Blind Separation of Bounded Sources Based on Sparse Representation

by

Eren Babataş

A Dissertation Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Doctor of Philosophy

in

Electrical and Electronics Engineering

KOÇ ÜNİVERSİTESİ

January 7, 2020

Blind Separation of Bounded Sources Based on Sparse Representation

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a doctoral dissertation by

Eren Babataş

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

examining committee have been made.

Committee Members:

Prof. Alper T. Erdoğan

Prof. Alper Demir

Prof. Engin Erzin

Prof. Özgür Erçetin

Prof. Fatih Alagöz

Date:

Dedicated to my beloved wife, Nadezda and my daughter, Lidya.

ABSTRACT

Blind Separation of Bounded Sources Based on Sparse Representation
Eren Babataş
Doctor of Philosophy in Electrical and Electronics Engineering
January 7, 2020

In this thesis, a new class of Bounded Component Analysis (BCA) algorithms based on sparsity assumption, and their applications on some practical problems are introduced. As an application, in a radar setting, the BCA algorithm proposed by Erdogan is demonstrated to separate the direct and reflected echoes which can improve the performance of classical direction of arrival estimation methods based on free space propagation theory. Following, we propose a new BCA framework for the separation of the instantaneous mixtures of sparse and bounded sources. Based on this framework, our first proposed algorithm is named Sparse Bounded Component Analysis (SBCA) which is derived from a geometric objective function defined over a completely deterministic setting. Since the framework is not related to statistical properties of source signal model, it is applicable to sources which can be statistically independent or dependent in both spatial and temporal domains. Then, SBCA framework is extended to the convolutive mixtures of sparse sources, and time and frequency domain convolutive signal separation algorithms are proposed. Finally, a space-time analysis tool is provided to detect and identify brain activity signals that have a non-stationary nature. This tool mainly relies on the short time convergence property of our SBCA framework.

ÖZETÇE

Seyrek Temsile Dayalı olarak Sınırlı Kaynakların Kör Ayrılması
Eren Babataş
Elektrik ve Elektronik Mühendisliği, Doktora
7 Ocak 2020

Bu tezde, Sınırlı Bileşen Analizi (BCA) algoritmalarından oluşan yeni bir sınıf ve bu algoritmaların bazı pratik problemler için uygulamaları tanıtılmaktadır. Bir uygulama olarak, Erdoğan tarafından önerilen BCA algoritmasının doğrudan ve yansıyan yankıları birbirinden ayırmak için nasıl kullanıldığı ve bu şekilde serbest alan yayılım teorisine dayanan klasik geliş yönü kestirim yöntemlerinin performansının iyileştirilebildiği gösterilmektedir. Ardından, seyrek ve sınırlı kaynakların anlık karışımlarının ayrılması için yeni bir BCA sistemi önerilecektir. Bu sisteme dayanarak önerilen ilk algoritmamız, tamamen belirleyici bir ayarda tanımlanan geometrik bir amaç fonksiyonundan türetilmiş olan, Seyrek Sınırlı Bileşen Analizi (SBCA) olarak adlandırılmıştır. Önerilen algoritma kaynak sinyal modelinin istatistiksel özellikleri ile ilişkili olmadığı için, hem mekansal hem de zamansal alanlarda istatistiksel olarak hem bağımsız hem de bağımlı kaynak sinyallerine uygulanabilir. Ardından, SBCA sistemi seyrek kaynakların sarmal karışımlarına genişletilerek, zaman ve frekans alanında sarmal sinyal ayrıştırma algoritmaları önerilecektir. Son olarak, durağan olmayan doğaya sahip beyin aktivite sinyallerini tespit ve teşhis etmek için SBCA yöntemini kullanan bir uzay-zaman analiz aracı sunulmaktadır. Bu araç, temel olarak SBCA yöntemimizin kısa süreli yakınsama özelliğine dayanmaktadır.

ACKNOWLEDGMENTS

I would like to gratefully acknowledge the exemplary supervision of Dr. Alper T. Erdoğan throughout the development of this thesis. Many thanks to his unique mentoring style, which pushes you to realize your full potential. I will always feel very proud and fortunate to be one of his students.

I would like to thank the thesis committee members Dr. Alper Demir and Dr. Pekcan Ungan for their constructive suggestions and comments that drove this thesis further. I am also thankful for the time and interest of the thesis jury members Dr. Engin Erzin, Dr. Özgür Erçetin and Dr. Fatih Alagöz for taking part in my thesis jury.

I would like to specially thank to my former colleague Dr. Muharrem Arık for his invaluable support.

Last but not the least, I would like to express my deepest gratefulness to my parents Emine and Cemal, my wife Nadezda, my daughter Lidya and my brother Emrah for supporting me, caring me and loving me all the time.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

BSS	**B**lind **S**ource **S**eparation
ICA	**I**ndependent **C**omponent **A**nalysis
BCA	**B**ounded **C**omponent **A**nalysis
SCA	**S**parse **C**omponent **A**nalysis
SBCA	**S**parse **B**ounded **C**omponent **A**nalysis
CSBCA	**C**onvolutive **S**parse **B**ounded **C**omponent **A**nalysis
FCSBCA	**F**requency based **C**onvolutive **S**parse **B**ounded **C**omponent **A**nalysis
DCT	**D**iscrete **C**osine **T**ransform
DFT	**D**iscrete **F**ourier **T**ransform
STFT	**S**hort **T**ime **F**ourier **T**ransform
ERP	**E**ffective **R**adiation **P**ower
DOA	**D**irection **O**f **A**rrival
ULA	**U**niform **L**inear **A**rray
BDML	**B**eamspace **D**omain **M**aximum **L**ikelihood
CMLM	**C**lassical **M**aximum **L**ikelihood **M**ethod
RMS	**R**oot **M**ean **S**quare
BP	**B**asis **P**ursuit
LTI	**L**inear **T**ime **I**nvariant
PSDM	**P**ower **S**pectral **D**ensity **M**atrix

Chapter 1

INTRODUCTION

1.1. Motivation

Usually it is not possible to directly measure the properties of a system of interest, sensors may only measure the system outputs related to the relevant properties. These outputs are the mixtures of underlying sources of interest corrupted by unwanted interference and noise. Extracting the sources of interest from their mixtures with no knowledge about the mixing process is named Blind Source Separation (BSS). This problem occurs in many irrelevant applications such as digital communications, radar signal processing, biomedical imaging and others.

The blindness, i.e., learning with no training data, knowledge about signal sources and mixing channels, is the main advantage of BSS based approaches. The difficulty arising from the lack of training data and statistical information is overcome via making use of assumptions or constraints on mixing channel and source models. After J. Herault and C. Jutten [Herault and Jutten, 1986] introduced the BSS concept, a variety of algorithms have been developed, ranging from instantaneous to convolutive ones. The main challenges for BSS algorithm developers can be listed as noisy environment, time varying and high order channel responses, moving sources, unknown number of sources, less number of channels than the number of sources.

BSS methods are increasingly applied to signal processing problems in which the multichannel observations are generally assumed to be a linear mixture of sources. The most common assumption to solve the BSS problem is the mutual statistical independence of source signals. This assumption can be used by designing optimization settings in accordance to different statistical descriptors. The methods exploiting statistical independence can be grouped under the name Independent Component Analysis which is among the most successful BSS methods [Comon, 1994, Ikeda and Murata, 1999]. Independent Component Analysis is a technique for extracting sources from their mixtures by searching for the unmixing process that results in maximally independent sources.

Besides the statistical independence, other source signal characteristics such as non-stationarity, non-whiteness or sparsity can be exploited. Magnitude-boundedness is also a practically valid signal characteristic that can be exploited for source separation. S. Cruces and I. Duran [Cruces, 2010] showed that magnitude-boundedness could be used as a side information to separate temporally or spatially dependent sources beside independent sources. This framework is called Bounded Component Analysis (BCA). BCA scheme is mainly based on domain separability assumption that is less restrictive than statistical independence assumption. In other words, BCA can be interpreted as a generalized method involving ICA as a subset.

Sparsity was firstly proposed as an alternative side information for BSS. Then, Bofill and Zibulevsky [Bofill and Zibulevsky, 2000] proposed a BSS approach based on the sparsity assumption for the undetermined case which refers to less number of channels than number of sources. Source separation methods exploiting the sparsity assumption are called Sparse Component Analysis (SCA) methods. Numerous approaches are available which address this framework [Bofill and Zibulevsky, 2000, Zibulevsky and Pearlmutter, 2001, Li et al., 2003, Georgiev and Cichoki, 2004, O'grady et al., 2005, Cotter et al., 2005, He and Cichocki, 2006, He et al., 2007, Javanmard et al., 2008, Mourad and Reilly, 2010]. SCA methods seek to find signal vectors whose entries are predominantly zero, i.e., it is aimed to find a solution for which sources contain as many zeros as possible. Combining this sparsity assumption with the domain separability assumption of BCA framework in order to cover sparse and bounded sources motivated us for a new algorithm development.

The simplest mixing model in BSS problems is a memoryless weighted sum of source samples which is referred as instantaneous mixture model. However, in practice, we face more challenging problems in which the sources are mixed both in space and in time. In the corresponding convolutive mixture problem, measured signals are linear mixtures of temporally filtered versions of sources. Convolutive mixture problem appears usually in acoustic signal processing [He et al., 2007, Parra and Spence, 2000, Douglas et al., 2007, Makino et al., 2005], biomedical applications such as brain imaging [Hansen, 2003] and electrocardiogram measurement [Vayá et al., 2007], radio communications and astronomical measurements. There are mainly two ways to solve the convolutive mixture problem. The first one is to develop a separation framework in time domain [Douglas et al., 2007]. The second one is to map the mixtures to frequency spectrum data and apply instantaneous BSS algorithms within each frequency bin separately [He et al., 2007, Parra and Spence, 2000]. The frequency domain methods have two important problems to consider: scaling and permutation inconsistencies that cause unequal scaling of the spectral components in different bins and the spectral mixing of the source components. Developing a convolutive BSS method with good practical performance which is free of such inconsistencies motivated us to extend our novel instantaneous BSS method to convolutive setting.

As mentioned, brain imaging is a typical BSS problem where recorded signals from different neural regions are mixed in sensor channels. The main goal in the brain imaging is to relate the recorded activity to hypothesized functions of the corresponding brain regions and to the related behaviour pattern, e.g., [Ziv et al., 2013, Patel et al., 2015]. For this purpose, the individual neurons and their activities need to be extracted from the recordings. What would be, of course, desirable is to accomplish this goal through an automated analysis and identification process. There are numerous methods in the literature that pretend to separate and identify neural activity signals. Separation based on temporally or spatially sparse nature of neural component activities without resorting to any independence assumption motivated us for the application of our proposed SBCA method to neuron sorting problem.

1.2. Outline of the Thesis

This thesis mainly addresses the separation problem of sparse sources without resorting to statistical independence assumption, and the application of the proposed approach to the practical BSS problems. It has a goal for extending the BCA approach to cover sparse sources and exploring the application of this extension to image separation, acoustic signal processing, and brain imaging. In Chapter 2, we review the BCA framework proposed by Erdogan for non-sparse bounded sources [Erdogan, 2013] and then put forward a novel application of this framework for direction of arrival estimation problem and demonstrate its usefulness. In the beginning of Chapter 3, we provide a comprehensive introduction to existing SCA methods in literature. Then, we propose a geometrical scheme to separate sparse bounded signals for instantaneous mixture case and derive an iterative update algorithm called Sparse BCA (SBCA) for separator matrix estimation. In this framework, the sources are assumed to be sparse in the sense that sample vectors are restricted to lie in an ℓ_1-norm ball. ℓ_1-norm has been a successful surrogate for ℓ_0-norm which is the actual indicator of sparsity, and many SCA approaches are based on the use of ℓ_1 norm minimization to promote sparsity. In our approach, we propose an objective function to be maximized, and this means to minimize ℓ_1-norm ball volume. That's why we classify our approach as an SCA method. In chapter 4, we extend SBCA algorithm to convolutive mixture case and propose three different methods in time, time-frequency and frequency domains. We exhibit a practical application of frequency based convolutive SBCA for audio signal separation and compare it to some well known convolutive BSS algorithms. In chapter 5, we develop an analysis tool that uses instantaneous SBCA algorithm to separate and identify neural components whose activities are recorded with functional Magnetic Resonance Imaging (fMRI) and Ca+2 imaging methods.

Chapter 2

AN APPLICATION OF BOUNDED COMPONENT ANALYSIS: DIRECTION OF ARRIVAL ESTIMATION IN DENSE MULTIPATH

2.1. Bounded Component Analysis Framework

Bounded Component Analysis (BCA) is a BSS scheme relying on domain separability assumption that is less strict than statistical independence assumption. By this way, it enables to separate statistically dependent sources as well as statistically independent sources. As a consequence, BCA can be interpreted to be an extension of ICA to dependent sources. Firstly, Cruces [Cruces, 2010] introduced a BCA criterion for the source extraction from a determined mixture. Then, Cruces et al [Cruces et al., 2011] extended this criterion to handle under-determined case by using semi-supervised learning. After Cruces proposed the BCA scheme, Erdogan [Erdogan, 2013] introduced a novel geometric BCA framework where instantaneous mixing is taken into consideration. This framework is completely deterministic, because the objective function is directly defined in terms of mixture samples. Moreover, the framework is based on a geometric assumption that the source samples are enclosed by a hyper-rectangle volume. This volume is formulated as the Cartesian product of the relevant source domains that fits well with many signals in practice, such as digital communication constellations, acoustic signals, natural images, etc. In this chapter, we give a short summary of Erdogan's BCA method and introduce an application of it on a practical problem, direction of arrival (DOA) estimation of a low altitude target in dense multipath [Babatas and Erdogan, 2017].

2.1.1 Blind Source Separation Signal Setup

In the instantaneous BSS setup defined in [Erdogan, 2013],

- There are p sources and L samples from each sources. Source signals are denoted by the set $\mathcal{S} = \{\mathbf{s}(n) \in \mathbb{R}^p, n = 1, \ldots L\}$.

- Source signals are magnitude-bounded and source samples lie inside a hyper-rectangle volume $\mathcal{B}_s = \{\mathbf{r} \in \mathbb{R}^p : \hat{\mathbf{l}}_s \leq \mathbf{r} \leq \hat{\mathbf{u}}_s\}$ where $\hat{\mathbf{l}}_s$ and $\hat{\mathbf{u}}_s$ represent the vectors of minimum and maximum values for the source sample set. Note that there is no stochastic assumption about the sources.

- The mixing channel response is assumed to be linear and memoryless denoted by the matrix $\mathbf{A} \in \mathbb{R}^{q \times p}$. $\mathbf{A}$ is a full rank matrix, in other words, there are more receiver channels than sources.

- The output of mixing system is formulated as

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n), \quad n = 1, \ldots, L. \tag{2.1}$$

- The separator filter is denoted by $\mathbf{W} \in \mathbb{R}^{p \times q}$, and the separator output is formulated as

$$\mathbf{y}(n) = \mathbf{W}\mathbf{x}(n), \quad n = 1, \ldots L. \tag{2.2}$$

- Giving the total system response as $\mathbf{F} = \mathbf{W}\mathbf{A}$, the separator output is expressed with

$$\mathbf{y}(n) = \mathbf{F}\mathbf{s}(n), \quad n = 1, \ldots, L. \tag{2.3}$$

- Perfect separation is achieved by a separator filter for which the total response satisfies

$$\mathbf{F} = \mathbf{P}\mathbf{D}, \tag{2.4}$$

where $\mathbf{P}$ is a permutation matrix and $\mathbf{D}$ is a full rank diagonal scaling matrix.

In [Erdogan, 2013], perfect separation is achieved by maximizing the volume ratio of two geometric objects. These objects, called as principal hyper-ellipsoid $\mathcal{E}_y$ and bounding hyper-rectangle $\mathcal{B}_y$, are defined over the separator output samples. The geometric framework is illustrated in Fig.2.1.

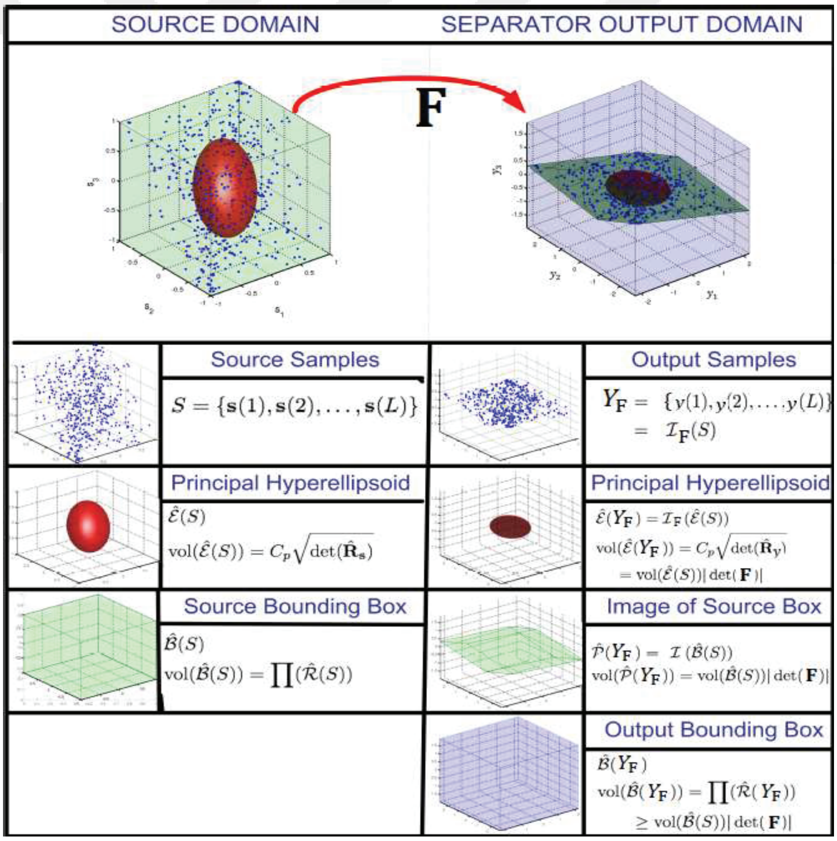


Figure 2.1: Geometric setting used in BCA framework.

2.1.2 Objective Function and Iterative Algorithm

The BCA objective is formulated as a volume ratio based upon the BSS signal setup in the previous section:

$$\bar{J}(\mathbf{W}) = \text{Volume}(\mathcal{E}_y)/\text{Volume}(\mathcal{B}_y) \tag{2.5}$$

$$= C_p \frac{\sqrt{\det(\hat{\mathbf{R}}_{\mathbf{y}})}}{\prod(\hat{R}(\mathcal{Y}_{\mathbf{F}}))}. \tag{2.6}$$

$\mathcal{E}_y$ and $\mathcal{B}_y$ represent the volumes of the principal hyper-ellipsoid and the bounding hyper-rectangle respectively. $C_p = \pi^{p/2}/\Omega(p/2+1)$, $\hat{\mathbf{R}}_{\mathbf{y}}$ is the output covariance matrix, and $\hat{R}(\mathcal{Y}_{\mathbf{F}})$ is the range vector. After eliminating the dimension dependent constant C_p, the simplified objective function is

$$J(\mathbf{W}) = \frac{\sqrt{\det(\hat{\mathbf{R}}_{\mathbf{y}})}}{\prod(\hat{R}(\mathcal{Y}_{\mathbf{F}}))}. \tag{2.7}$$

To prove that the global maximums of (2.7) give perfect separation, the following **local dominance** assumption is introduced:

Assumption(**A1**): The vertices of the bounding hyper-rectangle $\mathcal{B}_s$ are included in the source sample set.

Theorem-1: If $\mathbf{A}$ is full-rank and (**A1**) is correct, all global maximums of (2.7) correspond to perfect separators.

For a detailed proof of this theorem, interested readers are referred to [Erdogan, 2013].

Noting that the objective function (2.7) is non-convex and not differentiable everywhere, Clarke sub-differential method [Bagirov et al., 2014] can be applied to obtain an iterative update equation. The iterative update equation by steepest ascent method is

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} + \mu^{(t)} \left((\hat{\mathbf{R}}_{\mathbf{y}}^{(t)})^{-1} \mathbf{W}^{(t)} \hat{\mathbf{R}}_{\mathbf{x}} - \sum_{m=1}^{p} \frac{1}{\hat{R}_m(\mathcal{Y}_{\mathbf{F}^{(t)}})} \mathbf{e}_m (\mathbf{x}(k_{m,+}^{(t)}) - \mathbf{x}(k_{m,-}^{(t)})) \right), \tag{2.8}$$

where $\hat{\mathbf{R}}_{\mathbf{x}} = \dfrac{1}{L}\sum_{n=1}^{L}\mathbf{x}(n)\mathbf{x}(n)^T - \mu_x\mu_x^T$ with $\mu_x = \dfrac{1}{L}\sum_{n=1}^{L}\mathbf{x}(n)$, $\mathbf{x}(k_{m,+}^{(t)})$ and $\mathbf{x}(k_{m,-}^{(t)})$ represent the mixture signals that provide respectively maximum and minimum separator output values at the $(t)^{th}$ iteration.

2.2. *Extension of BCA Algorithm to Complex Signals*

The objective function (2.7) is formulated for real sources. We can extend this function to complex case easily. In the complex case, there are p complex sources having real and imaginary parts with finite support. To obtain the complex extension of the iterative algorithm, the isomorphism based approach is applied as explained in [Erdogan, 2013]. In this approach, the operator $\Upsilon : \mathbb{C}^p \rightarrow \mathbb{R}^{2p}$ defines an isomorphism between complex and real signal vectors, formulated as

$$\Upsilon(\mathbf{u}) = \left[\mathcal{R}e\{\mathbf{u}^T\} \quad \mathcal{I}m\{\mathbf{u}^T\}\ \right]^T. \tag{2.9}$$

For simplification, $\acute{\mathbf{u}}$ will denote real isomorphic counterpart of a complex vector $\mathbf{u}$, i.e., $\acute{\mathbf{u}} = \Upsilon(\mathbf{u})$.

We also introduce the operator $\Omega : \mathbb{C}^{p\times q} \rightarrow \mathbb{R}^{2p\times 2q}$ formulated as

$$\Omega(\mathbf{Z}) = \begin{bmatrix} \mathcal{R}e\{\mathbf{Z}\} & -\mathcal{I}m\{\mathbf{Z}\} \\ \mathcal{I}m\{\mathbf{Z}\} & \mathcal{R}e\{\mathbf{Z}\}, \end{bmatrix} \tag{2.10}$$

for complex matrix mapping. From these definitions, the following equations are derived for complex mixtures and separator outputs:

$$\acute{\mathbf{x}}(n) = \Omega(\mathbf{A})\acute{\mathbf{s}}(n), \tag{2.11}$$

$$\acute{\mathbf{y}}(n) = \Omega(\mathbf{W})\acute{\mathbf{x}}(n). \tag{2.12}$$

$\acute{\mathcal{X}} = \{\acute{\mathbf{x}} : \mathbf{x} \in \mathcal{X}\}$ and $\acute{\mathcal{Y}} = \{\acute{\mathbf{y}} : \mathbf{y} \in \mathcal{Y}\}$ correspond to real-isomorphic counterparts of mixture and separator output sample sets.

After giving the necessary mathematical definitions, we formulate the following objective and iterative update equation for complex source signals:

$$J_c(\mathbf{W}) = \frac{\sqrt{\det(\hat{\mathbf{R}}_{\mathbf{y}'})}}{\prod(\hat{R}(\mathcal{Y}_{\mathbf{F}}))}. \tag{2.13}$$

Using algebraic manipulations, the corresponding iterative update equation for (2.13) turns out to be

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} + \mu^{(t)}(\mathbf{W}^{(t)}_{logdet} - \mathbf{W}^{(t)}_{subg}), \tag{2.14}$$

where

$$\begin{aligned}\mathbf{W}^{(t)}_{logdet} = \frac{1}{2}\Bigg(&\begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}\hat{\mathbf{R}}^{-1}_{\mathbf{y}^{(t)}}\Omega(\mathbf{W}^{(t)})\hat{\mathbf{R}}_{\mathbf{x}'}\begin{bmatrix}\mathbf{I}\\ \mathbf{0}\end{bmatrix}\\ &+\begin{bmatrix}\mathbf{0} & \mathbf{I}\end{bmatrix}\hat{\mathbf{R}}_{\mathbf{y}^{(t)}}{}^{-1}\Omega(\mathbf{W}^{(t)})\hat{\mathbf{R}}_{\mathbf{x}'}\begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}\\ &+j\Bigg(\begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}\hat{\mathbf{R}}^{-1}_{\mathbf{y}^{(t)}}\Omega(\mathbf{W}^{(t)})\hat{\mathbf{R}}_{\mathbf{x}'}\begin{bmatrix}\mathbf{0}\\ -\mathbf{I}\end{bmatrix}\\ &+\begin{bmatrix}\mathbf{0} & \mathbf{I}\end{bmatrix}\hat{\mathbf{R}}_{\mathbf{y}^{(t)}}{}^{-1}\Omega(\mathbf{W}^{(t)})\hat{\mathbf{R}}_{\mathbf{x}'}\begin{bmatrix}\mathbf{I}\\ \mathbf{0}\end{bmatrix}\Bigg)\Bigg),\end{aligned} \tag{2.15}$$

and

$$\mathbf{W}^{(t)}_{subg} = \sum_{m=1}^{2p} \frac{1}{\hat{R}_m(\mathcal{Y}_{\mathbf{F^{(t)}}})}\mathbf{v}_m(\mathbf{x}(k^{(t)}_{m,+}) - \mathbf{x}(k^{(t)}_{m,-})), \tag{2.16}$$

$$\mathbf{v}_m = \begin{cases} \mathbf{e}_m & m \leq p, \\ -i\mathbf{e}_{m-p} & m > p \end{cases}. \tag{2.17}$$

2.3. DOA Estimation of a Low Altitude Target in Dense Multipath

Estimating the arrival directions of low-altitude targets Has been a main research problem of interest for a few decades. In the corresponding problem, sensors receive two or more coherent echoes due to the reflection paths of the signal reflected from low-altitude targets. This results in the periodic fading and strengthening of the combined signal along the propagation path. The reflected echoes consist of two fundamental components referred as specular and diffused components. The specular component is dominant in the case of small grazing angle and smooth reflecting surface. This case leads to the following problems:

1. The direct and specular echoes are almost same. The specular echo's phase and amplitude differ from those of the direct-path echo by quantities related to range and altitude; so they are consistent.

2. The two echoes are inside the beam width, so that the angular distance between them is little.

3. The two echoes cover almost the same length of propagation path, thus they are super-positioned and can not be distinguished over time axis.

The diffused component additionally complicates the multipath problem by adding inconsistent interferences to the detected signal in the case of rough reflecting surface and large grazing angle.

The problem of multipath propagation arises in the context of tracking low elevation targets and can cause severe degradation in the measured elevation angle of the target. Many researchers have investigated the negative effects of multipath reflections on target parameter estimation and proposed numerous methods to minimize the estimation error resulting from these effects. These methods can be classified into three main categories: Monopulse methods, parametric methods and beam-space methods [Zoltowski and Lee, 1991, Saruwatari et al., 2003, Cardoso and Souloumiac, 1993]. Classical Maximum Likelihood Method (CMLM) that aims to estimate target parameters by correlation maximization between array manifold matrix and received data [Stoica and Kenneth C., 1990] is categorized as a parametric method. However, the CMLM does not perform well if the multipath signal is also received coherent with the direct-path signal and fades the total received signal.

In this section, we introduce a new method proposed in [Babatas and Erdogan, 2017] to reduce multipath effect via separation of direct and reflected echoes from each other by using Bounded Component Analysis (BCA). We illustrate performance improvement over CMLM in the experimental results. The proposed algorithm reduces the effect of attenuation caused by multipath signals and improves the DOA estimation performance.

2.3.1 Multipath Signal Model

In our signal setup, Griesser and Balanis' multipath model [Griesser and Balanis, 1987] is considered. Propagation geometry by the model is given in Figure 2.2 below. As shown in Figure 2.2, a uniform linear array (ULA) and a point target

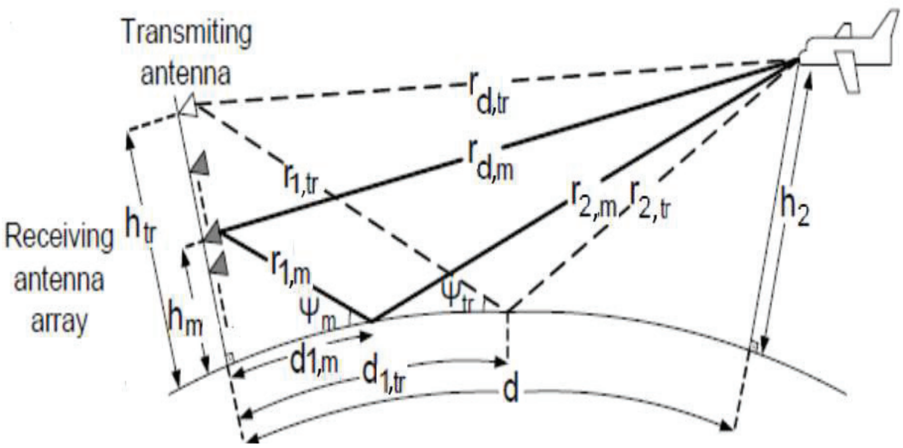


Figure 2.2: Geometrical model for direct-path and reflected-path echo propagation.

are used. There are two ways in transmission and reception of radio waves, direct-way ($r_{d,tr}, r_{d,m}$) and reflection-way ($r_{1,tr} + r_{2,tr}, r_{1,m} + r_{2,m}$). Reflection angles are represented with ψ_{tr}, ψ_m. The earth model is a sphere with radius R_e. $\rho(n, \psi)$ is the reflection coefficient of the earth, and complex valued.

We calculate the reflection angles ψ_{tr}, ψ_m with the approximation formulas proposed in [Griesser and Balanis, 1987]. The reflection coefficient includes two components named specular reflection coefficient (ρ_s) and diffuse reflection coefficient (ρ_d).

$$\rho(n, \psi) = \rho_s(\psi) + \rho_d(n) \tag{2.18}$$

The subsequent equations are used for specular reflection calculation:

$$\rho_s(\psi) = \rho_0 R_s(\psi) = \Gamma D R_s(\psi), \tag{2.19}$$

$$R_s(\psi) = \exp(-8\pi^2 \delta(\psi)^2) I_0(-8\pi^2 \delta(\psi)^2), \tag{2.20}$$

$$\delta(\psi) = (\sigma_h \sin \psi)/\lambda. \tag{2.21}$$

Γ is called the Fresnel reflection coefficient, D is the divergence factor, $I_0(z)$ is the modified Bessel function, σ_h is the effective surface height, and λ is the wavelength for the operational frequency. The equation

$$\begin{aligned} \rho_d(n) &= \rho_0 R_d(n) \\ &= \Gamma_{v,h} D R_d(n) \\ &= \Gamma_{v,h} D (R_{dx}(n) + j R_{dy}(n)) \end{aligned} \tag{2.22}$$

defines more complicated diffused reflection coefficient where R_{dx} and R_{dy} denote random processes with zero mean normal distribution. Finally, based on the **narrowband assumption**, the signal that arrives from a point target and is received by an M-element Uniform Linear Array (ULA) is expressed as

$$\mathbf{x}(n) = \left[\begin{array}{c} x_1(n), x_2(n), \cdots, x_M(n) \end{array} \right]^T, \tag{2.23}$$

$$x_k(n) = o(n)\{\exp(-jKr_{d,k}) + \rho(n,\psi)\exp(-jK(r_{1,k}+r_{2,k}))\}, \tag{2.24}$$

where $o(n)$ represents any waveform modulation and $K = \frac{2\pi}{\lambda}$ is the wave number.

The signal model in (2.23) looks similar with the BSS mixture model in (2.1). We aim to modify (2.23) in order to clarify this similarity. The overall reflection coefficient in (2.18) depends on two parameters, time and grazing angle. Let us define $\mathbf{r}_d = [r_{d,1}, r_{d,2}, \ldots, r_{d,M}]^T$ and $\psi = [\psi_1, \psi_2, \ldots, \psi_M]^T$. Note that $r_{1,k}$ and $r_{2,k}$ depend on ψ_k as can be seen in figure 2.2. As a result, the total received echo is defined with respect to three arguments: sampling time n, $\mathbf{r}_d$ and ψ.

$$\mathbf{x}(n,\mathbf{r}_d,\psi) = o(n)\begin{bmatrix} \exp(-jKr_{d,1}) + \rho(n,\psi)\exp(-jK(r_{1,1}+r_{2,1})) \\ \vdots \\ \exp(-jKr_{d,M}) + \rho(n,\psi)\exp(-jK(r_{1,M}+r_{2,M})) \end{bmatrix} \tag{2.25}$$

$$\mathbf{A}(\mathbf{r}_d,\psi) = \begin{bmatrix} \exp(-jKr_{d,1}) & \exp(-jK(r_{1,1}+r_{2,1})) \\ \vdots & \\ \exp(-jKr_{d,M}) & \exp(-jK(r_{1,M}+r_{1,M})) \end{bmatrix} \tag{2.26}$$

$$\mathbf{s}(n,\psi) = \begin{bmatrix} o(n) \\ \rho(n,\psi)o(n) \end{bmatrix} \tag{2.27}$$

$$\mathbf{x}(n,\mathbf{r}_d,\psi) = \mathbf{A}(\mathbf{r}_d,\psi)\mathbf{s}(n,\psi) \qquad n = 1,\ldots,L. \tag{2.28}$$

where $\mathbf{A}(\mathbf{r}_d,\psi)$ and $\mathbf{s}(n,\psi)$ resemble mixing channel response and source signal in the BSS signal model respectively. In order to simplify (2.28), the following assumption is put forward.

Assumption (**A2**): Variation in $\mathbf{A}(\mathbf{r}_d,\psi)$ with respect to the parameters $\mathbf{r}_d$ and ψ throughout receiving L samples from the target is negligible. According to the assumption (**A2**), (2.28) turns into the linear mixture model in BSS problem.

2.3.2 Direction of Arrival Estimation Estimation for a Low Altitude Target

Under the assumption (**A2**), we propose a four-step approach to estimate DOA of a target at low-altitude based on the analogy between the multipath model (2.28) and the BSS model (2.1) .

I- Estimate the separator matrix $\mathbf{W}$ using the received signal matrix ($M \times L$ size) in (2.14).

II- Separate the direct-path echo ($y_1(n)$) and reflected-path echo ($y_2(n)$) given in (2.27).

III- Multiply the separated echoes with the array manifold in order to obtain the received signals that include phase differences between the array elements. Let us denote these vectors with $\mathbf{x}_1(n)$ and $\mathbf{x}_2(n)$:

$$\mathbf{x}_1(n) = (\mathbf{W}(1,:)^H \mathbf{W}(1,:))^{-1} \mathbf{W}(1,:)^H y_1(n) \qquad n = 1, \dots, L. \tag{2.29}$$

$$\mathbf{x}_2(n) = (\mathbf{W}(2,:)^H \mathbf{W}(2,:))^{-1} \mathbf{W}(2,:)^H y_2(n) \qquad n = 1, \dots, L. \tag{2.30}$$

IV- Estimate DOAs for both $\mathbf{x}_1(n)$ and $\mathbf{x}_2(n)$ by using any estimation method based on free-space model. We selected the CMLM in [Stoica and Kenneth C., 1990] for this task. The larger angle estimate is accepted to be target's arrival angle.

2.3.3 Numerical Results

In the analysis, the proposed method's performance is evaluated compared to CMLM [Stoica and Kenneth C., 1990] and Beamspace Domain Maximum Likelihood (BDML) [Zoltowski and Lee, 1991]. In the first experiment, we evaluate the DOA estimation performance with respect to target altitude. Target distance is constant at 5500 meters and free space input SNR=20 dB. The other selected parameters are: i)-75 antenna elements linearly spaced with 0.5λ distance, ii)-Sensor Height (38th element)=16 m, iii)-Frequency = 10 GHz, iv)-Swerling-0 case (Constant target RCS), v)-Sea state = 0 (no diffuse reflection).

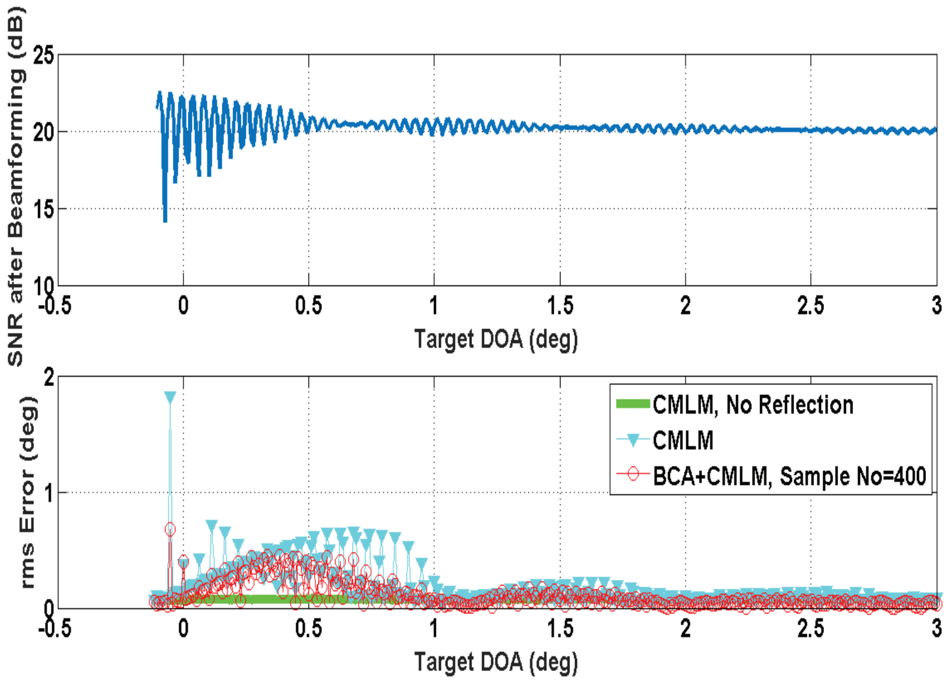


Figure 2.3: Performance improvement over CMLM provided by BCA. 1)-Output SNR versus target's arrival angle at distance=5500m, 2)-RMS error versus target's arrival angle (target distance=5500m).

The first plot of Fig. 2.3 illustrates total input SNR versus target arrival angle. The second plot of Fig. 2.3 illustrates rms errors for CMLM and BCA+CMLM with and without multipath effect for different target DOAs. It can be interpreted that BCA+CMLM provides a significant improvement over CMLM, especially under multipath effect at low angles. At high altitudes, the reflected echo is so weak that CMLM is not much worse than BCA+CMLM anymore.

In Fig. 2.4, the performances of BCA+CMLM and BDML methods are compared. Note that small number of samples causes the integration output SNR to be low. Sample no= 50 is considered a small amount so that the total SNR after integration is not enough for BDML to solve ambiguities in likelihood pattern and it suffers from gross estimation errors because of these ambiguities as illustrated in Fig. 2.4.

The second experiment evaluates the estimation performance with respect to target range. In this experiment, the target altitude is constant at 30 m while it

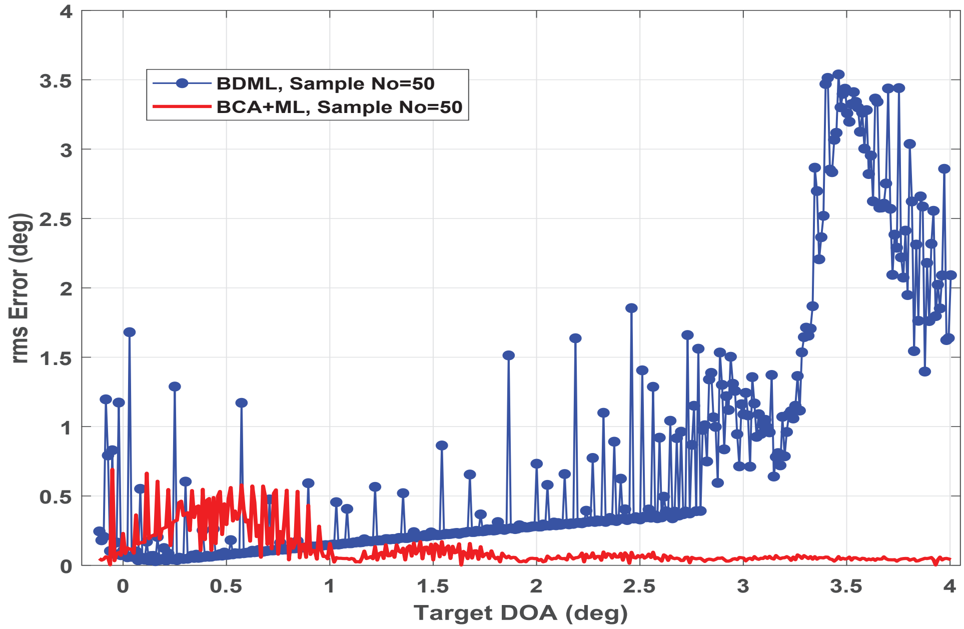


Figure 2.4: Performance comparison between BCA+CMLM and BDML. RMS error versus target arrival angle (target distance=5500m).

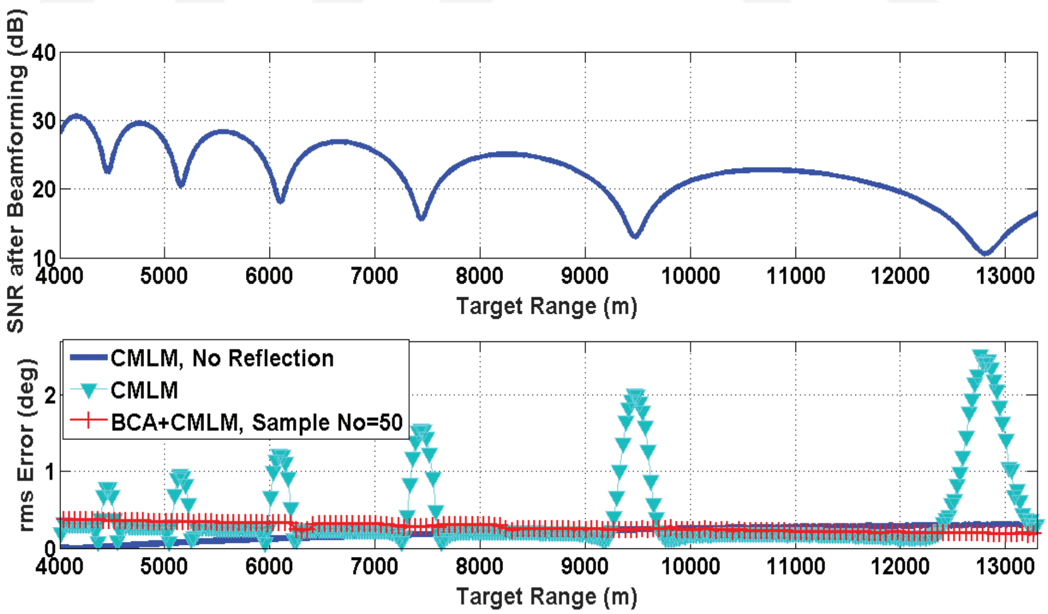


Figure 2.5: Performance improvement over CMLM provided by BCA. 1)-SNR versus target distance (altitude=30m), 2)-RMS error versus target distance (altitude=30m).

moves toward the sensor array. Free-space input SNR is 20 dB at 13.3 km. Total input SNR with respect to target distance is depicted Fig. 2.5. Latest L samples are acquired by using sliding window method and then fed to BCA algorithm. Selected parameters for this experiment are i)-Target range=[13300 m:-(Target Velocity * Sampling Time) m:4000 m], ii)-Target distance=30 m, iii)-Target Velocity=300 m/s, iv)-Sampling Time=10e-3 s, v)- ML search interval=$-\tan^{-1}(\frac{Radar-altitude}{Target-range})$: 0.1° : 5°. Fig. 2.5 reveals the advantage of BCA+CMLM over CMLM at the echo fading ranges. And Fig. 2.6 depicts that BDML yields a very rippled performance for the same reason as in the first example.

