1-1-1990

An Experimental Study on Reject Ratio Prediction for VLSl Circuits: Kokomo Revisited

Dharam Vir Das
University of Nebraska - Lincoln

Sharad Seth
University of Nebraska - Lincoln, seth@cse.unl.edu

Paul T. Wagner
Delco Electronics, Kokomo, Indiana

John Anderson
Delco Electronics, Kokomo, Indiana

Vishwani Agrawal
ATBT Bell Laboratories, Murray Hill, New Jersey

Proceedings, International Test Conference, 1990. doi: 10.1109/TEST.1990.114087

An Experimental Study on Reject Ratio Prediction for VLSI Circuits: Kokomo Revisited

Dharam Vir Das and Sharad C. Seth

University of Nebraska, Lincoln, Nebraska 68588

Paul T. Wagner and John C. Anderson

Delco Electronics, Kokomo, Indiana 46904

Vishwani D. Agrawal

AT&T Bell Laboratories, Murray Hill, New Jersey 07974

ABSTRACT

Assuring product quality is becoming increasingly more important for the semiconductor chip manufacturers. The reject ratio (defect level) provides a simple and accurate measure of a product's quality. However, measuring the reject ratio of tested chips is often not feasible or accurate. Statistical techniques for reject ratio prediction provide a possible way out of this dilemma. In this paper, we report on an experiment to verify the accuracy of reject ratio predictions by the available approaches. The data collection effort includes instrumenting the wafer probe test to obtain chip failures as a function of applied vectors and running a fault simulator to obtain the cumulative fault coverage of these vectors. The accuracy of reject ratio predictions is judged by assuming earlier stopping points for the wafer probe thereby gaining a measure of confidence in the final predicted value. The results of five different analysis are reported for over 70,000 tested die of a CMOS VLSI device manufactured at Delco.

Introduction

In a world of increasing industrial competition, manufacturers are becoming ever more conscious of product quality. Perhaps, no where is this quality consciousness more evident than in the semiconductor industry where the quality levels projected for the coming decade were unthinkable just a few years ago. An accurate measure of product quality is rather easily defined and is variously denoted by the terms *reject ratio*, *PQL* (product quality level), and *DPM* (defects per million). The basic notion is captured by the following definition of chip quality:

$$\textit{Reject ratio} = \frac{\textit{number of bad units tested as good}}{\textit{number of all units tested as good}}$$

This simple definition belies the great difficulty of its use for semiconductor devices. Estimating the reject ratio requires an extensive and sophisticated monitoring system to collect and analyze the devices that fail past the wafer-probe stage. This presupposes a degree of cooperation amongst groups of people with scarce or nonexistent communication links. For this reason, the direct approach is rarely attempted or attempted only in limited ways. The results from even the limited experiments are generally not available because of their sensitive nature.

The reject ratio of VLSI devices can be improved by increasing the fault coverage of the functional test set. However, the dependence of the reject ratio on fault coverage is complex and involves process-dependent parameters in the equation. Several statistical models (see Reject Ratio Computation) have been proposed to answer the question, "How much fault coverage is enough for a desired (predicted) reject ratio?" With so many competing models for reject ratio computation, two questions naturally arise: (a) How good are they? (b) Are they significantly different in their predictive value?

In 1981, a study of actual VLSI chip production data obtained at AT&T was published [1]. In this study, a model for reject ratio and fault coverage was proposed. Another significant study of chip data was conducted at Kokomo, Indiana, in 1980 jointly by Delco Electronics and Motorola [2]. Since then, several other models

have been proposed. A recent paper [3] concludes that extremely high fault coverage (in excess of 99%) is required by all models for a high quality level (less than 0.1%). We must, however, point out that such conclusion should not be taken in the absolute sense. According to our previous work, the reject ratio and fault coverage relation is a function of the technology, process maturity, etc. Actual assessment of quality should therefore be based upon experiments. In this paper, we demonstrate an experimental procedure.

Our results are based on experimental data collected on a large volume VLSI device at Delco Electronics. This is the second such data collection effort carried out by Delco at Kokomo for determining fault coverage requirements. Although the data collection aspects of the two experiments are similar, the objective of analysis reported here is quite different. Our goals are to obtain the final predicted values of the reject ratio by multiple approaches and to verify the accuracy of predictions over a range of fault coverages.

