Making Split Fabrication Synergistically Secure and Manufacturable

Lang Feng*, Yujie Wang†¶, Jiang Hu*, Wai-Kei Mak§, Jeyavijayan (JV) Rajendran‡
*Department of Electrical & Computer Engineering, Texas A&M University
†College of Electronic Information and Optical Engineering, Nankai University
§Department of Computer Science, National Tsing Hua University
‡Department of Electrical and Computer Engineering, The University of Texas at Dallas
¶Institute of Computing Technology, Chinese Academy of Sciences
flwave@tamu.edu; tjwangyj@hotmail.com;
jianghu@tamu.edu; wkmak@cs.nthu.edu.tw; jv.ee@utdallas.edu;

***Abstract*—Split fabrication is a promising approach to security against attacks by untrusted foundries. While existing split fabrication methods consider the overhead of conventional objectives such as wirelength and timing, they mostly neglect manufacturability—an unavoidable challenge in nanometer technologies. Observing that security and manufacturability can be addressed in a synergistic manner, this work introduces routing techniques that can simultaneously improve both security and manufacturability in terms of either Chemical Mechanical Planarization (CMP) uniformity or Self-Aligned Double Patterning (SADP) compliance. The effectiveness of these techniques is confirmed by experiments on benchmark circuits.**

I. Introduction

The sustained progress of VLSI technology brings challenges at multiple fronts. Transistor feature size of the latest technologies is only several nanometers while the lithography wavelength is still at 193nm. This mismatch gives rise to the IC manufacturability challenge and entails process technologies such as multi-patterning, immersion lithography, and electron-beam lithography [1]. Consequently, semiconductor fabrication becomes drastically complex and expensive, and its business is increasingly concentrated to a few high-end foundries, which are often offshore. Offshore fabrication leads to various attacks such as piracy and Trojan insertion.

Recently, the concept of split fabrication is proposed for security against untrusted foundries. By separating the fabrications of Front-End-Of-Line (FEOL) and Back-End-Of-Line (BEOL) at different foundries, the difficulty for attacks at a single foundry is considerably increased [2], [3]. Even with the increased difficulty, however, a foundry may still be able to reverse engineer an entire design according to design conventions and therefore continue to launch security attacks [2], [4], [5]. Cell placement [4] and routing perturbation [6] techniques have been developed to further enhance security for split fabrication. Existing security methods address the associated overhead, mostly in terms of conventional design objectives, such as wirelength and circuit timing [2], [4]–[8].

However, almost none of the previous works on split fabrication considers the manufacturability issue, which is fundamentally important and cannot be ignored in practice. Usually, the security for split fabrication is improved by modifying circuit layout such that the original design intention is disguised. Thus, reverse engineering by an attacker becomes more difficult. At the same time, modification affects manufacturability as semiconductor manufacturing process is usually sensitive to layout patterns. Indeed, most design-for-manufacturability (DFM) works are based on layout optimization. Manufacturability-oblivious security enhancement may inadvertently degrade lithography printability, decrease manufacturing yield, increase manufacturing cost or even cause unsolvable manufacturing hot-spots [1]. On the one hand, as shown in our results (see Section VI), all the previous techniques that ensure the security of split manufacturing leads to DFM violations, making them not manufacturable. On the other hand, designs that do not have DFM violations need not be secure, because attackers can use DFM hints to retrieve the missing parts, similar to using routing and placement hints [2], [4], [5]. Thus, it is imperative to consider manufacturability in conjunction with security-driven layout modification.

In fact, a layout modification can benefit both manufacturability and security , e.g., some routing detour can make a layout to be deceptive to attackers and improve the uniformity of wire density, which facilitates CMP. Likewise, wire extension can help to improve SADP compliance and simultaneously confuse attackers. Concurrent consideration of security and manufacturability in the layout is not only a convenience but also a necessity. Sequentially handling the two objectives one after another is inadequate for solving both problems. If security is first addressed and then manufacturability is taken care later in a separate step, the manufacturability-driven layout changes may overwrite the security obtained in earlier steps, and vice versa.

In this work, we make the following contributions:

- A routing modification method is proposed to improve both the security of split fabrication and CMP uniformity. This technique is targeted to IC products using ordinary modern process technologies.
- A wire extension technique is developed to simultaneously improve security and compliance with SADP. This approach is for products whose FEOL is fabricated with the latest high-end process technologies.
- An improvement is made to the state-of-the-art attack

[0] This work is partially supported by NSF (CNS-1618824), NSF (CNS-1618797), SRC (2016-TS-2688), SRC (2016-TS-2689), MOST (104-2628-E-007-003-MY3), and MOST (106-2918-I-007-008).

978-1-5386-3093-8/17/$31.00 ©2017 IEEE

method. It can reduce connection errors and Hamming distance[1] from 63% to 37% and from 17% to 14%, respectively.

- Through experiments, we show that DFM-oblivious security techniques can significantly degrade CMP uniformity or cause more SADP violations.
- The proposed techniques are validated on benchmark circuits using the state-of-the-art attack framework [4]. The results indicate that our CMP-friendly defense routing can reduce CMP variations by 37% and 25% for FEOL and BEOL layers, respectively. Our simultaneous security- and SADP-driven wire extension can reduce 97% of SADP violations. At the same time, our techniques achieve similar security as the latest previous work under the state-of-the-art attack [4]. In addition, the delay and wirelength overheads for our techniques are less than 1% and 4%, respectively.

To the best of our knowledge, this is the first work on simultaneously improving security and manufacturability for split fabrication.

