A Statistical and Circuit Based Technique for Counterfeit Detection in Existing ICs

Rashmi Moudgil

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Computer Engineering

Leyla Nazhandali, Chair

Michael S. Hsiao

Patrick Robert Schaumont

May 3, 2013
Blacksburg, Virginia

Keywords: Aging, Counterfeit, Process

Copyright 2013, Rashmi Moudgil

A Novel Statistical and Circuit-Based Technique for Counterfeit Detection in existing ICs

Rashmi Moudgil

(ABSTRACT)

Counterfeit Integrated Circuits (ICs) are previously used ICs that are resold as new. They have become a serious problem in modern electronic devices. They cause lower performance, reduced life span and even catastrophic failure of systems and platforms. To prevent counterfeiting and the associated revenue loss, there is need for non-invasive and inexpensive techniques to establish the authenticity of devices. We describe a technique to detect a counterfeit IC that does not have any special anti-counterfeiting mechanisms built-in prior to deployment. Our detection criterion is based on measuring path delays. The experiments show that a single path delay cannot directly reveal the age, as it is also greatly influenced by process variation and this could result in large error in classifying ICs as authentic or counterfeit. Instead, we establish that the relationship between the delays of two or more paths is a great indicator for the age of device. The idea is to project ICs from different age groups onto the space of the path delays and train a trusted reference hyper-surface for each age group. Ideally, the hyper-surfaces do not overlap. In this way, an IC under test can be assigned to one hyper-surface based on the distance of its footprint with respect to these hyper-surfaces, thus predicting its age. In our simulations, we observe over 97% correct prediction of identifying an aged IC from a new IC.

// Acknowledgements

Firstly, I extend sincere thanks to my research advisor, Dr. Leyla Nazhandali, for giving me an opportunity to work in the SePAC group, and keeping confidence in me throughout my engagement with the group. Her objective approach to work has served as a great motivation to accomplish higher goals, and her continued guidance and feedback has helped me improve on many fronts.

Thanks to Dr. Michael Hsiao and Dr. Patrick Schaumont for serving on my thesis committee.

I would like to thank the faculty members of the project – Dr. Michael Hsiao, Dr. Chao Wang, Dr. Simin Hall, Dr. Nazhandali, and the student members Avinash Desai, John Mulheren, Bing Liu, for providing an interesting platform for brainstorming and discussion of ideas while working as a team.

I also thank my lab-mates Meeta Srivastava, Mehrdad Khatir and Dinesh Ganta, for providing me an excellent environment within the lab, and for the great discussions on both technical and non-technical topics.

Finally, I express my sincere gratitude to my parents and family for keeping faith in me and in their willingness to let me come forward for this step in my career.

Contents

1. Introduction 1

1.1 Contributions 4

1.2 Organization of Thesis 4

2. Background and Motivation 6

2.1 Aging 6

2.1.1 BTI Effect 7

2.1.2 HCI Effect 9

2.1.3 TDDB Effect 9

2.1.4 More on variation in Path Delay with Aging 10

2.2 Process Variation 12

2.3 Intertwined Effect of PV and Aging 13

2.4 Related Work 16

2.5 Alternate Approaches 17

2.5.1 Using Burn-In testing 17

2.5.2 Use of Characterized trim bits as Process Identifiers 19

3. Proposed Anti-Counterfeit Method 21

3.1 Base Method 21

3.2 Enhancement - Just-In-Time Voltage Reduction 25

4. Simulation Framework 28

4.1 Predictive Technology Model (PTM) 28

4.2 MOS Reliability Analysis (MOSRA) 30

4.3 Process Variation Modeling 31

4.4 Test Circuits 35

4.5 Statistical Data Analysis 36

5. Results, Analysis and Improvements 37

5.1 Basic 2-Path Method 38

5.2 Comparison across different levels of Process Variation 41

5.3 Detection within lower range of Age 42

5.4 Sensitivity Analysis and Improvements 43

5.4.1 Sensitivity Analysis 43

5.4.2 Further Improvement: 3-path Approach 45

5.5 Benchmark circuit results and extension to real ICs 49

6. Conclusion 51

7. Bibliography 53

List of Figures

Figure 1-1 ICs scavenged and sorted 2

Figure 1-2 Anti-Counterfeit techniques 3

Figure 2-1 Transistor Degradation with time 7

Figure 2-2 Degradation in V_{th} and I_{on} due to NBTI effect in PMOS 8

Figure 2-3 Carrier generation under the effect of electric field 9

Figure 2-4 Delay change with aging for different gates 10

Figure 2-5 Impact of workload (switching activity) on aging rate 11

Figure 2-6 Summarizing the effect of process variations on circuit characteristics 12

Figure 2-7 Spatial variation of Process 13

Figure 2-8 Impact of PV and Aging on an IC 14

Figure 2-9 Delay Distribution of a path in a new vs. old IC 15

Figure 2-10 Bath-tub curve of Failure rate versus time 18

Figure 3-1 Delay based method for Counterfeit Detection in existing chips 22

Figure 3-2 Basic 2-path Method 23

Figure 3-3 Variation in delay in presence of process variation and supply voltage change 26

Figure 3-4 Delay vs. Age for diff supply voltage used during post-stress phase 27

Figure 4-1 The MOSRA flow in HSPICE 30

Figure 4-2 Modeling of typical delay with load variation 35

Figure 4-3 Circuit pruning to extract only paths of interest from the original circuit 36

Figure 5-1 Relationship curves between 2 delays, across different age groups, at different post-stress voltages of 0.5V (top) and 1.2V (bottom) 39

Figure 5-2 Sensitivity of 2-D prediction method (new/old) 44

Figure 5-3 Sensitivity of 2-D prediction method (Exact Age) 45

Figure 5-4 3-path method – hyper surface training and mapping 46

Figure 5-5 Sensitivity with 3-path prediction method (new/old) 47

Figure 5-6 Sensitivity with 3-path prediction method (Exact Age) 48

Figure 5-7 Sensitivity analysis in 2-D and 3-D for PV2 category 48

List of Tables

Table 4-1 Different Process Variation Classes used for Simulation................................. 34

Table 5-1 Prediction results for Basic 2-Delay method.. 38

Table 5-2 Comparison of prediction for various age groups across different voltages 40

Table 5-3 2-path method results for different levels of process variation 41

Table 5-4 2-path method within lower age groups (0, 3,6,12 months)............................. 42

Table 5-5 Results for standard circuits 49

1. Introduction

Use of electronic devices has become an essential part of our day to day life. They are found in applications from tiny embedded devices to advanced computers. With the pervasive nature of such applications, the concept of using off-the-shelf discrete Integrated Circuits (IC) has become a commonplace. Use of discrete IC components minimizes design time and is cost effective due to the large scale manufacturing of the components. Further, debug and maintenance of platforms become easy as faulty components can be replaced with new parts. As a result, most modern systems are built out of discrete Integrated Circuits (ICs).

Globalization has provided us with a vast choice of hardware suppliers at various levels of the design and manufacturing flow. Designs and fabricated chips can now come from practically anywhere in the world. However, this has not come without costs. First and foremost, the trustworthiness of the received device is no longer a guarantee. The reliability of a system is directly dependent on the reliability of the components that it is built of. One unreliable IC in a system possibly can result in a catastrophic system failure. The area where this is particularly critical is in defense and medical related systems where the reliability of a system is of utmost importance.

Figure 1-1 ICs scavenged and sorted

One of the main disadvantages with the proliferation of electronic devices is the associated electronic waste. A major source of this waste comes from the discarded boards and systems. Recently, many cases have come to light, which show ICs being scavenged from electronic waste, repackaged and sold in markets as genuine new ICs [1]. The article [2] states that the used or defective products being sold as new or working account for 80-90% of all counterfeits being sold worldwide. A report on the counterfeit ICs by IHS iSuppli estimates a $169 billion in potential annual risk to the global electronics business [2, 3, 4, 5]. These studies suggest that IC counterfeiting is a serious and growing problem. Some of the direct impacts of IC counterfeiting are enumerated here:

1. Consumers do not get what they pay for. The devices get slower and are more prone to failure. Both security and reliability are highly compromised.
2. Semiconductor manufacturers incur a significant loss due to lost sale, which eventually hurts everyone related to this industry including consumers.
3. System manufacturers suffer bad reputation when the components fail. In addition, they endure loss for providing warranty.
4. In case of critical applications such as avionics, defense systems, and medical

devices, untimely IC failures can result in catastrophic events.

As a result, it is extremely important to be able to distinguish new authentic ICs from counterfeits and/or used parts. For the future designs we might have the luxury of inserting specialized circuitry to the chip in order to help establish the age or usage of the chip (e.g., ROM-based fuses can be implemented to act as IC seals or special protected registers can be employed to produce the manufacturing date of the chip). However, determining whether an existing chip is a counterfeit is much harder. For such cases, statistical and testing techniques are needed.

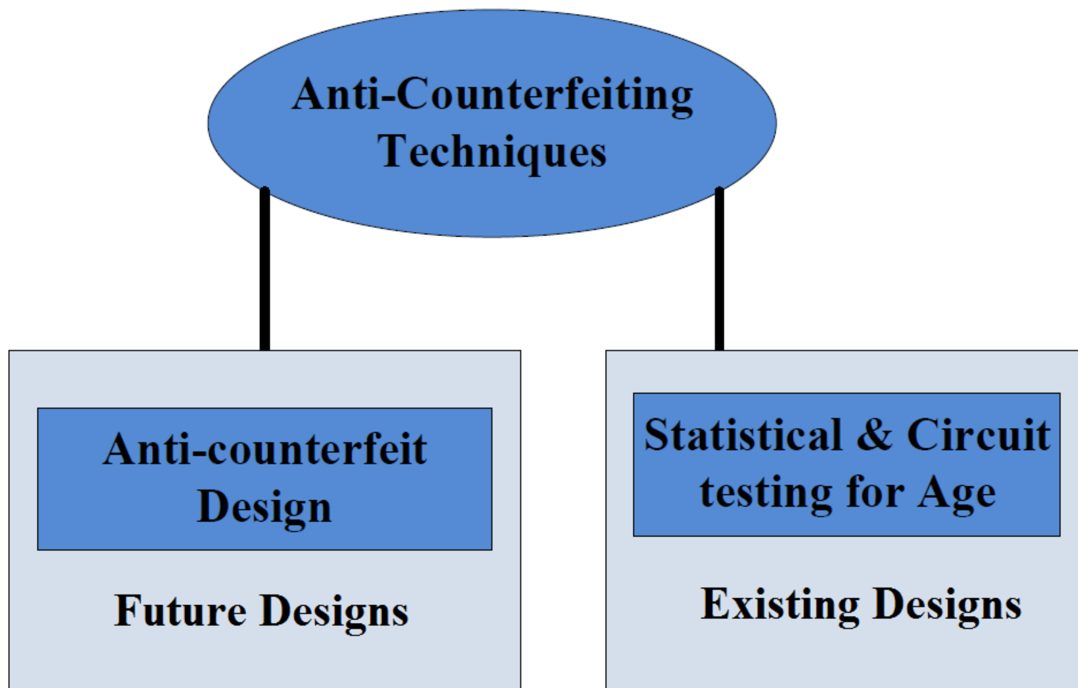

Figure 1-2 Anti-Counterfeit techniques

Figure 1-2 illustrates these two categories of counterfeit detection methods. Although finding anti-counterfeit methods for future designs is not a trivial problem, it poses a less challenge compared to detecting counterfeits amongst existing chips, where one has to depend on just the original circuit itself to establish the age of the device. This thesis focuses on age-prediction for existing ICs.

1.1 Contributions

The major contributions of this thesis are as follows:

1. Based on detailed simulation results, we show that although the delay of a circuit is greatly affected by aging, it cannot be used as an age indicator by itself since delay is also significantly affected by process variation.
2. We show that the relationship between two or more paths in the circuit can be used as an age indicator and we propose a statistical method based on this fact, which can correctly predict the age of an IC.
3. We provide an extensive analysis of our proposed method and show its accuracy by implementing on variety of sample circuits, including few cores and ISCAS benchmarks.

1.2 Organization of Thesis

The rest of this thesis is organized as follows:

Chapter 2 presents the background. We introduce the concept of Aging in ICs. An overview of the underlying physical mechanisms and their manifestations on circuit characteristics is provided. The discussion of the interference from process variation gives insight into the underlying challenges to counterfeit detection and therefore about the motivation for the thesis work.

Chapter 3 introduces the delay based anti-counterfeiting techniques proposed in this work. Combined with Statistical analysis, it is able to successfully detect counterfeits. Enhancement to this method by using “Just-in-Time” voltage reduction is shown to improve the performance of the base method.

Chapter 4 highlights the simulation framework which is used for the proof of concept of the proposed method. Details of the modeling of aging and process variation are provided.

Chapter 5 presents the simulation results of the technique. The detailed analysis of the sensitivity of the detection method to the nature of paths selected is provided. Improvements to the method for reducing this sensitivity are also proposed.

Chapter 6 concludes the thesis work.

2. Background and Motivation

This chapter provides an overview of the physical mechanisms behind transistor aging, and process variation. Their collective effect on circuit behavior establishes the challenge of the problem statement and forms the basis of the motivation behind the proposed technique. Then we discuss the related work in this field.

2.1 Aging

In simple terms, aging can be defined as slow but eventually permanent variations that generally deteriorate circuit performance over time. In a matter of few months, a transistor undergoes changes both within the gate di-electric and at the boundary of silicon and gate oxide. Examples of these changes include broken Si-H and Si-O bonds, and injection of charge carriers into gate dielectric from the drain end of the channel. These result in continuous degradation of transistor characteristics, as pictorially depicted in Figure 2-1. The process of aging is investigated widely in literature, and is observed to cause slower operation of circuits, irregular-timing characteristics, increase in power consumption and sometimes even functional failures [6, 7, 8].

