

# Blind Separation of Nonstationary Sources Based on Spatial Time-Frequency Distributions

Yimin Zhang and Moeness G. Amin

*Wireless Communications and Positioning Lab, Center for Advanced Communications, Villanova University, Villanova, PA 19085, USA*

Received 1 January 2006; Revised 24 July 2006; Accepted 13 August 2006

Blind source separation (BSS) based on spatial time-frequency distributions (STFDs) provides improved performance over blind source separation methods based on second-order statistics, when dealing with signals that are localized in the time-frequency (t-f) domain. In this paper, we propose the use of STFD matrices for both whitening and recovery of the mixing matrix, which are two stages commonly required in many BSS methods, to provide robust BSS performance to noise. In addition, a simple method is proposed to select the auto- and cross-term regions of time-frequency distribution (TFD). To further improve the BSS performance, t-f grouping techniques are introduced to reduce the number of signals under consideration, and to allow the receiver array to separate more sources than the number of array sensors, provided that the sources have disjoint t-f signatures. With the use of one or more techniques proposed in this paper, improved performance of blind separation of nonstationary signals can be achieved.

Copyright © 2006 Y. Zhang and M. G. Amin. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. INTRODUCTION

Several methods have been proposed to blindly separate independent narrowband sources [1–8]. When the spatial (mixing) signatures of the sources are not orthogonal, blind source separation (BSS) methods usually employ at least two different sets of matrices that span the same signal subspace. One set is used for whitening purpose, whereas the other set is used to estimate rotation ambiguity so that the spatial signatures and the source waveforms impinging on a multiantenna receiver can be recovered. Different methods have been developed for blind source separation based on cyclostationarity, spectral or/and higher-order statistics of the source signals, linear and quadrature time-frequency (t-f) transforms.

In this paper, we focus on the blind separation of nonstationary sources that are highly localized in the t-f domain (e.g., frequency modulated (FM) waveforms). Such signals are frequently encountered in radar, sonar, and acoustic applications [9–11]. For this kind of nonstationary signals, quadrature time-frequency distributions (TFDs) have been employed for array processing and have been found successful in blind source separations [12–16]. Among the existing methods, typically, the spatial time-frequency distribution (STFD) matrices are used for source diagonalization and

antidiagonalization, whereas the whitening matrix remains the signal covariance matrix. The STFD matrices are constructed from the auto-TFDs and cross-TFDs of the sensor data and evaluated at different points of high signal-to-noise ratio (SNR) pertaining to the t-f signatures of the sources. Joint diagonalization, antidiagonalization, or a combination of both techniques can be applied, depending on t-f point selections and the structure of the source TFD matrices.

Existing methods, however, only apply STFD matrices to recover the data from the unitary mixture, while the covariance matrices are still used in the whitening process. Therefore, the inherent advantages of STFD, for example, SNR enhancement and source discrimination, are not fully utilized. In particular, for an underdetermined problem where the number of sources is larger than the number of array sensors, signal whitening using the covariance matrix becomes inappropriate and impractical.

Several different approaches have been proposed to recover nonstationary signal waveforms based on t-f masking followed by signal waveform synthesis or inverse t-f transformations. In [17], the TFD is first averaged over different array sensors to identify the autoterm region. This across-sensor averaging provides significant reduction of the cross-term TFDs. The sources are separated in the t-f domain from the autoterms only using a vector classification approach.

In [18], the waveform of each source signal is synthesized from its t-f signature averaged over multiple array sensors. By applying appropriate t-f masking to the averaged t-f signatures, the autoterm of each source signal can be independently extracted, and the corresponding waveform can be synthesized. The approach presented in [19] considered t-f masking on linear t-f distributions (e.g., short-time Fourier transform and Gabor expansions) to separate signals with disjoint t-f signatures. Because the TFD is linear, waveform recovery is relatively simple compared to synthesis of bilinear TFDs. There are also some BSS methods that use nonorthogonal joint diagonalization procedure to eliminate the whitening process [20, 21]. More detailed information about BSS can be found in books and survey papers (e.g., [6–8]).

In this paper, we propose a source separation technique that employs STFDs for both phases of whitening and unitary matrix recovery. In essence, instead of using the covariance matrix for signal whitening, we apply multiple STFD matrices over the source t-f signatures, incorporating the t-f localization properties of the sources in both the whitening and joint estimation steps of source separation. The proposed method leads to noise robustness of subspace decompositions and, thereby, enhances the unitary mixture representations of the problem. When the number of array sensors is larger than the number of sources, t-f masking is optional. If it is possible to separate the impinging sources into several disjoint groups in the t-f domain, then t-f masking can be used to improve the source separation performance by allowing the selection of subsets of sources. As such, t-f masking allows the proposed technique to accurately estimate the spatial signatures and synthesize signal waveforms in the presence of high number of sources which exceeds the number of array sensors. These situations are often referred to as the underdetermined blind source separation problems and have been considered in [22–24].

Another important contribution of this paper is to propose a new method for selecting autoterm t-f points. Autoterm point selection is key in maintaining the diagonal structure of the source TFD matrix which is the fundamental assumption of source separation via diagonalization. The proposed method only requires the calculation of the autoterms of the whitened STFD matrix. It is simpler and more effective than the methods developed in [14, 25] which require the calculation of either the norm or the eigenvalues of the whitened STFD matrices and, therefore, rely on both auto- and cross-terms of whitened matrix elements. With effective autoterm selections, sources in the field of view can be disallowed from consideration by the receiver, leading to improved subspace estimation. This paper also discusses the selection of cross-terms.

This paper is organized as follows. Section 2 introduces the signal model and briefly reviews STFD and the STFD-based blind source separation methods [12–14]. In Section 3, the new methods for auto- and cross-term t-f point selection are addressed. Section 4 introduces the idea of t-f grouping and proposes the use of STFD whitening matrix in the source separation. Section 5 considers the scenarios where the num-

ber of source signals is larger than the number of array sensors. Simulation results are presented in Section 6.

## 2. BLIND SOURCE SEPARATION BASED ON SPATIAL TIME-FREQUENCY SIGNATURES

### 2.1. Signal model

In narrowband array processing, when  $n$  signals arrive at an  $m$ -element array, the linear data model

$$\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{d}(t) + \mathbf{n}(t) \quad (1)$$

is commonly used, where  $\mathbf{A}$  is the mixing matrix of dimension  $m \times n$  and is assumed to be full column rank,  $\mathbf{x}(t) = [x_1(t), \dots, x_m(t)]^T$  is the sensor array output vector, and  $\mathbf{d}(t) = [d_1(t), \dots, d_n(t)]^T$  is the source signal vector, where the superscript  $T$  denotes the transpose operator.  $\mathbf{n}(t)$  is an additive noise vector whose elements are modelled as stationary, spatially, and temporally white, zero-mean complex random processes, independent of the source signals.

The source signals in this paper are assumed to be deterministic nonstationary signals which are highly localized in the time-frequency domain. In the original source separation method proposed in [12], the source signals are assumed uncorrelated and their respective autoterms are free from cross-term contamination. In the proposed modification, only the second condition is required. In addition, if the t-f signatures of the sources are amendable to disjoint grouping, then it is possible to separate more sources than the number of array sensors, that is, the full column rank requirement of the mixing matrix  $\mathbf{A}$  is no longer necessary.

### 2.2. Spatial time-frequency distributions

The discrete form of Cohen's class of STFD of the data snapshot vector  $\mathbf{x}(t)$  is given by [12],

$$\mathbf{D}_{\mathbf{xx}}(t, f) = \sum_{l=-\infty}^{\infty} \sum_{\tau=-\infty}^{\infty} \phi(l, \tau) \mathbf{x}(t+l+\tau) \mathbf{x}^H(t+l-\tau) e^{-j4\pi f \tau}, \quad (2)$$

where  $\phi(l, \tau)$  is a t-f kernel and the superscript  $H$  denotes conjugate transpose. Substituting (1) into (2), we obtain

$$\mathbf{D}_{\mathbf{xx}}(t, f) = \mathbf{D}_{\mathbf{yy}}(t, f) + \mathbf{D}_{\mathbf{yn}}(t, f) + \mathbf{D}_{\mathbf{ny}}(t, f) + \mathbf{D}_{\mathbf{nn}}(t, f). \quad (3)$$

Under the uncorrelated signal and noise assumption and the zero-mean noise property,  $E[\mathbf{D}_{\mathbf{yn}}(t, f)] = E[\mathbf{D}_{\mathbf{ny}}(t, f)] = \mathbf{0}$ . It follows

$$\begin{aligned} E[\mathbf{D}_{\mathbf{xx}}(t, f)] &= \mathbf{D}_{\mathbf{yy}}(t, f) + E[\mathbf{D}_{\mathbf{nn}}(t, f)] \\ &= \mathbf{A}\mathbf{D}_{\mathbf{dd}}(t, f)\mathbf{A}^H + E[\mathbf{D}_{\mathbf{nn}}(t, f)]. \end{aligned} \quad (4)$$

Similar to the well-known and commonly used mathematical formula (see (6)), which relates the signal covariance matrix to the data spatial covariance matrix, (4) provides the

basis for source separation by relating the STFD matrix to the source TFD matrix,  $\mathbf{D}_{dd}(t, f)$ , through the mixing matrix  $\mathbf{A}$ .

