Modeling the Probability of Defect Excitation for a Commercial IC with Implications for Stuck-at Fault-Based ATPG Strategies *

Jennifer Dworak, Michael R. Grimaila, Sooryong Lee,
Li-C. Wang, and M. Ray Mercer

Department of Electrical Engineering, Computer Engineering Group
Texas A&M University, College Station, Texas 77843-3128
E-mail: mercer@ee.tamu.edu

Abstract

If many potential defects exist at each site in an integrated circuit, then as the number of applied test patterns increases, the number of defects which remain undetected decreases monotonically. Modeling this rate of decline in defective part level is a non-trivial problem. We show that the number of times each site is observed serves as a significantly superior basis for modeling this phenomenon– when contrasted with the number of faults detected. This "site observation-based" predictor not only increases the accuracy of defective part level prediction, it also provides the first quantitative method for comparing the effectiveness of various ATPG strategies to reduce the defective part level.

1 Introduction

The ultimate objective of manufacture testing is to reduce the defective part level (DL) from the value resulting at the end of the manufacturing process (1 - Yield) to a level which is acceptably low for a customer's application. Today, a typical customer requirement for DL would be on the order of 200 defective integrated circuits (chips) per million (delivered to the customer).

Every defective digital integrated circuit has at least one defect located some place on that chip. We define a site to be any location where a single stuck-at fault can be conveniently modeled, and we assume that there are many different defects which can produce an error at any given site. An error due to one of these defects appears at a given site only if that defect is excited, and that error can be detected on the tester only if the error at that site is observed. If a given test pattern detects a given defect, then all chips which contain that defect are discarded, and future patterns which detect only that defect will not further lower the defective part level (DL). With these assumptions, it is reasonable to expect that in the limit the number of undetected defects at each site will approach zero, and therefore DL will asymptotically approach (but never actually reach) zero.

The model above differs from all of those in current use today. For example, the Williams-Brown [WILL81] model is different because it assumes that only one defect exists at each site. In this case, the probability of a defect remaining at a site after the corresponding stuck-at fault is detected drops to zero. Therefore, the defective part level will be zero when the single stuck-at fault coverage is 100%. Except for this difference, the two models are based upon a very similar set of assumptions, including "the initial probability of a defect occurrence at every site is equal, and all of these probabilities are statistically independent."

Defective part level models traditionally have been used in two ways. First, a given test pattern set has been analyzed to predict whether the resulting DL will be sufficiently low. For example, Williams-Brown [WILL81] proposes that a target DL infers some minimal stuck-at fault coverage which must be achieved by the test pattern set. Second, two test pattern sets can be compared by predicting their resulting DL values, and the superior test pattern set can be selected over another which is inferior with respect to predicted defective part level [GRIM99]. This second application allows test pattern sets to be optimized by repeated comparison of the "current best test pattern set" and some minor variation which becomes the "possibly superior test pattern set."

In this paper, we introduce a third application for

*This work was supported by the Texas Advanced Technology Program – Project No. 036327-152.

our DL model. Because of the character and superior accuracy of our novel DL predictor, it is possible to quantitatively compare the effectiveness of competing ATPG strategies (in terms of their ability to produce test pattern sets which effectively reduce DL). Thus, for the first time, we introduce a tool to quantitatively compare ATPG strategies in terms of their ability to produce patterns which quickly reduce the defective part level.

2 A Commercial IC and Two Different Test Pattern Sets

In the first part of the paper, we focus on results from an actual commercial experiment which produced actual tester data. The commercial integrated circuit, consisting of more than 75,000 two input NAND equivalent logic gates, was tested using two different test pattern sets. We will designate the first test pattern set as COMMERCIAL because it is exactly what is used in the standard manufacturing test flow. Because the chip was designed with 100% scannable flip-flops, ATPG was performed using the Mentor Graphics FASTSCAN program. The COMMERCIAL test pattern set consisted of approximately 3,000 test patterns where each test pattern was applied using one scan chain load/unload. The stuck-at fault coverage for this test pattern set was just above 96.7%.

