Layout-Synthesis Techniques for Yield Enhancement

Venkat K. R. Chiluvuri, *Student Member*, and Israel Koren, *Fellow, IEEE*

Abstract—**Several yield-enhancement techniques are proposed for the last two stages of VLSI design, i.e., topological/symbolic and physical layout synthesis. Our approach is based on modifications of the symbolic/physical layout to reduce the sensitivity of the design to random point defects without increasing the area, rather than fault tolerance techniques. A layout compaction algorithm is presented and the yield improvement results of some industrial layout examples are shown. This algorithm has been implemented in a commercial CAD framework. Some routing techniques for wire length and via minimization are presented, and the results of wire length reduction in benchmark routing examples are shown. We demonstrate through topological optimization for PLA-based designs that yield enhancement can be applied even at a higher level of design abstraction. Experimental results show that it is possible to achieve significant yield improvements without increasing the layout area by applying the proposed techniques during layout synthesis.**

I. Introduction

CONTINUED advances in VLSI technology, along with the development of more sophisticated CAD tools, enable an increase in the level of integration of silicon chips. By integrating more and more circuits on a single chip, system performance improvements can be achieved. However, due to unavoidable manufacturing process variations/imperfections, chips must be limited to a certain size, beyond which the yield of the chip is so low that the product is no longer commercially viable. Processing imperfections such as point defects and line registration errors are introduced into the IC layers during the lithography process. The main sources of point defects are dust and other unwanted particles in the atmosphere of the clean room, chemicals, bulk gases and unionized water, lithography and other fabrication equipment. Some of these defects may result in missing patterns or open circuits, while other defects may result in extra patterns or short circuits.

During the past 15 years, feature sizes have diminished drastically from a few microns to submicrons, allowing integration of over a million transistors on a single chip. At the same time, manufacturing process complexity (i.e., number of lithography levels) has increased significantly. Die sizes of high-performance general-purpose microprocessors have already crossed 2 cm^2. For example, the die size of the Digital's Alpha 21064 microprocessor chip is 2.34 cm^2. Such large-area chips became a reality because defect densities dropped almost one order of magnitude during this period. It may be unlikely to have similar substantial improvements in manufacturing facilities in the near future to improve the yield. Therefore, further increase in the level of integration may result in chips with low manufacturing yields. For example, it may be extremely difficult to achieve a yield of over 40–50% for a 5-cm^2 chip even under mature manufacturing conditions.

Such yields cannot be expected by achieving factory performance goals alone [12]. Thus, new design techniques must be found in order to achieve further improvements in the yield of large-area chips. Localized changes in the layout generated by the general design rules may allow us to increase the yield without any area overhead. Our approach to yield enhancement is based on modifications of the layout to reduce the sensitivity of the design to point defects without increasing the area, rather than on fault tolerance techniques. Unlike the more traditional defect tolerance approach, which requires the development of special redundancy techniques for the given design, the proposed approach is applicable to all design styles and does not require any additional resources in terms of silicon area. Another important property is that the layout changes can be automated and made part of the physical design tools (e.g., compaction, routing) making them transparent to the designer.

A. Design for Yield

Several design techniques have been proposed for many stages of design synthesis for yield enhancement. A commonly-used design technique for yield enhancement is through defect tolerance, i.e., the incorporation of redundant circuits. A variety of fault-tolerant techniques have been proposed for memory IC's, PLA-based designs, and Wafer Scale Integrated Systems. These fault-tolerant techniques have proved to be very effective in certain situations (e.g., memory chips), but involve a cost of additional area and design effort. Also, defect tolerance techniques have been developed only for highly regular designs. There is still no general approach to the incorporation of defect tolerance in random logic design.

IC device parameters are very sensitive to unavoidable variations in manufacturing process. It is very important to maintain these device parameters within acceptable limits in order to guarantee the circuit performance. Yield degradation due to these global process variations is known as parametric yield. Several statistical design centering techniques have been developed for parametric yield optimization [28]. The objective of statistical design centering is to maximize the

Manuscript received March 1994; revised October 14, 1994. This work was supported by the National Science Foundation under Contract MIP-9305912.

I. Koren with the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA 01003 USA.

V. K. R. Chiluvuri was with the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA 01003 USA. He is now with the Semiconductor Systems Design Technology Group, Motorola, Austin, TX 78735 USA.

IEEE Log Number 9409814.

parametric yield of a circuit with respect to manufacturing process parameters.

For the physical layout design stage, the concept of 'Design for Yield' has been applied very successfully through global design rules and area minimization [32] from the very beginning. These design rules are optimized for minimizing the process yield losses due to global variations. However, in a mature manufacturing line, random point defects are the major source of yield losses. The effect of a defect on the chip is strongly dependent on where the defect is located. Therefore, the susceptibility of a chip depends on the layout density, where a denser layout is more susceptible to defects. So far, only limited attention has been given to point defects during the layout design. To maintain the yield of future chips with complexities exceeding 10 million transistors, the distribution of point defects and the sensitivity of the design to these defects must be taken into account during the layout synthesis. Therefore, by including defect sensitivity as an additional criterion at the physical design level for yield enhancement, substantially better results can be achieved.

Several mathematical models have been proposed for very accurate yield predictions. This high degree of accuracy is achieved, to a large extent, by replacing the chip area with critical area (which represents the defect sensitivity of the layout) in the yield models [14], [19], [20], [27], [33]. Recently, several methods have been proposed for estimating the yield of a chip from its layout detail. *Xlaser* and other yield analysis CAD tools are based on analytical models [15], [26]. Monte Carlo simulation-based yield prediction tools were also developed [35]. Although these tools are useful for yield analysis, they cannot be used to modify layouts for yield enhancement.

Only recently have researchers started reporting their work in the area of layout synthesis for yield enhancement. The first significant work in the area of layout modifications for yield improvement has been reported by Allan [1]. A set of local rules have been proposed for contacts, metal and polysilicon layers for yield enhancement. However, these techniques are not general enough to be applied in the regular physical layout-synthesis stages such as routing and compaction.

Some routing techniques for yield enhancement have been developed for two layer routing [22], [30]. These routing algorithms are based on minimization of the defect sensitivity of the layout. In the routers proposed in [30], the defect sensitivity of the second layer is not considered while minimizing the defect sensitivity of the first layer. Moreover, the criteria chosen for defect sensitivity were not based on any analytical models reported in the yield literature. These shortcomings were addressed to some extent in [22]. The major drawback is that if the layouts generated by these routers are compacted, their yield criteria becomes invalid. Some results have been reported for the placement and floorplan stages of layout synthesis as well [21].

