Interconnect Optimization in Deep Sub-micron Design under the Transmission Line Model

by

Youxin Gao, M.A.; M.S.; B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2000

Interconnect Optimization in Deep Sub-micron Design under the Transmission Line Model

Approved by
Dissertation Committee:

Martin D.F. Wong (Supervisor)

Charles J. Alpert

Donald S. Fussell

Stephen W. Keckler

Brian Mulvaney

Dedicated to my loving wife, Haiyun

Acknowledgments

I would like to thank my supervisor Prof. Martin D.F. Wong for his encouragement, careful guidance and invaluable support throughout these years.

Youxin Gao

The University of Texas at Austin
August 2000

Interconnect Optimization in Deep Sub-micron Design under the Transmission Line Model

Publication No. _____

Youxin Gao, Ph.D.
The University of Texas at Austin, 2000

Supervisor: Martin D.F. Wong

As the VLSI technology has been scaled down to $0.18\mu m$ in recent years and is expected to be scaled down to $0.05\mu m$ in the near future, interconnect delay becomes an important factor in achieving high performance. In deep sub-micron design, interconnect delay is shown to be 10 to a few hundred times bigger than the intrinsic gate delay for a global interconnect, and thus dominates the circuit delay. To reduce interconnect delay, wire-sizing and buffer insertion/sizing are two effective techniques. One of the approaches to wire-sizing is continuous wire-sizing. In continuous wire-sizing, the shape of a wire is described by a continuous function, and the objective is to find a shape function which minimizes delay or minimizes area subject to a delay bound.

In the first part of this dissertation, we present some continuous wire-sizing results under the Elmore delay model. Comparing with previous work, our algorithm can not only deal with uni-directional wires but also bi-directional wires, where both solutions are in closed form. We also develop an efficient algorithm to determine the

optimal wire shape for wires with one or two neighboring wires. These results are useful in optimization under the transmission line model, since they can give very good initial solutions.

However, since it is well known that the Elmore delay model is not an accurate delay model, those optimal results determined under the Elmore delay model may be inferior. In the second part of this dissertation, we present some wire-sizing results under the transmission line model. For a special case where fringing capacitance and inductance are not considered, we solve the transmission line equations analytically and derive a closed form solution on the transient response for an exponential wire shape $f(x) = ae^{-bx}$. We then determine a and b such that either delay is minimized or area is minimized subject to a target delay bound. For a general case where fringing capacitance and inductance are considered, we solve the transmission line equations by using the Picard-Carson method. We then develop a three pole based delay model. Analytical expressions for estimating delay at any threshold voltage and overshoot/undershoot voltage are further derived. The optimal wire shape is determined to minimize delay or area subject to undershoot voltage constraint. To calculate delays for interconnects with buffers inserted, we combine the wire delay model under the transmission line model with a buffer macromodel (k-factor equations) to provide a fast and accurate delay estimation method.

In the third part of this dissertation, we present a graph based algorithm for optimal buffer insertion under accurate delay models. The algorithm determines the number of buffers and their locations on a wire such that some optimization objective is satisfied. Two typical examples of such optimization objectives are minimizing the 50% threshold delay and minimizing the transition time. Both can be easily determined in our algorithm. We show that the buffer insertion problem is a shortest path problem. Our algorithm can be easily extended for simultaneous buffer insertion and wire-sizing, and complexity is still polynomial. The algorithm

can also be extended to deal with problems such as buffer insertion subject to transition time constraints at any position along the wire.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

As it is predicted by National Technology Roadmap for Semiconductors (NTRS) [3], the feature size of VLSI devices is gradually scaled down from 180nm in 1999, to 130nm in 2002, and is expected to be further reduced to 50nm in 2011. In Table 1.1, we list some main characteristics for each technology generation in NTRS and some estimated parasitic parameters for Metal 4 interconnects. These parasitic parameters (e.g., C_a and C_f) are estimated by using Sakurai's approximated formulas [62]. As the feature size decreases, chips become more congested and many new challenging issues arise. One challenging issue is that interconnect delay has dominated over gate delay in determining circuit performance. According to the simple scaling rule described in [4], if devices and interconnects are all scaled down by a factor S, the intrinsic gate delay is reduced by a factor S, and the delay of local interconnects remains almost the same, but the delay of global interconnects is increased by a factor of S^2. It is estimated that at least 20% of the cycle time is consumed by interconnect delay [21].

As technology goes on to deep sub-micron, characterizing interconnect delay is also becoming more and more complicated. Roughly speaking, interconnect delay is proportional to driver resistance, wire resistance, wire capacitance as well as load

Year	1999	2002	2005	2008	2011
Technology(nm)	180	130	100	70	50
Metal height/ width aspect ratio	1.8	2.1	2.4	2.7	3
On-chip clock frequency (GHz)	1.25	2.1	3.5	6	10
Wavelength(μm)	24000	14286	8571	5000	3000
Metal 4 interconnect					
Minimum W(nm)	450	325	250	175	125
$R(\Omega/\mu m)$	0.082	0.14	0.2	0.36	0.64
$L(pH/\mu m)$	0.99	1.02	1.05	1.08	1.10
$C_a(fF/\mu m)$	0.040	0.040	0.040	0.040	0.039
$C_f(fF/\mu m)$	0.11	0.11	0.12	0.12	0.12
$C_c(fF/\mu m)$	0.062	0.070	0.079	0.088	0.097
$C(fF/\mu m)$	0.21	0.22	0.23	0.25	0.26

Table 1.1: NTRS predicted technology trend and our estimated parasitic parameters for Metal 4 interconnects. R is the unit length resistance, L is the unit length inductance, and C_a, C_f, C_c are the unit length area, fringing, coupling capacitances, respectively. The total unit length capacitance is $C = C_a + C_f + C_c$.

capacitance. Wire capacitance is undoubtedly a major component in determining interconnect delay. Traditional methods only consider interconnect area capacitance as the unique source of capacitance when the feature size is around $1\mu m$. As the feature size decreases, the wire becomes thinner but thicker, i.e., the wire width is smaller than the wire thickness. Such trend is clearly reflected as the increasing wire height/width aspect ratio in Table 1.1. As a result, fringing capacitance becomes more important than area capacitance. As it is shown in Figure 1.1, area capacitance is due to the uniform electrical field between wire plate and substrate, and fringing capacitance is due to the non-uniform electrical field between side walls and substrate. On the other hand, since the spacing between neighboring wires also shrinks, coupling capacitance between two adjacent wires becomes another major component of total capacitance. Coupling capacitance is influenced not only by the decreased spacing but also by the increased height/width aspect ratio (Figure 1.1).

