Triangulating to a Defect's Physical Coordinates Using Multiple Supply Pad I_{DDQ}s: Test Chip Results

Jim Plusquellic, Dhruva Acharyya, Mohammad Tehranipoor and Chintan Patel
plusquel, adhruva1, abhishek, tehrani, cpatel2@umbc.edu
*Chips designed while on sabbatical at IBM Austin Research Laboratory

Department of CSEE, University of Maryland, Baltimore County

Abstract

Quiescent Signal Analysis (QSA) is an I_{DDQ} method for detecting defects that is based on the analysis of multiple simultaneous measurements of supply port I_{DDQ}s. The nature of the information in the multiple I_{DDQ}s measurements also allows for the localization of the defect to physical coordinates in the chip. In previous work, we derived a hyperbola-based method from simulation experiments that is able to "triangulate" the position of the defect in the layout. In this paper, we evaluate the accuracy of this method using data collected from 12 chips fabricated in a 65 nm process.

1.0 Introduction

Diagnosis is a process designed to identify the location of the fault in chips that have failed in the field or at production test. It is a key component to failure analysis. The information gleamed from failure analysis is used to tune the fabrication process for the purpose of improving reliability and yield.

Hardware-based fault localization is challenged by increases in chip complexity as well as additional interconnection levels and the limitations on the spatial resolution of imaging technology [1]. The increase in difficulty and cost of performing hardware physical failure analysis is likely to move it into a sampling/verification role. These trends continue to increase the importance of developing alternative software-based fault localization procedures.

Several "software-based" diagnostic methods have been proposed based on I_{DDQ} measurements [2-9]. These methods can be classified as static, quasi-static and dynamic diagnostic test paradigms. For static, the diagnostic test set and test response are precomputed and stored in a fault dictionary. The quasi-static paradigm, the test set is pre-computed but the fault dictionary is eliminated. Instead, the test response is computed dynamically. Under the dynamic paradigm, both the diagnostic test set and response are computed dynamically during response analysis.

The method that we propose in this work is a new approach to diagnosis and cannot be classified under these paradigms. It is complementary to these strategies and can be used in combination with them as a means of further improving diagnostic resolution. Moreover, QSA is more robust to the detrimental effects of increasing background leakage currents than these methods. This is true because the total (chip-wide) leakage current distributes across the multiple supply ports of the chip, and is reduced in magnitude in each of the supply port measurements.

In previous works, we developed statistical defect detection and localization methods using simulation experiments. In [10], a hyperbola-based diagnostic method is proposed that is able to "triangulate" a defect's location to a physical position in the layout of the chip. The method accomplishes this by computing the parameters for a pair of hyperbolas from the I_{DDQ}s measured at neighboring supply ports. The intersection of the hyperbolas identifies the predicted location of the defect in the layout.

A calibration circuit (CC) is proposed in [10] as a means of solving several problems associated with this type of localization scheme. Calibration circuits are inserted into the design at positions directly beneath the supply ports, and through scan chain control, allow for the controlled insertion of a short between the power and ground rails. The I_{DDQs} measured under the calibration tests define the upper and lower bounds of the hyperbola parameters and allow the equations to be solved. Secondly, the CC data is used to *calibrate* the I_{DDQ}s measured under a failing I_{DDQ} pattern as a means of diminishing the detrimental effects introduced by non-uniformities in the resistance of the power supply connections to the chip.

In this work, we elaborate and expand on the method described in [10] and evaluate its accuracy on twelve copies of a test chip. The test chips are fabricated in a 65 nm, 10 metal layer technology and incorporate an array of test structures that allow a defect to be emulated in one or more of 4,000 distinct locations on the chip. The design permits control over the magnitude of the emulated defect current and leakage current. The results of our analysis show that most defects can be located to a region less than 100 microns in diameter.

