

Optimized iterative shrinkage threshold generalized inverse beamforming for sound source localization

Huihui He(何辉辉), Xinyu Wang(王欣宇), Zeyu Yang(杨泽宇), Xiaofei Wu(吴晓飞), and Shengguo Shi(时胜国)
Citation: Chin. Phys. B, 2026, 35 (**5**): 054301. DOI: 10.1088/1674-1056/ae118d
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Ultra-broadband acoustic logic gate based on passive phase manipulation

Yu-Han Xia(夏宇涵), Nai-Qi Pang(庞乃琦), Yin Wang(王垠), Long-Xu Wang(王龙旭), and Yong Ge(葛勇)
Chin. Phys. B, 2026, 35 (**1**): 014301. DOI: 10.1088/1674-1056/ade667

Pipeline thickness estimation using the dispersion of higher-order SH guided waves

Zhengchen Dai(代政辰), Jinxia Liu(刘金霞), Yunfei Long(龙云飞), Jianhai Zhang(张建海), Tribikram Kundu, and Zhiwen Cui(崔志文)
Chin. Phys. B, 2024, 33 (**7**): 074301. DOI: 10.1088/1674-1056/ad3033

Fast estimation of distance between two hydrophones using ocean ambient noise in multi-ship scenarios

Xuefeng Liu(刘雪枫), Zhi Xia(夏峙), Qi Li(李琪), and Ye Ding(丁烨)
Chin. Phys. B, 2023, 32 (**12**): 124301. DOI: 10.1088/1674-1056/acf91b

Beam alignments based on the spectrum decomposition of orbital angular momentums for acoustic-vortex communications

Gepu Guo(郭各朴), Xinjia Li(李昕珈), Qingdong Wang(王青东), Yuzhi Li(李禹志), Qingyu Ma(马青玉), Juan Tu(屠娟), and Dong Zhang(章东)
Chin. Phys. B, 2022, 31 (**12**): 124302. DOI: 10.1088/1674-1056/ac686a

Measuring orbital angular momentum of acoustic vortices based on Fraunhofer's diffraction

Chao-Fan Gong(龚超凡), Jing-Jing Li(李晶晶), Kai Guo(郭凯), Hong-Ping Zhou(周红平)†, and Zhong-Yi Guo(郭忠义)‡
Chin. Phys. B, 2020, 29 (**10**): 104301. DOI: 10.1088/1674-1056/ab9c11

Optimized iterative shrinkage threshold generalized inverse beamforming for sound source localization

Huihui He(何辉辉)[3], Xinyu Wang(王欣宇)[3], Zeyu Yang(杨泽宇)[3], Xiaofei Wu(吴晓飞)[4], and Shengguo Shi(时胜国)[1,2,3,†]

[1] *National Key Laboratory of Underwater Acoustic Technology, Harbin Engineering University, Harbin* 150001, *China*
[2] *Key Laboratory of Marine Information Acquisition and Security (Harbin Engineering University), Ministry of Industry and Information Technology, Harbin* 150001, *China*
[3] *College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin* 150001, *China*
[4] *China Ship Scientific Research Center, Wuxi* 214082, *China*

(Received 12 July 2025; revised manuscript received 26 September 2025; accepted manuscript online 10 October 2025)

This work addresses underwater noise source localization. Leveraging the spatial sparsity of sound sources, we employ an optimized iterative shrinkage-thresholding generalized inverse beamforming (OISTA-GIB) method to localize noise sources. Firstly, the sparsity of sound sources is exploited by introducing the λ_1 norm, resulting in an objective function that combines the λ_1 norm, with generalized inverse beamforming. This function is solved using an iterative shrinkage-thresholding algorithm (ISTA) to obtain sound source positions. Secondly, we note that when ISTA solves this objective function, the penalty strength applied by the identity matrix to all scanning points on the sound source surface is uniform. This uniformity reduces positioning accuracy. To enhance localization accuracy and spatial resolution, we propose an iterative regularization matrix-optimized ISTA to solve the objective function. Here, the result from the previous iteration is used to construct a regularization matrix that increases the penalty strength in non-source regions during the current iteration. This process iteratively narrows the mainlobe width in source regions until termination conditions are met, yielding refined sound source positions. Finally, simulations and experimental data processing show that the proposed OISTA-GIB method achieves higher accuracy and spatial resolution in noise source localization compared to existing methods.

Keywords: generalized inverse beamforming, iterative regularization matrix, iterative shrinkage threshold, fast iterative shrinkage threshold

PACS: 43.60.+d, 43.58.+z, 43.50.+y

DOI: 10.1088/1674-1056/ae118d **CSTR:** 32038.14.CPB.ae118d

1. Introduction

Underwater acoustic stealth holds significant scientific research value. Identifying the spatial locations of major noise sources through underwater target sound source localization studies is a prerequisite for noise control.[1,2] Beamforming technology based on hydrophone array measurements provides an efficient method for noise source localization. Since the number of actual sound sources is often sparse relative to the scanned spatial area, scholars have investigated beamforming methods leveraging sound source sparsity.[3–7]

Generalized inverse beamforming (GIB)[8–10] effectively reduces sidelobe levels. It improves sound source localization resolution through fewer iterative inversion processes and is widely used to locate and identify different noise sources.[11,12] However, measurement errors are unavoidable, causing bias in noise source localization and identification results, and still ill-posed problems persist in the inversion process.[13–15] Regarding the ill-posed problem of generalized inverse beamforming, relevant scholars have conducted various research. To improve source localization accuracy, Prezniak[16] proposed a weighted generalized inverse beamforming method that uses a weighted pseudo-inverse approach and an optimization process, achieving higher accuracy compared to standard generalized inverse beamforming. Zavala[17] proposed an automatic regularization factor to enhance the method and employed a virtual target grid to obtain source mapping and intensity estimation. By constructing a new iterative regularization matrix, Zhang[14] proposed an optimize generalized inverse beamforming method with superior noise source localization performance. Takao[8] proposed the λ_1 norm constrained generalized inverse beamforming, but this method fails to eliminate the noise interference and exhibits poor robustness. The CVX toolkit solves this method, but its lack of portability in engineering software motivates this paper proposes to use the iterative shrinkage thresholding algorithm (ISTA) for solving λ_1 norm constraint generalized inverse beamforming.

Daubechies[18] proposed the ISTA and used it to solve linear inversion problems in signal and image processing. Since the ISTA algorithm converges slowly, Beck[19] proposed the

†Corresponding author. E-mail: shishengguo@hrbeu.edu.cn

© 2026 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

fast iterative shrinkage thresholding algorithm (FISTA), which retains the computational simplicity of the ISTA while significantly improving global convergence speed. Wang[20] proposed the improved fast iterative shrinkage thresholding algorithm (IFISTA) to optimize vibration signal reconstruction, achieving higher reconstruction accuracy and faster convergence. Shen[21] proposed a Fourier-based improved fast iterative shrinkage thresholding algorithm (FFT-IFISTA), which selects appropriate weighting coefficients to reduce the sidelobes for resolution enhancement.

In summary, relevant scholars have conducted extensive research on the sparse nature of sound sources and the ill-posed problem of generalized inverse beamforming, respectively. However, fewer studies address generalized inverse beamforming methods based on acoustic sparsity. In this paper, we mainly formulate it as a Lasso (least absolute shrinkage and selection operator) problem solved using the ISTA method. This approach applies uniform penalty strength via the identity matrix on all scanning points of the sound source surface, resulting in low positioning accuracy. To improve localization accuracy and spatial resolution, we optimize an iterative shrinkage thresholding algorithm with an iterative regularization matrix to solve sparse-source generalized inverse beamforming. Simulation and experimental data processing verify that the proposed method achieves the highest noise source localization accuracy and resolution.

The remainder of this paper is organized as follows. Section 2 gives the objective function combining the λ_1 norm with the generalized inverse beamforming, the Iterative threshold algorithm, and the optimized iterative threshold generalized inverse beamforming method. To verify the performance of the optimized iterative shrinking generalized inverse beamforming method the simulation and experimental data processing is described in Section 3 and Section 4, respectively. The work of this paper is the conclusions summarized in Section 5.

