# Performance optimization for gridded-layout standard cells

Jun Wang[a], Alfred K. Wong[b], and Edmund Y. Lam[a]

[a] Dept. of Electrical and Electronic Engineering, Univ. of Hong Kong, Pokfulam Road, Hong Kong;

[b] Fortis Systems, Inc., U.S.A.

## ABSTRACT

The grid placement of contacts and gates enables more effective use of resolution enhancement techniques, which in turn allow a reduction of critical dimensions. Although the regular placement adds restrictions during cell layout, the overall circuit area can be made smaller and the extra manufacturing cost can be kept to the lowest by a careful selection of the grid pitch, using template-trim lithography method, allowing random contact placement in the vertical direction, and using rectangular rather than square contacts. The purpose of this work is to optimize the gridded-layout-based process. The trade-off between the layout area and manufacturing cost, and the determination of the minimum grid pitch are discussed in this paper. We demonstrate that it is a 1-D scaling instead of the conventional 2-D scaling for standard cells and the narrow MOSFETs inside after the application of the gridded layout on the contact and gate levels. The corresponding effects on circuit performances, including the leakage current, are also explored.

**Keywords:** Design for manufacturability (DFM), Resolution Enhancement Techniques (RETs), Low-$k_1$ lithography, Gridded layout, Design rules, Template lithography, Multiple exposures, Standard cells, MOSFETs scaling, Leakage current, Circuit performance

## 1. INTRODUCTION

### 1.1. Resolution enhancement techniques (RETs)

The resolution limit of a conventional optical lithography system with on-axis coherent illumination can be expressed as

$$P_{min} = \frac{\lambda}{\mathrm{NA}}, \tag{1}$$

where $P_{min}$ is the minimum resolvable feature pitch, $\lambda$ is the illumination wavelength, and NA is the numerical aperture. The minimum resolvable feature pitch is proportional to the illumination wavelength $\lambda$ and decreases with increasing numerical aperture (NA) of the lithography system. Smaller dimensions can be printed by decreasing the wavelength or increasing the numerical aperture. Unfortunately, following Moore's Law,[1] the resolution demands of the IC industry have outpaced the introduction of more advanced lithography systems with a smaller $\lambda$ or a larger NA. As dimensions of leading edge devices continue to shrink, it becomes increasingly difficult to print certain levels such as contact and gate. In addition to the enhancement in lithography hardware, resolution enhancement techniques (RETs), new photomask technologies, and new layout design methodologies are required to print features with the minimum pitch approaching to or below the above limit.

According to the 2001 ITRS Roadmap,[2] optical lithography is still projected to be the dominant technology through at least the 2007 65 nm node. However, the required minimum resolvable pitches for these nodes are far below the limit in Eqn.1. RETs such as attenuated phase-shifting mask (Att-PSM)[3] and optical proximity correction (OPC)[4] only enable lithography with $P_{min}$ reaching or slightly below the limit in Eqn.1. Strong RETs pushing resolution significantly exceeding the conventional limit are necessary for these technology nodes.

Further author information: (Send correspondence to Jun Wang)
Jun Wang: E-mail: junwang@eee.hku.hk, Telephone: 852 2859-2698, Fax: 852 2559-8738, http://www.eee.hku.hk
Alfred K. Wong: E-mail: awong@fortis-systems.com, Telephone: 1 408 368 9782
Edmund Y. Lam: E-mail: elam@eee.hku.hk, Telephone: 852 2241-5942, Fax: 852 2559-8738

24th Annual BACUS Symposium on Photomask Technology,
edited by Wolfgang Staud, J. Tracy Weed, Proceedings of SPIE Vol. 5567
(SPIE, Bellingham, WA, 2004) 0277-786X/04/$15 · doi: 10.1117/12.569398

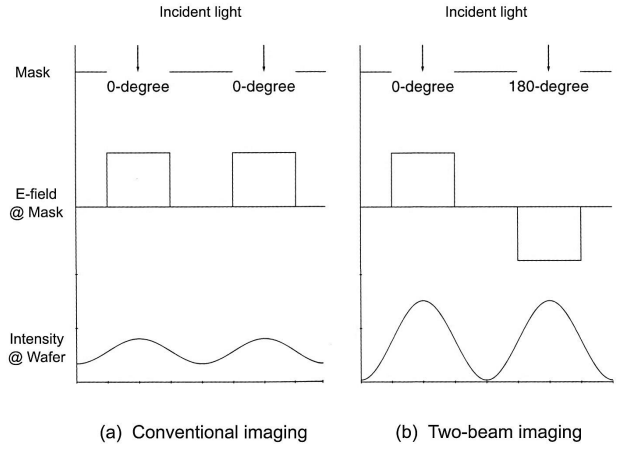



**Figure 1.** Two-beam imaging lithography provides higher resolution than conventional lithography by eliminating the zeroth order of diffracted light.

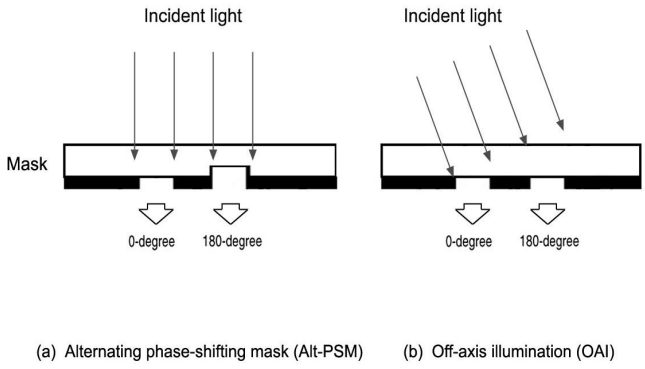



**Figure 2.** Two means of achieving two-beam imaging: (a) alternating phase-shifting mask (Alt-PSM) and (b) off-axis illumination (OAI).

