Enhanced Testing Performance via Unbiased Test Sets [*]

Li-C. Wang
University of Texas, Austin

M. Ray Mercer
University of Texas, Austin

Thomas W. Williams
System Technology Division, IBM

Abstract

The test generation task involves two separate questions which are: 1) what should the next test be?, and 2) Have enough tests been selected to achieve an acceptable defective part level? Historically, the same fault set (usually the stuck-at-fault set) has been used to answer both questions. When both questions use the same fault set, a statistical bias is introduced to the answer of the second question. In this paper, we propose the use of independent models for answers to the two questions above, and we show, via probabilistic analysis as well as experiments, that the result is a superior test set selection method.

1 Introduction

Testing is performed to weed out the defective parts coming out of the manufacturing process. Traditionally, test generation targets on a specific fault model to produce tests that are expected to identify defects such as unintended shorts and opens. There can be an enormous number of possible defects in a circuit. To do a good job, most of them should be detected by the test set. Since the test generation and the test application are limited by available resources like memory and time, generating tests for all defects is infeasible. Instead, a relatively small set of abstract defects, namely faults, are constructed and these faults are targeted to generate the tests. Usually, the test generation stops when all target faults are detected by the tests produced. With this approach, the test quality relies on fortuitous detection of the non-target defects [[BUTL90] [BUTL91a] [BUTL91b]].

As the quality demands and circuit sizes increase, the effectiveness of test generation on a single fault model becomes questionable. For instance, [KAPU92] showed that for most commonly used model, the single stuck-at fault, the range in defective part levels can spread over several orders of magnitude. [PARK94] did extensive studies on this issue and demonstrated that a high fault coverage was hard to predict an equally high quality.

The weakness of the single stuck-at fault model for obtaining a good test quality reveals that one model may bias the selection of tests to miss some non-target defects. One possibility to remedy this bias is to use more fault models to generate more tests. For instance, [MAX92] presents a combination of functional, I_{DDQ}, and scan tests and reported better results than merely the single stuck-at fault tests. Since each fault model represents a different perspective toward the total defect test space, it is likely that an undetected defect left by one test set can be captured by another.

Here, we attack the problem differently. First, we devise a simple probabilistic model to provide some insight on the testing process. We show analytically that for a circuit with a very large number of defects, the quality of tests generated with specific fault models in mind is limited. Then, from the theoretical model, we propose an unbiased test selection method which can break such limitation and also has the nice property of stably predicting the defective part level. For very large circuits, we show that an unbiased test set is more likely to be better than the set from explicit fault models. In addition to the probabilistic results, we experimented on two benchmark circuits, C432 and C499 (circuits larger than these two are hard to analyze due to their large test set spaces). Encouraging results that strengthen the theoretical analysis were obtained.

Plan

Section 2 explains the concept of fortuitous detection of defects in detail. A review of the William-Brown model for defective part level prediction is pre-

---

[*] This work was supported in part by the Semiconductor Research Corporation under contract SRC 93-DP-142

sented. In section 3, we illustrate the probabilistic model used to understand the testing process. Based on the result derived from the model, we address the question of how to improve the test quality. We also demonstrate in theory that using fault dropping is unlikely to produce desired high test quality. In section 4, the tool and the method for experiments are described. Additional data with respect to [PARK94] are reported in section 5, which demonstrate the insufficiency of using the single stuck-at fault model for testing and the infeasibility of using fault coverage in place of the defect coverage on defective part level prediction. An unbiased test selection method is proposed in the following section. There, we compare the effectiveness between the new idea and the traditional one using fault models via the probabilistic model constructed in section 2. Experimental results on the unbiased tests are reported in section 7.

2 Basic Concepts

A defect is a flaw in a circuit. A fault model is a hypothesis of how defects affect the circuit behavior. Given a fault model, a set of faults is derived, called target faults. Then, tests are generated on these faults. Usually, target fault coverage is used as an estimator for defect coverage. Defects are categorized into those which can be mapped directly onto modeled faults and the others which cannot. We call the former target defects, and the later non-target defects. Since detecting a particular non-target defect is not ensured, the accuracy of the estimation for defect coverage depends on fortuitous detection of the non-target defects. Note that there can be an enormous number of defects in a circuit. While lacking a clear uniform model to capture all defects, to study the fortuitous detection, in practice we assume surrogates. Surrogates are different faults from those target ones for test generation, which represent a different perspective of possible defects. In our study, we used two sets of surrogates, non-feedback AND bridging and transition (gate delay faults). We chose non-feedback AND bridging faults since feedback bridging faults are easier to detect [[MILL88] [MEI74]].