2. Sound source localization algorithm

2.1. The λ_1 norm constrained GIB

Assume that there exists a planar array of M-element sound pressure hydrophones in the sound field space (Figure 1), whose received sound field information consists of the sound pressure matrix $\boldsymbol{P}$. Where $\boldsymbol{R}^{(p)} = \boldsymbol{P}\boldsymbol{P}^H$, the eigenvalue decomposition of the cross-spectral matrix $\boldsymbol{R}^{(p)}$ of the received signal can be obtained:

$$\boldsymbol{R}^{(p)} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H = \sum_{m=1}^{M} \sigma_m \boldsymbol{u}_m \boldsymbol{u}_m^H, \tag{1}$$

where $\boldsymbol{U}$ is a unitary matrix containing the orthogonal eigenvectors, and $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues. $\boldsymbol{u}_m$ represents the m-th column vector in the matrix $\boldsymbol{U}$ and its corresponding eigenvalue is σ_m. Therefore, the expression for the m-th order eigenmode vector can be obtained

$$\boldsymbol{p}_m = \sqrt{\sigma_m}\boldsymbol{u}_m, \quad (m = 1, 2, \ldots, M). \tag{2}$$

In the process of obtaining modal vectors, threshold truncation filtering is applied to eliminate smaller eigenvalues. Typically, the truncation threshold is set between 0.1% and 10% of the maximum eigenvalue. Therefore, equation (2) is converted to

$$\boldsymbol{p}_i = \sqrt{\sigma_i}\boldsymbol{u}_i \quad (i < M). \tag{3}$$

The view of the GIB method is to use the eigenmodes to reconstruct the acoustic energy distribution information on the scanning domain of the sound source (Fig. 1), and therefore the following acoustic transfer equation can be established:

$$\boldsymbol{p}_m = \boldsymbol{G}_{\text{green}}\boldsymbol{q}_m, \tag{4}$$

where $\boldsymbol{q}_m$ is the sound source amplitude vector with dimension $N \times 1$, which describes the acoustic energy distribution of N grid points on the scan domain in the m-th eigenmode. $\boldsymbol{G}_{\text{green}}$ is the free-field Green's function transfer matrix between the grid points to each array element with dimension $M \times N$. The m-th order source amplitude vector $\boldsymbol{q}_m$ can be obtained by solving Eq. (4).

In most cases, the number of array elements is much smaller than the number of scanning points on the sound source surface, so equation (4) is ill-conditioned, and the number of sound sources is often sparse compared to the sound source surface. Therefore, combined with the λ_1 norm constraint, the expression is as follows:

$$\min_{\boldsymbol{q}_m} \|\boldsymbol{q}_m\|_1 \text{ s.t. } \left\|\boldsymbol{p}_m - G_{\text{green}}\boldsymbol{q}_m\right\|_2^2 \le \delta. \tag{5}$$

According to Eqs. (2)–(4):

$$\min_{\boldsymbol{q}_i} \|\boldsymbol{q}_i\|_1 \quad \text{s.t.} \quad \left\|\sqrt{\sigma_i}\boldsymbol{u}_i - G_{\text{green}}\boldsymbol{q}_i\right\|_2^2 \le \delta, \tag{6}$$

where δ is a constraint parameter whose magnitude is related to the signal-to-noise ratio (SNR), frequency and test distance, which interferes with Eq. (6), so equation (6) is transformed into a regularized optimization model:

$$F(\boldsymbol{q}_i) = \arg\min \left\{ \left\|\sqrt{\sigma_i}\boldsymbol{u}_i - G_{\text{green}}\boldsymbol{q}_i\right\|_2^2 + \lambda \left\|\boldsymbol{L}q_i\right\|_1 \right\}, \tag{7}$$

where λ is the λ_1 norm regularization parameter and $\boldsymbol{L}$ is the regularization matrix, both of which are employed to adjust the reconstruction accuracy and sparsity. When the regularization matrix satisfies $\boldsymbol{L} = I$, equation (7) is the Lasso[22–24] problem, which is usually solved by interior point.[25–27] The algorithmic complexity of the interior point method is $O\left(N^3\right)$, which requires a large amount of computation when the data dimension is large, which reduces the calculation efficiency.

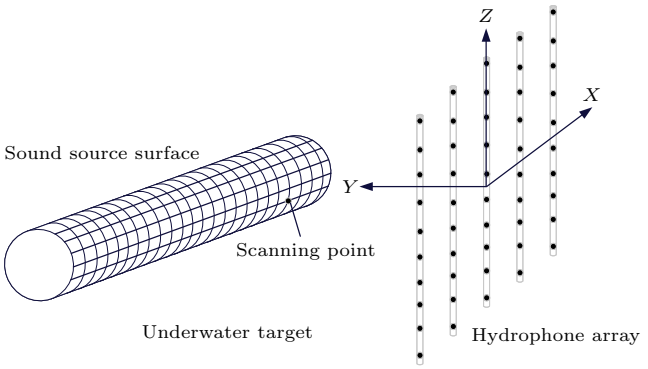


Fig. 1. Diagram of sound source localization.

2.2. Iterative thresholding algorithm

2.2.1. Iterative shrinking threshold algorithm

A large number of researchers have solved Eq. (7) by gradient-based methods.[28–30] The iterative shrinking threshold algorithm updates $\boldsymbol{q}_i$ through a shrinking threshold operation in each iteration. Equation (7) is solved as follows:

$$\boldsymbol{q}_i^{(k+1)} = \mathrm{soft}_{\lambda t}\left(\boldsymbol{q}_i^{(k)} - 2t^{(k)}\boldsymbol{G}_{\mathrm{green}}^{\mathrm{T}}\left(\boldsymbol{G}_{\mathrm{green}}\boldsymbol{q}_i^{(k)} - \sqrt{\sigma_i}\boldsymbol{u}_i\right)\right), \quad (8)$$

where $t^{(k)} > 0$ is the step size, satisfies

$$t^{(k)} \in \left(0, 1/\left\|\boldsymbol{G}_{\mathrm{green}}\right\|_2^2\right),$$

where k is the number of iterations, $\mathrm{soft}_{\lambda t}(\bullet)$ is the shrinking operator

$$\mathrm{soft}_{\lambda t}(q_i) = (|q_i| - \alpha)_+ \,\mathrm{sgn}(q_i), \quad (9)$$

where $|\,|$ is the absolute value, $(\,)_+$ represents the maximum value between $|q_i| - \alpha$ and 0. sgn is a sign function, which means that if $q_i > 0$, then $\mathrm{sgn}(q_i) = 1$; $q_i = 0$, then $\mathrm{sgn}(q_i) = 0$; $q_i < 0$ then $\mathrm{sgn}(q_i) = -1$. Equation (8) is the iterative shrinkage threshold algorithm for solving λ_1 norm constrained generalized inverse beamforming (ISTA-GIB).

2.2.2. Fast iterative shrinkage threshold

Since the ISTA algorithm converges slowly, Beck[19] proposed the FISTA, which retains the computational simplicity of the ISTA algorithm, resulting in a significant improvement in the algorithm's global convergence speed. The initialization of FISTA is $\boldsymbol{q}_i^{(k)} = 0$, $\boldsymbol{y}_i^{(1)} = \boldsymbol{q}_i^{(0)}$, $t_i^{(1)} = 1$. The steps from the result $\boldsymbol{q}^{(k)}$ of the k-th iteration to the result $\boldsymbol{q}^{(k+1)}$ of the $(k+1)$-th iteration is

$$\begin{cases} \boldsymbol{q}_i^{(k)} = P_+\left(\boldsymbol{y}_i^{(k)} - \nabla f(\boldsymbol{y}_i^{(k)})/L\right), \\ t^{(k+1)} = \dfrac{1}{2}\left(1 + \sqrt{1 + 4\left(t^{(k)}\right)^2}\right), \\ \boldsymbol{y}_i^{(k+1)} = \boldsymbol{q}_i^{(k)} + \dfrac{t^{(k)} - 1}{t^{(k+1)}}\left(\boldsymbol{q}_i^{(k)} - \boldsymbol{q}_i^{(k-1)}\right), \end{cases} \quad (10)$$

where P_+ is a Euclidean projection of a non-negative quadrant, combining Eq. (7)

$$\begin{cases} \boldsymbol{q}_i^{(k)} = P_+\left(\boldsymbol{q}_i^{(k-1)} - 2t^{(k-1)}\boldsymbol{G}_{\mathrm{green}}^{\mathrm{T}} \right. \\ \qquad \left. \times \left(\boldsymbol{G}_{\mathrm{green}}\boldsymbol{q}_i^{(k-1)} - \sqrt{\sigma_i}\boldsymbol{u}_i\right)\right), \\ t^{(k+1)} = \dfrac{1}{2}\left(1 + \sqrt{1 + 4\left(t^{(k)}\right)^2}\right), \\ \boldsymbol{q}_i^{(k+1)} = \boldsymbol{q}_i^{(k)} + \dfrac{t^{(k)} - 1}{t^{(k+1)}}\left(\boldsymbol{q}_i^{(k)} - \boldsymbol{q}_i^{(k-1)}\right). \end{cases} \quad (11)$$

Equation (11) is the fast iterative shrinkage threshold algorithm for solving λ_1 norm constrained generalized inverse beamforming (FISTA-GIB).