Data Collection

The process of data collection began with selecting the most appropriate device for meeting our objective of analyzing the relationship between the reject ratio and stuck-at fault coverage. The device needed to be purely digital, have a stuck-at fault coverage well in excess of 99%, and be in high-volume production. Very high fault coverage was required to minimize the error in the reject ratio prediction at fault coverage levels lower than the final fault coverage. The device needed to be in high volume production to allow the data to be collected in a reasonable amount of time. Based on these criteria, a 3 micron digital CMOS IC with 99.7% fault coverage was selected. This device had a size of 136 x 139 sq-mils, 7750 device count, and was manufactured in a Class 100 clean room (i.e., no more than 100 particles larger than 1/2 micron in size per cubic foot).

Wafer Test Results. A wafer test flow diagram for the selected IC is shown in Figure 1 where functional (stuck-at) testing was performed first. If the device passed functional test, it was recorded as a good die or a parametric failure depending on the outcome of the parametric test. If the device failed functional test, the device was recorded as a functional or a continuity failure depending of the outcome of the continuity test. If the device was a functional failure, the clock step where the failure occurred was recorded. The results of data collection for 72912 devices at wafer test are shown in Figure 2.

	Total	Percent
Die Tested:	72912	100.00
Failed Parametric:	847	1.16
Failed Continuity:	7669	10.56
Failed Functional:	18476	25.34
Passed Wafer Test:	45890	62.94

Figure 2: The Wafer Test Data Collection Results

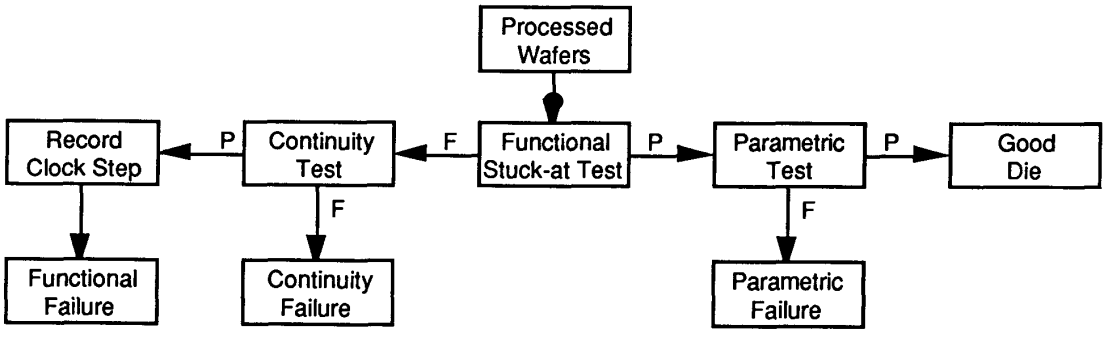

Figure 1: A Wafer Test Flow Diagram for the Selected IC

Data Analysis. Functional test for the selected IC consisted of 12188 clock steps where the cumulative fault coverage at each clock step had been precomputed by a fault simulator. Therefore, the capability existed of knowing exactly how many chip failures occurred up to any given level of fault coverage. For example, if testing was only performed until a 99% fault coverage was achieved, then the failures that occurred between 99% and 99.7% fault coverage can be used to predict the reject ratio. The error in this methodology exists in the fact that the IC did not have 100% fault coverage. Therefore, the additional failures that may have occurred after 99.7% fault coverage could not be determined.[1]

Reject Ratio Computation

Five statistical models for reject ratio estimation have appeared in the recent literature. Each model incorporates the chip yield as a parameter to characterize the fabrication process. Beyond this, the models differ substantially in their assumptions regarding the relationship between physical defects and logical faults, distribution of defects/faults, detection and occurrence probabilities of faults, etc. For a detailed understanding of each model, the reader is referred to the original sources (see References). We shall provide a brief review of the five models and state the essential equations used in reject ratio computation.