We will designate the second test pattern set as RESEARCH because it was produced using a new ATPG method we call REDO (Random Excitation and Deterministic Observation) [GRIM99]. The RESEARCH set of test patterns was also produced by the Mentor Graphics FASTSCAN program; the length of this test pattern set was exactly the same as the COMMERCIAL set, and its stuck-at fault coverage was 97.0%. However, the RESEARCH set differed from standard commercial practice in that the number of observations at sites which were difficult to observe was maximized.

In the commercial experiment, 6,986 die passed all parametric tests, and these were tested using the two test patterns sets described above; 220 were declared defective using the COMMERCIAL test pattern set, and 229 were declared defective using the RESEARCH test pattern set. If we arbitrarily assume that two defective die were never detected by either of the test pattern sets, then the two plots for DL as a function of the number of test patterns applied is shown in Figure 1.

The remainder of this paper focuses on methods for accurately predicting DL data such as that shown in

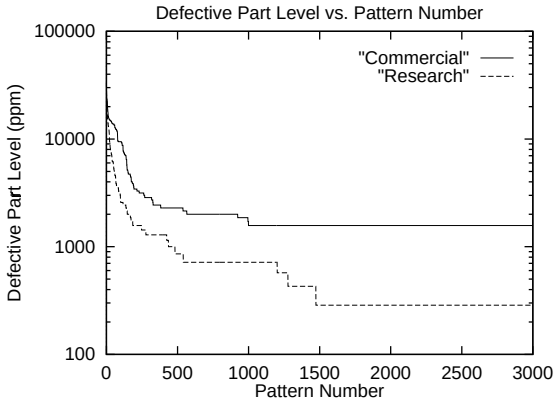

Figure 1: Comparing measured defective part level for two ATPG methods

Figure 1.

3 MPG-0 Defective Part Level Model

Since every site can be affected by many defects, the MPG(Mercer, Park, Grimaila) model assumes that the chip's defective part level can be calculated based upon the defects at each individual site. In particular:

$$DL \cong 1 - \prod_{i=1}^{Sites} [1 - (1 - P_{EXCITE})^{\#OBS_i} \times P_{DEFECT}] \quad (1)$$

P_{EXCITE} is the probability that at least one (as yet undetected) defect at site i is excited by one test pattern. The term $(1 - P_{EXCITE})^{\#OBS_i}$ gives the probability that no new defect is excited and observed after $\#OBS_i$ different test patterns have produced an observation of site i. A defect is missed if no new defect is excited and observed ($(1 - P_{EXCITE})^{\#OBS_i}$) and if the defect actually occurs (P_{DEFECT}). It is very reasonable to assume that these events are statistically independent, so the joint probability is simply calculated as the product of the two events. The result is the probability that a defect remains undetected at site i. One minus this quantity gives the probability that no defect remains undetected at this site. Assuming that each of these probabilities for every site is independent, then the product of these probabilities over all sites is the probability that the chip is defect free. DL is simply one minus that probability.

Values at each site for $\#OBS_i$ can easily be obtained by single stuck-at fault simulation of the input test pattern sequence against a logic level description of the chip being analyzed. According to this model, more observations of a site will exponentially reduce

the probability that some defect at that site remains undetected. Thus, a significant new goal for ATPG is to increase the number of observations achieved for sites which are not regularly observed. For example, primary outputs will be observed by every test pattern, and the chances of a defect remaining undetected at a primary output site are vanishingly small. In contrast, some sites deep within the circuit are observed very rarely, and chances of a defect remaining undetected there are much higher.

We assume that at every site, the probability of a defect's existence is the same constant value (P_{DEFECT}), and we calculate this probability from the chip yield using:

$$Yield = (1 - P_{DEFECT})^{\#SITES} \tag{2}$$

Thus,

$$P_{DEFECT} = 1 - (Yield)^{(1/\#SITES)} \tag{3}$$

The central focus of this paper is selection of the proper form for P_{EXCITE} so as to maximize the accuracy of the DL predictions produced by the MPG model. In the past, we have assumed this term to be a constant, and we will characterize that model as MPG-0. For comparison, we use the Williams-Brown model which predicts DL based upon fault coverage:

$$DL = 1 - Yield^{(1-FaultCoverage)} \tag{4}$$

Figure 2 compares predictions by the Williams-Brown model with those from the MPG-0 DL model for the case of our experimental chip. Note that neither of the models provide a very close match to actual data.

