

# Enhancing Test Compression With Dependency Analysis for Multiple Expansion Ratios

Taehee Lee, *Student Member, IEEE*, Nur A. Touba, *Fellow, IEEE*, and Joon-Sung Yang, *Member, IEEE*

**Abstract**—Scan test data compression is widely used in industry to reduce test data volume (TDV) and test application time (TAT). This paper shows how multiple scan chain expansion ratios can help to obtain high test data compression in system-on-chips. Scan chains are partitioned with a higher expansion ratio than normal in scan compression mode and then are gradually concatenated based on a cost function to detect any faults that could not be detected at the higher expansion ratios. It improves the overall test compression ratio since it potentially allows faults to be detected at the highest expansion ratio. This paper introduces a new cost function to choose scan chain concatenation candidates for concatenation for multiple expansion ratios. To avoid TDV and TAT increase by scan concatenation, the proposed method takes a logic structure and scan chain length into consideration. Experiment results show the proposed method reduces TAT and TDV by 53%–64% compared with a traditional scan compression method.

**Index Terms**—Compression ratio, test application time (TAT), test data compression, test data volume (TDV).

## I. INTRODUCTION

AS THE complexity of system-on-chip (SoC) designs keeps increasing, test data volume (TDV) and test application time (TAT) are growing significantly in scan test. Test cost can be reduced by decreasing TDV and TAT. In scan testing, a high TDV is required for scan vectors in automatic test equipment memory and a long scan data shift results in a long TAT. TDV and TAT are proportional to the scan chain length as well as the number of test patterns. In order to reduce TDV and TAT overhead, the number of test patterns and the longest scan chain length need to be reduced. The goal of this paper is to reduce the number of test patterns and the scan chain length with a high test compression ratio.

Manuscript received May 15, 2016; revised August 9, 2016; accepted October 11, 2016. Date of publication March 10, 2017; date of current version August 18, 2017. This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea by the Ministry of Education under Grant NRF-2015R1D1A1A01058856, in part by the Korea Institute for Advancement of Technology (KIAT) by the Korean Government (Motie:Ministry of Trade, Industry & Energy, HRD Program for Software-SoC Convergence) under Grant N0001883, in part by Samsung Research Fund, and in part by the National Science Foundation under Grant CCF-1617665. This paper was recommended by Associate Editor N. Nicolici.

T. Lee is with the Department of Semiconductor and Display Engineering, Sungkyunkwan University, Suwon 16419, South Korea.

N. A. Touba is with the University of Texas at Austin, Austin, TX 78712 USA.

J.-S. Yang is with the Department of Semiconductor Systems Engineering, Sungkyunkwan University, Suwon 16419, South Korea (e-mail: js.yang@skku.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCAD.2017.2681063

Using full scan design can achieve a high fault coverage since all scan cells are controllable and any desired values can be assigned by serially shifting required values through a scan chain. However, full scan design increases TDV and TAT. Scan compression techniques [1]–[20] have been proposed to reduce test cost based on using a test stimulus decompressor and response compressor.

Various scan compression related subjects have been researched. To reduce TAT, pipelined flip-flops are added to solve a delay problem on a long wire running between the scan ports and compressor/decompressor [16]. Efficient tester channel allocation techniques are proposed to further compress test data with a small number of scan test channels [6], [17], [18]. Wohl *et al.* [19] compressed the test data both at core level and top level for large size SoCs. Chandra *et al.* [20] introduced an efficient scan test compression technique called Zscan which reduces scan test channels. Zscan needs additional registers to convert a few inputs to multiple parallel inputs.

Fault coverage can be limited by the following reasons in test compression. First, most circuit states are not reachable when scan input pins are shared in a scan compression structure since the circuits driven by the scan chains are dependent. Fig. 1(a) gives an example of controllability loss by a shared scan input. Because the scan chain 1 shares scan input pins with scan chain 2, the stuck at 0 (S-A-0) fault at output pin of AND gate cannot be detected. Second, there is an observability loss by unknown ( $X$ ) values generated from non-scannable elements, uninitialized memories, timing-sensitive paths, IP cores without test modes and etc. A large number of  $X$  values in output response compression logic can degrade fault coverage since the unknown values may block fault propagation to external output pins [11], [12]. Fig. 1(b) illustrates an example that the unknown value in the overlapping fan-in logic cone between scan cells  $SC_4$  and  $SC_6$  reaches to  $XOR_1$  and  $XOR_2$  in the compression logic. It may result in a fault coverage loss. Third, a fault may not be detected if multiple fault values mask each other. In Fig. 1(c), if a stuck at 1 (S-A-1) fault exists in the overlapping fan-in logic cone between  $SC_2$  and  $SC_4$  and the fault on two different scan chains simultaneously propagates to  $XOR_1$ , the output value of  $XOR_1$  in the faulty design is masked. This limits the fault coverage.

In order to cover the fault coverage loss in scan compression mode, top-off test vectors are used in serial mode. Because the cost of top-off serial vectors is much higher than that of vectors in compression mode, it is desirable to keep the number of top-off serial vectors low.

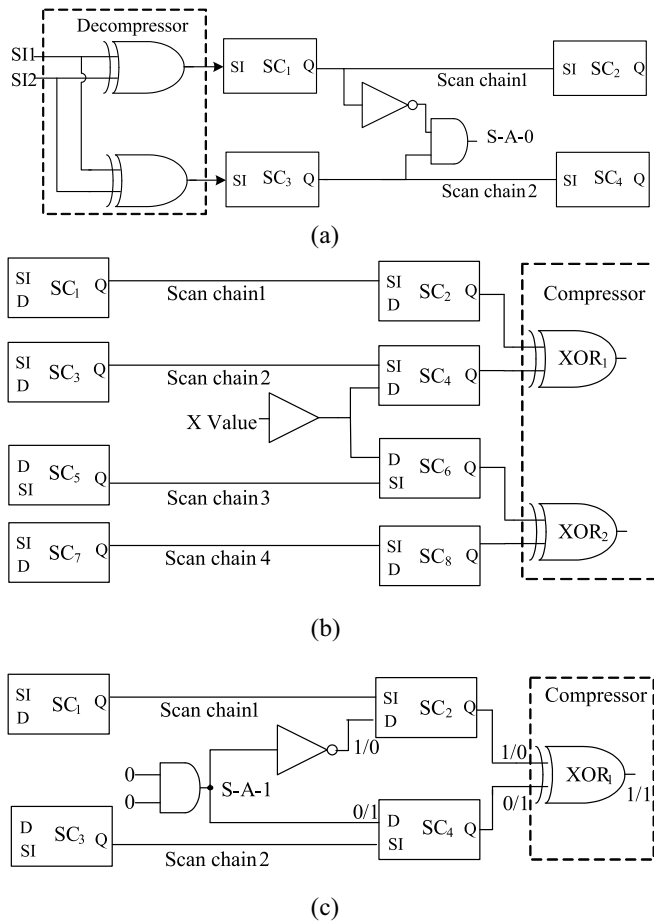


Fig. 1. Three types of fault coverage loss in scan compression structure. (a) Controllability loss induced by shared scan input pins. (b) Observability loss induced by  $X$  propagation. (c) Observability loss induced by aliasing.

