Portland State University
PDXScholar

Dissertations and Theses | Dissertations and Theses

Fall 10-20-2014

Training Set Design for Test Removal Classication in IC Test

Nagarjun Hassan Ranganath
Portland State University

Let us know how access to this document benefits you.

Follow this and additional works at: http://pdxscholar.library.pdx.edu/open_access_etds

Part of the Electrical and Computer Engineering Commons

Recommended Citation
Hassan Ranganath, Nagarjun, "Training Set Design for Test Removal Classication in IC Test" (2014). *Dissertations and Theses.* Paper 2028.

10.15760/etd.2028

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.

Training Set Design for Test Removal Classification in IC Test

by

Nagarjun Hassan Ranganath

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Electrical and Computer Engineering

Thesis Committee:
W.Robert Daasch, Chair
C.Glenn Shirley
Mark Faust

Portland State University
2014

Abstract

This thesis reports the performance of a simple classifier as a function of its training data set. The classifier is used to remove analog tests and is named the Test Removal Classifier (TRC).

The thesis proposes seven different training data set designs that vary by the number of wafers in the data set, the source of the wafers and the replacement scheme of the wafers. The training data set size ranges from a single wafer to a maximum of five wafers. Three of the training data sets include wafers from the Lot Under Test (LUT). The training wafers in the data set are either fixed across all lots, partially replaced by wafers from the new LUT or fully replaced by wafers from the new LUT.

The TRC's training is based on rank correlation and selects a subset of tests that may be bypassed. After training, the TRC identifies the dies that bypass the selected tests.

The TRC's performance is measured by the reduction in over-testing and the number of test escapes after testing is completed. The comparison of the different training data sets on the TRC's performance is evaluated using production data for a mixed-signal integrated circuit.

The results show that the TRC's performance is controlled by a single parameter - the rank correlation threshold.

Acknowledgements

I would like to express my gratitude to many wonderful people who have encouraged me during my graduate study at Portland State University. First, I would like to thank my research advisor Dr. Robert Daasch for his guidance and encouragement. His extensive knowledge and research experience in the field of IC test and statistics helped me not only to achieve the goal but also improve my critical thinking and presentation skills. I am very thankful to our industrial liasons at Texas Instruments (TI), Kenneth Butler, John Carulli, Amit Nahar and Chaitrali Chandorkar for their valuable industrial insights in this project.

Also, I would like to express my gratitude to Dr. Glenn Shirley and Professor Mark Faust for serving on my M.S. Defense Examination Committee. I would like to thank Professor Shirley and Professor John Acken for all the help and guidance they provided me in the two years of my association with ICDT (Integrated Circuit Design and Test) laboratory. Special thanks to my ICDT lab colleagues Neerja, Yu Shan, Navneeth, Chris and Kris for all the support and encouragements. Special thanks to my friends Srivarshini, Manas and Rakesh for their support and motivation. They helped boost my confidence during difficult times.

Finally, I would like to thank my family for supporting me and believing in me. They helped me stay focused and motivated throughout my course at Portland State University.

Contents

Abstract i

Acknowledgements ii

List of Figures iv

1 Introduction and Previous Work 1

1.1 Introduction to IC Test 1

1.2 Introduction to Adaptive Test 2

1.3 Introduction to Machine Learning 2

1.4 Brief Description and Contribution of the Thesis 4

1.5 Previous Work 6

1.6 Organization of the Thesis 8

2 Background Concepts 9

2.1 Rank 9

2.2 Properties of Rank order 9

2.3 Correlation 10

2.4 Kendall's tau rank correlation 11

2.5 Mixed Signal Test 14

2.6 Test Limits 15

2.7 Binning dies 16

2.8 Queue and Server 17

2.9 Buffer 18

3 Test Removal Classifier **20**

3.1 Typical Two Insertion Test Flow 21

3.2 Correlation among Tests 22

3.3 Test Data and Matched Test Pairs 25

3.4 Matched Test Pair Selection Scenarios 27

3.5 Test Removal Limits 28

3.6 Adaptive Test Flow 31

4 Training Strategies **34**

4.1 Window Training Strategy 34

4.2 Accumulate Training Strategy 37

4.3 Single LUT Wafer Training Strategy 39

4.4 Dual Lot and Single Lot Training Strategy 40

4.5 Special Case Training Strategies 41

5 Results using Production Data **43**

5.1 Classification of Die 43

5.2 Figures of Merit 45

5.2.1 Product Engineering Figures of Merit 47

5.2.2 Test Engineering Figures of Merit 48

5.3 Production Data Set 49

5.4 Results 51

5.5 Cross Validation of the Training Strategies 58

5.6 TRC Performance Dependence on Threshold 60

6 Conclusion and Future Work **63**

6.1 Conclusions 63

6.2 Future Work 64

References **65**

List of Figures

1.1 Adaptive Test flow 5

2.1 Concordance-Discordance plot 13

2.2 Test Limits 16

2.3 Queue and Server Model 18

2.4 Data Buffer - FIFO Configuration 19

3.1 Typical Two Insertion Test flow 21

3.2 Joint distribution in the Parameter Space 23

3.3 Joint distribution in the Rank Space 24

3.4 Test data and Matched Test Pairs 25

3.5 Matched Test Pair Selection Scenarios 27

3.6 Matched Test Pair Joint Distribution 29

3.7 TRC Limits for a Matched Test Pair 30

3.8 Test Removal Classifier 32

4.1 Window training strategy 36

4.2 Accumulate training strategy 38

4.3 Single LUT Wafer training strategy 39

4.4 Dual and Single training strategy 40

5.1 Classification of Die 44

5.2 CDF of Matched Test Pair Utilization - Intersection 55

5.3 CDF of Matched Test Pair Utilization - Union 56

5.4 PDF of Matched Test Pair Utilization - Intersection 57

5.5 PDF of Matched Test Pair Utilization - Union 58
5.6 Matched Test Pair Use - Errorbar plots 59
5.7 Dependence of FOM on τ Threshold 60

Chapter 1

Introduction and Previous Work

1.1 Introduction to IC Test

Manufactured dies have to be tested before they are shipped to the customer so as to ensure that the customer does not receive dies outside their datasheet specifications. IC testing is the process of differentiating the good dies from the defective ones. IC testing is carried out in two stages. The first stage is called wafer sort wherein dies are tested at the wafer level. Dies passing this stage are packaged and tested at the second stage called the package level. Dies which pass this final stage are shipped to the customer. The dies passing package level are sometimes subjected to burn-in and stress tests. This testing phase needs to be comprehensive in order to meet high expectations (typically less than 100 DPPM) of the customer [1].

With continual advancement in the field of semiconductors, the complexity of ICs has increased but the cost of fabrication has downgraded. The need for a comprehensive testing technique for such complex dies has led to significantly increased costs of testing them. Researchers predicted that the cost of testing a die would equal the cost of fabricating it [2] [3]. The industry was not able to drive down the test cost per fabricated transistor. To overcome the increasing costs of testing an IC dies, test engineers adopted many techniques such as multisite testing, design for test (DFT), structural test among others [1]. One method used to reduce the manufacturing test cost is called adaptive testing. A technique used

in adaptive testing is called machine learning. The thesis focuses on machine learning used in adaptive test.

1.2 Introduction to Adaptive Test

Adaptive test, as described by the ITRS 2013 [1], is an industry term for the practice of going beyond the use of fixed limits and invariant test flows and operations, by using data that has been collected during the process of IC manufacturing to influence, change, or "adapt" how a device or system is tested, or to alter its manufacturing flow. The adaptive test alters test limits or changes test rules in response to any manufacturing changes at the die, wafer or lot level [4]. Adaptive test using machine learning tools has been applied at many decision levels. The machine parametrized model makes classification decisions about tests already run or prediction decisions about future events.

1.3 Introduction to Machine Learning

In the context of this thesis, machine learning is a data-driven technique that exploits the correlation among tests carried out at one temperature to predict the outcomes of tests carried out at another temperature. At its root, machine learning sifts example data, called training data, which has been classified some other way. The training data is processed to determine parameters of a classification model based on a statistical model of the system. After the training, new data (assumed to be from the same statistical ensemble but without the classification label) is input to the parametrized model with the output a prediction of the class the new data was from.

The application of machine learning techniques to digital and analog test data has attracted considerable attention and is about to enter its second decade [5] [6]. In test, a common decision scenario is the trade-off between a test's cost and a test's benefit to later testing, product quality or reliability, or other manufacturing costs. The cost/benefit example this thesis uses is the removal of a fraction of tests from wafer testing where the benefit is measured in test removal and the cost is failing die escaping to an expensive (downstream) insertion. Test classifies each manufactured die as pass or fail [7]. During test, the pass or fail classification uses combinations of the die's measured response and categorical (binned) response. In addition to using the data to classify the die, decisions are made about the disposition of the wafer the die is on, the lot the wafer was manufactured in, or the entire production run.

The machine learning performance applied to adaptive test depends critically on the choice of training data [5]. Imbalanced training data is when the population is practically all one class (e.g. >99%). Imbalance is almost always the case in test and extra care in training set selection is needed [8]. The training data selection and use has to meet two requirements to ensure consistent performance of trained and parametrized model. First, training data must represent as completely as possible the response range of each class of interest. Second, there has to be a built-in monitor mechanism to detect and respond to slowly varying trends or sudden excursions in the manufacturing process. This thesis explores different training strategies that meet these requirements and highlights the impact of these requirements on the performance of the machine learning methods.

From the semiconductor test perspective, machine learning can be used in two different ways. When machine learning tools are used as classifiers, the goal is to process the die's test response data from a subset of the tests and assign one or more new labels to the die [9]. A typical binary classifier assigns a label to the die such as pass or fail. When machine learning tools are used as predictors, the goal is to use the data in-hand to anticipate a future result. For prediction the data from many die may be combined to predict something other than a pass/fail label [3] [10]. For example, a predictor may estimate the test response value of a die to an, as yet, unexecuted test in the test flow. Predicted results are used to evaluate the die's actual response to the test once it is executed. Often the goal of machine learning prediction of future test outcomes is to control the risk of eliminating a future test for one, some, or all manufactured die so that the benefit outweighs the risk [11].

Some machine learning test schemes use the predictor as well as the classifier [5]. The Test Removal Classifier described in this thesis is one such scheme. It is based on a parametrized model for predicting the pass/fail bin of a die for a future test. The die's predicted value is followed by a classification step where the die's future tests are scheduled.

1.4 Brief Description and Contribution of the Thesis

The goal of this thesis is to review a key factor common to all machine learning methods, the selection and use of imbalanced training data [8] [12]. The thesis demonstrates the advantage of matching the machine learning output to the level of decision making. The thesis will assume that the machine learning tool is a

decision about a single die.

The study reported in this thesis highlights the selection and dependence of machine learning performance on the training data sets. The study is limited to a simple machine learning classifier called the Test Removal Classifier (TRC). The TRC is purposely chosen to be simple so as to not distract from the central point that training is a degree of freedom which must be well characterized in and of itself.

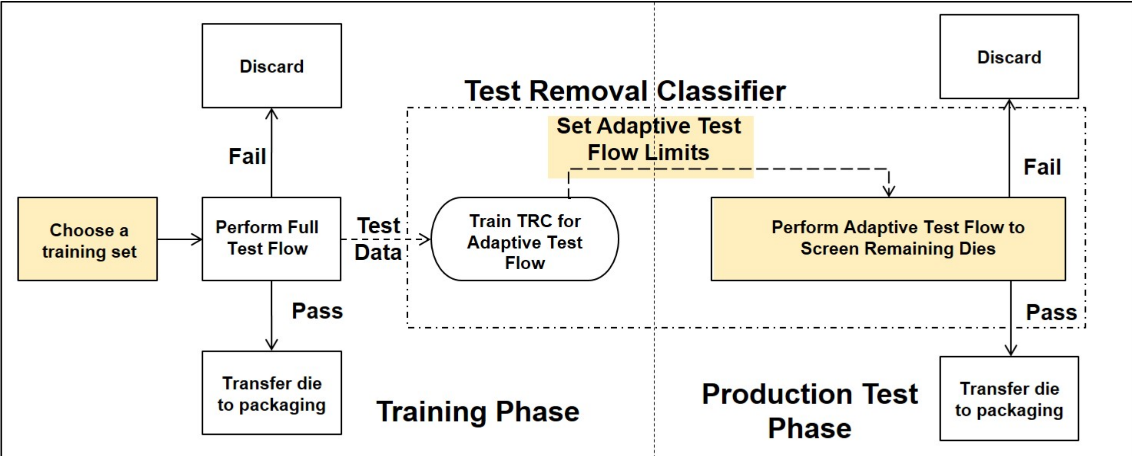

Figure 1.1: Flowchart describing the adaptive test flow applied to a test manufacturing process using machine learning. Full test flow indicates that all tests must be run. Adaptive test flow indicates that depending on the test limits set by the test removal classifier, tests can be eliminated on a per-die basis.

Figure 1.1 depicts the typical adaptive test flow used in this thesis. The test flow is designed on a per-die basis based on the limits set in the training phase. Failing dies are discarded and passing dies are transferred to packaging. The test flow consists of two phases - the training phase and the production test phase. In the training phase, a subset of the dies from wafer sort is chosen to be the training set.

A full test flow is executed for the training set. The test data from the full test flow is used to train the TRC, enclosed by the dashed box. Rank correlation among tests is exploited to identify tests that could be eliminated. Using the training data, limits are set for the adaptive test flow. In the testing phase, the new dies which are not used in training are tested by the adaptive test flow.

