

# Fault Simulation Model for $i_{DDT}$ Testing: An Investigation

Abhishek Singh, Chintan Patel and Jim Plusquellic

Department of Computer Engineering, University of Maryland, Baltimore County

## Abstract

*In today's technologies, resistive shorting and open defects are becoming more predominant. Conventional fault models, and tools based on these models are getting inadequate in addressing these new failure mechanisms. In prior works  $i_{DDT}$  testing techniques have been shown to be sensitive to such subtle resistive defects. Expensive transient simulations are required to perform ATPG and Fault simulation for  $i_{DDT}$  testing due to the analog nature of the test observable. In this paper we investigate a fault simulation model that may enable the development of tools required for the integration of  $i_{DDT}$  test into existing production test flows. We propose a partitioning scheme for the CUT (chip under test) and path isolation strategy for core-logic as a means of reducing the computational complexity involved in deriving  $i_{DDT}$  signals. More specifically, a unique Impulse Response based method is derived to eliminate the need for transient simulations of the entire CUT.*

## 1.0 Introduction

Existing fault models are becoming increasingly ineffective in modeling defect mechanisms in UDSM technologies. For example, the change from a subtractive aluminum process to damascene Cu may lead to more particle-related blocked-etch resistive opens [1]. Also, technology scaling increases the probability of resistive vias caused by incomplete etch. This suggests the importance or need of screening mechanisms that can target such resistive type of defects.

Previous research on  $i_{DDT}$  test, such as [2-8], [11-12], shows that changes in the circuit configuration caused by a defect, manifests itself as anomalies in the transient signals measured at the power supply ports. This provides test methods based on  $i_{DDT}$  analysis a benefit of global observability due to the global nature the power grid. Although the concepts proposed in this paper are applicable to  $i_{DDT}$  test methods in general, we are using Transient Signal Analysis (TSA) as a vehicle for screening defects in the proposed fault simulation flow. TSA is an  $i_{DDT}$  test technique based on cross correlation of multiple  $i_{DDT}$  signals measured at power supply ports in a CUT.

Integration of TSA or other  $i_{DDT}$  test techniques into existing production test flows requires tools for performing ATPG and fault simulations. A fault simulation engine, for such techniques, requires generation of  $i_{DDT}$  signals that are processed using the defect screening procedure. This implies the need for performing transient simulations on the entire CUT. The memory and time requirements of such simulations are prohibitive. In this paper we propose a

model that can be used to implement a practical fault simulator for  $i_{DDT}$  testing. This is achieved by decomposing the CUT into two systems and individually analyzing the means of reducing the simulation complexity of these systems. The two systems involve, 1) a linear constituent, namely, the power grid and 2) a nonlinear core logic circuit. A unique method based on impulse response (IR) and convolution is proposed for the linear power grid component. This method allows simulation-less computation of the power grid response to transient inputs from the nonlinear core logic component. Furthermore, a path isolation scheme is proposed to address the simulation complexity of the nonlinear component.

## 2.0 Background

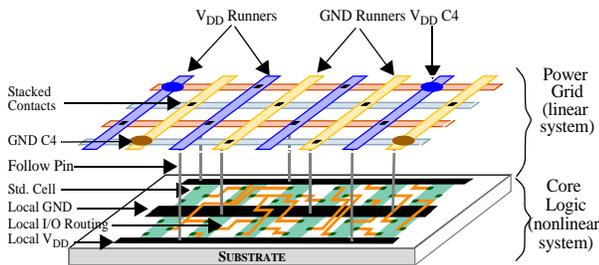
Testing methods based on the analysis of power supply transient signals are described in [2-6], [11-12] for digital circuits and in [7-8] for analog circuits. However, we have not come across any prior work that proposes models that can enable fault simulation of  $i_{DDT}$  test vectors. We believe that this work is inherently related to research in power verification and signal integrity analysis. This is due to the fact that both power verification and  $i_{DDT}$  fault simulation requires transient response of the CUT. The former analyzes dynamic IR,  $Ldi/dt$  drop, and package/on-chip resonance whereas, the latter processes these signals to identify anomalies caused by the defect.

Due to signal integrity problems caused by aggressively increasing device densities, simulation of transient power distribution in a chip has become an essential step in power verification. To meet this growing requirement several static and transient simulation techniques have been proposed in the past decade. Mathematical tools that speed-up the power grid simulations, such as [15], [18-20] and methods based on Transient Current Simulation of logic circuits [9], [13] and [16] have been proposed. These allow derivation of transient currents drawn by logic circuits without running SPICE like transient simulations. Both power verification and fault simulations based on  $i_{DDT}$  test share similar needs of performing transient simulations of a CUT.