Cross coupling between signals can affect the propagation delay significantly, and also cause fault switch (crosstalk noise), which is a potential problem in high speed circuits [64]. Inductance effect is another important factor which can affect interconnect delay significantly. While such effect is not prominent in local interconnects, it is more evident in longer interconnects, such as those in power supply distribution, clock distribution and Multi-chip module (MCM) interconnects [21, 41]. It is necessary to take all these effects into account when we analyze interconnect delay.

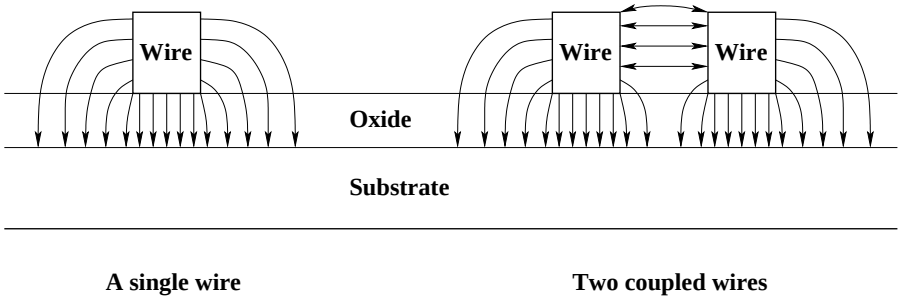


Figure 1.1: Capacitance models for a single wire and two adjacent wires.

Not only there is a trend in technology generation, but also interconnect delay models used in layout design have evolved from simple models to sophisticated models. Interconnect delay modeling is crucial in layout design, because it is the fundamental part of layout optimization techniques such as wire-sizing optimization, buffer insertion and performance driven routing [13, 53]. The Elmore delay model [22] is the most commonly used delay model in VLSI layout design. The advantage of the Elmore delay is that it has simple algebraic form and linear computational complexity in terms of the number of nodes in the circuit. The Elmore delay is also easy to be represented as a function of circuit design parameters such as wire width and length.

However, since the Elmore delay model has limited accuracy, some other more accurate delay models are necessary and they have already been proposed. In

all these models, calculating high order moments or poles in the transfer function is the essential part. Depending on the method on calculating moments/poles, these delay models can be classified into two main categories, one is the lumped circuit approximation based, and the other is the transmission line model based. In the lumped circuit approximation, each wire is divided into several identical segments, then each segment is modeled as a lumped RC/RLC circuit segment. Algorithms such as those in [26, 45, 53, 56] are proposed to calculate moments/poles for such lumped circuits. In general, the more number of segments the more accurate the model is, but it takes longer in computation. Another disadvantage of the lumped circuit approach is that it can produce spurious oscillations that appear due to the resonant behavior of segments. Since the number of segments is related to the signal wavelength λ ($\lambda = cT_r$, where c is the speed of light and T_r is the signal transition time), the lumped circuit approximation is efficient and accurate in dealing with electrically short wires, where the wire length is much smaller than the signal wavelength. The transmission line model, on the other hand, describes voltages and currents at any position on a wire by telegraph's equations [36]. The transfer function is obtained by solving these differential equations. Comparing with the lumped circuit approximation, the transmission line model does not depend on signal wavelength or rising time and also has higher accuracy.

Another important technology trend is that, as the feature size is scaled down, the operating frequency is also gradually increased from around 1GHz today to 10GHz in 2011. As a result, the interconnect wire length becomes comparable to the signal wavelength. Some estimated signal wavelengths are listed in Table 1.1, where we assume the signal rising time is one-tenth of the cycle time. Note that some signals in reality have faster rising time than our estimations. As a result, the inductive impedance becomes comparable to wire parasitic resistance, especially in wide interconnects, such as upper metal lines and MCM interconnects

[4, 48]. Therefore the transmission line effects such as reflections, dispersion and ringing become evident. These effects can not be characterized by the lumped circuit approximation.

A direct impact on layout design due to technology scaling is that, techniques which are aiming at reducing interconnect delay and crosstalk effects become necessary and important [13]. Wire-sizing optimization is one of such techniques which can reduce interconnect delay. The algorithms shown in [16, 17] divide a wire into several segments and assign each wire segment with a proper width such that the Elmore delay [22] is minimized. As the number of segments becomes large, it is further found that continuous wire-sizing is more efficient. In continuous wire-sizing, the wire shape is described as a continuous function $f(x)$, and the objective is to find $f(x)$ such that delay is minimized. In [7] and [25], it is found that the optimal shape under the Elmore delay model is an exponential function if area capacitance is the only source of capacitance. With fringing capacitance consideration, it is shown in [9] that the optimal wire shape function can be expressed in terms of the $Lambert's$ W function.

Buffer insertion is another effective technique which can not only reduce interconnect delay, but also improve the signal waveform by reducing the transition time. In [68], the buffer's position is determined by a dynamic programming algorithm, where it uses the Elmore delay model and a linear gate model. Other similar algorithms based on higher order delay models are proposed in [2, 52, 54]. Buffer insertion is also found an efficient technique in minimizing coupled noise [1, 8, 50]. Moreover, algorithms for simultaneous buffer insertion and wire-sizing have also been proposed. Under the Elmore delay model and a linear gate model, closed form solutions for delay minimization by simultaneous wire-sizing and buffer insertion/sizing are proposed in [10, 11].

This thesis presents some optimization results on both continuous wire-sizing

and buffer insertion. In Chapter 2, we present some wire-sizing results under the Elmore delay model. A closed form solution is presented in section 2.1. Our method can not only deal with uni-directional wires but also bi-directional wires. Since a uni-directional wire is a special case, we present the solution for bi-directional wires. In section 2.2, we determine the optimal wire shape for a wire with one or two neighboring wires. Although our goal in interconnect optimization is wire-sizing under the transmission line model, the wire-sizing results under the Elmore delay model are still useful. They can provide very good initial solutions for optimizations under the transmission line model. In Chapter 3, we consider a special case, where fringing capacitance and inductance are not considered. In Chapter 4, both fringing capacitance and inductance are taken back into consideration, but a different technique is used to solve the transmission line equations. We also derive analytical formulas for estimating overshoot/undershoot voltages. An extension to transmission line trees is presented in section 4.8. In Chapter 5, a fast and accurate delay estimation method under the transmission line model and k-factor equations is developed for buffered interconnects. In Chapter 6, we present a graph based algorithm for optimal buffer insertion under accurate delay models.