2.2.3. Optimized iterative threshold algorithm

The penalty strength of the unit matrix on all scanning points of the sound source surface is the same, leading to the noise source's minimal positioning accuracy. To improve the accuracy and spatial resolution of sound source localization, an iterative shrinkage threshold algorithm optimized by an iterative regularization matrix is proposed, and the result of the previous iteration is constructed as a regularization matrix, i.e., when calculating the $(k+1)$-th iteration result $\boldsymbol{q}_i^{(k+1)}$, the regularization matrix constructed by the k-th iteration result $\boldsymbol{q}_i^{(k)}$ is employed to increase the penalty strength for the non-sound source area.

$$\boldsymbol{L}_i^{(k+1)} = \mathrm{diag}\left(\left|\boldsymbol{q}_i^{(k)}\right| \Big/ \left\|\boldsymbol{q}_i^{(k)}\right\|_\infty\right). \quad (12)$$

When equation (12) is substituted into Eq. (7) and solved using Eq. (8):

$$\begin{cases} \boldsymbol{q}_i^{(k+1)} = \mathrm{soft}_{\lambda t}\left(\boldsymbol{q}_i^{(k)} - 2t^{(k)}\left(\boldsymbol{H}_i^{(k)}\right)^{\mathrm{T}} \right. \\ \qquad \left. \times \left(\boldsymbol{H}_i^{(k)}\boldsymbol{q}_i^{(k)} - \sqrt{\sigma_i}\boldsymbol{u}_i\right)\right), \\ \boldsymbol{H}_i^{(k)} = \boldsymbol{G}_{\mathrm{green}}\boldsymbol{L}_i^{(k)}. \end{cases} \quad (13)$$

Equation (13) is the optimized iterative shrinkage threshold algorithm for solving λ_1 norm constrained generalized inverse beamforming (OISTA-GIB).

When equation (12) is substituted into Eq. (7) and solved using Eq. (11):

$$\begin{cases} \boldsymbol{q}_i^{(k)} = P_+\left(\boldsymbol{q}_i^{(k-1)} - 2t^{(k-1)}\left(\boldsymbol{H}_i^{(k-1)}\right)^{\mathrm{T}} \right. \\ \qquad \left. \times \left(\boldsymbol{H}_i^{(k-1)}\boldsymbol{q}_i^{(k-1)} - \sqrt{\sigma_i}\boldsymbol{u}_i\right)\right), \\ \boldsymbol{H}_i^{(k-1)} = G_{\mathrm{green}}\boldsymbol{L}_i^{(k-1)}, \\ t^{(k+1)} = \dfrac{1}{2}\left(1 + \sqrt{1 + 4\left(t^{(k)}\right)^2}\right), \\ \boldsymbol{q}_i^{(k+1)} = \boldsymbol{q}_i^{(k)} + \dfrac{t^{(k)} - 1}{t^{(k+1)}}\left(\boldsymbol{q}_i^{(k)} - \boldsymbol{q}_i^{(k-1)}\right). \end{cases} \quad (14)$$

Equation (14) is the optimized fast iterative shrinkage threshold algorithm for solving λ_1 norm constrained generalized inverse beamforming (OFISTA-GIB).

The optimized iterative shrinkage threshold algorithm for solving λ_1 norm constrained generalized inverse beamforming process is as follows:

(i) Obtain the covariance matrix $\boldsymbol{R}^{(p)}$ of the sound field signal $\boldsymbol{P}$.

(ii) The m-order eigenmode vector equation (2) is obtained by calculating Eq. (1). The eigenmode vectors with eigenvalues less than 0.05 times the corresponding maximum eigenvalue are eliminated to obtain i eigenmode vectors, where $i < M$, i.e., Eq. (3).

(iii) According to the view of the GIB method, equations (2)–(4) are combined with the λ_1 norm constraint to obtain Eq. (6), and equation (6) is transformed into the Lasso problem, i.e., Eq. (7).

(iv) Equation (7) is solved by the ISTA algorithm to obtain the sound source localization results on the scanning surface corresponding to the i-th order modes.

(v) Repeat step (iv) to obtain the sound source localization results corresponding to all i eigenmode vectors and superimpose the results to obtain the noise source localization.

(vi) The positioning accuracy of the noise source obtained by Eq. (8) is poor. Therefore, this paper proposes to introduce the regularization matrix optimization ISTA algorithm with strong penalty intensity for the non-sound source area and uses it to solve the λ_1 norm constrained generalized inverse beamforming.

(vii) The result of the first iteration calculation of the i-th order mode in step (iv) is taken as the initial input to construct the regularization matrix $\boldsymbol{L}_i^{(1)}$ and obtain Eq. (12). Substitute Eq. (12) into Eq. (7) to obtain a new objective function, and then use Eq. (8) to solve the model to obtain Eq. (13).

