

A Class of Bounded Component Analysis Algorithms for the Separation of Both Independent and Dependent Sources

Alper T. Erdogan
Koc University
Istanbul, Turkey

Abstract—Bounded Component Analysis (BCA) is a recent approach which enables the separation of both dependent and independent signals from their mixtures. In this approach, under the practical source boundedness assumption, the widely used statistical independence assumption is replaced by a more generic domain separability assumption. This article introduces a geometric framework for the development of Bounded Component Analysis algorithms. Two main geometric objects related to the separator output samples, namely Principal Hyper-Ellipsoid and Bounding Hyper-Rectangle, are introduced. The maximization of the volume ratio of these objects, and its extensions, are introduced as relevant optimization problems for Bounded Component Analysis. The article also provides corresponding iterative algorithms for both real and complex sources. The numerical examples illustrate the potential advantage of the proposed BCA framework in terms of correlated source separation capability as well as performance improvement for short data records.

Index Terms—Bounded Component Analysis, Independent Component Analysis, Blind Source Separation, Dependent Source Separation, Finite Support, Subgradient.

I. INTRODUCTION

Blind Source Separation (BSS) has been a major research area in both signal processing and machine learning fields. The problem of extracting original sources from their mixtures appear as the central problem for several unrelated applications in digital communications, biomedical imaging, pattern recognition, finance and many more.

The blindness, i.e., the ability to adapt without requiring training information, is the key to the flexibility of BSS based approaches which leads to their widespread use. The blindness feature also makes BSS a challenging problem to solve. The hardship caused by the lack of training data and relational statistical information is generally overcome by exploiting some side information/assumptions about the model.

The most common assumption is the mutual statistical independence of sources. The BSS framework founded upon this assumption is referred to as Independent Component Analysis (ICA) and it is the most widespread BSS approach (see [1], [2], [3] and the references therein). The underlying

Copyright (c) 2013 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

Alper T. Erdogan is with Electrical-Electronics Engineering Department of Koc University, Sariyer, Istanbul, 34450, Turkey. Contact Information: Email: alper@stanfordalumni.org, Phone:+90 (212) 338 1490, Fax: +90 (212) 338 1548. This work is supported in part by TUBITAK 112E057 project.

principal of all existing ICA methods is to achieve mutual independence among separator outputs. The relative success of this approach is mainly due to the convenience of the corresponding mathematical setting, provided by the independence assumption, for the algorithm development. Another important factor is the applicability of the independence assumption on a relatively large subset of BSS application domains.

In addition to the independence assumption, several other data model assumptions or features have been exploited to produce a variety of BSS approaches. Among those, we can list time-sample structure (e.g., [4]), sparseness (e.g., [5]) and special distribution structure of communication signals such as constant modulus and finite alphabet properties (e.g., [6], [7]). During the last decade, source boundedness was also introduced as a practically valid assumption to be utilized for the separation of signals. Probably the earliest work in general BSS area is by Puntonet et.al. [8], where a geometric method based on identification of edges of parallelepiped is proposed for the separation of bounded sources. The pioneering work about exploiting boundedness in ICA framework is the reference [9], where Pham proposed the use of quantile function based approximation for the mutual information cost function in the ICA framework. In this work, he showed that under the assumption about the boundedness of the sources, the corresponding cost function can be formulated in terms of the ranges of separator outputs.

As another important step in exploiting source boundedness, in [10], Cruces and Duran posed the source extraction problem as the output support length minimization problem through an alternative path based on Renyi's entropy. In references [11], [12], [13], Vrins et.al. presented source separation algorithms based on range minimization within the ICA framework. In his thesis, Vrins also proposed the use of range based approach potentially for the separation of correlated sources [14].

As an extension of the blind equalization approach in [15], [16], the source separation based on infinity-norm minimization for magnitude bounded sources was introduced in [17]. The parallel extraction approach based on symmetrical orthogonalization procedure in this work later extended to a deflationary approach in [18] whose global convergence is proven. Due to the legacy of the digital communication signals, the output infinity norm minimization approaches assumed the equivalence of positive and negative peak magnitudes. This assumption was later abandoned in [19].

The aforementioned approaches exploiting source boundedness were introduced as a part of independent component analysis framework. Recently, Cruces in [20] showed that the source boundedness side information can be used to replace the source independence assumption with a weaker assumption about domain separability. The corresponding framework is named as Bounded Component Analysis (BCA), and it enables separation of both dependent (even correlated) and independent sources.

There actually exists some BSS approaches that address the issue about potential dependence among sources. For example, the multidimensional ICA approach was introduced to organize sources into subgroups that are internally dependent but externally independent among themselves [21], [22], [23]. There are also various approaches that target to achieve the separation of dependent sources by making use of some structural assumptions related to sources. For example, in [24], a special form of dependence, namely spatio-temporal dependence among variances, is assumed and exploited for separation. The reference [25] poses the dependence among sources/components as a practical concern in a variety of real world problems. The same reference proposes a subband decomposition based approach to handle potential dependency among sources under the assumption that some of the subband components are independent and they can be exploited to achieve global separation. The reference [26] proposes a method which exploits non-overlapping regions in the time-frequency plane and the positivity of signals to extract partially correlated data from their mixtures. Nascimento and Dias question the applicability of the ICA in hyper-spectral unmixing application in [27] and propose a dependent component separation approach imposing the convexity constraint for sources in the hyper-spectral unmixing application. Moreau et al. introduced a nonorthogonal joint diagonalization source separation approach applicable to correlated sources [28]. Chan et.al, in [29], proposed a convex optimization based geometric approach, called CAMNS-LP, which is capable of separating potentially correlated non-negative sources from the convex combination mixtures.

The dependent source separation capability of the BCA approach in [20] and the corresponding framework proposed in this article differentiates itself from the existing dependent component analysis approaches based on the following property: Under the standing boundedness assumption, BCA can be considered as the replacement of ICA with a more generic framework, which is obtained by relaxing the independence assumption with less restricting domain separability assumption. In other words, **BCA can be considered as a more general approach, covering ICA as a special case for bounded sources.**

Regarding the existing algorithmic framework for BCA, the reference [20] proposes a minimum perimeter criterion based complex source extraction algorithm. In reference [30], total range minimization based two step separation approach, where whitening step is followed by a unitary separator step, is posed as a BCA approach for uncorrelated sources. The same article also provides a convergence analysis results for the corresponding symmetrically orthogonalized algorithm.

In this article, we present a novel deterministic framework for the development of BCA algorithms which is presented in part in [31]. This framework is based on optimization settings where the objective functions are directly defined in terms of mixture samples rather than some stochastic measures or their sample based estimates. These optimization settings are derived from a geometric interpretation of separator output samples for which we define two geometric objects, namely “Principal Hyper-Ellipsoid” and “Bounding Hyper-Rectangle”. The maximization of the volume ratio of these objects is introduced as a relevant optimization approach for BCA. We show the perfect separation property of the global optima for this setting and its extensions, under a simple sample based (rather than ensemble based) assumption on sources. We also provide iterative algorithms corresponding to this framework for both real and complex sources. The potential benefits of the proposed framework are mainly twofold:

- Ability to separate both statistically dependent (even correlated) and independent sources: this property is due to the replacement of the independence assumption with a less restricting domain separability assumption.
- Desirable short-data-record performance: The proposed framework is completely deterministic, where the optimization settings and the assumption qualifying their global optima as perfect separators are all sample based. Despite the fact that the underlying ensemble structure may contain statistically independent sources, the sample realizations, especially for short data records, may not reflect this behavior. Therefore, even for the independent sources case, the proposed approach may provide better performance than some ICA approaches due to the failure of ergodicity assumption for finite data records.

The organization of the article is as follows: In Section II, we summarize the BSS setup assumed throughout the article. The proposed geometric BCA approach is introduced in Section III. The iterative algorithms based on this geometric setup are introduced in Section IV. The extension of the proposed approach for complex signals is provided in Section V. Numerical examples illustrating the benefits of the proposed approach in terms of dependent source capability and the short data record performance are provided in Section VII. Finally, Section VIII is the conclusion.

Notation: Let $\mathbf{Q} \in \Re^{p \times q}$, $\mathbf{U} \in \mathcal{C}^{p \times q}$ and $\mathbf{q} \in \mathcal{C}^{p \times 1}$ be arbitrary. The notation used in the article is summarized in Table I

Indexing: m is used for (source, output) vector components, k is the sample index and t is the algorithm iteration index.

II. BOUNDED COMPONENT ANALYSIS SETUP

The instantaneous BSS setup assumed throughout the article is show in Figure 1. In this setup

- There are p sources which are represented by the vector $\mathbf{s} = [s_1 \ s_2 \ \dots \ s_p]^T \in \Re^p$. We assume that sources are bounded in magnitude and \mathcal{S}_{s_m} represent the convex support of source m .
- The mixing system is a linear and memoryless system, which is represented by the matrix $\mathbf{H} \in \Re^{q \times p}$, where

Notation	Meaning
$\mathbf{U}_{m,:} (\mathbf{U}_{:,m})$	m^{th} row (column) of \mathbf{U}
$\Re\{\mathbf{U}\} (\Im\{\mathbf{U}\})$	the real (imaginary) part of \mathbf{U}
$\{\mathbf{Q}\}_+ (\{\mathbf{Q}\}_-)$	converts negative (positive) components of \mathbf{Q} to zero while preserving others.
$\text{abs}(\mathbf{Q})$	all elements of \mathbf{Q} are replaced by their absolute values.
$\text{sign}\{\mathbf{Q}\}$	replaces the positive entries of \mathbf{Q} with 1 and the negative entries of \mathbf{Q} with -1
$\text{signc}\{\mathbf{U}\}$	the complex sign operator defined by $\text{sign}\{\Re\{\mathbf{U}\}\} + j\text{sign}\{\Im\{\mathbf{U}\}\}$
$\mathcal{S}_{\mathbf{q}}$	the convex support for random vector \mathbf{q} : convex hull of the sample space region where the joint pdf of \mathbf{q} is non-zero.
$\mathcal{I}_{\mathbf{G}}(S)$	The image of the set S under a linear mapping defined by the matrix \mathbf{G} .
$\prod(\mathbf{q})$	$q_1 q_2 \dots q_p$, i.e. the product of the elements of \mathbf{q} .
$\ \mathbf{q}\ _r$	Usual r-norm given by $(\sum_{m=1}^p q_m ^r)^{1/r}$.

TABLE I: Notation used in the article.

