Compact Dictionaries for Fault Diagnosis in BIST

Chunsheng Liu and Krishnendu Chakrabarty

Department of Electrical & Computer Engineering
Duke University, Durham, NC 27708
E-mail: {mliu, krish}@ee.duke.edu

Abstract

We present a new technique for generating compact dictionaries for cause-effect diagnosis in BIST. This approach relies on the use of three compact dictionaries: (i) $\mathcal{D}_1$, containing compacted LFSR signatures for a small number of patterns and faults with high detection probability, (ii) an interval-based pass/fail dictionary $\mathcal{D}_2$ for the BIST patterns and for faults with relatively lower detection probability, and (iii) $\mathcal{D}_3$, containing compacted LFSR signatures for clean-up ATPG vectors and random-resistant faults. We show that $\mathcal{D}_2$, which is two orders of magnitude smaller than a maximal-resolution pass/fail dictionary, provides nearly the same diagnostic resolution as an uncompacted dictionary. We also show that by using a 16-bit LFSR signature for $\mathcal{D}_1$ and $\mathcal{D}_2$, we obtain three orders of magnitude reduction in dictionary size, yet nearly no loss in diagnostic resolution.

1 Introduction

Fault diagnosis is necessary for the identification of manufacturing defects and for yield learning. One approach to diagnosis is based on the use of fault dictionaries, which alleviate the need for repeated fault simulation [1, 2]. However, as designs grow in complexity, dictionary-based diagnosis becomes infeasible due to prohibitively large dictionary sizes. A full-response dictionary can require Gbits of storage for today's integrated circuits (ICs). An alternative approach is to record only pass/fail information in a pass/fail dictionary.

A number of techniques have recently been proposed for reducing dictionary size [1-6]. An important consideration in dictionary compaction is that of diagnostic resolution, which is determined by the number of modeled faults that correspond to the same entry in a compact fault dictionary. An effective dictionary compaction scheme should yield a small dictionary without significant loss in diagnostic resolution.

Dictionary compaction is achieved through the selection of appropriate dictionary organization and encoding schemes [1, 2, 4, 5, 6, 8]. A drawback of these techniques is that they are based on complex organization and encoding schemes, which increase computation and post-processing costs. In addition, these techniques rely on irregular data structures and data representation, which render bit-packing schemes ineffective. Finally, the loss in diagnostic resolution is often disproportionately large compared to the amount of compaction achieved.

The above problems have recently motivated new approaches to reduce dictionary size and increase diagnostic resolution [8]. However, it is difficult to analytically predict the amount of compaction relative to the diagnostic resolution for these approaches. Most prior work on cause-effect diagnosis has been targeted at fault dictionaries based on deterministic vectors for external testing. Recently, built-in self-test (BIST) has gained increasing acceptance as a test solution. However, a problem with this approach is that a BIST signature does not contain enough diagnostic information.

Cause-effect analysis for BIST is also challenging due to prohibitively large dictionary sizes. The number of pseudorandom test vectors used in BIST is at least an order of magnitude larger than the compacted deterministic test sets used for external testing. Recently, the use of intervals of test vectors was proposed to identify failing vectors in BIST [9]. Here, we exploit the concept of intervals of test vectors to generate compact fault dictionaries.

The proposed approach attemps to make diagnosis more efficient by classifying the faults in the underlying fault model and using a separate dictionary for each fault type. First, we create a compacted full-response dictionary $\mathcal{D}_1$ for a small number of vectors, e.g. the set of vectors in the first interval. We use this dictionary to diagnose faults with high detection probability. For a circuit with O outputs (including scan cells) and N vectors, a typical full-response dictionary contains an ON-bit entry for every fault. This makes a full-response dictionary prohibitively large. We overcome this problem by simulating an LFSR to compact each ON-bit entry in $\mathcal{D}_1$ to an S-bit signature during dictionary creation. For example, for O = 100, N = 1000, and S = 16, each dictionary entry is reduced from 10^5 bits to 16 bits, which represents several orders of magnitude compaction. During fault diagnosis, an S-bit LFSR is used to generate a signature that is compared with the entries in $\mathcal{D}_1$.

We next use the interval-based technique to generate a highly-compacted dictionary $\mathcal{D}_2$ over all BIST vectors to target the random-testable faults that have relatively low detection probability. We use an interval-based dictionary instead of an LFSR-based dictionary here because the former provides valuable information on failing vectors, which in turn allows us to perform targeted effect-cause analysis. Moreover, an LFSR-based dictionary is generated using all the output values hence it is practical only for a small number of test vectors.

Finally, a third dictionary $\mathcal{D}_3$ is used to diagnose the remaining random-resistant faults that are targeted by "clean-up" ATPG vectors. This dictionary has the same organization as $\mathcal{D}_1$.

