Lens Aberration Aware Placement for Timing Yield

ANDREW B. KAHNG and CHUL-HONG PARK
University of California at San Diego
PUNEET SHARMA
Freescale Semiconductor, Inc.
and
QINKE WANG
Magma Design Automation, Inc.

Process variations due to lens aberrations are to a large extent systematic, and can be modeled for purposes of analyses and optimizations in the design phase. Traditionally, variations induced by lens aberrations have been considered random due to their small extent. However, as process margins reduce, and as improvements in reticle enhancement techniques control variations due to other sources with increased efficacy, lens aberration-induced variations gain importance. For example, our experiments indicate that delays of most cells in the Artisan TSMC 90nm library are affected by 2–8% due to lens aberration. Aberration-induced variations are systematic and depend on the location in the lens field. In this article, we first propose an aberration-aware timing analysis flow that accounts for aberration-induced cell delay variations. We then propose an aberration-aware timing-driven analytical placement approach that utilizes the predictable slow and fast regions created on the chip due to aberration to improve cycle time. We study the dependence of our improvement on chip size, as well as use of the technique along with field blading which allows partial reticle exposure. We evaluate our technique on two testcases, *AES* and *JPEG* implemented in 90*nm* technology. The proposed technique reduces cycle time by 4.322% (80ps) at the cost of 1.587% increase in trial-routed wirelength for AES. On JPEG, we observe a cycle time reduction of 5.182% (132ps) at the cost of 1.095% increase in trial-routed wirelength.

Categories and Subject Descriptors: B.7.2 [**Lutegrated Circuits**]: Design Aids—*Layout*

General Terms: Algorithms, Design, Performance, Verification

Additional Key Words and Phrases: Layout, lithography, design for manufacturing, timing yield

16

1. INTRODUCTION

Low k-factor lithography drives many new process-design interactions that must be comprehended early in the development process to ensure rapid yield ramp-up and acceptable steady-state yield. Modern lithography tools can image a complex chip pattern with billions of pixels, within an exposure time of a fraction of a second. However, all optical projection systems used for microlithography depart from perfection because of various lens aberrations, especially when large image field size is combined with high numerical aperture (NA).

Aberrations [Levinson 2001] can be described as the departure from ideal imaging induced by an imperfect lens system, as shown in Figure 1. Aberrations cause optical path differences among the rays, resulting in wavefront deviation from a reference sphere at the exit pupil; this induces blur and distortion of images. Undesirable imaging artifacts from aberration are uncorrectable and, indeed, are sometimes exacerbated through use of resolution enhancement techniques (RETs) such as phase-shift mask and off-axis illumination [Brunner 1997]. The effects of lens aberrations on lithographic imaging [Gortych and Williamson 1991; Toh and Neureuther 1987] include shifts in the image position, image asymmetry, reduction of the process window, and the appearance of undesirable imaging artifacts. *Zernike's coefficients* capture the deviation from ideal imaging and may be used during lithography simulation to predict the impact of lens aberration on critical dimension (CD) [Levinson 2001; Progler and Wong 2000]. CD variation caused by lens aberration is relatively small compared to that caused by defocus and pattern proximity. However, most CD error caused by proximity can be corrected by RETs. Thus, lens aberration has turned out to be a major source of residual errors in across-field linewidth variation (AFLV) [Flagello et al. 1999].

Recent studies of lens aberration control have focused on measurement systems [Shiode et al. 2002; Farrar et al. 2001] and pattern sensitivity of aberration [Wong 2002], as well as lens mounting systems to compensate for the aberration [Matsuyama et al. 2002]. However, despite these efforts, the impact of lens aberration on CD will be an ever-present barrier to manufacturing yield as minimum design rules are pushed ever closer to fundamental resolution limits. From the design perspective, variations in CD affect the delays, slews, input capacitances and leakage of a given logic cell. We also observe that the maximum difference in delays of all timing arcs in a cell (delay skew) increases significantly with lens aberration, as different MOS devices in the layout are affected differently by aberration.

Progler et al. [2004] studied the impact of lens aberration on statistical timing behavior and observed that certain aberration coefficients are associated with large timing error. Orshansky et al. [1999] found that spatial gate CD variation
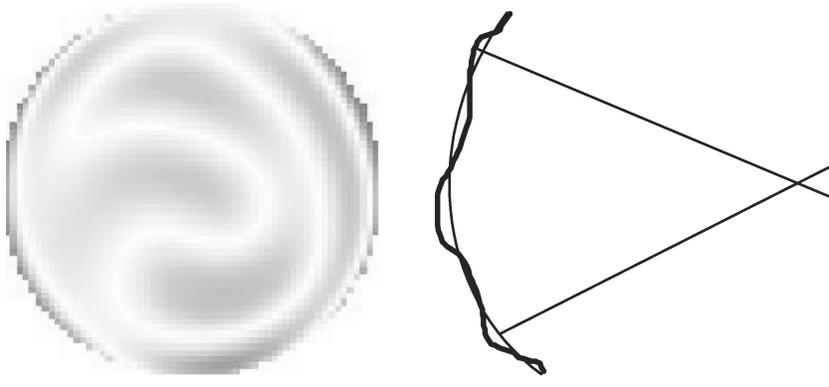
Fig. 1.   An imperfect lens system.

leads to a large variation in the raw speed of CMOS logic. Misleading timing results are obtained, which lead to slower and/or malfunctioning circuits because the simulation of a circuit's behavior has ignored the spatial CD information. The systematic variability of gate CD caused by lens aberration can be modeled in order to achieve better performance by way of accurate timing analysis at all stages of physical implementation [Orshansky et al. 2002, 2004]. However, more accurate analysis of gate delay impact is required as the scaling of lithographic features makes the impacts of lens aberrations even more complex.

In this paper, we use lens aberration-aware global placement for timing improvement. It is worth discussing why we use global placement, as opposed to OPC or detailed placement (for example), as the appropriate "knob" for this compensation and optimization. First, lens aberration can slightly change according to lens heating and lens contamination. OPC is very sensitive to changes of lens aberration because it embodies direct perturbations of mask shape; on the other hand, placement does not directly change any mask shape, but only rearranges cell instances according to fast and slow regions in a lens field. Thus, a placement-based solution is less sensitive to variations of aberration parameters than an OPC-based solution. Second, lens aberration globally changes cell characteristics within a lens field. That is, cells at the range of a few microns may use the same Zernike's coefficients, while cells at the range of a millimeter must use different Zernike's coefficients. On the other hand, detailed placement is effective for compensation of micron-ranged variability (i.e., proximity effects of resist and photo processes). Gupta et al. [2007] proposed a detailed placement technique to avoid forbidden pitch between cells, which is an optimization on the length scale of $0.5 \sim 2 \ \mu$m. The detailed placement approach may also increase wirelength significantly, and thus has no clear advantage (in terms of convergent flow, etc.) over global placement. We believe that the use of a global placer to minimize total wirelength with respect to global bins can more efficiently handle the lens aberration problem, as compared to a detailed placement approach.

In the following, we first describe a novel aberration-aware static timing analysis flow that integrates (i) results of lithography simulation to measure

CD across the lens field, (ii) SPICE simulation-based library performance characterization that captures variant CD combinations in library cell instances, and (iii) placement information. We also propose an aberration-aware timing-driven analytical placement framework that utilizes the aberration-aware timing analysis flow to minimize clock cycle time and avoid hold-time violations, without significantly increasing total wirelength. The placer is driven by models that capture the impact of lens position on timing arc delays in cells, and by weighted-wirelength models. Essentially, we preferentially place cells that are setup-time (resp. hold-time) critical at lens field locations where aberrations cause the cell delay to decrease (resp. increase).

The contributions of our work are as follows.

—Using industry OPC recipes, aberration parameters, and design testcases, we show that the variation in timing due to lens aberration can be significant. Over the cells in a 90nm foundry library, we observe cell delay (averaged over all timing arcs) to change by 2%–8%. The maximum difference in delays over all timing arcs of a cell (*delay skew*) increases significantly.

—We develop a novel aberration-aware timing analysis flow that affords more accurate timing analysis, taking into account the position of the chip in the lens field. It also considers the increase in delay skew caused by aberration.

—We propose a novel aberration-aware, timing-driven analytical placer that considers the impact of lens aberrations on timing to minimize clock period and avoid hold-time violations without significant total wirelength increase. Averaged over our two testcases, worst-case cycle time and total negative slack respectively reduce by $\sim$4.749% (116ps) and $\sim$7.535% at the cost of $\sim$1.341% increase in wirelength, with no hold-time violations—a very substantial performance improvement.

The remainder of this article is organized as follows. In Section 2, we describe lens aberration and study its impact on CD and gate delay. Section 3 proposes a novel aberration-aware timing analysis and an accompanying flow. Section 4 describes our aberration-aware analytical placement formulation and implementation details. Test designs, experimental conditions and experimental results are described in Section 5. We conclude in Section 6 with directions for ongoing research.

2. DESIGN IMPACT OF LENS ABERRATION

In this section we briefly describe how lens aberration impacts CD and consequently circuit delay.

