Portland State University

PDXScholar

Dissertations and Theses

Dissertations and Theses

Winter 6-10-2013

Data Driven Feed Forward Adaptive Testing

Chaitrali Santosh Chandorkar
Portland State University

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

Part of the [VLSI and Circuits, Embedded and Hardware Systems Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Chandorkar, Chaitrali Santosh, "Data Driven Feed Forward Adaptive Testing" (2013). *Dissertations and Theses*. Paper 1049.

<https://doi.org/10.15760/etd.1049>

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. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Data Driven Feed Forward Adaptive Testing

by

Chaitrali Santosh Chandorkar

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
2013

© 2013 Chaitrali Chandorkar

Abstract

Test cost is a critical component in the overall cost of the product. Test cost varies in direct proportion with test time. This thesis introduces a data driven feed forward adaptive technique for reducing test time at wafer sort while maintaining the product defect level. Test data from first insertion of wafer is statistically analyzed to make a decision about adaptive test flow at subsequent insertions.

The data driven feed forward technique uses a statistical screen to analyze test data from first probe of wafer and provides recommendations for test elimination at second insertions. At the second insertion dies are subjected to only the optimum number of tests for a reduced test flow. This technique is applicable only for the products which are tested at two or more insertions.

The statistical screen identifies the dies for reduced test flow based upon correlation of tests across insertions. The tests which are repeated at both the insertions and are highly correlated are the candidates of elimination at second insertion.

The feed forward technique is applied to a mixed signal analog product and figures of merit are evaluated. Application of technique to the production data shows that there is an average 55% test time reduction when a single site is tested per touchdown and up to 10% when 16 sites are tested in parallel per touchdown.

Acknowledgements

I take this opportunity to thank many wonderful people for their guidance and encouragement during my study at Portland State University. First of all, I would like to thank my research adviser Dr. W. Robert Daasch. This research would not have been possible without his guidance, constant motivation and deep insight in IC test and statistics. He is the person who has taught me how to solve any problem systematically and with perfection.

I am very thankful to our industrial liaisons at Texas Instruments (TI), Kenneth Butler, John Carulli and Amit Nahar for their valuable contribution in this project and giving me an opportunity to work as a summer intern at TI. I would like to thank Dr. Glenn Shirley for all the help and guidance he provided in the two years of my association with the ICDDT laboratory. I express my gratitude to Professor Mark Faust for serving on my M.S. Committee.

Special thanks to Kapil for introducing me to the field of IC test and for all his help and guidance throughout my graduate study. I cannot thank Rajesh and Prachi enough for their support in difficult times and without whom this thesis would not have been possible. I would like to thank my dearest friend Poorva for always being there for me, in tough times and in happy times.

Finally, thanks to my parents, my grandparents and my brother for being such a great support. Special thanks to my mother who taught me to achieve perfection in everything I do. I would also like to take a moment to remember my grandmother who breathed her last during my stay in the U.S.

Contents

Abstract	i
Acknowledgements	ii
List of Figures	v
1 Introduction: Need, Contribution and Organization	1
1.1 Introduction to IC Testing	1
1.2 Brief Description of Thesis and Contribution	3
1.3 Organization of Thesis	7
2 Background and Previous Work	8
2.1 Rank	8
2.2 Properties of Rank order	8
2.3 Correlation	9
2.4 Kendall's tau	10
2.5 Introduction to Mixed-signal Analog Test	11
2.6 Test Limits	13
2.7 Binning Dies	14
2.8 Previous Work	15
3 The New Feed Forward Adaptive Testing Method	18
3.1 Selection of Tests for Elimination at Insert2	18
3.2 Bounds Selection	21

3.3	Compact Test Limits at Insert2	22
4	Results and Figures of Merit	26
4.1	Product Data	26
4.2	Compact Test Limits at Insert2	29
4.3	Cross Validation	30
4.4	Test Bounds for Insert1	34
4.5	Site to Site Variations in Test Data	35
4.6	Handling Inverted Test Bounds at Insert1	39
4.7	Categories of Dies for a Test	40
4.8	Categories of a Die for a Test Flow	42
4.9	Figures of Merit	43
4.10	Evaluation 1: Figures of Merit using Datasheet Test Limits	46
4.11	Evaluation 2: Figures of Merit using Compact Test Limits	47
4.12	Evaluation 3: Figures of Merit using Compact Test Limits without Lot6	48
4.13	Multisite Adaptive Test	50
4.14	Test Time Reduction for ADC, VREF Bin	52
4.15	Effect of Parallel, Serial tests on Test Time Reduction	54
5	Conclusion	57
5.1	Contribution of this work	57
5.2	Recommendations and Future work	59
	References	60

List of Figures

1.1	Conventional Test flow	4
1.2	Adaptive test flow with a statistical screen	4
1.3	Reduced test flow at Insert2	5
1.4	Correlated test pair	6
2.1	Concordance-Discordance plot	12
2.2	Test limits	14
3.1	Adaptive test flow	18
3.2	Correlated matched test pair	19
3.3	Selection of bounds at first insertion	20
3.4	Datasheet and compact test limits at Insert2	23
3.5	Selection of compact test limits at Insert2	24
3.6	Compact test limits	25
4.1	Multisite Testing for 16 sites	27
4.2	Leave one out cross validation	31
4.3	ADC test data-cross validation	32
4.4	Distribution of site5 for ADC Test - Insert2	34
4.5	Wafer plot - cross validation	35
4.6	Histogram-site to site variation	36
4.7	Site to site variations when offsets are not correlated	37
4.8	Site to site variations when offsets are correlated	38
4.9	Figures of Merit-1	40

4.10	Categories of a die for test flow	43
4.11	Figures of Merit	44
4.12	TTR for ADC bin	53
4.13	TTR for VREF bin	54
4.14	TTR - Serial/Parallel Tests	55

Chapter 1

Introduction: Need, Contribution and Organization

1.1 Introduction to IC Testing

IC testing is a process of differentiating good devices from bad devices to ensure the quality of product shipped to customer. Testing involves many steps. The first step is wafer sort where wafers are probed multiple times and tested under different environmental conditions. The passing dies from wafer sort are packaged and tested at package level to remove dies with assembly problems or possible wafer-sort test escapes. The devices passing tests at package level are sometimes subjected to high stress or burn-in tests.

With the growth of telecommunication and automobile market, there is an increase in the demand for mixed signal devices [1]. These devices integrate the digital and analog components on a single die to increase performance and reduce the die size. Dies to be used in critical applications such as automotive, medical etc. are subjected to exhaustive testing.

In the production of mixed signal devices, test is a major factor contributing to manufacturing cost. The main challenge is that most analog mixed-signal circuits are tested by functionality, which is both expensive and time-consuming [2]. Consider measuring the integral non linearity (INL) of an ADC. For a 10-bit ADC, this would require locating 1024 input voltages which cause transitions in the output between codes, at multiple temperatures. Such large number of long tests can limit throughput during production testing. These products are usually tested at

multisite where multiple devices are tested in one touchdown of the prober.

Yield (Y) is the ratio of good dies to the total dies tested. Quality of parts shipped to a customer is measured by Defect Level (DL). Defect level is a function of Fault Coverage (FC) and Yield. Fault coverage is defined as the number of modeled faults tested to total number of the possible modeled faults [3]. Equation 1.1 shows relation between defect level, yield and fault coverage [4].

$$DL = 1 - Y^{(1-FC)} \quad (1.1)$$

Test cost is a critical component in the overall cost of the product. Test cost varies in direct proportion with test time. Test time for a failing die is the time required to first find a defect (test time bad). Test time for a good die is the time required to execute all the tests (test time good). As the yield (Y) increases test time good becomes a major contributor to the overall test time.