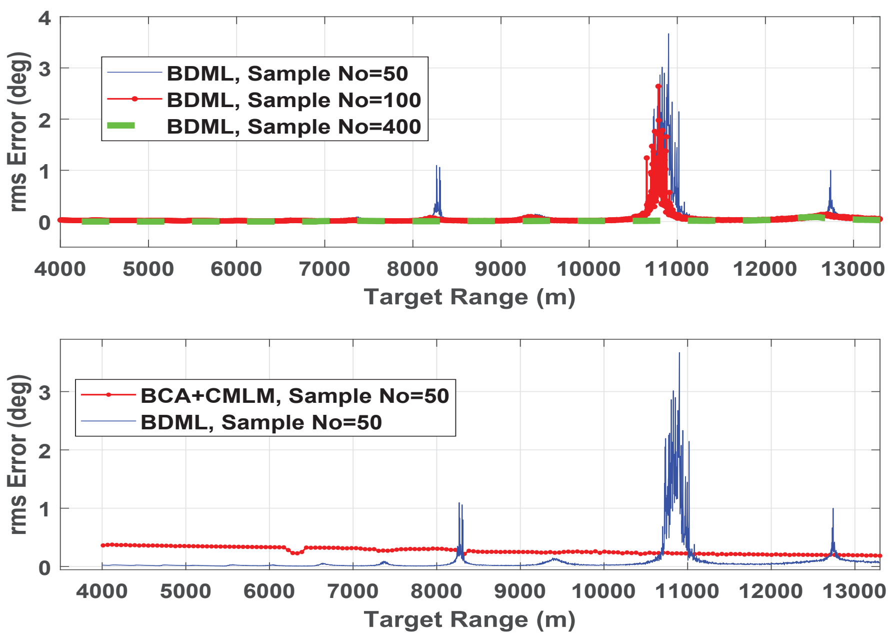


Figure 2.6: Performance comparison between BCA+CMLM and BDML. RMS error versus target distance (altitude=30m).

In the third experiment, the same simulation parameters with those in the second experiment except reflection coefficient are used. In this experiment, we take into account diffused reflection (rms height σ_h in (2.21)=50 cm). BDML assumes specular reflection, but BCA does not. Therefore, we expect worse performance from BDML in this case. This is clearly illustrated in Fig. 2.7.

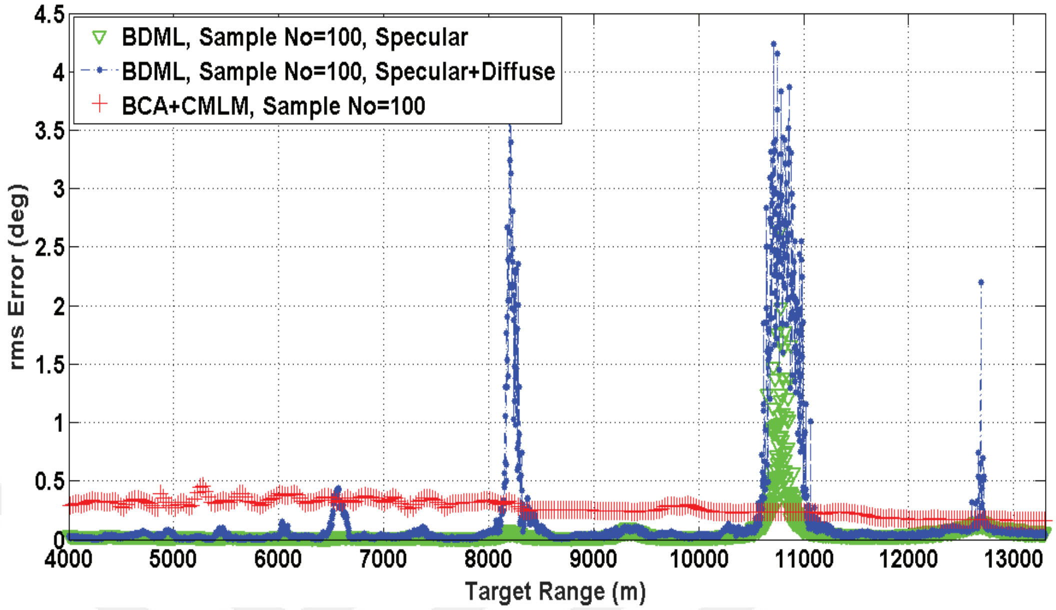


Figure 2.7: BCA+CMLM compared to BDML under specular+diffuse reflection effect. RMS error versus target distance (altitude=30m, σ_h =50cm).

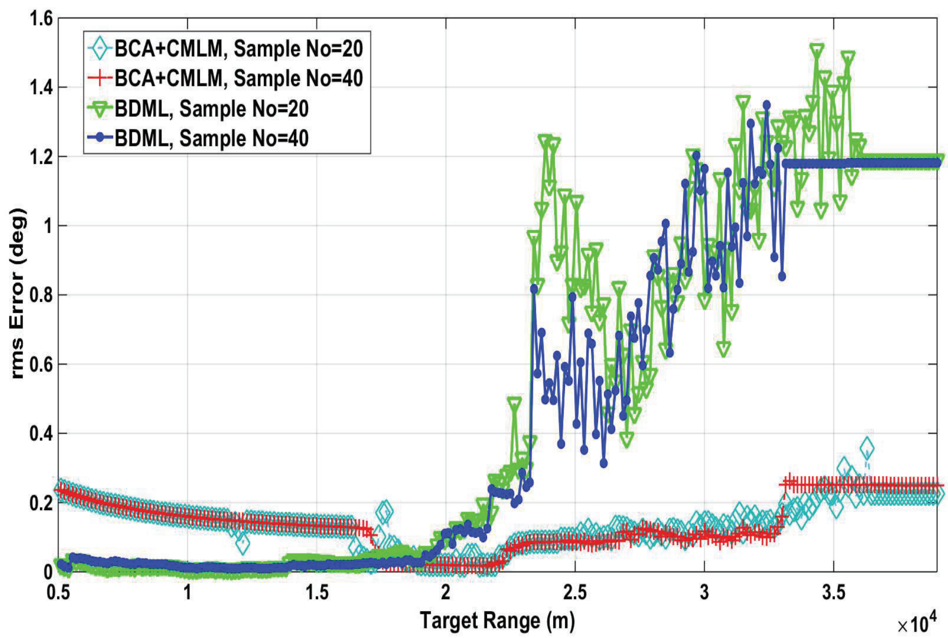


Figure 2.8: BCA+CMLM compared to BDML in Swerling-1 case. RMS error versus target distance (altitude=30m).

In the fourth experiment, Swerling-1 is used as the target RCS model. The target moves with 1500 m/s speed at constant altitude=30 m. Because of the high target speed, the maximum range is increased up to 39 km where free-space input SNR is 20 dB. The reflection is assumed to be specular only. The other simulation parameters are i)-Target range=[39 km:-(Target Velocity * Sampling Time) m:4 km], ii)-Target altitude=30 m, iii)-Target Velocity=1500 m/s, iv)-Sampling Time=100e-3 s, v)- ML search interval= $-\tan^{-1}(\frac{Radar-altitude}{Target-range})$: 0.1° : 5°,vi)- Frequency=10 GHz. In Fig. 2.8, BCA+CMLM seems not to be effected by RCS fluctuations in Swerling-1 case. On the other hand, BDML gets worse performance at far ranges where the reflection and the RCS fluctuation reduces the input SNR too much.

Chapter 3

BOUNDED COMPONENT ANALYSIS FOR INSTANTANEOUS MIXTURES BASED ON SPARSITY

3.1. Survey of Sparse Component Analysis Methods

BSS problem, in the simplest case, means to find the optimal solution to the following system of equations

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n) + \mathbf{v}(n), \qquad n = 1, \ldots, L. \tag{3.1}$$

without any knowledge about sources, environment and mixing system. In (3.1), $\mathbf{A}$ denotes mixing filter response, $\mathbf{s}(n)$ denotes source signals, and $\mathbf{v}(n)$ is additive noise. This representation is referred as instantaneous mixture model.

Since there is no a priori information to retrieve source signals by solving (3.1), we need some restrictions and assumptions. Among all, the mutual independence of sources is a powerful assumption which is used by the popular Independent Component Analysis (ICA) methods [Hyvarinen et al., 1999, Bingham and Hyvärinen, 2000]. To replace or relax the independence assumption, some other signal properties have been used by various approaches. Sparsity is such a useful measure that can be used with/without the independence assumption. The methods using sparsity assumption are classified with the name 'Sparse Component Analysis (SCA) '.

A sparse signal is most of the time zero or very close to zero relative to its variance, and occasionally gets non-zero values. Such a signal has a probability density function with a sharp peak at zero and heavy tails. The sparsity assumption in BSS means that the probability that more than one source are active at the same time is low. With respect to this definition, there are two cases assumed for sparse sources: Single dominant case and multiple dominant case. In single dominant case, only one source is active at each instant. In multiple dominant case, number of simultaneously active sources is at most equal to the size of the mixture vector. These case definitions enable SCA methods to solve blind source separation even in under-determined case where there are more sources than mixtures. In addition, SCA methods can improve performance for the (over)determined case that means there are equal or less sources than mixtures. Under-determined case becomes very challenging for large scale problems. Some methods consider large scale under-determined problems in single dominant case as in [He and Cichocki, 2006, Mourad and Reilly, 2010]. On the other hand, multiple dominant case is more usual in practice, that's why researchers show more interest in this case. Although different methods use different approaches to solve multiple dominant under-determined case, a common way is to use clustering techniques such as variants of weighted K-means [Févotte and Doncarli, 2004, Abrard and Deville, 2005, Arberet et al., 2006, Javanmard et al., 2008]. However, recursive processes in clustering increase the processing load. Therefore, alternative methods that do not use recursive processes can be seen more attractive. As an example, Theis et al. [Theis et al., 2004] present a new method that uses a generalized version of Hough Transform. Hough transform is used to detect hyperplanes where mixture data points lie and to extract lines lying at the intersections of these hyperplanes.

In practice, many signals are not often sparse in time domain. This problem can be overcome by signal transformation to different domains. Discrete Fourier Transform (DFT) [Attias, 2003], Short Time Fourier Transform (STFT) [Belouchrani and Amin, 1998, Bofill and Zibulevsky, 2000], Wavelet Transform [Candès and Donoho, 1999, Candes and Donoho, 2000, Do and Vetterli, 2005] are some of the most frequently used sparsification transforms. Another sparsification method is to decompose sources by using a signal dictionary [Lee et al., 1999b, Chen et al., 2001, Zibulevsky and Pearlmutter, 2001]. Many natural signals can be transformed to sparse counterparts using a properly defined signal dictionary. The signal representation by using a dictionary can be formulated as

$$s_i(n) = \sum_{m=1}^{M} c_{im}\varphi_m(n) \tag{3.2}$$

where scalar functions $\varphi_m(n)$ are called atoms or elements of the dictionary. Among numerous dictionaries, the most remarkable ones are wavelet-related dictionaries [Mallat, 2008] and learned dictionaries [Lewicki and Sejnowski, 2000].

There are three mixing models considered by BSS methods, and naturally, by SCA methods. The simplest model is called instantaneous mixture model where source signals arrive instantly at all receiver channels. The generative model for this case is given in eqn. (3.1). The second one is anechoic signal model where delays because of the distances between sensors are considered. A well-known and successful SCA technique is DUET introduced by Jourjine et al [Jourjine et al., 2000] and further explored by Yilmaz and Rickard [Hayvaci et al., 2013]. DUET is the first practical separation algorithm for anechoic signal model. The last mixing model is echoic case where multipath effect is considered [He et al., 2007]. The generative models for the anechoic and echoic cases are given as

$$\mathbf{x}(n) = \mathbf{A}(n) * \mathbf{s}(n) + \mathbf{v}(n), \quad n = 1, \ldots, L, \tag{3.3}$$

$$h_{ij}(n) = \sum_{k=1}^{K} c_{ij}^{k}\delta(n - \delta_{ij}^{k}), \quad n = 1, \ldots, L, \tag{3.4}$$

where operator $\mathbf{A}$ denotes channel impulse response, h_{ij} is the ij^{th} element of the channel impulse response, $*$ is convolution operator and $K = 1$ for the anechoic case.

To solve the BSS problem, SCA methods use various diversities. Among all, ℓ_0 norm provides the most powerful and simplest representation of sparsity. As an example approach based on ℓ_0 norm minimization, Mishali and Eldar [Mishali and Eldar, 2009] propose a method that assumes orthonormal mixing matrix and k-sparsity. The mixing matrix and sources are estimated alternatively using an ℓ_0 norm constrained objective function. However, the approach is based on noise-free linear mixing model and it should be reminded ℓ_0 norm is vulnerable to noise. Although ℓ_0 norm is the best representation of sparsity, its minimization is NP-hard and can be solved only by greedy algorithms. The sparsity constraint problem can be solved via a convex relaxation, in other words, a weaker sparsity measure is chosen to find approximately sparse solutions. ℓ_p norms for $0 < p \leq 1$ are of most interest in the mathematical analysis of sparsity. Unfortunately, $0 < p < 1$ leads to non-convex optimization which is very difficult to solve. On the other hand, the choice $p = 1$ can be solved by convex optimization methods. SCA approaches which are based on optimization techniques mostly use ℓ_1 norm to promote sparsity. In [Li et al., 2003], Li et al investigate ℓ_0 versus ℓ_1 norm minimization comparison and the superiority of ℓ_1 norm in the noisy case. They use K-means clustering and ℓ_1 norm minimization for estimating the mixing matrix and sources in the under-determined case. This method is a good example of ℓ_1 norm constraint in source estimation stage.

Instead of estimating the mixing matrix by various clustering algorithms and then sources in two steps as done in [Li et al., 2003], another approach is to use an iterative procedure to estimate them alternatively at a single step. This would remove disadvantages of using clustering for mixing matrix estimation. Gorodnitsky and Rao [Gorodnitsky and Rao, 1997] propose a single-step approach called FOCal Under-determined System Solver (FOCUSS) considering more general $\ell_{(p<1)}$ norm as the sparsity measure in optimization process. However, FOCUSS works with single measurement vector, that's why Cotter et al. [Cotter et al., 2005] propose extension of FOCUSS algorithm named as MFOCUSS for multiple measurement vectors. MFOCUSS estimates a separation matrix instead of the mixing matrix. An iterative procedure is employed to estimate the separation matrix and the source matrix alternatively. MFOCUSS is useful in the noiseless case. For the noisy case, an algorithm called Regularized MFOCUSS is introduced in [Cotter et al., 2005]. Another single step approach proposed by He et al [He et al., 2007] is used for anechoic and echoic mixture cases. In this approach, mixture signals are mapped to spectral data using DTFT and it is assumed that the transformed sources are sparse in frequency domain. Since convolution is transformed to multiplication, anechoic and echoic mixture cases are converted to instantaneous mixture case. As the last single-step approach to be mentioned here, Relative Newton [Zibulevsky, 2003] minimizes a constrained quasi maximum likelihood function that is obtained using the joint probability density function of the sources, i.e.,

$$L(\mathbf{W};\mathbf{X}) = -\log|det\mathbf{W}| + \frac{1}{T}\sum_{i,n} h(\mathbf{W}_i\mathbf{x}(n)), \tag{3.5}$$

in order to estimate the separation matrix $\mathbf{W}$ where $h(.) = -\log f(.)$, and $f(.)$ is the probability density function of the sources. The function is named "quasi" because an approximate function is used for $h(.)$ instead of $-\log f(.)$. Since the sources are sparse, $h(.)$ is approximated with an absolute value function. A sequential optimization with the Relative Newton method is used to solve (3.5). The Relative Newton method provides the convergence of the sequential optimization to be invariant to the value of the mixing matrix.

In source separation problem, we essentially aim to estimate the source matrix. The mixing matrix is only a fundamental ingredient for this estimation. Some researchers considered pre-constructed dictionaries to identify a redundant representation of the mixing matrix [Kreutz-Delgado et al., 2003]. Such an over-complete dictionary could then be used as a fixed dictionary and only the sources could be estimated instead of the alternating estimation. Chen et al [Chen et al., 2001] developed an optimization based approach for multiple-dictionary representation problem. In this approach, the minimization problem is formulated as

$$\min \sum_i \|\alpha_i\|_1 \quad \textit{subject to} \quad \mathbf{s}(n) = \sum_i \mathbf{\Phi}_i\alpha_i \tag{3.6}$$

where $\mathbf{\Phi_i}, i = 1, 2, \ldots$ are different dictionary elements to represent the sources as sparse signals. The authors used Basis Pursuit (BP) to find signal representations in over-complete dictionaries by convex optimization. There are different options for generating the over-complete dictionaries. A widely used way is to use some theoretical models such as wavelet packets [Chen et al., 2001, Mallat, 2008], cosine packets [Qayyum et al., 2015], curvelets, contourlets [Rubinstein et al., 2010], etc. Among all, a novel single step SCA approach that uses such theoretical models for sparse representation in instantaneous mixing case is named Generalized Morphological Component Analysis (GMCA) [Bobin et al., 2007]. GMCA relies on morphological diversity beside sparsity. In morphological diversity assumption, sources $\{s_i\}_{n=1,\ldots,N.}$ are linear combinations of the morphological components, i.e.,

$$s_i = \sum_{k=1}^{D} \phi_{ik} \tag{3.7}$$

and the morphological components are sparsely represented in an associated basis. The representation bases of different morphological components must be incoherent, and this criterion can be provided with an orthonormal sparsifying transform set such as wavelets. GMCA estimates the morphological components by using an alternating minimization and iterative threshold algorithm. One of the main bottlenecks that may prevent the potential applications of the dictionary based SCA methods is the computational complexity of their optimization algorithms.

3.2. Sparse Bounded Component Analysis for Instantaneous Mixtures

In this section, a novel Bounded Component Analysis approach that promotes sparsity assumption by using magnitude boundedness is introduced [Babatas and Erdogan, 2016, Babatas and Erdogan, 2018a]. Bounded Component Analysis is a recent BSS framework that is based on domain separability assumption. This assumption is less restrictive than the statistical independence assumption in ICA. Cruces showed in [Cruces, 2010] that the source boundedness can be utilized to separate temporally or spatially dependent sources as well as independent sources. In [Erdogan, 2013], Erdogan presented a deterministic framework in which the objective function is defined with respect to two geometric objects regarding only separator output samples, called as principal hyper-ellipsoid and bounding hyper-rectangle. All of the bounded source samples lie in the bounding hyper-rectangle whose side lengths are defined with ℓ_∞ norm. We explain this BCA framework in 2.1.

Our current scheme for SBCA adopts a geometrical setting as in BCA. In this setting, the source signals are bounded in an ℓ_1-norm-ball while the BCA framework sets an ℓ_∞-norm-ball. The highlights of the proposed approach are,

- Completely deterministic framework,
- Usability to separate both statistically dependent (even correlated) and independent sources,
- Good performance for short data record because of its completely deterministic structure,
- A single step estimation approach (estimation of only separator matrix).

3.2.1 Blind Source Separation Signal Setup

In the instantaneous BSS signal setup,

- There are p sources denoted by the sample set $\mathcal{S} = \{\mathbf{s}(n) \in \mathbb{R}^p, n = 1, \dots L\}$, and the source samples lie in an ℓ_1 norm ball, i.e., $\mathbf{s}(n) \in \mathcal{B}_s, \quad n = 1, \dots, L$ where

$$\mathcal{B}_s = \{\mathbf{q} \in \mathbb{R}^p | \|\mathbf{q}\|_1 < 1\}. \tag{3.8}$$

  Despite the assumption that the sources have identical unity range, there is no loss of generality. In the more general case, $\mathcal{B}_s$ can be replaced with the weighted ℓ_1 norm ball.

- Source signals are mixed in channels which have linear and memoryless time response denoted by $\mathbf{A} \in \mathbb{R}^{q \times p}$. It is assumed that $\mathbf{A}$ is full rank, in other words, $q \geq p$.

- The mixture signals are formulated as

$$\mathbf{x}(n) = \mathbf{A}(n)\mathbf{s}(n), \quad n = 1, \ldots, L. \tag{3.9}$$

- The separator outputs are formulated as

$$\mathbf{y}(n) = \mathbf{W}\mathbf{x}(n), \qquad n = 1, \ldots L, \tag{3.10}$$

  where $\mathbf{W} \in \mathbb{R}^{p \times q}$ indicates separator filter response.

- The total system response is written as

$$\mathbf{F} = \mathbf{W}\mathbf{A}, \tag{3.11}$$

$$\mathbf{y}(n) = \mathbf{F}\mathbf{s}(n) \quad n = 1, \ldots, L. \tag{3.12}$$

For perfect separation, the separator matrix must provide the total system response to satisfy the equality

$$\mathbf{F} = \mathbf{P}\mathbf{D}. \tag{3.13}$$

It means the total system response matrix must be transparent. In this formula, $\mathbf{P}$ is a permutation matrix and $\mathbf{D}$ is a full rank diagonal scaling matrix.

3.2.2 Objective Function

To create an optimization structure, the volume ratio of the two geometric objects called principal hyper-ellipsoid and bounding ℓ_1 norm ball is used.

Fig.3.1 illustrates the so called geometric objects' volumes. In this figure,

- On the left side:
  - the source samples are denoted by the blue dots,

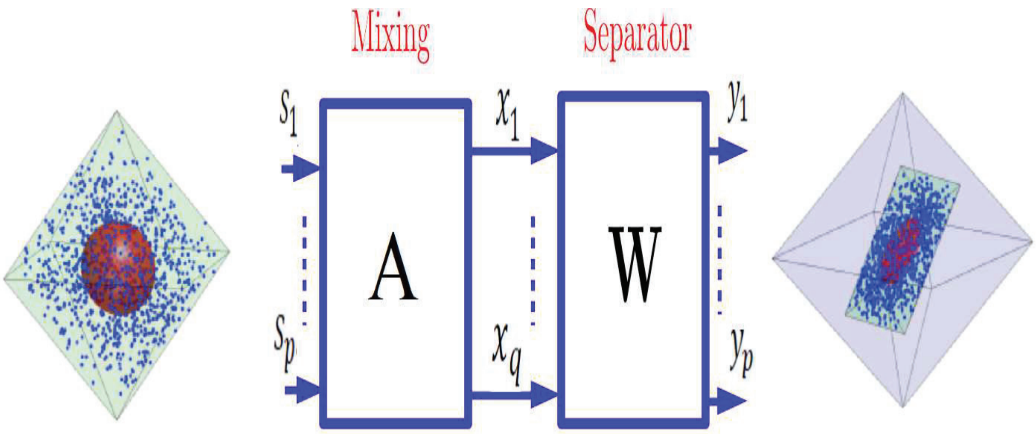


Figure 3.1: Geometric objects used in the SBCA framework.

- green diamond shaped box indicates the bounding ℓ_1 norm ball as defined in (3.8),
- red hyper-ellipsoid indicates the principal hyper-ellipsoid defined as $\mathcal{E}_s = \{\mathbf{q} | (\mathbf{q} - \hat{\mu}_s)^T \hat{\mathbf{R}}_{\mathbf{s}}^{-1} (\mathbf{q} - \hat{\mu}_s) \leq 1\}$ where $\hat{\mu}_s = \frac{1}{L}\sum_{n=1}^{L} \mathbf{s}(n)$ and $\hat{\mathbf{R}}_{\mathbf{s}} = \frac{1}{L}\sum_{n=1}^{L} \mathbf{s}(n)\mathbf{s}(n)^T - \hat{\mu}_s \hat{\mu}_s^T$.

- On the right side:
  - the separator output samples are denoted by the blue dots,
  - purple diamond shaped box indicates the bounding ℓ_1 norm ball defined as

$$\mathcal{B}_y = \{\mathbf{q} | \|\mathbf{q}\|_1 \leq \max_{n \in \{1,\ldots,L\}} \|\mathbf{y}(n)\|_1\}, \tag{3.14}$$

  - red hyper-ellipsoid indicates the principal hyper-ellipsoid defined as $\mathcal{E}_y = \{\mathbf{q} | (\mathbf{q} - \hat{\mu}_y)^T \hat{\mathbf{R}}_{\mathbf{y}}^{-1} (\mathbf{q} - \hat{\mu}_y) \leq 1\}$ where $\hat{\mu}_y = \frac{1}{L}\sum_{n=1}^{L} \mathbf{y}(n)$ and $\hat{\mathbf{R}}_{\mathbf{y}} = \frac{1}{L}\sum_{n=1}^{L} \mathbf{y}(n)\mathbf{y}(n)^T - \hat{\mu}_y \hat{\mu}_y^T$,
  - green polytope indicates the image of $\mathcal{B}_s$ under the mapping $\mathbf{F}$.

In accordance to these definitions, the SBCA objective is defined as

$$\bar{J}(\mathbf{W}) = \frac{\text{Volume}(\mathcal{E}_y)}{\text{Volume}(\mathcal{B}_y)} \tag{3.15}$$

$$= \frac{C_e\sqrt{\det(\hat{\mathbf{R}}_{\mathbf{y}})}}{C_l(\max_{n\in\{1,\dots,L\}} \|\mathbf{y}(n)\|_1)^p}, \tag{3.16}$$

where $C_e = \pi^{p/2}/\Gamma(p/2+1)$ and $C_l = 2^p/p!$ are the dimension dependent constants, and can be dropped. Consequently, the resulting objective function is

$$J(\mathbf{W}) = \frac{\sqrt{\det(\hat{\mathbf{R}}_{\mathbf{y}})}}{(\max_{n\in\{1,\dots,L\}} \|\mathbf{y}(n)\|_1)^p}. \tag{3.17}$$

3.2.3 Global Optimality of Perfect Separators

Before starting the proof for the proposed objective function, we define the following assumption which ensures that the global maximums of (3.17) give perfect separation.

Assumption (A3): Source sample set $\mathcal{S}$ contains the vertices of the bounding ℓ_1 norm ball $\mathcal{B}_s$.

Based on the assumption (**A3**), we propose our theorem for perfect separation.

Theorem-2: If (**A3**) is correct, then all global maximums of (3.17) correspond to perfect separators that satisfy the equality in (3.13).

Proof: Firstly, we write the objective function according to the argument $\mathbf{F} = \mathbf{WA}$:

$$J(\mathbf{F}) = \frac{|\det(\mathbf{F})|\sqrt{\det(\hat{\mathbf{R}}_{\mathbf{s}})}}{(\max_{n\in\{1,\dots,L\}} \|\mathbf{Fs}(n)\|_1)^p}. \tag{3.18}$$

Using the assumption (**A3**), we can rewrite the denominator of (3.18) as

$$(\max_{n\in\{1,\dots,L\}} \|\mathbf{Fs}(n)\|_1)^p \leq \|\mathbf{F}\|_{1,1}^p, \tag{3.19}$$

where $\|\mathbf{F}\|_{1,1}$ is the induced matrix norm [Trefethen and Bau III, 1997]. $\|\mathbf{F}\|_{1,1}$ can be turned out to be

$$\|\mathbf{F}\|_{1,1} = \left\| \begin{bmatrix} \|\mathbf{F}_{:,1}\|_1 & \|\mathbf{F}_{:,2}\|_1 & \dots & \|\mathbf{F}_{:,p}\|_1 \end{bmatrix} \right\|_\infty . \tag{3.20}$$

If (**A3**) holds, then (3.19) turns into an equality. This result yields

$$J(\mathbf{F}) = \frac{|\det(\mathbf{F})|\sqrt{\det(\hat{\mathbf{R}}_\mathbf{s})}}{\left\| \begin{bmatrix} \|\mathbf{F}_{:,1}\|_1 & \|\mathbf{F}_{:,2}\|_1 & \dots & \|\mathbf{F}_{:,p}\|_1 \end{bmatrix} \right\|_\infty^p} . \tag{3.21}$$

We can write on the basis of this statement as

$$J(\mathbf{F}) \leq \frac{|\det(\mathbf{F})|\sqrt{\det(\hat{\mathbf{R}}_\mathbf{s})}}{(\left\| \begin{bmatrix} \|\mathbf{F}_{:,1}\|_1 & \|\mathbf{F}_{:,2}\|_1 & \dots \|\mathbf{F}_{:,p}\|_1 \end{bmatrix} \right\|_1 /p)^p}, \tag{3.22}$$

$$\leq \frac{|\det(\mathbf{F})|\sqrt{\det(\hat{\mathbf{R}}_\mathbf{s})}}{\|\mathbf{F}_{:,1}\|_1 \|\mathbf{F}_{:,2}\|_1 \dots \|\mathbf{F}_{:,p}\|_1}, \tag{3.23}$$

$$\leq \frac{|\det(\mathbf{F})|\sqrt{\det(\hat{\mathbf{R}}_\mathbf{s})}}{\|\mathbf{F}_{:,1}\|_2 \|\mathbf{F}_{:,2}\|_2 \dots \|\mathbf{F}_{:,p}\|_2}, \tag{3.24}$$

$$\leq \sqrt{\det(\hat{\mathbf{R}}_\mathbf{s})}, \tag{3.25}$$

where

- The inequality between ℓ_1 and ℓ_∞ norms in (3.22) and the arithmetic-geometric mean inequality in (3.23) are accomplished only if there is the same ℓ_1 norm in all the columns of $\mathbf{F}$,

- The inequality between ℓ_1 and ℓ_2 norms in (3.24) is accomplished only if there is only one non-zero entry in each column of $\mathbf{F}$,

- Hadamard's Inequality in (3.25) is accomplished only if each column of $\mathbf{F}$ is orthogonal to the other columns.

As a consequence, the upper limit in (3.25) is accomplished only if $\mathbf{F} = \alpha\mathbf{P}$ where $\mathbf{P}$ is a permutation matrix and $\alpha \neq 0$. If the sources have varying ranges, the global optimizers should use more general form $\mathbf{F} = \mathbf{PD}$ because of the scaling uncertainty of sources.

3.2.4 Comparison of SBCA to ICA with Sparsity Promoting Marginals

SBCA objective (3.17) is converted into a difference form by taking the logarithm

$$\mathcal{J}(\mathbf{W}) = \log(J(\mathbf{W})) = \frac{1}{2}\log(\det(\hat{\mathbf{R}}_{\mathbf{y}})) - p\log(\max_{n\in\{1,\ldots,L\}} \|\mathbf{y}(n)\|_1). \tag{3.26}$$

When the constant term $\frac{1}{2}\log(\det(\hat{\mathbf{R}}_{\mathbf{s}}))$ is neglected and amount of sources is equal to amount of mixtures, this objective is simplified to

$$\mathcal{J}(\mathbf{W}) = \log(|\det(\mathbf{W})|) - p\log(\max_{n\in\{1,\ldots,L\}} \|\mathbf{y}(n)\|_1). \tag{3.27}$$

(3.27) is analogous to the Maximum Likelihood (ML) expression

$$\mathcal{L}(\mathbf{W}) = \log(|\det(\mathbf{W})|) + \sum_{k=1}^{p}\sum_{n=1}^{L} \log(f_s(y_k(n))), \tag{3.28}$$

where $f_s(\cdot)$ denotes the source marginal density. If the density is Laplace, it provides a more simplified likelihood expression as

$$\mathcal{L}(\mathbf{W}) = \log(|\det(\mathbf{W})|) - \sum_{n=1}^{L} \|\mathbf{y}(n)\|_1. \tag{3.29}$$

The likelihood function in (3.28) can be expressed in a more general form

$$\mathcal{L}(\mathbf{W}) = \log(|\det(\mathbf{W})|) + \sum_{k=1}^{p}\sum_{n=1}^{L} v(y_k(n)). \tag{3.30}$$

(3.30) is a quasi-ML expression (e.g. [Zibulevsky and Pearlmutter, 2001]) where v is not necessarily derived from $\log(f_s(\cdot))$.

The expressions in (3.28) and (3.30) (and therefore in (3.29)) are based on the source independence assumption reflected by the summation term. However, the

SBCA objective in (3.27) significantly does not take benefit of such assumptions. Thus, our SBCA scheme can be applied to dependent sources having non-separable joint densities.

3.2.5 Iterative Algorithm for Separator Estimation

Noting that the SBCA objective in (3.26) is non-convex and not differentiable everywhere, an iterative update equation can be still obtained by using Clarke sub-differential method [Bagirov et al., 2014].

$\mathcal{J}(\mathbf{W})$ involves two parts:

$$\mathcal{J}(\mathbf{W}) = \underbrace{\frac{1}{2}\log(\det(\hat{\mathbf{R}}_{\mathbf{y}}))}_{\mathcal{J}_1(\mathbf{W})} - p \underbrace{\log(\max_{n\in\{1,\ldots,L\}} \|\mathbf{y}(n)\|_1)}_{\mathcal{J}_2(\mathbf{W})}. \tag{3.31}$$

$\mathcal{J}_1(\mathbf{W})$ is convex differentiable, and $\mathcal{J}_2(\mathbf{W})$ is convex non-smooth. Let us consider $\mathcal{J}_1(\mathbf{W})$ and $\mathcal{J}_2(\mathbf{W})$ separately in order to derive an iterative update rule.

- The gradient term to maximize the volume of $\mathcal{E}(\mathcal{Y})$ is described as

$$\nabla\mathcal{J}_1(\mathbf{W}) = \left(\mathbf{W}\hat{\mathbf{R}}_{\mathbf{y}}\mathbf{W}^T\right)^{-1}\mathbf{W}\hat{\mathbf{R}}_{\mathbf{x}}. \tag{3.32}$$

- If $\mathcal{J}_2(\mathbf{W})$ is turned out to be

$$\mathcal{J}_2(\mathbf{W}) = \max_{n\in\{1,\ldots,L\}} f_n(\mathbf{W}), \tag{3.33}$$

  where $f_n(\mathbf{W}) = \|\mathbf{W}\mathbf{x}(n)\|_1$, the sub-differential set for $\mathcal{J}_2(\mathbf{W})$ yields to be [Boyd and Vandenberghe, 2004]

$$\partial\mathcal{J}_2(\mathbf{W}) = \mathbf{Co} \cup \{\partial f_n(\mathbf{W}) : f_n(\mathbf{W}) = \mathcal{J}_2(\mathbf{W})\}. \tag{3.34}$$

  For all $\mathbf{v} = \begin{bmatrix} v_1 & \cdots & v_N \end{bmatrix} \in \mathbb{R}^N$, one has $\|\mathbf{v}\|_1 = |v_1| + |v_2| + \cdots + |v_N|$.

Therefore the derivative of the map $\mathbf{v} \to \|\mathbf{v}\|_1$ can be written as,

$$\frac{\partial}{\partial v_j}(\mathbf{v} \to \|\mathbf{v}\|_1) = \frac{v_j}{|v_j|} = sign(v_j), \quad 1 \leq j \leq N, \tag{3.35}$$

provided that $v_j \neq 0$. If $v_j = 0$, then we have

$$\frac{\partial}{\partial v_j}(\mathbf{v} \to \|\mathbf{v}\|_1) = \alpha_j \in [-1, 1]. \tag{3.36}$$

As a result, the sub-differential set of $f_n(\mathbf{W})$ is obtained as

$$\begin{aligned}
\partial f_n(\mathbf{W}) &= \partial(\| \begin{bmatrix} \mathbf{W}_{1,:}\mathbf{x}(n) & \cdots & \mathbf{W}_{p,:}\mathbf{x}(n) \end{bmatrix}^T \|_1), && (3.37)\\
&= \| \begin{bmatrix} \frac{\partial(\mathbf{W}_{1,:}\mathbf{x}(n))}{\partial(\mathbf{W}_{1,:})} & \cdots & \frac{\partial(\mathbf{W}(p,:)\mathbf{x}(n))}{\partial(\mathbf{W}_{p,:})} \end{bmatrix}^T \|_1, && (3.38)\\
&= \{\mathbf{q}\mathbf{x}(l)^T : q_i = \text{sign}\{y_i(l)\} + 1_{y_i(l)=0}\alpha_i, && (3.39)\\
&\quad \alpha_i \in [-1, 1]\}.
\end{aligned}$$

sign$\{\mathbf{y}(l)\}\mathbf{x}(l)^T$ is a member of sub-gradient set $\partial f_n(\mathbf{W})$. By using this sub-gradient choice, $\partial \mathcal{J}_2(\mathbf{W})$ is expressed as

$$\partial \mathcal{J}_2(\mathbf{W}) = \sum_{l \in \mathcal{I}_{\mathbf{W}}} \lambda_l \text{sign}\{\mathbf{y}(l)\}\mathbf{x}(l)^T, \tag{3.40}$$

where $\mathcal{I}_W = \{l : \|\mathbf{W}\mathbf{x}(l)\|_1 = \mathcal{J}_2(\mathbf{W})\}$ denotes the index set for maximum ℓ_1-norm separator outputs, and λ_n's are convex combination coefficients satisfying $\lambda_l \geq 0$ and $\sum_{l \in \mathcal{I}_{\mathbf{W}}} \lambda_l = 1$.

Based on (3.32) and (3.40), an iterative update equation to maximize $\mathcal{J}(\mathbf{W})$ is given as

$$\begin{aligned}
\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} + \mu^{(t)}(\hat{\mathbf{R}}_{\mathbf{y}^{(t)}})^{-1}\mathbf{W}^{(t)}\hat{\mathbf{R}}_{\mathbf{x}} - \\
\frac{p}{\max_{n \in \{1,\ldots,L\}} \|\mathbf{y}^{(t)}(n)\|_1} \sum_{l \in \mathcal{I}_{\mathbf{W}^{(t)}}} \lambda_l^{(t)} \text{sign}\{\mathbf{y}^{(t)}(l^{(t)})\}\mathbf{x}(l^{(t)})^T.
\end{aligned} \tag{3.41}$$

If we ignore all convex combination coefficients but only one, by randomly selecting an index position from $\mathcal{I}_{\mathbf{W}^{(t)}}$ at each iteration, this simplifies (3.41) to

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} + \mu^{(t)}(\hat{\mathbf{R}}_{\mathbf{y}^{(t)}})^{-1}\mathbf{W}^{(t)}\hat{\mathbf{R}}_{\mathbf{x}} - \frac{p}{\max\limits_{n\in\{1,\ldots,L\}} \|\mathbf{y}^{(t)}(n)\|_1} \text{sign}\{\mathbf{y}^{(t)}(l^{(t)})\}\mathbf{x}(l^{(t)})^T. \quad (3.42)$$

3.2.6 Algorithmic Extensions for SBCA

In this section, at first it is described how to sparsify signals with an appropriate linear transformation. Then, the extension of the SBCA iterative algorithm for complex sources is proposed. The iterative algorithm is modified for the noisy case in the last subsection.

- **Sparsifying Transformations**:

  Lots of practical signals are not naturally sparse, but an appropriate linear transformation can provide sparse counterparts for most of them. The mixture model is defined as

$$\mathbf{X} = \mathbf{A}\mathbf{S}, \quad (3.43)$$

  where matrices for source and mixture samples are formulated as

$$\mathbf{S} = \begin{bmatrix} \mathbf{s}(1) & \mathbf{s}(2) & \ldots & \mathbf{s}(L) \end{bmatrix}, \quad (3.44)$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}(1) & \mathbf{x}(2) & \ldots & \mathbf{x}(L) \end{bmatrix}. \quad (3.45)$$

  After sparsification, we can express the transformed sources as

$$\mathbf{S}_T = \mathbf{S}\boldsymbol{\Phi}. \quad (3.46)$$

  where $\boldsymbol{\Phi} \in \mathbb{R}^{L\times L}$ denotes the sparsifying transform matrix. However, in practice, sparsifying transform can not be applied to sources, but mixtures. If we denote the sparsified mixtures as

$$\mathbf{X}_T = \mathbf{X}\boldsymbol{\Phi}, \tag{3.47}$$

substituting (3.43) gives

$$\begin{aligned} \mathbf{X}_T &= \mathbf{A}\mathbf{S}\boldsymbol{\Phi}, & (3.48) \\ &= \mathbf{A}\mathbf{S}_T. & (3.49) \end{aligned}$$

(3.49) results that transforming mixtures instead of sources is equivalent to mixing the transformed sources under linear transformation. As a result, we can apply SBCA algorithm to transformed mixtures $\mathbf{X}_T$ in order to obtain the separator matrix for original mixtures. As an example of this procedure, the natural images are sparsified using Morlet transformation in section 3.2.8.

- **Extension to Complex Signals**:

The isomorphism based approach in [Erdogan, 2013] will be used for complex extension of the SBCA approach. Before the proposition, we need to make some definitions.

