Enhanced Hotspot Detection Through Synthetic Hotspot Generation and Design of Experiments

Gaurav Rajavendra Reddy, Constantinos Xanthopoulos and Yiorgos Makris
gaurav.reddy@utdallas.edu, constantinos.xanthopoulos@utdallas.edu, yiorgos.makris@utdallas.edu
Department of Electrical Engineering, University of Texas at Dallas

Continuous technology scaling and the introduction of advanced technology nodes in Integrated Circuit (IC) fabrication has exposed new manufacturability issues. Lithographic hotspots are one of such problems which are a result of complex process interactions. These hotspots are known to vary from design to design and foundries expect such hotspots to be predicted early and corrected in the design stage itself, as compared to a process fix for every hotspot, which would be intractable. Various efforts have been made in the past to address this issue by using a known database of hotspots as a source of information. Most of these works use either Machine Learning (ML) or pattern matching techniques to identify and predict hotspots in new incoming designs. Most of these methods suffer from high false alarm rates and some of the main reasons for this are: these methods are oblivious to the root cause of the hotspots and there is a lack of availability of a large hotspot database to learn from. In this work, we try to address these issues by using novel hotspot Design Of Experiments (DOE) and synthetic hotspot generation approaches. We analyze the effectiveness of proposed methods against the state-of-the-art on a 45nm process, using industry standard tools and designs.

I. Introduction

Continued technology scaling and the introduction of every advanced technology node in Integrated Circuit (IC) fabrication brings in new challenges for foundries. Lithography is one such major challenge during technology development. As shown in figure 1, in early technology nodes, the wavelength of light used in lithography was much smaller than the features being printed. It is vice-versa in the latest nodes. While in the above wavelength region, patterning was relatively easier as shown in figure 2a. In the sub-wavelength region, patterning is extremely challenging due to complex light interactions, depicted in figure 2b. To mitigate some of the patterning related issues and ensure reliable manufacturing, various Resolution Enhancement Techniques (RETs) like Optical Proximity Correction (OPC), Multi-patterning, Phase-shifting masks etc., are used. In spite of employing RETs, complex designs from different designers give rise to various issues during fabrication which are often not foreseen by the foundry. One such issue is lithographic hotspots. Certain areas in the design which show abnormal and unexplained variation despite passing Design Rule Check (DRC) and complying Design For Manufacturability Guidelines (DFMGs), are termed as hotspots. The cause of these hotspots is mostly attributed to the neighborhood of design patterns which causes complex interaction of light during the lithography process.

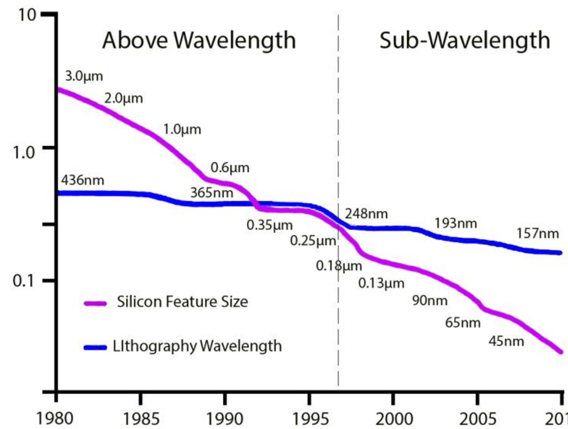

Figure 1: Comparison of light wavelengths with silicon feature sizes [11]

Since these hotspots vary from design to design, the process of identifying the root cause of such abnormal effects of light and finding a fix for all such hotspots through changes in the process is extremely difficult, time consuming and expensive. Thus, in most cases, foundries create a database of such known hotspots and restrict their presence in incoming customer designs. Foundries usually populate such a database through either Failure Analysis (FA) data, Inline inspections, through Lithographic simulations using well calibrated lithographic models [1] etc. If a design pattern turns out to be a hotspot in a later stage of fabrication, especially, if after the mask set has been fabricated, it may result in a huge loss to the foundry. Hence, there is a great need to identify these problematic patterns (hotspots) early, and correct them in the design stages itself.

