Self-Heating-Aware Optimal Wire Sizing under Elmore Delay Model

Min Ni and Seda Ogrenci Memik
Electrical Engineering and Computer Science
Northwestern University, Evanston, IL
{mni166,seda}@ece.northwestern.edu

Abstract

Global interconnect temperature keeps rising in the current and future technologies due to self-heating and the adiabatic property of top metal layers. The thermal effects impact adversely both reliability and performance of the interconnect wire, shortening the interconnect lifetime and increasing the interconnect delay. Such effects must be considered during the process of interconnect design. In this paper, one important argument is that the traditional linear dependence between wire resistance and wire width is no longer adequate for high layer interconnects due to the adiabatic property of these wires. By using curve fitting technique, we propose a quadratic model to represent the resistance of interconnect, which is aware of the thermal effects. Based on this model and the Elmore delay model, we derived a linear optimal wire sizing formula in form of $f(x) = ax + b$. Compared to non-thermal-aware exponential wire sizing formula in form of $f(x) = ae^{-bx}$, we observed a 49.7% average delay gain with different choices of physical parameters.

1 Introduction

Interconnect wire consumes I^2R power, which translates into heat. This is known as interconnect self-heating. This effect becomes significant in current technologies, and will become even more significant in future for high-layer high-switching global interconnections, such as clock networks. The reason is: first, the resistance of interconnect increases due to the reduction of wire size; second, the current in interconnect increases due to the increasing of switching frequency. The resulting heats are difficult to dissipate because unlike the substrate, which attaches directly to the heat sink, interconnects, especially the high-layer interconnects are more isolated. Furthermore, introducing low dielectric constant (low-k) materials to reduce cross-talk exacerbates the thermal condition. This is due to the fact that low-k materials usually have lower thermal conductivity than silicon dioxide [3]. It has been predicted that the maximum interconnect temperature in a global line may reach up to 210^{o}C for 50-nm technology node [19].

The analysis of interconnect self-heating [5, 4] is not only important in the context of electro-migration induced reliability concerns [7], but also due to its impact on performance, *i.e.*, delay. Ajami et al. presented a detailed analysis and modeling of interconnect performance degradation due to non-uniform temperature profiles [2]. In this work, a maximal 35% increase of delay due to thermal effects was reported. The main reason behind this is that the resistance of an interconnect increases linearly with the temperature.

The thermal effects on interconnect design has gained a lot of attentions in the past few years. Chen proposed an interconnect thermal model considering thermal coupling between different interconnects wires [12]. Kapur also considered thermal effects as one of major concerns when discussing the technology and reliability constrained future copper interconnects [20]. Recently, Casu et al. optimized the global interconnect under thermal effects by using RLC model [8].

The above papers focus mainly on how to model the thermal effects on the interconnection design. However, the discussion about how to adjust the design techniques of interconnect, such as wire sizing [15, 16], buffer insertion [23] and sizing, and combination of these [9, 13, 14], is still lacking in the literature. The only related work we are aware of is the optimal buffer insertion under substrate thermal gradient effects [1] proposed by Ajami. The thermal conditions in their analysis are based on the temperature profile of the substrate, which could involve inaccuracy because of the adiabatic property of high-layer global interconnects (while the generated heat of interconnects is hard to dissipate, the heat of the substrate is also difficult to transfer to high interconnect layers. The only thermal flow is through via, which affects only the temperature of the ends of the wires) , as well as the significant self-heating effects of these high switching, high resistance global interconnects.

In this paper, we consider the thermal effects, specifically, self-heating effects, on one of the most important interconnect design techniques, namely, wire sizing technique. Our self-heating-aware technique can be combined with the thermal-aware interconnect design methodology considering only substrate temperature conditions as shown in [1] to obtain more accurate optimal results.

It was first shown in [15] and [16] that when wire resistance becomes significant , as in DSM designs, proper wire sizing can effectively reduce the interconnect delay. The basic idea behind wire-sizing is that increasing the interconnect wire width has two different effects: it decreases the wire resistance while increasing the wire capacitance. Since the Elmore delay can be formulated by RC constant time, there must exist an optimal wire width that achieves least RC constant time. On the other hand, if we divide the wire into several segments, each segment would have different optimal wire width with different position along the wire.

Finding the optimal wire width for each segment is the problem of discrete wire sizing [15, 16]. Another formulation of wire-sizing optimization is to determine the continuous wire shaping functions. The closed-form wire shaping functions were derived to minimize the Elmore delay, first without considering fringing capacitance [11] and later with fringing capacitance [10] and for a bidirectional wire [18]. Menezes et al. also presented the wire-sizing optimization under a higher-order RC delay model [21].