The objective of the thesis is to study the impact of the training set selection on the test removal limits and the performance of the TRC, highlighted by the green boxes in Figure 1.1. Seven different strategies are designed to select the training set. The strategies vary by the number of wafers in the data set, the source of the wafers and the replacement scheme of the wafers. Chapter 4 describes in detail the seven different training strategies used in this thesis. Each strategy involves using rank correlation to identify tests that can be bypassed in the production test phase. This is discussed in detail in Chapter 3. The TRC performance results are reported in Chapter 5.

The results of the study show that the TRC performance and its dependence on training is controlled by a single parameter - the rank correlation threshold. The study also reports that stable predictions are found by using a filtering scheme that samples old, recent and new training data.

1.5 Previous Work

Researchers have proposed various different strategies [9] [13] [14] [11] [15] to select training data for the machine learning techniques. But, there is no detailed study in the literature on the aspect of the dependence of the performance of machine learning on the training data selection. Some of the different training strategies

proposed by researchers are discussed in this section.

H. Ayari et al [13] studied the influence of the training set size on the accuracy and reliability of prediction models and concluded that the prediction accuracy degrades with the reduction of the training set size. They proposed a new implementation of predictive alternative testing based on model redundancy. This approach permits to achieve accurate prediction results independent of the training set size.

Mingjing Chen and A. Orailoglu [11] explored an adaptive test strategy that included pruning the test set based on test correlation analysis. Pruning involves elimination and reordering of the test set. The training phase included random selection of chips on which the full test was performed. The analysis of the statistical features of the characterized data obtained by this step led to the pruning of the test set.

S. Benner and O. Boroffice [14] devised a method by which optimal production test times could be obtained using adaptive test programming. They suggested choosing a proper sample size for training based on various criteria such as Acceptable Quality Level (AQL) sampling plans, internal sampling plans, and historical knowledge of the particular device under test (DUT) among others. Whichever way is chosen, the sample size must be adequate enough to produce a good estimate for the mean and standard deviation for the rest of the population.

S. Biswas and R.D. Blanton [15] used stratified random sampling to train the model used for test compaction. When the accuracy of the model dipped below a preset tolerable threshold, this training data was updated using three different

methods of sampling. The scheme was called adaptive re-learning. This scheme was applied to test data from two in-production integrated systems. It was found that the difference between the actual functional accuracy and its estimate using stratified sampling was less than 2%.

H.G Stratigopoulos et al [9] proposed a machine learning approach that used a multi-objective genetic algorithm.The machine learning approach searches a subset of appropriate subsets and a classifier that makes pass/fail decisions based solely on these subsets. The multi-objective genetic algorithm was trained by a fixed set of 1/4N devices, where N is the total number of devices, and validated by the remaining 3/4N devices. To reduce the variance of the reported test error, cross validation was implemented with a uniform sample of devices at random in each run.

1.6 Organization of the Thesis

The thesis is organized into five additional chapters. The next chapter, Chapter 2, briefly describes the background concepts used in the thesis. Chapters 3 and 4 discuss the strategy used to study the dependence of machine learning on training. Chapter 3 outlines the simple machine learning technique called the Test Removal Classifier (TRC). Chapter 4 provides the details of the seven training strategies used in the study. Chapter 5 discusses the five figures-of-merit used in the study and presents TRC results for different training strategies which were described in Chapter 4. Chapter 6 offers conclusions and some recommendations for research in the application of machine learning tools to test.

Chapter 2

Background Concepts

This chapter introduces basic concepts to help better understand the ideas presented in the thesis. This chapter begins with the description of statistical terms which are central to the thesis. The discussion continues with the concept of mixed-signal test followed by the idea of test limits and binning of dies. Finally, the queue and server model is discussed followed by the concept of data buffer.

2.1 Rank

In statistics, ranking refers to the data transformation in which numerical values are replaced by their index when the data is sorted [16]. The rank of a value in a distribution is simply its numbered place in the list of ordered values.

$$\text{S} = 34, 56, 12, 99, 23 \quad \text{Rank-order(S)} = 3, 4, 1, 5, 2$$

Statistics calculated using rank order are called ordinal invariant statistics or non-parametric statistics.

2.2 Properties of Rank order

Rank order does not change if any monotonic transformation is applied to values in the original set. Monotonic transformation can be addition, subtraction, multiplication, division, logarithm etc [17]. For example: Consider the monotonic transformation $S' = S \cdot 100$ where all the elements of S are multiplied by 100 to obtain S'. The rank order of S' remains the same as S after the monotonic transformation.

$$S' = 3400, 5600, 1200, 9900, 2300 \quad \text{Rank-order}(S') = 3, 4, 1, 5, 2$$

When two or more observations in a set have the same values they are called ties. Tied data values have to be resolved while ranking the data. There are several different methods to handle ties. The method used in the thesis is fractional ranking. In fractional ranking, tied values receive the same ranking which is the mean of ordinal ranks.

$$S' = 23, 56, 12, 99, 23 \quad \text{Rank-order}(S') = 2.5, 4, 1, 5, 2.5$$

The two numbers in S' have same value 23. Hence, in the rank-order, the tied values are assigned the rank 2.5 which is the mean of ranks 2 and 3.

2.3 Correlation

The concept of statistical correlation is central to this thesis. Correlation is a statistical measure of dependence between two variables. Consider two random variables X and Y. If an increase in X is associated with increase in Y then they have a dependence and are said to be positively correlated. If an increase in X is associated with a decrease in Y then they have anti-dependence and said to be negatively correlated.

Correlation between two variables is measured in terms of the correlation coefficient. Correlation gives information regarding relationship between variables. The correlation coefficient is a measure of the strength of the relationship between the two variables. In the thesis, correlation is used to study the relationship between tests at one temperature and tests at a higher temperature.

Pearson's correlation coefficient (linear correlation) between two random variables X and Y with mean μ_x and μ_y and standard deviations σ_x and σ_y is defined as

$$\rho = \frac{cov(X,Y)}{\sigma_x \cdot \sigma_y} = \frac{E[(X-\mu_x)(Y-\mu_y)]}{\sigma_x \cdot \sigma_y} \tag{2.1}$$

If X and Y are perfectly correlated then ρ=1. If they are perfectly anti-correlated then ρ=-1 and if X and Y are independent, ρ=0.

Correlation between two variables can also be measured after transforming the variables in the rank domain. Two non-parametric statistics measuring rank correlation are Spearman's rho (ρ) and Kendall's tau (τ).

Spearman's rho is the correlation coefficient of the ranks. The data is transformed into ranks and the correlation coefficient is evaluated using Equation 2.1 for the transformed values. The thesis uses Kendall's τ rank correlation statistics to measure the strength of correlation between tests at different temperatures. The Kendall's τ rank correlation is discussed in detail in the following section.

2.4 Kendall's tau rank correlation

Kendall's τ rank correlation is a non-parametric statistic measuring the association between two or more variables. Kendall's τ coefficient varies between -1 and 1. Kendall's τ rank correlation is independent of the scale and the location of the random variables.

Let {X} = X_1, X_2.., X_n and {Y} =Y_1, Y_2..,Y_n be two sets of n observations. A pair of observations (X_i, Y_i) and (X_j, Y_j) is concordant, if

$$(X_i - X_j)\cdot(Y_i - Y_j) > 0.$$

On the other hand, a pair of measurements (X_i, Y_i) and (X_j,Y_j) is discordant, if

$$(X_i - X_j)\cdot(Y_i - Y_j) < 0.$$

The number of unique pairs between N observations is N(N-1)/2. The Kendall's τ rank correlation coefficient between the two data sets X and Y of each sample size N is calculated as shown.

$$\tau = \frac{\text{(number of concordant pairs} - \text{number of discordant pairs)}}{\text{N(N-1)/2}} \tag{2.2}$$

For perfect correlation $\tau = 1$, for perfect anti-correlation $\tau = -1$, and for independence $\tau = 0$.

Figure 2.1 shows the concordance-discordance plot to explain Kendall's τ. The x-axis and y-axis are the ranks of X and Y respectively. Each point on the graph is an observation of X and Y.

{X}= 32, 20, 41, 1, 88, 63, 50, 74	{Y}= 0.5, 0.3, 0.4, 0.6, 0.7, 0.2, 0.6, 1.0
Rank order(X) = 3, 2, 4, 1, 8, 6, 5, 7	Rank order(Y) = 4, 2, 3, 1, 5, 6, 8, 7

Consider the first two observations in the graph which are colored in green and labeled 'Concordant'. An increase in the rank of the X variable corresponds to an increase in the rank of the Y variable. Hence, the pair of observations are concordant. In general, pairs whose x-intercepts and y-intercepts do not intersect are concordant. In the figure, there are 22 concordant pairs out of 28 possible pairs..

Consider the next pair of observations colored in red and labeled 'Discordant'. An increase in the rank of the X variable corresponds to a decrease in the rank of the Y variable. Hence, the pair of observations are discordant. In general, pairs whose x-intercepts or y-intercepts intersect are discordant pairs. In the figure, there are 6 discordant pairs out of 28 possible pairs.

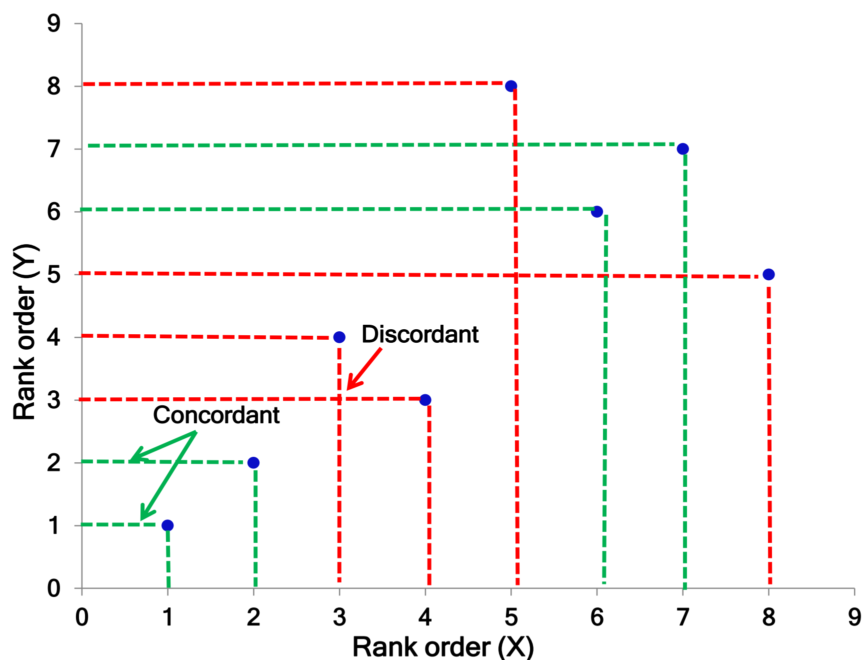

Figure 2.1: Concordance-discordance plot for random variables X, Y

Using Equation 2.2, the value of τ is calculated as shown below:

$$\tau = \frac{(22 - 6)}{8(8\text{-}1)/2}$$

$$\tau = \frac{16}{28}$$

$$\tau = 0.57$$

Now that the idea of rank statistics and Kendall's τ rank correlation have been established, the next three sections discuss the concept of mixed-signal test, the idea of test limits and binning of dies.

2.5 Mixed Signal Test

The thesis uses production data for mixed-signal dies. Mixed-signal dies consist of analog circuitry(operational amplifiers, comparators, filters and PLL) as well as digital circuitry(datapaths, control logic). Several different schemes are required to test mixed signal die [18]. A typical strategy for testing them involves first testing the digital and analog components, followed by some system tests to check the at-speed interaction among the components. The digital parts of the chip would be tested with standard tests such as IDDQ, path-delay fault tests, stuck-at fault tests etc. Analog parts are usually tested for circuit functionality. For example, ADC and DAC are tested for static parameters such as accuracy, resolution, dynamic range, offset, gain, differential nonlinearity (DNL), integral nonlinearity (INL) and dynamic parameters such as SNR, total harmonic distortion etc.

Broadly tests can be classified into two categories:

- Parametric Tests - Parametric tests are executed to decide whether the chip pins meet various specifications for rise and fall times, setup and hold times, low and high voltage thresholds, and low and high current thresholds. Parametric tests are carried out for both the DC and AC components of the die. DC parametric tests include shorts test, opens test, maximum current test, leakage test, output drive current test, and threshold levels test. AC parametric tests include propagation delay test, setup and hold test, functional

speed test, access time test, refresh and pause time test, and rise and fall time test [19].

- Functional Tests - Functional test vectors verify the proper operation of the dies. Input vectors are applied and output responses are recorded and compared to expected responses. Functional tests cover a very high percentage of modeled (e.g., stuck type) faults in logic circuits. Functional tests determine whether the internal digital logic and analog sub-systems in the die behave as intended. The results are usually recorded by pass/fail or go/no-go binary result [19].

2.6 Test Limits

During mixed-signal testing, many different parametric measurements such as current, voltage, frequency etc are taken. The tester compares the measured test responses with limits specified in the datasheet and then labels the die either a 'Pass' or a 'Fail'. These limits are called as test limits. Test limits can either be one-sided (upper/lower) or two-sided. Figure 2.2 depicts one-sided and two-sided limits. For tests with one-sided limits with an upper bound, the die is labeled as a 'Pass' if its test response is greater than the limit as shown in Figure 2.2a. For tests with one-sided limits with a lower bound, the die is labeled as a 'Pass' if its test response is lower than the limit as shown in Figure 2.2b. On the other hand, for tests with two-sided limits with an upper bound, the die is labeled as a 'Pass' if its test response is within the limits as shown in Figure 2.2c.