B. Defect Sensitivity of a Layout

Researchers have proposed several yield models [11], [19] to predict the manufacturing yield. The three-parameter generalized negative binomial yield model given in (1) was found to match empirical results better than other yield models.

$$Y = Y_0(1 + dA\theta/\alpha)^{-\alpha} \tag{1}$$

where Y is the yield of the die, Y_0 is the gross yield factor, d is the average number of defects per unit area, A is the area of the die, θ is the probability that a defect will result in a circuit fault and α is the clustering parameter. In this model, A represents the total area of the die, while the product $A\theta$ (also called the *critical area*) represents the portion of the chip area that is sensitive to defects. In other words, not every defect results in a circuit failure. The effect of a defect on the chip is strongly dependent on where the defect is located. Therefore, the susceptibility of a chip depends on the layout density, where a denser layout is more susceptible to defects. Thus, layout design rules for a given fabrication process have a strong impact on the yield. These design rules are formulated in such a way that global disturbances, such as misalignment of the masks, linewidth variations of the poly and diffusions, lateral diffusion on the diffusion line, etc., may have a minimal effect, and the amount of logic per chip is maximized.

Since all of the above disturbances are mainly the result of process variations, these layout rules are targeted to maximize the gross yield, Y_0. So far, only limited attention has been given to point defects while formulating the design rules. The contribution of point defects to yield losses will be very high in a mature manufacturing process of submicron technologies. Therefore, to further improve the yield of large chips, the distribution of point defects and the sensitivity of the design to these defects must be taken into account while designing the final layout.

The probability that a defect will cause a failure, θ, depends on the size of the defect relative to the dimensions of the layout patterns. Several analytical models were proposed to calculate the critical area from layout details [14], [27], [33]. The critical area for defects of size x is defined as the area in which the center of a defect must fall in order to cause a circuit failure. The expected value of the critical area, A_C, is computed using

$$A_C = \int_0^\infty A(x)f(x)dx \tag{2}$$

where $A(x)$ is the critical area for defects of size x and $f(x)$ is the defect size probability density function.

General layout design rules, like minimum width and spacing for individual layers, have to be maintained with respect to a specific process. Local variations might be possible in some layers in such a way that the sensitivity of a layer to point defects is reduced. For example, the spacing of some lines can be increased so that the total critical area of that layer is reduced. When these changes are made in the interconnect logic, they do not introduce any functional/parametric changes in the circuit. The *RC* characteristics remain almost the same. However, when similar changes are made in the active logic, special attention should be paid to maintaining the functional and performance requirements.

The effect of reduction in the critical area on the yield of a chip depends on its size. This is shown in Fig. 1. In larger chips yield improvements will be proportional to

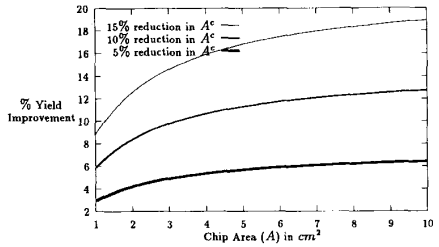


Fig. 1. Effect of critical area reduction on yield improvement.

critical area reduction. For example, the yield of a 2.5-cm^2 chip can be improved by 14% with a 15% reduction in the critical area. (Yield in Fig. 1 is calculated using the negative binomial model with $Y_0 = 0.95$, $\alpha = 1.5$ and $\lambda = 1/\text{cm}^2$.) This illustrates the significance of the critical area reduction for yield enhancement in large-area chips.

The future challenges of ever-increasing complexity of VLSI systems can be met only if the manufacturing yield of these very large chips is maintained at a profitable level. All aspects of yield enhancement need special attention in order to offset the yield losses due to increasing process complexity and diminishing feature size. To achieve higher yield for very complex VLSI/ULSI systems, they have to be designed for yield. It is not good enough just to look for area minimization. It is essential to consider how various elements are arranged within the given area. Existing CAD systems need to be supplemented by yield optimization tools.

II. Compaction Strategies for Yield Enhancement

The importance of chip area minimization in increasing the manufacturing yield cannot be overemphasized. Special CAD tools such as compactors are developed exclusively to perform area minimization [2]–[4]. While the primary goal of all the compactors is to minimize the area, they include some secondary objectives like minimizing the total wire length, minimizing the number of jogs, etc. Most of these secondary objectives are oriented toward performance improvements. Though the importance of yield enhancement is recognized [2], [4], [25], so far very little attention has been paid to it in physical layout synthesis.

Compactors generate actual layouts that occupy minimum area either from symbolic layouts or from actual layouts generated by other layout-synthesis tools. In constraint graph-based compaction algorithms [4], [23], physical connectivity and separation constraints between the elements are represented by a directed graph. The minimum achievable size of the layout is determined by the longest path (critical path) of the constraint graph. The elements on the critical path are placed at the minimum distances to minimize the area; therefore, they have no freedom to move. However, elements that do not lie on the critical path can be placed in a variety of ways. So far this additional freedom has been utilized very effectively to optimize the performance by way of wire length minimization [3].

A new compaction algorithm is presented in this section to improve the yield without increasing the layout area. This new compaction algorithm improves the yield of the final design by distributing the spacing between noncritical elements so as to minimize the total defect sensitivity for given particular manufacturing conditions, i.e., defect size distribution and defect densities for different layers of the layout. The defect sensitivity of the open-circuit type faults is minimized by increasing the width of several noncritical elements in the layout.

A. Relocation of Noncritical Elements

Compactors place various circuit elements as close as the design rules permit. Though this helps in minimizing the area, it unnecessarily packs many noncritical elements very close together, resulting in layers with a large critical area for short-circuit faults. When relocating the wire segments, the compactor may stretch them in order to maintain the original topology, resulting in longer nets and layers with a large critical area for open-circuit faults. In the SPARCS compactor [4], noncritical elements can be placed either on the top or bottom (left or right) optionally. If we also take the defect size distribution and the additional wires introduced in relocating the elements into consideration, much better results from the yield point of view can be obtained.

Since the defect size distribution is inversely proportional to the defect size raised to the power of three [33], change in the critical area will be nonuniform. For example, by increasing the spacing from 3 to 4 μm between the two wire segments of length 100 μm, the critical area of short-circuit faults can be reduced by 4 μm^2, whereas for the same amount of increase in spacing from 10 to 11 μm, the corresponding reduction will be only 0.35 μm^2. Any increase in spacing beyond the largest defect size (about 20 μm) does not decrease the critical area at all. It is interesting to note that when changes are made in the layout to minimize the sensitivity of the design to one type of defects, the sensitivity to other defect types may increase. For example, when the width of the metal/active lines is increased to minimize the sensitivity of the design to open-circuit faults, its sensitivity to short-circuit faults and pinhole faults might increase. Therefore, critical area of open- as well as short-circuit faults should be considered while finding an optimal location for noncritical elements.