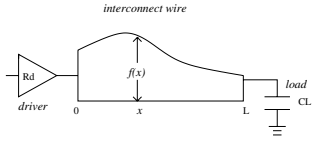

Figure 1: Wire sizing function

In this paper, we also consider the continuous wire-sizing. We are trying to find the width f of an interconnect as a function of its position x, also known as the shape formula [11]. The sizing function is illustrated in Figure 1.

When considering self-heating effect, one important observation is that narrowing down the wire not only increases wire resistance by the amount of $R = r_0/w_m$, but also exacerbates the thermal condition of the wire. Because the heat generation of the wire, which is proportional to I^2R, will increase; and the heat dissipation ability of that wire will decrease due to the shrinking of its surface area. This implies a temperature rise on the wire, which can increase the wire resistance further. Therefore, in place of the former linear relation between wire resistance and wire width $R(w_m) = r_0/w_m$ we

978-3-9810801-2-4/DATE07 © 2007 EDAA

propose a more accurate quadratic relation $R(w_m) = \alpha/{w_m}^2$ to capture thermal effects. It is necessary to note that such relation between wire width(w_m) and wire resistance(R) holds only for resistors in adiabatic environment such as high-layer global interconnect wires, in which the thermal positive feedback effects become significant.

Besides using the quadratic relation between resistance and wire width, we also make some reasonable simplification and assumption when deriving the optimal wire-sizing formula.

- We only consider single end global wires, which has been adopted for most of the continuous wire sizing problems [11, 10, 18].
- Elmore delay is adopted as an accurate metric for our delay calculation because of its simplicity. More accurate models considering higher moments can be used, but no analytical formulation can be obtained.
- We do not consider fringing capacitance. A simple uniform distribution of capacitance is assumed, which has been adopted in most of the wire sizing literature [11, 18, 15, 16, 14, 9].
- We assume a uniform substrate temperature.

The remainder of this paper is organized as follows. Section 2 describes in detail our self-heating model. The Elmore delay model is described in Section 3. Section 4 illustrates the derivation of the optimal non-uniform self-heating-aware wire sizing function. After that, we discuss in Section 5 the self-heating effects on discrete wire sizing. Section 6 presents experimental results based on our model, followed by Section 7, which gives our conclusion.

2 Self-Heating Model

Thermal effects are an inseparable aspect of electrical power distribution and signal transmission due to *self-heating* (also known as *joule heating*) caused by the flow of current [4]. Since the interconnects, especially the global interconnects, are far away from the substrate, which is attached to the heat sink, the heat generated by I^2R power dissipation cannot be efficiently removed and therefore causes an increase in interconnect temperature.

The self-heating temperature of the interconnects can be expressed as a function of wire width [5],

$$\Delta T_{self-heating}(w_m) = T_m - T_{ref} = {I^2}_{rms} \cdot R(w_m) \cdot R_\theta(w_m) \quad (1)$$

where R is the interconnect resistance, I_{rms} is the root mean square current, ${I^2}_{rms}R$ represents the power dissipation on the interconnects, and R_θ is the thermal impedance of the interconnect line to the chip, which has the form of [5],

$$R_\theta(w_m) = \frac{t_{ox}}{K_{ox} \cdot L \cdot W_{eff}} \quad (2)$$

This equation of thermal impedance is based on a quasi-1-D heat conduction model, where t_{ox} is the total thickness of the underlying dielectric, as illustrated in Figure 2, K_{ox} is the thermal conductivity normal to the plane of the dielectric, L is the length of the interconnect, and $W_{eff} = w_m + 0.88t_{ox}$, valid for $w_m/t_{ox} > 0.4$ and is accurate to within 3%.

In Equation (1), we only consider the heat conduction between metal wire and the surrounding dielectric materials (y direction and z direction). Although metal has much higher thermal conductivity than silicon dioxide, the heat transfer rate along the wire (x direction) is much smaller compared to those in y and z directions mainly due to the difference of the power dissipation area of these three directions.

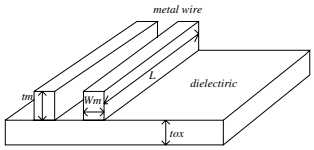

Figure 2: Interconnect wire parameters

The heat transfer rate in x direction is defined by the *Fourier law of thermal conduction* [6],

$$q_x = -k_x A_x \frac{\partial T}{\partial x} \quad (3)$$

Where q_x is the heat transfer rate in x direction, k_x is the thermal conductivity, A_x is the intersection area in x direction. Therefore,

$$\frac{q_x}{q_{yz}} = \frac{k_x}{k_{yz}} \frac{t_m w_m}{L(t_m + w_m)} \frac{\frac{\partial T}{\partial x}}{\frac{\partial T}{\partial y}} \quad (4)$$

The thermal conductivity of aluminum and silicon dioxide at 100^{o}C are 206W/m·K and 1.15W/m·K,respectively. Assume $t_m \cong w_m$, and $\frac{\partial T}{\partial x} \cong \frac{\partial T}{\partial y}$ for a global wire with length of 3000μm, we have $\frac{q_x}{q_{yz}} = 0.06 \ll 1$. Therefore, self-heating temperature Equation (1) is adequate for long interconnect wire by considering only the heat transfer between wire and dielectric materials.

