

Pulse propagation for the detection of small delay defects

M. Favalli
DI - Univ. of Ferrara

C. Metra
DEIS - Univ. of Bologna

Abstract

This paper addresses the problems related to resistive opens and bridging faults which cannot be detected using delay fault testing because they lie out of the most critical paths. Even if the induced defect is not large enough to result in timing violations, these faults may give rise to reliability problems. To detect them, we propose a testing method that is based on the propagation of pulses within the faulty circuit and that exploits the degraded capability of faulty paths to propagate pulses. The effectiveness of the proposed method is analyzed at the electrical level and compared with the use of reduced clock period which can detect the same class of faults. Results show similar performance in the case of resistive opens and better performance in the case of bridgings. Moreover, the proposed approach is not affected by problems on the clock distribution network.

1. Introduction

Resistive opens (ROPs) may produce timing degradations [1] that, in synchronous ICs, can be detected if the size of the induced delay defect exceeds the slack allowed for the faulty signal(s). As a consequence, ROPs can be detected by delay fault (DF) testing. This kind of testing can also be used to detect those resistive bridgings which do not result in functional errors [2].

Unfortunately, faults affecting only slow paths may be not detectable because the available slack is larger than the defect size. Conversely, these faults have been shown to possibly result in reliability problems [3].

In addition, undetectable defects which are not detected by production testing may give rise to functional errors during the IC's operating life because of the timing performance degradation due to aging [4].

In a combinational circuit, ROPs and bridgings can be detected using a clock period smaller than that fixed by the critical path, flip-flops' timing parameters and clock skews [5, 6, 7]. In this way, the available slack is reduced and the transitions of the affected primary outputs may occur after the sampling instant, thus resulting in fault detection.

These techniques deal in different ways with the problems related to delay fluctuations. In [6], hazard-free delay tests are applied by sensitizing group of paths featuring similar delays under nominal conditions. The output values sampled using reduced clock period(s), however, are sensitive to uncertainties on timing parameters which make the delay of a

path vary within the same wafer and lot.

To approach this problem, in [5] the rate of the clock test is progressively increased within a given interval, thus resulting in the capture of multiple data. Faulty devices are identified by comparing the current results with those of devices which are expected to produce a similar behavior. The uncertainties on timing parameters are kept under control by comparing neighboring dies and considering the same test conditions.

In order to reduce the number of data, multiple sampling times are still used in [7], but the author exploits the ordering of the transition times of the signals belonging to the logic block under test. A DF is detected if the switching order of any two outputs is opposite to that evaluated by means of fault-free simulation. This method must use signal transitions which are not too close: a too fine ordering may be impaired by timing fluctuations. In addition, fault effects can be masked by the presence of multiple path DFs.

Let us also note that DF testing should account not only for the uncertainties on the path's delays, but also for the uncertainties related to the timing of the clock distribution network. In fact, the buffers used to regenerate the clock signals may be affected by delay fluctuations. In particular, in [5] we have to match two samples of clock distribution networks belonging to neighboring dies. In [7], instead, we have to match two (partially) different clock distribution sub-networks in the same die. In this regard, it is well known that in deep-submicron devices, both within-die and die-to-die timing fluctuations are expected to grow [8].

In this work, we propose an alternate technique to detect this kind of faults. It requires a smaller amount of test data and it may reduce some of the problems related to previous techniques. In particular, it is based on the propagation of pulses through the faulty circuit. ROPs and bridging faults (BFs), in fact, do not affect only paths' delays, but also the paths' capability to propagate pulses¹. We will show that a pulse which is propagated through a fault-free circuit may be dampened in the faulty circuit, thus exposing the presence of a fault.

Our method requires to: a) sensitize a path including the target fault; b) inject a suitable pulse at the path's input; c) verify whether such a pulse is propagated to the path's output or not.

¹Note that we do not refer to the inertial delay of single gates (which of course are impaired), but to pulse propagation along a path which is a complex phenomenon involving more gates.

