Built-in Self-Test for Inter-Layer Vias in Monolithic 3D ICs*

Arjun Chaudhuri*, Sanmitra Banerjee*, Heechun Park†, Bon Woong Ku†,
Krishnendu Chakrabarty*, and Sung-Kyu Lim†
*Department of Electrical and Computer Engineering, Duke University
†Department of Electrical and Computer Engineering, Georgia Institute of Technology

***Abstract*—Monolithic 3D integration provides massive vertical integration through the use of nanoscale inter-layer vias (ILVs). However, high integration density and aggressive scaling of the inter-layer dielectric make ILVs especially prone to defects. We present a low-cost built-in self-test (BIST) method to detect opens, stuck-at faults (SAFs), and bridging faults (shorts) in ILVs. Two test patterns—all-1s and all-0s—are applied to the input side of a set of ILVs (e.g., making up a bus between two tiers). On the adjacent tier (the output side of the ILVs), the test responses are compacted to a 2-bit signature through space compaction. We prove that this compaction solution does not introduce any fault aliasing. Simulations results using HSPICE and M3D benchmark designs show that the proposed BIST method requires low area overhead and test time, but provides effective fault localization and the detectability of a wide range of resistive faults.**

I. Introduction

The advent of 3D integration has extended Moore's law by enabling novel architectures through high-density vertical interconnects [1]. 3D integrated circuits (ICs) are typically realized by one of three integration approaches, namely die stacking, wafer stacking, and monolithic integration [2]. The diameter and pitch of the high-density vertical interconnects differ for these approaches, typically ranging from a few microns for die/wafer stacking to only a few nanometers in the case of monolithic integration [3].

The nanoscale inter-layer vias (ILVs) used in monolithic 3D designs enable massive vertical integration, resulting in far denser vertical connections compared to conventional TSV-based 3D ICs [4]. However, high integration density and aggressive scaling of the inter-layer dielectric make ILVs especially prone to defects [5], [6]. ILV testing is therefore needed to ensure effective defect screening and quality assurance. While ILVs can conceivably be tested together with the M3D logic/memory tiers, defect isolation and yield learning require a test solution that can exclusively target the ILVs in an M3D IC. Design-for-testability (DfT) for ILVs is therefore a promising step towards this direction.

In this paper, we present a low-cost built-in self-test (BIST) method to detect opens, stuck-at faults (SAFs), and bridging faults (shorts) in ILVs. In the proposed method, two test patterns—all-1s and all-0s—are applied to the input side of a set of ILVs (e.g., making up a bus between two tiers). On the adjacent tier (the output side of the ILVs), the test responses are compacted to a 2-bit signature through space compaction. Since the proposed scheme requires only two test patterns and generates a 2-bit signature for each ILV bus, the test time is negligible. The key contributions of this paper are as follows.

- We present a low-cost BIST architecture for detecting SAFs, hard shorts, and hard opens in ILVs. We prove that compaction does not introduce any fault aliasing.
- We investigate the probability of ILV fault masking due to faults in the BIST hardware. We propose an extended BIST architecture that guarantees zero ILV fault masking due to single BIST faults and negligible ILV fault masking probabilities for multiple BIST faults.
- We explore the detectability of resistive faults and present an enhanced BIST design that can detect a wide range of resistive shorts and opens.
- We evaluate resistive fault detectability through HSPICE simulations using the 45 nm Nangate open-cell library. We also evaluate the BIST overhead for M3D benchmarks and compare it to a baseline DfT method that uses flip-flops at the two ends of an ILV for controllability and observability.

The remainder of the paper is organized as follows. Section II presents an overview of M3D technology and related prior work on interconnect BIST and ILV testing. Section III describes the proposed BIST solution in its simplest form. Section IV describes an extended BIST architecture that minimizes the probability of ILV fault masking. The enhanced BIST architecture for resistive faults is described in Section V. Section VI presents evaluation results and Section VII concludes the paper.

II. Background

A. M3D Fabrication Process

The first step in M3D fabrication involves a standard high-temperature process to integrate the transistors and interconnects in the bottom layer. A thin inter-layer dielectric is then created over the bottom layer and low temperature molecular bonding of the silicon-on-insulator (SOI) substrate is used to obtain the top layer [7], [8]. The ILVs are finally fabricated to connect the top and bottom layers. The above steps are repeated for the fabrication of additional layers.

*This research is funded by the DARPA 3DSOC project under Award HR001118C0096.

