

Fast azimuth–frequency sparse beamforming with adaptive filtering for tonal enhancement under low signal-to-noise ratio conditions

JIANG Fengdan^{1,2,3} CHI Cheng^{1,2†} LIU Chonglei^{1,2} JU Donghao^{1,2}
ZHANG Xiaowei^{1,2} HOU Minglei^{1,2} ZHANG Yangfan^{1,2}
LI Yu^{1,2} HUANG Haining^{1,2}

1 Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China

2 Key Laboratory of Science and Technology on Advanced Underwater Acoustic Signal Processing, Chinese Academy of Sciences, Beijing 100190, China

3 University of Chinese Academy of Sciences, Beijing 100049, China

Received January 11, 2026; accepted April 30, 2026

Abstract Enhancing weak tonal signals under low signal-to-noise ratio (SNR) conditions is a fundamental yet challenging problem in underwater acoustic signal processing. In this paper, a sparse fast azimuth–frequency beamforming with adaptive filtering (SFAFBF-AF) method is proposed for robust tonal enhancement using array measurements. The proposed method exploits the sparsity of tonal signals in the joint azimuth–frequency domain to achieve improved resolution. By introducing a whitening transformation, the sparse azimuth–frequency spectrum can be efficiently recovered using a closed-form soft-thresholding operator, thereby avoiding computationally expensive iterative sparse optimization. A simple adaptive filtering is then employed to suppress nonstationary noise components that are unstable over time while further reinforcing temporally stable tonal structures, leading to high enhancement gain. Owing to this design, SFAFBF-AF achieves a favorable balance between performance and computational efficiency. It provides high enhancement gain and good angular resolution while maintaining a relatively fast computation speed. Simulations and experiments demonstrate that the proposed method outperforms conventional wideband beamforming and other reference approaches, particularly for weak or closely spaced tonal targets under low-SNR conditions.

Keywords Tonal enhancement, Sparse beamforming, Adaptive filtering, Spatial-frequency analysis

PACS: 43.30, 43.60

DOI: 10.3724/0217-9776.3026007

CSTR: 32050.14.11-2066.3026007

1 Introduction

In passive underwater acoustic sensing, tonal signals generated by periodic mechanical processes are critical for target detection, identification, and tracking^[1–4]. In practical scenarios, however, such tonal components are often embedded in low signal-to-noise ratio (SNR) conditions, where their energy is insufficient to form a stable contrast against ambient noise within a finite observation interval. As a result, tonal enhancement continues to face persistent challenges in both resolution and robustness.

In many representative ocean environments, background noise can be reasonably approximated as a stationary

Gaussian random process, particularly when it arises from the superposition of a large number of weak and mutually independent sources^[5]. This assumption forms the theoretical foundation of classical tonal enhancement techniques. Under Gaussian noise backgrounds, traditional approaches primarily rely on second-order statistical modeling, employing power spectral analysis, covariance matrix estimation, and beamforming to improve the observability of tonal components^[6–9]. While these methods perform reliably at moderate to high SNRs, their effectiveness degrades significantly in low-SNR conditions. Finite-sample fluctuations of Gaussian noise tend to elevate the spectral floor and introduce spurious peaks, making it difficult to separate weak

This work was supported by the National Natural Science Foundation of China (E3111301).

† Corresponding author: CHI Cheng, chicheng@mail.ioa.ac.cn

tonal lines in the frequency domain and simultaneously leading to degraded spatial resolution. High-resolution subspace-based techniques, such as multiple signal classification (MUSIC)^[10] and estimation of signal parameters via rotational invariance techniques (ESPRIT)^[11], partially alleviate spectral leakage; however, their performance critically depends on accurate covariance estimation and stable signal–noise subspace separation^[12], which remain challenging under low-SNR conditions.

The fundamental difficulty of tonal enhancement at low SNR does not primarily stem from deficiencies in noise statistical modeling, but rather from the insufficient exploitation of the structural characteristics of tonal signals in the azimuth–frequency domain. Unlike Gaussian noise, which exhibits a continuous energy distribution in the frequency domain, tonal signals typically occupy only a limited number of discrete frequency bins and are concentrated in a small set of spatial directions, resulting in an inherently sparse representation in the joint azimuth–frequency domain^[13,14]. This sparsity persists even under low-SNR conditions and provides a promising avenue for overcoming the resolution limits of conventional methods^[15]. Nevertheless, in extremely low-SNR scenarios, sparse azimuth–frequency spectrum estimation based solely on sparsity constraints remains problematic. Noise-dominated random fluctuations may be falsely identified as valid sparse supports, giving rise to numerous isolated artifacts or spurious sparse components^[16,17], which in turn undermine the stability and interpretability of the estimated spectral structure.

Against this background, this paper focuses on tonal enhancement under low-SNR conditions in Gaussian noise environments and proposes an SFAFBF-AF algorithm. The proposed approach first performs the SFAFBF method to jointly model the observed data in the azimuth–frequency domain, explicitly exploiting the sparse structure of tonal components while globally suppressing the noise-dominated spectral floor. By introducing a whitening transformation, this sparse representation can be efficiently obtained using a closed-form soft-thresholding operator, thereby avoiding computationally expensive iterative sparse optimization and yielding an initial high-resolution azimuth–frequency representation.

To further improve robustness, an element-wise adaptive filtering stage is subsequently applied to the estimated sparse azimuth–frequency spectrum. This stage employs a simple adaptive filtering strategy to suppress unstable noise artifacts arising from sparse estimation and background noise, while preserving the continuity and stability of tonal components across both frequency and spatial domains. Owing to the combination of closed-form sparse recovery and lightweight adaptive updates, the proposed method achieves fast convergence with moderate computational cost.

As a result, SFAFBF-AF achieves simultaneous improvements in enhancement gain and azimuth–frequency resolution while maintaining relatively fast computation compared with other high-performance tonal enhancement methods. The subsequent sections detail the implementation of the SFAFBF-AF algorithm and validate its effectiveness through numerical simulations and experimental data under Gaussian noise backgrounds.

2 Data model

As illustrated in Fig. 1, we consider a far-field scenario in which a uniform linear array (ULA) of M sensors is deployed, with an inter-element spacing of d . To represent all possible arrival angles, an overcomplete spatial dictionary is

$$\boldsymbol{\theta} = [\theta_0, \dots, \theta_{R-1}], \quad (1)$$

where R is the number of discretized directions, and is typically much larger than the number of true tonal sources. Each tonal component is assumed to originate from one of these directions, and the signal arriving from direction θ_r is denoted by $x_r(t)$. If a direction θ_r contains no tonal signals, then $x_r(t) = 0$.

Under the far-field assumption, the impinging wavefronts can be approximated as planar. The received signal $y_m(t)$ at the sensor m is

$$y_m(t) = \sum_{r=0}^{R-1} x_r \left(t + \frac{md \cos \theta_r}{c} \right) + n_m(t), \quad (2)$$

where c denotes the speed of sound in water, and $n_m(t)$ represents the noise received by the sensor m .

Each sensor collects a snapshot of N time-domain samples $y_m[n]$, $n = 0, \dots, N-1$. Applying the discrete Fourier transform (DFT) and restricting the analysis to a frequency band of interest yields:

$$Y_m[k] = \sum_{r=0}^{R-1} X_r[k] e^{j2\pi f_k \frac{md \cos \theta_r}{c}} + N_m[k], \quad k = 0, \dots, K-1, \quad (3)$$

where $X_r[k]$ is the DFT of $x_r[n]$. The frequency grid is

$$f_k = \frac{k}{N} f_s + f_0, \quad k = 0, \dots, K-1, \quad (4)$$

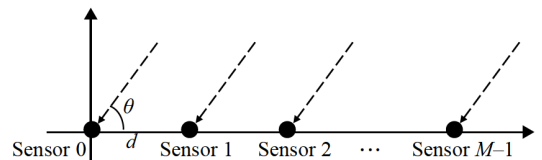


Fig. 1 Far-field signal reception model for a ULA.

where f_s is the sampling frequency, and $f_0 - f_{K-1}$ specify the processing band.

