

Article

A High-Resolution Time Reversal Method for Target Localization in Reverberant Environments

Huiying Ma ^{1,2,*}, Tao Shang ^{1,2}, Gufeng Li ^{1,2} and Zhaokun Li ^{1,2}

¹ State Key Laboratory of Integrated Service Network, Xidian University, Xi'an 710071, China; tshang@xidian.edu.cn (T.S.); ligufeng@xidian.edu.cn (G.L.); zhkli@xidian.edu.cn (Z.L.)

² Collaborative Innovation Center of Information Sensing and Understanding, Xidian University, Xi'an 710071, China

* Correspondence: huiyingma@xidian.edu.cn

Abstract: Reverberation in real environments is an important factor affecting the high resolution of target sound source localization (SSL) methods. Broadband low-frequency signals are common in real environments. This study focuses on the localization of this type of signal in reverberant environments. Because the time reversal (TR) method can overcome multipath effects and realize adaptive focusing, it is particularly suitable for SSL in a reverberant environment. On the basis of the significant advantages of the sparse Bayesian learning algorithm in the estimation of wave direction, a novel SSL is proposed in reverberant environments. First, the sound propagation model in a reverberant environment is studied and the TR focusing signal is obtained. We then use the sparse Bayesian framework to locate the broadband low-frequency sound source. To validate the effectiveness of the proposed method for broadband low-frequency targeting in a reverberant environment, simulations and real data experiments were performed. The localization performance under different bandwidths, different numbers of microphones, signal-to-noise ratios, reverberation times, and off-grid conditions was studied in the simulation experiments. The practical experiment was conducted in a reverberation chamber. Simulation and experimental results indicate that the proposed method can achieve satisfactory spatial resolution in reverberant environments and is robust.

Keywords: broadband low frequency; target localization; time reversal; sparse Bayesian learning; reverberant environments



Citation: Ma, H.; Shang, T.; Li, G.; Li, Z. A High-Resolution Time Reversal Method for Target Localization in Reverberant Environments. *Sensors* **2024**, *24*, 3196. <https://doi.org/10.3390/s24103196>

Academic Editor: Riccardo Carotenuto

Received: 15 April 2024

Revised: 11 May 2024

Accepted: 14 May 2024

Published: 17 May 2024



Copyright: © 2024 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Broadband low-frequency target sound source localization (SSL) is a long-standing and widely researched topic in reverberant environments [1,2]. Thus far, this remains a highly challenging issue. Reverberation generally occurs in enclosed spaces, including indoor spaces, cabins, and car interiors. SSL has numerous practical applications, such as in speech enhancement, noise control, human–robot interaction, and smart homes. This study mainly focuses on rectangular indoor rooms.

The traditional SSL method is based on the signal/channel model and signal processing and has made significant progress over the years. Currently, mainstream SSL methods include beamforming [3], time delay of arrival [4], and high-resolution spectrum estimation [5]. All of these methods achieved good localization performance in free space. However, a common problem is that the physical model assumed in these methods only considers direct sound, requires the sound source to be stable and incoherent, and has poor robustness in an environment of low signal-to-noise ratio (SNR) and reverberation. Moreover, the spatial resolution is limited, particularly at low frequencies [6].

SSL in reverberant environments is challenging because of the multipath nature of the received signals. This effect degrades SSL performance. Thus, it is important to develop SSL methods that are robust to reverberation [7]. The time reversal (TR) method derived from optical phase conjugation has attracted the attention of researchers because it can

overcome the multipath effect and realize adaptive focusing without a large array [8]. The TR technique was first applied in the field of ultrasound and then gradually extended to medical diagnosis, seismology, underwater acoustics, etc. In recent years, it has been introduced in the field of indoor SSL. Fink et al. [9–12] demonstrated that the TR method can achieve SSL in reverberant environments in both time and space, but the focal spot size obtained by the TR method is larger than $\lambda/2$ (λ is a wavelength) for low-frequency sound sources. Draeger et al. [13,14] achieved and demonstrated, for the first time, the effectiveness of SSL in a reverberant silicon cavity using the TR technique with both simulation and experimentation. The experiment adopted one microphone and a pulse signal with a center frequency of 1 MHz, and the focusing resolution reached 1.1 mm, corresponding to $1/2 \lambda$ of the signal. Fink et al. [15] systematically studied the validity of the TR method in SSL from both theoretical and experimental perspectives and demonstrated the effectiveness of the TR technique based on source–receiver reciprocity. Fink et al. [16] then applied the TR method to locate the sound source with a center frequency of 975 Hz using 20 microphones and obtained a resolution of $1/2 \lambda$ in a more common reverberant environment, which is rooms. Sprik R. [17] then used the TR method based on two microphones to perform the SSL experiment on a pulse with a center frequency of 1750 Hz, and the results demonstrate its applicability in reverberant environments. Researchers have proposed several solutions from different perspectives to the half-wavelength restriction. Fink et al. [18] proposed the concept of an acoustic sink to break through the $\lambda/2$ limitation. Conti et al. [19] surrounded an audible sound source between 100 and 200 Hz with an array of microphones to realize near-field TR, and $\lambda/20$ resolution was obtained. Catheline et al. [20] studied an SSL model in the special reverberant environment of the skull based on the TR method and predicted that the model could be applied to microphone arrays. Padois et al. [21] implemented the localization of a sound source with a bandwidth of 1.5–18 kHz in a wind tunnel. However, there is no echo inside the tunnel. Mimani [22] proposed a TR method on aeroacoustic sources with 3000 Hz and 500–3000 Hz for the presence of a reflecting surface and demonstrated its effectiveness by comparing the simulation with the traditional beamforming method. Ma et al. [23] achieved $1/4 \lambda$ resolution for a pulse with a center frequency of 9000 Hz by placing subwavelength scatters. Li [24] proposed a cross-spectrum TR method corrected in the wavenumber domain. Numerical simulation and experimental results show that the spatial resolution of the proposed method can reach $1/34 \lambda$ at 100 Hz under a low SNR (0 dB) in the near field.

Currently, methods that break through the $1/2 \lambda$ limit and significantly improve spatial resolution are applied in the near field. When the sound source is several wavelengths or even far away from the microphone array, i.e., far field, although the SSL method has broken through the limit of $1/2 \lambda$, it has not improved much. Therefore, it is crucial to seek new breakthroughs.

Signals in nature can basically or approximately be expressed in sparse form. The reconstruction theory of sparse signals has been studied in depth by Donoho [25] and Candès et al. [26]. Subsequently, Baraniuk proposed the concept of compressive sensing [27]. Recently, a sparse signal representation method based on compressive sensing (CS) was introduced to realize DOA estimation when the number of sound sources is unknown in a complex environment [28]. The signal was compressed and sampled at a low rate to considerably reduce the number of samples and improve computational efficiency [25]. In general, sound sources are considered point sources in reverberant environments, and the number of sound sources is limited. The problem of SSL in reverberant environments can be expressed by sparse approximation model [29]. Chardon [30] proposed a sparse model for a narrowband SSL in an unknown reverberant field. A spatial resolution of $3/5 \lambda$ is obtained for a sound source at 30,631 Hz. Jiang et al. [31] used CS to accurately estimate the position of the sinusoidal signal source at 16 kHz, and the accuracy was higher than that of the generalized cross-correlation method. Simard et al. [32] constructed an SSL sparse model using Green's function. The experimental results show that for a sound source with 5000 Hz, the CS method can obtain a spatial resolution of $7/20 \lambda$, and the localization

results of the beamforming algorithm are fuzzy at this resolution. Koyama et al. [33,34] demonstrated that the problem of broadband SSL can be transformed into the problem of joint sparse support recovery in terms of frequency. Chu et al. [35] further extended the sparse constraint to achieve SSL with a bandwidth of 2400–2600 Hz and achieved a spatial resolution of $9/25 \lambda$. Tipping et al. [36] proposed a sparse Bayesian learning (SBL) algorithm based on a correlation vector machine for a sparse linear regression model. Zhang et al. [37] demonstrated the effectiveness of the SBL method for DOA estimation of quasi-stationary signals at an off-grid position by simulation. Bai et al. [38] solved the reverberation problem using the SBL approach and demonstrated by simulation that the method is more robust than other methods in noisy environments and with grid errors. Zheng et al. [39] proposed an SBL method to resolve coherent sources by simulation in impulsive noise environments. Qin et al. [40] completed audible SSL using L_1 minimization in reverberant and noisy environments. For the music signal, the array with 12 microphones has a root-mean-square error of 0.2 m.

The existing research results show that although the SSL method based on TR is suitable for reverberant environments, the localization performance of sound sources with low-frequency bands needs to be improved. There are few studies based on CS in a strong reverberant environment for wideband SSL, and the common signals in reality are broadband. Therefore, the study of broadband SSL in a reverberant environment is also a hot issue. In summary, the joint use of TR and CS theory has attracted the attention of scholars. It was first introduced into multiple-input multiple-output radar Doppler estimation and has been demonstrated to have excellent performance [41]. In the field of underwater acoustic communication, Jiang et al. [42] simulated the channel as a sparse set consisting of constant and slowly time-varying supports to derive the Kalman Filtered CS channel estimation algorithm for TR communication under the framework of dynamic CS. Experimental results demonstrate the validity of the algorithm. SBL is the most commonly used signal reconstruction method, which employs a grid model to construct an overcomplete dictionary and then uses the expectation maximization (EM) principle to update parameters iteratively [43]; it has better performance and robustness than conventional SSL methods under less snapshots and low SNR [44].

Scholars have used TR and SBL methods to conduct localization research on broadband signals, but they either focus on the high-frequency range or the localization results in the low-frequency range are not extremely satisfactory. Currently, there is a lack of systematic and comprehensive research on broadband low-frequency SSL in reverberation environments. Therefore, we propose a novel SSL method based on TR and SBL, which is represented as TR-SBL, to overcome the limitations of poor spatial resolution in a reverberant environment and low SNR. The process of TR and retransmission is carried out on the computer; thus, the method is passive localization. The study is organized as follows. Section 2 discusses the SSL model in reverberant environments. Section 3 describes the numerical simulations performed to validate the proposed method. Section 4 reports the actual data experiments in a reverberation chamber and the analysis of the results. Finally, Section 5 presents the conclusions.