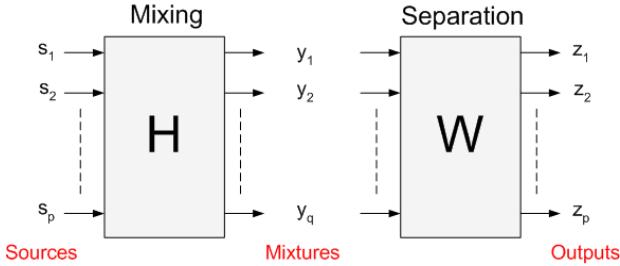


Fig. 1: Blind Source Separation Setup.

$q \geq p$. Therefore, we consider the (over)determined BSS problem.

- The mixtures are represented with $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_q]^T \in \mathbb{R}^q$, where the relation between mixtures and sources can be written as

$$\mathbf{y} = \mathbf{Hs}. \quad (1)$$

- $\mathbf{W} \in \mathbb{R}^{p \times q}$ is the separator matrix, whose output is represented by the vector $\mathbf{z} \in \mathbb{R}^p$. Therefore, we can write

$$\mathbf{z} = \mathbf{Wy}.$$

We also represent the overall mapping from sources to the separator outputs with $\mathbf{G} \in \mathbb{R}^{p \times p}$, i.e.,

$$\mathbf{z} = \mathbf{Gs}.$$

Therefore, $\mathbf{G} = \mathbf{WH}$, i.e., it is the cascade of the separator and the mixing systems. The ideal separation can be described as obtaining a \mathbf{W} matrix whose corresponding \mathbf{G} matrix has only one nonzero entry in each row and it is an invertible matrix.

The most popular approach in literature for separating sources from their mixtures is undoubtedly the ICA approach, where sources are assumed to be mutually independent. Therefore, the problem of obtaining separator matrix \mathbf{W} is typically posed as an optimization setting where an objective (contrast)

function measuring level of independence among separator outputs is maximized.

Recently, in [20], Cruces showed that if the sources are known to be bounded the mutual independence assumption among sources can be replaced with a weaker assumption. The corresponding framework, BCA, is based on the following three assumptions:

- (i.) The mixing process is invertible (i.e., \mathbf{H} is a full rank matrix with $q \geq p$),
- (ii.) The sources are non-degenerate random variables whose distributions have finite support,
- (iii.) For the support set of the distributions we can write $\mathcal{S}_{\mathbf{s}} = \mathcal{S}_{s_1} \times \mathcal{S}_{s_2} \times \dots \times \mathcal{S}_{s_p}$,

where \times is the Cartesian product.

We should note that the domain separability assumption in (iii) refers to the condition that the convex support of joint density of sources can be written as Kronecker product of the convex supports of individual source marginals. The domain separability assumption essentially implies that the boundaries of the values that a source can take is not dependent on the values of other sources, and it is a necessary condition for the mutual independence of sources. Therefore, the knowledge of source boundedness can be exploited to remove the requirement on the equivalence of the joint density and the product of the marginals. In other words, the source boundedness enables the replacement of the independence assumption with a more broadly applicable assumption allowing the separation of dependent sources. In this article, we introduce a geometric framework for the construction of BCA algorithms, which is outlined in the following sections.

III. GEOMETRIC APPROACH FOR BOUNDED COMPONENT ANALYSIS

In BSS, the main task is to adapt the separator matrix based on a finite set of mixture samples $\mathbf{Y} = \{\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(L)\}$. Let $S = \{\mathbf{s}(1), \mathbf{s}(2), \dots, \mathbf{s}(L)\}$ denote the corresponding set of unobservable source samples. For a given separator matrix \mathbf{W} with a corresponding overall mapping \mathbf{G} , we define the corresponding set of separator outputs as

$$\begin{aligned} Z_{\mathbf{G}} &= \mathcal{I}_{\mathbf{W}}(\mathbf{Y}) = \{\mathbf{Wy}(1), \mathbf{Wy}(2), \dots, \mathbf{Wy}(L)\} \\ &= \mathcal{I}_{\mathbf{G}}(S) = \{\mathbf{Gs}(1), \mathbf{Gs}(2), \dots, \mathbf{Gs}(L)\}. \end{aligned}$$

The approach proposed in this article is based on the geometric objects defined over the separator output set, $Z_{\mathbf{G}}$. Therefore, Section III-A introduces these geometric objects and some basic properties related to them. In Section III-B, a volume ratio based approach is introduced and later extended to another family of criteria in Section III-C.

A. Geometric Objects

The BCA framework proposed in this article is based on the two geometric objects related to $Z_{\mathbf{G}}$:

- **Principal Hyper-ellipsoid:** This is the hyper-ellipsoid

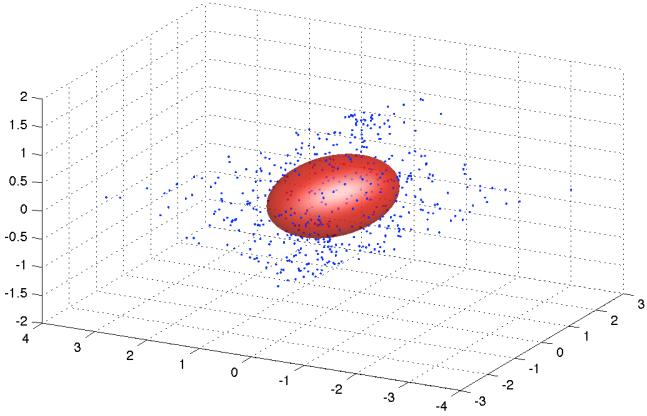


Fig. 2: A realization of separator output samples in 3-sources case and the corresponding hyper-ellipsoid.

- whose center is given by the sample mean of the set Z_G , which is defined as

$$\begin{aligned}\hat{\mu}(Z_G) &= \frac{1}{L} \sum_{\mathbf{z} \in Z_G} \mathbf{z} \\ &= \mathbf{W} \underbrace{\frac{1}{L} \sum_{k=1}^L \mathbf{y}(k)}_{\hat{\mu}(Y)}, \\ &= \mathbf{G} \underbrace{\frac{1}{L} \sum_{k=1}^L \mathbf{s}(k)}_{\hat{\mu}(S)},\end{aligned}$$

- whose principal semiaxes directions are determined by the eigenvectors of the sample covariance matrix corresponding to Z_G , which is given by,

$$\begin{aligned}\hat{\mathbf{R}}(Z_G) &= \frac{1}{L} \sum_{\mathbf{z} \in Z_G} (\mathbf{z} - \hat{\mu}(Z_G))(\mathbf{z} - \hat{\mu}(Z_G))^T \\ &= \mathbf{W} \underbrace{\frac{1}{L} \sum_{k=1}^L (\mathbf{y}(k) - \hat{\mu}(Y))(\mathbf{y}(k) - \hat{\mu}(Y))^T}_{\hat{\mathbf{R}}(Y)} \mathbf{W}^T, \\ &= \mathbf{G} \underbrace{\frac{1}{L} \sum_{k=1}^L (\mathbf{s}(k) - \hat{\mu}(S))(\mathbf{s}(k) - \hat{\mu}(S))^T}_{\hat{\mathbf{R}}(S)} \mathbf{G}^T\end{aligned}$$

- and whose principal semiaxes lengths are equal to the principal standard deviations, i.e., the square roots of the eigenvalues of $\hat{\mathbf{R}}(Z_G)$.

Therefore, we can define the corresponding set as

$$\hat{\mathcal{E}}(Z_G) = \{\mathbf{q} : (\mathbf{q} - \hat{\mu}(Z_G))^T \hat{\mathbf{R}}(Z_G)^{-1} (\mathbf{q} - \hat{\mu}(Z_G)) \leq 1\}.$$

An example, for a case of 3-sources to enable 3D picture, is provided in Figure 2. In this figure, a realization of

separator output samples and the corresponding principal hyper-ellipsoid are shown.

Note that we can also define the principal hyper-ellipsoid for the unobservable source samples as

$$\hat{\mathcal{E}}(S) = \{\mathbf{q} : (\mathbf{q} - \hat{\mu}(S))^T \hat{\mathbf{R}}(S)^{-1} (\mathbf{q} - \hat{\mu}(S)) \leq 1\}.$$

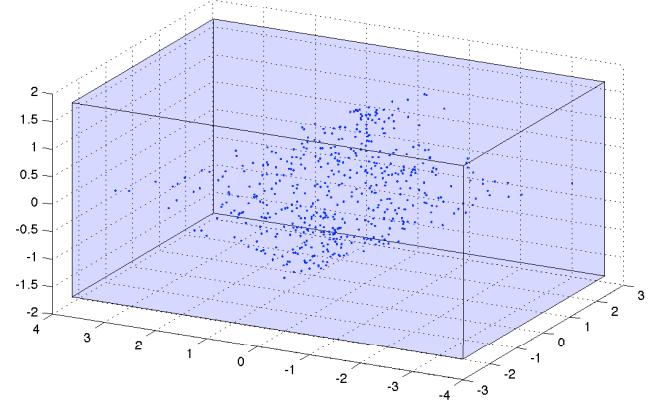


Fig. 3: Bounding hyper-rectangle for the output samples in Figure 2.

- **Bounding Hyper-rectangle:** This is defined as the minimum volume box covering all the samples in Z_G and aligning with the coordinate axes. Therefore, if we define

$$\hat{\mathbf{l}}(Z_G) = \begin{bmatrix} \min_{k \in \{1, \dots, L\}} z_1(k) \\ \min_{k \in \{1, \dots, L\}} z_2(k) \\ \vdots \\ \min_{k \in \{1, \dots, L\}} z_p(k) \end{bmatrix}, \quad (2)$$

$$\hat{\mathbf{u}}(Z_G) = \begin{bmatrix} \max_{k \in \{1, \dots, L\}} z_1(k) \\ \max_{k \in \{1, \dots, L\}} z_2(k) \\ \vdots \\ \max_{k \in \{1, \dots, L\}} z_p(k) \end{bmatrix}, \quad (3)$$

as the vectors containing minimum and maximum values for the components of the separator output vectors in Z_G respectively, we can define the bounding hyper-rectangle corresponding to the set Z_G as

$$\hat{\mathcal{B}}(Z_G) = \{\mathbf{q} : \hat{\mathbf{l}}(Z_G) \leq \mathbf{q} \leq \hat{\mathbf{u}}(Z_G)\}. \quad (4)$$

Figure 3 shows the bounding hyper-rectangle corresponding to the output samples provided in the previous example.