For convenience, we define the vector

$$\mathbf{y}_k = [Y_0[k], Y_1[k], \dots, Y_{M-1}[k]]^T \quad (5)$$

as the array observation at the frequency f_k . Similarly,

$$\mathbf{x}_k = [X_0[k], X_1[k], \dots, X_{R-1}[k]]^T, \quad (6)$$

represents the spectral amplitudes at the frequency f_k associated with all R candidate directions in the dictionary

$$\mathbf{A}_k = \begin{bmatrix} 1 & \dots & 1 \\ \exp\left(j2\pi f_k \frac{d\cos\theta_1}{c}\right) & \dots & \exp\left(j2\pi f_k \frac{d\cos\theta_R}{c}\right) \\ \vdots & \ddots & \vdots \\ \exp\left(j2\pi f_k \frac{(M-1)d\cos\theta_1}{c}\right) & \dots & \exp\left(j2\pi f_k \frac{(M-1)d\cos\theta_R}{c}\right) \end{bmatrix}. \quad (8)$$

We further define a block-diagonal matrix:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_{K-1} \end{bmatrix}, \quad (9)$$

which leads to the compact vectorized model:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \quad (10)$$

where $\mathbf{n} = \text{vec}(\mathbf{N}) \in \mathbb{C}^{RK \times 1}$ denotes the vectorized noise. Although the array manifold matrix $\mathbf{A}$ includes multiple discrete frequency bins, each submatrix is constructed under a narrowband assumption. Accordingly, the formulation in Eq. (10) represents a multi-frequency narrowband model that is intended for enhancing narrowband tonal components rather than broadband sources. From a physical perspective, the proposed joint azimuth–frequency model provides a signal-domain representation of tonal components jointly characterized in the azimuth and frequency dimensions, without aiming to describe broadband wave propagation.

It should be emphasized that the bold symbols $\mathbf{x}$ and $\mathbf{y}$ do not correspond to the time-domain signals $x_r(t)$ and $y_m(t)$ in Eq. (2). Instead, they denote the vectorized frequency-domain representations of $\mathbf{X}$ and $\mathbf{Y}$. Accordingly, x_i and y_i refer to the i th elements of $\mathbf{x}$ and $\mathbf{y}$, respectively, and are unrelated to the original time-domain signals. The time-domain notation $x_r(t)$ and $y_m(t)$ will not be used in the remainder of this paper.

$\mathbf{0}$. By stacking these column vectors $\mathbf{x}_k$ and $\mathbf{y}_k$ over all frequency bins, we form the point-source signal matrix $\mathbf{X} = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{K-1}] \in \mathbb{C}^{R \times K}$ and the received signal matrix $\mathbf{Y} = [\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{K-1}] \in \mathbb{C}^{M \times K}$. Their vectorized forms are

$$\begin{aligned} \mathbf{x} &= \text{vec}(\mathbf{X}) = [\mathbf{x}_0^T, \mathbf{x}_1^T, \dots, \mathbf{x}_{K-1}^T]^T \in \mathbb{C}^{RK \times 1}, \\ \mathbf{y} &= \text{vec}(\mathbf{Y}) = [\mathbf{y}_0^T, \mathbf{y}_1^T, \dots, \mathbf{y}_{K-1}^T]^T \in \mathbb{C}^{MK \times 1}. \end{aligned} \quad (7)$$

At frequency f_k , the array manifold matrix is

The objective is to accurately recover the useful tonal components $\mathbf{x}$ embedded in the noisy observation matrix $\mathbf{y}$. By exploiting both the sparsity and stability of tonal signals, we develop the SFAFBF-AF method presented in the following section.

3 Sparse fast azimuth–frequency beamforming with adaptive filtering

3.1 Sparse fast azimuth–frequency beamforming

In many practical configurations, the number of formed beams R is much larger than that of the array elements M . The forward model in Eq. (3) then becomes severely ill-conditioned, with a large condition number and an unstable inverse mapping from the measurements to the underlying signal. A common way to reconstruct the azimuth–frequency matrix $\mathbf{x}$ is to solve a least-squares problem:

$$\hat{\mathbf{x}}_{\text{MLE}} = \underset{\mathbf{x} \in \mathbb{C}^{RK \times 1}}{\text{argmin}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2, \quad (11)$$

where $\|\cdot\|_2$ is the ℓ_2 -norm. This formulation coincides with the maximum-likelihood estimator when the noise is assumed Gaussian, since it minimizes the Euclidean distance between the observed data $\mathbf{y}$ and the model output $\mathbf{A}\mathbf{x}$.

However, under low-SNR conditions, this approach often exhibits unsatisfactory performance. When the target signal is extremely weak, directly minimizing $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2$ leads the maximum-likelihood estimate $\hat{\mathbf{x}}_{\text{MLE}}$ to favor fitting the noise-dominated background, thereby masking or even

completely ignoring the underlying tonal components. To mitigate this issue, we explicitly constrain the admissible noise power and formulate a regularized estimation problem as

$$\hat{\mathbf{x}}_{\phi}(\varepsilon) = \underset{\mathbf{x} \in \mathbb{C}^{RK \times 1}}{\operatorname{argmin}} \phi(\mathbf{x}), \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \varepsilon, \quad (12)$$

where ε is an upper bound on the expected interference level, and $\phi : \mathbb{C}^{RK \times 1} \rightarrow \mathbb{R}$ is a penalty term chosen to encode prior knowledge of the signal structure.

In passive sonar applications, tonal components are known to be sparse in the frequency domain. Together with the limited number of targets and the overcompleteness of the spatial dictionary $\mathbf{0}$, this implies that $\mathbf{x}$ should be sparse in both azimuth and frequency dimensions, making a sparsity-promoting penalty a natural choice. Accordingly, we set

$$\phi(\mathbf{x}) = \|\mathbf{x}\|_0. \quad (13)$$

Such sparse optimization not only suppresses noise but also provides improved angular resolution, thereby alleviating the wide-mainlobe problem encountered with conventional beamforming.

Direct minimization of the ℓ_0 pseudo-norm in Eq. (12) is non-deterministic polynomial (NP)-hard and computationally prohibitive. A standard relaxation is therefore adopted, replacing ℓ_0 -norm with the ℓ_1 -norm, yielding the tractable formulation:

$$\hat{\mathbf{x}}_{\phi}(\varepsilon) = \underset{\mathbf{x} \in \mathbb{C}^{RK \times 1}}{\operatorname{argmin}} \|\mathbf{x}\|_1, \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \varepsilon. \quad (14)$$

Next, we consider the constrained optimization problem in Eq. (14). The Lagrangian associated with Eq. (14) is

$$L(\mathbf{x}, \lambda) = \|\mathbf{x}\|_1 + \frac{\lambda}{2} (\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 - \varepsilon^2), \quad (15)$$

where λ is the Lagrange multiplier. Taking the gradient of Eq. (15) with respect to $\mathbf{x}$ and setting it to zero yields:

$$\nabla \|\mathbf{x}\|_1 - \lambda (\mathbf{A}^H(\mathbf{y} - \mathbf{A}\mathbf{x})) = 0. \quad (16)$$

When the array manifold $\mathbf{A}_k$ satisfies $\mathbf{A}_k^H \mathbf{A}_k = \mathbf{I}$, we have $\mathbf{A}^H \mathbf{A} = \mathbf{I}$. Let $\mathbf{c} = \mathbf{A}^H \mathbf{y} \in \mathbb{C}^{RK \times 1}$, then, Eq. (16) can be written in an element-wise form as

$$\frac{\partial \|\mathbf{x}\|_1}{\partial x_i} - \lambda(c_i - x_i) = 0. \quad (17)$$

According to classical results in sparse optimization^[18,19], the solution to Eq. (14) is given by the soft-thresholding operator:

$$x_i = \left(|c_i| - \frac{1}{\lambda}, 0 \right)_+ \times \operatorname{sgn}(c_i), \quad (18)$$

where $|\cdot|$ denotes the modulus, $(x)_+ = \max(x, 0)$, and $\operatorname{sgn}(\cdot)$ is the sign function.

