Chapter 5

Yield-Optimal Cell Layout Synthesis for CMOS Logic Cells

5.1 Introduction

The recent improvement of VLSI process technologies enables us to integrate a large number of transistors on one chip, and significantly improves the circuit performance. On the other hand, the methodology of VLSI design becomes more and more complex and some new problems, such as Design For Manufacturability (DFM) have arisen. Due to the very high costs associated with the manufacturability of sub-micron integrated circuits, even a modest yield improvement can be extremely significant. In order to achieve the high yield, a standard-cell layout synthesis considering the DFM is required since standard-cells are the most basic elements of the cell-based design methodology as described in Chapter 1.

This chapter describes a comprehensive CMOS logic cell layout synthesis technique for yield optimization by minimizing the sensitivity to wiring faults due to spot defects. Spot defect is one of the main sources of electrical failure in VLSI integrated circuits. We modeled the sensitivity to wiring faults on intra-cell routings with consideration to the spot defects size distribution and the end effect of critical areas. Critical area is defined as the area in which the center of a spot defect must fall to cause a fault. We comprehensively generate the minimum-width layouts of CMOS logic cells and the exact optimal layouts are selected from all the possible minimum-width layouts by using our model of the sensitivity to wiring faults as a cost metric. Although the critical area used for the sensitivity calculation is extracted from the original layout patterns, the feasibility of the proposed sensitivity model to the practical lithography system is discussed. Moreover, the adequacy of the proposed sensitivity to the other cost metrics such as the cell delay and the total intra-cell wire length is demonstrated. The impact of the sensitivity reduction on the yield is also discussed in this chapter.

At first, Section 5.2 explains our yield cost metric for intra-cell routings and our layout styles are described in Section 5.3. Our comprehensive layout synthesis method and overall system are explained in Section 5.4 and Section 5.5 respectively, and experimental results are shown in Section 5.6. Section 5.7 discusses the relation of the critical area between before and after lithography, and the adequacy of the proposed sensitivity. In addition, the impact of the cost metric used in the proposed method on yield is also discussed in this section. Finally, Section 5.8 summarizes this chapter.

5.2 Wiring Fault Model for Yield Cost Function

Critical area is defined as the area in which the center of a spot defect must fall to cause a fault. Random defects caused by particulate contamination have been typically the dominant reason for yield losses[32]. Therefore, the correct estimation of the critical areas plays an important role in layout sensitivity to spot defects and yield prediction. The spot defect size distribution and the critical area are defined as $D(x)$ and $A(x)$ respectively, where x is the spot defect size. Assuming the area distribution of defect density is uniform and written as P_0, the fault probability P is expressed as:

$$P = P_0 \int_{min}^{max} D(x)A(x)dx \tag{5.1}$$

where *min* and *max* are the minimum and the maximum defect size[33]. The spot defects size distribution function can be assumed to be

$$D(x) = \frac{{X_0}^2}{x^3} \tag{5.2}$$

where X_0 is the peak defect size of the distribution[34]. We consider only a bridging fault of two wire segments on the same layer due to a spot defect, which is called fault type OE (One-layer Extra-material defect). Since it is impossible to define the shape of a real spot defect, the shape of defect is assumed to be rectangular for simplicity. The critical areas of two wire segments whose width are w and spaced by d with the defect size x are illustrated in Figure 5.1. The slashed area of Figure 5.1 (a) illustrates the critical area between two parallel wire segments. This type of critical area is defined as $L(x)$. Assume that the length of overlapped section is l, the critical area is expressed as $l \times (x - d)$ when the end effect is neglected[33]. We take the end effect into account so that the model can be applicable to more complex intra-cell routings, and the back-slashed areas are added to the critical area

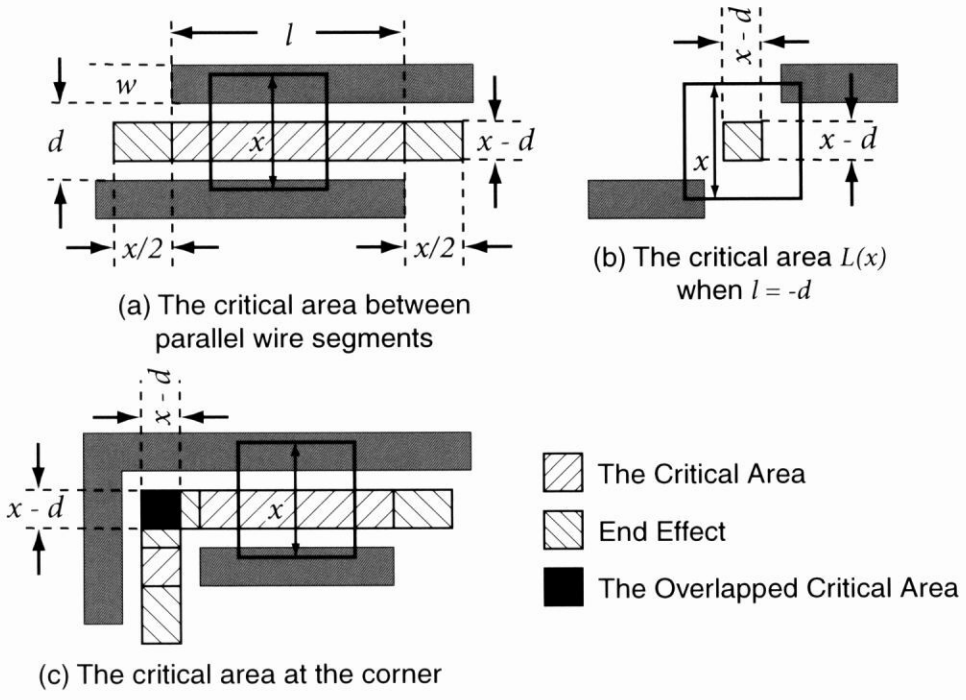


Figure 5.1 Critical areas between two wire segments spaced by d with the defect size x.

