Fast source independent estimation of lithographic difficulty supporting large scale source optimization*

David DeMaris*[a], Maria Gabrani[b], Sankha Subhra Sarkar+[b], Nathalie Casati[b], Ronald Luijten[b], Kafai Lai[c], Kehan Tian[c]

[a]IBM Systems & Technology Group, Austin, TX
[b]IBM Zurich Research Laboratory, Rueschlikon, Switzerland
[c]IBM Semiconductor Research and Development , Hopewell Junction, NY

**This is an uncorrected preprint: visit SPIE.org* Proc. of SPIE Vol. 8326 *for final version*

ABSTRACT

Many chip design and manufacturing applications including design rules development, optical proximity correction tuning, and source optimization can benefit from rapid estimation of relative difficulty or printability. Simultaneous source optimization of thousands of clips has been demonstrated recently, but presents performance challenges. We describe a fast, source independent method to identify patterns, which are likely to dominate the solution. In the context of source optimization the estimator may be used as a filter after clustering, or to influence the selection of representative cluster elements. A weighted heuristic formula identifies spectral signatures of several factors contributing to difficulty. Validation methods are described showing improved process window and reduced error counts on 22 nm layout compared with programmable illuminator sources derived from hand picked patterns, when the formula is used to influence training clip selection in source optimization. We also show good correlation with fail prediction on a source produced with hand picked training clips with some level of optical proximity correction tuning.

Keywords: printability analysis, source mask optimization, lithographic difficulty estimation

1. INTRODUCTION

As on-wafer features continue to shrink with 193 nm illumination staying constant, a wide variability is introduced in the imaging quality of patterns at a scale of a few pitches. Many tasks in technology can be managed by focusing processor time and human attention on the most difficult to print patterns as resolution enhancement techniques evolve through the technology learning cycle. Source mask optimization (SMO) technology allows improved process window and variability for small numbers of critical clips and canonical layout configurations, and is emerging as a tool for technology development [1]. Large scale SMO has been demonstrated recently [2], but processing thousands of patterns still presents performance challenges. Since lithographic performance is dominated by the weakest features within a few patterns [3], typically only a few patterns of the represented layout determine the source; adding additional patterns to the optimization does not change the result, so we would like to predict the subset of patterns likely to determine the source. We call this prediction problem lithographic difficulty estimation (henceforth LDE). Given an estimation of the difficulty of printing a pattern, we can rank patterns and use the rank order as a filter choosing only the most difficulty patterns for training the source. Alternatively, we may select patterns which span as much of the design space as possible, and choose the representative of a set of similar patterns by choosing the highest ranked. A more typical practice in such unsupervised classification or clustering process is to choose the most typical (centroid in the representation space). For low compression ratios (many clusters), the choice of representative may not be critical as all patterns are very similar. For higher compression ratios, the need to choose the more difficult patterns in larger sets is more critical. A fast polygon-based FFT produces a complex-valued spectral representation (diffraction order weights) serving as an input to several functions corresponding to different aspects of imaging difficulty: Near Rayleigh-limit spatial frequency, iso-dense patterns, frequency and phase diversity, and 2D vs. 1D. Each term contributes to an overall weighted formulation of estimated difficulty.

*demaris@us.ibm.com +now at Intel Corp, Hillsboro, OR USA. Color copies of the figures are available from the author upon request.

The organization of this paper is as follows: we first describe the current measures of printing difficulty and how pattern features contribute to poor printability across process variations. We then describe the spatial frequency representation of patterns and how we use it to generate the general form of the LDE formulation. We present some indirect results showing how the formulation improves the common process window of large scale optimizations, when the optimization is performed on an algorithmic selected subset for faster turnaround time. Finally we examine some direct correlations of the LDE measure with simulation hotspots. The hotspots (fail markers) were computed given a previously determined source, litho recipe, and hotspots determined by very stringent line and spacing checks to observe patterns which are relatively weak through defocus.

2. ASPECTS OF LITHOGRAPHIC DIFFICULTY

Overview of patterning difficulty

We consider patterns to be a localized layout region on the order of optical radius (5 λ) or smaller. These may be extracted from actual layouts of circuits, routing, or other structures or synthesized algorithmically.

Patterns are considered printable if their critical dimension (CD) measurements conform to expectation across the range of dose and focus conditions that manufacturing considers feasible. Until the source is known and models are calibrated by scanning electron microscope (SEM) measurements, printability of individual patterns is difficult to assess exactly. The source is generally determined by imaging quality (or direct source optimization) for hand-selected critical patterns such as static RAM cells, allowed pitch gratings, and a range of two dimensional calibration patterns.