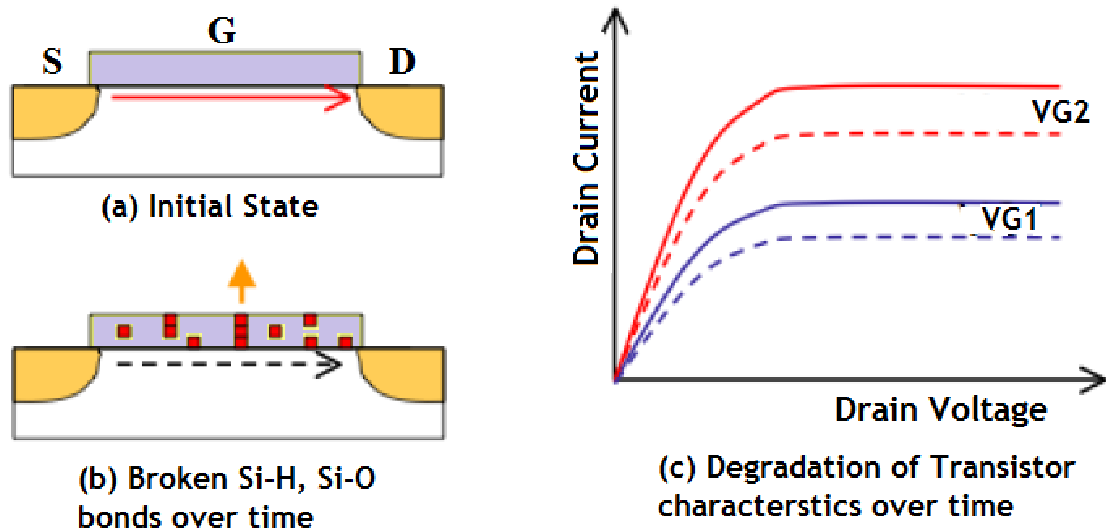

Figure 2-1 Transistor Degradation with time

Although there is no single physical mechanism that is comprehensive enough to explain all the behaviors, the major physical mechanisms behind aging in ICs are listed below:

1. Bias Temperature Instability (BTI)
2. Hot Carrier Injection (HCI)
3. Time Dependent Dielectric Breakdown (TDDB)

This chapter covers only a brief overview of these physical effects. Detailed studies on the factors that contribute to aging of transistors including their modeling are presented in [6, 9, 10, 11, 12] .

2.1.1 BTI Effect

Negative Bias Temperature Instability (NBTI) is prominent in PMOS transistors, and can shift (degrade) the PMOS threshold voltage by more than 50mV over ten years. This translates to more than 20% degradation in circuit speed [13, 14]. Similarly, Positive Bias Temperature Instability (PBTI) effect occurs in NMOS. In both BTI cases, the amount of

charges in the gate dielectric changes with the gate bias, because of charge trapping and de-trapping. Under the effect of a constant bias, the amount of trapped charges increases continuously, causing increase in the threshold voltage, $V_{TH,}$ and decreasing channel carrier mobility. This effect is also strongly proportional to the device operating temperature.

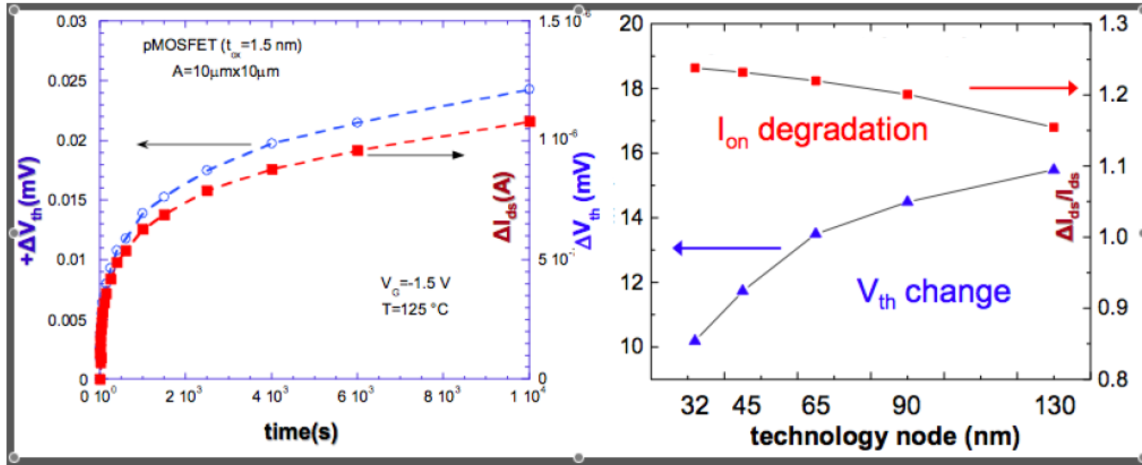

Figure 2-2 Degradation in V_{th} and I_{on} due to NBTI effect in PMOS

It has also been established that there is a partial recovery effect of the degradation caused to the circuit. And this is dependent on the duty cycle of the stimulus to the circuit. Basically, under changing gate voltage, the trapped charges can get de-trapped, and this causes some recovery from the degradation. When the 'recovery' phase is not accounted in the model, it is called 'static' NBTI effect, but if both the 'stress' and 'recovery' phases are accounted in the analysis, it is called as 'dynamic' NBTI effect. Figure 2-2 shows the degradation due to static NBTI effects on the threshold voltage and saturation current of a PMOS over time on the left side. The graph on the right hand side is from study in [11], which compares the impact of dynamic NBTI effects on the circuit parameters, across the technology nodes. A fixed aging period of one year is used for this analysis.

2.1.2 HCI Effect

In the presence of high electric fields, carriers are injected from the drain end of the channel into the gate dielectric, changing its electrical properties over time. The longer the transistor is in operation, the higher is the number of new carriers that are generated. This is why HCI effect is more dependent on the dynamic switching activity of the transistors, circuit structure, fan-out and input waveform. HCI phenomenon causes increase in threshold voltage for the NMOS. As a consequence, they produce less current which translates into slower switching speeds. The effect on PMOS is slightly different - it is evaluated in terms of the gate current. PMOS current increases a result of HCI effect, due to V_{th} shift.

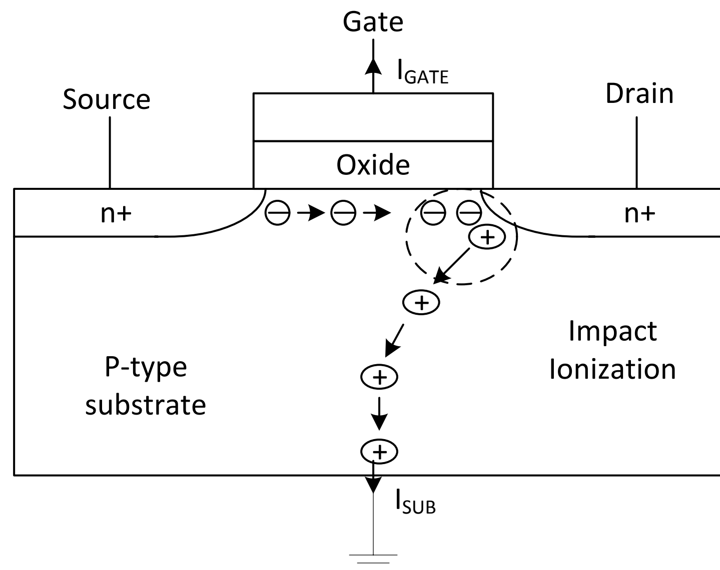

Figure 2-3 Carrier generation under the effect of electric field

2.1.3 TDDB Effect

TDDB is also a consequence of aging which results in shorting of gate dielectric leading to gate failure. This physical mechanism occurs as a result of long-time application of low electric field, and causes the gate oxide to break down over time.

All these effects are particularly prominent at technologies with smaller feature sizes, due to increasing electric fields in the devices, and decreasing thickness of oxide. Recently, there has been more focus on developing countermeasures for NBTI due to the fact that unlike HCI that occurs only during dynamic switching, NBTI is caused during static stress on the oxide even without current flow. This situation has become more challenging as many current and future digital systems tend to support longer stand by modes in their functionality [11].

2.1.4 More on variation in Path Delay with Aging

Few important observations in regard to aging of transistors, specifically to its manifestation in path delay are discussed in this sub-section.

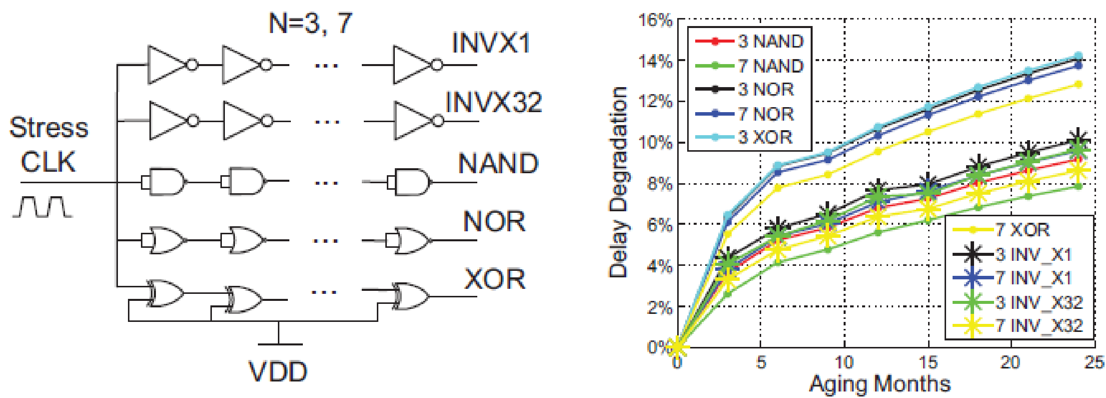

Figure 2-4 Delay change with aging for different gates

1. Different types of gates undergo different amount of aging. As a consequence, different paths in the digital circuit may age differently by virtue of their construction itself. For example, complex gates like XOR gates are found to age faster as compared to a basic inverter gate. Figure 2-4 is pulled from a study done in [15], and it shows that chains constituted from different gate types (45nm), undergo different amount of degradation in delay over time. It can also be

observed that the smaller gates (INV_X1) age faster than the larger gates (INV_X32) for same amount of aging stress.

2. The amount of activity of the transistors has a direct effect on the aging process. Both the value at the input and the switching frequency dictates the rate of aging of the path. A high activity path is expected to age faster, and hence undergoes more degradation in delay.

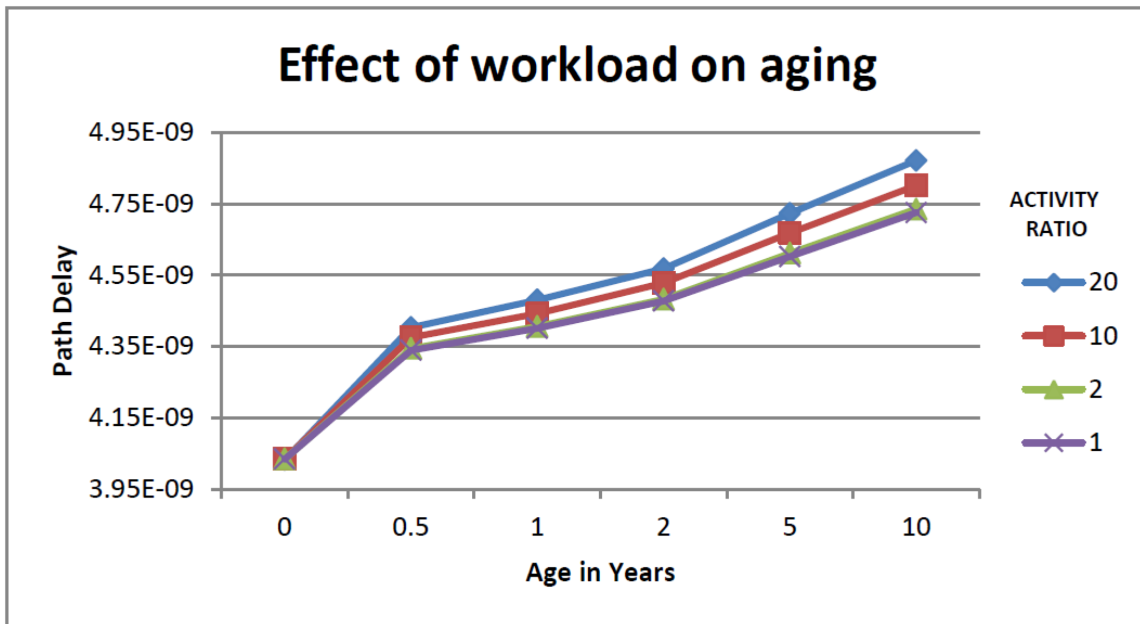

Figure 2-5 Impact of workload (switching activity) on aging rate

Plots in Figure 2-5 are obtained from simulation of delay paths (90nm), whose input stimulus is controlled, to arrive at different switching activity on each path. In this example, we simulate four identical paths with varying level of activity. The activity ratio of all paths is fixed w.r.t to one of the paths – for e.g. '20x' activity ratio implies that the path undergoes 20 times more number of transitions than the reference path, hence 20 times more active. As we can observe, at an activity ratio of 2x, difference in the aging rate is not very distinct, but at 10x and

20x, the difference is obvious i.e. the higher the transition activity on the path, more is the degradation in delay it sees over time.

2.2 Process Variation

Process variation (PV) is the random and permanent deviation from the designed, nominal value of a circuit structure, caused by random effects during manufacturing [16]. PV can be separated into two categories: The first category covers variations in process parameters, such as impurity concentration densities, oxide thicknesses, and diffusion depths. These result from non-uniform conditions during the deposition and/or the diffusion of the dopants. The second category covers variations in the dimensions of the devices. These result from limited resolution of the photolithographic process, which in turn causes width and length variations in transistors.

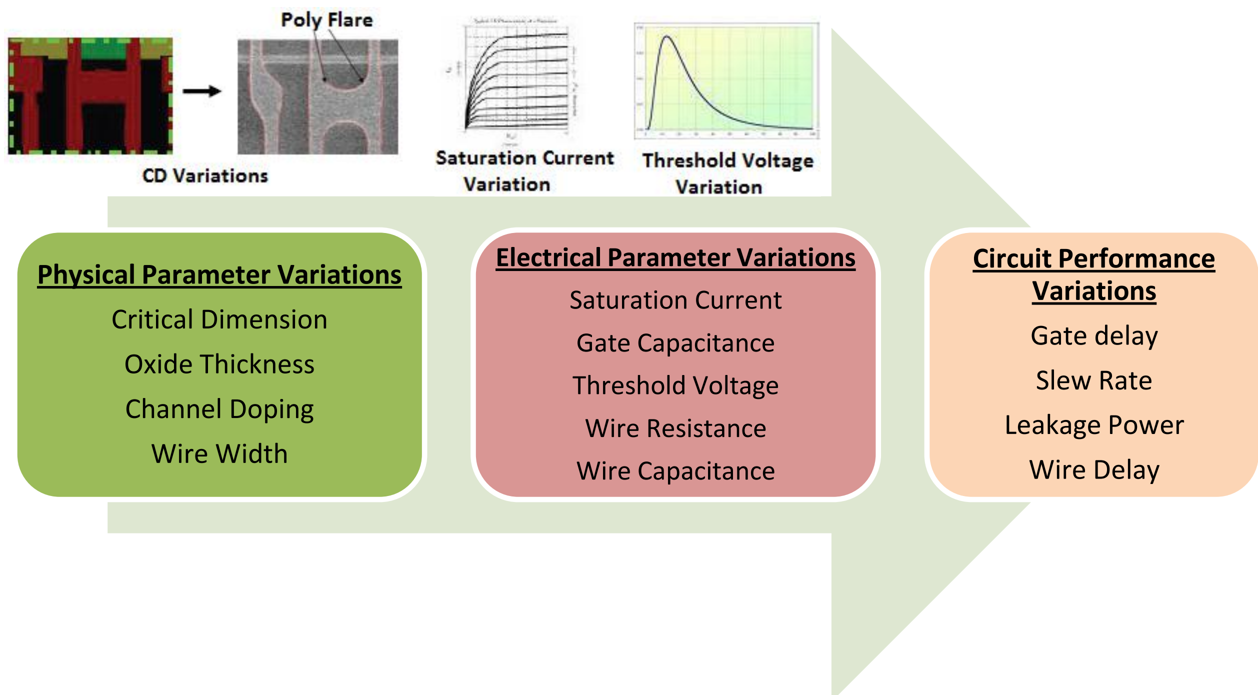

Figure 2-6 Summarizing the effect of process variations on circuit characteristics

Figure 2-6 provides a vivid depiction of how the process variation during manufacturing process introduces variability in the electrical parameters and eventually in the circuit performance of the end device.

