A Unified Framework for Simultaneous Layout Decomposition and Mask Optimization

Yuzhe Ma¹, Jhih-Rong Gao², Jian Kuang², Jin Miao², Bei Yu¹

¹The Chinese University of Hong Kong 2Cadence Design Systems

Outline

[Introduction](#page-2-0)

[Algorithms](#page-13-0)

[Experimental Results](#page-27-0)

[Conclusion](#page-34-0)

Outline

[Introduction](#page-2-0)

[Algorithms](#page-13-0)

[Experimental Results](#page-27-0)

VLSI Chip Design Flow

VLSI Chip Design Flow

Layout Decomposition (LD)

 \triangleright Conflict: two features with the same color, while distance $\lt d_{min}$

Problem Formulation

Input: layout and *dmin* Output: decomposed layout, minimizing conflict #

Mask Optimization (MO)

- \blacktriangleright The quality of printed image may be poor due to the diffraction effect of the light.
- \triangleright Optical Proximity Correction(OPC): Refine the mask to compensate the diffraction effect.
- Method for OPC:
	- rule-based [Park+,ISQED'2010];
	- model-based [Kuang+,DATE'2015][Su+,TCAD'2016];
	- inverse lithography technique [Gao+,DAC'2014].

Mask Optimization (cont.)

- \triangleright Edge Placement Error (EPE): Geometric displacement between the image contour and the edge of target image on the layout.
- \triangleright EPE Violation: The perpendicular displacement is greater than an EPE threshold value.

Problem Formulation

Input: target layout Output: refined mask, minimizing EPE violation #.

Two-Stage Flow for Layout Optimization

Two stages:

- ▶ Layout Decomposition (LD)
- \triangleright Mask Optimization (MO)

Issues

Solution 1: $#EPE$ Violation = 1 Solution 2: $#EPE$ Violation = 3

Options?

- Exhaustive MO for all LD solutions.
	- Running time overhead due to thousands of LD solutions.

Options? (cont.)

- **F** Heuristic selection among LD solutions.
	- Local region density [Yu+,ICCAD'13]: balance the pattern density on each mask.

- Spacing vector [Chen+,ISQED'13]: maximize minimum distance between patterns.

- Limited effectiveness.

Motivation

How about combining LD and MO together?

- \blacktriangleright It is an open problem.
- \blacktriangleright It is expected to be more effective and more efficient.

[Introduction](#page-2-0)

[Algorithms](#page-13-0)

[Experimental Results](#page-27-0)

Preliminaries

- \blacktriangleright Lithography model:
	- The aerial image is formed by a series of convolution operation between mask **M** and lithography kernel **h**.

$$
\mathbf{I} = f_{\text{optical}}(\mathbf{M}) = \sum_{k=1}^{K} w_k \cdot |\mathbf{M} \otimes \mathbf{h}_k|^2
$$

- \blacktriangleright Photo-resist model
	- Set a threshold *Ith* to binarize aerial image.

$$
\mathbf{Z}(x, y) = f_{resist}(\mathbf{I}) = \begin{cases} 1, & \text{if } \mathbf{I}(x, y) \ge I_{th}, \\ 0, & \text{otherwise.} \end{cases}
$$

Problem Formulation

 L^{DMO} : Given a target image \mathbf{Z}_t , find two masks \mathbf{M}_1 and \mathbf{M}_2 which can form printed image with high fidelity.

$$
\min_{\mathbf{M}_1, \mathbf{M}_2} F = \|\mathbf{Z}_t - \mathbf{Z}\|_2^2
$$
\n
$$
\text{s.t.} \quad \mathbf{M}_1(x, y) \in \{0, 1\}, \quad \forall x, y,
$$
\n
$$
\mathbf{M}_2(x, y) \in \{0, 1\}, \quad \forall x, y,
$$
\n
$$
\mathbf{I}_1 = \sum_{k=1}^K w_k \cdot |\mathbf{M}_1 \otimes \mathbf{h}_k|^2,
$$
\n
$$
\mathbf{I}_2 = \sum_{k=1}^K w_k \cdot |\mathbf{M}_2 \otimes \mathbf{h}_k|^2,
$$
\n
$$
\mathbf{Z} = f_{resist}(\mathbf{I}_1) \vee f_{resist}(\mathbf{I}_2).
$$

Overall Flow

Overall Flow

Grid Construction

- \blacktriangleright Extract target pattern.
- \blacktriangleright Add bounding box.
- \blacktriangleright Construct grid.
- \blacktriangleright Merge grid.