For the purpose of deriving the optimal wire sizing formula, we need to express the wire resistance as a function of wire width, *i.e.*, $R = f(w_m)$. Basically, the unit length resistance of interconnect wire is in a linear relation with the wire width,

$$R = \frac{\rho_0}{w_m \cdot t_m}(1 + \beta T_m) \quad (5)$$

where ρ_0 is the unit length resistance at the reference temperature, and β is the temperature dependence coefficient of resistance($1/^oC$).

Note that in Equation (5), T_m is also a function of wire width w_m as expressed in Equation (1). Substituting Equation (2) and (5) into Equation (1), we obtain the expression of T_m as a function of wire width w_m,

$$T_m - T_{ref} = \frac{{I^2}_{rms} \cdot t_{ox} \cdot \rho_0 \cdot (1 + \beta(T_m - T_{ref}))}{K_{ox} \cdot t_m \cdot w_m(w_m + 0.88t_{ox})} \quad (6)$$

Without loss of generality, we assume the reference temperature T_{ref} to be 0. After rearranging Equation (6), we have,

$$T_m = \frac{\eta/\beta}{({w_m}^2 + 0.88t_{ox}w_m) \cdot \gamma - \eta} \quad (7)$$

where $\eta = {I^2}_{rms} \cdot t_{ox} \cdot \rho_0 \cdot \beta$ and $\gamma = t_m \cdot K_{ox}$.

Substitute Equation (7) into (5), we obtain the unit length wire resistance as a function of wire width w_m, as shown in Equation (8).

$$R = \frac{\rho_0}{t_m} \cdot \frac{w_m + a_0}{{w_m}^2 + b_1 w_m + b_0} \quad (8)$$

where $a_0 = b_1 = 0.88 \cdot t_{ox}$ and $b_0 = -\frac{\eta}{\gamma}$.

Equation (8) has a standard form of a system function. Note that this system function has one positive pole and one negative pole as shown in Equation (9) , because $b_0 < 0$.

$$R = \frac{\rho_0}{t_m} \cdot \left(\frac{k_1}{w_m - p_1} - \frac{k_2}{w_m - p_2}\right) \quad (9)$$

where,

$$\begin{cases} p_1 = \frac{-b_1 - \sqrt{{b_1}^2 - 4b_0}}{2}, & k_1 = \frac{p_1 + a_0}{p_1 - p_2} \\ p_2 = \frac{-b_1 + \sqrt{{b_1}^2 - 4b_0}}{2}, & k_2 = \frac{p_2 + a_0}{p_2 - p_1} \end{cases}$$

The positive pole will be a singularity as shown in Figure 3. When the wire width tends towards the positive pole, the wire resistance will become infinite. This can be explained by the positive feedback relation between wire resistance (Equation(5)) and self-heating temperature (Equation(1)). As the wire width becomes smaller than the pole value, the resistance jumps to negative values, which indicates the occurrence of thermal-runaway. So, the pole value is the minimum valid wire width under a certain current value(I_{rms}) and thermal condition(R_θ). Figure 3 also illustrates several resistance curves under different current conditions, *i.e.*, $I_{rms} = 50mA$, $I_{rms} = 90mA$, and $I_{rms} = 140mA$. It is shown that a larger current value always implies a bigger positive pole, which means the minimum valid wire width under self-heating consideration becomes larger.

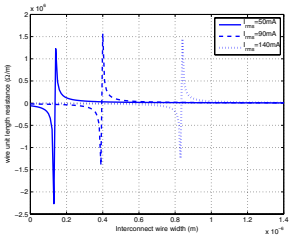

Figure 3: Wire resistance as a function of wire width and current value. The physical parameters are taken from [5], also listed in Table 1 for reference.

Table 1: Physical parameter values used in the interconnect self-heating model.

