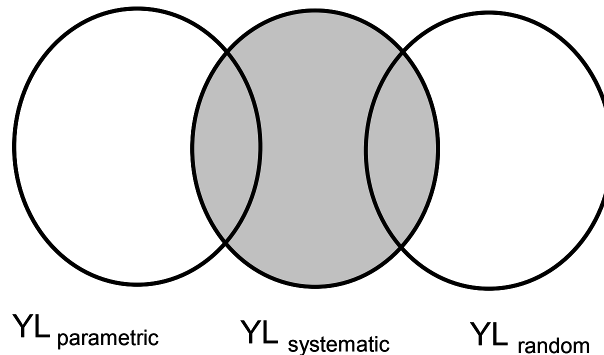


**Fig. 10. Yield loss due to parametric, systematic and random defects.**

It must be emphasized that not all defects lead to faults or functional failure. This can be conceptually illustrated with the help of Fig. 11. It shows a particle defect between two metal lines. Depending on the size of the particle and its location relative to the wires, it may or may not short the wires. If the particle barely touches the wires (subtle defect), it may not cause functional failure but remain latent until the chip is shipped. In some cases, it may cause delay fault [28].

Although there are infinite numbers of possible defects, many have a similar effect on the electrical (logical) behavior of a chip, and, therefore, a finite number of fault models are sufficient. Fault models enable us to quantify the quality of test suites and compare them. This is done using a *fault coverage* or *test coverage* metric, which is the percentage of faults detected by a test out of the total detectable faults in the circuit. Under the same fault model a test with higher fault coverage is better than one with lower fault coverage as it has a higher probability of detecting a defect. The most popular fault model is the stuck-at fault model that assumes that all circuit failures occur in such a way that a logic gate node (input/output) behaves as if it is clamped to logic 1 or 0. Since stuck-at fault model is not accurate for many actual defect

behaviors [29], other fault models like pseudo stuck-at, bridging, open, transistor stuck-on/off, etc. have been used [28], [30]. To model delay faults (which become more important as chip frequencies increase) path delay, gate delay and transition delay fault models have been suggested [31]. These models have their own merit and are successful in varying degrees, however, the single stuck-at fault model continues to be the most popular fault model owing to its simplicity and effectiveness [32] and investment in tools [33].

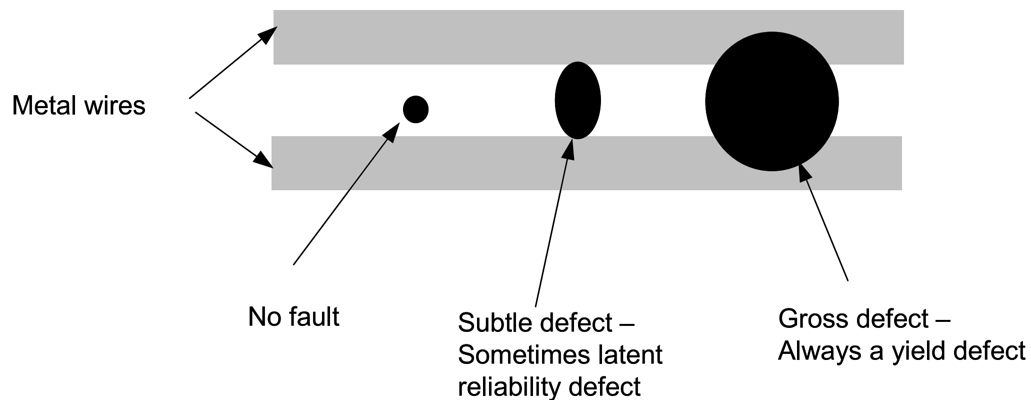


**Fig. 11. Subtle and gross defects.**

Once a chip is manufactured, there is no direct access to its internal nodes. Hence, for fault detection, a test pattern must excite the fault as well propagate its effect to an output. These conditions, called *excitation* (or *sensitization*) and *propagation* of the fault, form the basis of test generation [34], [35]. All *automatic test pattern generator* (ATPG) tools essentially try to generate a test vector (or a pair of test vectors) that excites a fault and makes its presence detectable at an output [36], [37].

## 2.5 Test Quality

Tests themselves are not perfect and cannot catch all defects. Since test cost increases roughly by an order of magnitude with each integration level [10], there is a strong motivation to screen defective chips at the wafer-level before they are packaged. This becomes increasingly important as packaging costs escalate. Finding a defective chip at the customer site is expensive as it can result in a customer return. Moreover, it damages the manufacturer’s reputation about quality. Today’s microprocessors demand quality levels on the order of 100 DPM [38]. To achieve these

quality targets given tests are not perfect, manufacturers need to use a combination of different test methods.

## 2.6 Test Method Classification

A simple method to test a chip is to verify that it performs its intended operation. This is what is called a *functional* test. Obviously, to guarantee a chip is indeed fault-free one must exercise all possible input conditions and verify outputs for correctness. As the number of inputs increases, this results in an exponential growth of test patterns (test vectors) and long test time. Therefore, such *exhaustive testing* is prohibitively expensive and impractical. Moreover, generation of functional test vectors requires thorough understanding of the design and hence long development time. For these reasons, complete functional testing is impractical. The alternative approaches use knowledge of circuit structure for testing. One way to achieve this is to ensure that all internal circuit nodes can assume both binary values and propagate them to (at least) an output. This so-called *structural* test gives some confidence about the integrity of a chip. To ensure ease of structural test requires certain design modifications that enhance controllability and observability of a circuit's internal nodes [39]. These modifications need to have little, if any, impact on the design. Such design modifications are termed *design for testability* (DFT) [40] and are a topic of great importance for VLSI chips [41][42].

Another class of test methods is *parametric tests* which measures various parameters like supply current or chip speed. $F_{max}$ test is a type of parametric test. $F_{max}$ is the maximum frequency a chip is capable of operating at. Microprocessor companies often speed bin their products to optimize revenues (a faster chip has a higher price). Die-to-die and within-die fluctuations in IC parameters significantly impact the result of the $F_{max}$ test [43].

A popular parametric test method is to plot two test parameters and measure their variation across wafers/lots. Such plots, called *shmoo plots*, represent how a particular test passes or fails when parameters like frequency, voltage, or temperature are varied and the test is executed repeatedly [44]. It is a method to visualize how the performance of an IC changes with changes in external environment like temperature or voltage, as well as process parameters. A shmoo plot for power supply and chip speed for an Intel Pentium® 4 processor is shown in Fig. 12. It shows that at a lower supply voltage chips operate at a lower frequency.

The impracticality of functional test and inadequacy of structural test to detect all defects led to *defect based test* (DBT) [45]. In this type of test, a chip is tested for the presence of defects by

comparing the chip's behavior if a defect were present to its defect-free behavior. Examples of this type of test include stuck-at test, pseudo stuck-at ($I_{DDQ}$) test, bridge fault test, etc. It is observed that the addition of DBT is useful for achieving high quality goals [46].

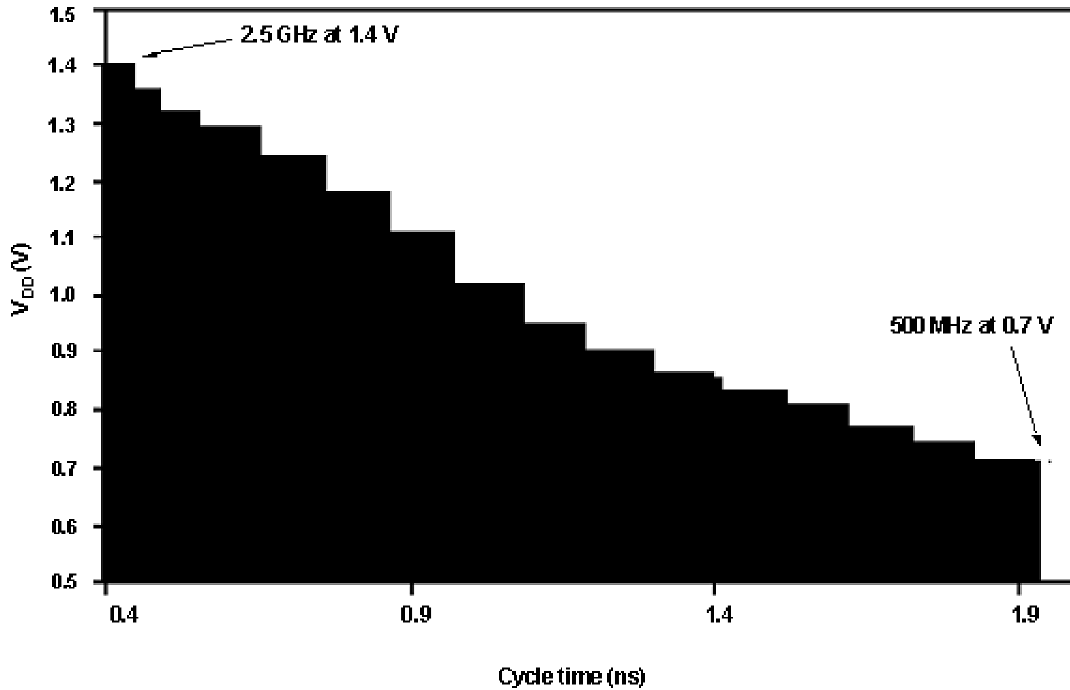


**Fig. 12. Shmoo plot showing relationship between supply voltage and chip speed.**

## 2.7 Principle of $I_{DDQ}$ Testing

When the inputs of CMOS circuits are static, there is no direct conduction path from $V_{DD}$ to $V_{SS}$ (ground/GND). Thus, ideally no current flows through the circuit. Fig. 13(a) shows a static CMOS inverter circuit having PMOS and NMOS transistors in series. When the input is at logic zero (one) the PMOS (NMOS) transistor is ON, NMOS (PMOS) transistor is OFF and the output is at logic one (zero). Ideally, current flows from $V_{DD}$ to GND only for a brief period of time during switching when both transistors are partially ON as shown in Fig. 13(b). However, in practice, a small amount of leakage or quiescent current ($I_{DDQ}$) flows even when inputs are stable. In the presence of a defect, (e.g. source-to-drain short as shown in Fig. 13(a)), the current flows through the direct path from $V_{DD}$ to GND comprised of the defect (short) and ON transistor. This current is much higher than the leakage current as shown in Fig. 13(b). Thus by

measuring the leakage current flowing through the chip for different input vectors, defective circuits can be identified. This observation has been used for screening defective chips and a corresponding fault model is proposed [47][48]. Note that $I_{DDQ}$ is large only when the defect shown in Fig. 13(a) is excited, when the input is high (i.e. PMOS transistor is OFF). Such a defect is called a pattern-dependent or *active* defect. If the defect current is independent of the input pattern, it is referred to as a pattern-independent or *passive* defect. An example of a passive defect is a bridge between $V_{DD}$ and GND [49].

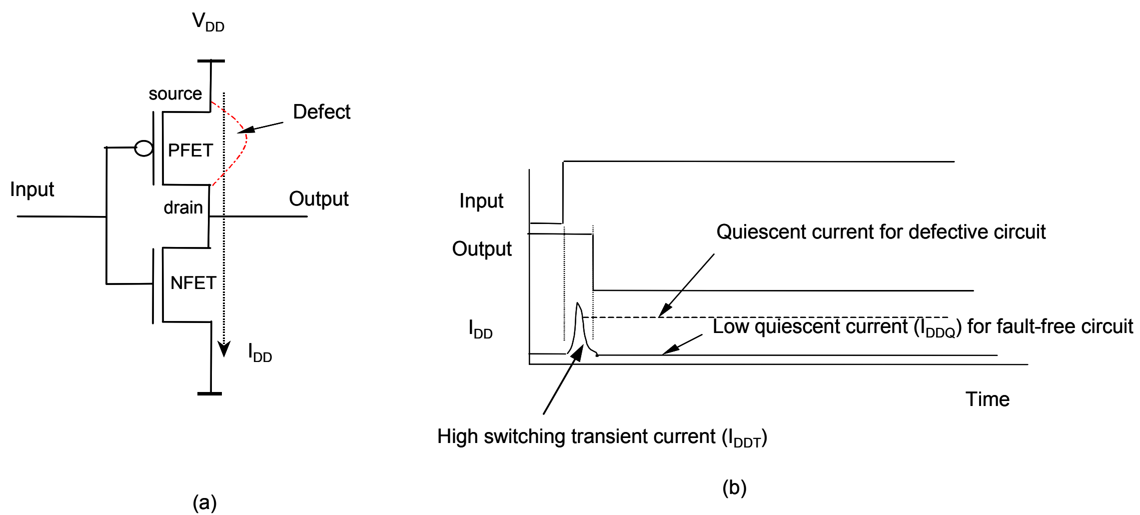


**Fig. 13. (a) CMOS inverter circuit and (b) faulty and fault-free waveforms.**

## 2.8 $I_{DDQ}$ Testing Benefits

$I_{DDQ}$ testing is attractive to semiconductor manufacturers due to its simplicity as well as many other benefits it offers [50]. Since power supply current can always be observed, $I_{DDQ}$ test does not have fault propagation requirements like stuck-at test. This is equivalent to saying that $I_{DDQ}$ test has 100% observability. Several studies have shown that $I_{DDQ}$ test detects certain defects not detected by any other test method [51–53] thus improving outgoing quality [54]. Examples of these defects include gate-oxide short (GOS), bridging defects, some stuck-open faults and certain delay faults [55]. As achieving very high stuck-at fault coverage becomes difficult, $I_{DDQ}$ testing has been used to improve coverage with non-observable test locations and relatively few measurements [56].

Some defects do not cause functional failure, but can result in low reliability of the chip. Such chips fail after they are put in the system. Reduction in the number of such failures (called *infant mortality*) can be achieved by subjecting chips to stress conditions before they are shipped. Semiconductor manufacturers have employed *burn-in* (BI) test[2] in which chips are subjected to higher temperature and voltage. The Arrhenius equation is used to normalize failure rate predictions at a system operation temperature (voltage) [57]. This model assumes the degradation of the performance parameter is linear with time. The temperature dependence is taken to be an exponential function that defines the probability of occurrence. The acceleration of defects due to increased temperature ($A_T$) and voltage ($A_V$) are given by the Arrhenius equation as follows [58]:

$$A_T = e^{\frac{E_a}{k}\left(\frac{1}{T_{use}} - \frac{1}{T_{stress}}\right)} \quad (2.1)$$

$$A_V = e^{\frac{E_a}{k}\left(\frac{1}{V_{use}} - \frac{1}{V_{stress}}\right)} \quad (2.2)$$

where $T_{use}$ ($V_{use}$) is the normal operating temperature (voltage), $T_{stress}$ ($V_{stress}$) is the stress temperature (voltage), $E_a$ is the activation energy, and $k$ is Boltzmann's constant ($1.38 \times 10^{-23}$ J/K) . Due to reduced operating voltage in newer technologies, the maximum ΔV ($V_{stress} - V_{use}$) that can be achieved without destroying the part is reduced. In addition, increased power levels cause high operating temperature, thus reducing the ΔT ($T_{stress} - T_{use}$) that can be obtained. This makes BI less effective for advanced process technologies or requires economically prohibitive BI time. Below 90 nm, BI may become economically infeasible. Moreover, it is a destructive test [59] and components failing BI represent lost revenues. The motivation for reduction in DPM due to BI is becoming outweighed by unacceptable (apparent) yield loss. Thermal runaway due

---

[2] BI is used for high-performance and high-reliability chips and is not part of the normal test flow.

to inadequate temperature control can destroy some fast chips. More cost-effective alternatives to BI are sought. It has been shown that $I_{DDQ}$ testing is useful for detecting many defects that lead to reliability hazards [60–64]. These include gate-oxide shorts, punchthrough and leaky transistors [65]. An independent study from SEMATECH showed that the chips having higher leakage have higher BI failure rate [66]. As $I_{DDQ}$ testing takes a fraction of the time it would take for BI, assembly and test cycle time can be reduced.

## 2.9 $I_{DDQ}$ Test in Practice

$I_{DDQ}$ test essentially involves setting a pass/fail threshold value for leakage current. A chip whose leakage exceeds the threshold is considered defective. The maximum $I_{DDQ}$ is estimated by circuit simulations [67–70] or by developing analytical models [71]. The pass/fail threshold value can be decided by simulation [72–74] or empirical analysis of data [75]. Transistor-level simulation of the circuit is performed for deciding the threshold through simulation. To account for vector-to-vector variation in $I_{DDQ}$ that results from which paths are turned ON/OFF, simulation is carried out for different vectors and maximum $I_{DDQ}$ is estimated. Empirical analysis involves plotting histograms of $I_{DDQ}$ for chips from different wafers and lots. Faulty chips having high leakage current appear in the tail of the distribution. Typically, the threshold is decided by using a value a few standard deviations above the mean. Characterization data is used to estimate maximum fault-free leakage. A sample of chips may be examined to ensure outgoing defect level is within the acceptable limit. Research indicates that a majority of chips that fail only $I_{DDQ}$ test contain defects that go undetected by other test methods [76].

The parametric measurement unit (PMU) is used to measure $I_{DDQ}$. However, it is slow and often inflexible. Other methods for high-speed $I_{DDQ}$ measurement include use of load board sensors [77], QuiC-Mon [78][79], Built-In Current Sensors (BICS) [80] or other proprietary methods.

## 2.10 Components of $I_{DDQ}$

Fig. 14 shows various components of the leakage current for DSM transistors as suggested by Keshavarzi et al. [81]. PN junction reverse bias current ($I_1$) is the result of minority carrier diffusion/drift near the edge of the depletion region and due to electron-hole pair generation in the depletion region of the reverse bias junction. Weak inversion or sub-threshold leakage ($I_2$) occurs when the gate voltage is below the threshold voltage and is the result of carrier movement along the channel surface. Drain-Induced Barrier Lowering or DIBL ($I_3$) is due to the interaction

between the depletion region of the drain with the source that results in lower source potential barrier. Gate-Induced Drain Leakage or GIDL ($I_4$) is the result of high electric field under the gate/drain overlap region that causes deep depletion and results in thin depletion width of the drain-to-well PN junction. Punchthrough ($I_5$) occurs when the drain and source depletion regions electrically "touch" deep in the channel. This is a space-charge condition that causes the gate to lose control of the channel region. The narrow width effect ($I_6$) is seen for transistor geometric gate widths of the order of ≤ 0.5 μm. Gate oxide tunneling ($I_7$) is the result of Fowler-Nordheim (FN) tunneling through the oxide. This becomes an important issue for very thin gate oxides. Hot carrier injection ($I_8$) is due to injection of hot carriers into the oxide. This increases in amplitude as $L_{eff}$ is reduced unless $V_{DD}$ is scaled accordingly.

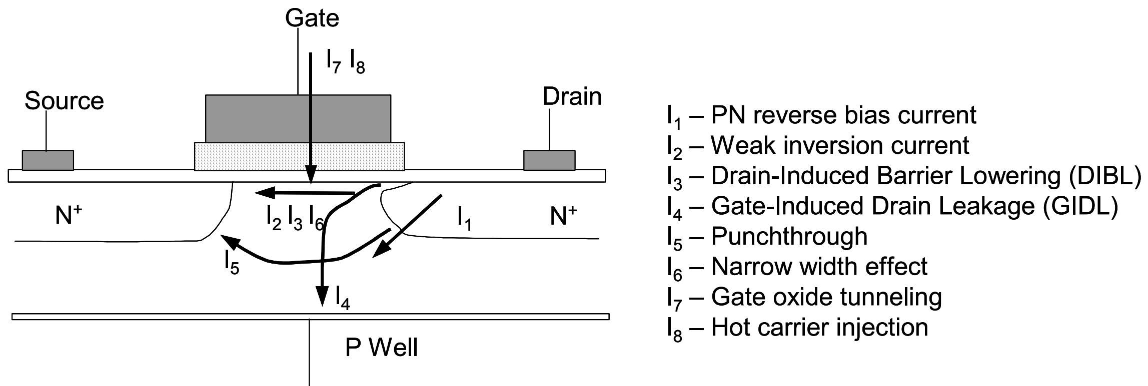


**Fig. 14. Leakage current mechanisms for DSM transistors.**

The dominating leakage current components for advanced bulk CMOS technologies include DIBL and sub-threshold weak inversion current. GIDL dominates for elevated voltages during BI. For higher reliability and lower power consumption, it is necessary to keep leakage current as low as possible. Leakage reduction, as will be explained later, helps $I_{DDQ}$ test to a certain extent, but does not completely solve the outlier identification issue.

## 2.11 Impact of Technology on $I_{DDQ}$ Test

As transistor geometries are reduced, it is necessary to reduce the supply voltage to avoid electrical breakdown of the gate oxide. However, to retain or improve the performance it is necessary to reduce the threshold voltage ($V_{TH}$) as well to maintain the gate overdrive. The sub-

threshold leakage current is given by

$$I_{sub} = \mu.C_{ox}.\frac{W}{L}V_t^2 e^{V_{GS}-V_{TH}/\eta V_t}.(1-e^{-V_{DS}/V_t}) \qquad (2.3)$$

where $\mu$ is the carrier mobility, $C_{OX}$ is the gate capacitance per unit area, $W$ is the channel width, $L$ is the channel length, $V_{GS}$ is the gate-to-source voltage, $V_{TH}$ is the threshold voltage and $\eta$ is the technology dependent parameter [82]. The thermal voltage $V_t$ is given by $kT/q$ where $k$ is Boltzmann's constant, T is the absolute temperature and $q$ is the electron charge ($1.6 \times 10^{-19}$ C). Thus, the reduction in threshold voltage ($V_{TH}$) causes an exponential increase in the sub-threshold leakage current [83] as illustrated in Fig. 15.

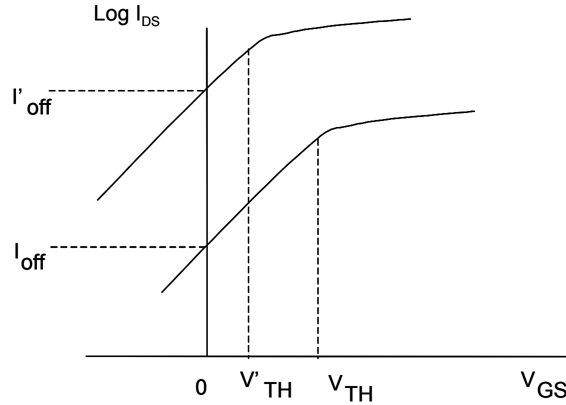


**Fig. 15. Reduction in threshold voltage increases leakage current exponentially.**

DSM technology chips are characterized by what is known as the *short channel effect* (SCE) [84]. As the distance between source and drain is reduced, the gate begins to lose control over the charge in the channel. Thus, even with no bias a large number of charge carriers are able to cross the channel resulting in appreciable leakage current. The International Technology Roadmap for Semiconductors (ITRS) projections for $I_{DDQ}$ of high-performance microprocessor circuits shown in Table I indicate that this trend will continue in the future due to pronounced SCE [2][85][86].

**TABLE I. ITRS PROJECTIONS FOR $I_{DDQ}$ OF HIGH-PERFORMANCE ICS.**

| Year | Maximum $I_{DDQ}$ |
|---|---|
| 2001 | 30-70 mA |
| 2003 | 70-150 mA |
| 2005 | 150- 400 mA |
| 2008 | 400 mA-1.6 A |
| 2011 | 1.6 - 8 A |
| 2014 | 8-20 A |

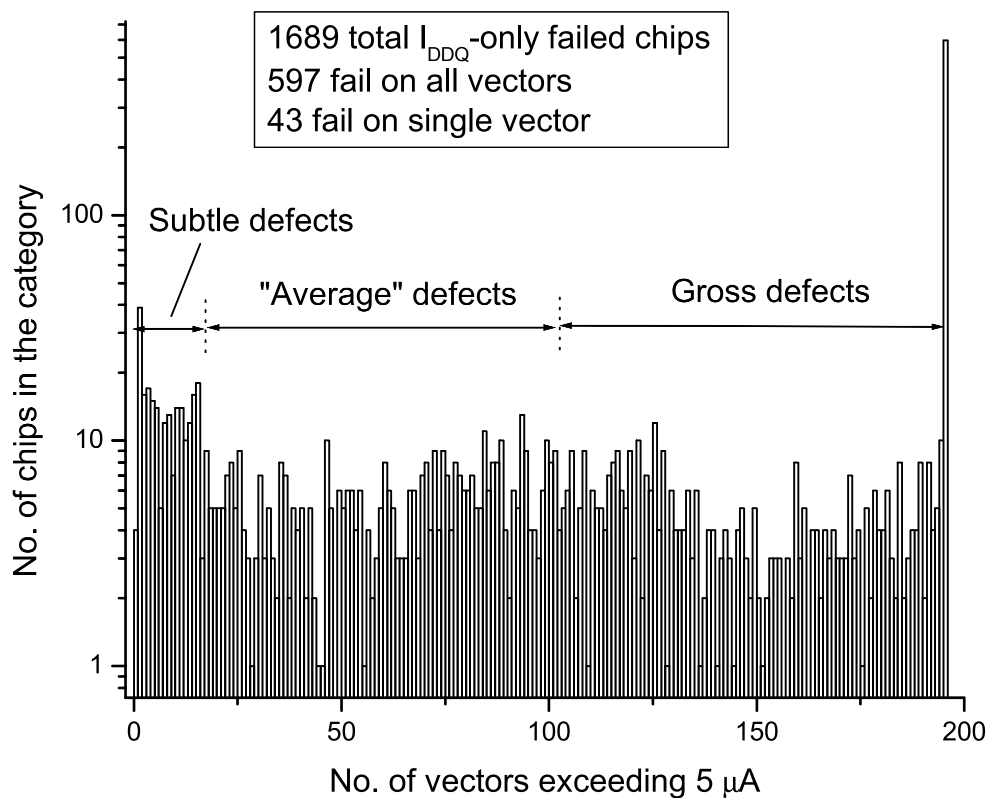


**Fig. 16. Defect detection using $I_{DDQ}$ test depends on defect severity.**

For smaller transistors, it is difficult to control geometries precisely. This results in increased relative process variations both within a die (intra-die) as well as between dice (inter-die). The threshold voltage is dependent on the effective channel length ($L_{eff}$). A small variation in $L_{eff}$ of a

transistor can therefore result in large variation in $I_{DDQ}$. The maximum defective $I_{DDQ}$ depends on the nature and severity of a defect. This eventually decides how many vectors are capable of detecting the defect. This is illustrated in Fig. 16 using $I_{DDQ}$ failed chips from IBM/SEMATECH data. This experiment used a pass/fail threshold of 5 μA and a total of 195 measurements per chip. Fig. 16 shows the number of chips and number of failing vectors. Thus, 43 chips had above threshold leakage current for a single vector and 597 chips failed on all 195 vectors. There is a continuum in between showing increasing severity of the defect. Of course, this analysis assumes that all vectors are equally likely to excite a defect, if it exists.

In earlier technologies, the distributions of $I_{DDQ}$ for fault-free and faulty chips were assumed to be separate as shown in Fig. 17(a). Thus, a single threshold value was capable of differentiating fault-free and faulty chips. With increasing magnitude and variation in the fault-free leakage, the two distributions begin to overlap as shown in Fig. 17(b). Hence, any single threshold results in yield loss and/or test escapes. Increasing spread in leakage as projected by the ITRS (Table I) suggests that the overlap will increase for future technologies [2][87]. It must be emphasized here that ITRS projections are often superceded by advances in technology. The leading edge microprocessors like Intel Pentium® 4 already have leakage of the order of several amperes in year 2003.

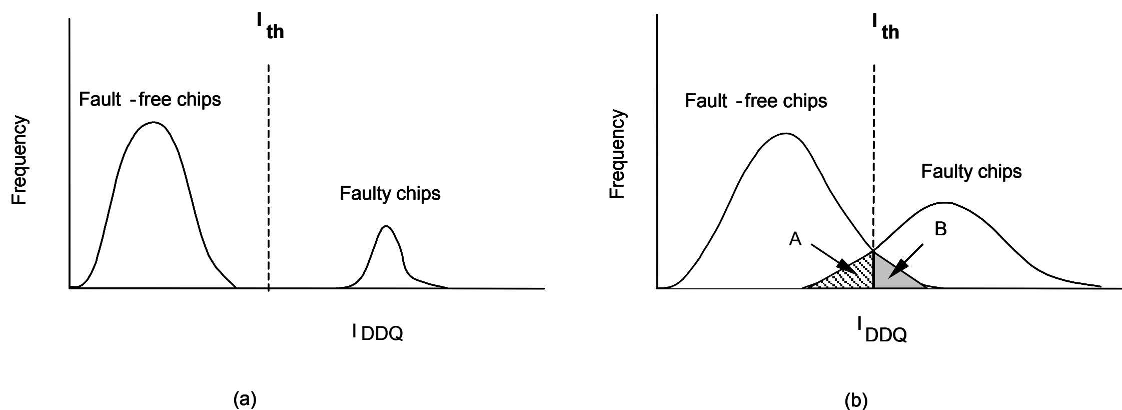


**Fig. 17. (a) Single threshold test for older and (b) for DSM technologies.**

Not only leakage current magnitude is increasing with each technology node, but there is increased variation in the leakage as well [88]. Fig. 18 shows variation in $I_{DDQ}$ for a vector across five wafers. It is obvious that setting a single pass/fail $I_{DDQ}$ threshold will reject several chips.

Thus, single threshold $I_{DDQ}$ test is obsolete for high-performance technologies. It becomes necessary to isolate systematic and random variations in order to identify defective chips. The next section reviews several methods that have been reported in the literature to solve this issue.

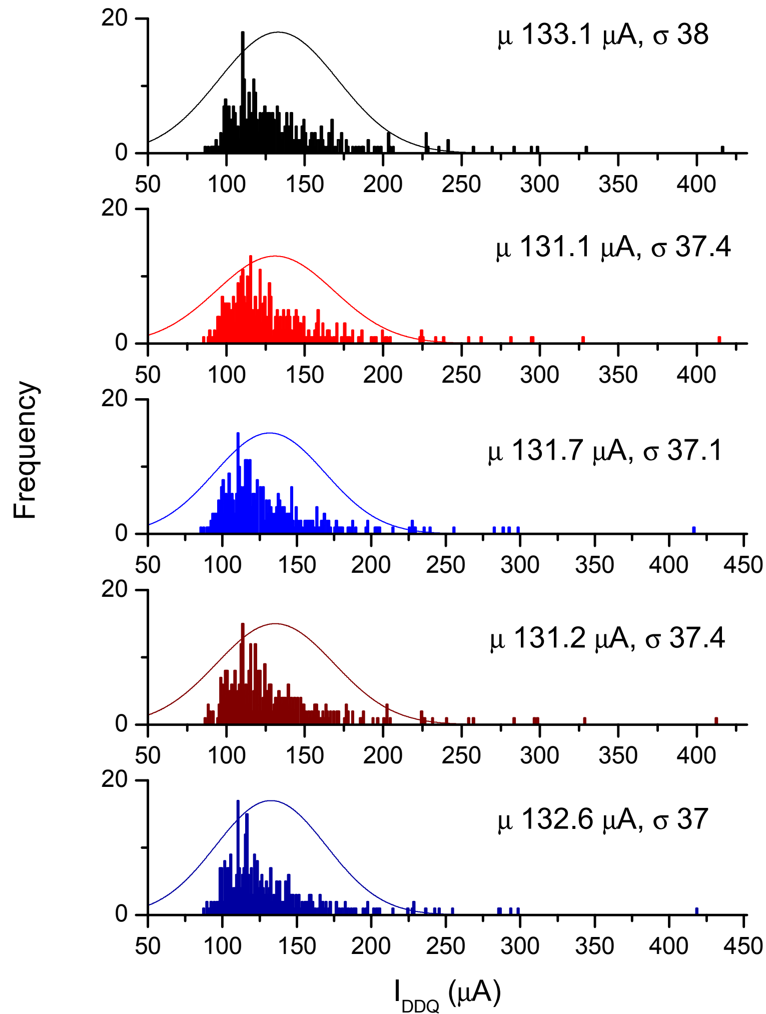


**Fig. 18. Chip-to-chip $I_{DDQ}$ variation for a vector for five wafers.**

# 3. REVIEW OF $I_{DDX}$ –BASED TEST METHODS

## 3.1 Introduction

Due to its several unique benefits mentioned in Section 2, $I_{DDQ}$ test forms an important component of a test suite [89]. However, there has been a growing concern about the applicability of $I_{DDQ}$ test to future technologies [90] due to unacceptable yield loss it incurs [91]. While quality is undoubtedly important, yield is perhaps the most sensitive factor for semiconductor manufacturers as it translates directly into revenue. For a yield manager, rejecting chips simply because they have increased leakage may not be justifiable especially if the percentage of such chips is high.

To retain the effectiveness of $I_{DDQ}$ test in production without causing much yield loss, several solutions have been proposed in the literature. Some test methods use transient current ($I_{DDT}$) measurement instead of quiescent leakage ($I_{DDQ}$) [92]. The goal of both classes, collectively known as $I_{DDX}$ test methods, is to determine whether a chip is faulty or fault-free through power supply current measurement. This section provides a review of these test methods. It should be noted, however, that the main goal of the dissertation is to distinguish faulty and fault-free chips using *any* parametric test data; $I_{DDQ}$ test data is used only as an example.

The problem with $I_{DDQ}$ test is straightforward: faulty and fault-free currents are indistinguishable due to increased magnitude and variance in fault-free $I_{DDQ}$. Fundamentally, there are two ways to solve the problem:

1. Reduce fault-free $I_{DDQ}$ magnitude and/or variance itself, or
2. Use data analysis methods to distinguish faulty $I_{DDQ}$ from fault-free $I_{DDQ}$.

The first class of solutions invariably involves modification in the design or processing steps. To be cost effective, such approaches need to provide a long-term solution (in terms of technology nodes). The second class of solutions attempts to find patterns in the data that make fault-free and faulty chips distinguishable. Many data variance-reduction techniques use wafer level data for analysis. Although outlier detection is not feasible in real-time [93], the data analysis can be completed either during shadow time when another wafer is being loaded on the tester or before wafers are diced so that the investment in packaging can be reduced. A variation of the second class of methods involves the use of the dependence of $I_{DDQ}$ on other test parameters like temperature or exploitation of correlation between $I_{DDQ}$ and clock frequency. In the remainder of the section, we will provide a brief review of some these methods. Some of the

methods will be illustrated by examples using empirical data. In the later part of the section, empirical data will be used to motivate the use of outlier detection.

### 3.2 $I_{DDQ}$ Reduction Techniques

There are two ways of reducing leakage itself: either by changing the design or technology parameters, or by changing the process technology itself. In general, both these methods require significant investment and planning. A review of sub-threshold leakage modeling and reduction techniques can be found in [94] and [95].

### 3.3 Reverse Body Biasing (RBB) Technique

In conventional CMOS circuits, the P-substrate is connected to $V_{SS}$ (ground) and N-wells are connected to $V_{DD}$. For an N-channel transistor, reverse body bias results when the substrate gets a negative voltage with respect to the source. The same effect is achieved for a P-channel transistor by applying positive voltage to the N-well. Reverse (back) biasing results in $V_{TH}$ modulation as described by the following equation [83]:

$$\Delta V_{TH} = V_{TH}(V_{BS}) - V_{TH}(V_{BS}=0) \tag{3.1}$$

$$= K\left\{\sqrt{2\Psi_B + V_{BS}} - \sqrt{2\Psi_B}\right\} \tag{3.2}$$

where K is the body effect coefficient, $\psi_B$ is the potential difference between actual and intrinsic fermi level for a given process and $V_{BS}$ is the substrate (bulk) to source voltage. Typical values for a sub-micron process are a K of 0.59 $V^{1/2}$ and $2\psi_B$ approximately 0.8 V [96]. An application of 1.2V back bias increases $V_{TH}$ by 310 mV, thus reducing sub-threshold leakage by approximately four orders of magnitude. The net effect of RBB is shifting the $I_{DS}$-$V_{GS}$ curve towards the right as shown in Fig. 16 in Section 2. For a 0.35-micron process, a reverse bias of 2 to 3V resulted in 2500x to 4400x reduction in leakage current [97].

Although RBB is a powerful technique, there are primarily two issues associated with it. First, it requires significant change in the cell library. Logic gates need to be redesigned such that separate N-well and substrate connections are routed in the same fashion as for $V_{DD}$ and $V_{SS}$. Extra supply lines can result in increased chip area, estimated to be 3 to 8% [83]. Secondly, the effectiveness of RBB diminishes for technologies below 0.25 microns. The body effect coefficient is not constant, but reduces with each technology node. GIDL also reduces the benefit

obtained by RBB. Thus, with each technology node RBB becomes less effective approximately by an order of magnitude in reducing the leakage current [97], [98].

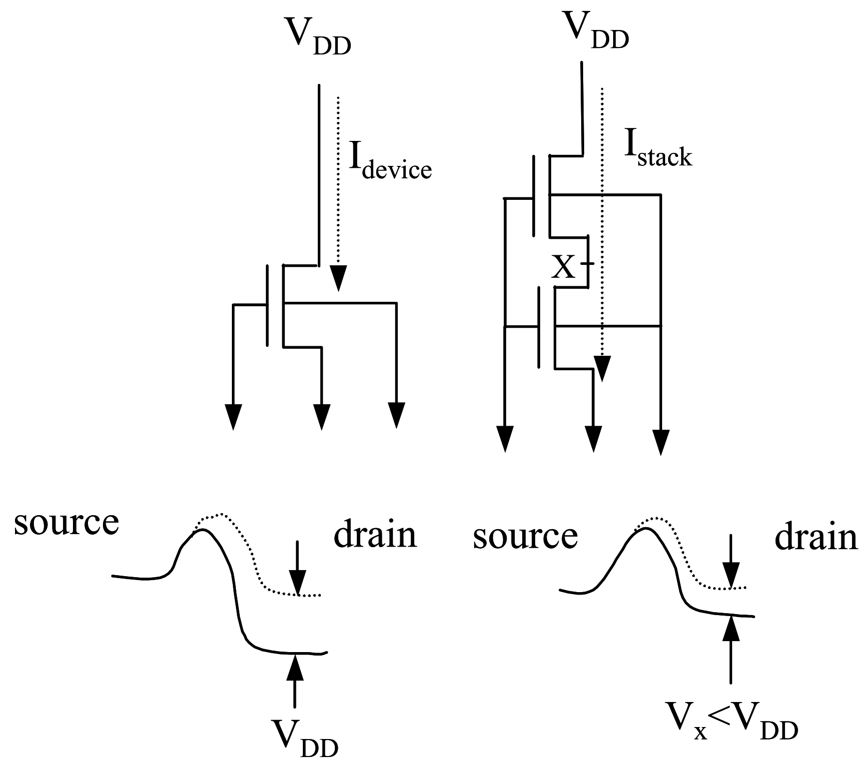


**Fig. 19. Leakage current reduction by stacking of transistors.**

3.4 Stack Effect

The leakage current depends on the potential difference between the source and the drain and can be reduced by lowering the potential difference. An alternative to reduce leakage is to use transistor stacks within the logic gates [99], [100]. As shown in Fig. 19 two series-connected OFF transistors have lower leakage compared to a single OFF device due to the self-reverse biasing effect. The energy band diagram in Fig. 19 shows that the barrier height is modulated to be higher for the two-stack due to smaller drain-to-source voltage resulting in reduced leakage. By selecting an appropriate input vector the leakage current can be reduced by maximizing the number of “off” transistors connected in series [101]. GIDL limits the potential benefit of stacking.

## 3.5 Multiple-$V_{TH}$ CMOS Transistors

Another alternative to reduce leakage is to trade speed for leakage current. By incorporating transistors with low threshold voltage in only critical paths, leakage current can be substantially reduced. Today's performance-optimized chips contain many critical paths and a large fraction of paths needs to be implemented with low-$V_{TH}$ devices. This reduces the effectiveness of this methodology. CAD tools that can be integrated into the design flow for design partitioning are required.

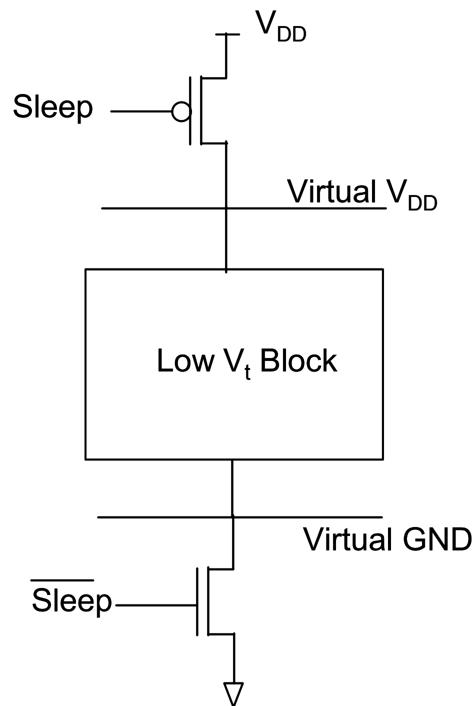


**Fig. 20. Basic principle of MTCMOS circuits showing sleep control transistors.**

MTCMOS (Multi-threshold CMOS) is a dual-$V_{TH}$ partitioning technique [102]. The leakage can be reduced by inserting a high-$V_{TH}$ leakage control transistor between the power supply and the pull-up network *or* between ground and the pull-down network. These leakage control transistors (called *sleep transistors* or *power gates*) are turned ON/OFF in such a way that no leakage path passes through more than one transistor which is turned OFF [103][104] (see Fig. 20). For benchmark circuits a 2x–5x reduction in leakage was observed with minimal delay impact [105]. The area overhead for additional wiring and transistors was estimated to be 18.3%.

MTCMOS offers the greatest reduction in leakage (on the order of 1000x) but the price paid for this is an increase in delay from 5% to 64% [105].

## 3.6 Use of Alternative Technology

One possible alternative for reducing $I_{DDQ}$ is to use Silicon-On-Insulator (SOI) technology [106]. In conventional bulk CMOS technology transistors are constructed on a crystalline active semiconductor while in SOI CMOS the devices are placed on an insulator that is grown on silicon as shown in

Fig. 21. Thus, two transistors placed on the same substrate do not have any electrical influence on each other. In case of long channel transistors, the PN-junction leakage (which contributes the maximum in the case of conventional CMOS) is drastically reduced, as no well is needed to separate P-channel and N-channel transistors. The drain leakage for SOI is reported to be 15 to 100 times smaller than bulk CMOS [107]. SOI devices also have better sub-threshold swing (60 mV/decade) than bulk CMOS (75-85 mV/decade). The sub-threshold swing essentially decides the sensitivity of leakage current to gate to source voltage ($V_{GS}$).

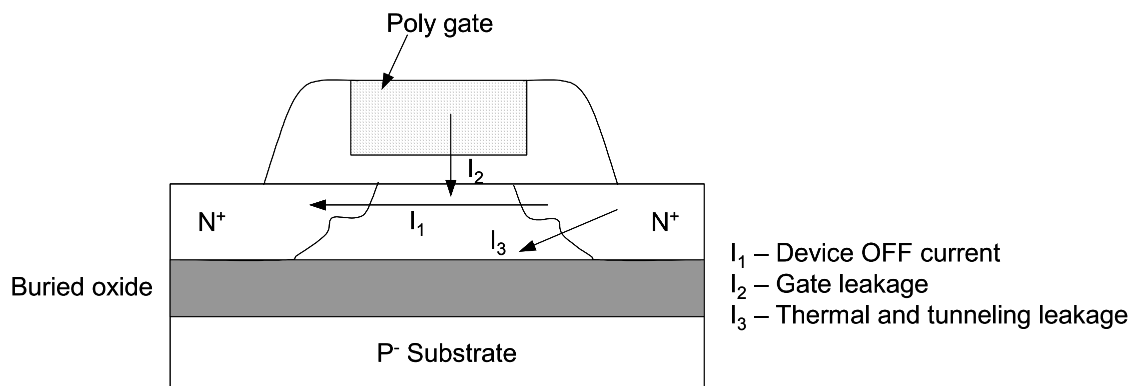


**Fig. 21. Cross section of SOI transistor and leakage components.**

Although SOI offers many advantages over conventional CMOS, it involves a significant change in the process technology and therefore requires huge initial investment. Secondly, SOI offers only a one-time benefit for leakage reduction.

### 3.7 Partitioning of Power Supply Network

Power supply partitioning has been proposed to increase test resolution by partitioning the power supply network, such that each partition has a relatively small fault-free $I_{DDQ}$ level [108], [109]. Every partition has a BICS that measures the current in that single partition. Thus the fault-free leakage current in each partition is low enough to achieve adequate $I_{DDQ}$ test resolution. External partitioning of the power supply is feasible only for the current technology node and that too for low power chips. Analysis in [110] shows that the only feasible long-term test approach would be to combine power supply partitioning with resolution enhancement methods discussed in the next subsection. Design of BICS itself is a difficult challenge [111].

### 3.8 $I_{DDQ}$ Variance Reduction/Data Analysis Techniques

The overall goal of variance-reduction techniques is to post-process parametric test data for screening defective chips. The main advantage of these methods is very little investment. In some cases, Automatic Test Equipment (ATE) needs to be modified or a new tester board may be needed. In other cases, ATE software may need to be reconfigured. Compared to design or process change these changes take much less time, money and are relatively easy to implement. The downside is that the effectiveness of these methods depends on process maturity so that normal process variations can be characterized.

### 3.9 Current Signature

Proposed by Gattiker and Maly [112], the current signature method relies on the graphical display of $I_{DDQ}$ readings sorted in ascending order. It relies on the premise that $I_{DDQ}$ for an active defect is higher (for vectors that excite it) than normal leakage [113]. Thus, the presence of a “step” or “jump” in a signature indicates the presence of at least two distinct leakage paths or an active defect. In the case of a passive defect, this assumption is violated, as defect excitation is independent of the input pattern. Thus, the current signature of a chip with a passive defect does not show any “steps” and can be indistinguishable from the fault-free current signature. If the background leakage is small, to a certain extent the step size is indicative of the severity of the defect. Fig. 22 shows the current signatures for three chips. Chip ‘A’ is fault-free and has a smooth signature (small intra-die variance). Chip ‘B’ has an active defect as indicated by a large step in the signature. Chip ‘C’ shows a smooth signature similar to that of chip ‘A’, but all $I_{DDQ}$ values are higher by two orders of magnitude. This chip is therefore likely to contain a passive defect.

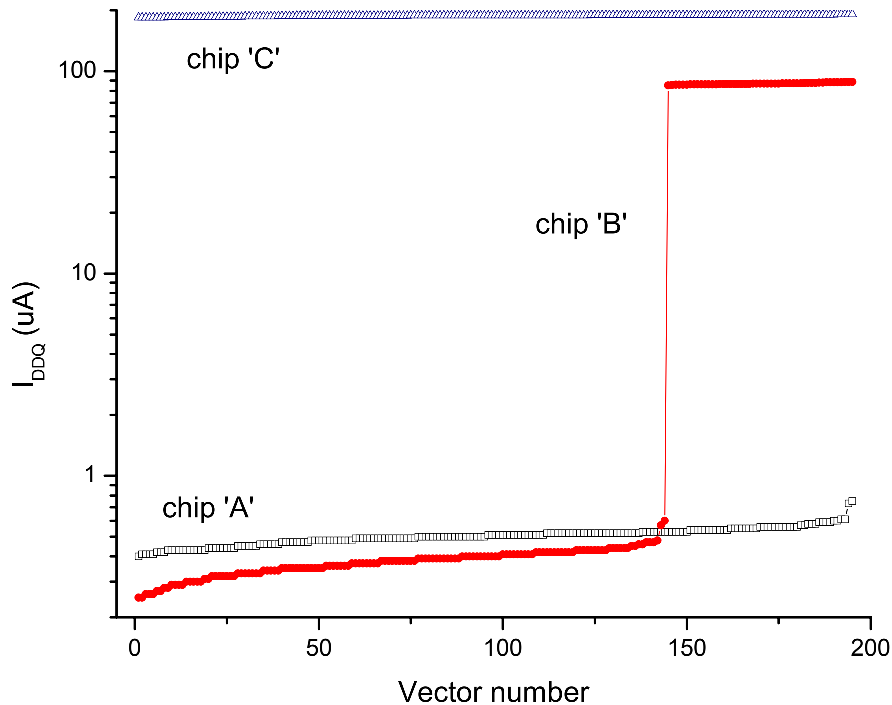


**Fig. 22. Current signatures for fault-free and faulty chips.**

To be effective, current signature requires several $I_{DDQ}$ measurements. Since quiescent leakage can be measured only after internal circuit activity has settled down, $I_{DDQ}$ is a slow-speed test and measuring $I_{DDQ}$ for many vectors may not be practical. An alternative way of sorting for production implementation of a signature-based approach suggests making a measurement and deciding a guard band around it [114]. If any later measurement falls outside this guard band, the chip is rejected.

In spite of its simplicity, current signature is very effective. A lot of information (*circuit personality*) can be gleaned from the analysis of current signatures [115] and is useful for failure diagnosis [116]. Two practical issues, however, must be dealt with. The first one is deciding how many measurements are enough. The second issue is deciding the maximum fault-free step size. Both these issues directly impact test time (and hence, test cost), yield loss and test escapes (quality).

## 3.10 Delta $I_{DDQ}$

The main goal of the delta $I_{DDQ}$ method is to cancel background leakage to reduce variance in fault-free leakage in order to make defective leakage distinguishable [117]. In this method differences (deltas) between $I_{DDQ}$ values for successive vectors for a chip are obtained [118][119]. Thus, delta $I_{DDQ}$ is defined as

$$\Delta I_{DDQ}(i) = I_{DDQ}(i) - I_{DDQ}(i\text{-}1) \quad (3.3)$$

where $I_{DDQ}(i)$ and $I_{DDQ}(i\text{-}1)$ are $I_{DDQ}$ readings for the $i^{th}$ and $(i\text{-}1)^{th}$ vectors. For a fault-free chip, only intrinsic variation in $I_{DDQ}$ causes the mean delta $I_{DDQ}$ to be close to or equal to zero and the variation in deltas to be small. This method assumes that at least one vector excites the defect and at least one vector does not and the defective $I_{DDQ}$ is much higher than the fault-free $I_{DDQ}$. This method is shown to be superior to the conventional single threshold approach [120]. In case of a passive defect, since all readings are elevated, deltas are small. Hence, this method is unable to screen chips with a passive defect. Fig. 23 illustrates the histograms of delta $I_{DDQ}$ for three chips. All chips passed all Boolean tests. In each case, 194 deltas are obtained by subtracting readings for two consecutive vectors. Fig. 23(a) is a histogram for a fault-free chip that exhibits small mean value and variation. Fig. 23(b) illustrates the histogram for a chip with an active defect. Such a chip typically exhibits large variation in delta $I_{DDQ}$. Fig. 23(c) underscores the difficulty in screening a chip with a passive defect, as the variation in deltas is not that large.

Other variations of the delta $I_{DDQ}$ method that use differences in $I_{DDQ}$ at different temperatures [121], voltages [122], etc. have been proposed. The production implementation of delta-$I_{DDQ}$ may consist of making a measurement and setting a guard band around this value. If any later reading falls outside this guard band, the chip is considered defective and is rejected. The width of the guard band needs to be determined through characterization of fault-free chips.

Although delta $I_{DDQ}$ is intuitively simple and easy to implement, it also suffers from the same issues as the current signature method. The number of measurements limits the defect screening resolution of delta $I_{DDQ}$. Deciding the maximum fault-free delta is not trivial as it requires elaborate vector sensitivity analysis (which paths are turned ON/OFF) as well as process sensitivity analysis (impact of process variations on intrinsic $I_{DDQ}$). Furthermore, the future of delta $I_{DDQ}$ is questionable for DSM technologies [123].

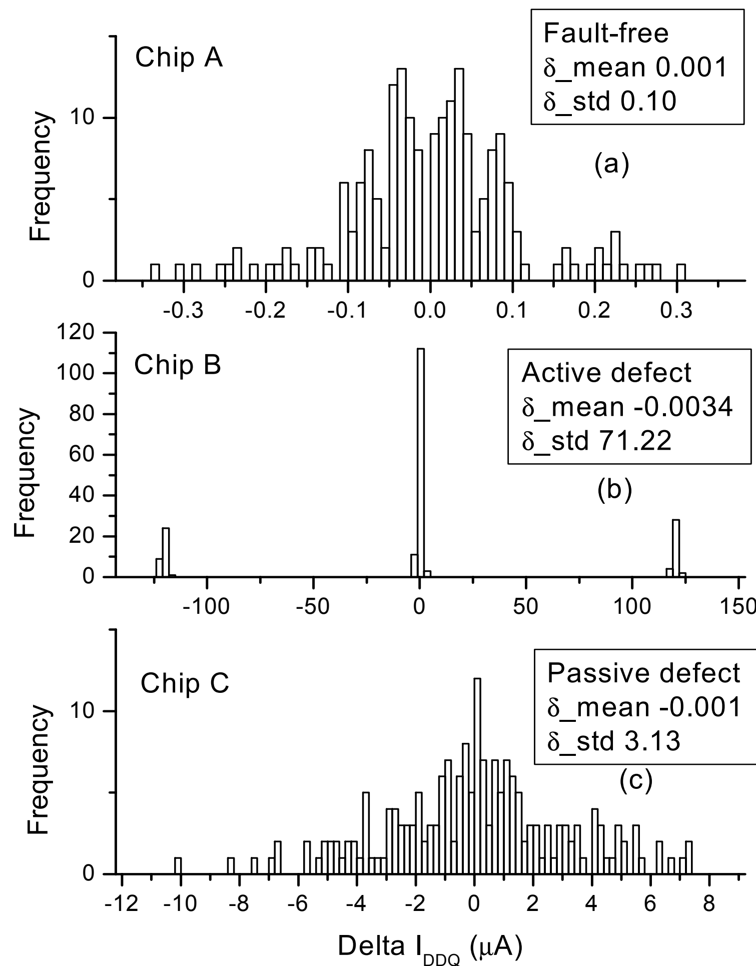


**Fig. 23. Histograms of delta-$I_{DDQ}$ for fault-free and faulty chips.**

## 3.11 Statistical Clustering

Clustering is a statistical procedure of sorting data into groups such that the degree of "natural association" is high among members of the same group and low between members of different groups. It uses correlation or other such measure of association for classifying the data into groups. In a loose sense, it can be considered as multi-dimensional regression. Some experiments on application of clustering techniques to $I_{DDQ}$ testing have been reported [124], [125]. Fig. 24 shows a typical result of clustering. The chips are divided into four clusters. Notice that chips having similar $I_{DDQ}$ can be clustered into different groups. Thus, seemingly fault-free chips can be grouped with defective chips and vice versa.

