Globally Convergent Deflationary Instantaneous Blind Source Separation Algorithm for Digital Communication Signals

Alper Tunga Erdogan

***Abstract*—Recently an instantaneous blind source separation (BSS) approach that exploits the bounded magnitude structure of digital communications signals has been introduced. In this paper, we introduce a deflationary adaptive algorithm based on this criterion and provide its convergence analysis. We show that the resulting algorithm is convergent to one of the globally optimal points that correspond to perfect separation. The simulation examples related to the separation of digital communication signals are provided to illustrate the convergence and the performance of the algorithm.**

***Index Terms*—Adaptive filtering, blind source separation (BSS), independent component analysis, multiple-in multiple-out (MIMO) blind equalization, subgradient.**

I. Introduction

THE separation of sources from their mixtures is a problem of interest particularly in the area of communications. The existing methods to solve this problem exploit different pieces of available information that can be utilized to achieve separation. The use of the presumed non-gaussianity and the independence of sources is a popular and general approach that is applicable to a wide range of problems [2]–[9]. The algorithms in the BSS field were rooted form the pioneering works in blind deconvolution/equalization area such as [10]–[12].

The most widely used approach is the maximization of high order cumulants of the separator outputs (such as kurtosis) while keeping the power constant (which originated from [11]). This would be equivalent to the maximization of the high order norms of the linear mapping vectors between the sources and the separator outputs while keeping their 2-norms constant. The maximum value for the high order norm is achieved when it is equal to the 2-norm, i.e., when the mapping is a delta function, which corresponds to the perfect separation condition.

When the sources are bounded magnitude as in the case of digital communication signals, we can exploit this information, in addition to the independence of sources. The first BSS approach based on the boundedness of the input sources is due to Pham [13]. In his work, Pham reformulated the mutual information cost function in terms of quantile functions which enables the approximation of the cost function using order statistics. In the bounded magnitude case, this approximate cost function can be simplified to the (logarithm of) the difference between maximum and the minimum values of the separator output. In [13], an adaptive algorithm for simultaneous recovery of sources corresponding to the proposed cost function is also provided.

In [14], Vrins *et al.* proposed the minimization of the effective support length by for the simultaneous separation of bounded magnitude sources. Around the same time frame, in [1] and, [15], the minimization of the infinity norm of the separator output is proposed where it's proven that all the global minima for the cost function are the desired separators. In the same reference an adaptive algorithm based on subgradient optimization is proposed and through digital communication examples it is illustrated that the proposed method can achieve high signal to distortion ratio (SDR) levels for relatively short window sizes. The proposed minimization of the infinity norm of the output is essentially equivalent to minimization of the 1-norms of the overall linear mapping vectors between the sources and the separator outputs, under constant 2-norm constraint. Similar to cumulant based approach the minimum for the 1-norm is achieved for the delta vector. This observation provides the connection between the cumulant based approach widely used in literature and the approach proposed in [1] and [15]. The infinity norm minimization method proposed for bounded magnitude signals can be considered as a complementary approach to cumulant maximization based algorithms.

We should note that the minimization of the infinity norm of the output proposed in [15] is essentially equivalent to minimization of the support length (proposed in [13], [14]) for the case where the source distribution's minimum and maximum values are symmetrically located with respect to the origin.

The adaptive algorithm proposed in [1] is a two step procedure where the mixture data is whitened in the first step, and in the next step unitary separator is updated based on a subgradient search. The unitary constraint on the separator is imposed by projecting the subgradient updated matrix to the nearest unitary matrix (in Frobenius norm sense). Although this (symmetrically orthogonalized) adaptive algorithm is successful in achieving the simultaneous separation of multiple sources, its convergence analysis is not trivial due to the complication caused by the minimum distance (also referred as the symmetric) projection to unitary matrices.

In this paper, we propose the use of the deflation procedure for the imposition of the unitary constraints. The use of the deflation method enables relatively easier global convergence anal-

Manuscript received January 23, 2006; revised July 24, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Luc Vandendorpe. This work was supported in part by a TUBITAK Career Award under Contract 104E073.

The author is with the Electrical and Electronics Engineering Department, Koc University, 34450 Sariyer, Istanbul, Turkey (e-mail: alperdogan@ku.edu.tr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2007.893214

ysis for the overall adaptive separation algorithm. In particular, we can show that the proposed adaptive algorithm always converges to one of the globally optimal points, where each optimal point corresponds to a perfect separation condition. Only assumptions used in this convergence analysis are the richness of the available data window in terms of reflecting the ensemble behavior for the separator output and also a generic condition on the structure of the initial mapping between sources and the separator outputs.

The global convergence of the proposed algorithm is a desirable feature from the reliability point of view. Due to this property, achieving perfect separation via use of the algorithm is guaranteed. The issue of reliability is a current research focus in the BSS field and the convergence behaviors of the existing adaptive algorithms are under investigation [16]–[18]. For example, recently, the existence of false minima for information theoretic BSS approaches in the case of multimodal distributions is shown [16].

The organization of the paper is as follows: Section II provides the notation and the assumed setup for the blind source separation problem for bounded magnitude sources. The deflationary adaptive algorithm is provided in Section III. The convergence analysis of this algorithm is provided in Section IV. In Section V, we provide a numerical simulation example to illustrate the convergence behavior of the deflationary algorithm. Finally, Section VI is the conclusion.

II. Blind Source Separation Approach for Bounded Magnitude Sources

In the blind source separation setup that we consider throughout the paper:

- $s_1(k), s_2(k), \ldots, s_p(k)$ are the source signals. It is assumed that they are unity variance (without loss of generalization) stationary signals with zero mean, and they are mutually independent of each other. It is not required that each s_l is an i.i.d. sequence. We assume that each s_l has sufficiently rich variations such that
  - — autocorrelation matrix estimate obtained by time averaging reflects the true correlation matrix;
  - — the maximum value of the separator's output in time reflects the ensemble maximum.

  Furthermore, the source signals are considered to be bounded and complex symmetric in the sense that

$$\begin{aligned}\sup \Re e\{s_l\} &= \sup \mathrm{Im}\{s_l\} = -\inf \Re e\{s_l\} \\ &= -\inf \mathcal{I}m\{s_l\} = M \end{aligned} \tag{1}$$

  over the ensemble of s_l and M is the minimum upper-bound on the magnitude of real (and imaginary) parts of the sources. We assume the existence of corner points $(\pm M \pm jM)$ in the closure of the domains of the source distributions. The complex sources that satisfy these conditions are quite typical in digital communications applications (e.g., QAM constellations).
- The source signals are mixed with the memoryless multiple-in multiple-out (MIMO) system with transfer matrix $\mathbf{H}$. $\mathbf{H}$ is a $q \times p$ matrix which implies that this mixing system has q outputs denoted by $y_1(k), y_2(k), \ldots, y_q(k)$, i.e., at any instant k, we can write

$$\underbrace{\begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_q(k) \end{bmatrix}}_{\mathbf{y}(k)} = \mathbf{H} \underbrace{\begin{bmatrix} s_1(k) \\ s_2(k) \\ \vdots \\ s_p(k) \end{bmatrix}}_{\mathbf{s}(k)}. \tag{2}$$

  We assume that $q \geq p$, i.e., the number of mixtures is greater than or equal to the number of sources.
- The purpose of blind source separation is to estimate the source signals (independent components) $s_1(k), s_2(k), \ldots, s_p(k)$ from the observation sequences $y_1(k), y_2(k), \ldots, y_q(k)$ using a linear system with transfer matrix $\mathbf{W}^T$, i.e.

$$\mathbf{z}(k) = \mathbf{W}^T \mathbf{y}(k) \tag{3}$$

  where $\mathbf{z}(k) = [z_1(k) \;\; z_2(k) \;\; \ldots \;\; z_p(k)]^T$ contains the estimates of the original sources. $\mathbf{W}$ is obtained adaptively from the time samples (realizations) of $y_1(k), y_2(k), \ldots y_q(k)$. No *a priori* knowledge of $\mathbf{H}$ and no training sequences are assumed.

