Localizing Faults in Digital Chips using Steady-State Current Measurements

Jim Plusquellic and Dhananjay Phatak
Department of CSEE, University of Maryland, Baltimore County

Abstract

Quiescent Signal Analysis (QSA) is a novel electrical-test-based diagnostic technique that uses I_{DDQ} measurements made at multiple chip supply pads as a means of locating shorting defects in the layout. The use of multiple supply pads reduces the adverse effects of leakage current by scaling the total leakage current over multiple measurements. In previous work, a resistance model for QSA was developed and demonstrated on a small circuit. In this paper, the weaknesses of the original QSA model are identified, in the context of a production power grid (PPG) and probe card model, and a new model is described. The new QSA algorithm is developed from the analysis of I_{DDQ} contour plots. A "family" of hyperbola curves is shown to be a good fit to the contour curves. The parameters to the hyperbola equations are derived with the help of inserted calibration transistors. Simulation experiments are used to demonstrate the prediction accuracy of the method on a PPG.

1.0 Introduction

I_{DDQ} has been a main-stream supplemental testing method for defect detection for more than a decade with many companies. With the advent of deep submicron technologies, the use of single-threshold I_{DDQ} technique results in unacceptable yield loss. Setting an absolute pass/fail threshold for I_{DDQ} testing has become increasingly difficult due to the increasing subthreshold leakage currents [1]. Current signatures [2], delta-I_{DDQ} [3] and ratio-I_{DDQ} [4] have been proposed as a means for calibrating for these high subthreshold leakages. These techniques rely on a self-relative or differential analysis, in which the average I_{DDQ} of each device is factored into the pass/fail threshold. However, these proposed forms of calibration are expected to become less effective over successive technology generations.

An alternative calibration strategy that may have better scaling properties is to distribute the total leakage current across a set of measurements. This is accomplished by introducing probing hardware that allows the measurement of I_{DDQ} at each of the supply ports. The method proposed in this work, called Quiescent Signal Analysis (QSA), is designed to exploit this type of leakage calibration for defect detection and as a means of providing information about the defect's location in the layout [5][6]. This latter diagnostic attribute of QSA may provide an alternative to image-based physical failure analysis procedures that are challenged by the increasing number of metal layers and flip chip technology.

This work is supported by a Faculty Partnership Award from IBM's Austin Center for Advanced Studies (ACAS) Program and by an NSF grant, award number 0098300.

A resistance-based diagnostic model for QSA was developed in previous works and simulation experiments were used to demonstrate the diagnostic capabilities of the QSA method on a small circuit [5][6]. In this paper, several weaknesses of the resistance-based model are uncovered from simulations of a production power grid (PPG). A current-ratio-based model is proposed and demonstrated to improve on defect localization accuracy of the original method [7]. The new method requires the insertion of calibration transistors (CT), one under each of the supply pads in the design, that permit the shorting of the power and ground supply rails at points close to the substrate. The state of the CTs are controlled by scan chain flip-flops. The I_{DDQ}s obtained when one of the CTs is turned on are used to calibrate the I_{DDQ}s measured under a failing I_{DDQ} pattern. The calibration technique is shown to address several weaknesses of the previous model including non-zero probe card resistance and irregular supply grid topologies. Current ratios, as opposed to absolute currents, are proposed as a means of reducing the dependence of the localization algorithm on the value of the defect current. SPICE simulation experiments demonstrate that the maximum prediction error is 650 units in a 30,000 by 30,000 unit area.

It is not possible to evaluate the QSA algorithm on the entire 80,000 by 80,000 unit area of the PPG using SPICE due to the large size of the R model. Instead, a specialized power grid simulation engine called ALSIM is used [8]. The anomalies in the grid's structure in this larger area increase the maximum prediction error to 1,340 units. Although the prediction accuracy is good for most cases, an alternative "lookup table" approach (in contrast to the hyperbola-based approach) is likely to be more accurate for irregular grid regions or configurations. The enhanced simulation capabilities of ALSIM enable this strategy, alone or in combination with the hyperbola-based approach described in this paper.

The remainder of this paper is organized as follows. Section 2.0 describes related work. Section 3.0 gives a brief description of the original resistance-based QSA technique, identifies its weaknesses and describes the basis of a new

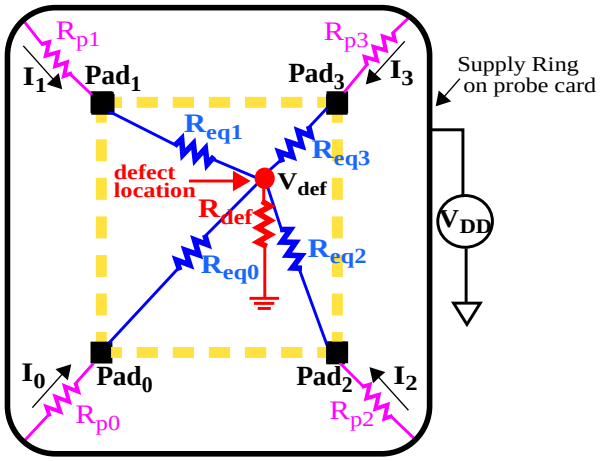

Figure 1. Equivalent resistance model of the power grid with a shorting defect.

model. Section 4.0 presents the details of the current-ratio-based QSA model. Section 5.0 gives experimental results. Section 6.0 gives our conclusions and areas of future research.

2.0 Background

Several diagnostic methods have been proposed based on I_{DDQ} measurements. In general, these methods produce a list of candidate faults from a set of observed test failures using a fault dictionary. The likelihood of each candidate fault can be determined by several statistical algorithms. For example, signature analysis uses the Dempster-Shafer theory, which is based on Bayesian statistics of subjective probability [9]. Delta-I_{DDQ} makes use of the concepts of differential current probabilistic signatures and maximum likelihood estimation [10]. Although these methods are designed to improve the selection of fault candidates, in many cases, they are unable to generate a single candidate. Other difficulties of these methods include the effort involved in building the fault dictionary and the time required to generate the fault candidates from the large fault dictionary using tester data.

The QSA procedure can help prune the candidate list produced by I_{DDQ} and other voltage based diagnostic algorithms. The physical layout information generated by our method can be used with information that maps the logical faults in the candidate lists to positions in the layout. In addition, it may be possible to use the (*x*,*y*) location information provided by QSA as a means of reducing the search space for likely candidates in the original fault dictionary procedure. This can reduce the processing time and space requirements significantly.