2. SSL Model in a Reverberant Environment

Consider a rectangular enclosed space to simulate the sound source propagation model in a reverberant environment, as shown in Figure 1. The acoustic signal received by the m -th microphone is as follows:

$$y_m(t) = x(t) \otimes h_m(r_s, r_m, t) + n(t) \quad (1)$$

where $x(t)$ is the sound source signal, $h_m(r_s, r_m, t)$ is the room impulse response (RIR) from the sound source position r_s to the microphone position r_m , $\otimes$ is the convolution, m is the m -th microphone, and $n(t)$ is the background noise, which is generally assumed to be independent of the sound source.

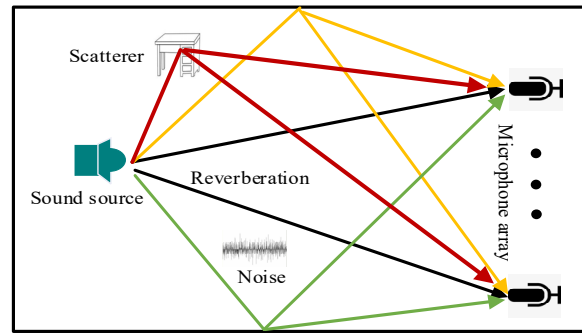


Figure 1. Sound propagation model in a reverberant environment.

As can be seen from Figure 1, RIR has not only direct sound (black line) but also reflected sound (green line, yellow line, and red line), which contains all the information of the sound field and can accurately reflect the acoustic environment (direction, distance, and sense of space). Therefore, RIR is a specific manifestation of reverberation. The reverberant environment is time-varying; obstacles, people, etc., may be different from the simulated environment, which requires strong robustness of the SSL algorithm.

When the received data are a broadband signal, the broadband processor can use the sound source information contained in multiple frequencies compared with the narrowband processor, thereby improving the SSL accuracy. The frequency domain form of Equation (1) is as follows:

$$\mathbf{Y}_m(f) = \mathbf{S}(f)\mathbf{H}_m(r_m, r_s, f) + \mathbf{N}(f) \quad (2)$$

where f is the frequency, $\mathbf{Y}_m(f) = [Y_m(f_1), Y_m(f_2), \dots, Y_m(f_F)]$, F stands for the number of frequency bands, and $\mathbf{S}(f) = [S(f_1), S(f_2), \dots, S(f_F)]$.

The TR processing of Equation (2) can be obtained according to the frequency domain phase conjugation TR as follows:

$$\mathbf{Y}_m^*(f) = \mathbf{S}^*(f)\mathbf{H}_m^*(r_m, r_s, f) + \mathbf{N}(f) \quad (3)$$

The signal $\mathbf{Y}_m^*(f)$ is reemitted back into the medium along the virtual channel $\mathbf{V}_m(r_{vs}, r_m, f)$. The overall diagram of time reversal is shown in Figure 2.

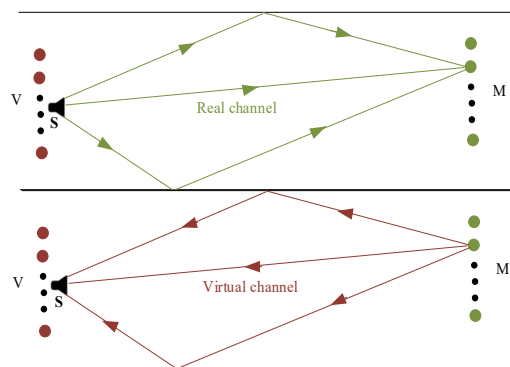


Figure 2. Diagram of time reversal.

As shown in Figure 2, there is another microphone array $\mathbf{V}$ at the initial sound source for receiving the signal propagated through the virtual channel after time reversal, and the signal received by the array $\mathbf{V}$ is

$$\begin{aligned} \mathbf{Z}_m(f) &= \mathbf{Y}_m^*(f)\mathbf{V}_m(r_{vs}, r_m, f) + \mathbf{N}(f) \\ &= \mathbf{S}^*(f)\mathbf{H}_m^*(r_m, r_s, f)\mathbf{V}_m(r_{vs}, r_m, f) + \mathbf{N}(f) \end{aligned} \quad (4)$$

Virtual channel refers to the channel propagated by the time reversal signal. According to the spatial reciprocity theorem, the time reversal signal will be refocused on the initial sound source position along the original path, but in practical application, the process of time reversal and retransmission is carried out on the computer; thus, the method is passive localization. Since the position of the sound source is unknown, so is the virtual channel. The possible position of the sound source is defined as the area of interest, and several grids are divided in this area. It is assumed that the position of the center of each grid is the position of the sound source, and the propagation channel of the microphone array to each grid is a virtual channel, which is obtained by the image method.

Once the virtual channel matches the real channel, $\mathbf{H}_m^*(r_m, r_s, f) \mathbf{V}_m(r_{vs}, r_m, f) = \mathbf{H}_m^*(r_m, r_s, f) \mathbf{H}_m(r_m, r_s, f)$, it is a real, even, and positive function with the maximum space gain of each element. This is equivalent to the spatial matching filtering of each array element. Finally, the focusing energy of each array element is added to obtain the SSL result graph, as shown in Equation (5), which is similar to the beamforming technology. Beamforming achieves spatial filtering through time delay or phase matching search and compensation of each array, whereas the TR method achieves spatial filtering through TR channel matching of each array.

$$R(f) = \sum_{m=1}^M \sum_{f=1}^F Z_m(f) Z_m^*(f) \quad (5)$$

Let us take into account a single microphone as an example of the TR process. The room size was set to 4 m × 2 m × 3 m, and the sound absorption coefficient of the six walls was 0.2. A Gaussian pulse signal with a center frequency $f_0 = 9.5$ kHz and bandwidth of 750 Hz was used as the original sound source, and the waveform is shown in Figure 3a. The coordinates of the sound source and microphone are (2, 1, 1.5) m and (3, 1, 1.5) m, respectively. The signal received by the microphone is shown in Figure 3b. The received signal is reversed along the time axis, and a TR signal is obtained, as shown in Figure 3c. Subsequently, we follow the principle that the first arrived signal is sent later, the last arrived signal is transmitted first, and it is retransmitted back to the medium. The second received signal is processed according to Equation (5). When the signal returns along the initial path, a focused signal can be formed at the original position, as shown in Figure 3d. Suppose that the signal after the TR process does not return according to the initial path, then the return path is between (3, 1, 1.5) m, (1.9, 1, 1.5) m, and (0.05, 0.5, 1.0) m. The focused signal is shown in Figure 3e,f. It can be seen from the last three subgraphs that (1) when the virtual channel is the same as the initial channel, there is an obvious focus peak and (2) when the virtual channel is not the same as the initial channel but is extremely close to the initial channel, the focused signal will also have a peak value. However, the second peak value in Figure 3e is larger than that in Figure 3d; therefore, (3) when the virtual channel is far from the initial channel, the peak value is not obvious, and even the focus is achieved in other positions.

To solve the problem that the sidelobe value of the TR method is relatively large, we use the SBL method for postprocessing. The area of interest in reverberant environments is divided into L grids, and it is assumed that there may be a sound source on each grid. There are N sound sources, and with $L \gg N$, the signal satisfies the sparse hypothesis, as shown in Figure 4. The matrix formed by the RIR of all sound sources to the microphone array forms an overcomplete basis. The more grids there are, the smaller the spacing between them, the higher the localization accuracy, but the larger the corresponding calculation amount.

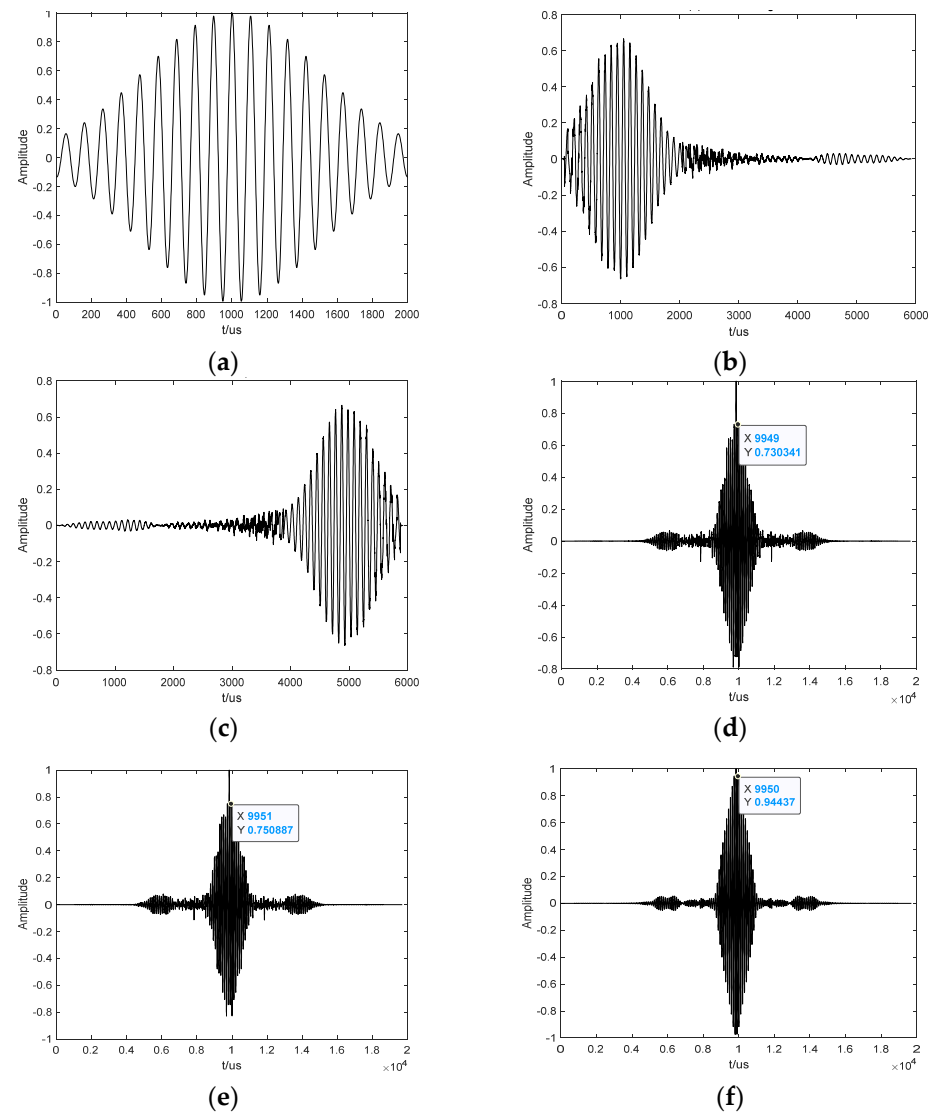


Figure 3. The entire process of traditional TR. (a) Gaussian pulse; (b) received signal; (c) TR signal; (d) focused signal at a non-sound-source position; (e) focused signal at a non-sound-source position; (f) focused signal at the position of the sound source.