II. Prior Work

A. Security of Split Fabrication

The vulnerability of split fabrication alone to attacks was first demonstrated in [2]. This work considered hierarchical designs, where the BEOL wires that connect different blocks can be easily guessed by attackers according to the proximity of the corresponding pins. Later, a set of obfuscation techniques [9] is suggested to improve the security for split fabrication. A placement perturbation technique is proposed in [4] to confuse attackers by moderately violating common design conventions. This work also describes an advanced network flow based attack method for flattened designs. Along the same direction, routing perturbation techniques are introduced in [5], [6] to enhance the security for split fabrication. Improvement to the proximity attack [2] is also discussed in [5]. The partitioning between FEOL and BEOL layers for security is studied in [7]. Security of memory and analog IP blocks in split fabrication is investigated in [8].

It is noticed in [10] that 3D IC can play a similar role as split fabrication for hardware security. The concept of K-security and accordingly the wire lifting technique are introduced in [11] for 3D IC chips; this work has a different threat model compared to others—the attacker has access to the golden netlist. The work of [12] studies layer partitioning for 3D ICs such that the difficulty of attacks is increased. It further suggests a placement technique to enhance the security.

Note that none of these security solutions addresses the manufacturability of their solutions. In fact, our results indicate that many of these solutions are not manufacturable, as they have DFM violations. Hence, we focus on developing solutions that are both manufacturable and secure.

B. Design for Manufacturability

Design for manufacturability (DFM) has been an active research area for almost two decades. This section summarizes several works on layout design considering CMP and SADP, which are relevant to our methods. An early work on CMP-aware routing is [13], which shows that optimizing wire density for CMP is not equivalent to minimizing wire congestion in conventional routing. It modifies a conventional global router by incorporating wire density and the impact on timing into consideration. Another CMP routing work is [14]. Its contribution is the use of Voronoi diagram for accurately estimate wire density for CMP-driven global routing. In [15], an SADP-aware detailed routing is developed in conjunction with layout decomposition. The work of [16] is an SADP-driven routing method considering cut process. SADP routing for 1D gridded designs is studied in [17] and [18].

III. Preliminaries

A. Definitions

Some technical terms used in this paper are:

Dangling wire: A wire on the topmost FEOL layer, with one end connected to driver/sink through a via to lower layers and the other end connected to a via to BEOL layers.

Sink wire: A dangling wire that is connected to sink through lower metal layers. Its connection to the driver is through BEOL layers. An example is the wire from b to B in Figure 1.

Source wire: A dangling wire connected to the driver through lower metal layers. Its connection with sinks is through BEOL layers, e.g., the wire from A to a in Figure 1.

Complete wire: A wire at the topmost FEOL layer and connected with the driver/sink through lower metal layers, i.e., without using BEOL layers, such as the wire from c to d in Figure 1.

Up-via: A via connecting the topmost FEOL layer and the bottom BEOL layer. In Figure 1, a and b are two up-vias.

Down-via: A via connecting the topmost FEOL layer with its lower metal layers.

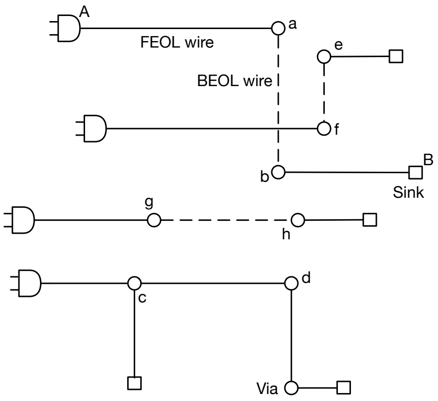

Fig. 1. Illustration for definitions.

B. Attack Method

In this work, we assume the FEOL is fabricated at a high-end untrusted foundry, and the BEOL is fabricated at a trusted foundry. The attack by untrusted foundry is the state-of-the-art network flow method [4] as it is one of the most effective attack techniques. In the network model, there is an edge between a source wire and a sink wire. The edge cost is defined according to hints from common design conventions, e.g., if the up-vias of the two wires are close to each other like a and

[1] The hamming distance is calculated between the combinational logic output of the original design and that of the attacked design

e in Figure 1, the corresponding edge cost tends to be small. Thus, they are more likely to be connected by the attack. We improved this method by considering preferred routing direction on related layers. Suppose the horizontal solid lines in Figure 1 indicate wires on the topmost FEOL layer. In other words, the preferred routing direction of the topmost FEOL layer is horizontal, in this example. In such scenario, vertical alignment between two up-vias (like a and b in Figure 1) is much more common than horizontal alignments like g and h. In chip layout, the distance between to points (x_1, y_1) and (x_2, y_2) is in Manhattan space as $|x_1 - x_2| + |y_1 - y_2|$. To improve the attack, we changed the distance to be a weighted form $w_x \cdot |x_1 - x_2| + w_y \cdot |y_1 - y_2|$, where w_x and w_y are the horizontal and vertical weights, respectively. For the scenario of Figure 1, where the preferred routing direction of the topmost FEOL layer is horizontal, we make $w_x > w_y$ such that the horizontal distance between two up-vias is emphasized. In other words, we use different horizontal and vertical weighting factors to encourage connections conforming to preferred routing directions in layout. Similarly, the horizontal alignment of a layer can be preferred by altering the weights.

C. Framework of Routing for Security and Manufacturability

Two routing-based defense methods are proposed for split fabrication. One considers the CMP-friendliness and the other addresses SADP-compliance. Both methods share the same framework, although they have significant differences. Taking a fully placed and routed circuit as the input, the framework consists of two steps.

- Step 1: Layer elevation. This is to selectively move some FEOL wires to BEOL layers such that they become invisible to attackers and manufacturability is benefited.
- Step 2: Rerouting. Some FEOL wires are rerouted to improve both security and manufacturability.

The details of these steps are elaborated in sections IV and V.

IV. CMP-Friendly Routing Defense

A. Background on CMP

Chemical Mechanical Planarization (CMP) is an important semiconductor manufacturing step. After one layer of metal and Inter-Layer Dielectric (ILD) is finished, CMP is performed so that the surface is flat enough for fabricating another layer. The effect of planarization depends on metal wire density as illustrated in Figure 2. If wire density is not uniform, the surface corresponding to the sparse region is lower than that in the dense region. The unevenness causes not only manufacturing difficult for upper metal layers but also ILD thickness variation, which worsens circuit timing variability.