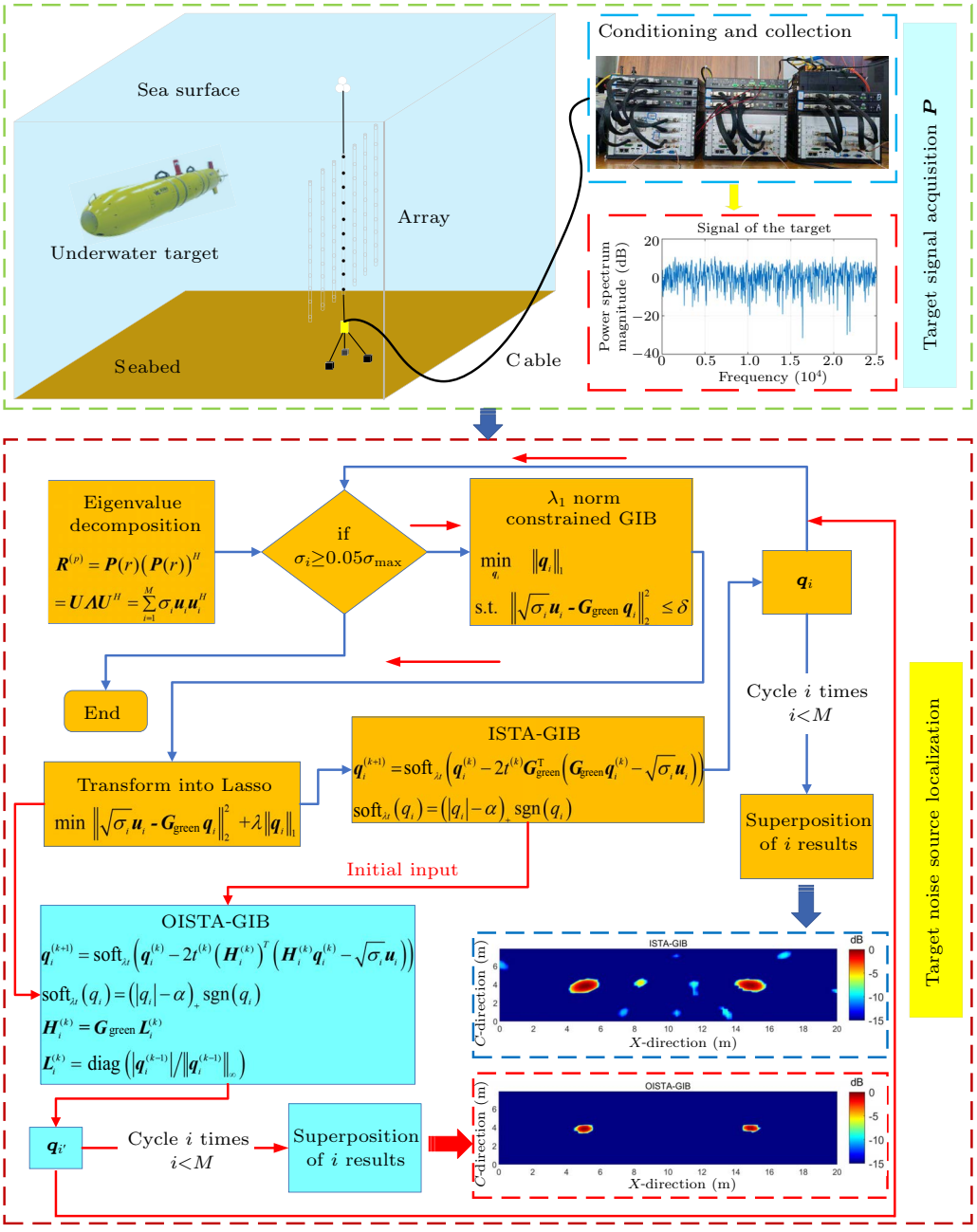


Fig. 2. Flowchart of noise source location.

(viii) The result of the 2nd iteration on the scan area corresponding to the i-th order mode is obtained by solving step (vii) and is used to construct the regularization matrix $\boldsymbol{L}_i^{(2)}$ and obtain Eq. (12). Substitute Eq. (12) into Eq. (7) to obtain a new objective function, and then use Eq. (8) to solve the model to obtain Eq. (13) to get the result of the third iteration, then the result is used to construct the regularization matrix $\boldsymbol{L}_i^{(3)}$. The iteration is terminated until the error between the k-th iteration and the $(k-1)$-th iteration is less than 10^{-6}, and the noise source positioning result $\boldsymbol{q}_i$ on the scan plane corresponding to the i-th mode is obtained.

(ix) Repeat step (viii) to obtain the sound source localization results corresponding to all eigenmode vectors and superimpose the results to obtain the noise source localization. The process is shown in Fig. 2, and the optimized fast iterative shrinkage threshold generalized inverse beamforming is similar to the process, which is not repeated.

3. Simulated validation

To validate the performance of the optimized iterative shrinkage thresholding generalized inverse beamforming method, this section compares the noise source localization capabilities of various methods through positioning errors and spatial resolution assessments using incoherent dual sound sources, varying SNRs, and different frequencies. Both simulation and experimental data were processed on a computer equipped with an Intel Xeon E3-1226 v3 processor running at 3.30 GHz with 32 GB of RAM.

3.1. Dual sound sources

The position of the source surface and the array is shown in Fig. 1, there are two equal-intensity non-coherent sound sources on the source surface. The sound source positions are $(x_1, z_1) = (5, 0)$ m and $(x_2, z_2) = (15, 0)$ m respectively, and the SNR is 3 dB. The source surface is a semi-cylindrical surface with a radius of 2.5 m and an axis length of 20 m. The scanning surface is discretized into 31×81 scanning points. The array is a 10×10 planar array with a spacing of 0.75 m between the array elements, and the sound source surface is 10 m away from the array. In order to compare and analyze the results of the methods, the acoustic imaging results were normalized to 0 dB, and their dynamic display range was 15 dB. Noise source localization was performed using GIB, ISTA-GIB, FISTA-GIB, OISTA-GIB, and OFISTA-GIB respectively. In Eq. (3), the truncation threshold is set to 0.05 times the maximum eigenvalue. For the algorithms ISTA-GIB, FISTA-GIB, OISTA-GIB, and OFISTA-GIB, the iteration count was fixed at 1000, with a termination tolerance of 10^{-6} to ensure convergence.

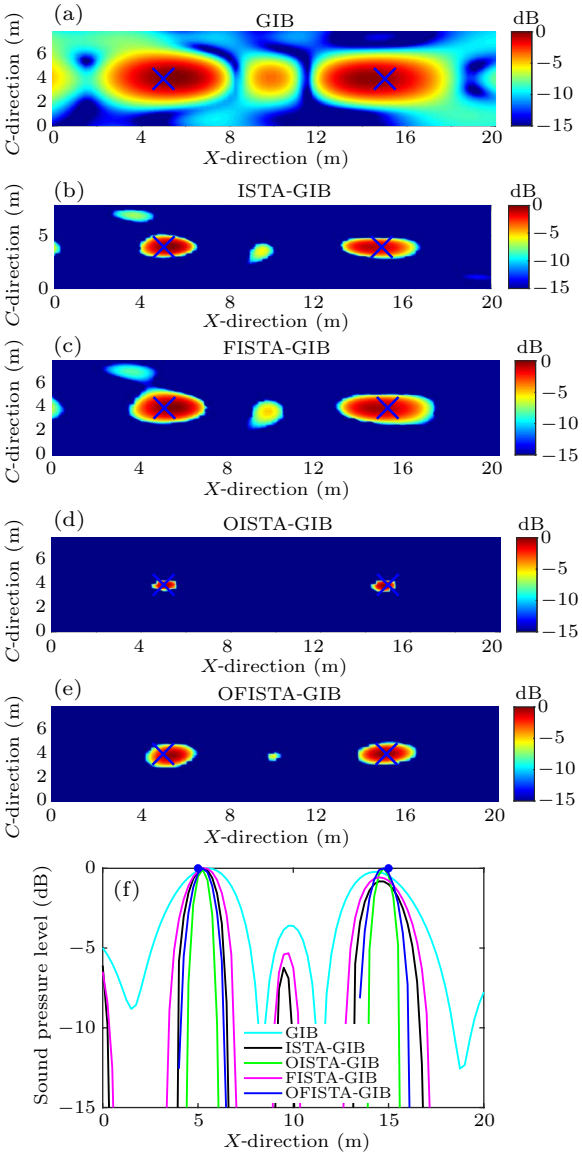


Fig. 3. The sound source localization results of 630 Hz: (a) GIB, (b) ISTA-GIB, (c) FISTA-GIB, (d) OISTA-GIB, (e) OFISTA-GIB, and (f) profile of imaging results.

Figures 3 and 4 respectively display the acoustic imaging results of 630-Hz and 1000-Hz sound sources obtained using different noise source localization methods. In the figures, the blue "×" and "•" denote the true source positions, while the localization errors and 3-dB mainlobe widths for each method are listed in Table 1. Where figures 3(a)–3(e) show the localization results of GIB, ISTA-GIB, OISTA-GIB, FISTA-GIB, and OFISTA-GIB, respectively, and figure 3(f) shows the profile of the imaging results. From the comparisons of Figs. 3(a), 3(b), 3(c), and 3(f) in conjunction with Table 1, it can be observed that the GIB method exhibits significant localization errors and broader mainlobes. Specifically, the localized positions deviate from the actual positions of sound source 1 and sound source 2 by 0.5 m and 0.75 m, respectively, with corresponding 3-dB mainlobe widths of 4.35 m and 4.44 m. The localization results of ISTA-GIB deviate from the actual positions of sound source 1 and sound source 2 by 0.25 m and

0.5 m, respectively, with corresponding 3-B mainlobe widths of 1.96 m and 2.63 m. The localization results of FISTA-GIB deviate from the actual positions of sound source 1 and sound source 2 by 0.25 m and 0.5 m, respectively, with corresponding 3-dB mainlobe widths of 1.5 m and 3.1 m. ISTA-GIB exhibits the same localization errors as FISTA-GIB, but demonstrates narrower mainlobe widths. This performance difference stems from FISTA-GIB's faster convergence rate and fewer iteration counts, along with its weaker penalty strength on off-source regions compared to ISTA-GIB. The localization results of OISTA-GIB deviate from the actual positions of sound source 1 and sound source 2 by 0.25 m and 0.25 m, respectively, with corresponding 3-dB mainlobe widths of 1.05 m and 1.2 m. The localization results of OFISTA-GIB deviate from the actual positions of sound source 1 and sound source 2 by 0.25 m and 0.25 m, respectively, with corresponding 3-dB mainlobe widths of 1.55 m and 1.5 m. The OISTA-GIB and OFISTA-GIB methods exhibit identical localization errors. However, OFISTA-GIB achieves faster convergence and fewer iteration counts, resulting in weaker penalty strength on off-source regions compared to OISTA-GIB along with broader mainlobes. Comprehensive analysis of localization errors and 3-dB mainlobe widths reveals that the OISTA-GIB method delivers the minimal localization error and highest resolution. Identical conclusions were reached at 1000 Hz.