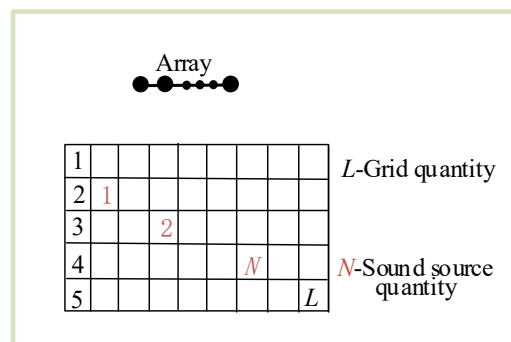


Figure 4. Spatial sparse representation.

The frequency domains of the TR signals received by the microphone array are as follows:

$$\mathbf{Z}_f = \begin{bmatrix} \mathbf{S}_1^*(f) \\ \mathbf{S}_2^*(f) \\ \vdots \\ \mathbf{S}_L^*(f) \end{bmatrix}^T \cdot \begin{bmatrix} H_{11}^*(f) & H_{12}^*(f) & \cdots & H_{1L}^*(f) \\ H_{21}^*(f) & H_{22}^*(f) & \cdots & H_{2L}^*(f) \\ \vdots & \vdots & \ddots & \vdots \\ H_{M1}^*(f) & H_{M2}^*(f) & \cdots & H_{ML}^*(f) \end{bmatrix}^T \cdot \begin{bmatrix} V_{11}(f) & V_{12}(f) & \cdots & V_{1L}(f) \\ V_{21}(f) & V_{22}(f) & \cdots & V_{2L}(f) \\ \vdots & \vdots & \ddots & \vdots \\ V_{M1}(f) & V_{M2}(f) & \cdots & V_{ML}(f) \end{bmatrix} + \mathbf{N}(f) \quad (6)$$

$$= \mathbf{A}_f \cdot \bar{\mathbf{S}}_f + \mathbf{N}_f$$

where $\mathbf{A}_f = (\mathbf{H}^*(f))^T \cdot \mathbf{V}(f)$ and $\bar{\mathbf{S}}_f = [\mathbf{S}_1^*(f), \mathbf{S}_2^*(f), \dots, \mathbf{S}_L^*(f)]^T$. The dimensions of each of these terms are as follows: $\mathbf{Z}_f \in \mathbb{C}^{L \times 1}$, $\mathbf{A}_f \in L \times L$, and $\bar{\mathbf{S}}_f \in L \times 1$. $\mathbf{A}_f$ is a dictionary composed of redundant RIR. Equation (6) is an underdetermined equation.

1. Noise model

The noise is assumed to conform to a complex normal distribution with the precision α_0 , so the likelihood function of Equation (6) is as follows:

$$\mathbf{Z}_f | \bar{\mathbf{S}}_f, \alpha_0 \sim \mathcal{CN}(\mathbf{A}_f \bar{\mathbf{S}}_f, \alpha_0^{-1} \mathbf{I}) \quad (7)$$

where α_0 is an unknown parameter that is assumed to obey the distribution of gamma with hyperparameters a and b [45], the symbol $\mathcal{CN}$ denotes the complex normal distribution, and $\mathbf{I}$ is an identity matrix. A gamma prior with hyperparameters a and b is further imposed on α_0

$$p(\alpha_0 | a, b) = \Gamma(\alpha_0 | a, b) \quad (8)$$

2. Sparse signal model

For the sparse signal $\bar{\mathbf{S}}_f$, we adopt a two-stage hierarchical sparse prior model to describe it:

$$\bar{\mathbf{S}}_f | \boldsymbol{\alpha} \sim \mathcal{CN}(0, \boldsymbol{\Lambda}^{-1} \mathbf{I}) \quad (9)$$

$$p(\boldsymbol{\alpha} | \varphi) = \prod_{n=1}^N \Gamma(\alpha_n | 1, \varphi) \quad (10)$$

where $\boldsymbol{\Lambda} = \text{diag}(\boldsymbol{\alpha})$ and $\{\bar{\mathbf{S}}_f\}_{f=1:F}$. According to conventional SBL, parameters a , b , and φ in Equations (8) and (10) are commonly set to be 10^{-6} [36].

According to the above analysis, the joint probability density function can be expressed as:

$$p(\{\bar{\mathbf{S}}_f\}_{f=1:F}, \{\mathbf{Z}_f\}_{f=1:F}, \alpha_0, \boldsymbol{\alpha}) = p(\{\mathbf{Z}_f\}_{f=1:F} | \{\bar{\mathbf{S}}_f\}_{f=1:F}, \alpha_0) p(\{\bar{\mathbf{S}}_f\}_{f=1:F} | \boldsymbol{\alpha}) p(\boldsymbol{\alpha}) p(\alpha_0) \quad (11)$$

3. Bayesian Inference

By treating $\bar{\mathbf{S}}_f$ as a hidden variable, whose posterior distribution is as follows:

$$p(\bar{\mathbf{S}}_f | \mathbf{Z}_f, \alpha_0, \boldsymbol{\alpha}) = \prod_{n=1}^N \mathcal{CN}(\bar{\mathbf{S}}_f | \boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (12)$$

with

$$\begin{aligned} \bar{\mathbf{S}}_f &= \boldsymbol{\mu} \\ \boldsymbol{\mu} &= \alpha_0 \boldsymbol{\Sigma} \mathbf{A}_f^H \mathbf{Z}_f \\ \boldsymbol{\Sigma}_f &= (\Delta + \alpha_0 \mathbf{A}_f^H \mathbf{Z}_f)^{-1} \end{aligned} \quad (13)$$

calculations of μ and Σ_f , need estimates of the hyper-parameters α_0 and α . The maximum posterior probability is used to calculate α_0 and α .

$$p(\mathbf{Z}_f, \alpha_0, \alpha) = p(\alpha_0, \alpha | \mathbf{Z}_f) p(\mathbf{Z}_f) \quad (14)$$

Each parameter in $p(\mathbf{Z}_f)$ is independent, and maximizing $p(\alpha_0, \alpha | \mathbf{Z}_f)$ is the same as maximizing $p(\mathbf{Z}_f, \alpha_0, \alpha)$.

Since $p(\bar{\mathbf{S}}_f, \alpha_0, \alpha | \mathbf{Z}_f)$ cannot be explicitly calculated, we adopt variational Bayesian inference (VBI) to approximate the posterior and infer hidden variables. The optimization problem within a distribution $q(\Delta)$ with a factorized form is assumed to approximate the true posterior, and based on Kullback–Leibler divergence (KL), true posterior $p(\Delta | \mathbf{z})$ can be expressed as follows:

$$\begin{aligned} q^*(\Delta) &= \arg \min_{q(\Delta)} \text{KL}(q(\Delta) | p(\Delta | \mathbf{z})) \\ &= \arg \min_{q(\Delta)} \left(\int q(\Delta) \log \frac{q(\Delta)}{p(\Delta | \mathbf{z})} d\Delta \right) \end{aligned} \quad (15)$$

where $q^*(\Delta)$ stands for estimated value.

$q(\Delta)$ can be written as follows:

$$q(\Delta) = q\left(\left\{\left\{\bar{\mathbf{S}}_f\right\}_{f=1,\dots,F}, \alpha, \alpha_0\right\}\right) = q(\alpha_0)q(\alpha) \prod_f q(\bar{\mathbf{S}}_f) \quad (16)$$

On the basis of the above derivation, the optimal distribution of Equation (15) can be expressed as follows:

$$\ln q^*(\Delta_k) = E\left(\ln p(\mathbf{z}_f, \Delta)\right)_{q(\Delta \setminus \Delta_k)} \quad (17)$$

where $E\left(\ln p(\mathbf{z}_f, \Delta)\right)_{q(\Delta \setminus \Delta_k)}$ denotes the expectation with respect to $q(\Delta \setminus \Delta_k)$, and $\Delta \setminus \Delta_k$ denotes the set Δ without Δ_k .

4. Hidden variables updating

Hidden variables are updated using the VBI technique. The updating rule for each type of variable is derived from Equation (17). The updated estimations of α and α_0 are as follows:

$$\begin{aligned} \alpha^{new} &= \frac{c+F}{d+\sum_f \text{diag}(\mu_f \mu_f^H + \Sigma_f)} \\ \alpha_0^{new} &= \frac{a+MF}{b+\sum_f \left\{ \|\mathbf{z}_f - \mathbf{A}_f \mu_f\|_2^2 + \text{Tr}(\mathbf{A}_f \Sigma_f \mathbf{A}_f^H) \right\}} \end{aligned} \quad (18)$$

where Tr is the operation of trace.

The parameters are always updated and stopped when the maximum error between two consecutive values is quite small or the number of iterations has reached the preset value. We can obtain final μ , so the positions of the sound sources can be estimated from it.

