A Framework for Early and Systematic Evaluation of Design Rules

Rani S. Ghaida and Puneet Gupta

Electrical Engineering Dept., University of California, Los Angeles

{rani,puneet}@ee.ucla.edu

Abstract—Design rules have been the primary contract between technology and design and are likely to remain so to preserve abstractions and productivity. While current approaches for defining design rules are largely unsystematic and empirical in nature, this paper offers a novel framework for early and systematic evaluation of design rules and layout styles in terms of major layout characteristics of area, manufacturability, and variability. Due to the focus on co-exploration in early stages of technology development, we use first order models of variability and manufacturability (instead of relying on accurate simulation) and layout topology/congestion-based area estimates (instead of explicit and slow layout generation). The framework is used to efficiently co-evaluate several debatable rules (evaluation for a 104-cell library takes 20 minutes). Results show that: a) diffusion-rounding mainly from diffusion power-straps is a dominant source of variability, b) cell-area overhead of fixed gate-pitch implementation compared to 1D-poly implementation is tolerable (5%) given the improvement in variability, and c) 1D-poly restriction, which improves manufacturability and variability, has almost no area overhead compared to 2D-poly. In addition, we explore gate-spacing rules using our evaluation framework. This exploration yields almost identical values as those of a commercial 65nm process, which serves as a validation for our approach.

I. INTRODUCTION

The semiconductor industry is likely to see several radical changes in the fabrication and device technologies in the next decade. On the patterning front, disruptive changes include adoption of one or more of candidate next-generation lithography techniques such as nanoimprint, electron beam direct write, and extreme ultraviolet [1–4]. Each of these has challenging implications for layout methodologies and design rules (DRs). Resolution enhancement techniques (RETs) and other patterning solutions such as immersion and double-patterning lithography (DPL), off-axis illumination (OAI), sub-resolution assist features (SRAFs), and phase-shift mask (PSM) require additional layout-restrictive DRs [5–11]. Therefore, *early assessment of design restrictions imposed by technological choices is absolutely essential*.

DRs are the biggest design-relevant quality metric for a technology. Even small changes in DRs can have significant
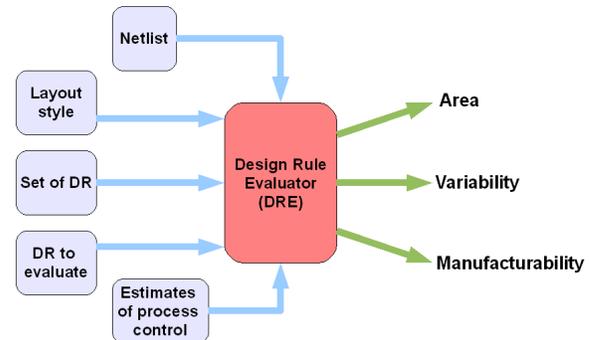
Figure 1: Overview-diagram of DRE framework.

impact on manufacturability [12] as well as circuit characteristics including layout area, variability, power, and performance [13, 14]. Unfortunately, even after decades of existence, design-rule evaluation and exploration is largely unsystematic and empirical in nature. Several published works have done "one-at-a-time" evaluation of design rules empirically [12, 15]. For example, [16] evaluates line-end extension rule electrically to conclude that it may be too conservative. Other recent works [17, 18] offer solutions to explore DRs from a pure printability perspective and do not examine the effects of DRs on circuit characteristics. Moreover, none of these methods account for layout topology changes that may happen when DR values change significantly. They also ignore several practical constraints imposed on layouts by the standard-cell design methodology (e.g., cell width and height being quantized). Finally, these approaches are simulation and/or explicit layout generation-based, which makes them slow and dependent on model accuracy.

To the best of our knowledge, this paper proposes the first framework to explore area-manufacturability-variability tradeoffs of design rules systematically and in a quantitative manner. Rather than fine-tuning DRs, our goal is to make early decisions *before* exact process and design technologies are known. At this stage, accurate evaluation methods and models are unlikely to be available and the return on investment of using them is fairly low. As a result, we use simple but justified approximations for manufacturability and variability unlike [19, 20] that rely on layout generation or perturbation. Since design rule space is very large, we further use fast layout topology generation methods to estimate area as opposed to full-blown layout generation. The accuracy of the former is surprisingly good and allows for explicit "layout style"

Figure 2: Illustration of major DRs, their notations and values in FreePDK 45nm process.
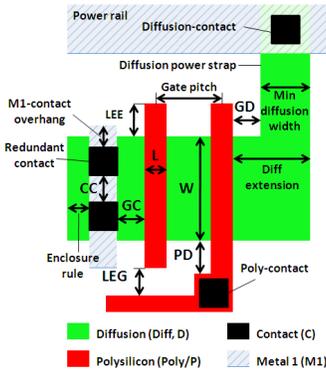
DR	Value[nm]
L	≥ 50
M1 1/2 pitch	≥ 65
Gate pitch	≥ 190
Min diff width	≥ 90
LEE	≥ 55
LEG	≥ 75
GC	≥ 35
CC	≥ 75
GD, PD	≥ 50
M1-C overhang	≥ 35
Diff enclosure	≥ 5
Poly enclosure	≥ 5
Diff extension	≥ 70

Figure 3: Techniques and notations used in layout topology generation.

Figure 6: S/D-to-gate interconnections are routed on M1 and poly layers and S/D-to-S/D interconnections are routed on M1 only. We assume a single-trunk Steiner tree for routing.

guidelines as we show later in this paper.