Unfortunately, in our problem, the number of spatial grid points R is typically much larger than the number of array elements M , so the array manifold $\mathbf{A}_k$ does not satisfy $\mathbf{A}_k^H \mathbf{A}_k = \mathbf{I}$. As a result, Eq. (18) cannot be directly applied. To address this issue, we first whiten the array manifold by defining

$$\mathbf{W}_k = (\mathbf{A}_k^H \mathbf{A}_k)^{1/2}, \quad (19)$$

and its corresponding block-diagonal matrix is

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{W}_{K-1} \end{bmatrix}. \quad (20)$$

The whitened array manifold is

$$\mathbf{B}_k = \mathbf{A}_k \mathbf{W}_k^{-1}, \quad (21)$$

which satisfies

$$\mathbf{B}_k^H \mathbf{B}_k = \mathbf{W}_k^{-1} \mathbf{A}_k^H \mathbf{A}_k \mathbf{W}_k^{-1} = \mathbf{I}. \quad (22)$$

The corresponding block-diagonal matrix is

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B}_{K-1} \end{bmatrix}. \quad (23)$$

Let $\mathbf{s}_k = \mathbf{W}_k^{-1} \mathbf{x}_k$ and define $\mathbf{s} = [\mathbf{s}_0^T, \mathbf{s}_1^T, \dots, \mathbf{s}_{K-1}^T]^T \in \mathbb{C}^{RK \times 1}$, so that $\mathbf{s} = \mathbf{W}^{-1} \mathbf{x}$. Substituting into Eq. (14) yields

$$\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 = \|\mathbf{y} - \mathbf{B}\mathbf{s}\|_2. \quad (24)$$

If $\|\mathbf{x}\|_1$ and $\|\mathbf{s}\|_1$ are equivalent norms, then Eq. (14) is equivalent to^[20,21]

$$\hat{\mathbf{s}}_{\phi}(\varepsilon) = \underset{\mathbf{s} \in \mathbb{C}^{RK \times 1}}{\operatorname{argmin}} \|\mathbf{s}\|_1, \quad \|\mathbf{y} - \mathbf{B}\mathbf{s}\|_2 \leq \varepsilon. \quad (25)$$

Once $\hat{\mathbf{s}}_{\phi}(\varepsilon)$ is obtained, the solution $\hat{\mathbf{x}}_{\phi}(\varepsilon)$ to Eq. (14) can be recovered via the inverse transformation. Since $\mathbf{B}_k^H \mathbf{B}_k = \mathbf{I}$, the solution $\hat{\mathbf{s}}_{\phi}(\varepsilon)$ to Eq. (25) can be computed directly using Eq. (18), where $\mathbf{c}$ is replaced by $\mathbf{B}^H \mathbf{y}$. However, two issues remain: whether $\|\mathbf{x}\|_1$ and $\|\mathbf{s}\|_1$ are equivalent norms, and how to determine the threshold $1/\lambda$.

We first establish the equivalence of the two norms. For any invertible matrix $\mathbf{T}$, the induced ℓ_1 -norm is

$$\|\mathbf{T}\|_1 = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{T}\mathbf{x}\|_1}{\|\mathbf{x}\|_1}. \quad (26)$$

It follows that

$$\|\mathbf{W}^{-1}\|_1 \|\mathbf{x}\|_1 \geq \|\mathbf{W}^{-1}\mathbf{x}\|_1 = \|\mathbf{s}\|_1, \quad (27)$$

for all $\mathbf{x}$ and by the properties of the ℓ_1 -norm,

$$\|\mathbf{x}\|_1 = \|\mathbf{W}\mathbf{s}\|_1 \leq \|\mathbf{W}\|_1 \|\mathbf{W}^{-1}\mathbf{x}\|_1. \quad (28)$$

Combining Eqs. (27) and (28) yields

$$\frac{\|\mathbf{x}\|_1}{\|\mathbf{W}\|_1} \leq \|\mathbf{s}\|_1 \leq \|\mathbf{W}^{-1}\|_1 \|\mathbf{x}\|_1. \quad (29)$$

From Eq. (19), it follows that for a fixed array manifold, which is specified by the array geometry and the angular grid $\boldsymbol{\theta}$, both $\|\mathbf{W}\|_1$ and $\|\mathbf{W}^{-1}\|_1$ are finite and nonzero constants. Consequently, $\|\mathbf{x}\|_1$ and $\|\mathbf{s}\|_1$ are equivalent norms up to constant scaling factors determined by $\|\mathbf{W}\|_1$ and $\|\mathbf{W}^{-1}\|_1$. This equivalence does not imply exact pointwise preservation of the support set. Instead, it allows us to quantify the effect of replacing $\|\mathbf{x}\|_1$ with $\|\mathbf{s}\|_1$: the solution obtained by minimizing $\|\mathbf{s}\|_1$ is guaranteed to be within a factor $\|\mathbf{W}\|_1 \|\mathbf{W}^{-1}\|_1$ of the optimal solution measured in $\|\mathbf{x}\|_1$. A detailed derivation is provided in the Appendix.

It is also worth noting that, although the whitening transformation does not guarantee that the transformed variable is sparse in a strict ℓ_0 sense, applying an ℓ_1 -norm regularization in the whitened domain remains theoretically justified, as it corresponds to a weighted ℓ_1 constraint on the original variable and preserves the sparsity-promoting behavior up to constant scaling. In practice, solving for $\mathbf{s}$ suffices, and the solution for $\mathbf{x}$ can be obtained accordingly.

Next, we describe the procedure for determining the threshold $1/\lambda$. Suppose that $1/\lambda$ is given and $\hat{\mathbf{s}}_\phi(\varepsilon)$ has been obtained via Eq. (18). For notational simplicity, we denote $\hat{\mathbf{s}}_\phi(\varepsilon)$ by $\hat{\mathbf{s}}$, and define the residual between the observation $\mathbf{y}$ and its reconstruction as

$$\mathbf{r} = \mathbf{y} - \mathbf{B}\hat{\mathbf{s}}. \quad (30)$$

This can be rewritten as

$$\mathbf{r} = \mathbf{B}(\mathbf{c} - \hat{\mathbf{s}}) + (\mathbf{I} - \mathbf{B}\mathbf{B}^H)\mathbf{y}. \quad (31)$$

Since $\mathbf{B}^H(\mathbf{I} - \mathbf{B}\mathbf{B}^H) = \mathbf{0}$, we obtain

$$\|\mathbf{r}\|_2^2 = \|\mathbf{B}(\mathbf{c} - \hat{\mathbf{s}})\|_2^2 + \|(\mathbf{I} - \mathbf{B}\mathbf{B}^H)\mathbf{y}\|_2^2. \quad (32)$$

Using $\mathbf{B}^H\mathbf{B} = \mathbf{I}$, it follows that

$$\|\mathbf{B}(\mathbf{c} - \hat{\mathbf{s}})\|_2^2 = \|\mathbf{c} - \hat{\mathbf{s}}\|_2^2. \quad (33)$$

Substituting into Eq. (32) and expanding element-wise gives

$$\|\mathbf{r}\|_2^2 = \sum_{i=0}^{RK-1} (c_i - \hat{s}_i)^2 + \|(\mathbf{I} - \mathbf{B}\mathbf{B}^H)\mathbf{y}\|_2^2. \quad (34)$$

From Eq. (18), we have

$$|c_i - \hat{s}_i| = \min\left(|c_i|, \frac{1}{\lambda}\right), \quad (35)$$

which leads to

$$\|\mathbf{r}\|_2^2 = \sum_{i=0}^{RK-1} \min\left(|c_i|, \frac{1}{\lambda}\right)^2 + \|(\mathbf{I} - \mathbf{B}\mathbf{B}^H)\mathbf{y}\|_2^2. \quad (36)$$

The second term in Eq. (36) does not represent noise energy alone. Instead, it corresponds to the portion of the observation signal $\mathbf{y}$ that cannot be represented in the subspace spanned by the matrix $\mathbf{B}$. This includes both noise components projecting onto the orthogonal complement of the dictionary $\mathbf{B}$ and signal components that are not covered by the assumed model, which are modeling errors. This term may vary significantly when the data are highly unstable.

Due to the presence of overfitting as analyzed previously, a smaller residual norm does not necessarily indicate a better solution. In practice, we aim to achieve $\|\mathbf{r}\|_2$ smaller than ε . To this end, we sort $d_i = |c_i|$ ($i = 0, \dots, RK-1$) in descending order to obtain $\tilde{\mathbf{d}}$. If $1/\lambda \in [\tilde{d}_{j+1}, \tilde{d}_j)$, where $\tilde{d}_j$ is the j th element of $\tilde{\mathbf{d}}$, then,

$$\|\mathbf{r}\|_2^2 = \frac{j+1}{\lambda^2} + \sum_{i=j+1}^{RK-1} \tilde{d}_i^2 + \|(\mathbf{I} - \mathbf{B}\mathbf{B}^H)\mathbf{y}\|_2^2. \quad (37)$$

Setting $\|\mathbf{r}\|_2 = \varepsilon$ yields

$$\frac{1}{\lambda} = \sqrt{\frac{\varepsilon^2 - \sum_{i=j+1}^{RK-1} \tilde{d}_i^2 - \|(\mathbf{I} - \mathbf{B}\mathbf{B}^H)\mathbf{y}\|_2^2}{j+1}}. \quad (38)$$

If the resulting $1/\lambda$ does not satisfy the interval constraint, we update $j \leftarrow j+1$ and repeat the procedure.