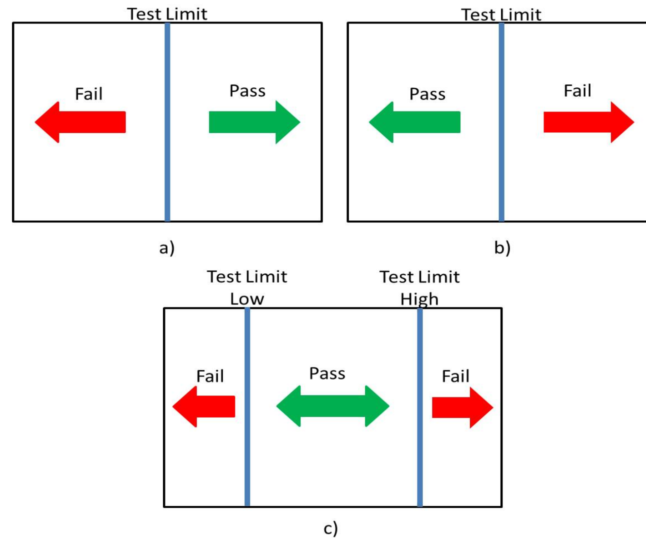

Figure 2.2: a)Single-sided test limits with an upper bound b)Single-sided test limits with a lower bound and c)Two-sided test limits

2.7 Binning dies

Tests in the test program are divided into different groups depending upon their nature or the section of the circuit they are testing. If a die fails in any one test from a group of tests, the die is assigned a failure bin number and testing is terminated. Generally different types of bins in which the dies are divided are open/shorts, IDDQ fails, delay test fails, etc. The number of dies in various bins provides information about the common failure modes of the chip. The order of the groups of tests with a common bin number often depends on the complexity of the test.

The previous three sections dealt with mixed-signal testing and binning of dies. The following two sections discuss the concept of queue and server and the concept of data buffer. The queue and server model is used to depict tests in the adaptive

test flow which is discussed in Chapter 3. The data buffer concept helps understand the training strategies which are discussed in detail in Chapter 4.

2.8 Queue and Server

The queue and server model is used in Chapter 3 to describe the tests in an adaptive test flow. The queue consists of the dies to be tested and the server indicates the test to be executed on the die input from the queue.

A queue is a collection of entities which are stored in the same order which they are received. The addition of new entities occur at the rear terminal position whereas the removal of entities occur at the front terminal position. A queue can be either unbounded or bounded. A bounded queue is limited to a fixed number of entities. On the other hand, an unbounded queue does not have a specific capacity. A queue can also be empty, at which point removing an entity will be impossible until a new entity has been added again.

A server is a device that executes a routine on the entities stored in a queue. The Figure 2.3 depicts a single queue and single server model. The circle represents the server and the rectangular box represents a queue with each slot occupied by an entity. The entities are input to the server one at a time. The routine R_i first services q_1 followed by q_2 and so on. The left blue arrow indicates the flow of entities into the queue. The right green arrow indicates the flow of entities out of the server after being operated by it.

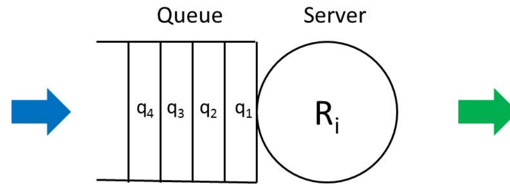

Figure 2.3: The figure depicts the queue and server model. The blue arrow indicates the flow of entities into the queue and the green arrow indicates the flow of entities out of the server.

2.9 Buffer

In the thesis, a buffer is used to store the training data set. A buffer is a region of a physical memory storage used to temporarily store data while it is being moved from one place to another [20]. Typically, the data is stored in a buffer as it is retrieved from an input device or just before it is sent to an output device. However, a buffer may be used when moving data between processes within a system.

A buffer can be implemented in many ways. The training strategies, described in Chapter 4, that use a data buffer to store the training data set are implemented in the FIFO configuration. A FIFO is a method for organizing and manipulating a data buffer where the oldest entry is processed first and removed from the buffer. Figure 2.4 shows the working of the FIFO implementation for a buffer of size 4.

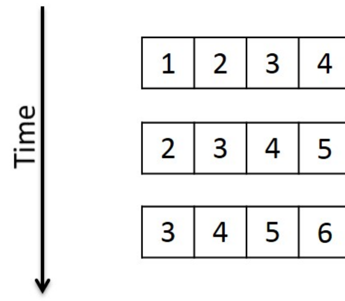

Figure 2.4: Data buffer implementations for FIFO configuration

This concludes the discussion of the Background Concepts. Chapters 3 and 4 discuss the strategy used to study the dependence of machine learning on training. Chapter 3 outlines the simple machine learning technique called the Test Removal Classifier (TRC). Chapter 4 provides the details of the six training strategies used in the study.

Chapter 3

Test Removal Classifier

This chapter introduces a test removal tool called the Test Removal Classifier (TRC) which is a simple machine learning scheme designed to reduce over-testing. The TRC predicts a die's pass/fail test response for a sub-set of tests. The TRC's simple structure means it is easy to compare the TRC model parametrization and the model performance with different training sets. Limiting the TRC prediction of tests to pass/fail simplifies the training. Pass/fail predictions control the flow of die and if correct, do not require a prediction of the die's response to the test.

When the same test is repeated at different environmental conditions, there is an opportunity for test removal. The challenge is to isolate from all possible tests the tests which can safely be eliminated from one environmental condition or the other. Complete removal of a test from one flow is generally easier to manage than selectively picking and choosing the die to be screened at one or the other environmental condition. But, the elimination of a test has an increased risk of test escapes being shipped. The TRC seeks a tradeoff between single test elimination and a sensible scheme for picking and choosing die that should be screened by exploiting the concept of Kendall's τ rank correlation discussed in Chapter 2.

The section below briefly describes the typical two insertion test flow originally used in the study.

3.1 Typical Two Insertion Test Flow

A two insertion test flow is originally used in the study to test the dies at wafer sort. Figure 3.1 displays the configuration of a typical two insertion full test flow. Each insertion is carried out at a different temperature. Tests are executed at different temperatures because different temperatures may expose different failing parts as outliers [21]. Tests in the first insertion (Insert1) are executed at room temperature. Tests in the second insertion (Insert2) are executed at high temperature (85 °C). The insertions are wafer tests and are Stop-on-First-Fail (SoF). In the Stop-on-First-Fail strategy, tests are terminated early once a fail is discovered so as to save test time [4]. A third package test completes the production testing.

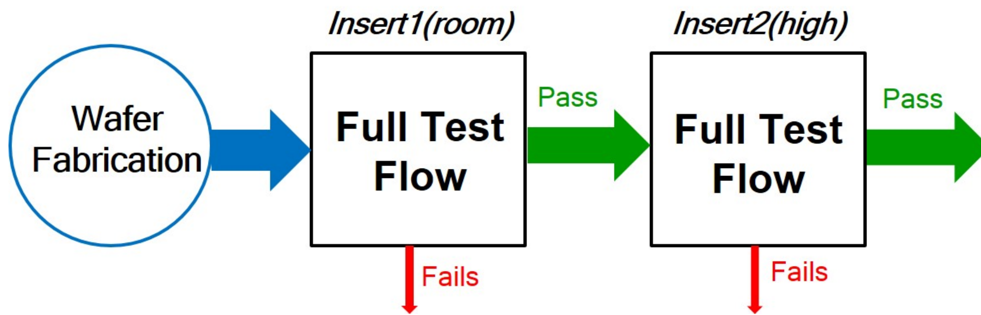

Figure 3.1: Typical two insertion full test flow

The original full test flows consist of defect screens as well as tests to verify datasheet specifications. The full test flows at Insert1 and Insert2 have a subset of tests in common. The commonality of tests between the two insertions is an opportunity for test removal. The next section describes how the potential tests for removal on a die-by-die basis were identified.

3.2 Correlation among Tests

The initial step involved in identification of potential tests to be removed was to analyze the correlation between tests at Insert1 and tests at Insert2 since strongly correlated tests are highly likely to result in consistent test conclusions [11]. The strength of correlation was measured using Kendall's τ rank correlation as discussed in Chapter 2.

Figure 3.2 shows the joint distributions of tests at Insert1 and tests at Insert2 in the performance space organized in a three-by-three matrix. The joint distributions are plotted for a single wafer and each point corresponds to an individual die belonging to the wafer. The X and Y axes of each panel are the test responses for a test at Insert1 and a test at Insert2 respectively. The notation T_{1i} denotes the ith test in the Insert1 test flow and the notation T_{2i} denotes the ith test in the Insert2 test flow. T_{1i},T_{1u} and T_{1j} correspond to the ith, u^{th} and j^{th} tests at Insert1. Similarly, T_{2i},T_{2u} and T_{2j} correspond to the ith, u^{th} and j^{th} tests at Insert2. The Kendall's τ correlation value is reported at the top left corner for each panel in the figure. The diagonal plots from upper left to lower right are for homogeneous test pairs and the off-diagonal plots are for heterogeneous test pairs. Each distribution corresponds to a test pair which includes a test from Insert1 and a test from Insert2.

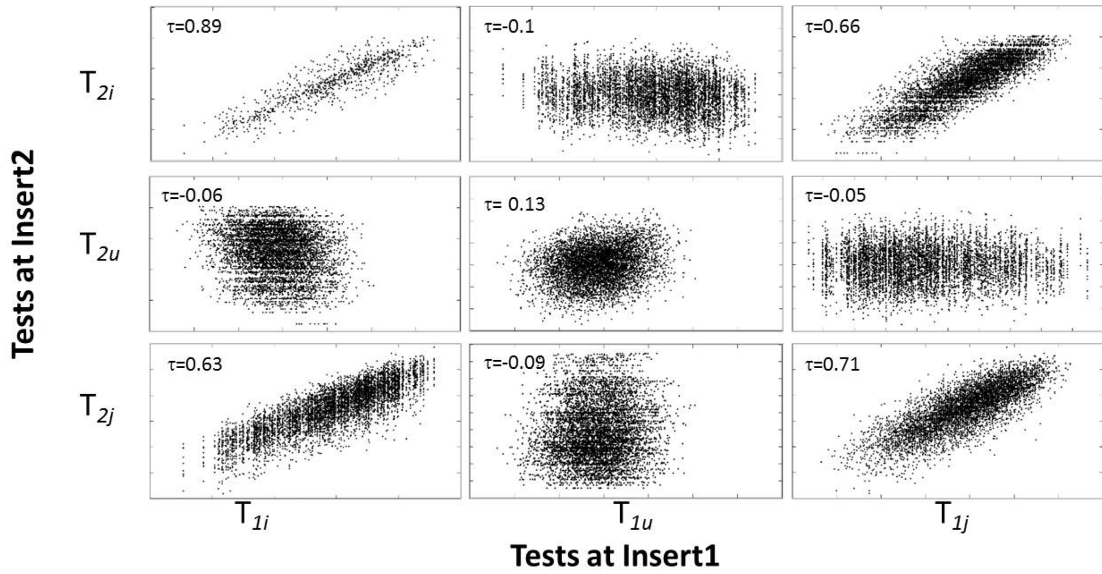

Figure 3.2: Joint distributions of tests between Insert1 and Insert2 in the parameter space. Each joint distribution plot is between a test at Insert1 (x-axis) and a test at Insert2 (y-axis). The diagonal plots from upper left to lower right are for homogeneous test pairs and the off-diagonal plots are for heterogeneous test pairs.

Figure 3.3 shows the joint distributions of the same tests at Insert1 and the same tests at Insert2 in the rank space organized in the same order as Figure 3.3. The rank space plots are included to ensure that the correlation is independent of scale and location of the test responses. The X and Y axes of each panel are the rank of test responses of a test at Insert1 and a test at Insert2. The scales of both axes are equal because they are measured in the rank space.

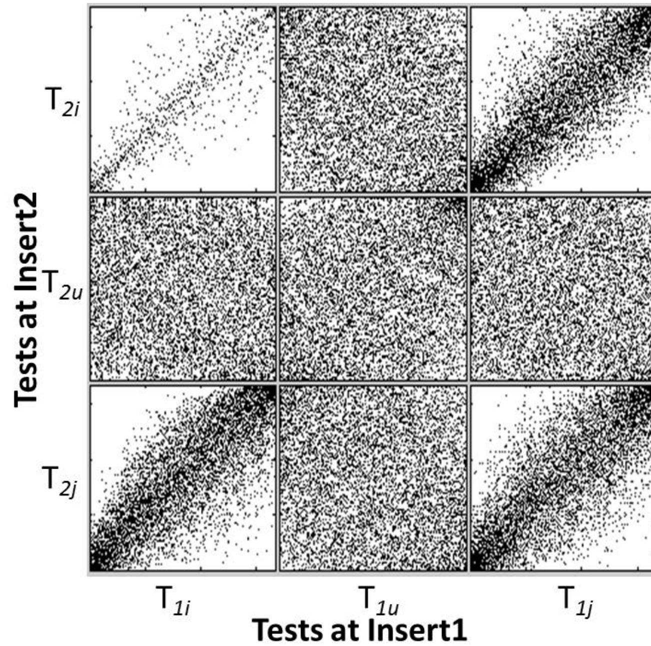

Figure 3.3: Joint distributions of tests between Insert1 and Insert2 in the rank space. Each joint distribution plot is between a test at Insert1 (x-axis) and a test at Insert2 (y-axis). The diagonal plots from upper left to lower right are for homogeneous test pairs and the off-diagonal plots are for heterogeneous test pairs.

From Figures 3.2 and 3.3, the following different types of test pairs were identified:

- Homogeneous test pairs that have a strong positive correlation. For example: (T_{1i},T_{2i}) has a strong positive correlation of $\tau = 0.89$.

- Heterogeneous test pairs that have a strong positive correlation. For example: (T_{1i},T_{2j}) has a strong positive correlation of $\tau = 0.63$.