3. Simulations and Analysis

In this section, first, the effectiveness of the proposed TR–SBL method is verified by comparing it with the TR method and L_1 minimization. Second, the SSL performance of the proposed TR–SBL method is verified under different parameter settings, including different bandwidths, a different number of microphones, SNRs, reverberation times, and off-grid conditions. These methods are implemented on the MATLAB R2022a platform installed on a computer with the following parameters: Intel (R) Core (TM) i7-13700KF CPU @3.40 GHz. The computer manufacturer is Intel, from Xi'an, China.

First, select a rectangular room with a size of $6\text{ m} \times 4\text{ m} \times 3\text{ m}$ as the reverberant environment. The array comprises 20 omnidirectional microphones with intervals of 0.1 m . The array is distributed along the x -axis from 1.0 to 2.9 m , and the y -axis and z -axis coordinates are 1.0 m . We have selected the area of interest at a certain distance from the array, aiming to position it as centrally as possible within the entire room. This is performed to avoid proximity to walls or corners, which could lead to destructive interference or phase addition effects. The coordinates of the area of interest are 2.55 – 3.50 m on the x -axis, 2.0 – 2.95 m on the y -axis, and 1.0 m on the z -axis. The area of interest is divided into 400 grids with intervals of 0.05 m . The spatial impulse response from the center of each grid to the microphone array is simulated using the Image method. The bandwidths of the sound source are 125 – 250 , 250 – 500 , 500 – 1000 , and 125 – 1000 . The SNR is 20 dB . The coordinates of the sound source are $(2.55, 2.75, 1.00)\text{ m}$. A sketch of the computational domain with the position of the microphone array and sound source is shown in Figure 5.

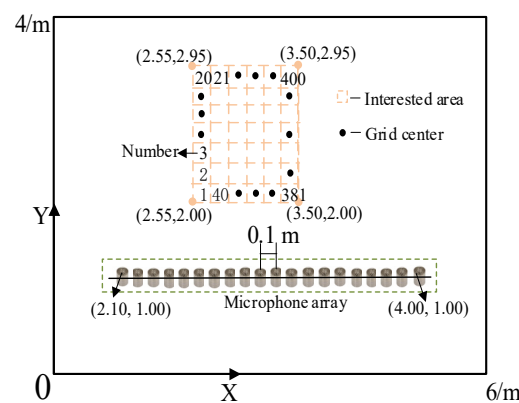


Figure 5. Sketch of the computational domain with the position of the mic and sound source.

The reverberation time is usually denoted by T_{60} , which is defined as when the sound field reaches a steady state and the sound source stops sounding. It takes time for the sound pressure level to decrease by 60 dB [46]. The unit is second. The reverberation time T_{60} of an ordinary room is determined by the volume V , surface area S , and sound absorption coefficient α .

$$T_{60} = 0.161 \frac{V}{-S \log_e(1 - \bar{\alpha})} \quad (19)$$

where $\bar{\alpha}$ is the average sound absorption coefficient. In this study, we assume that the sound absorption coefficient of all walls is the same, which is α .

3.1. Comparison with TR and L_1 Norm Minimization

To verify the effectiveness of the TR–SBL method, the conventional TR method and L_1 norm minimization are compared. All algorithms are implemented on the MATLAB platform. The bandwidth of the sound source is 125 – 1000 Hz , and the sound source is located at $(3.00, 2.90)\text{ m}$. The sound absorption coefficient α is 0.5 . T_{60} is 153.9 ms . Figure 6 shows the SSL results obtained using the three methods. The SSL results are represented by the focusing gain FG :

$$FG = 10 \log_{10} \left(\sum_{i=1}^M R(f)_i / R(f)_{imax} \right) \quad (20)$$

The actual position of the sound source is indicated by a red circle. The SSL results of the TR and TR–SBL methods are shown in Figures 6a and 6f, respectively, and the rest of the subfigures are SSL results of L_1 norm minimization under $\lambda = 10^{-10}, 10^1, 10^{-1}, 10^{-8}$.

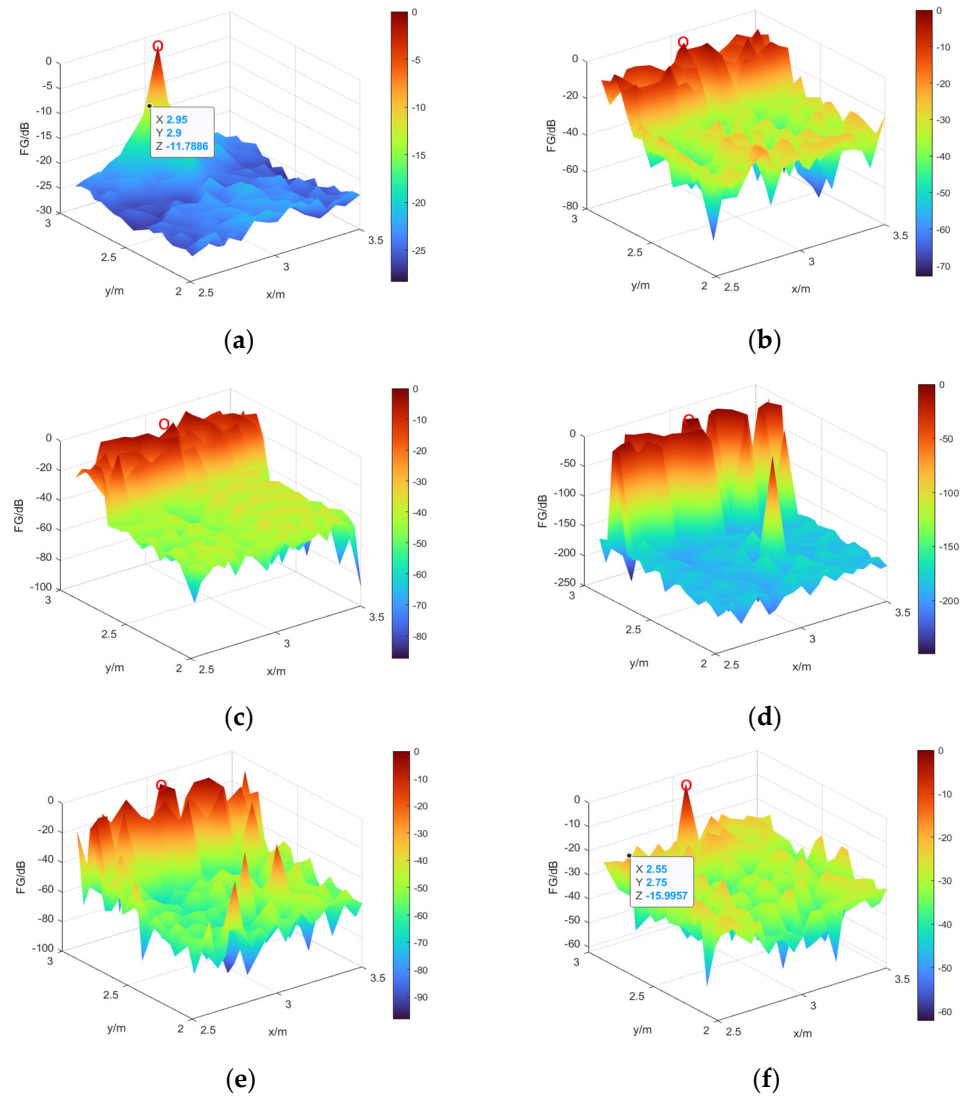


Figure 6. SSL results of TR method, L_1 norm minimization, and TR-SBL method. (a) TR; (b) L_1 norm, $\lambda = 10^{-10}$; (c) L_1 norm, $\lambda = 10^1$; (d) L_1 norm, $\lambda = 10^{-1}$; (e) L_1 norm, $\lambda = 10^{-8}$; (f) TR-SBL.

As can be seen from Figure 6, the SSL results obtained using different methods vary greatly. In terms of the obtained position, both the traditional TR method and the TR-SBL method can obtain accurate positions. The difference between the two methods is mainly the value of the second peak and the remaining sidelobe. The second peak value of the proposed TR-SBL method is 26% lower than that of the TR method, and the performance is greatly improved. The SSL position obtained from L_1 norm is greatly affected by the value of parameter λ . When the parameter λ selection is appropriate, such as 10^{-8} , an accurate SSL position can be obtained, but the sidelobe value is quite large. When the parameter λ selection is not appropriate, such as 10^{-10} , 10^1 , and 10^{-1} , the positions of the sound source obtained by L_1 norm are all (3.05 and 2.85), and the corresponding SSL error is 0.07 m. The selection of parameter λ is irregular; thus, we can only keep trying. From the peak value perspective, the proposed TR-SBL method obtains the highest performance and the lowest sidelobe value.

We take into account the root-mean-square error (RMSE) as an evaluation criterion for further analysis, which is defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (r_{s,i} - r_{a,i})^2} \quad (21)$$

where N is the number of Monte Carlo experiments, $r_{s,i} = (x_{s,i}, y_{s,i}, z_{s,i})$ is the position obtained by simulation calculation, and $r_{a,i} = (x_{a,i}, y_{a,i}, z_{a,i})$ is the actual position.

Subsequently, we compare the RMSEs of the three algorithms at different SNRs, mainly in the case of low SNR. Twenty Monte Carlo experiments were performed for each SNR, and $\lambda = 10^{-8}$; the results are shown in Table 1.

Table 1. RMSEs of three methods.

SNR/dB RMSE/m	TR	L_1 Norm	TR-SBL
−30	0.75	0.58	0
−20	0.62	0.52	0
−15	0	0.55	0
−10	0	0.49	0

It can be seen from the results of the RMSEs that under the same SNR, the traditional TR method has the largest RMSE, whereas the proposed TR-SBL method has the smallest; all are 0 cm. In the case of low SNR, −30 dB, the result of the TR method is 0.75 m, and that of the L_1 norm method is 0.58 m. With an increase in SNR, the RMSE of the two traditional methods is reduced. When the SNR is greater than −20, the RMSE of the TR method decreases to 0. The RMSE of the L_1 norm method decreases with an increase in the SNR, but it does not decrease to 0.