3.0 QSA Models

QSA analyzes a set of I_{DDQ} measurements, each obtained from individual supply pads, to predict the location of a shorting defect. The resistive element of the power grid causes the current drawn by the defect to be non-uniformly distributed to each of the supply pads. In particular, the defect draws the largest fraction of its current from supply pads topologically "nearby". The same is true of the leakage currents. However, only the leakage currents in the vicinity of the defect contribute to the measured current in these pads. The smaller background leakage component improves the accuracy of the defect current measurement. As described in previous works, QSA also proposes the use of regression analysis as a means of eliminating the remaining leakage component from the measured values [5][6].

3.1 The Resistance-based QSA Model

The fraction of the defect current provided by each of the pads in the region of the defect is proportional to the equivalent resistance between the defect site and each of the pads. The differences in these values can be used to localize the defect using a method based on triangulation. For example, Figure 1 shows a shorting defect in an equivalent resistance model of a simple power grid. Here, R_{eq0} through R_{eq3} represent the equivalent resistances between each of the supply pads, Pad_i, and the defect site shown in the center of the figure. The following set of equations describe the relationship between the power supply branch currents, I_0 through I_3, and V_{def}, the voltage at the defect site.

$$I_i \times (R_{eqi} + R_{pi}) = V_{DD} - V_{def} \quad \text{for i = 0,1,2,3} \quad (1)$$

In Eq. 1, the I are the measured I_{DDQ}s. The R_p represent the probe card's resistances, which we assume are very small with respect to the R_{eq} and can be ignored (this assumption is addressed below). This leaves the R_{eq} and V_{def} as unknowns. Without additional information, it is not possible to solve these equations since there are 4 equations and 5 unknowns. However, for the purpose of diagnosis, only the relationships between the R_{eq} are needed. Relative equivalent resistances can be computed with respect to a reference equivalent resistance, R_{eqj}, as given by Eq. 2.

$$I_i \times (R_{eqi} + R_{pi}) = I_j \times (R_{eqj} + R_{pj})$$

solving for R_{eqi} in terms of R_{eqj} gives

$$R_{eqi} = \frac{I_j}{I_i} \times R_{eqj} + \frac{I_j}{I_i} \times R_{pj} - R_{pi} \quad (2)$$

with $i \neq j$

Under the condition that $R_p << R_{eq}$ (otherwise the model shown if Figure 1 is not complete), it is possible to obtain an accurate prediction of the defect's location by solving the circle expressions given in Eq. 3 for a common point of intersection given by *x* and *y*. The parameters h_i and k_i represent the x and y coordinates of the center of the

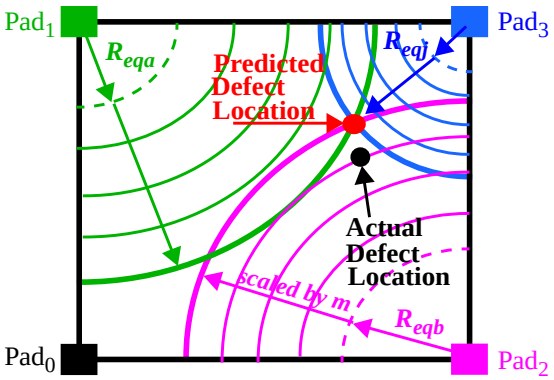

Figure 2. Triangulation under resistance model.

$$\begin{aligned} m \times R_{eqj} &= \sqrt{(x-h_j)^2 + (y-k_j)^2} \\ m \times R_{eqa} &= \sqrt{(x-h_a)^2 + (y-k_a)^2} \\ m \times R_{eqb} &= \sqrt{(x-h_b)^2 + (y-k_b)^2} \end{aligned} \qquad (3)$$

ith circle. The three circle equations are related to corresponding equations from the set described by Eq. 2 through the R_{eq}. Here, R_{eqj} is assumed to be 1.0 and R_{eqa} and R_{eqb} are computed from Eq. 2 using the I_{DDQ} measurements. Parameter m is used to map the resistances given on the left in Eq. 3 to distances in the layout.

The choice of the supply pads to be used in the triangulation procedure is based on two criteria. First, the supply pads are sorted according to the magnitude of their corresponding I_{DDQ}. The supply pad, j, with the largest I_{DDQ} is selected followed by two orthogonally adjacent supply pads, a and b, to pad j sourcing the next two largest values. Note that this model is based on two simplifying assumptions: a uniform resistance-to-distance mapping function and negligible values for R_p. A uniform resistance-to-distance mapping function is used to describe power grids in which the equivalent resistance and Euclidean distance between any two points on the grid are proportional.

An example application of this triangulation-based method is shown in Figure 2. Three dotted circles are shown whose centers are defined by the positions of the Pad_1, Pad_2 and Pad_3. The radii are labeled with the appropriate R_{eq} values as given in Eq. 3. For example, Pad_3 defines the center of the circle with smallest radius, i.e., it is the supply pad with the largest I_{DDQ}. Its radius is labeled with R_{eqj} in the figure. The initial radii of the three circles are then multiplied by a common factor, m, to a common point of intersection. This point is labeled as "Predicted Defect Location" in the figure to contrast it with the "Actual Defect Location".

3.2 Weaknesses of the Resistance-based Model

Unfortunately, the assumptions of the resistance-based model are not valid in many situations. Here, it is assumed that the R_p are small relative to the R_{eq}. Under this assumption, the measured I_{DDQ}s are related to the R_{eq} as given by Eq. 4 (derived from Eq. 2). Therefore, the resistance-based

$$R_{eqi} = \frac{I_j}{I_i} \times R_{eqj} + \underbrace{\frac{I_j}{I_i} \times R_{pj} - R_{pi}}$$

If these terms are negligible then

$$R_{eqi} \cong \frac{I_j}{I_i} \times R_{eqj} \quad \text{or} \quad \frac{R_{eqi}}{R_{eqj}} \cong \frac{I_j}{I_i} \qquad (4)$$

QSA model assumes that the current ratios are inversely proportional to the resistance ratios. If the values of R_p are similar to or larger than the R_{eq}, then the relationship given by Eq. 4 is weakened and the accuracy of the triangulation approach is correspondingly reduced.

In the next section, we present a more complete equivalent resistance model of the CUT that better represents an actual probe card model in which the R_p are significant. The new model requires additional information in order to solve for unknowns such as the R_{eq} and R_p. A new QSA method is proposed that obtains this information from calibration transistors measurements. However, it should be noted at this point that large values of R_p will adversely affect the precision required in the measurement of the I_{DDQ}s under any proposed strategy. This follows from a numerical analysis of Eq. 2, that shows the convergence of all current ratios to the ratios defined by the R_p as the magnitude of the R_p are increased to and above the R_{eq}.

