

Article

# Solution Strategies for Linear Inverse Problems in Spatial Audio Signal Processing

Mingsian R. Bai <sup>1,\*</sup>, Chun Chung <sup>1</sup>, Po-Chen Wu <sup>1</sup>, Yi-Hao Chiang <sup>1</sup> and Chun-May Yang <sup>2</sup>

<sup>1</sup> Department of Power Mechanical Engineering, National Tsing Hua University, No. 101, Section 2, Kuang-Fu Road, Hsinchu 30013, Taiwan; pups0409@yahoo.com.tw (C.C.); ricky82826@gmail.com (P.-C.W.); precious199382@gmail.com (Y.-H.C.)

<sup>2</sup> Department of Electrical Engineering, National Chiao Tung University, No. 1001, Ta-Hsueh Road, Hsinchu 30013, Taiwan; cmyang@cn.nctu.edu.tw

\* Correspondence: msbai63@gmail.com; Tel.: +886-3-5742915

Academic Editors: Woon Seng Gan and Jung-Woo Choi

Received: 30 March 2017; Accepted: 26 May 2017; Published: 5 June 2017

**Abstract:** The aim of this study was to compare algorithms for solving inverse problems generally encountered in spatial audio signal processing. Tikhonov regularization is typically utilized to solve overdetermined linear systems in which the regularization parameter is selected by the golden section search (GSS) algorithm. For underdetermined problems with sparse solutions, several iterative compressive sampling (CS) methods are suggested as alternatives to traditional convex optimization (CVX) methods that are computationally expensive. The focal underdetermined system solver (FOCUSS), the steepest descent (SD) method, Newton's (NT) method, and the conjugate gradient (CG) method were developed to solve CS problems more efficiently in this study. These algorithms were compared in terms of problems, including source localization and separation, noise source identification, and analysis and synthesis of sound fields, by using a uniform linear array (ULA), a uniform circular array (UCA), and a random array. The derived results are discussed herein and guidelines for the application of these algorithms are summarized.

**Keywords:** inverse problems; Tikhonov regularization; compressive sensing (CS); convex optimization (CVX); focal underdetermined system solver (FOCUSS); steepest descent (SD); Newton's method (NT); conjugate gradient (CG); golden section search (GSS)

## 1. Introduction

Numerous inverse problems exist in the field of acoustics. For example, nearfield acoustic holography (NAH) is a noise source identification method that reconstructs a surface field of the source on the basis of sound pressure measured in the nearfield of the source [1–5]. The deconvolution approach for the mapping of acoustic sources (DAMAS) is also a method for noise source identification [6]. Another example is the source signal separation problem, where an individual source signal is to be extracted from a mixed array of signals [7]. Inverse problems can also be found in source sound field synthesis (SFS) problems, where the sound field produced by secondary sources is to be matched with a target field [8,9]. Other examples include sound field control [10,11], crosstalk cancellation in binaural audio rendering [12], noise reduction in speech enhancement [13], room response equalization, and dereverberation [14,15]. In linear range acoustics, each of these problems can be formulated as a linear system ( $\mathbf{Ax} = \mathbf{b}$ ). The current study focused on the solutions of farfield noise source identification problems, sound source localization and separation problems, and sound field analysis (SFA) and synthesis (SFS) problems. Although inverse solutions of acoustic problems have long been investigated by researchers, according to our review of the literature, no conclusive results that give solution strategies and parameter choice guidelines can be found in the

literature. Furthermore, although audio quality is the chief concern in practical audio reproduction, most previous research has examined general numerical accuracy and stability. This study explored these problems from the perspective of reproduced signal quality. Solution strategies were compared in a unified framework, and guidelines of parameter choice are summarized herein.

In general, inverse problems can be divided into two categories: overdetermined and square systems, and underdetermined systems. To solve overdetermined systems, least-squares methods, Tikhonov regularization (TIKR) [16], and truncated singular value decomposition (TSVD) are commonly used. Traditionally, Morozov's discrepancy principle, generalized cross-validation (GCV), and the L-curve method can be used to choose the regularization parameter in the TIKR method [17–20]. However, solution methods that are better suited to audio applications than conventional approaches are proposed in this work. In particular, golden section search (GSS) [21] is applied to find optimal regularization parameters. To solve underdetermined problems, compressive sampling (CS) [22,23] solved by using convex optimization (CVX) [24–26] is a widely used approach that is known to be computationally expensive. In the present study, computationally efficient iterative approaches that incorporate sparsity constraints, including the focal underdetermined system solver (FOCUSS) [27], steepest descent (SD), Newton's (NT) method, and conjugate gradient (CG), were developed. These algorithms were compared for several audio application scenarios. The first scenario was sound source localization and separation using a uniform linear array (ULA) and a uniform circular array (UCA). To assess the separation quality, perceptual evaluation of audio quality (PEAQ), perceptual evaluation of speech quality (PESQ), and segmental signal-to-noise ratio (segSNR) were adopted [13,28]. The second scenario was concerned with analyses and syntheses of sound fields. Recently, an integrated array system was developed on the basis of a freefield model for spatial audio recording and reproduction [8,9]. This study extended the previous work to a reverberant environment; a live room was fitted with reflective walls. For the SFA, a 24-element circular microphone array was utilized to encode the sound field based on plane-wave decomposition, whereas in the SFS, a 32-element rectangular loudspeaker array was employed to decode the encoded sound field using three approaches. The third scenario was sound source localization and separation using a random array.

## 2. Inverse Solution Algorithms

In this section, an array model is given, along with its assumptions. Assume that the sound sources are at the farfield of the microphone array such that sound waves impinging on the array can be regarded as plane waves. In the following array model, time-harmonic dependence  $e^{j\omega t}$ ,  $j = \sqrt{-1}$  and  $\omega$  as angular frequency, is assumed so that the model is essentially formulated in the frequency domain.  $M$  microphones and  $N$  sources are considered. The array pressure vector can be expressed as [29].

$$\mathbf{p} = \mathbf{A}\mathbf{s} + \mathbf{v}, \quad (1)$$

where  $\mathbf{p} \in \mathbb{C}^M$  is the sound pressure vector received by the microphone,  $\mathbf{s} \in \mathbb{C}^N$  is the source amplitude vector,  $\mathbf{v}$  is the noise vector, and  $\mathbf{A} \in \mathbb{C}^{M \times N}$  is the steering matrix associated with the sources. Therefore, given the pressure measurement  $\mathbf{p}$  and the steering matrix  $\mathbf{A}$ , solving the problem of Equation (1) for the unknown source amplitude vector  $\mathbf{s}$  is a linear inverse problem. Linear inverse problems can be divided into three categories: square systems ( $M = N$ ), overdetermined systems ( $M > N$ ), and underdetermined systems ( $M < N$ ). In the following, solution strategies are presented for these categories.

### 2.1. Overdetermined or Square Systems

#### 2.1.1. TSVD and Least-Squares Problems

The most basic approach [16–18] to solve linear inverse problems is the least-squares method in which the following cost function is minimized:

$$J = \|\mathbf{e}\|_2^2 = \mathbf{e}^H \mathbf{e}, \tag{2}$$

where  $\mathbf{e} = \mathbf{p} - \mathbf{A}\mathbf{s}$  denotes the error vector and “ $\|\cdot\|_2$ ” denotes the vector 2-norm. If matrix  $\mathbf{A}$  is of full-column rank, the least-squares solution can be written as

$$\hat{\mathbf{s}} = [\mathbf{A}^H \mathbf{A}]^{-1} \mathbf{A}^H \mathbf{p}. \tag{3}$$

More generally, by the TSVD of

$$\begin{aligned} \mathbf{A} &= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H \\ \hat{\mathbf{s}} &= \mathbf{A}^+ \mathbf{p}, \end{aligned} \tag{4}$$

where  $\mathbf{\Sigma}$  represents a diagonal matrix with singular values at its diagonal entries,  $\mathbf{U}$  and  $\mathbf{V}$  represent unitary matrices, and  $\mathbf{A}^+$  represents the pseudoinverse of the matrix  $\mathbf{A}$  defined as

$$\mathbf{A}^+ = \mathbf{V}\mathbf{\Sigma}^+ \mathbf{U}^H, \tag{5}$$

where  $\mathbf{\Sigma}^+ = \text{diag}[\sigma_1^{-1}, \dots, \sigma_r^{-1}, \mathbf{0}, \dots, \mathbf{0}] \in \mathbb{C}^{N \times M}$ ,  $r = \text{rank}(\mathbf{A})$  [30,31]. Note that the expression of Equation (4) is sufficiently general that it always provides the minimum-norm least-squares solution, even when the matrix  $\mathbf{A}$  is rank-deficient. The square of the residual error becomes

$$\mathbf{e}_{LS}^2 = \|\mathbf{p} - \mathbf{A}\hat{\mathbf{s}}\|_2^2 = \|(\mathbf{I} - \mathbf{A}\mathbf{A}^+) \mathbf{p}\|_2^2 = \sum_{i=r+1}^M \left| \mathbf{u}_i^H \mathbf{p} \right|^2 \tag{6}$$

with  $\mathbf{u}_i$  being the  $i$ th column of  $\mathbf{U}$  and  $\mathbf{I}$  being the identity matrix.

In practice, the matrix  $\mathbf{A}$  can contain small singular values and can be very ill-conditioned, which leads to numerical instability during the inversion of  $\mathbf{A}$ . Two common methods to cope with the ill-posedness of the problem are TSVD and TIKR. Briefly, the TSVD method is simply to discard small singular values of the matrix  $\mathbf{A}$ , whereas TIKR involves attempts to minimize the following cost function [16]:

$$J = \|\mathbf{A}\hat{\mathbf{s}} - \mathbf{p}\|_2^2 + \beta^2 \|\mathbf{s}\|_2^2, \tag{7}$$

where the regularization parameter  $\beta$  weights the residual norm and the solution norm. After some manipulation, we derive with the following optimal solution:

$$\hat{\mathbf{s}} = \left( \mathbf{A}^H \mathbf{A} + \beta^2 \mathbf{I} \right)^{-1} \mathbf{A}^H \mathbf{p} \tag{8}$$

This result can be rewritten in terms of the TSVD of  $\mathbf{A}$  as follows:

$$\hat{\mathbf{s}} = \left( \mathbf{A}^H \mathbf{A} + \beta^2 \mathbf{I} \right)^{-1} \mathbf{A}^H \mathbf{p} = \sum_{i=1}^N f_i \sigma_i^{-1} \alpha_i \mathbf{v}_i, \tag{9}$$

where  $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^H = \sum_{i=1}^N \sigma_i \mathbf{u}_i \mathbf{v}_i^H$ , with  $\mathbf{u}_i$  and  $\mathbf{v}_i$  being the  $i$ th column partition of the matrices  $\mathbf{U}$  and  $\mathbf{V}$ ,  $\alpha_i = \mathbf{u}_i^H \mathbf{p}$ , where

$$f_i(\beta) = \frac{\sigma_i^2}{\sigma_i^2 + \beta^2} = \frac{1}{1 + (\beta/\sigma_i)^2} \tag{10}$$

denotes the filter function.

It can also be shown that the minimum residual vector can be written as

$$\mathbf{p} - \mathbf{A}\hat{\mathbf{s}} = \sum_{i=1}^N (1 - f_i) \alpha_i \mathbf{u}_i + \mathbf{r}_\perp \tag{11}$$

where  $\mathbf{r}_\perp = (\mathbf{I} - \mathbf{A}\mathbf{A}^+) \mathbf{p} = \sum_{i=M+1}^M \alpha_i \mathbf{u}_i$  is the residual vector of the components of  $\mathbf{p}$  orthogonal to  $\{\mathbf{u}_1 \cdots \mathbf{u}_n\}$ . The residual norm can be written as

$$\|\mathbf{p} - \mathbf{A}\hat{\mathbf{x}}\|_2^2 = \sum_{i=1}^n (1 - f_i)^2 |\alpha_i|^2 + \sum_{i=n+1}^m |\alpha_i|^2 = \sum_{i=1}^n (1 - f_i)^2 |\alpha_i|^2 + \|\mathbf{r}_\perp\|_2^2 \tag{12}$$

From Equation (9), the solution 2-norm can be written as

$$\|\hat{\mathbf{s}}\|_2^2 = \sum_{i=1}^N f_i^2 \sigma_i^{-2} |\alpha_i|^2 \tag{13}$$

### 2.1.2. Choice of Regularization Parameter $\beta$

In traditional solution strategies for inverse problems, methods are available for choosing regularization parameters, such as Morozov’s discrepancy principle, the Generalized Cross-Validation (GCV) method, and the L-curve method [19,20]. The first two methods have been discussed extensively in the literature. Therefore, we subsequently focus on only the L-curve method for brevity.