Patterns which are printable but still considered relatively difficult have high variability in line width, or have a relatively small process window in which critical dimension is acceptable.

Patterns may also fail the printability test after source determination and should be prohibited by design rule checks or other direct matching methods. Difficulty in these cases could be CD violations, failure to achieve adequate image performance in image parameters (exposure quality measures) such as maximum intensity, minimum intensity, image log slope.

Traditionally, difficult patterns have been identified by rigorous and time consuming testing of candidate clips and layouts performed by engineers trained in the art. Optical Rule Checks (ORC), image quality parameters, process variability bands, (pv-band) and failure analysis are among the tools used. The time needed for such a task depends on the experience of the engineer and the tools performance but mainly from the size of the candidate clips list. With the technology shrinking in dimensions and patterns becoming more complex, reducing this list already poses a significant challenge.

Previous work in pattern printability assessment has been performed in the context of a particular optical system and aperture (source) for which the modulation transfer function (MTF) can be computed, and is primarily oriented at analyzing printability at nominal dose and focus. In Yenikaya and Sezginer [4], requirements for intensity in the sampled intensity plane are derived from the polygons and expressed as a set of inequalities which can be checked for feasibility against the imaging system, eliminating some patterns for which no feasible mask can be generated.

A more limited case of direct comparison of thresholded MTF and non-zero diffraction order coefficients is shown to be effective in examining feasibility of process shrinks [5].

Since the objectives of the present work are to assist in the determination of the aperture and optimize the MTF, we make no assumptions about the nature of the mask and MTF, except that the frequencies of interest are bounded by the NA during the sampling process turning polygons into diffraction order vectors. We also consider defocus related spectral signatures in order that the highest process variability (PV-band) conditions can be observed and managed in the optimization process.

Representation of layout

In the evaluation of layout difficulty we rely on a frequency domain description of layouts. A method of fast FFT evaluation from polygons avoids conversion to image domain. A vector of sample points is determined based on the window size, partial coherence measure σ and NA parameters, and the maximum frequency to be observed. We have used layouts processed by retargeting, so features less than the minimum pitch are present, and typically we use features with half pitch spanning window borders; thus oversampling beyond the minimum pitch of a particular layer is required.

Rayleigh Limit

The minimum resolvable pitch is determined by the wavelength λ, numerical aperture NA, and resist effects captured in the k_1 factor.

$$\text{Minimum Feature} = k_1 \frac{\lambda}{NA}$$

Contrast is degraded substantially as the limit is approached, with the flattened intensity profiles resulting in linewidth error [6].

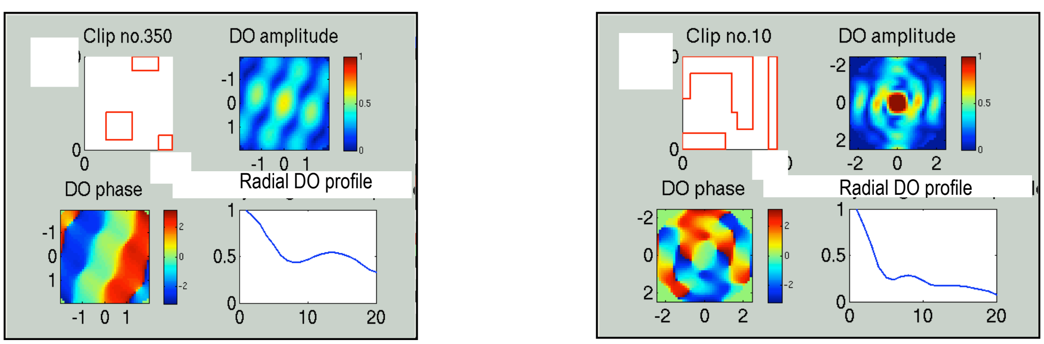
Through focus printing: Iso-dense patterns

Dense and isolated lines are subject to edge error in opposite directions, captured by the lithography adage "dense lines smile, isolated lines frown" [7,8]. While isolated lines are typically corrected by sub-resolution assist features (SRAF), in intermediate configurations containing some semi-dense or iso-dense frequencies, the positioning of SRAFs is complex and may be impossible to obtain good process window. Similarly it may be difficult to find optimal parameters of OPC for such configurations giving a stable result.