978-1-7281-1173-5/19/$31.00 ©2019 IEEE

B. ILV Fault Models

The target fault models for ILVs are the same as those for interconnects in an IC with one active (device) layer because the dimensions of an ILV are comparable to that of vias in today's ICs [9]. During fabrication, the ILVs are treated as back-end-of-line vias that are susceptible to open and short defects [10]. Therefore, typical fault models for an ILV are shorts, opens, and SAFs [6], [11]. ILV faults can be further classified into hard and resistive categories based on the type and size of the underlying defects. Hard shorts can occur due to imperfect design and circuit synthesis, or particle contamination during fabrication [12]. A resistive short has a resistance in the intermediate range, typically ranging from a few KΩs to several hundred KΩs [13]. Resistive shorts can occur when the ILV metal diffuses through the ILD to make a partial contact with another nearby ILV [14], or due to defects at the interface between two tiers after vertical integration of the top tier's device layer [6].

A hard open occurs when an ILV fails to land on a contact pad. This gap in connection leads to a very high open resistance, typically in the order of MΩs [6], [13]. A resistive open is of relatively smaller size and has a resistance in the intermediate range, typically ranging from a few KΩs to several hundred KΩs [13]. Resistive opens can occur due to bonding defects [6], mechanical stress-induced striations on the underside of an ILV [15], air-voids inside an ILV due to imperfect electro-chemical deposition of metal [16], hairline cracks, and pinhole defects [17].

C. Related Prior Work

The testing of M3D ICs in general, and ILVs in particular, has remained largely unexplored thus far. Due to the high ILV integration density (30 million per mm^2 [9]), retrofitting of conventional interconnect BIST approaches can introduce significant overhead. Methods such as [18], [19] use dedicated scan elements (test points) for test access. However, these solutions require large test application time since the number of test patterns required for high fault coverage can become prohibitively large for high ILV density [20]. Moreover, the number of required test points is directly proportional to the ILV count. ATPG-based interconnect test methods, such as [21], are likely be less effective for ILV testing because I/O pins are available only on one layer in an M3D IC; either test data or test responses—or both in the case of ILVs that do not land on the bottom tier—must be propagated through multiple tiers and the associated ILVs. This requirement adds significantly to the propagation constraints for ATPG. Even if tests can be found by an ATPG tool, additional ILV faults on test paths, which is a likely scenario due to high ILV density, will impede testability. Commercial ATPG tools tend to target single faults for test-pattern generation. However, multiple faults are likely for dense ILV layouts; hence, test escapes might occur if tests are generated under the single-fault assumption. The proposed BIST approach alleviates these problems by using only two test patterns that exhaustively test for single or multiple ILV fault scenarios with test-output compaction and negligible fault-masking probability.

One potential test solution is to extend pre-bond and post-bond TSV testing methods to ILVs. However, pre-bond TSV testing methods such as [22] are not applicable to ILVs since bare ILVs cannot be exposed and the current wafer-probe technology cannot support the pitch requirement for ILVs (100 nm to 200 nm) [5]. While post-bond TSV testing techniques can be extended to post-assembly M3D testing, recently proposed methods such as [23] need a die-wrapper register cell on both ends of the ILV for controllability and observability. This, in turn significantly increases the associated area overhead due to the large ILV count. In [5], an inter-layer ILV BIST solution is presented using interface scan cells and a twisted ring counter. However, it mandates a dedicated test layer, which can have a significant impact in terms of the number of fabrication steps and area overhead. This technique also assumes that the number of upward-facing ("up") ILVs is equal to the number of downward-facing ("down") ILVs between the two tiers. However, in real designs, this assumption is unlikely to hold and dummy ILVs must be added to equalize the ILV counts. A novelty of the proposed ILV BIST method is that faults can be detected without an additional test layer and without an equal number of up and down ILVs between two tiers.

III. Proposed BIST Approach and Analysis of Fault Detectability

A. BIST Architecture

The BIST architecture for testing shorts, opens, and SAFs in ILVs is shown in Fig. 1. The BIST design is partitioned into two segments. The first segment corresponds to the tier (Tier 1) that includes the driving side of the ILVs. The second segment corresponds to the tier (Tier 2) that is driven by the ILVs under test. This segmentation enables testing of both bidirectional and unidirectional ILVs. On the output side of the ILVs, we insert 2-input XOR gates between adjacent ILVs. During ILV planning and placement in the design phase of an M3D IC, ILVs of the same bus tend to be placed closer to each other; every ILV can therefore be shorted to at most two adjacent ILVs. This is accounted for by our XOR placement to reduce gate count. There are $(N-1)$ XOR gates in Tier 2 for a set of N ILVs. The XOR outputs feed a space compactor. This compactor is an optimally balanced AND tree with $(N-1)$ inputs and a 1-bit output signature Y_1. We can determine if there is a fault in the given set of ILVs by observing Y_1.