There are three possible relationships between tests for target faults and those for non-target defects. Figure 1 illustrates the three cases. For instance, F2 dominates S3, F1 overlays S1, and S2 is disjoint from both F1 and F2. s3 is ensured to be detected by a detection of f2, s1 can be fortuitous detected by a test in F1, and s2 cannot be detected by any target fault test. The interaction among the target fault test spaces and surrogate test spaces affects the accuracy of the estimation for non-target defect coverage using target fault coverage.

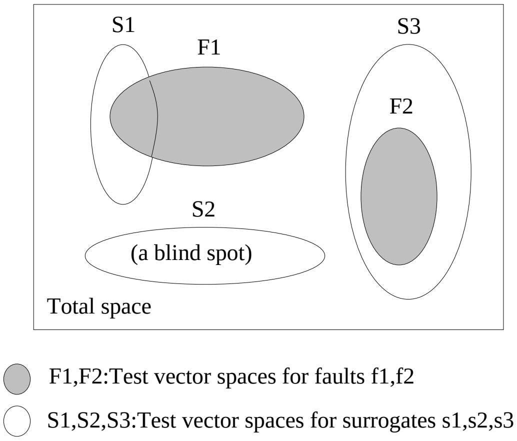

Figure 1: An example of distribution of test spaces for target faults and surrogates

Test generation involves mainly two issues — the selection of tests on target faults and the number of tests selected. For selection of tests, a good test should be able to detect more defects. Without knowing the test spaces for defects, this goal is hard to achieve. For the test size, target fault simulation is used to compute the current fault coverage and usually test generation stops when a criteria like 99% coverage is met. During the process of test generation, for the case with fault dropping, only those target faults currently undetected are considered for the generation of the next test, and for the case without fault dropping, each test is generated for every target fault. There are two drawbacks with the single fault model approach for test generation. First, the selection of tests biases toward the target test space so some test space is not reachable and some others have reduced chances to be selected. Second, if 100% fault coverage does not produce an adequate quality, the next step remains in question.

Our unbiased test selection method attacks both issue at the same time. Before we get into that, let us review a method for defective part level prediction. William and Brown [WILL81] has the following model

$$DL = 1 - Y^{(1-d)}$$

where DL is the defective part level, d defect coverage, and Y yield. Usually, the yield comes from empirical data on manufacturing process. In practice, fault coverage f is used in place of the defect coverage. For us to be able to achieve desired defective part level using f, it requires $f \approx d$. If f differs from d which is usual for single stuck-at fault model, then even f being close to 1 cannot ensure a particular DL.

3 A Simple Probabilistic Model for Testing

In general, let us assume N defects, n target faults, and a number m which represents the average number of defects detected by a test generated against a target fault. Also assume that there is no fault dropping involved so that the number of tests generated is n. Then, we can model the detection of defects by the group allocation problem defined below.

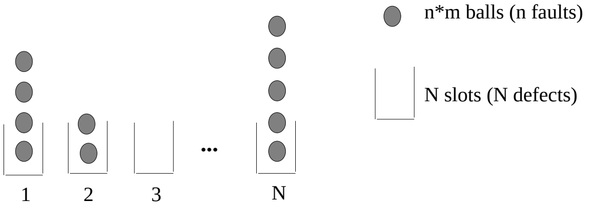

Figure 2: Illustration of the Group Allocation Problem

Definition 1 Group Allocation Problem *Suppose n groups of m balls are allocated independently in N slots such that for each group all* $\begin{pmatrix} N \\ m \end{pmatrix}$ *ways of allocating the balls are equiprobable, i.e. for each group allocation, no two balls go to the same slot. Then, we are interested in knowing the number of empty slots after allocating all n groups of balls.*

[PARK81] showed that the asymptotic distribution of the number of empty slots is normal with mean $N(1-\frac{m}{N})^n$ for sufficient large N, n and a fixed n/N. In particular, the slot, the group, the ball correspond to the defects, the test and one of the defects detected by a given test, respectively. Figure 2 illustrates the setting of the problem. From the result of this problem, we have

$$\text{Ave. \# of undetected defects} = N(1-\frac{m}{N})^n \qquad (1)$$

For the case of generating tests based upon a target fault model, assume N_1 faults and N_2 defects. Because of the targeting, the average number of detected faults by a test, denoted as m_1 is different from the average number of detected defects, denoted as m_2. We note that $m_1 \geq 1$ and if each fault models at least one defect, then we can also say that $m_2 \geq 1$. Further, let $l = \frac{N_2}{N_1}$ so that if all defects are targeted, $l = 1$ and in general l grows toward N_2 as fewer faults are sampled. In practice, we can assume $l \geq 1$ since normally the size of the non-target set is much larger than the size of the fault set. Since all faults are targeted, at the end of testing when 100% fault coverage is achieved, some non-target defects can be left undetected and if without fault dropping, N_1 tests should have been applied. Then for sufficient large N_2 and a fixed m_2, by replacing N, m and n in equation 1 by N_2, m_2, and N_1, we have the number of defects left undetected as (DC denotes the defect coverage)

$$N_2(1-DC) = N_2(1-\frac{m_2}{N_2})^{N_1}$$

$$= N_1 l(1-\frac{m/l}{N_1})^{N_1} \sim l\frac{N_1}{e^{m/l}} = N_2 e^{-\frac{m}{l}} \qquad (2)$$

$$1 - DC \sim e^{-(m/l)} \qquad (3)$$

We note that equation 2 holds for the case of multiple fault models as well by replacing m_2 with its average value over all models and N_1 the sum of all target faults.