* This research was supported in part by the National Science Foundation under grants CCR-9875324 and CCR-0204077.

0-7695-1881-8/03 $17.00 © 2003 IEEE

Compared to a traditional pass/fail dictionary, the interval-based approach leads to two orders of magnitude reduction in storage with negligible loss of diagnostic resolution. As interval length increases, we show that there is no loss of resolution until a compression threshold is reached on most faults. This threshold serves as an important parameter in the design of interval-based compact fault dictionaries.

The organization of the rest of the paper is as follows. In Section 2, we describe the proposed interval-based approach for diagnosis. In Section 3, we present the 3-stage BIST diagnosis approach and outline its use with scoring algorithms for diagnosing unmodeled faults. Finally, in Section 4, we demonstrate the effectiveness of the proposed diagnosis approach for BIST by determining dictionary sizes and diagnostic resolution for the large ISCAS-89 benchmark circuits.

2 Interval-based dictionary

Let $\mathcal{T}$ be an ordered set of patterns that is applied to the circuit under test. An interval I corresponds to a subset of consecutive vectors from $\mathcal{T}$. We divide $\mathcal{T}$ into a set of intervals I_1, I_2, …, I_N such that $\mathcal{T} = I_1 \cup I_2 \cup \cdots \cup I_N$. As described in detail in [9], the MISR used to collect signatures is reset at the start of every interval such that a faulty interval does not affect the pass/fail status of subsequent intervals.

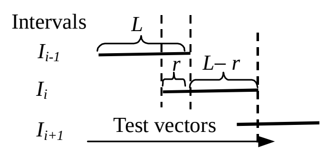
Figure 1(a) illustrates three consecutive intervals. The test sequence is split into intervals of length L and overlap r. Note that I_i represents the i^{th} interval. We assume that error masking (aliasing) can be neglected such that if an interval contains one or more failing vectors, the corresponding BIST signature is different from the fault-free signature.

We illustrate the interval-based method via a compact 2-D pass/fail dictionary in Figure 1(b). Let $\mathcal{F} = \{F_1, F_2, \ldots, F_M\}$ be the set of modeled faults. These faults correspond to rows in the compact dictionary, and the intervals make up the columns. If a fault F_i is detected by a vector in interval I_j, i.e. interval I_j fails due to fault F_i, the corresponding (i,j) entry in the fault dictionary is set to 1, otherwise it is set to 0. A given set of failing and non-failing intervals corresponds to a particular bit pattern that can often be mapped to a set of candidate faults from $\mathcal{F}$.

Since we set the number of intervals to be one to two orders of magnitude smaller than the number of test vectors, the interval-based pass-fail dictionary is significantly smaller than a maximal-resolution pass/fail dictionary. The savings in storage must now be weighed against the possible loss of diagnostic resolution. We address this issue in the next two subsections.

2.1 Analytical results

Let the number of test vectors applied to the circuit under test be N and let the number of modeled faults be M. The size (in bits) of a maximal-resolution 2-D pass/fail dictionary is then simply NM. If we use an interval-based 2-D pass/fail dictionary with interval length L and zero overlap, the size (in bits) of the compact dictionary is NM/L. Thus we are able to reduce the size of the dictionary by a factor of L. Since the processing time for dictionary-based diagnosis is proportional to the number of entries in the dictionary, the proposed approach significantly improves the efficiency of the diagnosis process.



Fault \ Interval	I_1	I_2	• • •	I_N
F_1	0	0	• • •	1
F_2	0	1	• • •	1
• • •			• • •	
F_M	1	1	• • •	0

(a) Test patterns in intervals (b) An interval-based 2-D dictionary

Figure 1. Interval-based dictionary.

We now analyze the diagnostic resolution of the interval-based compact dictionary. First, we note that even in a 2-D pass/fail dictionary based on the N test vectors, not all faults in F can be distinguished from each other. Let M_E ($M_E \leq M$) be the number of distinct rows in this maximal-resolution pass/fail dictionary. Each distinct row of this dictionary corresponds to an *equivalence class*; the faults in any equivalence class cannot be distinguished from each other using a pass/fail dictionary.

An interval-based compact dictionary is most effective if the number of equivalence classes M'_E ($M'_E \leq M_E$) in it is only slightly less than M_E. Note that M'_E depends to a large extent on the number of intervals N_I, which in turn depends on the length of the interval L. As L is increased, i.e. as N_I is decreased, the number of columns in the dictionary decreases, thereby leading to greater compaction. At first, there is no loss in diagnostic resolution since the use of longer intervals simply eliminates the redundant information in the dictionary based on pseudorandom vectors. However, loss of diagnosis resolution is inevitable ($M'_E \leq M_E$) after L exceeds a certain threshold. This property is formalized by the following theorem.

