

# FAR-DS: Full-plane AWE Routing with Driver Sizing \*

Jiang Hu and Sachin S. Sapatnekar

Department of Electrical and Computer Engineering  
University of Minnesota, Minneapolis, MN 55455, USA  
jhu, sachin@mountains.ee.umn.edu

## Abstract

We propose a Full-plane AWE Routing with Driver Sizing (FAR-DS) algorithm for performance driven routing in deep sub-micron technology. We employ a fourth order AWE delay model in the full plane, including both Hanan and non-Hanan points. Optimizing the driver size simultaneously extends our work into a two-dimensional space, enabling us to achieve the desired balance between wire and driver cost reduction, while satisfying the timing constraints. Compared to SERT, experimental results showed that our algorithm can provide an average reduction of 23% in the wire cost and 50% in the driver cost under stringent timing constraints.

## 1 Introduction

As the VLSI feature sizes keep shrinking, the interconnect contributes more and more as an obstruction to advancement. As the result, many efforts [1–7], have been carried out in recent years to improve the interconnect performance. A major advance was made a few years ago when the wire length minimization was augmented by delay reduction as the objective for interconnect optimization. Initial approaches used the Elmore delay [8] model extensively because of its simplicity. One major method of this kind, SERT [1], grows the routing tree in a greedy fashion to minimize the source-sink delay and shows much better results than any preceding work. However, it tends to result in star-like topologies and large area overheads. A second Elmore delay based routing algorithm, the P-tree [4] algorithm, first searches for a good permutation of the sinks and then limits the solution space to the topologies induced by this permutation. The objective of the P-tree method is same as that of [6], which is to minimize the cost subject to timing constraints on the sinks. Other researchers address the limitations of Elmore delay and apply second [3] and third order [5] models. Most recently, [2] suggested table-lookup methods to remedy the deficiencies of the Elmore delay model.

A classification of the three crucial aspects of performance driven routing research was presented in [2]: solution space, objective formulation and delay model. So far, minimizing the cost subject to timing constraints is widely recognized [3–7] as the best objective formulation. The solution space has long been restricted to the Hanan grid and [7] presented a result showing that the use of non-Hanan points can offer significant benefits for the problem of minimizing the cost subject to timing constraints. In this paper, we

propel this effort forward by using a fourth order AWE [9] delay model on the full plane that includes both Hanan and non-Hanan points. Therefore in each of the three classes in the taxonomy of [2]: the complete exploration of the solution space, the objective function formulation and the accuracy of the delay model, our approach constitutes the best feasible formulation for the performance driven routing problem. As shown in [7], the objective inspires the need for routing on non-Hanan grids. The use of higher order delay models on the full non-Hanan plane ensures that the optimal solution is found accurately.

In addition to the above, our work takes the driver sizing issue into consideration simultaneously. In this paper we will show the curvature properties of delay function vs. connection location and driver resistance in a two-dimensional space under the Elmore model. Though the Elmore model may be poor for specific points, it still provides a valid prediction of qualitative properties [1]. According to the solution region properties, we have developed the FAR-DS (Full-plane AWE Routing with Driver sizing) algorithm that can reach the desired balance between the minimization of the wire length cost and the driver cost, while satisfying the timing constraints.

Our algorithm was tested and compared the results with SERT and MVERT on both  $0.18\mu\text{m}$  IC and MCM technology. In the case when SERT or MVERT cannot meet the timing constraints, our algorithm can usually generate a routing tree that satisfies the timing constraints; when SERT and MVERT succeed, our algorithm can often further reduce the wire cost or driver cost, or both, significantly.

## 2 Preliminaries

### 2.1 The motivation for fourth order AWE

As interconnect wires become increasingly thinner and longer, the interconnect resistance may overshadow the driver resistance. Consequently, the net capacitance of sinks and downstream capacitance are shielded to the driver resistance by the interconnect resistance. This effect is called resistive shielding [10]. The Elmore delay does not correctly take the resistive shielding effect into account and tends to overestimate the delay. This error can be remarkably large, especially for the stub situation (i.e., when a sink that is close to the source co-exists with a much longer wire), where the Elmore delay can be several times larger than the actual delay.