Chapter 2

Wire-sizing under the Elmore Delay Model

2.1 Closed Form Solution on Wire-sizing

Wire-sizing optimization is one of the effective techniques which can reduce interconnect delay. In [7] and [25], it is found that the optimal shape under the Elmore delay model is an exponential function, if the area capacitance is the unique source of wire capacitance. With fringing capacitance consideration, it is shown in [9] that the optimal wire shape function can be expressed in terms of the *Lambert's W* function [1] in a closed form. At the same time, Fishburn in [24] independently presented a power series solution to the optimal shape function.

In this section, we present a closed form solution on wire-sizing by using calculus of variations. Comparing with previous studies, our method can not only deal with uni-directional wires but also bi-directional wires (bi-directional wires are

[1] The *Lambert's W* function [20] was first introduced by Euler in 1779 when he studied Lambert's transcendental equation. $W(x)$ function is defined as the value of w that satisfies $we^w = x$. The Lambert's W function is a smooth function. It is available as a standard library function in Maple V [38].

present in interconnects with multiple sources [12, 13, 18]). Since a uni-directional wire is a special case of bi-directional wires, we present results on bi-directional wires and show how to reduce to uni-directional wires. Given a bi-directional wire of length L, let $f(x)$ be the width of the wire at position x, $0 \leq x \leq L$. Let T_{DR} be the right-to-left delay. Let T_{DL} be the left-to-right delay. Let $T_{BD} = \alpha T_{DR} + \beta T_{DL}$ be the total weighted delay where $\alpha \geq 0$ and $\beta \geq 0$ are given weights such that $\alpha + \beta = 1$. We determine $f(x)$ such that T_{BD} is minimized. Our study shows that, if $\alpha = \beta$, the optimal shape function is $f(x) = c$, for some constant c; if $\alpha \neq \beta$, the optimal shape function can be expressed in terms of the *Lambert's W* function as $f(x) = -\frac{c_f}{2c_0}(\frac{1}{W(-ae^{-bx})} + 1)$, where c_f is the unit length fringing capacitance, c_0 is the unit area capacitance, a and b are constants in terms of the given circuit parameters. If $\alpha = 0$ or $\beta = 0$, our result gives the optimal shape function for a uni-directional wire [29, 31].

2.1.1 Capacitance Model

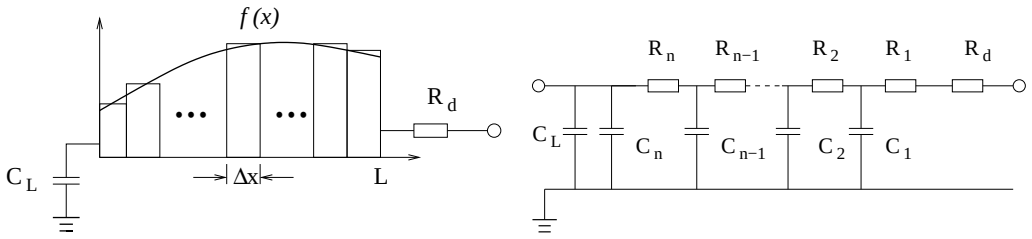
For a uniform wire segment with width W and thickness T, Sakurai's formula [62] gives a good approximation for calculating the unit length capacitance,

$$\frac{C}{\epsilon_{ox}} = 1.15\Big(\frac{W}{T_{ox}}\Big) + 2.28\Big(\frac{T}{T_{ox}}\Big)^{0.222} \tag{2.1}$$

where $\epsilon_{ox} = 3.9 \times 8.855 \times 10^{-14} F/cm$ is the dielectric constant of the insulator (e.g., SiO_2), T_{ox} is the thickness of silicon dioxide. The formula can be simplified as:

$$C_{tot} = c_0 W + c_f \tag{2.2}$$

where $c_0 = \frac{1}{T_{ox}} 1.15 \epsilon_{ox}$ is the unit area capacitance, and $c_f = \epsilon_{ox} 2.80 (\frac{T}{T_{ox}})^{0.222}$ represents the unit length fringing capacitance.



(a) (b)

Figure 2.1: The distributed circuit model for calculating the Elmore delay. R_d is the driver resistance, and C_l is the load capacitance. R_i and C_i are the resistance and capacitance of segment i, respectively. (a) A nonuniform wire whose width is given by a function $f(x)$; (b) its corresponding distributed RC network.

2.1.2 The Elmore Delay Model for Non-uniform Wires

Given a non-uniformly sized wire whose width is given by a function $f(x)$ as shown in Figure 2.1(a), we partition it into n equal-length wire segments, each of length $\Delta x = \frac{L}{n}$, and approximate it as a distributed RC network shown in Figure 2.1(b). Let $x_i = i\Delta x$, $1 \leq x \leq n$. The capacitance and resistance of segment i can be approximated by $C_i = (c_0 f(x_i) + c_f)\Delta x$ and $R_i = r_0 \Delta x / f(x_i)$, respectively, where c_0 is unit area capacitance, c_f is unit length fringing capacitance, and r_0 is unit resistance. The Elmore delay [22] of the distributed RC network [61] is given by

$$T_D = \sum_{i=1}^{n} \frac{r_0 \Delta x}{f(x)} \Big(\sum_{j=i}^{n} (c_0 f(x_j) + c_f)\Delta x + C_l \Big) + R_d \Big(C_l + \sum_{i=1}^{n} (c_0 f(x_i) + c_f)\Delta x \Big)$$

where T_D is equivalent to the sum over all segment resistance (R_i) multiplied by its down stream capacitance ([22, 61]). As $n \to \infty$, T_D thus becomes

$$T_D = \int_0^L \frac{r_0}{f(x)} \Big(\int_0^x c_0 f(t)dt + \int_0^x c_f dt + C_l \Big) dx + R_d \Big(C_l + \int_0^L (c_0 f(x) + c_f)dx \Big) \quad (2.3)$$

Moreover, T_D can be written in a slightly different way, which is given by the following lemma.

