Diagnosing Scan Chain Timing Faults through Statistical Feature Analysis of Scan Images

Mingjing Chen  Alex Orailoglu
CSE Department  UC San Diego
La Jolla, CA 92093
Email: {mjchen, alex}@cs.ucsd.edu

Abstract—Excessive test mode power-ground noise in nanometer scale chips causes large delay uncertainties in scan chains, resulting in a highly elevated rate of timing failures. The hybrid timing violation types in scan chains, plus their possible intermittent manifestations, invalidate the traditional assumptions in scan chain fault behavior, significantly increasing the ambiguity and difficulty in diagnosis. In this paper, we propose a novel methodology to resolve the challenge of diagnosing multiple permanent or intermittent timing faults in scan chains. Instead of relying on fault simulation that is incapable of approximating the intermittent fault manifestation, the proposed technique characterizes the impact of timing faults by analyzing the phase movement of scan patterns. Extracting fault-sensitive statistical features of phase movement information provides strong signals for the precise identification of fault locations and types. The manifestation probability of each fault is furthermore computed through a mathematical transformation framework which accurately models the behavior of multiple faults as a Markov chain. The fault model utilized in the proposed scheme considers the effect of possibly asymmetric fault manifestation, thus maximally approximating the realistic failure behavior. Simulations on large benchmark circuits and two industrial designs have confirmed that the proposed methodology can yield highly accurate diagnosis results even for complicated fault manifestations such as multiple intermittent faults with mixed fault types.

I. INTRODUCTION

Scan chain insertion has been widely considered one of the most effective design-for-testability techniques, and is incorporated in a majority of industrial designs. Despite its great benefits in improving fault coverage and easing test generation, the scan chain itself constitutes an important source of test failures. It has been reported that the scan circuitry occupies around 30% of the chip area [1] and possibly contributes up to 50% of the chip failures [2]. The diagnosis of scan chain failures thus becomes highly important for the yield learning of production chips.

Among various failure mechanisms, scan chain timing faults appear increasingly prevalent in nanometer scale designs with high clock frequencies [3]. Such a trend is mainly attributed to the elevated test-mode power-ground noises and the rapid shrinkage of design margins. Our characterization results of several 65nm chips show that the maximum IR-drop during scan phase can exceed 200mV, around 17% of the nominal operating voltage. With such a high noise level, the operational condition of the chip falls outside of the voltage range within which the designers close timing, thus introducing high delay uncertainties in scan chains. As a result, multiple faults with mixed timing violation types can exist in the failing scan chains. The mitigation of such failures necessitates accurate identification of the fault locations and precise understanding of the associated timing violation behaviors.

The manifestation of scan chain timing faults is in nature a probabilistic behavior, as multiple factors such as voltage and temperature fluctuation can disturb the timing in unpredictable ways. This imposes critical challenges in fault diagnosis, as the syndrome of intermittent faults becomes more ambiguous than permanent ones. With intermittent fault manifestation, the fault exclusion strategy employed in permanent fault diagnosis is no longer valid, and the number of possible fault hypotheses for each individual diagnostic pattern increases significantly. The existence of multiple fault locations and types further blurs the observed syndrome. Therefore, applying traditional approaches based on permanent fault assumptions to timing failure diagnosis would lead to inaccurate and even misleading results.

Although the intermittent timing faults distort each scan pattern in an unpredictable manner, the statistical characteristics of multiple failure patterns still provide strong signals that can be utilized to guide the fault diagnosis. From a statistical perspective, a scan chain timing fault would move the affected scan image forward or backward, depending on the specific type of the timing violation. Although such an image movement effect becomes less strict in an intermittent fault scenario, a strong correlation can still be observed between the original image and the distorted one with a shifted phase. Therefore, any abrupt phase shift between the two images would signify the location and the type of a new timing fault in the scan chain. In light of this observation, we propose, in this paper, a methodology capable of diagnosing both permanent and intermittent scan chain timing faults. A metric is defined to extract clear signals of phase shift from the correlation information of scan images, enabling the accurate identification of fault sites through an efficient algorithm.