Parameter	Value	
t_{ox}	3	μm
t_m	2	μm
ρ_0	1.67×10^{-6}	$\Omega \cdot cm$
β	6.8×10^{-3}	$^oC^{-1}$
K_{ox}	1.15	$W/m \cdot K$

Usually, wire sizing is performed in a range of wire width. Besides the current value in wire, this range is also decided by the technology, *e.g.*, the aspect ratio of the interconnects, the dielectric material and the thickness of the dielectric layer. In our example, we use the temperature as an indicator of this range. If we require that the temperature of the interconnects cannot exceed a certain value, for instance, 150^oC, we can find the minimum valid wire width to be about $1.4\mu m$ in Figure 4. Figure 4 is part of the curve shown in Figure 3 with I_{rms} equals 140mA. The maximum width is extended to $3.2\mu m$. In this figure, the y-axis represents the temperature, the resistance counterpart is shown in Figure 5. Here we also depict two additional curves, one is in the form of $R = \alpha_1/w_m$, which we will call *linear relation* in the remainder of our discussion. The other is $R = \alpha_2/{w_m}^2$, which we will call *quadratic relation*. In the conventional wire sizing function, the linear relation was adopted, as shown in Equation (5). However, this can underestimate the wire resistance and temperature significantly as illustrated in Figure 5. Next, we will show that the square relation can be a good approximation, at least in the valid wire width range.

The most accurate way to represent the resistance of a wire as a function of wire width is to substitute Equation (9) into the delay model. However, this makes it difficult to solve the formula of optimal wire sizing analytically. On the other hand, in a practical scenario only a relatively small range of

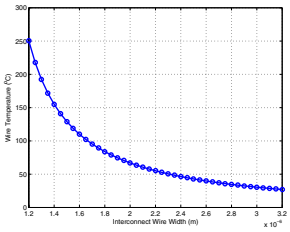

Figure 4: Self-heating temperature as a function of wire width ($I_{rms} = 140mA$).

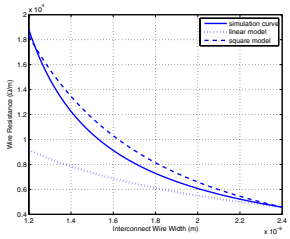

Figure 5: Wire resistance as a function of wire width ($I_{rms} = 140mA$) in the temperature range from 50^oC to 150^oC.

the wire width is of importance as opposed to the whole range from 0 to $+\infty$. Therefore, we can actually partition the resistance curve into multiple regions and use a simpler formula to fit the curve within the region of interest. Note that if we admit that Equation (9) is the most accurate relation between wire resistance and wire width when considering thermal effects, the resistance model $R = \alpha_1/w_m$ used in conventional wire sizing [11] is also a result of curve fitting. Moreover, the fitting is only accurate in the low temperature range. Therefore, we need a more general function, that is,

$$R = \frac{\alpha}{w_m^\gamma} \tag{10}$$

where R is the resistance value of unit length interconnect at reference temperature, and γ can be any value no less than 1. Depending on the temperature range in which the interconnect wire will most likely operate (the higher working temperature is, the bigger γ value will be), we can always find the best γ value by curve fitting method.

One way to decide γ is that we first estimate the working temperature range or wire width range of the wire by using the I_{rms} value derived from the power budget. In our work, we assume that the working temperature range is given. If the wire width range is given instead, the analysis is in the similar way. Then we can fit the general form exactly in some wire width range (w_{min}, w_{max}) corresponding to that temperature range. For example, in Figure 4, if given the temperature range is from 50^oC to 150^oC, then the corresponding wire width range is $(1.4, 2.2)$. One method to find the best fitting γ is the *method of least squares*,

$$D = \sum_{i=1}^{n} (R(w_i) - \frac{\alpha}{w_i^\gamma})^2 \tag{11}$$

where $w_{min} \le w_i \le w_{max}$. n is the number of sample points on the curve. The larger n is, the more accurate γ will be. $R_0(w_i)$ is the unit length resistance in Equation (8).

In order to obtain the value of coefficient α, we set the accurate resistance value in Equation (8) and the resistance value in simpler γ model to be equal at some wire width, for example, w_{max}. Therefore,

$$\alpha = w_{max}^\gamma \cdot R(w_{max}) \tag{12}$$

Substitute this equation into Equation (11), we obtain,

$$D = \sum_{i=1}^{n} (R_0(w_i) - (\frac{w_{max}}{w_i})^{\gamma} \cdot R_0(w_{max}))^2 \qquad (13)$$

Differentiating both sides of this equation, we have,

$$\frac{dD}{d\gamma} = \sum_{i=1}^{n} ((R_0(w_i) - (\frac{w_{max}}{w_i})^{\gamma} R_0(w_{max})) \cdot ln\frac{w_{max}}{w_i} \cdot (\frac{w_{max}}{w_i})^{\gamma}) \qquad (14)$$

In order to determine the optimal γ value, we set $\frac{dD}{d\gamma}$ to zero. We solved this nonlinear equation by secant method [22]. The results are the curves of γ value as a function of interconnect current I_{rms} and other physical properties such as t_{ox}, as illustrated in Figure 6.

