Transformation of Min-Max Optimization to Least-Square Estimation and Application to Interconnect Design Optimization

Jimmy Shinn-Hwa Wang
Wayne Wei-Ming Dai

UCSC-CRL-95-03
March 6, 1995

Baskin Center for
Computer Engineering & Information Sciences
University of California, Santa Cruz
Santa Cruz, CA 95064 USA

ABSTRACT

This paper describes a novel approach to find a tighter bound of the transformation of the Min-Max problems into the one of Least-Square Estimation. It is well known that the above transformation of one problem to the other can lead to the proof that their target functions linearly bound each other. However, this linear bound is not a tight one. In this paper, we prove that if we transform the Min-Max problem into two Least-Square Estimation problems, where one minimizes the Root-Mean-Square (RMS) of the original function and the other one minimizes the RMS of the difference between the original function and an arbitrary constant, one can obtain a tighter bound between their target functions. The tighter bound given by this novel approach depends on the outcome of the second Least-Square Estimation problem, so there is a great incentive to choose the arbitrary constant which gives the smallest RMS of the second Least-Square Estimation problem. For a problem with a large number of variables, this novel tighter bound can be two to three order tighter than the old one.

Keywords: Min-Max, Least-Square Estimation, Linear Bound, Root-Mean-Square

Contents

1 Introduction 2

2 Previous Work : Transform the Min-Max Problem to One Least-Square Estimation Problem 4

2.1 Physical Example 4

2.2 Problem Definition 5

2.3 Transformation Formulation 5

2.4 Physical Meaning 7

2.5 Practical Considerations 9

3 Novel Transformation of the Min-Max Problem into Two Least-Square Estimation Problems 10

3.1 Problem Definition 10

3.2 Novel Transformation Formulation 10

3.3 Physical Meaning 17

3.4 Practical Considerations 19

4 Implementation and Practical Consideration 20

5 Experimental Results 21

5.1 Example 5.1: A Tree Network 22

5.2 Example 5.2: A Network with a Loop 23

5.3 Example 5.3: A Clock Tree Network 24

6 Concluding Remarks 24

References 26

1 Introduction

For most of the Computer-Aided Design software, they usually ask for an optimization as part of their heuristic algorithm. The most common optimization problem encountered is the minimization of the maximum of all the observable outputs, which is often referred to as the Min-Max optimization problem. Because the Min-Max optimization problem is generally a nonlinear programming problem, it is not only hard to solve but also takes a long computation time. One Alternative to solve the Min-Max optimization problem and to obtain a good solution in a relative short period of time is to transform the original Min-Max optimization problem into a Least-Square Estimation (LSE) problem, which is not only very easy to solve but also has a number of well defined methods for solving it. It can be shown that the target function of the original Min-Max optimization problem and that of the transformed LSE problem linearly bound each other. Although these two problems have different objective functions, optimizing the transformed problem can produces a solution to the original Min-Max problem. By optimizing the LSE problem, one tries to minimize a *target function* which is the Root-Mean-Square (RMS) value of a given function. In this way, one obtains a *solution function* which is a set of values of a given function. By the definition of the LSE problem, when evaluating the Root-Mean-Square value of all the values of a solution function, the RMS value is at its minimum. The solution to the original Min-Max problem is found from the maximum value of the solution function of the LSE problem. However, this does not mean that solving the LSE problem can lead to the exact solution to the original Min-Max problem, it mealy states that solving the LSE problem gives one possible solution to the original Min-Max problem. The term: “linearly bound each other” means after one obtains the solution function of the LSE problem, one can translate the maximum value of the solution function at hand to a solution of the original Min-Max problem. It dose not say anything about the quality of this solution compared with other methods of solving the same original Min-Max problem.

As illustrate in Section 2, for any single observable output Min-Max problem, the maximum value of the solution function of the LSE problem is the same as the solution of the original Min-Max problem. However, for multiple observable outputs Min-Max problem, the Root-Mean-Square value of the solution function of the LSE problem linearly bounds the solution of the original Min-Max problem with a range. The solution to the original Min-Max problem can still be found.

However, because of this range, one cannot be sure about the quality of the solution of the original Min-Max problem. Worst of all, for Min-Max problems with large number of observable outputs, the solution of the original Min-Max problem is bounded by a huge range, between one and several tens multiplies the solution of the LSE problem. For example, if a Min-Max problem has 1000 observable output variables, then the solution of the Min-Max problem can lie between one and 31.62 multiplies the Root-Mean-Square value of the transformed LSE problem, since the square root of 1000 is 31.62. The drawback of this huge range is that it gives rise to such a huge uncertainty in transferring the maximum value of the solution function of the LSE problem back to the Min-Max solution. This renders the idea of the transformation and solving it through x LSE optimization much less useful.