The Springfield Model (SPR). This model was first described at a workshop in Springfield, Massachusetts [4]. It assumes that a chip failure due to an applied test vector is a random event caused by the occurrence of a fault detectable by the test vector. A distinguishing feature of this approach is that the required data can be readily derived from wafer probe. This data is the number c of chips tested, the number N of applied test vectors, and the number c_i of chips failing at vector i. From this data, the yield y_n after n vectors and the true yield y are estimated as shown in Equations (1) and (2) and the reject ratio is computed as in Equation (3).

$$y_n = y + (1 - y - \frac{1}{c}\sum_{i=1}^{N} c_i)\,\frac{N+1}{N+n+1} + \frac{1}{c}\sum_{i=1}^{N} c_i\,\frac{i(i+1)}{n+i}\,(n+i+1) \qquad (1)$$

$$y = 1 - \frac{1}{c}\sum_{i=1}^{N} c_i - \frac{2N+1}{N}\,\frac{1}{c}\sum_{i=1}^{N} c_i\,\frac{i(i+1)}{(N+i)(N+i+1)} \qquad (2)$$

$$r_n = (y_n - y) / y_n \qquad (3)$$

The JSSC Model. This model derives its name from an article that appeared in the *IEEE Journal of Solid State Circuits* (JSCC) in 1982 [5]. It is based on the assumption that the number of stuck-type faults on a faulty chip is a random variable having a Poisson distribution. In this model, f is the fault coverage, y is the yield, and n_0 is the average number of faults on a faulty chip. The first and second parameters are assumed to be known and the third must be estimated from experimental data. For this purpose, the chip-failure data (fraction of chips failed vs. test vector number) is combined with the fault simulation data (fault coverage vs. test vector number) to eliminate the vector number and obtain the fraction of chips failed vs. fault coverage. Equation (4) is made to fit the actual data by choosing a suitable value of n_0. Equation (5) is then used to estimate the reject ratio. P(f) is fraction of chips rejected by test patterns with cumulative fault coverage f.

$$P(f) = (1-y)\,[1 - (1-f)\,e^{-(n_0-1)f}] \qquad (4)$$

$$r(f) = \frac{(1-f)(1-y)\,e^{-(n_0-1)f}}{y + (1-f)(1-y)\,e^{-(n_0-1)f}} \qquad (5)$$

[1] We must caution, however, that a perfect fault coverage is an illusive goal in practice because of the ideosyncratic treatment of certain fault classes (e.g., faults causing races and oscillations, potential detection, hyperactivity, etc.) by a fault simulator, the limitations of the fault model itself, and the very high incremental cost of either determining a remaining fault to be redundant or finding a test for it.

The CAD Model. This model is a refinement of the JSCC model and requires the same experimental data. It was described in an article in the *IEEE Transactions on Computer Aided Design* [6]. The number of logical faults, caused by physical defects, is assumed to have a Poisson distribution. Further, the number of physical defects is assumed to have a clustered negative binomial distribution. Thus, three parameters characterize the model: two describe the clustering of defects and the third relates these defects to the logical faults. The relationship between the yield and the fault coverage is the key to computing the reject ratio. This is described by Equation (6) where A is the chip area, f is the fault coverage, and a, b, and c are the three model parameters which are estimated by fitting the y(f) vs. f to the experimental data.

$$y(f) = [1 + Ab(1 - e^{-cf})]^{-a} \qquad (6)$$

Then the reject ratio is given by r(f) = [y(f) - y]/y.

The Wadsack Model. Wadsack [7] derives the following simple relation for the reject ratio as a function of the chip yield and fault coverage:

$$r(f) = (1-f)(1-y) \qquad (7)$$

This result is based on the assumption that multiple faults on a chip are independent of each other. Later work by Wadsack [8] modified this model to produce results equivalent to the Williams model described next

The Williams Model. Williams and Brown [9] make an assumption that faults occur independently on a chip. This model, described by Equation (8), provides a rather simple method for estimating the reject ratio where **y** is the chip yield and **f** is the fault coverage.

$$r(f) = 1 - y^{(1-f)} \qquad (8)$$

Experimental Results and Comparison

Our results are based on data collected from 72912 die with the yield of 62.94%. Since each of the five analysis models are based on just functional test, we removed the 847 chips that failed the parametric test and the 7699 chips that failed the continuity test. Thus, the total number of die used was 72912 - (847+7699) or 64366. Of these, 18476 failed