The size of the injected pulse should be large enough to avoid the rejection of fault-free circuits in the presence of random fluctuations of IC's parameters. Of course, this poses some limitations on the range of detectable resistances. It should be noted, however, that also DF testing with reduced clock period has to trade-off test quality for yield.

Differently from the clock signal used in [5, 6, 7], the input pulses are locally generated and the output pulses are locally detected, thus avoiding the problems related to the clock distribution network. This is achieved with some hardware overhead. However, our method exploits well known circuits for the generation of input pulses. For the detection of pulses at the outputs of the tested circuit it uses circuits [9] that were introduced to on-line detect transient faults originated by ionizing particles. Therefore, an useful synergy between off-line and on-line reliability indicators can be obtained.

Since the proposed method is completely independent of synchronization constraints, it can also be used to test bus lines using handshake protocols to transfer data.

2. Target faults

We will target ROPs and bridgings inside combinational blocks. We analyze the effects of these faults from the point of view of DF testing and the proposed method.

ROPs may be due to partial breaks or resistive vias affecting a logic cell or its output interconnects. These cases will be referred to as *internal* and *external*, respectively. In both cases, depending on the resistance value, the propagation delay along paths including the fault may be increased. In addition, the path's capability to propagate pulses is reduced.

These faults are characterized by an additional resistance (R) affecting a conductive path.

In the *internal* case, the driving capabilities of the pull-up or pull-down network of a CMOS gate are impaired, thus affecting only one kind of gate output transition. In this regard, Fig. 1a shows an example of ROP that slows down any rising transition of the gate output (B). From the point of view of DF testing, this delay defect increases the delay of any path propagating a rising transition through the affected gate. When considering the propagation of pulses, the size of any 010 transition of B will be shrunk because the fault affects the rising transition, but not the falling one.

Fig. 2, shows the faulty waveforms of the circuit in Fig. 1a for $R = 18k\Omega$ when a pulse is propagated through the affected path. These waveforms are compared to the fault-free case. As can be seen, the rising transition of signal B is delayed by the fault and the pulse is dampened in a few logic levels.

Fig. 1b shows an *external* ROP between the gate output B and its fan-out branch $B.C$. As a consequence of this fault, the propagation delay along a path including these signals will be increased for both kinds (rising/falling) of transitions.

From the point of view of pulse propagation, it should be

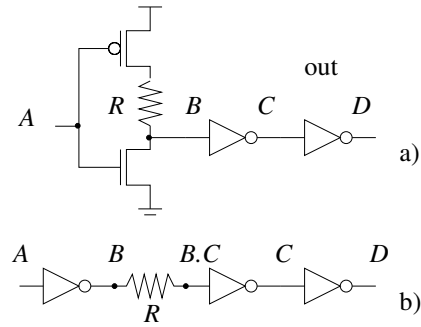


Figure 1. Internal (a) and external ROPs (b).

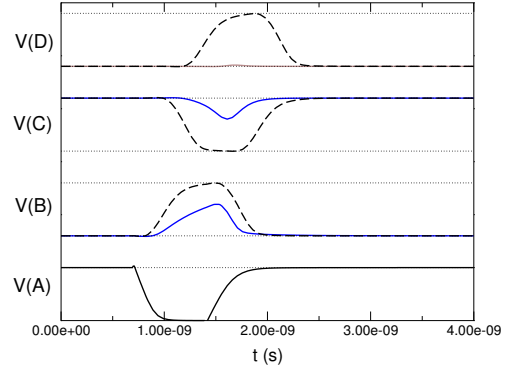


Figure 2. Faulty (solid lines) and fault-free (dashed lines) voltage waveforms in the faulty circuit in Fig. 1a.

noted that a pulse is less likely to be dampened than in the *internal* case, because both pulse edges are affected in the same way. As an example, Fig. 3 shows the propagation of a pulse in the circuit of Fig. 1b for $R = 18k\Omega$. As can be seen, the slopes of the affected transitions (of $B.C$) are consistently decreased. Depending on the width of the initial pulse, two possible behaviors may be in order: 1) if the pulse is much larger than the transition time of $B.C$, its width (measured, for instance, at $0.5V_{DD}$) will not be decreased; 2) otherwise, the second transition of the pulse starts when the first one is not yet exhausted, thus resulting in an incomplete pulse that is likely to be dampened (Fig. 3).