In scan testing, the expansion ratio is the ratio between scan chains and external scan input pins. The higher the selected expansion ratio, the more TDV and TAT are reduced since the length of internal scan chains decreases and fewer clock cycles are needed to load/unload scan data. However, as the expansion ratio increases, the controllability and observability generally become worse. In this paper, we propose a scan architecture supporting multiple expansion ratios to increase test compression. The scan chains supporting the multiple expansion ratios start with a higher expansion ratio and then the expansion ratio is gradually decreased to detect any faults which are not detected at the higher expansion ratio. Scan chains are concatenated to reduce the expansion ratio.

The scan chains with maximal free-variable dependency and structural dependency are selected for concatenation so that after concatenation, the overall dependency among scan chains is minimized. A free variable represents each bit stored on the tester and applied through an external scan input pins which can be loaded as either 0 or 1. If the same free variables are uploaded into two scan cells, free variable dependency between the scan cells exists. The structural dependency analysis investigates the number of overlapping cells between each pair of scan chains. There are two types of structure dependencies. One is the structure dependency of fan-in cone

and the other is the structure dependency of fan-out cone. Analyzing free variable and structural dependencies improves the probability of detecting additional faults as scan chains are concatenated.

The proposed approach considers the length of scan chains in scan concatenation. Because scan data has to be shifted in for all the scan cells in the scan chain, TDV and TAT are determined by the longest scan chain length. The length of scan chain should be taken into account before concatenating scan chains since it can minimize a potential increase for TDV and TAT. This paper proposes an enhanced dependency cost function considering free-variable dependency, structural dependency of fan-in and fan-out logic cones, and scan chain length for scan chain concatenation (it extends and improves the preliminary results in [1]). The proposed approach does not change scan cell ordering or stitching.

This paper is organized as follows. Section II provides an overview of the scan chain architecture for multiple expansion ratios. Section III discusses details of dependency analysis and dependency cost function. Experimental results and the conclusion are given in Sections IV and V, respectively.

## II. SCAN CHAIN ARCHITECTURE FOR MULTIPLE EXPANSION RATIOS

To minimize a fault coverage loss in compression mode, various approaches have been studied [13], [14]. A topology-based circuit analysis method is proposed to achieve a high fault coverage [13]. It generates scan chains more optimally by solving the input dependency issues.

Lee and Han [14] introduced an approach to reduce fault coverage loss of scan compression with a continuous decompression flow using unique cells. It analyzes the structural dependency of fan-in cones for each scan cell between each pair of scan chains, and then finds unique cells which are logic gates not belonging to other fan-in cones. If there are more unique cells between each pair of scan chains, there is a higher probability that the detectable faults would be scattered over scan chains. The scan chain pairs holding a number of unique cells are not considered as concatenation candidates. However, fault coverage loss in compression mode comes from the number of overlapping cells between two scan chain pairs rather than the number of unique cells. Hence, it does not always minimize a fault coverage loss in the compression mode.

In this paper, test compression is increased through using multiple expansion ratios. They can be achieved by the following manner. First, a higher expansion ratio can be achieved by dividing a scan chain into a large number of shorter scan chains. Lockup-latches are inserted at the end of each shorter scan chain to prevent clock skew problems. Then, the scan decompressor logic and the scan compressor logic are inserted in front of and after internal scan chains, respectively. In order to reconfigure the scan chains into different multiple expansion ratios, multiplexers are required. If the multiplexers are already placed to support serial mode, they can be used. The control registers are used to change the expansion ratio by concatenating the scan chains. In this architecture, the expansion

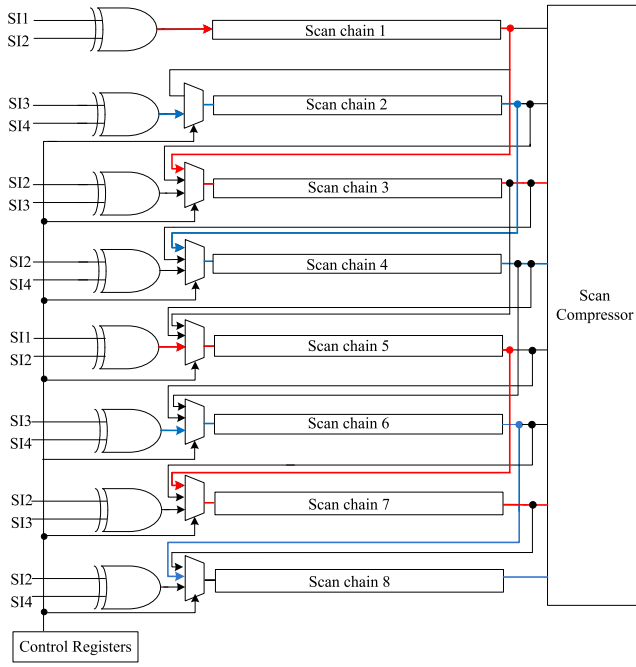


Fig. 2. Example of implementing multiple expansion ratio.

ratio is progressively reduced by concatenating scan chains through multiplexers.

Fig. 2 illustrates how scan chain concatenation is implemented to support multiple expansion ratios. There are four external scan input pins, SI1–SI4, driving eight scan chains. The control registers select the multiplexer input to determine which scan chains are concatenated. As scan chains are concatenated, the expansion ratio is decreased. To achieve a higher fault coverage, it is necessary to select an optimal scan chain concatenation among a number of possible scan chain concatenation combinations. Let us assume that scan chain 1 and 2 are concatenated and the scan chain 3 and 4 are joined together in Fig. 2. The first concatenated scan chain is driven by SI1 and SI2 and the second combined scan chain is driven by SI2 and SI3. The first and second scan chains are dependent on the same free-variable which is SI2 scan input. Because the same scan input, SI2, drives two scan chains, they are not independently controllable and it may cause a fault coverage loss. To increase the number of faults detected after concatenation, scan chain 1 and 3 need to be combined together and scan chain 2 and 4 together, since there is no free-variable dependency among four scan chains.