2.1 CD Impact of Lens Aberration

Several manufacturing process steps are involved in the transfer of the pattern on the mask to the photoresist, and then to the wafer. Lens aberration comes into play when the photoresist is exposed to light during lithography. Broadly speaking, a lithography setup includes one or more illumination sources, a mask, several lenses, and photoresist applied to the wafer. Modern lithography systems use step-and-scan to expose small portions of the wafer at a time, and
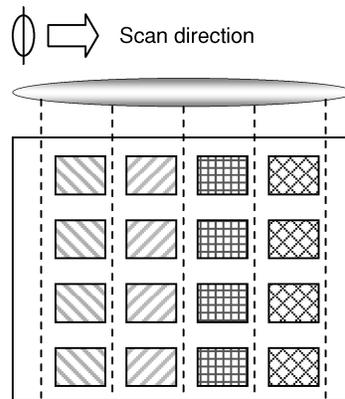
Fig. 2.   Different CD qualities of chips in a reticle due to aberration across the lens field.

then shift to the next region. The portion of the wafer that gets exposed in a step is called the *lens field*, or simply *field*. In each step, the photoresist is exposed to light through a slit that is scanned from one side of the field to the other [Wong 2001].

Lens aberration parameters (Zernike's coefficients), which capture the divergence from ideal behavior of light, change as the slit translates horizontally. Hence, the CD error induced by lens aberration varies along the horizontal direction but stays constant along the vertical direction. While the variation in CD along the horizontal direction is continuous, it is reasonable to discretize it and assume it to remain constant over small regions as shown in Figure 2. Based on industry-supplied Zernike's coefficients at multiple locations in the lens field, we run lithography simulation on some frequently-used standard cells from a 90nm foundry library, and study the impact on CD. Figure 3 shows average CD variation of devices in BUFX4, INVX2, NAND2X4 and NOR2X1 cell instances as their position within the lens field is varied. For example, average gate CD variation of NAND2X4 at 100nm worst defocus is up to 8nm across the entire lens field. In addition, we investigate the *CD skew* (maximum difference in CD over all devices in a cell) of different cells. Large CD skew can unbalance the timing arcs of a cell, as we discuss in greater detail in Section 3. Figure 4 shows the CD skew for NAND2X4 as its position in the lens field is changed. It is evident from these studies that the aberration impact on CD error is large across the lens field, and must be modeled to reduce guardbanding and overdesign.

2.2 Delay Impact of Lens Aberration

Variations in CD directly and indirectly affect circuit delay. At the device level, increase in gate CD causes an approximately linear decrease in saturation on-current of the device, which partially determines delay. Since lens aberration affects different devices in a cell differently, each of the cell's timing arcs can be affected differently. Most standard cells are designed such that the maximum difference in delays of timing arcs (*delay skew*) is small. Due to lens aberration,
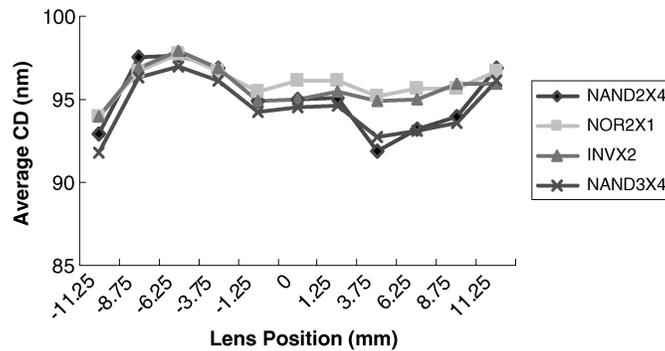
Fig. 3.   Average gate CD varies across the lens field; the range of this variation for the NAND2X4 cell is 8nm.
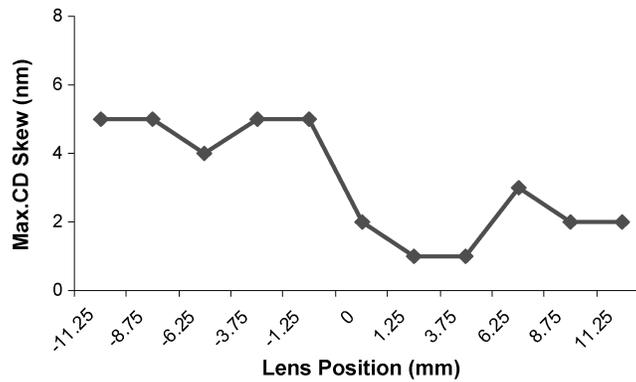
Fig. 4.   Maximum CD skew among all gates in NAND2X4 cell.

however, this delay skew can increase—that is, arcs that are governed by larger-than-nominal CDs will be slowed down, while those governed by smaller-than-nominal CDs will be sped up. Figure 5 shows how the delay, averaged over all timing arcs, changes for four cell masters as the cell instance location is varied from the lens center. Figure 6 shows the aberration-induced increase in delay skew with respect to the delay skew of the nominal (or drawn) cell as the location of cell NAND2X4 is varied in the field.[1]

CD variations also cause variations in cell input capacitance and output slews (transition times). Input capacitance affects the loading of fanin cells and consequently their delays; interconnect delays are also affected. Similarly, slews affect the output slews and delays of cells in the fanout cone. Again, to avoid unnecessary guardbanding, the performance analysis flow (library model characterization, timing/SI analysis, etc.) must comprehend these systematic variations.

---

[1]In the figure, the increase is always over 40% because in computing nominal delay skew, library characterization applies an equal CD error to all devices at worst-case process conditions. To compute aberration-induced delay skew, however, lithography simulation is performed at the worst-case process corner and all devices get different CD errors.
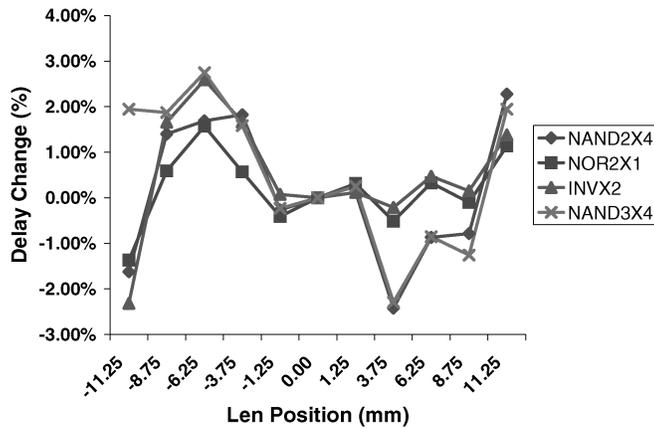
Fig. 5.   Change in average delay with lens position, with respect to the center of the lens.
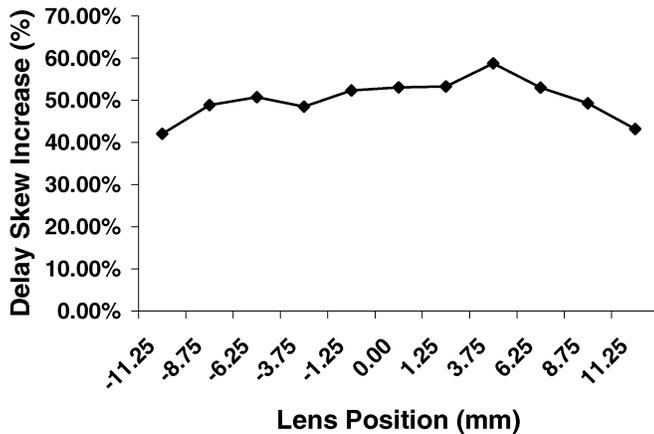
Fig. 6.   Percentage increase in delay skew (maximum difference in delays of all timing arcs) of the NAND2X4 cell, relative to the maximum delay skew of nominal (or drawn) cell, as lens position is changed. In the figure, the increase is always over 40% because in computing nominal delay skew, library characterization applies an equal CD error to all devices at worst-case process conditions. To compute aberration-induced delay skew, however, lithography simulation is performed at the worst-case process corner and all devices get different CD errors.

3. ABERRATION-AWARE TIMING ANALYSIS

In this section we describe our aberration-aware timing analysis flow. While the flow is complete and self-contained, it is at the same time designed for, and will be used by, the analytical placement framework described in Section 4. Our aberration-aware timing analysis flow involves two main steps: (1) constructing timing libraries of all standard cells for different locations in the lens field; and (2) using placement information of the design to compute the location of all cell instances in the lens field, then using this location information to look up appropriate models in the timing library for use with off-the-shelf static timing analysis (STA) tools.
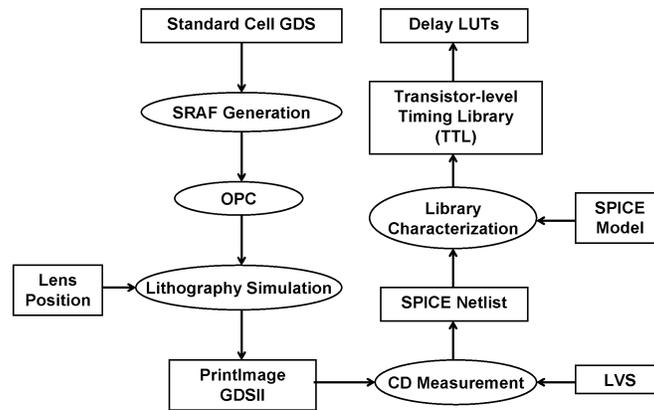
Fig. 7.   Aberration-aware timing analysis and its flow.