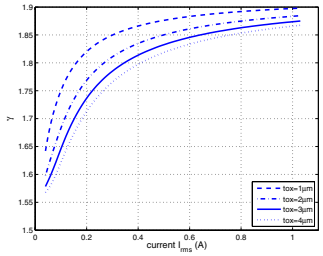

Figure 6: Optimal fitting for γ as a function of current I_{rms} and other physical parameters.

In this figure, the x-axis represents the current value and the y-axis is the least square fitting for γ. As the interconnect current value increases, the γ value also increases. However, γ will not exceed a bound near 2. We also show the γ curve under different t_{ox} values. It can be observed that γ is not sensitive to t_{ox}. We also changed other physical parameters, the result is similar to the case of t_{ox}.

This set of results are important, because they validate that our quadratic relation (γ=2) can be a good upper bound for wire sizing. That is one of the most important reasons why we are interested in solving the optimal wire sizing formula under quadratic resistance model. At the same time, $\gamma = 1$ yields the lower bound formula. The second reason to choose quadratic relation for analysis is that it is possible to obtain an analytical form of optimal wire sizing function when using quadratic model. In other words, we can have both the upper bound and the lower bound as explicit analytical formulae, making it much easier to see the trends of wire sizing (from exponential sizing to linear sizing as we will see) when thermal issues become more and more severe. Such trends will also eventually serve as good guidelines for the more practical discrete wire sizing. We will elaborate on this further in Section 5.

3 Elmore Delay Model

We use the Elmore delay model [17] to approximate the interconnect delay. Suppose the wire is partitioned into n equal-length wire segments, each has a length of $\Delta x = \frac{L}{n}$. Let x_i be $i\Delta x$, $1 \le i \le n$. The capacitance and resistance of wire segment i can be approximated by $c_0 f(x_i)$ and $\frac{r_0 \Delta x}{f(x_i)^2}$ respectively, where $f(x_i)$ is the wire width in position x_i. The Elmore delay through the entire wire can be expressed by,

$$D_n = R_d(C_L + \sum_{i=1}^{n}(c_0 f(x_i)\Delta x)) + \sum_{i=1}^{n} \frac{r_0 \Delta x}{f(x_i)^2}(\sum_{j=i}^{n}(c_0 f(x_j)\Delta x + C_L))$$

where R_d is the drive resistance, C_L is the load capacitor, as illustrated in Figure 1, c_0 represents the unit width capacitor, and r_0 is the α coefficient in Equation (10). Here we use r_0 to maintain consistency with the existing wire sizing terminology. The first term is the delay of the driver, which is given by the driver resistance R_d multiplied by the total capacitance of the wire and C_L. The second term is the sum of the delay in each wire segment i. The square relation is used here to take the self-heating effects into account. As $n \to \infty$, $D_n \to D$ as shown in Equation (15)

$$D = R_d(C_L + \int_0^L c_0 f(x)dx) + \int_0^L \frac{r_0}{f(x)^2}(\int_x^L c_0 f(x)dt + C_L)dx \qquad (15)$$

4 Optimal Self-Heating-Aware Wire Sizing

In this section, we will show that the optimal self-heating-aware wire shape function $f(x)$ which minimizes the Elmore delay can be solved by an analytical method.

Lemma 1 *Let f be an optimal wire-sizing function. We have,*

$$f^3(x) = \frac{2r_0(C_L + c_0\int_x^L f(t)dt}{c_0(R_d + r_0\int_0^x \frac{1}{f^2(t)}dt)} \qquad (16)$$

Proof: Similar proof can be found in literature [11] to derive a exponential wire sizing function.

Let $x \in [0, L]$. Assume f is continuous at x. We consider $\hat{f}$, which is a local modification of f in a small region $[x - \frac{\delta}{2}, x + \frac{\delta}{2}]$. The function $\hat{f}$ is defined as follows,

$$\hat{f} = \begin{cases} y & x - \frac{\delta}{2} \le t \le x + \frac{\delta}{2} \\ f(t) & otherwise \end{cases}$$

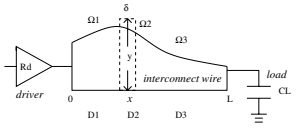

Figure 7: Local modification of an optimal wire-sizing function.