Figure 1 shows the structure of the DR Evaluator (DRE). Given SPICE netlists of cells (possibly scaled down from a previous technology generation), layout style and preferences (e.g., redundant contacts), design rules and their values (see Figure 2), and estimates of process control (e.g., overlay error distribution), only the values of DRs to be evaluated are modified while all other rules remain unchanged. This modified set of DRs is then used to estimate the layout and determine major metrics of area, manufacturability, and variability[1].

We make the following contributions.

- We offer a framework for *early* and *systematic* co-evaluation and exploration of DRs, layout styles, and library architectures.
- Fast runtime of the proposed framework permits its use to narrow down on reasonable DR choices for conventional compute-expensive DR generation and optimization methods.
- We evaluate debatable DRs and layout style decisions such as: 1D/2D poly, fixed gate-pitch, diffusion/M1 power-straps, cell-height, and I/O pin-access requirement.
- We demonstrate the usage of the framework for comparison between processes and DR exploration through case studies.

The remaining paper is organized as follows. Section II describes the methods used for layout topology generation and congestion estimation. Section III and IV provide details on models and metrics of manufacturability and variability evaluation[2]. In Section V, comparative evaluations of several debatable DRs are performed in 45nm process. In addition, we analyze area-manufacturability-variability tradeoffs of commercial standard and low power 65nm process and illustrate the use of our framework for DR co-exploration. Finally, Section VI concludes and discusses future work.

II. AREA ESTIMATION

The number of design rules is growing tremendously and design rule manuals (DRM) are becoming unmanageable as we move toward smaller feature sizes [21, 22]. In addition, DRs need to be evaluated individually as well as collectively over a wide range of values. As a result, our framework was designed for *fast* evaluation necessary to enable DR exploration/optimization.

This section describes the methods for fast layout topology generation and congestion estimation used in DRE framework.

A. Layout Topology Generation

Major transistor placement techniques used for layout-area reduction are highlighted in Figure 3. Transistor pairing consists of placing two inter-connected transistors, one pMOS and another nMOS transistors, on the same column to minimize wire length and facilitate routing as well as to ensure more layout regularity. The coupled pMOS/nMOS transistors are referred to as transistor pairs. Transistor folding consists of replacing a large transistor by equivalent multiple transistors of smaller sizes connected in parallel. Transistor chaining is the process of connecting transistors of the same type by sharing the same diffusion area. Non-isolated transistors of the same active region form a transistor chain. A transistor stack refers to two transistors sharing a diffusion area that is not connected to any other parts of the circuit (i.e. contact-free diffusion).

Figure 4 outlines the flow of transistor placement used for layout estimation and describes the algorithms used at each step. We illustrate the application of these steps on a standard cell in Figure 5.

Exact transistor and pin locations along the horizontal direction are then determined based on minimum DR dimensions. As for transistor locations along the vertical direction, we consider three possibilities: a) as near as possible to power rails, b) exactly in the center of p/n networks, and c) as near as possible to p/n interface. The choice of vertical location of transistors is regarded as a layout style, which can also be evaluated by the framework[3].

---

[1]DR choices also affect delay, power, reliability, and other circuit metrics along with some notion of designability. For brevity and continuity, we do not consider these aspects of DR evaluation in this work.

[2]For the purpose of readability and brevity, we skip certain details about the methods and algorithms that were used.

[3]This decision has implications on M1-congestion as well as the impact of stress and well-proximity effect on performance [13, 14].
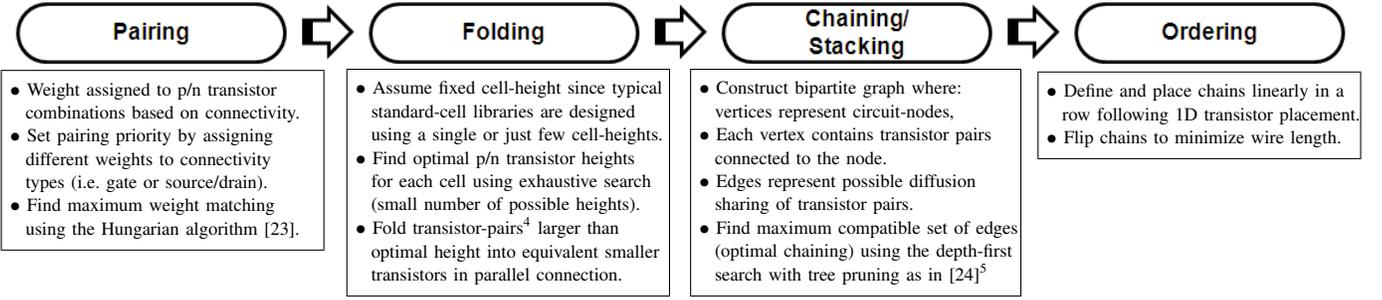
Figure 4: Flow of layout topology generation in DRE framework.
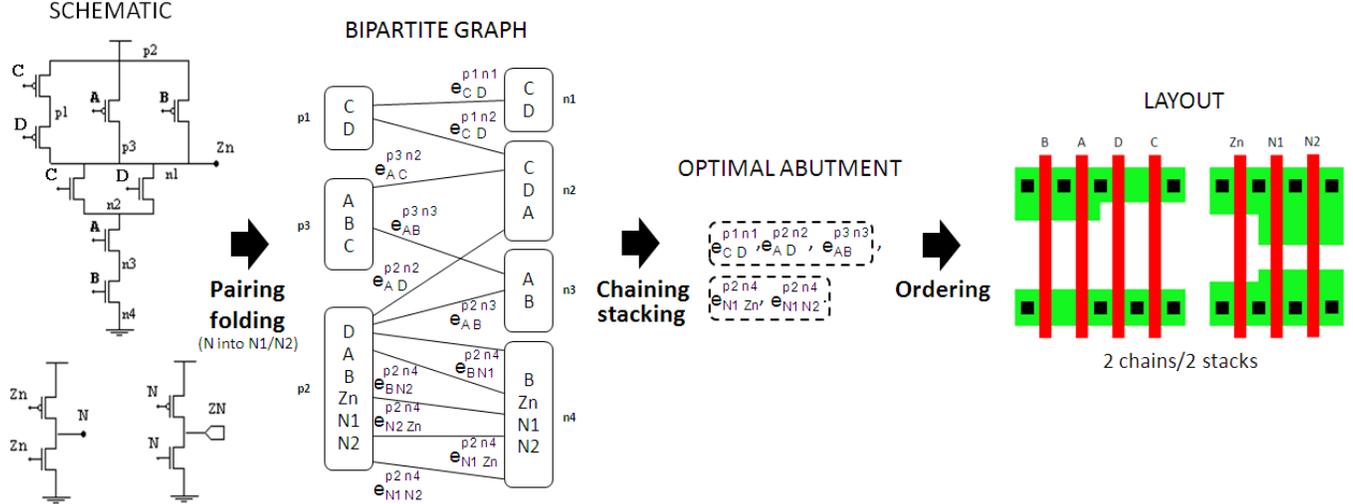
Figure 5: Example that illustrates our layout topology generation for 4-input OAI standard-cell.