Similarly, we define the bounding hyper-rectangle corresponding to the set S as

$$\hat{\mathcal{B}}(S) = \{\mathbf{q} : \hat{\mathbf{l}}(S) \leq \mathbf{q} \leq \hat{\mathbf{u}}(S)\},$$

where $\hat{\mathbf{l}}(S)$ and $\hat{\mathbf{u}}(S)$, i.e., the vectors containing minimum and maximum values for the components of the source vectors in S respectively, are defined in the same way as (2-3).

Figure 4 summarizes the objects in source and separator output domains and their features in relation with the mapping from source samples to the separator output

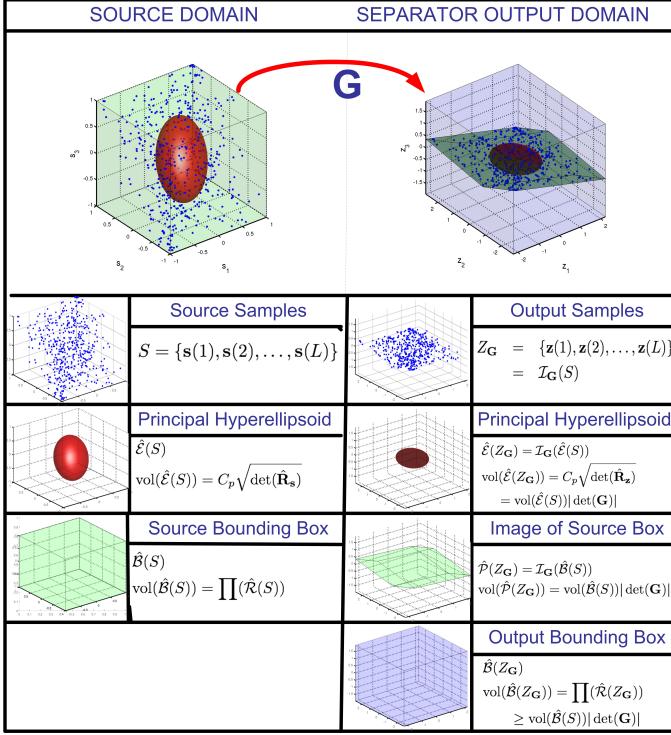


Fig. 4: Input and Output Objects related to mapping G from Sources to Separator Outputs.

samples. We make the following observations about this picture :

- The volumes of principal hyper-ellipsoids in source and separator output domains are given by

$$\text{vol}(\hat{\mathcal{E}}(S)) = C_p \sqrt{\det(\hat{\mathbf{R}}(S))}, \text{ and} \quad (5)$$

$$\text{vol}(\hat{\mathcal{E}}(Z_G)) = C_p \sqrt{\det(\hat{\mathbf{R}}(Z_G))}, \quad (6)$$

respectively, where

$$C_p = \frac{\pi^{\frac{p}{2}}}{\Gamma(\frac{p}{2} + 1)},$$

where $\Gamma(\cdot)$ is the Gamma function.

- The principal hyper-ellipsoid in the output domain is the image (with respect to the mapping defined by G) of the principal hyper-ellipsoid in the source domain, i.e.,

$$\hat{\mathcal{E}}(Z_G) = \mathcal{I}_G(\hat{\mathcal{E}}(S)). \quad (7)$$

- Based on the volume expressions in (5-6) where

$$\det(\hat{\mathbf{R}}(Z_G)) = |\det(G)|^2 \det(\hat{\mathbf{R}}(S)),$$

and also implied by the relation in (7), we can write

$$\text{vol}(\hat{\mathcal{E}}(Z_G)) = |\det(G)| \text{vol}(\hat{\mathcal{E}}(S)). \quad (8)$$

Therefore, the overall map G causes $|\det(G)|$ scaling in the volume of principal hyper-ellipsoid.

- The volumes of the bounding hyper-rectangles for the source and separator output vectors can be written as

$$\text{vol}(\hat{\mathcal{B}}(S)) = \prod(\hat{\mathcal{R}}(S)), \quad (9)$$

$$\text{vol}(\hat{\mathcal{B}}(Z_G)) = \prod(\hat{\mathcal{R}}(Z_G)), \quad (10)$$

where the sample based range operator $\hat{\mathcal{R}}$ used in these expressions is defined as

$$\hat{\mathcal{R}}(S) = \hat{\mathbf{u}}(S) - \hat{\mathbf{l}}(S), \quad (11)$$

$$\hat{\mathcal{R}}(Z_G) = \hat{\mathbf{u}}(Z_G) - \hat{\mathbf{l}}(Z_G). \quad (12)$$

Therefore, each component of the vector $\hat{\mathcal{R}}(S)$ ($\hat{\mathcal{R}}(Z_G)$) is the range of the corresponding component of the vectors in S (Z_G). In more geometric terms, the range vector $\hat{\mathcal{R}}(S)$ ($\hat{\mathcal{R}}(Z_G)$) contains the side lengths of the bounding hyper-rectangle for the source samples in S (the separator output samples in Z_G). Note that $\hat{\mathcal{R}}(S)$ ($\hat{\mathcal{R}}(Z_G)$) can be also perceived as the main diagonal vector for the bounding hyper-rectangle of the set S (Z_G).

- The image of the bounding hyper-rectangle in the source domain is a hyper-parallelepiped, represented by

$$\hat{\mathcal{P}}(Z_G) = \mathcal{I}_G(\hat{\mathcal{B}}(S)) \quad (13)$$

$$= \{\mathbf{q} : \mathbf{q} = \mathbf{Gr}, \mathbf{r} \in \hat{\mathcal{B}}(S)\}. \quad (14)$$

We note that $\hat{\mathcal{P}}(Z_G)$ is not necessarily equal to $\hat{\mathcal{B}}(Z_G)$ as illustrated by Figure 4.

- We introduce the following assumption:

Assumption: S contains the vertices of its (non-degenerate) bounding hyper-rectangle $\hat{\mathcal{B}}(S)$ **(A1)**.

Under the assumption **(A1)**, $\mathcal{I}_G(\hat{\mathcal{B}}(S))$ is a hyper-parallelepiped which is a subset of the bounding hyper-rectangle $\hat{\mathcal{B}}(Z_G)$, i.e.,

$$\mathcal{I}_G(\hat{\mathcal{B}}(S)) \subset \hat{\mathcal{B}}(Z_G),$$

which implies $\text{vol}(\hat{\mathcal{B}}(Z_G)) \geq \text{vol}(\mathcal{I}_G(\hat{\mathcal{B}}(S)))$. Therefore, the volume relation for the source and separator output bounding boxes is given by the inequality

$$\text{vol}(\hat{\mathcal{B}}(Z_G)) \geq |\det(G)| \text{vol}(\hat{\mathcal{B}}(S)). \quad (15)$$

As an important observation, under the assumption **(A1)**, the equality in (15) holds if and only if G is a perfect separator matrix, i.e. G can be written as

$$\mathbf{G} = \mathbf{DP}, \quad (16)$$

where \mathbf{D} is a diagonal matrix with non-zero diagonal entries, and \mathbf{P} is a permutation matrix. (The set of G matrices satisfying (16) will be referred as *Perfect Separators*.) This statement is formalized by the proof in Section III-B, however, it is easy to conclude by simple geometrical reasoning that the bounding box in the source domain is mapped

to another hyper-rectangle (aligning with coordinate axes) if and only if \mathbf{G} satisfies (16).

B. Volume Ratio Maximization as a BSS Approach

Based on the input/output volume relationships given by (8) and (15), we propose the following optimization setting to separate sources from their mixtures:

$$\text{maximize} \quad \frac{\text{vol}(\hat{\mathcal{E}}(Z_{\mathbf{G}}))}{\text{vol}(\hat{\mathcal{B}}(Z_{\mathbf{G}}))}$$

The related objective function can be more explicitly written as

$$J_1^{(\mathbf{W})}(\mathbf{W}) = C_p \frac{\sqrt{\det(\hat{\mathbf{R}}(Z_{\mathbf{G}}))}}{\prod(\hat{\mathcal{R}}(Z_{\mathbf{G}}))}. \quad (17)$$

The same objective function can be written as a function of \mathbf{G} , more explicitly, as

$$J_1^{(\mathbf{G})}(\mathbf{G}) = C_p \frac{|\det(\mathbf{G})| \sqrt{\det(\hat{\mathbf{R}}(S))}}{\prod(\hat{\mathcal{R}}(Z_{\mathbf{G}}))}. \quad (18)$$

Following theorem shows that the global maxima for the objective function in (17) are the perfect separators:

Theorem 1: Given \mathbf{H} in (1) is a full rank matrix and the assumption (**A1**) about the set S holds, then the set of global maxima for J_1 in (18) is equal to the set of perfect separator matrices satisfying (16).

Proof: When assumption (**A1**) holds, we can write the maximum and minimum vectors as

$$\begin{aligned} \hat{\mathbf{u}}(Z_{\mathbf{G}}) - \hat{\mu}(Z_{\mathbf{G}}) &= (\{\mathbf{G}\}_+ (\hat{\mathbf{U}} - \hat{\mathbf{M}}) + \{\mathbf{G}\}_- (\hat{\mathbf{L}} - \hat{\mathbf{M}})) \mathbf{1}, \\ \hat{\mathbf{l}}(Z_{\mathbf{G}}) - \hat{\mu}(Z_{\mathbf{G}}) &= (\{\mathbf{G}\}_- (\hat{\mathbf{U}} - \hat{\mathbf{M}}) + \{\mathbf{G}\}_+ (\hat{\mathbf{L}} - \hat{\mathbf{M}})) \mathbf{1}, \end{aligned}$$

where $\hat{\mathbf{M}} = \text{diag}(\hat{\mu}(S))$, $\hat{\mathbf{U}} = \text{diag}(\hat{\mathbf{u}}(S))$, $\hat{\mathbf{L}} = \text{diag}(\hat{\mathbf{l}}(S))$, and $\mathbf{1} = [1 \ 1 \ \dots \ 1]^T$. Therefore, the range vector for the separator outputs can be written as

$$\begin{aligned} \hat{\mathcal{R}}(Z_{\mathbf{G}}) &= \hat{\mathbf{u}}(Z_{\mathbf{G}}) - \hat{\mathbf{l}}(Z_{\mathbf{G}}) \\ &= (\{\mathbf{G}\}_+ \hat{\mathbf{U}} + \{\mathbf{G}\}_- \hat{\mathbf{L}}) \mathbf{1} - (\{\mathbf{G}\}_- \hat{\mathbf{U}} + \{\mathbf{G}\}_+ \hat{\mathbf{L}}) \mathbf{1} \\ &= (\{\mathbf{G}\}_+ (\hat{\mathbf{U}} - \hat{\mathbf{L}}) - \{\mathbf{G}\}_- (\hat{\mathbf{U}} - \hat{\mathbf{L}})) \mathbf{1} \\ &= (\{\mathbf{G}\}_+ - \{\mathbf{G}\}_-) (\hat{\mathbf{U}} - \hat{\mathbf{L}}) \mathbf{1}. \end{aligned}$$