By comparing the waveforms of signal D in Figs. 2 and 3 it can be noted that, for the same values of R , the effects of *internal* ROPs are more relevant than those of *external* ROPs.

The behavior induced by resistive bridging faults is slightly more complex. Depending on the bridging resistance and the faulty network topology, BFs may give rise to: 1) functional errors and/or oscillations; 2) additional delays; 3) changes in the static and dynamic current.

Low resistance BFs give rise to functional errors or oscillations (in case they close inverting feedback loops) and they are supposed to be detected by functional testing. Therefore,

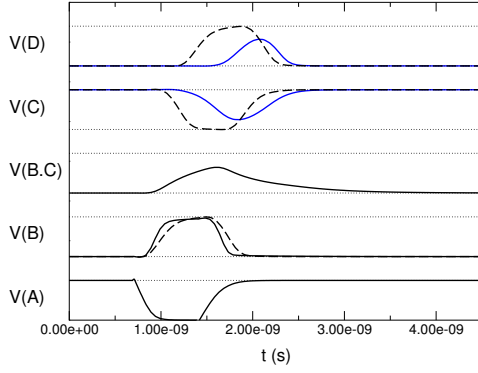


Figure 3. Faulty (solid lines) and fault-free (dashed lines) waveforms in the circuit in Fig. 1b.

we will focus on resistive BFs that provoke a voltage degradation of the affected signal(s) which is not large enough to result in functional errors, but which may result in, possibly significant, additional delays.

The kind of transition delayed by a BF mainly depends on the position of the bridged nodes which may belong to either the same logic gate (*internal BF*) or different logic gates (*external BF*). As an example, we will restrict our attention to non-feedback *external* BFs affecting gate outputs (Fig. 4). We will also use test vectors that propagate a transition or a pulse through one of the bridged gates while the output of the other one remains steady.

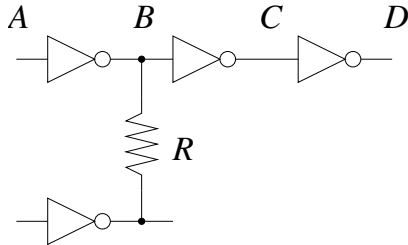


Figure 4. Example of external bridging fault.

Fig. 5 shows the propagation of a pulse through the path illustrated in Fig. 4 under nominal conditions. As can be seen, an incomplete pulse is produced, which is dampened in a few logic levels.

The case of *internal* BFs is slightly more complex and it is not considered here for the sake of brevity.

3. Testing environment

As shown in the previous section, ROPs and bridgings can be detected by verifying whether or not a pulse is propagated along a path including the faulty circuit. These operations can be performed using suitable pulse generators at the PIs of the combinational block under test. The POs of such a

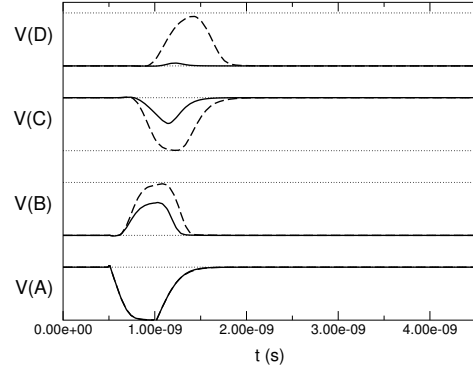


Figure 5. Faulty (solid lines) and fault-free (dashed lines) waveforms in the circuit in Fig. 4.

block, instead, have to be monitored by means of sensing circuits able to detect the presence of transitions.

Circuits of this kind have been introduced in order to on-line detect the possible presence of delayed transitions or transient faults. In the presence of transitions occurring when signals are expected to be steady, they produce an error indication [9]. Their use in the proposed approach is dual, because the presence of transitions indicates the absence of the fault, while the presence of a fault is denoted by the absence of transitions.

