Predictive Analysis of Safety Subsystems in Memories for Diagnostic Coverage

Student Name: Arvind Srinivasan
M.Tech-Electronics and Communication Engineering-2013-2015

June 2015

Indraprastha Institute of Information Technology, New Delhi

Advisors

Dr. Mohammad S. Hashmi

Mr. Dhori Kedar Janardan

Submitted in Partial fulfilment of the requirements
For the degree of M.Tech in Electronics and Communication Engineering

© 2015 Arvind Srinivasan
All Rights Reserved

Student's Declaration

I declare that the dissertation titled "Predictive Analysis of Safety Subsystems in Memories for Diagnostic Coverage" submitted by Arvind Srinivasan for the partial fulfilment of the requirements for the degree of Master of Technology in Electronics and Communication Engineering is carried out by me under the guidance and supervision of Dr. M. S. Hashmi at Indraprastha Institute of Information Technology, Delhi and Mr. Dhori Kedar Janardan at STMicroelectronics, Greater Noida. Due acknowledgements have been given in the report to all material used. This work has not been submitted anywhere else for the reward of any other degree.

………………………………
Arvind Srinivasan

Place and Date: ……………………...........

CERTIFICATE

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

……………………………..
Dr. Mohammad. S. Hashmi

ABSTRACT

Automobile safety is becoming an important subject because of the large number of electronic systems being integrated within a System on Chip (SoC) and with the use of current state-of-the-art technology, safety is becoming an even bigger challenge. ISO 26262 is a mandatory standard to ensure the safety performance in electronic systems in automobile. This standard also provides the required target of diagnostic coverage for various subsystems.

For any SoC, the proper working of the memory subsystem is very critical. Faults, occurring due to aging of devices, soft errors or manufacturing defects can lead to memory operation failures. Generally, memory system failures manifest themselves as Single Bit Failures (SBFs) and Multi Bit Failures (MBFs). The Single Error Correct-Double Error Detect (SEC-DED) Error Correcting Code (ECC) is very effective for detecting SBFs, but extremely limited in detecting MBFs, owing to implementation issues. Therefore, an On-chip Safety Circuit (OSC) or Safety Circuits (SC), which is basically a hardwired mechanism, is deployed to detect these MBFs. This dissertation discusses Online Monitoring Circuits, implemented in CMOS M40 technology, in conjunction with an Error Correcting Code, to detect both Single and Multi Bit Faults.

This dissertation also proposes algorithms of Predictive Analysis (PA) to provide the Diagnostic Coverage (DC) for the combination of ECC–On-chip Safety Circuit. This DC value is essentially a measure of the ability of the system to detect faults in field and can thus be a measure of the effectiveness of the safety subsystem in general. The first algorithm is used to predict DC for faults caused by manufacturing defects while the second algorithm predicts DC for in field faults.

ACKNOWLEDGEMENTS

The work for this dissertation was carried out at STMicroelectronics, Greater Noida, India, during the year 2014-2015.

Firstly, I would like to thank my advisors Dr. Mohammad. S. Hashmi and Mr. Dhori Kedar Janardan for providing excellent guidance and encouragement throughout the journey of this work. Without their guidance, this work would never have been a successful one. I also take this opportunity to thank my friends, Mr. Gundu Anil Kumar and Mr. Aditya Vasisth from IIIT-D for the technical discussion whenever I was in need of any. Last but not the least, I would like to thank my parents for supporting me spiritually and emotionally.

Publications:

1. K.J. Dhori, A. Kumar, A. Srinivasan and M.S. Hashmi, "Predictive Analysis of Diagnostic Coverage for Online Hard Fault Monitoring Circuits in Safe Memories", Submitted to IEEE Journal on Reliability. (Under Review)

2. A. Srinivasan and M.S. Hashmi, "Prediction of Local Spot Defects and Diagnostic Coverage of Safety Mechanisms in SRAMs", Submitted to IEEE Transactions on Instrumentation and Measurement. (Under Review)

LIST OF FIGURES

Figure 1: Block Diagram of SRAM architecture 11
Figure 2: Block diagram of SRAM architecture with online monitoring circuit (shown in dotted box) 13
Figure 3: Waveform for read operation 14
Figure 4: Physical representation of A(R) for a layout 16
Figure 5: Block diagram of safety mechanism and algorithm 19
Figure 6: Predictive analysis algorithm for manufacturing defects 21
Figure 7: Probability of occurrence of delay faults 22
Figure 8: Threshold resistance values 23
Figure 9: Detection probability variation with resistance 24
Figure 10: PDF of fault resistance 25
Figure 11: Coverage probability for various nodes/nets 26
Figure 12: Coverage probabilities for various safety mechanisms 27
Figure 13: Predictive Analysis flow diagram for in field defects 28
Figure 14: Resistance degradation profile 30
Figure 15: Coverage breakup for WL 33
Figure 16: Detect Times as a fraction of device lifetime 33
Figure 17: Coverage probabilities for various nodes 34
Figure 18: Coverage comparison for various safety mechanisms 34

CONTENTS

Certificate 1

ABSTRACT 2

ACKNOWLEDGEMENTS..3

Publications: 4

List of Figures..5

1 Introduction 7

1.1 Motivation..7

1.2 Thesis organization 8

2 Literature Review 10

3 SRAM and Safety Circuit architecture 11

3.1 SRAM Architecture 11

3.2 Causes of MBFs and OSC working 13

3.3 Limitations of the OSC 15

4 Critical Area Analysis..16

4.1 Introduction To Critical Area Analysis 16

4.2 Modelling Spot defects and Computation of Critical Area 17

5 Proposed Detection Mechanism and Diagnostic Coverage Algorithms 19

5.1 Proposed detection mechanism 19

5.2 Predictive Analysis Algorithm for Manufacturing Defects and Initial Testing..20

5.3 Predictive Analysis Algorithm for In Field Defects..28

6 Conclusions and Future work 35

References 36

1 INTRODUCTION

1.1 MOTIVATION

Faults, which are defined as improper functioning of a device, are bound to occur for any SoC , due to defects introduced through the manufacturing process , reliability related degradation as well as external effects like soft errors due to radiation. Unmitigated faults may cause may cause severe accidents. For example, failure of power brakes in car will lead to an accident. Hence, there is an increasing requirement of rigorous fault test regimes to ensure safety and reliability. Introduction of automotive standards such as ISO26262 [1] amongst others shows a clear trend of increasing awareness towards the importance of reliability analysis in device design. Furthermore, ISO26262 also advocates the use of reliability data such as FMEDA [2], requiring automotive suppliers to provide mandatory disclosures regarding the Diagnostic Coverage (DC) of the safety mechanisms provided by them. Requirements regarding Automotive Safety Integrity Levels (ASILs) warrant safety mechanisms capable of handling a much broader range of faults. ASILs are used to classify faults into several categories, based on the likelihood of occurrence of these faults, as well as the severity of the accident due to occurrence of these faults. A simple empirical formula for an ASIL may be calculated as

$$ASIL = Severity * (Likelihood\ of\ ocurrence) \tag{1}$$

ASILs are classified as A, B, C and D [1]. ASIL A class faults are said to be least hazardous and are generally not considered to be significant enough to require diagnostic coverage values for most of the safety subsystems. ASIL D is said to be the most hazardous and require diagnostic ranges of 99% for MBFs.

Memory IPs consume a significant area of the total SoC and are hence much more prone to manufacturing errors and reliability issues owing to their sheer sizes. This vulnerability requires robust fault identification and redundancy mechanisms to ensure optimal memory performance. Faults in memories cause the overall effect of bit failures which may be categorized into Single Bit Faults (SBFs) and Multi Bit Faults (MBFs). SBFs can be diagnosed by the basic Single Error Correct-Double Error Detect (SEC-DED) Error Correcting Code (ECC) [3] whereas MBFs require hardwired safety mechanisms [4] for detection.

Furthermore, faults, categorized as hard fault and soft fault that also transforms into hard fault, can cause permanent failures at device level and yield losses. All of these are generally caused by reliability related ageing phenomena, such as Electro Migration (EM) [5]-[7], Time Dependent Dielectric Breakdown (TDDB) [8]-[10], Hot Carrier Injection (HCI) [11]-[13] etc., and are usually modelled at layout level using spot defects through the use of resistance . There are mechanisms, such as BIST [14] and ATPG [15], for detection of Hard Faults. However, delay faults, a type of latent hard fault, are often much harder to detect through test patterns and other conventional techniques. Therefore, the estimation and detection of delay faults using known safety mechanisms fall below the diagnostic coverage requirements of ASIL D standards [1]. Hence, there is a need of a comprehensive detection mechanism, that takes into account both delay and hard faults. This need serves as the motivation for this dissertation.

This dissertation proposes a fault detection mechanism to identify the delay faults and algorithms to find the diagnostic coverage of the proposed mechanism under various testing scenarios. The new approach incorporates an On-chip Safety Circuit (OSC), which is a hardwired circuit and works in conjunction with ECC to detect delay defects which is then complemented by an effective algorithm to find the diagnostic coverage. The first algorithm determines diagnostic coverage in manufacturing error type delay faults, where fault effects are much more prominent. The subtler, lifetime degradation effects, are handled by the second algorithm.

1.2 THESIS ORGANIZATION

The dissertation is organized as follows. Chapter 2 briefly describes the various fault detection and DC evaluation techniques found in literature. Chapter 3 describes the basic SRAM memory and the OMC. Section 3.1 briefly describes the SRAM architecture. Section 3.2 describes the causes of MBFs and the working of the OSC. Chapter 4 describes Critical Area Analysis (CAA), a reliability technique, which is used to predict the probability of occurrence of spot defects. Section 4.1 gives a brief introduction to Critical Area Analysis. Section 4.2 describes several techniques to model spot defects and various approaches to compute critical area.

Chapter 5 presents the delay fault detection mechanism and the algorithms to predict diagnostic coverage. A detailed discussion on the algorithms is provided along with the stepwise results. Section 5.1 describes the fault detection carried out by using SEC-DED ECC and the OSC.