### 1.2. Two-beam imaging and challenges

Based on two-beam imaging,[5] which introduces a 180 degree phase shift in the transmitted light between adjacent regions to eliminate the zeroth order of diffracted light, strong-RETs provide solutions decreasing the minimum resolvable pitch by 50%, as shown in Fig.1. There are two means to achieve two-beam imaging (Fig.2): alternating phase-shifting mask (Alt-PSM) and off-axis illumination (OAI). To shift the phase of transmitted light 180 degrees relative to light in adjacent regions, thickness of transmitting (clear) regions on an alternating phase-shifting mask are modulated according to the illumination wavelength and the material of the mask substrate[6] [Fig.2 (a)]. Off-axis illumination achieves the same effect for a particular pitch by illuminating a conventional mask at an appropriate angle[7] [Fig.2 (b)]. The improved resolution limit is now given by[6]:

$$P_{min} = 0.5\frac{\lambda}{\mathrm{NA}}. \tag{2}$$

Although the new resolution limit satisfies the requirement of the 65 nm and 45 nm technology nodes, challenges associated with the alt-PSM and OAI, not only for electronic design automation (EDA) solutions, but also for manufacturing cost, prevented them from a broad implementation in the past.

Phase assignment conflicts[8] and sub-resolution assist feature (SRAF) design[9] are major challenges associated with Alt-PSM and OAI respectively. They impact layout design and make it difficult to provide EDA solutions, including design rule checking (DRC), design tools and methodologies, that guarantee layout-compliant designs.[8–11] The RET-embedded design flow, which is required by the strong-RET-imposed layout restrictions, are significantly more complicated than conventional flows.[5] Layout designers must have a good knowledge of the strong-RETs used in a process. An easily established and non-RET-specific design for manufacturability (DFM) methodology will be more attractive for layout designers. Furthermore, the implementation of RETs are pushing the fabrication cost of leading edge photomasks to an alarming level. For example, the cost for a full 28-30 mask set for a 90 nm node is projected to be well over 1 million US dollars.[12] The high mask set cost has induced a decline recently for low volume production market, such as ASICs. A low cost process will be more likely adopted by ASIC foundries for future technology nodes.

### 1.3. Gridded layout

Gridded layout is a possible solution to the above challenges. Designs are restricted to allow critical dimension features only on a grid, which simplifies the layout-compliance design flow and is optimized for many strong-RETs.[5] Many gridded-layout-based lithography approaches have been proposed in the last few years for the contact and the gate level,[13–20] pushing the critical feature pitch to about its minimum value. The gridded

layout also enables a prior fabrication and reuse of photomasks, thereby reducing manufacturing cost.[12] This makes gridded layout a very attractive option for ASIC foundries.

However, the radical restrictions of the gridded layout impose significant effects on layout design. Although the features can be designed smaller and packed closer in gridded layout, the excessive lithography friendliness may be so restrictive on layout compaction that circuit area increases unacceptably. The grid pitch and the layout topology should be carefully selected according to the new design rules to seek a fine balance between the lithographic optimization and layout compaction.

We have previously reported our efforts for a methodology to apply gridded contacts and gates on layout of standard cells—an elementary block of ASICs.[21–23] Although regular placement of contacts and gates adds restrictions during cell layout, the overall circuit area can be made smaller by keeping contacts and gates randomly in the cell height direction while placing regularly in the cell width direction with $\frac{1}{2}$ transistor pitch* as the grid pitch. Furthermore, by using template lithography, the cost for the mask set does not increase because no extra non-reusable mask is introduced in the process. Although the manufacturing cost per wafer increases because the yield decreases with the extra exposures, the manufacturing cost per die may decrease with the scaling of MOSFETs.

To achieve the best performance, the process for the gridded layout, including design methodology, lithographic method, and electrical ramifications such as leakage current increase needs to be studied. As the detail of the gridded layout methodology and lithographic method for standard cells have been discussed previously,[23] this paper briefly summarizes their salient points in Sec.2 and focuses on the design rule optimization to minimize the cell area in Sec.3 and the circuit performance of the gridded layout standard cells in Sec.4 respectively.

## 2. LITHOGRAPHIC METHOD FOR GRIDDED LAYOUT STANDARD CELLS

To fabricate a layout with gridded placed contacts and gates as shown in Fig. 3 for an example, we need one reusable template mask (a chromeless alternating phase-shifting mask is used in this example, as shown in Fig. 4) and two trim masks for contacts (Fig. 6) and gates (Fig. 9) separately. All contacts and gates are placed randomly in the vertical direction while regularly in the horizontal direction with $\frac{1}{2}$ transistor pitch as the grid pitch. A $\lambda = 193\,\text{nm}$, NA=0.75 lithography system and a 245 nm horizontal pitch are used in this investigation. After an exposure, the opposite phase shift of patterns on the chromeless template mask creates periodic unexposed dark lines at the boundary of 0° and 180° regions, as shown in Fig. 5. The period of the 0° and 180° regions on the chromeless phase-shifting template mask is designed to be one transistor pitch so that the period of the dark lines is half of that. Exposures of the contact and gate trim mask (Fig. 7 and Fig. 10) on these periodic dark lines remove the unwanted parts of the dark lines and the cuts of the dark lines left form the final images of regularly-placed contacts (Fig. 8) and gates (Fig. 11).

The horizontal dimension and position of contacts and gates are determined by the template mask, while the vertical dimension and position are determined by the trim mask. Because the features in the trim mask are placed randomly, the positions of contacts and gates are random in the height direction. Determined by the exposures of the different masks, contacts have different size in different directions. By using the regular placement and the chromeless phase-shifting mask in the horizontal direction, the horizontal contact size is smaller than the vertical contact size which is determined by the resolution of the contact trim mask. The horizontal size of a contact can be as small as that of a gate. For example, a 70 nm horizontal contact size can be reached using a 193 nm lithography.[14] Although a binary trim mask can be used for gates because of their relatively larger dimension in the vertical direction, an advanced trim mask should be used to get a vertical contact size the same as the minimum contact size in traditional one-exposure approaches, which is 160 nm in a 193 nm lithography.