- Test pairs that are uncorrelated. For example: T_{2u} is uncorrelated with all three tests at Insert2 in the figure.

The strong correlation among tests executed at different temperatures reveals an opportunity to predict binning limits and dynamically adjust test flow on a die-by-die basis. The opportunity is exploited to design an adaptive test flow. The adaptive test flow is specific to tests pairs called matched test pairs which have a strong correlation. Hence, matched test pairs are test pairs that have a Kendall's $|\tau|$ greater than a preset threshold. The absolute value is chosen to take advantage of both positive and negative correlation. Tests such as T_{2u} which are uncorrelated with any test in Insert1 are inapplicable for an adaptive test flow. The method to determine the matched test pairs is discussed in detail in Section 3.3 and Section 3.4.

3.3 Test Data and Matched Test Pairs

In training, the TRC uses per wafer die test response data. Figure 3.4 summarizes the training data model. The test data table on the left consists of the test response of every die for all the tests in the test flow at Insert1 and Insert2. The notation T_{1i} denotes the ith test in the Insert1 test flow and the notation T_{2i} denotes the ith test in the Insert2 test flow. TRC training data is die test response for the complete test flow at both insertions.

Wafer Test Data						
Die	Insert1			Insert2		
	T_{1i}	T_{1u}	T_{1j}	T_{2i}	T_{2u}	T_{2j}
1						
2						
3						
4						

Kendall's τ Rank Correlation			
	T_{1i}	T_{1u}	T_{1j}
T_{2i}	$\tau_{1i,2i}$	$\tau_{1u,2i}$	$\tau_{1j,2i}$
T_{2u}	$\tau_{1i,2u}$	$\tau_{1u,2u}$	$\tau_{1j,2u}$
T_{2j}	$\tau_{1i,2j}$	$\tau_{1u,2j}$	$\tau_{1j,2j}$

Figure 3.4: Test Removal Classifier training data and Kendall's τ Rank Correlation matrix

The training data is processed by test pairs. A test pair (T_{1i},T_{2j}) consists of a test at Insert1 (T_{1i}) and a test at Insert2 (T_{2j}). Kendall's tau rank correlation value is computed for test pairs, $\tau_{1i,2j}$ and tabulated in a correlation matrix as shown in the table on the right. Kendall's tau rank correlation is used to reduce the effects of data outliers on the correlation estimate since it is scale and location invariant. The correlation matrix table is asymmetric about the diagonal as seen in Figures 3.2 and 3.3.

A matched test pair (T_{1i}, T_{2j}) must satisfy three rules. First, both the tests forming the test pair T_{1i} and T_{2j} must have more than 10 unique test responses. This rule eliminates test pairs for which the computed Kendall's τ has a high standard error estimate due to a large number of ties. Second, there must be at least one failing die from either test T_{1i} or T_{2j}. This eliminates test pairs which are only used for data-sheet verification or are ineffective screens. Finally, the Kendall's τ for the test pair must be larger than a pre-determined absolute threshold (example: $|\tau| \geq 0.6$). Absolute value permits matched test pairs to be both positively and negatively correlated. The number of matched test pairs is a sub-set of the total

number of test pairs the TRC can take advantage of for test removal and varies by the training strategy used.

For strategies involving multiple training wafers, there are two different scenarios in which matched test pairs can be selected. They are discussed in detail in the next section.

3.4 Matched Test Pair Selection Scenarios

Three training strategies described in the next chapter use multiple wafers in the training set. There are two different scenarios for the selection of matched test pairs for the training strategies involving multiple wafers - the intersection scenario and the union scenario. In the intersection scenario, test pairs which meet the requirements on all the training wafers are included in the final set of matched test pairs. On the other hand, in the union scenario, test pairs which meet the requirements on any one of the training wafers are included in the final set of matched test pairs. Figure 3.5 shows the two different scenarios to choose the final set of matched test pairs for two wafers in the training set. For strategies involving only one wafer, the training data set for both the scenarios are equal.

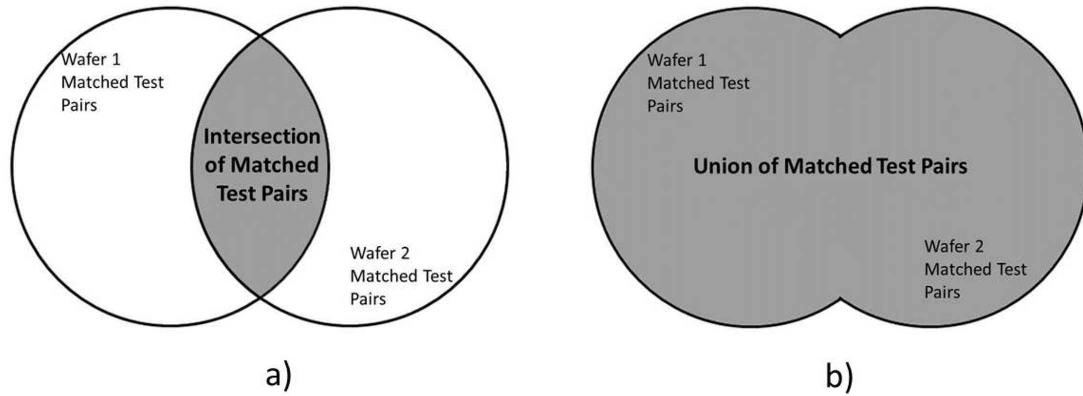

Figure 3.5: Final set of Matched Test Pairs - a) Intersection scenario b) Union scenario

For both the scenarios, matched test pairs that fall into two general classes are considered for test removal. Matched test pairs that fall into Class 1 are homogeneous matched test pairs (T_{1i},T_{2i}). In other words, these test pairs consist of the same test executed at different temperatures. Matched test pairs that fall into Class 2 are heterogeneous test pairs, i.e. $i \neq j$. In addition, Class 2 requires the tests bin failing die in identical soft bins. This constraint means the two tests are evaluating similar parts of the die and because the parts are mixed-signal, the tests have similar input and output.

Once the final set of matched test pairs is determined, the test removal limits for each matched test pair is determined. This is described in the following section.

3.5 Test Removal Limits

This section describes the process involved in setting the test removal limits for the TRC once the candidate matched test pairs are discovered. The method involved

is demonstrated by using the matched test pair (T_{1i},T_{2i}).

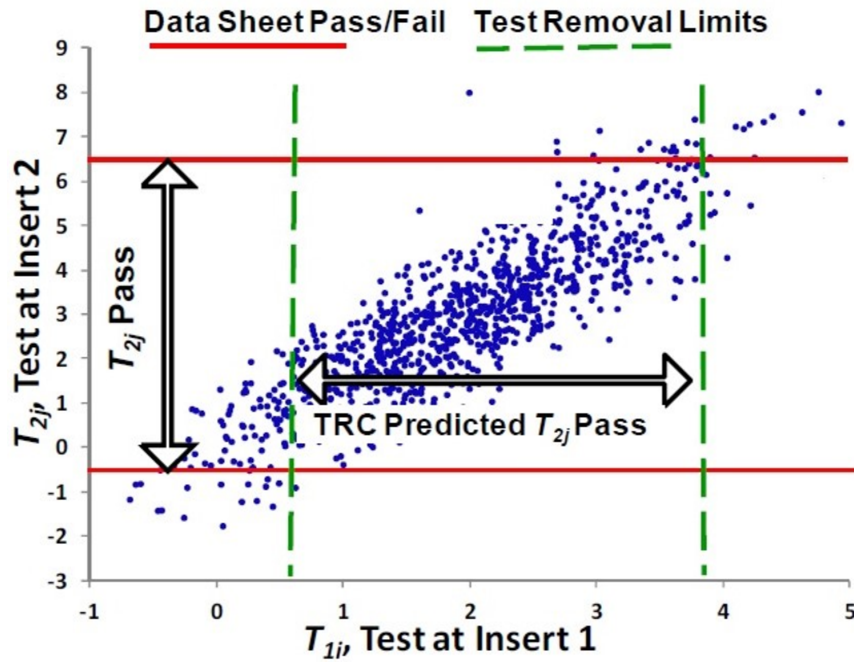

Figure 3.6: Scatter plot of the joint distribution of test responses of die at Insert1 (x-axis) and Insert2 (y-axis) and containing the data-sheet and test removal limits. The points used in this plot are synthetic data points

Figure 3.6 shows the joint distribution of the matched test pair (T_{1i},T_{2j}). Each point (x,y) on this scatterplot represents a single die from the training data set. The x-value represents the die's response to the Insert1 test T_{1i} and the y-value represents the die's response to the Insert2 test T_{2j}. The points used in this plot are synthetic data points. The test removal limits are shown as vertical dashed lines. The datasheet limits are shown as horizontal solid lines. The setting of the test removal limits follows a three step procedure:

1. All passing die in the training set are collected within the T_{2j} data-sheet limits. The die passing the test T_{2j} are in the area spanned by the vertical

arrow in the scatter plot above.

2. The unique test responses for the sample die within the area spanned by the vertical arrows are collected. The unique test responses are considered to make the choice of test removal limits insensitive to the response of a particular die.

3. The T_{1i} test removal limits are the 5% and 95% quantiles of these die.

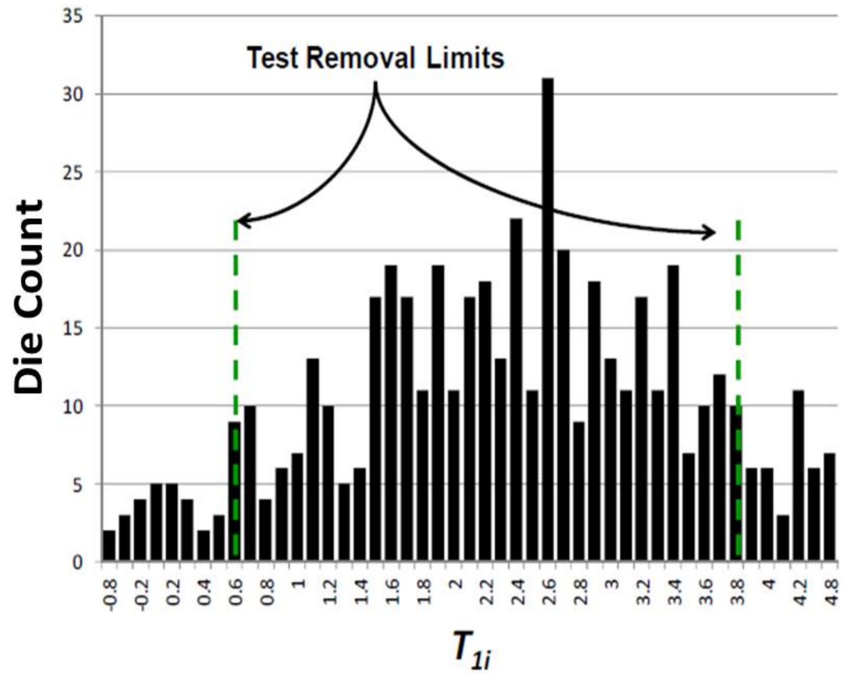

Figure 3.7: Histogram plot with the X-axis representing the unique test response of test T_{1i} and the Y-axis representing the die count. The vertical dashed limits are the 5th and 95th quantile limits for test removal.

Figure 3.7 shows the vertical TRC limits set for T_{1i}. The X-axis is the unique responses of test T_{1i} and the Y-axis is the die count for each unique test response. The dies within the vertical limits exploit the strong correlation between the two

tests to bypass T_{2j} in the Insert2 test flow. Dies outside the vertical limits have to be retested at T_{2j}. This method is followed for every matched test pair. The matched test pairs and the set of limits for each pair complete the process of training.

Once the final set of matched test pairs are discovered and the TRC limits are set for each matched test pair, the adaptive test flow is designed. The adaptive test flow with the TRC is discussed in detail in the next section.

3.6 Adaptive Test Flow

This section briefly describes the adaptive test flow which incorporates the TRC for the matched test pairs. Figure 3.8 shows the two insertion adaptive test flow highlighting a test that uses a TRC (Figure 3.8b) and another test that does not use the TRC (Figure 3.8c). Each test is represented by a queue and server model discussed in detail in Chapter 2. The queue contains the dies scheduled for the corresponding test. The server tests each die input from the queue. The output of each test is a vector of responses for the dies tested by the server. Dies flow from left to right following the arrows.

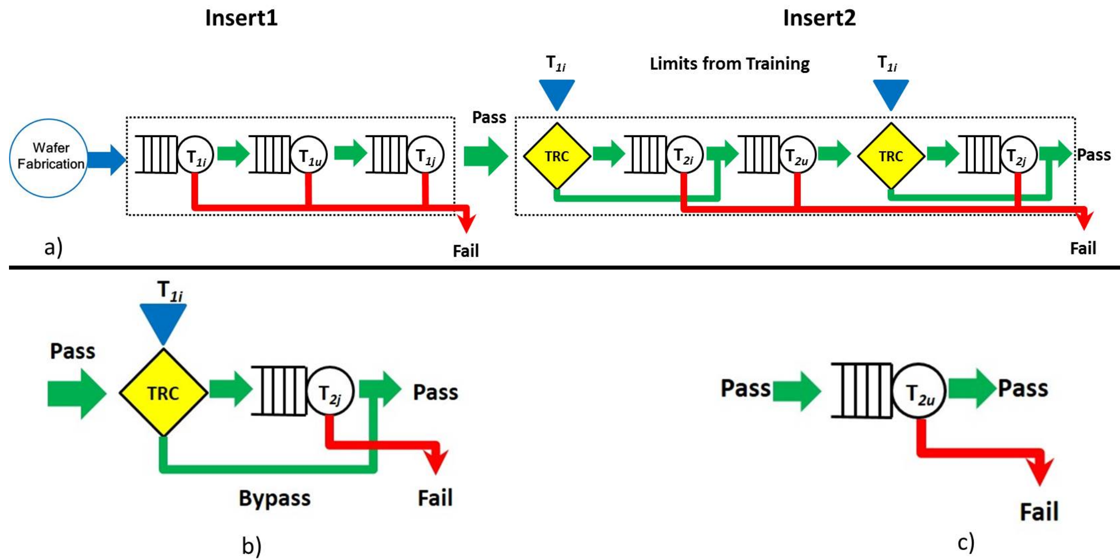

Figure 3.8: a) Test Flow for Insert1 and Insert2 respectively. b) and c) Expanded view of the test flow at Insert2 indicating a test that contains a TRC and a test that does not. The green arrows indicate the flow of passing die and the red arrows indicate the flow of failing die.