As is well known, the more grids there are, the greater the computational workload. In practical applications, the time-consuming nature of the algorithm is a factor that cannot be ignored, which determines whether it is suitable for engineering applications. Therefore, Table 2 compares the elapsed time of different methods under the same conditions and different grid numbers. The SNR is 20 dB and λ is 10^{-8} .

Table 2. Elapsed time required by different grid numbers.

Grid Time/s	TR	L_1 Norm	TR-SBL
6×6	1.02	1.73	1.20
8×8	1.05	1.74	1.22
9×9	1.10	1.77	1.23
10×10	1.12	1.78	1.23
10×20	1.29	1.79	1.25
20×20	1.66	1.87	1.51

As can be seen from Table 2, as the number of grids increases, so does the corresponding elapsed time; however, the increasing range is not significant. When the number of grids is the same, the TR method has the optimal performance, that is, the shortest elapsed time; L_1 norm minimization has the worst performance, that is, the longest elapsed time; and the elapsed time required by the proposed method falls between the two.

3.2. Analysis under Different Parameter Settings

We investigated the SSL performance of TR-SBL methods under different bandwidths, SNRs, reverberation times, and off-grid conditions in this section. In addition to the RMSE mentioned above, we added the SSL probability of accurate localization index to evaluate the performance, which is defined as

$$\Theta = \frac{C(\text{Correct localization number})}{M(\text{Monte Carlo experiments number})} \quad (22)$$

3.2.1. Different Bandwidths

The types of broadband signals in a reverberant environment are not fixed, but they are basically low- and medium-frequency signals. Therefore, we conducted research on SSL performance at different bandwidths to verify its robustness.

Broadband signals with bandwidths of 125–250, 250–500, 500–1000, and 125–1000 Hz were selected. Figure 7 shows the SSL results of the three methods under different bandwidths.

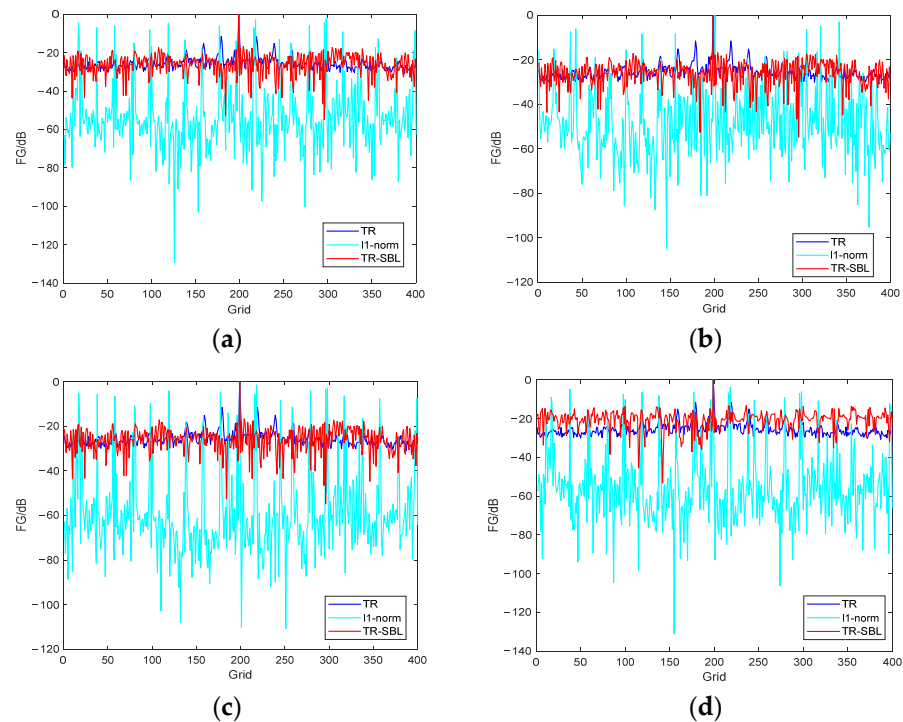


Figure 7. SSL results of the three methods under different bandwidths. (a) 125–250 Hz; (b) 250–500 Hz; (c) 500–1000 Hz; (d) 125–1000 Hz.

As can be seen from Figure 7, that the maximum and minimum sidelobe values come from the L_1 norm method for the same sound source. Although the L_1 norm can obtain the correct position, the sidelobe values are quite large, close to 0 dB. The maximum sidelobe value obtained by the TR method is approximately -10 dB, and the maximum sidelobe value obtained by the proposed TR-SBL method is approximately -20 dB. The overall sidelobe value of the proposed method is also lower than that of the TR method. In general, the proposed method has better SSL performance.

We further investigated the performance of the method when there are two sound sources in space, and the results are shown in Figure 8. Among the methods, the λ value of the L_1 norm method is 10^{-8} . The actual position coordinates of the two sound sources are (3.00, 2.90, 1.00) and (3.25, 2.90, 1.00) m, respectively.

As can be seen from Figure 8 that the proposed TR-SBL method can accurately find the position of dual sound sources, and the sidelobe value is significantly lower than that of the TR method. The maximum sidelobe difference between the two methods is approximately 10 dB. The sidelobe value of the L_1 norm method is too high, resulting in an unclear display of the sound source position.

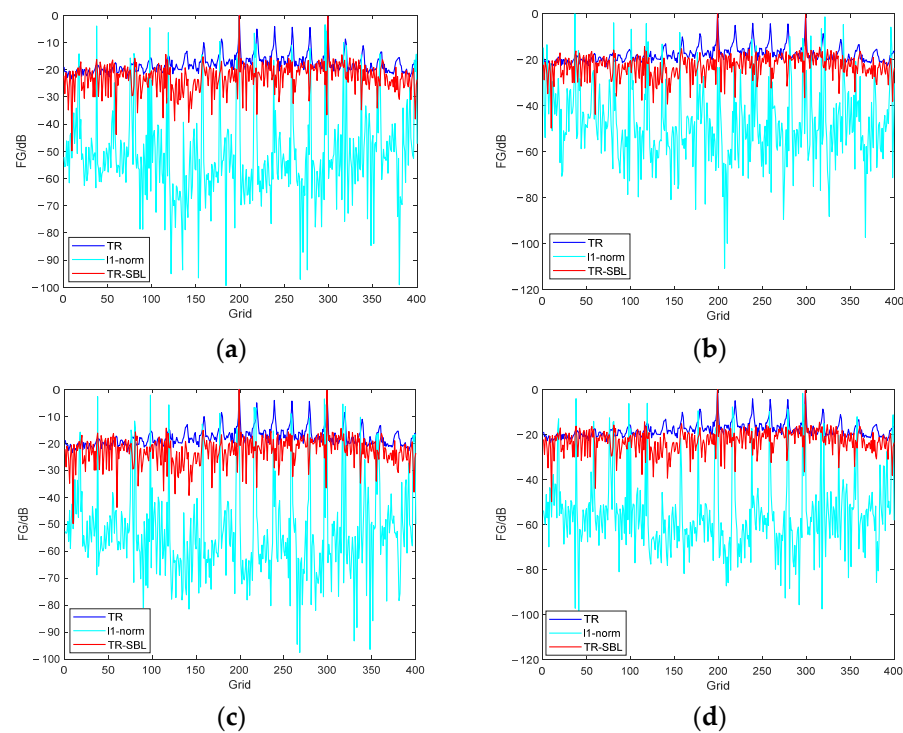


Figure 8. Comparison of SSL results of different methods for dual sound sources. (a) 125–250 Hz; (b) 250–500 Hz; (c) 500–1000 Hz; (d) 125–1000 Hz.

3.2.2. Different Numbers of Microphones

Generally, the more microphones in the array element, the greater the spatial gain and the better the SSL result. However, in engineering applications, what we seek is to obtain better results with as few microphones as possible. Therefore, we further investigate the robustness of the proposed method with different numbers of microphones. We take into account the number of microphones as $M = 20, 15, 10$, and 5 ; the SNR is 20 dB; and the sound source is 125–250 Hz. The SSL results are shown in Figure 9.

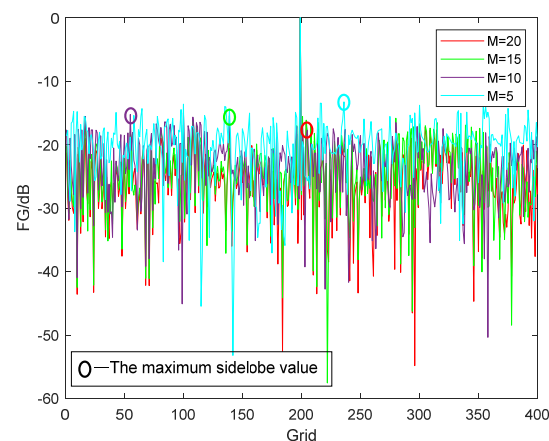


Figure 9. SSL results with different numbers of microphones.

Figure 9 shows that even if the number of microphones is reduced to five, the method can obtain the exact position of the sound source, which demonstrates that it has high robustness. The SSL results of different microphones are different in the sidelobe value. The smaller the number of microphones is, the higher the sidelobe value is.

3.2.3. Different SNRs

The noise in reverberant environments is also an important factor affecting the high spatial resolution of the SSL method. Therefore, the SSL performance of the proposed TR-SBL method is studied under different SNRs; -35 , -30 , -20 , -15 , and 0 dB are selected. The SSL probability of accurate localization Θ is used as an evaluation index. Twenty Monte Carlo experiments are conducted, and the SSL results are illustrated in Figure 10.

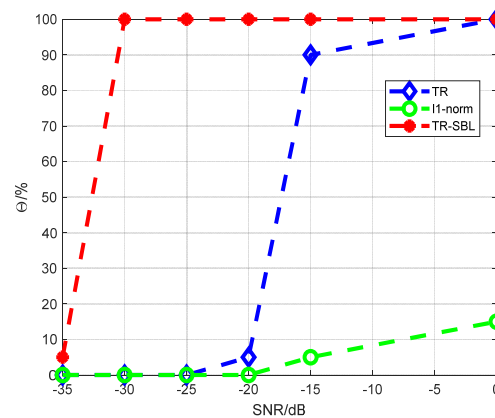


Figure 10. SSL probability of accurate localization Θ under different SNRs.