Scan chain 5 and 7, and 6 and 8 are concatenated together in the same manner as scan chain 1 and 3, 2 and 4 as shown in Fig. 2. Red line illustrates concatenated scan chains driven by input pins SI1 and SI2. Blue line marks concatenated scan chains driven by SI3 and SI4. Even though a scan chain concatenation is performed to avoid free-variable dependencies, SI1 and SI2 simultaneously drive the first concatenated scan chain (scan chain 1 and 3) and the third concatenated scan chain (scan chain 5 and 7). There is a free variable dependency between them. The state of the two concatenated scan chains can be uncontrollable and they can have undetected

### Linear Equations for Scan Chains

$$\begin{aligned} SC_1 &= SI1 \oplus SI2 & SC_2 &= SI2 \oplus SI3 \\ SC_3 &= SI3 \oplus SI4 & SC_4 &= SI3 \oplus SI4 \end{aligned}$$

### Degree of Free-Variable Overlap

$$\begin{aligned} F_{1,3} &= F_{1,4} = 0 \\ F_{1,2} &= F_{2,3} = F_{2,4} = 1/3 & F_{3,4} &= 1 \end{aligned}$$

Fig. 3. Example of linear equations and degree of free-variable overlap.

faults. To minimize the possibility of undetected faults, the number of logical overlap cells in their functional output path needs to be analyzed. In combining scan chains for multiple expansion ratios, scan chain candidates for concatenation have to be chosen to reduce the number of undetected faults.

## III. DEPENDENCY ANALYSIS AND DEPENDENCY COST FUNCTION

This section describes how the proposed multiple expansion ratio is achieved via dependency analysis and dependency cost function. By analyzing the dependency, the proposed method determines which scan cells to concatenate in order to minimize potential conflicts during automatic test pattern generation. This increases the probability of detecting more faults at higher expansion ratios.

To detect a fault, scan cells need to be loaded with certain values. In a serial mode, the conflict does not arise since a unique external scan input pin drives each scan chain. However, a free-variable dependency would exist among scan chains in compression mode. Consequently, because all of the circuit states would not be achievable when scan input pins are shared, conflicts may arise and faults cannot be detected.

The conflict may happen mainly by three types of dependencies. We define each dependency as follows.

*Definition 1:* The free-variable dependency is as a ratio between internal scan chains and external scan input pins.

For example, when two scan cells on different scan chains are driven by the same scan input pins, the conflict can happen as shown in Fig. 1(a). In other words, they depend on the same free-variable and this is referred as free-variable dependency.

*Definition 2:* The structural dependency of fan-out cone is defined as fan-out logic cone overlaps between two scan cells on different scan chains.

For example, when the fan-out logic cones of two scan cells on different scan chains overlap, the faults in the overlapping logic cone may not be controllable if scan input pins are shared. This dependency causes a controllability loss as illustrated in Fig. 1(a).

*Definition 3:* The structural dependency of fan-in cone is defined as fan-in logic cone overlaps between two scan cells on different scan chains.

For example, the fan-in logic cones of two scan cells on different scan chains overlap and the faults or X values in the overlapping logic cone propagate to a compressor through fan-out networks. The dependency causes observability loss as shown in Fig. 1(b) and (c).

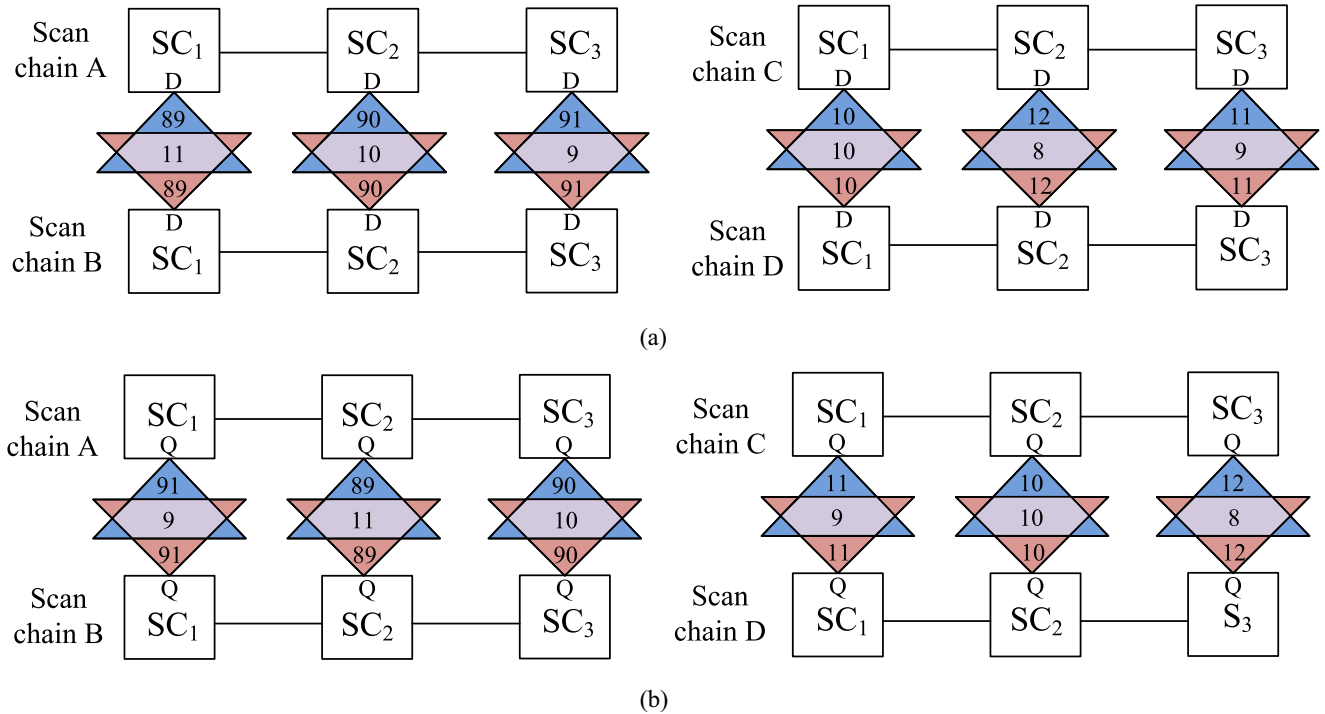


Fig. 4. Examples of (a) fan-in and (b) fan-out cone analysis with four scan chains.

For scan chain concatenation, the dependency analysis needs to be performed to minimize potential conflicts. This paper proposes a dependency cost function for multiple expansion ratios. The free-variable dependency and the structural dependency are considered to introduce the cost function. Scan chains are concatenated based on the dependency cost function.