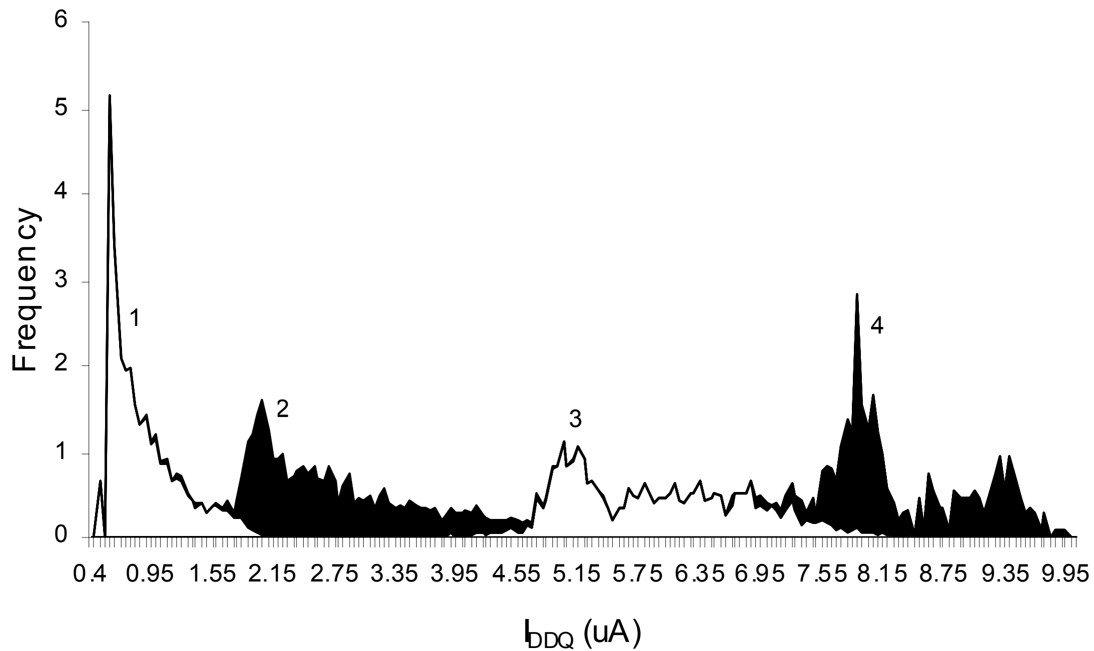


**Fig. 24. Normalized histogram for four clusters of $I_{DDQ}$ test data.**

In principle, clustering can be applied to any continuous parameter. For example, clustering $I_{DDQ}$ data combined with chip speed can be helpful in finding outliers with greater confidence. Due to its nature of grouping elements, clustering inherently accounts for process variations. However, it requires a number of readings to be meaningful. In addition, after data is divided among different clusters, it is up to the user to decide which groups represent faulty and fault-free chips. This is not trivial as the number of clusters increases. If the number of clusters is reduced, distinction between fault-free and faulty chips fades. It has been suggested that clustering can be used to decide appropriate pass/fail limits for each lot/wafer.

### 3.12 Current Ratio

In spite of the increased magnitude and variation in $I_{DDQ}$ in new technologies, it was observed that the ratio of maximum $I_{DDQ}$ to minimum $I_{DDQ}$ for fault-free chips (called *current* ratio) is relatively the same. A leaky chip will leak proportionately more for *all* vectors and therefore its max/min ratio will be comparable to fault-free max/min ratios for a chip with lower

leakage. This is illustrated in Fig. 25 (courtesy Peter Maxwell). It shows sorted $I_{DDQ}$ readings for two dice. One die has almost six times $I_{DDQ}$ of the other. The measurements made on lower current die for the same vector order are shown in gray color. In spite of having different magnitudes they have similar intra-die variance.

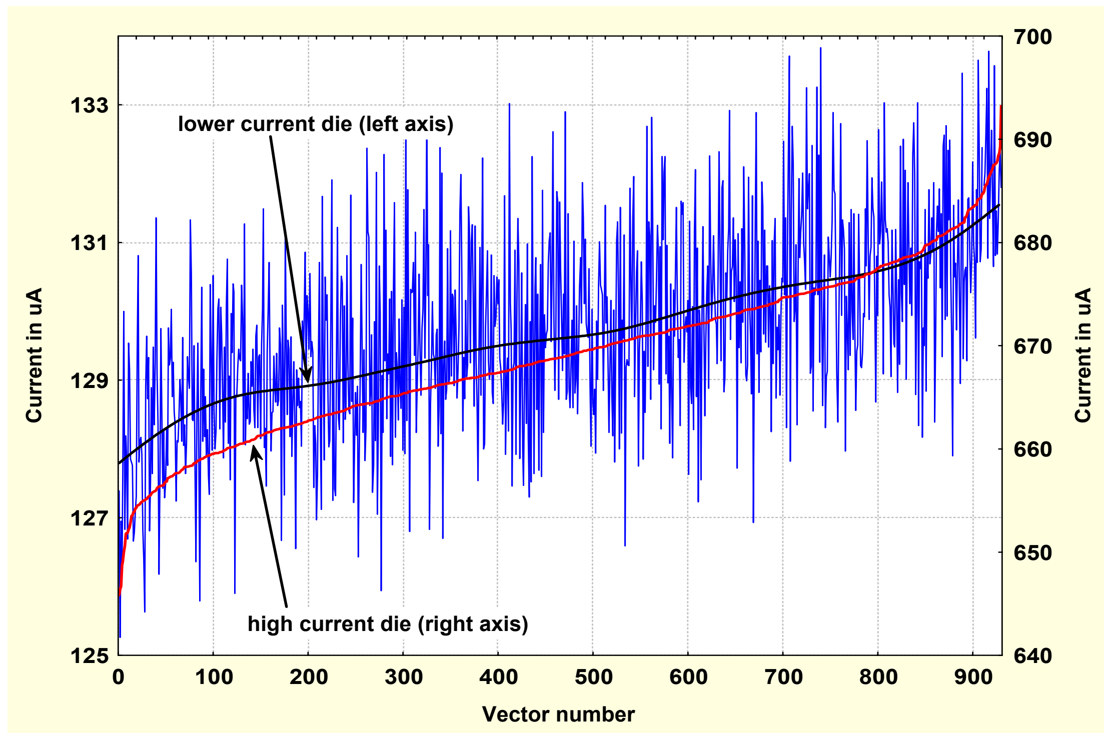


**Fig. 25. Two chips having different $I_{DDQ}$ magnitude but similar CR.**

Through characterization, the input vectors that cause minimum and maximum $I_{DDQ}$ are determined and current ratio (CR) is obtained. To account for process variation a guard band is added. In production, $I_{DDQ}$ is measured for all vectors and current ratio is obtained. If the current ratio exceeds a predetermined threshold, the chip is rejected [126]. Vector-to-vector correlation can be used to improve the performance of this method by selecting a vector pair with the highest correlation [127], [128].

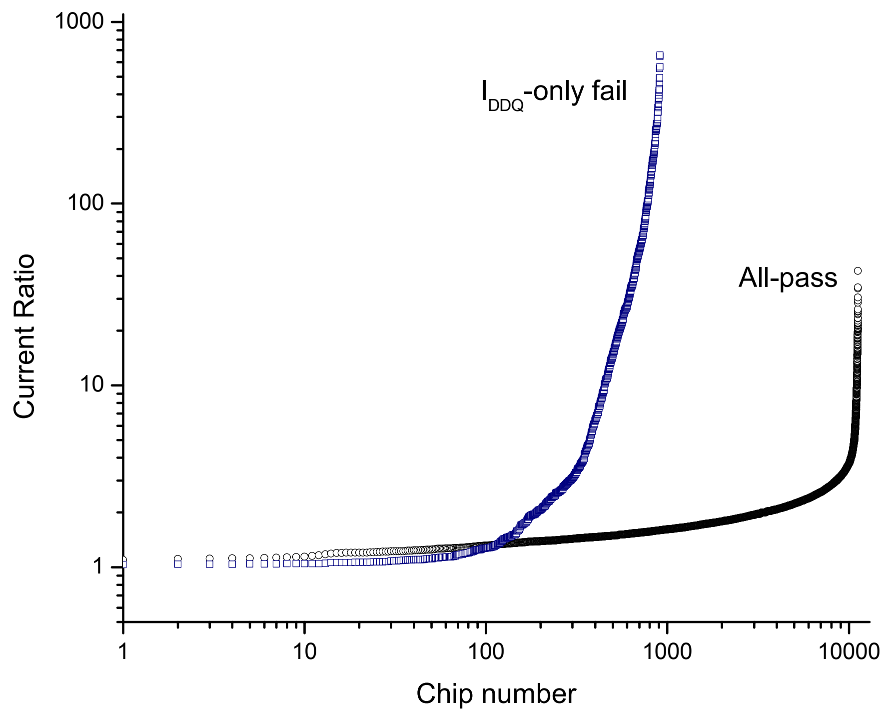


**Fig. 26. Current ratios of fault-free and faulty chips from a wafer.**

Fig. 26 shows current ratios sorted in ascending order for several fault-free (passed all SEMATECH tests) and faulty (failed the 5 μA threshold $I_{DDQ}$ test) chips. Clearly, the faulty chips exhibit more spread in current ratios than the fault-free chips. However, even for fault-free chips current ratios show variation of an order of magnitude. Many defective chips have current ratios comparable to fault-free chip current ratios. Therefore, deciding the current ratio threshold can be challenging. As passive defect current increases, current ratio reduces as shown in Fig. 27. As background current rises, CR reduces from its nominal value 3 and to a value close to 1. This indicates that CR is unable to screen passive defects unless a lower limit is set. As shown in Fig. 26, many $I_{DDQ}$ fail chips have low CRs and can contain a passive defect.

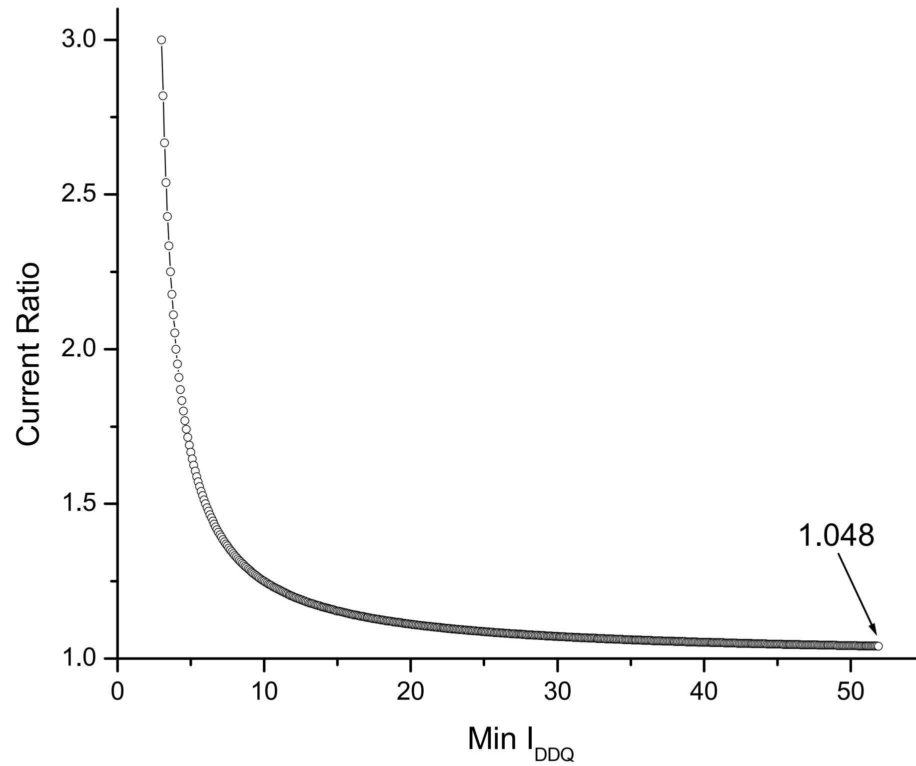


**Fig. 27. Reduction in CR with increasing background leakage.**

3.13 Eigen Signatures

Okuda suggested exploiting regularity of fault-free $I_{DDQ}$ to make it distinguishable from defective current [129][130]. Fig. 28 shows that fault-free chips exhibit regularity in $I_{DDQ}$ while a defective chip shows irregular variation. Okuda observed that for fault-free chips the maximum leakage was proportional to the mean $I_{DDQ}$ value. Because of deterministic variations in fault-free $I_{DDQ}$, the ratio of variance to mean $I_{DDQ}$ was suggested as an *Eigen signature*. However, we observed a large variation in mean and standard deviation of chips for the SEMATECH test data (see Fig. 29). This indicates that the success of an Eigen signature-based approach depends on the manufacturing process. It is also important to have a large number of $I_{DDQ}$ measurements for successful implementation.

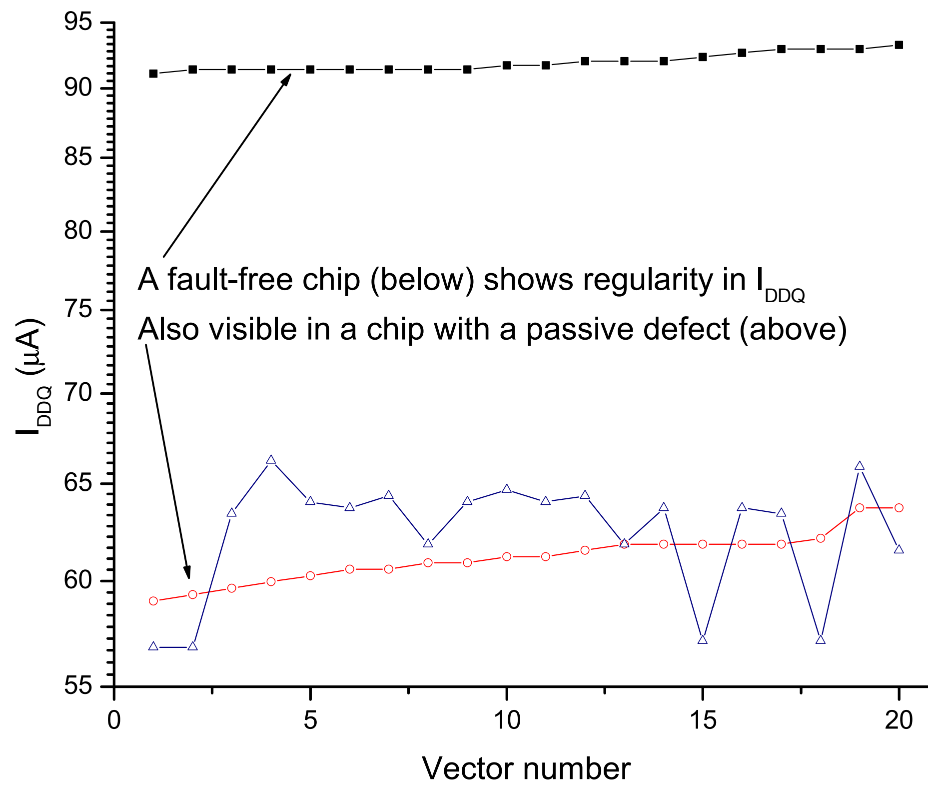


**Fig. 28. Fault-free chips show regular variation in $I_{DDQ}$.**

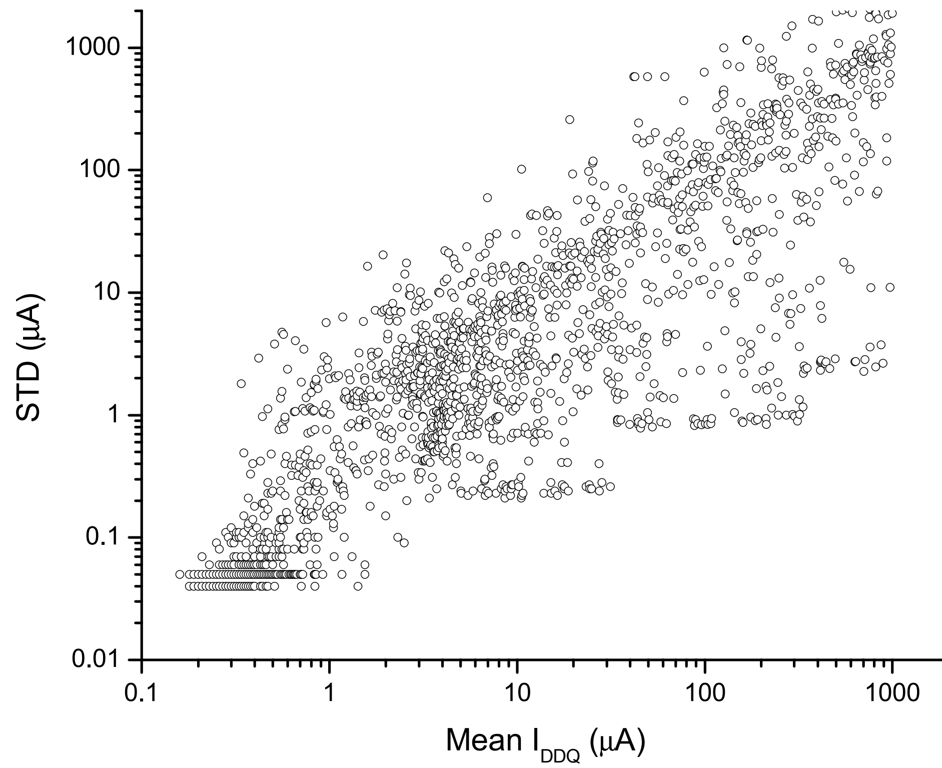


**Fig. 29. Scatter plot of mean and standard deviation of $I_{DDQ}$ for SEMATECH data.**

### 3.14 $I_{DDQ}$ versus Temperature

The leakage current equation in Section 2 suggests that fault-free $I_{DDQ}$ is a function of temperature. Defective current does not show such relationship. The dependence of fault-free leakage on temperature can be exploited by making current measurements at two temperatures [131].

Fig. 30 shows $I_{DDQ}$ values for two temperatures. The fault-free leakage exhibits an exponential relationship with temperature [132]. The defective current may remain the same or decrease (due to positive temperature coefficient for a resistive metal short) with an increase in temperature. This makes differentiation between fault-free and faulty chips possible. Testing can be performed by measuring $I_{DDQ}$ at room temperature and at a reduced temperature. However, such an approach is usually impractical in production due to high cost [133]. Fig. 31 shows a

similar plot for chips that have higher leakage. Note that a majority of the chips lie along the line having slope 1. A major component of this leakage can stem from defect. A leaky fault-free chip would exhibit exponential dependence with temperature. However as this example illustrates the effectiveness of temperature-based test reduces with increasing leakage.

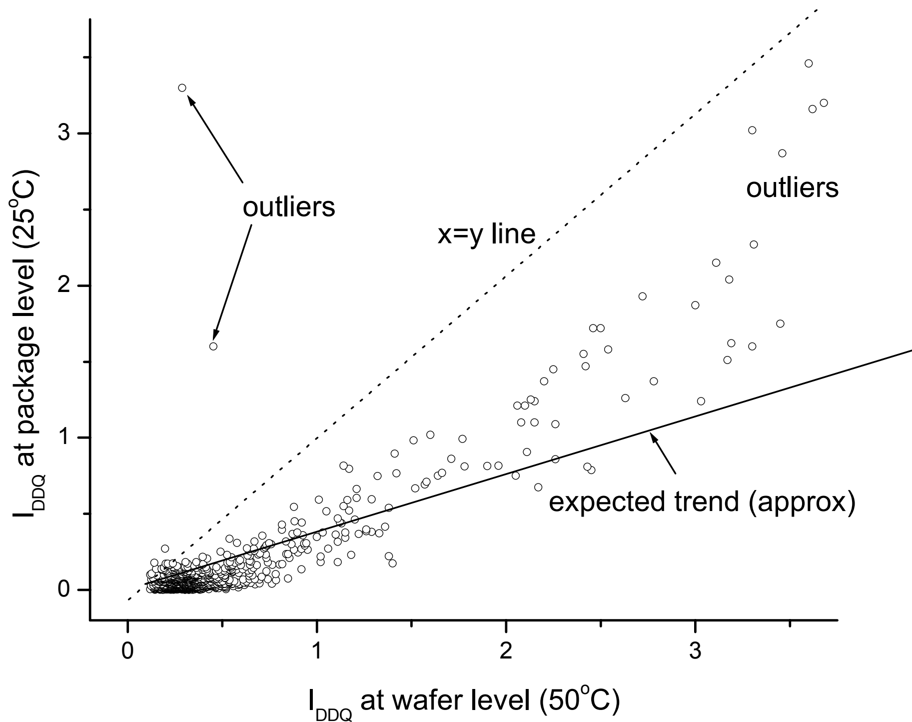


**Fig. 30. $I_{DDQ}$ values at two different temperatures for outlier screening.**

It is possible to use the temperature-based outlier screening by cooling chips and taking a measurement at a reduced temperature. However, it is very expensive and impractical for production implementation, and, therefore, mostly confined to academic interest.

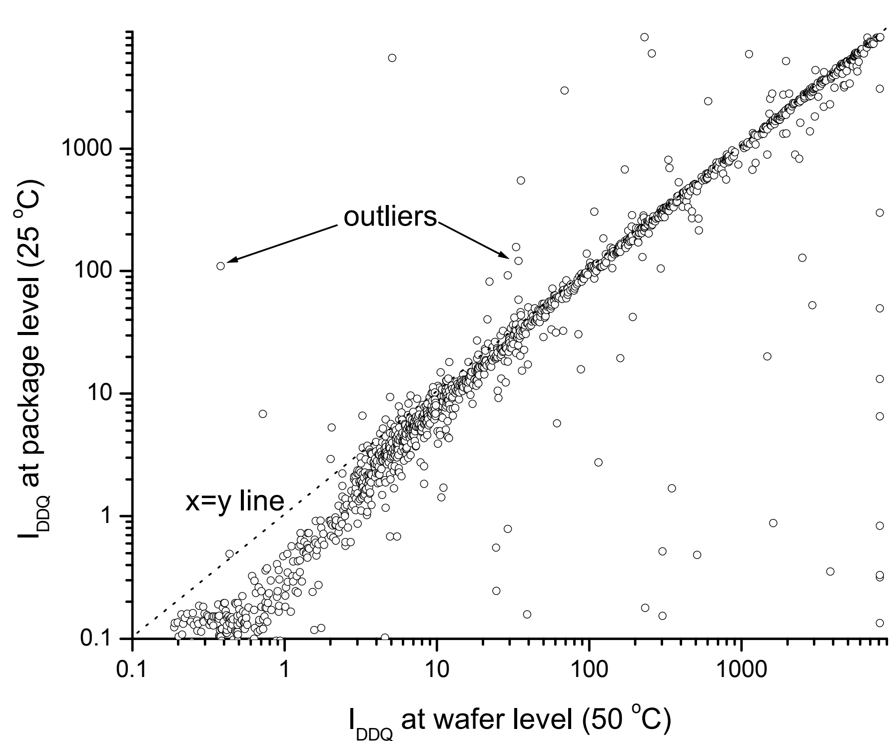


**Fig. 31. Temperature-based outlier screening.**

3.15 $I_{DDQ}$ versus $F_{max}$

$I_{DDQ}$ and transistor delay depend on effective channel length ($L_{eff}$). The smaller the channel length, the faster the transistor can switch (since charge carriers can cross the channel in less time) and the higher is the leakage. The maximum frequency ($F_{max}$) a chip is capable of operating is, therefore, related to the intrinsic leakage current. Fig. 32 shows strong correlation between $I_{DDQ}$ and $F_{max}$ for microprocessor circuits. Some researchers have used the correlation between these two factors to screen defective chips [134]. An estimate of chip frequency ($F_{max}$) can be obtained by using test structures, like ring oscillators. Since test structures are often used for yield learning or improvement [135], the test measurement data for test structures is mostly available at no extra cost. The advantage of exploiting this correlation is that fast and leaky chips can be distinguished from defective chips. As shown in Fig. 33 how the pass/fail thresholds

could be adjusted to reduce yield loss. For Level Sensitive Scan Designs (LSSD) flush delay is shown to have correlation with leakage current [136].

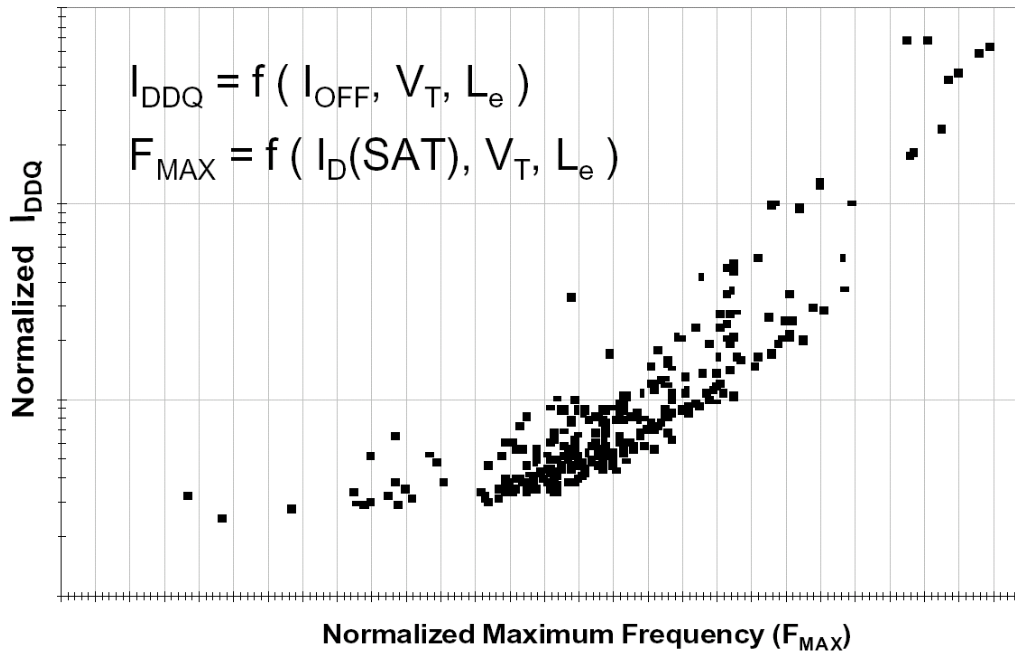


**Fig. 32. Relationship between $I_{DDQ}$ and Fmax.**

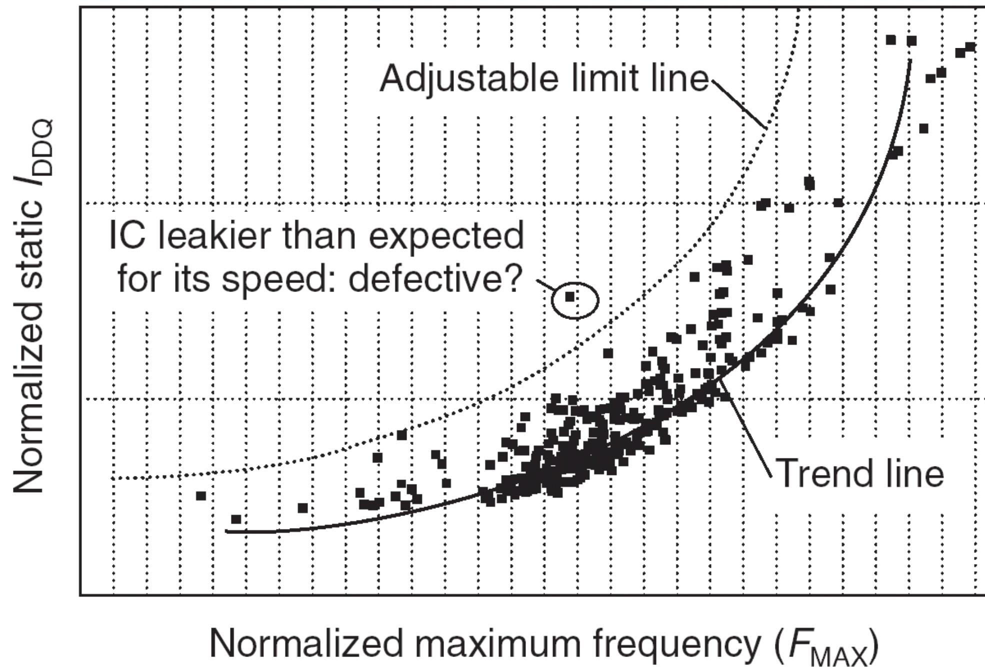


**Fig. 33. Adjustable limit setting for leakier and faster chips.**

## 3.16 $I_{DDQ}$ versus Voltage

Since the fault-free leakage current depends on the power supply voltage, $I_{DDQ}$ measurements at different voltages can be used for distinguishing defective and fault-free chips [137]. The idea itself is not new and has been investigated earlier [116]. The basic idea is to measure leakage current for a vector at different supply voltages and use linear regression to build a model from a sample of chips. In a production test similar measurements are obtained for the chips to be tested. Fault-free chips show conformance to the model and are accepted. Fig. 34 shows a example I-V curves for some defective chips that were uniquely detected using this method. The units on the Y-axis are suppressed to protect sensitive information. The figure shows that the slope of the I-V curve is different for outlier chips than that for the "good" chips. Note that distinction can be misled if only few voltage levels are used. This constraint limits the usability of this technique as chip voltages are scaled down for future technologies resulting in smaller voltage spread. Moreover, production implementation of this method is not straightforward.

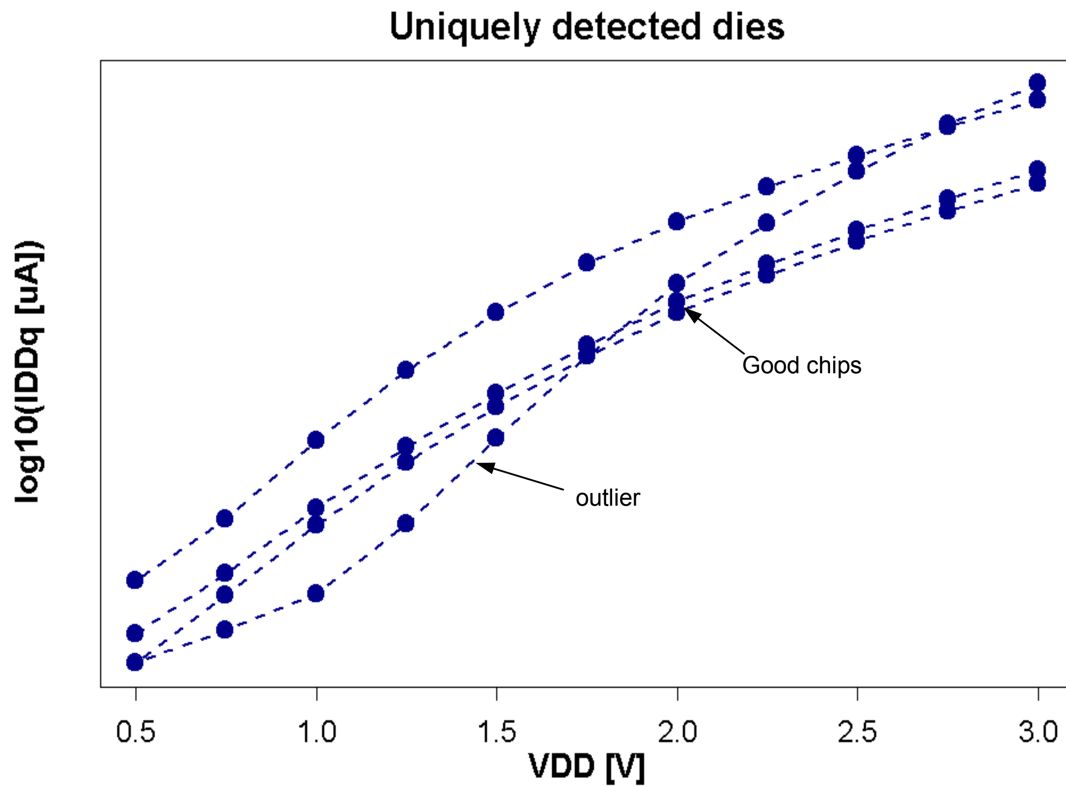


**Fig. 34. I-V curve for some uniquely detected defective chips.**

### 3.17 Principal Component Analysis-Based Linear Prediction

The use of a Principal Component Analysis (PCA)-like method to estimate fault-free $I_{DDQ}$ has been reported [138]. This method exploits the fact that the values of fault-free $I_{DDQ}$ vectors are correlated to each other through an underlying set of process parameters as shown in Fig. 35. It uses a sample of data to train a model and then uses this model for analyzing other chips. Thus, the $I_{DDQ}$ value of one test vector can be predicted from the $I_{DDQ}$ values of one or more other test vectors. For a fault-free chip, the residual values are small. Thus, faulty chips can be identified. Although this method does have certain merit, it can address only deterministic (systematic) process variations. Moreover, the learning sample used for PCA extraction ultimately decides the accuracy of prediction. It also requires a reasonable sample size to be effective.

Factor-analysis (FA) based approach can be used to understand and explain the “intrinsic” variance in $I_{DDQ}$ data. However, these methods do not create new data and interpretation of factor

analysis results is often left to test engineer's discretion [139]. The production worthiness of these methods is often debatable.

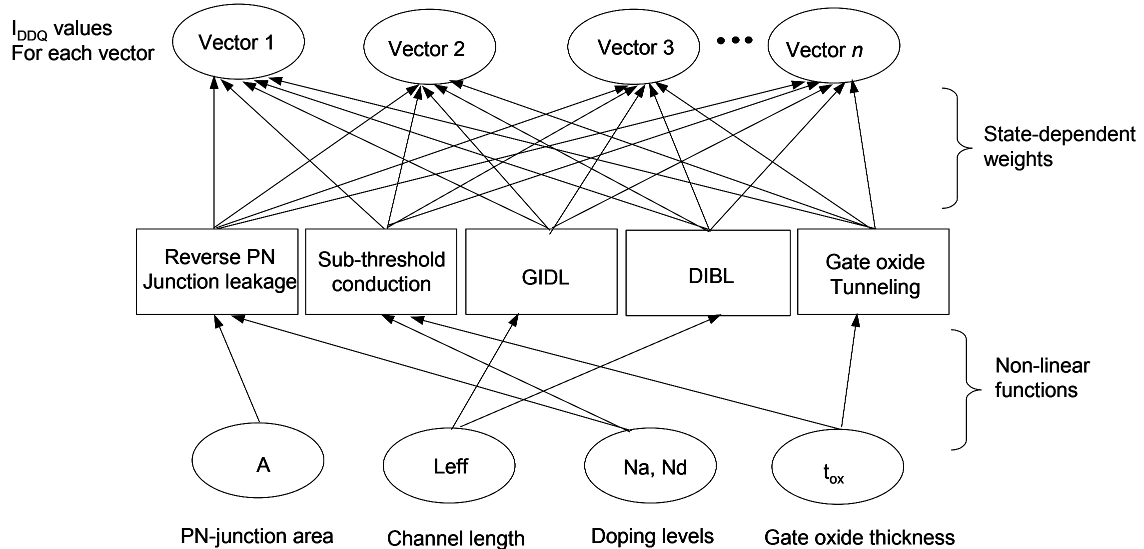


**Fig. 35. Test vector to fabrication process correlation.**

## 3.18 Transient Current ($I_{DDT}$)-Based Test Methods

The use of transient current ($I_{DDT}$) [140–163] test or a combination of $I_{DDQ}$ and $I_{DDT}$ tests for screening defective chips has been explored. $I_{DDT}$ test is the counterpart of $I_{DDQ}$ test and is based on the observation that fault-free CMOS chips draw large current during transition. It enjoys all the advantages of $I_{DDQ}$ test such as no propagation requirement and high fault coverage to vector ratio. Moreover, since the measurement circuit does not require internal circuit activity to settle down, $I_{DDT}$ tests are faster than $I_{DDQ}$ tests. Unlike $I_{DDQ}$ tests, $I_{DDT}$ tests are not restricted to static CMOS circuits and have high resolution for large ICs. They are capable of detecting certain delay faults and open defects as well [151]. However, all $I_{DDT}$ methods necessarily require high-speed measurement circuitry with high accuracy. In addition, similar to $I_{DDQ}$ test methods, defining fault-free $I_{DDT}$ limits is difficult.

## 3.19 $I_{DD}$ Pulse Response Testing (PRT)

In PRT [141][142], both power supply rails are pulsed simultaneously from the midpoint voltage to their nominal values (e.g. $V_{DD}$ rail from $V_{DD}/2$ to $V_{DD}$ and $V_{SS}$ rail from $V_{DD}/2$ to 0)

while bias voltages to all inputs are set to their midpoint values. During pulsing, transistors enter either sub-threshold, linear or saturation regions and their current characteristics can be observed. This method is vector independent, suitable for both digital and analog circuits and capable of detecting gate oxide shorts, opens, and poly or metal bridges [142]. The authors showed that analysis could be done in the time or frequency domain [145].

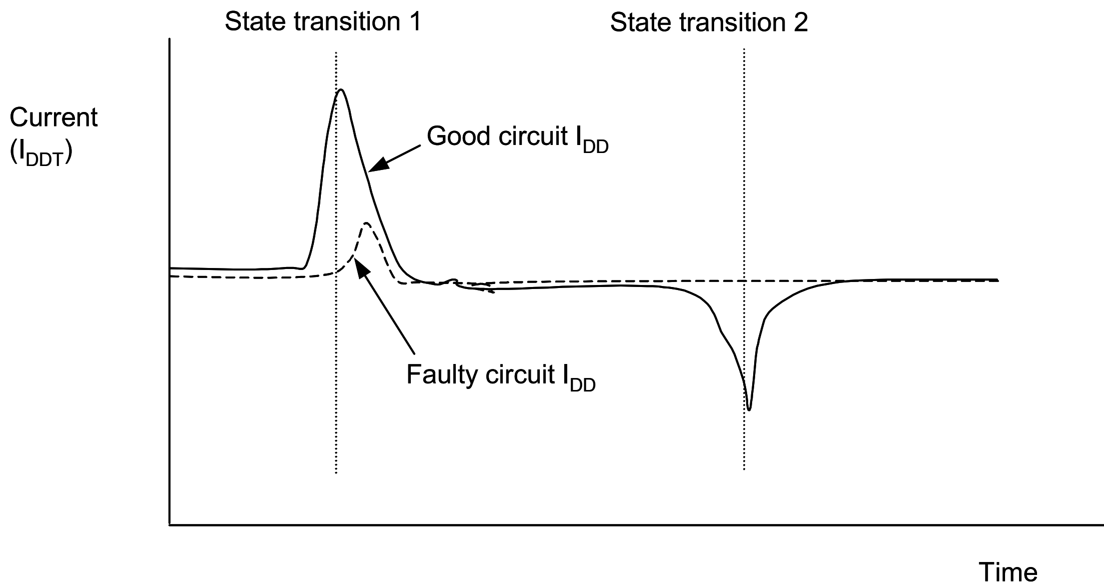


**Fig. 36. Comparison of good circuit $I_{DDT}$ response to faulty circuit $I_{DDT}$ response.**

3.20 Dynamic Power Consumption Current-Based Testing

Fig. 36 shows a typical $I_{DDT}$ response for a good and a faulty circuit. Note that the defect is not detectable by $I_{DDQ}$ measurement alone since there is no appreciable change in the quiescent value of the current. The power consumption peaks occur because for a brief period when both PMOS and NMOS transistors conduct simultaneously, thus short-circuiting the power supply. The current decays exponentially due to capacitive charge/discharge in the circuit. In this method, an integrator circuit is used to measure the dynamic power consumption of the circuit. The output voltage of the integrator is proportional to the dynamic power consumption of the circuit. If this value exceeds the predetermined threshold, the chip is considered defective. $I_{DDT}$ testing is capable of detecting certain open and parametric defects not detectable by $I_{DDQ}$ test, as they do not result in increased leakage [144]. Similar techniques are shown to be capable of

detecting bridging defects in static CMOS [148] as well as domino CMOS circuits [153].

### 3.21 Transient Signal Analysis (TSA)

TSA is based on a measurement of the contribution to the transient response of a circuit by physical characteristics such as substrate, power supply, parasitic capacitive coupling, etc [147]. In this method, transients are analyzed at multiple test points. Under the assumption that process variation is uniform across a die, TSA can distinguish between the changes in the transient response caused by defects and those caused by process variation. If changes are caused by process variation, the transients are correlated for fault-free devices. On the contrary, the presence of a defect alters transients at test points closer to the location of the defect. Recent studies of using TSA for defect localization [154][155] and for detecting delay defects [156] have been reported.

### 3.22 Frequency Spectrum Analysis of Dynamic Current

Analysis of the frequency spectrum of transient waveform for fault detection has been investigated [158]. Thibeault proposed sampling the $I_{DD}$ waveform several times per clock cycle in order to extract more information from the signal than a simple DC level used in $I_{DDQ}$ test. Fig. 37 shows the overview of this method called $I_{DDF}$ testing. The basic theme is to sensitize a given path to make defects along the path alter the $I_{DD}$ waveform and detect the alterations using frequency spectrum analysis. The first component harmonic is usually different for fault-free and faulty $I_{DD}$ waveforms. Frequency spectrum analysis using an 8-point FFT was shown to be sufficient for detecting significant current waveform alterations caused by a defect [158]. Frequency domain analysis is shown to be useful in detecting single and multiple faults as well as distinguishing between different faults [159]. $I_{DDF}$ testing has better noise immunity than $I_{DDQ}$ test. However, as $I_{DD}$ frequency increases, higher sampling rates are required. This requires expensive test equipment.

### 3.23 On the Fly Depowering

Transient current measurement while depowering a chip was proposed in [160]. In this case the detection of abnormal current is done by the whole vector set rather than vector-by-vector. The approach is based on Keating-Meyer method [161]. In this method, a CUT is depowered using a switch and the rate at which the capacitor connected to $V_{DD}$ discharges while input

vectors are being changed is monitored. A defective circuit consumes more energy and discharges faster than a fault-free circuit. The resolution depends on the decoupling capacitance.

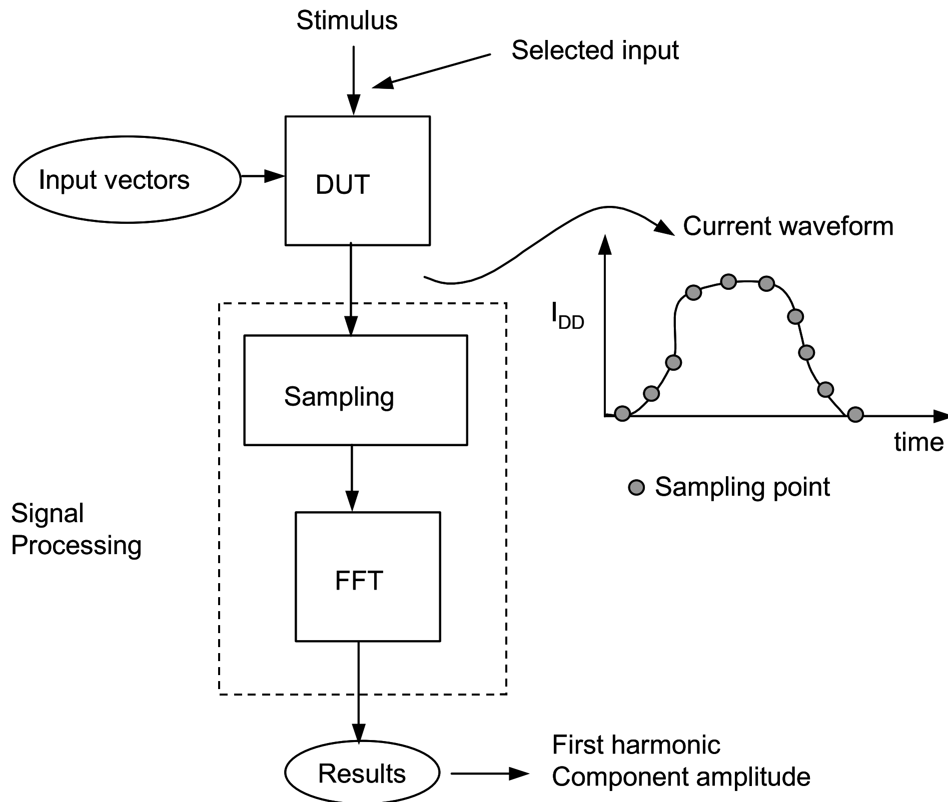


**Fig. 37. Overview of $I_{DDF}$ testing method.**

3.24 Energy Consumption Ratio (ECR)

Recently use of average dynamic currents consumed by the circuit for fault detection has been reported [162][163]. The advantage of this method is that average dynamic currents are larger and, therefore, easier to measure than static currents. ECR relies on the fact that a fault alters the number and location of signal transitions that occur due to a change in input. In other words, a fault can alter the energy consumed by the circuit. It uses two pairs of transitions, which are alternated at the input of the circuit, and ECR is the ratio of currents (or energies) consumed by the two transitions. ECR is immune to process variations to the first order as the effect of process changes affects both the numerator and the denominator and gets canceled due to the ratio. Like all other $I_{DDX}$ methods, deciding the fault-free ECR threshold is not trivial. The use of

statistical techniques like PCA is useful for this purpose [138].

A summary of $I_{DDQ}$ and $I_{DDT}$ test methods is tabulated in Table II and Table III, respectively.

### 3.25 Fault-Free Parameter Value

The definition of "fault-free" is very subjective in the case of continuous parameter data. Fig. 38 shows the distribution of $I_{DDQ}$ values for a vector across different chips. All these chips passed Boolean tests. Notice that the distribution has a long tail. Obviously, the chips in the tail have a different leakage mechanism than those in the central part of the distribution. However, determining whether these chips are so flawed that they must be rejected is extremely difficult. From a statistical perspective, the chips exhibiting abnormal behavior are *outliers*. Thus, it is possible to use outlier rejection methods to screen defective chips [164]. In the next section, we discuss the philosophy behind outlier rejection and its applicability to $I_{DDQ}$ data analysis.

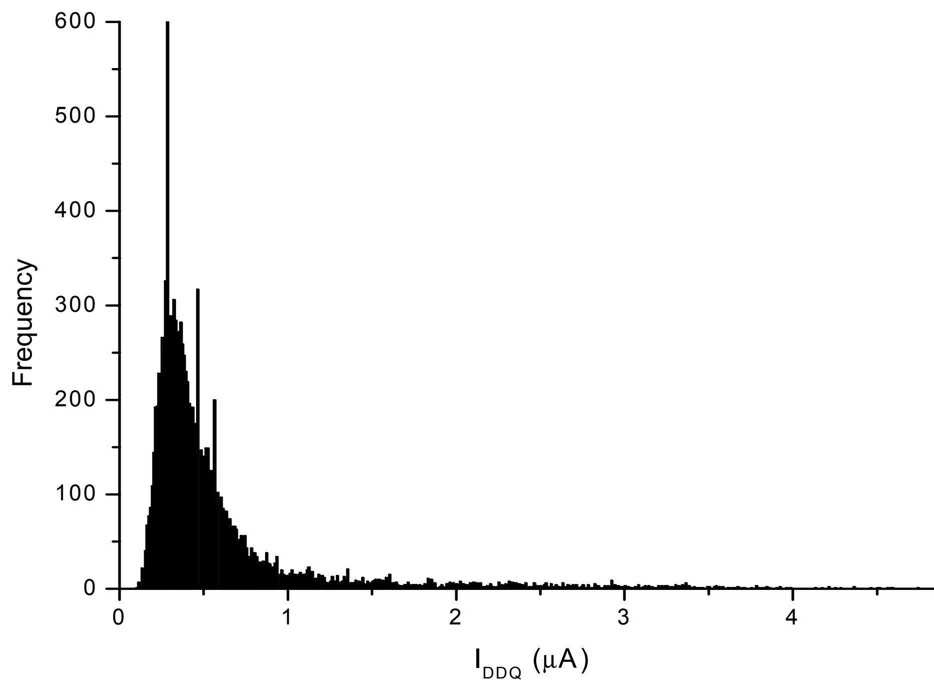


**Fig. 38. Distribution of $I_{DDQ}$ values from different chips shows a long tail.**

**TABLE II. SUMMARY $I_{DDQ}$-BASED TEST METHODS.**

| Method | Features | Advantages | Limitations |
|---|---|---|---|
| $I_{DDQ}$ Vs Temperature | Fault-free leakage current increases with temperature. Differences in leakage currents measured at two temperatures can distinguish faulty chips. | Simple | Low temperature measurement is too expensive. High temperature measurement may not be cost effective. |
| Delta $I_{DDQ}$ | Differences (deltas) between two $I_{DDQ}$ readings for consecutive vectors show near-zero mean and small variance for a fault-free chip. | Simple Ease of implementation. | Deciding fault-free variance is not easy Cannot screen passive defects |
| Current Ratios | Ratio of maximum $I_{DDQ}$ to minimum $I_{DDQ}$ of fault-free chips is relatively constant. | Simple Easy to implement | Difficult to set low threshold for screening passive defects. |
| Statistical Clustering | Chips are grouped using statistical clustering method that groups data such that chips in a group have natural association with chips from the same group. | Accounts for process variation | Needs many readings for good clustering. Labeling clusters (faulty or fault-free) is difficult. |
| Statistical Outlier Rejection | Statistical outlier rejection methods are applied for screening defective chips. | Simple Many outlier rejection methods are available. | Outliers affect the distribution properties Difficult to distinguish true outliers in continuous distribution. |
| Current Signature | Sorted $I_{DDQ}$ values show steps or "jumps" for active defects and smooth signature for fault-free chips. | Simple, intuitive Considers intra-die variations. | Number of measurements determines screening resolution. Deciding fault-free step size is difficult. Cannot screen passive defects. |
| Eigen Signatures | Relationship between mean $I_{DDQ}$ and variance is explored to distinguish between faulty and fault-free chips; leakage variance is allowed to be proportional to the mean value. | Accounts for process variation. | Elaborate analysis may be needed Outlier rejection treatment is subjective. |

**TABLE III. CONTINUED.**

| Method | Features | Advantages | Limitations |
|---|---|---|---|
| Principal Component Analysis (PCA) | Exploits the vector-to-vector correlation between chips due to process variation. | Accounts for process variations. | Threshold setting is difficult. |
| Nearest Neighbor Residual (NNR), Spatial fit, Neighbor Current Ratio (NCR) | Neighboring chips on the same wafer are used for estimating maximum fault-free $I_{DDQ}$ to identify wafer-level "spatial outliers" (NNR, spatial fit, NCR, INDIT). | Accounts for process variation.<br>No model building needed.<br>Scaleable to new technologies. | Prior knowledge of wafer patterns may be necessary.<br>Gross outliers must be rejected up front which requires threshold setting to identify "true" outliers. |
| $I_{DDQ}$ Vs $F_{max}$/ Flush Delay | The fact that the fault-free leakage and delay or $F_{max}$ are correlated is exploited to distinguish between faulty and fault-free chips. | Simple<br>Accounts for process variations. | May be applicable to mature, well-controlled processes only.<br>Cannot do per-chip analysis. |

**TABLE IV. SUMMARY OF $I_{DDT}$-BASED TEST METHODS.**

| Method | Features | Advantages | Limitations |
|---|---|---|---|
| Pulse Response Testing (PRT) | Pulses both $V_{DD}$ and $V_{SS}$ power rails while applying fixed bias to input. Temporal and spectral analysis of $I_{DD}$ is used to differentiate faulty and fault-free chips. | Test vector independent. Applicable to both digital and analog circuits. | May be impractical in production. Characterization and model building is difficult. Effectiveness falls for DSM technologies due to reduced supply voltage. |
| Dynamic Power Consumption Measurement | Integrate the $I_{DD}$ value and convert to voltage, if voltage exceeds the threshold the chip is rejected. | Simple | Slow. Does not account for process variations. |
| Transient Signal Analysis (TSA) | Uses multiple test points to sample $I_{DD}$. In the presence of a defect only $I_{DD}$ sampled at test points closer to defect site are affected, thus distinguishing fault-free and faulty $I_{DD.}$ | Accounts for process variation Scaleable to new technologies | Requires multiple test points. |
| Frequency Spectrum Analysis | $I_{DDT}$ waveform is sampled at multiple points. Harmonic analysis is used to differentiate between faulty and fault-free chips. | Insensitive to noise More robust technique | Signal processing involved. Difficult and expensive for high-frequency chips. |
| On the Fly Depowering | Time required to discharge a capacitor connected to $V_{DD}$ as inputs are changed is used to distinguish between fault-free and faulty circuits. | Simple Small hardware overhead | Need to calibrate for dynamic current. No defect localization. Does not account for process variations. |
| Energy Consumption Ratio (ECR) | Applies a pair of vectors at the input and are alternated. The ratio of average currents (energies) consumed is used for screening. | Simple, insensitive to process variations | Requires more test generation effort to maximize effectiveness. Pass/fail threshold selection issue exists. |

# 4. OUTLIER IDENTIFICATION AND PARAMETRIC TEST DATA

## 4.1 Introduction

The basic premise behind fault detection is that a fault when excited manifests itself by a change in at least one of the test parameters; voltage for Boolean tests and current or speed for parametric tests. The parameter can be static or transient as discussed in the previous section. If no test parameter changes or is not observable at the outputs, the fault cannot be detected and is called a *redundant fault*. Therefore, a chip is considered faulty only if its behavior is different from that of fault-free chips. Statistically, a chip is considered an *outlier* if its behavior is *different* from that of the rest of the chips (assumed fault-free). The distinction between a defective chip and an outlier chips comes from the fact that a chip is outlier from a parametric test perspective and passes all functional tests, unlike a defective chip which fails at least one functional test.

