Evaluation of Statistical Outlier Rejection Methods for I_{DDQ} Testing

Sagar Sabade Hank Walker
Department of Computer Science
Texas A&M University
College Station, TX-77843-3112 USA
E-mail: {sagars, walker}@cs.tamu.edu

Abstract

The quiescent current testing (I_{DDQ} testing) for CMOS ICs provides several advantages over other testing methods. However, the future of I_{DDQ} testing is threatened by increased sub-threshold leakage current for new technologies. The conventional pass/fail limit setting methodology cannot survive in its present form. In this paper we evaluate two statistical outlier rejection methods – the Chauvenet's criterion and the Tukey test – for their applicability to I_{DDQ} testing. They are compared with the static-threshold method. The results of the analysis of application of these methods to the SEMATECH data[1] are presented.

1. Introduction

I_{DDQ} testing has been recognized to be an important testing method for IC industry [1,2]. The continuous advancement in IC fabrication technology and the requirements for high reliability need better testing methods. I_{DDQ} testing has been shown to provide several benefits and ensure high reliability of components [3,4,5]. Since I_{DDQ} testing is based on the root cause of problem to identify a defective part [6] it can detect several defect types [7]. Traditionally, higher leakage current in CMOS was believed to indicate the presence of a defect. The test engineers used a current threshold above which all chips were considered faulty (see Figure 1). This threshold could be estimated using elaborate circuit models or empirically generated [8].

However, due to shrinking geometry of transistors, leakage current is increased [9,10]. Therefore, it has become difficult to distinguish between the good and the faulty chips based on a static threshold [11]. Rejecting good chips having higher I_{DDQ} results in *yield loss* (region B in Figure 1) and accepting bad chips with low I_{DDQ} (region A in Figure 1) can result in a *customer return*. Therefore, the approach based on pass/fail limit setting cannot survive in its present form.

[1] This data comes from the work of the Test thrust at SEMATECH, Project S121. The analysis here is the work of this university, the conclusions are our own and do not necessarily represent the views of SEMATECH or its member companies.

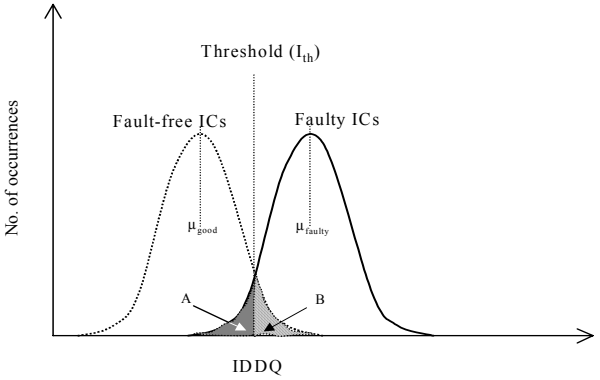

Figure 1: Typical IDDQ distributions

Several techniques have been proposed in the literature to extend I_{DDQ} testing for deep sub-micron (DSM) technologies. Each technique has its own merits and limitations. In this work, we evaluate the Chauvenet's criterion and the Tukey test for I_{DDQ} pass/fail limit setting. The remainder of this paper is organized as follows. In the next section, we describe the motivation to solve this problem. Section 3 reviews different approaches reported in the literature to solve this problem. Section 4 describes the statistical methods for data rejection mentioned above. Section 5 discusses methodology for the application of these methods to I_{DDQ} testing. The results based on the analysis of the SEMATECH data are presented in section 6. Finally, section 7 concludes the paper.

