IMPROVED BLIND SEPARATIONS OF NONSTATIONARY SOURCES BASED ON SPATIAL TIME-FREQUENCY DISTRIBUTIONS

Yimin Zhang[†], *Moeness G. Amin*[†], *and Alan R. Lindsey*[‡]

[†] Center for Advanced Communications
Villanova University, Villanova, PA 19085, USA
[‡] Air Force Research Laboratory, Rome, NY 13441, USA

ABSTRACT

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 localizable in the time-frequency (t-f) domain. In this paper, we introduce a simple method for autoterm and crossterm selection, and propose the use of STFD matrices for both pre-whitening and mixing matrix recovery. T-f grouping is also proposed for improved blind separation of nonstationary signals. This method provides robust performance to noise and allows reduction of the number of sources considered for separation.

1. INTRODUCTION

Several narrowband blind source separation (BSS) methods have been proposed in the literature [1, 2, 3, 4]. Generally, blind source separations for independent sources are performed based on the employment of at least two different sets of matrices that span the same signal subspace. One matrix is used for whitening purpose, while others are jointly used to estimate the spatial signatures and source waveforms impinging on a multi-antenna receiver. Covariance matrices with different time-lags, or cumulants matrices with different orders, are typically used for the above purpose.

Time-frequency distributions (TFDs) have been recently employed for array processing and found successful in blind separations of nonstationary signals [5, 6, 7, 8]. These time-frequency (t-f) methods are particularly effective when the signals are highly localized in the t-f domain. Typically, the spatial time-frequency distribution (STFD) matrices are used for source diagonalization and anti-diagonalization, 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. Although this approach improves the performance over that based on solely the covariance matrices, yet it does not fully utilize all inherent advantages of STFD.

This work was supported by the Office of Naval Research under Grant N00014-98-1-0176 and by the Air Force Research Laboratory under Grant F30602-00-1-0515.

In this paper, we propose an improved source separation method based on STFDs. Employing the proposed method leads to robustness of subspace decompositions to noise and, thereby, enhances the unitary mixture representations of the problem. The contribution of this paper is three-fold. (1) A new method for selecting autoterm (t, f) points is first proposed. Autoterm points selection are very important in the underlying source separation problem to maintain the diagonal structure of the source TFD matrix. Existing methods [7, 9] require the calculation of either the norm or the eigenvalue of the whitened STFD matrices, which in turn, require both autoterms and crossterms between whitened elements. The proposed method, however, only requires the autoterms of the whitened STFD matrix and the decision process is both simpler and effective. (2) T-f grouping is proposed for the case when subsets of the source signals have clear and distinct t-f signatures. Based on effective autoterm selections, the proposed t-f grouping allows the consideration of fewer sources, leading to improved subspace estimation. (3) In the proposed source separation method, instead of using the covariance matrix for signal pre-whitening, multiple STFD matrices over the source t-f signatures are used and, thereby, making full use of the t-f localization properties of the sources in both the whitening step and joint estimation step of the source separation procedure.

2. SIGNAL MODEL

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

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

is commonly used, where $\mathbf{A}$ is the mixing matrix of dimension $m \times n$ and is assumed to be full column rank, $\mathbf{x}(t) = [x_1(t), ..., x_m(t)]^T$ is the sensor array output vector, and $\mathbf{d}(t) = [d_1(t), ..., d_n(t)]^T$ is the source signal vector, where the superscript ^{T} denotes the transpose operator. $\mathbf{n}(t)$ is an additive noise vector whose elements are modeled 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 source separation method proposed in [5], the source signals should be uncorrelated and their autoterm TFD regions are not contaminated by crossterm TFDs. In the proposed modified method, only the latter is required.

3. BLIND SOURCE SEPARATION BASED ON SPATIAL TIME-FREQUENCY SIGNATURES

3.1. Spatial Time-Frequency Distributions

The discrete form of Cohen's class of STFD of the data snapshot vector $\mathbf{x}(t)$ is given by [5]

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

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

$$\mathbf{D_{xx}}(t,f) = \mathbf{D_{yy}}(t,f) + \mathbf{D_{yn}}(t,f) + \mathbf{D_{ny}}(t,f) + \mathbf{D_{nn}}(t,f). \tag{3}$$

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

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

Similar to the typical mathematical formula (see Eqn. (6)), which relates the signal covariance matrix to the data spatial covariance matrix, Eqn. (4) provides the basis for source separation by relating the STFD matrix to the source TFD matrix, $\mathbf{D_{dd}}(t,f)$.