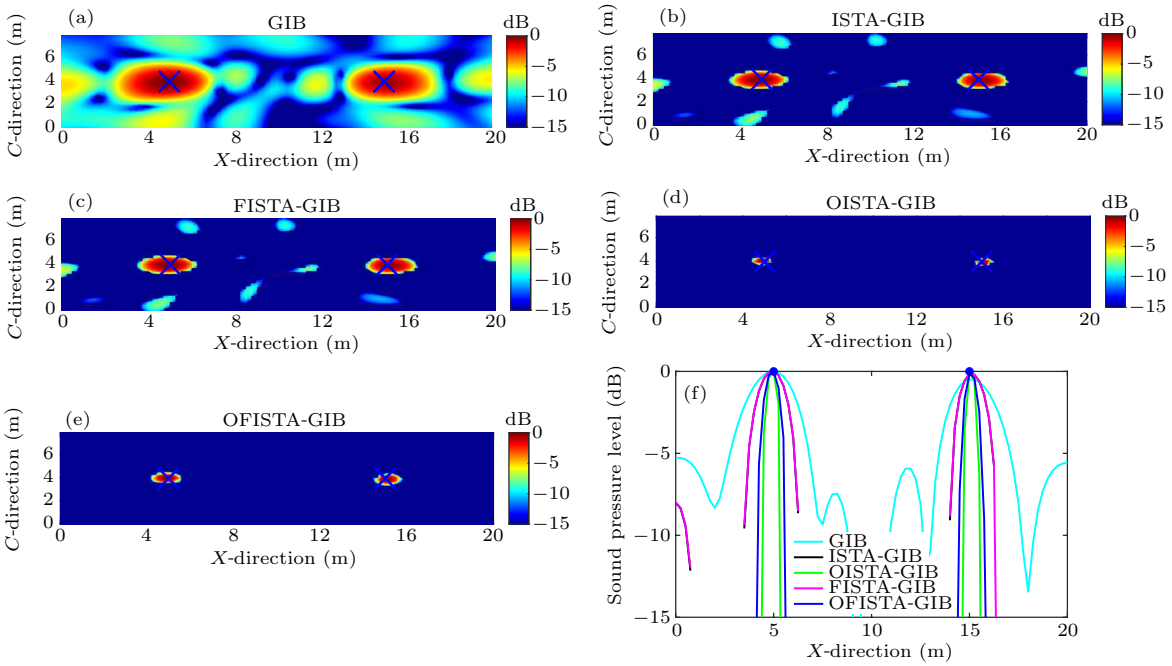


Fig. 4. The sound source localization results of 1000 Hz: (a) GIB, (b) ISTA-GIB, (c) FISTA-GIB, (d) OISTA-GIB, (e) OFISTA-GIB, and (f) profile of imaging results.

Table 1. Localization errors and axial 3-dB beamwidths.

Methods	Frequencies							
	630 Hz				1000 Hz			
	Localization error (m)		Beamwidth of 3 dB (m)		Localization error (m)		Beamwidth of 3 dB (m)	
	S1	S2	S1	S2	S1	S2	S1	S2
GIB	0.50	0.75	4.35	4.44	0	0	3.25	3.19
ISTA-GIB	0.25	0.5	1.96	2.63	0	0	1.95	1.75
FISTA-GIB	0.25	0.5	2.5	3.10	0	0	1.95	1.75
OISTA-GIB	0.25	0.25	1.05	1.2	0	0	0.68	0.61
OFISTA-GIB	0.25	0.25	1.55	1.5	0	0	1.0	0.95

3.2. SNR

To verify the performance of OISTA-GIB, this section compares the sound source localization errors of multiple methods at varying SNRs in a single source source.

The signal-to-noise ratio (SNR) used in this section is defined as the ratio of the signal power at the corresponding frequency point to the noise power, expressed in decibels (dB) as follows:

$$\mathrm{SNR}_{\mathrm{dB}} = 10\log 10\left(\frac{P(f_i)_{\mathrm{signal}}}{P(f_i)_{\mathrm{noise}}}\right), \tag{15}$$

where $P(f_i)_{\mathrm{signal}}$ represents the power of the signal at the frequency point f_i, and $P(f_i)_{\mathrm{noise}}$ represents the power of the noise at f_i .

Let (x_i, z_i) be the position coordinates identified in the i-

th trial, and $(\dot{x},\dot{z})$ be the true coordinates of the sound source. Based on n Monte Carlo trials, the relative average error (RAE) of the sound source localization is calculated as

$$\begin{cases} \mathrm{RAE}(x) = \left|\dfrac{E(x_i)-\dot{x}}{\dot{x}}\right|, \\ \mathrm{RAE}(z) = \left|\dfrac{E(z_i)-\dot{z}}{\dot{z}}\right|, \\ \mathrm{Error}(r) = \dfrac{\mathrm{RAE}(x)+\mathrm{RAE}(z)}{2}\times 100\%. \end{cases} \tag{16}$$

In the simulation, the single sound source is positioned at coordinates $(x_1,z_1) = (1.5,\pi/2)$ m with a frequency of 500 Hz. The SNR range is configured from $[-4,20]$ dB with a step size of 2 dB, and 20 Monte Carlo trials are performed. The source surface is a semi-cylindrical surface with a radius of 1 m and an axis length of 3 m. The scanning surface is discretized into 61×61 scanning points, and the sound source surface is 5 m away from the array. All other simulation parameters remain consistent with those defined in Subsection 3.1. In Eq. (3), the truncation threshold is set to 0.05 times the maximum eigenvalue. For the algorithms ISTA-GIB, FISTA-GIB, OISTA-GIB, and OFISTA-GIB, the iteration count was fixed at 1000, with a termination tolerance of 10^{-6} to ensure convergence. The simulation results are presented in Fig. 5.

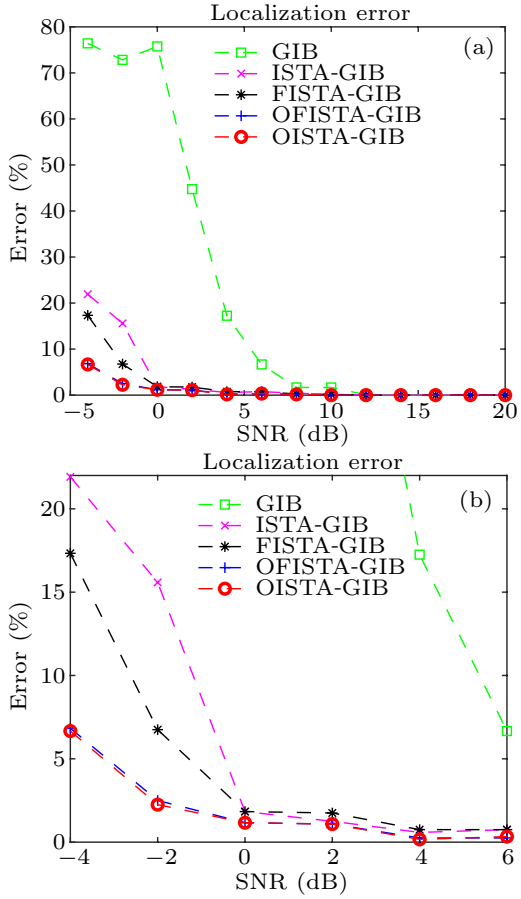


Fig. 5. Localization error: (a) localization errors under different SNRs and (b) magnified view.

Subfigure 5(a) depicts the localization errors across varying SNRs, while subfigure 5(b) provides a magnified view of the $[-4,6]$ dB range. Comparative analysis demonstrates that the localization errors progressively decrease with increasing SNR. Comparative analysis with the GIB method reveals that at an SNR of -4 dB, the four proposed methods achieve smaller localization errors. Among these, OISTA-GIB demonstrates the minimal error of 6.67%, indicating its superior localization performance under low-SNR conditions.