2. Motivation

I_{DDQ} testing offers several advantages [3]. Since the supply current can be monitored easily, it provides excellent observability [12]. High I_{DDQ} parts are also shown to have reliability concerns [13]. Moreover, I_{DDQ} testing is capable of detecting many defects that are not testable by other test methods [15]. As SIA roadmap indicates testing VLSI chips is one of the most important challenges for IC industry [14]. In order to achieve high reliability, higher yield or low DPM (defective parts per million) target, it is necessary that I_{DDQ} test be a part of testing strategy [15]. Therefore, the IC industry cannot afford to lose this testing method [16]. However, higher chip complexity and smaller

transistor widths have forced reduced threshold voltage (V_t). The reduced V_t causes corresponding increase in the leakage current. SIA roadmap indicates that I_{DDQ} testing is a crucial challenge in the coming years. Not only the leakage current is increasing, but the variation in fault-free and faulty I_{DDQ} is increasing as well. The process variations cause drift in transistor parameters (viz. L_{eff} and V_t) and cause a drift in I_{DDQ}. While some process variations are deterministic and can be modeled, some are random and cannot be predicted [17]. These random fluctuations are one of the biggest hurdles for I_{DDQ} testing. Figure 2 shows the plot of number of rejects for the SEMATECH data for various threshold values. The SEMATECH experiment used 5 μA threshold resulting in 1689 dice I_{DDQ}-only failures. Many of these dice passed all tests or failed only I_{DDQ} test after burn-in (BI). This demonstrates that any particular threshold value is not justifiable.

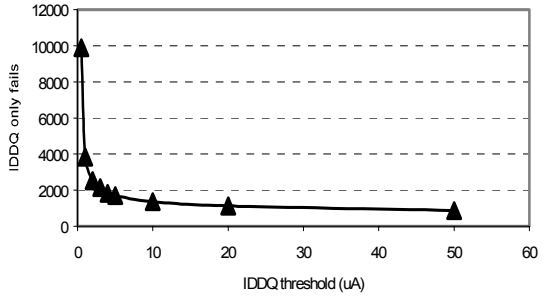

Figure 2: Rejected chips for static thresholds

The majority of techniques that propose to extend I_{DDQ} for deep sub-micron (DSM) chips rely on reducing the variance of I_{DDQ} of fault-free chips. The key idea is to increase the S/N ratio so that elevated I_{DDQ} due to a defect stands out and can be detected. In the current signature approach [18] the I_{DDQ} readings for a number of vectors are arranged in the ascending order. An I_{DDQ} signature for a defective chip shows a noticeable jump. This technique necessarily involves a series of measurements. The production-worthiness of this method is discussed in the literature [19].

3. Previous Work

In this section, we review some methods reported in the literature. This includes

- Delta I_{DDQ}
- Current Ratios
- Clustering Techniques
- Nearest Neighbor Residual

Delta I_{DDQ} [20,21]

The delta I_{DDQ} technique relies on the fact that the variance of the difference between adjacent vectors is much less than the absolute values. In this method, the difference between two consecutive readings is calculated. If a chip is defective, a vector that sensitizes the defect has I_{DDQ} an order of magnitude higher than the one that does not. The histogram of difference between readings is used to decide whether a die is defective.

Current Ratios [22]

Even if the leakage current is increasing, the ratio of maximum to minimum I_{DDQ} reading on a chip remains relatively constant. This is the basic idea exploited by current ratios method. It involves taking several I_{DDQ} measurements on various functionally fault-free chips and finding ratio of maximum I_{DDQ} to minimum I_{DDQ}. Then linear regression is performed on the minimum and maximum values to find the best fitting line for ratio. The slope of this line represents the current ratio of maximum to minimum I_{DDQ}. To account for non-deterministic process variations a guard band is added. If any I_{DDQ} ratio is more than that allowed by the guard band, the chip is considered defective and is discarded. This method necessarily needs characterization of the data and "tuning up" the current ratio to account for lot-to-lot variations.

Clustering Techniques [23,24]

Clustering is a statistical technique that groups data in a multi-dimensional space so that member in a group has more natural relation to other members in the group. This helps to understand the hidden patterns in the data. Clustering for I_{DDQ} testing examines all the I_{DDQ} measurements of a device instead of observing them in isolation. The inherent key benefit of clustering is that it accounts for process variations because it looks at all the measurements for forming clusters. This approach does eliminate the need for a static threshold to certain extent. However, a significant number of measurements are necessary for better clustering resolution.

