*International Journal of Innovative Computing, Information and Control*
Volume **8**, Number **5(B)**, May **2012**
 ISSN 1349-4198
pp. 3781–3792

# A TOTAL LEAST SQUARES TECHNIQUE FOR THE DESIGN OF A WIRELESS BEAMFORMING SYSTEM

Zhi Guo Feng[1,2], Ka Fai Cedric Yiu[2] and S. Nordholm[3]

[1]College of Mathematics
Chongqing Normal University
No. 12, Tianchen Road, Shapingba, Chongqing 400047, P. R. China
z.feng.scholar@gmail.com

[2]Department of Applied Mathematics
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong, P. R. China
macyiu@polyu.edu.hk

[3]School of Electrical and Computer Engineering
Curtin University
Perth, W.A., Australia
S.Nordholm@curtin.edu.au

Received October 2010; revised May 2011

Abstract. *In this paper, a new model of broadband microphone array system is proposed, where wireless transmission is employed to relay the signals to a fast server nearby in order to find the optimal filter coefficients continuously before the filter weights are sent back to the device. By using sequences of calibration signals, we formulate a total least squares problem and transform it into an equivalent optimization problem. By exploiting the structure of the problem, analytic expressions for the optimal solution can be obtained. For illustration, two numerical results are presented.*
**Keywords:** Total least squares, Microphone array, Beamforming

1. **Introduction.** In modern hands-free communication systems, there are many disturbance sources that cause unwanted sound [4], which may degrade the comprehension of the wanted speech. These disturbances vary depending on the environmental preliminaries. In this particular acoustic environment, the microphone array is usually deployed to suppress the noise as well as the echo from the hands-free loudspeaker, while leaving the distortion of the speech to a minimum [8]. If the environment is complicated, this problem is very difficult to be described by *a priori* models [13]. With the use of calibration signals, it is possible to design the beamformer satisfactorily [6]. In the design process, the minimum square errors and the maximum signal-to-noise plus interference power ratio are often used [7], and a multi-criteria decision problem is applied in [11] to optimize on the level of distortion, the level of noise suppression, and the level of interference suppression. Other filter design techniques (such as [1, 3, 10]) are also possible for designing beamformers.

With smart devices proliferating in the home and becoming universally networked, microphone array is often incorporated into the devices. It is getting popular to include noise reduction capability into blue-tooth headsets, mobile phones and other devices with communications or voice control functionality. On the other hand, in adapting the algorithms real time, high computational powers are required. Power consumption of portable devices become a major concern. The adaptation is impossible to be achieved inside the portable devices. Nevertheless, because of the advance in wireless technology [5, 9], it is possible to make use of the computational powers of nearby computer servers. For example, a wireless local area network (WLAN) links devices via a wireless distribution method (typically spread-spectrum or OFDM) and usually provides a connection through an access point to the wider network. This gives users the mobility to move around within a local coverage area and still be connected to the network. Wireless LANs have also become popular at home due to ease of installation and the increasing popularity of laptop computers.

The wireless transmission of multimedia signals requires a large amount of bandwidth. For transmitting such signals, lossy protocols (e.g., UDP protocol) are often used. Besides, for real time adaptation of filter coefficients, it is impossible to resend the signals again if certain packets are lost. However, for voice control systems and for human hearing system, they are usually fairly tolerant to some perturbation to the original signals. Signal restoration algorithms [2] can be applied to recovering some of the missing signals. Under this framework, it is advantageous to design the beamforming system such that small perturbation to the signals will not affect the overall quality too much. When this procedure is carried out, the original least squares or signal-to-noise ratio formulation is no longer optimal for the design. In this paper, we propose a novel configuration, where the received data is transferred wirelessly to a remote high power device, which is then used for calculating the optimal filter coefficients. Under this framework, there exists another perturbation to the signal data due to the lossy feature and signal restoration. Consequently, a total least squares technique must be used to formulate the problem. The resultant optimization problem has a special Toeplitz structure. By exploiting this property, the analytic expression for the optimal solution can be obtained. Compared with the ordinary least squares method, the performances both on SNR and segmental SNR are improved and the complexity of the algorithm is the same.

The rest of the paper is organized as follows. In Section 2, we propose a new model of broadband beamformer design and formulate it as a total least squares optimization problem. In Section 3, we show that this optimization problem can be simplified into an equivalent optimization problem. In Section 4, the optimal solution to the problem is derived analytically. Two numerical examples are illustrated in Section 5. Conclusion is given in Section 6.

## 2. The Problem.

A new model of microphone array design can be found in Figure 1, where the voice signal is received by the microphone array and processed by the filters behind the array. It is required to design the beamformer to filter out the background noise. Using a calibration signal, the filter coefficients are designed such that the error between the output signal and the source signal is minimized.

In order to adapt to the changing background noise, the beamformer filter coefficients need to be updated continuously. This will also require a very fast processor to recalculate the coefficients in real time. Assuming the device is running inside a battery-driven system, the adaptation process demands very high power consumption and will quickly drain the batteries. To overcome this difficulty, wireless communications can be established to transmit the received signals to a nearby computer server with very fast processors. Then, after updating the optimal beamforming filters, the coefficients are transmitted back to the wireless device for filtering.

During the transmission between the wireless devise and the computer server, the signal noise due to lost packets should not be neglected. Hence, the beamforming filters should be designed to minimize the background noise and the signal noise at the same time.

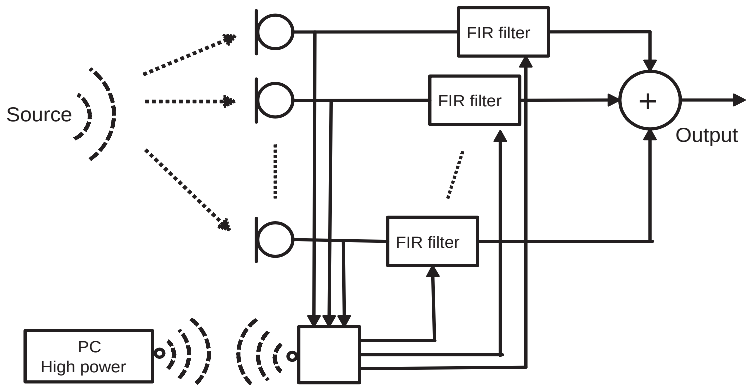


FIGURE 1. Model of the microphone array

In this model, we assume that there are $M$ elements in the microphone array. In general, the signals received by the $i$-th microphone element can be represented by

$$x_i(n) = s_i(n) + v_i(n), \quad i = 1, 2 \dots, M, \quad n = 1, 2, \dots, m, \tag{2.1}$$

where $s_i(n)$ is the source signal and $v_i(n)$ is the background noise signal.