Another weakness of the resistance-based QSA model is with regard to the uniform resistance-to-distance mapping function. Most supply topologies are poorly modeled as uniform. In previous work, we proposed a mapping function based on resistance contours to deal with complicated irregular topologies [5]. In this work, we propose a second strategy based on the use of a current ratio lookup-table. Both techniques require resistance and current profiles of the grid to be derived in advance through simulations, and should be avoided, if possible, in cases involving more regular topologies.

The topology of the PPG under investigation in this work fits between the totally regular and totally irregular extremes. The mapping function for it is not strictly uniform but, because the physical structure of the grid is regular in many places, it is possible to model the resistance per unit distance between each pairing of supply pads using a constant. The new hyperbola-based QSA method described in this paper is able to calibrate for this type of power grid resistance-to-distance profile using measured data only. Therefore, it provides a simpler alternative to a

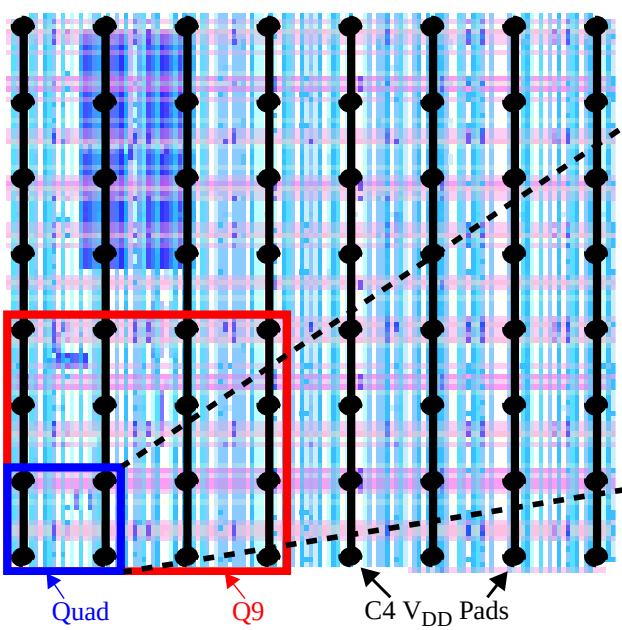

Figure 3. Layout of the PPG.

lookup-table approach.

3.3 The PPG's Physical Characteristics

Figure 3 shows the 80,000 by 80,000 unit layout of the PPG. The PPG interfaces to a set of external power supplies through an area array of V_{DD} and GND C4 pads. A C4 pad is a solder bump for an area array I/O scheme. The PPG has 64 V_{DD} C4s and 210 GND C4s (not shown in Figure 3). The 64 V_{DD} C4s divide the PPG into 49 different regions called Quads. ALSIM simulations experiments were run on the entire PPG. However, due to space and time constraints, it was not possible to run SPICE simulations on the entire PPG. Rather, a portion of the PPG consisting of 9 quads was simulated using SPICE. This portion consists of the lower left 9 Quads as shown in Figure 3, and is subsequently referred to as the Q9. The Q9 occupies a 30,000 by 30,000 unit area.

In order to derive an electrical model of the PPG, we focused our analysis on the portion shown in the lower left of Figure 3 identified as the Quad. Figure 4(a) expands on this view by showing a more detailed diagram of this 10,000 by 10,000 unit region. This is again expanded in Figure 4(b) which shows a stacked four metal layer configuration, with m1 and m3 running vertically and m2 and m4 running horizontally. The C4s are connected to wide runners of vertical m5, shown in the top portion of Figure 4(a), that are, in turn, connected to the m1-m4 grid. In each layer of metal, the V_{DD} and GND rails alternate. In the vertical direction, each m1 rail is separated by a distance of 432 units. The alternating vertical V_{DD} and GND rails are connected together using alternating horizontal metal runners. Stacked contacts are placed at the appropriate cross-

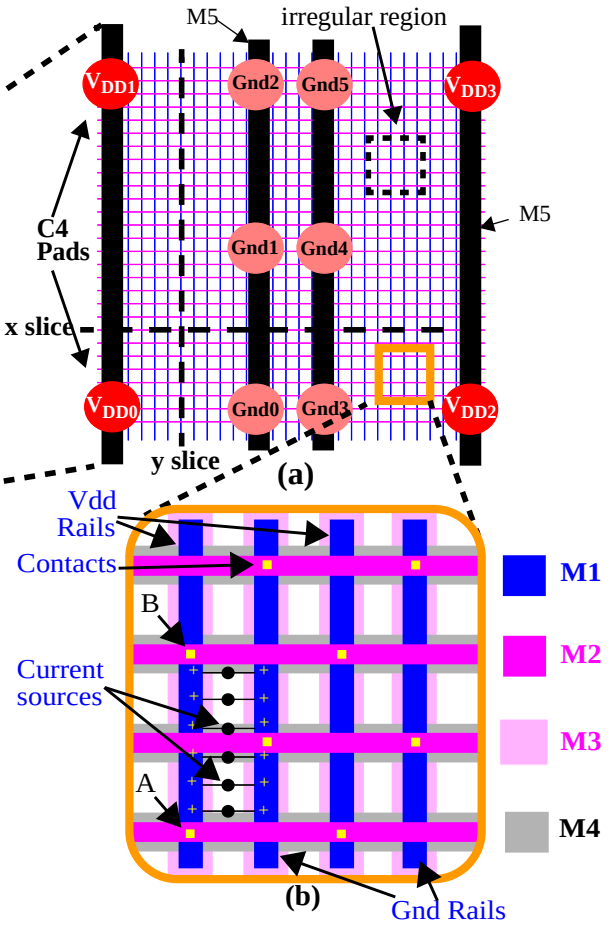

Figure 4. Layout details of the PPG.

ings of the horizontal and vertical rails. The grid is fairly regular except in the region labeled "irregular region" in the upper right corner of Figure 4(a). The m1 in this region of the layout varies from the regular pattern shown in Figure 4(b).

The R model of the Quad was obtained from an extraction script which uses process parameters from the TSMC's 0.25µm process [11]. 1Ω resistances were inserted between the power supplies and the R model of the grid to model the tester power supply(s) and probe card contact resistances to the chip. (Although our simulation model uses 1Ω for all probe card resistances, the analytical model that we derive below accommodates a more realistic probe card model in which probe card resistance is different from one pad to another.) The combined resistance network contains approximately 27,000 resistors.