Nearest Neighbor Residual [25]

It is observed that the defects tend to appear in clusters. Therefore on a wafer, a die is surrounded by defective chips is more likely to be faulty. This hypothesis was validated using SEMATECH data [26,27]. In [25], the authors showed that the I_{DDQ} readings of the neighboring dice on a wafer could be used for variance reduction. It involves estimation of fault-free I_{DDQ} for a die based on the I_{DDQ} readings of its neighboring dice. When neighboring dice are not available (e.g. on edge of a wafer) dice at a greater distance in all directions are used for estimation.

4. Statistical Data Rejection Methods

Sometimes in an experiment, a reading that seems to deviate from the normal trend of measurements is not

legitimate. Often the root cause of error cannot be determined or is beyond control of the researcher. Whether to reject such a reading or not is a debatable issue. The rejection of an "apparently" abnormal reading is an important decision in statistical analysis. A group of statisticians advocates the risk of rejecting "seemingly" outliers and questions the validity of results based on the remaining data. Yet another group supports such rejection, provided the trend of readings is expected or known [28]. The outlier rejection is highly subjective and must be performed with proper discretion. For some great discoveries have been made by searching for the root cause of the outliers. However, there are cases where such rejection of data is unquestionably needed. Chauvenet's criterion and Tukey test are two such methods for data rejection.

A. Chauvenet's Criterion [29]

Chauvenet's criterion assumes the Normal distribution of the data. It determines the probability of occurrence of a measurement. A probability threshold (P_{th}) is set such that all the readings having probability of occurrence less than P_{th} are considered unlikely and rejected. If there are N measurements $x_1, x_2, ...x_N$ in the data set, we compute the mean (μ_x) and the standard deviation (σ_x) for this data. If a reading (x_{sus}) is suspected to be an outlier, we calculate the number of standard deviations (t_{sus}) it is away from the mean.

$$t_{sus} = \frac{x_{sus} - \mu_x}{\sigma_x}$$

From the normal error integral table[2] we find the probability $P(outside\ t_{sus}.\sigma_x)$ that a reading will be t_{sus} or more standard deviations away from the mean. Finally, we multiply this probability by total number of measurements N, to obtain

$$n(\text{worse than } x_{sus}) = N.\ P(\text{outside } t_{sus}.\ \sigma_x)$$

The LHS of the above equation is the number of measurements expected to be at least as bad as x_{sus}. If n is less than a predetermined threshold, x_{sus} fails Chauvenet's criterion and is rejected.[3]

The conventional Chauvenet's criterion uses threshold of 0.5. It is possible to control the rigidity of control over the distribution by changing the threshold. It may appear that the pass/fail limit setting problem is simply translated into another domain. However, note that conventional pass/fail limit setting does *not* consider the probability of occurrence of elevated I_{DDQ}. A conventional method, for example, would take a measurement and consider a guard band around this measurement. Any chip that has I_{DDQ} outside this band is considered defective and is rejected. The application of Chauvenet's criterion, however, allows us to look at the entire data set as a whole. Thus a die that appears to have "high" I_{DDQ} could be a local maximum and need not be rejected if die-to-die (or wafer-to-wafer) variations are considered.

B. Tukey Test [30]

Tukey test also assumes the Normal distribution. Two points Lower Quartile (LQ) and Upper Quartile (UQ) are defined such that $1/4^{th}$ of total values are below than LQ and $1/4^{th}$ of total values are above UQ. The difference between these two values is called Inter Quartile Range (IQR).

$$IQR = UQ - LQ$$

Then the Upper Quartile Limit (UQL) is defined as

$$UQL = k.IQR + UQ$$

where k is the scaling factor.