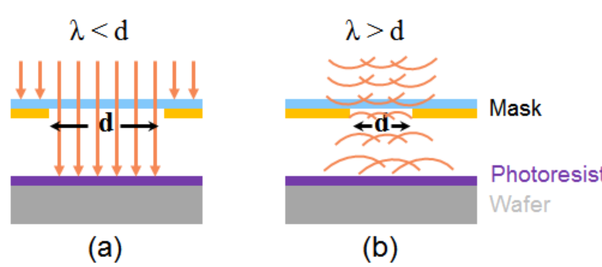

Figure 2: (a) Pattering in above wavelength region (b) Sub-wavelength region [2]

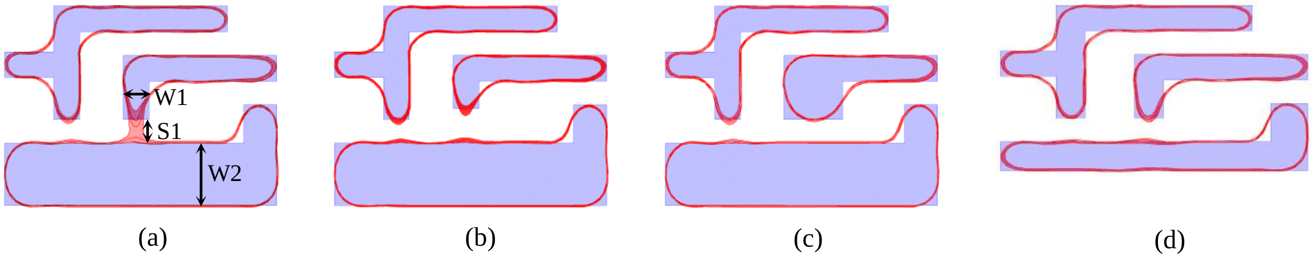

Figure 3: (a) A hotspot pattern (b-d) variations of pattern (a) which are non-hotspots

Many researchers have suggested various techniques to identify and predict hotspots in new incoming designs. Authors of [2] use pattern matching techniques, where, a new design is compared to a database of previously seen hotspots and any potential hotspots are flagged. While these techniques do a good job in identifying known hotspots, they often fail in predicting unknown hotspots. To address this issue, machine learning techniques using Support Vector Machines [3], Artificial Neural Networks [4], multiple classifiers [5] etc. were proposed which essentially 'learn' from a database of known hotspots and classify new patterns as either a hotspot or a non-hotspot.

Over time, various flavors of these techniques have been proposed which have provided better accuracy, lesser false-alarms, faster runtimes etc., [6][7]. However, most of techniques suffer from two main issues which cause increased false-alarms: (a) the ML models only try to learn the difference between hotspots and non-hotspots and do not try to learn more about the actual feature(s) of hotspots which really makes them hotspots. (b) Lack of availability of a large database of hotspots which limits the learning capabilities. Often the existing hotspots are replicated to match the non-hotspot sample numbers [2] [3].

We propose a novel approach to address these problems through synthetic hotspot generation and design of experiments, which enables the Machine Learning entity to train effectively by learning more about the actual geometries causing the hotspot.

The rest of the paper is organized as the following: The problem (source of false alarms) and proposed methods are explained in detail in section II. The complete flow incorporating these techniques is presented in section III. Experimental results are discussed in section IV and section V concludes the paper.

II. Problem Definition and Proposed Methodology

A. *Source of false-alarms*

The state-of-the-art machine learning based techniques for hotspot detection suffer from high false-alarm rates. The source of these false alarms can be understood using figure 3. Figure 3a shows a hotspot pattern, where, the contours (PV bands) obtained from litho simulations indicate a short between two different polygons. Figures 3b-d show patterns which are slightly different from the pattern shown in figure 3a.

Scenario 1 - Let us assume that, a machine learning based classifier was trained to detect hotspots and among the patterns shown in figure 3, only pattern (a) was part of its training dataset. During testing, if pattern (b) is presented to the classifier, it tends to classify it as a hotspot due of its close similarity to pattern (a). But, the litho simulation of pattern (b) shows that it is not a hotspot, thanks to the relaxed spacing 'S1'. The classifier fails to recognize the importance of this feature.