Initially, the operator $\Upsilon : \mathbb{C}^p \to \mathbb{R}^{2p}$

$$\Upsilon(\mathbf{x}) = \left[\mathcal{R}e\{\mathbf{x}^T\}\ \mathcal{I}m\{\mathbf{x}^T\}\right]^T \tag{3.50}$$

represents isomorphism between a complex vector of size p and a real vector of size $2p$. For simplification, $\acute{\mathbf{u}}$ will denote real isomorphic counterpart of the complex vector $\mathbf{u}$. We also define the operator $\Gamma : \mathbb{C}^{q\times p} \to \mathbb{R}^{2q\times 2p}$ for mapping complex matrices to real isomorphic matrices as

$$\Gamma(\mathbf{U}) = \begin{bmatrix} \mathcal{R}e\{\mathbf{U}\} & -\mathcal{I}m\{\mathbf{U}\} \\ \mathcal{I}m\{\mathbf{U}\} & \mathcal{R}e\{\mathbf{U}\} \end{bmatrix}. \tag{3.51}$$

Based on these definitions, the real isomorphic mixture and separator signals turns out to

$$\acute{\mathbf{x}}(n) = \Gamma(\mathbf{A})\acute{\mathbf{s}}(n), \tag{3.52}$$

$$\acute{\mathbf{y}}(n) = \Gamma(\mathbf{W})\acute{\mathbf{x}}(n). \tag{3.53}$$

The notation $\acute{\mathcal{X}} = \{\acute{\mathbf{x}} : \mathbf{x} \in \mathcal{X}\}$ represents the set of the real-isomorphic mixture vectors and $\acute{\mathcal{Y}} = \{\acute{\mathbf{y}} : \mathbf{y} \in \mathcal{Y}\}$ represents the set of the real-isomorphic separator outputs. After giving the necessary mathematical definitions, we formulate the following objective and iterative update equation for complex source signals as

$$J_c(\mathbf{W}) = \frac{\sqrt{\det(\hat{\mathbf{R}}_{\acute{\mathbf{y}}})}}{\left(\max_{n\in\{1,\ldots,L\}} \|\acute{\mathbf{y}}(n)\|_1\right)^{2p}}. \tag{3.54}$$

To derive an iterative algorithm for complex sources, the ratio form in (3.54) is transformed to difference form by taking logarithm:

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} + \mu^{(t)}(\mathbf{W}^{(t)}_{logdet} - \mathbf{W}^{(t)}_{subg}), \tag{3.55}$$

where

$$\begin{aligned}
\mathbf{W}^{(t)}_{logdet} = \frac{1}{2}\Bigg(&\left[\begin{array}{cc}\mathbf{I} & \mathbf{0}\end{array}\right]\hat{\mathbf{R}}^{-1}_{\acute{\mathbf{y}}^{(t)}}\Gamma(\mathbf{W}^{(t)})\hat{\mathbf{R}}_{\acute{\mathbf{x}}}\left[\begin{array}{c}\mathbf{I}\\ \mathbf{0}\end{array}\right] + \\
&\left[\begin{array}{cc}\mathbf{0} & \mathbf{I}\end{array}\right]\hat{\mathbf{R}}_{\acute{\mathbf{y}}^{(t)}}{}^{-1}\Gamma(\mathbf{W}^{(\mathbf{t})})\hat{\mathbf{R}}_{\acute{\mathbf{x}}}\left[\begin{array}{c}\mathbf{0}\\ \mathbf{I}\end{array}\right] + \\
&j\Bigg(\left[\begin{array}{cc}\mathbf{I} & \mathbf{0}\end{array}\right]\hat{\mathbf{R}}^{-1}_{\acute{\mathbf{y}}^{(t)}}\Gamma(\mathbf{W}^{(\mathbf{t})})\hat{\mathbf{R}}_{\acute{\mathbf{x}}}\left[\begin{array}{c}\mathbf{0}\\ -\mathbf{I}\end{array}\right] + \\
&\left[\begin{array}{cc}\mathbf{0} & \mathbf{I}\end{array}\right]\hat{\mathbf{R}}_{\acute{\mathbf{y}}^{(t)}}{}^{-1}\Gamma(\mathbf{W}^{(t)})\hat{\mathbf{R}}_{\acute{\mathbf{x}}}\left[\begin{array}{c}\mathbf{I}\\ \mathbf{0}\end{array}\right]\Bigg)\Bigg),
\end{aligned} \tag{3.56}$$

and

$$\mathbf{W}_{subg}^{(t)} = \frac{2p}{\max_{n\in\{1,\ldots,L\}} \|\mathbf{y}^{(t)}(n)\|_1} \text{signc}\{\mathbf{y}^{(t)}(l^{(t)})\}\mathbf{x}(l^{(t)})^H. \tag{3.57}$$

Alternatively, we can simplify the cost function (3.54) as

$$J_{ca}(\mathbf{W}) = \frac{\sqrt{\det(\hat{\mathbf{R}}_{\mathbf{y}})}}{\left(\max_{n\in\{1,\ldots,L\}} \|\mathbf{y}(n)\|_1\right)^p}, \tag{3.58}$$

where the covariance term in the numerator is turned out to be the complex covariance in accordance with the equality $|\det(\Gamma(\mathbf{F}))| = |\det(\mathbf{F})|^2$ [Erdogan, 2013], and $2p$ is changed to p. Finally, the update equation is simplified to

$$\mathbf{W}_{logdet}^{(t)} = \hat{\mathbf{R}}_{\mathbf{y}^{(t)}}^{-1}\mathbf{W}^{(t)}\hat{\mathbf{R}}_{\mathbf{x}}, \tag{3.59}$$

$$\mathbf{W}_{subg}^{(t)} = \frac{p}{\max_{n\in\{1,\ldots,L\}} \|\mathbf{y}^{(t)}(n)\|_1} \text{signc}\{\mathbf{y}^{(t)}(l^{(t)})\}\mathbf{x}(l^{(t)})^H. \tag{3.60}$$

- **Extension to Noisy Case**:

If the noisy version of the mixture model is written as

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \mathbf{n}(k) = \mathbf{x}_{noiseless}(k) + \mathbf{n}(k), \tag{3.61}$$

where $\mathbf{n}(k)$ denotes noise, the corresponding separator output is given by

$$\mathbf{y} = \mathbf{W}\mathbf{x}_{noiseless}(k) + \mathbf{W}\mathbf{n}(k) \tag{3.62}$$

$$= \mathbf{y}_{noiseless}(k) + \mathbf{v}(k). \tag{3.63}$$

The main goal is to find the locations of the maximum ℓ_1 norm of $\mathbf{y}_{noiseless}(k)$ from the noisy separator outputs. The presence of noise makes it a stochastic estimation problem for which each sample is assigned a probability. More clearly, let the probability of index k for being in the set $\mathcal{I}_{\mathbf{W}}$ be denoted by

$$p_{\mathcal{I}_{\mathbf{W}}}(k) = \Pr(k \in \mathcal{I}_{\mathbf{W}}). \tag{3.64}$$

The iterative update equation in (3.41) is modified to include the pmf values:

$$\begin{aligned}\mathbf{W}^{(t+1)} &= \mathbf{W}^{(t)} + \mu^{(t)}((\hat{\mathbf{R}}_{\mathbf{y}^{(t)}})^{-1}\mathbf{W}^{(t)}\hat{\mathbf{R}}_{\mathbf{x}} - \\ &\qquad \frac{p}{\max_{n\in\{1,\ldots,L\}} \|\mathbf{y}^{(t)}(n)\|_1} \sum_{l=1}^{L} p_{\mathcal{I}_{\mathbf{W}}^{(t)}}(l)\text{sign}\{\mathbf{y}^{(t)}(l)\mathbf{x}(l)^T\}. \end{aligned} \tag{3.65}$$

Briefly, $p_{\mathcal{I}_{\mathbf{W}^{(t)}}}$ is substituted with convex combination coefficients $\lambda^{(t)}$. In other words, $\mathbf{W}_{subg}^{(t)}$ term is calculated by weighted averaging of all samples.

Deriving the exact formula of $p_{\mathcal{I}_{\mathbf{W}}}(k)$ for every different problem is not trivial, even for Gaussian noise scenario. To overcome this difficulty, the pmf definition can be simplified. As the first step, an estimate of $\mathcal{I}_{\mathbf{W}}$ is defined as

$$\hat{\mathcal{I}}_{\mathbf{W}} = \{n : \|\mathbf{y}(n)\|_1 \geq \beta \times \max_{k\in\{1,\ldots,L\}} \|\mathbf{y}(k)\|_1, \tag{3.66}$$

where $0 < \beta \leq 1$ is a parameter to be specified. This corresponds to choosing index points, for which the output ℓ_1-norm value is at some neighbourhood of the maximum output ℓ_1-norm. As a result, the approximated pmf is expressed as

$$\hat{p}_{\mathcal{I}_{\mathbf{W}}}(k) = \begin{cases} \frac{1}{|\hat{\mathcal{I}}_{\mathbf{W}}|} & k \in \hat{\mathcal{I}}_{\mathbf{W}}, \\ 0 & \text{otherwise.} \end{cases} . \tag{3.67}$$

As an alternative, we could potentially compute the probability value $\hat{p}_{\mathcal{I}_{\mathbf{W}}}(k)$ in proportion to the value of $\|\mathbf{y}(k)\|_1$. On the other hand, the approximation in (3.67) yields satisfactory performance as illustrated in Section 3.2.8.

3.2.7 Algorithm Acceleration

We offer two major improvements for algorithm acceleration.

- **Weighted Update**:

  BSS algorithm developers attended natural gradient based learning methods due to its improved convergence behaviour [Amari, 1998]. The update rule of SBCA can be changed through this approach by weighting with positive matrix $\mathbf{W}^T\mathbf{W}$ which yields

$$\begin{aligned}\mathbf{W}_{logdet}\mathbf{W}^T\mathbf{W} &= \left(\left(\mathbf{W}\hat{\mathbf{R}}_{\mathbf{x}}\mathbf{W}^T\right)^{-1}\mathbf{W}\hat{\mathbf{R}}_{\mathbf{x}}\right)\mathbf{W}^T\mathbf{W} \\ &= \mathbf{W}, \end{aligned} \tag{3.68}$$

  and

$$\begin{aligned}\mathbf{W}_{subg}\mathbf{W}^T\mathbf{W} &= \frac{p}{\max\limits_{n\in\{1,\ldots,L\}}\|\mathbf{y}(n)\|_1}\text{sign}\{\mathbf{y}(l)\}\mathbf{x}(l)^T\mathbf{W}^T\mathbf{W}, \\ &= \frac{p}{\max\limits_{n\in\{1,\ldots,L\}}\|\mathbf{y}(n)\|_1}\text{sign}\{\mathbf{y}(l)\}\mathbf{y}(l)^T\mathbf{W}. \end{aligned} \tag{3.69}$$

  As a consequence, the equation in (3.41) turns into

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} + \mu^{(t)}\left(\mathbf{I} - \frac{p}{\max\limits_{n\in\{1,\ldots,L\}}\|\mathbf{y}^{(t)}(n)\|_1}\text{sign}\{\mathbf{y}^{(t)}(l^{(t)})\}\mathbf{y}^{(t)}(l^{(t)})^T\right)\mathbf{W}^{(t)}. \tag{3.70}$$

  It should be emphasized that (3.70) is analogous with the update formula of Infomax algorithm for super-Gaussian sources

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} + \mu^{(t)}(\mathbf{I} - 2\tanh\{\mathbf{y}^{(t)}(l^{(t)})\}\mathbf{y}^{(t)}(l^{(t)})^T)\mathbf{W}^{(t)}, \tag{3.71}$$

  with the difference that Infomax uses all sample points while SBCA selects some of them with respect to ℓ_1 norm value.

- **Nesterov Update Rule**:

  Nesterov update rule is a widely known method to shorten convergence time

[Nesterov, 2013] and was firstly proposed for smooth and strongly convex functions. Here, we attempt to use it for a non-smooth and non-convex function. Our empirical trials to improve the convergence time of SBCA updates demonstrated the usability of Nesterov update rule for our problem.

Let us express the original SBCA update rule here again:

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} + \mu^{(t)}\mathbf{U}^{(t)}, \tag{3.72}$$

In this equation, it is possible to replace $\mathbf{U}^{(t)}$ with any sub-gradient term derived previously. By incorporating Nesterov update, the SBCA equation is reformulated as

$$\mathbf{u}^{(t+1)} = \mathbf{W}^{(t)} + \upsilon^{(t)}\mathbf{U}^{(t)}, \tag{3.73}$$

$$\mathbf{W}^{(t+1)} = \mathbf{u}^{(t+1)} + \kappa^{(t)}(\mathbf{u}^{(t+1)} - \mathbf{u}^{(t)}), \tag{3.74}$$

where $\mathbf{u}^{(t)}$ is an intermediate algorithm variable, $\kappa^{(t)} = \frac{t-1}{t+2}$ and $\upsilon^{(t)}$ is the algorithm parameter to be selected. Note that the update rule essentially makes use of previous update expressions as seen in (3.73).

3.2.8 Numerical Results

For numerical analysis, we use Signal to Interference and Noise power Ratio (SINR) measure as a performance criterion. SINR is computed with

$$SINR = \frac{S}{I+N} \tag{3.75}$$

where $S = \| \begin{bmatrix} \mathbf{F}(m_1,1), & \cdots & ,\mathbf{F}(m_p,p) \end{bmatrix} \|_2^2$ represents the separator output signal power, $I = \left\| \begin{bmatrix} \mathbf{F}(m_1,:)^T, & \cdots & ,\mathbf{F}(m_p,:)^T \end{bmatrix}^T \right\|_F^2 - \left\| \begin{bmatrix} \mathbf{F}(m_1,1), & \cdots & ,\mathbf{F}(m_p,p) \end{bmatrix} \right\|_2^2$ represents the interference power, $N = \sigma_{noise}^2 \|\mathbf{W}\|_F^2$ represents the noise power, and the indices $\{m_1, \cdots, m_p\}$ represent indices of the maximums in the columns of the overall transfer matrix $\mathbf{F}$.

We choose, as benchmark algorithms, Extended Infomax [Lee et al., 1999a], Relative Newton [Zibulevsky and Pearlmutter, 2001], Generalized Morphological Component Analysis (GMCA) [Bobin et al., 2007], Bounded Component Analysis (BCA) [Cruces, 2010],[Erdogan, 2013] and non-negative Least-Correlated Component Analysis (nLCA) [Wang et al., 2010]. In addition to their success, another reason for choosing these methods is that the benchmark codes of these algorithms are available to all researchers (the code links are given in the related references).

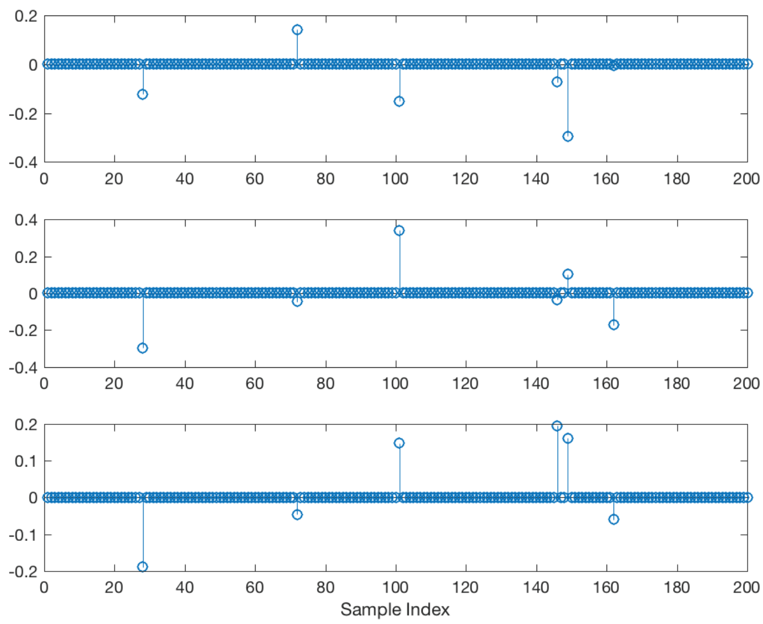


Figure 3.2: Synthetically generated sparse sequences in Experiment 1.

Synthetic Sparse Signals:

In this experiment, sparse signals are generated from i.i.d. uniform random vectors $\mathbf{u} \in [-1, 1]^p$ via the mapping

$$\mathbf{s} = \begin{cases} \mathbf{u} & \mathbf{u} \in \mathcal{U}_r \\ \mathbf{0} & \text{otherwise}, \end{cases} \tag{3.76}$$

where $\mathcal{U}_r = \{\mathbf{z} : \|\mathbf{z}\|_r \leq 1\}$ with $0 < r \leq 1$. For the analyses, the following setup is

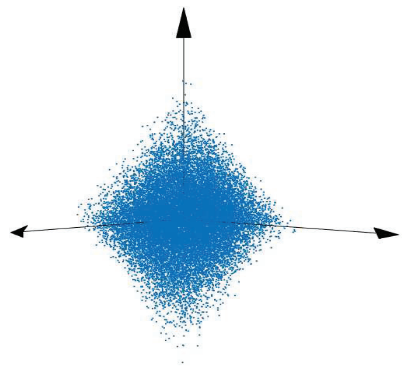


Figure 3.3: Scatter diagram for synthetically generated sparse sequences in Experiment 1.

used:

- Mapping coefficient, $r = 0.8$.
- Additive White Gaussian Noise (AWGN) with Input SNR=20 dB.
- Mixing matrix $\mathbf{A}$ with zero mean and unit variance Gaussian distribution.

Three of the discrete source sequences and their corresponding scatter diagram are depicted in Fig. 3.2 and Fig. 3.3. For comparison, output SINR values of each algorithm are analysed. Fig. 3.4 illustrates output SINR performance versus data length for 4 sources and 8 mixtures. As illustrated in this figure, SBCA provides $3-5$dB improvement over the compared algorithms.

To analyse the performance for dependent sources, we generate dependent uniform signals using Copula distribution with correlation constant $\rho = 0.7$, and apply again the rule in (3.76) to sparsify them. As seen in Fig. 3.5, source dependency seems to cause performance degradation of Extended Infomax and Relative Newton, particularly for large data sizes, while it does not effect SBCA's performance

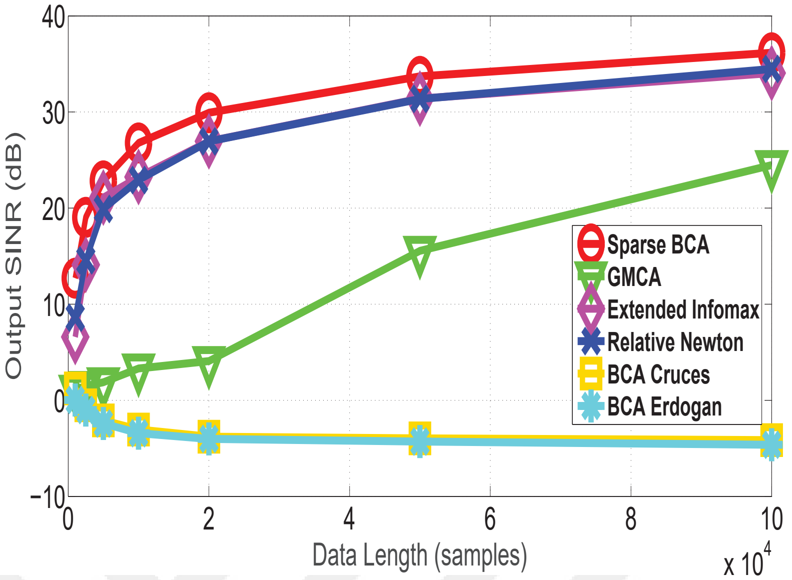


Figure 3.4: Output SINR versus data length for independent random variables in Experiment 1.

significantly.

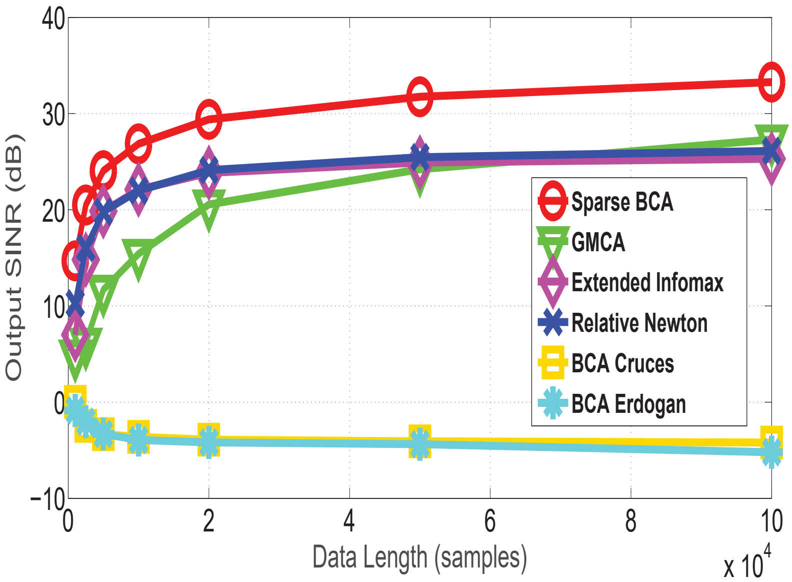


Figure 3.5: Output SINR vs data length for correlated sources in Experiment 1.

In order to evaluate complex extension of SBCA, complex sources were generated by multiplying the Copula signals in the previous experiment with i.i.d.(uniformly distributed phase terms in the interval $[-\pi, \pi]$) unit magnitude complex signals. Fig. 3.6 demonstrates the satisfactory performance of our algorithm for this experiment.

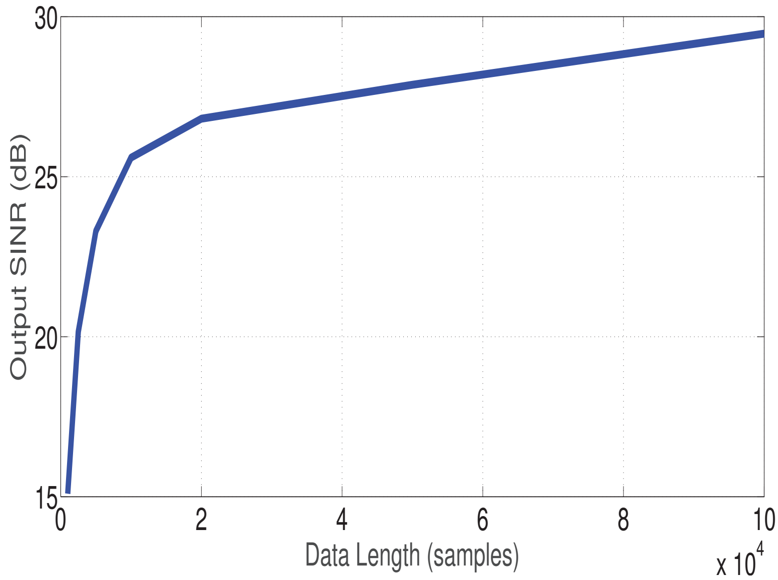


Figure 3.6: Performance of complex SBCA for correlated random variables.

The results given in this section show that SBCA provides BSS with improvement even in the case of statistically independent sources. Especially, the short data length performance of SBCA is remarkable since it defines the objective functions directly in terms of mixture samples while others exploit a stochastic structure.

Natural Images:

The next source separation example is handled with images. Image separation is a good application area to see how SCA performs in practice.

A well known fact about images is that high frequency components (details) of common images are sparse. It means that images can be modelled as a superposition of sparse components in a transformed domain. Sparsification can be accomplished by using a suitable linear transformation. There are many choices for the sparsification transform, such as Discrete Cosine Transform, Discrete Fourier Transform, Wavelet Transform (Haar, Daubechies, Fractional Spline, Gabor, etc.), Contourlets, Shearlets, etc. A sparsification transform is said to provide complete representation for an image if the image could be represented as the superposition of the corresponding transformation's basis functions. Some transforms (for ex. Haar) have orthonormal bases and images can be written as the linear superposition of their transform coefficients. Orthogonal transforms are a practical and easy way for sparse representation. On other hand, there are many non-orthogonal and redundant transforms which claim the superiority over each other and non-orthogonal transforms.

Recently, Gabor wavelets have been promoted as a multi-orientation and multi-resolution wavelet transform. Gabor wavelets is said to provide optimal resolution in both time (or spatial) and frequency domains among all wavelet transforms. The basic expression of the 2-D Gabor wavelet function in time (or spatial) domain is

$$\psi(a,b) = \frac{f^2}{\pi\gamma\eta}\exp\left(-\left(\frac{f^2}{\gamma^2}a_r^2 + \frac{f^2}{\gamma^2}b_r^2\right)\right)\exp(j2\pi f a_r) \tag{3.77}$$

$$a_r = a\cos\theta + b\sin\theta,\ b_r = -a\sin\theta + b\cos\theta \tag{3.78}$$

where f is the modulating sinus wave frequency, and θ is the orientation parameter. 2-D Gabor wavelet transform definition in (3.77) is generalized as

$$\psi_\theta(x_a, x_b, a, b, a_0, b_0) = \frac{1}{\sqrt{x_a x_b}}\psi_\theta\left(\frac{a-a_0}{x_a} + \frac{b-b_0}{x_b}\right) \tag{3.79}$$

where $\psi_\theta(a,b)$ is the 2-D mother wavelet, x_a and x_b are the scaling parameters, a_0 and b_0 are the spatial shifts. The scaling and orientation parameters are important for the sparsification performance of Gabor wavelets.

In this experiment, we use Morlet wavelet transform which is a modified version of Gabor wavelet transform [Morlet and Grossman, 1984].

Three 204×204 images are mixed in Gaussian channels. There are 3 mixing channels for signal observation. Fig. 3.7 illustrates the natural and mixed images. Since the natural images are not sparse, Morlet wavelet transform is applied to the mixture image patches (8×8) as a preprocessing step for sparsification. Fig.

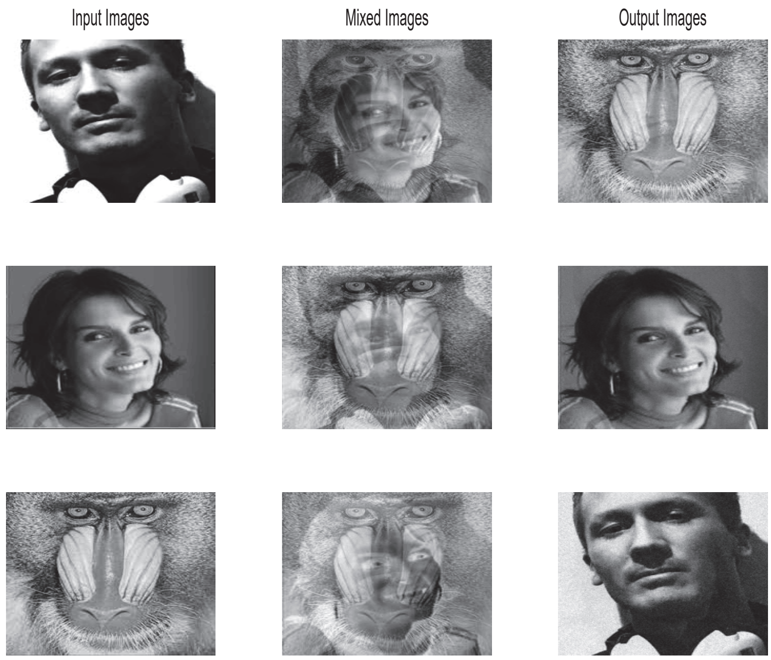


Figure 3.7: Original, mixed and output images for image separation example.

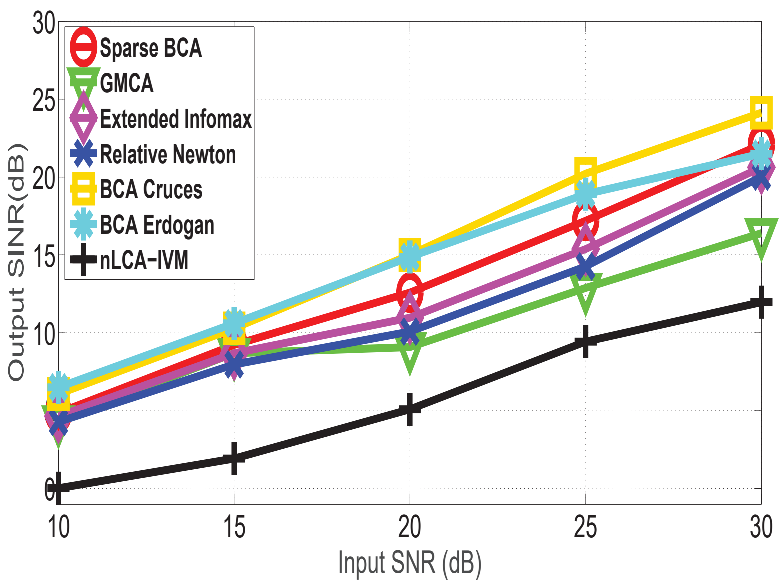


Figure 3.8: Output SINR vs Input SNR for image separation example.

3.8 illustrates mean output SINR (averaged over random mixture matrices) versus input SNR performance. The graphs in Fig. 3.8 reveal that SBCA outperforms the benchmark methods for every input SNR value. SBCA is not better than non-sparse methods (Erdogan's BCA and Cruces' BCA) because they directly process natural images and do not need sparsification process step that might cause information degradation. Also, Cruces' BCA extracts only one source signal at a time while the others extract all simultaneously. Thus, mean SINR comparison in Fig. 3.8 is not fair for the other methods and we provide maximum output SINR graphs in Fig. 3.9 that demonstrates the superiority of the other methods over Cruces' BCA. We also note that nLCA-IVM assumes non-negative signals and adding Gaussian noise to the mixtures might cause negative values that corrupt this non-negativity assumption. This may be the reason why nLCA-IVM did not show a satisfactory performance as illustrated in Fig. 3.8.

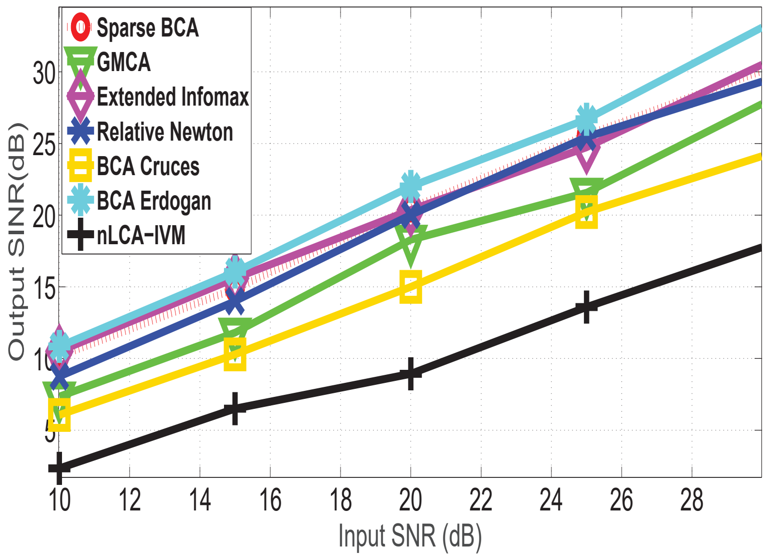


Figure 3.9: Maximum Output SINR vs Input SNR for image separation example.

Neuroimaging:

In this experiment, we deal with neuron sorting from a CA+2 imaging records from which neuron components and their temporal activities are identified. For the analysis, we used the video recording given by Yaniv Ziv et al in the annex of their article [Ziv et al., 2013]. The recording contains hippocampal cell activities of the freely moving mouse (This movie file is available in `https://www.nature.com/neuro/journal/v16/n3/full/nn.3329.html`.)

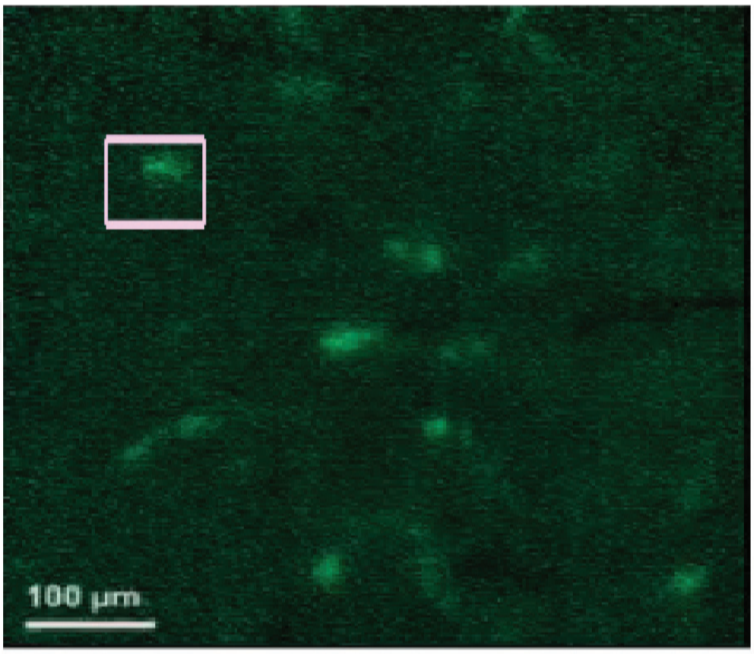


Figure 3.10: A frame of the Ca^{2+} movie of [Ziv et al., 2013].

As the first step, we form 32×32 subframes from movie frames. Fig. 3.10 illustrates a frame and a subframe inside this frame. The subframes are vectorized to obtain observation vectors that are the input to SBCA algorithm. The processed recording includes 1339 frames. SBCA is applied to each subframe consecutively. Fig. 3.11.(a) illustrates a detected neuron component's image. To clarify the image more, a denoising process based on Symlet wavelet is applied. Fig. 3.11.(b) illustrates the denoised version of the component in Fig. 3.11.(a). Fig. 3.12 shows some denoised neuron components and related temporal activities. We find out that

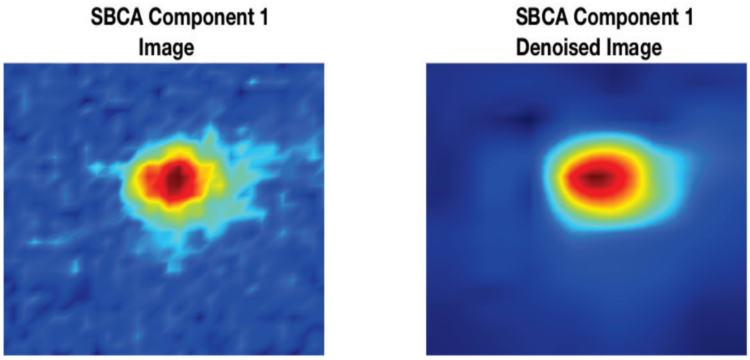


Figure 3.11: A component map output of SBCA algorithm applied to the sub-window marked in Fig. 3.10, and wavelet denoised version of the component map.

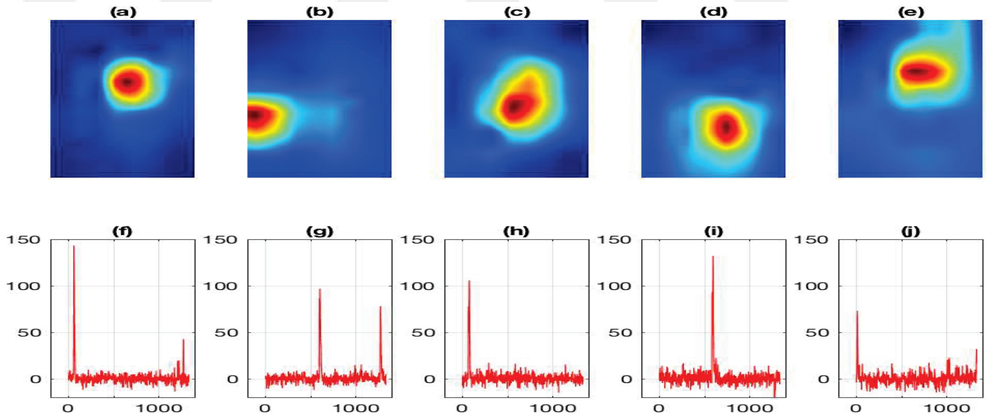


Figure 3.12: Some (denoised) component maps extracted by SBCA corresponding to the subframe in Fig. 3.10 (in the first row (a)-(e)), and their activations (in the second row (f)-(i))

the activity structures confirm the sparse nature of neuron activity. This analysis approach will be further developed and a detailed neural activity detection method for full frame analysis will be proposed in chapter 5.

Chapter 4

CONVOLUTIVE BOUNDED COMPONENT ANALYSIS BASED ON SPARSITY

4.1. Convolutive Blind Source Separation Methods

BSS algorithms consider different mixing models to extract source signals. The simplest mixing model is called instantaneous mixing model which is a sum of weighted source samples. However, in practice, we face more challenging BSS problems where the source signals are mixed in time or in space. These problems can be described by convolutive mixing model.

Convolutive BSS is a generic inverse problem for which the applications span a wide range including acoustic signal separation [Pedersen et al., 2007, Douglas et al., 2007, Douglas et al., 2005a, Makino et al., 2005, Parra and Spence, 2000], digital communications [Oktem et al., 2010, Randel et al., 2011], electrocardiogram measurements [Vayá et al., 2007], radar signal processing [Nion and Sidiropoulos, 2009], etc. While there are numerous different applications, the convolutive BSS approaches used in these applications can be organized in two main groups in accordance with processing domain. *Time domain methods* constitute the first group where the solution is based on the convolution with Multiple Input Multiple Output (MIMO) FIR filter [Douglas et al., 2005b, Douglas et al., 2007, Koldovský and Tichavský, 2007, Thomas et al., 2006, Aichner et al., 2003]. The main difficulty with these methods is that the mixing channel equalization with adequate performance may be possible only with very long FIR separator filters. In particular, audio signal separation in a highly reverberant environment [Makino et al., 2007] would create such a difficult case. High filter order also means many filter coefficients to be learned for these methods. *Frequency domain methods* constitute the second group where temporal mixtures are mapped to spectral data so that convolution in time domain is converted to multiplication in frequency domain [Smaragdis, 1998]. Hereby, it enables spectral data in each frequency bin to be processed by using instantaneous BSS algorithms. Frequency transformation is executed using different methods such as sliding window DFT [Attias, 2003] and STFT [Yilmaz and Rickard, 2004]. The bottlenecks of the frequency domain methods are scaling and permutation inconsistencies that cause unequally scaled spectral components in different bins and spectral mixing of source components. This issue complicates the inverse frequency transform of the extracted source components. On the other hand, the frequency domain methods take benefit of improved efficiency and better convergence.

Among various different assumptions on sources and mixing systems to solve BSS problem, the mutual source independence is one of the most powerful and useful assumptions that was firstly proposed for the instantaneous mixtures [Comon, 1994]. In order to extend this powerful setting to convolutive mixtures, Douglas [Douglas et al., 2007] proposed a spatio-temporal extension of Fast-ICA [Hyvärinen and Oja, 1997]. Similarly, Koldovský converted the instantaneous ICA algorithm called EFICA [Koldovský et al., 2006] to a time-domain convolutive algorithm in [Koldovský and Tichavský, 2007]. Thomas et al [Thomas et al., 2006] extended the Fast-ICA algorithm [Hyvärinen and Oja, 1997] to convolutive mixing case where the source signals are moving average processes. Except statistical independence, there are many other source signal characteristics to be exploited by the convolutive BSS approaches. For example, non-stationarity of sources is a second order statistic and a very common characteristic for speech signals [Parra and Spence, 2000, Aichner et al., 2003, Souloumiac, 1995, Rahbar and Reilly, 2005]. In another category, temporal correlation of signals is utilized to degrade convolutive mixing problem to instantaneous mixing problem in order to solve it by using additional features of the signals [Gorokhov and Loubaton, 1996, Mansour et al., 2000].

Sparsity has been recently a useful signal characteristic to solve convolutive mixing problem, particularly because of the ability of SCA methods to handle under-determined mixing case [Blin et al., 2005, Rickard et al., 2005]. Beside being the main assumption, sparsity can also be used as a post processing step after the main BSS step. For example, in [Kolossa and Orglmeister, 2004, Pedersen et al., 2005], Time-Frequency Masking is used to obtain a higher SIR value via magnitude comparison of the BSS outputs. Sparsity is not an inherent property of data, it is related with the representation domain of data. Although many source signals are not sparse in time domain, their sparse counterparts could be generated by transforming them into frequency domain or time-frequency domain. That's why frequency based SCA methods have been attended more in the last years [He et al., 2007, Bofill, 2003].

In chapter 3, a new instantaneous BCA approach called as Sparse BCA (SBCA) [Babatas and Erdogan, 2016, Babatas and Erdogan, 2018a] was introduced for sparsely natured bounded signals. SBCA uses a geometrical signal setup such that the sources are bounded in an ℓ_1-norm ball. Many researchers used similar geometrical constraints as side or main assumption in their optimization based separation frameworks [Craig, 1994, Chan et al., 2009]. Here, we introduce the extension of the instantaneous SBCA to more general convolutive mixtures which is called Convolutive SBCA (CSBCA) [Babatas and Erdogan, 2018b, Babatas and Erdogan,].

The main contributions of this study are,

- An algorithm that uses a time based objective function to separate sources from convolutive mixtures and its extension to complex signal case.

- A modified version of the objective function and iterative algorithm for stationary sources which combines time domain separation criteria and frequency domain computations for a faster convergence than the completely time-based computations.

- A new algorithm that uses a frequency based objective function in order to sparsify source signals by mapping them to frequency spectrum data. We note that the algorithm maximizes the objective function in accordance with a temporal separator matrix.

- A practical application on audio signal separation in a reverberant scene.

The chapter is organized as follows: In section 4.2, time domain convolutive BSS setup is defined and then the time-domain-sparsity based convolutive SBCA approach is introduced. In section 4.3, the time-frequency domain version of the temporal objective function for stationary sources is formulated. In section 4.4, we map convolutive mixtures to frequency spectrum data via STFT and describe a new signal setup in frequency domain. Then, we propose a frequency-domain-sparsity based SBCA optimization setting. Finally, there are numerical examples in Section 4.5, also including speech and musical signal separation applications.

We define some notations for $\mathbf{U} \in \mathbb{R}^{p \times p}$, $\mathbf{u} \in \mathbb{R}^p$, $a, b, k \geq 1 \in \mathbb{Z}$ in Table 4.1:

Notation	**Explanation**
$\|\mathbf{u}\|_a$	ℓ_a-norm formulated as $\sqrt[a]{\sum_{n=1}^{p} \|u_n\|^a}$
$\|\mathbf{U}\|_{a,b}$	Induced matrix norm formulated as $\sup_{\|\mathbf{x}\|_b \leq 1} \|\mathbf{U}\mathbf{x}\|_a$
$\tilde{\mathbf{U}}$	convolutive channel response defined as $\tilde{\mathbf{U}} = [\mathbf{U}(0),$ $\mathbf{U}(1), \ldots, \mathbf{U}(K-1)]$ where K is the channel order
$\Gamma_N(\mathbf{U})$	Block Toeplitz matrix of $\mathbf{U}$
$Co\{\mathbf{U}\}$	Convex Hull of $\mathbf{U}$
sign$\{\mathbf{U}\}$	extracts the sign of matrix entries
$\mathcal{R}e\{\mathbf{U}\}$	real part of $\mathbf{U}$
$\mathcal{I}m\{\mathbf{U}\}$	imaginary part of $\mathbf{U}$
signc$\{\mathbf{U}\}$	complex sign operator: $\mathcal{R}e\{\text{sign}\{\mathbf{U}\}\} + i\mathcal{I}m\{\text{sign}\{\mathbf{U}\}\}$
$\Upsilon(\mathbf{u})$	transform operator, $\mathbb{C}^p \rightarrow \mathbb{R}^{2p}$
$\mathbf{u}(m, f)$	Short Time Fourier Transform of $\mathbf{u}(n)$
$\mathbf{P}_{\mathbf{u}}(f)$	Power Spectral Density Matrix of $\mathbf{u}(f)$
$\mathbf{P}_{\mathbf{u}}(m, f)$	Power Spectral Density Matrix of $\mathbf{u}(m, f)$

Table 4.1: Notation Table.