Lemma 1

$$T_D = \int_0^L (c_0 f(x) + c_f)\Big(R_d + \int_x^L \frac{r_0 dt}{f(t)}\Big)dx + C_l\Big(R_d + \int_0^L \frac{r_0 dx}{f(x)}\Big) \qquad (2.4)$$

Lemma 1 can be proved by changing the order of integration in (2.3). It follows from Lemma 1 that T_D is also equivalent to the sum over all segment capacitance multiplied by its upstream resistance. In the next subsection, we will start from the delay expression in (2.4) and give a closed form solution for the optimal shape function. The reason of not using the delay expression in (2.3) will be explained later.

2.1.3 Optimal Wire Shape Function

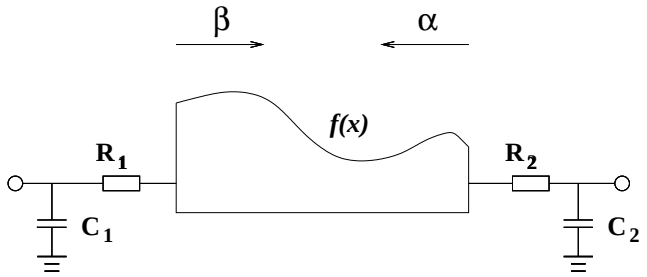


Figure 2.2: A bi-directional wire, with R_1 and R_2 as the driver resistances, and C_1 and C_2 as the load capacitances.

For a bi-directional wire shown in Figure 2.2, R_1 and R_2 are the driver resistances, and C_1 and C_2 are the load capacitances. Let T_{DR} be the delay when the signal is transmitted from right to left. Let T_{DL} be the delay when the signal is transmitted from left to right. According to Lemma 1, we have

$$T_{DR} = \int_0^L (c_f + c_0 f(x))\Big(\int_x^L \frac{r_0 dt}{f(t)} + R_2\Big)dx + C_1\Big(R_2 + r_0 \int_0^L \frac{dx}{f(x)}\Big)$$

$$T_{DL} = \int_0^L (c_f + c_0 f(x))\Big(\int_0^x \frac{r_0 dt}{f(t)} + R_1\Big)dx + C_2\Big(R_1 + r_0 \int_0^L \frac{dx}{f(x)}\Big)$$

We define the total weighted delay T_{BD} as

$$T_{BD} = \alpha T_{DR} + \beta T_{DL} \tag{2.5}$$

where $\alpha \geq 0$ and $\beta \geq 0$ are given weights such that $\alpha + \beta = 1$.

In the rest of this subsection, we will derive the optimal wire shape function so that the weighted delay T_{BD} is minimized.

Lemma 2 *For a bi-directional wire, if $\alpha = \beta$, the optimal shape function is the constant function*

$$f(x) = \sqrt{\frac{r_0 c_f L + r_0(C_1 + C_2)}{c_0(R_1 + R_2)}} \tag{2.6}$$

Proof: We introduce $u(x) = \int_x^L \frac{dx}{f(x)}$, so that $f(x) = -\frac{1}{u'(x)}$, $u(L) = 0$, and $\int_0^x \frac{dx}{f(x)} = u(0) - u(x)$. T_{BD} in (2.5) can be rewritten as an integration over $x, u(x), u'(x)$ as

$$T_{BD} = \alpha R_2(C_1 + c_f L) + \beta R_1(C_2 + c_f L) + \int_0^L F(u(x), u'(x))dx$$

where

$$\begin{aligned} F(u(x), u'(x)) &= -r_0(\alpha C_1 + \beta C_2)u'(x) - \frac{c_0}{u'(x)} \times (\alpha R_2 + \beta R_1) \\ &+ r_0(c_f - \frac{c_0}{u'(x)}) \times (\beta u(0) + (\alpha - \beta)u(x)) \end{aligned}$$

Note that T_{BD} is a functional (i.e., a function of function). Based on calculus of variations [23], the necessary condition to minimize T_{BD} is that $u(x)$ must satisfy the Euler's differential equation:

$$F_u(x, u(x), u'(x)) = \frac{d}{dx} F_{u'}(x, u(x), u'(x)) \tag{2.7}$$

Since $F(x, u(x), u'(x))$ in our problem does not contain x explicitly, the Euler's equation becomes

$$F - u' F_{u'} = c$$

i.e.,

$$-\frac{2c_0}{u'(x)}R_{12} + r_0(c_f - \frac{2c_0}{u'(x)})(\beta u(0) + \eta u(x)) = c \tag{2.8}$$

where c is a constant to be determined, $R_{12} = \alpha R_2 + \beta R_1$, and $\eta = \alpha - \beta$. By defining

$$c^* = c - r_0 c_f \beta u(0) \tag{2.9}$$

$$R_{12}^* = R_{12} + r_0 \beta u(0) \tag{2.10}$$

where $u(0)$ is a constant, equation (2.8) can be simplified as

$$\frac{c^* - r_0 c_f \eta u(x)}{-2c_0(R_{12}^* + r_0 \eta u(x))} u'(x) = 1 \tag{2.11}$$

Since $\alpha = \beta$, i.e., $\eta = 0$, equation (2.11) thus becomes

$$\frac{c^*}{-2c_0 R_{12}^*} u'(x) = 1$$

Obviously, the solution is $f(x) = -\frac{1}{u'(x)} = \frac{c^*}{2c_0 R_{12}^*}$. This means the optimal shape function is a uniform width function. In fact, the constant can be determined directly from the expression of T_{BD}. If we let $f(x) = f$, where f is a constant, then the delay can be written as,

$$\begin{aligned} T_{BD} &= \alpha R_2(C_1 + c_f L) + \beta R_1(C_2 + c_f L) + \frac{r_0 L}{f}(\alpha C_1 + \beta C_2) \\ &+ c_0 f L(\alpha R_2 + \beta R_1) + \frac{L^2 r_0}{2f}(c_f + c_0 f) \end{aligned}$$