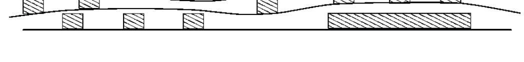

Fig. 2. Uneven surface after CMP due to non-uniform wire density [19]. The shaded rectangles indicate cross sections of metal wires.

If the oxide density before CMP at location (x, y) is $\rho_0(x, y)$, the ILD thickness z at this location can be estimated by [19]–[21]

$$\begin{cases} z = z_0 - [K_i t/\rho_0(x,y)] & t < (\rho_0 z_1/K_i) \\ z = z_0 - z_1 - K_i t + \rho_0(x,y) z_1 & t > (\rho_0 z_1/K_i) \end{cases}$$

where t is the polish time, and K_i, z_0 and z_1 are constant parameters. Since the oxide density $\rho_0(x, y)$ is directly determined by wire density, the surface uniformity or the variability of z also depends on the variability of wire density. Wire density also affects metal thickness t_i [13], [14] as

$$t_i = \alpha(1 - \frac{m_i^2}{\beta})$$

where m_i is the wire density at region i, and α and β are constant parameters. It is shown in [13], [14] that wire density must be explicitly considered in routing algorithms in order to reduce the CMP related variations.

B. Layer Elevation

Given a routed circuit, the first step of the CMP-friendly routing defense is layer elevation, where some wire segments are moved from an FEOL layer to a BEOL layer. Such move makes the elevated wires from visible to invisible to attackers and hence improves security. At the same time, it affects wire density and CMP as well. The wire segments to be elevated are selected according to several principles.

1) The wire segment has a significant logic difference from its neighboring wires. As such, an incorrect connection in attacking this wire may lead to more signal differences.
2) This wire segment has large observability so that an attack error can easily affect the circuit primary output signals.
3) This wire segment is originally at a wire-dense region. The wire density of this region would be reduced by the layer elevation and makes the corresponding FEOL layer have more uniform wire density.
4) The BEOL region where the wire segment is elevated to has low wire density so that the density of the corresponding BEOL layer is more uniform.

Items 1 and 2 are for security enhancement like in [6]. Items 3 and 4 are the new constraints, which intend to improve the uniformity of wire density and facilitate improved CMP.

C. Motivation and Wire Selection

After layer elevation, a set of wire segments is selected for rerouting. The rerouting has two purposes: CMP-friendliness and security improvement. For CMP-friendliness, one wishes to select wires in dense regions so that they can be rerouted into sparse regions. When rerouting a segment, the mechanism for security enhancement is two-fold. This is illustrated by an example in Figure 3, where we consider to reroute wire segment from driver B. The original layout is given in (a), and the rerouting result is in (b). The FEOL and BEOL wires are represented by solid and dashed lines, respectively. By the rerouting, the wire detour makes via point b is closer to c' than a. Such proximity can mislead an attacker to think b should be connected with c'. Then, via c' serves as a decoy to the net driven by A. The blue region in Figure 3 is designated as a target-decoy region for the net driven by A.

The additional consideration of CMP also helps security. The rerouting in Figure 3(b) is a detour to avoid the high wire density region. As such, an attacker would regard such detour as CMP-driven, and the defense purpose is disguised. In contrast, if a security-driven detour is around a sparse region,

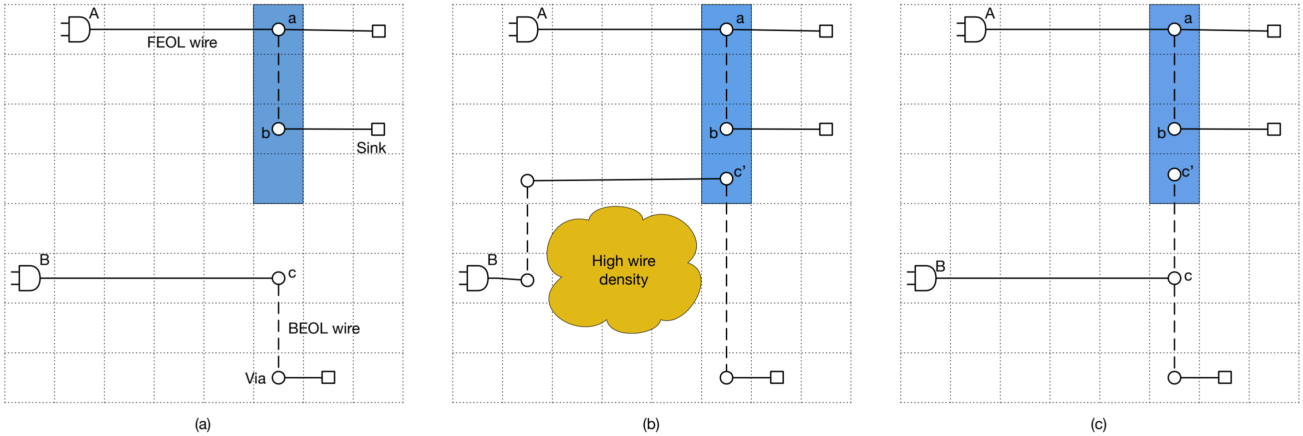

Fig. 3. (a) The original routing; (b) security and CMP-driven rerouting. (c) dangling stub for security. The blue squares indicate the decoy region for the net driven by gate A.

which is allowed in [6], an attacker would feel suspicious and may realize that the detour is a defense measure.

Besides detour, we consider another defense approach, which is the dangling stub in Figure 3(c) and not used in previous routing defense works [5], [6]. The dangling via point c' in Figure 3(c) may also mislead an attacker to connect b with c'. Although such layout is against common design convention, its wire and timing overhead may be less than that of a detour.