B. Routing Estimation

Once transistor placement is complete, locations of gates and contacts to transistor source/drain (S/D) terminals are determined. S/D contacts connected to power supply are located as close as possible to the power rail while meeting DR requirements. All other S/D contacts are located near p/n interface to reduce the length of wires connecting p-to-n type transistors.

Rather than performing the time-consuming step of actual routing, DRE estimates routing to approximate the wire length and models congestion with the intention of considering its effect on layout area. Hereafter, the term "routing" denotes "estimated routing" and not actual routing.

Transistor interconnections, i.e. intra-cell routing, are assumed to be performed using polysilicon (poly) and first metal (M1) layers only. There are three such interconnections: gate-to-gate, S/D-to-gate, and S/D-to-S/D. The way gate-to-gate interconnections are performed depends on poly-routing restrictions, which are characterized by the layout style and will be evaluated later in this paper. Details of S/D-to-gate and S/D-to-S/D routing are shown in Figure 6.

---

[4]We force transistors to be folded in pairs.

[5]The algorithm is made faster by limiting the number of iterations (to six). This has negligible effect on the quality of results since the optimal solution is among the first few examined solutions in almost every case. Moreover, the algorithm is modified to favor solutions with more transistor stacks over regular chaining or vice versa.

C. M1-Congestion Estimation

Once all routes are estimated, DRE calculates M1 wire length in x and y directions including via and contact-landing pads for I/O pins. Occupied track-length in a particular routing-direction is then determined as the sum of wire length, line-end spacing DR between wire segments, and track-length blocked by wires in the orthogonal direction. At intersections of tracks with orthogonal wires, a track-length equal to the minimum line-width plus spacing DRs is blocked. However, if the intersection actually forms a corner-connection between vertical and horizontal wires, no track-length is blocked.

M1 track-congestion in one direction is defined as the ratio of occupied to available track-length. To accommodate M1 wiring, cell-area is increased if M1 track-congestion is larger than a certain threshold. This threshold depends on the intra-cell routing efficiency and empty space required on M1 to access the cell I/O pins. Furthermore, routing efficiency is a function of the proportion of non-preferred direction wire length to total wire length. If wires are mostly in one direction, routing is efficient and increasing the cell-area is only necessary for very high congestion. In contrast, if wires are evenly distributed in the two directions, routing is difficult and increasing cell-area is expected for relatively low M1-congestion. To capture these effects, we model track-congestion threshold as follows:

$$C_{threshold} = \alpha + \left| \frac{U_x - U_y}{U_x + U_y} \right| \times \beta - \gamma, \qquad (1)$$
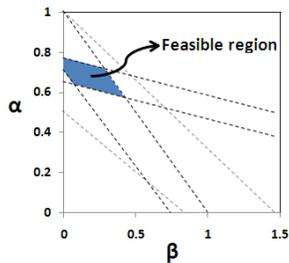
Figure 7: Illustrating example for extraction of α and β parameters of Equation 1 from cell congestion data.
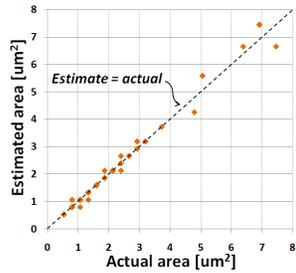
Figure 8: Estimated versus actual area for the entire Nangate Open Cell Library [25] (104 standard-cells).

Figure 9: Virtual artwork representation for (a) horizontal and (b) vertical M1 wires.

where U_x and U_y are the track utilization in x and y directions. Here, track utilization is defined as the ratio of occupied track-length without consideration for track-blockage from the orthogonal direction wiring to the available track-length. α and β parameters are a function of intra-cell routing efficiency and γ is a function of empty space left for inter-cell router to access cell-I/O pins. The values of all these parameters are set by the user based on the router specs. Figure 7 depicts one method to extract α and β parameters either from trial routes of few cells or from cells of a previous generation library. Every single cell implementation adds one lower and one upper bound line that narrow down the feasible solution space. Thus, the more cells are used, the more precise the solution is. The lower bound line is defined by cell congestion. If cell-area is increased to accommodate for M1-wiring, the upper bound line is defined by cell congestion *after* the increase; otherwise, the upper bound line corresponds to a congestion of 1. In the end, exact values of α and β are approximated by coordinates of the feasible region's geometric centroid.