For a continuous-flow linear decompressor, each scan cell is dependent on a linear combination of the free-variables on the tester. The degree of free-variable dependency is calculated by dividing the same free-variables of two scan cells by the total number of free-variables. Consider the example given in Fig. 3. Scan cells,  $SC_1$  and  $SC_2$ , have one common free-variable,  $SI_2$ . They have a total of three distinct free-variables ( $SI_1$ ,  $SI_2$ , and  $SI_3$ ) in the linear equations. Hence, the degree of free-variable overlap is calculated as  $1/3$ .  $SC_3$  and  $SC_4$  have two common free-variables and their total free-variable is also two, thus a degree of free-variable overlap is 1 which is the maximum amount of free-variable overlap. Since  $SC_1$  and  $SC_3$  do not have any common free-variable, their free-variable overlap is 0. The free-variable overlap of  $SC_1$  and  $SC_4$  is also 0. As the free-variable overlap of two scan cell increases, it becomes less flexible to control the value of two scan cell to detect the fault.

The dependency cost function in (1) relies on both the degree of free-variable overlap ( $F_{ij}$ ) and the structural dependency of fan-in cone ( $SI_{ij}$ ).  $SI_{ij}$  is the amount of logic cone overlap across functional data input path of the of scan cell  $i$  in one scan chain and scan cell  $j$  in the other scan chain. The cost function is given

$$\text{cost\_function 1} = \sum_{i,j} (f_{ij}^e * SI_{ij}^k). \quad (1)$$

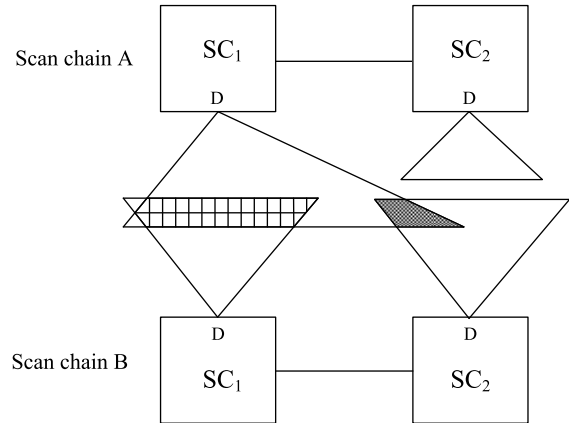


Fig. 5. Example of fan-in logic cone.

The variables  $e$  and  $k$  are used as tuning parameters to adjust the relative weights of the two parts of the cost function to optimize the results for different circuits and different test compression schemes.

The structural dependency of fan-in logic cone is calculated by measuring the number of overlapping gates in the fan-in cone of two scan cells on different scan chains. For example, each scan chain  $A$ ,  $B$ ,  $C$ , and  $D$  has three scan cells as shown in Fig. 4(a). The structural dependency of fan-in cone is determined by measuring logic cone overlaps between two scan cells on different scan chains. In a similar fashion, the structural dependency of fan-out cone can also be found. The example in Fig. 4 shows the fan-in and fan-out cone analysis with their sizes. In Fig. 4(a), the fan-in cone of  $SC_1$  in scan chain  $A$  shares 11 cells with the fan-in cone of  $SC_1$  in scan chain  $B$ . Via fan-in cone analysis, total 30 overlapping cells



TABLE I  
DESIGN DETAILS OF EIGHT IPs

IP	Num of Flip-flops	Total Faults (ea)	Compression Ratio		
			Flow (1)	Flow (2)	Flow (3)
IP 1	58464	2,490,340	3.5	6.8	9.7
IP 2	104173	3,172,200	2.9	4.8	6.7
IP 3	148271	4,515,040	2.7	4.3	5.9
IP 4	177925	5,418,048	2.7	4.3	5.8
IP 5	213510	6,043,394	3.0	5.2	7.4
IP 6	123538	3,496,685	18.8	36.2	48.3
IP 7	158239	4,700,011	22.6	44.4	54.2
IP 8	190027	5,644,237	25.4	49.5	57.9

are found between scan chain *A* and *B*, and 27 overlapping cells are detected between scan chain *C* and *D*. Cost\_function1 would merge the scan chain *A* and *B* since there are more overlapping cells located between the scan chain *A* and *B* than *C* and *D*.

An absolute number of overlapping cells may not represent the structural dependency since the size of fan-in cones differs. A more accurate structural dependency can be found when a size of overlapping cone is considered with a size of logic cones. The proposed method considers a ratio of overlapping cells in the total fan-in cones. In Fig. 4(a), the ratio of overlapping cells of SC<sub>1</sub> in scan chain *A* and *B* is 11/200 [11/(89 + 11 + 89 + 11)]. The ratio from SC<sub>2</sub> and SC<sub>3</sub> are 10/200 and 9/200, respectively. Hence, the total overlapping ratio between scan chain *A* and *B* is 30/200. In the same manner, the total ratio from scan chain *C* and *D* is found as 27/40 (10/40 + 8/40 + 9/40 = 27/40). The overlapping cell ratio analysis finds that the scan chain *C* and *D* has a higher structural dependency. To reflect the accurate structural dependency, the proposed method updates the cost function with a ratio of overlapping cells instead of an absolute size of overlaps.

The overlapping cell ratio of fan-out cone between scan chain pairs should also be investigated to minimize a controllability loss. In Fig. 4(b), the total ratio of overlapping fan-out cone between scan chains *A* and *B* is calculated as 30/200 (9/200 + 11/200 + 10/200 = 30/200) and 27/40 (9/40 + 10/40 + 8/40 = 27/40) is found between scan chain *C* and *D*. From the structural dependency perspective of a fan-out logic cone, a concatenation of scan chain *C* and *D* would detect more faults than merging *A* and *B*.

SC<sub>1</sub>'s in scan chain *A* and *B* are located at the same position of each scan chain and overlapping cells only exist between SC<sub>1</sub>'s in Fig. 4(a) to illustrate a simple example. However, as shown in Fig. 5, there can be overlapping cells between SC<sub>1</sub> in scan chain *A* and SC<sub>2</sub> in scan chain *B* highlighted in gray. Because they can cause an observability loss, we count the number of overlapping cells between SC<sub>1</sub> in scan chain *A* and SC<sub>2</sub> in scan chain *B* as well as the overlap between SC<sub>1</sub>'s in scan chain *A* and *B*.