We therefore identify such configurations by matching against a spectral template based on the fact that the relative ratio of low frequency component to high frequency component decreases as pitch, hence density, increases.

In an idealized case, we could measure the correlation of the FFT of the pattern, assuming a blank context round the pattern window, with a Mexican Hat function representing the spectral signature of an iso-dense pattern. For rapid computation we divide the weights into three frequency bands and compute a ratio measure of low and high frequency component to the center band. The band transition points are adjusted depending on the window size of the patterns.

The following figure illustrates the FFT spectrum arrays used to compute all the measures, plotting the amplitude and phase arrays and illustrating the diffraction order of an iso-dense array (Fig. 1a) vs. an easier dense pattern (Fig 1b).



a) b)

Figure 1. FFT (also termed diffraction order or DO) magnitude and phase plots for two synthetic layout patterns. The leftmost 200x200 nm pattern from a contact layer would be ranked more difficult due to the presence of low and high frequencies, matching the Mexican Hat template. The rightmost 240x240 nm pattern from a metal interconnect layer has higher entropy measure. The two patterns can not be directly ranked by virtue of different window sizes.

Two Dimensionality

At lowest k factor, only a single direction can be imaged well, necessitating decomposition and two exposures for two dimensional patterns. For less aggressive interconnect layers two dimensional patterning is possible but more difficult. Programmable sources with isolated pixels give best imaging, but create challenges for convergence of optical proximity correction (OPC) relative to one dimensional sources. We reflect this by a term comparing the iso-dense correlation measures in each direction. If the measure is equal in the two dimensions, the difficulty will be considered twice that of a pattern which is difficult in only one dimension.

Frequency and Phase Entropy

Simple source configurations can be optimized for the minimum pitch, with appropriate symmetry selected based on single or double exposures. For more general two dimensional imaging the optimal source problem becomes more complex and the existence of forbidden pitches will result in complex image quality measures across focus. Thus, probabilistically speaking, the greater the number of strong frequency components in the pattern distribution, the greater risk of optimization failing to solve some problems.

Since the pattern spectra are generally sparse, rather than a bandwidth description we take the Shannon entropy measure, treating the diffraction order vector as a random variable with frequency component weights $\{f_1,..f_n\}$:

$$H(DO) = \sum_{1}^{n} e(f_i)\log\frac{1}{e(f_i)} \tag{1}$$

In addition we take an entropy measure of the phase in the upper frequency band. The entropy in the case of phase is taken over a histogram of unwrapped phase values rather than directly from the phase values of the vector.

Summary and Combined Formulation

Table 2: Summary of contributing printability factors. E_{hf}, E_{lf} and E_{cf} are RMS energy in the high, low and center frequency bands, while E_{tot} is the total RMS energy. Threshold is an empirical parameter. Band transitions are adjusted based on window size. H is the Shannon entropy.

Summary of Difficulty Contributions	
Printability Factor	*Description*
Minimum pitch (MP)	$E_{hf} > threshold * E_{tot}$
Iso-Dense Correlation (IDC)	$E_{hf}/E_{cf} + E_{lf}/E_{cf}$
Two Dimensionality (2D)	$\min(IDC_h, IDC_v)/\max(IDC_h, IDC_v)$
Spatial Frequency Diversity H(E)	$H(mag(DO))$
Phase Diversity H(P)	$H(hist(imag(DO)))$

The contributing factors are extracted separately from the vector of diffraction order components, and combined in the following heuristic formulation:

$$(1+\alpha)MP * (1+2D) * \max(IDC_h, IDC_v)^{(H(P)*H(E))^{\beta}} \tag{2}$$

where α is an empirical constant and β is an exponential, used primarily to match the scales of the entropy and other terms to give a reasonable dynamic range.

The dimension of the underlying weight vector change as a function of the maximum spatial sampling frequency and window size, (larger windows containing lower frequencies). Because the values depend on this dimension, difficulty ranking is only possible between patterns of the same window size.

3. VALIDATION

Validation of the methodology and software has proceeded in several stages. All the results presented here are for 22 nm technology node, during intermediate stages of the development and learning cycles. Initially, we reviewed the results on typical one and two dimensional test structures and subjected the rankings to scrutiny of lithographers, while comparing simulation results for a particular source and OPC formulation tuned for the source. Where patterns were weakest (giving hotspots at extreme defocus, or more stringent spacing width specifications) the formulation was adjusted so that these patterns were elevated in the ranking. For patterns of intermediate complexity and no errors, it may not be surprising that consensus on difficulty of two dimensional patterns is lacking.