We assume that $\mathbf{W}$ is decomposed into two operators

$$\mathbf{W} = \mathbf{W}_{\mathrm{pre}} \boldsymbol{\Theta} \tag{4}$$

where $\mathbf{W}_{\mathrm{pre}}$ is a $q \times p$ whitening matrix such that

$$\underbrace{\begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_p(k) \end{bmatrix}}_{\mathbf{x}(k)} = \mathbf{W}_{\mathrm{pre}}^T \mathbf{y}(k) \tag{5}$$

is a white vector, i.e.

$$E(\mathbf{x}(k)\mathbf{x}^T(k)) = \mathbf{I}. \tag{6}$$

It is assumed that the whitening matrix $\mathbf{W}_{\mathrm{pre}}$ is obtained adaptively possibly through the factorization of the empirical covariance matrix $\hat{\mathbf{R}}_{\mathbf{y}}$ of $\mathbf{y}(k)$.

As a result, we can write

$$\begin{aligned} \mathbf{z}(k) &= \mathbf{W}^T \mathbf{y}(k) && (7) \\ &= \boldsymbol{\Theta}^T \underbrace{\mathbf{W}_{\mathrm{pre}}^T \mathbf{H}}_{\mathbf{C}} \mathbf{s}(k) && (8) \\ &= \boldsymbol{\Theta}^T \mathbf{x}(k). && (9) \end{aligned}$$

Here

$$\mathbf{C} = \mathbf{W}_{\mathrm{pre}}^T \mathbf{H} \tag{10}$$

is the combined mixing and whitening transfer functions, and therefore, we can write

$$\mathbf{x}(k) = \mathbf{C}\mathbf{s}(k). \tag{11}$$

Note that $\mathbf{x}(k)$ is a white vector with identity covariance and $\mathbf{C}$ is a unitary matrix. Although the elements of $\mathbf{x}(k)$ are uncorrelated, they are not necessarily independent. Therefore, our goal is to obtain a unitary matrix $\mathbf{\Theta}$ as some column permuted version of $\mathbf{C}$ multiplied by a diagonal matrix with unity magnitude complex entries such that

$$\begin{aligned} \mathbf{z}(k) &= \mathbf{\Theta}^T\mathbf{x}(k) && (12)\\ &= (\bar{\mathbf{C}}\mathbf{E}\mathbf{D})^T\mathbf{C}\mathbf{s}(k) && (13)\\ &= \mathbf{D}\mathbf{E}^{\mathbf{T}}\mathbf{s}(k) && (14)\end{aligned}$$

where $\mathbf{E}$ is a permutation matrix and $\mathbf{D}$ is a diagonal matrix with unit-magnitude complex entries, both of which reflect the unavoidable ambiguities in blind source separation [8].

In the BSS approach for bounded magnitude sources in [1], the following optimization problem is proposed to obtain $\mathbf{\Theta}$

$$\begin{aligned} &\text{minimize } \sup \|\mathcal{R}e\{\mathbf{z}(k)\}\|_\infty \quad (\text{Problem 1})\\ &\text{s.t. } \mathbf{\Theta}^H\mathbf{\Theta} = I\end{aligned}$$

which is the minimization of the supremum of the infinity norm of the real part of the output vector $\mathbf{z}(k)$ over the ensemble. The following theorem [1] shows that the global minimizers of the above optimization setting are the desired solutions for the source separation problem.

Theorem 1: Given the setup above and let $\mathbf{G} = \mathbf{\Theta}^T\mathbf{C}$ be defined as the overall mapping between $\mathbf{s}$ and $\mathbf{z}$, then $\mathbf{\Theta}_{\text{opt}}$ is a global minimizer of the Problem 1 if and only if the corresponding $\mathbf{G}_{\text{opt}}$ has the form

$$\mathbf{G}_{\text{opt}} = \mathbf{\Theta}_{\text{opt}}^T\mathbf{C} = \mathbf{D}\mathbf{E} \quad (15)$$

where $\mathbf{E}$ is a permutation matrix and $\mathbf{D}$ is a diagonal matrix of the form

$$\mathbf{D} = \begin{bmatrix} e^{j\frac{\pi}{2}k_1} & 0 & \dots & 0\\ 0 & e^{j\frac{\pi}{2}k_2} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ \dots & \dots & \dots & e^{j\frac{\pi}{2}k_p}\end{bmatrix} \quad (16)$$

where $k_1, k_2, \dots, k_p$ are integers.

III. Deflationary Adaptive Algorithm

In the previous section, we introduced the optimization problem (Problem 1) whose global minimizers are the desired separators. In this section we provide a deflationary adaptive algorithm which is based on this criterion. The deflationary approach is essentially the sequential recovery of the sources, and in [19] and [4] it was used for kurtosis based cost function. In this paper, we use the deflationary approach for the infinity norm based cost function proposed for magnitude bounded sources.

In the adaptive implementation, we assume that we have a finite window of whitened mixture samples, i.e., $\{\mathbf{x}(k) : k \in \{0, 1, \dots, \Omega-1\}\}$. Here we make a certain ergodicity assumption that these time samples are rich enough to reflect the ensemble behavior in terms of generating the maximum magnitude sample for any choice of separator.

For the deflationary approach, as we consider sequential recovery of each source, the algorithm consists of p consecutive phases where at the ith phase $\mathbf{\Theta}_i$, which is the ith column of $\mathbf{\Theta}$, is obtained through subgradient iterations:

For the first phase, we consider the optimization problem

$$\begin{aligned} &\text{minimize } f_1(\mathbf{\Theta}_1)\\ &\text{s.t. } \mathbf{\Theta}_1^H\mathbf{\Theta}_1 = 1\end{aligned}$$

where

$$f_1(\mathbf{\Theta}_1) = \max_{n\in\{0,1,\dots,\Omega-1\}} \left| \mathcal{R}e\left\{ \underbrace{\mathbf{\Theta}_1^T\mathbf{x}(n)}_{z_1(n)} \right\} \right|. \quad (17)$$

The subdifferential set corresponding to this cost function is given by

$$\begin{aligned}&\partial f_1(\mathbf{\Theta}_1)\\ &\quad = \mathbf{Co}\Big\{\text{sign}(\mathcal{R}e\{z_1(n)\})\bar{\mathbf{x}}(n) \quad | \quad |z_1(n)| = f_1(\mathbf{\Theta}_1)\Big\}. \quad (18)\end{aligned}$$

Here Co stands for convex hull operation. In the subgradient optimization approach, the search vector is updated using any vector chosen from the subgradient set, where we pick one of the vectors given in (18) (whose convex hull is the subdifferential set). Therefore, we can write the algorithm iterations to obtain $\mathbf{\Theta}_1$ as

$$\underline{\mathbf{\Theta}}_1^{(i+1)} = \mathbf{\Theta}_1^{(i)} - \mu_1^{(i)}\text{sign}\left(\mathcal{R}e\left\{z_1\left(n^{(i)}\right)\right\}\right)\bar{\mathbf{x}}\left(n^{(i)}\right) \quad (19)$$

$$\mathbf{\Theta}_1^{(i)} = \frac{\underline{\mathbf{\Theta}}_1^{(i)}}{\left\|\underline{\mathbf{\Theta}}_1^{(i)}\right\|} \quad (20)$$

where

- $n^{(i)}$ is the index for which maximum magnitude real output is achieved at the ith iteration;
- $\mu^{(i)} = (f_1(\mathbf{\Theta}_\mathbf{1}^{(i)}) - M)/(2\|\mathbf{x}(n^{(i)})\|^2)$ is the step-size. The choice of step size is based on the relaxation rule proposed in the framework of the subgradient optimization [20]–[24]. As we will show later, the selection of step size in this form plays a crucial role in the global convergence analysis.

Once $\mathbf{\Theta_1}$ is obtained through the subgradient iterations above, we form the projection matrix $\mathbf{V}^{(1)} = \mathbf{I} - \mathbf{\Theta}_1\mathbf{\Theta}_1^H$, which projects a given vector to the space orthogonal to $\mathbf{\Theta}_1$.

