Equivalent Capacitance Guided Dummy Fill Insertion for Timing and Manufacturability

Sheng-Jung Yu[1], Chen-Chien Kao[2], Chia-Han Huang[2] and Iris Hui-Ru Jiang[12]

[1]Department of Electrical Engineering, National Taiwan University, Taipei 10617, Taiwan
[2]Graduate Institute of Electronics Engineering, National Taiwan University, Taipei 10617, Taiwan
{b04901018, b04901026, b04901022, huirujiang}@ntu.edu.tw

ABSTRACT

To improve manufacturability, dummy fill insertion is widely adopted for reducing the thickness variation after chemical mechanical polishing. However, inserted metal fills induce significant coupling to nearby signal nets, thus possibly incurring timing degradation. Existing timing-aware fill insertion strategies focus on optimizing induced coupling capacitance instead of resultant equivalent capacitance. Therefore, the impact on timing cannot be fully captured. In contrast, in this paper, we analyze equivalent capacitance friendly regions for dummy fills. The analysis can wisely guide dummy fill insertion to prevent unwanted and unnecessary increase in the resultant equivalent capacitance of timing critical nets. Experimental results based on the ICCAD 2018 CAD Contest benchmark suite show that our solution outperforms the contest winning teams and state-of-the-art work. Moreover, our analysis results are highly correlated to actual equivalent capacitance values and indeed provide accurate guidance for timing-aware dummy fill insertion.

1 INTRODUCTION

Chemical mechanical polishing (CMP) is a predominant planarizing technique in metallization of semiconductor manufacturing process. Ideally, CMP should satisfy the planarity requirement by smoothing the wafer surface topography with both chemical and mechanical forces [1, 2]. The CMP quality, however, highly relies on the density uniformity of pattern features in a layout. To alleviate metal density variations, dummy fill insertion has been widely adopted [3–8]. Nevertheless, dummy fill insertion adds unwanted coupling to nearby signal nets, thus causing potential timing degradation [9, 10].

To reduce the timing risk, recent timing-aware fill insertion strategies mainly fall into three categories: (1) *Constraint Setting* sets constraints to bound induced coupling capacitance [11–15]. (2) *Overlay Reduction* minimizes the overlay area to reduce the induced coupling capacitance [16–18]. (3) *Capacitance Analysis* minimizes the induced coupling capacitance and the estimated timing impacts [19–21].

Because of the complicated computation of equivalent capacitance, all of these strategies focus on minimizing only induced coupling capacitance. However, propagation delay and slew on a signal net depend on the resultant equivalent capacitance of its capacitance network. The impact of a dummy fill on equivalent capacitance is determined by the fill location in the overall capacitance network. Minimizing induced coupling capacitance does not guarantee resolving the timing risk. Take Figure 1 as an example. Figure 1 (c) shows that if induced coupling capacitance is optimized without considering the fill location in the capacitance network, equivalent capacitance may increase significantly. In contrast, Figure 1 (e) illustrates that minimizing equivalent capacitance does not require minimizing induced coupling capacitance at the same time. This example shows that timing degradation may be severe even when the induced coupling capacitance has been optimized. Therefore, to reduce the timing risk, the key is to control equivalent capacitance. On the other hand, most prior works do not consider all types of intra- and inter-layer induced coupling capacitance during fill optimization. The incomplete considerations may even further degrade the fill quality.
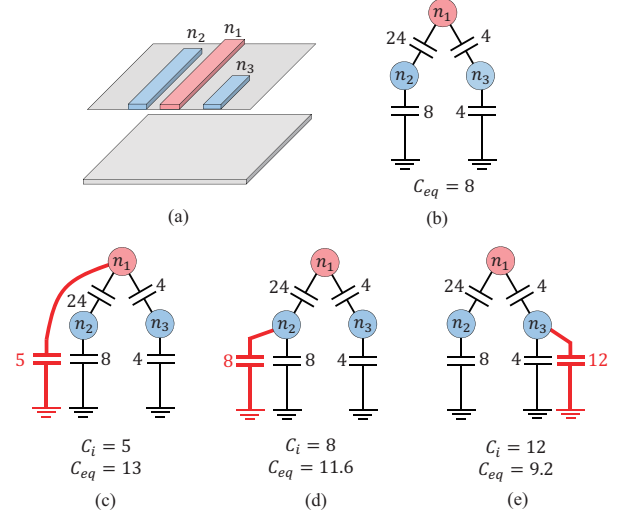
Figure 1: Induced coupling capacitance vs. equivalent capacitance. The red capacitance (highlighted by bold lines) indicates the induced coupling capacitance C_i of a dummy fill. (a) The layout before dummy fill insertion. (b) The capacitance network before dummy fill insertion. (c) The worst resultant equivalent capacitance, while the least induced coupling capacitance. (d) The resultant equivalent capacitance is in between the other two cases. (e) The best resultant equivalent capacitance, while the most induced coupling capacitance.

Motivated by these issues, we propose an equivalent capacitance guided dummy fill insertion framework. Our contributions are summarized as follows:

- We aim at equivalent capacitance instead of induced coupling capacitance. To the best of our knowledge, this is the *first* work that can measure the impact of a fill candidate region on equivalent capacitance.

- We propose a highly accurate equivalent capacitance estimation scheme. Compared with actual values, the error is only 5.5%, and the correlation is 0.997.

- We consider all types of capacitance in our flow. Our results reveal that fringe capacitance is not negligible during optimization. The neglect may cause a degradation of fill quality.