In addition to the fault locations and types, the manifestation probabilities of intermittent faults are of great interest to designers, as such information helps prioritize the failure spots for cost-effective design optimization. We further propose an accurate and computationally-efficient mathematical framework for estimating the fault probabilities. The proposed technique models the scan image distortion process as a Markov chain, with the transition matrices capturing the impact of fault probabilities on scan pattern distributions. The fault model employed in the proposed framework maximally approximates the realistic fault behavior by taking into account the possibly asymmetric manifestation probabilities of rising and falling faults. An accurate estimation of the manifestation probabilities can be attained by solving the equation sets associated with the proposed Markov model.

The proposed technique requires neither hardware modification nor special test equipment, and can therefore be applied to a wide range of designs and test environments. The failure syndromes can be collected by applying diagnostic patterns using the standard scan-based test schemes. The algorithms for fault diagnosis can be fully automated, thus enabling cost-effective integration into industrial failure analysis tools.

The remainder of this paper is organized as follows. Section II provides a brief review of the previous approaches in scan chain diagnosis. The preliminary information for the proposed diagnosis
scheme is introduced in Section III. The algorithm that identifies the fault locations and types through image correlation analysis is detailed in Section IV. Section V discusses the mathematical framework for calculating the fault manifestation probabilities. Simulation results on a number of large benchmarks and two industrial designs are reported in Section VI. A brief set of conclusions is provided in Section VII.

II. PREVIOUS WORK

A considerable amount of approaches have been proposed in the literature to address the scan chain diagnosis problem. A comprehensive survey of this area has been provided in [4]. In general, these approaches can be categorized into three groups.

- **Tester-based diagnosis:** Techniques in this category utilize physical failure analysis equipment to identify the failing locations in the scan chain [5], [6]. These approaches typically provide good diagnostic quality. Nonetheless they necessitate a time-consuming debugging process and possibly expensive equipments, thus having a limited application range.

- **Hardware-modification based diagnosis:** A number of techniques incorporate specialized scan chain designs to enhance diagnosability [7], [8]. These techniques reduce the algorithmic complexity of scan chain diagnosis, yet at the cost of extra hardware overhead and increased design complexity.

- **Software-based diagnosis:** This category of techniques aims to identify fault sites through algorithmic analysis of the failing data collected during regular scan operation [3], [9], [10], [11], [12], [13], [14]. Such techniques typically incur much lower cost compared to the tester/hardware-based approaches, thus having a wider application in general designs. Due to this reason, the proposed work focuses on developing innovative techniques in this domain.

Most of the software-based techniques target permanent faults in the scan chain. Such approaches identify the candidate faulty cells through fault simulation of regular test patterns [9], [10], [13] or special diagnostic patterns [12]. The observed values on the fault-sensitive bits of the pattern can indicate the upper and lower bounds of each fault site. Good diagnostic results can be attained with these techniques under the permanent fault assumption.

Due to the ambiguity induced by the intermittent manifestation of timing faults, the traditional simulation-based techniques might lead to inaccurate and even misleading results. This problem is illustrated in the example shown in Figure 1. Let us assume that an intermittent hold-time fault exists between scan cells 4 and 3. The three scan bits highlighted with underlines are sensitive to hold-time. After scanning out the pattern, sensitive bit 7 is corrupted by the fault, whereas sensitive bits 5 and 2 match the expected values. This indicates a candidate fault range between cells 7 and 5 under the permanent fault assumption, which fails to cover the actual fault site.

Only a few proposals have been presented to resolve the challenge imposed by intermittent faults. The approach proposed in [3] employs a Bayesian decision model to identify the location that has the highest fault probability. However, this approach can only be applied to scan chains with a single fault. The approach in [11] utilizes the signal probability computation to search for the fault locations that maximally account for the observed pattern. Nonetheless it makes assumptions about the fault manifestation probability a priori, which might slightly degrade the diagnostic quality. Another technique [14] uses error count as an indicator of fault locations. But this technique provides no information about fault types and manifestation probabilities. In order to gain a deeper understanding of the increasingly complex failure conditions seen in today’s large designs, a comprehensive methodology is needed to analyze scan chains with mixed fault types and various manifestation probabilities.

