On Diagnosis of Multiple Faults Using Compacted Responses

Jing Ye†‡ Yu Hu† Xiaowei Li†
†Key Laboratory of Computer System and Architecture,
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, P.R. China, 100190
‡Graduate University of Chinese Academy of Sciences, Beijing, P.R. China, 100049
{yejing, huyu, lxw}@ict.ac.cn

Abstract — **With the exponential growth in the number of transistors, not only test data volume and test application time may increase, but also multiple faults may exist in one chip. Test compaction has been a de-facto design-for-testability technique to reduce the test cost. However, the compacted test responses make multiple-fault diagnosis rather difficult. When there is no space compactor, the most likely suspect fault is considered producing the failing responses most similar to the failing responses observed from the automatic test equipment. But when compactor exists, those suspect faults may no longer have the same high possibility of being the actual faults. To address this problem, we introduce a novel metric *explanation necessity*. By using both of the new metric and the traditional metric *explanation capability*, we evaluate the possibility of a suspect fault to be the actual fault. For ISCAS'89 and ITC'99 benchmark circuits equipped with extreme space compactors, experimental results show that 98.8% of the top-ranked suspect faults hit the actual faults, outperforming a previous work by 11.3%.**

I. INTRODUCTION

Fault diagnosis plays an important role in speeding up the yield learning process and reducing the time-to-market [1]. It analyzes the observed responses of the *Circuit-Under-Test* (CUT) and provides the most likely suspect faults for further analysis. With the exponential growth in the number of transistors, two issues of diagnosis are encountered. Firstly, single fault model may not be adequate for diagnosing defects in modern devices where multiple defects may exist and affect several metal wires [2]. Experiments in [3] showed that more than 41% of defects cannot be diagnosed using the single stuck-at fault model. Therefore, methods for multiple-defect diagnosis [2-11] and compound-defect diagnosis [12, 13] have been proposed in recent years. They use non-compacted responses for diagnosis. Secondly, both of test data volume and test application time increase, thus some compaction techniques [14-18] are proposed to reduce the test cost. However, the compaction techniques aggravate the interaction of multiple-fault effects and make the diagnosis rather difficult.

For the CUT without compactor, fault diagnosis methods can be classified into two categories: the effect-cause based diagnosis methods and the cause-effect based diagnosis methods. The effect-cause based diagnosis methods [3-9] usually have four steps: marking, simulation, evaluation, and ranking of suspect faults. It dynamically calculates the responses of a suspect fault. The cause-effect based diagnosis methods [10-11] use a static mechanism. They normally pre-compute a fault dictionary to store the responses of possible faults. Both methods compare the responses of suspect faults with the responses of the CUT to choose the most likely suspect faults. Due to the large storage space requirement of the cause-effect based diagnosis methods, the effect-cause based diagnosis methods play a dominant role in diagnosis.

Using non-compacted responses, the most likely suspect fault is considered producing the failing responses most similar to the failing responses of the CUT. However, when space compactor functions, those suspect faults may no longer have the same high possibility of being the actual faults. For the CUT equipped with space compactor, fault diagnosis approaches can be classified into three categories as follows.

(1) Bypassing diagnosis: This kind of methods reloads test patterns and tests the CUT again in the context of bypassing space compactor. Then conventional fault diagnosis using non-compacted responses are performed. During volume diagnosis, reloading test patterns to the *Automatic Test Equipment* (ATE) for a second run of test increases the test cost, but twice tests may not be desired.

(2) Indirect diagnosis: This kind of methods [15-18] tries to mathematically decompress the compacted responses, and then applies conventional fault diagnosis methods. The problem is that the rebuilt responses may be distorted, thus the succeeding diagnosis based on them may be inaccurate, too. Moreover, the indirect diagnosis techniques are structure-dependent. This limits their applications.

(3) Direct diagnosis: The direct diagnosis methods [19, 20] are similar to the effect-cause based diagnosis methods. Since the space compactor structure is known, we can use fault simulation to get the compacted responses of a single suspect fault. These simulated responses are compared with the responses of the CUT to select the most likely suspect faults. Compared with indirect diagnosis, this kind of methods is structure-independent. In [19], the single fault assumption is used, thus the authors do not specify how to estimate the match degree between the simulated responses and the responses of the CUT. In ideal situations when single fault exits, a perfect match between these two kinds of responses can be surely found. In [20], the authors present a method of estimating the match degree when multiple faults exist. However, the accuracy of top-ranked suspect faults is not very high.

The proposed multiple-fault diagnosis method using compacted responses belongs to the direct diagnosis methods. It can be used when the space compactor structure in the CUT is known. We not only consider match degree between the simulated responses of a suspect fault and the responses of the CUT, but also estimate how much necessity the suspect fault is for explaining failing observation points. By using both of the proposed metric ***explanation necessity*** and the traditional metric ***explanation capability***, we rank all the suspect faults and select the most likely suspect faults.

The rest of the paper is organized as follows. Section 2 presents the issue of multiple-fault diagnosis using compacted

This work is supported in part by National Natural Science Foundation of China (NSFC) under grant No. (61076018, 60803031, 60633060, 60921002), and in part by National Basic Research Program of China (973) under grant No. 2011CB302503.