This thesis presents a method to reduce the test time of good dies. Test data from first insertion at wafer sort is statistically analyzed to make a decision about adaptive test flow at subsequent insertions. The aim is to reduce test time and still maintain the same product quality.

Process variation results in increasing statistical diversity in manufactured dies. Variation in test measurements appears as lot-to-lot variation, wafer-to-wafer variation and die-to-die variation. Test plans developed without this diversity in mind are bound to result in poor test yield or long test times. Adaptive testing is used to tailor the test strategy to accommodate a wide range of variation in the statistical characteristics of manufactured devices [5].

Adaptive test is recognized as a key driver for future of semiconductor test and was formally defined in the International Technology Roadmap of Semiconductors (ITRS), 2009, [6] “adaptive test is a broad term used to describe methods that change test conditions, test flow, test content and test limits (potentially at the die/unit or sub-die level) based on manufacturing test data and statistical data analysis. This includes feed-forward data from inline test and early test steps to later test steps and feed-back data from post-test statistical analysis that is used to optimize testing of future products.”

The key to adaptive test is to realize the optimum test set for screening defects. This optimum test set is different for every die. The key to adaptive test is to utilize data generated from the tester or relevant data from previous processes or measurements in predicting the process for the future tests in order to reduce or increase testing when required. The ultimate goal is to apply only the minimum set of tests required to screen the ICs that will fail in the system either as shipped or over time [7].

1.2 Brief Description of Thesis and Contribution

This thesis uses the correlation among test data to implement feed forward adaptive testing. Data collected from first insertion of the wafer at wafer sort is statistically analyzed to make adaptive test flow changes at subsequent insertions. Test flow changes means subjecting a die to only a required number of tests. The test flow changes are done on per die basis.

A wafer is tested at two insertions. In this thesis, the first insertion of a wafer is called as *Insert1* and the second insertion is called as *Insert2*. The two insertions

usually are at different temperatures or different environmental settings. Figure 1.1 shows the conventional test flow.

Figure 1.1: Conventional test flow

A wafer is first tested at *Insert1*. Dies failing at *Insert1* are discarded. The dies which pass *Insert1* are further tested at *Insert2*. At *Insert2* failing dies are discarded and passing dies are shipped to the customer.

Figure 1.2: Adaptive test flow with a statistical screen

This thesis introduces a new statistical screen in conventional test flow as shown in the Figure 1.2. The statistical screen analyzes test data from *Insert1* and identifies

the dies which will pass test at insertion *Insert2*. Such dies will not undergo a full test flow but a tailored reduced test flow at *Insert2*. This technique will achieve significant test time savings at *Insert2* while not compromising product quality.

Figure 1.3 shows test flow at *Insert2* for 3 sample dies. The tests are marked in green if they are eliminated at *Insert2* and in red if they are executed at the *Insert2*. As seen from the figure, Die 1 is subjected to a full test flow as it is identified as a susceptible fail by statistical screen. Die 2 and Die 3 undergo a reduced test flow at *Insert2*. Die2 is tested with Test3, Test4 and Test5 while Die3 is tested with Test2, Test3 and Test5.

Figure 1.3: Test flow for 3 sample dies at Insert2

The statistical screen identifies the dies for reduced test flow based upon correlation of tests across insertions. The tests which are repeated at both the insertions and record continuous parametric values are the candidates of elimination at *Insert2*. If the measurement of the same test at two different insertions are highly correlated then the test at *Insert2* can be eliminated for a population of good dies.

Data from correlated test pair is shown in Figure 1.4. The plot shows reference voltage measurements for a mixed signal automotive device. Each point on the scatter plot represents the test measurement taken at two different temperatures. The two test measurements are correlated with Kendall's $\tau = 0.9$.

Figure 1.4: Production data: Scatter plot of correlated test pair with Kendall's $\tau = 0.9$

If one of these tests from the correlated test pair indicates that a die is good die it will pass the test at *Insert2*. Eliminating test at *Insert2* will result in test time reduction.

1.3 Organization of Thesis

This thesis is organized in four chapters. Background concepts about rank order statistics and mixed signal analog test are explained in Chapter 2. Chapter 2 also outlines the previous work done in adaptive test. The feed forward adaptive testing model proposed in this thesis is presented in the Chapter 3. Chapter 4 contains the results of application of the proposed new technique to production data available from a leading semiconductor manufacturer. Contributions of this work and recommendations for future work are presented in the Chapter 5.

Chapter 2

Background and Previous Work

This chapter gives a brief background of statistical concepts used in this thesis and introduction to mixed signal analog device testing. Statistical terms like rank, correlation and Kendall's tau correlation coefficient are illustrated. Literature survey of previous work done in adaptive testing is also presented.

2.1 Rank

Rank is data transformation in which numerical or ordinal values are replaced by their rank when the data is sorted [8]. The rank of a value in a distribution is simply its numbered place in the list of ordered values.

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

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 numbers in the original set. Monotonic transformation can be addition, subtraction, multiplication, division, logarithm etc [9]. E.g.: $S' = S \cdot 100$

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

If two or more observations in a set have the same values, they are called as ties. Tied data values have to be handled while ranking the data. There are several different methods to handle ties such as standard competition ranking, dense ranking, ordinal ranking, fractional ranking etc.

The method used in this thesis is fractional ranking. In fractional ranking, tied values receive the same ranking number, which is the mean of what they would have under ordinal rankings.

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

In the above example, there are two numbers with the same value twenty three. In ranking set S, tied values twenty three have been assigned 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 used to select dies for reduced test flow at *Insert2*. Many other statistics such as chi-square are used to quantify how well the test can separate passing dies from failing dies.

Correlation is a statistical measure of dependence between two variables. Consider two random variables X and Y. If increase in the values of X is associated with increase in the values of Y then they have a strong dependence and are said to be positively correlated. If with increase in the values of X, the values of Y decrease then they have anti-dependence and said to be negatively correlated.

Correlation between two variables is measured in terms of the correlation coefficient. It gives information regarding relationship between variables. Correlation coefficient tells if the relationship is positive or negative and the strength of relationship.

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

$$\rho = \frac{cov(X, Y)}{\sigma_x \cdot \sigma_y} \quad (2.1)$$

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

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 in to ranks and the correlation coefficient is evaluated using equation 2.1 for the transformed values. In this thesis Kendall's tau is used as a measure of correlation.

2.4 Kendall's tau

Kendall's τ is a non-parametric statistic measuring the association between two or more variables. τ varies between -1 and 1. Let $X = X_1, X_2, \dots, X_n$ and $Y = Y_1, Y_2, \dots, Y_n$ be two set of n measurements. A pair of measurements (X_i, Y_i) and (X_j, Y_j) is concordant, if

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

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 measurements is $n(n-1)/2$. Rank correlation coefficient τ between two data sets X and Y of each sample size n is calculated as shown.

$$\tau = \frac{2 \cdot (\text{number of concordant pairs} - \text{number of discordant pairs})}{\text{total number of pairs}} \quad (2.2)$$

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

Calculation of τ using discordance plot in Figure 2.1.

$$X = 32, 20, 41, 1, 88, 63, 50, 74 \quad Y = 0.5, 0.3, 0.4, 0.6, 0.2, 0.7, 0.6, 1.0$$

$$\text{Rank order}(X) = 3, 2, 4, 1, 8, 6, 5, 7 \quad \text{Rank order}(Y) = 4, 2, 3, 5, 1, 6, 8, 7$$

Tau is calculated by counting the number of concordant and discordant pairs from the discordance plot and substituting values in equation 2.2.

$$\tau = 0.72$$

2.5 Introduction to Mixed-signal Analog Test

Mixed-signal chips include analog circuitry (operational amplifier (Op Amp), comparators, filters and PLL) and digital circuitry (data paths, control logic). The

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

interfaces between the analog and digital circuits are digital-to-analog converters (DAC) and analog-to-digital converters (ADC).