Implications from the probabilistic model

Based upon equation 2, the obvious way to improve the test quality is to reduce l. Note that m_2 depends on the fault model used but how to select a good model to produce a larger m_2 is unclear without knowing the the defect test space. Since a particular fault model has its "blind spot" over the whole test space as shown in figure 1, continuing test generation on the same fault model after a high fault coverage is less effective than starting a new one. With a fixed N_2, generating more tests via a new fault model implies a larger N_1 and a smaller l. [MAX92] has demonstrated the success following this approach.

Equation 2 has another implication. That is by using William-Brown model and replacing the term $(1-DC)$ by its asymptotic approximation $e^{-(m/l)}$, we obtain

$$DL = 1 - Y^{(e^{(-\frac{mN_1}{N_2})})} \qquad (4)$$

Many researchers have suggested different refinements for DL with respect to the William-Brown model, for instance [[AGR82] [SETH84]]. The difference between equation 4 and those in [[AGR82], [SETH84]] is that no assumption is made by us about the distribution of fault occurrence. Our purpose here is to study the testing behavior and hence nondeterminism was applied to the quality of tests.

Since we expect that for target faults, the average number of "detectability" is larger than that for defects, as the fault coverage approach 100%, the defective part level should be higher than that predicted by the William-Brown model by misusing fault coverage as the defect coverage. More formally, Let T, N be the fault and defect sets and $N_1 =| T |$, $N_2 =| N |$. Let n be the number of tests generated for T. For each test applied, again let m_1, m_2 be the mean numbers of faults/defects detected in sets T, S, respectively. Then, for sufficient large N_1, N_2 and given n, m_1, m_2, using equation 2 we compute the fault coverage as

$$FC = 1 - e^{-(m_1 n/N_1)} \tag{5}$$

and defect coverage (DC) as

$$DC = 1 - e^{-(m_2 n/N_2)} \tag{6}$$

Then, using the William-Brown model, we can obtain the defective part level with respect to $(1 - FC)$. By changing n, this gives us a curve showing how effectiveness applying n tests varies according to the given m_1, m_2.

Definition 2 *We define the detectability for T to be $d_t = m_1/N_1$, for N to be $d_n = m_2/N_2$, and the* **test effectiveness** *(ε) for a test set to be $\varepsilon = \frac{d_n}{d_t} = \frac{m2N_1}{m1N_2}$.*

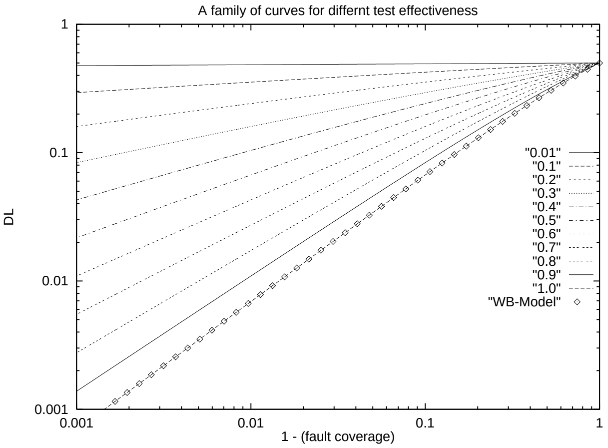

Figure 3: A family of effectiveness curves, $\varepsilon = 0.01 \ldots 1.0$

Normally, we should have $\varepsilon \leq 1$. Figure 3 shows a family of curves as ε ranges from 0.01 to 1. We obtained this figure by assuming that $N_2 = 5000, N_1 = 500$ and m_1 is fixed at 2. Then, ε depends on only m_2. For $\varepsilon = 1$, $m_2 = 20$, $\varepsilon = 0.9$, $m_2 = 18$, and so on. We also assumed that there is no limit on the number of tests applied so that fault coverage can be as close to 100% as possible at the end. Two things can be observed from the figure. First, as the **test effectiveness** reduces, the defective part level increase, i.e. the tests do a worse job (consider the top curve with $\varepsilon = 0.01$). Second, when $\varepsilon = 1$, the curve matches to the William-Brown model, which under the assumption $\varepsilon \leq 1$ represents the optimal case. We call a test set with $\varepsilon = 1$ the **unbiased tests**. In section 5, we will develop a test selection scheme based upon the idea of the unbiased test set.