Previous researches that utilizes the transformation include those of Zhu et al. [5] and Wang et al. [3]. Both of these researches lack the ability to precisely translate the maximum value of the solution function of the LSE problem back to the original Min-Max problem.

In this paper, a novel transformation is formulated, follow immediately by a proof that this transformation indeed gives the tightest bound for the ideal case. Because of this tightest bound, the maximum value of the solution function of the LSE problem can be transformed back as the solution of the original Min-Max problem even for the case of the multiple observable output variables.

As demonstrated in Example 5.3 in Section 5, where the best case happens when the difference between all the values of a given function can be minimized. This makes this novel approach especially effective in solving the problem of minimizing the delay of the equal path length clock tree.

Section 2 presents the original Min-Max optimization problem, the transformation into one LSE problem, and the proof of their target functions linearly bound each other. The drawback of this type of transformation for the case of the multiple observable output variables is also discussed in detail. Section 3 presents the novel transformation into two LSE problems and the proof of the tighter linear bound between their target functions. The advantage of this novel transformation and an ideal case of the solution translation is discussed in detail in Section 3. Section 4 shows the details about the implementation of the optimization of the LSE problems. Section 5 demonstrates the

input variables. There are a total of m outputs in this interconnect graph and are called the observable outputs. The delay is a function controlled by the widths of all the branches and are observed at the m outputs. The optimization problem is to minimize the maximum of all the delays observed at the m outputs through manipulating the n widths.

2.2 Problem Definition

Given a *positive* function $G(W, J)$ where $W = \{w_i | i = 1 \ldots n\}$ and $J = \{j | j = 1 \ldots m\}$, the Min-Max problem wants minimize the maximum of all its values. The set of W is called the controllable input variables, and the set of J is called the observable outputs. There are n controllable input variables, w_i, and m observable output variables, j, for this optimization problem. Denote $G(W, j)$ to be the discrete value of $G(W, J)$ at j. The Min-Max optimization problem is to control the w_1 through w_n input variables so that maximum value of the $G(W, J)$ is minimized.

If one defines the target function of the Min-Max problem to be $F(W, J)$, one can write:

$$\min_W [F(W, J)] \triangleq \min_W \left[\max_J [G(W, J)] \right]. \tag{2.1}$$

This is a general nonlinear programming problem and usually takes a long computation time to solve it.

From any general optimization textbook, an easier alternative of solving this Min-Max problem is to transform it into a Least-Square Estimation (LSE) problem. In order to solve this LSE problem, one has to minimize the RMS of the given function $G(W, J)$. Solving the LSE problem gives a solution function whose maximum can be translated into a solution of the original Min-Max problem.

2.3 Transformation Formulation

The detail formulation of the transformation is as follows. Let $G(W, j)$ be the observed $j - th$ output, and let the column vector

$$\Theta(W, J) \triangleq \{G(W, 1), G(W, 2), \ldots, G(W, j), \ldots, G(W, m)\}^T, \tag{2.2}$$

represents the estimation vector, where T denotes the transposition operation, $j \in J$ and $1 \leq j \leq m$, and m is the number of the outputs. The summation of all squares of the estimations $\Phi(W, J)$ is:

$$\Phi(w_1, w_2, \ldots, w_n, 1, \ldots, m) \triangleq \Theta(W, J)^T \Theta(W, J) = \sum_{j=1}^{m} [G(W, j)]^2 \tag{2.3}$$

If one defines the Root-Mean-Square (RMS) of the estimation as:

$$\varphi(w_1, w_2, \ldots, w_n, 1, \ldots, m) \triangleq \sqrt{\frac{\Phi(W, J)}{m}} = \sqrt{\sum_{j=1}^{m} \frac{[G(W, j)]^2}{m}}, \tag{2.4}$$

then new optimization problem becomes the minimization of the RMS of the given function $G(W, J)$. One can write the new target function of the optimization as:

$$\min_W [\varphi(W, J)] \triangleq \min_W \left[\sqrt{\sum_{j=1}^{m} \frac{[G(W, j)]^2}{m}} \right]. \tag{2.5}$$

The new target function (Equation (2.5)) does not correspond to the original target function (Equation (2.1)). However, it can be shown that the minimization result of the new target function linearly bounds the original optimization solution. As can be seen later, if the linear bound is tight, the maximum value of the optimization result of the new solution function can be translated to be the solution of the original Min-Max problem.

The following is the proof of the target functions of the transformation linearly bound each other.

