# Clustering Based Evaluation of IDDQ Measurements: Applications in Testing and Classification of ICs

Sri Jandhyala[*], Hari Balachandran[*], Manidip Sengupta*, Anura P. Jayasumana[**]

[*]Texas Instruments Inc., P.O. Box 660199, Dallas, TX 75243
[**]Colorado State University, Dept. of Electrical and Computer Engr., Fort Collins, CO 80523
email: srij@ti.com

## Abstract

*Effectiveness of the clustering based approach in detecting devices with abnormal $I_{DDQ}$ values is evaluated using data from the SEMATECH test methods experiment. The results from clustering are compared to the results obtained on actual silicon during the SEMATECH study. The differences between the results obtained in each case are analyzed. The clustering approach is also compared to two common $I_{DDQ}$ test techniques, the single-threshold approach and the delta-$I_{DDQ}$ approach, and the results are presented.*

## 1. Introduction

The inapplicability of existing $I_{DDQ}$ test techniques to differentiate between normal and abnormal quiescent currents has become a major concern for manufacturers. With the advent of more advanced design, process and packaging technologies, this problem is bound to get worse [1]. It has become increasingly accepted that $I_{DDQ}$ testing in its present form cannot be applied in the near future. In an effort to address this growing concern, several innovative and alternative approaches have been proposed [2-10]. More such newer ways of looking at the problem of decreasing distinction between defective and defect-free quiescent currents of an IC are essential for the continued use of $I_{DDQ}$ testing in the future.

One such approach presented was the Cluster analysis based approach [5, 8]. Rather than looking at the $I_{DDQ}$ measurements of a device in isolation, the process and benefits of applying statistical cluster analysis techniques to arrive at device quality predictions was introduced. In this approach, the current distribution of an IC, which is the set of $I_{DDQ}$ measurements for all applied vectors, is compared to those of other ICs. Based on the similarities of these current distributions, calculated using cluster analysis techniques, ICs are assigned to different “groups” or “clusters”. The benefits to this grouping, in contrast to existing approaches where each IC is considered in isolation, is in using the natural classification of the devices to eliminate the need for a static threshold. It also eliminates the need for setting a static threshold, which is required for other existing techniques. There are several other benefits including the “binning” of devices, identifying a smaller set of candidates for physical failure analysis and achieving rapid ramp-to-volume during production by allowing for the faster identification and diagnosis of early process issues.

For the clustering based approach, each IC is represented as a point in the n-dimensional space of $I_{DDQ}$ measurements, where n is the number of vectors applied. The output of this process are groups of devices that are in close proximity to each other with respect to $I_{DDQ}$ values measured across all the applied vectors. In a typical scenario, one large group containing the majority of devices with normal $I_{DDQ}$ values and several smaller groups with abnormal $I_{DDQ}$ values can be expected. The goal is to allow the clustering process to classify the devices as opposed to using an arbitrary or pre-determined threshold.

The paper begins with the objectives behind the work in section 2. Brief background information on the data used for this study is presented next. In section 4, results from the application of the clustering based approach to this data are presented. Correlation between the results obtained from the clustering technique and those obtained from the SEMATECH study is done and the results are presented in section 5. A comparison of the results from clustering to single threshold and delta $I_{DDQ}$ is included in section 6 along with an explanation on the reasons for the inconsistencies in the results obtained from each technique. Finally, conclusions from the study and on-going work are outlined in section 7.

## 2. Objectives

The main objective is to evaluate the effectiveness of the clustering approach. This is accomplished by performing two comparisons. First, the results from clustering are compared to the results obtained during the SEMATECH test methods experiment on real silicon [11,12]. The primary reason behind electing to make this comparison with the SEMATECH data was both the availability of this data to a large audience of

0-7695-0613-5/00 $10.00 © 2000 I

SEMATECH member companies and the level of understanding and acceptance of this data due to prior work and involvement by a number of companies. It also allows for the continuation of such analyses in the future when newer techniques are available and, for scrutiny of the results presented here. Since some of the devices were subjected to burn-in during the SEMATECH study, as part of our analysis, we also hope to identify potential candidates for burn-in through the clustering technique. Results obtained when clustering is performed on the same data set with different number of clusters formed are presented. The motive here was to analyze the differences in device quality predictions when better resolution is desirable and/or feasible.