978-3-9810801-7-9/DATE11/©2011 EDAA

responses. Section 3 elaborates the proposed diagnosis methodology. Experimental results are shown in Section 4, followed by the conclusion section.

II. Issue of Diagnosis with Compacted Responses

In this work, the compacted observable signals are called *Observation Point*s (OPs). Before illustrating the issue of multiple-fault diagnosis using compacted responses, we introduce two concepts first as follows.

A suspect fault is considered ***explaining*** a failing observation point if the single suspect fault can produce the signal same to the failing observation point. The number of the failing observation points, which can be explained by a suspect fault, is represented as "$\boldsymbol{\sigma}$".

A suspect fault is considered ***contaminating*** a passing observation point if the single suspect fault produces a signal opposite to the passing observation point. The number of the passing observation points, which are contaminated by a suspect fault, is represented as "$\boldsymbol{\iota}$".

For example, in Fig.1, assuming there are 7 observation points from OP_1 to OP_7. The responses expected to be observed are (0000000), and the actually observed responses are (0011111). Five suspect faults exist, and their simulated compacted responses are also shown in the figure. According to above definitions, F_2 can explain OP_3, OP_4, and OP_5, meanwhile it contaminates OP_1, thus its σ is 3, and its ι is 1.

The explanation capability of a suspect fault reflects how much degree the suspect fault has to explain the failing observation points while not contaminating the passing observation points. A suspect fault with a high σ and a low ι reflects its high explanation capability.

Without space compactor, when multiple faults exist, the suspect faults with the highest explanation capability often hit the actual faults [9]. The situation, that the actual faults do not show the highest explanation capability, rarely occurs. However, this situation becomes serious after a space compactor is equipped in the circuit. For example, in Fig.2-(1), the actual faults are a/0 (a-stuck-at-0) and c/0. Without space compactor, failing observation points are the Scan $Cell_{11}$ and the primary output q. In this situation, no suspect faults have higher explanation capability than the actual faults. If a simple space compactor shown in Fig.2-(2) is equipped, the failing observation points become q and the output of the space compactor ($Cell_{11}$-XOR-$Cell_{21}$, we use "OP_1" to represent it). In this situation, b/0 can explain both of OP_1 and q, thus it has the highest explanation capability.

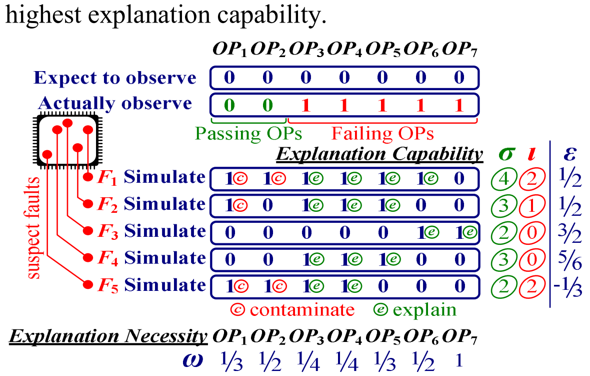

Figure 1. An example of fault diagnosis

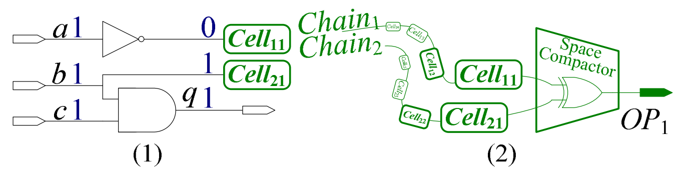

Figure 2. An example issue of multiple-fault diagnosis using compacted responses

We can see from above example that after the responses are compacted, the suspect fault with the highest explanation capability may not hit any actual faults. In other words, only using the explanation capability metric to evaluate and rank suspect faults is not enough anymore for achieving high diagnosis accuracy.

To address this issue, we introduce a novel metric explanation necessity**.** By using both of the explanation necessity and the explanation capability, each suspect fault is evaluated and ranked. The explanation necessity of a suspect fault reflects how important it is for explaining the failing observation points. For example, in Fig.1, F_4 can explain OP_3, OP_4, and OP_5. If it is not an actual fault, OP_3 and OP_4 can still be explained by F_1, F_2, or F_5, and OP_5 can still be explained by F_1 or F_2. However, if F_3 is not an actual fault, no other suspect faults can explain OP_7. In other words, the explanation necessity of F_3 is higher than F_4, because if F_3 is not an actual fault, there is one failing observation point cannot be explained.

To quantify the explanation necessity, we give each observation point a **weight** represented as "$\boldsymbol{\omega}$". It is defined as $1/N_{ec}$, where N_{ec} is the number of suspect faults which can explain/contaminate the failing/passing observation point. For example in Fig.1, the passing observation point OP_1 is contaminated by three suspect faults, thus it gets the weight 1/3, and the failing observation point OP_3 is explained by four suspect faults, thus it gets the weight 1/4. The strategy of combining the explanation necessity and the explanation capability for fault diagnosis will be given in the next section.