In mixed-signal chip, both analog and digital circuitry is involved. Several different schemes are required to test a mixed-signal chip [10]. A typical strategy for testing a mixed-signal chip involves, first testing the digital and analog components, followed by some system tests to check the at-speed interaction among 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's functionality. ADC and DAC are tested for static parameters such as accuracy, resolution, dynamic range, offset, gain, differential non-linearity (DNL),

integral non-linearity (INL) and dynamic parameters such as SNR, Total harmonic distortion etc.

Broadly tests can be classified in to two categories: Parametric tests, Functional tests.

- Parametric tests: 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 [3]. Parametric tests are necessary to decide whether the chip meets various rise and fall times, setup and hold times, low and high voltage thresholds, and low and high current specifications.
- Functional tests: Functional test vectors verify the functionality of the chips. Input vectors are applied and output responses are recorded. Functional tests cover a very high percentage of modeled faults in logic circuits. Functional tests determine whether the internal digital logic and analog sub-systems in the chip behave as intended. The results are usually recorded by pass or fail or go or no-go binary result [3].

2.6 Test Limits

At test many different measurements such as current, voltage, frequency etc. are taken. The tester compares the measured values with test limits and then marks the die as pass or fail. If a parameter is outside a predetermined limit or threshold, the die is marked as fail.

Figure 2.2: (a)(b) Single value test limit (c) Test limit defined as interval

Test limits are defined by a single value or an interval. Figure 2.3 shows an example of three types of test limits.

Figure 2.3(a) shows a single test limit. The single test limit is not to exceed limit. If test measurement is greater than the test limit, die is marked as a fail. Similarly as shown in Figure 2.3(b) the single test limit is not to fall below a certain value. If the test measurement is less than the test limit, die is marked as a failing die.

Figure 2.3(c) shows an example of test limit which is defined as an interval. If the test measurement value lies within test limit min and test limit max, then the die is a good die. If the measurement lies outside the interval the die is marked as a fail.

2.7 Binning Dies

Tests in a test program are divided in different categories or groups depending upon their nature or the section of the circuit they are testing. If a die fails in any

one test from a set 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, stuck at fails etc.

The number of die in various bins provides information about the common failure modes of the chip. Usually gross failures, like shorts between power and ground, are checked first, and followed by, tests for defects in digital components. The last set of tests usually relates to checking the performance of the entire system [1].

2.8 Previous Work

Forward prediction to eliminate tests at final test and at burn-in is studied by [11] [12]. Sumikawa et. al [11] focus on modeling test data after 10 hours of burn in to predict parts which are likely to fail after 48 hours of burn-in. The authors have developed a technique which separates passing parts from failing parts using Principal Component analysis (PCA). Chi-square statistics was used for feature (test) selection to quantify how much the test can separate passing parts from failing parts. Entire socket was eliminated for good parts instead of reduced test flow. Half of the total available fails were analyzed and multivariate models were developed to screen similar fails. The method does not handle unique fails and authors suggest that rule learning is required to handle unique fails that only occur after 24 or 48 hours of burn-in.

Sumikawa et. al [12] have analyzed parametric wafer sort data to predict fails at final test or customer returns. The separation of failing parts from passing parts is done using a binary classifier model and an outlier model. The method learns from failing parts to find the test perspective where the failing parts can be separated

from most of the passing parts. Multivariate test analysis was used to predict parts that would fail in the field, i.e. customer returns.

Biswas et. al [13] analyzed the parametric test data and applied enhanced binary decision forests to identify redundancies in the test set. The authors identified redundant tests belonging to one insertion and suggested that more expensive tests could be replaced with models built from those in less expensive insertions. They have utilized multivariate test analysis to predict failing parts.

Kupp et. al [14] identifies a set of inline (etest) parameters correlating to module final test parameters. Final test values are predicted by exploiting the existing correlations using regression models.

The work by Madge et. al [7] introduces the concept of using correlation between parameters to identify only the minimum set of tests required to screen the ICs that will fail in the system. Different types of adaptive tests for defect coverage, outlier screening, test pattern re-ordering are explored as a first step in paradigm shift away from full testing of every part.

Authors in [15] explore an adaptive strategy pruning the test set based on test correlation analysis. A graph model is developed to represent the correlation relationship among the candidate tests, based on which an algorithm is proposed to generate the minimum test set that covers all the correlated tests. The reduced test set is then ordered in the second stage using statistical models to maximize the overall probability of early fault detection. Tests with higher Cpk have a reduced probability of the performance falling out of the acceptance region and a lower

defect detection probability. Cpk is used to make decision to drop a test from the test flow that has a less quality impact.

Chapter 3

The New Feed Forward Adaptive Testing Method

Figure 3.1: Adaptive test flow with statistical screen

Figure 3.1 shows the new feed forward adaptive test flow introduced in section 1.2. Statistical screen is trained by test data from *Insert1* and *Insert2*. Correlation among tests at both insertions is the underlying principle for identifying dies for reduced test flow at *Insert2*.

3.1 Selection of Tests for Elimination at Insert2

Not all tests executed at the second insertion are eligible for elimination in the reduced test flow. There must be some criteria to select tests for elimination at *Insert2*.

Figure 3.2: Production data: Scatter plot of correlated matched test pair

A test should record continuous parametric values and must be repeated at both the insertions. The test measurements at both insertions should be highly correlated. Such a pair of tests is called as a matched pair. The electrical test in the matched pair is same at both the insertions except for the change in environmental condition.

Figure 3.2 shows scatter plot of correlated matched pair from production data available for analysis. The data shows reference voltage measurements for operational amplifiers in the analog design.

The screen has to identify good dies based upon their values at *Insert1* which will pass the test *Insert2*. These dies will be eligible for reduced test flow at *Insert2*.

Figure 3.3: Step by step procedure for setting Upper Bound and Lower Bound

3.2 Bounds Selection

Figure 3.3 shows data scatter plot for a correlated pair of random variables with Kendall's tau= 0.7. All plots in Figure 3.3 show the same synthetic data. Assume that the two random variables are test data measurements of matched test pair at *Insert1* and *Insert2*.

In Figure 3.3(b), horizontal red lines show arbitrary test limits at *Insert2*. All the dies which have *Insert2* test measurement values greater than limit high or less than limit low will fail the test at *Insert2*. We are interested in identifying dies which will pass test at *Insert2* based upon test response at *Insert1*. Test bounds are defined at *Insert1* to identify die for reduced test flow.

Find the minimum test measurement at *Insert2* greater than limit low. If there are many dies having the same minimum test measurement at *Insert2* (ties) then choose the die which has the maximum test measurement at *Insert1*. Ties are resolved using the test measurement at *Insert1*. In Figure 3.3(c) there are 3 dies having minimum test measurement of -0.9 greater than limit low. To resolve the ties, we choose the die which has maximum *Insert1* test measurement. This is done so that we choose a point deeper in distribution of test at *Insert1*. The test measurement of this die at *Insert1* is selected as the lower bound.

Similarly, find the die with maximum test measurement at *Insert2* less than limit high. Ties are resolved using the minimum *Insert1* test measurement. The test measurement of this die at *Insert1* is selected as upper bound.