Thus $\frac{dT_{BD}}{df} = 0$ gives

$$f = \sqrt{\frac{r_0 c_f L + r_0(C_1 + C_2)}{c_0(R_1 + R_2)}}$$

□

Lemma 3 *Given a bi-directional wire, if $\alpha \neq \beta$, the optimal shape function is*

$$f(x) = -\frac{c_f}{2c_0}(\frac{1}{W(-ae^{-bx})} + 1) \tag{2.12}$$

where $W(x)$ is the Lambert's W function, and

$$a = \frac{R_{12}^* c_f}{R_{12}^* c_f + c^*} \cdot exp(\frac{2r_0 c_0 \eta L - R_{12}^* c_f}{R_{12}^* c_f + c^*}) \tag{2.13}$$

$$b = \frac{2\eta r_0 c_0}{R_{12}^* c_f + c^*} \tag{2.14}$$

Proof: Under the condition $u(L) = 0$, the solution to equation (2.11) is

$$u(x) = \frac{R_{12}^*}{\eta r_0}\Big[exp(\frac{\eta r_0 (c_f u - 2c_0(x-L))}{R_{12}^* c_f + c^*}) - 1\Big] \tag{2.15}$$

The wire shape function $f(x)$ is thus

$$f(x) = -\frac{1}{u'(x)} = \frac{R_{12}^* c_f + c^*}{R_{12}^* + r_0 u(x)} \frac{1}{2c_0} - \frac{c_f}{2c_0} \tag{2.16}$$

Rearranging the terms in equation (2.15), we get

$$\begin{aligned} & exp\Big[-\frac{2c_0 r_0 (x-L) + R_{12}^* c_f}{R_{12}^* c_f + c^*}\Big] \cdot exp\Big[\frac{r_0 c_f u + R_{12}^* c_f}{R_{12}^* c_f + c^*}\Big] \\ &= \frac{R_{12}^* c_f + c^*}{R_{12}^* c_f} \frac{r_0 c_f u + R_{12}^* c_f}{R_{12}^* c_f + c^*} \end{aligned} \tag{2.17}$$

If we let

$$\begin{aligned} y &= \frac{r_0 c_f u + R_{12}^* c_f}{R_{12}^* c_f + c^*} \\ A &= \frac{R_{12}^* c_f}{R_{12}^* c_f + c^*} \cdot exp\Big[-\frac{2c_0 r_0 (x-L) + R_{12}^* c_f}{R_{12}^* c_f + c^*}\Big] \end{aligned}$$

we can rewrite equation (2.17) as

$$Ae^y = y, \text{ i.e., } (-y)e^{-y} = -A \tag{2.18}$$

From the definition of the *Lambert's W* function, $W(x)$ is defined by $We^W = x$. In terms of W function, equation (2.18) becomes $y = -W(-A)$. Expanding y in terms of u, we have

$$u(x) = -(\frac{R_{12}^*}{r_0} + \frac{c^*}{r_0 c_f})W(-A) - \frac{R_{12}^*}{r_0} \tag{2.19}$$

Substituting (2.19) into (2.16) and observing that $A = ae^{-bx}$, we get

$$f(x) = -\frac{c_f}{2c_0}(\frac{1}{W(-ae^{-bx})} + 1)$$

□

Remark 1 $f(x)$ depends on two parameters a and b. As it can be found from the definitions in (2.9), (2.10), (2.13) and (2.14), in order to get the values of a and b, we should know the values of the other two constants c and $u(0)$ first. In Lemma 5, we will derive the corresponding nonlinear equations for solving c and $u(0)$.

Remark 2 It is well known that $W(x)$ is monotonically increasing in the interval $[-e^{-1}, \infty]$. Because $a > 0$ and when $\alpha > \beta$, $b > 0$, $W(-ae^{-bx})$ increases as x increases. Therefore, $f(x)$ is monotonically increasing in the interval $[0, L]$. Similarly, when $\alpha < \beta$, $f(x)$ is a monotonically decreasing function in the interval $[0, L]$. Based on the above analysis and Lemma 2, a schematic solution of $f(x)$ can be sketched out as shown in Figure 2.3.

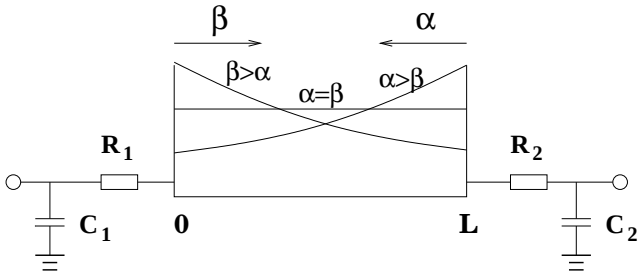


Figure 2.3: The optimal shape function for a bi-directional wire. When $\alpha = \beta = 0.5$, the optimal shape is a uniform width function. When $\alpha > \beta$, $f(x)$ is monotonically increasing in $[0, L]$. When $\alpha < \beta$, $f(x)$ is monotonically decreasing in $[0, L]$.

Without loss of generality, we assume $\alpha > \beta$ in the rest of the section. Before we derive Lemma 5, first we give the following relations which will be used in deriving Lemma 5.

Lemma 4 *Given the optimal wire shape function* $f(x) = -\frac{c_f}{2c_0}(\frac{1}{W(-ae^{-bx})} + 1)$, *we have*

$$\int_0^L \frac{dx}{f(x)} = \frac{2c_0}{bc_f}(w_1 - w_0) \tag{2.20}$$

$$\int_0^L \int_x^L \frac{dt}{f(t)} dx = \frac{2c_0}{bc_f} w_1 L + \frac{c_0}{b^2 c_f}(w_1^2 - w_0^2) + \frac{2c_0}{b^2 c_f}(w_1 - w_0) \quad (2.21)$$

where

$$w_0 = W(-a) \quad (2.22)$$

$$w_1 = W(-ae^{-bL}) = -\frac{R_{12}^* c_f}{R_{12}^* c_f + c^*} \quad (2.23)$$

Proof: Let $Y = W(-ae^{bx})$. From the definition of W function, we have

$$Ye^Y = -ae^{-bx}$$

Differentiating the above, we get

$$\frac{dY}{dx} = -\frac{bY}{1+Y}$$

By changing the variable from x to Y, the left hand side of equation (2.20) becomes

$$\begin{aligned} \int_x^L \frac{dx}{f(x)} &= \frac{2c_0}{bc_f} \int_{Y(x)}^{Y(L)} dY \\ &= \frac{2c_0}{bc_f}[W(-ae^{-bL}) - W(-ae^{-bx})] \end{aligned}$$

Therefore,

$$\int_0^L \frac{dx}{f(x)} = \frac{2c_0}{bc_f}(w_1 - w_0)$$

Now, we derive the second equation in the lemma statement.

$$\begin{aligned} \int_0^L \int_x^L \frac{dt}{f(t)} dx &= \frac{2c_0}{bc_f} \int_0^L [w_1 - W(-ae^{-bx})] dx \\ &= \frac{2c_0}{bc_f} w_1 L - \frac{2c_0}{bc_f} \int_0^L -\frac{Y+1}{b} dY \\ &= \frac{2c_0}{bc_f} w_1 L + \frac{c_0}{b^2 c_f}(w_1^2 - w_0^2) + \frac{2c_0}{b^2 c_f}(w_1 - w_0) \end{aligned}$$

□

To solve the two unknown parameters c and $u(0)$, we have the following lemma.