III. Diagnosis Methodology

The proposed diagnosis method uses four steps including marking, simulation, evaluation, and ranking of suspect faults. These steps are similar to the steps of the effect-cause based diagnosis methods. The biggest difference is that we introduce the explanation necessity metric into the suspect fault evaluation step. This improves the accuracy of diagnosis using compacted responses.

A. Suspect Fault Marking

According to the failing observation points, we can mark possible suspect faults by searching the related logic cones. For example, in Fig.2, if we observe a failing bit at OP_1, at least one actual fault should exist in the union of the logic cones of $Cell_{11}$ and $Cell_{21}$.

B. Suspect Fault Simulation

With the marked suspect faults, we simulate each single suspect fault under all the failing patterns and the passing patterns to record their compacted responses. This step is practical because the structure of the space compactor is known.

C. Suspect Fault Evaluation

After getting the simulated compacted responses of each suspect fault, we compare the simulated responses with the responses of the CUT. Two metrics (1) the explanation necessity and (2) the explanation capability are used to evaluate each suspect fault. Every suspect fault is given a score to reflect its possibility of being the actual faults, represented as "ε". The scoring equation is given as follows.

$$\varepsilon = \sum_{i=0}^{\sigma} \omega \text{ of explained } OP_i - \sum_{i=0}^{\iota} \omega \text{ of contaminated } OP_i$$

If a suspect fault can explain a failing observation point, then ε of this suspect fault is updated by adding the weight of the explained failing observation point. The larger weight the failing observation point gets, the less suspect faults can explain this failing observation point, and the higher explanation necessity the suspect fault has. Thus ε of this suspect fault pluses a larger value.

If a suspect fault contaminates a passing observation point, then ε of this suspect fault is updated by subtracting the weight of the contaminated passing observation point. The larger weight the passing observation point gets, the less suspect faults can contaminate this passing observation point, and the lower explanation necessity the suspect fault has. Thus ε of this suspect fault minuses a larger value.

For example, in Fig.1, the score of F_2 is $1/4_{(\text{explain } OP3)} + 1/4_{(\text{explain } OP4)} + 1/3_{(\text{explain } OP5)} - 1/3_{(\text{contaminate } OP1)} = 1/2$. From this equation, we can see that both of the explanation necessity and the explanation capability are considered.

D. Suspect Fault Ranking

In this step, suspect faults are ranked according to their scores. A higher score leads to a higher rank. The top-ranked suspect faults are considered having the highest possibility of being the actual faults. In the example of Fig.1, the top-ranked suspect fault is F_3 with the highest score ε=1.5.

From the example in Fig.1, we can also see that using the proposed evaluation mechanism, the suspect fault with a high explanation capability may not get a high rank. For example, either of F_1, F_2, and F_4 can explain no less than three failing observation points, but F_3 can only explain only two failing observation points. This indicates F_1, F_2, and F_4 have higher explanation capability than F_3. But considering both of the explanation necessity and the explanation capability, F_1, F_2, and F_4 get lower ranks than F_3.

In addition, the top-ranked suspect faults are expected that they cannot be distinguished by the given patterns. That means they should have the same σ and ι. However, when using the proposed equation to calculate the scores, two distinguishable faults may get the same scores. For example, in Fig. 1, F_1 and F_2 with different σ and ι get the same scores. In practice, this kind of cases rarely occurs. If this case indeed occurs, we first use σ to rank the suspect faults which have the same ε, and then use ι to rank the suspect faults which have the same σ. With this supplement, the top-ranked suspect faults are indistinguishable under given patterns. To further reduce the number of the top-ranked suspect faults, more patterns or other related information such as the design layout need to be provided for distinguishing them.

IV. Experimental Results

Experiments are conducted using ISCAS'89 and ITC'99 benchmark circuits. Test patterns are generated by a commercial ATPG tool. The information of benchmark circuits and test patterns is shown in Table I.

The works in [15, 20] propose a space compactor namely ***extreme response compactor***. It has small impact on fault coverage and can significantly reduce response data volume and test application time. In our experiments, we utilize the extreme response compactor to compact responses. An example of extreme space compactor is shown in Fig.3. The outputs of four scan chains are directly input into a XOR-GATE, and only the output of the XOR-GATE can be observed. Since four signals are compacted into one signal, the compaction ratio of this space compactor is 4.

In practice, though it is possible that the chip could contain multiple defects, it is impractical for physical failure analysis to analyze all the suspect faults [21]. The accuracy of top-ranked suspect faults is very important. Therefore, in our experiments, a diagnosis is considered successful if the top-ranked suspect faults hit at least one actual fault.

Another commonly used metric is resolution. It can be defined as the number of top-ranked suspect faults. As we mentioned before, the reported top-ranked suspect faults are indistinguishable under given patterns, thus the resolution of our diagnosis results is affected by the given patterns but not the diagnosis method itself. Because our experiments are conducted to evaluate our diagnosis method, we only give the experimental results of success rate, and do not list the results of resolution. In addition, both experiments in [19, 29] have proved that with more failing responses and test patterns, higher resolution can be achieved.

