Digital Signal Processing 52 (2016) 25–34

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

DOA estimation of closely-spaced and spectrally-overlapped sources using a STFT-based MUSIC algorithm ☆

Haijian Zhang [a,b,*], Guoan Bi [b], Yunlong Cai [c], Sirajudeen Gulam Razul [d,e], Chong Meng Samson See [d,e]

[a] School of Electronic Information, Wuhan University, China
[b] School of EEE, Nanyang Technological University, Singapore
[c] Department of Information Science and Electronic Engineering, Zhejiang University, China
[d] Temasek Laboratories@NTU, Singapore
[e] DSO National Laboratories, Singapore

A R T I C L E   I N F O

Article history:
Available online 23 February 2016

Keywords:
DOA estimation
MUSIC
Closely-spaced sources
Spectrally-overlapped sources
Short-time Fourier transform

A B S T R A C T

The multiple signal classification (MUSIC) algorithm based on spatial time-frequency distribution (STFD) has been investigated for direction of arrival (DOA) estimation of closely-spaced sources. However, the limitations of the bilinear time-frequency based MUSIC (TF-MUSIC) algorithm lie in that it suffers from heavy implementation complexity, and its performance strongly depends on appropriate selection of auto-term location of the sources in time-frequency (TF) domain for the formulation of a group of STFD matrices, which is practically difficult especially when the sources are spectrally-overlapped. In order to relax these limitations, this paper aims to develop a novel DOA estimation algorithm. Specifically, we build a MUSIC algorithm based on spatial short-time Fourier transform (STFT), which effectively reduces implementation cost. More importantly, we propose an efficient method to precisely select single-source auto-term location for constructing the STFD matrices of each source. In addition to low complexity, the main advantage of the proposed STFT-MUSIC algorithm compared to some existing ones is that it can better deal with closely-spaced sources whose spectral contents are highly overlapped in TF domain.
© 2016 Elsevier Inc. All rights reserved.

1. Introduction

The direction of arrival (DOA) estimation has received tremendous research attention in different areas, such as radar, sonar, speech, and communication [1–6]. The multiple signal classification (MUSIC) algorithm and its variants [7–18] based on the eigen-structure of the observed covariance matrix have been efficient estimation techniques. Nevertheless, the MUSIC algorithm has limitations with regard to resolving closely-spaced sources in low SNR environments. In addition, it can only work when the number of sensors is larger than the number of sources. With the development of spatial time-frequency distributions (STFDs) [19–24], the conventional MUSIC based on space-time processing has been developed by constructing STFD matrices instead of the covariance matrices for narrowband DOA estimation [25–32]. It has also been extended for wideband scenarios [33]. Compared to the temporal MUSIC, the bilinear time-frequency based MUSIC (TF-MUSIC) can provide noise-robust direction finding since the noise power is spread over the time-frequency (TF) domain while the source energy is localized [34]. Moreover, the selection of TF region belonging to a specific set of sources allows the improved DOA estimation with a small number of sensors in an underdetermined case. By separately dealing with the auto-source TF points of each source, the DOAs of closely-spaced sources can be estimated as long as they are distinctly distributed in TF domain.

However, the TF-MUSIC[1] is computationally complicated compared to the temporal MUSIC, and it is highly dependent on the selection of auto-term location of the sources in TF domain[2] to

☆ This work was supported in part by the National Natural Science Foundation of China (No. 61501335, 61471319), the Natural Science Foundation of Hubei Province (No. 2015CFB202) and Zhejiang Provincial Natural Science Foundation of China under Grant LY14F010013.

* Correspondence to: School of Electronic Information, Wuhan University, Wuhan 430072, China.

E-mail addresses: haijian.zhang@whu.edu.cn (H. Zhang), egbi@ntu.edu.sg (G. Bi), ylcai@zju.edu.cn (Y. Cai), esirajudeen@ntu.edu.sg (S.G. Razul), samsonsee@ntu.edu.sg (C.M.S. See).

[1] Herein the TF-MUSIC means the MUSIC algorithm based on pseudo Wigner–Ville distribution (PWVD) unless otherwise specified.

[2] We term the auto-term location as auto-source TF points in this paper.

http://dx.doi.org/10.1016/j.dsp.2016.01.015
1051-2004/© 2016 Elsevier Inc. All rights reserved.

mitigate the noise effect and ensure the full rank property of the STFD matrix. Multiple auto-source TF points [25,27] or cross-source TF points [26] have been considered to construct the STFD matrix. Most of previous studies assume that the source signals are sparse in TF domain and the TF signature of each source is known. However, the precise information of auto-source TF points in practical situations might be unavailable, and the estimation of the TF signature of individual source is required. This problem becomes significant for scenarios involving a large number of sources, whose spectral contents are overlapped in TF domain. Therefore, an estimation of appropriate auto-source TF points becomes practically crucial. Some efficient selection methods have been reported. In [35], Heidenreich et al. applied morphological image processing to detect instantaneous frequency (IF) segments of each source. The detected IF segments are combined based on a bootstrap resampling technique, and the linking IF segments belonging to a single source are then used for DOA estimation. Other TF point selection methods based on quadratic STFDs can refer to [29,36,37]. Due to the computational complexity and the cross-terms of quadratic time-frequency distributions (TFDs), the linear TFD-based DOA estimation techniques, e.g., the fractional Fourier transform (FRFT) in [38] and the short-time Fourier transform (STFT) in [39], have also been investigated in the literature. Some methods of selecting signal-source TF points based on linear TFDs of mixtures have been reported in [40–44]. However, they have some limitations, i.e., the method in [40,41] is only suitable for two mixtures, and requires the sources to be W-disjoint orthogonal in TF domain. The methods in [42–44] are designed in the case of real-valued mixing matrix. More sophisticated selection methods in the presence of complex-valued mixing matrix are strongly required particularly for the case where the sources are spectrally-overlapped in a low SNR environment.

To ease the limitations of the temporal MUSIC and the bilinear TF-MUSIC, herein we propose a short-time Fourier transform based MUSIC (STFT-MUSIC) algorithm, aiming at solving the challenging scenarios associated with a large number of closely-spaced and spectrally-overlapped sources. The research emphasis is on the STFT because it is simple to implement and the cross-term is avoided. The main contribution in the paper is that an efficient method to automatically select single-source TF points (i.e., the TF points where only one source exists) is proposed. Instead of detecting the TF points along the IF trajectory, we propose a subspace projection based method to detect single-source TF points of each source in the presence of complex-valued mixing matrix. Although the linear STFT suffers low resolution in TF domain, appropriate single-source TF points can be selected for accurate DOA estimation since the proposed algorithm is not focused on the signal's IF laws. Compared to the temporal MUSIC and the bilinear TF-MUSIC, the advantages of the proposed STFT-MUSIC algorithm are fourfold:

- The ability to resolve the closely-spaced sources, which is the limitation of the temporal MUSIC.
- The feasibility in underdetermined cases, which is the limitation of the temporal MUSIC.
- The feasibility and efficiency to deal with spectrally-overlapped sources, which are the limitations of the TF-MUSIC.
- The low complexity, which is the limitation of the TF-MUSIC.

The remainder of this paper is structured as follows. We describe the STFT-MUSIC algorithm in Section 2. In Section 3, the details of the proposed method of how to choose single-source TF points are given. In Section 4, the STFT-MUSIC algorithm with selected single-source TF points is evaluated on two scenarios, and the comparison with existing algorithms is presented. The discussion of the proposed algorithm with existing work is given in Section 5. Finally, Section 6 concludes the paper.

2. The proposed STFT-MUSIC algorithm

In our study, a linear and equispaced sensor array having M elements is considered, i.e., the spatial distribution of sensors is uniform linear array (ULA). Let $s_i(t), i = 1, \dots, N$, denote the unknown sources, where N is the number of sources impinging on an M-dimensional ULA from N distinct directions $\theta_1, \dots, \theta_N$. The output signals $x_m(t), m = 1, \dots, M$ are modeled as

$$\mathbf{x}(t) = \mathbf{A}(\theta)\mathbf{s}(t) + \mathbf{n}(t), \tag{1}$$

where $\mathbf{A}(\theta) = \left[\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_N)\right]$ denotes the complex-valued mixing matrix, and $\mathbf{a}(\theta_i)$ is the steering vector of the ith source. $\mathbf{x}(t) = \left[x_1(t), \dots, x_M(t)\right]^T$ are the received mixtures, $\mathbf{s}(t) = \left[s_1(t), \dots, s_N(t)\right]^T$ are the signal sources, and $\mathbf{n}(t)$ is the additive white Gaussian noise vector. $[\cdot]^T$ is the transpose operator.