In Tier 1, test data is fed to the ILVs from an input source V_{in}, which provides complementary signals to adjacent ILVs in the test mode via an inverter chain. Therefore, for N ILVs, there are $(N-1)$ inverters in Tier 1. A 2:1 multiplexer (MUX) is present at the input of every ILV to switch between functional and test modes based on the *Launch* signal. Note that the inverter chain is not on the functional path, therefore it does not impact timing closure in functional mode. Nevertheless, a long inverter chain can impact the frequency at which V_{in} can be switched. In such a scenario, the inverter chain can be broken into sub-chains and the test inputs to the

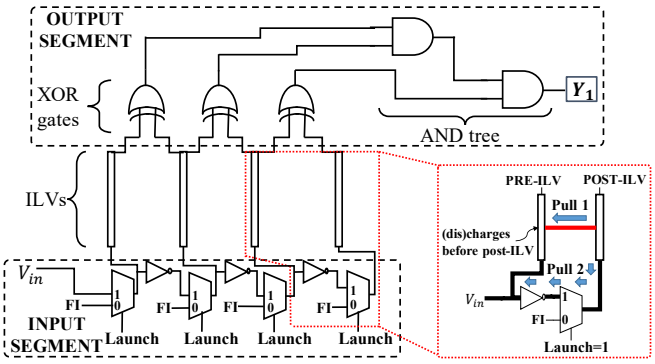

Fig. 1. Simplest form of the BIST architecture and illustration of the hard-short behavior.

ILVs adjusted accordingly to ensure that adjacent ILVs receive complementary values.

B. Detection of Hard Faults

The ILVs are tested in two clocks cycles and V_{in} is switched between these cycles. In the first (second) clock cycle, V_{in} is set to 1 (0). The resulting test patterns to the ILVs are therefore "101..." in the first cycle and "010..." in the second cycle. We next show that these test patterns and the space compactor can together detect all single and multiple hard faults in the ILVs.

Theorem 1: A set of ILVs contains no hard faults if and only if Y_1 is 1 in both clock cycles.

Proof: If Y_1 is 1 in both the clock cycles, it implies that all the XOR outputs must be 1 in these two cycles. This implies that every XOR gate gets opposite values at its two inputs. Therefore, there is no hard short present between adjacent pair(s) of ILVs. If an ILV is stuck-at-d (s/d, $d \in \{0,1\}$), the XOR output will flip to 0 in the input cycle when the adjacent ILV is given d as input, which will force Y_1 to 0. If the ILV contains a hard open, the ILV output will be held at d for both cycles; this causes the XOR output to be 0 in the input cycle when the adjacent ILV is given d as input, which forces Y_1 to 0. Therefore, if Y_1 is 1 in both cycles, no hard faults are present in the ILVs. We next prove the "only if" part of the theorem. If Y_1 is 0 in either one or both cycles, we can infer that at least one of the XOR outputs is 0 in each cycle. If a hard short is present, it will force Y_1 to 0 in both cycles. If there is a stuck-at fault or hard open, Y_1 will be 0 in either of the two cycles, as explained earlier. If multiple concurrent faults are present, they will also be detected since their effects propagate through the XOR present between every adjacent ILV pair to Y_1 in a mutually exclusive manner; a hard short will force Y_1 to 0 in both cycles and a hard open or stuck-at fault will force Y_1 to 0 in one of the cycles, irrespective of the presence of other faults. □

The manner in which the ILVs are driven in the test mode leads to a deterministic hard-short behavior; this is illustrated in the inset of Fig. 1. If two ILVs are shorted, the ILV that appears first (pre-ILV) in the path of the incoming test signal from V_{in} will drive the other ILV (post-ILV). This is because the post-ILV will experience opposite pulls via two paths: one through the highly conductive short to the pre-ILV (pull 1) and the other through the MUX and inverter to the test signal (pull 2). HSPICE simulations show that the latter is of higher resistance and hence offers weaker pull.

IV. Avoidance of ILV Fault Masking

A. Proposed Architecture

The BIST design of Section III.A is also susceptible to SAFs. Some of these faults may mask fault(s) in the ILVs under test. If this happens, Y_1 will be 1 in both clock cycles even if the ILVs contain fault(s). To minimize the probability of ILV fault masking due to faults in the BIST design (referred to as BIST-A), we add a parallel propagation path from the ILV outputs to a second 1-bit signature Y_2 (we refer to this part of the BIST logic as BIST-B). The physical structure of this parallel path to Y_2 (PY_2) is identical to that of the path from the XOR inputs to Y_1 (PY_1). The XOR and AND gates in PY_1 are replaced with their dual counterparts (XNOR and OR, respectively) in PY_2. This "dual" BIST architecture is shown in Fig. 2.