Naturally we expect to leverage existing lithographic analysis and simulation tools to assess our predictions, but there are challenges and risks associated with this. Such software is evolving and possibly contains bugs which could distort the picture, and we must use some reference "plan of record" source and OPC recipe to evaluate our results. If we produce a new source using our difficulty estimates, the OPC will likely benefit from tuning to the new source. We also do not have any *a priori* way to measure the quality of an existing OPC as it evolves through the technology cycle.

With a ranked most difficult 10% subset of 1000 synthetic contact patterns (240 nm), we used joint optimization to design a source, then performed process window evaluation within the SMO subsystem to estimate common process window (CPW), a measure of the area in the dose-focus plane giving acceptable imaging parameters (i_{max}, i_{min} and slope) across all edges and evaluation points in the pattern representation.

In addition to ranking with the contact data, we used the LDE function to choose the representative elements for clusters. Previous work demonstrated quick turnaround validation of process and RET recipe changes can be performed with sampling and clustering of large layouts [9,10]; and recently clustering methods have been extended to choose patterns for "training" a source based on analysis of many patterns [2]. While many strategies are possible for choosing representative elements, we show that using the LDE metric gives improved results. It is also possible to simply sort by LDE rank, and with some experience one may choose a safe threshold value below which patterns are unlikely to fail and will not influence the optimization for the source.

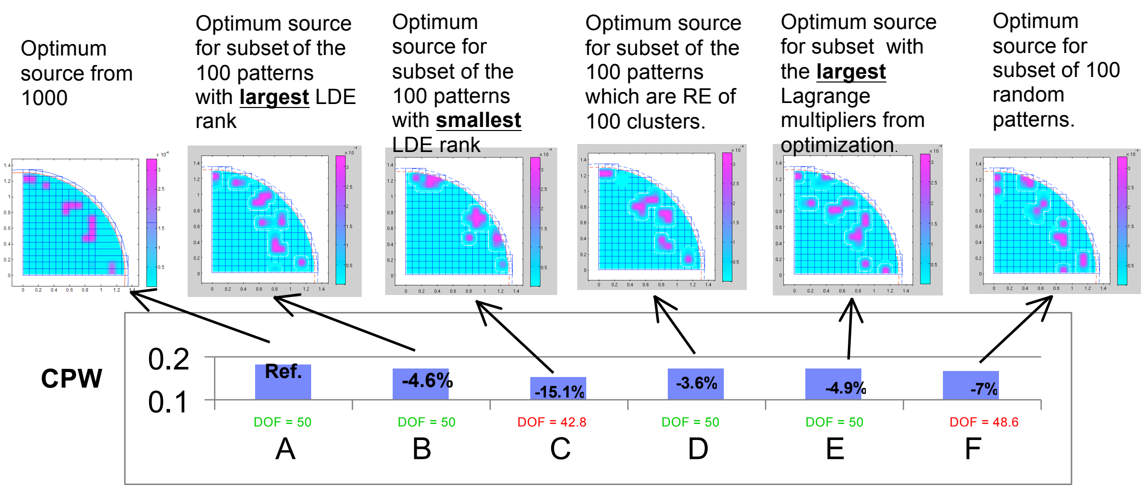


Figure 2. Results of Common process window measure for various sources. The reference A optimized the source using all 1000 patterns. Using only patterns with top 10% LDE values gave a source with common window 4.6% lower, while using LDE representative elements gave a common window 3.6% lower.

In subsequent experiments we focused on a single exposure interconnect level, using synthetic patterns with window size 3 pitches (240 nm) and 4 pitches (400 nm) square. These layers present some difficulty due to a greater degree of two dimensionality and more pitch and spacing flexibility. Clustering was performed on 1510 240x240 synthetic metal patterns, with parameters setting resulting in a compression ratio 1:8. We compared choosing the representative by the typical centroid (pattern nearest center in DO vector space) with using the highest LDE rank. Using Common Window (CW) as reported by the SMO tool, the pattern with the smallest CW was improved by 14% compared to the CW window obtained with the centroid pattern for training. We then evaluated the entire set (1510 layout clips) with the POR source and the source obtained by training on LDE selection. 2 sigma process assumption errors (nominal and process window space and width errors) were reduced by 93% from the POR source. In addition, we checked the source against some early auto-routed layouts and found that the source generalized well, reducing process window errors by 66%.