- We guide fill insertion by critical region analysis which determines equivalent capacitance friendly regions for dummy fills. Experimental results based on the ICCAD 2018 CAD Contest benchmark suite [22] show that our approach outperforms all contest winning teams and the state-of-the-art work [21], achieving more than 27% reduction in total equivalent capacitance.

The remainder of this paper is organized as follows: Section 2 introduces induced capacitance and equivalent capacitance. Section 3 formulates the timing-aware fill insertion problem and gives an overview of our flow. Section 4 proposes our critical region analysis algorithm. Section 5 presents dummy fill insertion. Section 6 analyzes and shows experimental results. Finally, Section 7 concludes this work.
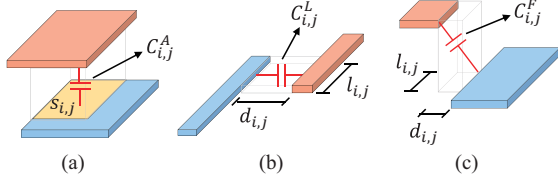
Figure 2: Three major types of induced coupling capacitance. (a) Area capacitance. (b) Lateral capacitance. (c) Fringe capacitance.

2 PRELIMINARIES

This section introduces capacitance evaluation.

2.1 Capacitance in Parallel and in Series

For two capacitors in parallel with capacitance C_1 and C_2, the equivalent capacitance can be computed as follows:

$$C_{par} = C_1 + C_2. \tag{1}$$

For two capacitors in series with capacitance C_1 and C_2, the equivalent capacitance can be computed as follows:

$$C_{ser} = \frac{1}{\frac{1}{C_1} + \frac{1}{C_2}}. \tag{2}$$

2.2 Induced Coupling Capacitance

In this paper, dummy fills and signal nets are composed of rectangular metals [22]. Three types of induced coupling capacitance between metals are generated according to their relative positions (see Figure 2):

- *Area Capacitance*: Two metals p_i and p_j in different layers with overlapped area form area capacitance $C_{i,j}^A$:

$$C_{i,j}^A = P_{m,n}^A(s_{i,j}) \cdot s_{i,j}, \tag{3}$$

  where $s_{i,j}$ denotes the overlapped area, and $P_{m,n}^A$ denotes the area capacitance per unit area when the area is $s_{i,j}$ for metals in the mth layer and the nth layer.

- *Lateral Capacitance*: Two metals p_i and p_j with parallel edge in the same layer form lateral capacitance $C_{i,j}^L$:

$$C_{i,j}^L = P_m^L(d_{i,j}) \cdot l_{i,j}, \tag{4}$$

  where $l_{i,j}$ denotes the parallel edge length of the two metals, $d_{i,j}$ denotes the distance between the two metals, and P_m^L denotes the lateral capacitance per unit length when the distance is $d_{i,j}$ for metals in the mth layer.

- *Fringe Capacitance*: Two metals p_i and p_j in different layers without overlaps and without being shielded by intermediate metals form fringe capacitance $C_{i,j}^F$:

$$C_{i,j}^F = P_{m,n}^F(d_{i,j}) \cdot l_{i,j} + P_{n,m}^F(d_{j,i}) \cdot l_{j,i}, \tag{5}$$

  where $l_{i,j}$ and $l_{j,i}$ denote the horizontal overlap length, $d_{i,j}$ and $d_{j,i}$ denote the horizontal distance between two metals and $P_{m,n}^F$ and $P_{n,m}^F$ denote the fringe capacitance per unit length for metals in the mth layer and the nth layer, which are functions of $d_{i,j}$ and $d_{j,i}$, respectively.

In this work, $P_{m,n}^A$, P_m^L and $P_{m,n}^F$ are specified by capacitance look-up tables, which increase the flexibility for designers. Previous work considered only area and/or lateral capacitance in fill optimization. Later, our results will reveal that fringe capacitance is not negligible for equivalent capacitance optimization.

All induced coupling capacitance determines a capacitance circuit graph $G = (V, E)$. Here V denotes the set of vertices, where each vertex v_i represents a signal net n_i, and E contains weighted edges.
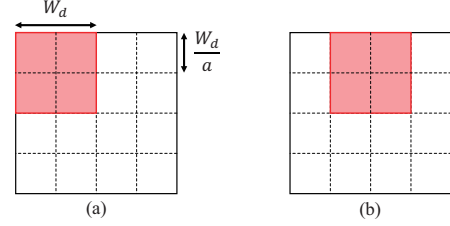

Figure 3: Two examples of $W_d \times W_d$ windows and tiles when $a = 2$.

The weight $C_{i,j}$ of an edge $e_{i,j} = (v_i, v_j)$ is defined as follows:

$$C_{i,j} = \sum_{p_x \in n_i} \sum_{p_y \in n_j} (C_{x,y}^A + C_{x,y}^L + C_{x,y}^F), \tag{6}$$

which is the total induced capacitance between metals of both nets. *Ground net* n_0 is a special net connected to ground (v_0 in G).

2.3 Equivalent Capacitance

The total capacitance of a timing critical net is measured by its equivalent capacitance to ground. The *equivalent capacitance* $C_i^{G,k}$ of signal net n_i with respect to critical net n_k is a capacitance that can replace all induced capacitance connected to n_i without altering the physical effect between n_i and ground (n_0) when only critical net n_k is applied with a voltage source. Based on the definition of capacitance, the physical effect is the ratio of the quantity of electric charge on a signal net to the electric potential. The equivalent capacitance $C_i^{G,i}$ determines the delay/slew of signal net n_i. The larger $C_i^{G,i}$, the longer delay/slew.

3 TIMING-AWARE FILL INSERTION

This section first formulates the timing-aware fill insertion problem and then gives the overview of our timing-aware fill insertion flow.