As shown in Figure 10, in the case of an extremely low SNR of -35 dB, although the SSL Θ of the proposed TR-SBL method is only 5%, the other two methods have a lower Θ , which is 0%. When the SNR rises to -30 dB, the proposed method can obtain 100% SSL Θ , whereas the Θ of the other two methods is 0%. When the SNR is 0 dB, the SSL Θ of the TR method reaches 100%. The accuracy of the L_1 norm method has been extremely low, and the highest is only 15%. In general, the proposed method can obtain a high Θ at a lower SNR.

3.2.4. Different Reverberation Times

The localization performance of the algorithm is greatly affected by reverberation. The reverberation time is generally determined by the sound absorption coefficient α and the room size. Rooms include walls, ceilings, and floors, and each of them may have different sound absorption coefficients. In the present study, we assume that each wall has the same sound absorption coefficient α .

We take into account 0.2, 0.4, and 0.6 as sound absorption coefficients for research, and their corresponding reverberation times are 0.3330, 0.1455, and 0.0811 s, respectively. Four kinds of broadband SSL simulation experiments are completed in the three reverberant environments. After 20 Monte Carlo experiments on 4 broadband sound sources, the final probability of accurate localization Θ is 100%, indicating that the proposed TR-SBL method is robust in a reverberant environment.

We also reviewed several SSL methods that have been applied in typical environments with noise and reverberation over the past two years, with some of them presented as examples: Jiajun Yan et al. [47] fed the data extracted from a time-domain feature-vector based on the sample covariance of microphone array data. The simulation results showed that the SSL error was 0.2 m when the SNR was 10 dB and the reverberation time T_{60} was 0.1 s. Guillermo et al. [48] proposed exploiting spatial diversity as an alternative approach to achieve robustness against reverberation. When the reverberation time was 0.2 s, the SSL error of 0.3 m was obtained. There is no mention of the SNR in this study. Compared with these two methods, the proposed method in this paper has satisfactory performance.

3.2.5. Off-Grid Conditions

In practical engineering applications, the position of the sound source is random and uncertain, meaning that the sound source may appear at any position in the room, but

if each position is searched once, the computational complexity is huge. Therefore, we need to study the robustness of the proposed TR-SBL method when the sound source is not in the center of the grid, which is an off-grid condition. Figure 11 provides a detailed description of the position of off-grid conditions in different situations. Positions that are not in the center of the grid are generally divided into two categories: (1) outside the area of interest, as shown by the $\blacklozenge$ in Figure 11, and (2) within the area of interest, as shown by the $\ast$ in Figure 11.

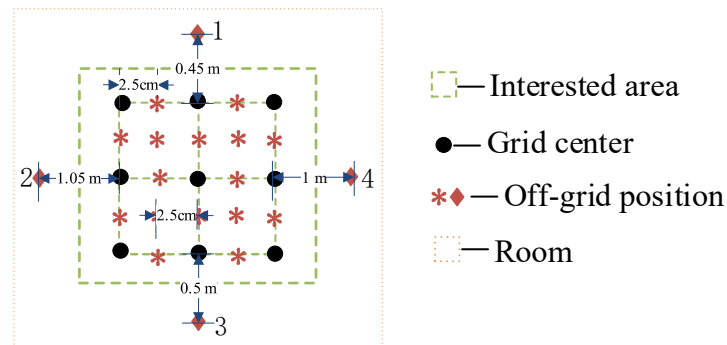


Figure 11. Diagrammatic sketch of different sound source positions.

The sound source bandwidth is 125–1000 Hz, the sound absorption coefficient α is 0.4, and the SNR is 20 dB. The distance from the position obtained by the TR-SBL algorithm to the actual position is used to represent the SSL result, as shown in Figure 12. Points 2 and 4 located outside the area of interest have the largest SSL results from the actual position: 1.9 and 1.44 m, respectively. The SSL results for points 1 and 3 are 0.56 and 0.72 m, respectively. For the points within the area of interest, the distance of points 8, 10, 14, 15, and 18 obtained by the TR-SBL algorithm is more than 0.5 m, and the farthest distance is 0.82 m. The distance of points 8, 10, 14, 15, and 18 is more than 0.5 m; the farthest distance is 0.82 m; and the distance of points 13, 19, and 20 is between 0.1 and 0.3 m. The remaining eight points are less than 0.03 m, which is basically negligible. In general, points outside the area of interest have greater distances than points inside the area of interest.

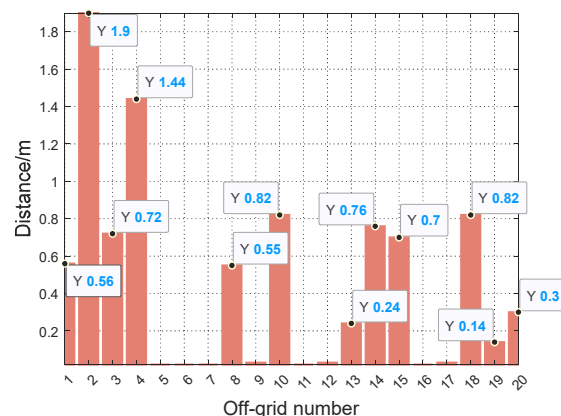


Figure 12. SSL results for off-grid conditions.

4. Real Data Experiments

To demonstrate the effectiveness of the proposed TR-SBL method, real data experiments are conducted in a real rectangular room with strong reverberation, i.e., a reverberation chamber. The corresponding details and results of these experiments are given below.

The sound sources are recorded by a uniform linear array consisting of 15 omnidirectional microphones. The microphone type is 4938 (Brüel & Kjær, Virum, Denmark), and the frequency response range is 4 Hz to 70 kHz. A photograph of the equipment in

a reverberation chamber is shown in Figure 13. The spherical sound source serves as the sounding device with a diameter of 0.5 m and the Pulse 3060 module as the receiving device (both from Brüel & Kjær).

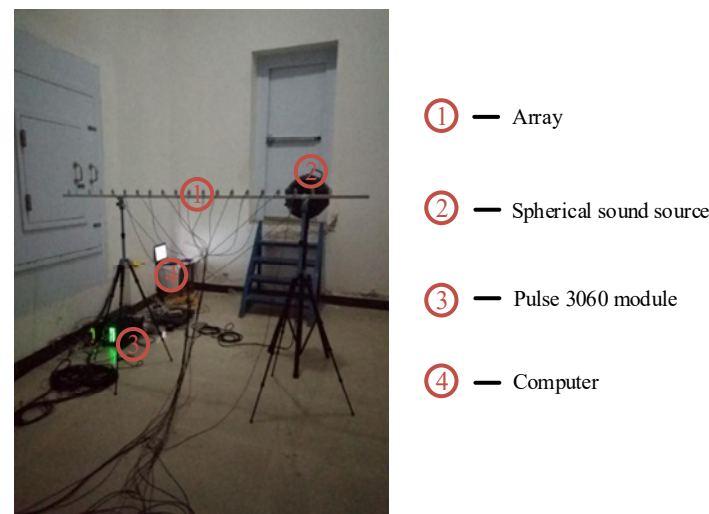


Figure 13. Photograph of equipment in a reverberation chamber.

The coordinates of each device and sound source are shown in Figure 14. The reverberation chamber size is $3.24 \text{ m} \times 5.73 \text{ m} \times 4.80 \text{ m}$. The array is arranged along $1.0\text{--}2.4 \text{ m}$ with a spacing of 0.1 m along the x -axis and 1.64 m along the y -axis. The z -axis coordinate, i.e., height, is 1.40 m . There are 15 microphones in total. The computer and pulse are placed in the corner, with negligible coordinates. The orange square represents the designated area of interest, has a height of 1.4 m , and is at the same level as the array. The range of the area of interest in the x -axis direction is $1.42\text{--}2.02 \text{ m}$, and the y -axis is $2.11\text{--}4.51 \text{ m}$. The area is divided into grids at intervals of 0.1 m , totaling 175 grids, which are represented by the orange circle. We collect signals from six positions inside the area of interest with coordinates: $(2.02, 2.11, 1.40)$, $(1.42, 2.11, 1.40)$, $(2.02, 3.31, 1.40)$, $(1.42, 3.31, 1.40)$, $(2.02, 4.51, 1.40)$, and $(1.42, 4.51, 1.40) \text{ m}$. We select signals with a bandwidth of $125\text{--}250$, $250\text{--}500$, $500\text{--}1000$, and $125\text{--}1000 \text{ Hz}$ as the sound source. Considering that the SNR in real-world environments is often much higher than in ideal conditions, the SNR is set to 20 dB .

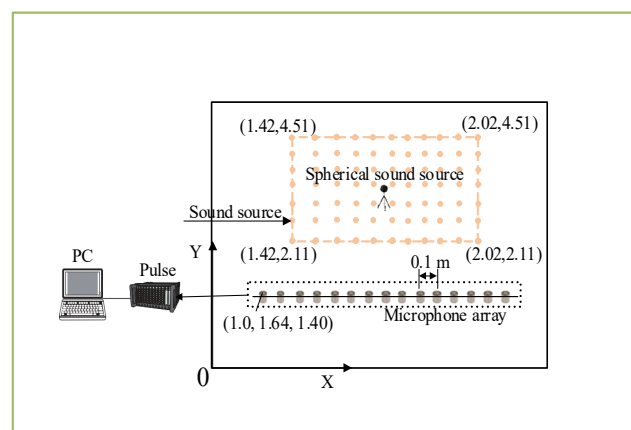


Figure 14. Schematic diagram of the array and sound source coordinates.