$L(x)$ which is newly expressed as

$$L(x) = (x + l) \times (x - d), \quad (l \geq -d). \tag{5.3}$$

$L(x)$ is calculated in the range of $l \geq -d$. When $l = -d$, the shape of the critical area $L(x)$ becomes square as illustrated in Figure 5.1 (b). Figure 5.1 (c) illustrates the critical area at the "L" type corner. Because the end sections of two critical areas are overlapped here, the area illustrated as a black square is counted twice. Therefore, this area is subtracted once from the critical area. If the corner type is "T" or "+", the area of the square is subtracted two or three times respectively. The black square area $R(x)$ is expressed as

$$R(x) = (x - d)^2. \tag{5.4}$$

The total critical area is calculated using these equations when the size of the spot defect is x. The probabilities of fault which results from each critical area is calculated using the equation (5.1) and the defect size distribution function (5.2) as follows:

$$P_L = P_0 {X_0}^2 \int_{min}^{max} \frac{1}{x^3} L(x) dx \tag{5.5}$$

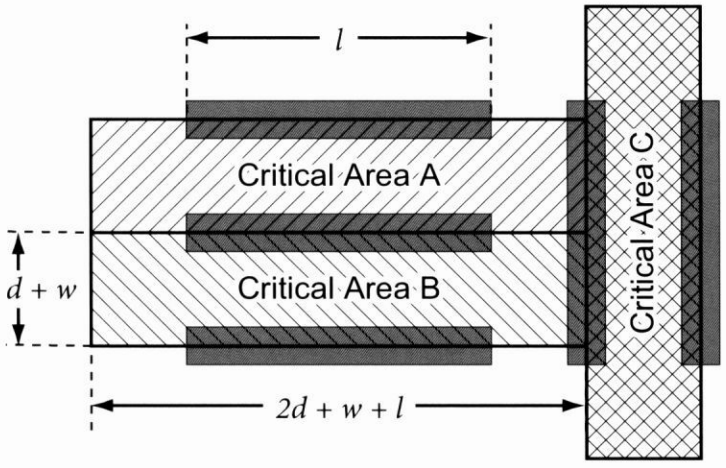


Figure 5.2 Critical areas between two wire segments spaced by d when the defect size x is larger than $2d + w$.

where P_L is the probability of fault which results from the critical area $L(x)$. Since P_0 and X_0 are process-dependent constant in this equation, these factors are excluded in the following discussion. The integral value without $P_0 X_0^2$ is defined as a *Sensitivity* to spot defect and written as S_L. If the defect size x is smaller than d, the defects will not cause any fault because of the zero critical area. If $d \leq x < 2d + w$, the critical area $L(x)$ is expressed as the equation (5.3). The critical areas when the defect size x is equal to $2d + w$ are illustrated in Figure 5.2. For $2d + w < x$, the adjacent critical areas have overlapped sections and these sections are calculated redundantly. Since a standard-cell layout commonly has dense wiring, this causes overestimation of the sensitivity. Therefore, we assume the critical area is saturated to $(2d + w + l) \times (d + w)$ for $2d + w \leq x$. As a result, the sensitivity is expressed as

$$S_L = \int_d^{2d+w} \frac{(x + l)(x - d)}{x^3} dx + \int_{2d+w}^{max} \frac{(2d + w + l)(d + w)}{x^3} dx. \tag{5.6}$$

By setting max to ∞, we obtain

$$S_L = \ln \frac{2d + w}{d} + \frac{l - d}{2}\left(\frac{1}{d} - \frac{1}{2d + w}\right). \tag{5.7}$$

By calculating in the same manner, we also obtain the sensitivity S_R which results from the critical area $R(x)$ as follows:

$$S_R = \ln \frac{2d + w}{d} - \frac{d + w}{2d + w}. \tag{5.8}$$

These sensitivities are calculated between two different wire segments only when these are placed on the neighboring grids as a result of a grid-based intra-cell routing. The total sensitivity is calculated as follows:

Table 5.1 Our layout styles of the proposed comprehensive cell layout synthesis method.

1. Static CMOS logic circuits.
2. Transistors are drawn up in two horizontal rows, the upper row for P type transistors and the bottom row for N type transistors.
3. Two transistors which have the same gate input signals are vertically aligned.
4. All transistors are uniform-sized.
5. The intra-cell routing uses only first metal layer.
6. VDD are connected from the top of P-diffusion to the top boundary by the vertical first metal.
7. GND are connected from the bottom of N-diffusion to the bottom boundary by the vertical first metal.
8. A single contact hole is assumed to be enough to connect metal and diffusion or polysilicon.

1. Enumerate the adjacent parallel wire segments and calculate the sensitivity using the equation (5.7) for each wire segment.

2. Enumerate the overlapped critical areas at the corner and calculate the redundant sensitivity using the equation (5.8) for each overlapped critical area.

3. Subtract the sum of the redundant sensitivities from the sum of the sensitivities of the adjacent parallel wire segments.

The sensitivity calculated using our model is more accurate than that using the conventional model which does not take the end effect of the critical area into consideration. The conventional model neglects about 40% of the sensitivities compared with our model. Therefore, our model is more suitable for the cell layout which has dense and complex intra-cell wirings.

5.3 Layout Styles

Our layout styles are described in Table 5.1. The layout styles No. 1 through 3 for the transistor placement are based on the transistor placement styles defined in Chapter 2, since the cell layout synthesis method in this chapter uses the minimum-width transistor placement method proposed in Chapter 2. Although we assumed that all transistors are uniform-sized for simplicity in this chapter, it is easy to extend it to handle multiple-sized transistors. The

layout styles No. 5 through 8 in Table 5.1 are for intra-cell routing. We use only first metal layer for intra-cell routing since it is suitable for standard-cell layouts. A single contact hole is assumed to be enough to connect the metal and diffusion or polysilicon due to the commonly used silicidation.

5.4 Comprehensive Cell Layout Synthesis

5.4.1 Transistor Placement

Our method utilizes the procedure proposed in Chapter 2 to generate the minimum width transistor placements. In Chapter 2, the minimum width placements are generated one by one until a routable placement is found via Boolean satisfiability. To generate all possible minimum-width placements comprehensively, we repeatedly generate a minimum-width placement until no other placement is found for the width of the minimum-width placement. Although the layout style used in this chapter assumes that two transistors which have the same gate input signals are vertically aligned for simplicity, the comprehensive intra-cell routing method explained in the following section is applicable to the placement which has a vertically aligned P and N type transistor pair with different gate input signals. Therefore, we can also use the minimum-width transistor placement method for non-dual cells proposed in Chapter 4 to generate all possible minimum-width placements, though the proposed comprehensive cell layout synthesis method uses the transistor placement method only for dual cells in this chapter.

