MULTICHANNEL BLIND ESTIMATION TECHNIQUES:
BLIND SYSTEM IDENTIFICATION AND BLIND SOURCE SEPARATION

BY

KAMRAN RAHBAR

NOVEMBER 2002

A THESIS
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING
AND THE SCHOOL OF GRADUATE STUDIES
OF MCMASTER UNIVERSITY
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Multichannel Blind Estimation Techniques:

Blind System Identification and Blind Source Separation

Doctor of Philosophy (2002)	McMaster University
(Electrical & Computer Engineering)	Hamilton, Ontario

TITLE:	Multichannel Blind Estimation Techniques: Blind System Identification and Blind Source Separation
AUTHOR:	Kamran Rahbar B.Sc. (Electrical Engineering) AmirKabir University (Tehran PolyTechnique), Tehran , Iran
SUPERVISOR:	Dr. James P. Reilly
NUMBER OF PAGES:	xiv, 151

Abstract

The focus of this thesis is on blind identification techniques for multi-input, multi-output (MIMO) systems. In this respect we study three problems:

1. The joint diagonalization problem:

   Joint diagonalization is an efficient tool for blind identification techniques for MIMO systems. In this thesis we discuss new adaptive joint orthogonal diagonalization algorithms based on optimization methods over the Stiefel manifold.

2. Blind identification of MIMO systems:

   We demonstrate that by using the second-oder statistics of the system outputs, by exploiting the non-stationarity of sources, and some mild conditions on the sources and the system, the impulse response of the MIMO system can be identified up to an inherent scaling and permutation ambiguity. An efficient two-step frequency domain algorithm for identifying the MIMO system then has been proposed. Numerical simulations verify the theoretical results and the performance of the new algorithm.

3. Real room blind source separation problem:

   The final part of the thesis focuses on the practical problem of blind source separation of mixed audio signals in a real room. The new proposed algorithm exploits the non-stationarity of audio signals to separate them from their mixtures recorded in a reverberant environment. This method has successfully been applied to real data acquired during extensive recording experiments done in different office rooms on the McMaster campus.

Acknowledgements

I would like to express my gratitude to Dr. Jim Reilly for his excellent supervising and having confidence in my work throughout this past four years. His invaluable comments and support during the making of this thesis are highly appreciated.

I am also very grateful to Dr. Jonathan Manton for his assistance and insightful comments which have considerably improved the material presented in Chapter 3 of this thesis.

I would like to thank Dr. Tim Davidson and Dr. Marcel Joho for many insightful discussions. I am also grateful to Dr. Tom Luo who introduced me, through a course project, to the optimization techniques, used in Chapter 2 of this thesis.

I would like to acknowledge my fellow ECE graduate students, more notably Jean-René, Sabrina, Kaywan and Amin. I am grateful to them for their assistance and all the moments that we shared.

Further thanks go to ECE staff (Terry, Grace, Helen, Barb) and particulary Cheryl whom I am deeply grateful to and I really respect all her efforts and the superb job that she does as the ECE graduate secretary.

I am very grateful to my family, specially my parents and my brother Ali and sister-in-law Ramesh for their support and being there for me. Last but not least, I would like to thank Eliana for being such wonderful company and friend during this past four years. Her encouragement, support and kindness meant very much for me and are greatly appreciated.

The research reported in this thesis has been made possible by the financial support from the National Sciences and Engineering Research Council of Canada, CITO (Communications and Information Technology Ontario) and Mitel Corporation.

Notation and Acronyms

NOTATIONAL CONVENTIONS	
a	Scalar
$\boldsymbol{a}$	Vector
$\boldsymbol{A}$	Matrix
$\boldsymbol{A}^T$	Matrix transpose
$\boldsymbol{A}^\dagger$	Hermitian transpose
$\det(\mathbf{A})$	Determinant of **A**
$Tr(\mathbf{A})$	Trace of **A**
$E\{\cdot\}$	Expectation Operator
$\mathrm{diag}(\mathbf{a})$	Forms a diagonal matrix from the vector **a**
$\mathrm{diag}(\mathbf{A})$	Forms a column vector from the diagonal elements of **A**
$\mathrm{ddiag}(\mathbf{A})$	Forms a diagonal matrix from the diagonal elements of **A**
$\mathrm{vec}\{\mathbf{A}\}$	Forms a column vector by stacking the columns of **A**
$\mathrm{mat}\{\mathbf{a}\}$	Forms a $J \times J$ matrix from a $J^2 \times 1$ column vector **a**
$\mathbf{A}^+$	Pseudo Inverse of the matrix **A**
$\mathrm{Off}(\mathbf{A})$	Sum of squared off-diagonal values of **A**

$\|\|\mathbf{A}\|\|_F$	Frobenius norm of matrix $\mathbf{A}$
$\|\|\mathbf{a}\|\|_2$	Euclidean norm of vector $\mathbf{a}$
$\mathbb{R}$	Set of real numbers
$\mathbb{C}$	Set of complex numbers
$\mathcal{Z}$	Set of integer numbers
$\Re\{c\}$	Real part of the complex variable c
$\Im\{c\}$	Imaginary part of complex variable c

PRINCIPAL SYMBOLS

J	Number of observed signals
N	Number of sources
$\mathbf{H}(t)$	Impulse response of a MIMO system
$\mathbf{H}(\omega)$	Discrete time Fourier transform of $\mathbf{H}(t)$
$\mathbf{\Pi}$	Permutation matrix
$\mathbf{n}(t)$	Noise Vector
σ_n^2	Noise Variance
$\mathbf{s}(t)$	Vector of sources
$\mathbf{x}(t)$	Vector of observed signals
$\mathbf{y}(t)$	Vector of output signals
$\mathbf{R}_x$	Covariance matrix of the observations $\boldsymbol{x}$
$\mathbf{P}_x(\omega, m)$	Cross power spectral density matrix of $\mathbf{x}(t)$ evaluated at time epoch m

ABBREVIATIONS

ALS	Alternating least-squares
ALSP	Alternating least-squares with projection
AR	Autoregressive
BSS	Blind source separation
CDMA	Code division multiple accesss
CPSD	Cross Power Spectral Density

CRLB	Cramér-rao lower bound
DFT	Discrete Fourier transform
DTFT	Discrete time Fourier transform
FFT	Fast Fourier transform
FIR	Finite-duration impulse response
HOS	Higher-order statistics
Hz	Hertz
iid	Independent and identically distributed
IIR	Infinite-duration impulse responce
ISR	Interference-to-signal ratio
MIMO	Multi-input multi-output
ML	Maximum likelihood
mse	Mean squared error
pdf	Probability density function
s	Second
SDMA	Space division multiple access
SIMO	Single-input multi-output
SISO	Single-input single-output
SIR	Signal-to-interference ratio
SNR	Signal-to-noise ratio
SOS	Second-order statistics
SVD	Singular value decomposition
STFT	Short time Fourier transform
TITO	Two-input Two-output

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Multichannel blind estimation techniques are of great interest in many fields of study including signal processing, communication, biomedicine etc., mainly due to their potentially wide range of applications in all of these fields. In general these techniques address the following problem. Consider the unknown multi-input, multi-output system **A**, shown in Figure 1.1 where $s_1(t), \ldots, s_N(t)$, the inputs to the system, are inaccessible and only $x_1(t), \ldots, x_J(t)$, the outputs of the system, are observable. The objective is to identify the system **A** or its inputs or both using only the observed signals. The problem is called "blind" because no knowledge is available about the system **A** nor its input, and also to distinguish it from standard identification methods where either the system **A** or the inputs are known. Most

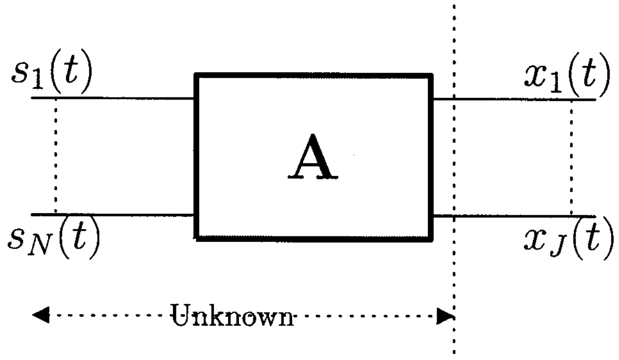

Figure 1.1: Blind estimation problem: **A** and $s_1(t), \ldots, s_N(t)$ are unknown.

of the known techniques for blind identification are presented under one of the following categories:

- *Blind Signal Separation:* In blind signal separation (Cardoso, 1998) the objective is to separate the sources $s_1(t), \ldots, s_N(t)$ which are mixed through an unknown system $\mathbf{A}$. For the blind source separation we require $N \geq 2$, and although in general there is no requirement on the number of output signals, the most common one, as is the case in this thesis, is that $J \geq N$.

- *Blind Equalization:* Blind equalization (Ding and Li, 2001) (Tong *et al.*, 1994)(Proakis, 2001), in its most common form is defined for the case where there is one source, $N = 1$, but the number of observed signals is one or more than one; i.e., $J \geq 1$. A common application of blind equalization is in communication systems. The objective of blind equalization is to recover the original source, $s(t)$, given only the observed signals $x_1(t), \ldots, x_J(t)$. For $N \geq 2$ the objectives of blind source separation and blind equalization become similar. Note that the desired objective in blind source separation is more flexible compared with the objective in blind equalization. More specifically, in blind source separation we may allow the outputs, although separated, to be a filtered version or a non-linearly distorted version of the original sources, while in blind equalization it is desired that the output be at worst a scaled and delayed version of the original sources.