Scenario 2 - Let us assume that the classifier's training dataset included both the patterns 3a and 3b. In that case, the classifier would easily understand that the constrained space 'S1' for such a pattern would make it a hotspot and a relaxed space would make it a non-hotspot. Then, if pattern (c), which is very similar to patterns (a & b), while having a constrained space 'S1', is presented to the classifier, the classifier tends to classify it as a hotspot. But, the litho simulation of pattern (c) shows that it is not a hotspot, thanks to the relaxed width 'W1'. The cause of this error is the fact that, during training, the classifier had only recognized 'S1' as an important feature, but not 'W1'. Similarly, the feature 'W2' is also responsible in turning a pattern into a hotspot or a non-hotspot.

From the above scenarios, we can clearly see that, unless otherwise trained with many variants of a known hotspot, the ML entity assumes that all shapes and their positions in a hotspot pattern equally contribute towards making it a hotspot. It remains oblivious to the features of the pattern which increase or decrease the variation (drive the pattern more towards being a hotspot or a non-hotspot). Lacking this information, the trained model might classify some of the testing patterns shown in figure 3b-d as hotspots, resulting in false-alarms.

To address this issue we propose to enrich the hotspot database through design of experiments and synthetic hotspot generation.

B. *Synthetic hotspot generation and DOEs*

A known hotspot is taken from the database and multiple variations of the same hotspot are created by changing one or more features at a time. Figure 6(a) shows one such hotspot and figures 6(b-h) show some of its variations. Various features which could possibly create variation in a pattern, such as (a) corner to corner distances (b) jogs (c) line end positions (d) metal spacing (e) metal area etc., are varied to produce multiple variations of a known hotspot. A time efficient

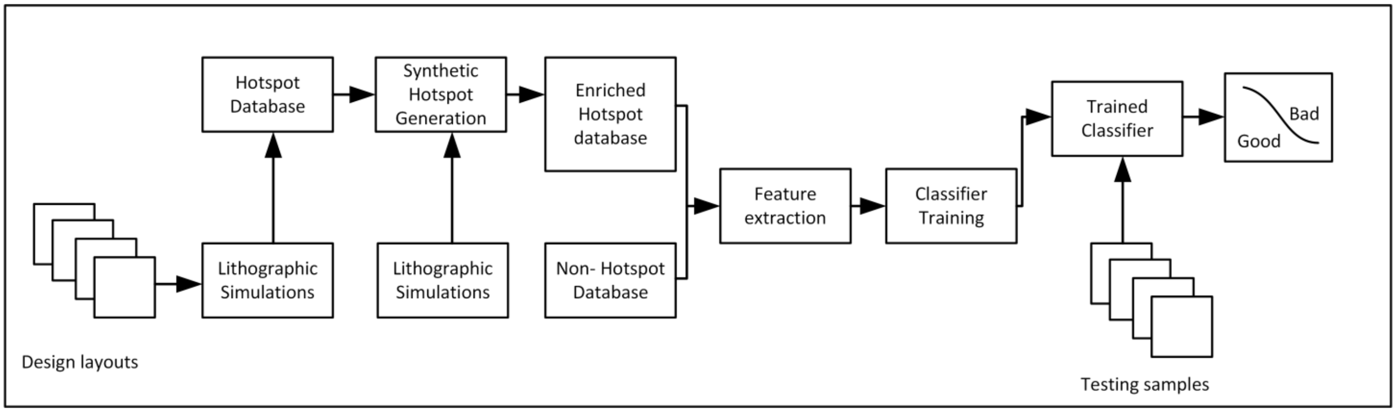

Figure 4: Complete flow of machine learning based hotspot detection using synthetic hotspot generation

method for varying these features relies on perpendicularly moving the edges of a given polygon in each snippet by a random distance. This approach allows the fast generation of multiple variants whose variance can be easily controlled by two parameters. The first parameter is the probability of any given edge to move or remain stationary. By increasing this probability, we effectively increase the number of polygons and their edges in the input snippets that will be altered. The second parameter is a distribution of distances which is sampled for every polygon edge selected by the first parameter and the sampled value denotes the distance by which the edge will be displaced. A simple change to this distribution of distances changes the amount of variation between generated patterns.