The L-curve method is widely used for choosing regularization parameters in inverse solutions. In the curve, the solution norm is plotted versus the residual norm by varying the regularization parameter. From Equations (9) and (11), it is straightforward to find the solution norm

$$\|\hat{\mathbf{s}}\|_2^2 = \sum_{i=1}^N f_i^2 \sigma_i^{-2} |\alpha_i|^2 \tag{14}$$

and the residual norm

$$\|\mathbf{p} - \mathbf{A}\hat{\mathbf{s}}\|_2^2 = \sum_{i=1}^N (1 - f_i)^2 |\alpha_i|^2 + \sum_{i=n+1}^M |\mathbf{r}_\perp|^2 \tag{15}$$

Regularization helps to improve robustness against system perturbation and measurement noise. Insights can be gained by writing the solution error as

$$\mathbf{s} - \hat{\mathbf{s}} = \mathbf{s} - \mathbf{A}^\# \mathbf{b} = (\mathbf{s} - \mathbf{A}^\# \bar{\mathbf{p}}) - \mathbf{A}^\# \mathbf{e} = \left[ \sum_{i=1}^N (1 - f_i) \frac{\mathbf{u}_i^H \bar{\mathbf{p}}}{\sigma_i} \mathbf{v}_i \right] - \sum_{i=1}^N f_i \frac{\mathbf{u}_i^H \mathbf{e}}{\sigma_i} \mathbf{v}_i, \tag{16}$$

where  $\mathbf{A}^\# = (\mathbf{A}^H \mathbf{A} + \beta^2 \mathbf{I})^{-1} \mathbf{A}^H$ ,  $(\mathbf{s} - \mathbf{A}^\# \bar{\mathbf{p}})$  is the regularization error and  $\mathbf{A}^\# \mathbf{e}$  is the perturbation error.

Hence, when  $\beta \rightarrow 0, f_i \rightarrow 1$  for very well-conditioned problems with high signal-to-noise ratio (SNR) measurements, the solution error is dominated by the perturbation error and a few high-order modes are filtered out. In this case, the solution norm is sensitive to the choice of  $\beta$ . The solution tends to be undersmoothed and susceptible to measurement noise. Conversely, when  $\beta \rightarrow \infty, f_i \rightarrow 0$  for very ill-conditioned problems with low SNR measurements, the solution error is dominated by regularization error and numerous high-order modes are filtered out. The solution tends to be oversmoothed and fine details such as resolution are thus lost. In this case, the residual norm is sensitive to the choice of  $\beta$ .

The parameter  $\beta$  acts as a weighting factor between the residual norms and the solution norm. Choosing an appropriate  $\beta$  to strike the balance between these two terms is vital. However, conventional approaches such as GCV and the L-curve method do not always provide satisfactory results in this situation. In this paper, a new method is proposed to facilitate the choice of the regular parameter  $\beta$  for the TIKR method.

Setting the gradient of the cost function of the TIKR method,  $J = \|\mathbf{p} - \mathbf{A}\mathbf{s}\|_2^2 + \beta^2 \|\mathbf{s}\|_2^2$ , to zero leads to the normal equation

$$(\mathbf{A}^H \mathbf{A} + \beta^2 \mathbf{I}) \mathbf{s} = \mathbf{A}^H \mathbf{p} \tag{17}$$

Without loss of generality, assume  $\mathbf{A}$  is of full-column rank. Note that

$$\mathbf{A}^H \mathbf{A} + \beta^2 \mathbf{I} = \mathbf{V} \left[ \Sigma^H \Sigma + \beta^2 \mathbf{I} \right] \mathbf{V}^H = \mathbf{V} \text{diag} \left[ (\sigma_1^2 + \beta^2), \dots, (\sigma_N^2 + \beta^2) \right] \mathbf{V}^H$$

which has an effective condition number  $\sqrt{(\sigma_1^2 + \beta^2)/(\sigma_N^2 + \beta^2)}$ . Therefore, if we want the condition number to be  $\tau$  after regularization, we must require

$$\tau^2 = \frac{\sigma_1^2 + \beta^2}{\sigma_N^2 + \beta^2} \tag{18}$$

Let  $\kappa$  be the condition number of  $\mathbf{A}$ ; that is,  $\kappa = \sigma_1/\sigma_M$ . In general, for very ill-conditioned systems,  $\kappa \gg \tau \gg 1$ , and

$$\beta = \sqrt{\frac{\sigma_1^2 - \tau^2 \sigma_M^2}{\tau^2 - 1}} \approx \sqrt{\frac{\sigma_1^2 - \tau^2 (\sigma_1^2 / \kappa^2)}{\tau^2}} = \sqrt{\frac{\sigma_1^2}{\kappa^2 \tau^2} (\kappa^2 - \tau^2)} \approx \frac{\sigma_1}{\tau} \tag{19}$$

Therefore, the regularization parameter  $\beta$  can be chosen to be the maximal singular value  $\sigma_1$  of  $\mathbf{A}$  divided by the regularized condition number  $\tau$ . For example, one may choose that  $\tau = 100$ , which causes 40-dB potential loss of SNR in the inverse solution. Normally,  $\mathbf{A}$  tends to be ill-conditioned at low frequencies. Choosing a frequency-independent  $\beta$  may suffice for the worst-case scenario. Thus, the regularization parameter  $\beta$  is chosen according to the maximal condition number at a selected low frequency (100 Hz is selected in the following simulation). Next, a coarse search is performed by varying  $\beta$  in orders of 10. A potential interval in which an optimal  $\beta$  may exist is located by observing how an objective function, such as Perceptual Evaluation of Speech Quality (PESQ) [29], varies with  $\beta$ . Finally, a fine-grained search is performed in the potential interval by using the Golden Section Search (GSS) algorithm [21].

GSS is an optimization technique suited for finding the extremum of a unimodal function. It is a simple bracketing method that does not require evaluation of the gradient of the cost function. In each search, a probe point must be selected within the left and right brackets according to the golden ratio. The golden ratio can be defined as

$$\varphi = \frac{1 + \sqrt{5}}{2} = 1.618, \varphi \text{ is the golden ratio}$$

Let  $f(\beta)$  be the objective function (PESQ in our case) for which we wish to find the optimal  $\beta$  that maximizes  $f(\beta)$ . First,  $f(\beta)$  has been evaluated at two points,  $\beta_1$  and  $\beta_3$ . The maximizing value is between  $\beta_1$  and  $\beta_3$ . The golden ratio can be used to find  $\beta_2$  and  $\beta_4$ .  $\beta_2$  and  $\beta_4$  can be written as

$$\beta_2 = \beta_3 - (\beta_3 - \beta_1)/\varphi, \beta_4 = \beta_1 + (\beta_3 - \beta_1)/\varphi$$

If  $f(\beta_2)$  is larger than  $f(\beta_4)$ , a maximum clearly lies in the interval between  $\beta_1$  and  $\beta_4$ . Therefore, the new  $\beta_3$  is equal to  $\beta_4$ . If  $f(\beta_2)$  is smaller than  $f(\beta_4)$ , a maximum lies in the interval between  $\beta_2$  and  $\beta_3$ . Therefore, the new  $\beta_1$  is equal to  $\beta_2$ . Figure 1 shows the Schematic of golden section search. The process is repeated until the gap between  $\beta_2$  and  $\beta_4$  is small. The iteration process stops when the beta converges within a prespecified tolerance window (0.0001 in our case). The optimal beta  $\beta_{opt}$  can be written as

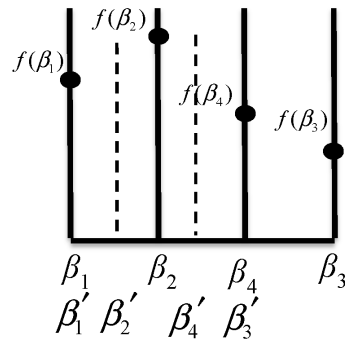
$$\beta_{opt} = (\beta_2 + \beta_4)/2$$

The algorithm is summarized as the following pseudocode:

```

 $\beta_2 = \beta_3 - (\beta_3 - \beta_1) / \varphi$ 
 $\beta_4 = \beta_1 + (\beta_3 - \beta_1) / \varphi$ 
while  $|\beta_2 - \beta_4| > \varepsilon$ 
  if  $f(\beta_4) < f(\beta_2)$ 
     $\beta_3 = \beta_4$ 
  else
     $\beta_1 = \beta_2$ 
end
 $\beta_{opt} = (\beta_2 + \beta_4) / 2$ 

```



**Figure 1.** Schematic of golden section search. If  $f(\beta_2)$  is higher than  $f(\beta_4)$ ,  $\beta'_3$  is equal to  $\beta_4$  in the next iteration.

Therefore, this study developed a procedure for choosing optimal regularization parameter beta. The procedure involves four steps as follows:

- Step 1. Select  $\tau$  according to a condition number threshold.
- Step 2. Select a constant  $\beta = \sigma_1 / \tau$  as an initial guess. For a frequency-domain design, it may be necessary to choose a frequency-independent  $\beta$  for the worst-case scenario.
- Step 3. Perform a coarse search by running the simulation forward and backward with 10 s powers of  $\beta$ . Locate a potential interval in which an optimal  $\beta$  may exist by observing the trend of an objective function, such as PESQ, with respect to  $\beta$ .
- Step 4. Perform a fine search by using optimization methods such as the GSS algorithm to find the optimal regularization parameter  $\beta$ .

## 2.2. Underdetermined Systems

In this section, algorithms are presented for underdetermined problems, where the number of microphones ( $M$ ) is lower than the number of potential sources ( $N$ ). In this case, the solution is generally not unique unless we impose constraints. Although a pseudoinverse gives a unique minimum-norm least-squares solution, the resolution of the solution is generally not favorable because the solution error tends to be evenly distributed among all entries. Instead, we impose sparsity as the constraint to limit the cardinality (nonzero entries) of the solutions in the study, which suggests that pruning procedures of some sort must be incorporated into the iteration process.

### 2.2.1. CVX Algorithms

An underdetermined problem with sparse solutions can be written as the following CS problem:

$$\min_{\mathbf{s}} \|\mathbf{s}\|_1 \quad \text{st.} \quad \|\mathbf{A}\mathbf{s} - \mathbf{p}\|_2 \leq \varepsilon, \tag{20}$$

where  $\|\cdot\|_1$  denotes the vector 1-norm. Numerous methods are available for solving this constrained optimization problem [22,23]. Suppose that the noise energy is constrained within a threshold  $\varepsilon$  that

can be selected with reference to the aforementioned least-squares solution. This problem can be solved numerically by CVX. Freeware was adopted to conduct CVX in this study [24–26].

### 2.2.2. Iterative Approaches

In the following, we apply four iterative algorithms to solve underdetermined problems. The first method is Focal Underdetermined System Solver (FOCUSS) [27], which is an iterative technique well suited for finding sparse solutions to underdetermined linear systems. The algorithm has two integral parts: a low-resolution initial estimate of the real signal and the iteration process that refines the initial solution to the final localized solution. Because the system is underdetermined, the sensors are more numerous than the sources. In this case, we assume our dictionary contains 36 sources. These sources are located at 5° intervals from the  $x$  label. Actually, this case has only three sources. Therefore, if the result is perfect, our final solution has only three nonzero solutions.

The FOCUSS algorithm can be summarized in three steps,

$$\mathbf{W}_k = [\text{diag}(\mathbf{s}_{k-1})] \tag{21}$$

$$\mathbf{q}_k = (\mathbf{A}\mathbf{W}_k)^+ \mathbf{p} \tag{22}$$

$$\mathbf{s}_k = \mathbf{W}_k \mathbf{q}_k \tag{23}$$

In Equation (21),  $[\text{diag}(\mathbf{s}_{k-1})]$  converts the vector  $\mathbf{s}_{k-1}$  into a diagonal weight matrix. The TIKR solution is used as the initial condition. Similar to other fixed-point iteration methods, the algorithms converge within finite numbers of iterations to the sparse solution with appropriate initial conditions.

The large term in the weighting reduces the 2-norm of  $\mathbf{q}$

$$\|\mathbf{W}^+ \mathbf{s}_p\|^2 = \|\mathbf{q}\|^2 = \sum_{i=1, \omega_i \neq 0}^n \left(\frac{\mathbf{s}_p^i}{w_i}\right)^2 \tag{24}$$

The relatively large entries in  $\mathbf{W}$  reduce the contributions of the corresponding elements of  $\mathbf{s}_p$  to the cost, and the solution is nonzero in the source direction. The pseudoinverse in Equation (22) can also be implemented by using the TIKR method. Therefore, it can also be written as

$$\mathbf{q}_k = (\mathbf{W}_{k-1}^H \mathbf{A}^H \mathbf{A} \mathbf{W}_{k-1} + \beta^2 \mathbf{I})^{-1} \mathbf{W}_{k-1}^H \mathbf{A}^H \mathbf{p} \tag{25}$$

The FOCUSS-TIKR method is robust to noise because of the regularization parameter  $\beta$ . The iteration process stops when the solution converges within a prespecified tolerance window (0.0001 in our case).

### 2.2.3. Iterative Approaches: Promote Sparsity by Pruning

CVX algorithms can solve CS problems, but these algorithms are known to be very computationally expensive, which prevents their use in real-time processing. To address this challenge, several iterative approaches are proposed as follows.