All values exceeding UQL are considered illegitimate and rejected. The scaling factor (k) can be varied to control the threshold value. Obviously, the higher the value of k, the fewer is the number of chips rejected.

5. Application to I_{DDQ} Testing

Both methods mentioned earlier assume the Normal distribution of data. However, I_{DDQ} data does not follow the Normal distribution.

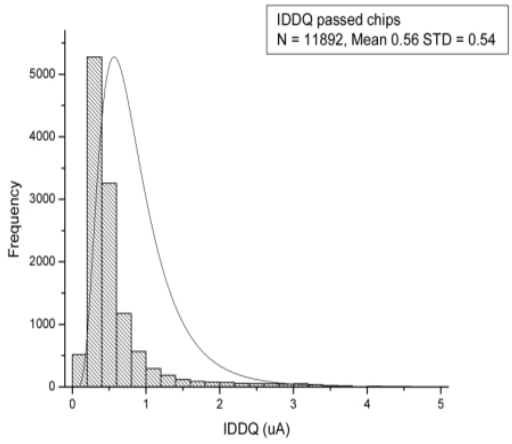

Figure 3: IDDQ distribution of good chips

Any I_{DDQ} reading has two components: (1) background current component which is the leakage current flowing through reverse biased PN junctions [9] and (2) defect current component which is the result of

[2] Such tables can be easily found in any standard book on statistics. Refer to [29], for example.

[3] An example of use of Chauvenet's criterion for data rejection is shown in the appendix.

defect in the chip (like a gate oxide short or a metallic bridge). Fault-free I_{DDQ} or the leakage current is a function of effective channel length (L_{eff}) and threshold voltage (V_T) of transistors, which in turn depends on the technology. Due to process variations, the effective channel length and threshold voltage follow a Normal distribution. Since the relation between I_{DDQ} and L_{eff} is exponential [10], the I_{DDQ} distribution is expected to be Lognormal[4] rather than Normal. The I_{DDQ} due to defective chips, however, does not follow any standard distribution. Its value depends on the nature of the defect, state of the device (i.e. input vector) and the resistance of the path taken by the current. Since we are interested in separating fault-free chips from the faulty chips, we convert the Lognormal distribution to the Normal distribution. This is achieved by taking the natural logarithm of the original data.

There are several sources of errors in the I_{DDQ} measurement. If I_{DDQ} readings are collected using a wafer probe, noise and contact resistance can corrupt the readings. The resolution of the measurement unit is another constraining factor. Thus the actual distribution will not exactly follow any standard distribution. For our purpose of analysis, we are interested in determining which distribution is closest to the actual distribution. The distribution of I_{DDQ} of die that passed 5μA threshold test is shown in Figure 3. It follows the Lognormal distribution quite closely. The distribution of all die on the same wafer is as shown in Figure 4. Notice that no single distribution can fit.

6. Results

In the SEMATECH experiment four different tests were conducted on a sample of 18466 chips [31]. In total, it contained data from 75 wafers. The I_{DDQ} testing was performed using a static threshold of 5 μA. For each die 195 I_{DDQ} readings were taken at the wafer level. If any of these 195 readings exceeded 5 μA current threshold, the die was considered to be an I_{DDQ} fail. Note that this threshold does not imply any good manufacturing limit [31]. All these chips were tested again after 6, 72 and 144 hours of burn-in as well.

[4] The standard Lognormal distribution has probability distribution function given by