## 3.0 System Overview

Any digital CMOS chip can be modeled as a combination of two complementary electrical systems, a linear RC system formed by the power ( $V_{DD}$  and GND) grid and a nonlinear RC-Transistor system formed by the underlying CMOS logic circuit. For example, Figure 1 shows a portion

of a typical row based standard-cell design. The upper half



**Figure 1. Linear and non-linear system representation of a chip.**

of the figure depicts the power grid, laid out as a stack of uniformly spaced metal runners at alternating layers connected using stacked contacts. For simplicity only a two layer structure is shown. The parasitic resistance and capacitance of the power grid along with the on-chip decoupling capacitors form a linear RC system. This linear system is referred to as *Power Grid Circuit (PGC)*.

The lower half of the figure shows the core logic comprising the standard-cells (MOS-transistors) and the I/O routing of the chip. Since, MOS-transistors are inherently non-linear devices this portion forms a non-linear electrical system comprising RC and transistor elements, referred to as *Core Logic Circuit (CLC)*. The local  $V_{DD}$  and GND runners on the standard-cells are connected to the global power grid at various points through special nets/vias called *follow pins*, as indicated in the figure.

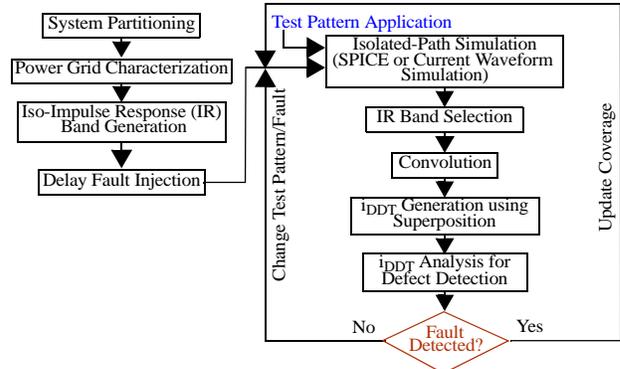
A transition sequence applied at the primary inputs (or outputs of scan latches) causes the gates along a sensitized path to switch in a temporal sequence. Each switching gate draws transient current ( $i_{DS}$ ) from the power grid, sourced by the external power supply pads or C4s. The PGC transforms these  $i_{DS}$  signals into composite current transients ( $i_{DDT}$ ) that are measured at the C4s. Therefore, the CUT can be modeled as a cascade of two electrical systems wherein, the outputs the CLC feeds the inputs of PGC. The transient current drawn by a CMOS gate is modeled as a PWL current source connected between  $V_{DD}$  and GND grid. The metal layer at which the CUT is partitioned into the above two systems, determines the sparseness of the input and output ports in the PGC and CLC respectively.

#### 4.0 Fault Simulation Flow for $i_{DDT}$ Testing

The fault simulation procedure for  $i_{DDT}$  testing can be described using a flow diagram shown in Figure 2. The subsequent sections of this paper elaborate on each part of the flow diagram.

**Step1: System Partitioning:** The most important step in the flow is the partitioning of the CUT into a linear (PGC) and a non-linear (CLC) electrical system.

**Step2: Power Grid Characterization:** Power grid characterization involves the derivation of IR functions between each input-output port of the power grid. The IR functions, by definition, completely characterize the power



**Figure 2. Flow diagram of the fault simulation procedure for  $i_{DDT}$  testing.**

grid and can be used to derive its response to any arbitrary input using convolution.

**Step3: Generation of Iso-IR Bands:** An iso-IR band defines a physical region in the grid layout consisting of input locations that generate response at a output within a user defined margin. Such a categorization helps in significantly reducing the number of convolution operations required to generate the power grid response.

**Step4: Fault Injection:** A fault is injected between two nodes in the layout of the CUT by introducing resistive connection modeling resistive short/bridge or open. The advantages of using layout information are 1) the fault model closely represents the physical defects and 2) only physically adjacent nets are considered as short/bridge candidates. We believe that this fault injection scheme closely follows the defect based test paradigm.

**Step5: Isolated-Path Transient Simulation:** An isolated path corresponds to a sensitized path that is physically separated from the chip layout by breaking connections with the unsensitized logic and the power grid, however, preserving the fanout loads and the I/O connections of the gates within the sensitized path. The transient current waveform generated by the core logic at each input location in the CLC can be derived through SPICE or time-domain Current Waveform Simulations of these isolated paths.

**Step6: Iso-IR Band Selection:** Based on the physical location of the input current source, the iso-IR band corresponding to that input is identified.