Theorem 1: *Given a function $G(W, J)$, the minimum of the Root-Mean-Square (RMS) as defined in Equation (2.5) and the minimum of the maximum as defined in Equation (2.1) linearly bound each other.*

Proof:

Given that the largest of all the $G(W, J)$ is equal to $F(W, J)$, $F(W, J) \triangleq \max_J G(W, J)$. For all $G(W, j)$, $G(W, j) \leq F(W, J)$, $(1 \leq j \leq m)$, one can obtain Equation (2.6 from Equation (2.5)

$$\varphi(W, J) \triangleq \sqrt{\sum_{j=1}^{m} \frac{[G(W, j)]^2}{m}} \leq \sqrt{\sum_{j=1}^{m} \frac{[\max_J G(W, J)]^2}{m}} = \max_J G(W, J) = F(W, J). \tag{2.6}$$

between the $Area_{RMS} = \varphi$ and the $\sqrt{m} \cdot Area_{RMS} = \sqrt{m} \cdot \varphi$. By minimizing the area of $Area_{RMS}$, one also brings down the area of $\sqrt{m} \cdot Area_{RMS}$, since the area of $Area_{max}$ lies in between them, so it is also minimized. However, for the multiple output observable output variable case shown in the figure, the range between the $Area_{RMS}$ and $\sqrt{m} \cdot Area_{RMS}$ can be quite large. This gives rise to the uncertainty about the minimization of the original target function when one attempts to optimize the LSE problem.

2.5 Practical Considerations

When applied to a case of the single observable output, $m = 1$, Equation (2.8) becomes $\varphi \leq F \leq \sqrt{1}\varphi$. This means the two target functions are equal, so minimizing one is the same as minimizing the other. So the solution of the LSE problem, φ, can be taken as the solution of the Min-Max Problem. However, for a case of the multiple observable output variable, Equation (2.8) is $\varphi \leq F \leq \sqrt{m}\varphi$. This means the *target function* of the LSE problem, φ, linearly bounds the solution of the original Min-Max problem, F, with a range. This means the two target functions are not the same, so minimizing one does not guarantee the minimization of the other. This range is defined as from φ to $\sqrt{m}\varphi$. For example, if a Min-Max problem has 1000 observable outputs which is common for a global clock distribution net, then the solution of the original Min-Max problem is bounded between one and $\sqrt{1000} = 31.62$ multiplies the *target function* of the LSE problem. This huge range give rise to one's hesitation about the quality of the solution to the Min-Max problem . This is the drawback when one tries to use the transformation to solve the Min-Max problem on the case of the multiple observable output. The huge linear bound range and the uncertainty in the solution translation renders the idea of the transformation and solving through LSE optimization much less useful.

In Section 3, a novel transformation approach is presented. This novel transformation can be shown to have a tighter linear bound than that of Equation (2.8) and, in the best case, can make the solution translation of the case of the multiple observable output identical to that of the case of the single observable output.

3 Novel Transformation of the Min-Max Problem into Two Least-Square Estimation Problems

The transformation presented in Section 2 changes the target function of the Min-Max problem to the RMS of a different LSE problem. The above transformation works only because the minimization the RMS of the different LSE problem is identical to the minimization of the original problem. In the case of the multiple observable output variable, the minimization of the LSE problem cannot be proved to be identical to the minimization of the Min-Max problem. This prevents the use of the transformation as a method of solving the original Min-Max problem.

In this Section, a novel transformation which gives a tighter linear bound so that it works even in the case of the multiple observable output is presented. This novel transformation transforms the original Min-Max problem into two LSE problem. One of the LSE problem is the minimization of the RMS of the given function. This LSE problem is the same as the one presented to in Section 2. The second LSE problem is the minimization of the RMS of the difference between an arbitrary constant and the given function $G(W, J)$. Solving these two LSE problems together gives two *target function values* and one *solution function.* The maximum value of the solution function of the two LSE problem can be translated into a solution of the original Min-Max problem. The solution to the second LSE problem only serves as an assurance that when minimizing the first LSE problem, one is indeed minimizing the original Min-Max problem.

3.1 Problem Definition

The problem definition is identical to the one that presented in Section 3.1. Given a *positive* function $G(W, J)$ where $W = \{w_i | i = 1 \ldots n\}$ and $J = \{j | j = 1 \ldots m\}$, one can define the target function of the Min-Max problem to be $F(W, J)$, and write:

$$\min_W [F(W, J)] \triangleq \min_W \left[\max_J [G(W, J)] \right]. \tag{3.1}$$

3.2 Novel Transformation Formulation

The detail formulation of the novel approach is as follows.