According to the mechanism in Figure 3, the security aspects of the wire selection is to see if a wire near the decoy region of another net. Also, the logic difference between the net to be selected (net driven by B in Figure 3) and the net to be decoyed (net driven by A in Figure 3) should be large, and so is the observability of the net to be decoyed.

Please note we use a different strategy from the K-security in [11]. The concept of K-security is to provide $K-1$ additional connection options, which are *equally good* as the original connection, to attackers. Then, it is very difficult for an attacker to make the correct connection among K options, especially when K is large. In contrast, our defense just provides one additional connection option (decoy), which looks *better* than the original connection. For example, in Figure 3(b), c' is closer to b than a and therefore looks better than the original connection between a and b. Since rerouting usually comes with wirelength and delay overhead; our strategy requires less rerouting and lower overhead.

D. Wire Rerouting Method

For the wire segments selected according to Section IV-C, we reroute them one at a time. Our approach has a key difference from the routing perturbation in [6]. As CMP is not considered, the rerouting of wire segment in [6] can be solely focused on security. For example, in rerouting the wire segment connected to driver B in Figure 3, only nets driven by A and B are considered. By contrast, our method needs to consider wire density in addition. Therefore, we divide layout area into an array of tiles like in Figure 3 and perform a coarse rerouting on the tiles, which is further refined as like detailed routing. Moreover, we consider the application of dangling stub like Figure 3(c) while [6] does not.

In our methods, a routing graph is built according to the routing tiles as in Figure 4. In this graph, each node corresponds to one tile and there is a pair of directed edges between two adjacent tile nodes. Please note this is different from conventional global routing where there is only one undirected edge between two adjacent tiles. This difference is due to the fact that wire congestion in conventional global routing is evaluated across a boundary between two neighboring tiles while the effective oxide density [19] is evaluated within each tile. To capture the preference to wires from high-density tiles to low-density tiles, such directed edges are needed.

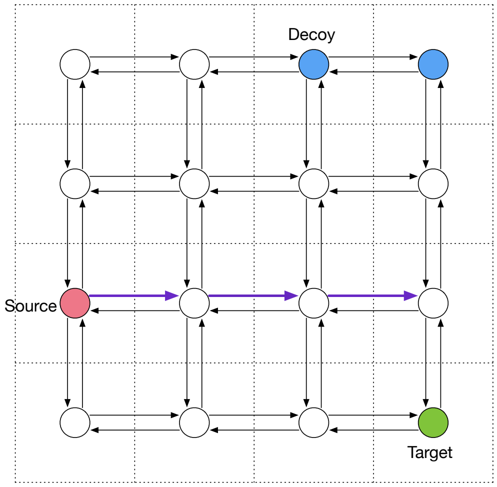

Fig. 4. Routing graph.

Considering an edge $(a \rightarrow b)$ from tile a to b, where the effective oxide densities are d_a and d_b, respectively. The edge is associated with a density weight $w_d(a \rightarrow b)$ defined as

$$w_d(a \rightarrow b) = \begin{cases} \frac{1}{d_a - d_b + \delta}, & \text{if } d_a \geq d_b \\ K \cdot (d_b - d_a) + \frac{1}{\delta}, & \text{if } d_a < d_b \end{cases} \quad (1)$$

where δ and K are two constant parameters. This weight definition implicitly addresses routing congestion and wirelength as well. If tile b is very congested, its oxide density and edge weight $w_d(a \rightarrow b)$ are both high. As a result, this weight definition resists connection through tile b. As a density weight is always non-trivially positive, a long wire detour or wirelength is also penalized.

The rerouting is to find a new wire connection among the source node, like B, the target node like c and decoy nodes like the blue tiles in Figure 3. This is equivalent to constructing a Steiner tree on the graph shown in Figure 4, which is a well-known NP-hard problem. Therefore, we design a heuristic based on the Dijkstra's shortest path algorithm.

If the Dijkstra's algorithm is performed using the density weight w_d for edges, we can obtain the minimum weight paths from the source to the target and decoy nodes, respectively. However, such approach neglects the benefit of sharing the two paths. In order to encourage sharing between the two paths, which will result in the dangling stub, like Figure 3(c), we augment edge weight with sharing weight, which is defined as follows. First, we draw two bounding boxes—one is between the source and the target, and the other is between the source and decoy nodes. If there is no edge overlap between the two boxes, no change is made to edges. If an edge is on the shared boundary between the two boxes and along the direction from the source to the target and decoy nodes, this edge is called a *sharing edge*, which is indicated as the purple thickened edges in Figure 4. If an edge is k hops away from any sharing edges, its weight is incremented by $k \cdot \epsilon$, where ϵ is a small constant. With the augmented weight, the Dijkstra's algorithm is performed on the routing graph to obtain paths from the source to the target and decoy nodes as the rerouting results.

V. SADP-Compliant Routing Defense

A. Background on SADP

For sub-16nm technology nodes, self-aligned double patterning (SADP) is an excellent option for fabricating the lower metal layers of a design to achieve the required fine metal pitches. In addition, unidirectional routing is usually advocated in these layers for its higher manufacturing yield and uniformity compared to 2D routing. In such case, the SADP process will first print a sea of parallel tracks as Figure 5(a), and then wire-end cuts are printed using a cut mask to cut up the tracks into the target wire segments and some dummy wire segments. Due to the constraints on the cut mask of self-aligned double patterning, two cuts cannot be too close to each other, i.e., there is the minimum spacing rule between cuts. Some wire-end of the target wire segments is extended after initial routing in order to avoid violation of this rule, or make a layout SADP compliant [17], [18].

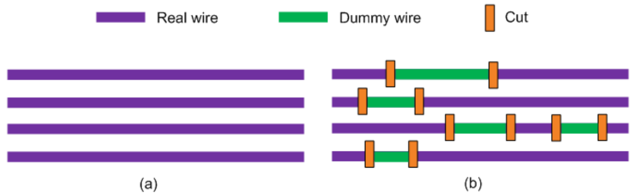

Fig. 5. SADP starts with dense lines generation as in (a). By cutting, the desired patterns are the purple parts in (b).