3.1 Problem Formulation

Timing-aware fill insertion is desired to reduce the timing risk while improving the CMP quality. The timing-aware fill insertion problem addressed in this work can be stated as follows:

PROBLEM 3.1 (**TIMING-AWARE FILL INSERTION**). *Given a routed layout, a set of signal nets with their timing criticalities, capacitance look-up tables, the metal density requirement, and design rules, insert dummy fills into the layout such that the weighted sum of increased equivalent capacitance of signal nets is minimized, and the density requirement and design rules are satisfied.*

Signal nets can be classified into *critical nets* and *non-critical* nets. Critical nets are associated with non-zero (positive) timing criticalities, while non-critical nets are associated with zero ones. The density requirement and design rules are hard constraints:

- *Density requirement* (D_{min}^i, D_{max}^i): The layout is divided into overlapping windows of dimension $W_d \times W_d$. Each window covers $a \times a$ tiles of dimension $(W_d/a) \times (W_d/a)$, as shown in Figure 3. The area density of signal nets and dummy fills in each window in the ith layer must be in the range $[D_{min}^i, D_{max}^i]$.
- *Spacing rule* (S_{min}^i): The minimum spacing between any two metals in the ith layer must be at least S_{min}^i.
- *Sizing rule* (L_{min}^i, L_{max}^i): The width/length of any dummy fill in the ith layer must be in the range $[L_{min}^i, L_{max}^i]$.

3.2 Overview

As shown in Figure 4, our timing-aware fill insertion flow consists of two stages: (1) *Critical Region Analysis* extracts all candidate regions from empty space for dummy fills and evaluates the impact of each
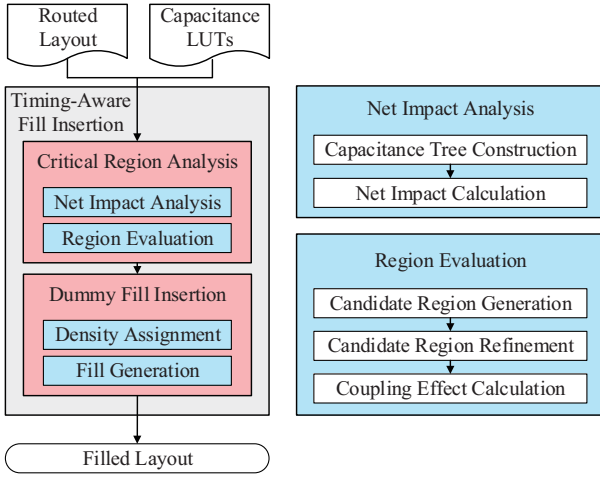
Figure 5: Circuit tree construction. (a) The capacitance circuit graph, where vertex v_i represents net n_i. (b) The corresponding circuit tree, where removed edges are approximated by compensation capacitance.

Algorithm 1 Net Impact Calculation Algorithm

Require: Circuit Tree $G' = (V, E_T)$, the removed edge set E_M, the vertex set of critical nets $V_c \in V$, and the number of iterations T.
Ensure: I the net impact of all nets
1: **for all** vertex v_i in V **do**
2:    $I_i \leftarrow 0$
3: **for all** vertex v_k in V_c **do**
4:    $V_{DFS} = DFS(G', v_k)$
5:    **for** $t = 0, t < T, t = t + 1$ **do**
6:       **for** $i = 0, i < |V_{DFS}|, i = i + 1$ **do**
7:          $capacitanceCompute(G', V_{DFS}[i])$
8:          $V_k^k \leftarrow 1$
9:       **for** $i = |V_{DFS}| - 1, i >= 0, i = i - 1$ **do**
10:          $voltageCompute(G', V_{DFS}[i])$
11:       $G' = compensate(G', E_M)$
12:    **for** $i = 0, i < |V_{DFS}|, i = i + 1$ **do**
13:       $capacitanceCompute(G', V_{DFS}[i])$
14:    **for** $i = |V_{DFS}| - 1, i >= 0, i = i - 1$ **do**
15:       $I_i = I_i + netImpactCompute(G', V_{DFS}[i])$
16:    $G' = reset(G')$
17: **return** I

---

region on equivalent capacitance. (2) According to the guidance of Critical Region Analysis, *Dummy Fill Insertion* assigns a target fill density for each candidate region and generates fills accordingly.

4 CRITICAL REGION ANALYSIS

Minimizing the induced capacitance does not imply minimizing equivalent capacitance. Therefore, measuring the impact of each fill candidate region on equivalent capacitance is crucial for fill insertion. In this section, we propose a critical region analysis algorithm. Critical Region Analysis contains two parts: (1) *Net Impact Analysis* analyzes induced capacitance and estimates equivalent capacitance for each net. (2) *Region Evaluation* creates candidate regions for dummy fills and evaluates each region according to Net Impact Analysis.

4.1 Net Impact Analysis

The resultant equivalent capacitance depends on not only the value of individual induced capacitance but also its location in the capacitance circuit graph. Therefore, we identify the nets inducing coupling capacitance and estimate the equivalent capacitance after fill insertion. We propose the concept of *net impact* to estimate the change on equivalent capacitance. The net impact I_i of net n_i is defined as follows:

$$I_i = \sum_{n_k \in N_c} w_k \cdot I_{i,k}, \quad I_{i,k} = \frac{\partial C_k^{G,k}}{\partial C_i^{G,k}}, \tag{7}$$

where w_k denotes the criticality of the critical net n_k, $C_i^{G,k}$ denotes the equivalent capacitance of n_i with respect to the critical net n_k, and N_c denotes the set of critical nets. The partial derivative $I_{i,k}$ is the linear approximation ratio of the change in $C_k^{G,k}$ to the change of $C_i^{G,k}$, i.e., $I_{i,k}$ reflects the change of the equivalent capacitance of n_k caused by the change of the equivalent capacitance on n_i. *Net Impact Analysis* computes all net impacts.