3.2. Blind Source Separation

In the TFD based blind source separation method proposed in [5], the following data covariance matrix is used for pre-whitening,

$$\mathbf{R_{xx}} = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbf{x}(t)\mathbf{x}^H(t), \tag{5}$$

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

$$\mathbf{R_{xx}} = \mathbf{R_{yy}} + \sigma\mathbf{I} = \mathbf{A}\mathbf{R_{dd}}\mathbf{A}^H + \sigma\mathbf{I}, \tag{6}$$

where $\mathbf{R_{dd}} = \lim_{T\to\infty}(1/T)\sum_{t=1}^{T}\mathbf{d}(t)\mathbf{d}^H(t)$ is the signal correlation matrix which is assumed diagonal, σ 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 of blind source separation based on TFDs is whitening of the signal $\mathbf{x}(t)$ of the observation. This is achieved by estimating the noise power [1] and applying a whitening matrix $\mathbf{W}$ to $\mathbf{x}(t)$, i.e., an $n\times m$ matrix satisfying:

$$\mathbf{W}\mathbf{R_{yy}}\mathbf{W}^H = \mathbf{W}(\mathbf{R_{xx}} - \sigma\mathbf{I})\mathbf{W}^H = \mathbf{W}\mathbf{A}\mathbf{A}^H\mathbf{W}^H = \mathbf{I}. \tag{7}$$

Therefore, the whitening matrix is estimated from the eigen-decomposition of $\mathbf{R_{xx}}$ [5]. Let λ_i denote the ith descendingly sorted eigenvalue of $\mathbf{R_{xx}}$, and $\mathbf{q}_i$ as the corresponding eigenvector, then the ith row of the whitening matrix is obtained as

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

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,

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

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

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

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, $\{\mathbf{D_{zz}}(t_i,f_i)|i = 1,...,K\}$. The source signals and the mixing matrix can be, respectively, estimated as $\hat{\mathbf{d}}(t) = \hat{\mathbf{U}}\hat{\mathbf{W}}\mathbf{x}(t)$ and $\hat{\mathbf{A}} = \hat{\mathbf{W}}^{\#}\hat{\mathbf{U}}$, where superscript $^{\#}$ denotes pseudo-inverse.

4. AUTOTERM SELECTION AND TIME-FREQUENCY GROUPING

4.1. Autoterm Selection

Existing Methods

The selection of autoterm (t,f) points has been considered in [7, 9]. In [7], it is pointed out that, at the crossterm (t,f) points, there are no source autoterms, that is, $trace(\mathbf{D}_{dd}(t,f)) = 0$. It was also shown that

$$\begin{aligned} trace(\mathbf{D_{xx}}(t,f)) &= trace(\mathbf{U}\mathbf{D_{dd}}(t,f)\mathbf{U}^H) \\ &= trace(\mathbf{D_{dd}}(t,f) \approx 0, \end{aligned} \tag{11}$$

and the following testing procedure was proposed:

$$\text{if } \frac{trace(\mathbf{D_{zz}}(t,f))}{norm(\mathbf{D_{zz}}(t,f))} \begin{matrix} < \\ > \end{matrix} \epsilon \rightarrow \text{decide that } (t,f) \text{ is } \begin{matrix} \text{crossterm} \\ \text{autoterm} \end{matrix}$$

where ϵ is a small positive real scalar. In [9], single autoterm locations are selected by noting the fact that $\mathbf{D_{dd}}(t,f)$ is diagonal with only one non-zero 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 a STFD matrix in the above two methods, all the autoterms and crossterms of the whitened vector $\mathbf{z}(t)$ are required. In the following, we propose an alternative method that only require the autoterm TFDs.

[1] The noise power can be estimated only when $m > n$ [5]. If $m = n$, the estimation of the noise power becomes unavailable and $\sigma = 0$ will be assumed.