III. PRELIMINARIES

A set of preliminary information related to the proposed work is outlined in this section.

A. Fault models

The proposed technique analyzes the scan chain timing failures based on two fault models, namely the setup-time violation and hold-time violation. Since a timing fault can cause incorrect toggles in two possible directions, we assign to each fault type two parameters that define the manifestation probabilities of rising and falling faults, respectively. Table I provides a summary of the proposed fault model. Compared to traditional models which typically assume unidirectional or equal probability for both directions, the proposed model approximates the realistic fault behaviors better, as it is able to reflect the asymmetric manifestation probabilities induced by nonideal design and fabrication.

B. Application scheme of diagnostic patterns

The failing scan chains can be easily identified by performing a flush test [4]. Hence the proposed work mainly focuses on the fine-grained diagnosis down to the scan cell level after attaining the failing scan chain information.

As outlined in the introduction, the proposed technique identifies the faulty cells by analyzing the scan pattern image of the failing scan chain. Therefore, the proposed diagnostic scheme first applies a number of timing violation immune vectors to the chips-under-diagnosis in order to create the scan image. The application of these vectors follows the standard scan-based test application scheme, namely, the scan-in→capture→scan-out flow. The timing violation immune vectors consist of all 0’s or all 1’s in the failing scan chains and random patterns in fault-free scan chains. The application of such vectors introduces no errors in the scan-in and capture phases. Hence the expected scan-out patterns can be attained without ambiguity through logic simulation. As the capture step creates in the failing scan chains toggling patterns that are sensitive to timing faults, the captured responses are no longer immune to the scan chain faults and would be corrupted during scan-out. The difference between the observed scan-out patterns and the expected ones provides information regarding the faulty cells that would be identified during the post-analysis step.
the response values that are actually observed at the scan-out pin of the observed image of a scan chain is formed by the correct response values before the scan-out step introduces any errors.

**Expected image:** The expected image of a scan chain is formed by the response values that are actually observed at the scan-out pin of this chain.

**C. Scan chain images**

Applying multiple diagnostic vectors would generate a set of distinct responses in each failing scan chain. The union of these response patterns forms an image of the associated failing scan chain. Formally, the image of a scan chain can be represented in a matrix form, with each column denoting a scan cell and each row corresponding to the response of a diagnostic vector. Figure 2 presents an illustrative scan image of a scan chain. Several images can be defined to facilitate the analysis of the failure syndromes.

**Expected image:** The expected image of a scan chain is formed by the correct response values before the scan-out step introduces any errors.

**Observed image:** The observed image of a scan chain is formed by the response values that are actually observed at the scan-out pin of this chain.

**IV. FAULT LOCATION AND TYPE IDENTIFICATION**

**A. Basic idea**

The impact of scan chain timing faults can be extracted by analyzing the behavior of each type of timing violations. The manifestation of a hold-time fault would speed up by one cycle the propagation of the scan value that passes through the fault site as the input of the affected scan cell toggles too fast to meet the hold-time constraint, whereas a setup-time violation delays the propagation of the scan value by one cycle due to the slow toggling of the affected signal. From a scan image perspective, such a speedup or delay effect can be interpreted as a forward or backward movement of the affected portion of the scan image. For instance a scan sequence of “101001011” can be converted to “X10100111” by a hold-time fault between the second and third bits from the right. It can be observed that the phase of the sub-sequence “1010011” is moved one bit forward. If a portion of the scan image passes through multiple faults during the scan-out stage, its phase would be moved back and forth multiple times. Such an accumulative distortion effect transforms the expected image to the observed image. If we compare these two images by traversing them from the scan-out to the scan-in, the detection of any change of the phase skew between them signifies the existence of a new timing fault. The type of the timing fault can furthermore be identified by examining the direction of the phase movement.