The output of the beamformer is given by

$$y(n) = \sum_{i=1}^{M} \sum_{j=0}^{L-1} w_i(j) x_i(n-j), \quad n = 1, 2, \dots, m, \tag{2.2}$$

where $L$ is the length of the filters and $w_i(j)$, $j = 0, 1, \dots, L-1$, are the coefficients of the $i$-th FIR filter. Basically, (2.2) can be rewritten as

$$\boldsymbol{y} = \boldsymbol{H}(\boldsymbol{x})\boldsymbol{w}, \tag{2.3}$$

where

$$\boldsymbol{y} = [y(1), y(2), \dots, y(m)]^\intercal,$$
$$\boldsymbol{x} = (\boldsymbol{x}_1^\intercal, \dots, \boldsymbol{x}_M^\intercal)^\intercal, \quad \boldsymbol{x}_i = (x_i(1), \dots, x_i(m))^\intercal, \quad i = 1, \dots, M,$$
$$\boldsymbol{w} = (\boldsymbol{w}_1^\intercal, \dots, \boldsymbol{w}_M^\intercal)^\intercal, \quad \boldsymbol{w}_i = (w_i(0), \dots, w_i(L-1))^\intercal, \quad i = 1, \dots, M,$$

and $\boldsymbol{H} : \mathbb{R}^{mM} \to \mathbb{R}^{m \times ML}$ is an injective function defined by

$$\boldsymbol{H}(\boldsymbol{x}) = \begin{pmatrix} x_1(1) & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_1(L) & x_1(L-1) & \cdots & x_1(1) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(m) & x_1(m-1) & \cdots & x_1(m-L+1) \end{pmatrix} \cdots$$
$$\cdots \begin{pmatrix} x_M(1) & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_M(L) & x_M(L-1) & \cdots & x_M(1) \\ \vdots & \vdots & \ddots & \vdots \\ x_M(m) & x_M(m-1) & \cdots & x_M(m-L+1) \end{pmatrix}. \tag{2.4}$$

Denote $\boldsymbol{s}_r$ as the reference signal, which is selected from a good separate channel. Then, the least squares problem can be formulated as

$$\begin{aligned} &\min_{\boldsymbol{w}, \boldsymbol{\eta}} \ \mathcal{E}\{||\boldsymbol{\eta}||^2\} \\ &\text{s.t.} \ \ \boldsymbol{y} = \boldsymbol{H}(\boldsymbol{x})\boldsymbol{w} = \boldsymbol{s}_r + \boldsymbol{\eta}, \end{aligned} \tag{2.5}$$

where $\boldsymbol{\eta} = [\eta(1), \eta(2), \ldots, \eta(m)]^\mathsf{T}$ is the difference between the output signal and the reference signal, and $\mathcal{E}\{\cdot\}$ is the expectation operation.

However, since there exists signal noise from the wireless device to the high power device due to lost packets, we need to restore the missing signal $\boldsymbol{x}$. With some algorithms, such as interpolation method [2], we can estimate the signal $\bar{\boldsymbol{x}}$, which is used for calculating the coefficients. The estimated signal $\bar{\boldsymbol{x}}$ can be expressed by

$$\bar{\boldsymbol{x}} = \boldsymbol{x} + \boldsymbol{\varepsilon},$$

where $\boldsymbol{x}$ is the true signal and $\boldsymbol{\varepsilon}$ is the estimation noise given by

$$\boldsymbol{\varepsilon} = [\boldsymbol{\varepsilon}_1^\mathsf{T}, \boldsymbol{\varepsilon}_2^\mathsf{T}, \ldots, \boldsymbol{\varepsilon}_M^\mathsf{T}]^\mathsf{T}, \quad \boldsymbol{\varepsilon}_k = [\varepsilon_k(1), \varepsilon_k(2), \ldots, \varepsilon_k(m)]^\mathsf{T}, \quad k = 1, \ldots, M.$$

Then, the optimization problem (2.5) becomes

$$\begin{aligned} &\min_{\boldsymbol{w},\boldsymbol{\eta}} \ \mathcal{E}\{||\boldsymbol{\eta}||^2\} \\ &\text{s.t.} \ \ \bar{\boldsymbol{y}} = \boldsymbol{H}(\bar{\boldsymbol{x}})\boldsymbol{w} = \boldsymbol{s}_r + \boldsymbol{\eta}. \end{aligned} \tag{2.6}$$

Hence, since the additional estimation noise exists, the optimal coefficients obtained by (2.6) is not suitable for (2.5). Then, we need to estimate $\boldsymbol{x}$ from $\bar{\boldsymbol{x}}$ by eliminating the estimation noise. That is, the problem is formulated as to find the coefficients of FIR filters $\boldsymbol{w}$, such that

$$\boldsymbol{H}(\bar{\boldsymbol{x}} - \boldsymbol{\varepsilon})\boldsymbol{w} = \boldsymbol{H}(\bar{\boldsymbol{x}})\boldsymbol{w} - \boldsymbol{H}(\boldsymbol{\varepsilon})\boldsymbol{w} = \boldsymbol{s}_r + \boldsymbol{\eta}. \tag{2.7}$$

There are two error vectors $\boldsymbol{\varepsilon}$ and $\boldsymbol{\eta}$ to be minimized in (2.7). Since they are minimized in the same time, the vector $\boldsymbol{\varepsilon}$ should have some relations with the vector $\boldsymbol{\eta}$. That is, if $\boldsymbol{\varepsilon}$ is chosen properly and $||\boldsymbol{\varepsilon}||$ increases, then $||\boldsymbol{\eta}||$ should decreases. On the other hand, $||\boldsymbol{\varepsilon}||$ should not be very large, since it will cause aliasing of the signal and the solution will become very poor.

Since there are $M$ elements in the microphone array, there are $M$ estimation noise vectors $\boldsymbol{\varepsilon}_i$ to be minimized. To adjust the amplitude of $\boldsymbol{\varepsilon}_i$, we add a positive weight vector $\boldsymbol{c} = (c_1, \ldots, c_M)$ to $M$ channels. Then, we choose the cost function to be minimized as

$$f(\boldsymbol{w}, \boldsymbol{\varepsilon}, \boldsymbol{\eta}) = \mathcal{E}\left\{||\boldsymbol{\eta}||^2 + \sum_{i=1}^{M} c_i ||\boldsymbol{\varepsilon}_i||^2\right\}. \tag{2.8}$$

The vector $\boldsymbol{c}$ is used to adjust the amplitude $||\boldsymbol{\varepsilon}||$. That is, if $c_i$ is chosen very small, the amplitude of $\boldsymbol{\varepsilon}_i$ can be large and if $\boldsymbol{c}$ is large, the amplitude of $\boldsymbol{\varepsilon}$ must be very small. Hence, $\boldsymbol{c}$ is a weight vector to control the amplitude of signal noise.

The total least squares problem can be formulated as

**Problem 1.** *Find the coefficients of FIR filters $\boldsymbol{w}$, the noise vector $\boldsymbol{\varepsilon}$ and the vector $\boldsymbol{\eta}$, such that the cost function (2.8) is minimized, subject to the constraint (2.7).*

Problem 1 is a general optimization problem. However, $m$, the dimension of the estimation noise vector $\boldsymbol{\varepsilon}$, is always very large in practical application and the computation will be very expensive. Since the vectors $\boldsymbol{\varepsilon}$ and $\boldsymbol{\eta}$ are random, it's not necessary to find these error vectors for each sample. For this, we can transform this problem into a deterministic problem as follows.