2.0 Test Chip Design

A block diagram of the test chip design is shown in Figure 1(a). It consists of a 80x50 array of test circuits (TCs) that occupies an area 558 μm in width and 380 μm in height. Each TC consists of three FFs connected in a scan

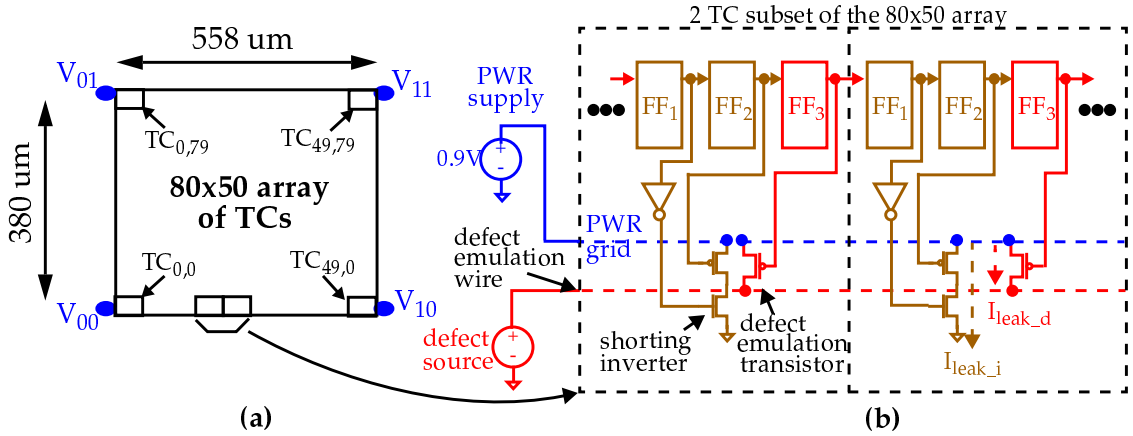

Figure 1. (a) Block diagram of the test structure and (b) details of the test cells (TC).

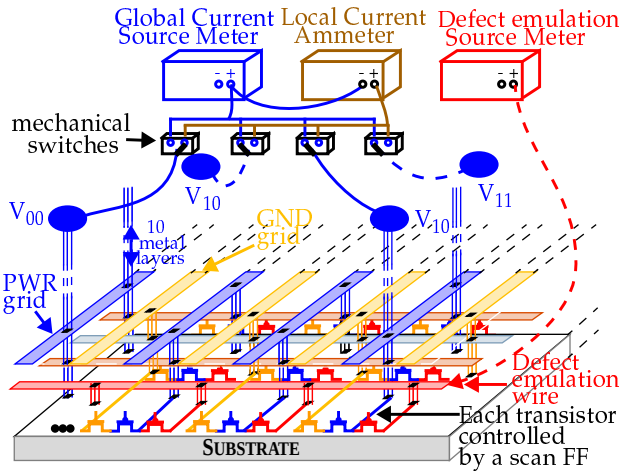

Figure 2. External Instrumentation Setup.

chain configuration, a *shorting inverter*, and a *defect emulation transistor* connected to a globally routed *defect emulation wire*. A schematic diagram of two adjacent TCs is shown in Figure 2(b). The shorting inverters and defect emulation transistors within each TC connect to the same point on the power grid.

The connection of the shorting inverters and the defect emulation transistors to point sources on the power grid enable two types of shorts to be introduced within any one (or more) of the 4,000 TCs. The first type shorts the power grid to ground through the inverter using FF_1 and FF_2 and the second type shorts the power grid to the defect emulation wire using FF_3. For the first type, the magnitude of the shorting current is defined by the external power supply voltage, labeled *PWR supply* in Figure 1(b).[1] For the second type, the magnitude of the shorting current is controlled through an external voltage source, labeled *defect source*. Given this configuration, a defect can be emulated at any point in the array by setting the defect source to a value less than the PWR supply voltage and scanning a bit pattern into the scan chain such that exactly one FF_3 contains a 0 and the remaining 11,999 FFs contain 1's.

In addition to controlling the magnitude of the defect current, the defect source also influences the magnitude of the background leakage current, as measured through the PWR supply. The total leakage current can be decomposed into two types, shown as $I_{leak\_i(nverter)}$ and $I_{leak\_d(efect)}$ in the right-most TC of Figure 1(b). Given the defect emulation wire connects to the drains of 4,000 defect emulation transistors, only one of which is enabled in a particular experiment, the remaining 3,999 transistors will source leakage current from the PWR supply proportional to the magnitude of the defect source voltage. This leakage, I_{leak_d}, adds to the leakage current already present through the shorting inverters, I_{leak_i}. Therefore, it is possible to analyze a variety of shorting and leakage current configurations by controlling the states of the defect emulation transistors and voltage on the defect emulation wire.