The SSL results obtained by processing signals ranging from 125 to 250 Hz are shown in Figure 15a, which are extremely blurry. To obtain clearer localization results, the received signals are whitened.

$$\hat{Z}(\omega) = Z(\omega) / |Z(\omega)| \quad (23)$$

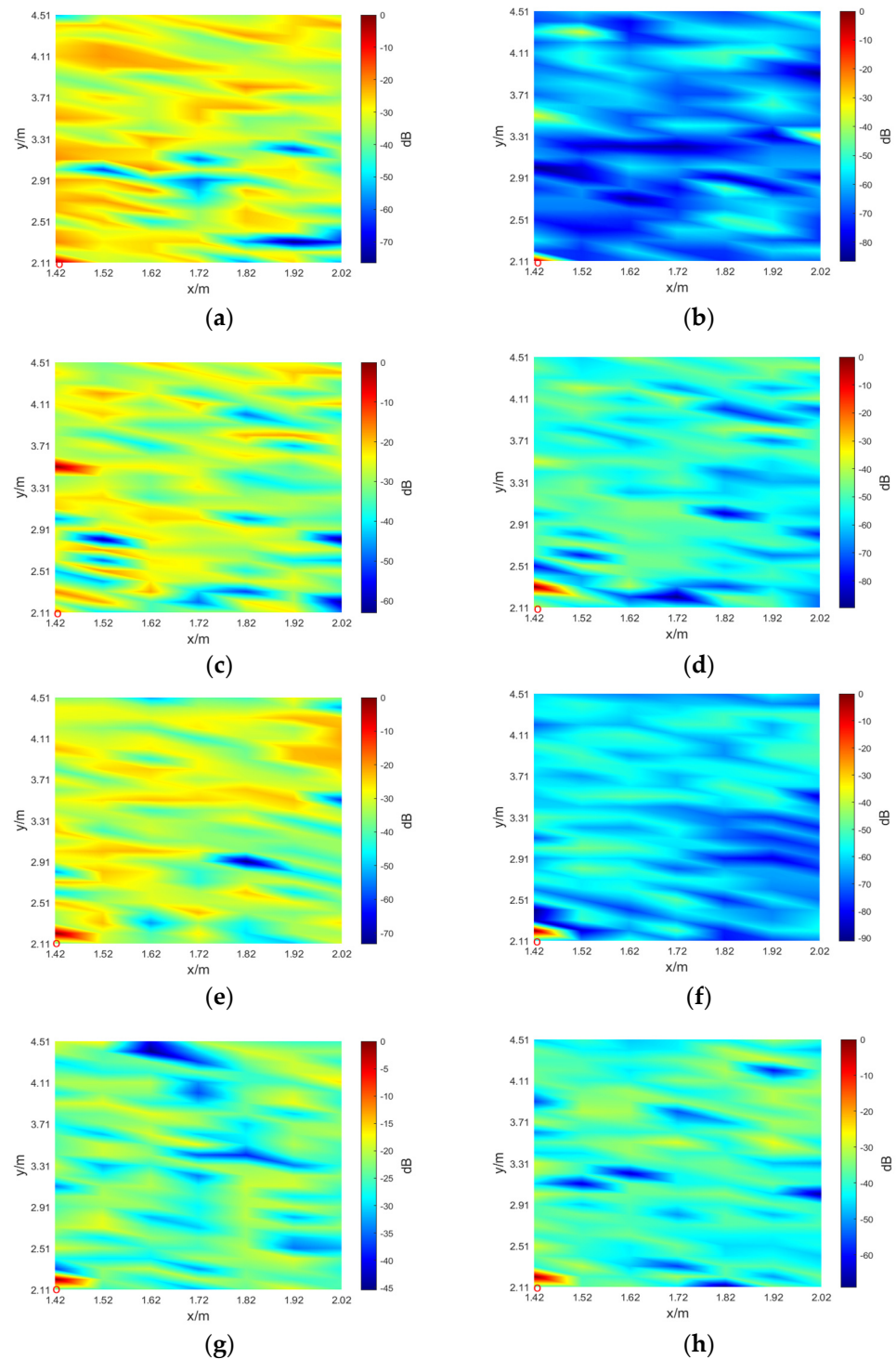


Figure 15. SSL results of the TR-SBL method. (a) Result of non-whitening for 125–250 Hz signal; (b) result of whitening for 125–250 Hz signal; (c) result of non-whitening for 250–500 Hz signal; (d) result of whitening for 250–500 Hz signal; (e) result of non-whitening for 500–1000 Hz signal; (f) result of whitening for 500–1000 Hz signal; (g) result of non-whitening for 125–1000 Hz signal; (h) result of whitening for 125–1000 Hz signal.

Figure 15 shows the SSL results of different sound sources before (figures on the left) and after (figures on the right) whitening. The red circle represents the true position of the sound source.

As can be seen from Figure 15, the SSL results of the proposed method for the non-whitening signals are far from ideal. In Figure 15a, the experimental result of the 125–250 Hz

signal is too fuzzy to find the true position of the sound source. After whitening, Figure 15b is much clearer, and the position of the sound source can be directly seen. The nonwhitening experimental results of the 250–500 Hz signal are far from the actual position, which is 1.4 m. However, after whitening, the spatial resolution can reach 0.2 m. The spatial resolution of the sound source at 500–1000 and 125–1000 Hz is the same, which is 0.1 m, regardless of whether it is nonwhitening or whitening, and the experimental results of the whitened signal are clearer and the sidelobe value is lower.

Subsequently, we further compared the SSL performance of the three methods. Taking into account the sound source of 250–500 Hz as an example, we demonstrate the results of the TR method and the L_1 norm method, where λ is 1, as shown in Figure 16.

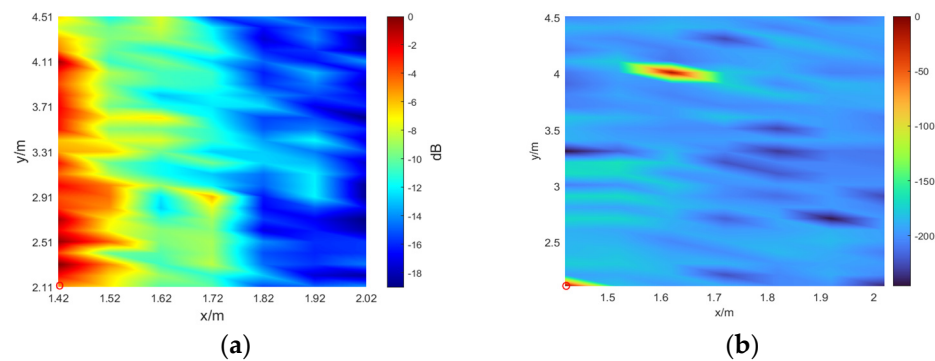


Figure 16. The SSL results of the TR method and the L_1 norm method. (a) TR; (b) L_1 norm.

From Figure 16, it can be seen that there is a certain deviation in the position of the sound source identified by the TR method, with a spatial resolution of 0.2 m. The L_1 norm method can obtain the accurate position of the sound source. It is obvious from Figure 16 that the common limitation of these two methods is that the sidelobe value is extremely large, meaning that the SSL result graph cannot be clearly displayed.

We process the collected signals from the six positions and take into account the RMSE as the index to evaluate the SSL performance of the three methods, as shown in Table 3.

Table 3. RMSE of different methods under different bandwidths.

Bandwidth/Hz RMSE/m	TR	L_1 Norm	TR-SBL
125–250	0.50	0.37	0.37
250–500	0.72	0.36	0.17
500–1000	0.59	0.38	0.10
125–1000	0.64	0.10	0.10

Overall, the RMSE of the proposed TR-SBL method is the smallest, and the RMSE of the TR method is the largest. At the low-frequency band of 125–250 Hz, the RMSE of the TR method is 0.50 m, corresponding to $7/25 \lambda$, whereas the other two methods are both 0.37 m, corresponding to $1/5 \lambda$, an improvement of 26%. For the frequency band of 250–500 Hz, the RMSE of the proposed method is reduced to 0.17 m, corresponding to $19/100 \lambda$, the TR method is increased, and the L_1 norm method has a slight change, which can be ignored. For the frequency band of 500–1000 Hz, the proposed method further reduces the RMSE to 0.10 m, corresponding to $11/50 \lambda$, whereas the TR method is reduced to 0.59 m, corresponding to 1.3λ , and the L_1 norm method has slight changes. For the frequency band of 125–1000 Hz, the RMSE of the proposed method and the L_1 norm method are both 0.10 m, corresponding to $9/50 \lambda$, whereas the TR method has an RMSE of 0.64 m, corresponding to 1.3λ , which is significant.

5. Conclusions

This study reformulates the SSL problem under TR and SBL, which aims to achieve high performance and strong robustness at broadband low-frequency sound sources in reverberant environments. The proposed TR-SBL method fully utilizes the antireverberation properties of the TR method and the super-resolution characteristics of the SBL method. By comparing it with the traditional TR method and L_1 norm minimization, it has been demonstrated that this method performs better and has stronger robustness. Simulation results indicate that the proposed TR-SBL method obtains an accurate position for a broadband low-frequency sound source successfully in a low SNR and reverberant environment. Real-data experimental results indicate that satisfactory SSL performance is attained by the TR-SBL method. When the frequency of the sound source is low, a spatial resolution of 0.1 m can be obtained by this method. In the future, problems such as the localization performance over the number and positions of the microphones should be investigated.