The upper and the lower bounds decide an interval based upon the *Insert1* test measurements as shown in the Figure 3.3(d). The dies which have test measurements within the upper and lower bound at *Insert1* are likely to pass the test at *Insert2*. Dies which lie within the bounds will have a reduced test flow at *Insert2*.

The bounds must be computed for every correlating matched test pair. The tailored test flow for a die will be decided based upon its response at *Insert1*.

The upper and lower bounds at *Insert1* and the test limits at *Insert2* divide the scatter plot in four different categories. Bad dies which are not tested at *Insert2* because of incorrect decision of the screen are called Escape Screen (ES). ES contribute to the end use defect level as these parts will be shipped to the customer without testing at *Insert2*. Good dies which are tested at *Insert2* are called as Over Screen (OS). Good dies which are not tested at *Insert2* are called Pass Screen (PS). Dies which are tested at *Insert2* and fail the test are Correct Screen (CS).

3.3 Compact Test Limits at Insert2

Both plots in Figure 3.4 show matched correlated test pair with test limits at *Insert2*. Figure 3.4(a) shows the datasheet limits at *Insert2* and Figure 3.4(b) shows compact test limits at *Insert2*.

The datasheet test limits are set to accommodate variations from different factories, different testers etc. Datasheet test limits are wide as compared to the distribution of data from one factory. The wide datasheet specification limits affect the decision about lower and upper bounds at *Insert1*. The wide datasheet limits lead to

Figure 3.4: (a) Datasheet test limits at *Insert2* (b) Compact test limits at *Insert2*

selecting wide bounds at *Insert1*.

Wide bounds at *Insert1* means that almost all the dies in Figure 3.4(a) are within the bounds. Very few dies will be tested with the matched test at *Insert2*. The probability of Escape Screen (ES) dies increases. The prediction accuracy is affected by the wide datasheet limits. There is a need for compact test limits to select proper bounds at *Insert1*.

Figure 3.4 (b) shows compact test limits at *Insert2*. The compact test limits at second insertion lead to tighter bounds at *Insert1*. Compact test limits at *Insert2* give a better idea about the distribution of test data. Test bounds at *Insert1* due to compact limits reduces the probability of failing dies escaping the screen. However, tight bounds at *Insert1* imply that more dies will be tested at *Insert2* which will increase the number of Over Screen (OS) dies.

Figure 3.5: Selection of compact test limits at Insert2

To select compact test limits for test data from one particular factory, a training dataset must be analyzed and limits within distribution of the test data must be chosen. In this work, test data available for 7 lots of a product was analyzed and compact test limits were set. The tenth minimum rank value of test measurement ranked in ascending order is chosen as compact test limit low. Similarly 10th maximum rank value of test measurement ranked in descending order is chosen as compact test limit high.

Figure 3.5 shows distribution of a correlated matched test from production data at *Insert2*. The red lines show the compact test limits based upon the 10th rank selection.

Figure 3.6: Scatter plot of correlated matched test pair with compact test limits

The scatter plot in Figure 3.6 shows the new compact test limits with respect to the distribution of test data.

Chapter 4

Results and Figures of Merit

4.1 Product Data

Wafer sort test response data for a mixed signal automotive product is used to develop data driven feed forward adaptive testing technique. Dataset contains a total of 122 wafers divided in 7 lots with 6800 die per wafer. Lots 1-7 have 13, 19, 15, 22, 22, 22 and 9 wafers respectively.

Product was tested at two insertions *Insert1* and *Insert2* at wafer sort. *Insert1* is at room temperature of 25° C and *Insert2* is at high temperature of 85° C. Dies were first tested at room temperature in *Insert1*. Failing dies at room temperature were discarded and not tested at insertion *Insert2*. At *Insert2*, passing dies from *Insert1* were re-tested at elevated temperature.

Wafers are tested using multisite testing technique where multiple dies are tested on every touchdown of prober. For the product, sixteen dies were tested per touchdown. Each touchdown was a 4x4 template and each position on that template is called a site. A die was tested at the same site or the same position at both insertions to minimize the test measurement instrument offset errors. Figure 4.1 shows the wafer being probed as 16 sites per touchdown. The order of sites in the touchdown are as seen in the product data.

Every die is uniquely identified by a set of attributes. They are lot id, wafer-id, X location, Y location, site, touchdown. Lot-id and wafer-id are used to identify the

Table 4.1: Product data description

	Insert 1	Insert 2
Total Test Records	780	396
Tests- Parametric Measurements	480	286
Tests- Pass/Fail	300	110

wafer. X and Y locations on the wafer are used to locate the die on the wafer. Site is the position of the die in the touchdown template. The site count goes from one to sixteen for this product. Touchdown is number of times a wafer is probed. The dies which are tested in one touchdown of the prober have the same touchdown number.

Figure 4.1: Multisite testing for 16 sites

Table 4.1 shows the details of the tests executed at *Insert1* and *Insert2*. The first row shows the total number of test records for every die.

Parametric tests at *Insert1* and *Insert2* are measurements of propagation delay, voltage reference, leakage current, rise time, analog to digital converter values or pass/fail kind of digital functional tests. The pass/fail tests only record the result if the part meets a given specification or not. Pass/fail tests are not used in this analysis. The tests which record some kind of continuous parametric values are chosen for analysis.

Test data presented in all the plots is normalized between -1 and 1. The tests have been referred by generic names according to which bin they map in to. The names do not reflect any actual test executed in the test program.

Any defect level (DPPM) value is for the training data analyzed in this thesis. Generally automotive products are expected to have end use defect level in the range of zero to ten DPPM. In order to maintain the quality of the product shipped to customer the adaptive test flow must limit the defect level below ten DPPM.

As discussed in section 3.1 we are interested in finding tests which record some continuous parametric measurements and are highly correlated. Such matched pairs of test are found to be the tests in the voltage reference (VREF) bin and analog to digital converter (ADC) bin. Tests in these two bins are used for analysis in this thesis.

ADC bin has total 56 tests and VREF bin has 27 tests. Out of these, 44 tests in ADC bin and 23 tests in VREF bin are correlated with Kendall's tau greater than

Table 4.2: Correlated test pairs

	ADC Bin	VREF Bin
Tests with Kendalls tau > 0.5	44	23
Total Tests	56	27

0.5. These 67 highly correlated matched pairs are chosen for analysis. The tests associated with VREF bin will be referred as VREF Test 1-23 and those in ADC bin will be referred as ADC Test 1-44.

The new feed forward adaptive testing technique described in section 3.2 is applied to production data and figures of merit are evaluated.

4.2 Compact Test Limits at Insert2

Compact test limits as discussed in section 3.3 introduce additional fails. There are two types of fails.

- Compact fails. Dies with test measurements outside compact test limits are named as compact fails.
- Datasheet fails. Dies with test measurements outside the datasheet test limits are named as datasheet fails. The datasheet fails are the original fails in the product data.

Compact test limits introduce additional fails in the test data. Total number of fails introduced by the compact test limits is shown in the Table 4.3.

Bin	Datasheet Fails	Additional Compact Fails	Total Fails
ADC Bin	40	105	145
VREF Bin	5	90	105
Total	45	195	240

Table 4.3: Fails introduced by compact test limits

The table 4.3 shows the datasheet fails and compact fails for ADC bin and VREF bin. The compact fails in the table are the additional failing dies introduced by the compact test limits. Essentially all the 240 dies have test measurements outside compact test limits and are compact fails.

4.3 Cross Validation

To study the effect of sample size on the new adaptive test limits, an experiment of cross validation was carried out.

Cross validation is a technique for assessing how the results of a statistical analysis will generalize to an independent data set [16]. It is mainly used when the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice.