We evaluated larger patterns to understand whether additional memory and compute time is justified. Clustering 1000 400x400 synthetic metal patterns, we experimented with various compression ratios using the LDE representative selection. Best performance in reducing the errors was with 1:29, the highest compression ratio evaluated. 62% reduction was obtained in 2 sigma error markers (however the OPC had improved the baseline since the previous experiment). We tested the autorouted layout and now achieved only 44% reduction in errors. Analysis of the remaining errors in the full set indicated that most failing synthetic patterns were too dense to support SRAF insertion.

Beyond the quality improvement, a major contribution of using LDE can be found in the TAT savings. Selecting clips for the POR source is a process that requires engineers skilled in the art and significant time investment, due to the full cycle validation process. The execution time of LDE, in our Matlab prototype code, depends on the number and the size of clips, but it is in range of 300-500 clips/minute on a 300Ghz Xeon CPU. Reducing the set of clips used for training the source improves significantly the source optimization time, since this is exponentially dependent on the number of clips.

In a final experiment we examined the distribution of 1 sigma markers (i.e. for drawn spacing we detect errors where the CD appears at 1 sigma off nominal) at the union of defocus conditions evaluated for optical rule checking in real design layout for the metal routing layer. We used a 320x320 window with regular stepping on the layout, which was examined for presence or absence of markers. The diffraction order vector for each layout sample pattern and LDE values were computed, and the conditional probability distribution of presence of fail markers was plotted as in Figure 4.

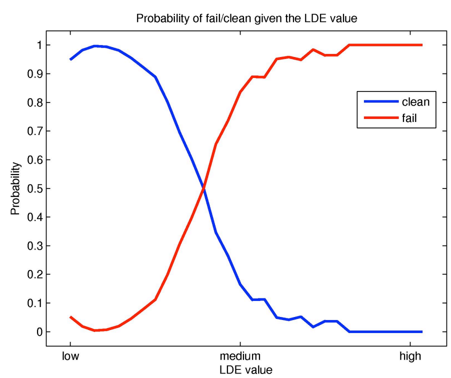


Figure 4. LDE value vs. conditional probability of fail in a set of 90,000 320x320 metal patterns

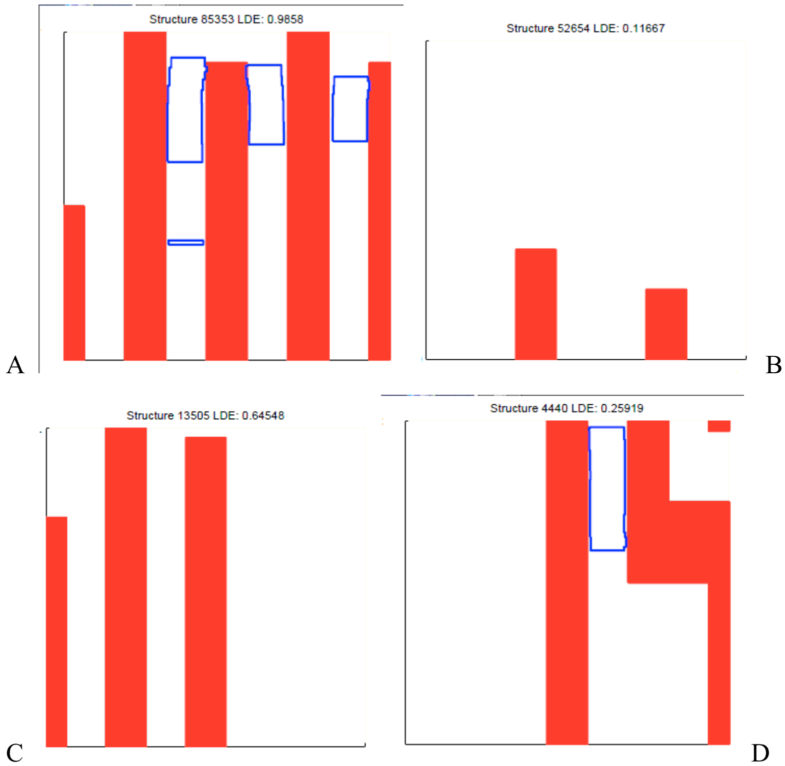


Figure 5. Examples of high and low ranked patterns (labeled 'Structures') from test site layout samples. Top A and B are from extreme high and low ranges of distribution. C is the highest ranked non failing clip, D is the lowest ranked failing clip in a small sample. Examining these extrema helps to refine the algorithm and test various window sizes.