5.4.2 Intra-Cell Routing

The proposed cell layout synthesis method uses a comprehensive intra-cell router to realize an exhaustive pattern generation for intra-cell routing. Conventionally, several exact routing algorithms have been proposed[23, 35, 36]. Among them, [35] can generate all routing patterns exactly, but can not solve larger problems than 4×4 grids. On the other hand, [36] can solve larger problems, but this method has some restrictions and the search space of this method is too small for general cell generation. In this chapter, we propose a new exact routing algorithm that can solve problems which is large enough to generate standard cell routing patterns and its search space is much larger than that of [36] by using more practical restrictions for standard cells. In our method, a generated placement is transformed into a grid model problem illustrated in Figure 5.3 (a). When it is given to our router, the grids for

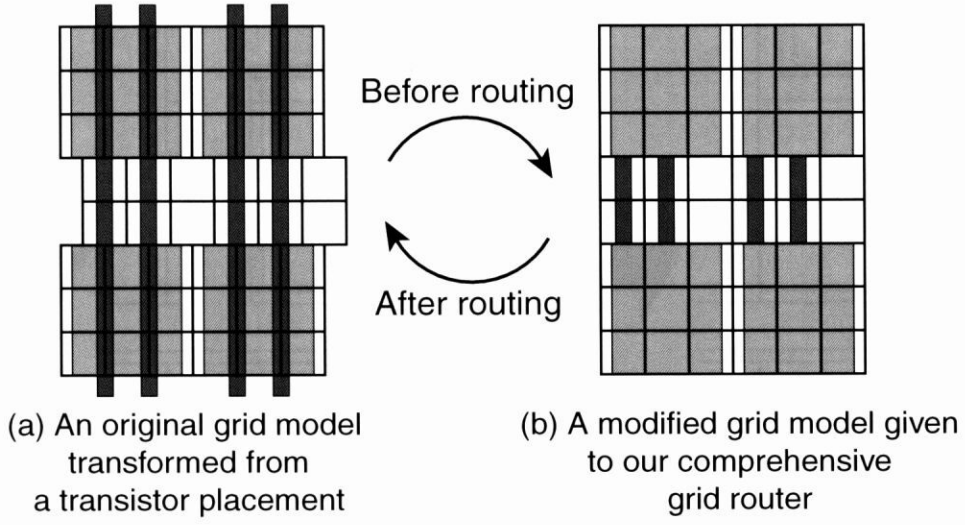
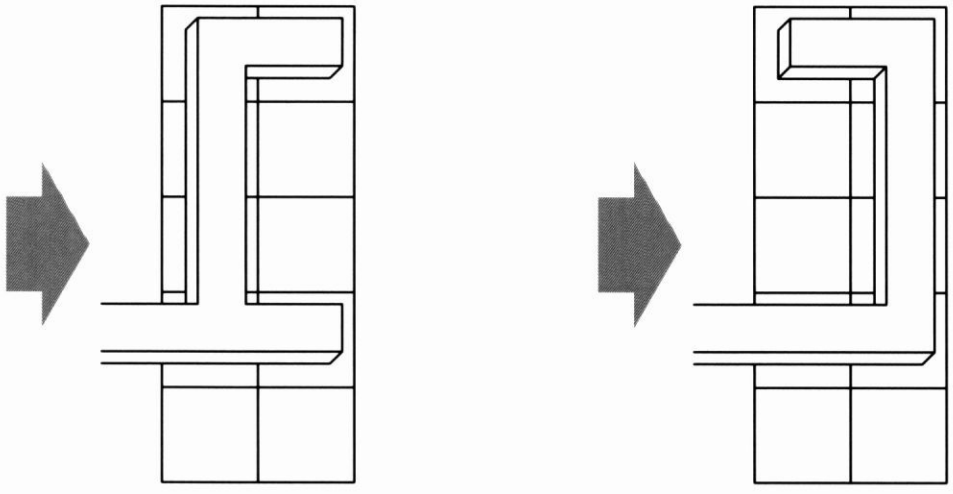


Figure 5.3 Grid models used in our comprehensive intra-cell routing system.

gate, between the P and N diffusions, are shifted half grid to the left, since our method routes the problems of complete grid model as illustrated in Figure 5.3 (b). Our method routes a given grid model problem from left to right, column by column. Inside each column, all possible patterns are generated for each pattern of the previous column considering VDD/GND terminals, diffusion silicides, gate, and I/O. VDD terminals must be connected from the top of P diffusion to the top boundary of routing area, and GND terminals from the bottom of N diffusion to the bottom. A single contact hole is assumed to be enough to connect the metal and the terminals on the diffusion or polysilicon because of the silicides. When redundant via holes are needed, it is easy to add them to all generated routing patterns after routing is finished and select the optimal one. If a gate terminal is an input port and is not connected to other terminals by metal layer, at least one grid of this terminal must be reserved so that an I/O contact can be placed on it. It also considers the design rules so as not to generate the routing patterns which violate metal design rules when the grids are transformed back to the original grid position as shown in Figure 5.3 (a). The proposed intro-cell router has a constraint for search space reduction as follows:

- A single net can not fork when it extends to the subsequent column.

Figure 5.4 illustrates examples about this restriction. The routing pattern of Figure 5.4 (a) is not generated by the proposed method since a single net forks to extend to the subsequent column and this pattern violates the above restriction. On the other hand, the pattern of Figure 5.4 (b) is allowed because two nets, which is not connected yet, are extended to the



(a) Constrained pattern. (b) Allowed pattern.

Figure 5.4 Wire branching constraint of the proposed comprehensive intra-cell routing method.

subsequent column individually and connected inside the subsequent column.

Figure 5.5 illustrates the flow of our intra-cell routing. The detail of each procedure is described as follows. The numbers below corresponds to the numbers in Figure 5.5. Each procedure is explained using a sample problem illustrated in Figure 5.6. This figure shows that the nets on the left are extended to the leftmost column. The thick rectangles in this figure indicate the terminals on source or drain, and the narrow rectangles indicate those on gate. The nets and terminals that have the same number have to be connected to each other on at least one grid.

1. At first, the pattern to be routed in the current column is determined by the terminal pattern of this column and the net pattern given from the previous column. If there are some terminals which are already connected, these terminals are removed from this pattern. Figure 5.7 (a) shows an example of this procedure. The terminal number 4 on the bottom row is removed because this terminal is already connected to the net by silicides.

2. Next, all the terminals which are not connected yet are found and all possible connection patterns are generated considering the silicides and space of I/O contact. In Figure 5.7 (a), the terminals number 5 and 6 can be connected inside this column. The patterns illustrated in Figure 5.7 (b), (c), (d), and (e) show all the possible connection patterns generated from Figure 5.7 (a).

