

Quiescent-Signal Analysis: A Multiple Supply Pad I_{DDQ} Method

Jim Plusquellic, Dhruva Acharyya, Abhishek Singh,
Mohammad Tehranipoor, and Chintan Patel

University of Maryland, Baltimore

Increasing leakage current makes single-threshold I_{DDQ} testing ineffective for differentiating defective and defect-free chips. Quiescent-signal analysis is a new detection and diagnosis technique that uses I_{DDQ} measurements at multiple chip supply ports, reducing the leakage component in each measurement and significantly improving detection of subtle defects. The authors apply regression and ellipse analysis to data collected from 12 test chips to evaluate the technique.

using transient-signal measurements.⁷ These techniques analyze multiple simultaneous measurements to accomplish three goals: detect local signal variations introduced by defects at point sources in the layout, reduce the adverse impact of background leakage current, and diminish the adverse effects of within-die and between-die

■ **AS SILICON TECHNOLOGY MOVES** forward, background leakage current continues to increase. This trend reduces the effectiveness of traditional I_{DDQ} testing methods and poses a challenge for newer, alternative strategies.¹ Alternative methods rely on a self-relative or differential analysis, which factors each chip's average I_{DDQ} into the pass-fail threshold. Although application of these techniques to low-power chips will continue, they will be increasingly less effective for high-performance ASICs with high background leakage currents.

An alternate strategy that could have better scaling properties is to measure the I_{DDQ} from each of the individual power ports. In this case, the total leakage current of the chip is distributed across a set of simultaneous measurements. Our method, called quiescent-signal analysis (QSA), exploits this type of measurement scheme to increase the ratio of defect current to leakage current. Other publications describe a secondary diagnostic benefit of this technique.^{2,5}

In previous work, we developed several statistic-based methods for processing data collected from simultaneous measurements. We developed a linear regression analysis procedure and applied it to simulation data obtained from a commercial power grid.⁶ A hyperbola-based method performs defect detection

process variations.

In this article, we apply linear regression analysis and a new technique called ellipse analysis to the data collected from a set of 12 test chips to illustrate QSA's defect detection capabilities and limitations. The test chips, which Jim Plusquellic designed while on sabbatical at IBM Austin Research Laboratory, were fabricated in a 65-nm, 10-metal-layer technology. They incorporate an array of test structures that let us emulate a defect in one or more of 4,000 distinct chip locations. The design permits control over the magnitude of the emulated-defect current and the leakage current.

In particular, we show that regression analysis applied to the data from 21,600 emulated defects detected 99.4 percent of the emulated defects with less than 0.9 percent yield loss. Regression performed better than ellipse analysis, with results of 94.6 percent and 4.8 percent for detection and yield loss, respectively. However, with a restricted defect-free data set, ellipse analysis detected 99.7 percent of the emulated defects with no yield loss. Detection sensitivity as well as the level of confidence in the detection decision strongly correlate with the emulated defect's position in the layout. Both of these measures relate inversely to the distance

Related Work

The single-threshold I_{DDQ} technique relied on the steady-state current's distribution of defect-free chips being distinct from that of the defective chips. A chip that draws current exceeding the defect-free current distribution by a fixed threshold is deemed defective. In deep-sub-micron technologies, however, the distributions overlap, and it is not possible to set an absolute pass-fail threshold that distinguishes defect-free and defective chips. The increase in subthreshold and gate leakage currents in newer technologies can result in background leakage currents significantly higher than the defect current. Thus, the chip-manufacturing industry needs alternative techniques to reduce the adverse effects of high background leakage currents on defect current resolution. Researchers have proposed the following techniques based on self-relative or differential analysis as a solution to this problem:

- a current signature method that looks for discontinuities in the curve obtained by sorting I_{DDQ} measurements in ascending order;¹
- a differential I_{DDQ} method (Delta I_{DDQ}) in which differences between successive I_{DDQ} measurements are compared with a threshold;²
- a current ratio method that derives chip-specific thresholds by using vectors that produce minimum and maximum I_{DDQ} values;³
- a clustering technique that groups good chips separately from bad chips;⁴
- a method that predicts device I_{DDQ} using the spatial proximity correlations among chips on a wafer;⁵
- a linear prediction-based technique in which each I_{DDQ} value among a set of values for a given chip is predicted from the remaining I_{DDQ} values in the set;⁶
- a method using I_{DDQ} readings of the neighboring die on a wafer to reduce variance and identify wafer-level spatial outliers;⁷
- methods based on wafer-level spatial correlation analysis, which derive a maximum defect-free I_{DDQ} threshold from analysis of neighboring dies.^{8,9}

Many of these process-tolerant I_{DDQ} methods use relative pass-fail thresholds instead of absolute thresholds,

and all use global, or chip-wide, I_{DDQ} measurements. As the variance in I_{DDQ} increases, it tends to increase the threshold bands in many of these techniques, thus decreasing their sensitivity to defects. Quiescent-signal analysis (QSA) differs from these methods by cross-correlating local, or within-chip, I_{DDQ} measurements obtained from multiple, individual supply ports on the chip. In addition to the benefits identified in the introduction of this article, local I_{DDQ} measurements eliminate the adverse effects of vector-to-vector variations inherent in global I_{DDQ} measurement strategies.

References

1. A.E. Gattiker and W. Maly, "Current Signatures," *Proc. 14th VLSI Test Symp. (VTS 96)*, IEEE Press, 1996, pp. 112-117.
2. C. Thibeault, "On the Comparison of Delta I_{DDQ} and I_{DDQ} Test," *Proc. 17th VLSI Test Symp. (VTS 99)*, IEEE Press, 1999, pp. 143-150.
3. P. Maxwell et al., "Current Ratios: A Self-Scaling Technique for Production I_{DDQ} Testing," *Proc. Int'l Test Conf. (ITC 99)*, IEEE Press, 1999, pp. 738-746.
4. S. Jandhyala, H. Balachandran, and A.P. Jayasumana, "Clustering Based Techniques for I_{DDQ} Testing," *Proc. Int'l Test Conf. (ITC 99)*, IEEE Press, 1999, pp. 730-737.
5. W.R. Daasch et al., "Variance Reduction Using Wafer Patterns in I_{DDQ} Data," *Proc. Int'l Test Conf. (ITC 2000)*, IEEE Press, 2000, pp. 189-198.
6. P.N. Variyam, "Increasing the I_{DDQ} Test Resolution Using Current Prediction," *Proc. Int'l Test Conf. (ITC 2000)*, IEEE Press, 2000, pp. 217-224.
7. A. Singh, "A Comprehensive Wafer Oriented Test Evaluation (WOTE) Scheme for the I_{DDQ} Testing of Deep Sub-Micron Technologies," *Proc. Int'l Workshop on I_{DDQ} Testing*, IEEE Press, 1997, pp. 40-43.
8. S. Sabade and D.M.H. Walker, "Improved Wafer-Level Spatial Information for I_{DDQ} Limit Setting," *Proc. Int'l Test Conf. (ITC 01)*, IEEE Press, 2001, pp. 82-91.
9. S. Sabade and D.M.H. Walker, "Neighbor Current Ratios (NCR): A New Metric for I_{DDQ} Data Analysis," *Proc. 17th Int'l Symp. Defect and Fault Tolerance in VLSI Systems (DFT 02)*, IEEE Press, 2002, pp. 381-389.

between the emulated defect and the nearest neighboring V_{DD} . (The "Related work" sidebar reviews other proposed techniques for handling the background leakage current problem.)

Test chip design

Figure 1a shows a block diagram of the test chip design. It consists of an 80×50 array of test circuits (TCs), which occupies an area 558 microns wide and

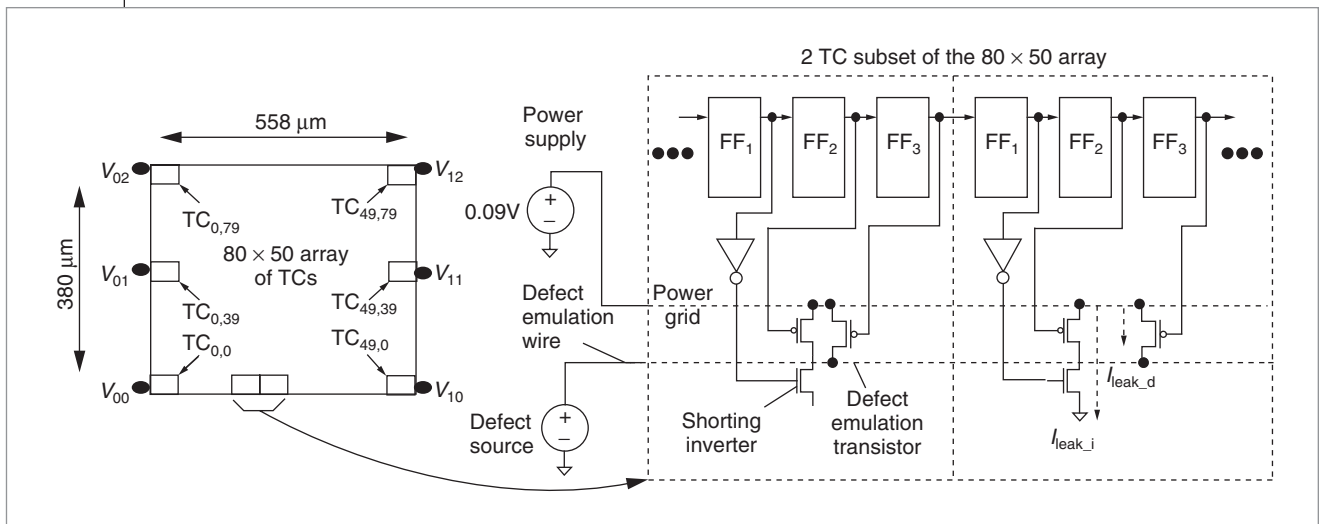


Figure 1. Block diagram of test structure (a) and test circuit (TC) details.