Before describing our analysis flow, we describe two alternative flows and our reasons for not using them. In the first alternative flow, variants of each cell are created such that the CD of all devices in the cell is different for each variant, but the same for all devices in a given variant. A timing library can be created using SPICE models for all the variants. Since all devices in a cell variant have the same CD, we call this library a *cell-level* granularity library. To perform timing analysis on a placed design, lithography simulation is performed to obtain CDs of all devices in all cells. For each cell, the CDs of its devices can be averaged, and the closest-matching available cell variant in the timing library then fed to off-the-shelf STA. However, as CD skews can be large, averaging of device CDs can introduce inaccuracy in the estimated impact of aberration. In other words, the effect of nonuniform CDs is non-uniformity in timing arc delays, rather than average increase or decrease in the delays of all timing arcs. Our experiments have found that the cell-level library-based approach is very inaccurate compared to the approach that we adopt.

The second alternative flow creates a priori variants for each cell master, such that there is one variant for every possible assignment of CDs to devices. This means that given any assignment of CDs to devices, an exactly matching, precharacterized cell variant can be found. After lithography simulation provides CDs of all devices in all cells, a correctly matching variant can be picked for use in timing analysis. Though this flow is very accurate, it requires a very large number of cell variants (exponential in the number of devices in the cell); this is infeasible with respect to both characterization time and library size.

In our proposed flow, variants are created for each cell for different lens field locations. Figure 7 illustrates our timing library construction flow. We begin with standard-cell GDSIIs and use Mentor Graphics *Calibre* (v 9.3_5.11).[2] for subresolution assist feature (SRAF) generation and model-based OPC. We use Zernike's coefficients for eight sampling positions in the lens field (data provided by a major chip maker), and compute the other coefficients at 19 different locations with 1.5mm stepsize on the field using linear interpolation.

---

[2]http://www.mentor.com.
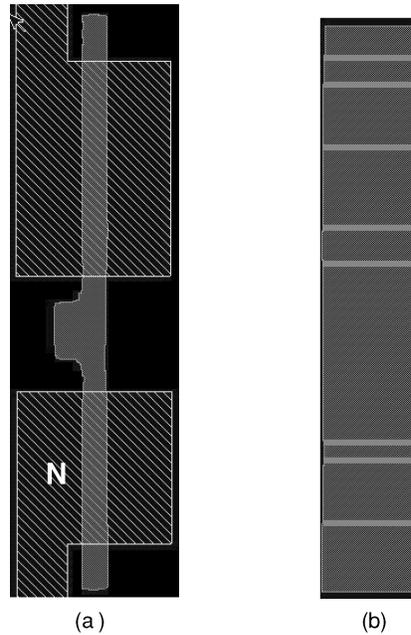
Fig. 8. Polygon generation for CD measurement: (a) result of PrintImage simulation of an inverter, and (b) rectilinearized polygon representation of a gate device in the region N of (a).

Using the post-OPC standard-cell GDSIIs and Zernike's coefficients, we perform lithography simulation at 19 different field locations with wavelength λ = 193nm, numerical aperture NA = 0.75, and annular aperture σ = 0.75/0.50. After lithography simulation, we have 19 PrintImage GDSII results for each standard cell; we then measure the CD of each of the MOS devices in each GDSII result.

Figure 8(a) shows the PrintImage contour generated by *Mentor Graphics PrintImage v9.3_5.11* for one device.[3] To measure the CD of the PrintImage contours, we first take an intersection with the active layer to obtain the contour of the gate. Contours are rectilinearized and split into rectangles in a staircasing fashion. The lengths of all rectangles are then averaged with rectangle widths as weights to compute the CD of the gate (i.e., $CD_{gate} = \sum^n l_i \times w_i / \sum^n w_i$ where n is the number of rectangles into which the contour is split, and l_i and w_i are the length and width of the ith rectangle).

The measured CDs are then used to alter SPICE netlists of standard cells, preparatory to running library characterization. A complication arises because GDSII typically does not have device names, while SPICE netlists only reference devices by device names. We solve this problem by applying LVS (layout vs. schematic) to obtain a mapping between device locations and device names.

---

[3]Mentor Graphics PrintImage produces rectilinear contours; our approach, however, is generic enough to be used for arbitrary polygonal contours.

After modifying the SPICE netlists, we run Cadence SignalStorm (v 4.1)[4] to perform library characterization. Since lens aberrations affect different devices in a cell differently, the altered SPICE netlists may no longer have equal CD for all devices. We call our characterized library a *transistor-level timing library* (TTL); it accurately captures the delay skew induced by CD skew while incurring manageable added complexity of characterization effort and library size. The choice of the number of field locations to use depends on the extent and rate of change of aberration-induced CD. A larger number of field locations improves the accuracy but also increases the number of cell variants in the cell library.

Our test library contains 50 combinational cells. For each we create 19 variants corresponding to 19 field locations. Library characterization requires approximately 6 hours (wall time) running on 18 CPUs ranging from Intel *Xeon* 1.4GHz to AMD *Opteron* 2.2GHz. We do not create variants for the 13 sequential cells in our library due to large CPU time (estimated at 60 hours on our machines) required by their characterization. We note that while the characterization time can be significant, it is a one-time task for each process.

4. ABERRATION-AWARE TIMING-DRIVEN PLACEMENT

Because of lens aberrations, a cell placed at different locations within the reticle will exhibit varying performance characteristics. In order to improve timing yield after manufacturing, we propose a lens aberration aware timing-driven placement formulation that minimizes total timing-weighted delays of cells in conjunction with common timing-driven placement objectives such as minimizing total timing-weighted wirelength. We implement our method based on a general analytical placement framework and describe implementation details in this section.

4.1 Introduction of Analytical Placement

Analytical placement methods have recently received increased attention from both academia and industry [Eisenmann and Johannes 1998; Etawil et al. 1999; Hu and Marek-Sadowska 2002; Kahng and Wang 2005; Naylor et al. 2001; Viswanathan and Chu 2004]. Specifically, recent work implements *APlace*, a general analytic placement framework [Kahng and Wang 2004b, 2005; Kahng et al. 2005a, 2005b], which has high solution quality and strong extensibility. Here we briefly introduce the APlace analytic placement framework, upon which we build our proposed aberration-aware timing-driven placement method.

APlace formulates global placement as a *constrained nonlinear optimization problem*: the layout area is uniformly divided into global bins and APlace minimizes total half-perimeter wirelength (HPWL) while maintaining equalized cell area in each global bin (i.e., uniform density). A formal problem formulation is as follows:

$$
\begin{aligned}
&min \ \ HPWL(\mathbf{x}, \mathbf{y})\\
&s.t. \ \ \ D_g(\mathbf{x}, \mathbf{y}) = D \ \ \text{for each global cell } g
\end{aligned}
\tag{1}
$$

---

[4]http://www.cadence.com.

where $(\mathbf{x}, \mathbf{y})$ is the vector of center coordinates of cells, $HPWL(\mathbf{x}, \mathbf{y})$ is the total HPWL of the current placement, $D_g(\mathbf{x}, \mathbf{y})$ is a density function that equals the total cell area in a global bin g, and D is the average cell area over all global bins.

APlace applies smooth approximations of the HPWL and density functions and solves the constrained optimization problem in Equation (1) using the simple *quadratic penalty method*. For example, the placer solves a sequence of unconstrained minimization problems of the form

$$min \ \ HPWL(\mathbf{x}, \mathbf{y}) + \frac{1}{2\mu} \sum_g (D_g(\mathbf{x}, \mathbf{y}) - D)^2 \qquad (2)$$

for a sequence of values $\mu = \mu_k \to 0$, with the solution of each unconstrained problem being used as an initial guess for the next one. A *Conjugate Gradient* (CG) solver is employed to optimize the objective function in Equation (2). The conjugate gradient method is quite useful in finding an unconstrained minimum of a high-dimensional function. Also, the memory required is only linear in the problem size, which makes the approach adaptable to large-scale placement problems.

The general APlace framework has been extended to address a variety of placement tasks across many aspects of physical implementation, including mixed-size placement, timing-driven placement, power-aware placement, voltage-drop aware placement and I/O-core co-placement; it has been shown to be competitive in a wide variety of contexts [Cheon et al. 2005; Kahng et al. 2005; Kahng and Wang 2005].