Basically, we assume the estimation noise vector $\boldsymbol{\varepsilon}$ satisfies the following assumptions:

(A1) $\boldsymbol{\varepsilon}$ is a zero mean vector.
(A2) For each $i \neq j$, $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\varepsilon}_j$ are uncorrelated.

(A3) For each $i$, the normalized autocorrelation matrix is known from *a priori* information and is given by

$$\boldsymbol{R}_{\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i} = \begin{pmatrix} r_{\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i}(0) & r_{\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i}(1) & \cdots & r_{\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i}(L-1) \\ r_{\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i}(1) & r_{\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i}(0) & \cdots & r_{\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i}(L-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i}(L-1) & r_{\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i}(L-2) & \cdots & r_{\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i}(0) \end{pmatrix},$$

where $r_{\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i}(k)$ is normalized by

$$r_{\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i}(k) = \mathcal{E}\left\{\sum_{n=k+1}^{m}(\boldsymbol{\varepsilon}_i(n-k))(\boldsymbol{\varepsilon}_i(n))\right\}/\mathcal{E}\{||\boldsymbol{\varepsilon}_i||^2\}, \text{ if } k \geq 0.$$

(A4) The intensities of $\boldsymbol{\varepsilon}_i$ in different channels are proportional, that is,

$$\mathcal{E}\{||\boldsymbol{\varepsilon}_1||^2\} : \mathcal{E}\{||\boldsymbol{\varepsilon}_2||^2\} : \cdots : \mathcal{E}\{||\boldsymbol{\varepsilon}_M||^2\} = d_1 : d_2 : \cdots : d_M,$$

where $\{d_i, i = 1, \ldots, M\}$ are given positive constants.

For the assumption (A1), the average value of the noise is always zero for interpolation methods. The assumption (A2) states that the noise data of different channels are mutually independent. This is reasonable in practice. For the assumption (A3), $\boldsymbol{\varepsilon}_i$ can be treated as a white noise vector in general, and then $\boldsymbol{R}_{\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i}$ is the identity matrix. The assumption (A4) is motivated by the setting of the microphones, where the noises of some channels are high, while the noises of other channels are small.

Then, by the assumption (A2), we can obtain the covariance matrix of $\boldsymbol{\varepsilon}$ as

$$\begin{aligned}\boldsymbol{R}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} =& \mathcal{E}\{(\boldsymbol{H}(\boldsymbol{\varepsilon}))^{\intercal}\boldsymbol{H}(\boldsymbol{\varepsilon})\} \\ =& \begin{pmatrix} \mathcal{E}\{||\boldsymbol{\varepsilon}_1||^2\}\boldsymbol{R}_{\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_1} & 0 & \cdots & 0 \\ 0 & \mathcal{E}\{||\boldsymbol{\varepsilon}_2||^2\}\boldsymbol{R}_{\boldsymbol{\varepsilon}_2\boldsymbol{\varepsilon}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathcal{E}\{||\boldsymbol{\varepsilon}_M||^2\}\boldsymbol{R}_{\boldsymbol{\varepsilon}_M\boldsymbol{\varepsilon}_M} \end{pmatrix}.\end{aligned}$$

Since the intensities of $\boldsymbol{\varepsilon}_i$ in different channels are proportional, we can choose to minimize the noise error of a reference channel. For this, suppose that the noise vector of the reference channel is $\boldsymbol{\varepsilon}_r$. Then, we rewrite the second term of (2.8) as

$$\sum_{i=1}^{M} c_i||\boldsymbol{\varepsilon}_i||^2 = ||\boldsymbol{\varepsilon}_r||^2 \sum_{i=1}^{M} c_i||\boldsymbol{\varepsilon}_i||^2/||\boldsymbol{\varepsilon}_r||^2 = ||\boldsymbol{\varepsilon}_r||^2 \sum_{i=1}^{M} c_i d_i/d_r.$$

Let $c_r = \sum_{i=1}^{M} c_i d_i/d_r$. The cost function (2.8) is equivalent to

$$f(\boldsymbol{w}, \boldsymbol{\varepsilon}, \boldsymbol{\eta}) = \mathcal{E}\{||\boldsymbol{\eta}||^2 + c_r||\boldsymbol{\varepsilon}_r||^2\}. \tag{2.9}$$

Since $\boldsymbol{\eta}$ is a random vector, the cost function is required to be transformed. To remove these random vectors, we denote $\boldsymbol{p} = \boldsymbol{H}(\boldsymbol{x})\boldsymbol{w} - \boldsymbol{s}_r$, then

$$\boldsymbol{\eta} = \boldsymbol{p} - \boldsymbol{H}(\varepsilon)\boldsymbol{w}. \tag{2.10}$$

Let $\boldsymbol{H}(\boldsymbol{\varepsilon})$ be denoted by

$$\boldsymbol{H}(\boldsymbol{\varepsilon}) = [\boldsymbol{H}_{\boldsymbol{\varepsilon}1}, \boldsymbol{H}_{\boldsymbol{\varepsilon}2}, \ldots, \boldsymbol{H}_{\boldsymbol{\varepsilon}M}],$$

where for each $k = 1, \ldots, M$, $\boldsymbol{H}_{\boldsymbol{\varepsilon}k} \in \mathbb{R}^{m\times L}$ is a sub-matrix in $\boldsymbol{H}(\boldsymbol{\varepsilon})$. Furthermore, let $\boldsymbol{H}_{\boldsymbol{\varepsilon}k}$ be denoted by

$$\boldsymbol{H}_{\boldsymbol{\varepsilon}k} = [\boldsymbol{h}_{\boldsymbol{\varepsilon}k}(0), \boldsymbol{h}_{\boldsymbol{\varepsilon}k}(1), \ldots, \boldsymbol{h}_{\boldsymbol{\varepsilon}k}(L-1)],$$

where for each $j = 0, \dots, L-1$, $\boldsymbol{h}_{\boldsymbol{\varepsilon}k}(j)$ is the corresponding column in $\boldsymbol{H}_{\boldsymbol{\varepsilon}k}$. Then, (2.10) can be rewritten as

$$\boldsymbol{\eta} = \boldsymbol{p} - \sum_{i=1}^{M}\sum_{j=0}^{L-1} \boldsymbol{h}_{\boldsymbol{\varepsilon}i}(j) w_i(j).$$

Next, since the estimation noise exists and is not zero, we have $||\boldsymbol{\varepsilon}_i|| > 0$, $i = 1, 2, \dots, M$. Then, we express $\boldsymbol{p}$ in the subspace spanned by the vectors $\{\boldsymbol{h}_{\boldsymbol{\varepsilon}i}(j), i = 1, \dots, M, j = 0, \dots, L-1\}$ as follows.

$$\boldsymbol{p} = \left(\sum_{k=1}^{M}\sum_{l=0}^{L-1} \alpha_{kl} \boldsymbol{h}_{\boldsymbol{\varepsilon}k}(l)\right) + \bar{\boldsymbol{h}} = \boldsymbol{H}(\boldsymbol{\varepsilon})\boldsymbol{\alpha} + \bar{\boldsymbol{h}}, \tag{2.11}$$

where

$$\boldsymbol{\alpha} = (\alpha_{10}, \dots, \alpha_{1,L-1}, \alpha_{20}, \dots, \alpha_{2,L-1}, \dots, \alpha_{M0}, \dots, \alpha_{M,L-1})^{\mathsf{T}},$$

and $\bar{\boldsymbol{h}}$ is a vector orthogonal to the subspace spanned by the vectors $\{\boldsymbol{h}_{\boldsymbol{\varepsilon}i}(j), i = 1, \dots, M, j = 0, \dots, L-1\}$.