**Step7: Convolution:** The  $i_{DDT}$  or  $i_{DS}$  waveforms measured at specific nodes during the isolated path simulations are convolved with the IR functions of the iso-IR band selected in the previous step.

**Step8: Linear Superposition:** By the virtue of superposition property of linear time invariant (LTI) systems the power grid response to individual current input source can be combined through linear superposition to derive the overall response of the grid at a given C4. This property allows us to simulate all the sensitized paths independently. Another major advantage of this property is to limit the maximum number of convolution operations required to derive the overall response of the PGC. This is detailed further in the paper.

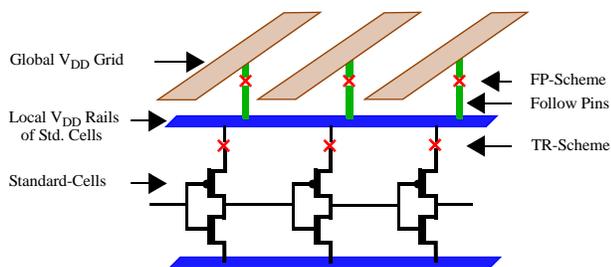
**Step9:  $i_{DDT}$  analysis for fault detection:** The  $i_{DDT}$ s thus obtained at the C4s, includes the current drawn by the defect and the defect-free logic. These  $i_{DDT}$ s can be processed using the detection criteria of any given  $i_{DDT}$  test

technique.

It is emphasized that, the outcome of the  $i_{DDT}$  screening procedure can be used a) to determine the  $i_{DDT}$ -test coverage of a given set of test patterns, b) to guide the ATPG algorithm to generate vectors that can enhance the detectability of the defect and c) to determine the range of resistance values for which a resistive short, open or bridging defect is detectable.

## 5.0 System Partitioning

The CUT is decomposed into a linear and non-linear system by breaking the physical connections between the power grid and the core logic cells. The metal layer at which the connections are broken defines the partitioning scheme. For example, Figure 3 shows two possible partitioning schemes. In one case the partitioning is done at the



**Figure 3. Partitioning schemes of a chip into linear and non-linear systems.**

follow pins whereas, in the second case the partitioning is done at the nodes where MOS transistors in each standard-cell connects to the local  $V_{DD}/GND$  rails. The “cross” symbol indicates the nodes at which the connections are broken in each scheme. The same is true for GND grid not shown in the figure. In the first case, referred to as FP-scheme (for follow-pins), the linear system (PGC) consist of the global power-grid and the nonlinear system (CLC) consists of the standard-cells, including their local  $V_{DD}/GND$  rails and the I/O routing. In the second case, referred to as the TR-scheme (for transistors), the PGC includes the global power grid as well as the local  $V_{DD}/GND$  rails in each row of the core logic and the CLC includes the standard-cells (without  $V_{DD}/GND$  rails) and the I/O routing. The locations (or nodes) at which the systems are partitioned represent the input-ports of the PGC and output-ports of the CLC. For the TR scheme, when the source nodes of more than one MOS-transistor are connected directly to the local  $V_{DD}/GND$  rails, which may be true for complex standard cells, any one node can be considered as an input. This is a reasonable assumption given that the impedance between any two points on the local  $V_{DD}$  (or GND) rail in a standard-cell is sufficiently small. The FP-scheme offers fewer number of input-output ports, reducing the complexity involved in the PGC characterization but may increase the complexity of core-logic simulation due to inclusion of local  $V_{DD}/$

GND metal rails in the CLC. In contrast, the TR-scheme offers larger number of input-output ports which increases the complexity of PGC characterization but simplifies the core-logic simulations. These trade-offs will become clear to the reader in subsequent sections.

## 6.0 Power Grid Circuit (PGC) Characterization

In the current technology the  $V_{DD}/GND$  routing can occupy more than 25% of the total routing area on a chip and thus consist of millions of linear elements. SPICE simulations on these grids can be very expensive and therefore, several techniques, such as [15], [18], [19] and [20], have been proposed for fast power grid simulation. Even these tools may prove infeasible as the base simulation engine for an  $i_{DDT}$  fault simulator, due to the large number of faults and test vectors. We demonstrate a convolution procedure based on the linearity property of the PGC that enables us to compute the power-grid response to the input switching transients from the core logic without running simulations.