The transformed optimization problem consists of two LSE problem. The first one is the minimization of the RMS of the given function $G(W, J)$. The second LSE problem is the minimization of the RMS of the difference between an arbitrary constant c and the given function $G(W, J)$. For this new LSE problem, one has to create a new Min-Max problem. Define a new function $\Delta(W, J) \triangleq [\max_J[c - G(W, J)]]$, where c is an arbitrary constant which remains the same during the entire LSE optimization process. The choice of c affects the outcome of the the solution translation a great deal and will be discussed in detail in Section 3.4. Define the target function of the new Min-Max problem to be $\Delta(W, J)$, one can write:

$$\min_W [\Delta(W, J)] \triangleq \min_W \left[\max_J [c - G(W, J)] \right]. \tag{3.2}$$

Denote $G(W, j)$ to be the discrete value of $G(W, J)$ at j. Define the estimation vector for the first LSE problem to be $\Theta(W, J) \triangleq \{G(W, 1), G(W, 2), \ldots, G(W, j), \ldots, G(W, m)\}^T$, where T denotes the transposition operation, $j \in J$ and $1 \leq j \leq m$, and m is the number of the observable outputs. Define the summation of all squares of the first estimations, $\Phi(W, J)$, to be:

$$\Phi(w_1, w_2, \ldots, w_n, 1, \ldots, m) \triangleq \Theta(W, J)^T \Theta(W, J) = \sum_{j=1}^{m} [G(W, j)]^2 \tag{3.3}$$

Define the Root-Mean-Square (RMS) of the first estimation as:

$$\varphi(w_1, w_2, \ldots, w_n, 1, \ldots, m) \triangleq \sqrt{\frac{\Phi(W, J)}{m}} = \sqrt{\sum_{j=1}^{m} \frac{[G(W, j)]^2}{m}} \tag{3.4}$$

Similarly, denote $\Delta(W, j)$ to be the discrete value of $\Delta(W, J)$ at j. Define the estimation vector for the second LSE problem to be:

$$\Omega(W, J) \triangleq \{c - G(W, 1), c - G(W, 2), \ldots, c - G(W, j), \ldots, c - G(W, m)\}^T, \tag{3.5}$$

where T denotes the transposition operation. Define The summation of all squares of the second estimations, $\Psi(W, J, c)$, to be:

$$\Psi(w_1, w_2, \ldots, w_n, 1, \ldots, m, c) \triangleq \Omega(W, J)^T \Omega(W, J) = \sum_{j=1}^{m} [c - G(W, j)]^2 \tag{3.6}$$

Define the Root-Mean-Square (RMS) of the second estimation as:

$$\psi(w_1, w_2, \ldots, w_n, 1, \ldots, m, c) \triangleq \sqrt{\frac{\Psi(W, J, c)}{m}} = \sqrt{\sum_{j=1}^{m} \frac{[c - G(W, j)]^2}{m}} \tag{3.7}$$

The transformed optimization problem consists of two LSE problem. The first one is the minimization of the RMS of the given function $G(W, J)$. The second LSE problem is the minimization of the RMS of the difference between an arbitrary constant c and the given function $G(W, J)$. One can write the new target function of the first optimization as:

$$\min_W [\varphi(W, J)] \triangleq \min_W \left[\sqrt{\sum_{j=1}^{m} \frac{[G(W, j)]^2}{m}} \right]. \tag{3.8}$$

Likewise, one can write the new target function of the second optimization as:

$$\min_W [\psi(W, J, c)] \triangleq \min_W \left[\sqrt{\sum_{j=1}^{m} \frac{[c - G(W, j)]^2}{m}} \right]. \tag{3.9}$$

The two new target functions (Equation (3.8)) and (Equation (3.9)) do not correspond to the original target function (Equation (3.1)). However, it can be shown that the linear combination of the minimization results of the two new target functions linearly bound the original optimization solution. As can be seen in Section 3.4, if the linear bound is the tightest for the idea case, the optimization result of the maximum value of the *solution function* of the two LSE optimization can be translated to be the solution of the original Min-Max problem.

The following is the proof of the target functions of the transformation linearly bound each other. From Theorem 1, one has:

$$\varphi(W, J) \leq F(W, J) \leq \sqrt{m} \cdot \varphi(W, J). \tag{3.10}$$

Similarly, for the new Min-Max problem and the new transformed LSE problem, one can prove their target functions linearly bound each other.