380 microns high. Each TC consists of three flip-flops (FFs) connected in a scan chain configuration, a shorting inverter, and a defect emulation transistor connected to a globally routed defect emulation wire. Figure 1b shows a schematic diagram of two adjacent TCs. 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 power grid point sources enables introduction of two types of shorts in any one (or more) of the 4,000 TCs. The first type shorts the power grid to ground through the inverter using FF₁ and FF₂; the second type shorts the power grid to the defect emulation wire using FF₃. For the first type, the external power supply voltage (see Figure 1a) defines the shorting current's magnitude. (In our experiments, we held the power supply constant at 0.9 V.) For the second type, an external voltage source ("Defect source") controls the shorting current's magnitude. Given this configuration, we can emulate a defect at any point in the array by setting the defect source to a value less than the power supply voltage and by scanning a bit pattern into the scan chain such that exactly one FF₃ contains a 0, and the remaining 11,999 FFs contain 1s.

In addition to controlling the defect current's magnitude, the defect source influences the background leakage current's magnitude, as measured through the power supply. As Figure 1b shows, the total leakage current consists of two types: I_{leak_i} (inverter) and I_{leak_d} (defect). 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 sink leakage current from the power 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, we can analyze various shorting and leakage current configurations by controlling the states of the defect emulation transistors and voltage on the defect emulation wire.

Figure 2 shows the external instrumentation setup. Power ports V_{00} through V_{12} wire out of the chip on separate pins in the package. Individual power pins are each wired to a low-resistance mechanical switch, which can be configured in one of three positions: left, middle, or right. The left and right outputs of the six switches connect to a common wire that routes to the global current source meter (GCSM) and the local current ammeter (LCA), respectively. The middle switch position's output floats, allowing experiments in which a subset of V_{xx} ports connects to neither the GCSM nor the LCA.

The GCSM provides 0.9 V to the power grid and can also measure current with accuracies less than 100 nA. The LCA is wired in series with the GCSM and allows measurement of individual power port (local) currents at the same accuracy level. For example, the switch configuration in Figure 2 allows measurement of the local V_{00} current, I_{00} , as well as the global current. The defect emulation source meter (DESM) sets the voltage of and measures current I_{def} through the defect emulation wire on a separate pin in the package (not shown).

The experiments described in this article tested two switch configurations: The four- V_{DD} configuration uses only four V_{DD} s to power up the grid; that is, switches con-

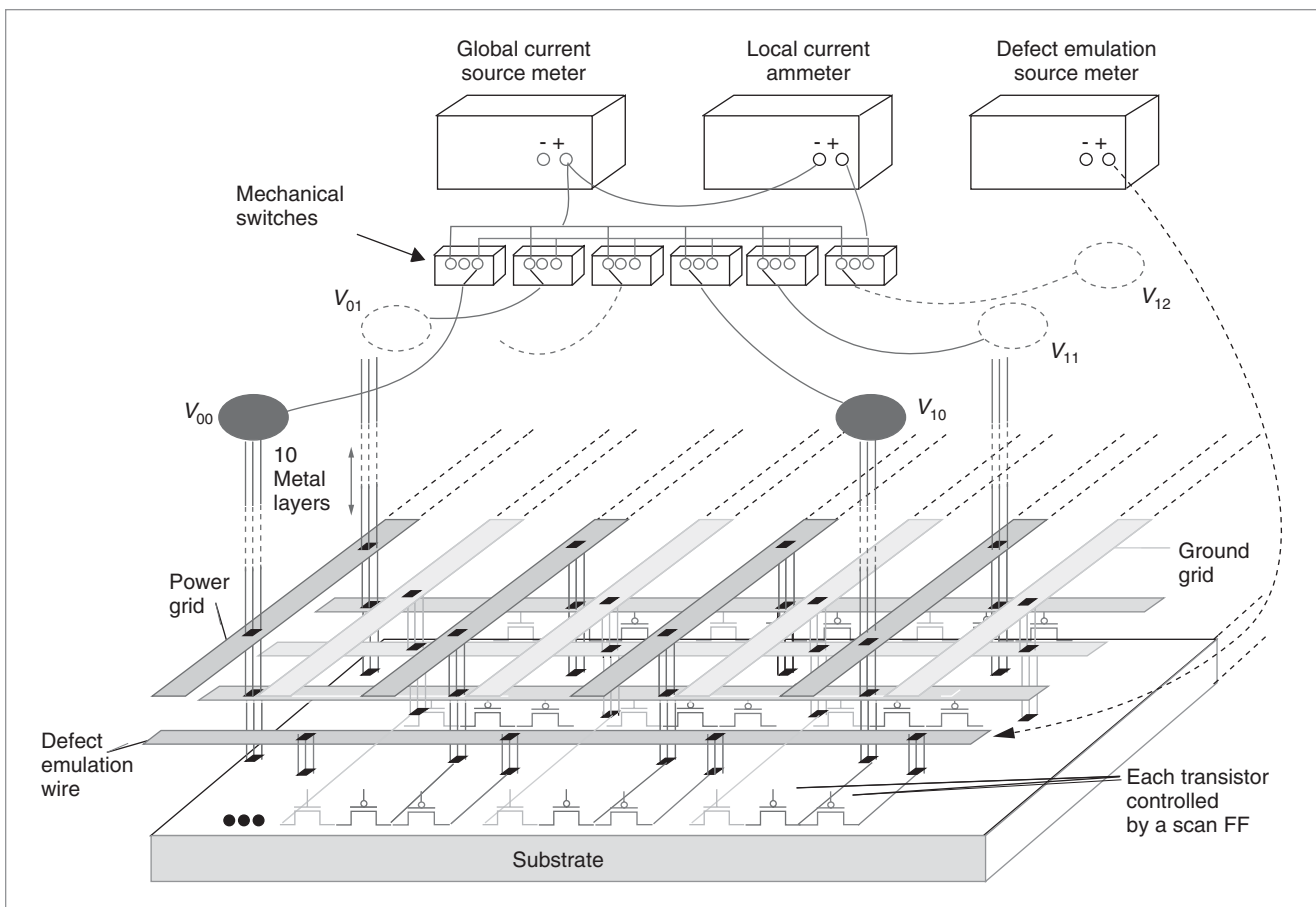


Figure 2. External instrumentation setup.

trolling the connection of V_{01} and V_{11} to the GCSM are set to their high-impedance (middle) positions. The six- V_{DD} configuration uses all six V_{DD} s.

Power grid characterization experiments

We designed the first set of experiments to determine how grid resistance influences local currents' magnitude. In these experiments, we disconnected the defect emulation wire, disabled the defect emulation transistors, and used the shorting inverters instead to provide stimulus to the grid. In theory, either structure could serve this purpose, but the shorting inverters minimally change the array's leakage current when enabled and provide a larger current than the defect emulation transistors, thus enhancing the signal-to-noise ratio.

We enabled each of the 4,000 shorting inverters from one of the chips, one at a time, and measured global and local currents. Because we were interested in characteristics of the grid resistance and its influence on local current distributions from layout point sources, we

also performed the following steps: After testing each array element, we disabled the shorting inverter of the TC under test, measured global and local leakage currents, and subtracted them from the values measured with the shorting inverter enabled. (Although it may appear that only one set of global and local leakage measurements are necessary, chip temperature variations cause leakages to vary over time. To minimize this source of variation, we made leakage measurements immediately following the shorting inverter current measurements for each TC.) We then normalized these current differences by dividing them by the global current. This type of normalization virtually eliminates variations in transistor current magnitudes caused by process variations.

Figure 3a shows the current profile derived from the normalized local currents, I_{norm_00} , under a four- V_{DD} configuration; that is, V_{00} , V_{02} , V_{10} , and V_{12} are powered from the GCSM. The x - and y -axes represent the TC array's (x , y) plane, and the z -axis represents I_{norm_00} . Figure 3b shows the I_{norm_11} profile under a six- V_{DD} configuration.