Array Averaging

In [12], array average in the context of TFDs is discussed. Averaging of the auto-sensor TFDs across the array introduces a weighing function in the t-f domain that decreases the noise levels, reduces the interactions of the source signals, and mitigates the crossterms. 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 ith array sensor, $x_i(t) = \sum_{k=1}^{n} a_{ki} s_k(t)$, where a_{ki} is the ith element of mixing vector $\mathbf{a}_k$, is expressed as

$$D_{x_i x_i}(t,f) = \sum_{k=1}^{n}\sum_{l=1}^{n} a_{ki} a_{li}^{*} D_{d_k d_l}(t,f). \tag{12}$$

The averaging of $D_{x_i x_i}(t,f)$ for $i = 1, ..., m$ yields the array averaged TFD of the data vector $\mathbf{x}(t)$, defined as [12]

$$\begin{aligned} \bar{D}_{\mathbf{xx}}(t,f) &= \frac{1}{m}\sum_{i=1}^{m} D_{x_i x_i}(t,f) \\ &= \frac{1}{m}\sum_{i=1}^{n}\sum_{k=1}^{n} \mathbf{a}_k^H \mathbf{a}_i D_{d_i d_k}(t,f). \end{aligned} \tag{13}$$

Autoterm Selection

At a pure auto-source (t,f) point, where no cross-source terms are present, the TFD at the ith sensor is

$$D_{x_i x_i}(t,f) = \sum_{k=1}^{n} |a_{ki}|^2 D_{d_k d_k}(t,f), \tag{14}$$

which is consistently positive for all values of i. Accordingly $\mathbf{D}_{x_i x_i}(t,f) = |\mathbf{D}_{x_i x_i}(t,f)|$, $i = 1, ..., m$. Defining the following criterion function [2],

$$C_x(t,f) = \frac{\left|\sum_i^m D_{x_i x_i}(t,f)\right|}{\sum_i^m |D_{x_i x_i}(t,f)|}. = \frac{\left|trace\left(\bar{\mathbf{D}}_{\mathbf{xx}}(t,f)\right)\right|}{\sum_i^m |D_{x_i x_i}(t,f)|}. \tag{15}$$

The autoterm points can be identified as $C_x(t,f) > \alpha_1 \approx 1$.

For a pure cross-source (t,f) point, 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, $\mathbf{a}_k^H \mathbf{a}_i << 1$ for $k \neq i$ in (13), we have $C_x < \beta_1 \approx 0$.

For moderate values of $C_{\mathbf{x}}(t,f)$ between α_1 and β_1, the (t,f) point has both autoterm and crossterm present. Such a point should be excluded in computing the STFD matrix for unitary matrix estimation.

Note that, to avoid the inclusion of noise-only (t,f) points, autoterm point selection is limited to those (t,f) points where $\bar{D}_{\mathbf{xx}}(t,f)$ exceeds a certain threshold level. We denote the ratio of the threshold to the peak value of $\bar{D}_{\mathbf{zz}}(t,f)$ as γ_1.

[2]The use of absolute operator in the numerator allows us to include the crossterms of different component of the same source. They can be treated as autoterms with regard to the spatial information [13].

Orthogonalization

Although the array averaging is simple, it is likely to identify some false autoterm locations when the spatial correlation between the sources is high, i.e., the sources are closely spaced. In this case, the performance can be improved by orthogonalizing the received signal. Contrary to the prewhitening, the orthogonalization process does not change the relative signal powers. The $n \times m$ orthogonalization matrix $\mathbf{T}$ contains the conjugate transpose of the eigenvectors corresponding to the signal subspace, i.e., its ith row is given by [3]

$$\mathbf{t}_i = \mathbf{q}_i^H, i = 1, ..., n.$$

Denote $\mathbf{x}' = \mathbf{Tx}$ as the $n \times 1$ vector and x_i' as the ith element of $\mathbf{X}'$. With the (almost) prefect orthogonalization of the steering vectors (i.e., columns of $\mathbf{A}' = \mathbf{TA}$), Eqn. (13) becomes

$$\bar{D}_{\mathbf{x}'\mathbf{x}'}(t,f) = \frac{1}{n}\sum_{i=1}^{n} |\mathbf{a}_i'|^2 D_{d_i d_i}(t,f). \tag{16}$$

As such, the cross-source component are substantially mitigated. With orthogonalization, the criterion for autoterm selection then becomes

$$C_{x'}(t,f) = \frac{\left|\sum_i^n D_{x_i' x_i'}(t,f)\right|}{\sum_i^n |D_{x_i' x_i'}(t,f)|} = \frac{\left|trace\left(\bar{\mathbf{D}}_{\mathbf{x}'\mathbf{x}'}(t,f)\right)\right|}{\sum_i^n |D_{x_i' x_i'}(t,f)|}. \tag{17}$$

We denote the threshold levels for the orthogonalization method as α_2, β_2, and γ_2, respectively.