It was analytically shown in [26] that, when the STFD matrices are constructed using the autoterm points with localized signal energy, the estimated subspace based on these matrices is more robust to noise perturbation than that obtained from the covariance matrices because of the enhancement of the signal power. Such advantage is particularly useful when the noise effect is large, and it becomes more attractive when dealing with fewer selected sources. These facts apply to the performance of blind source separation as the performance is directly related to the robustness of the estimated signal subspace.

### 2.3. Blind source separation

In the STFD-based blind source separation method proposed in [12], the following data covariance matrix is used for prewhitening:

$$\mathbf{R}_{xx} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbf{x}(t)\mathbf{x}^H(t). \quad (5)$$

Under the assumption that the source signals are uncorrelated to the noise, we have

$$\mathbf{R}_{xx} = \mathbf{R}_{yy} + \sigma \mathbf{I} = \mathbf{A}\mathbf{R}_{dd}\mathbf{A}^H + \sigma \mathbf{I}, \quad (6)$$

where  $\mathbf{R}_{dd} = \lim_{T \rightarrow \infty} (1/T) \sum_{t=1}^T \mathbf{d}(t)\mathbf{d}^H(t)$  is the source correlation matrix which is assumed diagonal,  $\sigma$  is the noise power at each sensor, and  $\mathbf{I}$  denotes the identity matrix. It is assumed that  $\mathbf{R}_{xx}$  is nonsingular, and the observation period consists of  $N$  snapshots with  $N > m$ .

In blind source separation techniques, there is an ambiguity with respect to the order and the complex amplitude of the sources. It is convenient to assume that each source has unit norm, that is,  $\mathbf{R}_{dd} = \mathbf{I}$ .

The first step in TFD-based blind source separations is whitening (orthogonalization) of the signal  $\mathbf{x}(t)$  of the observation. This is achieved by estimating the noise power<sup>1</sup> and applying a whitening matrix  $\mathbf{W}$  to  $\mathbf{x}(t)$ , that is, an  $n \times m$  matrix satisfying

$$\mathbf{W}\mathbf{R}_{yy}\mathbf{W}^H = \mathbf{W}(\mathbf{R}_{xx} - \sigma \mathbf{I})\mathbf{W}^H = \mathbf{W}\mathbf{A}\mathbf{A}^H\mathbf{W}^H = \mathbf{I}. \quad (7)$$

The whitening matrix is estimated using the signal subspace obtained from the eigendecomposition of  $\mathbf{R}_{xx}$  [12]. Let  $\lambda_i$  denote the  $i$ th descendingly sorted eigenvalue of  $\mathbf{R}_{xx}$  and  $\mathbf{q}_i$  the corresponding eigenvector. Then, the  $i$ th row of the whitening matrix is obtained as

$$\mathbf{w}_i = (\lambda_i - \sigma)^{-1/2} \mathbf{q}_i^H, \quad 1 \leq i \leq m. \quad (8)$$

It is clear that the accuracy of the whitening matrix estimate depends on the estimation accuracy of the eigenvectors and eigenvalues corresponding to the signal subspace. The whitened process  $\mathbf{z}(t) = \mathbf{W}\mathbf{x}(t)$  still obeys a linear model:

$$\mathbf{z}(t) = \mathbf{W}\mathbf{x}(t) = \mathbf{W}\mathbf{A}\mathbf{d}(t) + \mathbf{W}\mathbf{n}(t) = \mathbf{U}\mathbf{d}(t) + \mathbf{W}\mathbf{n}(t), \quad (9)$$

where  $\mathbf{U} \triangleq \mathbf{W}\mathbf{A}$  is an  $n \times n$  unitary matrix.

The next step is to estimate the unitary matrix  $\mathbf{U}$ . The whitened STFD matrices in the noise-free case can be written as

$$\mathbf{D}_{zz}(t, f) = \mathbf{W}\mathbf{D}_{xx}(t, f)\mathbf{W}^H = \mathbf{U}\mathbf{D}_{dd}(t, f)\mathbf{U}^H. \quad (10)$$

In the autoterm regions,  $\mathbf{D}_{dd}(t, f)$  is diagonal, and an estimate  $\hat{\mathbf{U}}$  of the unitary matrix  $\mathbf{U}$  may be obtained as a joint diagonalizer of the set of whitened STFD matrices evaluated at  $K$  autoterm t-f points,  $\{\mathbf{D}_{zz}(t_i, f_i) \mid i = 1, \dots, K\}$ . The source signals and the mixing matrix can be, respectively, estimated as  $\hat{\mathbf{d}}(t) = \hat{\mathbf{U}}^H \hat{\mathbf{W}} \mathbf{x}(t)$  and  $\hat{\mathbf{A}} = \hat{\mathbf{W}}^{\#} \hat{\mathbf{U}}$ , where superscript  $\#$  denotes pseudoinverse.

In [13], higher-order TFDs are used to replace the bilinear TFDs used in [12]. In [14], cross-term t-f points were allowed to take part in the separation process by incorporating an antidiagonalization approach. However, the key concept remains the same as that introduced in [12] and summarized above.

## 3. AUTO- AND CROSS-TERM SELECTION

### 3.1. Existing methods

The selection of auto- and cross-term t-f points has been considered in [14, 25, 27]. It is pointed out in [14] that, at the cross-term  $(t, f)$  points, there are no source autoterms, that is,  $\text{trace}(\mathbf{D}_{dd}(t, f)) = 0$ . It was also shown that

$$\begin{aligned} \text{trace}(\mathbf{D}_{zz}(t, f)) &= \text{trace}(\mathbf{U}\mathbf{D}_{dd}(t, f)\mathbf{U}^H) \\ &= \text{trace}(\mathbf{D}_{dd}(t, f)) \approx 0, \quad (t, f) \in \text{cross-term}. \end{aligned} \quad (11)$$

Subsequently, the following testing procedure was proposed:

$$\text{if } \frac{\text{trace}(\mathbf{D}_{zz}(t, f))}{\text{norm}(\mathbf{D}_{zz}(t, f))} < \epsilon \rightarrow \text{decide that } (t, f) \text{ is cross-term,} \\ \text{if } \frac{\text{trace}(\mathbf{D}_{zz}(t, f))}{\text{norm}(\mathbf{D}_{zz}(t, f))} > \epsilon \rightarrow \text{decide that } (t, f) \text{ is autoterm,} \quad (12)$$

where  $\epsilon$  is a small positive real scalar. In [27], single autoterm locations are selected by noting the fact that  $\mathbf{D}_{dd}(t, f)$  is diagonal with only one nonzero diagonal entry. Therefore,  $\mathbf{D}_{zz}(t, f)$  is rank one, and the dominant eigenvalue of  $\mathbf{D}_{zz}(t, f)$  is close to the sum of all eigenvalues.

In calculating the norm or eigenvalues of an STFD matrix in the above two methods, all the auto- and cross-terms of the whitened vector  $\mathbf{z}(t)$  are required. In the following, after reviewing the concept of array averaging, we propose a simple alternative method for auto- and cross-term selection which only requires the autoterm TFDs.

<sup>1</sup> The noise power can be estimated only when  $m > n$  [12]. If  $m = n$ , the estimation of the noise power becomes unavailable and  $\sigma = 0$  will be assumed.

### 3.2. Array averaging

In [18], array average in the context of TFDs is proposed. Averaging of the autosensor TFDs across the array introduces a weighing function in the t-f domain which decreases the noise levels, reduces the interactions of the source signals, and mitigates the cross-terms. This is achieved independent of the temporal characteristics of the source signals and without causing any smearing of the signal terms.

The TFD of the signal received at the  $i$ th array sensor,  $x_i(t) = \sum_{k=1}^n a_{ki}s_k(t)$ , where  $a_{ki}$  is the  $i$ th element of mixing vector  $\mathbf{a}_k$ , is expressed as

$$D_{x_i x_i}(t, f) = \sum_{k=1}^n \sum_{l=1}^n a_{ki} a_{li}^* D_{d_k d_l}(t, f). \quad (13)$$

The averaging of  $D_{x_i x_i}(t, f)$  for  $i = 1, \dots, m$  yields the array averaged TFD of the data vector  $\mathbf{x}(t)$ , defined as [18],

$$\begin{aligned} \bar{D}_{\mathbf{xx}}(t, f) &= \frac{1}{m} \sum_{i=1}^m D_{x_i x_i}(t, f) \\ &= \frac{1}{m} \sum_{i=1}^n \sum_{k=1}^n \mathbf{a}_k^H \mathbf{a}_i D_{d_i d_k}(t, f) \\ &= \sum_{i=1}^n \sum_{k=1}^n \beta_{k,i} D_{d_i d_k}(t, f), \end{aligned} \quad (14)$$

where

$$\beta_{k,i} = \frac{1}{m} \mathbf{a}_k^H \mathbf{a}_i \quad (15)$$

is the spatial correlation between source  $k$  and source  $i$ .

The average of the TFDs over different array sensors is the trace of the corresponding STFD matrix  $\mathbf{D}_{\mathbf{xx}}$ , up to the normalization factor  $m$ . However, with the introduction of the spatial signature between two source signals, it becomes clear that  $\beta_{k,i}$  is equal to unity for the same source signal (i.e.,  $k = i$ , corresponding to the autoterm t-f points), and is smaller than unity for two different source signals (i.e.,  $k \neq i$ , corresponding to the cross-term t-f points). With this fact in mind, it becomes much simpler and more effective to select the threshold for auto- and cross-term selection based on array averaging.

