Process-variability aware delay fault testing of delta VT and weak-open defects

Citation for published version (APA):

Arumi-Delgado, D., Rodriguez-Montanes, R., Pineda de Gyvez, J., & Gronthoud, G. (2003). Process-variability aware delay fault testing of delta VT and weak-open defects. In *Proceedings of the Eighth IEEE European Test Workshop, 2003, 25-28 May 2003* (pp. 85-90). Institute of Electrical and Electronics Engineers. <https://doi.org/10.1109/ETW.2003.1231673>

DOI:

[10.1109/ETW.2003.1231673](https://doi.org/10.1109/ETW.2003.1231673)

Document status and date:

Published: 01/01/2003

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Process-Variability Aware Delay Fault Testing of ΔV_T and Weak-Open Defects

Daniel Arumí-Delgado
Rosa Rodríguez-Montañés
Polytechnical University of Catalunya
Department of Electronics Engineering
Barcelona, Spain
e-mail: arumi@eel.upc.es

José Pineda de Gyvez
Guido Gronthoud
Philips Research Laboratories
Prof. Holstlaan 4, 5656 AA Eindhoven,
The Netherlands
e-mail: jose.pineda.de.gyvez@philips.com

Abstract

Circuits are tested for both functionality and performance. As opposed to circuits with large delay faults, circuits with small delay faults are difficult to diagnose and as such are potential test escapes. This paper presents a strategy to diagnose these types of defective circuits. Our approach takes into account the effect of inter and intra die process variability in the detection method. Unlike conventional delay fault testing, our approach propagates a delay fault in such a way that the mean of the statistical delay distribution of the faulty circuit is further increased with respect to the mean of a fault-free circuit. This is a two-step strategy that is derived from a simple fault model. The paper further presents simulation results supporting the delay fault testing.

1. Introduction

Modern deep submicron circuits are affected not only by spot defects but also by process variations [1]. Process variations can have both global and local effects on an IC. The interconnect resistance increasing (or decreasing) radial gradient across the wafer is an example of global variation, while threshold voltage mismatch can be considered a local variation because of its random nature. Faults arising from process variability result in a correct circuit's output but without meeting performance specs. This type of faults is important to the semiconductor industry because of their direct association to test escapes.

Figure 1 shows the path delay distribution of a batch of DSP like circuits. This is a module of more than 250k transistors used for experimental yield ramp of a CMOS 0.18 μ m technology. The horizontal axis shows the delay, the left vertical axis shows the percentage of chips with a given delay; the right axis shows the test escapes in ppm. Notice that the distribution's right tail can be quite long due to the presence of faulty circuits. This tail is shown as the second curve. To determine whether a circuit is faulty or not, a delay threshold is used and tuned to minimize test escapes, i.e. it is a trade off between i) the IC performance

and the slow corner of the process, ii) the desired test cost and the test quality, and iii) the IC reliability. For the distribution of Figure 1 one has the choice of testing the circuit for delays above 11ns depending upon the required ppm target level.

Sometimes, when a fault does not add a large delay, the circuit can be found in the neighborhood of the delay threshold region. It follows then that it is difficult to differentiate if a circuit found in the shaded area of Figure 1 is a faulty one. In other words, it is not clear which circuits must be rejected. Therefore, the objective of this work is to propose a method that amplifies the delay difference between defect-free and defective circuits to easier decide which circuits must be rejected.

Figure 1. Typical delay distribution of a set of circuits in wafer batch.

The circuit's delay can be tested according to a delay fault model. Several models have appeared in the literature, e.g. the transition delay fault model [8], the gate delay fault model [6, 10], the path delay fault model [5-7, 9, 11-12], the line delay fault model [5], and the segment delay fault model [2, 11]. The main drawback of transition delay fault model is that it only detects large delay variations, and taking into account that defects can provoke a wide range of delay values, many faults would not be detected. The gate delay fault model has as advantage the limited number of modeled faults, but on the other hand, the main drawback is that the required values for the delays are not always available per individual gate. In the path delay fault model faults are modeled as the sum of small delays on each connection along the signal path. Its main drawback is

that it is not possible to enumerate all paths. In this way, any path with a total delay exceeding the system clock interval is said to have a path delay fault. The line delay fault model is a combination between the transition and the path delay fault model. The segment delay fault model combines the advantages of the single models while avoiding some of their limitations.