Theorem 1: Let M_E be the number of fault equivalence classes in the maximal-resolution pass/fail dictionary. Let M'_E be the number of equivalence classes in an interval-based dictionary with interval length L and $N_I = N/(L - r)$ intervals. If $N_I < \lceil \log_2(M_E + 1) \rceil$ then $M'_E \leq M_E$. In other words, $M'_E \leq M_E$ if $L > N/\lceil \log_2(M_E + 1) \rceil + r$. (Proof omitted.)

The above bounds on L and N_I are sufficient conditions for loss of diagnostic resolution, but they are not necessary. In practice, we observe loss in diagnostic resolution for values of L that are smaller than the bound given by Theorem 1.

Theorem 2: If two faults f_1, $f_2 \in \mathcal{F}$ belong to the same equivalence class in a maximal-resolution pass/fail dictionary, they belong to the same equivalence class in an interval-based pass/fail dictionary. (Proof omitted.)

Theorem 2 leads us to conclude that if $L' > L$, the number of candidate faults obtained with interval length L' is at least equal to the number of candidate faults obtained with interval length L. This leads to the somewhat expected result that the diagnostic resolution for $\mathcal{ID}$ can never be better than the diagnostic resolution for $\mathcal{MD}$. However, an interesting observation is that this relationship does not hold for two interval-based dictionaries.

We next investigate the variation of diagnostic resolution with change in interval length L. Consider two faults f_1 and f_2 with detection probabilities α and β, respectively. A high

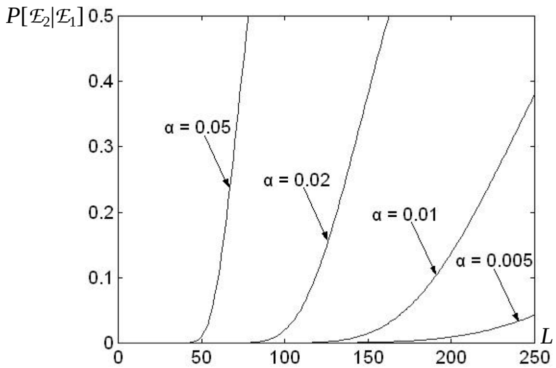


Figure 2. Variation of $P[\mathcal{E}_2|\mathcal{E}_1]$ with interval length L for several values of detection probability α ($\alpha = \beta$).

detection probability indicates an easy-to-detect fault and a low detection probability indicates a hard-to-detect fault. Consider the following three events:

1. $\mathcal{E}_1$: f_1 and f_2 are not in the same equivalence class in $\mathcal{MD}$;
2. $\mathcal{E}_2$: f_1 and f_2 are in the same equivalence class in $\mathcal{ID}$;
3. $\mathcal{E}_3$: f_1 and f_2 are in the same equivalence class in $\mathcal{MD}$.

Let the corresponding probabilities of occurrence of these events be $P[\mathcal{E}_1]$, $P[\mathcal{E}_2]$ and $P[\mathcal{E}_3]$, respectively. These probabilities can be expressed as follows:

$$P[\mathcal{E}_1] = 1 - P[\mathcal{E}_3] \qquad (1)$$

$$P[\mathcal{E}_2] = ((1-(1-\alpha)^L)(1-(1-\beta)^L) + (1-\alpha)^L (1-\beta)^L)^{N_I} \qquad (2)$$

$$P[\mathcal{E}_3] = (\alpha\beta + (1-\alpha)(1-\beta))^N \qquad (3)$$

Next we consider the conditional probability $P[\mathcal{E}_2|\mathcal{E}_1]$, which is defined as the probability that f_1 and f_2 are in the same equivalence class in $\mathcal{ID}$, given that they are not in the same equivalence class in the corresponding $\mathcal{MD}$. Note that $P[\mathcal{E}_2|\mathcal{E}_1]$ reflects the likelihood that two faults are distinguishable in $\mathcal{MD}$ but indistinguishable in $\mathcal{ID}$. A smaller value of $P[\mathcal{E}_2|\mathcal{E}_1]$ indicates better diagnostic resolution.