The noise bound ε can be chosen according to the background noise level. Since the proposed method is designed for low-SNR scenarios, we estimate

$$\varepsilon^2 = \beta \sum_{i=0}^{K-1} |y_{i:M}|^2, \quad (39)$$

where β is a scaling factor depending on the array configuration, frequency band, and detection false-alarm rate. Larger β leads to stronger sparsity, while smaller β yields solutions closer to the original data.

In extreme cases where ε is either very large or very small, Eq. (37) may not admit a valid solution. Since Eq. (37) decreases monotonically with j , we set $1/\lambda = \tilde{d}_0$ when $\varepsilon^2 > \sum_{i=0}^{RK-1} \tilde{d}_i^2 + \|(\mathbf{I} - \mathbf{B}\mathbf{B}^H)\mathbf{y}\|_2^2$, and $1/\lambda = 0$ when $\varepsilon^2 < RK\tilde{d}_{RK-1}^2 + \|(\mathbf{I} - \mathbf{B}\mathbf{B}^H)\mathbf{y}\|_2^2$.

Finally, given the estimated $1/\lambda$, the solution of Eq. (25) is

$$\hat{s}_i = \left(|c_i| - \frac{1}{\lambda}, 0\right)_+ \times \text{sgn}(c_i), \quad (40)$$

and the solution to Eq. (14) is recovered by the inverse transformation:

$$\hat{\mathbf{x}}_\phi = \mathbf{A}^H \mathbf{A} \mathbf{W}^{-1} \hat{\mathbf{s}}. \quad (41)$$

Reshaping $\hat{\mathbf{x}}_\phi$ according to Eq. (7) yields the recovered sparse azimuth–frequency spectrum $\hat{\mathbf{X}}_\phi$.

3.2 Adaptive filtering

Due to the presence of noise, the recovered sparse azimuth–frequency spectrum $\hat{\mathbf{X}}_\phi$ still contains a number of spurious components, which become particularly pronounced under low-SNR conditions. Fortunately, these noise-induced components are generally unstable over time, whereas the true tonal components exhibit relatively stable temporal behavior. Motivated by this observation, we incorporate the idea of adaptive line enhancement (ALE) and apply an element-wise adaptive filtering process to $\hat{\mathbf{X}}_\phi$ to suppress noise while reinforcing stable tonal components.

Specifically, a block-processing framework^[22,23] is adopted, in which the sensor observations are divided into overlapping frames to preserve temporal continuity. For l th frame, the sparse azimuth–frequency spectrum $\hat{\mathbf{x}}_\phi^l$ is obtained using the beamforming method described in Sec. 3.1. An adaptive weight matrix $\mathbf{w}^l \in \mathbb{C}^{RK \times N_L}$ is then introduced to represent the adaptive filter coefficients at the l th update, where N_L denotes the filter order.

We assume that, for the current estimate $\hat{\mathbf{x}}_\phi^l$, the noise components are uncorrelated with the previous N_L estimates $\hat{\mathbf{x}}_\phi^{l-i}$, $i = 1, \dots, N_L$, while the tonal components remain temporally correlated. Under this assumption, minimizing the mean squared error (MSE) between the adaptive filter input $\hat{\mathbf{x}}_\phi^l$ and its output $\mathbf{u}_\phi^l$ effectively suppresses noise and enhances stable tonal components. The output of the adaptive filtering stage is

$$\mathbf{u}_\phi^l = \sum_{i=1}^{N_L} \bar{\mathbf{w}}_i^l \otimes \hat{\mathbf{x}}_\phi^{l-i}, \quad (42)$$

where $\otimes$ denotes element-wise multiplication, $\bar{(\cdot)}$ denotes

the complex conjugate of elements, and $\mathbf{w}_i^l$ represents the i th column of $\mathbf{w}^l$.

For updating $\mathbf{w}^l$, we adopt the least mean square (LMS) algorithm, a fundamental and efficient method in ALE-based filtering. The update rule is

$$\mathbf{w}_i^{l+1} = \mathbf{w}_i^l + \mu \bar{\mathbf{e}}^l \otimes \hat{\mathbf{x}}_\phi^{l-i}, \quad (43)$$

where $\mathbf{e}^l = \hat{\mathbf{x}}_\phi^l - \mathbf{u}_\phi^l \in \mathbb{C}^{RK \times 1}$ represents the error tensor between the adaptive filter output and the adaptive reference signal, and μ is the LMS step-size. It should be noted that the sparse estimate obtained in the previous stage is not regarded as an exact desired signal, but serves as an adaptive reference that emphasizes temporally stable tonal components.

Finally, the element-wise output $\mathbf{u}_\phi^l$ is reshaped to obtain the adaptively filtered sparse azimuth–frequency spectrum $\mathbf{U}_\phi^l$. In the proposed SFAFBF method, the sparsity constraint significantly improves the spatial and spectral resolution of tonal components, while the subsequent adaptive filtering further refines the output by suppressing nonstationary noise and enhancing tonal gain. Moreover, the entire procedure does not rely on prior knowledge such as the number of sources or the statistical characteristics of the noise, which substantially improves the robustness of the proposed SFAFBF-AF method. A flowchart of the algorithm is shown in Fig. 2, and the effectiveness of the proposed approach is validated through simulations and experimental results in the following section.

4 Simulations

4.1 Single weak tonal target

We first evaluate the capability of the proposed SFAFBF-AF algorithm in enhancing weak tonal targets.

A ULA consisting of 32 sensors with an inter-element spacing of 5 m is considered. The sampling frequency is set to 300 Hz. The target emits a tonal signal at 50 Hz from an azimuth angle of 70°. Additive white Gaussian noise is added

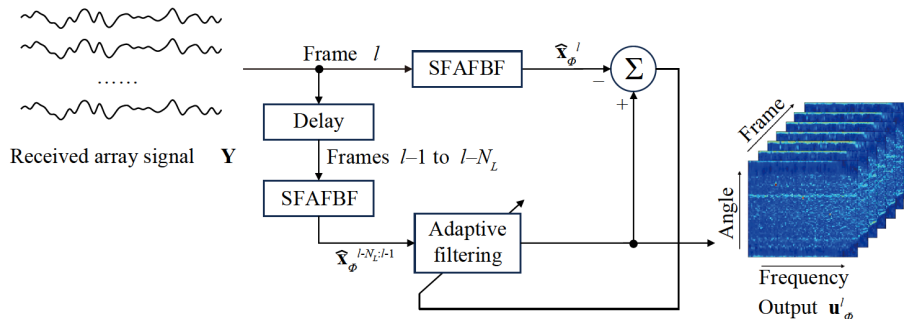


Fig. 2 Flow diagram of the SFAFBF-AF algorithm.

to the received array signals, resulting in a single-channel SNR of approximately -25 dB within a 1 Hz bandwidth. The time-domain waveform and time-frequency representation of the signal received by sensor 0 are shown in Fig. 3.

A total duration of 900 s of array data is used for processing. The data are segmented into overlapping frames with a frame length of 10 s and an update interval of 1 s. The processing frequency band is selected as 45–55 Hz with a frequency resolution of 0.1 Hz, and the azimuth grid spacing is set to 1° . The sound speed in water is assumed to be constant at 1500 m/s.

To evaluate performance, three reference algorithms are selected for comparison: conventional wideband beamforming (CWBF), the previously introduced SFAFBF method, and a hybrid approach termed conventional wideband beamforming with adaptive filtering (CWAF), which combines CWBF with the adaptive filtering described in Sec. 3.2. In the CWAF framework, the SFAFBF module in the SFAFBF-AF pipeline (Fig. 2) is replaced by CWBF, allowing the contribution of SFAFBF within the overall SFAFBF-AF framework to be isolated and assessed. Comparative analyses under different conditions are conducted to evaluate the effectiveness of the proposed SFAFBF-AF algorithm. For CWAF and SFAFBF-AF, the adaptive filter order is set to 150, and the parameter β in SFAFBF-AF is set to 28.5. The processing results of different algorithms are shown in Figs. 4 and 5.