B. Masking of ILV Faults due to a SAF in BIST

In the dual-BIST architecture, let the 2-bit signature for the ILVs be $\{Y_1, Y_2\}$, where Y_1 and Y_2 represent the outputs from the XOR-AND and XNOR-OR compactors, respectively. The XOR-AND and XNOR-OR compactors are denoted as BIST-A and BIST-B, respectively. The proposed BIST method cannot detect the fault scenarios in which all the ILVs are stuck at 0's and 1's alternately (010101... or 101010...). This is because the ILVs are stuck at values identical to the test patterns and $Y_1 = 1$ in both the test clock cycles, resulting in fault masking. We next prove that if this low-probability scenario does not occur, ILV faults cannot be masked by a single BIST fault. Hence, in the more likely event that not all the ILVs are stuck at alternate values, we can present the following theorem:

Theorem 2: The ILVs under test and the dual-BIST engine are fault-free if and only if $Y_1 = 1$ and $Y_2 = 0$ in both cycles.

Proof: We present the proof by explicitly enumerating all possible scenarios. Consider the case where $Y_1 = 0$ in either or both cycles. This observation implies at least one of the following three scenarios: (i) two or more ILVs are shorted; (ii) there is a stuck-at-0 fault in either BIST-A or BIST-B; (iii) one or more ILVs are either stuck-at-0/stuck-at-1 or have an

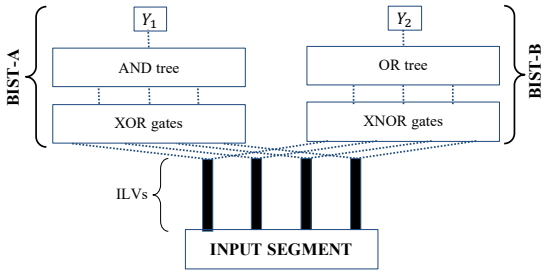

Fig. 2. Illustration of the "dual" BIST architecture.

open fault. It is clear that all the above cases are indicative of faults in the ILVs and/or the dual-BIST architecture. Next we consider the case where $Y_1 = 1$ in both the cycles. Let us take any two neighbouring ILVs, say V_1 and V_2. Suppose V_1 and V_2 are the two inputs to the XOR gate and its counterpart XNOR gate denoted by X_1 and XN_1, respectively. The possible events that can give rise to this scenario are as follows:

- *Event I:* V_1 and V_2 are shorted (or both are s/d; $d \in \{0, 1\}$); there is a node s/1 in the path from X_1's output to Y_1. This implies that BIST-A is faulty and it masks the short between V_1 and V_2.
- *Event II:* V_1 (V_2) is s/d. This implies that BIST-A is faulty since the other input of X_1, namely V_2 (V_1) is s/$\bar{d}$.
- *Event III:* V_1 is s/d and V_2 is s/$\bar{d}$. Therefore, the two inputs of X_1 and XN_1 are s/d and s/$\bar{d}$, respectively.
- *Event IV:* BIST-A is fault free; V_1 and V_2 are fault free.

If *Event I* occurs, error detection is provided by Y_2. Since there is a fault in BIST-A, BIST-B is fault-free based on our single fault assumption for BIST. Hence, the short will produce a 1 at XN_1's output and $Y_2 = 1$ in both cycles. Thus $Y_1.Y_2 = 1$ in both cycles. If *Event II* occurs, error detection is provided by Y_2 again. Similar to the previous case, BIST-B is fault-free due to the single fault assumption. The open/stuck-at will produce a 1 at XN_1's output and $Y_2 = 1$ in the two cycles since V_2 (V_1) will toggle with the applied input pattern despite V_1 (V_2) being unchanged. Thus $Y_1.Y_2 = 1$ in the two cycles. For the analysis of *Event III*, consider V_0 and V_3 to be the ILVs adjacent to V_1 and V_2 respectively. If the ILVs are fault-free, they can propagate the effect of open/stuck-at faults. In *Event IV*, BIST-A is fault-free, thus we can consider a single fault in BIST-B. *Event IV* may arise in two possible scenarios:

- Any node in BIST-B is s/1, hence $Y_2 = 1$ in both cycles. Therefore, $Y_1.Y_2 = 1$ in both cycles.
- BIST-B contains either no fault or a s/0 fault, and thus $Y_2 = 0$ in both cycles. No fault masking occurs since V_1 and V_2 are fault-free.