4.1.1 Circuit Tree Construction. The net impacts depend on the structure of the capacitance circuit graph. However, they cannot easily be calculated in a general graph. To resolve this issue, we propose an approximation scheme by removing insignificant edges to create another tree-structured capacitance circuit graph, as shown in Figure 5. In the *circuit tree*, we replicate vertex v_0 for each edge connected to v_0 (ground net). With the tree structure, all characteristics, including equivalent capacitance and its derivatives, can be calculated efficiently.

Given the routed layout, we first construct the capacitance circuit graph. For efficiency, instead of checking all pairs of metals, we ignore small induced capacitance. First, we split the layout into uniform
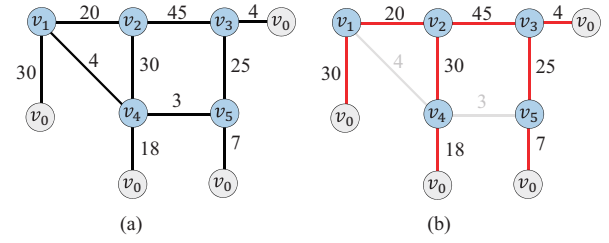
grids. Then, we evaluate the induced capacitance of metal pairs in nearby grids. The grid size W_{grid} is a user-defined parameter, which trades off accuracy for efficiency.

If the capacitance circuit graph is a tree, then we are done. Otherwise, we approximate the capacitance circuit graph by a circuit tree. The circuit tree is constructed by applying the Maximum Spanning Tree (MST) algorithm to the initial circuit graph. The Maximum Spanning Tree (MST) algorithm separates the edges into two sets E_T and E_M, where E_T includes the maximum spanning tree edges and all edges connected to v_0, and E_M includes the remaining (removed) edges. The resultant circuit tree is $G' = (V, E_T)$.

If a net has multiple induced capacitance, the smaller one can be discarded without largely affecting the net. Therefore, the circuit tree is constructed to reflect the characteristics of initial circuit graph with small changes. The edges in E_M are used to further compensate the change, which will be detailed in the next stage.

4.1.2 Net Impact Calculation. With the circuit tree, Algorithm 1 calculates the net impact of each signal net. Lines 1-2 initialize all net impacts to zero. Line 4 determines the calculation order. Lines 6–11 compensate the removed edges to increase the accuracy, the compensation iterates T times. Lines 12–15 calculates the net impacts of all nets on the critical net n_k. Lines 16 resets the compensation from the circuit tree. Lines 3–16 are repeated until all net impacts are calculated. Then, Line 17 outputs the net impacts of all signal nets. The calculation order is decided by the post-order and pre-order traversals of the circuit tree G' by Depth First Search (DFS).

Before calculating the impact, we compensate the difference by adding compensation capacitance on the circuit tree. For every edge $e_{i,j} \in E_M$, we append two compensation capacitance $C_{i,0}^{M,k}, C_{j,0}^{M,k}$ from both vertices to v_0 (ground net):

$$C_{i,0}^{M,k} = \max\left(0, C_{i,j} \cdot \left(1 - \frac{V_j^k}{V_i^k}\right)\right), \tag{8}$$

Figure 4: The overview of our timing aware-fill insertion flow.

$$C_{j,0}^{M,k} = \max\left(0, C_{i,j} \cdot \left(1 - \frac{V_i^k}{V_j^k}\right)\right), \qquad (9)$$

where V_i^k (V_j^k) is the voltage of v_i (v_j) with respect to critical net n_k. Equations (8) and (9) are derived from *Miller Theorem* [23], which approximates a capacitor by capacitors to ground. The approximation is accurate when $C_{i,j}$ is small enough compared to its adjacent edges. To avoid large approximation errors and negative capacitance, we set the lower bound of compensation capacitance as 0. After compensation, each compensation capacitance is connected to a distinct vertex v_0. No edges are created between the vertices from the graph G, which guarantees that *the compensated circuit tree G' remains a tree structure*. The equivalent capacitance and voltage with respect to critical net n_k can be efficiently obtained from G' by exploiting the tree structure as follows:

$$C_i^{G,k} = C_{i,0} + C_{i,0}^{M,k} + \sum_{v_j \in children(v_i)} \frac{C_{i,j} \cdot C_j^{G,k}}{C_{i,j} + C_j^{G,k}}, \qquad (10)$$

$$V_i^k = V_j^k \cdot \frac{C_{i,j}}{C_{i,j} + C_i^{G,k}}, v_j = parent(v_i). \qquad (11)$$

Note that in Equation (11), because the graph is a tree, v_j is unique for each vertex v_i, and thus this equation is well-defined. With these two equations, we can obtain the equivalent capacitance and voltage of each net sequentially following the post-order traversal. This compensation process can be applied repeatedly for higher accuracy.