3.3. Spatial resolution

To verify the performance of OISTA-GIB, this section uses a single sound source to compare the spatial resolution of different methods at different frequencies. Spatial resolution is a critical metric for evaluating algorithm performance, as it characterizes the method's capability for multi-target localization and identification. Based on the characteristics of beamforming, the spatial resolution of noise sources depends on the beamwidth. Specifically, sound sources can only be resolved if their separation exceeds the beamwidth. In this study, the 3-dB beamwidth is adopted to characterize the spatial resolution.

In the simulation, the monopole point source is positioned at coordinates $(x_1,z_1) = (10,7\pi/2)$ m with SNR $= 5$ dB. The frequency range spans $[100,2000]$ Hz at 100-Hz intervals, with 20 Monte Carlo simulations conducted per frequency point. The source plane is configured as a semi-cylindrical surface with a radius of 7 m and axial length of 20 m, discretized into a 1001×1100 grid of scanning points, and the sound source surface is 5 m away from the array. All other simulation parameters remain consistent with those defined in Subsection 3.1. For the algorithms ISTA-GIB, FISTA-GIB, OISTA-GIB, and OFISTA-GIB, the iteration count was fixed at 1000, with a termination tolerance of 10^{-6} to ensure convergence. The spatial resolutions of the five methods at different frequency points are comparatively analyzed, with simulation results presented in Fig. 6. Figure 6(a) displays the axial resolution, while figure 6(b) presents the circumferential resolution. Comparative analysis demonstrates that the OISTA-GIB method achieves the narrowest mainlobe widths in both dimensions, indicating its superior spatial resolution. At the 100-Hz frequency point, the mainlobe widths of GIB, ISTA-GIB, FISTA-GIB, OISTA-GIB, and OFISTA-GIB are measured at 9.7 m, 4.31 m, 4.88 m, 2.97 m, and 3.12 m, respectively. Consequently, when two sound sources are spaced 3 m apart in the sound field, only the OISTA-GIB method can successfully resolve and localize them, whereas the other methods fail to achieve discernible localization.

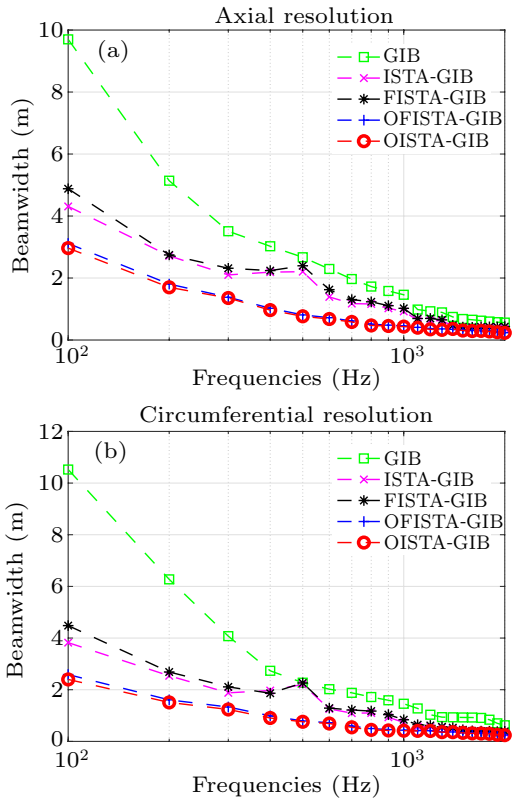


Fig. 6. Spatial resolution of different frequencies: (a) axial resolution and (b) circumferential resolution.

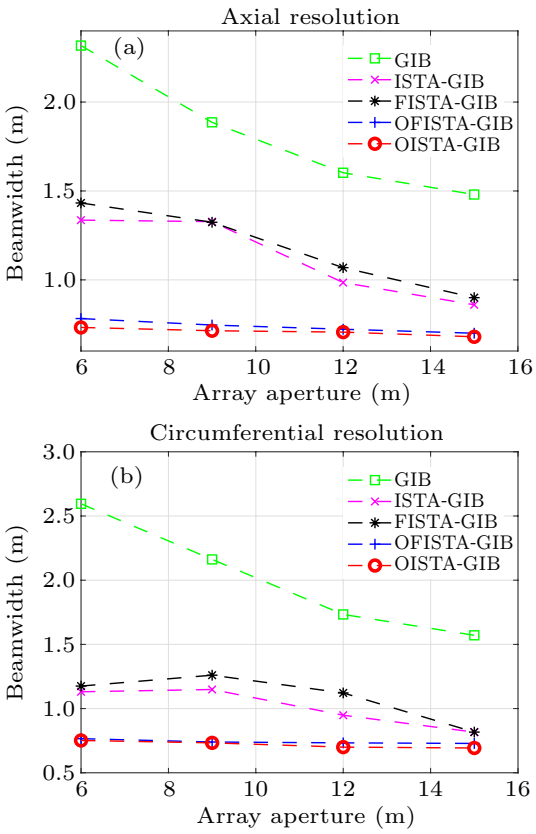


Fig. 7. Spatial resolution under different array apertures: (a) axial resolution and (b) circumferential resolution.

3.4. Array aperture

To verify the performance of OISTA-GIB, this section compares the spatial resolution of different methods under various array apertures using a single sound source. In the simulation, the monopole point source is positioned at coordinates $(x_1, z_1) = (6, 2\pi)$ m with SNR = 5 dB. The array's vertical aperture is set to 6 m, 9 m, 12 m, and 15 m (with the horizontal aperture being identical to the vertical one). For each aperture, 20 Monte-Carlo trials were conducted. The source plane is configured as a semi-cylindrical surface with a radius of 4 m and axial length of 12 m, discretized into a 1201×1257 grid of scanning points, and the sound source surface is 5 m away from the array. All other simulation parameters remain consistent with those defined in Subsection 3.1. For the algorithms ISTA-GIB, FISTA-GIB, OISTA-GIB, and OFISTA-GIB, the iteration count was fixed at 1000, with a termination tolerance of 10^{-6} to ensure convergence. The spatial resolutions of the five methods are compared for different array apertures, as shown in the simulation results of Fig. 7. Figure 7(a) displays the axial resolution, while figure 7(b) presents the circumferential resolution. The comparison reveals that a larger aperture yields higher resolution for all methods. The OISTA-GIB method achieves the smallest main lobe width in both the axial and circumferential directions, signifying the highest spatial resolution.

4. Experimental validation

Numerical simulation results show that the OISTA-GIB algorithm can effectively eliminate the sidelobe and improve the accuracy of source localization. To further verify the feasibility and practicality of the method, the corresponding sound source localization test is constructed. To simulate the testing situation of underwater targets, a hollow cylindrical cabin section model was designed with a length of about 18.6 m and an outer diameter of about 5.35 m. During the test, the water pump pipeline system operating in the cabin section was used as the sound source. To facilitate identifying the location of the noise source, each rib position inside the cabin section was numbered. The side profile of the cabin section model and the rib numbers where each sound source was located are shown in Fig. 8.

The test was carried out in a large pool, with a water depth of 15 m, and the measurement array was a 10-element hydrophone linear array. The linear array was arranged parallel to the water surface. The spacing between each array element was 0.75 m, and the array aperture was 6.75 m. The array is arranged about 13-m underwater, and the upper surface of the cabin section intrudes into the water. The positional relationship of test deployment is shown in Fig. 9. During the test, a winch was used to ensure that the cabin section moved horizontally and uniformly in the water, with a movement speed of 0.35 m/s. During the test, the cabin section passed over

the linear array in a direction perpendicular to the linear array. According to the relativity of motion and the inverse synthetic aperture technique,[31] it can be assumed that the cabin section is stationary, while the line array is moving at a steady speed from far to near and then from near to far in a horizontal uniform motion, and as a result, a virtual surface array can be formed spatially, so that the aperture perpendicular to the direction of the array is virtually extended, thus improving the resolution of the array in the direction of motion for the localization of the sound source.