Theorem 2: *Given a function $\Delta(W, J)$, the minimum of the Root-Mean-Square (RMS) defined in Equation (3.7) and the minimum of the maximum defined in Equation (3.2) linearly bound each other.*

Proof:

Given that the largest of all the $\Delta(W, J)$ is $H(W, J)$. $H(W, J) \triangleq \max_J \Delta(W, J)$. For all $\Delta(W, j)$, $\Delta(W, j) \leq H(W, J)$, $(1 \leq j \leq m)$. From Equation (3.7), one has:

$$\psi(W, J, c) \triangleq \sqrt{\sum_{j=1}^{m} \frac{[c - G(W, j)]^2}{m}} \leq \sqrt{\sum_{j=1}^{m} \frac{\max_J [c - G(W, J)]^2}{m}} = \max_J \Delta(W, J) = H(W, J). \quad (3.11)$$

On the other hand, $H(W, J)$ is defined as

$$\begin{aligned} H(W, J) \triangleq \max_J [c - G(W, J)] = \sqrt{\max_j [c - G(W, J)]^2} &\leq \sqrt{\sum_{j=1}^{m} [c - G(W, j)]^2} \\ = \sqrt{m \cdot \sum_{j=1}^{m} \frac{[c - G(W, j)]^2}{m}} &= \sqrt{m} \cdot \psi(W, J, c). \end{aligned} \quad (3.12)$$

From Equation (3.11) and (3.12), one has

$$\psi(W, J, c) \leq H(W, J) \leq \sqrt{m} \cdot \psi(W, J, c). \quad (3.13)$$

This concludes the proof that $\psi(W, J, c)$ and $H(W, j)$ linearly bound each other. □

From the solution of the two LSE problem, one can transform them back to a solution of the original Min-Max problem.

Theorem 3: *Given a function $G(W, J)$, the minimum of the Root-Mean-Square (RMS) as defined in Equation (3.4), and the sum of the minimum of the RMS as defined in Equation (3.4) and sqrtm multiplies the minimum of the RMS as defined in Equation (3.7) linearly bound the target function of the Min-Max problem defined in Equation (3.1). i. e.*

$$\varphi(W, J) \leq F(W, J) \leq \varphi(W, J) + \sqrt{m} \cdot \psi(W, J, c). \quad (3.14)$$

Proof:

From the Theorem 1, one has:

$$\varphi(W,J) \leq F(W,J). \tag{3.15}$$

The second half of the equation is:

$$F(W,J) \leq \varphi(W,J) + \sqrt{m} \cdot \psi(W,J,c). \tag{3.16}$$

Substitute the definitions of φ and ψ into the above equation, one has:

$$\max_j [G(W,J)] \leq \sqrt{\sum_{j=1}^{m} \frac{[G(W,j)]^2}{m}} + \sqrt{m} \cdot \sqrt{\sum_{j=1}^{m} \frac{[c - G(W,j)]^2}{m}}, \tag{3.17}$$

or

$$\max_j [G(W,J)] \leq \sqrt{\sum_{j=1}^{m} \frac{[G(W,j)]^2}{m}} + \sqrt{\sum_{j=1}^{m} [c - G(W,j)]^2}, \tag{3.18}$$

The remaining of the proof of the theorem is by first establishing the extreme value is in fact the global minimum, and later prove that this extreme value is equal to the left hand side of the Equation (3.18). If the global minimum is indeed equal to the left hand side of the Equation (3.18), then this Equation (3.18) holds for all value of c. Throughout the proof, because the case of the multiple outputs is analyzed here, it is assumed that there are at least *two* observable output variables.

It is apparent from Equation (3.14), the choice of c determines how tight the bound will be. It can be shown that the c, which makes $\psi(W,J,c)$ assumes the smallest value, gives the tightest bound. In order to find such a constant c, one takes the partial derivative of the upper bound function w.r.t. c, one has:

$$\frac{\partial}{\partial c}\left[\varphi(W,J) + \sqrt{m} \cdot \psi(W,J,c)\right] = \frac{\partial}{\partial c}\varphi(W,J) + \sqrt{m} \cdot \frac{\partial}{\partial c}\psi(W,J,c) = 0 + \sqrt{m} \cdot \frac{\partial}{\partial c}\psi(W,J,c). \tag{3.19}$$

Setting the partial derivative equal to zero and solve, one can find the minimum of the right hand side of Equation (3.18). From the definition:

$$\psi(W,J,c) \triangleq \sqrt{\sum_{j=1}^{m} \frac{[c-G(W,j)]^2}{m}}. \tag{3.20}$$

In order to find the extreme value of $\psi(W,J,c)$ w.r.t. c, one sets the first partial derivative of $\psi(W,J,c)$ w.r.t. c equal to zero and solve. To find out the whether it si a global minimum of global maximum, one needs to find out the sign of the second partial derivative of $\psi(W,J,c)$ w.r.t. c. Take the first and second partial derivative of $\psi(W,J,c)$ w.r.t. c, one has:

$$\frac{\partial\psi}{\partial c} = \frac{\sum_{j=1}^{m} \frac{2\cdot[c-G(W,J)]}{m}}{\sqrt{\sum_{j=1}^{m} \frac{[c-G(W,j)]^2}{m}}} = 0, \tag{3.21}$$

$$\frac{\partial^2\psi}{\partial c^2} = \frac{\sum_{j=1}^{m} \frac{2}{m}}{\left(\sqrt{\sum_{j=1}^{m} \frac{[c-G(W,j)]^2}{m}}\right)^3}. \tag{3.22}$$