This proves that the ILVs and BIST are fault-free if and only if $Y_1 = 1$ and $Y_2 = 0$ in both cycles. Thus, we can detect ILV faults in the presence of a single BIST fault. □

C. Masking of ILV Faults due to Multiple SAFs in BIST

For analyzing ILV fault masking due to multiple BIST faults, we only need to address the cases when we observe $Y_1 = 1$ and $Y_2 = 0$ in both cycles. We first define a faulty ILV pair as a pair of adjacent ILVs that are either shorted, or either one or both of them are stuck-at or open, or there is a combination of these faults. Let F be the event that a given ILV pair contains fault(s) and M be the event that these faults are masked by faults in the BIST logic. Let p_m be the probability $P[F]$ that any given ILV pair is faulty. We next calculate $P[F \cap M]$, i.e., the probability that a given ILV pair is faulty but masking occurs due to BIST faults. Let p_b be the probability that a node in the BIST logic is faulty. Let p_s be the probability that there is a short between adjacent ILVs. Let p_o be the probability that an ILV has a stuck-at fault, with s/0 (open) and s/1 being equiprobable. Therefore, p_m can now be defined as: $p_m = 1 - (1 - p_s)(1 - p_o)^2$. We know from basic probability theory that $P[F \cap M] = P[F] \cdot P[M|F]$; here, $P[M|F]$ is the probability that, given a faulty ILV pair, masking is caused by BIST faults.

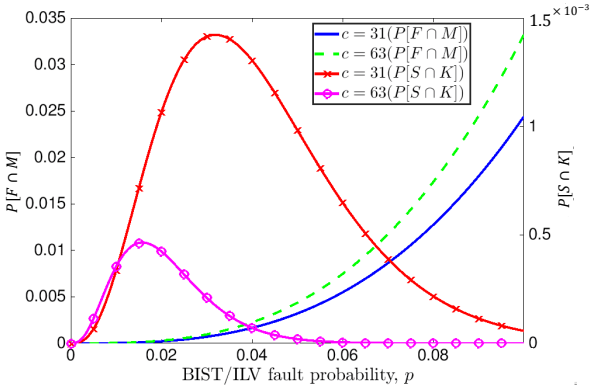

Fig. 3. Multiple fault-masking probability.

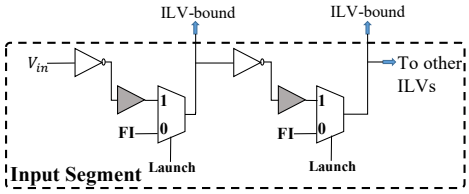

Fig. 4. Enhanced dual-BIST for resistive faults.

It is shown in [24] that $P[F \cap M] \approx (p_m p_b^2 N_e^2)/4$, where $N_e = \lceil \log_2 c \rceil + 1$ and c is the number of adjacent ILV pairs in a set. Therefore, $P[M|F] \approx (p_b^2 N_e^2)/4$. The probability that any one out of the c ILV pairs contains fault(s) (event S) and the fault(s) is (are) masked by faults present in the BIST engine (event K) is given by: $P[S \cap K] = c \cdot P[F \cap M](1 - p_m)^{c-1}$ [24]. Fig. 3 illustrates the variation of $P[F \cap M]$ and $P[S \cap K]$ with p, where $p = p_s = p_o = p_b$ and $c \in \{31, 63\}$. We observe that the probability of ILV fault(s) being masked by multiple BIST faults is very low.

V. Enhanced Dual-BIST for Resistive Faults

The range of detectable defect size (i.e., magnitude of open or short resistance) can extended by the addition of a delay element—an N_s-stage inverter (where N_s is the number of cascaded single-stage CMOS inverters)—to the input of every ILV in the test mode. Fig. 4 describes this design for $N_s = 2$; the buffers are shaded. We next show how the addition of the delay stage facilitates the detection of resistive faults caused by defects of intermediate size.