$P[\mathcal{E}_2|\mathcal{E}_1]$ can be derived as follows using the notion of joint probabilities and marginal probabilities:

$$\begin{aligned} P[\mathcal{E}_2|\mathcal{E}_1] &= \frac{P[\mathcal{E}_2\mathcal{E}_1]}{P[\mathcal{E}_1]} = \frac{P[\mathcal{E}_1|\mathcal{E}_2]P[\mathcal{E}_2]}{P[\mathcal{E}_1]} = \frac{(1-P[\mathcal{E}_3|\mathcal{E}_2])P[\mathcal{E}_2]}{P[\mathcal{E}_1]} \\ &= \frac{(1-P[\mathcal{E}_3\mathcal{E}_2]/P[\mathcal{E}_2])P[\mathcal{E}_2]}{P[\mathcal{E}_1]} = \frac{(1-P[\mathcal{E}_3]/P[\mathcal{E}_2])P[\mathcal{E}_2]}{P[\mathcal{E}_1]} \\ &= \frac{P[\mathcal{E}_2] - P[\mathcal{E}_3]}{1-P[\mathcal{E}_3]} \end{aligned}$$

Here we have utilized the relationship $P[\mathcal{E}_3\mathcal{E}_2] = P[\mathcal{E}_3]$, which follows from Theorem 2. Note that although this analysis addresses only a pair of faults, the conclusion appears to be valid for larger sets of faults, as is justified by experimental results later in this section.

Figure 2 shows the variation of $P[\mathcal{E}_2|\mathcal{E}_1]$ with L for various values of α. We assume here that the faults f_1 and f_2 have the same detection probability, i.e. $\alpha = \beta$. We also assume that a total of 1,500 pseudorandom vectors are applied to the circuit under test, i.e. $N = 1{,}500$. As expected, we find that for given values of α and N, using a smaller interval results in better diagnostic resolution. The surprising observation is that $P[\mathcal{E}_2|\mathcal{E}_1]$ remains almost unchanged for a wide range of L up to a certain threshold. This is potentially significant since it implies that L can be increased (the dictionary can be compacted) without significantly affecting diagnostic resolution. For example, for $\alpha = 0.01$, we reach this threshold at about $L = 140$, which implies that no diagnostic resolution is lost with a compaction factor of 140 using our interval-based method.

Circuit	$T_{\mathcal{MD}}$ (hours)	$T_{\mathcal{ID}}$ (minutes)			
		$L = 40$	$L = 80$	$L = 120$	$L = 160$
s13207	4	4.5	2.5	1.5	1.3
s15850	10	7.5	5.6	2.2	1.8
s35932	89.7	18.5	12	5.2	4.6
s38417	54.2	55	25	14.8	9.5
s38584	75.7	57	28	14.3	9.4

Table 1. Dictionary generation time for maximal-resolution and interval-based dictionaries.

2.2 Initial experimental results

We now present some initial simulation results on interval-based dictionaries. For each benchmark circuit, we first create a maximal-resolution pass/fail dictionary for 1,500 pseudorandom vectors and all single stuck-at faults. We then generate interval-based pass/fail dictionaries for various values of L. We use a Sun Ultra 10 workstation with a 333 MHz processor and 256 MB of RAM.

In order to evaluate diagnostic resolution for a benchmark, we randomly inject a stuck-at fault into the circuit and determine the dictionary entries for the maximal-resolution dictionary $\mathcal{MD}$ and each interval-based dictionary $\mathcal{ID}$. We consider a total of 50 randomly-chosen faults, out of which we report results here for a set of eight representative faults. Similar results were obtained for the remaining faults.

Table 1 shows the CPU time needed to generate maximal-resolution and interval-based dictionaries for the benchmark circuits. These times are referred to as $T_{\mathcal{MD}}$ and $T_{\mathcal{ID}}$, respectively. Note that the time needed to generate $T_{\mathcal{MD}}$ is excessive in many cases; this clearly makes maximal-resolution pass/fail dictionaries impractical. On the other hand, interval-based dictionaries are generated in at most an hour for $L = 40$, and within a few minutes of CPU time for larger values of L.

Figure 3 shows the diagnostic resolution using an interval-based dictionary for various values of L. We show results on only two benchmarks due to lack of space. We present results for eight representative faults, injected one at a time. For each injected fault, we determine the size of the equivalence class obtained from the dictionary. The diagnostic resolution is measured by the size of this equivalence class. Clearly, a smaller equivalence class indicates higher resolution. We make the following two important observations, which corroborate the analytical results of Section 2.1.

First, compared to $\mathcal{MD}$, the diagnostic resolution for $\mathcal{ID}$ does not decrease significantly with increase in L until a threshold is reached. This threshold is lower than the threshold $TH = \lceil \log_2(M_E + 1) \rceil$ predicted by Theorem 1, and it depends on the circuit and the particular fault. We find that in every case, the dictionary can be compacted considerably with negligible impact on diagnostic resolution. Once this threshold on L is exceeded, the diagnostic resolution decreases for some faults. However, for many faults we continue to get high diagnostic resolution for larger values of L.