[PARK94] demonstrated the questionability of fault dropping by experiments on C432. Here, we will address the question using the probabilistic results obtained. When fault dropping is involved in testing, we can no longer model the fault coverage as in equation 5. This is because the behavior of testing becomes more deterministic, i.e. at least a new fault will be taken out with each test application. Without fault dropping, a newly applied test can detect no faults currently left due to the fact that the particular fault targeted by that test was fortuitously detected by an earlier test(s).

Hence, we modify 5 as the following, assuming $m_1 > 1$ and n tests, $n \leq N_1$, are applied. Note that m_1 can be any real number close to 1.

$$FC = \frac{[N_1 - N_1 e^{-((m_1-1)n/N_1)}] + n}{N_1} \tag{7}$$

The numerator in equation 7 consists of two terms, n which accounts for the detection of at least n undetected faults and $N_1 - N_1 e^{-((m_1-1)n/N_1)}$ for fortuitous detection of other faults. We use $m_1 - 1$ instead of m_1 in the second term since one fault has been counted deterministically by the term n for each test. Since $0 \leq FC \leq 1$, there is an upper bound imposed on n, the number of tests applied, by equation 7. Suppose $n = N/q$ where $q \geq 1$ is a real number. By a straightforward calculation, we obtain $q \ln(q) > m_1 - 1$ or in other word, $q^q > e^{(m_1-1)}$. Hence, for a large m_1, i.e. an easily detected set of target faults, q should be large and n will be small. With a smaller n plugged into equation 6 for defect coverage, we see that DC becomes smaller, i.e. the quality becomes worse. Therefore, we can conclude that with fault dropping, the easier the faults to be detected, the worse the testing quality will be.

Using equation 7 for the case of fault dropping, we drew the defective part level curve and showed in figure 4 with two other curves with $\varepsilon = 0.7$ and 1 for comparison. These three curves represent three different models for testing — fault dropping, no fault dropping (with $\varepsilon = 0.7$), and unbiased (with $\varepsilon = 1$).

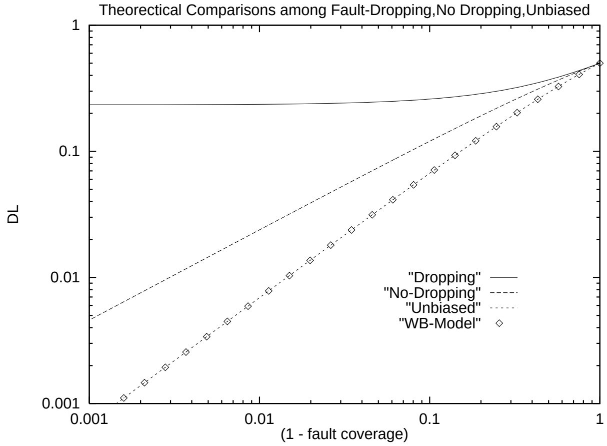

Figure 4: Comparison in theory among fault dropping, no dropping, and unbiased

4 The Tool and the Method for Experiments

Ordered binary decision diagram (OBDD) [BR86] was used to compute the whole test space for each target fault. Due to the large memory requirements, currently OBDD is unable to handle circuits larger than C499. Since our goal was to study the testing behavior in general, test generation should not be biased toward a particular algorithm or implementation. Instead, thousands of test sets are computed and the average results and their variances were obtained for each case study. This procedure was done for test generation using the single stuck-at fault model, as well as on the unbiased selection of tests described later. Average results are used when comparing these two.

To evaluate the test quality, we are interested in knowing the defective part level resulting from a test set. As described before, we assumed surrogates. Suppose that the number of surrogates assumed is N and the number of faults considered is n. After testing, let N' and n' be the surrogates and faults detected. We compute defect coverage to be the surrogate coverage and use the William-Brown model to approximate the "real" defective part level with respect to the surrogate coverage. Although this is not necessary the true defective part level, it is a good estimator for the purpose of studying testing behavior since usually, surrogates perceive a different test space from faults, where this difference is the main reason for the reduction of test quality. To obtain the defective part level, we also assumed that the yield is 0.5 for all experiments.