Table 1: Elmore and AWE delays vs. SPICE

Dist.	SPICE	Elmore	Error	4th AWE	Error
370	13.6	52.5	286%	12.8	-6%
600	9.5	39.8	319%	8.9	-6%
900	10.7	40.5	279%	10.5	-2%
1100	26.2	77.4	195%	25.5	-3%
12000	283.2	257.5	-9%	282.4	-0.3%

Table 1 shows a pathological example of a net with five sinks to illustrate the inaccuracy of the Elmore delay. The load capacitance is the same for each sink. The delays on all sinks are computed

\*This work is supported in part by the NSF under contract CCR-9800992 and the SRC under contract 98-DJ-609.

using the Elmore formula, fourth order AWE and a SPICE transmission line model, and the percentage errors relative to SPICE are calculated. The Manhattan distance from each sink to the source are also listed for reference. We can see that the error of Elmore delay can be over 300% and the delay from the AWE model is clearly superior. In fact, as the wire size shrinks, this trend will be more and more severe.

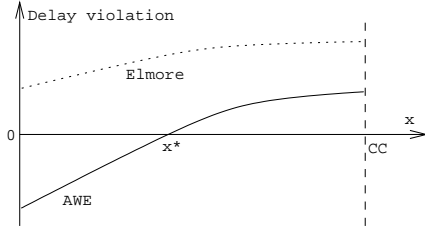


Figure 1: An example that Elmore delay and higher order AWE delay may result in different connection choice.

To see how this will affect the non-Hanan routing, consider the graph in Figure 1. The graph plots the delay violation function against the location of the connection point,  $x$ , as pictured in Figure 2. The dotted curve indicates the Elmore delay while the solid curve represents the fourth order AWE result. The solution corresponds to the point closest to  $CC$  where the delay violation function is negative or zero. For the Elmore delay, which overestimates the delay near the source, no solution is found, whereas an actual solution exists and corresponds to  $x^*$ .

On the other hand, we have observed that the Elmore model tends to under-estimate delay at sinks far from the source<sup>1</sup>. This may lead to the opposite error, as can be seen in the last row of Table 1. This under-estimation may result in over-reduction of cost while the timing constraints have not been satisfied yet. On the whole, a higher order model is greatly superior to the Elmore model in handling non-Hanan points.

The reason that we choose fourth order instead of a second or third order model is that second order gives less accuracy and for many examples that we tried, we found that the third order model induces positive poles more often.

In the computation of fourth order AWE delay, we first use RICE [13] to obtain the moments. We solve the denominator of the Padé approximation result, which is a fourth order polynomial, using a closed form formula to obtain the poles. After an inverse Laplace transformation, the time domain exponential functions are expanded about the Elmore delay to fourth order Taylor series polynomials to reach the delay values. Since a closed form solution to a fourth order polynomial exists and may be used to ease the computation, the additional computation cost of fourth order AWE as compared to a second order model is not large. This process is iterated until convergence, and we found that we always converged within 3 iterations. This method is related to the Newton-Raphson root-finding method: the Newton-Raphson method uses a first order Taylor series in each iteration, and our method uses a fourth order expansion instead.

## 2.2 Problem formulation

The objective of our approach is to minimize the cost of the routing tree, subject to a timing constraint at each sink. In contrast with [7], we extend the cost here to include both wire cost and driver cost,

<sup>1</sup>The Elmore delay is theoretically proven to be an upper bound on the delay of an RC network in [12]. However, in practice, greater accuracies are obtainable by multiplying the Elmore delay formula of [11] by a factor of  $\ln 2$ , and we refer this quantity as the “Elmore delay” in our discussion, and this may be either optimistic or pessimistic.

i.e., we perform topology optimization and driver sizing simultaneously. The rationale behind this is to permit the driver to share the task of delay optimization with the interconnect by sizing it, thereby obtaining a better result than optimizing the driver size and interconnect topology separately. Our method can also be extended to buffer sizing easily.

A list of notational terms used in this work is as follows:

- $Q_i$ : required arrival time for sink  $i$ .
- $D_i$ : calculated delay for sink  $i$  in the routing tree.
- $V_i$ : delay violation of sink  $i$ , given by  $V_i = D_i - Q_i$ .
- $W$ : total wire length for a routing tree.
- $R_d$ : driver resistance.
- $R_{dMAX}$ : largest allowable driver resistance.
- $R_{dmin}$ : smallest allowable driver resistance.
- $r$ : resistance per unit length for interconnect.
- $c$ : capacitance per unit length for interconnect.