When injecting multiple faults, we choose stuck-at faults as an example of the static fault model, and choose transition faults as an example of the dynamic fault model for experiments. We randomly inject multiple stuck-at faults or multiple transition faults into the circuit. By comparing the simulated responses of suspect stuck-at/transition faults with the responses of the faulty circuit, we use both of the explanation necessity and the explanation capability to select the most likely suspect stuck-at/transition faults.

TABLE I. Information of Benchmark Circuits & Test Patterns

ISCAS'89 Circuit	Number of gates	Number of cells	Number of patterns	ITC'99 Circuit	Number of gates	Number of cells	Number of patterns
s9234	5597	228	142	b17	22645	1415	502
s13207	7951	638	275	b20	8875	490	450
s15850	9772	597	132	b22	14282	735	477
s35932	16065	1728	24				
s38417	15106	1636	104				
s38584	21353	1426	147				

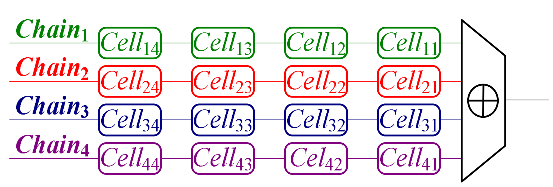

Figure 3. An example of extreme space compactor

We also compare our diagnosis method with the method in [20]. In [20], suspect faults are firstly ranked by σ, and then the suspect faults with the same σ are ranked by ι. We can see that [20] only considers the explanation capability, but not utilizes the explanation necessity.

Our experiments focus on two aspects: (1) the success rate when space compactors with different compaction ratios are equipped and (2) the success rate when different numbers of multiple faults are injected.

A. Success Rate vs Compaction Ratio

In this subsection, we study the relation between the success rate and the compaction ratio. We randomly inject two faults, three faults, and four faults into the circuit equipped with the extreme space compactor under different compaction ratios (2, 4, 8, 16, and 32), respectively. For each circuit, each number of multiple faults, each kind of faults (stuck-at faults and transition faults), and each compaction ratio, we conduct experiments 100 times for a total of 9×3×2×5×100=27000 fault injections and fault diagnosis.

Experimental results are shown in Table II and Fig.4. In Table II, we show the success rate (S), the improved success rate compared with [20] (ΔS), and the run time (T), respectively. In Fig.4, we show the average success rate for each compaction ratio. From left to right, the nine triangular marks within each compaction ratio bar represent the average success rates of the circuits s9234, s13207, s15850, s35932, s38417, s38584, b17, b20, and b22, respectively.

When the compaction ratio is 2, 99.94% of all the cases (injecting two, three, and four faults into each circuit) are successfully diagnosed using the proposed method. Note 4.96% of all the diagnosis cases (compaction ratio=2), which are failed using the method in [20], are successful using our method. As the compaction ratio increases to 32, our method can still successfully diagnose 98.64% of all the cases, and the success rate improvement increases to 12.17%. We can see from this data that, by using explanation necessity and explanation capability together, high accuracy of top-ranked suspect faults is achieved even when the extreme space compactor has a large compaction ratio.

TABLE II. EXPERIMENTAL RESULTS OF SUCCESS RATE FOR DIFFERENT COMPCTION RATIOS & COMPARISON WITH [20]

R_c: compaction ratio; M: fault model; N_f: number of injected faults; T (seconds): average run time of diagnosis in this work; S (%): success rate in this work; ΔS (%) = (success rate in this work - success rate in [20]);