The second comparison undertaken was the clustering approach to other existing $I_{DDQ}$ test techniques. Two popular $I_{DDQ}$ techniques in use were chosen and each of these methods were applied to the SEMATECH data. The results obtained from this application were compared in each case and the differences are detailed. The two methods that were finally chosen for this comparison were the single threshold approach and the delta $I_{DDQ}$ approach [3, 6, 7, 9, 10]. The reasoning behind the selection of these two approaches was the wide use of these techniques for $I_{DDQ}$ testing in production environments. In addition, the delta $I_{DDQ}$ technique has garnered a lot of interest and is believed to hold a lot of promise for the continued use of $I_{DDQ}$ testing for future technologies. It also forms the basis for a number of the newer techniques proposed including the current signatures approach for production testing [3].

## 3. Background on Data

The data used for the analyses here was obtained from the SEMATECH test methods experiment where approximately 18,000 devices were subjected to four main kinds of testing – Stuck-at, $I_{DDQ}$, Functional and Path Delay. Irrespective of whether a die failed one of the earlier tests, all 4 tests were performed. For each die, 195 $I_{DDQ}$ measurements were taken regardless of the magnitude of the reading. The measurements were clamped at 8mA. For determining whether a die passed the $I_{DDQ}$ test, a single threshold of 5uA was used during the SEMATECH study. This 5uA limit was purely arbitrary and yield loss was not a consideration while determining it. The 5uA limit was established only because a reasonable limit was needed to define IDDQ-fails for the purposes of choosing die to undergo post-wafer testing. Depending on the kind of analysis being done and the objective, this limit could vary quite widely. From the initial 18,000 devices, approximately 4,000 devices were identified for burn-in. The burn-in sample included a small percentage of good devices along with a representative sample of different kinds and combinations of fails, with a higher number selected from $I_{DDQ}$-only fails. All these 4,000 devices were subjected to at least 6 hours of burn-in. After this, around a fourth of the die (~1000) were carefully selected for further burn-in of 72-hours and 144-hours, around 2/3rds of which were $I_{DDQ}$-only failures.

For the purposes of this study, from the initial set of 18,000 devices, all die that had $I_{DDQ}$ measurements over 80uA were discarded. From our analysis and the data, it was felt that 80uA was a large enough $I_{DDQ}$ threshold beyond which any die would be considered defective. Table 1 shows the number of die that "pass" at different single threshold levels. In addition, only 20 out of the total 195 $I_{DDQ}$ measurements were considered. Though a thorough analysis was not done on whether any of the 20 $I_{DDQ}$ measurements would provide the same coverage and pass/fail numbers, a preliminary analysis indicated that they did not vary by much for any 20 chosen at random. Due to this, the first 20 $I_{DDQ}$ measurements were chosen for this study. Thus, the final data presented in this paper includes 13,162 die and the first 20 $I_{DDQ}$ measurements for each of these die - in essence, a matrix 13162x20.

| **Pass/Fail Threshold** (in uA) | **# of Devices Passing** |
|---|---|
| 1 | 9949 |
| 5 | 12213 |
| 8 | 12441 |
| 10 | 12545 |
| 15 | 12714 |
| 20 | 12800 |
| 25 | 12856 |
| 30 | 12906 |
| 80 | 13162 |

Table 1: "Good" devices at different thresholds

## 4. Results from Cluster Analysis

All the devices were subjected to clustering analysis across the 20-dimensional space of $I_{DDQ}$ measurements considered. One of the inputs to the clustering technique are the number of clusters to be formed. The results from forming 5 clusters and 10 clusters are shown in figures 1 & 2 respectively. The horizontal axis in the figures indicates the cluster number and the number of devices in each cluster. Within each cluster, the $I_{DDQ}$ distributions for each of the 20 vectors are shown as the vertical bars. The marker on these bars represents the median values of $I_{DDQ}$ measurements in the corresponding vector.