4.2. Time-Frequency Grouping

In reference [10], the subspace analyses of STFD matrices are presented for signals with clear t-f signatures, such as frequency modulated (FM) signals. It was shown that the offerings of using a STFD matrix instead of the covariance matrix are two-fold. First, the selection of autoterm t-f points, e.g., points on the source instantaneous frequency, where the power is concentrated, enhances the 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 matrix construction. In the presence of noise, the consideration of a subset of signal arrivals reduces perturbation in matrix eigen-decomposition, and it becomes essential when there is insufficient number of sensors.

In this section, we introduce the notion of signature grouping in the t-f domain. For sources whose TFDs are highly localized, the grouping of different autoterm t-f points are usually straightforward. Roughly, a TFD group is a continued or cluttered region of the autoterm (t,f) points obtained in Section 4.1. It is noted that, in a typical situation where there are more sensors than sources, the t-f

[3]Although the use of eigenvectors obtained from the STFD matrices may provide even better result, using the results obtained from the data covariance matrix is much simpler in this stage. Autoterm selection does not require high accuracy of the subspace estimation.

grouping is optional. On the other hand, the t-f grouping becomes essential if there are more sources than the number of array sensors.

5. MODIFIED SOURCE SEPARATION METHOD

In the method proposed in [5] and summarized in Section 3.2, although STFD matrices are used to estimate the unitary matrix $\mathbf{U}$, the covariance matrix is still used in the whitening process. As we discussed earlier, an estimate of the covariance matrix is often not as robust to noise as a well-defined STFD matrix. Particularly, when 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 are required to provide complete pre-whitening based on the covariance matrix. In this paper, we propose to use the STFD matrix in place of the covariance matrix $\mathbf{R}_{\mathbf{xx}}$ for whitening [11].

Denote $\mathbf{D}_{\mathbf{xx}}(t_1, f_1), ..., \mathbf{D}_{\mathbf{xx}}(t_K, f_K)$ as the STFD matrices constructed from K autoterm points being defined over a t-f region Ω and belonging to fewer $n_o \leq n$ signals. Also, denote, respectively, $\mathbf{d}^o(t)$ and $\dot{\mathbf{d}}(t)$ as the n_o and $n - n_o$ sources being present and absent in the t-f region Ω. 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 have

$$\mathbf{x}(t) = \mathbf{A}^o\mathbf{d}^o(t) + \dot{\mathbf{A}}\dot{\mathbf{d}}(t) + \mathbf{n}(t), \tag{18}$$

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

The incorporation of multiple (t, f) points through the joint diagonalization or t-f averaging process reduces the noise effect on the signal subspace estimation, as discussed in [5, 10]. For example, let $\tilde{\mathbf{D}}_{\mathbf{xx}}$ be the average STFD matrix of a set of STFD matrices defined over the same region Ω, but may incorporate a different t-f kernel, and denote $\hat{\sigma}^{tf}$ as the estimation of the noise-level eigenvalue of $\tilde{\mathbf{D}}_{\mathbf{xx}}$. Then,

$$\begin{aligned}\tilde{\mathbf{W}}\tilde{\mathbf{D}}_{\mathbf{yy}}\tilde{\mathbf{W}}^H &= \tilde{\mathbf{W}}(\tilde{\mathbf{D}}_{\mathbf{xx}} - \hat{\sigma}^{tf}\mathbf{I})\tilde{\mathbf{W}}^H \\ &= \tilde{\mathbf{W}}\mathbf{A}^o\tilde{\mathbf{D}}^o_{\mathbf{dd}}(\tilde{\mathbf{W}}\mathbf{A}^o)^H = \mathbf{I}.\end{aligned} \tag{19}$$