Section 5.2 describes the first algorithm, to predict the diagnostic coverage for delay faults caused by spot defects during the manufacturing process. Section 5.3 describes the second algorithm to predict the diagnostic coverage for in-field faults. Chapter 6 describes the possible future work and concludes the dissertation.

2 LITERATURE REVIEW

Test vector based approaches to detect presence of fault and hence find the net diagnostic coverage are commonly described in literature [16]-[17]. The Layout Under Test (LUT) is analyzed to determine all possible sites which are vulnerable to hard faults. Critical Area Analysis (CAA) is used to determine the probability of occurrence of these hard faults. Subsequently, a one-to-many functional mapping for these hard faults is carried out, to obtain fault functions which describe the effects of shorts and opens in the form of Boolean expressions. The diagnostic coverage of the test vector system is then calculated using the probability of occurrence of the hard faults. Since, not all faults are detectable by test vectors, coverage of 100% is not possible. Such techniques are useful for hard fault prediction for smaller designs as hard faults cause clear violation in the form of wrong logic outputs. However, hard faults, though catastrophic, are either removed during the initial testing phase or occur after the eventual useful life of the device, thereby being limited in effect during normal device runs. Moreover, test vectors do not predict presence of delay faults, since delay faults do not cause functional logic flips.

The use of Critical Area Analysis, in context of TDDB based shorting between various metal levels due to dielectric breakdown, to predict the occurrence of delay faults is described in [18] while [19] describes a mechanism of introduction of resistive changes to model open type defects within a smaller device design using a fault dictionary and predict overall coverage.[19] also uses a PDF of resistive fault occurrence in analysis.

All these techniques, have been carried out for smaller designs and have severe portability issues, on larger sized memory cuts owing to huge fault dictionary sizes. Techniques for diagnostic coverage, described in Chapter 5, use critical area analysis similar to [18], but work on open type defects. The PDF based resistance approach is incorporated within the first algorithm to model the variation of fault resistances, similar to [19], but uses CAA to reduce the size of fault dictionaries, a technique based on [16] . Furthermore, the second algorithm of Chapter 5 uses a lifetime estimation of Electro-migration to predict coverage, a technique not found in literature.

3 SRAM AND SAFETY CIRCUIT ARCHITECTURE

In chapter 3, the basic SRAM architecture and the On-chip Safety Circuit (OSC) are discussed. This chapter starts with the basic SRAM block diagram and briefly describes the various blocks which constitute the memory IP. The latter part of the chapter discusses the cause of MBFs, the OSC working and its limitations.

3.1 SRAM ARCHITECTURE

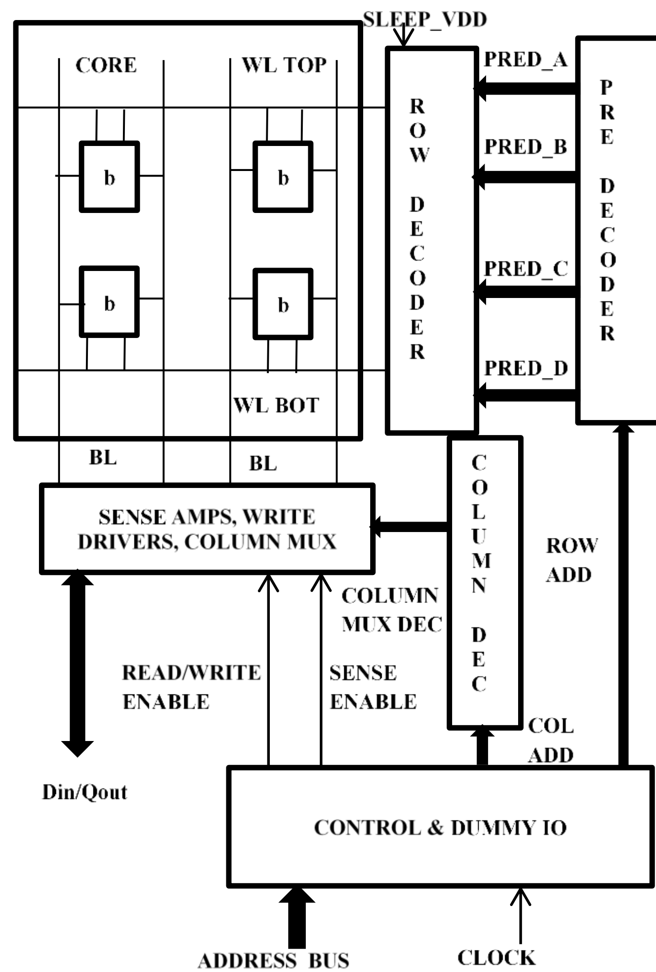

Figure 1: Block Diagram of SRAM architecture

Figure 1 shows the basic SRAM architecture. It consists of ROW DECODER block, memory array or CORE, Input / Output (IO) blocks and the CONTROL & DUMMY IO block. The ADDRESS BUS is subdivided into row and column addresses, ROW ADD and COLUMN

ADD, for proper memory aspect ratios by the CONTROL & DUMMY I/O block. The row addresses are further decoded into four pre-decoder signal buses, namely PRED_A, PRED_B, PRED_C and PRED_D using pre-decoder blocks. The primary task of the pre-decoder block is to reduce load on subsequent WORD LINE (WL) and reduce area of the WL generation block [20]. PRED_A is a clock qualified pre-decoded bus, while PRED_B, PRED_C and PRED_D are not qualified by clock. Pre-decoders of orders 2X4 and 3X8 have been observed, with PRED_A, PRED_B and PRED_C being generated by 2X4 decoders while PRED_D is generated by a 3X8 decoder. The ROW DECODER blocks decode the pre-decoded bus signals to WL signals which selects one row of the memory array for multiple word selection. Normally, ROW DECODER have high branching efforts, and are designed using trees of NAND gates and inverters [20]. The SLEEP_VDD signal provides power supply to the ROW DECODER in active state. The 6T SRAM cells are represented by the bit cells, b, in Figure 1. Only one of the set of multiple words are selected by the ROW DECODER is passed to the output by use of the COLUMN DEC block. The COL ADD bus is used to generate the COLUMN MUX DEC signal bus for selection of COLUMN MUX.

The bit lines associated with the pass transistors of the bit cells are passed to the I/O blocks which contain the sense amplifier (SENSE AMPS) circuits needed for read operation, as well as the WRITE DRIVERS, used for writing into the memory cell during write operation. The SENSE AMPS are regenerative feedback based circuits, which “amplify” the low differential voltages, generated across the bit cells during the read cycle, to proper logic levels. Clocked SENSE AMPS [20] are generally used, which require a “self-timing” clock signal for activation, mostly using a “dummy” bit cell read using “dummy” rows. The SENSE ENABLE signal, generated after the self-time duration, is used to generate the clock signal for the SENSE AMPs, which is then used to activate the sense amplifier to amplify the differential signal generated across the bit lines during read operation. Other signals required for both read and write operations, such as READ/WRITE ENABLE are also generated by the CONTROL & DUMMY IO. Din/Qout represents the input/output data.

3.2 CAUSES OF MBFS AND OSC WORKING

Spot defects only on some signals, such as WL or pre-decoder buses PRED_A, PRED_B, PRED_C, PRED_D and COLUMN MUX DEC, can manifest themselves as MBFs in the system as they directly affect a large number of words. In addition, special signals such as READ and SENSE ENABLE are also responsible for read failures for entire words. Hence, a comprehensive mechanism to identify MBFs in the system, should check for the proper generation of all these signals.

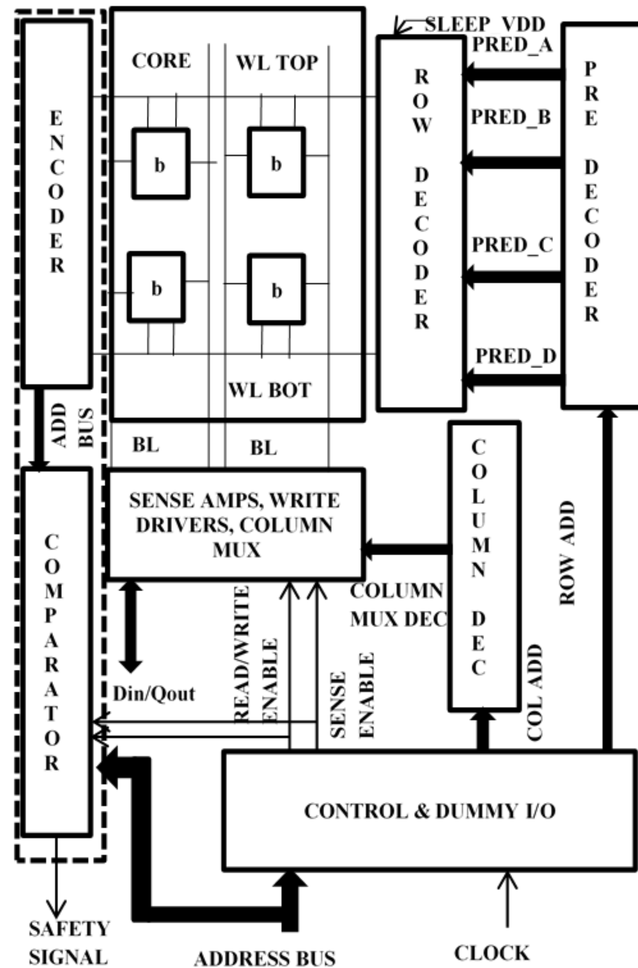

Figure 2: Block diagram of SRAM architecture with online monitoring circuit (shown in dotted box)

Figure 2 depicts the block diagram of the on-chip safety circuit incorporated in the SRAM. MBFs are identified by the online monitoring circuit using a simple address encoding procedure by the ENCODER. Here, WL/Column selection signals generated by address decoding are re-

encoded back using a ROM cell array inside the ENCODER. This generates back a copy of the address bus called ADD BUS as shown in Figure 2. This is then compared with the actual input address bus, ADDRESS BUS, latched in the same cycle, by the COMPARATOR. The presence of MBFs in the WL is signified by a signal low. A similar comparison is carried out in COMPARATOR for the COLUMN MUX DEC bus to check for its correct generation within the cycle.