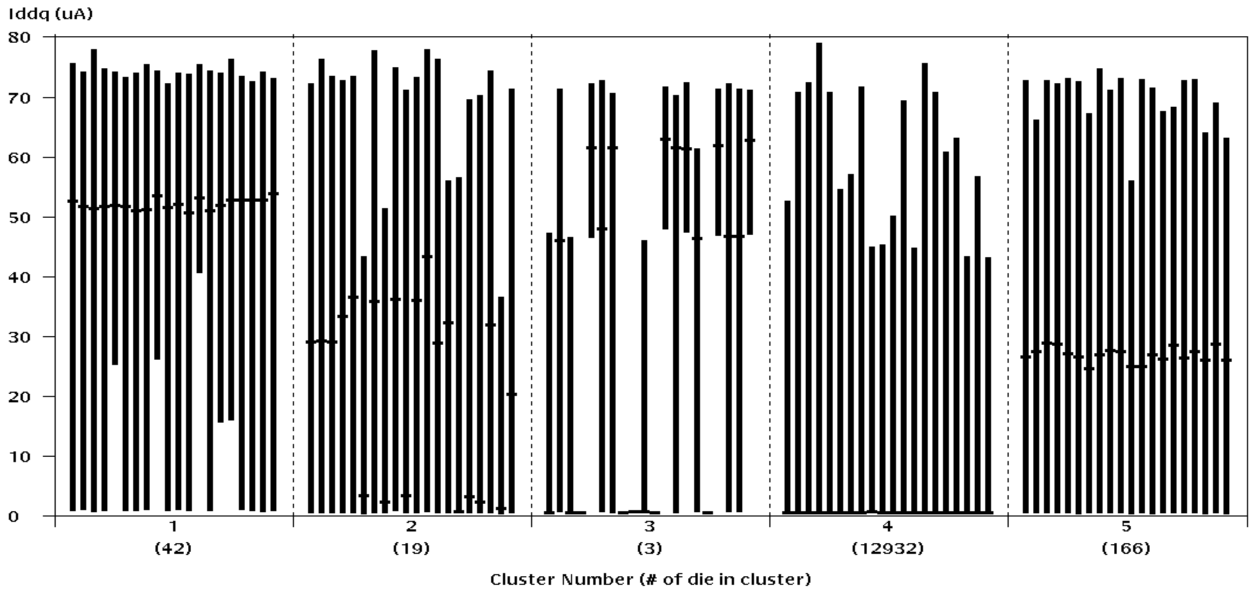


Figure 1: Current Ranges of Different Vectors for Devices in Each Cluster – 5 Final Clusters

When 5 final clusters were specified, 12932 out of the total 13162 devices were classified into cluster 4. The devices in this cluster can be categorized as having passed $I_{DDQ}$ testing when the clustering technique is applied, without using a threshold level. The remainder of devices were grouped into the other four clusters and exhibit characteristics different from cluster 4. Due to the variation in the distribution across different vectors and the median value for each vector, all of the nineteen devices in cluster 2 and the three devices in cluster 3 can be conclusively tabulated into the defective class. Note that the median $I_{DDQ}$ values for different vectors in cluster 3 differ by as much as 60uA whereas the variation is only a couple of uA for devices in clusters 1 and 5. There is a good likelihood that the 166 devices in cluster 5, with median values ranging from 27uA to 32uA, are defect-free. The higher current measurements are likely due to excess leakage current in relation to the other devices but not due to the presence of random manufacturing defects. A similar argument could be made about the 42 devices in cluster 1 though this cluster has median values ranging from 50uA to 52uA. A case could be made that though these devices exhibit higher median values, since they are consistent across all the 20 applied vectors, they are likely to be defect free. However, the more judicious conclusion would be to classify them as being indeterminate from the first round of clustering and perform an additional iteration on just these devices. This would further separate the devices by weeding out high current causing devices which were