To simplify the study of fault detection, we will suppose that it is possible: 1) to sensitize a path starting from a PI and ending to a PO which includes the fault location; 2) to inject a pulse in such a path and to verify whether it propagates to the PO or not.

As regards sensitization, we will suppose that all the side inputs of the path are set to non controlling values.

In delay fault testing, the clock frequency is the main parameter to be set in order to detect faults. In our case, instead, we have to determine the width of the pulse to be injected in the path containing the fault.

Both operations are affected by ICs' parameters fluctuations. In particular, DF testing must consider: a) the skew between the clock signal triggering input transitions and that sampling output signal; b) the uncertainty in the path's timing; c) the uncertainty in the flip-flops timing. The proposed method, instead, must account for: a) the uncertainty in the width of the input pulse; b) the uncertainty in the width of the pulse which can be propagated by the path; c) the uncertainties in the timing of the sensing circuit.

From this point of view, it should be noted that in DF testing the uncertainty in gate propagation delays are combined with the uncertainty in path delay. This cumulative effect is only partially present in the case of pulse propagation. This can be easily explained by considering the logic

level model of the path. In this case, the filtering capabilities of a path depend only on the largest between the inertial delays of the gates in the path. In practice, this is only an approximation and the capability to propagate pulses typically depends on small segments [10]. However, the standard deviation on path's propagation delay is larger than that on the size of pulses which can be propagated.

Moreover, in DF testing, we have to account for the skew between the clock signals sampling PIs and POs that is due to the different distribution networks.

4. Fault detection

In this section, we consider paths containing ROPs and bridgings. For these circuits, we compare the proposed approach with the use of variable clock period.² To deal with IC's parameter fluctuations, this comparison has been performed at the electrical level using a Monte Carlo (MC) approach.

Under nominal conditions, a sensitized path p is characterized by $d_{p,r}^0$ and $d_{p,f}^0$ which denote the propagation delays of input rising and falling transitions, respectively.

As for pulse propagation, we can inject two different kind of pulses (which will be referred to as $h = 010$ and $l = 101$, respectively). In this case, the sensitized path is characterized (under nominal conditions) by two functions $w_{out} = f_{p,h}^0(w_{in})$ and $w_{out} = f_{p,l}^0(w_{in})$ which relate the width of the output pulse (w_{out}) to the size of an input pulse (w_{in}) of kind h or l , respectively. To simplify the notations, we will thereafter omit the indication of the kind of transition or pulse.

Nominal parameters cannot characterize the actual behavior of ICs in DSM circuits. Therefore, we considered a sample (S) of circuits. In a given circuit $s \in S$, p is characterized by its delay d_p^s and by its relationship (f_p^s) between w_{in} and w_{out} . In the faulty circuit, both these quantities are a function of the faulty resistance R : $d_p^s = d_p^s(R)$ and $w_{out} = f_p^s(w_{in}, R)$.

In DF testing, the test circuitry is supposed to include a flip-flop (FF_0) feeding the input of the path and a flip-flop (FF_1) sampling the output of the path. The DF can be detected by applying an input transition and comparing the sampled value with the fault-free one. Under nominal conditions, let t_0 be the transition instant of the clock signal triggering the input change and t_1 be the sampling instant of the output. Therefore, $T^0 = t_1 - t_0$ is the nominal value of the clock period used to test the path. Note that, in the considered kind of DF testing, T^0 is typically smaller than the operating clock period. The uncertainties on the clock distribution network make the clock period T which is actually used to test p be different from T^0 .

At the logic level, a faulty circuit instance (s) is detected if $T < d_p^s(R) + \tau_{CQ}^s + \tau_{DC}^s$, where τ_{CQ}^s and τ_{DC}^s are the delay and the setup time of FF_0 and FF_1 .

In the proposed method instead, the testing circuitry is characterized by the nominal width of the injected pulse (ω_{in}^0) and the nominal width of the minimal pulse (ω_{th}^0) which can be detected by the sensing circuit. Also in this case, let ω_{in} and ω_{th} be the values of such parameters in the actual circuit. In this case, the fault affecting a circuit instance s is detected if $w_{out}(R) = f_p^s(\omega_{in}) < \omega_{th}$.