We state the problem formulation as follows:

Given a source  $s_0$ , a set of sinks  $S = \{s_1, s_2, \dots, s_n\}$ , timing specifications  $Q = \{Q_1, Q_2, \dots, Q_n\}$  for all sinks, and resistance bounds  $R_{dMAX}$  and  $R_{dmin}$ , construct a Steiner routing tree and find an  $R_d$  such that:

$$\begin{aligned} & \text{minimize} && rW - e^\gamma R_d \\ & \text{subject to:} && \max_{i \in \text{set of sinks}} (V_i) \leq 0 \\ & && \text{and} \quad R_{dmin} \leq R_d \leq R_{dMAX}. \end{aligned} \quad (1)$$

The objective function can be interpreted as a minimization of the total wire length and maximization driver resistance, i.e., minimization of the driver size. The parameter  $\gamma$  is a user-specified weighting factor. The purpose of the exponential form is to appropriately weight the importance of driver and wire cost<sup>2</sup>. A positive value of  $\gamma$  tends to reduce the cost of driver more than that of wire, and a negative value has the opposite effect. This provides a nice flexibility to the user, because in some situations the driver cost is more crucial while in others the wire cost may be more important. A compromised weight may also give a similar priority to reducing both wire and driver at the same time. The reason for choosing the parameter  $r$ , the resistance per unit length of the wire, as the weight for total wire length will be explained in the next section.

## 2.3 Properties of solution regions

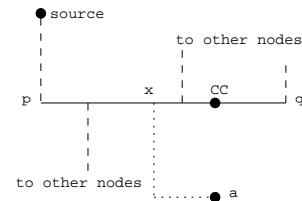


Figure 2: A general situation of connection sink  $a$  to a maximal segment  $pq$

Consider a general form of a routing tree such as that shown in Figure 2. We define a segment to be a contiguous set of straight edges in a routing tree that are either all horizontal or all vertical.

<sup>2</sup>It was found that a constant coefficient of  $kR_d$  required large variations in  $k$  before it affected the solution point, and hence an exponential was used for convenience.

A maximal segment is a segment that is not properly contained in any other segment. In Figure 2,  $pq$  is a maximal segment and  $p$  is the root of this maximal segment. Consider the connection for sink  $a$  and its downstream subtree, if any, to  $pq$ . Let  $x$  denote the distance from the connection point to  $p$  and let  $CC$  be the closest connection from  $a$  to  $pq$ . We also overload the notation for all other nodes by using their name to denote their distance to  $p$ .

Parallel to the derivation in [1], we can obtain a general form of delay violation to any sink  $s_i$  as:

$$V_i = f(x, R_d) = -krcx^2 - cxR_d + k_2x + k_1R_d + k_0 \quad (2)$$

where

$$k = 0 \text{ or } 1, \quad 0 \leq x \leq CC < \frac{k_1}{c}, \quad R_{dmin} \leq R_d \leq R_{dMAX}, \quad (3)$$

with  $k_0, k_1$  and  $k_2$  being constants.

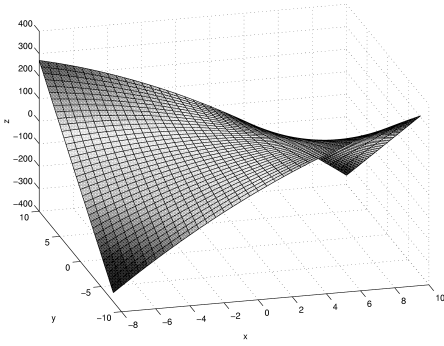


Figure 3: Example of surface of function  $V_i = f(x, R_d)$ .

If we use  $y$  to denote the coordinate of  $R_d$  and let  $z$  to denote the coordinate of  $V_i$ , the function of (3) is a surface in the three dimensional space of  $(x, y, z)$ , which is illustrated by Figure 3. To investigate the curvature property of this surface, we can examine the Hessian matrix of  $V_i$  in this space:

$$\mathbf{H} = \nabla^2 V_i(x, R_d) = \begin{pmatrix} -2krc & -c \\ -c & 0 \end{pmatrix}. \quad (4)$$

A function is concave if its Hessian matrix is negative semi-definite. One way to know if  $\mathbf{H}$  is negative semi-definite or not is to check the eigenvalues of  $\mathbf{H}$ . When all the eigenvalues are non-positive, the Hessian matrix is negative semi-definite and the function it corresponds to is concave. Similarly, all the eigenvalues of a Hessian matrix being non-negative implies that the function is convex. The eigenvalues of the matrix in (4) are:

$$\lambda_1 = -rc + c\sqrt{r^2 + 1} > 0, \quad \lambda_2 = -rc - c\sqrt{r^2 + 1} < 0, \quad (5)$$

and the corresponding eigenvectors are:

$$\mathbf{e}_1 = \begin{pmatrix} kr - \sqrt{k^2 r^2 + 1} \\ 1 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} kr + \sqrt{k^2 r^2 + 1} \\ 1 \end{pmatrix}. \quad (6)$$

Note that since one eigenvalue is positive and the other is negative, the function is neither concave nor convex over the entire space. This is in accordance with the shape of the surface in Figure 3. However, there are still properties that we can exploit.