As we can expect, the above mentioned procedure results in a plethora of variants, many of which might not even pass the design rule check. To ensure that these are valid layout topologies that can provide meaningful information for training a classifier, a minimal Design Rule Check (DRC) engine is implemented and executed after each variant generation. Design rule clean, synthetic hotspot variants are subjected to lithographic simulations. Run-time of these simulations, on small layout snippets is negligible and is a one-time procedure. Hotspot variants along with their litho simulation results are added into the hotspot database to obtain an enhanced

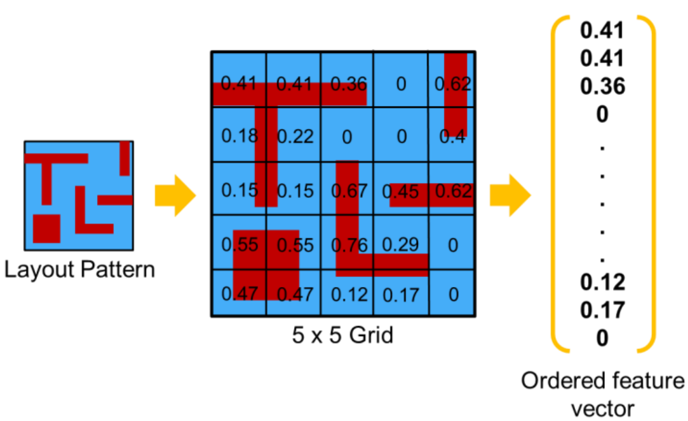

Figure 5: Density based feature extraction [2]

This research has been partially supported by the Semiconductor Research Corporation (SRC) Task 2709.001.

database.

C. *Feature extraction*

In all proposed machine learning based hotspot detection schemes, hotspot and non-hotspot patterns are stored in the form of layout snippets which are subjected to feature extraction, where, the image snippet is transformed into a feature vector which can be used as a training/testing vector. Various feature extraction methods like density based [2], bounded rectangle region based [4], fragment based [3], etc. have been proposed. We have used density based feature extraction for our analysis. In this method, as shown in figure 5, an n*n grid is overlapped on a pattern and the density of the metal within each block of the grid is computed. The ordered vector of such densities is used as the feature vector. While density captures information about the presence of a material, the vector ordering captures the location information. The size of the grid is a parameter decided through experimentation. A very fine grid results in a large number of features which leads to over-fitting and vice-versa.

III. The Overall Flow

The complete flow is shown in figure 4. The database of hotspot patterns is initially populated through full chip lithographic simulations on a wide variety of layouts [1]. Considering that this is a one-time effort, the more the merrier. This step is not necessary if the foundry already has an initial hotspot database from prior learning. The hotspot database is enriched as described in section II(b). Hotspots from the enriched hotspot database and some non-hotspots sampled from placed & routed layouts are used as the training dataset. The training patterns are subjected to feature extraction and the resultant feature vectors are used to train a two-class classifier.

Subsequently, whenever the foundry receives a new customer design/Intellectual Property (IP) for fabrication, the entire layout is scanned layer by layer and each layer is converted into image snippets of a pre-determined size. All such snippets are converted into feature vectors and are then tested for hotspots. For those patterns/feature-vectors flagged as hotspots, their corresponding locations in the layout are inspected and design fixes are requested if necessary. Hence,