Solving for $\frac{\partial\psi}{\partial c} = 0$, one has:

$$\sum_{j=1}^{m} \frac{2\cdot[c-G(W,J)]}{m} = 0. \tag{3.23}$$

In order for the above equation to be true, the value of c must be

$$c = avg_j\left[G(W,J)\right]. \tag{3.24}$$

To find out whether this extreme value of $\psi(W,J,c)$ is a minimum or a maximum, one looks at the sign of $\frac{\partial^2\psi}{\partial c^2}$. Simplify $\frac{\partial^2\psi}{\partial c^2}$, one has:

$$\frac{\partial^2\psi}{\partial c^2} = \frac{\sum_{j=1}^{m} \frac{2}{m}}{\left(\sqrt{\sum_{j=1}^{m} \frac{[c-G(W,j)]^2}{m}}\right)^3} = \frac{2}{\left(\sqrt{\sum_{j=1}^{m} \frac{[c-G(W,j)]^2}{m}}\right)^3} \tag{3.25}$$

Because $[c-G(W,j)]^2$ is always greater than or equal to zero, so $\frac{\partial^2\psi}{\partial c^2} > 0$, which means $\psi(W,J,c)$ is a concave upward function, and its value at $c = avg_j\left[G(W,J)\right]$ is a global minimum.

The following three cases list all the possible values of the choice of constant c:

- Case I: $c \leq avg_j\left[G(W,J)\right]$
- Case II: $c \geq \max_j\left[G(W,J)\right] \geq avg_j\left[G(W,J)\right]$
- Case III: $\max_j\left[G(W,J)\right] \geq c > avg_j\left[G(W,J)\right]$

Because $c = avg_j\left[G(W,J)\right]$ is a global minimum, one only need to find out the extreme value of the right hand side of Equation (3.18) in Case I.

Let

$$avg_j\left[G(W,J)\right] \triangleq \frac{\sum_{j=1}^m G(W,j)}{m}. \tag{3.26}$$

Assume $c \leq avg_j\left[G(W,J)\right]$. It is known that Root-Mean-Square of a function is greater than or equal to the Mean of the function, that is:

$$\sqrt{\sum_{j=1}^m \frac{[c-G(W,j)]^2}{m}} \geq \frac{\sum_{j=1}^m G(W,j)}{m} = avg_j\left[G(W,J)\right]. \tag{3.27}$$

So

$$\sqrt{\sum_{j=1}^m \frac{[c-G(W,j)]^2}{m}} + \sqrt{\sum_{j=1}^m [c-G(W,j)]^2} \geq \frac{\sum_{j=1}^m G(W,j)}{m} + \sqrt{\sum_{j=1}^m [c-G(W,j)]^2}. \tag{3.28}$$

Since $c \leq avg_j\left[G(W,J)\right]$, thus

$$\begin{aligned} \max_j\left[G(W,J)\right] - c &\geq \max_j\left[G(W,J)\right] - avg_j\left[G(W,J)\right], \\ \sqrt{\sum_{j=1}^m [c-G(W,j)]^2} &\geq \max_j\left[G(W,J)\right] - c, \\ \sqrt{\sum_{j=1}^m [c-G(W,j)]^2} &\geq \max_j\left[G(W,J)\right] - avg_j\left[G(W,J)\right]. \end{aligned}$$

Rearrange, one has:

$$\sqrt{\sum_{j=1}^m [c-G(W,j)]^2} + avg_j\left[G(W,J)\right] \geq \max_j\left[G(W,J)\right]. \tag{3.29}$$

From Equation (3.28) and (3.29), if $c \le avg_j\,[G(W,J)]$, one has:

$$\sqrt{\sum_{j=1}^{m}\frac{[c-G(W,j)]^2}{m}}+\sqrt{\sum_{j=1}^{m}[c-G(W,j)]^2} \ge \max_j\,[G(W,J)]\,. \tag{3.30}$$

From Equations (3.30), one can write:

$$\varphi(W,J)+\sqrt{m}\cdot\psi(W,J,c) = \sqrt{\sum_{j=1}^{m}\frac{[G(W,j)]^2}{m}}+\sqrt{m}\cdot\sqrt{\sum_{j=1}^{m}\frac{[c-G(W,j)]^2}{m}}$$
$$\ge \max_j\,[G(W,J)] = F(W,J) \tag{3.31}$$

