

Robust Adaptive Wideband Beamforming Based on Time Frequency Distribution

Yaqi Liu , Chengcheng Liu, Dexiu Hu , and Yongjun Zhao

Abstract—Many robust adaptive wideband beamforming algorithms have been carried out to preserve the signal of interest, but their performance can be still degraded by the mismatches of steering vector and covariance matrix. Therefore, in this paper, a novel robust adaptive wideband beamformer based on the time-frequency distribution is proposed, which can achieve high output signal-to-interference-plus-noise ratio close to the optimal one even with perturbed array manifold and limited snapshots. It transforms the received signal into the time-frequency domain, and gives a method to select the single-source points of wideband signals. With the single-source points, it reconstructs the interference-plus-noise covariance matrix and estimates the steering vector without using the imprecise prior information about the array manifold, guaranteeing high estimation accuracy. Simulation results demonstrate that the proposed method outperforms other adaptive wideband beamformers, and achieves excellent output performance over a broad range of direction and sensor position errors even with small number of snapshots.

Index Terms—Covariance matrix reconstruction, steering vector estimation, time-frequency distribution, wideband robust adaptive beamforming.

I. INTRODUCTION

ADAPTIVE wideband beamforming has been widely used for enhancing the signal of interest (SOI) and suppressing interferences in many applications ranging from wireless communications to radar and sonar [1]. Most of traditional beamforming techniques can work effectively and achieve satisfactory output signal-to-interference-plus-noise ratio (SINR) with the accurate covariance matrix and steering vector (SV) of SOI [2], [3]. However, in practice, there always exist mismatches between the nominal SV and the actual one due to the direction of arrival (DOA) error, sensor position error, local scattering effect, etc [4]–[7]. The performance of beamformer can be degraded severely by these mismatches especially the DOA error and sensor position error. This is because the accuracy of SV depends strongly on the propagation delays, and the propagation delays

are determined by the DOA and sensor position. Besides, for the limited snapshots, the mismatch between the ideal covariance matrix and the sample covariance matrix (SCM) can also result in dramatic performance degradation.

To alleviate the effect of mismatches, many robust adaptive wideband beamforming techniques have been proposed, which cover the diagonal loading (DL) algorithm [8], norm-constrained algorithm [9], derivative constraint algorithm [10], [11], worst-case performance-optimization (WC-PO) algorithm [12], [13], etc. The DL algorithm achieves robustness by adding a scaled identity matrix to SCM. However, there is no explicit guideline to choose the optimal DL factor in different scenarios. The norm-constrained algorithm, which applies generalized sidelobe canceller (GSC) structure, preserves the SOI by constraining the norm of adaptive filter vector at each update. Similar to DL algorithm, the constrained norm value varying with the environment is difficult to determine. The derivative constraint algorithm, which forms a wider main beam by additional derivative constraints, can alleviate the attenuation to the SOI caused by DOA error, whereas it has poor robustness against sensor position error. Besides, the robust algorithms above share a common drawback that the output SINR deviates from the optimal one gradually with the increase of input signal-to-noise ratio (SNR), and even a very small mismatch can result in severe performance degradation in the case of high input SNR. This is because the SOI component is contaminated in the SCM, then the mismatches of SV and SCM will make the beamformer erroneously treat the SOI as an interference to be suppressed, and higher SNR leads to stronger suppression. In WC-PO algorithm, the SOI self-nulling is avoided by imposing constraints that the absolute value of the array response is not smaller than one for all SVs belonging to an uncertainty set of SVs, and the WC-PO method can be further combined with the response variation (RV) element to achieve better performance. Nevertheless, in practice, the uncertainty set is unknown. If the predefined uncertainty set is not large enough, the actual SV cannot be involved and thus the robustness will not be achieved. Conversely, if the predefined uncertainty set is too large, the ability of suppressing interferences and noise will be reduced, also leading to the degradation of output SINR.

Recently, some new robust narrowband beamforming algorithms based on the interference-plus-noise covariance matrix (IPNCM) reconstruction were proposed and achieved excellent performance [14], [15]. They reconstruct the IPNCM by integrating the Capon spectrum over an angular sector separated from the DOA of SOI. Thus, the SOI is removed from the

Manuscript received September 3, 2018; revised February 17, 2019 and June 16, 2019; accepted July 5, 2019. Date of publication July 24, 2019; date of current version August 1, 2019. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Yimin D. Zhang. This work was supported by the National Natural Science Foundation of China under Grant 61401469. (Corresponding author: Dexiu Hu.)

Y. Liu is with the Beijing Remote Sensing Institute, Beijing 100086, China (e-mail: yuyang911026@126.com).

C. Liu, D. Hu, and Y. Zhao are with the National Digital Switching System Engineering and Technological Research Center, Zhengzhou 450086, China (e-mail: luchylcc079@126.com; paper_hdx@126.com; zhaoyjzhaozeya@126.com).

Digital Object Identifier 10.1109/TSP.2019.2929924

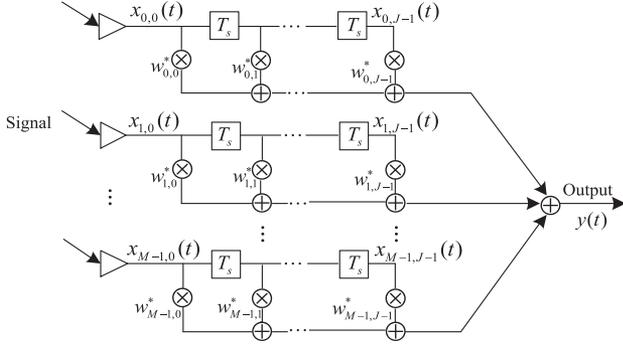


Fig. 1. A general wideband beamforming structure.

SCM and will not be suppressed even in the case of SCM and SV errors. However, most of these algorithms are only robust against DOA error. Although some modified methods have been proposed, the improvement of performance is limited [16], [17]. The IPNCM reconstruction technique can be easily extended to wideband case by integrating the Capon spectrum over both angular sector and frequency sector. Unfortunately, it cannot work effectively with sensor position error. Hence we propose a novel IPNCM reconstruction method based on time-frequency distribution (TFD), which can obtain accurate IPNCM with limited snapshots and estimate the SV of SOI precisely in the presence of DOA error and sensor position error. It can always achieve high output SINR close to the optimal one.

The remainder of this paper is organized as follows. In Section II, we formulate the problem of interest by introducing wideband array signal model and conventional wideband beamformer briefly. Then we present our proposed algorithm based on TFD in Section III, and discuss some points relative to the proposed algorithm in Section IV. In Section V, the performance of the proposed algorithm is examined by the simulation results. Conclusions are drawn in Section VI.

II. PROBLEM BACKGROUND

A. Signal Model

A general wideband beamforming structure based on tapped delay-lines (TDLs) is shown in Fig. 1, where J denotes the number of taps associated with each of the M sensors. Suppose that K wideband signals $s_k(t)$ ($k = 0, 1, \dots, K-1$) impinge on the linear array. $s_0(t)$ is the SOI, and the remaining $K-1$ signals are interferences. Then the far-field signal model can be written as

$$x_{m,j}(t) = \sum_{k=0}^{K-1} s_k(t - \tau_{k,m} - jT_s) + n_{m,j}(t) \quad (1)$$

$$m = 0, 1, \dots, M-1, j = 0, 1, \dots, J-1$$

where $x_{m,j}(t)$ is received signal of the m th sensor at the j th tap, and $\tau_{k,m} = d_m \sin(\theta_k)/c$ is the propagation delay of s_k from sensor m to sensor 0 (the reference sensor) with θ_k , c and d_m denoting the DOA of $s_k(t)$, speed of propagation, and position of sensor m , respectively. T_s is the sampling period in the attached TDLs. $n_{m,j}(t)$ is the additive noise, which is assumed to be spatially and temporally white Gaussian.

Equation (1) can be rewritten in vector form as

$$\mathbf{x}(t) = \sum_{k=0}^{K-1} \mathbf{x}_k(t) + \mathbf{n}(t) \quad (2)$$

where $\mathbf{x}(t) = [x_{0,0}(t), \dots, x_{M-1,0}(t), \dots, x_{0,J-1}(t), \dots, x_{M-1,J-1}(t)]^T$ is the received signal vector with the superscript $(\cdot)^T$ denoting the transpose operator. $\mathbf{x}_k(t) = [s_k(t - \tau_{k,0}), \dots, s_k(t - \tau_{k,M-1}), \dots, s_k(t - \tau_{k,0} - (J-1)T_s), \dots, s_k(t - \tau_{k,M-1} - (J-1)T_s)]^T$ is the received signal vector of $s_k(t)$, and $\mathbf{n}(t)$ is the additive noise vector.

The covariance matrix is defined as

$$\mathbf{R}_x = E[\mathbf{x}(t)\mathbf{x}^H(t)] \quad (3)$$

where $(\cdot)^H$ denotes the Hermitian transpose operator. Similarly, the covariance matrix of $\mathbf{x}_k(t)$ is

$$\mathbf{R}_{x_k} = E[\mathbf{x}_k(t)\mathbf{x}_k^H(t)] \quad (4)$$

Then the IPNCM is defined as

$$\mathbf{R}_{i+n} = \sum_{k=1}^{K-1} \mathbf{R}_{x_k} + \delta^2 \mathbf{I} \quad (5)$$

where δ^2 represents the noise power, and \mathbf{I} denotes the identity matrix.