Any linear time invariant (LTI) system can be completely characterized by its Impulse Response (IR) function denoted as  $h(t)$ . The impulse response  $h(t)$ , is the output of the system to a unit impulse function,  $\delta(t)$ . Once the IR of a linear system is known, we may construct the response of a the system to an arbitrary input signal as a sum of suitably delayed and scaled impulse responses. This process is called convolution and is mathematically described using Eq. 1. Here,  $f(t)$  is the input signal,  $g(t)$  is the output signal

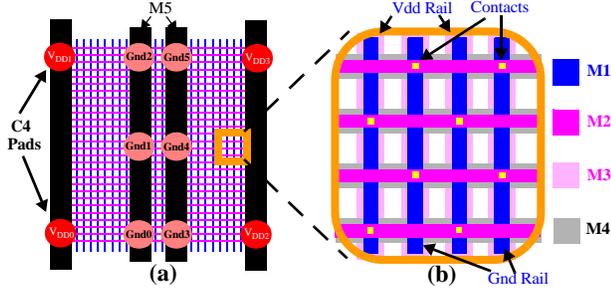
$$\begin{aligned} g(t) &= h(t) \otimes f(t) = \int_{-\infty}^{\infty} h(u)f(t-u)du & \text{Eq. 1.} \\ &= \int_{-\infty}^{\infty} h(t-u)f(u)du \end{aligned}$$

and  $h(t)$  is the IR function. The response of a linear system to an arbitrary input signal can thus be computed by convolution using the IR function in time domain.

The PGC represents a multi-input and multi-output linear system. Each input on the grid sees a different RC network to each of the outputs (C4s) therefore, there exists a unique IR function for each input-output pair, denoted as  $h_{ij}(t)$ , where  $i$  and  $j$  represents the input and output port of the PGC respectively. A set of such IR functions,  $h_{ij}(t)$ , can be used to characterize the PGC. Once the grid is characterized, its response to transient inputs can be determined by convolving the transients with the corresponding IR functions. Superposition and shift-invariance properties of a LTI system are used to determine the outputs due to 1) multiple switching gates within a sensitized path and 2) multiple sensitized paths under the same input sequence. This can eliminate the need to simulate millions of RC elements with a fixed number of multiplication and addition operations during the fault simulation.

## 7.0 Experimental Setup

Figure 4(a) shows a portion of the commercial power grid subsequently referred to as the Quad. The Quad occupies a 10,000 by 10,000 unit area and interfaces to a set of external power supplies through an area array of  $V_{DD}$  and GND C4 pads. As indicated in the figure, there are 4  $V_{DD}$  C4s and 6 GND C4s in this portion of the grid. Figure 4(b) shows that the grid is constructed over 4 layers of metal, with metal 1 (M1) and metal 3 (M3) running vertically and M2 and M4 running horizontally. The C4s are connected to wide runners of vertical M5, shown in 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. Stacked contacts are placed at the appropriate crossings of the horizontal and vertical rails.



**Figure 4.** The “Quad”: A portion of the commercial power grid used in the simulation experiments.

We derived an RC model of the Quad using an extraction script that preserves the physical structure of the metal interconnect in the topology of the RC network, i.e. no network reduction heuristics are applied. The resistance per square and the overlap capacitances per unit area of TSMC’s 0.25 $\mu\text{m}$  5 metal process used in the extraction process were obtained from published parameters by MOSIS [21].

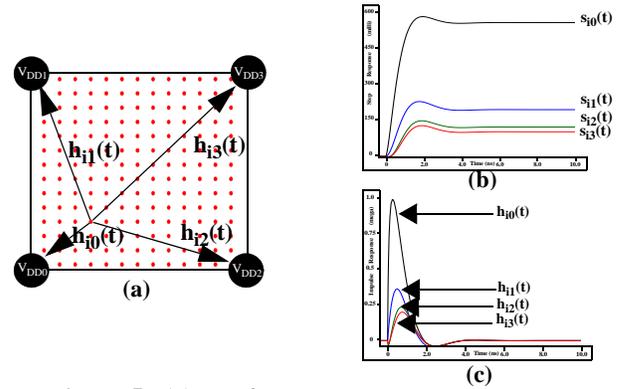
In this experiment we consider the Quad as the PGC and a custom designed 16-bit logarithmic adder as the CLC. We inserted  $\sim 3000$  labels at the M1-M2 crossovers in the grid layout to represent possible input locations on the PGC.

The impulse response from each of the input ports to the C4s are obtained by first stimulating the grid at each of the input ports with a unit step input. This gives the step response,  $s_{ij}(t)$ , of the grid with respect to input location  $i$  and output  $j$ . The impulse response,  $h_{ij}(t)$ , can be derived from the step response,  $s_{ij}(t)$ , using the differentiation in time domain. This is again due to the linearity property of the system and can be described using Eq. 2, where  $\delta(t)$  and  $u(t)$  represent unit impulse and step input functions, respectively and  $h(t)$  and  $s(t)$  represent the unit impulse and step response of the linear system, respectively.