D. Runtime and Validation of Area Estimation

In order to validate our layout estimation method and its efficiency, we use DRE framework to estimate the topology of the entire Nangate 45nm Open Cell Library [25] and estimate cell-area. Comparison between estimated and actual areas is depicted in Figure 8. Results show very good accuracy of the layout estimation method (2.4% average error). The source of imperfect area estimation is from using different layout styles[6].

The runtime of DRE evaluation procedure for the entire cell-library is 20 minutes real time on a single processor of 2GHz clock speed and 2MB cache.

III. MANUFACTURABILITY

Our manufacturability index for evaluating DRs is the probability of survival (POS) from three major sources of failure[7]:

1) contact-defectivity (a.k.a. contact-hole failure);
2) overlay error (i.e. misalignment between layers) coupled with lithographic line-end shortening (a.k.a. pull-back);

3) random particle defects.

POS associated with contact-hole failure is equal to the number of non-redundant contacts in the layout times contact-hole failure rate. In case contact-redundancy is implemented, duplicated contacts are assumed to always yield since the probability for *two* contacts connected to the *same* pin to fail is negligible.

Overlay vector components in x and y directions are described by a normal distribution with zero mean and process-specific 3σ estimate. Since overlay vector is roughly the same across the die, overlay-induced failure analysis needs to consider only the most critical spot where the least amount of overlay is required to cause a failure. We compute POS from overlay causing: failure to connect between contact and poly/M1/diffusion, gate-to-contact short defect, and always-on device caused by poly-to-diffusion overlay error. Connection failure at contacts occurs when the area of overlap with top/bottom connecting layers is smaller than a certain threshold-value. Thus, we consider overlay in both x and y directions in this analysis. In gate-related failure analysis, overlay in just one direction is considered since gates are presumably unidirectional. Moreover, we assume all layers are aligned to a reference alignment mark on substrate[8] and overlay between different layers and the reference layer to be independent[9]. The overall POS from overlay is then calculated as the product of POS from independent overlay errors. If overlay is assumed to be completely a die-to-die variation, then POS of the die is p (equal to POS of the most overlay-critical spot in layout). On the other extreme, if overlay is completely random within-die variation, then POS of the die is p^n, where n is the total number of critical spots in the design. Reality is closer to the former situation (since field and wafer level components dominate intra-field components [26]), which is our assumption in this paper.

To capture failure caused by random particles, we perform critical area analysis for open and short defects at M1/poly/contact layers and short defects between gates and diffusion-contacts. For fast analysis, we use the virtual artwork approach proposed in [27]. Poly and contact layers are represented by strips separated by spacing-DRs; whereas for M1 layer, this separation corresponds to the spacing that makes the wires as far apart as possible (see example of Figure 9). The virtual artwork representation allows quick calculation of critical area as a function of defect size by applying a closed-form equation model. The probability of failure from random particles is then inferred from average critical area

---

[6]In particular, Nangate-cells are generated with preference for transistor chaining over pairing; whereas, our method enforces pairing of transistors with same gate-signal. The latter is better approach for modern 1D-poly and fixed gate-pitch implementations.

[7]More involved models of lithography induced failures are part of our ongoing work.

[8]This can be modified to conform with the process alignment strategy.

[9]In reality overlay of different layers with the reference layer have some degree of correlation. This can be dealt with by reducing the amount of overlay (i.e. use smaller 3σ for overlay distribution).

Table I: TEST BENCHES USED IN OUR EXPERIMENTS AND THEIR CORRESPONDING NUMBER OF CELL INSTANCES AND UNIQUE CELL-TYPES.

Circuit	Description	Cell instances	Cell types
nova	video compression decoder	43156	90
vga	VGA/LCD controller core	36097	63
mips	processor core	17091	57
ae18	processor core	4358	51

for all defect sizes and average defect density (i.e. random particles per unit area) following standard methods [28]. In this calculation, we use the following defect size distribution model [29, 30]:

$$f_s(r) = \begin{cases} \frac{2(n-1)r}{(n+1)r_0^2} & \text{if } 0 \le r \le r_0, \\ \frac{2(n-1)r_0^{n-1}}{(n+1)r^n} & \text{if } r > r_0. \end{cases} \quad (2)$$

where r is the defect size, r_0 is the defect size with peak density (a.k.a. critical defect size), and n is a parameter related to the cleanliness of the fabrication process and ranges between 2 and 4. Overall POS (i.e. complement of probability of failure) from all sources is then calculated as the product of POS from individual sources.

IV. VARIABILITY

In sub-wavelength lithography regime, three sources of printing imperfection causing gate-dimension variation are dominant (depicted in Figure 10):

- diffusion and poly corner-rounding;
- line-end tapering under overlay error and line-end pull-back;
- CD variability.

The contribution of each source to gate length and width variations (ΔW and ΔL) is modeled independently. The total change in drive current is set as our variability index for evaluating and comparing DRs and is calculated using the following equation:

$$\Delta(\frac{W}{L}) = \frac{\sum_{allgates} \left| \Delta(\frac{W}{L})_i \right|}{(\frac{W_{tot}}{L})_{ideal}}, \quad (3)$$

where i represents the source of variability[10].

Since the resulting ΔW and ΔL are not across the entire gate, we quantify their contribution to $\Delta(\frac{W}{L})$ by modeling devices as parallel slices of transistors[11].

Diffusion rounding at corners formed by diffusion power-straps and unleveled abutment of transistors (as depicted in Figure 10) induces width variation at the gate-edge. In addition, poly corner-rounding in bends and contact-pads near the gate represents an important source of gate-length variation. The shape of the rounding is a function of the corner dimensions and is modeled as $\Delta H = K_1 \Delta Y / \sqrt[n]{1 + (\frac{\Delta Y}{K_2})^n}$ where $K_1 = Ce^{D\Delta X}$ and $K_2 = A\Delta X + B$. In this model,

---

Figure 10: Illustration of slicing model, rounding model parameters, and the sources of gate length and width variability considered in DRE framework.