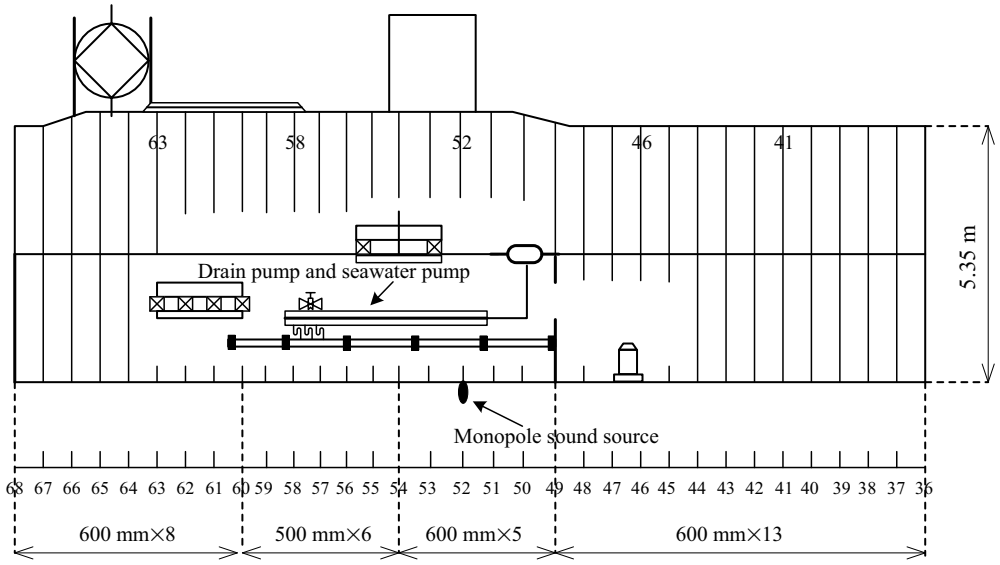


Fig. 8. Lateral profile of cabin.

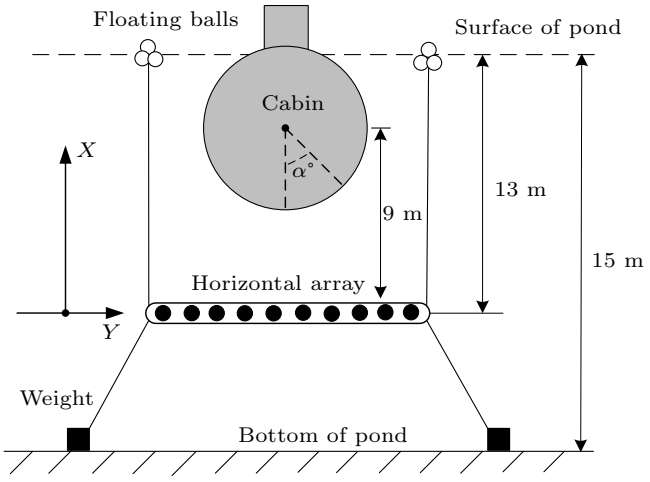


Fig. 9. Test deployment positions of cabin sections and linear arrays.

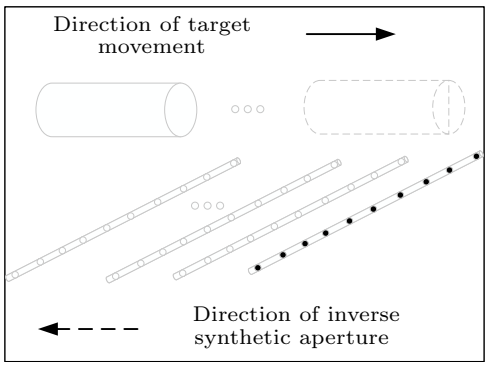


Fig. 10. The inverse synthetic aperture.

According to the inverse synthetic aperture technology, a "virtual plane array" with an array element spacing of 0.75 m, an aperture of 6.75 m, and a total number of array elements of $10 \times 10 = 100$ is generated in the opposite direction of the target motion. The results are shown in Fig. 9. This study employs an acoustic exciter to transmit a source signal covering the frequency range of 500 Hz–2000 Hz. Single-frequency components at 630 Hz and 1000 Hz were extracted using a filter for subsequent data processing. The GIB, ISTA-GIB, OISTA-GIB, FISTA-GIB, and OFISTA-GIB methods (in Eq. (3), the truncation threshold is set to 0.05 times the maximum eigenvalue. For the algorithms ISTA-GIB, OISTA-GIB, FISTA-GIB, and OFISTA-GIB, the iteration count was fixed at 1000, with a termination tolerance of 10^{-6} to ensure convergence.) were used to process the 630-Hz and 1000-Hz test data to obtain the results shown in Figs. 11 and 12. The corresponding localization errors and 3-dB mainlobe widths for each method are summarized in Table 2. Figures 11(a)–11(e) show the localization results of each method, respectively, and figure 11(f) shows the profile of imaging results. By combining the data in Table 2 and comparing Figs. 11(a), 11(b), 11(d), and 11(f), it can be concluded that under the condition of sparse sound sources on the source plane, the ISTA-GIB method produces a cleaner background compared to the conventional GIB method. It achieves a localization error of 0.46 m and a main lobe width of 2.70 m, indicating higher resolution. The OISTA-GIB method, based on the iteratively regularized matrix, yields a cleaner background compared to both ISTA-GIB and GIB. It achieves a localization error of 0.4 m

and a main lobe width of 1.31 m, demonstrating a narrower main lobe and higher spatial resolution.

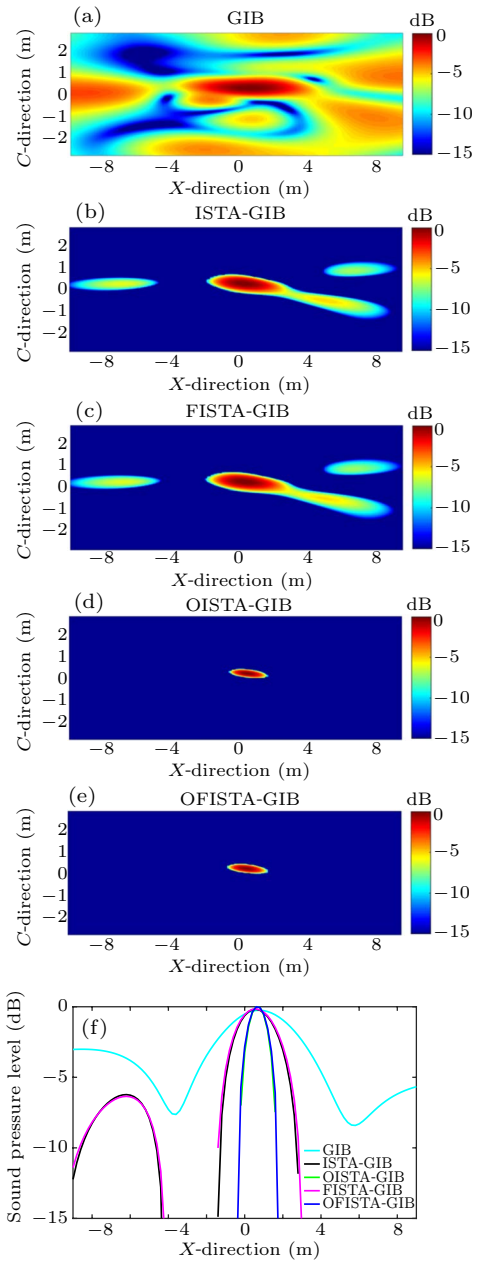


Fig. 11. The sound source localization results of 630 Hz: (a) GIB, (b) ISTA-GIB, (c) FISTA-GIB, (d) OISTA-GIB, (e) OFISTA-GIB, and (f) profile of imaging results.

By combining Table 2 and comparing Figs. 11(b) and 11(c), as well as Figs. 11(d) and 11(e), it can be concluded that ISTA-GIB achieves higher spatial resolution than FISTA-GIB, and OISTA-GIB surpasses OFISTA-GIB. These results are consistent with the conclusions drawn from the simulated data processing, and the same findings hold at a frequency of 1000 Hz.