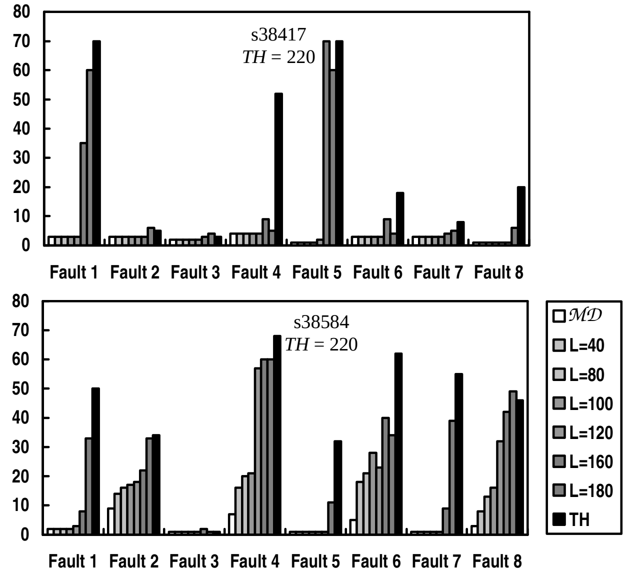


Figure 3. Diagnostic resolution (size of equivalence class) for the two largest ISCAS-89 benchmarks.

Our second observation is that the diagnostic resolution tends to decrease for most faults after the threshold predicted by Theorem 1 is exceeded. Beyond this threshold, the dictionary simply does not contain any redundancy to allow further compaction.

3 Fault dictionaries for BIST diagnosis

In Section 2, we did not consider the effect of the detection probability of a fault on diagnostic resolution. Experimental results presented in Section 2 were for faults with low detection probability. If an entry is made in an interval-based dictionary for an easy-to-detect fault, it is likely that all intervals will fail due to the fault, and the dictionary entry will consist of all 1s. As a result, a large number of easy-to-detect faults will map to the same dictionary entry of all 1s. Hence the interval-based dictionary is only one component of a complete diagnosis procedure.

3.1 A three-step diagnosis approach

First we note that in a diagnosis method based on modeled faults, most of the computation effort is expended on the hard-to-detect faults. We have seen during simulation that over 90% of the pseudorandom vectors are used to cover the relatively hard-to-detect faults, which comprise a significantly smaller proportion of the set of modeled faults. The interval-based dictionary provides the best performance for the hard-to-detect faults that are detected by a small number of pseudorandom vectors.

We create an interval-based dictionary using an appropriate interval length, usually determined by the diagnostic threshold, and two smaller full-response dictionaries. As a pre-processing step, the set of modeled faults is partitioned into three categories: easy-to-detect, hard-to-detect, and undetectable by the pseudorandom patterns used for BIST. The interval-based dictionary is created for the hard-to-detect faults such that the resolution is not adversely affected by the easy-to-detect faults. An LFSR-based compacted full-response dictionary $\mathcal{D}_1$, as

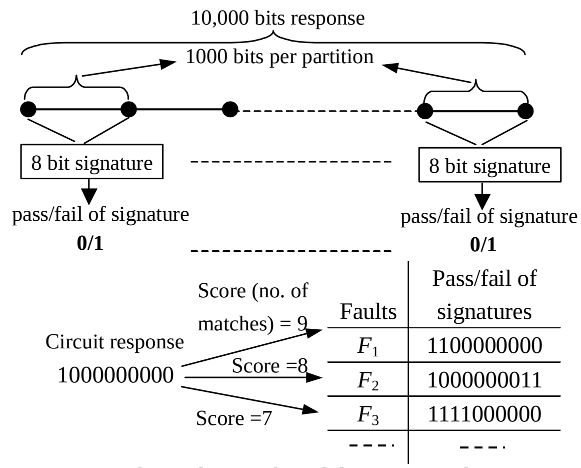


Figure 4. Enhanced LFSR-based dictionary with scoring algorithm using *vector-based-partition*.

described in Section 1, is created for vectors in the first interval and for the easy-to-detect faults. We have seen that the vectors in the first interval are usually adequate for these faults. Faults in the third category are not detected by the BIST vectors. Additional "clean-up" ATPG vectors are typically used for these faults and dictionary $\mathcal{D}_3$ is used for this step. We have considered LFSRs with primitive polynomials in our work since they appear to be the most effective in reducing dictionary size with little impact on diagnostic resolution.