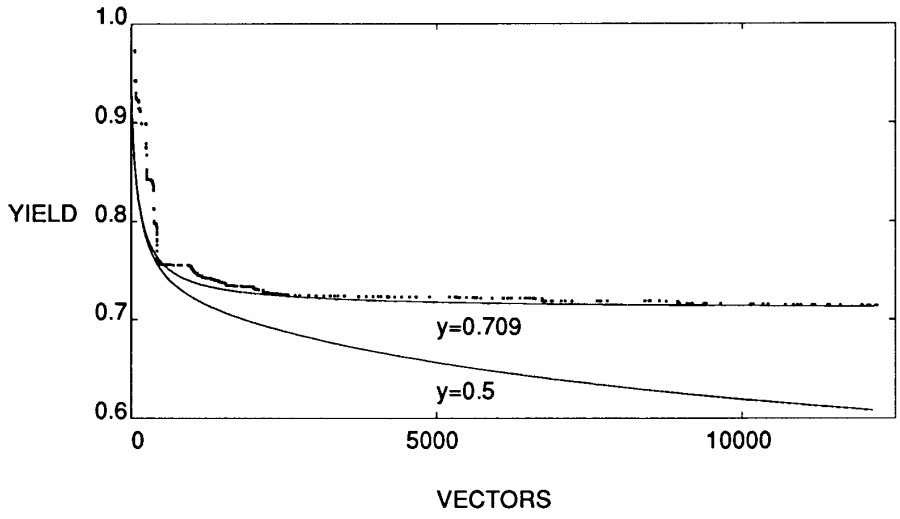

Figure 3: Yield Determination From The Measured Data

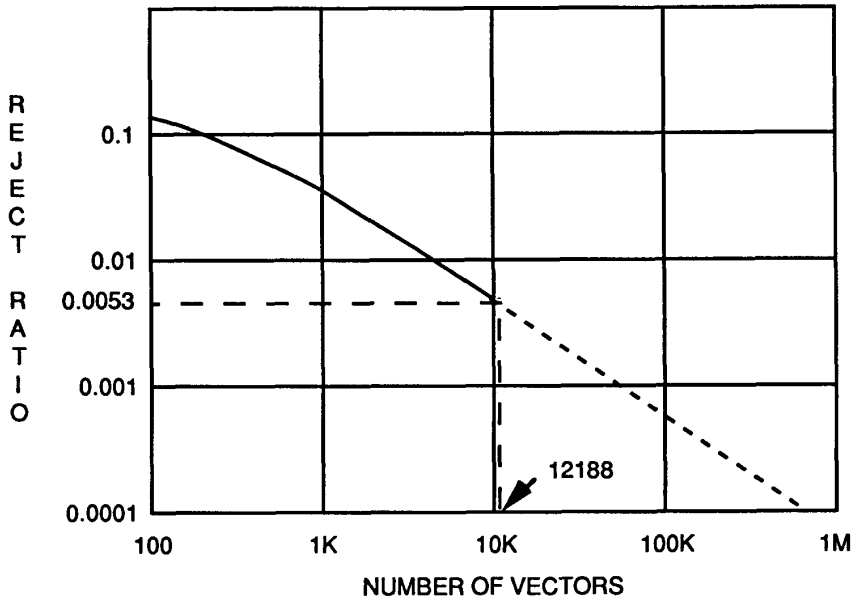

Figure 4: Computed Reject Ratio

functional test and the *measured* functional yield was 71.30%. Our determination of *true* yield is based on the SPR model because it allows yield estimation without fault-coverage data.

The true yield was estimated as 70.92% by fitting Equation (2) to the measured data. This is illustrated in Figure 3 showing the measured yield by as a function of vector number (dots) and the predicted true yield that best fits the experimental data. In order to demonstrate the sensitivity of the model to the yield parameter, another curve for a true yield value of 0.5 is also shown. Equation (3) allows the reject ratio to be estimated as a function of the vector number. The result is shown in Figure 4 as well as in the row marked SPR of Table 1 (where the vector number has been replaced by the corresponding cumulative fault coverage.) The estimated reject ratio after all test vectors (12188) have been applied is 0.00532.

The model [4] allows extrapolation of reject ratio values to larger number of vectors. For example, if the testing were to continue to 100,000 vectors, the reject ratio would drop by one order of magnitude (Figure 4).