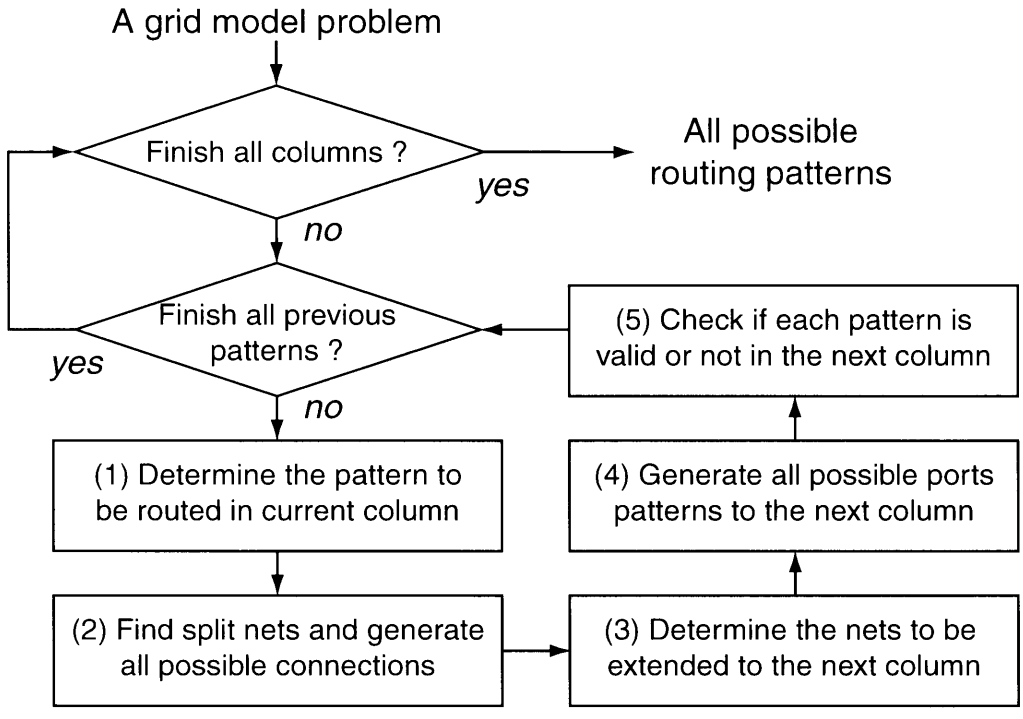


Figure 5.5 The flow diagram of our comprehensive intra-cell router.

3. Then, the nets which have to be extended to the subsequent column, *i.e.*, the nets with unconnected terminals, are determined for each pattern generated in the former step. For example of the pattern of Figure 5.7 (d), the nets number 3, 4, and 6 have other unconnected terminals in the following columns and the net number 5 is not finished connection yet. Therefore, these nets have to be extended to the next column.

4. All possible net patterns extended to the subsequent column are generated for each pattern generated in the step 3. Figure 5.8 illustrates all the possible patterns generated from the pattern of Figure 5.7 (d). Because of the constraint explained before, a single net can not fork when it extends to the subsequent column. Each net in this figure has at most one extension to the subsequent column.

5. Finally, check if all the generated patterns to the subsequent column are valid or not. If some nets of a generated pattern conflict to other connections of nets or I/Os, this pattern is impossible and removed from the group of generated patterns. For example of Figure 5.8, (d-1) and (d-2) have nets which conflict all the terminals of terminal number 4 in the subsequent column in Figure 5.6. Therefore, these patterns are removed and others are given to the subsequent column.

These procedures are applied for each pattern of each column and generate all possible routing patterns of a given grid-model problem under our layout styles.

5.5 Overall System

Using the procedures described in Section 5.4, we can generate all possible layouts comprehensively. Then calculate the sensitivity using the model described in Section 5.2 for each layout. To obtain the exact yield-optimal layouts, the proposed method does not use heuristic approaches. The proposed method selects the yield-optimal layout from all the possible minimum-width layouts. Although the sensitivity can be reduced by inserting redundant spaces, this chapter targets to generate the yield-optimal minimum-width cell layout. It is possible to optimize the yield further by adjusting the space between wires in a grid-less manner after the yield-optimal layouts are selected. Moreover, if there are spaces between cells after placing the generated cells, we can also improve the yield by increasing the width of a cell without any area overhead. The next chapter proposes a further yield optimization by inserting redundant spaces through a cell layout de-compaction in a grid-less manner. Overall flow of our yield-optimized cell layout synthesis is illustrated in Figure 5.9 and described as follows.

1. For a given transistor netlist, generate a minimum-width transistor placement using the procedure explained in Chapter 2. The number of diffusion gaps of this placement is defined as G.

2. Comprehensively route the generated placement using the procedure explained in Section 5.4.2. If this placement is routable, generate routed layouts. Otherwise no solution.

3. Find another placement with G gaps using the procedure explained in Chapter 2. If there is one, go to step 2 again. Otherwise go to step 4.

4. If there is no routed layout, increment G and generate a transistor placement with G gaps using the procedure explained in Chapter 2. Then go to step 2 again. Otherwise go to step 5.

5. Place the I/O contact holes for each generated layout if needed. If several possible I/O placement patterns can exist for one layout, generate all possible I/O patterns.

6. Calculate the sensitivity for each layout and find the optimal solution.

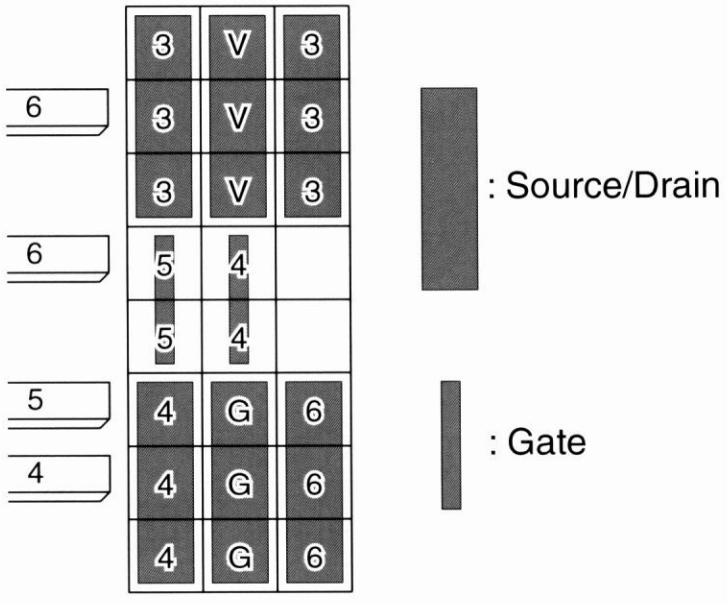


Figure 5.6 A sample problem for explaining our comprehensive intra-cell router.