- *Blind System Identification:* Blind system identification(Abed-Meraim *et al.*, 1997a) deals with the case where the objective is to identify the system $\mathbf{A}$ using only its output, without any access to the inputs. The commonly discussed form for blind identification is when $N = 1$ and $J \geq 1$. Fewer works discuss the more general form when $N > 1$ and $J > 1$.

All the problems discussed above are closely related to each other and solving one can often help to solve the others. In the following we discuss in detail the blind source separation and blind system identification problems for the case where $J \geq N \geq 2$. Note that blind equalization for this case can be considered as one of the solutions to the blind source separation problem.

1.1 The Blind Source Separation (BSS) Problem

The blind source separation problem can be explained through the following cocktail party example:

1.1.1 A cocktail party scenario

Consider a room where several people are talking simultaneously to each other and there is background music and other sources of sound such as moving fans etc. The sounds in the room are recorded using multiple microphones located randomly around the room. Note that the recorded sound is a mixture of different speech signals, music and sound, caused for example by the moving fans. In this example room acoustics can be considered as the unknown system **A**, the speech, music and noise generated by the moving fans are considered as the input to this system, and the recorded signal can be considered as the output of the system. For this scenario the objective of blind source separation is to separate all the different sounds from each other, given only the recorded signals and without any knowledge about the characteristics of the room nor the original sound sources.

1.1.2 Models

The most common model used in the blind source separation problem is the linear model; in other words, the system **A** is linear. The linear model by itself can be categorized into two major models:

1. *Instantaneous mixing*: In the instantaneous mixing model (Comon, 1994) (Bell and Sejnowski, 1995) (Cardoso and Laheld, 1996) we assume the system **A** is a matrix of real or complex scalars. In this case the outputs of the system **A** are a linear combination of the inputs and we use the following to describe the relationship between

the inputs and outputs of $\mathbf{A}$

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{1.1}$$

where $\mathbf{A} \in \mathbb{C}^{J \times N}$ is the mixing system, $\mathbf{s}(t) = (s_1(t), \ldots, s_N(t))^T$ are the sources, $\mathbf{n}(t) = (n_1(t), \ldots, n_N(t))^T$ represents the additive noise and $\mathbf{x}(t) = (x_1(t), \ldots, x_J(t))^T$ represent the observed signals. The most common assumptions used in instantaneous mixing models are

- $J \geq N \geq 2$.
- $\mathbf{A}$ has full column rank.
- The sources $s_1(t), \ldots, s_N(t)$ are statistically independent from each other.
- The noise $\mathbf{n}(t)$ is spatially and temporally white and is independent from the sources[1].

The objective in the instantaneous BSS problem is to find a separating matrix $\mathbf{W}$ whose outputs

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t) \tag{1.2}$$

are an estimate of $\mathbf{s}(t)$, the vector of source signals. In recovering the sources in the instantaneous mixing problem some indeterminacies may arise depending on the amount of a-priori knowledge available about the sources or the mixing system. If no a-priori knowledge is available about the ordering of sources and their power nor about the structure of the mixing system then one at best can estimate the sources up to a scaling and permutation ambiguity (Cardoso, 1998) (Tong *et al.*, 1991); in other words, the outputs of the separating matrix $\mathbf{W}$ are given as

$$y_i(t) = \alpha_{ij} s_j(t) + \eta_i(t) \quad i = 1, \ldots, N, \;\; j = 1, \ldots, N, \;\; \alpha_{ij} \in \mathbb{C} \tag{1.3}$$

where α_{ij} represent the scaling ambiguities that exist in recovering the sources and $\eta_i(t)$ is the additive noise due to $\mathbf{n}(t)$. This also can be expressed as

$$\mathbf{WA} = \mathbf{\Pi D} \tag{1.4}$$

[1] For simplicity most often it is assumed the additive noise has zero power which corresponds to a noise free model.

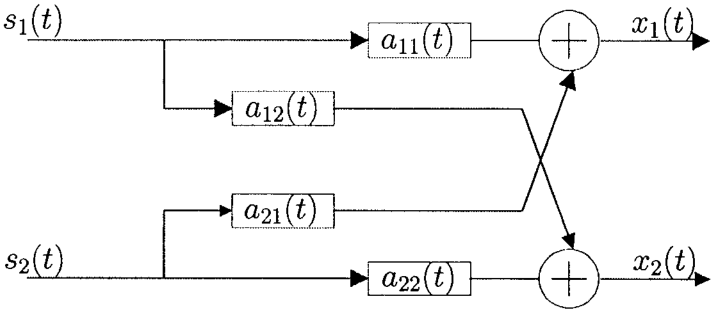

Figure 1.2: A 2 × 2 Convolutive mixing BSS problem.

where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{D}$ is a diagonal matrix.

2. *convolutive mixing*: In the convolutive mixing model (Sahlin and Broman, 2000) (Yellin and Weinstein, 1994) (Weinstein *et al.*, 1993) (Tugnait, 1999) the mixing system $\mathbf{A}$ is a matrix of complex or real filters and it can be written as:

$$\mathbf{A}(t) = \begin{pmatrix} a_{11}(t) & \dots & a_{1N}(t) \\ \vdots & \ddots & \vdots \\ a_{J1}(t) & \dots & a_{JN}(t) \end{pmatrix} \tag{1.5}$$

where $a_{ij}(t)$ is the impulse response of the ij_{th} filter in the mixing system. Figure 1.2 shows the convolutive blind source separation problem for a two-input, two-output system. The relationship between the outputs and inputs of the mixing system in a convolved BSS problem can be written as

$$\mathbf{x}(t) = \sum_{\tau=-\infty}^{\infty} \mathbf{A}(\tau)\mathbf{s}(t-\tau) + \mathbf{n}(t). \tag{1.6}$$

As can be seen, in the convolutive mixing problem the sources are convolved with the elements of the mixing system. Note that instantaneous mixing BSS can be considered as a special case of convolutive mixing BSS when $a_{ij}(t)$ are constant for all t.

The common assumptions used in convolutive mixing models are

- $J \geq N \geq 2$
- $a_{ij}(t)$ is causal and has a finite impulse response (FIR mixing model).

- $\mathbf{A}(\omega)$, the DTFT of $\mathbf{A}(t)$, has full column rank for all $\omega \in [0, 2\pi)$.
- The sources $s_1(t), \ldots, s_N(t)$ are statistically independent from each other.
- The noise $\mathbf{n}(t)$ is spatially and temporally white and is independent from the sources.

Similar to the instantaneous case, the objective in the convolutive BSS case is to find a separating matrix $\mathbf{W}(t)$ whose outputs

$$\mathbf{y}(t) = \sum_{\tau=-\infty}^{\infty} \mathbf{W}(\tau)\mathbf{x}(t-\tau) \tag{1.7}$$

are an estimate of the original sources. In the ideal case, the set of indeterminacies are similar to instantaneous case; i.e., without any a-priori knowledge, at best we can recover the sources up to a scaling and permutation ambiguity as given in (1.3). A less restrictive set of indeterminacies also exists by allowing the outputs to be a permuted and filtered version of the original sources; i.e.,

$$y_i(t) = \sum_{\tau=-\infty}^{\infty} h_{ij}(\tau)s_j(t-\tau) + \eta_i(t) \tag{1.8}$$

where $h_{ij}(\tau)$ represents the filter ambiguity which exists in recovering the source $s_j(t)$. The above makes the convolutive BSS case distinct from a multichannel blind equalization problem where the objective is to recover the sources up to a scaling and permutation ambiguity.

1.1.3 Applications

Blind source separation has a wide range of applications in audio, communication, mechanical vibration analysis, biomedical signal processing and computer imaging. In the audio application, blind source separation can be used for speech enhancement to separate the unwanted signals from the desired speech signals. Hands-free telephony, teleconferencing, music recording and hearing aid devices are some of the examples where blind source separation can be useful. See also the references in (Parra and Spence, 2000), (Torkkola, 1999), (Schobben and Sommen, 1998) for applications of blind source separation in audio. Blind source separation can also be used as a preprocessing stage in speech recognition devices to enhance the quality of the input speech. In communication systems in the case of a multi-antenna receiver, blind source separation can be used for separating SDMA signals coming from different locations but using the same frequency band and time slot(Feng and Kammeyer, 1999). See also (Sidiropoulos *et al.*, 1998) for applications of BSS in direct sequence CDMA systems. In biomedical signal processing, blind source separation also seems to have lots of applications. For example BSS can be used to remove the artifacts from noninvasive measurements of bioelectrical process such as EEG(electroencephalograph) or MEG(magnetoneurography). The main source of artifact in these measurements is the heart beat signal which is usually an order of magnitude higher than the signal of interest. Using blind source separation methods, the multichannel measurement vector can be transformed into a representation of independent components which then allow us to distinguish between the signal of interest and the artifacts. More information about this subject can be found in (Ziehe *et al.*, 1998). Also see (Makeig *et al.*, 2000) and (Jung *et al.*, 2000) for more applications of BSS in biomedical signal processing.