Along with the address comparison, a comparison process to check proper generation of read cycle and write cycle signals is also carried out. During the read cycle, READ and SENSE ENABLE should be high while in write cycle they should be low. Output of this comparison process together with the comparison results of proper WL and COLUMN MUX DEC selection generates a final SAFETY SIGNAL, as shown in Figure 3. Presence of level high for SAFETY SIGNAL signifies the absence of any MBFs.

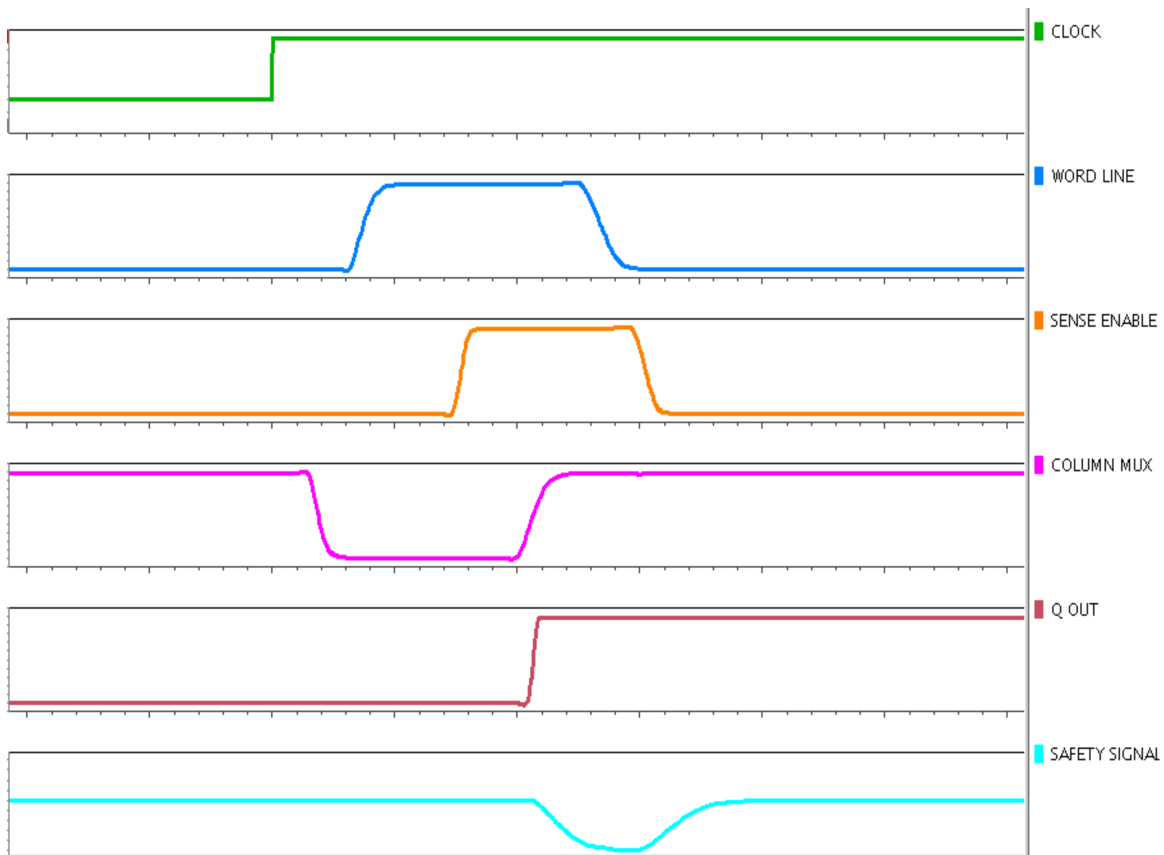

Figure 3: Waveform for read operation

3.3 LIMITATIONS OF THE OSC

The On-chip Safety Circuit however is not sensitive enough to smaller resistances (delay defects) variation along the nets during the life cycle of the device, which has been verified through simulations. These resistive (and capacitive) variations through the life of the device represent changes caused by aging effects such as EM, TDDB etc. Thus, the OSC is not able to match the 99% detection requirement criterion throughout the life of the device, required by the ISO 26262 standard, considering that it is rather insensitive to smaller delay faults and is also prone to small delay faults by itself. Furthermore, few nets tend to mask the effect of resistance changes at output (SAFETY SIGNAL) and hence virtually become undetectable by the online monitoring circuitry till a large defect resistance sets in. Hence, it becomes very essential for this limitation to be overcome, a task which is accomplished by using an ECC in conjunction with this circuitry.

4 Critical Area Analysis

Chapter 4 is dedicated to the description of Critical Area Analysis. The basic idea of critical area analysis, spot defect modelling and popular approaches to compute critical area are briefly mentioned in this chapter. The primary use of the Critical Area approach in this dissertation is to determine the probability of occurrence of delay faults, which when coupled with the probability of detection of faults by the safety mechanism, gives an overall coverage probability of the mechanism.

4.1 Introduction To Critical Area Analysis

Critical Area, a measure of sensitivity of a design to spot defects [21], is defined as the expectation of the area, $A(R)$, where the centre of the defect of radius R should occur to cause a fault[22]. It is defined mathematically according to (2), with D(R) being the defect density function of the spot defects, a function of the defect radius R [22].

$$A = \int A(R)D(R)dr \tag{2}$$

A physical representation of $A(R)$ is obtained from the layout diagram of Figure 4, with $A(R)$ highlighted.

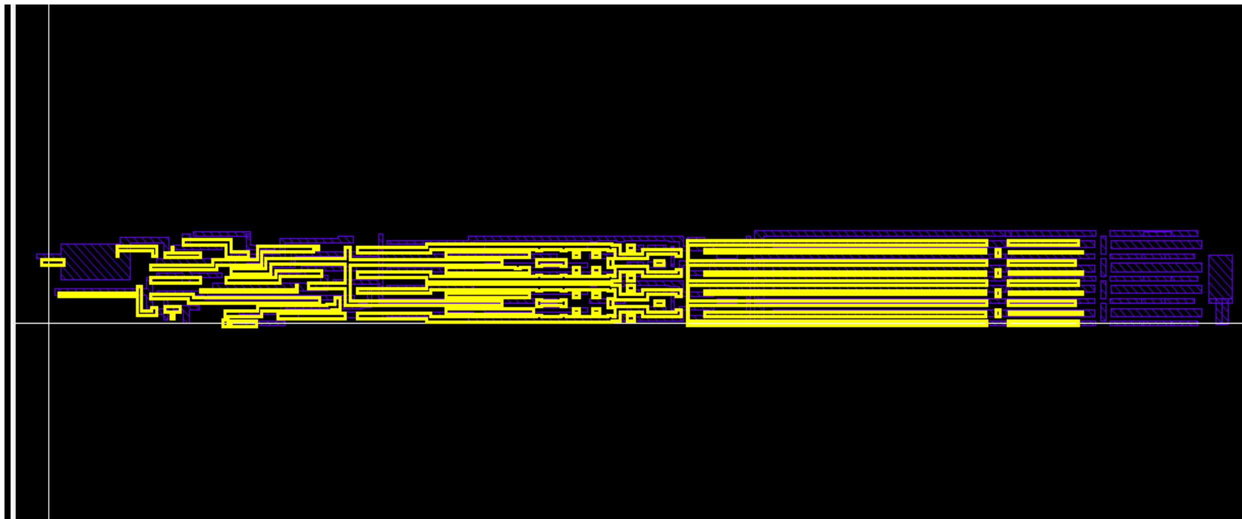

Figure 4: Physical representation of A(R) for a layout

The method of Critical Area Analysis is widely accepted for analysis of spot defect based reliability and yield phenomena [23].

4.2 MODELLING SPOT DEFECTS AND COMPUTATION OF CRITICAL AREA

The spot defects, though of irregular shape, may be modelled as circles or polygons. The polygon is supposed to be oriented so as to just touch the points of vulnerability (to be referred henceforth as P and Q) which determine the extent of the critical area parameter A, (R). Based on type of model used, distance between 2 points may be found by several metrics. For circular defects, the distance between two points P and Q, with co-ordinates (x_p, y_p) and (x_q, y_q), is found using the simple Euclidean distance estimation equation (3). This "distance" is the random variable R and is used calculate the value of A(R).

$$d_{p,q} = ((x_p - x_q)^2 + (y_p - y_q)^2)^{1/2} \tag{3}$$

For square type models, two types of distance metrics are available in literature [22], referred to as the L_∞ metric and the Manhattan or the L_1 metric. The L_∞ metric calculates the distance between two points, P and Q, is calculated using (4) and this distance serves as the upper bound of the defect size.

$$d_{p,q} = \max(|x_p - x_q|, |y_p - y_q|) \tag{4}$$

The other metric, called the L1 or the Manhattan metric, is based on a square rotated by 45^0 with respect to the centre. The distance between the points, P and Q, is calculated using (5) and serves as the lower bound of the defect size.

$$d_{p,q} = |x_p - x_q| + |y_p - y_q| \tag{5}$$

The square model, however, does suffer from problems. This approximation is unrealistic especially for 90^0 polygon bends and leads to overestimation of critical area and hence underestimation of reliability by about 25% as compared to circular defect [24].

Several approaches have been reported in literature to determine the Critical Area of a layout. Statistical approaches include Monte Carlo Analysis and layout sampling techniques. Monte-Carlo approach generates a large number of defects, varying defect radii according to the defect density function and fitting these defect radii into the layout to check the presence of faults. This

technique has the easiest implementation, but is also computationally intensive.

The layout sampling technique, generates random samples of the layout and computes Critical Area for those samples. These samples are combined together to generate the total, chip level, critical area. The CADENCE EYES system uses this approach to generate critical area data. This mechanism is however dependent on the ability of the system to generate accurate samples.