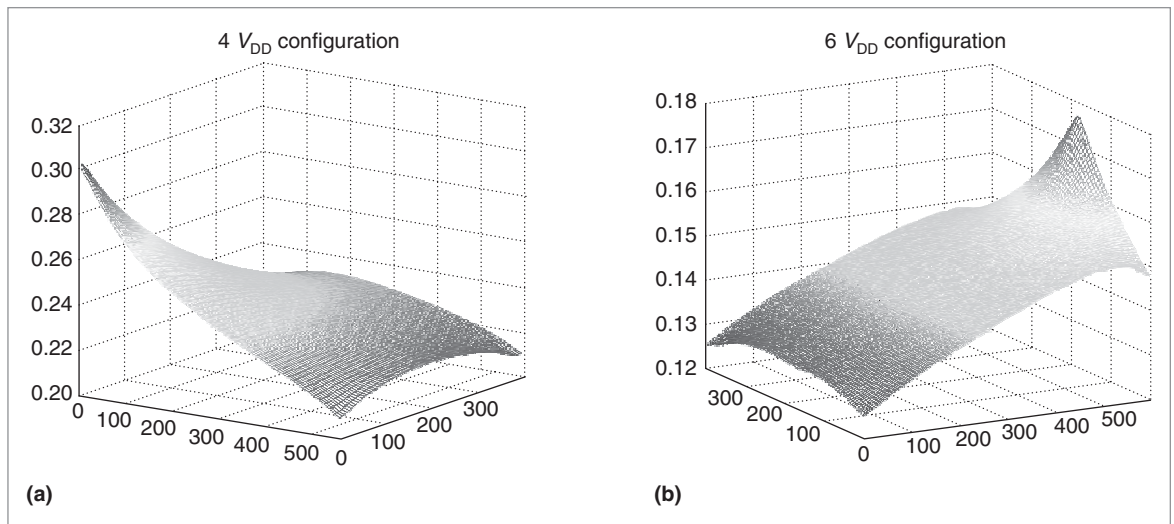


Figure 3. I_{norm_00} (a) and I_{norm_11} (b) profiles.

Local currents are largest near V_{00} and V_{11} in each plot, respectively, because TCs near these locations draw a larger fraction of current from V_{00} and V_{11} (maximums are approximately 31 percent and 17 percent, respectively) than TCs further removed. The range of values in each plot of approximately 11 percent and 4.5 percent shows the effect of grid resistance on current distribution to the V_{DD} s. The smooth, monotonically decreasing nature of the surfaces from largest to smallest provides the basis for building defect detection methodologies.

I_{DDQ} defect detection experiments

The purpose of these experiments was to investigate our defect detection methodologies and their sensitivity to defects that draw only small amounts of current. To best meet these objectives, we used the defect emulation transistors and the corresponding defect emulation wire because we could control both position and magnitude of the emulated-defect current.

Data collection procedure

Unlike the power grid characterization experiments, which tested all 4,000 TC array elements, these experiments tested only a 100-TC subset. Figure 4 shows the set of randomly selected TCs in the 80×50 array. The numbered positions are the TCs under investigation.

For each of the 12 test chips, we performed a series of measurements for each TC under different voltage configurations of the DESM—the source meter that drives the defect emulation wire. The first experiment for each chip is the defect-free experiment. In this experi-

ment, we set the state of all scan chain FFs to 1, which disables both the shorting inverters and the defect emulation transistors in all TCs in the array. We then swept the DESM 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, we measured a set of local and global currents through the V_{DD} s—four local and four global under the four- V_{DD} configuration, and six under the six- V_{DD} configuration. We performed the same operation sequence with each enabled defect emulation transistor, one at a time.

Ideally, all local currents are measured simultaneously. This minimizes signal variations that can occur—for example, from thermal drift—if the measurements are sequential. In our experiments, the limited number of ammeters prevented simultaneous local current measurements. Instead, we measured each local current simultaneously with the global current. A simple correction procedure corrected local currents for drift. The correction procedure computes corrected local currents I_{corr_xx} by scaling the measured I_{xx} by the ratio of global currents I_{glob_yy} and I_{glob_xx} , where I_{glob_yy} is a global current that serves as the reference: $I_{\text{corr}_xx} = (I_{xx})(I_{\text{glob}_yy}/I_{\text{glob}_xx})$.

Data sets and pairings

For each chip, the data collection procedure produces 1,919 data sets, of which 19 represent data from the defect-free experiments and 1,900 (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

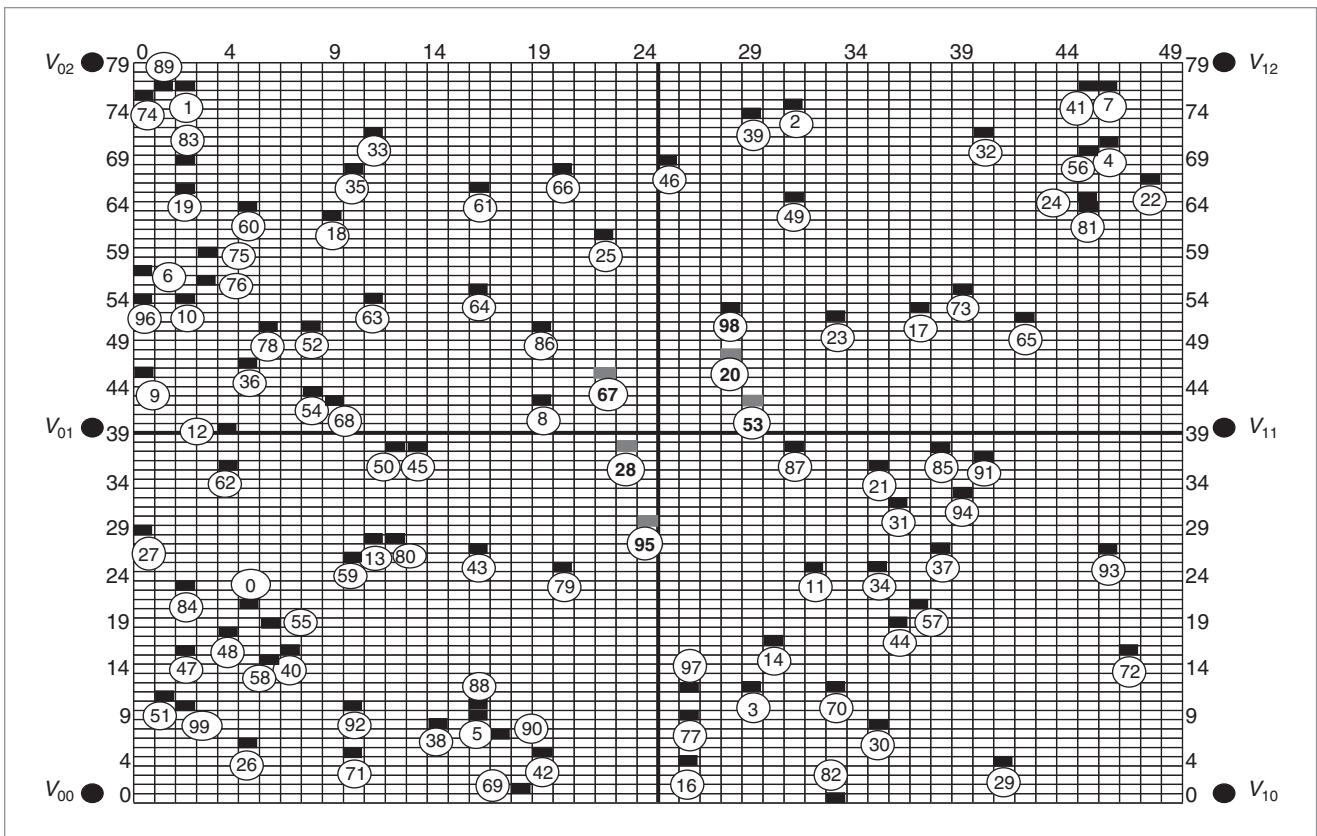


Figure 4. Positions of 100 randomly selected TCs.

because there is no voltage drop across the defect emulation transistor. Therefore, we treat only 18 of the 19 data sets as emulated defects. With 12 chips, there are $12 \times 19 = 228$ defect-free data sets and $12 \times 1,800 = 21,600$ emulated-defect data sets.

We performed the analysis on pairs of local currents in each data set. Figure 5 displays the V_{DD} ports for the four- V_{DD} and six- V_{DD} configurations and lists the possible pairings as two subsets in each configuration. For example, the four- V_{DD} configuration (Figure 5a) has an orthogonal-neighbors subset (V_{00} - V_{02} , V_{00} - V_{10} , and so on) and a cross-neighbors subset. For the six- V_{DD} configuration (Figure 5b), 11 of the pairings fall into the orthogonal- and cross-neighbors subset, and the remaining four fall into the nonneighbors subset. We performed the analysis on the entire set (all pairings) and on subsets identified as orthogonal neighbors and orthogonal and cross neighbors to determine the impact of the “pairings” parameter on defect detection sensitivity.