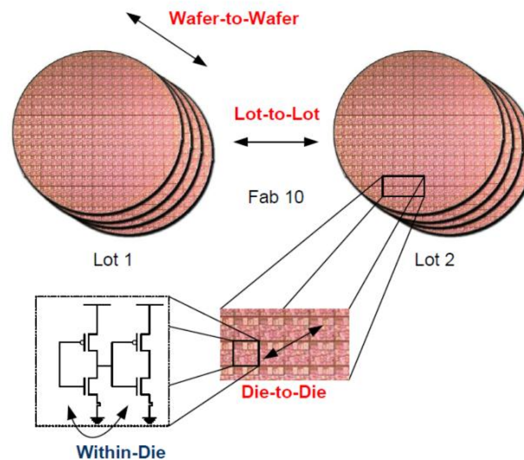

Figure 2-7 Spatial variation of Process

It is important to note, that variation in manufacturing, can introduce mismatch between two transistors sitting within a single die, and on a higher level of abstraction, from one wafer to other, and between manufacturing lots. This spatial classification of process variation (Figure 2-7) is widely grouped as intra-die (within the die), and inter-die (across dies). In general, inter-die variation has more pronounced variability in the design as compared to intra-die variation.

2.3 Intertwined Effect of PV and Aging

So far, we have looked at the individual effects of process variation and aging. An integrated chip that has been used in field is bound to exhibit shift in parameters by virtue of the aging effects. Since the used ICs are impacted by the discussed aging mechanisms, we expect the path delay of a used IC to be different from the path delay of a new IC. If

we could rule out the presence of any environmental factors, and process variation, we could easily use delay of a single path from the chip, and use its degradation to quantify the age of the chip, but as we saw, the process variation has a significant impact, which cannot be left out from the analysis.

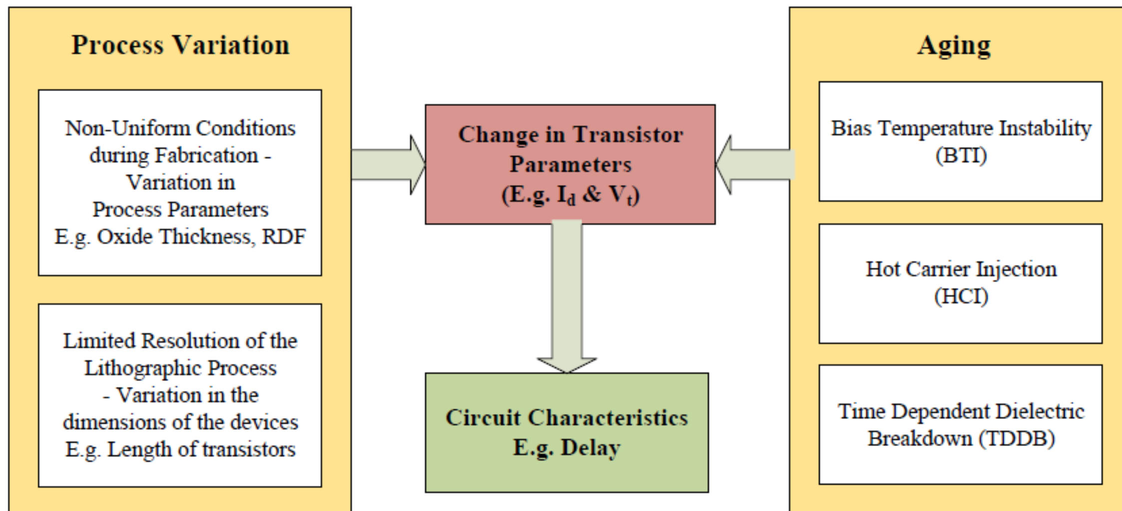

Figure 2-8 Impact of PV and Aging on an IC

Bringing the two of them together, Figure 2-8 summarizes how both PV and aging impact common electrical parameters such as threshold voltage, saturation current and eventually circuit characteristics like delay. This intertwined effect of PV and aging is very hard to isolate from each other. It makes the counterfeit detection a very challenging problem.

Figure 2-9 shows this intertwined effect from simulation point of view. A simple 50-stage inverter chain is simulated both as a new circuit (light-blue histogram) and as a 5-year old circuit (dark-blue histogram). In both cases, the circuit is also subjected to process variation using Monte Carlo simulation. More details of our simulation setup are provided in Chapter 4. Although the aged circuits are typically slower than new circuits, there is a significant overlap in the range of delay values for the two groups. In fact, more than 70% of both the new and old circuits have delay between 0.565ns to 0.615ns. Suppose we pick a delay value of 0.6ns, it would be impossible to tell if it corresponds to

an aged path – which was originally faster and now has increased delay due to aging or to an un-aged one, which is just slower due to process variation.

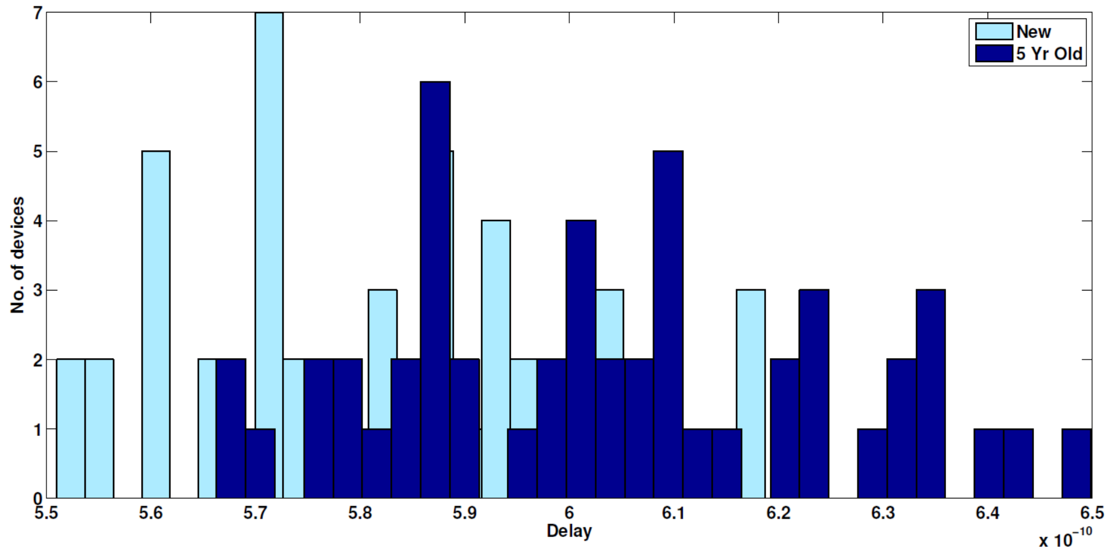

Figure 2-9 Delay Distribution of a path in a new vs. old IC

The above observation demonstrates that delay of a single path from the chip is seemingly not effective as a direct indicator of the age of an IC. The path delay degradations caused by aging effects in used ICs must be separated from those caused by process variations in new ICs. To make matters more complicated, there is recent submicron data showing that aging rate itself can be affected by process variation [17]. Specifically, there is significant dependence of NBTI and PBTI on device geometry, e.g., channel length, which is variable as virtue of process also. These effects have begun to be modeled by commercial tools [17], but reverse engineering the cumulative effect to isolate the factor of age becomes a tedious problem. Therefore, for an unknown device, there are two unknown variables that we need to deal with –

- Age
- Process

In this thesis, we propose a method that overcomes this challenge by combining the characteristics of more than one path and isolating the effect of aging.

2.4 Related Work

In regard to counterfeit prevention and detection, traditional techniques such as printing serial numbers or barcodes to prevent counterfeiting in integrated circuits are not effective as they can be easily cloned or faked. IC Metering [18, 19, 20] provides techniques that allow post fabrication control on the ICs. Although metering techniques can be leveraged to prevent counterfeiting, most such techniques require modifying the design significantly [9] and are not applicable to existing designs. In [21], the authors present a technique to prevent piracy of ICs, which renders the ICs inoperative upon fabrication, and it requires a device unique key to take the IC to the functional mode. Recently, some anti-counterfeiting methodologies were presented which use Physical Unclonable Functions (PUFs) [22]. An odometer based technique to estimate age of ICs has been proposed in [9]. Although, most of these methods are designed to monitor reliability of ICs, they can be deployed for anti-counterfeiting purposes as they provide information on the age of the ICs. A method in [23] presents identification of recovered ICs by the use of an on-chip light weight sensor. An important limitation with all techniques mentioned above is that they are only applicable to designs for the future. Many applications including a wide range of critical applications heavily depend on designs that were created years ago due to the fact that the test cost and time can be prohibitive of they keep deploying newer designs. As a result, it is necessary to distinguish between unused and used (counterfeit) for ICs already manufactured with no anti-counterfeiting capabilities. There are few companies, which deploy inspection techniques using X-ray and Infrared microscopy for this category of ICs [24]. Special high magnification machinery is needed for it and with more refined ways of re-packaging and cleaning; the image based techniques may be ineffective. Instead, our work aims to investigate cost-effective anti-counterfeiting techniques for this purpose. The most recent related work is presented in [15], where path delay finger prints from ICs

are used, and combined with Principal Component Analysis techniques to authenticate an IC. The volume of the data that needs to be measured for analysis is very high. In our work, we show that with selection of appropriate paths, we can deduce the aging information by looking at a set of just 2 or 3 paths. Our method is successful in detection of an old counterfeit from a new authentic IC in presence of high levels of process variation. Moreover, we can also predict the age of an IC with a high success rate.

2.5 Alternate Approaches

This sub-section is dedicated to present the alternate ideas which can be explored for countering the problem of counterfeit detection that we intend to solve.

2.5.1 Using Burn-In testing

For most of the products the reliability curve is in shape of a bath-tub, hence the name "bathtub curve" (Figure 2-10). It is a plot of the failure rate against time. The initial region is called the 'infant mortality' or 'early failure' region. It is characterized by a high failure rate, which rapidly decreases over time. System with weak components, tend to fail during this phase (i.e. the start of the life-cycle of product).

After this phase, the failure rate tends to settle off to a fairly constant value for majority of what is called the 'Useful Operating Life' of the product. Most of the systems spend majority of their lifetimes operating in this flat region marked by low failure rate. Finally, if product is used long enough, the failure rate begins to rise owing to the wear out of the device. This marks the end of life of the product.

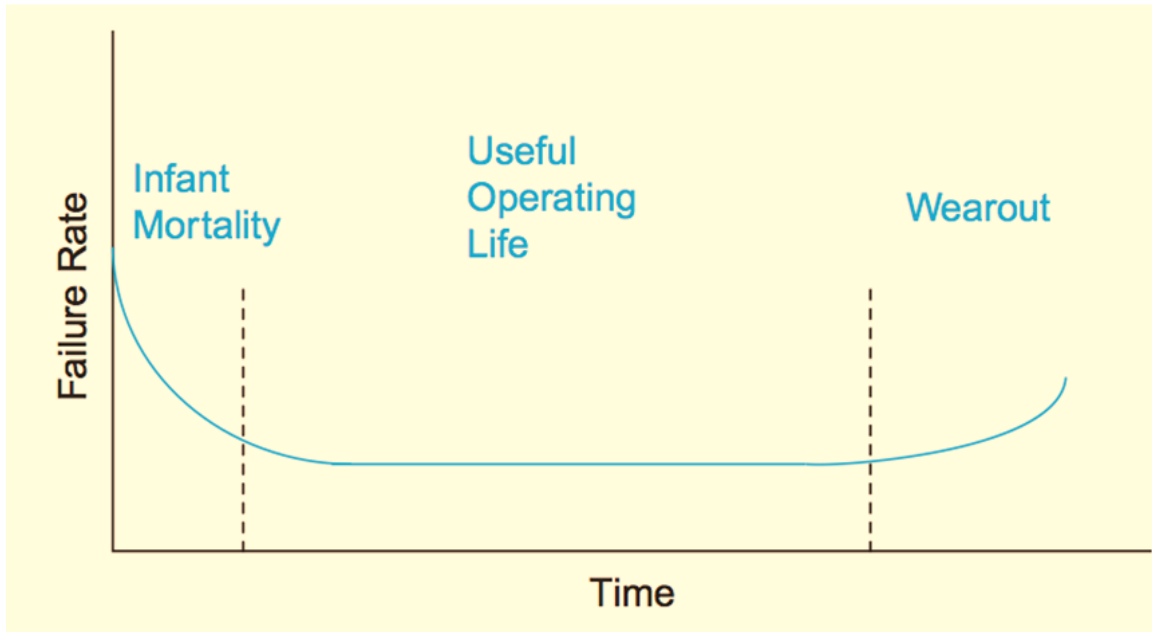

Figure 2-10 Bath-tub curve of Failure rate versus time

Production Burn-in Tests: Typically, some form of stress testing is necessarily deployed before shipping the product to the customer, as it is important to age the system beyond the 'early failure' region. The ICs are subjected to extreme voltage and temperature conditions for long periods of time, to emulate accelerated aging. These are also called 'burn-in' conditions. The application of burn-in during production is intended to screen out early failures, and it comes at the expense of a reduced yield caused by the burn-in process. So, in a typical high volume production, these tests are continued to be used only until various root causes for early failures are identified and eliminated.