### 3.3. Selection based on unwhitened data

At a pure autosource  $(t, f)$  point, where no cross-source terms are present, the TFD at the  $i$ th sensor is

$$D_{x_i x_i}(t, f) = \sum_{k=1}^n |a_{ki}|^2 D_{d_k d_k}(t, f), \quad (16)$$

which is consistently positive for all values of  $i$ . Accordingly  $D_{x_i x_i}(t, f) = |D_{x_i x_i}(t, f)|$ ,  $i = 1, \dots, m$ . Define the following

criterion:<sup>2</sup>

$$C_x(t, f) = \frac{\sum_{i=1}^m D_{x_i x_i}(t, f)}{\sum_{i=1}^m |D_{x_i x_i}(t, f)|} = \frac{\text{trace}(\mathbf{D}_{\mathbf{xx}}(t, f))}{m \mathcal{D}_{\mathbf{xx}}(t, f)}, \quad (17)$$

where

$$\mathcal{D}_{\mathbf{xx}}(t, f) = \frac{1}{m} \sum_{i=1}^m |D_{x_i x_i}(t, f)| \quad (18)$$

is the averaged absolute value of TFD, referred to as the absolute average TFD at  $(t, f)$  point.

For a pure cross-source t-f point,<sup>3</sup> on the other hand, the TFD is oscillating and it changes its value for different array sensor. Therefore, provided that the spatial correlation between different sources is small, that is,  $\mathbf{a}_k^H \mathbf{a}_i \ll 1$  for  $k \neq i$  in (14), we have  $C_x(t, f) < \alpha_2 \approx 0$ .

When  $C_x(t, f)$  takes a moderate value between  $\alpha_2$  and  $\alpha_1$ , where  $\alpha_2 < \alpha_1$ , the  $(t, f)$  point has both auto- and cross-terms present. Such a point should be avoided in computing the STFD matrix for unitary matrix estimation.

Therefore, the auto- and cross-term points can be identified as

$$\begin{aligned} C_x(t, f) &> \alpha_1 \longrightarrow \text{decide that } (t, f) \text{ is autoterm,} \\ &< \alpha_2 \longrightarrow \text{decide that } (t, f) \text{ is cross-term,} \end{aligned} \quad (19)$$

where we use two different threshold levels for auto- and cross-terms to have more flexibility for different situations. Because  $C_x(t, f)$  is upper bounded, the value of  $\alpha_1$  is usually chosen to be close to unity.

It is important to note that, to avoid the inclusion of noise-only t-f points, selection of meaningful auto- and cross-term points should be limited only among those t-f points where the TFD has certain strength. We use the absolute average TFD to measure the TFD strength. Denote

$$\mathcal{D}_{\mathbf{xx}, \max} = \max_{(t, f)} [\mathcal{D}_{\mathbf{xx}}(t, f)] \quad (20)$$

as the maximum value of the absolute average of TFD, then the selection of meaningful t-f points of certain TFD strength amounts to the following condition:

$$F_x(t, f) = \frac{\mathcal{D}_{\mathbf{xx}}(t, f)}{\mathcal{D}_{\mathbf{xx}, \max}} > \begin{cases} \gamma_1, & \text{for autoterm selection,} \\ \gamma_2, & \text{for cross-term selection,} \end{cases} \quad (21)$$

<sup>2</sup> Alternatively, the criterion can be defined as follows:  $|C_x(t, f)| = |\sum_{i=1}^m D_{x_i x_i}(t, f)| / \sum_{i=1}^m |D_{x_i x_i}(t, f)| = |\text{trace}(\mathbf{D}_{\mathbf{xx}}(t, f))| / m \mathcal{D}_{\mathbf{xx}}(t, f)$ . The use of absolute value allows us to exclude the cross-terms of different signal components of the same source. The cross-component terms of a multicomponent source signal are actually autosource terms from the source separation perspective [28] (notice that cross-term TFD takes both positive and negative values). The difference between the use of  $C_x(t, f)$  and  $|C_x(t, f)|$  will be demonstrated through numerical examples in Section 6.

<sup>3</sup> Although the cross-term points are not directly used in the proposed BSS method, they can be incorporated for the purpose of BSS as well as for direction finding [14, 29]. Therefore, the selection of cross-term and mixed auto- and cross-term regions is an important issue in the underlying topic.

where  $\gamma_1$  and  $\gamma_2$  are the respective threshold values for auto- and cross-term selection.

### 3.4. Selection based on whitened data

Although the array averaging is simple, it is likely to identify some false autoterm locations when the spatial correlation between the sources is high, that is, the sources have close signatures. In this case, the performance can be improved by averaging the whitened STFDs instead. When the array averaging of the whitened STFD matrices  $\mathbf{D}_{\mathbf{z}\mathbf{z}}(t, f)$  is considered, as depicted in (10), the unitary matrix  $\mathbf{U}$  becomes the effective mixing matrix that relates an STFD matrix and its corresponding source TFD matrix. Therefore, the whitening amounts to force the spatial correlation between any pair of different source signals to be zero, whereas the spatial correlation of the same source remains unity. When the whitened STFDs are used, the above autoterm selection procedure is represented by the following equations:

$$C_z(t, f) = \frac{\sum_{i=1}^n D_{z_i z_i}(t, f)}{\sum_{i=1}^n |D_{z_i z_i}(t, f)|} = \frac{\text{trace}(\mathbf{D}_{\mathbf{z}\mathbf{z}}(t, f))}{n\mathcal{D}_{\mathbf{z}\mathbf{z}}(t, f)}, \quad (22)$$

where

$$\mathcal{D}_{\mathbf{z}\mathbf{z}}(t, f) = \frac{1}{n} \sum_{i=1}^n |D_{z_i z_i}(t, f)|. \quad (23)$$

The auto- and cross-term points are identified as<sup>4</sup>

$$C_z(t, f) = \frac{\text{trace}(\mathbf{D}_{\mathbf{z}\mathbf{z}}(t, f))}{n\mathcal{D}_{\mathbf{z}\mathbf{z}}(t, f)} > \alpha_3 \rightarrow \text{decide that } (t, f) \text{ is autoterm,} \\ < \alpha_4 \rightarrow \text{decide that } (t, f) \text{ is cross-term.} \quad (24)$$

We also use a threshold level of the averaged absolute value of the TFD for meaningful auto- and cross-term selection. When the whitened data are used, we can define  $\mathcal{D}_{\mathbf{z}\mathbf{z}, \max}$  in a similar manner to  $\mathcal{D}_{\mathbf{z}\mathbf{z}, \max}$ , and the associated condition becomes

$$F_z(t, f) = \frac{\mathcal{D}_{\mathbf{z}\mathbf{z}}(t, f)}{\mathcal{D}_{\mathbf{z}\mathbf{z}, \max}} > \begin{cases} \gamma_3, & \text{for autoterm selection,} \\ \gamma_4, & \text{for cross-term selection,} \end{cases} \quad (25)$$

where  $\gamma_3$  and  $\gamma_4$  are the respective threshold values for auto- and cross-term selection when the whitened data are used for this purpose.

Therefore, (24) differs from (12) only on the denominator. While the computation of a matrix norm requires all the auto- and cross-sensor terms, the computation of the average absolute term used in the proposed method only requires autosensor terms. Moreover, because  $C_z(t, f)$  is upper-bounded by unity and the physical meaning of  $C_z(t, f) = 1$  is very clear, it becomes much easier to determine the threshold values.

## 4. MODIFIED SOURCE SEPARATION METHOD

### 4.1. Time-frequency grouping

In [26], the subspace analysis of STFD matrices was presented for signals with clear t-f signatures, such as frequency modulated (FM) signals. It was shown that the offerings of using an STFD matrix instead of the covariance matrix are basically two folds. First, the selection of autoterm t-f points, that is, points on the source instantaneous frequencies, where the signal power is concentrated, enhances the equivalent input SNR. Second, the difference in the t-f localization properties of the source signals permits source discrimination and allows the selection of fewer sources for STFD matrix construction. In the presence of noise, the consideration of a subset of signal arrivals reduces perturbation in matrix eigendecomposition. T-f grouping becomes essential to recover the source waveforms when there is insufficient number of sensors, provided that the TFD of the different subgroups is disjoint.

In this section, we introduce the notion of t-f signature grouping to process a subclass of the sources which have disjoint t-f signatures. The use of STFD for improved whitening performance is considered in the next section.

With the effective selection of autoterm-only t-f points, sources with disjoint (orthogonal) t-f supports can be classified into different groups. For example, if  $n_o < n$  sources occupy t-f support  $\Omega_1$  (i.e.,  $D_{d_i d_k}(t, f) \neq 0$  if and only if  $(t, f) \in \Omega_1, i, k = 1, \dots, n_o$ ), and the remaining  $n - n_o$  sources occupy t-f support  $\Omega_2$  (i.e.,  $D_{d_i d_k}(t, f) \neq 0$  if and only if  $(t, f) \in \Omega_2, i, k = n_o + 1, \dots, n$ ), then Group 1 of the first  $n_o$  sources and Group 2 of the remaining  $n - n_o$  are said to be disjoint in the t-f domain if  $\Omega_1 \cap \Omega_2 = \emptyset$ . The number of sources included in a t-f group can be estimated by examining the rank of the STFD matrix defined over the t-f support of this group [17, 26].