B. Conventional Wideband Beamforming Algorithm

Based on the structure in Fig. 1, many adaptive beamforming algorithms have been proposed. The most well-known one is the linearly constrained minimum variance (LCMV) beamformer [3], which is formulated as

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{R}_x \mathbf{w} \quad \text{subject to} \quad \mathbf{C}^H \mathbf{w} = \mathbf{h} \quad (6)$$

where $(\cdot)^H$ denotes the Hermitian transpose operator. \mathbf{C} and \mathbf{h} represent the constraint matrix and the response vector, respectively, which are set as

$$\mathbf{C} = [\mathbf{a}_0(f_1), \mathbf{a}_0(f_2), \dots, \mathbf{a}_0(f_P)]_{MJ \times P} \quad (7)$$

$$\mathbf{h} = [e^{j2\pi f_1/2}, e^{j2\pi f_2/2}, \dots, e^{j2\pi f_P/2}]_{P \times 1}^T \quad (8)$$

in which $f_p \in [f_l, f_h]$ ($p = 1, 2, \dots, P$) with f_l and f_h denoting the low and high bound of the desired frequency band, and P is the number of the selected frequencies. For a given frequency f_p , the $MJ \times 1$ SV $\mathbf{a}_0(f_p)$ of SOI is written as

$$\mathbf{a}_0(f_p) = [e^{-j2\pi f_p \tau_{0,0}}, \dots, e^{-j2\pi f_p \tau_{0,M-1}}, \dots, e^{-j2\pi f_p (\tau_{0,0} + (J-1)T_s)}, \dots, e^{-j2\pi f_p (\tau_{0,M-1} + (J-1)T_s)}]^T \quad (9)$$

The beamformer weight vector \mathbf{w} can be easily obtained by solving (6) as

$$\mathbf{w}_{opt} = \mathbf{R}_x^{-1} \mathbf{C} (\mathbf{C}^H \mathbf{R}_x^{-1} \mathbf{C})^{-1} \mathbf{h} \quad (10)$$

Observing (10), to obtain optimal \mathbf{w}_{opt} , we need get accurate \mathbf{R}_x and \mathbf{C} (herein \mathbf{h} is already precisely known since it only depends on frequency). For \mathbf{R}_x , it is usually replaced by the SCM $\hat{\mathbf{R}}_x = \sum_{n=0}^{N-1} \mathbf{x}(n)\mathbf{x}^H(n)/N$ with N snapshots. When $N \rightarrow \infty$, $\hat{\mathbf{R}}_x$ converges to the theoretical covariance matrix \mathbf{R}_x . However, in the case of small size N , there will be a large bias between $\hat{\mathbf{R}}_x$ and \mathbf{R}_x . For \mathbf{C} , it is mainly determined by the propagation delays $\tau_{0,0}, \tau_{0,1}, \dots, \tau_{0,M-1}$ of the SOI, and $\tau_{0,0}$,

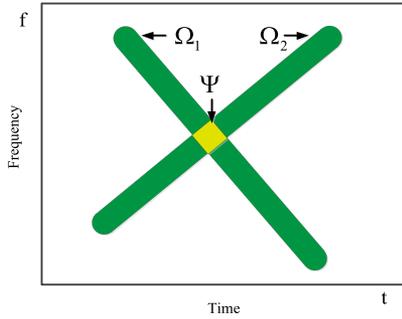


Fig. 2. TFDs of a received signal involving two source signals s_1 and s_2 .

$\tau_{0,1}, \dots, \tau_{0,M-1}$ are determined by the DOA θ_0 and sensor positions d_0, d_1, \dots, d_{M-1} . In the presence of DOA error or sensor position error, the nominal \mathbf{C} will have a bias compared with the actual one. The biases of \mathbf{R}_x and \mathbf{C} can dramatically deteriorate the performance, especially when the input SNR is high.

III. PROPOSED ALGORITHM

To avoid the undesired suppression, we propose a new adaptive beamforming algorithm based on TFD in this section. It modifies the beamformer from two aspects. Firstly, it can estimate the propagation delays more accurately even with DOA error and sensor position error. Secondly, it removes the SOI from the covariance matrix by reconstructing the IPNCM, which can preserve the SOI effectively.

A. Propagation Delays Estimation Based on TFDs

Generally, TFD can be classified as linear or quadratic TFD depending on whether it satisfies the linear superposition principle. Compared with quadratic TFD, the linear TFD is simple to implement with lower computational complexity and the cross-term can be avoided. The short-time Fourier transform (STFT) is the most widely used linear TFD. Hence STFT is chosen in this paper. Note that the proposed algorithm is also suitable for other linear TFD. Besides, it can be further applied to quadratic TFD only by a small change.

For a signal $z(t)$, its STFT is defined as [22]

$$Z(t, f) = \int_{-\infty}^{+\infty} z(\tau) g(\tau - t) e^{-j2\pi f\tau} d\tau \quad (11)$$

where $g(t)$ is a window function, and it is selected as Hamming window in this paper. Using (11), we can transform $z(t)$ into the TF domain. In the TF domain, a single-source point is defined as a TF point at which only one source signal contributes to the energy, while an overlapping point is a TF point at which there is an energy contribution of more than one source signals. Fig. 2 shows the TFDs of a received signal involving two source signals s_1 and s_2 . Ω_1 and Ω_2 represent the sets of single-source points of s_1 and s_2 , respectively, and Ψ represents the set of overlapping points.

Under assumed noise-free environment, taking STFT of (1) yields

$$X_{m,j}(t, f) = \sum_{k=0}^{K-1} S_k(t, f) e^{-j2\pi f(\tau_{k,m} + (j-1)T_s)} \quad (12)$$

$$m = 0, 1, \dots, M-1, j = 0, 1, \dots, J-1$$

where $X_{m,j}(t, f)$ and $S_k(t, f)$ denote the TFDs of $x_{m,j}(t)$ and $s_k(t)$, respectively.

Stacking (12) for the first M elements of $\mathbf{x}(t)$ (i.e. $\tilde{\mathbf{x}}(t) = [x_{0,0}(t), x_{1,0}(t), \dots, x_{M-1,0}(t)]^T$) yields

$$\tilde{\mathbf{X}}(t, f) = \sum_{k=0}^{K-1} \tilde{\mathbf{a}}_k(f) S_k(t, f) \quad (13)$$

where $\tilde{\mathbf{X}}(t, f) = [X_{0,0}(t, f), X_{1,0}(t, f), \dots, X_{M-1,0}(t, f)]^T$ and $\tilde{\mathbf{a}}_k(f) = [e^{-j2\pi f\tau_{k,0}}, \dots, e^{-j2\pi f\tau_{k,M-1}}]^T$ are both $M \times 1$ vectors, and $\tilde{\mathbf{a}}_k(f)$ denotes the SV of $s_k(t)$ at frequency f . In this paper, we call $\tilde{\mathbf{a}}_k(f)$ as no-tap SV.

At a single-source point (t_a, f_a) of the SOI $s_0(t)$, (13) can be simplified to

$$\tilde{\mathbf{X}}(t_a, f_a) = \tilde{\mathbf{a}}_0(f_a) S_0(t_a, f_a) \quad (t_a, f_a) \in \Omega_0 \quad (14)$$

in which Ω_0 denotes the set of single-source points corresponding to $s_0(t)$. Then we can calculate $\tilde{\mathbf{a}}_0(f_a)$ by

$$\tilde{\mathbf{a}}_0(f_a) = \frac{\tilde{\mathbf{X}}(t_a, f_a)}{S_0(t_a, f_a)} \quad (15)$$

Since sensor 0 is the reference sensor, we have $\tau_{0,0} = 0$, $\tilde{\mathbf{a}}_0(f_a, 1) = 1$, and $\tilde{\mathbf{X}}(t_a, f_a, 1) = S_0(t_a, f_a)$, where $\tilde{\mathbf{a}}_0(f_a, 1)$ and $\tilde{\mathbf{X}}(t_a, f_a, 1)$ denote the first element of $\tilde{\mathbf{a}}_0(f_a)$ and $\tilde{\mathbf{X}}(t_a, f_a)$, respectively. Hence (15) is rewritten as

$$\tilde{\mathbf{a}}_0(f_a) = \frac{\tilde{\mathbf{X}}(t_a, f_a)}{\tilde{\mathbf{X}}(t_a, f_a, 1)} \quad (16)$$

Finally, the delay vector $\boldsymbol{\tau}_0(t_a, f_a) = [0, \tau_{0,1}, \dots, \tau_{0,M-1}]^T$ can be obtained by

$$\boldsymbol{\tau}_0(t_a, f_a) = \frac{\text{unwrap}(\text{angle}(\tilde{\mathbf{a}}_0(f_a)))}{-2\pi f_a} \quad (17)$$

where $\text{unwrap}(\cdot)$ is used to correct the radian phase angles in the vector by adding multiples of $\pm 2\pi$ when absolute jumps between consecutive elements of the vector are greater than or equal to the default jump tolerance of π radians, and $\text{angle}(\tilde{\mathbf{a}}_0(f_a))$ returns the phase angles for each element of $\tilde{\mathbf{a}}_0(f_a)$.

For two different single-source points (t_{a1}, f_{a1}) and (t_{a2}, f_{a2}) of $s_0(t)$, the delay vectors satisfy

$$\boldsymbol{\tau}_0(t_{a1}, f_{a1}) = \boldsymbol{\tau}_0(t_{a2}, f_{a2}) = [0, \tau_{0,1}, \dots, \tau_{0,M-1}]^T \quad (18)$$

$$\forall (t_{a1}, f_{a1}), (t_{a2}, f_{a2}) \in \Omega_0$$

Equations (16) and (17) indicate that the propagation delays can be estimated using a single-source point. However, under noise environment in practice, the estimated delay vector may have some biases. Equation (18) suggests that a more accurate estimation can be obtained by averaging the delay vectors at different single-source points belonging to Ω_0 . Similarly, we can also estimate the propagation delays of interferences by the same way. Besides, we can find that the propagation delays are calculated without using any prior information of array

manifold, which guarantees high estimation accuracy even in the presence of DOA error and sensor position error.

Now a necessary task is to select the single-source points of the SOI and interferences, which can be realized by the following procedures.

1) *SOI and Interferences Separation*: Transform $\tilde{\mathbf{x}}(t)$ into the TF domain using (13). For each TF point (t, f) , calculate $\tilde{\mathbf{a}}(f)$ using (16). Then the TF point of the SOI satisfies

$$\frac{\|\tilde{\mathbf{a}}_0^H(f) \tilde{\mathbf{a}}(f)\|}{\|\tilde{\mathbf{a}}_0(f)\| \|\tilde{\mathbf{a}}(f)\|} \geq \varepsilon_1 \quad (19)$$

where $\tilde{\mathbf{a}}_0(f) = [1, e^{-j2\pi f \sin(\bar{\theta}_0) \bar{d}_1/c}, \dots, e^{-j2\pi f \sin(\bar{\theta}_0) \bar{d}_{M-1}/c}]^T$ is the pre-assumed SV of the SOI calculated with the perturbed DOA $\bar{\theta}_0$ and sensor positions $\bar{d}_1, \bar{d}_2, \dots, \bar{d}_{M-1}$. ε_1 is a threshold (typically, $0.5 < \varepsilon_1 < 1$), and $\|\cdot\|$ denotes the Euclidean norm. Based on (19), we can obtain the TF points set associated with the SOI and denote it as Ω_d . The remaining TF points are corresponding to interferences and noise and denoted as Ω_{i+n} .

2) *Noise Thresholding*: Ω_d mainly consists of the single-source points of the SOI. However, it may also involve some noise points and overlapping points. For the noise points, we can remove them as following: at each time-slice (t_p, f) of the TFDs, apply (20) for all the frequency f_q points belonging to this time-slice

$$\text{if } \frac{\|\tilde{\mathbf{X}}(t_p, f_q)\|}{\max_{(t_p, f) \in \Omega_d} \|\tilde{\mathbf{X}}(t_p, f)\|} \leq \varepsilon_2, \quad \text{remove } (t_p, f_q) \quad (20)$$

where ε_2 is a small threshold (typically, $0 < \varepsilon_2 < 0.1$), and $\max_{(t_p, f) \in \Omega_d} \|\tilde{\mathbf{X}}(t_p, f)\|$ represents the maximum matrix norm at the time-slice (t_p, f) .