Since each vector corresponds to a certain cumulative fault coverage as determined by fault simulation, the reject ratio can also be represented as a function of fault coverage (see SPR result in Table 1). The SPR technique also allows estimation of the *failure profile* of the circuit, that is the distribution of failure probabilities of all the faults in the circuit. The Failure probability (x) refers to the combined probability of fault occurrence and detection The density function $f(x)$ of this distribution is shown in Figure 5. It has two components: a delta function at the origin of height equal to the chip yield and a failure probability function $p(x)$ with the area under the curve equal to $(1-y)$. It is the shape of $p(x)$ near the origin that determines the reject ratio for high fault-coverage values.

Next, we show the reject ratio computation according to the JSCC model [5]. Here, we need the fault simulator data (fault coverage vs. vector number) in addition to the chip failure data. Before the reject ratio can be computed, we must estimate an additional parameter from the experimental data. This parameter represents the average number of logical (stuck type) faults produced by a physical defect on a chip and is denoted by n_0. Figure 6 shows that n_0 is approximately 2 for our

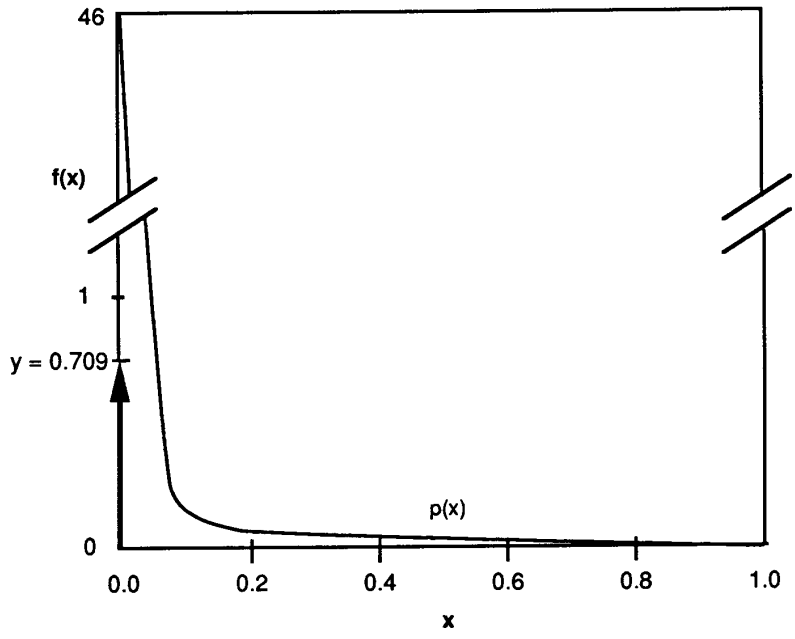

Figure 5: Failure Profile of the Circuit

data. The resulting reject ratio is shown in the row marked JSCC in Table 1.

The third row in the Table 1 presents results from the CAD model due to Seth and Agrawal [6]. This is a refinement of the n_o approach. The number of logical faults per defect is assumed to be a random variable with a Poisson distribution. The three parameters used