ΔX, ΔY, and ΔH are depicted in Figure 10 and A, B, C, D, and n parameters are fitted to give $< 0.8nm$ error with measured data from printed-image simulations on a fairly wide range of practical corner-dimensions, i.e. $\Delta X = 30 \rightarrow 70nm$ and $\Delta Y = 10 \rightarrow 200nm$ (for brevity concerns, we do not discuss the derivation of this model). It is worth noting that approximate predictive rounding-models fitted from tentative simulation models, which are typically available in early stages of technology development, could be used in lieu of the current model.

Line-end tapering can affect the length of the gate at its edge. This effect becomes more significant when considering line-end pull-back and poly-to-diffusion overlay error. The tapered shape and gate-length at the edge are described using the model offered in [16][12] while accounting for line-end pull-back (mean value) and overlay errors (from distribution). Line-ends are assumed to extend beyond the gate as far as possible unless the user enforces minimum line-end extension (LEE) rule for the entire layout.

CD uniformity (CDU) is another major contributor to the change in drive current. In our framework, CDU is described by a distribution, which captures the dependency on dose and focus variations. Pattern-dependency is captured by using different CDU 3σ values for each poly-patterning style including 1D/2D patterning and fixed/non-fixed pitch, which can seriously impact CDU [11, 35, 36].

After determining all $\Delta(\frac{W}{L})$ terms from different sources, we compute the absolute sum of all terms for the entire layout with the intention of highlighting actual gate variability. Finally, the drive current variability index is calculated using Equation 3.

V. EXPERIMENTAL SETUP AND RESULTS

In this section, we evaluate and analyze major debatable DRs and layout styles for 45nm open-source FreePDK process [37]. DRE framework is also used to compare standard and low power 65nm process from a commercial vendor. In

---

[10]We realize that this estimate is approximate as effects from different sources can interfere. Nevertheless, it is a good indicator of worst-case variability and process control requirement.

[11]More accurate slicing models of [31–34] can also be embedded in the framework if they are available.

[12]$L_i = 2a\left(1 - |\frac{h_i - k}{b}|^n\right)^{\frac{1}{n}}$, where l_i is the gate-length at i location in the line-end extension, h_i is the distance from i to gate-edge, a is half the nominal gate-length, b is the line-end extension, and k and n parameters describe the taper-shape. In our experiments, we use $k = 0$ and $n = 3$.
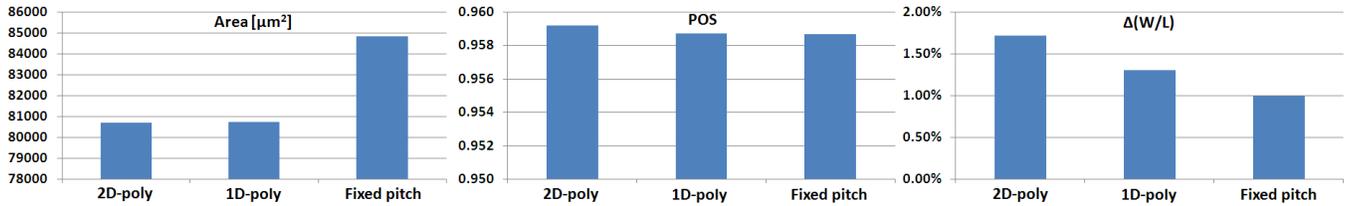
Figure 11: Evaluation of restrictive poly-patterning rules on 45nm FreePDK process.

Table II: PROCESS CONTROL PARAMETERS USED IN THE EXPERIMENTS.

Parameter	45nm	65nm
Avg defect density [$faults/m^2$]	1395	1757
Critical defect size [nm]	34	45
Max defect size [nm]	250	250
Fab cleanliness parameter	3	3
Contact-holes rate [ppm]	0.00004	0.00004
Overlay (3σ) [nm]	13	15
Line-end pull-back (mean) [nm]	10	14
Gate CDU (3σ) [nm]	2.6	3.3
Critical M1 line-width [nm]	10	15
Critical poly line-width [nm]	15	20
Critical contact-width [nm]	10	15

another experiment, we co-explore two gate-spacing related DRs collectively.

A. Testing Setup

Throughout the experiments, we use four benchmark designs from [38] synthesized using Nangate 45nm Open Cell Library (scaled for testing with 65nm process). Table I describes all designs and lists their number of cell-instances and unique cell-types.

Experiments were performed using 45nm open-source FreePDK process and 65nm process from a commercial vendor. Estimates of process control parameters associated with each process are summarized in Table II. We use projected values from ITRS technology roadmap [1] and typical values for critical M1 and poly line-width and critical contact-width, which represent the minimum acceptable width for the defect not to be considered a failure. CDU value in the table is for 2D-poly patterning. For fixed pitch 1D-poly, we use CDU 3σ improvement factor of 47% over 2D-poly reported by IBM in [11] and assume half the improvement is from unidirectional patterning.

α and β parameters of the congestion threshold model (Equation 1) are extracted from Nangate cells using the method discussed in Section II-C and γ parameter of the model is set to zero (i.e. no extra space requirement for I/O pin-access).

Since the area of the benchmark designs is relatively small, we normalize POS values to a $100mm^2$ chip-area. We determine for the base case in each experiment the number of design copies that can fit in $10 \times 10mm$ chip size with 80% cell-area utilization and find the corresponding number of contacts and critical area.

Results of DR evaluations are a strong function of the base set of rules, layout styles, library architecture, and design type and, hence, they are *not generalizable*. First, we perform studies on 45nm FreePDK process and later we perform studies on a 65nm commercial process as an example.