The method presented in this paper is based on the path delay fault model. However, our approach sensitizes the target gate in an optimum way such that the nominal path delay difference between the fault free circuit and the faulty one is maximized. All other gates are sensitized such that their delay spread is statistically compensated to minimize its impact in the total path delay.

2. Qualitative Gate Delay Model

To qualitatively explain the delay behavior of a digital gate, let us take a look at the equivalent RC network. When a transistor is *on*, its resistance can be calculated as

$$R_{on} = \frac{V_{DD}}{\beta \frac{W}{L} (V_{DD} - V_T)^\alpha} \quad (1)$$

where α is the velocity saturation factor, V_T the threshold voltage and β the transconductance. Without loss of generality consider now an N input *nand* gate. Consider the parallel connection of identical PMOS transistors of this gate. Using the equivalent digital transistor model the intrinsic time constant for a low to high transition at the output is given as

$$T_{LH} = \frac{R_p}{N} (NC_{op}) = R_p C_{op} \quad (2)$$

where C_{op} is the transistor's output capacitance. With an external load capacitance C_L the delay becomes

$$T_{LH} = \frac{R_p}{N} (NC_{op} + C_L) \quad (3)$$

This assumes that the transistors are identical, e.g. free of process variations, and that there is no input signal skew, i.e. all gate signals arrive at the same time. Obviously a V_T mismatch will result in a change of the equivalent R_p . A variation in the transistor geometries yields a change in both R_p and C_{op} . Likewise, input signal skew results in an off resistance much greater than the transistor's R_p . In summary the circuit's delay varies.

Now let us take a look at the series connected NMOS transistors of the *nand* gate. The intrinsic delay time constant for a high to low transition taking into account a load capacitance is given as

$$T_{HL} = NR_n \left(\frac{C_{on}}{N} + C_L \right) \quad (4)$$

where C_{on} is the NMOS transistor's output capacitance. This neglects the internal node capacitance between transistors. A first order approximation is to assume that the N series connected transistors behave as a transmission line with time constant

$$T_d = 0.35 C_n R_n N^2 \quad (5)$$

where C_n is the parallel connection of a transistor's input and output capacitance, R_n is the transistor's *on* resistance. It follows then that the delay of a high to low transition is given as

$$T_{HL} = NR_n \left(\frac{C_{on}}{N} + C_L \right) + 0.35 R_n C_n (N-1)^2 \quad (6)$$

Let us label the inputs of our *nand* as A and B, respectively. Figure 2 shows actual transistor-level simulation results using a full process description of this *nand*. To analyze the impact of threshold voltage mismatch let us consider a constant V_{TPB} and a varying V_{TPA} . If only one input signal is being switched at a time, the change in delay ΔT_{pLH} with respect to the nominal value (T_{pLHNOM}) depends only on the transistor being switched. This delay is larger than when multiple transistors are switched on; see eqn. (3). Assuming a constant V_{TPB} but greater than its nominal value, the $T_{pLH} - T_{pLHNOM}$ difference for the B input transition only is a constant value (a horizontal line in Figure 2). Conversely, the $T_{pLH} - T_{pLHNOM}$ behavior, for a transition in the A input only, varies along the x axis for the various values of V_{TPA} spread; the delay increases if the PMOS_A transistor's V_{TPA} increases since its on resistance increases too (see eqn. (1)) resulting in a slightly larger RC constant, as seen in eqn. (2). For the case of two simultaneous transitions (A & B inputs), an intermediate delay variation is found (see Figure 2). More details on ΔV_T fault modeling are given in [13].