We can write the kth phase of the deflation algorithm as follows.

- Initialize the $\mathbf{\Theta}_k$ vector using

$$\mathbf{\Theta}_k^{(0)} = \mathbf{V}_1^{(k-1)}$$

which is the first column of the $\mathbf{V}^{(k-1)}$ matrix, and this choice guarantees that $\mathbf{\Theta}_k^{(0)}$ is orthogonal to previous $\mathbf{\Theta}_i$'s for $i = 1, \dots, k-1$.

- Iteratively obtain $\boldsymbol{\Theta}_k$, which minimizes the cost function

$$f_k(\boldsymbol{\Theta}_k) = \max_{n\in\{0,1,\ldots,\Omega-1\}} \left| \mathcal{R}e\left\{ \underbrace{\boldsymbol{\Theta}_k^T \mathbf{x}(n)}_{z_k(n)} \right\} \right| \quad (21)$$

using subgradient iterations

$$\underline{\boldsymbol{\Theta}}_k^{(i+1)} = \boldsymbol{\Theta}_k^{(i)} - \mu_1^{(i)} \mathrm{sign}\left(\mathcal{R}e\left\{z_k(n^{(i)})\right\}\right) \times \mathbf{V}^{(k-1)} \bar{\mathbf{x}}^{(k)}\left(n^{(i)}\right) \quad (22)$$

$$\boldsymbol{\Theta}_k^{(i)} = \frac{\underline{\boldsymbol{\Theta}}_k^{(i)}}{\left\|\underline{\boldsymbol{\Theta}}_k^{(i)}\right\|} \quad (23)$$

where

$$z_k\left(n^{(i)}\right) = \boldsymbol{\Theta}_k^{(i)^T} \mathbf{x}^{(k)}\left(n^{(i)}\right) \quad (24)$$

$n^{(i)}$ is the index for maximum real magnitude z_k and

$$\mu^{(i)} = \frac{f_k\left(\boldsymbol{\Theta}_\mathbf{k}^{(i)}\right) - M}{2\left\|\bar{\mathbf{V}}^{(k-1)}\mathbf{x}\left(n^{(i)}\right)\right\|^2}. \quad (25)$$

Note that multiplication with $\mathbf{V}^{(k-1)}$ constrains $\boldsymbol{\Theta}_k$ to the space orthogonal to $\boldsymbol{\Theta}_1, \ldots \boldsymbol{\Theta}_{k-1}$.
- Update the orthogonal projection matrix

$$\mathbf{V}^{(k)} = \mathbf{V}^{(k-1)}\left(\mathbf{I} - \boldsymbol{\Theta}_k \boldsymbol{\Theta}_k^H\right). \quad (26)$$

The algorithm presented above is repeated for $k = 1$ to $k = (p-1)$. After the $p-1$th phase, $\boldsymbol{\Theta}_p$ can be simply chosen as equivalent to a column of the $\mathbf{V}^{(p-1)}$ matrix.

We should note that the deflationary approach has inherent error accumulation problem in practical applications. Since we proceed from one phase from to the other without complete convergence due to finite number of iterations, existence of noise and nonlinear factors, the assumption about elimination of previous sources will fail, and this will create extra noise effect for the following stage, which will accumulate throughout the multiple stages of the algorithm. Despite this undesired feature, the deflation approach we outlined above leads to framework where we can obtain global convergence under "ideal" conditions.

IV. Convergence Analysis of the Deflationary Adaptive Algorithm

In this section, we analyze the convergence behavior of the adaptive algorithm we introduced in the previous section. In particular, we show that the algorithm converges to one of the global minima corresponding to perfect separation.

We first start by writing the update equation for the kth phase

$$\underline{\boldsymbol{\Theta}}_k^{(i+1)} = \boldsymbol{\Theta}_k^{(i)} - \mu^{(i)} \mathrm{sign}\left(z_k\left(n^{(i)}\right)\right) \mathbf{V}^{(k-1)} \bar{\mathbf{x}}^{(k)}\left(n^{(i)}\right) \quad (27)$$

$$= \boldsymbol{\Theta}_k^{(i)} - \mu^{(i)} \mathrm{sign}\left(z_k\left(n^{(i)}\right)\right) \bar{\mathbf{C}}^{(k)} \bar{\mathbf{s}}\left(n^{(i)}\right) \quad (28)$$

where

$$\bar{\mathbf{C}}^{(k)} = \mathbf{V}^{(k-1)} \bar{\mathbf{C}}. \quad (29)$$

We should note that $\mathbf{C}^{(k)}$ is equivalent to $\mathbf{C}$ except $k-1$ of its columns are equal to zero

$$\bar{\mathbf{C}}^{(k)} = \mathbf{V}^{(k-1)} \bar{\mathbf{C}} \quad (30)$$

$$= \left(\mathbf{I} - \boldsymbol{\Theta}_{k-1}\boldsymbol{\Theta}_{k-1}^H\right)\ldots\left(\mathbf{I} - \boldsymbol{\Theta}_2\boldsymbol{\Theta}_2^H\right)\left(\mathbf{I} - \boldsymbol{\Theta}_1\boldsymbol{\Theta}_1^H\right)\bar{\mathbf{C}}. \quad (31)$$

Here, as we will prove

$$\boldsymbol{\Theta}_l = \bar{\mathbf{C}}_{m_l} e^{j\frac{\pi}{4}u_l} \quad (32)$$

where $m_l \in 1, 2, \ldots, p$ and u_l is an integer. Therefore, plugging this expression for $\boldsymbol{\Theta}_l$ in (31), we can obtain that $\mathbf{C}^{(k)}$ is equal to $\mathbf{C}$ with the columns having indexes $\{m_l, l = 1, \ldots, k-1\}$ equal to zero.

We define the effective channel between the sources and the separator $\boldsymbol{\Theta}_k$'s output as

$$\mathbf{G}_k = \boldsymbol{\Theta}_k^T \mathbf{C}^{(k)}. \quad (33)$$

Note that $\mathbf{G}_k$ is a $1 \times p$ vector and its components at index locations $\{m_l, l = 1, \ldots, k-1\}$ are zero. If we define

$$\mathcal{G}_k = [\mathcal{R}e\{\mathbf{G}_k\} \quad \mathcal{I}m\{\mathbf{G}_k\}] \quad (34)$$

then the real component of the output can be written as

$$\mathcal{R}e\{z_k(n)\} = \mathcal{R}e\{\mathbf{G}_k \mathbf{s}(n)\} \quad (35)$$

$$= \mathcal{G}_k \underbrace{\begin{bmatrix} \mathcal{R}e\{\mathbf{s}(n)\} \\ -\mathcal{I}m\{\mathbf{s}(n)\} \end{bmatrix}}_{\breve{\mathbf{s}}(n)}. \quad (36)$$

Multiplying both sides of (28) (from the left) with $\mathbf{C}^{(k)^T}$ (and taking transpose of both sides), we obtain a recursion for $\mathbf{G}_k$

$$\underline{\mathbf{G}}_k^{(i+1)} = \mathbf{G}_k^{(i)} - \mu^{(i)} \mathrm{sign}\left(\mathcal{R}e\left\{z_k\left(n^{(i)}\right)\right\}\right) \mathbf{s}^H\left(n^{(i)}\right) \times \underbrace{\mathbf{C}^{(k)^H}\mathbf{C}^{(k)}}_{\mathbf{J}^{(k)}} \quad (37)$$

where $\mathbf{J}^{(k)}$ is equivalent to a modified form of identity matrix where the diagonal entries with indexes $\{m_l, l = 1, \ldots, k-1\}$ are equal to zero. This update equation is equivalent to

$$\underline{\mathcal{G}}_k^{(i+1)} = \mathcal{G}_k^{(i)} - \mu^{(i)} \mathrm{sign}\left(\mathcal{R}e\left\{z_k\left(n^{(i)}\right)\right\}\right) \breve{\mathbf{s}}^T\left(n^{(i)}\right) \mathcal{J}^{(k)} \quad (38)$$

where

$$\mathcal{J}^{(k)} = \begin{bmatrix} \mathbf{J}^{(k)} & 0 \\ 0 & \mathbf{J}^{(k)} \end{bmatrix}. \quad (39)$$