B. Security Enhancement under SADP

We introduce an approach to modify a 1-D layout on the topmost FEOL layer to facilitate SADP-compliance and enhance security simultaneously. The context is that FEOL foundry uses high-end process technology including SADP while BEOL foundry operates with the conventional manufacturing process without SADP. Our approach follows the same 2-step framework of Section III-C. The first step is layer elevation, which is similar to the one in CMP routing (Section IV-B) except that wire density is not considered, as SADP layout with wire-end cuts has near uniform wire density.

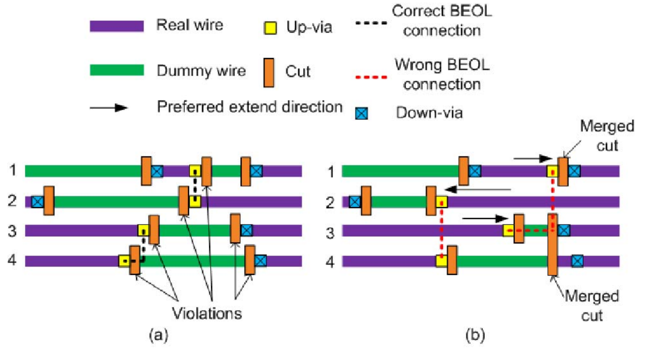

Fig. 6. (a) Original layout with SADP violations. (b) Wire extension for both SADP-compliance and security.

The second step of rerouting is actually wire extension of FEOL wires as in [17]. Please note the wire extension of FEOL wires inevitably causes rerouting of connected BEOL wires. We use the example in Figure 6 to illustrate how such wire extension can simultaneously help SADP-compliance and security. Figure 6(a) shows a part of the original topmost FEOL layer. Recall that an up-via is a via connecting the topmost FEOL layer to the bottom BEOL layer. A down-via is a via connecting the topmost FEOL layer to the layer below it. In Figure 6(a), there are three pairs of cuts too close to each other, and each causes an SADP rule violation. Wire extension where four cuts are relocated is shown in Figure 6(b). After the wire extension, there is no SADP rule violation. Moreover, the proximity of the up-vias is changed by the wire extension. As a result, an attacker is misled to the wrong (red) connection in (b) while the original design is the black BEOL connections in (a). Note that when a wire-end with an up-via is extended, the corresponding up-via is moved accordingly. On the other hand, a wire-end with a down-via can be extended, but the corresponding down-via will not be moved as shown in the bottom right corner of Figure 6(b). Finally, we allow two cuts to be merged into a single one as the two cuts in the top line and the two cuts near the bottom right corner in Figure 6.

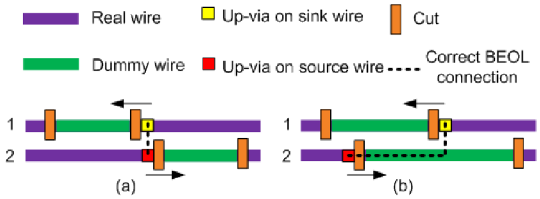

Fig. 7. Wire extension enhances security in (a) but degrades security in (b).

The wire extension for simultaneous SADP-compliance and security is realized using Integer Linear Programming

(ILP) like [17], but the ILP here is quite different from [17] and needs to handle more complicated situations. The ILP in [17] attempts to minimize total wire extension (or overhead) and eliminate the SADP violations. At the same time, it is subject to: (1) SADP manufacturability constraints, (2) maximum wire extension constraint. When security is considered, the problem is more complex as wire extension can degrade security as well. In Figure 7(a), the extension enhances security as it moves the two up-vias apart. By contrast, the extension in Figure 7(b) makes the two up-vias closer and then the correct connection is easier to be figured out by an attacker. Therefore, whether or not to maximize or minimize an extension in our ILP depends on layout scenarios. In Figure 8, we summarize if wire extension is good or bad for security in eight different scenarios.

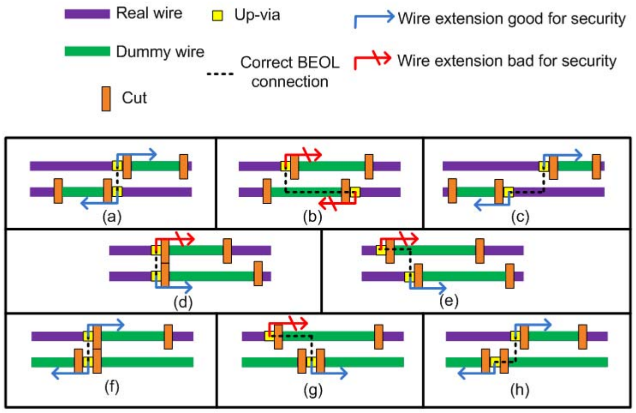

Fig. 8. Scenarios of different security implications by wire extension.

When security is considered, another complicated situation is multi-pin net. In Figure 9, the source wire in the middle (associated with the red via) and the four sink wires on the other lines belong to the same net. For the sink wires, whether or not line extension is good for security is indicated by the blue and red arrows. For the source wire, wire extension toward right makes a connection to the sink wire on line 5 easier, i.e., the security of the source and the 5th line sink connection is weakened. However, the extension of the source wire increases the security for connections with the other sink wires. In this situation, we decide if increasing a source wire extension is preferred or not according to majority vote. More specifically, we prefer to increase (decrease) the extension of a source wire if the number of sink wire connections with security improvement is more (less) than the number of sink wire connections with security degradation.