Deterministic Iterative Techniques to determine Critical Area include Shape Shifting techniques, which start with an approximate value of critical area, undergo iterative area expansion processes to find overlap areas among different nets to obtain a final value of critical area. Several variation of this shape shifting technique exist for different layout styles, including a maximum Critical Area Rectangle (CAR) technique for Manhattan style layouts.[22]. However, all these iterative techniques suffer from intensively long computational times.

Non iterative techniques of critical area determination include the Voronoi approach. Such techniques, though more accurate, have much harder implementation [22]. The primary use of the Critical Area approach in this dissertation is to determine the probability of occurrence of delay faults, which when coupled with the probability of detection of faults by the safety mechanism, gives an overall coverage probability of the mechanism.

5 PROPOSED DETECTION MECHANISM AND DIAGNOSTIC COVERAGE ALGORITHMS

This chapter describes a mechanism to detect delay faults and explains algorithms to predict the diagnostic coverage of this mechanism for various testing scenarios.

5.1 PROPOSED DETECTION MECHANISM

The OSC has its own limitation which can be overcome if it is complemented by the ECC in the determination of the coverage.

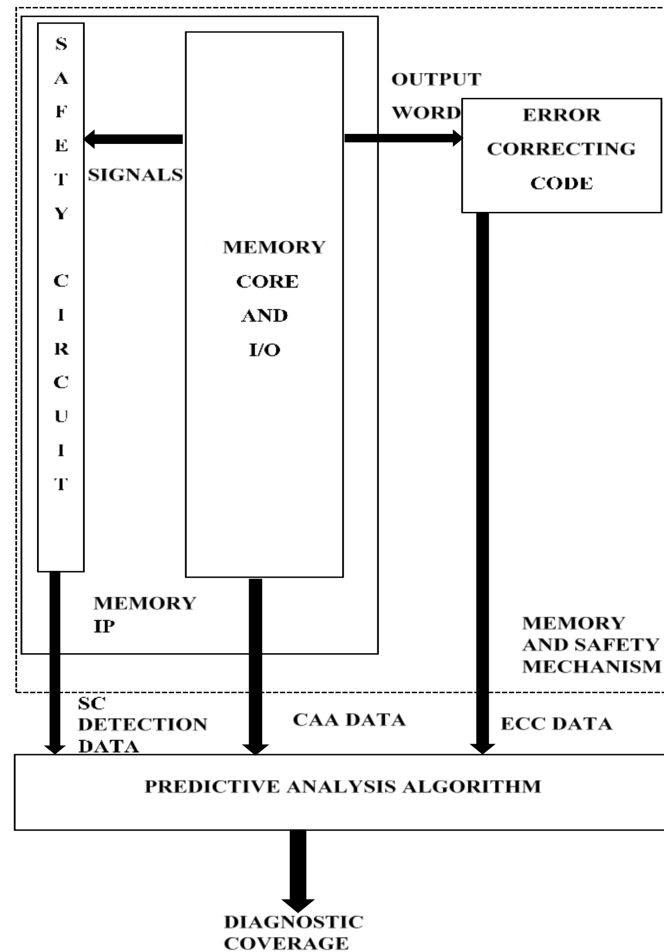

FIGURE 5: Block diagram of safety mechanism and algorithm

5.1.1 Fault Prediction by OSC

MBFs, as discussed previously, are best detected by use of OSCs which are hardwired circuit and can detect faults with very high probability once triggered. MBFs, which occur due to delay faults, tend to delay the generation of OSC input signals, which result in a delay in generation of the SAFETY SIGNAL. A delay of over 50% to the nominal SAFETY SIGNAL value can tide over any time borrowing mechanisms incorporated within the system and can hence be used as the lower threshold to predict activation of the OSC. Once triggered, detection probabilities of 99.9% can be achieved using the OSC.

5.1.2 Fault Prediction by ECC

The OSC, however, has a certain threshold resistance value below which it is unable to predict the presence of a fault. Hence, up to this resistive threshold, there is a virtual blind zone of fault prediction. ECC based detection can come into play here, to reduce the size of this blind zone of no information. Therefore the region where there is little idea of detectability is reduced greatly by use of the ECC based mechanism in conjunction with the SC. This adds on to the diagnostic coverage of the SC and improves the diagnostic coverage of the system in general. The overall block diagram of the proposed methodology is given in Figure 5.

5.2 Predictive Analysis Algorithm for Manufacturing Defects and Initial Testing

Predictive analysis algorithm proposed in this section is detailed in flow diagram of Figure 6. Essentially, this algorithm can be used for determining DC of the proposed safety mechanism for situations, where the memory suffers from manufacturing defects and has passed through the "Stuck At Fault" testing mechanisms such as ATPG or BIST, and has been declared safe. The Algorithm consists of five steps: 1) Critical Area Analysis (CAA) on the memory instance to identify the probability of occurrence of delay fault and identification of set of nodes/nets which are most critical, 2) determination of the lower threshold of OSC, 3) determination of diagnostic coverage by ECC, 4) determination of node wise coverage using PDF of fault resistance, and 5) prediction of final cumulative diagnostic coverage.

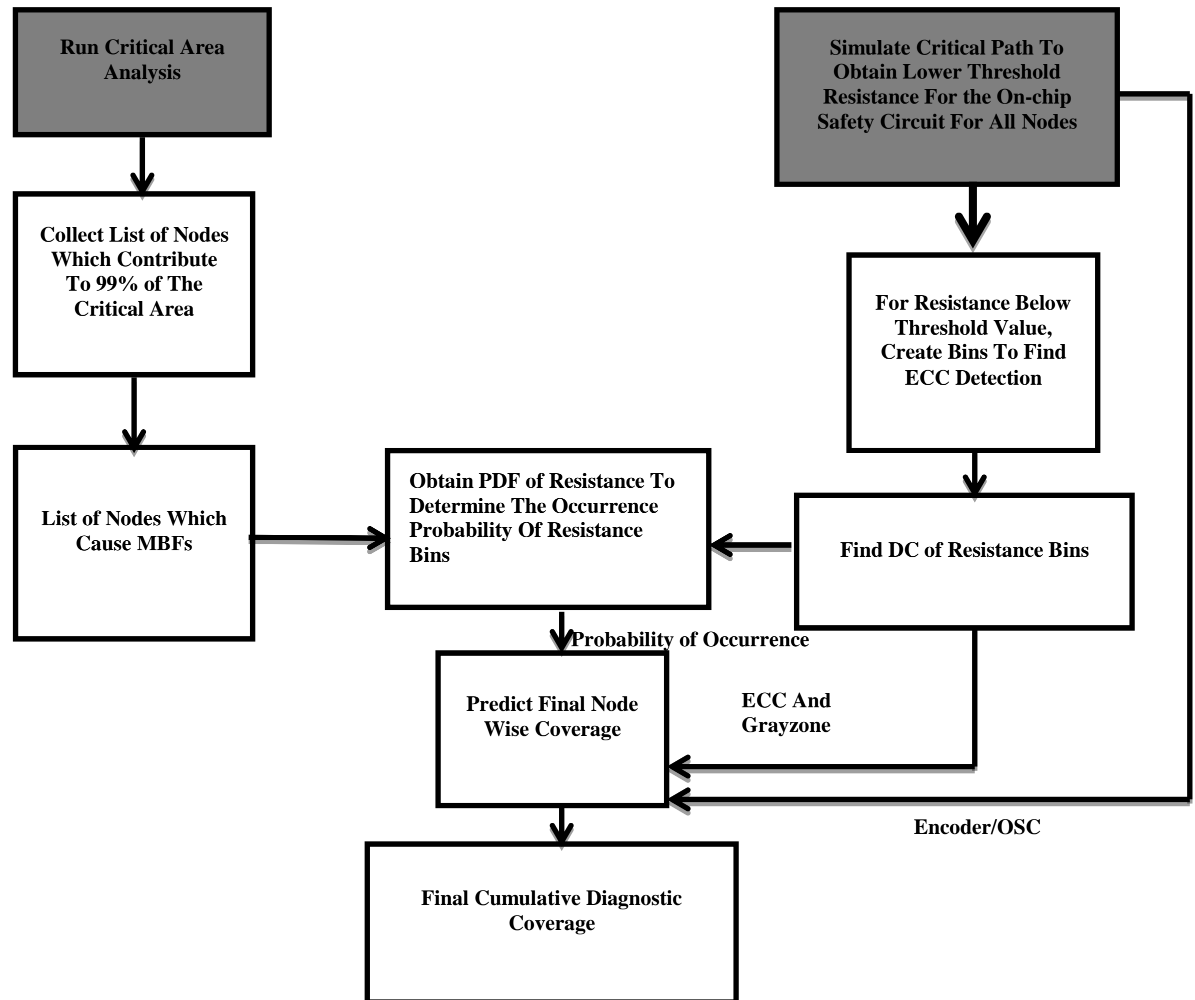

Figure 6: Predictive analysis algorithm for manufacturing defects

1. CAA on memory instance/layout

Critical Area is defined as statistical expectation of the region where if the center of the defect of a particular radius occurs, it is sure to cause a short or open defect. CAA determines the sensitivity of a layout to short or open type faults and can be used to prune down the list of nodes which are most hazardous for a layout. A database of nodes with the critical area associated with each node is obtained for a given defect density and defect radius. MBF causing faults which amount to over 99% of the total critical area are recorded. The net output of this step is a set of nodes and the probability of occurrence of delay faults for these nodes. Figure 7 gives the plot of probability of occurrence of delay faults, P_{Delay_fault}. CAA was executed on a 1024X 8 MUX16 instance of SRAM of SPHD compiler for M40 technology , using Calibre DFM tool with deck

written for CAA. Results include only those nodes which are responsible for MBFs and whose critical areas have summed up to 99% of the total critical area. As expected, it is clear from the result that WL which occupies the maximum area has the maximum probability of occurrence of delay fault.