Leave One Out Cross Validation

Leave-One-Out Cross-Validation (LOOCV) involves using a single observation from the original sample as the validation data, and the remaining observations as the training data. This is repeated such that each observation in the sample is used once as the validation data.

Figure 4.2: Leave one out cross validation

For a dataset with N examples, N trials are performed. For each trial $N-1$ examples are used for training and the remaining example for testing. Figure 4.2 shows N trials. In each trial the green block is used for validation and remaining blocks are used for training.

The production data available for analysis has 122 wafers which are divided in seven different lots. In order to find the sensitivity of test limits at second insertion *Insert2* to training sample size, LOOCV is used. For each trial one lot is used as the validation lot and the remaining 6 lots are used as training lots to find the test limits. Hence seven experiments are carried out. The test limits for all the 67 test parameters from VREF bin and ADC bin are calculated in each experiment.

The results of LOOCV show that for 49 out of the 67 parameters, leaving one lot out has little effect on the test limit. The test limits remain constant even if we leave one lot out of the analysis.

There are 18 tests which are sensitive to elimination of a particular lot from analysis. All these 18 tests are from the ADC bin. These tests have a significantly different test limit low when Lot6 is not used in the training set. Lot6 is different or an outlier lot as compared to others.

Further analysis is done to find out why Lot6 behaves differently than other lots for these 18 tests. It is seen that dies which are tested at site5 on the wafers from Lot6 have an outlier behavior as compared to other dies.

Figure 4.3: Distribution of test data for example wafer from Lot6

Figure 4.3 shows the distribution of test data for ADC test at *Insert2* which is one of the 18 tests sensitive to Lot6 elimination. Test data from site5 has been colored in red. As it can be clearly seen from the Figure 4.2, all the dies tested at site5 have low values for ADC test and lie in the left tail of the distribution. There is

an offset in the analog to digital measurements taken at site5. All 18 tests have a similar distribution. No other lot from the data shows a similar behavior.

The measurement offset at site5 has a significant effect on test limits. When Lot6 is excluded from the training set, the test limit low is the 10th rank value in the blue distribution. Number of fails increase when Lot6 is not used as a training lot.

All the Lot6 fails come from the lower tail of the distribution i.e. have measurement values less than test limit low. Further drill in shows that failing units come from one particular portion of the wafer.

Figure 4.4 shows that the distribution of data from site5 is bi-modal. The two parts red and blue come from two different halves of the wafer. The red part of the distribution comes from dies which have X location greater than 50 and the blue part of the distribution comes from dies which have X location less than 50. So there is a spatial component present in the distribution of test data for site5. All the fails are from one half of the wafer where X co-ordinate is greater than 50.

Figure 4.4: Distribution of site5 for ADC test at Insert2

Figure 4.5 shows the wafer map for an example wafers from Lot6. The fails marked by red dots come from that part of wafer which has $X > 50$.

4.4 Test Bounds for Insert1

Test bounds for each correlating test pair are selected from data as described in section 3.2. Bounds are then applied to identify dies for reduced test flow at *Insert2*.

To find the bounds test data at *Insert2* should be available. The idea is to test first wafer from a lot with full test flow at both the insertions and then use the test data to identify dies for reduced test flow in remaining wafers of the lot. One wafer is used to train the bounds at *Insert1*. A lot typically consists of 20-25 wafers.

Figure 4.5: Wafer plot of a wafer from Lot6

Testing one wafer completely per 25 wafers ensures a valid feedback from the data and reset of the test bounds.

The product undergoes multisite testing at *Insert2* and *Insert1*. Analysis of test data shows that there are site to site offsets in the test measurements. It is necessary to understand the effect of site to site variations on calculation of bounds.

4.5 Site to Site Variations in Test Data

Multisite testing is done at probe or package level where a single ATE tests multiple devices at the same time. Each device in a single touchdown is called a site. Multisite testing is done to parallelize parametric testing which results in overall test time and test cost reduction. Measurement instruments are replicated in the

tester. Measurements taken from different sites may have different offsets due to calibration of the instruments, probe sensitivity, etc.

Figure 4.6: Site to site variations of 4 sites for example wafer

Figure 4.6 shows the distribution of ADC test at *Insert2* for an example wafer from production data. The four different colors show the distribution for four different sites. Every site has an offset in the test measurement as compared to others. Every site has a similar shape of distribution but it is centered about a different mean.

Site to site variations can be misleading when looking at correlation of two tests. The offsets in test measurements can be correlated or miscorrelated. There are two conditions where:

The offsets are not correlated but test measurements taken at each site are highly correlated

Consider the scatter plot for four sites shown in Figure 4.7. Each individual site is highly correlated ($\tau > 0.85$). The overall correlation between two parameters considering data from all sites is poor ($\tau = 0.4$). This is because the offset in the measurements at four sites are not correlated.

Figure 4.7: Production data: Site to site variations when offsets are not correlated

The offsets are correlated but each test measurement taken at each site is not correlated

Figure 4.8 from [11], shows a scatter plot for correlated tests, Test 1 and Test 2. Test data measurements from entire wafer shows a clear high correlation. If we

Figure 4.8: Site to site variations - offsets are correlated (Figure 8 from [11])

focus on only the samples of a single site, we see that the two tests are actually uncorrelated. The offsets between test data of different sites are correlated and not the test data. Misleading linear trend can be created by site-to-site variation.

Test data analysis shows that there are site to site variations in test data. On every touchdown of the prober sixteen sites are tested. Offsets in measuring instruments cause the test data tested at a single site to be more correlated with test data from same site than entire wafer. Test data at site 1 at *Insert1* is highly correlated to test data at site 1 at *Insert2*.

We choose test bounds for every site for per test than for every wafer per test. Every test has sixteen sets of (Upper bound, Lower bound) for sixteen sites. The site wise bounds are then applied to particular sites while identifying dies for test

elimination.

4.6 Handling Inverted Test Bounds at Insert1

For a highly correlated test pair it is expected that the upper bound is greater than lower bound. All the tests chosen for analysis as described in section 3.1 have a Kendall's tau greater than 0.5. A single miscorrelated test measurement can sometimes lead to the case where upper bound is lower than lower bound (Inverted bounds).

Inverted bounds at *Insert1* can result in testing all the dies for that test at *Insert2*. Over screening of the good dies at *Insert2* incurs additional test time. The solution to over screening of dies and getting rid of inverted bounds is choosing a different set of bounds.

If in the first iteration of calculation, bounds are inverted, then the process of finding the bounds is repeated. The next greater test data value greater than test limit low is chosen to find the lower bound and the next less value than test limit high is chosen. This removes the dependence on the last significant digit of the test data. The process is repeated until the bounds are not inverted.

A check to get rid of inverted bounds is to find the correlation coefficient for the test data. If tau for a test pair is less than 0.5, then we choose to test all the dies at *Insert2*.

4.7 Categories of Dies for a Test

The upper bound, lower bound and test limits at *Insert2* introduce 9 cells in the scatter diagram of correlated matched tests as seen in Figure 4.9. They can be classified in to different categories as shown in Figure 4.9 below.

	Screened @ Insert 2	Not Screened @ Insert 2	Screened @ Insert 2
Test at Insert 2	Correct Screen	Escape Screen	Correct Screen
	Over Screen	Pass Screen	Over Screen
	Correct Screen	Escape Screen	Correct Screen
	Test at Insert 1		

Figure 4.9: Categories of dies for a test

Figures of merit measure the accuracy with which the the statistical screen predicts that a particular die should be tested at *Insert2* or should not be tested. The prediction outcome is positive if *Insert1* test data correctly predicts *Insert2* fail and negative if *Insert1* test data incorrectly predicts *Insert2* fail.