3) *TF Points Clustering*: For the overlapping points, we can remove them by TF points clustering: calculate $\tau(t, f)$ using (16) and (17) for the remaining points in Ω_d . Based on (18), (t_1, f_1) and (t_2, f_2) belong to the same signal, if they satisfy

$$\frac{\|\tau(t_1, f_1) - \tau(t_2, f_2)\|}{0.5(\|\tau(t_1, f_1)\| + \|\tau(t_2, f_2)\|)} \leq \varepsilon_3 \quad (21)$$

where ε_3 is a small threshold (typically, $0 < \varepsilon_3 < 0.5$), we will discuss $\varepsilon_1, \varepsilon_2$, and ε_3 further in Section IV). Using (21), we cluster the points in Ω_d into P ($P \geq 1$) classes. Since the single-source points are always much more than the overlapping points, the largest class Ω_0 is corresponding to the set of single-source points.

Similarly, we can remove the noise points in Ω_{i+n} through noise thresholding procedure (20), and cluster the remaining points into Q ($Q \geq K - 1$) classes based on (21). The $K - 1$ largest classes Ω_k ($k = 1, 2, \dots, K - 1$) are corresponding to the sets of single-source points for $K - 1$ interferences.

Up to now, we have obtained the single-source points of the SOI and interferences. The final delay vectors are estimated as the centroid vectors of these classes, i.e.,

$$\hat{\tau}_k = \frac{1}{D_k} \sum_{(t, f) \in \Omega_k} \tau(t, f), \quad k = 0, 1, \dots, K - 1 \quad (22)$$

where D_k is the number of points in Ω_k . With $\hat{\tau}_0$, the constraint matrix $\hat{\mathbf{C}}$ can be easily obtained by (7) and (9).

B. IPNCM Reconstruction for Wideband Signal

Observing (5), to reconstruct the IPNCM, we need obtain the covariance matrix $\mathbf{R}_{\mathbf{x}_k}$ ($k = 1, 2, \dots, K - 1$) associated with the wideband interference $s_k(t)$. For $s_k(t)$, it can be regarded as multiple narrowband signals

$$s_k(t) = \sum_{f_i \in [f_l, f_h]} s_{k, f_i}(t) \quad (23)$$

where $s_{k, f_i}(t)$ represents the narrowband component of $s_k(t)$ at f_i . Then the received signal vector of $s_k(t)$ can be rewritten as

$$\mathbf{x}_k(t) = \sum_{f_i \in [f_l, f_h]} \mathbf{a}_k(f_i) s_{k, f_i}(t) \quad (24)$$

in which $\mathbf{a}_k(f_i)$ is the SV of $s_k(t)$ at f_i and can be written as

$$\mathbf{a}_k(f_i) = \mathbf{g}(f_i) \otimes \tilde{\mathbf{a}}_k(f_i) \quad (25)$$

where \otimes denotes the Kronecker product, $\tilde{\mathbf{a}}_k(f_i)$ is the no-tap SV at f_i , $\mathbf{g}(f_i)$ is a $J \times 1$ vector defined as

$$\mathbf{g}(f_i) = [1, e^{-j2\pi f_i T_s}, \dots, e^{-j2\pi f_i T_s (J-1)}]_{J \times 1}^T \quad (26)$$

which depends on the frequency and can be easily obtained.

Substituting (24) into (4) yields

$$\mathbf{R}_{\mathbf{x}_k} = \sum_{f_i \in [f_l, f_h]} \sum_{f_j \in [f_l, f_h]} \mathbf{a}_k(f_i) \mathbf{a}_k^H(f_j) E[s_{k, f_i}(t) s_{k, f_j}^*(t)] \quad (27)$$

Since $E[s_{k, f_i}(t) s_{k, f_j}^*(t)] = 0$ ($i \neq j$), (27) can be rewritten as

$$\begin{aligned} \mathbf{R}_{\mathbf{x}_k} &= \sum_{f_i \in [f_l, f_h]} \mathbf{a}_k(f_i) \mathbf{a}_k^H(f_i) E[s_{k, f_i}(t) s_{k, f_i}^*(t)] \\ &= \sum_{f_i \in [f_l, f_h]} \mathbf{a}_k(f_i) \mathbf{a}_k^H(f_i) \delta_{k, i}^2 \\ &= \sum_{f_i \in [f_l, f_h]} \mathbf{R}_{\mathbf{x}_k, i} \end{aligned} \quad (28)$$

where $\delta_{k, i}^2 = E[s_{k, f_i}(t) s_{k, f_i}^*(t)]$ is the power of $s_{k, f_i}(t)$, and $\mathbf{R}_{\mathbf{x}_k, i} = \mathbf{a}_k(f_i) \mathbf{a}_k^H(f_i) \delta_{k, i}^2$ is the covariance matrix component of $\mathbf{R}_{\mathbf{x}_k}$ at f_i . Equation (28) indicates that if we obtain the covariance matrix components at different frequencies, $\mathbf{R}_{\mathbf{x}_k}$ can be reconstructed easily. Herein we propose two methods to estimate the covariance matrix components at different frequencies.

1) *The First One*: As introduced above, we can estimate the propagation delays using (22). With the estimated $\hat{\tau}_k$, the no-tap SV $\tilde{\mathbf{a}}_k(f_i)$ is calculated by

$$\tilde{\mathbf{a}}_k(f_i) = [1, e^{j2\pi f_i \hat{\tau}_{k,1}}, \dots, e^{j2\pi f_i \hat{\tau}_{k, M-1}}]_{J \times 1}^T \quad (29)$$

Substituting (29) into (25), we can obtain $\mathbf{a}_k(f_i)$. Furthermore, $\delta_{k, i}^2$ is estimated by Capon spatial spectrum estimator as

$$\delta_{k, i}^2 = \frac{1}{\mathbf{a}_k^H(f_i) \mathbf{R}_{\mathbf{x}}^{-1} \mathbf{a}_k(f_i)} \quad (30)$$

Thus, the covariance matrix component can be easily obtained as

$$\mathbf{R}_{\mathbf{x}_k, i} = \frac{\mathbf{a}_k(f_i) \mathbf{a}_k^H(f_i)}{\mathbf{a}_k^H(f_i) \mathbf{R}_{\mathbf{x}}^{-1} \mathbf{a}_k(f_i)} \quad (31)$$

Finally, $\mathbf{R}_{\mathbf{x}_k}$ is reconstructed as

$$\hat{\mathbf{R}}_{\mathbf{x}_k} = \sum_{f_i \in [f_l, f_h]} \frac{\mathbf{a}_k(f_i) \mathbf{a}_k^H(f_i)}{\mathbf{a}_k^H(f_i) \mathbf{R}_{\mathbf{x}}^{-1} \mathbf{a}_k(f_i)} \quad (32)$$

2) *The Second One:* Based on (25), $\mathbf{R}_{\mathbf{x}_{k,i}}$ can be rewritten as

$$\begin{aligned} \mathbf{R}_{\mathbf{x}_{k,i}} &= \mathbf{a}_k(f_i) \mathbf{a}_k^H(f_i) \delta_{k,i}^2 \\ &= [\mathbf{g}(f_i) \otimes \tilde{\mathbf{a}}_k(f_i)] [\mathbf{g}(f_i) \otimes \tilde{\mathbf{a}}_k(f_i)]^H \delta_{k,i}^2 \\ &= [\mathbf{g}(f_i) \mathbf{g}^H(f_i)] \otimes [\tilde{\mathbf{a}}_k(f_i) \tilde{\mathbf{a}}_k^H(f_i)] \delta_{k,i}^2 \\ &= [\mathbf{g}(f_i) \mathbf{g}^H(f_i)] \otimes \tilde{\mathbf{R}}_{\mathbf{x}_{k,i}} \end{aligned} \quad (33)$$

in which $\tilde{\mathbf{R}}_{\mathbf{x}_{k,i}} = \tilde{\mathbf{a}}_k(f_i) \tilde{\mathbf{a}}_k^H(f_i) \delta_{k,i}^2$ represents the no-delay covariance matrix component at f_i . If $\tilde{\mathbf{R}}_{\mathbf{x}_{k,i}}$ is known, $\mathbf{R}_{\mathbf{x}_{k,i}}$ can be obtained. Hence we develop a method to obtain $\tilde{\mathbf{R}}_{\mathbf{x}_{k,i}}$.

Based on (14), at a single-source point (t, f) of $s_k(t)$, we have

$$\begin{aligned} \tilde{\mathbf{X}}(t, f) \tilde{\mathbf{X}}^H(t, f) &= \tilde{\mathbf{a}}_k(f) \tilde{\mathbf{a}}_k^H(f) S_k(t, f) S_k^*(t, f) \quad (t, f) \in \Omega_k \\ &= \tilde{\mathbf{a}}_k(f) \tilde{\mathbf{a}}_k^H(f) \rho_k(t, f) \end{aligned} \quad (34)$$

where $\rho_k(t, f) = S_k(t, f) S_k^*(t, f)$ is power of $s_k(t)$ at (t, f) .

Summing (34) for all the TF points at frequency-slice (t, f_i) in Ω_k yields

$$\begin{aligned} &\sum_{(t,f) \in \Omega_k, f=f_i} \tilde{\mathbf{X}}(t, f) \tilde{\mathbf{X}}^H(t, f) \\ &= \sum_{(t,f) \in \Omega_k, f=f_i} \tilde{\mathbf{a}}_k(f) \tilde{\mathbf{a}}_k^H(f) \rho_k(t, f) \\ &= \tilde{\mathbf{a}}_k(f_i) \tilde{\mathbf{a}}_k^H(f_i) \sum_{(t,f) \in \Omega_k, f=f_i} \rho_k(t, f) \\ &= \tilde{\mathbf{a}}_k(f_i) \tilde{\mathbf{a}}_k^H(f_i) E_{k,f_i} \end{aligned} \quad (35)$$

in which $E_{k,f_i} = \sum_{(t,f) \in \Omega_k, f=f_i} \rho_k(t, f)$ is the energy of $s_{k,f_i}(t)$ in TF domain. Denote the energy of $s_{k,f_i}(t)$ in time domain as $E_{k,i}$, based on Parseval's Theorem [18], $E_{k,i}$ satisfies

$$E_{k,i} = \frac{1}{N_f} E_{k,f_i} \quad (36)$$

where N_f is the number of frequencies at one time slice. Besides, we have $E_{k,i} = N \delta_{k,i}^2$ with N representing the number of snapshots. Then

$$E_{k,f_i} = N_f N \delta_{k,i}^2 \quad (37)$$

Substituting (37) into (35) yields

$$\begin{aligned} \sum_{(t,f) \in \Omega_k, f=f_i} \tilde{\mathbf{X}}(t, f) \tilde{\mathbf{X}}^H(t, f) &= \tilde{\mathbf{a}}_k(f_i) \tilde{\mathbf{a}}_k^H(f_i) N_f N \delta_{k,i}^2 \\ &= N_f N \tilde{\mathbf{R}}_{\mathbf{x}_{k,i}} \end{aligned} \quad (38)$$