4.2 Aberration-Aware Placement Formulation

We now propose a novel aberration-aware timing-driven placement objective for improved timing yield after manufacturing, and describe its integration into the analytical placement framework. We perform aberration-aware timing-driven placement by optimizing a hybrid placement objective. Besides the typical objective of minimizing total timing-weighted net wirelength, we also minimize the sum of timing-weighted delays of timing-critical cells. The aberration-aware timing-driven placement formulation is as follows:

$$
\begin{aligned}
min \ \ & WWL(\mathbf{x}, \mathbf{y}) + W_a \sum_v w(v) \cdot g_{t_v}(x_v) \\
s.t. \ \ & D_g(\mathbf{x}, \mathbf{y}) = D \text{ for each global bin } g \\
& \text{and } g_{t_v}(x_v) = MAX\{g^1 t_v(x_v), \dots, g^n t_v(x_v)\}
\end{aligned}
\qquad (3)
$$

where $WWL(\mathbf{x}, \mathbf{y})$ is the sum of timing-weighted net HPWL of the current placement and W_a is the weight for the aberration-aware timing-driven objective function terms, which is the sum of timing-weighted delays of timing-critical cells.[5] In the formulation, $g_{t_v}(x_v)$ is the delay function, obtained from the TTL

---

[5]We divide the objective into two parts since we consider the aberration-induced variation in only the cell delay. Aberration-induced CD variation of the gates affects timing yield, while aberration-induced CD variation of wires may be neglected in comparison to the impact of HPWL in wire delay. Note that larger CD of a wire increases the capacitance, but decreases the resistance, and vice versa.

timing library described above, for cell instance v's timing model t_v; it is a function of v's horizontal position x_v in the chip. In the situation where there are multiple copies ($n > 1$) of chips in the reticle, we let $g^i t_v(x_v)$ be the delay function for the ith chip,[6] and we consider the maximum delay of cell instance v over all copies so that the performance of the slowest chips is improved. We note that this is a pessimistic approximation of cells' delays, since not all timing-critical cells may exhibit their maximum delays on the same chip copy. However, we do not consider this pessimism to be significant, since the impact of aberration on delays of all cells is similar, and a chip copy that has large delay for one cell likely has large delays for other cells as well. For example, cells except INV1 and INV2 (which have only one isolated line) have similar cell delay behavior, as we saw previously in Figure 5. This is because all cells share similar parameters of pitch, width and design rules, and because linewidth variation due to lens aberration is not a function of cell type, but rather a function of pattern geometry. We thus believe that our delay upper-bounding is not significantly pessimistic.

Our problem formulation applies to the single lens as well as multiple chips on a wafer. A modern fab may employ multiple lithography lens systems. For high-volume, cutting-edge designs such as microprocessors, it is already common practice to have stepper-specific masks. Stepper-specific masks are tuned according to the stepper "signature" as part of the RET/mask data preparation flow. Our methodology brings aberrations upstream in the design and is easily adoptable when stepper-specific masks are used. Further, recent studies of lens aberration enable quick measurement of Zernike's parameters to capture lens aberrations [Farrar et al. 2001; Shiode et al. 2002]. Even when identical masks are used on multiple steppers, it is preferable and common practice to use steppers from the same manufacturer to reduce stepper-to-stepper variations. Steppers from the same manufacturer have very similar aberrations and our methodology can use the Zernike's coefficients from any one stepper to optimize the design. It is also possible to extract the systematic aberration components from a database of aberration measurements of various lithography systems, and then generate a lens aberration map incorporating "universal" Zernike's parameters which can be applied to our aberration-aware placement flow. All of these scenarios leverage the basic design optimization that is proposed in this section.

As with traditional net weighting methods, we assign timing weights to cells based on timing criticality and path sharing. First, a cell along a timing-critical path should receive a heavy weight. Second, a cell with many timing-critical paths passing through should have a large weight as well. Therefore, we assign to cell v the weight $w(v)$, given as,

$$w(v) = \sum_{v \in \pi} (D_s(slack_s(\pi), T_s) \cdot D_h(slack_h(\pi), T_h) - 1), \tag{4}$$

---

[6]The critical-path delay of the ith copy of the chip depends on the horizontal position of that copy in the reticle. We assume that the chip size can be determined using any initial placement optimization, and that the horizontal position of a copy of the chip can then be obtained from a reticle floorplan. Our aberration-aware placement optimization thus incorporates the chip size and the reticle floorplan.

where

$$D_s(slack_s(\pi), T_s) = \begin{cases} (1 - s/T)^\delta & s \leq 0 \\ 1 & s \geq 0 \end{cases} \tag{5}$$

and

$$D_h(slack_h(\pi), T_h) = \begin{cases} (1 + s/T)^\delta & s \leq 0 \\ 1 & s \geq 0. \end{cases} \tag{6}$$

Here, δ is the criticality exponent, and u is the expected improvement of the longest (or shortest) path delay after this timing-driven iteration. T is $T_s = (1 - u) \cdot \max_\pi \{delay(\pi)\}$ for setup-critical paths or $T_h = (1 + u) \cdot \min_\pi \{delay(\pi)\}$ for hold-critical paths. Additionally, $slack_s(\pi) = T_s - delay(\pi)$ is the slack of a setup-critical path π, while $slack_h(\pi) = delay(\pi) - T_h$ is the slack of a hold-critical path π. In Equation (4), we compute a weight for each timing-critical path based on its slack, and obtain the timing weight of a cell by summing up the weights of timing-critical paths passing through it.

For timing-driven edge weights, existing approaches can be broadly divided into two classes, *path-based* and *net-based*. The path-based approach is based on mathematical programming techniques, and can maintain an accurate timing view during optimization. But, its drawback is relative high complexity. We use a net-weighting based approach which assigns weight to nets based on their timing criticality [Marquardt et al. 2000; Kong 2002; Kahng and Wang 2004a].[7] The basic idea is that a timing-critical net should receive a heavy weight, and an edge with many paths passing through it should have a heavy weight as well. We thus assign to edge e the weight $w(e)$, given as

$$w(e) = 1 + \sum_{e \in \pi}(D_s(slack_s(\pi), T_s) \cdot D_h(slack_h(\pi), T_h) - 1) \tag{7}$$

where $D_s(slack_s(\pi), T_s)$ and $D_h(slack_h(\pi), T_h)$ have the same formulations as in Equations (5) and (6). Note that the balance of timing weights between wire and cell is determined by the weight W_a for consideration of the aberration-aware timing-driven objective. Note also that the constant 1 in Equation (7) means that the weight for nontiming critical nets is 1 (whereas timing-critical nets will have a weight > 1). Equation (4) does not require the constant 1 because the total wirelength of nontiming critical nets is optimized at the same time besides the timing-related objectives.

4.3 Placement Flow

Our aberration-aware timing-driven placement and evaluation flow is shown in Figure 9. In addition to the design netlist, we also inputs the delay functions of cell models, which represent how the delays of given cell models change with their horizontal position in the chip.

The timing-driven process in our placer may include several iterations. As shown in Figure 9, during each iteration, we send the intermediate placement

---

[7]Note that in the timing analysis step, we use commercial extraction tools for accurate wire delay estimation.
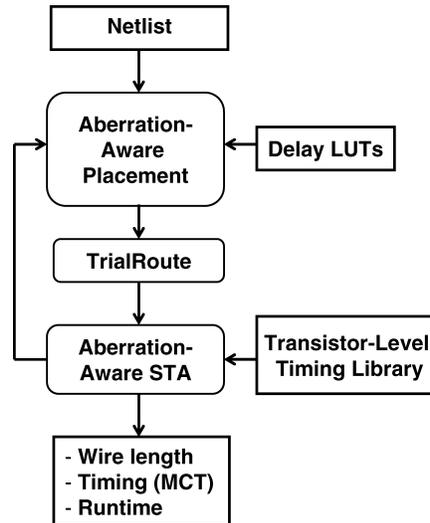
Fig. 9.   Aberration-aware timing-driven placement and evaluation flow.

to *TrialRoute* (Cadence *SOC Encounter* v 2004.10) to perform a fast global and detailed routing, and extract RC parasitics.[8] We then change the type of each cell in the netlist according to its horizontal position within the lens field and use Synopsys *PrimeTime* (v W-2004.12-SP2)[9] to perform accurate aberration-aware static timing analysis (STA) with the transistor-level timing libraries (TTLs) described in Section 3. The resulting critical paths are imported into the placer to decide timing weights for nets and cells. The total timing-weighted cell delay is then minimized using the Conjugate Gradient solver, together with the timing-weighted wirelength objective, and subject to density constraints.

4.4 Implementation Details

As mentioned above, for each master cell, we create 19 different variants according to 19 lens field locations. Through the recticle floorplan, we can extract the position of the ith chip in the field and (in the timing analysis) instantiate timing model variants corresponding to the actual position of each instance of the given master cell. Thus, there is no need to create variants for different copies. In Equation (3), $g^i t_v(x_v)$ is the delay function of the ith chip, which is generated using (interpolation of) the position—specific delay model variants.

We compute the weight of the aberration-aware objective W_a in Equation (3) according to the x-gradients derived from the wirelength and delay terms, so that the scaled gradients of delay functions are comparable to the wirelength

---

[8]Separately, we have verified that TrialRoute results give the same conclusions as final detailed routing results. We use TrialRoute because of runtime constraints for our large number of experiments.

[9]http://www.synopsis.com.
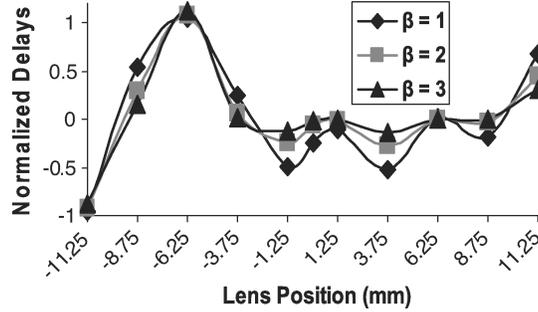
Fig. 10. Delay curves of NOR2X1 with a variety of smoothing factors (β's).

gradients, that is,

$$W_a = \alpha \cdot \left(\sum_v \left| \frac{\partial WWL}{\partial x_v} \right| \right) \bigg/ \left(\sum_v \left| \frac{\partial g_{t_v}}{\partial x_v} \right| \right) \qquad (8)$$

The delay ratio α decides the ratio of the delay gradients to the wirelength gradients, and must be carefully tuned according to the impact of reduced cell delay and increased net wirelength on design performance.