Based on (36), we deduce that at index $n^{(i)}$ where the maximum magnitude $|\mathcal{R}e\{z_k\}|$ is achieved

$$\breve{\mathbf{s}}\left(n^{(i)}\right) = \pm M \mathrm{sign}\left(\mathcal{G}_k^{(i)}\right)^T \quad (40)$$

where $\pm$ in the above equation would be the sign of $\mathcal{R}e\{z_k(n^{(i)})\}$, sign( ) of a vector is the sign() of its elements and $\text{sign}(0) = 0$. Note that the components of $\breve{\mathbf{s}}$ corresponding to the zero components of $\mathbf{G}_k$ at index locations $\{m_l, l = 1, \ldots, k-1\}$ could actually be arbitrarily as they have no effect on the output, and they have no effect on the update since they are to be converted to zero after multiplication with $\mathcal{J}^{(k)}$. Using (40), we can write

$$\underline{\mathcal{G}}_k^{(i+1)} = \mathcal{G}_k^{(i)} - \mu^{(i)}\text{sign}\left(\mathcal{R}e\{z_k\}\left(\pm M\text{sign}\left(\mathcal{G}_k^{(i)}\right)\right)\mathcal{J}^{(k)}\right. \tag{41}$$

$$= \mathcal{G}_k^{(i)} - \mu^{(i)}M\text{sign}\left(\mathcal{G}_k^{(i)}\right)\mathcal{J}^{(k)}. \tag{42}$$

Based on this equation we observe that, for the vth component of $\underline{\mathcal{G}}_k$, where $v \notin \{m_l, 2m_l, l = 1, \ldots, k-1\}$, we can write

$$\left|\underline{\mathcal{G}}_{k,v}^{(i+1)}\right| = \left|\left|\mathcal{G}_{k,v}^{(i)}\right| - \mu^{(i)}M\right|. \tag{43}$$

Here

$$\mu^{(i)} = \frac{\left|\mathcal{R}e\left\{z(n^{(i)})\right\}\right| - M}{2\left\|\bar{\mathbf{V}}^{(k-1)}\mathbf{u}\left(n^{(i)}\right)\right\|_2^2} \tag{44}$$

$$= \frac{M\left\|\mathcal{G}_k^{(i)}\right\|_1 - M\left\|\mathcal{G}_k^{(i)}\right\|_2}{2\left\|\mathbf{C}^{(k)}\mathbf{s}\left(n^{(i)}\right)\right\|_2^2} \tag{45}$$

$$= \frac{M\left(\left\|\mathcal{G}_k^{(i)}\right\|_1 - \left\|\mathcal{G}_k^{(i)}\right\|_2\right)}{4(p-k+1)M^2} \tag{46}$$

$$= \frac{\left\|\mathcal{G}_k^{(i)}\right\|_1 - \left\|\mathcal{G}_k^{(i)}\right\|_2}{4(p-k+1)M} \tag{47}$$

$$\leq \frac{\left\|\mathcal{G}_k^{(i)}\right\|_1 - \left\|\mathcal{G}_k^{(i)}\right\|_\infty}{4(p-k+1)M} \tag{48}$$

$$= \frac{\left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_1}{4(p-k+1)M} \tag{49}$$

$$\leq \frac{\left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty}{2M} \tag{50}$$

where $\tilde{\mathcal{G}}_k^{(i)}$ is the vector obtained by deleting the maximum magnitude element of $\mathcal{G}_k^{(i)}$. In the manipulations above, the numerator of (45) follows from (36) and (40) and from the fact that

$$\left\|\mathcal{G}_k^{(i)}\right\|_2 = \left\|\mathbf{G}_k^{(i)}\right\|_2 = 1. \tag{51}$$

The denominator of (46) can be justified as

$$\begin{aligned}&\left\|\mathbf{C}^{(k)}\mathbf{s}\left(n^{(i)}\right)\right\|_2^2\\ &\quad= \mathbf{s}^H\left(n^{(i)}\right)\mathbf{C}^{(k)^H}\mathbf{C}^{(k)}\mathbf{s}\left(n^{(i)}\right) && (52)\\ &\quad= \mathbf{s}^H\left(n^{(i)}\right)\mathbf{J}^{(k)}\mathbf{s}\left(n^{(i)}\right) && (53)\end{aligned}$$

and since $\mathbf{J}^{(k)}$ has $p-k+1$ nonzero diagonal entries and $s(n^{(i)})$ consists of entries $\pm M \pm Mj$ [due to (40)]

$$\left\|\mathbf{C}^{(k)}\mathbf{s}\left(n^{(i)}\right)\right\|_2^2 = 2M^2(p-k+1) \tag{54}$$

easily follows.

Inequality in (48) is due to norm ordering $\|\mathcal{G}_k^{(i)}\|_\infty \leq \|\mathcal{G}_k^{(i)}\|_2$, the equality in (49) follows directly from the definition of $\tilde{\mathcal{G}}_k^{(i)}$ and the inequality in (50) is due to inequality between 1 and ∞ norms.

Note that $\|\tilde{\mathcal{G}}_k^{(i)}\|_\infty$ is nothing but the second peak value of the vector $\mathcal{G}_k^{(i)}$. Since $\mu^{(i)}M$ is bounded by the halve of the second peak value, i.e., $(\|\tilde{\mathcal{G}}_k^{(i)}\|_\infty/2)$, after the update in (43), which subtracts the same constant from all of the freely changing components, the peak location and the second peak value location will not change, and the reduction level for both will be the same. However, when we look at their ratios, we obtain

$$\frac{\left\|\underline{\mathcal{G}}_k^{(i+1)}\right\|_\infty}{\left\|\tilde{\underline{\mathcal{G}}}_k^{(i+1)}\right\|_\infty} = \frac{\left\|\mathcal{G}_k^{(i)}\right\|_\infty - \mu^{(i)}M}{\left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty - \mu^{(i)}M} > \frac{\left\|\mathcal{G}_k^{(i)}\right\|_\infty}{\left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty} \tag{55}$$

for $\mu^{(i)} > 0$. Note that since

$$\frac{\left\|\mathcal{G}_k^{(i+1)}\right\|_\infty}{\left\|\tilde{\mathcal{G}}_k^{(i+1)}\right\|_\infty} = \frac{\left\|\underline{\mathcal{G}}_k^{(i+1)}\right\|_\infty}{\left\|\tilde{\underline{\mathcal{G}}}_k^{(i+1)}\right\|_\infty} \tag{56}$$

we can write

$$\frac{\left\|\mathcal{G}_k^{(i+1)}\right\|_\infty}{\left\|\tilde{\mathcal{G}}_k^{(i+1)}\right\|_\infty} > \frac{\left\|\mathcal{G}_k^{(i)}\right\|_\infty}{\left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty}. \tag{57}$$

We now show that given

$$\frac{\left\|\mathcal{G}_k^{(0)}\right\|_\infty}{\left\|\tilde{\mathcal{G}}_k^{(0)}\right\|_\infty} > 1 \tag{58}$$

which is the generic and stable case, $(\|\mathcal{G}_k^{(i)}\|_\infty/\|\tilde{\mathcal{G}}_k^{(i)}\|_\infty)$ grows unboundedly, and therefore, $\mathcal{G}_k^{(i)}$ converges to a delta function:

Defining

$$\rho^{(i)} = \frac{\left\|\mathcal{G}_k^{(i)}\right\|_\infty}{\left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty} \tag{59}$$

we write

$$\rho^{(i+1)} = \frac{\left\|\mathcal{G}_k^{(i)}\right\|_\infty - \mu^{(i)}M}{\left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty - \mu^{(i)}M} \tag{60}$$

$$= \rho^{(i)}\underbrace{\frac{1 - r^{(i)}/\rho^{(i)}}{1 - r^{(i)}}}_{t(r^{(i)})} \tag{61}$$

where

$$r^{(i)} = \frac{\mu^{(i)}M}{\left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty} \tag{62}$$

$$= \frac{\left\|\mathcal{G}_k^{(i)}\right\|_1 - \left\|\mathcal{G}_k^{(i)}\right\|_2}{4(p-k+1)\left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty} \tag{63}$$

$$\leq \frac{1}{2}. \tag{64}$$

Here, the inequality (64) is due to the fact that [from (50)]

$$\mu^{(i)} M \leq \frac{\left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty}{2}. \tag{65}$$

Letting $\grave{\mathcal{G}}_k^{(i)}$ denote the $1 \times (2p-2)$ vector obtained by removing peak and the second peak from $\mathcal{G}_k^{(i)}$, we can write

$$\begin{aligned}
&\left\|\mathcal{G}_k^{(i)}\right\|_1 - \left\|\mathcal{G}_k^{(i)}\right\|_2 \\
&= \left\|\mathcal{G}_k^{(i)}\right\|_\infty + \left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty + \left\|\grave{\mathcal{G}}_k^{(i)}\right\|_1 - \left\|\mathcal{G}_k^{(i)}\right\|_2 && (66)\\
&\geq \left\|\mathcal{G}_k^{(i)}\right\|_\infty + \left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty + \left\|\grave{\mathcal{G}}_k^{(i)}\right\|_1 \\
&\quad - \left(\sqrt{\left\|\mathcal{G}_k^{(i)}\right\|_\infty^2 + \left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty^2} + \left\|\grave{\mathcal{G}}_k^{(i)}\right\|_2\right) && (67)\\
&= \left\|\mathcal{G}_k^{(i)}\right\|_\infty + \left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty - \sqrt{\left\|\mathcal{G}_k^{(i)}\right\|_\infty^2 + \left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty^2} \\
&\quad + \left\|\grave{\mathcal{G}}_k^{(i)}\right\|_1 - \left\|\grave{\mathcal{G}}_k^{(i)}\right\|_2 && (68)\\
&\geq \left\|\mathcal{G}_k^{(i)}\right\|_\infty + \left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty - \sqrt{\left\|\mathcal{G}_k^{(i)}\right\|_\infty^2 + \left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty^2} && (69)
\end{aligned}$$

where the inequality (67) is due to the triangle inequality, and the inequality in (69) is due to the inequality between 1 and 2 norms.

Therefore

$$r^{(i)} \geq \frac{\left\|\mathcal{G}_k^{(i)}\right\|_\infty + \left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty - \sqrt{\left\|\mathcal{G}_k^{(i)}\right\|_\infty^2 + \left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty^2}}{4(p-k+1)\left\|\tilde{\mathcal{G}}_k^{(i)}\right\|_\infty} \tag{70}$$

$$= \frac{1 + \rho^{(i)} - \sqrt{1 + {\rho^{(i)}}^2}}{4(p-k+1)} \tag{71}$$

from which we can write

$$\rho^{(i+1)} \geq \rho^{(i)} \underbrace{\left(\frac{1 - \frac{1+\rho^{(i)} - \sqrt{1+{\rho^{(i)}}^2}}{(4p-k+1)\rho^{(i)}}}{1 - \frac{1+\rho^{(i)} - \sqrt{1+{\rho^{(i)}}^2}}{4(p-k+1)}}\right)}_{\gamma\left(\rho^{(i)}\right)} \tag{72}$$

since $t(r^{(i)})$ is an increasing function of $r^{(i)}$ for $r^{(i)} \in [0, 1/2]$.

Note that, for $\rho > 1, \gamma(\rho)$ is an increasing function since its derivative, see (73) at the bottom of the page, is positive for $\rho > 1$.

Moreover, $\gamma(\rho^{(0)}) > 1$ for $\rho^{(0)} > 1$. Therefore

$$\gamma\left(\rho^{(i)}\right) > \gamma\left(\rho^{(0)}\right) > 1. \tag{74}$$

Consequently

$$\rho^{(i)} > \rho^{(0)} \gamma\left(\rho^{(0)}\right)^i \tag{75}$$

and, therefore

$$\lim_{i \to \infty} \rho^{(i)} = \infty. \tag{76}$$

Since we are normalizing the norm of the $\mathbf{G}$ to 1, this implies that the second peak value of the $\mathcal{G}$ and all its remaining components converge to zero which means that $\mathcal{G}_k^{(i)}$ converges to

$$\sigma \mathbf{e}_{n_k} \tag{77}$$

where $\mathbf{e}_{n_k}$ is a standard basis vector, i.e., the n_kth column of the $2p \times 2p$ identity matrix, and σ is equal to either 1 or -1. As a result, $\mathbf{G}_k^{(i)}$ converges to a vector with only one nonzero entry whose value is from the set $\{1, -1, j, -j\}$. Consequently, $\mathbf{\Theta}_k$ converges to $\bar{\mathbf{C}}_{m_k} e^{j(\pi u_k/2)}$, where u_k is an integer. Furthermore, by construction, $\mathbf{\Theta}_k$'s form an orthonormal set.

Note that the condition in (IV) means that $\mathcal{G}_k^{(0)}$ has a single peak. This is generically true as assuming a random channel $\mathbf{C}$ and a random initialization for $\Theta_k^{(0)}$ having a resulting $\mathcal{G}_k^{(0)}$ has zero probability. Even if the assumption in () is not met, having more than one multiple peak points (with equivalent levels) is not a stable condition and a slight disturbance in $\mathbf{\Theta}_k$ would also disturb the peak balance condition and the algorithm would converge to one of its global points afterwards.

Note that the algorithm that we provided above assumed no noise. In [25] it is shown that the stochastic version of the bounded magnitude cost function can be defined as

$$E\left(\max_{n \in \{0,1,\ldots,\Omega-1\}} \left|\mathcal{R}e\left\{\underbrace{\mathbf{\Theta}_k^T \mathbf{x}(n)}_{z_k(n)}\right\}\right|\right) \tag{78}$$

and it can be shown that the perfect separation solutions are the stationary points of this cost function. This fact is useful in terms of evaluating the behavior of the proposed approach in noisy environments. In the next section, we also provided some examples incorporating the effects of background noise.

$$\gamma'(\rho) = \frac{\frac{\sqrt{1+\rho^2}-1}{(4p-k+1)\rho^2\sqrt{1+\rho^2}} \frac{1+\rho-\sqrt{1+\rho^2}}{4p-k+1} + \frac{1}{4p-k+1}\left(1 - \frac{\rho}{\sqrt{1+\rho^2}}\right)\frac{1+\rho-\sqrt{1+\rho^2}}{(4p-k+1)\rho}}{\left(\frac{1+\rho-\sqrt{1+\rho^2}}{4p-k+1}\right)^2} \tag{73}$$