When the number of sources does not exceed the number of array sensors, t-f grouping is optional, and we can rely only on the autoterm points for blind source separation. In this case, we can simplify the problem by examining only the autoterm points obtained in Section 3. When the number of sources exceeds the number of array sensors, t-f grouping is essential, and we must carefully consider all the auto- and cross-term information within each group for signal synthesis. We will discuss such situations in more detail in Section 5.

Subgrouping has been studied in, for example, [17, 22], depending on the closeness of the spatial signatures in a group, or on the potential function as the sum of the individual contributions in the space of directions. In this paper, we consider a subgroup simply as a region determined by continued or cluttered autoterm t-f points. The subgrouping procedure is summarized below.

(1) Compute  $\mathcal{D}_{\mathbf{z}\mathbf{z}}$  or  $\mathcal{D}_{\mathbf{x}\mathbf{x}}$  and the corresponding  $C_z(t, f)$  or  $C_x(t, f)$  function.

(2) Perform two-dimensional low-pass filtering in both the time and frequency domains. (It is an optional operation to reduce the cross-terms, which may show higher peak value than the autoterms, by taking advantage of the oscillating

<sup>4</sup> Similar to  $|C_x(t, f)|$ , we can also use  $|C_z(t, f)|$  for auto- and cross-term identification.

nature of the cross-terms whereas the autoterms are positive and less variant).

(3) Find the peak of the autoterm and its connected autoterm region. A mask is then identified as the polygon spanned by the autoterm region.

(4) Repeat this process until no significant autoterm regions are identified.

In selecting the autoterm t-f points, a moderate  $\gamma_1$  or  $\gamma_3$  value can be used to ensure the selection of t-f points with high energy localization and to reduce the set size of autoterm points so that the computational complexity can be managed. It is often effective to select high SNR autoterm t-f points that achieve local maxima [16].

#### 4.2. Modified source separation method

In the method proposed in [12] and summarized in Section 2.3, STFD matrices are used to estimate the unitary matrix  $\mathbf{U}$ . However, the whitening process is still based on the covariance matrix. An estimate of the covariance matrix is often not as robust to noise as a well-defined STFD matrix. Particularly, when the source signals can be separated in the t-f domain but fail to separate in the time domain, then at least the same number of sensors as the number of sources is required to provide complete whitening based on the covariance matrix, whereas fewer array sensors could do the job if the STFD matrices are used. Below, we use the STFD matrix in place of the covariance matrix  $\mathbf{R}_{\mathbf{x}\mathbf{x}}$  for whitening [30].

Denote  $\mathbf{D}_{\mathbf{x}\mathbf{x}}(t_1, f_1), \dots, \mathbf{D}_{\mathbf{x}\mathbf{x}}(t_K, f_K)$  as the STFD matrices constructed from  $K$  autoterm points being defined over a t-f region  $\Omega_1$  and belonging to fewer  $n_o \leq n$  signals. Also, denote, respectively,  $\mathbf{d}^o(t)$  and  $\mathbf{d}(t)$  as the  $n_o$  and  $n - n_o$  sources being present and absent in the t-f region  $\Omega_1$ . The  $n - n_o$  sources could be undesired emitters or sources to be separated in the next round of processing. The value of  $n_o$  is generally unknown and can be determined from the eigenstructure of the STFD matrix. Using the above notations, we obtain

$$\mathbf{x}(t) = \mathbf{A}^o \mathbf{d}^o(t) + \mathbf{A} \mathbf{d}(t) + \mathbf{n}(t), \quad (26)$$

where  $\mathbf{A}^o$  and  $\mathbf{A}$  are the  $m \times n_o$  and the  $m \times (n - n_o)$  mixing matrices corresponding to  $\mathbf{d}^o(t)$  and  $\mathbf{d}(t)$ , respectively.

The incorporation of multiple t-f points through the joint diagonalization or t-f averaging reduces the noise effect on the signal subspace estimation, as discussed in [12, 26]. For example, let  $\tilde{\mathbf{D}}_{\mathbf{x}\mathbf{x}}$  be the average STFD matrix of a set of STFD matrices defined over the same region  $\Omega_1$  using a different t-f kernel, and denote  $\hat{\sigma}^{tf}$  as the estimation of the noise-level eigenvalue of  $\tilde{\mathbf{D}}_{\mathbf{x}\mathbf{x}}$ . Then:

$$\begin{aligned} \tilde{\mathbf{W}} \tilde{\mathbf{D}}_{\mathbf{y}\mathbf{y}} \tilde{\mathbf{W}}^H &= \tilde{\mathbf{W}} (\tilde{\mathbf{D}}_{\mathbf{x}\mathbf{x}} - \hat{\sigma}^{tf} \mathbf{I}) \tilde{\mathbf{W}}^H \\ &= \tilde{\mathbf{W}} \mathbf{A}^o \tilde{\mathbf{D}}_{\mathbf{d}\mathbf{d}}^o (\tilde{\mathbf{W}} \mathbf{A}^o)^H = \mathbf{I}. \end{aligned} \quad (27)$$

In (27), due to the ambiguity of signal complex amplitude in BSS, we have assumed for convenience and without loss of generality that the averaged source TFD matrix  $\tilde{\mathbf{D}}_{\mathbf{d}\mathbf{d}}^o$  corresponding to  $\mathbf{d}^o(t)$  is  $\mathbf{I}$  of  $n_o \times n_o$ . Therefore, the whitening

matrix  $\tilde{\mathbf{W}}$  is obtained as

$$\tilde{\mathbf{W}} = [(\lambda_1^{tf} - \hat{\sigma}^{tf})^{-1/2} \mathbf{h}_1^{tf}, \dots, (\lambda_{n_o}^{tf} - \hat{\sigma}^{tf})^{-1/2} \mathbf{h}_{n_o}^{tf}]^H, \quad (28)$$

where  $\lambda_1^{tf}, \dots, \lambda_{n_o}^{tf}$  are the  $n_o$  largest eigenvalues of  $\tilde{\mathbf{D}}_{\mathbf{x}\mathbf{x}}$  and  $\mathbf{h}_1^{tf}, \dots, \mathbf{h}_{n_o}^{tf}$  are the corresponding eigenvectors of  $\tilde{\mathbf{D}}_{\mathbf{x}\mathbf{x}}$ . Note that  $\tilde{\mathbf{D}}_{\mathbf{d}\mathbf{d}}^o$  and  $\tilde{\mathbf{D}}_{\mathbf{y}\mathbf{y}}$  are of reduced rank  $n_o$  instead of rank  $n$ , as a result of the source discrimination performed through the selection of the t-f points or specific t-f regions. Therefore,  $\tilde{\mathbf{W}} \mathbf{A}^o = \tilde{\mathbf{U}}$  is a unitary matrix, whose dimension is  $n_o \times n_o$  rather than  $n \times n$ . The whitened process  $\tilde{\mathbf{z}}(t)$  becomes

$$\begin{aligned} \tilde{\mathbf{z}}(t) &= \tilde{\mathbf{W}} \mathbf{x}(t) = \tilde{\mathbf{W}} \mathbf{A}^o \mathbf{d}^o(t) + \tilde{\mathbf{W}} \mathbf{A} \mathbf{d}(t) + \tilde{\mathbf{W}} \mathbf{n}(t) \\ &= \tilde{\mathbf{U}} \mathbf{d}^o(t) + \tilde{\mathbf{W}} \mathbf{A} \mathbf{d}(t) + \tilde{\mathbf{W}} \mathbf{n}(t). \end{aligned} \quad (29)$$

In the t-f region  $\Omega_1$ , the TFD of  $\mathbf{d}(t)$  is zero and, therefore, the averaged STFD matrix of the noise-free components becomes an identity matrix, that is,

$$\tilde{\mathbf{D}}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} = \tilde{\mathbf{W}} \tilde{\mathbf{D}}_{\mathbf{x}\mathbf{x}} \tilde{\mathbf{W}}^H = \tilde{\mathbf{U}} \tilde{\mathbf{D}}_{\mathbf{d}\mathbf{d}}^o \tilde{\mathbf{U}}^H = \mathbf{I}. \quad (30)$$

Equation (30) implies that the auto- and cross-term TFDs averaged over the t-f region  $\Omega_1$  become unity and zero, respectively, upon whitening with matrix  $\tilde{\mathbf{W}}$ .  $\tilde{\mathbf{U}}$  as well as the mixing matrix and source waveforms are estimated following the same procedure of Section 3. It is noted that, when  $n_o = 1$ , source separation is no longer necessary and the steering vector of the source signal can be obtained from the received data at a single or multiple t-f points in the respective t-f region [31].

In the method developed in [12], the number of sources included in the STFD matrices may be smaller than that included in the covariance matrix, if the STFD is constructed from a subset of signal arrivals. As such, the signal subspace spanned by the STFD matrices is not identical to that spanned by the covariance matrix. For the modified method, both sets of STFD matrices are based on the number of sources.

Selection of the same number of sources,  $n_o$ , should be done at both whitening and joint diagonalization stages, otherwise mismatching of the corresponding sources will result. While our proposed modified blind source separation method provides the mechanism to satisfy this condition, the covariance matrix-based whitening approach does not lend itself to avoid any mismatching.