In (19), due to the ambiguity of signal complex amplitude in BSS, we have assumed for convenience and without loss generality that the averaged source TFD matrix $\tilde{\mathbf{D}}^o_{\mathbf{dd}}$ 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

$$\tilde{\mathbf{W}} = \left[(\lambda_1^{tf} - \hat{\sigma}^{tf})^{-1/2}\mathbf{h}_1^{tf}, ..., (\lambda_{n_o}^{tf} - \hat{\sigma}^{tf})^{-1/2}\mathbf{h}_{n_o}^{tf}\right]^H, \tag{20}$$

where $\lambda_1^{tf}, ..., \lambda_{n_o}^{tf}$ are the n_o largest eigenvalues of $\tilde{\mathbf{D}}_{\mathbf{xx}}$ and $\mathbf{h}_1^{tf}, ..., \mathbf{h}_{n_o}^{tf}$ are the corresponding eigenvectors of $\tilde{\mathbf{D}}_{\mathbf{xx}}$. Note that $\tilde{\mathbf{D}}^o_{\mathbf{dd}}$ and $\tilde{\mathbf{D}}_{\mathbf{yy}}$ are of reduced rank n_o instead of rank n, due to 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

$$\begin{aligned}\tilde{\mathbf{z}}(t) &= \tilde{\mathbf{W}}\mathbf{x}(t) = \tilde{\mathbf{W}}\mathbf{A}^o\mathbf{d}^o(t) + \tilde{\mathbf{W}}\dot{\mathbf{A}}\dot{\mathbf{d}}(t) + \tilde{\mathbf{W}}\mathbf{n}(t) \\ &= \tilde{\mathbf{U}}\mathbf{d}^o(t) + \tilde{\mathbf{W}}\dot{\mathbf{A}}\dot{\mathbf{d}}(t) + \tilde{\mathbf{W}}\mathbf{n}(t),\end{aligned} \tag{21}$$

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

$$\tilde{\mathbf{D}}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} = \tilde{\mathbf{W}}\tilde{\mathbf{D}}_{\mathbf{xx}}\tilde{\mathbf{W}}^H = \tilde{\mathbf{U}}\tilde{\mathbf{D}}^o_{\mathbf{dd}}\tilde{\mathbf{U}}^H = \mathbf{I}. \tag{22}$$

Eqn. (22) implies that the autoterm and crossterm TFDs averaged over the t-f region Ω 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 III. 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 [14].

6. SEPARATION OF MORE SOURCES THAN THE NUMBER OF SENSORS

When there are more sources than array sensors, orthogonalization of all signal mixing vectors becomes impossible. Therefore, even though the mixing vector can be estimated by using the source discrimination introduced in Section 5 by choosing $n_0 \leq m$, the signal waveforms remain inseparable. 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.

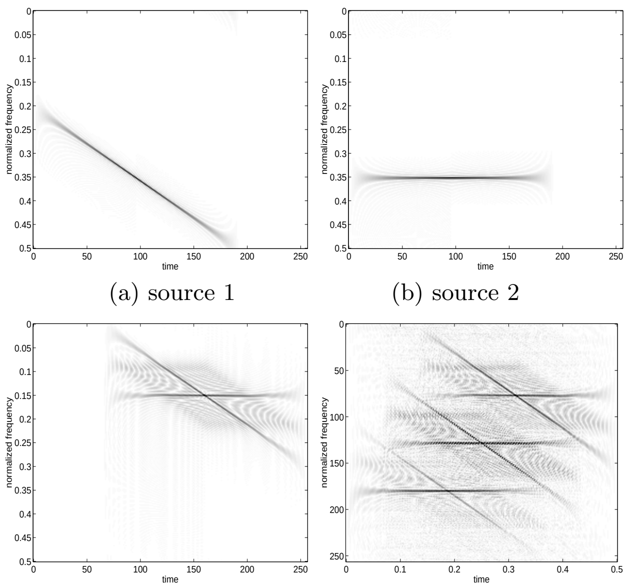
The grouping for this purpose is different from that discussed in Section 4.2. In Section 4.2, we only need to select several autoterm (t, f) points that provide sufficient information for the estimation of the mixing matrix of the sources. The selected autoterm region may not contain the full source waveform information. In this section, 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-source terms should be included for this purpose. This is achieved by constructing proper t-f masks. The mask at the kth t-f group, denoted as $M_k(t, f)$ should include the autoterm of the signals in this group and the crossterm between them, whereas the autoterms and crossterms of the signals not included in the group, and the crossterms between in-group and out-group signals, should be excluded.