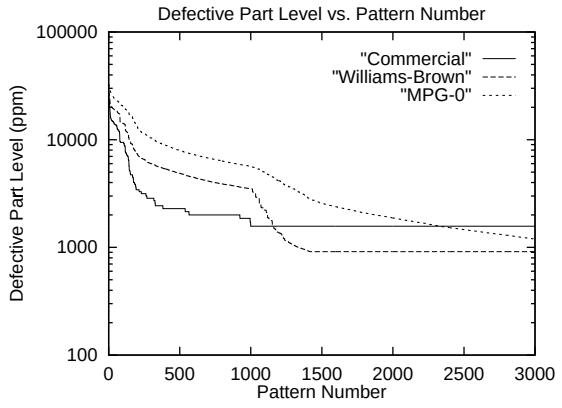

Figure 2: Comparison of defective part level using the empirical, the MPG-0, and the Williams-Brown model

4 A Refined Defective Part Level Model – MPG-1

In fact, the value of P_{EXCITE} should be a monotonically decreasing function of $\#OBS_i$ because each time site i is observed, some defects at site i are detected, and all chips with those defects will be eliminated from the remainder of the testing process. Thus, future observations at site i have a smaller probability of exciting a defect – since some of the defects have been removed. Figure 3 shows this idea in pictorial form.

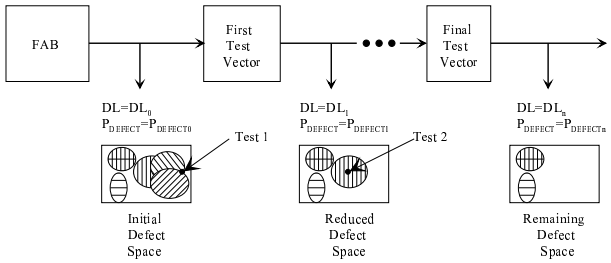

Figure 3: Probability of defect excitation decreases as more tests are applied

Before a chip enters the testing process, call the defective part level

$$DL_0 = 1 - factoryYield = 1 - Y_0 \tag{5}$$

At site i, none of the defects have been detected, and the probability of exciting at least one new defect, $P_{EXCITE_{i,0}}$, is shown on the Venn diagram (representing the set of all possible tests partitioned based upon whether a particular test excites at least one new defect).

After the first test pattern has been applied, the detected defective parts are removed from the chip set, and a new defective part level, DL_1, results. Since some defective chips have been removed, the set of all possible tests which will excite at least one new defect is reduced – as shown by the middle Venn diagram in Figure 3. This process continues as each test pattern is applied, and when the testing process is finished, the defective part level, DL_n, and the probability of exciting at least one new defect at site i, $P_{EXCITE_{i,n}}$, have been significantly reduced. Based upon these observations, we refine MPG-0 to form the new DL model MPG-1 where:

$$P_{EXCITE_{i,j}}(\#OBS_{i,1\rightarrow(j-1)}) = A{*}e^{(-\#OBS_{i,1\rightarrow(j-1)}/\tau)} \tag{6}$$

In this paper, A is fixed at one for simplicity. The time constant, τ, remains to be determined based

upon a best possible fit for empirical data. A higher value for τ indicates that the probability of excitation decreases less rapidly as the site is observed more times. Thus, the point of diminishing returns (when more site observations will not significantly reduce the defective part level) is reached less quickly. Note that τ is a function of the type of ATPG used, the type of defects present, the initial yield of the device, and circuit structure.