When selecting T^0 or ω_{in}^0 and ω_{th}^0 , we have to trade-off the test quality for yield by accounting for the fluctuations of circuit parameters. In fact, by lowering T^0 or increasing ω_{th}^0 , the range of detectable resistances and the number of detected faults increase, but it is possible to produce false positives, thus decreasing yield.

To determine testing parameters, we performed Monte Carlo (MC) fault-free simulations. In such simulations, a sample S of 100 circuit instances (including the path and the testing circuitry, but not the clock distribution network and the sensing circuit) has been generated accordingly to a normal distribution of main circuit parameters with a 10% standard deviation.

In a first step, we used MC simulations to find a value of T^0 ensuring that no false positive is produced even if T is decreased by 10% with respect to its nominal value ($T = 0.9T_0$). In this way we accounted for clock skews and uncertainties on the clock distribution network. Let us note that this choice is more optimistic than a 10% clock skew design tolerance, because we refer to a clock period which is smaller than the nominal one. In the proposed method, we have used MC simulations to select a pair of nominal values ($\omega_{in}^0, \omega_{th}^0$) ensuring that no false positive is produced for 10% worst case variations of the sensing circuit sensitivity (i.e. $w_{th} = 1.1\omega_{th}^0$). Let us note that we used a conservative approach giving priority to yield. Different strategies can be used to enhance test quality.

Based on such configuration, we performed MC simulations by injecting the fault with different values of resistance. In such an experiment, we have considered three possible values of T ($0.9T^0$, T^0 and $1.1T^0$) and w_{th} ($0.9\omega_{th}^0$, ω_{th}^0 and $1.1\omega_{th}^0$).

To summarize the achieved results, we define a DF coverage (C_{del}) as the fraction of IC instances that do not pass DF testing for a given value of T and R . In our method, the fault coverage (C_{pulse}) can be defined in the same way and it is a function of ω_{th} and R .

In the case of opens, we have considered an *external* ROP which, as noted in section 2, is expected to represent the worst case for our method. In particular, we considered a path including 7 gates and a fault affecting the output of the second one. Fig. 6 shows the fault coverage achieved by DF testing as a function of the open resistance R for different values of T .

Fig. 7, instead, shows the results achieved by means of the proposed method.

²Note that we do not perform a direct comparison with the methods in [5, 6, 7] because of the lack of experimental data.

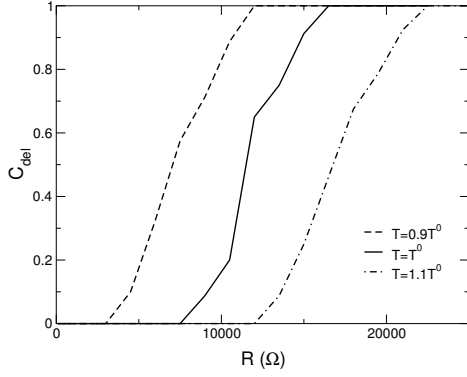


Figure 6. $C_{del}(R)$ for a ROP.

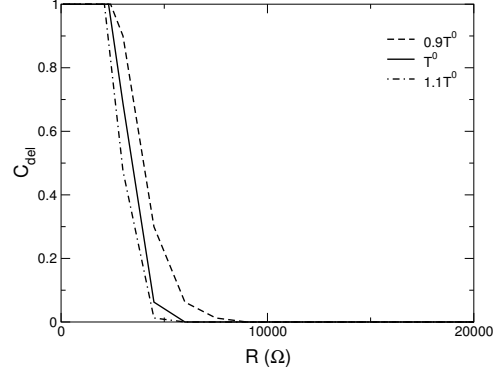


Figure 8. $C_{del}(R)$ for a resistive bridging.