The external instrumentation setup is shown in Figure 2. As indicated above, the power ports, labeled V_{00} through V_{11} wire out of the chip on separate pins in the package. The individual power pins are each wired to a low resistance mechanical switch as shown along the top portion in Figure 2. The switch can be configured in a left or right position. The left and right outputs across the switches connect to a common wire that routes to the *Global Current Source Meter* (GCSM) and *Local Current Ammeter* (LCA), respectively.

The GCSM is configured to provide 0.9 Volts to the PWR grid and is also able to measure current with accuracies less than 100 nA. The LCA is wired in series with the GCSM and allows the individual power port (*local*) currents to be measured at the same level of accuracy. As an

1The PWR supply is held constant in our experiments at 0.9 V.

example, the configuration of the switches as shown in Figure 2 allow the local V_{00} current, I_{00}, as well as the global current to be measured. The *Defect Emulation Source Meter* (DESM) is used to set the voltage of and measure the current, I_{def}, through the defect emulation wire on a separate pin in the package (not shown).

3.0 Power Grid Characterization Experiments

The first set of experiments is designed to determine how the grid resistance influences the magnitude of the local currents. In these experiments, the defect emulation wire is disconnected and the defect emulation transistors are disabled. Instead, the shorting inverters are used to provide the stimulus to the grid.

In these experiments, each of the 4,000 shorting inverters from one of the chips is enabled, one at a time, and the global and local currents are measured. Given that we are interested in the characteristics of the grid resistance and its influence on the local current distributions from point sources in the layout, the following steps are also performed. After testing each element of the array, the shorting inverter of the TC under test is disabled and the global and local leakage currents are measured and subtracted from the values measured with the shorting inverter enabled. These *difference currents* are then *normalized* by dividing by the global current. This type of normalization virtually eliminates the variations in the transistor current magnitudes caused by process variations.

Figure 3 shows the current profile derived from the normalized local currents, I_{norm_00}. Here, the *x* and *y* axes represent the (*x*, *y*) plane of the TC array and I_{norm_00} is plotted on the *z* axis. The local currents are largest near V_{00} because TCs near this location draw a larger fraction of their current from V_{00} (maximum is approximately 31%) than TCs that are further removed. The degree to which the grid resistance influences the distribution of currents to the V_{DD}s is reflected in the range, which is approximately 11%.

The smooth monotonically decreasing nature of the surface in Figure 3 provides important "local" information that can be leveraged for the purpose of detecting and localizing defects. Although the current profiles for the remaining three V_{DD}s are not shown, the same characteristic shape exists in their surfaces except the orientation is different. A key feature embedded across these profiles is the unique mapping that exists between the supply port I_{DDQ}s and the position of the enabled TC in the 2-D array. In other words, the set of I_{DDQ}s measured can be used to distinguish among the TCs in the array. This is the basis on which our localization algorithm is based.

Assuming that the power grid resistance characteristics remain relatively consistent from one chip to another, i.e., they are minimally affected by process variations, a straightforward way to perform defect localization is to create a *lookup table* that maps each position in the layout to a set of supply port I_{DDQ}s. This can be accomplished using simulation experiments and fast power grid simulators, such as ALSIM [11]. For defect localization, a reverse lookup is performed using the I_{DDQ}s measured from a defective chip. Reference [12] gives the results of applying this type of methodology to a set of chips.

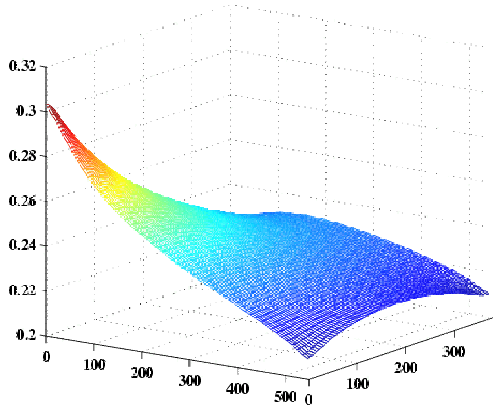

Figure 3. (a) I_{norm_00} profile.

The method investigated in this work is based on an analytical model, and therefore does not require a lookup table. The analytical model is derived from the current profile characteristics of the power grid, such as those shown in Figure 3. Although it is possible to derive a model directly from the normalized I_{DDQ}s as shown in the figure, the global normalization performed to produce this data is likely to be less effective in large power grids in which the magnitude of the total current is much larger and subject to a wider range of process variations. Therefore, a "local" form of normalization is needed to establish a more robust model, as described in the next section.