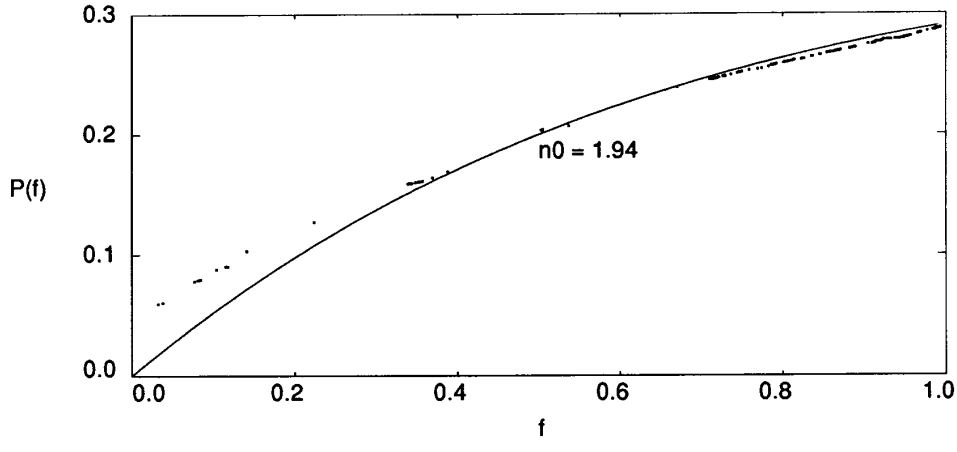

Figure 6: Determination of n_0 from experimental data

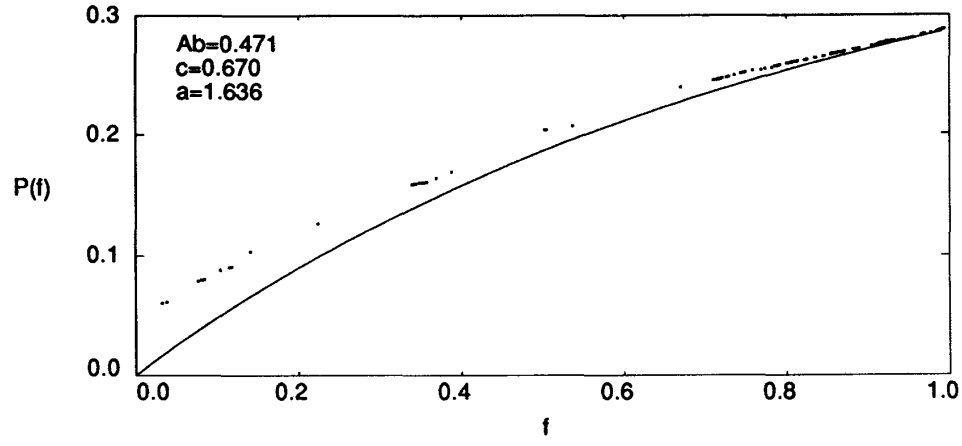

Figure 7: Determination of a, Ab, and c from experimental data

Table 1: Predicted Reject Ratio

Model	Faullt Coverage						
	20%	50%	80%	91%	95%	98%	99.70%
SPR	0.11291	0.08005	0.03531	0.02160	0.00927	0.00702	0.00532
JSSC	0.21383	0.11373	0.03730	0.01548	0.00834	0.00362	0.00048
CAD	0.21714	0.12439	0.04556	0.01985	0.01090	0.00432	0.00064
Wadsack	0.23267	0.14542	0.05817	0.02617	0.01454	0.00582	0.00087
Williams	0.24038	0.15788	0.06642	0.03046	0.01704	0.00685	0.00103

in this model were estimated to be as follows:

$$a = 1.636,\ Ab = 0.471,\ \text{and}\ c = 0.670.$$

These were obtained by weighted fitting of Equation (6) to the experimental data as shown in Figure 7 [10]. Finally, we also show the results according to the analysis presented by Wadsack [7] and by Williams & Brown [9]. These are shown in the rows labeled *Wadsack* and *Williams* in Table 1.

Summary of Results

The predicted reject ratios in Table 1 differ by as much as an order of magnitude. Unfortunately, there is no way of determining the accuracy of these predictions in an absolute way except by obtaining substantially large amount of accurate data on *field rejects,* a formidable procedure fraught with its own inaccuracies. Each of the five methods can predict the reject ratio at the end of testing. In order to determine the effectiveness of the predictions we suggest a normalization of the results given in Table 1 that allows simple comparison of predicted reject ratios against the actual values *for all but the last column of the table.* This is performed by reducing each value by the amount shown in the last column in the same row. The last row marked "actual" assumes that the yield at 99.7% coverage is the true yield. The resulting values in a particular column can now be compared with the actual chip rejects between the fault coverage of the column under consideration and the final fault coverage. We know, for example, that