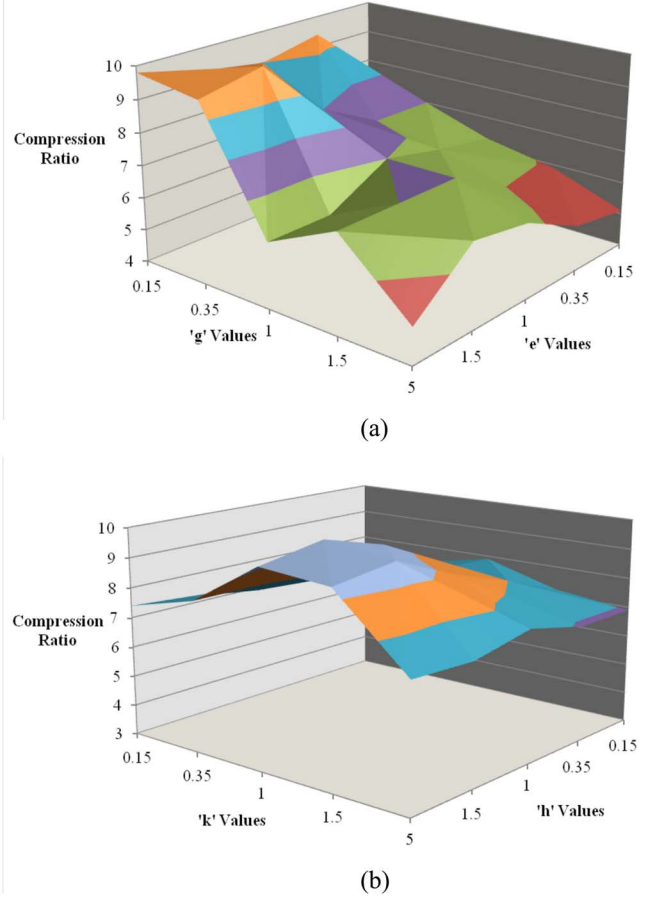


Fig. 6. Plot of compression ratio for IP1 using different parameters. (a) Compression ratio with respect to *e* and *g*. (b) Compression ratio with respect to *k* and *h*.

To enhance the controllability and observability loss in scan chain concatenation, the cost function is proposed as follows:

$$\text{proposed\_cost\_function} = \sum_{i,j} \left( f_{ij}^e \times (SI_{ij}/TI_{ij})^g \right. \\ \left. \times (SO_{ij}/TO_{ij})^h \right) \times (1/L)^k \quad (2)$$

where  $TI_{ij}$  is the total number of cells in one fan-in cone of scan cell<sub>*i*</sub> and scan cell<sub>*j*</sub> belonging to different scan chains.  $SO_{ij}$  is the overlapping logic cone size across functional data output path of scan cell<sub>*i*</sub> and scan cell<sub>*j*</sub>.  $TO_{ij}$  is the total number of cells in the fan-out cone of scan cell<sub>*i*</sub> and scan cell<sub>*j*</sub>. The degree of free-variable overlap ( $F_{ij}$ ) and the structural dependency of fan-out cone  $(SO/TO)^h$  express the possibility of controllability loss and the structural dependency of fan-in cone  $(SI/TI)^g$  considers the possibility of observability loss.  $L$  is the total chain length of two scan chain pairs. Because the longest scan chain length is a dominant factor increasing TDV and TAT, the proposed cost function takes a scan chain length into consideration to minimize a TDV and TAT overhead by scan chain concatenations. The variables *e*, *g*, *h*, and *k* are tuning parameters which alter the relative weights of the four parts in the proposed cost function.

TABLE II  
COMPARISONS OF TEST PATTERNS, LONGEST CHAIN LENGTH, TDV, AND TAT FOR IPs (EXPANSION RATIO = 18)

IP	Scan Compression Flow	Number of Expansion Ratio	Number of Test patterns	Longest Chain Length	Test Data Volume (bit)	Test Application Time (ps)	
IP 1	Flow (1)	1st Expan. Ratio	1060	83	8,973,960	3,519	
		Serial Top Off	237	1720	50,541,864	19,820	
	Flow (2) (Multiple Expansion Ratios Using Equation 1)	1st Expan. Ratio	1060	83	8,973,960	3,519	
		2nd Expan. Ratio	410	166	15,916,080	6,242	
		3rd Expan. Ratio	137	332	20,544,160	8,057	
		4th Expan. Ratio	46	664	23,629,547	9,267	
		Serial Top Off	11	1720	25,558,859	10,023	
	Flow (3) (Multiple Expansion Ratios Using Equation 2)	1st Expan. Ratio	1060	83	8,973,960	3,519	
		2nd Expan. Ratio	373	114	13,312,367	5,221	
		3rd Expan. Ratio	120	160	15,271,647	5,989	
		4th Expan. Ratio	42	235	16,264,800	6,379	
		Serial Top Off	10	1720	18,018,720	7,066	
	IP 2	Flow (1)	1st Expan. Ratio	2120	154	33,300,960	13,059
			Serial Top Off	479	3064	182,997,561	71,764
		Flow (2) (Multiple Expansion Ratios Using Equation 1)	1st Expan. Ratio	2120	154	33,300,960	13,059
2nd Expan. Ratio			1074	308	67,041,744	26,291	
3rd Expan. Ratio			370	554	87,984,300	34,504	
4th Expan. Ratio			123	1109	101,946,003	39,979	
Serial Top Off			26	3064	110,071,497	43,165	
Flow (3) (Multiple Expansion Ratios Using Equation 2)		1st Expan. Ratio	2120	154	33,300,960	13,059	
		2nd Expan. Ratio	1063	203	55,341,797	21,703	
		3rd Expan. Ratio	367	274	65,602,187	25,726	
		4th Expan. Ratio	122	390	70,458,772	27,631	
		Serial Top Off	25	3064	78,271,747	30,695	
IP 3		Flow (1)	1st Expan. Ratio	2756	263	73,932,456	28,993
			Serial Top Off	629	4361	353,586,389	138,661
		Flow (2) (Multiple Expansion Ratios Using Equation 1)	1st Expan. Ratio	2756	263	73,932,456	28,993
	2nd Expan. Ratio		1378	500	144,168,289	56,537	
	3rd Expan. Ratio		431	1000	188,092,039	73,762	
	4th Expan. Ratio		125	2000	213,592,039	83,762	
	Serial Top Off		33	4361	228,270,868	89,518	
	Flow (3) (Multiple Expansion Ratios Using Equation 2)	1st Expan. Ratio	2756	263	73,932,456	28,993	
		2nd Expan. Ratio	1367	341	121,479,450	47,639	
		3rd Expan. Ratio	427	474	142,132,676	55,738	
		4th Expan. Ratio	129	687	151,207,578	59,297	
		Serial Top Off	33	4361	165,886,407	65,054	
	IP 4	Flow (1)	1st Expan. Ratio	3583	316	115,334,631	45,229
			Serial Top Off	760	5233	521,238,949	204,408
		Flow (2) (Multiple Expansion Ratios Using Equation 1)	1st Expan. Ratio	3583	316	115,334,631	45,229
2nd Expan. Ratio			1543	600	209,709,572	82,239	
3rd Expan. Ratio			514	1199	272,626,200	106,912	
4th Expan. Ratio			171	2279	312,473,397	122,539	
Serial Top Off			37	5233	332,223,094	130,284	
Flow (3) (Multiple Expansion Ratios Using Equation 2)		1st Expan. Ratio	3583	316	115,334,631	45,229	
		2nd Expan. Ratio	1540	410	179,737,431	70,485	
		3rd Expan. Ratio	531	541	209,051,809	81,981	
		4th Expan. Ratio	177	763	222,829,567	87,384	
		Serial Top Off	38	5233	243,113,040	95,339	
IP 5		Flow (1)	1st Expan. Ratio	4599	395	185,059,161	72,572
			Serial Top Off	1070	6280	870,181,068	341,248
		Flow (2) (Multiple Expansion Ratios Using Equation 1)	1st Expan. Ratio	4599	395	185,059,161	72,572
	2nd Expan. Ratio		1785	789	328,712,391	128,907	
	3rd Expan. Ratio		595	1578	424,481,211	166,463	
	4th Expan. Ratio		198	3156	488,327,091	191,501	
	Serial Top Off		45	6280	517,150,973	202,805	
	Flow (3) (Multiple Expansion Ratios Using Equation 2)	1st Expan. Ratio	4599	395	185,059,161	72,572	
		2nd Expan. Ratio	1678	504	271,316,644	106,399	
		3rd Expan. Ratio	579	675	311,173,550	122,029	
		4th Expan. Ratio	199	966	330,776,644	129,716	
		Serial Top Off	45	6280	359,600,527	141,020	