In the case of series-connected transistors, the same current flows through each one of the transistors. If more than one transistor switches *on* at the same time, the output response takes longer than if only one does (and the rest are already *on*) as predicted by eqn. (6). In the former case, there is more than one transistor with an initial high resistance while in the latter only one transistor has an initial high resistance.

Figure 2. Delay for V_{TPB} constant and greater than the nominal value (V_{TPBNOM}) and varying V_{TPA}

3. Testing Process Induced and Weak Open Defects

According to the results shown in the previous sections, the effect of V_T mismatch on the timing behavior of CMOS combinational gates depends on the amount of switching activity at their input signals. Indeed, for parallel-connected networks, this effect (delay variation) is maximized by switching as less input signals as possible. On the contrary, the excitation of more than one input signals (parallel-connected networks) helps in the statistically compensation of the delay variation at the output of the gate. In the case of serial-connected networks, the timing behavior is qualitatively similar although not as advantageous as in the parallel case. In this section, a strategy is proposed taking into account the previous dependence between the network nature (parallel or series) and the maximum or minimum delay variation at the output of each gate.

Let us assume a general circuit with a target gate that can be excited by means of the path illustrated in Figure 3. The effects of process parameter variation on the delay of the target gate are maximized by individually switching its inputs.

Figure 3. Delay path in a circuit

Let us assume that this delay variation is named ΔT_G . The delay at the external output of the circuit results from adding the delay of every gate in the path. In general, these gates may also present a delay variation (ΔT_P). This delay variation has been divided into ΔT_{P1} for the gates previous to the target gate and ΔT_{P2} for the gates driven by the target gate. To maximize the effect of the ΔT_G on the total delay of the path, the relationship between ΔT_G and ΔT_P should be as big as possible. This is achieved if the target gate is excited according to a maximizing excitation and the rest of the gates in the path are excited according to a minimizing excitation pattern (see Figure 3). The goal of our strategy is thus to separate the fault free delay distribution from the faulty one. There are two possible ways to achieve this; one way is by shifting one distribution with respect to the other (Figure 4), that is to say, increasing the value of $\mu'_d - \mu_d$ which is the difference between the mean values of the defective and non-defective distributions. The other way is by decreasing the distribution spread; that is to say, decreasing the sigmas of both distributions, see also Figure 4. In both cases, the overlapping area is reduced and consequently, it will be easier to distinguish between them.

In the next subsections, our approach to achieve a reduction of the spread and an increase of $\mu'_d - \mu_d$ are presented.

Figure 4. Increasing $(\mu'_d - \mu_d)$ and decreasing the spread.

3.1. Shifting the delay distribution

Let us assume the circuit shown in Figure 3 with a range of process and electrical parameters belonging to a set of acceptable values. Under this condition the path is expected to have a delay distribution as illustrated in Figure 5 a). Now, let us assume that the target gate has an important $+\Delta V_T$ variation in the transistor being switched by the test pattern. The new delay distribution (denoted by (μ'_d, σ_d)) of this defective circuit is shifted to the right due to the effect of the increment in the threshold voltage of the target transistor (see Figure 5 b). The standard deviation of the distribution is due to the parameter variation of the rest of the transistors belonging to the path.

For a given path, the parameters of the delay distribution (μ_d, σ_d) depend on the applied test pattern. In Section 2 we observed that the mean value of the distribution is increased if single transitions are provoked in parallel networks. A similar effect is obtained if multiple transitions are applied in series-connected networks with internal parasitic capacitors initially pre-charged. Thus, the first step to distinguish between the “acceptable” circuit from the “defective” circuit is to significantly make their delay distributions shift one with respect to the other (i.e. to separate their mean values).

Figure 5. Delay path distribution after a Monte Carlo simulation for a) all the parameters and b) all the parameters except the target transistor affected by a fixed $+\Delta V_T$

3.2. Narrowing of the delay distribution

Once the delay distribution is shifted as much as possible from the “expected” case, the intersection between both distributions has to be minimized (to decrease their overlap). This can be accomplished by decreasing the standard deviation of both distributions. Indeed, for a fixed difference between both mean values, i.e., $\Delta\mu = \mu'_d - \mu_d$, the intersection depends on the σ_d value; the smaller the σ_d , the smaller the intersection area. The sigma reduction is achieved by switching as many signals as possible in all the gates of the delay path except the target gate. In this case the sigma contribution of each gate is minimum (due to the compensation of effects presented in previous sections) and the final addition of all values is minimum too.