For an arbitrary vector  $\mathbf{v} = (v_x, v_y)^T$  in the  $(x, y)$  space, we define the slope of the vector as  $h = v_y/v_x$ . We will show through the following two theorems that the function has concave and convex properties along specific vector directions.

**Theorem 1.** Along the direction defined by  $h \geq -kr$ , the function  $V_i = f(x, R_d)$  is concave.

**Proof:** Any vector  $\mathbf{v}$  in  $\{x, R_d\}$  space can be expressed as a linear combination of the eigenvectors:

$$\mathbf{v} = \alpha \mathbf{e}_1 + \beta \mathbf{e}_2. \quad (7)$$

To simplify the expressions, we introduce the following new variable:

$$\xi = kr + \sqrt{k^2 r^2 + 1} > kr \geq 0 \quad (8)$$

The given condition of  $h \geq -kr$  can be written as:

$$h = \frac{\alpha + \beta}{-\alpha/\xi + \beta\xi} \geq -kr \quad (9)$$

Regarding the denominator in the above inequality, there are three cases that we will consider separately.

*Case 1:* The denominator is negative, which can be shown to be impossible.

*Case 2:* Consider the situation that the denominator in (9) is positive. Therefore,

$$\alpha/\beta < \xi^2 \quad (10)$$

By rearranging (9) and substituting (8), we have:

$$\frac{\alpha}{\beta} \geq -\frac{1 + kr\xi}{1 - kr/\xi} = -\frac{1 + kr\xi}{\xi - kr} \cdot \xi = -\xi^2 \quad (11)$$

*Case 3:* When the denominator equals zero, which corresponds to the vector with slope of  $\infty$ , we may say that

$$\alpha/\beta = \xi^2. \quad (12)$$

Integrating the results of (10), (11) and (12) leads to  $(\alpha/\beta)^2 \leq \xi^4$  from which we can show, after some algebra, that

$$\mathbf{v}^T \mathbf{H} \mathbf{v} = \lambda_1 \alpha^2 \|\mathbf{e}_1\|^2 + \lambda_2 \beta^2 \|\mathbf{e}_2\|^2 \leq 0, \quad (13)$$

For any pair of vectors  $\tilde{\mathbf{v}}$  and  $\mathbf{v}_0$ , letting  $\delta \mathbf{v} = \tilde{\mathbf{v}} - \mathbf{v}_0$ , according to Taylor's theorem<sup>3</sup> [14] we have

$$f(\tilde{\mathbf{v}}) = f(\mathbf{v}_0) + \nabla f(\mathbf{v}_0) \delta \mathbf{v} + \frac{1}{2} \delta \mathbf{v}^T \mathbf{H}(\mathbf{v}_0 + \varepsilon \delta \mathbf{v}) \delta \mathbf{v} \quad (14)$$

for some  $\varepsilon$ ,  $0 \leq \varepsilon \leq 1$ . If along the direction where the slope of  $\delta \mathbf{v}$  is no less than  $-kr$ , the third term of *RHS* is non-positive from (13). Hence

$$f(\tilde{\mathbf{v}}) \leq f(\mathbf{v}_0) + \nabla f(\mathbf{v}_0) (\tilde{\mathbf{v}} - \mathbf{v}_0) \quad (15)$$

and we can get the conclusion that  $f(x, R_d)$  is concave. This completes the proof.

Similarly, we can get the following theorem.

**Theorem 2.** Along the direction defined by  $h \leq -kr$ , the function  $V_i = f(x, R_d)$  is convex.

A positive slope belongs to the scenario of Theorem 1 and usually a steep negative slope is the scenario for Theorem 2. Both scenarios can be seen in Figure 3.

To meet the constraint in (1) which states that  $\max_{i \in \text{set of sinks}} (V_i) \leq 0$ , the solution space must be on or below the plane of  $V_i(x, R_d) = 0$ . Hence, we need study the delay violation functions only on this zero plane where the critical

<sup>3</sup>Note that Taylor's theorem differs from a truncated Taylor's expansion.