To analyze how tuning parameters affect the compression ratio, we study the trend of parameters by varying their values between 0.15 and 5. We plot a compression ratio for

IP1 [which is the intellectual property (IP) used in the experimental results discussed in Section IV] using different values of  $e$  and  $g$  in Fig. 6(a). The compression ratio tends to be

TABLE III  
COMPARISONS OF TEST PATTERNS, LONGEST CHAIN LENGTH, TDV, AND TAT FOR IPs (EXPANSION RATIO = 70)

IP	Scan Compression Flow	Number of Expansion Ratio	Number of Test patterns	Longest Chain Length	Test Data Volume (bit)	Test Application Time (ps)	
IP 6	Flow (1)	1st Expan. Ratio	2636	128	14,171,136	13,496	
		Serial Top Off	91	8824	47,858,671	45,580	
	Flow (2) (Multiple Expansion Ratios Using Equation 1)	1st Expan. Ratio	2636	128	14,171,136	13,496	
		2nd Expan. Ratio	293	256	17,320,277	16,496	
		3rd Expan. Ratio	146	512	20,469,419	19,495	
		4th Expan. Ratio	67	1024	23,332,274	22,221	
		Serial Top Off	4	8824	24,814,730	23,633	
	Flow (3) (Multiple Expansion Ratios Using Equation 2)	1st Expan. Ratio	2636	128	14,171,136	13,496	
		2nd Expan. Ratio	223	156	15,629,582	14,885	
		3rd Expan. Ratio	117	208	16,653,053	15,860	
		4th Expan. Ratio	65	312	17,505,945	16,672	
		Serial Top Off	3	8824	18,617,787	17,731	
	IP 7	Flow (1)	1st Expan. Ratio	3387	168	23,897,261	22,759
			Serial Top Off	89	11303	66,207,038	63,054
Flow (2) (Multiple Expansion Ratios Using Equation 1)		1st Expan. Ratio	3387	168	23,897,261	22,759	
		2nd Expan. Ratio	282	336	27,880,138	26,553	
		3rd Expan. Ratio	94	672	30,535,389	29,081	
		4th Expan. Ratio	29	1344	32,194,921	30,662	
		Serial Top Off	3	11303	33,619,072	32,018	
Flow (3) (Multiple Expansion Ratios Using Equation 2)		1st Expan. Ratio	3387	168	23,897,261	22,759	
		2nd Expan. Ratio	186	206	25,508,903	24,294	
		3rd Expan. Ratio	64	282	26,269,672	25,019	
		4th Expan. Ratio	25	434	26,719,991	25,448	
		Serial Top Off	2	9891	27,550,835	26,239	
IP 8		Flow (1)	1st Expan. Ratio	4567	200	38,362,800	36,536
			Serial Top Off	104	13573	97,534,617	92,890
	Flow (2) (Multiple Expansion Ratios Using Equation 1)	1st Expan. Ratio	4567	200	38,362,800	36,536	
		2nd Expan. Ratio	304	400	43,477,840	41,408	
		3rd Expan. Ratio	87	800	46,400,720	44,191	
		4th Expan. Ratio	29	1600	48,349,307	46,047	
		Serial Top Off	3	13573	50,059,550	47,676	
	Flow (3) (Multiple Expansion Ratios Using Equation 2)	1st Expan. Ratio	4567	200	38,362,800	36,536	
		2nd Expan. Ratio	189	250	40,344,878	38,424	
		3rd Expan. Ratio	54	350	41,137,709	39,179	
		4th Expan. Ratio	22	550	41,636,060	39,654	
		Serial Top Off	2	13573	42,776,222	40,739	

higher where  $e$  is greater than or equal to 1 and  $g$  is smaller than or equal to 1. We find that the compression ratio increases when the possibility of controllability loss ( $f$ )<sup>e</sup> has more weight than the possibility of observability loss ( $SI/TO$ )<sup>g</sup>. The compression ratio is plotted with  $h$  and  $k$  in Fig. 6(b). It tends to be better where  $h$  is greater than or equal to 1 and  $k$  becomes close to 1. The results tells that the compression ratio becomes better when the possibility of controllability loss ( $SO/TO$ )<sup>h</sup> has more weight than chain length ( $1/L$ )<sup>k</sup>. Fig. 6 shows that the compression ratio of IP1 can be enhanced by minimizing controllability loss than reducing observability loss or chain length. The highest compression ratio is found when  $e$ ,  $g$ ,  $h$ , and  $k$  are 1.5, 0.35, 1.5, and 1, respectively. The same tuning parameters are used for other seven IPs in Section IV.

#### IV. EXPERIMENTAL RESULTS

Three scan compression flows are compared. The details are given as follows.

- 1) *Traditional Scan Compression Flow*: Scan chains are compressed with a single fixed expansion ratio.

Any undetected faults in the compression mode are detected with top-off vectors in serial mode.