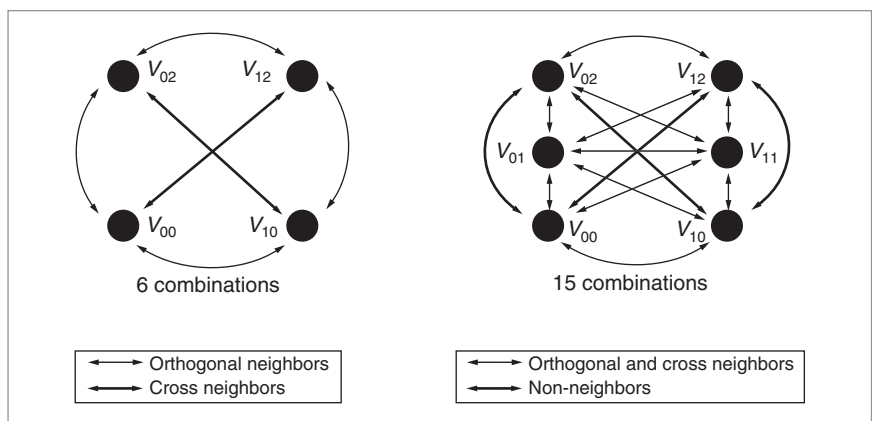


Figure 5. Pairing combinations under four- V_{DD} (a) and six- V_{DD} (b) configurations.

Correlation analysis for variation

The primary purpose of sweeping the DESM across 19 different values is to answer questions such as “At what DESM voltage levels can we detect emulated defects?” and “How high is each positive detection’s confidence level?” A secondary purpose is to solve problems associated with applying statistical methods to small

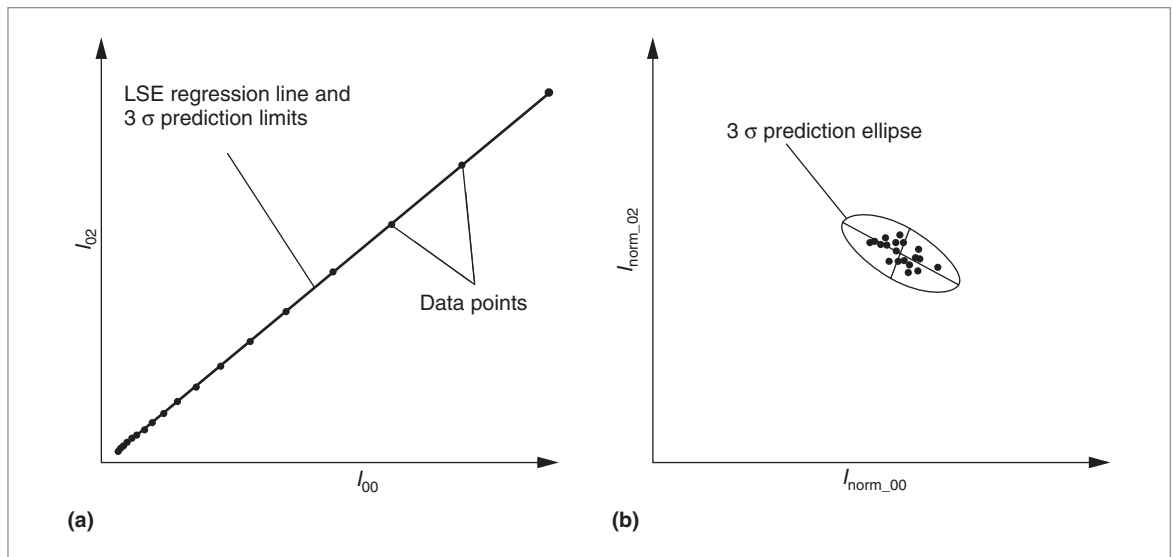


Figure 6. Chip C1 scatter plot of defect-free data: I_{00} vs. I_{02} (a) and I_{norm_00} vs. I_{norm_02} (b). Change LSE in figure to least squares estimate.

data sets—that is, the 12 chips. For this, we used the 19 defect-free data sets obtained from each chip under different DESM voltages to increase the sample size. To eliminate any concerns that this might unfairly bias the results, we also give the results of an analysis that used only one set of defect-free data from each chip.

A third purpose is the analysis of variance in the data. By using the defect-free data sets obtained from a single chip, we can decompose several sources of variation that affect the results, such as variations introduced by test apparatus and power grid parameters. A comparative analysis using data sets from other chips lets us identify the relative magnitude and significance of these variation sources.

Correlation analysis of defect-free data from a single chip

Perhaps the most challenging aspect of hardware experiments, in comparison with simulation experiments, is understanding and accounting for various sources of signal variations. The measured parameter in our experiments, I_{DDQ} , is analog in nature and is subject to variations associated with the measurement instrumentation and test apparatus. These variations include noise and series parasitic resistances, as well as variations in the chip itself, such as pin and routing parasitics and within-die and between-die process variations. It is important to understand the relative effect of these signal variation sources as well as to have a means of calibrating for them.

In the context of our test methods, the most meaningful approach to decomposing signal variation sources is analysis of scatter plot variance. Two types of scatter plots are of interest in our analysis. The first type is constructed with absolute local currents; the second type is constructed with normalized local currents. For example, Figure 6a plots absolute local currents I_{00} along the x -axis against the corresponding I_{02} on the y -axis for chip C1. The plot includes 19 pairs of values, one pair for each DESM voltage. In contrast, the scatter plot in Figure 6b shows the same data except that each local current is first divided by global values measured simultaneously, as described earlier. The normalization operation's effect is to remove the absolute current's magnitude from consideration. In other words, the dispersion of the data points along the line as portrayed in Figure 6a is eliminated, and the data points are effectively clustered together in a blob as shown in Figure 6b.

We apply standard variance analysis methods to these scatter plots. For the data in Figure 6a, we applied linear regression analysis by computing a best-fit line through the data points and a set of 3-sigma prediction limits. For the data in Figure 6b, we used a prediction ellipse method, computing the elliptical bound around the data points from the eigenvalues of their covariance matrix and a 3-sigma X^2 (chi-square) distribution statistic.

The data points in Figure 6a are nearly colinear, yielding very narrow prediction limits, expectable because the data is derived from a single chip. Therefore, several important parameters that introduce

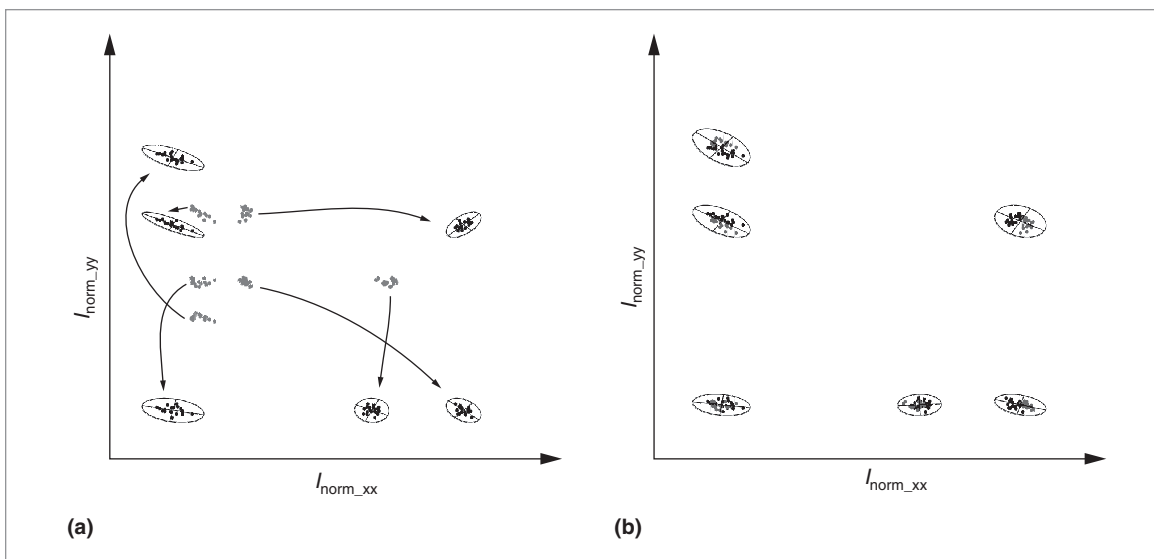


Figure 7. Chips C1 and C2: All pairings of I_{norm} , uncalibrated (a) and all pairings of I_{norm} , calibrated (b).

dispersion of the data points are held constant, such as series resistance between the power supply and the V_{DD} ports and on-chip process variation parameters. The remaining variation sources are environmental changes such as temperature and noise. We minimize temperature variations by collecting both data points in each pair closely together in time as described earlier. Therefore, most variation is due to noise. The noise floor in the existing test setup is approximately 300 nA.

We draw similar conclusions from the data displayed in Figure 6b, except that the dispersion is more pronounced. This is due, in part, to the difference in scaling factors used to plot the data. However, the dispersion is actually larger than that present in the scatter plot in Figure 6a. In this case, the normalization operation is responsible for increasing the dispersion because the divisor—the global current measurement—is also subject to noise. Moreover, the measurement’s global context—the entire array—is subject to a wider range of process variations than the regional context associated with two local measurements. As will become evident in the defect sensitivity analysis given in the following sections, these elements can reduce detection sensitivity of small defect currents.