3.4 The Quad's Electrical Characteristics

Figure 4(b) also shows a set of current sources that were inserted individually in a sequence of simulations as a means of evaluating the electrical behavior of the resistance model at the V_{DD} C4s. The current sources, which model the presence of a shorting defect, were placed at regular

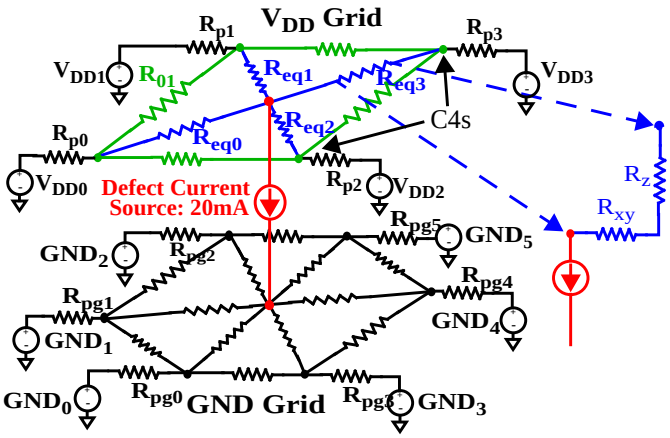

Figure 5. Equivalent resistance model of the Quad.

intervals between m1 V_{DD} and GND runners. An equivalent resistance model of the Quad is shown in Figure 5 with one of the current sources inserted. The four grid equivalent resistances, R_{eq}, in the upper center portion of the figure are the source of resistance variation as seen from the power supplies, as the current source is moved in the layout. The strength of the correspondence of these resistances to the position of the defect determines the accuracy of the triangulation procedure used in QSA. It is therefore prudent to evaluate this relationship for the Quad.

There are several significant differences between this model and the model shown in Figure 1. First, under the assumption that the values of the R_p are non-zero, the grid resistances between the C4s, e.g. R_{01} shown on the top left of Figure 5, are needed in any complete equivalent resistance expression such as that given by Eq. 2. Second, the R_{eq} are actually three dimensional in nature and can be modeled as R_z and R_{xy} as shown on the right side of Figure 5. R_z adversely impacts the accuracy of the triangulation procedure for the same reasons given earlier for R_p.

A third glitch in our original resistance-based model is related to the resistance profile that characterizes the PPG under investigation in this research. Our analysis reveals that the variation in equivalent resistance over small vertical intervals of the Quad, e.g., along the interval between two contact points in m1, is on order with the variation across the entire Quad. For example, the segment length given between points A and B in Figure 4(b) is approximately 630 units. Using the m1 resistance parameter for TSMC's 0.25µm process yields a value of 5.6Ω. Therefore, in m1 alone, the resistance varies from 0Ω at the contact to 5.6 || 5.6 = 2.8Ω in the center. On the other hand, the average resistance from the center of the Quad (shown in Figure 4(a)) to any of the V_{DD}s (distance of ~7,000 units) is less than 6Ω. The increasing width of the metal runners from m1 to m5 is responsible for these resistance to distance anomalies.

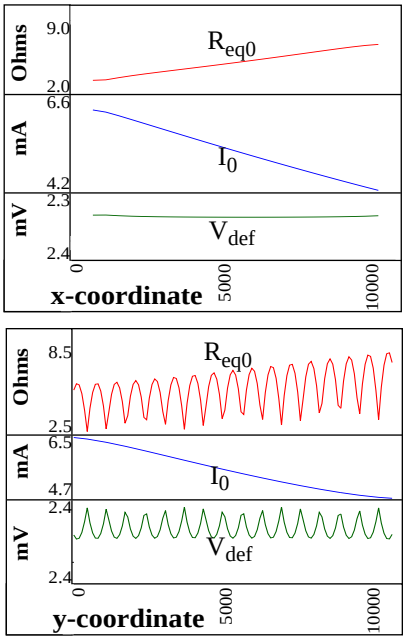

Figure 6. Network variable plots for sources along x-slice (top) and y-slice (bottom) lines of Figure 4(a).

In order to gain insight into other alternative diagnostic strategies, we first derived the profiles of the network variables including R_{eq}, V_{def} (the voltage at the defect site) and the I_{DDQ}s at the V_{DD} C4s. The profiles were derived from the results of 2,600 SPICE simulation experiments of the Quad. In each simulation, a 20mA current source was placed between m1 V_{DD} and GND rails at different locations in the layout. Figure 6 shows the curves for R_{eq0} and I_0 (at $C4_0$) and V_{def} (the current source's terminal voltage at the connection point on the m1 V_{DD} rail) for a set of simulations run along the lines identified as x-slice and y-slice in Figure 4(a). The R_{eq0} values were computed using Eq. 5.

$$R_{eq0} = \frac{(V_{DD} - V_{def})}{I_0} \qquad (5)$$

V_{DD} = 2.5V
V_{def} = voltage at the defect site
I_0 = current through V_{DD0}

It is clear from these graphs that the variations in R_{eq0} and V_{def} along the *y* dimension are significantly larger than those along the *x* dimension. In contrast, the currents are well behaved along either dimension. The staggered arrangement of V_{DD} and GND grids, as shown in Figure 4(b), causes the *total* resistance between V_{DD} and GND to change slowly across the grid, through the exchange of nearly equal resistance fragments between the V_{DD} and GND grids. This keeps the currents well behaved while the resistances to, and voltages at, the defect site oscillate inversely with each other.

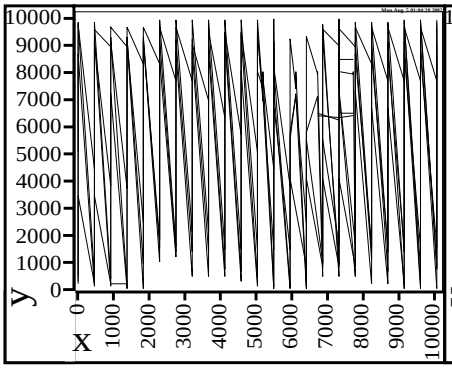

Figure 7. R_{eq0} contours of Quad.

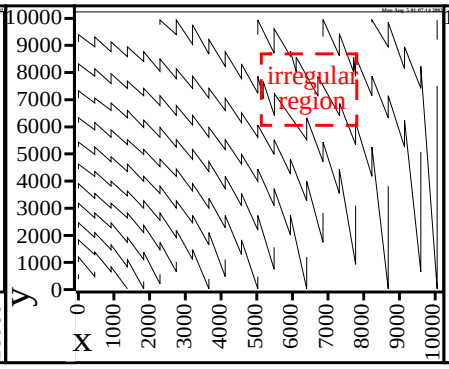

Figure 8. I_0 contours of Quad.