## 5. SEPARATION OF MORE SOURCES THAN THE NUMBER OF SENSORS

When there are more sources than array sensors, the mixing matrix  $\mathbf{A}$  is wide, and orthogonalization of all signal mixing vectors becomes impossible. Therefore, even though the mixing vector, or the spatial signature, can be estimated for each source signal by using the source discrimination introduced in Section 4 and choosing  $n_o \leq m$ , the signal waveforms remain inseparable by merely multiplying the (pseudo) inverse of the mixing matrix to the received data

vector. For the sources to be fully separable, they have to be partitioned into groups such that the number of sources in each group does not exceed the number of array sensors.

For this purpose, it is important to emphasize that, while the same grouping procedure described in Section 4.1 can be used to construct the masks, special consideration should be taken to solve the underdetermined source separation problems. For the scenario discussed in Section 4.1, where the number of sources is less than the number of sensors, we only need to select several autoterm t-f points that provide sufficient information for the estimation of the mixing matrix of the sources. It was not required for the selected autoterm region to contain the full source waveform information. When we consider the situation with more source signals than the number of array sensors, however, the selected autoterm regions must contain as much as possible the full information of the signal waveforms. In particular, the regions with mixed auto- and cross-terms of the sources of the interested group should be included for this purpose.

We consider to achieve this purpose by constructing proper t-f masks. The mask at the  $k$ th t-f group, denoted as  $M_k(t, f)$ , should include the autoterm of the signals in this group and the cross-term among them, whereas the auto- and cross-terms of the signals not included in the group, and the cross-terms between in-group and out-group signals, should be excluded. Fortunately, as the cross-terms are located between autoterms, a group region is usually bounded by the signatures of its autoterm components. Cross-terms located between two groups can be simply considered as cross-group terms and thus can be removed for this purpose. To preserve the waveform information, a relatively small value of  $\gamma$  should be chosen. It is also noted that perfect prewhitening using the covariance matrix cannot be realized with the number of array sensors smaller than the number of sources.

Once the sources are successfully partitioned into several groups, the masked TFD,  $D_{x_i x_i}(t, f)M_k(t, f)$ , at the  $i$ th sensor is used to synthesize the (mixed) signal waveforms at the  $k$ th group [32–34]. The method proposed in Section 4 is then applied to each group, and  $\hat{\mathbf{U}}^{(k)}$  and  $\hat{\mathbf{W}}^{(k)}$  corresponding to the  $k$ th group can be obtained. Notice that, because the synthesized signal  $\tilde{x}_i^{(k)}(t)$  is phase blind, the phase information should be recovered by projecting the original signal  $x_i(t)$  onto the signal subspace that  $\tilde{x}_i^{(k)}(t)$  spans, that is,

$$\underline{x}_i^{(k)} = \underline{\tilde{x}}_i^{(k)} \left[ \left( \underline{\tilde{x}}_i^{(k)} \right)^H \underline{\tilde{x}}_i^{(k)} \right]^{-1} \left( \underline{\tilde{x}}_i^{(k)} \right)^H \underline{x}_i, \quad (31)$$

where the underbar is used to emphasize the fact that each variable used here is a vector constructed over a period of time, for example,  $t = 0, \dots, T$ . The source signals are recovered at the  $k$ th group from

$$\hat{\mathbf{d}}^{(k)}(t) = \left( \hat{\mathbf{U}}^{(k)} \right)^H \hat{\mathbf{W}}^{(k)} \mathbf{x}^{(k)}(t), \quad (32)$$

where  $\mathbf{x}^{(k)}(t) = [x_1^{(k)}(t), \dots, x_{n^{(k)}}^{(k)}(t)]^T$ , with  $n^{(k)}$  denoting the number of sources at the  $k$ th subgroup.

## 6. SIMULATION RESULTS

### 6.1. Autoterm selection and grouping

In the first part of our simulations, we consider a three-element linear array with a half-wavelength spacing. Three source signals are considered. The first two are windowed single-component chirp signals, whereas the third one is a windowed multicomponent chirp signal. All the chirp components have the same magnitude. Therefore, the third signal with two chirp components has three dB higher SNR. The data length is 256. For simplicity, the three signals arrive from respective directions-of-arrival of 45, 15, and  $-10$  degrees, although a structured mixing matrix is not assumed. The WVDs of the three signals are plotted in Figures 1(a)–1(c). The WVD of the mixed signal at the first array sensor is shown in Figure 1(d) with input SNR = 5 dB.

In Figure 2, the results of pure autoterm selection are illustrated. While both plots show clear identification of the autoterm regions, the orthogonalization result is much “cleaner”. From these results, we can form two disjoint groups with one including sources 1 and 2, and the other including only source 3. For comparison, we have shown the results based on  $C_x(t, f)$  and  $C_z(t, f)$  as well as their absolute value counterparts. The use of  $C_x(t, f)$  and  $C_z(t, f)$  allows the exclusion of cross-terms with large negative values, whereas their absolute value counterparts do not discriminate the negative cross-term values.

In Figure 3, the results of pure cross-term selection are illustrated. It is noted that the cross-terms between sources 1 and 2 are cross-source terms, whereas the cross-terms between the two components of source 3 are autoterm terms. When comparing the use of  $C_x(t, f)$  and  $C_z(t, f)$  with their absolute value counterparts, the difference is very evident. Results based on  $C_x(t, f)$  and  $C_z(t, f)$  include cross-component terms of source 3, whereas such cross-component terms are clearly removed in the results obtained from  $|C_x(t, f)|$  and  $|C_z(t, f)|$ . Therefore, the later results are closer to the actual situation. As for the effect of orthogonalization, it is evident that the orthogonalization reduces the cross-term components in general. The results obtained before orthogonalization are closer to the real situation.

### 6.2. Source separation

The performance of source separation is evaluated by using the mean rejection level (MRL), defined as [12],

$$\text{MRL} = \sum_{p \neq q} E \left| \left( \hat{\mathbf{A}}^\# \mathbf{A} \right)_{pq} \right|^2, \quad (33)$$

where  $\hat{\mathbf{A}}$  is the estimate of  $\mathbf{A}$ . A smaller value of the MRL implies better source separation performance. An MRL lower than  $-10$  dB is considered satisfactory [12].

Figure 4 shows that the MRL versus the input SNR of the three sources. The curves are calculated by averaging 100 independent trials with different noise sequences. The dashed line corresponds to method [12] where the covariance matrix

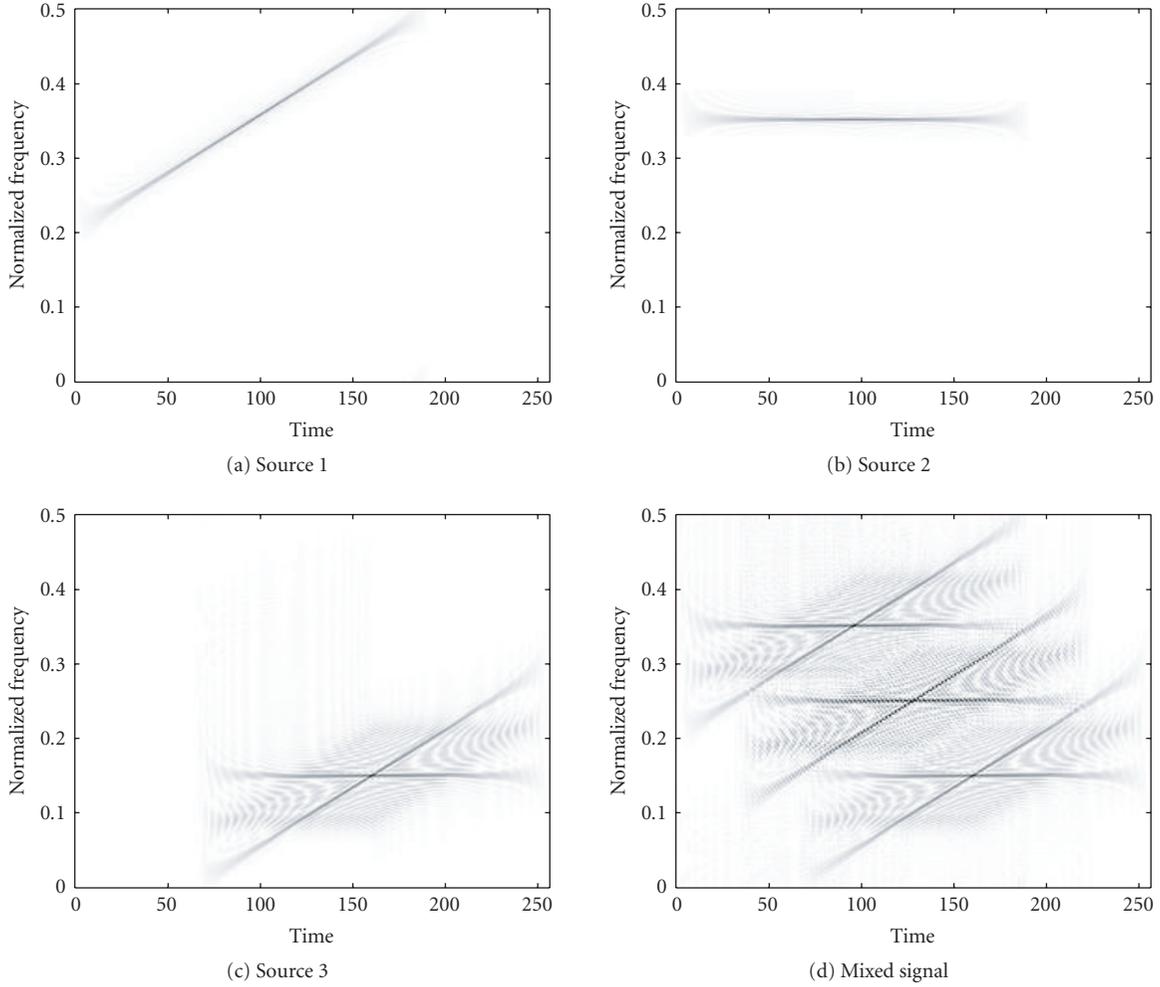


FIGURE 1: WVDs of the three source signals and the mixed signal at the first sensor with input SNR = 5 dB.