The double-exposure lithographic approach in the example only forms vertically-oriented fine features on a poly-silicon layer. To include other features on the layer such as the horizontally-oriented poly-silicon connection paths and contact-landing pads, whose dimensions are not critical for cell area and can be designed larger, many

*A transistor pitch, also called a "contacted pitch", is the minimum pitch between two gates with a contact between them.

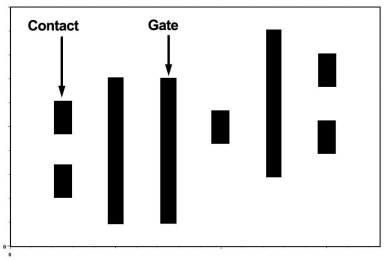



**Figure 3.** A test layout block (contact and gate levels only).

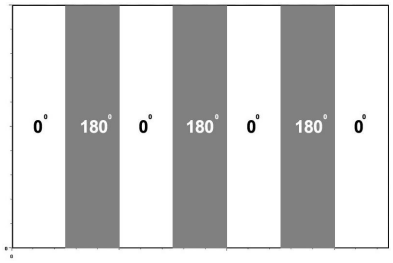



**Figure 4.** The chromeless alt-PSM template.

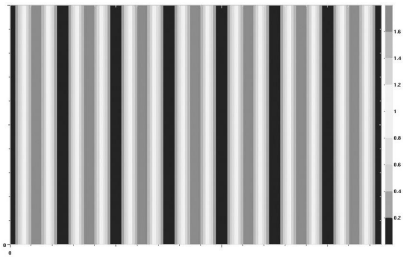


**Figure 5.** The exposure image of the chromeless alt-PSM template.

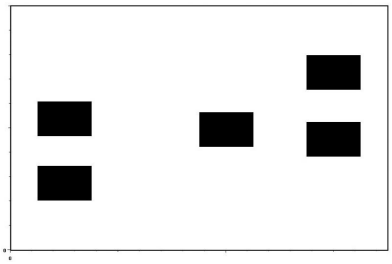


**Figure 6.** The trim mask for the contact level.

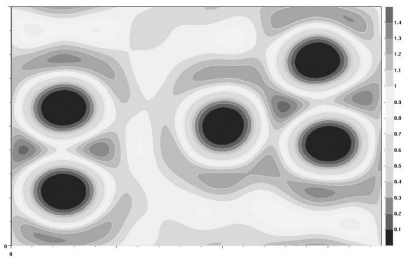


**Figure 7.** The exposure image of the trim mask for the contacts.

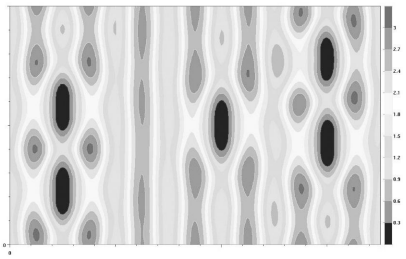


**Figure 8.** The final image of the contacts.

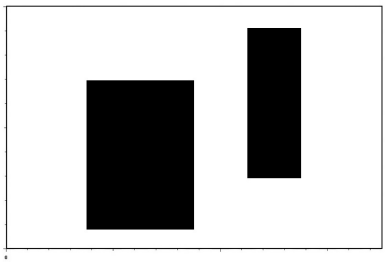


**Figure 9.** The binary trim mask for the gate level.

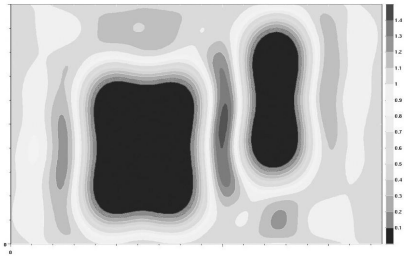


**Figure 10.** The exposure image of the binary trim mask for the gates.

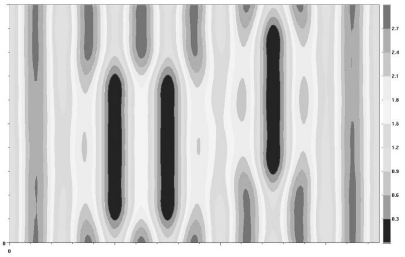


**Figure 11.** The final image of the gates.

other template-trim approaches can be used[12, 14, 15, 18] and the final number of exposures, masks, and resistor layers might be different. For example, since most of the gate are vertically-orientated in standard cells, it is a suitable application of Canon IDEAL method.[15] All coarse features can be formed by the trim mask and the total number of masks and exposures can be kept to two. Altogether 3 masks (1 reusable template mask, 1 trim mask for contacts, and 1 trim mask for gates) and 4 exposures (2 for a contact layer and 2 for a poly-silicon layer) are needed to fabricate the contact and poly-silicon layers of a fabrication-friendly standard cell. The extra cost is kept to the lowest because no extra non-reusable mask is needed.