By (2.11), we have

$$\begin{aligned}
||\boldsymbol{\eta}||^2 &= (\boldsymbol{p} - \boldsymbol{H}(\boldsymbol{\varepsilon})\boldsymbol{w})^{\mathsf{T}}(\boldsymbol{p} - \boldsymbol{H}(\boldsymbol{\varepsilon})\boldsymbol{w}) \\
&= ||\boldsymbol{p}||^2 - 2\boldsymbol{w}^{\mathsf{T}}(\boldsymbol{H}(\boldsymbol{\varepsilon}))^{\mathsf{T}}\boldsymbol{p} + \boldsymbol{w}^{\mathsf{T}}(\boldsymbol{H}(\boldsymbol{\varepsilon}))^{\mathsf{T}}\boldsymbol{H}(\boldsymbol{\varepsilon})\boldsymbol{w} \\
&= \boldsymbol{\alpha}^{\mathsf{T}}(\boldsymbol{H}(\boldsymbol{\varepsilon}))^{\mathsf{T}}\boldsymbol{H}(\boldsymbol{\varepsilon})\boldsymbol{\alpha} + ||\bar{\boldsymbol{h}}||^2 - 2\boldsymbol{w}^{\mathsf{T}}(\boldsymbol{H}(\boldsymbol{\varepsilon}))^{\mathsf{T}}\boldsymbol{H}(\boldsymbol{\varepsilon})\boldsymbol{\alpha} + \boldsymbol{w}^{\mathsf{T}}(\boldsymbol{H}(\boldsymbol{\varepsilon}))^{\mathsf{T}}\boldsymbol{H}(\boldsymbol{\varepsilon})\boldsymbol{w} \\
&= (\boldsymbol{w} - \boldsymbol{\alpha})^{\mathsf{T}}(\boldsymbol{H}(\boldsymbol{\varepsilon}))^{\mathsf{T}}\boldsymbol{H}(\boldsymbol{\varepsilon})(\boldsymbol{w} - \boldsymbol{\alpha}) + ||\bar{\boldsymbol{h}}||^2.
\end{aligned}$$

Then, we have

$$\mathcal{E}\{||\boldsymbol{\eta}||^2\} = (\boldsymbol{w} - \boldsymbol{\alpha})^{\mathsf{T}}\boldsymbol{R}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}(\boldsymbol{w} - \boldsymbol{\alpha}) + \beta, \tag{2.12}$$

where $\beta = \mathcal{E}\{||\bar{\boldsymbol{h}}||^2\} \geq 0$.

Define $\bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} = \boldsymbol{R}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}/||\boldsymbol{\varepsilon}_r||^2$. Then, (2.12) becomes

$$\mathcal{E}\{||\boldsymbol{\eta}||^2\} = \mathcal{E}\{||\boldsymbol{\varepsilon}_r||^2\}(\boldsymbol{w} - \boldsymbol{\alpha})^{\mathsf{T}}\bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}(\boldsymbol{w} - \boldsymbol{\alpha}) + \beta, \tag{2.13}$$

and the cost function (2.8) becomes

$$\begin{aligned}
f(\boldsymbol{w}, \boldsymbol{\varepsilon}, \boldsymbol{\eta}) &= \mathcal{E}\{||\boldsymbol{\varepsilon}_r||^2\}(\boldsymbol{w} - \boldsymbol{\alpha})^{\mathsf{T}}\bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}(\boldsymbol{w} - \boldsymbol{\alpha}) + \beta + c_r\mathcal{E}\{||\boldsymbol{\varepsilon}_r||^2\} \\
&= \mathcal{E}\{||\boldsymbol{\varepsilon}_r||^2\}\left((\boldsymbol{w} - \boldsymbol{\alpha})^{\mathsf{T}}\bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}(\boldsymbol{w} - \boldsymbol{\alpha}) + c_r\right) + \beta \\
&= \bar{f}(\boldsymbol{w}, \boldsymbol{\alpha}, \beta, ||\boldsymbol{\varepsilon}_r||^2).
\end{aligned} \tag{2.14}$$

Furthermore, by simplifying

$$\mathcal{E}\{||\boldsymbol{H}(\boldsymbol{x})\boldsymbol{w} - \boldsymbol{s}_r||^2\} = \mathcal{E}\{||\boldsymbol{p}||^2\} = \mathcal{E}\{||\boldsymbol{H}(\boldsymbol{\varepsilon})\boldsymbol{\alpha} + \bar{\boldsymbol{h}}||^2\},$$

we have

$$\boldsymbol{w}^{\mathsf{T}}\boldsymbol{R}_{xx}\boldsymbol{w} - 2\boldsymbol{w}^{\mathsf{T}}\boldsymbol{r}_s + \boldsymbol{r}_c = \mathcal{E}\{||\boldsymbol{\varepsilon}_r||^2\}\boldsymbol{\alpha}^{\mathsf{T}}\bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}\boldsymbol{\alpha} + \beta, \tag{2.15}$$

where $\boldsymbol{r}_c = \mathcal{E}\{||\boldsymbol{s}_r||^2\}$ and $\boldsymbol{R}_{xx}$ is defined by

$$\boldsymbol{R}_{xx} = \mathcal{E}\{(\boldsymbol{H}(\boldsymbol{x}))^{\mathsf{T}}\boldsymbol{H}(\boldsymbol{x})\}, \tag{2.16}$$

and similarly has the structure of

$$\boldsymbol{R}_{xx} = \begin{pmatrix} \boldsymbol{R}_{x_1x_1} & \boldsymbol{R}_{x_1x_2} & \cdots & \boldsymbol{R}_{x_1x_M} \\ \boldsymbol{R}_{x_2x_1} & \boldsymbol{R}_{x_2x_2} & \cdots & \boldsymbol{R}_{x_2x_M} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{R}_{x_Mx_1} & \boldsymbol{R}_{x_Mx_2} & \cdots & \boldsymbol{R}_{x_Mx_M} \end{pmatrix},$$

where

$$\boldsymbol{R}_{x_i x_j} = \begin{pmatrix} r_{x_i x_j}(0) & r_{x_i x_j}(-1) & \cdots & r_{x_i x_j}(1-L) \\ r_{x_i x_j}(1) & r_{x_i x_j}(0) & \cdots & r_{x_i x_j}(2-L) \\ \vdots & \vdots & \ddots & \vdots \\ r_{x_i x_j}(L-1) & r_{x_i x_j}(L-2) & \cdots & r_{x_i x_j}(0) \end{pmatrix},$$