3.1 Current Fraction Contour Analysis of the Power Grid

The smooth surface of the current profile shown in Figure 3 suggests that normalization is effective at eliminating current magnitude as a parameter in the characterization of the power grid's resistance. A similar argument can be made with regard to localizing defects, i.e., the magnitude of the defect current should not influence the position predicted by a localization algorithm. In previous work, we *normalized* the supply port I_{DDQ}s of neighboring V_{DD} ports by computing a current ratio, e.g. I_{00}/I_{01} [10]. Unfortunately, current ratios introduce non-linearity in the analytical model that complicates matters. An alternative that does not have this effect is to compute *current fractions*,

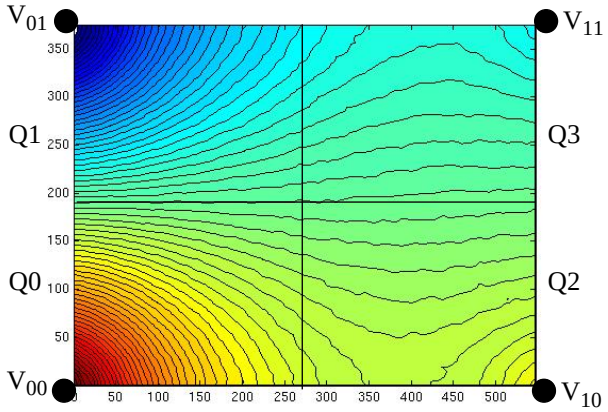

Figure 4. C3 V_{00}-V_{01} current fraction contours

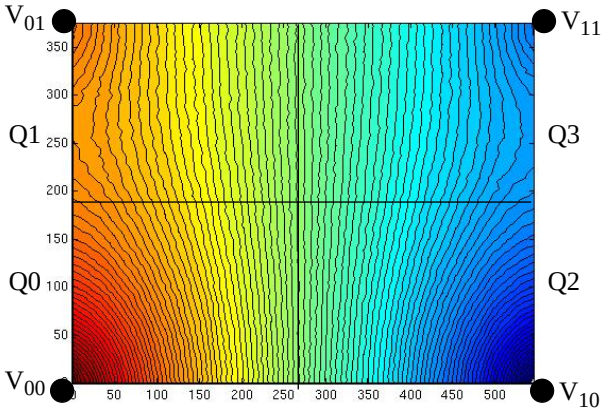

Figure 5. C3 V_{00}-V_{10} current fraction contours

e.g., $I_{00}/(I_{00} + I_{01})$.

Figures 4 and 5 plot the current fraction contours for chip C3 for V_{DD} port pairings V_{00}-V_{01} and V_{00}-V_{10}, respectively. The x-y plane represents the TC array. The lines within each plot delimit the iso-current fraction regions, i.e., positions in the layout that posses a similar current fraction. From these curves, it is clear that the analytical model that describes the entire surface is complicated and must incorporate parameters for every V_{DD} port in order to be comprehensive. However, the region surrounding the I_{DDQ} used in the numerator, I_0, is fairly "well-behaved". This region is identified as *Q0* in both of the figures. The contour curves in this region are best characterized as ellipses or hyperbolas. The model developed in the next section assumes the appropriate model is one based on hyperbolas. However, on-going work continues to investigate this hypothesis.

3.2 Hyperbola-based Model

Equation 1 and Figure 6 define and illustrate "horizontally-oriented" hyperbolas, as a model for the contours curves shown in Figure 5. The V_{DD} ports V_{00} and V_{10} map onto the graph at the foci, labeled F_1 and F_2. The space between the foci represents the 2-D space of the TC array.

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1 \qquad (1)$$

Figure 6 graphically portrays the definition of the *a* and *b* parameters used in Equation 1and defines an additional parameter, *c*, that is used to define the relationship among the sets or "families" of contour curves shown in Figures 4 and 5. A family of hyperbolas is defined as a set that share a common center and focus. The *h* and *k* parameters in Equation 1 define the center of the hyperbolas, which corresponds to the midpoint between the foci.

The *a* and *b* parameters of the hyperbolas need to be defined in terms of the current fractions. Fortunately, the nature of the contours allow an alternative formulation of the equation for a hyperbola, given by Equation 2. Here, b^2 is replaced with $(c^2 - a^2)$. Since *c* is fixed for all hyperbolas

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{c^2 - a^2} = 1 \qquad (2)$$

in the family as the distance between their center, (*h*, *k*), and the coordinates of the power ports, this makes *b* dependent on *a*. Therefore, only *a* needs to be defined in terms of current fractions.