We note that $(\hat{\mathbf{U}} - \hat{\mathbf{L}}) = \text{diag}(\hat{\mathcal{R}}(S))$. Defining $\mathbf{Q} = \mathbf{G}(\hat{\mathbf{U}} - \hat{\mathbf{L}})$, the range vector for the separator outputs can be rewritten as

$$\begin{aligned} \hat{\mathcal{R}}(Z_{\mathbf{G}}) &= (\{\mathbf{Q}\}_+ - \{\mathbf{Q}\}_-) \mathbf{1} \\ &= \text{abs}(\mathbf{Q}) \mathbf{1} \\ &= [\|\mathbf{Q}_{1,:}\|_1 \ \|\mathbf{Q}_{2,:}\|_1 \ \dots \ \|\mathbf{Q}_{p,:}\|_1]^T. \end{aligned}$$

Therefore, the objective J_1 can be more explicitly (in terms of argument \mathbf{Q}) written as

$$\begin{aligned} J_1^{(\mathbf{Q})}(\mathbf{Q}) &= C_p \frac{|\det(\mathbf{Q}(\text{diag}(\hat{\mathcal{R}}(S)))^{-1})| \sqrt{\det(\mathbf{R}(S))}}{\prod_{m=1}^p \|\mathbf{Q}_{m,:}\|_1} \\ &= \frac{|\det(\mathbf{Q})|}{\prod_{m=1}^p \|\mathbf{Q}_{m,:}\|_1} \frac{C_p \sqrt{\det(\hat{\mathbf{R}}(S))}}{\det(\text{diag}(\hat{\mathcal{R}}(S)))} \\ &= \frac{|\det(\mathbf{Q})|}{\prod_{m=1}^p \|\mathbf{Q}_{m,:}\|_1} \frac{C_p \sqrt{\det(\hat{\mathbf{R}}(S))}}{\prod(\hat{\mathcal{R}}(S))} \\ &= \frac{|\det(\mathbf{Q})|}{\prod_{m=1}^p \|\mathbf{Q}_{m,:}\|_1} \frac{\text{vol}(\hat{\mathcal{E}}(S))}{\text{vol}(\hat{\mathcal{B}}(S))}. \end{aligned} \quad (19)$$

Note that the right part of the expression in (19), which is $\frac{\text{vol}(\hat{\mathcal{E}}(S))}{\text{vol}(\hat{\mathcal{B}}(S))}$, is the volume ratio for the object of interest at the source domain. Therefore, according to (19), the volume ratio at the separator output domain is the scaled version of the volume ratio at the source domain where the scaling constant is $\frac{|\det(\mathbf{Q})|}{\prod_{m=1}^p \|\mathbf{Q}_{m,:}\|_1}$. Consequently, the maximization of the J_1 is equivalent to the maximization of this volume ratio scaling. Regarding this scaling, we note that

$$|\det(\mathbf{Q})| \leq \prod_{m=1}^p \|\mathbf{Q}_{m,:}\|_2 \quad (20)$$

$$\leq \prod_{m=1}^p \|\mathbf{Q}_{m,:}\|_1, \quad (21)$$

where

- (20) is the Hadamard inequality, which states that the volume of the hyper-parallelepiped defined by the rows of \mathbf{Q} is bounded by the volume of the hyper-rectangle with the same side-lengths. The equality in this inequality is achieved if and only if the rows of \mathbf{Q} are perpendicular to each other.
- (21) is due to the ordering $\|\mathbf{q}\|_1 \geq \|\mathbf{q}\|_2$ for any \mathbf{q} . The equality in (21) is achieved if and only if each row of \mathbf{Q} has only one non-zero element.

As a result, the ratio

$$\frac{|\det(\mathbf{Q})|}{\prod_{m=1}^p \|\mathbf{Q}_{m,:}\|_1} \quad (22)$$

is maximized if and only if \mathbf{Q} has orthogonal rows and only one non-zero element for each row. Consequently, a globally maximizing \mathbf{Q} is a perfect separator matrix defined by (16). As a result, this implies that J_1 is maximized if and only if \mathbf{G} is a perfect separator.

Remarks:

- We note that the proposed approach does not assume independence among sources. In fact, the sources can be correlated.
- Furthermore, the proposed optimization setting is directly based on sample based quantities such as sample covariances and sample ranges, rather than stochastic (or ensemble based) quantities. The proposed approach works (i.e., the global maxima are equal to perfect separators) without any requirement that the sample covariance for mixtures is equal to or even close to the ensemble

covariance, as long as the sample based assumption **(A1)** holds.

C. Extension of Volume Ratio Criterion

We can extend the approach introduced in the previous subsection by replacing the volume of the box with alternative measures reflecting the "size" of the box. Examples of such measures can be obtained by considering the "length" of the main diagonal of the corresponding box. As a result, we define a family of alternative objective functions in the form

$$J_{2,r}^{(\mathbf{W})}(\mathbf{W}) = \frac{C_p \sqrt{\det(\hat{\mathbf{R}}(Z_{\mathbf{G}}))}}{\|\hat{\mathcal{R}}(Z_{\mathbf{G}})\|_r^p}, \quad (23)$$

where $r \geq 1$. In terms of scaled overall mapping \mathbf{Q} , the corresponding objective expression can be rewritten as (by modifying (19))

$$\begin{aligned} & J_{2,r}^{(\mathbf{Q})}(\mathbf{Q}) \\ &= \frac{|\det(\mathbf{Q})|}{\left\| [\|\mathbf{Q}_{1,:}\|_1 \ \|\mathbf{Q}_{2,:}\|_1 \ \dots \ \|\mathbf{Q}_{p,:}\|_1]^T \right\|_r^p} \frac{\text{vol}(\hat{\mathcal{E}}(S))}{\text{vol}(\hat{\mathcal{B}}(S))}. \end{aligned} \quad (24)$$

If we analyze this objective function, for some special r values:

- $r = 1$ Case: In this case, we can write

$$\begin{aligned} & \left\| [\|\mathbf{Q}_{1,:}\|_1 \ \|\mathbf{Q}_{2,:}\|_1 \ \dots \ \|\mathbf{Q}_{p,:}\|_1]^T \right\|_1^p \\ &= \left(\sum_{m=1}^p \|\mathbf{Q}_{m,:}\|_1 \right)^p \end{aligned} \quad (25)$$

$$= p^p \left(\frac{1}{p} \sum_{m=1}^p \|\mathbf{Q}_{m,:}\|_1 \right)^p \quad (26)$$

$$\geq p^p \prod_{m=1}^p \|\mathbf{Q}_{m,:}\|_1, \quad (27)$$

where the last line is due to Arithmetic-Geometric-Mean Inequality, and the equality is achieved if and only if all the rows \mathbf{Q} have the same 1-norm. In consequence, we can write

$$J_{2,1}^{(\mathbf{Q})}(\mathbf{Q}) \leq \frac{|\det(\mathbf{Q})|}{\prod_{m=1}^p \|\mathbf{Q}_{m,:}\|_1} \frac{1}{p^p} \frac{\text{vol}(\hat{\mathcal{E}}(S))}{\text{vol}(\hat{\mathcal{B}}(S))}. \quad (28)$$

As a result, \mathbf{Q} is a global maximum of $J_{2,1}^{(\mathbf{Q})}$ if and only if it is a perfect separator matrix of the form

$$\mathbf{Q} = d\mathbf{P}\text{diag}(\boldsymbol{\sigma}), \quad (29)$$

where d is a non-zero value, $\boldsymbol{\sigma} \in \{-1, 1\}^p$ and \mathbf{P} is a permutation matrix. This implies \mathbf{G} is a global maximum of $J_{2,1}^{(\mathbf{G})}(\mathbf{G})$ if and only if it can be written in the form

$$\mathbf{G} = d\mathbf{P}(\hat{\mathbf{U}} - \hat{\mathbf{L}})^{-1}\text{diag}(\boldsymbol{\sigma}). \quad (30)$$

As a result, all the members of the global optima set share the same relative source scalings, unlike the set of global optima for J_1 which has arbitrary relative scalings.

- $r = 2$ Case: In this case, using the basic norm inequality

$$\begin{aligned} & \left\| [\|\mathbf{Q}_{1,:}\|_1 \ \|\mathbf{Q}_{2,:}\|_1 \ \dots \ \|\mathbf{Q}_{p,:}\|_1]^T \right\|_2 \geq \\ & \frac{1}{\sqrt{p}} \left\| [\|\mathbf{Q}_{1,:}\|_1 \ \|\mathbf{Q}_{2,:}\|_1 \ \dots \ \|\mathbf{Q}_{p,:}\|_1]^T \right\|_1, \end{aligned} \quad (31)$$

where the equality is achieved if and only if all the rows of \mathbf{Q} have the same 1-norm, we can write

$$J_{2,2}^{(\mathbf{Q})}(\mathbf{Q}) \leq \frac{|\det(\mathbf{Q})|}{\prod_{m=1}^p \|\mathbf{Q}_{m,:}\|_1} \frac{1}{p^{p/2}} \frac{\text{vol}(\hat{\mathcal{E}}(S))}{\text{vol}(\hat{\mathcal{B}}(S))}. \quad (32)$$

Therefore, $J_{2,2}$ has the same set of global maxima as $J_{2,1}$.

- $r = \infty$ Case: This time we utilize the norm inequality

$$\begin{aligned} & \left\| [\|\mathbf{Q}_{1,:}\|_1 \ \|\mathbf{Q}_{2,:}\|_1 \ \dots \ \|\mathbf{Q}_{p,:}\|_1]^T \right\|_2 \geq \\ & \frac{1}{p} \left\| [\|\mathbf{Q}_{1,:}\|_1 \ \|\mathbf{Q}_{2,:}\|_1 \ \dots \ \|\mathbf{Q}_{p,:}\|_1]^T \right\|_1, \end{aligned} \quad (33)$$

where the equality is achieved if and only if all the rows of \mathbf{Q} have the same 1-norm. Based on this inequality, we obtain

$$J_{2,\infty}^{(\mathbf{Q})}(\mathbf{Q}) \leq \frac{|\det(\mathbf{Q})|}{\prod_{m=1}^p \|\mathbf{Q}_{m,:}\|_1} \frac{\text{vol}(\hat{\mathcal{E}}(S))}{\text{vol}(\hat{\mathcal{B}}(S))}. \quad (34)$$

Therefore, $J_{2,\infty}$ also has same set of global optima as $J_{2,1}$ and $J_{2,2}$ which is a subset of perfect separators defined by (16).