5 The Limitations on Single Stuck-at-fault

Previously, [PARK94] studied the limitations of the single stuck-at fault model. Their conclusions about the single stuck-at fault model include 1) test quality varies as fault coverage approaches 100%, 2) fault dropping is questionable, and 3) using fault coverage in place of defect coverage with the William-Brown model is also questionable. Their results were obtained using C432 and non-feedback bridging AND surrogates. Here, we extend their studies by including transition surrogates (gate delay faults). Besides, we implemented a more efficient OBDD tool to analyze C499. In addition to confirming their conclusions, we show that transition surrogates consistently result in a higher uncertainty with respect to test quality and test prediction than bridging surrogates. This is expectable since usually the transition test space is larger (which require two tests for a fault). Figures 5 to 8 show the results.

Assuming no fault dropping, for C432 there are 524 stuck-at fault tests and for C499 there are 758 tests. 650 and 1162 bridging surrogates are randomly picked for C432 and C499, respectively. The numbers of transition surrogates for C432 and C499 are 320 and 404 which are twice the numbers of gates since each gate output can have a slow-to-rise and a slow-to-fall transition. We chose the number of bridging surrogates $1162 = 758 + 404$ for the purpose of evaluating the unbiased tests presented later. 650 bridging faults for C432 are selected to match that in [PARK94]. As we can observe from figures 5 to 8, previous conclusions are confirmed on the new surrogates and circuit. Besides, if we compare figures 5 to 7 with the test effectiveness figure 3, we observe that the curves in figures 5 to 7 go from high effective region to low effective region. This indicates that as the fault coverage approach 100%, the effectiveness of a test to detect non-target defects reduces.

6 An Unbiased Test Generation

In section 3, we defined the **test effectiveness** as $\varepsilon = \frac{d_n}{d_t} = \frac{m_2 N_1}{m_1 N_2}$ (see Definition 2). We see there that an unbiased test set ($\varepsilon = 1$) is a better choice to enhance test quality. Based on the test generation with specific fault models in mind, we need to ensure $d_t \approx d_n$ to obtain a good quality. Without knowing the defect test space, this can be very hard. Instead, in this section, we propose a different approach from trying to refine the fault models in use.

Remember that for an input vector to be a test for a defect, this vector must excite the defect (produce a difference at the defective site from its normal behavior) , as well as make the difference observable at one of the primary outputs. Without knowing how a defect can change the logical behavior at the local site (for instance, a logical gate), it is hard to achieve the excitation criteria. However, no matter what excitation requirements are, the observation criteria remain the same. To achieve an unbiased testing, we propose to use a combination of *random excitation* and *deterministic observation*. In other words, there will be no effort made in test generation to ensure fault excitation. Only the observation condition is satisfied on each test, and since no fault model is used, a good test will be evaluated based upon the number of observable sites.

What advantages do we gain by taking this approach? First, the test generation can be easier because one of the requirements has been dropped. Now assume that each time, C sites are observed by a test, and with a probability p, a defect on an observed site is detected. Note that p should depend on the fanin to the site and the defect type. Also assume that n tests are generated and $N = N_1 + N_2$ defects are presented, where N_1 is the number of defects to be used for test quality evaluation and N_2 are all others. For test quality evaluation, we mean that the set is used to compute the fault coverage so that test generation stops when a preset number is reached. Then, we have

$$\text{Average number of defects left} = N(1 - \frac{C \cdot p}{N})^n \quad (8)$$

Suppose we will generate tests until N_1 defects are all detected. In this case, we have

$$N_1(1 - \frac{C \cdot p}{N_1})^n < 1 \Longrightarrow n > \frac{-\log(N_1)}{\log(1 - Cp/N_1)} = n_s \quad (9)$$

We can plug in n_s to obtain the formula $N_2(1 - \frac{C \cdot p}{N_2})^{n_s}$ as the prediction of the number of defects left. Assuming $N_1 \times l = N_2, l \geq 1$, the closer the N_1 is to N_2, the better the test quality will be. From equation 8 and 9, it is not hard to derive that the number of defects left when all N_1 defects are detected approximates l. Comparing this result with equation 2 before, we see that the difference comes from the term $\frac{N_1}{e^{(m/l)}}$. If l is fixed as N_1 and N_2 grows, equation 3 implies the quality of testing remains fixed given a fixed fault model. However, the quality of the unbiased testing becomes better. Fixing l while N_2 grows requires a larger number of sampling faults. Since usually, the number of faults is in the linear order of the circuit size but the number of defects is in a higher order, keeping l fixed is hard. Suppose N_1 is fixed instead, which equals to the maximal test length we can afford. Then, as N_2 grows, equation 3 says that the test quality becomes worse, but an unbiased testing can deliver a fixed testing quality, i.e. $1 - DC = \frac{1}{N_1}(\frac{l}{l+1}) < \frac{1}{N_1}$. In both cases discussed above, unbiased testing is better than that based upon explicit fault models.

From the viewpoint of test effectiveness in Definition 2, it is more likely for the unbiased testing to achieve $d_t = d_n$ since now T, the target fault set, can be selected freely to cover a wide spectrum of the perspectives toward the whole test space. We note that the selection of T is independent of the test generation process. T only tells when to stop the process.