The left half of Figure 3.8a is a reduced Insert1 full test flow indicating only three tests T_{1i}, T_{1u} and T_{1j}. Each test consists of a queue with a buffer and a server labeled appropriately. Passing die flow from the buffer of the current test to the buffer of the next test in the test flow. The test flow is a SoF design and hence, failing die are removed from the test flow as indicated by the red arrow.

The right half of Figure 3.8a is a reduced test flow at Insert2 that includes the TRC for tests that are part of the matched test pair set. The TRC is test specific and used only for tests that are part of matched test pairs such as (T_{1i},T_{2j}). The TRC statistical screen schedules the flow at the T_{2j} of each matched test pair (T_{1i},T_{2j}). Figure 3.8b and 3.8c are the expanded views of the Insert2 test flow for passing

die highlighting two tests T_{2j} and T_{2u} respectively. T_{2j} is part of the matched test pair (T_{1i}, T_{2j}) and eligible for a reduced test flow and hence contains a TRC queue scheduler. T_{2u} is not a member of a matched test pair and not eligible for an adaptive test flow and hence does not contain a TRC queue scheduler. The TRC functions by the following three steps:

1. For each matched test pair (T_{1i},T_{2j}), a die's T_{1i} test response is combined with T_{1i}'s training limits to make a pass/fail prediction at T_{2j}. This step is denoted by the TRC diamond.

2. If the TRC predicts a die to pass the test T_{2j}, the bypasses the queue for T_{2j}. This is denoted by the arrow originating from the bottom of the TRC diamond.

3. For each die that bypasses a test T_{2j} due to predicted pass by the TRC, its test time decreases.

The steps are repeated as many times as there are matched test pairs. The location of the TRC in the test flow is dependent on the matched test pairs discovered during training. The limits for each TRC is unique to the matched test pair. For unmatched tests which do not include a TRC, such as T_{2u}, all passing die reaching the queue are tested.

Designing the adaptive test flow involves two steps in training - selection of matched test pairs (T_{1i},T_{2j}) and setting the test removal limits. The TRC is trained by performing a full test flow on a training set as discussed in Section 1.4. The next chapter discusses seven different strategies to select the training set.

Chapter 4

Training Strategies

This chapter describes the seven training strategies used to train the Test Removal Classifier (TRC). They span the range of likely strategies in production environments. Each strategy is wafer based and uses data from Stop-on-First-Fail (SoF) production lots. The wafers that contribute to the training set are tested with the Full Test Flow (FTF) at both insertions. The training data is stored in a data buffer discussed in detail in Chapter 2. The first three training strategies include effects from within the lot under test (LUT). The final two training strategies are included for contrast and comparison of the first four training strategies. Training strategies which include multiple wafers in the training data set are implemented in both the intersection and the union scenarios as discussed in the previous chapter.

4.1 Window Training Strategy

Windowing is a common strategy. A sample window acts as a low pass filter on the incoming material to dampen the effects of any single value. Used across the production lots, the window training strategy is a low pass filter for broad product/process effects. In other words, it acts as a low pass filter to dampen the effects of a rogue training wafer. The effects in the context of this strategy refer to the consistency in the choice of matched test pairs. This strategy selects test pairs that are consistently matched within the training set for test removal by the TRC.

The window training strategy consists of a data buffer which holds the training data of a set of N training wafers, where N is the buffer size. The buffer can be implemented in a number of ways. The FIFO implementation is chosen to dampen the process effects from the most recent production lots. The buffer consists of single training wafers each from (N-1) previously tested lots and a training wafer from the Lot under Test (LUT). For a new LUT, the oldest training wafer in the buffer is replaced with a single training wafer from the current LUT. The revised buffer is reanalyzed for matched test pairs and TRC limits. The buffer size in this thesis is set to N=5 wafers.

Figure 4.1 shows the occupancy of the training set buffer when the LUT is changed from Lot 5 to Lot 6 and finally to Lot 7. It is assumed that testing is completed for Lots 1-4. In this figure, L1 W1 refers to the first fully tested wafer from Lot 1; L2 W1 refers to the first fully tested wafer from Lot 2 and so on. The wafers are highlighted in the buffer - the darkest shade being the wafer that was most recently included in the buffer and the lightest shade being the wafer that was least recently included in the wafer. The arrows indicate the direction of flow of the wafers' test data into and out of the buffer respectively. The highlighted wafer of the LUT refers to the fully tested wafer that is in the buffer. For a new LUT, the least recently included wafer is cleared from the buffer and replaced with a fully tested wafer from the LUT. For example, when test flow changes from Lot 5 to Lot 6, the training wafer from Lot 1, L1 W1, is replaced with a fully tested wafer from Lot 6, L6 W1. The same process repeats while transitioning from Lot 6 to Lot 7 and so on.

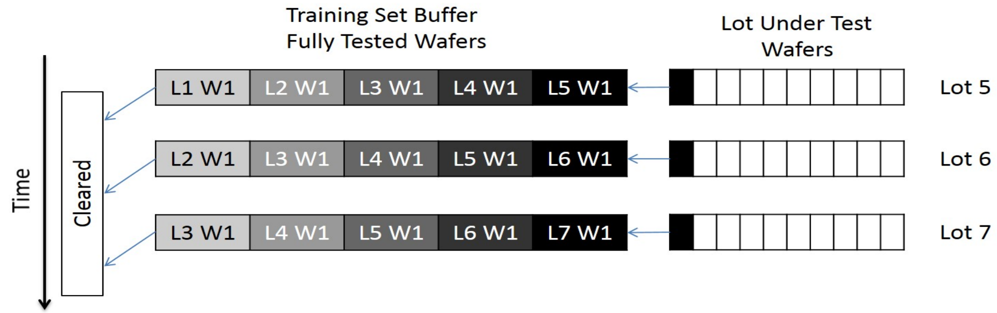

Figure 4.1: Timeline showing the occupancy of the training set buffer for the window training strategy as test flow moves from sample lot 5 to sample lot 6 and finally to sample lot 7. The highlighted wafers are the fully tested wafers used in the FIFO buffer.

This strategy designed for the intersection scenario is similar to the dynamic part average testing (DPAT) [7]. Dies with parameters that are within specification but not within the normal distribution are inked out by the DPAT. In the same way, test pairs which are matched for all the wafers within the buffer are only considered by the TRC for test removal by this strategy. If a test pair is not matched for even one of the wafers in the buffer, it is “inked” out.

This strategy can also be considered similar to stratified random sampling [15] where the population of wafers is divided into strata called lots and the training set consists of one random training wafer selected from each stratum. The concept of stratified random sampling is constant to a training wafer and not a random sample from the lot.

4.2 Accumulate Training Strategy

The accumulate training strategy is a different version of the window training strategy. For this strategy, the first N fully tested wafers from the current LUT accumulate in the buffer. Once the buffer is full, the training data is analysed for matched test pairs and test removal limits for the TRC. To ease comparison, the size of the buffer is set to N=5.

Figure 4.2 shows the occupancy of the training set buffer when the LUT is changed from Lot 5 to Lot 6 and finally to Lot 7. It is assumed that testing is completed for Lots 1-4. In this figure, L5 W1 refers to the first fully tested wafer from Lot 5; L5 W2 refers to the second fully tested wafer from Lot 5 and so on. The double arrow indicates the flow of the highlighted wafers' test data from the LUT into and out of the buffer respectively. This figure clearly depicts the use of multiple wafers from the LUT in contrast to a single wafer used in the window training strategy. Unlike the window strategy, all the wafers in the buffer are highlighted by the same shade of gray because they are included from the same LUT. Once Lot 5 has been tested, the buffer is cleared and populated with fully tested wafers from Lot 6 as shown. The same process repeats while transitioning from Lot 6 to Lot 7 and so on.

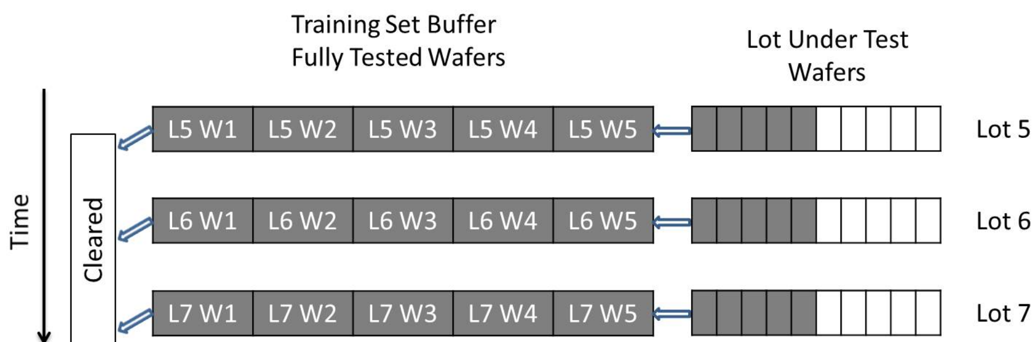

Figure 4.2: Timeline showing the occupancy of the training set buffer for the accumulate training strategy as test flow moves from sample lot 5 to sample lot 6 and finally to sample lot 7. The highlighted wafers are the fully tested wafers used in the FIFO buffer.

This strategy differs from the window training strategy in two aspects. First, the buffer does not consist of any training wafers from the previously tested lots. Second, for a new LUT, the buffer is cleared and filled with training data of N fully tested wafers from the current LUT. Hence, this strategy requires more testing of the LUT than the window training strategy. The accumulate training strategy can be considered to be low-pass filter with the emphasis on dampening the effect of a rogue training wafer within the LUT.

Unlike the window training strategy, this strategy is not similar to stratified random sampling for the following reason. Considering the population of dies from the LUT to be divided into strata called wafers, all the dies from the first five wafers of the lot are chosen and no dies from the remaining wafers of the lot are chosen. This is a violation of the rule of stratified random sampling where equal random samples must be chosen from all the strata.

4.3 Single LUT Wafer Training Strategy

The single LUT wafer training strategy uses a single wafer from the lot under test. For a new LUT, the training wafer in the buffer is replaced with a fully tested wafer from the current LUT. The single LUT wafer training strategy can be considered to be a special case of the accumulate training strategy with the buffer size set to N=1. With one wafer from the LUT, this strategy also overlaps with the window training strategy without the low-pass filtering from the previous tested lots. Hence, this strategy provides an assessment of the benefits of the addition of a low-pass filter within the LUT or across lots.

Figure 4.3 shows the occupancy of the training set buffer when the LUT is changed from Lot 5 to Lot 6 and finally to Lot 7. It is assumed that testing is completed for Lots 1-4. The figure shows that while moving to a new LUT, the buffer is cleared and filled with a fully tested wafer from the LUT.

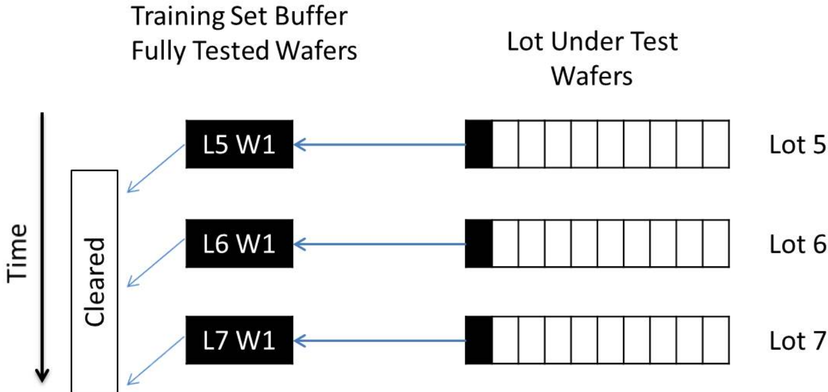

Figure 4.3: Timeline showing the occupancy of the training set buffer for the single LUT wafer training strategy as test flow moves from sample lot 5 to sample lot 6 and finally to sample lot 7. The highlighted wafers are the fully tested wafers used in the FIFO buffer.

4.4 Dual Lot and Single Lot Training Strategy

The dual lot training strategy uses two training wafers, one from each of two previously tested lots. The single lot training strategy on the other hand uses a training wafer from only one previously tested lot. Figure 4.4a and 4.4b show the occupancy of the training set buffer for the dual lot training strategy and single lot training strategy respectively when the LUT is changed from Lot 5 to Lot 6 and finally to Lot 7. It is assumed that testing is completed for Lots 1-4. From the figure, it is observed that these strategies do not contain any wafers from the current LUT and differ from the window and the accumulate training strategies.