The optimal location for a noncritical element is calculated by minimizing the function given below

$$\lambda(y) = \int_{W_{\min}}^{D_{\max}} A_{op}^{c}(y) D_0^{op} + \int_{W_{\min}}^{D_{\max}} A_{sh}^{c}(y) D_0^{sh} \qquad (3)$$

where $\lambda(y)$ is the number of defects which can affect the functionality of the element, $A_{sh}^{c}(y)$ ($A_{op}^{c}(y)$) is the average critical area of short-circuit (open-circuit) faults, D_0^{sh} (D_0^{op}) is the defect density of short-circuit (open-circuit) faults, $W_{\min}$ is the minimum design rule for spacing, $D_{\max}$ is the maximum defect size, and y represents the width and the location of the element.

The optimal location for a layout element depends on its length, and the spacing between the element and the elements above and below. In addition, elements connected on both sides and their widths also influence the optimal location of

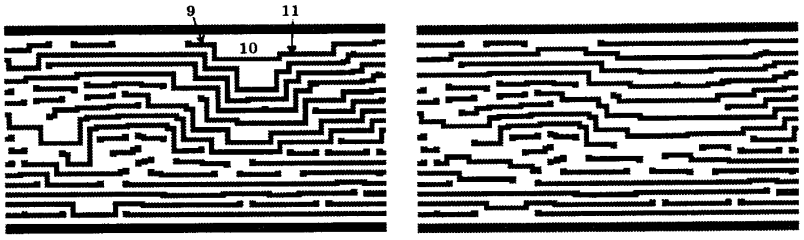


Fig. 2. Minimization of short-circuit type faults. (a) Original layout. (b) Modified layout.

an element. For example, in Fig. 2(a), element 10 is connected to the two elements 9 and 11, which are above. When element 10 is moved up by one unit (Λ), the length of each jog will be reduced by one unit. This reduction in jog length will result in fewer open-circuit faults as well as short-circuit faults between adjacent jogs. On the other hand, if an element is connected to other elements which are below, its upward movement will result in extra jog lengths and the associated faults. For a wire segment, only elements in the same layer are considered, whereas for a contact, characteristics of the elements connected in the other layer must also be considered.

The advantage of this method is that a layout can be optimized for any given manufacturing conditions. Details of the algorithm and the yield improvements achieved in two-layer channel layouts by using this compaction algorithm are presented below.

```
Step 1. Initialization
    { Identify all non-critical elements in the graph and arrange them in BFS order;
      Initialize the node weights (nodewt-ref);
      term-flag = true;  }
Step 2. Optimization
      while (term-flag)
      {  term-flag = false;
         for each non-critical element
         {  nodewt = nodewt-ref;
            new-nodewt = 0;
            while (new-nodewt < nodewt)
              if the element can be moved do
              {   move the element upward by StepSize;
                  calculate new-nodewt;
                  if (new-nodewt < nodewt)
                  {  update node positions
                     nodewt = new-nodewt;
                     new-nodewt = 0;
                     term-flag = true; }
              }
         }
      }
```

Fig. 3. Compaction algorithm.

B. Compaction Algorithm

We propose a constraint graph-based algorithm to find optimal locations for all noncritical elements of a given layout, so that its sensitivity to short-circuit type faults is reduced. The input to the algorithm is the directed-graph representation of the compacted layout. The defect size distribution $f(x)$, where x is the size of the defect, and the defect densities, d (per unit area), for different layers of the layout are the other inputs to the algorithm.

In the compacted layout, all elements are placed at the minimum possible distances from one end edge of the layout, top or bottom, left or right depending on the direction of compaction. We assume vertical compaction from top to bottom to describe the algorithm. Each node in the graph represents an element or a group of elements, and each edge represents the spacing constraint between two adjacent elements. The edge weight represents the distance between two adjacent wire segments and the node weight represents its defect sensitivity. The algorithm is shown in Fig. 3.

First, node weights are calculated for all noncritical elements in the layout. The critical area of the rectangular elements or regions is calculated using the analytical models presented in [14], [33]. Then, each noncritical element is searched in the breadth-first order for an optimal location. Elements can be moved only in the upward direction, since all the elements are initially at the minimum possible location. If the current spacing between the node and each of its incident nodes is more than the minimum design rule specification, then that node can be moved upwards. The maximum distance that an element can be moved is known as critical slack. An element is moved only if the critical slack is positive. An element is moved upwards by a prespecified amount (step size), e.g., one Λ, and the node weight is recalculated. If the current node weight is smaller than the previous one, then the element is moved by another step size. This procedure continues until the node weight in the new position starts to increase. Then the next element is processed.

Once all the noncritical elements are searched, the search is repeated. The optimal location for an element cannot be found in one iteration. This is due to the fact that the conditions under which an optimal location is found for an element may alter while processing its adjacent elements. It may take a few iterations before final optimal positions are found for all the elements. For example, the optimal location for element 10 in Fig. 2(a) is 11Λ from the top edge in the first iteration. The elements below it are not yet moved in this iteration. When these elements are moved subsequently, the location of element 10 may not be optimal due to changes in the spacing. Therefore, the final optimal location for this element is 6Λ from the top edge and is found in the fifth iteration.

C. Minimization of Open-Circuit Type Faults

The compaction algorithm presented in the previous section for yield optimization is based mainly on optimal distribution of spacing among noncritical elements without increasing their widths. Therefore, most of the yield enhancement is due to reduction in short-circuit type faults. In a similar way, further

yield improvements can be achieved by modifying the width of some noncritical elements.

Layout-synthesis tools such as routers are designed to generate layouts with minimum width elements to minimize the area. (Special layout requirements of power and ground and other critical signals may be met by manual layout or special tools). By increasing the width of some noncritical elements, open-circuit type faults can be minimized without any area penalties. These width changes do not affect the performance, since the *RC*-characteristics of the interconnecting wires remain almost the same. However, similar changes in the active logic, clock, and other long global nets require careful analysis of performance characteristics. The possible increase in the width depends on the electrical characteristics of the layer. The width of the element is increased in such a way that the total average number of faults λ as given in (3) is minimized.