Circuit			s9234			s13207			s15850			s35932			s38417			s38584			b17			b20			b22		
R_c	M	N_f	2	3	4	2	3	4	2	3	4	2	3	4	2	3	4	2	3	4	2	3	4	2	3	4	2	3	4
2	Stuck-At	S	100	100	100	100	100	100	100	100	100	100	100	100	100	100	99	100	100	100	99	99	100	100	100	100	100	100	100
		ΔS	+3	+2	+7	+1	+2	+8	+2	+6	+7	+15	+33	+57	+1	+1	-1	+3	+16	+28	+1	+4	+6	+4	+10	+5	+4	+6	+5
		T	0.46	0.63	0.79	2.19	3.61	4.06	1.97	1.79	2.86	0.93	0.69	1.69	2.20	3.47	4.17	3.01	5.12	6.83	99.5	134	179	29.9	35.0	40.0	48.4	61.1	64.1
	Transition	S	100	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
		ΔS	+1	+3	+2	+1	+4	+4	+1	0	0	+2	+1	+2	0	0	+1	0	0	+2	+2	+1	0	+2	0	+2	+2	0	0
		T	0.38	0.57	1.37	1.81	2.21	3.55	1.12	1.63	1.26	0.44	0.56	2.62	1.78	2.92	4.13	2.99	3.93	4.50	86.1	124	149	23.5	31.0	35.2	34.8	49.3	60.8
4	Stuck-At	S	100	100	100	100	100	99	100	100	99	99	100	100	100	100	99	100	100	100	99	99	100	100	100	100	100	100	100
		ΔS	+5	+2	+7	+4	+6	+12	+2	+6	+7	+11	+31	+49	0	0	0	+2	+16	+31	+3	+4	+2	+5	+12	+5	+4	+5	+3
		T	0.54	0.70	0.88	2.79	3.85	5.04	2.21	2.75	3.06	1.14	0.94	1.90	2.17	3.48	4.22	4.43	7.47	10.2	92.4	133	142	28.0	32.9	37.2	45.5	57.4	60.6
	Transition	S	100	99	100	100	100	100	100	100	100	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
		ΔS	+2	+2	+5	+1	+7	+6	+1	0	+2	0	0	+2	+1	0	0	0	0	+2	+4	+2	+1	+2	0	0	0	+1	+3
		T	0.43	0.64	1.35	2.28	2.92	5.83	1.23	1.78	1.59	0.54	0.73	2.86	1.76	3.00	4.26	4.30	5.93	7.07	81.5	115	147	22.8	29.4	33.3	33.5	47.4	57.3
8	Stuck-At	S	100	100	99	100	100	100	99	100	97	99	100	99	100	100	99	100	100	100	99	98	98	100	99	100	100	100	100
		ΔS	+12	+5	+7	+8	+15	+19	+1	+7	+6	+10	+28	+55	0	+1	0	+4	+18	+31	+4	+4	+3	+6	+10	+7	+4	+6	+8
		T	0.62	0.83	1.04	3.84	5.48	6.78	2.19	3.54	3.56	1.69	2.13	2.62	2.45	3.88	4.87	6.96	11.3	0.43	92.4	124	136	28.0	32.8	36.8	45.4	56.8	61.0
	Transition	S	100	99	99	100	100	100	100	99	99	98	100	100	100	100	100	100	100	100	100	99	99	100	100	100	100	100	100
		ΔS	+2	+4	+5	+1	+8	+6	+1	-1	+3	+2	0	+3	0	+1	0	+2	+2	+2	+4	+1	0	+2	0	+2	0	0	+9
		T	0.52	0.80	1.73	3.17	3.95	5.97	1.44	1.99	2.05	0.79	1.12	3.28	1.98	3.40	4.85	6.85	9.00	10.6	81.3	117	141	23.3	29.7	33.4	34.5	48.4	58.4
16	Stuck-At	S	99	98	96	100	100	98	99	99	97	98	99	97	100	100	99	100	100	100	99	98	96	99	98	100	100	100	100
		ΔS	+14	+12	+9	+15	+26	+33	+2	+7	+12	+16	+37	+60	+1	+2	+1	+7	+21	+39	+6	+3	+5	+6	+13	+13	+5	+9	+11
		T	0.78	1.05	1.30	4.28	7.81	9.53	2.11	3.96	4.13	2.23	4.27	4.18	3.01	4.87	6.23	10.7	16.3	14.0	99.8	123	131	29.0	34.0	37.6	47.4	59.1	65.8
	Transition	S	100	99	99	100	100	100	100	100	97	97	99	100	100	99	100	100	100	100	100	97	97	100	100	100	100	100	100
		ΔS	+2	+7	+5	+3	+7	+9	+3	+1	+5	+1	+1	+3	0	0	+1	+3	+3	+4	+6	+3	0	+6	0	+6	+2	+2	+13
		T	0.69	0.99	1.94	4.55	5.79	8.18	1.68	2.32	3.19	1.28	1.90	3.73	2.58	4.25	6.23	10.3	13.5	15.7	85.9	117	152	24.4	30.7	34.5	37.5	51.8	61.5
32	Stuck-At	S	98	98	94	98	99	98	99	98	96	98	97	96	100	100	99	100	100	100	99	96	94	99	98	99	99	100	99
		ΔS	+22	+15	+19	+26	+34	+39	+4	+8	+22	+18	+40	+60	+4	+6	+13	+18	+31	+49	+7	+8	+5	+8	+13	+20	+7	+17	+16
		T	1.14	1.48	1.80	8.87	11.3	11.4	2.76	4.77	5.22	3.38	6.96	7.17	4.64	6.98	9.21	16.0	23.4	31.4	106	125	146	30.6	36.0	39.4	51.6	63.4	62.3
	Transition	S	100	99	99	100	100	100	100	100	94	97	98	100	100	100	100	99	100	100	100	96	94	100	100	100	100	100	100
		ΔS	+4	+8	+7	+2	+11	+12	+5	+2	+5	+2	+4	+4	0	+1	+1	+5	+4	+8	+3	-1	+1	+4	+2	+19	+2	+3	+10
		T	1.07	1.36	2.56	6.49	8.27	11.0	2.38	3.15	5.11	2.30	3.53	5.26	4.00	6.41	9.13	15.4	19.7	23.0	91.4	132	166	26.3	32.5	47.1	41.5	56.2	65.6