Correlation analysis of defect-free data from multiple chips

The data in Figure 6 is drawn from a single chip and therefore does not represent an actual testing environment, in which data defining the chip’s defect-free behavior would be drawn from a far larger sample. With

such a sample, sources of variations not present in the single-chip analysis will affect the dispersion level of scatter plot data points.

Ellipse analysis illustrates this more clearly than regression analysis. Figure 7 plots normalized currents from two chips, C1 and C2, for all six pairings of the four V_{DD} s. We derived the ellipses in the scatter plot in Figure 7a from the data collected from C1. The tails of the arrows identify corresponding data sets for C2. Although dispersion in either set is approximately the same, the position of the clusters in the 2D plane is significantly different. The arrows indicate where the C2 clusters should appear to be consistent with the C1 clusters.

The displacement of the clusters from each other is caused primarily by variations in series resistance between the power supply and each power grid attachment point. Series resistance variations can occur in either of the wire segments that define these paths—namely, the segment between the power supply and the V_{DD} ports on the packaged chips and the segment between the V_{DD} ports and the power grid in the package and chip. Because we use the same test apparatus to collect data from both of these chips, resistance variations along the first segment are nearly zero. Therefore, series resistance differences must occur along the second segment—within the package and chip—or as contact resistance variations in the clamshell ZIF (zero insertion force) socket on the test board.

Clearly, the area enclosed by the ellipse would increase significantly if we didn’t correct for this type of variation. In previous work, we developed and demon-

strated a probe card calibration (PCC) technique.⁷ The method uses data collected from a special set of calibration circuits similar in design to the TCs shown in Figure 1b, with only the shorting inverters present. For PCC, we place one copy of the TC underneath each V_{DD} port. Figure 1a shows the positions of the TCs ($TC_{0,0}$, $TC_{0,39}$, and so on) used for PCC in these experiments. We collect the data for PCC by enabling the shorting inverter in each TC, one at a time, and measuring local currents. We also take leakage measurements as described earlier and subtract them from the shorting-inverter currents. (Peak values in Figures 3a and 3b correspond to the shorting-inverter currents of $TC_{0,0}$ and $TC_{49,39}$, respectively. By placing the TCs at points directly beneath the power ports, we minimize the signal-to-noise ratio. This choice of position is also beneficial for diagnostic purposes.⁸) We normalize the shorting-inverter currents by dividing them by the global current, both with leakage subtracted.

We use the data matrix collected in the PCC tests to calibrate the local currents measured in other tests. We do this by using the matrices obtained from two chips: a chip whose data is to be corrected, such as C2, and a reference chip, such as C1. For each chip, the matrix is 4×4 under the four- V_{DD} configuration. Equation 1 gives the expression for computing the transformation matrix, X :

$$X = C2^{-1} * C1 \quad (1)$$

$$\begin{bmatrix} x_{00} & x_{01} & x_{02} & x_{03} \\ x_{10} & x_{11} & x_{12} & x_{13} \\ x_{20} & x_{21} & x_{22} & x_{23} \\ x_{30} & x_{31} & x_{32} & x_{33} \end{bmatrix} = \text{inv} \begin{bmatrix} a_{00} & a_{01} & a_{02} & a_{03} \\ a_{10} & a_{11} & a_{12} & a_{13} \\ a_{20} & a_{21} & a_{22} & a_{23} \\ a_{30} & a_{31} & a_{32} & a_{33} \end{bmatrix} \times \begin{bmatrix} r_{00} & r_{01} & r_{02} & r_{03} \\ r_{10} & r_{11} & r_{12} & r_{13} \\ r_{20} & r_{21} & r_{22} & r_{23} \\ r_{30} & r_{31} & r_{32} & r_{33} \end{bmatrix}$$

Once we obtain X , we calibrate any set of measured currents from C2, such as the data shown in Figure 7a, using the linear transformation operator defined by Equation 2:

$$N_1 = T_1 * X \quad (2)$$

$$\begin{bmatrix} n_0 & n_1 & n_2 & n_3 \end{bmatrix} = \begin{bmatrix} t_0 & t_1 & t_2 & t_3 \end{bmatrix} \times \begin{bmatrix} x_{00} & x_{01} & x_{02} & x_{03} \\ x_{10} & x_{11} & x_{12} & x_{13} \\ x_{20} & x_{21} & x_{22} & x_{23} \\ x_{30} & x_{31} & x_{32} & x_{33} \end{bmatrix}$$

Figure 7b shows the transformation result. The C2 clusters have been linearly displaced to locations corresponding to the C1 clusters. We derive prediction ellipses at 3 sigma using the combined data sets shown

in the figure. The area enclosed by the new ellipses in Figure 7b is similar to that shown for C1 in Figure 7a. This is a desirable characteristic because defect detection sensitivity strongly correlates to the dispersion level in the defect-free data.

Defect current and background leakage in test chips

An important advantage of measuring local currents at each supply port is the increased observability this method provides. Here, *observability* refers to the ability to distinguish between current drawn by a defect and normal background leakage current. For example, if a chip contains 100 power ports, the leakage current through any one V_{DD} port will be approximately 100 times smaller, on average, than the leakage current measured globally. Therefore, measuring local currents greatly increases the probability of detecting a defect, particularly in large chips with many V_{DD} ports.

The chips used in our research contain four or six V_{DD} ports, depending on the configuration. Although we can't reveal the actual magnitudes of the defect and leakage currents, we report the ratio of defect current to leakage current in Table 1. The ratio demonstrates the claim that measuring local currents enhances defect observability.

Table 1 lists mean values of the ratios of total defect current to global background leakage for 18 DESM voltage levels. We measured the defect currents used in the ratios' numerators through the DESM as follows: First, we turned off all defect emulation transistors and recorded the DESM current. This represents the leakage current through the defect emulation transistors at a particular DESM voltage. Second, we measured the DESM current as each defect emulation transistor in the 100 TCs was enabled, one at a time, and subtracted the leakage value. The difference yields only the portion of DESM current that passes through the enabled defect emulation transistor, which is the value we are interested in. The mean value used in the numerator is the average across the 100 TCs and the 12 chips. The denominator in each ratio is the average global leakage value (all defect emulation transistors turned off) as measured through the GCSM across the 12 chips.

The range of mean ratios varies from 0.72 (harder to detect) to 2.94 (easier to detect). At first glance, it appears that larger mean ratios should be easy to detect with traditional I_{DDQ} methods; for example, the ratio 2.94 is nearly 3-to-1 defect current to leakage current. However, in the context of a larger chip, this would not

be the case. To illustrate this, the third column of Table 1 lists the projected ratios for a chip approximately 1 cm² in size.

The hypothetical 1-cm² chip contains 468 copies of the TC array and contains a 27 × 19 power port area array. As a conservative estimate, the ratios given in the second column are scaled in the third column by 468/4 = 117. The factor of 4 provides an allowance for redistribution of defect and leakage current in the larger array that adversely affects detection sensitivity. (We chose the value 4 assuming that redistribution causes defect current to be reduced by half and leakage current to double over that measured in a chip with one copy of the array, as we do here.) Under these assumptions, the smallest projected ratio is 0.006, as given in the first row of the table for the DESM voltage of 0.85. This indicates that the defect current measured in our experiments would be more than 160 times smaller than the leakage current in the hypothetical chip. In specific cases in our experiments, we could detect defect current in ratios smaller than the values shown in the table by a factor of 4, and we believe this factor could be as large as 10 with improved signal-to-noise ratios.

Emulated-defect detection results

This section presents the results of applying regression and ellipse analysis to the I_{DDQ} data collected from 12 chips, as we described in the previous section. We report the detection sensitivity as well as the number of test escapes and yield loss for several pairing subsets.

Data analysis

We formulated our defect detection procedures on the analysis of scatter plots, such as those shown in Figures 8a and 8b for regression and ellipse analysis, respectively. Under the four- V_{DD} configuration, local currents from any one chip and experiment can be paired in six distinct combinations, given as V_{00} - V_{02} , V_{00} - V_{12} , and so on, as described earlier.

Table 1. Ratios of defect current to leakage current in the test chips.