$\mathbf{R}_{xx}$  is used for whitening, and the solid line corresponds to the modified method where the averaged STFD matrix  $\tilde{\mathbf{D}}_{xx}$  is used instead. The dashed-dotted line shows the results using the proposed method and the three signals are partitioned into two groups, where the first group contains the first two sources, and the second group contains the third source signal. In the proposed method, the average of spatial pseudo-Wigner-Ville distributions (SPWVDs) of window size 33 is applied to estimate the whitening matrix. For the estimation of the unitary matrix for both methods, the spatial Wigner-Ville distribution (SWVD)<sup>5</sup> matrices using the entire data record are computed. The number of points used to perform the joint diagonalization for unitary matrix estimation is  $K = 32$  for each signal, and the points are selected at the t-f autoterm locations. Figure 4 clearly shows the improvement when STFDs are used in both phases of source separations, specifically for low SNRs. To satisfy the  $-10$  dB MRL, the

required input SNR is about 12.1 dB for the method developed in [12], and is about 2.4 dB and 5.1 dB for the modified method with and without t-f grouping. The advantages of using the proposed method, particularly with the t-f grouping, are evident from the results shown in this figure.

### 6.3. Separation of more sources than the number of sensors

In the second part of simulation, we use the same parameters used in Section 6.1, but the number of sensors is now only 2. The input SNR is fixed to 5 dB. In this case, covariance matrix-based method cannot whiten the three-source data vector. To separate the three signal arrivals using the proposed method, we need to partition the t-f domain so that the maximum number of sources contained in each group does not exceed two. In this example, we construct a mask that contains the first two sources and the procedure described in Section 5 is followed.

Figure 5 illustrates the construction of the masks. We determine the autoterm regions based on unwhitened criterion function  $C_x(t, f)$  where, as we explained earlier, a small value

<sup>5</sup> The method proposed here is not limited to use specific TFDs and the SPWVD and SWVD are chosen for simplicity. Other TFDs can also be used.

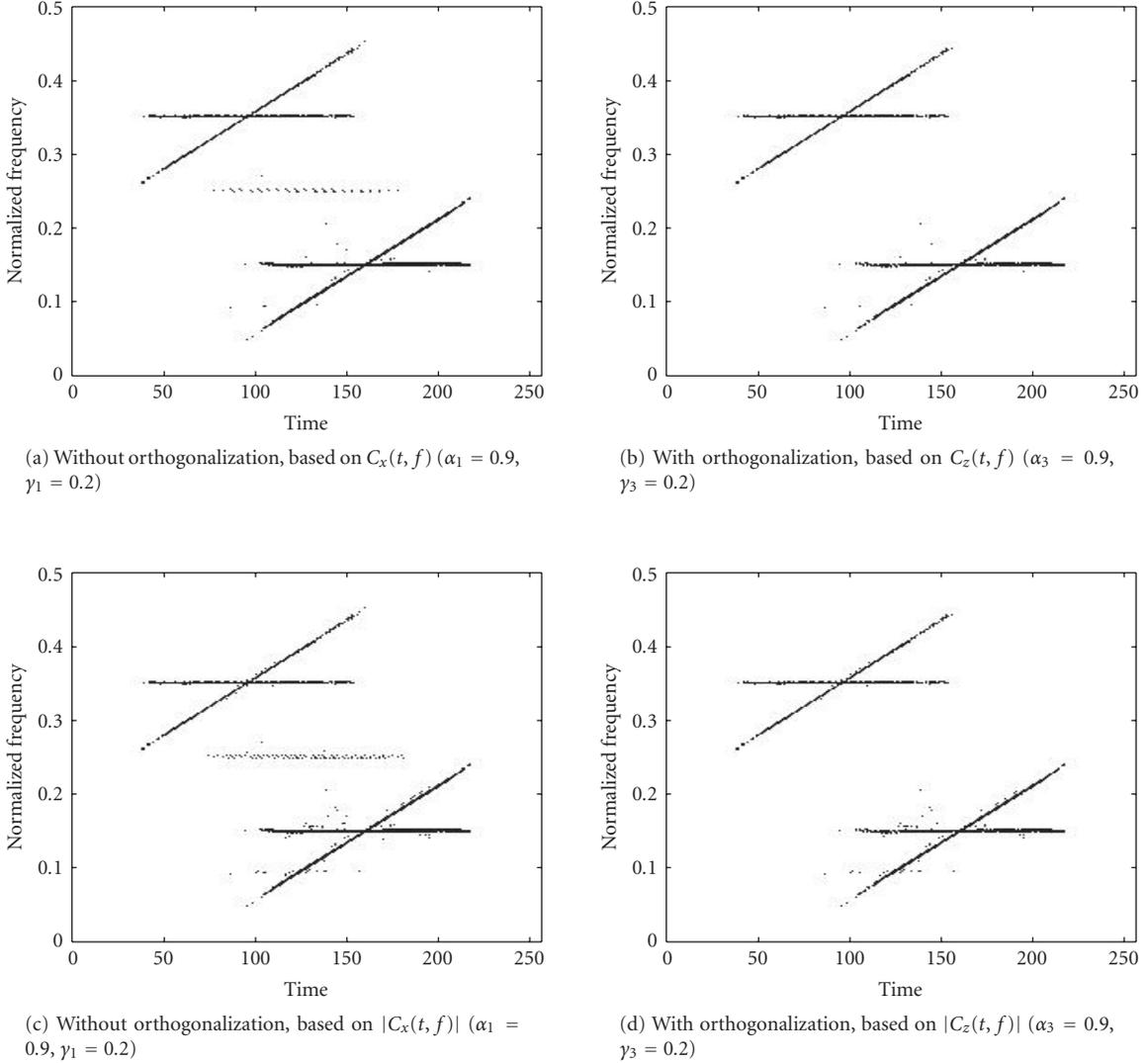


FIGURE 2: Selected autoterm regions.

of  $\gamma = 0.05$  is used, which coincides with the threshold level for noise reduction in [17]. The estimated result of the autoterm regions is depicted in Figure 5(a). Figure 5(b) shows the two masks constructed from the mask construction process illustrated in Section 4.1, one includes sources 1 and 2, whereas the other includes source 3. Source separation and waveform recovery are performed within each masked region separately.

From the discussion in Section 5, we know that the performance index alone, when the number of sources exceeds the number of sensors, does not explain how the separated signal waveforms are close to the original source waveforms. For this reason, we plot in Figure 6 the WVDs of the two separated signals (source 1 and source 2). They are very close to the original source TFDs. The MRL, computed from the spatial signatures of the selected two sources averaged for 200 independent trials, is  $-19.5$  dB, compared to  $-20.5$  dB corresponding to the case in which only the two source signals

are present and, therefore, no mask is applied. The WVD of source 3 estimate is also included for reference. Note that the estimation of source 3 does not require separation because it is the only source in the group. It is reconstructed from masking, waveform synthesis at each sensor, and the combining of the synthesized waveforms at the sensors.

## 7. CONCLUSION

In this paper, we have addressed several important issues in STFD-based BSS problems. First, a simple method for auto- and cross-term selection was introduced which requires only the autosensor TFDs. Second, the STFD-based BSS method has been modified to use multiple STFD matrices for prewhitening. Third, t-f grouping and masking for source discrimination are introduced for performance improvement and to separate more sources than the number of sensors.