## 3. DESIGN RULE OPTIMIZATION FOR GRIDDED LAYOUT STANDARD CELLS

### 3.1. Design rules for gridded layout standard cells

Besides the introduction of the placement restriction for contacts and gates, design rules for the gridded layout should be optimized according to the relevant lithography methods to achieve a balance among a variety of quality metrics, including layout area, manufacturing cost, and circuit performance. For example, fabricated by different masks, the vertical features and the horizontal features on a poly-silicon layer should have different design rules and can be separated into different layers, such as a gate layer and a poly layer, to facilitate the

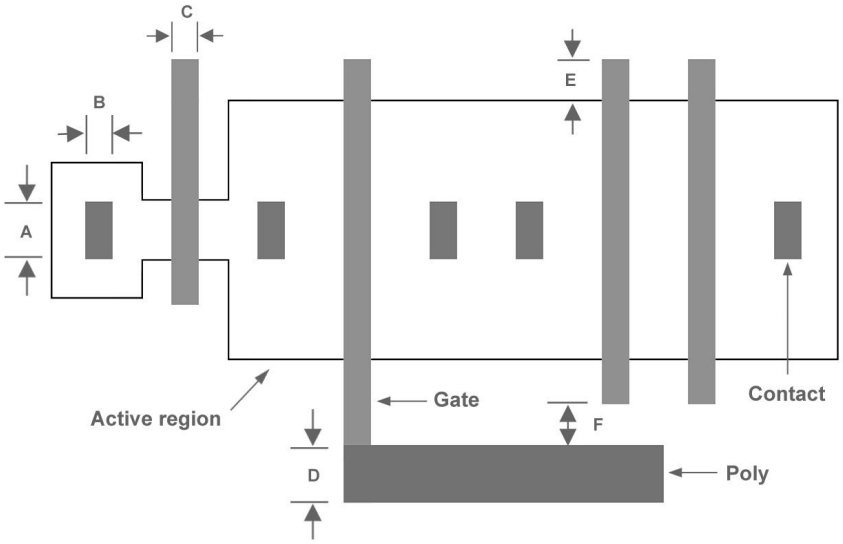



**Figure 12.** The design rules that are affected by the gridded layout and the relevant lithography methods.

**Table 1.** The list of the design rules that should be modified or introduced by the gridded contacts and gates in standard cells. The possible rule values for a 193 nm lithography are listed also as a demonstration.

| Label | Description | Rule |
|---|---|---|
| R | All contacts and gates must be placed on a grid in the cell width direction. | pitch=210 nm |
| A | Minimum size of a contact region in the cell height direction. | 150 nm |
| B | Fixed size of a contact region in the cell width direction. | 70 nm |
| C | Fixed size of a gate region in the cell width direction. | 70 nm |
| D | Minimum width of a poly region. | 130 nm |
| E | Minimum overlap of a gate region extended into field oxide (endcap). | 70 nm |
| F | Minimum clearance of a gate region to a poly region not connected. | 70 nm |

layout design and design rule check. The gate layer contains all of the vertically-oriented features, such as all gates and vertical poly-silicon connections, which are placed regularly in the horizontal direction and imaged by the template mask. The poly layer contains all other features on a poly-silicon layer, such as contact-landing pads and horizontal poly-silicon connections, which are imaged by the trim mask or a third mask. Fig.12 illustrates the design rules that are affected by the gridded layout and the relevant lithography methods. The detailed descriptions are listed in Table.1.

The rules R is introduced for the gridded layout to restrict all contacts and gates on a grid in the horizontal direction. With different size in different directions, rectangular contacts have an extra size rule B in the horizontal direction to distinguish it from the vertical contact size A. Separated into different levels, the gate level and the poly level on a poly-silicon layer have different size rules, C and D, respectively. The fixed horizontal size of contacts and gates (rules B and C) can be decreased to as low as 70 nm $\sim$ 80 nm in a 193 nm lithography by using template-trim approaches, while the minimum vertical contact size (rule A) and poly width (rule D) are about as same as those of the original design rules.

Furthermore, fabricated by double exposures, thereby, without the problem of line shorting, gates are expected to have a gate extension (rule E) smaller than that in conventional design rules. The gate extension is then only limited by the vertical alignment error between the gate level and the active level.

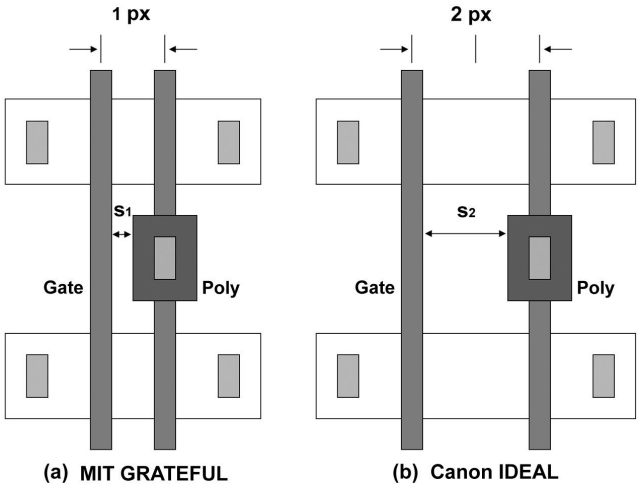



**Figure 13.** Different lithographic methods for the poly-silicon layer lead to different impact on layout area.

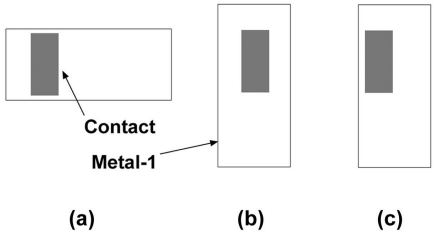


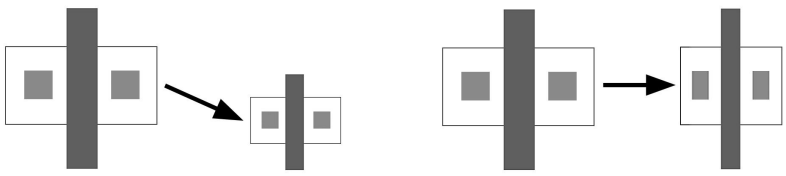

**Figure 14.** Scenarios for the connection between the contact and the metal-1 level.

The determining of the rule F in Table.1 is related to the balance between layout area and manufacturing cost, while the determining of the grid pitch in the rule R in is related to layout scaling style. They are discussed in detail in the following Sec.3.2 (rule F) and Sec.3.3 (rule R) respectively.