From the diagram shown in Figure 6, *a* defines the point of intersection of the $I_{00}/(I_{00} + I_{10})$ hyperbola with the horizontal line defined between the center (*h*, *k*) = (279,0) and V_{00} (F_1 in the figure). Therefore, *a* varies from 0, at the center, to L/2 at V_{00}, where L is defined as the distance between V_{00} and V_{10} (558 for the TC array). The current ratios at points along this line increase from 0.5 at the center to the maximum current ratio, e.g. 0.60 for $I_{00}/(I_{00} + I_{10})$ in the array.

Figure 7 shows the current fractions computed from chip C3 as each of the shorting inverters are enabled in the TCs, one at a time, along the lines between V_{00}-V_{01} and V_{00}-V_{10}. The nearly linear shape of the curve reflects the uniform spacing of the contour curves along these lines in Figures 4 and 5. The linear nature of the curves allows a close approximation for *a* to be derived if the endpoints or bounds of the curves are known. Equation 3 can be used to derive *a* under these assumptions. Here, I_{ub} and I_{lb} represent the upper and lower bounds on the current fractions at the endpoints and *L* represents the distance between the V_{DD} ports.

The calibration circuits that we propose are positioned

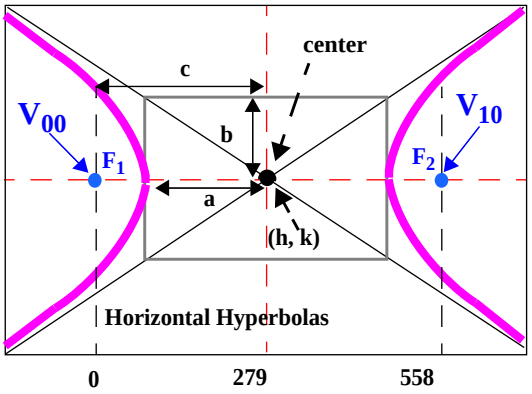

Figure 6. Definitions of the hyperbola parameters

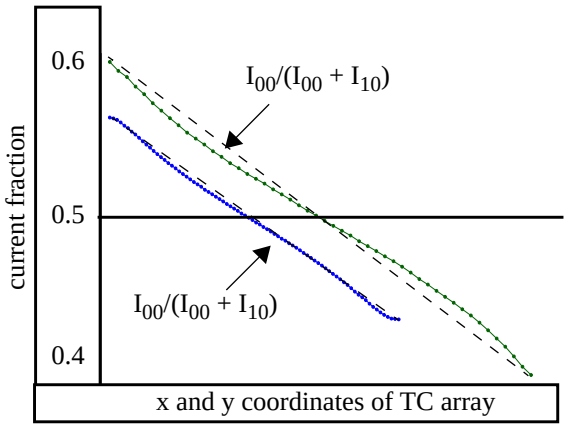

Figure 7. C3 current fractions along lines between V_{00}-V_{01} and V_{00}-V_{10}.