Hence $\tilde{\mathbf{R}}_{\mathbf{x}_{k,i}}$ can be calculated by

$$\tilde{\mathbf{R}}_{\mathbf{x}_{k,i}} = \frac{1}{N N_f} \sum_{(t,f) \in \Omega_k, f=f_i} \tilde{\mathbf{X}}(t, f) \tilde{\mathbf{X}}^H(t, f) \quad (39)$$

Substituting (39) into (33), we obtain the covariance matrix component $\mathbf{R}_{\mathbf{x}_{k,i}}$ as

$$\begin{aligned} \mathbf{R}_{\mathbf{x}_{k,i}} &= [\mathbf{g}(f_i) \mathbf{g}^H(f_i)] \otimes \left[\frac{1}{N N_f} \sum_{(t,f) \in \Omega_k, f=f_i} \tilde{\mathbf{X}}(t, f) \tilde{\mathbf{X}}^H(t, f) \right] \\ &= \frac{1}{N N_f} \sum_{(t,f) \in \Omega_k, f=f_i} [\mathbf{g}(f) \mathbf{g}^H(f)] \otimes [\tilde{\mathbf{X}}(t, f) \tilde{\mathbf{X}}^H(t, f)] \end{aligned} \quad (40)$$

Then the covariance matrix $\mathbf{R}_{\mathbf{x}_k}$ can be reconstructed by summing (40) for all frequencies $f_i \in [f_l, f_h]$, i.e.,

$$\begin{aligned} \hat{\mathbf{R}}_{\mathbf{x}_k} &= \frac{1}{N N_f} \sum_{f_i \in [f_l, f_h]} \sum_{(t,f) \in \Omega_k, f=f_i} [\mathbf{g}(f) \mathbf{g}^H(f)] \\ &\quad \otimes [\tilde{\mathbf{X}}(t, f) \tilde{\mathbf{X}}^H(t, f)] \\ &= \frac{1}{N N_f} \sum_{(t,f) \in \Omega_k} [\mathbf{g}(f) \mathbf{g}^H(f)] \otimes [\tilde{\mathbf{X}}(t, f) \tilde{\mathbf{X}}^H(t, f)] \end{aligned} \quad (41)$$

Generally, both of the two methods can reconstruct the covariance matrix effectively. However, they also have some drawbacks. The first method reconstructs $\tilde{\mathbf{a}}_k(f)$ using the propagation delays, hence it can collect the spatial information at any frequency. But its performance depends on the estimation accuracy of propagation delays. By contrast, the second method collect the spatial information using $\tilde{\mathbf{X}}(t, f)$. The accuracy of propagation delays has no influence on its performance. But it is sensitive to overlapping points, since it cannot collect the spatial information at these points. To improve the performance, we combine the two methods. The final covariance matrix of \mathbf{x}_k is reconstructed by averaging the two covariance matrixes estimated by (32) and (41) as

$$\begin{aligned} \hat{\mathbf{R}}_{\mathbf{x}_k} &= \frac{1}{2} \left\{ \sum_{f \in [f_l, f_h]} \frac{\mathbf{a}_k(f) \mathbf{a}_k^H(f)}{\mathbf{a}_k^H(f) \mathbf{R}_{\mathbf{x}}^{-1} \mathbf{a}_k(f)} \right. \\ &\quad \left. + \frac{1}{N N_f} \sum_{(t,f) \in \Omega_k} [\mathbf{g}(f) \mathbf{g}^H(f)] \otimes [\tilde{\mathbf{X}}(t, f) \tilde{\mathbf{X}}^H(t, f)] \right\} \end{aligned} \quad (42)$$

Then the IPNCM is reconstructed as

$$\hat{\mathbf{R}}_{i+n} = \sum_{k=1}^{K-1} \hat{\mathbf{R}}_{\mathbf{x}_k} + \hat{\sigma}_n^2 \mathbf{I} \quad (43)$$

where the noise power $\hat{\sigma}_n^2$ is approximately estimated by the smallest eigenvalue of $\hat{\mathbf{R}}_{\mathbf{x}}$.

Up to now, we have obtained both $\hat{\mathbf{R}}_{i+n}$ and $\hat{\mathbf{C}}$. Substituting them back into (10), the final beamformer weight vector is

$$\mathbf{w} = \hat{\mathbf{R}}_{i+n}^{-1} \hat{\mathbf{C}} \left(\hat{\mathbf{C}}^H \hat{\mathbf{R}}_{i+n}^{-1} \hat{\mathbf{C}} \right)^{-1} \mathbf{h} \quad (44)$$

The proposed algorithm is summarized in Table I.

TABLE I
THE PROPOSED ALGORITHM

-
- 1) Transform the no-tap received signal vector $\mathbf{x}(t) = [x_{0,0}(t), x_{1,0}(t), \dots, x_{M-1,0}(t)]^T$ into TF domain by STFT;
 - 2) Obtain Ω_d and Ω_{i+n} using (19), and remove the noise points in Ω_d and Ω_{i+n} using (20);
 - 3) Based on (21), cluster the remaining TF points in Ω_d and Ω_{i+n} into P classes and Q classes, respectively. Select Ω_0 corresponding to the SOI and Ω_k ($k = 1, 2, \dots, K-1$) corresponding to interferences, respectively;
 - 4) Calculate the delay vectors using (22), and reconstruct the IPCM using (42) and (43);
 - 5) Obtain the final beamformer weight vector using (44).
-

IV. DISCUSSION

In this section, we will discuss some points relative to the proposed algorithm.

1) *Parameters $\varepsilon_1, \varepsilon_2$, and ε_3 setting:* For ε_1 , it is used to select the TF points of the SOI. In this paper, we do not discuss the case of close DOAs of SOI and interference. Hence the value of ε_1 mainly depends on the potential maximum bias between the pre-assumed SV and the actual one. Denote the potential maximum bias as $\Delta \mathbf{a}_0(f_h)$, then ε_1 can be set to $\|(\Delta \mathbf{a}_0(f_h) + \mathbf{a}_0(f_h))^H \mathbf{a}_0(f_h)\| / (\|\Delta \mathbf{a}_0(f_h) + \mathbf{a}_0(f_h)\| \|\mathbf{a}_0(f_h)\|)$. For ε_2 , it is used to remove the noise points. However, big ε_2 will also filter out some TF points of the source signals. Hence ε_2 is generally set to a small value, such as $\varepsilon_2 = 0.05$ in this paper. With small ε_2 , some noise points may not be removed. Nevertheless, it has little effect on the final performance, since the remaining noise points will be removed by the *TF points clustering* procedure due to the fact that noise points can hardly be clustered into the classes of source signals. For ε_3 , it is used to cluster the TF points, and its value can be influenced by the SNR and interference-to-noise ratio (INR) at the single-source point. Generally, higher SNR and INR lead to smaller ε_3 . For example, under noise-free environment, we can set $\varepsilon_3 = 0$. In fact, the SNR and INR at the single-source point are always high, since the noise spans its energy over the whole TF domain while the sources mainly concentrate their energy at the single-source points. Thus, $\varepsilon_3 = 0.1$ is selected in this paper. Note that if the SNR and INR at the single-source point are low, it means the energy of source signal is small and has limited influence on the output performance.

2) *Assumption:* The proposed algorithm is based on the assumption that the SOI and interferences do not overlap completely in the TF domain, which means that the SOI and interferences can overlap partly in the TF domain, but they cannot have the same TFDs. This assumption can be easily met in practice

[19]–[26]. If two array signals have the same TFDs, they must share the same frequency components, time of arrival (TOA) and DOA. This condition is extremely restrictive in practice. Even one signal and its multipath signals will have different TFDs for different TOAs and DOAs.

3) *Advantages as compared to TF-based Blind source separation (BSS) methods:* Firstly, the proposed algorithm can be applied to both narrowband signal and wideband signal, while the BSS methods are limited to the narrowband case where the SV is constant [19]–[24]. The BSS methods recover the source signals by TF synthesis or the inverse of mixing matrix. In the wideband case, the mixing matrix varies with frequency, and cannot be obtained by these BSS methods. Similarly, they cannot obtain the complete TFDs of each source, either. This is because the TF points clustering uses $\mathbf{a}_k(f)$ as the feature, but $\mathbf{a}_k(f)$ is not constant, making the clustering fail. Without the mixing matrix or complete TFDs, they cannot achieve good performance of source separation. Although some BSS approaches were carried out for convolutive mixing case [25], [26], they are still unsuitable for wideband signal due to the signal model mismatch. Secondly, even though the mixing matrices at different frequencies are obtained by our proposed method, we may not separate the sources by inverse of mixing matrix, since the separation procedure must be repeated at every frequency, leading to huge computation complexity. There also exists the problem of phase discontinuity, when different narrowband components are combined.

In [27], a method that uses the combined information of spatial and time-frequency signatures is proposed, and it can achieve good performance. However, in the TF-joint case, it is impossible to obtain the complete TF signature of each source, resulting in performance degradation. Besides, the method is based on the assumption that the frequency is constant in one symbol duration. Nevertheless, the frequencies are different from symbol to symbol. Thus, to obtain the complete output signal, the source recovery procedures in [27] must be repeated for every symbol, leading to large computation complexity. In comparison, the proposed method can recover the SOI for one time. Moreover, it estimates the SV and reconstructs the IPNCM only using a part of TFDs. Hence the fragmentary TFDs have little effect on the performance.

4) *Extension to quadratic TFD:* For quadratic TFD, the spatial time-frequency distribution (STFD) matrix of the received signal vector $\tilde{\mathbf{x}}(t)$ is defined as

$$\mathbf{D}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(t, f) = \sum_{i=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} \tilde{\mathbf{x}}(t+i+l) \tilde{\mathbf{x}}^H(t+i-l) \phi(i, l) e^{-j4\pi fl} \quad (45)$$

where $\phi(i, l)$ is the Doppler-lag kernel. Under assumed noise-free environment, the wideband $\tilde{\mathbf{x}}(t)$ can be divided into multiple narrowband components as

$$\tilde{\mathbf{x}}(t) = \tilde{\mathbf{x}}_{f_{as}}(t) + \sum_{f_i \in [f_l, f_h], f_i \neq f_{as}} \tilde{\mathbf{x}}_{f_i}(t) \quad (46)$$

in which $\tilde{\mathbf{x}}_{f_{as}}(t) = \sum_{k=0}^{K-1} \tilde{\mathbf{a}}_k(f_{as}) s_{k,f_{as}}(t)$ is the narrowband component of $\tilde{\mathbf{x}}(t)$ at frequency f_{as} , and $\sum_{f_i \in [f_l, f_h], f_i \neq f_{as}} \tilde{\mathbf{x}}_{f_i}(t)$ represents other narrowband components whose frequencies differ from f_{as} . The single-source auto-source point is defined as a TF point at which only one source contributes to the energy. Thus, at a single-source auto-source point (t_{as}, f_{as}) of $s_k(t)$, substituting (46) to (45) yields

$$\mathbf{D}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(t_{as}, f_{as}) = \rho_{s_k s_k}(t_{as}, f_{as}) \tilde{\mathbf{a}}_k(f_{as}) \tilde{\mathbf{a}}_k^H(f_{as}) \quad (47)$$

which $\rho_{s_k s_k}(t_{as}, f_{as})$ denotes the quadratic TFD of $s_{k,f_{as}}(t)$ at (t_{as}, f_{as}) .