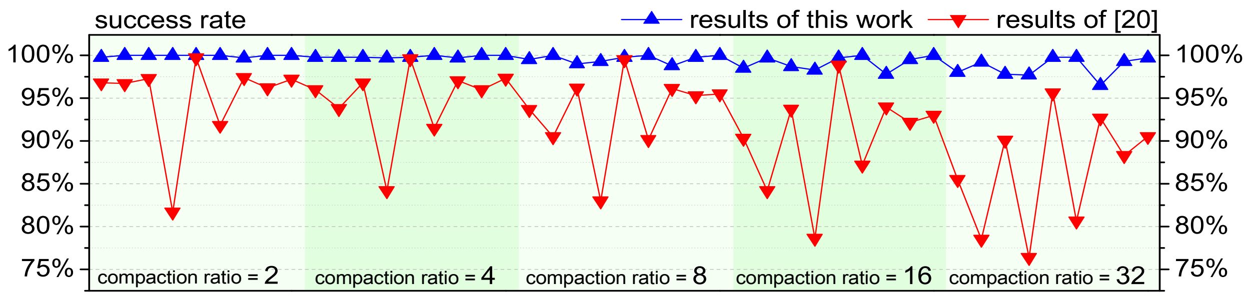

Figure 4. Comparison of success rate for different compaction ratios

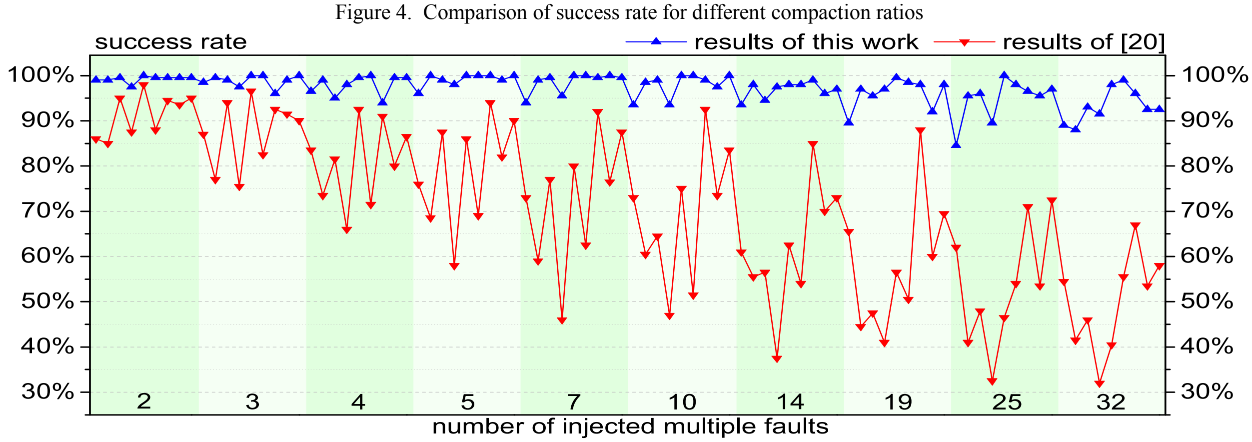

Figure 5. Comparison of success rate for different numbers of faults

TABLE III. EXPERIMENTAL RESULTS OF SUCCESS RATE FOR DIFFERENT NUMBERS OF FAULTS & COMPARISON WITH [20]

S (%): success rate in this work; ΔS (%) = (success rate in this work - success rate in [20]); T (seconds): average run time of diagnosis in this work;

Number of faults		5			7			10			14			19			25			32		
Circuit	Fault-Type	S	ΔS	T	S	ΔS	T	S	ΔS	T	S	ΔS	T	S	ΔS	T	S	ΔS	T	S	ΔS	T
s9234	Stuck-At	96	+30	1.94	91	+33	2.19	89	+31	2.49	87	+36	2.64	80	+30	2.67	77	+25	2.81	84	+40	2.89
	Transition	96	+10	1.80	97	+9	2.00	98	+10	2.23	100	+29	2.46	99	+18	2.62	92	+20	2.70	94	+29	2.75
s13207	Stuck-At	100	+42	15.2	98	+58	17.5	97	+51	15.7	96	+55	17.4	97	+61	19.5	91	+66	19.7	78	+52	20.7
	Transition	100	+21	11.7	100	+22	13.6	100	+25	16.0	100	+30	18.0	97	+44	19.0	100	+43	19.0	98	+41	19.8
s15850	Stuck-At	98	+22	6.59	99	+41	7.86	98	+62	8.67	90	+63	10.8	91	+80	12.2	92	+79	13.5	88	+76	14.2
	Transition	100	+1	4.40	100	+4	5.09	100	+7	6.38	99	+13	7.56	100	+16	8.37	100	+17	9.22	98	+18	9.79
s35932	Stuck-At	98	+76	9.59	98	+91	11.3	89	+80	14.4	95	+95	16.3	96	+94	19.6	83	+79	20.2	86	+85	17.2
	Transition	98	+4	4.94	93	+8	6.90	98	+13	8.96	100	+25	11.0	98	+18	12.6	96	+35	14.1	97	+34	15.9
s38417	Stuck-At	100	+26	10.9	100	+38	14.9	100	+47	18.8	97	+60	24.0	99	+74	29.0	100	+81	33.5	98	+92	37.4
	Transition	100	+2	10.94	100	+2	13.6	100	+3	17.6	99	+11	21.9	100	+12	27.1	100	+26	31.9	98	+23	35.2
s38584	Stuck-At	100	+56	29.2	100	+67	35.9	100	+83	40.1	98	+76	48.6	97	+84	53.9	96	+70	60.7	98	+73	64.0
	Transition	100	+6	26.2	100	+8	31.4	100	+14	39.2	98	+12	43.3	100	+12	49.4	100	+18	56.5	100	+14	60.0
b17	Stuck-At	100	+10	221	99	+13	255	100	+11	276	99	+24	301	98	+17	303	95	+48	314	96	+51	329
	Transition	100	+2	156	100	+2	204	98	+2	232	99	+4	256	98	+3	271	98	+3	276	96	+7	286
b20	Stuck-At	98	+32	40.5	100	+39	43.8	95	+38	45.4	94	+50	46.6	86	+48	47.5	91	+66	47.7	85	+58	48.0
	Transition	100	+2	38.1	100	+8	40.4	100	+10	43.3	98	+2	48.5	98	+16	45.9	100	+18	46.6	100	+20	47.1
b22	Stuck-At	100	+18	79.5	99	+20	86.0	100	+29	97.5	94	+40	100	96	+50	104	96	+36	107	90	+48	108
	Transition	100	+2	71.7	100	+4	81.7	100	+4	88.9	100	+8	96.6	100	+7	101	98	+13	103	95	+21	107