DESM (V)	Mean I_{def}/I_{leak} ratio	Projection for a 1-cm ² chip
0.85	0.72	0.006
0.80	1.33	0.011
0.75	1.83	0.016
0.70	2.24	0.019
0.65	2.55	0.022
0.60	2.75	0.024
0.55	2.89	0.025
0.50	2.94	0.025
0.45	2.93	0.025
0.40	2.85	0.024
0.35	2.72	0.023
0.30	2.55	0.022
0.25	2.34	0.020
0.20	2.11	0.018
0.15	1.87	0.016
0.10	1.64	0.014
0.05	1.42	0.012
0.00	1.21	0.010

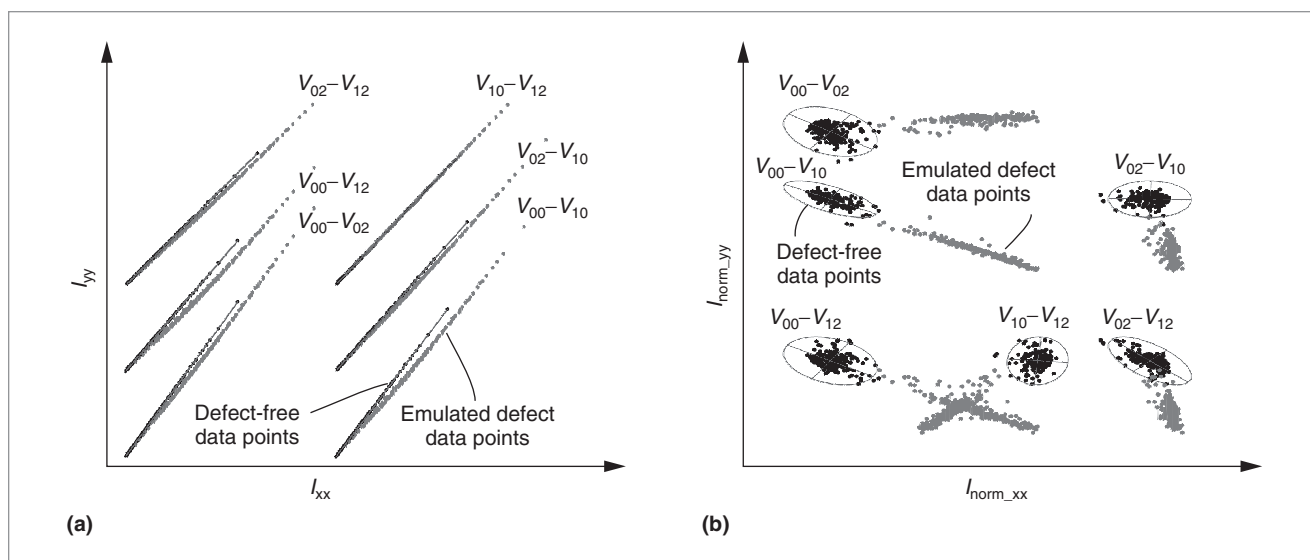


Figure 8. Regression (a) and ellipse (b) analysis of defect #0, four- V_{DD} configuration.

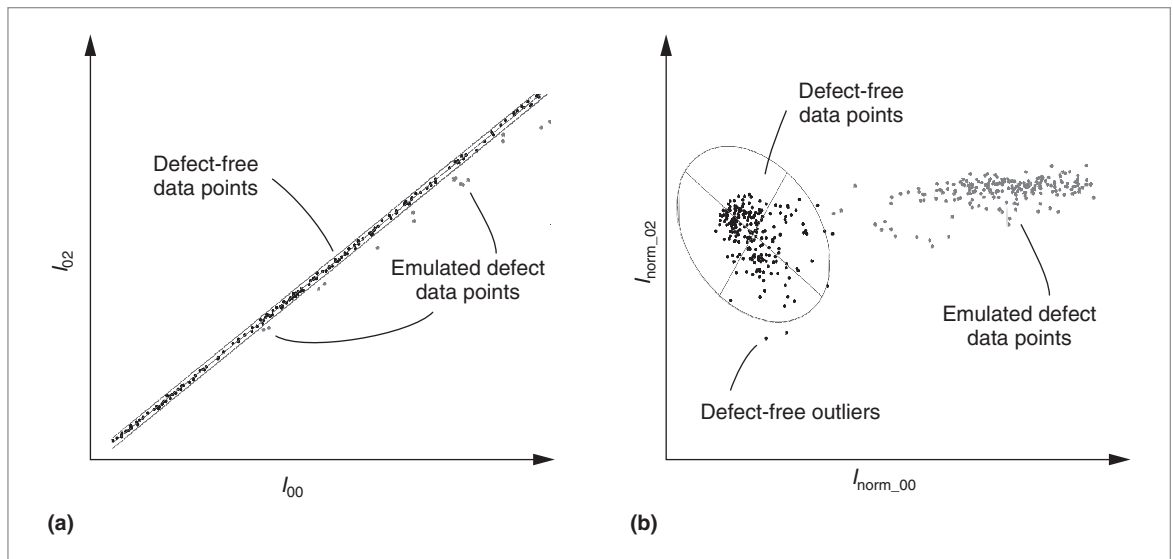


Figure 9. Blowup of regression (a) and ellipse (b) analysis of V_{00} - V_{02} in Figure 8.

We derived the scatter plots shown in Figure 8 from experiments that investigate defect #0, as shown at position $(x, y) = (5, 21)$ in Figure 4. We drew the defect-free data used to compute the prediction limits in Figure 8a and the prediction ellipses in Figure 8b from 12 chips at each of 19 DESM voltages for a total of 228 data points. The emulated-defect data also comes from the 12 chips at each of 18 DESM voltages for a total of 216 data points. (The same defect-free data and limits are used in the analysis of all 100 emulated defects.)

The scatter plots for each of the six pairings in Figure 8a are offset along the x and y axes to assist the visual presentation of the data. The prediction limits in Figure 8a are extremely narrow and appear as a single line. Figure 9a shows a blowup of the lower left corner of scatter plot pairing V_{00} - V_{02} . Here, it is clearer that the defect-free data points are within the two hyperbolas defining the 3-sigma prediction limits, and the emulated-defect data points are outside the limits.

Figure 9b shows an enlarged view of the data from pairing V_{00} - V_{02} in Figure 8b. In contrast to the regression analysis, some defect-free data points, labeled “defect-free outliers,” fall outside the 3-sigma prediction ellipse. The chips represented by these data points correspond to yield loss, as discussed later. Similar to the regression analysis, all the emulated-defect data points for this pairing fall outside the prediction ellipse.

Given the position of the defect in the lower left quadrant shown in Figure 4, we expect emulated-defect #0 to draw a larger fraction of its total current from supply port V_{00} . Therefore, we expect scatter plots that

include V_{00} as an element of the pairing to yield a higher confidence level for detecting the emulated defect. This is revealed in Figures 8a and 8b as a more distinct separation of the defect-free data and emulated-defect data, particularly for pairings V_{00} - V_{10} and V_{00} - V_{12} .

Detection criterion

Our criteria for deciding whether an emulated defect is detected is the same for either regression or ellipse analysis and is given as follows: We use the defect-free data points to derive 3-sigma prediction limits and prediction ellipses for each pairing considered in the analysis. We consider a defect detected if one or more of its data points within any pairing falls outside the prediction limits (for regression analysis) or the prediction ellipse (for ellipse analysis). It follows that an emulated defect has been missed—that is, is a test escape—if all its data points fall within the limits across all pairings. Under this criterion, it is straightforward to conclude that increasing the number of pairings increases the chances of detecting the emulated defect. However, increasing the number of pairings also increases the chance that defect-free data points fall outside the 3-sigma limits and contribute to yield loss.

To evaluate the impact of the number of pairings, we performed the analysis over several pairing sets. As described earlier, the number of pairings possible under the four- and six- V_{DD} configurations is six and 15, respectively. In addition to the analysis over these sets, Figure 5 identifies two subsets analyzed as orthogonal neighbors for the four- V_{DD} configuration, and orthogonal and cross neighbors for the six- V_{DD} configuration.

Table 2. Test escape (maximum 21,600) and yield loss (maximum 228) results.

Analysis	Four-VDD configuration				Six-VDD configuration			
	Orthogonal neighbors (4 pairings)		All pairings (6 pairings)		Orthogonal and cross pairings (11 pairings)		All pairings (15 pairings)	
	Test escapes	Yield loss	Test escapes	Yield loss	Test escapes	Yield loss	Test escapes	Yield loss
Regression	169	1	128	2	197	2	124	4
Ellipse	1,195	11	1,149	11	1,526	21	1,505	22

Test escape and yield loss analysis using all defect-free data

Our regression and ellipse analysis used defect-free and emulated-defect data sets from 12 chips. As discussed, we can “artificially” increase the sample size of the defect-free data sets by considering the tests performed at each of the 19 distinct DESM voltages as a separate chip. Under these conditions, the number of defect-free data sets increases from 12 to 228.

For regression analysis, the inclusion of the additional data sets improves defect sensitivity because the wider range of leakage currents produced across the different DESM voltages increases the data point spread along the *x*-axis in the scatter plots. As long as dispersion is small around the regression line, this characteristic keeps the prediction limits small along the regression line’s entire length. Figure 9a, in which the hyperbola-shaped prediction limits actually appear straight and parallel with the regression line, illustrates this characteristic.

In contrast, for ellipse analysis, including the additional defect-free data sets weakens the method’s defect sensitivity, particularly when we include data sets associated with lower signal-to-noise ratios—those associated with the larger DESM voltages. This is true because the dispersion captured by the ellipse is defined by the worst-case dispersion across the entire range of values measured. The larger dispersion added by the lower defect-free currents adversely affects the detection sensitivity of defects that produce larger currents. This does not happen with regression, because separate limits are defined and preserved along the *x*-axis for each current level in the defect-free data.