As the wearout mechanisms are known to depend exponentially on voltage and temperature, using burn-in tests the time to wearout is calculated. These results can also be extrapolated to normal operating conditions, to predict the duration of useful life. Such estimated circuit failure prediction approach like burn-in testing, can be used for the purpose of detecting counterfeits as well. We can re-perform burn-in tests on the suspected ICs under test. The results from the tests can be used to estimate the remaining

useful life of the IC, and using the acceptable thresholds, we can screen out the older ICs that are not fit enough to be part of the system.

The downside of this idea is the test cost and test time associated with this form of testing. However, if Burn-in tests were part of the original production test suite of the IC, the required infrastructure (special burn-in boards) can be leveraged to mitigate the test cost. Another important consideration is the possible damage it can cause to the IC. The re-use of burn-in tests, on a device under test, later in the product cycle, will definitely run risks of destroying the device permanently.

2.5.2 Use of Characterized trim bits as Process Identifiers

As we saw, the problem with use of path delay directly in order to find the age of the IC is the unknown variable of process. We could wrongly classify the device under test as old based on larger delays, although in reality it may have just belonged to a slower spectrum of the new manufactured ICs. The idea in this section explores the possibility of extracting the process information from the IC itself in order to solve this part of the problem.

Typically, most chips have analog components like RC Oscillators, Voltage regulators/controllers. Variations in manufacturing process result in deviation of the analog circuits from their specification. After the chip is manufactured, all these components are characterized across different lots (process corners). To optimize the performance of the systems, it is important to "trim" the interface circuitry to match specific analog circuit. Some of this information is stored or fused-in, in form of trim bits. And these trim bits are based on characterization results, and expected to have strong correlation with the process corner of the IC. If we are able to read out these trim bits, and extrapolate it to find the process variable (or make use of look up tables to map it back to the process variable), half of the problem is solved. Once the process corner of the unknown IC gets identified, we can measure path delays from this unknown chip

under test, and the delays can be directly compared to ones measured from an authentic reference chip belonging to the same process corner. The shift in the delays can then be directly used to find the amount of aging it has undergone.

This concludes this sub-section. The next chapters pertain to our proposed method of detection, which can be deployed in already existing integrated circuits, and the simulation results are shared in detail to provide proof of concept. Also discussed is the feasibility of the method on real ICs.

3. Proposed Anti-Counterfeit Method

This chapter discusses the method we propose for detecting old counterfeit ICs, which lack special anti-counterfeiting. We begin with the goal of just being able to distinguish an old device from a new one and later extend the methodology to accurately predict the age of a chip.

3.1 Base Method

As we saw in Chapter 2, although delay of a single path is significantly affected by aging, it is a poor indicator of the age of the chip as it is also greatly affected by process variation. Therefore, instead of using a single path, we propose the use of a set of paths from within a chip, such that each path individually has a characteristic aging behavior. *The relationship between the delays of the chosen paths is used as an indicator of the age of the chip.* The premise of our method is that the pattern that governs the relationship between two delay paths for new ICs are distinguishably different from the pattern governing that of old ICs. This means, if we have a trusted set of new chips and we are able to study them before and after (accelerated) aging, we can characterize these relationship patterns at different ages. Now, when an untrusted chip is under test, the relationship of the same paths is evaluated and it is determined whether it follows the pattern belonging to the new chips or older chips. The flow of steps is illustrated in Figure 3-1.

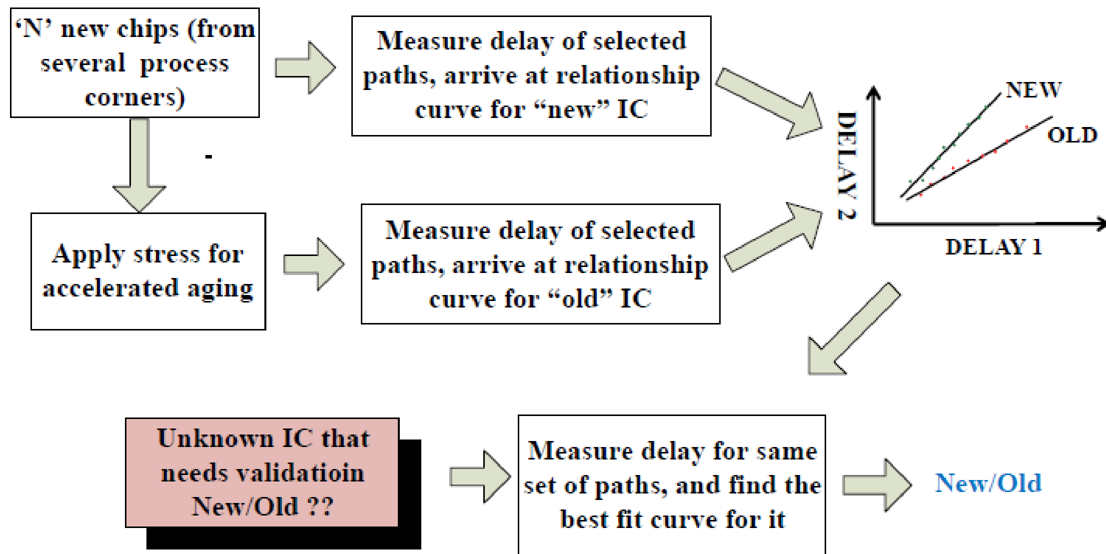

Figure 3-1 Delay based method for Counterfeit Detection in existing chips

Our base method when used with only two path candidates is called 2-path method. The relationship between the two delay variables is used for determining the age. The details of our technique are as follows:

1. Two different paths are selected from the chip. The difference in the aging rate of the two candidates is pivotal for the success of our method. While selecting the best candidates, we exploit the fact the amount of *activity (workload)* of paths significantly impacts its aging rate. The right selection of candidate paths is discussed in more detail in coming sections.

2. The delays of the two paths are measured for a variety of trusted new chips. Ideally the trusted chip pool has to be large enough to present a realistic mix of chips affected by process variation. Curve fitting is used to identify a pattern relating the two paths to each other, similar to the "AGE0" line shown in Figure 3-2.

3. The trusted chips undergo accelerated aging by subjecting the chips to more than nominal operating voltages and high temperatures.

4. The delays of the two paths are again measured for all the trusted chips. Once more curve fitting is used for relating the paths to each other, similar to the "AGE5" line in Figure 3-2. These reference linear-fit curves that are obtained can also be called the 'characterized' database.

5. When an untrusted chip is under test, the delays of the same two paths are measured. Based on mathematical methods (e.g. shortest perpendicular distance, or least residual), it is determined whether the untrusted chip better fits to the known curve of "new ICs" or the "older ICs". In other words, the position of the unknown chip is determined with respect to the pre-existing relationship curves, to identify which curve it most probably belongs to.

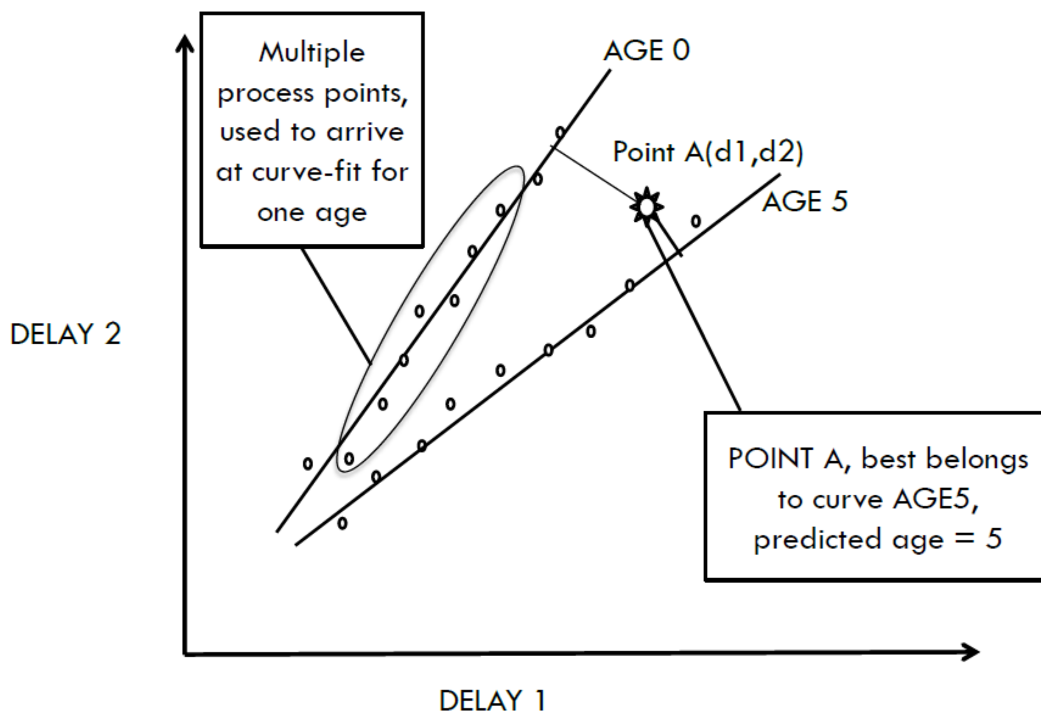

Figure 3-2 Basic 2-path Method

This can be extended to a 3-path or 4-path method, which uses more than two paths, and the relationship between their delays is fitted to a surface plot instead.

In order for this method to work, trusted new chips should be obtainable, from the design house. These new chips will be used to generate the reference curves at Age0. Then these chips will be exposed to stress conditions to emulate aging, and arrive at these relationship curves that shall serve as the reference database. Path delay information from the trusted and untrusted ICs is obtained by performing test procedures on the ICs. A process similar to speed binning using ATPG test patterns can be deployed. Already existing scan patterns from the test suite can be used to determine the path delay of the selected candidates. The delay test patterns are applied at different clock frequencies, and the maximum frequency at which a particular delay pattern passes is noted. The idea is, as long as the time period of the clock cycle used in the pattern is larger than the path delay, it will pass. Hence, the max passing frequency of clock gives the delay of the path. This technique is widely used in industry for purpose of speed binning.

As already stated, the important aspect is to decide on the kind of paths to choose. Favorable results are obtained if the paths chosen resemble each other in nature of their delay, e.g., paths with fairly similar capacitive loading, and similar length, but have a *distinct aging profile*. The latter is achieved by using paths of dissimilar activity – i.e. paths undergoing different number of transitions. The path with higher activity is more affected by age, which means the relationship of the two paths changes significantly over time. As shown pictorially in Figure 3-2, the AGEX relationship curve drifts to the right of the AGE0 curve, as one delay variable is increasing more than the second delay variable, due to difference in their aging rate. Practically, to find such path candidates in on a real chip is not unrealistic, as some portion of the chip logic is not as active during its lifetime, as other portions. For example, the test paths around the memory test logic will be quite less active or totally inactive, during the lifetime, relative to a functional clock tree path or highly active functional data path. Such paths, which significantly differ in their activity, are good candidates for this scheme. We provide a detailed

sensitivity analysis in Chapter 5, in which we show what two paths can be the best candidates for this method.

If we limit the problem to just distinguishing between a new and an old IC, the above conditions can achieve only up to 87% accuracy for detecting counterfeit ICs, but almost 1 out of every 3 new chips get falsely detected as old (Chapter 5). And if we want to use the technique to determine exact age of the IC, the correct prediction rate is not much over 50%. In other words, the pattern that governs the relationship between two delay paths for new devices are fairly different from the pattern governing that of old devices, but they are not sufficiently different to uniquely identify the exact age of the device. So, the basic technique needs to be strengthened to improve the detection rate, and also extend it to accurately predict the age of the unknown device. Some more control parameters need to be identified, so that the changes in delay are made more distinguishable for different age groups. One such promising parameter is variation in supply voltage during the measurements for the chip under test. We use it to enhance our detection method, as described in the next section.

3.2 Enhancement - Just-In-Time Voltage Reduction

In order to improve our method, we propose an augmentation technique that we call *Just-in-Time Voltage Reduction*. In this approach, we reduce the operating voltage of the chip at the time of testing. The rationale is that the effect of aging can be more easily detected at lower voltages. This is explained in more detail in the rest of this section.

The first order equations for gate delay (inverter with symmetrical nmos and pmos), shown in equations (1) and (2), provide the basis for modeling delay in both strong-inversion and sub-threshold regions [25]. The dependence of the delay on the difference of VDD and V_T changes from a linear relation to exponential one as we move into sub-threshold region.

$$t_d = \frac{K C_G V_{DD}}{(V_{DD} - V_T)^{\alpha}} \quad (1)$$

$$t_{d,sub} = \frac{K C_G V_{DD}}{I_o \exp(\frac{V_{DD} - V_T}{n V_{th}})} \quad (2)$$

On the other hand, the equations of aging models [12] show how HCI and BTI result in increase to the threshold voltage of the transistor over time. So, it is expected that operating the circuit at lower voltages, especially near or lower than threshold voltage will result in more drop in circuit speed after aging, compared to a circuit that is run at nominal voltage. It is also known that the effect of process variation is more prominent when operating the circuit at a reduced voltage. However, as we will show, the effect of aging is more magnified than the effect of process variation at reduced voltages. This means that eventually it is more beneficial to our method to operate the test chips in reduced voltage.

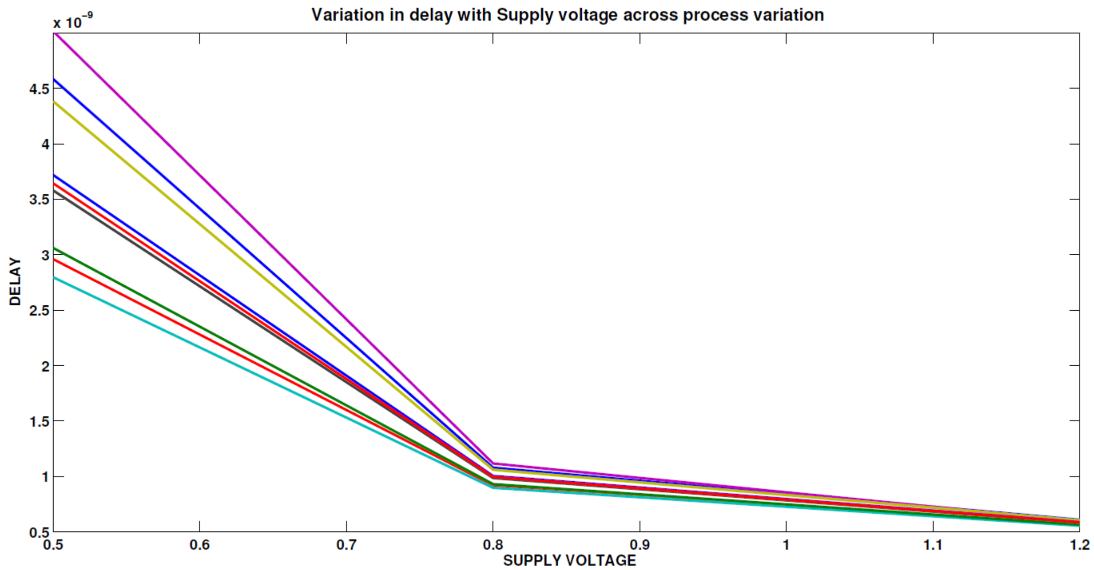

Figure 3-3 Variation in delay in presence of process variation and supply voltage change

The following simulation results illustrate the effect of voltage reduction on process variation and aging effects. Plot in Figure 3-3 shows the effect of supply voltage on the delay, across 10 different process points. It can be noticed that in the lower voltage range,

the distance between the respective delay lines is evidently more than in the higher voltage region. In other word, the spread of delay numbers across process is larger at lower voltage than at higher voltages.