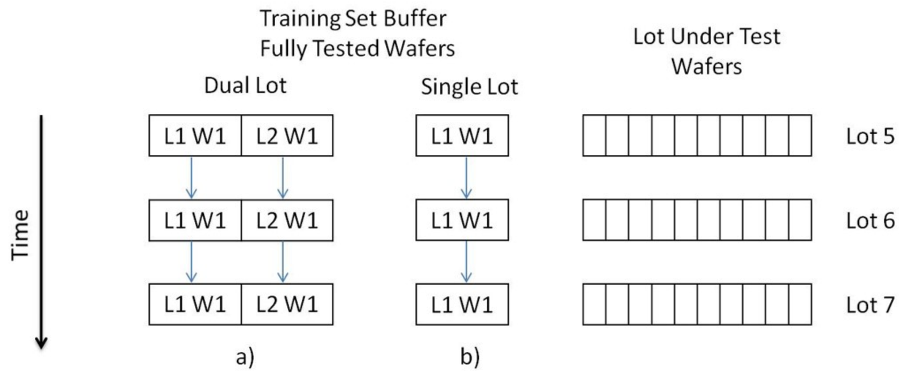

Figure 4.4: Timeline showing the occupancy of the training set buffer for the a) dual lot training strategy and b) single lot training strategy as test flow moves from sample lot 5 to sample lot 6 and finally to sample lot 7.

These two strategies are based on the common practice of test characterization. Test characterization assumes that a fixed sample of some appropriate size reflects the statistics of the product's response to process variation. Many machine learning test data tools are based on selecting a characterization data set for training. For such strategies to function, two requirements need to be met. First, dies

from other lots are suitable representatives of the dies in the LUT. Second, rogue events or process excursions in the training data or the LUT are captured by other components of the test flow.

4.5 Special Case Training Strategies

The special case training strategies are the random training strategy and the zero training strategy. The zero training strategy takes a contrarian view of rank correlation. The zero training strategy randomly selects without replacement test pairs that are distinctly not matched in rank correlation. In other words, this strategy selects test pairs with rank correlation $|tau| \leq 0.05$. The random training strategy, on the other hand, ignores rank correlation and randomly selects without replacement a predetermined number of test pairs from all possible test pairs in the pool. Unlike the zero training strategy, the random training strategy includes test pairs across the entire range of rank correlation represented. Randomly selected test pairs may also include some of the matched test pairs (For example $|\tau_{ij}| \geq 0.6$) used in the other strategies. Typically, the average rank correlation for the random training strategy of all test pairs used is $|\tau_{avg}| \geq 0.1$.

Random and zero test pairs are input to the TRC and it follows the same procedure as do the previously described strategies. These two strategies are included for contrast and comparison of the more plausible training strategies described earlier.

This concludes the discussion of the different training strategies. To summarize, training involves the following steps:

- Select the training set depending on the choice of training strategy as discussed in this chapter.

- Select either the intersection or the union scenario for the training strategy as discussed in Section 3.3. For strategies involving only one training wafers, both the scenarios correspond to the same training set.

- Select the final set of matched test pairs from the data set based on the preset rank threshold as discussed in Section 3.4. For example, choose final set of matched test pairs which satisfy the condition $|\tau| \geq 0.6$.

- For each matched test pair, set the test removal limits as discussed in Section 3.5.

- Install the TRC with the test removal limits for each matched test pair as shown in Figure 3.8.

The next chapter compares the training strategies based on the performance of the TRC. The performance is evaluated using figures of merit which are discussed in detail in the next chapter.

Chapter 5

Results using Production Data

This chapter discusses the figures-of-merit (FOM) used in the study followed by the results obtained when each of the seven training strategies, described in the previous chapter, were used to train the TRC, discussed in Chapter 3. The results are expressed in terms of the FOM described in detail in Section 5.2. But first, the section below discusses the classification of die in order to clearly establish the figures of merit.

5.1 Classification of Die

Figure 5.1 shows a scatter plot containing the joint distribution of test responses T_{1i} and T_{2j} at Insert1 and Insert2 respectively. The scatter plot is divided into 9 sub-areas, defined by the Test Removal Limits (dashed lines) and the Data-sheet Pass/Fail limits (solid lines). The die that fall into the area between the solid lines pass the test T_{2j} at Insert2. The die that fall into the area between dashed lines are predicted to pass test T_{2j} at Insert2 by the TRC.

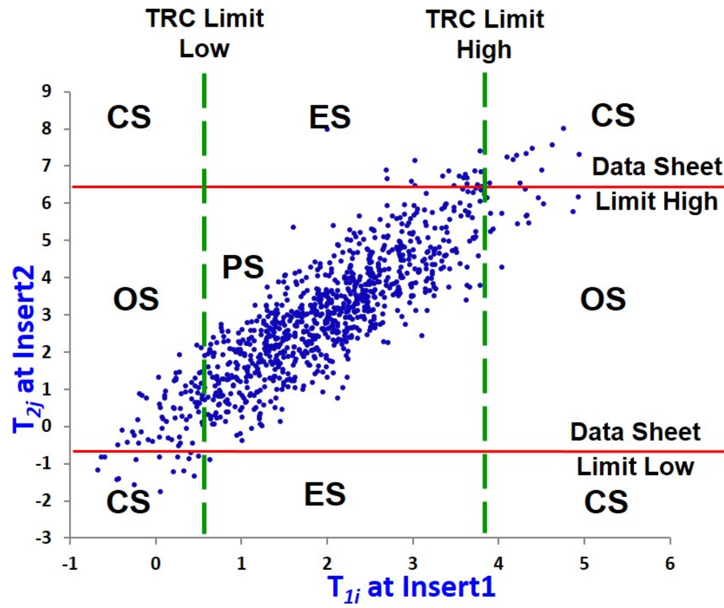

Figure 5.1: Classification of die based on the TRC limits(dashed lines) and the Datasheet limits(solid lines)

The sub-area labeled PS consists of Pass Screen dies. These dies are predicted to be passes at T_{2j} and if they were tested, they would pass T_{2j}. On the other hand, the dies that fall into the sub-area labeled ES, called the Escape Screen dies, are predicted to pass at T_{2j} but if they were tested, they would fail at T_{2j}. The escape screens are test escapes.

The sub-area labeled CS consists of Correct Screen dies. These dies are predicted to fail at T_{2j} and after testing at T_{2j}, they indeed fail. The dies in the sub-area labeled OS, called the Over Screen dies, are predicted to pass at T_{2j} but once they are tested, these dies pass at T_{2j}. These dies contribute to the yield loss and are missed opportunities for test removal.

5.2 Figures of Merit

This section describes the figures of merit used in the thesis. The figures of merit can be expressed in terms of three indicator variables R_{jd}, B_{jd} and Z_{jd}.

- R_{jd} describes the per die test outcomes in the production. $R_{jd} = 1$ indicates that the die d passes test T_{2j} and $R_{jd} = 0$ indicates that the die d fails test T_{2j}.
- B_{jd} describes the results of TRC prediction and scheduling for the matched test pairs. $B_{jd} = 1$ indicates that die d was not inserted in the test queue of test T_{2j} and $B_{jd} = 0$ indicates that die d was inserted into the test queue of test T_{2j}.
- Z_{jd} is an indicator variable to select the dies tested in production by the test T_{2j} in the matched test pair. $Z_{jd} = 1$ indicates that die d has a pass/fail record for test T_{2j} in the production data. On the other hand, $Z_{jd} = 0$ indicates that the die d is not tested by test T_{2j} because this die has failed and been removed by a test earlier in the test flow. Recall that the production data is a Stop-on-First-Fail.

The FOM use four sums for the four types of die in the regions PS, OS, CS and ES as described in the previous section.

Pass Screen (PS) This sum counts the die which are not scheduled by test T_{2j} and which the die would have passed had test T_{2j} been executed. This count reflects the TRC correctly predicting the T_{2j} passes.

Over Screen (OS) This sum counts the dies which are queued by the TRC to be tested by test T_{2j} and which the die passes the scheduled test. This count reflects the TRC incorrectly predicting T_{2j} fail but recovered from yield-loss because the dies pass T_{2j} when executed.

Correct Screen (CS) - This sum counts the die the TRC schedules the dies for test T_{2j} and which failed T_{2j} when executed. This count reflects the TRC correctly predicting T_{2j} fail.

Escape Screen (ES) This sum counts the die which are not scheduled by the TRC to be tested by T_{2j} but would have failed had the test been executed. This count reflects the TRC incorrectly predicting T_{2j} pass and contributes to a failure at a later test insertion which would be more expensive to test.

The equations below express these sums in terms of the three indicator variables as described earlier. The counts are summed over the total number of tests j and total number of die d.

$$PS = \sum_{j} \sum_{d} B_{jd} \cdot R_{jd} \cdot Z_{jd} \tag{5.1}$$

$$OS = \sum_{j} \sum_{d} (1 - B_{jd}) \cdot R_{jd} \cdot Z_{jd} \tag{5.2}$$

$$CS = \sum_{j} \sum_{d} (1 - B_{jd}) \cdot (1 - R_{jd}) \cdot Z_{jd} \tag{5.3}$$

$$ES = \sum_{j} \sum_{d} B_{jd} \cdot (1 - R_{jd}) \cdot Z_{jd} \tag{5.4}$$

The FOM are used to compare the results from different training strategies on the TRC use of matched test pairs. The FOM are mainly of concern to two engineers

involved in the test flow - product engineer and the test engineer.

5.2.1 Product Engineering Figures of Merit

The FOM described in this section are of primary concern to the product engineer. The main goal of the product engineer is to ensure that there are no test escapes transferred to the downstream insertion because it would be more expensive to test and filter them out. The following three FOM are of interest to the product engineer.

Correct Screen PV

The Correct Screen PV measures the predictive value of TRC fails. This FOM is the fraction of failing dies which would be screened by the TRC to the total number of failing dies which should be screened.

$$CorrectScreenPV = \frac{CS}{CS + ES} \tag{5.5}$$

High values of this FOM reflect the success rate of the TRC in correctly identifying the failing dies. On the other hand, low values of this FOM indicate low success rate of the TRC in identifying failing dies and hence, a more expensive downstream test.

Pass Screen PV

The Pass Screen PV measures the predictive value of the TRC passes. This FOM is the fraction of dies which are correctly predicted to be passes by the TRC to

the total number of passing dies.

$$PassScreenPV = \frac{PS}{PS + OS} \tag{5.6}$$

High values of this FOM indicates that the die which TRC bypasses the test T_{2j} queue have a high probability of passing the test had the test been executed. This reflects the success rate of the TRC to obtain test removal without adding cost to the downstream package test.

Escape Screen Best Estimate - 50% UCL

Upper confidence limits are computed to provide test engineers with a broader sense of TRC performance based on the sample size. The TRC interpretation of the UCL is as follows. Assuming the observed value correct screen PV is the population average, and then with subsequent use, fifty out of every 100 uses the test engineer observes a sample correct screen PV greater than the reported 50% UCL. The 50% confidence level reflects the Best Estimate of population average. The UCL Fraction is reported in terms of Devices per million (DPM). The UCL is computed using standard methods. UCL depends on the selected confidence level and the sample size. In the tables reported in Section 5.3, the sample size varies with the strategy used because some LUT die are used to train the TRC.

5.2.2 Test Engineering Figures of Merit

The other engineer involved in the adaptive test flow is the test engineer. The figures of merit described in this section are of primary concern to the test engineer. The main goal of the test engineer is in the reduction of the test time.

Matched Test Pairs Count

This figure of merit indicates the number of matched test pairs (T_{1i}, T_{2j}) discovered during training. The matched test pairs discovered provides a simple metric for gauging the effects of different training strategies. Recall that the matched test pairs discovered is restricted to homogeneous test pairs (Class 1 matched test pairs) and heterogeneous test pairs that fall into identical soft bins (Class 2 matched test pairs). Matched test pairs that do not fall into these categories are not considered.

Average Tests Bypassed

The average tests removed is computed by multiplying the matched test pairs discovered by the average number of test queues a die bypasses. This figure of merit tracks the success of the matched test pair selection process. Averages close to the matched test pairs discovered means that the training provided the TRC a set of matched test pairs which are frequently used in test removal.

5.3 Production Data Set

The TRC was evaluated using a production test data set for a small mixed-signal design. The data set is the output from wafer production testing. The production tests included a room temperature insertion (Insert1) and a high temperature insertion (Insert2). In the evaluation, single site testing is assumed. The test times for each test were not available and each test is assumed to take unit time to be executed per die.

Wafers from six production lots were included in the data and a total of ~750K dies. The number of wafers in each lot varied from a low of 9 to a maximum of

22. For the study, the lots were numbered from 1 to 6. Lot 6 was selected as the Lot Under Test (LUT) for TRC performance evaluation for the different training strategies. A total of 59797 dies were available for evaluation in the LUT. For three of the training strategies, data from the LUT used in training varied from a minimum of one wafer to a maximum of five wafers. The table reports the number of dies (and corresponding wafers) used in training and in evaluation from the LUT respectively.

Training Strategy	Used in Training LUT Wafer(Die)	Used in Evaluation LUT Wafer(Die)
Window	6688(1)	53109(8)
Accumulate	33382(5)	26415(4)
Single LUT	6705(1)	53092(8)
Dual Lot	0(0)	59797(9)
Single Lot	0(0)	59797(9)

Table 5.1: Dies (and corresponding wafers) used in training and in evaluation respectively from the LUT for each of the training strategies

Each row of Table 5.1 corresponds to a training strategy. The first column reports the number of dies used in training from the LUT and the second column reports the number of dies used in evaluation from the LUT. Among these strategies, the accumulate training strategy has the largest training set from the LUT which corresponds to the smallest evaluation data set from the LUT. This shows that there is a large test time overhead involved in the training phase of this training strategy. On the other hand, the dual and single lot training strategies do not use any dies from the LUT in the training data set. Hence, all 9 wafers from the LUT can take advantage of the adaptive test flow and contribute to a reduced test time. The special case training strategies are not included in this table because they do

not involve any training.

5.4 Results

This section reports the TRC performance results for each of the training strategies. Tables 5.2 and 5.3 report the test engineering FOM and product engineering FOM for the LUT respectively. The results are reported for both the intersection and the union scenarios discussed in Chapter 3. Each row in these tables corresponds to a training strategy sorted in the same order in each table for easy comparison.