Then the SV $\tilde{\mathbf{a}}_k(f_{as})$ of $s_k(t)$ at frequency f_{as} can be obtained by

$$\tilde{\mathbf{a}}_k(f_{as}) = \frac{\mathbf{u}_P(t_{as}, f_{as})}{\mathbf{u}_P(t_{as}, f_{as}, 1)} \quad (48)$$

where $\mathbf{u}_P(t_{as}, f_{as})$ is the principal eigenvector of $\mathbf{D}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(t_{as}, f_{as})$, and $\mathbf{u}_P(t_{as}, f_{as}, 1)$ denotes the first element of $\mathbf{u}_P(t_{as}, f_{as})$.

The observation (48) indicates that the SV can be obtained using the STFD matrix at the single-source auto-source point. Besides, since the STFD matrix in (47) is a rank-1 matrix under noise-free environment, the single-source auto-source points can be selected by

$$\text{if } \frac{|\lambda_{\max}(t, f)|}{\sum_{m=0}^{M-1} |\lambda_m(t, f)|} \geq \varepsilon_4, \quad \text{keep}(t, f) \quad (49)$$

where $\lambda_0(t, f), \lambda_1(t, f), \dots, \lambda_{M-1}(t, f)$ are the eigenvalues of $\mathbf{D}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}(t, f)$, and $\lambda_{\max}(t, f)$ denotes the maximum one. ε_4 is a threshold (typically, $0.8 < \varepsilon_4 < 1$).

Finally, we can estimate the delay vector by (17) and reconstruct the IPNCM by (42) and (43).

V. SIMULATION RESULTS

In this part, we first show the simulation results of the proposed method in different cases, then evaluate its performance by comparing it with other algorithms and the optimal algorithm which calculates the weighted vector using the accurate IPNCM \mathbf{R}_{i+n} and SV $\mathbf{a}_0(f)$ of SOI, i.e.,

$$\mathbf{w}_{opt} = \mathbf{R}_{i+n}^{-1} \mathbf{C} (\mathbf{C}^H \mathbf{R}_{i+n}^{-1} \mathbf{C})^{-1} \mathbf{h} \quad (50)$$

where both $\mathbf{C} = [\mathbf{a}_0(f_1), \mathbf{a}_0(f_2), \dots, \mathbf{a}_0(f_P)]$ and \mathbf{R}_{i+n} have no biases.

A. Simulation of the Proposed Beamformer in A General Case

Assuming that three wideband signals s_0, s_1 and s_2 impinge on an array of $M = 6$ sensors, and each sensor is attached with $J = 20$ delay taps. The SOI s_0 from $\theta_0 = 10^\circ$ is a sinusoidal frequency modulated signal with SNR = 0 dB. The interferences s_1 and s_2 from $\theta_1 = -20^\circ$ and $\theta_2 = 35^\circ$ are linear chirp signal and frequency hopping signal, respectively, with the same INR equal to 20 dB. Their frequency ranges are [2.4, 3.6] GHz, [2.6, 3.8] GHz, and [2.5, 3.5] GHz, sequentially. Each sensor has 512 snapshots with the sample frequency $f_s = 2$ GHz. The nominal array is assumed to be a uniform linear array with the inter-sensor spacing $\Delta d = 0.5\lambda_{\min}$ corresponding to the maximum

frequency $f_{\max} = 3.8$ GHz. However, for the sensor position error, the actual inter-sensor spacing $\Delta d_{j-1,j}$ ($j = 1, 2, \dots, 5$) is a random value falling in $[\Delta d - \varepsilon_p \Delta d, \Delta d + \varepsilon_p \Delta d]$, herein ε_p is set to 0.2. Besides, the nominal DOA $\bar{\theta}_0$ is 7° , which has an estimation error of 3° .

In Fig. 3, the top three plots show the TFDs of the source signals s_0, s_1 and s_2 , respectively. It can be seen that they will have overlapping blocks in the TF domain. Fig. 3(d) gives the TFDs of the received signal, which contain the single-source points, overlapping points, and many noise points. After SOI and interferences separation, we can get the TF points associated with the SOI shown in Fig. 3(e), and the remaining TF points shown in Fig. 3(f) are corresponding to the interferences and noise. Finally, the single-source points of s_0, s_1 and s_2 are obtained by removing the noise points and clustering the residual points, which are shown in Fig. 3(g)–(i), sequentially. Compared with Fig. 3(a)–(c), the distributions of the single-source points are fragmentary, since the overlapping points are filtered out. However, its effect can be ignored due to the fact that the propagation delays estimation and IPNCM reconstruction only need some single-source points of the SOI and interferences but not all. This can be verified by Fig. 4(a), where the biases between the estimated delays and the actual ones are very slight. The beam pattern is given in Fig. 4(b), where nulls are formed at the DOAs of interferences and distortionless response is maintained at the DOA of the SOI, indicating that the beamformer can enhance the SOI and suppress the interferences effectively.

B. Sources With Low Energy

If different sources have different energies, the source with low energy might be confused with noise. Firstly, we discuss the SOI with low energy. The SNR of s_0 is set to -20 dB, and other conditions are the same as *Simulation A*. Fig. 5(a) gives the TFDs of received signal. In Fig. 5(b), the proposed method fails to select the single-source points of SOI. Hence the estimated delay vector has big bias shown in Fig. 5(c). However, the beamformer can still achieve satisfactory performance, due to the fact that the SOI has been removed from covariance matrix and no null will be formed at the direction of SOI as shown in Fig. 7(a). We can see that the beamformer only suffers 0.4385 dB performance degradation.

Then we discuss the interference with low energy. The INR of s_2 is set to -20 dB, and other conditions are the same as *Simulation A*. The TFDs of received signal and single-source points of s_2 are given in Fig. 6(a) and Fig. 6(b), respectively. It can be seen that the single-source points cannot be selected, and the estimated delay vector of s_2 has big bias shown in Fig. 6(c). The resulting beam pattern does not form null at the direction of s_2 shown in Fig. 7(b). However, since the energy of s_2 is too small to influence the performance, the beamformer can still achieve high SINR. To indicate this, the output SINR versus INR of s_2 is given in Fig. 7(c), where high SINR is achieved over the whole range of INR. Overall, the interference with high energy has big influence on the output performance, but its delay vector

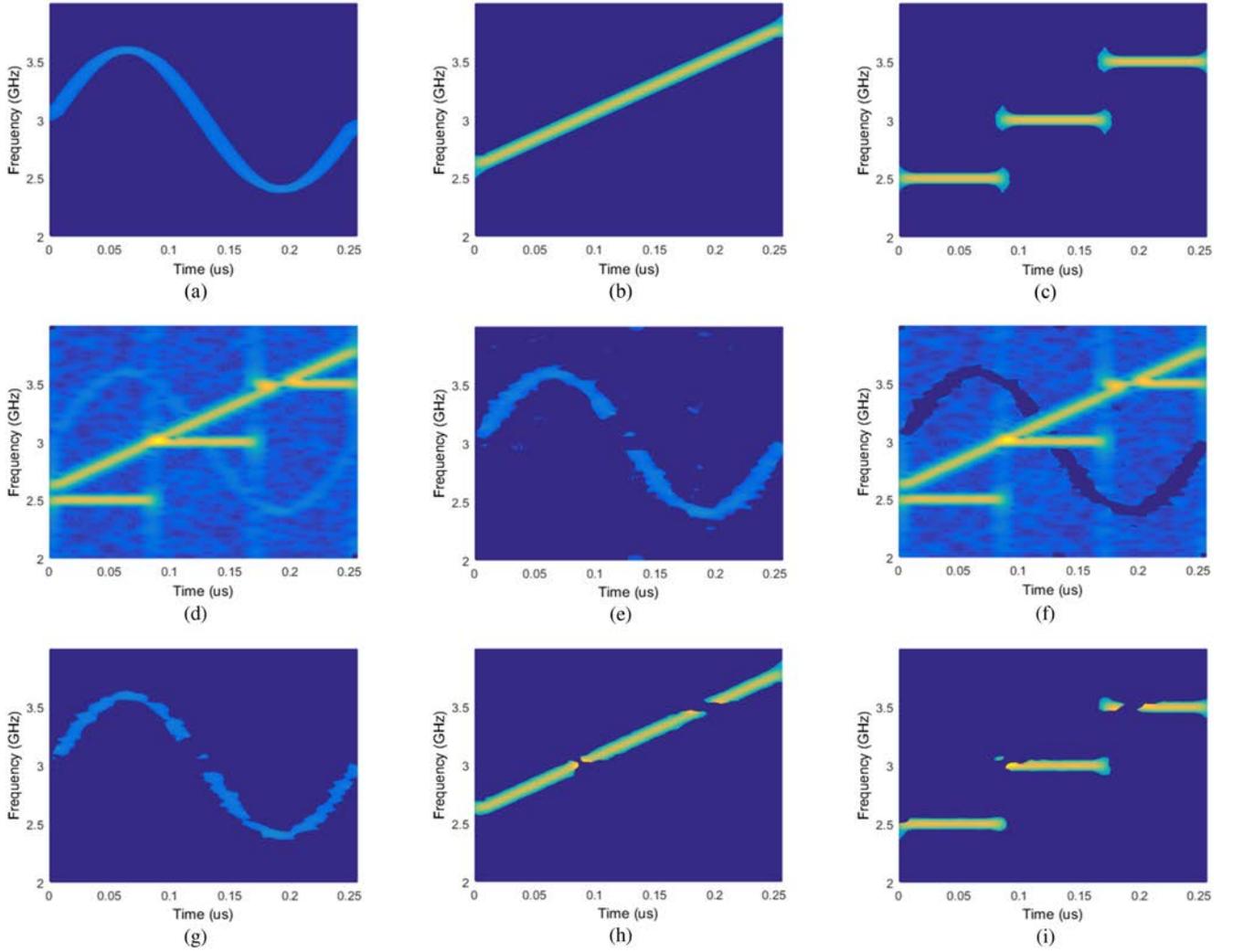


Fig. 3. TFDs of the source signals (a)–(c) and received signal (d), TF points of the SOI (e) and interferences-plus-noise (f), single-source points of the SOI and interferences (g)–(h).