For the variable of aging, only one-process point is observed: Figure 3-4 shows the plot of the delay of a path, as it is aged from zero to five years. While all circuits are operated at nominal voltage during their lifetime, each is subjected to a different voltage, i.e., 0.5, 0.8 and 1.2V at the time of delay measurement. As expected, the delays are higher at lower voltages. In addition, we observe that at 0.5V, the effect of aging is more prominent, as seen by the clear increase in delay after aging. We intend to use this observation to augment our proposed method in order to more accurately detect counterfeit ICs.

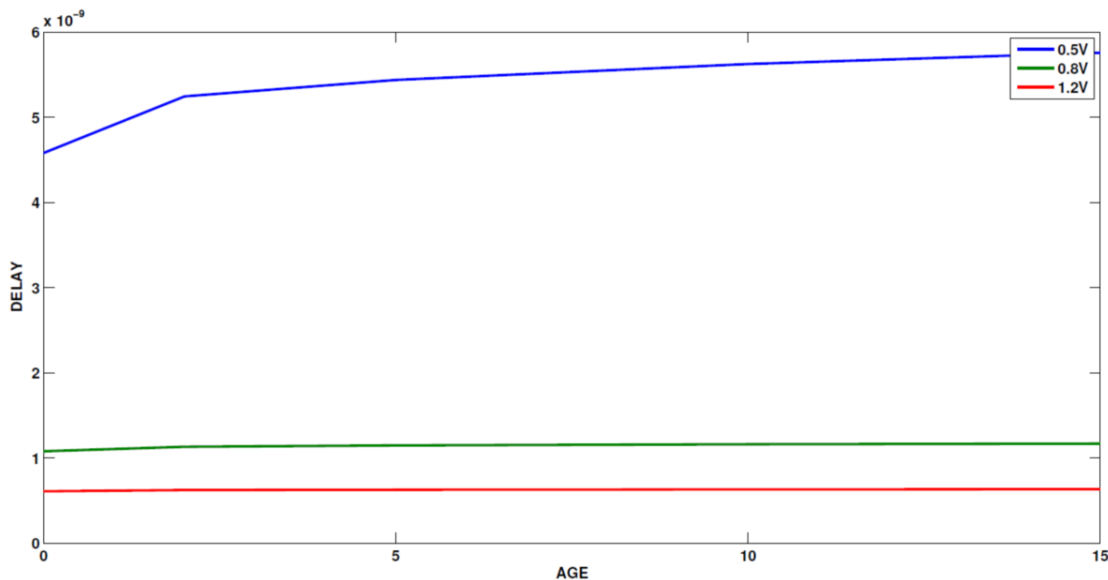

Figure 3-4 Delay vs. Age for diff supply voltage used during post-stress phase

Chapter 5 shows the improvement in results, when the basic method is complemented with use of reduced supply voltage during the post-stress phase, i.e. only at the time of measurement analysis of the aged chip.

4. Simulation Framework

All the circuits are simulated in HSPICE simulator. A brief overview of the technology models and simulation models for aging and process variation is provided in this chapter. It also describes the type of test circuits used for the experiments, and the statistical method used for data analysis.

4.1 Predictive Technology Model (PTM)

For this thesis, publicly available PTM transistor models (http://ptm.asu.edu/) at 90nm technology node are used. PTM provides accurate, customizable, and predictive model files for future transistor and interconnect technologies. These predictive model files are compatible with standard circuit simulators, such as SPICE, and scalable with a wide range of process variations. Built upon previous Berkeley Predictive Technology Model (BPTM), PTM includes the correlations between process parameters and closely matched the data from different foundries [26]. PTM models are based on standard BSIM4 for bulk CMOS. The model cards corresponding to 90nm technology node are generated from the website. Two fundamental parameters from the model card are discussed below:

1. **Threshold Voltage**

   For a long channel transistor, with uniform doping assumption, the equation for threshold voltage is as presented in *(3)*. The more detailed equation for that

HSPICE uses for V_{th} which accounts for all short channel and doping effects can be found in the BSIM4 User Manual.

$$V_{th} = V_{TH0} + \gamma(\sqrt{\varphi_S - V_{bs}} - \sqrt{\varphi_S}) \qquad (3)$$

where V_{TH0} is the threshold voltage with substrate bias $V_{bs} = 0$, and γ is the body bias coefficient given by equation in *(4)*.

$$\gamma = \frac{\sqrt{2q\varepsilon_{Si}N_{substrate}}}{C_{oxe}} \qquad (4)$$

where $N_{substrate}$ is the uniform substrate doping concentration.

2. Effective Gate Length

The effective transistor length is specified by the equation in *(5)*.

$$L_{eff} = L_{drawn} + XL - 2dL \qquad (5)$$

where, *XL* is the channel length offset due to the mask/etch effect, and dL is calculation by using *(6)*.

$$dL = LINT + \frac{LL}{L^{LLN}} + \frac{LW}{W^{LWN}} + \frac{LL}{W^{LWN}L^{LLN}} \qquad (6)$$

where *LINT* is the channel length offset parameter, and

LL LW, are coefficients of length and width dependence for length offset. Finally, *LLN and LWN* are power of length dependence for length offset

The V_{TH0} and *LINT* are the two main parameters that are varied in our simulations to incorporate process variations.

4.2 MOS Reliability Analysis (MOSRA)

The effect of aging is simulated through MOS Reliability Analysis (MOSRA) model provided by HSPICE. MOSRA accurately models the HCI and BTI aging mechanisms and analyzes their impact on circuit performance using actual circuit operation and stimulus [17].

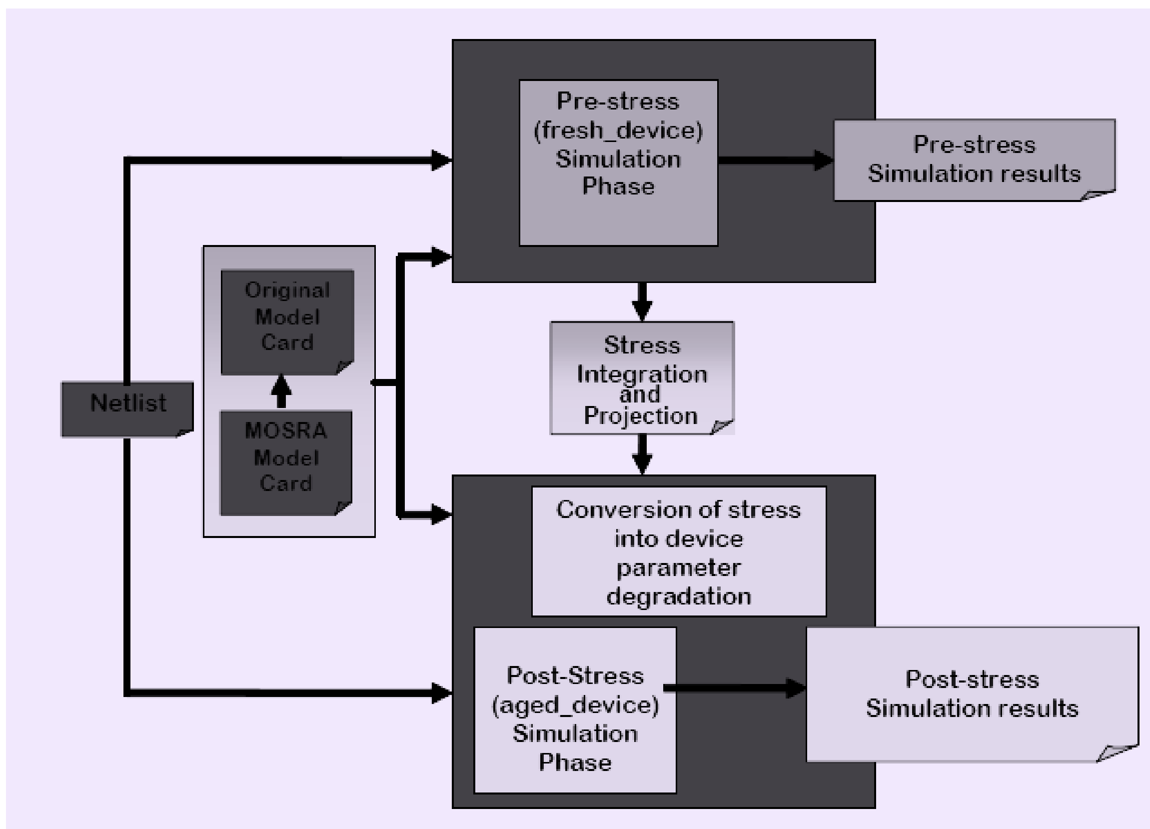

Figure 4-1 The MOSRA flow in HSPICE

It operates in two phases of simulation: pre-stress and post-stress, as shown in Figure 4-1. During the pre-stress simulation phase, the simulator computes the electrical stress of selected MOSFETs in the circuit based on the MOSRA models. The calculation depends on the electrical simulation conditions of each targeted device. The stress value from the MOSRA equation is integrated over a specified simulation time interval, through the duration of the transient analysis. The result of the integration is then extrapolated to

calculate the total stress after a specified time of circuit operation (age). During the post-stress phase, a second simulation is launched. The degradation of device characteristics is therefore translated to performance degradation at the circuit level.

Different operating voltages can be used for each phase. In our analysis, we use the nominal voltage for the pre-stress phase, which represents the typical usage of the circuit. We use different voltages for the post-stress phase, which corresponds to testing stage. MOSRA model is annotated for both NMOS and PMOS, and both HCI and NBTI effects are incorporated.

4.3 Process Variation Modeling

HSPICE supports several ways for inclusion of Process Variation during circuit simulation. In the past, the conventional method to specify global variations was through process corner files, and the other types of variations mentioned above were either estimated manually or not accounted for. In recent years, statistics blocks can be added to the model files. They allow specification of variations in terms of distributions on device model parameters.

Using these techniques, we can model both Global and Local variations. Global Variations are variations in device characteristics from lot to lot, wafer to wafer, and chip to chip. They are caused by variations in starting material and differences between manufacturing procedures. These variations affect all devices with the same model name alike. Local Variations are defined as variations between devices in proximity residing on the same chip; they are caused by microscopic variations in materials and geometry, and affect different devices differently. A brief summary of some options available in HPSICE are discussed here.

1. **Variations Specified Using DEV and LOT**

This method can be used for parameters in the model card. LOT is keyword for global distribution and DEV for local distribution. So, in effect, different distributions can be specified for DEV and LOT. The distribution number corresponds to the 3-sigma value. The format followed is:

parameterName=parameterValue LOT/distribution LotDist
+DEV/distribution DevDist

Example: vth0=0.6 lot/agauss 0.1 dev/agauss 0.02
This provides variation of 0.1V and 0.02V on parameter vth0, around the mean value of 0.6V, for global and local respectively.

2. Definition of Variation Block

It provides a clear way of specifying global, local and spatial variations. All three classes can be described in the Variation Block in a flexible way by user-defined expressions. A subset of these variation types can also be selected in a simulation.

Example:

.Variation
    Define options
    Define common parameters that apply to all subblocks
.Global_Variation
    Define the univariate independent random variables
    Define additional random variables through tranformations
    Define variations of model parameters
.End_Global_Variation
.Local_Variation
    Define the univariate independent random variables
    Define additional random variables through transformations

Define variations of model parameters

.End_Local_Variation

.Spatial_Variation

Define the univariate independent random variables

Define additional random variables through tranformations

Define variation of model parameters

.End_Spatial_Variation

.End_Variation

3. Definition of Variation on parameters

Variations can be specified on individual parameter basis, in form of distributions – uniform or normal. This can be done for parameters at several levels – model, instance or sub-circuit. The instance parameter definition overrides the model definition for parameter.

Example1:

.param var=AGAUSS (20, 1.2, 3)

// Format: AGAUSS (nominal_val,absolute_var, num_sigmas)

AGAUSS distributions will vary around the nominal_val according to the number of standard deviations (num_sigmas) this absolute variation represents.

Example2

.param globp = agauss(-0.356,0.021,3)

.param globpvth = globp

.param localp = agauss(0,0.02,3)

.param pfet_var = 'globpvth+localp'

Two kinds of distributions can be used to derive a single parameter also. As shown in example to the final value for 'pfet_var' is the sum of the 'globpvth' and 'localp'. Therefore, the instances on which this parameter is used will have

different values of V_{th} using a larger common variation due to 'globpvth' and a separate narrower variation due to 'localp'.

Our Approach: The fast/typical/slow corners of the MOS transistors models are generated from the PTM website by incorporating 10% change in Leff and a 3σ change (30mV) in V_{th}. Instead of using the global corners directly, we pick the models corresponding to typical corner. In order to model the process variation, we tune the parameters 'V_{TH0}' (for threshold voltage variation) and '*LINT*' (for length variation), around the values of nominal corner using the DEV/LOT method. We have used different process variation rates for our simulation results, as shown in Table 4-1.

Process Variation class	**Global Variation (absolute value of 3σ variation)**		**Local Variation (absolute value of 3σ variation)**	
	V_{TH0}	**LINT**	**V_{TH0}**	**LINT**
PV1	30mV	1.75nm	None	None
PV2	30mV	1.75nm	10mV	0.58nm
PV3	30mV	1.75nm	20mV	1.16nm

Table 4-1 Different Process Variation Classes used for Simulation

In our simulations using Monte Carlo, 100 sample chips are evaluated for each of the age groups. In each Monte Carlo sample, for every parameter with Global distribution, a different random value is used. In addition, if Local variation is allowed, different instances of the same model can have different values for the parameters, from the values dictated by the local distribution. The characterization step that involves obtaining the reference delay curves uses 70 samples chips for every age group. For the testing phase, the remaining 30 sample chips (in all 120 unknown chips) are used for validation of prediction scheme.