After compensating the circuit tree, we calculate the net impact with respect to critical nets based on the chain rule of derivatives:

$$I_{i,k} = \frac{\partial C_k^{G,k}}{\partial C_j^{G,k}} \cdot \frac{\partial C_j^{G,k}}{\partial C_i^{G,k}} = I_{j,k} \cdot \left(\frac{C_{i,j}}{C_{i,j} + C_i^{G,k}}\right)^2, \qquad (12)$$

where v_i, v_j are adjacent vertices, and v_j is the parent vertex of v_i. Equation (12) is derived from differentiating the formula of capacitance in series. The net impact with respect to net n_k can be obtained sequentially following the pre-order traversal of G' using Equation (12). After iterating through all critical nets in N_c and accumulating the result for each critical net, we complete our net impact analysis.

THEOREM 1. *Algorithm 1 computes the net impact in $O(k|E| + |V| \lg |V| + k|V|)$ time, where $|E|$ is the size of the edge set in G, $|V|$ is the size of the vertex (net) set in G, and k is the number of critical nets.*

4.2 Region Evaluation

In this stage, we first extract empty space in a layout and generate candidate regions, which are possible regions for dummy fills. Then, we evaluate the impact on equivalent capacitance and densities of each region.

Net Impact Analysis has measured the impact on equivalent capacitance when the induced capacitance on a net changes. Here, we further compute the change on equivalent capacitance when a candidate region is inserted with dummy fills.

Since checking the density requirement and handling the spacing constraints are nontrivial if a candidate region crosses several tiles. For simplicity, we generate each candidate region within a single tile.

4.2.1 Candidate Region Generation. For facilitating capacitance extraction, our idea is to generate candidate regions so that each region has uniform coupling effect (i.e., induced coupling capacitance with the same set of signal nets).

We first split each tile by extending the expanded boundaries along the reverse preferred direction, as shown in Figures 6 (a) and (b). The expanded boundaries are formed by the existing metals with their boundaries expanded outwards by $S_{min}^i/2$. Then, we merge abutting regions if they share the same split lines, as shown in Figure 6 (c).
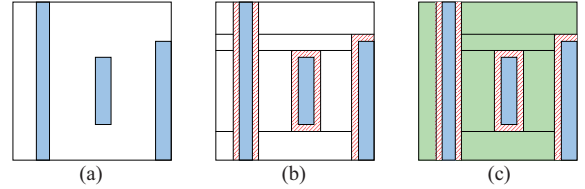
Figure 6: Coarse candidate region generation. (a) A tile with no candidate regions; blue regions are existing metals. (b) The region after splitting; shaded regions are forbidden regions. (c) The resultant candidate regions; green regions are candidate regions.
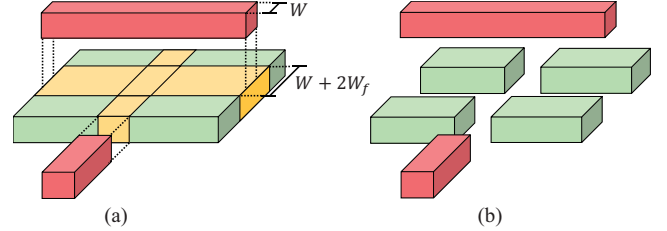
Figure 7: Candidate region refinement. (a) A candidate region may be partially affected by metals not in the preferred direction and those in other layers. Yellow regions indicate the affected region. (b) The result of splitting the candidate region to avoid these metals.

The resultant regions are candidate regions. The merging prevents regions coupling with same signal nets from being separated, thus increasing available area for dummy fills.

Meanwhile, the spacing rule is satisfied. The expended $S_{min}^i/2$ boundaries form forbidden regions, where no legal dummy fills can be located. If dummy fills keep a distance of at least $S_{min}^i/2$ to the region boundaries, any two metals in different candidate regions naturally satisfy the spacing rule, S_{min}^i. In the following steps, we maintain the above distance to meet the spacing constraint.

4.2.2 Candidate Region Refinement. In Section 4.2.1, candidate regions are generated according only to metals in the same layer. However, a candidate region may couple with metals not in the preferred direction or in other layers. Hence, we refine the candidate regions considering metals in different layers and in the non-preferred direction in this step, as shown in Figure 7. We cut candidate regions along the affected regions, as shown in Figure 7 (a). The affected regions are defined by the extension line of metal boundaries, with an adjustable offset W_f to avoid large fringe capacitance from the metals in different layers.

To ensure that the resultant available area for dummy fills is sufficient to satisfy the density requirement, we refine the coarse candidate regions in the descending order of their net impacts. The refinement terminates when the available area in the window cannot satisfy the density requirement. After the refinement, any location within a resultant candidate region has induced capacitance with the same nets, as shown in Figure 7 (b).

4.2.3 Coupling Effect Calculation. With the computed net impact and the generated candidate regions for dummy fills, we can evaluate the coupling effect (on equivalent capacitance) of a candidate region. The coupling effect is defined as follows:

$$E_x = \frac{\sum_{n_i \in S_x} C_{x,i}^c \cdot I_i}{A_x}, \qquad (13)$$

where S_x denotes the set of nets that couple with the candidate region r_x, $C_{x,i}^c$ denotes the capacitance induced by the signal net n_i, I_i denotes the net impact of the signal net n_i, and A_x is the area of the region r_x. The coupling effect of a candidate region represents the average increase of equivalent capacitance resulted from dummy fills.

5 DUMMY FILL INSERTION

After critical region analysis, the relation between a dummy fill location and the equivalent capacitance is found. In this section, we demonstrate how the critical region analysis results guide dummy fill insertion.