We derive the delay of a cell at a specific horizontal field position by averaging the rise and fall delays of all timing arcs with zero wire load, according to the transistor-level timing libraries. Thus, the delay functions represent how gate delays vary with horizontal locations and gate CDs. Due to simulation limits, delay functions have accurate values only at discrete horizontal coordinates, and consequently are expressed as look-up tables (LUTs). We obtain delay at continuous positions using linear interpolation and compute gradients accordingly.

A smoothing technique [Gu and Huang 1994] can be applied to smooth the delay curves. To reduce the effect of local minima, we use a local search method with search space smoothing technique. The smoothing technique transforms the given problem into a series of problem instances with different terrain structures. Initially, a simplified instance with a smooth terrain surface is solved using the local search algorithm [Gu and Huang 1994]. Then, the solution of the problem instance is then taken as the initial solution for the next problem instance that has a slightly more complicated search space. The problem is again solved using the same algorithm. The above procedure is repeated until the final problem instance having the original search space is solved. Given a normalized delay function, a smooth function is a pre-defined smoothing factor $\beta \geq 1$ as follows:

$$g' = \begin{cases} \overline{g} + (g - \overline{g})^\beta & if \quad g \geq \overline{g} \\ \overline{g} - (\overline{g} - g)^\beta & if \quad g \leq \overline{g} \end{cases} \qquad (9)$$

where $\overline{g}$ is the average value of the delay function. Figure 10 shows delay curves with a variety of smoothing factors β for NOR2X1. A delay function generated from a larger β exhibits a smoother curve, while a delay function generated from a smaller β exhibits a more rugged curve.

Table I. Design Characteristics of Two Benchmark Circuits

Design	Utilization (%)	Chip Size (mm)	#Cells	#Nets
AES	60	0.50	17304	17465
JPEG	60	1.41	118321	125036

5. EXPERIMENTS

In this section, we empirically test our aberration-aware placement approach on two designs within a standard design flow using commercial design automation tools. We assess the impact on timing, wirelength, and runtime.

Experimental setup. We use two designs from OpenCores[10] as our test cases. The circuits are synthesized using Synopsys *Design Compiler* (v W-2004.12-SP3) with tight timing constraints and a set of 63 most commonly used standard cells (50 combinational, 13 sequential) from Artisan TSMC 90nm library, then floorplanned in Cadence *SOC Encounter* (v 2004.10). The design characteristics are summarized in Table I. The experimental flow is shown in Figure 9. The inputs for each design include synthesized netlists, floorplan, timing constraints, aberration-aware timing libraries, delay look-up tables derived from the libraries for convenience of the placer, and physical libraries in LEF format. The placer executes iteratively with STA to improve and converge on timing.

We evaluate the following three timing-driven placers.

—*TradPl_TD*: Analytical timing-driven placer, APlace, with the traditional (or standard) STA during the placement optimization. This is the traditional timing-driven analytical placer.

—*APlace_TD*: Timing-driven APlace with aberration-aware STA. Aberration-aware STA accounts for aberration-induced cell delay changes, and therefore computes more accurate timing slacks which are used in the timing-driven placer objective function.

—*AberrPl_TD*: Aberration-aware timing-driven placer, with timing-driven wirelength and aberration objectives, and aberration-aware STA. This improves upon *APlace_TD* by explicitly accounting for aberration-induced cell delay changes in the placement objective function.

We use aberration-aware STA to compare the three placers for circuit delay.

We expect larger chips to benefit more from our aberration-aware placement technique since they will have larger CD and delay variation induced by an imperfect lens system across the layout region. However, our testcases are not sufficiently large to witness the effect of lens aberration that may be observed in real-world systems on chip. Hence, in our studies we scale the aberration map, which captures the impact of aberration at every chip location, along the horizontal direction to mimic the aberration that is observed in larger modern designs.

We perform three sets of experiments to evaluate the performance improvement under different die size and field size scenarios: (1) when there is only

---

[10]http://www.opencores.org/projects/.
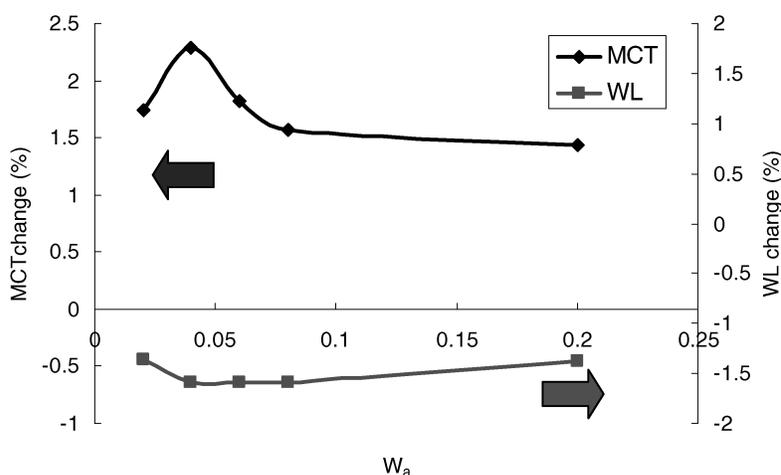
Fig. 11. MCT change of AberrPl_TD according to the weight of the aberration-aware objective W_a for testcase AES.

one copy of the chip in the lens field; (2) when there are multiple copies (the number of which is determined by a scaling factor, with a variety of scaling factors); and (3) when field blading is performed for partial reticle exposure. We compute timing weights with criticality exponent $\delta = 4$ and expected improvement $u = 10\%$. Note that we only perform APlace_TD and AberrPl_TD for experiments (2) and (3) since the result of TradPl_TD is always the same as the result of (1). Figure 11 shows the minimum cycle time (MCT) change of AberrPl_TD according to the weight of the aberration-aware objective W_a for testcase AES. With $W_a = 0.04$, MCT and trial-routed wirelength are optimized, and we use this value of W_a in our experiments. In general, MCT improvement results in increase of wirelength.

After each placement, we perform global and detailed routing, RC extraction, and finally aberration-aware timing analysis using Synopsys *PrimeTime*. MCT of the slowest chip in the reticle is reported by aberration-aware STA to measure performance of timing-driven placements. We also report HPWL and runtime for placement, and routed wirelength and the number of viasafter routing. All experiments are conducted on Linux machines with 2.4GHz CPU and 4GB memory.

Experimental results. Table II summarizes the results of TradPl_TD, APlace_TD, and AberrPl_TD on our two test cases, AES and JPEG, when there is one die in a reticle. In comparison to TradPl_TD, APlace_TD reduces MCT by 2.585% (48ps) with 0.892% HPWL increase and 0.307% increase of trial-routed wirelength for AES, and reduces MCT by 1.289% (38ps) with 0.687% HPWL increase and 0.7% increase oftrial-routed wirelength for JPEG. Our aberration-aware placer (AberrPl_TD), in comparison to traditional timing-driven placement (TradPl_TD), reduces MCT by 5.667% (105ps) with 1.909% HPWL increase and 1.902% increase of trial-routed wirelength for AES, and reduces MCT by 5.13% (150ps) with 1.673% HPWL increase and

Table II. Comparison of Traditional Timing Driven Placement (*TradPl_TD*) Versus APlace_TD
Placement or AberrPl_TD Placement for AES and JPEG

Design	Method	Place		TrialRoute		STA	
		HPWL **(e9 um)**	**CPU** **(s)**	**WL** **(e5 um)**	**#Vias**	**MCT** **(ns)**	**TNS** **(ns)**
AES	TradPl_TD	1.1699	1432	6.521	1.2521	1.8491	156.3829
	APlace_TD	1.1803	1457	6.541	1.2531	1.8013	150.8231
	Impr. (%)	**−0.8919**	**−1.7458**	**−0.3067**	**−0.0743**	**2.5850**	**3.5525**
	AberrPl_TD	1.1922	1471	6.645	1.2542	1.7443	144.9321
	Impr. (%)	**−1.9090**	**−2.7235**	**−1.9016**	**−0.1629**	**5.6668**	**7.3223**
JPEG	TradPl_TD	6.2980	23598	3.717	6.1762	2.9252	213.4321
	APlace_TD	6.3312	23791	3.743	6.1874	2.8875	206.3124
	Impr. (%)	**−0.6871**	**−0.8179**	**−0.6995**	**−0.1809**	**1.2879**	**3.3357**
	AberrPl_TD	6.3932	24139	3.780	6.1938	2.7751	196.8943
	Impr. (%)	**−1.6731**	**−2.2926**	**−1.6949**	**−0.2846**	**5.1296**	**7.7484**
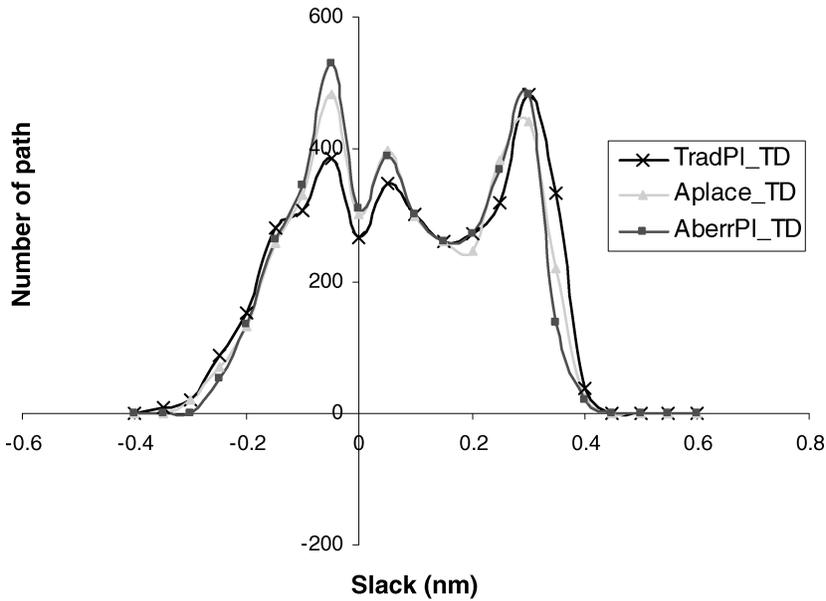
Fig. 12.   Slack distributions of TradPl_TD, APlace_TD and AberrPl_TD for AES.