The STFT (denoted by $\mathbf{S}$ hereafter) of the array output vector $\mathbf{x}(t)$ in (1) without noise is computed as

$$\mathbf{S_x}(t, f) = \mathbf{A}(\theta)\mathbf{S_s}(t, f) = \left[\mathbf{a}(\theta_1) \quad \dots \quad \mathbf{a}(\theta_N)\right] \begin{bmatrix} \mathbf{S}_{s_1}(t, f) \\ \vdots \\ \mathbf{S}_{s_N}(t, f) \end{bmatrix}, \tag{2}$$

where $\mathbf{S}_{s_i}(t, f)$ denotes the STFT value of the ith source. The steering vector of the ith source denotes $\mathbf{a}(\theta_i) = [a_{i1}, \dots, a_{iM}]^T$, and its mth element is expressed as

$$a_{im} = \frac{1}{\sqrt{M}} e^{-j\frac{2\pi}{\lambda}d(m-1)sin(\theta_i)}, \quad m \in \{1, \dots, M\}, \tag{3}$$

where d is the inter-element spacing, λ denotes the wavelength, and θ_i is the DOA of the ith source to be estimated.

Define $\mathbf{D_{xx}}(t, f)$ to be formulated as

$$\begin{aligned}\mathbf{D_{xx}}(t, f) &= \mathbf{S_x}(t, f)\mathbf{S_x}^H(t, f) = \mathbf{A}(\theta)\mathbf{D_{ss}}(t, f)\mathbf{A}^H(\theta) \\ &= \mathbf{A}(\theta) \begin{bmatrix} \mathbf{S}_{s_1}(t, f)\mathbf{S}_{s_1}^*(t, f) & \dots & \mathbf{S}_{s_1}(t, f)\mathbf{S}_{s_N}^*(t, f) \\ \vdots & \vdots & \vdots \\ \mathbf{S}_{s_N}(t, f)\mathbf{S}_{s_1}^*(t, f) & \dots & \mathbf{S}_{s_N}(t, f)\mathbf{S}_{s_N}^*(t, f) \end{bmatrix} \\ &\quad \times \mathbf{A}^H(\theta),\end{aligned} \tag{4}$$

where $\mathbf{D_{ss}}(t, f)$ denotes the source STFD matrix whose (i, j)th element is $\mathbf{D}_{s_i s_j}(t, f) = \mathbf{S}_{s_i}(t, f)\mathbf{S}_{s_j}^*(t, f)$. The conjugate operator is denoted by $[\cdot]^*$.

The STFT-MUSIC algorithm estimates the DOA by determining the N peaks of the spatial spectrum

$$P(\theta) = \frac{\mathbf{a}^H(\theta)\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\mathbf{U}_n\mathbf{U}_n^H\mathbf{a}(\theta)}, \tag{5}$$

where $\mathbf{U}_n$ denotes the noise eigenvectors of the STFD matrices in (4), which are replaced by the covariance matrices of $\mathbf{x}(t)$ in the conventional MUSIC algorithm [9]. The conjugate transpose operator is denoted by $[\cdot]^H$. The STFD matrices lead to an enhanced SNR, and therefore to an improved accuracy of DOA estimation [27]. Assuming a group of single-source TF point sets $\Omega_i, i = 1, \dots, N$, for individual sources are determined, the averaged STFD matrices are obtained by

$$\overline{\mathbf{D}}_i = \frac{1}{\sharp\Omega_i} \sum_{(t,f)\in\Omega_i} \mathbf{D_{xx}}(t, f), \quad i = 1, \dots, N, \tag{6}$$

where $\sharp\Omega_i$ denotes the number of TF points in the set Ω_i, and $\mathbf{D_{xx}}(t, f)$ is the STFD matrix at the TF point (t, f). In the STFT-MUSIC algorithm, the noise subspace $\mathbf{U}_n$ in (5) is therefore computed by the eigen-decomposition of the matrix $\overline{\mathbf{D}}_i$.

It is noted that the premise of implementing the STFT-MUSIC algorithm lies in the available single-source point sets to construct the STFD matrices in (6). The single-source point set of each source is separately used for implementing the STFT-MUSIC algorithm via (6) and (5), which offers the potential ability to independently estimate the DOA of each source with a smaller number of sensors in underdetermined cases, i.e., $N > M$. In the next section, we elaborate a proposed method to determine a group of effective single-source point sets Ω_i, $i = 1, \ldots, N$, for the N sources, respectively.

3. Selection of single-source TF points

The TF-MUSIC algorithm generally assumes that different sources are approximately disjoint in TF domain, and then it implements the sequential DOA estimation of individual source by masking the TF region corresponding to a single source or a subset of sources. The desirable TF points for accurate DOA estimation of one source are the single-source TF points associated to this source. However, it becomes difficult to determine the single-source TF points when different sources are non-disjoint in TF domain. The multiple-source TF points (the TF points where multiple sources exist) and the cross-term TF points, if they are chosen, will severely impact the accuracy of DOA estimate. That is the reason why we use the STFT to avoid the cross-terms and investigate a feasible solution to select single-source TF points by getting rid of multiple-source TF points and noise TF points.

In our study, the proposed single-source selection method complies with the following three relaxed assumptions:

1. In practice, different sources might be closely localized in space and highly overlapped in TF domain. As shown in Fig. 1 (a), four frequency-modulated (FM) sources are non-disjoint in TF domain.
2. The number of sources presenting at each auto-source TF point is not larger than two. Certain auto-source TF points may involve more than two sources, however, this seldom happens and can be neglected.
3. For every source, there always exist some TF points that are associated with this source only. This assumption is practical because it has a very low probability that different sources completely overlap in TF domain.

First of all, a set of auto-source TF points are determined by detecting the TF points with high energy concentration in TF domain, which is implemented by using the following criterion at each time-instant [45]

$$\frac{||\mathbf{S}_{\mathbf{x}}(t,f)||}{\max_{\nu} ||\mathbf{S}_{\mathbf{x}}(t,\nu)||} > \epsilon_0, \tag{7}$$

where $||\cdot||$ is the norm operator, ν denotes the frequency index, and ϵ_0 is an empirical threshold value for the selection of TF points with strong energy. Assuming the TF points which satisfy the criterion in (7) are contained in the set Ω. Fig. 1 (b) displays the auto-source TF points with strong energy in Fig. 1 (a) when SNR=10 dB by assigning $\epsilon_0 = 0.2$.

Actually, the selection strategy of auto-term points in (7) based on the ratio between the norm and the maximum norm at the specific time instant, has already been applied in existing work, e.g., eq. (18) in [45]. In contrast, the selection of auto-terms was proposed in [36] and [49] using the ratio between the trace and norm of the instantaneous STFD matrix. The reason why choosing this kind of selection strategy in (7) is that the STFT is free of cross-term, and a TF point with a non-zero STFT value is ideally an auto-term TF point. Therefore, the auto-term TF points can be selected by applying a noise thresholding procedure. In STFT domain, the effects of spreading the noise energy while localizing the source energy will keep only TF points whose energy is significant. In a word, the two different strategies for spatial STFT and spatial bilinear TFD, have the same objective and capability, i.e., effectively detecting auto-term TF points having sufficient energy.

Under the first two assumptions, each auto-source TF point in Ω is either a single-source point or a double-source point. Besides, some TF points where only noise exists might be detected as spurious auto-source points due to an inappropriate choice of the threshold value ϵ_0. Therefore, the auto-source TF point set Ω consists of three types of TF point: single-source TF point, double-source TF point and noise TF point. The essential idea of the proposed method is to extract the entire or a subset of single-source TF points from the set Ω. For identifying a set of effective single-source TF points, our next step is to estimate a matrix ($\mathbf{A}_0$ shown in (10)), which helps to identify single-source TF points.

For each TF point in Ω obtained by (7), we then compute its normalized spatial vector by making its first element real and positive

$$\mathbf{v}(t,f) = \begin{bmatrix} \frac{S_{x_1}(t,f)}{||\mathbf{S}_{\mathbf{x}}(t,f)||} \\ \vdots \\ \frac{S_{x_M}(t,f)}{||\mathbf{S}_{\mathbf{x}}(t,f)||} \end{bmatrix} \cdot \frac{||S_{x_1}(t,f)||}{S_{x_1}(t,f)}, \quad (t,f) \in \Omega. \tag{8}$$

Next, the k-means clustering method is applied to classify all the spatial vectors computed in (8) given a fixed number of clusters N_0 (how to choose this parameter is discussed in the following), which is usually assigned a value larger than N. The TF points with similar spatial vectors will be grouped into the same cluster. The k-means clustering result by setting $N_0 = 8$ on the set Ω is shown in Fig. 1 (b), where we use different colors to represent different clusters, e.g., the yellow color represents the TF points at the intersections of overlapped signals, which are automatically classified into one cluster by the k-means method. Assuming the k-means clustering method gives rise to N_0 clusters $\{\Gamma_k | k = 1, \ldots, N_0\}$, where Γ_k denotes the TF point set of the kth cluster. A direction vector can be then obtained by averaging the spatial vectors of all the TF points belonging to the same cluster

$$\widehat{\mathbf{v}}_k = \frac{1}{\sharp\Gamma_k} \sum_{(t,f)\in\Gamma_k} \mathbf{v}(t,f), \qquad k = 1, \ldots, N_0. \tag{9}$$

This linear combination of spatial vectors gives an estimation of the steering vector of some source if the TF points in the same cluster belong to this source, which is explained in the Appendix.