4.2. Time Domain Sparsity Based Convolutive SBCA Framework

4.2.1 Convolutive Blind Source Separation Setup in Time Domain

Before introducing the new method, we should define the considered convolutive BSS setup. In this setup,

- The set $\mathcal{S} = \{\mathbf{s}(n) \in \mathbb{R}^p\}$ represents the samples of p sources with the assumption that they are bounded and lie in an ℓ_1-norm-ball given by

$$\mathcal{B}_s = \{\mathbf{s} \in \mathbb{R}^p | \|\mathbf{s}\|_1 < 1\}, \tag{4.1}$$

  Here, p is a priori information and an input to the proposed algorithm. If we like to generalize the ℓ_1-norm-ball description for sources with different ranges, we can replace it with a weighted ℓ_1-norm-ball. However, we will go on with (4.1) to simplify expressions without any loss of generality.

- The MIMO convolutive mixing channel output is described with

$$\mathbf{x}(n) = \sum_{k=0}^{K-1} \mathbf{A}(k)\mathbf{s}(n-k), \tag{4.2}$$

  where $\{\mathbf{A}(k) \in \mathbb{R}^{q \times p}, k \in \{0, \ldots, K-1\}\}$ denotes the impulse response of the equalizable mixing channel [Inouye and Liu, 2002] with order $K-1$. Equalizable channels allow transmission of zeros at $z = 0$ in the z-plane that enables source signals to arrive at the sensors with different delays. Moreover, a necessary and sufficient condition for a channel $\mathbf{A} \in \mathbb{R}^{q \times p}$ to be equalizable is that $rank(\mathbf{A}) = p$ for $q \geq p$.

- The output of the FIR separator filter is defined as

$$\mathbf{y}(n) = \sum_{m=0}^{M-1} \mathbf{W}(m)\mathbf{x}(n-m), \tag{4.3}$$

  where $\{\mathbf{W}(m) \in \mathbb{R}^{p \times q}, m \in \{0, \ldots, M-1\}\}$ denotes the separator impulse response with order $M-1$.

- Combining (4.2) and (4.3) yields the total system response

$$\mathbf{F}(k) = \sum_{m=0}^{M-1} \mathbf{W}(m)\mathbf{A}(k-m), \quad k = 0, \ldots, P-1, \tag{4.4}$$

where $P-1 = K+M-2$.

- The mixtures and the separator outputs can be redefined by replacing the convolutive mixing channel response matrix with $\tilde{\mathbf{A}} = [\mathbf{A}(0), \mathbf{A}(1), \ldots, \mathbf{A}(K-1)]$, separator filter response matrix with $\tilde{\mathbf{W}} = [\mathbf{W}(0), \mathbf{W}(1), \ldots, \mathbf{W}(M-1)]$ and system response matrix with $\tilde{\mathbf{F}} = [\mathbf{F}(0), \mathbf{F}(1), \ldots, \mathbf{F}(P-1)]$:

$$\mathbf{x}(n) = \tilde{\mathbf{A}}\tilde{\mathbf{s}}_K(n), \quad n = 1, \ldots L, \tag{4.5}$$

$$\mathbf{y}(n) = \tilde{\mathbf{W}}\tilde{\mathbf{x}}_M(n), \quad n = 1, \ldots, L+M-1, \tag{4.6}$$

$$\mathbf{y}(n) = \tilde{\mathbf{F}}\tilde{\mathbf{s}}_P(n), \quad n = 1, \ldots, L+M-1. \tag{4.7}$$

In these equations, $\tilde{\mathbf{s}}_K(n) = [\mathbf{s}^T(n), \mathbf{s}^T(n-1), \ldots, \mathbf{s}^T(n-K+1)]^T$, $\tilde{\mathbf{x}}_M(n) = [\mathbf{x}^T(n), \mathbf{x}^T(n-1), \ldots, \mathbf{x}^T(n-M+1)]^T$ and $\tilde{\mathbf{s}}_P(n) = [\mathbf{s}^T(n), \mathbf{s}^T(n-1), \ldots, \mathbf{s}^T(n-P+1)]^T$.

- Before giving the objective function formula, two more terms should be described. The first one is the extended separator output vector which is formulated as $\tilde{\mathbf{y}}_N(n) = [\mathbf{y}^T(n), \mathbf{y}^T(n-1), \ldots, \mathbf{y}^T(n-N+1)]^T$. The second term is the block-Toeplitz matrix $\Gamma_N(\tilde{\mathbf{F}})$ given by the equation

$$\Gamma_N(\tilde{\mathbf{F}}) = \begin{bmatrix} \mathbf{F}(0), \mathbf{F}(1), \ldots, \mathbf{F}(P-1), \ldots, \mathbf{0} \\ \vdots \qquad \ddots \qquad \ddots \qquad \vdots \\ \mathbf{0}, \ldots, \mathbf{F}(0), \mathbf{F}(1), \ldots, \mathbf{F}(P-1) \end{bmatrix}, \tag{4.8}$$

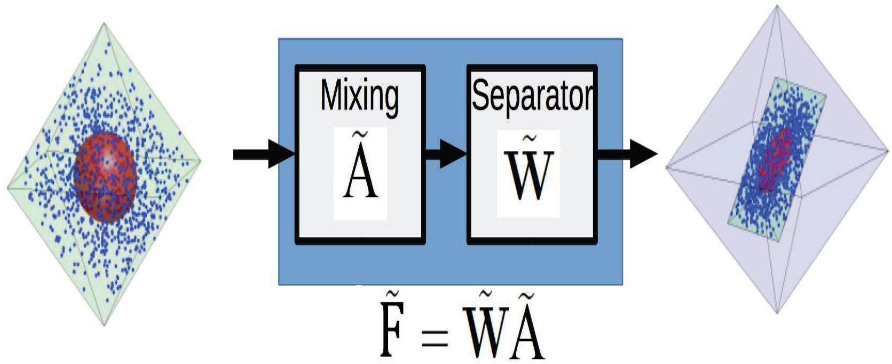


Figure 4.1: Geometric objects used in the convolutive SBCA setup.

where $N \geq P$. This yields the following equation for the extended separator output vector:

$$\tilde{\mathbf{y}}_N(n) = \Gamma_N(\tilde{\mathbf{F}})\tilde{\mathbf{s}}_{N+P-1}(n). \tag{4.9}$$

Defining the source sample set $S_{N+P-1} = \{\tilde{\mathbf{s}}_{N+P-1}(M+N-1), \tilde{\mathbf{s}}_{N+P-1}(M+N), \ldots, \tilde{\mathbf{s}}_{N+P-1}(L)\}$, the following assumption is provided as the basis of our geometrical optimization setting:

Assumption (A4): The vertices of the bounding ℓ_1-norm-ball $\mathcal{B}_s$ are contained in S_{N+P-1}.

4.2.2 Objective Function

As in chapter 3, we formulate the objective function as a volume ratio described with respect to the samples $\{\tilde{\mathbf{y}}_N(n), n \in \{N, \ldots, L_1\}\}$ where $L_1 = L + M - 1$.

- *Bounding ℓ_1-Norm Ball*: It encloses all separator output samples, and its volume is formulated using the maximum ℓ_1 norms of the extended separator outputs, i.e., $\mathcal{B}_{\mathbf{Y}} = \{\mathbf{q} | \|\mathbf{q}\|_1 \leq \max_{n \in \{N, \ldots, L_1\}} \|\tilde{\mathbf{y}}_N(n)\|_1\}$.

- *Principal Hyper-ellipsoid* : Its volume is formulated using the covariance of the extended separator outputs, i.e., $\mathcal{E}_{\mathbf{Y}} = \{\mathbf{q} | (\mathbf{q} - \hat{\mu}_{\tilde{\mathbf{y}}_N})^T \hat{\mathbf{R}}_{\tilde{\mathbf{y}}_N}^{-1} (\mathbf{q} - \hat{\mu}_{\tilde{\mathbf{y}}_N}) \leq 1\}$

where $\hat{\mu}_{\tilde{\mathbf{y}}_N} = \frac{1}{L_2}\sum_{n=N}^{L_1}\tilde{\mathbf{y}}_N(n)$, $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_N} = \frac{1}{L_2}\sum_{n=N}^{L_1}(\tilde{\mathbf{y}}_N(n)-\hat{\mu}_{\tilde{\mathbf{y}}_N})(\tilde{\mathbf{y}}_N(n)-\hat{\mu}_{\tilde{\mathbf{y}}_N})^T$ and $L_2 = L_1 - N + 1$.

Fig.4.1 illustrates the geometrical objects in the proposed convolutive framework.

- On the left side,
  - diamond shaped box is the sources' bounding ℓ_1-norm-ball given in (4.1),
  - red hyper-ellipsoid is the sources' principal hyper-ellipsoid formulated as $\mathcal{E}_{\tilde{\mathbf{s}}_{L_3}} = \{\mathbf{q} | (\mathbf{q}-\hat{\mu}_{\tilde{\mathbf{s}}_{L_3}})^T\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}^{-1}(\mathbf{q}-\hat{\mu}_{\tilde{\mathbf{s}}_{L_3}}) \leq 1\}$ where $\hat{\mu}_{\tilde{\mathbf{s}}_{L_3}} = \frac{1}{L_2}\sum_{n=N}^{L_1}\tilde{\mathbf{s}}_{L_3}(n)$, $\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}} = \frac{1}{L_2}\sum_{n=N}^{L_1}(\tilde{\mathbf{s}}_{L_3}(n)-\hat{\mu}_{\tilde{\mathbf{s}}_{L_3}})(\tilde{\mathbf{s}}_{L_3}(n)^T-\hat{\mu}_{\tilde{\mathbf{s}}_{L_3}})^T$, and $L_3 = N + P - 1$.
- On the right side,
  - diamond shaped box is the output bounding ℓ_1-norm-ball defined above,
  - red hyper-ellipsoid is the output principal hyper-ellipsoid defined above,
  - green polytope is the image of $\mathcal{B}_{\mathbf{s}}$ under the mapping $\tilde{\mathbf{F}}$.

Based on these geometric objects, the CSBCA objective is defined as a volume ratio

$$J(\tilde{\mathbf{W}}) = \frac{\sqrt{\det(\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_N})}}{\left(\max_{n\in\{N,\ldots,L_1\}}\|\tilde{\mathbf{y}}_N(n)\|_1\right)^{Np}}, \tag{4.10}$$

to be maximized for source separation.

The following theorem ensures that perfect source separation is achievable via the maximization of (4.10).

Theorem-3: Provided that the assumption (**A4**) is correct and the mixing channel is equalizable by an FIR separator filter, all global maximums of (4.10) provide perfect separation.

The proof of **Theorem-3** is provided in Appendix A.

4.2.3 Iterative Algorithm for Separator Estimation

The logarithm transform of (4.10) would be more convenient for the algorithm derivation:

$$\begin{aligned}\mathcal{J}(\tilde{\mathbf{W}}) &= \log(J(\tilde{\mathbf{W}})) \\ &= \underbrace{\frac{1}{2}\log\left(\det(\Gamma_N(\tilde{\mathbf{W}})\hat{\mathbf{R}}_{\tilde{\mathbf{x}}_{\mathbf{N+M-1}}}\Gamma_N(\tilde{\mathbf{W}})^T)\right)}_{\mathcal{J}_1(\tilde{\mathbf{W}})} - Np\log(\underbrace{\max_{n\in\{N,\ldots,L_1\}}\|\tilde{\mathbf{y}}_N(n)\|_1)}_{\mathcal{J}_2(\tilde{\mathbf{W}})}\end{aligned} \tag{4.11}$$

In this function, $\mathcal{J}_1(\tilde{\mathbf{W}})$ is convex differentiable, and $\mathcal{J}_2(\tilde{\mathbf{W}})$ is convex non-smooth.

Talking about $\mathcal{J}_2(\tilde{\mathbf{W}})$, it is possible to use Clarke sub-differentials method [Bagirov et al., 2014] for computing the derivative. If $f_n(\Gamma_N(\tilde{\mathbf{W}})) = \|\tilde{\mathbf{y}}_N(n)\|_1$ denotes the output ℓ_1 norm ball, the sub-differential set of $f_n(\Gamma_N(\tilde{\mathbf{W}}))$ with respect to $\Gamma_N(\tilde{\mathbf{W}})$ is written as

$$\begin{aligned}\partial f_n(\Gamma_N(\tilde{\mathbf{W}})) &= \mathbf{o} \\ &= \mathbf{r}\tilde{\mathbf{x}}_{N+M-1}(l) : r_i = \operatorname{sign}\{(\tilde{\mathbf{y}}_N)_i(l)\} + 1_{(\tilde{\mathbf{y}}_N)_i(l)=0}\alpha_i,\end{aligned} \tag{4.12}$$

where $\alpha_i \in [-1, 1]$.

(4.12) is transformed into a new derivative equation with respect to $\tilde{\mathbf{W}}$ because our goal is to obtain an iterative update equation for $\tilde{\mathbf{W}}$:

$$\acute{\mathbf{o}} = \sum_{k=0}^{N-1} \mathbf{o}_{kp+1:(k+1)p,kq+1:(k+M)q} \tag{4.13}$$

The chain rule is used to obtain the update term for $\mathcal{J}_2(\tilde{\mathbf{W}})$ as

$$\partial\tilde{\mathbf{W}}_{subg} = Np\frac{\sum\limits_{l\in\mathcal{I}_{\tilde{\mathbf{W}}}}\sum\limits_{k=0}^{N-1}\lambda_l\mathbf{o}_{kp+1:(k+1)p,kq+1:(k+M)q}}{\max\limits_{n\in\{N,\ldots,L_1\}}\|\tilde{\mathbf{y}}_N(n)\|_1} \tag{4.14}$$

where $\lambda_l \geq 0$, $\sum\limits_{l\in\mathcal{I}_{\tilde{\mathbf{W}}}} \lambda_l = 1$. $\mathcal{I}_{\tilde{\mathbf{W}}}$ is the index subset of maximum ℓ_1-norm at the separator output.

As in (4.13), the gradient of $\mathcal{J}_1(\tilde{\mathbf{W}})$ with respect to $\tilde{\mathbf{W}}$ is written as a summation term

$$\partial\mathcal{J}_1(\tilde{\mathbf{W}}) = \sum_{k=0}^{N-1} \mathbf{Z}_{kp+1:(k+1)p,kq+1:(k+M)q}, \tag{4.15}$$

where $\mathbf{Z} = \left(\Gamma_N(\tilde{\mathbf{W}})\hat{\mathbf{R}}_{\tilde{\mathbf{x}}_{N+M-1}}\Gamma_N(\tilde{\mathbf{W}})^T\right)^{-1}\Gamma_N(\tilde{\mathbf{W}})\hat{\mathbf{R}}_{\tilde{\mathbf{x}}_{N+M-1}}$,
$\hat{\mathbf{R}}_{\tilde{\mathbf{x}}_{N+M-1}} = \frac{1}{L_2}\sum_{n=N}^{L_1}(\tilde{\mathbf{x}}_{N+M-1}(n) - \hat{\mu}_{\tilde{\mathbf{x}}_{N+M-1}})(\tilde{\mathbf{x}}_{N+M-1}(n) - \hat{\mu}_{\tilde{\mathbf{x}}_{N+M-1}})^T$, and
$\hat{\mu}_{\tilde{\mathbf{x}}_{N+M-1}} = \frac{1}{L_2}\sum_{n=N}^{L_1}\tilde{\mathbf{x}}_{N+M-1}(n)$.

Adding up the update terms corresponding to $\mathcal{J}_1(\tilde{\mathbf{W}})$ and $\mathcal{J}_2(\tilde{\mathbf{W}})$ returns the intended iterative update equation. This equation is simplified more by choosing a random index location l from the index set $\mathcal{I}_{\tilde{\mathbf{W}}}$ at every iteration:

$$\begin{aligned}\tilde{\mathbf{W}}^{(t+1)} &= \tilde{\mathbf{W}}^{(t)} + \\ &\sigma^{(t)}\left(\sum_{k=0}^{N-1}\mathbf{Z}^{(t)}_{kp+1:(k+1)p,kq+1:(k+M)q} - Np\frac{\sum_{k=0}^{N-1}\mathbf{o}^{(t)}_{kp+1:(k+1)p,kq+1:(k+M)q}}{\max_{n\in\{N,\ldots,L_1\}}\|\tilde{\mathbf{y}}_N(n)\|_1}\right)\end{aligned} \tag{4.16}$$

4.2.4 Extension of CSBCA to Complex Case

The convolutive algorithm for complex sources can be developed again by using the isomorphism based approach explained in chapter 3. Some new transformations for the convolutive terms need to be defined before beginning the derivation. As in chapter 3, $\Upsilon : \mathbb{C}^p \to \mathbb{R}^{2p}$ maps a complex vector to a real isomorphic vector:

$$\Upsilon(\mathbf{u}) = \begin{bmatrix}\mathrm{Re}(\mathbf{u}^T) & \mathrm{Im}(\mathbf{u}^T)\end{bmatrix}^T, \tag{4.17}$$

The notation $\acute{\mathbf{u}}$ will be used to refer $\Upsilon(\mathbf{u})$ in the following derivations. However, a mapping operator for the complex convolutive signal vector $\tilde{\mathbf{u}}_K(n)$ is necessary for the convolutive case. This operator is written as

$$\acute{\mathbf{u}}_K(n) = [\mathrm{Re}(\mathbf{u}^T(n)),\ \mathrm{Im}(\mathbf{u}^T(n)),\ \ldots \\ \mathrm{Re}(\mathbf{u}^T(n-K+1)),\ \mathrm{Im}(\mathbf{u}^T(n-K+1))]^T. \tag{4.18}$$

For mapping a complex convolutive matrix to its real isomorphic counterpart, the operator $\mathbb{C}^{q\times Kp} \rightarrow \mathbb{R}^{2q\times 2Kp}$ is defined as

$$\Omega(\tilde{\mathbf{A}}) = \begin{bmatrix} \mathrm{Re}(\mathbf{A}(0)) & -\mathrm{Im}(\mathbf{A}(0)) & \ldots & -\mathrm{Im}(\mathbf{A}(K-1)) \\ \mathrm{Im}(\mathbf{A}(0)) & \mathrm{Re}(\tilde{\mathbf{A}}(0)) & \ldots & \mathrm{Re}(\mathbf{A}(K-1)) \end{bmatrix}. \tag{4.19}$$

As a consequence, the complex convolutive mixture and separator output signals turn out to

$$\acute{\mathbf{x}}(n) = \Omega(\tilde{\mathbf{A}})\acute{\mathbf{s}}_K(n), \tag{4.20}$$

$$\acute{\mathbf{y}}_N(n) = \Omega(\tilde{\mathbf{W}})\acute{\mathbf{x}}_{N+M-1}(n). \tag{4.21}$$

Based on the above definitions, the suggested objective function for the complex extension is proposed to be

$$J_c(\tilde{\mathbf{W}}) = \frac{\sqrt{\det(\hat{\mathbf{R}}_{\acute{\mathbf{y}}_N})}}{(\max_{n\in\{N,\ldots,L_1\}} \|\acute{\mathbf{y}}_N(n)\|_1)^{2Np}} \tag{4.22}$$

Again the logarithm of (4.22) is computed for iterative algorithm derivation:

$$\begin{aligned} \mathcal{J}_c(\tilde{\mathbf{W}}) &= \log(J_c(\tilde{\mathbf{W}})) \\ &= \frac{1}{2}\log\left(\det(\hat{\mathbf{R}}_{\acute{\mathbf{y}}_N})\right) - 2Np\log(\max_{n\in\{N,\ldots,L_1\}} \|\acute{\mathbf{y}}_N(n)\|_1) \\ &= \frac{1}{2}\log\left(\det(\Gamma_N(\Omega(\tilde{\mathbf{W}}))\hat{\mathbf{R}}_{\acute{\mathbf{x}}_{\mathbf{N+M-1}}}\Gamma_N(\Omega(\tilde{\mathbf{W}}))^T)\right) - 2Np\log(\max_{n\in\{N,\ldots,L_1\}} \|\acute{\mathbf{y}}_N(n)\|_1) \end{aligned} \tag{4.23}$$

The iterative update equation for the separator matrix at time lag k is formulated as

$$\mathbf{W}^{(t+1)}(k) = \mathbf{W}^{(t)}(k) + \sigma^{(t)} \Big(\mathbf{T}_{1:p,2kq+1:(2k+1)q} + \mathbf{T}_{p+1:2p,(2k+1)q+1:2(k+1)q} + j(\mathbf{T}_{p+1:2p,2kq+1:(2k+1)q} - \mathbf{T}_{1:p,(2k+1)q+1:2(k+1)q}) - 2Np\frac{\sum_{m=0}^{N-1} \check{\mathbf{o}}_{mp+1:(m+1)p,(m+k)q+1:(m+k+1)q}}{\max_{n\in\{N,\ldots,L_1\}} \|\acute{\tilde{\mathbf{y}}}_N(n)\|_1}\Big) \tag{4.24}$$

where $\check{\mathbf{o}} = \text{signc}\{\tilde{\mathbf{y}}_N(l^{(t)})\}\tilde{\mathbf{x}}_{N+M-1}(l^{(t)})^H$, $\mathbf{T} = \sum_{r=0}^{N-1} \Omega(\mathbf{Z})^{(t)}_{2rp+1:2(r+1)p,2rq+1:2(r+M)q}$ and $\Omega(\mathbf{Z}) = \left(\Gamma_N(\Omega(\tilde{\mathbf{W}}))\hat{\mathbf{R}}_{\acute{\tilde{\mathbf{x}}}_{\mathbf{N+M-1}}}\Gamma_N(\Omega(\tilde{\mathbf{W}}))^T\right)^{-1}\Gamma_N(\Omega(\tilde{\mathbf{W}}))\hat{\mathbf{R}}_{\acute{\tilde{\mathbf{x}}}_{\mathbf{N+M-1}}}$. $l^{(t)}$ represents the time index for which maximum ℓ_1-norm at the separator output is achieved.

4.3. Time Domain Sparsity Based Convolutive SBCA Framework for Stationary Signals

Here, a new SBCA framework is defined in both time and frequency domains. The time-domain framework in Fig. 4.1 is used while some computations are executed in frequency domain. Proof for perfect separation, and an iterative algorithm for obtaining the separator matrix are also supplied in the following subsections. Lastly, we extend the algorithm to complex source case.

4.3.1 Objective Function

As a constraint, we assume that the source vectors are zero-mean wide sense stationary vector processes represented by $\mathcal{S} = \{\mathbf{s}(n) \in \mathbb{R}^p\}$ with a power spectral density (PSD) $\{\mathbf{P}_s(f) \in \mathbb{C}^{p \times p},\ \ f \in \left[\frac{-1}{2}, \frac{1}{2}\right]\}$.

We modify the objective function in (4.10) as

$$J(\tilde{\mathbf{W}}) = \frac{N}{2} \int_{-1/2}^{1/2} \log(\det(\mathbf{P}_{\mathbf{y}}(f)))df - Np\log(\max_{n \in \{N,\ldots,L_1\}} \|\tilde{\mathbf{y}}_N(n)\|_1), \quad (4.25)$$

where $L_1 = L + M - 1$, $\tilde{\mathbf{y}}_N(n) = [\mathbf{y}^T(n), \mathbf{y}^T(n-1), \ldots, \mathbf{y}^T(n-N+1)]^T$, and $\mathbf{P}_{\mathbf{y}}(f)$ is the separator output PSD. Accordingly, the volume definition of the principal hyper-ellipsoid is changed and extended to random process based auto-covariance which is a function of time lag between separator output samples.

Assumption (**A4**) is again utilized as the basic assumption. Based on this assumption, the following theorem is put forward.

Theorem-4: Provided that $\tilde{\mathbf{A}}$ can be equalized by an FIR extractor filter of order $M-1$ and the assumption (**A4**) is correct, all global maximums of (4.25) give perfect separation.

The proof of **Theorem-4** is provided in Appendix B.

4.3.2 Iterative Algorithm for Separator Estimation

Since there is a finite set of observations, $\{\mathbf{x}(1), \ldots, \mathbf{x}(L)\}$, we modify the objective in (4.25) as

$$J(\tilde{\mathbf{W}}) = \frac{N}{2\beta} \sum_{l=-L_1}^{L_1} \log(\det(\hat{\mathbf{P}}_{\mathbf{y}}(l))) - Np \log(\max_{n \in \{N,\ldots,L_1\}} \|\tilde{\mathbf{y}}_N(n)\|_1), \tag{4.26}$$

where $L_1 = L + M - 1$ and $\beta = 2L_1 + 1$ is the DFT length. The PSD estimate is given by

$$\hat{\mathbf{P}}_{\mathbf{y}}(l) = \sum_{k=-L_1}^{L_1} \hat{\mathbf{R}}_{\mathbf{y}}(k) e^{-j2\pi kl/\beta}, \tag{4.27}$$

$$\hat{\mathbf{R}}_{\mathbf{y}}(k) = \frac{1}{L_1 - |k|} \sum_{q=\max(1,-k)}^{\min(L_1, L_1-k)} \mathbf{y}(q)\mathbf{y}^T(q+k). \tag{4.28}$$

The derivative of the first part of (4.26) is

$$\frac{N}{2\beta} \frac{\partial(\sum_{l=-L_1}^{L_1} \log(\det(\hat{\mathbf{P}}_{\mathbf{y}}(l))))}{\partial \mathbf{W}(k)} = \frac{N}{\beta} \sum_{l=-L_1}^{L_1} \mathrm{Re}\{\hat{\mathbf{P}}_{\mathbf{y}}(l)^{-1} \mathbb{W}(l) \hat{\mathbf{P}}_{\mathbf{x}}(l) e^{j2\pi kl/\beta}\}, \tag{4.29}$$

where

$$\mathbb{W}(l) = \sum_{m=-L_1}^{L_1} \mathbf{W}(m) e^{-j2\pi ml/\beta}, \tag{4.30}$$

$$\hat{\mathbf{P}}_{\mathbf{x}}(l) = \sum_{m=-L_1}^{L_1} \hat{\mathbf{R}}_{\mathbf{x}}(m) e^{-j2\pi ml/\beta}. \tag{4.31}$$

The derivative of the second part is

$$\frac{\partial(Np \log(\max_{n \in \{N,\ldots,L_1\}} \|\tilde{\mathbf{y}}_N(n)\|_1))}{\partial \mathbf{W}(k)} = Np \frac{\sum_{r=0}^{N-1} \mathbf{o}_{rp+1:(r+1)p,(r+k)q+1:(r+k+1)q}}{\max_{n \in \{N,\ldots,L_1\}} \|\tilde{\mathbf{y}}_N(n)\|_1}, \tag{4.32}$$

where $\mathbf{o} = \text{sign}\{\tilde{\mathbf{y}}_N(m^{(t)})\}\tilde{\mathbf{x}}_{N+M-1}(m^{(t)})^T$, and $m^{(t)}$ is chosen arbitrarily from the index set of maximum output ℓ_1-norm at the t^{th} iteration. Combining the derivatives of the two parts results the iterative update equation

$$\begin{aligned}\mathbf{W}^{(t+1)}(k) = \mathbf{W}^{(t)}(k) + \\ \sigma^{(t)}\Bigg(\frac{N}{\beta}\sum_{l=-L_1}^{L_1}\text{Re}\{\hat{\mathbf{P}}_{\mathbf{y}}(l)^{-1}\mathbb{W}(l)\hat{\mathbf{P}}_{\mathbf{x}}(l)e^{j2\pi kl/\beta}\} - \\ Np\frac{\sum_{r=0}^{N-1}\mathbf{o}_{rp+1:(r+1)p,(r+k)q+1:(r+k+1)q}}{\max_{n\in\{N,\ldots,L_1\}}\|\tilde{\mathbf{y}}_N(n)\|_1}\Bigg)\end{aligned} \tag{4.33}$$

4.3.3 Extension of Stationary CSBCA to Complex Case

In this case, the source signals are considered to be complex with finite support in the complex domain. For complex extension, the mathematical terms and isomorphism based approach described in section 4.2.4 are again applied. Accordingly, the objective function in (4.25) is modified to

$$J_c(\tilde{\mathbf{W}}) = \frac{N}{2}\int_{-1/2}^{1/2}\log(\det(\mathbf{P}_{\acute{\tilde{\mathbf{y}}}}(f)))df - 2Np\log(\max_{n\in\{N,\ldots,L_1\}}(\|\acute{\tilde{\mathbf{y}}}_N(n)\|_1)) \tag{4.34}$$

For the finite set of observations, the objective function turns into,

$$J_c(\tilde{\mathbf{W}}) = \frac{N}{2\beta}\sum_{l=-L_1}^{L_1}\log(\det(\hat{\mathbf{P}}_{\acute{\tilde{\mathbf{y}}}}(l))) - 2Np\log(\max_{n\in\{N,\ldots,L_1\}}(\|\acute{\tilde{\mathbf{y}}}_N(n)\|_1)) \tag{4.35}$$

The derivative of the first part of $J_c(\tilde{\mathbf{W}})$ is

$$\frac{N}{2\beta}\frac{\partial(\sum_{l=-L_1}^{L_1}\log(\det(\hat{\mathbf{P}}_{\acute{\tilde{\mathbf{y}}}}(l))))}{\partial\mathbf{W}(k))} = \mathbf{Z}_{11} + \mathbf{Z}_{22} + j(\mathbf{Z}_{21} - \mathbf{Z}_{12}), \tag{4.36}$$

where $\mathbf{Z}_{ij} \in \mathbb{R}^{p\times q}$ and is obtained from the derivative with respect to the real isomorphic matrix $\Omega(\mathbf{W}(k))$:

$$\frac{N}{\beta}\sum_{l=-L_1}^{L_1} \mathrm{Re}\{\hat{\mathbf{P}}_{\acute{\mathbf{y}}}(l)^{-1}\Omega(\mathbb{W})(l)\hat{\mathbf{P}}_{\acute{\mathbf{x}}}(l)e^{j2\pi kl/\beta}\} = \begin{bmatrix} \mathbf{Z}_{11} & \mathbf{Z}_{12} \\ \mathbf{Z}_{21} & \mathbf{Z}_{22} \end{bmatrix} \tag{4.37}$$

where $\Omega(\mathbb{W})(l) = \sum_{m=-L_1}^{L_1} \Omega(\mathbf{W}(m))e^{-j2\pi ml/\beta}$. The derivative of the second part of $J_c(\tilde{\mathbf{W}})$ is

$$\frac{\partial(2Np\log(\max_{n\in N,\ldots,L_1}\|\acute{\tilde{\mathbf{y}}}_N(n)\|_1))}{\partial\mathbf{W}(k)} = 2Np\frac{\sum_{r=0}^{N-1}\dot{\mathbf{o}}_{rp+1:(r+1)p,(r+k)q+1:(r+k+1)q}}{\max_{n\in\{N,\ldots,L_1\}}\|\acute{\tilde{\mathbf{y}}}_N(n)\|_1}, \tag{4.38}$$

where $\dot{\mathbf{o}} = \mathrm{signc}\{\tilde{\mathbf{y}}_N(m^{(t)})\}\tilde{\mathbf{x}}_{N+M-1}(m^{(t)})^H$.

Finally, the corresponding iterative update equation for the separator matrix at time lag k turns out to be

$$\begin{aligned}\mathbf{W}^{(t+1)}(k) = \mathbf{W}^{(t)}(k) + \sigma^{(t)}(&\mathbf{Z}_{11}^{(t)} + \mathbf{Z}_{22}^{(t)} + j(\mathbf{Z}_{21}^{(t)} - \mathbf{Z}_{12}^{(t)}) - \\ &2Np\frac{\sum_{r=0}^{N-1}\dot{\mathbf{o}}_{rp+1:(r+1)p,(r+k)q+1:(r+k+1)q}}{\max_{n\in\{N,\ldots,L_1\}}\|\acute{\tilde{\mathbf{y}}}_N(n)\|_1}).\end{aligned} \tag{4.39}$$

4.4. Frequency Domain Sparsity Based Convolutive SBCA Framework

In this section, a frequency-domain-sparsity based SBCA objective function for convolutive mixtures is introduced and the corresponding iterative algorithm is derived. The fact that many non-sparse signals are mapped to sparse forms in frequency domain motivated us to define a new objective function in frequency domain. In addition to sparsification, mathematical expressions are simplified since convolutive mixtures transform to instantaneous mixtures. However, we do not attempt to derive an algorithm for estimating instantaneous separator matrices for different frequency bins which suffer from scaling and permutation inconsistencies between different frequency bins. Instead, we focus on deriving an iterative update algorithm for a single temporal separator matrix.

4.4.1 BSS Setup in Frequency Domain

For the spectral BSS signal setup,

- Time-frequency spectrum representations of p magnitude-bounded sources via Short Time Fourier Transform (STFT) are formulated as

$$s_i(m,f) = \sum_n s_i(n)v(n-mT)e^{-j2\pi fn}, \quad i = 1,\ldots,p, \tag{4.40}$$

  where $f \in [-1/2, 1/2]$ is the frequency index, $v(n)$ is a time-window function (Hamming, Kaiser, etc.) of length R, $R \in \mathbb{Z}$ is the STFT frame length, $T \in \mathbb{Z}$ is the hop size, in samples, between successive DTFTs. Let L_v denote the index set of STFT frames taken over the sample set $\{1,\ldots,L\}$. STFT computation maps $p \times L$ size temporal data to $p \times dim\{L_v\}$ size data at the frequency bin f.

- Since STFT creates complex source vectors, the isomorphism based approach described in Section 4.2.4 will be used. $2p$ dimensional real isomorphic counterpart of the spectral data is written as

$$\acute{\mathbf{s}}(m,f) = \left[\mathrm{Re}(\mathbf{s}(m,f))^T \quad \mathrm{Im}(\mathbf{s}(m,f))^T\right]^T. \tag{4.41}$$

- For each frequency bin, the unity-ℓ_1-norm ball that bounds the real isomorphic source sample vectors is defined by

$$\mathbb{B}_{\acute{\mathcal{S}}} = \{\acute{\mathbf{s}}(m,f) \in \mathbb{R}^{2p} | \quad \|\acute{\mathbf{s}}(m,f)\|_1 \leq 1, \forall m\}. \tag{4.42}$$

- The convolutive mixing channel is linear, time invariant (LTI) and equalizable [Inouye and Liu, 2002].

- The mixtures of the source STFTs are formulated as the multiplication of the source STFTs and the mixing filter response's DTFT:

$$\mathbf{x}(m,f) = \mathbb{A}(f)\mathbf{s}(m,f). \tag{4.43}$$

$\mathbb{A}(f)$ denotes the DTFT of the mixing filter elements at frequency bin f. Computing the inverse STFTs of $\mathbf{x}(m,f)$ and adding them up with the corresponding overlaps gives exactly the same result with (4.2). However, the mixture STFTs can not be computed with (4.43) in practice.

- In practice, the STFT of the mixtures can be computed by

$$\tilde{x}_k(m,f) = \sum_n x_k(n)v(n-mT)e^{-j2\pi fn}, \tag{4.44}$$

$$\tilde{\mathbf{x}}(m,f) = [\tilde{x}_1(m,f), \tilde{x}_2(m,f), \ldots, \tilde{x}_q(m,f)]^T. \tag{4.45}$$

Due to boundary effects [Mei et al., 2008], $\tilde{\mathbf{x}}(m,f)$ and $\mathbf{x}(m,f)$ in (4.43) are not equal. On the other hand, the amount of different samples, that is equal to the mixing filter order, can be ignored if a sufficiently long STFT window type is used and the mixing channel has a fast decaying response. If these conditions are realized, it results

$$\mathbf{x}(m,f) \approx \tilde{\mathbf{x}}(m,f). \tag{4.46}$$

Under these circumstances, the observed frequency elements of the mixing channel outputs can be assumed to be instantaneous mixtures of the frequency elements of the sources in accordance to (4.43).

- By using (4.41) and the mapping operator $\Gamma : \mathbb{C}^{q\times p} \to \mathbb{R}^{2q\times 2p}$ defined in (3.51), (4.43) is transformed to real terms:

$$\acute{\mathbf{x}}(m,f) = \Gamma(\mathbb{A}(f))\acute{\mathbf{s}}(m,f). \tag{4.47}$$

- The separator output for mixture STFTs is produced by the multiplication of the separator matrix' DTFT and mixture STFTs

$$\mathbf{y}(m,f) = \mathbb{W}(f)\mathbf{x}(m,f), \tag{4.48}$$

where $\mathbb{W}(f), f \in [-1/2, 1/2]$ is the separator frequency response of $p \times q$ size.

- The separator output's STFT is computed with the separator output samples as

$$\tilde{y}_k(m,f) = \sum_n y_k(n)v(n-mT)e^{-j2\pi fn}, \tag{4.49}$$

$$\tilde{\mathbf{y}}(m,f) = \left[\tilde{y}_1(m,f), \tilde{y}_2(m,f), \ldots, \tilde{y}_p(m,f)\right]^T. \tag{4.50}$$

$\mathbf{y}(m,f)$ and $\tilde{\mathbf{y}}(m,f)$ are approximately equal due to (4.46).

- The real isomorphic counterpart of the separator output is defined in terms of the source signal vector and the total system frequency response $\mathbb{F}(f) =$

$\mathbb{W}(f)\mathbb{A}(f), \quad f \in [-1/2, 1/2]$ as

$$\acute{\mathbf{y}}(m,f) = \Gamma(\mathbb{F}(f))\acute{\mathbf{s}}(m,f). \tag{4.51}$$

- If perfect separation is accomplished, the real isomorphic counterpart of the system frequency response turns out to

$$\Gamma(\mathbb{F}(f)) = \mathbf{PD}(f), \tag{4.52}$$

where $\mathbf{P}$ is a permutation matrix and $\mathbf{D}(f)$ is a diagonal scaling matrix. It means that the total system response is transparent if perfect separation is realized.

The basic assumption for the frequency domain sparsity based optimization setting in terms of the source sample set $\acute{\mathbf{S}}(f) = \{\acute{\mathbf{s}}(m,f) \in \mathbb{R}^{2p}, m \in L_v\}$ is given below.

Assumption(**A5**): For each frequency bin, the vertices of the bounding ℓ_1-norm-ball $\mathcal{B}_{\acute{s}}$ are included in the source sample set $\acute{\mathbf{S}}(f)$.

4.4.2 Objective Function

The frequency domain BSS setup resembles the instantaneous BSS setup introduced in section 3.2.1. That's why a similar geometrical optimization setting can be offered for this setup. To make it possible, the temporal principal hyper-ellipsoid definition in section 3.2 should be adapted to the spectral data. This can be provided via invoking the linear-time invariance property of the mixing and separator filters. Power Spectral Density Matrix (PSDM) of source processes is formulated as

$$\mathbf{P_s}(f) = \lim_{L\to\infty} E\left[\mathbf{s}(f)\mathbf{s}^H(f)\right], \tag{4.53}$$

where $\mathbf{s}(f)$ is the source DTFT, L is the sample size of $\mathbf{s}(n)$. Source PSDM can be redefined in accordance with the relation between the source DTFTs and STFTs as

$$\mathbf{P}_{\mathbf{s}}(f) = \lim_{L_v \to \infty} \sum_{m \in L_v} \mathbf{P}_{\mathbf{s}}(m, f), \tag{4.54}$$

where $\mathbf{P}_{\mathbf{s}}(m, f)$ is the instant PSDM of $\mathbf{s}(m, f)$ [Mei et al., 2008]. The separator output PSDM is linked to the source PSDM as

$$\mathbf{P}_{\mathbf{y}}(f) = \mathbb{F}(f)\mathbf{P}_{\mathbf{s}}(f)\mathbb{F}(f)^H, \tag{4.55}$$

where $\mathbf{P}_{\mathbf{y}}(f) \in \mathbb{C}^{p\times p}$. For a given frequency bin, this is equal to the covariance matrix of the separator output spectrum. Thus, the determinant of the separator output PSDM can be used to define the principal hyper-ellipsoid volume in frequency domain.

After the volume definition extension, the frequency domain objective function is proposed to be

$$J(\tilde{\mathbf{W}}) = \int_{f=-1/2}^{1/2} \left(\frac{1}{2} \log\left(\det(\Gamma\left(\mathbf{P}_{\mathbf{y}}(f)\right))\right) - 2p \log\left(\max_{m \in L_v} \|\acute{\mathbf{y}}(m, f)\|_1 \right) \right) df, \tag{4.56}$$

where $\Gamma(\mathbf{P}_{\mathbf{y}}(f))$ is the corresponding real isomorphic matrix. The objective function is defined as a $\tilde{\mathbf{W}}$ dependent function through the use of the integral term in (4.56). Without this integral term, it would not be possible to estimate a single temporal separator matrix for the whole frequency band, and it would be necessary to estimate a different separator matrix for each frequency bin.

The objective function in (4.56) can be simplified by replacing the real isomorphic PSD term $\Gamma\left(\mathbf{P}_{\mathbf{y}}(f)\right)$ with the complex PSD, $\mathbf{P}_{\mathbf{y}}(f)$, based on $|\det(\Gamma\left(\mathbb{F}(f)\right))| = |\det(\mathbb{F}(f))|^2$ [Erdogan, 2013], and replacing $2p$ with p:

$$J_{alt}(\tilde{\mathbf{W}}) = \int_{f=-1/2}^{1/2} \left(\frac{1}{2} \log\left(\det(\mathbf{P}_{\mathbf{y}}(f))\right) - p \log\left(\max_{m \in L_v} \|\acute{\mathbf{y}}(m, f)\|_1 \right) \right) df. \tag{4.57}$$

The following theorem ensures that the maximums of (4.56) accomplishes blind separation of temporally convolutive mixtures.

Theorem-5: Provided that (**A5**) is correct and the mixing channel $\mathbb{A}(f)$ can be equalized by an FIR filter, all global maximums of (4.56) correspond to perfect separators.

The proof of **Theorem-5** is provided in Appendix C.