Pass Screen (PS). Dies whose *Insert1* test data value correctly predicts *Insert2* pass are classified as pass screen. *Insert2* test value of such dies lie within the

bounds and also within the test limits for *Insert2* test. In general these dies would not be tested at *Insert2* and still function correctly at use.

Over Screen (OS). Dies whose *Insert1* test data value incorrectly predicts as *Insert2* fail are the over screen dies. Over Screen dies are misclassified by the screen as bad dies because their test value at *Insert1* lies outside the bounds. Testing these dies at *Insert2* is an additional overhead because of miscorrelation between test data values.

Escape Screen (ES). Dies whose *Insert1* test data value incorrectly predicts them as *Insert2* pass are called as escape screen. These dies have *Insert1* test values within bounds but fail at *Insert2*. In general, these dies are shipped to the customer without additional screening at *Insert2*. Screen escapes contribute to end use defect level.

Correct Screen (CS). Dies which are correctly identified by the screen as bad parts and fail *Insert2* test are called as correct screen. These dies lie outside the bounds of *Insert1* test as well as they lie outside *Insert2* test limits. Correct screens are tested at *Insert2*.

Figure 4.9 shoes categories of dies for a single correlated matched test pair. The dies would be classified in one of the four categories for each correlating test pair. Each die needs to be classified in one of the four categories considering multiple tests in the test flow.

4.8 Categories of a Die for a Test Flow

Figure 4.9 categorizes a die as either a pass screen, over screen, screen escape or a correct screen for every matched test pair in the test flow. Each die needs to be classified in one of the four categories considering multiple tests in the test flow. A precedence order is needed to put a die in one of the four categories based upon entire test flow.

The precedence order by which a die is classified in one of the four categories is:

1. If a die is correct screen for even a single test in the test flow, it is categorized as a Correct Screen. Highest importance is given to correctly screening failing dies at *Insert2*.
2. If a die is escape screen for one test and not a correct screen for any other test in test flow then it is categorized as a Escape Screen. The die is not tested by the correct test at *Insert2* and a faulty die is shipped to customer.
3. If a die is not a screen escape or a correct screen, it is a good die. If a good die is tested at *Insert2* by any test i.e. it is over screen for one of the tests then it is categorized as Over Screen.
4. If a good die is not tested at *Insert2* even by a single test i.e. it is pass screen for all the tests then it is categorized as Pass Screen.

Figure 4.10 shows four dies which are classified as correct screen, escape, over screen or pass screen.

Precedence order	Test 1	Test 2	Test 3	Category
1	Correct Screen	Over Screen	Escape Screen	Correct Screen
2	Pass Screen	Pass Screen	Escape Screen	Escape Screen
3	Pass Screen	Over Screen	Over Screen	Over Screen
4	Pass Screen	Pass Screen	Pass Screen	Pass Screen

Figure 4.10: Categories of a die for test flow

The effectiveness of the feed forward method is measured in terms of end use defect level (failing parts at customer) and test time reduction (tester utilization) in the factory. The precedence order is decided such that the customer is given priority than the factory resources.

4.9 Figures of Merit

The 9 cells shown in Figure 4.9 can be merged in only 4 categories as described above. Figure 4.11 shows the four categories. N1 to N4 represent the total number of dies in each cell.

Figures of merit(FOM) help making decisions about test set points. Performance of the screen is correctly predicting the reduced test flow at *Insert2*. The performance of the screen can be varied by the test limit set point at *Insert2*. The prediction outcome is positive when *Insert1* test correctly predicts a *Insert2* fail and negative when *Insert1* test incorrectly predicts *Insert2* fail.

Test at Insert 2	Screened @ Insert 2	Not Screened @ Insert 2	Total
	Over Screen N1	Pass Screen N2	N1+N2
	Correct Screen N3	Escape Screen N4	N3+N4
	N1+N3	N2+N4	N1+N2+N3+N4
Test at Insert 1			

Figure 4.11: Figures of Merit

Figures of merit are generally designed as not to exceed set point. FOM are defined in the interval from zero to one. If $FOM = 1$ then the screen has the worst performance else when $FOM = 0$ performance of the screen is best.

Figures of merit are defined in terms of Predictive Value (PV) for each of the four categories.

Escape Screen PV is defined on the condition that the dies fail at *Insert2*. It is the ratio of escape screen dies to the total dies failing test at *Insert2*.

$$\text{Escape Screen PV} = \frac{N4}{(N3 + N4)} \quad (4.1)$$

Over Screen PV is defined on the condition that the dies pass at *Insert2*. It is the ratio of over screen dies to the total dies passing test at *Insert2*.

$$\text{Over Screen PV} = \frac{N1}{(N1 + N2)} \quad (4.2)$$

The other two predictive values namely Correct Screen PV and Pass Screen PV are complements of Escape Screen PV and Over Screen PV. The FOM defined in equations 4.1 and 4.2 are used since they are not to exceed limits. Higher the value worst is the performance.

There is a tradeoff between Escape Screen PV and Over Screen PV. Practically both the figures of merit can never go to one or zero simultaneously. When we want to perfectly screen all the failing dies, the test limits at *Insert2* must be very compact. As the test limits at *Insert2* start becoming compact or narrow, more dies are tested at *Insert2* and the number of over screened dies starts increasing. Only if the screen is ideal (hypothetical) both the figures of merit will be zero.

End user defect level or the customer perceived defect level is defined as the fraction of parts failing at the customer to total parts shipped to the customer. From the Figure 4.11, this fraction is defined as:

$$\text{End use defect level} = \frac{\text{Parts failing at customer}}{\text{Total parts shipped to customer}} \quad (4.3)$$

$$\text{End use defect level} = \frac{N4}{N4 + N1 + N2} \quad (4.4)$$

End use defect level should be ideally 0. No bad parts should be shipped to the customer. The end use defect level is generally expressed as number of defective

parts per million (dppm) shipped to the customer. The quality requirements for a product are expressed in terms of defect level in dppm.

Yield loss is not defined with respect to our technique. The technique makes a decision to execute tests or not. It does not make a decision about a die to be a pass/fail or discard failing dies. A good figure of merit which can be similar to yield loss would be the effectiveness of the screen to identify the passing parts at *Insert2*. The quality of screen can be measured as the ratio of over screen dies to the correct screen and the pass screen dies. This figure of merit is not evaluated in this thesis.

The figures of merit of concern to the customer are the Escape Screen PV and the End Use Defect Level. The Over Screen PV and Pass Screen PV describe the effectiveness of the factory and the tester utilization.

The new feed forward adaptive testing method is evaluated by using the datasheet test limits and compact test limits at *Insert2*. The results for both the evaluations are presented in the following sections.

4.10 Evaluation 1: Figures of Merit using Datasheet Test Limits

In this experiment, the datasheet test limits for *Insert2* are used to calculate bounds for *Insert1* tests. The wide datasheet specification limits affect the decision about lower and upper bounds at *Insert1*. The wide limits leads to choosing wide bounds at *Insert1*.

The results are summarized by the figures of merit as seen in the Table 4.4.

Table 4.4 shows that in ADC bin 13 dies which are not screened by the bounds. These are the escape screen not tested at *Insert2* and fail at end use. The end use defect level is expressed in DPPM that is defective parts per million. End use defect level for ADC bin is 16 DPPM and for VREF bin is 4 DPPM. The total adds to 20 DPPM which is much higher than what is expected for an automotive product.

Escape Screen PV value should be ideally 0. The high Escape Screen PV reflects the number of screen escapes.