The run time of diagnosis is determined by several factors such as the number of test patterns, the number of faults, the compaction ratio, and so on. During the entire diagnosis process, the most time-consuming step is the fault simulation for each suspect fault. When both of the compaction ratio and the number of faults are 2, the average run time is about 20.9 seconds. As the compaction ratio and the number of faults increase, more signals are compacted into one signal and more failing observation points may appear, thus more suspect faults are marked. Therefore when the compaction ratio and the number of faults increase to 32 and 4, respectively, the average run time becomes 37.2 seconds. Since some high-speed fault simulation techniques have been proposed in recent years, the diagnosis speed can be further improved.

B. *Sucess Rate vs Mutliple Faults*

In this subsection, we study the relation between the success rate and the number of injected faults. We fix the compaction ratio to 32, and inject different numbers (5, 7, 10, 14, 19, 25, and 32) of multiple faults. For each circuit, each number of multiple faults, and each kind of faults, we conduct experiments 100 times for a total of 9×7×2×100=12600 fault injections and fault diagnosis.

Experimental results are shown in Table III and Fig.5. In Table III, we also show the success rate (S), the improved success rate compared with [20] (ΔS), and the run time (T), respectively. In Fig.5, we show the average success rate for different numbers of injected multiple faults. The nine triangular marks within each bar represent the success rates of the nine benchmark circuits.

When the compaction ratio is fixed to 32, the average success rate decreases from 99.2% (2 faults) to 93.3% (32 faults) using the proposed method, but it decreases from 91.4% (2 faults) to 49.8% (32 faults) using the method in [20]. An improvement of 7.8% is achieved in terms of 2 faults, and an improvement of 43.5% is achieved in terms of 32 faults. From this data, we can see that our method is more accurate than the method in [20] when multiple faults exist.

Considering all the experimental cases (27000 in the previous subsection and 12600 in this subsection), 98.8% of cases are successfully diagnosed using the proposed method. Note 11.3% of all the diagnosis cases, which are failed using the method in [20], are successful using our method. Only 0.0075% of all the diagnosis cases, which are successful using the method in [20], are failed using our method. The reason why sometimes our method fails is that the fault simulations of suspect faults are based on the single-fault assumption. An actual fault may contaminate some passing observation points when it exists in the circuit alone, but it may not contaminate any passing observation points when other actual faults exist. Since we compare the responses of the CUT with the simulated responses of the single suspect fault, false ranking may occur. Coincidentally in those 0.0075% diagnosis cases, there do exist some actual faults having the highest explanation capability, thus the method in [20] succeeds. Overall, from this data, we can see that, by introducing the explanation necessity metric, the proposed diagnosis method achieves a significant improvement of the diagnosis accuracy.

V. DISCUSSION

In this paper, we compare the simulated responses of a suspect fault with the responses of the CUT for all the patterns, including the passing patterns and the failing patterns. This is efficient when we know the actual fault models. However, in practice, we cannot know the most suitable fault models for the defects in the CUT. So the rankings may be misleading. For example, assume there are two suspect faults at net *b* and net *c*, and the actual defect is at net *b*. The actual defect is a short defect. If we use stuck-at fault model for diagnosis, we may find that the stuck-at fault at *b* contaminates some passing observation points under a passing pattern, but the stuck-at fault at *c* does not. In this situation, *b* may get a lower rank than *c*. This wrong ranking happens because we assume the defect is excited as a stuck-at fault under that passing pattern, but actually, the short defect is not excited at all under that.

To confront this problem, we may use the failing patterns only. That is because under those patterns, we can be sure that at least one actual defect is indeed excited. Though this may decrease the diagnostic resolution, other information such as the design layout can be utilized to make up the resolution loss. This is what we will study in the future.