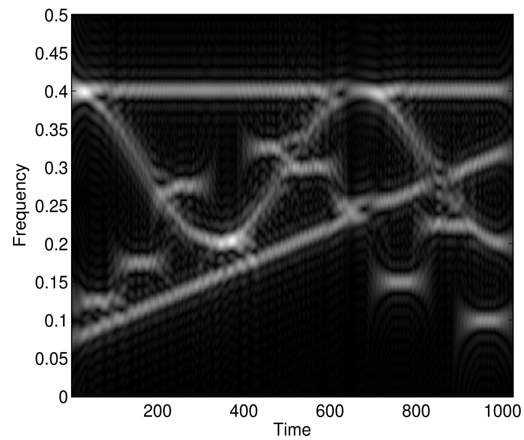
As a consequence, we obtain an $M \times N_0$ matrix $\mathbf{A}_0$ as below

$$\mathbf{A}_0 = [\widehat{\mathbf{v}}_1 \quad \widehat{\mathbf{v}}_2 \quad \cdots \quad \widehat{\mathbf{v}}_{N_0}]. \tag{10}$$

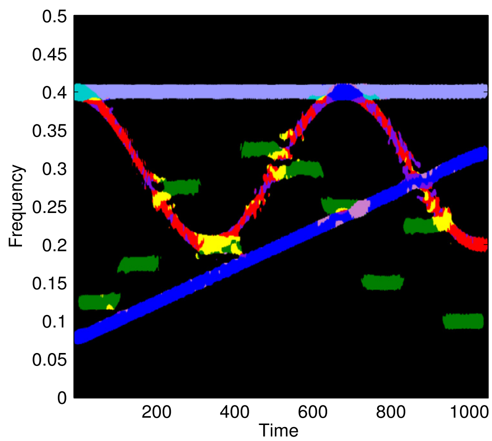
Mathematically, the column vectors of $\mathbf{A}_0$ estimated in (10) can be approximated as combinations of three different types of spatial vectors

$$\mathbf{A}_0^{\flat} = [\overbrace{\mathbf{a}_1 \quad \mathbf{a}_2 \cdots \mathbf{a}_N}^{\text{The 1st part: pure}} \quad \overbrace{\mathbf{a}_{12} \quad \mathbf{a}_{13} \cdots \mathbf{a}_{(N-1)N}}^{\text{The 2nd part: mixed}} \quad \overbrace{\text{other cases}}^{\text{The 3rd part: random}}], \tag{11}$$

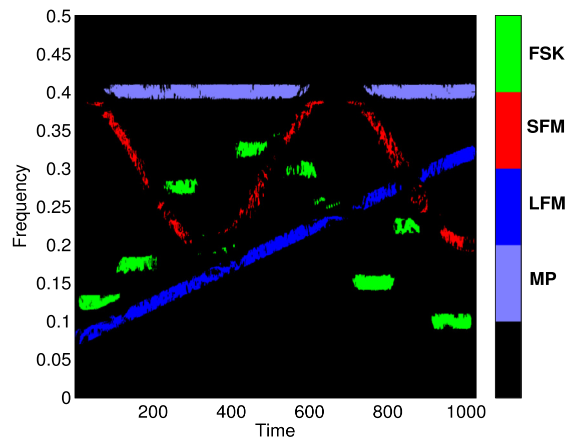
where $\mathbf{A}_0^{\flat}$ is the variant of $\mathbf{A}_0$ by reordering its columns. It implies that $\mathbf{A}_0$ is made up of three components: single-signal steering vectors of N sources, compound spatial vectors among N sources, and other situations, e.g., random spatial vectors due to white noise. This estimated $\mathbf{A}_0$ is helpful for the identification of single-source TF points in Ω.



(a) STFT image free of noise.



(b) Auto-source points $\in \Omega$ with N_0=8 classes.



(c) Single-source points $\in \Omega^*$ with N=4 classes.

Fig. 1. Selected single-source TF points of four spectrally-overlapped FM sources (MP: monopulse; LFM: linear FM; SFM: sinusoidal FM; FSK: frequency-shift-keying) with DOAs: $\theta_1 = -10°$, $\theta_2 = -5°$, $\theta_3 = 0°$, and $\theta_4 = 5°$ ($N = 4$ and $M = 10$). (a) STFT image. (b) Clustering result in different colors of detected auto-source TF points by (7) at SNR $= 10$ dB ($\epsilon_0 = 0.2$). (c) Clustering result of detected single-source TF points by (14) using $\lambda_0 = 5$ and $l_0 = 0.5$.

According to assumption 2, we define that the energy at each TF point in the set Ω is contributed by two STFT values: $S_1(t, f)$ and $S_2(t, f)$, which can be derived from (2)

$$\mathbf{S_s}(t, f) = \begin{bmatrix} S_1(t, f) \\ S_2(t, f) \end{bmatrix} = \mathbf{A}_2^{\natural}\mathbf{S_x}(t, f), \quad (t, f) \in \Omega, \tag{12}$$

where $\natural$ is the Moore–Penrose pseudoinverse operator. $\mathbf{A}_2 = [\mathbf{a}_{n_1}, \mathbf{a}_{n_2}]$ denote the steering vectors of two most possible sources present at every point $\in \Omega$. For single-source points, the two STFT values in (12) stem from a specific source and noise, respectively. For every point in the set Ω, the optimal $\mathbf{a}_{n_1}$ and $\mathbf{a}_{n_2}$ are determined based on $\mathbf{A}_0$ by minimizing the following subspace projection

$$\{\mathbf{a}_{n_1}, \mathbf{a}_{n_2}\} = \arg \min_{\mathbf{a}_{m_1}, \mathbf{a}_{m_2}} \left\{ ||\mathbf{Q}\mathbf{S_x}(t, f)|| \right\}, \quad (t, f) \in \Omega, \tag{13}$$

where $\mathbf{Q} = \mathbf{I} - \widetilde{\mathbf{A}}_2(\widetilde{\mathbf{A}}_2^H\widetilde{\mathbf{A}}_2)^{-1}\widetilde{\mathbf{A}}_2^H$ means the orthogonal projection matrix into the noise subspace of $\widetilde{\mathbf{A}}_2$ [45], $\mathbf{I}$ denotes the identity matrix, and $\widetilde{\mathbf{A}}_2 = [\mathbf{a}_{m_1}, \mathbf{a}_{m_2}]$. The elements $\mathbf{a}_{m_1}$ and $\mathbf{a}_{m_2}$ are two random columns of the matrix $\mathbf{A}_0$, where $m_1, m_2 \in \{1, \ldots, N_0\}$.

It should be discussed that the two steering vectors of $\mathbf{A}_2$ in (12) for the TF points in Ω can be efficiently selected from the column vectors of $\mathbf{A}_0$ by the minimization operation in (13). The optimal $\mathbf{a}_{n_1}$ and $\mathbf{a}_{n_2}$ shall be selected among the 1st part of $\mathbf{A}_0^{\flat}$ because the minimization operation will automatically choose pure steering vectors. More specifically, for double-source TF points with sources i and j, $i, j \in \{1, \ldots, N\}$, the resultant $\mathbf{a}_{n_1}$ and $\mathbf{a}_{n_2}$ by (13) will be the purest steering vectors of source i and source j chosen from the 1st part of $\mathbf{A}_0^{\flat}$. For the TF points with a single source i, $i \in \{1, \ldots, N\}$, one of $\mathbf{a}_{n_1}$ and $\mathbf{a}_{n_2}$ will be chosen from the 1st part of $\mathbf{A}_0^{\flat}$, while the other one can be chosen from any column vector of $\mathbf{A}_0^{\flat}$ because of the property of random noise. We should also mention that the selection of the value of N_0 has a wide range of space, because the two STFT values in (12) can be successfully computed as long as the estimated $\mathbf{A}_0$ contains at least one pure steering vector for each source. In our study, the value of N_0 is set as around twice the value of the true number of sources, i.e., N.