4.11 Evaluation 2: Figures of Merit using Compact Test Limits

In Experiment 2, compact test limits calculated from all 7 Lots as described in section 3.2. The upper and lower bounds calculated using the new compact test limits screen almost all the ADC bin fails except one.

The Table 4.5 shows two numbers for screen escapes. The screen escapes according to compact test limits are the dies which violate the test limits set by the data. The datasheet screen escapes are the dies which are failing dies according to the datasheet test limits.

For ADC bin, there are 18 dies which are compact fails but only 1 of them is a datasheet fail. The 1 datasheet fail counts towards the end use defect level. VREF bin also has 18 compact fails and 2 datasheet fails.

The new compact test limits results in better bounds at the first insertion. The number of datasheet fails is significantly lower than in evaluation 1 where datasheet

specification limits are used. The compact test limits thus prove to give a better idea about the test bounds at *Insert1* as well as reduce defect level significantly. Using compact test limits, the overall end use defect level is 4 DPPM which is within the limits for automotive product.

4.12 Evaluation 3: Figures of Merit using Compact Test Limits without Lot6

The experiment of cross validation as described in section 4.2 shows that 18 tests from ADC bin had an outlier behavior for Lot6. The compact test limits calculated without including wafers from Lot6 are tighter for these 18 ADC tests. When these limits are applied to data from Lot6 there 970 dies which fail in these 18 parameters.

Adaptive test limits excluding data from Lot6 are used to see the effect of absence of an outlier lot on the number of screen escapes.

Results after evaluation show that there is no change in the end use defect level after the compact test limits excluding Lot6 were applied. The count of over screen dies increased by a few hundreds but there is no significant change in the Over Screen PV.

The fact that there is no increase in DPPM in presence or absence of an outlier lot demonstrates that the method is immune for some extent to variations in data sample.

Table 4.4: Datasheet test limits

	Escape Screen Datasheet Fails	End use defect level DPPM	Escape Screen PV	Over Screen PV
ADC BIN	13/40	16	0.32	0.90
VREF BIN	3/5	4	0.60	0.92

Table 4.5: Compact test limits

	Escape Screen Compact Fails	Escape Screen Datasheet Fails	End use defect level DPPM	Escape Screen PV	Over Screen PV
ADC BIN	18	1/40	1	0.025	0.95
VREF BIN	18	2/5	2	0.4	0.97

4.13 Multisite Adaptive Test

In multisite testing strategy, multiple devices are tested at the same time. For the product, 16 sites are tested in one touchdown of the prober.

A faulty site cannot be replaced from the head of the ATE until the testing of all sites in one touchdown is completed. Individual sites cannot be replaced due to limitations in the operation of the head for the touchdown process [19].

In the feed forward adaptive testing method, every site has a different reduced test flow. If one of the 16 sites is not eligible for reduced test flow because of its response at *Insert1* then all the 16 sites have to remain on the tester for the full test flow. In multisite testing, the reduced test flow for entire set of 16 sites must be decided rather than individual dies due to the limitation of operation of ATE head. The combined test flow is the union of all the tests needed to test the dies.

Table 4.6 shows a scenario where four sites are tested per touchdown. All tests are assumed to be executed in parallel.

	Test1	Test2	Test3	Test4
Site 1	1	0	0	0
Site 2	0	1	0	0
Site 3	0	0	1	0
Site 4	0	1	0	0
One touchdown	1	1	1	0

Table 4.6: Multisite test flow at Insert2

Site 1, 2, 3 and 4 are the four sites which are tested in one touchdown of the tester. Consider that there are 4 tests in the test program (Test1, Test2, Test3, Test4). The test is marked as 1 if it is executed at *Insert2* for a particular site and marked as 0 if it is not executed at *Insert2*. Site1 is tested with Test1, Site2 with Test2 etc. The reduced test flow for the touchdown (4 sites), is the union of all the test flow for each site. The four sites have to remain on the tester head until all the tests for that touchdown are executed.

Test4 is not executed for all the four sites in the touchdown. So it is not present in the test flow for the touchdown. For a test to be eliminated from one touchdown, it should be eliminated for all the sites in the touchdown.

Even though every site has a reduced test flow consisting of only one test, there is no test time reduction when these dies are tested at multisite. As number of sites increase, the reduced test flow starts looking like full test flow. Test time savings expected from reduced test flow are diminished due to the idle time the sites have to sit on the tester.

In multisite testing, a test program can be divided into two parts; parallel tests that can be run concurrently on all sites and serial tests which are executed sequentially over the N sites in a single touchdown [20]. The effects of dividing a multi-site test program into distinct parallel and serial tests is described by the well-known Amdahl's law. Let P be the time for parallel testing and S the time for serial testing. Time to test N dies in parallel with conventional test flow is given by equation 4.5.

$$\text{Test Time for N sites } T_{conv} = P + N \cdot S \quad (4.5)$$

$$\text{Average test Time for 1 site} = \frac{P}{N} + S \quad (4.6)$$

The new reduced test flow for site with N sites will have P' tests in parallel and S' tests in series. The total test time is given as

$$\text{Test Time for N sites with reduced test flow } T_{new} = P' + N \cdot S' \quad (4.7)$$

Test time reduction is the ratio of reduction in new test time (4.7) w.r.t the conventional test time (4.5).

$$\text{Test time reduction(TTR)} = \frac{T_{conv} - T_{new}}{T_{conv}} \quad (4.8)$$

Test time reduction for a wafer is calculated by summation of test time for each site. For M touchdowns of prober on a wafer test time reduction is given by equation 4.9.

$$\text{TTR for a wafer} = \frac{\sum_{i=1}^M (T_{conv} - T_{new})}{\sum_{i=1}^M T_{conv}} \quad (4.9)$$

4.14 Test Time Reduction for ADC, VREF Bin

The feed forward adaptive testing method is applied to all wafers from 7 lots available for the product. Test time reduction at *Insert2* is evaluated. Test time reduction for ADC bin and VREF bin is evaluated separately.

Each test is assumed to have unity test time and all the tests from ADC and VREF bin are assumed to be executed in parallel. Test time reduction is same as the percentage of tests eliminated from the test flow.

Figure 4.12: Test time reduction for ADC bin tests

Figure 4.12 shows the percentage test time reduction for ADC bin tests on the Y-axis and Lots 1-7 on the X-axis. The error bars show the maximum and minimum test time reduction obtained for wafers in each lot. The bold mark shows the average test time reduction for each lot.

The product is tested at 16 sites per touchdown at *Insert2*. Test time reduction when the dies are tested at 1-site and 4-site is evaluated for finding the boundary condition. The three different colors show the percentage test time reduction at 1-site, 4-site and 16-site for all the lots.

Maximum test time reduction is observed when a single site is tested per touchdown. As number of sites per touchdown increases the percentage test reduction decreases. Maximum percentage test time reduction is seen for Lot 4. Reduction

is 82% 1-site, 53% for 4-sites and 23% for 16 sites.

Similar results are seen for the VREF bin tests as shown in the Figure 4.13. The percentage test time reduction obtained for VREF bin tests is less as compared to ADC bin tests.

Figure 4.13: Test time reduction for VREF bin tests

4.15 Effect of Parallel, Serial tests on Test Time Reduction

To study the effect of number of serial tests and parallel tests in test program on the test time reduction, multiple simulations are carried out. All the simulations are carried out for 16-sites at *Insert2*. Figure 4.14 shows the percentage test time reduction for 9 wafers in Lot 7 in five different simulation environments.

Figure 4.14: Test time reduction- serial/parallel Tests

The ADC bin has 44 correlating test pairs which are candidates for elimination at second insertion. The five different simulations are carried out by assuming a varying number of serial and parallel tests in the 44 tests. The least test time reduction is seen when all the tests are assumed to be executed in parallel. Maximum test time reduction is seen when all tests are assumed to be executed serially which is essentially one die tested per touchdown.