4.4 Test Circuits

A variety of test circuits are used in this thesis. The results for proof of concept are based on two sets of circuits. The first set consists of a variety of inverter chains that differ from each other in their typical delay and activity. Typical delay is the characteristic delay of the path, which has not been affected by either process variation or aging. We modify the depth of the chain and the load capacitance at gate links in order to arrive at different typical delays (Figure 4-2). Load can be modeled as a simple capacitive load or transistor load. The input stimulus is controlled to bring desired variation in the activity for different paths.

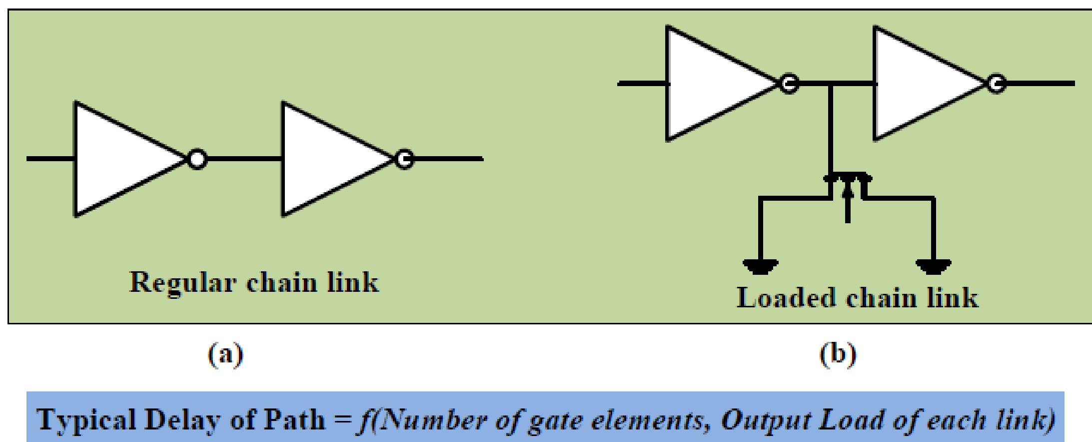

Figure 4-2 Modeling of typical delay with load variation

The second set of test circuits consists of typical circuits such as dividers, discrete-cosine-transform circuits and a few ISCAS benchmarks. The SPICE level simulation of the entire circuit of these sizes can be very time consuming, and in fact, infeasible when combined with aging analysis. Therefore, a circuit-pruning approach is adopted.

Circuit-Pruning: With use of Synthesis and Timing tools, the paths of interest e.g. top critical paths are listed. Then a smaller circuit is built just using these paths. The fan-ins of the paths are pulled in as the primary inputs, and justification vector is identified, and

plugged in as static stimulus to the new circuit. While doing so, attention is paid to the load of the gates along the path in the original circuit. In fact, the value of these loads is extracted and appended in form of equivalent capacitances. Now Monte Carlo Analysis can be done on this smaller circuit.

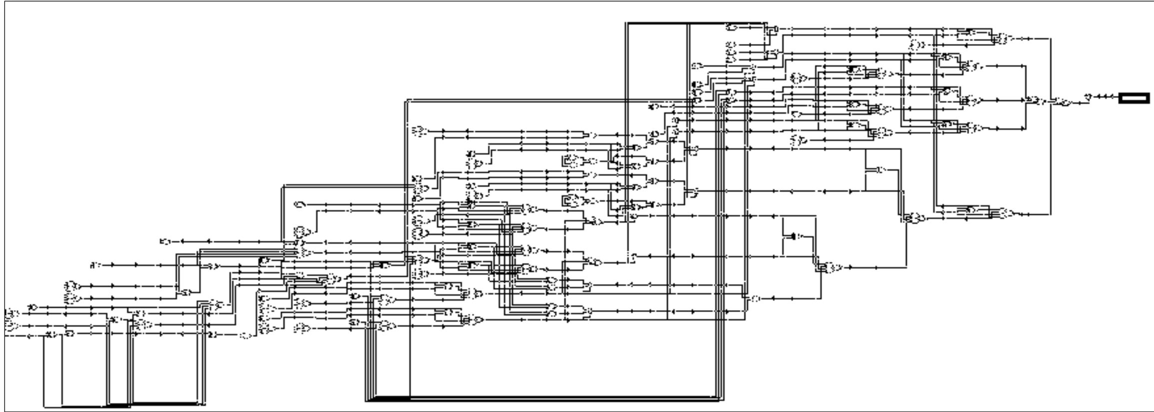

Figure 4-3 Circuit pruning to extract only paths of interest from the original circuit

4.5 Statistical Data Analysis

Regression Analysis and Curve-fitting tool-box of MATLAB are used for the statistical data analysis. Initial correlation analysis of the delay value of two paths is used to determine if the two quantities are fit for a linear regression. MATLAB functions ‘polyfit’ and ‘polyval’ are used to fit data to a model that is linear. The Model and the Data can be plotted on the same plot to visually see the accuracy of the fit. This is done in the characterization step. During the testing phase, the unknown delay point is identified to belong to a certain reference curve using methods like least perpendicular distance or least residual. Both work equally well for the 2-path method. For method involving more than 2 paths, advanced polynomial fit functions are used, and only residual based method is used for testing phase analysis.

5. Results, Analysis and Improvements

This chapter is organized in three parts. In the first part, we use the set of inverter-based circuits for establishing the results of our base method, and the augmented reduced voltage method. In the second part, using the same set of circuits, we show a detailed sensitivity analysis of the method of detection to the type of paths selected. As a remedy to this sensitivity, we introduce 3-path approach, and show the improved results. In the third part, we share our results on some standard benchmark circuits.

Test results for the design under test are categorized into two sets:

1. Plain New/Old detection

   Our aim is to identify whether the IC under test is old or new, instead of finding out how old it is. Correct prediction in this test category implies the power to distinguish between a new and old device.

2. Exact Prediction of age

   For exact age analysis, we start with pre-defined number of age bins, and then the IC under test is classified into one of them. For most of our experiments we consider 4 possible age bins: 0 (new), 2, 5 and 10 years. In some sections, the lower aging range from zero to one year is also considered. Correct prediction in this case implies correct match between the predicted and actual age.

Different Process Variation classes as illustrated in Table 4-1 are used in different sections of the results, and comparison between the detection rates obtained in each of them is also provided.

5.1 Basic 2-Path Method

Table 5-1 shows the results achieved for the 2-path method. It also provides some basic information about the two paths involved in this experiment.

Candidate Paths	Transition Activity Ratio	**100:1**
	Typical Delay Difference	**None**
	Post-Stress Voltage	**1.2V**
	Process Variation Class	**PV1**
Prediction of New/Old	Overall Correct Detection	**87%**
	False Negatives	**6 out of 90**
	False Positives	**10 out of 30**
Exact Prediction of Age	Overall Correct Prediction	**58%**

Table 5-1 Prediction results for Basic 2-Delay method

The result shows that if the two paths have an equal typical delay (in the absence of PV effect) and one of them is 100 times more active than the other, we are able to predict with 87% accuracy, whether a chip is old or new. The terms "False Positive" and "False Negative" are w.r.t true detection of an old IC. This method is seen to have a high false positive rate, which means we will be discarding more than 30% authentic chips as potential counterfeits. It is also shown that this method does not do very well in determining the age of the IC as the exact age prediction is correct only 58% of the time. There exists a lot of scope for improvement.

Continuing with our investigation, for the same experimental setup, instead of making the post-stress measurements at 1.2V, we use an operating voltage of 0.5V. As illustrated in Figure 5-1, the relationship curves obtained at 0.5V are more distinct from one age group to another, and therefore, result in higher correct prediction rate. With "Just-in-time voltage reduction" method, both false positives and false negatives are reduced to zero, and we are able to achieve 100% accurate detection of old counterfeits from new ICs. Also, the exact age prediction is also accurate 100% of the times, for the candidate paths of above characteristics.

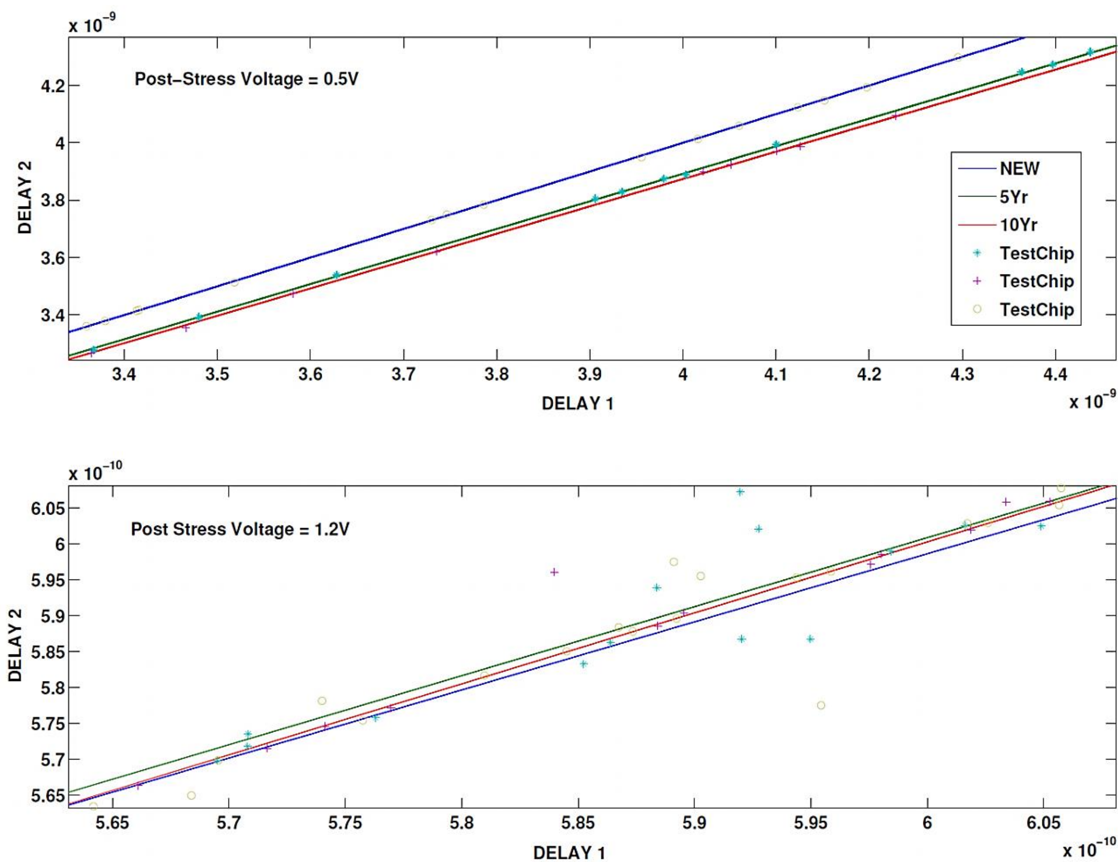

Figure 5-1 Relationship curves between 2 delays, across different age groups, at different post-stress voltages of 0.5V (top) and 1.2V (bottom)

The results in Table 5-2 show the exact numbers for correct predictions across different age groups (0, 2, 5, and 10) and across different post-stress voltages. There are 30 ICs of

each group in the unknown pool of chips under test. The number in the columns, denote the number of correct matches between the calculated and predicted age. And the last column is the overall correct detection in percentage, named as "HIT" percentage. For the test category of plain OLD/NEW detection, an IC is considered correctly detected, if it is distinguishable from AGE 0.

SUPPLY VOLTAGE	**# of correct predictions out of 30 chips**					**HIT (%)**
	AGE in years	**0**	**2**	**5**	**10**	
1.2V	OLD/NEW	20	29	28	27	**86.6**
	Exact Age	20	16	11	17	57.5
0.8V	OLD/NEW	30	30	30	30	**100**
	Exact Age	30	26	24	26	88.3
0.5V	OLD/NEW	30	30	30	30	**100**
	Exact Age	30	30	30	30	100

Table 5-2 Comparison of prediction for various age groups across different voltages

Here are observations on the results presented in Table 5-2:

1. For plain old/new detection test at 1.2V, it seems that the number of correct predictions have a random pattern. As we saw in Figure 5-1, at 1.2V, the relationship curves of different age groups are very close for this voltage. Therefore, the mathematical method of residuals or least perpendicular distance is not able to categorize the unknown device correctly against the reference delay curves. So, the observed random trend is caused due to error of finding best fit curve among highly overlapping ones, which results in more false detections and misses occur. To reduce this error, we move to lower voltages, where different age groups have more curves.

2. The detection rates improve with lowering of post-stress supply voltage, for both the test categories. Even at 0.8V, the plain detection improves to 100%, but there

is still scope for improvement in the exact age prediction, which is achieved by further reduction of the supply voltage.

Based on the above experiments, we decided to perform the rest of our analysis with 0.5V in the post-stress stage. Although the detection numbers are appealing, this method is found to be sensitive to the characteristics of the two paths under test. Details of the sensitivity trends along with a proposed remedy are provided in a later section.

5.2 Comparison across different levels of Process Variation

In accordance with the information provided in Table 4-1, different process variation classes are set through different Monte Carlo simulations. PV1 incorporates only inter-die variations of 3-sigma magnitude on threshold voltage and gate-length parameters. PV2 and PV3 include both inter and intra die variations using DEV/LOT method. The intra-die variations in PV3 are larger in magnitude than in PV2. The detection results, corresponding to these three classes are shown in Table 5-3. The number of age groups for this simulation result is limited to four (0, 2, 5 and10 years).

Candidate Paths Characteristics	**Transition Activity Ratio**	20:1
	Typical Delay Difference	None
	Post-Stress Voltage	0.5V
PV Class	**Old/New Detection (%)**	**Exact Age prediction (%)**
PV1	100	100
PV2	97	61
PV3	81	45

Table 5-3 2-path method results for different levels of process variation

As expected, the increase in degree of process variation makes it difficult to isolate the impact of aging. It can be seen that when the intra-die variation is introduced, the percentage detection reduces for both test categories. As the magnitude of intra-die

mismatch grows, the detection rates suffer further. However, PV2 can be considered as a more realistic estimate for covering both kinds of variations. So, during the sensitivity analysis of the 2-path method in the later section, we restrict our analysis to categories PV1 and PV2 only.

It is also observed that even though the typical delay difference between the candidate paths is none, the success rate has dependence on the delay magnitude of the candidate paths. The higher the delays of the paths, the better are the prediction results. The above table uses 4 path-pairs, and the quoted value is the average prediction rate achieved over the 4 different pairs. The typical delay between sets of pairs is varied by controlling the output loading along the paths. Output loading of each pair is normalized against one base pair, in order to create the 4 grades in typical delay.

5.3 Detection within lower range of Age

Most of the aging is expected to happen in the early phases of the IC's life, and the simulation results support the same. Here, we show the distinguishing power of the 2-path method in the lower age range that is to 1 year of age only. Four age groups are considered - 0, 3, 6, and 12 months. Three sets of Monte Carlo simulations are done to cover the different process variation categories.