### 3.2. Layout area vs. manufacturing cost

The last rule (F) in Table.1, the minimum clearance of a gate region to a poly region that is not connected, varies with different lithographic method for the poly-silicon layer. Because there are many template-trim approaches available[12, 14, 15, 18] for the poly-silicon layer, the final number of exposures, masks, and resistor layers might be different, leading to different impact on design rules and manufacturing cost.

For example, the MIT GRATEFUL[12] method uses different resistor layers and exposures for gate and poly levels respectively. Therefore, the minimum clearance between a gate to a poly region can be much smaller than the conventional minimum space between two regions on a poly-silicon layer. Fig.13 (a) illustrates a layout enabled by the MIT GRATEFUL. The minimum clearance between the gate and the poly region, $s_1$, is small enough to enable the placement of a gate and a poly region on two neighboring grid lines (1/2 transistor pitch). Another lithographic method, Canon IDEAL,[15] uses only one resistor layer. The exposure of the trim mask images the features on the gate level and the poly level at the same time. The minimum clearance between the gate and the poly region is then much large than that in the MIT GRATEFUL method. One extra grid line has to be introduced between the two neighboring MOSFETs to satisfy the rule for the minimum clearance, $s_2$ (one transistor pitch), as shown in Fig.13 (b). In our estimation, MIT GRATEFUL leads to an average layout area about $2\% \sim 3\%$ smaller than that enabled by Canon IDEAL for standard cells. However, on the other hand,



(a) Conventional 2-D scaling of MOSFETs. (b) One-D scaling of MOSFETs.

**Figure 15.** It is a 1-D scaling, instead of the conventional 2-D scaling enabled by the gridded placement for minimum-width MOSFETs in standard cells.

the extra resistor layer and exposure in the MIT GRATEFUL increase the manufacturing cost and decrease the yield. The balance between layout area and manufacturing cost should be taken into consideration when choosing the lithographic method for the poly-silicon layer.

Another example for the balance between layout area and manufacturing cost is the connection between the contact level and the metal-1 level. The different size of a rectangular contact in different directions enables three different connection styles with a metal-1 path, as shown in Fig.14. Although the connection style in Fig.14 (c) affects the yield and is not preferred by manufacturers, it alleviates the restrictions for metal-1 placement and leads to a smaller layout area. The minimum extension of a metal-1 path over a rectangular contact for the connection in Fig.14 (c) should be determined carefully to make a balance between the yield and layout area.

## 3.3. One-D scaling and minimum horizontal grid pitch

### 3.3.1. One-D scaling

The contact and gate levels only determine the width of a standard cell. The metal-1 layer determines the height[†]. When applying the gridded layout on multiple layers, the matching between different grids on different layer adds extra restrictions beyond the gridded layout restrictions on each layer. At the same time, the multiple exposures associated with the gridded layout increase the manufacturing cost.[21] These make it impractical to apply the gridded layout on the metal-1 layer after the application on the contact and gate levels from both design and economic points of view. Therefore, the height of cells does not scale with the scaling of cell width,[23] which is enabled by the gridded-placed contacts and gates. It is a 1-D scaling instead of the conventional 2-D scaling for boundary of standard cell after the application of the gridded layout on the contact and gate levels.

Furthermore, although the minimum gate length can be decreased by the gridded gate placement, the minimum gate width cannot, which is determined by the minimum width of an active region. Therefore, with only the scaling of their length, minimum-width MOSFETs in standard cells get also the 1-D scaling after the application of the gridded layout on the contact and gate levels, as shown in Fig.15.

For 1-D scaled standard cells, cell area has a linear relationship with the transistor pitch. A small transistor pitch is preferred to get a smaller cell area. As twice the horizontal grid pitch, the transistor pitch can be calculated by the equation below:

$$P_{transistor} = 2 \cdot S_{c \to g} + L_g + W_c, \tag{3}$$

where $P_{transistor}$ is the transistor pitch, $S_{c \to g}$ is the minimum clearance of a contact to a gate, $L_g$ is the minimum length of a gate, and $W_c$ is the horizontal dimension of a contact. Although $L_g$ and $W_c$ can be reduced by the grid layout placement, the minimum clearance of a contact to a gates $S_{c \to g}$ which is determined by the misalignment between two masks will not decrease accordingly with the decrease of $L_g$ and $W_c$. The original minimum transistor pitch of a 130 nm technology is about 510 nm. The improved minimum transistor pitch is determined by the relevant lithographic method. Using a template-trim approach, the size of contacts and gates can be decreased by 50% to as low as 70 nm, thereby reduces the minimum transistor pitch to about 420 nm and the minimum grid pitch to about 210 nm.

### 3.3.2. Minimum horizontal grid pitch

However, layout area does not always scaled proportionally with the scaling of the transistor pitch. Fig.16 shows an estimation of the average area change of a gridded layout standard cell library migrated from a 130 nm library. The technology-enabled minimum transistor pitch is assumed to be as low as 360 nm. The average cell area decreases with the decrease of the transistor pitch except at 420 nm, which is the transistor gap of the technology.

The transistor gap, as shown in Fig.17, is the minimum pitch between the terminal contacts of two isolated transistors which are adjacent in the layout but not connected in the circuit. The transistor gap is usually a little smaller than a transistor pitch and can be treated as one transistor pitch in the gridded layout. The transistor

[†]The typical height of a standard cell is 10 metal-1 pitches: 3 pitches for power supply paths and 7 pitches for intra-cell routing.