Each noncritical element is processed in a breadth-first order. First, the optimal width increase at the top edge is found and then at the lower edge. The algorithm is similar to the one presented in Section I, and is omitted for the sake of brevity. While determining the optimal width for an element, the tradeoff between short- and open-circuit faults is evaluated using (3). The significant difference is that the optimal widths for the elements will be found in one iteration.

D. Examples

The results of the layout optimization are shown in Fig. 2(b) and 4. The uncompacted layout of Fig. 2 is a part of the layout generated by the router of MAGIC [31] layout editor from the netlist of $example3b$ of [36]. The layout shown in Fig. 2(a) is generated by the compactor PLOW [31] with automatic jog insertion, straightening, and with minimum horizontal length of 12Λ. The layout generated by our algorithm is shown in Fig. 2(b). To characterize the impact of the layout optimization on manufacturing yield, the yield analysis tool *Xlaser* is used. Probability of failure versus defect size plots are shown in Fig. 4, for both short-circuit faults and open-circuit faults of metal-1 layer. The probability of a short-circuit type failure in metal-1 layer of the original compacted layout is 0.0118. This failure probability is reduced by 21% to 0.00932. Since the area and the total number of defects remain the same, the average number of faults is reduced by 21%, due to this layout optimization. The reduction in probability of open-circuit faults is only 4%, since the widths of the elements are not altered. This marginal decrease is due to reduction in jog lengths.

The optimized layout shown in Fig. 2(b) is then processed for minimizing the defect sensitivity of the open-short faults and the results are shown in Figs. 5 and 6. The layout shown in Fig. 5(a) is optimized for a manufacturing environment where defect densities of both open- and short-circuit type are the same. Probability of failure versus defect size plots for both types of defects are shown in Fig. 6. As expected, the probability of failure for short-circuit faults is increased by 7.4% and the probability for open-circuit faults is decreased by 21%. The overall decrease in the average number of faults is approximately 5% from the previous stage. The two stages of layout optimization for yield enhancement resulted in an approximately 17% reduction in the number of faults in the metal-1 layer. Similar improvements are possible in the metal-2 layer as well. The same layout was optimized for conditions where the short-circuit type defect density is 10 times more than that of open-circuit type; it is shown in Fig. 5(b). It is interesting to see the significant differences in the final layout under these two different manufacturing conditions.

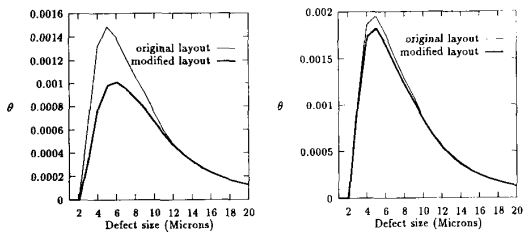


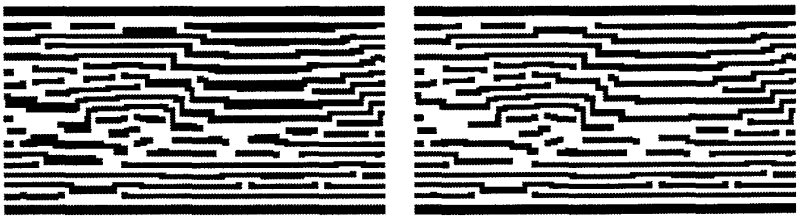
Fig. 4. Probability of failure versus defect size. (a) Short-circuit faults. (b) Open-circuit faults.

E. Results

The yield-enhancement algorithms proposed in this section have been prototyped in the IBM CircuitBench Compactor [3]. Two large circuits were analyzed for yield improvements that can be achieved by these techniques. The layouts are scaled to 0.5-μm technology. These designs consist of several thousands of active devices and two metal layers are used as interconnect layers. The layouts were first compacted in the vertical direction without enabling yield optimization feature. The defect sensitivity (in terms of probability of failure) of each interconnect layer (including polysilicon) for open- and short-circuit faults is measured using *Xlaser*. About 75% of the area is occupied by cells which consist of predominantly active regions. During compaction and yield optimization, layout patterns in these cells are not moved; however, cells as a unit can be moved. As a result, defect sensitivity of the layers belonging to the active regions remains almost unchanged. Therefore, the results of these layers are omitted.

The defect-size distribution model by Stapper and Ferris-Prabhu [14], [33] is used for yield optimization as well as defect-sensitivity measurements. We have assumed the following values for different parameters of the model: x_o = 0.5, p = 3.0, and q = 1.0. The average probability of failure (*pof*) for each interconnect layer is shown in the third and sixth columns of Table I. The layout area of these two circuits, after compaction, is 0.38 and 0.425 mm^2.

The layouts were then compacted by enabling the yield-enhancement option. The area of the layout remains unchanged during yield optimization phase. The *pof* for each layer of these layouts is shown in the fourth and seventh columns. The percentage reduction in the *pof* is shown in the fifth and eighth columns. In both these examples, the *pof* of metal-1 layer for short-circuit faults is reduced by 8.2%. The *pof* for open-circuit faults is reduced very marginally. In these examples the metal-1 layer consists of predominantly

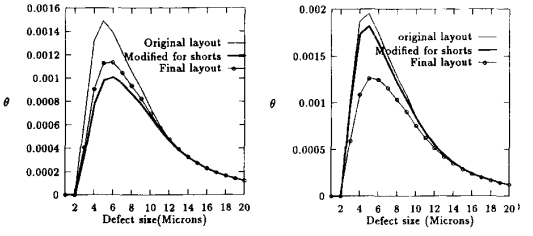


(a) (b)

Fig. 5. Minimization of open-circuit type faults. (a) Modified for same defect density. (b) Modified for higher short-circuit defect density.

TABLE I
EFFECT OF LAYOUT MODIFICATIONS ON YIELD

Layer	Defect Type	YKT-I (area 0.38 sq. mm.) Simple Compaction POF	YKT-I Compaction for Yield Optim. POF	YKT-I % Red.	YKT-II (area 0.425 sq. mm.) Simple Compaction POF	YKT-II Compaction for Yield Optim. POF	YKT-II % Red.
Metal-1	Shorts	0.003382	0.003103	8.25	0.004319	0.003964	8.22
	Opens	0.010093	0.010133	-0.40	0.012352	0.012389	-0.30
Metal-2	Shorts	0.000625	0.000595	4.80	0.001158	0.001159	-0.09
	Opens	0.002957	0.002818	4.70	0.004269	0.004266	0.07
Polysilicon	Shorts	0.002786	0.002786	0.00	0.002878	0.002871	0.24
	Opens	0.006000	0.006011	-0.18	0.005733	0.005721	0.21



(a) (b)

Fig. 6. Probability of failure versus defect size. (a) Short-circuit faults. (b) Open-circuit faults.

horizontal patterns and the metal-2 layer of vertical patterns. Since the layouts were compacted in the vertical direction without auto-jogging, changes in the wire length of metal-1 layer are minimal. Therefore, the percentage reduction in the *pof* is also negligible. We suspect that the small increase in the *pof* of open-circuit faults is due to the proximity effect [33].