We conclude this section by making a final remark. Note that pure random testing is also likely to achieve a test effectiveness close to 1. However, the term Cp implied by a purely random testing process can be very small. As a result, an enormous number of tests will be generated to obtain the desired quality. In practice, this is not economically feasible.

7 Experiments

We obtained the unbiased tests by computing the whole boolean difference space for each site using the OBDD tool we built. As usual, thousands of test sets are selected randomly. However, for each test set, we ensure that if possible, every site should be observed at least once. The cardinality of the test set is equal to the number of tests generated before using the single stuck-at fault model. Again, the average results and their variances for defective part level curves were plotted. The average results are presented in figures 9 to 10, with those average curves from stuck-at fault tests and curve from William-Brown model for comparison. One question which remains is "What target fault set should we select to compute the fault coverage?". In the figures, for C432, the fault coverage is obtained using the single stuck-at fault model. For C499, when the defect is bridging, the fault coverage is based upon stuck-at and transition faults, and when the defect is transition, upon stuck-at and bridging faults.

Figure 9 to 12 confirm our previous conclusion that an unbiased test set should be superior in terms of reducing the defective part level. If we compare those figures with the effectiveness curves shown in figure 3, we see that the unbiased approach is more effective (closer to 1) in all cases. Perhaps the most surprising result is the case when using transition surrogates for C432, where a test effectiveness greater than 1 is observed.

As shown in section 5 as well as in [PARK94], test quality results using the stuck-at fault tests varies as the fault coverage approach to 100%. This situation is much worse for transition surrogates. In figures 13 and 14, we compare the variances for the stuck-at fault tests and the unbiased tests using transition surrogates. For each testing approach, 2 curves were drawn in each figure, where the upper curve is obtained by computing $Mean + 2 \times StandardDeviation$ and the lower one $Mean - 2 \times StandardDeviation$. In both figures, the variance of quality (defective part level) using the unbiased tests is consistently smaller than the single stuck-at fault approach. By this, we conclude that the unbiased selection of tests can reduce the uncertainty in testing quality as well.

8 Conclusion

Many researchers have questioned using a single fault model, in particular the single stuck-at fault model, on test generation for high quality testing. The conclusions made before by experiments include 1) fault dropping can degrade test quality seriously, 2) testing quality can vary significantly for a given high fault coverage, and 3) multiple fault models may be required to get a desired testing quality. In this paper, we model the testing process by the Group Allocation Problem, and provide, for the first time, mathematical explanations to the three conclusions above. From our probabilistic analysis, we propose an unbiased testing method of which the test effectiveness is higher for high quality testing of large circuits. The superiority of the unbiased testing, predicted by our analysis, is then confirmed by extensive experiments on benchmark circuits.

References

[AGR82] Agrawal, V. D., Seth, S. C. and Agrawal, P., “Fault Coverage Requirement in Production Testing of LSI Circuits”, IEEE Journal of Solid-State Circuits, Vol SC-27, 1982, pp. 57-61.

[BR86] R.E. Bryant, “Graph-based algorithms for Boolean function manipulation,” IEEE Trans. on Computers, vol C-35, no. 8, Aug. 1986, pp. 677-692.

[BUTL90] K. M. Butler and M.R. Mercer, ”The Influences of Fault Type and Topology on Fault Model Performance and the Implications to Test and Testable Design,” *Proc. 27th ACM/IEEE Design Automation Conference*, Orlando, FL, June 24- 28, 1990, pp. 673-678.

[BUTL91a] K. M. Butler and M. R. Mercer, *Assessing Fault Model and Test Quality*, Kluwer Academic Publishers, 1991, ISBN 0 - 7923 - 9222 - 1.

[BUTL91b] K. M. Butler and M. R. Mercer, ”Quantifying Non- Target Defect Detection by Target Fault Test Sets,” *Proc. of The European Test Conference*, Munich, Germany, April 10- 12, 1991, pp. 91-100.

[KAPU92] R. Kapur, J. Park, and M. R. Mercer, ”All Tests for a Fault are Not Equally Valuable for Defect Detection,” *Proc. 1992 International Test Conference*, Baltimore, MD, September 20-24, 1992. pp. 762-769.

[PARK94] Jaehong Park, Mark Naivar, Rohit Kapur, M. Ray Mercer and T. W. Williams, ”Limitations in Predicting Defect Level Based on Stuck-at Fault Coverage,” the VLSI Test Symposium 1994.

[MAX92] Maxwell, P. C. and Aitken, R. C. “ I_{DDQ} Testing as a Component of a Test Suite: The Need for Several Fault Coverage Metrics”, Journal of Electronic Testing 3, 1992, pp. 305-316.