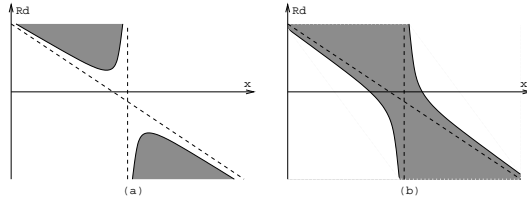


Figure 4: Intersection of  $V_i(x, R_d)$  and zero violation plane.

condition lies. By letting  $V_i = f(x, R_d) = 0$ , we can obtain the intersection between the  $V_i(x, R_d)$  surface and the zero delay violation plane which is defined by:

$$R_d = -krx + \frac{(-kk_1^2r + k_0c + k_1k_2)/c^2}{x - k_1/c} + \frac{k_2 - kk_1r}{c}. \quad (16)$$

This represents a hyperbola in the  $\{x, R_d\}$  space. Two scenarios of such a hyperbola are illustrated in Figure 4. From Theorem 1 and 2, the shaded regions are where the  $f(x, R_d)$  surface is above the  $V_i = 0$  plane, i.e., values of  $V_i$  are positive. In Figure 4, the dashed lines are the asymptotic lines defined by:

$$x = \frac{k_1}{c} \quad (17)$$

$$R_d = -krx + \frac{k_2 - kk_1r}{c}. \quad (18)$$

From conditions in (3) we know that  $x < k_1/c$ , thus we need only care about the regions to the left of the vertical asymptotic line  $x = k_1/c$ , where we can reach the following conclusion.

**Theorem 3.** *If  $x < k_1/c$ ,  $V_i(x, R_d)$  is a monotone increasing function of  $R_d$ .*

This conclusion is consistent with the intuition that delay always goes up as the driver resistance increases.

In fact, the feasible region can be refined further as a rectangle bordered by  $CC \geq x \geq 0$  and  $R_{dMAX} \geq R_d \geq R_{dmin}$ . We refer to this rectangle as a search sheet. From now on, we will restrict our discussions only to the search sheet.

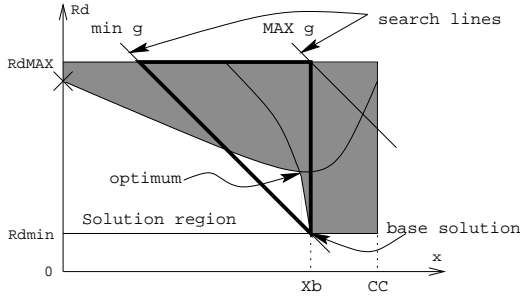


Figure 5: Graphical representation of the regions and search lines on a search sheet.

In order to obtain a better intuitive view of the region on this search sheet that satisfies the timing constraint of (1), we introduce a shading process. For a sink  $i$ , if any part of the  $V_i(x, R_d)$  surface is above the zero delay violation plane, the projection of this part on the search sheet is shaded. The border of this shading region is the same curve as that described in (16). This shading is repeated for all sinks, and the finally unshaded regions are the solution regions. Figure 5 shows an example of this shading process for two sinks. One sink belongs to the situation in Figure 4 (a), the other corresponds to the situation in Figure 4 (b). The shaded area from

these two sinks are overlapped in Figure 5. We can obtain a direct corollary of Theorem 3 as follows.

**Corollary 1.** *If maximum  $V_i$  for all sinks are positive everywhere along a segment on the search sheet, there is no solution above this segment along the positive  $R_d$  direction that satisfies the constraints in (1).*

This property provides us with a convenient method for detecting the existence of solution region. We need only search on the lower border of  $R_d = R_{dmin}$ , which we will call the ground segment. If the solution region is not empty, there must be solutions on this lower border and we define the rightmost solution point on this lower border as a base solution. This point is denoted as  $(x_b, R_{dmin})$ , as shown in Figure 5. From Theorem 1, we know that  $V_i$  for all sinks are concave functions with respect to  $x$  along this lower border. Therefore, we can apply the binary search technique introduced in [7] to our algorithm to find the base solution.

If the solution region is nonempty, the next step will be to find the optimum solution that satisfies the objective specified in (1). For the local situation of Figure 2, we reframe the objective (1) to an equivalent form and restate the problem as:

$$\begin{aligned} &\text{maximize} && rx + e^\gamma R_d \\ &\text{subject to} && (x, R_d) \in \text{the unshaded region.} \end{aligned} \quad (19)$$

This objective can be described by the straight line:

$$rx + e^\gamma R_d = g. \quad (20)$$

The optimum solution corresponds to the straight line in this family that intersects solution region with the maximum  $g$ . By varying  $g$  from a large value to a small value, this search line makes a sweep from the upper right corner towards the lower left corner of the search sheet. During the sweep, the position where a solution is first encountered corresponds to the location of the optimum solution.

One observation is that the lower bound of the search range does not need to be at the lower left corner of the search sheet. Since the base solution is optimal on the ground segment, any solution from search line with smaller  $g$  than that of the base solution would be sub-optimal. Therefore, we can begin with a minimum  $g$  corresponding to this base solution. Another observation is that there is no solution above the ground segment between  $x_b$  and  $CC$ . Hence, by Corollary 1, we can begin with a maximum  $g$  at  $(x_b, R_{dMAX})$ . The search lines with minimum  $g$  and maximum  $g$  are demonstrated in Figure 5. The actual search range is within the darkened triangle.