1.1.4 Approaches

In this section we give an introductory discussion on approaches that can be used for solving the blind source separation problem. One general approach is shown in Figure 1.3 where $\mathbf{s}$, $\mathbf{x}$ and $\mathbf{y}$ are random vectors representing respectively the sources, observed signals and the outputs of the separating system $\mathbf{W}$. As can be seen from the figure the separating system $\mathbf{W}$ is calculated by minimizing the contrast function $\phi[\mathbf{y}]$. Note that $\phi[\mathbf{y}]$ is a function of the statistics of $\mathbf{y}$ rather than the instantaneous value of random variable $\mathbf{y}$. In BSS terminology a contrast function is a real valued function of the statistics of the output signals such that its value is minimized when the outputs have been separated. In an instantaneous mixing model, if $\mathbf{A}$ is the mixing matrix and $\mathbf{W}$ is the separating matrix and we define the global system $\mathbf{C} = \mathbf{WA}$, then if $\phi[.]$ is a contrast function we expect

$$\phi[\mathrm{Cs}] \geq \phi[\mathrm{s}] \qquad (1.9)$$

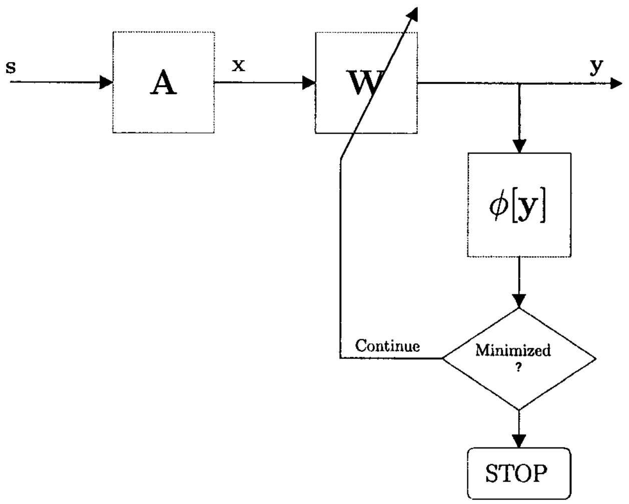

Figure 1.3: Approach I to the BSS problem.

with equality when $\mathbf{C} = \mathbf{\Pi D}$ where $\mathbf{D}$ is a diagonal matrix and $\mathbf{\Pi}$ is a permutation matrix. As can be seen, a contrast function in general is invariant to any scaling of the sources or any permutation of the order of the sources. Refer to (Comon, 1994) for a rigorous definition of the contrast functions. A good example for a contrast function in a blind source separation context is a probabilistic measure of the statistical independence of the outputs. The motivation behind this is based on the works in (Comon, 1994) where it is shown that if the random variables in s are statistically independent from each other and at most one of them is Gaussian then the random variables in $\mathbf{y} = \mathbf{Cs}$ become independent when $\mathbf{C} = \mathbf{\Pi D}$. In other words the independence of the outputs is equivalent to them being separated. Let $p(y_i)$ represent the probability density function of random variable y_i. From statistics we know that the random vector $\mathbf{y} = (y_1, \ldots, y_N)^T$ has mutually independent components if and only if

$$p(\mathbf{y}) = \prod_{i=1}^{N} p(y_i) \tag{1.10}$$

where $p(\mathbf{y})$ represents the joint probability density function of $y_1, \ldots, y_N$. Based on the relationship in (1.10) we can check the independence of the output random variable $\mathbf{y}$ by measuring how close the joint probability distribution of the outputs is to the product of their marginal probability distributions. In statistics, to measure such a distance between two probability density functions $p(\mathbf{x})$ and $q(\mathbf{x})$, we use the Kullback-Leibler distance defined as

$$\mathcal{K}(p|q) = \int_{\mathbf{x}} p(\mathbf{x}) \log\Big(\frac{p(\mathbf{x})}{q(\mathbf{x})}\Big) d\mathbf{x}. \tag{1.11}$$

An important property of the Kullback-Leibler distance is that $\mathcal{K}(p|q) \geq 0$ with equality if and only if $p(\mathbf{x})$ and $q(\mathbf{x})$ are equal. For an independence check we can choose $p(\mathbf{x}) = p(\mathbf{y})$ and $q(\mathbf{x}) = \prod_{i=1}^{N} p(y_i)$ and then using the Kullback-Leibler distance the independence measure can be defined as

$$I(\mathbf{y}) = \int_{\mathbf{y}} p(\mathbf{y}) \log\Big(\frac{p(\mathbf{y})}{\prod_{i=1}^{N} p(y_i)}\Big) d\mathbf{y}. \tag{1.12}$$

The above measure in information theory is known as *mutual information* and for random variables $y_1, \ldots, y_N$ is notated by the symbol $I(y_1, \ldots, y_N)$. Notice that $I(y_1, \ldots, y_N) \geq 0$ with equality if and only if $y_1, \ldots, y_N$ are mutually independent. Also note that mutual information is invariant with respect to scaling of random variables y_i and the permutation of their order. As can be seen the mutual information $I(\mathbf{y})$ has all the characteristics of a contrast function and it can be used as the $\phi[\mathbf{y}]$ in Figure 1.3, as long as the independence and non-Gaussianity assumptions are satisfied.

To calculate the mutual information one needs to estimate first the marginal probability distributions of the outputs and express them in terms of separating parameters of $\mathbf{W}$. In (Comon, 1994) the author suggests an approximate way of calculating the mutual information based on a fourth-order Edgeworth expansion of the probability density function. The method in (Pham, 1996) also uses the criterion in (1.12) to separate independent sources by replacing the unknown density functions in (1.12) with their kernel estimates.

Another approach for solving blind source separation problem is illustrated in Figure 1.4. This approach shows clearly the link between the blind source separation and the the blind system identification problem. As can be seen from the figure here we first estimate the mixing system $\mathbf{A}$ from the observed data $\mathbf{x}$ by minimizing an estimation function $F(\hat{\mathbf{A}}, x)$ with respect to $\hat{\mathbf{A}}$, an estimate of $\mathbf{A}$. Having the estimated value of $\mathbf{A}$ and assuming that $\mathbf{A}$ is invertible, we then separate the sources by applying the inverse of $\hat{\mathbf{A}}$ to the observed signals. An example of an estimation function for the instantaneous mixing scenario is the

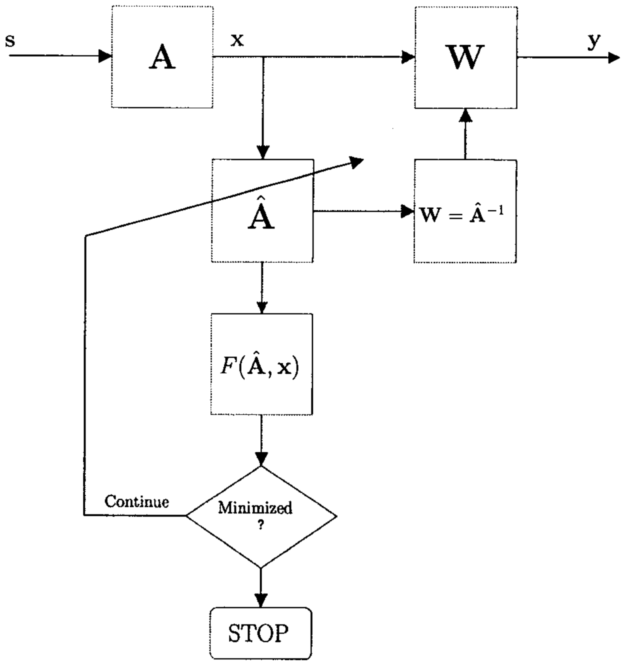

Figure 1.4: Approach II to the BSS problem.

following least-squares criterion

$$F(\hat{\mathbf{A}}, \mathbf{x}) = \sum_{l=0}^{L-1} ||\hat{\mathbf{R}}_x(l) - \hat{\mathbf{A}}\mathbf{\Lambda}(l)\hat{\mathbf{A}}^\dagger||_F^2 \tag{1.13}$$

where $\mathbf{\Lambda}(l)$ are unknown diagonal matrices with real diagonal values and $\hat{\mathbf{R}}_x(l)$ is an estimate of the covariance matrix of the observed signals evaluated at time lag l and is calculated from

$$\hat{\mathbf{R}}_x(l) = \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{x}(t)\mathbf{x}^\dagger(t+l), \tag{1.14}$$

where T is the observation time. Notice that $F(\hat{\mathbf{A}}, \mathbf{x}) \geq 0$. Also in the above criterion since both $\hat{\mathbf{A}}$ and $\mathbf{\Lambda}(l)$ are unknown, any scaling or permutation exchange between the columns of $\hat{\mathbf{A}}$ and diagonal values of $\mathbf{\Lambda}(l)$ does not change the value of F. When an exact estimate of $\mathbf{R}_x(l)$ is available then it can be shown that, under the conditions listed below, when $F(\hat{\mathbf{A}}, \mathbf{x}) = 0$ then $\hat{\mathbf{A}} = \mathbf{A\Pi D}$ where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{D}$ is a diagonal matrix. The identifiability conditions of $\mathbf{A}$ are as follows (see Chapter 2):

1. The mixing system $\mathbf{A}$ has full column rank.

2. Noise $\mathbf{n}(t)$ is spatially and temporally white with known variance.

3. $\mathbf{R}_s(l)$, the covariance matrix of the sources, is diagonal for all $l = 0, \dots, L-1$ and the vectors $\mathbf{d}(l) = \mathrm{vec}\{\mathbf{R}_s(l)\}$ are linearly independent.

Based on Condition 3, to identify $\mathbf{A}$, we require that the sources be uncorrelated over the range of $l = 0, \dots, L-1$, and also that their autocorrelation coefficients be linearly independent from each other. In practice only an approximate estimate of $\mathbf{R}_x(l)$ is available; in this case the optimum point of (1.13) is an approximate estimate of $\mathbf{A}$ up to a permutation and scaling ambiguity.