Training Strategy	Intersection		Union	
	Matched Test Pair Count	Avg. Tests Bypassed (per die)	Matched Test Pair Count	Avg. Tests Bypassed (per die)
Window	43	36	136	121
Accumulate	58	46	126	107
Single LUT	79	64	79	64
Dual Lot	63	53	87	74
Single Lot	73	52	73	52
Random*	63	52	101	69
Zero*	63	54	101	69

Table 5.2: Test engineering FOM for the intersection and the union scenarios respectively

Table 5.2 reports the test engineering FOM for the training strategies. The first two columns are the results for the intersection scenario and the final two columns are the results for the union scenario. The special case training strategies are denoted by an asterisk because they do not involve training. In such a case, the matched test pair count is chosen to be the average of the matched test pair counts of the other strategies.

The variation in the matched test pairs for each strategy signifies that the performance of the TRC is dependent on the choice of training. This variation is observed to be independent of the choice of scenario. The matched test pair count in the union scenario is higher when compared to the intersection scenario for strategies that use multiple training wafers. This is due to the fact that the intersection scenario chooses test pairs that are matched for all the training wafers in the training set whereas the union scenario chooses test pairs that are matched for any one of the training wafers in the training set. However, for strategies that use only a single training wafer, such as single lot and single LUT, the number of matched test pairs are the same for both the scenarios.

The matched test pair counts for the window and accumulate training strategies is lower than the remaining strategies in the intersection scenario. The fact that the average tests bypassed count is close to the matched test pair count suggests that even though a smaller subset of matched test pairs are discovered during training, they are effectively utilized in the adaptive test flow. The results also show that the utilization of matched test pairs for the special case training strategies is comparable to the utilization of matched test pairs for the single and dual lot training strategies. This result questions the need for rank correlation in training which can be clarified by Table 5.3.

Table 5.3 reports the product engineering FOM results for the training strategies. The first three columns report the FOM results for the intersection scenario whereas the last three columns report the FOM results for the union scenario.

Training Strategy	Intersection			Union		
	Correct Screen PV	Pass Screen PV	Escape Screen Best Estimate(DPM)	Correct Screen PV	Pass Screen PV	Escape Screen Best Estimate(DPM)
Window	0.95	0.84	50	0.91	0.89	88
Accumulate	1.00	0.78	26	0.95	0.85	64
Single LUT	0.89	0.82	107	0.89	0.82	107
Dual Lot	0.94	0.84	61	0.94	0.85	61
Single Lot	0.94	0.77	61	0.94	0.77	61
Random	0.88	0.68	95	0.71	0.68	245
Zero	0.83	0.68	145	0.71	0.68	262

Table 5.3: Product engineering FOM for the intersection and the union scenarios respectively

The effect of matching with the union or the intersection scenario can be understood by the value of the correct screen PV and the DPM levels in Table 5.3. The decrease in correct screen PV which translates to increase in escape screens would add test costs by forcing such die to be screened at a later and more expensive insertion or more critically released to the customer and lowering product quality. Strategies that use one or more wafers from the LUT have higher correct screen PV when compared to the other strategies and this is consistent in either scenario - intersection or union. The accumulate strategy has a correct screen PV value of '1.00'. This is due to the large training set from the LUT as reported in Table 5.1. The window strategy which uses only one wafer from the LUT in the training set has a correct screen pv value comparable to that of the accumulate strategy.

The importance of including the UCL is best seen by the comparison of strategies which use rank correlation with strategies that do not use rank correlation in the selection of matched test pairs. For strategies using rank correlation, the UCL is significantly lower than the special case training strategies which do not use rank correlation. Hence, even though the utilization of matched test pairs for special

case training strategies is comparable to dual and single lot strategies as seen in Table 5.2, the use of rank correlation in training ensures that there is a better judgement of the die that can bypass the matched test pairs in the adaptive test flow.

The UCL of the single LUT training strategy is higher than the other two training strategies that include the effects of the LUT because the single LUT training strategy does not include the low pass effects of either the LUT, as in the case of the accumulate training strategy, or the low pass effects of the previous lots, as in the case of the window training strategy.

Tables 5.4 and 5.5 show the escape screen die generated per wafer by the TRC for the intersection and the union scenario respectively. For comparison, the random strategy reveals that there is an escape screen die on almost every wafer and the utilization of the matched test pairs is considerably lower.

		Random Strategy			Window Strategy		
Wafer ID	No. of Die	Correct Screen	Escape Screen	Avg. Test Bypassed	Correct Screen	Escape Screen	Avg. Test Bypassed
5	6705	4	0	51	6	0	39
6	6677	4	1	51	5	0	38
7	6660	4	0	53	4	0	38
8	6659	4	1	55	7	0	32
9	6641	4	0	53	4	0	35
11	6616	4	0	55	6	0	32
12	6653	4	1	50	5	1	36
13	6498	4	1	53	4	1	38
Totals	53109	32	4	52	41	2	36

Table 5.4: TRC performance comparison by wafer for the random and window training strategies - intersection scenario

		Random Strategy			Window Strategy		
Wafer ID	No. of Die	Correct Screen	Escape Screen	Avg. Test Bypassed	Correct Screen	Escape Screen	Avg. Test Bypassed
5	6705	3	3	86	6	0	129
6	6677	4	1	87	5	0	126
7	6660	3	1	88	4	0	125
8	6659	4	4	90	7	1	118
9	6641	4	0	88	4	0	123
11	6616	4	2	90	6	0	119
12	6653	4	1	82	5	1	123
13	6498	4	1	87	4	2	125
Totals	53109	30	13	87	41	4	123

Table 5.5: TRC performance comparison by wafer for the random and window training strategies - union scenario

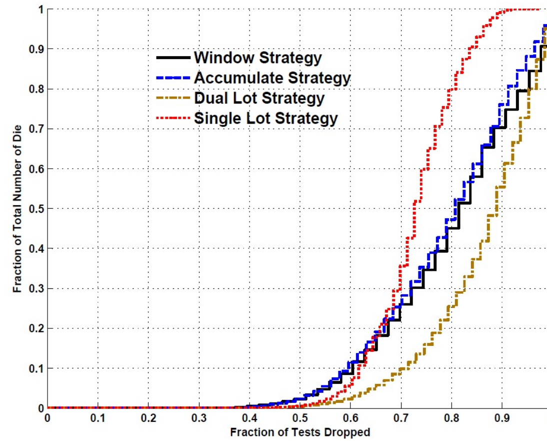

Figure 5.2: Training tests fraction of die cumulative probability vs fraction of TRC bypassed tests for the Intersection scenario

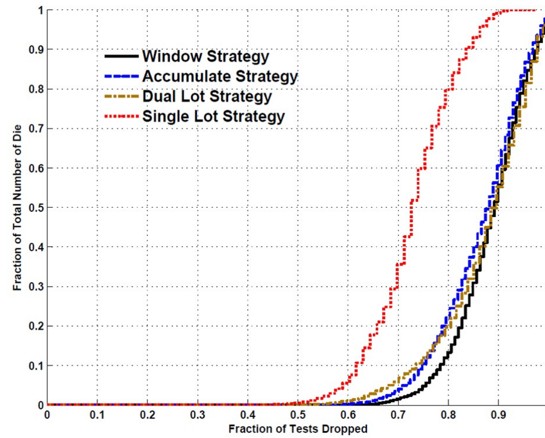

Figure 5.3: Training tests fraction of die cumulative probability vs fraction of TRC bypassed tests for the Union scenario

Figures 5.2 to 5.5 compare the utilization of the test pairs for the intersection and union scenarios respectively. The data is plotted for four training strategies as indicated. The training strategies are color coded for ease of comparison. The difference between the window and single lot training strategies points again the need to include the effects of the current LUT. The single lot strategy selected the matched test pairs from a single training wafer of an entirely different lot. However, as the data indicates, the matched test pairs are used inefficiently. The effect is more clearly seen in Figures 5.4 and 5.5.

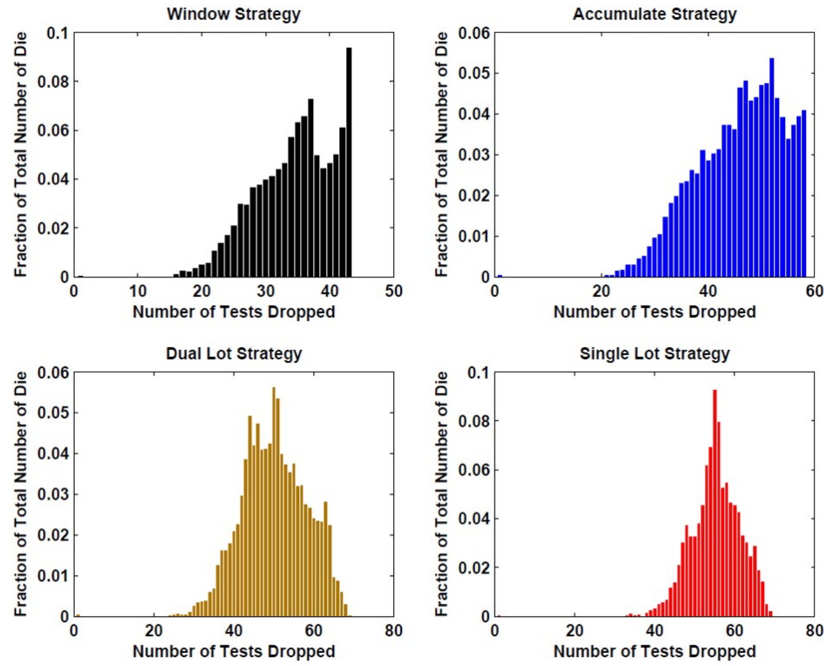

Figure 5.4: PDF of the Fraction of die bypassed by the TRC for tests T_{2j} for the Intersection scenario

Figures 5.4 and 5.5 show the PDF of matched test pair utilization for each training strategy for the intersection and the union scenarios respectively. Note the difference in the right tails of the PDFs. For the window and the accumulate training strategies, the fraction of the matched test pairs grows and hits a hard wall of the maximum number. On the other hand, for the single and dual lot training strategies, the matched test pair fraction peaks well short of the maximum and trails off to an insignificant number of die at the maximum. This shows that strategies that include filtering schemes have a better utilization of matched test pairs.

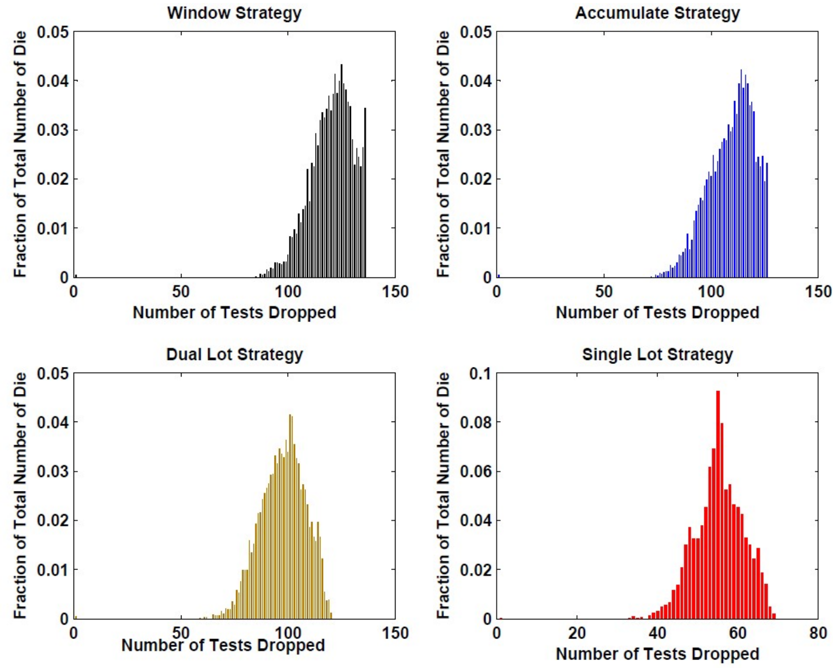

Figure 5.5: PDF of the Fraction of die bypassed by the TRC for tests T_{2j} for the Union scenario

5.5 Cross Validation of the Training Strategies

This section reports the TRC matched test pair selection and use for a cross validation study of the training strategies. For each run of the cross validation, a different set of training wafers are selected as per the rules set by the training strategy. This method is followed for each of the five training strategies.

Figure 5.6 summarizes the TRC selection and use of matched test pairs. The intervals report the range of maximum test pairs and their use for different selections of training wafers. The intervals in red (circle) are reported for the union scenario and the intervals in black (square) are reported for the intersection scenario. Since single lot and single LUT wafer training strategies have a single wafer in the

training set, they are indicated by intervals in blue (triangle).