VI. CONCLUSION

In this paper, to address the challenge of multiple-fault diagnosis using compacted responses, we propose a novel metric ***explanation necessity***. Four steps are performed: (1) suspect fault marking by searching the related logic cones of the failing observation points, (2) suspect fault simulation with the knowledge of the space compactor structure, (3) suspect fault evaluation considering both of the explanation necessity and the explanation capability, and (4) suspect fault ranking according to the score of each suspect fault. Experimental results show that the proposed diagnosis method can provide highly accurate top-ranked suspect faults, and with the introduction of explanation necessity, a significant improvement is achieved, compared with a latest work.

REFERENCES

[1] Semiconductor Industry Association, "The International Technology Roadmap for Semiconductors," 2005 edition.

[2] R.C. Aitken, "Modeling The Unmodelable: Algorithmic Fault Diagnosis," IEEE Design and Test of Computers, vol. 14, no. 3, pp. 98-103, Jul.-Sep. 1997.

[3] L.M. Huisman, "Diagnosing Arbitrary Defects in Logic Designs Using Single Location at A Time (SLAT)," IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems, vol. 23, no. 1, pp. 91-101, Jan. 2004.

[4] J.A. Waicukauski and E. Lindbloom, "Failure Diagnosis of Structured VLSI," IEEE Design and Test of Computers, vol. 6, no. 4, pp. 49-60, 1989.

[5] S. Venkataraman and S.B. Drummonds, "POIROT: A Logic Fault Diagnosis Tool and Its Applications," Proc. of International Test Conference (ITC), pp. 253-262, 2000.

[6] D.B. Lavo, I. Hartanto, and T. Larrabee, "Multiplets, Models, and The Search for Meaning Improving Per-Test Fault Diagnosis," Proc. of International Test Conference (ITC), pp. 250-259, 2002.

[7] R. Desineni, O. Poku, and R.D. Blanton, "A Logic Diagnosis Methodology for Improved Localization and Extraction of Accurate Defect Behavior," Proc. of International Test Conference (ITC), paper 12.3, 2006.

[8] X. Yu and R.D. Blanton, "An Effective and Flexible Multiple Defect Diagnosis Methodology Using Error Propagation Analysis," Proc. of International Test Conference (ITC), paper 17.1, 2008.

[9] J. Ye, Y. Hu, and X. Li, "Diagnosis of Multiple Arbitrary Faults with Mask and Reinforcement Effect," Proc. of Design, Automation, and Test in Europe (DATE), pp. 885-890, 2010.

[10] I. Pomeranz, S.M. Reddy, "A Same/Different Fault Dictionary: An Extended Pass/Fail Fault Dictionary with Improved Diagnostic Resolution," Proc. of Design, Automation, and Test in Europe (DATE), pp. 1474-1479, 2008.

[11] X. Yu and R.D. Blanton, "Multiple Defect Diagnosis Using No Assumptions on Failing Pattern Characteristics," Proc. of Design Automation Conference (DAC), pp. 361-366, 2008.

[12] Y. Huang, W.-T. Cheng, K.-H. Tsmi, G. Crowell, and C. Mcmahon, "Diagnosing DACS (Defects that Affect Scan Chain and System Logic)," Proc. of International Symposium Testing and Failure Analysis (ISTFA), pp. 291-296, 2004.

[13] F. Wang, Y. Hu, Y. Huang, H. Li, X. li, and J. Ye, "Deterministic Diagnostic Pattern Generation (DDPG) for Compound Defect," Proc. of International Test Conference (ITC), paper 14.1, 2008.

[14] J. Rajski, M. Kassab, N. Mukherjee, N. Tamarapalli, J. Tyszer, and J. Qian, "Embedded Deterministic Test for Low-Cost Manufacturing", IEEE Design and Test of Computers, Vol. 20, no. 5, pp. 58-66, Sep.-Oct. 2003.

[15] H.P.E. Vranken, S.K. Goel, A. Glowatz, J. Schloffel, and F. Hapke, "Fault Detection and Diagnosis with Parity Trees for Space Compaction of Test Responses," Proc. of Design Automation Conference (DAC), pp. 1095-1098, 2006.

[16] S. Mitra and K.S. Kim, "X-Compact: An Efficient Response Compaction Technique for Test Cost Reduction," Proc. of International Test Conferece (ITC), pp. 311-320, 2002.

[17] J. Rajski, J. Tyszer, C. Wang, and S.M. Reddy, "Convolutional Compaction of Test Responses," Proc. of International Test Conferece (ITC), pp. 745-754, 2003.

[18] P.H. Bardell, W.H. McAnney, and J. Savir, "Built-In Self-Test For VLSI: Pseudorandom Techniques," John Widley and Sons, 1987.

[19] W.-T. Cheng, K.-H. Tsai, Y. Huang, N. Tamarapalli, and J. Rajski, "Compactor Independent Direct Diagnosis," Proc. of Asian Test Symposium (ATS), pp. 204-209, 2004.

[20] S. Holst, H.-J. Wunderlich, "A Diagnosis Algorithm for Extreme Space Compaction," Proc. of Design, Automation, and Test in Europe (DATE), pp. 1355-1360, 2009.

[21] L.-T. Wang, C.-W. Wu, and X. Wen, "VLSI Test Principles and Architectures: Design for Testability," Morgan Kaufmann, San Fransico, 2006.