IV. ITERATIVE BCA ALGORITHMS

In the previous section, a new geometric approach for BSS is presented, where the separation problem is cast as an optimization problem involving ratio of relative sizes of two geometrical objects namely the principal hyper-ellipsoid and the bounding hyper-rectangle. In this section, we derive iterative algorithms aiming to solve the corresponding optimization problems.

First, we start with the following observation: the size of the bounding hyper-rectangle is determined by the ranges of the separator outputs and the range operator is not a differentiable function of the separator matrix. As a result, the non-differentiable optimization methods, such as subgradient search, can be utilized to obtain iterative separation algorithms. For this purpose, we explicitly define the components of the range, maximum, minimum and range operators as

$$\hat{u}_m(Z_{\mathbf{G}}) = \mathbf{e}_m^T \hat{\mathbf{u}}(Z_{\mathbf{G}}) \quad m = 1, \dots, p, \quad (35)$$

$$\hat{l}_m(Z_{\mathbf{G}}) = \mathbf{e}_m^T \hat{\mathbf{l}}(Z_{\mathbf{G}}) \quad m = 1, \dots, p, \quad (36)$$

$$\hat{\mathcal{R}}_m(Z_{\mathbf{G}}) = \mathbf{e}_m^T \hat{\mathcal{R}}(Z_{\mathbf{G}}) \quad m = 1, \dots, p, \quad (37)$$

and their subdifferential sets as

$$\partial \hat{u}_m(Z_{\mathbf{G}}) = \mathbf{Co}\{\mathbf{y}(k) : k \in \mathcal{K}_{m,+}(Z_{\mathbf{G}})\}, \quad (38)$$

$$\partial \hat{l}_m(Z_{\mathbf{G}}) = \mathbf{Co}\{\mathbf{y}(k) : k \in \mathcal{K}_{m,-}(Z_{\mathbf{G}})\}, \quad (39)$$

$$\begin{aligned} \partial \hat{\mathcal{R}}_m(Z_{\mathbf{G}}) &= \mathbf{Co}\{\mathbf{y}_{\max} - \mathbf{y}_{\min} : \mathbf{y}_{\max} \in \partial \hat{u}(Z_{\mathbf{G}}), \\ &\quad \mathbf{y}_{\min} \in \partial \hat{l}(Z_{\mathbf{G}})\}, \end{aligned} \quad (40)$$

where \mathbf{Co} is the convex hull operator and the index sets $\mathcal{K}_{m,+}(Z_{\mathbf{G}}), \mathcal{K}_{m,-}(Z_{\mathbf{G}})$ are defined as

$$\mathcal{K}_{m,+}(Z_{\mathbf{G}}) = \{k : \mathbf{e}_m^T \mathbf{z}(k) = \hat{u}_m(Z_{\mathbf{G}})\}, \quad (41)$$

$$\mathcal{K}_{m,-}(Z_{\mathbf{G}}) = \{k : \mathbf{e}_m^T \mathbf{z}(k) = \hat{l}_m(Z_{\mathbf{G}})\}. \quad (42)$$

Secondly, instead of maximizing J_1 and J_2 function families, we maximize their logarithms, mainly due the convenience of the conversion of the ratio expressions to the difference expressions, which reflects as simplified expressions for the update components in the iterative algorithms. Therefore, the new objective functions can be written as

$$\bar{J}_1 = \log(J_1) \quad (43)$$

$$= \log(C_p) + \frac{1}{2} \log(\det(\hat{\mathbf{R}}(Z_{\mathbf{G}}))) \\ - \log(\prod(\hat{\mathcal{R}}(Z_{\mathbf{G}}))) \quad (44)$$

$$= \log(C_p) + \frac{1}{2} \log(\det(\mathbf{W}\hat{\mathbf{R}}(Y)\mathbf{W}^T)) \\ - \log(\prod(\hat{\mathcal{R}}(Z_{\mathbf{G}}))), \quad (45)$$

and

$$\bar{J}_{2,r} = \log(J_1) \quad (46)$$

$$= \log(C_p) + \frac{1}{2} \log(\det(\hat{\mathbf{R}}(Z_{\mathbf{G}}))) \\ - p \log(\|\hat{\mathcal{R}}(Z_{\mathbf{G}})\|_r) \quad (47)$$

$$= \log(C_p) + \frac{1}{2} \log(\det(\mathbf{W}\hat{\mathbf{R}}(Y)\mathbf{W}^T)) - \\ p \log(\|\hat{\mathcal{R}}(Z_{\mathbf{G}})\|_r). \quad (48)$$

It is interesting to remark at this point that, the objective function in (45), in square \mathbf{W} case simplifies to Pham's objective in [9] for bounded independent signals, which was obtained by approximating mutual entropy cost function via use of quantiles in the ICA framework. In the previous section, we just proved that the same objective can be derived based on geometric arguments and can be used for the separation of dependent (even correlated) bounded signals. Note that, the aforementioned equality only holds in the square case, where \mathbf{W} decouples from the correlation matrix. In the convolutional extension of this proposed framework, for example, where \mathbf{W} corresponds to a rectangular block convolution matrix such decoupling is not possible.

Based on the subdifferential sets explicitly written in (38-40), we can write the iterative update expressions for the algorithms corresponding to the modified objective functions as follows:

- *Objective Function $\bar{J}_1(\mathbf{W})$:*

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} \\ + \mu^{(t)} \left(\left(\mathbf{W}^{(t)} \hat{\mathbf{R}}(Y) \mathbf{W}^{(t)T} \right)^{-1} \mathbf{W}^{(t)} \hat{\mathbf{R}}(Y) \right. \\ \left. - \sum_{m=1}^p \frac{1}{\mathcal{R}_m(Z_{\mathbf{G}^{(t)}})} \mathbf{e}_m \mathbf{b}_m^{(t)T} \right), \quad (49)$$

with

$$\mathbf{b}_m^{(t)} = \sum_{k_{m,+} \in \mathcal{K}_{m,+}(Z_{\mathbf{G}^{(t)}})} \lambda_{m,+}^{(t)}(k_{m,+}) \mathbf{y}(k_{m,+}) \\ - \sum_{k_{m,-} \in \mathcal{K}_{m,-}(Z_{\mathbf{G}^{(t)}})} \lambda_{m,-}^{(t)}(k_{m,-}) \mathbf{y}(k_{m,-}), \quad (50)$$

where

- $\mathbf{W}^{(t)}$ is the separator matrix at the t^{th} iteration,
- $\mu^{(t)}$ is the step size at the t^{th} iteration,
- $\mathcal{K}_{m,+}(Z_{\mathbf{G}^{(t)}})$ is the set of indexes where the m^{th} separator output reaches its maximum value at the t^{th} iteration,
- $\mathcal{K}_{m,-}(Z_{\mathbf{G}^{(t)}})$ is the set of indexes where the m^{th} separator output reaches its minimum value at the t^{th} iteration,
- $\{\lambda_{m,+}^{(t)}(k_{m,+}) : k_{max} \in \mathcal{K}_{m,+}(Z_{\mathbf{G}^{(t)}})\}$ is the convex combination coefficients, used for combining the input vectors causing the maximum output, at the t^{th} iteration, which satisfy

$$\lambda_{m,+}^{(t)}(k_{m,+}) \geq 0, \quad k_{m,+} \in \mathcal{K}_{m,+}(Z_{\mathbf{G}^{(t)}}), \quad (51)$$

$$\sum_{k_{m,+} \in \mathcal{K}_{m,+}(Z_{\mathbf{G}^{(t)}})} \lambda_{m,+}^{(t)}(k_{m,+}) = 1, \quad (52)$$

$$\text{for } m = 1, \dots, p,$$

- $\{\lambda_{m,-}^{(t)}(k_{min}) : k_{min} \in \mathcal{K}_{m,-}(Z_{\mathbf{G}^{(t)}})\}$ is the convex combination coefficients, used for combining the input vectors causing the minimum output, at the t^{th} iteration, which satisfy

$$\lambda_{m,-}^{(t)}(k_{m,-}) \geq 0, \quad k_{m,-} \in \mathcal{K}_{m,-}(Z_{\mathbf{G}^{(t)}}), \quad (53)$$

$$\sum_{k_{m,-} \in \mathcal{K}_{m,-}(Z_{\mathbf{G}^{(t)}})} \lambda_{m,-}^{(t)}(k_{m,-}) = 1, \quad (54)$$

$$\text{for } m = 1, \dots, p.$$

We note that the update equation in (49) can be simplified by selecting only one input vector from each of the maximum and minimum index sets, so that we obtain

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} \\ + \mu^{(t)} \left(\left(\mathbf{W}^{(t)} \hat{\mathbf{R}}(Y) \mathbf{W}^{(t)T} \right)^{-1} \mathbf{W}^{(t)} \hat{\mathbf{R}}(Y) \right. \\ \left. - \sum_{m=1}^p \frac{1}{\mathcal{R}_m(Z_{\mathbf{G}^{(t)}})} \mathbf{e}_m \left(\mathbf{y}(k_{m,+}^{(t)}) - \mathbf{y}(k_{m,-}^{(t)}) \right)^T \right), \quad (55)$$

where $k_{m,+}^{(t)} \in \mathcal{K}_{m,+}(Z_{\mathbf{G}^{(t)}})$ and $k_{m,-}^{(t)} \in \mathcal{K}_{m,-}(Z_{\mathbf{G}^{(t)}})$.

- *Objective Function $\bar{J}_{2,r}$:*

- For $r = 1, 2$, we can write the update equation as

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} \\ + \mu^{(t)} \left(\left(\mathbf{W}^{(t)} \hat{\mathbf{R}}(Y) \mathbf{W}^{(t)T} \right)^{-1} \mathbf{W}^{(t)} \hat{\mathbf{R}}(Y) \right. \\ \left. - \sum_{m=1}^p \frac{p \hat{\mathcal{R}}_m(Z_{\mathbf{G}^{(t)}})^{r-1}}{\|\mathcal{R}(Z_{\mathbf{G}^{(t)}})\|_r^r} \mathbf{e}_m \mathbf{b}_m^{(t)T} \right), \quad (56)$$

where $\mathbf{b}_m^{(t)}$ is as defined in (50).