So far we have discussed methods for blind source separation for the instantaneous mixing. In general, blind source separation of convolved mixtures is more challenging compared to its instantaneous counterpart. At the same time relative to instantaneous mixing, the convolutive BSS algorithms are of higher interest because they are more likely to be used in a practical application. For example, separation of audio signals in a room environment is a convolutive mixing problem, due to the reflection of sound from the walls, furniture etc.

Here are some key differences between instantaneous and convolutive mixing algorithms:

- The number of unknown parameters in a convolved BSS problem is much higher than that in an instantaneous BSS problem. Note for an N-input, J-output mixing system, for instantaneous mixing the unmixing system is an $N \times J$ matrix with JN unknown elements. For convolutive mixing if we model the mixing system as a matrix of FIR filters, then JNL, where L is length of the mixing filters, parameters are needed to identify the mixing system. Note for the FIR mixing system, if one decides to directly estimate the unmixing system, then the number of unknown parameters is usually higher than JNL. This can explain one motivation to first estimate the mixing system and then use its inverse to recover the sources in a convolutive mixing problem.

- In instantaneous mixing at each time instant the observed signals are a linear combination of the sources, while in convolutive mixing the observed signals are given by the convolution of the mixing system and the sources. To recover the sources they

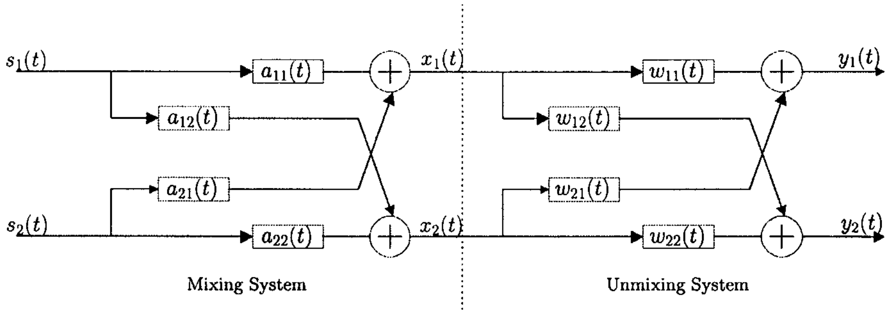

Figure 1.5: Example of a 2 × 2 convolutive mixing, unmixing systems.

need to be deconvolved from the unknown mixing system. Because of this, compared to the instantaneous mixing BSS methods, the convolutive BSS algorithms are more complicated with usually slower convergence and a higher chance of a suboptimal solution.

Most time-domain convolutive BSS methods assume an FIR model for the mixing and unmixing system (Chan *et al.*, 1996)(Reilly and Mendoza, 1999)(Gorokhov and Loubaton, 1997). This is shown in Figure 1.5 for a 2 × 2 mixing and unmixing system where $a_{ij}(t)$ and $w_{ij}(t)$ are FIR filters. Note that in general if the mixing system is an FIR system, one may not recover the sources up to a scaling and permutation ambiguity using an FIR unmixing system. Nevertheless it can be shown that one can always separate the sources up to a permutation and a filter ambiguity when using an FIR separating system. Let's assume that $\mathbf{A}(z)$ represents the transfer function of an N-input, N-output FIR mixing system with L being maximum order of its elements. Also let $\mathbf{A}_{adj}(z)$ represent the *adjoint* of $\mathbf{A}(z)$. Then

$$\mathbf{A}_{adj}(z)\mathbf{A}(z) = \mathbf{D}(z) \tag{1.15}$$

where $\mathbf{D}(z)$ is a diagonal matrix for all z and with diagonal elements all equal to $\det(\mathbf{A}(z))$, the determinant of $\mathbf{A}(z)$. Since $\mathbf{A}(z)$ is a polynomial FIR matrix its *adjoint* is also an FIR polynomial matrix with maximum element order equal to $L(N-1)$. Based on equation (1.15), we can see that for any polynomial mixing system $\mathbf{A}(z)$, there exists a polynomial unmixing system $\mathbf{A}_{adj}(z)$ which can separate the sources up to a filter ambiguity given by diagonal elements of $\mathbf{D}(z)$.

Only under the condition that the FIR mixing system $\mathbf{A}(z)$ has full column rank for all $z \in \mathbb{C}$ ($\mathbf{A}(z)$ is irreducible), can it be shown there exists an FIR unmixing system which not only separates the sources, but also recovers them up to some scaling and delay ambiguities (Gorokhov and Loubaton, 1997). This condition on $\mathbf{A}(z)$ guaranties the FIR invertibility of an FIR multi-input, multi-output system. For the special case where $\mathbf{A}(z)$ is a square polynomial matrix, irreducibility of $\mathbf{A}(z)$ can easily be verified from (1.15) by noticing that the above condition corresponds to $\det(\mathbf{A}(z))$ being constant for all $z \in \mathbb{C}$.

Time-domain algorithms for blind source separation under convolutive mixing can be designed using similar concepts discussed for instantaneous mixing algorithms. In this instance one can use a contrast function at the outputs of the unmixing system $\mathbf{W}(t)$ and minimize the value of the chosen contrast function with respect to all the elements of the unmixing filters $\mathbf{w}_{ij}(t)$. Some of contrast functions used in instantaneous mixing BSS can be extended to convolutive case. In (Yellin and Weinstein, 1994) the authors show that for blind source separation of convolved non-Gaussian *iid* sources, in a fashion similar to the instantaneous case, one can use the independence of the output signals to achieve separation. Note that in general not all the contrast functions used in the instantaneous mixing problem can be extended to the convolutive case. For example for instantaneously mixed sources, it is known that under the condition that the sources are uncorrelated colored signals with linearly independent autocorrelation coefficients, then the sum squared of the cross-correlation functions of the output signals of unmixing system, evaluated at different lags, can be used as a contrast function. In general assuming the same set of conditions on the sources, the above contrast function is not suitable for a convolutive mixing problem unless the mixing system has a specific structure (e.g it is column-wise coprime as has been discussed in (Hua and Tugnait, 2000)).

The main disadvantage of a time domain convolutive BSS approach is its slow convergence and computational cost. For a large scale convolutive mixing problem such as blind source separation of audio signals in a real reverberant environment where the size of the mixing filters can be a few thousand taps, time domain BSS algorithms are inefficient and impractical. A rather convenient way of solving the convolutive BSS problem is to use a frequency domain approach. The advantage of using a frequency domain rather than a time domain approach is that using the frequency domain method, one can decompose a time domain estimation problem, with a large number of parameters, into multiple, independent estimation problems, with much fewer parameters to be estimated at each frequency bin. As a result, in general the frequency domain estimation algorithms have a simpler implementation and better convergence properties.

Based on the model (1.6), in the frequency domain we have:

$$\mathbf{x}(\omega, m) = \mathbf{A}(\omega)\mathbf{s}(\omega, m) + \mathbf{n}(\omega, m) \tag{1.16}$$

where $\mathbf{x}(\omega, t)$, $\mathbf{s}(\omega, t)$ and $\mathbf{n}(\omega, t)$ are respectively the short-time Fourier transform (STFT) of $\mathbf{x}(t)$, $\mathbf{s}(t)$ and $\mathbf{n}(t)$ at time m and $\mathbf{A}(\omega)$ is the DTFT of $\mathbf{A}(\tau)$[2]. As can be seen using a frequency domain transformation, at each frequency we have an instantaneous BSS problem where the mixing matrix is given by the complex matrix $\mathbf{A}(\omega)$. Note that $\omega \in [0, 2\pi)$ is a continuous variable. In practice we use a discretize version of ω given as $\omega_k = \frac{2\pi k}{K}, \quad k = 0, \ldots, K-1$ where K is total number of frequency bins.

As shown in Figure 1.6, at each frequency bin we can use an instantaneous blind source separation algorithm to separate the sources. The final output then is calculated by taking an inverse Fourier transform of each of the separated outputs. The suggested frequency domain approach, although simple, has some serious drawbacks, which unless they are treated properly, can make the frequency domain algorithm impractical.

The main drawback of a frequency domain convolutive BSS approach is that at each frequency bin, the permutation and scaling of the separated outputs can be different from those of other frequency bins. Because of the random permutations across the frequency spectrum, even if the frequency domain outputs are separated at each frequency bin, after being transformed back to the time domain, the resulting outputs may not be separated at all. In addition, the effect of random scaling of the outputs across the frequency spectrum

[2] In practice we use the DFT(FFT) rather than the DTFT; in this case, (1.16) only approximately holds when the number of data samples (and as a result number of DFT samples) is much higher than the maximum order of elements of $\mathbf{A}(z)$.

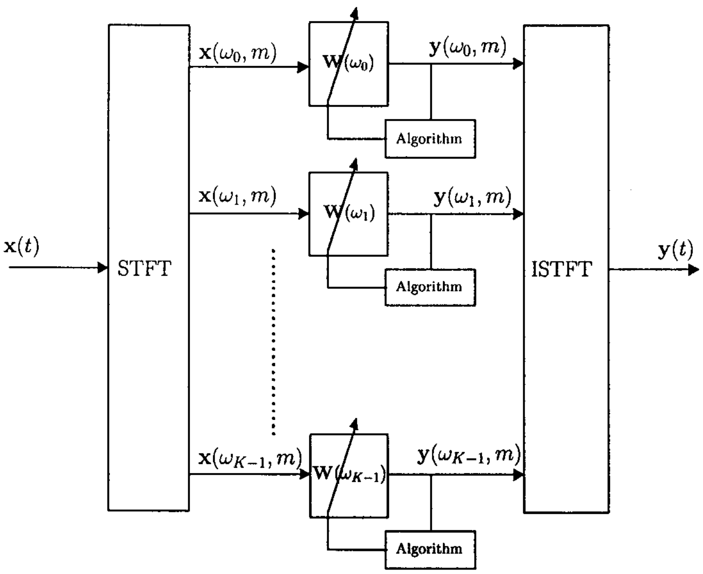

Figure 1.6: Frequency domain approach to BSS problem.

will appear as an arbitrary filtering of the output signals in the time domain. For audio sources these random filtering may severely distort the audio quality of the output signals.