It was shown in Section 3 that the leakage current distribution has a long tail and chips in the tail of the distribution are essentially outliers. This section provides some statistical background for outlier identification. Outlier detection in itself is a vast topic in statistics and some excellent texts are available [19][165]. The goal here is to provide background information in this area and then evaluate its applicability to VLSI testing using parametric test data.

## 4.2 What Are Outliers?

Instrumentation is the science of measurement. Measurement is a fundamental task to many scientific observations. Sometimes certain observations in data do not appear to belong with the rest. These observations could be a result of improper calibration, human error or uncontrollable external event. These anomalous observations are called *outliers*. It is often of interest to understand the underlying mechanism that produced these anomalies. However, usually only partial information on mechanisms underlying observations is available, and hence we may not know about unusual circumstances that affect the data. In this case, we may have to rely on the data in judging which observations are outliers. Statistical procedures that enable us to achieve this are called outlier detection or rejection methods.

## 4.3 Outlier Identification, Detection and Rejection

Whether to simply reject outliers or retain them in the data set is a subjective decision. Scientists, engineers and statisticians often face this challenge. In certain cases, the deduction

could be meaningful only if some data points are discarded. On the other hand, several important discoveries in science have occurred because scientists relentlessly attempted to find the origin of outliers. There are proponents of both the viewpoints – those that advocate rejection of outliers and those who advise inclusion of outliers no matter how abnormal they appear. Therefore, outlier detection is an important topic in statistics.

Before we delve into this topic further, it is necessary to make a distinction between the terms ‘identification’, ‘detection’ and ‘rejection’ of outliers. Identification and detection of outliers both involve examination of data to check whether some suspicious data exist. This process can start with certain assumptions about the underlying distribution and hypothesize that the data is polluted by outliers [165] or can be an impartial examination of data. Either way the purpose is to examine the change in the properties of the data after screening outliers. Identification or detection does not necessarily imply rejection. In outlier rejection, the purpose is to “improve” the data to draw meaningful conclusions from it. This “improvement” is achieved by rejecting data marked suspicious by the identification process.

This statistical distinction takes a different form when applied to VLSI testing. The purpose of using outlier detection is to screen apparently defective chips. Thus, once a chip is marked an outlier by a certain outlier detection method, it must be considered defective. In principle, it is possible to alter the test strategy for optimum cost. Thus depending on how outlying a chip is, it may be rejected or subjected to BI or more stringent testing as shown in Fig. 39. However, for our purpose outlier detection/identification also means outlier rejection. Henceforth, we use the terms ‘outlier detection’ and ‘outlier rejection’ interchangeably as we assume that the purpose of outlier detection is to reject data (chips) after certain qualification. Of course, outlier detection essentially involves some form of threshold setting. The threshold decides how flawed a chip’s parameters need to be before it can be considered an outlier.

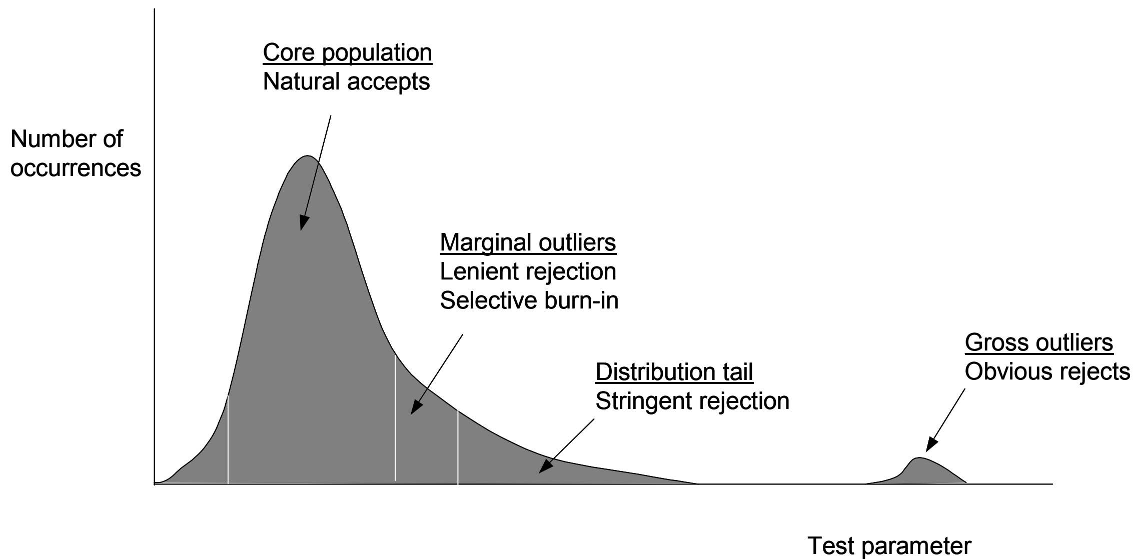


**Fig. 39. Test strategy can be modified depending on "outlier-ness".**

## 4.4 Why Do Outliers Exist?

There are two reasons why some chips are outliers. Firstly, a chip may be defective and appears to be an outlier as the defect alters its parameters. Secondly, the variation in the processing conditions can also affect chip parameters and, an unfortunate combination of statistical parameters can make it appear to be an outlier [166]. Since variability occurs due to process changes, outliers are inevitable. If the parameter variation is due to processing conditions and is legitimate, a chip cannot be considered faulty. This is conceptually illustrated in Fig. 40. It shows outlier behavior as a result of four possible reasons. Therefore, one of the goals of outlier detection is to understand whether outlier behavior stems from a defect or other reason. A secondary yet important goal of outlier detection is to explain whether outlier behavior is benign or fatally flawed from a reliability perspective.

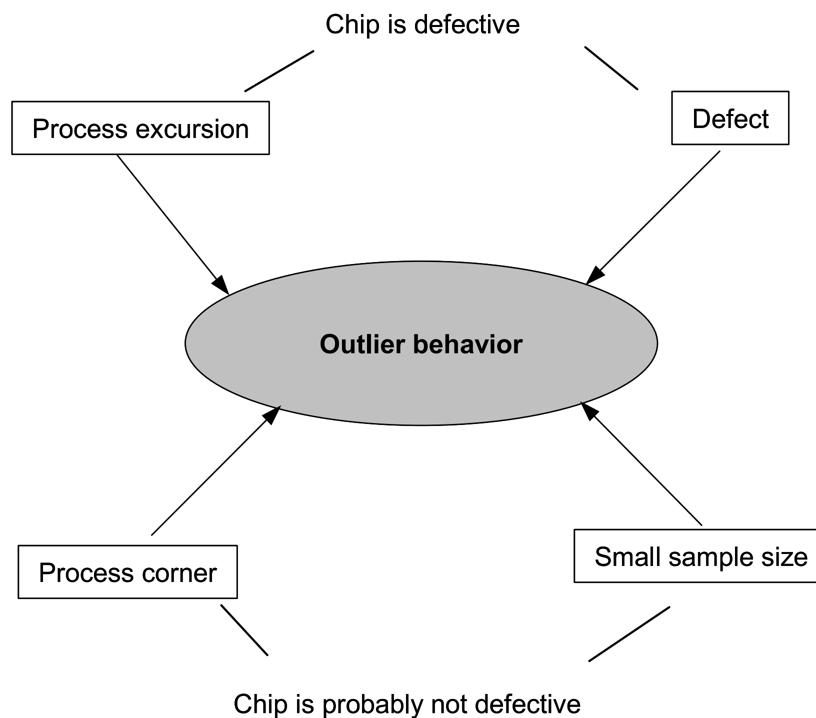


**Fig. 40. Outlier behavior can stem from different reasons.**

## 4.5 Challenges in Outlier Rejection

### 4.5.1 Outlier definition is relative

Defining an observation to be an outlier itself is not trivial. The definition of an outlier is dependent on “normal” variation in data. This variation is usually expressed in terms of the mean and variance of standard distributions. Many natural phenomenon and random variations can be approximated by the Normal distribution, which is a bell-shaped curve with characteristic parameters mean and standard deviation. The symmetrical Normal distribution has two interesting characteristics: first, the mean ($\mu$) is at the center of the distribution and 99.67% of all values lie within three standard deviations ($\pm 3\sigma$) of the mean value. Thus, a data point outside these limits has very low probability of occurrence and can be considered an outlier. A symmetric distribution shown in Fig. 41(a) is skewed due to outliers as shown in Fig. 41(b). Unfortunately, the very presence of outliers alters the properties like the mean and standard deviation of the distribution. In other words, the presence of outliers affects what can be considered “normal” variation. This makes outlier detection a chicken-and-egg problem.

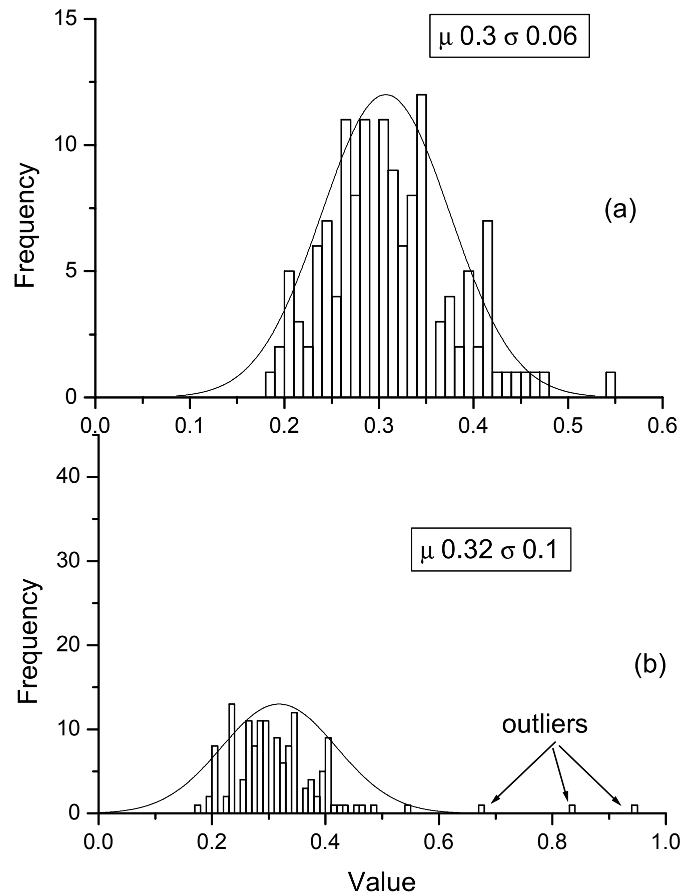


**Fig. 41. (a) A symmetrical distribution gets skewed (b) due to outliers.**

### 4.5.2 Outliers can be misinterpreted

If enough data is not available, some data points may appear to be outliers. As the sample size increases, however, these "apparent" outliers disappear and may get grouped with the core of the population. Therefore, it is important to examine whether outliers appear because of too little data; otherwise the analysis could be misleading [167]. It must be emphasized that several important discoveries have occurred because scientists relentlessly pursued the origin of the presence of outliers [168].

### 4.5.3 Outliers can be hidden

While too few samples can label legitimate readings as outliers, another problem can come from too many samples. If the data size is too large and outliers are large in numbers, they can appear as "normal" variation of parameters. The distributions of $I_{DDQ}$ values for a wafer and a lot are illustrated in Fig. 42 and Fig. 43, respectively. Notice that several chips that appear to be outliers in Fig. 42 do not seem to be outliers in Fig. 43.

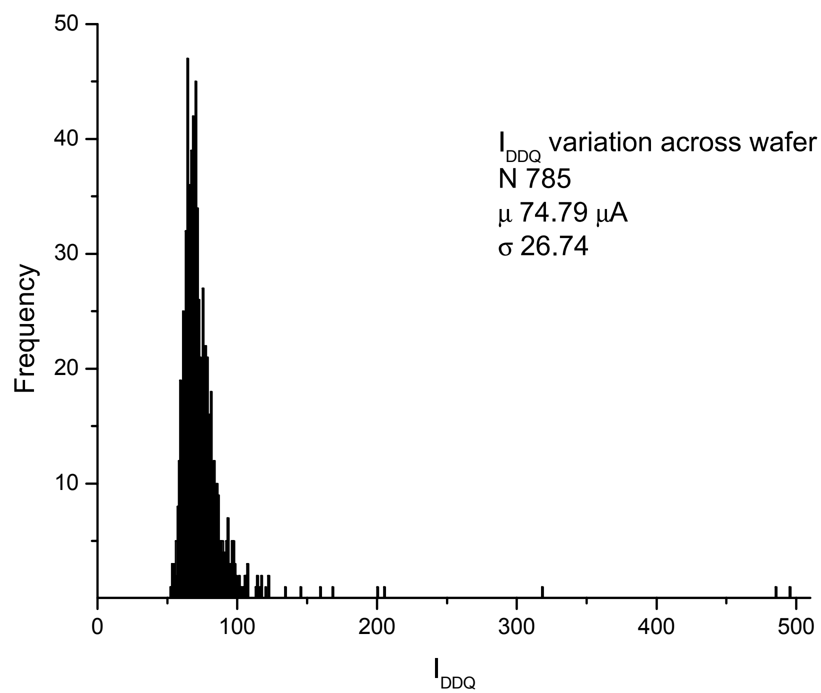


**Fig. 42. $I_{DDQ}$ variation across a wafer.**

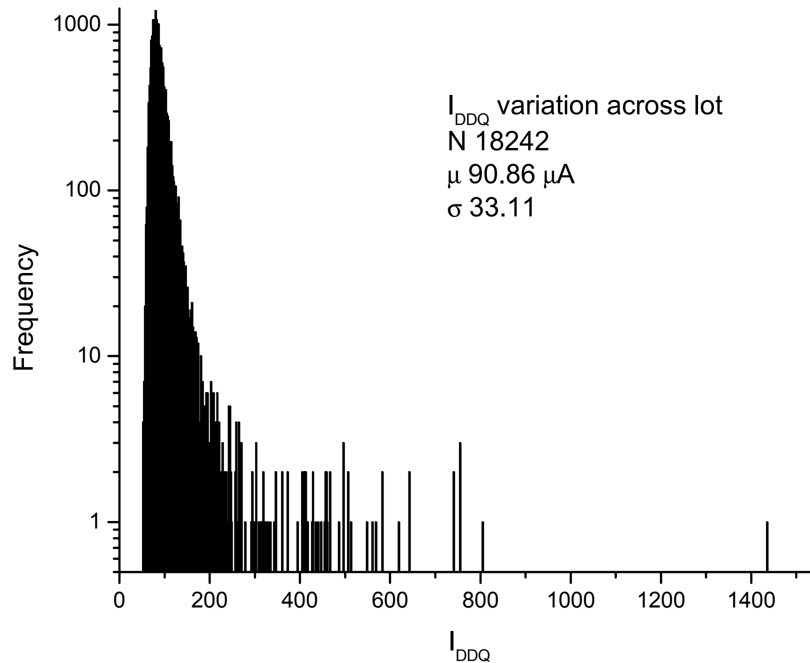


**Fig. 43. $I_{DDQ}$ variation across a lot.**

4.5.4 Successive outlier rejection can nibble entire data set

Since the presence of outliers affects the properties of a standard distribution, as some outliers are rejected these properties change. Thus, in the new distribution, some other data points appear as outliers. The net result is that with each successive pass of outlier rejection, some part of the data appears as outliers. This is illustrated with the help of histograms of $I_{DDQ}$ test data in Fig. 44(a) through Fig. 44(h). In each case, chips having leakage current more than three standard deviations ($3\sigma$) above mean ($\mu$) were rejected. The properties of the data are shown in Table V. Note that with each pass the mean and standard deviation reduces. Also, notice that there is no clear stopping criterion. Outlier rejection therefore requires judicious use of statistical methods and understanding of the underlying mechanisms that produce the data.

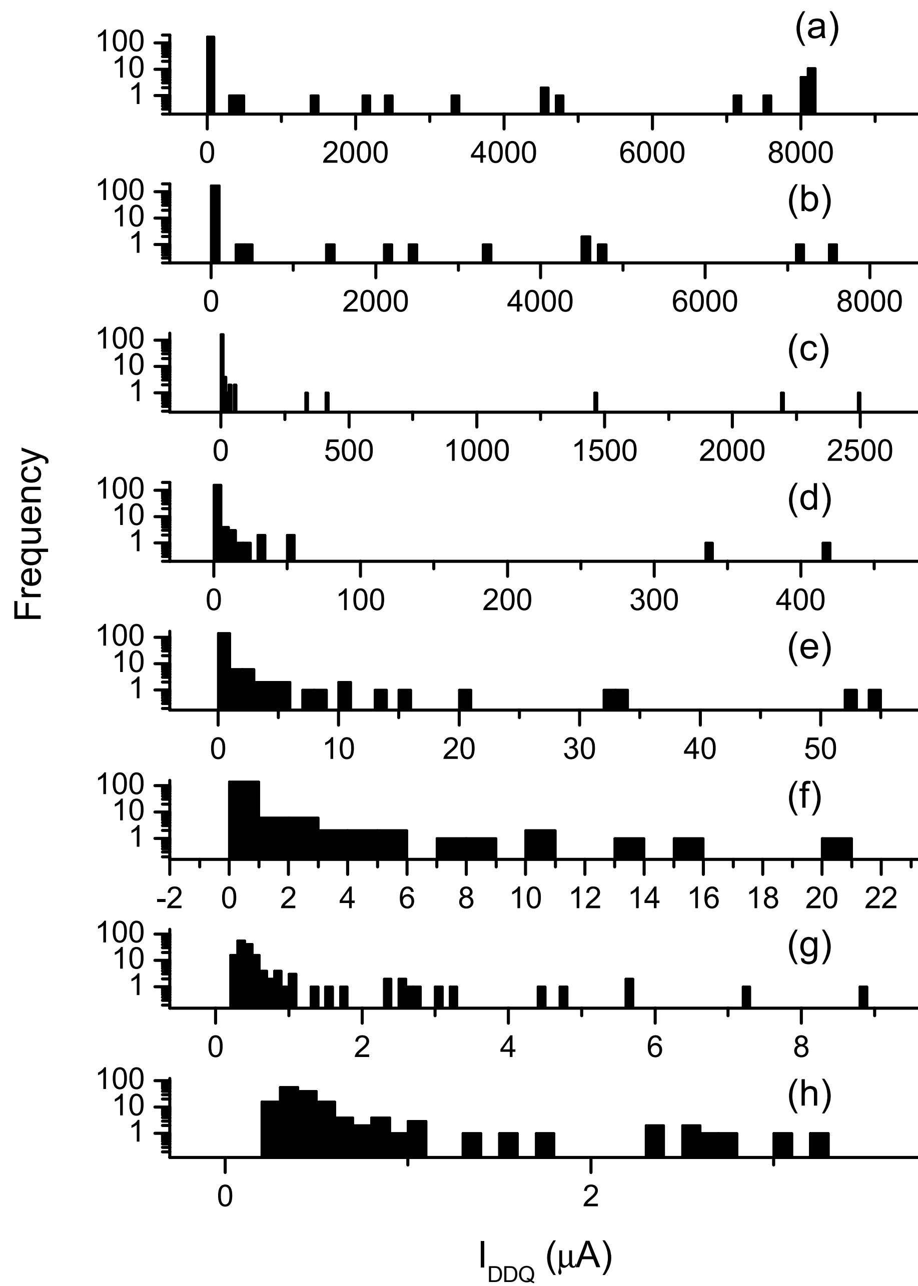


**Fig. 44. Successive outlier rejection can nibble away data.**

**TABLE V. PROPERTIES OF DISTRIBUTIONS SHOWN IN FIG. 44.**

| Figure | N | Mean (μ) | SD (σ) | Limit (μ+3σ) | No. of rejects |
|---|---|---|---|---|---|
| (a) | 197 | 856.87 | 2365.93 | 7954.66 | 16 |
| (b) | 181 | 216.32 | 1020.05 | 3276.47 | 6 |
| (c) | 175 | 41.51 | 273.85 | 863.06 | 3 |
| (d) | 175 | 6.53 | 41.13 | 129.92 | 2 |
| (e) | 170 | 2.16 | 7.02 | 23.22 | 4 |
| (f) | 166 | 1.18 | 2.67 | 9.19 | 5 |
| (g) | 161 | 0.77 | 1.21 | 4.4 | 6 |
| (h) | 155 | 0.57 | 0.54 | 2.19 | 8 |

### 4.5.5 Outliers are subjective

It should be noted that the test parameter being monitored is an analog quantity and therefore has a continuous distribution. The pass/fail distinction is binary. Thus, a pass/fail threshold imposes a binary value on an analogue quantity. Since the binary distinction is artificially created and externally imposed on the analog quantity, it is always subjective.

### 4.5.6 Data transformations may be hard to find

Most outlier detection methods assume that the data follows a certain distribution. For example, Chauvenet's criterion – a method described later – assumes that data has a Normal distribution. If outlier-free data does not conform to this assumption, an appropriate transformation must be applied before outlier analysis. For example, a Lognormal distribution can be converted to a Normal distribution by taking the natural logarithm of scaled values. Fig. 45 illustrates the effect of such a transformation on a typical $I_{DDQ}$ distribution. After applying an outlier rejection method to the transformed data, it is necessary to apply the reverse transformation to get the original data. Although such transformations are well defined for standard distributions, they are extremely hard to find for non-standard distributions.

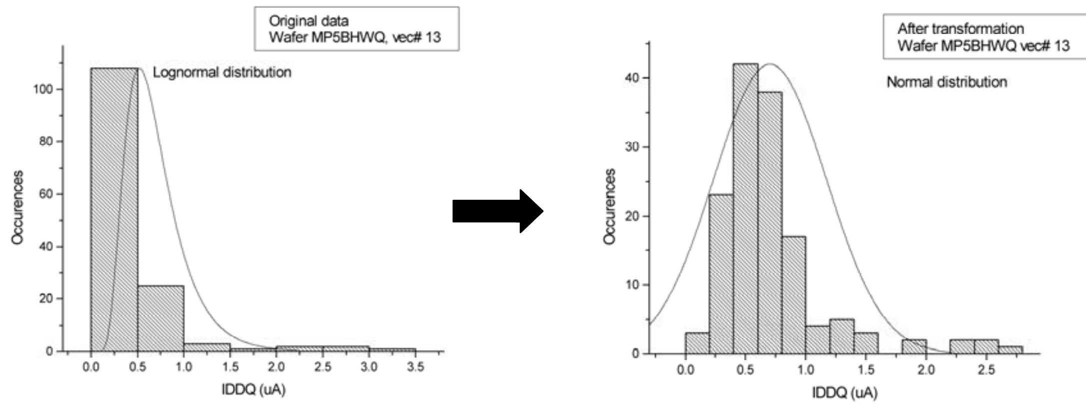


**Fig. 45. Transformation of Lognormal data to Normal distribution.**

4.5.7 Statistics is a tricky business!

Drawing meaningful conclusions from numbers is not easy [169][170]. Many times visual interpretation of data can be misleading. Use of averages, medians, and standard deviations must be judicially examined for meaningful analysis.

4.6 Outlier Detection Methods

There are many outlier detection methods available. In this section, we provide a brief overview of philosophy behind outlier detection methods and discuss three methods – Chauvenet's criterion, Tukey method, and Median of Absolute Deviations (MAD) about medians. These methods are later used for data analysis. Although many outlier rejection methods are available, the principle of parsimony or Occam's razor, which says, "what can be done with fewer assumptions is done in vain with more", remains applicable [171].

It must be emphasized here that no outlier detection method offers a panacea. Each method has its own drawbacks, and which method is best suited for an application depends on the distribution of the data. The use of outlier rejection must be accompanied by an understanding of the underlying physical mechanisms that cause the variation in parameters. Even if a certain reading is not rejected by the outlier rejection method, if it appears suspicious and no physical mechanism can explain such behavior, user discretion must be used for accepting or rejecting the reading.

## 4.7 Resistant Outlier Detection

The sensitivity of a statistical procedure to one or two outlying observations is an important criterion for selecting an outlier rejection method. A statistical procedure is considered *resistant* to outliers if it does not change very much when a small part of the data changes even drastically [171]. For example, consider the hypothetical sample: 10, 20, 30, 50, and 70. The sample mean is 36, and the sample median is 30. If we change the value 70 to 700, the sample mean becomes 162 but the median remains 30. The sample mean is not a resistant statistic because it is influenced by a single observation. The median, however, is resistant. Later in this section, we review an outlier rejection method based on this resistant property of the median.

## 4.8 Philosophy Behind Outlier Rejection

If normal variation in parameters is known from a theoretical model or by any other means, any abnormal variation cannot be considered fault-free. In general, theoretical models cannot account for all conditions that introduce variation. Thus, the normal variation is not known a priori and must be deduced from the available data through empirical analysis. Therefore, most, if not all, statistical procedures use certain parameters obtained from the distribution of the data and use these values as a starting point to determine normal distribution properties. Methods such as linear regression, Analysis of Variance (ANOVA) or multiple analyses of variance (MANOVA), covariance matrix, etc. invariably attempt to extract the central tendency of the population. As lot-level variations increase for DSM technologies, it is important to determine primary source of variation [172] for better SPC and outlier rejection.

## 4.9 Chauvenet's Criterion

This method was proposed by the American astronomer Chauvenet in 1863 [173] and remains one of the simplest criteria to reject outliers [174]. It determines the probability that a seemingly illegitimate reading can occur in a data set. If this probability is less than a threshold value, it is discarded. The threshold probability used is usually 0.5.

Assume a data set having *n* readings with the mean and standard deviation $\mu$ and $\sigma$, respectively. Whether a reading $k_{sus}$ is illegitimate or not is decided as follows:

$$t_{sus} = \frac{k_{sus} - \mu}{\sigma} \tag{4.1}$$

where $t_{sus}$ is the number of standard deviations by which $k_{sus}$ differs from $\mu$. The probability

P(outside $t_{sus}.\sigma$) is obtained from the standard probability tables for a Normal distribution and is multiplied by *n*.

$$K = n \cdot \mathrm{P}(\text{outside } t_{sus} \cdot \sigma) \quad (4.2)$$

If *K* is less than 0.5, the reading is rejected. An example of the application of Chauvenet's criterion is illustrated in Appendix E. According to Chauvenet's criterion, an observation is rejected if it lies outside the lower and upper 100/(4*n*) percentiles of the null distribution. In a sample size *n*, the probability that an arbitrary observation is rejected is 1/(2*n*). Hence with this procedure the chance of *wrongly* rejecting a reasonable sample value is $1-e^{-0.5}$ or about 40%!

4.10 Tukey Method

The Tukey method assumes that the distribution is Normal. Two quartile points (*Q1* and *Q3*) are defined such that $1/4^{th}$ of the readings are less than *Q1* and $1/4^{th}$ of the readings are greater than *Q3*. Then the Inter Quartile Range (*IQR*) is defined as

$$\mathrm{IQR} = \mathrm{Q3} - \mathrm{Q1} \quad (4.3)$$

The lower quartile limit (*LQL*) and upper quartile limit (*UQL*) are defined as

$$\mathrm{LQL} = \mathrm{Q1} - 3\mathrm{IQR} \quad (4.4)$$

$$\mathrm{UQL} = \mathrm{Q3} + 3\mathrm{IQR} \quad (4.5)$$

The readings outside of this range are considered outliers (see Fig. 46). Of course, for $I_{DDQ}$ testing, only the *UQL* need be used. The multiplying factor in the LQL and UQL formulae is a parameter of choice and other values can be used.

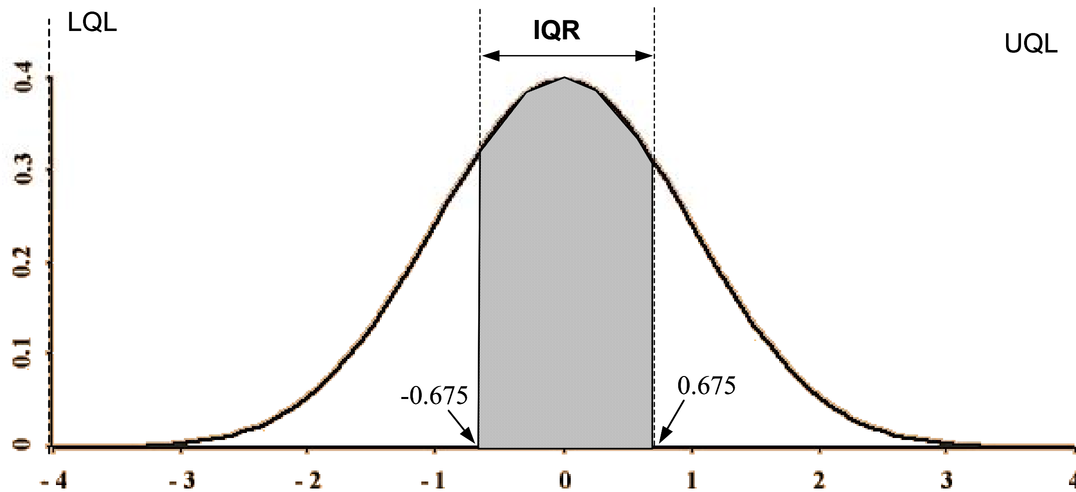


**Fig. 46. Parameter definition for Tukey method.**

## 4.11 Median of Absolute Deviations (MAD) About Medians

Many methods for outlier rejection like Chauvenet's criterion [174], the Tukey test [175] or Z-scores [19] rely on distribution properties like mean and variance. The presence of outliers in the data causes a shift in the mean and variance. Thus, many "true" outliers are not detected. Furthermore, many of these methods assume the data has Normal distribution. A typical $I_{DDQ}$ distribution has a long tail due to outliers. The $I_{DDQ}$ distribution for fault-free chips can be approximated by a lognormal distribution. Outliers do not follow any standard distribution.

For successful outlier detection, we need a resistant estimator that should not be unduly affected by outliers in the sample. The Median of the Absolute Deviations about the median (MAD) is such an estimator [19]. It is defined as:

$$MAD = median_i \left\{ \left| x_i - \tilde{x} \right| \right\} \tag{4.6}$$

where $\tilde{x}$ is the sample median. Then the MAD score ($M_i$) is defined as:

$$M_i = \frac{0.6745\left(x_i - \tilde{x}\right)}{MAD} \tag{4.7}$$

The constant 0.6745 is used because for large N for a Normal distribution E(MAD) =

0.6745σ. The $M_i$ is similar to Z-scores. Any observation is labeled as an outlier and rejected when $|M_i| > D$ where D is the maximum permissible MAD score. For large N and a Normal distribution, a value of 3.5 for D is suggested in the literature [19]. To clarify the MAD approach an example of MAD-based outlier rejection is illustrated in Appendix E.

Since outliers do not change the median appreciably, MAD-based rejection has a higher *breakdown point*. The breakdown point of an estimator is defined as the largest proportion of the data that can be replaced by arbitrary values without causing the estimated value to become infinite [19]. The sample mean and standard deviation have breakdown points of zero, as one observation moved to infinity would make these estimators infinite. The sample median has a breakdown point of approximately 50%. The exact percentage depends on whether the number of data points is odd or even.

### 4.12 Applicability of Outlier Rejection Methods to VLSI Testing

The VLSI manufacturing flow consists of hundreds of process steps and thousands of variables that affect the circuit performance. The process parameters have inevitable variation. The variation in one or more process parameters leads to variation in performance. Since process variation is a physical phenomenon, parameters show a continuous distribution. Since the parameter distributions are continuous, predictable and can be modeled, any unfortunate statistical variation can be seen as an outlier. Therefore, in principle, outlier rejection methods are applicable to VLSI testing.

Successful application of any outlier rejection method depends on the following:

- Choice of outlier rejection method
- Knowledge of outlier-free distribution of the (fault-free) parameter
- Assumptions of the outlier rejection method (e.g. Normality requirement)
- Judicial selection of the threshold for the outlier rejection method

The knowledge of the fault-free parameter distribution requires an understanding of physical mechanisms that govern the variation. The channel length and width have a Normal distribution across a chip or a wafer and leakage current has an exponential relationship with channel length.

As a result, the fault-free leakage current distribution is expected to be lognormal in nature. Unfortunately, a distribution is invariably polluted by outliers – some of which are quite subtle (called *marginal* outliers [165]). This results in a long tail of the distribution[3]. This point is reemphasized using Fig. 47.

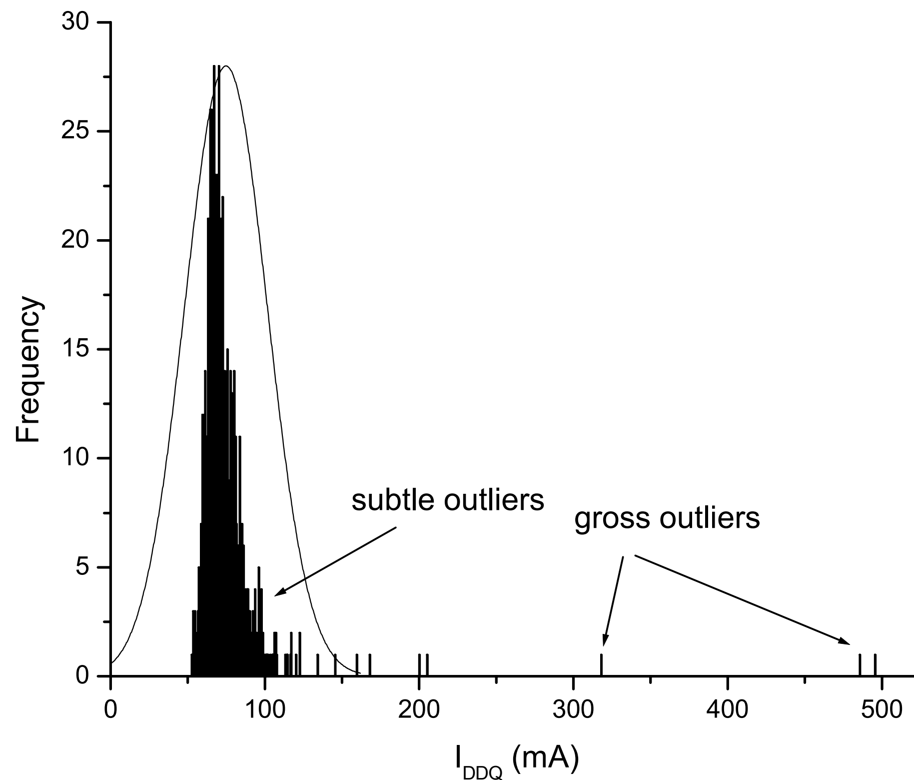


**Fig. 47. Outliers can be subtle or gross. Gross outliers hide the subtle outliers.**

Certain outlier rejection methods like Chauvenet's criterion and the Tukey method assume that the fault-free parameter distribution is Normal. Suitable goodness-of-fit tests need to be used

---

[3] It is incorrect to assume that a long-tailed distribution always contains outliers. Some outlier-free distributions do have a characteristic long tail (for example, the lognormal distribution).

to verify the Normality assumption. Such tests include the Chi-square test, Anderson-Darling test, Kolmogorov-Smirnov test and Shapiro-Wilk test [176]. When the distribution is not Normal, it is necessary to use a proper Normalizing transformation. In some instances, it may be quite hard or even impossible to find an appropriate transformation, thus making outlier rejection method inapplicable.

Deciding the appropriate thresholds for an outlier rejection method can prove to be very challenging. It has to be a combination of engineering judgment, statistical insight and acceptability criteria. If a certain threshold accepts many chips but results in a high BI fallout rate, the use of the outlier rejection threshold is not judicious. On the other hand, rejecting too many chips may not be economical. These conflicting goals must be balanced by examining what is an acceptable solution. However, this should not be interpreted as use of outlier rejection translates the threshold setting problem (illustrated in Fig. 5) from one domain ($I_{DDQ}$ pass/fail limit setting) to another (outlier threshold setting). Outlier rejection, of course, does not offer a panacea, but another means to screen defective chips. However, statistical techniques provide an effective way of screening outliers.

Since a single parameter variation may not be enough to cause a chip failure, it is necessary to consider statistical variation of multiple parameters and how their combination affects the chip performance. In the next section, we will discuss how correlation between various parameters can be exploited for screening outliers.

### 4.13 Measuring Outlier Screening Effectiveness

The main motivation behind outlier screening is to screen likely defective chips. Like any other test method, there are two types of errors associated with it: (1) Type-I errors in which good chips are rejected unintentionally (region ‘B’ in Fig. 17(b)), and (2) Type-II errors in which bad chips get accepted unintentionally (region ‘A’ in Fig. 17(b) [177]. Both these errors are quantified in terms of conditional probabilities as follows:

Type–I error probability = P (chip is rejected | chip is fault-free) (4.8)

Type–II error probability = P (chip is accepted | chip is defective) (4.9)

Both these errors are costly and, therefore, to quantify the effectiveness of an outlier screening method two metrics must be used. The first one, *defect level* (DL), measures how many defective chips were misclassified by the method as good chips (Type-II error). It is a measure of outgoing quality. The second metric, *yield loss* (YL), measures how many good chips

were scrapped by the method as bad chips (Type-II error) [178]. It measures revenue loss due to incorrect classification. We will measure defect level as percentage of bad chips from all accepted chips and yield loss as percentage of good chips from all chips. We will not compute exact probabilities in this work.

As described earlier, defining a “good” or a “bad” chip is difficult for parametric tests. One oracle to define this is to consider post BI results. A further qualification that may be used is to consider the nominal parameter value distribution and set $\pm 3\sigma$ limits. Throughout this dissertation, the primary thesis that is followed can be summarized as follows: *if the variation cannot be explained by any fault-free mechanisms, it must be due to a defect.*

In the remainder of this dissertation, we use empirical data to support this argument.

# 5. MULTIPLE PARAMETER CORRELATION

## 5.1 What Is Parameter Correlation?

Correlation is a statistical measure of the tendency of two or more parameters to vary together. If the two parameters vary in the same direction (when one increases, the other one also increases), they are said to be *positively correlated* (Fig. 48(a)). On the other hand, if they change in the opposite direction, they are said to be *negatively correlated* (Fig. 48(b)). If the change in one parameter is unrelated to a change in the other parameter, they are said to be *uncorrelated* (Fig. 48(c)). When two parameters are correlated, one may be the function of the other through a third (independent) parameter. It is necessary to understand the relationship between two correlated parameters to use it for multi-parameter-based outlier screening. For example, leakage current is a function of channel length (negative correlation). The channel length also modulates the transistor delay (positive correlation). Therefore, leakage current and transistor delay are negatively correlated. Since delay and switching speed are inversely proportional, leakage current and chip speed are positively correlated. This relationship is exploited for outlier identification later in this section.

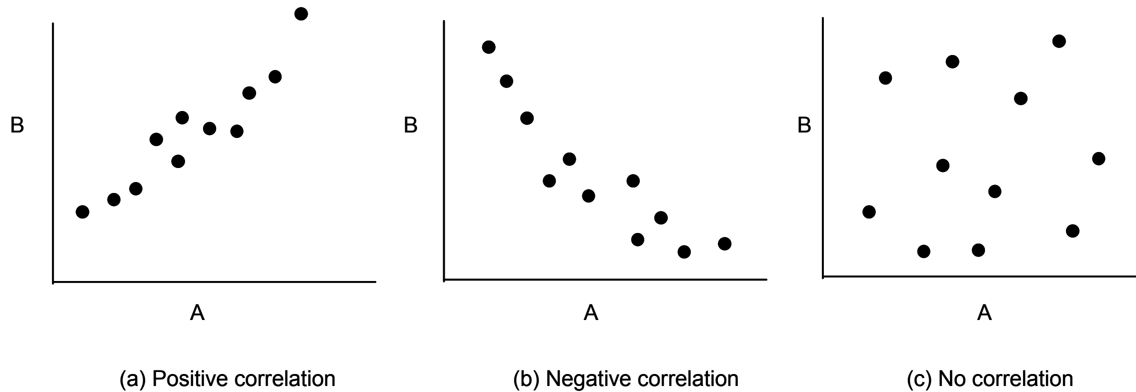


**Fig. 48. Illustrations for positive (a), negative (b) and no (c) correlation.**

## 5.2 Why Correlate Multiple Parameters?

A chip is an outlier if any parameter shows suspicious variation (variation that is outside the predetermined bound) or a combination of two or more parameters seems unlikely for a fault-free chip. Due to interactive effects, two or more parameters may have a statistical variance that leads to failure. For example, larger channel length and higher interconnect impedance can lead

to delay failure, although individual variation may not cause delay failure. It is important to screen chips that show such behavior earlier in the test cycle. Since single parameter test data analysis cannot screen such chips, it is necessary to explore a combination of multiple test parameters. In the remainder of this section, we explore various methods to achieve this.

It is understood that the parameters that are correlated must be governed by similar underlying physical process mechanisms. In Section 3, it was mentioned that $F_{max}$ can be used to set adjustable limits for $I_{DDQ}$. This is possible because both $F_{max}$ and $I_{DDQ}$ vary due to the same underlying factors. The main idea behind using multiple parameters stems from the fact that parameters have different process sensitivities. This is useful for deciding whether "marginal" outliers are indeed defective and improves confidence in outlier rejection.

## 5.3 Correlating Multiple Factors

### 5.3.1 Wafer-level spatial correlation

Fig. 49 shows a gray scale map for a wafer using $I_{DDQ}$ readings for a vector. The chips that failed functional tests or were gross outliers are marked with dots. Two observations can be made from this figure. The variation in $I_{DDQ}$ across a wafer is smooth and neighboring chips have similar $I_{DDQ}$ values. Secondly, defects clusters are noticeable. Analysis of wafer-level patterns is shown to be useful for estimating fault-free parameters more accurately [179].

The fact that neighboring chips on a wafer undergo similar changes in the process parameters and have similar fault-free parameters can be exploited to estimate fault-free $I_{DDQ}$ and screen outlier chips. Apart from test parameters measured off chip, test structures embedded in the scribe line can be used to analyze spatial dependence [180]. For large chips manufactured using modern semiconductor processes, wafer-level die coordinate information is usually available through electronic die identifiers. Hence it is easier to trace a die through the test flow and to reject a die before packaging. Since this analysis can be done at the wafer-level, packaging cost of defective (outlier) parts is saved, thus reducing overall test cost.

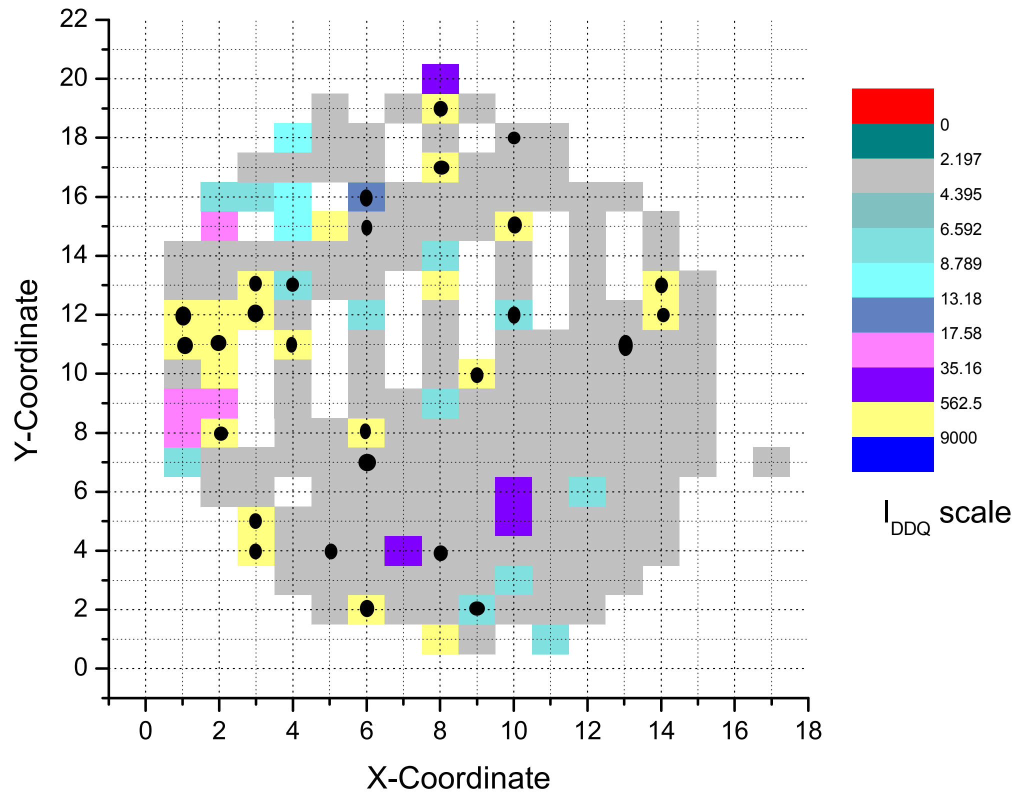


**Fig. 49. Two dimensional grayscale map for $I_{DDQ}$ across a wafer.**

### 5.3.2 Spatial fit method

In this method, linear regression [181] is carried out for die XY-coordinates and $I_{DDQ}$ as the Z-coordinate. This is equivalent to finding the equation of the best-fitting plane with die XY coordinates and $I_{DDQ}$ data for the adjacent die. In a simple approach, eight neighboring dice[4] (see Fig. 50) are used for estimation. However, dice at longer distances may be used if sufficient

---

[4] This definition will be used throughout the remainder of this dissertation unless mentioned otherwise.

neighboring die information is not available. The residual is obtained by subtracting actual $I_{DDQ}$ from the predicted $I_{DDQ}$ (plane estimate). This procedure is repeated for each vector and the distribution of residuals is used to set threshold limits. Before performing linear regression according to the procedure described in [182], it is necessary to reject gross outliers to avoid overestimation of $I_{DDQ}$. Gross outliers can be rejected using any suitable statistical outlier rejection method like Chauvenet's criterion [174] or the Tukey method [175]. Of course, the challenge lies in using appropriate thresholds for outlier screening methods.

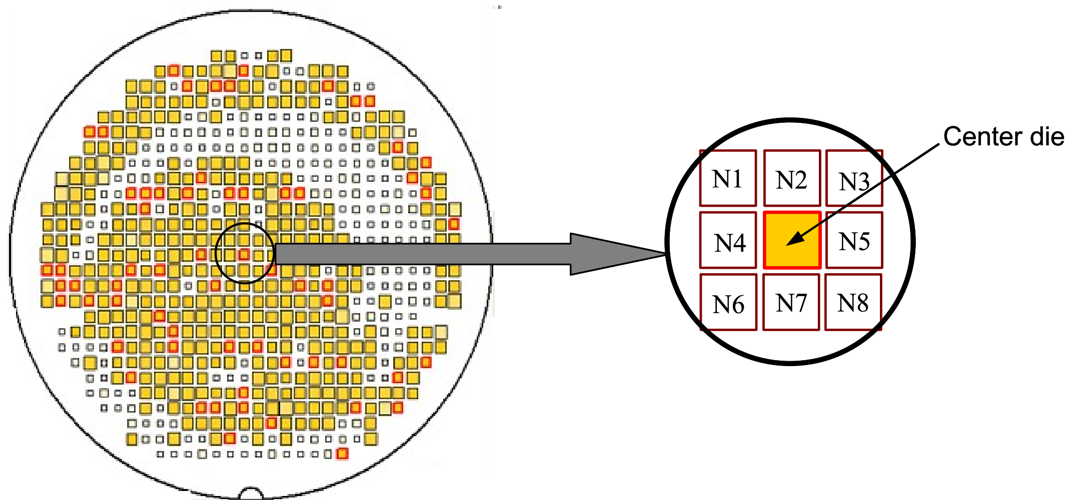


**Fig. 50. Neighborhood die definition.**

Fig. 51 shows a wafer surface plot for $I_{DDQ}$ values for a vector across a wafer. It shows that except for a few spatial outliers, chips have similar $I_{DDQ}$ values. Fig. 52 shows surface plot for $I_{DDQ}$ values estimated using spatial fit method. The residual values were computed by subtracting estimated values from the actual values. The wafer surface plot of the residual values is shown in Fig. 53. Chips having positive residual values above the plane are defective chips. Some chips (spatial dips) have negative residual values. In some cases, the overestimation could be a result of outliers in the neighborhood. These chips may be discarded for reliability concerns due to defect clustering.

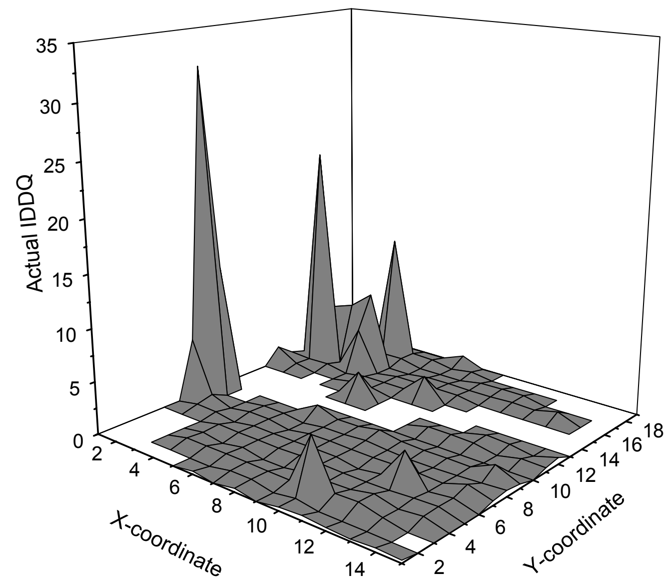


**Fig. 51. Wafer surface plot for a single $I_{DDQ}$ vector.**

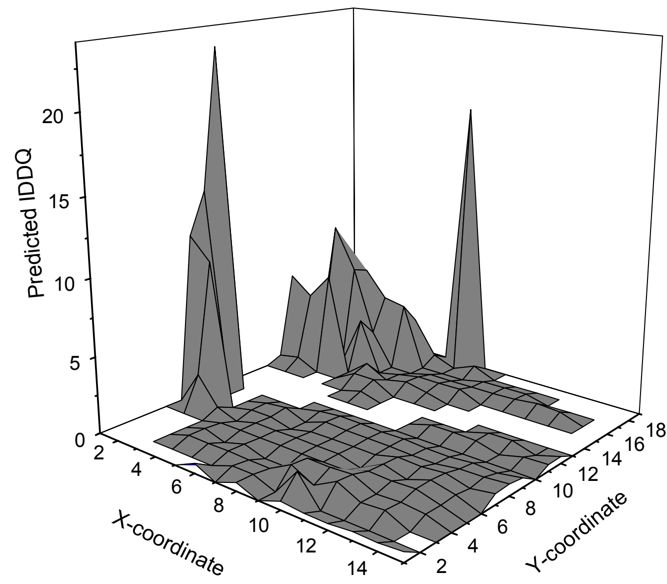


**Fig. 52. Wafer surface plot for estimated $I_{DDQ}$ values.**

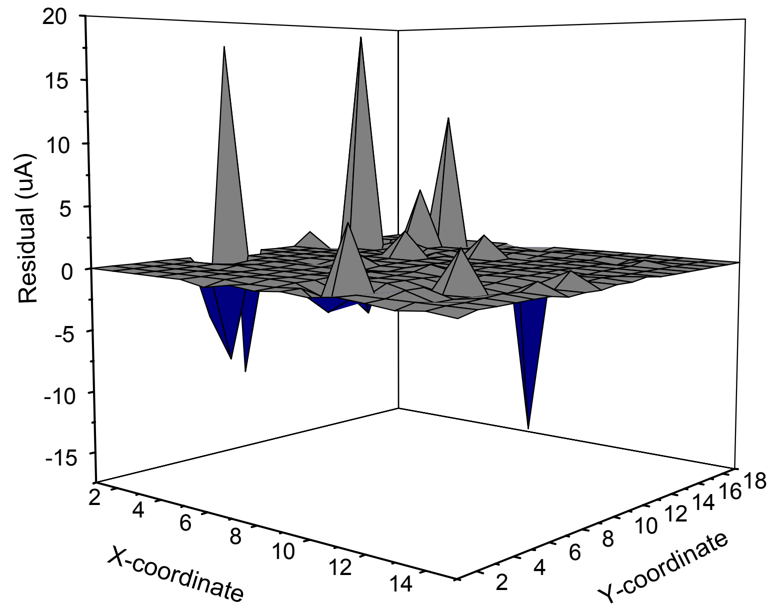


**Fig. 53. Wafer surface plot for residual values for a vector.**

5.3.3 Nearest Neighbor Residual (NNR) method

Another method to exploit wafer-level correlation was suggested in [183]. In this method, called *location averaging*, the adjacent dice $I_{DDQ}$ is used for estimating $I_{DDQ}$ of the center die. However, instead of linear regression as in the spatial fit method, the median $I_{DDQ}$ is used as an estimate of $I_{DDQ}$ [184]. This gives less weight to outliers in the vicinity. The dice at longer distances are considered when data for immediate neighbors is not available. It is important to ensure those dice used for estimation are correlated to the die whose parameters are being estimated. Studies of wafer-level patterns in the data can be useful [185]. In a similar study for predicting parametric yield, it was found that the dice that correlate with die on the wafer-edge are other die on the wafer-edge either from the same wafer (diametrically opposite) or from different wafers in the same lot or batch at the same XY locations [186].

5.3.4 Neighbor Current Ratio (NCR)

The concept of Current Ratios described in Section 3 relies on the assumption that intra-die

variance in fault-free $I_{DDQ}$ is deterministic [126]. This assumption fades away for DSM technologies. Although the current ratio method does hold promise for present and future technologies, determining the appropriate pass/fail threshold for current ratios is not trivial. Fig. 54 shows the variation in the minimum and maximum $I_{DDQ}$ for SEMATECH chips [52] that passed all tests or failed only $I_{DDQ}$ test (5 μA threshold) at the wafer level. Fig. 55 shows a similar spread in the minimum and the maximum $I_{DDQ}$ values for 949,753 chips from 1342 different wafers across different lots for LSI chips. The ratio of ordinate to abscissa (slope) gives the CR.