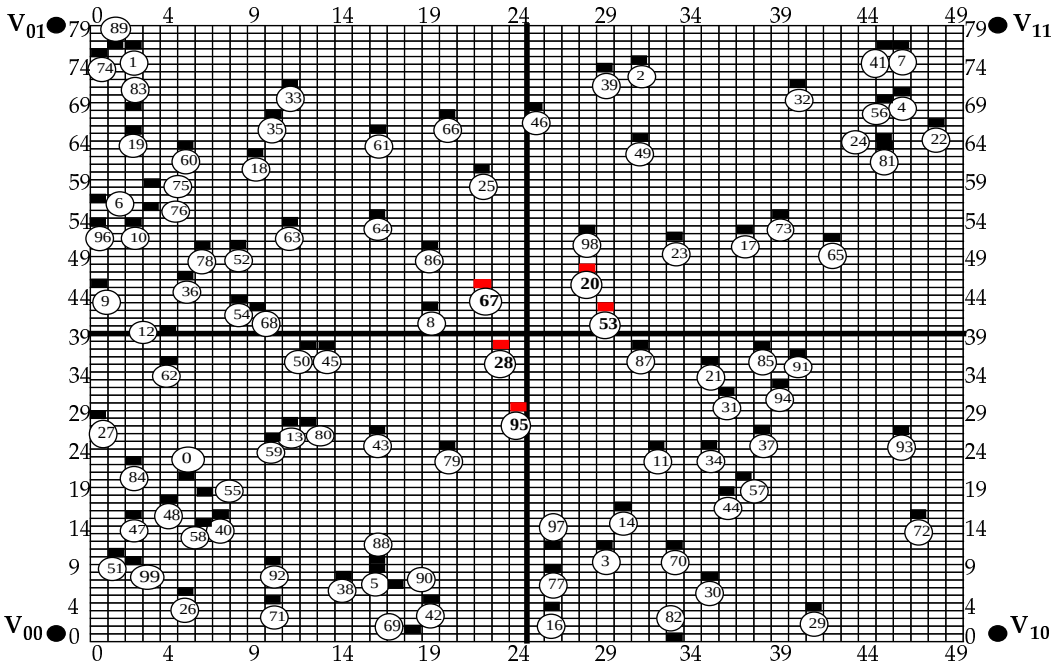

Figure 8. Positions of the 100 randomly selected TCs.

$$a = \frac{L}{2} - L\left(\frac{(I_{ub} - I_{00})}{(I_{ub} - I_{lb})}\right) \qquad (3)$$

underneath the V_{DD} ports and provide the information needed to compute the lower and upper bounds. With information about the bounds available, defect localization is performed by calculating the parameter *a* for two pairings of orthogonally positioned V_{DD} ports, e.g., V_{00}-V_{01} and V_{00}-V_{10}. The *a* parameters are used in Equation 2 to derive a vertical and horizontal hyperbola. The intersection of the two hyperbolae defines the predicted position of the defect in the layout.

4.0 I_{DDQ} Defect Localization Experiments

These experiments are designed to investigate the accuracy of our defect localization methodology for defect currents and leakages that vary over a wide range of values. These objectives are better met through the use of the defect emulation transistors and corresponding defect emulation wire because both the position and magnitude of the emulated defect current can be controlled.

4.1 Data Collection Procedure

Unlike the power grid characterization experiments which tested all 4,000 elements of the TC array, only a subset of 100 TCs are tested in these experiments. The set of randomly selected TCs within the 80x50 array are shown in

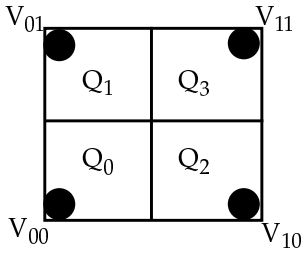

Figure 9. Quarters, Q_x, defined for the TC array.

Figure 8. The numbered positions are the TCs under investigation in these experiments.

For each of the 12 chips, a series of measurements were made for each of the TCs under different voltage configurations of the DESM, i.e., the source meter that drives the defect emulation wire. The first experiment for each chip is referred to the *defect-free* experiment. In this experiment, the state of all scan chain FFs is set to 1, which disables both the shorting inverters and the defect emulation transistors within all TCs in the array. The DESM is then swept across a sequence of voltages, from 0.9 V to 0.0 V in 50 millivolt intervals, for a total of 19 steps. At each DESM voltage, a set of 4 local and 4 global currents are measured. The same sequence of operations is performed with each of the defect emulation transistors enabled, one at a time.

4.2 Data Sets and Quarter Selection Algorithm

For each chip, the data collection procedure produces 1,919 data sets. 19 of these data sets represent data from the defect-free experiments and 1900 (19 * 100 emulated defects) represent data from the emulated defect experiments. However, the emulated defect experiment with the DESM voltage set to 0.9 V is not meaningful because there is no voltage drop across the defect emulation transistor. Therefore, only 18 of the 19 data sets are treated as emulated defects. With 12 chips, there is a total of 12*19 = 228 defect-free data sets and 12*1800 = 21,600 emulated defect data sets.

For each emulated defect, the analysis is performed by selecting a pair of orthogonally positioned supply ports. The supply ports selected are those surrounding the supply port with the largest measured I_{DDQ}, called the *primary port*. The supply port with the largest I_{DDQ} identifies the *quarter* in which the defect lies. Figure 9 displays the four quarters of the TC array. For example, if the I_{DDQ} in V_{01} is the largest in an emulated defect experiment, then V_{01} is selected as the primary port and V_{00} and V_{11} are the orthogonal neighbors. Figure 10 shows several examples of the method applied to predict emulated defects #0, #1, #4 and #27, as identified in Figure 8.