- 2) *Scan Compression Flow With Multiple Expansion Ratios Using Cost Function 1*: Scan chains are compressed with a higher expansion than a traditional flow. To detect any undetected faults, an expansion ratio is progressively reduced based on the `cost_function1` in (1).
- 3) *Scan Compression Flow With Multiple Expansion Ratios Using Cost Function 2*: Scan chains are compressed with a higher expansion than a traditional flow. The expansion ratio is progressively reduced based on the `proposed_cost_function` in (2).

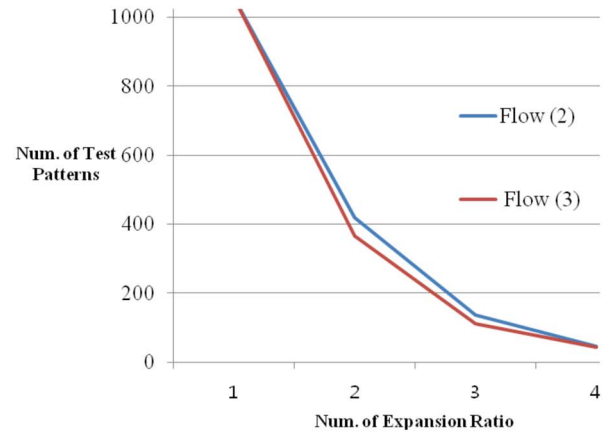
To demonstrate the effectiveness of the proposed methodology, eight industrial IP blocks from state-of-the-art SoC are used. Design details of IPs are given in Table I. The number of scan cells, total faults, and compression ratio are given in the second, third, and last three columns, respectively. The compression ratio is defined as the ratio between TDV of the conventional full scan design and TDV of compression mode. For the flow (2) and (3), the compression ratio is measured after fourth expansion.

The compression ratio of flow (2) is 1.9 times higher than flow (1) since progressive scan chain concatenation based on `cost_function1` can minimize undetected faults. The last column shows a compression ratio by the flow (3). The compression ratio by the flow (3) is 2.4 and 1.3 times higher than the flow (1) and (2).

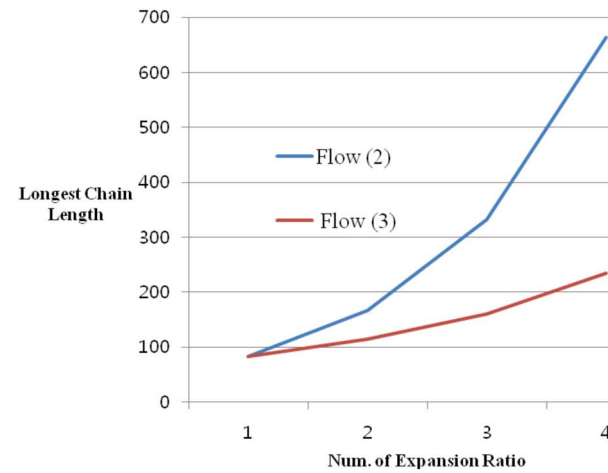
Table II lists the comparisons of test patterns, longest chain length, TDV and TAT for three compression flows with first five IPs given in Table I. The initial expansion ratio of the five IPs are 18. Table III shows the comparisons of TDV and TAT for three flows with last three IPs in Table I. They have the expansion ratio of 70. Since IP6, IP7, and IP8 have higher expansion ratio than IP1–IP5, the compression ratio of IP6–IP8 is higher than others as given in Table I.

The first and second column in Tables II and III shows IPs and the three flows evaluated. The third column gives an expansion ratio for each flow. The number of test patterns for each expansion ratio and serial top off patterns are given in the fourth columns. The serial top off patterns are generated to detect any undetected faults with different scan chain configurations. The fifth column denotes a longest chain length. In IP1, the longest scan chain length is 83 and the serial top off pattern requires a length of 1720 with a flow (1). The flow (2) concatenates scan chains using `cost_function1` and the longest scan chain length keeps increasing with a higher expansion ratio. Because the `cost_function1` does not consider the scan chain length, the longer scan chains are generally selected and concatenated. In the flow (2), with a higher expansion ratio, the longest scan chain length generally becomes twice as long as that of the previous expansion ratio. In IP1, from the first expansion ratio to the second expansion ratio, the longest scan chain length increases from 83 to 166. 166 becomes 332 and it becomes 664 as the expansion ratio changes from second, third and fourth. This trend is also found from other IPs. The flow (2) tends to generate longer and longer scan chains via scan concatenation. However, the flow (3) considers the scan chain length along with controllability and observability loss. Hence, as can be seen, the scan chain length does not linearly increase with a longest scan chain length. TDV and TAT are determined by the longest scan chain. The flow (3) concatenates the scan chains considering a scan chain length and this significantly reduces TDV and TAT. The last two columns list TDV and TAT for each method. TDV and TAT values are accumulated and the final TDV and TAT values for each flow can be found at the serial top off row. The TDV and TAT of IPs are reduced by 53%–64% and 14%–30% compared with the flow (1) and (2), respectively.

Fig. 7 plots the number of test patterns and the longest scan chain length for IP1. In Fig. 7(a), the slope of test pattern reduction in flow (2) is steeper than flow (3) between the second and fourth expansion ratio. The flow (3) requires 373 and 120 test patterns with second and third expansion ratios. It has 253 (373-120) test pattern differences while the flow (2) has 273 (410-137). From third to fourth expansion ratios, there are 78 (120-42) and 91 (137-46) for the flow (3) and (2), respectively. As shown in Fig. 7(b), this happens because the flow (2) has longer scan chains than the flow (3) and the longer scan



(a)



(b)

Fig. 7. Test patterns and longest chain length for IP1 per each expansion ratio. (a) Number of test patterns versus number of expansion ratio. (b) Longest chain length versus number of expansion ratio.

chains generally lead a higher possibility of leaving smaller number of undetected faults.

## V. CONCLUSION

This paper shows the benefits of compressing higher than a traditional scan compression ratio initially and then progressively concatenating scan chains together to reduce the expansion ratio. The scan chain concatenation is performed with a cost function which considers free-variable dependencies, structural dependencies, and scan chain lengths. The proposed method considerably enhances a compression ratio, TDV, and TAT by detecting faults with multiple expansion ratios based on the enhanced cost function.

## REFERENCES

- [1] R. Putman and N. A. Touba, "Using multiple expansion ratios and dependency analysis to improve test compression," in *Proc. 25th IEEE VLSI Test Symp.*, Berkeley, CA, USA, May 2007, pp. 211–218.
- [2] N. A. Touba, "Survey of test vector compression techniques," *IEEE Des. Test. Comput.*, vol. 23, no. 4, pp. 294–303, Apr. 2006.
- [3] A. Dutta and N. A. Touba, "Using limited dependence sequential expansion for decompressing test vectors," in *Proc. Int. Test Conf.*, Santa Clara, CA, USA, Oct. 2006, pp. 1–9.