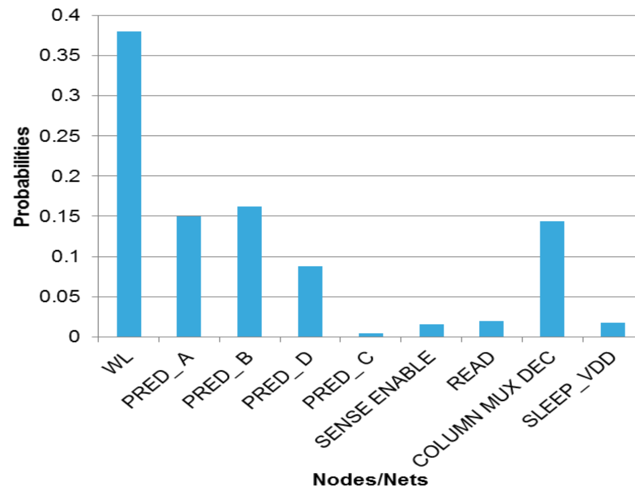

Figure 7: Probability of occurrence of delay faults

2. Minimum threshold of resistance for OSC

There is a definite threshold resistance, $R_{threshold}$, at which sufficient delay is obtained in generation of the SAFETY SIGNAL to indicate the occurrence of MBFs. This can be simulated by modelling the defect as resistance across nodes/nets determined through the CAA step, and by checking whether the SAFETY SIGNAL has been delayed sufficiently to the extent of 50%. The final output of this step is values of $R_{threshold}$ for all CAA dominating nodes. The plot for the threshold resistance of various nodes/nets is shown in Figure 8. For this step, resistances are added in extracted net list of the memory instance, which represent the layout level spot defects. The extracted net list, with added fault resistance is simulated on XA simulator and the value of the SAFETY SIGNAL output is checked and value of resistance which accounts for a 50% delay in SAFETY SIGNAL is noted. Non synchronous nodes/nets such as pre decoder nodes, PRED_B, PRED_C and PRED_D generally have much higher threshold resistance values as compared to other nodes. These nodes tend to stabilize much before the arrival of the clock triggered PRED_A node and small resistance changes across these nodes have little impact. At

very large resistance values, the steady state values of output voltage across these nodes begins to fall, and eventually cause word line to fail. This clearly shows a limitation in the OSC's sensitivity to non-clocked nodes, and can be a latent reliability problem for this mechanism.

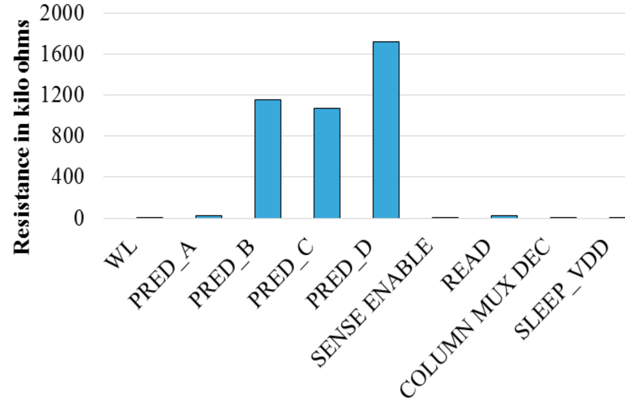

Figure 8: Threshold resistance values

3. Determination of diagnostic coverage by ECC for resistances smaller than $R_{threshold}$

There may be instances, such as random dopants or presence of aging effects, which cause generation of weak bits inside the memory core if certain signals such as WL are impacted by those effects. This phenomenon can be detected by failure of the SENSE AMP circuit. Access transistors of SRAM being directly linked to the bit lines are used to find the probability of generation of weak and failing bits. A simple monte-carlo (MC) analysis can be carried out to find the value of the standard deviation threshold voltage, V_{th}, of the access transistors. Introducing an offset, δ_{vto}, in the V_{th}, the values of standard deviation, σ, qualified by the SENSE AMP for this offset can be estimated. The normalized value of δ_{vto}, which causes a 6σ of SENSE AMP to fail, can then be used to determine the probability P_{bit_fail} of the weak bit failure using (6). Here, X is a random variable which defines the normalized standard deviation of access transistor, $\sigma_{access_transistor}$, of 6T SRAM cell qualified by the defect. Here it is also assumed that δ_{vto}has a Gaussian distribution.

$$P_{bit_fail} = P\left(X > {}^{\delta_{vto}}/_{\sigma_{access_transistor}}\right) \quad (6)$$

Here, the entire resistance from 0Ω to $R_{threshold}$ is divided into suitable sized resistance bins,

and for each resistance bin, the value of P_{bit_fail} is calculated. Subsequently, equation (2) gives the value of the parameter $P_{ECC\text{-}detect}$ which gives the probability of up to n bits in error out of a k bit word. It is the simple discrete binomial distribution formula with i as an index parameter representing the number of failing bits.

$$P_{ECC-detect} = \sum_{i=0}^{n} {}^{k}_{i}C * \left(P_{bit-fail}^{i} * (1 - P_{bit-fail})^{k-i}\right) \tag{7}$$

It can be inferred from (6) that the ECC detection drops as the resistance bins move closer to $R_{threshold}$. Such a resistance range where there is a fall in ECC detect from values of 99% to significantly low values constitutes the Gray zone region as shown in Figure 9. It demonstrates that the Gray zone after using ECC along with the OMC is substantially smaller.

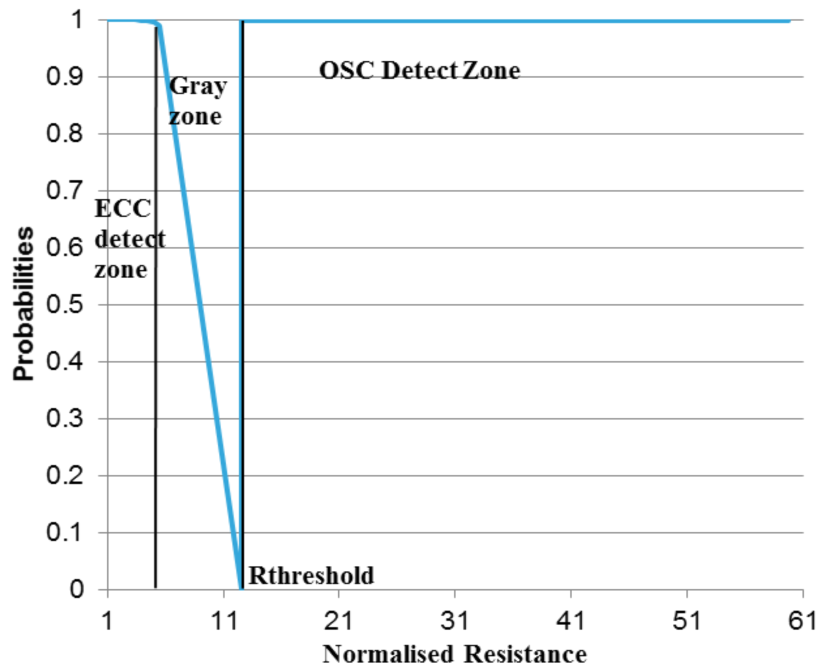

Figure 9: Detection probability variation with resistance

4. Determination of Node-Wise coverage using PDF of resistance

To find the probability of occurrence of the bins, where bins are obtained by dividing 0Ω to $R_{threshold}$ into suitable sized resistance units, there is a need of probability density function (PDF) of fault resistance occurrence. A typical fault resistance PDF is shown in Figure 10 [19]. It

is clear from Figure 10 that resistances beyond 100 MΩ have very little probability of occurrence and it is logical as well for practical systems. Now the bin wise resistance occurrence probability using ECC together with the detection probability of SC can be used to calculate the coverage for a single node/net.

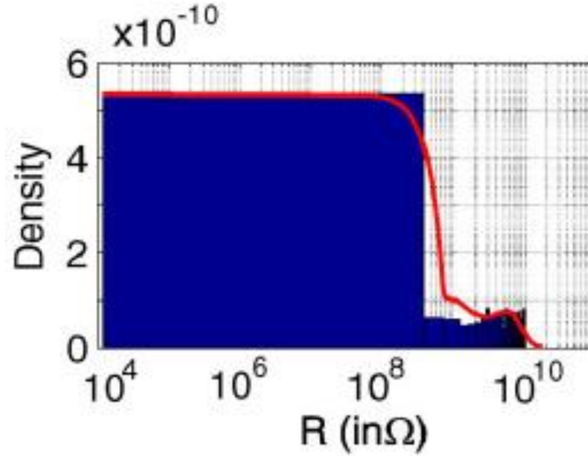

Figure 10: PDF of fault resistance

Now the bin wise resistance occurrence probability using ECC together with the detection probability of OSC can be used to calculate the coverage for a single node/net. Following equation (8) can be used to determine the net ECC coverage probability, $P_{ECC_Coverage}$. Here, $P_{resistance_bin_j}$ is the probability of occurrence (which is identified by utilizing Figure 10) and $P_{ECC_detect_j}$ is the ECC detection probability of bin j.

$$P_{ECC_coverage} = \sum_{j=1}^{n}\left(P_{ECC_detect_j} * P_{resistance_bin_j}\right) \tag{8}$$

The OSC coverage probability, $P_{OSC_coverage}$, can be determined using (9). In this expression, P_{OSC_zone} is the probability of occurrence (which is identified by utilizing Figure. 10) of zone of resistance from $R_{threshold}$ to the upper threshold resistances for delay faults and P_{OSC_detect} is the detection probability of the OSC which is known *a priori*.

$$P_{OSC_coverage} = P_{OSC_detect} * P_{OSC_zone} \tag{9}$$

Finally, the node wise coverage probability, $P_{coverage}$, is determined using (10).

$$P_{coverage} = P_{ECC_coverage} + P_{OSC_coverage} \tag{10}$$

Figure 11 gives the coverage probability of nodes/nets, obtained by this approach. As seen from the results, including ECC together with OSC improves coverage by a great extent. Nodes, with very high values of threshold resistance have much lower coverage probabilities because with the increase in threshold resistance the range of resistances for which the OSC is active tends to reduce. An exception to this observation is seen in the coverage probability of node PRED_A. This node does not seem to cause degradation in SENSE AMP performance and thus resistance changes are never detected. This may seem to be a highly dangerous situation but as long as the output word seems to be unaffected it can be safely ignored. It is to be noted however that the coverage values thus obtained are pessimistic in nature, as they are calculated for a $6\,\sigma$ sense amplifier, a rather stringent standard. For lower qualification standards, such as $3\sigma,$ much better coverage data can be obtained.