Apart from random permutations and scaling factors, other problems may arise when one wants to implement the frequency domain structure shown in Figure 1.6 using off-the-shelf instantaneous BSS algorithms. For example, the group of instantaneous BSS algorithms which use independence and non-Gaussianity of the source signals to achieve separation may not perform well in the frequency domain structure shown in Figure 1.6, due to the fact that the signals at the output of an FFT process tend to become more Gaussian as the number of the FFT points increases (Serviere, 1998). The algorithms used in (Westner, 1998) are examples of the frequency domain scheme described above, used for separation of speech signals in real reverberant environments. In this case, as shown by experimental results, for some instances the quality of the output signals not only has not improved, but it has even been degraded compared to the quality of the original input mixed audio signals.

In Chapter 4 of this thesis we propose an alternative frequency domain approach shown in Figure 1.7. In this approach we use the time varying second-order cross spectral parameters of the observed signal to estimate the mixing system $\mathbf{A}(\omega_k)$ at each frequency bin.

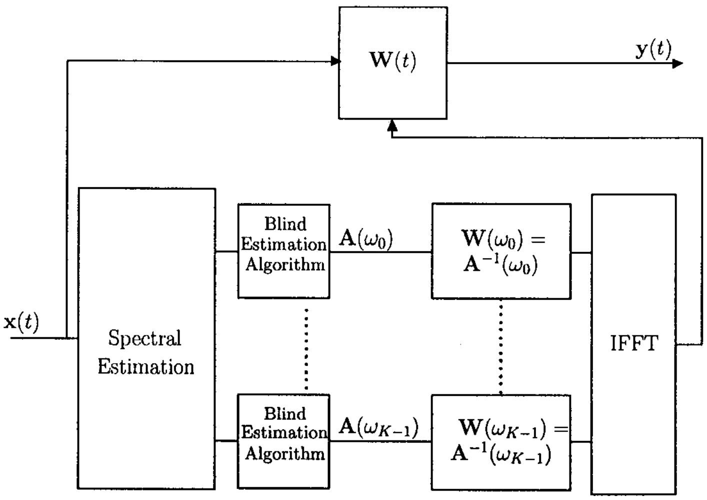
W(t)
y(t)
x(t)
Spectral Estimation
Blind Estimation Algorithm
A(ω_0)
W(ω_0) = A^{-1}(ω_0)
Blind Estimation Algorithm
A(ω_{K-1})
W(ω_{K-1}) = A^{-1}(ω_{K-1})
IFFT

Figure 1.7: Proposed frequency domain BSS algorithm for convolutive mixing.

The separating matrix $\mathbf{W}(t)$ is then obtained by calculating the inverse Fourier transform of $\mathbf{A}^{-1}(\omega_k)$. In this proposed algorithm, as has been discussed in Chapter 4 of thesis, we use an efficient method to prevent frequency dependent, arbitrary permutations of the columns of $\mathbf{A}(\omega_k)$. Also we propose a novel initialization procedure to alleviate the effect of frequency dependent scaling ambiguities in recovering the columns of $\mathbf{A}(\omega_k)$.

1.2 MIMO System Blind Identification

The multi-channel blind identification problem is closely related to the blind source separation problem, since in practice one can always identify the channel first and then use its inverse to recover the sources. Note that the blind identification problem is more general compared to the blind source separation problem. For example blind source separation does not apply to the case when there is only one source and one or multiple observed signals. On the other hand, one can apply blind identification techniques to identify the channel between each of the observed signals and each source in this case. Note that in the literature, some of the proposed methods for blind identification of multi-input, multi-output (MIMO) systems are extensions of the methods for single-input, multiple-output (SIMO) or single-input, single-output (SISO) blind identification methods, and in this respect they

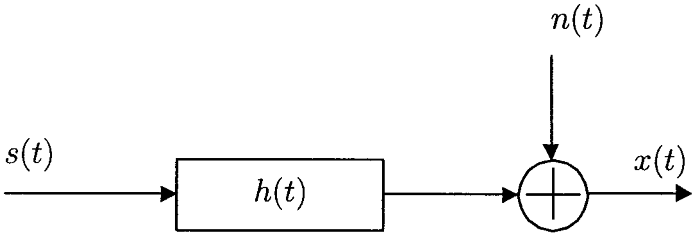

Figure 1.8: Example of a SISO blind system identification problem: objective is to identify $h(t)$ using only $x(t)$.

use a different approach than the commonly used methods for the blind source separation problem. In this section we first discuss the blind identification problem in its general form including some examples of its applications. We also discuss in more detail some of the existing approaches for solving this problem.

1.2.1 Problem Description

In its most common form, the objective of a blind system identification problem is to estimate the impulse response of a linear, time-invariant system $h(t)$, shown in Figure 1.8, using only the system output without access to the system input. Note that this is in contrast to classical system identification where both the input and the output of the unknown system are accessible. The most common system models in the blind identification literature are the SISO model (Tong *et al.*, 1994) (Giannakis and Mendel, 1989) (see Figure 1.8) and the SIMO model (Abed-Meraim *et al.*, 1997b) (Hua and Wax, 1996) (see Figure 1.9). Note that the SIMO model can be used for blind identification of a SISO system. For example, in digital communication, one known approach for blind identification of a SISO channel is to sample the received signal at a higher rate than the baud rate. In this case the processed received signal shows *cyclostationary* behavior, and equivalently it can be represented as a stationary process with an underlying SIMO model (Tong *et al.*, 1994).

There are two advantages of using a SIMO model rather than a SISO model. The first advantage is that it can be shown that a non-minimum phase[3] FIR SIMO system

[3] For blind identification of minimum phase FIR channels, second-order statistics are known to be sufficient even for SISO systems.

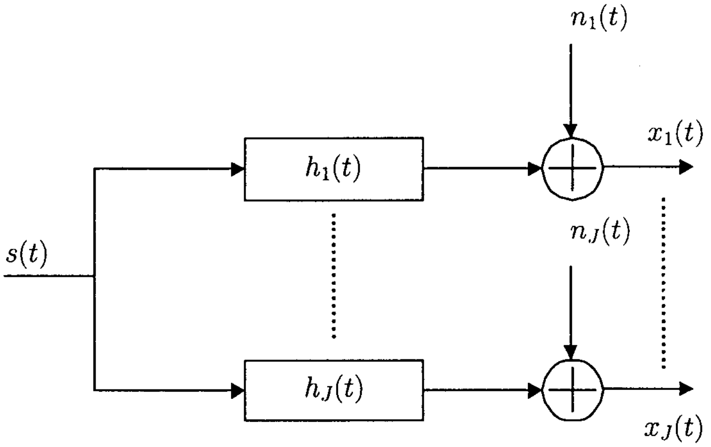

Figure 1.9: Example of a single-input, multiple-output blind system identification problem: objective is to identify $h_1(t), \dots, h_J(t)$ using $x_1(t), \dots, x_J(t)$.

can be identified using the second-order statistics of the observed signals if the channels $h_1(t), \dots, h_J(t)$ do not share common zeros. Another advantage of using the SIMO model is that under the above condition it is FIR invertible; in other words, if $h_1(t), \dots, h_J(t)$ are FIR filters which do not share common zeros, then based on the Bézout's identity (Kailath, 1980) (Vaidyanathan, 1993) there exists a multi-input, single-output system with set of FIR filters $g_1(t), \dots, g_J(t)$ such that the combined impulse response of the two systems is a pure delay. In literature there are a few works related to the more general case of the MIMO system. These works are mostly an extension from the SIMO case.

1.2.2 Problem Formulation

We consider the following N-input, J-output FIR mathematical model for the MIMO blind identification problem

$$\mathbf{x}(t) = \sum_{l=0}^{L-1} \mathbf{H}(l)\mathbf{s}(t-l) + \mathbf{n}(t) \tag{1.17}$$

where $\mathbf{H}(l) \in \mathbb{R}^{J\times N}$ is the unknown FIR system with a maximum element order of L, $\mathbf{s}(t) \in \mathbb{R}^{N\times 1}$ is the source vector, $\mathbf{x}(t) \in \mathbb{R}^{J\times 1}$ is the vector of observed signals, and $\mathbf{n}(t)$ is the additive noise vector. The objective is to identify the system $\mathbf{H}(l)$ using only the observed signals $\mathbf{x}(t)$ and without any knowledge about $\mathbf{s}(t)$.