Errors are worst case process window and more conservative spacing dimensions than those used to set groundrules, and were observed in the middle of yield learning and OPC and auto router rules development cycles thus should not be taken as a reflection of production capability.

The following table summarizes the results of the various experiments.

Table 3. Summary of validation results

Data Source	Measure	Result
Lithography test patterns	Lithographer's intuition, hotspot	Rankings improved
1000 Contacts, 200 nm ,100 training clips	Common Window on full set	Top 10% LDE sort within 5% of CW from full set
1510 metal synthetic pattern, 240 nm 193 training clips (1:8 compression ratio)	Common Window on full set	14% improved vs. centroid 7.3% improved vs. random
1510 metal synthetic patterns, 240 nm, 193 LDE training	2 σ CD errors on full set 2 σ CD , autorouted validation set	93% reduced from POR 66% reduced from POR
1000 metal synthetic patterns, 400 nm, 34LDE training set (1:29 Compression)	2 σ CD errors on full set 2 σ CD, auto routed validation set	75% reduced from POR 44% reduction from POR
500 sampled metal autoroute, 400x400	Correlation of LDE with 1 σ hotspots for mid-cycle source	80% of fail/not fail predicted by LDE

4. CONCLUSIONS

We have demonstrated a method of estimating and ranking printability difficulty which is not derived from a specific aperture and shown that it functions well in a large scale source mask optimization flow, providing runtime and memory improvements over processing a larger collection when used as a filter or as representative element selection. We have demonstrated up to 93% improvement in quality and orders of magnitude (minutes from days) improvement in turn around time. Solution quality approaches more complex parallel programming optimization methods [2].

Further it may be used as an easily computed metric to allow designers to improve layout printability or in physical synthesis methods, to direct design rule checking developers to patterns of interest.

The method currently has some limitations. For patterns which are too dense to support correction by SRAF insertion, the rank may be too low and unless topology breaking mask optimization is used the source may be biased by trying to solve such patterns. Adding some overall density correction factor could improve identification of such.

It may not correctly account for patterns which are not maximally difficult but have incompatible spatial frequencies, i.e. solving one will increase the difficulty of the other by driving it into a forbidden pitch zone. This can be addressed by computing joint spectra of multiple patterns, or assessing a difference of each pattern from a normalized aggregate spectrum.

REFERENCES

[1] A. E. Rosenbluth et al., "Intensive optimization of mask and sources for 22nm lithography", Proc. SPIE 7274 Optical Microlithography XXII, 727409 (2009).

[2] K. Lai et al., "Design specific joint optimization of masks and sources on a very large scale", Proc. SPIE 7973, 797308 (2011).

[3] D. O. Melville et al., "Demonstrating the benefits of source-mask optimization and enabling technologies through experiment and simulations", Proc. SPIE 7640, 764006 (2010).

[4] B. Yenikaya, A. Sezginer, "A rigorous method to determine printability of a target layout", Proc. SPIE 6521, 652112 (2007).

[5] J. A. Torres, O. Otto, F.G. Pikus, "Challenges for the 28 nm half node : is the optical shrink dead?", Proc. SPIE 7488, 74882A, 381047 (2009).

[6] H. J. Levinson, Principles of Lithography 2nd Ed., SPIE Press Bellingham, 17-24 (2004).

[7] A. K. Wong, Resolution Enhancement Techniques in Optical Lithography, SPIE Press, Bellingham, 65 (2001).

[8] http://www.si2.org/openeda.si2.org/dfmdictionary/index.php/Bossung_Plots

[9] D. DeMaris et al., "Automated Regression Test Selection for Optical Proximity Correction", ISMI Symposium on Manufacturing Effectiveness, (2006).

[10] D. DeMaris et al., System and method for testing pattern sensitive algorithms for semiconductor design, US Patent 7,353,472, (2008).

ACKNOWLEDGEMENTS

The authors thank Saeed Bagheri for early involvement with the formulation; David Melville and Alan Rosenbluth, for their assistance with validation via SMO methods; Phil Strenski and Mike Monkowski of IBM SRDC for providing layout patterns, J. Andres Torres of Mentor Graphics for insight into how pixelated sources interact with and impact OPC.