In these iterative techniques, the quadratic residual function

$$J(\mathbf{s}) = \frac{1}{2} \|\mathbf{A}\mathbf{s} - \mathbf{p}\|_2^2 = \frac{1}{2} [\mathbf{s}^H \mathbf{A}^H \mathbf{A} \mathbf{s} - \mathbf{p}^H \mathbf{A} \mathbf{s} - \mathbf{s}^H \mathbf{A}^H \mathbf{p} + \mathbf{p}^H \mathbf{p}] \tag{26}$$

is to be minimized. The key step that executes the “compressive sampling” is a pruning process that must be incorporated into each iteration to promote sparsity, as illustrated in Figure 2. First, several main peaks as well as sidelobes may appear in the source diagram. We reset all elements in the source vector  $\mathbf{s}$  below a prespecified threshold ( $s_{\max} - D$ ) to zero.  $D$  is initially set to a very small value  $D_0$ ; it is then increased incrementally by  $\Delta D$  in each iteration, typically by the same  $\Delta D$  in every step.

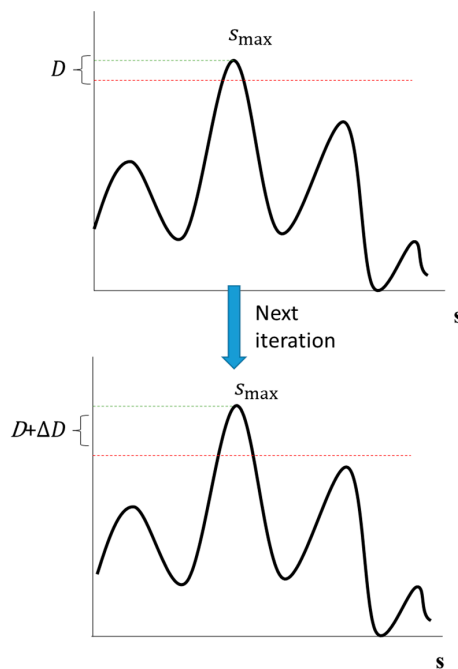


Figure 2. Pruning process to promote sparsity of inverse solution.

The iterative pruning process is summarized with a flowchart in Figure 3. The stopping condition is

$$D > D_{\max} \text{ or } \|\nabla J(\mathbf{s})\|_2^2 \leq 0.001 \tag{27}$$

in which  $\nabla J(\mathbf{s})$  is the gradient of the quadratic residual function  $J(\mathbf{s})$ . Three approaches were employed in this study to update the solutions in the iterative CS algorithms such that each quadratic residual function  $J(\mathbf{s})$  is minimized.

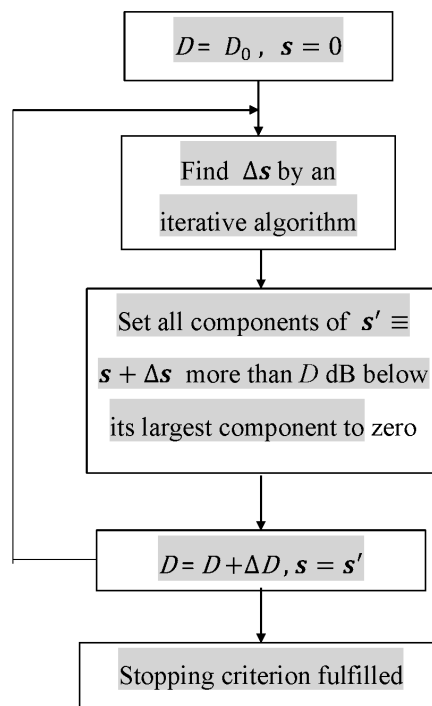


Figure 3. Flowchart of iterative compressive sampling algorithms (adapted from reference [5]).



### Steepest Decent (SD) Method

The SD algorithm is based on the notion that the search direction at each iteration is the negative gradient of the cost function in Equation (26) for minimization problems. The gradient vector of the quadratic residual function is given by

$$\mathbf{w} = -\nabla J(\mathbf{s}) = -(\mathbf{A}^H \mathbf{A} \mathbf{s} - \mathbf{A}^H \mathbf{p}) = \mathbf{A}^H \mathbf{r}, \tag{28}$$

where  $\mathbf{r} = \mathbf{p} - \mathbf{A} \mathbf{s}$  is the residual vector.

The new solution  $\mathbf{s}'$  is updated as

$$\mathbf{s}' = \mathbf{s} + \Delta \mathbf{s} = \mathbf{s} + \mu \mathbf{w} \tag{29}$$

where  $\mu$  is the step size. To determine the optimal step size  $\mu$ , let the vector  $\mathbf{g}$  be

$$\mathbf{g} = \mathbf{A} \mathbf{w} \tag{30}$$

It can be shown after some algebraic manipulations that

$$F(\mathbf{s}') = \frac{1}{2}(\mu^2 \mathbf{g}^H \mathbf{g} - 2\mu \mathbf{g}^H \mathbf{r} + \mathbf{r}^H \mathbf{r}) \tag{31}$$

From Equation (31), the step size  $\mu$  to minimize  $F(\mathbf{s}')$  along the direction  $\mathbf{w}$  can be found by setting the derivative of  $F(\mathbf{s}')$  with respect to  $\mu$  to zero. Consequently, we obtain

$$\mu = \frac{\|\mathbf{w}\|_2^2}{\|\mathbf{g}\|_2^2} \tag{32}$$

### Newton's Method

The NT method is also an iterative method gradient search. Recall the gradient of the cost function is

$$\nabla F(\mathbf{s}) = \mathbf{v} = (\mathbf{A}^H \mathbf{A} \mathbf{s} - \mathbf{A}^H \mathbf{p}) = -\mathbf{A}^H (\mathbf{p} - \mathbf{A} \mathbf{s}) = -\mathbf{A}^H \mathbf{r}, \tag{33}$$

where  $\mathbf{r}$  is as defined before. Setting this gradient to zero leads to the optimal solution

$$\mathbf{s}' = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{p} = \mathbf{A}^+ \mathbf{p}, \tag{34}$$

where  $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ . Next, multiplying  $(\mathbf{A}^H \mathbf{A})^{-1}$  on both sides of  $\mathbf{v}$  yields

$$\begin{aligned} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{v} &= (\mathbf{A}^H \mathbf{A})^{-1} (\mathbf{A}^H \mathbf{A} \mathbf{s} - \mathbf{A}^H \mathbf{p}) = \mathbf{s} - (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{p} = \mathbf{s} - \mathbf{A}^+ \mathbf{p} \\ &= -(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{r} = -\mathbf{A}^+ \mathbf{r} \end{aligned}$$

or

$$-\mathbf{A}^+ \mathbf{r} = \mathbf{s} - \mathbf{A}^+ \mathbf{p} = \mathbf{s} - \mathbf{s}'$$

Hence, the solution can be updated as

$$\mathbf{s}' = \mathbf{s} + \mathbf{A}^+ \mathbf{r} \tag{35}$$

In practical implementation, the pseudoinverse  $\mathbf{A}^+$  is usually of the form of TIKR, namely  $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A} + \beta^2 \mathbf{I})^{-1} \mathbf{A}^H$ , because  $\mathbf{A}^H \mathbf{A}$  is singular for underdetermined problems. As a refinement of the algorithm, a step size  $\mu$  can be used to rewrite the update equation as

$$\mathbf{s}' = \mathbf{s} + \mu \mathbf{A}^+ \mathbf{r} \tag{36}$$

### Conjugate Gradient (CG) Method

The CG method is an iterative algorithm well suited for the numerical solution of systems of linear equations, associated with symmetric and positive-definite matrices. Instead of the negative gradient used in the SD method, which occasionally causes zigzag convergence, the search direction of the CG method is a linear combination of the current negative gradient and the previous search direction. Development of the CG algorithm is based on nested Krylov subspaces. For details, we refer the interested readers to Ref. [32–34]. For brevity, we only summarize the CG algorithm with the following pseudocode:

```

 $\mathbf{r}_0 = \mathbf{p} - \mathbf{A}\mathbf{s}_0$ 
 $\mathbf{p}_0 = \mathbf{r}_0$ 
 $k = 0$ 
for
 $\alpha_k = \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{p}_k^T \mathbf{A} \mathbf{p}_k}$ 
 $\mathbf{s}_{k+1} = \mathbf{s}_k + \alpha_k \mathbf{p}_k$ 
 $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{A} \mathbf{p}_k$ 
 $\beta_k = \frac{\mathbf{r}_{k+1}^T \mathbf{r}_{k+1}}{\mathbf{r}_k^T \mathbf{r}_k}$ 
 $\mathbf{p}_{k+1} = \mathbf{r}_{k+1} + \beta_k \mathbf{p}_k$ 
 $k = k + 1$ 
end

```

To put the system of equations in Equation (1) into a more tractable form, we multiply by  $\mathbf{A}^H$  on both sides, which leads to the normal equations

$$\mathbf{A}^H \mathbf{p} = \mathbf{A}^H \mathbf{A} \mathbf{s}_p \quad (37)$$

This equation is equivalent to finding the vector  $\mathbf{s}$  for which the gradient of  $F(\mathbf{s})$  equals zero.

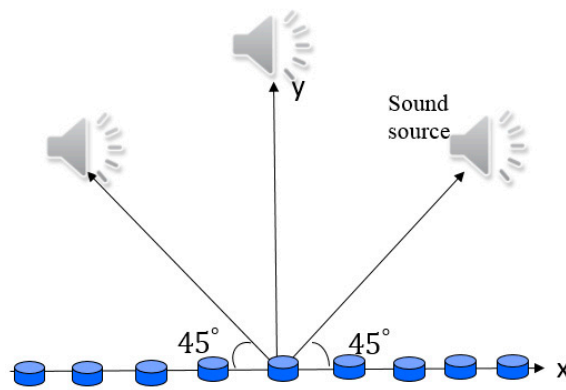
$$\nabla F(\mathbf{s}) = -\mathbf{A}^H (\mathbf{p} - \mathbf{A} \mathbf{s}) = 0 \quad (38)$$

### 3. Comparison of Algorithms

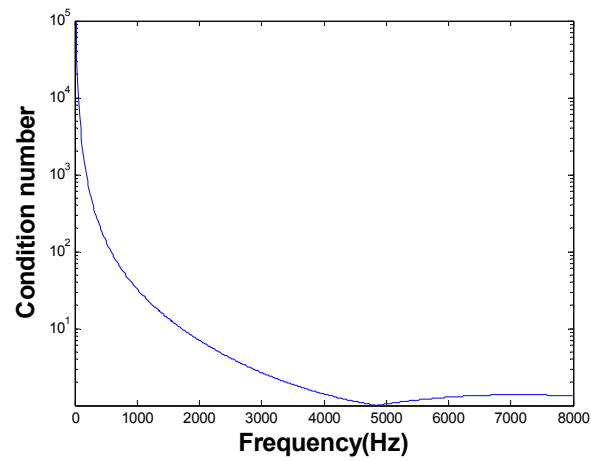
This section presents the application of the preceding algorithms to acoustic source localization and separation problems through numerical simulation. These algorithms were compared for three example problems, with the aid of a uniform linear array (ULA) and a random array. In addition, an inverse solution involved in sound field analysis (SFA) and sound field synthesis (SFS) in spatial audio was investigated. Microphone data are synthetic and generated by the model of Equation (1).

#### 3.1. Uniform Linear Array

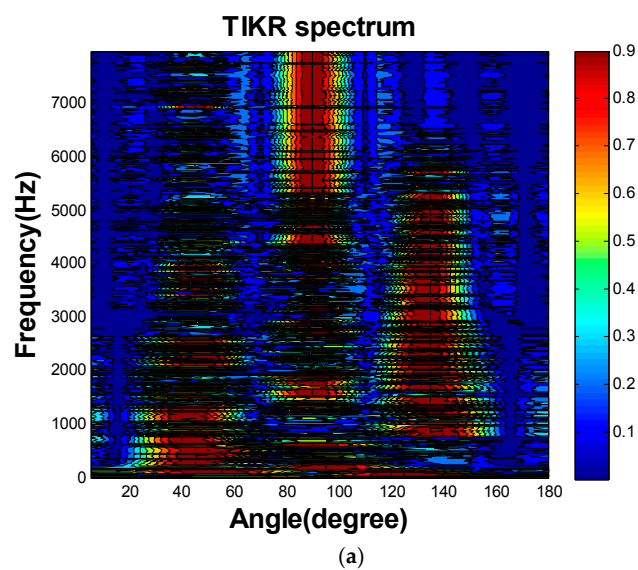
In the numerical simulation shown in Figure 4, 10-microphone ULA was utilized to separate the signals emitted by three sources. The sources were located at the far field such that the plane wave assumption was valid. The spacing between adjacent microphones was 10 cm. This simulated underdetermined system contained 36 sources as our dictionary. We separated the signals and localized the signals in one stage. TIKR, FOCUS, and CS-CVX algorithms were used to solve these inverse problems. Figure 5 shows the condition numbers of different frequencies. The problem is ill-conditioned at low frequencies. Source localization results are shown in Figure 6.