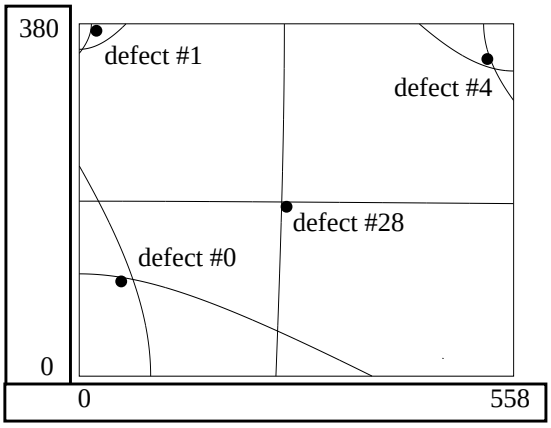

Figure 10. Example hyperbola based predictions of emulated defects.

4.3 Separating Defect Current and Background Leakage

The hyperbola-based localization algorithm maps the defect currents measured through each of the supply ports to an (x,y) position in the layout. Unfortunately, the currents actually measured also posses leakage current, which, if not subtracted, acts to “wash out” the current fractions described above and reduces the level of accuracy of the prediction algorithm.

One approach of eliminating leakage current from the measured values is through backtracking. Here, leakage data from defect-free chips is used to build a profile of the leakage relationship among the V_{DD} ports. Regression analysis is applied to derive a slope of the best fit line through the pair-wise plots of leakage data. The leakage component of the measured currents in a defective chip are obtained by selecting a V_{DD} port far from the defect site, i.e., a V_{DD} port whose defect current component is small or zero. The current measured from this port is then used in a backtracking operation to obtain the leakage currents in the V_{DD} ports surrounding the defect site.

It was not possible to apply this method directly because of the small size of the test chip under investigation in this work. In other words, it was not possible to select a V_{DD} port in the defect experiments that possessed only leakage current. Instead, an equation was derived (not shown) that approximated the leakage component in the V_{DD} port with the smallest defect current component and backtracking was applied to obtain the other leakage values.

5.0 Prediction Results

The results of applying the hyperbola-based defect location methods to chip C3 are shown in Figures 11 and 12[1]. The emulated defect number is given along the x-axis,

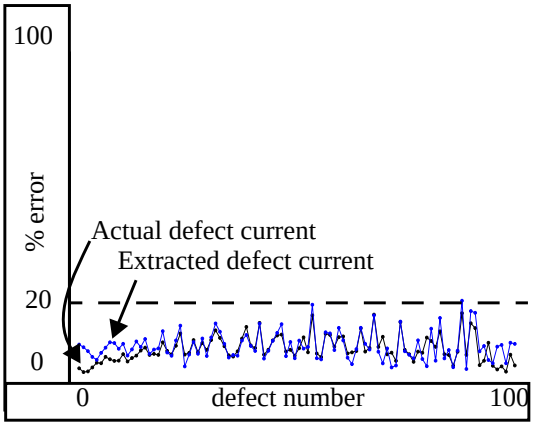

Figure 11. Chip C3: Linear equation prediction results.

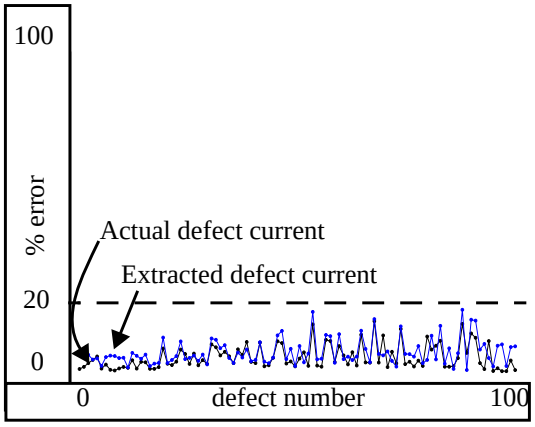

Figure 12. Chip C3: Curve derived prediction results.

as it is defined in Figure 8. The prediction error is plotted along the y-axis as a percentage. Prediction error is computed as the Euclidean distance between the actual defect location and the location given by the intersection of the hyperbolas. This error is converted into a percentage by dividing by the Euclidean distance across the diagonal of the TC array. The diagonal distance is 675 um. Therefore, a 1% error is equivalent to an actual error of 6.75 um. The percent error values given in Figures 11 and 12 is the average error computed for each emulated defect across all DESM voltages.