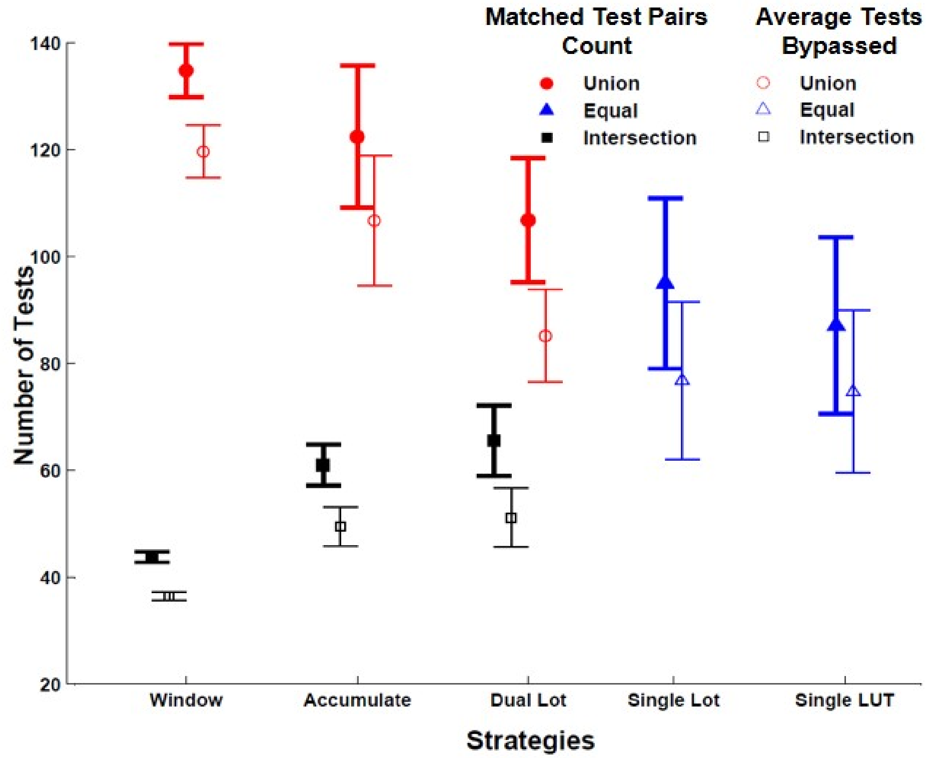

Figure 5.6: TRC test pair training and bypass performance per strategy

Even though the maximum number of tests removed is lower than the other strategies, the stability of the window training strategy is notable for both matched test pair selection and their use. The effects of the low pass filter are evident for the window training strategy and the accumulate training strategy. The figure also indicates that the low pass filtering action with N=5 is substantially better than the smaller N=2 as in the case of the dual lot strategy.

5.6 TRC Performance Dependence on Threshold

This section discusses the TRC performance dependence on the Kendall's τ rank correlation threshold. The dependence is studied by evaluating two contrasting FOM across the range of Kendall's $|\tau|$ threshold.

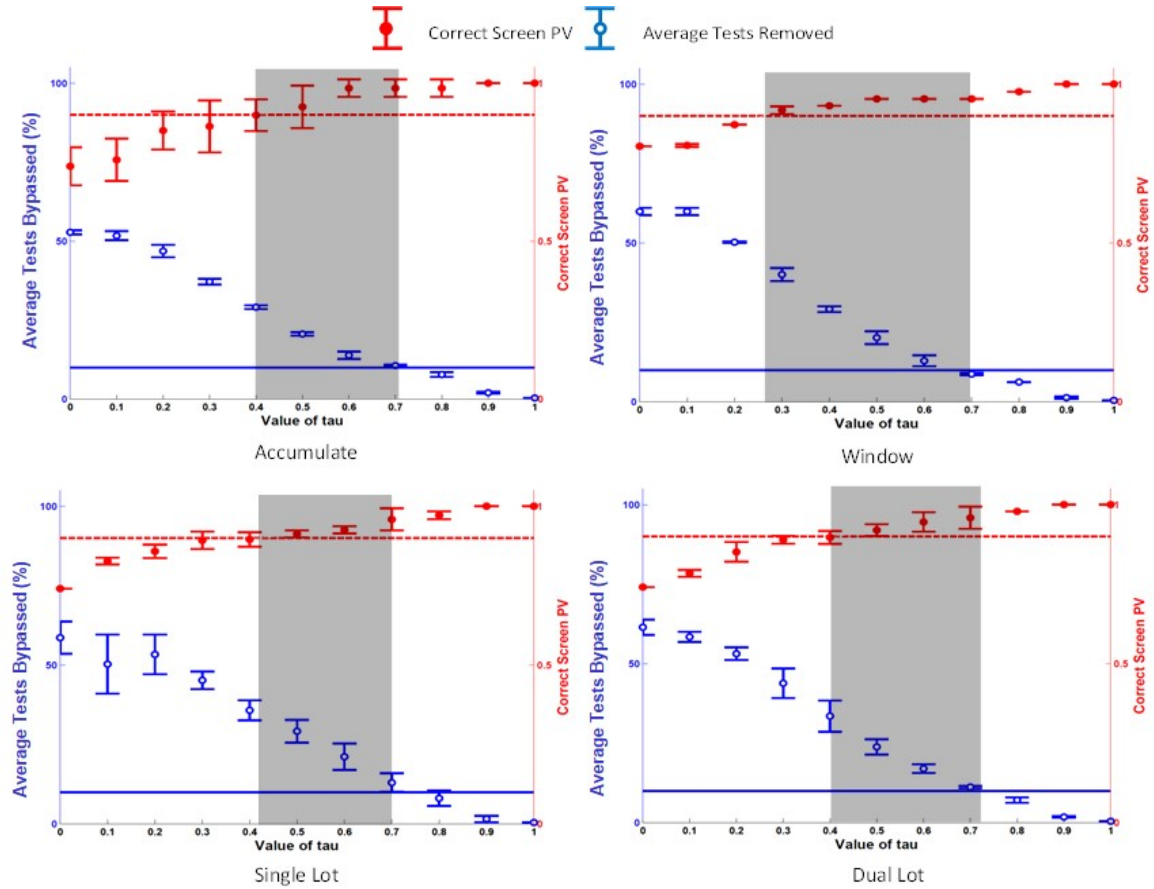

Figure 5.7: FOM versus τ threshold plots for the four different training strategies. The blue hollow circles and the red solid circles represent average tests bypassed and the correct screen PV respectively. The blue solid line and the red dashed line represent the target goals for the average tests bypassed and the correct screen PV respectively.

Figure 5.7 compares two contrasting FOM, the average tests bypassed and the correct screen PV plotted as a function of the threshold set for the selection of the matched test pairs (labeled “Value of tau”), for the four different training strategies. The FOM are evaluated for an LUT consisting of 25 wafers. The average tests bypassed is plotted in blue with hollow circles and the correct screen PV is plotted in red with solid circles. The contrasting nature of these two figures of merit is evident from the figure. For achieving a higher correct screen PV, the average tests bypassed is low and vice versa. The intervals report the range of the FOM for the different selection of the training wafers. The high variation observed for the single and dual lot strategies in Figure 5.6 is also observed in Figure 5.7. Another observation to be made is the variation in the correct screen PV for the accumulate training strategy. This is due to the larger sample size of the training data set.

Each FOM is of concern to a particular engineer involved in the testing phase. The average tests bypassed is of concern to the test engineer whose primary goal is the reduction of the test time. On the other hand, the correct screen PV is of concern to the product engineer whose primary goal is to make sure there are no test escapes transferred to a more expensive downstream insertion. These engineers need to achieve some target goals. The red dashed line and the blue solid line represent the target goals set for the correct screen PV and the average tests bypassed respectively. Both the engineers work towards achieving their respective target goals and set the threshold for the selection of the matched test pairs accordingly.

A compromise in the choice of this threshold has to be made to ensure that the target goals are met for both the concerned engineers. This is highlighted by the

grey box in 5.7. This box is called the Region of Compromise (ROC). The right boundary of the ROC is set by the target goal for the average tests bypassed and the left boundary is set by the target goal for the correct screen PV. As seen from the plots, the right boundary is almost equal for all the training strategies. On the other hand, the left boundary varies from one strategy to the other.

The similarity of the right boundary reflects the consistency in the selection and use of matched test pairs and the similarity in the performance of the TRC for the four training strategies at high values of τ as seen in Figure 5.7. At higher values of τ, the performance of the TRC is independent of the choice of training strategy. The price to be paid is the inability to meet the target goal for the average tests bypassed.

The difference in the left boundary is due to the use of these matched test pairs. This is highlighted by the variation in the correct screen PV with the choice of training strategy. Recall that the correct screen PV is a measure of the ability of the TRC to correctly screen a failing die. This variation of the correct screen PV is dependent on the limits set by training for the identification of failing die by the TRC. At lower values of τ, the performance of the TRC is sensitive to the choice of the training strategy.

The seven different training strategies, described in Chapter 4, were compared based on the performance of the TRC in this chapter. The next chapter discusses some conclusions from the study and scope of future work in the study.

Chapter 6

Conclusion and Future Work

This chapter briefly discusses the conclusions derived from the study and ends the discussion with the scope for future work in the study.

6.1 Conclusions

The application of machine learning tools is well established. The study compared the TRC performance with five different training strategies. The effectiveness of the TRC was further assessed by adding special case strategies. The results highlight three main conclusions from this study:

- Stable predictions are found by designing a filtering scheme through sampling of old, recent and new data into the training. The results of the cross validation study in Figure 5.6 supports this conclusion. The TRC matched test pair selection and use was stable for strategies that used low pass filtering such as the window and accumulate training strategies.

- The dependence of TRC performance on training is reduced to a single parameter - the rank correlation threshold. This is highlighted by Figure 5.6. The TRC performance dependence on the training strategy varies with the choice of rank correlation threshold. For high values of τ, the performance of the classifier was independent of the choice of training strategy whereas for lower values of τ, the performance of the classifier depended on the choice of the training strategy.

- The TRC performance for special case training strategies demonstrates the significance of rank correlation in training.

6.2 Future Work

The dependence of machine learning on training is studied at wafer sort in this thesis. This study can be extended from wafer sort to final test, including packaging and burn-in testing. Additionally, this thesis is focused on single-site testing. This study could also be explored for multisite testing.

References

[1] International Technology Roadmap for Semiconductors 2013, Test and Test Equipment. [Online]. Available: http://www.itrs.net/links/2013itrs/home2013.htm

[2] P. Gelsinger, "Discontinuities driven by a billion connected machines," *Design Test of Computers, IEEE*, vol. 17, no. 1, pp. 7–15, Jan 2000.

[3] P. O'Neill, "Statistical test: A new paradigm to improve test effectiveness & efficiency," in *Test Conference, 2007. ITC 2007. IEEE International*, Oct 2007, pp. 1–10.

[4] E. Marinissen, A. Singh, D. Glotter, M. Esposito, J. Carulli, A. Nahar, K. Butler, D. Appello, and C. Portelli, "Adapting to adaptive testing," in *Design, Automation Test in Europe Conference Exhibition (DATE), 2010*, March 2010, pp. 556–561.

[5] N. Sumikawa, D. Drmanac, L.-C. Wang, L. Winemberg, and M. Abadir, "Forward prediction based on wafer sort data - a case study," in *Test Conference (ITC), 2011 IEEE International*, Sept 2011, pp. 1–10.

[6] S. Biswas and R. Blanton, "Statistical test compaction using binary decision trees," *Design Test of Computers, IEEE*, vol. 23, no. 6, pp. 452–462, June 2006.

[7] L. Milor, "A tutorial introduction to research on analog and mixed-signal circuit testing," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 45, no. 10, pp. 1389–1407, Oct 1998.

[8] H. He and E. Garcia, "Learning from imbalanced data," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 21, no. 9, pp. 1263–1284, Sept 2009.

[9] H. Stratigopoulos, P. Drineas, M. Slamani, and Y. Makris, "Rf specification test compaction using learning machines," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 18, no. 6, pp. 998–1002, June 2010.

[10] J. Brockman and S. Director, "Predictive subset testing: optimizing ic parametric performance testing for quality, cost, and yield," *Semiconductor Manufacturing, IEEE Transactions on*, vol. 2, no. 3, pp. 104–113, Aug 1989.

[11] M. Chen and A. Orailoglu, "Test cost minimization through adaptive test development," in *Computer Design, 2008. ICCD 2008. IEEE International Conference on*, Oct 2008, pp. 234–239.

[12] G. E. A. P. A. Batista, R. C. Prati, and M. C. Monard, "A study of the behavior of several methods for balancing machine learning training data," *SIGKDD Explor. Newsl.*, vol. 6, no. 1, pp. 20–29, Jun. 2004. [Online]. Available: http://doi.acm.org/10.1145/1007730.1007735

[13] H. Ayari, F. Azais, S. Bernard, M. Comte, V. Kerzerho, O. Potin, and M. Renovell, "Making predictive analog/rf alternate test strategy independent of training set size," in *Test Conference (ITC), 2012 IEEE International*, Nov 2012, pp. 1–9.

[14] S. Benner and O. Boroffice, “Optimal production test times through adaptive test programming,” in *Test Conference, 2001. Proceedings. International*, 2001, pp. 908–915.

[15] S. Biswas and R. Blanton, “Maintaining accuracy of test compaction through adaptive re-learning,” in *VLSI Test Symposium, 2009. VTS ’09. 27th IEEE*, May 2009, pp. 257–263.

[16] Wikipedia- Ranking. [Online]. Available: http://en.wikipedia.org/wiki/Ranking

[17] M. Kendall, *Rank correlation methods*. C. Griffin, 1955. [Online]. Available: http://books.google.com/books?id=TJAuAAAAMAAJ

[18] L. Wang, C. Wu, and X. Wen, *VLSI Test Principles and Architectures: Design for Testability*, ser. Systems on Silicon. Elsevier Science, 2006. [Online]. Available: http://books.google.com/books?id=P1ea4znZhGsC

[19] M. L. Bushnell and V. D. Agrawal, *Essentials of Electronic Testing for Digital, Memory, and Mixed-Signal VLSI Circuits*. Boston: Kluwer Academic, 2000.

[20] Data Buffer. [Online]. Available: http://en.wikipedia.org/w/index.php?title=Data_buffer&oldid=603650601

[21] N. Sumikawa, L.-C. Wang, and M. Abadir, “An experiment of burn-in time reduction based on parametric test analysis,” in *Test Conference (ITC), 2012 IEEE International*, Nov 2012, pp. 1–10.