For the MPG-0 model, P_{EXCITE_i} is a constant and therefore independent of the number of observations, so a single calculation can be performed from DL_0 to DL_n. In contrast, for the MPG-1 model, P_{EXCITE_i} changes as the number of test patterns (and number of site observations) increases. Therefore, the defective part level calculation becomes iterative on a pattern by pattern basis. From DL_0 we find DL_1 , we continue to find DL_j from DL_{j-1} until, in the end, we have calculated DL_n using DL_{n-1}. More formally:

$$DL_j \cong 1 - \prod_{i=1}^{Sites} [1-(1-P_{EXCITE_{i,j}})^{\#OBS_{i,j}} \times P_{DEFECT_j}] \quad (7)$$

where: DL_j is the estimated defective part level after the application of pattern j, $P_{EXCITE_{i,j}} = e^{(-\#OBS_{i,1\rightarrow(j-1)}/\tau)}$, $\#OBS_{i,1\rightarrow(j-1)}$ is the number of times site i has been observed during the first $(j-1)$ patterns, $\#OBS_{i,j}$ is the number of times site i is observed during pattern j (either 0 or 1), and $P_{DEFECT_j} = 1-(1-DL_{j-1})^{(1/\#SITES)}$

This modification is a natural result of the changing of the probability of a defect. After each test vector is run, the process essentially begins again. Since the current defect level determines the probability of a defect, each test vector applied can be considered the first vector in a new test process. The effects of the application of previous test vectors are contained in the new P_{DEFECT}. The only way to possibly decrease the defective part level further is to observe sites with the application of the current vector.

It is also possible to see that this new model gives correct values for the endpoint condition. If the probability of excitation at every site is equal to zero for vector j, then the defect level is:

$$DL_j \cong 1 - \prod_{i=1}^{Sites} [1 - P_{DEFECT_j}] \quad (8)$$

Thus,

$$DL_j \cong 1 - (1 - P_{DEFECT_j})^{\#SITES} = DL_{j-1} \quad (9)$$

and the defective part level remains the same as would be expected.

5 Comparision of Actual DL with Predicitions by MPG-1

Figure 4 shows the actual and predicted DL for the experimental chip tested using the COMMERCIAL test pattern set. Here, the optimal value of τ is 2.59. Similarly, Figure 5 shows the same information for the RESEARCH test pattern set, and the value of τ is 4.42.

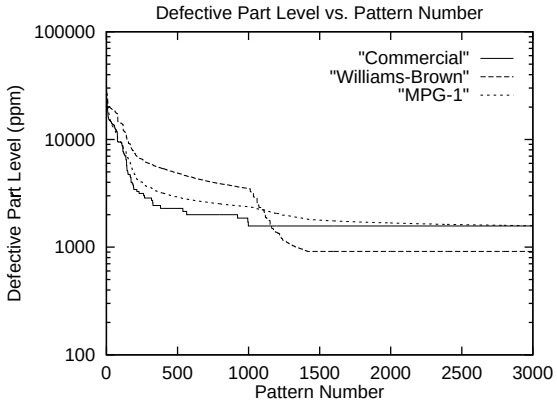

Figure 4: Predicted Defective Part Level when using τ=2.59 and COMMERCIAL vectors

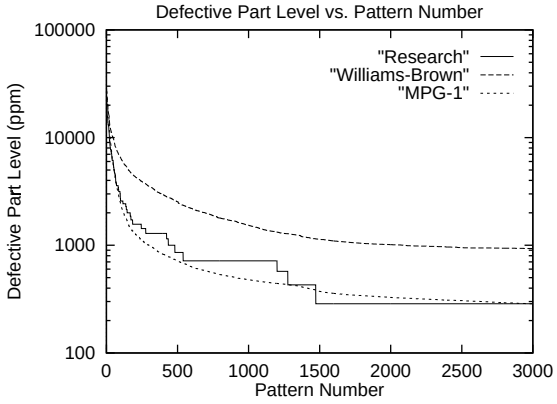

Figure 5: Predicted Defective Part Level when using τ=4.42 and RESEARCH vectors

In addition to this experimental data, we have also compared the DL predictions by MPG-1 to surrogate simulations on the C432 ISCAS benchmark circuit [BRGL85] All basic stuck-at one and stuck-at zero faults, after equivalent fault collapse, were used as the basic fault set, and statistics on the number of detections of each fault were monitored during the test

pattern generation and fault simulation process performed using ATALANTA [LEE93]. During the test pattern generation process, only stuck-at faults were monitored and no surrogate data affected test pattern selection. The yield was assumed to be 96.7%. A total of 45,000 AND and OR bridges between lines in the C432 circuit were modeled as surrogates, and the defective part level was calculated based upon the number of bridges which remained undetected.