PV Class	Old/New Detection (%)	Exact Age prediction (%)
PV1	100	93.3
PV2	84	48
PV3	67	38

Table 5-4 2-path method within lower age groups (0, 3,6,12 months)

As noted in the previous sub-section, the absolute typical delay of the chosen identical paths also matter. And in the lower age range, the sensitivity of the prediction rate is even higher to this aspect. Using similar approach, 4 grades of typical delays are evaluated,

and the result is averaged. As shown in Table 5-4, the distinguishability in the lower age bracket, is not as precise as in the upper age brackets. And in presence of intra-die variations, it can become worse than a unbiased coin-toss (50%).

5.4 Sensitivity Analysis and Improvements

In this section, we have analyzed the sensitivity of our detection method to variation in two characteristics of the paths under test:

1) Activity ratio between the selected paths
2) Difference in the typical delays of selected paths

5.4.1 Sensitivity Analysis

Difference in the activity of the paths is the key reason behind the success of our method. We use the fact that the activity directly influences the aging profile of a path or the degree of aging in a path. One goal is to find out the minimum degree by which they should be different, in order to yield high predictability. In addition, we would like to relax the constraint that the chosen paths should be of same typical delay.

For a comprehensive view of the sensitivity to the two factors together, a variation matrix is plotted to show the impact on both test types (plain detection and exact age prediction). For the variation matrix in Figure 5-2, the typical delay is varied between 0 to 10% and the activity ratio is varied from 2x to 20x. The percentage detection is recorded across this variation matrix. This result set corresponds to PV1 of Process Variation Class, which includes only inter-die variations.

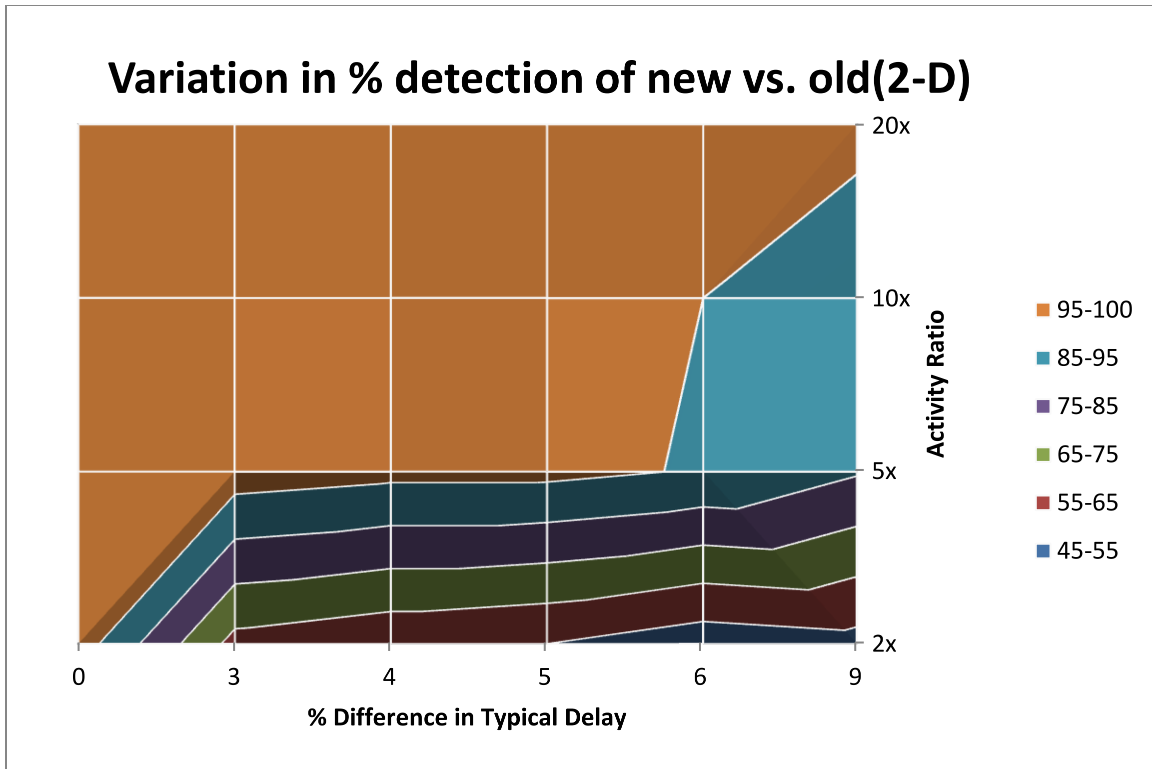

Figure 5-2 Sensitivity of 2-D prediction method (new/old)

Along the activity axis, it is clear that for a fixed set of paths, the prediction result improves as we move from a lower activity difference to a higher one. As for variation to difference in typical delay, it is seen that deviation from similar typical delay causes deterioration of the detection rate. This variation is very limited in the region of high activity difference, but drops the detection rates from 100% to about 50%, in the region of very low activity difference.

Similarly we can plot the exact-age prediction percentages along the variation matrix as shown in Figure 5-3. Similar trends are observed along the activity axis and typical-delay axis, but for this test category, the variations in prediction rates are more distinctly spread across the entire variation matrix. While in plain detection, an activity ratio of 10x was good enough to provide 100% for complete typical delay difference range of 10%, in case of exact-age prediction, even at activity ratio of 20x, the prediction rate can go down to 60%.

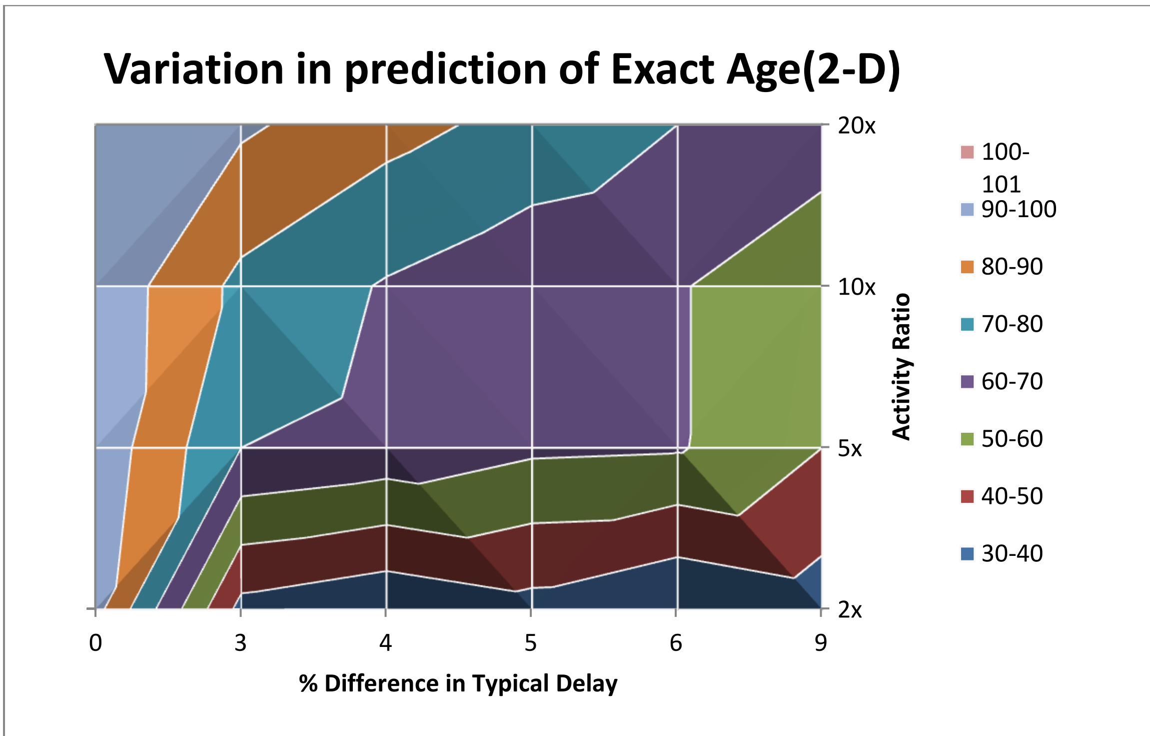

Figure 5-3 Sensitivity of 2-D prediction method (Exact Age)

In conclusion, for the 2-path method, for a reasonably good and stable prediction rate, minimum activity ratio should be 5x. But in order to reduce the sensitivity to typical delay difference, we need to explore alternate approaches. One simple idea we used was to extend the method to include more number of delay paths. By doing so, we extract additional information from the IC under and evaluate if this mitigates the variation in detection rate. In the following section we discuss about addition of just one extra path, and the sensitivity evaluation is repeated for variation matrix, leaving out lower range of activity ratio and looking at 5x and beyond on the activity axis.

5.4.2 Further Improvement: 3-path Approach

In the 3-path method, an additional path is chosen. Similar to our 2-path method, the relationship curve is now defined between the three delay variables using surface fit techniques. And for the unknown test point, minimum residual is used to identify the

best-fit surface it belongs to. Figure 5-4 shows a 3-D hyper-surface training for different age-groups, and mapping of the unknown IC under test to one of them.

Since the search space for how these three paths should be related to each other in terms of their typical delay and activity is very large, we impose some restrictions by emulating a scenario in a typical design. In this scenario, we assume that we are able to evaluate two critical paths with similar (but not the same) typical delay from a circuit with some moderate to high activity and then, a third path from a test circuit with much lower activity is chosen.

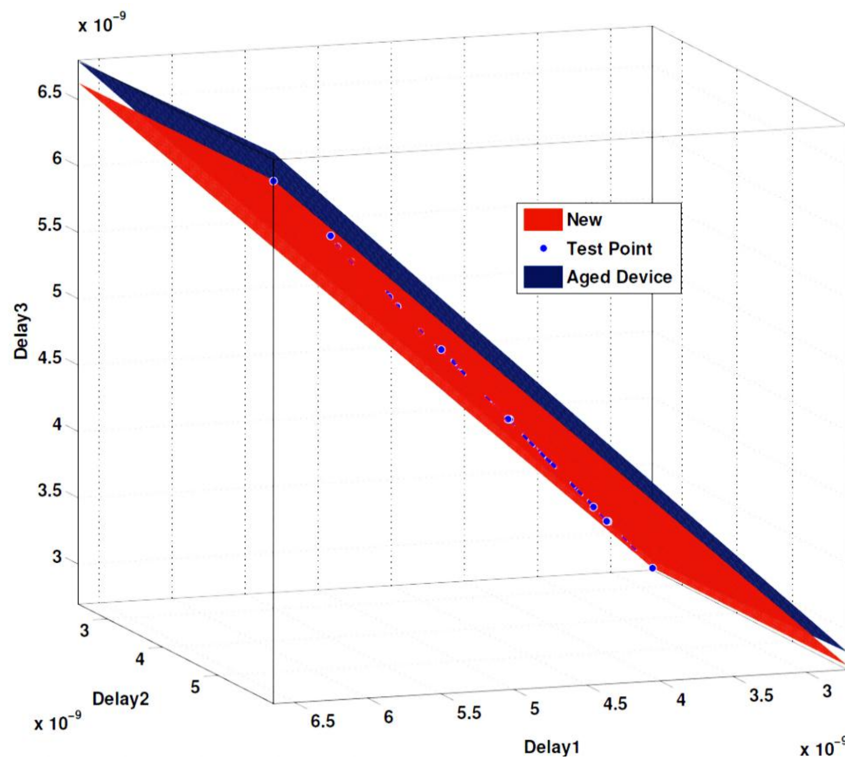

Figure 5-4 3-path method – hyper surface training and mapping

The improvement in the results for both test types is shown in Figure 5-5 and Figure 5-6. These can be directly compared to Figure 5-2 and Figure 5-3 that corresponded to 2-path method. In 2-D, we used a pair of paths, which differed in their transition activity (high-activity and low activity path pair), now a third path is added to the original couplet. It is

observed that it is more beneficial to add a higher activity path, such that the following is satisfied in the chosen triplet:

1) Minimum activity ratio between the high activity paths to low activity path is 5 or more. The activity ratio between the higher activity paths should be minimal.

2) In terms of typical delay, all candidate paths must be within 10% difference in typical delay to each other.

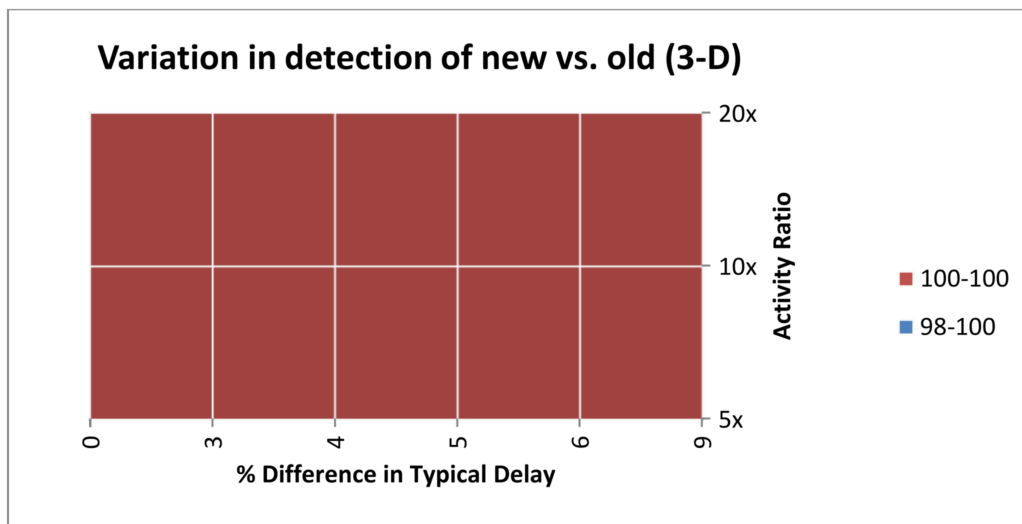

Figure 5-5 Sensitivity with 3-path prediction method (new/old)

In Figure 5-5 and Figure 5-6, for any data cell of the variation matrix, 8 possible triplets are evaluated. Each data cell value records an average of the %prediction over the chosen triplets. On average, large improvements are observed in the prediction results for all triplets, except for a few outliers. Even in 2-path method for activity ratio 5x and beyond the variation in prediction results was not very high, and whatever little sensitivity existed, is now reduced to zero for 3-path method. Significant improvements are seen in the exact-age prediction. For an activity ratio of 20x, correct prediction rate of as high as 96% is guaranteed over the complete range of typical delay variation, which is relatively a high improvement over the results from 2-path method.