In YKT-I, the *pof* for short- and open-circuit faults is reduced by 4.8 and 4.7%, respectively, in the metal-2 layer. However, in YKT-II, the reduction is negligible. In this example most of the metal-2 wires are bounded by contacts outside the boundary, hence, they don't have freedom to move. In both these examples, the *pof* reduction in the polysilicon layer is minimal because it is not used as an interconnect layer except for few interconnects. The reduction in the defect sensitivity of individual layers of the circuit can be directly translated into yield improvement with additional information on defect densities for short- and open-circuit faults, clustering factor data, etc. Our sample calculations show that an 8-10% improvement, in defect sensitivity on 2 or 3 interconnect layers, on a chip of 1 cm^2 can result in a 5-10% improvement in chip yield.

III. ROUTING STRATEGIES FOR YIELD ENHANCEMENT

Since compaction is the last stage of the layout synthesis, its effectiveness is highly dependent on the quality of the layout synthesis of the previous stages. For example, the quality of the routers has a major impact on compaction. Therefore, further yield improvements can be achieved through new strategies for routing, layer assignment and alike.

A. Via and Wire Length Minimization

Most of the existing routers try to minimize the number of vias in the layout. Since the minimum width and spacing requirements for vias are larger than the wire segments, generally more compact designs are possible with fewer vias [5], [13], [17]. Sometimes, just to avoid a via, routers may introduce very long wire segments, which certainly results in layers with higher critical areas. On the contrary, in certain situations it may be worthwhile to introduce some vias (or leave some vias intact) to avoid unnecessary additional wiring. This situation is illustrated in Fig. 7. By shifting a part of a horizontal segment of net-1 from track-5 to another free track (track-2), the total net length is reduced by 48% at the cost of one additional via. With a similar reassignment in net-2 from track-2 to track-7, the net length is reduced by half. The A_c of the metal-1 layer is reduced by 25% with these two modifications.

Analogous to minimum wire length requirements for introducing additional vias in the preferred layer maximization problem [9], some criteria to remove/add a via in terms of wire length/critical area must be introduced. For example, for the defect densities reported in [10], the fault probability of one metal-1/poly contact is equivalent to that of a poly wire segment of length 15 μm and width 1.5 μm. By adding a via, which can eliminate more than 15 μm of polysilicon, critical area/fault probability can be reduced. This additional criterion in routing, for the tradeoff between wire length and via, will result in layouts with better yield characteristics.

In the routing techniques proposed in [30] only the adjacency information of horizontal tracks is considered as a criterion for defect sensitivity. Neither defect size distributions, nor analytical models, were used to characterize yield. Since the vertical layer is not considered, we have seen an increase in the overall defect sensitivity in some of the examples generated by these routers. If the routing area is compacted, the spacing between horizontal tracks is changed by the compactor.

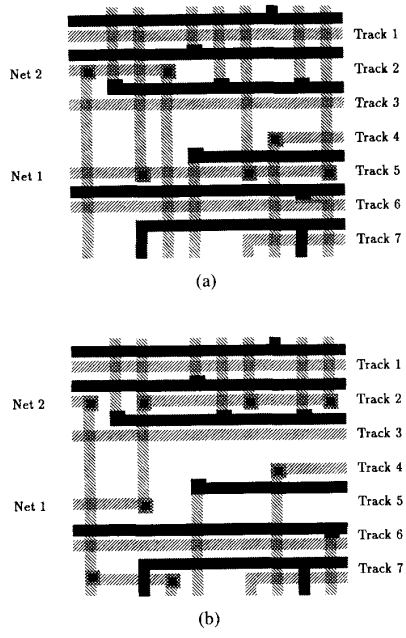


Fig. 7. Via and wire length minimization. (a) Original layout. (b) After reassigment of two tracks.

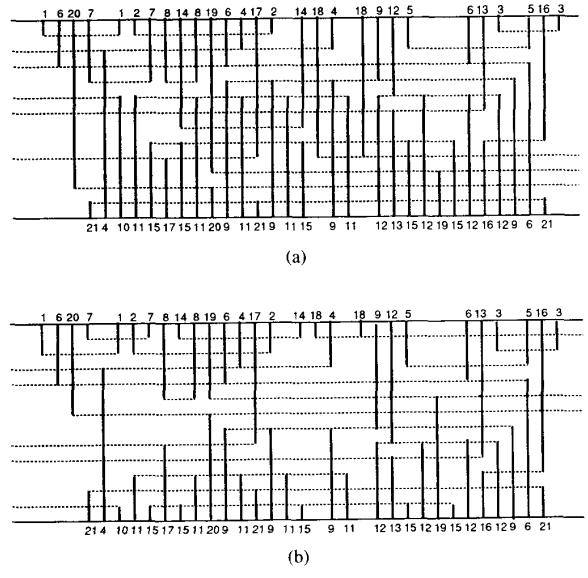


Fig. 8. (a) Example 1 original routing (WL 310). (b) Example 1 optimized for wire length (WL 222).

TABLE II
WIRE LENGTH AND VIA REDUCTION DUE TO TOPOLOGICAL OPTIMIZATION IN TWO-LAYER ROUTING

Examples in [36]	Chan. Den.	# of col.	# of nets	Org. Routing WL	Org. Routing Vias	Opt. Routing WL	Opt. Routing Vias	% Reduction Wl	% Reduction Vias
ex1	12	35	21	310	57	222	36	28.4	36.8
ex9a	15	62	45	583	91	511	59	12.3	35.2
ex9b	17	61	47	818	107	736	78	10.0	27.1
ex9c	18	79	54	976	125	832	92	14.5	26.4
ex4b	17	119	54	1150	179	1113	116	3.2	35.2
ex5	20	119	60	1309	150	981	102	25.1	32.0
	Average Reducton							14.6	31.9

Therefore, at the routing stage, if the vertical layer is targeted, better layouts can be generated. Since the compactors can not alter the topological order of the nets, defect sensitivity of the vertical layers can be reduced efficiently during routing. Then the yield-enhancement techniques presented in the previous section can be applied during compaction for further yield enhancement.