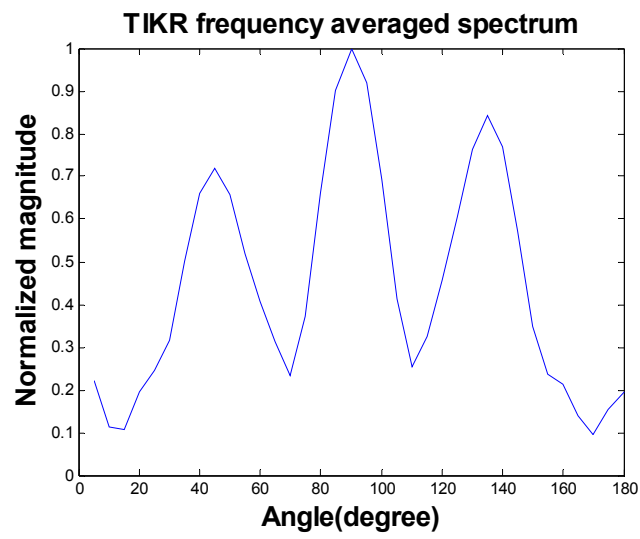
**Figure 4.** Numerical simulation of 10-microphone uniform linear array utilized to separate the signals emitted by three sources. The first source, located at  $\theta = 45^\circ$ , was broadcasting a male speech signal. The second source, located at  $\theta = 90^\circ$ , was broadcasting a female speech signal. The third source, located at  $\theta = 135^\circ$ , was broadcasting a music signal.



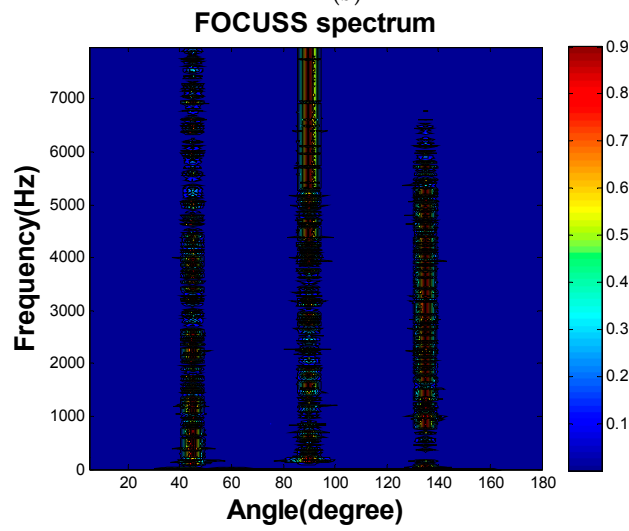
**Figure 5.** Condition number versus frequency.



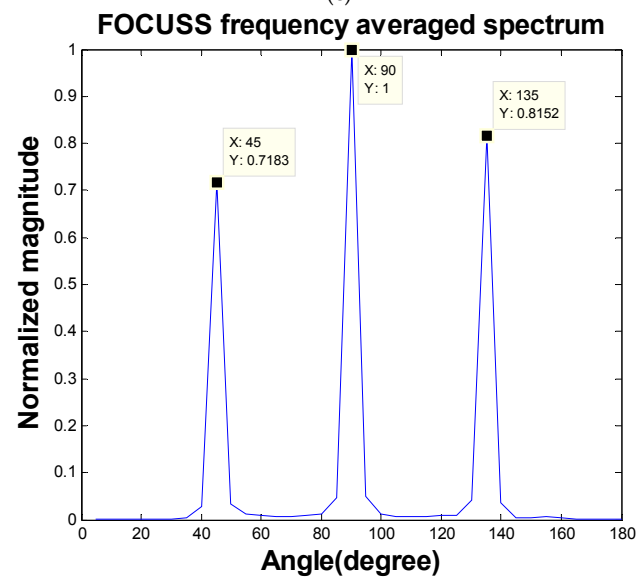
**Figure 6.** Cont.



(b)

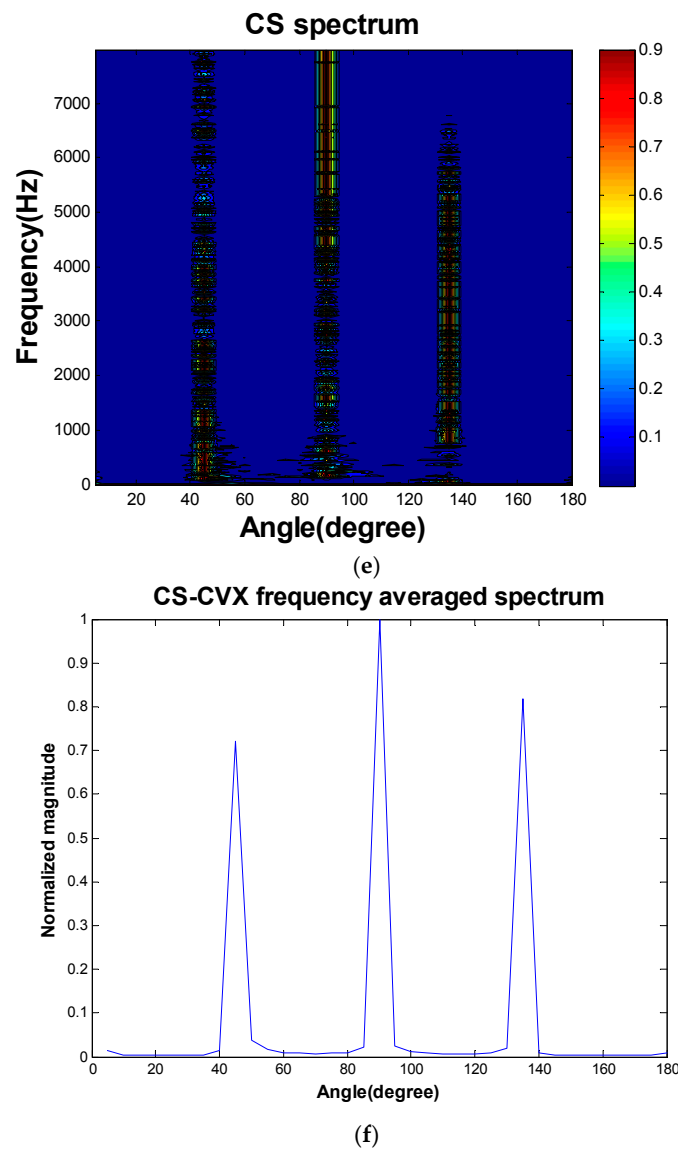


(c)



(d)

Figure 6. Cont.



**Figure 6.** Source localization results obtained using the focal underdetermined system solver (FOCUSS) algorithm and a uniform linear array. (a) Tikhonov regularization (TIKR) spectrum (b) TIKR averaged spectrum (c) FOCUSS spectrum (d) FOCUSS frequency-averaged spectrum. (e) Compressive sampling-convex optimization (CS-CVX) spectrum (f) CS-CVX frequency-averaged spectrum.

The separation results obtained using TIKR, FOCUSS, and CS-CVX are summarized in Table 1. PESQ is an objective test measure for speech quality evaluation. It is a full-reference algorithm and analyzes the speech signal sample-by-sample after a temporal alignment of corresponding excerpts of reference and test signal. The mean opinion score (MOS) is calculated on the basis of PESQ ranging from 1 to 5; MOS signifies the difference in speech quality between the clean and the separated signals, which is affected by separation performance and signal distortion. The segmental SNR (segSNR) is defined as

$$SNR_{seg} = \frac{1}{N} \sum_{k=1}^N 10 \log_{10} \left[ \frac{\sum_{n \in frame_k} |\mathbf{s}(n)|^2}{\sum_{n \in frame_k} |\hat{\mathbf{s}}(n) - \mathbf{s}(n)|^2} \right] \quad (39)$$

The segSNR correlates with the effect of noise reduction. The FOCUSS-PINV algorithm was observed to achieve the highest score in PESQ and segSNR (Table 1), although it required more computation time than TIKR and FOCUSS-TIKR.

**Table 1.** Separation performance of the Tikhonov regularization (TIKR) and focal underdetermined system solver (FOCUSS) methods for three sources in the underdetermined system.

Methods		TIKR	FOCUSS-PINV	FOCUSS-TIKR	CS-CVX
PESQ	Source 1	2.034	3.966	3.350	3.251
	Source 2	1.696	3.146	2.879	2.783
	Source 3	1.912	3.818	3.394	3.347
segSNR	Source 1	0.554	11.92	11.27	8.817
CPU time (s)		53	810	487	16302

In our previous simulation, we simulated microphones that had no noise; therefore, our regularization parameter was very close to zero. In the current simulation, our microphones did have white noise with a magnitude equal to the magnitude of the microphone signals divided by 100. Therefore, the potential loss of SNR was 40 dB.

The regularization parameter is chosen by the maximal singular  $\sigma_1$  value dividing 100 in 100 Hz. In our case, the maximal singular value was equal to 5.32. Next, a coarse search was performed by varying  $\beta$  in orders of 10 and using the GSS algorithm to find the optimal regularization parameter. In this case, the optimal regularization parameter was 0.0174, and the FOCUSS-PINV of the robustness to the noise was very ineffective because PINV artifacts caused discontinuities in regularization. CS-CVX was sensitive to the noise, and the segSNR and PESQ achieved lower scores than FOCUSS-TIKR did, as listed in Tables 2 and 3 (20dB). Therefore, noise was present, and FOCUSS-TIKR was the best choice in this case. It outperformed PESQ and segSNR.

**Table 2.** Separation performance of the Tikhonov regularization (TIKR), focal underdetermined system solver (FOCUSS), and compressive sampling-convex optimization (CS-CVX) methods for three sources with additive white noises (40dB SNR).

Methods		TIKR	FOCUSS-PINV	FOCUSS-TIKR	CS-CVX
PESQ	Source 1	2.135	1.170	2.941	2.412
	Source 2	1.810	1.121	2.234	1.321
	Source 3	1.723	1.168	2.841	2.241
segSNR	Source 1	0.52	-10.00	1.558	0.946
CPU time (s)		58	817	497	16431

**Table 3.** Separation performance of the TIKR, FOCUSS and CS-CVX methods for three sources with additive white noises (20dB SNR).

Methods		TIKR	FOCUSS-PINV	FOCUSS-TIKR	CS-cvx
PESQ	Source 1	1.512	1.168	2.512	1.436
	Source 2	1.401	1.118	2.012	1.712
	Source 3	1.489	1.044	2.487	1.324
segSNR	Source 1	-0.12	-13.11	0.12	-1.112
CPU time (s)		57	815	494	16421

### 3.2. Random Array

A simulation was conducted for localization and separation of two point sources located at (0, 0, -1 m) and (0.8, 0.3, -1 m), both of which were emitting clean speech signals, as illustrated

in Figure 7. A 30-element random array with aperture dimension  $0.48\text{ m} \times 0.4\text{ m}$  situated at  $z = 0$  was utilized to capture the signals emitted by these two sources. To set up the propagation matrix, 100 ( $10 \times 10$ ) equivalent sources were distributed on the image plane located 1 m away from the array. Figure 8 shows the condition numbers of different frequencies. The problem is ill-conditioned at low frequencies.

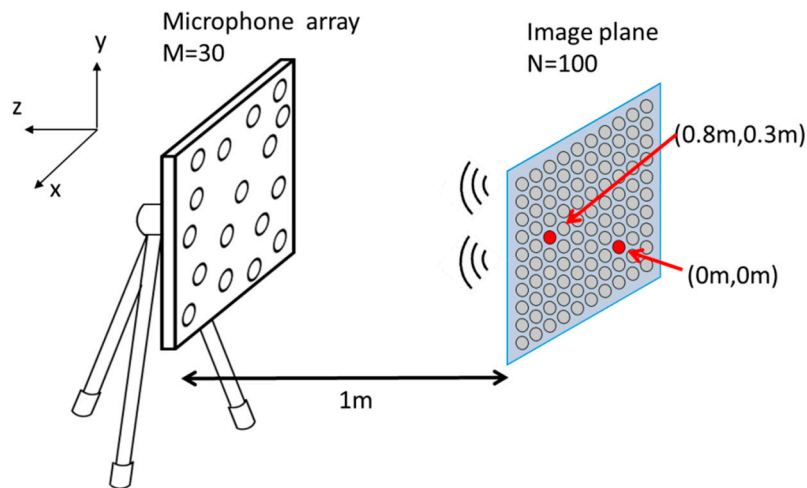


Figure 7. Arrangement for simulation of localization and separation of two sources.