- For $r = \infty$, the update equation has the form

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} \quad (57)$$

$$+ \mu^{(t)} \left(\left(\mathbf{W}^{(t)} \hat{\mathbf{R}}(Y) \mathbf{W}^{(t)T} \right)^{-1} \mathbf{W}^{(t)} \hat{\mathbf{R}}(Y) \right. \\ \left. - \sum_{m \in \mathcal{M}(Z_{\mathbf{G}^{(t)}})} \frac{p\beta_m^{(t)}}{\|\hat{\mathcal{R}}(Z_{\mathbf{G}^{(t)}})\|_\infty} \mathbf{e}_m \mathbf{b}_m^{(t)T} \right), \quad (58)$$

where $\mathbf{b}_m^{(t)}$ is as defined in (50) and $\mathcal{M}(Z_{\mathbf{G}^{(t)}})$ is the set of indexes for which the peak range value is achieved, i.e.,

$$\mathcal{M}(Z_{\mathbf{G}^{(t)}}) = \{m : \hat{\mathcal{R}}_m(Z_{\mathbf{G}^{(t)}}) = \|\hat{\mathcal{R}}(Z_{\mathbf{G}^{(t)}})\|_\infty\}, \quad (59)$$

and $\beta_m^{(t)}$'s are the convex combination coefficients satisfying

$$\beta_m^{(t)} \geq 0 \quad \forall m \in \mathcal{M}(Z_{\mathbf{G}^{(t)}}), \quad (60)$$

$$\sum_{m \in \mathcal{M}(Z_{\mathbf{G}^{(t)}})} \beta_m^{(t)} = 1. \quad (61)$$

V. EXTENSION TO COMPLEX SIGNALS

In the complex case, we consider p complex sources with finite support, i.e., real and complex components of the sources have finite support. We define the operator $\Upsilon : C^p \rightarrow \Re^{2p}$,

$$\Upsilon(\mathbf{x}) = [\Re\{\mathbf{x}^T\} \quad \Im\{\mathbf{x}^T\}]^T \quad (62)$$

as an isomorphism between p dimensional complex space and $2p$ dimensional real space. For a given complex vector \mathbf{x} , we refer to the corresponding isomorphic real vector as $\hat{\mathbf{x}}$, i.e., $\hat{\mathbf{x}} = \Upsilon\{\mathbf{x}\}$. We also define the operator $\Psi : C^{p \times q} \rightarrow \Re^{2p \times 2q}$

$$\Psi(\mathbf{A}) = \begin{bmatrix} \Re\{\mathbf{A}\} & -\Im\{\mathbf{A}\} \\ \Im\{\mathbf{A}\} & \Re\{\mathbf{A}\} \end{bmatrix}. \quad (63)$$

In the complex case, both mixing and separator matrices are complex matrices, i.e., $\mathbf{H} \in C^{q \times p}$ and $\mathbf{W} \in C^{p \times q}$. The set of source vectors S and the set of separator outputs Z are subsets of C^p and the set of mixtures Y is a subset of C^q . We also note that since

$$\mathbf{y}(k) = \mathbf{H}\mathbf{s}(k), \quad (64)$$

we have

$$\hat{\mathbf{y}}(k) = \Psi(\mathbf{H})\hat{\mathbf{s}}(k). \quad (65)$$

A. Complex Extension of the Volume Ratio Approach

We extend the approach introduced in the Section III to the complex case, by simply applying the same approach to the real vector $\hat{\mathbf{z}}$. The subset of \Re^{2p} vectors which are isomorphic to the elements of Z is defined as

$$\hat{Z} = \{\hat{\mathbf{z}} : \mathbf{z} \in Z\}. \quad (66)$$

Similarly, we define

$$\hat{S} = \{\hat{\mathbf{s}} : \mathbf{s} \in S\}, \quad (67)$$

$$\hat{Y} = \{\hat{\mathbf{y}} : \mathbf{y} \in Y\}. \quad (68)$$

Using these definitions, J_1 objective function in (17) can be modified for the complex case as

$$Jc_1^{(\mathbf{W})}(\mathbf{W}) = C_p \frac{\sqrt{\det(\hat{\mathbf{R}}(\hat{Z}_{\mathbf{G}}))}}{\prod(\hat{\mathcal{R}}(\hat{Z}_{\mathbf{G}}))}. \quad (69)$$

Since the mapping between $\Upsilon(\mathbf{s})$ and $\Upsilon(\mathbf{z})$ is given by $\Psi(\mathbf{G})$, Theorem 1 implies that the set of global maxima for the objective function has the property that the corresponding $\Psi(\mathbf{G})$ satisfies (16), in addition to the structure imposed by (63). Therefore, the global optima for (69) (in terms of \mathbf{G}) is given by

$$O_c = \{\mathbf{G} = \mathbf{P}\mathbf{D} : \mathbf{P} \in \Re^{p \times p} \text{ is a permutation matrix,} \\ \mathbf{D} \in C^{p \times p} \text{ is a full rank diagonal matrix with} \\ D_{ii} = \alpha_i e^{\frac{j\pi k_i}{2}}, \alpha_i \in \Re, k_i \in Z, i = 1, \dots, p\}, \quad (70)$$

which corresponds to a subset of complex perfect separators with discrete phase ambiguity.

The corresponding (for the logarithm of Jc_1) iterative update equation for \mathbf{W} can be written as

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} + \mu^{(t)}(\mathbf{W}_{logdet}^{(t)} - \mathbf{W}_{subg}^{(t)}), \quad (71)$$

where

$$\mathbf{W}_{logdet}^{(t)} = \frac{1}{2} \left([\mathbf{I} \quad \mathbf{0}] \hat{\mathbf{R}}(\hat{Z}_{\mathbf{G}^{(t)}})^{-1} \Upsilon(\mathbf{W}^{(t)}) \hat{\mathbf{R}}(\hat{Y}) \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix} \right. \\ \left. + [\mathbf{0} \quad \mathbf{I}] \hat{\mathbf{R}}(\hat{Z}_{\mathbf{G}^{(t)}})^{-1} \Upsilon(\mathbf{W}^{(t)}) \hat{\mathbf{R}}(\hat{Y}) \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \right. \\ \left. + i([\mathbf{I} \quad \mathbf{0}] \hat{\mathbf{R}}(\hat{Z}_{\mathbf{G}^{(t)}})^{-1} \Upsilon(\mathbf{W}^{(t)}) \hat{\mathbf{R}}(\hat{Y}) \begin{bmatrix} \mathbf{0} \\ -\mathbf{I} \end{bmatrix} \right. \\ \left. + [\mathbf{0} \quad \mathbf{I}] \hat{\mathbf{R}}(\hat{Z}_{\mathbf{G}^{(t)}})^{-1} \Upsilon(\mathbf{W}^{(t)}) \hat{\mathbf{R}}(\hat{Y}) \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix} t) \right), \quad (72)$$

$$\mathbf{W}_{subg}^{(t)} = \sum_{m=1}^{2p} \mathbf{v}_m \frac{1}{2\hat{\mathcal{R}}_m(\hat{Z}_{\mathbf{G}^{(t)}})} \mathbf{b}_m^{(t)H}, \quad (73)$$

$\mathbf{b}_m^{(t)}$ is as defined in (50), and

$$\mathbf{v}_m = \begin{cases} \mathbf{e}_m & m \leq p, \\ i\mathbf{e}_{m-p} & m > p. \end{cases} \quad (74)$$

A variety on the volume ratio approach applied to the complex case can be obtained based on the following observation: we note that (part of) the numerator of the Jc_1 objective function

$$\sqrt{\det(\hat{\mathbf{R}}(\hat{Z}_{\mathbf{G}}))} = |\det(\Psi(\mathbf{G}))| \sqrt{\det(\hat{\mathbf{R}}(\hat{S}))}, \quad (75)$$

i.e., it is proportional to $|\det(\Psi(\mathbf{G}))|$. We also note that

$$|\det(\Psi(\mathbf{G}))| = |\det(\mathbf{G})|^2. \quad (76)$$

Therefore, if we define an alternative objective function

$$Jc_{1a}^{(\mathbf{W})}(\mathbf{W}) = C_p \frac{\det(\hat{\mathbf{R}}_c(Z_{\mathbf{G}}))}{\prod(\hat{\mathcal{R}}(\hat{Z}_{\mathbf{G}}))}, \quad (77)$$

where

$$\hat{\mathbf{R}}_c(Z_{\mathbf{G}}) = \sum_{k=1}^L (\mathbf{z}(k) - \hat{\mu}(Z_{\mathbf{G}}))(\mathbf{z}(k) - \hat{\mu}(Z_{\mathbf{G}}))^H, \quad (78)$$

and

$$\hat{\mu}(Z_{\mathbf{G}}) = \sum_{k=1}^L \mathbf{z}(k), \quad (79)$$

it would be proportional to J_{C_1} , and therefore, would have the same set of global optima. The convenience of J_{C_1} is in terms of the simplified update expression which is same as (71) except that the expression in (72) is replaced with the more compact expression

$$\mathbf{W}_{logdet}^{(t)} = \hat{\mathbf{R}}_c(Z_{\mathbf{G}^{(t)}})^{-1} \mathbf{W}^{(t)} \hat{\mathbf{R}}_c(Y). \quad (80)$$

B. Complex Extension of the Ellipse Volume-Box Norm Ratio Approach

Similar to the complex extension of J_1 provided in the previous subsection, we can extend the J_2 family by defining

$$J_{C_{2,r}}^{(\mathbf{W})}(\mathbf{W}) = C_p \frac{\sqrt{\det(\hat{\mathbf{R}}(\hat{Z}_{\mathbf{G}}))}}{\|(\hat{\mathcal{R}}(\hat{Z}_{\mathbf{G}}))\|_r^{2p}}, \quad (81)$$

or alternatively,

$$J_{C_{2,r}}^{(\mathbf{W})}(\mathbf{W}) = C_p \frac{\det(\hat{\mathbf{R}}(Z_{\mathbf{G}}))}{\|(\hat{\mathcal{R}}(\hat{Z}_{\mathbf{G}}))\|_r^{2p}}. \quad (82)$$

The update equation is still given by (71) where $\mathbf{W}_{logdet}^{(t)}$ is given by either (72) or (80) depending on the choice between (81) and (82), respectively, and the $\mathbf{W}_{subg}^{(t)}$ part depends on the selection of r , e.g.,

- $r = 1, 2$ Case: In this case

$$\mathbf{W}_{subg}^{(t)} = \sum_{m=1}^{2p} \mathbf{v}_m \frac{p \hat{\mathcal{R}}_m(\hat{Z}_{\mathbf{G}^{(t)}})^{r-1}}{\|(\hat{\mathcal{R}}(\hat{Z}_{\mathbf{G}^{(t)}}))\|_r^r} \mathbf{b}_m^{(t)H}, \quad (83)$$

where \mathbf{v}_m is as defined in (74) and $\mathbf{b}_m^{(t)}$ is as defined in (50).