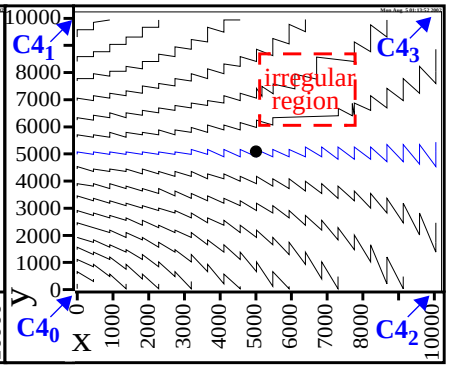

Figure 9. I_0/I_1 contours of Quad.

3.5 Contour Profiles of the Quad

Another useful view of the behavior of these network variables is through contour plots. A line within a contour plot is defined as the parameter values over which the value of the function remains constant. Contours are particularly useful when data is to be fit to a function. Figures 7 and 8 show the equivalent resistance and current contours of the Quad for V_{DD0} (only every 3rd contour curve is shown.) The x and y axes correspond to the (*x*,*y*) coordinates of the Quad as shown in Figure 4(a). The jagged nature of the curves as shown in Figure 8 models a band whose width is defined by the vertical line segments in the curves. It is clear that the equivalent resistance contour plot is difficult to make use of. The same is true of the V_{def} contour plot (not shown). In contrast, the current contours are elliptical in shape, (except for a region in the upper right hand corner, identified as "irregular region" given earlier in reference to Figure 4(a)). Similar patterns are present in the current contour plots of the other V_{DD}s.

Therefore, a diagnostic method based on currents is likely to yield the best results. However, unlike equivalent resistance, the disadvantage of using the currents directly is the dependency that is created between the contours and the magnitude of the defect's shorting current. Current ratios are an alternative that reduce this dependency since different values of defect current are reflected as the same ratio in the C4 I_{DDQ}s.

The contour plot for I_0/I_1 is shown in Figure 9. Like the I_0 contour plot, the contour lines are well behaved. However, the elliptical curves characterizing the I_0 plot now appear as hyperbola curves, particularly in the region to the left of x=5000. The set or "family" of hyperbolas is centered at the midpoint between the position of $C4_0$ (lower left) and $C4_1$ (upper left). The contour curve that passes through this midpoint (y=5000) on the y axis is nearly linear along a line to the center of the Quad (shown by the 'dot' in the center of the figure). This curve defines the points in the layout that are expected to produce an I_0/I_1 current ratio closely approximated by 1.0. The I_0/I_1 increase to a maximum in the lower left corner. The maximum I_0/I_1 current ratio is largely determined by the R_z component of resistance at $C4_0$ and R_{p0}. As an example, the I_0/I_1 maximum for the Quad is 1.55 and the I_0/I_2 maximum is 1.84. These maximum current ratios can be determined experimentally using a simple test circuit. We describe this test circuit and its other benefits after we derive the analytical model for the new QSA procedure.

4.0 The Current Ratio Model for QSA

The density of the contour curves in the lower left quarter of Figure 9, i.e. the region with *x* and *y* coordinates less than 5000, is higher than the density in other regions of the Quad. For example, the number of contour curves below the y=5000 is 10 while the number above this point is 6. Therefore, the I_0/I_1 and I_0/I_2 current ratios are expected to provide the best resolution for defects that occur in this region. Under the assumption that the C4s with largest I_{DDQ}s are closest to the defect site, it is straightforward to identify the relevant region and to compute the appropriate current ratios from the measured data. (This assumption is later removed.) The more challenging problem is to determine how to use these ratios to identify the location of the defect in the selected region.

The most straightforward method is to use the measured ratios to select two contour curves. For example, Figures 10 and 11 show the I_0/I_1 and I_0/I_2 contour curves obtained from simulations performed on the lower left quarter of the Quad (only every other curve is shown). The point of intersection of two curves, one from each figure, identifies the position of the defect in the layout. This is the general idea behind the lookup-table method referred to above. The drawback of this method is the large number of simulations that are needed (one for each candidate position in the layout) to build the table. Power grid simulators such as ALSIM make this practical and we expect this approach to be useful for irregular grid topologies. However, a simpler method is possible in many situations.

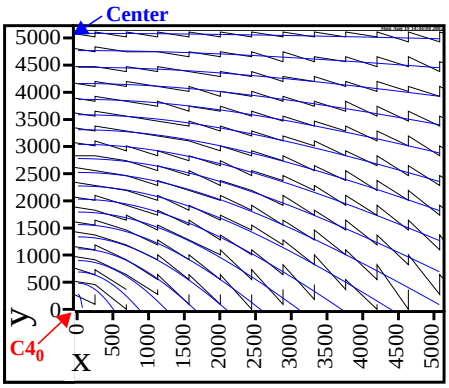

Figure 10. I_0/I_1 contours (jagged) and hyperbolas (smooth curves) for lower left quadrant of Quad.

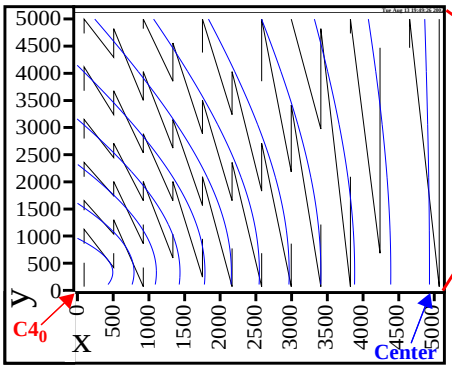

Figure 11. I_0/I_2 contours (jagged) and hyperbolas (smooth curves) for lower left quadrant of Quad.

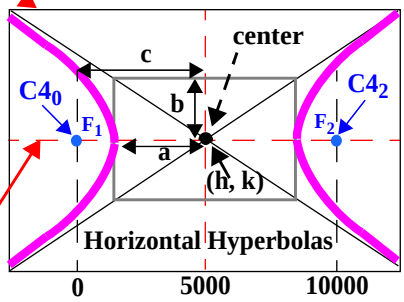

Figure 12. Definitions of the hyperbola parameters

An alternative strategy is to derive a function that approximates the contour curves using the measured quantities, i.e. the current ratios, as parameters. As noted above, the current ratio contour curves are similar in shape to hyperbolas. Figures 10 and 11 show a set of curves derived from hyperbolas superimposed on the contour curves for illustration. In order to realize this mapping, it is necessary to derive expressions for the hyperbola parameters. Eq. 6 and Figure 12 define and illustrate "horizontally-oriented" hyperbolas, such as those shown in Figure 11. The arrows on the right of Figure 11 illustrate the region in which these curves are represented in Figure 12.