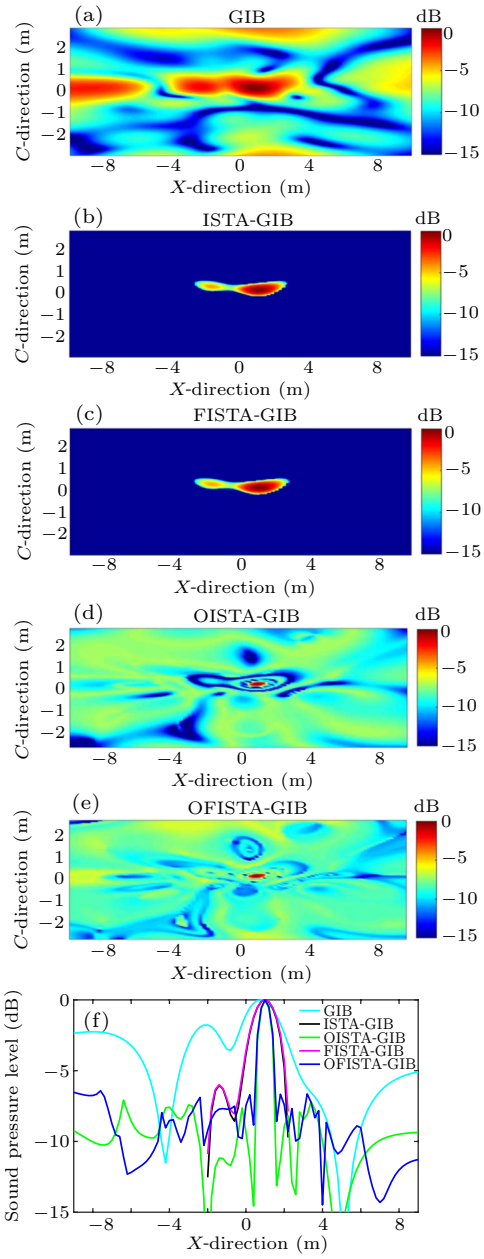


Fig. 12. The sound source localization results of 1000 Hz: (a) GIB, (b) ISTA-GIB, (c) FISTA-GIB, (d) OISTA-GIB, (e) OFISTA-GIB, and (f) profile of imaging results.

Table 2. Localization errors and axial 3-dB beamwidths.

Methods	Frequencies			
	630 Hz		1000 Hz	
	Localization error (m)	Beamwidth of 3 dB (m)	Localization error (m)	Beamwidth of 3 dB (m)
GIB	0.50	5.03	0.2	2.74
ISTA-GIB	0.46	2.70	0	1.72
FISTA-GIB	0.46	2.82	0	1.78
OISTA-GIB	0.4	1.31	0	0.80
OFISTA-GIB	0.4	1.40	0	0.84

5. Conclusion

In this paper, an optimized iterative shrinkage threshold generalized inverse beamforming method is proposed for the underwater target sound source localization problem. Firstly, the objective function combining the λ_1 norm with the generalized inverse beamforming is obtained. The iterative shrinkage threshold algorithm is employed to solve the function and get the position information of the sound source. Secondly, when the iterative shrinkage threshold algorithm is used to solve the above objective function, the penalty strength of the unit matrix on all scanning points of the sound source surface is the same, leading to the noise source's minimal positioning accuracy. To improve the accuracy and spatial resolution of sound source localization, an iterative shrinkage threshold algorithm optimized by an iterative regularization matrix is proposed and used to solve the objective function. A comparison of the performance of FISTA and ISTA is considered when solving the objective function, in which OFISTA-GIB source localization accuracy is slightly worse than OISTA-GIB because the convergence speed of FISTA is faster than that of ISTA, and the number of iterations is less resulting in a lower number of penalties for the non-source area than that of ISTA. If more attention is paid to computational efficiency, the FISTA algorithm can be used to solve the objective function. The highest sound source localization accuracy and resolution of the OISTA-GIB method is verified by simulation and experimental data processing, and the conclusion OISTA-GIB $>$ OFISTA-GIB $>$ ISTA-GIB $>$ FISTA-GIB $>$ GIB is obtained.

Acknowledgement

Project supported by the National Key Scientific Instrument and Equipment Development Projects of China (Grant No. 52327901).

References

[1] Gao Y, Yang B Q and Shi S G 2023 *Chin. Phys. B* **32** 044302
[2] Zhang H Y, Li F Q, Wang P, Xin W and Liu Y Z 2025 *The Journal of the Acoustical Society of America* **157** 3402
[3] Yue Y X, Zhang Z Y and Shi Z G 2024 *IEEE Transactions on Aerospace and Electronic Systems* **60** 5663
[4] Jiang S Y, Jiang R X, Liu X S, Gu B X and Chen Y W 2024 *IEEE Journal of Oceanic Engineering* **49** 340
[5] Zhao D D, Chen P, Hu Y T, Liang R H, Wang H X and Guo X X 2021 *IEEE Journal of Oceanic Engineering* **46** 1356
[6] Firat U and Akgül T 2024 *Signal Processing* **214** 109221
[7] Yu Y C, Shi P C, Krynkin A and Horoshenkov K V 2024 *Measurement* **238** 115361
[8] Suzuki T 2011 *Journal of Sound and Vibration* **330** 5835
[9] Guo X, Wu X, Liu X Q and Tang L J 2023 *Journal of Mechanical Science and Technology* **37** 43
[10] Shi S G, Gao Y, Yang D S, Shi J and Tian D Y 2021 *IEEE Sensors Journal* **21** 16222
[11] Zamponi R, Chiariotti P, Battista G, Schram C and Castellini P 2020 *Applied Acoustics* **163** 107229
[12] Merino-Martínez R, Salil Luesutthiviboon S, Zamponi R, Alejandro Rubio C A, Ragni D, Sijtsma P, Snellen M and Schram C 2020 *Journal of Sound and Vibration* **470** 115176
[13] Wu Y X, Zhang H C K, Kang J and Boctor E M 2020 *Ultrasonics* **103** 106098
[14] Zhang Z F, Chen S, Xu Z M, He Y S and Li S 2017 *Journal of Sound and Vibration* **396** 108
[15] Li W, S. Zhao S, Zhou C, Qin Y M, Zhu H R and Li S L 2025 *Measurement* **242** 116238
[16] Presezniak F, Zavala P A G, Steenackers G, Janssens K, Arruda J R F, Desmet W and Guillaume P 2012 *Mechanical Systems and Signal Processing* **32** 349
[17] Zavala P A G, Roeck W D, Janssens K, Arruda J R F, Sas P and Desmet W 2011 *Mechanical Systems and Signal Processing* **25** 928
[18] Daubechies I, Defrise M and Mol C D 2004 *Communications on Pure and Applied Mathematics* **57** 1413
[19] Bredies K 2009 *Journal of Inverse and Ill-Posed Problems* **17** 19
[20] Wang Q, Meng C, Ma W N, Wang C and Yu L 2019 *Measurement* **142** 68
[21] Shen L B, Chu Z G, Zhang Y X and Yang Y 2020 *Journal of Low Frequency Noise Vibration and Active Control* **39** 866
[22] Hu M S, Zheng G Q, Su Z, Kong L R and Wang G D 2024 *Energy* **303** 131951
[23] Liu S M, Chen H W, Liu P X, Qin F Z and Fars A 2023 *International Journal of Hydrogen Energy* **48** 34486
[24] Ghazali S M and Baleghi Y 2023 *Knowledge-Based Systems* **275** 110683
[25] Karmarkar N 1984 *Combinatorica* **4** 373
[26] Malyuta D, Reynolds T P, Szmuk M, Lew T, Bonalli R and Pavone M 2022 *IEEE Control Systems Magazine* **42** 40
[27] Zhou X W, Liu F T, Yin Z Y, Jin Y F and Zhang C B 2022 *Computers and Geotechnics* **145** 104701
[28] Li J, Li X R and Zhao L Y 2017 *International Journal of Wavelets Multiresolution and Information Processing* **15** 1750059
[29] Dokuz Y and Tüfekci Z 2022 *Multimedia Tools and Applications* **81** 9969
[30] Zheng P C, Lyu X T and Gong Y 2023 *IEEE Wireless Communications Letters* **12** 1781
[31] Liu Q, Chu N, Yu L, Ning Y and Wu P 2022 *IEEE Transactions on Instrumentation and Measurement* **71** 6501612