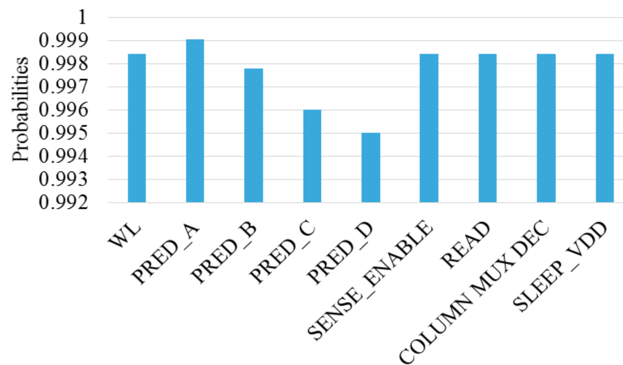

Figure 11: Coverage probability for various nodes/nets

5. Prediction of final cumulative coverage

Once the coverage data for various nodes/nets have been obtained, the total coverage probability can be determined by the theorem of total probability given in equation (11). Here the coverage, $P_{coverage_m}$, for each net m is multiplied by the probability of occurrence of delay fault, $P_{delay_fault_m}$, for that net. This weighted net wise sum gives the final coverage for the instance using both the OSC and ECC, with the probability of delay fault having been already obtained through Critical Area Analysis (CAA) explained in step 1.

$$P_{total\ coverage} = \sum_{set\ of\ nets(i)} (P_{coverage_m} * P_{delay_fault_m}) \quad (11)$$

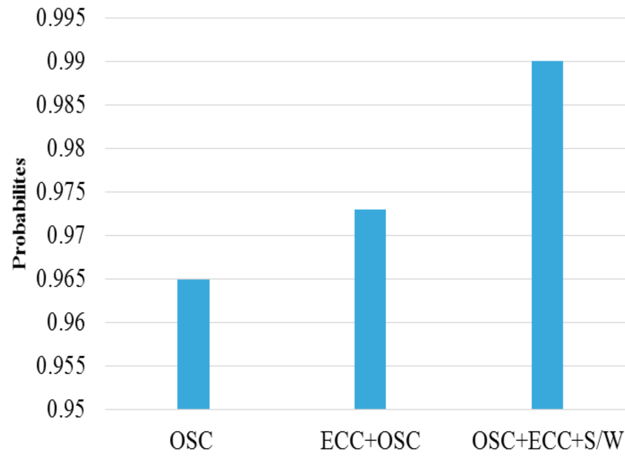

FIGURE 12: Coverage probabilities for various safety mechanisms

The coverage probabilities for various mechanisms such as OSC only, ECC and OSC and OSC, ECC and software hit per cycle is shown in Figure 12. Clearly, ECC provides improvement in coverage for on the basic OSC. For further coverage improvement, a software check per ECC hit can further confirm the presence of fault and improve diagnostic coverage.

The primary uncertainty in the above approach is the probability density function of the resistances that occur in the field. The distribution used in this approach, although closely monitors occurrence probabilities for situation where there may be initial high resistance faults, certainly does not reflect the situation, as observed during ageing effects such as electro migration. These effects cause gradual small resistance changes, which according to the described PDF never occur. As a result of this complication, this algorithm results in highly optimistic coverage results for in field, ageing related degradation effects. Other resistance PDFs may be described in the form of Gaussian distributions, with mean shifted as per degradation caused by aging effects. While, more accurate than the PDF of Figure 10, a new problem is encountered. Such a PDF is very complicated to generate, and requires a net integration over entire lifetime data. Such an effort is not justified due to large calculation times as well as lack of entire lifetime data availability. As a result, a new algorithm is described in the subsequent section, which also takes into account real life resistance variations, owing to degradations caused by EM.

5.3 PREDICTIVE ANALYSIS ALGORITHM FOR IN FIELD DEFECTS

Predictive analysis algorithm proposed in this section is detailed in flowchart of Figure 13.The algorithm has mainly six steps: 1) Critical Area Analysis on the instance to identify the probability of occurrence of delay fault and identification of set of nodes/nets which are most critical, 2) Assessment of the impact of EM on metal tracks which leads to increases in resistance, 3) Determination of the lower threshold of online monitoring circuit, 4) Determination of the diagnostic coverage of ECC for resistance less than the threshold of online monitoring circuit, 5) Prediction of time interval for various regions with EM/life aging, and 6) Prediction of final cumulative diagnostic coverage. While steps 1, 3 and 4 and 6 remain same as steps 1, 2 and 3 and 5 of the previous algorithm, in order to take into consideration the slow degradation of metal tracks due to electro migration, two new steps have been incorporated. Step 2 assesses the impact of Electro migration on wire track, while step 5 estimates the probability of

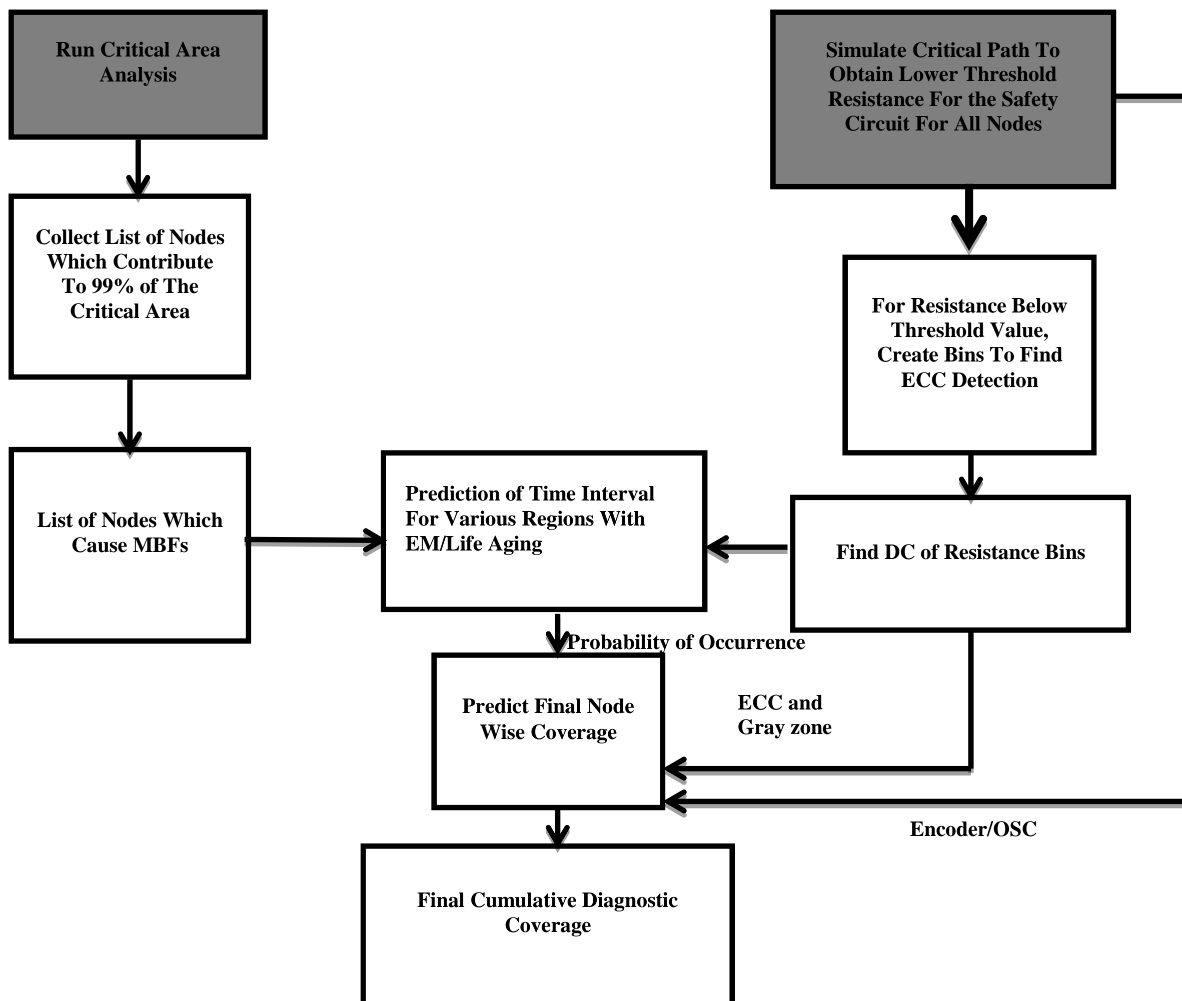

Figure 13: Predictive Analysis flow diagram for in field defects

occurrence of electro migration based resistance changes in the lifetime of the device. Also, an average detection probability of 0.5 for ECC is assumed for gray zone in this algorithm.

2. *Assessment of the impact of Electro migration based aging effect on metal tracks*

Electro migration is defined as the flow of conductor metal atoms or mass transfer of conductor metal atoms due to presence of impressed DC electric field in the direction of electron flow. Sufficient evidences is presented in [25] to predict a more or less general trend of metal wire degradation due to electro migration. The metal wires experience no initial resistance increase with high applied field, during which one or more than one voids nucleate within the material because of the absence of the metal atom. However, with time, in presence of such conditions these voids tend to coalesce together creating "pockets" which tend to expose TaN/Ta barrier material surrounding the metal tracks, thereby suddenly increasing the resistance of the wire in consideration, detected by a step growth in resistance varying in size from 15 to 30% of the nominal metal line resistance. Further degradation causes a more or less linear increase in metal resistance with time, eventually leading to creation of opens in the line [25].