3.2 Diagnosis of unmodeled faults using scoring algorithms

The single stuck-at fault model has often been shown to be inadequate in practice [10]. In order to handle unmodeled faults in dictionary-based diagnosis, scoring algorithms have been proposed in the literature [2, 6]. Scoring can be easily applied to an interval-based dictionary $\mathcal{D}_2$ since it contains valuable pass/fail status of failing vectors. However, it is difficult to directly use scoring on LFSR-based dictionaries $\mathcal{D}_1$ and $\mathcal{D}_3$ because the information of failing vectors and failing outputs in the responses is 'scrambled' in the S-bit LFSR signature.

An LFSR-based dictionary (either $\mathcal{D}_1$ or $\mathcal{D}_3$) can be made useful for scoring through two partition schemes described below. These schemes are currently being implemented, hence we describe them here without presenting experimental results.

In *vector-based-partition*, we perform a partition on the test response information over all the outputs. For each ON-bit entry in a full response dictionary, where O is the length of scan chain and N is the number of test vectors, we divide the entry into J_v partitions of equal length. We apply LFSR compaction to each ON/J_v-bit long partition. The resulting entry in such an LFSR-based dictionary is therefore a composite signature, which contains J_v signatures concatenated together. Since each signature is derived for only ON/J_v bits of test response data, we can now reduce the length of the S-bit LFSR. Let the LFSR size now be Q_v, where $Q_v < S$. The resulting dictionary is therefore J_vQ_v/S times as large as the original $\mathcal{D}_1$ or $\mathcal{D}_3$. However, this new

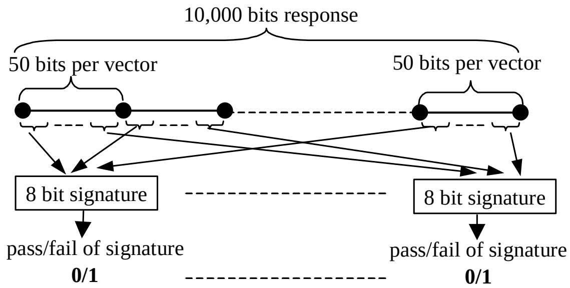


Figure 5. Enhanced LFSR-based dictionary using *output-based-partition*.

dictionary contains pass/fail status information for each partition of ON/J_v bits, which makes it suitable for scoring algorithms. For example, if $O = 50$, $N = 200$, $J_v = 10$, $Q_v = 8$ and $S = 64$, each dictionary entry is reduced from 10^4 bits to 80 bits, which still represents two orders of magnitude compaction, and since the length of each dictionary entry is increased, the diagnostic resolution also increases. The pass/fail status of the 10 signatures can be deemed as a 10-bit entry which can be used by a scoring algorithm. We illustrate this procedure in Figure 4.

An alternative *output-based-partition* can be performed on the output responses over all the test vectors. In this method, we split the O-bit outputs into J_o partitions and we also use J_o LFSRs, each Q_o bits long, where Q_o is chosen to be much smaller than S. For each test vector, the output response from the i^{th} partition is scanned into the i^{th} LFSR, and a signature is produced over all the test vectors. As a result, the length of each dictionary entry is now J_oQ_o/S times as large as the original LFSR-based dictionary. We note that this scheme can be implemented easily in a scan-BIST architecture, where each scan chain can be associated with one LFSR. For the above example, if $O = 50$, $N = 200$, $J_o = 10$, $Q_o = 8$ and $S = 64$, each dictionary entry is reduced from 10^4 bits to 80 bits, which still represents almost three orders of magnitude compaction. The pass/fail status of the 10 signatures can be viewed as a 10-bit dictionary entry that can be used by scoring algorithm. We illustrate this procedure in Figure 5; the scoring procedure is similar to that in Figure 4.

4 Experimental results

In this section, we present simulation results for the six largest full-scan ISCAS-89 benchmark circuits. In Section 2.2, we reported results for some randomly injected faults; we now examine the resolution of an interval-based dictionary for *all* single stuck-at faults using an average measure *DE*, the *diagnostic expectation*, which is defined as the average size of an equivalence class. It can be calculated as follows:

$$DE = \frac{\sum_{\forall \text{ faults } f} \text{Size of } f\text{'s equivalence class}}{\text{number of faults}}$$

A smaller value of *DE* indicates higher diagnostic resolution.

For each benchmark circuit, we considered a total of 10,000 pseudorandom vectors to construct a pass/fail dictionary for all the "hard-to-detect" faults that are not

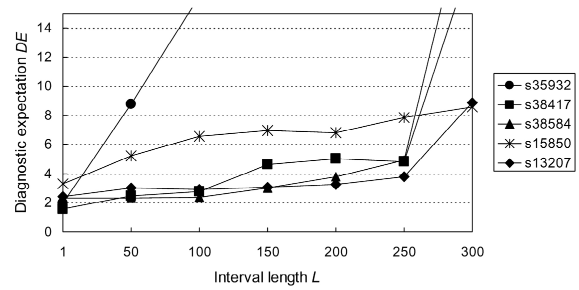


Figure 6. Variation of *DE* with interval length for pass/fail dictionaries on larger ISCAS-89 circuits.