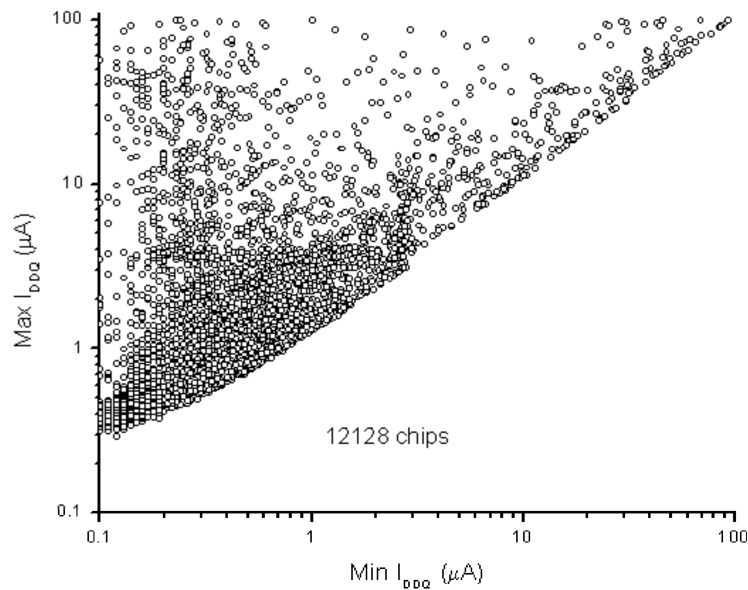


**Fig. 54. Variation in minimum and maximum $I_{DDQ}$ for SEMATECH chips.**

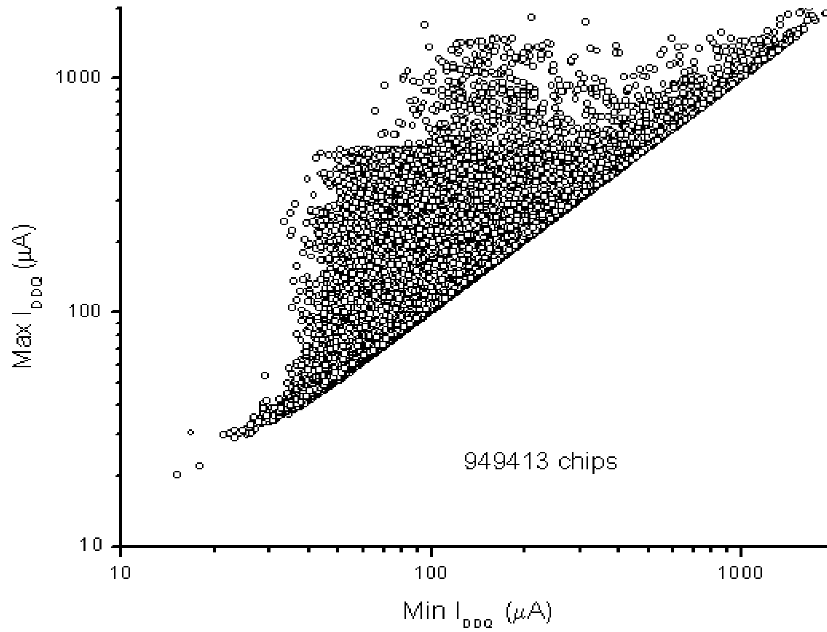


**Fig. 55. Variation in minimum and maximum $I_{DDQ}$ for LSI chips.**

It can be clearly seen that deciding a "fault-free" CR threshold is not easy. Too small a threshold would result in tremendous yield loss. Notice that many chips cannot be conclusively termed outliers. Hence determining an appropriate CR threshold is difficult. Some of the chips having nominal CR are "spatial outliers" – chips exhibiting much higher current than their immediate neighbors on the wafer. Such chips are likely to fail (reliability risk) even if they pass all Boolean tests.

The NCR is defined as the ratio of leakage current of a chip to the leakage current of a neighboring chip for the same vector [187], [188]. NCR can be computed for each vector pair for the center die and all its neighbors. Mathematically, NCR is defined as follows:

$$NCR_{ji} = \{\frac{I_{ci}}{I_{ji}}\} \quad \forall i \, and \, 1 \le j \le 8 \tag{5.1}$$

where $I_{ci}$ is the $I_{DDQ}$ of the center die for the $i^{th}$ vector and $I_{ji}$ is the $I_{DDQ}$ of the $j^{th}$ neighboring die for the $i^{th}$ vector.

Since neighboring chips have similar fault-free leakage currents for the same vector, the

nominal value of NCR is 1. Of course, owing to process variations NCR values for a chip vary, generally having a Normal distribution with the mean value of one (assuming fault-free neighbors and smooth wafer-level process variations). The maximum of all NCR values is used for screening chips since it is most sensitive to defects.

**TABLE VI. VARIOUS POSSIBILITIES AS IMPLIED BY NCR.**

**(DIE 'A' $I_{DDQ}$/DIE 'B' $I_{DDQ}$).**

| Die A | Die B | NCR |
|---|---|---|
| Fault-free | Fault-free | ~1 |
| Fault-free | Passive Defect | <1 |
| Fault-free | Active Defect | <1 |
| Passive Defect | Fault-free | >1 |
| Active Defect | Fault-free | >1 |
| Active Defect | Passive Defect | Depends on the nature of defect |
| Passive Defect | Active Defect | |
| Passive Defect | Passive Defect | |
| Active Defect | Active Defect | |

The nominal value of the maximum NCR is more than one since at least one vector on a neighbor will have a lower value for most chips. Table VI lists different defect possibilities as implicated by NCR values. A passive defect such as a $V_{DD}$-GND short has elevated $I_{DDQ}$ for all vectors. Active defects produce elevated $I_{DDQ}$ for some but not all vectors. When both the chips have passive defects, NCR depends on the relative magnitudes of defect currents for the two chips.

NCR has the capability to screen chips with passive defects. Since CR depends on the relative magnitudes of defect current and background (fault-free) leakage, NCR is more advantageous than CR. Fig. 56 and Fig. 57 show variation in CR and NCR values across a wafer. Several spatial outliers are clearly visible in the NCR plot that are not seen in the CR plot the illustrating advantage of NCR over CR.

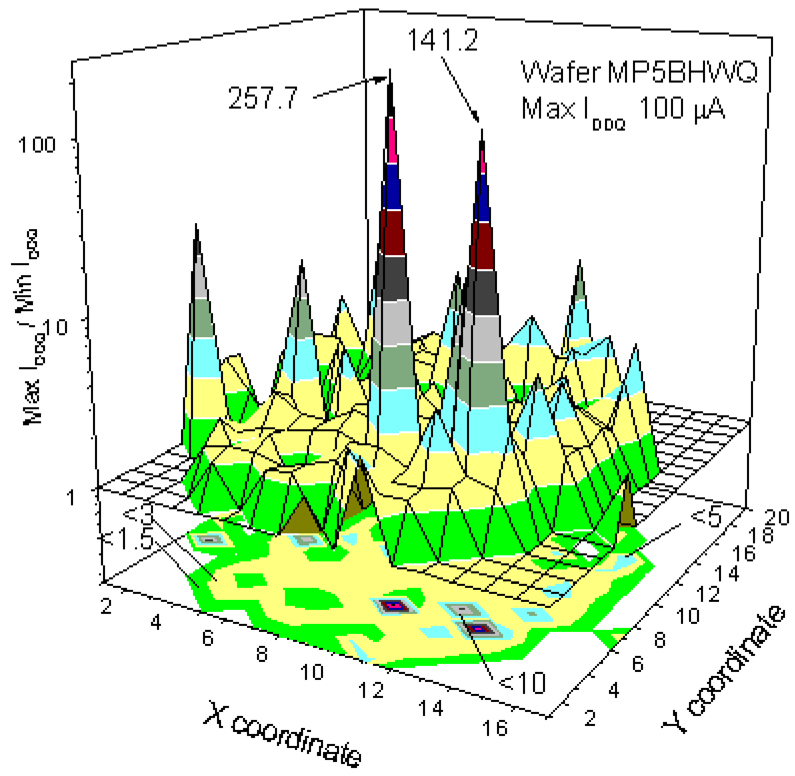


**Fig. 56. Variation in CRs across a wafer.**

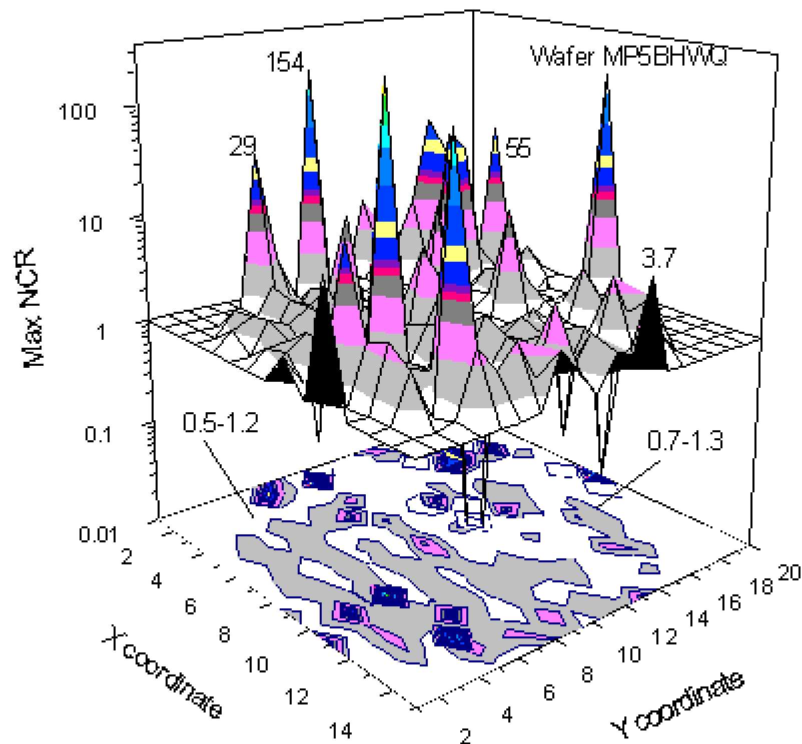


**Fig. 57. Variation in NCRs across a wafer.**

5.3.5 Immediate Neighbor Difference $I_{DDQ}$ Test (INDIT)

A variation of the delta-$I_{DDQ}$ method by considering adjacent dice was investigated in [189]. In this method, differences (deltas) between $I_{DDQ}$ of a die and that of its adjacent dice are used for screening instead of deltas of a chip for different vectors (*self deltas*) as in conventional delta-$I_{DDQ}$ method [119]. Since fault-free chips have similar $I_{DDQ}$ values, differences would be close to zero. These deltas are referred to as *neighbor deltas*. The maximum of neighbor deltas is used for screening outlier chips. The thresholds can be set by using wafer or lot-level distributions of neighbor deltas. The INDIT algorithm is shown in Fig. 58. INDIT essentially combines wafer-level information with $I_{DDQ}$ data for identifying outliers. The addition of this information provides an advantage over the conventional delta-$I_{DDQ}$ method. This can improve the defect level and/or overkill without compromising the yield. INDIT has similar or better capability to screen chips with passive defects similar to NCR. Fig. 59 shows wafer level variation in maximum self-deltas. The wafer level variation in maximum neighbor deltas for the same wafer is shown in Fig. 60. It can be observed that several outlier chips that are hidden in

Fig. 59 become visible in Fig. 60 when neighbor deltas are considered. These chips likely contain passive defects that cannot be detected by self-delta alone.

Screen all dice that fail Boolean test
For each remaining dice on a wafer
Find number of immediate neighbors ($I_n$) available
If ($I_n$>0){
For each neighbor $N_i \in \{N_1..N_8\}$
For each vector *j*
Nbr Delta(*j*) = Center die $I_{DDQ}(j)$- $N_i$ $I_{DDQ}$ (*j*)
Find maximum neighbor-delta $N_i$(max) for $N_i$
Find maximum of all neighbor-deltas
$\delta_{max}$=max(Ni(max)), $N_i \in$ {N1..N8}
}
If $\delta_{max}$ > threshold, reject the die

**Fig. 58. INDIT algorithm.**

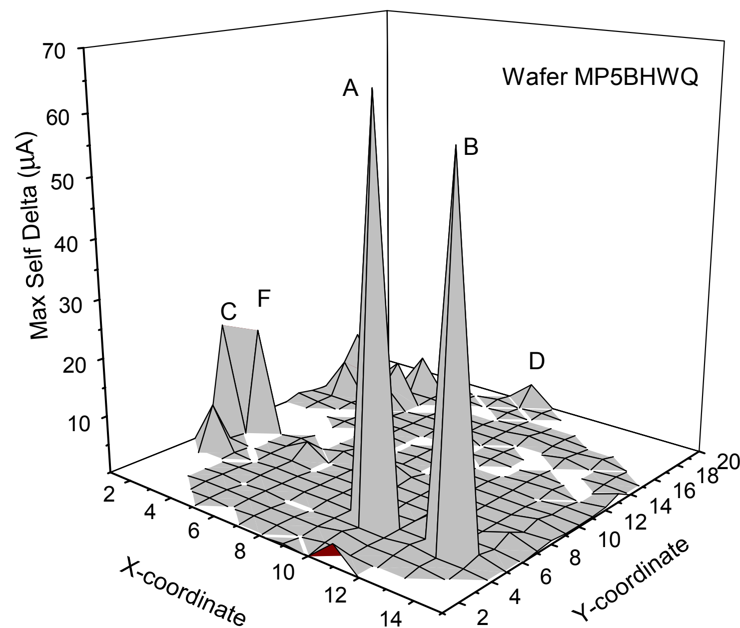


**Fig. 59. Wafer surface plot of maximum self-deltas for a wafer.**

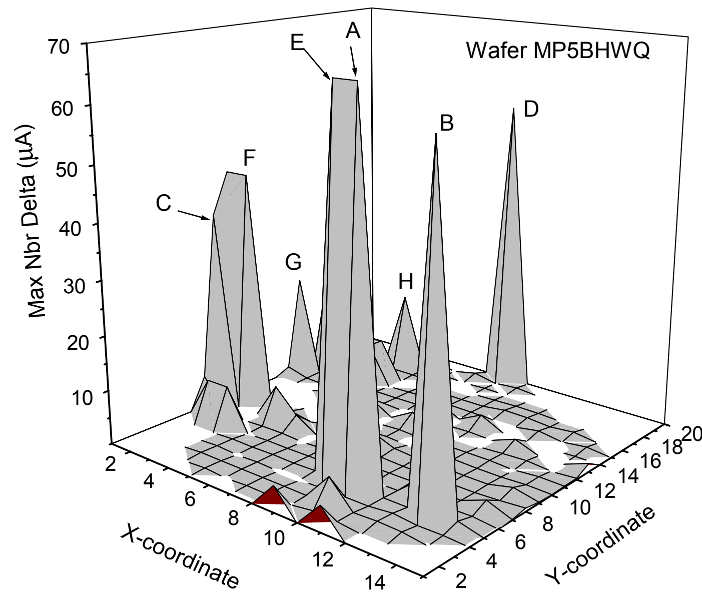


**Fig. 60. Wafer surface plot of maximum neighbor-deltas for a wafer.**

### 5.3.6 Wafer signature analysis

A wafer signature is obtained by sorting $I_{DDQ}$ readings of all dice on a wafer for identical vectors. This is similar to a current signature [112], with the exception that readings from different chips make a wafer signature. Thus, a wafer signature reveals inter-die variance. The jumps or steps in a wafer signature can exist either due to outliers or due to missing data [190]. The step size, number of readings before/after the step, and the magnitude of the current after the step can be used to identify whether a jump is due to missing data or a defect. The wafer signature approach can differentiate chips with passive defect as they consistently appear in the tail of wafer signatures. The advantage of wafer signature is that it uses the same test data and does not require any additional measurements. Moreover, even with few $I_{DDQ}$ measurements wafer signatures can screen chips with passive defects that are not easily screened by current signature [112] or delta-$I_{DDQ}$ methods [118].

Fig. 61 shows wafer signatures for two wafers for two $I_{DDQ}$ readings. Note that the numbers on the abscissa do not correspond to the same dice due to sorting. The jumps in the signatures also occur at different locations, thus identifying possibly different outliers. A similar wafer

signature for a wafer from LSI data is shown in Fig. 62. It shows a smooth *head* (beginning part of the signature) and some breaks. Relatively bigger breaks in the *tail* of the signature indicate wafer-level outliers for this vector. As LSI data comes from state-of-the-art technology, it also shows the applicability of wafer signature to DSM technology.

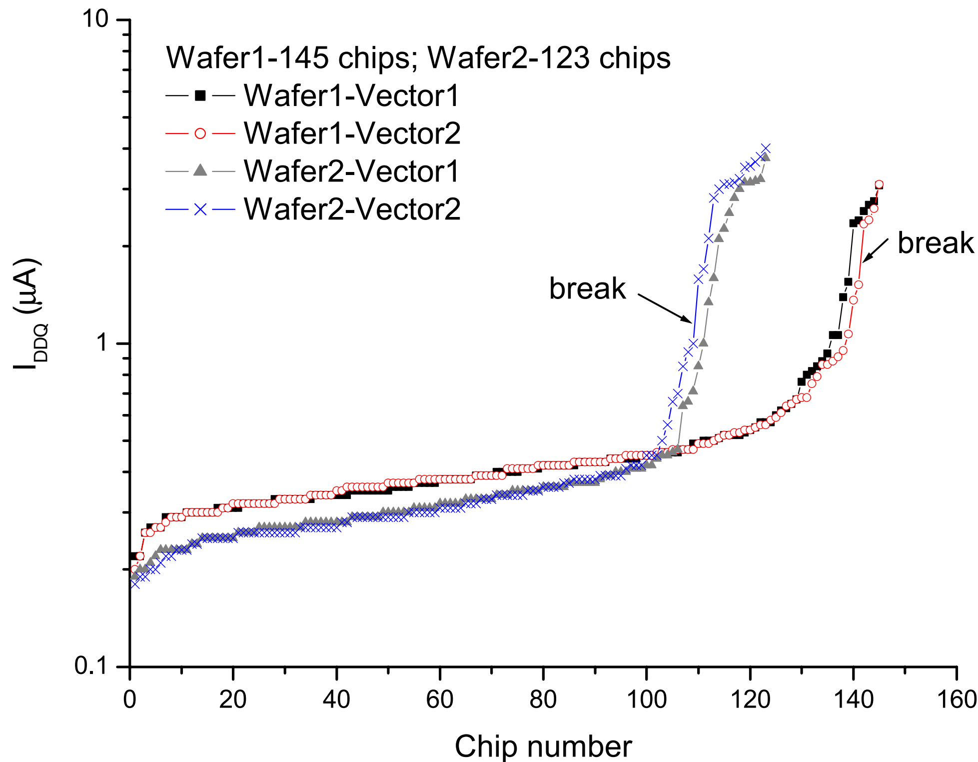


**Fig. 61. Wafer signatures for two wafers for two vectors.**

Of course, the resolution of the wafer signature approach depends on the number of dice on a wafer and on the number of $I_{DDQ}$ readings. If a wafer shows systematic, pattern (e.g. stepper field), the wafer signature approach can result in higher yield loss unless the wafer pattern information is used for threshold setting.

5.3.7 Threshold setting for wafer signatures

It can be argued that the wafer signature approach combines best of delta $I_{DDQ}$ [118] and current signature [112] methods and at the same time is capable of screening passive defects. However, threshold setting for wafer signature can be challenging. The simplest way to set a pass/fail limit is to select a single threshold value for delta. The threshold could be derived

through empirical analysis. However, this method does not account for wafer-to-wafer variation in $I_{DDQ}$ and can result in significant yield loss for some wafers. It is clear that the delta threshold must be adjusted to account for wafer and lot level variations in $I_{DDQ}$. Different possibilities exist to achieve this. One way is to set a threshold for each wafer by observing deltas across all signatures. Alternatively, a different threshold can be used for each vector. This threshold should be set for each wafer. Data from several wafers may be combined to decide the nominal delta size. Assuming a pass/fail decision is required immediately after wafer probe, wafer-level analysis is used. Since fault-free and faulty delta values are not known a priori, we try to maximize the number of points in a wafer signature. Therefore, to minimize the number of "dummy" breaks that occur due to missing data, all chips that do not fail functional tests should be used to obtain a wafer signature.

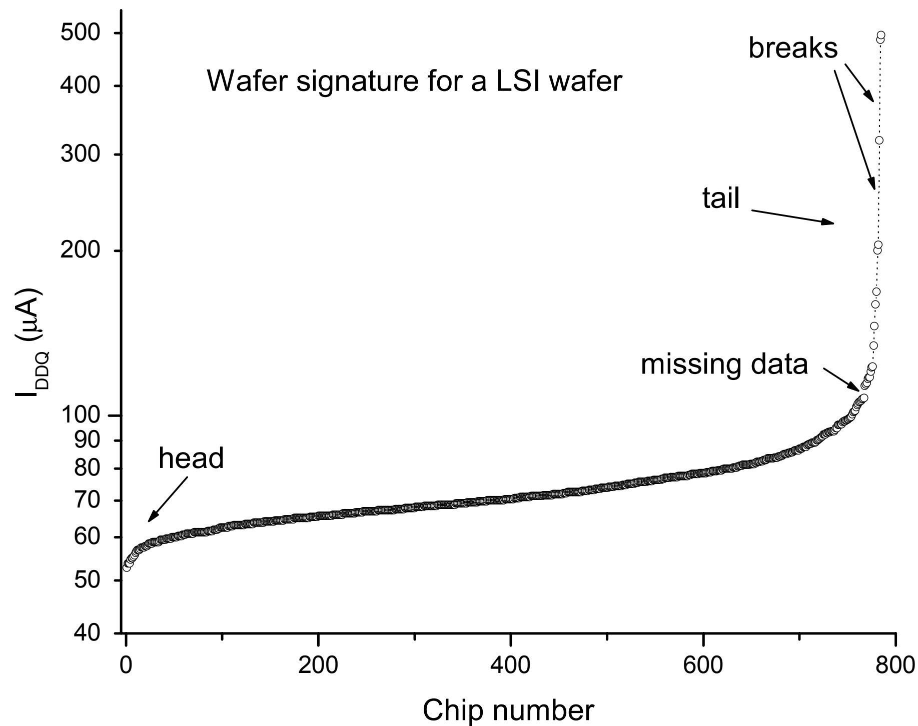


**Fig. 62. Wafer signature for a wafer from LSI data.**

## 5.4 Multiple Test Parameter Correlation

All outliers cannot be identified by exploiting wafer-level spatial correlation alone. Due to defect clustering [191] some outliers are hidden and cannot be distinguished, especially if a spatial fit-like method (that gives equal weight to data) is used. Correlating multiple test parameters can glean valuable information that can be used for identifying outliers.

There exists a correlation between $I_{DDQ}$ and delay failures for certain bridging defects [192]. This fact can be utilized by correlating $I_{DDQ}$ and delay data for outlier identification. The chip delay information can be obtained from the test structures embedded in the chip or the *kerf* (the area between chips on the wafer). The correlation between the transistor saturation current, $I_{dsat}$, and the flush delay (time required to propagate a transition through a LSSD scan chain with all scan clocks turned ON) has been used for system performance estimation [193]. The correlation between critical and near critical path delays is useful for performance prediction of an IC [194] and test cost reduction [195].

### 5.4.1 $I_{dsat}$ and $F_{max}$ correlation

The maximum frequency a chip is capable of operating at ($F_{max}$) depends on the critical path in the circuit. Generally due to process variation, $F_{max}$ shows a Normal distribution [38]. Since the transistors are affected by the underlying process variation, the switching speed and leakage current are inversely related. The correlation between saturation current, $I_{dsat}$, and $F_{max}$, has been shown to be useful to identify outliers [196].

### 5.4.2 $I_{DDQ}$ and flush delay correlation

Flush delay is obtained by turning all scan clocks simultaneously thus making the scan chain like a long wire with buffers and inverters [194]. The flush delay is the time it takes for a rising or falling transition to traverse the entire chain. For a full scan design the scan chain practically traverses across the entire chip. Every chip has effective channel lengths of transistors normally distributed around the mean value [197]. Thus flush delay measurement is indicative of process variations. By knowing chip-to-chip and wafer-to-wafer process drifts, it is possible to refine the estimate of fault-free $I_{DDQ}$.

Fig. 63 shows a scatter plot of $I_{DDQ}$ and flush delay values for SEMATECH chips. A reasonable correlation is observed because smaller transistors have higher leakage and switch faster. Similar correlation is also observed in test structure (ring oscillator) frequency and leakage currents for LSI data as shown in Fig. 64.

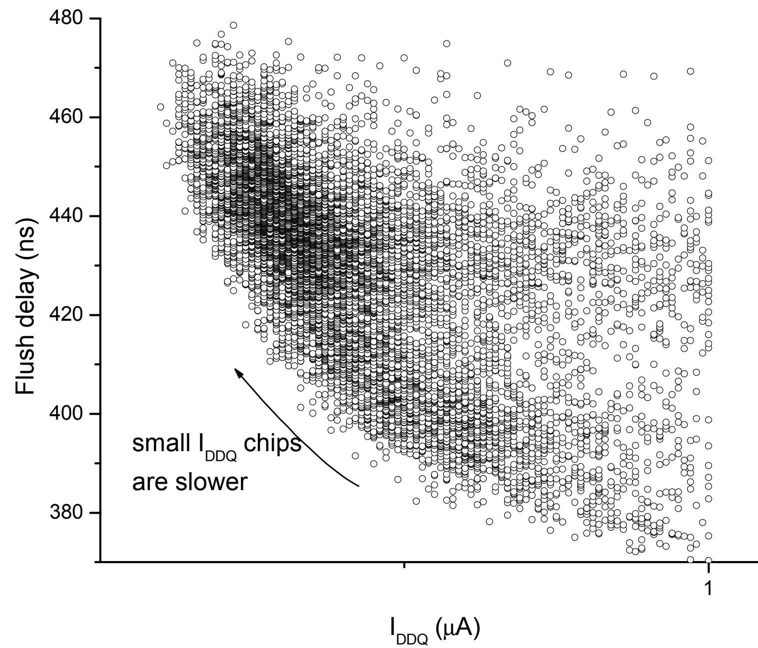


**Fig. 63. Flush delay and $I_{DDQ}$ correlation for SEMATECH data ($R^2$ = 0.40).**

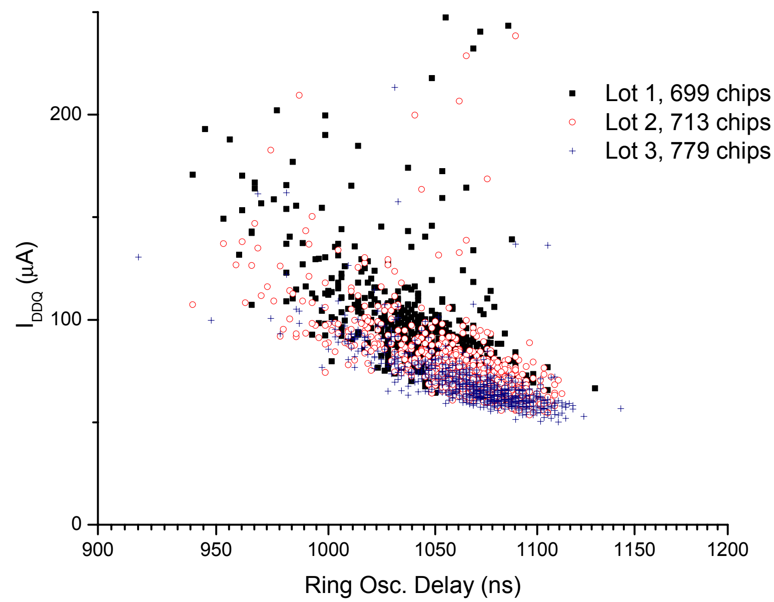


**Fig. 64. Correlation between ring oscillator frequency and $I_{DDQ}$ for LSI data ($R^2$ = 0.48).**

The correlation between the static leakage current ($I_{DDQ}$) and the flush delay of a chip is shown to be useful for estimating the leakage current [136] to avoid excessive yield loss. In this method, a threshold is altered to account for fast and leaky chips. However, flush delay tends to average out the effect of process variations across a chip. As the number of transistors per chip increases, the correlation between these two factors is not strong enough to improve the estimate of the fault-free $I_{DDQ}$ obtained by other methods. One possible solution is to use test structures embedded in the chip. These test structures measure local variation accurately because of their proximity to the circuit.

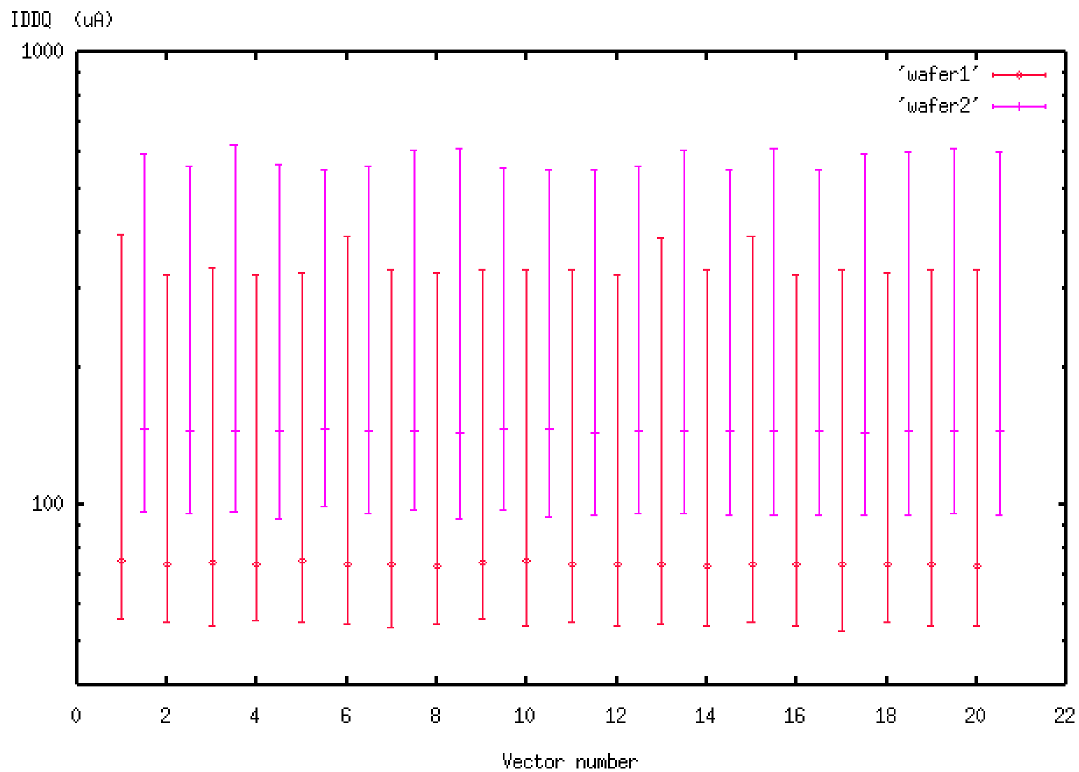


**Fig. 65. Leakage current variation for 20 vectors for two wafers.**

The variation in leakage current can be decomposed by considering lot-to-lot and vector-to-vector variations differently. Fig. 65 shows $I_{DDQ}$ variations for 20 vectors for all chips from two wafers each having more than 700 chips. A systematic shift between two wafers is noticeable. The lot-to-lot variations can be accounted for by adjusting pass/fail thresholds using delay data. An example in Fig. 66 shows a systematic shift in $I_{DDQ}$ between two lots for the first two vectors

in Fig. 65. Notice that variation within different vectors for a lot is smaller than that between two lots. Correlation between delay with $I_{DDQ}$ can be used for adaptive threshold setting in order to reduce yield loss and test escapes.

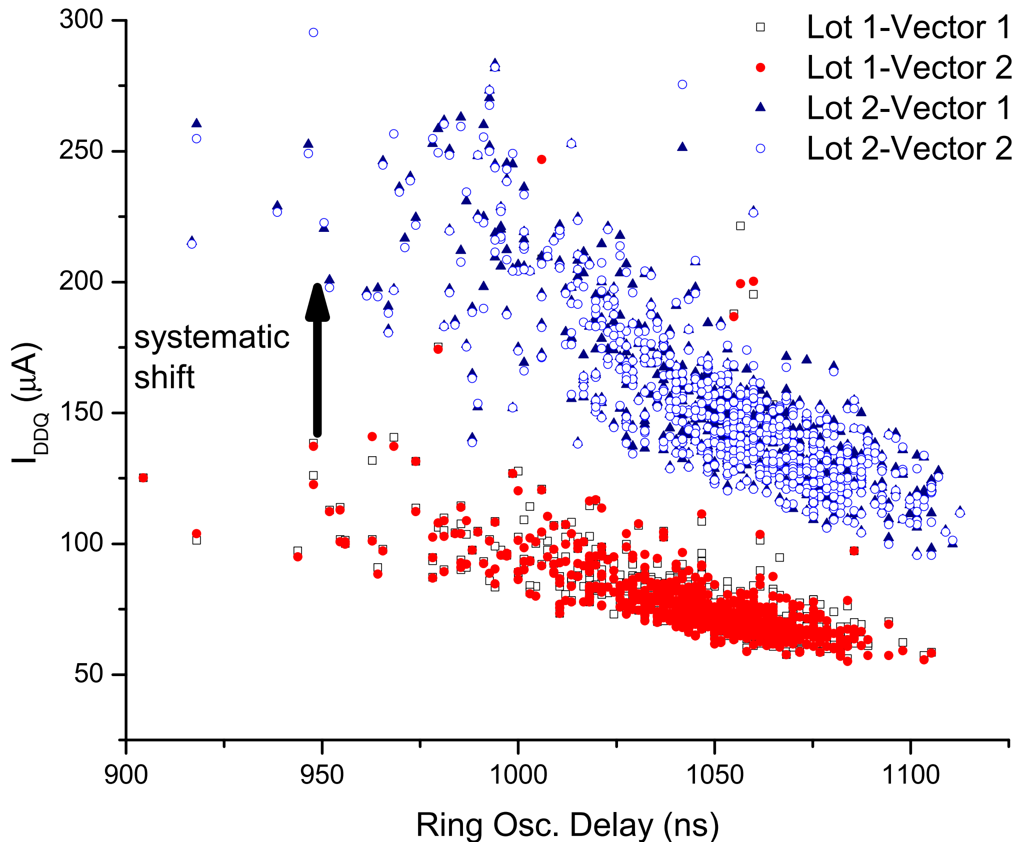


**Fig. 66. Systematic shift in $I_{DDQ}$ across lots.**

5.4.3 Radial shift in $I_{DDQ}$

Wafers show systematic variation in parameters like $L_{eff}$, poly thickness, etc. This results in alteration of their fault-free parameters and seemingly outlier behavior. For example, consider the variation of $I_{DDQ}$ with wafer radius as shown in Fig. 67. For the vectors shown in Fig. 66, it shows that there is a systematic shift in leakage current within a wafer with die position. Some obvious outliers on the wafer edge are noticeable. However, setting a single threshold for any

wafer will result in unjustifiable yield loss. An adaptive test strategy can consider die position on a wafer and use different pass/fail criterion.

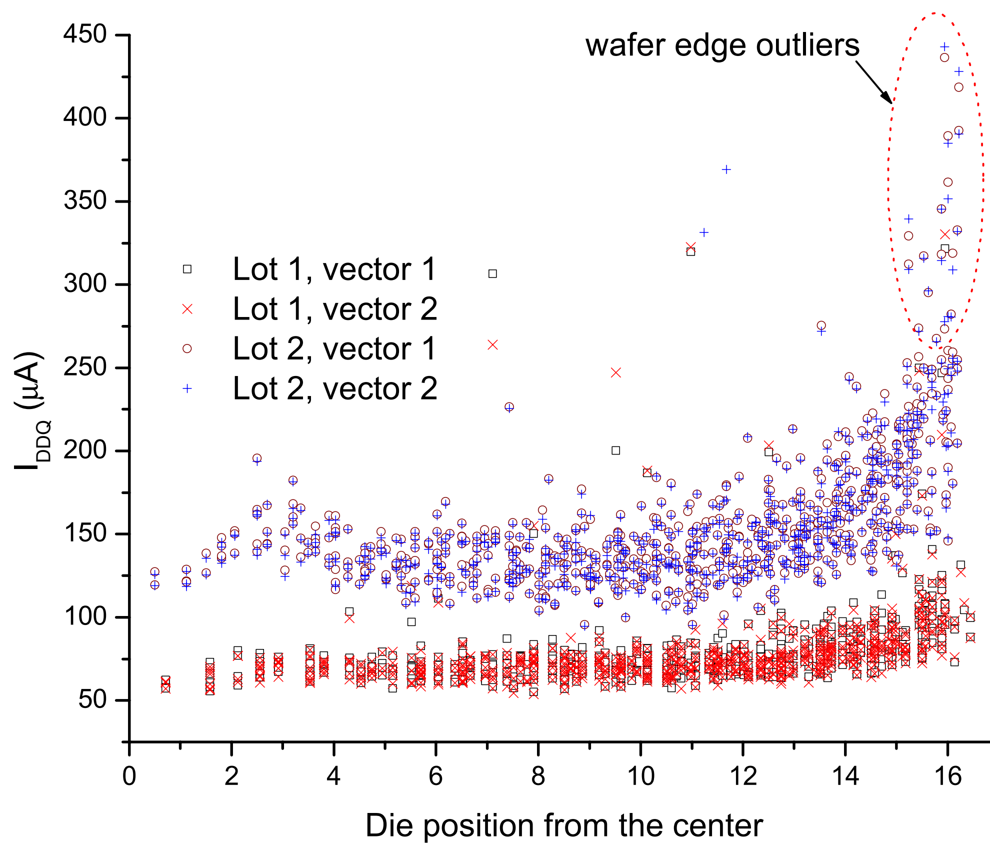


**Fig. 67. Radial variation in $I_{DDQ}$.**

5.4.4 $I_{DDQ}$, $F_{max}$ and RBB

Outlier identification can be improved by adding additional test parameters. In one example, Keshavarzi et al. [97] showed that the shift in the plots of $I_{DDQ}$ and frequency due to the application of reverse body bias (RBB) was useful for identifying certain outliers as shown in Fig. 68 (courtesy Ali Keshavarzi). The application of RBB results in reduction in $I_{DDQ}$ by more than an order of magnitude. However, collection of data for RBB as used in [196] is time consuming and, therefore, expensive. In addition, it may not be cost effective to implement RBB.

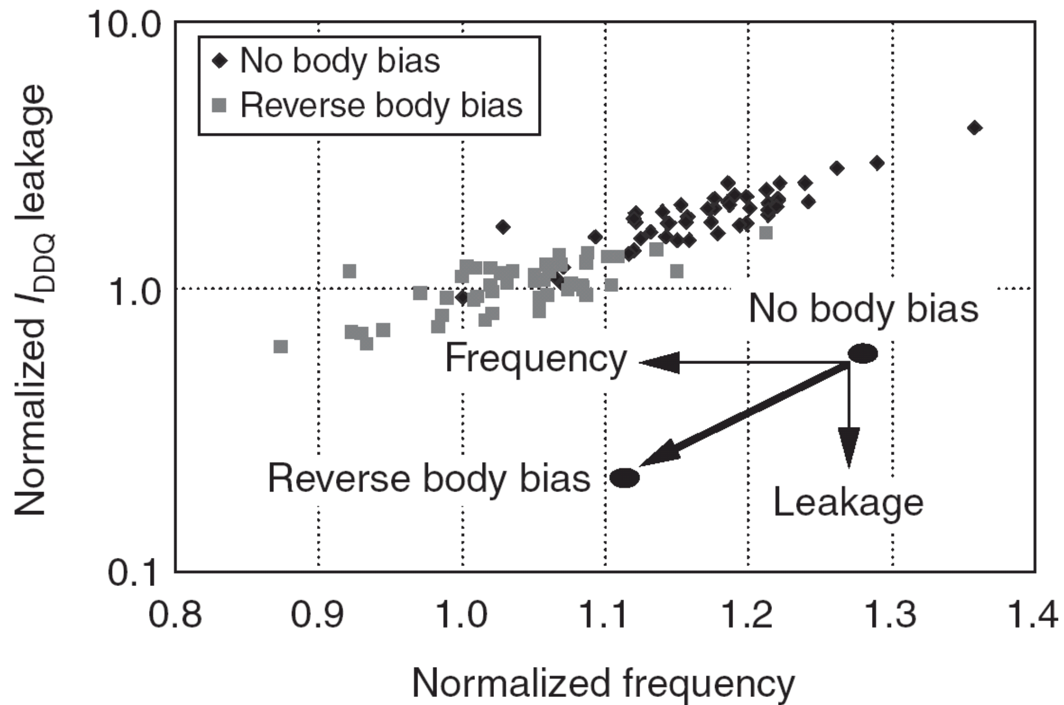


**Fig. 68. IC leakage with and without RBB of 0.5 V at room temperature.**

### 5.5 Use of Multiple Test Metrics

Generating new test metrics with the existing test data is often cost effective. The approaches suggested earlier like wafer signature, NCR and INDIT used the same theme. In the following subsections we discuss use of multiple test metrics for identifying outlier chips.

### 5.6 CR-NCR Combination

The Current Ratio method [126] exploits the regularity in fault-free leakage current that causes within-die variation to screen outliers. However, as intra-die and inter-die $I_{DDQ}$ variances increase, the distinction between faulty and fault-free chips is difficult. One possible solution is to combine multiple test metrics like CR and NCR [198]. As shown in Fig. 69, the combination of CR and NCR can be used to divide chips into five regions. Chips having very small CR values (close to 1) are shown as 'passive defects'. Chips having CR and NCR smaller than the respective thresholds are denoted as 'Good chips' in region D. Of course, some of these chips

will contain faults that do not cause elevated $I_{DDQ}$. Certain subtle active defects cause higher NCR even though their CR may not increase. Such chips fall in region A. In case of a strong active defect, both CR and NCR values are high. Such chips fall in region B. As mentioned earlier, NCR may not be very effective for screening chips that are surrounded by many defective chips. However, it is unlikely that all the neighboring chips have similar defective currents for all input vectors. Outliers in a bad neighborhood and in a fast wafer region would appear in region C. Previous studies have shown such chips to be a reliability risk [199].

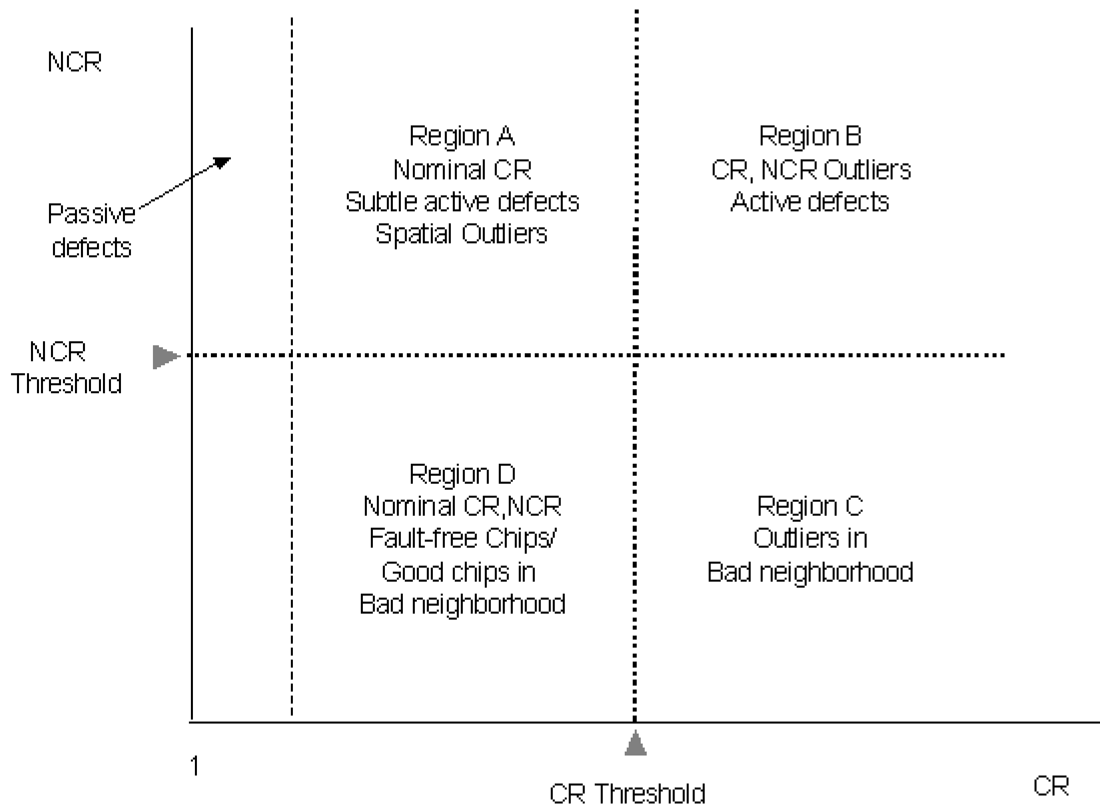


**Fig. 69. Categorization of chips based on CR and NCR.**

## 5.7 Combination of CR, NCR and Delay Data

When both metrics agree on fault-free or faulty behavior of a chip, it is easy to accept or reject it. However, a test engineer is in a dilemma when they disagree. In this situation, different approaches can be taken as indicated in Fig. 70. It is possible to identify the reasons for *seemingly* outlier behavior by considering multiple test metrics. We illustrate this point

conceptually by considering flush delay data in addition to CR and NCR [200]. For a combination of CR, NCR and delay data, one can construct Table VII to discriminate between fast chips and defective chips. A typical analysis flow for outlier screening using CR, NCR and delay data is as shown in Fig. 70. Note that this flow is generic and a variety of combinations and sequences are possible. As expected, addition of a test parameter is helpful in identifying some outliers that are not detected previously. Fig. 71 shows a scatter plot for CR and NCR values for IBM/SEMATECH chips. We considered a sample of these chips having CR and NCR values less than 10 and used additional flush delay data to plot Fig. 72. Consider two chips marked 'A' and 'B' that appear to be outliers in Fig. 71. When flush delay data is considered, chip 'A' still appears to be outlier, but chip 'B' cannot be considered outlier.

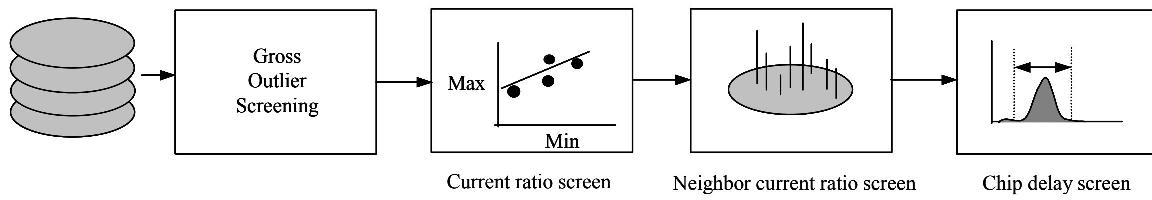


**Fig. 70. Typical analysis flow for multiple test metric screening.**

**TABLE VII. TEST STRATEGY WHEN TEST METRICS DISAGREE.**

<table>
<tr><th>Metric 1</th><th>Metric 2</th><th>Test Goal</th><th>Result</th></tr>
<tr><td>Pass</td><td>Pass</td><td>Any</td><td>Accept</td></tr>
<tr><td>Fail</td><td>Fail</td><td>Any</td><td>Reject</td></tr>
<tr><td>Pass</td><td>Fail</td><td rowspan="2">No customer return</td><td rowspan="2">Reject</td></tr>
<tr><td>Fail</td><td>Pass</td></tr>
<tr><td>Pass</td><td>Fail</td><td rowspan="2">Yield Maximization</td><td rowspan="2">Consider additional test parameters</td></tr>
<tr><td>Fail</td><td>Pass</td></tr>
</table>

**TABLE VIII. COMBINATION OF CR, NCR AND DELAY DATA.**

| CR | NCR | Delay | Comments |
|---|---|---|---|
| Low | Low | Small | Fast wafer region |
| | | Large | Resistive short/defective? |
| Low | High | Small | A chip with passive defect in good neighborhood |
| | | Large | |
| High | Low | Small | A chip with an active defect in bad neighborhood (cluster) |
| | | Large | |
| High | Low | Small | A chip with an active defect in good neighborhood |
| | | Large | |

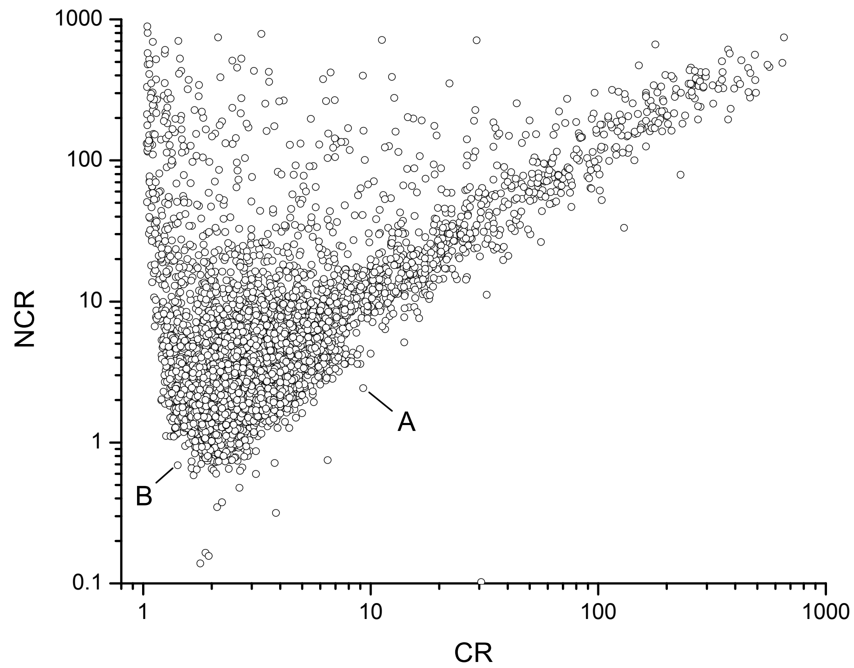


**Fig. 71. CR-NCR scatter plot for chips that passed all wafer tests or failed only $I_{DDQ}$ test.**

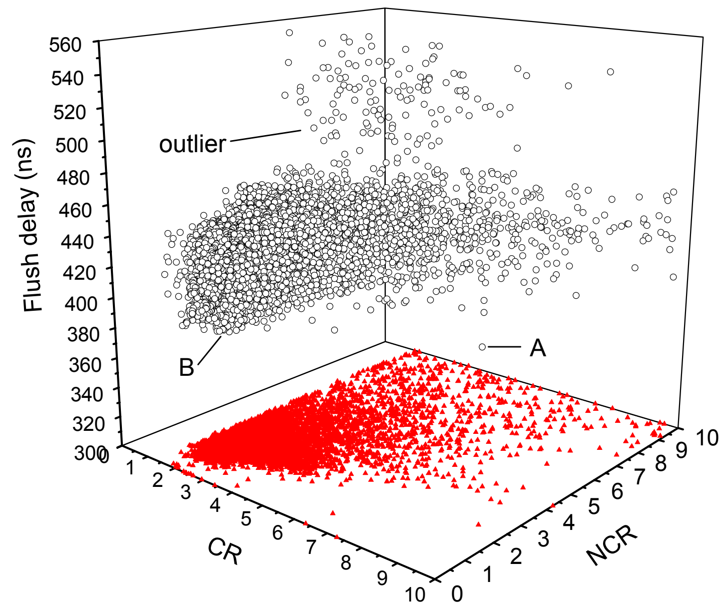


**Fig. 72. CR, NCR and flush delay scatter plot for chips having CR≤10 and NCR≤10.**

## 5.8 Conclusion

No single metric alone is sufficient to screen all outliers. The addition of a second degree of freedom also comes at the risk of additional yield loss. The thresholds for two metrics must be carefully selected to optimize yield loss/defect level. Use of additional test parameters (e.g. chip speed) is useful to resolve the disagreement between two metrics. Whether the accompanying reduction in the defect level is cost-effective is decided by yield loss and eventually depends on the quality to cost ratio [201].

Maintaining stringent process control will prove to be challenging for DSM technologies. Understanding underlying process variations and their impact on test parameters will be crucial for yield requirements [202][203]. As each parametric test loses its effectiveness, it will be necessary to correlate multiple test metrics in the future. A combination of multiple outlier screening methods may be needed [204]. Scalability of the metrics used for outlier rejection is crucially important. The basic physical mechanisms that govern CR and NCR will not change in the future. Therefore, the CR/NCR combined metric should be scaleable to future technologies. NCR is not a variance reduction technique like other methods [183]. However, a combination of CR/NCR with other test parameters can be useful for screening low-reliability chips. It will be interesting to see whether the variation in CR/NCR thresholds helps screen chips with different severities of defect currents. We did not investigate the relationship between the vector pairs yielding CR and NCR. It will be interesting to examine if such correlation exists and whether that can be useful for production implementation of NCR-based outlier rejection. Analysis of wafer patterns can be useful to reduce number of vector pairs that may be required [185][186].