[MEI74] Mei, K.C. “Bridging and stuck-at faults”, IEEE Tran. on Computers, Vol C-23, no. 7, 1974.

[MILL88] Millman, S. D. and McCluskey, E. J. “ Detecting bridging faults with stuck-at test sets”, Proc. International Test Conference 1988.

[PARK81] Park, C. J. “On the distribution of the number of unobserved elements when M-samples of size N are drawn from a finite population”, Comm. Statist. – Theory and Method, A10(4), 1981, pp. 371-383.

[SETH84] Seth, S. C. and Agrawal, V. D. “Characterizing the LSI Yield from Wafer Test Data”, IEEE. Tran. on CAD, Vol. CAD-3, 1984, pp. 123-126.

[WILL81] William, T. W. and Brown, N. C. “Defect level as a function of fault coverage”, IEEE Trans. on Computers, Vol. C-30, no. 12, 1981.

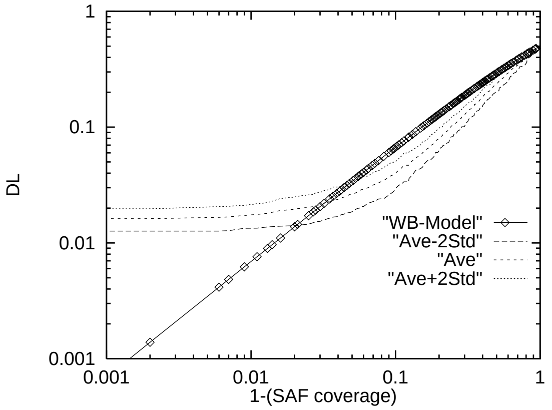

Figure 5: Results on bridging surrogates, C432 (Ave is the mean, Std is the standard deviation)

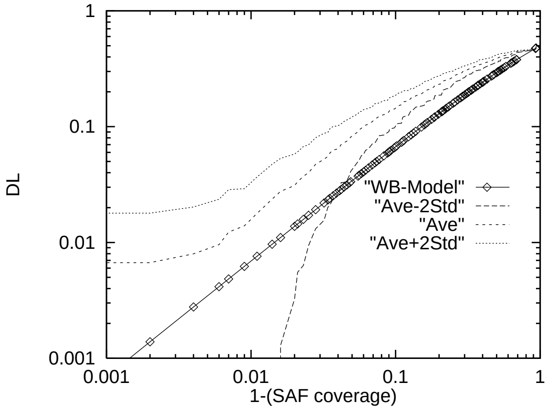

Figure 6: Results on transition surrogates, C432 (Ave is the mean, Std is the standard deviation)

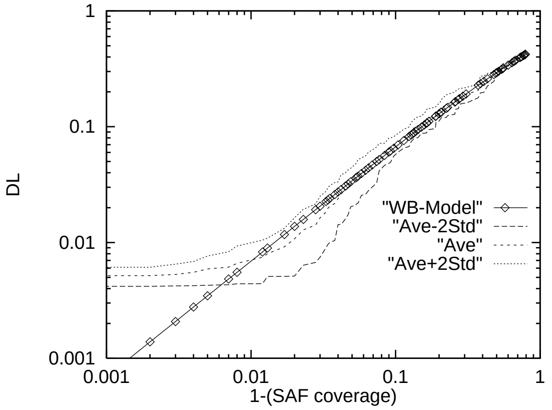

Figure 7: Results on bridging surrogates, C499 (Ave is the mean, Std is the standard deviation)

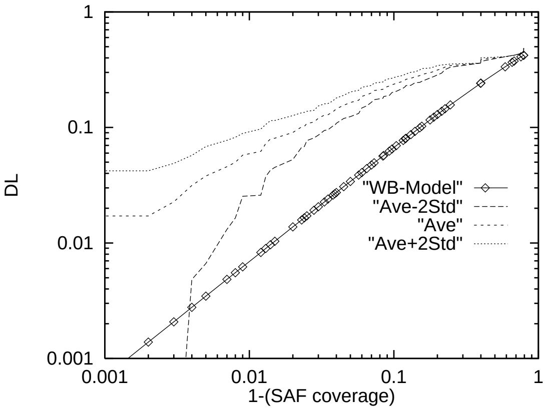

Figure 8: Results on transition surrogates, C499 (Ave is the mean, Std is the standard deviation)