The wire W would be divided into three regions Ω_1, Ω_2, and Ω_3, as illustrated in Figure 7. We denote the signal delay through Ω_i by D_i. Hence, the total interconnect delay is equal to $\sum_{i=1}^{3} D_i$. We represent the wire resistance (capacitance) of Ω_i by $R_i(C_i)$. We have $R_2 = \frac{r_0\delta}{y^2}$ and $C_2 = c_0\delta y$. The driver delay is given by $R_d(C_1 + C_2 + C_3 + C_L)$. Hence, the signal delay through the wire and driver can be calculated as follows,

$$\begin{aligned} &D = R_d(C_1 + C_3 + C_L + c_0\delta y) \\ &+ \int_0^{x-\frac{\delta}{2}} \frac{r_0}{f^2(t)}(\int_t^{x-\frac{\delta}{2}} c_0 f(s)ds + c_0\delta y + C_3 + C_L)dt \\ &+ \frac{r_0\delta}{y^2}(c_0\delta y + C_3 + C_L) + \int_{x+\frac{\delta}{2}}^{L} \frac{r_0}{f^2(t)}(\int_t^L c_0 f(s)ds + C_L)dt \end{aligned}$$

By setting $\frac{dD}{dy} = 0$, we obtain,

$$R_d\delta c_0 + c_0\delta\int_0^{x-\frac{\delta}{2}} \frac{r_0}{f^2(t)}dt - \frac{2r_0\delta(C_3 + C_L)}{y^3} - \frac{r_0\delta^2}{y^2} = 0$$

As $\delta \to 0$, we have,

$$y^3{}_{min} = \frac{2r_0(C_L + c_0\int_x^L f(t)dt}{c_0(R_d + r_0\int_0^x \frac{1}{f^2(t)}dt)} \qquad (17)$$

Since $f(x)$ is the optimal wire sizing function, we have $y_{min} = f(x)$. Therefore, we obtain $f(x)$ as shown in Equation (16). □

Theorem 1 *Let $f(x)$ be an optimal wire-sizing function. Then,*

$$\frac{f(x) \cdot f''(x)}{f'^2(x)} = 0 \tag{18}$$

Proof: Direct results of first multiplying Equation (16) by the denominator of its right hand side, then differentiating both sides with respect to x. □

Corollary 1.1 *Let $f(x)$ be an optimal wire-sizing function. we have,*

$$f(x) = ax + b \tag{19}$$

Before we discuss more general optimal wire-sizing formula, one thing should be mentioned first. In the process of previous derivation of optimal formula, we have assumed that function f has second order derivative. This does not count the situation when the wire shape function is a constant value. As mentioned in section 1, the wire-sizing methodology is only beneficial when the interconnect resistance becomes significant. Otherwise, constant wire width can have the optimal RC constant time, hence optimal delay. In essential, even constant wire width can be regarded as one special case of wire sizing. Therefore the optimal wire sizing formula has to take constant wire width into account.

$$f_o(x) = \begin{cases} w_c & if\ D_c \leq D_l \\ ax + b & if\ D_c > D_l \end{cases}$$

where D_c and D_l are the elmore delay value of constant width wire and linear shape wire respectively. In order to calculate the actual delay value, we need to substitute the shape function $f(x)$ into Equation (15) and use numerical integral to solve it. Newton-cotes method [22] can be good candidate for such purpose.

Theorem 2 *The general optimal wire-sizing formula considering self-heating effect must satisfy the equation,*

$$f''(x)f(x) = (2 - \gamma)f'^2(x) \tag{20}$$

where γ is the exponent value in the relation between wire width and wire resistance as given in Equation (10).

Proof: Similar to the proof of Lemma 1. □

Remark 1 *When γ value equals to 1, our general optimal wire sizing formula reduces to the one existing in the literature [11]. When γ equals to 2, it reduces to the formula in Theorem 1. On the other hand, for any γ value that is larger than 2, the right-hand side of Equation (20) becomes negative. Because $f(x)$ cannot be negative, $f''(x)$ has to be negative, which implies a concave wire sizing formula. Therefore, $\gamma = 2$ is the boundary between convex wire-sizing ($f'(x) < 0 \wedge f''(x) > 0$) and concave wire-sizing ($f'(x) < 0 \wedge f''(x) < 0$). That is also one of the reasons why we are interesting in solving the optimal wire-sizing formula when given $\gamma = 2$.*

After solving the general function of optimal wire sizing, we need to determine the coefficient values, *i.e.*, the values of a and b in $f(x)$.