4.4.3 Iterative Algorithm for Separator Estimation

Because the observation set $\{\mathbf{x}(1), \ldots, \mathbf{x}(L)\}$ is finite, the integral term in the objective function is replaced with a summation term. This yields the expression

$$J(\tilde{\mathbf{W}}) = \frac{1}{\beta} \sum_{l=-R_1+1}^{R_1-1} \left(\frac{1}{2} \log \left(\det(\Gamma(\hat{\mathbf{P}}_{\mathbf{y}}(l))) \right) - 2p \log \left(\max_{m \in L_v} \|\acute{\hat{\mathbf{y}}}(m,l)\|_1 \right) \right), \quad (4.58)$$

where $\beta = 2R_1 - 1$ is the DFT size and $\hat{\mathbf{P}}_{\mathbf{y}}(l)$ is the output PSDM estimate. Note that $\acute{\hat{\mathbf{y}}}(m,l)$ is substituted with $\acute{\mathbf{y}}(m,l)$ in accordance to (4.46). $\hat{\mathbf{P}}_{\mathbf{y}}(l)$ can be expressed in terms of the mixture PSDM estimate as

$$\hat{\mathbf{P}}_{\mathbf{y}}(l) = \mathbb{W}(f) \hat{\mathbf{P}}_{\mathbf{x}}(l) \mathbb{W}(f)^H, \quad (4.59)$$

where $\hat{\mathbf{P}}_{\mathbf{x}}(l)$ is given by

$$\hat{\mathbf{P}}_{\mathbf{x}}(l) = \sum_{m \in L_v} \hat{\mathbf{P}}_{\mathbf{x}}(m,l). \quad (4.60)$$

$\hat{\mathbf{P}}_{\mathbf{x}}(m,l)$ is computed using the spectrogram method [Martin and Flandrin, 1985], i.e.,

$$\hat{\mathbf{P}}_{\mathbf{x}}(m,l) = \sum_{k} u(k) \tilde{\mathbf{x}}_k(m,l) \tilde{\mathbf{x}}_k(m,l)^H, \quad (4.61)$$

where $u(k)$ is a time-window of length R_2 that causes a weighted averaging of R_2 spectra so that the PSDM estimation variance gets lower. $\tilde{\mathbf{x}}_k(m,l)$ is given by

$$\tilde{\mathbf{x}}_k(m,l) = \sum_n \mathbf{x}(n+k)v(n+k-mT)e^{-j2\pi(n+k)l/\beta}. \tag{4.62}$$

The derivative of the first part of $J(\tilde{\mathbf{W}})$ with respect to $\mathbf{W}(k)$ for $k = 0, \ldots, M-1$ is

$$\begin{aligned} \frac{1}{2\beta}\frac{\partial\left(\sum\limits_{l=-R_1+1}^{R_1-1}\log\det(\Gamma(\hat{\mathbf{P}}_{\mathbf{y}}(l)))\right)}{\partial\mathbf{W}(k)} &= \partial\mathbf{W}_{logdet} \\ &= \mathrm{Re}\{\mathbf{Z}_{11}+\mathbf{Z}_{22}+j(\mathbf{Z}_{21}-\mathbf{Z}_{12})\}. \end{aligned} \tag{4.63}$$

The components $\mathbf{Z}_{ij} \in \mathbb{R}^{p\times q}$ are defined as follows:

$$\frac{1}{\beta}\sum_{f=-R_1+1}^{R_1-1}\Gamma(\hat{\mathbf{P}}_{\mathbf{y}}(l))^{-1}\Gamma(\mathbb{W}(l))\Gamma(\hat{\mathbf{P}}_{\mathbf{x}}(l))\mathbf{O}^T = \begin{bmatrix}\mathbf{Z}_{11} & \mathbf{Z}_{12}\\ \mathbf{Z}_{21} & \mathbf{Z}_{22}\end{bmatrix}, \tag{4.64}$$

where $\mathbf{O} = \begin{bmatrix}\cos 2\pi kl/\beta & \sin 2\pi kl/\beta \\ -\sin 2\pi kl/\beta & \cos 2\pi kl/\beta\end{bmatrix}$.

$\mathbb{W}(l)$ in (4.64) is formulated as

$$\mathbb{W}(l) = \sum_{k=-R_1+1}^{R_1-1}\mathbf{W}(k)e^{-j2\pi kl/\beta} \tag{4.65}$$

The derivative of the second part is formulated as

$$\begin{aligned} \frac{2p}{\beta}\frac{\partial\left(\sum\limits_{l=-R_1+1}^{R_1-1}\log(\max\limits_{m\in L_v}\|\acute{\tilde{\mathbf{y}}}(m,l)\|_1)\right)}{\partial\mathbf{W}(k)} &= \partial\mathbf{W}_{subg} \\ &= \frac{2p}{\beta}\frac{\sum\limits_{l=-R_1+1}^{R_1-1}\mathrm{Re}\{\mathrm{signc}\{\tilde{\mathbf{y}}(u_l,l)\}\tilde{\mathbf{x}}(u_l,l)^H e^{j2\pi kl/\beta}\}}{\max\limits_{m\in L_v}\|\acute{\tilde{\mathbf{y}}}(m,l)\|_1}, \end{aligned} \tag{4.66}$$

where $u_l \in L_v$ denotes the frame index of maximum output ℓ_1-norm in frequency bin l.

Combining $\partial\mathbf{W}_{logdet}$ and $\partial\mathbf{W}_{subg}$ gives an iterative update equation for temporal separator matrix estimation, i.e.,

$$\mathbf{W}^{(t+1)}(k) = \mathbf{W}^{(t)}(k) + \sigma^{(t)}\left(\partial\mathbf{W}^{(t)}_{logdet} - \partial\mathbf{W}^{(t)}_{subg}\right), \tag{4.67}$$

where $^{(t)}$ denotes iteration index.

For the alternative simplified objective in (4.57), we follow the same way. The derivative of the first part becomes

$$\begin{aligned}\frac{1}{2\beta}\frac{\partial\left(\sum_{l=-R_1+1}^{R_1-1}\log\det(\hat{\mathbf{P}}_{\mathbf{y}}(l))\right)}{\partial\mathbf{W}(k)} &= \partial\mathbf{W}_{logdet}\\ &= \frac{1}{\beta}\sum_{l=-R_1+1}^{R_1-1}\mathrm{Re}\{\hat{\mathbf{P}}_{\mathbf{y}}(l)^{-1}\mathbb{W}(l)\hat{\mathbf{P}}_{\mathbf{x}}(l)e^{j2\pi kl/\beta}\}.\end{aligned} \tag{4.68}$$

The derivative of the second part is the same as (4.66) except the exponential degree p:

$$\begin{aligned}\frac{p}{\beta}\frac{\partial\left(\sum_{l=-R_1+1}^{R_1-1}\log(\max_{m\in L_v}\|\acute{\tilde{\mathbf{y}}}(m,l)\|_1)\right)}{\partial\mathbf{W}(k)} &= \partial\mathbf{W}_{subg}\\ &= \frac{p}{\beta}\frac{\sum_{l=-R_1+1}^{R_1-1}\mathrm{Re}\{\mathrm{signc}\{\tilde{\mathbf{y}}(u_l,l)\}\tilde{\mathbf{x}}(u_l,l)^H e^{j2\pi kl/\beta}\}}{\max_{m\in L_v}\|\acute{\tilde{\mathbf{y}}}(m,l)\|_1}.\end{aligned} \tag{4.69}$$

4.5. Numerical Results

In the first and second experiments, we use synthetic sparse signals and compare the proposed time-domain-sparsity based algorithm to some benchmark convolutive BSS algorithms that have publicly available codes. In addition, we evaluate the performances of the complex CSBCA and the CSBCA for stationary signals in the first experiment. The third experiment is a practical application for the blind source separation (BSS) of speech signals mixed in a reverberant environment to analyse our frequency based algorithm labelled as FCSBCA. The fourth experiment is another practical application for the separation of recorded musical instruments.

In the experiments, Kurtosis Maximization algorithm (MaxKurtosis) optimizing kurtosis based contrast function [Castella and Moreau, 2010], the Joint-Approximate-Diagonalization (JAD) based methods proposed in [Parra and Spence, 2000, Rahbar and Reilly, 2005] labelled respectively as "convBSS", "ALS", Spatio-Temporal algorithm (STFICA) [Douglas et al., 2007] which is an extension of Fast-ICA method, time-based Convolutive Bounded Component Analysis (CBCA) [Inan and Erdogan, 2015], and Cho's under-determined convolutive BSS [Cho and Yoo, 2015] are used as benchmark. CBCA is evaluated only in the first and second experiments to prove CSBCA's superiority for sparse signals. And Cho's method is examined only in the third experiment because it is an under-determined BSS method and its publicly available code is optimized for audio source separation.

For performance evaluation, a well known method proposed in [Vincent et al., 2006, Fevotte et al., 2005] particularly for Blind Audio Signal Separation applications is used. A MATLAB toolbox for the method is publicly available. This tool quantifies performance by computing two values called Signal to Distortion Ratio (SDR), Signal to Interference Ratio (SIR). SDR is the ratio of signal power to the sum of interference, noise and artifacts power. SIR is the ratio of signal power to interference power. For simplification, only SDR values are illustrated in the first two experiments since SDR quantifies the total performance.

4.5.1 Synthetic Sparse Signals

In the first experiment, time-domain-sparsity based CSBCA is examined by analysing the synthetic signals of the benchmark dataset published in [A. Cichocki and Siwek,]. The first part of the experiment evaluates SDR performance of each method versus input SNR by running a scenario with 6 sources, 12 mixtures and mixing channel of order 3. The channel noise is assumed additive white Gaussian (AWGN). The second part of the experiment evaluates SDR performance versus number of mixing channels and mixing order by running a scenario with 5 sources, 10 mixtures and

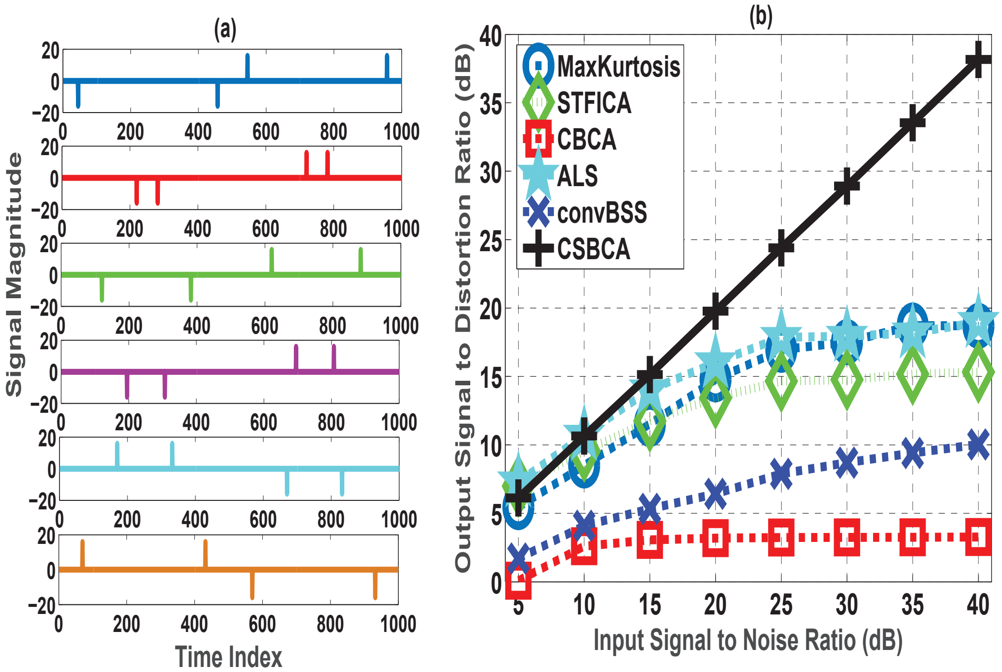


Figure 4.2: a)- Synthetic signals from RIKEN Laboratory website. b)- Output SDR vs input SNR.

input SNR=20 dB (additive white Gaussian noise). The synthetic signals from the dataset are illustrated in Fig. 4.2-(a). The performance curves against the input SNR are illustrated in Fig. 4.2-(b) and it can be interpreted that CSBCA yields better or equal performance almost at all SNRs. The results of the second part seem more complex to interpret. CBCA algorithm fails to extract the source signals satisfactorily, even at high SNR case as seen in Fig. 4.2-(b). Thus, number of mixtures and mixing channel order does not effect its performance any more as shown in Fig. 4.3. ALS and convBSS algorithms take benefit of source non-stationarity and perform a joint diagonalization on the cross power spectral density matrices of sources. The ALS algorithm differs from convBSS in the respect that it attempts to solve scaling problem inherent with frequency domain approaches while convBSS does not use such a method. This may be the main reason why convBSS fails to separate the sources satisfactorily as shown in Fig. 4.3 and ALS has a superiority over most of the other methods.

To evaluate the performances of CSBCA's complex extension and time-frequency version for stationary signals, we again use the synthetic sparse signals. In the com-

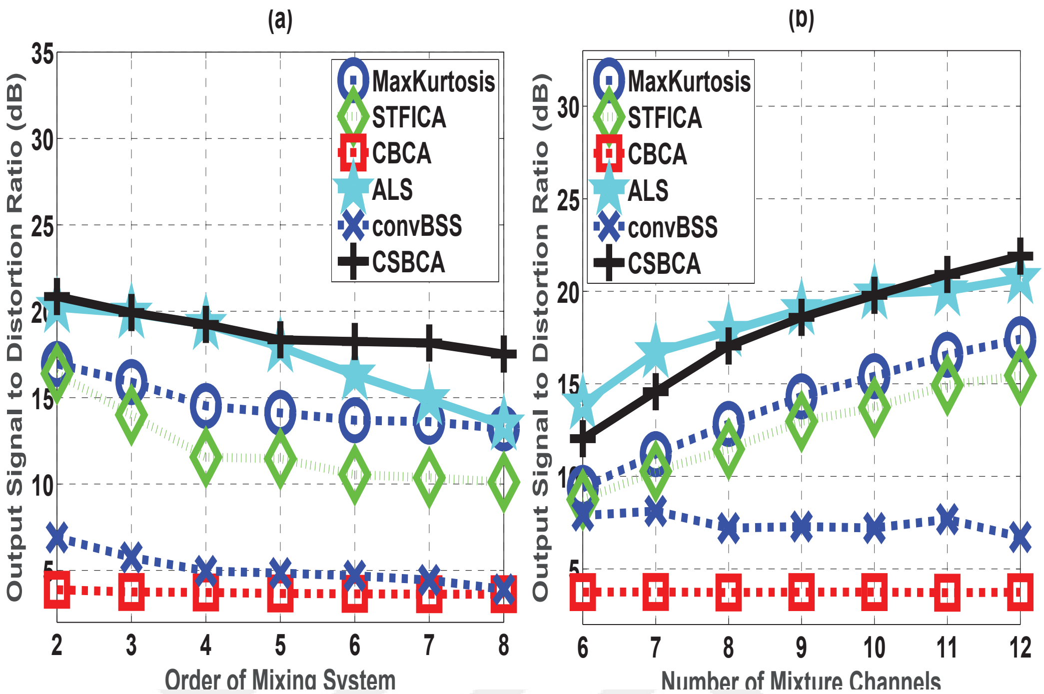


Figure 4.3: a)- Output SDR vs mixing order. b)- Output SDR vs number of mixing channels.

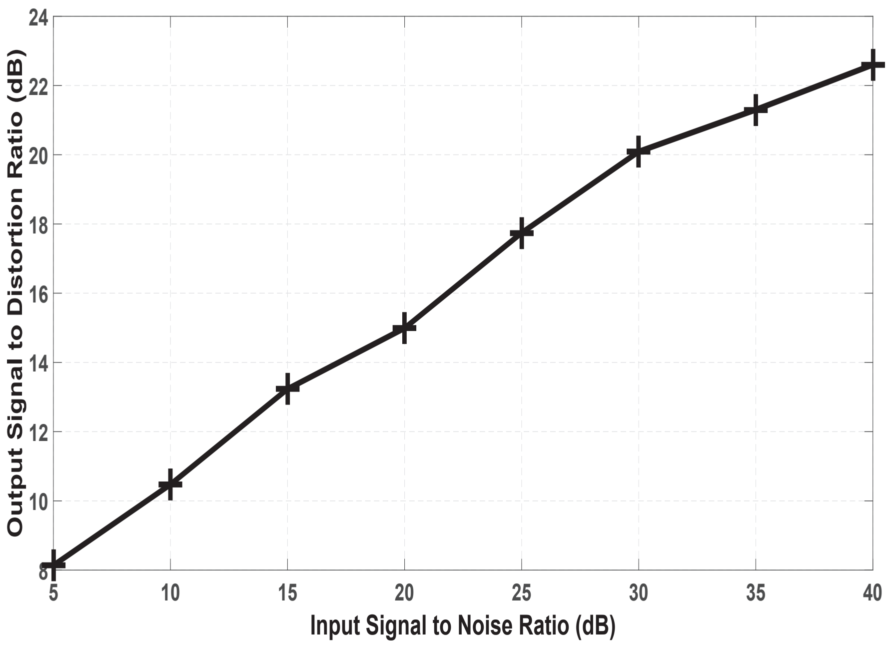


Figure 4.4: Performance of Complex CSBCA for 3 sources and 6 mixtures.

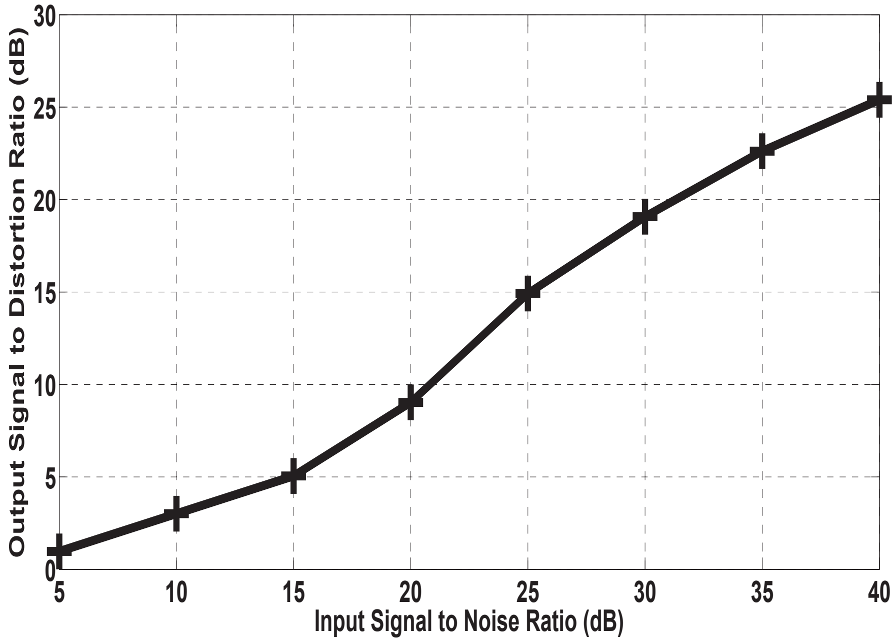


Figure 4.5: Performance of Stationary CSBCA for 5 sources and 10 mixtures.

plex CSBCA experiment, 3 complex sources are generated from 6 synthetic sources. Fig. 4.4 illustrates the SNR performance for complex CSBCA. In the stationary CSBCA experiment, we consider a scenario with 5 sources and 10 mixtures. Fig. 4.5 depicts stationary CSBCA's performance versus input SNR.

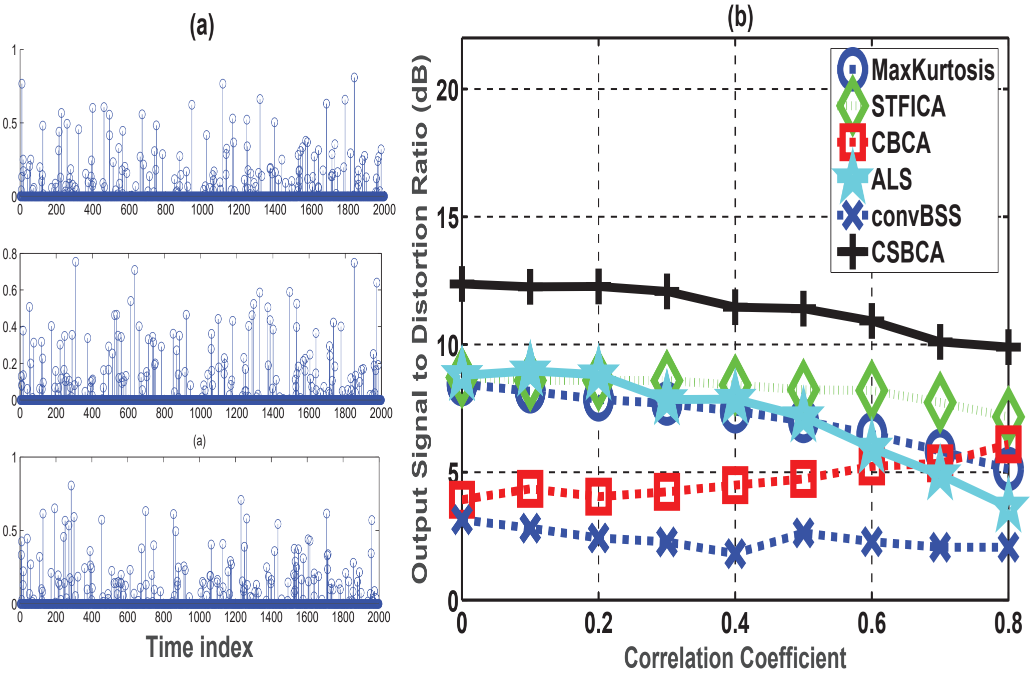


Figure 4.6: a)- Copula-T distributed sparse signals. b)- Output SDR vs correlation under SNR=20 dB.

4.5.2 Statistically Dependent Sources

In this experiment, statistically dependent sparse sources are generated with Copula-T distribution and random vector transformation

$$\mathbf{s} = \begin{cases} \mathbf{u}, & \mathbf{u} \in \mathbf{U}_r \\ \mathbf{0}, & otherwise \end{cases} \tag{4.70}$$

where $\mathbf{u} \in [-1, 1]^p$ and $\mathbf{U}_r = \mathbf{z} : \|\mathbf{z}_r\| \leq 1$ with $0 \leq r \leq 1$. The scenario includes 3 sources and 6 mixtures. It can be commented from the results in Fig. 4.6 that time based CSBCA performs better for all the correlation degrees. It seems that ALS is most effected by the source dependency. ALS and convBSS employ joint diagonalization of cross power spectral density matrices as the first stage in blind source separation. The source dependency can deteriorate this stage, that's why ALS may show a worse performance as the correlation increases. The performance degradation of MaxKurtosis and STFICA with increasing correlation is very reasonable because these methods assume statistical mutual independence of sources.

4.5.3 Speech Signal Separation

The third experiment evaluates the performance of the frequency-domain-sparsity based CSBCA compared to the benchmark methods. For this, three recorded speech signals (2 males and 1 female) and the measured channel impulse responses in MARDY (Multichannel Acoustic Reverberation Database at York) Database [Speech and Group,] are utilized. The records include the voices of different persons who read different texts. These signals are convolved with the channel responses from MARDY database which is an open library for dereverberation researchers. The method of recording the channel responses is explained in [Wen et al., 2006]. The original speech signals and their mixtures are depicted in Fig. 4.7. The config-

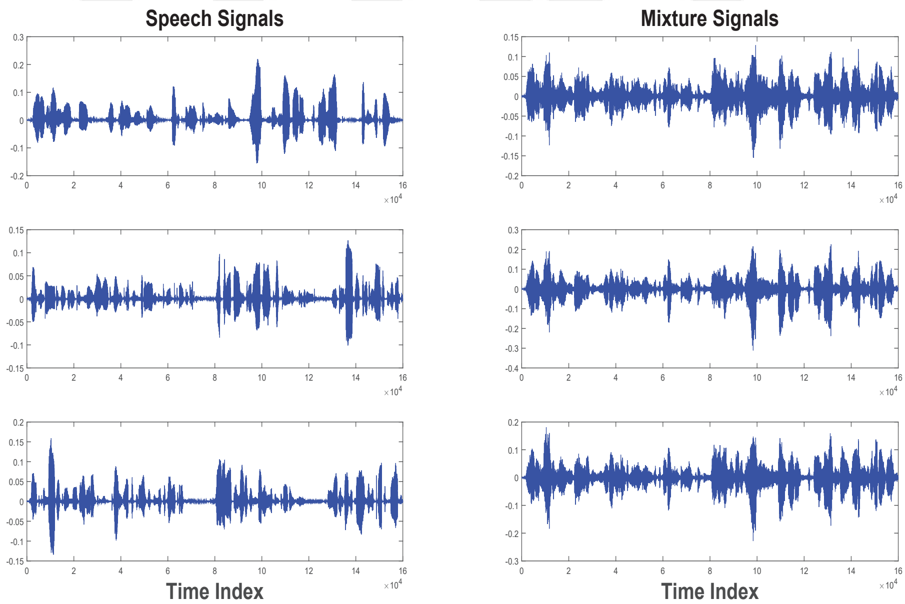


Figure 4.7: Speech signals and their mixtures.

uration parameters for each algorithm are listed below:

FCSBCA: Separation Filter Order= 80, STFT Window Length= 150, Overlap Ratio of STFT Windows= 0.8

ALS: Epoch Length= 1000, FFT Length= 512, Overlap Ratio of FFT Windows= 0.7

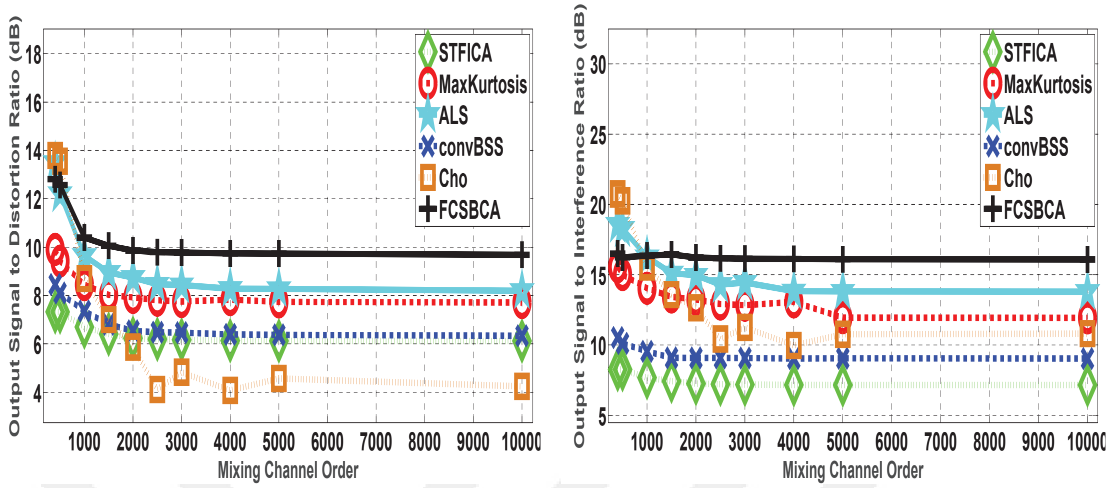
STFICA: Separator Filter Order= 100

convBSS: FFT Length= 512, Separator Filter Order= 100, Number of matrices to diagonalize= 5

Cho: FFT Length= 4096, Shift Length=512

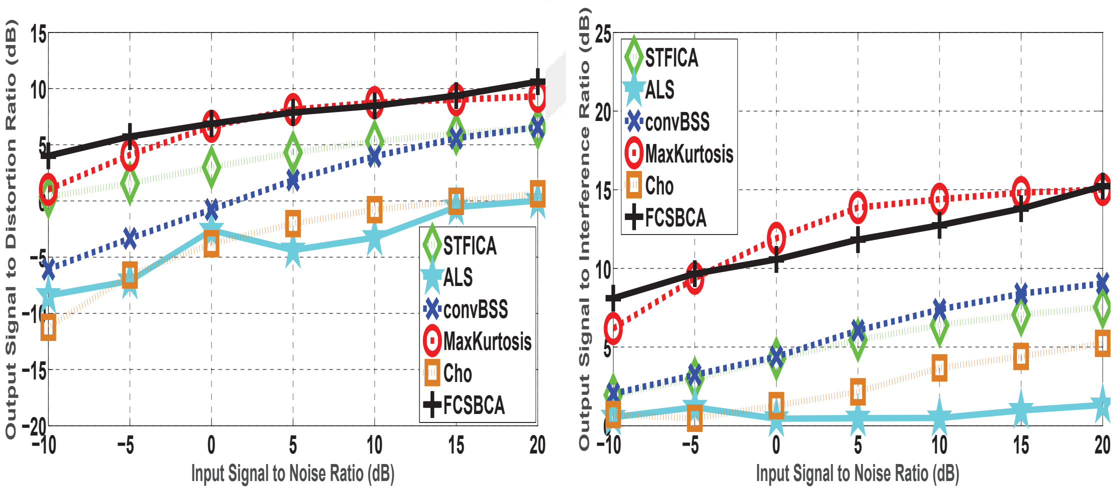
MaxKurtosis: Separator Filter Order= 100

Note that the term "Epoch" used in ALS method means FFT frame. The performance quantifications for different mixing channel orders and input SNRs are shown in Fig. 4.8, Fig. 4.9, Fig. 4.10 and Fig. 4.11. In the channel order analysis (Fig. 4.8, Fig. 4.10), microphones with noiseless channel responses (SNR=Inf) are used. In the input SNR analysis (Fig. 4.9, Fig. 4.11), microphones with channel order = 500 are used. It is clear that FCSBCA yields better performance than all the others with respect to both SDR and SIR values for the most of different cases. As a result, it can be said that the results show a practical merit of FCSBCA.



(a) (b)

Figure 4.8: 2 speech signals in a reverberant scene. (a)-Output SDR versus mixing channel order. (b)-Output SIR versus mixing channel order.

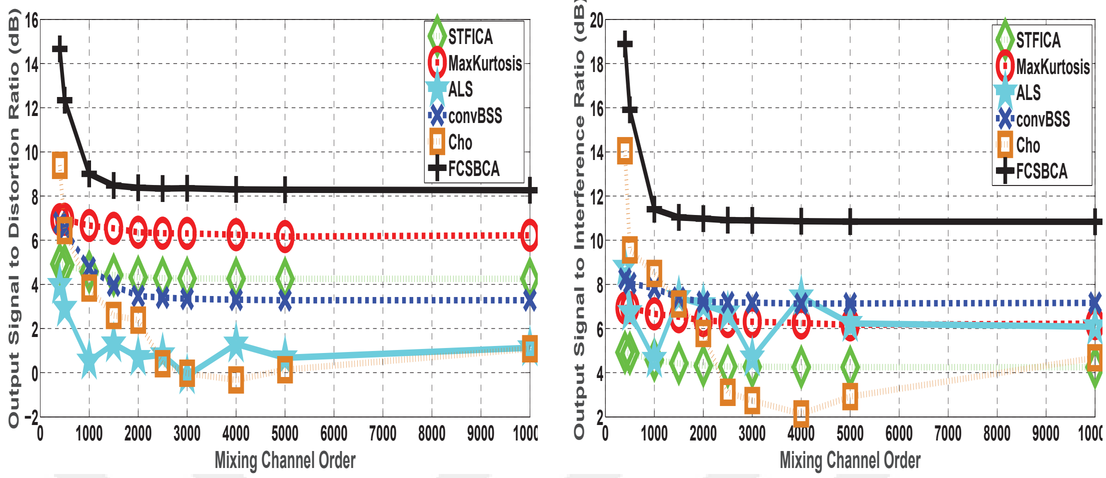


(a) (b)

Figure 4.9: 2 speech signals in a reverberant scene. (a)-Output SDR versus input SNR. (b)-Output SIR versus input SNR.

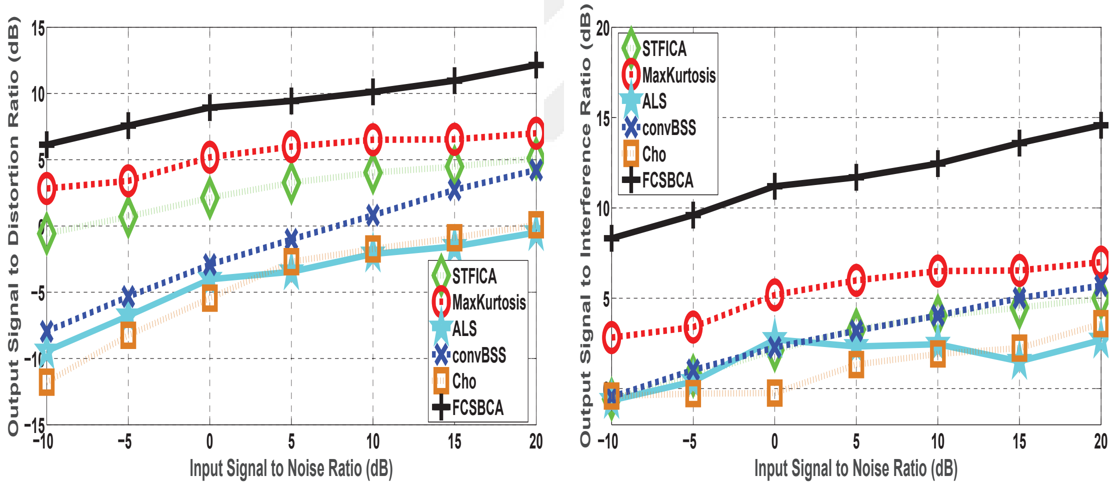
4.5.4 Musical Signal Separation

In the fourth example, we analyse the performance of the frequency-domain-sparsity based CSBCA for musical signal separation. Test signals are provided from SMAR database which consists of 960 audio recordings for 48 different loudspeaker and microphone array configurations. The detailed information about recording setup



(a) (b)

Figure 4.10: 3 speech signals in a reverberant scene. (a)-Output SDR versus mixing channel order. (b)-Output SIR versus mixing channel order.



(a) (b)

Figure 4.11: 3 speech signals in a reverberant scene. (a)-Output SDR versus input SNR. (b)-Output SIR versus input SNR.

can be found in [Nielsen et al., 2014]. In order to obtain mixture signals, we simply sum the channel recordings of the 3 musical instruments (flute, quartet and violin) as can be seen in Fig. 4.12. Selected configuration parameters for the algorithms are the same as in section 4.5.3.

The performances of the algorithms for different numbers of mixtures and input

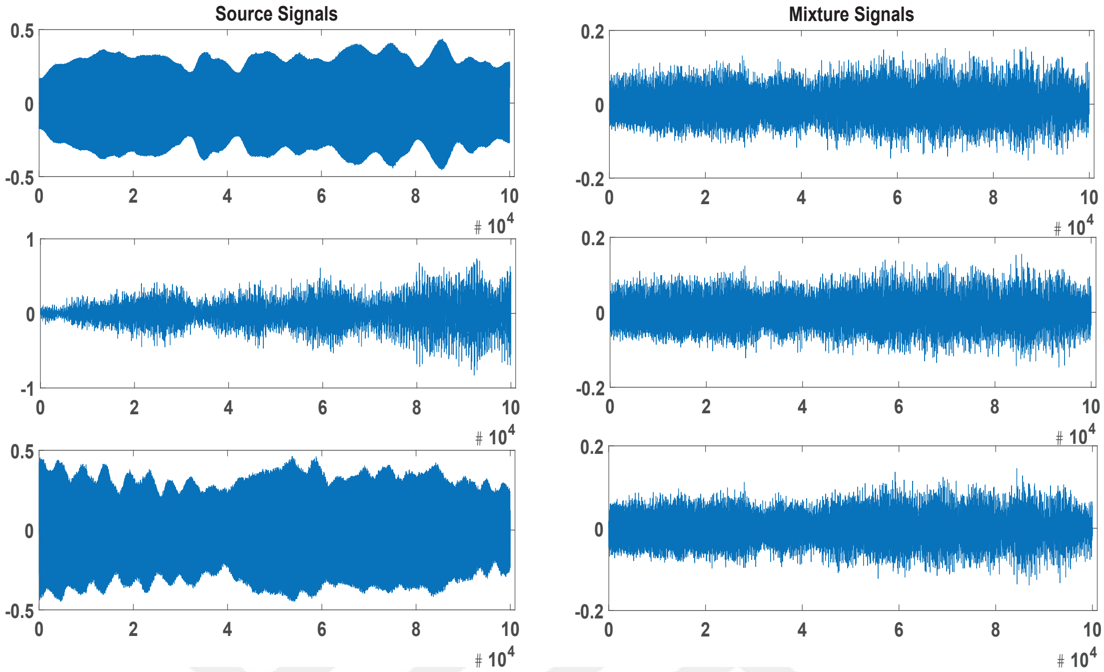
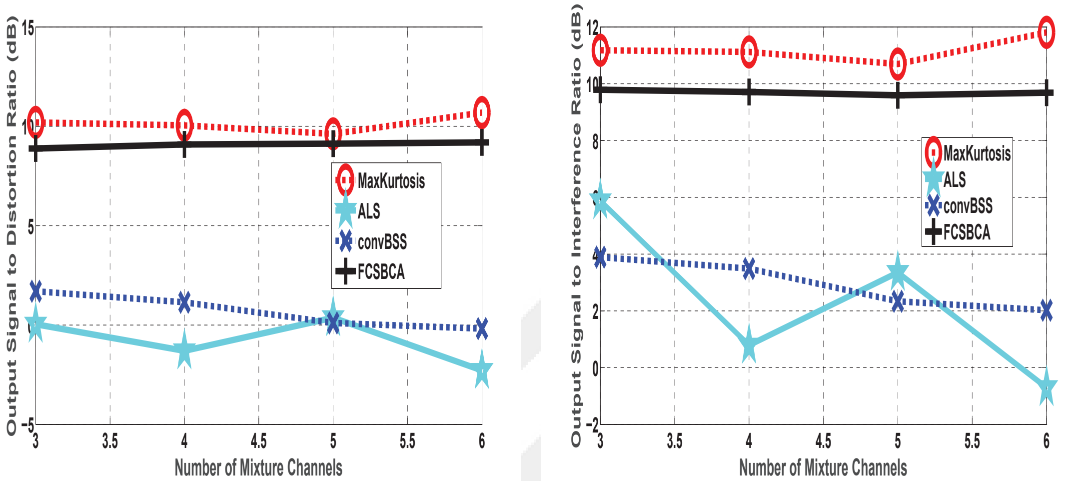


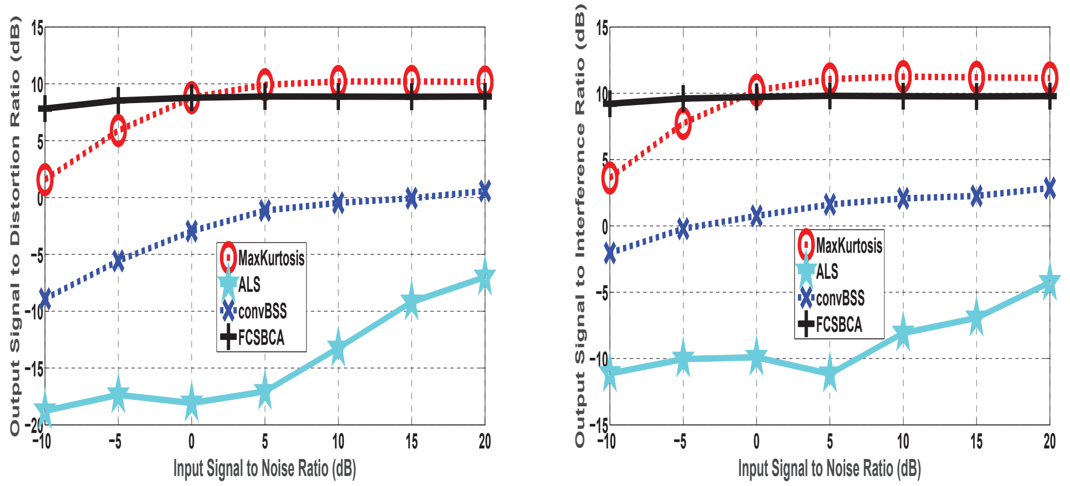
Figure 4.12: Signal Separation Setup: 3 musical signals (flute, quartet and violin), and 3 channel recordings.

SNRs are examined. In Fig. 4.13 and Fig. 4.14, the results for the 3 musical signals are illustrated. In the mixture number analysis, noiseless channel recordings (SNR=Inf) are used. In the input SNR analysis, 3 channel recordings are used. It yields that MaxKurtosis' outperforms FCSBCA for high input SNRs. But still, it can be said that FCSBCA provides performance improvement over the compared algorithms, particularly at low input SNRs.



(a) (b)

Figure 4.13: 3 musical signals (flute, quartet and violin) in a reverberant scene. (a)-Output SDR vs mixing channel order. (b)-Output SIR vs mixing channel order.



(a) (b)

Figure 4.14: 3 musical signals (flute, quartet and violin) in a reverberant scene. (a)-Output SDR vs input SNR. (b)-Output SIR vs input SNR.

Chapter 5

FUNCTIONAL BRAIN IMAGING WITH SPARSE BOUNDED COMPONENT ANALYSIS

Functional data refer data type that can be modelled as a function and can be illustrated as smooth curves, surfaces, or hyper-surfaces in a finite interval. Statistical analysis methods for studying functional data are called 'Functional Data Analysis'(FDA). Among numerous FDA methods, functional neuro-imaging [Marqui and Lehmann D., 1994, Michela et al., 2004, Wendel et al., 2009] has revolutionized in neuroscience. Localizing the different modules of the functional brain network in a given mental task by observing neural activity and predicting behaviour from the observations are the principal aims of functional neuro-imaging studies.

Over the last decades, functional neuro-imaging has been more commonly used to define the neuron activations during specific behaviours and cognitive tasks. Among all, some functional neuro-imaging techniques measure neural activity by directly observing changes in electrical activity (electroencephalography: EEG) or changes in magnetic fields related to electrical activity changes (magneto-encephalography:MEG). Some indirect observations can also be used to detect neural activities. As an example, Positron Emission Tomography (PET) measures changes in metabolism driven by energy requirement of neural activities.

As another functional neuro-imaging method, functional Magnetic Resonance Imaging (fMRI) monitors brain activities via observing blood flow changes that relies on a background result that couples cerebral blood flow and neural activity. There are many constituents in brain. Some of them : Gray Matter, White Matter, Cerebro Spinal Fluid Fat, Muscle/Skin, Glial Matter and so on, which exhibit unique properties under the effect of magnetic field. However, while scanning, we get one MRI image of the entire brain. Therefore, fMRI scans can be considered as the mixtures of brain activity regions. Some techniques, e.g., correlation and time-frequency analysis techniques have analysed changes in fMRI signal due to temporal characteristics of brain tasks. Correlation techniques rely on the assumption that voxels involved in brain tasks must generate different signal levels in performing these tasks [Bandettini et al., 1993]. These techniques take benefit of the expected time period information for task related signal level changes to specify the spatial extent and the intensity of relevant task. The main drawback of these techniques is that some non-task related processes correlated with task-related ones may be able to mask the task related activation signals. As another technique, time-frequency analyses describe a useful way to identify signals recorded from different voxels in frequency domain so that physiological pulsatiles and recurrent artifacts can be distinguished. Time-frequency analysis techniques mostly assume that temporal activities are periodic, however this might be invalid in some cases. Wavelet techniques can relieve this requirement [Brammer et al., 1997].

Although blind source separation has been used to separate physiologically different components, it has not received much attention for fMRI. McKeown et al [McKeown et al., 1998] introduced a BSS algorithm based on the ICA algorithm of Bell and Sejnowski [Bell and Sejnowski, 1995] to detect fMRI activations and distinguish between task related and non-task related components originated from movements and other artifacts. Nakai et al [Nakai et al., 2004] used Independent Component Analysis to extract physiologically independent components in fMRI scans. They reported that they were successful in increasing contrast for grey and white matter that are useful for brain tumour diagnosis. Snoussi and Calhoun [Snoussi and Calhoun, 2005] presented a maximum likelihood method based on second order statistics and using inter-source diversity to blindly separate brain activations in fMRI data.