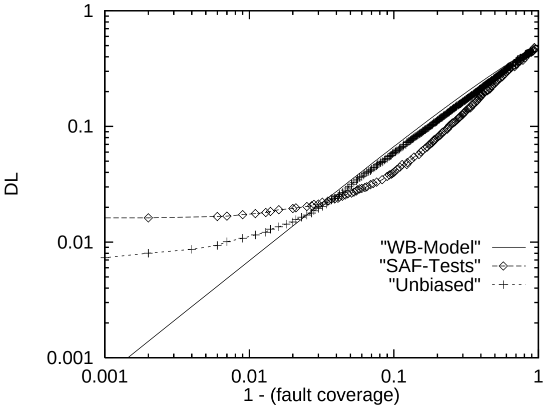

Figure 9: Comparison between unbiased and SAF tests using bridging surrogates and C432

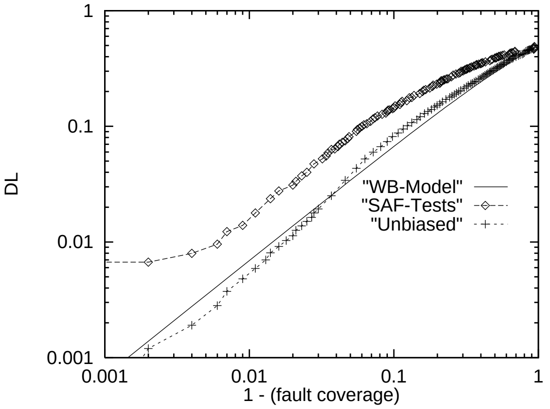

Figure 10: Comparison between unbiased and SAF tests using transition surrogates and C432

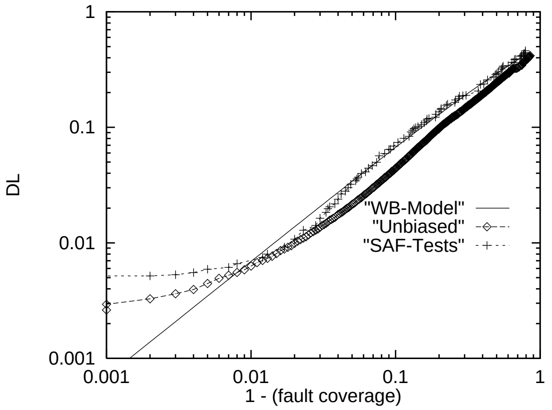

Figure 11: Comparison between unbiased and SAF tests using bridging surrogates and C499

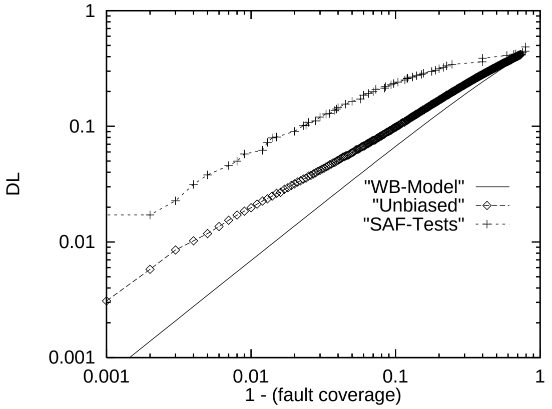

Figure 12: Comparison between unbiased and SAF tests using transition surrogates and C499

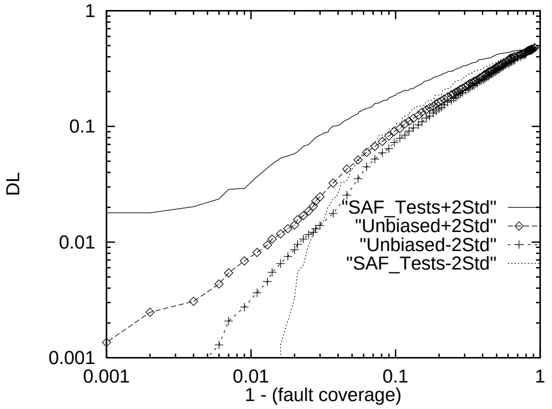

Figure 13: Comparison of test quality uncertainty between unbiased and SAF tests using transition surrogates and C432

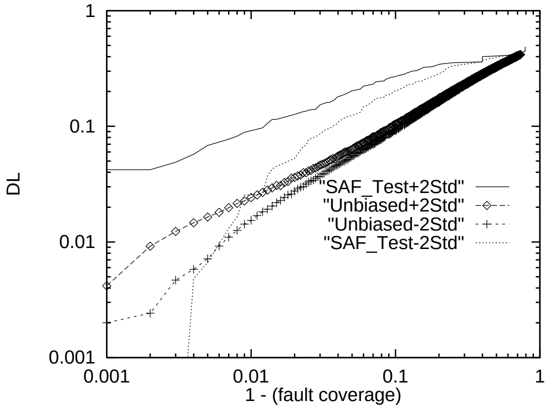

Figure 14: Comparison of test quality uncertainty between unbiased and SAF tests using transition surrogates and C499