# 6. ANALYSIS OF INDUSTRIAL TEST DATA

## 6.1 Introduction

This section presents the results of application of various outlier screening methods to industrial test data. The data used for the analysis comes from IBM (SEMATECH), LSI Logic and Texas Instruments (TI). The details of this data are provided in Appendices A, B and C, respectively. It must be noted that the aims of data collection at IBM, LSI Logic and TI were different. At IBM, data collection was part of the research experiment S-121 sponsored by SEMATECH while at LSI Logic and TI data collection was part of the conventional manufacturing test flow. The test vehicles used at these companies also differed drastically. The IBM/SEMATECH data is from an older technology (0.6 μm) and for relatively fewer chips (18,466 total) than the LSI data (180 nm technology, 968,387 chips). Although the TI data is for fewer chips (11347 chips) it comes from a recent technology node (130 nm). A sample of SEMATECH chips was subjected to various levels of BI and the same tests done before BI were conducted after BI. LSI and TI chips were not subjected to BI. TI chips underwent voltage stress test and the post-stress test data is available.

Defining a "good" chip is straightforward if a chip passes *all* tests. However, as it must have become clear to the reader by now, it is not so easy to define "good" for parametric failures. One oracle to use is the post-BI test result. Our analysis of defect level (DL) and yield loss (YL) uses the BI test result for validation of our approaches. We consider chips that fail tests pre-BI and pass tests after BI (called *healers*) to be unreliable[5]. For this reason, we do not include them in DL/YL calculations. Clearly, such analysis is not possible when BI data is not available. In such cases, the effectiveness of a metric is decided by its ability to screen a chip based on its "outlier-ness". Unfortunately, this is a subjective term and to a certain extent relies on the visual inspection of data.

---

[5] It is possible to regard some of these chips to be good by considering the "amount" of healing. However, pre-BI failures are always rejected upfront in a conventional test flow. Therefore, such analysis would have only academic interest, if any.

This research mostly focused on wafer level spatial correlation. The methods we developed as part of this research are as follows:

- Spatial fit method
- Statistical outlier rejection methods
- Correlation of two test parameters
- Neighbor Current Ratio (NCR)
- Immediate Neighbor Difference $I_{DDQ}$ Test (INDIT)
- Combination of CR and NCR
- Combination of CR, NCR and flush delay (CROWNE)
- Wafer signature

These methods were developed over a period of time. Like any research, our knowledge about the unknown improved as we explored different methods. Moreover, the data sets we used were not all available at the beginning of this study. The IBM/SEMATECH data was available since September 2000. But the LSI data was obtained in June 2002 while the TI data was obtained in August 2003 and later in Jan 2004. As data became available over a period of time, we could not compare all of our methods. We have already described these methods in earlier sections. In the following subsections, we first describe data used for evaluating an approach, followed by the results.

## 6.2 Analysis Using Spatial Fit Method

As described earlier, we did linear regression fit using die XY-coordinates and $I_{DDQ}$ as the Z-coordinate to obtain the best fitting plane. IBM/SEMATECH data was used for validating this approach. We rejected functional, stuck-at and delay fail chips. Gross $I_{DDQ}$-only failed chips were rejected using Chauvenet's criterion with a threshold of 0.5. Fig. 73 shows $I_{DDQ}$ projections on the XY plane for a wafer before outlier removal. The projection plot after removing gross outliers using Chauvenet's criterion is shown in Fig. 74. Note that a chip near the center of the wafer that is off-the Z-axis scale is still an outlier. This chip passed all tests including SEMATECH $I_{DDQ}$ test. However, for all practical purposes, it is likely to be defective.

Before applying Chauvenet's criterion a Normalizing transform was used. We assumed that the original $I_{DDQ}$ distribution is approximately lognormal. To convert a lognormal distribution to

the Normal distribution a lognormal transform was used. That is, each $I_{DDQ}$ reading from the set of readings for a vector is divided by the minimum $I_{DDQ}$ for that vector on the wafer ($x_{min}$) and the logarithm of the ratio is obtained.

$$x(Normal) = \log\left(\frac{x(\log normal)}{x_{\min}}\right) \quad (6.1)$$

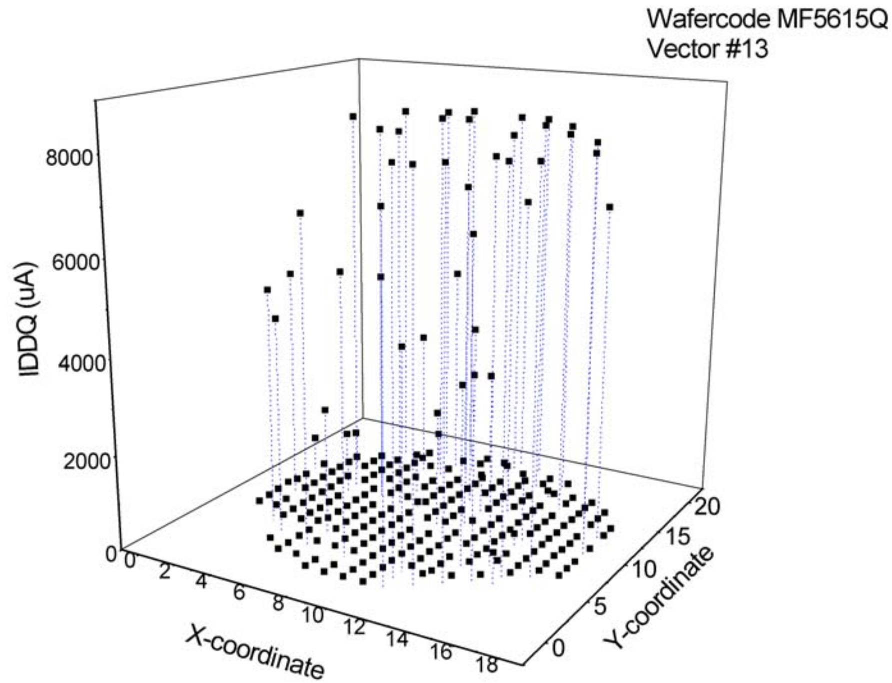


**Fig. 73. $I_{DDQ}$ projections on wafer XY plane.**

The data was reverse transformed after outlier rejection before obtaining the best-fitting plane. The residuals ($R_i$) were computed by subtracting the predicted $I_{DDQ}$ from the actual value for each vector *i*.

$$R_i = I_{DDQ\text{-actual (i)}} - I_{DDQ\text{-estimated (i)}} \quad (6.2)$$

We then determined the standard deviation of the residuals across the wafer for each vector. The dice that are above any predicted $I_{DDQ}$ value by more than 3σ were considered high risk and

rejected. Dice having actual $I_{DDQ}$ below the predicted value by more than $3\sigma$ are not rejected. However, these chips may be a reliability risk being in a bad neighborhood [205]. The execution flow for this procedure is shown in Fig. 75.

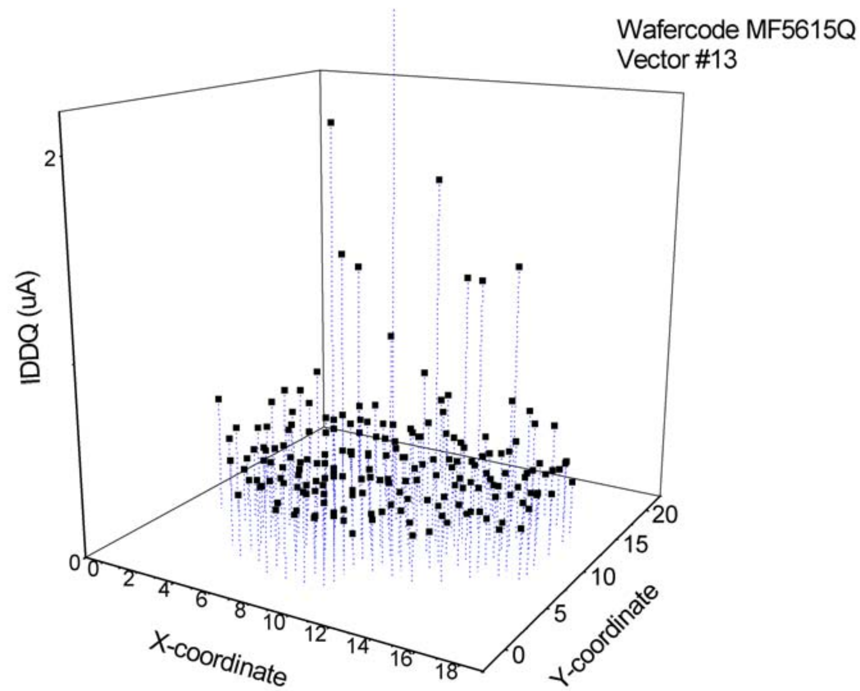


**Fig. 74. $I_{DDQ}$ projections on wafer XY plane after outlier removal.**

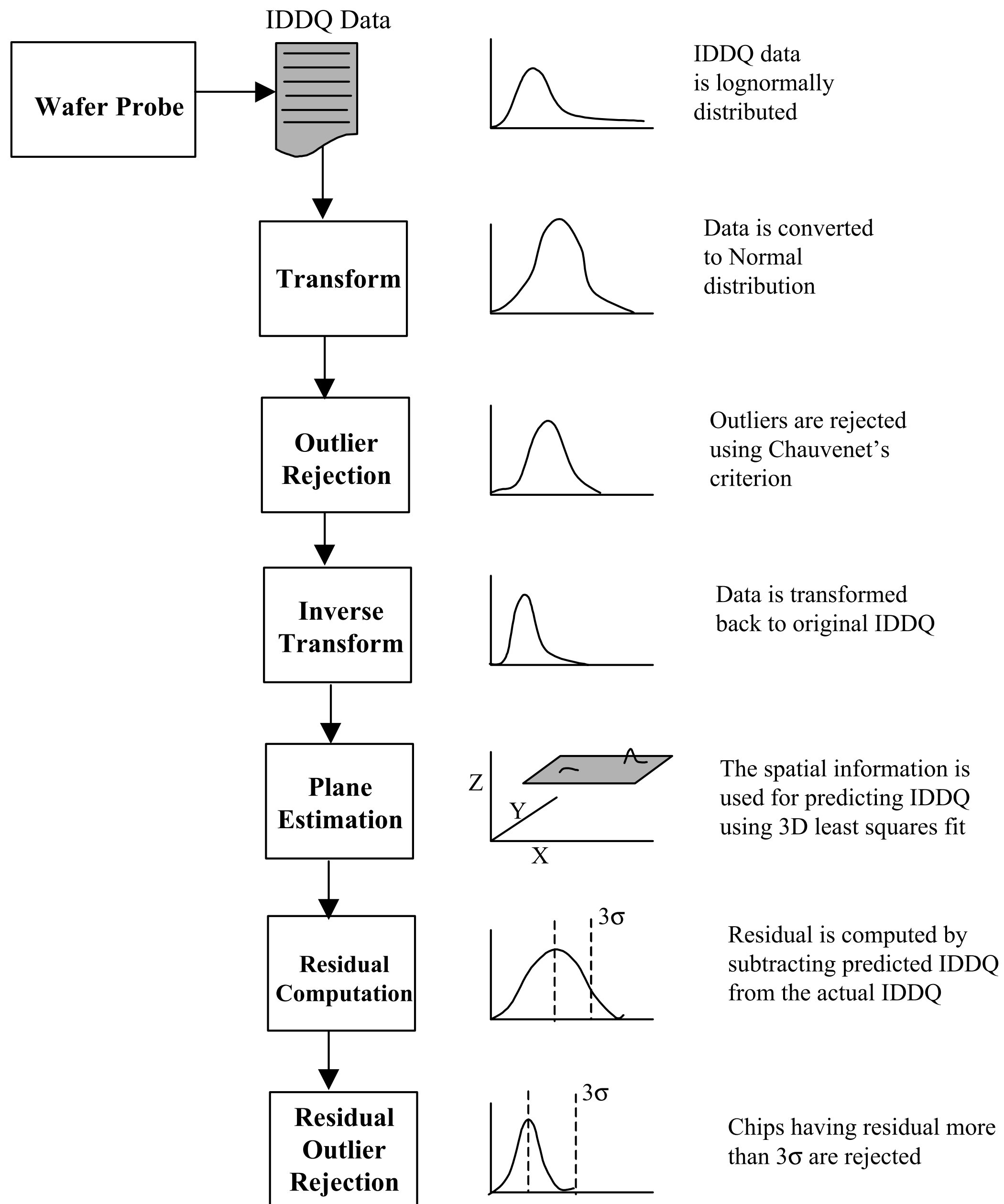


**Fig. 75. Execution flow for outlier screening in spatial fit method.**

Table IX shows a sample of typical cases on a wafer. This wafer had a total of 197 dice out of which 140 passed all the tests and 21 failed only the 5 μA threshold $I_{DDQ}$ test at the wafer level. Chauvenet's criterion with threshold 0.5 rejected 44 dice out of which 13 failed $I_{DDQ}$-only test at wafer probe, 28 failed all tests and 3 failed all but functional test. In the accepted dice, 8

dice failed only $I_{DDQ}$ test. Out of these, 4 failed only $I_{DDQ}$ test again after BI, one exhibited power failure, and BI results were not available for the remaining 3 dice. The three dice had fewer than three neighbors and could not be predicted using the plane estimate. The number of neighbors that passed all the tests or failed only $I_{DDQ}$ test at wafer probe are listed in the second column, the remaining were either voltage fails or missing. The predicted values and actual $I_{DDQ}$ values are shown for min, max and mean for that particular dice across all 195 vectors. Note that die 1003 predictions are lower than the actual values, as all good dice surround this die. Similarly, the actual mean for die 0416 is lower than the predicted value. If the difference is large, one must suspect the neighboring dice. One of the neighbors for this die had failed all the tests. It is likely that a defect in this area has affected this die too. The post-BI result is not available for this die.

**TABLE IX. COMPARISON OF PREDICTED AND ACTUAL $I_{DDQ}$**

| Die ID | Pass/ | Predicted | | | Actual | | |
|---|---|---|---|---|---|---|---|
| | $I_{DDQ}$-fail | Min | Max | Mean | Min | Max | Mean |
| 0316 | 2/3 | 2.62 | 8.44 | 4.12 | 4.24 | 7.36 | 5.68 |
| 0416 | 3/2 | 2.48 | 4.60 | 3.44 | 0.41 | 7.60 | 1.84 |
| 0415 | 2/2 | 1.15 | 3.76 | 1.85 | 3.62 | 7.66 | 5.42 |
| 0206 | 2/1 | 4.85 | 6.04 | 5.52 | 0.32 | 7.94 | 1.00 |
| 1003 | 7/0 | 0.38 | 0.74 | 0.47 | 4.20 | 5.40 | 4.73 |

Fig. 76 shows the distribution of the residuals for the plane fitting approach. These residuals are plotted against XY coordinates in Fig. 77. Fault-free chip residual values form a cluster near zero while defective dice have much higher residuals (e.g. a die having residual near 4 μA) and float above the cluster.

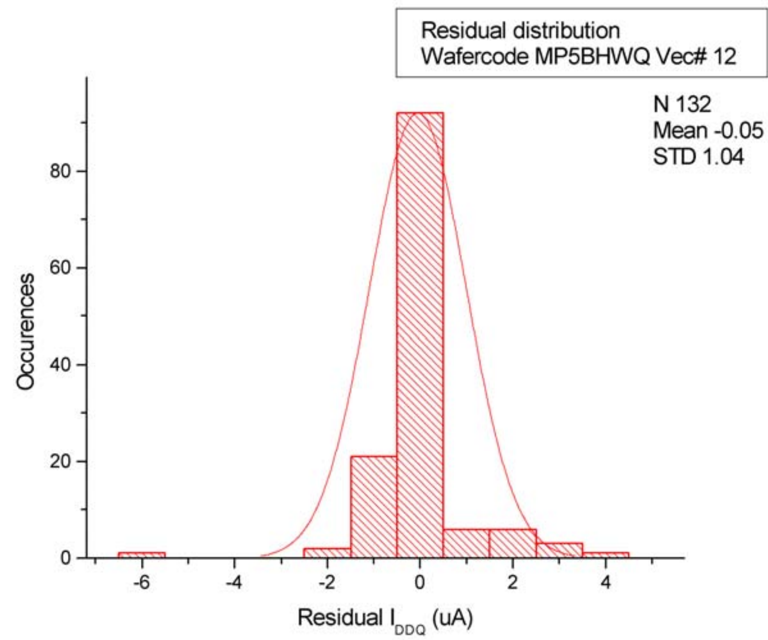


**Fig. 76. Distribution of residuals for spatial fit approach for a vector across wafer.**

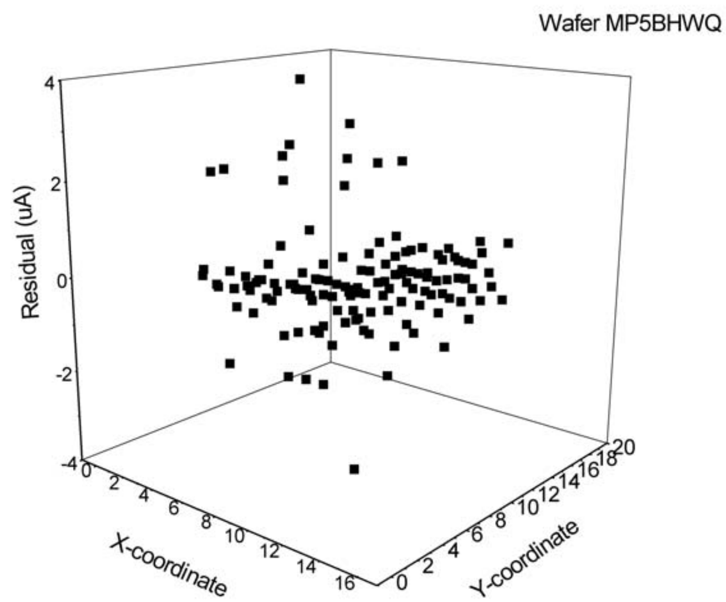


**Fig. 77. A three-dimensional view of residual variation for a vector across a wafer.**

Table IX and Table X summarize a comparison of several limit-setting methods. For the static threshold approach, we used the 5 μA threshold value. For the wafer median approach, the median $I_{DDQ}$ across the wafer for each vector was determined. The threshold value for $I_{DDQ}$ used was computed as:

$$I_{th(i)} = \text{median}_{(i)} + 3\sigma_i \quad (6.3)$$

For the delta-$I_{DDQ}$ method, we computed the medians and the standard deviations of the differences between consecutive readings of all dice. If any delta was more than 3σ above the median, the die was rejected. For our implementation of the current signature approach, we followed the same procedure as for delta-$I_{DDQ}$ after sorting all the readings in ascending order. For the plane-fit approach, Chauvenet's criterion was first used to reject gross outliers, a plane-fit estimate was computed for $I_{DDQ}$-only failed dice, residuals were computed for all 195 vectors for all dice on a wafer, and the dice having residuals outside 3σ were rejected. We then used SEMATECH test results to compute the BI fallout, overkill and the defect levels for the different methods. Because we are not certain whether dice with $I_{DDQ}$ above the 5 μA threshold are actually faulty or not, we considered post-BI $I_{DDQ}$-only failures first as fail and then as pass as shown in Table IX and Table X, respectively. The actual values will presumably lie in between.

**TABLE X. COMPARISON OF DIFFERENT LIMIT SETTING SCHEMES CONSIDERING POST-BI $I_{DDQ}$ FAILS**

| Test Result | Approach | | | | |
|---|---|---|---|---|---|
| | Single Threshold | Wafer Median | Delta $I_{DDQ}$ | Current Signature | Plane Fit |
| Accept | 16644 | 16001 | 14464 | 15993 | 11099 |
| BI | 2215 | 3652 | 3143 | 3075 | 1419 |
| Fail BI | 664 | 1821 | 1660 | 1358 | 252 |
| % fail | 29.97 | 49.86 | 52.81 | 44.16 | 17.76 |
| Reject | 1822 | 2465 | 4002 | 2473 | 7367 |
| BI | 1674 | 237 | 746 | 814 | 2470 |
| Pass BI | 313 | 33 | 381 | 147 | 697 |
| % pass | 18.69 | 13.92 | 51.07 | 18.05 | 28.22 |
| % Overkill | 1.84 | 1.85 | 11.06 | 2.41 | 11.26 |
| % Yield | 90.13 | 86.65 | 78.32 | 86.61 | 60.11 |
| % DL | 27.02 | 43.21 | 41.36 | 38.25 | 10.67 |

**TABLE XI. COMPARISON OF DIFFERENT LIMIT SETTING SCHEMES IGNORING POST-BI $I_{DDQ}$ FAILS**

| Test Result | Approach | | | | |
|---|---|---|---|---|---|
| | Single Threshold | Wafer Median | Delta $I_{DDQ}$ | Current Signature | Plane Fit |
| Accept | 16644 | 16001 | 14464 | 15993 | 11099 |
| BI | 2215 | 3652 | 3143 | 3075 | 1419 |
| Fail BI | 539 | 512 | 356 | 297 | 220 |
| % fail | 24.33 | 14.02 | 11.32 | 9.65 | 15.50 |
| Reject | 1822 | 2465 | 4002 | 2473 | 7367 |
| BI | 1674 | 237 | 746 | 814 | 2470 |
| Pass BI | 1555 | 91 | 528 | 454 | 1392 |
| % pass | 92.89 | 38.39 | 70.77 | 55.77 | 56.36 |
| % Overkill | 9.16 | 5.12 | 15.33 | 7.46 | 22.48 |
| % Yield | 90.13 | 86.65 | 78.32 | 86.61 | 60.11 |
| % DL | 21.93 | 12.14 | 8.87 | 8.36 | 9.32 |

The yield values are obtained simply by dividing the number of accepted chips by the total number of chips (18466) in the data set. We use the BI sample to estimate misclassification of chips. The overkill is the yield loss due to rejecting chips that pass BI. This figure includes healing defects. For example, in Table X, all of the chips that are rejected due to the single threshold test at wafer level would also be expected to fail after BI. But as can be seen, 313 pass. In practical production, chips that fail wafer test would be rejected, since too few would heal to justify the packaging and additional testing cost, and healers are a reliability hazard. Thus overkill figures are higher than if computed using standard practice.

The defect level is assumed to be the BI failure rate of accepted chips scaled by the yield. This permits comparison of the absolute number of defective parts for each method. Since the SEMATECH experiment focused on test method evaluation, the sample selected for BI was biased towards failed dice, particularly $I_{DDQ}$-only dice. This explains why defect levels are abnormally high. The high BI fallout rate of accepted chips for the wafer median, delta-$I_{DDQ}$ and current signature test can be explained as follows. Some of the $I_{DDQ}$ readings are several mill

amperes, which results in pass/fail thresholds much higher than the other two approaches. This causes many dice with $I_{DDQ}$ above 5 μA to be accepted, which subsequently fail the 5 μA threshold after BI. This effect can also explain the difference in accepted and rejected fallout rates between the static threshold and plane fit methods.

Large jumps in deltas between $I_{DDQ}$ values are eliminated in the current signature approach due to sorting, so it accepts more dice than delta-$I_{DDQ}$. Many of the dice accepted by current signatures have elevated $I_{DDQ}$ for all vectors. Many of these devices fail BI, resulting in a lower percentage of accepted chips passing post-BI tests.

It can be seen that each method has its own limitations. The results also underscore how the static-threshold approach is the worst choice for $I_{DDQ}$ testing. When $I_{DDQ}$-only failures are ignored, it has the highest defect level. Although the yield of the plane-fit method is lower than the other methods, it has by far the lowest defect level in Table X, and a relatively low one in Table X, indicating that the dice rejected by this method are indeed defective. However this comes at the cost of the highest overkill. By selecting a less stringent probability threshold and/or residual rejection limit, it is possible to achieve higher yield with a low defect level. As part of a BI minimization strategy, thresholds can be set so that accepted dice have sufficiently low defect levels to avoid BI, while dice rejected by this method are burned in, rather than being counted as yield loss.

### 6.3 Evaluation of Statistical Outlier Rejection Methods

We evaluated Chauvenet's criterion and the Tukey test method for rejecting outliers using IBM/SEMATECH data. Our analysis flow was as follows. All the analysis was done on a per-wafer basis. There are 195 measurements for each die corresponding to 195 different vectors. For each vector we converted the data to the Normal distribution using the lognormal transform described in the previous subsection. Then we found the mean ($\mu_i$) and standard deviation ($\sigma_i$). Then we applied both data rejection methods independently. In case of Chauvenet's criterion, the probability threshold was changed to find the effect of the threshold on the number of dice rejected. For Tukey test method, we computed lower and upper quartile pointes (Q1 and Q3), determined IQR and changed the UQL by changing the multiplying factor $k$ in the equation UQL+$k$.IQR.

The results for Chauvenet's criterion are summarized in Table XII and those for Tukey test are shown in Table XIII. An accepted chip is considered failed after BI if it failed any test including the 5 μA threshold test. A rejected chip is considered passed after BI if it passes all the tests. Table XIV shows the results for the static threshold method.

The yield loss, overkill and defect level are computed as follows:

$$Yield \quad Loss \; (YL) = \frac{\left(\frac{P}{K}\right)R + \left(\frac{F}{B}\right)A}{T}.100 \qquad (6.4)$$

$$Overkill = \frac{P}{K}(1 - \frac{A}{T}).100 \qquad (6.5)$$

$$Defect\,Level(DL) = \frac{F}{B}.\frac{A}{T}.100 \qquad (6.6)$$

where

A = total number of accepted chips

B = number of accepted chips burned-in (BI)

F = number of accepted chips that failed BI

R = total number of rejected chips

K = number of rejected chips burned-in

P = number of rejected chips that passed after BI

T = A + R = total chips = 18466.

Thus, P/K represents the fraction of the rejected chips passed after BI and F/B represents the fraction of the accepted chips failed after BI. Both contribute to the yield loss. The defect level is a measure of bad parts that got shipped while overkill is a measure of good chips that were considered bad. If a die failed at wafer level and passed all tests after BI, it is counted as overkill. Since we used the SEMATECH test result after BI (which includes the 5 μA $I_{DDQ}$ test), the actual defect level and overkill for other static thresholds would be worse than those presented here.

**TABLE XII. RESULTS FOR CHAUVENET'S CRITERION**

| Threshold | A | B | F | R | K | P | YL % | Overkill % | DL % |
|---|---|---|---|---|---|---|---|---|---|
| 0.1 | 13777 | 3051 | 315 | 4689 | 838 | 90 | 10.42 | 2.72 | 7.7 |
| 0.2 | 13282 | 2802 | 277 | 5184 | 1087 | 109 | 9.92 | 2.81 | 7.11 |
| 0.3 | 13022 | 2617 | 251 | 5444 | 1272 | 126 | 9.90 | 2.92 | 6.76 |
| 0.4 | 12857 | 2486 | 238 | 5609 | 1403 | 134 | 9.57 | 2.90 | 6.66 |
| 0.5 | 12680 | 2330 | 224 | 5786 | 1559 | 156 | 9.74 | 3.13 | 6.60 |
| 0.6 | 12531 | 2201 | 213 | 5935 | 1688 | 175 | 9.89 | 3.33 | 6.57 |

**TABLE XIII. RESULTS FOR TUKEY TEST**

| Threshold[6] | A | B | F | R | K | P | YL % | Overkill % | DL % |
|---|---|---|---|---|---|---|---|---|---|
| 0.5 | 11550 | 1715 | 234 | 6916 | 2042 | 366 | 15.25 | 6.71 | 8.5 |
| 1 | 11752 | 1774 | 268 | 6714 | 1983 | 341 | 15.86 | 6.25 | 9.6 |
| 1.5 | 11870 | 1826 | 290 | 6596 | 1981 | 319 | 15.96 | 5.75 | 10.2 |
| 2 | 11969 | 1874 | 311 | 6497 | 1975 | 293 | 15.98 | 5.21 | 10.7 |
| 2.5 | 12039 | 1915 | 337 | 6427 | 1934 | 278 | 16.47 | 5.00 | 11.4 |
| 3 | 12093 | 1953 | 357 | 6373 | 1896 | 260 | 16.70 | 4.73 | 11.9 |

**TABLE XIV. RESULTS FOR STATIC THRESHOLD METHOD**

| Threshold μA | A | B | F | R | K | P | YL % | Overkill % | DL % |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 14627 | 2174 | 745 | 3839 | 1715 | 435 | 32.41 | 5.27 | 27.1 |
| 2 | 15933 | 2258 | 753 | 2533 | 1631 | 359 | 31.79 | 3.01 | 28.8 |
| 5 | 16777 | 2331 | 759 | 1689 | 1558 | 292 | 31.29 | 1.71 | 29.6 |
| 10 | 17119 | 2653 | 911 | 1347 | 1236 | 122 | 32.55 | 0.72 | 31.8 |
| 20 | 17350 | 2871 | 1093 | 1116 | 1018 | 86 | 36.28 | 0.51 | 35.8 |
| 50 | 17590 | 3087 | 1292 | 876 | 802 | 69 | 40.27 | 0.41 | 39.9 |

[6] Threshold is UQL + k•IQR.  The scaling factor ‘k’ is specified in this column.

Statistical data rejection methods provide a simple yet powerful way to sustain $I_{DDQ}$ testing for future technologies. Nevertheless various parameters used in their application must be carefully selected and justified. These methods are like a double-edged sword – if used carefully they provide a powerful means of data analysis but improper selection of parameters can abuse them. Especially Chauvenet's criterion has the lowest yield loss and defect level. However, true effectiveness of any method can be verified only with a sufficient sample of data. The use of these methods is not justified if the sample space is limited. The data collected during wafer probe or characterization can be used for analysis and can prevent bad chips from being shipped. Of course, in order to account for lot-level and wafer-level variations it is necessary to perform the analysis more often and refine the threshold values using methods like regression fit. As a process becomes more mature a better goodness of fit can be obtained. Clearly, outlier rejection methods do not provide a push-button solution to the problem, but if used intelligently, they can alleviate costly customer returns in the long term.

## 6.4 Correlating Two Test Parameters

We exploited correlation between $I_{DDQ}$ and flush delay in order to estimate $I_{DDQ}$. The approach was evaluated using IBM/SEMATECH data. We limited our data set to chips having six hours BI data (2660 chips). A large percentage of chips from SEMATECH data fail only $I_{DDQ}$ test after BI. There was no clear oracle to define which of these chips are indeed defective, as all chips did not go through multiple BI cycles. It was necessary to eliminate gross outliers while forming a spatial estimate. To decide the outlier rejection limit, we obtained the histogram of wafer level minimum $I_{DDQ}$ for these chips as shown in Fig. 78. It reveals that a small percentage of chips (~2.5%) have *minimum* $I_{DDQ}$ of more than 100 μA. Since minimum $I_{DDQ}$ is mostly[7] an intrinsic component, such a high value is likely indicative of a gross defect. So chips with $I_{DDQ}$ above 100 μA were rejected even though they pass all voltage tests (1773 chips remain).

---

[7] A part of minimum $I_{DDQ}$ can stem from a subtle defect.

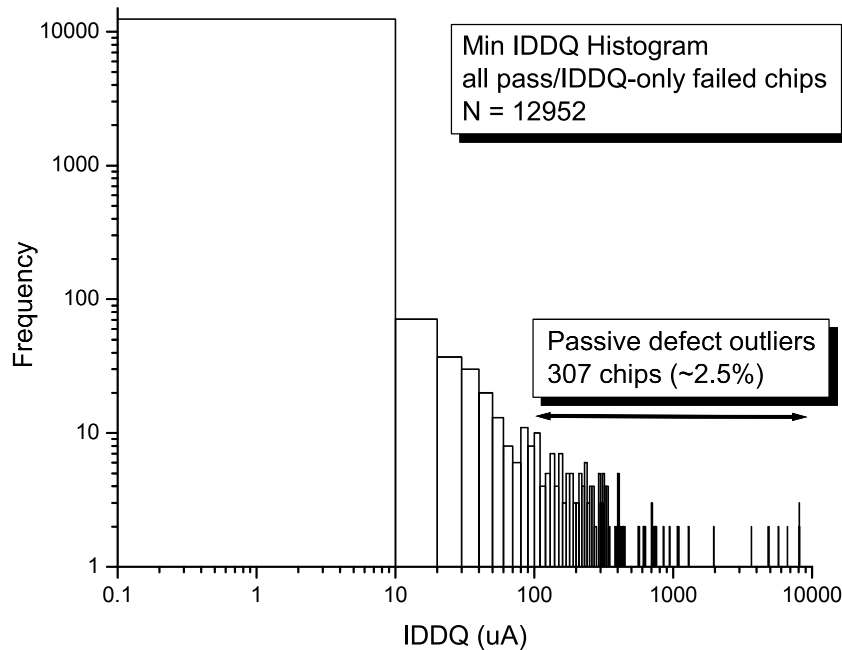


**Fig. 78. Wafer-level minimum $I_{DDQ}$ distribution for SEMATECH chips.**

We obtained two estimates of $I_{DDQ}$ for each vector for each die. The first estimate was obtained by using 3D linear regression between XY coordinates of the neighboring chips' $I_{DDQ}$ as the Z coordinate. Only functional dice were used for plane formation. Since at least three dice are needed to define a plane, it was not possible to estimate $I_{DDQ}$ for dice surrounded by more than five defective/missing chips. Such dice were excluded (total 187) from the analysis. The center die readings were not considered to avoid bias.

The second estimate was obtained by linear regression between neighboring chips' $I_{DDQ}$ and flush delay values. Only functional dice were used in linear regression. This relationship was then used with the center die flush delay to predict the center die $I_{DDQ}$. The final estimate was the average of these two estimates. To account for random variations, a guard band of 20% was added to obtain an upper bound on estimated $I_{DDQ}$ ($I_E$). Such analysis was performed for each vector.

The total leakage current ($I_T$) consists of two components: an intrinsic leakage component ($I_L$) and a defective component ($I_D$). If neighboring chips are fault-free, a high correlation between actual ($I_T$) and estimated $I_{DDQ}$ ($I_E$) can be observed. However, defective component $I_D$ depends on the nature and severity of the defect, input vector and several other factors. Thus for a defective chip actual values deviate considerably from the estimate. This is conceptually illustrated by a scatter plot of estimated and actual $I_{DDQ}$ for a die as shown in Fig. 79. For a defect-free chip, actual values would lie within the guard band obtained from the estimated values. If a defect exists, the elevated $I_{DDQ}$ values form a cluster as shown. The approximate defective component of the current is the residual value give by:

$$\delta = I_T - I_E \tag{6.7}$$

While a positive $\delta$ indicates a spatial outlier and a likely defective die, a negative $\delta$ could signify a good die in the bad neighborhood or simply a test escape.

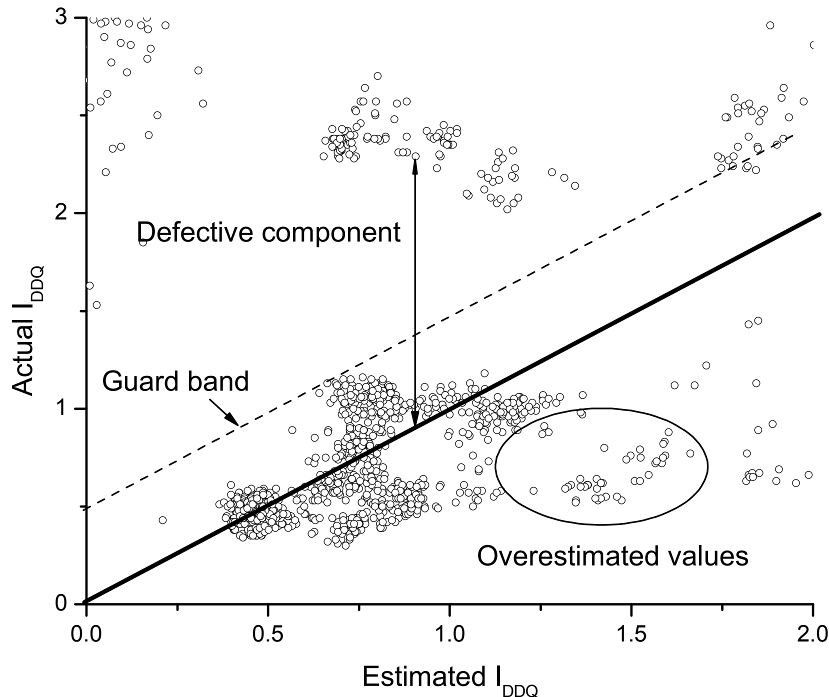


**Fig. 79. Correlating actual and estimated $I_{DDQ}$ values.**

Fig. 80 shows the actual and estimated $I_{DDQ}$ values for 50 vectors for a chip. Both the spatial-fit estimate and the one obtained by correlating flush delay information are shown. It shows that for many vectors the estimate obtained by using both flush delay and spatial-fit is more accurate than that obtained by spatial-fit method alone. Note that for defective leakage residual value increases than that for spatial-fit (plane) estimate. This makes identifying *apparent* outlier (fast and leaky) chips easier.

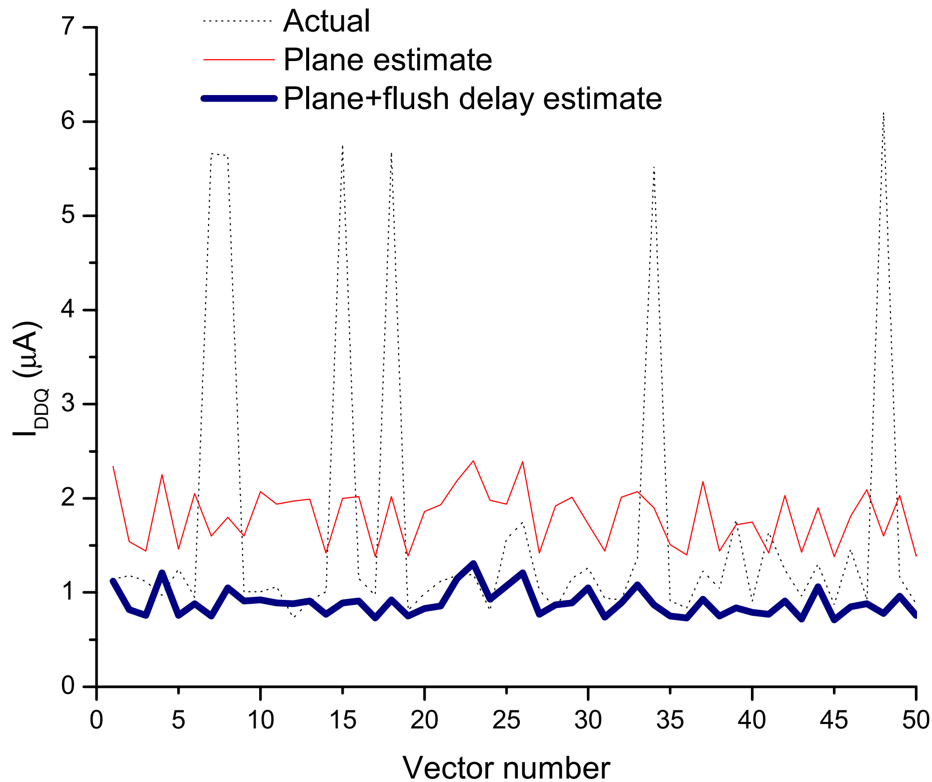


**Fig. 80. Actual and estimated $I_{DDQ}$ values for a chip.**

Fig. 81 shows the histogram of residual values for all chips (1773) for all 195 vectors. Approximately 80% of the residual values are within ± 25 μA. Therefore, any chip having $\delta > 25$ μA is rejected. Any chip having $\delta < -25$ μA is considered a reliability risk and is rejected. Since the 20% guard band is included, this is a relatively loose threshold.

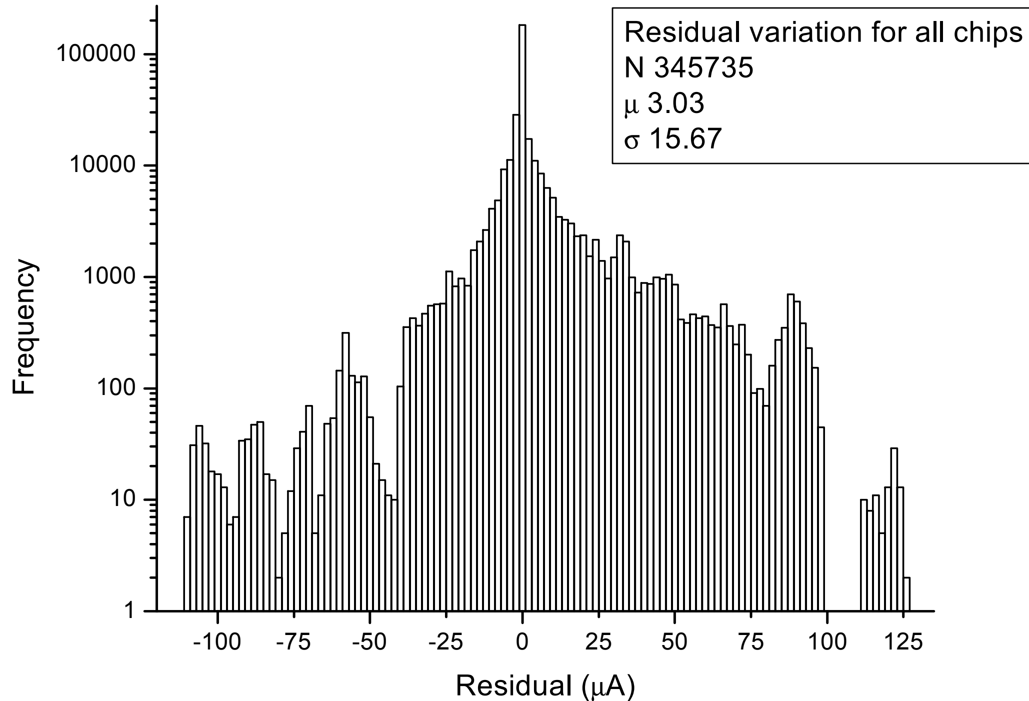


**Fig. 81. Residual distribution for all chips.**

We performed wafer level analysis for all chips that either passed all tests or failed only $I_{DDQ}$ test after removing outliers as explained earlier. A total of 1773 dice, of which 1044 passed all wafer tests and 729 failed only $I_{DDQ}$ test, were analyzed. Out of 1044 dice, 1003 passed all tests after BI, 24 failed only $I_{DDQ}$ test and 17 failed voltage test(s). Out of 729 $I_{DDQ}$-only fail dice, 524 failed only $I_{DDQ}$ test after BI, 187 passed all tests (i.e. were healers) and 18 failed voltage test(s). Since 5 μA is not necessarily a good manufacturing limit, all 729 $I_{DDQ}$-only fail dice are not necessarily defective.

With $I_{DDQ}$ plane fit alone, 1487 chips are accepted and 286 chips are rejected. Out of 1487 accepted chips, 1135 passed all tests after BI, 323 failed only $I_{DDQ}$ test and 29 failed voltage test. Out of 286 rejected chips, 55 passed all tests after BI, 225 failed only $I_{DDQ}$ test and 6 chips failed voltage test.

When flush delay information is combined with the $I_{DDQ}$ plane estimate, 1538 chips are

accepted and 235 chips are rejected. Out of 1538 accepted chips, 1176 passed all tests after BI, 332 failed only $I_{DDQ}$ test and 30 failed voltage test. Out of 235 rejected chips, 14 passed all tests after BI, 216 failed only $I_{DDQ}$ test and 5 chips failed voltage test. Table XV shows the distribution of chips accepted by both methods, rejected by either or both. In each category chips are divided according to their post BI result.

**TABLE XV. DISTRIBUTION OF CHIPS FOR TWO METHODS**

| Method | $I_{DDQ}$-only accept | $I_{DDQ}$-only reject | Post BI Result |
|---|---|---|---|
| $I_{DDQ}$ + Flush delay accept | 1132 | 44 | Pass all |
| | 302 | 30 | Fail $I_{DDQ}$ |
| | 29 | 1 | Voltage fail |
| $I_{DDQ}$+ Flush delay reject | 3 | 11 | Pass all |
| | 21 | 195 | Fail $I_{DDQ}$ |
| | 0 | 5 | Voltage fail |

When only $I_{DDQ}$ information is used the yield is 83.8%. If we assume all $I_{DDQ}$-only failed chips are fault-free (defective), the defect level is 1.95% (23.67%). When flush delay information is also used and a chip is rejected only if rejected by both methods, the yield is 88%. Assuming all $I_{DDQ}$-only failed chips are fault-free (defective), the defect level becomes 1.92% (22.6%). Thus there is almost a 5% improvement in yield and a slight reduction in the defect level by using flush delay information. The reduction in defect level from 1.95% to 1.92% is not statistically significant considering the sample size. However, it can be argued that when flush delay information is combined with $I_{DDQ}$, an increase in yield is not accompanied by an increase in defect level.

### 6.5 Exploiting Wafer Spatial Correlation Using NCR

To evaluate the effectiveness of NCR in screening wafer-level spatial outliers, we applied this method to IBM/SEMATECH data. We considered only those chips that passed all tests (1102) or failed only 5 μA threshold $I_{DDQ}$ test (1558) at the wafer level and underwent six hours of BI. If any $I_{DDQ}$ reading exceeded 100 μA the chip was assumed to contain a gross defect. Therefore, all

chips for which $I_{DDQ}$ exceeded 100 μA were removed from the analysis. This reduced the dataset from 1558 $I_{DDQ}$-only fails to 858 $I_{DDQ}$-only failed chips. The distribution of the entire dataset according to the number of available neighbors is shown in Table XVI. Even though the BI sample was not random, it is interesting to note that a large number of dice that passed all wafer and post-BI tests or failed only $I_{DDQ}$ test at both levels have 3 or more neighbors. Although NCRs can be obtained even if a single neighboring die is available, confidence in outlier rejection is improved if more neighboring dice are present.

**TABLE XVI: DISTRIBUTION OF DICE IN THE ORIGINAL DATASET**

| Wafer Probe | Post-BI | Number of available neighbors | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| All-pass | All pass | 4 | 16 | 39 | 82 | 120 | 185 | 233 | 225 | 152 |
| | $I_{DDQ}$-fail | 0 | 0 | 3 | 3 | 4 | 2 | 7 | 7 | 1 |
| | Other | 0 | 1 | 1 | 0 | 3 | 1 | 6 | 5 | 2 |
| $I_{DDQ}$-fail | All pass | 7 | 9 | 27 | 47 | 48 | 37 | 25 | 23 | 9 |
| | $I_{DDQ}$-fail | 7 | 19 | 51 | 99 | 95 | 105 | 107 | 82 | 40 |
| | Other | 1 | 0 | 1 | 5 | 4 | 4 | 5 | 0 | 1 |

Since we considered only immediate neighboring dice, the dice having zero adjacent neighbors (19 dice shaded in Table XVI) could not be considered for analysis. Thus the total dice in the dataset were reduced from 1960 to 1941 chips shown in Table XVII (1098 all pass, 843 $I_{DDQ}$ fail). For each available neighbor a total of 195 NCRs were obtained and the maximum NCR value (across all neighbors) was used for the pass/fail decision. Note that this means that NCR-based rejection was more stringent towards chips surrounded by more good neighbors.

To compare the effectiveness of NCR with current ratios, we considered the same dataset and used current ratios for the pass/fail decision. An important criterion was to select threshold values for a fair comparison. We observed the distribution of CR for all pass and $I_{DDQ}$-only failed chips, shown in Fig. 82. It can be observed that chips that passed all tests but have a current ratio of more than 4 seem to be “outliers”. We therefore considered a CR of 4 to be an appropriate pass/fail threshold. Then the NCR threshold was adjusted so as to match the same defect level obtained by the CR method. While computing the defect level, all post-BI $I_{DDQ}$-only failed chips

were considered fault-free. For the same defect level as obtained by the CR threshold of 4, the NCR threshold value was 21.

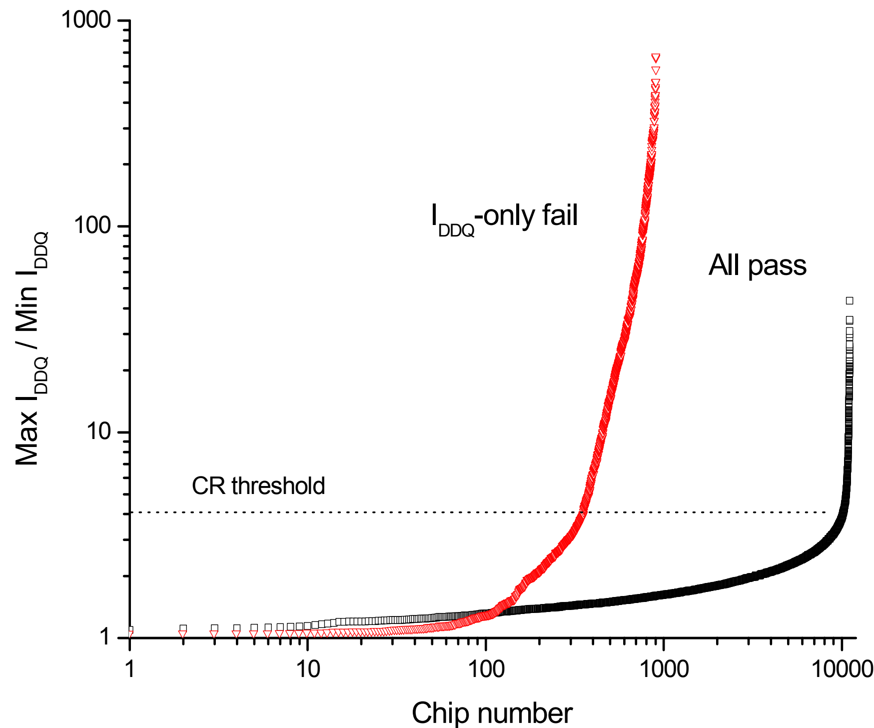


**Fig. 82. Current Ratios for IBM/SEMATECH chips in the dataset.**

Table XVII shows the distribution of chips in different categories. All chips are divided into two main categories: Chips accepted by current ratio (CR accept) and chips rejected by current ratio (CR reject). They are further subdivided into two categories: chips accepted by neighbor current ratio (NCR accept) and those rejected by neighbor current ratio (NCR reject). These four categories are divided depending on their wafer probe result: (a) chips that passed all SEMATECH tests ("All pass") and (b) chips that failed the 5 µA threshold $I_{DDQ}$ test but passed other tests ("$I_{DDQ}$-only fail"). Each category is subdivided based on the post-BI SEMATECH test result. This distinction is made to understand the distribution of NCRs and CRs in different categories and understand if certain chips get detected by one method but not by the other and the rejection rate of healer chips. As mentioned earlier, the 5 µA test limit is not necessarily a good manufacturing limit. So the difference in $I_{DDQ}$-only failed chips detected by one method but

not the other may not be insignificant.

**TABLE XVII. DISTRIBUTION OF CHIPS FOR DIFFERENT TEST METHODS**

| SEMATECH Wafer Probe Test Result | CR Accept | | CR Reject | | Post 6-hour BI SEMATECH Test Result |
|---|---|---|---|---|---|
| | NCR accept (1153) | NCR reject (149) | NCR accept (280) | NCR reject (359) | |
| All pass | 949 | 0 | 100 | 3 | All pass |
| | 19 | 0 | 7 | 1 | $I_{DDQ}$ fail |
| | 17 | 0 | 2 | 0 | Other |
| $I_{DDQ}$-fail | 69 | 26 | 76 | 54 | All pass |
| | 94 | 119 | 90 | 295 | $I_{DDQ}$ fail |
| | 5 | 4 | 5 | 6 | Other |

The overkill and defect level are computed as follows:

$$Overkill = \frac{Number\ of\ chips\ that\ pass\ after\ BI}{Total\ number\ of\ chips\ rejected}.100 \qquad (6.8)$$

$$Defect\ Level = \frac{Number\ of\ chips\ that\ fail\ after\ BI}{Total\ number\ of\ chips\ accepted}.100 \qquad (6.9)$$

These values are scaled appropriately considering the entire population that was not burned in. Table XVIII shows overkill and defect level (DL) values for both methods and their combination. In the combined method, a chip is rejected if it is considered faulty by either method. Since $I_{DDQ}$-only fail chips are not conclusively defective, overkill and DL are computed by considering all such chips fault-free and then by considering all such chips faulty. The columns headed “Good” (“Faulty”) have overkill or defect levels values computed by considering all $I_{DDQ}$-only failed chips fault-free (faulty). The actual values would lie between these two extremes. Table XVII reveals many interesting findings. Since both NCR and CR thresholds are high, only 3 of the all pass chips at wafer probe are rejected by both methods. The NCR test rejects fewer chips that pass all wafer tests than the CR test. The SEMATECH data has many healer chips that have reduced $I_{DDQ}$ after BI (and thus pass all tests). NCR test accepts fewer healer chips than CR tests. Since healers represent unstable or unreliable chips, they would typically get rejected up front in the test flow and will not be subjected to BI. NCR test rejects

more chips that fail wafer level and post-BI $I_{DDQ}$ test than the CR test. Considering the high NCR threshold of 21, these chips are more likely to be defective.

**TABLE XVIII. OVERKILL AND DL FOR DIFFERENT TEST METHODS**

| Metric | Overkill % | | Defect Level % | | Effective Yield % |
|---|---|---|---|---|---|
| | Good | Faulty | Good | Faulty | |
| CR | 97.97 | 36.46 | 2.00 | 19.82 | 67.08 |
| NCR | 98.03 | 16.34 | 2.02 | 16.68 | 77.83 |
| CR+NCR | 97.84 | 32.87 | 1.91 | 11.71 | 59.40 |

Fig. 83 shows the distribution of post-BI failures of chips according to their maximum NCR. The healer chips are shown separately. As expected, the bins with maximum NCR values less than 2.5 have a high percentage of fault-free chips and the bins for high NCR values have higher failure rates. Since the BI sample was non-uniform, some higher NCR bins have very few chips, resulting in low failure rates. But for all practical purposes, it is safe to assume that NCR values higher than 10 would have high failure rate.

The bins with NCR values much less than 1 represent good dice in a bad neighborhood (*spatial dips*). Previous work has shown that the probability of failure of such dice is high [199]. Although several chips having maximum NCR less than 1 pass all post BI tests, such chips are more likely to contain subtle defects and fail sooner.