- [4] S. Mitra and K. S. Kim, "XPAND: An efficient test stimulus compression technique," *IEEE Trans. Comput.*, vol. 55, no. 2, pp. 163–173, Feb. 2006.
- [5] P. Wohl *et al.*, "Minimizing the impact of scan compression," in *Proc. 25th IEEE VLSI Test Symp.*, Berkeley, CA, USA, May 2007, pp. 67–74.
- [6] P. Wohl, J. A. Waicukauski, J. E. Colburn, and M. Sonawane, "Achieving extreme scan compression for SoC designs," in *Proc. Int. Test Conf.*, Seattle, WA, USA, Oct. 2014, pp. 1–8.
- [7] V. Tenentes and X. Kavousianos, "High-quality statistical test compression with narrow ATE interface," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 32, no. 9, pp. 1369–1382, Sep. 2013.
- [8] P. Wohl, J. A. Waicukauski, and T. Finklea, "Increasing PRPG-based compression by delayed justification," in *Proc. Int. Test Conf.*, Austin, TX, USA, Nov. 2010, pp. 1–10.
- [9] G. Mrugalski, N. Mukherjee, J. Rajski, D. Cyszcz, and J. Tyszer, "Compression based on deterministic vector clustering of incompatible test cubes," in *Proc. Int. Test Conf.*, Austin, TX, USA, Nov. 2009, pp. 1–10.
- [10] S. Chebiyam, A. Chandra, and R. Kapur, "Designing effective scan compression solutions for industrial circuits," in *Proc. Int. Symp. Qual. Electron. Design*, Santa Clara, CA, USA, Mar. 2015, pp. 167–172.
- [11] P. Wohl, J. A. Waicukauski, F. Neuveux, and J. E. Colburn, "Hybrid selector for high-X scan compression," in *Proc. Int. Test Conf.*, Anaheim, CA, USA, Nov. 2012, pp. 1–10.
- [12] P. Wohl, J. A. Waicukauski, F. Neuveux, and E. Gizdarski, "Fully X-tolerant, very high scan compression," in *Proc. Design Autom. Conf.*, Anaheim, CA, USA, 2010, pp. 362–367.
- [13] S. Donglikar, M. Banga, M. Chandrasekar, and M. S. Hsiao, "Fast circuit topology based method to configure the scan chains in Illinois scan architecture," in *Proc. Int. Test Conf.*, Austin, TX, USA, Nov. 2009, pp. 1–10.
- [14] D. H. Lee and T. H. Han, "An improvement technique for the test compression ratio and application time of multiple expansion scan chain based SoC using new cost function," in *Proc. Int. Midwest Symp. Circuits Syst.*, College Station, TX, USA, Aug. 2014, pp. 129–132.
- [15] F. F. Hsu, K. M. Butler, and J. H. Patel, "A case study on the implementation of the Illinois scan architecture," in *Proc. IEEE Int. Test Conf.*, Baltimore, MD, USA, Oct. 2001, pp. 538–547.
- [16] A. Sanghani, B. Yang, K. Natarajan, and C. Liu, "Design and implementation of a time-division multiplexing scan architecture using serializer and deserializer in GPU chips," in *Proc. IEEE VLSI Test Symp.*, Dana Point, CA, USA, 2011, pp. 219–224.
- [17] W.-T. Cheng *et al.*, "Scan test bandwidth management for ultralarge-scale system-on-chip architectures," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 23, no. 6, pp. 1050–1062, Jun. 2015.
- [18] J. Janicki *et al.*, "EDT bandwidth management—Practical scenarios for large SoC designs," in *Proc. IEEE Int. Test Conf.*, Anaheim, CA, USA, 2013, pp. 1–10.
- [19] P. Wohl *et al.*, "Two-level compression through selective reseeding," in *Proc. IEEE Int. Test Conf.*, Anaheim, CA, USA, 2013, pp. 1–10.
- [20] A. Chandra, S. Kulkarni, S. Chebiyam, and R. Kapur, "Designing efficient combinational compression architecture for testing industrial circuits," in *Proc. VLSI Design Test (VDAT)*, Ahmedabad, India, 2015, pp. 1–6.



**Taehee Lee** (S'15) received the B.S. degree from Chungnam National University, Daejeon, South Korea, in 1997, the M.S. degree from the University of Southern California, Los Angeles, CA, USA, in 2003, and the Ph.D. degree from the Department of Semiconductor and Display Engineering, Sungkyunkwan University, Suwon, South Korea, in 2017.

He joined the SoC Design Team with Samsung Electronics, Suwon, in 2003. He was engaged in development of Samsung mobile SoC product. From 2006 to 2010, he was with Apple, Cupertino, CA, USA, where he joined application processor development. He is currently a Principal Engineer with Samsung Electronics, where he is researching on physical implementation for Samsung application processors.



**Nur A. Touba** (SM'05–F'09) received the B.S. degree from the University of Minnesota, Minneapolis, MN, USA, in 1990, and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, USA, in 1991 and 1996, respectively, all in electrical engineering.

He is currently a Professor with the Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX, USA.

Dr. Touba was a recipient of the National Science Foundation Early Faulty CAREER Award in 1997, the Best Paper Award at the 2001 VLSI Test Symposium, and the 2008 Defect and Fault Tolerance Symposium. He served as the Program Chair for the 2008 International Test Conference and the General Chair for the 2007 Defect and Fault Tolerance Symposium. He currently serves on the program committee for the Design Automation and Test in Europe Conference, International On-Line Test Symposium, European Test Symposium, Asian Test Symposium, and Defect and Fault Tolerance Symposium.



**Joon-Sung Yang** (S'05–M'09) received the B.S. degree from Yonsei University, Seoul, South Korea, in 2003, and the M.S. and Ph.D. degrees from the University of Texas at Austin, Austin, TX, USA, in 2007 and 2009, respectively, all in electrical and computer engineering.

He was with Intel Corporation, Austin, TX, USA, for four years. He is currently an Associate Professor with Sungkyunkwan University, Suwon, South Korea. His current research interests include very large scale integration (VLSI) testing, silicon debug, and nanometer scale test and design methodologies.

Dr. Yang was a recipient of the Korea Science and Engineering Foundation Scholarship in 2005, the Best Paper Award at 2008 IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems and at 2016 IEEE International SoC Design Conference, and was nominated for the Best Paper Award at 2013 IEEE VLSI Test Symposium.