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1 \qquad (6)$$

Figure 12 portrays the role of the *a* and *b* parameters in a graph and defines an additional parameter, *c*, that is used to define the relationship among the sets or "families" of hyperbolas in Figures 10 and 11. A family of hyperbolas is defined as a set that share a common center and focus. The *h* and *k* parameters in Eq. 6 define the center of the hyperbolas. The centers of the hyperbola curves shown in Figures 10 and 11 are identified at (*h*,*k*)=(0,5000) and (5000,0), respectively, and represent the midpoint between the foci. The foci of the hyperbolas are given by F_1 and F_2 in Figure 12. These points represent the (*x*,*y*) coordinates of the C4 V_{DD}s.

The *a* and *b* parameters of the hyperbolas need to be defined in terms of the current ratios. Fortunately, the nature of the contours defined by the grid allow an alternative formulation of Eq. 6 as given by Eq. 7. Here, b^2 is replaced with (c^2 - a^2). Since *c* is fixed for all hyperbolas in

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{c^2-a^2} = 1 \qquad (7)$$

the family as the distance between their center, (*h*,*k*), and the coordinates of the C4 supply pad, this makes *b* dependent on *a*. Therefore, only *a* needs to be defined.

From the diagram shown in Figure 12, *a* defines the point of intersection of the I_0/I_2 hyperbola with the horizontal line defined between the center (*h*,*k*) = (5000,0) and $C4_0$ (F_1 in the figure). Therefore, *a* varies from 0, at the center, to L/2 at $C4_0$, where L is defined as the distance between $C4_0$ and $C4_2$ (10,000 for the Quad). The current ratios at points along this line increase from 1.0 at the center to the maximum current ratio, e.g. 1.84 for I_0/I_2 in the Quad. If this maximum current ratio is known, then the function that defines *a* can be derived from the lumped R model shown in Figure 13 as follows.

Eqs. 8 and 9 give the expressions for the current ratio β_{02} and *a* without proof (please refer to [12] for proofs and other details of the analytical model presented in this paper). R_{eqT} (total resistance) is equal to the sum of the

$$\beta_{02} = \frac{I_0}{I_2} = \frac{R_{eqT} \times (L-m) + L \times R_{p2}}{m \times R_{eqT} + L \times R_{p0}}$$
$$\text{and} \quad a = \frac{L}{2} - m \qquad (8)$$

Substituting and solving for *a* yields

$$a = \frac{L}{2} - \frac{L}{R_{eqT}}\left(\frac{R_{p2} - (\beta_{02} \times R_{p1}) + R_{eqT}}{(1+\beta_{02})}\right) \qquad (9)$$

equivalent resistances, i.e. R_{eq0} + R_{eq2}, between the defect site and each of the two C4s. R_{p0} and R_{p2} are the probe card resistances at $C4_0$ and $C4_2$, respectively.

4.1 Calibration Transistors

As pointed out earlier, R_{eq0} and R_{eq2}, and therefore R_{eqT}, cannot be obtained in the defective chip. However, under the special case where the defect shown in Figure 13 is positioned on a line between $C4_0$ and $C4_2$, we can obtain a close approximation of R_{eqT} experimentally. This is accomplished by inserting a calibration transistor (CT_0)

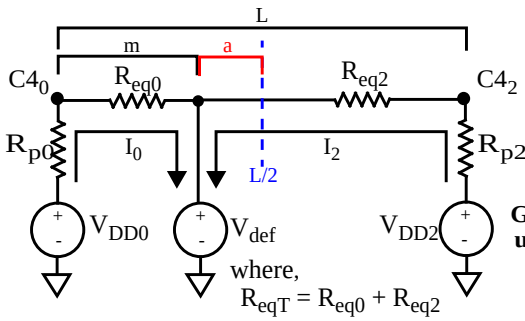

Figure 13. Lumped R model for the hyperbola *a* parameter.

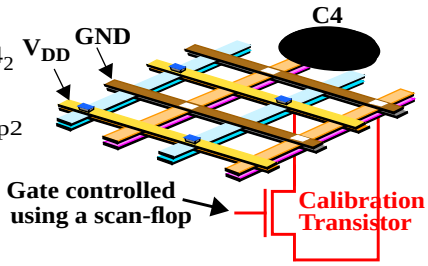

Figure 14. Calibration Transistor and controlling scan-chain FF.

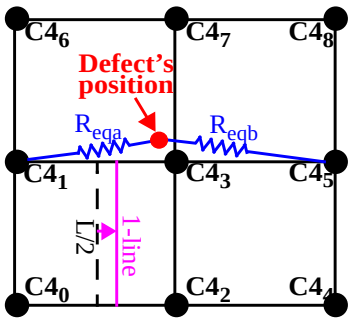

Figure 15. Anomalies in complex grids.

under $C4_0$, as shown in Figure 14. The source and drain of the CT_0 connect to V_{DD} and GND in m1 and provide a way to conditionally short these nodes together. By positioning the CT_0 directly under $C4_0$ (at the lowest resistance position from m1 to $C4_0$), the maximum current ratio, $\beta_{02(CT0)} = I_{0(CT0)}/I_{2(CT0)}$, can be obtained. This is accomplished by placing the chip into a state that does not provoke the defect and turning on CT_0 using the scan chain flip-flop driving its gate.

The measured values of $I_{0(CT0)}$ and $I_{2(CT0)}$ resolve several issues related to the application of this technique. First, $\beta_{02(CT0)}$ allows Eq. 8 to be solved under the boundary condition m=0. If the same process is repeated using a calibration transistor CT_2, positioned under $C4_2$, then Eq. 8 can be solved under a second boundary condition, m=L, using $\beta_{02(CT2)}$. With three equations, R_{eqT}, R_{p0}, and R_{p2} in Eq. 8 can now be eliminated, allowing a to be expressed as a function of the measured current ratios, β_{02}, $\beta_{02(CT0)}$ and $\beta_{02(CT2)}$. This is possible because the values of R_{eqT}, R_{p0}, and R_{p2} are nearly invariant across the three tests. Eq. 10 and 11 gives the expressions for a and b in terms of the current ratios derived from $C4_0$ and $C4_2$ CT tests.

$$m = \frac{L(\beta_{02(CT0)}(1+\beta_{02(CT2)})-\beta_{02}(1+\beta_{02(CT2)}))}{(1+\beta_{02})(\beta_{02(CT0)}-\beta_{02(CT2)})}$$

$$a = \frac{L}{2} - m \qquad (10)$$

with,

β_{02} = current ratio I_0/I_2 at state with defect provoked
$\beta_{02(CT0)}$ = current ratio $I_{0(CT0)}/I_{2(CT0)}$ with CT_0 on
$\beta_{02(CT2)}$ = current ratio $I_{0(CT2)}/I_{2(CT2)}$ with CT_2 on
L= distance between two adjacent C4 V_{DD}s

$$b = \sqrt{c^2 - a^2} = \sqrt{\left(\frac{L}{2}\right)^2 - a^2} \qquad (11)$$

Thus, a nice feature of this calibration technique is that it is independent of the R_p, which are likely to vary from touch-down to touch-down of the probe card.