- **The SIMO Model:**

For the SIMO model we have $N = 1$ and $J \geq 2$. For a block of Q consecutive samples (1.17) can be represented as

$$\mathbf{x}_q(t) = \mathbf{H}_q \mathbf{s}_q(t) + \mathbf{n}_q(t) \tag{1.18}$$

where $\mathbf{x}_q(t) \in \mathbb{R}^{JQ\times 1}$, $\mathbf{s}_q(t) \in \mathbb{R}^{(Q+L)\times 1}$, $\mathbf{n}_q(t) \in \mathbb{R}^{JQ\times 1}$ are respectively the observed signal, source and noise data matrices given as

$$\mathbf{x}_q(t) = \begin{pmatrix} \mathbf{x}(t) \\ \vdots \\ \mathbf{x}(t-Q+1) \end{pmatrix}, \quad \mathbf{s}_q(t) = \begin{pmatrix} s(t) \\ \vdots \\ s(t-Q-L+1) \end{pmatrix}, \quad \mathbf{n}_q(t) = \begin{pmatrix} \mathbf{n}(t) \\ \vdots \\ \mathbf{n}(t-Q+1) \end{pmatrix}, \tag{1.19}$$

and $\mathbf{H}_q \in \mathbb{R}^{QJ\times(L+Q)}$ is the system matrix having the following Sylvester structure

$$\mathbf{H}_q = \begin{pmatrix} \mathbf{h}(0) & \ldots & \mathbf{h}(L) & & \\ & \ddots & & \ddots & \\ & & \mathbf{h}(0) & \ldots & \mathbf{h}(L) \end{pmatrix} \tag{1.20}$$

where $\mathbf{h}(t) = (h_1(t), \ldots, h_J(t))^T$ is the impulse response vector of the SIMO system.

Most of the methods for blind identification of SIMO channels use one or all of the assumptions listed below (Tong and Perreau, 1998):

1. The subchannels $h_1(t), \ldots, h_J(t)$ are coprime; in other words, their z-transforms $h_1(z), \ldots, h_J(z)$ do not share common zeros.
2. Persistence of excitation of the sources, as defined in (Tong and Perreau, 1998)
3. The noise $\mathbf{n}(t)$ is zero mean, white with known variance.
4. The channel has known order.

Under the conditions 1 and 2 it can be shown that $\mathbf{h}(t)$ can be identified up to a constant factor from the noiseless observation $\mathbf{x}(t)$.

- **The MIMO Model**

The MIMO model, where $N \geq 2$ and $J \geq 2$, for a block of W consecutive samples can be represented as

$$\mathbf{x}_w(t) = \mathbf{H}_w \mathbf{s}_w(t) + \mathbf{n}_w(t) \tag{1.21}$$

where $\mathbf{x}_w(t) \in \mathbb{R}^{JW \times 1}$, $\mathbf{s}_w(t) \in \mathbb{R}^{(W+L)N \times 1}$, $\mathbf{n}_w(t) \in \mathbb{R}^{JW \times 1}$ are respectively observed signals, source and noise vectors and $\mathbf{H}_w \in \mathbb{R}^{JW \times (W+L)N}$ is the channel matrix given as

$$\mathbf{H}_w = \begin{pmatrix} \mathbf{H}(0) & \dots & \mathbf{H}(L) & & \\ & \ddots & & \ddots & \\ & & \mathbf{H}(0) & \dots & \mathbf{H}(L) \end{pmatrix} \tag{1.22}$$

where $\mathbf{H}(t)$ is the impulse response of the MIMO system.

Most of the MIMO blind identification methods make one or all of these assumptions on the system (Abed-Meraim *et al.*, 1997a)

1. $J \geq N$.
2. $\mathbf{H}(z)$ the z-transform of $\mathbf{H}(t)$ has full column rank for all z except $z = 0$.
3. $[\mathbf{h}_1(M_1)\ \mathbf{h}_2(M_2)\ \dots\ \mathbf{h}_N(M_N)]$ has full column rank where $\mathbf{h}_i(t)$ is the i_{th} column of $\mathbf{H}(t)$ with M_i being the maximum element order of $\mathbf{h}_i(z)$.
4. The elements of each column of $\mathbf{H}(z)$ do not share common zeros (column-wise coprimeness assumption).

Note that the first three assumptions also guarantee the invertibility of the MIMO FIR system $\mathbf{H}(z)$. Based on the last assumption (column-wise coprimeness), in (Hua and Tugnait, 2000) it is shown that when the sources are uncorrelated colored signals with distinct power spectra, then a MIMO channel can be identified up to a constant diagonal scaling and permutation matrix using only the second-order statistics of the observed signals. When the input signals are stationary white signals then second-order statistics are not enough to identify the system. Nevertheless it can be shown that under the first three assumptions in this case the MIMO system can be identified up to a scalar mixing system (Giannakis *et al.*, 2001).

1.2.3 Applications

The need for blind system identification arises from many applications, including data communication, echo cancellation for hands-free telephony, teleconferencing, speech recognition systems, image restoration and seismic signal processing.

The main application of blind system identification to data communication systems is to remove the intersymbol interference (ISI) caused by the communication channel's finite bandwidth. For an application which eliminates ISI, the channel response is first identified without using any training sequences. The identified channel is then used to equalize the received data and remove ISI. The fact that no training sequence is required is a major advantage of blind identification methods since non-blind methods for channel identification and equalization require a significant fraction of the channel capacity for sending training sequences.

For speech and audio signals the main application of blind system identification is for echo cancellation and dereverberation. In a hands-free telephone application, the speech signals received by the telephone's microphone may get distorted due to a reverberant room environment. To remove the reverberation effects an equalization process is used which by itself requires knowledge of the impulse response of the room. Since the impulse response of the room can change depending on the room characteristic and on the location of the handset inside the room, a blind system identification can be very useful in identifying the impulse response of a room which then can be used to equalize the received speech signal.

In image processing the main application of blind system identification is to restore a distorted image. In applications such as astronomy and medical imaging, blurring effects may occur which can be the result of camera motion during exposure or inaccurately focused lenses. The blurring effect can be represented by the convolution of the original undistorted image with the point spread function of the blurring system. Having the point spread function, one can restore the original image. Unfortunately in many practical situations the point spread function is unknown and so a blind system identification algorithm can be used to identify the blurring system that is used to restore the image.

Blind system identification has also application to seismology where the objective is to identify the physical characteristics of various layers in the earth. In this application, an explosion in the earth is used to create excitation signals and the resulting reflection and diffraction signals caused by different layers of the earth are measured through installed geophones. Since the exact waveform of the excitation signal responsible for generating the signal received through the geophones are unknown, identifying the impulse response of the system representing the various layers of earth is a blind system identification problem.

1.2.4 Approaches

The main focus of this section is on the theoretical approaches used for blind identification of MIMO systems. Since some of the MIMO blind identification algorithms are an extension of blind identification algorithms for SIMO systems, we start with describing some of the methods for blind identification of SIMO systems.

- **SIMO blind identification**: Many recent SIMO blind estimation techniques exploit the subspace structure of the observation signals. The key idea in subspace methods is that the channel vector lies in a unique direction specified by observation statistics or a block of noiseless observations. One of the attractive features of subspace methods is that most often a closed from solution can be found. The disadvantage is that they may not be robust against modelling error, especially when the channel matrix $\mathbf{H}_q$ is close to being singular. Another disadvantage of subspace methods is that, compared to other methods, they are usually computationally expensive.

  One of the frequently used approaches is the signal-noise subspace decomposition. From equation (1.18) we can write

$$\begin{aligned}\mathbf{R}_x &= \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbf{x}_q(t)\mathbf{x}_q^T(t)\\ &= \mathbf{H}_q\mathbf{R}_s\mathbf{H}_q^T + \sigma^2\mathbf{I}\end{aligned} \tag{1.23}$$

where $\mathbf{R}_s$ is the covariance matrix of the sources and σ^2 is the power of the noise. Note that if $h_1(t), \ldots, h_J(t)$ do not share common zeros then $\mathbf{H}_q$ has full column rank and its range space, which is the same as the signal subspace, is represented by the $L+Q$ dominant eigenvectors of $\mathbf{R}_x$ and is orthogonal to the noise subspace, spanned by remaining eigenvectors of $\mathbf{R}_x$. The motivation behind using a signal-noise subspace decomposition is based on the results in (Moulines *et al.*, 1995), where it is shown that if the channels $h_1(t), \ldots, h_J(t)$ do not share common zeros, then for $Q \geq L+1$ the range space of $\mathbf{H}_q$ uniquely determines $\mathbf{h}(t)$ up to a scaling factor. In practice, when the exact estimate of $\mathbf{R}_x$ is not available, using the orthogonality between the range space of $\mathbf{H}_q$ and the noise subspace, the channel vector can be estimated from the following least-squares criterion (Abed-Meraim *et al.*, 1997a)

$$\hat{\mathbf{h}} = \arg \min_{||\mathbf{h}||=1} ||\mathbf{E}_n^T \mathbf{H}_q||^2 \tag{1.24}$$

where $\mathbf{h} = (h_1(0), \ldots, h_1(L), h_2(0), \ldots, h_2(L), \ldots, h_J(0), \ldots, h_J(L))^T$ and $\mathbf{E}_n$ is a matrix which consists of the noise subspace eigenvectors of $\hat{\mathbf{R}}_x$, the sample estimate of $\mathbf{R}_x$.