Figure 11 gives the prediction errors if the hyperbolas are derived using the linear model for parameter *a*, as defined in Section 3.2. The curve labeled *actual defect current* gives the best achievable result under the linear model. Here, the predictions are derived using the measured I_{DDQ}s with leakage subtracted. It is possible to derive the exact defect current in our experiments because the scan chain allows control over whether the emulated defect is on or off. This is clearly difficult or impossible to do in a real design. Therefore, in most cases, it will be necessarily to separate defect current from leakage current in the measured values. Section 4.3 describes a method that allows defect current to be *estimated* from the measured values. The curve labeled *extracted defect current* gives the results under these conditions. Figure 12 gives the prediction errors in an analogous way under the condition that the *a* parameters for the hyperbolas are derived using the curve data, such as those shown in Figure 7.

The figures include a horizontal dashed line that identifies 20% error. The prediction errors for all but one of the emulated defects are below this line. Moreover, in general, the prediction errors using the curve data are smaller than those obtained under the linear model. This is expected since the curve data gives a more precise value for the hyperbola parameter *a* than the linear model. It is interesting to note that the prediction error using the *extracted defect current* is sometimes smaller than the error obtained using the *actual defect current*. This indicates that the hyperbola model used here is capable of only providing an estimate of the emulated defect's location. A more complete model is under investigation.

1.The other chips produce similar results.

6.0 Conclusion

A hyperbola-based defect localization method is applied to the data collected from a set of chips fabricated in a 65 nm technology. The test chips incorporate an array of test structures that permit the controlled insertion of a shorting defect. The method uses multiple supply port I_{DDQ}s to triangulate the physical location of the defect in the layout. The results of the analysis indicate that good diagnostic resolution is achievable in most cases, with average prediction errors less than 15% or 100 um.

7.0 Acknowledgments

We thank Dr. Sani Nassif, Chandler McDowell and Dr. Anne Gattiker of IBM Austin Research Laboratory for their support of this research.

7.1 References

[1] ITRS (http://public.itrs.net/)
[2] Gong Yiming, S. Chakravarty, "Using Fault Sampling to Compute I_{DDQ} Diagnostic Test Sets", VTS, pp. 74-79, 1997.
[3] S. Chakravarty, M. Liu, "I_{DDQ} Measurement Based Diagnosis of Bridging Faults", JETTA, Vol. 3, pp. 377-385, 1992.
[4] I. Pomeranz and S. M. Reddy. "On the Generation of Small Dictionaries for Fault Location", ICCAD, pp. 272-279, 1992.
[5] Aitken, R.C., "A Comparison of Defect Models for Fault Location with Iddq Measurements", ITC, pp. 1051-1060, 1993.
[6] Y. Gong and on S. Chakravarty, "On adaptive Diagnostic Test Generation, ICCAD, pp. 181-184, 1995.
[7] P. Nigh, D. Forlenza, F. Motika, "Application and Analysis

of IDDQ Diagnostic Software", ITC, pp. 319-327, 1997.

[8] Christopher L. Henderson and Jerry M.Soden, "Signature Analysis for IC Diagnosis and Failure Analysis", ITC, pp. 310-318, 1997.

[9] C. Thibeault, L. Boisvert, "Diagnosis method based on delta IDDQ probabilistic signatures: Experimental results", ITC, pp. 1019-1026, 1998.

[10] C. Patel, E. Staroswiecki, S. Pawar, D. Acharyya and J. Plusquellic. "Defect Diagnosis using a Current Ratio based Quiescent Signal Analysis Model for Commercial Power Grids", *JETTA*, Volume 19, Issue 6, pp. 611-623, Dec 2003.

[11] S. Nassif and J. Kozhaya, "Fast Power Grid Simulation", DAC, pp. 156-161, 2000.

[12] D. Acharyya and J. Plusquellic, "Hardware Results Demonstrating Defect Localization using Power Supply Signal Measurements", ISTFA, pp. 58-66, Nov. 2004.

[13] D. Acharyya and J. Plusquellic, "Hardware Results Demonstrating Defect Detection Using Power Supply Signal Measurements", *VTS*, 2005.