A. Resistive Shorts

Fig. 5(a) illustrates the lumped circuit model for the testing of a resistive short. The buffer's equivalent resistance is denoted by R_{BUF}, the ILV self-resistance is denoted by R_{ILV}, and R_S denotes the resistance of the short. Complimentary test inputs are applied at the inputs of the buffers—$V_{DD}(1)$ and Ground (0). The differential input voltage of the 2-input XOR is given by $\Delta V_{out} = V_{out,1} - V_{out,2} = V_{DD} R_S / (2R_{ILV} + 2R_{BUF} + R_S)$. A short is detected if and only if it forces the XOR output to 0 instead of 1 (fault-free

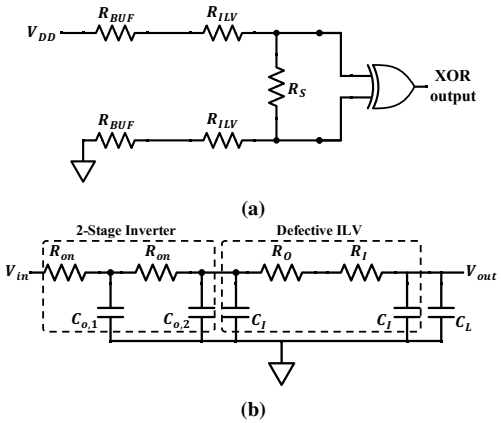

Fig. 5. Circuit models for resistive (a) short, and (b) open.

case). The XOR output is 0 when its differential input voltage is less than a preset threshold. The addition of the buffer stage decreases the magnitude of this differential input voltage since ΔV_{out} for $R_{BUF} > 0$ is less than ΔV_{out} for $R_{BUF} = 0$. Therefore, the maximum detectable short resistance can be increased by increasing R_{BUF}.

B. Resistive Opens

For a given functional clock period T_{clk}, let the range of opens detected be $[R_{o,min}, \infty)$. The enhanced dual-BIST can detect open resistances lower than $R_{o,min}$ by adding a buffer delay to the small delay of the open defect. Fig. 5(b) illustrates the lumped circuit model of the buffer aiding the test of a resistive open. Here, R_{on} is the ON-resistance of a transistor, $C_{o,x}$ $(x = 1, 2)$ are the output capacitances of the two inverter stages in the buffer, C_I is the parasitic ILV-to-substrate capacitance, C_L is the ILV load (fan-out) capacitance, and R_O denotes the open resistance. Using the Elmore-Delay model for an RC ladder, the total delay, D, of a signal from the buffer input V_{in} to the ILV output V_{out} can be expressed as:

$$D = (\ln 2) \times (R_{on}(C_{o,1} + 2C_{o,2} + 4C_I + 2C_L)) + (\ln 2) \times (R_O + R_I)(C_I + C_L)$$

A sufficient condition for the detection of an open is given by: $D + \Delta_{XOR} > T_{clk}$, where Δ_{XOR} is the XOR gate delay. Therefore, without decreasing T_{clk}, the minimum detectable open resistance can be decreased beyond $R_{o,min}$ by increasing R_{on}, i.e., by decreasing transistor width or increasing N_s.

VI. Experimental Results

A. Experimental Set-Up

We evaluated the detection of resistive faults through HSPICE simulations using the Nangate 45 nm open-cell library. Our test-bench comprised of a set of 11 ILVs and three test clock frequencies—500 MHz, 1 GHz, and 2 GHz. We considered a power supply voltage of 1 V.

We also evaluated the impact of BIST on the power-performance-area (PPA) metrics of four M3D benchmarks (Table I), where f_m denotes the maximum operating frequency in MHz. Synthesis was performed using the TSMC 28 nm library.

TABLE I. M3D benchmarks used for PPA comparison.

Name	Footprint (μm×μm)	Cell count	ILV count
Rocketcore [25] ($f_m = 350$)	700×700	301219	1200
AVC-Nova [26] ($f_m = 366$)	420×420	134547	317
AES-128 (I) [26] ($f_m = 2273$)	350×350	102278	425
AES-128 (II) [26] ($f_m = 1250$)	350×350	90233	426

B. Detection of Hard and Resistive Faults

Four scenarios (I-IV) were chosen to validate the effectiveness of our enhanced dual-BIST solution. Scenario I is the fault-free scenario; Scenario II contains two hard shorts (3 Ω) between adjacent ILVs of two distinct pairs; Scenario III contains two hard opens (1 MΩ) in two distinct ILVs; Scenario IV contains a resistive short (5 KΩ) and a resistive open (20 KΩ) in two distinct ILVs. The signatures Y_1 and Y_2 are captured at the rising edge of the second clock cycle. The test clock frequency is 2 GHz. Figure 6 shows that an error is detected for each multiple-fault scenario. Table II(a) shows the minimum detectable resistive open ($R_{o,min}$) for different buffer sizes and clock frequencies (0.5 GHz, 1 GHz, and 2 GHz). The results confirm that for all three clock frequencies, the detection range of resistive opens increases with the decrease in transistor widths of the buffer—we are able to detect a resistive open as small as 25 KΩ with a clock frequency of 2 GHz. The impact of buffer size on the maximum detectable resistive short (R_S) is presented in Table II(b). The range of detectable shorts is independent of the clock frequency, and it expands with a decrease in buffer size. Hence, we are able to detect a resistive short as large as 12 KΩ. The transistor lengths are fixed at 50 nm for both open- and short- detection experiments. The maximum lengths of appropriately-sized inverter chains in the input segment that provided detection at 0.5 GHz, 1 GHz, and 2 GHz were 21, 16, and 12, respectively.