We have formulated wire length of the vertical layer in a two-layer routing as an Integer Linear Programming problem. To illustrate the wire length minimization, a simple example (*example 1* of [36]) is shown in Fig. 8. The original channel has 12 tracks and the total wire length of the vertical layer is 310 units; by reassigning the nets to different tracks, it has been reduced to 222 units resulting in a reduction of 28%. This reduction in wire length results in similar reductions in the defect sensitivity of open- and short-circuit type faults. Wire length minimization achieved in the examples given in [36] is shown in Table II. The average wire length reduction in these benchmark examples is 14.6%. This can result in proportional improvements in defect sensitivity of the vertical layer. Due to this wire length reduction, the number of vias is also reduced significantly (32%). In many examples, the via-reduction results are better than the results reported by several via-minimization algorithms [34]. Via reduction will further improve the defect sensitivity of the layouts. Yield-improvement results of these examples due to wire length and via reduction have been reported in [8].

IV. TOPOLOGICAL LAYOUT DESIGN TECHNIQUES

During the last decade, many structured design techniques have been developed to minimize the design cycle time of VLSI systems. PLA's, gate arrays, and standard-cell designs are some of the popular design styles. Two-level PLA-based logic synthesis is well developed, and commercial automatic synthesis tools/silicon compilers are now available. One of the major drawbacks of these design styles, as of now, is the large area overhead (due to sparsity of the personality matrices of most designs) and the attendant yield and performance degradation. A variety of fault-tolerant techniques have been proposed for PLA-based designs in order to enhance yield [24] which involve a cost of additional area and design effort.

Several optimization techniques have been proposed to minimize the area of PLA's at various stages of the design, starting from functional design to physical design. PLA folding techniques and the corresponding software tools have been developed to optimize the topological representation of PLA's [16], [18]. The primary objective of all these techniques is to reduce the area of the PLA. Significant yield enhancement can also be achieved by minimizing the defect sensitivity of the design that is already optimized for area.

We have proposed a topological optimization technique for yield enhancement of PLA-based designs in [7]. In our approach, the topological representation of the PLA is altered so that the critical area of the generated layout is minimized. This reduction in critical area is achieved primarily by minimizing the wire lengths in one or more layers of the layout.

TABLE III
EFFECT OF TOPOLOGICAL OPTIMIZATION ON THE PRIMARY LAYER OF Misex1 PLA

Parameter	Original layout	After topological optimization		After physical optimization		Total reduction
			reduc.		reduc.	
Poly wire length	5.83mm	4.73mm	19.0%	—	—	19.0%
A_c of short-ckt faults	144μm^2	129μm^2	10.4%	145μm^2	-11.0%	-0.7%
A_c of open-ckt faults	665μm^2	542μm^2	18.5%	467μm^2	13.8%	29.8%
A_c total	809μm^2	671μm^2	17.1%	612μm^2	8.8%	24.4%
Area	0.024mm^2	0.021mm^2	12.5%	—	—	12.5%
Delay	1.93ns	1.93ns	—	1.93ns	—	—

By modifying the topological representation of the PLA, wire length in the physical layout can be minimized without increasing the area, resulting in layers with reduced critical area. The amount of wire reduction in various layers depends on the folding level, technology and constraints imposed on the input/output positions, etc. For example, in simple column folding with constraints on the inputs, higher reduction in wire length is possible. We show the yield improvement and wire length reduction results achieved through topological optimization with the help of some MCNC benchmark PLA examples in the following section.

A. Examples and Results

The PLA example, *misex1*, consists of 8 input lines (16 lines with complements), 7 output lines, and 12 product terms. Layout of this PLA is generated using OCT and MAGIC systems. The original symbolic representation of the circuit and the corresponding physical layout are shown in Figs. 9(a) and Fig. 10, respectively. The inputs run from both the top and bottom. The optimized symbolic layout is shown in Fig. 9(b). The wire length and the critical area of open- and short-circuit type faults of polysilicon layer are shown in Table III. The performance of the modified layout has been verified using *crystal*. There is no change in the maximum delay of 1.93 ns in the *misex1* PLA circuit. The length of the input polysilicon lines is reduced by permuting the product terms in row positions. This rearrangement of product terms does not have any negative impact on the performance.

Due to this optimization, 19% wire length reduction is achieved in the polysilicon layer of the AND plane. Consequently, the critical area of this layer is reduced by 17%. It is interesting to observe the incidental reduction in the wire length of the other layers, e.g., metal-1 and diffusion layers. The reduction in wire length of the individual layers may result in better compaction, thereby reducing the overall area. In this example, the area is reduced by 12.5%. Therefore, the effective yield of the die also increases by approximately that percentage. However, this percentage of area reduction may not always be possible when performing yield optimization, if better folding tools are used [18]. In such a case, the percentage reduction in wire length and in the attendant critical area may be less. The reduction in the wire length of several layers facilitates the implementation of yield-enhancement techniques during physical layout synthesis. After implementing these changes in the layout, the critical area is further reduced by 8.8%. Therefore, the overall reduction in the critical area is about 24% in the polysilicon layer and about 11% in metal-1

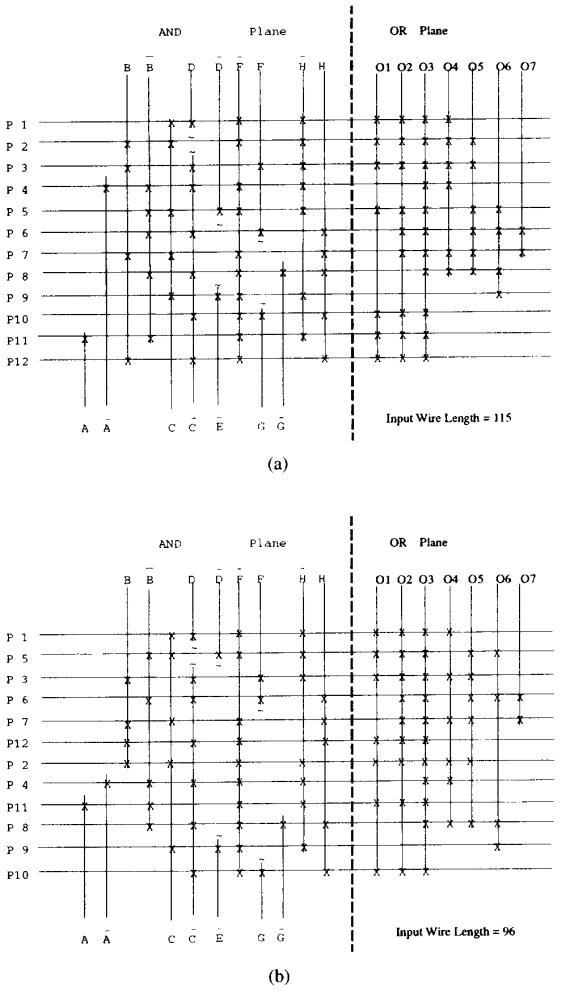


Fig. 9. (a) Mixex1 PLA original topological representation. (b) Misex1 PLA optimized topological representation.

and diffusion layers. In larger chips, these reductions in the critical areas and the chip area result in a yield improvement of up to 15%.