Lemma 2 *In order to solve for the coefficients a and b in the optimal wire sizing function $f(x)$, we need to solve the following mixed system of equalities and inequalities.*

$$\begin{cases} \frac{r_0(C_L + c_0\int_0^L (ax+b)dt}{c_0 R_d} = b^3 \\ \frac{r_0 C_L}{c_0(R_d + r_0\int_0^L \frac{1}{(aL+b)^2}dt)} = (aL+b)^3 \\ b > 0 \\ a < 0 \\ aL + b > 0 \end{cases}$$

The first two equations represent the boundary conditions at positions $x = 0$ and $x = L$. We require $b > 0$ because the value of b is the wire width at position x. A better lower bound can be found for b, that is the minimum valid wire width, or the positive pole value discussed in Section 2. The value of a should be negative, because in order to obtain the optimal interconnect delay, the wire width should decrease from position 0 along the direction of the current flow. The last inequality is similar to $b > 0$. We can also use a better wire width lower bound to replace the zero bound.

This system of equation can be solved efficiently by first finding the lower bound and upper bound of b, then doing incremental search in that range. Due to space limitation, details will not be elaborated here.

5 Uniform discrete wire sizing on self-heating effect

Generally, non-uniform continuous wire sizing helps define the contour of the uniform discrete counterpart, as illustrated in Figure 8.

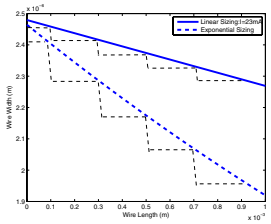

Figure 8: non-uniform wire sizing predicts the contour of uniform discrete wire sizing.

In this figure, the discrete wire sizing consists of five segments. Non-uniform continuous wire sizing and uniform discrete wire sizing become identical when the number of segments of uniform discrete wire sizing becomes infinite. In past works [11] addressing wire-sizing, it has been argued that when the number of segments is large, implementing discrete wire sizing would substantially increase the number of variables, hence, result in long runtime and large storage. Hence, non-uniform wire sizing is necessary to provide guidelines for the discrete solution. It can at least give an evaluation on how good the discrete wire sizing is by observing how far it deviates from the continuous optimal function. In practice, there are only several possible wire widths that can be chosen. Continuous optimal wire sizing function provides us hints on how the design library would be changed accordingly when considering self-heating effect, on the wire width range as well on the step size between two consecutive wire widths. After updating the library according to these guidelines, the continuous solution can be segmented into discrete portions using the given number of wire types in the library. Figure 8 depicts a segmented solution to our linear wire-sizing solution.

6 Results

We implemented and tested our algorithm in C on a Sun Blade150 workstation with 512 MB memory. The parameters are taken from the literature [11], which is also listed here in Table 2.

One major difference between our linear wire sizing approach and the conventional optimal solution is that we take

the current value and thermal conditions of the interconnect as parameters. In other words, conventional optimal wire

Table 2: Interconnect RC parameters

Unit Capacitance (c_0)	6×10^{-17}	$F/\mu m$
Driver Resistance (R_d)	25	Ω
Load Capacitance (C_L)	1×10^{-12}	F
Wire Length (L)	1000	μm

Table 3: Delay of linear wire-sizing compared to normalized exponential wire-sizing with different driver resistance.

$R_d(\Omega)$	Delay(D_l/D_e)
5	0.77
25	0.63
50	0.57

Table 4: Delay of linear wire-sizing compared to normalized exponential wire-sizing with different load capacitance.

$C_L(fF)$	Delay(D_l/D_e)
1	0.13
10	0.29
100	0.63

sizing generates a single solution if the wire dimension and other physical parameters are given. The wire sizing will not change even if the amount of current flow in that wire is expected to be significantly different. Such a solution is not practical in the face of thermal effects for future technologies. As illustrated in Figure 9, the exponential shape will be highly inaccurate when the wire width is small. The reason is that in narrow wires, self-heating effect becomes very significant. In this figure, the dashed line represents the exponential wire function, while the solid line is the corresponding linear sizing function. Additional lines in this figure illustrate the self-heating-aware linear wire sizing solution under different current conditions.

Another part of our experiment is to compare the actual delay of both linear wire sizing $f(x) = ax + b$ and exponential sizing $f(x) = ae^{-bx}$. We use Elmore delay in Equation (15) to represent the actual delay of the interconnect. Its value is calculated numerically by Newton-cotes method. It can be seen from the results that the ratio between linear sizing and exponential sizing depends on physical parameters. We change both driver resistance and load capacitance to achieve different results. Table 3 is the delay value for different drive resistance R_d. Table 4 is the delay value for different load capacitance. The delay value of exponential(D_e) wire-sizing has been normalized to one. We can observe 34.3% delay gain from continuous linear wire sizing on average with driver resistances range from 5Ω to 50Ω. And we also observed 65% delay gain on average with load capacitances from $1fF$ to $100fF$.