Lemma 5 *The two constants $u(0)$ and c are the roots of the following two nonlinear equations*

$$\frac{1}{\eta}(\frac{2C_{21}}{c_f}+\beta L+\frac{\eta w_0}{b}+\frac{\eta}{b})\Big\{[\frac{ab}{2\eta}(2\eta L+c')-\frac{c^*a\beta}{c_0R^*_{12}}\frac{du(0)}{dz}]\frac{dw_0}{da}-w_0\Big\}$$
$$-(\frac{2C_{21}}{c_f}+\alpha L+\frac{\eta w_1}{b}+\frac{\eta}{b})\frac{\beta c_f}{c_0\eta}\frac{du(0)}{dz}+L+\frac{1}{b}(w_1-w_0)=0 \qquad (2.24)$$

$$u(0)=\frac{R^*_{12}}{r_0\eta}\Big[exp(\frac{\eta r_0}{R_{12}c_f+c}(c_fu(0)+2c_0L))-1\Big] \qquad (2.25)$$

where

$$c' = c^*/(r_0c_0) \qquad (2.26)$$
$$R'_c = R^*_{12}c_f/(r_0c_0) \qquad (2.27)$$
$$C_{21} = \alpha C_1+\beta C_2 \qquad (2.28)$$
$$\frac{dw_0}{da} = \frac{w_0}{a(1+w_0)} \qquad (2.29)$$
$$\frac{du(0)}{dz} = -\frac{R_{12}+r_0\alpha u(0)}{R'_c+c'}\times\frac{R'_cu(0)+2LR^*_{12}/r_0}{R_{12}c'-R'_cr_0\alpha u(0)} \qquad (2.30)$$

Proof: Equation (2.25) is derived directly from (2.15) by setting $x=0$, where we make use of the relation $R^*_{12}c_f+c^*=R_{12}c_f+c$. To determine the constant c, we substitute $f(x)$ back into the delay expression in (2.5). With the help of the Euler's equation (2.8), T_{BD} in (2.5) can be written as

$$T_{BD} = \alpha R_2(C_1+c_fL)+\beta R_1(C_2+c_fL)+\frac{1}{2}cL$$
$$+ \int_0^L\frac{1}{2}r_0c_f\Big(\beta\int_0^L\frac{dt}{f(t)}+\alpha\int_x^L\frac{dt}{f(t)}\Big)dx+(\alpha C_1+\beta C_2)\int_0^L\frac{r_0}{f(x)}dx$$

Let $T'_{BD}=T_{BD}/(r_0c_0)$, $z=c/(r_0c_0)$, and make use of Lemma 4 and (2.26-2.28), then the delay can be simplified as

$$T'_{BD} = R_{12}\frac{(C_1+c_fL)}{r_0c_0}+\frac{\beta R_1(C_2-C_1)}{r_0c_0}+\frac{1}{2}Lz+\frac{2C_{21}}{bc_f}(w_1-w_0)$$
$$+ \frac{L}{b}(\alpha w_1-\beta w_0)+\frac{\eta}{2b^2}(w_1^2-w_0^2)+\frac{\eta}{b^2}(w_1-w_0) \qquad (2.31)$$

Equation (2.24) comes from $\frac{dT'_{BD}}{dz} = 0$. The derivative $\frac{dT'_{BD}}{dz}$ can be calculated as

$$\begin{aligned}\frac{dT'_{BD}}{dz} &= \frac{\partial T'_{BD}}{\partial z} + \frac{\partial T'_{BD}}{\partial (w_1/b)}\frac{d(w_1/b)}{dz} + \frac{\partial T'_{BD}}{\partial b}\frac{db}{dz} \\ &+ \frac{\partial T'_{BD}}{\partial (w_1/b^2)}\frac{d(w_1/b^2)}{dz} + \frac{\partial T'_{BD}}{\partial (w_0/b)}\frac{d(w_0/b)}{dz} + \frac{\partial T'_{BD}}{\partial (w_0/b^2)}\frac{d(w_0/b^2)}{dz}\end{aligned}$$

where

$$\begin{aligned}\frac{db}{dz} &= -\frac{1}{2}b^2 \\ \frac{d(w_1/b)}{dz} &= -\frac{\beta c_f}{2c_0\eta}\frac{du(0)}{dz} \\ \frac{d(w_1/b^2)}{dz} &= \frac{1}{b}\frac{d(w_1/b)}{dz} + \frac{w_1}{2\eta b} \\ \frac{d(w_0/b)}{dz} &= -\frac{1}{2\eta}\Big\{[\frac{ab}{2\eta}(2\eta L + c') - \frac{c^* a\beta}{c_0 R^*_{12}}\frac{du(0)}{dz}] \times \frac{dw_0}{da} - w_0\Big\} \\ \frac{d(w_0/b^2)}{dz} &= \frac{1}{b}\frac{d(w_0/b)}{dz} + \frac{w_0}{2\eta b}\end{aligned}$$

$\frac{du(0)}{dz}$ in equation (2.30) is directly derived from (2.25). Substituting all these into $\frac{dT'_{BD}}{dz} = 0$, we then get equation (2.24).

□

In practice, we employ bisection method to solve these two nonlinear equations in Lemma 5, because we found it more efficient than the Newton Raphson method in our experiments. Algorithms for calculating the W function can be found in [20]. The efficiency of bisection method depends on the initial root range, which can be estimated from the following lemma.