Figure 5(a) shows a subset of the possible input loca-

$$\delta(t) = \frac{d}{dt}u(t) \Rightarrow h(t) = \frac{d}{dt}s(t) \quad \text{Eq. 2.}$$

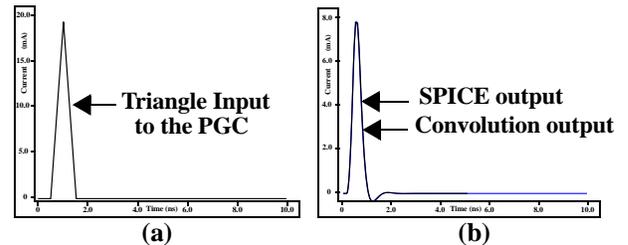
tions in the quad. It also indicates the IR functions that



**Figure 5.** (a) IRs from a sample current source location,  $i$ , to all the C4s. (b) Step response. (c) Impulse Response.

exists between a source location,  $i$ , to each of the four outputs. The step response curves,  $s_{ij}(t)$ , obtained using SPICE simulations are shown in Figure 5(b) and the IR functions obtained after the differentiation operation on  $s_{ij}(t)$  are shown in Figure 5(c).

To verify the convolution based procedure the PGC was stimulated with a triangle current source shown in Figure 6(a). The results obtained using the SPICE simulations are superimposed with the response obtained by convolving the triangle input with the IR function corresponding to that current source location in Figure 6(b). As shown in the



**Figure 6.** (a) Triangle input current to the PGC (b) Superimposed SPICE and Convolution responses.

figure the two curves are almost identical.

### 7.1 Iso-IR Contours

Eq. 3 gives the expression for convolution sum of discrete signals where  $g[i]$ ,  $h[i]$  and  $f[i]$  represents the output, IR and input of the linear system respectively. If  $f[i]$  is a N

$$g[i] = h[i] \otimes f[i] = \sum_{j=0}^{M-1} h[j]f[i-j] \quad \text{Eq. 3.}$$

$$= \sum_{j=0}^{N-1} f[j]h[i-j]$$

point signal,  $h[i]$  is a M point signal then  $g[i]$  is a N+M-1 point signal as given by the above equation. Eq. 3 shows

that the convolution based method requires  $N \cdot M$  additions and multiplications. Thus, the complexity involved in the computation of grid response is significantly reduced as compared to a SPICE simulation based approach that involves solving several partial differential equations.

Comparison of IR function curves from adjacent input locations suggests that their amplitude and shape characteristics vary slowly as a function of distance. Thus IR functions from adjacent input locations, within a user defined threshold, can be grouped into regions or bands. The threshold is chosen based on tolerable difference in the output waveforms obtained using Eq. 3. However, selection of a suitable threshold requires a means of quantifying the similarity between two waveforms in terms of their shape characteristics. This similarity analysis is performed using cross-correlation and auto-correlation operations.

Cross-correlation of two waveforms results into a third waveform, the amplitude of which indicates the degree of similarity between the two waveforms. Also, the location of its peak indicates the time-shift required in the second signal to obtain the maximum match with the first waveform. This is mathematically expressed by Eq. 4, where  $r_{xy}[i]$  represents cross-correlation of two waveforms  $x[i]$  and  $y[i]$ . When  $x[i]$  and  $y[i]$  are identical, the operation is

$$r_{xy}[i] = x[i] \oplus y[i] = \sum_{j=0}^{M-1} x[j]y[j-i] \quad \text{Eq. 4.}$$

$$a_x[i] = x[i] \oplus x[i] = \sum_{j=0}^{M-1} x[j]x[j-i]$$

termed as auto-correlation,  $a_x(t)$ . The peak value of an auto-correlation function provides us the maximum expected value for the degree of similarity in the IR functions.

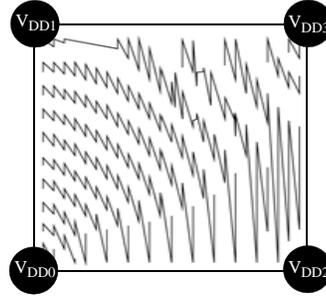
First, the input locations are sorted in an ascending order of their euclidian distance with respect to a reference location. The creation of a new iso-IR band begins with the selection of an representative location, referred to as *focus* of the iso-IR band, which is the first input location in the distance-sorted list and is not a part of any previously identified iso-IR band. The IR function of the focus is auto-correlated to obtain the maximum expected degree of similarity,  $\max(a_{f_i}(t))$ . The algorithm then searches the entire grid space to find all input locations that have a peak cross-correlation value not exceeding 5% (selected difference tolerance) of the peak auto-correlation value of the focus. This condition is given by Eq. 5.

$$\max \left\{ r_{f_i, h_j}(t) \right\} \leq 0.05 \cdot \max \left\{ a_{f_i}(t) \right\} \quad \text{Eq. 5.}$$

Where the quantity on the left represents the cross-correlation of the IR function of focus with the IR function of any other location  $h_j(t)$ . The quantity on the right represents the auto-correlation of focus's IR function.