$$r_{x_i x_j}(k) = \begin{cases} \sum\limits_{n=k+1}^{m} \mathcal{E}\{x_i(n-k)x_j(n)\} & \text{if } k \geq 0 \\ \sum\limits_{n=1}^{m+k} \mathcal{E}\{x_i(n-k)x_j(n)\} & \text{if } k \leq 0 \end{cases},$$

and

$$\boldsymbol{r}_s = (\boldsymbol{r}_1^\top, \boldsymbol{r}_2^\top, \ldots, \boldsymbol{r}_M^\top)^\top, \quad \boldsymbol{r}_i = (r_i(0), r_i(1), \ldots, r_i(L-1))^\top, \quad i = 1, \ldots, M,$$

with

$$r_i(k) = \sum_{j=1}^{m-k} \mathcal{E}\{x_i(n)s_r(n+k)\}.$$

Then, by (2.15), we have

$$\mathcal{E}\{||\boldsymbol{\varepsilon}_r||^2\} = \frac{\boldsymbol{w}^\top \boldsymbol{R}_{xx}\boldsymbol{w} - 2\boldsymbol{w}^\top \boldsymbol{r}_s + \boldsymbol{r}_c - \beta}{\boldsymbol{\alpha}^\top \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}}\boldsymbol{\alpha}},$$

and (2.14) becomes

$$\hat{f}(\boldsymbol{w}, \boldsymbol{\alpha}, \beta) = \frac{\boldsymbol{w}^\top \boldsymbol{R}_{xx}\boldsymbol{w} - 2\boldsymbol{w}^\top \boldsymbol{r}_s + \boldsymbol{r}_c - \beta}{\boldsymbol{\alpha}^\top \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}}\boldsymbol{\alpha}} \left((\boldsymbol{w}-\boldsymbol{\alpha})^\top \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}}(\boldsymbol{w}-\boldsymbol{\alpha}) + c_r\right) + \beta. \tag{2.17}$$

Hence, we transform Problem 1 into a deterministic problem as

**Problem 2.** *Find the coefficients $\boldsymbol{w}$, $\boldsymbol{\alpha}$, and $\beta \geq 0$, such that the cost function (2.17) is minimized.*

3. **Method.** Problem 2 can be solved by any optimization methods. However, we can simplify this problem by decreasing the number of variables as follows.

In general, the matrix $\boldsymbol{R}_{\boldsymbol{\varepsilon\varepsilon}}$ or $\bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}}$ are always nonsingular in real applications. Then, we have the lemma as follows.

**Lemma 3.1.** *For a given $\boldsymbol{w} \neq \boldsymbol{0}$ and $\beta$, the optimal $\boldsymbol{\alpha}$ to minimize the cost function (2.17) is given by $\boldsymbol{\alpha} = b \cdot \boldsymbol{w}$, where*

$$b = 1 + c_r / \boldsymbol{w}^\top \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}}\boldsymbol{w}. \tag{3.1}$$

**Proof:** For a given $\boldsymbol{w} \neq \boldsymbol{0}$ and $\beta$, the minimization of the function $\hat{f}(\boldsymbol{w}, \boldsymbol{\alpha}, \beta)$ is equivalent to the minimization of the function

$$g(\boldsymbol{\alpha}) = \left((\boldsymbol{w}-\boldsymbol{\alpha})^\top \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}}(\boldsymbol{w}-\boldsymbol{\alpha}) + c_r\right) / \boldsymbol{\alpha}^\top \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}}\boldsymbol{\alpha}. \tag{3.2}$$

Denote $\bar{\boldsymbol{w}} = (\bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}})^{1/2}\boldsymbol{w}$ and $\bar{\boldsymbol{\alpha}} = (\bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}})^{1/2}\boldsymbol{\alpha}$, (3.2) becomes

$$\bar{g}(\bar{\boldsymbol{\alpha}}) = \left(||\bar{\boldsymbol{w}} - \bar{\boldsymbol{\alpha}}||^2 + c_r\right) / ||\bar{\boldsymbol{\alpha}}||^2. \tag{3.3}$$

Then, the first order necessary condition of $\bar{g}(\bar{\boldsymbol{\alpha}})$ is given by

$$\frac{\partial \bar{g}(\bar{\boldsymbol{\alpha}})}{\partial \bar{\alpha}_i} = \frac{2(\bar{\alpha}_i - \bar{w}_i)||\bar{\boldsymbol{\alpha}}||^2 - 2\bar{\alpha}_i(||\bar{\boldsymbol{w}} - \bar{\boldsymbol{\alpha}}||^2 + c_r)}{||\bar{\boldsymbol{\alpha}}||^4} = 0, \quad i = 1, \ldots, ML.$$

Since the optimal $\bar{\boldsymbol{\alpha}}$ cannot be zero, we obtain

$$(\bar{\alpha}_i - \bar{w}_i)||\bar{\boldsymbol{\alpha}}||^2 = \bar{\alpha}_i(||\bar{\boldsymbol{w}} - \bar{\boldsymbol{\alpha}}||^2 + c_r), \quad i = 1, \ldots, ML. \tag{3.4}$$

Multiplied by $\bar{\alpha}_i$ on both sides of (3.4) and sum up from $i = 1$ to $ML$, we have

$$\left(||\bar{\boldsymbol{\alpha}}||^2 - \sum_{i=1}^{ML} \bar{\alpha}_i \bar{w}_i\right) ||\bar{\boldsymbol{\alpha}}||^2 = ||\bar{\boldsymbol{\alpha}}||^2 \left(||\bar{\boldsymbol{w}} - \bar{\boldsymbol{\alpha}}||^2 + c_r\right).$$

It can be simplified as

$$||\bar{\boldsymbol{\alpha}}||^2 - \sum_{i=1}^{ML} \bar{\alpha}_i \bar{w}_i = ||\bar{\boldsymbol{w}} - \bar{\boldsymbol{\alpha}}||^2 + c_r. \tag{3.5}$$

Multiplied by $\bar{w}_i$ on both sides of (3.4) and sum up from $i = 1$ to $ML$, we have

$$\left(\sum_{i=1}^{ML} \bar{\alpha}_i \bar{w}_i - ||\bar{\boldsymbol{w}}||^2\right) ||\bar{\boldsymbol{\alpha}}||^2 = \left(||\bar{\boldsymbol{w}} - \bar{\boldsymbol{\alpha}}||^2 + c_r\right) \sum_{i=1}^{ML} \bar{\alpha}_i \bar{w}_i. \tag{3.6}$$

Substituting (3.5) into (3.6) yields

$$\left(\sum_{i=1}^{ML} \bar{\alpha}_i \bar{w}_i - ||\bar{\boldsymbol{w}}||^2\right) ||\bar{\boldsymbol{\alpha}}||^2 = \left(||\bar{\boldsymbol{\alpha}}||^2 - \sum_{i=1}^{ML} \bar{\alpha}_i \bar{w}_i\right) \sum_{i=1}^{ML} \bar{\alpha}_i \bar{w}_i.$$