Results of wire length minimization on benchmark examples are shown in Table IV. The average wire length reduction is 15.4% for the nonfolded PLA's and 10.9% for the column-folded case. This wire length reduction can significantly improve the defect sensitivity of the corresponding layers of the circuits. Yield improvement results due to the topological optimization on larger PLA's have been presented in [7].

We illustrated the significance of topological optimization for yield enhancement in PLA-based designs. This new approach for yield enhancement has many attractive features compared to conventional methods. Yield enhancement in PLA's through redundancy and reconfiguration is proposed in [24]. In this approach, extra resources such as spares, testing, reconfiguration etc., require additional area of up to 25%. Due to this additional area requirement, effective yield will go down beyond a certain optimal level of redundancy. In this approach, performance degradation is another concern. In

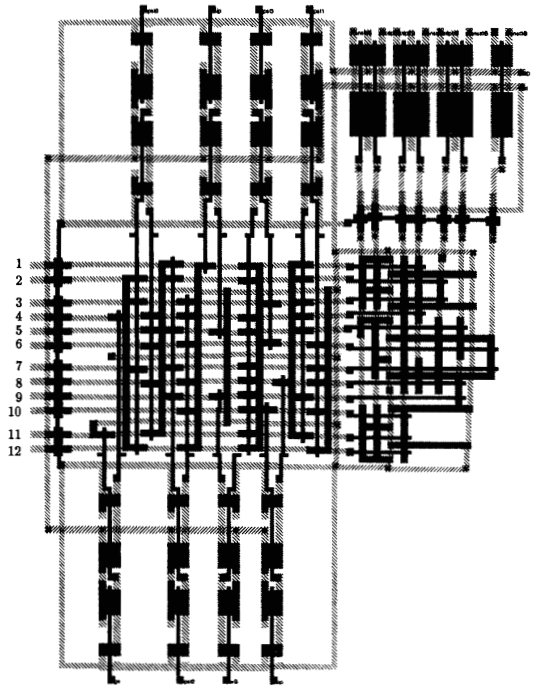


Fig. 10. Original layout of Misex1 PLA.

TABLE IV
INPUT WIRE LENGTH REDUCTION DUE TO TOPOLOGICAL OPTIMIZATION IN THE PLA'S

PLA Parameters				Wire Length					
Circuit				No Folding			Column Folding		
Name	I/P	O/P	Pterms	Org	Opt	%Red.	Org	Opt	%Red.
bw	5	28	22	202	182	9.9	185	172	7.0
clip	9	5	118	1973	1590	19.4	1864	1551	16.8
con1	7	2	9	64	64	0.0	45	38	15.5
duke2	22	29	86	2608	2074	20.5	2146	1991	7.2
inc	7	9	30	334	292	12.6	341	287	15.8
misex1	8	7	12	120	114	5.0	115	96	16.5
misex2	25	18	28	527	486	7.8	585	542	7.4
sao2	10	4	58	983	864	12.1	949	831	12.4
table3	14	14	175	4842	3966	18.1	4464	3870	13.3
table5	17	15	158	4776	4169	12.7	4442	4191	5.7
vg2	25	8	110	3213	2814	12.4	3302	2853	13.6
Average Wire Length Reduction						15.4			10.9

the new approach proposed above, no additional resources are required to achieve the yield improvement.

V. CONCLUSIONS

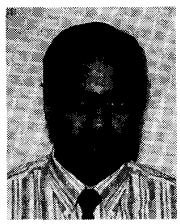
The most significant aspect of the proposed yield-enhancement techniques is that no additional area is required. The yield enhancement can be realized almost at no additional cost, except for a marginal increase in the computational time of the CAD tools. Another important aspect is that a layout can be optimized for given manufacturing conditions. Further yield improvements can be achieved by layer reassignment during the routing and compaction stages of layout design. Incorporation of additional defect types for the yield criterion will improve the accuracy of yield estimation.

The proposed layout-design techniques for yield enhancement should supplement rather than replace the more traditional defect-tolerance techniques. The complexity of future products will be too high to achieve the yield targets with either of these two approaches alone. The effectiveness of these two approaches is highly dependent on the design structure, complexity, and the process yield. In very regular architectures, most notably memory units, defect-tolerance techniques are expected to have a higher contribution toward yield improvement. As the design becomes less regular, the contribution of the layout techniques is expected to increase. The use of redundancy is especially beneficial for processes and circuits which have low yields. By using both these techniques, the yield targets can be achieved with lesser area overheads.