$$f(x) = \frac{e^{-\frac{(\log x)^2}{2\sigma^2}}}{x\sigma\sqrt{2\pi}}$$

The Normal distribution has the probability distribution function given by

$$f(x) = \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$$

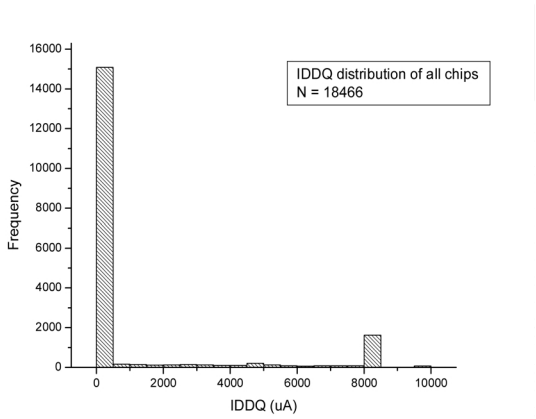

Figure 4: IDDQ distribution of all chips

Our analysis flow is as follows. All the analysis is done on per-wafer basis. There are 195 measurements for each die corresponding to 195 different vectors. For each vector we convert the data to the Normal distribution and find mean (μ_i) and standard deviation (σ_i). Then we apply both data rejection methods independently. In case of Chauvenet's criterion the probability threshold is changed to find the effect of the threshold on the number of dice rejected. In a similar manner, the UQL is changed for the Tukey test and effect was monitored.

The results for Chauvenet's criterion are summarized in Table 1 and those for Tukey test are shown in Table 2. An accepted chip is considered failed after burn-in if it fails any test including 5μA threshold test. A rejected chip is considered passed after burn-in if it passes all the tests. Table 3 shows results. The yield loss, overkill and defect level are computed as:

$$Yield\ \ Loss\ (YL) = \frac{\left(\frac{P}{K}\right)R + \left(\frac{F}{B}\right)A}{T}.100$$

$$Overkill = \frac{P}{K}(1-\frac{A}{T}).100$$

$$Defect\ Level(DL) = \frac{F}{B}.\frac{A}{T}.100$$

A = total number of accepted chips
B = number of accepted chips burned-in (BI)
F = number of accepted chips that failed BI
R = total number of rejected chips
K = number of rejected chips burned-in
P = number of rejected chips that passed after BI
T = A + R = total chips = 18466.

Thus, P/K represents the fraction of the rejected chips passed after burn-in and F/B represents the fraction of

the accepted chips failed after burn-in. Both contribute to the yield loss. The defect level is a measure of bad parts that got shipped while overkill is a measure of good chips that were considered bad. If a die failed at wafer level and passed all tests after BI, it is counted as overkill. Since we used SEMATECH test result after BI (which includes 5 μA I_{DDQ} test), the actual defect level and overkills for other static thresholds would be worse than those presented here.

Table 1. Results for Chauvenet's criterion

Threshold	A	B	F	R	K	P	YL %	Overkill %	DL %
0.1	13777	3051	315	4689	838	90	10.42	2.72	7.7
0.2	13282	2802	277	5184	1087	109	9.92	2.81	7.11
0.3	13022	2617	251	5444	1272	126	9.90	2.92	6.76
0.4	12857	2486	238	5609	1403	134	9.57	2.90	6.66
0.5	12680	2330	224	5786	1559	156	9.74	3.13	6.60
0.6	12531	2201	213	5935	1688	175	9.89	3.33	6.57

Table 2. Results for Tukey test

Threshold[5]	A	B	F	R	K	P	YL %	Overkill %	DL %
0.5	11550	1715	234	6916	2042	366	15.25	6.71	8.53
1	11752	1774	268	6714	1983	341	15.86	6.25	9.61
1.5	11870	1826	290	6596	1981	319	15.96	5.75	10.21
2	11969	1874	311	6497	1975	293	15.98	5.21	10.76
2.5	12039	1915	337	6427	1934	278	16.47	5.00	11.47
3	12093	1953	357	6373	1896	260	16.70	4.73	11.97

Table 3. Results for static threshold method

Threshold μA	A	B	F	R	K	P	YL %	Overkill %	DL %
1	14627	2174	745	3839	1715	435	32.41	5.27	27.14
2	15933	2258	753	2533	1631	359	31.79	3.01	28.77
5	16777	2331	759	1689	1558	292	31.29	1.71	29.58
10	17119	2653	911	1347	1236	122	32.55	0.72	31.83
20	17350	2871	1093	1116	1018	86	36.28	0.51	35.77
50	17590	3087	1292	876	802	69	40.27	0.41	39.87