Then, we obtain

$$\left|\sum_{i=1}^{ML} \bar{\alpha}_i \bar{w}_i\right| = ||\bar{\boldsymbol{w}}|| \cdot ||\bar{\boldsymbol{\alpha}}|| = \left(\sum_{i=1}^{ML} \bar{\alpha}_i^2\right)^{1/2} \left(\sum_{i=1}^{ML} \bar{w}_i^2\right)^{1/2}. \tag{3.7}$$

Since $\boldsymbol{w} \neq \mathbf{0}$ and $\bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}}$ is nonsingular, we have $\bar{\boldsymbol{w}} \neq \mathbf{0}$. Then, the equal sign of (3.7) is true if and only if there is a constant $b$ such that $\bar{\boldsymbol{\alpha}} = b \cdot \bar{\boldsymbol{w}}$. Without lost of generality, we suppose that $\bar{w}_i \neq 0$. Then, (3.4) becomes

$$(b-1)b^2 \bar{w}_i ||\bar{\boldsymbol{w}}||^2 = b\bar{w}_i \left((b-1)^2 ||\bar{\boldsymbol{w}}||^2 + c_r\right), \quad i = 1, \ldots, ML.$$

It can be simplified as

$$b = 1 + c_r/||\bar{\boldsymbol{w}}||^2 = 1 + c_r/\boldsymbol{w}^\intercal \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}} \boldsymbol{w}.$$

Furthermore, since $\bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}}$ is nonsingular, it follows from $\bar{\boldsymbol{\alpha}} = b \cdot \bar{\boldsymbol{w}}$ that $\boldsymbol{\alpha} = b \cdot \boldsymbol{w}$. This completes the proof.

Hence, by Lemma 3.1, we only need to minimize the cost function

$$\begin{aligned}
\tilde{f}(\boldsymbol{w}, \beta) &= \hat{f}(\boldsymbol{w}, \boldsymbol{\alpha}^*, \beta) \\
&= \frac{\boldsymbol{w}^\intercal \boldsymbol{R}_{xx} \boldsymbol{w} - 2\boldsymbol{w}^\intercal \boldsymbol{r}_s + \boldsymbol{r}_c - \beta}{\left(1 + \frac{c_r}{\boldsymbol{w}^\intercal \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}} \boldsymbol{w}}\right)^2 \boldsymbol{w}^\intercal \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}} \boldsymbol{w}} \left(\left(\frac{c_r}{\boldsymbol{w}^\intercal \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}} \boldsymbol{w}}\right)^2 \boldsymbol{w}^\intercal \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}} \boldsymbol{w} + c_r\right) + \beta \\
&= (\boldsymbol{w}^\intercal \boldsymbol{R}_{xx} \boldsymbol{w} - 2\boldsymbol{w}^\intercal \boldsymbol{r}_s + \boldsymbol{r}_c) \frac{c_r}{c_r + \boldsymbol{w}^\intercal \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}} \boldsymbol{w}} + \frac{\beta \boldsymbol{w}^\intercal \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}} \boldsymbol{w}}{c_r + \boldsymbol{w}^\intercal \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}} \boldsymbol{w}}.
\end{aligned} \tag{3.8}$$

Since $\beta \geq 0$ and $\boldsymbol{w}^\intercal \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}} \boldsymbol{w}/(c_r + \boldsymbol{w}^\intercal \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}} \boldsymbol{w}) \geq 0$, the optimal $\beta$ is $\beta^* = 0$ and we have

$$\tilde{f}(\boldsymbol{w}, \beta^*) = (\boldsymbol{w}^\intercal \boldsymbol{R}_{xx} \boldsymbol{w} - 2\boldsymbol{w}^\intercal \boldsymbol{r}_s + \boldsymbol{r}_c) \frac{c_r}{c_r + \boldsymbol{w}^\intercal \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}} \boldsymbol{w}}.$$

Thus, Problem 2 is equivalent to a simplified problem as

**Problem 3.** *Find the coefficients* $\mathbf{w}$*, such that the cost function*

$$F(\boldsymbol{w}) = \frac{c_r(\boldsymbol{w}^\intercal \boldsymbol{R}_{xx} \boldsymbol{w} - 2\boldsymbol{w}^\intercal \boldsymbol{r}_s + \boldsymbol{r}_c)}{c_r + \boldsymbol{w}^\intercal \bar{\boldsymbol{R}}_{\boldsymbol{\varepsilon\varepsilon}} \boldsymbol{w}} \tag{3.9}$$

*is minimized.*

Problem 3 is an optimization problem and can be solved by any gradient based optimization algorithms. Basically, we can solve this problem very easily by some basic matrix computations. For this, we denote

$$\hat{\boldsymbol{w}} = [-1\ \boldsymbol{w}^\intercal]^\intercal, \quad \hat{\boldsymbol{H}}(\boldsymbol{x}) = [\boldsymbol{s}_r\ \boldsymbol{H}(\boldsymbol{x})],$$
$$\hat{\boldsymbol{R}}_{xx} = \mathcal{E}\{(\hat{\boldsymbol{H}}(\boldsymbol{x}))^\intercal \hat{\boldsymbol{H}}(\boldsymbol{x})\}, \quad \hat{\boldsymbol{R}}_{\varepsilon\varepsilon} = \begin{pmatrix} c_r & 0 \\ 0 & \bar{\boldsymbol{R}}_{\varepsilon\varepsilon} \end{pmatrix}.$$

Then, by ignoring the multiplier $c_r$, the objective function (3.9) can be rewritten as a simplified form:

$$\hat{F}(\hat{\boldsymbol{w}}) = \frac{\hat{\boldsymbol{w}}^\intercal \hat{\boldsymbol{R}}_{xx} \hat{\boldsymbol{w}}}{\hat{\boldsymbol{w}}^\intercal \hat{\boldsymbol{R}}_{\varepsilon\varepsilon} \hat{\boldsymbol{w}}}. \tag{3.10}$$

To minimize the cost function (3.10), we first do the Cholesky factorization of $\hat{\boldsymbol{R}}_{\varepsilon\varepsilon}$ as

$$\hat{\boldsymbol{R}}_{\varepsilon\varepsilon} = \boldsymbol{U}^\intercal \boldsymbol{U}, \tag{3.11}$$

where $\boldsymbol{U}$ is an upper triangular matrix. Denote

$$\tilde{\boldsymbol{w}} = \boldsymbol{U}\hat{\boldsymbol{w}}, \quad \tilde{\boldsymbol{R}}_{xx} = (\boldsymbol{U}^{-1})^\intercal \hat{\boldsymbol{R}}_{xx} \boldsymbol{U}^{-1},$$

then (3.10) is transformed into

$$\tilde{F}(\tilde{\boldsymbol{w}}) = \frac{\tilde{\boldsymbol{w}}^\intercal \tilde{\boldsymbol{R}}_{xx} \tilde{\boldsymbol{w}}}{||\tilde{\boldsymbol{w}}||^2}. \tag{3.12}$$

Considering the cost function (3.12), the optimal value $\tilde{F}^*$ is given by the minimal eigenvalue of $\tilde{\boldsymbol{R}}_{xx}$ and the optimal $\tilde{\boldsymbol{w}}^*$ is given by the corresponding eigenvector. Then, the optimal $\hat{\boldsymbol{w}}^*$ and $\boldsymbol{w}^*$ are given by

$$\hat{\boldsymbol{w}}^* = \boldsymbol{U}^{-1}\tilde{\boldsymbol{w}}^*, \quad \begin{pmatrix} -1 \\ \boldsymbol{w}^* \end{pmatrix} = -\frac{\tilde{\boldsymbol{w}}^*}{\tilde{\boldsymbol{w}}_1^*},$$

where $\tilde{\boldsymbol{w}}_1^*$ is the first element of $\tilde{\boldsymbol{w}}^*$.