Figure 7 shows the Quad with the iso-IR band limit

contours. A band is enclosed within two iso-IR limit contours. For the purpose of clarity only alternate contours are shown in the figure. Using the difference tolerance of 5%, totally 28 iso-IR contours were obtained. This shows a factor of 100 reduction in the number of IR functions (~3000 to ~30) required to generate the response of the PGC within the given difference tolerance.



**Figure 7. Iso-IR contours depicting the regions with similar impulse response at C4  $V_{DD0}$ .**

The categorization of IR functions into iso-IR bands reduces the maximum number of convolution operations to the total number of identified iso-IR bands. This is due to the superposition property of a linear system explained using Eq. 6. Where B represents the total number of iso-IR

$$y_{C40}[n] = \sum_{j=0}^B y_j[n] \quad \text{Eq. 6.}$$

$$y_j[n] = \sum_{i=0}^S h_i[n] \otimes f_i[n] = h[n] \otimes \sum_{i=0}^S f_i[n]$$

Where  $y_{C40}[n]$  is the overall response ( $i_{DDT}$ ) measured at C40,  $y_j[n]$  is response due the inputs in an iso-IR band (j) at C40,  $h_i[n]$  is the IR function from each input in an iso-IR band to C40,  $f_i[n]$  is the input signal inside an iso-IR band and  $h[n]$  represents the IR function of the focus.

bands on the grid and S represents the total number of input locations inside a given iso-IR band. As the equation suggests the output  $y_{C40}[n]$  is a linear superposition of responses due to each iso-IR band,  $y_j[n]$ . Ideally, computation of each  $y_j[n]$  requires S convolutions, however, due to creation of iso-IR bands the complexity involved in the computation of each  $y_j[n]$  reduces to a single convolution operation. This helps to reduce the total number of convolutions to B.

The accuracy of the iso-IR band based fault simulation procedure is verified using a full-custom designed 16-bit logarithmic adder as a representative of the sensitized portion in the CLC. Figure 8(a) shows the approximate location of the adder in the lower left corner of the PGC. The layout for the 16-bit logarithmic adder is shown in Figure 8(b). The gates in the layout consist of transistors with W/L ratios ranging from 2 to 5 for NMOS and 3 to 7 for PMOS (however, most are minimum size). The power rails of the

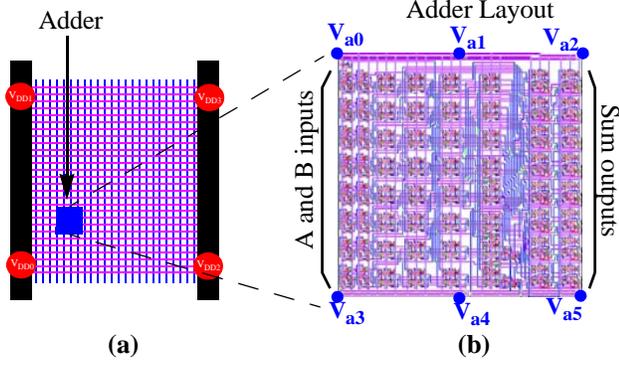


Figure 8. (a) A 16-bit logarithmic adder connected to the PGC. (b) Layout of the adder.

adder are connected to SPICE voltage sources at the six labeled points,  $V_{a0}$  through  $V_{a5}$ . These points are determined by locating the intersection of each label placed in the PGC with the local  $V_{DD}$  and GND rails of the adder. Therefore, it may be considered a case of partitioning using the FP-scheme as the local rails are considered part of the CLC. The GND connection points (not shown) are adjacent to the  $V_{DD}$  connection points.