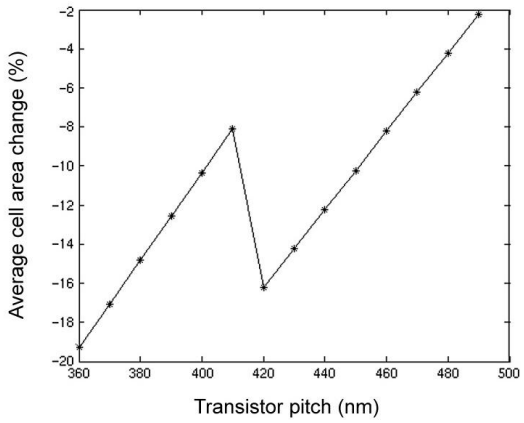



**Figure 16.** Decreasing the transistor pitch does not always lead to a smaller cell area.

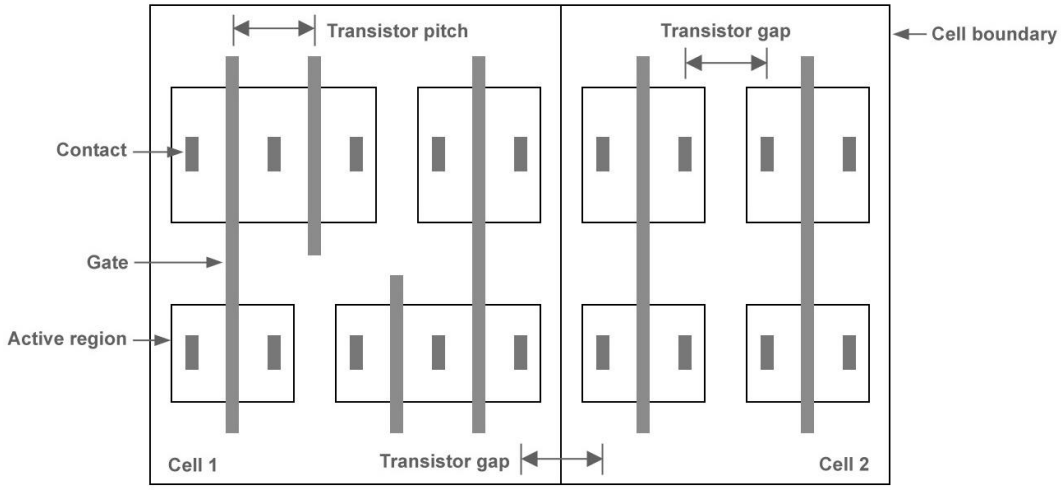



**Figure 17.** The transistor gap and the transistor pitch (Only active, gate, and contact layer are plotted here.).

gap scales slower than the transistor pitch because the minimum space between two neighboring active regions cannot be decreased by the application of the gridded layout. When the minimum transistor pitch decreases with the application of the gridded layout, it can be smaller than the transistor gap in some technologies. In that case, if a half minimum transistor pitch is used as the grid pitch, the transistor gap is slightly larger than 2 grid pitches (a transistor pitch) and must be increased to 3 grid pitches ($1\frac{1}{2}$ transistor pitches). Furthermore, extra space must be introduced between the adjacent standard cells during cell placement in the physical design to avoid the adjacent transistors in different cells from being placed too close. The overall layout area may not decrease with the transistor pitch shrinking under the transistor gap.

For the 130 nm technology under study, the technology-enabled minimum transistor pitch is equal to the transistor gap at 420 nm coincidentally. Therefore, 210 nm, as half of the minimum transistor pitch, can be used as the grid pitch. However, as discussed above, for a technology with a technology-enabled minimum transistor pitch smaller than the transistor gap, whatever to use half of the technology-enabled minimum transistor pitch or half of the transistor gap as the grid pitch should be determined by an examine. Use the examine result shown in Fig.16 for the 130 nm technology under study as an example. If the technology-enabled minimum transistor pitch is smaller than the transistor gap (420 nm) and lies between 380 nm to 420 nm, half of the transistor gap should be used as the grid pitch to get a smaller layout area. For other cases, half of the technology-enabled

minimum transistor pitch should be used as the grid pitch.

## 4. CIRCUIT PERFORMANCE OF GRIDDED-LAYOUT STANDARD CELLS

### 4.1. Scaling effects

As mentioned above, it is a 1-D scaling instead of the conventional 2-D scaling for narrow MOSFETs in standard cells after applying the gridded layout on the contact and gate levels. This results in a 1-D scaled circuit performance for standard cells with mainly narrow MOSFETs, such as cells with small load capacitance. For cells with mainly wide MOSFETs, such as cells with large load capacitance in which wide MOSFETs are needed to provide large driving current, the scaling effects lie between the 1-D and 2-D scalings. Scaling effects of the 1-D scaling on device and circuit parameters are summarized in Table.2. Effects of the 2-D scaling are also listed as a comparison. It can be found that the two scalings have the same scaling factor for intrinsic delay while different scaling factors for layout area and active power consumption.

Using the studied 130 nm technology as an example, the estimated performance of the 1-D scaling is listed in Table.3 comparing with that of the conventional 2-D scaling. The intrinsic delay $t_p$ for both two scalings are scaled to 40% of the original value. Although the layout area of a 1-D scaled standard cell is larger than that of

**Table 2.** Scaling relationships for the 1-D and 2-D scalings. The scaling factor for critical dimension and voltage are $S$ and $U$ respectively for both the 1-D and the 2-D scaling. Since the scaling factor for layout area in a gridded layout is not determined by the factor $S$, it is represented by $E$ in the list.

| Parameter | Relation | One-D scaling | Two-D scaling |
|---|---|---|---|
| $W$ | | 1 | $1/S$ |
| $L$, $t_{ox}$ | | $1/S$ | $1/S$ |
| $V_{DD}, V_T$ | | $1/U$ | $1/U$ |
| $A$ (cell area) | | $1/E$ | $1/S^2$ |
| $t_p$(intrinsic) | $C_{Load}V/I_{av} = L^2/V$ | $U/S^2$ | $U/S^2$ |
| $P_{av}/Device$ | $C_{Load}V^2/t_p = C_{ox}V^3(W/L)$ | $S^2/U^3$ | $S/U^3$ |
| $P_{av}/Area$ | $C_{Load}V^2/t_p/A = (C_{ox}V^3/A)(W/L)$ | $S^2E/U^3$ | $S^3/U^3$ |