Lemma 6 *If c' is the root of nonlinear equation (2.24), it must satisfy*

$$c' \geq \frac{2\eta L - R'_c}{W^*} - R'_c \tag{2.32}$$

In terms of z, where $z = c/(r_0 c_0)$, this relation is

$$z \geq \frac{2\eta L - R_{12}c_f/(r_0 c_0)}{W^*} - \frac{\beta u(0) c_f}{W^* c_0} - \frac{R_{12}c_f}{r_0 c_0} \tag{2.33}$$

where $W^* = W(e^{-1}\frac{2\eta L - R_c'}{R_c'})$ *and* R_c' *is as defined in (2.27).*

Proof: This idea originates from the definition of the W function. For the function $W(x)$, the range of x is not from $-\infty$ to $+\infty$, but from $-e^{-1}$ to $+\infty$ instead. There are two different W function values in (2.24) need to be calculated, namely w_0 and w_1, i.e., $W(-a)$ and $W(-ae^{-bL})$. Therefore, the constraint on x in W function gives a limit on the values of a and b, which depend on the root c' that we are finding. Obviously, $-a < -ae^{-bL}$, thus we need only consider the constraint on w_0, i.e., we need to know what kind of constraint on c' can satisfy $-a \geq -e^{-1}$. In terms of c', it is

$$\frac{R_c'}{R_c' + c'} \cdot exp(\frac{2\eta L - R_c'}{R_c' + c'}) \leq e^{-1}$$

If we let $y = \frac{2\eta L - R_c'}{R_c' + c'}$, then the above inequality is simplified as

$$ye^y \leq e^{-1}\frac{2\eta L - R_c'}{R_c'}$$

i.e.,

$$y \leq W(e^{-1}\frac{2\eta L - R_c'}{R_c'}) \tag{2.34}$$

There is no problem in calculating the W function in (2.34), since L is always positive and thus the term inside this W function is always larger than $-e^{-1}$. Substituting c' back into (2.34), then we find the constraint on root c',

$$c' \geq \frac{2\eta L - R_c'}{W^*} - R_c' \tag{2.35}$$

To get (2.33), we substitute (2.26) and (2.9) into (2.35). In the above, we have implicitly assumed that $2\eta L - R_c' > 0$. In fact, the same inequality can be proved if $2\eta L - R_c' < 0$.

□

So far, we have determined the optimal shape function for a bi-directional wire. Uni-directional wire optimization, which has been studied in [9, 24], can be

thought of as a special case of bi-directional wire optimization with either $\alpha = 0$ or $\beta = 0$. The authors in [9] give a closed form optimal wire shape function, and the author in [24] gives a power series solution. Based on Lemmas 3 and 5, it is easy to get the following results for a uni-directional wire by setting $\alpha = 1$, $\beta = 0$.

Lemma 7 *Given a uni-directional wire with driver resistance R_d and load capacitance C_l, the optimal shape function is*

$$f(x) = -\frac{c_f}{2c_0}\left(\frac{1}{W(-ae^{-bx})} + 1\right) \tag{2.36}$$

where

$$\begin{aligned} a &= \frac{R_d c_f}{R_d c_f + c} \cdot exp\left(\frac{2r_0 c_0 L - R_d c_f}{R_d c_f + c}\right) \\ b &= \frac{2r_0 c_0}{R_d c_f + c} \end{aligned}$$

The constant c is the root of the following nonlinear equation

$$-\left(\frac{2C_l}{c_f} + \frac{w_0}{b} + \frac{1}{b}\right)\left(w_0 - \frac{1}{2}ab\left(2L + \frac{c}{r_0 c_0}\right)\frac{dw_0}{da}\right) + L + \frac{1}{b}(w_1 - w_0) = 0 \tag{2.37}$$

where

$$w_0 = W(-a) \tag{2.38}$$

$$w_1 = W(-ae^{-bL}) = -\frac{R_d c_f}{R_d c_f + c} \tag{2.39}$$

Remark 3 The wire shape function in (2.36) is equivalent to but looks different from the one in [9] (where $f(x) = \frac{-c_f}{2c_0}\left(\frac{1}{W\left(\frac{-c_f}{ae^{-bx}}\right)} + 1\right)$, and a and b are coefficients). This is because, the derivation in [9] was based on the assumption that the driver is on the left end of the wire, whereas here we assume it is on the right end. However, after we substitute x with $x' = L - x$ into equation (2.36) and rearrange the terms, (2.36) becomes $f(x') = \frac{-c_f}{2c_0}\left(\frac{1}{W\left(\frac{-c_f}{a'e^{-b'x'}}\right)} + 1\right)$, where a' and b' are coefficients. $f(x')$ is now consistent with $f(x)$ in [9]. Furthermore, we find that the two coefficients in [9] rely on solving two nonlinear equations, but ours rely on solving only one nonlinear equation.

Remark 4 Comparing with the optimal shape function of bi-directional wire in Lemma 3, the optimal uni-directional wire shape function has the same form of solution. But the uni-directional wire shape function only depends on one constant c, whereas the bi-directional wire shape function depends on two constants a and b.

Remark 5 Fishburn in [24] also used calculus of variations to determine the optimal function for a uni-directional wire. But his result is based on the delay expression in (2.3). The Euler's differential equation in [24] is

$$c_0 u(x) + \frac{1}{2} c_f x + C_l = c u'(x) \tag{2.40}$$

where $u(x)$ is defined as $u(x) = \int_0^x f(t)dt$, which is different from our definition. However, it is difficult to obtain a closed form solution to equation (2.40), and hence Fishburn had to rely on series expansion instead. However, if we start from the delay expression (2.4), and define $u(x)$ in the same way as what we did in Lemma 2, the Euler's equation will be

$$-2c_0(R_d + r_0 u(x)) + r_0 c_f u(x) u'(x) = c u'(x) \tag{2.41}$$

which is equivalent to (2.11). This can be solved analytically, and the solution has been shown in (2.36).