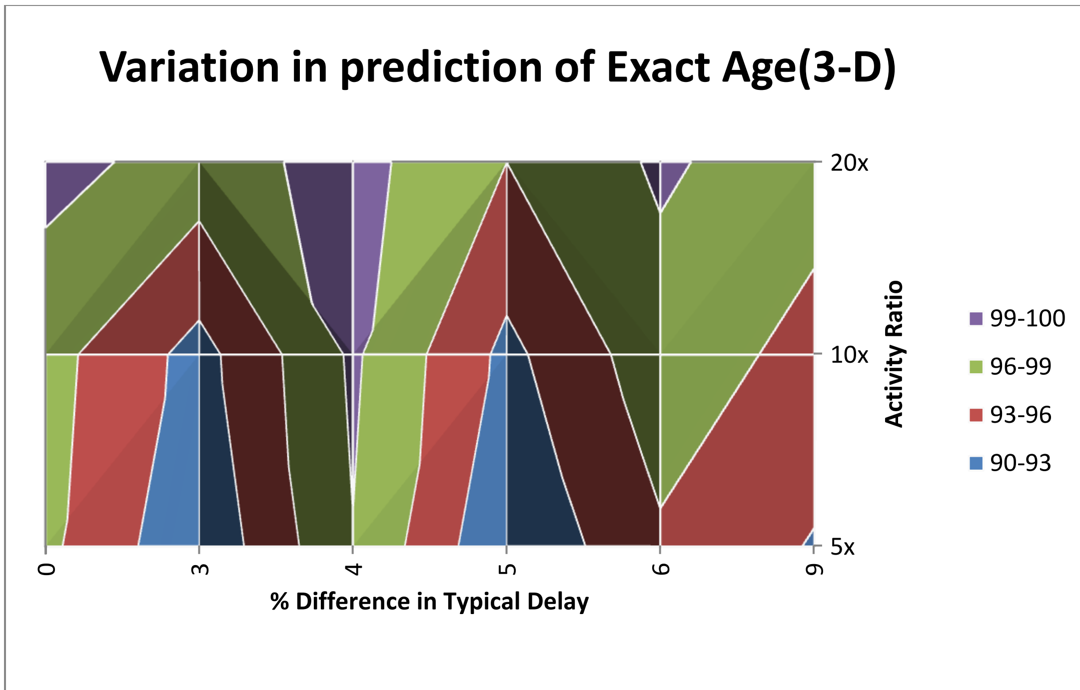

Figure 5-6 Sensitivity with 3-path prediction method (Exact Age)

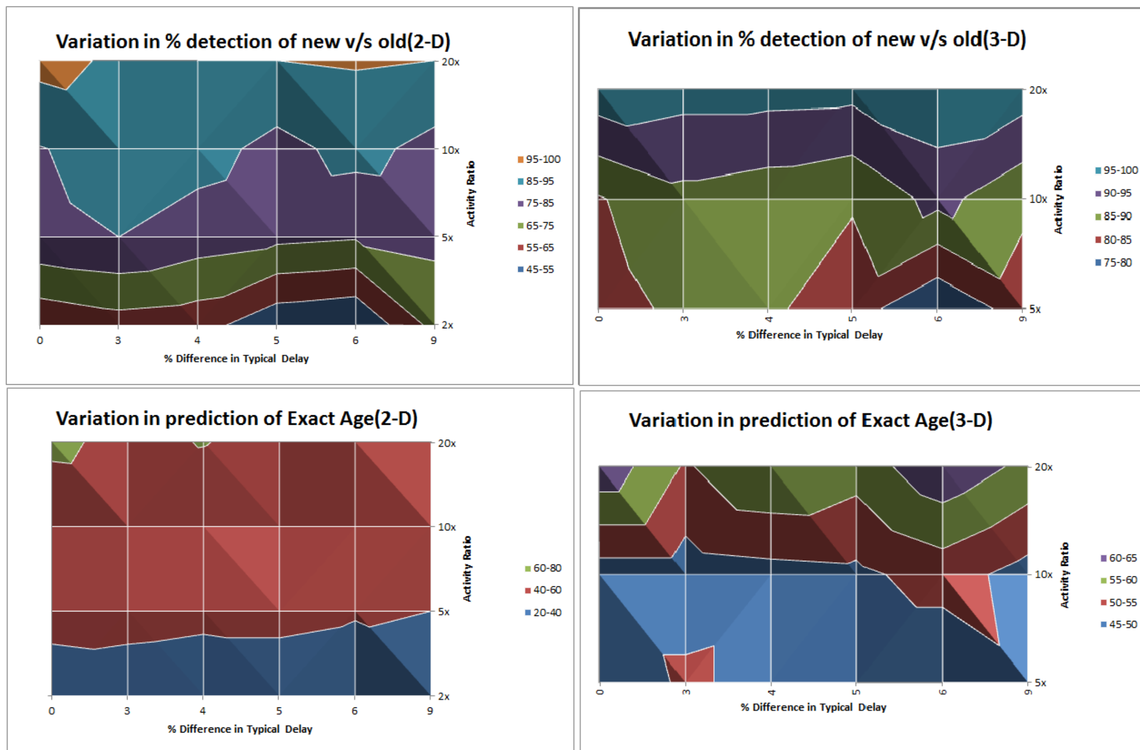

Figure 5-7 Sensitivity analysis in 2-D and 3-D for PV2 category

In order to see if this 3-path method is effective in other categories of higher process variation, Figure 5-7 shows the sensitivity graphs for process variation class PV2, which includes both inter and intra die variations. In the plain detection test category, the percentage detection improves to 90% and above for activity ratio of 10x and beyond. The relative improvement in exact-age prediction is not that large, when compared to the PV1 category. When operating in higher process variation bracket, it is definitely more favorable to choose high activity difference between the paths that form the base pair.

This concludes the sub-section of sensitivity analysis. The idea is extendible to include larger number of paths into the analysis. We restrict it to only three paths, as it is sufficient to yield a reasonably good prediction results. The 3-path technique is validated further on some circuit benchmarks, using the circuit pruning method as discussed in Chapter 4.

5.5 Benchmark circuit results and extension to real ICs

Circuit	**% Diff in typical delay**	**Detection of Old/New**	**Exact Age Prediction**
DCT	6-7%	100%	97.5%
Divider	<1%	100%	98.3%
C6882 (ISCAS)	15%	97.5%	91%

Table 5-5 Results for standard circuits

The 3-path method is deployed on a few circuits like multiplier, divider, discrete-cosine-transform (DCT) module, and combinational circuits from ISCAS benchmark. Top critical paths were extracted from the circuits. As expected, the degree of variation in

typical delays among the critical paths is not more than 10%. The degree of process variation included for this simulation set is PV1. Some sample results for each circuit, is shown in Table 5-5. The typical delay difference between the 3 chosen paths is also shown. Minimum activity ratio is 10x for this experiment.

These results are appealing and establish confidence in the detection method. Here the chosen candidate paths are within 15% of delay difference. This idea is easily extendible to real designs of larger sizes too. Finding the candidate paths, as per the requirement of the detection scheme should not be very difficult. Critical paths from a functionally active block can be chosen. And the third path can be chosen from a test path of the chip. If the test paths are found not to be as long, we can re-choose the high activity paths such that the typical delay is in the range of the critical test path. Eventually, the aim is to not have more than 10-15% typical delay difference between the paths. If the typical delays of the chosen paths are almost equal, then even 2 paths of different activity are sufficient for successful detection.

6. Conclusion

In this thesis, we presented a technique for successful detection of counterfeits among existing ICs. It is important to determine the authenticity of the ICs, as counterfeit ICs pose a threat to the security and reliability of the whole system. The goal is to be able to predict the age of an unknown IC, and deterministically distinguish an old counterfeit from a new authentic IC. This detection is non-trivial for existing ICs, as they do not have in-built age-detection circuits, and we can use only existing circuit characteristics for this purpose. Apart from X-Ray based methods, there is little work done in this direction.

First, it is important to understand the physical aging mechanisms, and their influence on circuit characteristics. Most of the existing aging models represent the degradation as a shift in threshold voltage, which manifests as degradation in delay. During the operating life of the IC, there is degradation in circuit delay due to underlying aging mechanisms, but we cannot simply map the shift in the delay to the age of the device. The reason is that the parameter of circuit delay is also impacted as virtue of process variability – non-uniformity in feature size, and doping concentration etc. This makes it important to isolate the effect of process variability from effect of aging.

In our method, we make use of two or more paths, and use the relationship between their delay variables. Reference relationship patterns are created for different age groups using authentic ICs in the characterization phase. When the unknown IC under test is evaluated,

the same path delays are extracted, and the new relationship between the delays is mapped against the known reference curves, to predict the age of the device.

When using path delay, it is also important to take account of other factors such as path activity, gate type, and gate strength on the aging rate. For our method to work successfully it is important to select path candidates that have distinct aging rate. It is not possible to know the activity that an IC might have undergone during their lifetime. We use the fact that a normal functional path would be more active than a relatively dormant test logic path. Such paths with difference in their activity are great candidates as they exhibit different aging rates. We also augment the method by using reduced supply voltage at the time of test, which is shown to magnify the effects on delay.

Under ideal conditions where path candidates have same magnitude of delay, even 2 paths are sufficient for a 100% prediction result, but as this may not be practically possible to find such paths, we present a sensitivity analysis of the 2-path method to difference in typical delay of path candidates to up to 10-15%. As a remedy to the observed sensitivity, we propose addition of a third path to the analysis. This is shown to improve the detection results significantly. With the 3-path method, our results that were based on industry-enabled simulation framework show a detection rate of over 97% for identifying an old IC from a new IC.

Future Work: As no circuit study is complete without real chip validation, as future work, we plan a chip tapeout. A variety of paths with varied typical delay will be fabricated. All the chips will be extensively characterized in terms of delay while running at different operating voltages. Later, the chips will undergo accelerated aging using heat and high operating voltage. The chips will be consequently characterized several times and the results will be aggregated. The proposed methodology can then be employed to detect aged versus new devices and prediction rates will be compared to that of simulation results.

7. Bibliography

[1] SMT Corporation, 2011. [Online]. Available: http://armed-services.senate.gov/statemnt/2011/11%20November/Sharpe%20Slides%2011-08-11.pdf.

[2] L. Kessler and T. Sharpe, "Faked Parts Detection," 2010. [Online]. Available: http://spectrum.ieee.org/riskfactor/computing/hardware/the-financial-risks-of-counterfeit-semiconductors.

[3] "IHS, ISuppli - Press Release," 2011. [Online]. Available: http://www.isuppli.com/Semiconductor-Value-Chain/News/pages/Top-5-Most-Counterfeited-Parts-Represent-a-$169-Billion-Potential-Challenge-for-Global-Semiconductor-Market.aspx.

[4] "Combating Countefeits with IHS," Dec 2011. [Online]. Available: http://www.ihs.com/info/sc/a/combating-counterfeits/index.aspx?tid=t4.

[5] "Fake chips from China sold to U.S. defense contractors - Reuters US Edition," 26 Oct 2010. [Online]. Available: http://www.reuters.com/article/2010/10/26/us-china-counterfeit-defence-idUSTRE69P3GH20101026.

[6] A. Baba and S. Mitra, "Testing for transistor Aging," in *VLSI Test Symposium*, 2009.

[7] X. Chen, Y. Wang and Y. Cao, "Variation-Aware supply voltage assignment for minimizing circuit degradation and leakage," in *ISLPED*, New York, USA, 2009.

[8] T. Douseki, M. Harada and T. Tsuchiya, "Ultra-low-voltage MTCMOS/SIMOX technology hardened to temperature variation," *Solid-State Electronics,* vol. 41, no.

4, pp. 519-525, 1997.

[9] T. Kim, R. Persaud and C. Kim, "Silicon odometer: An on-chip reliability monitor for measuring frequency degradation of digital circuits," *IEEE Journal of Solid-State Circuits,* vol. 43, pp. 874-880, 2008.

[10] D. Lorenz, G. Georgakos and U. Schlichtmann, "Aging analysis of circuit timing considering nbti and hci," in *15th IEEE International On-Line Testing Symposium*, June, 2009.

[11] R. Vattikonda, W. Wang and Y. Cao, "Modeling and Minimization of PMOS NBTI Effect for Robust Nanometer Design," in *Design Automation Conference(DAC)*, 2006.

[12] M. Alam and S. Mahapatra, "A comprehensive model of pmos nbti degradation," *Microelectronics Reliability,* vol. 45, no. 1, pp. 71-81, 2005.

[13] S. Borkar, "Electronics beyond nano-scale CMOS," in *DAC*, NY,USA, 2006.

[14] D. Schroder and J. Babcock, "Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing," *Jounrnal of Applied physics,* pp. 1-18, 2003.

[15] X. Zhang, N. Tuzzio and M. Tehranipoor, "Path-Delay Fingerprinting for Identification of Recovered ICs," in *IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems*, 2012.

[16] P. Gupta and A. Kahng, "Manufacturing-aware physical design," in *ICCAD*, Washington,DC,USA, 2003.

[17] B. Tudor, J. Wang, Z. Chen, R. Tan, W. Liu and F. Lee, "An Accurate and Scalable MOSFET Aging for Circuit Simulation," in *12th International symposium on Quality Electronic Design*, 2011.

[18] F. Koushanfar and G. Qu, "Hardware Metring," in *DAC*, 2001.

[19] F. Koushanfar, "Integrated circuits metering for piracy protection and digital rights management: An overview," in *GLSVLSI*, 2011.

[20] Y. Alkabani and F. Koushanfar, "Active hardware metering for intellectual property protection and security," in *16th USENIX Security Symposium*, Berkely,CA,USA,

2007.

[21] W. Griffin, A. Raghunathan and K. Roy, "Clip: Circuit level ic protection through direct injection of process variations," *Very Large Scale Integration Systems,* vol. 20, pp. 791-803, 2012.

[22] S. Devadas, E. Suh, S. Paral, R. Sowell, T. Ziola and V. Khandelwal, "Design and implementation of puf-based ”unclonable” rfid ics for anti-counterfeiting and security applications," in *RFID*, 2008.

[23] X. Zhang, N. Tuzzio and M. Tehranipoor, "Identification of Recovered ICs using Fingerprints from a Light-Weight On-Chip Sensor," in *DAC*, 2012.

[24] DAGE - A Nordstorm Company, "Detecting Counterfeit Components," *Chip Scale Review,* 2008.

[25] A. Wang, B. H. Calhoun and A. P. Chanrakasan, Sub-Threshold Design for Ultra Low-Power Systems, vol. XII, Springer, 2006, p. Pg. 75.

[26] W. Zhao and Y. Cao, "New Generation of Predictive Technology Model for Sub-45 nm Early Design Exploration," *IEEE Transactions on Electron Devices,* vol. 50, 2006.