REFERENCES

[1] G. A. Allan *et al.*, "An yield improvement technique for IC layout using local design rules," *IEEE Trans. Computer-Aided Design*, vol. 11, no. 11, pp. 1355-1362, Nov. 1992.
[2] D. G. Boyer, "Symbolic layout compaction review," in *Proc. 25th ACM/IEEE Design Automation Conf.*, 1988, pp. 383-389.
[3] EDA CircuitBench User's Guide, IBM Corporation, NY, 1994.
[4] J. L. Burns and R. Newton, "SPARCS: A new constraint-based IC symbolic layout spacer," in *Proc. IEEE Custom Integr. Circuits Conf.*, 1986, pp. 534-539.
[5] C. K. Cheng *et al.*, "Geometric compaction on channel routing," *IEEE Trans. Computer-Aided Design*, vol. 11, no. 1, pp. 115-127, Jan. 1992.
[6] V. K. R. Chiluvuri and I. Koren, "New routing and compaction strategies for yield enhancement," in *IEEE Int. Workshop on Defect and Fault Tolerance in VLSI Systems*, 1992, pp. 325-334.
[7] ——, "Topological optimization of PLA's for yield enhancement," in *IEEE Int. Workshop on Defect and Fault Tolerance in VLSI Systems*, 1993, pp. 175-182.
[8] ——, "A wire length minimization algorithm for channel routing," ECE Dept., University of Massachusetts, Amherst, Technical Report TR-94-CSE-10, 1994.
[9] M. J. Ciesielski, "Layer assignment for VLSI interconnect delay minimization," *IEEE Trans. Computer-Aided Design*, vol. 8, no. 6, pp. 702-707, June 1989.
[10] R. S. Collica *et al.*, "A yield enhancement methodology for custom VLSI manufacturing," *Dig. Tech. J.*, vol. 4, no. 2, pp. 83-99, 1992.
[11] J. A. Cunningham, "The use and evaluation of yield models in integrated circuit manufacturing," *IEEE Trans. Semicond. Manufact.*, vol. 3, no. 2, pp. 60-71, May 1990.
[12] R. B. Fair, "Challenges to manufacturing submicron, ultra-large scale integrated circuits," *Proc. IEEE*, vol. 78, no. 11, pp. 1687-1705, Nov. 1990.
[13] S. C. Fang *et al.*, " Via minimization with associated constraints in three-layer routing problem," in *Proc. Int. Symp. Circuits Syst.*, 1990, pp. 1632-1635.
[14] A. V. Ferris-Prabhu, "Role of defect size distribution in yield modeling," *IEEE Trans. Electron Devices*, vol. ED-32, no. 9, pp. 1727-1736, Sept. 1985.
[15] J. P. Gyvez and Chennian Di, "IC defect sensitivity for footprint-type spot defects," *IEEE Trans. Computer-Aided Design*, vol. 11, no. 5, pp. 638-658, May 1992.
[16] G. D. Hachtel, A. R. Newton, and A. Sangiovanni-Vincentelli, "An algorithm for optimal PLA folding," *IEEE Trans. Computer-Aided Design*, vol. CAD-1, no. 2, pp. 63-76, Apr. 1982.
[17] S. Haruyama, D. F. Wong, and D. S. Fussell, "Topological channel routing," *IEEE Trans. Computer-Aided Design*, vol. 11, no. 10, pp. 1177-1197, Oct. 1992.
[18] S. Y. Hwang, R. W. Dutton, and T. Blank, "A best-first search algorithm for optimal PLA folding," *IEEE Trans. Computer-Aided Design*, vol. CAD-5, no. 3, pp. 433-442, July 1986.
[19] I. Koren and C. H. Stapper, "Yield models for defect tolerant VLSI circuits: A review," in *Defect and Fault Tolerance in VLSI Systems*, vol. 1, I. Koren, Ed., New York: Plenum, 1989, pp. 1-21.
[20] I. Koren, Z. Koren, and C. H. Stapper, "A unified negative binomial distribution for yield analysis of defect tolerant circuits," *IEEE Trans. Comput.*, vol. 42, no. 6, pp. 724-734, June 1993.
[21] Z. Koren and I. Koren, "Does the floorplan of a chip affect its yield?" in *IEEE Int. Workshop on Defect and Fault Tolerance in VLSI Systems*, 1993, pp. 159-166.

[22] S. Y. Kuo, "YOR: A yield-optimizing routing algorithm by minimizing critical areas and vias," *IEEE Trans. Computer-Aided Design*, vol. 12, no. 9, pp. 1303-1311, Sept. 1993.
[23] T. Lengauer, *Combinational Algorithms for Integrated Circuit Layout*. Chichester, U.K.: John Wiley & Sons, 1990.
[24] R. Leveugle, M. Soueidan, and N. Wehn, "Defect tolerance in a 16-bit microprocessor," in *Defect and Fault Tolerance in VLSI Systems*, vol. 1, I. Koren Ed., New York: Plenum, 1990.
[25] M. Lorenzetti, "The effect of channel router algorithms on chip yield," presented at The MCNC Int. Workshop on Layout Synthesis, May 1990.
[26] A. R. Dalal *et al.*, "A layout-driven yield predictor and fault generator for VLSI," *IEEE Trans. Semicond. Manufact.*, vol. 6, no. 1, pp. 77-81, Feb. 1993.
[27] W. Maly, "Modeling of lithography related yield losses for CAD of VLSI circuits," *IEEE Trans. Computer-Aided Design*, vol. CAD-4, no. 3, pp. 166-177, July 1985.
[28] ——, "Computer-aided design for VLSI circuit manufacturability," *Proc. IEEE*, vol. 78, no. 2, pp. 356-392, Feb. 1990.
[29] *M^2C CAD Tools Distribution*, Release 6.0, 1990.
[30] A. Pitaksanonkul *et al.*, "DTR: A defect-tolerant routing algorithm," in *26st IEEE Design Automat. Conf.*, 1989, pp. 795-798.
[31] J. K. Ousterhout *et al.*, "MAGIC: A VLSI layout system," in *21st IEEE Design Automat. Conf.*. 1984, pp. 152-159.
[32] R. D. Rung, "Determining IC layout rules for cost minimization," *IEEE J. Solid-State Circuits*, vol. SC–16, no. 1, pp. 35-42, Feb. 1981.
[33] C. H. Stapper, "Modeling of defects in integrated circuit photolithographic patterns," *IBM J. Res. Develop.*, vol. 28, no. 4, pp. 461-474, July 1984.
[34] K. The, D. F. Wong, and J. Cong, "Via minimization by layout modification," in *Proc. 26th ACM/IEEE Design Automat. Conf.*, 1989, pp. 799-802.
[35] H. Walker and S. W. Director, "VLASIC: A catastrophic fault yield simulator for integrated circuits," *IEEE Trans. Computer-Aided Design*, vol. CAD–5, no. 4, pp. 541-556, Oct. 1986.
[36] T. Yoshimura and E. S. Kuh, "Efficient algorithms for channel routing," *IEEE Trans. Computer-Aided Design*, vol. CAD–1, no. 1, pp. 25-35, Jan. 1982.



Venkat K. R. Chiluvuri (S'93) received the B.Tech degree in electronics and communication engineering from J. Nehru Technological University, College of Engineering, Kakinada, India, in 1982, and the M.Tech degree in electrical engineering from the Indian Institute of Technology, Kharagpur, India, in 1984. He is currently working towards the Ph.D degree at the University of Massachusetts, Amherst.

From 1984 to 1990 he was with Wipro Information Technology Ltd., Bangalore, India. During 1993 he was with IBM's T. J. Watson Research Center, NY. His research interests include CAD for yield and manufacturability, and VLSI design.

Israel Koren (S'72–M'76–SM'87–F'91), for photograph and biography, see this issue, p. 138.