Figures 4(a), 4(c), 4(e), and 4(g) show the three dimensional (3D) time–frequency–azimuth point cloud results. The sparse point clouds are normalized and thresholded to suppress background noise. Figures 4(b), 4(d), 4(f), and 4(h) show the corresponding 2D frequency–azimuth spectrum U_ϕ^f . It can be observed that, in the 2D azimuth–frequency domain, only the SFAFBF-AF algorithm clearly reveals the target. In the 3D point-cloud representations, both CWAF and SFAFBF-AF are able to capture the target trajectory, whereas the CWAF results contain a considerable number of noise points and interferences. In contrast, the SFAFBF-AF

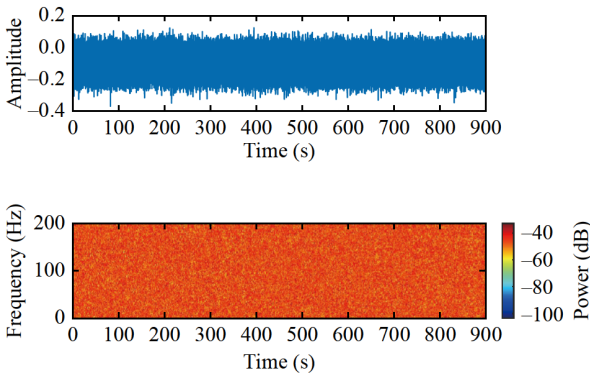


Fig. 3 Time-domain waveform and time–frequency spectrum of the received signal at sensor 0 in the single weak target scenario.

output is significantly cleaner.

To obtain the azimuth spectrum shown in Fig. 5, the frequency–azimuth spectrum U_ϕ^f is traversed across all beams, selecting the maximum value at each beam over all frequency bins. As shown in Fig. 5, SFAFBF-AF achieves the strongest processing gain. Although SFAFBF and CWAF exhibit peaks near 70° , these peaks are relatively weak and easily obscured by other clutter or interference. The CWBF, by contrast, fails to produce a distinct peak even near the true target direction.

It is worth noting that although CWBF alone cannot effectively enhance the target tonal line, the introduction of adaptive filtering allows CWAF to reveal a clearer target trajectory in the 3D point-cloud representation, demonstrating the effectiveness of adaptive filtering. The SFAFBF can provide an initial enhancement of the tonal signal; however, as discussed earlier, its processing gain is limited, and the result contains numerous spurious points, preventing a clear localization of the target in either the frequency–azimuth spectrum or the azimuth spectrum. By combining SFAFBF with adaptive filtering, the proposed SFAFBF-AF algorithm consistently produces the clearest target trajectories and most accurate localization results across the 2D frequency–azimuth spectrum, 3D point clouds, and azimuth spectra. Compared with the other methods, SFAFBF-AF achieves the largest processing gain and exhibits a clear advantage under high-noise conditions.

Specifically, the measured processing gains are 6.09 dB for CWBF, 14.35 dB for SFAFBF, 12.08 dB for CWAF, and 31.23 dB for SFAFBF-AF. These results demonstrate that the proposed method achieves the largest enhancement gain among the compared approaches under identical conditions. Notably, although CWAF exhibits a lower processing gain than SFAFBF, Fig. 4 shows that CWAF produces more continuous and visually coherent tonal trajectories. This indicates that visual separability in the azimuth–frequency domain is not solely determined by processing gain, especially under low-SNR conditions.

All simulations were conducted on a laptop equipped with an Intel Core i9-12900HK 3.20-GHz processor and 64 GB of random access memory (RAM). Each algorithm was executed multiple times, and the average computational time was recorded. The average processing times are 5.3218 s for CWBF, 10.4440 s for CWAF, 9.5340 s for SFAFBF, and 16.0568 s for SFAFBF-AF. From a computational complexity perspective, for each snapshot, CWBF mainly involves conventional beamforming operations with complexity on the order of $O(KRM)$, meaning that the computational cost increases approximately in proportion to KRM , with constant factors and lower-order terms omitted. The CWAF introduces additional adaptive filtering operations across multiple snapshots, resulting in increased computational

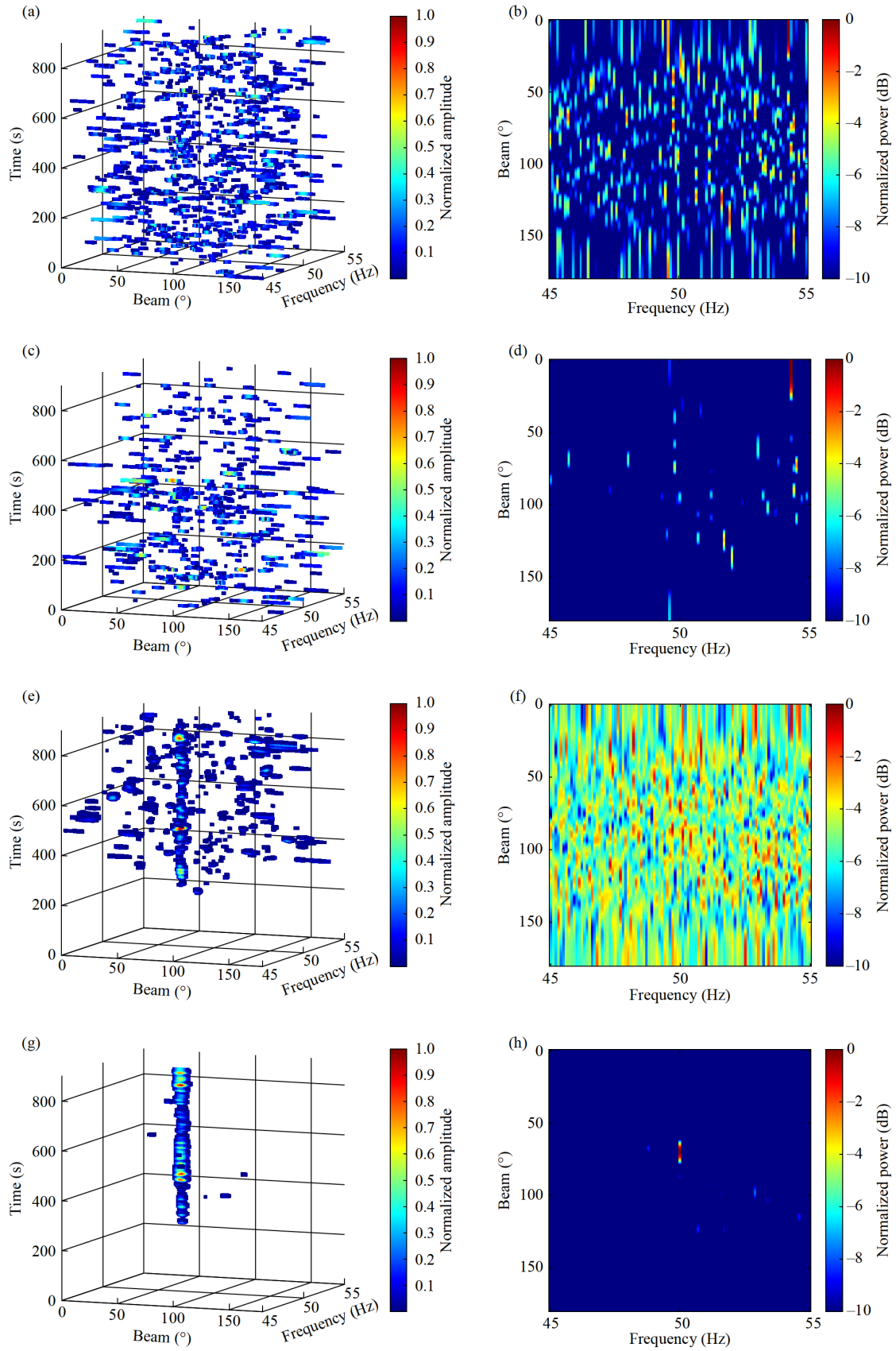


Fig. 4 2D frequency–azimuth spectra and 3D time–frequency–azimuth point clouds of different algorithms in the single weak target scenario: (a, b) CWBF; (c, d) SFAFBF; (e, f) CWAF; (g, h) SFAFBF-AF.

cost. The proposed SFAFBF-AL further incorporates sparse estimation and adaptive filtering stages, resulting in a higher computational complexity than both CWBF and CWAF. Consistent with this analysis, all four algorithms operate within the same order of magnitude in terms of computational cost. Although SFAFBF-AF requires a slightly longer computation time, the performance improvement it provides justifies the additional computational overhead.