7. Conclusions

This work has demonstrated that the use of a single threshold value for pass/fail limit setting is not judicial and can result in considerable yield loss and/or defect level. Statistical data rejection methods provide a simple yet powerful way to sustain I_{DDQ} testing for future technologies. Nevertheless various parameters used in their application must be carefully selected and justified. These methods are like a double-edged sword – if used carefully they provide powerful means of data analysis but improper selection of parameters can abuse them. Especially Chauvenet's criterion has lowest yield loss and defect level. However, true effectiveness of any method can be verified only with sufficient sample of data. The use of these methods is not justified if the sample space is limited. We believe that the data collected during wafer probe or characterization can be used for analysis and can prevent bad chips from being shipped. Of course, in order to account for lot-to-lot and wafer-to-wafer variations it may be necessary to perform the analysis more often and refine the threshold values using methods like regression fit. As a process becomes more mature a better goodness of fit can be obtained. Clearly, these methods do not provide a push-button solution to the problem, but if used intelligently, they can alleviate costly customer returns in the long term.

Acknowledgements

This research was funded in part by the National Science Foundation under grant CCR-9971102. Thanks to Phil Nigh of IBM for providing the SEMATECH data.

Appendix: Use of Chauvenet's criterion

Consider that the following readings were obtained from an experiment:

3.2, 3.4, 3.3, 3.6, 3.5, 2.8, 7.8, 2.9.

[5] Threshold is UQL + *k*. IQR. The scaling factor '*k*' is specified in this column.

We have $N = 8$, $\mu = 3.81$ and $\sigma = 1.53$. The reading 7.8 seems to be outside the normal expectation. In fact, it is $(7.8\text{-}3.81)/1.53 = 2.6\sigma$ away from the mean. From the error table [29], the probability that a reading can be 2.45σ away is 0.0142. Multiplied by 8, the number of readings in the data set, we get $n(\text{worse than } x_{sus}) = 0.11$. Thus probability of getting a reading so much away for such a small sample size is 1 in 10. So the chances of this reading being legitimate in a set of 8 readings are very poor. According to Chauvenet's criterion this reading is rejected. After rejection, we obtain $\mu_{new} = 3.24$ and $\sigma_{new} = 0.28$. Note that if the distribution is Normal, rejection of a single reading does not cause much reduction in the mean, but considerable reduction in the standard deviation of the data.

References

[1] Jerry Soden et al., "I_{DDQ} Testing: A Review," *Journal of Electronic Testing: Theory and Applications*, 1992.

[2] Roger Perry, "I_{DDQ} Testing in CMOS Digital ASICs," *Journal of Electronic Testing: Theory and Applications*, (*JETTA*), 1992.

[3] Steven McEuen, "I_{DDQ} Benefits," *Proc. of IEEE VLSI Test Symposium*, 1991.

[4] Steven McEuen, "Reliability Benefits of I_{DDQ}," *Journal of Electronic Testing: Theory and Applications*, 1992.

[5] K. Wallquist, "On the Effect of ISSQ Testing in Reducing Early Failure Rate," *Intl. Test Conf.*, 1995.

[6] Rochit Rajsuman, "I_{DDQ} Testing for CMOS VLSI," *Proceeding of the IEEE*, April 2000.

[7] M. Jacomino et al., "Fault Detection in CMOS Circuits by Consumption Measurement," *IEEE Trans. on Instrumentation and Measurement vol. 38, No. 3*, June 1989, pp.773-778.

[8] T. Aruna Unni and D. M. H. Walker, "Model-based I_{DDQ} Pass/Fail Limit Setting", *IEEE Intl. Workshop on I_{DDQ} Testing,* 1998.

[9] Ali Keshavarzi et al., "Intrinsic Leakage in Deep Submicron CMOS ICs – Measurement-Based Test Solutions," *IEEE Trans. on VLSI Systems*, Dec. 2000.