The nominal resistance of the metal/signal under consideration is calculated through a simple extraction process. Design for Manufacturing (DFM) data for metals for a given technology is studied to obtain the value of sheet resistance of the metal wire level associated with the given net in consideration. Slope of the resistance growth during the linear growth period is obtained from foundry data. Further, cases for reduced metal width owing to presence of spot defects are also considered. Simple calculation using the sheet resistance and metal width data along with defect size data helps to find a first order formula for increase in nominal resistance with spot defect caused by metal width reduction. The overall increased resistance in presence of spot defect coupled with the value of initial resistance, $R_{degraded}$, is given by (12), where x is the width of the spot defect size and w is the minimum width of the metal wire in consideration. All the lengths are in µm and R_{sheet} is in mΩ/µm^2. Availability of defect density file enables the determination of the probability of occurrence of defect of particular size on the net based on a simple defect density ratio formula given by (13), with P_{defect_i} signifying the probability of occurrence of spot defect of radius d_i on the metal track. The usual size of defects ranges from 17nm to 69nm (referred as 17 and 69 in (13)), ranging from nearly quarter of the metal width to

almost whole metal width size. Although it is a first order expression, it sufficiently determines the probability of occurrence of the defect.

$$R_{degraded} = R_{nominal} + R_{sheet} * \left(\frac{x}{w}\right)^2 / \left(1 - \left(\frac{x}{w}\right)\right) \tag{12}$$

$$P_{defect_i} = {1/d_i}^2 \Big/ \sum_{i=17}^{69} 1/d_i^2 \tag{13}$$

Figure 14 is a graphical representation of variation resistance degradation with EM for no defect condition as well as with defects. For the no defect situation, the graph clearly shows a very slow degradation of the wire due to electro migration. With the increase in size of defect, a much quicker rate of metal track degradation is clearly observed, causing a rapid resistance increase early in the life of the device. The final output of this stage is a set of resistances, each for particular defect radius and probability of occurrence of spot defect of particular radius on a particular line.

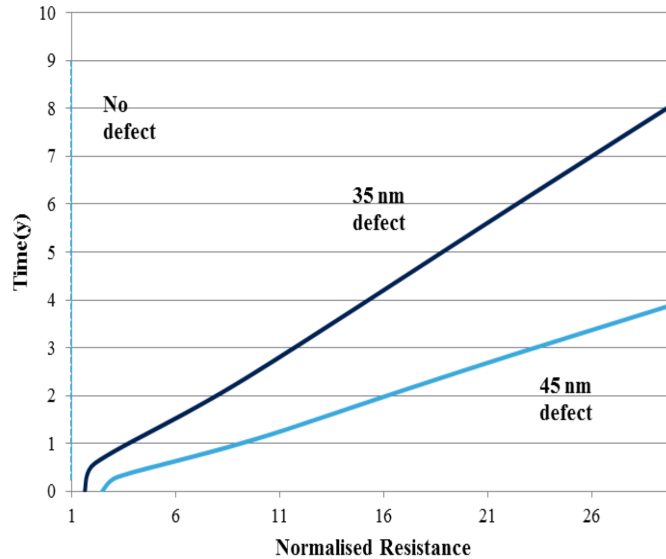

Figure 14: Resistance degradation profile

5. *Prediction of time interval for various regions with EM/life aging*

Mainly three regions describe the life of delay defects moving from its nominal value at initial time of operation to the End of Life (EoL). Time period/interval for which defect resistance

remains in these regions need to be determined to find the probability of occurrence of respective region.

Equation (14) is used to calculate the acceleration factor (AF) to compensate for the increased current densities found in defect degraded metal track as compared to the actual maximum permitted current density values through the metal track for a given technology. It is based on the Black's law approximation [26] for mean time to failure which places an inverse square law dependency of time to failure with current density, where J_2 is the increased current density under defects while J_1 is the minimum current density for the appropriate line without defect, obtained from technology data.

$$AF = \left(\frac{J_1}{J_2}\right)^{-2} \quad (14)$$

Now, as previously discussed, the bins of resistance from 0 to $R_{threshold}$ have probabilities of occurrence which needs to be estimated for given defect. It can be estimated through the time period for which this bin exists in the life of the device and that time can then be used to predict the probability of occurrence of this bin. The expected time period of bin existence in the lifetime, which is the weighted average of bin time period for various defect sizes, is estimated by equation (15). The term R_{slope_i} [25] is the rate of change of resistance with time due to EM degradation, for each defect radius, and is obtained from foundry data. The term R_{bin} is the bin resistance size. The index j is the total number of resistance bins in the ECC detect zone.

$$T_{bin_j} = \sum_{i=17}^{69}\left(\left(R_{bin_j}/R_{slope_i}\right) * (AF_i) * P_{defect_i}\right) \quad (15)$$

The probability of occurrence of bin_j is calculated by dividing the weighted average value of time of occurrence of bin_j by the total *device lifetime* given in (16).

$$P_{bin_j} = \frac{T_{bin_j}}{device\ \ lifetime} \quad (16)$$

The occurrence probability, denoted as P_{zone_1}, of ECC detect zone, defined as $zone_1$, can be expressed by equation (17). In this expression, L is the total number of bins within the $zone_1$. The lower resistance threshold for $zone_1$ is determined by $R_{degraded}$. The upper resistance

threshold is identified by R_L, the situation when $P_{ECC\text{-}detect}$ drops below 99%.

$$P_{zone_1} = \sum_{j=1}^{l} P_{bin_j} \tag{17}$$

Then eventually, the probability of coverage of *zone_1* can be obtained from equation (18).

$$P_{coverage\ zone_1} = \left(\sum_{j=1}^{l} P_{ECC-detect\ j} * P_{bin\ j}\right) \tag{18}$$

For gray zone the expected time of existence, T_{gray_zone}, is obtained from equation (19). The term R_{gray_zone} is the resistance range, lies between R_L and $R_{threshold}$, when detection probability is less than 99%.

$$T_{gray_zone} = \sum_{i=17}^{69} \left(\left(R_{gray\ zone}/R_{slope\ i}\right) * (AF_i) * P_{detect\ i}\right) \tag{19}$$

Now the probability of occurrence of gray zone, denoted as P_{zone_2}, can be found from equation (20).

$$P_{zone_2} = \frac{T_{gray_zone}}{device\ lifetime} \tag{20}$$

Then the coverage probability for any node/net, which is defined as the total detection probability of the ECC and OSC combination, can be found from equation (21).

$$P_{coverage} = P_{coverage\ \ zone_1} + 0.5 * P_{zone_2} + P_{OMCdetect} * P_{zone_3} \tag{21}$$

The term $P_{OSCdetect}$ in the above expression refers to coverage provided by the OSC while P_{zone_3} refers to the occurrence probability when the OSC comes into play for detection and is given by (22).

$$P_{zone_3} = 1 - (P_{zone_1} + P_{zone_2}) \tag{22}$$

Figure 15 shows the net breakup of coverage probability for a single net (WORD LINE) for different zones for the lifetime of the device. As can be clearly seen, there is a drop in ECC detect during the gray zone period.

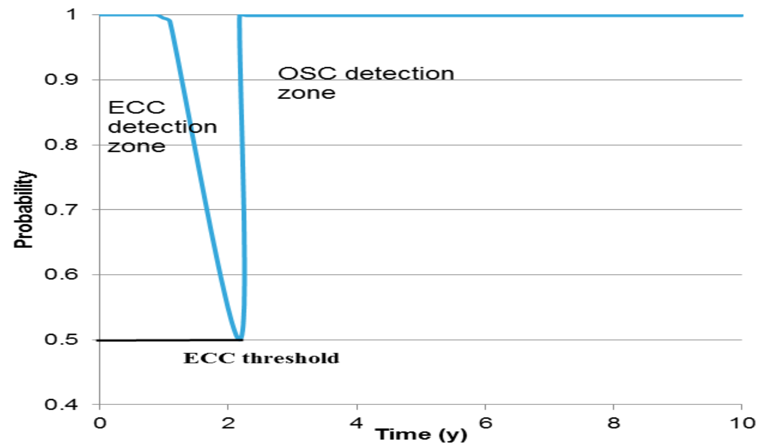

Figure 15: Coverage breakup for WL

The breakup of time of the 10 years useful life period in 3 zones i.e. ECC detect zone, Gray zone and OSC detect zone, of the memory for various nets is shown in Figure 16. It is apparent from the results that the gray zone forms a very small part of the lifetime of the net in consideration owing to the high detectability of this ECC-Circuitry combination.

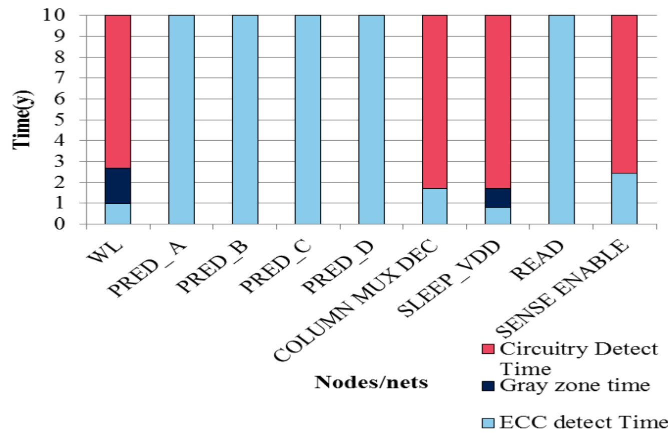

Figure 16: Detect Times as a fraction of device lifetime

The coverage probability data for various nets using ECC, Gray zone and On-chip Safety Circuit is given in Figure 17. It is obvious from the results in Figure 17 that a good value of detectability is obtained by including the ECC mechanism in conjunction with the Online Monitoring Circuit. It is also apparent that diagnostic coverage still does not reach 99%, as targeted for ASIL D

level, even after considering both the ECC and Online Monitoring Circuit. Therefore an additional mechanism is needed to improve the DC to the required target figures.