Table 2: Normalized Predicted Reject Ratio

Model	Faullt Coverage						
	20%	50%	80%	91%	95%	98%	99.70%
SPR	0.10760	0.07474	0.03000	0.01629	0.00396	0.00171	0
JSSC	0.21335	0.11335	0.03682	0.01500	0.00786	0.00278	0
CAD	0.21650	0.12375	0.04492	0.01921	0.01026	0.00368	0
Wadsack	0.23180	0.14455	0.05730	0.02530	0.01367	0.00495	0
Williams	0.23935	0.15685	0.06539	0.02943	0.01601	0.00582	0
Actual	0.18440	0.08340	0.02830	0.01330	0.00740	0.00210	0

if testing had stopped at 98% fault coverage, an additional 0.0021 fraction of the chips rejected beyond this point would have been included in the reject ratio. For a 95% fault coverage, this fraction would have been 0.0074 and so on. Thus, one way to compare the predicted values at less than the maximum fault coverage against the actual values is to normalize the reject ratio at the maximum coverage in all cases to zero. The resulting values are shown in Table 2 where the high end coverage of all methods provide comparable accuracy. However, the methods described by the first three rows are uniformly close to the actual values unlike the remaining values in the table.

It is not uncommon in wafer probe to find that no chips are rejected by the last few vectors even though the fault coverage continues to rise. For example, for the sample of chips considered in our experiment, no chips were rejected beyond the clock step number 12140 (fault coverage = 98.89%)even though, as stated earlier, the testing continued till clock step number 12188 (fault coverage = 99.7%). What value of fault coverage should one use in reject ratio calculations under such circumstances? It might be argued, for example, that the last 48 vectors are redundant, appearing to cover faults that never occur. However, we believe, there is another explanation for this commonly observed phenomenon, namely, that the remaining faults do occur, possibly clustered with faults that are detected by earlier test vectors. Such fault clustering is explicitly considered in the JSSC and CAD models.

Conclusion

Not surprisingly, the five methods discussed in this paper predicted values for the reject ratio that vary by an order of magnitude at high values of fault coverage. We have shown that with only an incremental effort during wafer probe, data collection is possible that can be used to compare the relative accuracy of different models over a range of fault coverage. We believe that more studies of this nature are necessary for further refinement and calibration of the existing approaches for reject ratio prediction. Of all the methods discussed in this paper, only the one marked SPR [4] can provide results without the fault simulation data. Since this model predicted results similar to those given by the other methods, this would be the method of choice in situations where the fault coverage data is either not available or too expensive to obtain.

References

[1] V. D. Agrawal, S. C. Seth, & P. Agrawal, "LSI Product Quality and Fault Coverage," *Proc. Design Automation Conf.*, pp. 196-203, June 1981.

[2] R. A. Harrison, R. W. Holzworth, P. R. Motz, R. G. Daniels, J. S. Thomas, and W. H. Wieman, "Logic Fault Verification of LSI: How it Benefits the User," *Proc. WESCON*, paper 34/1, September 1980.

[3] E. J. McCluskey, "IC Quality and Test Transparency," *Proc. Int. Test Conf.*, pp. 295-301, September 1988.

[4] S. C. Seth and V. D. Agrawal, "On the Probability of Fault Occurrence," in *Defect and Fault Tolerance in VLSI Systems*, ed. I. Koren, pp. 47-52, Plenum, New York, 1989.

[5] V. D. Agrawal, S. C. Seth, and P. Agrawal, "Fault Coverage Requirements in Production Testing of LSI Circuits," *IEEE Journal of Solid State Circuits*, Vol. SC-17, pp. 57-61, February 1982.

[6] S. C. Seth and V. D. Agrawal, "Characterizing the LSI Yield Equation From Wafer Test Data," *IEEE Trans. Computer Aided Design of Int. Circuits and Systems*, Vol. CAD-3, pp. 123-126, April 1984.

[7] R. L. Wadsack, "Fault Coverage in Digital Integrated Circuits," *Bell Syst. Tech. Jour.*, Vol. 57, pp. 1475-1488, May-June 1978.

[8] R. L. Wadsack, "VLSI: How Much Fault Coverage is Enough?," *Proc. IEEE Int. Test Conf.*, pp. 547-554, Philadelphia, October 1981.

[9] T. W. Williams and N. C. Brown, "Defect Level as a Function of Fault Coverage," *IEEE Trans. Computers*, Vol. C-030, pp. 987-988, December 1981.

[10] S. C. Seth and V. D. Agrawal, "Forecasting Reject Rate of Tested LSI Chips," *IEEE Elec. Dev. Letters*, Vol. EDL-2, No. 11, pp. 286-287, November 1981.