The six  $V_{DD}$  input locations in the PGC ( $V_{a0}$  through  $V_{a5}$ ) were found to traverse 4 different iso-IR bands. Therefore, the computation of power grid response for a given input sequence entails six convolution operations using four IR functions. Figure 9(a) shows the current transients measured at the six input locations using SPICE simulations. The response of the PGC at  $V_{DD0}$  obtained by con-

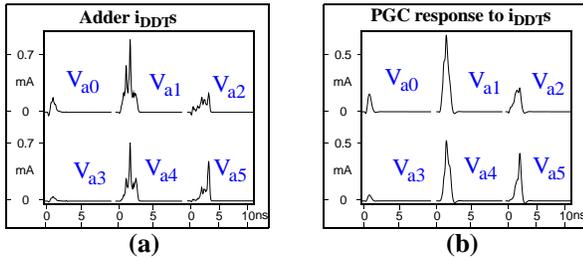


Figure 9. (a) Adder  $i_{DDT}$  waveforms that form inputs the PGC. (b) Response of the PGC to the individual  $i_{DDT}$ s in figure 9 (a).

volving each of the current waveforms with their corresponding IR functions is shown in Figure 9(b). The overall response of PGC to the sensitized adder logic measured at  $V_{DD0}$  is obtained by linear superposition of its response to individual current sources. Figure 10(a) shows the overall response of the PGC to the sensitized adder logic at  $V_{DD0}$  obtained using convolution based method overlaid with the response obtained using SPICE simulations.

The complexity involved in the computation of power grid response based on convolution operations can be fur-

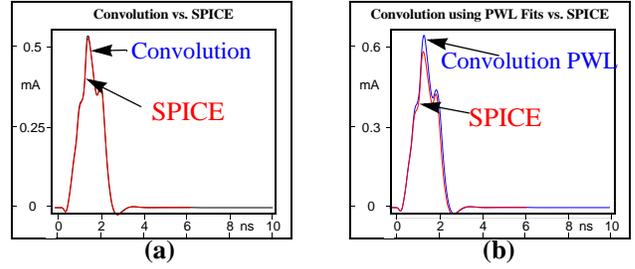


Figure 10. (a) Overall response using SPICE superimposed with convolution derived results. (b) Overall response using SPICE superimposed with convolution derived results of PWL fits to  $i_{DDT}$ s.

ther reduced by creating piece-wise-linear (PWL) abstractions of the input current waveforms. An algorithm based on detection of the change in polarity of slope in a waveform is used to derive its PWL abstraction. This significantly reduces the number of points in each input signal and thus the number of multiplication and addition operations. Figure 10(b) shows the overall response of the grid obtained using convolution based on PWL abstracted inputs. Overlaid with this curve is the SPICE generated response obtained using original SPICE waveforms as inputs. The peak amplitude and width values obtained using SPICE and the percentage difference error as compared to the convolution results on original SPICE and PWL inputs are listed in Table 1. The width of  $i_{DDT}$  is measured as the time-interval between the points at which the waveform attains 5% of its peak value.

Original $i_{DDT}$ s	PWL fitted $i_{DDT}$ s
5% width (spice) = 3.26ns	5% width (spice) = 3.26ns
peak (spice) = 0.594 mA	peak (spice) = 0.594 mA
% width error = 0.4%	% width error = 3%
% peak error = 1.8%	% peak error = 10%

Table 1: Amplitude and Peak Error.

## 8.0 Isolated Path (CLC) Simulation

The SPICE simulations of isolated paths can be used as a means of enabling the simulation of the CLC paths to derive the current input waveforms that feed the PGC. We can analyze the complexity involved in such a scheme by assuming the average number of gates sensitized under a given test sequence. If we assume the maximum fanout ( $F_O$ ) of 4, logic depth ( $D$ ) of 5 and maximum fanin ( $F_I$ ) of 3, the maximum number of sensitized gates can be computed using geometric progression given by Eq. 7. Assum-

$$Max\ Numer\ of\ gates = \frac{F_O^D - 1}{F_O - 1} = 341 \quad \text{Eq. 7.}$$

$$Max\ Numer\ of\ transistors = 2F_I \cdot \frac{F_O^D - 1}{F_O - 1} = 2046$$

ing that the minimum number of gates sensitized under a test sequence is 5 (30 transistors), the average number of

transistors per test sequence is 1038. This shows a significant reduction in the number of transistors that need to be simulated under a given test sequence compared to the entire CLC of the CUT. It must be realized that under any given test sequence several paths may get sensitized independently. All such independently sensitized paths can be isolated and simulated in parallel and their results can be combined during the PGC convolution process. Instead of using SPICE to simulate the isolated paths we may also make use of tools that perform Transient Current Simulations at the switch level verilog netlist. We are currently investigating the accuracy and complexity of Transient Current Simulations. Also results for paths already simulated when processing a previous input test vector can be reused for other test sequences.