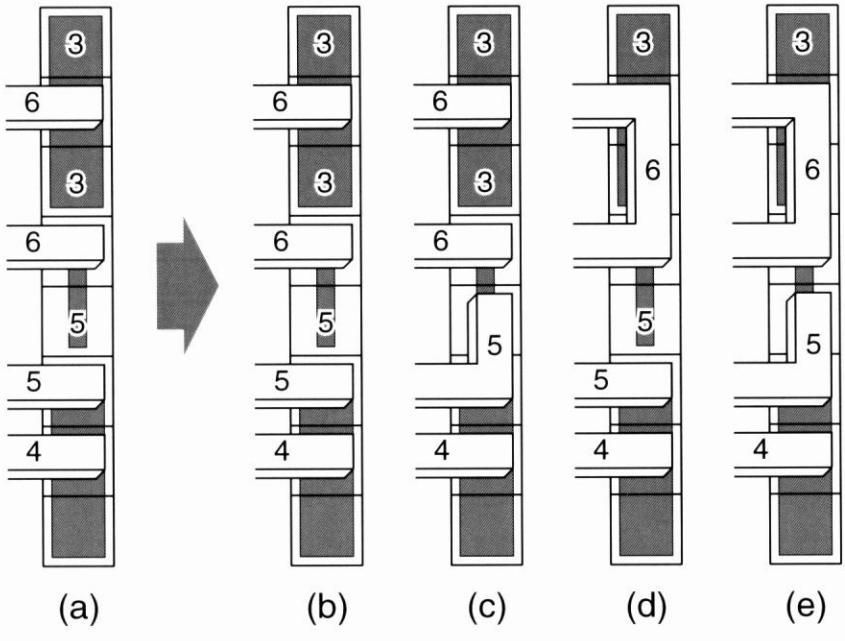


Figure 5.7 An example of all possible connection patterns inside a column.

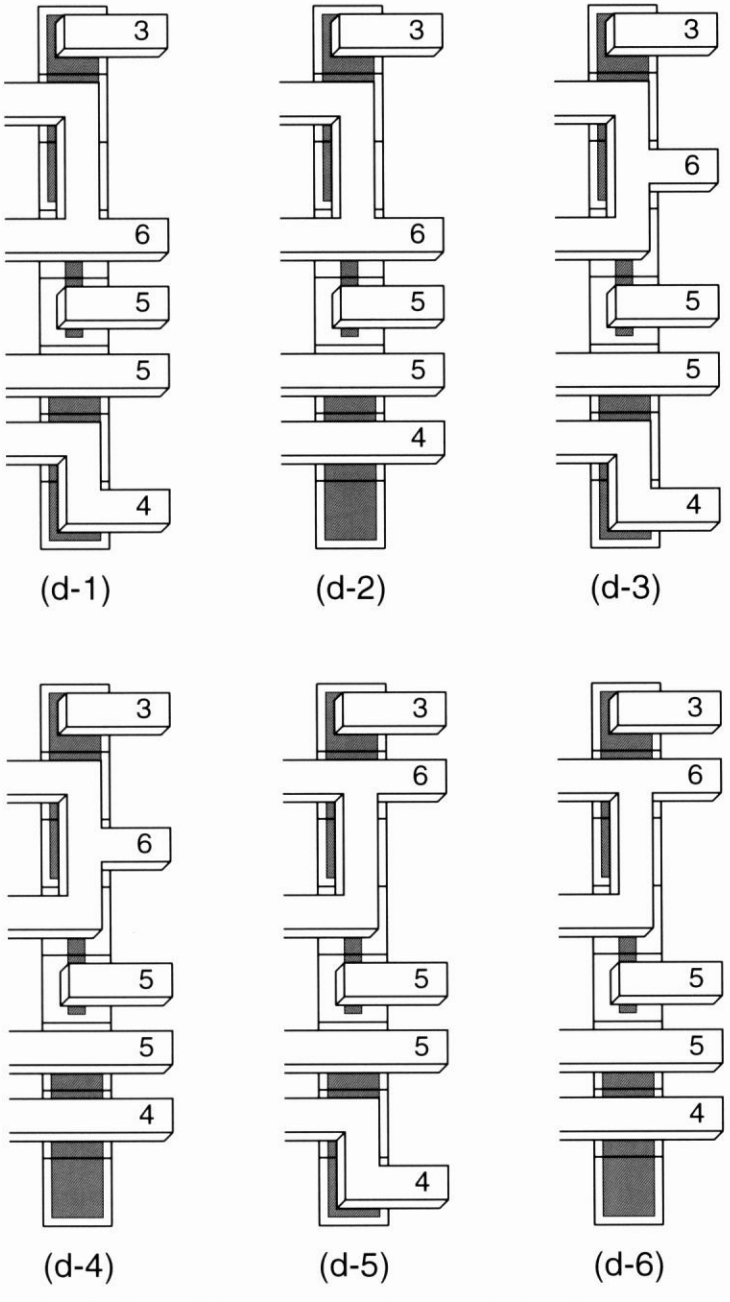


Figure 5.8 An example of all possible patterns to be extended to the subsequent column generated from the pattern (d) in Figure 5.7.