- $r = \infty$ Case: In this case

$$\mathbf{W}_{subg}^{(t)} = \sum_{m \in \mathcal{M}(\hat{Z}_{\mathbf{G}^{(t)}})} \frac{p \beta_m^{(t)}}{\|(\hat{\mathcal{R}}(\hat{Z}_{\mathbf{G}^{(t)}}))\|_\infty} \mathbf{v}_m \mathbf{b}_m^{(t)H}, \quad (84)$$

where \mathbf{v}_m is as defined in (74), $\mathbf{b}_m^{(t)}$ is as defined in (50),

$$\mathcal{M}(\hat{Z}_{\mathbf{G}^{(t)}}) = \{m : \hat{\mathcal{R}}_m(\hat{Z}_{\mathbf{G}^{(t)}}) = \|(\hat{\mathcal{R}}(\hat{Z}_{\mathbf{G}^{(t)}}))\|_\infty\}, \quad (85)$$

and $\beta_m^{(t)}$'s are the convex combination coefficients satisfying

$$\beta_m^{(t)} \geq 0 \quad \forall m \in \mathcal{M}(\hat{Z}_{\mathbf{G}^{(t)}}), \quad (86)$$

$$\sum_{m \in \mathcal{M}(\hat{Z}_{\mathbf{G}^{(t)}})} \beta_m^{(t)} = 1. \quad (87)$$

VI. ADVANTAGE OF THE BCA APPROACH

The proposed BCA approach has two main potential advantages:

- Dependent/Correlated Source Separation Capability: Main utility of the proposed BCA approach is based on the fact that it is free of stochastic assumptions about the distribution structure of the source signals. In consequence, the proposed approach doesn't assume and exploit independence or uncorrelatedness. In fact, under the practical boundedness assumption, it is a more generic tool than Independent Component Analysis which assumes independence of sources. Example 1 in the next section illustrates the dependent as well as independent source separation capability of the proposed BCA approach. As a result, the algorithms introduced in the previous section are especially useful when it is known that the sources are dependent or when the information about the source independence/dependence is not available.
- Short Data Record Performance: Some Blind Source Separation approaches are based on some presumed stochastic settings. For example, the ICA method assumes stochastic independence of sources and exploit this feature. In the adaptive implementation of such stochastic approaches, an ergodic property on data is assumed such that the sample realization behavior somehow reflect the presumed ensemble behavior. However, capturing the stochastic behaviors such as independence, which amounts to a decomposition assumption in the pdf level, may require "sufficiently long" data samples. In contrast, the proposed BCA optimization settings are not defined in terms of stochastic parameters, they are only defined in terms of sample based quantities. Furthermore, the equivalence of the global optima of these objectives to some perfect separators are valid under the sample based assumption (A1), without invoking any stochastic arguments. Therefore, as verified by the numerical examples in the next section, the proposed BCA approaches can be expected to have better short data record performance compared to BSS methods based on stochastic criteria, even for the cases where statistical independence assumption holds.

VII. NUMERICAL EXAMPLES

A. Correlated Component Separation

The first numerical example illustrates the dependent-correlated component analysis capability of the proposed framework. For this purpose, we consider the following setup:

- Sources: Sources are randomly generated based on copulas, which serve as perfect tools for the generation of sources with controlled correlation. We consider 10 sources where the samples are generated based on Copula-t distribution [32] with 4 degrees of freedom and with Toeplitz correlation matrix whose first row is $[1 \ \rho \ \rho^2]$, where the correlation parameter is varied in the range 0 to 1.

- *Mixtures:* We assume 15 mixtures. The coefficients of the 15×10 mixing matrix are randomly generated, based on i.i.d. Gaussian distribution. We consider two different sets of mixture lengths, namely $L = 5000$ and $L = 80000$.
- *Approach:* In the simulations, the algorithm in (56) corresponding to the $J_{2,1}$ objective function is used.
- *Simulation Procedure:* We sampled 10 uniformly spaced ρ values in $[0, 0.9]$ and for each ρ value Signal energy (summed over all components) to Interference energy (summed over all components) Ratio (SIR) is computed and averaged over 600 realizations. The same procedure is repeated for FastICA [3], [33], and JADE [34], [35] algorithms, as representative ICA approaches.

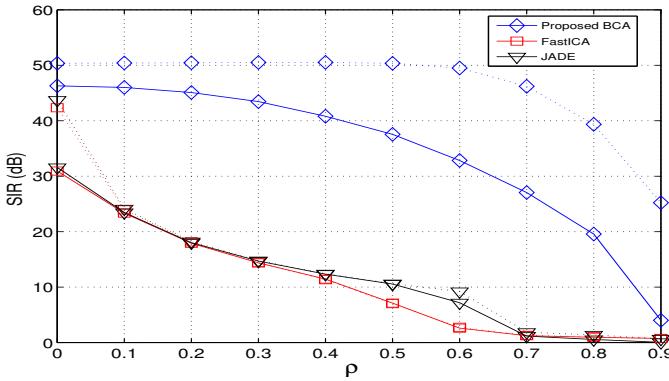


Fig. 5: Signal to Interference Ratio versus correlation parameter ρ for Example 1. Solid and Dashed Curves are for 5000 and 80000 samples long block lengths respectively.

- *Results:* The resulting Signal to Interference Ratio vs. correlation parameter curve is shown in Figure 5. Inspecting this figure, we can make the following observations:
 - Proposed BCA algorithm maintains high separation performance over a large range of ρ values, especially for the longer block length (i.e., $L = 80000$). On the other hand, both FastICA and JADE algorithms' performance degrades gracefully with increasing correlation, as expected due to the fact that independence assumption simply fails in the correlated case.
 - For the shorter block length $L = 5000$ and $\rho = 0$, the performance of the BCA algorithm is still better than FastICA, despite the fact that independence assumption still holds in this case. This can be attributed to the fact that a block of 5000 samples, in the uncorrelated case, is mostly sufficient for the assumption **A1** to hold, whereas it may fall short to reflect the independence feature of the ensemble from where it is drawn.
 - As proven in the previous sections, the proposed BCA approach achieves perfect separation even in the source correlation case as long as **A1** holds. The degradation of the performance of the BCA algorithm as ρ increases, especially in $L = 5000$ case, is due to the fact that assumption **A1** is more likely to fail when the samples are correlated and the data

length is not "long enough". However, as we increase the data size to $L = 80000$, performance remains pretty stable in the range $\rho \in [0, 0.6]$ and quite satisfactory for even higher covariances. Therefore, this is an indicator that with longer data records, the likeliness of **A1** to hold increases, and therefore, the BCA algorithm would become more successful in separating correlated sources.

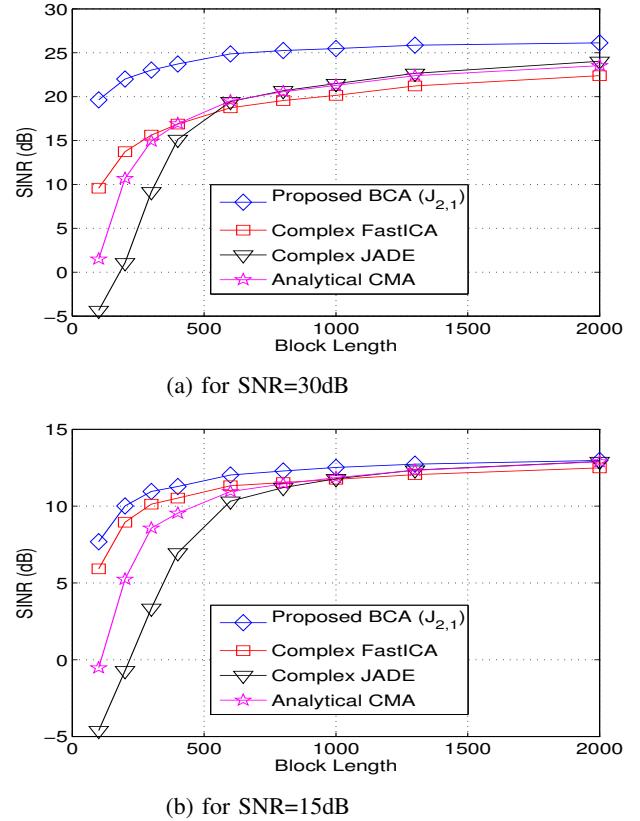


Fig. 6: Signal to Interference-Noise Ratio as a function of Block Length for Example 2 (SNR=30dB).

B. Digital Communications Example

The second example illustrates the performance of the proposed approach for a realistic digital communications scenario where the sources are complex valued signals. In the corresponding setup:

- *Sources:* We consider 8 complex QAM sources: four of the sources are 4-QAM signals and four of the sources are 16-QAM signals.
- *Mixtures:* The mixtures are signals received at a 16-antenna uniform linear array (ULA), where a narrowband model with $\lambda/2$ spacing is assumed. Each source arrives at the ULA at two paths with different gains and angles. The mixture signals are corrupted by an additive white Gaussian noise. For SNR levels two different scenarios, namely 15dB and 30dB, are considered. It is also assumed that the (symbol) clock frequencies at the receiver and the transmitters are within a range of

100ppm of a nominal value, i.e., the perfect synchronization is not assumed.

- *Approach:* In the simulations, the algorithm in (83) corresponding to the $J_{ca2,1}$ objective function is used. We compare this algorithm with complex FastICA [36], complex JADE [35]¹ and Analytical Constant Modulus Algorithm (ACMA)[7]. We also compare the performance of the proposed algorithm with the BCA source extraction algorithm by Cruces [20].