5.1 Density Assignment

Following the guide of coupling effects, we determine a target fill density for each candidate region by a linear programming problem:

$$\min \quad \sum_{r_x \in R} E_x \cdot D_x^F \cdot A_x, \tag{14a}$$

$$s.t. \quad D_{\min}^{i,j,k} \leq \frac{\sum_{r_x \in W_{i,j}^k} D_x^F \cdot A_x}{W_d{}^2} \leq D_{\max}^{i,j,k}, \forall i,j,k, \tag{14b}$$

$$0 \leq D_x^F \leq D_{x,\max}^F, \forall r_x, \tag{14c}$$

where R is the set of all candidate regions, D_x^F denotes the target fill density of region r_x, $W_{i,j}^k$ denotes the window in the ith row, j^{th} column, and the kth layer, $D_{\max}^{i,j,k}$ and $D_{\min}^{i,j,k}$ deduct the densities contributed by existing net metals from $D_{\max}^k$ and $D_{\min}^k$, and $D_{x,max}^F$ is the maximum available fill density of region r_x. The objective function is to minimize total coupling effect of dummy fills. The constraints describe that dummy fills should satisfy the density requirement. (The spacing and sizing rules have been handled during candidate region analysis.) After solving this optimization problem, each fill candidate region is assigned a target fill density.

5.2 Fill Generation

After density assignment, we generate dummy fills to fulfill the assigned densities. He *et al.* [24] proposed a guideline which indicates that dummy fills should be located in the middle of a region, and the number of dummy fills between parallel signal nets should be minimized. We follow the guideline to minimize the equivalent capacitance increase.

6 EXPERIMENTAL RESULTS

We implemented our approach in the C++ programming language. Gurobi [25] was selected as the linear programming solver. The evaluation platform was an AMD Ryzen 5 1600 six-core 3.2GHz Linux workstation with 16GB memory. Experiments were conducted on the ICCAD 2018 CAD Contest benchmark suite [22] as listed in Table 1. For fair comparison, the binaries provided by the contest winning teams and the authors of FIT [21] were executed on our platform. The resultant equivalent capacitance was evaluated by the official capacitance evaluation tool released by the contest organizer. Because the objective function of ICCAD 2018 CAD Contest is minimizing the summation of the equivalent capacitance of all critical nets, the criticality of each critical net was set to 1.0 in experiments. To show the completeness, effectiveness, and efficiency of our approach, we performed five experiments.

In the first experiment, we compared our complete timing-aware fill insertion flow with the winning teams of ICCAD 2018 CAD Contest and state-of-the-art work FIT [21]. Table 2 lists the comparison results, where 'no Fills' means the original equivalent capacitance (without dummy fills), 'TF' denotes the total equivalent capacitance (F), 'IF' denotes the increased equivalent capacitance (F), and 'RT' denotes the runtimes (sec). Note that 'IF' is the difference between 'TF' and 'no Fills'. Overall, our approach outperforms all contest winning teams and state-of-the-art work in both the total equivalent capacitance and the increased equivalent capacitance with reasonable runtimes. Compared with the first place team and FIT, we reduce the total equivalent capacitance by 33%, 27%, respectively. Their increased equivalent capacitance is 3.86X, 3.14X as large as

Table 1: ICCAD 2018 CAD Contest benchmark statistics including the number of fill layers (FL), chip size, the number of signal net metals (M), the number of signal nets (N), and the number of critical nets (CN).

	FL	Chip Size (nm^2)	M	N	CN
circuit1	9	600000×540000	305667	30266	105
circuit2	9	1050000×1110000	750166	64674	185
circuit3	9	270000×170000	64903	6683	55
circuit4	9	420000×210000	149464	16484	100
circuit5	9	480000×350000	275425	29726	171
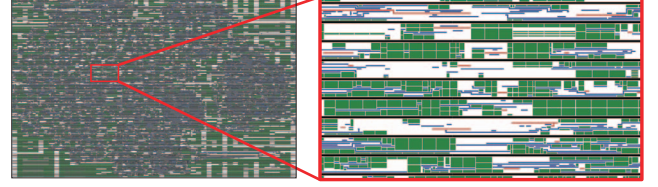
Figure 8: The layout of the first layer of circuit3, where red rectangles are critical nets, blue rectangles are non-critical nets, black rectangles are ground, and green rectangles are dummy fills.

ours, respectively. Figure 8 shows our filled layout of the 1st layer of circuit3. In the layout, our fill insertion flow automatically determines the locations of the dummy fills. The dummy fills are close to ground nets and non-critical nets which have less impacts on the equivalent capacitance. The empty space between ground nets indicates that high impact nets in other layers affect this region. This fact means inserting dummy fills in this region will degrade the timing, and our flow tries to avoid it naturally.

In the second experiment, we compared the estimated equivalent capacitance of our net impact analysis with the actual equivalent capacitance values reported by the official evaluation tool. Table 3 shows our net impact calculation algorithm accurately estimates the equivalent capacitance of critical nets in the original layout with only 5.5% error and 0.997 correlation. It can be observed that our estimated capacitance is slightly larger than the actual value. The over-estimation might be resulted from the exclusion of negative capacitance in Equations (8) and (9). If the Miller Theorem does apply to a capacitance, the negative capacitance would reduce the estimated capacitance. To avoid inaccurate approximation if the theorem does not apply, we remove all negative capacitance. The removal of negative capacitance results in a slight over-estimation. Even so, the results are still acceptable and provide good guidance because of the low error rates and the high correlations. This experiment demonstrates that our estimation approach is effective.

In the third experiment, to show the significance of fringe capacitance, we compared our complete fill insertion flow with the version without fringe capacitance consideration in Table 4. When the fringe capacitance is ignored, the estimated coupling effect is over-optimistic. Thus, the equivalent capacitance (TF and IF) dramatically increases, which confirms the importance of fringe capacitance on timing degradation.