As shown in section 4.13, the average test time for one site is given by equation 4.3. If P is the probability of elimination of a test for one site, then the probability of elimination of a parallel test for N sites in a touchdown is P^N . Series tests can be eliminated with probability P ($> P^N$) for each site. As the number of parallel tests increases, the potential test time reduction benefits decrease. The probability of removing parallel test for all the sites in a touchdown is less than a serial test.

If there are majority of serial tests in test program then the test time savings obtained by the method proposed in this thesis as compared to traditional testing is maximum. However, if all the tests in test program are executed are parallel the test time reduction is obtained and it is about 12-15%.

Chapter 5

Conclusion

5.1 Contribution of this work

This thesis introduces a data driven feed forward adaptive technique for test time reduction at wafer sort maintaining the product defect level. Feed forward technique for test time reduction at wafer sort is presented for the first time in this thesis. All earlier known works have developed forward prediction strategies to eliminate tests at final test and at burn-in.

Data driven feed forward technique introduces a statistical screen which analyses test data from first probe of wafer and provides recommendations for test elimination at subsequent insertions. The statistical screen uses correlation between tests at multiple insertions to identify good die. The technique focuses on identifying good dies and testing them with minimum test set.

The decision of elimination of tests is completely driven by the test data. The value of test measurement is not predicted but a decision regarding executing a particular test or not is made by the statistical screen.

A pilot wafer from each lot is tested with all the tests at both insertions at wafer sort. Testing one wafer completely and ensures a valid feedback from test data to adapt to process variations. However, the tester will be idle for the time when test data from first wafer of the lot is analyzed and bounds for the subsequent wafers are selected.

Application of the method to production data show encouraging results. Maximum test time reduction can be obtained if every die is tested serially. The test time reduction benefits go on decreasing as number of sites tested in parallel increase. The defect level using the new adaptive test limits is 1 DPPM for the ADC bin tests and 4 DPPM for VREF bin tests which is acceptable according to industry standards.

Test time reduction benefits using reduced test flow are summarized in the tables 5.1 and 5.3.

ADC Bin	$Site_{conventional}$	$Site_{reduced}$
	1	3.1
	4	5.9
	16	18.2

Table 5.1: Equivalent site count - all tests parallel

Table 5.1 shows the effective number of sites tested using the conventional test flow and reduced test flow for all the tests in the ADC bin. $Site_{conventional}$ are the effective number of sites tested per touchdown using the conventional test flow. $Site_{reduced}$ are the effective number of sites tested using reduced test flow. From the table we can infer that using reduced test flow, test time reduction benefits are equivalent of testing two additional sites per touchdown.

Table 5.2 shows the effective site count when 22 tests from ADC bin are assumed to be executed in parallel and remaining 22 tests are assumed to be executed in series. The effective site count using the conventional test flow for 1, 4 and 16 sites when half the tests are in parallel and half the tests are in series are 1, 1.6 and 1.8.

Using the reduced test flow, the effective site count is almost twice the site count with conventional test flow. Comparing the results when all tests are executed in parallel, test reduction benefits are much higher when only half of the tests in the test flow are serialized.

ADC Bin	$Site_{conventional}$	$Site_{reduced}$
	1	3.1
	1.6	3.2
	1.8	3.8

Table 5.2: Equivalent site count - 22 series tests, 22 parallel tests

5.2 Recommendations and Future work

The feed forward adaptive testing method developed in the thesis is validated only for wafer sort test time reduction. However, this technique can be applied for test time reduction where the wafers are tested at multiple insertions and have correlating tests across insertions. The feed forward of information can be from wafer sort to final test, system test or burn-in if each die can be uniquely identified by a die-id. The application of this technique for test time reduction to final test, burn-in etc. has to be explored.

References

- [1] L. Milor, “A tutorial introduction to research on analog and mixed-signal circuit testing,” in *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 45, no. 10, oct 1998, pp. 1389 –1407.
- [2] J.-L. Huang, C.-K. Ong, and K.-T. Cheng, “A bist scheme for on-chip adc and dac testing,” in *Design, Automation and Test in Europe Conference and Exhibition 2000. Proceedings*, 2000, pp. 216 –220.
- [3] M. L. Bushnell and V. D. Agrawal, *Essentials of Electronic Testing for Digital, Memory, and Mixed-Signal VLSI Circuits*. Boston: Kluwer Academic, 2000.
- [4] T. Williams and N. Brown, “Defect level as a function of fault coverage,” *Computers, IEEE Transactions on*, vol. C-30, no. 12, pp. 987 –988, dec. 1981.
- [5] E. Yilmaz, S. Ozev, O. Sinanoglu, and P. Maxwell, “Adaptive testing: Conquering process variations,” in *Test Symposium (ETS), 2012 17th IEEE European*, may 2012, pp. 1 –6.
- [6] International Technology Roadmap for Semiconductors 2009, Test and Test Equipment. [Online]. Available: <http://www.itrs.net/links/2009itrs/home2009.htm>
- [7] R. Madge, B. Benware, R. Turakhia, R. Daasch, C. Schuermyer, and J. Ruffler, “In search of the optimum test set - adaptive test methods for maximum defect coverage and lowest test cost,” in *Proceedings of International Test Conference*, Oct. 2004, pp. 203 – 212.

- [8] Wikipedia- Ranking. [Online]. Available: <http://en.wikipedia.org/wiki/Ranking>
- [9] M. Kendall, *Rank correlation methods*. C. Griffin, 1955. [Online]. Available: <http://books.google.com/books?id=TJAuAAAAMAAJ>
- [10] 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>
- [11] N. Sumikawa, L.-C. Wang, and M. S. 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.
- [12] N. Sumikawa, D. Drmanac, L.-C. Wang, L. Winemberg, and M. Abadir, “Forward prediction based on wafer sort data x2014; a case study,” in *Test Conference (ITC), 2011 IEEE International*, sept. 2011, pp. 1 –10.
- [13] S. Biswas and R. Blanton, “Reducing test execution cost of integrated, heterogeneous systems using continuous test data,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 30, no. 1, pp. 148 –158, jan. 2011.
- [14] N. Kupp, M. Slamani, and Y. Makris, “Correlating inline data with final test outcomes in analog/rf devices,” in *Design, Automation Test in Europe Conference Exhibition (DATE), 2011*, march 2011, pp. 1 –6.
- [15] M. Chen and A. Orailoglu, “Test cost minimization through adaptive test development,” in *IEEE International Conference on Computer Design*, Oct. 2008, pp. 234 –239.

- [16] Cross-validation (statistics). [Online]. Available: [http://en.wikipedia.org/wiki/Cross-validation_\(statistics\)](http://en.wikipedia.org/wiki/Cross-validation_(statistics))
- [17] Positive Predictive value. [Online]. Available: http://en.wikipedia.org/wiki/Positive_predictive_value
- [18] Negative Predictive value. [Online]. Available: http://en.wikipedia.org/wiki/Negative_predictive_value
- [19] H. Hashempour, F. Meyer, and F. Lombardi, “Analysis and evaluation of multisite testing for vlsi,” *Instrumentation and Measurement, IEEE Transactions on*, vol. 54, no. 5, pp. 1770 – 1778, oct. 2005.
- [20] N. Velamati and R. Daasch, “Analytical model for multi-site efficiency with parallel to serial test times, yield and clustering,” in *VLSI Test Symposium, 2009. VTS '09. 27th IEEE*, may 2009, pp. 270 –275.