The single-source points in Ω are determined according to the computed STFT values: S_1 and S_2 from (12) and (13). Based on the fact that the magnitude of the two STFT values at a single-source TF point is very different and the smaller value due to noise should be trivial, a set of single-source TF points are therefore extracted from the set Ω by defining proper thresholds according to the following criteria

$$\begin{cases} \dfrac{\max\{|S_1|, |S_2|\}}{\min\{|S_1|, |S_2|\}} & > \lambda_0, \\ \min\{|S_1|, |S_2|\} & < l_0, \end{cases} \tag{14}$$

where $|\cdot|$ denotes the absolute operator. λ_0 and l_0 are two thresholds empirically determined. Defining the set Ω^* which includes the single-source TF points in Ω to satisfy (14). The set Ω^* by setting $\lambda_0 = 5$ and $l_0 = 0.5$ is shown in Fig. 1 (c). Note that the

Table 1
Selection method of single-source TF points.

1)	Detect a set of auto-source TF points $\in \Omega$ by (7), and compute the spatial vectors of the points in Ω by (8).
2)	Apply k-means clustering method on the computed spatial vectors in Ω to obtain $\mathbf{A}_0$ in (10).
3)	Based on $\mathbf{A}_0$, extract a set of signal-source TF points $\in \Omega^*$ from Ω by sequentially implementing (13), (12) and (14).
4)	Apply k-means clustering method on Ω^* by giving the known number of clusters, N. The resultant clusters give the estimates of $\Omega_i, i = 1, \cdots, N$ in (6).

double-source TF points or noise points are well ignored, while a group of single-source points are kept. It is worth mentioning that the final DOA estimation is insensitive to the values of the thresholds ϵ_0, λ_0 and l_0, because different values of these thresholds mainly influence the quantity of selected single-source points from Ω.

Lastly, the k-means clustering method, given the number of clusters, N,[3] is carried out on the set Ω^*. From the clustering result in Fig. 1 (c), it is seen that the whole single-source TF point set Ω^* is categorized into individual single-source point subsets. The TF point sets Ω_i, $i = 1, \ldots, N$, in (6) belonging to N sources are therefore determined. The relationship between aforementioned point sets is expressed as

$$\cup\Omega_i = \Omega^* \subset \Omega, \quad i = 1, \ldots, N. \tag{15}$$

The procedure of the selection method of single-source TF points is summarized in Table 1.

4. Simulation

In this section, numerical results on two different scenarios are presented to show the potential advantages of the proposed STFT-MUSIC algorithm compared to the temporal MUSIC algorithm and the TF-MUSIC algorithm. The use of the proposed STFT-MUSIC algorithm is to be demonstrated in the following three aspects:

- The ability to resolve the closely-spaced sources.
- The feasibility in underdetermined cases.
- The feasibility and efficiency to deal with spectrally-overlapped sources.

We consider a uniform linear array whose sensors being separated by half wavelength spacing. It is assumed that all sources have the same power, and 1024 snapshots are used.

4.1. Scenario with two spectrally-overlapped linear frequency-modulated (LFM) sources

Let us consider two spectrally-overlapped LFM sources coming from different DOAs: $\theta_1 = -2.5°$ and $\theta_2 = 2.5°$. The window length of the STFT is 128, and $M = 5$ sensors are used. Different from temporal MUSIC, the TF-MUSIC and STFT-MUSIC based on known IF information assume single-source DOA estimation.

It is assumed that the two LFM sources are spectrally overlapped in TF domain. In order to obtain accurate DOA estimation of the two LFMs, we use the proposed method in Table 1 to eliminate the multi-source TF points, and keep the single-source TF points, which are utilized to estimate the DOA of each source. The spatial vectors of the detected single-source TF points of two spectrally-overlapped LFM sources are shown in Fig. 2 (a) at SNR $= 5$ dB, and the clustering result of the single-source TF points in Fig. 2 (a) is shown in Fig. 2 (b). Note that the multi-source TF points (overlapping TF points) are well eliminated. The estimated spatial spectra through different algorithms are given in Fig. 2 (c), which shows that the temporal MUSIC gives the whole DOA information, and cannot resolve the DOAs of closely-spaced targets. Since the TF-MUSIC and the STFT-MUSIC assume single-source DOA estimation, they can achieve desirable performance based on the true IFs of individual sources. In contrast, the STFT-MUSIC algorithm based on the selected single-source TF points can provide more accurate DOA estimation compared to other algorithms due to the elimination of the TF points at the intersection of two LFM sources.

Fig. 2 (d) shows the root-mean-square error (RMSE) of DOA estimation at each SNR level computed from 100 independent trials. For comparison purpose, the Cramér–Rao bound (CRB) [46,47] is also presented. Based on known IF information, the TF-MUSIC slightly outperforms the STFT-MUSIC, and they both outperform the temporal MUSIC, which fails for closely-spaced sources. The STFT-MUSIC algorithm based on selected single-source TF points can achieve the best performance at high SNR levels, while its performance is still desirable when SNR degrades. Lastly, the DOA estimation performance of the STFT-MUSIC with $\theta_1 = -0.5°$ and $\theta_2 = 0.5°$ is displayed in Fig. 3 (a). When the sources are very close, e.g., with DOA spacing 1°, it is still possible to classify different sources due to the following two reasons: The classification is implemented based on the detection results of single-source TF points. Although the spatial signatures are very close, they are still distributed in distinct clusters, as shown in Fig. 2 (a); We assume the number of sources is known, the k-means clustering method is used to classify the spatial vectors of the detected single-source TF points. The RMSE of DOA estimation with $\theta_1 = -0.5°$ and $\theta_2 = 0.5°$ at different SNR levels is shown in Fig. 3 (b). It is seen that although the two sources have a small space separation, the STFT-MUSIC using the estimated single-source points can still resolve their separate DOAs. However, there will be a performance limitation when DOA spacing becomes smaller (DOA spacing $< 1°$) or the SNR level decreases (SNR < 15 dB).

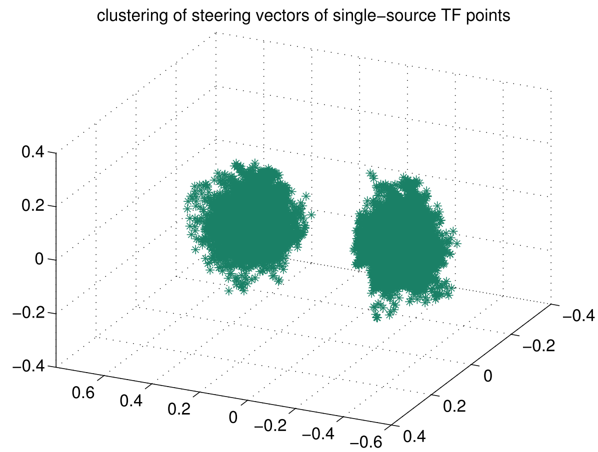
It should be explained that the STFT-MUSIC with selected single-source TF points can achieve more accurate DOA estimation performance since the selection method only extracts a group of single-source TF points (as shown in Fig. 2 (a)), which are suitable for single-source DOA estimation. In addition, the proposed method automatically removes the multi-source TF points, which appear at the intersection of the two spectrally-overlapped LFM signals, as shown in Fig. 2 (b). In contrast, TF-MUSIC and STFT-MUSIC based on known IF information give rise to less accurate DOA estimation due to two reasons: Only the TF points along the IF trajectory are used for DOA estimation, which might be insufficient for an accurate estimation; The estimation bias resulting from the interference between different sources at the intersections in STFT domain.

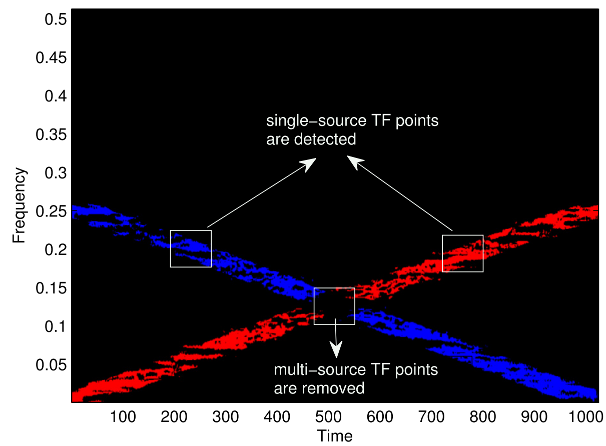
4.2. Scenario with a large number of sources

Let us next consider four FM sources shown in Fig. 1, which are located at different DOAs: $\theta_1 = -10°$, $\theta_2 = -5°$, $\theta_3 = 0°$, and $\theta_4 = 5°$. The window length of the STFT is 64, and $M = 10$ sensors are used. The estimated spatial spectra by different algorithms are shown in Fig. 4 (a) at SNR $= 5$ dB. Note that an estimation bias appears for TF-MUSIC based on true IFs due to the interferences between different sources at the intersections in TF domain, whereas the proposed STFT-MUSIC effectively reduces this bias.