A second problem addressed by the CTs is related to the procedure described in Section 3.1. Pad selection is accomplished by sorting the I_{DDQ} values and identifying the pad with the largest I_{DDQ} as the "primary" pad (pad *j*). Two (of the four) orthogonally adjacent supply pads to pad *j* are then selected from the top of the sorted list. Unfortunately, this algorithm fails to select the pads surrounding the defect under certain conditions. For example, Figure 15 shows a portion of the supply grid with 9 C4s. The defect is located in the upper left Quad and therefore, the algorithm should select $C4_3$ as the "primary" pad and $C4_1$ and $C4_7$ as the orthogonally adjacent pads. However, if $R_{eqa} > R_{eqb}$, the sorted list places $C4_5$ above $C4_1$ and the algorithm incorrectly selects $C4_5$. This type of resistance anomaly can occur, for example, if the power grid mesh is denser between $C4_3$ and $C4_5$ than it is between $C4_3$ and $C4_1$.

The CT data can be used to instrument a more robust pad selection algorithm. The current ratio $\beta_{31(CT3)} = I_{3(CT3)}/I_{1(CT3)}$ obtained by turning on the CT under $C4_3$ gives the upper bound on the current ratio between $C4_3$ and $C4_1$. The current ratio computed under the circuit state with the defect provoked, β_{31}, is necessarily less than the $\beta_{31(CT3)}$, since $\beta_{31(CT3)}$ is the maximum ratio. Therefore, an improved algorithm selects the correct secondary pads, e.g. $C4_1$ instead of $C4_5$, by using CT ratio data.

It is also possible in some grid configurations that the largest I_{DDQ} is not drawn from the pad that is closest to the defect site. In this case, the existing algorithm does not select the correct "primary" pad. We are currently investigating the use of CT data to solve this problem, and hope to describe a solution in a future work.

A third problem addressed by the CTs is related to the assumption that the unity current ratio line (the 1-line or center for the hyperbolas) is positioned midway between the C4s. This is only true for simple grids (such as the Quad shown in Figure 4) if the R_p are equal. If the R_p are not equal, Eq. 10 can be used to derive the offset, *c'*, of the 1-line by setting β_{02} to 1 and simplifying.

A similar shift occurs in more complex grids, such as that shown in Figure 15, but for a reason related to the degree of symmetry in the C4s surrounding a region. For example, the bottom portion of the grid in Figure 15 contains a row of three C4s, $C4_0$, $C4_2$ and $C4_4$. The 1-line in the lower left Quad is shown skewed to the right from the midpoint given by L/2. The asymmetry in the C4s surrounding this region, e.g. C4s 0, 1 and 6 on the left and C4s 2-5, 7 and 8 on the right, are responsible for this shift. We are currently evaluating more complex circuit models such as the one shown in Figure 5 as a means of formulating an expression that accounts for this shift. Experimentally, we determined that Eq. 12 yields a good approximation of the offset, *c'*, of 1-lines for Quads within the PPG.

$$c' = \frac{L}{2}(\beta_{02(CT0)}\beta_{02(CT2)}) \qquad (12)$$

4.2 Leakage Current

One element that we haven't addressed is the impact of leakage current. A second calibration method was proposed in previous work to deal with leakage [5]. Although calibrating for leakage is clearly an important issue, we do not focus on it in this work because of space limitations. The limited number of experiments conducted thus far involving leakage indicate that it has only a small impact on the accuracy of the predictions. The same is true for experiments conducted using different values of defect current. Current ratios are naturally robust to these variables but a quantitative analysis of their impact remains to be determined and will be addressed in a future work.

4.3 The QSA Procedure

The procedure to localize a defect follows from the discussion given in the previous section. Once a chip is identified as defective, e.g. from a Stuck-At or I_{DDQ} go-nogo test, the following tests are performed under the QSA procedure. First, the chip is set to a state that provokes the defect and the individual I_{DDQ} values are measured. The C4 pad, *j*, sourcing the largest I_{DDQ} and two orthogonally adjacent C4 pads, *x* and *y*, are identified as described in Section 4.1. The current ratios β_{jx} and β_{jy} are computed. The chip is then put into a state that doesn't provoke the defect. The CT for the *jth* pad is turned on and the current ratios $\beta_{jx(CTj)}$ and $\beta_{jy(CTj)}$ are computed. Similarly, the current ratios $\beta_{jx(CTx)}$ and $\beta_{jy(CTy)}$ are computed from measurements made with CT_x and CT_y turned on. Eq. 12 gives the offsets needed to derive the two centers of the hyperbolas, $(h',k)_x$ and $(h,k')_y$, along the x- and y-dimension, respectively, from pad *j*. Eq. 10 is then used to derive a_x and a_y parameters using $L_x = 2*c'_x$ and $L_y = 2*c'_y$ for L. The b_x and b_y parameters are computed using Eq. 11. These two pairs of *a* and *b* parameters define both the position and shape of one hyperbola from each of the two families, e.g. as illustrated in Figure 16 using the hyperbola curves from Figures 10 and 11. The intersection of these two hyperbolae gives the predicted location of the defect.

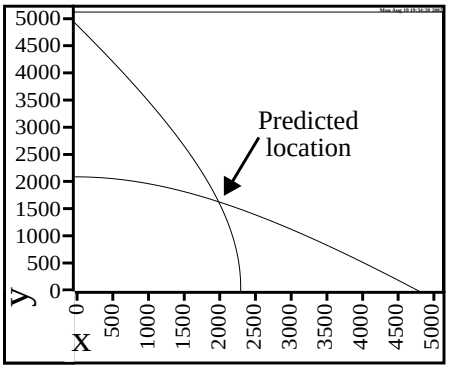

Figure 16. Example prediction using the hyperbola curves from Figures 10 and 11.

The algorithm, as stated, requires a change in the state of the CUT after the first set of I_{DDQ} measurements are made. Therefore, the contribution of leakage to the currents measured with the CTs turned on is different than the contribution under the state with the shorting defect provoked. The vector-to-vector leakage variation is likely to adversely affect the accuracy of the predictions. An alternative test procedure that does not change the CUT's state is to perform the CT tests with the defect provoked. The currents measured under the CT tests can be "adjusted" by subtracting the currents measured under the defect provoking test. Even though the presence of the defect's current is likely to change the equivalent resistances of the CUT under the CT tests, we expect the error introduced by this type of procedure to be smaller than the error introduced under test scenarios in which the vector-to-vector leakage variation is large.