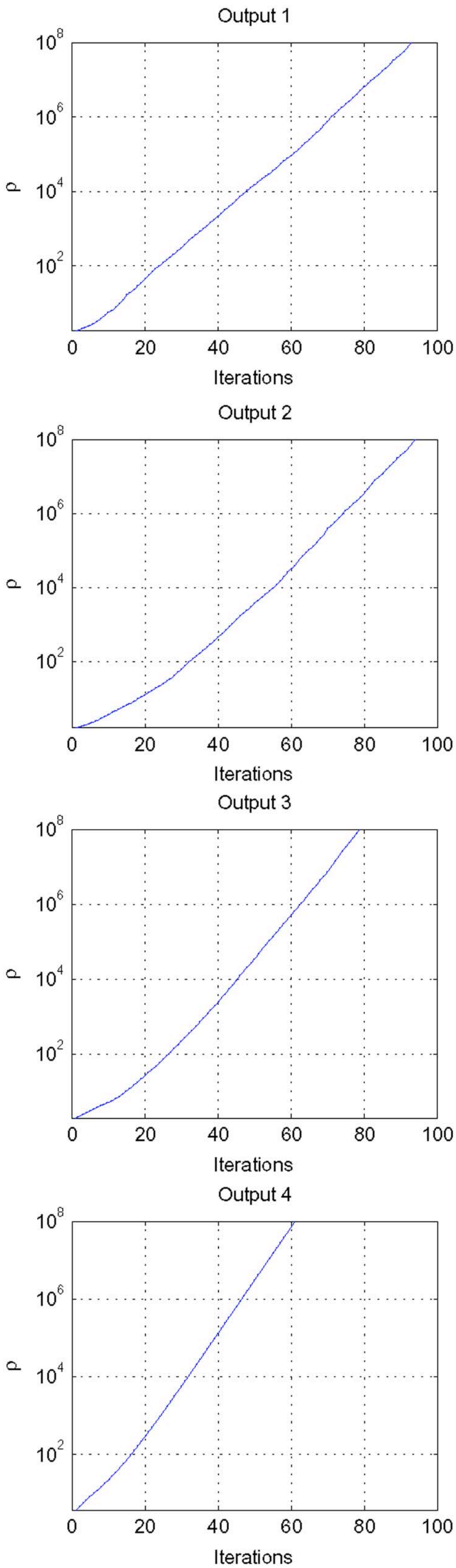

Fig. 1. ρ as a function of iterations for different separator outputs.

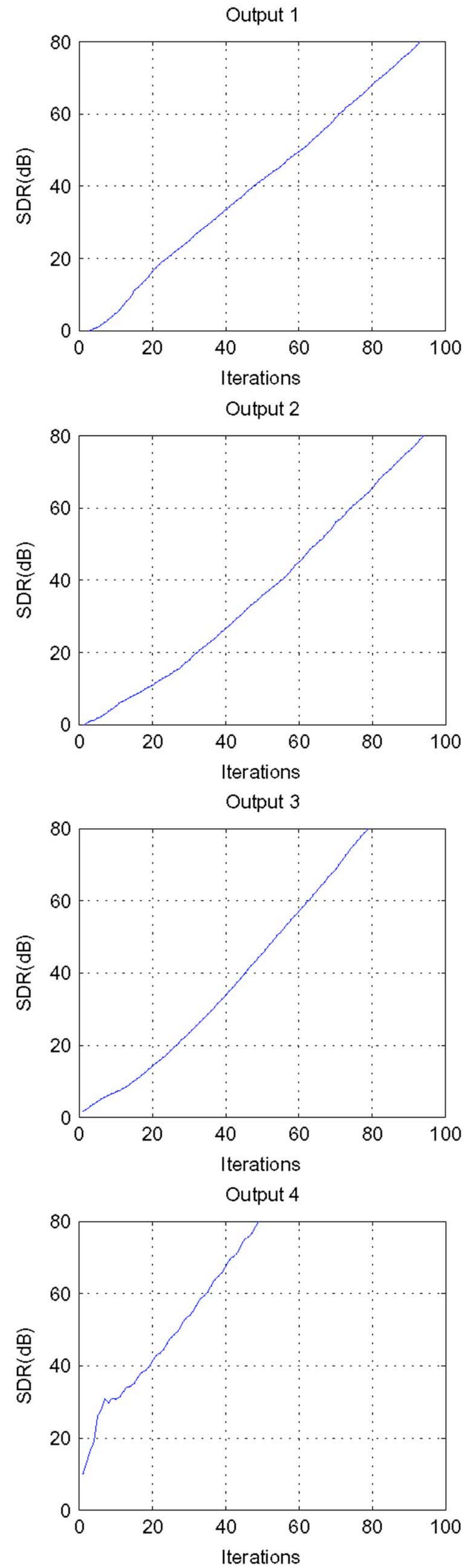

Fig. 2. SDR Convergence curves for the example.

V. Simulation Examples

In this section, we first provide a simulation example to illustrate the convergence of the proposed deflationary source separation algorithm. For the simulation setup, we consider five sources. Sources are assumed to be 64-QAM digital communications signals. We randomly generate a unitary channel for the effective channel $\mathbf{C}$. We used a sample size of $\Omega = 2000$.

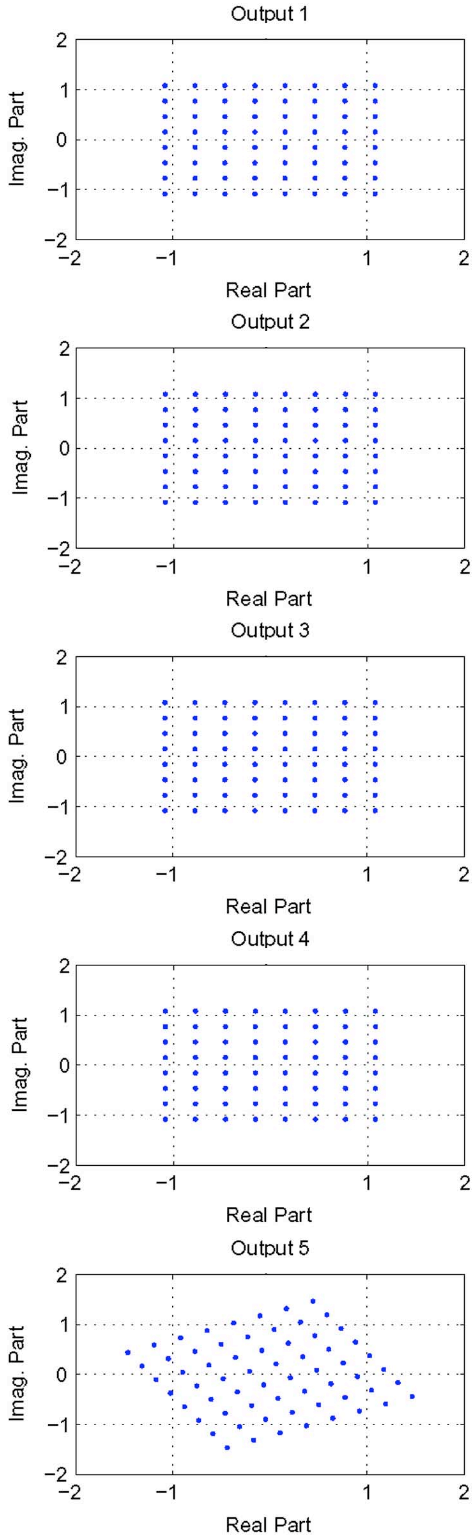

Fig. 3. Output constellations for the sources after convergence.

We record the signal to distortion ratio (SDR) for each output component, where SDR is defined as

$$\mathrm{SDR}^{(k)} = \frac{\|\mathbf{G}_k\|_\infty^2}{\|\mathbf{G}_k\|_2^2 - \|\mathbf{G}_k\|_\infty^2} \tag{79}$$

as a function of iterations.

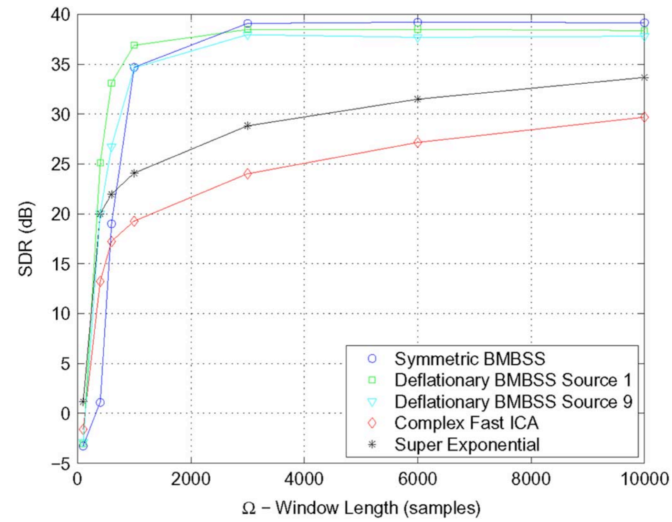

Fig. 4. Output SDR versus Window Length of Observations for the case with 64-QAM sources and 40 dB input SNR.