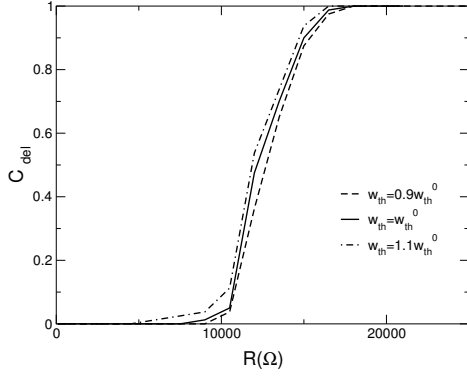


Figure 7. $C_{pulse}(R)$ for a ROP.

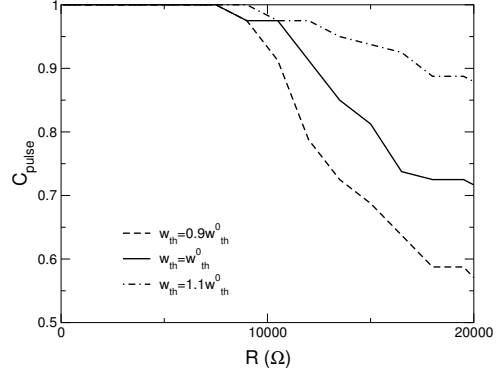


Figure 9. $C_{pulse}(R)$ for a resistive bridging.

Under nominal conditions, the two methods exhibit similar performance. Conversely, the performance of DF testing is affected by possible variations in the clock period, which are expected to be larger than in the considered method which performs a local analysis.

In the case of *external* BFs, we considered a fault affecting the second gate of the fault-free path used in the case of opens. Under nominal conditions, the critical resistance of such a fault is equal to 2100Ω. Above such a value, an additional delay is produced instead of a logic error. For the considered kind of load, this additional delay rapidly decreases with R [2]. As a consequence (Fig. 8), also C_{del} decreases rapidly with R . In the practice, the range of detectable resistances is slightly larger than that detectable under steady conditions.

In the case of pulses, instead, the injected pulse is likely to be dampened even if the additional delay produced when a transition is propagated through the faulty path is almost negligible.

Therefore (Fig. 7), the proposed method behaves much better than the considered kind of DF testing. Let us note that

for $R > 7500\Omega$ the size of the faulty pulse is very sensitive to fluctuations in the logic threshold of the fan-out gate, thus resulting in a significant sensitivity to variations in ω_{th}^0 .

5. Test generation and application issues

In order to detect a fault, we have to select a suitable kind of pulse (h or l) and a path including the fault site. The target is to optimize the pair $(\omega_{in}^0, \omega_{th}^0)$ which should maximize the range of detectable resistances while avoiding false positives.

This process strongly relies on the characterization of the set of candidate paths from the point of view of f_p under fault-free conditions and in the presence of IC's parameter fluctuations.

At this regard, Fig. 10 shows w_{out} as a function w_{in} for a path composed by 7 randomly selected gates with randomly selected load conditions. As can be seen, we have 3 different regions: 1) a region where the input pulse is completely dampened; 2) an asymptotic region exhibiting a linear behavior; 3) an attenuation region connecting regions 1) and 2).

When considering the fluctuations of IC's parameters, different values of w_{out} corresponding to different circuit sam-

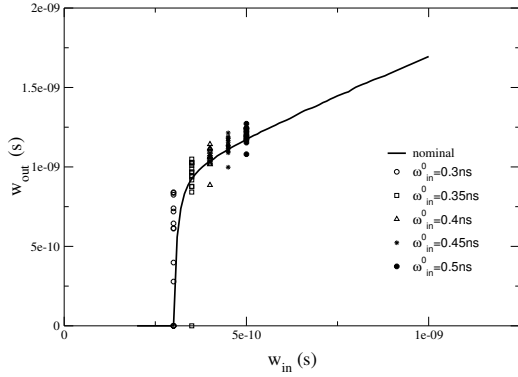


Figure 10. Relationship between w_{in} and w_{out} under nominal conditions (solid line) and for a set of different samples of the considered path.