So for $c \le avg_j\,[G(W,J)]$, the minimum of $\varphi(W,J)+\sqrt{m}\cdot\psi(W,J,c)$ is equal to $\max_j\,[G(W,J)] = F(W,J)$, so

$$F(W,J) \le \varphi(W,J)+\sqrt{m}\cdot\psi(W,J,c). \tag{3.32}$$

From Equations (3.15) and (3.32), one has $\varphi(W,J) \le F(W,J) \le \varphi(W,J)+\sqrt{m}\cdot\psi(W,J,c)$. □

3.3 Physical Meaning

Given the same *positive* function $G(W,J)$ where $W=\{w_i|i=1\ldots n\}$ and $J=\{j|j=1\ldots m\}$, the Min-Max problem wants to minimize the maximum of all its values. The physical meaning of the novel transformation is not only to minimize the RMS of the given function $G(W,J)$ but also to minimize the difference between the given function $G(W,J)$ and an arbitrary constant c. Preferably, the constant c equals to the average of the final results of the function $G(W,J)$. Plot the function $G(W,J)$ after optimization in Figure 3.1, one can see the difference between each $G(W,j)$ and the constant c has also been minimized. Plot the constant c, the maximum, minimum, average, and RMS value of the function $G(W,J)$ in Figure 3.2, one can tell the relative relations between them. Because the upper bound is now $Area_{RMS}+\sqrt{m}\cdot\psi$, it is much tighter than the range shown in Figure 2.3. This figure gives the insight into why one choose this new linear bound over the old one. Looking at Figure 3.2, one can see that $Area_{max}=F$ lies between the $Area_{RMS}=\varphi$ and the $Area_{RMS}=\varphi$ plus the area of some other function. One wants the area of some other function to

$\sqrt{m} \cdot \psi(W, J, c) = \varphi(W, J) + \sqrt{m} \cdot 0 = \varphi(W, J)$, which means $\varphi(W, J)$ is equal to $F(W, J)$. Because the above ideal case makes the two target function equal to one another, so the minimization of the LSE problem is indeed identical to the minimization of the original Min-Max problem. This is the tightest linear bound possible which makes the translation of the solution in the case of the multiple observable output variable exactly the same as that of the case of the single observable output variable.

4 Implementation and Practical Consideration

The Levenberg-Marquardt method is used to solve the Least-Square Estimation problem [2]. Theorems 1, 2, and 3 show the consistency between the minimization of the original problem and the minimization of the transformed Least-Square Estimation problem. Consider the physical example in Figure 2.1, starting with an arbitrary initial solution of width assignment $W^{(0)} = \{{w_1}^{(0)}, {w_2}^{(0)}, \ldots, {w_n}^{(0)}\}^T$, the next width assignment W, which according to Levenberg-Marquardt, is optimized according to the following formula:

$$W^{(k+1)} = W^{(k)} - (J^T J + \lambda\Lambda)^{-1} J^T \begin{bmatrix} \Theta|_{W^{(k)}} \\ \Omega|_{W^{(k)}} \end{bmatrix} \tag{4.1}$$

where k is the number of iteration, $\Theta|_{W^{(k)}}$ the column vector of delays from the source to all the receivers at the $k-th$ iteration, and $\Omega|_{W^{(k)}}$ the column vector of damping ratio errors from the source to all the receivers at the $k-th$ iteration. The $\Theta|_{W^{(k)}}$ is defined in Equation (2.2) and the $\Omega|_{W^{(k)}}$ is defined in Equation (3.5). J is the $2m \times n$ sensitivity matrix, J^T is the transposition matrix of J where the $(i,j)th$ element $J^T(i,j) = J(j,i)$, Λ is a diagonal matrix in which the values of its diagonal elements are the same as the diagonal elements of $J^T J$, and λ is the *Lagrange Multiplier* properly selected to speed up the convergence of the optimization process [2]. $J^T \begin{bmatrix} \Theta|_{W^{(k)}} \\ \Omega|_{W^{(k)}} \end{bmatrix}$ represents the gradient around the current width assignment $W^{(k)}$. To obtain the sensitivity matrix J, the $(i,j)th$ element is defined as:

$$J(i,j) = \begin{cases} \frac{\partial \Theta[i]}{\partial w_j}, & \text{if } 1 \le i \le m \\ \frac{\partial \Omega[i-m]}{\partial w_j}, & \text{if } m+1 \le i \le 2m \end{cases} \quad . \tag{4.2}$$

The partial derivatives are computed using a central difference method. The optimization continues until the maximum delay cannot be further improved, or the iteration number exceeds a preset limit. The convergence to the optimal values of Levenberg-Marquardt method is proved in [2].

5 Experimental Results

The examples that were tested are constructed with High Performance MCM process technologies published by Frye [1]. The important parameters of the MCM process are listed in Table 5.1. In the case of the uniform width, all of the widths are equal to $25\mu m$ for all of the examples tested. All the drivers are modeled with a step input voltage source in series with the parallel combination of a 12Ω resistor and a $4.3pF$ capacitor. All the receivers in Examples 5.1 and 5.3 are modeled using a $2.5pF$ capacitor. All the receivers in Example 5.2 are modeled using a $4.5pF$ capacitor.