The DOA estimation in the underdetermined case is feasible by the STFT-MUSIC, whereas the temporal MUSIC can only work when the number of sensors is larger than the number of sources. Next,

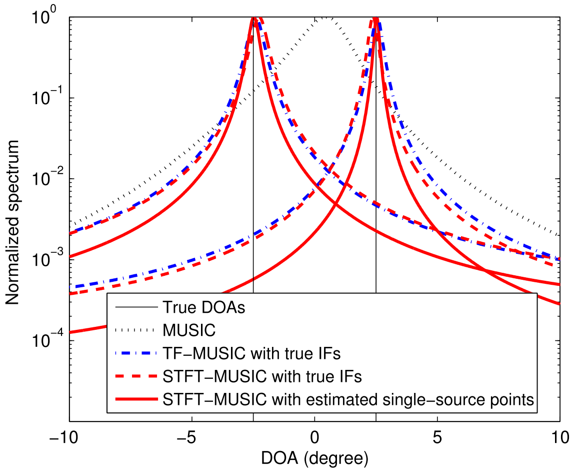
[3] Although N is assumed to be known, clustering methods without knowing N can be also used, and the number of clusters gives an estimated N.

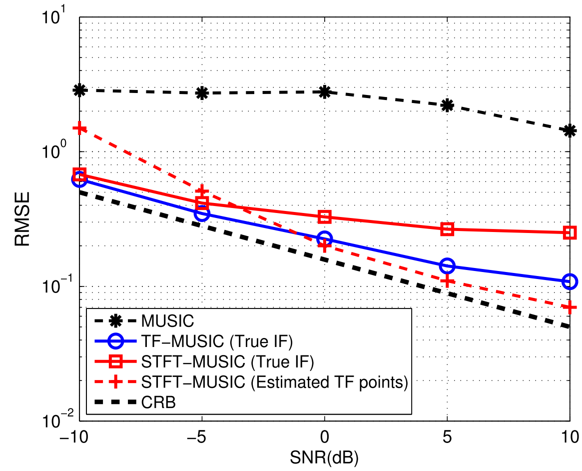




(a) Detected single-source TF points of two spectrally-overlapped LFM sources (DOA spacing: 5°).

(b) Classification results of single-source TF points in (a).

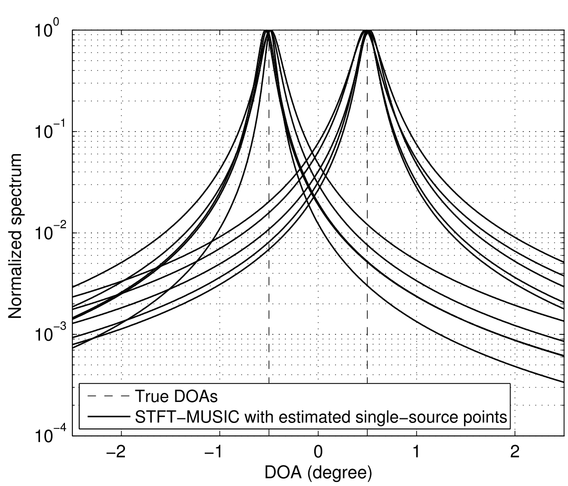


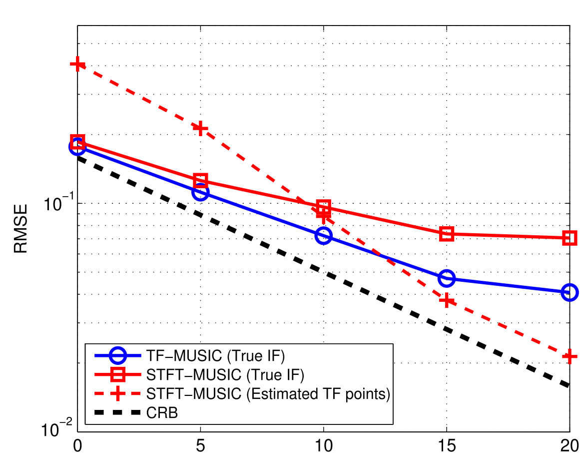


(c) Spatial spectra of two sources (DOA spacing: 5°).

(d) RMSE vs. SNR (DOA spacing: 5°).

Fig. 2. Numerical results of DOA estimation of two closely-spaced and spectrally-overlapped LFM sources with $\theta_1 = -2.5°$ and $\theta_2 = 2.5°$. (a) Detected single-source TF points of different LFM sources using the method in Table 1 and (b) their classification results (different colors represent different sources) in TF domain ($M = 5$, SNR = 5 dB, DOA spacing: 5°). (c) Estimated DOAs of different methods ($M = 5$, SNR = 5 dB, DOA spacing: 5°). (d) RMSE versus different SNR levels ($M = 5$, DOA spacing: 5°).

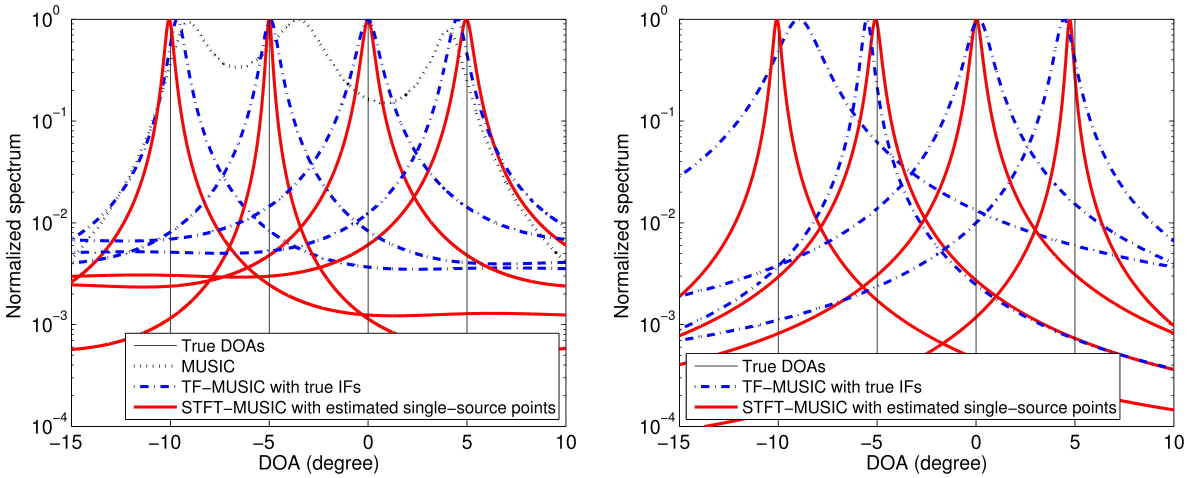




(a) Spatial spectra of two sources (DOA spacing: 1°).

(b) RMSE vs. SNR (DOA spacing: 1°).

Fig. 3. Numerical results of DOA estimation of two closely-spaced and spectrally-overlapped LFM sources with $\theta_1 = -0.5°$ and $\theta_2 = 0.5°$. (a) Estimated DOAs with five situations ($M = 5$, SNR = 15 dB, DOA spacing: 1°). (b) RMSE versus different SNR levels ($M = 5$, DOA spacing: 1°).



(a) Spatial spectra (M=10, N=4).

(b) Spatial spectra (M=3, N=4).

Fig. 4. Numerical results of DOA estimation of four closely-spaced and spectrally-overlapped sources with $\theta_1 = -10°$, $\theta_2 = -5°$, $\theta_3 = 0°$ and $\theta_4 = 5°$. (a) Estimated DOAs of different methods ($M = 10$, SNR $= 5$ dB). (b) Estimated DOAs in underdetermined case ($M = 3$, SNR $= 15$ dB).

we reduce the number of sensors into $M = 3$. The discrimination of TF points for different sources indicates that the DOA estimation can be processed for individual sources. The spatial spectra by the STFT-MUSIC as well as the TF-MUSIC with true IFs are shown in Fig. 4 (b). We observe that the detected single-source points and the separation of different sources well guarantee the estimation accuracy, while the estimation bias becomes worse in the underdetermined case when true IFs are used.

Although the performance of the STFT-MUSIC degrades at low SNRs compared to that of the TF-MUSIC with true IFs (see Fig. 2 (d)), accurate IF information is unavailable in practice. Furthermore, the IF estimation is a challenging issue in low SNR environments especially for spectrally-overlapped sources. Instead of detecting the signal signature along the IF trajectory, we try to identify the TF points associated with single sources for DOA estimation. This implies that it is unnecessary to put a high requirement on the quality and resolution of used TFDs.