Table 2 gives the results of regression and ellipse

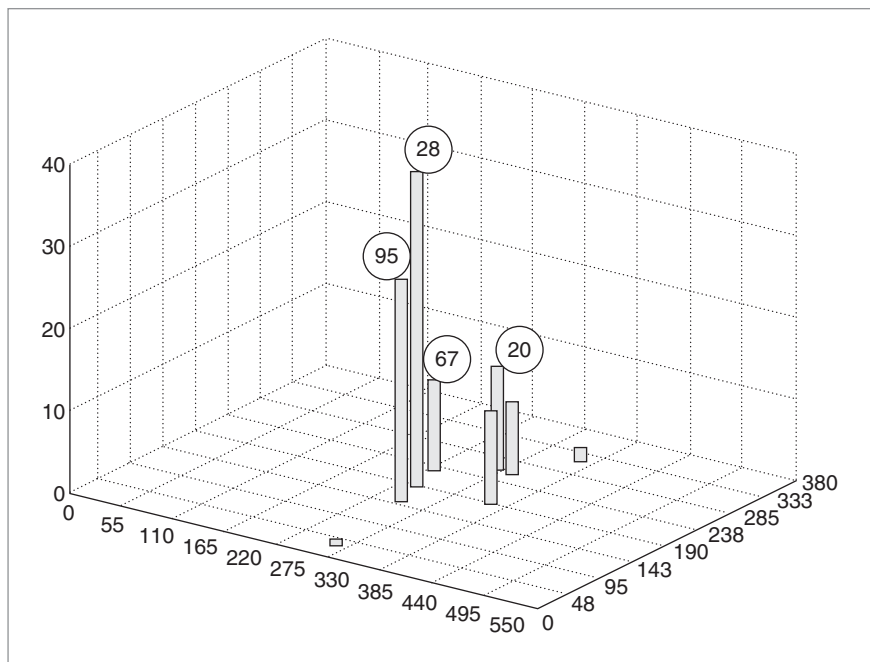


Figure 10. Location of test escapes for regression analysis under four-V_{DD} configuration, all-pairings set.

analysis using the pairing subsets described in the previous section under the four- and six-V_{DD} configurations. The pair of numbers reported in each cell of the table corresponds to the number of emulated defects that were not detected (test escapes) and the number of defect-free chips that failed the test (yield loss), respectively. As indicated previously, we tested 21,600 emulated defects and 228 defect-free chips. The ideal result is 0/0: no test escapes and no yield loss.

The regression analysis results are clearly superior to the ellipse analysis results. Overall, regression performs very well—for example, missing less than 1 percent of the emulated defects and having less than 1 percent yield loss under the all-pairings four-V_{DD} configuration. Figure 10 portrays test escape distribution in a 3D plot

in which the (x, y) plane represents the TC array, as shown in Figure 4. The bar height indicates the number of times that the emulated defect at that position was missed. Clearly, the bars are 0 in all cases except for the emulated defects in the center of the grid. The circled defect number identifiers (defined in Figure 4) above four of the bars are the emulated defects most often missed. (Because each defect is tested at 18 DESM voltages on 12 chips, the maximum z value for any defect is 216.) Moreover, these are the only defects missed at DESM voltages smaller than 0.85 V.

All the emulated defects missed under either regression or ellipse analysis are located in the center of the power grid. This region distributes the emulated defect's current almost equally among surrounding V_{DD} ports. The lack of an anomalous local current variation cannot be distinguished from global changes in background leakage current. However, in larger chips that incorporate more V_{DD} ports, the center portion of this grid would be asymmetric to V_{DD} ports outside this region. Therefore, the defects missed here may be detectable in pairings involving V_{DD} s outside this region in larger grids. We are currently conducting experiments that apply power to other subsets of V_{DD} ports (for example, V_{01} , V_{02} , and V_{12}) to determine whether it is possible to detect the defects missed under the four- and six- V_{DD} configurations investigated here.

The results shown in Table 2 for regression under the four- and six- V_{DD} configurations don't suggest that one is significantly better than the other. For example, the number of test escapes and yield loss are 128 and 2 versus 124 and 4, respectively. In contrast, the differences are more significant for ellipse analysis, with 1,149 and 11 versus 1,505 and 22. As discussed earlier, the noise in global currents adversely affects ellipse analysis. The noise effect is more pronounced under the six- V_{DD} configuration because local currents are smaller and closer to the noise floor than currents measured under the four- V_{DD} configuration.

It is clear under either analysis that larger subsets of pairings reduce the number of test escapes at the expense of increasing yield loss. For example, under the regression results in the second and third columns in Table 2, test escapes decrease from 169 for the orthogonal-neighbors subset to 128 for the all-pairings set, and yield loss increases correspondingly from 1 to 2. (Note that for the subsets in columns two and four, the number of test escapes can never be smaller and the yield loss can never be larger than the all-pairings set, because the latter includes the subset pairings in its

analysis.) Given the detection criteria, we expect a decrease in test escapes as more pairings are included in the analysis. The yield loss increase is caused by the larger magnitude of within-chip process variations in pairings more widely separated in the layout. The cross-correlation profile of defect-free data from nonneighboring V_{DD} ports has a higher dispersion level, which increases the chance that defect-free data points will become outliers.

Test escape and yield loss analysis using a subset of defect-free data

The analysis given in the preceding subsection indicates that regression analysis is superior in terms of reducing test escapes and yield loss. However, ellipse analysis performed under a special condition produces the best overall result with 71 test escapes and no yield loss. The special condition involves restricting the defect-free samples used to derive the prediction ellipses to those collected with the highest background leakage—that is, with the DESM voltage set to 0.0 V. Although this constraint is difficult to realize in production testing, it demonstrates an important advantage of ellipse analysis over regression for situations in which we can obtain good signal-to-noise ratios.

The discussion of Figure 9a indicated that the hyperbola prediction limits appear as straight lines because the distribution of defect-free currents is large. The distribution of the 228 data points along the x -axis keeps the prediction limits close to the regression line's entire length. If this did not hold true—if the distribution of defect-free data were clustered in a small region along the x -axis—the curves' hyperbolic nature would be far more pronounced, and the prediction limits would widen significantly around the regression line on either side of the cluster. This would reduce the method's sensitivity for emulated defects that produce data points on either side of the cluster. Therefore, regression works well in our experiments because the sum of the emulated-defect and leakage currents is of the same order as the leakage currents alone.

As discussed earlier, ellipse analysis is more sensitive to noise than regression because both the local and global current measurements that define the data point's position possess a noise element. Under the assumption that the noise floor is independent of the current's magnitude, it follows that the adverse impact of noise is smaller for larger-magnitude currents. Therefore, prediction ellipses derived using larger defect-free currents will be smaller.

Figure 11 illustrates these properties. The defect-free data is partitioned into 19 groups of 12 data sets. Each group consists of the currents measured on the 12 chips at a particular DESM voltage. We derived the prediction limits and ellipses independently for each of the 12-point data sets, and we performed regression and ellipse analysis. We derived the results shown in Figure 11 using the data from the six- V_{DD} configuration under the all-pairings set. (The same trend is present in the other pairing sets in Figure 5.) The x -axis gives the DESM voltage applied during the collection of defect-free data. The y -axis gives the number of test escapes out of 21,600 emulated defects. (Yield loss is 0 for all experiments using either regression or ellipse analysis.)

The downward trend from left to right in both curves indicates that detection sensitivity increases for both regression and ellipse analysis as background leakage current increases in defect-free data sets. However, we obtain the best result for regression when all defect-free data is used to derive the prediction limits, as indicated by the horizontal line labeled 124—the value listed in the last “Test escapes” column of Table 2. In contrast, the number of test escapes for ellipse analysis are fewer than 1,505 for all cases except DESM voltages of 0.85 and 0.75. Interestingly, the number of test escapes for ellipse analysis becomes less than that for regression using any combination of defect-free data for DESM voltages of 0.15, 0.10, and 0.0. The best result overall, 71, occurs at DESM voltage 0.0.

Residual analysis

The analysis just described provides a summary of the two methods’ overall effectiveness in terms of misses and yield loss. However, it does not give information about the confidence level associated with detected emulated defects. In our analysis, we express the confidence level numerically as the distance between the data point and the closest prediction limit, and we call this a residual. It follows that positive residuals correspond to data points that fall outside the limits, whereas negative residuals are associated with data points that fall within the limits.