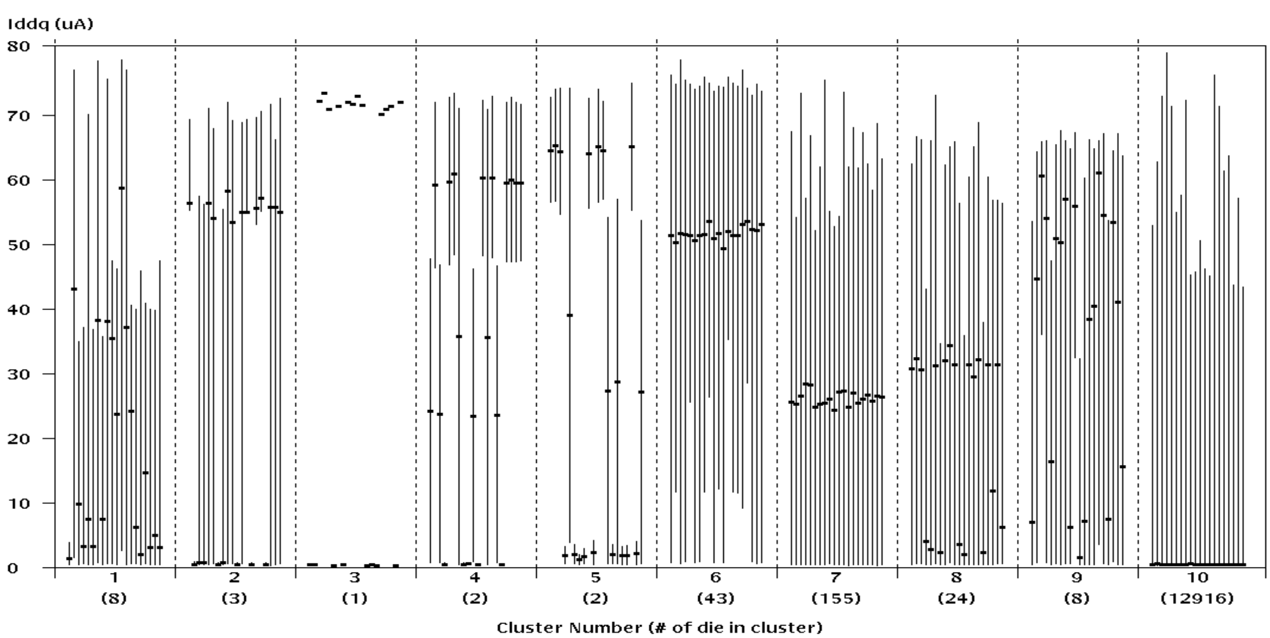


Figure 2: Current Ranges of Different Vectors for Devices in Each Cluster – 10 Final Clusters

grouped along with devices exhibiting lower currents due to the restriction placed by the number of clusters formed. Based on this analysis, in the case when five clusters were formed, 98.2% of the devices were binned into the "pass" group, 0.16% of the devices into the "fail" group, and 1.5% of devices flagged as inconclusive from clustering.

In the case when 10 final clusters were formed (figure 2), the majority of devices (~98%) fell into cluster 10. All these devices can be predicted to be good devices i.e. devices that pass $I_{DDQ}$ testing using the clustering approach. Of the remaining devices, more than half of them were in cluster 7. Though this cluster has a higher median for all devices, due to the fact that median values across all vectors are nearly identical, these can also be considered to be defect-free with a higher leakage current. If this same distribution resulted during the initial testing of a new product, the more prudent approach would be to consider a small sub-set of these devices and perform physical failure analysis in order to determine the cause of the elevated leakage currents. This higher current may have been due to leaky transistors as opposed to a defect being present across all devices. The same could be said about the devices in cluster 6 (43). However, since the median current is even higher now, until the reason behind the higher current is established no concrete result about the nature of these devices, as with those in cluster 7, can be conclusively made. All of the other clusters that contained the remaining devices can be labeled defective from $I_{DDQ}$ testing and hence, rejected. One interesting point to note is the current distribution of the 24 devices in cluster 8. Though there are a number of low-current causing vectors, there are also several vectors that caused a high current. It is more likely that these vectors did indeed elevate the quiescent current due to the presence of a defect. Clustering was able to conclusively determine that 98.13% of the devices "passed", 0.37% of the devices "failed", and the quality of the remaining 1.5% of devices (198 in total) could not be confidently established from $I_{DDQ}$ testing. As stated in earlier work, these can be put through another iteration of the clustering process after which higher confidence on the prediction of device quality can be achieved.