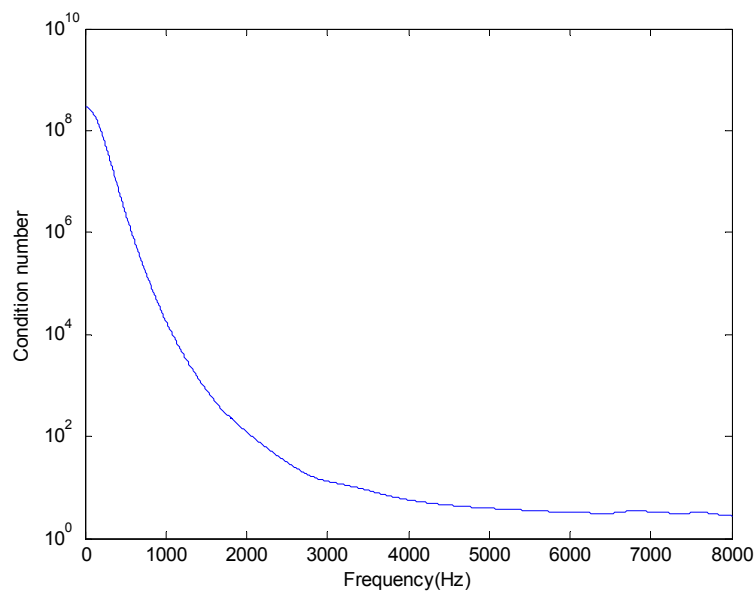
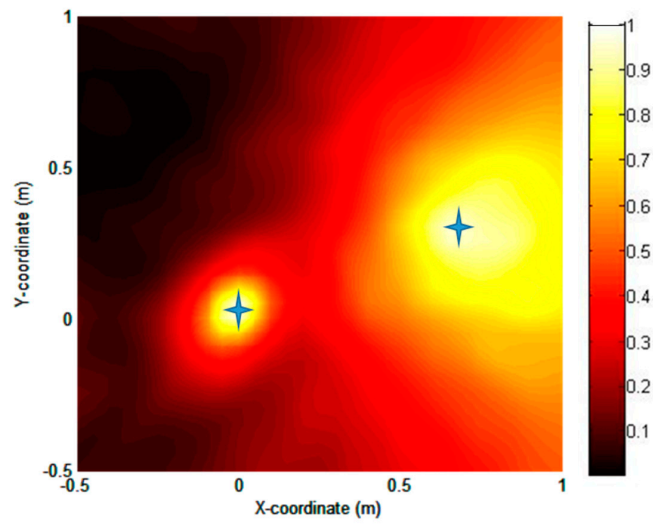
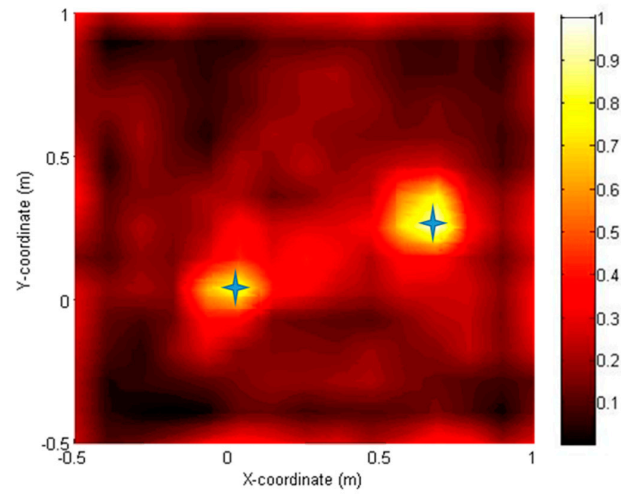


Figure 8. Condition number versus frequency.

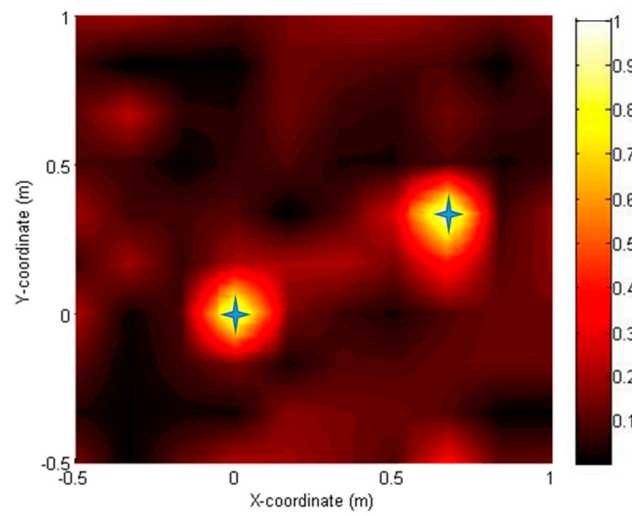
Figure 9a–f show the source localization results obtained using six approaches. Two sources were correctly located on the noise map with varying degrees of resolution by all methods. A conventional method delay and sum (DAS) method (a) gave the poorest resolution, whereas the CS-CVX method provided the highest resolution. The SD method (d) and CG method (f) were acceptable but did not perform quite as well as the CS-CVX method. The NT method (e) yielded accurate source locations with a slightly increased sidelobe level, but it was the most computationally efficient (Table 2).



(a)



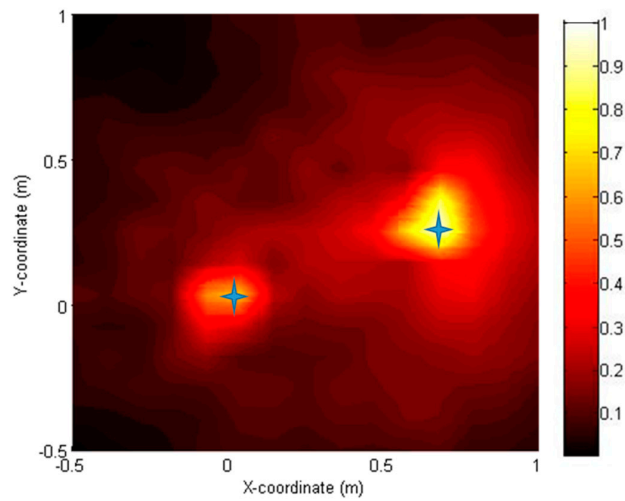
(b)



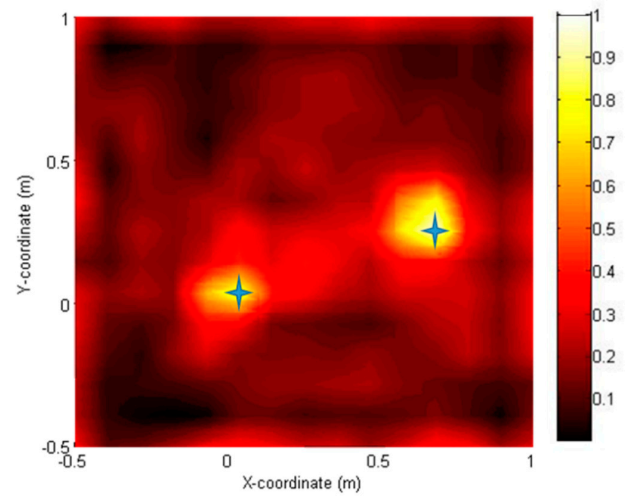
(c)

Figure 9. Cont.

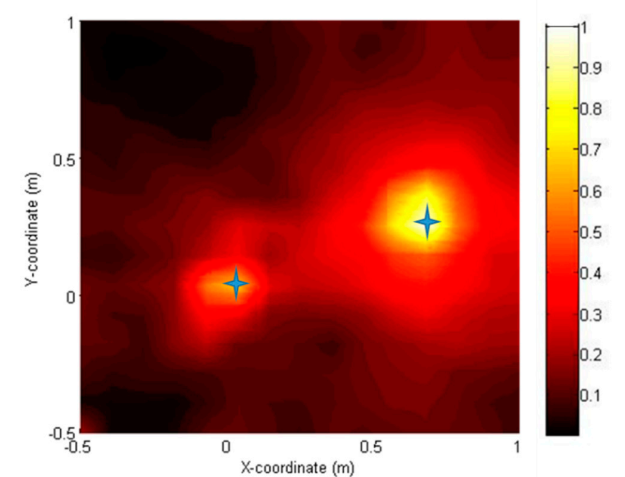




(d)



(e)



(f)

**Figure 9.** Localization results of two point sources. (a) Delay and sum algorithm, (b) Tikhonov regularization algorithm, (c) compressive sampling-convex optimization algorithm, (d) steepest descent algorithm, (e) Newton's algorithm, and (f) conjugate gradient algorithm.

Table 4 presents a comparison of the separation results obtained using five methods. CS-CVX displayed the highest scores for PESQ and segSNR, despite being extremely time-consuming. Iterative CS approaches were determined to be far more computationally efficient than the CS-CVX method. The CG method attained the highest PESQ, but the lowest segSNR. This suggests that the favorable separation performance of the CG method comes at the price of signal distortion. The SD and NT methods demonstrated acceptable PESQ and high segSNR. Although signals were not perfectly separated by using these two methods, the incurred distortion was minor. In general, methods present a trade-off between separation performance and signal distortion. Table 5 shows the separation results with additive noise (SNR = 28 dB). All the methods were observed to suffer from the interference of noise; consequently, the values of PESQ and segSNR were notably low. However, all the methods were determined to be robust to noise. The present study also considered mismatches between equivalent sources (dictionaries) and real sources. Table 6 shows the separation results of the NT method with different levels of mismatch. Mismatch means that the real source is not precisely on the deployed source location (dictionary). Extreme mismatch means the real source is exactly on the center of four nearest the deployed source location. More descriptions are added to the revised manuscript. Unless the source was just at the center of the near dictionaries, the separation performance was high and not influenced by noise.

**Table 4.** Separation performance of five algorithms for two speech sources.

Methods		TIKR	CS-CVX	SD	NT	CG
PESQ	Source 1	1.99	3.12	2.39	2.48	2.76
	Source 2	2.60	3.31	2.77	2.83	3.13
segSNR	Source 1	2.15	7.54	5.08	6.04	0.50
	Source 2	2.73	8.24	7.02	7.26	1.53
CPU time (s)		201	31065	377	296	386

**Table 5.** Separation performance of five algorithms for two speech sources with additive noise.

Methods		TIKR	CS-CVX	SD	NT	CG
PESQ	Source 1	1.82	2.83	2.22	2.38	2.59
	Source 2	2.12	3.20	2.53	2.53	2.88
segSNR	Source 1	1.40	2.11	2.03	4.13	−0.50
	Source 2	2.30	6.51	2.09	1.72	1.23
CPU time (s)		211	31348	358	290	307

**Table 6.** Separation performance of Newton’s algorithm for two speech sources with different mismatch conditions.

Methods		without Mismatch	with Mismatch	Extreme Mismatch
PESQ	Source 1	2.52	2.48	2.14
	Source 2	2.81	2.83	1.70
segSNR	Source 1	6.90	6.04	−0.40
	Source 2	6.33	7.26	0.18

### 3.3. SFA and SFS

Depending on the sparsity of the sound sources, the SFA stage can be implemented in several manners [35–38]. For the sparse-source scenario, a two-stage algorithm is utilized; the source bearings are estimated using the minimum power distortionless response (MPDR) [7] and the associated amplitudes of plane waves are estimated using the TIKR algorithm. For the nonsparse-source scenario, a one-stage algorithm based on the CS-CVX algorithm or the FOCUSS algorithm is employed.

The SFS stage is carried out using a loudspeaker array to reconstruct the sound field with the source bearing and amplitude obtained in the SFA stage. Pressure matching was employed for the SFS purpose in this study by sampling a large number of virtual control points in the interior area surrounded by the loudspeakers. The pressure matching procedure can be described as the following optimization problem:

$$\min_{\mathbf{s}_s(\omega)} \|\mathbf{B}(\omega)\mathbf{s}_p(\omega) - \mathbf{H}(\omega)\mathbf{s}_s(\omega)\|, \tag{40}$$

where  $\mathbf{s}_p(\omega) = [s_1(\omega) \ \dots \ s_p(\omega)]^T$  is the amplitude vector of the  $P$ th primary plane-wave component,  $\mathbf{s}_s(\omega) = [s_1(\omega) \ \dots \ s_L(\omega)]^T$  denotes the amplitude vector of the input signals to the  $L$  secondary loudspeaker sources,  $\mathbf{H}(\omega) \in \mathbb{C}^{K \times L}$  denotes the room response matrix, and  $\mathbf{b}_d = [e^{-jk_d \cdot y_1} \ \dots \ e^{-jk_d \cdot y_K}]^T$  is the steering vector for the  $d$ th primary plane-wave component to the  $n$ th control point,  $\mathbf{y}_n, n = 1, \dots, K$ .  $\mathbf{B}(\omega) = [\mathbf{b}_1 \ \dots \ \mathbf{b}_d] \in \mathbb{C}^{K \times P}$  is the steering matrix from the plane-wave components obtained in the preceding SFA stage to the control points. Therefore, the optimal solution can be written as

$$\mathbf{s}_s(\omega) = \mathbf{H}^\#(\omega)\mathbf{B}(\omega)\mathbf{s}_p(\omega), \tag{41}$$

where “#” symbolizes some type of inverse operation on the matrix  $\mathbf{H}(\omega)$ . In this study, TIKR was utilized to calculate the input signal amplitudes to the secondary sources. GSS can be used to find the optimal regularization parameter.