3.3. Low Voltage Testing for V_T Defects

Low voltage testing maximizes the sensitivity of delay faults. Taking into account the delay approximation formula

$$T \propto \frac{V_{dd}}{(V_{dd} - V_T)^\alpha} \quad (7)$$

where V_{dd} is the supply voltage, the sensitivity between the delay and V_T is found as

$$S_T^{V_T} \propto \frac{\alpha \cdot V_T}{V_{dd} - V_T} \quad (8)$$

One can observe that the sensitivity between T and V_T increases if the difference $V_{dd} - V_T$ decreases. This means that the lower the value of V_{dd} , the better. Thus, if V_{dd} decreases, the circuit is more sensitive to variations of V_T . Consequently, the delay distribution is shifted to the right with a bigger difference of $\mu'_d - \mu_d$. Naturally, V_{dd} must remain bigger than a certain value to ensure the correct operation of the circuit. To verify eqn. (8), a simple statistical simulation of a nand gate was carried out at $V_{dd}=1V$ and $V_{dd}=1.8V$ with and without a fault. We found that for $V_{dd}=1V$ the mean delay difference between the faulty and fault free circuit was around 300ps. For $V_{dd}=1.8V$ this difference was only 50ps.

4. Delay Fault Testing Methodology

The proposed approach comprises tracing a signal path from a primary input to a primary output through the site where the fault is. The primary output has a logic state that is dependent upon the logic value at the site of the fault. In this way, the method of test pattern generation using path sensitisation techniques can be divided into different steps as follows:

- Given the target line, specify the other inputs of the faulty gate to make its output response sensitive to the fault condition.
- Sensitize the path, by assigning non-controlling logic values to each off-path input, so that the output of every gate is sensitive to the fault condition. In this way, the effect of the fault is propagated to a primary output.
- Generate as many transitions as possible in all off-path-lines, except in the target gate, to sensitise the whole path in the optimum way. Indeed, the distribution effects are compensated in every gate, so the delay spread is the minimum possible.
- Determine the set of primary input values necessary to carry out the previous steps. The set of input values so derived constitutes the test for the fault condition.

We implemented a modified algorithm of [3] to sensitize the delay path. Once the target path, the enclosed paths and all the transitions for these paths are known, the next step consists of ensuring that all these transitions are properly generated. An auxiliary network, different for every pattern, is added to the circuit so that with the detection of a stuck-at-fault in the last gate of this auxiliary network, the off-path-lines and the on-path-lines are set to the required values. As this procedure is carried out for both patterns, the combination of these results leads to the required transitions for every line.

Figure 6: Auxiliary network for the first pattern

Imagine a path like in Figure 6. Suppose that the response for the target gate must be tested when its first input has a falling transition. Once the target transition is known, automatically the transitions for all the other on-path lines and off-path-lines are known. Thus, for the first pattern, every line with a rising transition must firstly be set to 0, and to 1 if the transition is a falling one. Similarly, for the second pattern, the lines with a rising transition must be initially set to 1, and the lines with a falling transition to 0, see Figure 7. It can be observed that both networks are complementary. If the auxiliary networks shown in the figures are added, and we ask for a stuck-at-one at the output of the last gate of these networks, the primary inputs will be controlled such that, the lines will be set to the required values. Accordingly, the two patterns necessary to test the fault will be obtained.

The idea of the auxiliary network is based on the stuck-at-fault model. Given an *or* gate, testing a stuck-at-one at its

output, means that the line must be set to 0. Therefore, all the inputs of the *or* gate must be set to 0 as well. Thereby, every circuit line that must be set to 0, becomes an input of an *or* gate in the auxiliary network. Similarly, testing a stuck-at-one fault at the output of a *nand* gate means that its output must be set to 0, that is, all its inputs must be set to 1.