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 ith sensor is used to synthesize the (mixed) signal waveforms at the kth group [15, 16]. The method proposed in Section 5 is then applied to each group. 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 $x_i(t)$ to the signal subspace that $\tilde{x}_i^{(k)}(t)$ spans.

7. SIMULATION RESULTS

7.1. Autoterm Selection and Grouping

In the first part of our simulation, 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 multi-component chirp signal. All the chirp components have the same magnitude. The number of data length is 256. The directions-of-arrival of the three signals are 45, 30, and 15 degrees, respectively. The WVDs of the three signals are plotted in Fig. 1(a)–(c). The WVD of the mixed signal at the first array sensor is shown in Fig. 1(d) with input SNR = 5 dB.



(a) source 1 (b) source 2

(c) source 3 (d) mixed signal

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

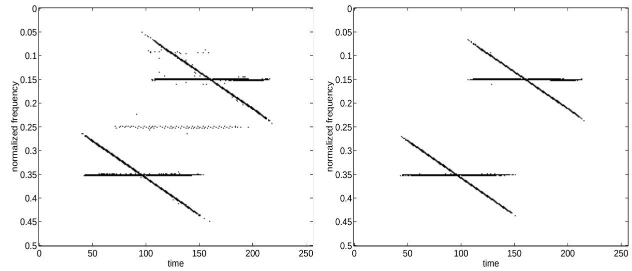
In Fig. 2, the results of pure autoterm selection are illustrated. While both plots show clear identification of the autoterm regions, the orthogonalization result is much "cleaner". From these results, we can form two groups with one including sources 1 and 2, and the other including only source 3.

7.2. Source Separation

The performance is evaluated by using the mean rejection level (MRL), defined as [5]

$$\text{MRL} = \sum_{p \neq q} E \left| (\hat{\mathbf{A}}^{\#} \mathbf{A})_{pq} \right|^2 \qquad (23)$$

where $\hat{\mathbf{A}}$ is the estimate of $\mathbf{A}$. A smaller value of the MRL implies better source separation results. An MRL lower than -10 dB is considered satisfactory [5].



(a) without orthogonalization ($\alpha_1 = 0.99$, $\gamma_1 = 0.15$)

(b) with orthogonalization ($\alpha_2 = 0.90$, $\gamma_2 = 0.15$)

Figure 2: Selected autoterm locations.

Fig. 3 shows the MRL versus the input SNR of the three sources (We assume that all chirp components have the same power. Therefore, the third signal with two chirp components has three dB higher SNR). The curves are calculated by averaging 100 independent trials with different noise sequences. The dashed line corresponds to the existing method where the covariance matrix $\mathbf{R_{xx}}$ is used for whitening, and the solid line corresponds to the modified method where the averaged STFD matrix $\tilde{\mathbf{D}}_{\mathbf{xx}}$ is used instead. The dash-dot 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) [4] 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. Fig. 3 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 existing method, 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.

7.3. Separating More Sources than Number of Sensors

In the second part of simulation, we use the same parameters used in Section 7.1, but the number of sensors is now only 2. The input SNR is fixed to 20 dB. To separate the three signal arrivals, we need to partition the t-f domain so that the maximum sources contained in each group does not exceed 2. In this example, we construct a mask that

[4] 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.

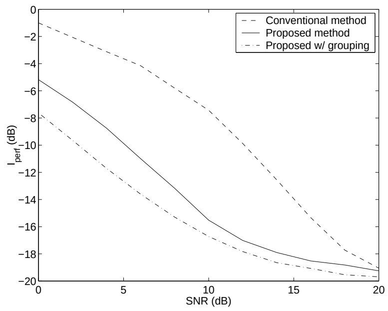
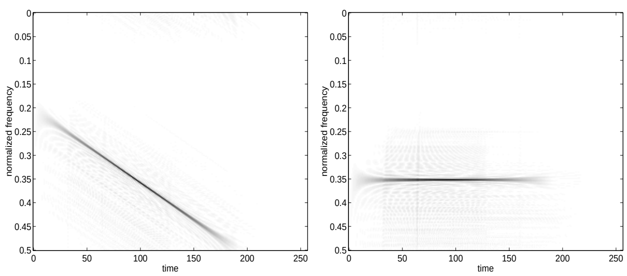


Figure 3: MRL versus input SNR (m=3).