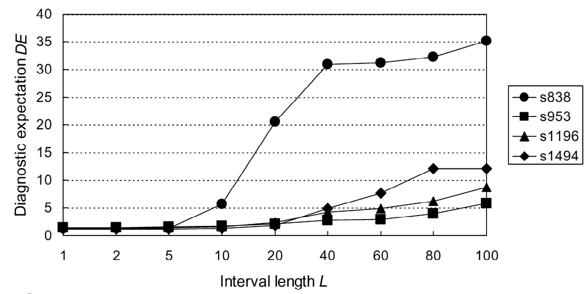


Figure 7. Variation of *DE* with interval length for full-response dictionaries on smaller ISCAS-89 circuits.

detected by the first interval. We constructed a number of interval-based dictionaries for varying interval lengths.

Our experimental results on diagnostic expectation are presented in Figure 6. As expected, *DE* increases with an increase in the interval length. However, unlike in Figure 3, there does not always exist a distinct threshold point. The reason is that our analytical results in Section 2 are based on detection probabilities of individual faults, and the experimental results are also presented for individual faults. Here, *DE* is calculated over a large number of faults with different detection probabilities and different thresholds. Nevertheless, for some circuits, e.g. s38417 and s38584, an abrupt increase in *DE* is noticed, and this can serve as a valuable guideline to choose an appropriate interval length.

We limit ourselves to 10,000 vectors because we also generate comparative data for full-response dictionaries. Even with only 10,000 patterns, the full-response dictionaries are gigantic, and it takes several days of CPU time to generate them. Moreover, we have observed in experiments that as we increase the number of BIST vectors to 100,000 and increase the interval length from 100 to 1,000 simultaneously, i.e., keep dictionary size the same, there is no loss of diagnostic resolution. Therefore, $\mathcal{D}_2$ can be efficiently generated for a much larger number of patterns. The generation of $\mathcal{D}_1$ and $\mathcal{D}_3$ takes relatively less time since these dictionaries are generated using only a small number of patterns. Experimental results on smaller ISCAS-89 benchmarks show that full-response dictionaries can also be compacted using this method with two orders of magnitude reduction in dictionary size and nearly no loss on resolution (Figure 7).

Next we examine the effectiveness of using a simulated LFSR to compact the first-stage full-response dictionary $\mathcal{D}_1$. In order to save computation time and storage for the uncompacted case, we use only 10 pseudorandom vectors to

Circuits	Size before compaction (Mbits)	*DE* before compaction	Size after compaction (Mbits)	*DE* after compaction	R_{fp} before compaction (%)	R_{fp} after compaction (%)
s9234	6.7	3.02	0.045	3.04	0.0755	0.0763
s13207	43.1	2.47	0.09	2.56	0.0253	0.0277
s15850	43.8	2.74	0.11	2.82	0.0272	0.0283
s38417	332	2.96	0.30	3.24	0.0103	0.0118
s38584	386	1.98	0.38	2.3	0.0044	0.0058
s35932	396	10.5	0.33	10.6	0.0489	0.0497

Table 2. Dictionary size and diagnostic expectation of $\mathcal{D}_1$ for larger ISCAS-89 circuits.

target the easy-to-detect faults. In practice, we expect the number of vectors for the first stage to be less than 100. Faults that are not detected by these vectors are handled in the next two stages. In Table 2, we list the dictionary sizes and the *DE* values before and after compaction based on a 16-bit LFSR. It can be seen that by storing LFSR signatures in the dictionary, we obtain three orders of magnitude reduction in size, yet there is almost no loss in diagnostic resolution.

In order to investigate the effectiveness of LFSR compaction further, we also use another measure of diagnostic resolution, which is referred to as "fault pairs" [2]. We use the parameter R_{fp} to refer to the fraction of pairs of faults (out of all pairs of faults in the dictionary) that cannot be distinguished using the dictionary. Once again, we attain almost the same diagnostic resolution after the compaction—the values of R_{fp} before and after compaction are almost the same in Table 2.