The number of possible case studies that DRE framework can perform is huge. For brevity, we only perform studies of several debatable DRs and layout styles including: 1D/2D-poly, fixed gate-pitch, diffusion/M1 power-straps, and 7/9/11-track cell-height. The baseline experiment to which all results are compared to is with the following setup:

- 1D poly (non-fixed pitch),
- M1 power-straps,
- and 9-track cell-height.

B. Evaluation of Poly-Restrictions

Three configurations of poly-patterning styles are investigated:

- no poly-routing, i.e. 1D-poly,
- limited poly-routing,
- and non-restricted poly-routing, i.e. 2D-poly.

In case of 1D-poly configuration, poly is used only to connect dual gates (i.e. gates of same transistor-pair). In case of limited poly-routing, it is also used to connect adjacent gates in the same p or n network. In case of 2D-poly, it is used to perform all gate interconnections unless blocked by previous routing or diffusion power-straps.

Figure 11 shows area, manufacturability, and variability tradeoffs associated with poly-patterning styles including 1D/2D-poly and fixed gate-pitch on 45nm process with M1 power-straps and 9-track cell-height.

We observe that 2D and 1D-poly results in almost identical cell-area, which has two reasons. First, cells have fairly simple poly-patterning as a result of pairing transistors with same gate-signal. Second, gate-alignment requirement for 1D-poly induces negligible area overhead in FreePDK process, which uses the same rule for minimum and contacted gate-pitch. On the other hand, 2D-poly leads to 32% larger variability compared to 1D-poly, which is mainly caused by CDU improvement associated with unidirectional patterning. In another experiment, we compare 2D-poly and limited poly-routing. Results are almost identical and, thus, allowing U-shape and W-shape poly-patterns with RET complications might not bring real benefits.

Fixed gate-pitch implementation leads to 23% less variability compared to 1D-poly implementation, but has 5% area overhead. This area overhead is relatively small because minimum gate-pitch is equal to contacted pitch in FreePDK process and, consequently, gate-spacing increase is necessary only for isolated-gates.

C. Evaluation of Layout Styles

Figure 12 shows area, manufacturability, and variability tradeoffs associated with M1/diffusion power-straps on 45nm
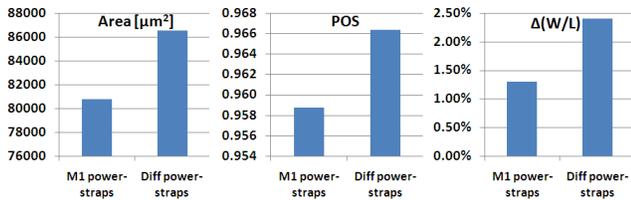
Figure 12: Evaluation of M1/diffusion power-straps styles on 45nm FreePDK process.
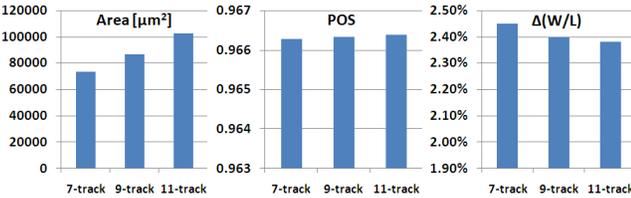
Figure 13: Evaluation of 7/9/11-track cell-height on 45nm FreePDK process.

process with 1D-poly and 9-track cell-height. Diffusion power-strap style results in much larger variability than in the case of M1 power-strap style (84% larger), which manifests the intensity of diffusion-rounding effect. The reason for this large effect is the fact that cells are packed in the horizontal direction to minimize cell-width and minimum design rules are used. In contrast, poly-rounding and line-end tapering effects are much less important because cells are normally relaxed in the vertical direction (cell-height being fixed).

Furthermore, a 7% area overhead is associated with diffusion power-strap style for the reason highlighted in Section V-B. In another experiment (not shown in Figure 12) with a smaller cell-height (7 tracks), diffusion power-strap style leads to 5.8% area improvement because it reduces M1-congestion, which affects cell-area seriously in this case.

Diffusion power-straps have some manufacturability benefits, however. Gate-to-contact shorts are reduced and contact redundancy for power connections is implemented at no cost since these contacts are placed on the power-rail in this case.

It is important to note that area-impact of diffusion/M1 power-straps style depends strongly on design rule values as Figure 15 would show in a later study that different DRs yield completely opposite results.

We also investigate different cell-height decisions. Figure 13 shows area, manufacturability, and variability tradeoffs associated with 7/9/11-track cell-heights on 45nm process with 1D-poly and M1 power-straps style. Results show a minor effect of cell-height decision on variability. This is because poly rounding and line-end tapering that are affected by cell-height
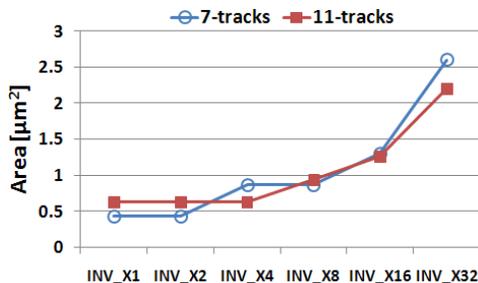
Figure 14: Increasing area with increasing transistor-width for 7-track and 11-track cell-height.
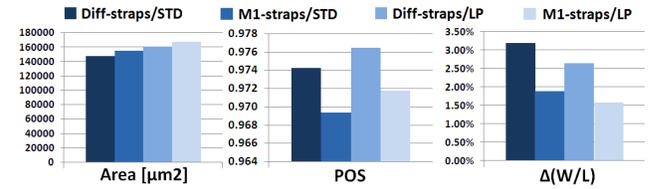
Figure 15: Restrictive DR study for commercial standard and low power 65nm-processes.

decision are second-order sources of variability as discussed earlier. The smallest cell-area of the benchmark designs is achieved with 7-track cell-height. However, this is not true for all cells as large cell-height is more suitable for cells with wide transistors (depicted in Figure 14), i.e. high-performance designs.