2.1.4 Extension to Minimizing the Maximum Delay

In previous subsections, we have determined the optimal shape which minimizes the total weighted delay $T_{BD} = \alpha T_{DR} + \beta T_{DL}$, where T_{DR} and T_{DL} are delays from two opposite directions. In this subsection, we show that our study can be extended to dealing with another optimization objective in which the maximum delay of T_{DR} and T_{DL} is minimized. The problem (which is sometimes called the primal problem) can be stated as follows:

$$\text{Minimize} \qquad \text{Max}\{T_{DR}, T_{DL}\}$$

By introducing a variable T, the problem is also equivalent to:

$$\begin{array}{ll} \text{Minimize} & T \\ \text{Subject to} & T_{DR} \leq T \\ & T_{DL} \leq T \end{array}$$

Following the Lagrangian relaxation technique [5], we introduce two non-negative variables α and β which are called Lagrangian multipliers for each constraint. The Lagrangian function associated with the primal problem is:

$$L = T + \alpha(T_{DR} - T) + \beta(T_{DL} - T) \tag{2.42}$$

$$= (1 - \alpha - \beta)T + \alpha T_{DR} + \beta T_{DL} \tag{2.43}$$

By Kuhn-Tucker conditions [5], let $f(x)$ be the optimal shape function for the primal problem. Then there exists α and β such that

$$\begin{cases} \nabla_T = \alpha + \beta - 1 = 0 \\ \nabla_{f(x)} = \nabla(\alpha T_{DR} + \beta T_{DL}) = 0 \\ \alpha \geq 0, \ \beta \geq 0 \\ \alpha(T_{DR} - T) = 0 \\ \beta(T_{DL} - T) = 0 \end{cases} \tag{2.44}$$

Since T_{DR} and T_{DL} are functionals, $\nabla_{f(x)}(\alpha T_{DR} + \beta T_{DL}) = 0$ implies the Euler's differential equation (2.7), i.e., $F - u'F_{u'} = c$. Therefore, $f(x)$ in minimizing the maximum delay still has the form:

$$f(x) = \begin{cases} -\frac{c_f}{2c_0}\left(\frac{1}{W(-ae^{-bx})} + 1\right) & \text{for } \alpha \neq \beta \\ constant & \text{for } \alpha = \beta \end{cases} \tag{2.45}$$

According to Lemma 7, $f(x)$ depends on a constant c, which can be determined from equations (2.44). Also note that in solving equation (2.44), not all two constraints will be active at the same time (an inequality constraint like $T_{DR} \leq T$ is said to be active if $T_{DR} = T$ and inactive if $T_{DR} < T$). In fact, there may be three different cases.

1. $\alpha \neq 0$ and $\beta \neq 0$: all constraints are active, i.e., $T_{DR} = T$ and $T_{DL} = T$ which yield $T_{DR} = T_{DL}$. According to Lemma 7, $T_{DR} = T_{DL}$ is in fact a nonlinear equation in c. c can thus be obtained by solving such nonlinear equation. The following equations are useful in solving equation $T_{DR} = T_{DL}$.

$$\begin{aligned} T_{DR} &= C_1 R_2 + c_f R_2 L + (\frac{c_f R_2}{2} + \frac{c_0 r_0 w_1}{b})\frac{M}{b} \\ &+ \frac{r_0 c_0}{b^2}(N + 2bw_1 L) + \frac{r_0 c_0 L}{b} \qquad (2.46) \\ T_{DL} &= C_2 R_1 + c_f R_1 L + (\frac{c_f R_1}{2} - \frac{c_0 r_0 w_1}{b})\frac{M}{b} \\ &- \frac{r_0 c_0}{b^2}(N + 2bw_0 L) - \frac{r_0 c_0 L}{b} \qquad (2.47) \end{aligned}$$

   where

$$\begin{aligned} M &= w_0 - w_1 - 2bL - \frac{1}{w_0} + \frac{1}{w_0} \qquad (2.48) \\ N &= \frac{1}{2}w_1^2 - \frac{1}{2}w_0^2 + w_1 - w_0 \qquad (2.49) \end{aligned}$$

   To obtain the Lagrangian function L for this case, α and β have to be solved through equations $\alpha \nabla T_{DR} + \beta \nabla T_{DL} = 0$ and $\alpha + \beta = 1$.

2. $\alpha = 0$ and $\beta \neq 0$ or $\alpha \neq 0$ and $\beta = 0$: in either case, one of the constraints is active. The optimization problem is reduced to uni-directional optimization.

Let's denote the Lagrangian functions in each of these three cases as L_1, L_2 and L_3. The final solution on $f(x)$ is the function where its Lagrangian function is the maximum among L_1, L_2 and L_3.

2.1.5 Experimental Results

In this subsection, we will show some experimental results. The parameters in our experiment are chosen as follows: $L = 30,000\mu m$, $r_0 = 0.03\Omega/\Box$, $c_0 = 0.2fF/\mu m^2$, $c_f = 20fF/\mu m$, $R_1 = 100\Omega$, $C_2 = 200pF$, $R_2 = 10\Omega$, and $C_1 = 20pF$. We choose

$R_1 \neq R_2$, $C_1 \neq C_2$ and $R_1 \times C_2 \neq R_2 \times C_1$ such that the circuit is asymmetric. We want to show by our experiments that in this case, the optimal shape is uniform if $\alpha = \beta$. If the circuit is symmetric, one intuition may guess that the wire shape is uniform just because of symmetry of the circuit. Further intuition may say that if the circuit is asymmetric, the optimal shape would look like a bowl shape, which is the average of two optimal shapes solved from each direction. We have already shown in Lemma 2 and Remark 2 that such intuition is not correct. For our experiments, we choose several values for α ranging from 0.5 to 1.0. The calculated minimum weighted delays T_{BD} are summarized in Table 2.1.

α	β	Weighted delay (ns)
0.9	0.1	27.27
0.8	0.2	37.94
0.7	0.3	48.20
0.6	0.4	58.15

Table 2.1: Weighted delays calculated for bi-directional wires.

Because the circuit is asymmetric and obviously the left-to-right delay dominates over the right-to-left delay, the minimum weighted delay is very sensitive to the weights. The minimum weighted delay for $\alpha = 0.6$ is twice bigger than delay for $\alpha = 0.9$. The increased delay is primarily contributed by the left-to-right delay.

The calculated optimal wire shape functions are shown in Figure 2.4. This figures shows clearly that when $\alpha = \beta$, the optimal shape is in fact uniform. When $\alpha = 1$ and $\beta = 0$, bi-directional wire reduces to uni-directional wire. We also observe that as α increases from 0.5 to 1.0, the driver-end width increases.