[10] Manoj Sachdev, "Deep Sub-micron I_{DDQ} Testing: Issues and Solutions," *Proc. of European Design & Test Conf.*, 1997.

[11] Yashwant Malaiya et al., "Resolution Enhancement in I_{DDQ} Testing for Large ICs," *VLSI Design vol. 1, no. 4*, 1994.

[12] Tsuyoshi Shinogi and Terumine Hayashi, "An Iterative Improvement Method for Generating Compact Tests for I_{DDQ} Testing of Bridging Faults," *Institute of Electronics, Information and Communication Engineers (IEICE) Trans. on Information and Systems, Vol. E81-D, No.7,* July 1998.

[13] Doug Josephson et al., "Microprocessor I_{DDQ} Testing: A Case Study," *IEEE Design & Test of Computers*, Summer 1995.

[14] Semiconductor Industries Association, *International Technology Roadmap for Semiconductors*, 1999.

[15] Peter Maxwell and Robert Aitken, "I_{DDQ} Testing as a Component of a Test Suite: The Need for Several Fault Coverage Metrics," *Journal of Electronic Testing: Theory and Applications"*, 1992.

[16] Charles Hawkins and Jerry Soden, "Deep Sub-micron IC Testing: Is There A Future?," *IEEE Design & Test of Computers*, Oct-Dec. 1999.

[17] Brian Stine et al., "Analysis and Decomposition of Spatial Variation in Integrated Circuit Processes and Devices," *IEEE Trans. on Semi. Manufacturing Vol. 10, No. 1*, Feb. 1997

[18] Anne Gattiker and Wojciech Maly, "Current Signatures," *IEEE VLSI Test Symposium*, 1996.

[19] Anne Gattiker and Wojciech Maly, "Current Signatures for Production Testing," *IEEE Intl. Workshop on I_{DDQ} Testing*, Oct. 1996.

[20] C. Thibeault, "A novel probabilistic approach for IC diagnosis based on differential quiescent current signatures," *IEEE VLSI Test Symposium*, 1997.

[21] C. Thibeault, "An Histogram based procedure for current testing of active defects," *Intl. Test Conference*, 1999.

[22] Peter Maxwell et al., "Current Ratios: A Self-scaling Technique for Production I_{DDQ} Testing," *Intl. Test Conference*, 2000.

[23] Sri Jandhyla et al., "Clustering Based Identification of Faulty ICs Using I_{DDQ} Tests," *I_{DDQ} Workshop*, 1998.

[24] Sri Jandhyla et al., "Clustering Based Techniques for I_{DDQ} Testing," *Intl. Test Conference*, 1999.

[25] W. R. Daasch et al., "Variance Reduction Using Wafer Patterns in I_{DDQ} Data," *Intl. Test Conference*, 2000, pp.189-198.

[26] A. D. Singh, "A Comprehensive Wafer Oriented Test Evaluation (WOTE) Scheme for the I_{DDQ} Testing of Deep Sub-Micron Technologies", *IEEE Intl. Workshop on I_{DDQ} Testing,* 1997.

[27] Adit Singh et al., "Binning for IC Quality: Experimental Studies on the SEMATECH Data", *Intl. Symp. on Defect and Fault Tolerance in VLSI Systems,* 1998.

[28] Stuart Meyer, "Data Analysis for Scientists and Engineers," *John Wiley & Sons,* 1975.

[29] J. R. Taylor, "An Introduction to Error Analysis," 1st Ed., *University Science Books, CA*, 1982.

[30] Ronald Richmond, "Successful Implementation of Structured Testing," *Intl. Test Conference,* 2000 Atlantic City, pp. 344-348

[31] Phil Nigh et al., "An experimental study comparing the relative effectiveness of functional, scan, IDDQ and delay fault testing," *IEEE VLSI Test Symposium*, 1997.