From Fig. 1 to Fig. 4, the proposed method is compared with the MUSIC-based methods. Next, the comparison between the proposed method and another joint estimation method reported in [48] is presented in Fig. 5. In [48], a subspace-based L1-SVD estimation method based on a sparse representation of sensor measurements was presented. The L1-SVD method exhibits some advantages over several existing DOA methods, including MUSIC method. The numerical results of the MUSIC and L1-SVD methods with different DOA spacing between each source are presented in Fig. 5 (a) to (c) at SNR $= 5$ dB. Note that these two methods can both achieve desirable estimation performance when DOA spacing equals to 10°. Compared to MUSIC, the L1-SVD obtains a sharp estimate of the spatial spectrum that exhibits super-resolution. When the DOA spacing is reduced from 10° to 8°, the two methods begin to become sensitive to the small spacing especially for the MUSIC method. When the sources get very close, i.e., DOA spacing is reduced from 8° to 5°, the two methods cannot successfully resolve closely-spaced sources. It is worth noting that, for both the above methods, the number of array sensors must be larger than the number of sources. In contrast, the proposed method is able to provide accurate estimations for closely-spaced sources, e.g., DOA spacing equals to 5°, as shown in Fig. 5 (d) where $\lambda_0 = 5$ is given.

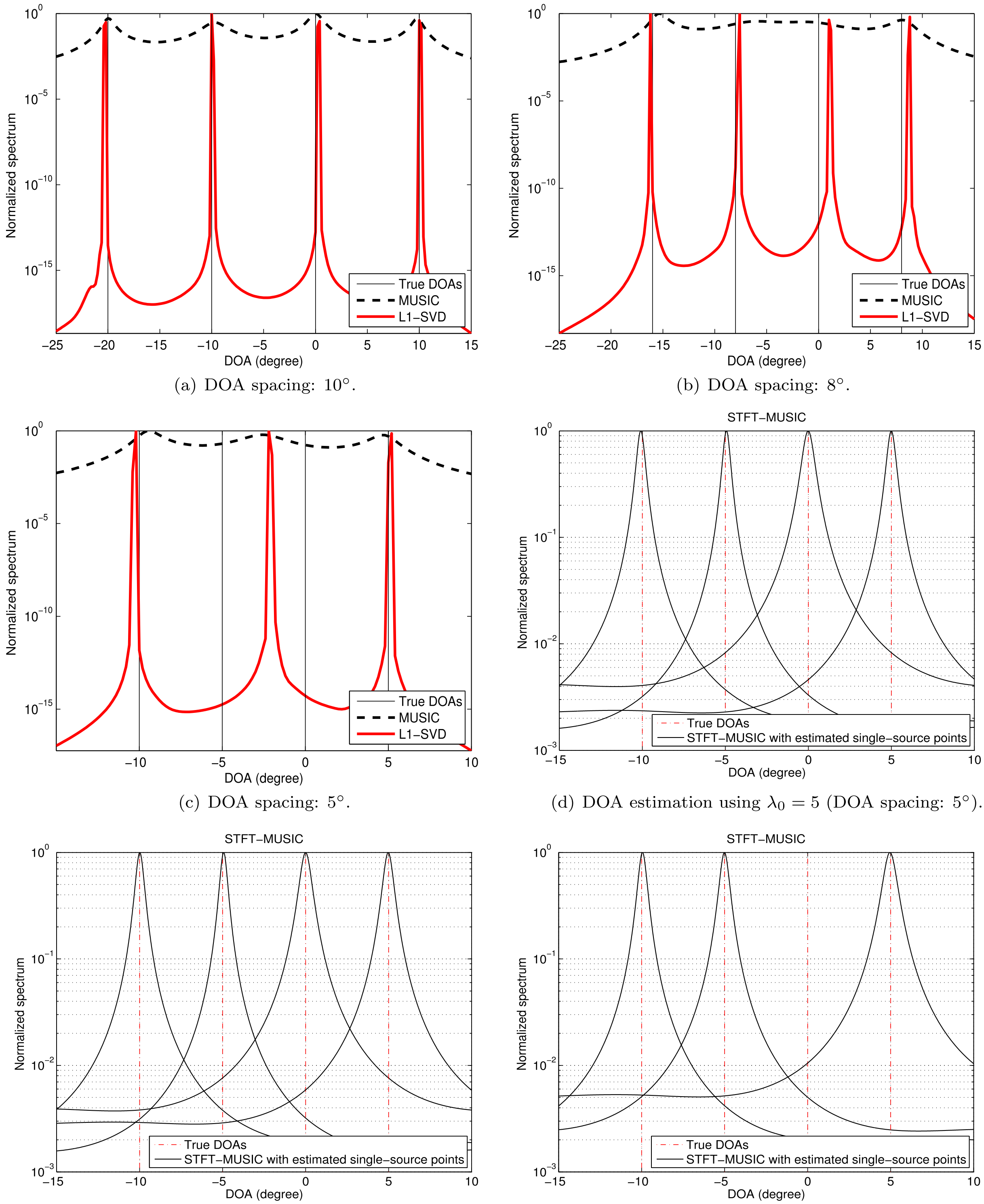
It should be discussed the STFT-MUSIC is not sensitive to the number of single-source points in the signal, i.e., the DOA of one source can be successfully estimated as long as some single-source points (no less than one) corresponding to this source are detected. In the method, the number of single-source points is mainly dependent on the parameter λ_0 in (14). Fig. 5 shows the DOA estimation results in cases of different values of λ_0. Note that the larger the value of λ_0 is, the less the detected single-source TF points are. It should be mentioned that the method with a larger value of λ_0 can achieve a little more accurate DOA estimation because the TF points with larger λ_0 have higher signal-to-noise ratios. However, in practice we cannot choose a very large value of λ_0, because some sources may not have any single-source point which satisfies the condition in (14), see Fig. 5 (f). To guarantee that we can always detect some single-source points for each source, we practically choose an empirical value $\lambda_0 = 5$. Note that we can both achieve accurate DOA estimation results when respectively evaluating $\lambda_0 = 5$ and $\lambda_0 = 10$. The value of l_0 is dependent on the STFT value due to white noise, which should be trivial. Based on the normalized STFT image, we choose an empirical value of $l_0 = 0.5$. Surely we can also choose a smaller or a larger value than 0.5, but different values of l_0 mainly influence the quantity of selected single-source points.

In our study, it is assumed that all sources have the comparative energy. Therefore, the possible limitation of the proposed STFT-MUSIC algorithm is that it cannot well deal with the case where the sources have very different energy. Specifically, the STFT-MUSIC algorithm might fail to detect any single-source TF points for the source with very weak energy, and therefore fails to estimate the DOA information of this source. This limitation can be relaxed by increasing the number of sensors or when a denoising preprocessing is implemented.

5. Discussion

In [45], the complex-valued mixing matrix of sources with overlapped (weak-sparseness) spectral contents was estimated by clustering the single-source TF points (i.e. the TF points associated with a single source), which are detected by selecting the TF points having sufficient energy (only the step 1 in Table 1). However, when the number of sources or the observation time increases, more multi-source TF points possessing strong energy will appear, which significantly influences the estimation accuracy of the mixing matrix. Therefore, the estimation performance of single-source TF points in [45] is limited by the cases where the spectral contents of sources are highly overlapped in TF domain.

The study in this paper is an extension of the method in [45]. The novelty of this paper lies in that the proposed method, which



(a) DOA spacing: 10°.

(b) DOA spacing: 8°.

(c) DOA spacing: 5°.

(d) DOA estimation using $\lambda_0 = 5$ (DOA spacing: 5°).

(e) DOA estimation using $\lambda_0 = 10$ (DOA spacing: 5°).

(f) DOA estimation using $\lambda_0 = 25$ (DOA spacing: 5°).

Fig. 5. Comparison of DOA estimation results of MUSIC and the L1-SVD method in [48] with the proposed STFT-MUSIC algorithm in cases of different values of λ_0 ($N = 4$, $M = 10$, SNR $= 5$ dB).

is of practical importance, is especially suitable for the sources which are significantly overlapped in STFT domain. Specifically, we propose an estimation method of single-source points based on the detection of TF points with dominant energy (four steps shown in Table 1), which are demonstrated to be appropriate candidates for accurate estimation of DOA. Although the detection of single-source TF points based on bilinear TFDs has been investigated in the literature, the detection of dominant TF points based on linear STFT is firstly addressed in our study for DOA estimation.

Another limitation of the method in [45] lies in the optimal selection of ϵ_0 in (7) for different applications. The proposed method avoids the optimal selection of ϵ_0 by detecting single-source TF points from an initial set Ω, which can be obtained by setting a relatively small value of ϵ_0. Although the single-source TF points

Table 2
Comparison of computational complexity.

STFT-MUSIC	TF-MUSIC
$O(MN_s^2 log_2 N_s)$	$O(M^2 N_s^2 log_2 N_s)$

are detected by evaluating a value of λ_0 in (14), however, final estimation performance is not very sensitive to the parameter λ_0, and it can allow an evaluation of wide dynamic range, as shown in Fig. 5.