The ratios of peak level to the second peak level, ρ, for the effective mapping as a function iterations for different outputs are shown in Fig. 1. As expected ρ parameters grow without bound. We should note that theoretical monotonic increase may sometimes be violated when we use a finite data window and use finite number of iterations in each phase of the algorithm. The use of finite window potentially causes the violation of the ergodicity assumption about the capture of the ensemble peak within the given data window. The use of finite iterations results in partial convergence for the separator and since the projection matrices and therefore the following phases are dependent on the resulting separator this can cause a deviation from the expected convergence behavior for ρ.

In Fig. 2, the SDR convergence curves corresponding to all four phases of the algorithm are shown. These curves confirm the successful convergence behavior of the deflationary algorithm.

In Fig. 3, the output constellations after the convergence are shown. Note that the fifth separator is simply chosen based on $\mathbf{V}^{(4)}$ therefore the corresponding phase ambiguity is not an integer multiple of $(\pi/2)$ and this phase ambiguity can be reduced if desired.

In order to illustrate the relative performance of the deflationary bounded magnitude BSS (BMBSS) in comparison to its symmetric orthogonalization based counterpart and FastICA and super exponential algorithms (with symmetric orthogonalization), we assumed a simulation setup with ten sources, random 10×10 unitary mapping whose output is corrupted by white Gaussian noise. In the first setting corresponding to this setup, we assumed sources are 64-QAM signals and SNR level of mixtures is 40 dB (i.e., relatively high SNR). The output SDRs (signal to distortion ratio in this case incorporates noise power at the denominator) are averaged over different random unitary channels, noise and signal paths.

In Fig. 4, the plot of the SDRs as a function of sample size used by the algorithms are shown (for one of the sources). It

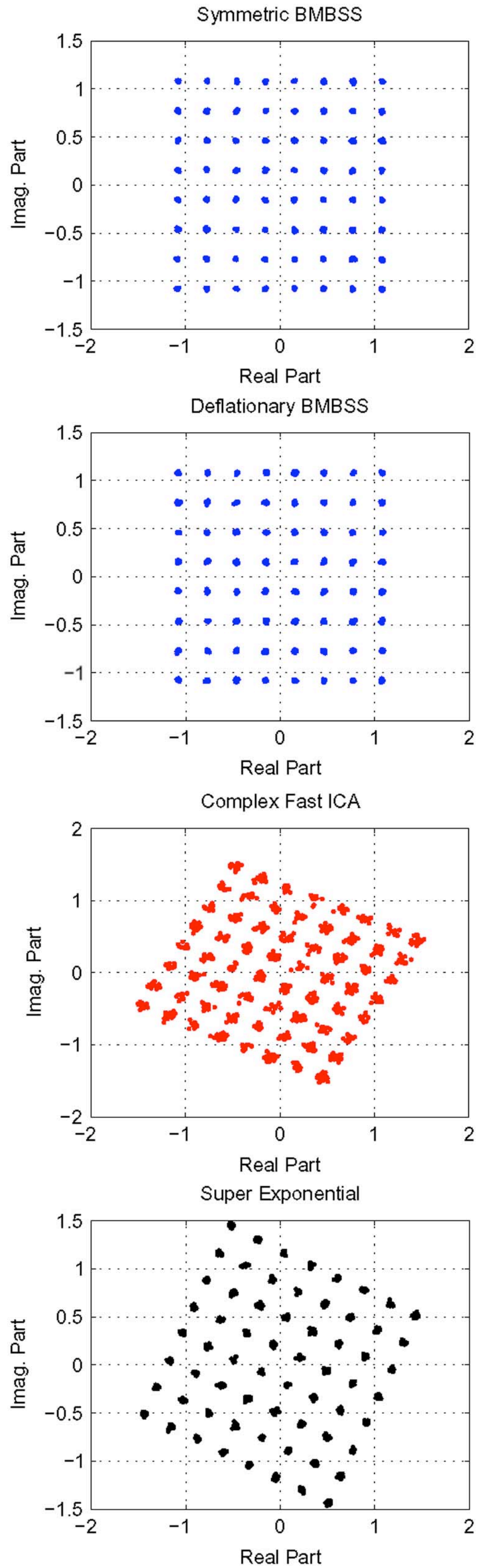

Fig. 5. The output constellations for the case with 64-QAM input sources and 40 dB input SNR.

can be seen from this plot that for this high SNR case, the proposed algorithm achieves higher SDR levels for shorter window lengths. This property of the algorithm is potentially useful for the communication links used for transmitting data (with low BER requirements and relatively high SNR) where the link is a time varying channel (for which convergence in a short burst(packet) is very valuable). Another observation we can make is that the error accumulation effects of the deflation procedure is negligible in this case as the Source 1 and Source 9 has similar performance. Furthermore, the performance of the deflation approach is close to the symmetrically orthogonoalized BMBSS algorithm.

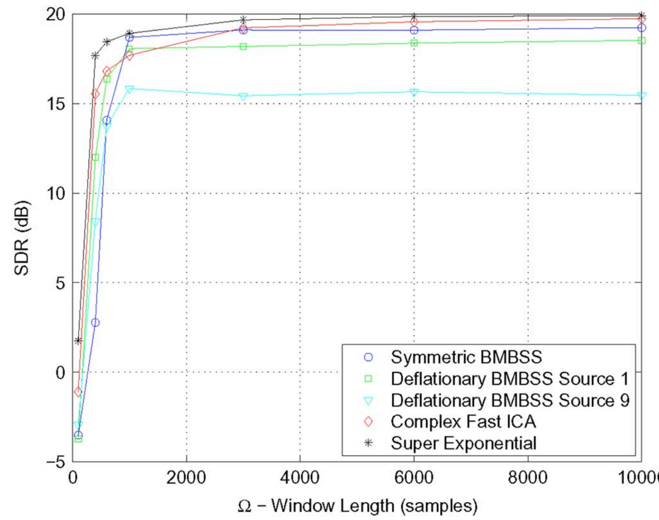

Fig. 6. Output SDR versus window length of observations for the case with 16-QAM input sources and 20 dB input SNR.

In Fig. 5, the sample output constellations obtained for a window length of 10000 are shown. Since BMBSS approach has integer multiples of $(\pi/2)$ as the phase ambiguity, the corresponding output constellation looks like the original source constellation.

For the next setting for the same setup, we consider 16-QAM sources and an input SNR level of 20 dB. In Fig. 6, the plot of the SDRs as a function of sample size used by the algorithms are shown (for one of the sources). For the low input SNR case, performance of BMBSS algorithm (for the first source) is slightly worse than the super exponential and FastICA algorithms. However, we see the effect of error accumulation in this case more clearly, as the Source 9 has about 2.5 dB performance degradation in comparison to Source 1. This degradation level can be reduced if we use more number of iterations at each phase of the deflationary algorithm. We should note that for the symmetrically orthogonalized BMBSS algorithm all outputs have SDR values around 19 dB.

In Fig. 7, the sample output constellations obtained for a window length of 10000 are shown.

VI. Conclusion

In this paper, we introduced a deflationary source separation algorithm that exploits the magnitude boundedness of the digital communication signals. Based on the following assumptions.