Optical neuro-imaging techniques have been extensively researched for measuring cell activations based on the light penetration into cells and tissues. Calcium (Ca^{2+}) imaging technique is popular among optical neuro-imaging techniques for observing activities of neuron populations in real time [Tian et al., 2009, Jin et al., 2012]. Through this imaging technique, fluorescent emissions reflect the neural activities.

Many researchers have showed that the observed temporal dynamics of fluorescent emissions are sufficient for discovering spatio-temporal properties of neuron populations [Ziv et al., 2013, Mohammed et al., 2016].

In this chapter, we target to achieve the neural activity detection task based on fMRI and Ca^{2+} imaging data without resorting to any independence assumption regarding neurons, but based solely on the sparsity of their activities. For this purpose, we use SBCA approach introduced in chapter 3. We also propose to utilize denoising techniques for further cleaning component maps and activities obtained through the source separation step.

5.1. Neural Activity Detection Using functional Magnetic Resonance Images

The resultant time series data recorded in fMRI experiments contain a complicated mixture of brain tissue activities, motion artifacts and machine artifacts. Among all, function-related, task-related and transiently task-related signal records are of interest. Task specific responses of brain may be sometimes irregular, e.g., the response may change over time when stimuli is repeated, resulting in a transiently task-related signal. The function-related signals are characterized by functional similarities between voxels [Biswal et al., 1995]. Non-task related signals are unwanted signals such as motion-related artifacts. These artifacts usually appear as separate components. Relatively low image contrast-to-noise ratio and the artifacts make it difficult to detect signal changes associated with tasks.

ICA is a successful tool to analyse single-subject [McKeown et al., 1998, Biswal and Ulmer, 1999] or multi-subject [Calhoun et al., 2001] fMRI data sets. ICA aims to decompose data as a product of spatial patterns and corresponding time courses. In this case, signals can be statistically independent in space or time. Signal features such as non-Gaussianity and spatial/temporal independence were initially discussed in [McKeown et al., 1998] for the application of ICA to fMRI data. McKeown et al suggested that a particular task-related brain activity could be a functional combination of activities in different regions and the multi-focal brain regions activated by a psycho-motor task could not be related to those brain regions affected by non-task related activities. In this approach, each activity is attempted to be represented by one or more spatially-independent components. These independent spatial components are assumed to have sparse distribution that is claimed to work well with spatial ICA. Beside the usage of spatial distribution to extract activities, the idea that spatial patterns are activated by independent temporal processes was firstly followed by Biswal and Ulmer [Biswal and Ulmer, 1999] whose analysis was based on Bell-Sejnowski's temporal ICA [Bell and Sejnowski, 1995]. Comparison of spatial and temporal ICA methods was also analysed by Petersen et al [Petersen et al., 2000], and they put forward that most of the extracted task-related brain activities by spatial and temporal ICA methods were consistent. In other words, spatial and temporal ICA methods produced very similar brain component maps.

In this section we describe a new technique for fMRI data analysis based on SBCA algorithm. We utilize the sparsity assumption for blind separation in spatial domain, in other words, we propose a spatial BSS method like spatial ICA. The most important reason for this is that the spatial size of fMRI records is much bigger than their temporal size. Our method potentially allows the extraction of both spatially independent and dependent brain activity components since SBCA uses weaker domain separability assumption instead of statistical independence assumption.

The proposed technique for fMRI data analysis consists of 3 main processing phases explained below.

5.1.1 Image Masking Phase

Suppose that the k^{th} spatial slice of fMRI dataset is denoted with $\mathbf{X}_k$. $\mathbf{X}_k$ is an $L \times M$ matrix where L is the number of time points and M is the number of voxels indicating brain regions. The first processing step is to create a masked image map so that the expected areas of activation are more evident. For the masking, we use the acquired fMRI data at the first time point, i.e., $\mathbf{I}_{mask} = \mathbf{X}_k(1,:)$. The indices for which the voxel amplitude is greater than mean of all the voxels are selected to insert SBCA algorithm, i.e., $\mathbf{n}_{mask} = \{m|\mathbf{I}_{mask}(m) > \mu_{\mathbf{I}_{mask}}\}$ where

$\mu_{\mathbf{I}_{mask}} = \frac{1}{M}\sum_{l=1}^{M}\mathbf{I}_{mask}(l)$. In figure 5.1, an example fMRI record is masked. In the record, the subject performs a visuo-motor task involving two identical but spatially off-set, periodic, visual stimulus, shifted by 20 seconds from one another. Detailed explanation of this recording can be found in [Calhoun et al., 2003]. After masking, the non-active regions are cleaned and the boundaries of the active regions appear more clearly.

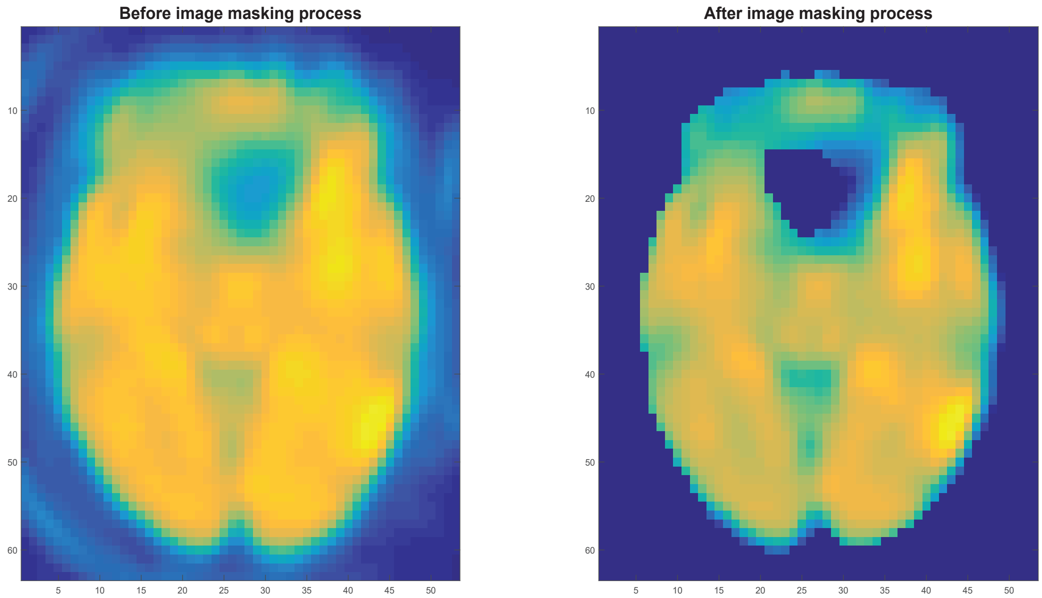


Figure 5.1: Image Masking Phase for a Visumotor fMRI record

5.1.2 Source Separation Phase

In this phase, we apply SBCA algorithm introduced in chapter 3 to each spatial slice of fMRI data separately. We assume that there are p components generating each slice of fMRI data which can be linearly separated using $\mathbf{Y}_k = \mathbf{W}_k\tilde{\mathbf{X}}_k$ where $\tilde{\mathbf{X}}_k = \mathbf{X}_k(:, \mathbf{n}_{mask})$ is the masked data of the k^{th} slice, $\mathbf{W}_k$ is a $p \times L$ unmixing matrix and $\mathbf{Y}_k$ is an $p \times \dim\{\mathbf{n}_{mask}\}$ matrix containing p independent or dependent components. The associated time courses are found in the columns of $\mathbf{W}_k^{-1}$. The SBCA algorithm iterations are applied to each slice of fMRI dataset separately to obtain the separator matrix $\mathbf{W}_k$:

$$\mathbf{W}_k^{(t+1)} = \mathbf{W}_k^{(t)} + \beta^{(t)}((\mathbf{R}_{\mathbf{y}_k}^{(t)})^{-1}\mathbf{W}^{(t)}\mathbf{R}_{\mathbf{x}_k} \\ - \frac{p}{\max_{i\in\mathbf{n}_{mask}} \|\mathbf{Y}_k^{(t)}(:,i)\|_1}\mathrm{sign}\{\mathbf{Y}_k^{(t)}(:,l^{(t)})\}\tilde{\mathbf{X}}_k(:,l^{(t)})^T) \tag{5.1}$$

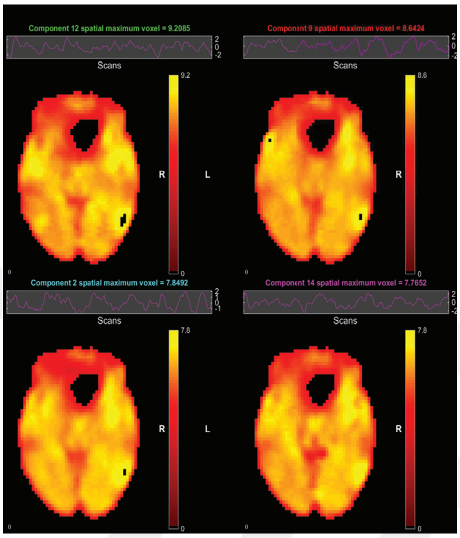
where

- $\mathbf{Y}_k^{(t)} = \mathbf{W}_k^{(t)}\tilde{\mathbf{X}}_k$ is the separator output vector for the k^{th} slice at iteration t.
- $\mathbf{R}_{\mathbf{x}_k} = \frac{1}{F}\tilde{\mathbf{X}}_k\tilde{\mathbf{X}}_k^T - \mu_{\mathbf{x}_k}\mu_{\mathbf{x}_k}^T$ is the sample covariance for mixtures, with $\mu_{\mathbf{x}_k}$ defined as the sample mean.
- $\mathbf{R}_{\mathbf{y}_k}^{(t)} = \mathbf{W}_k^{(t)}\mathbf{R}_{\mathbf{x}_k}\mathbf{W}_k^{(t)^T}$ is the sample covariance for the separator outputs at iteration t.
- $l^{(t)}$ is the voxel index corresponding to the column of $\mathbf{Y}_k^{(t)}$ with the maximum ℓ_1-norm. If there are multiple indices for which maximum ℓ_1-norm separator output is achieved, only one arbitrary index location is selected.
- $\beta^{(t)}$ is the algorithm step size at iteration t.

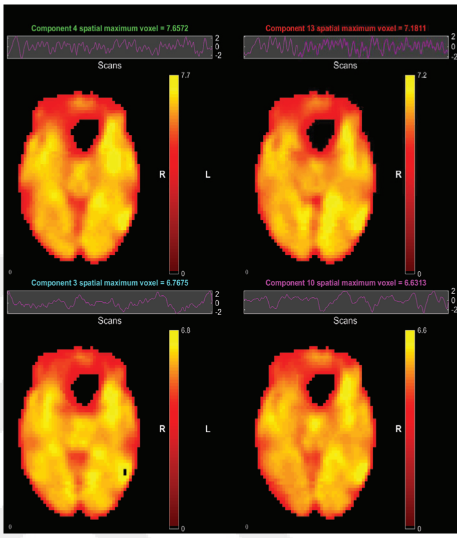
The masked fMRI data in Fig. 5.1 are inserted to SBCA algorithm and the component maps in Fig. 5.2 are obtained. Note that source number is taken to be 15 for this trial as a priori information obtained with GIFT software [Lab,].

5.1.3 Image Overlay Phase

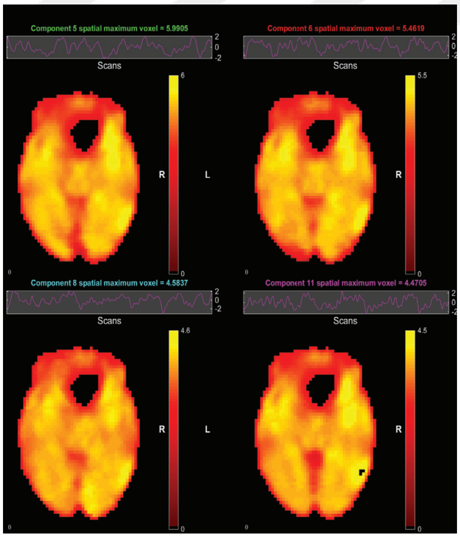
To implement component map visually better, we form overlaid component maps. For this, we need a baseline image referred as structural image in fMRI analysis. Again we use the fMRI data at the first time point as the structural image, i.e., $\mathbf{SI} = \mathbf{X}_k(1,:)$. In order to put the components on top of the structural image, we



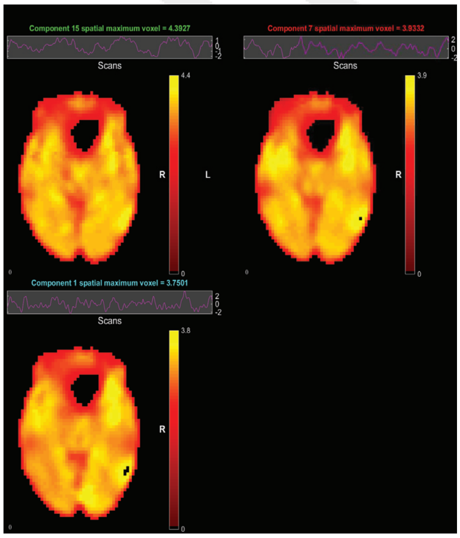
(a) Components 1-4



(b) Components 5-8



(c) Components 9-12



(d) Components 13-15

Figure 5.2: Extracted Component Maps and Associated Time Courses for the Masked Visuomotor fMRI record

scale the component images as follows:

$$c = \max \mathbf{Y}_k(i,:) / \left(\max \mathbf{Y}_k(i,:) - \min \mathbf{Y}_k(i,:)\right)$$
$$\hat{\mathbf{Y}}_k(i,:) = Bc \left(\mathbf{Y}_k(i,:) / \max \mathbf{Y}_k(i,:) + \min \mathbf{Y}_k(i,:) / \max \mathbf{Y}_k(i,:)\right) \quad (5.2)$$

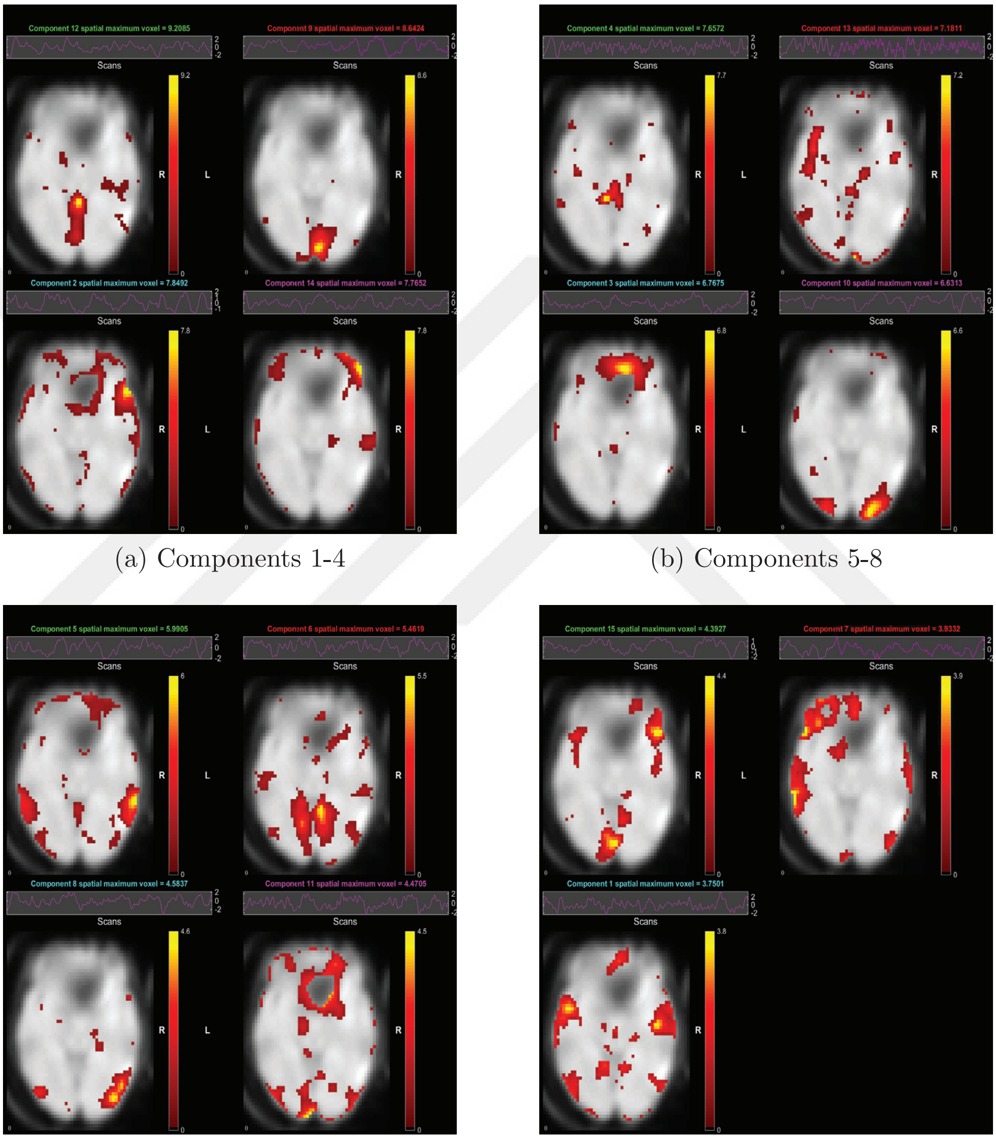
where B represents maximum of scaling interval that each component and the structural image are scaled between. Similarly, the structural image is scaled as

$$\mathbf{SI}_n = B((\mathbf{SI} - \min \mathbf{SI})/(\max \mathbf{SI} - \min \mathbf{SI})) + (\alpha - 1)B \quad (5.3)$$

where α is a positive integer in order to scale the structural image in a different interval. After scaling the component and structural images, the component images are overlaid on the structural image using (5.4).

$$\mathbf{SI}_{n,i} = \mathbf{SI}_n$$
$$\mathbf{SI}_{n,i}(\mathbf{n}_{mask}) = \hat{\mathbf{Y}}_k(i,:) \quad (5.4)$$

Finally, the detected component images overlaid on the structural image for one spatial slice of the fMRI data are illustrated in Fig. 5.3.



(a) Components 1-4

(b) Components 5-8

(c) Components 9-12

(d) Components 13-15

Figure 5.3: Overlaid Component Maps and Associated Time Courses for the Masked Visuo-motor fMRI record

5.2. Neural Activity Detection Using Calcium Images

Optical methods based on the activity of fluorescent molecules of neural circuits in response to the binding of calcium (Ca^{2+}) ions have become a standard tool for functional neuro-imaging. In Ca^{2+} imaging, the activity of neurons is recorded through a digital camera. The Ca^{2+} record is denoted by a three dimensional array $\mathbf{T}[m, n, f]$ where $(m, n, f) \in \{1, \ldots, M\} \times \{1, \ldots, N\} \times \{1, \ldots, F\}$, m, n are pixel location indices, row and column respectively, and f is the frame index. In other words, there are F frames of size $M \times N$.

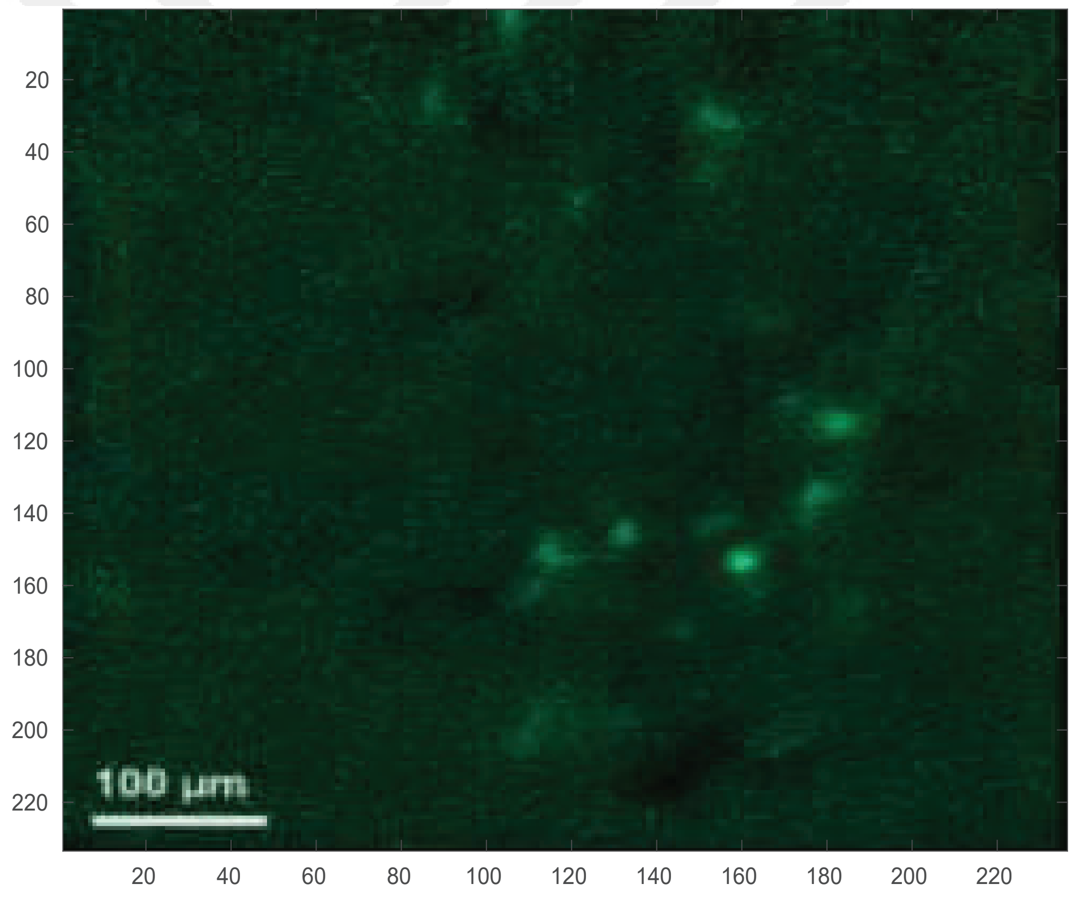


Figure 5.4: A frame from the Ca^{2+} movie of [Ziv et al., 2013] (used with the consent of Nature Publishing Group).

Fig.5.4 illustrates a frame in the movie corresponding to Ca^{2+} based hippocampal cell imaging study in [Ziv et al., 2013] (This movie file is available in https://www.nature.com/neuro/journal/v16/n3/full/nn.3329.html). In this figure, the activities of neurons are reflected as bright spots, while some neuron activities are buried under the background noise as well as the bright activities from the neighbouring neurons.

In order to predict behaviours from acquired Ca^{2+} imaging data, detection and identification processes that extract the individual neurons and their activities from the Ca^{2+} image recordings are needed. Unsupervised methods may be ideal for this purpose because they do not need training. In [Mukamel et al., 2009], Mukamel et al introduced an automated technique that combines spatio-temporal ICA and image segmentation for extracting active neuron locations. The technique relies on the sparsity of Ca^{2+} signals in both time and spatial domains. By using this attribute of Ca^{2+} signals, the algorithm decomposes data into statistically independent signals, each having a high sparsity. Then it examines these extracted signals to estimate the locations and activities of neurons. This neuron detection tool works well for image registration on small datasets. As another Ca^{2+} signal analysis tool, FluoroSSNAP (Fluorescence Single Neuron and Network Analysis Package) [Patel et al., 2015] utilizes Ca^{2+} imaging as the first step to identify active cells in an image frame, and then extracts a temporal fluorescence trace for the identified cells. After identification, it characterizes single-cell and network-level features of calcium dynamics. Another automatic neuron detection method proposed by Pachitariu et al in [Pachitariu et al., 2013] searches for the repeating motifs in accordance to a generative image model, and segments the detected 2-D images into their basis elements. These are the atoms in a dictionary which is used by a variant of convolutional matching pursuit for detection.

In this section, we propose a new neuron detection and identification tool along with the illustrations obtained from its application to the movie from [Ziv et al., 2013]. Our tool is completely automated and this can be considered as the main advantage against manual annotation techniques [Dombeck et al., 2007]. The corresponding algorithm consists of several consecutive phases and the following subsections describe them in the order of their execution.

5.2.1 Noise Normalization Phase

As a preprocessing step, background noise is approximated and time frames are normalized with this term. For this purpose, we apply vectorization to each frame by stacking columns through the vec(·) operator:

$$\mathbf{X} = \left[\; \mathrm{vec}(\mathbf{T}[:,:,1]), \ldots, \mathrm{vec}(\mathbf{T}[:,:,F]) \; \right], \tag{5.5}$$

where $\mathbf{X} \in \mathbb{R}^{MN \times F}$. The mean value for the time frames is computed as

$$\hat{\mathbf{m}} = \frac{1}{F} \sum_{f=1}^{F} \mathbf{X}(:, f), \tag{5.6}$$

where $\hat{\mathbf{m}}$ is an $MN \times 1$ vector. As the second step, the vectorized noise map is computed as

$$\mathbf{I}_n = \hat{\mathbf{m}} + median(\hat{\mathbf{m}}). \tag{5.7}$$

Then, the normalized time frames for $2D$ image are obtained as

$$\mathbf{T}[:,:,f] = \mathbf{T}[:,:,f] \oslash \left[\mathbf{I}_n(1:M), \ldots, \mathbf{I}_n(M(N-1)+1:MN) \right], \tag{5.8}$$

where $\oslash$ represents Hadamard division, i.e., element-wise division.

5.2.2 Subwindowing Phase

In the second phase of the method, the input movie is divided into multiple sub-movies by extracting sub-windows of size $L \times L$. The sub-windows are obtained by shifting an $L \times L$ window by integer multiples of S pixels in horizontal and vertical directions. Here, S is chosen to be smaller than L such that the corresponding sub-windows have overlapping regions. Therefore, we can represent sub-movies obtained through this procedure as

$$T_{(j,k)}[m, n, f] = T[(j-1)S + m, (k-1)S + n, f], \tag{5.9}$$

where $(j,k) \in \{1,\ldots,\lfloor\frac{M-L}{S}\rfloor+1\} \times \{1,\ldots,\lfloor\frac{N-L}{S}\rfloor+1\}$, and $m,n \in \{1,\ldots,L\}$.

The main rationale behind this procedure is to reduce the number of active neurons in a given sub-movie so that number of frames is sufficient for SBCA algorithm to extract the active neurons in the next step. It also enables to reduce computational load and to process different sub-movies concurrently. The overlapping structure is designated to alleviate the algorithmic issues caused by the border-shared neurons of two non-overlapping windows.

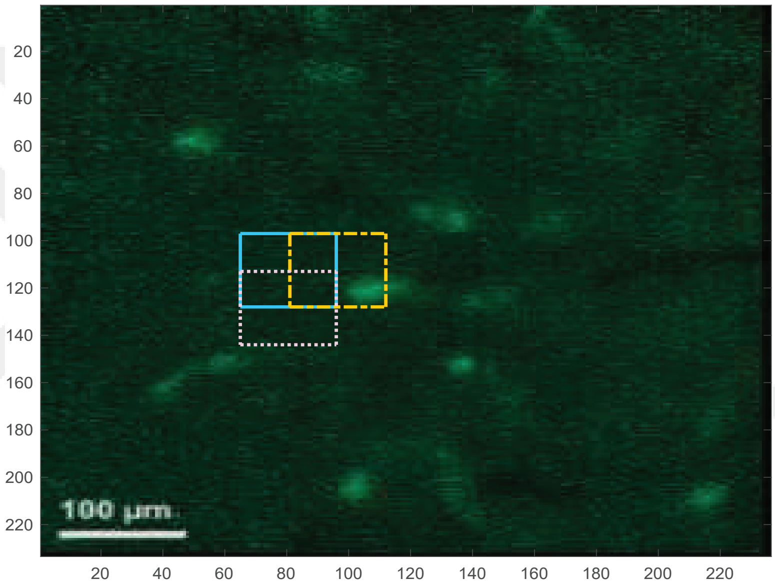


Figure 5.5: Overlapping sub-windows for a frame from the Ca^{2+} image movie of [Ziv et al., 2013] (Used by the permission of Nature Publishing Group).

For the movie in [Ziv et al., 2013], Fig. 5.5 illustrates three overlapping sub-windows, where window size is chosen as $L = 32$ and the shift size $S = 16$.

5.2.3 Source Separation Phase

In this phase, we apply the SBCA algorithm to each sub-movie separately. For this purpose, for each sub-movie $\mathbf{T}_{(j,k)}$, we apply vectorization to frames by stacking columns through the vec$(\cdot)$ operator

$$\mathbf{X}_{(j,k)} = \left[\text{ vec}(\mathbf{T}_{(j,k)}[:,:,1]), \ldots, \text{vec}(\mathbf{T}_{(j,k)}[:,:,F]) \right], \tag{5.10}$$

where $\mathbf{X}_{(j,k)} \in \mathbb{R}^{L^2 \times F}$. We assume that there are p components generating each sub-movie, which can be linearly separated using $\mathbf{Y}_{(j,k)} = \mathbf{W}_{(j,k)}\mathbf{X}_{(j,k)}$ where $\mathbf{W} \in \mathbb{R}^{p \times L^2}$ is the separator matrix, and l^{th}-row of $\mathbf{Y}_{(j,k)}$ denotes the l^{th} component's temporal activity in the regarded sub-movie.

After vectorization, the SBCA algorithm iterations are applied to each sub-movie data separately to obtain the separator matrix $\mathbf{W}_{(j,k)}$:

$$\begin{aligned}\mathbf{W}^{(t+1)}_{(j,k)} &= \mathbf{W}^{(t)}_{(j,k)} + \sigma^{(t)}((\mathbf{R}^{(t)}_{\mathbf{y}})^{-1}\mathbf{W}^{(t)}_{(j,k)}\mathbf{R}_{\mathbf{x}} \\ &- \frac{p}{\max\limits_{l \in \{1,\ldots,F\}} \|\mathbf{Y}^{(t)}[:,l]\|_1}\text{sign}\{\mathbf{Y}^{(t)}[:,i^{(t)}]\}\mathbf{X}_{(j,k)}[:,i^{(t)}]^T),\end{aligned} \tag{5.11}$$

where

- $\mathbf{Y}^{(t)}_{(j,k)} = \mathbf{W}^{(t)}_{(j,k)}\mathbf{X}_{(j,k)}$ denotes the separator output vector at iteration t.
- $\mathbf{R}_{\mathbf{x}} = \frac{1}{F}\mathbf{X}_{(j,k)}\mathbf{X}^T_{(j,k)} - \mu_{\mathbf{x}}\mu^T_{\mathbf{x}}$ denotes the sample covariance matrix for mixtures, with $\mu_{\mathbf{x}} = \frac{1}{F}\mathbf{X}_{(j,k)}\mathbf{1}$ defined as the sample mean vector.
- $\mathbf{R}^{(t)}_{\mathbf{y}} = \mathbf{W}^{(t)}_{(j,k)}\mathbf{R}_{\mathbf{x}}{\mathbf{W}^{(t)}_{(j,k)}}^T$ denotes the sample covariance for the separator outputs at iteration t.
- $i^{(t)}$ is the frame index corresponding to the column of $\mathbf{Y}^{(t)}_{(j,k)}$ with the maximum ℓ_1-norm. If maximum ℓ_1-norm separator output corresponds to more than one index location, only one arbitrary index is chosen.
- $\sigma^{(t)}$ is the algorithm step size at iteration t.

Once the source separation has been completed, the component images are obtained using Moore-Penrose pseudo-inverse:

$$\mathbf{Q}_{(j,k)} = \mathbf{W}^{\dagger}_{(j,k)}, \tag{5.12}$$

where the l^{th} column of $\mathbf{Q}_{(j,k)}$ corresponds to the component l. The image for that component can be generated by converting this $L^2 \times 1$ column vector into $L \times L$

matrix as the inverse operation of the vec($\cdot$) operator.

5.2.4 Background Estimation Phase

Neuropil signals cause a background activity in Ca^{2+} imaging that can be modelled as a low-rank matrix. Removing this activity provides enhanced spatial landmarks so that component identification can be executed more accurately. Let us model the observed data matrix for sub-movie (j, k) as

$$\mathbf{X}_{(j,k)} = \mathbf{Q}_{(j,k)}\mathbf{Y}_{(j,k)} + \mathbf{B}_{(j,k)}\mathbf{F}_{(j,k)} + \mathbf{E}_{(j,k)} \tag{5.13}$$

where $\mathbf{Q}_{(j,k)} \in \mathbb{R}^{L^2 \times p}$ and $\mathbf{Y}_{(j,k)} \in \mathbb{R}^{p \times F}$ are the spatial images and temporal activities of the detected components, $\mathbf{B}_{(j,k)} \in \mathbb{R}^{L^2 \times n_b}$ and $\mathbf{F}_{(j,k)} \in \mathbb{R}^{n_b \times F}$ are the spatial images and temporal activities of the regarded sub-movie's background. An approximate solution for background is

$$\hat{\mathbf{B}}_{(j,k)} = (\mathbf{X}_{(j,k)} - \mathbf{Q}_{(j,k)}\mathbf{Y}_{(j,k)})\hat{\mathbf{F}}^T_{(j,k)}(\hat{\mathbf{F}}_{(j,k)}\hat{\mathbf{F}}^T_{(j,k)})^{-1}, \tag{5.14}$$

$$\hat{\mathbf{F}}_{(j,k)} = (\hat{\mathbf{B}}^T_{(j,k)}\hat{\mathbf{B}}_{(j,k)})^{-1}\hat{\mathbf{B}}^T_{(j,k)}(\mathbf{X}_{(j,k)} - \mathbf{Q}_{(j,k)}\mathbf{Y}_{(j,k)}). \tag{5.15}$$

The spatial image and temporal activity of background are estimated using (5.14) and (5.15) iteratively. The estimated background is used for correlation computation in the component identification phase.

5.2.5 Component Denoising Phase

Before classification of extracted components, we denoise both component images and activities. The usefulness of denoising process can be explained using Fig. 5.6. Fig. 5.6 illustrates the component images of the first five SBCA components obtained for a sub-movie of the Ca^{2+} imaging example in the first row, and their activities in the second row. Based on this figure, we can judge that the SBCA components $2, 3$ and 5 correspond to real neurons and their activities. The activity

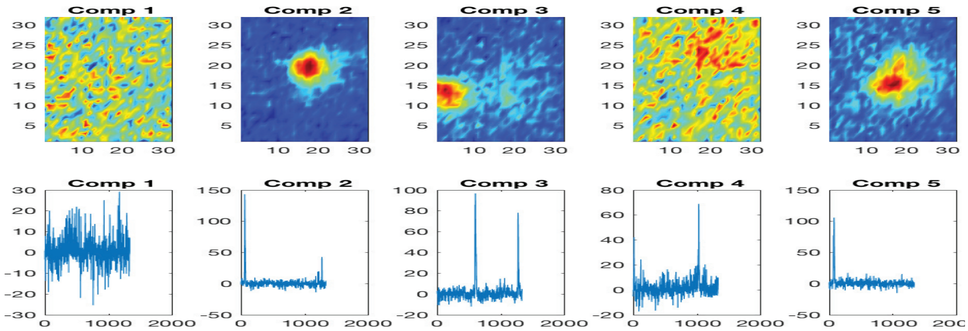


Figure 5.6: The images of the first five SBCA components (first row) and their activities (second row).

of the component 4 is spiky while its image does not seem like a neuron. It has been tested that denoising process improves particularly the accuracy of component registration by clearing such neurons as the component 4. On the other hand, the image of the component 1 is relatively dispersed and its activity is less spiky so that it is almost definitely not a neuron.

For component images, we use total variation based denoising [Chambolle et al., 2010] by applying the Matlab function provided in [Lourakis, 2016]. Let $\mathbf{I}_l$ represent $2D$ component image that corresponds to the l^{th} column of $\mathbf{Q}_{(j,k)}$, and $\mathbf{I}_{l,d}$ represent its denoised version. Total variation based denoising cost function can be written as

$$J_{TV}(\mathbf{I}_{l,d}) = \|\nabla \mathbf{I}_{l,d}\|_{2,1} + \|\mathbf{I}_l - \mathbf{I}_{l,d}\|_F, \tag{5.16}$$

where

$$\|\nabla \mathbf{I}_{l,d}\|_{2,1} = \sum_{i,j} \sqrt{|\nabla \mathbf{I}_{l,d}^{(x)}[i,j]|^2 + |\nabla \mathbf{I}_{l,d}^{(y)}[i,j]|^2}. \tag{5.17}$$

$\nabla \mathbf{I}_{l,d}^{(x)}$ ($\nabla \mathbf{I}_{l,d}^{(y)}$) is the image gradient component in the x (y) direction, obtained by evaluating the difference between the pixel value at the next row (column) and the current position, $\|\|_F$ is Frobenius norm.

In order to denoise component activities, we use wavelet denoising [Donoho, 1995]. For the parameters of the denoising operation, we choose $Symlet-4$ wavelet, soft thresholding, 4^{th} level of wavelet decomposition and rescaling based on level noise estimation. These parameters are mostly tuned by empirical experiments. Fig. 5.7 shows the denoised versions of the component images and activities in Fig.

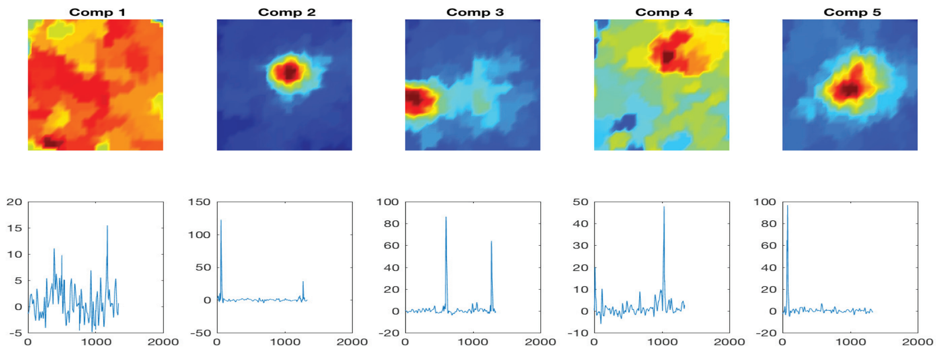


Figure 5.7: Denoised component images (first row) and their activities (second row).

5.6. All the component images except that of component 1 look more like neuron-shaped images with spiky activities after denoising. The denoised component images and the related denoised activities are used as input to the component registration phase.

5.2.6 Component Identification Phase

For the task of identifying SBCA components as “neuron ”or “not-neuron ”, we employ correlation method. Each component's image and temporal activity are correlated with the average frame of the regarded sub-movie. The average frame is computed over the interval when only the regarded component was active. Let $\mathbf{Q}_{(j,k)}[:,l]$, $\mathbf{Y}_{(j,k)}[l,:]$ denote the vectorized image and activity of the component l in the sub-movie (j,k). And let $\tilde{\mathbf{Q}}_{(j,k)}$, $\tilde{\mathbf{Y}}_{(j,k)}$ denote the spatial and temporal matrices when the component l has been removed. Then, the sub-movie dataset for the component l is computed by removing the background and all the components except l:

$$\mathbf{X}_{(j,k),l} = \mathbf{X}_{(j,k)} - \tilde{\mathbf{Q}}_{(j,k)}\tilde{\mathbf{Y}}_{(j,k)} - \mathbf{B}_{(j,k)}\mathbf{F}_{(j,k)}. \tag{5.18}$$

If the component l is a real neural component, $\mathbf{X}_{(j,k),l}$ and $\mathbf{Q}_{(j,k)}[:,l]\mathbf{Y}_{(j,k)}[l,:]$ must look similar at time intervals when the component l is active. To decide the similarity, we compute spatial and temporal correlations. To compute spatial correlation, we average $\mathbf{X}_{(j,k),l}$ across time over the intervals when the first N_p local peaks of $\mathbf{Y}_{(j,k)}[l,:]$ occur, and obtain the average image $\hat{\mathbf{I}}_l$. The spatial correlation between $\hat{\mathbf{I}}_l$ and $\mathbf{Q}_{(j,k)}[:,l]$ is then computed as

$$\rho_{s,l} = \frac{|\langle \mathbf{Q}_{(j,k)}[:,l], \hat{\mathbf{I}}_l \rangle|}{\|\mathbf{Q}_{(j,k)}[:,l]\|_2 \|\hat{\mathbf{I}}_l\|_2}. \tag{5.19}$$

To compute temporal correlation, we average $\mathbf{X}_{(j,k),l}$ across space, and obtain the average temporal activity $\hat{\mathbf{c}}_l$. The temporal correlation between $\hat{\mathbf{c}}_l$ and $\mathbf{Y}_{(j,k)}[l,:]$ over the time intervals of the first N_p local peaks is then computed as

$$\rho_{t,l} = \frac{|\langle \mathbf{Y}_{(j,k)}[l,:], \hat{\mathbf{c}}_l \rangle|}{\|\mathbf{Y}_{(j,k)}[l,:]\|_2 \|\hat{\mathbf{c}}_l\|_2}. \tag{5.20}$$

The component l is accepted to be a neural component if the spatial and temporal correlation coefficients are above pre-defined threshold values.

5.2.7 Component Registration Phase

This phase aims to combine the identified components that correspond to the same neuron in a single sub-movie and in neighbouring sub-movies. For this registration problem, we compare the spatial footprints of the pairs of components for a single sub-movie, and compare the activities of the pairs of components for neighbouring sub-movies. Comparison is performed through the calculation of the correlation coefficient. Let $\mathbf{Y}_{d,(j,k)}$ represent the denoised activities of the identified components for sub-movie (j,k). We define the temporal correlation coefficient between the identified component l in the sub-movie (j,k) and the identified component m in the sub-movie $(j, k+1)$ as

$$\rho t_{(j,k),l}^{(j,k+1),m} = \frac{|\langle \mathbf{Y}_{d,(j,k)}[l,:], \mathbf{Y}_{d,(j,k+1)}[m,:]\rangle|}{\|\mathbf{Y}_{d,(j,k)}[l,:]\|_2 \|\mathbf{Y}_{d,(j,k+1)}[m,:]\|_2}. \tag{5.21}$$

To compute the spatial correlation coefficient between the components in a single sub-movie, we generate a binary image from the denoised image matrix $\mathbf{Q}_{d,(j,k)}$, by thresholding values below 90% of the peak value to zero, and the others to 1. In Fig. 5.8.(b), the binary image obtained by thresholding the component image in Fig. 5.8.(a) is shown. Each component's binary image is later dilated 2 times by a 3×3 star pattern to obtain a binary mask for neuron image as shown in Fig. 5.8.(c). Lastly, the denoised image is masked by the binary dilated image. Let $\mathbf{Q}_{n,(j,k)}$ represent the dilated versions of the denoised component images for sub-movie (j, k). Then, we define the spatial correlation coefficient for the identified component l and m in this sub-movie as

$$\rho s_{(j,k),l}^{(j,k),m} = \frac{|\langle \mathbf{Q}_{n,(j,k)}[:,l], \mathbf{Q}_{n,(j,k)}[:,m]\rangle|}{\|\mathbf{Q}_{n,(j,k)}[:,l]\|_2 \|\mathbf{Q}_{n,(j,k)}[:,m]\|_2}. \tag{5.22}$$

Especially, for small frame rates, the activity patterns of neurons could closely resemble each other. Therefore, in addition to the correlation of activities in neighbouring sub-movies, we also check the overlap between the neural component images. The overlap factor is obtained by taking inner products of the dilated binary images in the shared region between the neighbouring sub-movies.