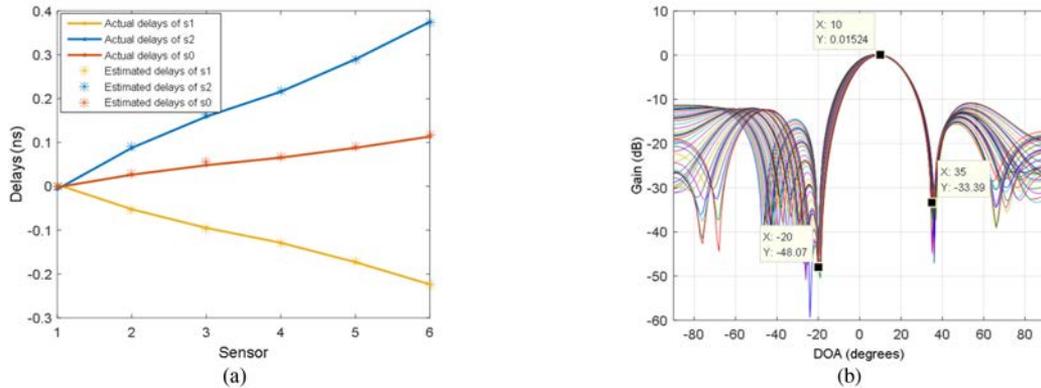


Fig. 4. Output results. (a) comparison between the estimated delays and the actual ones; (b) the beam pattern.

can be estimated accurately. The interference with low energy cannot obtain accurate delay vector, but it has limited effect on the output performance. Thus, the proposed beamformer can always achieve good performance.

C. Close DOAs of Interferences

In this section, we discuss the effect of close DOAs of interferences. The DOAs of interferences s_1 and s_2 are set to -30° and -32° , and other conditions are the same as *Simulation A*.

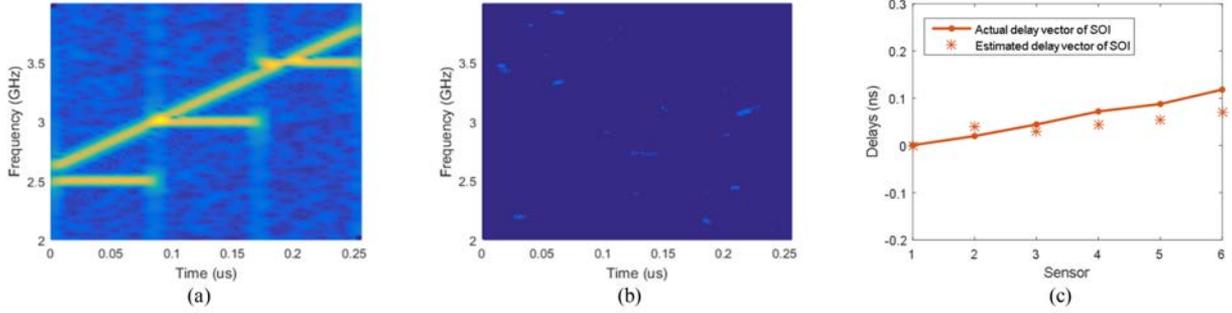


Fig. 5. In the case of $\text{SNR} = -20$ dB, (a), (b), and (c) represent the TFDs of the received signal, single-source points of the SOI, and comparison between the estimated delay and the actual one, respectively.

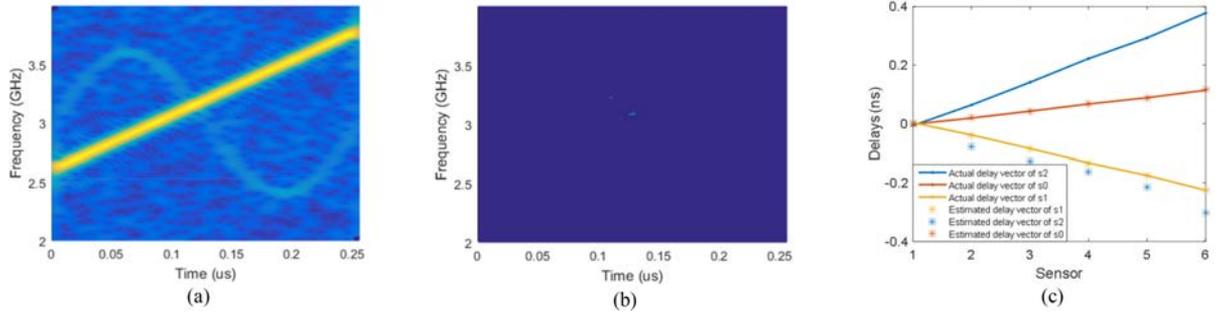


Fig. 6. In the case of $\text{INR} = -20$ dB, (a), (b), and (c) represent the TFDs of the received signal, single-source points of s_2 , and comparison between the estimated delay and the actual one, respectively.

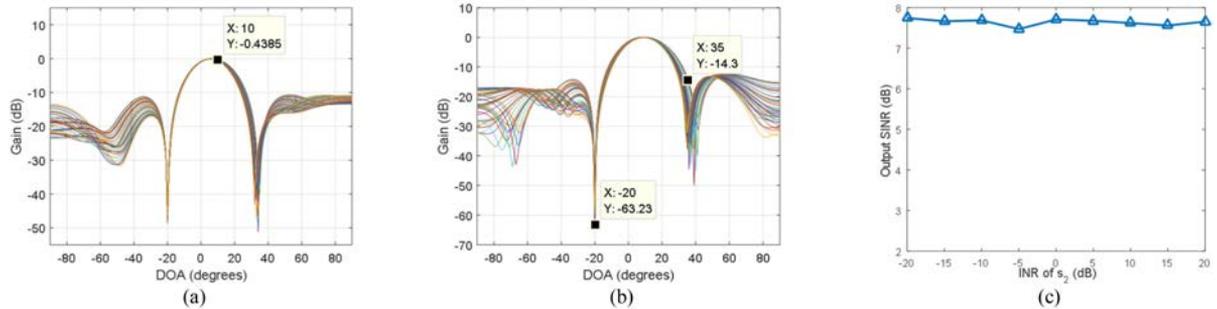


Fig. 7. Output results. (a) the beam pattern in the case of $\text{SNR} = -20$ dB, (b) the beam pattern in the case of $\text{INR} = -20$ dB, (c) output SINR versus the INR of s_2 .

Due to the close DOAs, the single-source points of s_1 and s_2 are clustered into the same class, as shown in Fig. 8(a). Thus, the estimated delay vector $\hat{\tau}$ is between the actual delay vectors τ_1 and τ_2 shown in Fig. 8(b), making the beamformer form null at $\hat{\theta}(-32^\circ < \hat{\theta} < -30^\circ)$. However, since the DOAs of interferences are very close to the null, the interferences can be also suppressed to some degree. More importantly, we reconstruct the IPNCM using two methods, among which the second method (41) reconstructs the IPNCM by the TFDs of received signal vector at single source points of interferences. Thus, the accuracy of delay vector has little influence on its performance, and it can collect the complete information of s_1 and s_2 , despite their single-source points are clustered into one class or two classes. Fig. 8(c) gives the final beam pattern, from which we can see the beamformer suppresses the interferences effectively.

D. Large Overlapping Blocks

In the case of large overlapping blocks, the overlapping points may be more than the single-source points. Next we discuss its effect on the performance. s_2 is set to a new signal shown in Fig. 9(a), and other conditions are the same as *Simulation A*. Fig. 9(b) shows the TFDs of received signal, where the overlapping block contains more points. However, these overlapping points will be clustered into many classes, since the estimated delay vectors at overlapping points vary with frequencies, while the estimated delay vectors at single-source points keep constant for the same signal. After clustering, the two biggest classes are shown in Fig. 9(c) and Fig. 9(d). Obviously, they are corresponding to the set of single-source points, which indicates that the overlapping points are not clustered into the same class. With the selected single-source points, the delay vectors are

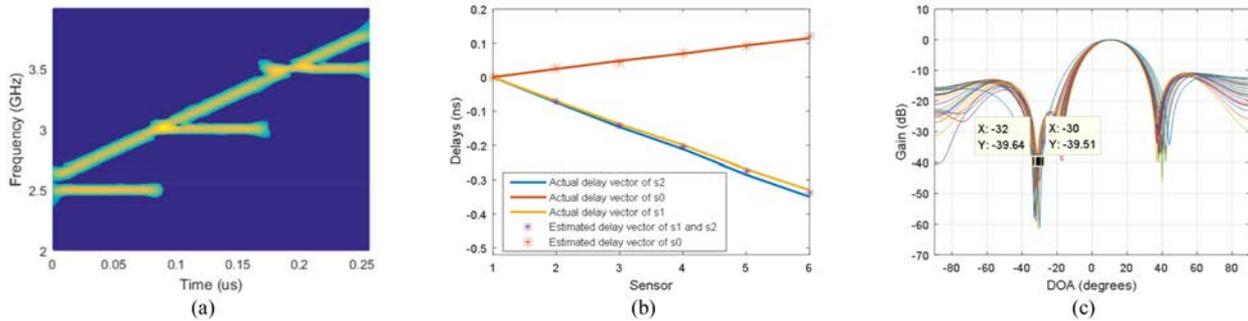


Fig. 8. In the case of close DOAs. (a), (b), and (c) represent the biggest single-source points class after clustering, comparison between the estimated delays and the actual ones, and the beam pattern, respectively.

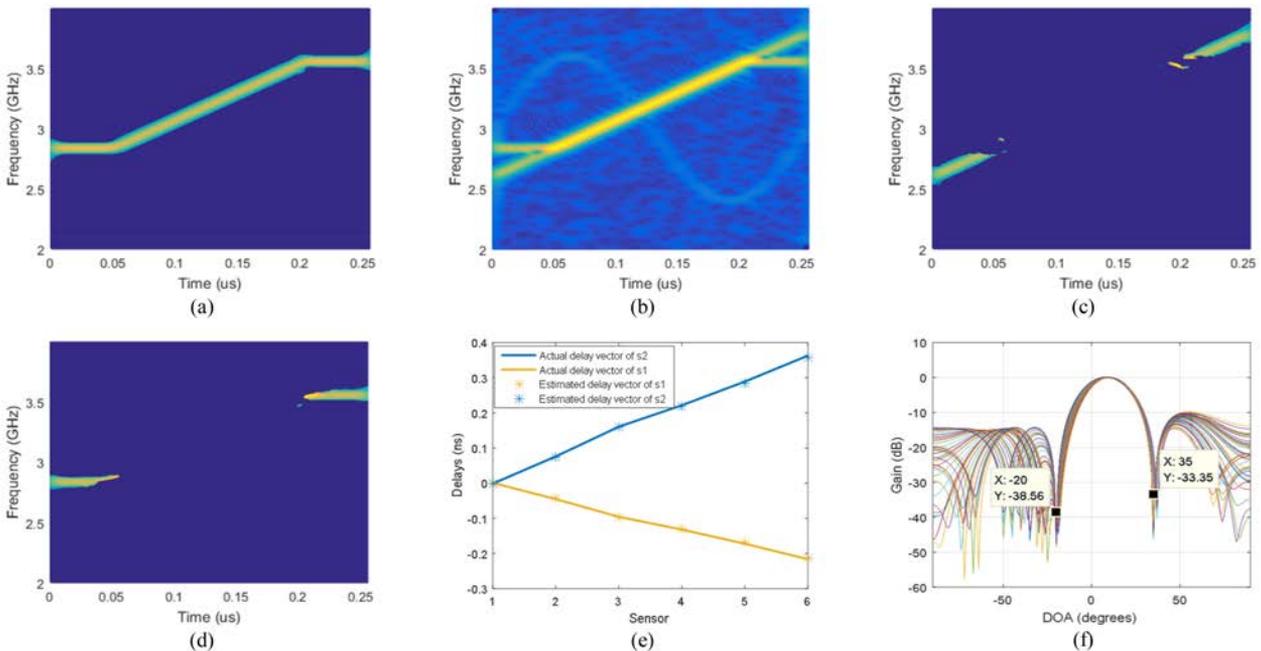


Fig. 9. In the case of large overlapping blocks. (a) and (b), and (c) represent the TFDs of s_2 and received signal, respectively. (c) and (d) are the two biggest single-source points classes after clustering. (e) and (f) denote the comparison between the estimated delays and the actual ones, and the beam pattern, respectively.

estimated accurately shown in Fig. 9(e), and good beamforming performance is achieved shown in Fig. 9(f).