- The available window is sufficiently rich such that sample peak reflects the ensemble supremum;

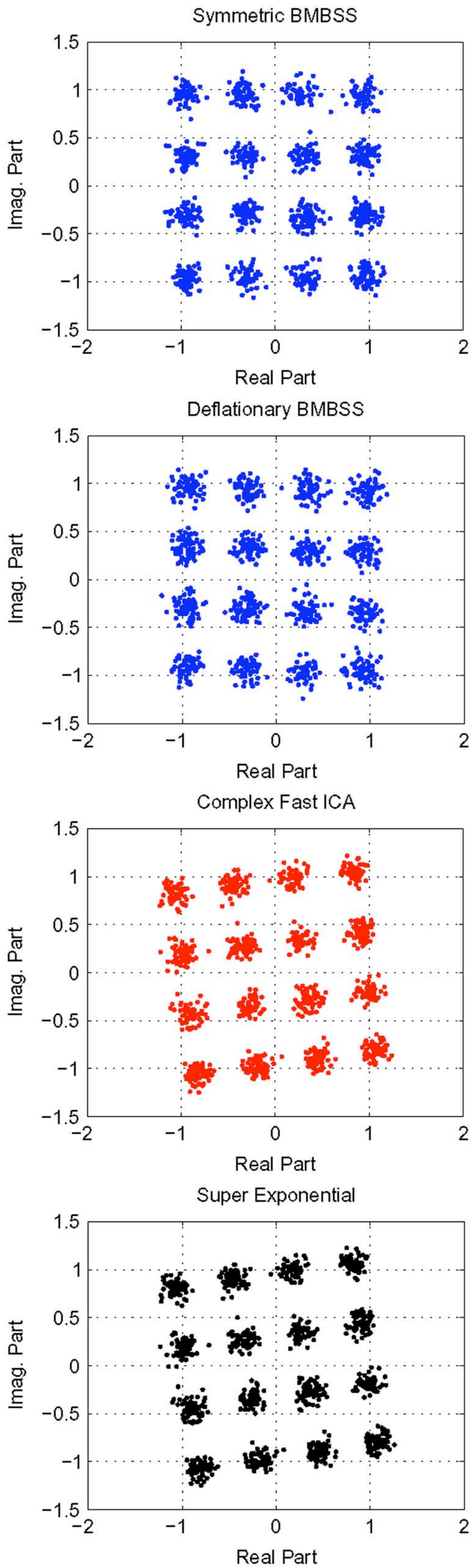

Fig. 7. The output constellations for the case with 16-QAM input sources and 20 dB input SNR.

- The initial choice of separator satisfies the generic condition that the initial overall mapping has a single peak.

we proved the global convergence of the algorithm to one of the global minima, where each minimum point corresponds to perfect separation of sources and a phase ambiguity that is an integer multiple of $(\pi/2)$.

The assumption about the richness of the window is generally satisfied for relatively short sample sizes especially for finite alphabet digital communications signals as illustrated by the simulation examples. However, if the algorithm is applied to sources with a tailing distribution the assumption about the availability of the corner points in the given data window may not hold and which results in the violation of the equivalence of the output of infinity norm minimization to minimization of the 1-norm of the combined channel.

The assumption about the single-peak initial condition holds with probability one. Even if this assumption is violated, having multiple peak points is not a stable condition. A slight disturbance on the initial separator vector will move it to another initial point (with a single peak) starting with which the algorithm will converge to a global minimum point.

ACKNOWLEDGMENT

The author would like to thank to the anonymous reviewers for their valuable comments which contributed to the quality of the paper.

REFERENCES

[1] A. T. Erdogan, "A simple geometric blind source separation method for bounded magnitude sources," *IEEE Trans. Signal Processing*, vol. 54, pp. 438–449, Feb. 2006.
[2] C. Papadias, "Globally convergent blind source separation based on a multiuser kurtosis maximization criterion," *IEEE Trans. Signal Process.*, vol. 48, pp. 3508–3519, Dec. 2000.
[3] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Netw.*, vol. 10, no. 3, pp. 626–634, 1999.
[4] N. Delfosse and P. Loubaton, "Adaptive blind separation of independent sources: A deflation approach," *Signal Process.*, vol. 45, pp. 59–83, Jul. 1995.
[5] P. Common, "Independent component analysis—A new concept?," *Signal Process.*, vol. 36, pp. 287–314, Apr. 1994.
[6] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Comput.*, vol. 9, pp. 1483–1492, Oct. 1997.
[7] D.-T. Pham, "Mutual information approach to blind separation of stationary sources," *IEEE Trans. Signal Process.*, vol. 48, pp. 1935–1946, Jul. 2002.
[8] A. Hyvärinen, J. Karhunen, and E. Oja*, Independent Component Analysis*. New York: Wiley, 2001.
[9] D.-T. Pham and P. Garat, "Blind separation of mixture of independent sources through a quasi-maximum likelihood approach," *IEEE Trans. Signal Process.*, vol. 45, pp. 1712–1725, Jul. 1997.
[10] A. Benveniste, M. Goursat, and G. Ruget, "Robust identification of a nonminimum phase system: Blind adjustment of a linear equalizer in data communications," *IEEE Trans. Autom. Control*, vol. 25, pp. 385–399, Jun. 1980.
[11] D. Donoho, "On minimum entropy deconvolution," in *Applied Time Series Analysis II*, D. Findley, Ed., 1981, pp. 565–608.
[12] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, pp. 1867–1875, Nov. 1980.
[13] D.-T. Pham, "Blind separation of instantaneous mixtures of sources based on order statistics," *IEEE Trans. Signal Process.*, vol. 48, pp. 363–375, Feb. 2000.
[14] F. Vrins, M. Verleysen, and C. Jutten, "SWM: A class of convex contrasts for source separation," in *Proc. IEEE ICASSP*, Philadelphia, PA, Mar. 2005, vol. V, pp. 161–164.

[15] A. T. Erdogan, "A blind separation approach for magnitude bounded sources," in *Proc. IEEE ICASSP*, Philadelphia, PA, Mar. 2005, vol. V, pp. 177–180.
[16] D.-T. Pham and F. Vrins, "Local minima of information-theoretic criteria in blind source separation," *IEEE Signal Process. Lett.*, vol. 12, pp. 788–791, Nov. 2005.
[17] F. Vrins and M. Verleysen, "On the entropy minimization of a linear mixture of variables for source separation," *Signal Process.*, vol. 85, pp. 1029–1044, May 2005.
[18] F. Vrins and M. Verleysen, "Information theoretic versus cumulant-based contrasts for multi-modal source separation," *IEEE Signal Process. Lett.*, vol. 12, pp. 190–193, Mar. 2005.
[19] P. Loubaton and P. Regalia, "Blind deconvolution of multivariate signals: A deflation approach," in *IEEE Int. Conf. Communications*, May 1993, vol. 2, pp. 1160–1164.
[20] D. P. Bertsekas, *Nonlinear Programming*. Nashua, NH: Athena Scientific, 1999.
[21] U. Brannlund, "On relaxation methods for nonsmooth convex optimization," Ph.D. dissertation, Dept. Mathematics, Royal Institute of Technology, Stockholm, 1993.
[22] A. Nedic and D. Bertsekas, "Convergence rate of incremental subgradient algorithms," in *Stochastic Optimization: Algorithms and Applications*, S. Ursayev and P. M. Pardalos, Eds. The Netherlands: Kluwer Academic, 2000.
[23] A. T. Erdogan and C. Kizilkale, "Fast and low complexity blind equalization via subgradient projections," *IEEE Trans. Signal Process.*, vol. 53, pp. 2513–2524, Jul. 2005.
[24] J. Goffin and K. Kiwiel, "Convergence of a simple subgradient level method," *Math. Programming*, vol. 85, no. 1, pp. 207–211, 1999.
[25] A. T. Erdogan, Performance of bounded magnitude blind source separation algorithm in noise (under preparation).

Alper Tunga Erdogan was born in Ankara, Turkey, in 1971. He received the B.S. degree from the Middle East Technical University, Ankara, Turkey, and the M.S. and Ph.D. degrees from Stanford University, CA, in 1993, 1995, and 1999, respectively, all in electrical engineering.

He was a Principal Research Engineer in Globespan-Virata Corporation (formerly Excess Bandwidth and Virata Corporations) from September 1999 to November 2001. He has been an Assistant Professor in the Electrical and Electronics Engineering Department of Koc University, Istanbul, Turkey, since January 2002. His research interests include wireless and wireline communications, adaptive signal processing, optimization, system theory and control, and information theory.