Experiments were conducted to validate the proposed audio analysis and synthesis system. In the SFA stage, a 24-element circular microphone array with a radius of 12 cm was utilized to capture and parameterize the sound field in an anechoic chamber (the recording room), as illustrated in Figure 10. In the SFS stage, a rectangular, 32-loudspeaker array was employed to reproduce in a live room (the reproduction room) the sound field previously encoded in the SFA stage. The walls of the room were lined with acoustically reflective boards (Figure 11).

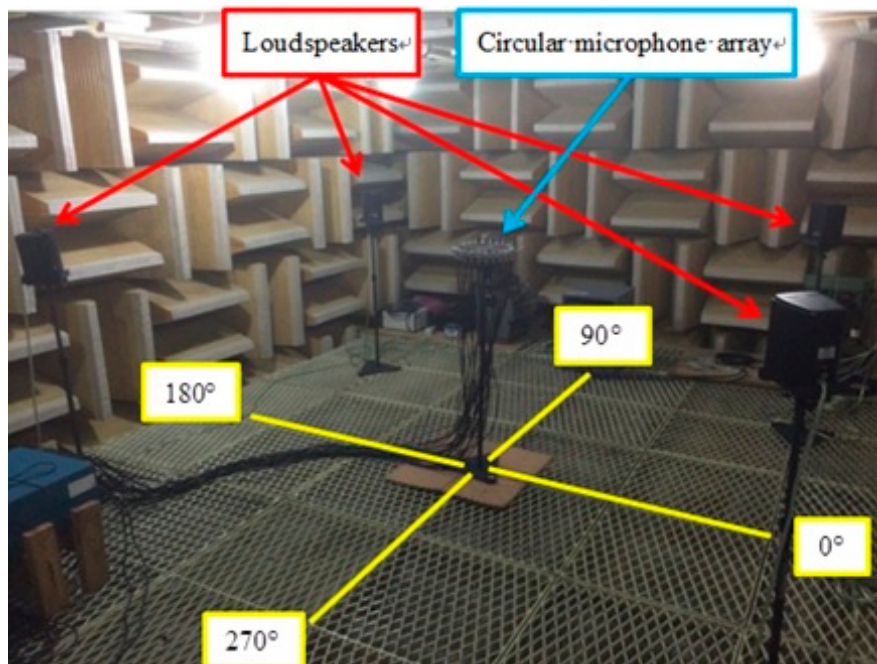


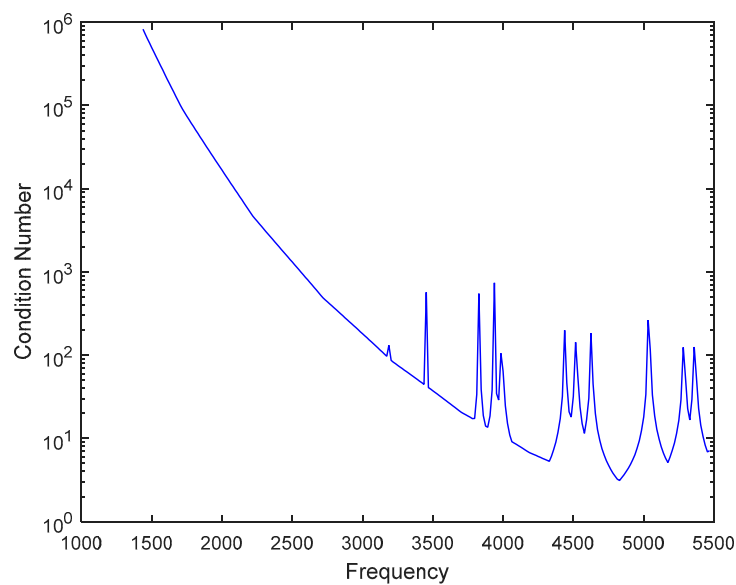
Figure 10. Sound field analysis experimental arrangement in a 5.4 m × 3.5 m × 2 m anechoic room.



**Figure 11.** Sound field synthesis experimental arrangement in a  $3.6 \text{ m} \times 3.6 \text{ m} \times 2 \text{ m}$  live room fitted with reflective walls.

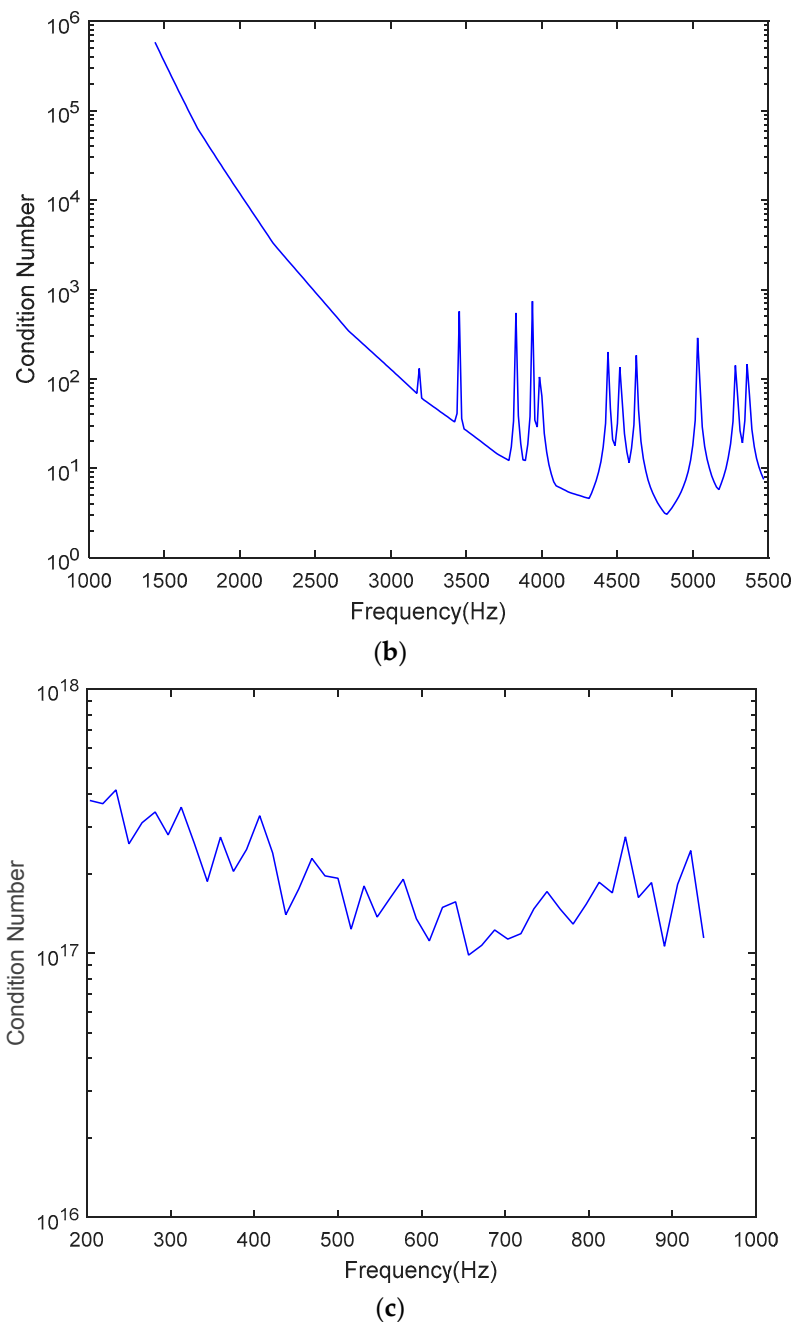
To process microphone output and loudspeaker input signals, multichannel analog-to-digital converters (M-32 AD) and digital-to-analog converters (M-32 DA) (RME, Haimhausen, Germany) were used with a sampling frequency of 16 kHz.

An audio codec system involves three inverse problems, namely the SFA stage, room response modeling, and the SFS stage. The condition numbers are plotted against the frequencies of three steering matrices in Figure 12a–c. Figure 12c indicates that the ill-posedness encountered in the room response modeling procedures must be addressed, with the aid of appropriate regularization methods. Large regularization parameters can increase the robustness of the inverse problem. In this study, we set the regularization parameter to 10.



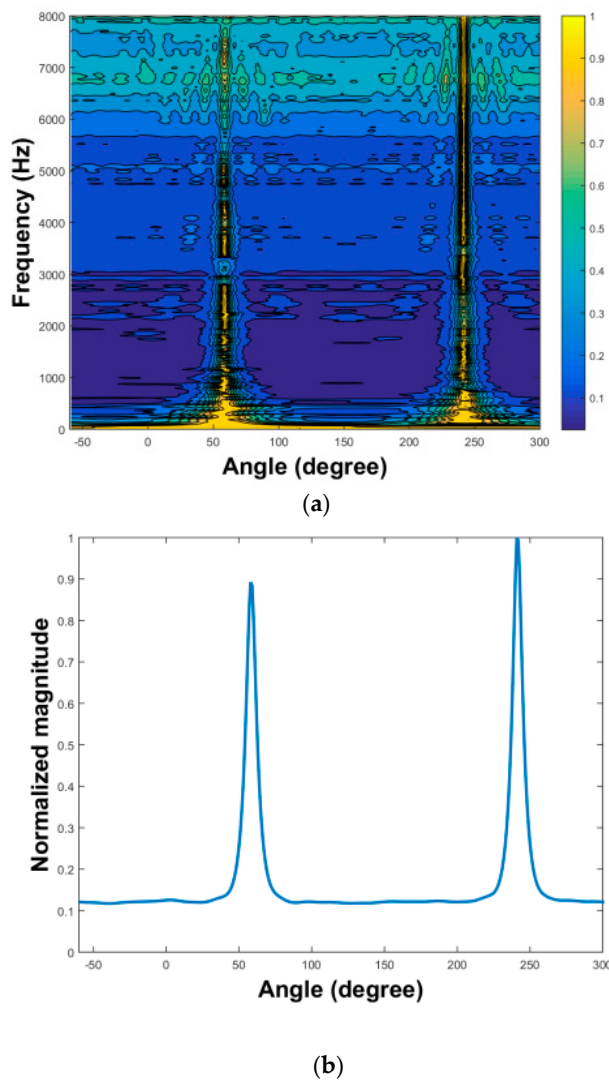
(a)

**Figure 12.** *Cont.*



**Figure 12.** Plots of condition numbers versus frequencies of each steering matrix in three inverse problems: (a) sound field analysis stage; (b) room response modeling; and (c) sound field synthesis stage.

In the SFA experiment, loudspeaker sources positioned at the angles  $\theta = 60^\circ, 240^\circ$  played two 10-s speech clips. After recording the source by CMA, we used three algorithms to extract the source signals. First, we applied the two-stage MPDR and TIKR algorithms. The MPDR spectrum is plotted as a function of angle and frequency in Figure 13a. The resulting frequency-averaged and normalized MPDR spectrum is illustrated in Figure 13b, which peaks at the angles  $\theta = 60^\circ, 240^\circ$  as desired. The results show that the source was accurately localized using MPDR. Next, the source signals were extracted using the TIKR algorithm. We also applied one-stage CS algorithms and one-stage FOCUSS algorithms, which located sources and separated their amplitudes in a single calculation.



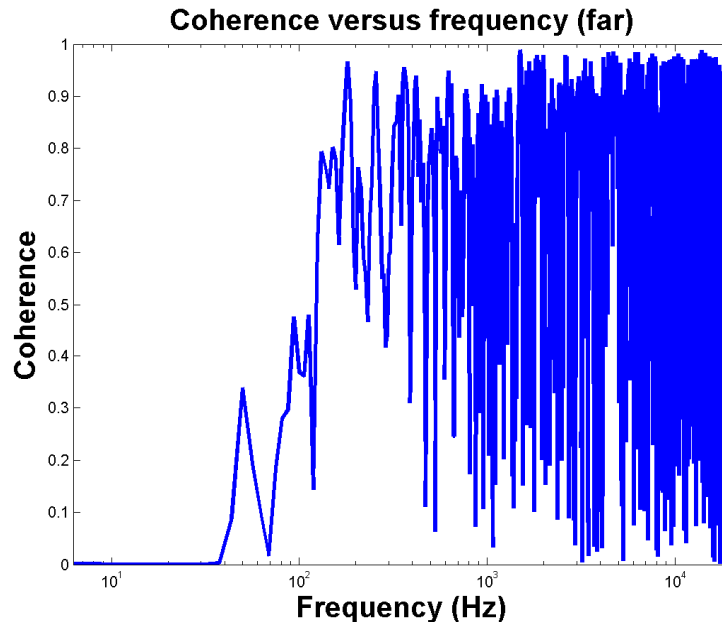
**Figure 13.** Localization of one music source signal with sampling rate 16 kHz, located at the angles 60°, 240° in an anechoic chamber. By using a uniform circular array with a radius of 12 cm, the source direction can be identified by the peak in the angular spectrum. (a) Minimum power distortionless response (MPDR) spectrum plotted versus angle and frequency. (b) Frequency-averaged and normalized MPDR spectrum.