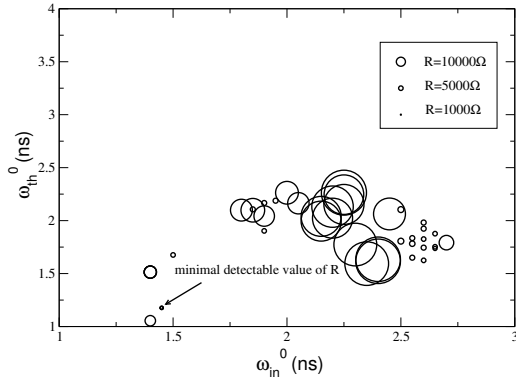


Figure 11. ω_{in}^0 , ω_{th}^0 and R_{min} for a ROP.

ples are related to the same value of w_{in} .

To analyze this problem, we used MC simulations with the same conditions considered in Sect. 4. In particular, a few values of w_{in} have been considered and the values of w_{out} corresponding to different samples of the same circuit are shown in Fig. 10. As can be seen, the attenuation region is rather sensitive to parameter fluctuations and it must be avoided if we do not want false positives. Therefore, we propose to use values of w_{in} at the beginning of region 3.

As an example, Fig. 11 shows the pairs $(\omega_{in}^0, \omega_{th}^0)$ computed accordingly to the above defined rule for a set of paths that include an *external* ROP in the ISCAS benchmark C432. For each path (and value of $(\omega_{in}^0, \omega_{th}^0)$), the figure shows a circle whose radius is proportional to the minimal value of R which can be detected by means of our method (R_{min}). The best path has a minimal detectable resistance of 35000 Ω and as shown in the figure, it should be searched between paths featuring low values of ω_{in}^0 and ω_{th}^0 .

In the case of more realistic circuits featuring several paths including the fault site, electrical level simulation is impractical and we need to operate at the logic level with timing accurate models such as that in [10] to study the propagation of pulses in a digital circuit.

Once w_{in} has been estimated for each path containing the fault site, the test generation process can sensitize the path providing the maximal range of detectable resistances. To this purpose, the basic algorithms used for path DF test generation can easily be modified.

These test conditions are ideal because they suppose that any value of ω_{in}^0 and ω_{th}^0 is available. In the practice, the on-chip testing circuitry will make available only a small number of values.

6 Conclusions

In this work we showed that pulse propagation can be used to detect the presence of resistive opens and bridgings affecting non critical paths. With respect to DF testing its accuracy does not depend on the clock distribution network. A logic level fault simulation tool is under development in order to apply our method to the case of large combinational networks.

References

- [1] Baker et al., "Defect-based delay testing of resistive vias-contacts," in *Proc. of ITC*, pp. 467 – 476, 1999.
- [2] M. Favalli et al., "Dynamic Effects in the Detection of Bridging Faults in CMOS ICs," *JETTA*, vol. 3, pp. 197 – 205, 1992.
- [3] P. Nigh, "The importance of on-line testing to enhance high-reliability performance," in *Proc. of ITC*, p. 1281, 2003.
- [4] B. P. Paul et al., "Temporal performance degradation under NBTI: estimation and design for improved reliability of nanoscale circuits," in *DATE*, 2006.
- [5] H. Yan and A. Singh, "Evaluating the effectiveness of detecting delay defects in the slack interval: a simulation study," in *Proc. of ITC*, pp. 242 – 248, 2004.
- [6] B. Kruseman et al., "On hazard-free patterns for fine-delay fault testing," in *Proc. of ITC*, pp. 213 – 222, 2004.
- [7] A. Singh, "A self-timed structural test methodology for timing anomalies due to defects and process variations," in *Proc. of ITC*, pp. (5.1) 1 – 6, 2005.
- [8] K. Bowman et al., "Impact of die-to-die and within-die parameter fluctuations for the maximum clock frequency distribution for Giga-scale integration," *IEEE JSSC*, vol. 37, no. 2, pp. 183 – 190, 2002.
- [9] C. Metra et al., "Self-checking detection and diagnosis for transient, delay and crosstalk faults affecting bus lines," *IEEE Trans. on Computers*, vol. 49, no. 6, pp. 560 – 574, 2000.
- [10] M. Omana et al., "A model for transient fault propagation in combinational logic," in *IEEE On-Line Test Symposium*, pp. 111 – 115, 2003.