C. Design and Synthesis Flow of BIST-inserted M3D IC

The M3D design flow *Shrunk-2D* [27] is used to synthesize the benchmarks listed in Table I. BIST is inserted tier-wise to generate a 3D BISTed design. Fig. 7 shows the die shots of the BISTed Rocketcore.

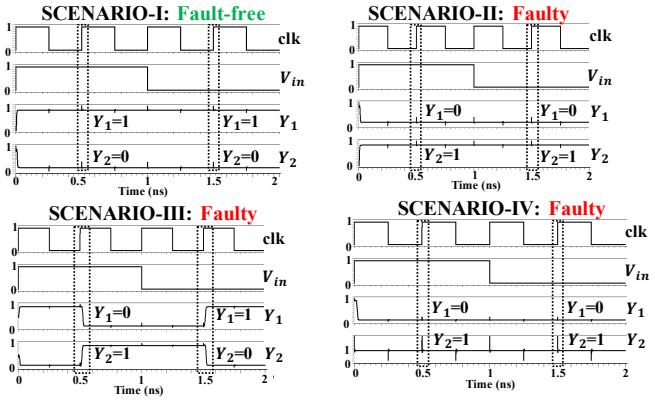

Fig. 6. Detection results for multiple-fault scenarios.

TABLE II. Impact of buffer (and buffer sizing) on (a) resistive open detection, and (b) resistive short detection.

(a)

Frequency (GHz)	Width (μm/μm) (PMOS/NMOS)	$R_{o,min}$ (KΩ)
0.5	Without buffer	118
	2/1	112
	1/0.5	110
1	Without buffer	56
	2/1	54
	1/0.5	52
2	Without buffer	29
	2/1	26
	1/0.5	25

(b)

Width (μm/μm) (PMOS/NMOS)	R_S (KΩ)
Without buffer	0.13
1/0.5	0.9
0.5/0.25	1.8
0.2/0.1	5
0.1/0.05	12

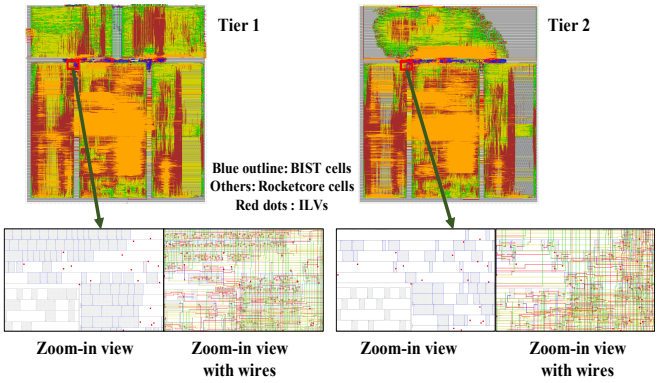

Fig. 7. BISTed Rocketcore layout.

D. Overhead for BIST

The dual-BIST architecture was inserted (with an inverter chain of length nine) in the four M3D benchmark circuits, as described in Section VI.C. Table III presents the power-consumption and area overheads of the BISTed designs (BI) with respect to the non-BISTed designs (N-BI). The impact of BIST on PPA is minimal. To further reduce the switching power consumption, transmission gate switches are connected between every ILV output and the output segment of the BIST to turn off BIST engine in functional mode; the additional area overheads are 0.19% for Rocketcore, 0.2% for AES (I), 0.23% for AES (II), and 0.12% for Nova.

The dual-BIST solution requires less area compared to a baseline DfT scheme where scan flops are added at both ends of an ILV for controllability and observability; the comparison is shown in Table IV.

VII. Conclusion

We have presented a low-cost BIST method that can detect opens, SAFs, and shorts in ILVs using only two test patterns. The proposed solution achieves bus-level fault localization via 2-bit signatures generated for every ILV bus under test. HSPICE simulations and evaluation results for four M3D benchmarks validate the effectiveness of the BIST solution in detecting a wide range of resistive faults with negligible impact on the PPA metrics.

TABLE III. Impact of BIST on PPA.