The signals extracted using different methods were evaluated by using the MOS of the PESQ test. Results confirmed that the TIKR performed well in signal separation with satisfactory audio quality. The results are summarized in Table 7.

**Table 7.** Mean opinion score of perceptual evaluation of speech quality for source signal separation at the angles 60°, 240° with speech signals using Tikhonov regularization (TIKR), compressive sampling-convex optimization (CS-CVX), and focal underdetermined system solver (FOCUSS).

Methods		Two-Stage TIKR	One-Stagecs-CVX	One-Stagefocuss
PESQ	Source 1	2.84	3.11	1.56
	Source 2	2.79	2.99	1.61
segSNR	Source 1	17.40	20.74	13.54
	Source 2	16.59	17.34	12.68
CPU time (s)		8	27,588	275

One sample coherence function between one loudspeaker and one microphone is shown in Figure 14, indicating the signal quality to be poor below 200 Hz. Therefore, band-limited processing was applied for all frequencies up to 200 Hz in the SFS stage. In this frequency range, pressure matching was used on the basis of the room response model.



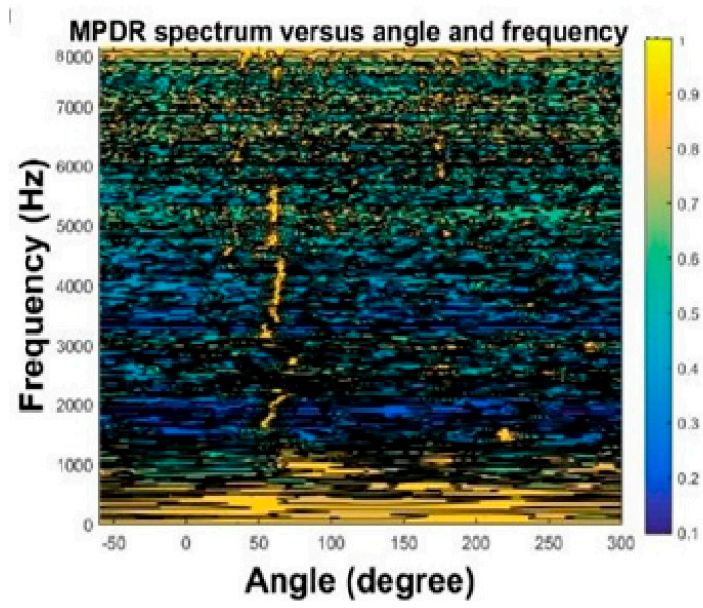
**Figure 14.** Coherence curve measured from one loudspeaker to one microphone in a live room. A PULSE analysis platform made from Brüel & Kjær was arranged to measure the coherence curve. A white noise signal with sampling rate 16 kHz was used as the driving signal of the loudspeaker.

The SFS stage was conducted for three different methods. The coherence between the loudspeaker and the microphone was poor below 200 Hz; therefore, the signals below 200 Hz were not processed. Method 1, band-limited processing, was applied from 200 Hz to the spatial aliasing frequency, 952 Hz, in the SFS stage. In this frequency range, pressure matching was performed on the basis of the room response model. Below 200 Hz, unprocessed audio signals were fed directly to the loudspeakers. Above 952 Hz, a simple vector panning [39] approach was adopted. The optimal regularization parameter  $\beta$  achieving the highest MOS in room response modeling was calculated using GSS [21] as  $\beta = 0.0008634$ .

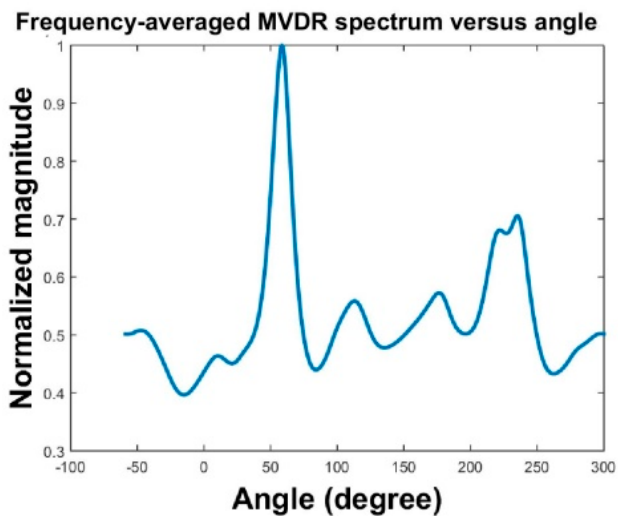
In the second method, instead of a vector panning method, we used DAS to process signals above 952 Hz. In the third method, we used pressure matching to obtain signals above 200 Hz. The use of different regularization parameters in pressure matching results in different levels of localization performance and audio quality.

Figure 15a,c shows the MPDR spectrum and the normalized MPDR spectrum obtained using the third method for  $\beta = 0.01$  and  $\beta = 10$ , respectively. Low values of the regularization parameter  $\beta$  yielded higher localization performance than high values did. These two signals were compared with the clean signal through the PESQ test. The results showed that the high  $\beta$  ensured satisfactory voice quality, whereas the low  $\beta$  impaired voice quality.

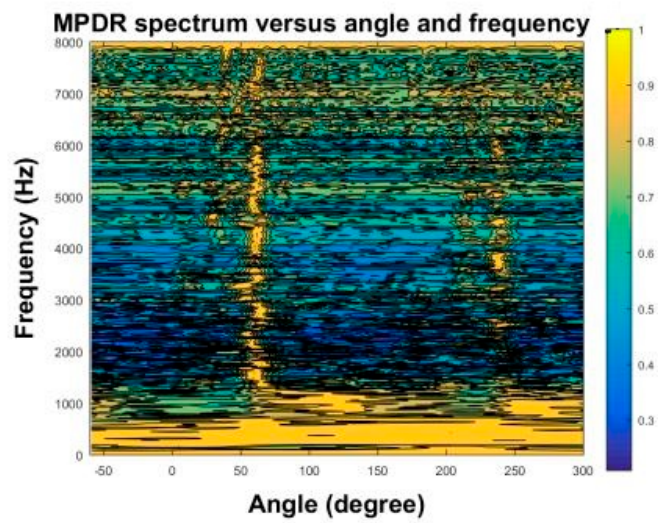
The results of three localization methods are presented in Figure 16a–f. The MPDR spectra are plotted as functions of angle and frequency in Figure 16a,c,f. The resulting frequency-averaged and normalized MPDR spectra are shown in Figure 16b,d,f.



(a)



(b)



(c)

Figure 15. Cont.



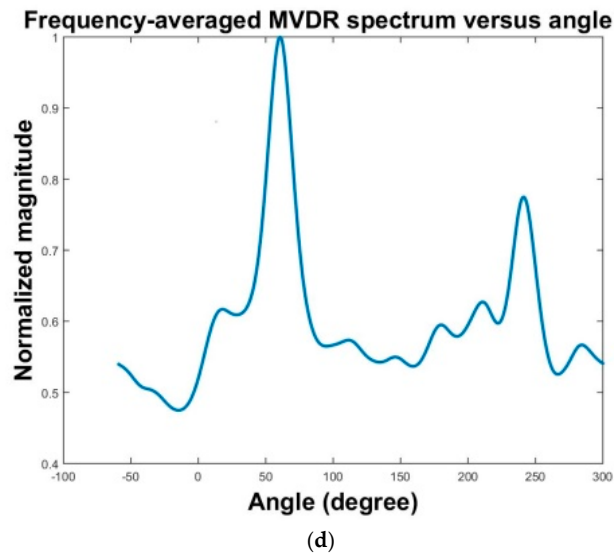


Figure 15. Localization results in the sound field synthesis experiment by the third method with different regularization parameters.  $\beta = 10$  (a,b) and  $\beta = 0.01$  (c,d).

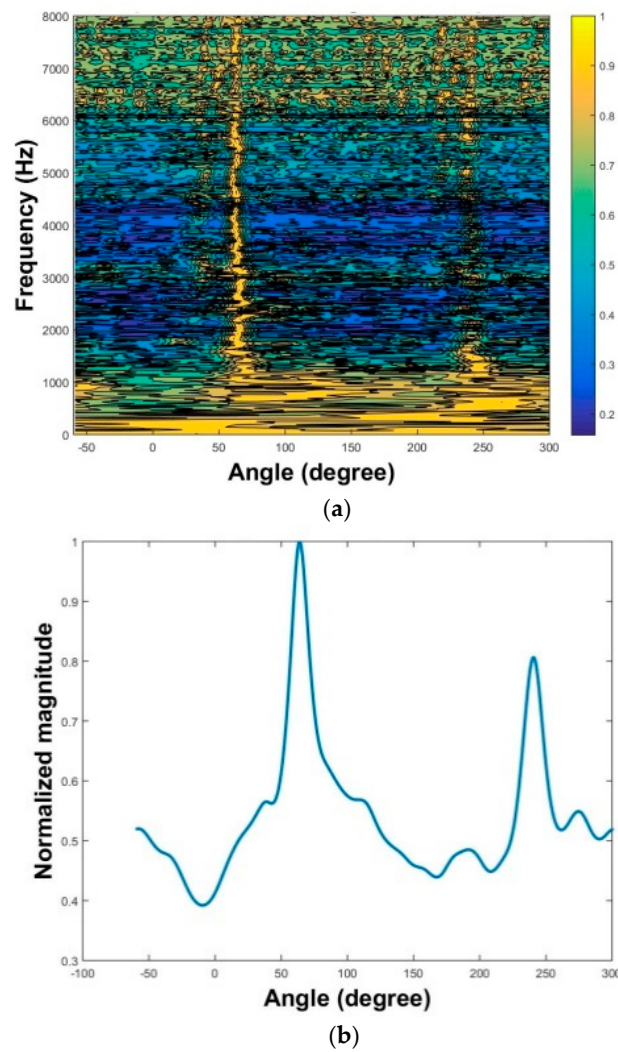
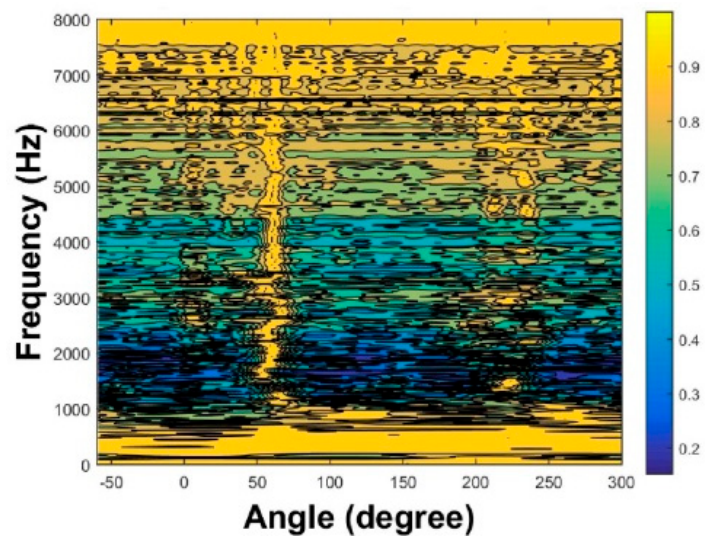
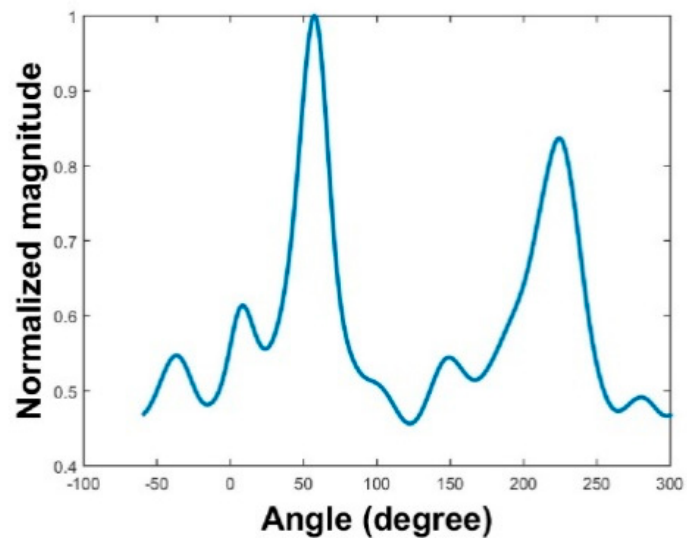


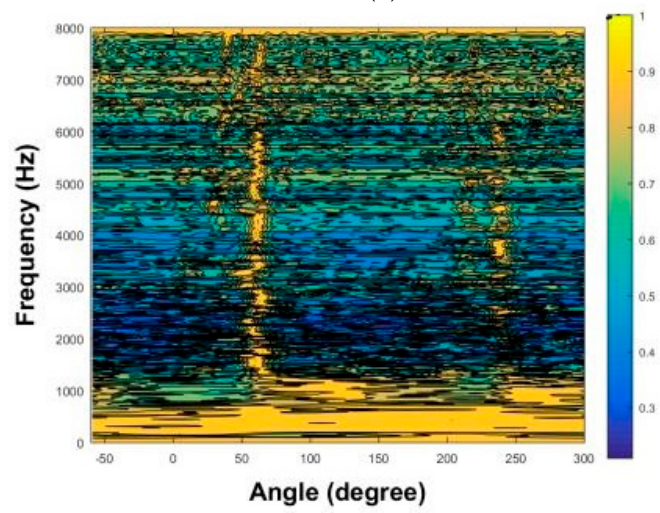
Figure 16. Cont.