## 9.0 Conclusions

This paper investigates the practical issues concerning the implementation of fault simulation methodology for  $i_{DDT}$  testing. A new model based on convolution based approach is proposed that can be used to compute the power grid response using the precomputed impulse response (IR) functions. This circumvents the need for running time and memory intensive transient simulations on the entire CUT. The categorization of IR functions into iso-IR bands is shown to further reduce the number of convolutions required to compute the transient response of the grid to the maximum number of iso-IR bands in the power grid. An approach based on isolation of sensitized paths from the layout is proposed as a means of enabling the simulation of core logic circuit. The accuracy and complexity of these methods are evaluated on part of a commercial power grid using a 16-bit logarithmic adder as the core logic.

## References

- [1] ITRS (<http://public.itrs.net/>).
- [2] J. F. Frenzel and P. N. Marinos, "Power supply current signature (PSCS) analysis: A new approach to system testing", IEEE proceedings of International Test Conference, pp. 125-135, 1987.
- [3] M. Hashizume, et al., "Fault detection of combinatorial circuits based on supply current", IEEE proceedings of International Test Conference, pp. 374-380, 1988.
- [4] A. P. Dorey, et al., "Reliability testing by precise electrical measurement", IEEE proceedings of International Test Conference, pp. 369-373, 1988.
- [5] F. Franco and E. J. McCluskey, "Delay testing of digital circuits by output waveform analysis", IEEE proceedings of International Test Conference, pp. 798-807, 1991.
- [6] J. S. Beasley, et al., "IDD pulse response testing of analog and digital CMOS circuits", IEEE proceedings of International Test Conference, pp. 626-634, 1993.
- [7] D. K. Papakostas and A. A. Hatzopoulos, "Analogue Fault Identification Based on Power Supply Current Spectrum", Electronics Letters, 7th January Vol. 2, No. 1, pp. 118-119, 1993.
- [8] G. Gielen, et al., "Fault Detection and Input Stimulus Determination for the Testing of Analog Integrated Circuits Based on Power-Supply Current Monitoring", IEEE/ACM ICCAD, pp. 495-498, 1994.
- [9] J. H. Wang, et al., "An Accurate Time-Domain Current Waveform Simulator for VLSI Circuits", Proceedings of EDAC, pp. 562 - 566, 1994.
- [10] A. Barber, et al., "A Bare-chip Probe for High I/O, High Speed Testing", HP Labs 1994 Technical Report-18, Hewlett Packard, 1994. (<http://www.hpl.hp.com/techreports/94/HPL-94-18.pdf>)
- [11] R. Z. Makki, et al., "Transient power supply current testing of digital CMOS circuits", IEEE proceedings of International Test Conference, pp. 892-901, 1995.
- [12] B. Vinnakota, "Monitoring Power Dissipation for Fault Detection", IEEE proceedings of VLSI Test Symposium, pp. 483-488, 1996.
- [13] A. Bogliolo, et al., "Gate-Level Power and Current Simulation of CMOS Integrated Circuits", IEEE Transactions on VLSI systems, Vol. 5, No. 4, 1997.
- [14] L. Smith, et al., "Power Distribution System Design Methodology and Capacitor Selection for Modern CMOS Technology", IEEE Transactions on Advanced Packaging, Vol. 22, pp 284-291, 1999.
- [15] S. R. Nassif and J. N. Kozhaya, "Fast Power Grid Simulation", Proceedings of the 37th conference on Design Automation, pp 156 - 161, 2000.
- [16] J. H. Wang, et al., "Current Waveform Simulation for CMOS VLSI Considering Even-Overlapping", IEICE Trans. Fundamentals, Vol. E83-A, No.1, 2000.
- [17] A. Singh, C. Patel, S. Liao, J. F. Plusquellic and A. Gattiker, "Detecting Delay Faults using Power Supply Transient Signal Analysis", IEEE Proceedings of International Test Conference, pp.704-712, 2001.
- [18] Y. Lee and C. C. Cheng, "Power Grid Transient Simulation in Linear Time Based on Transmission-Line-Modeling Alternating-Direction-Implicit Method", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, pp 1343-1352, Vol. 21, NO. 11, Nov. 2002.
- [19] Y. Lee and C. Chen "The Power Grid Transient Simulation in Linear Time Based on 3D Alternating-Direction-Implicit Method", Proceedings of the Design, Automation and Test in Europe Conference and Exhibition, 2003.
- [20] T. Wang, et al., "3D Thermal-ADI: An Efficient Chip-Level Transient Thermal Simulator", ACM International Symposium on Physical Design (ISPD), ACM 1581136501/ 03/ 0004 2003.
- [21] MOSIS at <http://www.mosis.edu/Technical/Testdata/tsmc-025-prm.html>.