The repeated implementation of the STFD-based MUSIC for each source inevitably increases the computation load. The computation of $M \times M$ bilinear TFDs is required in the TF-MUSIC algorithm for each source, whereas the STFT-MUSIC algorithm can be implemented by only M linear TFDs, which substantially relaxes the complexity of practical implementation. Consequently, the TF-MUSIC is implemented with a computational complexity $O(M^2 N_s^2 log_2 N_s)$, whereas the STFT-MUSIC is implemented with a computational complexity $O(MN_s^2 log_2 N_s)$, where N_s denotes the number of samples (see Table 2).

6. Conclusion

Traditional MUSIC algorithm cannot work for closely-spaced sources or in an underdetermined case, while the bilinear TF-MUSIC algorithm has the implementation issue and the difficulty to select auto-source TF points. To relax these limitations, this paper has developed a TF-MUSIC algorithm by using spatial STFT matrices. Two practical issues are solved. Firstly, the implementation complexity is relaxed due to the low implementation cost of STFT. Secondly, a selection method of single-source TF points in case of complex-valued mixing matrix is proposed. Numerical results show the efficiency and the robustness of the proposed STFT-MUSIC algorithm especially when the sources highly intersect in TF domain and are spatially close in space domain. The potential applications of the STFT-MUSIC algorithm include the DOA estimation of communication, radar, sonar and speech signals.

Appendix. The explanation of the linear combination in Equation (9)

It should be explained that the linear combination of the spatial vectors in (9) gives an estimation of the array steering vector of some source if the TF points in the same cluster belong to this source. Specifically, the spatial vector at a single-source TF point (t, f) is approximately equal to the steering vector of this source, because the expression (2) in the presence of noise is reduced to

$$\mathbf{S_x}(t,f) = \begin{bmatrix} \mathbf{S}_{x_1}(t,f) \\ \vdots \\ \mathbf{S}_{x_m}(t,f) \\ \vdots \\ \mathbf{S}_{x_M}(t,f) \end{bmatrix} = \begin{bmatrix} a_{i1} \\ \vdots \\ a_{im} \\ \vdots \\ a_{iM} \end{bmatrix} \mathbf{S}_{s_i}(t,f) + \begin{bmatrix} \mathbf{S}_{n_1}(t,f) \\ \vdots \\ \mathbf{S}_{n_m}(t,f) \\ \vdots \\ \mathbf{S}_{n_M}(t,f) \end{bmatrix}, \quad i = 1, \dots, N, \tag{16}$$

when only the ith source exists at this TF point, where $\mathbf{S}_{s_i}(t, f)$ denotes the STFT value of the ith source at the point (t, f), and $\mathbf{S}_{n_m}$ denotes the STFT value of the noise at the mth sensor. It is assumed that the noises, which have a common variance at all sensors, are uncorrelated among the M sensors. For a single-source TF point, its normalized spatial vector in (8) becomes

$$\mathbf{v}(t,f) = \begin{bmatrix} \frac{a_{i1}\mathbf{S}_{s_i}(t,f)+\mathbf{S}_{n_1}(t,f)}{||\mathbf{S_x}(t,f)||} \\ \vdots \\ \frac{a_{im}\mathbf{S}_{s_i}(t,f)+\mathbf{S}_{n_m}(t,f)}{||\mathbf{S_x}(t,f)||} \\ \vdots \\ \frac{a_{iM}\mathbf{S}_{s_i}(t,f)+\mathbf{S}_{n_M}(t,f))}{||\mathbf{S_x}(t,f)||} \end{bmatrix} \cdot \frac{||a_{i1}\mathbf{S}_{s_i}(t,f)+\mathbf{S}_{n_1}(t,f)||}{a_{i1}\mathbf{S}_{s_i}(t,f)+\mathbf{S}_{n_1}(t,f)}, \quad i = 1, \dots, N. \tag{17}$$

It is seen that the computed spatial vector in (17) is influenced by the white noise. For the TF points with strong source energy and negligible noise energy, the computed spatial vectors in (17) are approximately equal to the ideal steering vector. For the TF points where the noise energy is significant, the computed spatial vectors in (17) deviate from the ideal steering vector, and they are randomly distributed around the ideal steering vector because of the random property of noise. Therefore, the centroid of the computed spatial vectors belonging to the same cluster gives an estimation of the ideal steering vector. Estimating the steering vectors of sources by the linear combination of the normalized spatial vectors as in (8) and (9) has already been applied in the literature [44,45,49].

References

[1] H. Krim, M. Viberg, Two decades of array signal processing research: the parametric approach, IEEE Signal Process. Mag. 13 (4) (1996) 67–94.
[2] J. Steinwandt, R.C. de Lamare, M. Haardt, Beamspace direction finding based on the conjugate gradient and the auxiliary vector filtering algorithms, Signal Process. 93 (4) (2013) 641–651.
[3] W. Jiao, S. Yang, Y. Chang, DOA estimation of multiple convolutively mixed sources based on principle component analysis, in: Neural Information Processing, in: Lecture Notes in Computer Science, vol. 5863, 2009, pp. 340–348.
[4] J. Jensen, M. Christensen, S. Jensen, Nonlinear least squares methods for joint DOA and pitch estimation, IEEE Trans. Audio Speech Lang. Process. 21 (5) (2013) 923–933.
[5] Y. Wu, L. Amir, J. Jensen, G. Liao, Joint pitch and DOA estimation using the ESPRIT method, IEEE Trans. Audio Speech Lang. Process. 23 (1) (2015) 32–45.
[6] L. Wang, L. Zhao, G. Bi, C. Wan, L. Zhang, H. Zhang, Novel wideband DOA estimation based on sparse Bayesian learning with Dirichlet process priors, IEEE Trans. Signal Process. 64 (2) (2016) 275–289.
[7] G. Bienvenu, L. Kopp, Adaptivity to background noise spatial coherence for high resolution passive methods, in: IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '80, vol. 5, 1980, pp. 307–310.
[8] R.O. Schmidt, A signal subspace approach to multiple emitter location and spectral estimation, Thesis (Ph. D.), Stanford University, 1981.
[9] R.O. Schmidt, Multiple emitter location and signal parameter estimation, IEEE Trans. Antennas Propag. 34 (3) (1986) 276–280.
[10] A. Swindlehurst, T. Kailath, A performance analysis of subspace-based methods in the presence of model errors. I. The MUSIC algorithm, IEEE Trans. Signal Process. 40 (7) (1992) 1758–1774.
[11] X.L. Xu, K. Buckley, Bias analysis of the MUSIC location estimator, IEEE Trans. Signal Process. 40 (10) (1992) 2559–2569.
[12] A. Ferreol, P. Larzabal, M. Viberg, On the asymptotic performance analysis of subspace DOA estimation in the presence of modeling errors: case of MUSIC, IEEE Trans. Signal Process. 54 (3) (2006) 907–920.
[13] A.J. Barabell, Improving the resolution performance of eigenstructure-based direction-finding algorithms, in: IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '83, vol. 8, 1983, pp. 336–339.
[14] B. Rao, K. Hari, Performance analysis of root-MUSIC, IEEE Trans. Acoust. Speech Signal Process. 37 (12) (1989) 1939–1949.
[15] H. Krim, P. Forster, J. Proakis, Operator approach to performance analysis of root-MUSIC and root-min-norm, IEEE Trans. Signal Process. 40 (7) (1992) 1687–1696.
[16] M. Zoltowski, G. Kautz, S. Silverstein, Beamspace root-MUSIC, IEEE Trans. Signal Process. 41 (1) (1993) 344–364.
[17] R. Goossens, H. Rogier, S. Werbrouck, UCA root-MUSIC with sparse uniform circular arrays, IEEE Trans. Signal Process. 56 (8) (2008) 4095–4099.
[18] C. El Kassis, J. Picheral, C. Mokbel, Advantages of nonuniform arrays using root-MUSIC, Signal Process. 90 (2) (2010) 689–695.
[19] A. Belouchrani, M. Amin, Blind source separation based on time-frequency signal representations, IEEE Trans. Signal Process. 46 (11) (1998) 2888–2897.
[20] W. Sharif, M. Muma, A. Zoubir, Robustness analysis of spatial time-frequency distributions based on the influence function, IEEE Trans. Signal Process. 61 (8) (2013) 1958–1971.