The bins with NCR values greater than 10 essentially represent bad dice in good neighborhoods (*spatial peaks*). Such spatial outliers can be easily identified by the NCR test method. As Fig. 83 indicates the probability that such chips will pass BI falls with higher NCR values. Many of these chips are healers and hence unreliable.

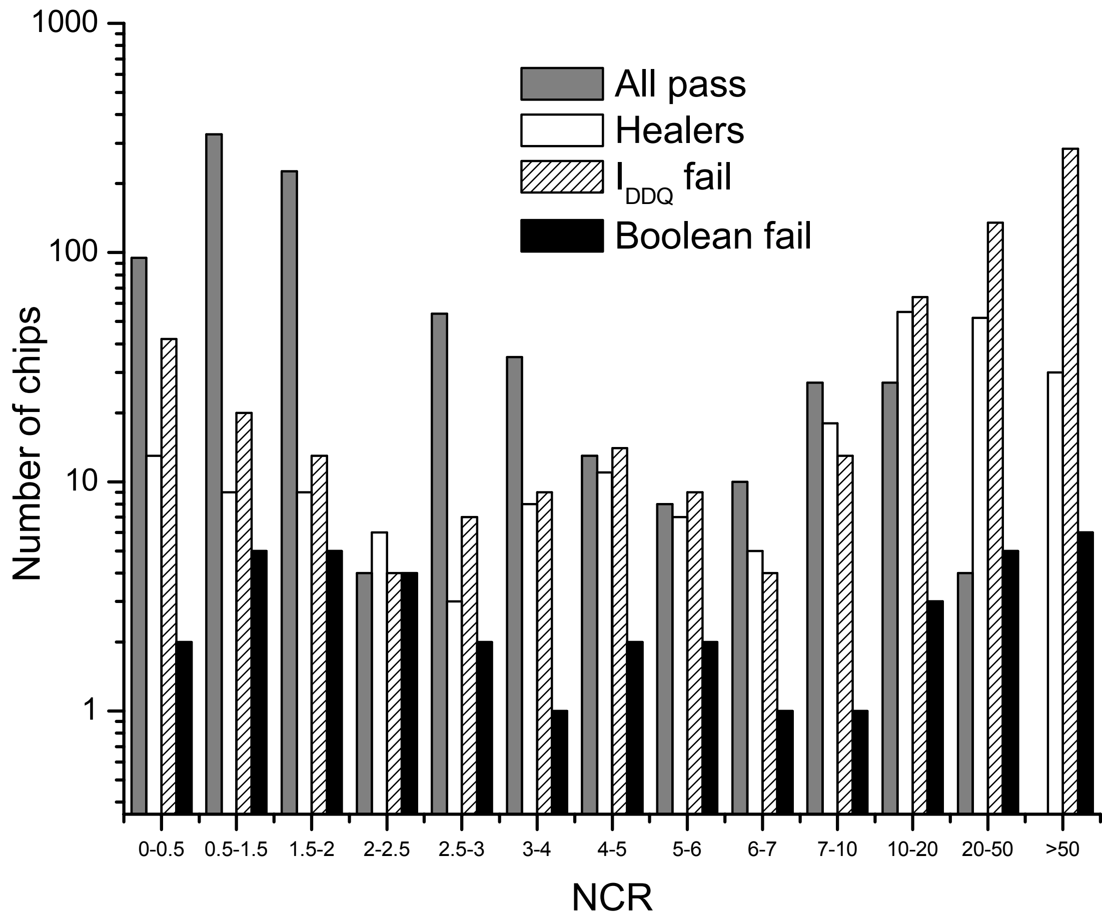


**Fig. 83. Distribution of post-BI results of chips for different maximum NCR values**

NCR test seems to rejects more $I_{DDQ}$-only failed chips than the CR test. These are mostly passive defects that cannot be screened with CR alone effectively. Another metric like flush delay can be combined with NCR values to reduce the yield loss. Even if a single die in the neighborhood yields high NCR for the center die, the center die should be regarded as defective. If this causes unacceptable yield loss, such dice could be selectively burned in.

6.6 Immediate Neighbor Difference $I_{DDQ}$ Test (INDIT)

We used IBM/SEMATECH data for evaluating INDIT. We limited our analysis to the 6-hrs BI sample. We screened chips that fail functional, stuck-at or delay tests at the wafer level. We

also screened chips that had $I_{DDQ}$ more than 100 μA for any vector. Chips having leakage current above 100 μA were assumed to contain a gross defect. These chips appear in the tail of the distribution and are screened due to reliability concerns.

The total number of chips in the data set is 1941. The distribution of these wafer level and post-BI results of these chips is shown in Table XIX. A total of 225 wafer-level $I_{DDQ}$-only failed chips pass $I_{DDQ}$ test after BI, exhibiting a healing defect. These chips are unreliable and are rejected in the test flow.

**TABLE XIX. DISTRIBUTION OF WAFER TEST AND POST BI RESULTS OF CHIPS IN THE DATASET**

| Wafer Test Result | Post BI result | | |
|---|---|---|---|
| | All Pass | $I_{DDQ}$ Fail | Other Fail |
| All pass | 1052 | 27 | 19 |
| $I_{DDQ}$ fail | 225 | 598 | 20 |

### 6.6.1 Deciding Pass/Fail Criterion

The comparison with neighboring chips makes identification of outlier (defective) chips easier. However, the appropriate threshold for neighbor-delta for rejecting defective devices must be determined.

Fig. 84 shows the cumulative distribution of maximum self-deltas and maximum neighbor-deltas. A total of 8 chips that passed all tests and had negative maximum neighbor-deltas (good chips in bad neighborhood) are not included while plotting the neighbor-delta CDF. Fig. 84 shows that 90% of maximum self-deltas are less than 16 μA. This point also marks a beginning of the tail of distribution. We therefore selected a self-delta pass/fail threshold of 16 μA. Fig. 85 highlights the distinction between CDFs for self and neighbor-deltas for chips that pass all tests and fail $I_{DDQ}$-only test at the wafer level. Chips that pass all tests have a noticeable sharp rise in both the CDFs since they have similar $I_{DDQ}$ values. Although by definition, $I_{DDQ}$-only failed chips have at least one reading greater than 5 μA, notice that ~5 % $I_{DDQ}$-only failed chips have maximum self-deltas less than 1 μA and neighbor-deltas less than 5 μA.

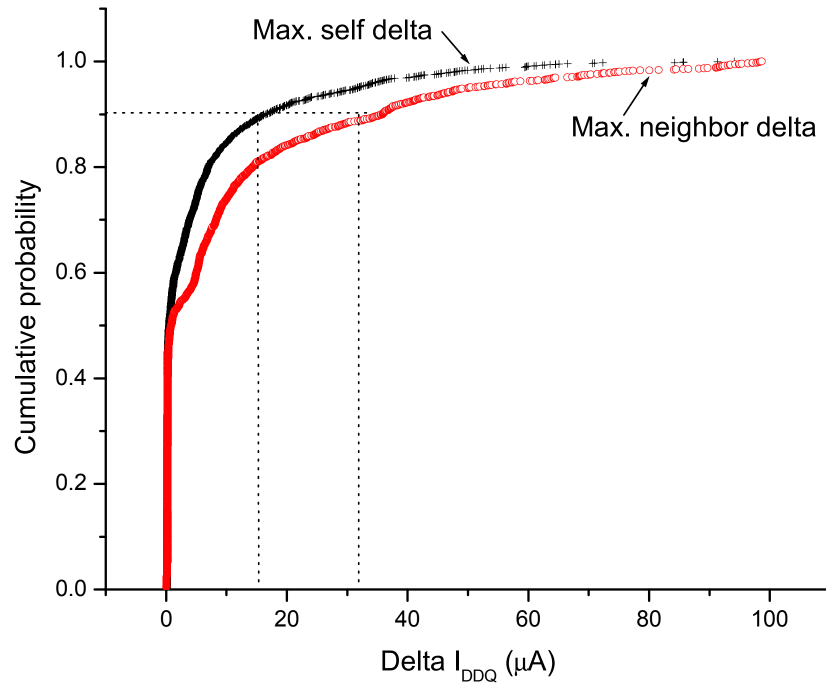


**Fig. 84. CDFs for self and neighbor deltas.**

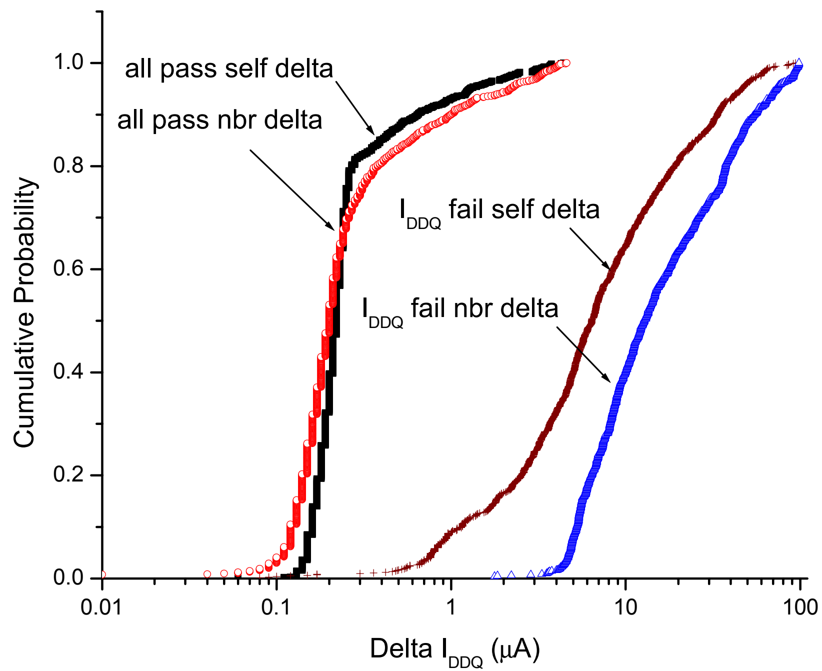


**Fig. 85. CDFs for chips with different wafer probe test results.**

The defect level (DL) and overkill were computed by observing post-BI results as follows:

$$DL = \frac{Number\ of\ accepted\ chips\ that\ fail\ boolean\ tests\ after\ BI}{Total\ number\ of\ chips\ accepted} \quad (6.10)$$

$$Overkill = \frac{Number\ of\ rejected\ chips\ that\ pass\ all\ tests\ after\ BI}{Total\ number\ of\ chips\ rejected} \quad (6.11)$$

The healers are not counted while computing overkill. To compare the effectiveness of the INDIT procedure in screening defective chips not detected by self-delta, the neighbor-delta threshold was varied so as to achieve nearly the same DL as self-delta. Due to the discontinuous distribution of neighbor-deltas, it is difficult to match DLs exactly. The neighbor threshold of 60 μA yielded the closest obtainable DL. Table XX shows the distribution of the chips in the two methods.

**TABLE XX. DISTRIBUTION OF CHIPS ACCORDING TO SEMATECH TEST RESULTS FOR SELF-DELTA AND NEIGHBOR-DELTA TEST METHODS**

| Wafer Probe Result | Self-delta accept | | Self-delta reject | | Post BI result |
|---|---|---|---|---|---|
| | N-delta accept | N-delta reject | N-delta accept | N-delta reject | |
| All pass | 1052 | 0 | 0 | 0 | All Pass |
| | 27 | 0 | 0 | 0 | $I_{DDQ}$ Fail |
| | 19 | 0 | 0 | 0 | Other Fail |
| $I_{DDQ}$ fail | 206 | 1 | 13 | 5 | All Pass |
| | 394 | 24 | 138 | 42 | $I_{DDQ}$ Fail |
| | 17 | 1 | 2 | 0 | Other Fail |

Table XXI shows DL and overkill values for these two methods. Since both thresholds are quite loose, none of the chips from the “All pass” category get rejected by either method. The distinction appears for $I_{DDQ}$-only failed chips. We ignore healers from the analysis due to their reliability concerns. Both methods accept a majority of $I_{DDQ}$-only failed chips. The agreement in the both metrics indicates that these chips are likely to be fault-free.

**TABLE XXI. DL AND OVERKILL COMPARISON OF TWO METHODS**

| Method | DL (%) | Overkill (%) | Yield % |
|---|---|---|---|
| Self-delta | 2.13 | 99 | 89.7 |
| Nbr-Delta | 2.03 | 98.6 | 96.2 |

Fig. 86 shows the distribution of post-BI results of chips for different maximum neighbor-delta values. A high percentage of chips having small neighbor-deltas (<2 μA) pass all post-BI tests. The healers do not exhibit any specific trend. A majority of chips having large neighbor-deltas (>10 μA) fail $I_{DDQ}$ test or Boolean tests after BI.

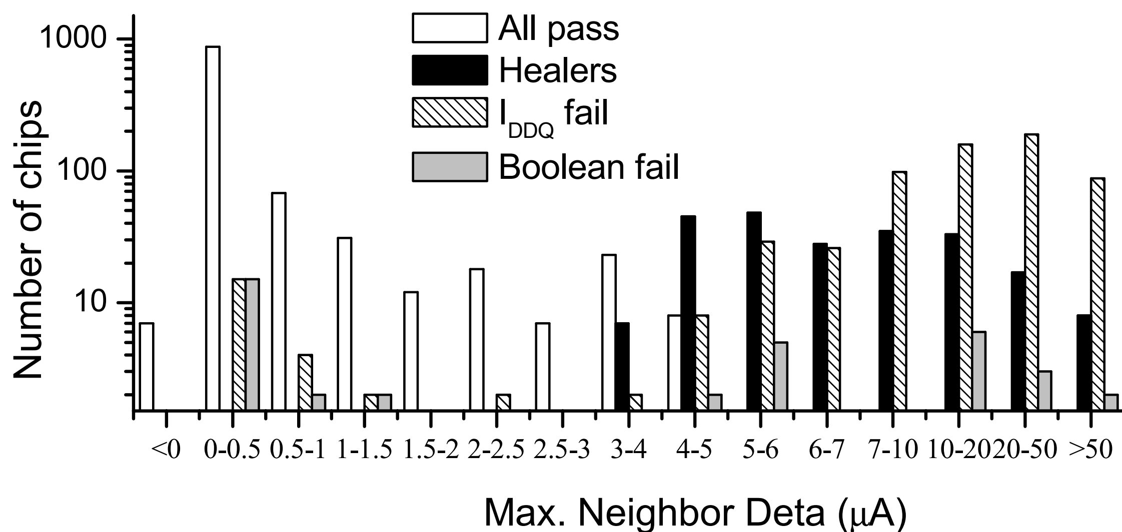


**Fig. 86. Distribution of post-BI results of chips for different maximum neighbor-delta values.**

## 6.7 Combination of CR and NCR

The scatter plot of CR and NCR values for IBM/SEMATECH chips that passed all tests or failed only $I_{DDQ}$ test at wafer probe is shown in Fig. 87. Chips having $I_{DDQ}$ of more than 100 μA are rejected as gross outliers and are not shown. Both CR and NCR values show long tails due to outliers. Considering more than an order of magnitude intra-die variation unlikely for a fault-free chip (since the SEMATECH test chip was $I_{DDQ}$ testable), chips can be divided in four regions as

shown in Fig. 69. The thresholds can be determined by observing the distributions of CR and NCR. Note that many chips that appear to be inliers with CR alone are outliers when NCR values are considered.

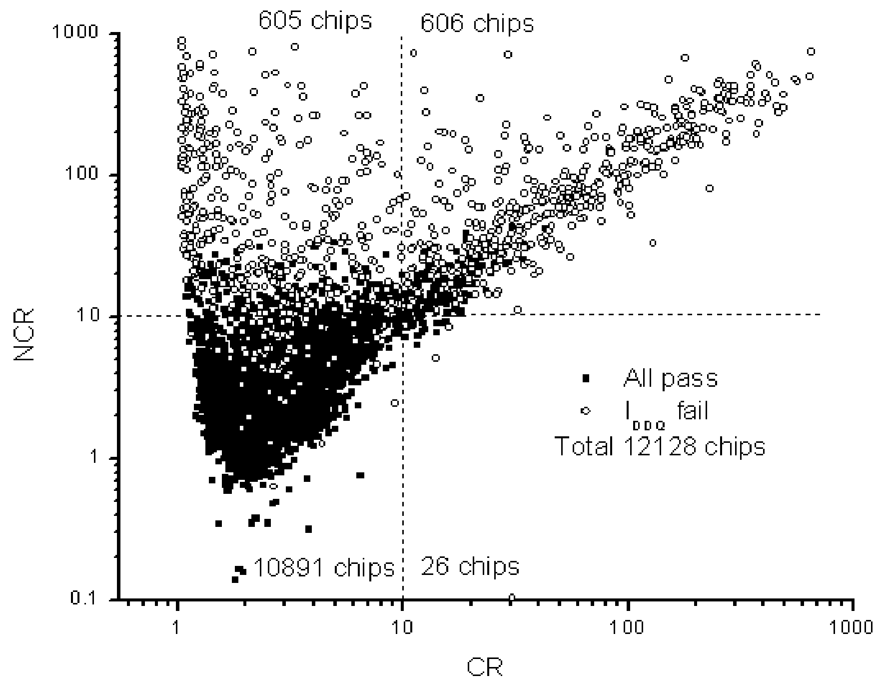


**Fig. 87. Scatter plot of CR/NCR values for all pass or $I_{DDQ}$-only fail SEMATECH chips.**

CR essentially considers intra-die leakage but comparison with other chips is implicit through the characterization process. It does not consider wafer-level variations at the test application time. Since several orders of magnitude variation is observed across wafers it is necessary to understand whether the increased intra-die leakage (high CR) is due to intra-die process variation and whether it is within the "tolerable" bound for a given wafer region. NCR achieves this by estimating inter-die variation using immediate neighboring chips to reveal nonconformance to the local neighborhood. If every vector activates a defect to some degree, intra-die variance in $I_{DDQ}$ can remain relatively low. In this case, the chip can have nominal CR but is likely to have high NCR.

Another CR/NCR scatter plot for LSI test data is shown in Fig. 88. The distribution of CR and NCR values is shown in Fig. 89 and Fig. 90, respectively. CR depends on minimum and maximum $I_{DDQ}$ current of a chip. The minimum $I_{DDQ}$ stems from intrinsic leakage and is

expected to have lognormal distribution. The maximum $I_{DDQ}$ is dependent on the nature of a defect and therefore does not fit any standard distribution. As a result, the CR distribution cannot be described by any standard distribution. In general, both CR and NCR distributions for chips from different wafers/lots have a long tail due to outliers. The thresholds are normally set so as to limit the number of chips accepted from the tail of the distribution and are always a trade off between quality and overkill. Due to the rapid changes in CR frequency near 1 and the long tail, the normality assumption of Chauvenet's criterion, the Tukey method or similar outlier rejection methods can result in tremendous overkill. For Fig. 87 the conventional $\mu+3\sigma$ thresholds would be 1.83 for CR and 6.43 for NCR. The skew and long tail of the distributions make it difficult to find normalizing transforms. A threshold of 1.8 for CR would reject 8643 (<1%) chips. A threshold of 6.5 for NCR would reject 1777 (0.1%) chips. We chose a slightly looser threshold for NCR than CR based on the assumption that CR will catch most defective chips, and the NCR threshold primarily targets more subtle and passive defects. For a total of 736 chips, both CR and NCR values are above the respective thresholds.

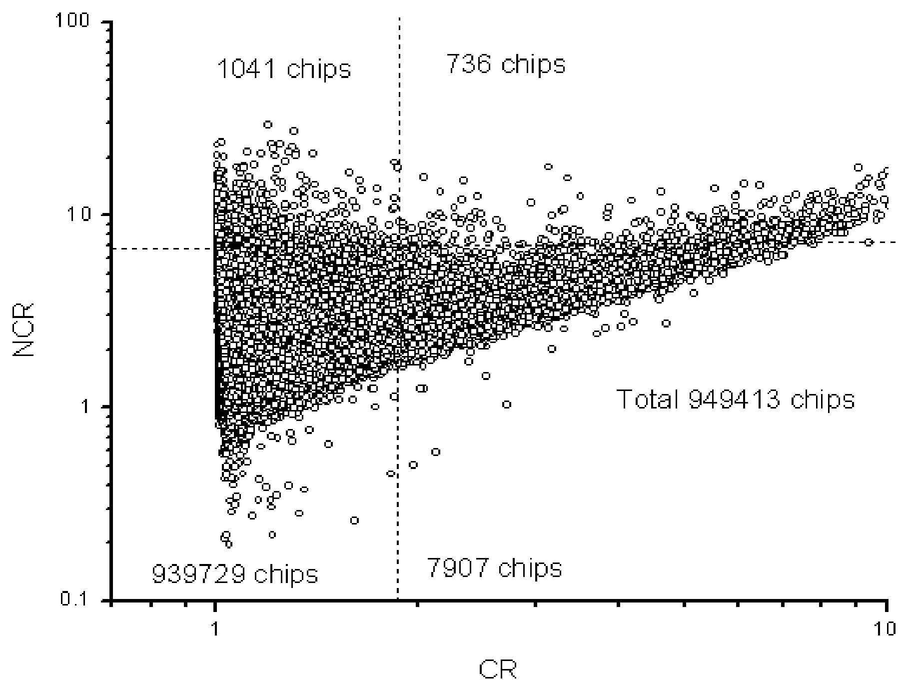


**Fig. 88. CR/NCR scatter plot for LSI data showing number of chips in each region.**

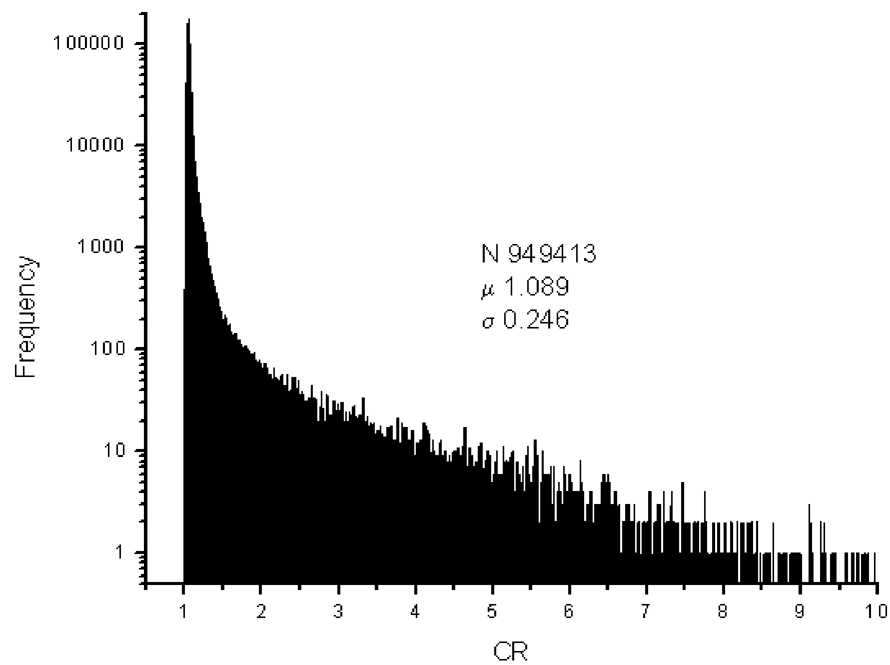


**Fig. 89. Distribution of CR values for LSI data.**

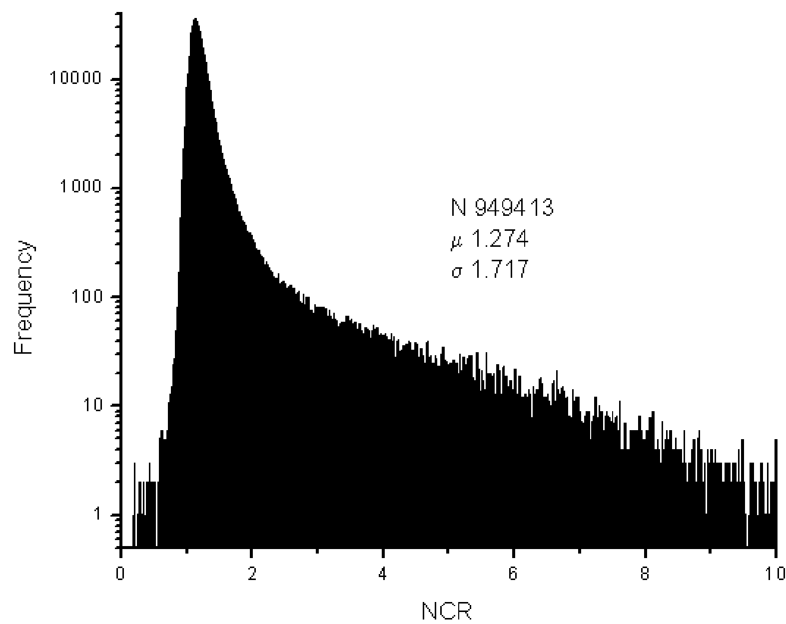


**Fig. 90. Distribution of NCR values for LSI data.**

It is relatively easy to reject chips with a strong active defect, as their CRs are high. Such chips also have high NCRs since high NCR requires only a single neighbor with one fault-free $I_{DDQ}$ reading. Notice that in Fig. 87 and Fig. 88 NCR is mostly in agreement with CR for chips having elevated leakage due to an active defect in the upper right region.

The presence of a passive defect lowers the CR or does not change it appreciably depending on the relative magnitudes of background leakage and defect currents. It was shown earlier that as the relative magnitude of passive defect current compared to intrinsic $I_{DDQ}$ increases, CR approaches 1. Use of a high threshold alone for CR would pass such devices. To reject such chips it is necessary to set a lower threshold for CR. This is difficult due to the steepness of the low end of the CR distribution, as shown in Fig. 91. Chips with a passive defect are likely to have high NCR since the probability that all neighboring chips have similar defective current for *all* vectors is small. The distribution of NCR for chips having CR less than 1.02 (6909 total) is shown in Fig. 92. Some of these chips must be passive defects and (assuming at least one fault-free neighboring vector) appear in the tail of the NCR distribution. A total of 447 chips show NCR above 6.5. Rejecting chips with passive defects is relatively easier with NCR. Chips with low NCR and low CR are likely fast chips. Of course, threshold setting is still a challenge. It is also possible to set a lower NCR limit for low CR chips, on the assumption that they are already suspicious. From Fig. 90 chips with NCR above 2 or 3 appear to be outliers.

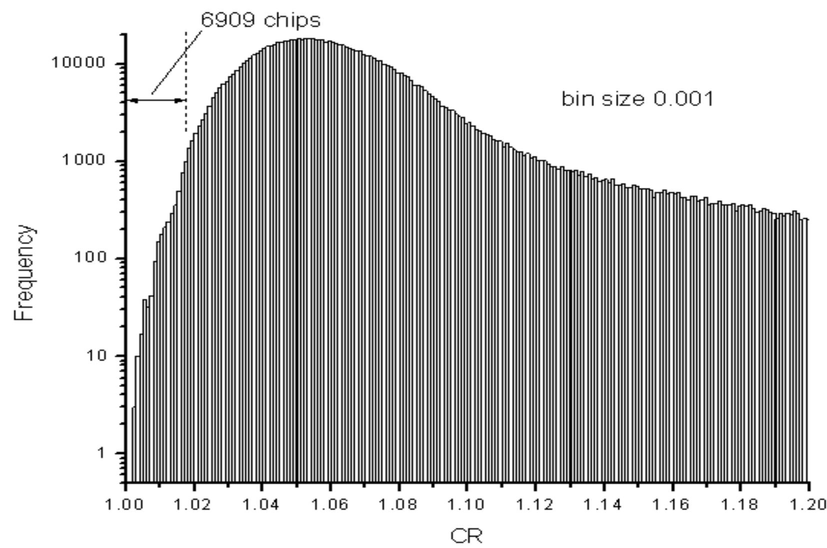


**Fig. 91. Distribution of CRs in the range 1-1.02 for LSI data.**

Since NCR is a relative metric, the reliability of its prediction depends on the degree of correlation between the chip and its neighbors. When a defective chip is surrounded by fault-free chips NCR provides a reliable (conclusive) answer for accepting or rejecting an otherwise suspicious chip. Unfortunately, any relative metric can be misled by the presence of outliers in the data. Many defects on a wafer are known to form clusters [205]. Thus if the neighboring chips are defective, NCR can mislead the outlier rejection, thereby increasing test escapes. But since the maximum NCR is used, this can only happen if all neighbors are similarly defective. Studies of systematic wafer-level variation can be helpful to identify neighboring chips that are better estimators. Note that these chips need not be adjacent and could be at longer distances. Further, note that NCR alone is not capable of screening all outliers (see Fig. 87 and Fig. 88), although there is a high probability that a CR outlier is also an NCR outlier. Therefore, CR and NCR should be used together. Correlating additional parameters can improve confidence or resolve disagreement between the two metrics as we will describe later [98].

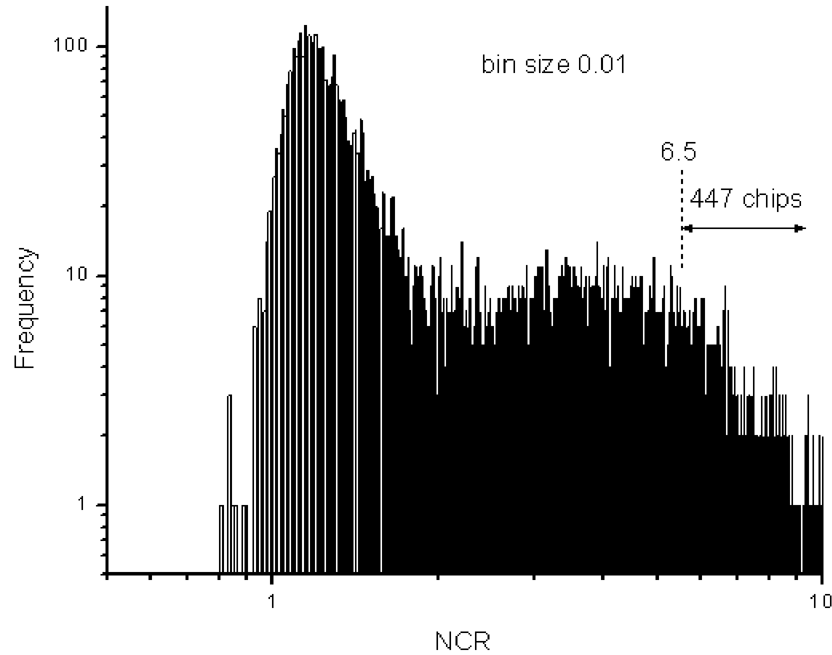


**Fig. 92. Distribution of NCRs for LSI chips having CR less than 1.02.**

We used IBM/SEMATECH data to confirm our hypothesis underlying the categorization shown in Fig. 69 in the previous section. We considered chips [206] that passed all tests or failed only $I_{DDQ}$ test at the wafer level (12128 chips). The chips having leakage current of more than 100 μA were considered gross outliers and discarded. NCR values were computed by considering all 195 vectors for each available neighbor that passed all wafer tests or failed only $I_{DDQ}$ test and the maximum of NCR values was used. The distributions of CR and NCR values are shown in Fig. 93. The gross CR and NCR outliers were discarded before plotting these distributions. As the distributions fall off very quickly after 1σ, the CR and NCR thresholds were selected ($\mu+\sigma$) as 4.81 and 5.67, respectively. A lower CR threshold of 1.1 was selected so that none of the all-pass chips from the BI sample are in the 'passive defect' region in Fig. 69. The distribution of chips in five categories (passive, A, B, C, D) is shown in Table XXII.

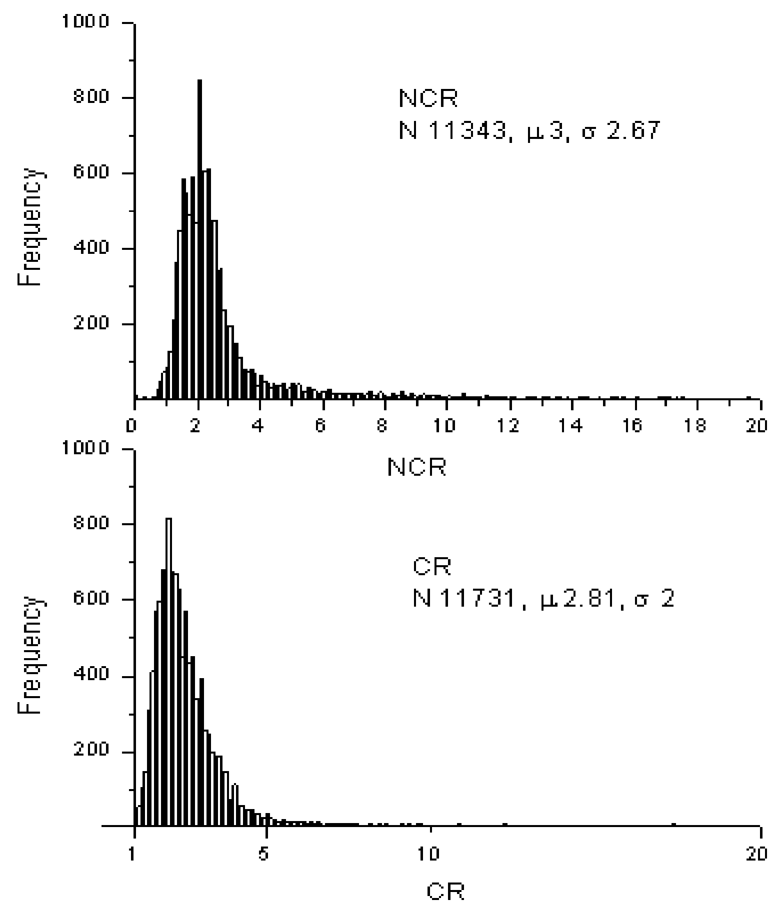


**Fig. 93. Distributions of CR and NCR values less than 20.**

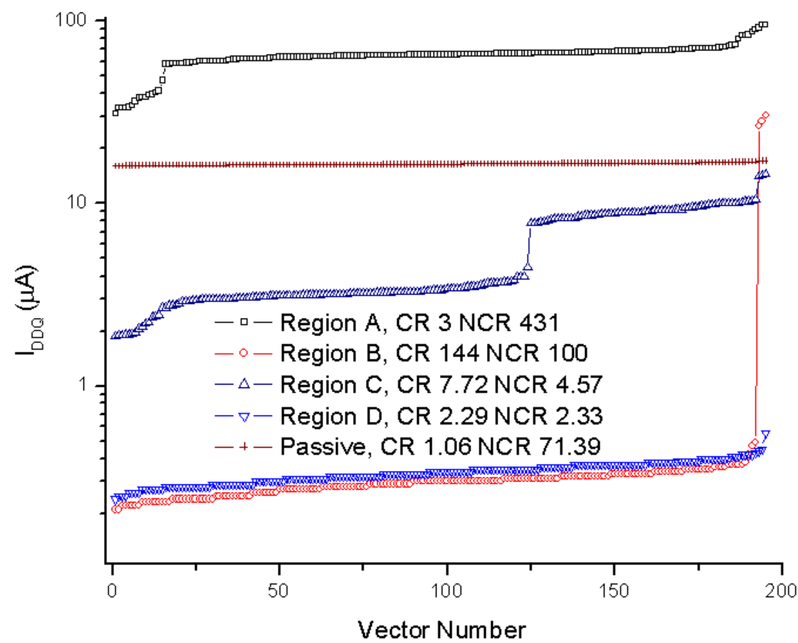


**Fig. 94. Current signatures for die from regions shown in Fig. 69.**

**TABLE XXII. DISTRIBUTION OF SEMATECH CHIPS.**

| Category (1941BI + 10187 no BI) | Wafer Probe Result | | Post BI Result |
|---|---|---|---|
| | All Pass (11220) | $I_{DDQ}$ Fail (908) | |
| Passive Defects (33+7) | - | - | All Pass |
| | - | 31 | $I_{DDQ}$ Fail |
| | - | 2 | Boolean Fail |
| | 1 | 6 | No BI |
| Region A (337+465) | 55 | 86 | All Pass |
| | 3 | 184 | $I_{DDQ}$ Fail |
| | 2 | 7 | Boolean Fail |
| | 447 | 18 | No BI |
| Region B (551+416) | 48 | 119 | All Pass |
| | 5 | 369 | $I_{DDQ}$ Fail |
| | - | 10 | Boolean Fail |
| | 380 | 36 | No BI |
| Region C (19+125) | 14 | 1 | All Pass |
| | - | 3 | $I_{DDQ}$ Fail |
| | - | 1 | Boolean Fail |
| | 125 | - | No BI |
| Region D (1001+9174) | 935 | 19 | All Pass |
| | 19 | 11 | $I_{DDQ}$ Fail |
| | 17 | - | Boolean Fail |
| | 9169 | 5 | No BI |

Fig. 94 shows the current signatures for a sample chip from the SEMATECH data from each of five regions shown in Fig. 69. Notice that chips in regions A, B and C all show the presence of an active defect with a varying degree of severity. The chip from region D shows a smooth fault-free signature. A die with a passive defect also shows similar smooth signature.

The SEMATECH BI test data can give us some insight about the reliability of the chips in the different regions. Chips in regions A, B, C and the passive defects are assumed to be fatally flawed and rejected. Chips in region D are assumed to be good. The defect level (DL) of the shipped lot (region D) and the yield loss (YL) incurred for rejecting the other regions are computed by using the following equations:

$$DL = \frac{Number\ of\ post-BI\ Boolean\ fails}{Total\ number\ of\ chips\ in\ the\ category}.100 \tag{6.12}$$

$$YL = \frac{Number\ of\ chips\ that\ pass\ all\ post-BI\ tests}{Total\ number\ of\ chips\ in\ the\ category}.100 \tag{6.13}$$

### 6.7.1 Chips that pass all tests at both levels

Since the SEMATECH definition of “all pass” implies $I_{DDQ}$ less than 5 µA, a majority of all pass chips have small CR and NCR values as shown in Fig. 95. However, it is interesting to note that some chips exhibit high CR and/or NCR. Even though these chips did pass all tests, they are a potential reliability risk. This underscores the fact that thresholds are getting fuzzier and “different behavior” is reason enough to reject chips [207].

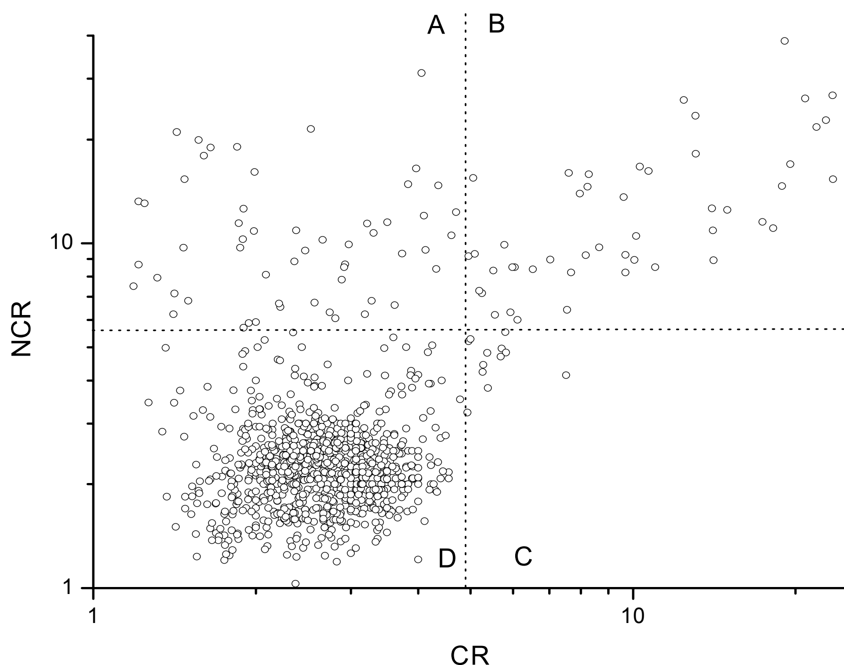


**Fig. 95. CR/NCR scatter plot for SEMATECH chips that passed all tests.**

Chips that failed only $I_{DDQ}$ test at wafer and post-BI test are considered either faulty or fault-free. Healer chips are considered defective. The results were then scaled up to the entire population based on the distribution of chips used in the BI sample. There are too few passive defects or region C chips to perform an analysis and so are not considered further. The results of the analysis are shown in Table XXIII. The results support our hypothesis that chips from region B are more likely to be defective than chips from other regions, but chips from region A also have a high defect level.

**TABLE XXIII . DL AND YL FOR VARIOUS CATEGORIES.**

| Category | Defect Level (%) | | Yield Loss (%) |
|---|---|---|---|
| | $I_{DDQ}$ fail = Fault-free | $I_{DDQ}$ fail = Faulty | |
| Region A | 3 | 58 | 15.9 |
| Region B | 1.8 | 69 | 8.38 |
| Region D | 1.69 | 4.69 | NA |

It is interesting to see the trends in CR and NCR values of the following sub-categories: chips that pass all tests at both levels, chips that fail only $I_{DDQ}$ test at both levels and chips that fail Boolean test after BI (independent of their wafer probe result). Fig. 95 through Fig. 97 show CR/NCR scatter plots of these chips.

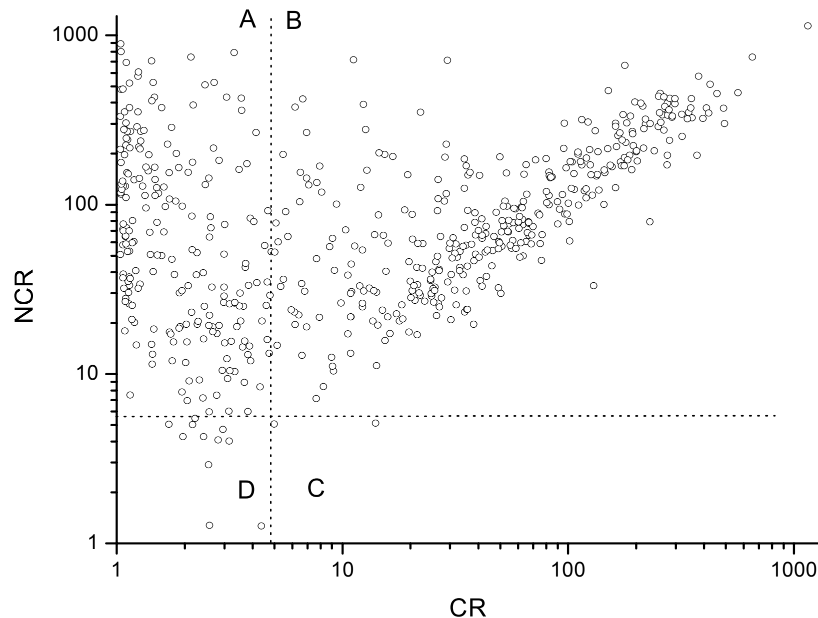


**Fig. 96. CR/NCR scatter plot for chips that failed only $I_{DDQ}$ test before and after BI.**

### 6.7.2 Chips that fail only $I_{DDQ}$ test

It is known that the $I_{DDQ}$ threshold of 5 μA was decided by empirical analysis and is not necessarily a good manufacturing limit. Understanding the behavior of chips that fail only $I_{DDQ}$ test has been a topic of interest [208] due to yield loss [91] as well as reliability issues [55], the latter becoming more important with technology scaling. SEMATECH chips show strange behavior in this regard. Chips spread over several orders of magnitude of NCR fail only $I_{DDQ}$ test (see Fig. 96). Most of these chips are in regions A and B and are likely to have active defects as depicted by current signatures in Fig. 94. Some of these chips may exhibit healing behavior later after extended BI [116], but for all practical purposes, these chips are a reliability risk and must be rejected up front in the test flow. Since such rejection at the wafer level in BI avoidance mode can result in excessive yield loss, the thresholds for CR/NCR must be appropriately selected.

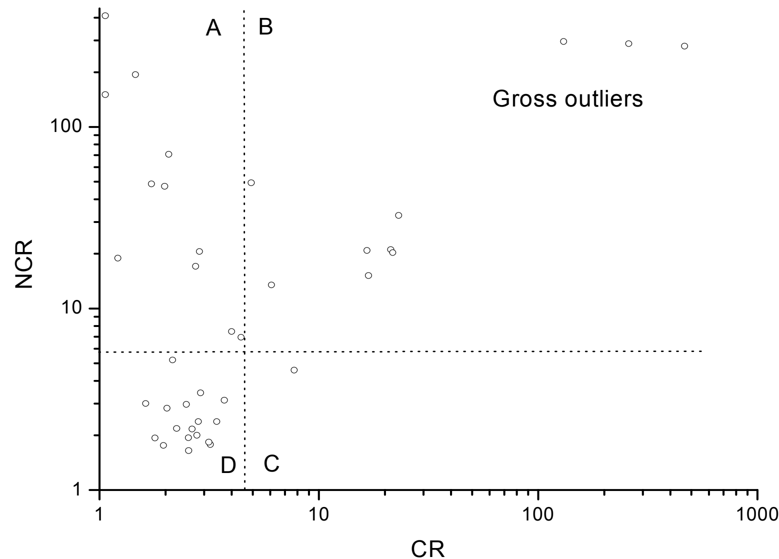


**Fig. 97. CR/NCR scatter plot for chips that failed Boolean test after BI.**

### 6.7.3 Chips that fail Boolean test after BI

Fig. 97 shows the CR/NCR scatter plot for chips that fail a Boolean test after BI. Such chips from region A generally have high NCR. This can occur if a chip has multiple active defects or a combination of active and passive defects resulting in a shift in $I_{DDQ}$ values, while limiting CR.

Chips that fail Boolean tests from region D exhibit random variation in NCR. A similar random nature is also observed for chips from region B. Unfortunately the SEMATECH BI sample was biased and the data set is too limited to draw meaningful conclusions in this regard. Some of the defects that lead to these Boolean fails are not $I_{DDQ}$-testable and, therefore, CR/NCR values are close to their mean values.

## 6.8 Combination of CR, NCR and Delay Data

To evaluate the effectiveness of an additional test parameter to the CR-NCR combination, we used SEMATECH test data. The analysis is carried out using wafer level test data. Functional and stuck-at fail are screened as their parametric test data is not reliable. We also screen gross $I_{DDQ}$ outliers. The threshold for gross outlier rejection is obtained from cumulative distribution of test data. We used a 100 µA limit for gross outlier screening. The CR is obtained by taking the

ratio of the maximum and minimum $I_{DDQ}$ values for each chip. The chips are then screened using NCR values. Chip flush delay is used for further screening. A flush delay threshold of 500 ns is used by observing the distribution shown in Fig. 98. The analysis flow shown in Fig. 70 was followed.

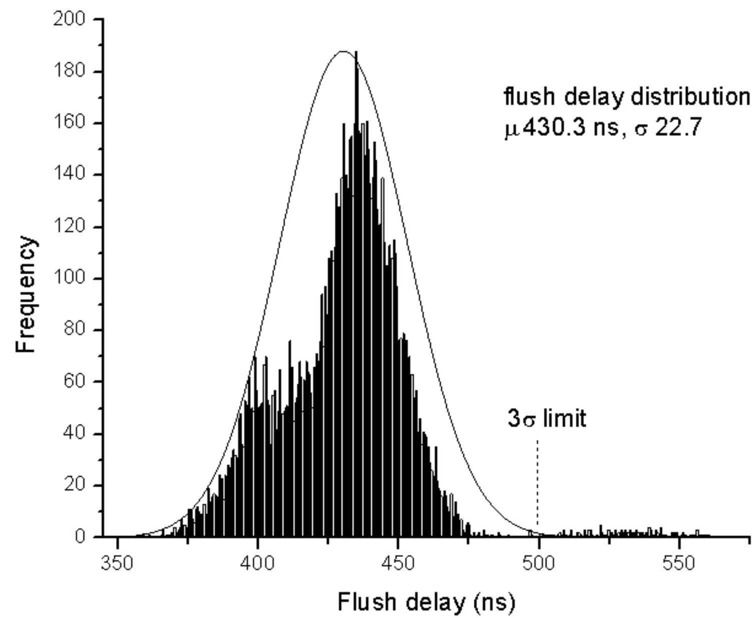


**Fig. 98. Flush delay distribution for the SEMATECH sample.**

The cumulative distributions for CR and NCR values are shown in Fig. 99 and Fig. 100, respectively.

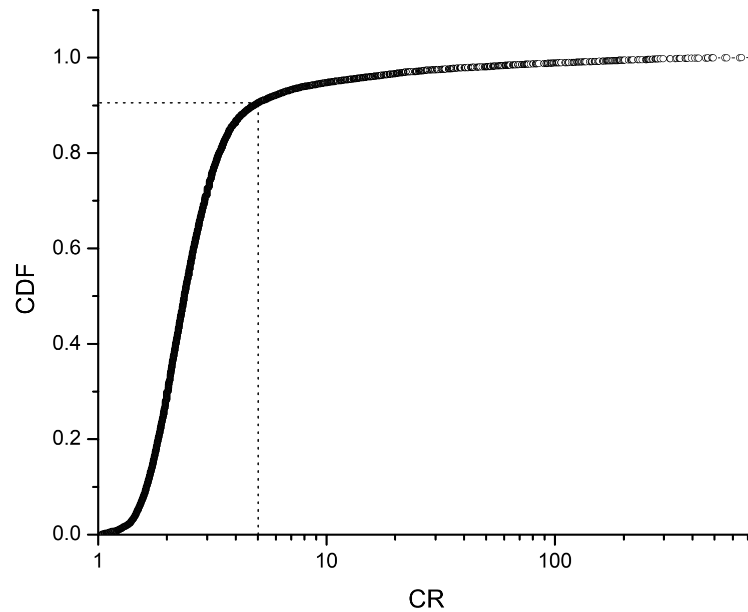


**Fig. 99. Cumulative distribution for CR values for the SEMATECH sample.**

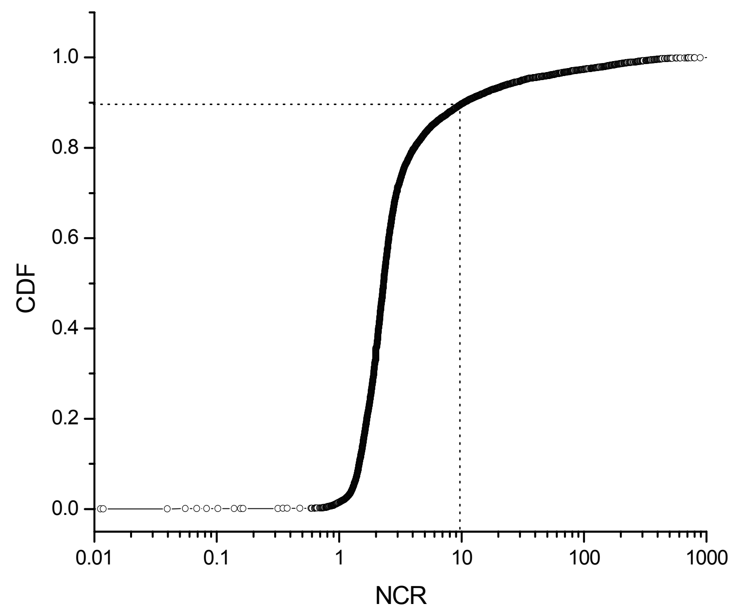


**Fig. 100. Cumulative distribution of NCR values for the SEMATECH sample.**

The characteristic long tail of a logarithmic distribution in $I_{DDQ}$ data is observed in the NCR distribution as well. The distribution of chips in the data set according to their wafer and post-BI

test results is shown in Table XXIV. They are further categorized based on their CR and NCR values using thresholds obtained from the respective cumulative distributions.

**TABLE XXIV. DISTRIBUTION OF SAMPLE ACCORDING TO CHIP TEST RESULTS (12521 CHIPS)**

| Wafer Probe | Post BI Result | CR < 5 | | CR > 5 | |
|---|---|---|---|---|---|
| | | NCR < 10 | NCR > 10 | NCR < 10 | NCR > 10 |
| AP | AP | 968 | 28 | 32 | 27 |
| | IF | 20 | 2 | 2 | 3 |
| | DF | - | - | - | - |
| | IDF | - | - | - | - |
| | BF | 19 | - | - | - |
| | NB | 9552 | 126 | 254 | 198 |
| IF, IF or IDF | AP | 196 | 80 | 89 | 125 |
| | IF | 29 | 207 | 4 | 372 |
| | DF | 2 | 8 | 3 | 8 |
| | IDF | 1 | 2 | - | 3 |
| | BF | 12 | 11 | 7 | 12 |
| | NB | 45 | 28 | 3 | 43 |

AP: All Pass, IF: $I_{DDQ}$ Fail, DF: Delay Fail, IDF: $I_{DDQ}$ + Delay Fail,
BF: Voltage (Boolean) Fail NB: No BI

### 6.8.1 Analysis of CROWNE Chips

Current Ratio Outlier With Neighbor Estimator (CROWNE) chips are those outliers that cannot be identified with current ratio alone, but can be detected when neighboring dice information is used. Most likely these chips contain passive defects that results in small NCR. The addition of spatial information is useful for detecting such defects. Such defects can escape the NCR screen only if all neighboring chips have similar or higher defective currents for *all* vectors. Although due to defect clustering there is increased possibility that neighboring chips are defective as well, there is no empirical evidence that all neighboring chips have similar defective currents for all vectors. Since the maximum of all NCR values is used for screening, the NCR screen is more sensitive to defects than CR. Chips with very low CR values (~1) have a strong passive defect. These chips will not cause functional failure at the wafer test, but may fail in the system. There are 138 chips with a CR of 1-1.5 but only one of these chips has a NCR<10.

Fig. 101 shows the scatter plot of CR and NCR values for these chips. Detection of passive defects with CR alone may be feasible by setting a lower bound on CR. However, as shown in the previous subsection, due to the sharp fall of the CR distribution such threshold setting is very difficult and results in excessive yield loss.