1.695% increase of trial-routed wirelength for JPEG. Moreover, Aberr_TD, in comparison to TradPl_TD, reduces total negative slack (TNS) by 7.322% for AES, and by 7.748% for JPEG. Figure 12 shows the slack distributions of TradPl_TD, APlace_TD and AberrPl_TD for AES.

Impact of scaling. Our second set of experiments evaluates the effect of chip size on performance improvement obtained with our aberration-aware placement method. We perform AberrPl_TD with a variety of scaling factors, such that the number of die copies within the reticle is 1x1, 2x2, 4x4, 6x6, and 8x8. The results for circuits AES and JPEG are presented in Table III and Table IV, respectively. We report the improvement of the slowest chips among the multiple copies of chips. Comparing with APlace_TD, we see that MCT of AberrPl_TD with

Table III. Results of Aberration-Aware Placement (AberrPl_TD) with a Variety of Scaling Factors for Testcase AES

Copies	Method	Place		TrialRoute		AberrSTA
		HPWL (e9 um)	CPU (s)	WL (e5 um)	#vias (e5)	MCT (ns)
1	APlace_TD	1.1803	1457	6.541	1.2531	1.8013
	AberrPl_TD	1.1922	1471	6.645	1.2542	1.7443
	Imp (%)	**−1.0081**	**−0.9609**	**−1.5900**	**−0.0886**	**3.1636**
2	APlace_TD	1.1814	1469	6.548	1.2531	1.8212
	AberrPl_TD	1.1923	1486	6.651	1.2545	1.7651
	Imp (%)	**−0.9210**	**−1.1572**	**−1.5730**	**−0.1085**	**3.0812**
4	APlace_TD	1.1813	1478	6.555	1.2531	1.8461
	AberrPl_TD	1.1927	1491	6.657	1.2544	1.7942
	Imp (%)	**−0.9677**	**−0.8796**	**−1.5561**	**−0.1037**	**2.8093**
6	APlace_TD	1.1814	1482	6.556	1.2532	1.8483
	AberrPl_TD	1.1926	1499	6.651	1.2545	1.7974
	Imp (%)	**−0.9503**	**−1.1471**	**−1.4490**	**−0.1061**	**2.7558**
8	APlace_TD	1.1814	1487	6.555	1.2532	1.8500
	AberrPl_TD	1.1929	1502	6.649	1.2545	1.7995
	Imp (%)	**−0.9759**	**−1.0087**	**−1.4340**	**−0.1021**	**2.7310**

Table IV. Results of Aberration-Aware Placement (AberrPl_TD) with a Variety of Scaling Factors for Testcase JPEG

Copies	Method	Place		TrialRoute		AberrSTA
		HPWL (e9 um)	CPU (s)	WL (e5 um)	#vias (e5)	MCT (ns)
1	APlace_TD	6.3312	23791	3.743	6.1874	2.8875
	AberrPl_TD	6.3932	24139	3.780	6.1938	2.7751
	Imp (%)	**−0.9792**	**−1.4627**	**−0.9885**	**−0.1036**	**3.8918**
2	APlace_TD	6.3340	23801	3.746	6.1881	2.9009
	AberrPl_TD	6.3988	24211	3.778	6.1940	2.8002
	Imp (%)	**−1.0236**	**−1.7226**	**−0.8542**	**−0.0952**	**3.4710**
4	APlace_TD	6.3381	23821	3.745	6.1891	2.9309
	AberrPl_TD	6.3918	24203	3.781	6.1943	2.8396
	Imp (%)	**−0.8474**	**−1.5396**	**−0.9612**	**−0.0835**	**3.1176**
6	APlace_TD	6.3379	23801	3.744	6.1892	2.9210
	AberrPl_TD	6.4021	24298	3.772	6.1944	2.8613
	Imp (%)	**−1.0124**	**−2.0881**	**−0.7479**	**−0.0837**	**2.9549**
8	APlace_TD	6.3379	23802	3.745	6.1881	2.9380
	AberrPl_TD	6.4020	24299	3.77	6.1942	2.8532
	Imp (%)	**−1.0104**	**−2.0881**	**−0.6676**	**−0.0989**	**2.8835**

the scaling factor improves by $2.731 − 3.164\%$ $(50 − 57\text{ps})$ for AES and by $2.884 − 3.892\%$ $(85 − 112\text{ps})$ for JPEG. Trial-routed wirelength increases by $1.434 − 1.59\%$ for AES and by $0.668 − 0.989\%$ for JPEG, which is negligible compared to the significant MCT improvement. Figure 13 shows the MCT and trial-routed wirelength improvement as a function of the scaling factor. We observe that the performance improvement obtained gradually decreases as the number of copies in the field increases. However, larger chip size may not always achieve better timing improvement compared to the smaller chip size with aberration-aware placement. For example, suppose that there are two
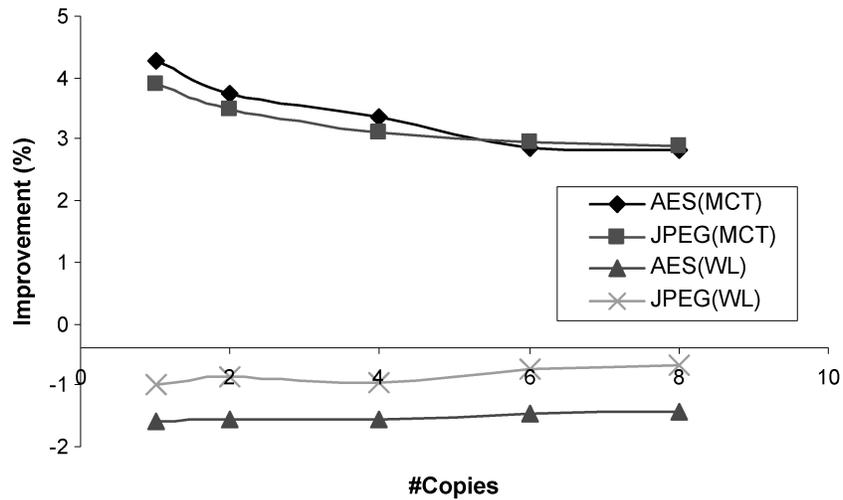
Fig. 13. Routed wirelength (WL) and MCT of AberrPl_TD as functions of the scaling factor for testcases AES and JPEG.
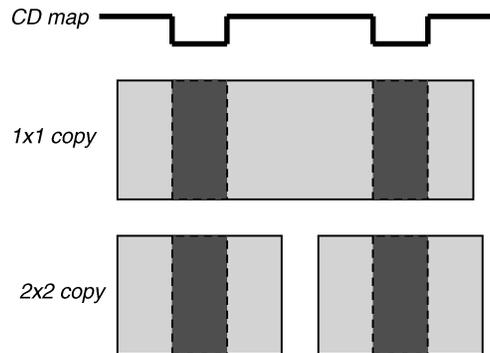

Fig. 14. Example showing nonmonotonicity of achievable MCT versus chip size. Red color represents fast cell delay regions in the lens aberration map. It is possible to achieve better MCT even with smaller chip size (e.g., 2x2 copies per field instead of 1x1 copy per field).

regions of the field which make gate CDs small (i.e., gate delay fast) due to aberration, as shown in Figure 14. In the case of 1x1 copy, aberration-aware placement will attempt to place timing-critical cells in these two regions to improve the gate delay. However, due to the limited size of the regions, not all timing-critical cells in a timing-critical path can be accommodated in one region. As a result, the separation of cells from a timing-critical path into two regions increases the wirelength, and consequently delay, of the timing-critical path. In the case of 2x2 copies, all timing-critical cells can be placed in one region or the neighborhood of the region. As a result, aberration-aware placement does not significantly affect wirelength, and 2x2 copies could have smaller delay than 1x1 copy.
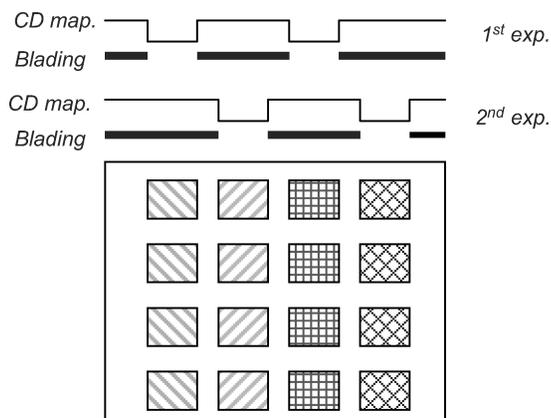
Fig. 15. An example of a *blading column technique*. For the first exposure, columns 1 and 3 in a lens field are used for chips in the reticle columns 1 and 3, while columns 2 and 4 are bladed. Columns 1 and 3 in a lens field can then be used for chips in the reticle columns 2 and 4 during a second exposure.