To further evaluate the robustness of the proposed method, we added two quantitative experiments in the revised manuscript. Except for the input SNR, all other simulation settings were identical to those described above.

First, Monte Carlo simulations were conducted under a range of input SNR conditions, and the average processing gain, together with error bars, was reported, as shown in Fig. 6. The circles denote the mean processing gain over 100 repeated trials, and the vertical error bars represent the corresponding standard deviations. The results show that the proposed method maintains stable performance across repeated noise realizations. As the input SNR increases, the average processing gain increases monotonically, while the standard deviation gradually decreases.

Second, the processing gain was compared as a function of input SNR for all reference methods, as shown in Fig. 7. The results show that the proposed method consistently achieves the highest processing gain over the entire tested

SNR range, and its advantage over the comparison methods becomes more pronounced as the input SNR increases.

These results provide statistical evidence that the superiority of the proposed method is consistently observed across repeated trials and different SNR conditions, rather than being limited to a single illustrative example or a specific noise realization.

4.2 Multiple closely spaced tonal targets

Next, we evaluate the resolution capability of the proposed SFAFBF-AF algorithm for closely spaced tonal targets.

Two tonal targets are assumed to be located at azimuth angles of 70° and 75° , emitting signals at frequencies of 50 and 50.5 Hz, respectively. As in the previous simulation, additive white Gaussian noise is added to the received signals, resulting in a single-channel SNR of approximately -20 dB within a 1-Hz bandwidth for both targets. The time-domain waveform and time-frequency representation of the signal received by sensor 0 are shown in Fig. 8.

A total duration of 600 s of array data is processed. The parameter β in SFAFBF-AF is set to 21, while all other parameters remain identical to those used in Sec. 4.1. The azimuth spectra obtained by different algorithms are shown in Fig. 9.

As can be observed, CWBF and SFAFBF are only able to enhance the target located at 75° , while CWAF enhances only the target at 70° . These methods therefore fail to resolve both targets simultaneously, which suppresses one target and enhances the other. In contrast, the proposed SFAFBF-AF algorithm successfully resolves both targets at distinct azimuth angles, demonstrating its superior angular resolution.

As in Sec. 4.1, the average computational times are also evaluated. The average processing times are 4.0600 s for CWBF, 8.1563 s for CWAF, 7.7159 s for SFAFBF, and 11.5893 s for SFAFBF-AF. The computational costs of the four algorithms remain comparable. Considering the significant performance advantage in terms of resolution and robustness, SFAFBF-AF offers clear benefits for practical applications.

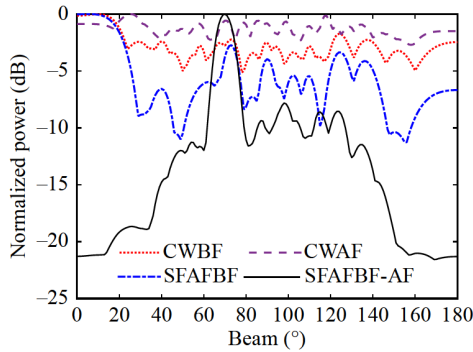


Fig. 5 Azimuth spectrum outputs of different algorithms in the single weak target scenario.

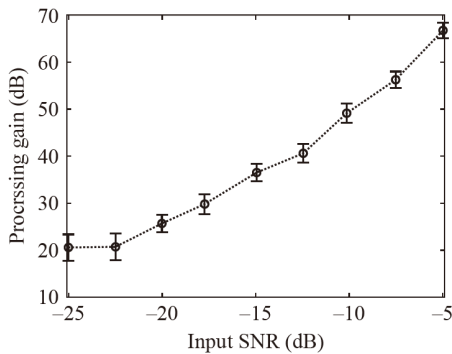


Fig. 6 Average processing gain of the proposed method under different input SNR conditions over 100 repeated trials.

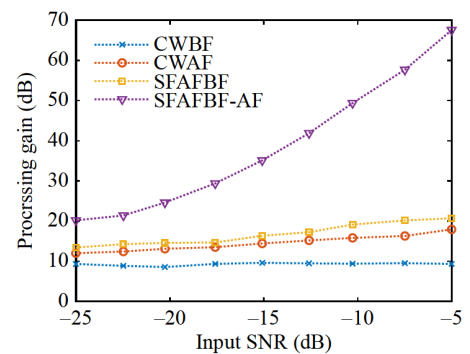


Fig. 7 Comparison of processing gain versus input SNR for the proposed method and the reference methods.

4.3 Coherent tonal targets

This scenario evaluates the robustness of each algorithm when processing multiple coherent tonal sources. In this scenario, two tonal components share the same frequency (50 Hz) and phase, with azimuths located at 70° and 115° , respectively. All processing parameters remain identical to those used in Sec. 4.2.

The results shown in Fig. 10 indicate that the proposed SFAFBF-AF method remains capable of enhancing the

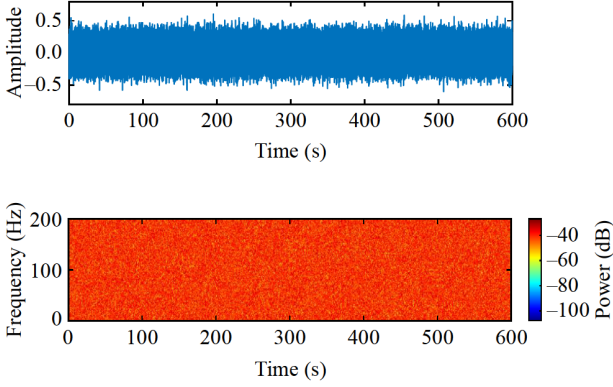


Fig. 8 Time-domain waveform and time-frequency spectrum of the received signal at sensor 0 in the multiple closely spaced tonal targets scenario.

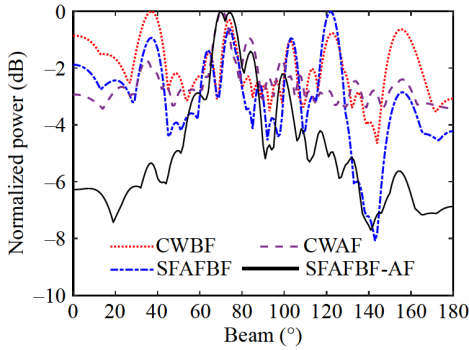


Fig. 9 Azimuth spectrum outputs of different algorithms in the multiple closely spaced tonal targets scenario.

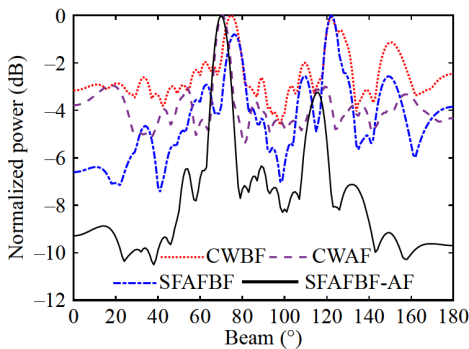


Fig. 10 Azimuth spectrum outputs of different algorithms for two coherent tonal targets under low-SNR conditions.

coherent target tonal components under low-SNR conditions, while baseline methods exhibit degraded separation and increased spectral leakage. These results provide direct empirical evidence that the claimed robustness of the proposed method extends to coherent source scenarios.

4.4 Selection and sensitivity of parameter β

To investigate the influence of β , we evaluated the processing gain of the proposed method as a function of β under three representative input SNR conditions, as shown in Fig. 11. Except β and the input SNR of the target signal, all other simulation parameters were identical to those used in Sec. 4.1.

The results show a consistent overall trend across the three conditions. As β increases, the gain first improves, then remains relatively high over a moderate interval, and finally decreases when β becomes too large. This behavior is consistent with the role of β in controlling the tolerance level of the data-fidelity constraint. When β is too small, the feasible set becomes overly restrictive, which may suppress the weak tonal components. When β is too large, the constraint is overly relaxed, allowing more residual noise and interference to enter the solution and thereby degrading the enhancement performance.