Our ILP to determine the wire extension of each wire for simultaneous SADP-compliance and security enhancement is constructed as follows. We incorporate the constraints similar to [17] to enforce (i) two wire-end cuts in the same track or nearby track have to maintain a minimum spacing min_s or have to be merged (as in Figure 6), and (ii) enforce a maximum allowed wire extension δ_i for each wire i due to timing consideration. Our objective function is to maximize the difference in the total security improvement and the total wire extension. When an extension of a wire-end with an up-via is good for security, a positive security improvement proportional to the extension length is accumulated. On the other hand, when an extension of a wire-end with an up-via is bad for security, a negative security improvement proportional to the extension length is incurred. Any non-dangling end of a wire (i.e., any end with a down-via as in Figure 6) can be extended, but the corresponding security improvement is always zero. Any wire-end cut conflict that cannot be resolved in the final layout is handled by an e-beam shot in [17], but we simply report it as an SADP violation in this paper.

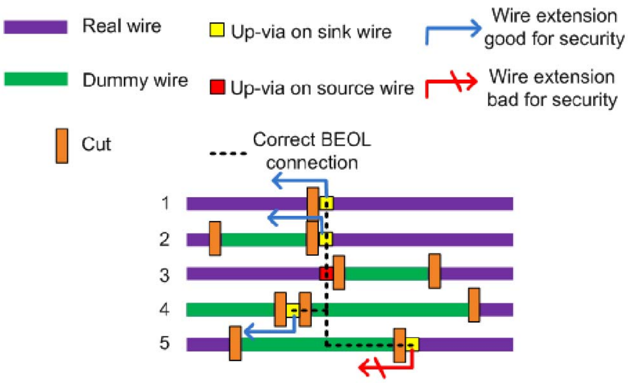

Fig. 9. The source wire in the middle line and the four sink wires on the other lines belong to the same net.

VI. Experiment Results

A. Setup and Comparison Methodology

The experiments are conducted on the five largest circuits in ISCAS'85 benchmark suites and the seven largest circuits in ITC'99 benchmark suites. These circuits are synthesized by Synopsys Design Compiler using 45nm standard cell library. The initial layout is generated by Cadence SoC Encounter. The timing analysis is obtained through Synopsys PrimeTime. The defense and attack algorithms are implemented and run on a PC with Intel 3.4GHz CPU with 16Gb memory. The ILPs are solved by solver Gurobi 7.0.2. The defense results are complete layouts with detailed routing and design rule checking are performed by Cadence SoC Encounter.

Comparisons are made to the following layout results.

- Original: The layout generated by Cadence SoC Encounter without routing defense or manufacturability improvement.
- Security-only: The latest previous work on routing based security for split fabrication [6], which is performed on the original layout.
- Security+CMP: Our CMP-friendly routing defense (Section IV) performed on the original layout.
- Security+SADP: Our SADP-compliant routing defense (Section V) performed on the original layout.

The layout results from all the above methods are attacked using the network flow method [4] with our improvement (Section III-B) to evaluate their security.

The results are evaluated with the following metrics.

1) Security
   - Connection error: the percentage of wrong connections by the attack (Section III-B) among all missing wires, which are on BEOL layers.
   - Hamming distance: the Hamming distance between output vectors of the original complete design and the reverse

TABLE I
RESULTS OF THE ORIGINAL DESIGN, DESIGN WITH SECURITY ONLY ROUTING [6], AND DESIGN WITH CMP-FRIENDLY SECURITY ROUTING.

Circuits	# nets	Original Design				Security Only [6]				Security + CMP			
		Attack [4]		Improved Attack (Sec. III-B)		Improved Attack (Sec. III-B)		FEOL $\Delta Var(den)$ (%)	BEOL $\Delta Var(den)$ (%)	Improved Attack (Sec. III-B)		FEOL $\Delta Var(den)$ (%)	BEOL $\Delta Var(den)$ (%)
		Connection error (%)	Hamming distance (%)	Connection error (%)	Hamming distance (%)	Connection error (%)	Hamming distance (%)			Connection error (%)	Hamming distance (%)		
c2670	607	51.9	14.5	38.0	11.9	68.6	20.3	-17.2	24.3	66.7	20.5	-9.2	-17.7
c3540	638	55.3	16.9	22.4	13.5	71.2	38.5	-7.4	68.9	88.5	35.0	-44.7	-1.4
c5315	997	54.7	15.9	41.2	13.3	57.9	24.1	-21.8	5.7	85.1	23.6	-59.7	-16.1
c6288	1921	6.1	2.4	0.0	0.0	72.3	44.3	3.8	-27.7	66.9	40.6	-0.4	-20.9
c7552	1041	59.5	18.6	45.2	15.5	75.7	30.7	-19.3	6.8	78.7	24.7	-59.4	-28.3
b14_1	3018	66.8	10.6	15.9	4.3	83.7	25.2	-17.7	98.9	68.5	21.7	-30.3	-28.1
b15	6018	73.0	10.2	48.6	9.5	75.7	19.7	-13.3	-21.8	80.1	20.3	-22.2	-8.6
b17	18613	78.3	17.3	54.8	13.8	93.8	33.7	-44.1	-71.0	82.3	22.4	-70.1	-71.4
b18	55029	87.9	21.6	67.8	18.1	94.5	34.2	-30.7	105.8	89.3	27.6	-48.0	38.2
b20	8109	84.6	30.7	44.3	26.5	89.9	37.3	-11.4	-14.2	77.0	32.2	-28.3	-46.1
b21	8153	84.3	33.1	47.8	26.1	67.5	36.3	-18.8	-11.7	79.3	36.5	10.4	-48.6
b22	12065	50.1	13.9	15.4	9.8	87.5	32.6	-13.9	55.8	93.2	29.8	-77.6	-45.6
Avg.		62.7	17.1	36.8	13.5	78.2	31.4	-17.7	18.3	79.6	27.9	-36.6	-24.5

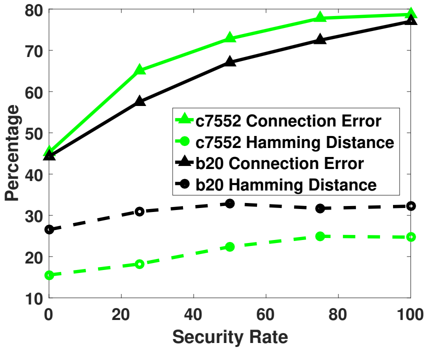

Fig. 10. Security versus % selected wires being rerouted.