**Table 3.** Scaling effects for a 130 nm technology. The minimum gate length $L$ is scaled from 130 nm to 70 nm by the gridded layout. The average cell area is about 80% of original cells.[23] Assuming that the supply voltage is decreased from 1.2 V to 0.9 V for the new library and the $V_T$ is scaled by the same factor, $S = 130/70$, $U = 1.2/0.9$, and $E = 1/0.8$.

| Parameter | One-D scaling | Two-D scaling |
|---|---|---|
| $W$ | 1 | 0.54 |
| $L$, $t_{ox}$ | 0.54 | 0.54 |
| $V_{DD}, V_T$ | 0.75 | 0.75 |
| $A$ (cell area) | 0.8 | 0.3 |
| $t_p$(intrinsic) | 0.4 | 0.4 |
| $P_{av}/Device$ | 1.4 | 0.8 |
| $P_{av}/Area$ | 1.8 | 2.7 |

**Table 4.** Three major components of leakage current. The magnitudes of each of these components depend strongly on the device geometry (namely, channel length ($L$), transistor width ($W$), and oxide thickness ($t_{ox}$)) and the doping profiles. In the equations, $K_g$, $K_s$, $\alpha$, and $n$ are determined by the technology parameters, such as the doping profiles, $V_\theta$ is the thermal voltage, $S$ is the surface of the source and drain to substrate junction, and $J(S)$ is the tunnelling current density through the surface.

| Leakage components | Description | Relation |
|---|---|---|
| $I_{gate}$ | Gate leakage | $= K_g W L (V/t_{ox})^2 e^{-\alpha t_{ox}/V}$ |
| $I_{sub}$ | Subthreshold leakage | $= K_s (W/L) e^{-V_{th}/nV_\theta}(1 - e^{-V/V_\theta})$ |
| $I_{BTBT}$ | Band-To-Band-Tunnelling (BTBT) leakage | $= \iint_S J(S) dS\big\vert_{source,drain}$ |

a 2-D scaled cell, the active power consumption $P_{av}$ per unit area is smaller for a 1-D scaled cell, which alleviates the challenge for the chip's cooling problem.

### 4.2. Leakage current

Although the supply voltage scaling has reduced the active power consumption $P_{av}$ per device, the leakage current increases drastically with technology scaling[24] and leakage power becomes a major contributor to the total power. The total leakage current ($I_{OFF}$) is influenced by many mechanisms and the leakage contributions from all of the sources must be taken into consideration to fully benefit from the 1-D scaling technique.

Among different leakage mechanisms in scaled devices, three major ones can be identified as: gate leakage, subthreshold leakage, and reverse biased source and drain to substrate junction Band-To-Band-Tunnelling (BTBT) leakage[25] as shown in Table.4. The magnitudes of each of these components depend strongly on the device geometry ($W$, $L$, and $t_{ox}$) and the doping profiles.[25]

Comparing with the 2-D scaling, the only disadvantage of the 1-D scaling is the increase of $W/L$ for all narrow MOSFETs because $W$ is not scaled with the scaling of $L$. Designers often use the combined $W/L$ ratio of all transistors as a convenient measure of total subthreshold leakage. The increased $W/L$ aggravates the increase of the subthreshold leakage $I_{sub}$ which originally results from the scaling of the threshold voltage $V_{th}$. However, since the total leakage current ($I_{OFF}$) is the sum of all components, a minimized $I_{OFF}$ can be reached by a carefully selection of the doping profile which has a smaller $I_{sub}$ per unit $W/L$. Plus the fact that the layout area of a 1-D scaled circuit is larger than that of a 2-D scaled circuit, the leakage current per area of the 1-D scaled circuit can be smaller than that of a 2-D scaled circuit.

## 5. CONCLUSIONS

The gridded layout enables more effective use of resolution enhancement techniques, which in turn allows a scaling of the technology. The effects on layout design, manufacturing cost, scaling performance, and circuit performance must be taken into consideration to fully benefit from the gridded-layout-based techniques.

Different gridded-layout-based lithographic methods have different effects on layout area and manufacturing cost. Designers have to make a choose between them. The application of the gridded layout on standard cells results in a 1-D scaling of the technology which leads to a linear relationship between cell area and the transistor pitch. However, layout area does not always scaled proportionally with the scaling of the technology-enabled minimum transistor pitch. Whatever to use half of the technology-enabled minimum transistor pitch or half of the transistor gap as the grid pitch should be determined by an examine.

Application of the grided layout on standard cells increase the $W/L$ ratio of narrow transistors which aggravates the increase of the subthreshold leakage $I_{sub}$ which originally results from the scaling of the threshold voltage $V_{th}$. A minimized $I_{OFF}$ for 1-D scaled standard cells can be reached by a carefully selection of the doping profile which has a smaller $I_{sub}$ per unit $W/L$.