## 5. Correlation of Clustering Results to SEMATECH Results

In order to verify the device quality predictions from clustering, the results were compared to physical data obtained from the SEMATECH study. The goal was to verify that devices that were predicted to be defect-free were indeed free from defects in actual silicon. For the purpose of comparison, all the cluster 10 devices (figure 2) that were considered "good" were used.

The results from this comparison are summarized in table 2. The first row of the table shows the number of devices that passed the clustering technique. In the second row, the number of "clustering pass" devices (from row one) that also passed all the SEMATECH tests at wafer probe are shown. Devices that failed the $I_{DDQ}$ tests in the SEMATECH study (using a single threshold of 5uA) were considered as having passed i.e. the $I_{DDQ}$ test classifications from the SEMATECH study were not used. Row 3 in table 2 shows the difference between the number of devices in row1 and row2. This indicates that there were a total of 845 devices that clustering deemed to be defect-free based on their $I_{DDQ}$ measurements but had failed one of the functional, stuck-at or path delay tests during the SEMATECH study. The fourth row represents those devices, out of the total 845, that passed all tests after burn-in during the SEMATECH study. The last row represents the number of devices that clustering considered to be defect-free but failed one of the three tests (stuck-at, path delay or functional) in the SEMATECH study even after burn-in.

| Devices deemed pass/fail | # of Devices |
|---|---|
| Pass from Clustering | 12916 |
| Passed at Wafer Probe | 12071 |
| Failed at Wafer Probe | 845 |
| Passed after Burn-In | 404 |
| Failed after Burn-in | 441 |

Table 2: Correlation of Clustering and SEMATECH Results

The results in table 2 show that of the 865 devices that failed one of the tests at wafer probe, almost half of them passed after burn-in. The majority of these 404 devices went from being delay-fails at wafer to delay-pass after burn-in. This change was most likely due to the change in test temperature in the two cases and not because they became defect-free after burn-in. This also indicates the likelihood that these devices exhibited lower $I_{DDQ}$ measurements across the 20 vectors, based on which the clustering technique grouped them into the "good" cluster. In essence, the clustering approach considered them as being defect-free based on their low $I_{DDQ}$ readings because of which this cannot be cited against the clustering technique when evaluating its effectiveness. The last row in table 2 also indicates that half of the devices that clustering thought to be defect free were in reality, defective. This implies that though these devices exhibited lower $I_{DDQ}$ distributions, due to which the clustering technique classified them as being "good", defects were present in these devices. This also suggests that the clustering technique was not very effective in screening out bad parts. From initial analysis of the reasons behind the clustering technique labeling defective devices to be defect-free, it was noticed that a common trait in a majority of these devices was low $I_{DDQ}$ measurements across the 20 vectors considered for our experiment. In figure 3, the $I_{DDQ}$ measurements across all 20 vectors for two devices which "passed" the clustering approach but failed in silicon are shown. Both these devices failed all tests in silicon (including the $I_{DDQ}$ test with a 5uA threshold). As can be seen, device-1 had $I_{DDQ}$ measurements for all 20 vectors under 1uA. Device-2 had measurements ranging from 3 to 8.5uA. This indicates that the vector(s) that caused device-1 to show a reading above 5uA was not present in the 20 vectors chosen. This is not to claim that none of the devices in the "good" set (from the clustering technique) did not exhibit high $I_{DDQ}$ measurements. But there is a reason behind the clustering approach classifying these devices with a high current along with devices that had very low currents into the same cluster.