Hence, we have solved this problem with some basic matrix transformations and it has a very low complexity of computation.

4. **Choice of Coefficient.** From Problem 2 and Problem 3, we can see that these problems are formulated with a given weight coefficient $c_r > 0$. However, what is the value of a suitable coefficient $c_r$? How to choose this value?

Basically, the coefficient $c_r$ varies according to the estimation noise level. If the intensity of the estimation noise changes, the coefficient $c_r$ should also be adjusted. From the cost function (2.9), we can see that when $c_r$ becomes larger, the weight of $\boldsymbol{\varepsilon}_r$ increases in (2.9). Then, the vector $\boldsymbol{\varepsilon}_r$ should not be too large and $||\boldsymbol{\varepsilon}_r||^2$ becomes smaller. If $c_r \to +\infty$, it means that $||\boldsymbol{\varepsilon}_r||^2$ will approach to zero. Then, the total least squares problem is equivalent to ordinary least squares problem. This can also be seen from the limit of the cost function (3.9) as $c_r \to +\infty$, that is,

$$\lim_{c_r\to\infty} F(\boldsymbol{w}) = \lim_{c_r\to\infty} \frac{c_r(\boldsymbol{w}^\intercal \boldsymbol{R}_{xx}\boldsymbol{w} - 2\boldsymbol{w}^\intercal \boldsymbol{r}_s + \boldsymbol{r}_c)}{c_r + \boldsymbol{w}^\intercal \bar{\boldsymbol{R}}_{\varepsilon\varepsilon}\boldsymbol{w}} = \boldsymbol{w}^\intercal \boldsymbol{R}_{xx}\boldsymbol{w} - 2\boldsymbol{w}^\intercal \boldsymbol{r}_s + \boldsymbol{r}_c.$$

On the other hand, if $c_r \to 0$, the weight of $\boldsymbol{\varepsilon}_r$ decreases in (2.9) and $||\boldsymbol{\varepsilon}_r||^2$ can become larger, while the other term $||\boldsymbol{\eta}||^2 \to 0$.

Hence, the choice of a suitable $c_r$ depends on the intensity of the estimation noise. It should never be a large number or it will approach to the ordinary least squares formulation. Also, it should not be too small, or the fluctuation of $||\boldsymbol{\varepsilon}_r||^2$ will change excessively and the solution may not be optimal. Then, a suitable $c_r$ is given such that the fluctuation of $||\boldsymbol{\varepsilon}_r||^2$ is of similar magnitude to the intensity of the estimation noise. This can be seen from the examples in the next section.

Basically, a best choice of $c_r$ can be obtained from *a priori* information with calibration data, that is, we can check the performance of many $c_r$ and choose the best one. Furthermore, whether the choice of $c_r$ is good depends on the criteria of performance. For different criteria of performance, the best $c_r$ can be different. That is, if the criterion of performance with respect to $\boldsymbol{w}$ is $g(\boldsymbol{w})$, the best coefficient $c_r$ can be optimized as

$$c_r^* = \min_{c_r} g(\boldsymbol{w}^*(c_r)),$$

where $\boldsymbol{w}^*(c_r)$ denotes the solution of Problem 3. Furthermore, the $c_r^*$ value also changes if the estimation noise changes. This relationship can be seen in the next section.

5. **Illustrative Examples.** In this section, we apply the model to two examples, where calibration signals are collected for the speech signal and the background noise. In the first example, the signals are collected in a Volvo station wagon environment with a multi-channel microphone array in a hands-free situation. An artificial talker was mounted in the passenger seat to simulate a real person leading a conversation. The desired sound source signal was created from a speaker's sound sequence in a non-moving car with the engine turned off. The background noise signal was created when the car was driving at a speed of 110km/h. In the second example, the signals are collected in an anechoic chamber where the speech signal and the noise signal are recorded separately.

In order to simulate the lossy nature in the wireless transmission of audio data, we artificially add random noise to the calibration signals resembling the situation that part of the signals are loss and restored mathematically. The proposed total least squares method is used and is compared with the LS method described in [7] by designing FIR filters with a filter length of 16. The computation is implemented in Matlab.

In the first example, the acoustic data was sampled with the rate of 12KHz and the duration of these signals was 4 seconds. The initial SNR values for two channels are both 2dB and the corresponding initial segmental SNR values for two channels are $-4.5812$dB and $-4.5808$dB. The SNR and segmental SNR values obtained by LS method and TLS method are given in Figure 2. If the SNR criterion is used, the best coefficient $c_r^*$ is 0.0380, where the SNR value can achieve 8.0945dB. If the segmental SNR criterion is used, the best coefficient $c_r^*$ is 0.0380, where the segmental SNR value can achieve $-1.7212$dB.

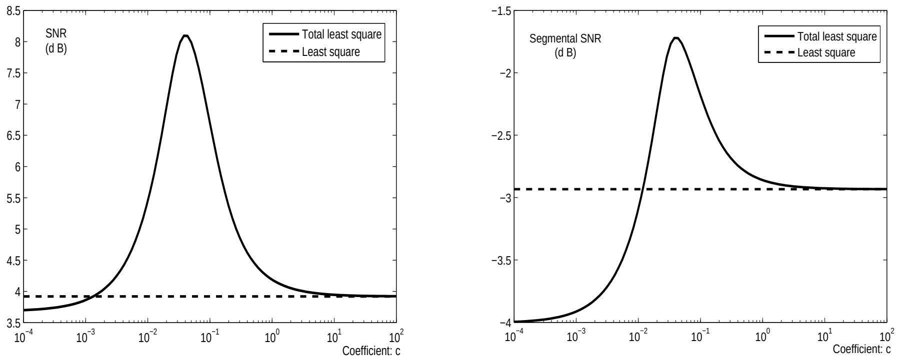


FIGURE 2. SNR and segmental SNR in Example 1

For the relation between the best coefficient $c_r^*$ and the estimation noise, we calculate the best coefficient $c_r^*$ for different scales of $\boldsymbol{\varepsilon}_r$. This relationship can be seen in Figure

3, where the SNR and segmental SNR criteria are used, respectively. It can be seen that the best coefficient $c_r^*$ decreases when the intensity of the estimation noise increases.