## REFERENCES


1. G. E. Moore, "Lithography and the future of Moore's law," in *Advances in Resist Technology and Processing XII*, R. D. Allen, ed., *Proc. SPIE* **2438**, pp. 2–17, 1995.
2. "2001 international technology roadmap for semiconductors," tech. rep., Semiconductor Industry Assn., San Jose, CA, U.S.A., 2001.
3. B. J. Lin, "The attenuated phase-shifting mask," in *Solid State Technology*, **35**(1), pp. 43–47, 1992.
4. J. Garofalo, K. K. Low, O. Otto, C. Pierrat, P. K. Vasudev, and C. Yuan, "Automatic proximity correction for 0.35 $\mu$m i-line photolithography," in *International Workshop on Numerical Modeling of Processes and Devices for Integrated Circuits*, pp. 92–94, 1994.
5. L. Liebmann, G. Northorp, J. Culp, L. Sigal, A. Barish, and C. Fonseca, "Layout optimization at the pinnacle of optical lithography," in *Design and process integration for microlithography manufacturing II*, A. Starikov, ed., *Proc. SPIE* **5042**, pp. 1–14, 2003.
6. M. Levenson, N. Viswanathan, and R. Simpson, "Improvement resolution in photolithography with a phase-shifting mask," in *IEEE Trans. on Electron Devices*, **29**(12), pp. 1812–1846, 1982.
7. S. Mansfield, "Lithographic comparison of assist feature design strategies," in *Proc. SPIE - Int. Soc. Opt. Eng. (USA)*, **4000**, pp. 63–76, 2000.
8. L. Liebmann, J. Lund, F. L. Heng, and L. Graur, "Enabling alternating phase shifted mask designs for a full logic gate level," in *Journal of Microlithography, Microfabrication and Microsystems*, **1**, pp. 31–42, April 2002.
9. S. M. Mansfield, L. W. Liebmann, A. F. Molless, and A. K. Wong, "Lithographic comparison of assist feature design strategies," in *Optical Microlithography XIII*, C. J. Progler, ed., *Proc. SPIE* **4000**, pp. 63–76, 2000.
10. J. Garofalo, C. J. Biddick, R. L. Kostelak, and S. Vaidya, "Mask assist off-axis illumination technique for random logic," in *J. Vac. Sci. Technol. B*, **11**, pp. 2651–2658, 1993.
11. J. Garofalo, O. Otto, R. Cirelli, R. L. Kostelak, and S. Vaidya, "Automated layout of mask assist-features for realizing 0.5 $k_1$ ASIC lithography," in *Optical/Laser Microlithography VIII*, T. A. Brunner, ed., *Proc. SPIE* **2440**, pp. 302–312, May, 1995.
12. M. Fritze, B. Tyrrell, R. D. Mallen, and B. Wheeler, "Dense only phase-shift template lithography," in *Design and Process Integration for Microelectronic Manufacturing*, A. Starikov, ed., *Proc. SPIE* **5042**, pp. 15–29, 2003.
13. A. Rosenbluth, S. Bukofsky, C. Fonseca, M. Hibbs, K. Lai, A. Molless, R.Singh, and A. K. Wong, "Optimum mask and source patterns to print a given shape," in *Optical Microlithography XIV*, C. J. Progler, ed., *Proc. SPIE* **4346**, pp. 486–502, 2001.
14. M. Fritze, B. Tyrrell, R. D. Mallen, B. Wheeler, P. Rhyins, and P. Martin, "Optical imaging properties of dense shift feature patterns," in *J. Vac. Sci. Technol. B*, **20(6)**, pp. 2589–2596, Nov/Dec 2002.
15. A. Suzuki, K. Saitoh, and M. Yoshii, "Multilevel imaging system realizing $k_1 = 0.3$ lithography," in *Optical Microlithography XII*, L. V. den Hove, ed., *Proc. SPIE* **3979**, pp. 396–407, 1999.
16. S. Nakao, A. Nakae, A. Yamaguchi, H. Kimura, Y. Ohno, Y. Matsui, and M. Hirayama, "$0.12\mu m$ hole pattern formation by KrF lithography for giga bit DRAM," in *Tech. Dig. - Int. Electron Devices Meet.*, pp. 61–64, 1996.
17. S. Nakao, A. Tokui, K. Tsujita, I. Arimoto, and W. Wakamiya, "Random pattern formation by attenuated non-phase-shift assist pattern mask," in *Optical Microlithography XIV*, C. J. Progler, ed., *Proc. SPIE* **4346**, pp. 778–786, 2001.
18. B. Tyrrell, M. Fritze, D. Astolfi, R. Mallen, B. Wheeler, P. Rhyins, and P. Martin, "Investigation of the physical and practical limits of dense-only phase shift lithography for circuit feature definition," in *Journal of Microlithography, Microfabrication and Microsystems*, **1(3)**, pp. 243–252, October 2002.
19. D. A. Chathey and J. B. Rolfson, "Method of phase shift lithography," U. S. Patent No. 5,766,829, 1998.
20. M. D. Levenson, J. S. Peterson, D. G. Gerold, and C. A. Mack, "Phase Phrist! An improved strong-psm paradigm," in *20th Annual BACUS Symposium on Photomask Technology*, B. J. Grenon and G. T. Dao, eds., *Proc. SPIE* **4186**, pp. 395–404, September 2000.

21. J. Wang, A. K. Wong, and E. Y. Lam, "Standard cell layout with regular contact placement," in *IEEE Trans. on Semiconductor Manufacturing*, **17(3)**, pp. 375–383, August 2004.
22. J. Wang, A. K. Wong, and E. Y. Lam, "Standard cell design with regularly-placed contacts and gates," in *Design and process integration for microelectronic manufacturing III*, L. W. Liebmann, ed., *Proc. SPIE* **5379**, pp. 55–66, Feb 2004.
23. J. Wang, A. K. Wong, and E. Y. Lam, "Standard cell design with RET-driven regularly-placed contacts and gates," in *SPIE Journal of Microlithography, Microfabrication and Microsystems*, 2004/2005.
24. K. Roy, "Leakage current mechanisms and leakage reduction techniques in deep-sunmicron CMOS circuits," in *Proceedings of IEEE*, Feb. 2003.
25. S. Mukhopadhya, A. Raychowdhury, and K. Roy, "Accurate estimation of total leakage current in scaled CMOS logic circuits based on compact current modeling," in *Proceedings of DAC 2003*, pp. 169–174, Jun. 2003.