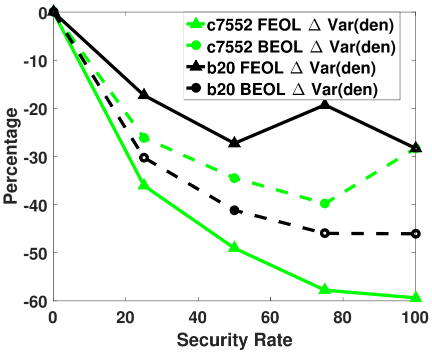

Fig. 11. Wire density variance change versus % selected wires being rerouted.

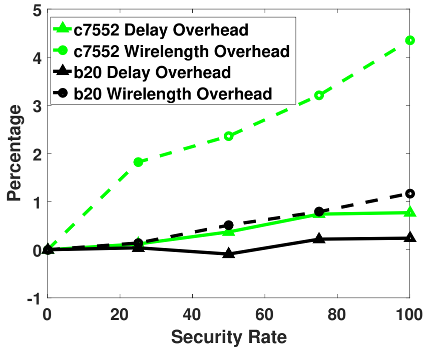

Fig. 12. Delay and wirelength overhead versus % selected wires being rerouted.

engineered design by the attack (Section III-B). Security is strong when Hamming distance is near 50%. The result is obtained by 5K runs of Monte Carlo simulation. We found that the 5K-run results are usually very close to 50K-run results, typically with less than 1% difference.

2) Manufacturability

- $\Delta Var(den)$: the change of wire density variance compared to the original layout. A negative change means variation decrease and is preferred.
- # SADP violations: the number of SADP rule violations. The violations can be solved by using electron-beam lithography at a certain expense. Thus, such violation is permitted but not preferred.

3) Overhead

- Delay overhead: the critical path delay change compared to the original layout according to Synopsys PrimeTime.
- Wirelength overhead: the total wirelength change compared to the original layout.

B. Results of CMP-Friendly Defense

The main results of CMP-friendly security routing are summarized in Table I along with those from the original layout and the previous work [6]. On average, the secure routing of [6] can increase attack connection errors and Hamming distance from 37% to 78% and from 14% to 31%, respectively. However, it increases BEOL wire density variance by 18%, which implies a significant degradation of manufacturability. By contrast, our security+CMP approach can reduce wire density variance for FEOL and BEOL by 37% and 25%, respectively. At the same time, the security of our approach is similar to that of the previous work [6]. Table I also compares the attack of [4] and that with our improvement (Section III-B) in columns 3-6. The results show that our improvement can reduce connection errors and Hamming distance from 63% to 37% and from 17% to 14%, respectively. The delay and wirelength overhead of both methods are shown in Figure 13. One can see that the overhead from our approach is equally small as that of [6]. The CPU runtime of our method is typically several seconds for each case.

We further study the impact of wire selection for two circuits c7552 and b20. The wires selected to be elevated (Section IV-B) and rerouted (Section IV-C) are sorted according to a weighted combination of potential wire detour distance and benefit to CMP, with wires of small distance and large CMP benefit at the top. Then, top α% of these selected wires are elevated and rerouted. We vary the value of α to see the impact on security, CMP uniformity and overhead. The results are plotted in Figure 10, 11 and 12. In general, increasing the number of elevated and rerouted wires improves security and CMP uniformity, and causes more overhead. There are a couple of non-monotone changes of $\Delta Var(den)$ in Figure 11. These are due to the heuristic nature of our approach.

C. Results of SADP-Compliant Defense

The SADP results are shown in Table II. The previous work of security-only routing [6] increases the number of SADP violations by 44%. In contrast, our SADP-compliant security routing can reduce the violations by 97%. The security of our method in terms of connection error and Hamming distance is about the same or even better than the previous work [6]. The delay and wirelength overhead by our method are increased, but still quite small. The ILP runtime is always within 2 minutes for each circuit.

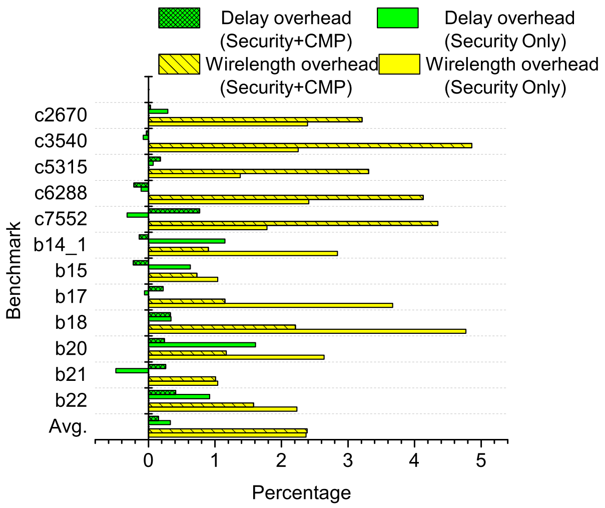

Fig. 13. Rerouting overhead.

TABLE II
RESULTS OF SADP VIOLATIONS AND SADP-COMPLIANT SECURITY ROUTING.