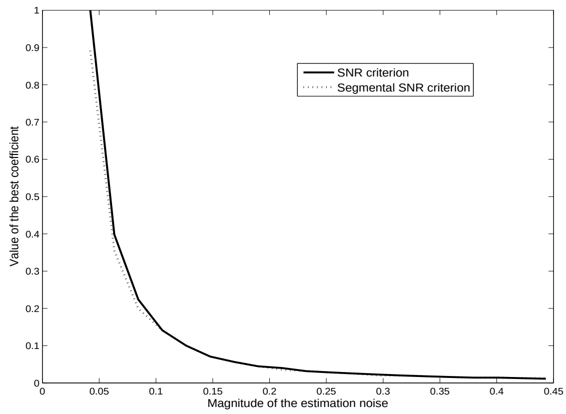


FIGURE 3. Relation between $||\boldsymbol{\varepsilon}_r||$ and the best coefficient $c_r$

In the second example, the acoustic data was sampled with the rate of 8KHz and the duration of these signals was 6 seconds. The initial SNR values for two channels are both 5dB and the corresponding initial segmental SNR values for two channels are $-0.6490$dB and $-0.6205$dB. The SNR and segmental SNR values obtained by LS method and TLS method are given in Figure 4. If the SNR criterion is used, the best coefficient $c_r^*$ is 0.0224, where the SNR value can achieve 10.0737dB. If the segmental SNR criterion is used, the best coefficient $c_r^*$ is 0.0355, where the segmental SNR value can achieve 3.3497dB.

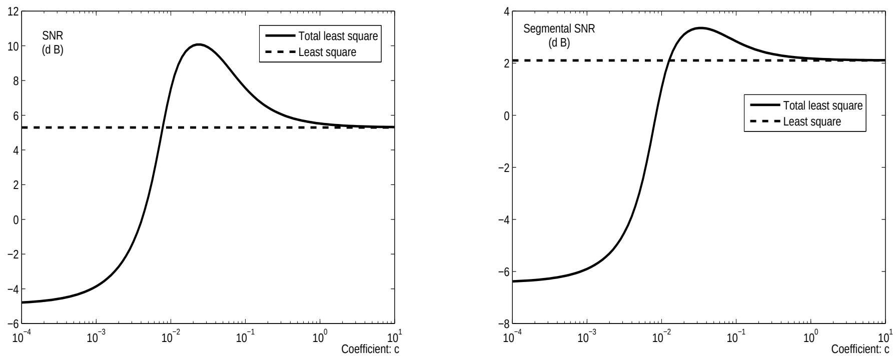


FIGURE 4. SNR and segmental SNR in Example 2

For the relation between the best coefficient $c_r^*$ and the estimation noise, we calculate the best coefficient $c_r^*$ for different scales of $\boldsymbol{\varepsilon}_r$. This relationship can be seen in Figure 5, where the SNR and segmental SNR criteria are used, respectively. It can be seen that the best coefficient $c_r^*$ decreases when the intensity of the estimation noise increases.

6. **Conclusion.** A novel total least squares design of a broadband beamforming system has been proposed. This problem is formulated as an optimization problem and shown to be equivalent to a simplified one. The analytic expression for the optimal solution can be derived. Compared with the ordinary LS method, this method can yield improved performance both on SNR and segmental SNR with a similar computation complexity.

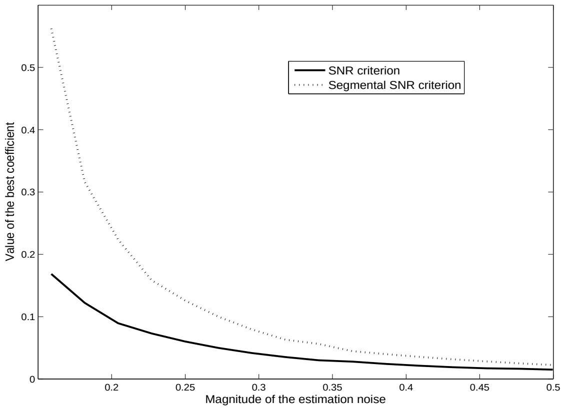


FIGURE 5. Relation between $||\varepsilon_r||$ and the best coefficient $c_r^*$

**Acknowledgement.** This paper is supported by RGC Grant PolyU. (5365/09E). The first and second authors would also like to thank the support of NSFC Grant (10901170) and the Research Committee of the Hong Kong Polytechnic University.

## REFERENCES

[1] Z. G. Feng and K. L. Teo, A discrete filled function method for the design of FIR filters with signed-powers-of-two coefficients, *IEEE Trans. on Signal Processing*, vol.56, pp.134-139, 2008.

[2] S. J. Godsill and P. J. W. Rayner, *Digital Audio Restoration: A Statistical Model Based Approach*, Springer, 1998.

[3] C. Y. F. Ho, B. W. K. Ling, Y. Q. Liu, P. K. S. Tam and K. L. Teo, Efficient algorithm for solving semi-infinite programming problems and their applications to nonuniform filter bank designs, *IEEE Trans. on Signal Processing*, vol.54, pp.4223-4232, 2006.

[4] C. Kyriakakis, P. Tsakalides and T. Holman, Surrounded by sound, *IEEE Signal Processing Magazine*, pp.55-66, 1999.

[5] I.-C. Lin, P.-Y. Chang and C.-C. Chang, A key management scheme for sensor networks using bilinear pairings and gap Diffie-Hellman group, *International Journal of Innovative Computing, Information and Control*, vol.6, no.2, pp.809-816, 2010.

[6] S. Nordholm, I. Claesson and M. Dahl, Adaptive microphone array employing calibration signals: An analytical evaluation, *IEEE Trans. on Speech Audio Processing*, vol.7, pp.241-252, 1999.

[7] S. Nordholm, I. Claesson and N. Grbić, Optimal and adaptive microphone arrays for speech input in automobiles, in *Microphone Arrays*, M. Brandstein and D. Ward (eds.), Berlin, Springer Verlag, 2001.

[8] B. D. van Veen and K. M. Buckley, Beamforming: A versatile approach to spatial filtering, *IEEE ASSP Magazine*, pp.4-24, 1988.

[9] L.-C. Shiu, The robot deployment scheme for wireless sensor networks, *International Journal of Innovative Computing, Information and Control*, vol.6, no.7, pp.2941-2953, 2010.

[10] K. F. C. Yiu, N. Grbić, S. Nordholm and K. L. Teo, A hybrid method for the design of oversampled uniform DFT filter banks, *Signal Processing*, vol.86, pp.1355-1364, 2006.

[11] K. F. C. Yiu, N. Grbić, K. L. Teo and S. Nordholm, A new design method for broadband microphone arrays for speech input in automobiles, *IEEE Signal Processing Letters*, vol.9, pp.222-224, 2002.

[12] K. F. C. Yiu, X. Q. Yang, S. Nordholm and K. L. Teo, Near-field broadband beamformer design via multidimensional semi-infinite-linear programming techniques, *IEEE Trans. on Speech and Audio Processing*, vol.11, pp.725-732, 2003.

[13] Z. Zhang, K. I. T. Miyake and T. Imamura, Three-dimension sound localization by binaural model using self organizing map, *International Journal of Innovative Computing, Information and Control*, vol.6, no.1, pp.361-371, 2010.