(a) signal 1 (b) signal2

Figure 4: WVD of separated signals (m=2, SNR = 20 dB).

contains the first two sources and the procedure described in Section 6 is followed.

From the discussion in Section 6, 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 Fig. 4 the WVDs of the two separated signals. They are very close to the original source TFDs. The performance index, computed from 100 independent trials, is −21.23 dB, compared to −21.12 dB corresponding to the case in which only the two source signals are present and, therefore, no mask is applied.

8. 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 auto-sensor TFDs. Second, the STFD based BSS method has been modified to use multiple STFD matrices for pre-whitening. Third, t-f grouping for source discrimination is introduced for performance improvement and to separate more sources than the number of sensors.

9. REFERENCES

[1] L. Tong, Y. Inouye, and R-W. Liu, "Waveform-preserving blind estimation of multiple independent sources," *IEEE Trans. Signal Proc.*, vol. 41, no. 7, pp. 2461–2470, July 1993.

[2] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *IEE Proc. F*, vol. 140, no. 6, pp. 362–370, Dec. 1993.

[3] A. Belouchrani, K. A. Meraim, H.-F. Cardoso, and E. Moulines, "A blind source separation techniques using second order statistics," *IEEE Trans. Signal Proc.*, vol. 45, no. 2, pp. 434–444, Feb. 1997.

[4] O. Grellier and P. Comon, "Blind separation of discrete sources," *IEEE Signal Processing Letters*, vol. 5, no. 8, pp. 212–214, Aug. 1998.

[5] A. Belouchrani and M. G. Amin, "Blind source separation based on time-frequency signal representation," *IEEE Trans. Signal Proc.*, vol. 46, no. 11, pp. 2888–2898, Nov. 1998.

[6] 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, July 2000.

[7] A. Belouchrani, K. Abed-Meraim, M. G. Amin, and A. M. Zoubir, "Joint anti-diagonalization for blind source separation," in *Proc. ICASSP*, Salt Lake City, UT, pp. 2789–2792, May 2001.

[8] L. Giulieri, N. Thirion-Moreau, and P.-Y. Arques, "Blind sources separation based on bilinear time-frequency representations: a performance analysis," in *Proc. ICASSP*, Orlando, FL, pp. 1649–1652, May 2002.

[9] A. Holobar, C. Fevotte, C. Doncarli, and D. Zazula, "Single autoterms selection for blind source separation in time-frequency plane," in *Proc. EUSIPCO*, Toulouse, France, pp. 565–568, Sept. 2002.

[10] Y. Zhang, W. Mu, and M. G. Amin, "Subspace analysis of spatial time-frequency distribution matrices," *IEEE Trans. Signal Proc.*, vol. 49, no. 4, pp. 747–759, April 2001.

[11] Y. Zhang and M. G. Amin, "Blind separation of sources based on time-frequency distributions," in *Proc. ICASSP*, Istanbul, Turkey, June 2000.

[12] W. Mu, M. G. Amin, and Y. Zhang, "Bilinear signal synthesis in array processing," *IEEE Trans. Signal Processing*, vol. 51, no. 1, Jan. 2003.

[13] Y. Zhang and M. G. Amin, "Spatial averaging of time-frequency distributions for signal recovery in uniform linear arrays," *IEEE Trans. Signal Processing*, vol. 48, no. 10, pp. 2892-2902, Oct. 2000.

[14] Y. Zhang, M. G. Amin, and G. J. Frazer, "A new approach to FM jammer suppression for digital communications", *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, Rosslyn, VA, Aug. 2002.

[15] G. Boudreaux-Bartels and T. Parks, "Time-varying filtering and signal estimation using Wigner distribution synthesis techniques," *IEEE Trans. ASSP*, vol. ASSP-34, no. 3, pp. 442–451, June 1986.

[16] F. Hlawatsch and W. Krattenthaler, "Bilinear signal synthesis," *IEEE Trans. Signal Processing*, vol. 40, no. 2, pp. 352–363, Feb. 1992.