In Table 3, we present results on the total dictionary size and the *DE* values when we take into account the dictionaries for all the three stages. We also compare our results to a recent compact dictionary approach based on output compaction [8]. We note here that our method is based primarily on BIST vectors, while the work in [8] was aimed at deterministic ATPG vectors. Column 2 in Table 3 shows the dictionary size of a pass/fail dictionary based on 10,000 BIST vectors. Clearly, for large circuits, even a pass/fail dictionary is impractical. For the interval-based dictionary used in the second stage, we choose the interval length L to be 100, which is less than the threshold value in most cases. This ensures that we achieve a low value of *DE*, hence high diagnostic resolution. We add the dictionary sizes for the three stages and list them in Column 3. In Column 4, we list the *DE* values of the proposed method. It is calculated over all the three dictionaries and all the detectable faults. In Columns 5 and 6, we list dictionary sizes and *DE* values obtained from the compaction method described in [8] using ATPG vectors. (Dictionary in [8] is based on pass/fail information together with output-compacted signatures, and the dictionary size is referred to as $f(v+o)$.) It can be seen that even though our method uses a larger amount of BIST vectors than the ATPG vectors in [8], the overall dictionary size is smaller, and diagnostic resolution is in most case comparable to that in [8].

Finally, in order to demonstrate the effectiveness of using LFSR-compaction for a dictionary based on ATPG vectors, we apply the compaction procedure on the dictionaries referred to in Columns 5 and 6 of Table 3. We show the results in terms of dictionary size and *DE* values in Columns 7 and 8. The results demonstrate that the proposed LFSR-based compaction is not only effective for pseudorandom vectors, but it is also effective for dictionaries based on ATPG vectors.

Circuits	Pass/fail dictionary size before compaction (Mbits)	Overall size after compaction (Mbits)	*DE* after compaction	Size of dictionary in [8] (Mbits)	*DE* for dictionary in [8]	Dictionary size in [8] after LFSR-based compaction (Mbits)	*DE* in [8] after compaction
s9234	58.1	0.36	2.8	2.30	1.87	0.11	1.97
s13207	89.2	0.41	3.7	9.84	1.64	0.19	1.77
s15850	106.8	0.50	2.7	8.83	1.7	0.19	1.86
s38417	274.3	1.10	2.86	55.9	1.5	0.52	1.96
s38584	340.0	1.46	2.24	63.5	1.22	0.59	1.75
s35932	351.1	0.84	8.03	60.9	3.9	0.5	4.3

Table 3. Overall dictionary size and *DE* values for the larger ISCAS-89 circuits and the corresponding results obtained using the output-compacted dictionary proposed in [8].

5 Conclusions

We have presented a new technique for generating compact dictionaries for diagnosis in scan-based BIST. This approach makes diagnosis more efficient by classifying the faults in the underlying fault model and using a separate dictionary for each fault type. We have shown that a combination of three compact dictionaries can be used to obtain two to three orders of magnitude reduction in memory requirements without compromising diagnostic resolution.

References

[1] B. Chess and T. Larrabee, "Creating Small Fault Dictionaries", *IEEE Trans. CAD.*, vol. 18, pp. 346-356, March 1999.

[2] I. Pomeranz and S. M. Reddy, "On Dictionary-Based Fault Location in Digital Logic Circuits", *IEEE Trans. Comp.*, vol. 46, pp. 48-59, Jan 1997.

[3] R. C. Aitken and V. K. Agarwal. "A Diagnosis Method Using Pseudo-random Vectors without Intermediate Signatures", *Proc. ICCAD*, pp. 574-580, 1989.

[4] V. Boppana, I. Hartanto and W. K. Fuchs, "Full Fault Dictionary Storage Based on Labeled Tree Encoding", *Proc. VTS*, pp. 174-179, 1996.

[5] S. Chakravarty and V. Gopal, "Techniques to Encode and Compress Fault Dictionaries", *Proc. VTS*, pp. 195-200, 1999.

[6] P. G. Ryan and W. K. Fuchs, "Dynamic Fault Dictionaries and Two-Stage Fault Isolation", *IEEE Trans. VLSI Systems*, vol. 6, pp. 176-180, Mar 1998.

[7] P. Song et al., "S/390 G5 CMOS Microprocessor Diagnostics", *Proc. Int. Test Conf.*, pp. 1073-1082,1999.

[8] D. Lavo and T. Larrabee, "Making Cause-Effect Effective: Low-Resolution Fault Dictionaries", *Proc. Int. Test Conf.*, pp. 278-286, 2001.

[9] C. Liu, K. Chakrabarty and M. Goessel, "An Interval-Based Diagnosis Scheme for Identifying Failing Vectors in a Scan-BIST Environment", *Proc. DATE Conf.*, pp 382-386, 2002.

[10] P. C. Maxwell, R. C. Aitken, R. Kollitz and A. C. Brown, "IDDQ and AC Scan: The War Against Unmodeled Defects", *Proc. Int. Test Conf.*, pp. 250-258, 1996.