In practice, timing critical designs may contain many critical nets. Critical nets are extracted from the timing analysis report. For these designs, preventing large increase in the equivalent capacitance is of great importance. To simulate such a scenario, in the fourth experiment, we increased the proportion of critical nets to 5% at random. Table 5 shows that our critical region analysis can handle designs with more critical nets. Compared with the first place team and FIT, we effectively reduce the total equivalent capacitance by 27%, 42%, respectively. In addition, FIT may even increase much more equivalent capacitance than the first place team in this scenario.

Finally, in the fifth experiment, we simulated all testcase circuits by SPICE to measure the wire delay increment, which is the main objective of timing-aware fill insertion. As shown in Table 6, compared with the first place team and FIT, we reduce the delay increment

Table 2: Comparisons with winning teams of ICCAD 2018 CAD contest and FIT [21] on the resultant total equivalent capacitance (TF) (F), the increased equivalent capacitance (IF) (F), and runtimes (RT) (sec).

Benchmark		1st place			2nd place			FIT [21]			Ours		
	no Fills	TF	IF	RT	TF	IF	RT	TF	IF	RT	TF	IF	RT
circuit1	1.764E-11	3.310E-11	1.546E-11	16.74	6.106E-11	4.341E-11	3.91	3.094E-11	1.330E-11	5.71	2.156E-11	3.915E-12	15.59
circuit2	4.076E-11	7.942E-11	3.867E-11	52.53	1.460E-10	1.053E-10	14.55	7.353E-11	3.277E-11	16.99	4.984E-11	9.081E-12	40.22
circuit3	7.483E-12	1.294E-11	5.452E-12	3.10	2.334E-11	1.586E-11	0.71	1.180E-11	4.313E-12	0.95	9.052E-12	1.569E-12	3.22
circuit4	1.519E-11	2.546E-11	1.027E-11	6.71	4.517E-11	2.998E-11	1.71	2.325E-11	8.061E-12	2.08	1.782E-11	2.635E-12	7.37
circuit5	2.949E-11	5.089E-11	2.140E-11	13.44	8.953E-11	6.003E-11	3.26	4.597E-11	1.647E-11	3.88	3.524E-11	5.742E-12	14.20
Ratios		1.486	3.862	1.039	2.683	10.925	0.259	1.365	3.137	0.328	1.000	1.000	1.000

by 72%, 65%, respectively. The results demonstrated that controlling equivalent capacitance is crucial to reduce the timing risk.

Our results of these experiments show that our equivalent capacitance guided dummy fill insertion is able to minimize its timing impact.

7 CONCLUSIONS

We have proposed an equivalent capacitance guided dummy fill insertion framework. To the best of our knowledge, this is the first work that can measure the impact of a fill candidate region on equivalent capacitance. With the accurate estimation of equivalent capacitance and full consideration of induced capacitance types, we can score the impact of each fill candidate region on equivalent capacitance. Therefore, our dummy fill insertion flow can prevent unwanted increase in total equivalent capacitance of timing critical nets. Experimental results have shown that our algorithm outperforms contest winning teams and state-of-the-art work in minimizing equivalent capacitance in the filled layout.

Table 3: Estimated total equivalent capacitance vs. actual value for each testcase before dummy fill insertion

	Our Estimation		Actual Value		Correlation	Error (%)
	TF	RT	TF	RT		
circuit1	1.828E-11	6.04	1.764E-11	593.06	0.994	6.20
circuit2	4.219E-11	21.85	4.076E-11	1547.50	0.995	4.99
circuit3	7.842E-12	0.70	7.483E-12	410.54	0.998	5.30
circuit4	1.591E-11	2.92	1.519E-11	257.39	0.999	4.92
circuit5	3.129E-11	8.66	2.949E-11	443.07	0.998	6.15
Average	–	–	–	–	0.997	5.51

Table 4: The impact of fringe capacitance on equivalent capacitance.

	w/o fringe consideration		w/ fringe consideration	
	TF	IF	TF	IF
circuit1	2.719E-11	9.550E-12	2.156E-11	3.915E-12
circuit2	5.366E-11	1.291E-11	4.984E-11	9.081E-12
circuit3	1.136E-11	3.872E-12	9.052E-12	1.569E-12
circuit4	2.195E-11	6.756E-12	1.782E-11	2.635E-12
circuit5	4.438E-11	1.488E-11	3.524E-11	5.742E-12
Ratios	1.217	2.297	1.00	1.000

Table 5: Fill insertion considering more critical nets.

benchmark		1st place	2nd place	FIT [21]	Ours
	no Fills	IF	IF	IF	IF
circuit1	4.426E-11	3.140E-11	8.605E-11	4.507E-11	8.642E-12
circuit2	1.027E-10	7.810E-11	2.127E-10	1.470E-10	2.089E-11
circuit3	1.258E-11	7.466E-12	2.128E-11	1.355E-11	2.577E-12
circuit4	2.609E-11	1.537E-11	4.388E-11	2.809E-11	4.822E-12
circuit5	4.960E-11	3.015E-11	8.640E-11	3.392E-11	1.020E-11
Ratios		3.282	9.194	5.332	1.000

Table 6: Wire delay increment (%) after fill insertion.

	1st place	2nd place	FIT [21]	Ours
circuit1	148.6	402.1	115.0	38.8
circuit2	133.6	358.3	105.4	33.7
circuit3	80.2	248.4	68.1	25.3
circuit4	68.6	214.5	60.1	19.9
circuit5	79.9	240.3	67.7	25.6
Ratios	3.507	10.195	2.890	1.000

REFERENCES

[1] A. B. Kahng and K. Samadi, "CMP fill synthesis: A survey of recent studies," *IEEE TCAD*, vol. 27, no. 1, pp. 3–19, 2008.