To precisely extract the phase movement information, we propose to monitor the column-wise correlation between the expected and observed images, as any phase movement would strongly change the correlation levels of the scan image columns under comparison. In a fault-free case, the expected image exactly matches the observed one, thus leading to a full correlation between the columns with the same index. However, when a column, i, of the expected image passes through a hold-time (setup-time) fault, the correlation between it and column i−1 (i+1) of the observed image would significantly increase as a result of the phase movement behavior. Such an abrupt change of the column correlation provides clear signals for identifying phase movement.

As the columns of scan images are essentially binary vectors, the metric proposed in [15] is employed to estimate the column-wise correlation. Such a metric has been widely utilized in resolving pattern recognition problems and has proven highly effective in assessing binary vector similarity. Formally, let $S_{ij}$ ($i, j \in \{0, 1\}$) be the number of occurrences of bit pair $(i, j)$ at the corresponding positions of two binary vectors. Then the correlation between the two vectors is defined as follows.

\[
\text{corr}(c_i, c_j) = \frac{S_{ii} + S_{jj} - 2S_{ij}}{\sqrt{(S_{00} + S_{01})(S_{00} + S_{11})(S_{00} + S_{10})(S_{00} + S_{01})}}^{1/2}
\]

Figure 3 shows an example of column correlation computation. The aforementioned binary vector correlation metric shares similar properties with the correlation coefficient of random variables. It has a value range of $[-1, 1]$, with the value 1 indicating a perfect positive correlation and −1 denoting the correlation between complementary vectors. If the two vectors under comparison are independent and random enough, the correlation between them would be around 0, as the four types of bit pairs examined in this metric would have similar occurrence probabilities, resulting in the numerator of the metric approximately equaling 0.

The principle of detecting phase movement through monitoring column-wise correlation can be generically utilized for both the permanent and intermittent timing fault diagnosis, as discussed in Sections IV-B and IV-C, respectively.

**B. Permanent fault diagnosis**

A permanent timing fault in the scan chain would result in strict phase movements of affected scan image columns, as the propagation of scan values at the fault site will always be sped up or delayed. This in turn results in an abrupt change of the column-wise correlations. The columns to the right of the fault location are not impacted by the fault during the scan-out stage. Therefore, such columns in the expected image should have a full correlation with the corresponding columns in the observed image. Nonetheless, the fault would impact the columns to the left of the fault site, thus changing the correlation relationship. More specifically, a single hold-time violation would move the affected column i in the expected image one bit forward to the position of $i - 1$ in the observed image, resulting in a full correlation between them. Similarly, a single setup-time violation would result in a full correlation between column i in the expected image and column i + 1 in the observed image. From the aforementioned analysis, it can be seen that the position where the correlation relationship starts to change is the fault location. If the correlation change at the fault site signifies a forward phase movement, the fault type can be ascertained to be a hold-time violation. Otherwise, a setup-time violation is indicated by a backward phase movement.

In the case of multiple permanent faults, the accumulative phase movements need to be considered. Figure 4 illustrates a double-fault example with mixed fault types. The two fault sites naturally
partition the scan image of this faulty scan chain into three segments. During the scan-out stage, these three segments pass through different numbers of faults, thus experiencing distinct fault accumulation effects. Segment 1 passes through no faults, thus exhibiting no phase skew between the expected and observed images. Segment 2 is moved forward by a hold-time fault, thus having a phase skew of 1. The phase skew returns back to 0 in Segment 3, as the setup-time fault would change these correlations the other way around. In light of this observation, a weighted correlation sum is defined to extract the phase movement information, as shown in the following equation.

\[ S_{corr} = -2 \times corr(i, i - v + 1) - 1 \times corr(i, i - v) + 1 \times corr(i, i - v - 1) + 2 \times corr(i, i - v - 2) \]  