Impact of blading. A third set of experiments validates the proposed method when used in conjunction with lens field blading which allows partial reticle exposure. Balasinski [2004] proposed a multilayer mask technology which relies on sharing the reticle space between multiple layers of the same design. Based on the concept, which cuts out parts of the lens field, we propose a new *blading column technique* (BCT) to further optimize MCT in conjuction with our aberration-aware placement. The technique avoids the use of those portions of the aberration map that induce a large, positive gate delay variation.

In our experiments we assume that there are four die copies in the field, as shown in Figure 15. BCT allows any two dies to be exposed, thereby only partially using the reticle. For example, if we blade columns (2, 4) at the first exposure, only columns (1, 3) in the lens field are exposed for the chips in reticle columns 1 and 3. Chips in columns (2, 4) can be exposed in a second exposure after moving the wafer stage to use columns (1, 3) in the lens field again. Note that we use only some columns for exposure of all chips, selectively blading the columns that have aberration that is unfavorable to chip performance. Unfortunately, BCT requires two exposure passes and thus the throughput is halved. When not all columns are used, our aberration-aware placement performs timing optimization for only the columns in the lens field that are used. We evaluate the use of our technique with blading by considering several blading schemes and assessing the impact on chip performance, HPWL, and trial-routed wirelength.

We assume that there are four columns, where column numbers increase from left to right, in the reticle with 4x4 die copies of chip. The results are summarized in Table V and Table VI. Three comparisons in MCT improvement are presented: (1) blading versus no blading (**Impr.1**), (2) blading column of AberrPl_TDversus blading column of APlace_TD (**Impr.2**) and (3) blading columnof AberrPl_TD versus no blading of APlace_TD (**Impr.3**). The results of APlace_TD and AberrPl_TD show the performance improvements obtained

Table V.
Results of timing driven APlace (APlace_TD) and aberration-aware placements (AberrPl_TD) with a variety of blading columns for testcase AES. Three comparisons in MCT improvement are presented: (1) blading versus no blading (**Impr.1**), (2) blading column of AberrPl_TD versus blading column of APlace_TD (**Impr.2**) and (3) blading column of AberrPl_TD versus no blading of APlace_TD (**Impr.3**)

Method	Blading Col.	Place		TrialRoute		AberrSTA	Impr.1	Impr.2	Impr.3
		HPWL	**CPU**	**WL**	**#vias**	**MCT**	**MCT**	**MCT**	**MCT**
		(e9 um)	**(s)**	**(e5 um)**	**(e5)**	**(ns)**	**(%)**	**(%)**	**(%)**
APlace_TD	No Blading	1.18128	1478	6.555	1.2531	1.8461	—	—	—
	2,4	1.18131	1479	6.556	1.2534	1.8260	1.0846	—	—
	1,3	1.18122	1477	6.557	1.2535	1.8268	1.0422	—	—
	3,4	1.18131	1479	6.556	1.2537	1.8461	0.0000	—	—
	1,2	1.18131	1481	6.556	1.2537	1.8171	1.5710	—	—
	1,4	1.18131	1480	6.557	1.2536	1.8142	1.7241	—	—
	2,3	1.18131	1479	6.556	1.2535	1.8260	1.0846	—	—
AberrPl_TD	No Blading	1.1927	1491	6.657	1.2544	1.7942	—	2.8093	—
	2,4	1.1929	1492	6.658	1.2549	1.7731	1.1746	2.8978	3.9509
	1,3	1.1929	1493	6.659	1.2550	1.7721	1.2329	2.9966	4.0075
	3,4	1.1929	1491	6.657	1.2550	1.7728	1.1931	3.9688	3.9688
	1,2	1.1931	1493	6.659	1.2553	1.7663	1.5560	2.7944	4.3215
	1,4	1.1930	1490	6.658	1.2551	1.7678	1.4691	2.5572	4.2371
	2,3	1.1930	1489	6.658	1.2550	1.7731	1.1746	2.8978	3.9509

Table VI.
Results of timing driven APlace (APlace_TD) and aberration-aware placements (AberrPl_TD) with a variety of blading columns for testcase JPEG. Three comparisons in MCT improvement are presented: (1) no blading versus blading (**Impr.1**), (2) blading column of AberrPl_TD versus blading column of APlace_TD (**Impr.2**) and (3) blading column of AberrPl_TD versus no blading of APlace_TD (**Impr.3**)

Method	Blading Col.	Place		TrialRoute		AberrSTA	Impr.1	Impr.2	Impr.3
		HPWL	**CPU**	**WL**	**#vias**	**MCT**	**MCT**	**MCT**	**MCT**
		(e9 um)	**(s)**	**(e5 um)**	**(e5)**	**(ns)**	**(%)**	**(%)**	**(%)**
APlace_TD	No Blading	6.3381	23821	3.745	6.1891	2.9309	—	—	—
	2,4	6.3374	23828	3.746	6.1892	2.9237	0.2428	—	—
	1,3	6.3387	23832	3.744	6.1893	2.9164	0.4951	—	—
	3,4	6.3390	23824	3.743	6.1892	2.9240	0.2356	—	—
	1,2	6.3391	23824	3.747	6.1893	2.9080	0.7835	—	—
	1,4	6.3391	23826	3.746	6.1892	2.9121	0.6455	—	—
	2,3	6.3383	23824	3.748	6.1893	2.9138	0.5866	—	—
AberrPl_TD	No Blading	6.3918	24203	3.781	6.1943	2.8396	—	3.1176	—
	2,4	6.3924	24213	3.783	6.1943	2.8292	0.3632	3.2345	3.4695
	1,3	6.3924	24211	3.782	6.1943	2.8274	0.4236	3.0480	3.5280
	3,4	6.3918	24206	3.781	6.1944	2.8281	0.4066	3.2837	3.5115
	1,2	6.3930	24214	3.786	6.1944	2.7789	2.1312	4.4337	5.1824
	1,4	6.3931	24217	3.781	6.1943	2.8076	1.1263	3.5865	4.2088
	2,3	6.3921	24211	3.782	6.1943	2.8318	0.2716	2.8107	3.3808

using BCT. We observe that for test case AES, APlace_TD and AberrPl_TD can respectively reduce MCT by 1.724% and 1.556% with 0.003% and 0.107% HPWL increase, and 0.031% and 0.03% increase in trial-routed wirelength. With (1, 2) blading columns for the JPEG test case, APlace_TD and AberrPl_TD can respectively reduce MCT by 0.784% and 2.131%, with 0.0162% and 0.0187% HPWL increase, and 0.0534% and 0.1322% increase in trial-routed wirelength.

The absolute MCT improvements achieved with BCT and AberrPl_TD for AES and JPEG are 28ps and 61ps, respectively.

We also compare MCT improvements of blading for AberrPl_TD versus corresponding improvements for APlace_TD (**Impr.2**), and MCT improvements for AberrPl_TD with blading versus corresponding improvement for APlace_TD with no blading (**Impr.3**).[11] For the AES testcase, AberrPl_TD (**Impr.2**) can reduce MCTs by $2.557 - 3.969\%$ (i.e., $46 - 73$ps), with $0.9676 - 1.074\%$ (resp. $1.541 - 1.556\%$) increase in half-perimeter (resp. trial-routed) wirelength. For the JPEG test case, AberrPl_TD (**Impr.2**) can reduce MCTs by $2.811 - 4.434\%$ (i.e., $82 - 129$ps), with $0.848 - 0.868\%$ (resp. $0.9072 - 1.0152\%$) increase in half-perimeter (resp. trial-routed) wirelength. **Impr.3** shows the maximum improvement of AberrPl_TD with the blading column technique. For AES, AberrPl_TD with $(1, 2)$ blading reduces MCT by 4.322%, (i.e., 80ps) with 1.00% (resp. 1.587%) increase in half-perimeter (resp. trial-routed) wirelength. For JPEG, AberrPl_TD with $(1, 2)$ blading reduce MCT by 5.182% (i.e., 152ps) with 0.866% (resp. 1.095%) increase in half-perimeter (resp. trial-routed) wirelength. Averaged over our two test cases, worst-case cycle time and total negative slack respectively reduce by $\sim 4.749\%$ (i.e., 116ps) and $\sim 7.535\%$.

We consider the observed MCT improvements (i.e., $80 - 152$ps) achieved by our aberration-aware placement and the blading column technique to be quite significant. Such MCT reductions can tremendously improve parametric yield and quicken timing closure. The penalties of HPWL, trial-routed wirelength, and the number of vias are less than 1.5%, and are practically negligible. We note that the concept of stepper-specific place-and-route has long been attractive to high-end, high-volume custom products—for example, Dr. N. Sherwani of Intel posed exactly this challenge to the physical design community at the 1999 International Symposium on Physical Design (ISPD). With the future of process module costs, inherent equipment variabilities, and exclusivity of fabless-foundry tie-ups all being unclear today, we believe that it will be important to have stepper-specific layout flows available going forward.