The search line example shown in Figure 5 is for a specific weighting factor  $\gamma$ . If the  $\gamma$  is much smaller, the search line would be steeper and the optimum solution is more likely to be at the base solution position. On the other hand, if the  $\gamma$  is much larger, the search line is closer to horizontal and the crossed point in Figure 5 would be more likely to be the optimum solution. Recall that one asymptotic line of the intersection hyperbola is defined by the equation (18) whose slope is  $-kr$ . When  $\gamma > 0$ , the slope of the search line will be greater than  $-kr$  and search result will tend to move toward the crossed point that emphasizes the driver cost. When  $\gamma < 0$ , the search line will be steeper than the line (18) and a more wire cost oriented result at the base solution has more chance to be obtained.

These properties are derived from Elmore delays. Though the Elmore delay may have large errors for specific points, its qualitative fidelity is still true [1] and able to serve as good strategic guide. Our experimental results also support this assertion.

### 3 Algorithm

From the properties above, it is straightforward to construct our algorithm. The algorithm consists of two phases. Phase I, called

SART (Steiner AWE Routing Tree) is similar to SERT except that the Elmore model is replaced by fourth order AWE. The output is a routing tree  $T'$ .

In SART, starting with a single source, a partial routing tree grows in a greedy fashion. In each growing step, a previously unconnected sink is selected and connected to a certain node in the partial tree such that the maximum delay is minimized.

---

**Algorithm: FAR-DS.**

**Input:**  $T'$  from SART.

**Output:** FAR-DS tree  $T$ .

1. sort sinks in ascending order according to distance to  $s_0$ ;
  2. for  $i = n; i > 0; i - -$
  3. default optimal tree = current tree;
  4. disconnect  $s_i$  and its downstream subtree  $T_{s_i}$ ;
  5. for each maximal segment  $m_j$  in  $T' \setminus T_{s_i}$
  6. if no solution exist on ground segment for  $(s_i, m_j)$  continue;
  7. else Sweep\_search\_line( $s_i, m_j, x_b$ ):
  8.  $g_{min} = rx_b + e^{\gamma} R_{dmin}$ ;
  9.  $g_{MAX} = rx_b + e^{\gamma} R_{dMAX}$ ;
  10. if  $\exists$  solution at  $g_{MAX}$
  11. if  $\exists$  improvement, update optimal tree;
  12. continue;
  13.  $g_{lower} = g_{min}; g_{upper} = g_{MAX}$ ;
  14. Binary\_search( $s_i, m_j, g_{lower}, g_{upper}$ ):
  15.  $g_{mid} = (g_{lower} + g_{upper})/2$ ;
  16. along search line  $g_{mid} = rx + e^{\gamma} R_d$ :
  17. find its intersection with border,  
 $p_l(x_l, R_{dl})$  and  $p_r(x_r, R_{dr})$ ;
  18.  $p_m = (p_l + p_r)/2$ ;
  19. compute  $V_i \forall i$  at  $p_l, p_m, p_r$ ;
  20. for each sink  $s_i$
  21. determine  $V_i$  curve via a quadratic fit;
  22. search  $x$  where  $V_i \leq 0$ ;
  23. if  $\exists$  solution
  24. if  $\exists$  improvement, update optimal tree;
  25. Binary\_search( $s_i, m_j, g_{mid}, g_{upper}$ );
  26. else Binary\_search( $s_i, m_j, g_{lower}, g_{mid}$ );
  27. connect  $s_i$  according to optimal tree;
- 

The above is the algorithm of Phase II. During Phase II, in the descending order of distance to source, each sink is reconnected to the routing tree to meet the objective. While reconnecting any sink  $s_i$ , the sink and its downstream  $T_{s_i}$  subtree is disconnected first. It is then connected to each maximal segment on the routing tree  $T \setminus T_{s_i}$  tentatively to find the optimum solution.

The tentative connection point to each maximal segment is determined by a searching process based on the properties described in section 2.3. At the beginning, the search is conducted along the lower border (see figure 5) to detect the existence of a solution. Since  $V_i$  is a concave function along this border, the implementation here is the same as the binary search in MVERT [7]. The output is the base solution  $(x_b, R_{dmin})$ .

If a solution is found on the lower border, we continue to sweep the search line defined by equation (20) to search for the optimum solution. This sweep is performed with  $g$  confined by  $g_{min} \leq g \leq g_{MAX}$ . The value of  $g_{min}$  is that causes the search line to pass through  $(x_b, R_{dmin})$  while  $g_{MAX}$  corresponds to a line incident with  $(x_b, R_{dMAX})$ . From Corollary 1, if we cannot find a solution on a search line for a specific  $g$ , we try a line with smaller  $g$ , otherwise we go on to search the region above, with the choice of  $g$  being made in a binary search fashion.