[2] L.-T. Wang, Y.-W. Chang, and K.-T. T. Cheng, *Electronic design automation: synthesis, verification, and test.* Morgan Kaufmann, 2009.

[3] C. Liu, P. Tu, P. Wu, H. Tang, Y. Jiang, J. Kuang, and E. F. Young, "An effective chemical mechanical polishing fill insertion approach," *ACM TODAES*, vol. 21, no. 3, p. 54, 2016.

[4] X. Wang, C. C. Chiang, J. Kawa, and Q. Su, "A min-variance iterative method for fast smart dummy feature density assignment in chemical-mechanical polishing," in *Proc. ISQED*, 2005, pp. 258–263.

[5] Y. Chen, A. B. Kahng, G. Robins, and A. Zelikovsky, "Closing the smoothness and uniformity gap in area fill synthesis," in *Proc. ISPD*, 2002, pp. 137–142.

[6] Y. Chen, A. B. Kahng, G. Robins, and A. Zelikovsky, "Area fill synthesis for uniform layout density," *IEEE TCAD*, vol. 21, no. 10, pp. 1132–1147, 2002.

[7] Y. Chen, A. B. Kahng, G. Robins, and A. Zelikovsky, "Monte-carlo algorithms for layout density control," in *Proc. ASPDAC*, 2000, pp. 523–528.

[8] Y. Chen, A. B. Kahng, G. Robins, and A. Zelikovsky, "Practical iterated fill synthesis for CMP uniformity," in *Proc. DAC*, 2000, pp. 671–674.

[9] B. E. Stine, D. S. Boning, J. E. Chung, L. Camilletti, F. Kruppa, E. R. Equi, W. Loh, S. Prasad, M. Muthukrishnan, D. Towery, M. Berman, and A. Kapoor, "The physical and electrical effects of metal-fill patterning practices for oxide chemical-mechanical polishing processes," *IEEE Transactions on Electron Devices*, vol. 45, no. 3, pp. 665–679, 1998.

[10] W.-S. Lee, K.-H. Lee, J.-K. Park, T.-K. Kim, Y.-K. Park, and J.-T. Kong, "Investigation of the capacitance deviation due to metal-fills and the effective interconnect geometry modeling," in *Proc. ISQED*, 2003, pp. 373–376.

[11] S. Sinha, J. Luo, and C. Chiang, "Model based layout pattern dependent metal filling algorithm for improved chip surface uniformity in the copper process," in *Proc. ASPDAC*, 2007, pp. 1–6.

[12] H.-Y. Chen, S.-J. Chou, and Y.-W. Chang, "Density gradient minimization with coupling-constrained dummy fill for CMP control," in *Proc. ISPD*, 2010, pp. 105–111.

[13] H.-J. K. Chiang, C.-Y. Liu, J.-H. R. Jiang, and Y.-W. Chang, "Simultaneous EUV flare variation minimization and CMP control by coupling-aware dummification," *IEEE TCAD*, vol. 35, no. 4, pp. 598–610, 2016.

[14] H. Xiang, L. Deng, R. Puri, K. Chao, and M. D. F. Wong, "Fast dummy-fill density analysis with coupling constraints," *IEEE TCAD*, vol. 27, no. 4, pp. 633–642, 2008.

[15] L. Deng, M. D. F. Wong, K. Chao, and H. Xiang, "Coupling-aware dummy metal insertion for lithography," in *Proc. ASPDAC*, 2007, pp. 13–18.

[16] R. O. Topaloglu, "ICCAD-2014 CAD contest in design for manufacturability flow for advanced semiconductor nodes and benchmark suite," in *Proc. ICCAD*, 2014, pp. 367–368.

[17] Y. Lin, B. Yu, and D. Z. Pan, "High performance dummy fill insertion with coupling and uniformity constraints," *IEEE TCAD*, vol. 36, no. 9, pp. 1532–1544, 2017.

[18] Y. Tao, C. Yan, Y. Lin, S.-G. Wang, D. Z. Pan, and X. Zeng, "A novel unified dummy fill insertion framework with sqp-based optimization method," in *Proc. ICCAD*, 2016, pp. 1–8.

[19] C. Feng, H. Zhou, C. Yan, J. Tao, and X. Zeng, "Efficient approximation algorithms for chemical mechanical polishing dummy fill," *IEEE TCAD*, vol. 30, no. 3, pp. 402–415, 2011.

[20] Y. Chen, P. Gupta, and A. B. Kahng, "Performance-impact limited area fill synthesis," in *Proc. DAC*, 2003, pp. 22–27.

[21] B. Jiang, X. Zhang, R. Chen, G. Chen, P. Tu, W. Li, E. F. Y. Young, and B. Yu, "Fit: Fill insertion considering timing," in *Proc. DAC*, 2019.

[22] B. Yang and S. Sridharan, "ICCAD-2018 CAD contest in timing-aware fill insertion," in *Proc. ICCAD*, 2018.

[23] J. M. Miller, "Dependence of the input impedance of a three-electrode vacuum tube upon the load in the plate circuit," *Scientific Papers of the Bureau of Standards*, vol. 15, no. 351, pp. 367–385, 1920.

[24] L. He, A. B. Kahng, K. H. Tam, and J. Xiong, "Variability-driven considerations in the design of integrated-circuit global interconnects," in *Proc. VMIC*, 2004, pp. 214–221.

[25] "Gurobi Optimization Inc. Gurobi Optimizier," http://www.gurobi.com.