Intermittent fault diagnosis can be efficiently performed with the guidance of the weighted correlation sum. If we sweep the scan images from scan-out to scan-in and extract \( S_{corr} \) for each column, it can be observed that the \( S_{corr} \) value remains relatively stable in scan chain portions that contain no faults. However, a sharp increase

\[ corr(i,j) \] correlation between column i in the expected image and column j in the observed image

\[ corr(i,i) \text{ correlation parameters:} \ corr(i,i-2), corr(i,i-1), corr(i,i) \]

\[ \text{Fault location: between scan cells 28 & 27} \]

\[ \text{Fault type: hold-time violation} \]

\[ S_{corr} \] weighted correlation sum

\[ corr(i,i-1), corr(i,i), corr(i,i-2), corr(i,i-3) \] correlation values

\[ corr(i,i-1) \text{ correlation values:} \ corr(i,i-2), corr(i,i-1), corr(i,i) \]
(decrease) in $S_{corr}$ would occur at a hold-time (setup-time) fault site, providing clear signals of fault locations and types.

It is important to note that the set of correlation parameters incorporated in the $S_{corr}$ computation needs to be updated whenever a newly detected fault changes the value range of the phase skew. The adaptive update of the correlation parameters enables the capture of the most evident change in correlation values, thus providing sufficient signal strength to differentiate the fault-induced signal from random fluctuations. Figure 6 presents the simulated curves for a scan chain that contains two intermittent timing faults. A sharp change of the curve can be clearly observed at the fault locations, providing high levels of diagnostic resolutions.

V. MANIFESTATION PROBABILITY COMPUTATION

In order to evaluate the criticality of an intermittent fault for cost-effective design optimization, the fault manifestation probability needs to be ascertained. Attaining such information necessitates mathematically modeling the impact of manifestation probability on statistical features that can be observed in the scan images. The distribution of two-bit patterns in each segment of the scan images constitutes a fault-sensitive feature, as the magnitude of distribution distortion is a function of fault manifestation probabilities. Moreover, if the manifestation probabilities of a fault are different in distinct directions, an asymmetric change in the two-bit pattern distribution can also be observed, enabling good characterization of the realistic fault behavior. Therefore, a mathematical framework is proposed in this work to extract the fault manifestation probability information from the two-bit pattern distribution feature.

The manifestation of any intermittent timing fault would distort the two-bit pattern distribution of a scan image by probabilistically transforming one type of two-bit pattern to another type. If a segment of the scan image passes through multiple faults, such a random distortion process is essentially a Markov chain, as the next distribution resulting from a new fault only depends on the current distribution. Such a process is illustrated in the double-fault example shown in Figure 7. Segment 3 of the expected image is first transformed to an intermediate state by the setup-time violation, and then to the observed image by the hold-time violation.

The impact of setup-time and hold-time violations on the two-bit pattern distribution can be modeled as two transition matrices of the Markov chain, as shown in Figure 8. These two matrices depict the probability of a fault transforming a certain two-bit pattern to another pattern. For example, in order for a hold-time fault to convert a 00 pattern to 10, the left neighboring bit of this pattern must have a value of 1 and a fast-to-rise type of hold-time violation must manifest itself when this sequence passes through the fault location. This results in a transition probability of $P_1 M_{ff}$. The probabilities for other types of transitions can be derived in a similar manner. The transition probabilities are organized in the form of right transition matrices, with the rows summing to 1.

In the example shown in Figure 7, Segment 2 of the scan image passes solely through the hold-time fault. If we represent the two-bit pattern distributions of expected and observed Segment 2 images as two row vectors, namely, $D_e(Seg_2) = [P00 \; P01 \; P10 \; P11]$ and $D_o(Seg_2) = [P00 \; P01 \; P10 \; P11]$, then the impact of the fault on Segment 2 can be approximated by a linear transformation.

$$D_e(Seg_2) \ast H = D_o(Seg_2)$$

Since the pattern distributions of the expected and observed images are known information, the only variables in this equation consist of the fault manifestation probabilities, $M_{sf}$ and $M_{ff}$, which can be resolved by minimizing the least-square norm of the difference between the two sides of the equation.

$$\min ||D_e(Seg_2) \ast H - D_o(Seg_2)||$$

s.t. $0 < M_{sf} < 1, 0 < M_{ff} < 1$

Solving this nonlinear programming problem enables us to identify the probability values that can maximally account for the observed distribution distortion effect in Segment 2. A number of numerical algorithms, such as the trust region approach [16], can be utilized to attain solutions with good accuracy.