E. Combination of the Two IPNCM Reconstruction Methods

In the paper, we give two IPNCM reconstruction methods. Generally, they can both reconstruct the IPNCM effectively. However, in some special cases, only applying one method may lead to performance degradation. Thus, we combine the two methods. Table II shows their output SINRs in different cases. In the general case of Fig. 3, the two method both achieve high output SINRs. Their combined method also obtains good performance. In the case of large overlapping blocks in Fig. 9, the second method cannot collect complete information of interferences, resulting in severe performance degradation. In contrast, the first method achieves good performance, as the delay vector can be estimated accurately. For the combined method, it fuses both advantage and disadvantage of the two

TABLE II
PERFORMANCE OF DIFFERENT IPNCM RECONSTRUCTION METHODS

Simulation scenarios	Output SINR (dB)		
	First method	Second method	Combined method
General case	7.6477	7.3720	7.5892
Large overlapping blocks	7.3510	-2.1295	7.1189
Close DOAs	6.0855	7.3104	7.2563

methods. The disadvantage of the second method is that its collected information is not complete. However, it does not contain wrong information. Besides, the better one (i.e., the first method) always dominates in the combined method. Hence the disadvantage has limited effect on the performance, and the

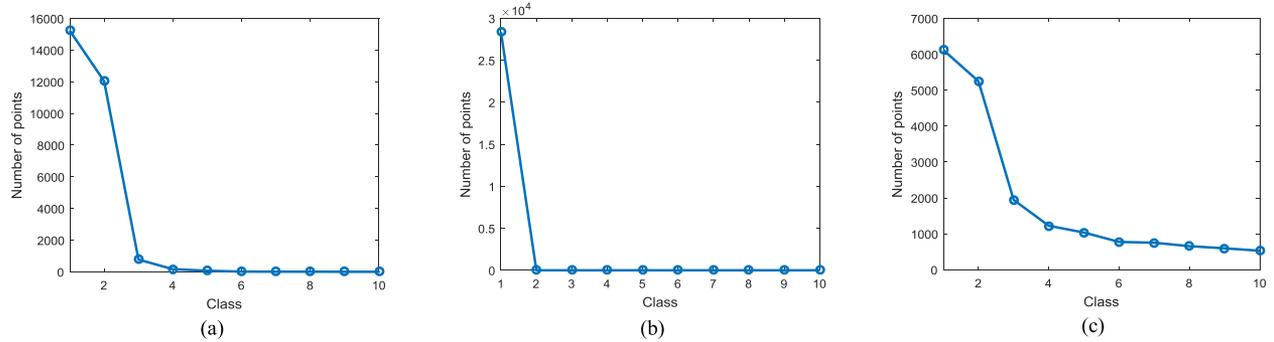


Fig. 10. The number of points in different classes after clustering in a general case (a), in the case of close DOAs (b), and in the case of large overlapping blocks (c).

TABLE III
OUTPUT SINRS WITH DIFFERENT ASSUMED NUMBERS OF SOURCES

Assumed number of sources	Output SINR (dB)		
	General case	Large overlapping blocks	Close DOAs
$K = 3$	7.5892	7.1189	7.2563
$K = M - 1$	7.5137	6.8454	7.2563

combined method achieves high output SINR. In the case of close DOAs, the delay vector cannot be estimated accurately. Thus, the performance of first method will degrade for its sensitivity to delay vector, whereas the second one can achieve good performance, as it can collect the complete information of s_1 and s_2 . Besides, the combined method also achieves satisfactory performance. This is because the first method can still suppress interferences to some degree, although its ability of filtering out interferences is reduced. Consequently, the disadvantages have limited effect on the combined method, and it can always achieve high output SINR.

F. The Number of Sources K

If the number of sources is known, we can achieve better performance. There are a lot of signal detection methods [28]–[30], which can be utilized to estimate K . Besides, we can also determine the source number based on the number of single-source points in each class after clustering. In the general case, Fig. 10(a) gives the number of points in different classes, where the two biggest classes contain much more points than other classes. Hence the number of interferences can be determined as $K_{in} = 2$, which means $K = K_{in} + 1 = 3$. Fig. 10(b) shows the number of points in the case of close DOAs, based on which we will get a wrong result $K_{in} = 1$. However, it has little effect on the performance, since the biggest class has contained the complete information of s_1 and s_2 . Fig. 10(c) plots the number of points in the case of large overlapping blocks. It can be seen that the two biggest classes also contain much more points. In fact, if the number of sources cannot be estimated, we can set $K = M - 1$ directly. This may degrade the performance to some degree, but the degradation is not severe shown in Table III. Note that in the case of close DOAs, all the single-source points

of interferences are clustered into one class. Hence the assumed number of sources has no influence on the performance.

G. Comparison With Other Beamforming Algorithms

To further indicate the effectiveness of the proposed beamformer, we compare it with the WC-PO beamformer [12], the DL beamformer [8], the derivative constraint beamformer [10], the Frost beamformer [3], the IPNCM reconstruction based beamformer [31] (referred to as reconstruction 1), and another IPNCM reconstruction based beamformer extended from the narrowband IPNCM reconstruction method (referred to as reconstruction 2), which reconstructs the IPNCM by

$$\hat{\mathbf{R}}_{i+n} = \frac{1}{D} \sum_{f \in [f_l, f_h]} \sum_{\theta \in \Theta_{\text{int}}} \frac{\mathbf{a}(\theta, f) \mathbf{a}(\theta, f)^H}{\mathbf{a}(\theta, f)^H \hat{\mathbf{R}}_{i-1}^{-1} \mathbf{a}(\theta, f)} \quad (51)$$

where $\mathbf{a}(\theta, f)$ denotes the SV at frequency f and DOA θ , Θ_{int} represents the potential DOA range of interferences, and D is the total number of DOA samples.

In the following simulations, we will evaluate their performance versus the number of snapshots, the input SNR, the DOA bias, and the position bias, respectively. For each scenario, 200 Monte-Carlo runs are performed. Note that the covariance matrix is calculated as $\hat{\mathbf{R}}_{\mathbf{x}} = \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{x}(n) \mathbf{x}^H(n)$ rather than $\hat{\mathbf{R}}_{\mathbf{x}} = \frac{1}{N} \sum_{k=0}^{K-1} \sum_{n=0}^{N-1} \mathbf{x}_k(n) \mathbf{x}_k^H(n) + \frac{1}{N} \mathbf{n}(t) \mathbf{n}^H(t)$, since the latter can be hardly achieved in practice for the mixture of source signals. By contrast, the former is more practical but needs more snapshots to achieve accurate covariance matrix.

Example 1. Number of snapshots: In this example, we evaluate the performance versus the number of snapshots with input SNR = 10 dB, and there is neither DOA error nor sensor position error. Fig. 11(a) shows the output SINRs, where the range of N is [100, 1000]. It can be seen that the proposed beamformer and the two IPNCM reconstruction based beamformers enjoy much faster convergence rate and achieve satisfactory performance even with small number of snapshots, such as $N = 300$. Instead, other traditional beamformers cannot achieve convergence, because they are sensitive to the accuracy of the covariance matrix, which needs to be calculated with enough snapshots. In Fig. 11(b), with the snapshots further increasing, the traditional beamformers achieve higher SINRs and converge when $N \geq 8000$. To guarantee satisfactory output SINR of the

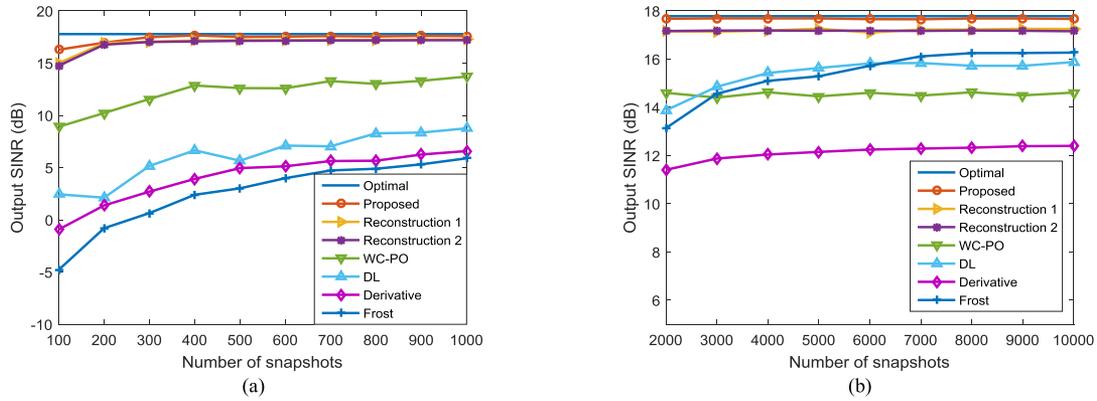


Fig. 11. Output SINR versus the number of snapshots. (a) the number varies from 100 to 1000; (b) the number varies from 2000 to 10^4 .

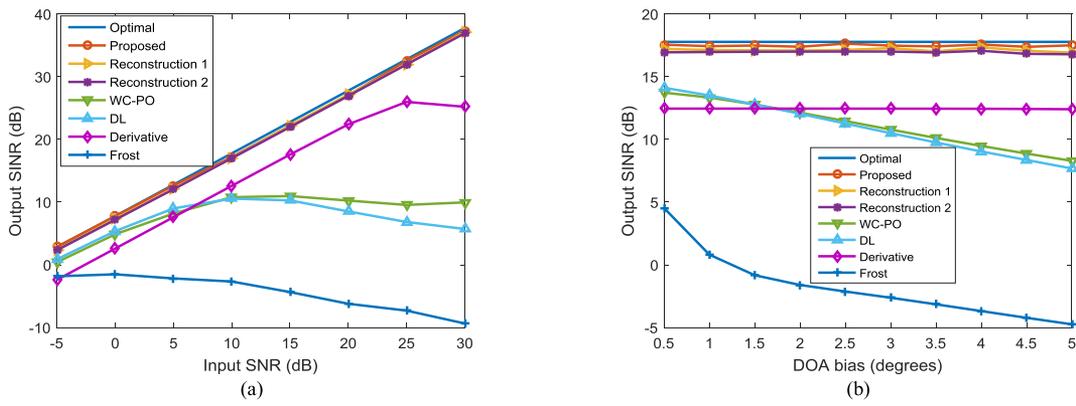


Fig. 12. In the case of DOA error, output SINR versus input SNR (a) and DOA bias (b).