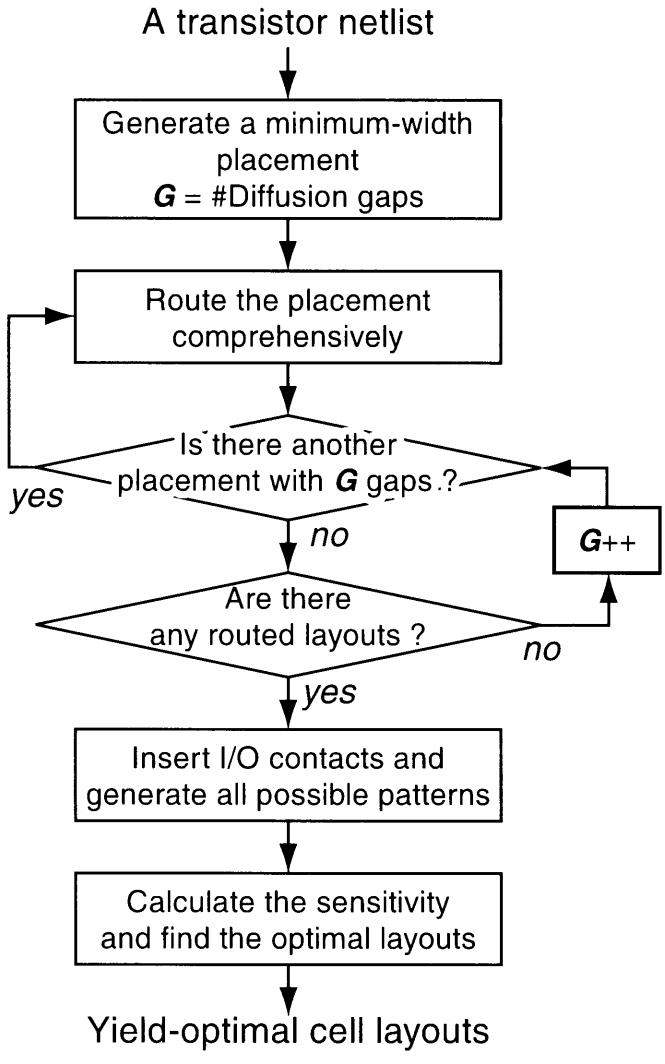


Figure 5.9 The flow diagram of our yield-optimal cell layout synthesis system.

5.6 Experimental Results

We applied our layout synthesis method to 8 CMOS logic circuits in a standard-cell library, which have up to 14 transistors. Tables 5.2 and 5.3 summarize the results. In this experiment, we assumed a 0.35μm process technology and used 8 rows for routing grids. For each circuit, these tables indicate the number of transistors, the number of columns of routing grids, the number of possible placements, the number of routable placements, the number of generated layouts, the CPU times for layout synthesis and for selecting the sensitivity-minimum layout, the minimum sensitivity value S_{min} and the average sensitivity value of all wire-length-minimum layouts S_{wire}. This table also shows the percentage of reduction in the sensitivity by selecting the sensitivity-minimum layout. This percentage is calculated by

Table 5.2 The results of the comprehensive cell layout synthesis.

Circuit name	*#tr.*	*#col.*	*#place*	*#route*	*#layout*	*Cell synth. CPU(sec.)*	*Selection CPU(sec.)*
ao222	14	9	32	8	948852	14685.04	323.70
aoi21	6	4	4	4	999	0.36	0.17
aoi211	8	5	4	4	8839	1.70	1.60
eno	10	6	2	1	19648	13.70	4.45
gen2	12	8	24	2	34120	419.98	10.38
mux2	12	9	144	4	392	6449.82	0.19
nand4	8	5	4	4	38366	3.01	6.92
xnor2	10	7	144	24	27414	2164.30	6.97

Table 5.3 The cell selection results of the comprehensive cell layout synthesis.

Circuit name	*Minimum sensitivity* S_{min}	*Average sens. wire-min.* S_{wire}	*Reduction ratio (%)*
ao222	18.04	20.55	12.19
aoi21	5.37	6.56	18.10
aoi211	6.08	7.13	14.77
eno	11.35	13.33	14.93
gen2	17.78	20.90	14.91
mux2	31.72	33.79	6.10
nand4	5.37	6.38	15.78
xnor2	11.18	15.27	26.81

$(S_{wire} - S_{min})/S_{wire} \times 100$. As shown in Table 5.2, our method generates the layout comprehensively in only a few seconds for circuits with up to 8 transistors, whereas the runtime is severely high for the circuit with 14 transistors. In the case of standard-cell layout synthesis, we are allowed to consume relatively long time to obtain high quality layouts. For larger circuits, however, we need to utilize some heuristic techniques such as netlist partitioning for transistor placement or branch-and-bound for intra-cell routing to solve them in reasonable time, even though these techniques do not guarantee the optimality.

In Figure 5.10, the minimum sensitivity value and the minimum, average, and maximum

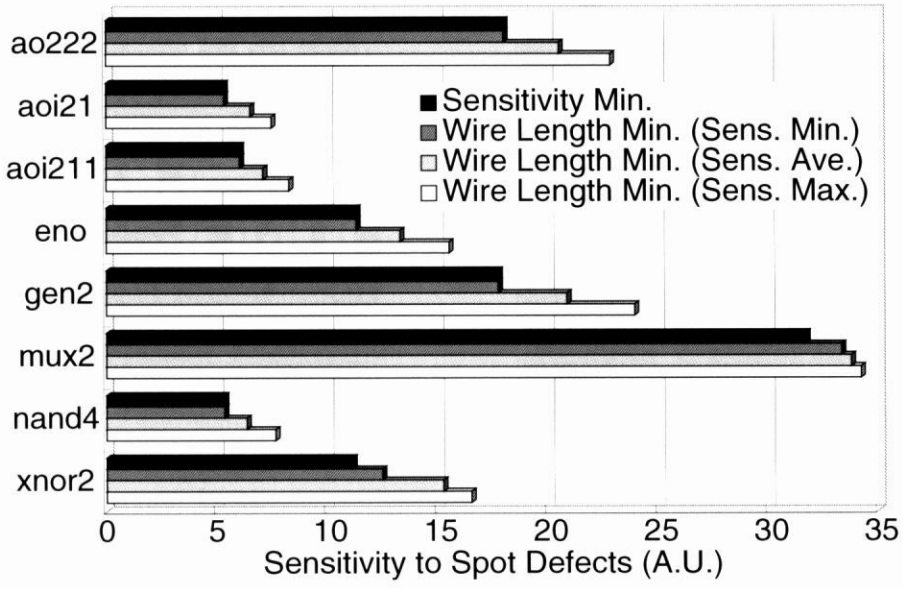
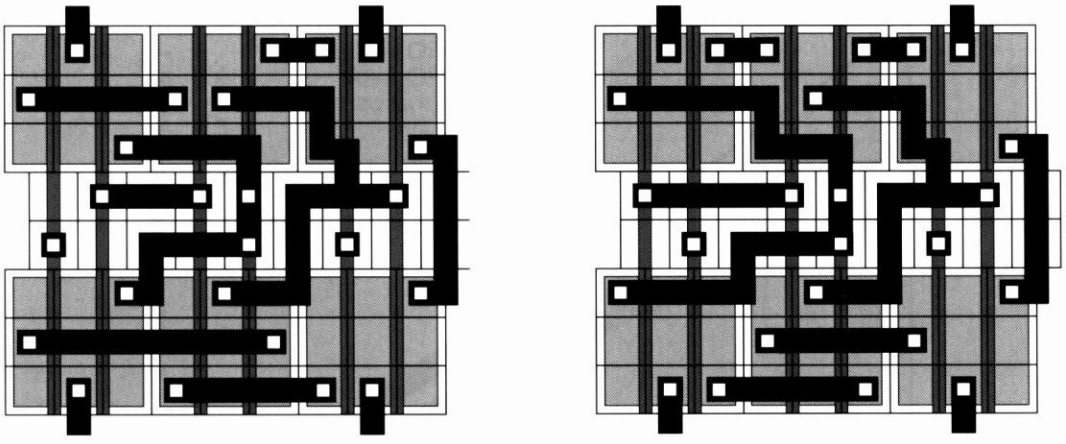