7 Conclusions

The major contribution of our work is describing a quadratic relation between wire resistance and wire width for high-layer global interconnect, and incorporating this model into the non-uniform continuous wire sizing optimization process. Our self-heating-aware optimal wire sizing formula is in a simple linear form $f(x) = ax + b$. This result helps clarify the trends of both continuous and discrete wire sizing, that the wire width slope should be flatter if self-heating effect becomes significant. Compared to non-thermal-aware exponential wire sizing, we obtain 49.7% delay gain on average.

8 Acknowledgments

This work has been partly funded by the NSF Grants CCF-0541337 and CNS-0546305.

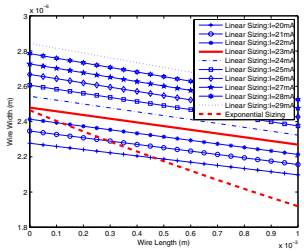

Figure 9: Comparison between conventional exponential wire-sizing and our linear wire-sizing function.

References

[1] A. H. Ajami, K. Banerjee, and M. Pedram. Analysis of substrate thermal gradient effects on optimal buffer insertion. In *ICCAD*, 2001.

[2] A. H. Ajami, K. Banerjee, M. Pedram, and L. P.P.P. van Ginneken. Analysis of Non-Uniform Temperature-Dependent Interconnect Performance in High Performance ICs. In *DAC*, Las Vega, NV, June 2001.

[3] K. Banerjee, A. Amerasekera, G. Dixit, and C. Hu. The effect of interconnect scaling and low-k dielectric on the thermal characteristics of the IC metal. In *IEDM*, 1996.

[4] K. Banerjee and A. Mehrotra. Global Interconnect Warming. *Circuit and Devices*, pages 16–32, 2001.

[5] K. Banerjee, A. Mehrotra, A. Sangiovanni-Vincentelli, and C. Hu. On Thermal Effects in Deep Sub-Micron VLSI Interconnects. In *DAC*, New Orleans, Louisiana, 1999.

[6] A. Bejan. *Heat Transfer*. John Wiley Sons, Inc, 1993.

[7] J. R. Black. Electromigration – A brief survey and some recent results. *IEEE Transactions on Electron Devices*, pages 338–347, 1969.

[8] M. R. Casu, M. Graziano, G. Piccinini, G. Masera, and M. Zamboni. Effects of Temperature in Deep-Submicron Global Interconnect Optimization. *Microelectronic Journal*, 2004.

[9] C. P. Chen, C. C. N. Chu, and D. F. Wong. Fast and exact simultaneous gate and wire sizing by Lagrangian relaxation. In *IEEE TCAD*, 1999.

[10] C. P. Chen and D. F. Wong. Optimal wire-sizing function with fringing capacitance consideration. In *DAC*, 1997.

[11] C. P. Chen Y. P. Chen and D. F. Wong. Optimal wire-sizing formula under the Elmore delay model. In *ICCAD*, 1996.

[12] D. Chen, E. Li, E. Rosenbaum, and S. Kang. Interconnect Thermal Modeling for Accurate Simulation of Circuit Timing and Reliability. *IEEE TCADICAS*, 19(2), February 2000.

[13] J. Cong, L. He, and Z. Pan. Global interconnect sizing and spacing with consideration of coupling capacitance. In *ICCAD*, 1997.

[14] J. Cong, C. K. Koh, and K. S. Leung. Simultaneous routing and wire sizing for performance and power optimization. In *ISLPED*, 1996.

[15] J. Cong and K. S. Leung. Optimal wiresizing under the distributed Elmore delay model. In *ICCAD*, 1993.

[16] J. Cong, K. S. Leung, and D. Zhou. Performance-driven interconnect design based on distributed RC delay model. In *DAC*, 1995.

[17] W. C. Elmore. The transient response of damped linear networks with particular regard to wide band amplifiers. *J. Applied Physics*, 1948.

[18] Y. Gao and D. F. Wong. Optimal shape function for a bidirectional wire under Elmore delay model. In *ICCAD*, 1997.

[19] S. Im and K. Banerjee. Full Chip Thermal Analysis of Planar(2-D) and Vertically Integrated(3-D) high-performance ICs. In *IEDM*, 2000.

[20] P. Kapur, J. P. McVittie, and K. C. Saraswat. Technology and Reliability Constrained Future Copper Interconnects – Part I: Resistance Modeling. *IEEE Transactions on Electron Devices*, 49(4), April 2002.

[21] N. Menezes, S. Pullela, F. Dartu, and L. Pillage. RC Interconnect Synthesis – A Moment Fitting Approach. In *ICCAD*, 1994.

[22] S. S. Rao. *Applied numerical methods for engineerings and scientists*. Prentice Hall, 2002.

[23] L. P. P. P. van Ginneken. Buffer placement in distributed RC-tree networks for minimal Elmore delay. In *ISCAS*, 1990.