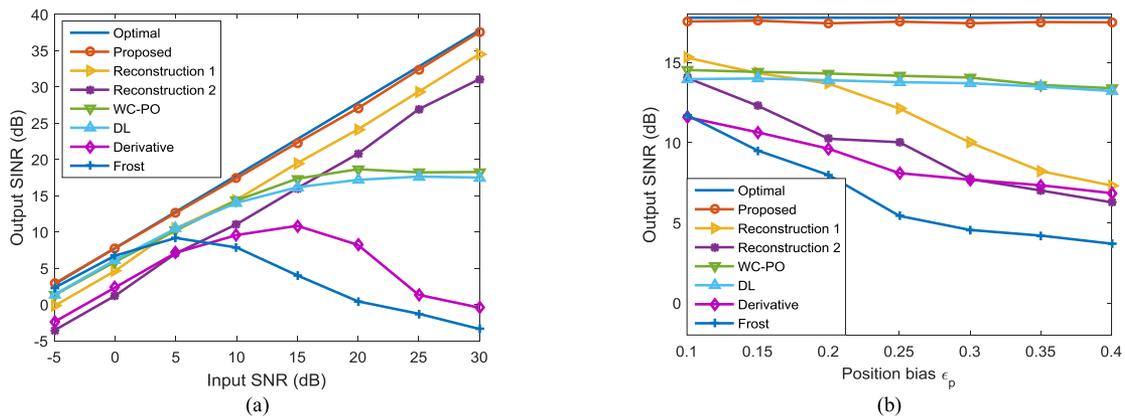


Fig. 13. In the case of sensor position error, output SINR versus input SNR (a) and position bias (b).

traditional beamformers, the number of snapshots is set to 10^4 for them in the following simulations, while $N = 300$ is set for the proposed beamformer and the IPNCM reconstruction based beamformers.

Example 2. DOA error: In the second example, we consider the influence of DOA error. Fig. 12(a) shows the output SINRs versus input SNR, where the DOA estimation bias is 3° . The Frost beamformer fails immediately since it has poor robustness. The DL beamformer, WC-PO beamformer and derivative constraint beamformer can obtain improved performance, but their output SINRs deviate from the optimal one gradually with

the SNR increasing. By contrast, the proposed beamformer and the IPNCM reconstruction based beamformers achieve high output SINR close to the optimal one over the whole SNR range, indicating their effectiveness in the case of high SNR and DOA error. Furthermore, the performance curves versus the DOA bias are plotted in Fig. 12(b). As expected, the proposed beamformer and the IPNCM reconstructed beamformers can always obtain satisfactory performance regardless of the DOA error, whereas the traditional beamformers, except the derivative constraint beamformer, have worse performance with the increase of DOA bias. The derivative constraint beamformer has strong robustness

against DOA error. However, it cannot achieve high enough output SINR, since its robustness is achieved at the expense of the ability of suppressing interferences and noise.

Example 3. Sensor position error: In the last example, a scenario with sensor position error is considered. The inter-sensor spacing $\Delta d_{j-1,j}$ is assumed to be a random value $\Delta d_{j-1,j} \in [\Delta d - \varepsilon_p \Delta d, \Delta d + \varepsilon_p \Delta d]$ ($j = 1, 2, \dots, 5$), and ε_p is set to 0.2 in Fig. 13(a). Compared with Fig. 12(a), the two IPNCM reconstructed beamformers suffer severe performance degradation for their poor robustness against sensor position error. As a comparison, the proposed beamformer can still achieve excellent performance. This is because it can estimate the SV and reconstruct the IPNCM accurately even in the presence of sensor position error. To further indicate this, Fig. 13(b) shows the output SINR versus the position bias. It can be seen that the proposed beamformer maintains high output SINR over the whole range of position bias, while other beamformers will decrease with the position bias.

VI. CONCLUSION

In this paper, we develop a robust wideband adaptive beamforming algorithm based on the STFT, which can reconstruct the IPNCM precisely with limited snapshots and estimate the steering vector accurately in the presence of DOA error and sensor position error. Unlike the existing beamformers, the proposed one use none of the perturbed array manifold information, leading to higher estimation and reconstruction accuracy. The simulation results have demonstrated that the proposed beamformer achieves satisfactory performance even with limited snapshots, and have strong robustness against direction error and sensor position error. The proposed algorithm is also suitable for other linear TFD. Besides, it can be further extended to quadratic TFD.

REFERENCES

- [1] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part IV: Optimum Array Processing*, New York, NY, USA: Wiley, 2002.
- [2] E. W. Vook and R. T. Compton, "Bandwidth performance of linear adaptive arrays with tapped delay-line processing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 901–908, Jul. 1992.
- [3] O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1972.
- [4] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Acoust., Speech, Signal Process. Mag.*, vol. 5, no. 2, pp. 4–24, Apr. 1988.
- [5] Y. Bucris, I. Cohen, and M. A. Doron, "Bayesian focusing for coherent wideband beamforming," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. 20, no. 4, pp. 1282–1296, May 2012.
- [6] Y. Feng, G. S. Gao, J. W. Xu, S. Q. Zhu, and C. Zeng, "Robust adaptive beamforming against large steering vector mismatch using multiple uncertainty sets," *Signal Process.*, vol. 152, pp. 320–330, Jun. 2018.
- [7] P. Chen, Y. X. Yang, Y. Wang, and Y. L. Ma, "Robust adaptive beamforming with sensor position errors using weighted subspace fitting-based covariance matrix reconstruction," *Sensors*, vol. 18, no. 5, pp. 1476.1–1476.12, May 2018.
- [8] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 397–401, Jul. 1988.
- [9] W. Liu and S. Weiss, *Wideband Beamforming: Concepts and Techniques*. Chichester, U.K.: Wiley, 2010.
- [10] M. H. ER and A. Cantoni, "Derivative constraints for broad-band element space antenna array processors," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. 31, no. 6, pp. 1378–1393, Dec. 1983.
- [11] S. Zhang and I. Thng, "Robust presteering derivative constraints for broadband antenna arrays," *IEEE Trans. Signal Process.*, vol. 50, no. 1, pp. 1–10, Jan. 2002.
- [12] Y. Zhao, W. Liu, and R. J. Langley, "Adaptive wideband beamforming with frequency invariance constraints," *IEEE Trans. Antennas Propag.*, vol. 59, no. 4, pp. 1175–1184, Apr. 2011.
- [13] Y. Zhao and W. Liu, "Robust wideband beamforming with frequency response variation constraint subject to arbitrary norm-bounded error," *IEEE Trans. Antennas Propag.*, vol. 60, no. 5, pp. 2566–2571, May 2012.
- [14] Y. Gu and A. Leshem, "Robust adaptive beamforming based on interference covariance matrix reconstruction and steering vector," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3881–3885, Jul. 2012.
- [15] Z. Zhang, W. Liu, W. Leng, A. Wang, and H. Shi, "Interference-plus-noise covariance matrix reconstruction via spatial power spectrum sampling for robust adaptive beamforming," *IEEE Signal Process. Lett.*, vol. 23, no. 1, pp. 121–125, Jan. 2016.
- [16] Z. X. Lu, Y. J. Li, M. G. Gao, and Y. R. Zhang, "Interference covariance matrix reconstruction via steering vectors estimation for robust adaptive beamforming," *Eletron. Lett.*, vol. 49, no. 22, pp. 1373–1374, Mar. 2016.
- [17] L. Huang, X. X. Zhang, and Z. Ye, "Robust adaptive beamforming with a novel interference-plus-noise covariance matrix reconstruction method," *IEEE Trans. Signal Process.*, vol. 63, no. 7, pp. 1643–1650, Apr. 2015.
- [18] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing Principle, Algorithm, and Applications*, 3rd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996.
- [19] Y. H. Luo, W. W. Wang, J. A. Chambers, S. Lambbotharan, and I. Proudler, "Exploitation of source nonstationarity in underdetermined blind source separation with advanced clustering techniques," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2198–2212, Jun. 2006.
- [20] S. Liu, Y. D. Zhang, T. Shan, and R. Tao, "Structure-aware Bayesian compressive sensing for frequency-hopping spectrum estimation with missing observations," *IEEE Trans. Signal Process.*, vol. 66, no. 8, pp. 2153–2166, Apr. 2018.
- [21] S. Liu, Z. Zeng, Y. D. Zhang, T. Fan, T. Shan, and R. Tao, "Automatic human fall detection in fractional Fourier domain for assisted living," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Shanghai, China, Mar. 2016, pp. 799–803.
- [22] A. Aissa-El-Bey, N. Linh-Trung, K. Abed-Meraim, A. Belouchrani, and Y. Grenier, "Underdetermined blind separation of nondisjoint sources in the time-frequency domain," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 897–907, Mar. 2007.
- [23] D. Z. Peng and Y. Xiang, "Underdetermined blind source separation based on relaxed sparsity condition of sources," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 809–814, Feb. 2009.
- [24] A. Belouchrani and M. G. Amin, "Blind source separation using time-frequency distributions: Algorithm and asymptotic performance," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Munich, Germany, Apr. 1997, pp. 3469–3472.
- [25] A. Aissa-El-Bey, K. Abed-Meraim, and Y. Grenier, "Blind separation of underdetermined convolutive mixtures using their time-frequency representation," *IEEE Trans. Signal Process.*, vol. 15, no. 5, pp. 1540–1550, Jul. 2007.
- [26] V. G. Reju, S. N. Koh, and I. Y. Soon, "Underdetermined convolutive blind source separation via time-frequency masking," *IEEE Trans. Audio Speech Lang. Process.*, vol. 18, no. 1, pp. 101–116, Feb. 2010.
- [27] Y. M. Zhang and M. G. Amin, "Array processing for nonstationary interference suppression in DS/SS communications using subspace projection techniques," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3005–3014, Dec. 2001.
- [28] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 387–392, Apr. 1985.
- [29] L. Huang, S. J. Wu, and L. Xia, "Reduced-rank MDL method for source enumeration in high-resolution array processing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5658–5667, Dec. 2007.
- [30] P. J. Chung, J. F. Böhme, C. F. Mecklenbräuker, and A. O. Hero, "Detection of the number of signals using the Benjamini-Hochberg procedure," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2497–2508, Jun. 2007.
- [31] P. Chen, Y. X. Yang, Y. Wang, Y. L. Ma, and L. Yang, "Robust covariance matrix reconstruction algorithm for time-domain wideband adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5658–5667, Dec. 2007.