[21] L. Stanković, I. Djurović, S. Stanković, M. Simeunović, S. Djukanović, M. Daković, Instantaneous frequency in time–frequency analysis: enhanced concepts and performance of estimation algorithms, Digit. Signal Process. 35 (2014) 1–13.
[22] B. Boashash, N.A. Khan, T. Ben-Jabeur, Time-frequency features for pattern recognition using high-resolution TFDs: a tutorial review, Digit. Signal Process. 40 (2015) 1–30.
[23] H. Zhang, G. Bi, S.G. Razul, C.M.S. See, Robust time-varying filtering and separation of some nonstationary signals in low SNR environments, Signal Process. 106 (2015) 141–158.
[24] H. Zhang, L. Yu, G.-S. Xia, Iterative time-frequency filtering of sinusoidal signals with updated frequency estimation, IEEE Signal Process. Lett. 23 (1) (2016) 139–143.
[25] A. Belouchrani, M. Amin, Time-frequency MUSIC, IEEE Signal Process. Lett. 6 (5) (1999) 109–110.
[26] M. Amin, Y. Zhang, Direction finding based on spatial time-frequency distribution matrices, Digit. Signal Process. 10 (4) (2000) 325–339.
[27] Y. Zhang, W. Mu, M. Amin, Subspace analysis of spatial time-frequency distribution matrices, IEEE Trans. Signal Process. 49 (4) (2001) 747–759.
[28] M. Khodja, A. Belouchrani, K. Abed-Meraim, Performance analysis for time-frequency MUSIC algorithm in presence of both additive noise and array calibration errors, EURASIP J. Adv. Signal Process. (2012), http://dx.doi.org/10.1186/1687-6180-2012-94.
[29] A. Belouchrani, M. Amin, N. Thirion-Moreau, Y. Zhang, Source separation and localization using time-frequency distributions: an overview, IEEE Signal Process. Mag. 30 (6) (2013) 97–107.
[30] M. Amin, Y. Zhang, DOA Estimation of Nonstationary Signals, Academic Press Library in Signal Processing, vol. 3, Academic Press, 2013, pp. 1–30, chap. 17.
[31] S. Ghofrani, M.G. Amin, Y.D. Zhang, High-resolution direction finding of non-stationary signals using matching pursuit, Signal Process. 93 (12) (2013) 3466–3478.
[32] S. Ghofrani, Matching pursuit for direction of arrival estimation in the presence of Gaussian noise and impulsive noise, IET Signal Process. 8 (5) (2014) 540–551.
[33] A. Gershman, M. Amin, Wideband direction-of-arrival estimation of multiple chirp signals using spatial time-frequency distributions, IEEE Signal Process. Lett. 7 (6) (2000) 152–155.
[34] H. Zhang, G. Bi, W. Yang, S.G. Razul, C.M.S. See, IF estimation of FM signals based on time-frequency image, IEEE Trans. Aerosp. Electron. Syst. 51 (1) (2015) 326–343.
[35] P. Heidenreich, L. Cirillo, A. Zoubir, Morphological image processing for FM source detection and localization, Signal Process. 89 (6) (2009) 1070–1080.
[36] Y. Zhang, M.G. Amin, Blind separation of nonstationary sources based on spatial time-frequency distributions, EURASIP J. Appl. Signal Process. 2006 (2006) 64785.
[37] E. Fadaili, N. Moreau, E. Moreau, Nonorthogonal joint diagonalization/zero diagonalization for source separation based on time-frequency distributions, IEEE Trans. Signal Process. 55 (5) (2007) 1673–1687.
[38] H. Qu, R. Wang, W. Qu, P. Zhao, Research on DOA estimation of multi-component LFM signals based on the FRFT, Wirel. Sens. Netw. 1 (3) (2009) 171–181.
[39] Y. Zhang, M. Amin, B. Himed, Direction-of-arrival estimation of nonstationary signals exploiting signal characteristics, in: 2012 11th International Conference on Information Science, Signal Processing and Their Applications, ISSPA, 2012, pp. 1223–1228.
[40] A. Jourjine, S. Rickard, O. Yilmaz, Blind separation of disjoint orthogonal signals: demixing N sources from 2 mixtures, in: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 5, 2000, pp. 2985–2988.
[41] O. Yilmaz, S. Rickard, Blind separation of speech mixtures via time-frequency masking, IEEE Trans. Signal Process. 52 (7) (2004) 1830–1847.
[42] F. Abrard, Y. Deville, A time-frequency blind signal separation method applicable to underdetermined mixtures of dependent sources, Signal Process. 85 (7) (2005) 1389–1403.
[43] Y. Li, S. Amari, A. Cichocki, D. Ho, S. Xie, Underdetermined blind source separation based on sparse representation, IEEE Trans. Signal Process. 54 (2) (2006) 423–437.
[44] S. Kim, C. Yoo, Underdetermined blind source separation based on subspace representation, IEEE Trans. Signal Process. 57 (7) (2009) 2604–2614.
[45] A. Aïssa-El-Bey, N. Linh-Trung, K. Abed-Meraim, A. Belouchrani, Y. Grenier, Underdetermined blind separation of nondisjoint sources in the time-frequency domain, IEEE Trans. Signal Process. 55 (3) (2007) 897–907.
[46] P. Stoica, A. Nehorai, MUSIC, maximum likelihood, and Cramer–Rao bound, IEEE Trans. Acoust. Speech Signal Process. 37 (5) (1989) 720–741.
[47] P. Stoica, A. Nehorai, MUSIC, maximum likelihood, and Cramer–Rao bound: further results and comparisons, IEEE Trans. Acoust. Speech Signal Process. 38 (12) (1990) 2140–2150.
[48] D. Malioutov, M. Cetin, A. Willsky, A sparse signal reconstruction perspective for source localization with sensor arrays, IEEE Trans. Signal Process. 53 (8) (2005) 3010–3022.
[49] N. Linh-Trung, A. Belouchrani, K. Abed-Meraim, B. Boashash, Separating more sources than sensors using time-frequency distributions, EURASIP J. Appl. Signal Process. 2005 (17) (2005) 2828–2847.

Haijian Zhang received the B.Sc. degree in electronic information engineering from Wuhan University, Wuhan, China, in 2006 and the joint Ph.D. degrees from Conservatoire National des Arts et Metiers (CNAM), Paris, France, and Wuhan University, in 2011. From 2011 to 2014, he was a research fellow with the school of electrical electronic engineering, Nanyang Technological University, Singapore. He is currently an Associate Professor with school of electronic information, Wuhan University, China. His main research interests include time-frequency analysis, array signal processing, Compressing sensing, and cognitive radio.

Guoan Bi received the B.Sc. degree in radio communications from Dalian University of Technology, P. R. China, 1982, and the M.Sc. degree in telecommunication systems, and the Ph.D. degree in electronics systems, both from Essex University, U.K., in 1985 and 1988, respectively. Since 1991, he has been with the school of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His current research interests include DSP algorithms and hardware structures, and signal processing for various applications including sonar, radar, and communications.

Yunlong Cai received the B.S. degree in computer science from Beijing Jiaotong University, Beijing, China, in 2004, the M.Sc. degree in electronic engineering from the University of Surrey, Guildford, U.K., in 2006, and the Ph.D. degree in electronic engineering from the University of York, York, U.K., in 2010. From February 2010 to January 2011, he was a Postdoctoral Fellow at the Electronics and Communications Laboratory of the Conservatoire National des Arts et Metiers (CNAM), Paris, France. Since February 2011, he has been with the Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China, where he is currently an Associate Professor. His research interests include spread spectrum communications, adaptive signal processing, multiuser detection, and multiple antenna systems. He has served on the Technical Program Committee (TPC) as member in various international conferences including for example IEEE GLOBECOM, IEEE ISWCS, and ITG WSA. He has also served as the Publicity Co-Chair for the fifth International Conference on Wireless Communications and Signal Processing (WCSP).

Sirajudeen Gulam Razul received his B.Eng and M.Eng in Electrical and Electronics Engineering from the Nanyang Technological University, Singapore in 1997 and 2000 respectively. He received his Ph.D from the University of Cambridge, UK in 2003. His Ph.D work was on Bayesian estimation using Markov Chain Monte Carlo techniques. He was a faculty member in the School of EEE, Nanyang Technological University from 2004 to 2009. He is currently a Senior Research Scientist with TL@NTU. His research interests include Bayesian signal processing, statistical signal processing and array processing.

Chong Meng Samson See received the Diploma in Electronics and Communications Engineering (with merit) from Singapore Polytechnic in 1988 and the M.Sc. in Digital Communication Systems and the Ph.D in Electrical Engineering both from Loughborough University of Technology, Loughborough, U.K., in 1991 and 1999, respectively. Since 1992, he has been with DSO National Laboratories, Singapore, where he is now a Distinguished Member of Technical Staff and is currently leading a team in the research and development of advanced array signal processing systems and algorithms. He also holds an adjunct appointment at TL@NTU as a Principal Research Scientist where he is Capability Manager on Sensor Array Research. His research interests include the area of statistical and array signal processing, communications, and bio-inspired systems. He has two issued patents on direction finding. Samson is a Member of IEEE and was Associate Editor of IEEE Transactions on Signal Processing and currently member of the IEEE Sensor Array and Multichannel Signal Processing Technical Committee.