- **MIMO blind identification:**

The subspace method described above can be extended to MIMO channels. Note that subspace MIMO channel blind identification algorithms require stricter conditions on the channel structure; e.g., $\mathbf{H}(z)$ should have full column rank for all z. This condition is equivalent to the Sylvester matrix $\mathbf{H}_w$ given in (1.21) having full column rank. Similar to subspace methods for the SIMO case, the left null space of $\mathbf{H}_w$ can be identified from the noise eigenspace of the observed signal covariance matrix. Note that contrary to the SIMO case where the null space of $\mathbf{H}_q$ is unique up to a constant scalar, the null space of $\mathbf{H}_w$ is unique only up to a constant invertible matrix. In other words if $\mathbf{A}$ is a constant invertible $N \times N$ matrix, then the left null space of $\mathbf{H}_w$ given in (1.22) and the matrix given below will be the same (Giannakis *et al.*, 2001):

$$\mathbf{H}_a = \begin{pmatrix} \mathbf{H}(0)\mathbf{A} & \ldots & \mathbf{H}(L)\mathbf{A} & & \\ & \ddots & & \ddots & \\ & & \mathbf{H}(0)\mathbf{A} & \ldots & \mathbf{H}(L)\mathbf{A} \end{pmatrix}. \tag{1.25}$$

Because of this, the subspace methods for MIMO blind identification can at best identify the MIMO system up to an unknown constant matrix $\mathbf{A}$. To identify $\mathbf{A}$, one can use the instantaneous blind source separation methods discussed in previous sections. Note that with this approach there may be problems, such as in a non-ideal situation $\mathbf{A}$ may not be a constant matrix but it could be time-variant. In this case, using an instantaneous BSS algorithm as the second stage will not be effective.

In (Hua and Tugnait, 2000) a method for blind identification of MIMO systems has been described which uses a bank of decorrelators to transform a MIMO system into a bank of SIMO systems. The resulting SIMO systems can be identified using the standard SIMO blind identification methods. The main assumptions that are used in this approach are that sources are colored signals with linearly independent spectra and the unknown system $\mathbf{H}(z)$ is column-wise coprime.

All the methods described above use the second-order statistics of the observed signal to identify the unknown system[4]. For the SIMO case, the second-order statistics subspace method described above is enough to completely reveal the underlaying system up to a scaling factor without making any assumptions on the sources. In other words the sources can be white or colored, Gaussian or non-Gaussian signals. For blind identification of MIMO systems on the other hand, second-order statistics may not be enough to completely identify the system without making assumptions on the sources or the system.

Higher-order statistics (HOS) methods are other alternatives which are usually used for blind identification of single-input, single-output systems. Compared with second-order statistics methods, the advantage of using higher-order statistics algorithms is that they usually require a less restrictive set of assumptions on the unknown system. Nevertheless

[4] For more references on SOS MIMO blind identification methods see also (Loubaton and Moulines, 2000), (Loubaton and Moulines, 1999), (Shen and Ding, 2001) and (Zhu *et al.*, 1998)

most of the higher-order statistical methods still require certain conditions on the statistics of the sources. Notice that in general non-Gaussianity is a necessary condition for all HOS methods. Some of HOS based blind identification methods exploit higher-order spectra to identify the system transfer function including the ones described in (Chen and Petropulu, 2001)(Shamsunder and Giannakis, 1997)(Tugnait, 1997).

The method in (Chen and Petropulu, 2001) uses the cross-bispectrum and cross spectrum of observed signal to identify the channel. The cross-bispectrum of a random process by definition is the two-dimensional discrete Fourier transform of the third order cumulant of that process[5]. Note that in general the third order cumulant of a symmetrically distributed random process is zero(Mendel, 1991). Due to this restriction the method in (Chen and Petropulu, 2001) requires that the sources be non-Gaussian, nonsymetrically distributed (for example the sources cannot be Gaussian or Laplacian distributed).

Apart from these limitations of HOS methods with respect to statistics of the sources, another disadvantage of the HOS methods is that they require large sample sizes for accurate time-averaged approximation of higher-order cumulants or spectra and they usually suffer from slow convergence due to the large estimation variance of the higher-order statistics (Abed-Meraim *et al.*, 1997b).

1.3 Non-stationarity, joint diagonalization methods and their applications to BSS and MIMO System Identification

In the previous section we discussed some of the second-order and higher-order statistics approaches for blind source separation and blind identification of MIMO systems. Note that in all the discussed methods, the sources are assumed stationary. In this thesis we discuss blind source separation and blind identification methods which exploit the non-stationarity of the observed signals.

- **Motivation for using a non-stationarity assumption**

[5] Third order cumulant of a random process $x(t)$ by definition is $C_{3,x}(\tau 1, \tau 2) = E[x(t)x(t+\tau 1)x(t+\tau 2]$. Also refer to (Mendel, 1991) for a tutorial on higher-order statistic (spectra) in signal processing.

A non-stationarity assumption can be justified by noticing that most real world signals including speech, biomedical signals etc. are inherently non-stationary. In communication systems non-stationarity in the form of *cyclostationarity* can be created by over-sampling the received signals (Tong *et al.*, 1994), or as suggested in (Serpedin and Giannakis, 1998), cyclostationarity can be induced in the transmitted signal.

We can use non-stationarity for instances where standard HOS or SOS blind system identification (or blind source separation) methods developed for stationary signals both fail; e.g., when the sources are both Gaussian distributed and temporally white.

In the previous section we mentioned some of the advantages of the second-order statistics methods compared to the higher-order statistics methods; e.g. SOS methods are insensitive to the sources' statistical distribution, they require less data samples and they usually have a simple implementation. The main disadvantage of SOS methods is that they require a more restrictive set of assumptions on the channel and the sources. By exploiting the non-stationarity of the input signals one can now develop second-order statistics solutions to problems where only higher-order statistics methods were previously applicable; e.g., blind identification of non-minimum phase MIMO systems driven by temporally white signals.

In large scale problems such as blind source separation of audio signals in a reverberant room, where the number of unknown parameters is usually very high, a frequency domain approach will be effective. As mentioned earlier, a major drawback of frequency domain approaches for blind source separation and blind MIMO system identification is the arbitrary frequency dependent permutation problem. As will be shown later in this thesis, non-stationarity of the sources can be used to eliminate this major problem of frequency domain methods.

- **Previous works**

In (Pham and Cardoso, 2001), (Souloumiac, 1995) and (Tsatsanis and Zhang, 2001), (Chang *et al.*, 2000) the authors propose non-stationary blind source separation algorithms for instantaneous mixing. In (Parra and Spence, 2000) a frequency domain algorithm is used for blind source separation of convolved non-stationary sources. The method presented in (Parra and Spence, 2000) has some limitations with respect to solving the frequency domain permutation problem. These limitations, as have been discussed recently in (Ikram and Morgan, 2000) and later on in (Ikram and Morgan, 2001) and (Araki *et al.*, 2001), degrade the performance of the algorithm in a long reverberant environment. The method in (Parra and Spence, 2000) also requires the diagonal elements of the convolutive mixing system to be constant. This is a rather strong condition on the mixing system and in practice will result in the separated outputs being a filtered version of the original sources.

- **Connection with the joint diagonalization problem**

The joint diagonalization problem, which has recently been discussed in (Cardoso and Souloumiac, 1993)(Pham, 2000)(Yeredor, 2000), has a close connection with blind source separation and blind MIMO identification problems. This connection is more evident when one uses non-stationarity to solve these blind identification problems.

The joint diagonalization problem can be expressed as follows. Assume that there exists a set of matrices $\mathbf{P}_1, \ldots, \mathbf{P}_M$ where $\mathbf{P}_i$ is a $J \times J$ real or complex matrix. Also assume that these matrices are related as

$$\mathbf{P}_m = \mathbf{A}\mathbf{\Lambda}_m\mathbf{A}^\dagger \quad m = 1, \ldots, M \tag{1.26}$$

where $\mathbf{A}$ is a $J \times N$ ($J \geq N$) unknown matrix and $\mathbf{\Lambda}(m)$ are $N \times N$ diagonal matrices which are also unknown. The objective in the joint diagonalization problem is to find a matrix $\mathbf{W}$ such that it jointly diagonalizes the set of matrices $\mathbf{P}_1, \ldots, \mathbf{P}_M$; i.e., we wish to find a $\mathbf{W}$ such that $\mathbf{W}\mathbf{P}_1\mathbf{W}^\dagger, \ldots, \mathbf{W}\mathbf{P}_M\mathbf{W}^\dagger$ are all diagonal matrices. Given $\mathbf{A}$ of full column rank, then one trivial solution to the above problem is to set $\mathbf{W} = \mathbf{A}^+$ where $\mathbf{A}^+$ is the pseudoinverse of $\mathbf{A}$. Note that in general there are many possible solutions to the above problem; e.g., if $\mathbf{W}$ is a joint diagonalizer of a set of $\mathbf{P}_1, \ldots, \mathbf{P}_M$ then $\mathbf{W}\mathbf{D}\mathbf{\Pi}$ where $\mathbf{D}$ is a diagonal matrix and $\mathbf{\Pi}$ is a permutation matrix will also be a joint diagonalizer of the same set.

Let $\mathbf{d}_m = \text{diag}\{\mathbf{\Lambda}(m)\}$ be vectors, organized from the diagonal elements of $\mathbf{\Lambda}(m)$. Then it can be shown that under the conditions that $\mathbf{A}$ has full column rank and the set of N-dimensional vectors $\mathbf{d}_m, \quad m = 1, \ldots, M$ span $\mathbb{R}^N$, then for any full column rank matrix $\mathbf{W}$ that jointly diagonalizes the set of matrices $\mathbf{P}_1, \ldots, \mathbf{P}_M$, defined in (1.26), we have[6]

$$\mathbf{WA} = \mathbf{\Pi D} \tag{1.27}$$

where $\mathbf{D}$ is a diagonal matrix and $\mathbf{\Pi}$ is a permutation matrix. We can also show that the same set of conditions guaranties a unique estimation of $\mathbf{A}$ up to some permutation and scaling ambiguity from the set of matrices $\mathbf{P}_1, \ldots, \mathbf{P}_M$. In other words if there is a matrix $\mathbf{B}$ and diagonal matrices $\tilde{\mathbf{\Lambda}}(m)$ such that

$$\mathbf{P}_m = \mathbf{B}\tilde{\mathbf{\Lambda}}(m)\mathbf{B}^\dagger \tag{1.28}$$

then we should have $\mathbf{B} = \mathbf{AD\Pi}$.