Overall Flow

Formulation Relaxation

EXECUTE: Relaxation on binary constraints with *sigmoid* function.

$$
\mathbf{M}_1(x, y) \in \{0, 1\} \to \mathbf{M}_1(x, y) = \mathbf{sig}(\mathbf{P}_1(x, y)) = \frac{1}{1 + \exp[-\theta_M \mathbf{P}_1(x, y)]}
$$

$$
\mathbf{Z}_1(x, y) = f_{resist}(\mathbf{I}_1) \to \mathbf{Z}_1(x, y) = \mathbf{sig}(\mathbf{I}_1(x, y)) = \frac{1}{1 + \exp[-\theta_Z(\mathbf{I}_1(x, y) - I_{th})]}
$$

▶ Relaxation on **Z**.

$$
\mathbb{Z} = f_{\text{resist}}(\mathbb{I}_1) \lor f_{\text{resist}}(\mathbb{I}_2) \to \mathbb{Z}(x, y) = \min\{\mathbb{Z}_1(x, y) + \mathbb{Z}_2(x, y), 1\}
$$

Gradient-Based Optimization

Algorithm 1 Gradient-Based Mask Update

- 1: **function** MaskUpdate(P_1 , P_2)
- 2: Initialize stepsize *t*;
- 3: Compute the relaxed masks M_1, M_2 ;
- 4: Compute **Z** according to current P_1 and P_2 ;
- 5: Compute the gradient $\nabla_{\mathbf{P}_1} F$, $\nabla_{\mathbf{P}_2} F$
- 6: $\mathbf{P}_1 \leftarrow \mathbf{P}_1 t \times \nabla_{\mathbf{P}_1} F$;
- 7: $\mathbf{P}_2 \leftarrow \mathbf{P}_2 t \times \nabla_{\mathbf{P}_2} F$;
- 8: **return** P_1 , P_2 , $\nabla_{P_1} F$, $\nabla_{P_2} F$;
- 9: **end function**

Overall Flow

Violation Graph

$$
w_{ij} = \begin{cases} 1, & \text{if } v_i \text{ and } v_j \text{ have conflict,} \\ \beta, & \text{if } v_i \text{ and } v_j \text{ have large } \text{#EPEV,} \\ 0, & \text{otherwise.} \end{cases}
$$

$$
\mathbf{W} = \begin{bmatrix} 0 & 0 & 0 & 0 & \beta \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \beta & 0 & 0 & 1 & 0 \end{bmatrix}
$$

Semidefinite Programming

 \blacktriangleright Use $\mathbf{x} = [x_1, x_2, \cdots, x_n]^\intercal$ to denote the grid assignment solution.

 \blacktriangleright Max-Cut:

$$
\max_{x_i} \sum_{(i,j)\in E} w_{ij} (1 - x_i x_j)
$$

s.t. $x_i \in \{-1, 1\}, \quad \forall v_i \in V$

Relax to Semidefinite Programming:

$$
\min_{\mathbf{X}} \mathbf{W} \bullet \mathbf{X}
$$

s.t. diag(**X**) = **e**,

$$
\mathbf{X} \succeq \mathbf{0}
$$

Semidefinite Programming (cont.)

▶ Randomized rounding [Goemans+,JACM'1995]

- Obtain **X** ∗ by solving SDP.
- Cholesky decomposition with **X** ∗ .

 $X^* = U^{\mathsf{T}}U$

- Get x_i as follows. \mathbf{u}_i is the *i*-th column of **U** and **r** is random unit vector.

$$
x_i = \text{sgn}(\mathbf{u}_i^{\mathsf{T}} \mathbf{r}) = \begin{cases} 1, & \text{if } \mathbf{u}_i^{\mathsf{T}} \mathbf{r} \ge 0, \\ -1, & \text{otherwise.} \end{cases}
$$

Pruning

- \triangleright Obtain multiple solutions by randomized rounding.
- \blacktriangleright Efficient pruning.

[Introduction](#page-2-0)

[Algorithms](#page-13-0)

[Experimental Results](#page-27-0)

#EPE Violation Convergence Curve

Comparison – EPE Violation Num

Comparison – Runtime

Distribution of #EPE violations

Examples of Printed Image

(a) [ICCAD'13] + [DAC'14]; (b) [ISQED'13] + [DAC'14]; (c) Ours.

Examples of Printed Image

```
(a) [ICCAD'13] + [DAC'14];
(b) [ISQED'13] + [DAC'14];
(c) Ours.
```


Outline

[Introduction](#page-2-0)

[Algorithms](#page-13-0)

[Experimental Results](#page-27-0)

[Conclusion](#page-34-0)

Conclusion

- \triangleright A unified framework is proposed for solving LDMO problem.
- \blacktriangleright Two collaborative flows are designed:
	- \blacktriangleright A gradient-based numerical optimization
	- \blacktriangleright A set of discrete optimization.
- \blacktriangleright Effectiveness and efficiency are verified.

Conclusion

- \triangleright A unified framework is proposed for solving LDMO problem.
- \blacktriangleright Two collaborative flows are designed:
	- \triangleright A gradient-based numerical optimization
	- \blacktriangleright A set of discrete optimization.
- \blacktriangleright Effectiveness and efficiency are verified.

Future Exploration

- More advanced lithography process, e.g., triple patterning lithography.
- More optimization targets, such as process variation band.

Thank You

Yuzhe Ma (yzma@cse.cuhk.edu.hk) Jhih-Rong Gao (jgao@cadence.com) Jian Kuang (jkuang@cadence.com) Jin Miao (jmiao@cadence.com) Bei Yu (byu@cse.cuhk.edu.hk)