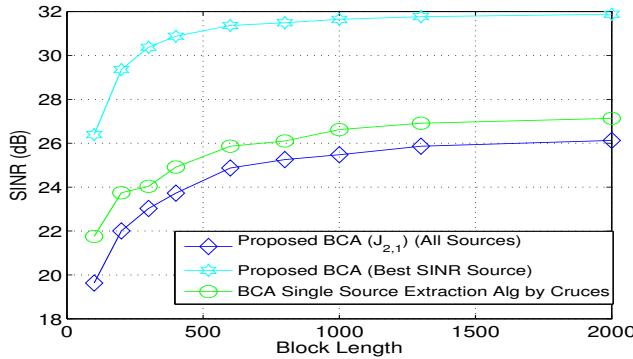


Fig. 7: Comparison of the proposed BCA algorithm with the BCA Source Extraction Algorithm by Cruces [20].

- *Results:* The resulting Signal to Interference-Noise Ratio vs. block length curves for the proposed algorihtm, FastICA, JADE and ACMA are shown in Figure 6. According to these figures, the proposed BCA approach achieves better Signal-to-Interference plus Noise Ratio (SINR) performance than other algorithms, where the relative gain is more pronounced at higher SNRs. This can be attributed to the non-stochastic nature of the proposed BCA scheme. As noted earlier, the proposed scheme doesn't assume/exploit statistical independence and it doesn't require estimation of stochastic parameters. The only impact of short data length is on accurate representation of source box boundaries. The simulation results suggest that the shorter data records may not be sufficient to reflect the stochastic independence of the sources, and therefore, the tested ICA algorithms require more data samples to achieve the same SINR level as the proposed approach.

Figure 7 compares the performance of the proposed approach by the Cruces's algorithm in [20]². Since the algorithm in [20] is a single source extraction algorithm, whereas the proposed approach is a source separation algorithm, Figure 7 shows SINR values for all sources as well as for the individual source with the best SINR performance for the proposed approach. Based on this Figure, it is clear that Cruces's BCA extraction algorithm also overperforms other ICA algorithms as well as ACMA algorithm, where the corresponding performance lies in the range the range determined by the performance

of the all sources and the source with the best performance for the proposed algorithm.

C. Separation of Images

The third example illustrates the benefit of correlated source separation capability of the proposed algorithms for the separation of images from their mixtures. We consider 5 mixtures, shown in Figure 8.(a), which are obtained through linear combination of original images shown in Figure 8.(b). In order to satisfy the assumptions of CAMNS-LP algorithm, the randomly generated mixing matrix has non-negative coefficients with unity row sum. This assumption is actually not needed for the proposed BCA algorithm. Based on this figure we can make the following observations: , both proposed BCA ($J_{2,1}$) algorithm (Figure 8.(c)) and CAMNS-LP algorithm ³ [29] (Figure 8.(d)) successfully separate original images, with 41.3 dB and 36.7 dB SIR levels, respectively (We should note here that proposed BCA approach doesn't require non-negative source and convex combination mixing assumptions made by CAMNS-LP approach). Furthermore, FastICA algorithm's performance (Figure 8.(e)) seems to suffer from the correlation among source images where the achieved SIR level is equal to only 3.4 dB.

VIII. CONCLUSION

In this article, we introduced a deterministic and geometric framework for BCA algorithm development, which can be used for the separation of both independent and dependent sources. The corresponding optimization setting is completely based on the deterministic measures derived from the geometric objects related to the samples. The numerical examples illustrate the proposed framework's potential in terms of ability to separate dependent sources as well as the short term data performance.

REFERENCES

- [1] P. Comon, "Independent component analysis- a new concept?", *Signal Processing*, vol. 36, pp. 287–314, April 1994.
- [2] Pierre Comon and Christian Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press, 2010.
- [3] Aapo Hyvärinen, Juha Karhunen, and Erkki Oja, *Independent Component Analysis*, John Wiley and Sons Inc., 2001.
- [4] M. Ohaya M. Matsuoka and M. Kawamoto, "A neural net for blind separation of nonstationary signals," *Neural Networks*, vol. 8, no. 3, pp. 411–419, March 1995.
- [5] F. Theis P. Georgiev and A. Cichocki, "Sparse component analysis and blind source separation of underdetermined mixtures," *IEEE Transactions on Neural Networks*, vol. 16, no. 4, pp. 992–996, July 2005.
- [6] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array. Part I: Algorithms," *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1184–1197, May 1996.
- [7] A. J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1136–1155, May 1996.
- [8] Carlos G. Puntonet and Alberto Prieto, "Neural net approach for blind separation of sources based on geometric properties," *Neurocomputing*, , no. 18, pp. 141–164, 1998.

¹Code available through <http://perso.telecom-paristech.fr/~cardoso/guidesepsou.html>

²Code available through <http://personal.us.es/sergio/algBCA.html>

³Code available in <http://www.ee.cuhk.edu.hk/~wkma/CAMNS/CAMNS.htm> that uses SeDuMi which is available through <http://sedumi.ie.lehigh.edu/>



Fig. 8: Image Source Separation Example

- [9] Dinh-Tuan Pham, "Blind separation of instantaneous mixtures of sources based on order statistics," *IEEE Trans. on Signal Processing*, vol. 48, pp. 363–375, February 2000.
- [10] S. Cruces and I. Duran, "The minimum support criterion for blind source extraction: a limiting case of the strengthened young's inequality," *Lecture Notes in Computer Science*, pp. 57–64, September 2004.
- [11] F. Vrins, M. Verleysen, and C. Jutten, "SWM: A class of convex contrasts for source separation," *Proceedings of IEEE ICASSP 2005*, vol. V, pp. 161–164, March 2005.
- [12] F. Vrins, John A. Lee, and M. Verleysen, "A minimum-range approach to blind extraction of bounded sources," *IEEE Trans. on Neural Networks*, vol. 18, no. 3, pp. 809–822, May 2007.
- [13] F. Vrins and D.T. Pham, "Minimum range approach to blind partial simultaneous separation of bounded sources: Contrast and discriminacy properties," *Neurocomputing*, vol. 70, pp. 1207–1214, March 2007.
- [14] F. Vrins, "Contrast properties of entropic criteria for blind source separation : a unifying framework based on information-theoretic inequalities," *Ph.D. Thesis*, 2007.
- [15] S. Vembu, S. Verdú, R. Kennedy, and W. Sethares, "Convex cost functions in blind equalization," *IEEE Trans. on Signal Processing*, vol. 42, pp. 1952–1960, August 1994.
- [16] Alper T. Erdogan and Can Kizilkale, "Fast and low complexity blind equalization via subgradient projections," *IEEE Trans. on Signal Processing*, vol. 53, pp. 2513–2524, July 2005.
- [17] Alper T. Erdogan, "A simple geometric blind source separation method for bounded magnitude sources," *IEEE Trans. on Signal Processing*, vol. 54, pp. 438–449, February 2006.
- [18] Alper T. Erdogan, "Globally convergent deflationary instantaneous blind source separation algorithm for digital communication signals," *IEEE Trans. on Signal Processing*, vol. 55, no. 5, pp. 2182–2192, May 2007.

- [19] Alper T. Erdogan, "Adaptive algorithm for the blind separation of sources with finite support," *Proceedings of EUSIPCO 2008, Lausanne*, August 2008.
- [20] S. Cruces, "Bounded component analysis of linear mixtures: A criterion for minimum convex perimeter," *IEEE Trans. on Signal Process.*, pp. 2141–2154, 2010.
- [21] J. F. Cardoso, "Multidimensional independent component analysis," *IEEE International Conf. on Acoustics, Speech and Signal Processing*, vol. 4, pp. 1941–1944, May.
- [22] A. Hyvärinen and P. Hoyer, "mergence of phase- and shift-invariant features by decomposition of natural images into independent feature subspaces," *Neural Computation*, vol. 12, no. 7, pp. 1705–1720, July 2000.
- [23] F. J. Theis, "Uniqueness of complex and multidimensional independent component analysis," *Signal Processing*, vol. 84, no. 5, pp. 951–956, May 2004.
- [24] A. Hyvärinen and J. Hurri, "Blind separation of sources that have spatiotemporal variance dependencies," *Signal Processing*, vol. 84, no. 2, pp. 247–254, February 2004.
- [25] T. Tanaka and A. Cichocki, "Subband decomposition independent component analysis and new performance criteria," *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 5, pp. 541–544, May 2004.
- [26] W. Naanaa and Jean-Marc Nuzillard, "Blind source separation of positive and partially correlated data," *Signal Processing*, vol. 85, no. 9, pp. 1711–1722, September 2005.
- [27] J.M.P. Nascimento and J.M. Bioucas-Dias, "Does independent component analysis play a role in unmixing hyperspectral data?," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 1, pp. 175–187, January 2005.
- [28] E.M. Fadili, N. Thirion-Moreau, and E. Moreau, "Non-orthogonal joint diagonalization/zero-diagonalization for source separation based on time-frequency distributions," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1673–1687, May 2007.
- [29] Tsung-Han Chan, Wing-Kin Ma, Chong-Yung Chi, and Yue Wang, "A convex analysis framework for blind separation of non-negative sources," *Signal Processing, IEEE Transactions on*, vol. 56, no. 10, pp. 5120–5134, November 2008.
- [30] Alper T. Erdogan, "On the convergence of symmetrically orthogonalized bounded component analysis algorithms for uncorrelated source separation," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 6058–6063, November 2012.
- [31] Alper T. Erdogan, "A family of bounded component analysis algorithms," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1881–1884, March 2012.
- [32] Stefano Demarta and Alexander J. McNeil, "The t copula and related copulas," *International Statistical Review*, vol. 73, no. 1, pp. 111–129, April 2005.
- [33] Aapo Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.
- [34] Jean-François Cardoso and Antoine Souloumiac, "High-order contrasts for independent component analysis," *Radar and Signal Processing, IEE Proceedings*, vol. 140, no. 6, 1993.
- [35] Jean-François Cardoso, "High-order contrasts for independent component analysis," *Neural computation*, vol. 11, no. 1, pp. 157–192, January 1999.
- [36] Ella Bingham and Aapo Hyvärinen, "A fast fixed-point algorithm for independent component analysis of complex valued signals," *International Journal of Neural Systems*, vol. 10, no. 1, pp. 1–8, 1997.



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

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

Dr. Erdogan is the recipient of several awards including TUBITAK Career Award (2005), Werner Von Siemens Excellence Award (2007), TUBA GEBIP Outstanding Young Scientist Award (2008) and TUBITAK Encouragement Award (2010). He served as an Associate Editor for the IEEE Transactions on Signal Processing, and he was a member of IEEE Signal Processing Theory and Methods Technical Committee.