Circuit	Original	Security Only [6]	Security + SADP				
			Improved Attack (Sec. III-B)			Delay	Wirelength
	#SADP violations	#SADP violations	Connection error (%)	Hamming distance (%)	#SADP violations	overhead (%)	overhead (%)
c2670	20	33	93.6	24.4	0	0.05	7.49
c3540	12	28	85.0	36.9	0	0.14	2.41
c5315	39	48	96.3	29.7	3	0.02	3.74
c6288	40	49	65.7	33.3	0	-0.01	0.65
c7552	27	64	93.0	33.4	1	0.08	4.53
b14_1	16	25	98.1	26.8	0	0.53	1.49
b15	42	50	96.2	22.4	1	5.45	3.07
b17	203	385	95.9	29.6	3	1.06	4.82
b18	740	928	97.3	29.7	27	0.55	4.64
b20	70	131	90.5	39.4	3	0.43	4.55
b21	46	62	89.9	40.2	2	0.95	4.89
b22	50	76	85.9	30.2	0	0.74	1.93
Avg.	109	157	90.6	31.3	3	0.83	3.68

VII. CONCLUSIONS

Although the security risk associated with untrusted foundries partially arises from the advanced process technology and related manufacturability challenge, existing works on split fabrication almost always focus on security while neglect manufacturability issues. In this work, we show that manufacturability and security in split fabrication can actually be addressed in a synergistic manner. In particular, two routing based security methods are developed, one is friendly with chemical mechanical planarization, and the other improves compliance with self-aligned double patterning. Comparison with the latest previous work indicates that the proposed methods can achieve the same security with significantly improved manufacturability.

REFERENCES

[1] D. Z. Pan, B. Yu, and J.-R. Gao, "Design for manufacturing with emerging nanolithography," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 32, pp. 1453–1472, Oct. 2013.

[2] J. Rajendran, O. Sinanoglu, and R. Karri, "Is split manufacturing secure?," *Proceeding of the ACM/IEEE Design, Automation Test in Europe Conference*, pp. 1259–1264, 2013.

[3] Intelligence Advanced Research Projects Activity, "Trusted Integrated Circuits Program." https://www.fbo.gov/utils/view?id=b8be3d2c5d5babbdffc6975c370247a6, 2011.

[4] Y. Wang, P. Chen, J. Hu, and J. Rajendran, "The cat and mouse in split manufacturing," *ACM/IEEE Design Automation Conference*, pp. 1–6, 2016.

[5] J. Magaña, D. Shi, and A. Davoodi, "Are proximity attacks a threat to the security of split manufacturing of integrated circuits?," *IEEE/ACM International Conference on Computer-Aided Design*, pp. 90:1–90:7, 2016.

[6] Y. Wang, P. Chen, J. Hu, and J. Rajendran, "Routing perturbation for enhanced security in split manufacturing," *IEEE Asia and South Pacific Design Automation Conference*, pp. 605–510, 2017.

[7] K. Vaidyanathan, P. B. Das, E. Sumbul, R. Liu, and L. Pileggi, "Building trusted ICs using split fabrication," *IEEE International Symposium on Hardware-Oriented Security and Trust*, pp. 1–6, 2014.

[8] K. Vaidyanathan, R. Liu, E. Sumbul, Q. Zhu, F. Franchetti, and L. Pileggi, "Efficient and secure intellectual property (ip) design with split fabrication," *IEEE International Symposium on Hardware-Oriented Security and Trust*, pp. 13–18, 2014.

[9] M. Jagasivamani, P. Gadfort, M. Sika, M. Bajura, and M. Fritze, "Split-fabrication obfuscation: Metrics and techniques," *IEEE International Symposium on Hardware-Oriented Security and Trust*, pp. 7–12, 2014.

[10] J. Valamehr, T. Sherwood, R. Kastner, D. Marangoni-Simonsen, T. Huffmire, C. Irvine, and T. Levin, "A 3-D split manufacturing approach to trustworthy system development," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 32, pp. 611–615, April 2013.

[11] F. Imeson, A. Emtenan, S. Garg, and M. Tripunitara, "Securing computer hardware using 3d integrated circuit (ic) technology and split manufacturing for obfuscation," *USENIX Security Symposium*, pp. 495–510, 2013.

[12] Y. Xie, C. Bao, and A. Srivastava, "Security-aware design flow for 2.5D IC technology," *International Workshop on Trustworthy Embedded Devices*, pp. 31–38, 2015.

[13] M. Cho, D. Z. Pan, H. Xiang, and R. Puri, "Wire density driven global routing for cmp variation and timing," *IEEE/ACM International Conference on Computer Aided Design*, pp. 487–492, 2006.

[14] H.-Y. Chen, S.-J. Chou, S.-L. Wang, and Y.-W. Chang, "A novel wire-density-driven full-chip routing system for CMP variation control," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 28, no. 2, pp. 193–206, 2009.

[15] J. R. Gao and D. Z. Pan, "Flexible self-aligned double patterning aware detailed routing with prescribed layout planning," *ACM International Symposium on Physical Design*, pp. 25–32, 2012.

[16] I.-J. Liu, S.-Y. Fang, and Y.-W. Chang, "Overlay-aware detailed routing for self-aligned double patterning lithography using the cut process," *ACM/IEEE Design Automation Conference*, pp. 1–6, 2014.

[17] Y. Ding, C. Chu, and W.-K. Mak, "Throughput optimization for SADP and e-beam based manufacturing of 1D layout," *ACM/IEEE Design Automation Conference*, pp. 1–6, 2014.

[18] J. Kuang, E. F. Y. Young, and B. Yu, "Incorporating cut redistribution with mask assignment to enable 1D gridded design," *IEEE/ACM International Conference on Computer-Aided Design*, pp. 48:1–48:8, 2016.

[19] R. Tian, D. F. Wong, and R. Boone, "Model-based dummy feature placement for oxide chemical-mechanical polishing manufacturability," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 20, no. 7, pp. 902–910, 2001.

[20] D. Ouma, D. Boning, J. Chung, G. Shin, L. Olsen, and J. Clark, "An integrated characterization and modeling methodology for cmp dielectric planarization," *IEEE International Interconnect Technology Conference*, pp. 67–69, 1998.

[21] N. Dhumane and S. Kundu, "Critical area driven dummy fill insertion to improve manufacturing yield," *IEEE International Symposium on Quality Electronic Design*, pp. 334–341, 2012.