Along each search line position, as  $V_i(x, R_d)$  may be either concave or convex, we use quadratic fitting to search a solution instead of binary search. First the intersection points between the

search line and the search sheet borders are found. The two intersection points serve as the end points of the search interval along the search line. The left end point is  $p_l(x_l, R_{dl})$ , the right end point is  $p_r(x_r, R_{dr})$  and the middle point is  $p_m$ . The connection and driver resistance are set to each of the three points and AWE are invoked to evaluate the delays at all the sinks. Then, for each sink  $s_i$ , its  $V_i$  at the three points are available for a quadratic fitting to get an approximate  $V_i$  curve with respect to  $(x, R_d)$  along the search line. If any of the  $V_i$  curves is positive over the entire interval between  $p_l$  and  $p_r$ , we can conclude that there is no solution along this search line position. Otherwise, the non-positive ranges for all the  $V_i$  are overlapped to check if a solution exists.

Each local optimum obtained for a specific  $s_i$  and a maximal segment will be compared with the current global solution, and if it is better according to the objective of (1), then the current "global" solution is updated and stored. After all of the maximal segments have been searched, the sink  $s_i$  is connected to the point corresponding to the global solution.

## 4 Experimental results

The experimental results are shown in Table 2 and Table 3 for IC and MCM technology, respectively. The parameters for MCM are from [1] and the IC parameters are for 0.18  $\mu\text{m}$  technology which are also scaled from [1]. The nets for test are generated randomly, and the number of sinks for each sink varies from 5 to 17. The value of  $(R_{dmin}, R_{dMAX})$  were chosen as (30, 600) for the IC technology and (5, 150) for the MCM technology. The leftmost column is the number of sinks for each test. The  $W$  is the total wire cost and the  $V_{MAX}$  is the maximum delay violation.

The results from SERT and MVERT are also listed for comparison. If the timing constraints are stringent, SERT cannot always satisfy the delay specifications. Even MVERT cannot always satisfy the timing constraints, though it usually makes significant improvement on wire cost. This is due to errors from the Elmore delay model, or because the specification is unachievable without driver sizing.

If violations from SERT or MVERT are not too large, they can usually be reduced to be smaller than zero by our algorithm. Besides this, the experimental result showed the adaptive nature of our algorithm, which is indicated by the fact that there is virtually no timing slack in the solutions.

Two values for the weighting factor  $\gamma$  are tested. When  $\gamma < 0$ , the objective is more wire length reduction oriented which result in an average of 20% reduction of wire length for IC and 25% for MCM. When  $\gamma > 0$ , our algorithm focuses more on reducing the driver size, or enlarging the driver resistance. In this case, the average improvement in the driver cost is about 42% for 0.18  $\mu\text{m}$  IC and 57% for MCM.

The right most column  $R_d^*$  shows results for the same nets, but with relaxed timing constraints. The wire cost reductions are similar to those before, while the driver costs are greatly reduced. This also shows the adaptive nature of our methods that can fulfill any potential cost reduction for real situation.

Another observation from our experience is that the choice of good  $\gamma$  depends on how large the net is. For a small net with fewer sinks and sinks close to the source, the value of  $\gamma$  should be smaller; on the other hand,  $\gamma$  should be larger for large nets to get similar effects as those of small nets.

In our experiments, the time cost of the computation is usually within a half minute for nets of up to 10 sinks. In the worst case for net with up to 20 sinks, the run time can be up to three minutes. On the whole, the computation cost of our algorithm is reasonable. These experiments were run at a SUN Ultra-10 station.

## 5 Conclusion

In this paper we propose a fourth order AWE model based routing algorithm that can explore both Hanan and non-Hanan points. We also suggest a new combined objective to accommodate the improvement of both wire and driver cost subject to timing constraints. Our algorithm achieves this objective in such a way that flexibility is provided to the user to emphasize either cost, depending on the desired situation. For pre-placed buffer positions, this algorithm can also be extended to buffer sizing directly. Experiments results shows significant improvements in delay and both kinds of cost.