These results also indicate that the optimal β varies moderately with the noise level and generally becomes smaller as the input SNR decreases. This trend can be explained by the fact that, under lower-SNR conditions, the reconstruction needs to remain more consistent with the observed data to avoid suppressing weak target components.

More importantly, the proposed method still maintains relatively high performance over a contiguous interval of β , rather than at a single isolated point. This suggests that β should be interpreted as a tunable tolerance parameter selected within a stable operating range, and that the performance improvement is not due to a narrowly tuned empirical choice.

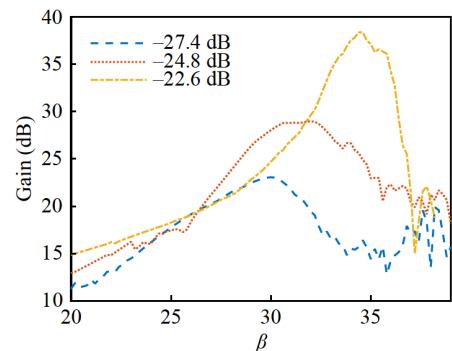


Fig. 11 The processing gain of the proposed method for the parameter β under three representative input SNR conditions.

5 Experiment

A field experiment was conducted in January 2024 at an ice-covered lake. The water depth in the experimental area was approximately 33 m, and the measured sound-speed profile is shown in Fig. 12. No fishing nets, aquatic vegetation, or other obstructions were present beneath the ice, and the lakebed was relatively flat, providing favorable conditions for acoustic signal propagation.

The horizontal receiving array used in the experiment consisted of 16 sensors with an inter-element spacing of 0.1875 m, deployed at a depth of approximately 15 m. The sampling frequency was set to 100 kHz. A 160-s segment of the recorded noise-only data was selected as background noise, which contained both natural environmental noise, such as flow-induced noise, and artificial interference. During processing, the noise data were first downsampled to 10 kHz. A tonal target was then synthetically added with an azimuth angle of 120° and a frequency of 3000 Hz, resulting in a single-channel SNR of approximately -25 dB within a 1-Hz bandwidth. The time-domain waveform and time-frequency representation of the signal received by channel 0 after adding the tonal target are shown in Fig. 13.

The processing frequency band was selected as 2900–

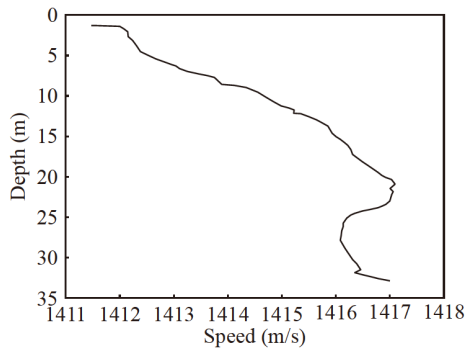


Fig. 12 Sound speed profile of the water column at the experimental site.

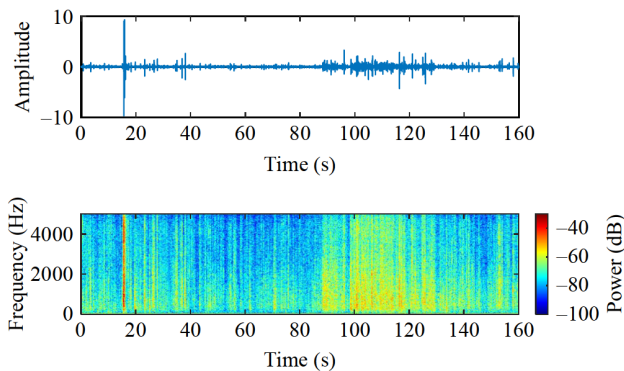


Fig. 13 Time-domain waveform and time-frequency spectrum of data in the sensor 0.

3100 Hz. The data were segmented into overlapping frames with a frame length of 5 s and an update interval of 0.5 s. The frequency resolution was set to 1 Hz, and the azimuth grid spacing was 1° . During processing, the sound speed was assumed to be constant at 1415 m/s. The results obtained using different algorithms are presented in Figs. 14 and 15. The average computational times are 5.5387 s for CWBF, 7.1976 s for CWAF, 13.8592 s for SFAFBF, and 16.3105 s for SFAFBF-AF. The relatively longer computation time in this experiment, compared with the simulation results, is mainly attributed to the larger number of frequency bins involved in the processing.

Consistent with the simulation results, neither CWBF nor SFAFBF is able to effectively enhance the tonal signal. Although CWAF can reveal the target trajectory in the 3D point-cloud representation, it suffers from slow convergence and produces a large number of spurious noise points, making it difficult to accurately identify the target azimuth and frequency in the 2D frequency–azimuth spectrum. In contrast, the proposed SFAFBF-AF algorithm achieves the best performance: clear target trajectories, accurate azimuth estimation, and precise frequency localization can be observed in both the 2D frequency–azimuth spectrum and the 3D point-cloud representation. These results demonstrate the effectiveness and robustness of the proposed method under real-world background noise conditions.

In particular, the background noise contains evident man-made interference in the 100–150 s interval. This scenario goes beyond the assumption of stationary Gaussian noise and implicitly demonstrates the stability and robustness of the proposed method under non-ideal noise environments.

6 Conclusion

This paper presented the SFAFBF-AF framework for enhancing weak tonal signals in challenging low-SNR underwater acoustic environments. By exploiting the sparsity of tonal signals in the azimuth–frequency domain and their temporal stability, the proposed method integrates sparse beamforming with adaptive filtering to effectively suppress noise, while achieving high enhancement gain and improved angular resolution. This enables reliable detection and separation of weak or closely spaced tonal targets under low-SNR conditions. Moreover, the use of a whitening-based closed-form sparse solver combined with a simple adaptive filtering strategy allows SFAFBF-AF to maintain a relatively low computational burden. As a result, SFAFBF-AF achieves strong performance while retaining relatively fast computation.

Simulation results show that SFAFBF-AF consistently outperforms CWBF, SFAFBF, and CWAF in terms of