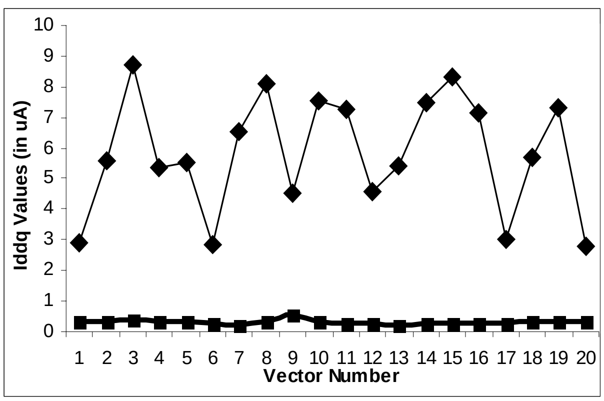


Figure 3: Iddq Values for 2 different devices

The results showed proof supporting two very important and generally accepted facts in the industry. One: $I_{DDQ}$ testing alone is not sufficient to achieve desired quality levels. As exemplified by the fact that devices that failed all tests in silicon did not exhibit high $I_{DDQ}$ measurements. Unless a device exhibits high $I_{DDQ}$ measurements in the presence of a defect, no $I_{DDQ}$ technique will be able to flag it from within a larger set. Two: even though a device may exhibit higher $I_{DDQ}$ currents, it does not necessarily translate into that device being defective. This was proved by the fact that some of the devices that exhibited high $I_{DDQ}$ currents, did not fail any of the tests even after extended hours of burn-in.

## 6. Comparison of Clustering to Existing Techniques

While comparing the clustering technique to the single threshold and delta-$I_{DDQ}$ approaches, one of the first issues that had to be overcome was the threshold to be used for making the pass/fail decisions. In order to make a fair comparison, what would be the right threshold value? There is no "right" answer to this question. For the single threshold approach in this comparison, based on the magnitude and the number of devices that were rejected based on this limit, a pass/fail limit of 15uA was chosen. The difference between the number of devices that were rejected with a threshold limit of 10uA did not vary by much from the number of devices rejected when a threshold of 30uA was used (table 1). While selecting the threshold value, another consideration was to keep the number of devices passing the single threshold method and the clustering approach to an equal number. A case could be made for setting this limit to a higher or to a lower value and surely, there would have been valid points to consider in either case. Similarly, for the delta $I_{DDQ}$ approach, two such limits had to be chosen. One upper limit, which could be relatively high, to be used as the absolute value beyond which an IC would be considered defective. A second, more tighter limit, which would be used for measuring the difference, or the delta, between two vectors of the same IC. The upper limit selected was 30uA and the delta limit chosen was 10uA. Using these limits, devices that were classified as defective and defect-free were compared to the case when 10 clusters were formed using the clustering technique and the results are shown in table 3.

For the single threshold approach, 12714 devices passed the 15uA limit. All of these devices were present in cluster 10 of figure 2. This also implies that there were 202 other devices that were present in cluster 10, but were rejected using the single threshold approach. Our analysis revealed that 147 of the 202 devices were passing in the SEMATECH data as shown in row 2 of table 3. In the delta $I_{DDQ}$ approach, with the loose upper limit of 30uA, 12906 devices were deemed defect-free. When these devices were subjected to the 10uA delta-limit, a total of 12677 devices were classified as "good" devices. All the 12677 devices that passed delta $I_{DDQ}$ were in cluster 10 – the same cluster that was deemed to be defect-free from the clustering analysis. There were 239 devices rejected by delta $I_{DDQ}$ that were also present in cluster 10. After correlation analysis against the SEMATECH data, it was found that 214 of these 239 devices were considered defect-free from the SEMATECH experiment (row 2 of table 3).