D. DR Comparison of Different Processes

Comparison of DR sets of different processes is another application of our framework. Here, we compare DRs of a standard and a low power 65nm process from the same commercial vendor. We do this comparison with diffusion/M1 power-strap style and 1D-poly patterning.

Results depicted in Figure 15 show an advantage of low power over standard process in terms of variability and manufacturability; on the other hand, standard process is more area-efficient (7.9% less area).

Compared to M1 power-straps, diffusion power-straps lead to larger variability but better manufacturability as in the previous study with 45nm process. Contrary to the previous experiment, diffusion power-straps lead to smaller cell-area in this case (by 5.2% and 4.2% for standard and low power processes respectively). This area improvement is due to *reduced* M1-congestion and gate-spacing at power connections according to design rule values of the 65nm commercial process.

E. DR Exploration

The framework is used for co-exploration of gate-to-diffusion (GD) and gate-to-contact (GC) rules in 65nm process. We perform the study for all benchmark designs and use diffusion power-straps and 1D-poly patterning styles, which lead to the most area-efficient solution for this process.

Results are depicted in Figure 16 where each data point represents a GD/GC value. Minimum-area and minimum variability to POS ratio solutions are shown on the plot. The solution corresponding to process GD/GC actual values falls very near the Pareto optimal frontier. Although quite simplistic, this example provides compelling evidence of our evaluation metrics fidelity and validates our approach.

VI. CONCLUSIONS AND FUTURE WORK

We proposed a novel framework for co-evaluation and exploration of design rules and technology decisions (**available for download at http://nanocad.ee.ucla.edu/Main/Projects**). By using first order models of variability and manufacturability and layout topology/congestion-based area estimation, our framework can evaluate big decisions *before* exact process

and design technologies are known. In this paper, we illustrated potential applications of our framework for collective DR evaluation and exploration as well as comparison of DRs from different processes. The framework can also be used to narrow down on reasonable DR choices for conventional DR generation and optimization methods. To the best of our knowledge, this is the first work that includes all area/manufacturability/variability metrics in the evaluation. Nevertheless, this is just the first step and our ongoing work pursues the following directions:

- addressing design rule effects on other layout and circuit characteristics including performance, power, reliability, and some notion of designability;
- introducing a 2D printability model (not based on field simulation), for example, derived from [39, 40];
- extrapolating DR evaluation to the chip level and include intermediate and global metal/via layers;
- studying interactions and tradeoffs of variability and area, as in [41] for example;
- improve chaining runtime for transistor with large number of folds (i.e. fingers).

REFERENCES

[1] International Technology Roadmap for Semiconductors, Report 2007.

[2] I. Yoneda et al., "Study of nanoimprint lithography for applications toward 22nm node cmos devices," in Proc. SPIE, vol. 6921, 2008, p. 692104.

[3] T. Maruyama et al., "Ebdw technology for eb shuttle at 65nm node and beyond," in Proc. SPIE, vol. 6921, 2008, p. 69210H.

[4] H. Meiling et al., "Euvl system: moving towards production," in Proc. SPIE, vol. 7271, 2009, p. 727102.

[5] K. Lai et al., "32 nm logic patterning options with immersion lithography," in Proc. SPIE, vol. 6924, 2008, p. 69243C.

[6] C. A. Mack, "Seeing double," IEEE Spectrum, pp. 46–51, 2008.

[7] G. Bailey et al., "Double pattern eda solutions for 32nm hp and beyond," in Proc. SPIE, vol. 6521, 2007, p. 65211K.

[8] A. Hazelton et al., "Double patterning requirements for optical lithography and prospects for optical extension without double patterning," in Proc. SPIE, vol. 6924, 2008, p. 69240R.

[9] A. Balasinski et al., "Patterning techniques for next generation ic's," in Proc. SPIE, vol. 6798, 2008, p. 679803.

[10] L. W. Liebmann, "Layout impact of resolution enhancement techniques: impediment or opportunity?" in Proc. IEEE Intl. Symp. on Physical Design ISPD'03, 2003, pp. 110–117.

[11] L. W. Liebmann et al., "High-performance circuit design for the ret-enabled 65nm technology node," in Proc. SPIE, vol. 5379, 2004, pp. 20–29.

[12] Y. Zhang, J. Cobb, A. Yang, J. Li, K. Lucas, and S. Sethi, "32nm design rule and process exploration flow," in Proc. SPIE, vol. 7122, 2008, p. 71223Z.

[13] V. Joshi, B. Cline, D. Sylvester, D. Blaauw, and K. Agarwal, "Leakage power reduction using stress-enhanced layouts," in Proc. ACM/IEEE Design Automation Conference DAC'08, vol. 7122, 2008.

[14] Y. Sheu et al., "Modeling well edge proximity effect on highly-scaled mosfets," in Proc. of IEEE CICC'05, 2005, pp. 831–834.

[15] L. Capodieci, P. Gupta, A. B. Kahng, D. Sylvester, and J. Yang, "Toward a methodology for manufacturability-driven design rule exploration," in Proc. IEEE/ACM Design Automation Conference DAC'04, 2004, pp. 311–316.

[16] P. Gupta, K. Jeong, A. B. Kahng, and C. Park, "Electrical metrics for lithographic line-end tapering," in Proc. SPIE, vol. 7028, 2008, p. 70283A.