Once the manifestation probabilities of the rightmost fault are resolved, the proposed scheme furthermore examines the second fault from the right by modeling the distribution variation in Segment 3. In the example shown in Figure 7, this segment sequentially passes through the setup-time and hold-time faults. Hence the transformation process can be depicted by the following equation.

$$D_e(Seg_3) \ast S \ast H = D_o(Seg_3)$$

Since the manifestation probabilities for the rightmost fault are already known, the manifestation probabilities for the second fault, $M_{sr}$ and $M_{sf}$, can be efficiently computed by solving the nonlinear programming problem constructed with Equation 4. Under a multiple fault situation, the manifestation probabilities of all the faults can be efficiently attained by iteratively applying the aforementioned technique.
VI. EXPERIMENTAL RESULTS

The proposed methodology has been implemented using the C and matlab programming languages. Fault diagnosis simulations have been performed on large ISCAS89 and ITC99 benchmarks and two industrial designs in order to evaluate the effectiveness of the proposed scheme. The two industrial designs employed in our simulation consist of a floating-point unit (FPU) and the top-level logic of a memory system (MEMSYS). The number of scan faults in these two designs is 13803 and 45766, respectively. A total of 20 scan chains are constructed in each benchmark, and 40 scan chains are inserted in each industrial design. Each circuit is assumed to have one faulty scan chain in our simulation. To examine the complicated fault behavior, a set of three timing faults with randomly assigned fault locations and types is injected into the faulty scan chain. For each circuit, 200 randomly generated timing violation immune test stimuli are utilized to create the scan image of the faulty scan chain. Hence the scan image consists of 200 rows. Both the permanent and intermittent fault scenarios have been examined in our simulation.

In the case of permanent faults being injected, the simulation shows that the proposed method always guarantees a perfect diagnostic resolution. This is because the incremental phase movement induced by permanent faults always has a step length of 1. Such a highly regular behavior can be captured by the proposed technique with no ambiguity, leading to a precise identification of the fault locations and types.

The proposed technique also delivers highly accurate diagnostic results for the more challenging case of intermittent faults, as summarized in Table II. The diagnosis result (Diag.) is compared to the fault injection parameters (Inj.) in terms of location, type and manifestation probability2. For most faults, a high diagnostic resolution is delivered, as the resulting candidate fault window typically contains fewer than 3 scan cells. The proposed technique always guarantees the correct identification of the fault type throughout the simulation. The estimated fault manifestation probabilities approximate the injected fault quite well, as the error between the injected and estimated probabilities is typically less than 0.1. More importantly, it can be seen that the estimated probability values preserve the relative magnitude relationship of the injected fault manifestation probabilities, providing insights regarding the relative criticality of different faults.

VII. CONCLUSION

In this work, we investigate the challenge of diagnosing multiple scan chain timing faults with mixed fault types and various manifestation probabilities. A novel approach based on statistically monitoring the phase skew of scan images is proposed to identify the fault locations and types. Column-wise correlations of scan images are extracted to generate strong signals regarding phase movement, thus maximally eliminating the ambiguity induced by probabilistic fault manifestations. We further propose a mathematical framework that models the impact of intermittent faults as a Markov chain, thus enabling an accurate estimation of the fault manifestation probabilities. The proposed technique is completely compatible with existing scan test schemes, imposing no design overhead whatsoever. Integration of the proposed methodology into an industrial scheme enables efficient silicon debugging, significantly speeding up the design and test optimization cycles.

2The indices of scan cells between which the injected/diagnosed fault resides are listed in the Location column. In the Type column, HT stands for hold-time fault, and ST for setup-time fault. In the Prob. column, the manifestation probabilities of each fault are presented in the form of \( P(\text{rising fault}) / P(\text{falling fault}) \).