Author Contributions: Conceptualization, H.M.; methodology, H.M.; software, H.M.; validation, H.M.; formal analysis, H.M.; investigation, G.L., Z.L. and T.S.; resources, G.L.; data curation, H.M.; writing—original draft preparation, H.M.; writing—review and editing, H.M.; visualization, H.M.; supervision, T.S.; project administration, T.S.; funding acquisition, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Natural Science Foundation of China, grant number 62001286.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Nehorai, A.; Paldi, E. Acoustic vector-sensor array processing. *IEEE Trans. Signal Process.* **1994**, *42*, 2481–2491. [\[CrossRef\]](#)
2. Grumiaux, P.A.; Kitić, S.; Girin, L.; Guérin, A. A survey of sound source localization with deep learning methods. *J. Acoust. Soc. Am.* **2022**, *152*, 107–151. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Ma, W.; Bao, H.; Zhang, C.; Liu, X. Beamforming of phased microphone array for rotating sound source localization. *J. Sound Vib.* **2020**, *467*, 115064. [\[CrossRef\]](#)
4. Wang, X.; Huang, G.; Benesty, J.; Chen, J.; Cohen, I. Time difference of arrival estimation based on a Kronecker product decomposition. *IEEE Signal Process. Lett.* **2020**, *28*, 51–55. [\[CrossRef\]](#)
5. Meng, X.T.; Cao, B.X.; Yan, F.G.; Greco, M.; Gini, F.; Zhang, Y. Real-Valued MUSIC for Efficient Direction of Arrival Estimation With Arbitrary Arrays: Mirror Suppression and Resolution Improvement. *Signal Process.* **2023**, *202*, 108766. [\[CrossRef\]](#)
6. Chardon, G. Gridless covariance matrix fitting methods for three dimensional acoustical source localization. *J. Sound Vib.* **2023**, *551*, 117608. [\[CrossRef\]](#)
7. Wu, Y.; Ayyalasomayajula, R.; Bianco, M.J.; Bharadia, D.; Gerstoft, P. SSLIDE: Sound source localization for indoors based on deep learning. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Toronto, ON, Canada, 6–12 June 2021.
8. Ikeda, O. An image reconstruction algorithm using phase conjugation for diffraction-limited imaging in an inhomogeneous medium. *J. Acoust. Soc. Am.* **1989**, *85*, 1602–1606. [\[CrossRef\]](#)
9. Cassereau, D.; Wu, F.; Fink, M. Limits of self-focusing using closed time-reversal cavities and mirrors-theory and experiment. In Proceedings of the IEEE Symposium on Ultrasonics (IUS), Honolulu, HI, USA, 4–7 December 1990; pp. 1613–1618.
10. Fink, M. Time-reversal mirrors. *J. Phys. D Appl. Phys.* **1993**, *26*, 1333–1350. [\[CrossRef\]](#)
11. Prada, C.; Thomas, J.L.; Fink, M. The iterative time-reversal process: Analysis of the convergence. *J. Acoust. Soc. Am.* **1995**, *97*, 62–71. [\[CrossRef\]](#)
12. Fink, M.; Cassereau, D.; Derode, A. Time-reversed acoustics. *Rep. Prog. Phys.* **2000**, *63*, 1933. [\[CrossRef\]](#)
13. Draeger, C.; Fink, M. One-channel time reversal of elastic waves in a chaotic 2d-silicon cavity. *Phys. Rev. Lett.* **1997**, *79*, 407–410. [\[CrossRef\]](#)
14. Draeger, C.; Fink, M. One-channel time-reversal in chaotic cavities: Theoretical limits. *J. Acoust. Soc. Am.* **1999**, *105*, 611–617. [\[CrossRef\]](#)
15. Draeger, C.; Cassereau, D.; Fink, M. Theory of the time-reversal process in solids. *J. Acoust. Soc. Am.* **1997**, *102*, 1289–1295. [\[CrossRef\]](#)
16. Yon, S.; Tanter, M.; Fink, M. Sound focusing in rooms: The time-reversal approach. *J. Acoust. Soc. Am.* **2003**, *113*, 1533–1543. [\[CrossRef\]](#) [\[PubMed\]](#)

17. Sprik, R. Time-reversed experiments with acoustics. *J.-Ned. Akoestisch Genoot.* **2005**, *174*, 1–10.
18. Bavu, E.; Besnainou, C.; Gibiat, V.; de Rosny, J.; Fink, M. Subwavelength sound focusing using a time-reversal acoustic sink. *Acta Acust. United Acust.* **2007**, *93*, 706–715.
19. Conti, S.G.; Roux, P.; Kuperman, W.A. Near-field time-reversal amplification. *J. Acoust. Soc. Am.* **2007**, *121*, 3602–3606. [\[CrossRef\]](#)
20. Catheline, S.; Fink, M.; Quieffin, N.; Ing, R.K. Acoustic source localization model using in-skull reverberation and time reversal. *Appl. Phys. Lett.* **2007**, *90*, 063902. [\[CrossRef\]](#)
21. Padois, T.; Prax, C.; Valeau, V.; Marx, D. Experimental localization of an acoustic sound source in a wind-tunnel flow by using a numerical time-reversal technique. *J. Acoust. Soc. Am.* **2012**, *132*, 2397–2407. [\[CrossRef\]](#)
22. Mimani, A.; Porteous, R.; Doolan, C.J. A simulation-based analysis of the effect of a reflecting surface on aeroacoustic time-reversal source characterization and comparison with beamforming. *Wave Motion* **2017**, *70*, 65–89. [\[CrossRef\]](#)
23. Ma, C.; Kim, S.; Fang, N.X. Far-field acoustic subwavelength imaging and edge detection based on spatial filtering and wave vector conversion. *Nat. Commun.* **2019**, *10*, 204. [\[CrossRef\]](#)
24. Li, Y.; Li, M.; Feng, D.; Pan, W.; Wei, L.; Yang, D. Low-frequency acoustic source localization based on the cross-spectral time reversal method corrected in wavenumber do-main. *Measurement* **2022**, *188*, 110579. [\[CrossRef\]](#)
25. Donoho, D.L. Compressed sensing. *IEEE Trans. Inform. Theory* **2006**, *52*, 1289–1306. [\[CrossRef\]](#)
26. Candès, E.J. Compressive sampling. In Proceedings of the International Congress of Mathematicians, Madrid, Spain, 22–30 August 2006.
27. Baraniuk, R.G. Compressive sensing. *IEEE Signal. Proc. Mag.* **2007**, *24*, 118–121. [\[CrossRef\]](#)
28. Cai, J.; Liu, W.; Zong, R.; Wu, B. Sparse array extension for non-circular signals with subspace and compressive sensing based DOA estimation methods. *Signal. Process.* **2018**, *145*, 59–67. [\[CrossRef\]](#)
29. Le Roux, J.; Boufounos, P.T.; Kang, K.; Hershey, J.R. Source localization in reverberant environments using sparse optimization. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 4310–4314.
30. Chardon, G.; Daudet, L. Narrowband source localization in an unknown reverberant environment using wavefield sparse decomposition. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Kyoto, Japan, 25–30 March 2012.
31. Jiang, H.; Mathews, B.; Wilford, P. Sound localization using compressive sensing. *arXiv* **2013**, arXiv:1302.7070.
32. Simard, P.; Antoni, J. Acoustic source identification: Experimenting the ℓ_1 minimization approach. *Appl. Acoust.* **2013**, *74*, 974–986. [\[CrossRef\]](#)
33. Koyama, S.; Shimauchi, S.; Ohmuro, H. Sparse sound field representation in recording and reproduction for reducing spatial aliasing artifacts. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Florence, Italy, 25–30 May 2014; pp. 4443–4447.
34. Koyama, S.; Matsubayashi, A.; Murata, N.; Saruwatari, H. Sparse sound field decomposition using group sparse Bayesian learning. In Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, Hong Kong, China, 16–19 December 2015.
35. Chu, N.; Picheral, J.; Mohammad-Djafari, A.; Gac, N. A robust super-resolution approach with sparsity constraint in acoustic imaging. *Appl. Acoust.* **2014**, *76*, 197–208. [\[CrossRef\]](#)
36. Tipping, M.E. Sparse Bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.* **2001**, *1*, 211–244.
37. Zhang, W.; Zhang, X.; Wu, S.; Huang, J.; Yuan, N. Off-grid underdetermined DOA estimation of quasi-stationary signals via sparse Bayesian learning. In Proceedings of the 2nd International Conference on Information Communication and Signal Processing, Weihai, China, 28–30 September 2019.
38. Bai, Z.; Jensen, J.R.; Sun, J.; Christensen, M.G. A sparse Bayesian learning based RIR reconstruction method for acoustic TOA And DOA estimation. In Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, USA, 20–23 October 2019.
39. Zheng, R.; Xu, X.; Ye, Z.; Dai, J. Robust sparse Bayesian learning for DOA estimation in impulsive noise environments. *Signal. Process.* **2020**, *171*, 107500. [\[CrossRef\]](#)
40. Qin, M.; Hu, D.; Chen, Z.; Yin, F. Compressive sensing-based sound source localization for microphone arrays. *Circuits Syst. Signal Process.* **2021**, *40*, 4696–4719. [\[CrossRef\]](#)
41. Sajjadih, M.; Asif, A. Compressive sensing time reversal MIMO radar: Joint direction and Doppler frequency estimation. *IEEE Signal. Process. Lett.* **2015**, *22*, 1283–1287. [\[CrossRef\]](#)
42. Jiang, W.; Tong, F. Exploiting dynamic sparsity for time reversal underwater acoustic communication under rapidly time varying channels. *Appl. Acoust.* **2021**, *172*, 107648. [\[CrossRef\]](#)
43. Liang, G.; Li, C.; Qiu, L.; Shen, T.; Hao, Y. State-updating-based DOA estimation using sparse Bayesian learning. *Appl. Acoust.* **2022**, *192*, 108719. [\[CrossRef\]](#)
44. Qi, Y.; Zhou, S.; Luo, Z.; Liu, C.; Du, S.; Dun, J.; Zhou, L. Passive source localization based on multipath arrival angles with a vertical line array using sparse Bayesian learning. *J. Acoust. Soc. Am.* **2023**, *153*, 773–791. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Wipf, D.P.; Rao, B.D. Sparse Bayesian learning for basis selection. *IEEE Trans. Signal. Process.* **2004**, *52*, 2153–2164. [\[CrossRef\]](#)
46. Sabine, W.C. Reverberation. In *The American Architect*; Ticknor and Company: Boston, MA, USA, 1938.

47. Yan, J.; Zhao, W.; Wu, Y.I.; Zhou, Y. Indoor sound source localization under reverberation by extracting the features of sample covariance. *Appl. Acoust.* **2023**, *210*, 109453. [[CrossRef](#)]
48. Garcia-Barrios, G.; Iglesias, E.L.; Gutierrez-Arriola, J.M.; Fraile, R.; Saenz-Lechon, N.; Osma-Ruiz, V.J. Exploiting spatial diversity for increasing the robustness of sound source localization systems against reverberation. *Appl. Acoust.* **2023**, *202*, 109138. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.