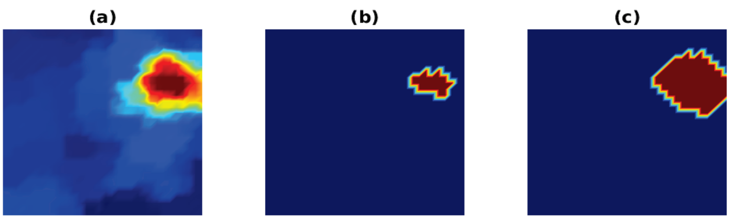


Figure 5.8: Morphological operations on the component image.

The correlation and overlap checking processes are iteratively completed starting from the top-left sub-movie. If the correlation coefficients are greater than a predetermined threshold τ_ρ, and the overlap factor is greater than another threshold τ_o, then the corresponding component is removed from the considered sub-movie, and combined into the component in the current sub-movie. The combination is achieved by averaging aligned neuron images and averaging their activities. This iterative procedure makes sure that component overlaps with neighbouring sub-movies and component replicas in current sub-movie have been removed.

5.2.8 Image Segmentation Phase

If the number of frames is limited, some neuron groups in the same sub-movie may fire together once, and therefore, they result in the same SBCA component. In order to discriminate these components, we need an extra processing on the registered component images. For this reason, we apply an image segmentation based on Otsu's method [Otsu, 1979]. The aim of this method is to correct non-uniform illumination in an image frame so that the different regions inside the frame can be identified better. First, we estimate the background illumination in the component image by morphological opening with a disk structuring element. Then we subtract

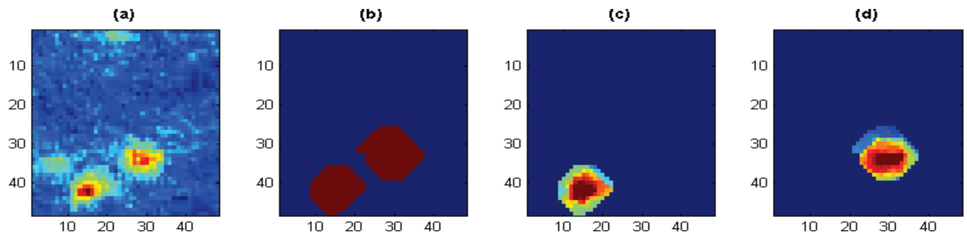


Figure 5.9: (a) Original component image (b) Adjusted and binarized component image (c-d) Separated neural components.

the estimated background image from the original image. As the second step, we binarize the adjusted image by thresholding, and remove noise by throwing away the objects smaller than a predefined size. Finally, we identify different neurons by grouping the connected objects in the binary image. As the final step, the component images are masked by the binary dilated images. An example of the segmentation process is illustrated in Fig. 5.9.

Fig.5.10 illustrates all the identified neural components (magenta contours) for the full movie recording from [Ziv et al., 2013]. The sub-window size was taken to be $L = 12$, and the overlap size $S = 6$. The background image was obtained by averaging the movie frames.

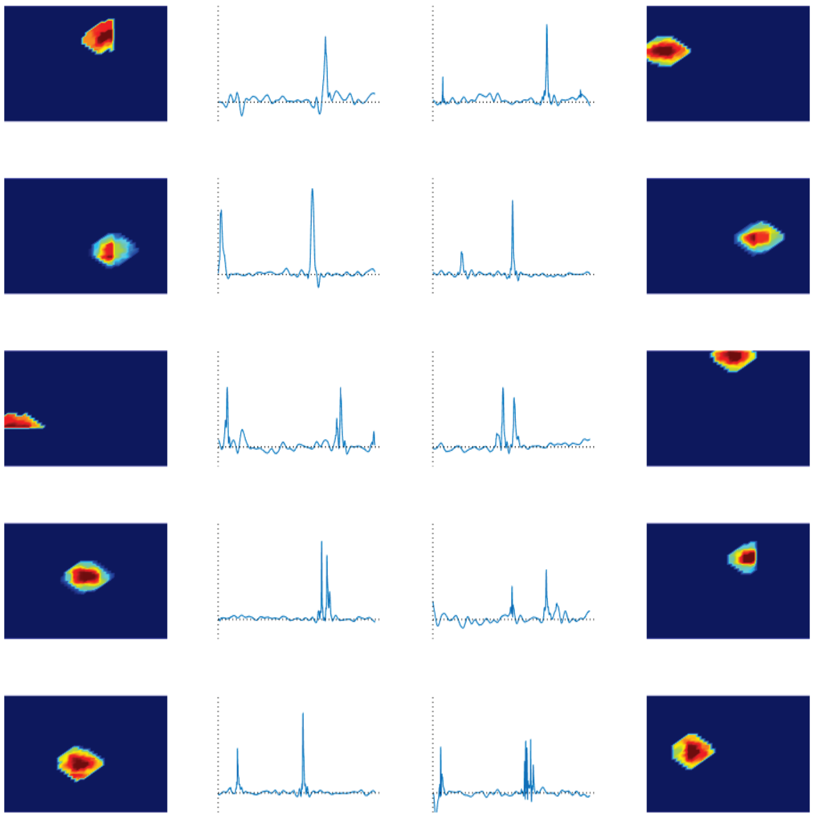


Figure 5.10: The component maps and the corresponding activities obtained at the final phase of the algorithm.

5.2.9 Evaluation of Performance

To quantify performance, we first extract temporal activities of ground truth components since labelled datasets include only ground truth spatial masks. To obtain the approximate ground truth activities, we insert binary masks and spatial centres of ground truth components to CAIMAN (CAlcium IMaging ANalysis) tool [Giovannucci et al., 2019] which is an open-source library for calcium imaging data analysis.

After extracting the temporal activities of ground truth components, performance quantification can be put forth with different parameters. Among all, most used ones are accuracy, precision, recall and F_1 score.

Precision: It quantifies the certainty that a detected component corresponds to a ground truth component.

$$precision = \frac{|\{l : \mathbf{I}_l \in \mathcal{I}_{gt}\}|}{|\mathcal{I}_{dt}|} \tag{5.23}$$

where $\mathbf{I}_l$ is the vectorized image of a detected component, $\mathcal{I}_{dt}$ and $\mathcal{I}_{gt}$ denote the sets of detected and ground truth components respectively.

Recall: It quantifies the certainty that a ground truth component has been detected.

$$recall = \frac{|\{l : \mathbf{I}_l \in \mathcal{I}_{gt}\}|}{|\mathcal{I}_{gt}|} \tag{5.24}$$

$\mathbf{F_1}$ **score**: It is harmonic mean between precision and recall parameters.

$$F_1 = 2 \times \frac{precision \times recall}{precision + recall} \tag{5.25}$$

5.2.10 Experimental Results

Ziv's Movie:

This experiment illustrates the performance of our neuron sorting tool using the movie frames in [Ziv et al., 2013]. In this experiment, we consider the movie frames (size of 233×236) in [Ziv et al., 2013]. The selected parameters for this analysis are given in Table 5.1.

The background image obtained from batches by taking the maximums is depicted in Fig. 5.11.

The registered neurons are denoted by contours in Fig. 5.12. It seems that some neuron regions have not been detected. Researches have concluded that there are two main reasons for such failures: 1) high background noise in the corresponding

$Sub - windowSize(L)$	12
$\rho_{s,thr}$	0.4
$\rho_{t,thr}$	0.4
ρs_{thr}	0.5
ρt_{thr}	0.9
N_{sbca}	4

Table 5.1: Input Parameters for SBCA Analysis Tool

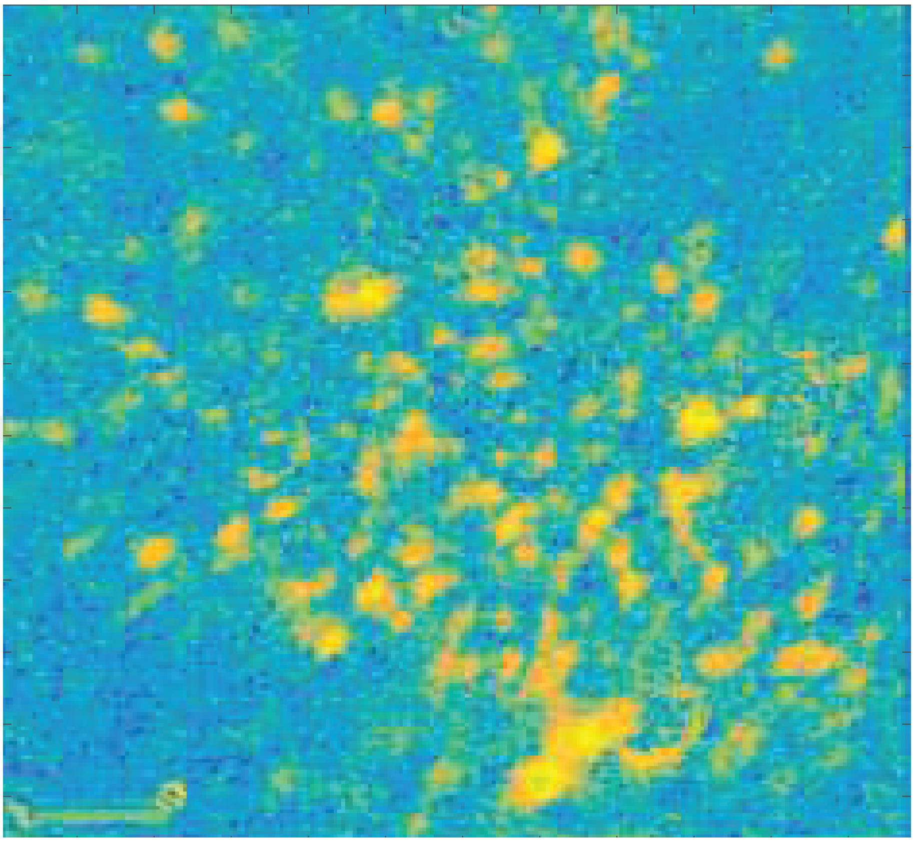


Figure 5.11: Background image for the full hippocampal movie from [Ziv et al., 2013].

sub-windows, 2) high detection threshold value assigned by the adaptive threshold algorithm. It was concluded that it would be better not to detect these neural components, since a more powerful denoising increases the amount of miss detection, and decreasing the threshold values increases the amount of false detection.

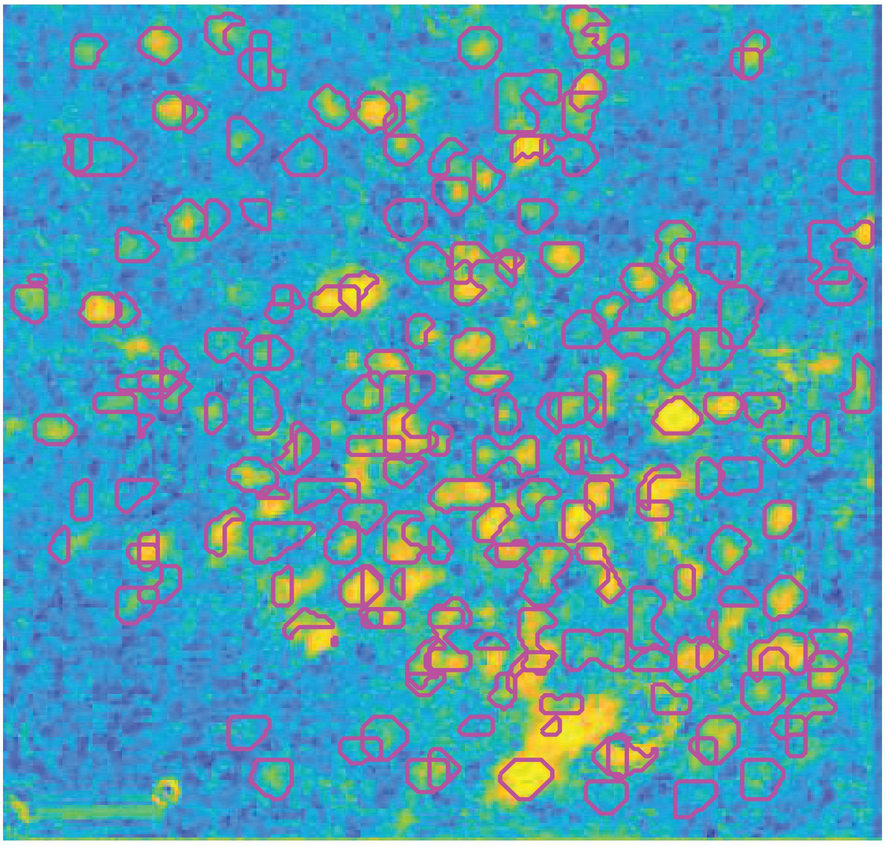


Figure 5.12: Detected neurons (magenta contours) by SBCA for the full movie from [Ziv et al., 2013].

Neurofinder Movies:

In the first part of this experiment, we compare SBCA and CAIMAN [Giovannucci et al., 2019] methods by using recorded data of two neurons' activities from [Freeman et al.,]. The shape of these neurons is called donut in literature, which means that neurons appear as ring shapes because proteins that make up calcium activity are confined outside the cell nucleus. The movie is of 48×20 size. The background image was generated with the correlation image generation module of CAIMAN tool [Giovannucci et al., 2019]. We used CAIMAN tool with the parameters suggested in the demo codes of [Giovannucci et al., 2019], except tuning the minimum SNR threshold value in the interval $[0 - 4]$ dB, and the space threshold value in the interval $[0.1-0.5]$. The selected SBCA parameters in this part are given in Table 5.2.

$Sub-windowSize(L)$	16
$\rho_{s,thr}$	0.3
$\rho_{t,thr}$	0.3
ρs_{thr}	0.5
ρt_{thr}	0.9
N_{sbca}	2

Table 5.2: Input Parameters for SBCA Analysis Tool

As can be seen in Fig. 5.13, SBCA captured the spatial footprints of the two neurons better than CAIMAN did while both the methods could resolve the closely located neurons.

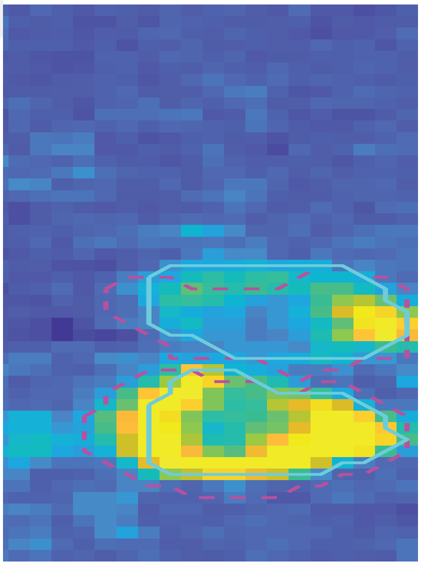
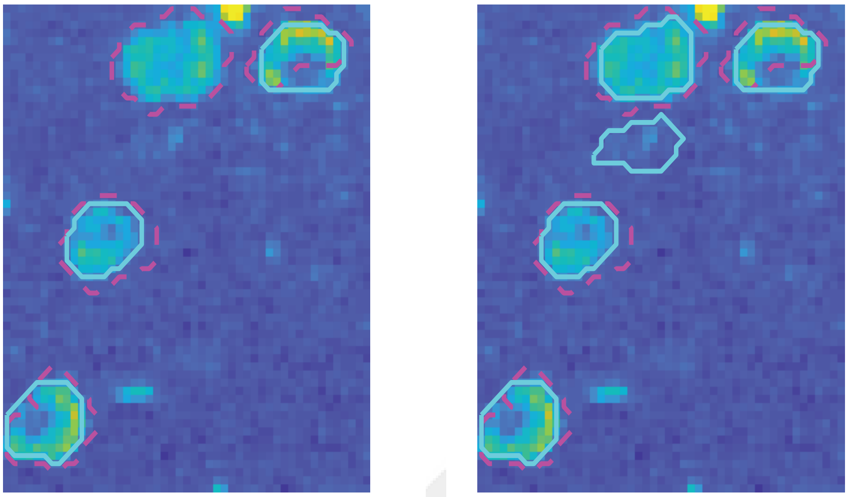


Figure 5.13: Resolving the images of two donut-shaped neurons in a 48×20 size image. Cyan: Components detected by CAIMAN, Dashed Magenta: Components detected by SBCA.

The second part of the experiment shows the detection ability for 4 spatially apart neurons. The movie is of 60×48 size. The selected SBCA parameters in this part are given in Table 5.2 except that sub-window size was taken to be 32. The



(a) (b)

Figure 5.14: Detection of 4 neurons in a 60×48 size image. Cyan: Components detected by CAIMAN, Dashed Magenta: Components detected by SBCA. a)-Min. SNR threshold for CAIMAN = 2dB, b)- Min. SNR threshold for CAIMAN = 4dB

minimum SNR threshold for CAIMAN was tuned in the interval $[0-4]$ dB, and the space threshold was tuned in the interval $[0.1-0.5]$. Decreasing the minimum SNR threshold for CAIMAN caused false detections and increasing it caused misdetections as illustrated in Fig. 5.14.

In the third part of the experiment, we compare the two methods against ground truth by using Neurofinder collection of Ca^{2+} imaging datasets [Freeman et al.,]. Table 5.3 shows the properties of the analysed datasets.

Name	Brain Area	Size	Imaging rate	# GT components
NF.01.00	Visual cortex	$512 \times 512 \times 2500$	7	345
NF.01.01	Visual cortex	$512 \times 512 \times 1825$	7	350
NF.02.01	Cortex	$512 \times 512 \times 2500$	30	178

Table 5.3: Properties of Neurofinder datasets. For each dataset brain area, size, imaging rate and number of ground truth components are given.

The datasets provide binary masks per ground truth components. We produced the ground truth spatial centres using these binary masks, and then run the seeded initialization procedure of CAIMAN tool [Giovannucci et al., 2019] to construct the ground truth spatial and temporal components restricted to the areas determined by manual annotation. After the spatial and temporal components were estimated by SBCA and CAIMAN methods, the estimations were registered against the set of ground truth components. For registration, spatial overlap and temporal correlation values between the detected and ground truth components were calculated and the components, whose temporal correlation and spatial overlap both exceeded 0.5, were accepted to be matched. Lastly, the performance of each method was quantified using these component pairs and the precision/recall framework defined in section 5.2.9. Note that the authors of CAIMAN generated their own manually annotated ground truth (they have not yet published this annotation work) and computed only spatial overlap for registration in [Giovannucci et al., 2019].

The selected SBCA parameters in this part are given in Table 5.2 except that sub-window size and number of sources per sub-window were taken to be 24 and 4 respectively. We again used CAIMAN tool with the parameters suggested in the demo codes of [Giovannucci et al., 2019], except decreasing the minimum SNR threshold in order to increase amount of detections (only 50 detections were extracted for dataset 01.00 when using the original parameter values). The minimum SNR threshold value was tuned in $[0 - 4]$ dB, and the space threshold value was tuned in $[0.1 - 0.5]$. The best performance was obtained for minimum SNR threshold=0.5 dB and space threshold=0.5 (original value). We had the maximum number of CNMF detections as 500 for all the datasets. This is fair for our performance comparison because CAIMAN method has a post-processing step for eliminating false detections and because we assumed that there were 4 sources in each of 1764 sub-movies (24×24) for SBCA which means that SBCA algorithm extracted 7056 sources in source separation phase.

Fig. 5.15 and Fig. 5.16 illustrate results for the Neurofinder movies 01.00 and 02.01. Yellow spots are the binarized spatial footprints of the ground truth components, cyan contours are the detected components by CAIMAN and magenta contours are the detected components by SBCA. As seen in these figures, both

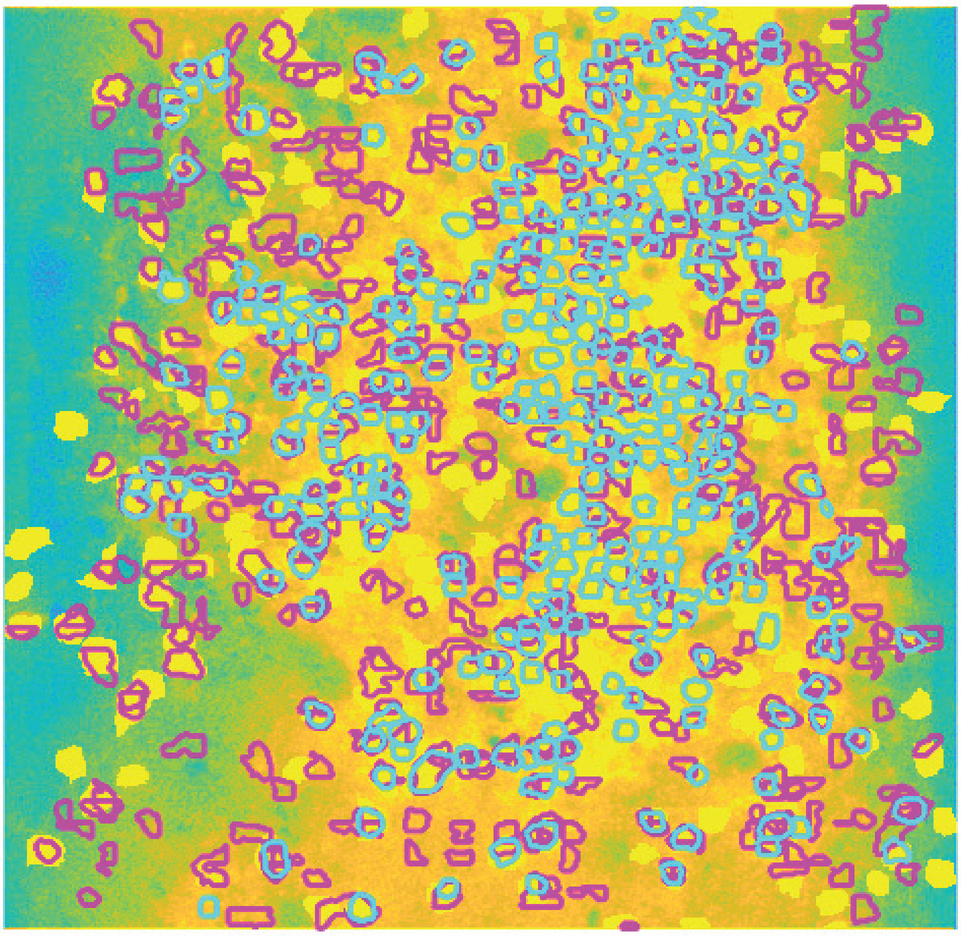


Figure 5.15: Sorted neurons for Neurofinder movie 01.00. Yellow: Ground truth components, Cyan: Components detected by CAIMAN, Magenta: Components detected by SBCA.

methods could not detect some ground truth components, but detected some components which were not labelled as ground truth. These unlabelled components may be related to neuropil activities or the neural activities that could not be manually annotated. CAIMAN seems to have failed to detect the ground truth components with weak activity signals (the components far from intensive activity regions) while SBCA achieves this issue better. On the other hand, it can be concluded that CAIMAN is more successful in the regions where active neurons are close to each other. In fact, the datasets 01.00 and 02.01 initially contained 3024 and 8000 frames. Since CAIMAN could not process data of these sizes even though using a powerful

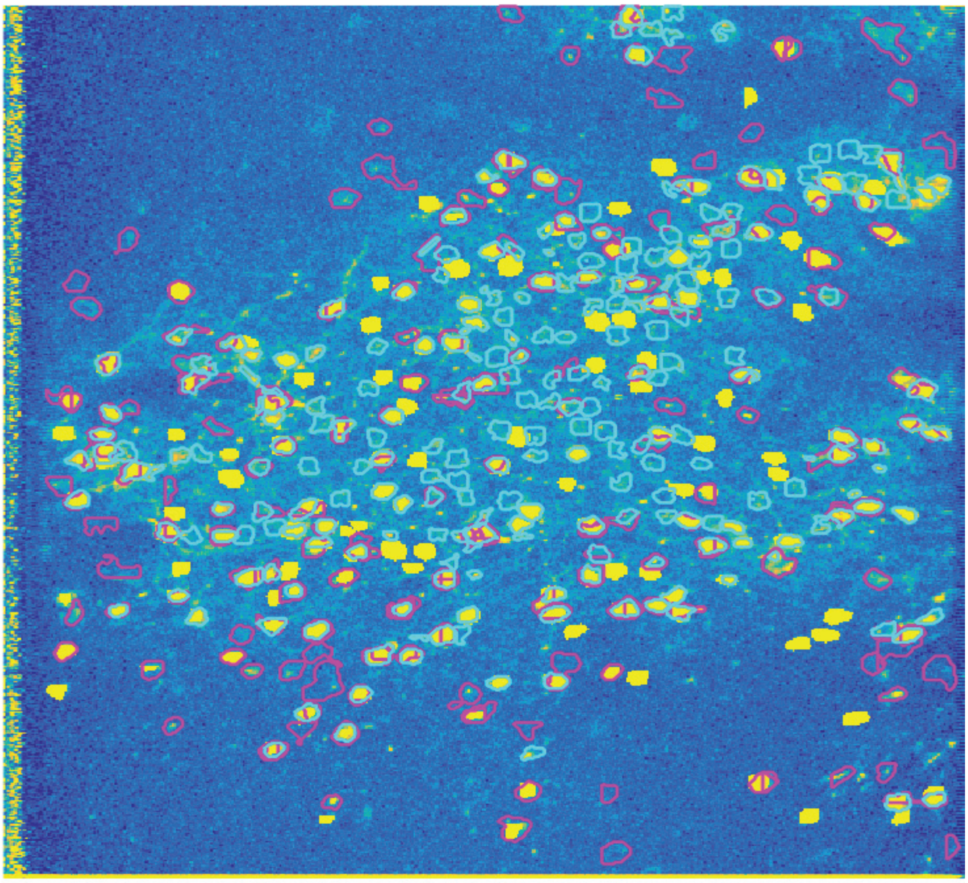


Figure 5.16: Sorted neurons for Neurofinder movie 02.01. Yellow: Ground truth components, Cyan: Components detected by CAIMAN, Magenta: Components detected by SBCA.

processor (Intel Xeon CPU E5-2650 @ 2.00 GHz (2 processors), 64 GB RAM, 64-bit Operating System), we had to reduce them to 2500 frames. This can be interpreted as a weakness of CAIMAN method with respect to processing load.

Table 5.4 shows the performance quantifications for the methods, and it results that SBCA and CAIMAN yield close detection and identification performance. We note that the authors achieved a higher performance ($F_1 > 0.7$) for many different datasets in [Giovannucci et al., 2019] by generating their own manually annotated ground truth and computing only spatial overlap for component registration. Also, we note that we extracted the spatial footprints and temporal traces of the ground truth components using a module of CAIMAN tool, and this is probably an advantage to CAIMAN method to better match the ground truth components. Beside it, CAIMAN method uses correlation and Convolutional Neural Network (CNN) methods together for neuron identification while our sorting method uses only a correlation based identification. We are also working on using a neural network classifier and expect it to improve our method's identification performance. Consequently, the close performance of the two methods and processing load issue make our SBCA based method an attractive alternative for neuron sorting applications.

Name	**CAIMAN BATCH**			**SBCA**		
	Precision	Recall	F_1	Precision	Recall	F_1
NF.01.00	0.4437	0.5507	0.4914	0.4520	0.6058	0.5177
NF.01.01	0.4797	0.5629	0.5180	0.4987	0.5314	0.5145
NF.02.01	0.5692	0.6067	0.5874	0.5576	0.5030	0.5289

Table 5.4: Results of CAIMAN BATCH and SBCA algorithms against ground truth components in the form precision, recall, F_1 score.

Chapter 6

CONCLUSIONS AND FUTURE WORK

6.1. Conclusions

In this dissertation, we have presented a new class of Blind Source Separation algorithms and their applications on some practical problems.

In chapter 2, we summarize the BCA framework introduced by Erdogan in [Erdogan, 2013] and propose a new approach based on this framework to solve the low altitude target's angle estimation problem. BCA acts as a preprocessing step, prior to the main DOA estimation method, to extract direct-path echo from total received signal (direct-path + reflected-path echoes) so that it decreases the multipath effect on DOA estimation. Because the framework is completely deterministic, it does not require any information about propagation path, detected object and multipath model. The numerical examples show improvement over CMLM, particularly at far ranges where echo fading effect is high. Furthermore, the practical relative merit of the proposed framework over BDML method is demonstrated in the numerical examples, especially for small amounts of samples.

In chapter 3, we introduce a novel BCA framework for sparse bounded sources and derive an iterative algorithm to estimate separator matrix based on this framework. We propose an objective function to be maximized with respect to two geometric objects called principal hyper-ellipsoid and bounding ℓ_1 norm ball. The main assumptions of the algorithm are **sparsity** and **source boundedness** which enable the construction of the optimization setting. Since the denominator of the objective, i.e., volume of the bounding ℓ_1 norm ball, is the maximum ℓ_1 norm of the separator output samples, maximization of the objective means minimization of the maximum of ℓ_1 norm which is one of the most well known cost functions for measuring sparsity. Thus, our proposed method can be regarded as a Sparse Component Analysis method. There is **no constraint about the analysis domain**, it is possible to define sparse components in time, frequency or time-frequency domains. The mixing model is **instantaneous** and the dimensionality of the problem is restricted to be **(over)determined**. Numerical results show that SBCA outperforms some state of the art methods if sources are sparse or can be sparsified via transformation. Additionally, the image separation and the neuron detection examples exhibit SBCA's usability in practice.

In chapter 4, we propose novel time and frequency domain sparsity based Bounded Component Analysis approaches for convolutive mixtures. The framework is very similar with that in chapter 3, the most important difference is the convolutive nature of the filter responses. The proposed approaches do not assume statistical independence in space, time or frequency domain. That's why, we can draw the conclusion that the convolutive approaches introduced here are the natural extensions of the SBCA algorithm in chapter 3. We demonstrate two practical applications for speech signal separation and musical signal separation by using the frequency domain sparsity based CSBCA. The reason why we selected the frequency based method for the practical applications is that audio signals generally are not sparse in time domain. We illustrate the performance improvement compared to some benchmark convolutive BSS methods. We also note that the frequency domain sparsity based CSBCA has the advantages associated with the frequency domain approaches but does not have the inherently known permutation and scaling ambiguities of these approaches because it updates a temporal separator matrix instead of instantaneous separator matrices for different frequency bins.

In chapter 5, we propose a space-time analysis tool to track the non-stationary nature of the neural activities. This analysis tool uses Ca^{2+} imaging records as input data, detects activations based on the instantaneous SBCA algorithm, and identifies them as 'neuron/not neuron 'based on spatial/temporal correlation. It rather offers to capture temporally sparse active neurons so that it enables to analyse some additional neural network dynamics. At this point, the SBCA based analysis method can provide superiority over PCA/ICA based analysis methods since it basically utilizes the temporal sparsity property of neural activities instead of spatial/temporal independence of neurons.

6.2. Contribution of the Thesis

The major contributions of this thesis are summarized as follows:

1. Use of BCA method in [Erdogan, 2013] to separate the direct and ground reflected echoes from a target is proposed. The application can improve the performance of the direction of arrival estimation methods especially for faded signal samples.

2. A new geometric framework called Sparse Bounded Component Analysis is proposed. This framework relaxes the statistical independence assumption used by well known BSS methods with a less stringent domain separability assumption in a sparse source setting. As a result, it can be applied to both independent and dependent source mixtures.

3. The instantaneous SBCA framework is extended to cover convolutive mixtures. Both time-based and frequency-based versions of the convolutive SBCA approach are proposed. It is also proven that the global optimums of the corresponding optimization problems are perfect separators.

4. A novel short time component analysis tool based on SBCA is developed targeting brain imaging applications.

6.3. Future Work

Although we consider that the goal of the thesis has been accomplished, there are some improvements and studies that could be done. Note that it is basically a compilation of the future works of each of the previous chapters. The future works are summarized as follows:

1. A performance improvement study in the identification stage of the neuron sorting tool is planned. Usage of a neural network algorithm instead of the correlation computations is believed to be able to increase the precision of the neuron identification stage.

2. A new article entitled "Sparse Bounded Component Analysis Based Neuron Sorting for Ca^{2+} Imaging" will be submitted.

3. Stationary point analysis of SBCA algorithm is planned. Inan, Erdogan and Cruces provided an analysis in [Inan et al., 2017] that the stationary points of the instantaneous BCA algorithm [Erdogan, 2013] correspond to either the global maximums of the objective function or unstable saddle points. It means that the stationary points of the BCA algorithm do not correspond to the local maximums of the objective function. A similar stationary point characterization is in our future research agenda and we expect it to lead to the same conclusion with [Inan et al., 2017].

BIBLIOGRAPHY

[A. Cichocki and Siwek,] A. Cichocki, S.-I. A. and Siwek, K. Icalab toolboxes. *http://www.bsp.brain.riken.jp/ ICALAB/ICALABSignalProc/benchmarks/.*

[Abrard and Deville, 2005] Abrard, F. and Deville, Y. (2005). A time–frequency blind signal separation method applicable to underdetermined mixtures of dependent sources. *Signal processing*, 85(7):1389–1403.

[Aichner et al., 2003] Aichner, R., Buchner, H., Araki, S., and Makino, S. (2003). On-line time-domain blind source separation of nonstationary convolved signals. In *Proc. Int. Symposium on Independent Component Analysis and Blind Signal Spearation*, pages 987–992.

[Amari, 1998] Amari, S.-I. (1998). Natural gradient works efficiently in learning. *Neural computation*, 10(2):251–276.

[Arberet et al., 2006] Arberet, S., Gribonval, R., and Bimbot, F. (2006). A robust method to count and locate audio sources in a stereophonic linear instantaneous mixture. In *International Conference on Independent Component Analysis and Signal Separation*, pages 536–543. Springer.

[Attias, 2003] Attias, H. (2003). New em algorithms for source separation and deconvolution. In *Acoustics, Speech, and Signal Processing, 2003 IEEE International Conference on*, volume 5, pages 297–300. IEEE.

[Babatas and Erdogan,] Babatas, E. and Erdogan, A. T. Time and frequency based sparse bounded component analysis algorithms for convolutive mixtures. *revised manuscript, Signal Processing.*

[Babatas and Erdogan, 2018a] Babatas, E. and Erdogan, A. T. (01 October 2018a). An algorithmic framework for sparse bounded component analysis. *IEEE Transactions on Signal Processing*, 66(19):5194–5205.

[Babatas and Erdogan, 2016] Babatas, E. and Erdogan, A. T. (2016). Sparse bounded component analysis. In *Machine Learning for Signal Processing (MLSP), 2016 IEEE 26th International Workshop on*, pages 1–6. IEEE.

[Babatas and Erdogan, 2017] Babatas, E. and Erdogan, A. T. (2017). Doa estimation of low altitude target based on bounded component analysis algorithm. In *International Conference on Radar Systems (Radar 2017).* IET.

[Babatas and Erdogan, 2018b] Babatas, E. and Erdogan, A. T. (2018b). Sparse bounded component analysis for convolutive mixtures. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2741–2745. IEEE.

[Bagirov et al., 2014] Bagirov, A., Karmitsa, N., and Mäkelä, M. M. (2014). *Introduction to Nonsmooth Optimization: theory, practice and software*. Springer.

[Bandettini et al., 1993] Bandettini, P. A., Jesmanowicz, A., Wong, E. C., and Hyde, J. S. (1993). Processing strategies for time-course data sets in functional mri of the human brain. *Magnetic resonance in medicine*, 30(2):161–173.

[Bell and Sejnowski, 1995] Bell, A. J. and Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural computation*, 7(6):1129–1159.

[Belouchrani and Amin, 1998] Belouchrani, A. and Amin, M. G. (1998). Blind source separation based on time-frequency signal representations. *IEEE Transactions on Signal Processing*, 46(11):2888–2897.

[Bingham and Hyvärinen, 2000] Bingham, E. and Hyvärinen, A. (2000). A fast fixed-point algorithm for independent component analysis of complex valued signals. *International journal of neural systems*, 10(01):1–8.

[Biswal et al., 1995] Biswal, B., Zerrin Yetkin, F., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic resonance in medicine*, 34(4):537–541.

[Biswal and Ulmer, 1999] Biswal, B. B. and Ulmer, J. L. (1999). Blind source separation of multiple signal sources of fmri data sets using independent component analysis. *Journal of computer assisted tomography*, 23(2):265–271.

[Blin et al., 2005] Blin, A., Araki, S., and Makino, S. (2005). Underdetermined blind separation of convolutive mixtures of speech using time-frequency mask and mixing matrix estimation. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 88(7):1693–1700.

[Bobin et al., 2007] Bobin, J., Starck, J.-L., Fadili, J., and Moudden, Y. (2007). Sparsity and morphological diversity in blind source separation. *IEEE Transactions on Image Processing*, 16(11):2662–2674.

[Bofill, 2003] Bofill, P. (2003). Underdetermined blind separation of delayed sound sources in the frequency domain. *Neurocomputing*, 55(3-4):627–641.

[Bofill and Zibulevsky, 2000] Bofill, P. and Zibulevsky, M. (2000). Blind separation of more sources than mixtures using sparsity of their short-time fourier transform. In *Proc. ica*, volume 2000, pages 87–92.

[Boyd and Vandenberghe, 2004] Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

[Brammer et al., 1997] Brammer, M., Wright, I., Woodruff, P., Williams, S., Simmons, A., and Bullmore, E. (1997). Wavelet analysis of periodic and nonperiodic experimental designs in functional magnetic resonance imaging of the brain. *Neuroimage*, 5:479.

[Calhoun et al., 2001] Calhoun, V. D., Adali, T., Pearlson, G. D., and Pekar, J. J. (2001). A method for making group inferences from functional mri data using independent component analysis. *Human brain mapping*, 14(3):140–151.

[Calhoun et al., 2003] Calhoun, V. D., Adali, T., Pekar, J. J., and Pearlson, G. (2003). Latency (in) sensitive ica: group independent component analysis of fmri data in the temporal frequency domain. *NeuroImage*, 20(3):1661–1669.

[Candès and Donoho, 1999] Candès, E. J. and Donoho, D. L. (1999). Ridgelets: A key to higher-dimensional intermittency? *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 357(1760):2495–2509.

[Candes and Donoho, 2000] Candes, E. J. and Donoho, D. L. (2000). Curvelets: A surprisingly effective nonadaptive representation for objects with edges. Technical report, Stanford Univ Ca Dept of Statistics.

[Cardoso and Souloumiac, 1993] Cardoso, J.-F. and Souloumiac, A. (1993). Blind beamforming for non-gaussian signals. In *IEE Proceedings F-Radar and Signal Processing*, volume 140, pages 362–370. IET.

[Castella and Moreau, 2010] Castella, M. and Moreau, E. (2010). A new method for kurtosis maximization and source separation. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 2670–2673. IEEE.

[Chambolle et al., 2010] Chambolle, A., Caselles, V., Cremers, D., Novaga, M., and Pock, T. (2010). An introduction to total variation for image analysis. *Theoretical foundations and numerical methods for sparse recovery*, 9(263-340):227.

[Chan et al., 2009] Chan, T.-H., Chi, C.-Y., Huang, Y.-M., and Ma, W.-K. (2009). A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing. *IEEE Transactions on Signal Processing*, 57(11):4418–4432.

[Chen et al., 2001] Chen, S. S., Donoho, D. L., and Saunders, M. A. (2001). Atomic decomposition by basis pursuit. *SIAM review*, 43(1):129–159.

[Cho and Yoo, 2015] Cho, J. and Yoo, C. D. (2015). Underdetermined convolutive bss: Bayes risk minimization based on a mixture of super-gaussian posterior approximation. *IEEE Transactions on Audio, Speech, and Language Processing*, 23(5):828–839.

[Comon, 1994] Comon, P. (1994). Independent component analysis, a new concept? *Signal processing*, 36(3):287–314.

[Cotter et al., 2005] Cotter, S. F., Rao, B. D., Engan, K., and Kreutz-Delgado, K. (2005). Sparse solutions to linear inverse problems with multiple measurement vectors. *IEEE Transactions on Signal Processing*, 53(7):2477–2488.

[Craig, 1994] Craig, M. D. (1994). Minimum-volume transforms for remotely sensed data. *IEEE Transactions on Geoscience and Remote Sensing*, 32(3):542–552.

[Cruces, 2010] Cruces, S. (2010). Bounded component analysis of linear mixtures: A criterion of minimum convex perimeter. *IEEE Transactions on Signal Processing*, 58(4):2141–2154.

[Cruces et al., 2011] Cruces, S., Aguilera-Bonet, P., and Durán-Díaz, I. (2011). Criterion for signal extraction in underdetermined mixtures of bounded support. *Signal Processing*, 91(10):2437–2442.

[Do and Vetterli, 2005] Do, M. N. and Vetterli, M. (2005). The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on image processing*, 14(12):2091–2106.

[Dombeck et al., 2007] Dombeck, D. A., Khabbaz, A. N., Collman, F., Adelman, T. L., and Tank, D. W. (2007). Imaging large-scale neural activity with cellular resolution in awake, mobile mice. *Neuron*, 56(1):43–57.

[Donoho, 1995] Donoho, D. L. (1995). De-noising by soft-thresholding. *IEEE transactions on information theory*, 41(3):613–627.

[Douglas et al., 2005a] Douglas, S., Sawada, H., and Makino, S. (2005a). Natural gradient multichannel blind deconvolution and speech separation using causal FIR filters. *IEEE Trans. on Speech and Audio Processing*, 13(1):92–104.

[Douglas et al., 2007] Douglas, S. C., Gupta, M., Sawada, H., and Makino, S. (2007). Spatio–temporal fastica algorithms for the blind separation of convolutive mixtures. *IEEE transactions on audio, speech, and language processing*, 15(5):1511–1520.

[Douglas et al., 2005b] Douglas, S. C., Sawada, H., and Makino, S. (2005b). Natural gradient multichannel blind deconvolution and speech separation using causal fir filters. *IEEE Transactions on Speech and Audio Processing*, 13(1):92–104.

[Erdogan, 2013] Erdogan, A. T. (2013). A class of bounded component analysis algorithms for the separation of both independent and dependent sources. *IEEE Trans. Signal Processing*, 61(22):5730–5743.

[Févotte and Doncarli, 2004] Févotte, C. and Doncarli, C. (2004). Two contributions to blind source separation using time-frequency distributions. *IEEE Signal Processing Letters*, 11(3):386–389.

[Fevotte et al., 2005] Fevotte, C., Gribonval, R., and Vincent, E. (2005). Bss eval toolbox user guide, irisa tech. report 1706. Technical report, Tech. Rep. Rennes, France.

[Freeman et al.,] Freeman, J., Sofroniew, N., Broxton, M., Grosenick, L., Trautmann, E., and Kitch, L. *http://neurofinder.codeneuro.org/*.