5.0 Experimental Results

This algorithm was applied to the data obtained from 200 SPICE simulations of the 30,000 by 30,000 unit region of the PPG referred to as Q9 in Figure 3. A three dimensional error map plotting the prediction error against the (*x*,*y*) coordinate of the inserted defect (modeled using a current source) is shown in Figure 17. The prediction error is computed as the directed distance between the predicted location and the actual location of the defect. The average and worst case prediction errors are 215 and 650 units, respectively.

The size of the simulation model for the entire PPG shown in Figure 3 made it impossible to perform SPICE simulations on it. Instead, the PPG was simulated using a specialized power grid simulation engine called ALSIM. The prediction error map from 500 ALSIM simulations is shown in Figure 18. The average and worst case prediction errors are 410 and 1,340 units, respectively. The increase in

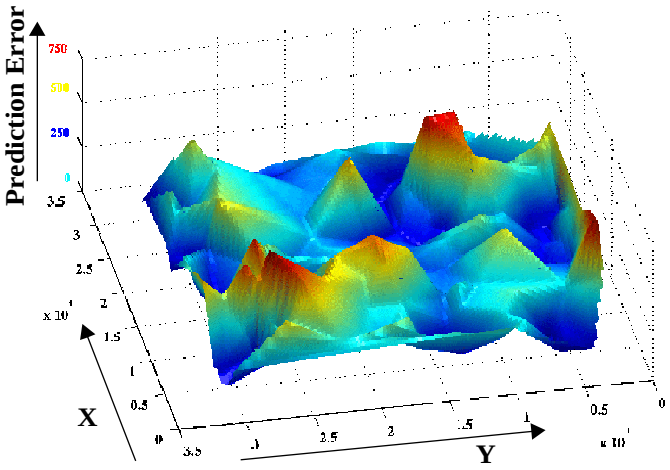

Figure 17. Prediction error map of the Q9.

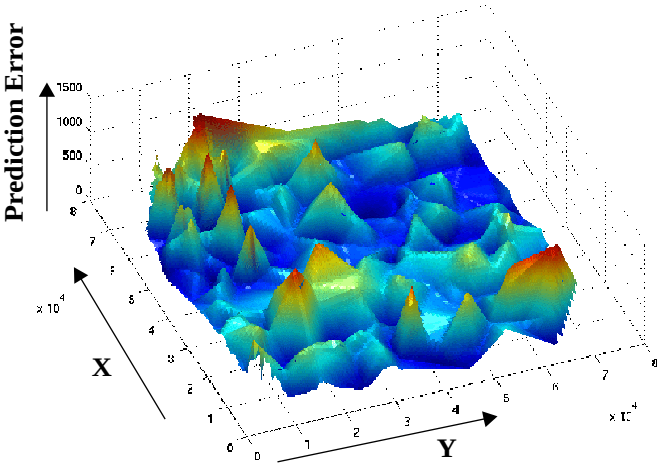

Figure 18. Prediction error map for the entire PPG.

prediction error is largely due to the more significant anomalies in the grid's structure over the larger region defined by the entire PPG.

6.0 Conclusions

The weaknesses of our previously derived resistance-based Quiescent Signal Analysis model are addressed in a new current-ratio-based technique. Calibration transistors are proposed to reduce the adverse effects of probe card resistance variations on the prediction accuracy of the new QSA technique. The calibration transistor data is also used to account for power grid resistance variations from one region to the next and asymmetrical or irregular arrangements in the positions of the power supply pads.

The current ratio contours derived through SPICE simulations of a commercial power grid are shown to be well approximated by "families" of hyperbola curves. An analytical framework is derived that allows the measured I_{DDQ} data to be translated to physical (x,y) layout coordinates, that represent the position of the defect.

Although the analytical model that we present in this work accounts for tester environment variables such as probe card resistance variations, the simulation data was derived from a simpler model. For example, the probe card resistance was held constant at 1Ω at every supply pad, 20mA was used for defect currents, and leakage currents were not included. As pointed out, current ratios are naturally robust to variations in defect current and the calibration transistors in combination with regression analysis are expected to be effective in dealing with leakages. Simulation experiments are currently underway to verify these hypotheses.

The last issue that remains to be explored is the effectiveness of this technique on other types of grid topologies. For significantly irregular grids, we expect a lookup-table approach to be more accurate than the hyperbola-based technique. We are investigating the use of power grid simulators such as ALSIM as a means of making this type of approach practical.

Acknowledgments

We thank Dr. Anne Gattiker and Dr. Sani Nassif at IBM Austin Research Lab for their support of this research.

References

[1] T.W.Williams, R.H.Dennard, R.Kapur, M.R.Mercer & W.Maly, "I_{DDQ} test: Sensitivity Analysis of Scaling", ITC, 1996, pp.786-792.

[2] A.E.Gattiker and W.Maly, "Current Signatures", VTS, 1996, pp.112-117.

[3] C. Thibeault, "On the Comparison of Delta IDDQ and IDDQ Test", VTS, 1999, pp. 143-150.

[4] Peter Maxwell, Pete O'Neill, Rob Aitken, Roland Dudley, Neal Jaarsma, Minh Quach, Don Wiseman, "Current Ratios: A self-Scaling Technique for Production IDDQ Testing", ITC, 1999, pp.738-746.

[5] Jim Plusquellic, "IC Diagnosis Using Multiple Supply Pad IDDQs", Design and Test, Vol. 18, No. 1, Jan/Feb 2001, pp. 50-61.

[6] Chintan Patel and Jim Plusquellic, "A Process and Technology-Tolerant IDDQ Method for IC Diagnosis", VTS, 2001, pp. 145-150.

[7] C. Patel, E. Staroswiecki, D. Acharyya, S. Pawar and J. Plusquellic," A Current Ratio Model for Defect Diagnosis using Quiescent Signal Analysis", Workshop on Defect Based Testing, VTS, 2002.

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

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

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

[11] MOSIS at http://www.mosis.edu/Technical/Testdata/tsmc-025-prm.html.

[12] Chintan Patel, Ernesto Staroswiecki, Dhurva Acharyya, Smita Pawar, and Jim Plusquellic, "Defect Diagnosis using a Current Ratio based Quiescent Signal Analysis Model for Commercial Power Grids", Accepted for publication in VLSI Design, an International Journal of Custom-Chip Design, Simulation, and Testing, 2002.