| Number of devices that "pass" Clustering: 12916 | Single Threshold | | Delta | |
|---|---|---|---|---|
| | **Pass** | **Fail** | **Pass** | **Fail** |
| **# of devices that "pass" or "fail" each Iddq technique** | 12714 | 202 | 12677 | 239 |
| **# of failing devices in each technique that passed during the SEMATECH study** | 147 | 55 | 214 | 25 |

Table 3: Comparison between Results from Clustering and Single Threshold & Delta Iddq Approaches

## 7. Conclusion

From the comparison results presented, we believe that the clustering approach proved to be the technique that held the most promise in predicting device quality based on the $I_{DDQ}$ measurements. It certainly did not reject, or classify as defective, devices that passed one of the other two techniques. If anything, it could be said that the clustering approach was more optimistic about the quality of the devices. The results of the tests that were done during the SEMATECH experiment further validate this. When compared to the other two approaches, single threshold and delta $I_{DDQ}$, it may appear that there are some test escapes in the clustering approach. After correlation with the results that were seen in actual silicon, however, it appears that the majority of the devices that were classified as defect-free were indeed $I_{DDQ}$-only fails. There were also some clustering-pass devices that failed in real silicon. But, these can be explained as defects that cannot be detected through the $I_{DDQ}$ measurements. This argument would certainly hold based on the wafer level $I_{DDQ}$ measurements that were taken during the SEMATECH study.

Work on analyzing these discrepancies is ongoing with the aim of presenting the results at a future forum. Analysis on the number of vectors that are required before a steady state is reached on the distribution of devices into a particular cluster is also being done. For future work, investigation into the reasons for the higher medians in some of the clusters and any information that can be gleaned from the vector to vector variations within a cluster would be very interesting. The practical issue of implementing the clustering technique in production is also being investigated.

## 8. References


1. T. W. Williams, et al, "$I_{DDQ}$ Test: Sensitivity Analysis of Scaling", Proc. 1996 IEEE International Test Conference, pp. 786-792.
2. A. Gattiker and W. Maly, "Current Signatures", Proc. 1996 IEEE VLSI Test Symposium, pp. 112-117.
3. Anne Gattiker, et al, "Current Signatures for Production Testing", Proc. 1996 IEEE International Workshop on $I_{DDQ}$ Testing, pp.25-28.
4. A. Gattiker and W. Maly, "Current Signatures: Application", Proc. 1997 IEEE International Test Conference, pp. 156-165.
5. S. Jandhyala, et al, "Clustering Based Identification of Faulty ICs Using $I_{DDQ}$ Tests", Proc. 1998 IEEE Int. Workshop on $I_{DDQ}$ Testing, pp. 48-53.
6. C. Thibeault, "A Histogram-based Procedure for Current Testing of Active Defects", Proc. 1999 IEEE International Test Conference, pp. 714-724.
7. A. C. Miller, "$I_{DDQ}$ Testing in Deep-submicron Integrated circuits", Proc. 1999 IEEE International Test Conference, pp. 730-738.
8. S. Jandhyala, H. Blachandran, A. Jayasumana. "Clustering Based Techniques for $I_{DDQ}$ Testing", Proc. 1999 IEEE International Test Conference, pp. 730-738.
9. P. Maxwell, et al, "Current Ratios: A Self-Scaling Technique for Production $I_{DDQ}$ Testing", Proc. 1999 IEEE International Test Conference, pp. 738-747.
10. Theo Powell, et al, "Delta $I_{DDQ}$ for Testing Reliability", Proc. 2000 IEEE VLSI Test Symposium.
11. P. Nigh, et al, "An Experimental Study Comparing the Effectiveness of Functional, Scan, $I_{DDQ}$, and Delay Testing", Proc. 1997 IEEE VLSI Test Symposium", pp. 459-464.
12. P. Nigh, et al, "Failure Analysis of Timing and $I_{DDQ}$-only Failures from the SEMATECH Test Methods Experiment", Proc. 1998 IEEE International Test Conference, pp. 43-52.