Thus, all the circuit lines that must be set to 1 become one of the inputs of a *nand* gate in the auxiliary network. If the auxiliary network consists only of these *or* and *nand* gates, whose inputs are the lines of the circuit, asking for multiple stuck-at-faults sets all the required values on the lines. In fact, we will need to ask for a stuck-at-fault at the output of every gate. To solve this problem, an *or* gate whose inputs are the outputs of all the gates of the auxiliary network is added. In this way, now it is only needed to ask for a stuck-at-one fault at the output of this gate.

When the path is longer, using one *or* gate only is not smart enough, since the number of gates in the auxiliary network could be big. Therefore, sometimes instead of adding one *or* gate, a group of *or* gates will be needed.

Figure 7: Auxiliary network for the second pattern

5. Simulation Results

Example C432. Let us consider the ISCAS 432 circuit illustrated in Figure 8 where only a part of it has been depicted. Let us assume that one of the NAND gates is affected by an increment in the threshold voltage of its PMOS_A transistor (parallel-network), i.e., $V_{TPA} = V_{TPANOM} + \Delta V_{TPA}$. The observable output used to detect the delay is shown in the Figure. If the target gate (“defective”) is not excited properly, the difference between the delay distribution for the fixed $V_{TPA} = V_{TPANOM} + \Delta V_{TPA}$ case and the “acceptable” case is negligible (see Figure 9). In the graph, the black histogram depicts the distribution for the first case while the gray case depicts the general case with no fixed V_{TPA} deviation. One possible way to excite the gate is illustrated in Figure 8 in which the delay path has been highlighted. The test pattern allows the maximum difference between the mean value of the defective and the non-defective delay distributions (maximum $\Delta\mu = \mu'_d - \mu_d$). For this test pattern all gates are forced to switch individual inputs (transistors). In this way, the target gate is excited in the optimum way to maximize its delay but the rest of the circuit is not because the delay spread of each gate is not minimized. The results are shown in Figure 10. The

optimum excitation of the fixed V_{TPA} deviation is based on the path presented in Figure 8. The narrowing of the delay distribution can be achieved by applying the test illustrated in Figure 11. In this case, the last level of NOR and NAND gates have been excited in a multiple way to statistically compensate the effect of their deviations. Figure 12 shows the results for this test in agreement with the expected behavior. Indeed, both histograms have diminished the overlap as predicted.

Figure 8: ISCAS432; Selected delay path to detect the ΔV_{TPA} affecting the target gate (optimum excitation for the target gate & non-optimum for the rest of the circuit)

Figure 9: Histograms for the defective and non-defective circuits (no-excitation for the target gate and non-optimum for the rest of circuit), $V_{DD}=1V$

Figure 10: Histograms for the “defective” (Bad) and non-defective (Good) circuits (optimum excitation for the target gate and non-optimum for the rest of the circuit) at $V_{DD}=1V$.

Example C880. In this example we take a look at the diagnosis of weak open defects [14]. For this purpose a distributed RC network is introduced in every branch. The defective line resistance is set to 100k Ω to model the presence of a weak open. Results are shown in Table 1 and Figure 13.

Table 1: Simulations results for line 290 of C880

C880	Target gate (inverter) --Target line 290		
Simulation Weak open (100KO)	Nominal Value (ns)	Mean (ns)	Spread (ps)
Fault free	0.93	0.94	32.11
Faulty	1.49	1.50	38.60

Figure 11: Selected delay path to detect the ΔV_{TPA} that affects the target gate (optimum excitation for both the target gate and the rest of the circuit) **Figure 12: Histograms for the defective and non-defective circuits (optimum excitation for the target gate and also for the rest of circuit), $V_{DD}=1V$.** **Figure 13. Delay distribution for C880.**