Four different approaches were used to produce four different test pattern sets, and each of these is discussed below. The test pattern sets are ordered from the set which requires the least ATPG computational effort to the most computationally expensive approach.

Figure 6 compares the surrogate-based DL calculation with the MPG-1 predictor. In this case, one test pattern was targeted to detect each fault in the fault set (SAF). No fault simulation or fault dropping was done, so a pattern was produced to detect fault f_i even if an earlier pattern targeting another fault fortuitously had already detected fault f_i . Here, no effort at all was expended to maximize the number of detections of every fault (and therefore the number of observations at every site). Because the total number of observations was relatively small, the optimal value for τ was 5.65. This time constant for $P_{EXCITE_{i,j}}(\#OBS_{i,1\rightarrow(j-1)})$ is smallest of all the cases – indicating the least effective set of test patterns in terms of defective part level reduction per test pattern.

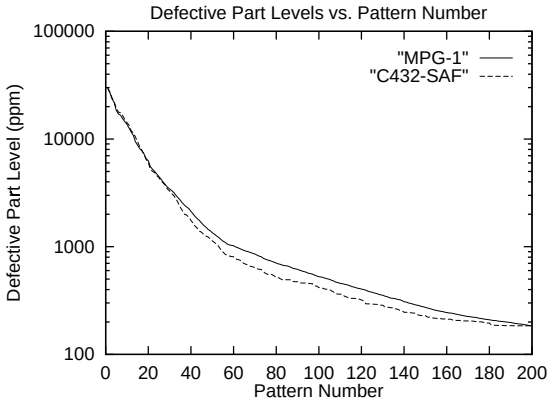

Figure 6: Predicted Defective Part Level for C432 when using τ=5.65 and SAF vectors

Figure 7 compares the surrogate-based DL calculation with the MPG-1 predictor for the second test pattern generation approach (DROP). τ's optimal value for this case is 6.14. This approach is essentially what is standard industry practice today and corresponds with the COMMERCIAL test pattern set for the experimental chip. Fault dropping was done, so no test pattern was produced to detect fault f_i if an earlier pattern targeting another fault fortuitously had already detected fault f_i. Since only about 70 test patterns were required to generate 100% stuck-at fault coverage, the process was repeated about three times to produce the a set of the same size – 200 test patterns – used in the other cases. Relative to the test pattern set of Figure 6, there were more fault detections (and therefore the number of observations at every site was larger). The resulting time constant for the $P_{EXCITE_{i,j}}(\#OBS_{i,1\rightarrow(j-1)})$ is larger – indicating that the test pattern set of Figure 7 is more effective than the set for Figure 6 in terms of defective part level reduction per test pattern.

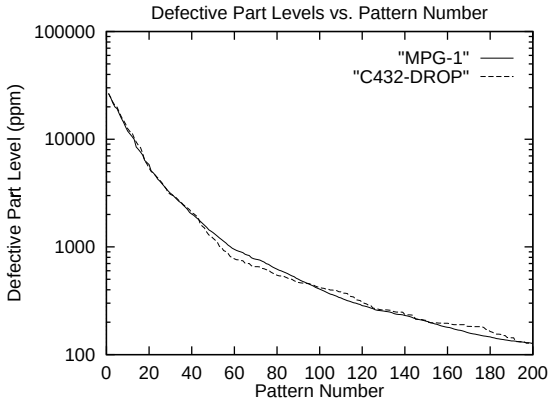

Figure 7: Predicted Defective Part Level for C432 when using τ=6.14 and DROP vectors

τ's optimal value for the set of Figure 8 is 6.26. This approach is essentially the strategy used to generate the RESEARCH test set (OPT). A super-set with about 1,000 test patterns was first produced, and then an optimization strategy was used to select the subset of 200 test patterns which produced the minimal defective part level (as estimated by MPG-0). The resulting time constant for the $P_{EXCITE_{i,j}}(\#OBS_{i,1\rightarrow(j-1)})$ is even larger than those for the previous three cases.