(c)

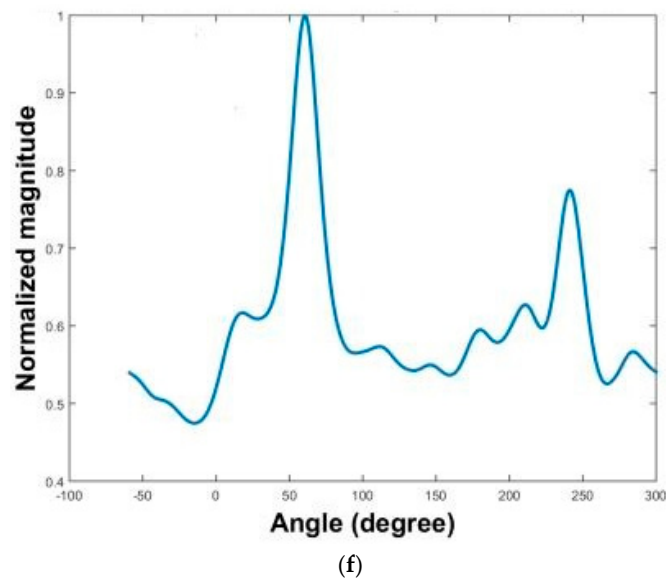


(d)



(e)

Figure 16. Cont.



**Figure 16.** Localization results in the sound field synthesis experiment with three different approaches. (a,b) first method, (c,d) second method, and (e,f) third method.

#### 4. Conclusions

This study developed algorithms for solving inverse problems generally encountered in spatial audio signal processing. The TIKR algorithm was shown to solve overdetermined problems. However, the regularization parameter in the TIKR method was not effectively chosen. This study thus presents a guideline for choosing the optimal regularization parameter  $\beta$  in the TIKR method. Specifically, choosing the optimal  $\beta$  involves dividing the maximal singular value at low frequency by the threshold and then running the simulation forward and backward for powers of 10 of  $\beta$ . Optimization methods such as GSS can be used by observing the trend of an objective function (such as PESQ). Some trade-offs must be made between localization performance and voice quality. In general, a high  $\beta$  results in a small solution norm with high voice quality, whereas a low  $\beta$  yields a small residual norm with high localization performance.

Inverse problems in noise sound source localization and separation problems can be solved by 1-stage and 2-stage (overdetermined and underdetermined), each of the 1-stage and 2-stage methods has its advantages and disadvantages. In general, the 1-stage methods provide both localization and separation results with good performance. The 2-stage methods give slightly better separation performance than the 1-stage methods. From our experience, PESQ correlates better with separability and segSNR correlates better with distortion.

For 1-stage (underdetermined) problems, iterative CS algorithms have been developed for solving acoustic inverse problems, with applications to localization and separation. The results demonstrate that the CS-CVX method was effective in solving CS problems, despite being computationally expensive. Iterative CS methods achieved comparable performance to the CS-CVX method for CS problems, in far less computation time. The FOCUSS-TIKR and CG methods attained high PESQ, whereas the SD and NT methods attained high segSNR. In general, iterative CS methods were determined to perform better than the TIKR method. For 1-stage methods, FOCUSS-TIKR attains the highest MOS value of PESQ for clean signals, while the Newton method performs the best. Both methods require less CPU time than the CS-CVX.

Inverse solution approaches are also useful in solving SFA and SFS problems. In this study, three inverse problems were solved for implementing an audio codec system. Because of the ill-posed yield at low frequencies, particularly in the room response modeling stage, choosing an appropriate regularization parameter  $\beta$  was crucial. Therefore, this stage required a larger regularization parameter

than the SFA and SFS stages required. In the analysis stage, the one-stage CS algorithm was determined to be more computationally expensive than the two-stage TIKR algorithm. In the synthesis stage, the first method performed well in localization, but did not perform well in reproduced voice quality. As compared with the third method, the second method reproduced signals with boosted high-frequency content above 952 Hz with poor localization. The third method had the highest performance in terms of voice quality and localization performance.

**Acknowledgments:** The work was supported by the Ministry of Science and Technology (MOST) of Taiwan, Republic of China, under project number 102-2221-E-007-029-MY3.

**Author Contributions:** M.R. Bai and C.-M. Yang conceived and designed the experiments; Y.-H. Chiang and P.-C. Wu performed the experiments; P.-C. Wu and C. Chung analyzed the data; C. Chung and Y.-H. Chinag wrote the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Kim, Y.; Nelson, P.A. Spatial resolution limits for the reconstruction of acoustic source strength by inverse methods. *J. Sound Vib.* **2003**, *265*, 583–608. [[CrossRef](#)]
- Nelson, P.A.; Yoon, S.H. Estimation of acoustic source strength by inverse methods: Part I, conditioning of the inverse problem. *J. Sound Vib.* **2000**, *233*, 639–664. [[CrossRef](#)]
- Kim, Y.; Nelson, P.A. Optimal regularisation for acoustic source reconstruction by inverse methods. *J. Sound Vib.* **2004**, *275*, 463–487. [[CrossRef](#)]
- Maynard, J.D.; Williams, E.G.; Lee, Y. Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH. *J. Acoust. Soc. Am.* **1985**, *78*, 1395–1413. [[CrossRef](#)]
- Hald, J. Fast wideband acoustical holography. *J. Acoust. Soc. Am.* **2016**, *139*, 1508–1517. [[CrossRef](#)] [[PubMed](#)]
- Brooks, T.F.; Humphreys, W.M. A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays. *J. Sound Vib.* **2006**, *294*, 856–879. [[CrossRef](#)]
- Bai, M.R.; Kuo, C.H. Deconvolution-based acoustic source localization and separation algorithms. *J. Acoust. Soc. Am.* **2014**, *135*, 2358. [[CrossRef](#)]
- Bai, M.R.; Hua, Y.H.; Kuo, C.H.; Hsieh, Y.H. An integrated analysis-synthesis array system for spatial sound fields. *J. Acoust. Soc. Am.* **2015**, *137*, 1366–1376. [[CrossRef](#)] [[PubMed](#)]
- Bai, M.R.; Hsu, H.S.; Wen, J.C. Spatial sound field synthesis and upmixing based on the equivalent source method. *J. Acoust. Soc. Am.* **2014**, *135*, 269–282. [[CrossRef](#)] [[PubMed](#)]
- Elliott, S.J.; Cheer, J.; Murfet, H.; Holland, K.R. Minimally radiating sources for personal audio. *J. Acoust. Soc. Am.* **2010**, *128*, 1721–1728. [[CrossRef](#)] [[PubMed](#)]
- Bai, M.R.; Hsieh, Y.H. Point focusing using loudspeaker arrays from the perspective of optimal beamforming. *J. Acoust. Soc. Am.* **2015**, *137*, 3393–3410. [[CrossRef](#)] [[PubMed](#)]
- Bai, M.R.; Tung, C.W.; Lee, C.C. Optimal design of loudspeaker arrays for robust cross-talk cancellation using the Taguchi method and the genetic algorithm. *J. Acoust. Soc. Am.* **2005**, *117*, 2802–2813. [[CrossRef](#)] [[PubMed](#)]
- Loizou, P.C. *Speech Enhancement: Theory and Practice*; Taylor & Francis: Park Drive, UK; Abingdon, UK, 2007.
- Shabtai, N.R. Optimization of the directivity in binaural sound reproduction beamforming. *J. Acoust. Soc. Am.* **2015**, *138*, 3118–3128. [[CrossRef](#)] [[PubMed](#)]
- Miyoshi, M.; Kaneda, Y. Inverse filtering of room acoustic. *IEEE Transac. Acoust. Speech Signal Process.* **1998**, *36*, 145–152. [[CrossRef](#)]
- Groetsch, C.W. *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*; Pitman Advanced Pub. Program: Boston, MA, USA, 1984.
- Hansen, P.C. *Rank-Deficient and Discrete Ill-Posed Problems*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1998.
- Bertero, M.; Poggio, T.; Torre, V. Ill-Posed Problems in Early Vision. *Proc. IEEE* **1988**, *76*, 869–889. [[CrossRef](#)]
- Hansen, P.C. *Analysis of Discrete Ill-Posed Problems by Means of the L-Curve*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1992.

20. Hansen, P.C.; O'leary, D.P. *The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1993.
21. Brent, R.P. *Algorithms for Minimization without Derivatives*; Prentice-Hall, Inc.: Englewood Cliffs, HJ, USA, 1973; pp. 48–75.
22. Candes, J.; Wakin, M.B. An introduction to compressive sampling. *IEEE Signal Process. Mag.* **2008**, *25*, 21–30. [[CrossRef](#)]
23. Edelmann, G.F.; Gaumont, C.F. Beamforming using compressive sensing. *J. Acoust. Soc. Am.* **2011**, *130*, 232–237. [[CrossRef](#)] [[PubMed](#)]
24. Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: New York, NY, USA, 2004; Chapters 1–7.
25. Bai, M.R.; Chen, C.C. Application of Convex Optimization to Acoustical Array Signal Processing. *J. Sound Vib.* **2013**, *332*, 6596–6616. [[CrossRef](#)]
26. Grant, M.; Boyd, S. CVX, Version 1.21 MATLAB Software for Disciplined Convex Programming. Available online: <http://cvxr.com/cvx> (accessed on 14 June 2013).
27. Gorodnitsky, I.F.; Rao, B.D. Sparse Signal Reconstruction from Limited Data Using FOCUSS: A Re-weighted Minimum Norm Algorithm. *IEEE Trans. Signal Process.* **1997**, *45*, 600–616. [[CrossRef](#)]
28. ITU-T Recommendation P.862. *Perceptual Evaluation of Speech Quality (Pesq): An Objective Method for End-to-End Speech Quality Assessment of Narrow-Band Telephone Networks and Speech Codecs*; International Telecommunication Union: Geneva, Switzerland, 2001; p. 21.
29. Bai, M.R.; Ih, J.G.; Benesty, J. *Acoustic Array Systems: Theory, Implementation, and Application*, 1st ed.; Wiley-IEEE Press: Singapore, 2013; Chapters 3–4.
30. Golub, G.H.; van Loan, C.F. *Matrix Computations*, 3rd ed.; Johns Hopkins University Press: Baltimore, MD, USA, 1989; Chapter 12.
31. Noble, B.; Daniel, J.W. *Applied Linear Algebra*; Prentice Hall: Englewood, NJ, USA, 1977.
32. Hestenes, M.R.; Stiefel, E. Methods of conjugate gradients for solving linear systems. *J. Res. Natl. Bur. Stand.* **1952**, *49*, 409–436. [[CrossRef](#)]
33. Fraysse, V.; Giraud, L. *A Set of Conjugate Gradient Routines for Real and Complex Arithmetics*; CERFACS Technical Report TR/PA/00/47; Cedex: Toulouse, France, 2000.
34. Ginn, K.B.; Haddad, K. Noise source identification techniques: Simple to advanced applications. *Proc. Acoust.* **2012**, *2012*, 1781–1786.
35. Capon, J. High-Resolution Frequency-Wavenumber Spectrum Analysis. *Proc. IEEE* **1969**, *57*, 1408–1418. [[CrossRef](#)]
36. Gomes, J.; Hald, J.; Juhl, P.; Jacobsen, F. On the Applicability of the Spherical Wave Expansion with a Single Origin for Near-Field Acoustical Holography. *J. Acoust. Soc. Am.* **2009**, *125*, 1529–1537. [[CrossRef](#)] [[PubMed](#)]
37. Candes, J.; Romberg, J.; Tao, T. Stable Signal Recovery from Incomplete and Inaccurate Measurements. *Commun. Pure Appl. Math.* **2006**, *59*, 1207–1223. [[CrossRef](#)]
38. Candes, J.; Romberg, J.; Tao, T. Robust Uncertainty Principles: Exact Signal Reconstruction Form Highly Incomplete Frequency Information. *IEEE Trans. Inf. Theory* **2006**, *52*, 489–509. [[CrossRef](#)]
39. Kim, Y.H.; Choi, J.W. *Sound Visualization and Manipulation*; Wiley: Singapore, 2013.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).