6. CONCLUSIONS AND ONGOING WORK

We have proposed an accurate aberration-aware timing analysis flow and a novel aberration-aware timing-driven placement technique, *AberrPl*, as a practical and effective approach to improve timing yield after manufacturing. We implement our method based on a general analytical placement framework and test it within a standard industry flow using leading-edge tools. We also study the dependence of our improvement on chip size, and when the technique is used along with field blading which allows partial reticle exposure. Averaged over our two test cases, worst-case cycle time and total negative slack respectively reduce by $\sim 4.749\%$ (116ps) and $\sim 7.535\%$ at the cost of $\sim 1.341\%$ increase in wirelength, with hold-time violations.

---

[11]There are no entries in **Impr.2** and **Impr.3** of APlace_TD as shown in Tables V and VI since we compare APlace_TD with AberrPl_TD and record the improvements in the AberrPl_TD column.

The benefits of AberrPl_TD are expected to increase in future technology nodes. We are currently engaged in further experimental validation and research. Our ongoing research is in the following directions.

—The proposed aberration-aware placement approach aims at improving performance of all design copies in the reticle field and hence is limited by the slowest ones. However, for many designs, chips of slower speeds can also be sold, albeit at a lower value (speed binning). We plan to improve our approach so that the total value of all chips is maximized.

—We also wish to enhance our placer to comprehend leakage constraints, since leakage is increasingly starting to determine yield and is exponentially affected by CD.

—We are researching the possibility of an aberration-aware OPC method which applies different OPC models for devices at different lens positions, instead of the simple OPC method with average Zernike's coefficients across the reticle, to improve pattern printability and lithographic process window. While we noted at the outset that global placement seems to be a more appropriate knob than OPC for compensation of lens aberration, we wish to pursue a clear confirmation or refutation of this intuition.

—For chip manufacturing, a modern fab can employ multiple lithography lens systems. Different lenses will have different lens aberrations. For very high-volume production of a chip (ASIC or microprocessor), it may be the case that multiple systems are used simultaneously. We plan to improve our placement engine to achieve "generic" aberration-aware placements that improve parametric yield in light of the systematic lens aberrations of all the lenses.

—A modern reticle may contain multiple chips (especially for ASICs), even outside the shuttle context. Different chips are located at different points in the reticle field. We are developing an aberration-aware placement to address such multiple concurrent instances of design optimization.

—Restricted design rules have been receiving increased attention from industry (e.g., relaxed pitch helps to reduce CD asymmetry caused by coma aberration). We intend to evaluate such approaches with AberrPl in terms of design and manufacturability metrics.

REFERENCES

BALASINSKI, A. 2004. Multi-layer and multi-product masks: Cost reduction methodology. In *Proceedings of the 24th BACUS Symposium on Photomask Technology*. 351–359.

BRUNNER, T. A. 1997. Impact of lens aberrations on optical lithography. *IBM J. Resear. Devel. 41*.

CHEON, Y.-S., HO, P.-H., KAHNG, A. B., REDA, S., AND WANG, Q. 2005. Power-aware placement. In *Proceedings of the ACM/IEEE Design Automation Conference*. 795–800.

EISENMANN, H. AND JOHANNES, F. M. 1998. Generic global placement and floorplanning. In *Proceedings of the ACM/IEEE Design Automation Conference*. 269–274.

ETAWIL, H., AREIBI, S., AND VANNELLI, A. 1999. Attractor-repeller approach for global placement. In *Proceedings of the IEEE International Conference on Computer-Aided Design*. 20–24.

FARRAR, N., SMITH, A., BUSATH, D., AND TAITANO, D. 2001. In-situ measurement of lens aberrations. In *Proceedings of the SPIE Conference on Optical Microlithography*. 18–29.

FLAGELLO, D. G., LAAN, H., SCHOOT, J., BOUCHOMS, I., AND GEHA, B. 1999. Understanding systematic and random CD variations using predictive modelling techniques. In *Proceedings of the SPIE Conference on Optical Microlithography*. 162–175.

GORTYCH, J. AND WILLIAMSON, D. 1991. Effects of higher-order aberrations on the process window. In *Proceedings of the SPIE Conference on Optical Microlithography*. 368–381.

GU, J. AND HUANG, X. 1994. Efficient local search with search space smoothing: A case study of the traveling salesman problem (tsp). *IEEE Trans. Syst. Man Cybern. 24,* 5, 728–735.

GUPTA, P., KAHNG, A. B., AND PARK, C.-H. 2007. Detailed placement for enhanced control of resist and etch CDs. *IEEE Trans. Comput.-Aid. Des. Integr. Circ. Syst. 26,* 12, 2144–2157.

HU, B. AND MAREK-SADOWSKA, M. 2002. FAR: Fixed-points addition & relaxation based placement. In *Proceedings of the IEEE International Symposium on Physical Design*. 161–166.

KAHNG, A. B., LIU, B., AND WANG, Q. 2005. Supply voltage degradation aware analytical placement. In *Proceedings of the IEEE International Conference on Computer-Aided Design*. 437–443.

KAHNG, A. B., PARK, C.-H., SHARMA, P., AND WANG, Q. 2006. Lens aberration-aware timing-driven placement. In *Proceedings of IEEE Design, Automation and Test in Europe*. 890–895.

KAHNG, A. B., REDA, S., AND WANG, Q. 2005a. Aplace: A general analytic placement framework. In *Proceedings of the IEEE International Symposium on Physical Design*. 233–235.

KAHNG, A. B., REDA, S., AND WANG, Q. 2005b. Architecture and details of a high quality, large-scale analytical placer. In *Proceedings of the IEEE International Conference on Computer-Aided Design*. 891–898.

KAHNG, A. B. AND WANG, Q. 2004a. An analytic placer for mixed-size placement and timing-driven placement. In *Proceedings of the IEEE International Conference on Computer-Aided Design*. 565–572.

KAHNG, A. B. AND WANG, Q. 2004b. Implementation and extensibility of an analytic placer. In *Proceedings of the IEEE International Symposium on Physical Design*. 18–25.

KAHNG, A. B. AND WANG, Q. 2005. Implementation and extensibility of an analytic placer. *IEEE Trans. Comput.-Aid. Des. Integr. Circ. Syst. 24,* 5, 734–747.

KONG, T. 2002. A novel net weighting algorithm for timing-driven placement. In *Proceedings of the IEEE International Conference on Computer-Aided Design*. 10–14.

LEVINSON, H. J. 2001. *Principles of Lithography*. SPIE Press.

MARQUARDT, A., BETZ, V., AND ROSE, J. 2000. Timing-driven placement for FPGAs. In *Proceedings of the ACM Symposium on FPGAs*. 203–213.

MATSUYAMA, T., SHIBAZAKI, Y., OHMURA, Y., AND SUZUKI, T. 2002. High NA and low residual aberration projection lens for duv scanner. In *Proceedings of the SPIE Conference on Optical Microlithography*. 687–695.

NAYLOR, W., DONELLEY, S., AND SHA, L. 2001. Non-linear optimization system and method for wire length and delay optimization for an automatic electric circuit placer. US Patent 6301693.

ORSHANSKY, M., MILOR, L., CHEN, P., KEUTZER, K., AND HU, C. 2002. Impact of spatial intrachip gate length variability on the performanceof high-speed digital circuits. *IEEE Trans. Comput.-Aid. Des. Integr. Circ. Syst. 21,* 5, 544–553.

ORSHANSKY, M., MILOR, L., AND HU, C. 2004. Characterization of spatial intrafield gate CD variability, its impact on circuit performance, and spatial mask-level correction. *IEEE Trans. Semiconduct. Manufact. 17,* 2, 2–11.

ORSHANSKY, M., MILOR, L., NGUYEN, L., HILL, G., PENG, Y., AND HU, C. 1999. Intra-field gate CD variability and its impact on circuit performance. In *Proceedings of the IEEE International Electron Devices Meeting*. 479–482.

PROGLER, C., BORNA, A., BLAAUW, D., AND SIXTA, P. 2004. Impact of Lithography Variability on Statistical Timing Behavior. In *Proceedings of the SPIE Conference on Design and Process Integration for Microelectronic Manufacturing*. 101–110.

PROGLER, C. J. AND WONG, A. K. 2000. Zernike coefficients: Are they really enough? In *Proceedings of the SPIE Conference on Optical Microlithography*. 40–52.

SHIODE, Y., OKADA, S., TAKAMORI, H., MATUSDA, H., AND FUJIWARA, S. 2002. Method of Zernike coefficients extraction for optics aberration measurement. In *Proceedings of the SPIE Conference on Optical Microlithography*. 138–147.

TOH, K. K. H. AND NEUREUTHER, A. 1987. Identifying and monitoring effects of lens aberrations in projection printing. In *Proceedings of the SPIE Conference on Optical Microlithography*. 202–209.

VISWANATHAN, N. AND CHU, C. C.-N. 2004. Fastplace: Efficient analytical placement using cell shifting, iterative local refinement and a hybrid net model. In *Proceedings of the IEEE International Symposium on Physical Design*. 26–33.

WONG, A. K. 2001. *Resolution Enhancement Techniques in Optical Lithography*. SPIE PRESS.

WONG, A. K. 2002. Theoretical discussion on reduced aberration sensitivity of enhanced alternating phase-shifting masks. In *Proceedings of the SPIE Conference on Optical Microlithography*. 395–368.