Finally, Figure 9 was produced using a test pattern set (XSIM) which required the largest amount of computational effort. The resulting optimal value for τ was 6.72. In this case, the test pattern for a fault could initially contain "don't cares" for some of the primary input values. All possible combinations of values were assigned to these "don't care" inputs, and the test pattern which maximized the number of site observations was selected from all the others. Note that

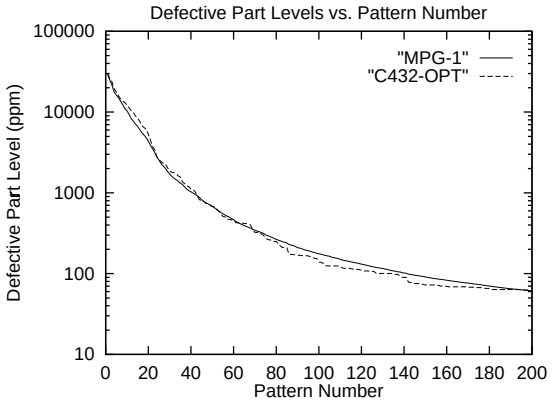

Figure 8: Predicted Defective Part Level for C432 when using τ=6.26 and OPT vectors

in the case of 10 inputs with don't care assignments, 1,024 different test patterns were simulated and only one was selected as part of the final set of 200. While this approach is not reasonable for commercial applications, it serves to indicate what might be achieved in the limit when the number of site observations is maximized. Of all cases, this one shows the smallest final defective part level, and the value for τ is also maximal.

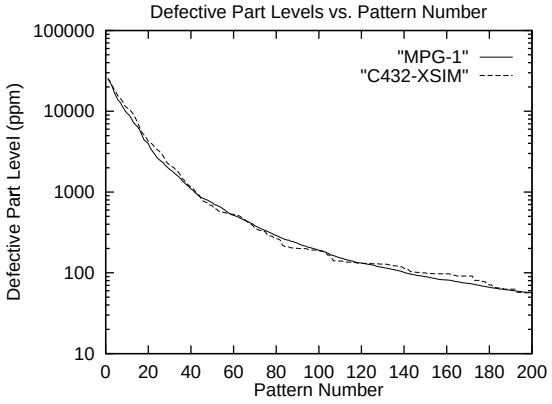

Figure 9: Predicted Defective Part Level for C432 when using τ=6.72 and XSIM vectors

6 Randomness of Excitation

Recall that the RESEARCH tests are selected to enhance observation counts at each individual site. Results in the previous section, however, indicate that by focusing ATPG efforts *only* on site observation, we also at the same time increase the chance of defect excitation. We can view τ as quantizing "the number of patterns for which defect excitation is *effective*" and hence, RESEARCH tests are more robust than COMMERCIAL tests in that sense.

For each test, individual defect excitation is a random variable with unknown probability distribution. Previous research [WMW96] suggested that given such unknown probabilities across the whole circuit, the best strategy to optimize *the least probability of defect excitation* is to maintain excitation uniformly random among all sites, i.e. maintain 0.5 probability for a site to be set to 0 or 1, for all sites. Evidences that more randomness during the test selection process help to detect more non-target faults can also be found in the early work of weighted random pattern generation, where experimental results clearly demonstrated the improvements of fortuitous detection [EIC91].

In today's commercial ATPG tools, sophisticated algorithms, such as smarter test selection criteria [MIR90], are commonly adopted to reduce the search effort for tests. This means that the test selection process tends to be biased in a certain way (in order to improve run-time efficiency) and hence, contradicts the "**uniform-randomness principle**" stated above.