Figure 5.10 Changes in layout sensitivity to spot defects by selecting the sensitivity-minimum layouts.

sensitivity value of wire-length-minimum layouts are plotted for comparison. These results show that the wire-length-minimum layouts are not always the optimal layout for the sensitivity to spot defects. Compared with the average sensitivity value of all the wire-length-minimum layouts, our method can improve the sensitivity about 15% on an average of 8 circuits by picking up the optimal layout of sensitivity from all the generated layouts.

The snapshots of the wire-length-minimum layout and sensitivity-minimum layout of mux2 are illustrated in Figure 5.11 for example. The sensitivity optimal layout has smaller sensitivity whereas its wire length are longer than the wire length optimal layout. We also compared the cell delay of these two layouts. The sum of the delay of all the timing arc with no output capacitance was calculated in simulation. A timing arc is defined as a signal flow from an input to an output on a cell, *e.g.*, A rise $\rightarrow$ Y fall. The comparison result shows that the delay of the sensitivity-minimum layout is slower only 0.23% than that of wire-length-minimum layout. This result shows that there is little performance degradation between these two layouts and the layouts selected by the sensitivity are acceptable also in terms of cell delay.



(a) Wire length minimum
Sensitivity: 33.34
Wire length: 60.0

(b) Sensitivity to spot defects minimum
Sensitivity: 31.72
Wire length: 63.0

Figure 5.11 The optimal layouts of "mux2" in Table 5.2 generated by our method and selected by (a)wire length and (b)sensitivity to spot defects.

5.7 Discussion of Yield Cost Metrics

5.7.1 Lithography Impact on the Critical Area

The critical area used for modeling the sensitivity through this section is calculated from the original layout pattern of the first metal layer. However, the actual patterns printed on the silicon wafer usually changes due to lithography and/or Optical Proximity Correction (OPC). Therefore, it is necessary to clarify the impact of the lithography processes on the critical area. We used a cell library of a 90nm technology to show the effect of the lithography. Although we used cell layouts of a 0.35μm technology in the experimental results section, a 90nm cell library which includes the subwavelength feature size is used to highlight the optical proximity effects in this section. Figure 5.12 shows the relations between the critical area of the original layout patterns and the patterns after lithographic simulation before and after OPC, respectively. The reference line in this figure indicates $y = x$. The critical area plotted in this graph is calculated only for a bridging fault of two wire segments on the same layer due to a spot defect, which is called fault type OE (One-layer Extra-material defect). This graph plots the critical area of 340 cells in a 90nm cell library using the defect size of 1.5 times larger value than the minimum spacing of the first metal layer. Let D denote the defect size, the OE type critical area calculation procedure in this experiment is written as follows.

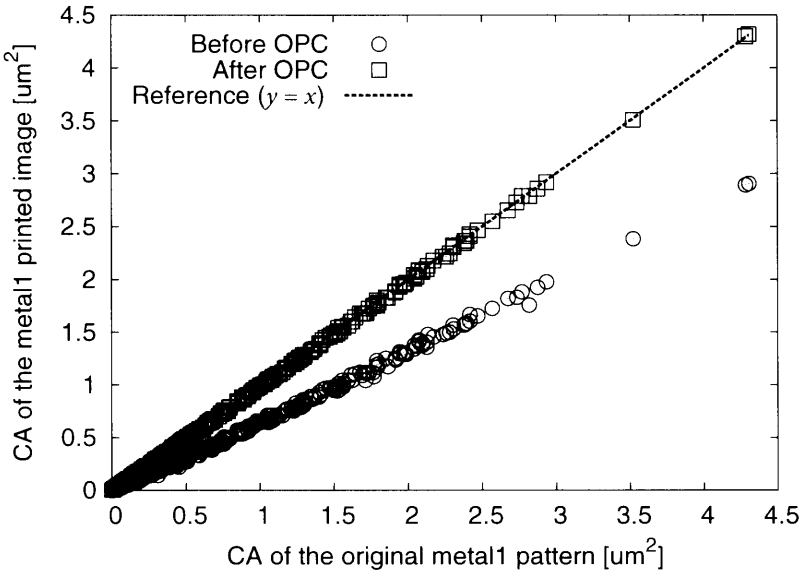


Figure 5.12 Lithography impact on the critical area before and after OPC for 340 cell layouts in a 90nm technology.

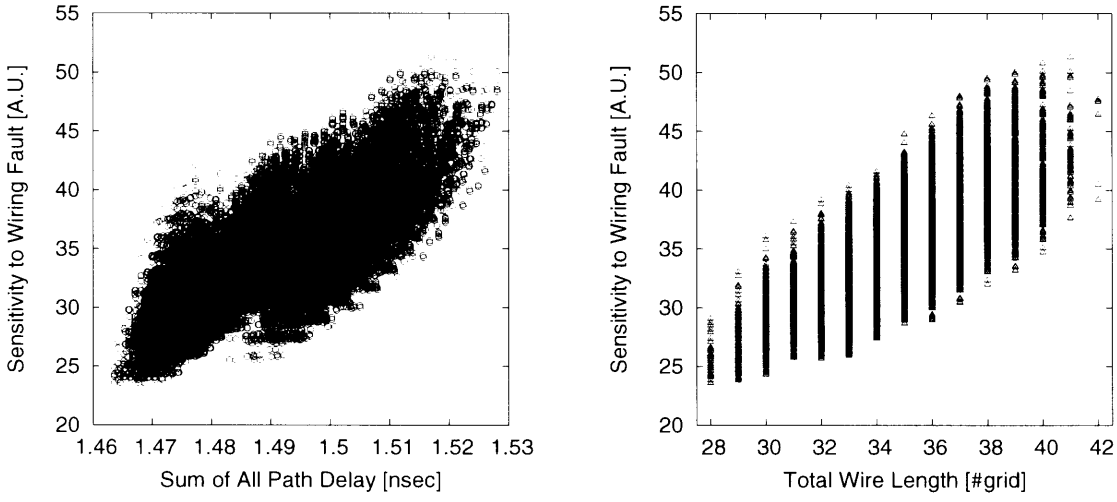
1. Expand the polygons on the first metal layer by $D/2$ moving the sides of the polygons perpendicularly.