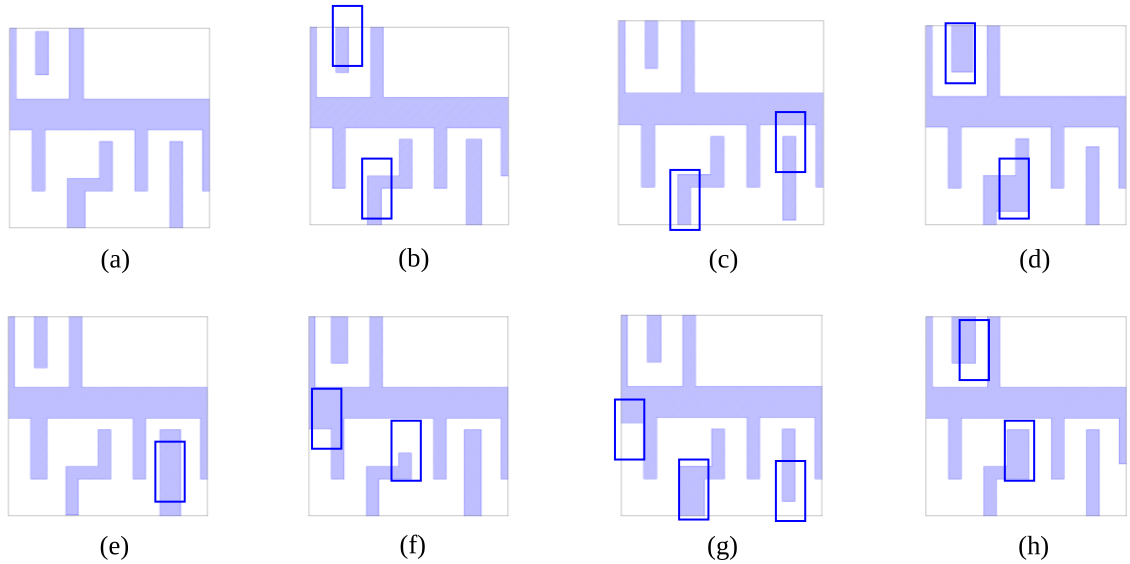

Figure 6: (a) Hotspot pattern obtained from a benchmark layout, (b-h) Synthetic patterns generated using pattern 'a' as the initial pattern. Blue markers indicate the subtle changes in them, in comparison to pattern 'a'

this methodology enables effective hotspot detection before fabrication.

IV. Results

A set of Register-Transfer-Level (RTL) benchmarks are placed and routed using the Nangate open cell library [8], which was developed based on a 45nm Product Design Kit (PDK) [9]. Patterns are captured from layouts using Calibre Pattern Match tool and Lithographic simulations are performed using Calibre Litho-Friendly-Design (LFD) tool-kit [10], using the litho models provided with the PDK. Lithographic simulations are necessary to find the ground truth about the captured patterns. Half of the dataset obtained from litho simulations is used for training and the other half is used for testing. A Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel is trained and used for classification. The testing dataset is presented to the trained classifier. Initial experiments show a reduction in False-alarms by about 50% in comparison to generic ML based hotspot detection methods. Large scale experiments are underway.

V. Conclusion

We have discussed the problem of lithographic hotspots in advanced technology nodes, analyzed the state-of-the-art in this domain and highlighted that they suffer from high false alarm rates. We have shown that this is partly due to these methods being oblivious to the root cause of the hotspots and the lack of availability of a large hotspot database to learn from. We have tried to address these issues by using novel hotspot Design Of Experiments (DOE) and synthetic hotspot generation approaches. Our initial evaluation of the effectiveness of the proposed methods using industry standard tools and designs on a 45nm process has shown promising results.

References

[1] V. Dai, L. Capodieci, J. yang, N. Rodriguez, "Developing DRC plus rules through 2D pattern extraction and clustering techniques" in Proceedings of SPIE, 2009

[2] W. Y. Wen, J. C. Li, S. Y. Lin, J. Y. Chen and S. C. Chang, "A Fuzzy-Matching Model With Grid Reduction for Lithography Hotspot Detection," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 33, no. 11, pp. 1671-1680, Nov. 2014.

[3] D. Ding, J. A. Torres and D. Z. Pan, "High Performance Lithography Hotspot Detection With Successively Refined Pattern Identifications and Machine Learning," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 30, no. 11, pp. 1621-1634, Nov. 2011.

[4] Duo Ding, Xiang Wu, J. Ghosh and D. Z. Pan, "Machine learning based lithographic hotspot detection with critical-feature extraction and classification," 2009 IEEE International Conference on IC Design and Technology, Austin, TX, 2009, pp. 219-222.

[5] Duo Ding, Bei Yu, J. Ghosh and D. Z. Pan, "EPIC: Efficient prediction of IC manufacturing hotspots with a unified meta-classification formulation," 17th Asia and South Pacific Design Automation Conference, Sydney, NSW, 2012, pp. 263-270.

[6] K. Madkour, S. Mohamed, D. Tantawy and M. Anis, "Hotspot detection using machine learning," 2016 17th International Symposium on Quality Electronic Design (ISQED), Santa Clara, CA, 2016, pp. 405-409.

[7] H. Zhang, Bei Yu and E. F. Y. Young, "Enabling online learning in lithography hotspot detection with information-theoretic feature optimization," 2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD), Austin, TX, 2016, pp. 1-8.

[8] http://www.nangate.com/?page_id=22

[9] https://www.eda.ncsu.edu/wiki/FreePDK45:Contents

[10] https://www.mentor.com/products/ic_nanometer_design/design-for-manufacturing/calibre-lfd/

[11] http://www.monolithic3d.com/blog/guest-contribution-entanglement-squared-by-zvi-or-bac