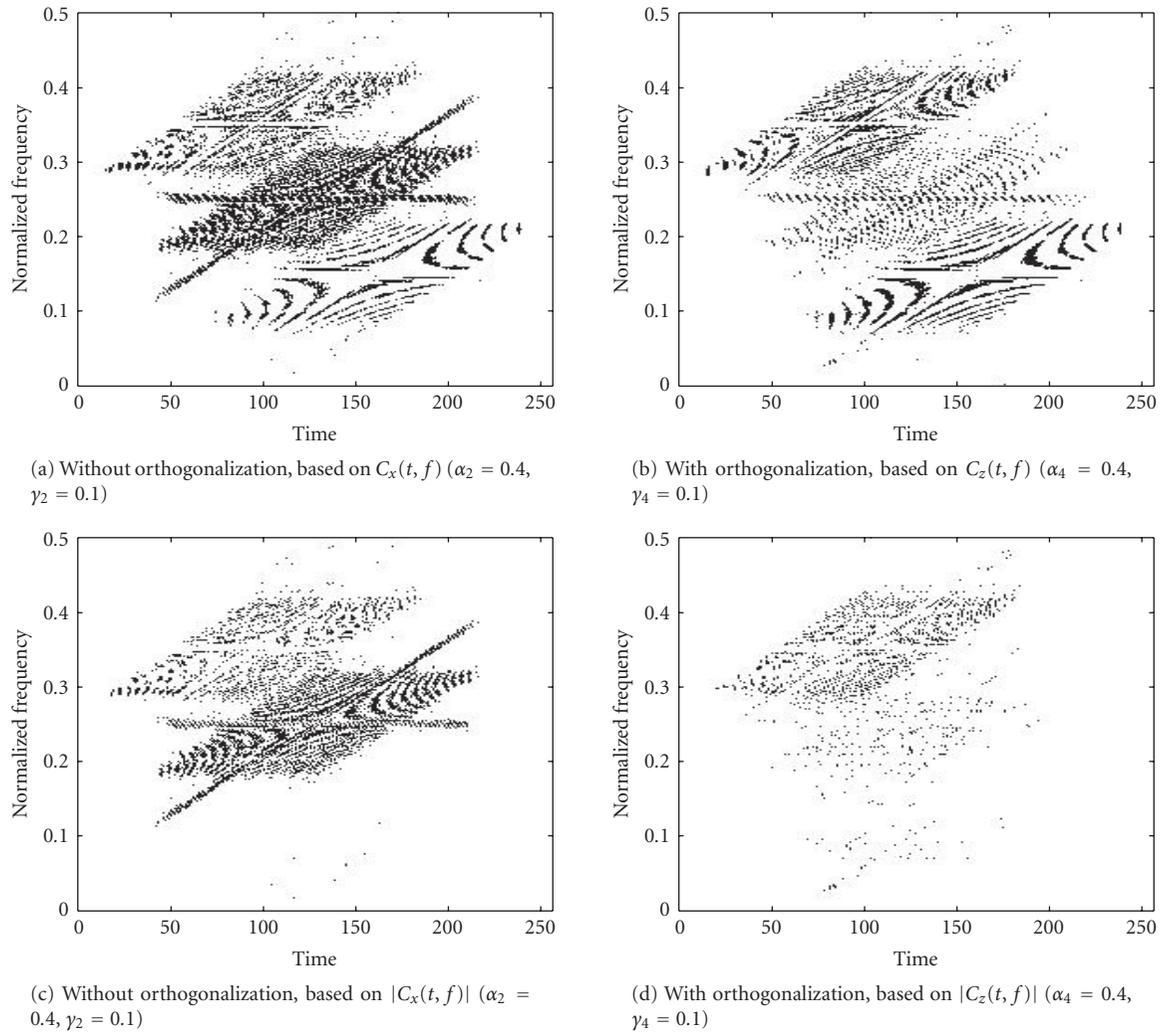
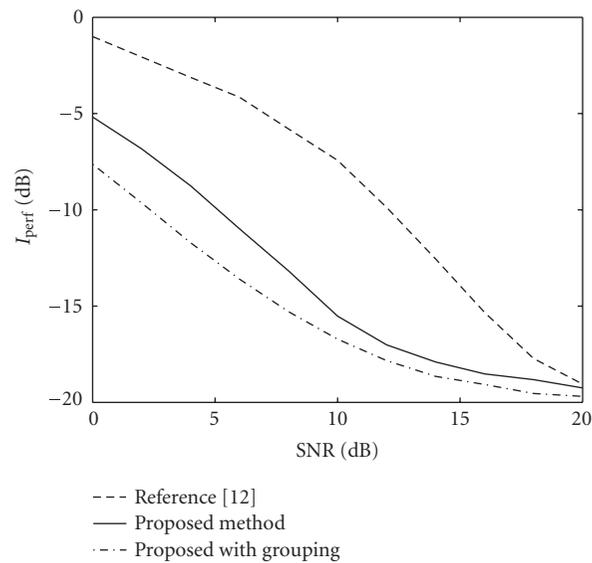


FIGURE 3: Selected cross-term regions.

FIGURE 4: MRL versus input SNR ( $m = 3$ ,  $n = 3$ ).

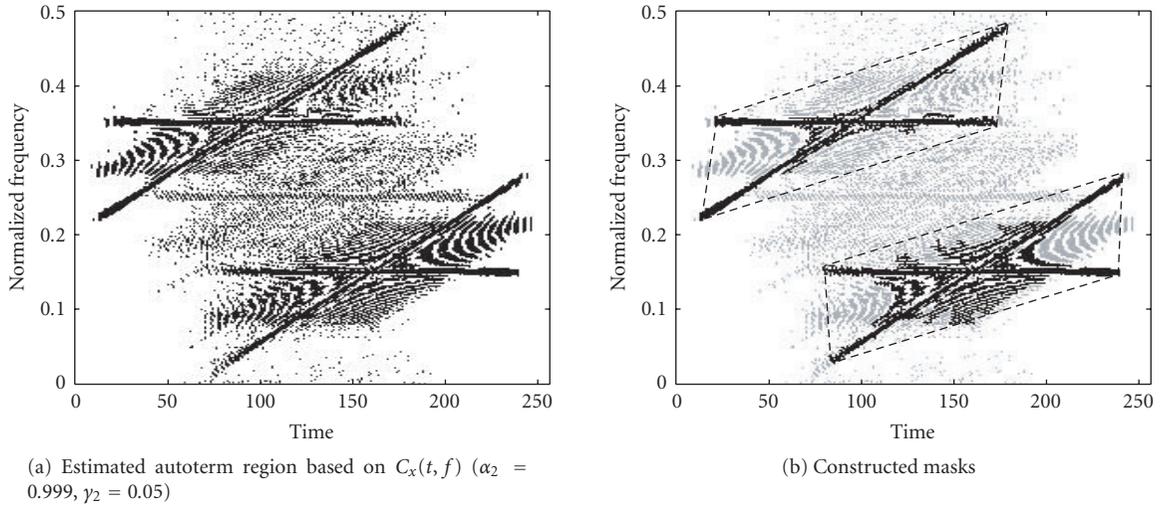


FIGURE 5: Construction of mask ( $m = 2, n = 3, \text{SNR} = 5 \text{ dB}$ ).

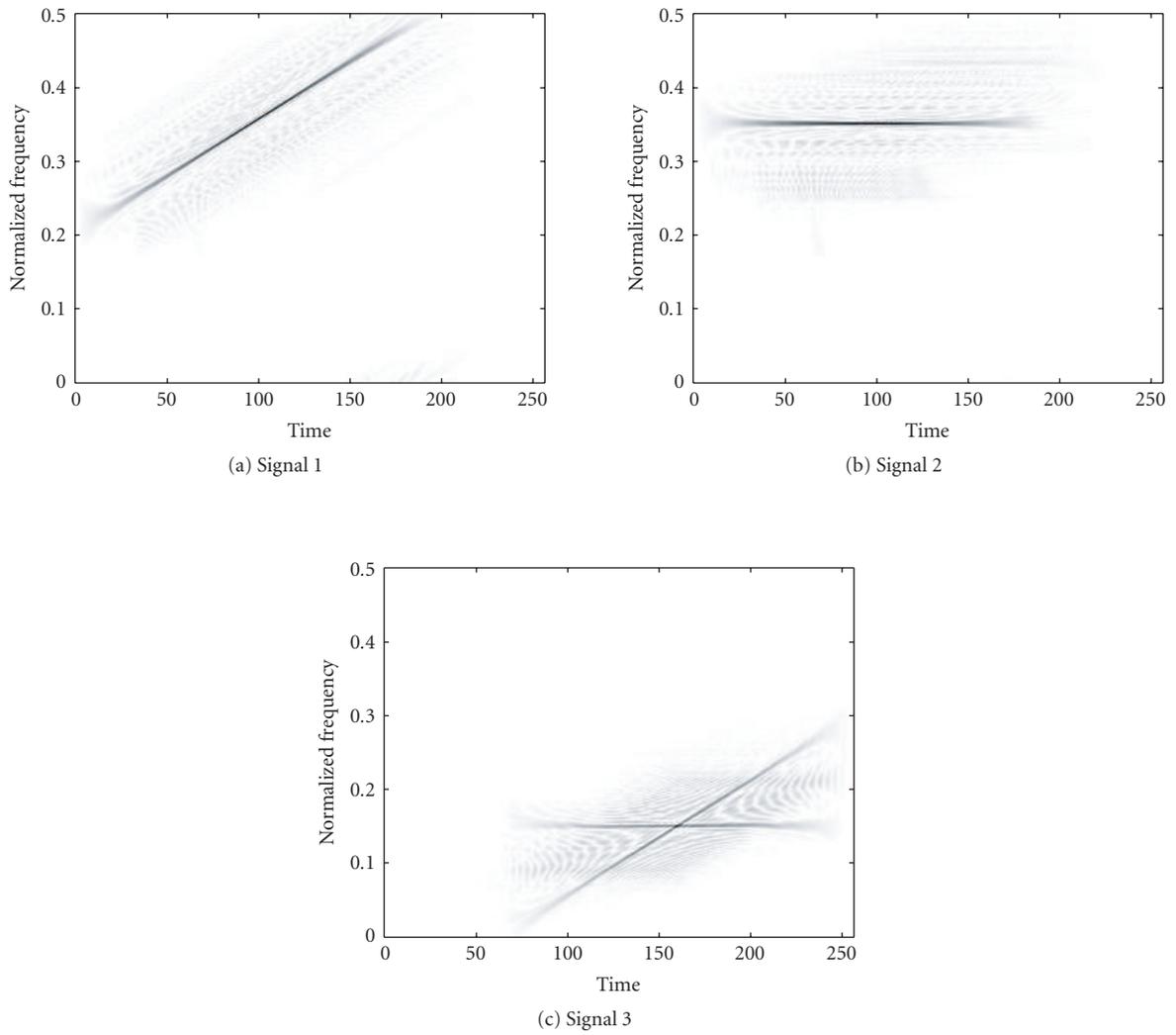


FIGURE 6: WVD of separated and reconstructed signals through masking ( $m = 2, \text{SNR} = 5 \text{ dB}$ ).