HIGH PERFORMANCE MCM-D	
Thickness of Dielectric (μm)	5
ε_{rel}	3.2
Thickness of Metal (μm)	2.5
R ($\Omega/\mu m$) for typical edge width	2.4
L ($nH/\mu m$) for typical edge width	2.9
C ($pF/\mu m$) for typical edge width	1.39
lower bound metal line width (μm)	10
typical metal line width (μm)	25
upper bound metal line width (μm)	50

Table 5.1: **The High Performance MCM technologies process parameters.**

	Maximum Path Delay (Uniform) (nS)	Lower Bound φ (nS)	Maximum Path Delay (Optimal) (nS)	Upper Bound $\sqrt{m}\varphi$ (nS)	Percent Improvement (%)
Example 5.1	0.9874	0.5833	0.7182	1.4288	27.26
Example 5.2	0.9354	0.6266	0.8245	1.2532	11.85
Example 5.3	1.2308	0.8510	0.8533	2.0841	30.67

Table 5.2: **Comparison between the Uniform Width Design and the Optimal Design.**

5.3 Example 5.3: A Clock Tree Network

Example 5.3 is a clock tree network shown in Zhu's ICCAD paper [5]. Figure 5.3 (a) shows the topology, lengths, and widths of the optimal design for all of the edges. All of the the simulation waveforms of the optimal design are shown in 5.3 (b), those of the old optimal design are shown in 5.3 (c), and those of the uniform width design are shown in Figure 5.3 (d). The maximum path delays, their respective bounds, and their percentage of improvements are listed in Tables 5.2 and 5.3.

The old optimization method presented by Wang et al. [3] does improve the performance of this clock tree by reducing the maximum path delay and skew through minimizing the delays and the damping ratio errors. Although the skew has been reduced to $96pS$, there is still room for improvement. This novel optimization method minimizes the RMS of all the delays and the RMS of all the differences between the delays and an arbitrary constant. The resulting skew is reduced to only $4.4pS$. This method not only guarantees the quality of the solution to the Min-Max delay problem but also implies the minimization of the skew. The results clearly show that this novel optimization method is most suitable to be used in the case when the differences between all the observable output variables can be minimized. The minimization of the maximum path delay of the equal-path-length clock tree such as **H** clock tree is an application with this kind of characteristic.

6 Concluding Remarks

The Min-Max optimization problem is often required in most of the Computer-Aided Design software, it is generally a nonlinear programming problem which is difficult to solve. An Alternative is to transform the Min-Max problem into one LSE problem and solve the LSE problem instead. However, this method has a big drawback when it comes to deal with the case of multiple observable variables because the quality of the solution of the Min-Max problem cannot be guaranteed because the uncertainty introduced by the loose linear bound.

This paper describes a novel approach which transform the original problem into two LSE problems. A tighter linear bound can be found through this novel approach. For large number of observable output variables, the novel tighter bound can be two to three order of magnitude tighter than the old one. In the best case where the choice of the constant c makes the target function

of the second LSE problem equals to zero. This tightest linear bound makes the quality of the solution for the multiple observable output case is as good as the single observable output case.

For all of the examples tested, the optimization of the two LSE problems gives a better solution to the original Min-Max problem as indicated in Section 5. Table 5.2 also demonstrates that the respective maximum value of a given function is indeed within its respective linear bound. The best example shown in the Example 5.3 where the target function of the second LSE problem is evaluated to be near zero, so the minimization of the two transformed LSE problem is identical to the minimization of the original Min-Max problem. This leads to the conclusion that this novel approach is best suitable to be used in the equal-path-length clock tree delay optimization.

References

[1] Robert C. Frye. Physical scaling and interconnection delays in multichip module. *IEEE Trans. on CHMT*, 17(1):30–37, Feb. 1994.

[2] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of Society of Indust. Appl. Math.*, 11:431–441, 1963.

[3] Jimmy Wang and Wayne Dai. Optimal design of self-damped lossy transmission lines for multichip modules. In *Proc. of IEEE ICCD-94*, October 1994.

[4] D. Zhou, F. Tsui, J. S. Cong, and D. S. . A distributed-RLC model for MCM layout. In *Proc. of IEEE Multi-Chip Module Conference MCMC-93*, pages 191–197, 1993.

[5] Qing Zhu and Wayne Dai. Optimal sizing of high speed clock networks based on distributed RC and lossy transmission line models. In *IEEE/ACM International Conference on Computer-Aided Design*, 1993.