6. Conclusions

A method to detect parametric faults has been presented. The proposed method takes into account the impact of inter and intra die process variations in the generation of test patterns to diagnose the fault. Our strategy uses a novel path sensitization method such that for the target gate, the statistical nominal delay difference between the fault free circuit and the faulty one is maximized. The other gates are sensitized in a way such that their delay spreads are compensated statistically to minimize the total path delay

spread. The concept of an auxiliary network has been introduced to obtain the testing patterns. Through the use of this network one can find the delay fault patterns by requesting stuck-at faults to the ATPG tool. Finally, simulations were carried out to check the usefulness of the proposed method. From these results one can conclude that the proposed methodology detects delay faults. It is able to capture delay faults caused by V_T variations, although the delays caused by this kind of faults are typically very small.

Acknowledgments

R. Rodríguez acknowledges the support of CICYT Project TIC2001-2246 .

References

- [1] M. Eisele, et al. "The Impact of Intra-Die Device Parameter Variations on Path Delays and on the Design for Yield of Low Voltage Digital Circuits", *IEEE Trans. on Very Large Scale Integ. (VLSI) Systems*, Vol. 5. No. 4, December 1997.
- [2] K. Heragu, J. H. Patel and V. D. Agrawal, "A Test Generator for Segment Delay Faults", *Twelfth Int. Conf. on VLSI Design*, pp. 484-491, 1999.
- [3] S. Kajihara, et al. "A Method for Identifying Robust Dependent and Functionally Unsensitizable Paths", *Int. conf. on VLSI Design*, pp. 82-87, January 1997.
- [4] W.-N. Li, S. M. Reddy, S. K. Sahni, "On Path Selection in combinational Logic Circuits", *Design Automation Confer., Proceedings, 25th ACM/IEEE*, pp. 142-147, 1988.
- [5] A. K. Majhi, et al., "Line Coverage of Path Delay Faults", *Trans. on Very Large Scale Integ. (VLSI) Syst., IEEE*, Vol 8, pp 610-614, Oct 2000..
- [6] V. H.-W. Meyer, W. Anheier and A. Sticht, "Non-Robust Delay Test Pattern Generation Based on Stuck-at TPG", *IEEE Int. Conf. on Electronics, Circuit and Systems, ICECS*, Volume 2, pp. 1007-1010, 2001.
- [7] M. Michael and S. Tragoudas, "Functional-Based ATPG for Path Delay Faults", *Southwest Symposium on Mixed-Signal Design, SSMSD*, pp. 159-164, 2000.
- [8] W. Moore, G. Gronthoud, K. Baker, M. Lousberg, "Delay-Fault Testing and Defects in Deep Sub-Micron ICs – Does Critical Resistance Really Mean Anything?", *Int. Test. Conf.*, pp. 95-104, 2000.
- [9] I. Pomeranz and S. M. Reddy, "A Flexible Path Selection Procedure for Path Delay Fault Testing", *VLSI Test Symp.*, pp. 152-159, 1999.
- [10] A. K. Pramanick and S. M. Reddy, "On the Fault Coverage of Gate Delay Fault Detecting Tests", *IEEE Trans. on Comp. Aided Des. of Integ. Circ. and Syst.*, Vol. 16 No. 1, Jan. 1997.
- [11] M. Sharma and J. H. Patel, "Testing of Critical Paths for Delay Faults", *Int. Test. Conf.*, pp. 634-641, 2001.
- [12] H. Wittmann and M. Henftling, "Path Delay ATPG for Standard Scan Design", *European Proceedings EURO-DAC, with EURO-VHDL*, pp. 202-207, 1995.
- [13] J. Pineda de Gyvez and R. Rodríguez-Montañés, "Threshold Voltage Mismatch ΔV_T Fault Modeling," *VLSI Test Symposium*, pp.145 -150, 2003.
- [14] R. Rodríguez-Montañés, P. Volf, J. Pineda de Gyvez, "Resistance Characterization for Weak Open Defects" *IEEE Design&Test of Computers*, pp. 18-26, Sept.-Oct., 2002