[17] V. Dai, L. Capodieci, J. Yang, and N. Rodriguez, "Developing drc plus rules through 2d pattern extraction and clustering techniques," in Proc. SPIE, vol. 7275, 2009, p. 727517.

[18] S. Chang et al., "Exploration of complex metal 2d design rules using inverse lithography," in Proc. SPIE, vol. 7275, 2009, p. 72750D.

[19] A. R. Subramaniam, R. Singhal, C. Wang, and Y. Cao, "Design rule optimization of regular layout for leakage reduction in nanoscale design," in Proc. Asia and South Pacific Design Automation ASP-DAC'08, 2008, pp. 474–479.

[20] S. Kobayashi et al., "Yield-centric layout optimization with precise quantification of lithographic yield loss," in Proc. SPIE, vol. 7028, 2008, p. 70280O.

[21] L. Liebmann, L. Pileggi, J. Hibbeler, V. Rovner, T. Jhaveri, and G. Northrop, "Simplify to survive, prescriptive layouts ensure profitable scaling to 32nm and beyond," in Proc. SPIE, vol. 7275, 2009, p. 72750A.

[22] J. Yang, L. Capodieci, and D. Sylvester, "Layout verification and optimization based on flexible design rules," in Proc. SPIE, vol. 6156, 2006, p. 61560A.

[23] J. P. S. Kuhn, "The hungarian method for the assignment problem," Naval Research Logistics Quarterly, vol. 2, pp. 83–97, 1955.

[24] C. Hwang, Y. Hsieh, Y. Lin, and Y. Hsu, "A fast transistor-chaining algorithm for cmos cell layout," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 9, no. 7, pp. 781–786, 1990.

[25] Nangate open cell library v1.2. [Online]. Available: http://www.si2.org/openeda.si2.org/projects/nangatelib

[26] B. Eichelberger et al., "32nm overlay improvement capabilities," in Proc. SPIE, vol. 6924, 2008, p. 69244C.

[27] W. Maly, "Modeling of lithography related yield losses for cad of vlsi circuits," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. CAD-3, no. 3, 1985.

[28] I. Koren and Z. Koren, "Defect tolerance in vlsi circuits: techniques and yield analysis," in Proc. of the IEEE, vol. 86, no. 9, 1998, pp. 1819–1837.

[29] d. G. J. P., "Yield modeling and beol fundamentals," in Proc. IEEE/ACM System Level Interconnect Prediction SLIP'01, 2001, pp. 135–163.

[30] C. H. Stapper, "Modeling of integrated circuit defect sensitivities," IBM Journal of Research and Development, vol. CAD-3, no. 3, pp. 549–557, 1983.

[31] P. Gupta, A. B. Kahng, Y. Kim, S. Shah, and D. Sylvester, "Investigation of diffusion rounding for post-lithography analysis," in Proc. Asia and South Pacific Design Automation ASP-DAC'08, 2008, pp. 480–485.

[32] R. Singhal, A. Balijepalli, A. Subramaniam, F. Liu, S. Nassif, and Y. Cao, "Modeling and analysis of non-rectangular gate for post-lithography circuit simulation," in Proc. of Design Automation Conference DAC'07, 2007, pp. 823–828.

[33] P. Gupta, A. B. Kahng, S. Nakagawa, S. Shah, and P. Sharma, "Lithography simulation-based full-chip design analyses," in Proc. SPIE, vol. 6156, 2006, p. 61560T.

[34] S. X. Shi, P. Yu, and D. Z. Pan, "A unified non-rectangular device and circuit simulation model for timing and power," in Proc. IEEE/ACM Int'l Conf. on Computer-Aided Design ICCAD'06, 2006.

[35] M. C. Smayling, H. Liu, and L. Cai, "Low k1 logic design using gridded design rules," in Proc. SPIE, vol. 6925, 2008, p. 69250B.

[36] L. Pileggi, H. Schmit, A. Strojwas, P. Gopalakrishnan, V. Kheterpal, A. Koorapaty, C. Patel, V. Rovner, and K. Tong, "Exploring regular fabrics to optimize the performance-cost trade-off," in Proc. IEEE/ACM Design Automation Conference DAC'03, 2003.

[37] Freepdk. [Online]. Available: http://www.eda.ncsu.edu/wiki/FreePDK

[38] [Online]. Available: http://www.opencores.org/

[39] A. B. Kahng, C. Park, and X. Xu, "Fast dual graph-based hotspot detection," in Proc. SPIE, vol. 6349, 2006, p. 63490H.

[40] M. Cho, K. Yuan, Y. Ban, and D. Z. Pan, "Eliad: efficient lithography aware detailed router with compact post-opc printability prediction," in Proc. Design Automation Conference DAC'08, 2008, pp. 504–509.

[41] K. Jeong, A. B. Kahng, and K. Samadi, "Quantified impacts of guardband reduction on design process outcomes," in Proc. Intl. Symp. Quality Electronic Design ISQED'08, 2008, pp. 790–797.
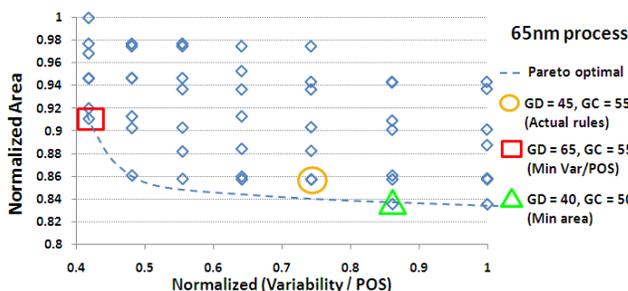
Figure 16: Co-exploration of GC and GD rules (see figure 2) for 65nm commercial process.