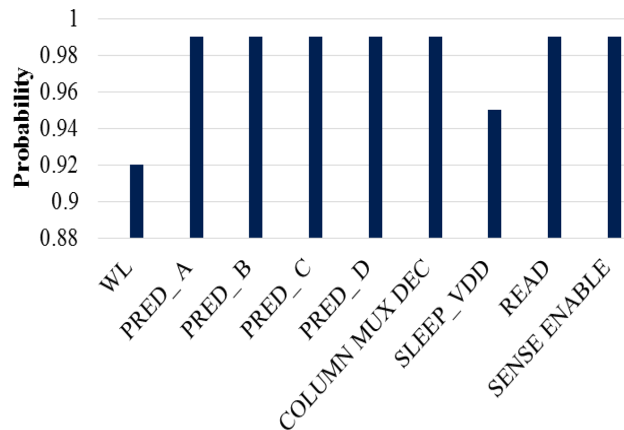

Figure 17: Coverage probabilities for various nodes

A small software check of writing and reading the same location after each ECC hit will ensure that there is a delay defect if ECC again gives a hit in subsequent cycle. This is true as soft errors probability in two consecutive cycles is almost negligible. Figure 18 gives the value of final coverage trends for all the nets for cases using 1) using OSC but without ECC, 2) the combination of OSC and ECC, and 3) the combination of ECC, OSC as well as a software routine. Just to reiterate, the software is used to provide additional coverage which thereby increases the detectability and hence diagnostic coverage to target values.

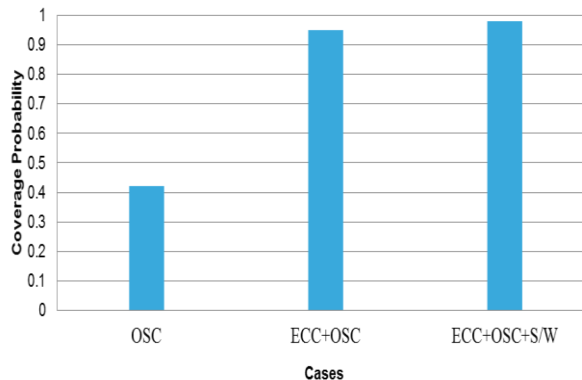

Figure 18: Coverage comparison for various safety mechanisms

6 CONCLUSIONS AND FUTURE WORK

A methodology to detect delay faults has been discussed in this dissertation. Algorithms to predict the diagnostic coverage of this methodology have also been discussed. The dissertation discusses OMCs as an MBF detection mechanism. Diagnostic Coverage analysis have been carried out for both manufacturing related faults as well as in field faults. Due to lower sensitivity, the OMC is not capable of fault detection below a certain threshold value. This limitation can be overcome by using the OMC in conjunction with ECC and therefore the diagnostic coverage can be improved significantly. Eventually, with additional software check of memory at each ECC hit one can reach the target diagnostic coverage of 99%.

This dissertation discusses algorithms only for the case of open defects, primary cause of which is Electro migration. Future work deals with modelling the effect of short defects and finding diagnostic coverage of the safety mechanism in case of short defects. The PDF for the short resistance [19] and the Algorithm of section 5.2 can still be used in this regard.

In field short defects require further analysis. Primarily, shorts between metals belonging to different layers, can be modelled through the Time Dependent Dielectric Breakdown (TDDB) effect. However, a proper short resistance growth mechanism for same metal layer level is much harder to analyze by this process and requires additional modelling for the metal atom transfer during void nucleation process, which may cause metal layer branching, causing metal shorts. The lifetime analysis of shorting mechanisms can then be plugged into Step 2 and Step 5 of the algorithm of section 5.3 to determine the diagnostic coverage of the proposed Safety Mechanism. Eventually, the final diagnostic effect, considering both open and short type defects can then be calculated.

REFERENCES

[1] ISO/DIS 26262, June 28 ,2009, ISO Standard (Draft International Standard).

[2] J.V. Bukowski and W.M. Goble, “Predictive analytics: Assessing failure rate accuracy & failure mode completeness,” in Proc. Reliability and Maintainability Symposium (RAMS), pp. 1-7, Jan. 2013.

[3] A. Neale and M. Sachdev, “A new SEC-DED error correcting code subclass for adjacent MBU tolerance in Embedded Memory,” IEEE Transaction on Device and Materials Reliability , Vol. 13, pp .223–230, Dec. 2012.

[4] D. Gizopoulos, M. Psarakis, S.V. Adve, P. Ramachandran, S.K.S. Hari, D. Sorin, A. Meixner, A. Biswas and X. Vera, “Architectures for online error detection and recovery in multicore processors,” in Proc. Of Design, Automation & Test in Europe Conference & Exhibition (DATE), pp. 1-6, March 2011.

[5] P.B. Ghate, “Electromigration induced failures in VLSI interconnects,” in Proc. 20th Annual Reliability Physics Symposium, pp. 292-299, Mar. 1982.

[6] Tan Ming Cher, and K.N.C. Yeo, “Reliability statistics perspective on standard wafer-level electromigration accelerated test(SWEAT), ” in Proc. Canadian Conference on Electrical and Computer Engineering , Vol. 1, Mar. 2000.

[7] K. N. Tu, “Recent advances on electromigration in very-large-scale integration of interconnects,” IEEE J. Appl. Phys., Vol. 94 , pp. 5451–5473, Nov. 2003.

[8] S.A. Sahhaf, R. Degraeve, Ph. Roussel, T. Kauerauf, B. Kaczer and G. Groeseneken , “TDDB reliability prediction based on statistical analysis of hard breakdown including multiple soft breakdown and wear out ,” in Proc. International Electron Devices Meeting (IEDM), pp. 504-507, Dec. 2007.

[9] E. Wu, E.J. Nowak, R.P. Vollertsen and L.K. Han “Weibull breakdown characteristics and oxide thickness uniformity,” IEEE Transactions on Electron Dev. , Vol. 47, pp. 2301-2309, Dec. 2000.

[10] B. Kaczer, R. Degraeve, R. O’ Conner, Ph. Roussel and G. Groeseneken, “Implications of progressive wear-out for lifetime extrapolation of ultra-thin (EOT= 1nm) SiON films,” IEDM Techn. Dig., pp. 713-716, 2004.

[11] K.L. Chen, S.A. Saller, I.A. Groves and D.B. Scott, “Reliability effect on MOS transistors due to Hot-Carrier Injection ,” IEEE Journal of Solid State Circuits , pp. 306-313, Feb. 1985.

[12] S. Sze, “Physics of Semiconductor Devices,” Wiley Publication.

[13] K.K. Ng and G.W. Taylor, “Effect of hot-carriers trapping in n- and p- channel MOSFETS,” IEEE Trans. Electron Devices, Vol. 30, pp. 871-876, Aug. 1983.

[14] Wen Xiaoqing and Wang Hsin-Po, “A flexible logic BIST scheme and its application to SoC designs,” in Proc. 10th Asian Test Symposium , pp. 463, Nov. 2001.

[15] K.H. Tsai, R. Tompson, J. Rajski and M. Marek-Sadowska, “STAR-ATPG: a high speed test pattern generator for large scan designs,” in Proc. International Test Conference (ITC), pp. 1021-1030, Sep. 1999.

[16] M. Blyzniuk, I. Kazymyra, W. Kuzmicz, W.A. Pleskacz, J. Raik and R. Ubar, “Probabalistic Analysis of CMOS defects in VLSI circuits for test coverage improvement,” Elsevier J. Microelectronic Reliability, pp. 2023-2040, Mar. 2001.

[17] M. Blyzniuk and I. Kazymyra, “Development of software tools for defect/fault analysis in complex gates from standard cell library,” in Proc. International Symposium on Defect and Fault Tolerance in VLSI Systems, pp. 375-383, Oct. 2001.

[18] T. Ouchi, K. Makabe, M. Ogasawara and E. Murakami, “Modelling of Cu IMD-TDDB caused by extrinsic defects,” in Proc. International Reliability Physics Symposium (IRPS), pp 120-124, May. 2010.

[19] K. Huang, H.G. Stratigopoulos, S. Mir, C. Hora, X. Yizi and B. Kruseman. “Diagnosis of local spot defects in analog circuits,” in IEEE Tran. Instr. And Meas., Vol. 61, pp. 2701-2712, May. 2012.

[20] N.H.E Weste and D.M. Harris ,“CMOS VLSI design ,” 4th Edition, Pearson.

[21] G.A. Allen and A.J. Walton, "Critical area extraction for soft fault estimation," IEEE Transaction on Semiconductor Manufacturing, Vol 11., pp. 146-154 , Feb.1998.

[22] P.Gupta and E. Papadoupoulu, "Yield analysis and optimization," in *The Handbook of Algorithms for VLSI Physical Design Automation,* CRC Press, 2010.

[23] A.V. Ferris-Prabhu, "Modelling of critical area in yield forecast," IEEE J. Solid State Circuits, vol. SC-20, pp. 874-880, Aug. 1985.

[24] J.H.N. Mattick, R.W. Kelsall and R.E. Miles "Reliability prediction through critical area analysis," Integrated Reliability Workshop, pp. 167, Oct. 1995.

[25] L. Arnaud, P. Lamontagne, R. Galand, E. Petitprez, D. Ney and P. Waltz, "Electromigration induced void kinetics in Cu interconnects for advanced CMOS nodes," in Proc. International Reliability Physics Symposium (IRPS), pp. 3E.1.1 - 3E.1.10, Apr. 2011.

[26] D.A. Hodges, H.G. Jackson and R.A. Saleh, "Analysis and design of integrated circuits," Third Edition, Tata McGraw-Hill.