## References

- [1] K. D. Boese, A. B. Kahng, B. A. McCoy and G. Robins, "Near-optimal critical sink routing tree constructions," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 14, No. 12, pp. 1417-36, Dec. 1995.
- [2] J. Lillis and P. Buch, "Table-lookup methods for improved performance-driven routing," *Proceedings of the ACM/IEEE Design Automation Conference*, pp. 368-373, 1998.
- [3] J. Cong and C. K. Koh, "Interconnect layout optimization under higher-order RLC model," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 713-720, 1997.
- [4] J. Lillis, C. K. Cheng, T. T. Lin and C. Y. Ho, "New performance driven routing techniques with explicit area/delay tradeoff and simultaneous wire sizing," *Proceedings of the 33rd ACM/IEEE Design Automation Conference*, pp. 395-400, Jun. 1996.
- [5] F. J. Liu, J. Lillis and C. K. Cheng, "Design and implementation of a global router based on a new layout-driven timing model with three poles," *Proceedings of the IEEE International Symposium on Circuits and Systems*, 1997.
- [6] S. S. Sapatnekar, "RC interconnect optimization under the Elmore delay model," *Proceedings of the ACM/IEEE Design Automation Conference*, pp. 392-396, 1994.
- [7] H. Hou and S. S. Sapatnekar, "Routing tree topology construction to meet interconnect timing constraints", *Proceedings of the International Symposium on Physical Design*, pp. 205-210, 1998.
- [8] W. C. Elmore, "The transient response of damped linear network with particular regard to wideband amplifiers," *Journal of Applied Physics*, Vol. 19, pp. 55-63, 1948.
- [9] L. T. Pillage and R. A. Rohrer, "Asymptotic waveform evaluation for timing analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 9, No. 4, pp. 352-366, Apr. 1990.
- [10] J. Qian, S. Pullela and L. T. Pillage, "Modeling the effective capacitance for the RC interconnect of CMOS gates," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 13, No. 12, pp. 1526-35, Dec. 1994.
- [11] J. Rubinstein, P. Penfield and M. A. Horowitz, "Signal delay in RC tree networks," *IEEE Transactions on Computer-Aided Design*, Vol. CAD-2, No. 3, pp. 202-211, July 1983.
- [12] R. Gupta, B. Krauter, B. Tutuianu, J. Willis and L. T. Pillage, "The Elmore delay as a bound for RC trees with generalized input signals," *Proc. 33rd ACM/IEEE Design Automation Conference*, 1995.
- [13] C. L. Ratzlaff, N. Gopal and L. T. Pillage, "RICE: rapid interconnect circuit evaluator," *Proc. 28th ACM/IEEE Design Automation Conference*, pp. 555-560, 1991.
- [14] D. G. Luenberger, "Linear and Nonlinear Programming," *Addison-Wesley Publishing Company, Inc.*, 1984.

Table 2: Experimental results on .18 $\mu$ m IC,  $R_d = 300$  for SERT and MVERT

n	SERT		MVERT		FAR-DS $\gamma < 0$			FAR-DS $\gamma > 0$			
	W	$V_{MAX}$	W	$V_{MAX}$	W	$R_d$	$V_{MAX}$	W	$R_d$	$V_{MAX}$	$R_d^*$
5	226	4.49	244	-1.54	187	30	-0.10	281	453	-0.00	600
5	221	-1.13	171	1.86	166	302	-0.01	207	422	-0.02	600
9	252	-3.37	250	3.08	212	141	-0.01	257	444	-0.00	600
9	367	3.20	270	-2.00	258	332	-0.05	266	413	-0.05	587
13	407	0.82	315	-1.11	281	319	-0.01	341	364	-0.04	524
13	452	-0.64	376	3.05	339	217	-0.09	371	422	-0.02	547
17	395	-1.87	323	-2.42	298	244	-0.13	341	466	-0.01	587
17	477	-0.53	417	-1.76	339	61	-0.03	368	426	-0.04	551

Table 3: Experimental results on MCM,  $R_d = 25$  for SERT and MVERT

n	SERT		MVERT		FAR-DS $\gamma < 0$			FAR-DS $\gamma > 0$			
	W	$V_{MAX}$	W	$V_{MAX}$	W	$R_d$	$V_{MAX}$	W	$R_d$	$V_{MAX}$	$R_d^*$
5	237	-0.75	200	0.81	182	35	-0.01	244	46	-0.00	150
5	277	-1.20	224	2.26	219	32	-0.00	295	38	-0.00	142
9	451	2.16	405	0.96	333	5	-0.01	468	38	-0.00	150
9	434	-0.51	376	-0.41	356	18	-0.07	418	31	-0.04	150
13	677	2.36	624	-0.48	591	19	-0.17	700	32	-0.01	134
13	626	0.80	529	-2.08	515	24	-0.08	593	38	-0.10	133
17	945	-2.25	852	1.24	804	34	-0.07	859	51	-0.14	118
17	783	-0.60	693	-0.76	613	12	-0.03	740	39	-0.01	101