Circuit	Metric	N-BI	BI	Overhead (%)
Rocketcore	Cell area (μm^2)	382037.6	389785.6	2.03
	Wire-length (m)	7.03	7.17	1.97
	Power (mW)	227.3607	232.5544	2.28
AVC-Nova	Cell area (μm^2)	196841.5	199038.7	1.12
	Wire-length (m)	3.02	3.04	0.61
	Power (mW)	87.46587	90.79729	3.81
AES-128 (I)	Cell area (μm^2)	132049.9	134771.2	2.06
	Wire-length (m)	1.68	1.7	1.24
	Power (mW)	178.5411	194.3433	8.85
AES-128 (II)	Cell area (μm^2)	103632.4	106324	2.6
	Wire-length (m)	1.61	1.64	1.35
	Power (mW)	110.96	119.42	7.63

TABLE IV. Area comparison of baseline DfT and dual-BIST.

Benchmark	Baseline DfT area (μm^2)	Dual-BIST area (μm^2)
Rocketcore	12288	7748
AVC-Nova	3246.1	2197.2
AES-128 (I)	4352	2721.3
AES-128 (II)	4362.2	2691.6

References

[1] S. Wong et al., "Monolithic 3D integrated circuits," in *VLSI-TSA*, 2007.
[2] J.-Q. Lu et al., "3D integration: Why, what, who, when?" in *Future Fab Int.*, 2007.
[3] S. S. Wong et al., "The prospect of 3D-IC," in *CICC*, 2009.
[4] M. M. Shulaker et al., "Monolithic 3D integration: a path from concept to reality," in *Proc. DATE*, 2015.
[5] A. Koneru et al., "A design-for-test solution based on dedicated test layers and test scheduling for monolithic 3D integrated circuits," in *TCAD*, 2018.
[6] ——, "Impact of electrostatic coupling and wafer-bonding defects on delay testing of monolithic 3D integrated circuits," in *JETC*, 2017.
[7] T. Yonehara, "Epitaxial layer transfer technology and application," in *SOI-3D*, 2015.
[8] P. Batude et al., "3DVLSI with CoolCube process: An alternative path to scaling," in *Symp. on VLSI Technology*, 2015.
[9] C. Liu et al., "A design tradeoff study with monolithic 3D integration," in *ISQED*, 2012.
[10] R. F. Hafer et al., "Full-wafer voltage contrast inspection for detection of BEOL defects," in *IEEE Trans. Semiconductor*, 2015.
[11] A. Koneru et al., "Impact of wafer-bonding defects on monolithic 3D integrated circuits," in *EPEPS*, 2016.
[12] N. Campregher et al., "Analysis of yield loss due to random photolithographic defects in the interconnect structure of FPGAs," in *Proc. ACM/SIGDA FPGA*, 2005.
[13] O. D. Patterson et al., "Detection of resistive shorts and opens using voltage contrast inspection," in *ASMC*, 2006.
[14] R. M. Geffken et al., "Method of forming a self-aligned copper diffusion barrier in vias," 1999, US Patent 5,985,762.
[15] W. Liebsch et al., "Process characterisation of Dupont MXA140 dry film for high resolution microbump application," in *EMPC*, 2009.
[16] T. Frank et al., "Reliability approach of high density through silicon via (TSV)," in *EPTC*, 2010.
[17] M. Tsai et al., "Through silicon via (TSV) defect/pinhole self test circuit for 3D-IC," in *IEEE Int. Conf. on 3D Sys. Integration*, 2009.
[18] R. Pendurkar et al., "Switching activity generation with automated bist synthesis for performance testing of interconnects," in *TCAD*, 2001.
[19] J. Rajski et al., "Fault diagnosis of TSV-based interconnects in 3-D stacked designs," in *ITC*, 2013.
[20] A. Jutman, "Shift register based TPG for at-speed interconnect BIST," in *24th Int. Conf. on Microelectronics*, 2004.
[21] D. Erb et al., "Multi-cycle circuit parameter independent ATPG for interconnect open defects," in *VTS*, 2015.
[22] B. Noia et al., "Pre-bond probing of TSVs in 3D stacked ICs," in *ITC*, 2011.
[23] E. J. Marinissen et al., "IEEE Std P1838: DfT standard-under-development for 2.5 D-, 3D-, and 5.5 D-SICs," in *ETS*, 2016.
[24] Supplementary Document. https://bit.ly/2Llxk1G.
[25] RISC-V Cores and SoC Overview. https://riscv.org/risc-v-cores/.
[26] OpenCore benchmark suite. http://www.opencores.org/.
[27] S. Panth et al., "Shrunk-2D: A Physical Design Methodology to Build Commercial-Quality Monolithic 3D ICs," in *TCAD*, 2017.