Residuals are quantities that vary widely and depend on the underlying data’s magnitudes. Therefore, comparing them directly is not meaningful. The textbook approach to creating a common ground on which residuals from different experiments can be compared is to convert them to standardized quantities. We can convert a residual to a standardized residual, $ZRES$, using the following equation:

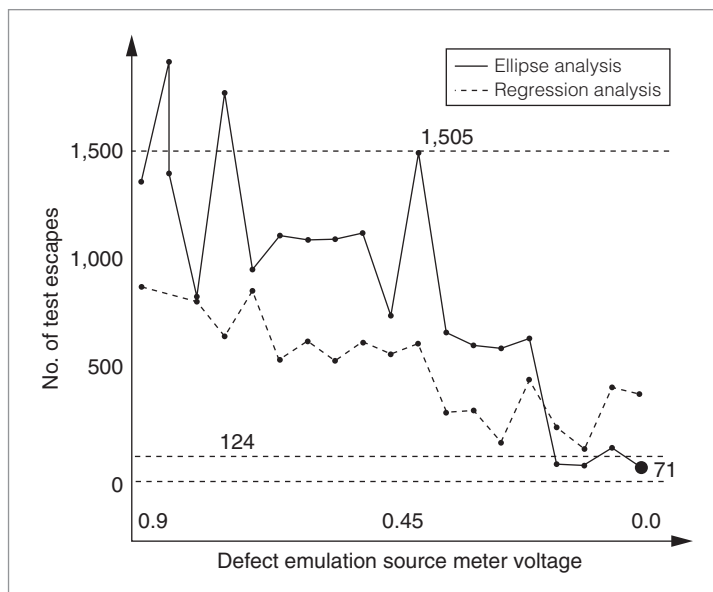


Figure 11. Number of test escapes for six- V_{DD} configuration, all-pairing set with defect-free set drawn from 12 chips at each DESM voltage.

$$ZRES = \frac{\text{residual}}{\sqrt{MSE}}$$

Here, MSE represents the mean square error or the variance of defect-free data points from their respective means. For regression, the mean for a particular value of x is the regression line y value. For ellipse analysis, the mean is the (x, y) center of the ellipse.

Given the definition of *residual*, it follows that the confidence level associated with detection of an emulated defect corresponds to the largest positive residual computed across all scatter plots used in the analysis. This process selects one residual for each emulated defect. However, the large number of emulated defects in our experiments requires another level of data compression to display the results meaningfully. The focus here is to determine the worst-case confidence—the smallest maximum residual, across the 12 chips for each of the 100 emulated defects. Intuitively, the smallest maximum residual for each emulated defect would occur when the defect is tested at a DESM voltage of 0.85 V, when the defect currents are smallest. Therefore, we describe the results at this DESM voltage.

Figure 12 shows the plotted standardized residuals from this subset of data for regression and ellipse analysis under the four- V_{DD} configuration, all-pairings set. The x -axis plots the defect number, sorted according to the distance of the emulated defect from the nearest V_{DD}

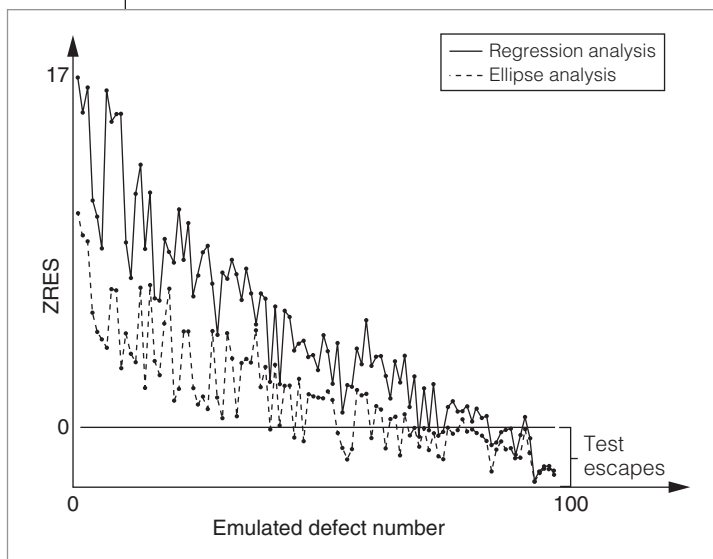


Figure 12. Residual analysis of four- V_{DD} configuration, all-pairings set, with DESM voltage at 0.85 V.

port. In other words, emulated defects such as 28 and 95, as given in Figure 4, are plotted in the rightmost positions on the x -axis because they are in the middle of the grid and furthest from the V_{DD} ports. The y -axis plots the standardized residuals, $ZRES$. The zero line plotted horizontally represents the prediction limits. Emulated defects with $ZRES$ values less than 0 are test escapes.

Several features are notable in the data. First, the ellipse curve is similar in shape to the regression curve, indicating that the same information is captured by either technique. But the ellipse curve appears shifted downward with respect to the regression curve, most prominently on the left side. The shift downward causes a larger number of data points to fall below the zero line. This feature is consistent with the larger number of test escapes reported for ellipse analysis in Table 2. Second, under regression analysis, the $ZRES$ magnitudes reach a maximum of approximately 17 for emulated defects close to a V_{DD} port. This reflects a high level of confidence that the emulated defect is present. Both curves steadily decrease from right to left, reflecting decreasing sensitivity for emulated defects closer to the center of the grid.

Our positive results for applying QSA for the detection of emulated defects suggests that it might be possible to continue using I_{DDQ} testing in advanced technologies, where leakage trends are expected to further erode the detection sensitivity of I_{DDQ} as currently practiced. Future work will focus on the hardware verification of QSA in large industrial designs and on chips with actual defects. ■

Acknowledgments

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

References

1. *International Technology Roadmap for Semiconductors*, 2005, <http://www.itrs.net/Common/2005ITRS/Home2005.htm>.
2. Y. Ouyang and J. Plusquellic, "IC Diagnosis Using Multiple Supply Pad I_{DDQ} s," *IEEE Design & Test*, vol. 18, no. 1, Jan.-Feb. 2001, pp. 50-61.
3. C. Patel and J. Plusquellic, "A Process and Technology-Tolerant I_{DDQ} Method for IC Diagnosis," *Proc. 19th VLSI Test Symp. (VTS 01)*, IEEE Press, 2001, pp. 145-150.
4. C. Patel et al., "A Current Ratio Model for Defect Diagnosis Using Quiescent Signal Analysis," *IEEE Int'l Workshop Current and Defect Based Testing (DBT 02)*, 2002; [http://domino.research.ibm.com/acas/w3www_acas.nsf/images/projects_01.02/\\$FILE/01plus.pdf](http://domino.research.ibm.com/acas/w3www_acas.nsf/images/projects_01.02/$FILE/01plus.pdf).
5. C. Patel et al., "Defect Diagnosis Using a Current Ratio Based Quiescent Signal Analysis Model for Commercial Power Grids," *J. Electronic Testing: Theory and Applications*, vol. 19, no. 6, Dec. 2003, pp. 611-623.
6. C. Patel, A. Singh, and J. Plusquellic, "Defect Detection under Realistic Leakage Models Using Multiple I_{DDQ} Measurements," *Proc. Int'l Test Conf. (ITC 04)*, IEEE Press, 2004, pp. 319-328.
7. D. Acharyya and J. Plusquellic, "Hardware Results Demonstrating Defect Detection Using Power Supply Signal Measurements," *Proc. 23rd VLSI Test Symp. (VTS 05)*, IEEE Press, 2005, pp. 433-438.
8. D. Acharyya and J. Plusquellic, "Hardware Results Demonstrating Defect Localization Using Power Supply Signal Measurements," *Proc. 30th Int'l Symp. Testing and Failure Analysis (ISTFA 04)*, ASM Int'l, 2004, pp. 58-66.



Jim Plusquellic is an associate professor of computer engineering at the University of Maryland, Baltimore. His research interests include defect-based testing; digital, mixed-signal, and analog VLSI design; and test structure design for process variability. Plusquellic has an MS and a PhD in computer science from the University of Pittsburgh. He is a member of the IEEE.

Dhruva Acharyya is an intern at IBM Austin Research Labs and a PhD candidate at the University of Maryland, Baltimore. His research interests include power supply testing and test structure design for measuring process variability. Acharyya has a MS in computer engineering the University of Maryland, Baltimore. He is a member of IEEE.

Abhishek Singh is a test engineer at nVidia. His research interests include power supply testing and fault simulation techniques. Abhishek has a PhD in computer engineering from the University of Maryland, Baltimore. He is a member of IEEE.



Mohammad Tehranipoor is an assistant professor of electrical and computer engineering at the University of Maryland, Baltimore. His research interests include CAD and test for CMOS VLSI designs and emerging nanoscale devices. Tehranipoor has a PhD in electrical engineering from the University of Texas at Dallas. He is a member of the IEEE Computer Society, the ACM, and ACM SIGDA.



Chintan Patel is a research assistant professor at the University of Maryland, Baltimore. His research interests include power supply testing and power supply monitor design. Chintan has a PhD in computer engineering from the University of Maryland, Baltimore. He is a member of IEEE.

■ Direct questions and comments about this article to Jim Plusquellic, Dept. of CSEE, Univ. of Maryland, Baltimore, MD 21250; plusquel@umbc.edu.

For further information on this or any other computing topic, visit our Digital Library at <http://www.computer.org/publications/dlib>.