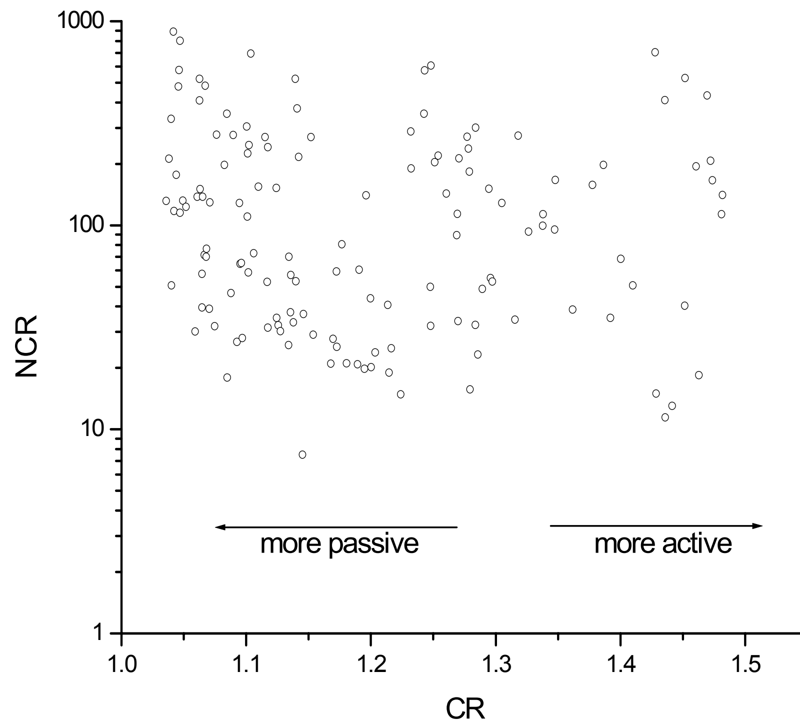


**Fig. 101. CR-NCR scatter plot for CROWNE chips.**

A total of 336 chips fail only delay tests and 43 chips fail both delay and $I_{DDQ}$ tests. Fig. 102 shows the CDF and distribution of flush delays for these chips and Fig. 103 shows the scatter plot of CR and NCR values. Some of these chips are NCR outliers although they did not fail $I_{DDQ}$ test. Apparently, these delay failures do not seem to be the result of wafer processing conditions. Delay-only failed chips are likely to be due to resistive open defects that do not lead to an increase in $I_{DDQ}$. However, a few delay-only failed chips do exhibit high CR and NCR values. The poor correlation between NCR and flush delay (Fig. 104) indicates that NCR alone cannot be used to screen all chips. The poor correlation can also be explained with the help of the self-scaling nature of NCR. For chips from a fast or leaky wafer region, NCR values scale accordingly. However, for defective chips, NCR values show poor correlation. This observation can be helpful for screening potential delay defects.

## 6.9 Conclusion

The central theme explored in this section is to use the *same* test data combined with statistical post-processing to identify outliers. We showed that some metrics are better than others. For example, NCR catches passive defects more easily than CR or delta $I_{DDQ}$. However, no single test method is a clear winner. This does not come as a surprise as we know that each method is trying to *see* data in some a way other methods do not. Combining multiple test methods is therefore useful in finding hidden outliers in the data and making a distinction between true and apparent outliers.

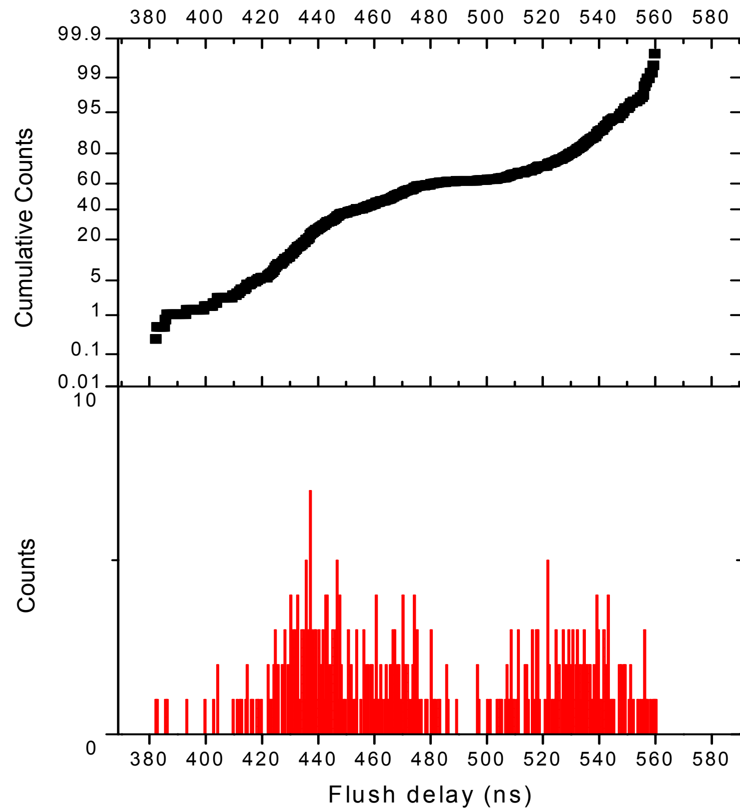


**Fig. 102. Distribution of flush delays for delay failed chips.**

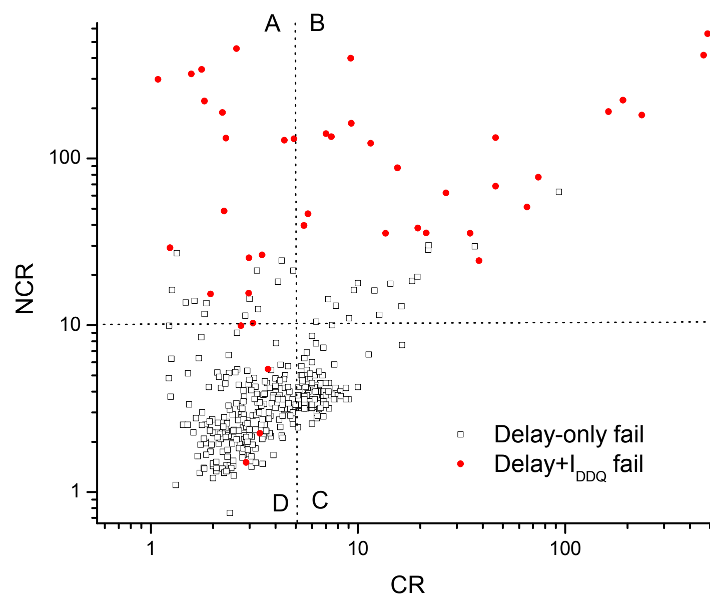


**Fig. 103. CR-NCR scatter plot for delay fail and/or $I_{DDQ}$ fail chips.**

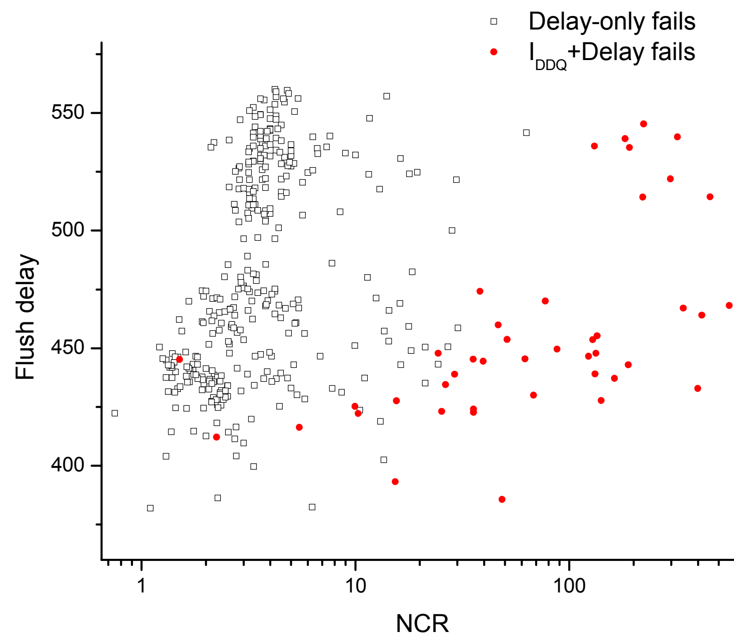


**Fig. 104. NCR and flush delay scatter plot for delay fail and/or $I_{DDQ}$ fail chips.**

# 7. APPLICATION TO BURN-IN REDUCTION

## 7.1 Introduction

As discussed in Section 1, testing is a “cost-avoidance” strategy adopted by semiconductor manufacturers. However, owing to the imperfect nature of tests, some defective chips do get shipped to customers. Although targeting zero DPM is impractical, a reduction in DPM (i.e. improvement in quality) is always sought. It was also mentioned earlier that the traditional method of reducing DPM by screening low-reliability chips before they are shipped to customers is burn-in (BI). As BI loses its effectiveness for DSM chips and BI costs escalate, alternatives to BI are sought. $I_{DDQ}$ test has been proved to have a potential to replace or avoid BI [61-64]. The relation between yield and BI failure rate has been modeled for high volume IC’s manufactured in several processes [209].

Outlier identification earlier in the test flow can be used for screening low-reliability chips. In Section 4, we mentioned that the test strategy could be altered depending on the “outlier-ness” of a chip in the distribution. For example, outlier chips may be subjected to reduced BI time (in *BI reduction mode*) or simply rejected without BI (in *BI avoidance mode*). Alternatively, inliers (chips conforming to “normal” variation) can be shipped without BI or subjected to reduced BI. This also results in reduction in BI cost. However, in this case note that the confidence in DPM can be achieved only after collecting and analyzing sufficient data. In this section, we evaluate the effectiveness of an outlier identification method for BI reduction.

## 7.2 Methodology

We evaluated the Median of Absolute Deviation (MAD) about medians for BI reduction. This method is evaluated using SEMATECH data (see Appendix A). The distribution of chips before and after 6 hours of BI is shown in Appendix A. Of particular interest here are the chips that pass all tests and those that fail only $I_{DDQ}$ test. Out of 1558 chips that failed only $I_{DDQ}$ test at wafer level, 1219 failed *only* $I_{DDQ}$ test after burn-in. Fig. 105 shows $I_{DDQ}$ for a sample of chips that passed all the tests at wafer probe. Fig. 106 shows $I_{DDQ}$ for a sample of chips that failed only $I_{DDQ}$ test at wafer probe. Note that since wafer probe and post burn-in tests were conducted at 50°C and 25°C respectively, $I_{DDQ}$ readings after BI are expected to be lower than those at wafer probe for a fault-free chip. For $I_{DDQ}$-only fail devices the spread in $I_{DDQ}$ values is clearly noticeable in Fig. 106. While some chips do exhibit appreciably increased $I_{DDQ}$ after burn-in (well above the trend line) and are high-risk devices, chips that did not have significant increase

in $I_{DDQ}$ may not be defective. This is because $I_{DDQ}$ failure simply means $I_{DDQ}$ exceeds 5 μA threshold. Chips having stable $I_{DDQ}$ above this threshold may be fast and leaky chips, not necessarily unreliable. Around 291 of 1558 $I_{DDQ}$-only failed chips at wafer probe heal after burn-in and move to the "all pass" category after burn-in. These unstable devices may be a high reliability risk devices.

### 7.2.1 MAD-based Outlier Rejection

We used the MAD outlier rejection method as discussed in Section 4.11. We considered only chips having six hours of BI data. In practice, no $I_{DDQ}$ data is usually available for functional fails. Therefore, wafer level functional, delay and stuck-at fails were screened. The resulting data set contained 1660 chips that have either passed all the tests or failed only $I_{DDQ}$ test at the wafer level. The gross outliers (several mA of $I_{DDQ}$) were then removed using Chauvenet's criterion [174] with a loose probability threshold of 0.1, following normalizing transform. This roughly corresponded to an $I_{DDQ}$ threshold of 500 μA. We used Chauvenet's criterion rather than MAD-based rejection of gross outliers to avoid bias. For these chips the maximum $I_{DDQ}$ for each vector was less than 500 μA. For each vector a 3σ limit was determined for post BI $I_{DDQ}$ pass/fail decision as described later.

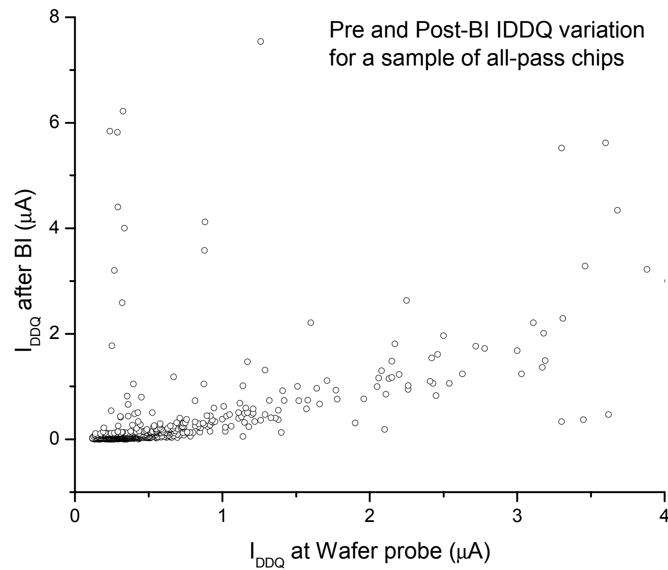


**Fig. 105. $I_{DDQ}$ before and after burn-in for a sample of chips that passed all tests at wafer probe; chips for which post burn-in $I_{DDQ}$ < 8 μA are shown.**

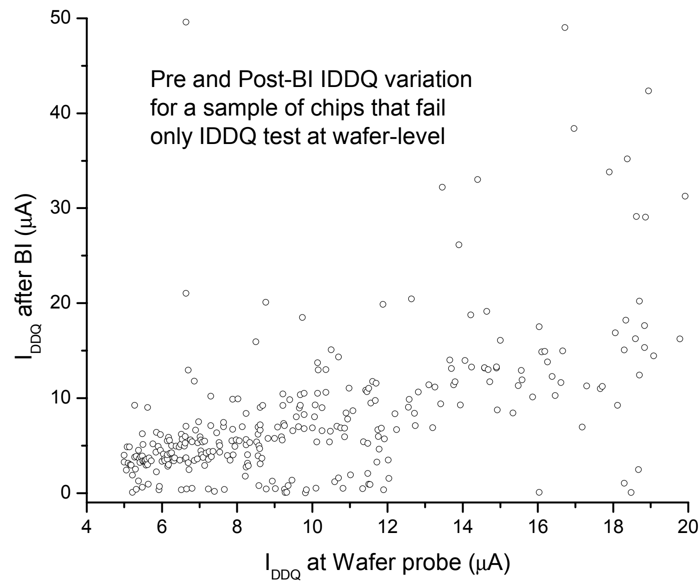


**Fig. 106. $I_{DDQ}$ before and after burn-in for a sample of chips that failed only $I_{DDQ}$ test at wafer probe; chips for which post burn-in $I_{DDQ}$ is less than 50 μA are shown.**

For each die a total of 195 $I_{DDQ}$ measurements are available at probe and after BI. Since the industry practice is to use 10-20 $I_{DDQ}$ measurements, we considered only the first 20 measurements for each die. The $I_{DDQ}$ distribution was converted to a Normal distribution by logarithmic transform as follows. For each vector the minimum nonzero value was found across all chips and all readings were divided by this value. Then the logarithm of the ratio was taken. The median $I_{DDQ}$ for each vector was obtained. Then MAD values for each vector were computed. Using these MAD values, MAD scores were computed for each reading.

### 7.3 Burn-in Pass/Fail Decision

If a chip failed any voltage test after BI it was considered to be defective. A chip passing all SEMATECH tests after BI was considered to be fault-free. Several chips fail only SEMATECH $I_{DDQ}$ (5 μA threshold) test after BI. Considering them all to be defective or defect-free would give misleading results for test escapes and/or yield loss. Hence in order to decide optimum post BI $I_{DDQ}$ threshold we used pre-BI $I_{DDQ}$ variation.

We used the $3\sigma$ limit obtained from $I_{DDQ}$ at wafer probe as the $I_{DDQ}$ pass/fail threshold for post-BI data. However, note that because the wafer probe was conducted at a higher temperature (50$^{o}$C) than the package level test (room temperature), this limit is not stringent. If any $I_{DDQ}$ reading (for the 20 vectors) exceeded this limit, a die was considered to be defective and rejected. There are four possible cases for each chip: (a) accepted chip passes all tests after BI, (b) accepted chip fails any test after BI, (c) rejected chip fails any test after BI and (d) rejected chip passes all tests after BI. Cases (a) and (c) are correct predictions. Cases (b) and (d) are incorrect predictions, the former causing defect level in the shipped lot and the later causing overkill or yield loss. These values were expressed as a percentage of the total number of chips (2534).

The defect level and overkill have an obvious inverse relationship. By tightening the pass/fail limit (e.g. MAD threshold), it is possible to increase the quality of the shipped product. The price paid for quality is increased overkill. The advantage in using a statistical technique like MAD-based rejection is rejection of only “true” outliers that deviate from the median. This causes a smaller change in overkill for a given quality than a static threshold technique. A MAD threshold of 10 was used for the analysis.

## 7.4 Experimental Results

To compare the effectiveness of the MAD-based scheme, we performed a similar analysis with delta-$I_{DDQ}$ and current signatures. Delta $I_{DDQ}$ was defined as the difference between two adjacent readings, thus yielding 19 delta values. We computed the mean ($\mu_\delta$) and standard deviation ($\sigma_\delta$) of deltas for each chip. If the absolute value of any delta exceeded the threshold $\mu_\delta + 3\sigma_\delta$, the chip was rejected. The current signature was obtained by sorting the $I_{DDQ}$ readings and then following a similar approach as delta $I_{DDQ}$. Fig. 107 shows the comparison of these methods. The defect level and yield loss values are shown in Table XXV. DL is expressed as a percentage of accepted chips and yield loss is expressed as a percentage of the total chips.

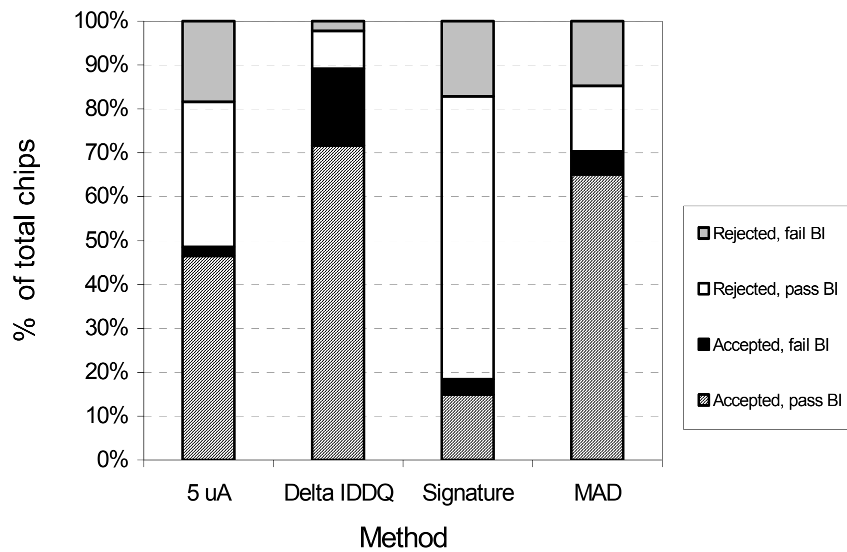


**Fig. 107. Comparison of effectiveness of various methods.**

**TABLE XXV. COMPARISON OF DL AND YL FOR VARIOUS METHODS.**

| Method | DL % | YL % |
|---|---|---|
| 5 μA Threshold | 4.2 | 33.12 |
| Delta-$I_{DDQ}$ | 19.58 | 8.88 |
| Current Signature | 18.7 | 60.87 |
| MAD | 7.4 | 14.19 |

## 7.5 Discussion of Results

Although the single 5 μA threshold approach has a lower DL, the corresponding YL would be clearly unacceptable. Delta-$I_{DDQ}$, on the other hand, has the highest overall yield but high DL as well. For chips having a $V_{DD}$ to ground short, small vector-to-vector variations in $I_{DDQ}$ result in smaller non-zero deltas. Unless an upper static threshold limit is used, these chips are accepted by delta-$I_{DDQ}$. The reduction in bridge resistance after burn-in would eventually cause increased post-BI $I_{DDQ}$. On the other hand some of these chips could exhibit healing behavior. For high reliability requirements, it would be necessary to screen these devices. The MAD scores for these devices would be higher and can reject many of them.

Very low yield of current signature can be explained as follows. After readings are sorted, the mean and variance of deltas are reduced. This results in a lower pass/fail threshold and rejects many chips that do not exhibit increased leakage after BI, thus showing a huge yield loss.

The static-threshold approach is not recommended for BI reduction. The yield loss and defect level figures can vary considerably. The MAD-based approach has comparable yield to delta $I_{DDQ}$ with lower defect level. Chips having passive defects have higher MAD scores for all the vectors and get rejected. MAD is also sensitive to active defects. Since the MAD technique is looking at the entire distribution instead of multiple measurements from a single chip, it has higher resolution in screening defects.

## 7.6 Conclusion

This section presented preliminary analysis of use of a statistical outlier rejection method for BI reduction. Unfortunately, IBM/SEMATECH data comes from an older technology and the BI sample was biased towards chips that failed one or more tests. The BI sample is also too small to draw meaningful conclusions that can be generalized. Due to the stratified nature of the sample that comes from an older technology it is difficult to extrapolate our conclusions to the present or future technologies. However, this analysis indicates that outlier screening has a potential to reduce BI test cost.

# 8. CONCLUSIONS AND FUTURE WORK

## 8.1 Summary

This research has primarily focused on identifying outlier chips using parametric test data. The basic purpose of outlier identification is two-fold. First, it should reduce the outgoing defect level or improve the quality of parts shipped. Secondly, outlier identification should reduce the overall test cost by screening the parts that are *deemed* defective or unreliable. Although a secondary goal of outlier identification is to achieve reliability targets, it is limited by the inherent reliability of the components. In this research, we were particularly interested in identifying those chips that fail Boolean test after they pass initial (wafer-level) screening. Thus, our goal was to identify "flawed" chips and separate out "seemingly flawed" chips. This research mainly exploited wafer-level spatial correlation. We used wafer-level test data in order to screen outlier chips earlier in the test flow to keep the test costs low.

We observed that:

Wafer-level spatial variation is primarily deterministic and can be useful to screen outlier (defective) chips.

Statistical outlier rejection methods are helpful to reject chips having a reliability risk, although extensive studies are needed to evaluate the long-term impact on reliability.

1. Outlier screening methods assume certain "standard" distributions, however, the presence of outliers alters these properties and hence it is difficult to screen some apparent outliers. Different data transformations can be used.
2. Multiple parameter correlation can be helpful to increase outlier screening resolution. It is necessary to understand the underlying physical mechanisms and use appropriate methods.
3. Use of NCR/INDIT can help increase defect-screening resolution without additional testing.
4. In methods like wafer signature, since the same test data is used in a different way, there is no additional cost. Yet, it proves to be an effective method to screen outliers.

Statistical tools provide a powerful means to analyze data and identify outliers. However, they are like a double-edged sword. In order to use them effectively for outlier screening, it is necessary to understand the underlying physics that governs the defect-free variation in the data. When the data does not conform to the requirements (e.g. has a non-Normal distribution), it is

necessary to use the appropriate transforms. Threshold setting is always a debatable issue and does not disappear with the use of statistical methods. However, "normal" variation, by definition, never appears in the tail of a distribution. Using a simple example, we showed how the presence of outliers alters the properties of the standard distribution and successive outlier rejection is not a reasonable strategy to follow.

As parametric test methods lose their resolution, it is necessary to combine multiple parameters. We observed that there are certain defects that cannot be caught by a single test metric. A combination of multiple test metrics like CR and NCR is useful for screening outliers not detected by CR or NCR alone. As transistor geometries scale further, test becomes even more challenging. However, fundamentally the wafer processing sequence still consists of a deposit-pattern-and-etch sequence. This means that even in the nanometer regime, test methods that rely on the underlying physics will remain valuable in the future. Nevertheless, the ability of a test metric to screen outliers will continue to fade and it will be necessary to use multiple test metrics.

## 8.2 Future Directions for Research

There are certain promising research directions based on this work.

### 8.2.1 Exploring vector-to-vector correlation

There is strong vector-to-vector correlation among different chips as they are influenced by similar process changes. Fig. 108 shows that $I_{DDQ}$ readings for vector 1 and vector 2 ($R^2$ = 0.9139) and vector 3 ($R^2$ = 0.9022) are highly correlated. This observation can be exploited to build a model for estimating the maximum parameter value ($I_{DDQ}$). An outlying observation shows poor correlation and can be identified. Wafer-level spatial information can be combined with such a model to account for deterministic wafer-level patterns.

### 8.2.2 Use of factor analysis

The basic goal of factor analysis (FA) is to reduce the number of variables by analyzing the covariance between them. A FA-based model can be built by analyzing the correlation between different die positions on a wafer as well as vector correlations. The residuals between the parameter value estimated by such model and the actual value can be subjected to outlier identification methods. The model should account deterministic variations. Hence, it is possible

to construct a *difference signature* for each die. The difference signature patterns can then be analyzed to spot outliers. It can also provide some insights about the nature of the defect.

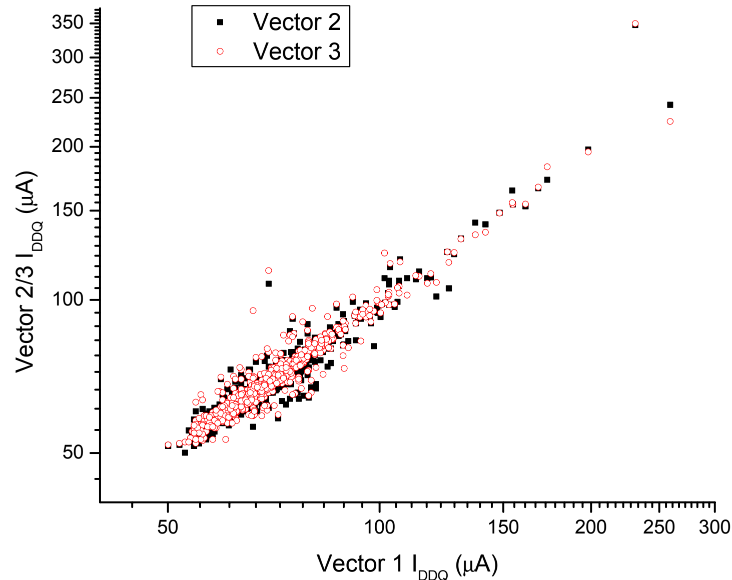


**Fig. 108. Vector-to-vector correlation between different chips.**

### 8.2.3 Evaluation of test cost reduction

Since test is unavoidable, test cost reduction will continue to be a top priority for semiconductor manufacturers. Using optimal mixture of wafer and package level tests is important to gain valuable information from test data while abating test costs [210]. It is possible to build a cost model to analyze the economic benefit of outlier identification. One such cost model is suggested in Appendix F. Using such a cost model it is possible to alter the test strategy depending on the "outlierness" of a chip. Thus, it is possible to BI only those outliers in the "fuzzy" region or when outlier behavior cannot be conclusively determined.

### 8.2.4 Reduced vector set

One way to reduce test cost is to improve the test efficiency of a test suite. If a defect is excited by many vectors, test metrics like NCR and INDIT can screen defective chips even with a reduced vector set. This fact can be exploited for test cost reduction using a reduced vector set by reducing the total test time.

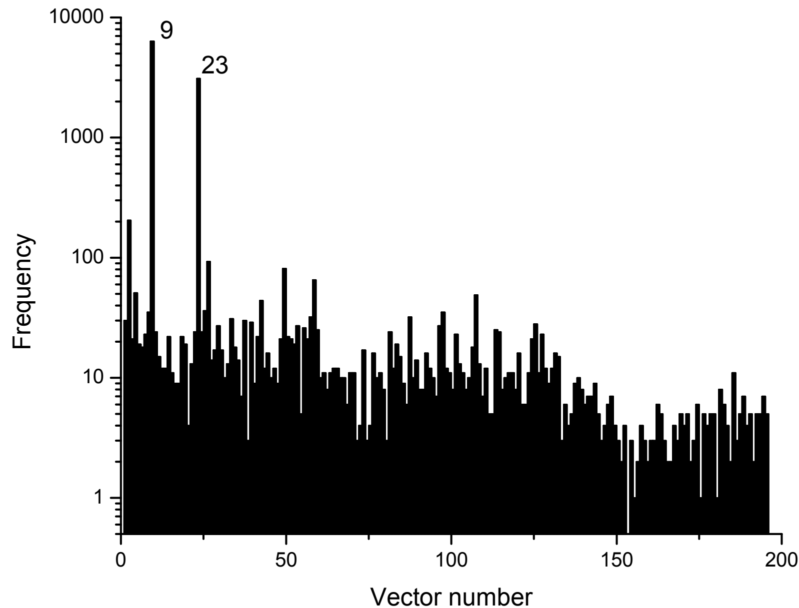


**Fig. 109. Histograms of vectors resulting in minimum $I_{DDQ}$.**

To evaluate this possibility we investigated the correlation between the vector pair that yield CR (one vector resulting in maximum $I_{DDQ}$ and the other resulting in minimum $I_{DDQ}$) and the vector pair that yields NCR. Minimum $I_{DDQ}$ is mostly due to intrinsic leakage. Which vector causes minimum $I_{DDQ}$ depends on which and how many paths are turned ON/OFF. Fig. 109 shows the distribution of vectors that resulted in minimum $I_{DDQ}$ for SEMATECH data. It shows that vector #9 yields the minimum $I_{DDQ}$ most often. Collectively, vectors #9 and #23 keep the device in the minimum $I_{DDQ}$ state 75% of the time (9423 of 12128 dice). A large component of the maximum $I_{DDQ}$ is likely to be due to defect current. Due to the random nature of defects, it is unlikely to find a single vector always having the maximum $I_{DDQ}$. In fact as Fig. 110 shows vectors #129 and #147 most frequently result in maximum $I_{DDQ}$, yet collectively account for less than 10% of the chips (1056 of 12128 dice). Since a large portion of maximum $I_{DDQ}$ stems from defect current and NCR is a *defect-oriented metric*, one would expect the vector causing maximum $I_{DDQ}$ also to result in maximum NCR. However, this need not always be true. NCR depends on which vector excites the defect for the center die and the relative magnitude of defect current compared to neighboring dice for that vector (assuming at least one neighboring die is fault-free and/or this vector does not excite a defect in a neighboring die). In general, which of

the high $I_{DDQ}$ vectors results in the highest NCR depends on the defect-free behavior of the neighboring dice. Fig. 111 shows the histogram of vectors producing the maximum NCR.

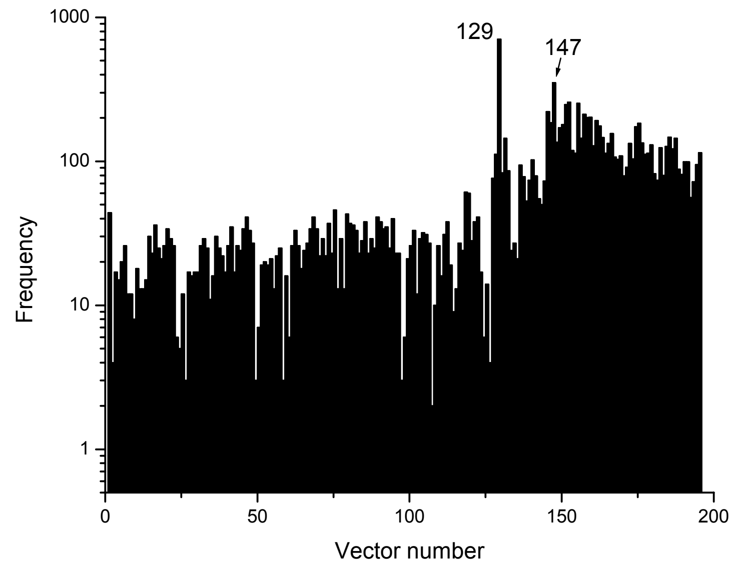


**Fig. 110. Histogram of vectors resulting in maximum $I_{DDQ}$.**

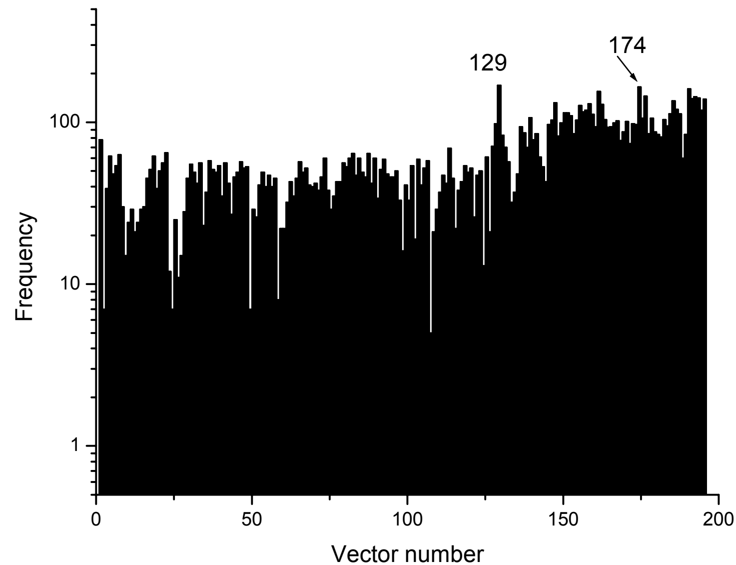


**Fig. 111. Histogram of vectors resulting in max NCR.**

As Fig. 110 and Fig. 111 show, the maximum CR vector and NCR vector distributions are not always in agreement. For example, although the vector (#129) causing maximum $I_{DDQ}$ most of the time is also the vector that results in the highest NCR most of the time, this is not true for other vectors. Thus, the second most common maximum $I_{DDQ}$ vector (#147) is not the second most common maximum NCR vector (#174). Due to the random nature of defects, each vector has a potential to excite a defect, although a vector that keeps the device already in a high $I_{DDQ}$ state is more likely to result in high NCR. This illustrates that vector reduction while employing CR alone can result in yield loss. It should be possible to reduce the vector set while achieving the same quality level (DPM) when NCR is used. Whether this is generally true or not and how our assumptions relate to actual fault coverage is left to future work.

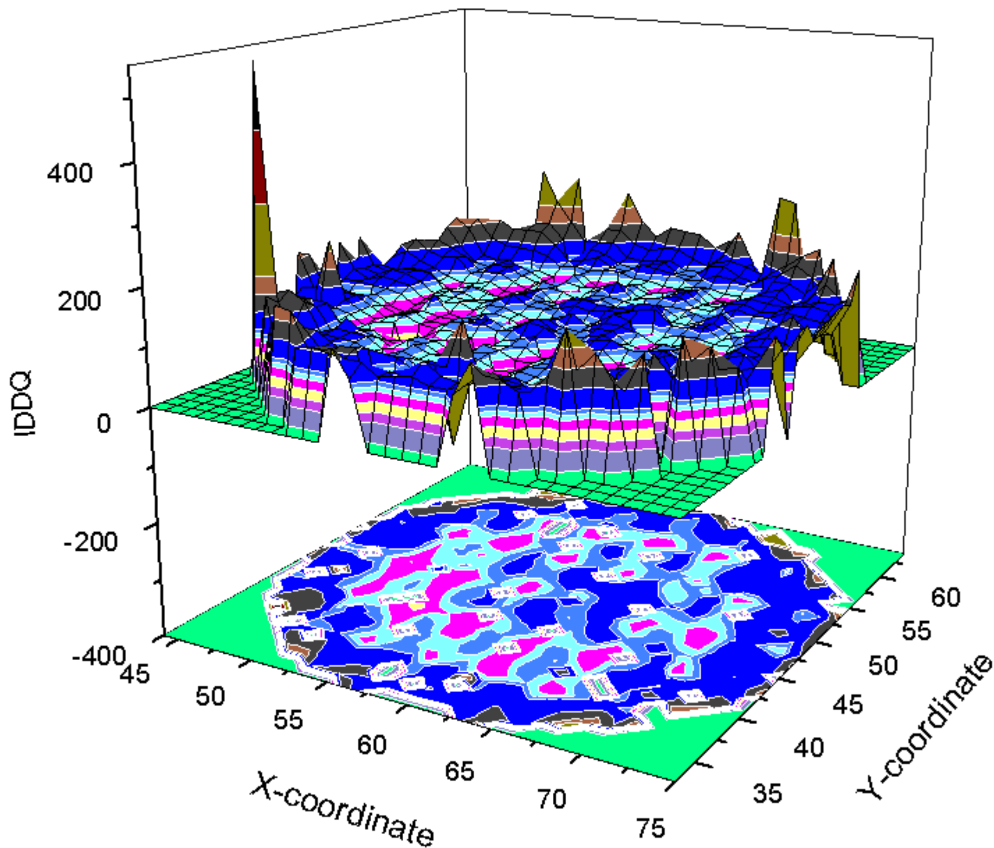


**Fig. 112. $I_{DDQ}$ surface plot for a wafer showing outlier behavior of wafer edge dice.**

### 8.2.5 Neighbor selection criteria

By analyzing wafer-to-wafer spatial patterns across a lot, it is possible to identify die positions that exhibit high correlation. This information about deterministic variations can then be used for computing NCR or INDIT. This will presumably lower yield loss and test escapes. The wafer edge dice often show different behavior than the rest of the dice on a wafer. Consider the surface plot and contour projections shown in Fig. 112. Many wafer edge dice *appear* to be outliers. However, there seems to be *deterministic* pattern to this behavior. By identifying die positions that correlate well with these chips, it is possible to reduce yield loss.

### 8.2.6 Multivariate outlier rejection

We showed that combining multiple test metrics [198] and/or parameters [200] is helpful for identifying outliers that are not detected by a single metric. Thus, a parameter (e.g. $I_{DDQ}$) can be expressed as a function of die position (or radial distance), flush delay, lot number, etc.

$$I_{DDQ} = f(\text{die XY, lot, delay, radial distance, process parameter}) \qquad (7.1)$$

Multivariate statistical outlier rejection methods can be used for outlier screening. Conceptually multivariate rejection using parametric data is logical extension of use of multiple test parameters like CR, NCR or any other combination that gives information not obtained by a single metric. Multivariate outlier rejection methods will hopefully provide better outlier screening capability or improve confidence in outlier rejection when two test metrics are not in agreement.

### 8.2.7 Process diagnosis

Statistical analysis methods have been used for process diagnosis [211]. Outlier identification can be used to provide feedback to process engineers to diagnose process drifts and glitches or to select chips for failure analysis.

### 8.2.8 Use of transient parameter data

In Section 3, we reviewed different transient current test methods. However, this research used only static parameter ($I_{DDQ}$). It is possible to use some of the techniques discussed here with transient parameter data.

### 8.2.9 Change in environmental condition

Valuable information about the circuit under test can be obtained by changing the environmental conditions in which a chip is tested. We discussed change of temperature for outlier detection. Similarly, it is possible to perform tests at very low voltage (VLV) [212], or temperature [213], or by exerting voltage stress [214] and combine that data for outlier screening. Such data should be used in conjunction with other test parameter data. Note that methods like NCR are applicable to any test parameter, not only $I_{DDQ}$. An example of altering test conditions was illustrated using post-stress data from TI. Similar analysis is possible by changing temperature, speed or other test parameters. Recent studies indicate that the hysteresis phenomenon in $I_{DDQ}$ can lead to outlier identification [215]. This is similar to changing test conditions.

## 8.3 Conclusion

This dissertation has evolved around the central theme of "if you cannot explain the variation, it must be due to a defect" and "the more, the better" (in a sense, more data catches more defects). As semiconductor technology advances, test becomes more challenging. These challenges will continue to be two-fold. The first challenge is related to the *magnitude of the problem* itself. As chips become more densely packed and have more transistors, wires and switching nodes, verifying their integrity in tandem becomes difficult. Several issues like crosstalk due to capacitive and inductive couplings need to be dealt with both at design and test levels. The second challenge comes from the *accuracy of testing*. The introduction of new materials, changes in process flows and shrinkage of transistor geometries result in inaccurate fault models. Thus, what was a subtle defect in earlier technology can result in a gross outlier in new technology (refer to Fig. 11). This means that as technology evolves, test engineers need to chase a moving target. This will continue to be true in the foreseeable future for *all* aspects of testing. Outlier identification in particular will continue to be more challenging and to be successful requires knowledge of underlying physics, semiconductor processing, the way a test is generated and a combination of different test methods.

## REFERENCES


[1] G. Moore, “Cramming More Components onto Integrated Circuits,” *Electronics*, vol. 38, no. 8, April 19, 1965.

[2] International Technology Roadmap for Semiconductors (ITRS), Semiconductor Industry Association, 2001 edition, [online] http://public.itrs.net.

[3] P. Burlison, “The LSI Manufacturers Testing,” in *Tutorial: LSI Testing*, 2nd Ed., Warren Fee, Ed., Piscataway, NJ: IEEE Press, 1978.

[4] A. Brown, *VLSI Circuit and Systems in Silicon*, London: McGraw Hill, 1991.

[5] R. J. Harvey, A. H. Bratt, and A. P. Dorey, “A Review of Testing Methods for Mixed Signal ICs,” *Microelectronics Journal*, vol. 24, no. 6, pp. 663-674, Oct. 1993.

[6] R. Kittler, “Challenges for Use of Statistical Software Tools in the Semiconductor Industry,” in *Proc. Joint Statistical Meetings*, Indianapolis, IN, Aug. 2000.

[7] J. S. Oakland and R. F. Followell, *Statistical Process Control: A Practical Guide*, London: Butterworth-Heinemann, 1986.

[8] T. Williams and N. Brown, “Defect Level as a Function of Fault Coverage,” *IEEE Trans. Computers*, vol. 30, no. 12, pp. 987-988, Dec. 1981.

[9] T. Foote et al., “Testing the 500-MHz IBM S/390 Microprocessor,” *IEEE Design and Test of Computers*, pp. 83-89, July-Sep. 1998.

[10] B. Davis, *The Economics of Automatic Testing*, London: McGraw-Hill, 1982.

[11] W. Needham, Personal Communication (former chief test engineer at Intel, CA), May 2003.

[12] W. Wolf, *Modern VLSI Design: System-on-Chip Design*, 3rd Ed., Upper Saddle River, NJ: Prentice Hall, 2002.

[13] Y. Cao, P. Gupta, A. Kahng, D. Sylvester, and J. Yang, “Design Sensitivities to Variability: Extrapolations and Assessments in Nanometer VLSI,” in *Proc. ASIC/SOC Conf.*, Rochester, NY, Sep. 2002, pp. 411-415.

[14] J. Rabaey, A. Chandrakasan, and B. Nokolic, *Digital Integrated Circuits: A Design Perspective*, 2nd Ed., Upper Saddle River, NJ: Prentice Hall, 2003.

[15] R. Madge, M. Rehani, K. Cota, and W. R. Daasch “Statistical Post-Processing at Wafersort – An Alternative to Burn-in and a Manufacturable Solution to Test Limit Setting for Sub-micron Technologies,” in *Proc. IEEE VLSI Test Symp.*, Monterey, CA, Apr. 2002, pp. 69-74.

[16] J. A. Loranger Jr., "The Case for Component Burn-in: The Gain is Well Worth the Price," *Electronics*, Jan. 23, 1975.

[17] A. M. Richardson, "An Assessment of Supply Current Measurement ($I_{DDQ}$) as a Reliability Indicator for CMOS Integrated Circuits," Ph.D. Dissertation, Lancaster University, UK, Oct. 1991.

[18] J. Segura, A. Keshavarzi, J. Soden, and C. Hawkins, "Parametric Failures in CMOS ICs – A Defect-based Analysis," in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Oct. 2002, pp. 90-99.

[19] B. Iglewicz and D. Hoaglin, *The ASQC Basic References in Quality Control: Statistical Techniques* in "How to Detect and Handle Outliers", vol. 16, E. F. Mykytka (Ed.), Milwaukee, WI: ASQC Quality Press, 1993.

[20] M. Burns and G. Roberts, *An Introduction to Mixed-Signal IC Test and Measurement*, New York: Oxford University Press, 2001.

[21] J. Pineda de Gyvez and D. Pradhan, *Integrated Circuit Manufacturability*, Piscataway, NJ: IEEE Press, 1999.

[22] C. F. Hawkins, H. T. Nagle, R. Fritzemeier and J. R. Guth, "The VLSI Circuit Test Problem – A Tutorial," *IEEE Trans. on Industrial Electronics*, vol. 36, no. 2, pp. 111-116, May 1989.

[23] J. Ferguson and J. Shen, "Extraction and Simulation of Realistic CMOS Faults Using Inductive Fault Analysis," in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Sep. 1988, pp. 475-484.

[24] H. Walker and S. Director, "VLASIC: A Catastrophic Fault Yield Simulator for Integrated Circuits," *IEEE Trans. on Computer Aided Design*, vol. 5, no. 4, pp. 541-556, Oct. 1986.

[25] A. Steininger, "Testing and Built-in Self-test – A Survey," *J. System Arch.*, vol. 46, no. 9, pp. 721-747, July 2000.

[26] W. Maly, A. Strojwas, and S. Director, "VLSI Yield Prediction and Estimation: A Unified Framework," *IEEE Trans. on Computer Aided Design*, vol. CAD-5, no. 1, pp. 114-130, Jan. 1986.

[27] R. Madge, "Aspects of Yield, Quality and Reliability," Invited talk, *Industrial Test Challenges Meeting*, Santa Clara, CA, May 2, 2003.

[28] F. J. Ferguson, M. Taylor, and T. Larrabee, "Testing for Parametric Faults in Static CMOS Circuits," in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Sep. 1990, pp. 436-443.

[29] E. McCluskey and C. Tseng, "Stuck-Fault Tests vs. Actual Defects," in *Proc. Intl. Test Conf.*, Atlantic City, NJ, Sep. 2000, pp. 336-343.

[30] R. Aitken, "Nanometer Technology Effects on Fault Models for IC Testing," *IEEE Computer*, vol. 32, no. 11, pp. 46-51, Nov. 1999.

[31] A. Majhi and V. Agrawal, "Delay Fault Models and Coverage," in *Proc. IEEE Intl. Conf. on VLSI Design*, Chennai, India, Jan. 1998, pp. 364-369.

[32] M. Bushnell and V. Agrawal, *Essentials of Electronic Testing for Digital, Memory and Mixed-Signal VLSI*, Norwell, MA: Kluwer Academic Publishers, USA, 2000.

[33] E. J. McCluskey, "Quality and Single-Stuck Faults," in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Oct. 1993, p. 597.

[34] J. P. Roth, "Diagnosis of Automata Failures: A Calculus and a Method," *IBM J. Research and Development*, pp. 278-291, July 1966.

[35] P. Goel and B. Rosales, "PODEM-X: An Automatic Test Generation System for VLSI Structures," in *Proc. IEEE Design Automation Conf.*, Nashville, TN, June 1981, pp. 260-268.

[36] S. Funatsu and M. Kawai, "An Automatic Test-Generation System for Large Digital Circuits," *IEEE Design and Test of Computers*, pp. 54-60, Oct. 1985.

[37] T. Kirkland and M. R. Mercer, "Algorithms for Automatic Test Pattern Generation," *IEEE Design and Test of Computers*, vol. 5, no. 3, pp. 43-54, June 1988.

[38] K. S. Kim, S. Mitra, and P. G. Ryan, "Delay Defect Characteristics and Testing Strategies," *IEEE Design and Test of Computers*, vol. 20, no. 5, pp. 8-16, Sept.-Oct. 2003.

[39] L. H. Goldstein, "Controllability/Observability Analysis of Digital Circuits," *IEEE Trans. on Circuits and Systems*, vol. CAS-26, no. 9, pp. 685-693, Sept. 1979.

[40] A. L. Crouch, *Design for Test for Digital IC's and Embedded Core Systems*, Upper Saddle River, NJ: Prentice Hall, 1999.

[41] T. Williams and K. Parker, "Design for Testability – A Survey," *IEEE Trans. on Computers*, vol. 71, no. 1, pp. 2-15, Jan. 1983.

[42] R. R. Fritzemeier, H. T. Nagle and C. F. Hawkins, “Fundamentals of Testability,” *IEEE Trans. on Industrial Electronics*, vol. 36, no. 2, pp. 117-128, May 1989.

[43] K. A. Bowman, S. G. Duvall, and J. D. Meindl, “Impact of Die-to-Die and Within-Die Parameter Fluctuations on the Maximum Clock Frequency Distribution for Gigascale Integration,” *IEEE J. Solid-State Circuits*, vol. 37, no. 2, pp. 183-190, Feb. 2002.

[44] K. Baker and J. Van Beers, “Shmoo Plotting: the Black Art of IC Testing,” *IEEE Design and Test of Computers*, vol. 14, no. 3, pp. 90-97, Jul-Sept. 1997.

[45] S. Sengupta, S. Kundu, S. Chakravarty, P. Parvathala, R. Galivanche et al., “Defect-Based Test: A Key Enabler for Successful Migration to Structural Test,” *Intel Technology Journal*, First Quarter, pp. 1-14, 1999.

[46] G. Nelson and W. Boggs, “Parametric Tests Meet the Challenge of High-density ICs,” *Electronics*, pp. 108-111, Dec. 11, 1975.

[47] M. Levi, “CMOS Is Most Testable,” in *Proc. IEEE Intl. Test Conf.*, Philadelphia, PA, Oct. 1981, pp. 217-220.

[48] Y. Malaiya and S. Su, “A New Fault Model and Testing Technique for CMOS Devices,” in *Proc. IEEE Intl. Test Conf.*, Philadelphia, PA, Sep. 1982. pp. 25-34.

[49] R. Rajsuman, “$I_{DDQ}$ Testing for CMOS VLSI,” in *Proc. of the IEEE*, vol. 88, no. 4, pp. 544-566, Apr. 2000.

[50] S. D. McEuen, “$I_{DDQ}$ Benefits,” in *Proc. IEEE VLSI Test Symp.,* Atlantic City, NJ, Oct. 1991, pp. 285-290.

[51] P. Nigh and A. Gattiker, “Test Method Evaluation Experiments and Data,” in *Proc. IEEE Intl. Test Conf.*, Atlantic City, NJ, Oct. 2000, pp. 454-463.

[52] P. Nigh, W. Needham, K. Butler, P. Maxwell, R. Aitken et al., “So What Is an Optimal Test Mix? A Discussion of the SEMATECH Methods Experiment,” in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Oct. 1997, pp. 1037-1038.

[53] P. C. Maxwell, R. C. Aitken, V. Johansen, and I. Chiang, “The Effectiveness of $I_{DDQ}$, Functional and Scan Tests: How Many Fault Coverages do We Need?,” in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Sep. 1992, pp. 168-177.

[54] R. Perry, “$I_{DDQ}$ Testing in CMOS Digital ASICs – Putting It All Together,” in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Sep. 1992, pp. 151-157.

[55] C. F. Hawkins, A. Keshavarzi, and J. M. Soden, "Reliability, Test and $I_{DDQ}$ Measurements," in *Proc. IEEE Intl. Workshop on $I_{DDQ}$ Testing,* Washington, DC, Oct. 1997, pp. 96-102.

[56] P. Wiscombe, "A Comparison of Stuck-at Fault Coverage and $I_{DDQ}$ Testing on Defect Levels," in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Oct. 1993, pp. 293-299.

[57] W. Kuo, W. K. Chien, and T. Kim, *Reliability, Yield, and Stress Burn-in: A Unified Approach for Microelectronics Systems Manufacturing & Software Development*, Boston: Kluwer Academic Publishers, 1998.

[58] J. van Horn, "Towards Achieving Reliability Expectations in a Future of GHz, Nanometers, Amps, Watts and Miles of Copper," Invited talk, *Industrial Test Challenges Meeting*, Santa Clara, CA, May 2002.

[59] R. Rajsuman, *Digital Hardware Testing*. Norwood, MA: Artech House, 1992.

[60] S. McEuen, "Reliability Benefits of $I_{DDQ}$," *J. Electronics Testing: Theory and Applications*, vol. 3, no. 4, pp. 41-49, 1992.

[61] T. Barrette, V. Bhide, K. De, M. Stover, and E. Sugaswara, "Evaluation of Early Failure Screening Methods," in *Proc. IEEE Intl. Workshop on $I_{DDQ}$ Testing*, Washington, DC, Oct. 1996, pp. 14-17.

[62] T. Henry and T. Soo, "Burn-in Elimination of a High Volume Microprocessor Using $I_{DDQ}$," in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Oct. 1996, pp. 242-249.

[63] S. Mallarapu and A. Hoffmann, "$I_{DDQ}$ Testing On a Custom Automotive IC," *IEEE J. Solid State Circuits*, vol. 30, no. 3, pp. 295-299, Mar. 1995.

[64] K. Wallquist, "On the Effect of $I_{SSQ}$ Testing in Reducing Early Failure Rate," in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Oct. 1995, pp. 910-915.

[65] J. Soden and C. Hawkins, "$I_{DDQ}$ Testing and Defect Classes – A Tutorial," in *Proc. Custom Integrated Circuits Conf.*, Santa Clara, CA, May 1995, pp. 633-642.

[66] SEMATECH Report 97023245A-TR, "Direct Drain Quiescent Current ($I_{DDQ}$) Burn-in Effectiveness Project (S88B) Report," 1997.

[67] H. Su, F. Liu, A. Devgan, E. Acar and S. Nassif, "Full Chip Leakage Estimation Considering Power Supply and Temperature Variations," in *Proc. IEEE Intl. Symp. on Low Power Electronics and Design*, Seoul, Korea, Aug. 2003, pp. 78-83.

[68] C. Thibeault and L. Boisvert, "On the Current Behavior of Faulty and Fault-free ICs and the Impact on Diagnosis," in *Proc. Intl. Symp. on Defect and Fault Tolerance in VLSI Systems*, Austin, TX, Nov. 1998, pp. 202-210.