2. Sum up the overlapped areas of the expanded polygons.

The printed images are obtained by the lithographic simulation using *Calibre OPC*[37]. The lithographic condition used in this experiment is summarized as follows.

- Lithography Wavelength: 193nm
- Numerical Aperture: 0.8
- Mask Reduction Factor: 4

As shown in this graph, the critical areas of the printed images after OPC are almost always equal to that of the original layouts under this condition. Both before and after OPC, the correlation coefficients between the critical area of the printed images and the original layouts are almost equal to 1. Therefore, the critical area of the printed images has strong correlation with the original critical area even without OPC. We can conclude from this result that the proposed model of sensitivity to wiring fault calculated from the critical area of the original layout pattern also has strong correlation with the sensitivity after lithography, and can be used to model the yield of the printed images on the silicon wafers.



(a) Cell delay and sensitivity (b) Wire length and sensitivity

Figure 5.13 The relation between two cost metrics in the case of the exhaustively-generated 27414 cell layouts of xnor2.

5.7.2 Relation Between Sensitivity and Performance

This section describes the relation between the proposed sensitivity to wiring faults and other cost metrics of cell layout, such as cell delay and total intra-cell wire length. Figure 5.13 (a) and (b) show the relation between cell delay and sensitivity, and wire length and sensitivity, respectively, in the case of the comprehensively-generated 27414 cell layouts of "xnor2" in Table 5.2. For each layout of xnor2, we simulated and summed up the delay of all the timing arcs, and plotted it in the graph (a). Both (a) and (b) show the positive correlation between two cost metrics, and the values of the correlation coefficient are 0.77 and 0.76, respectively. There are strong correlations both between cell delay and sensitivity, and wire length and sensitivity. Therefore, we can conclude that the selected layout by the sensitivity might have the acceptable quality in terms of the other cost metrics such as cell delay and total intra-cell wire length.

5.7.3 Relation Between Sensitivity and Yield

This section describes the relation between the sensitivity and yield. Given the critical area of the first metal layer, the yield of this layer can be modeled using a Poisson-based yield model[38]:

$$Y_{met1} = exp(-P_0 \int_{min}^{max} D(x)A(x)dx) \tag{5.9}$$

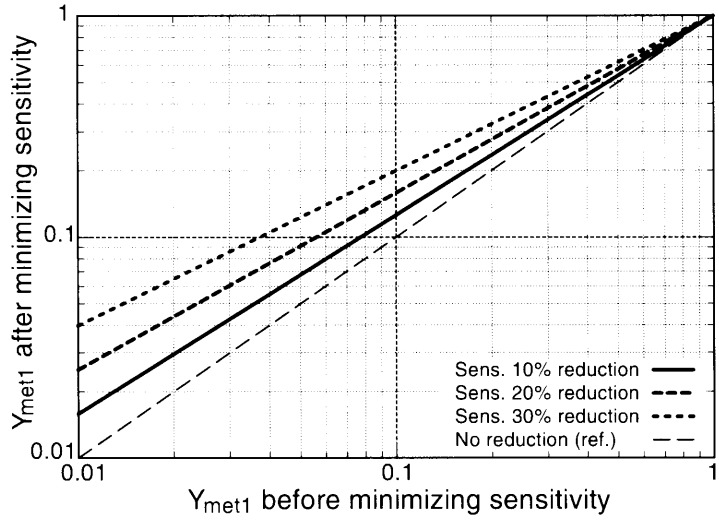


Figure 5.14 The impact of the sensitivity reduction on the yield of the first metal layer.

where Y_{met1} is the yield of the first metal layer. This equation can also be expressed as below:

$$Y_{met1} = exp(-P_0 X_0{}^2 \times S_{met1}) \tag{5.10}$$

where S_{met1} is the sensitivity of the first metal layer. We define the sensitivity reduction ratio σ and the yield improvement ratio ϕ as follows:

$$\sigma = \frac{S_{orig} - S_{met1}}{S_{orig}}, \tag{5.11}$$

$$\phi = \frac{Y_{met1} - Y_{orig}}{Y_{orig}} \tag{5.12}$$

where S_{orig} and Y_{orig} are the value of the sensitivity and yield before the sensitivity is minimized, respectively. Although $P_0 X_0{}^2$ is the process-dependent constant and is hard to be determined accurately, we can derive the relation between σ and ϕ using the equations (5.10), (5.11), and (5.12) as follows:

$$\phi = exp(-\sigma \log Y_{orig}) - 1. \tag{5.13}$$

Using this equation, the relation between the yield value before and after sensitivity minimization can be illustrated as Figure 5.14. This graph clearly shows that the impact of the sensitivity reduction on the yield improvement is greater when the original yield value Y_{orig} is small. For example of the sensitivity reduction is 20% case, the yield improvement ϕ is 0.20% when $Y_{orig} = 0.99$, while 4.56% when $Y_{orig} = 0.80$.

The total yield Y_{total} can be expressed as the product of all yield loss reasons like below[38]:

$$Y_{total} = \prod_{i=1}^{N} Y_i \tag{5.14}$$

where N is the number of yield loss reasons and Y_i is the yield value of each reason. Therefore, the yield improvement ratio ϕ is the same value for both the yield of the first metal layer and the total yield if the yield values of all the other reasons remain the same. From this point of view, we can conclude that our method is more effective when the total yield value of the manufacturing process is small, *i.e.*, the process is not mature yet.

5.8 Summary

An optimal cell layout synthesis technique to minimize the sensitivity to wiring faults has been presented in this chapter. The sensitivity to wiring fault due to spot defects for intra-cell routings was modeled considering the spot defects size distribution and the end effect of critical areas, and used as a cost function. The impact of the sensitivity reduction on the yield improvement is also described. Our cell layout synthesis technique generates the minimum width layouts of CMOS logic cells comprehensively, and selects the optimal layouts based on the cost functions. We applied our comprehensive layout synthesis method to 8 CMOS logic circuits which have up to 14 transistors and the results show that the fault sensitivities are reduced about 15% compared with the wire-length-minimum layouts. Our layout synthesis method will be applicable for deriving the optimal cell layouts by some other cost metrics, such as power, delay, and signal integrity, if reasonable cost functions are given.