[Georgiev and Cichoki, 2004] Georgiev, P. and Cichoki, A. (2004). Sparse component analysis of overcomplete mixtures by improved basis pursuit method. In *Circuits and Systems, 2004. ISCAS'04. Proceedings of the 2004 International Symposium on*, volume 5. IEEE.

[Giovannucci et al., 2019] Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B. L., Koay, S. A., Taxidis, J., Najafi, F., Gauthier, J. L., Zhou, P., Khakh, B. S., Tank, D. W., Chklovskii, D. B., and Pnevmatikakis, E. A. (2019). Caiman: An open source tool for scalable calcium imaging data analysis. *eLife*, 8:e38173.

[Gorodnitsky and Rao, 1997] Gorodnitsky, I. F. and Rao, B. D. (1997). Sparse signal reconstruction from limited data using focuss: A re-weighted minimum norm algorithm. *IEEE Transactions on signal processing*, 45(3):600–616.

[Gorokhov and Loubaton, 1996] Gorokhov, A. and Loubaton, P. (1996). Second order blind identification of convolutive mixtures with temporally correlated sources: A subspace based approach. In *EUSIPCO*, volume 96, pages 2093–2096. EURASIP.

[Griesser and Balanis, 1987] Griesser, T. and Balanis, C. (1987). Oceanic low-angle monopulse radar tracking errors. *IEEE Journal of Oceanic Engineering*, 12(1):289–295.

[Hansen, 2003] Hansen, L. K. (2003). Ica of fmri based on a convolutive mixture model. In *9th Annual Meeting of the Organization for Human Brain Mapping*.

[Hayvaci et al., 2013] Hayvaci, H. T., De Maio, A., and Erricolo, D. (2013). Improved detection probability of a radar target in the presence of multipath with prior knowledge of the environment. *IET Radar, Sonar & Navigation*, 7(1):36–46.

[He and Cichocki, 2006] He, Z. and Cichocki, A. (2006). K-evd clustering and its applications to sparse component analysis. In *International Conference of Independent Component Analysis and Signal Separation*, pages 90–97.

[He et al., 2007] He, Z., Xie, S., Ding, S., and Cichocki, A. (2007). Convolutive blind source separation in the frequency domain based on sparse representation. *IEEE Transactions on Audio, Speech, and Language Processing*, 15(5):1551–1563.

[Herault and Jutten, 1986] Herault, J. and Jutten, C. (1986). Space or time adaptive signal processing by neural network models. In *Proceedings of the American Institute for Physics Conference*, volume 151, pages 206–211.

[Hyvarinen et al., 1999] Hyvarinen, A., Cristescu, R., and Oja, E. (1999). A fast algorithm for estimating overcomplete ica bases for image windows. In *Neural Networks, 1999. IJCNN'99. International Joint Conference on*, volume 2, pages 894–899. IEEE.

[Hyvärinen and Oja, 1997] Hyvärinen, A. and Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. *Neural Computation*, 9(7):1483–1492.

[Ikeda and Murata, 1999] Ikeda, S. and Murata, N. (1999). A method of ica in time-frequency domain. In *in Proc. ICA*. Citeseer.

[Inan and Erdogan, 2015] Inan, H. A. and Erdogan, A. T. (2015). A convolutive bounded component analysis framework for potentially nonstationary independent and/or dependent sources. *IEEE Transactions on Signal Processing*, 63(1):18–30.

[Inan et al., 2017] Inan, H. A., Erdogan, A. T., and Cruces, S. (2017). Stationary point characterization for a class of bca algorithms. *IEEE Transactions on Signal Processing*, 65(20):5437–5452.

[Inouye and Liu, 2002] Inouye, Y. and Liu, R.-W. (2002). A system-theoretic foundation for blind equalization of an fir mimo channel system. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 49(4):425–436.

[Javanmard et al., 2008] Javanmard, A., Pad, P., Babaie-Zadeh, M., and Jutten, C. (August 2008). Estimating the mixing matrix in underdetermined sparse component analysis using consecutive independent component analysis. In *Signal Processing Conference, 16th European*, pages 1–5.

[Jin et al., 2012] Jin, L., Han, Z., Platisa, J., Wooltorton, J. R., Cohen, L. B., and Pieribone, V. A. (2012). Single action potentials and subthreshold electrical events imaged in neurons with a fluorescent protein voltage probe. *Neuron*, 75(5):779–785.

[Jourjine et al., 2000] Jourjine, A., S., R., and O., Y. (2000). Blind separation of disjoint orthogonal signals: Demixing n sources from 2 mixtures. In *ICASSP2000*, volume 5, pages 2985–2988.

[Koldovský and Tichavský, 2007] Koldovský, Z. and Tichavský, P. (August 2007). Time-domain blind audio source separation using advanced ica methods. In *8th Annual Conference of the International Speech Communication Association*.

[Koldovský et al., 2006] Koldovský, Z., Tichavský, P., and Oja, E. (2006). Efficient variant of algorithm fastica for independent component analysis attaining the cramér-rao lower bound. *IEEE Transactions on Neural Networks*, 17(5):1265–1277.

[Kolossa and Orglmeister, 2004] Kolossa, D. and Orglmeister, R. (2004). Nonlinear postprocessing for blind speech separation. In *International Conference on Independent Component Analysis and Signal Separation*, pages 832–839. Springer.

[Kreutz-Delgado et al., 2003] Kreutz-Delgado, K., Murray, J. F., Rao, B. D., Engan, K., Lee, T.-W., and Sejnowski, T. J. (2003). Dictionary learning algorithms for sparse representation. *Neural computation*, 15(2):349–396.

[Lab,] Lab, M. I. A. Gift software. *http://mialab.mrn.org/software/gift/*.

[Lee et al., 1999a] Lee, T.-W., Girolami, M., and Sejnowski, T. J. (1999a). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation*, 11(2):417–441.

[Lee et al., 1999b] Lee, T.-W., Lewicki, M. S., Girolami, M., and Sejnowski, T. J. (1999b). Blind source separation of more sources than mixtures using overcomplete representations. *IEEE signal processing letters*, 6(4):87–90.

[Lewicki and Sejnowski, 2000] Lewicki, M. S. and Sejnowski, T. J. (2000). Learning overcomplete representations. *Neural computation*, 12(2):337–365.

[Li et al., 2003] Li, Y., Cichocki, A., and Amari, S.-i. (April 2003). Sparse component analysis for blind source separation with less sensors than sources. In *ICA*, pages 89–94.

[Lourakis, 2016] Lourakis, M. (2016). TvL1denoise.m: TV-L1 image denoising algorithm. `https://www.mathworks.com/matlabcentral/fileexchange/57604-tv-l1-image-denoising-algorithm`.

[Makino et al., 2007] Makino, S., Sawada, H., and Lee, T.-W. (2007). *Blind Speech Separation*, volume 615. Springer.

[Makino et al., 2005] Makino, S., Sawada, H., Mukai, R., and Araki, S. (July 2005). Blind source separation of convolutive mixtures of speech in frequency domain. *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 88(7):1640–1655.

[Mallat, 2008] Mallat, S. (2008). *A wavelet tour of signal processing: the sparse way*. Academic press.

[Mansour et al., 2000] Mansour, A., Jutten, C., and Loubaton, P. (2000). Adaptive subspace algorithm for blind separation of independent sources in convolutive mixture. *IEEE Transactions on Signal Processing*, 48(2):583–586.

[Marqui and Lehmann D., 1994] Marqui, P. and Lehmann D., M. C. (1994). Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. *International Journal of psychophysiology*, 18(1):49–65.

[Martin and Flandrin, 1985] Martin, W. and Flandrin, P. (1985). Wigner-ville spectral analysis of nonstationary processes. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(6):1461–1470.

[McKeown et al., 1998] McKeown, M. J., Makeig, S., Brown, G. G., Jung, T.-P., Kindermann, S. S., Bell, A. J., and Sejnowski, T. J. (1998). Analysis of fmri data by blind separation into independent spatial components. *Human brain mapping*, 6(3):160–188.

[Mei et al., 2008] Mei, T., Mertins, A., Yin, F., Xi, J., and Chicharo, J. (2008). Blind source separation for convolutive mixtures based on the joint diagonalization of power spectral density matrices. *Signal Processing*, 88(8):1990 – 2007.

[Michela et al., 2004] Michela, C. M., Murraya, M. M., Lantza, G., Gonzaleza, S., Spinelli, L., and Peralta, R. G. d. (2004). Eeg source imaging. In *Clinical Neurophysiology*, volume 115, pages 2195–2222.

[Mishali and Eldar, 2009] Mishali, M. and Eldar, Y. C. (2009). Sparse source separation from orthogonal mixtures. In *ICASSP 2009*, pages 3145–3148.

[Mohammed et al., 2016] Mohammed, A. I., Gritton, H. J., Tseng, H., Bucklin, M. E., Yao, Z., and Han, X. (2016). An integrative approach for analyzing hundreds of neurons in task performing mice using wide-field calcium imaging. *Scientific Reports 6:20986.*

[Morlet and Grossman, 1984] Morlet, J. and Grossman, A. (1984). Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM journal on mathematical analysis*, 15(4):723–736.

[Mourad and Reilly, 2010] Mourad, N. and Reilly, J. P. (2010). Modified hierarchical clustering for sparse component analysis. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 2674–2677.

[Mukamel et al., 2009] Mukamel, E. A., Nimmerjahn, A., and Schnitzer, M. J. (2009). Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron*, 63(6):747–760.

[Nakai et al., 2004] Nakai, T., Muraki, S., Bagarinao, E., Miki, Y., Takehara, Y., Matsuo, K., Kato, C., Sakahara, H., and Isoda, H. (2004). Application of independent component analysis to magnetic resonance imaging for enhancing the contrast of gray and white matter. *Neuroimage*, 21(1):251–260.

[Nesterov, 2013] Nesterov, Y. (2013). *Introductory lectures on convex optimization: A basic course*, volume 87. Springer Science & Business Media.

[Nielsen et al., 2014] Nielsen, J. K., Jensen, J. R., Jensen, S. H., and Christensen, M. G. (2014). The single- and multichannel audio recordings database (smard). In *IWAENC*, pages 40–44.

[Nion and Sidiropoulos, 2009] Nion, D. and Sidiropoulos, N. D. (2009). Adaptive algorithms to track the parafac decomposition of a third-order tensor. *IEEE Transactions on Signal Processing*, 57(6):2299–2310.

[O'grady et al., 2005] O'grady, P. D., Pearlmutter, B. A., and Rickard, S. T. (2005). Survey of sparse and non-sparse methods in source separation. *International Journal of Imaging Systems and Technology*, 15(1):18–33.

[Oktem et al., 2010] Oktem, T., Erdogan, A. T., and Demir, A. (2010). Adaptive receiver structures for fiber communication systems employing polarization-division multiplexing. *Journal of Lightwave Technology*, 28(10):1536–1546.

[Otsu, 1979] Otsu, N. (1979). A threshold selection method from gray-level histogram. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66.

[Pachitariu et al., 2013] Pachitariu, M., Packer, A. M., Pettit, N., Dalgleish, H., Hausser, M., and Sahani, M. (2013). Extracting regions of interest from biological images with convolutional sparse block coding. In *Advances in Neural Information Processing Systems*, pages 1745–1753.

[Parra and Spence, 2000] Parra, L. and Spence, C. (2000). Convolutive blind source separation of non-stationary sources. *IEEE Trans. on Speech and Audio Processing, (Code available at http://bsp.teithe.gr/members/downloads.html)*, 8(3):320–327.

[Patel et al., 2015] Patel, T. P., Man, K., Firestein, B. L., and Meaney, D. F. (2015). Automated quantification of neuronal networks and single-cell calcium dynamics using calcium imaging. *Journal of Neuroscience Methods*, pages 26–38.

[Pedersen et al., 2007] Pedersen, M. S., Larsen, J., Kjems, U., and Parra, L. C. (2007). *A survey of convolutive blind source separation methods*. Springer.

[Pedersen et al., 2005] Pedersen, M. S., Wang, D., Larsen, J., and Kjems, U. (2005). Overcomplete blind source separation by combining ica and binary time-frequency masking. In *Proceedings of the MLSP workshop, Mystic, CT, USA*.

[Petersen et al., 2000] Petersen, K., Hansen, L. K., Kolenda, T., Rostrup, E., and Strother, S. (2000). On the independent components of functional neuroimages. In *Third international conference on independent component analysis and blind source separation*, pages 615–620.

[Qayyum et al., 2015] Qayyum, A., Malik, A. S., Naufal, M., Saad, M., Mazher, M., Abdullah, F., and Abdullah, T. A. R. B. T. (2015). Designing of overcomplete dictionaries based on dct and dwt. In *Biomedical Engineering & Sciences (ISSBES), 2015 IEEE Student Symposium in*, pages 134–139. IEEE.

[Rahbar and Reilly, 2005] Rahbar, K. and Reilly, J. P. (September 2005). A frequency domain method for blind source separation of convolutive audio mixtures. *IEEE Trans. on Speech and Audio Processing*, 13(5):832–844.

[Randel et al., 2011] Randel, S., Ryf, R., Sierra, A., Winzer, P. J., Gnauck, A. H., Bolle, C. A., Essiambre, R.-J., Peckham, D. W., McCurdy, A., and Lingle, R. (2011). $6\times$ 56-gb/s mode-division multiplexed transmission over 33-km few-mode fiber enabled by $6\times$ 6 mimo equalization. *Optics Express*, 19(17):16697–16707.

[Rickard et al., 2005] Rickard, S., Melia, T., and Fearon, C. (2005). Desprit - histogram based blind source separation of more sources than sensors using subspace methods. In *WASPAA05*, pages 5–8.

[Rubinstein et al., 2010] Rubinstein, R., Bruckstein, A. M., and Elad, M. (2010). Dictionaries for sparse representation modeling. *Proceedings of the IEEE*, 98(6):1045–1057.

[Rudin, 2006] Rudin, W. (2006). *Real and Complex Analysis.* McGraw-Hill Book Co.

[Saruwatari et al., 2003] Saruwatari, H., Kurita, S., Takeda, K., Itakura, F., Nishikawa, T., and Shikano, K. (2003). Blind source separation combining independent component analysis and beamforming. *EURASIP Journal on Advances in Signal Processing*, 2003(11):569270.

[Smaragdis, 1998] Smaragdis, P. (1998). Blind separation of convolved mixtures in the frequency domain. *Neurocomputing*, 22(1-3):21–34.

[Snoussi and Calhoun, 2005] Snoussi, H. and Calhoun, V. D. (2005). Bayesian blind source separation for brain imaging. In *IEEE International Conference on Image Processing 2005*, volume 3, pages III–581. IEEE.

[Souloumiac, 1995] Souloumiac, A. (1995). Blind source detection and separation using second order non-stationarity. In *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, volume 3, pages 1912–1915. IEEE.

[Speech and Group,] Speech and Group, A. P. Mardy database *http://www.commsp.ee.ic.ac.uk/ sap/resources / mardy-multichannel-acoustic-reverberation-database-at-york-database/*.

[Stoica and Kenneth C., 1990] Stoica, P. and Kenneth C., S. (1990). Maximum likelihood methods for direction-of-arrival estimation. *IEEE Trans. Acoust. Speech Signal Process.*, 38(7):1132–1143.

[Theis et al., 2004] Theis, F. J., Georgiev, P. G., and Cichocki, A. (2004). Robust overcomplete matrix recovery for sparse sources using a generalized hough transform. In *ESANN'2004 Proceedings - European Symposium on Artificial Neural Networks*, pages 343–348.

[Thomas et al., 2006] Thomas, J., Deville, Y., and Hosseini, S. (2006). Time-domain fast fixed-point algorithms for convolutive ica.

[Tian et al., 2009] Tian, L., Hires, S. A., Mao, T., Huber, D., Chiappe, M. E., Chalasani, S. H., Petreanu, L., Akerboom, J., McKinney, S. A., Schreiter, E. R., et al. (2009). Imaging neural activity in worms, flies and mice with improved gcamp calcium indicators. *Nature methods*, 6(12):875.

[Trefethen and Bau III, 1997] Trefethen, L. N. and Bau III, D. (1997). Numerical linear algebra. *SIAM*, 50.

[Vayá et al., 2007] Vayá, C., Rieta, J. J., Sánchez, C., and Moratal, D. (2007). Convolutive blind source separation algorithms applied to the electrocardiogram of atrial fibrillation: Study of performance. *IEEE Transactions on Biomedical Engineering*, 54(8):1530–1533.

[Vincent et al., 2006] Vincent, E., Gribonval, R., and Févotte, C. (2006). Performance measurement in blind audio source separation. *IEEE Trans. on Audio, Speech and Language Processing*, 14(4):1462–1469.

[Wang et al., 2010] Wang, F.-Y., Chi, C.-Y., Chan, T.-H., and Wang, Y. (2010). Nonnegative least-correlated component analysis for separation of dependent sources by volume maximization. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 32(5):875–888.

[Wen et al., 2006] Wen, J. Y., Gaubitch, N. D., Habets, E. A., Myatt, T., and Naylor, P. A. (2006). Evaluation of speech dereverberation algorithms using the mardy database. In *IWAENC 2006*.

[Wendel et al., 2009] Wendel, K., Vaisanen, O., Malmivuo, J., Gencer, N. G., and Vanrumste, B. (2009). Eeg/meg source imaging: Methods challenges and open issues. *Computational Intelligence and Neuroscience*.

[Yilmaz and Rickard, 2004] Yilmaz, O. and Rickard, S. (2004). Blind separation of speech mixtures via time-frequency masking. *IEEE Transactions on Signal Processing*, 52:1830 – 1847.

[Zibulevsky, 2003] Zibulevsky, M. (2003). Blind source separation with relative newton method. In *Proc. ICA*, volume 2003, pages 897–902.

[Zibulevsky and Pearlmutter, 2001] Zibulevsky, M. and Pearlmutter, B. A. (2001). Blind source separation by sparse decomposition in a signal dictionary. *Neural computation*, 13(4):863–882.

[Ziv et al., 2013] Ziv, Y., Burns, L. D., Cocker, E. D., Hamel, E. O., Ghosh, K. K., Kitch, L. J., Gamal, A. E., and Schnitzer, M. J. (2013). Long-term dynamics of ca1 hippocampal place codes. *Nature Neuroscience*, 16(3).

[Zoltowski and Lee, 1991] Zoltowski, M. D. and Lee, T.-S. (1991). Maximum likelihood based sensor array signal processing in the beamspace domain for low angle radar tracking. *IEEE Transactions on Signal Processing*, 39(3):656–671.

Appendix A

PROOF OF THEOREM-3

The first step is to rewrite (4.10) in terms of the arguments $\mathbf{F}(k) = \sum_{l=0}^{P-1} \mathbf{W}(l)$ $\mathbf{A}(k-l)$ (for $k = 0, \ldots, P-1$), $\tilde{\mathbf{F}}$, and $\Gamma_N(\tilde{\mathbf{F}})$ defined in the Table 4.1. $\Gamma_N(\tilde{\mathbf{F}})$ is of size $(Np) \times (N+P-1)p$ whose first block row is $[\mathbf{F}(0), \mathbf{F}(1), \ldots, \mathbf{F}(P-1), \mathbf{0}, \ldots, \mathbf{0}]$ and first block column is $[\mathbf{F}(0), \mathbf{0}, \ldots, \mathbf{0}]^T$ where the zero matrices are of size $p \times p$.

Writing the sample covariance matrix with respect to extended separator output vector as

$$\tilde{\mathbf{y}}_N(n) = \Gamma_N(\tilde{\mathbf{F}})\tilde{\mathbf{s}}_{L_3}(n) n = N, \ldots, L_1, \tag{A.1}$$

$$\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_\mathbf{N}} = \Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{\mathbf{L_3}}}\Gamma_N(\tilde{\mathbf{F}})^T, \tag{A.2}$$

the objective function turns out to

$$J(\tilde{\mathbf{F}}) = \frac{\sqrt{\det(\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T)}}{(\max_{n\in\{N,\ldots,L_1\}} \|\Gamma_N(\tilde{\mathbf{F}})\tilde{\mathbf{s}}_{L_3}(n)\|_1)^{Np}}. \tag{A.3}$$

Under the assumption (**A4**), the following inequality is valid for the denominator of (A.3):

$$(\max_{n\in\{N,\ldots,L_1\}} \|\Gamma_N(\tilde{\mathbf{F}})\tilde{\mathbf{s}}_{L_3}(n)\|_1)^{Np} \leq \|\Gamma_N(\tilde{\mathbf{F}})\|_{1,1}^{Np}, \tag{A.4}$$

where

$$\|\Gamma_N(\tilde{\mathbf{F}})\|_{1,1} = \| \left[\|\Gamma_N(\tilde{\mathbf{F}})_{:,1}\|_1, \ldots, \|\Gamma_N(\tilde{\mathbf{F}})_{:,L_3 p}\|_1 \right] \|_\infty. \tag{A.5}$$

(A.4) turns into an equality under the assumption (**A2**). Consequently, (A.3) becomes

$$J(\tilde{\mathbf{F}}) = \frac{\sqrt{\det(\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T)}}{\left\|\left[\ \|\Gamma_N(\tilde{\mathbf{F}})_{:,1}\|_1 \cdots \|\Gamma_N(\tilde{\mathbf{F}})_{:,L_3p}\|_1\ \right]\right\|_\infty^{Np}}. \tag{A.6}$$

Since we choose $N \geq P$, there is at least one block column inside $\Gamma_N(\tilde{\mathbf{F}})$ that contains $\tilde{\mathbf{F}}$ in upside down form. Therefore, we have

$$\begin{aligned}\left\|\left[\ \|\Gamma_N(\tilde{\mathbf{F}})_{:,1}\|_1 \quad \cdots \quad \|\Gamma_N(\tilde{\mathbf{F}})_{:,L_3p}\|_1\ \right]\right\|_\infty &= \left\|\left[\ \|\tilde{\mathbf{F}}_{1,:}\|_1 \quad \cdots \quad \|\tilde{\mathbf{F}}_{p,:}\|_1\ \right]\right\|_\infty \\ &= \left\|\left[\ \|\Gamma_N(\tilde{\mathbf{F}})_{1,:}\|_1 \quad \cdots \quad \|\Gamma_N(\tilde{\mathbf{F}})_{Np,:}\|_1\ \right]\right\|_\infty.\end{aligned} \tag{A.7}$$

Moving from the equality (A.7), the objective can be transformed to

$$J(\tilde{\mathbf{F}}) = \frac{\sqrt{\det(\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T)}}{\left\|\left[\ \|\Gamma_N(\tilde{\mathbf{F}})_{1,:}\|_1 \cdots \|\Gamma_N(\tilde{\mathbf{F}})_{Np,:}\|_1\ \right]\right\|_\infty^{Np}} \tag{A.8}$$

$$\leq \frac{\sqrt{\det(\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T)}}{(\|\left[\|\Gamma_N(\tilde{\mathbf{F}})_{1,:}\|_1 \cdots \|\Gamma_N(\tilde{\mathbf{F}})_{Np,:}\|_1\right]\|_1/(Np))^{Np}} \tag{A.9}$$

$$\leq \frac{\sqrt{\det(\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T)}}{\|\Gamma_N(\tilde{\mathbf{F}})_{1,:}\|_1\|\Gamma_N(\tilde{\mathbf{F}})_{2,:}\|_1 \cdots \|\Gamma_N(\tilde{\mathbf{F}})_{Np,:}\|_1} \tag{A.10}$$

$$\leq \frac{\sqrt{\det(\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T)}}{\|\Gamma_N(\tilde{\mathbf{F}})_{1,:}\|_2\|\Gamma_N(\tilde{\mathbf{F}})_{2,:}\|_2 \cdots \|\Gamma_N(\tilde{\mathbf{F}})_{Np,:}\|_2}. \tag{A.11}$$

To simplify the upper bound expression more, it is possible to rewrite the nominator of the objective by completing $\Gamma_N(\tilde{\mathbf{F}})$ to a full rank square matrix. For the resulting matrix's positions, we could use Schur complement. Let's define a $(P-1)p \times (L_3)p$ matrix $\mathbf{Z} = \mathbf{DP}$ where $\mathbf{D} = diag(\mathbf{a}_1; \mathbf{a}_2; \ldots; \mathbf{a}_{(P-1)p})$ is a full rank diagonal matrix and $\mathbf{P}$ is a permutation matrix such that $\det(\mathbf{ZBZ}^T) = 1$, $\mathbf{B} = \hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}} - \hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T(\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T)^{-1}\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}$.

The determinant in the denominator can be rewritten using the matrix $\mathbf{Z}$:

$$\begin{aligned}
&\det\left(\begin{bmatrix} \Gamma_N(\tilde{\mathbf{F}}) \\ \mathbf{Z} \end{bmatrix} \hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}} [\ \Gamma_N(\tilde{\mathbf{F}})^T \quad \mathbf{Z}^T\ ]\right) \\
&= \det\left(\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T\right)\det(\mathbf{Z}\left(\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}} - \hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T(\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T)^{-1}\right. \\
&\qquad \left.\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\right)\mathbf{Z}^T), \\
&= \det\left(\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T\right)\det(\mathbf{Z}\mathbf{B}\mathbf{Z}^T), \\
&= \det\left(\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T\right). \qquad \text{(A.12)}
\end{aligned}$$

$\mathbf{Z}\mathbf{B}\mathbf{Z}^T$ is called the Schur complement of $\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T$. Selecting an appropriate $\mathbf{Z}$ matrix with respect to Schur complement and using Hadamard's Inequality provides the inequality

$$\det\left(\Gamma_N(\tilde{\mathbf{F}})\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}\Gamma_N(\tilde{\mathbf{F}})^T\right) \leq \prod_{m=1}^{Np} \parallel \Gamma_N(\tilde{\mathbf{F}})_{m,:} \parallel_2^2 \prod_{n=1}^{(P-1)p} \parallel \mathbf{Z}_{n,:} \parallel_2^2 \det(\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}}). \quad \text{(A.13)}$$

A new upper bound can be produced from (A.11) using (A.13):

$$\begin{aligned}
J(\tilde{\mathbf{F}}) &\leq \frac{\prod_{m=1}^{Np} \|\Gamma_N(\tilde{\mathbf{F}})_{m,:}\|_2 \prod_{n=1}^{(P-1)p} \parallel \mathbf{Z}_{n,:} \parallel_2 \det(\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}})^{1/2}}{\prod_{l=1}^{Np} \|\Gamma_N(\tilde{\mathbf{F}})_{l,:}\|_2} \\
J(\tilde{\mathbf{F}}) &\leq \prod_{n=1}^{(P-1)p} \parallel \mathbf{Z}_{n,:} \parallel_2 \det(\hat{\mathbf{R}}_{\tilde{\mathbf{s}}_{L_3}})^{1/2} \qquad \text{(A.14)}
\end{aligned}$$

$\Gamma_N(\tilde{\mathbf{F}})$ is block-Toeplitz, thus each row of it corresponds to a row of $\tilde{\mathbf{F}}$. Moreover, we choose $N \geq P$, it guarantees that there is a block column of $\Gamma_N(\tilde{\mathbf{F}})$ that contains $[\mathbf{F}(0), \mathbf{F}(1), \ldots, \mathbf{F}(P-1)]$. Thus, we will consider $\tilde{\mathbf{F}}$ instead of $\Gamma_N(\tilde{\mathbf{F}})$ to comment the above inequalities. The inequalities in (A.9) and (A.10) are achieved only if all the rows of $\tilde{\mathbf{F}}$ have the same ℓ_1 norm. The inequality between ℓ_1 and ℓ_2 norms in (A.11) is achieved only if there is only one non-zero entry in each row of $\tilde{\mathbf{F}}$. Hadamard inequality in (A.13) turns into an equality only if the non-zero entries in the rows of $\tilde{\mathbf{F}}$ are in different positions with respect to mode p, in order to satisfy the orthogonality of the rows of $\Gamma_N(\tilde{\mathbf{F}})$ to each other and to the rows of $\mathbf{Z}$. To achieve all these restrictions and satisfy the upper bound in (A.14), we conclude the overall system transfer matrix as

$$\mathbb{F}(z) = diag\left(\alpha_1 z^{-d_1}, \alpha_2 z^{-d_2}, \ldots, \alpha_p z^{-d_p}\right)\mathbf{P}, \tag{A.15}$$

i.e., the total system response is transparent. In this formula, $\mathbf{P}$ is a permutation matrix.

Appendix B

PROOF OF THEOREM-4

We can write $\mathbf{P}_{\mathbf{y}}(f)$ explicitly as

$$\mathbf{P}_{\mathbf{y}}(f) = \mathbb{F}(f)\mathbf{P}_{\mathbf{s}}(f)\mathbb{F}(f)^H \tag{B.1}$$

Then the first term in (4.25) turns into

$$\begin{aligned}\int_{-1/2}^{1/2} \log(\det(\mathbf{P}_{\mathbf{y}}(f)))df &= \int_{-1/2}^{1/2} \log(|\det(\mathbb{F}(f))|^2 \det(\mathbf{P}_{\mathbf{s}}(f)))df \\ &= \int_{-1/2}^{1/2} \log(|\det(\mathbb{F}(f))|^2)df + \int_{-1/2}^{1/2} \log(\det(\mathbf{P}_{\mathbf{s}}(f)))df \end{aligned} \tag{B.2}$$

The first term in (B.2) can be written with respect to Hadamard inequality rule as

$$\begin{aligned}\int_{-1/2}^{1/2} \log(|\det(\mathbb{F}(f))|^2)df &\leq \int_{-1/2}^{1/2} \log(\prod_{m=1}^{p} \|\mathbb{F}_{m,:}(f)\|_2^2)df \\ &= \sum_{m=1}^{p} \int_{-1/2}^{1/2} \log(\|\mathbb{F}_{m,:}(f)\|_2^2)df \end{aligned} \tag{B.3}$$

It is possible to take the logarithm term outside the integral by using Jensen's inequality [Rudin, 2006]:

$$\int_{-1/2}^{1/2} \log(\|\mathbb{F}_{m,:}(f)\|_2^2)df \leq \log(\int_{-1/2}^{1/2} \|\mathbb{F}_{m,:}(f)\|_2^2 df) \tag{B.4}$$

After that, using Parseval's theorem, we obtain the equation

$$\int_{-1/2}^{1/2} \|\mathbb{F}_{m,:}(f)\|_2^2 df = \|\tilde{\mathbf{F}}_{m,:}\|_2^2 \tag{B.5}$$

Accordingly, replacing the equality in eqn. (B.5) into eqn. (B.4) results

$$\int_{-1/2}^{1/2} \log(\|\mathbb{F}_{m,:}(f)\|_2^2) df \leq \log(\|\tilde{\mathbf{F}}_{m,:}\|_2^2) \tag{B.6}$$

and

$$\int_{-1/2}^{1/2} \log(|\det(\mathbb{F}(f))|^2) df \leq \sum_{m=1}^{p} \log(\|\tilde{\mathbf{F}}_{m,:}\|_2^2) \tag{B.7}$$

If we replace this inequality into (B.2), it becomes

$$\int_{-1/2}^{1/2} \log(\det(\mathbf{P}_{\mathbf{y}}(f))) df \leq \sum_{m=1}^{p} \log(\|\tilde{\mathbf{F}}_{m,:}\|_2^2) + \int_{-1/2}^{1/2} \log(\det(\mathbf{P}_{\mathbf{s}}(f))) df \tag{B.8}$$

As a result of the successive inequality steps for the first term, we reach the following inequality for the objective function:

$$J(\tilde{\mathbf{F}}) \leq \frac{N}{2} \sum_{m=1}^{p} \log(\|\tilde{\mathbf{F}}_{m,:}\|_2^2) + \frac{N}{2} \int_{-1/2}^{1/2} \log(\det(\mathbf{P}_{\mathbf{s}}(f))) df - Np \log(\max_{n \in \{1,\ldots,L_1\}} \|\tilde{\mathbf{y}}_N(n)\|_1) \tag{B.9}$$

The second term in (B.9) can be explicitly written as

$$(\max_{n \in \{1,\ldots,L_1\}} \|\tilde{\mathbf{y}}_N(n)\|_1)^{Np} = (\max_{n \in \{1,\ldots,L_1\}} \|\Gamma_N(\tilde{\mathbf{F}})\tilde{\mathbf{s}}_{L_3}(n)\|_1)^{Np} \tag{B.10}$$

where $L_3 = N + P - 1$, $\tilde{\mathbf{s}}_{L_3}(n) = [\mathbf{s}^T(n), \ldots, \mathbf{s}^T(n - L_3)]^T$, and $\Gamma_N(\tilde{\mathbf{F}})$ is the block Toeplitz matrix described in (4.8). Remembering that $\tilde{\mathbf{s}}_{L_3}(n)$ samples are inside the unity ℓ_1 norm ball, we can write the following inequality for the maximum term:

$$(\max_{n\in\{1,\ldots,L_1\}} \|\Gamma_N(\tilde{\mathbf{F}})\tilde{\mathbf{s}}_{L_3}(n)\|_1)^{Np} \leq \|\Gamma_N(\tilde{\mathbf{F}})\|_{1,1}^{Np} \tag{B.11}$$

If the assumption (**A**1) holds, this inequality turns into an equality,

$$(\max_{n\in\{1,\ldots,L_1\}} \|\Gamma_N(\tilde{\mathbf{F}})\tilde{\mathbf{s}}_{L_3}(n)\|_1)^{Np} = \|\Gamma_N(\tilde{\mathbf{F}})\|_{1,1}^{Np} \tag{B.12}$$

$\|\Gamma_N(\tilde{\mathbf{F}})\|_{1,1}^{Np}$ can be explicitly written as,

$$\|\Gamma_N(\tilde{\mathbf{F}})\|_{1,1}^{Np} = \| \left[\|\Gamma_N(\tilde{\mathbf{F}})_{:,1}\|_1, \cdots, \|\Gamma_N(\tilde{\mathbf{F}})_{:,L_3p}\|_1\right] \|_\infty^{Np}. \tag{B.13}$$

If we again use the equality in (A.7), we can put forward the following inequalities:

$$\begin{aligned}
&\| \left[\|\Gamma_N(\tilde{\mathbf{F}})_{1,:}\|_1, \cdots, \|\Gamma_N(\tilde{\mathbf{F}})_{Np,:}\|_1\right] \|_\infty^{Np} \\
&\qquad \geq (\| \left[\|\Gamma_N(\tilde{\mathbf{F}})_{1,:}\|_1, \cdots, \|\Gamma_N(\tilde{\mathbf{F}})_{Np,:}\|_1\right] \|_1/(Np))^{Np} && \text{(B.14)} \\
&\qquad \geq \|\Gamma_N(\tilde{\mathbf{F}})_{1,:}\|_1 \|\Gamma_N(\tilde{\mathbf{F}})_{2,:}\|_1 \ldots \|\Gamma_N(\tilde{\mathbf{F}})_{Np,:}\|_1 && \text{(B.15)} \\
&\qquad \geq \|\Gamma_N(\tilde{\mathbf{F}})_{1,:}\|_2 \|\Gamma_N(\tilde{\mathbf{F}})_{2,:}\|_2 \ldots \|\Gamma_N(\tilde{\mathbf{F}})_{Np,:}\|_2 && \text{(B.16)}
\end{aligned}$$

As a result, the objective inequality in (B.9) turns out to

$$J(\tilde{\mathbf{F}}) \leq N \sum_{m=1}^{p} \log(\|\tilde{\mathbf{F}}_{m,:}\|_2) + \frac{N}{2} \int_{-1/2}^{1/2} \log(\det(\mathbf{P}_\mathbf{s}(f)))df - \log(\prod_{m=1}^{Np} \|\Gamma_N(\tilde{\mathbf{F}})_{m,:}\|_2) \tag{B.17}$$

We note that $\prod_{m=1}^{Np} \|\Gamma_N(\tilde{\mathbf{F}})_{m,:}\|_2 = (\prod_{m=1}^{p} \|\tilde{\mathbf{F}}_{m,:}\|_2)^N$ since $\Gamma_N(\tilde{\mathbf{F}})$ is a block Toeplitz matrix as described in (4.8). Consequently, (B.17) turns into

$$J(\tilde{\mathbf{F}}) \leq \frac{N}{2} \int_{-1/2}^{1/2} \log(\det(\mathbf{P_s}(f)))df \tag{B.18}$$

The inequality between ℓ_1 and ℓ_∞ in (B.14) and the arithmetic-geometric mean inequality in (B.15) are achieved only if there is the same ℓ_1 norm in all the rows of $\tilde{\mathbf{F}}$ since $\Gamma_N(\tilde{\mathbf{F}})$ is a block Toeplitz matrix. The norm inequality in (B.16) is achieved only if there is only one non-zero entry in every row of $\tilde{\mathbf{F}}$. The Hadamard Inequality in (B.3) is achieved only if all the rows of $\mathbb{F}(f)$ are orthogonal to each other. And, since $\mathbb{F}(f)$ is the Fourier Transform of $\{\mathbf{F}(l); l \in \{0, \ldots, P-1\}\}$, the non-zero elements in the rows of $\tilde{\mathbf{F}}$ must be in different positions with respect to mod p in order to achieve (B.3). Consequently, the upper bound for $J(\tilde{\mathbf{F}})$ in (B.18) is achieved only if $\mathbb{F}(f) = diag\left(\alpha_1 e^{-j2\pi f d_1}, \alpha_2 e^{-j2\pi f d_2}, \ldots, \alpha_p e^{-j2\pi f d_p}\right)\mathbf{P}$ where $\mathbf{P}$ is a permutation matrix.

Appendix C

PROOF OF THEOREM-5

The PSD term in the nominator of (4.56) is defined in terms of the real isomorphic source PSD as

$$\Gamma\left(\mathbf{P}_{\mathbf{y}}(f)\right) = \Gamma\left(\mathbb{F}(f)\right)\Gamma\left(\mathbf{P}_{\mathbf{s}}(f)\right)\Gamma\left(\mathbb{F}(f)\right)^{T}. \tag{C.1}$$

Then, the nominator of the objective turns out to

$$\int\limits_{f=-1/2}^{1/2} \log\det(\Gamma\left(\mathbf{P}_{\mathbf{y}}(f)\right))df =$$

$$\int\limits_{f=-1/2}^{1/2} \left(\log|\det(\Gamma\left(\mathbb{F}(f)\right))|^2 + \log(\det(\Gamma\left(\mathbf{P}_{\mathbf{s}}(f)\right)))\right) df. \tag{C.2}$$

The determinant term in this expression can be arranged in column wise due to Hadamard inequality rule:

$$\begin{aligned}\log(|\det(\Gamma\left(\mathbb{F}(f)\right))|^2) &\leq \log(\prod_{k=1}^{2p} \|\Gamma\left(\mathbb{F}(f)\right)_{:,k}\|_2^2)\\ &= \sum_{k=1}^{2p} \log(\|\Gamma\left(\mathbb{F}(f)\right)_{:,k}\|_2^2)\end{aligned} \tag{C.3}$$

Using the inequality in (C.3), we reach the following expression for the objective function:

$$J(\mathbb{F}(f)) \leq \int\limits_{f=-1/2}^{1/2} \left(\sum_{k=1}^{2p} \log(\|\Gamma\left(\mathbb{F}(f)\right)_{:,k}\|_2)\right) \tag{C.4}$$

$$+\frac{1}{2}\log(\det(\Gamma(\mathbf{P_s}(f)))) - 2p\log(\max_{m\in L_v}\|\acute{\mathbf{y}}(m,f)\|_1)\Big)\, df$$

The volume term for the ℓ_1-norm ball in (C.4) is explicitly written as

$$(\max_{m\in L_v}\|\acute{\mathbf{y}}(m,f)\|_1)^{2p} = (\max_{m\in L_v}\|\Gamma(\mathbb{F}(f))\,\acute{\mathbf{s}}(m,f)\|_1)^{2p} \tag{C.5}$$

The fact that all source samples lie inside as ℓ_1 norm ball realizes the following inequality for the denominator of (4.56):

$$(\max_{m\in L_v}\|\Gamma(\mathbb{F}(f))\,\acute{\mathbf{s}}(m,f)\|_1)^{2p} \leq \|\Gamma(\mathbb{F}(f))\|_{1,1}^{2p}. \tag{C.6}$$

If the assumption (**A**5) holds, this inequality transforms to an equality

$$(\max_{m\in L_v}\|\Gamma(\mathbb{F}(f))\,\acute{\mathbf{s}}(m,f)\|_1)^{2p} = \|\Gamma(\mathbb{F}(f))\|_{1,1}^{2p}. \tag{C.7}$$

The term $\|\Gamma(\mathbb{F}(f))\|_{1,1}^{2p}$ can be reformulated as

$$\|\Gamma(\mathbb{F}(f))\|_{1,1}^{2p} = \|\left[\|\Gamma(\mathbb{F}(f))_{:,1}\|_1, \cdots, \|\Gamma(\mathbb{F}(f))_{:,2p}\|_1\right]\|_\infty^{2p}. \tag{C.8}$$

By using the mathematical relations between ℓ_∞, ℓ_1 and ℓ_2 norms,

$$\begin{aligned}
&\|\left[\|\Gamma(\mathbb{F}(f))_{:,1}\|_1, \ldots, \|\Gamma(\mathbb{F}(f))_{:,2p}\|_1\right]\|_\infty^{2p} \\
&\qquad \geq (\|\left[\|\Gamma(\mathbb{F}(f))_{:,1}\|_1, \ldots, \|\Gamma(\mathbb{F}(f))_{:,2p}\|_1\right]\|_1/(2p))^{2p}, && \text{(C.9)}\\
&\qquad \geq \|\Gamma(\mathbb{F}(f))_{:,1}\|_1\|\Gamma(\mathbb{F}(f))_{:,2}\|_1\cdots\|\Gamma(\mathbb{F}(f))_{:,2p}\|_1, && \text{(C.10)}\\
&\qquad \geq \|\Gamma(\mathbb{F}(f))_{:,1}\|_2\|\Gamma(\mathbb{F}(f))_{:,2}\|_2\cdots\|\Gamma(\mathbb{F}(f))_{:,2p}\|_2. && \text{(C.11)}
\end{aligned}$$

If we replace the inequality (C.11) into (C.4), the objective upper bound becomes

$$J(\mathbb{F}(f)) \leq \frac{1}{2}\int_{f=-1/2}^{1/2} \log(\det(\Gamma(\mathbf{P_s}(f))))df \tag{C.12}$$

The norm inequality in (C.9) and the arithmetic-geometric mean inequality in (C.10) can be achieved only if there is the same ℓ_1 norm in all the columns of $\Gamma(\mathbb{F}(f))$. The norm inequality in (C.11) can be achieved only if there is only one non-zero entry in all the columns of $\Gamma(\mathbb{F}(f))$. The Hadamard's Inequality in (C.3) is achieved only if $\Gamma(\mathbb{F}(f))$ have orthogonal columns. As a consequence, to reach the upper limit in (C.12), we conclude the system frequency response as $\Gamma(\mathbb{F}(f)) = \alpha\mathbf{P}$ where $\mathbf{P}$ is a permutation matrix.