[69] R. Rao, A. Srivastava, D. Blaauw and D. Sylvester, "Statistical Estimation of Leakage Current Considering Inter- and Intra-Die Process Variation," in *Proc. IEEE Intl. Symp. on Low Power Electronics and Design*, Seoul, Korea, Aug. 2003, pp. 84-89.

[70] W. Mao, R. Gulati, D. Goel, an M. Ciletti, "QUIETEST: A Quiescent Current Testing Methodology for Detecting Leakage Faults," in *Proc. IEEE Intl. Conf. on Computer-Aided Design*, Santa Clara, CA, Nov. 1990, pp. 280-283.

[71] S. Mukhopadhyay and K. Roy, "Modeling and Estimation of Total Leakage Current in Nano-scaled CMOS Devices Considering the Effect of Parameter Variation," in *Proc. IEEE Intl. Symp. on Low Power Electronics and Design*, Seoul, Korea, Aug. 2003, pp. 172-175.

[72] P. Maxwell and J. Rearick, "A Simulation-based Method for Estimating Defect-free $I_{DDQ}$," in *Proc. IEEE Intl. Workshop on $I_{DDQ}$ Testing*, Washington, DC, Nov. 1997, pp. 80-84.

[73] P. Maxwell and J. Rearick, "Estimation of Defect-Free $I_{DDQ}$ in Sub-micron Circuits Using Switch Level Simulation," in *Proc. IEEE Intl. Test Conf.,* Washington, DC, Oct. 1998, pp. 882-889.

[74] T. A. Unni and D. M. H. Walker, "Model-based $I_{DDQ}$ Pass/Fail Limit Setting," in *Proc. IEEE Intl. Workshop on $I_{DDQ}$ Testing,* San Jose, CA, 1998, pp. 43-47.

[75] A. Miller, "$I_{DDQ}$ Testing in Deep Sub-micron Integrated Circuits," in *Proc. IEEE Intl. Test Conf.,* Atlantic City, NJ, Sep. 1999, pp. 724-729.

[76] T. Miller, J. Soden, and C. Hawkins, "Diagnosis, Analysis and Comparison of 80386EX $I_{DDQ}$ and Functional Test Failures," in *Proc. IEEE Intl. Workshop on $I_{DDQ}$ Testing*, Washington, DC, Oct. 1995, pp. 66-68.

[77] B. Laquai, H. Richter, and H. Wekmann, "A Production-Oriented Measurement of Fast and Exhaustive $I_{DDQ}$ Tests," in *Proc. European Design and Test Conf.*, Paris, France, Mar. 1997, pp. 279-286.

[78] K. M. Wallquist, A. W. Righter, and C. F. Hawkins, "A General Purpose $I_{DDQ}$ Measurement Circuit," in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Oct. 1993, pp. 642-651.

[79] K. M. Wallquist, “Achieving $I_{DDQ}/I_{SSQ}$ Production Testing With QuiC-Mon,” *IEEE Design and Test of Computers*, vol. 12, no. 5, pp. 62-69, Sept. 1995.

[80] S. Athan, “Method and Apparatus for Use in $I_{DDQ}$ Integrated Circuit Testing,” US Patent 6,444,214, Nov. 2000.

[81] A. Keshavarzi, K. Roy, and C. Hawkins, “Intrinsic Leakage in Low Power Deep Submicron CMOS ICs,” in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Oct. 1997, pp. 146-155.

[82] J. Figueras and A. Ferre, “Possibilities and Limitations of $I_{DDQ}$ Testing in Sub-micron CMOS,” *IEEE Trans. on Components, Packaging, and Manufacturing Technology*, part B, vol. 21, no. 4, pp. 352-359, Nov. 1998.

[83] M. Sachdev, “Current-Based Testing for Deep-Sub-micron VLSIs,” *IEEE Design and Test of Computers*, vol. 18, no. 2, pp. 76-84, March-April 2001.

[84] H. Iwai, “CMOS Technology – Year 2010 and Beyond,” in *IEEE J. Solid-State Circuits*, vol. 34, no. 3, pp. 357-366, Mar. 1999.

[85] T. W. Williams, R. H. Dennard, R. Kapur, M. R. Mercer, and W. Maly, “$I_{DDQ}$ Test: Sensitivity Analysis of Scaling,” in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Oct. 1996, pp. 786-792.

[86] T. W. Williams, R. Kapur, M. R. Mercer, R. H. Dennard and W. Maly, “$I_{DDQ}$ Testing for High-Performance CMOS – The Next Ten Years,” in *Proc. IEEE European Design and Test Conf.*, Paris, France, March 1996, pp. 578-583.

[87] K. Baker, “SIA Roadmaps: Sunset Boulevard for $I_{DDQ}$,” in *Proc. IEEE Intl. Test Conf.*, Atlantic City, NJ, Oct. 1999, p. 1121.

[88] C. Hawkins and J. Soden, “Deep Submicron CMOS Current IC Testing: Is There a Future?,” *IEEE Design and Test of Computers*, vol. 16, no. 4, pp. 14-15, Oct.-Dec. 1999.

[89] P. Maxwell and R. Aitken, “$I_{DDQ}$ Testing as a Component of a Test Suite: The Need for Several Fault Coverage Metrics,” *J. Electronics Testing: Theory and Applications*, vol. 3, no. 4, pp. 19-30, 1992.

[90] A. Ferré, E. Isern, J. Rius, R. Rodrígues-Montañés, and J. Figueras, “$I_{DDQ}$ Testing: State of the Art and Future Trends,” *Integration: the VLSI Journal*, vol. 26, pp. 167-196, Dec. 1998.

[91] S. Davidson, “Is $I_{DDQ}$ Yield Loss Inevitable?,” in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Oct. 1994, pp. 572-579.

[92] J. Beasley, H. Ramamurthy, J. Ramfrez-Angulo, and M. DeYong, “$I_{DD}$ Pulse Response Testing: A Unified Approach to Testing Digital and Analogue ICs,” *IEEE Electronic Letters*, vol. 29, no. 24, pp. 2101-2103, Nov. 1993.

[93] P. Buxton and P. Tabor, “Outlier Detection for DPPM Reduction,” in *Proc. Intl. Test Conf.*, Charlotte, NC, Sep. 2003, pp. 818-827.

[94] J. Kao, S. Narendra, and A. Chandrakasan, “Sub-threshold Leakage Modeling and Reduction Techniques,” in *Proc. IEEE Intl. Conf. on Computer-Aided Design*, San Jose, CA, Nov. 2002, pp. 141-148.

[95] Y-F. Tsai, D. Duarte, N. Vijaykrishnan, and M. J. Irwin, “Implications of Technology Scaling on Leakage Reduction Techniques,” in *Proc. IEEE Design Automation Conf.*, Anaheim, CA, June 2003, pp. 187-190.

[96] S. Sze, *Physics of Semiconductor Devices*. New York: John Wiley, 1981.

[97] A. Keshavarzi, S. Ma, S. Narendra, B. Bloechel, K. Mistry et al., “Effectiveness of Reverse Body Bias for Leakage Control in Scaled Dual-$V_t$ CMOS ICs,” in *IEEE Intl. Symp. on Low Power Electronics and Design*, Huntington Beach, CA, Aug. 2001, pp. 207-212.

[98] A. Keshavarzi, K. Roy, M. Sachdev, C. Hawkins, K. Soumyanath et al., “Multiple-Parameter CMOS IC Testing with Increased Sensitivity for $I_{DDQ}$,” in *Proc. IEEE Intl. Test Conf.*, Atlantic City, NJ, Oct. 2000, pp. 1051-1059.

[99] M. Johnson, D. Somasekhar, L. Chiou, and K. Roy, “Leakage Control with Efficient Use of Transistor Stacks in Single Threshold CMOS,” *IEEE Trans. on VLSI Systems*, vol. 10, no. 1, pp. 1-5, Feb. 2002.

[100] S. Narendra, S. Borkar, V. De, D. Antoniadis, and A. Chandrakasan, “Scaling of Stack Effect and its Application for Leakage Reduction,” in *Proc. IEEE Intl. Symp. on Low Power Electronics and Design*, Huntington Beach, CA, Aug 2001, pp. 195-200.

[101] Y. Ye, S. Borkar, and V. De, “A New Technique for Standby Leakage Reduction in High-Performance Circuits,” in *Proc. IEEE Symp. on VLSI Circuits,* Honolulu, HI, June 1998, pp. 40-41.

[102] S. Mutoh, T. Douskei, Y. Matsuya, T. Aoki, S. Shigematsu et al., “1-V Power Supply High-speed Digital Circuit Technology with Multithreshold-voltage CMOS,” *IEEE J. Solid-State Circuits*, vol. 30, no. 8, pp. 847-853, Aug. 1995.

[103] T. Karnik, S. Borkar, and V. De, “Sub-90 nm Technologies – Challenges and Opportunities for CAD,” in *Proc. IEEE Intl. Conf. on Computer Aided Design*, San Jose, CA, Nov. 2002, pp. 203-206.

[104] G. Sery, S. Borkar, and V. De, “Life Is CMOS: Why Chase the Life After?,” in *Proc. IEEE Design Automation Conf.*, New Orleans, LA, June 2002, pp. 78-83.

[105] L. Wei, Z. Chen, K. Roy, M. Johnson, Y. Ye et al., “Design and Optimization of Dual Threshold Circuits for Low Voltage Low Power Applications,” *IEEE Trans. on VLSI Systems*, vol. 7, no. 1, pp. 16-24, Mar. 1999.

[106] B. Iniguez, J. Raskin, P. Simon, D. Flandre and J. Segura, “Analysis and Future Trends of $I_{DDQ}$ Testing for Silicon On Insulator CMOS ICs,” in *Proc. IEEE Intl. Workshop on Defect Based Testing*, Los Angeles, CA, Apr. 2001, pp. 40-44.

[107] J. Colinge, *Silicon-On-Insulator Technology: Materials to VLSI*, Boston: Kluwer Academic Publishers, 1997.

[108] W. Maly and P. Nigh, “Built-in Current Testing – Feasibility Study,” in *Proc IEEE Intl. Conf. Computer Aided Design*, Santa Clara, CA, Nov. 1988, pp.340-343.

[109] W. Maly and M. Patyra, “Built-in Current Testing,” *IEEE J. Solid State Circuits*, vol. 27, no. 3, pp. 425-428, Mar. 1992.

[110] D. M. H. Walker, “Requirements for Practical $I_{DDQ}$ Testing of Deep Sub-micron Circuits,” in *Proc. IEEE Intl. Workshop on Defect and Current Based Testing*, Montreal, Canada, April 2000, pp. 15-20.

[111] Y. Gao and J. Lo, “Challenges of Built-In Current Sensor Designs,” in *Proc. Intl. Symp. on Defect and Fault Tolerance in VLSI Systems*, Austin, TX, Nov. 1998, pp. 192-200.

[112] A. Gattiker and W. Maly, “Current Signatures: Application,” in *Proc. IEEE Intl. Test Conf.,* Washington, DC, Oct. 1997, pp. 156-165.

[113] A. Gattiker and W. Maly, “Current Signatures for $I_{DDQ}$ Testing,” US Patent 6,175,244, Jan. 2001.

[114] A. Gattiker and W. Maly, “Current Signatures for Production Testing,” in *Proc. IEEE Intl. Workshop on $I_{DDQ}$ Testing,* Washington, DC, Oct. 1996, pp. 25-28.

[115] A. Gattiker, “Current Signatures for Test Strategy Advisor,” Ph.D. Dissertation, Carnegie Mellon University, Pittsburgh, PA, USA, 1998.

[116] A. Gattiker, P. Nigh, and W. Maly, "Current-Signature-Based Analysis of Complex Test Fails," in *Proc. IEEE Intl. Symp. on Test and Failure Analysis*, Santa Clara, CA, Nov. 1999, pp. 377-387.

[117] C. Thibeault, "Replacing $I_{DDQ}$ Testing: With Variance Reduction," in *J. Electronic Testing: Theory and Applications*, vol. 19, no. 3, pp. 325-340, June 2003.

[118] C. Thibeault, "A Novel Probabilistic Approach for IC Diagnosis Based on Differential Quiescent Current Signatures," in *Proc. IEEE VLSI Test Symp.*, Monterey CA, 1997, pp. 80-85.

[119] C. Thibeault, "An Histogram Based Procedure for Current Testing of Active Defects," in *Proc. IEEE Intl. Test Conf.*, Atlantic City, NJ, Sep. 1999, pp. 714-723.

[120] C. Thibeault, "On the comparison of $I_{DDQ}$ and $\Delta I_{DDQ}$ Testing," in *Proc. IEEE VLSI Test Symp.,* San Diego, CA, Apr. 1999, pp. 143-150.

[121] T. Powell, J. Pair, M. St. John, and D. Counce, "Delta $I_{DDQ}$ for Testing Reliability," in *Proc. IEEE VLSI Test Symp.,* Montreal, Canada, Apr. 2000, pp. 439-443.

[122] O. Semenov, A. Vassighi, and M. Sachdev, "Leakage Current in Sub-Quarter Micron MOSFET: A Perspective on Stressed Delta $I_{DDQ}$ Testing," *J. Electronic Testing: Theory and Applications*, vol. 19, no. 3, pp. 341-352, June 2003.

[123] B. Kruseman, R. van Veen, and K. van Kaarm, "The Future of Delta-$I_{DDQ}$ Testing," in *Proc. IEEE Intl. Test Conf.,* Baltimore, MD, Oct. 2001, pp. 101-110.

[124] S. Jandhyala, H. Balachandran, and A. Jayasumana, "Clustering Based Techniques for $I_{DDQ}$ Testing," in *Proc. IEEE Intl. Test Conf.,* Atlantic City, NJ, Sep. 1999, pp. 730-737.

[125] S. Jandhyala, H. Balachandran, M. Sengupta, and A. Jayasumana, "Clustering Based Evaluation of $I_{DDQ}$ Measurements: Applications in Testing and Classification of ICs," in *Proc. IEEE VLSI Test Symp.*, Montreal, Canada, Apr. 2000, pp. 444-449.

[126] P. Maxwell, P. O'Neill, R. Aitken, R. Dudley, N. Jaarsma et al., "Current Ratios: A Self-scaling Technique for Production $I_{DDQ}$ Testing," in *Proc. IEEE Intl. Test Conf.,* Atlantic City, NJ, Sep. 1999, pp. 738-746.

[127] X. Sun, L. Kinney, and B. Vinnakota, "Performance Improvement of Current-Ratio based Test Using Correlation-Oriented Test Vector Generation," in *Proc. IEEE Intl. Workshop on Defect Based Testing*, Napa Valley, CA, Apr. 2003, pp. 23-28.

[128] X. Sun, L. Kinney, and B. Vinnakota, "Test Vector Generation Based on Correlation Model for Ratio-$I_{DDQ}$," in *Proc. IEEE Intl. Test Conf.*, Charlotte, NC, Sep. 2003, pp. 545-554.

[129] Y. Okuda, "DECOUPLE: Defect Current Detection in Deep Submicron $I_{DDQ}$," in *Proc. IEEE Intl. Test Conf.*, Atlantic City, NJ, Oct. 2000, pp. 199-206.

[130] Y. Okuda, "Eigen-Signatures for Regularity-based $I_{DDQ}$ Testing," in *Proc. IEEE VLSI Test Symp.*, Monterey, CA, Apr. 2002, pp. 289-294.

[131] J. Kalb, Jr., "Method for Testing a Semiconductor Device by Measuring Quiescent Currents ($I_{DDQ}$) at Two Different Temperatures," US Patent 6,242,934, June 2001.

[132] S. Kundu, "$I_{DDQ}$ Defect Detection in Deep Submicron CMOS ICs," in *Proc. IEEE Asian Test Symp.*, Singapore, Dec. 1998, pp. 150-152.

[133] V. Szekely, M. Rencz, S. Torok, and B. Courtois, "Cooling as a Possible Way to Extend the Usability of $I_{DDQ}$ Testing," *IEEE Electronics Letters*, vol. 33, no. 6, pp. 2117-2118, Dec. 1997.

[134] A. Keshavarzi et al., "Feasibility of Current Measurements in Sub 0.25 micron VLSIs," in *Proc. IEEE Intl. Workshop on Defect Based Testing*, Montreal, Canada, Apr. 2000, pp. 3-8.

[135] C. Yu, T. Maung, C. Spanos, D. Boning, J. Chung et al., "Use of Short-Loop Electrical Measurements for Yield Improvement," *IEEE Trans. on Semiconductor Manufacturing*, vol. 8, no. 2, pp. 150-159, May 1995.

[136] S. Sabade and D. M. H. Walker, "Wafer-level Spatial and Flush Delay Correlation Analysis for $I_{DDQ}$ Estimation," in *Proc. IEEE Intl. Workshop on Defect Based Testing*, Monterey, CA, Apr. 2002, pp. 47-52.

[137] T. J. Vogels, "Effectiveness of I-V Testing in Comparison to $I_{DDQ}$ Tests," in *Proc. IEEE VLSI Test Symp.*, Napa Valley, CA, Apr. 2003, pp. 47-52.

[138] D. Bergman and H. Engler, "Improved $I_{DDQ}$ Testing with Empirical Linear Prediction," in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Oct. 2002, pp. 954-963.

[139] S. K. Kachigan, *Multivariate Statistical Analysis*: *A Conceptual Introduction*, New York: Radius Press, 1982.

[140] J. Beasley, H. Ramamurthy, J. Ramirez-Angulo, and M. Deyong, "$I_{DD}$ Pulse Response Testing of Analog and Digital CMOS Circuits," in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Oct. 1993, pp. 626-634.

[141] J. Beasley, A. Righter, C. Apodaca, S. Pour-Mozafari, and D. Huggett, "$I_{DD}$ Pulse Response Testing Applied to Complex CMOS ICs," in *Proc. IEEE Intl. Test Conf.*, Nov. 1997, Washington, DC, pp. 32-39.

[142] J. F. Frenzel and P. N. Marinos, "Power Supply Current Signature (PSCS) Analysis: A New Approach to System Testing," in *Proc. IEEE Intl. Test Conf.*, Nov. 1987, pp. 125-135.

[143] S. Su, R. Makki, and T. Nagle, "Transient Power Supply Current Monitoring – A New Test Method for CMOS VLSI Circuits," *J. Electronic Testing: Theory and Applications*, vol. 6, no. 1, pp. 23-43, Feb. 1995.

[144] J. Segura, M. Roca, D. Mateo, and A. Rubio, "An Approach to Dynamic Power Consumption Testing of CMOS ICs," in *Proc. IEEE VLSI Test Symp.*, Princeton, NJ, Apr. 1995, pp. 95-100.

[145] C. Thibeault, "Detection and Location of Faults and Defects Using Digital Signal Processing," in *Proc. IEEE VLSI Test Symp.*, Princeton, NJ, May 1995, pp. 262-267.

[146] C. Thibeault and A. Payeur, "Experimental Results from $I_{DDF}$ Testing," in *Proc. IEEE Intl. Symp. on Defect and Fault Tolerance in VLSI Systems*, Boston, MA, Nov. 1992, pp. 22-31.

[147] J. Plusquellic, D. Chiarulli, and S. Levitan, "Digital Integrated Circuit Testing Using Transient Signal Analysis," in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Oct. 1996, pp. 481-490.

[148] A. Walker and P. Lala, "An Approach for Detecting Bridging Fault-induced Delay Faults in Static CMOS Circuits Using Dynamic Power Supply Current Monitoring," in *Proc. IEEE Intl. Workshop on $I_{DDQ}$ Testing*, Apr. 1997, pp. 73-77.

[149] Y. Min, Z. Zhao, and Z. Li, "$I_{DDT}$ Testing," in *Proc. IEEE Asian Test Symp.*, Akita, Japan, Nov. 1997, pp. 378-382.

[150] B. Vinnakota, W. Jiang, and D. Sun, "Process-Tolerant Test with Energy Consumption Ratio," in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Oct. 1998, pp. 1027-1036.

[151] J. Rius and J. Figueras, "Exploring the Combination of $I_{DDQ}$ and $I_{DDT}$ Testing: Energy Testing," in *Proc. IEEE Design Automation and Test in Europe*, 1999, pp. 543-548.

[152] M. Ishida, D. Ha, T. Yamaguchi, Y. Hashimoto and T. Ohmi, "$I_{DDT}$ Testing: An Efficient Method for Detecting Delay Faults and Open Defects," in *Proc. IEEE Intl. Workshop on Defect Based Testing*, Marina del Rey, CA, 2001, pp. 23-28.

[153] A. Walker, A. Henry, and P. Lala, "An Approach for Detecting Bridging Faults in CMOS Domino Logic Circuits Using Dynamic Power Supply Monitoring," in *Proc. IEEE Intl. Workshop on Defect and Fault Tolerance in VLSI Systems*, Paris, France, Oct. 1997, pp. 272-280.

[154] A. Germida, Y. Zheng, J. Plusquellic, and F. Muradali, "Defect Detection Using Power Supply Transient Signal Analysis," in *Proc. IEEE Intl. Test Conf.*, Atlantic City, NJ, Sep. 1999, pp. 67-76.

[155] A. Singh, J. Plusquellic, and A. Gattiker, "Power Supply Transient Signal Analysis Under Real Process and Test Hardware Models," in *Proc. VLSI Test Symp.*, Monterey, CA, May 2002, pp. 357-362.

[156] A. Singh et al., "Detecting Delay Faults Using Power Supply Transient Signal Analysis," in *Proc. Intl. Test Conf.*, Baltimore, MD, Oct. 2001, pp. 395-404.

[157] J. Plusquellic, D. Chiarulli, and S. Levitan, "Identification of Defective CMOS Devices Using Correlation and Regression Analysis of Frequency Domain Transient Signal Data," in *Proc. Intl. Test Conf.*, Washington, DC, Nov. 1997, pp. 40-49.

[158] C. Thibeault, "Using Frequency Analyses to Enhance IC Testability," in *Proc. Defect and Fault Tolerance in VLSI Systems*, Montréal, Canada, Oct. 1994, pp. 280-288.

[159] M. Hashizume, K. Yamada, T. Tamesada, and M. Kawakami, "Fault Detection of Combinatorial Circuits Based on Supply Current," in *Proc. Intl. Test Conf.*, Washington, DC, Sep. 1988, pp. 374-380.

[160] J. Rius and J. Figueras, "$I_{DDQ}$ Fault Detection by On the Fly Depowering," in *Proc. IEEE Intl. Workshop on $I_{DDQ}$ Testing*, Washington, DC, Oct. 1995, pp. 40-44.

[161] M. Keating and D. Meyer, "A New Approach to Dynamic $I_{DD}$ Testing," in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Nov. 1987, pp. 316-319.

[162] B. Vinnakota and W. Jiang, "IC Test Using the Energy Consumption Ratio," in *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 19, no. 1, pp. 129-141, Jan. 2000.

[163] W. Jiang and B. Vinnakota, "Statistical Threshold Formulation for Dynamic $I_{DD}$ Test," in *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 21, no. 6, pp. 694-705, June 2002.

[164] S. Sabade and D. M. H. Walker, "Evaluation of Outlier Rejection Methods for $I_{DDQ}$ Limit Setting," in *Proc. IEEE VLSI Design/Asia South Pacific Design Automation Conf.*, Bangalore, India, Jan. 2002, pp. 755-760.

[165] V. Barnett and T. Lewis, *Outliers in Statistical Data*, 3rd Ed., New York: Wiley, 1994.

[166] C. Hawkins and J. Segura, "GHz Testing and Its Fuzzy Targets," in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Oct. 2002, p. 1228.

[167] H. Darrell, *How to Lie With Statistics*, New York: Norton, 1954.

[168] G. E. Box, W. G. Hunter, J. S. Hunter, and W. G. Hunter, *Statistics for Experimenters: An Introduction to Design, Data Analysis*, *and Model Building*, New York: Wiley, 1978.

[169] R. P. Runyon, *How Numbers Lie: A Consumer's Guide to the Fine Art of Numerical Deception*, Lexington, MA: Lewis Publishing Company, 1981.

[170] A. J. Jaffe and H. F. Spirer, *Misused Statistics: Straight Talk for Twisted Numbers*, New York: Marcel Dekker, 1987.

[171] F. L. Ramsey and D. W. Schafer, *The Statistical Sleuth – A Course in Methods of Data Analysis*, Belmont, CA: Duxbury Press, 1997.

[172] D. Turner, D. Abercrombie, J. McNames, R. Daasch, and R. Madge, "Isolating and Removing Sources of Variation in Test Data," in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Oct. 2002, pp. 464-471.

[173] W. Chauvenet, "Method of Least Squares," *Appendix to Manual of Spherical and Practical Astronomy*, vol. 2, Philadelphia: Lippincott, pp. 469-566.

[174] J. R. Taylor, *An Introduction to Error Analysis*, 2nd Ed., Herndon, MA: University Science Books, 1997.

[175] R. Richmond, "Successful Implementation of Structured Testing," in *Proc. IEEE Intl. Test Conf.,* Atlantic City, NJ, Oct. 2000, pp. 344-348.

[176] R. B. D'Agostino, and M. A. Stephens, *Goodness-of-Fit Techniques*, New York: Marcel Dekker, 1986.

[177] R. H. Williams and C. F. Hawkins, "Errors in Testing," in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Sep. 1990, pp. 1018-1027.

[178] R. H. Williams, R. G. Wagner, and C. F. Hawkins, "Testing Errors: Data and Calculations in an IC Manufacturing Process," in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Sep. 1992, pp. 352-361.

[179] B. Stine, D. Boning, and J. Chung, “Analysis and Decomposition of Spatial Variation in Integrated Circuit Processes and Devices,” *IEEE Trans. on Semi. Manufacturing*, vol. 10, no. 1, pp. 24-41, Feb. 1997.

[180] J. K. Kibarian and A. J. Strojwas, “Using Spatial Information to Analyze Correlations Between Test Structure Data,” *IEEE Trans. on Semi Manufacturing*, vol. 4, no. 3, pp. 219-225, Aug. 1991.

[181] J. Miller, *Statistics for Advanced Level,* London, UK: Cambridge University Press, 1994.

[182] S. Sabade and D. M. H. Walker, “Improved Wafer-level Spatial Analysis for $I_{DDQ}$ Limit Setting,” in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Oct. 2001, pp. 82-91.

[183] W. R. Daasch, J. McNames, D. Bockelman, and K. Cota, “Variance Reduction Using Wafer Patterns in $I_{DDQ}$ Data,” in *Proc. IEEE Intl. Test Conf.,* Atlantic City, NJ, Oct. 2000, pp. 189-198.

[184] K. Cota, “Quiescent Current Estimation in Deep Submicron CMOS Technologies Using Spatial Analysis of $I_{DDQ}$ Data Sets,” M.S. Thesis, Portland State University, 2001.

[185] W. R. Daasch, K. Cota, and J. McNames, “Neighbor Selection for Variance Reduction in $I_{DDQ}$ and Other Parametric Data,” in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Oct. 2001, pp. 92-100.

[186] R. B. Miller and W. C. Riordan, “Unit Level Predicted Yield: a Method of Identifying High Defect Density Die at Wafer Sort,” in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Oct. 2001, pp. 1118-1127.

[187] S. Sabade and D. M. H. Walker, “NCR: A Self-scaling, Self-calibrated Metric for Outlier Identification,” in *Proc. IEEE Midwest Symp. on Circuit and Systems*, Tulsa, OK, Aug. 2002, pp. 392-395.

[188] S. Sabade and D. M. H. Walker, “Neighbor Current Ratio (NCR): A New Metric for Outlier Identification,” in *Proc. IEEE Intl. Symp. on Defect and Fault Tolerance in VLSI Systems*, Vancouver, Canada, Nov. 2002, pp. 381-389.

[189] S. Sabade and D. M. H. Walker, “Immediate Neighbor Difference $I_{DDQ}$ Test (INDIT) for Outlier Identification,” in *Proc. IEEE Intl. Conf. on VLSI Design*, New Delhi, India, Jan. 2003, pp. 361-363.

[190] S. Sabade and D. M. H. Walker, “Wafer Signature Analysis of $I_{DDQ}$ Test Data,” in *Proc. IEEE Intl. Workshop on Defect Based Testing*, Napa Valley, CA, Apr. 2003, pp. 59-64.

[191] D. R. Lakin II and A. D. Singh, “Exploiting Defect Clustering to Screen Bare Die for Infant Mortality Failures: An Experimental Study,” in *Proc. IEEE Intl. Test Conf.*, Atlantic City, NJ, Sep. 1999, pp. 23-30.

[192] H. Vierhaus, W. Meyer, and U. Gläser, “CMOS Bridges and Resistive Transistor Faults: $I_{DDQ}$ Versus Delay Effects,” in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Oct. 1993, pp. 83-91.

[193] R. Rizzolo, G. Hinkel, S. Michnowski, T. McPherson, and A. Sutcliffe, “System Performance Management for the S/390 Parallel Enterprise Server G5,” in *IBM J. Research and Development*, vol. 43, no. 5/6, pp. 651-660, Sep./Nov. 1999.

[194] L. Huisman, “Correlations Between Path Delays and the Accuracy of Performance Prediction,” in *Proc. IEEE Intl. Test Conf.*, Washington, DC, Oct. 1998, pp. 801-808.

[195] J. Lee, D. M. H. Walker, L. Milor, Y. Peng, and G. Hill, “IC Performance Prediction for Test Cost Reduction,” in *IEEE Intl. Symp. on Semiconductor Manufacturing*, Santa Clara, CA, Nov. 1999, pp. 111-114.

[196] A. Keshavarzi, K. Roy, and C. Hawkins, “Intrinsic Leakage in Deep Submicron CMOS ICs – Measurement-Based Test Solutions,” *IEEE Trans. on VLSI Systems*, vol. 8, no. 6, pp. 717-723, Nov. 2000.

[197] M. Hargrove, A. Bryant, S. Ku, E. Nowak, and M. Shernoy, “A Methodology for Estimating Total Off-State Leakage Current ($I_{DDQ}$) in CMOS Technology,” *IBM MicroNews,* vol. 7, no. 1, First Quarter 2001, pp. 26-28.

[198] S. Sabade and D. M. H. Walker, “Use of Multiple $I_{DDQ}$ Test Metrics for Outlier Identification,” in *Proc. IEEE VLSI Test Symp.*, Napa Valley, CA, Apr. 2003, pp. 31-38.

[199] A. D. Singh, P. Nigh, and C. M. Krishna, “Screening for Known Good Die Based on Defect Clustering: An Experimental Study,” in *Proc. IEEE Intl. Test Conf.,* Washington, DC, 1997, pp. 362-369.

[200] S. Sabade and D. M. H. Walker, “CROWNE: Current Ratio Outliers With Neighbor Estimator,” in *Proc. IEEE Intl. Symp. on Defect and Fault Tolerance in VLSI Systems*, Boston, MA, Nov. 2003, pp. 132-139.

[201] R. Gayle, “The Cost of Quality: Reducing ASIC Defects With $I_{DDQ}$, At-speed Testing and Increased Fault Coverage,” in *Proc. IEEE Intl. Test Conf.*, Baltimore, MD, Oct. 1993, pp. 285-292.

[202] S. Nassif, "Modeling and Forecasting of Manufacturing Variations," in *Proc. IEEE Design Automation Conf.*, Las Vegas, NV, June 2001, pp. 145-149.

[203] S. Nassif, "Modeling and Analysis of Manufacturing Variations," in *Proc. IEEE Custom Integrated Circuits Conf.*, San Diego, CA, May 2001, pp. 223-228.

[204] R. Madge, B. Goh, V. Rajgopalan, C. Macchietto, R. Daasch et al., "Screening Min$V_{DD}$ Outliers Using Feed-Forward Voltage Testing," in *Proc. IEEE Intl. Test Conf.,* Baltimore, MD, Oct. 2002, pp. 673-682.

[205] A. D. Singh and C. M. Krishna, "On the Effect of Defect Clustering on Test Transparency and IC Test Optimization," *IEEE Trans. on Computers*, vol. 45. no. 6, pp. 753-757, June 1996.

[206] P. Nigh, W. Needham, K. Butler, P. Maxwell and R. Aitken, "An Experimental Study Comparing the Relative Effectiveness of Functional, Scan, $I_{DDQ}$ and Delay-Fault Testing," in *Proc. IEEE VLSI Test Symp.,* Monterey, CA, May 1997, pp. 459-464.

[207] P. Maxwell, "The Heisenberg Uncertainty of Test," Invited Address, in *Proc. IEEE Intl. Test Conf.,* Baltimore, MD, Oct. 2002, p. 13.

[208] A. Gattiker and W. Maly, "Toward Understanding $I_{DDQ}$-Only Fails," in *Proc. IEEE Intl. Test Conf.*, Washington DC, 1998, pp. 174-183.

[209] J. Van der Pol, F. Kuper, and E. Ooms, "Relation Between Yield and Reliability of Integrated Circuits and Application to Failure Rate Assessment and Reduction in the One-Digit FIT and PPM Reliability Era," *Microelectronics Reliability*, vol. 36, no, 11/12, pp.1603-1610, Nov.-Dec. 1996.

[210] P. C. Maxwell, "Wafer-package Test Mix for Optimal Defect Detection and Test Time Savings," *IEEE Design and Test of Computers*, vol. 20, no. 5, Sep.-Oct. 2003, pp. 84-89.

[211] J. K. Kibarian and A. J. Strojwas, "IC Manufacturing Diagnosis Based on Statistical Analysis Techniques," *IEEE Trans. on Components, Hybrids, and Manufacturing Technology*, vol. 15, no. 3, pp. 317-321, June 1992.

[212] J. Chang and E. McCluskey, "Quantitative Analysis of Very-Low-Voltage Testing," in *Proc. IEEE VLSI Test Symp.*, Princeton, NJ, Apr. 1996, pp. 332-337.

[213] C. Tseng, E. McCluskey, X. Shao, J. Wu, and D. Wu, "Cold Delay Defect Screening," in *Proc. IEEE VLSI Test Symp.*, Montreal, Canada, Apr. 2000, pp. 183-188.

[214] J. Chang and E. McCluskey, "SHOrt Voltage Elevation (SHOVE) Test for Weak CMOS ICs," in *Proc. IEEE VLSI Test Symp.*, Monterey, CA, Apr. 1997, pp. 446-451.

[215] Y. Okuda and N. Furukawa, "Hysteresis of Intrinsic $I_{DDQ}$ Currents," in *Proc. IEEE Intl. Test Conf.*, Charlotte, NC, Sep. 2003, pp. 555-564.

# APPENDIX A

## SEMATECH EXPERIMENT S-121

The general purpose of SEMATECH project S-121 ("Semiconductor Test Method Evaluation") was to evaluate various test methods [206]. The test vehicle used for this experiment was an IBM ASIC (120K gates) manufactured in 0.6 µm (0.45 µm $L_{eff}$) three metal layer CMOS technology. Some portions of the chip used 50 MHz and other portions used 40 MHz. The design used full-scan LSSD latches and was $I_{DDQ}$ testable. Four types of tests were conducted at the wafer level. A sample of parts was packaged and all four tests were conducted again. A sample of packaged parts was subjected to successive levels of BI and each BI cycle was followed by the same tests. These test methods and their brief description is tabulated in Table XXVI.

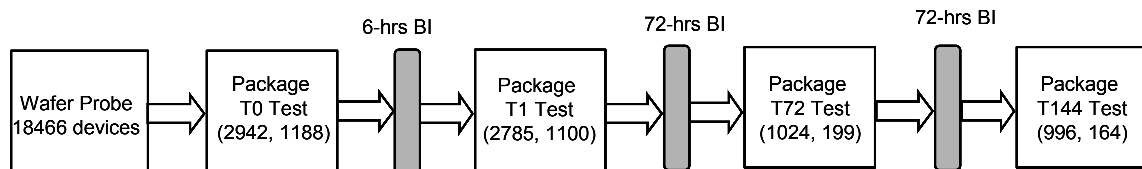


**Fig. 113. Test flow for SEMATECH experiment S-121.**

The wafer test was conducted at 50°C while other tests were conducted at room temperature (25°C). The BI temperature was 140°C and voltage was 1.5 times nominal $V_{DD}$. The chips that failed at least one but not all tests are referred to as *delta* chips. The chips that passed all tests were used as a reference and called *control* chips. The BI sample was biased towards delta chips. The test flow is illustrated in Fig. 113. The numbers in parenthesis are the number of delta chips, followed by the number of control devices. The distribution of chips according to their test results at each level is shown in Table XXVI.

**TABLE XXVI. DESCRIPTION OF SEMATECH TEST METHODS.**

| Test Method | Description |
| --- | --- |
| Functional test | Total 436,633 functional vectors<br>52% stuck-at fault coverage<br>Cycle time 35 ns (fast), 5 μs (slow)<br>Strobe time 25 ns (fast), 4 μs (slow) |
| Stuck-at fault test | 8023 vectors<br>99.6 % stuck-at fault coverage<br>(374339 out of 375485 total faults)<br>Cycle time 400 ns<br>Strobe time 390 ns |
| Scan-based delay test | 5232 test vectors<br>91% transitional fault coverage (estimated)<br>Untimed cycle time |
| $I_{DDQ}$ test | Total 195 vectors with 99.5% pseudo stuck-at coverage<br>Applied 125 vectors with 95.3% coverage<br>Additional 60 vectors same as stuck-at fault vectors<br>Additional 10 vectors from generic scan test<br>Pass/fail limit of 5 μA |
| Flush delay measurement | Average of propagation delays of transition down the scan chain from scan-in pin to scan-out pin for all scan chains for both transitions was measured |

**TABLE XXVII. DISTRIBUTION OF CHIPS AT DIFFERENT TEST LEVELS.**

| Test Result Code | Test level | | | | |
|---|---|---|---|---|---|
| | Wafer Test | Package Test | 6-hr BI test | 72-hr BI test | 144-hr BI test |
| $$ | 11263 | 1806 | 1864 | 168 | 187 |
| AF | 3103 | 275 | 102 | 33 | 86 |
| 1T | 18 | 9 | 9 | 1 | 1 |
| 1F | 89 | 7 | 6 | 4 | 5 |
| 1P | 339 | 16 | 22 | 17 | 16 |
| 1I | 1689 | 1469 | 1367 | 928 | 867 |
| 3T | 18 | 1 | 1 | 1 | 1 |
| 3F | 1234 | 253 | 254 | 8 | 6 |
| 3P | 161 | 90 | 85 | 2 | 1 |
| 3I | 90 | 36 | 36 | 0 | 0 |
| 2A | 51 | 0 | 0 | 0 | 0 |
| 2B | 54 | 53 | 49 | 6 | 5 |
| 2C | 48 | 31 | 27 | 4 | 4 |
| 2D | 73 | 7 | 5 | 5 | 5 |
| 2E | 78 | 46 | 43 | 0 | 0 |
| 2F | 85 | 0 | 0 | 0 | 0 |

$$: All pass; AF: All fail; 1T: Stuck-at-only fail; 1F: Functional-only fail; 1P: Delay-only fail; 1I: $I_{DDQ}$-only fail; 3T: Stuck-at pass other fail; 3F: Functional pass other fail; 3P: Delay pass other fail; 3I: $I_{DDQ}$ pass other fail; 2A: Functional and delay fail; 2B: Stuck-at and delay fail; 2C: Stuck-at and $I_{DDQ}$ fail; 2D: Functional and $I_{DDQ}$ fail; 2E: Functional and stuck-at fail; 2F: $I_{DDQ}$ and delay fail

The distribution of the 6-hr BI sample (2952 chips) based on test results is shown in Table XXVII. It also shows how test results of some chips change after BI. Chips that failed functional and/or stuck-at test during wafer probe are not shown, as in a typical production flow chips failing wafer test are rejected before packaging and never undergo BI. The chips that fail some tests earlier in the cycle but pass all tests after BI are called '*healer*' chips. These chips presumably contain some defect that healed during BI. For example, a metal sliver shorting two nodes may evaporate during BI. For all practical purposes, healers are unreliable chips. It can be observed that several chips fail only $I_{DDQ}$ test at all levels. Obviously not all of them are fatally flawed.

**TABLE XXVIII. DISTRIBUTION OF BI SAMPLE BASED ON TEST RESULTS.**

| Wafer Test Result | Package Test | 6-hrs BI Test | Number of chips |
|---|---|---|---|
| All Pass | All Pass | All Pass | 1051 |
| | | $I_{DDQ}$-only Fail | 23 |
| | | Delay Fail | - |
| | | Boolean Fail | 16 |
| | $I_{DDQ}$-only Fail | All Pass | - |
| | | $I_{DDQ}$-only Fail | 3 |
| | | Delay Fail | - |
| | | Boolean Fail | 1 |
| | Delay Fail | All Pass | 2 |
| | | $I_{DDQ}$-only Fail | - |
| | | Delay Fail | - |
| | | Boolean Fail | - |
| | Boolean Fail | All Pass | - |
| | | $I_{DDQ}$-only Fail | - |
| | | Delay Fail | - |
| | | Boolean Fail | 5 |
| $I_{DDQ}$-only Fail | All Pass | All Pass | 169 |
| | | $I_{DDQ}$-only fail | 36 |
| | | Delay Fail | - |
| | | Boolean Fail | 5 |
| | $I_{DDQ}$-only Fail | All Pass | 121 |
| | | $I_{DDQ}$-only fail | 1175 |
| | | Delay Fail | 2 |
| | | Boolean Fail | 33 |
| | Delay Fail | All Pass | - |
| | | $I_{DDQ}$-only Fail | 6 |
| | | Delay Fail | 1 |
| | | Boolean Fail | - |
| | Boolean Fail | All Pass | 2 |
| | | $I_{DDQ}$-only Fail | 2 |
| | | Delay Fail | - |
| | | Boolean Fail | 6 |
| Delay Fail | All Pass | All Pass | 255 |
| | | $I_{DDQ}$-only Fail | 1 |
| | | Delay Fail | - |
| | | Boolean Fail | 4 |
| | $I_{DDQ}$-only Fail | All Pass | - |
| | | $I_{DDQ}$-only Fail | 4 |
| | | Delay Fail | - |
| | | Boolean Fail | 1 |
| | Delay Fail | All Pass | 1 |
| | | $I_{DDQ}$-only Fail | - |
| | | Delay Fail | 8 |
| | | Boolean Fail | 1 |
| | Boolean Fail | All Pass | 3 |
| | | $I_{DDQ}$-only Fail | - |
| | | Delay Fail | - |
| | | Boolean Fail | 14 |

# APPENDIX B
## LSI DATA

LSI data was obtained in a production flow for LSI Logic Corporation's G12 (Twelfth Generation) ASIC Cell-based product. It uses 180 nm L-drawn (130 nm $L_{eff}$) CMOS process with 33M usable gates on a single chip. It uses a 6 metal layer 3.3V CMOS technology. The data comes from 79 lots and 1219 wafers for 949753 chips. For each chip, 20 $I_{DDQ}$ measurements are available. The information about the particular ASIC for which the data was collected cannot be disclosed due to proprietary reasons.

# APPENDIX C
# TI DATA

Due to proprietary nature of data, all details are not revealed in this dissertation. All figures that use TI data either do not show units or have been normalized. This data comes from an ASIC manufactured in 130-nm CMOS technology currently in production. Data was collected for 11348 chips from 7 lots and 30 wafers. For each chip, 12 $I_{DDQ}$ readings were obtained. Large on-chip memories used a different power supply. The chips were subjected to high-voltage stress and all 12 measurements were repeated. Each chip also underwent stuck-at scan test, delay test and functional test. The fault coverage of these tests is not known. For each chip bin data was available.

# APPENDIX D
## 3-D LINEAR REGRESSION

The least squares fit for two dimensions can be found in [181]. We use this method for three dimensions. The desired plane equation is of the form

$$z(x,y) = A + Bx + Cy \tag{C.1}$$

where $Z(x,y)$ is the parameter value for the $i^{th}$ vector for the die whose X and Y coordinates are $x$ and $y$. We wish to minimize the sum of the squares of errors, that is, perpendicular distances from each $Z(x_i,y_i)$ point to the plane. The error function to be minimized is given by:

$$E = \sum (z - A - Bx_i - Cy_i)^2 \tag{C.2}$$

Differentiating with respect to A, B and C and equating to zero we obtain:

$$\frac{\partial E}{\partial A} = \sum (z - A - Bx_i - Cy_i) = 0 \tag{C.3}$$

$$\frac{\partial E}{\partial B} = \sum (z - A - Bx_i - Cy_i)x_i = 0 \tag{C.4}$$

$$\frac{\partial E}{\partial C} = \sum (z - A - Bx_i - Cy_i)y_i = 0 \tag{C.5}$$

Rewriting,

$$\sum z_i = AN + B\sum x_i + C\sum y_i \tag{C.6}$$

$$\sum x_i z_i = A\sum x_i + B\sum x_i^2 + C\sum x_i y_i \tag{C.7}$$

$$\sum y_i z_i = A\sum y_i + B\sum x_i y_i + C\sum y_i^2 \tag{C.8}$$

where $N$ is the total number of points. If data is available for all the neighboring dice and all dice are tested fault-free, $N$ would be 8.

For simplicity, we use the following notation:

$$S_x = \sum x_i \text{ , } Sy = \sum y_i \text{ , } S_z = \sum z_i \quad \text{(C.9)}$$

$$S_{xx} = \sum x_i^2 \text{ , } S_{yy} = \sum y_i^2 \quad \text{(C.10)}$$

$$S_{xy} = \sum x_i y_i \text{ , } S_{xz} = \sum x_i z_i \text{ , } S_{yz} = \sum y_i z_i \quad \text{(C.11)}$$

Thus, the above equations can be written as:

$$S_z = AN + BS_x + CS_y \quad \text{(C.12)}$$

$$S_{xz} = AS_x + BS_{xx} + CS_{xy} \quad \text{(C.13)}$$

$$S_{yz} = AS_y + BS_{xy} + CS_{yy} \quad \text{(C.14)}$$

Solving simultaneously we obtain,

$$d = NS_{xx}S_{yy} + 2S_xS_yS_{xy} - S_{yy}S_x^2 - S_{xx}S_y^2 - NS_{xy}^2 \quad \text{(C.15)}$$

$$A = [S_{xz}(S_{xy}S_y - S_xS_{yy}) + S_{yz}(S_xS_{xy} - S_{xx}S_y) + S_z(S_{xx}S_{yy} - S_{xy}^2)]/d \quad \text{(C.16)}$$

$$B = [S_{xz}(NS_{yy} - S_y^2) + S_{yz}(S_xSy - NS_{xy}) + S_z(S_{xy}Sy - S_xS_{yy})]/d \quad \text{(C.17)}$$

$$C = [S_{xz}(SxS_y - NS_{xy}) + S_{yz}(NSxx - S_x^2) + S_z(S_xSxy - S_{xx}Sy)]/d \quad \text{(C.18)}$$

By substituting values of A, B and C the value of Z for a given (*x,y*) pair can be determined. This is the best fitting plane value.

# APPENDIX E

## CHAUVENET'S CRITERION

Chauvenet's criterion is a method to decide whether a data point is illegitimate. In this method, the accept/reject threshold is based on the probability of occurrence of such a data point in the given data set by calculating the standard deviation of the data set. Consider a data set containing $I_{DDQ}$ measurements as 3.8, 3.5, 3.9, 3.9, 5.8 and 3.4. The $I_{DDQ}$ measurement 5.8 seems illegitimate. To decide whether to accept or reject this chip, we compute the mean and standard deviation of this data. Here, the mean is 4.05 and standard deviation is 0.81. For each reading, we compute the standard scores $(x_i-\mu)/\sigma$. These are, -0.31, -0.68, -0.18, -0.18, 2.16 and -0.8. The probabilities of occurrence for these readings if data follows a Normal distribution are 0.75, 0.49, 0.85, 0.85, 0.03 and 0.42, respectively. If we set a minimum probability threshold of 0.4, reading 5.8 that has probability of occurrence of merely 0.03 gets rejected. After rejecting this reading, the new data set has mean of 3.7 and standard deviation of 0.21. Note that rejection of true outliers *may* not change the mean appreciably, but standard deviation does change.

This example is for the sake of illustration only. In practice, for proper use of Chauvenet's criterion, it is necessary to have a reasonable sample of data.

# APPENDIX F

# MAD SCORE COMPUTATION

This appendix provides an example of MAD score computation and outlier rejection. Consider the data shown in Table XXIX. Since there are a total of 10 values, we compute the median by averaging the 5th and 6th readings in the ordered data. Thus $\tilde{x}$ = (1+1.01)/2 = 1.005. The fourth column lists the ordered $|x_i - \tilde{x}|$. Thus MAD = (0.025+0.045)/2=0.035. The $M_i$ scores are computed as 0.6745($x_i$-1.005)/0.035.

**TABLE XXIX. MAD-BASED REJECTION EXAMPLE.**

| No. | Data ($x_i$) | Ordered Data ($x_i$) | Ordered $\|x_i - \tilde{x}\|$ | Mi |
|---|---|---|---|---|
| 1 | 1.03 | 0.76 | 0.005 | 0.48 |
| 2 | 0.96 | 0.89 | 0.005 | -0.87 |
| 3 | 1.11 | 0.96 | 0.015 | 2.02 |
| 4 | 0.76 | 0.98 | 0.025 | -4.72 |
| 5 | 1.02 | 1.00 | 0.025 | 0.29 |
| 6 | 0.98 | 1.01 | 0.045 | -0.48 |
| 7 | 0.89 | 1.02 | 0.105 | -2.21 |
| 8 | 2.34 | 1.03 | 0.115 | 25.72 |
| 9 | 1.01 | 1.11 | 0.245 | 0.09 |
| 10 | 1.00 | 2.34 | 1.335 | -0.09 |

The data points 0.76 and 2.34 have $M_i$ values of –4.72 and 25.72 respectively, and so are rejected when using a threshold of 3.5.

# APPENDIX G

## COST MODEL FOR OUTLIER REJECTION BENEFIT EVALUATION

A simplified cost model for evaluating benefit of outlier identification for test cost reduction is proposed.

**TABLE XXX. LIST OF VARIABLES USED IN COST MODEL.**

| Variable | Symbol |
|---|---|
| Wafer manufacturing cost/wafer | $T_w$ |
| Average Wafer Yield | $Y_w$ |
| Wafer probe cost per chip | $C_w$ |
| Additional test cost for outlier analysis per chip | $C_o$ |
| Packaging cost per chip | $C_p$ |
| Assembly test cost per chip | $C_a$ |
| Customer return cost including replacement per customer return | $C_r$ |
| Package Yield | $Y_p$ |
| Burn-in cost per chip | $C_b$ |
| No. of chips per wafer | N |
| Customer reject rate before outlier rejection | $R_1$ |
| Customer reject rate after outlier rejection | $R_2$ |
| BI failure rate before outlier screening (# of chips failing BI/total # of chips BI) | $B_{f1}$ |
| BI failure rate after outlier screening | $B_{f2}$ |
| Correct outlier screening ratio (# of chips screened and are bad/ total chips) | S |
| Incorrect outlier screening ratio (# of chips screened and are good/total chips) | YL |
| Correct outlier acceptance ratio (# of chips accepted and are good/total chips) | A |
| Incorrect outlier acceptance ratio (# of chips accepted and are bad/total chips) | DL |

Total functional chips/wafer = $NY_w$

Total chips tested good after packaging = $NY_wY_p$

Total chips tested good after BI = $NY_wY_p(1-B_f)$

Loss due to outlier screening used as BI replacement

= Customer return due to incorrect screening (defect level) + Profit lost due to incorrect screening (yield loss)

= $NY_wY_p(DL)C_a + NY_wY_p$

A. For a conventional test flow (no outlier screening) with BI

Total test cost

= wafer cost + wafer probe cost + packaging and assembly test cost + burn-in cost + customer return handling cost

= $T_w + NC_w + NY_w(C_p + C_a) + NY_wY_pC_b + NY_wY_pC_b\,(1-B_{f1})\,R_1.C_r$

B. For modified test flow with outlier screening and BI

Total test cost

= wafer cost + wafer probe cost + outlier analysis cost + incorrect outlier screening cost – correct outlier screening cost + packaging and assembly test cost + burn-in cost + customer return handling cost

$= T_w + NC_w + NY_wC_o + NY_w.(A+DL)(C_p+C_a) + NY_w(A+DL)Cb + NY_w(A+DL).(1-B_{f2})$

Loss due to incorrect outlier screening = $NC_wYL$(profit/chip)

Savings due to correct outlier screening

= saving due to reduced packaging and test

$= NC_wA(Cp+Ca)$

Using a model similar to this, it is possible to quantify cost savings due to early outlier identification. However, quantifying all numbers in these equations is not straightforward and may involve considerable engineering effort.

# VITA

Sagar Suresh Sabade.

136 Guruwar Peth,

Vedant Vidyapeeth

Satara, India 415002

E-mail: ssabade@yahoo.com

Mr. Sagar Suresh Sabade was born in Pune, India in 1973. He completed his pre-engineering education in Satara, India. He obtained a B.E. in instrumentation engineering from Pune University, Pune, India in August 1994, an M.Tech. in electronic design and technology from the Indian Institute of Science, Bangalore, India in January 1996 and a Ph.D. in computer engineering from Texas A&M University in May 2004. He worked as a Senior ASIC Design Engineer for Paxonet Communications (formerly CG-CoreEl Logic Systems), Pune, India from February 1996 to June 1999 before joining Texas A&M University for his doctoral studies in September 1999. At Paxonet, he designed Application Specific Integrated Circuits (ASICs) for ATM and SONET applications. He worked as a summer intern with Texas Instruments, Dallas in 2000 and 2001 where he was involved in production implementation of defect diagnosis tool and development of a web-based statistical test data analysis tool. His research interests are in the areas of VLSI testing, technology issues, and defect diagnosis. He is interested in understanding positive impact of technology of lives of common people that does not create a digital divide. His non-technical interests include reading, photography and poetry. He is a member of the IEEE, ACM, and ASEE.