## ACKNOWLEDGMENTS

This work was supported by the ONR under Grant N00014-98-1-0176 and ONR/NSWC under Contract no. N65540-05-C-0028.

## REFERENCES

- [1] L. Tong, Y. Inouye, and R.-W. Liu, "Waveform-preserving blind estimation of multiple independent sources," *IEEE Transactions on Signal Processing*, vol. 41, no. 7, pp. 2461–2470, 1993.
- [2] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *IEE Proceedings, Part F: Radar and Signal Processing*, vol. 140, no. 6, pp. 362–370, 1993.
- [3] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "Blind source separation technique using second-order statistics," *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 434–444, 1997.
- [4] O. Grellier and P. Comon, "Blind separation of discrete sources," *IEEE Signal Processing Letters*, vol. 5, no. 8, pp. 212–214, 1998.
- [5] B. Chen and A. P. Petropulu, "Frequency domain blind MIMO system identification based on second- and higher order statistics," *IEEE Transactions on Signal Processing*, vol. 49, no. 8, pp. 1677–1688, 2001.
- [6] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, New York, NY, USA, 2001.
- [7] A. Hyvärinen, "Survey on independent component analysis," *Neural Computing Surveys*, vol. 2, pp. 94–128, 1999.
- [8] P. D. O'Grady, B. A. Pearlmutter, and S. T. Rickard, "Survey of sparse and non-sparse methods in source separation," *International Journal of Imaging Systems and Technology*, vol. 15, no. 1, pp. 18–33, 2005, special issue on Blind Source Separation and Deconvolution in Imaging and Image Processing.
- [9] V. Chen and H. Ling, *Time-Frequency Transforms for Radar Imaging and Signal Analysis*, Artech House, Boston, Mass, USA, 2002.
- [10] B. Boashash, Ed., *Time-Frequency Signal Analysis and Processing*, Elsevier, Oxford, UK, 2003.
- [11] L. Debnath, Ed., *Wavelets and Signal Processing*, Birkhäuser, Boston, Mass, USA, 2003.
- [12] A. Belouchrani and M. G. Amin, "Blind source separation based on time-frequency signal representations," *IEEE Transactions on Signal Processing*, vol. 46, no. 11, pp. 2888–2897, 1998.
- [13] A. R. Leyman, Z. M. Kamran, and K. Abed-Meraim, "Higher-order time frequency-based blind source separation technique," *IEEE Signal Processing Letters*, vol. 7, no. 7, pp. 193–196, 2000.
- [14] A. Belouchrani, K. Abed-Meraim, M. G. Amin, and A. M. Zoubir, "Joint anti-diagonalization for blind source separation," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '01)*, vol. 5, pp. 2789–2792, Salt Lake, Utah, USA, May 2001.
- [15] L. Giulieri, N. Thirion-Moreau, and P.-Y. Arquès, "Blind sources separation based on bilinear time-frequency representations: a performance analysis," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '02)*, vol. 2, pp. 1649–1652, Orlando, Fla, USA, May 2002.
- [16] C. Févotte and C. Doncarli, "Two contributions to blind source separation using time-frequency distributions," *IEEE Signal Processing Letters*, vol. 11, no. 3, pp. 386–389, 2004.
- [17] N. Linh-Trung, A. Belouchrani, K. Abed-Meraim, and B. Boashash, "Separating more sources than sensors using time-frequency distributions," *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 17, pp. 2828–2847, 2005.
- [18] W. Mu, M. G. Amin, and Y. Zhang, "Bilinear signal synthesis in array processing," *IEEE Transactions on Signal Processing*, vol. 51, no. 1, pp. 90–100, 2003.
- [19] Ö. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Transactions on Signal Processing*, vol. 52, no. 7, pp. 1830–1847, 2004.
- [20] A. Yeredor, "Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation," *IEEE Transactions on Signal Processing*, vol. 50, no. 7, pp. 1545–1553, 2002.
- [21] L. Giulieri, N. Thirion-Moreau, and P.-Y. Arquès, "Blind sources separation based on quadratic time-frequency representations: a method without pre-whitening," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '03)*, vol. 5, pp. 289–292, Hong Kong, April 2003.
- [22] P. Bofill and M. Zibulevsky, "Blind separation of more sources than mixtures using sparsity of their short-time Fourier transform," in *Proceedings of the 2nd International Workshop on Independent Component Analysis and Blind Signal Separation*, pp. 87–92, Helsinki, Finland, June 2000.
- [23] A. Aissa-El-Bey, K. Abed-Meraim, and Y. Grenier, "Underdetermined blind source separation of audio sources in time-frequency domain," in *Proceedings of the Signal Processing with Adaptive Sparse Structured Representations (SPARS '05)*, Rennes, France, November 2005.
- [24] S. Rickard, T. Melia, and C. Fearon, "DESPRIT—histogram based blind source separation of more sources than sensors using subspace methods," in *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA '05)*, pp. 5–8, New Paltz, NY, USA, October 2005.
- [25] L. A. Cirillo and M. G. Amin, "Auto-term detection using time-frequency array processing," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '03)*, vol. 6, pp. 465–468, Hong Kong, April 2003.
- [26] Y. Zhang, W. Mu, and M. G. Amin, "Subspace analysis of spatial time-frequency distribution matrices," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, pp. 747–759, 2001.
- [27] A. Holobar, C. Févotte, C. Doncarli, and D. Zazula, "Single autoterms selection for blind source separation in time-frequency plane," in *Proceedings of 11th European Signal Processing Conference (EUSIPCO '02)*, pp. 565–568, Toulouse, France, September 2002.
- [28] Y. Zhang and M. G. Amin, "Spatial averaging of time-frequency distributions for signal recovery in uniform linear arrays," *IEEE Transactions on Signal Processing*, vol. 48, no. 10, pp. 2892–2902, 2000.
- [29] M. G. Amin and Y. Zhang, "Direction finding based on spatial time-frequency distribution matrices," *Digital Signal Processing*, vol. 10, no. 4, pp. 325–339, 2000.
- [30] Y. Zhang and M. G. Amin, "Blind separation of sources based on their time-frequency signatures," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '00)*, vol. 5, pp. 3132–3135, Istanbul, Turkey, June 2000.
- [31] Y. Zhang, M. G. Amin, and G. J. Frazer, "A new approach to FM jammer suppression for digital communications," in *Proceedings of IEEE Sensor Array and Multichannel Signal Processing Workshop*, Rosslyn, Va, USA, August 2002.

- [32] G. F. Boudreaux-Bartels and T. W. Parks, "Time-varying filtering and signal estimation using Wigner distribution synthesis techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 3, pp. 442–451, 1986.
- [33] W. Krattenthaler and F. Hlawatsch, "Time-frequency design and processing of signals via smoothed Wigner distributions," *IEEE Transactions on Signal Processing*, vol. 41, no. 1, pp. 278–287, 1993.
- [34] F. Hlawatsch and W. Krattenthaler, "Bilinear signal synthesis," *IEEE Transactions on Signal Processing*, vol. 40, no. 2, pp. 352–363, 1992.

**Yimin Zhang** received his Ph.D. degree from the University of Tsukuba, Tsukuba, Japan, in 1988. He joined the faculty of the Department of Radio Engineering, Southeast University, Nanjing, China, in 1988. He served as a Technical Manager at the Communication Laboratory Japan, Kawasaki, Japan, from 1995 to 1997, and was a Visiting Researcher at ATR Adaptive Communications Research Laboratories, Kyoto, Japan, from 1997 to 1998. Since 1998, he has been with the Villanova University, where he is currently a Research Associate Professor at the Center for Advanced Communications and the Director of the Radio Frequency Identification (RFID) Lab. His research interests are in the areas of array signal processing, space-time adaptive processing, multiuser detection, MIMO systems, ad hoc and cooperative communications, blind signal processing, digital mobile communications, time-frequency analysis, and RFID systems. He is an Associate Editor for the *IEEE Signal Processing Letters*.



**Moeness G. Amin** received his Ph.D. degree in 1984 from University of Colorado, Boulder. He has been in the faculty of Villanova University since 1985, where he is now a Professor in the Department of Electrical and Computer Engineering and the Director of the Center for Advanced Communications. He is the Recipient of the IEEE Third Millennium Medal, Distinguished Lecturer of the IEEE Signal Processing Society for 2003–2004, Member of the Franklin Institute Committee of Science and Arts, recipient of the 1997 Villanova University Outstanding Faculty Research Award, recipient of the 1997 IEEE Philadelphia Section Service Award. He has over 280 publications in the areas of wireless communications, time-frequency analysis, smart antennas, interference cancellation in broadband communication platforms, digitized battlefield, direction finding, over-the-horizon radar, radar imaging, and channel equalizations.