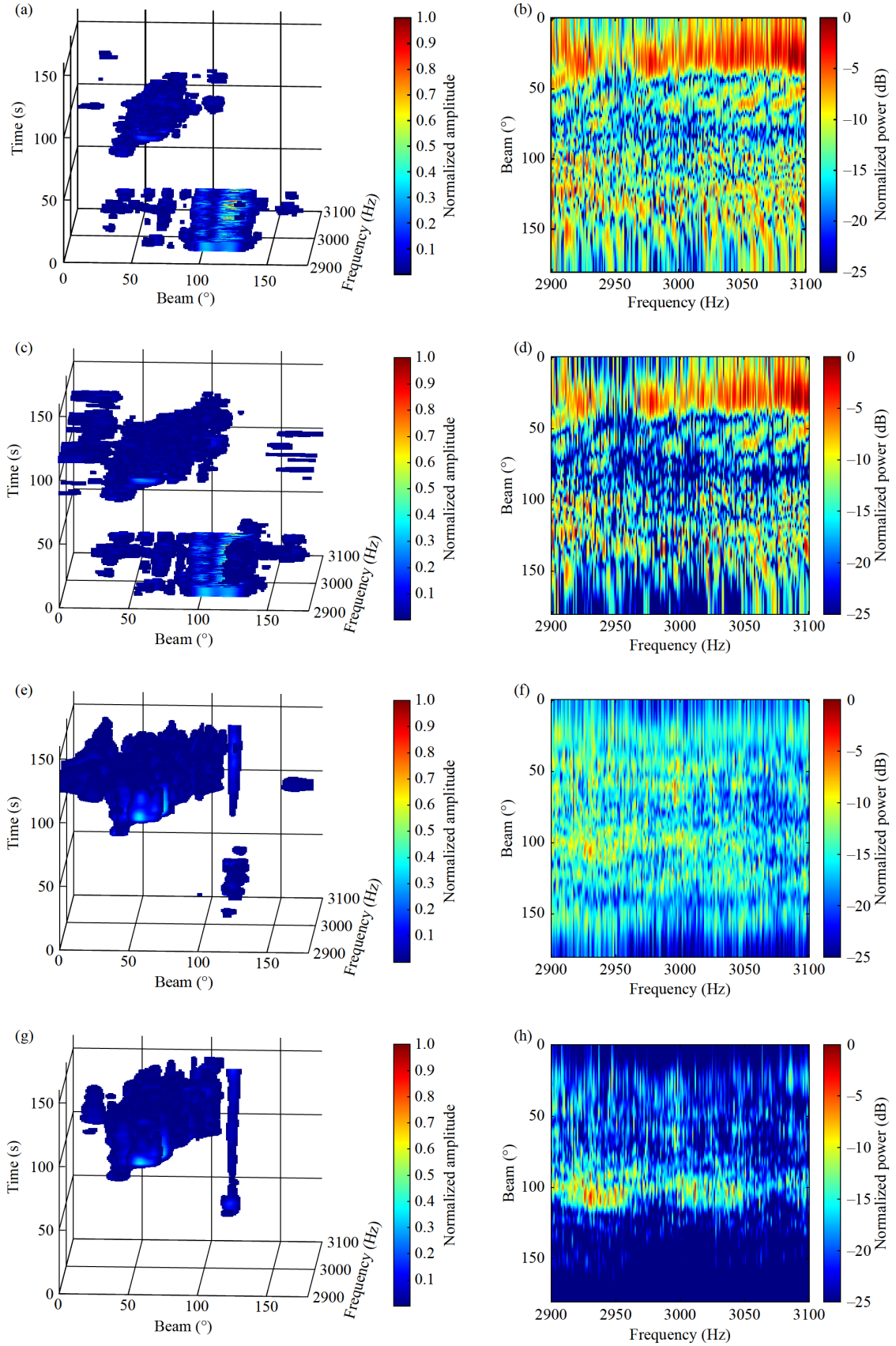


Fig. 14 2D frequency–azimuth spectra and 3D time–frequency–azimuth point clouds of different algorithms in the experiment: (a, b) CWBF; (c, d) SFAFBF; (e, f) CWAF; (g, h) SFAFBF-AF.

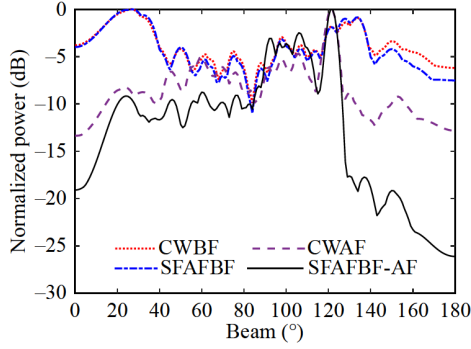


Fig. 15 Azimuth spectrum outputs of different algorithms in the experiment.

enhancement gain and resolution, and field experiments further confirm its robustness and effectiveness under real background noise conditions. Overall, these results demonstrate that SFAFBF-AF provides a practical and efficient solution for tonal enhancement, offering a favorable trade-off between performance and computational cost, and making it well-suited for real-time or near-real-time passive acoustic monitoring in complex underwater environments.

References

- [1] Li Q, Li M, Yang X. The detection of single frequency component of underwater radiated noise of target: theoretical analysis. *Acta Acust.*, 2008, **33**(3): 193–196 (in Chinese)
- [2] McKenna M F, Ross D, Wiggins S M, et al. Underwater radiated noise from modern commercial ships. *J. Acoust. Soc. Am.*, 2012, **131**(1): 92–103
- [3] Ju D H, Chi C, Li Z G, et al. Deep-learning-based line enhancer for passive sonar systems. *IET Radar Sonar Navig.*, 2022, **16**(3): 589–601
- [4] Wang Y, Li Z, Chi C, et al. Single vector hydrophone time-domain deconvolution bearing estimation method based on combined second-order statistics. *Acta Acust.*, 2023, **48**(4): 656–667 (in Chinese)
- [5] Wenz G M. Acoustic ambient noise in the ocean: spectra and sources. *J. Acoust. Soc. Am.*, 1962, **34**(12): 1936–1956
- [6] Stoica P, Moses R L. Spectral analysis of signals. Upper Saddle River, NJ, USA: Pearson Prentice Hall, 2005
- [7] Thomson D J. Spectrum estimation and harmonic analysis. *Proc. IEEE*, 1982, **70**(9): 1055–1096
- [8] Li Q. A new architecture of adaptive line enhancer for towed line array. *Acta Acust.*, 1988, **13**(3): 167–173 (in Chinese)
- [9] Wang G, Huang H, Chi C. Line spectrum target enhancement detection in the frequency–wavenumber domain. *Acta Acust.*, 2024, **49**(4): 647–655 (in Chinese)
- [10] Schmidt R. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas. Propag.*, 1986, **34**(3): 276–280
- [11] Roy R, Kailath T. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoust. Speech Signal Process.*, 1989, **37**(7): 984–995
- [12] Kaveh M, Barabell A. The statistical performance of the MUSIC and

the minimum-norm algorithms in resolving plane waves in noise.

- IEEE Trans. Acoust. Speech Signal Process.*, 1986, **34**(2): 331–341
- [13] Malioutov D, Cetin M, Willsky A S. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Trans. Signal Process.*, 2005, **53**(8): 3010–3022
- [14] Cotter S F, Rao B D, Engan K, et al. Sparse solutions to linear inverse problems with multiple measurement vectors. *IEEE Trans. Signal Process.*, 2005, **53**(7): 2477–2488
- [15] Yang Z, Xie L H, Zhang C S. Off-grid direction of arrival estimation using sparse Bayesian inference. *IEEE Trans. Signal Process.*, 2013, **61**(1): 38–43
- [16] Candès E J, Romberg J K, Tao T. Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.*, 2006, **59**(8): 1207–1223
- [17] Chen Y L, Gu Y T, Hero A O. Sparse LMS for system identification. International Conference on Acoustics, Speech and Signal Processing, IEEE, Taipei, China, 2009: 3125–3128
- [18] Daubechies I, Defrise M, De Mol C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.*, 2004, **57**(11): 1413–1457
- [19] Parikh N, Boyd S. Proximal algorithms. *Found. Trends Optim.*, 2014, **1**(3): 127–239
- [20] Boyd S, Vandenberghe L. Convex optimization. Cambridge, UK: Cambridge University Press, 2004
- [21] Bruckstein A M, Donoho D L, Elad M. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Rev.*, 2009, **51**(1): 34–81
- [22] Haykin S. Block-adaptive filters. In: Adaptive filter theory. 5th ed. Upper Saddle River, NJ, USA: Pearson Education, 2002: 357–393
- [23] Hao Y, Chi C, Qiu L H, et al. Sparsity-based adaptive line enhancer for passive sonars. *IET Radar Sonar Navig.*, 2019, **13**(10): 1796–1804

Appendix

Define

$$\begin{aligned} f(\mathbf{s}) &= \|\mathbf{s}\|_1, \\ g(\mathbf{s}) &= \|\mathbf{W}\mathbf{s}\|_1 = \|\mathbf{x}\|_1. \end{aligned} \quad (\text{A1})$$

From Eq. (29), there exist constants $\alpha_1 = 1 / \|\mathbf{W}\|_1$ and $\alpha_2 = \|\mathbf{W}^{-1}\|_1$ such that

$$\alpha_1 f(\mathbf{s}) \leq g(\mathbf{s}) \leq \alpha_2 f(\mathbf{s}), \quad (\text{A2})$$

for all $\mathbf{s}$. Defining the feasible set F :

$$F = \{\mathbf{s} : \|\mathbf{y} - \mathbf{B}\mathbf{s}\|_2 \leq \varepsilon\}, \quad (\text{A3})$$

and denoting

$$\begin{aligned} \mathbf{s}_f &= \underset{\mathbf{s} \in F}{\operatorname{argmin}} f(\mathbf{s}), \\ \mathbf{s}_g &= \underset{\mathbf{s} \in F}{\operatorname{argmin}} g(\mathbf{s}), \end{aligned} \quad (\text{A4})$$

it follows that

$$g(\mathbf{s}_f) \leq \alpha_2 f(\mathbf{s}_f) \leq \alpha_2 f(\mathbf{s}_g) \leq \frac{\alpha_2}{\alpha_1} g(\mathbf{s}_g). \quad (\text{A5})$$

This result provides a quantitative bound on how replacing $\|\mathbf{W}\mathbf{s}\|_1 = \|\mathbf{x}\|_1$ with $\|\mathbf{s}\|_1$ affects the solution in the original norm.