On the contrary, during the RESEARCH test selection process, more randomness was allowed in two ways and hence, contributed to the enhancement of defect excitation.

- ATPG heuristics that directed the test search process for run-time efficiency were turned off. As a result, more randomness was involved and search space was less restricted.
- Although the order of faults selected attempted to maximize site observation count, the use of random ATPG helped maintain the unformity of defect excitation across all sites in the circuit. The method also attempted to minimize the difference of fault excitation counts among all sites. For these reasons, the uniform-randomness principle was more fulfilled when compared to conventional stuck-at fault-based ATPG methods.

7 Conclusions

During the testing process, defects are detected at various sites and removed from future consideration. Thus the probability of defect detection decreases monotonically with the number of times that a particular site is observed. Incorporating this observation into the MPG-0 model leads to an iterative form with an exponentially decreasing probability of

excitation, MPG-1. The optimal value for τ depends on the application for which the DL model is being employed.

Figure 10 shows a comparison of the actual and MPG-1 predicted DL plots for the RESEARCH test pattern set when using different τ values. An upper bound is shown using $\tau = 2.59$ (the smallest value for τ used in this paper and the optimal value for the COMMERCIAL patterns). Similarly, a lower bound is shown using $\tau = 6.72$ (the largest value for τ used in this paper and the optimal value for the best test pattern set, XSIM, for the C432 circuit).

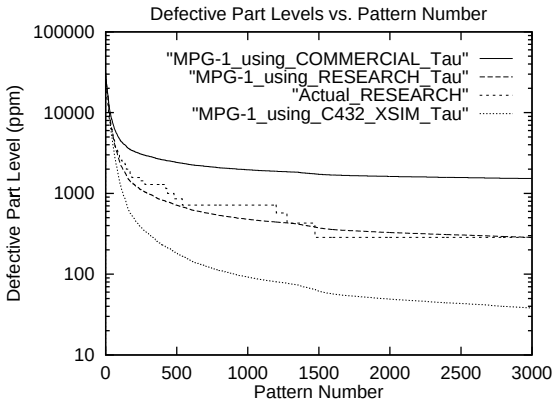

Figure 10: Comparing defective part level predictions for different τ values

The accuracy of defective part level prediction by MPG-1 represents a significant enhancement over previous approaches. Not only does the MPG-1 model allow relatively more accurate DL predictions, it also is the first tool capable of quantitatively comparing the effectiveness of an ATPG strategy in terms of defective part level reduction effectiveness.

References

[BRGL85] Brglez, F. and Fujiwara, "A Neutral Netlist of 10 Combination Benchmark Circuits and a Target Translator in FORTRAN," *Proc. Int. Symp. Circuits Syst.*, 1985.

[EIC91] E. B. Eichelberger et. al., *Structured Logic Testing,* Prentice Hall, 1991, Sec. 15.

[GRIM99] Grimaila, M.R., Lee, S., Dworak, J.,Butler, K.M., Stewart, B., Balachandran, H., Houchins, B., Mathur, V., Park, J. , Wang, L.C., Mercer, M.R., "REDO — Random Excitation and Deterministic Observation — First Commercial Experiment," *Proc. VLSI Test Symposium 1999*, pp. pending.

[LEE93] Lee, H.K. and Ha, D.S., "On the generation of test patterns for combinational circuits," *Technical Report No. 12, Department of Electrical Engineering,* Virginia Polytechnic Institute and State University, 1993.

[MIR90] M. Abramovici, M. A. Breuer, and A. D. Friedman, *Digital Systems Testing and Testable Design*, Computer Science Press, 1990, Section 6.2.1.3.

[WILL81] Williams, T.W. and Brown, N.C. "Defect level as a function of fault coverage," *IEEE Trans. on Computers*, Vol. C-30, no. 12, 1981, pp. 987-988.

[WMW96] Wang, L-C., Mercer, M.R., and Williams, T.W., "Using Target Faults to Detect Non-Target Defects," *Proc. International Test Conference 1996*, pp. 629-638.