In a blind source separation context, assuming that the sources are non-stationary with time-varying variances, the quantities $\mathbf{P}_1, \ldots, \mathbf{P}_M$ can be considered as the set of covariance matrices of the observed signals evaluated at different time instances; in other words, we can set

$$\mathbf{P}_m = E[\mathbf{x}(t_m)\mathbf{x}^\dagger(t_m)] = \mathbf{AR}_s(m)\mathbf{A}^\dagger \quad m = 1, \ldots, M \tag{1.29}$$

where $\mathbf{x}(t_m)$ is the noiseless observed data at time instance t_m described by (1.1) and $\mathbf{R}_s(m)$ is the covariance matrix of sources at epoch m. Assuming that the sources are statistically independent from each other, $\mathbf{R}_s(m)$ is diagonal for all m. Based on the previous discussion we can easily see that if $\mathbf{A}$ and $\mathbf{R}_s(m)$ satisfy the identifiability conditions described above then $\mathbf{W}$, the joint diagonalizer of set of $\mathbf{P}_1, \ldots, \mathbf{P}_M$, is also the separating matrix; i.e., $\mathbf{WA} = \mathbf{\Pi D}$.

We can also extend the concept of joint diagonalization techniques to blind identification of MIMO systems. This is done by using the frequency domain model for MIMO

[6] Refer to Chapter 2 for a related Theorem and its proof.

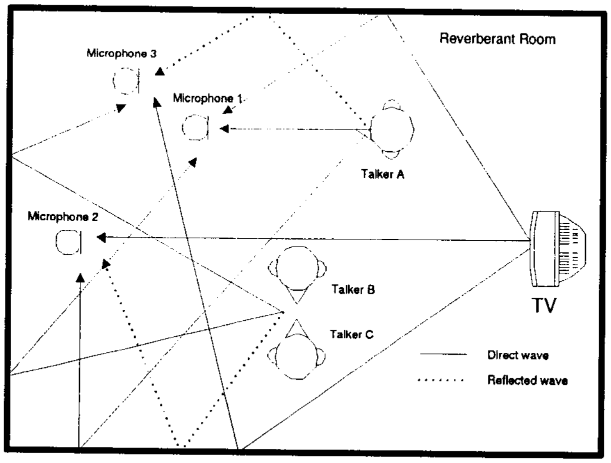

Figure 1.10: Illustration of a real room blind source separation problem with multipath effects between the talkers and the microphones.

blind identification. As will be discussed in this thesis, assuming a second-order non-stationary statistical model for sources, blind MIMO identification can be expressed as an extended version of the joint diagonalization problem where the objective is to estimate $\mathbf{H}(\omega)$, the DTFT of the channel matrix, from a set of functional matrices $\mathbf{P}_1(\omega), \dots, \mathbf{P}_M(\omega)$ related as

$$\mathbf{P}_m(\omega) = \mathbf{H}(\omega)\mathbf{\Lambda}(m)\mathbf{H}^{\dagger}(\omega), \qquad m = 1, \dots, M \tag{1.30}$$

where $\mathbf{\Lambda}(m)$ are diagonal matrices for all m.

1.4 Blind Source Separation for Real Acoustic Environments

In this section we consider blind source separation of audio signals in a real reverberant environment. This problem is of high practical interest. We start by explaining the problem and some of the difficulties that exist. We also discuss some of the limitations of the existing BSS algorithms in their application to real reverberant environments.

In many practical applications it is necessary to record the speech of a talker in a reverberant room. In some situations there may also be additional talkers or other sources of sound such as a TV inside the room (Figure 1.10). In this case the microphones will

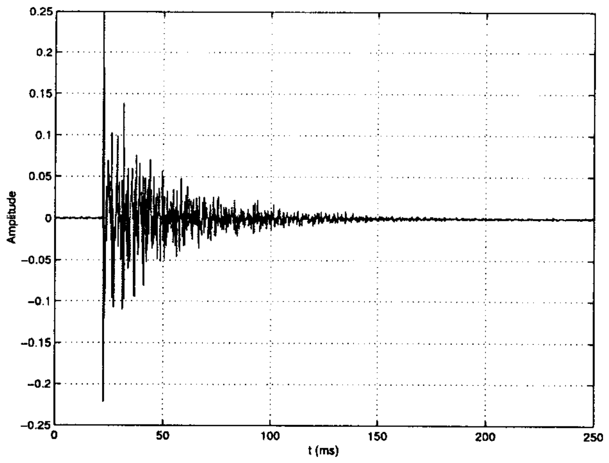

Figure 1.11: The measured impulse response between a sound source (speaker) and an omnidirectional microphone in an office room.

pick up a mixture of the direct sounds (if there are any) from the talkers and the TV, and also indirect sound waveforms caused by reflections from the walls or furniture inside the room. Figure 1.10 shows only one reflection between each sound source and each microphone. In practice there may be multiple reflections between the sound sources and the microphones. As the number of these reflections increases, it takes more time for the sound waves to reach the microphones and the energy of the sound waves reaching the microphone decreases. Due to these effects the impulse response between each sound source and microphone will appear as a decaying exponential. Figure 1.11 shows an example of a room acoustic impulse response, measured in a moderate reverberant room, between a sound source (a speaker) and an omnidirectional microphone with 1.5m spacing, using an 8.0 kHz sampling rate. Note that an omnidirectional microphone picks up the sounds from every direction, and as a result, the recorded signals are more reverberant compared to recordings done with a directional microphone. As can be seen from Figure 1.11, the measured impulse response is quite dense and rather long. To use blind source separation algorithms for mixed audio signals recorded in a reverberant room, there are a couple of issues that need to be considered, the most important ones being:

- The transfer functions between the sound sources and the microphones are typically non-minimum phase (Neely and Allen, 1979).

- The length of the impulse response even for a moderate reverberant environment is large. Even ignoring the small tails, these acoustic impulse responses can have a few thousand taps. Most of the existing blind source separation/blind MIMO identification methods cannot handle long impulse responses because of the computational cost, convergence problems, memory capacity etc. Also most of the blind source separation/MIMO blind identification methods assume an FIR model for the mixing filter, and for identifiability they need at least an over-estimate of the length of the channel. Notice that for room acoustics an over-estimate of the length of the mixing filters is not available because there is no way to measure when the impulse responses end. Of course one can always approximately measure the length of acoustics impulse responses by ignoring the very small tails. Note that the tails are most difficult to estimate and ignoring them as has been mentioned in (Wilbur, 2000), will greatly affect the perceptual quality of the recovered sound signals.

- The effect of sensor noise due to microphones or preamplifiers can affect algorithms that use a noise-free model. Also the microphones and preamplifiers are not perfectly linear. The non-linearity can degrade the performance of those algorithms that assume a linear model for BSS/MIMO blind identification problems.

- A real room environment can be a dynamic mixing system, caused e.g., when the sound sources are moving inside the reverberant room. This usually will put a constraint on the adaptation time of the algorithm and also on the number of data points needed for reliable identification of the mixing system or separation of the sources.

Above are some examples of the difficulties that may exist when one wants to apply a blind source separation method in a real reverberant environment. Most of the existing BSS algorithms only have been tested using computer simulations based on synthetically generated mixing systems and sources. Among these algorithms that do consider the convolutive mixing problem, the performance of the algorithms is demonstrated using synthetically generated convolutive mixing with the order of the mixing filters limited to few taps. There are only a few algorithms whose results are presented for a real reverberant room (Parra and Spence, 2000)(Ikram and Morgan, 2001). Nevertheless the reported performances of these methods are limited and do not exceed more than few dBs separation in a moderate reverberant office room using omnidirectional microphones.

1.5 Scope of the Thesis

This thesis contributes to the body of active research in blind source separation and blind MIMO identification problems. The main contribution of this thesis is to provide new insights in solving these two problems with some promising results in their real-world applications. Although the main focus of this thesis is on the blind source separation and MIMO blind identification problems, as side results, some additional contributed works are related to solving the algebraic problem of joint diagonalization, which has recently received a lot of attention and has immediate application to the two former problems.

A summary of the contributions can be listed as follows:

1. New algorithms for solving the joint diagonalization problem based on optimization methods over the Stiefel manifold have been proposed (Chapter 2).

    - New methods for solving the joint diagonalization problem based on gradient descent and conjugate gradient methods over the Stiefel manifold and their applications to second-order statistics blind source separation for the instantaneous mixing case.

    - Newton based algorithms for joint diagonalization of complex matrices with application to blind source separation of convolved and instantaneous mixtures.

    - A new maximum likelihood algorithm for joint orthogonal diagonalization using optimization methods over Stiefel manifold.

2. A novel frequency domain approach to blind identification of MIMO systems by exploiting the non-stationarity of sources has been developed (Chapter 3).

    - Sufficient identifiability conditions for blind identifiability of a MIMO system in the frequency domain under a second-order non-stationarity assumption of the inputs have been proved.
    - It has been proved that a limited number of frequency samples is sufficient to identify the channel and to this end an upper bound on the smallest number of frequency samples sufficient for blind identification of the MIMO system has been derived.
    - New frequency domain algorithms for blind identification of MIMO systems have been proposed.

3. New algorithms for blind source separation of convolved audio mixtures, which include the following features (Chapter 4).

    - New frequency domain algorithms based on the extension of joint diagonalization techniques to blind source separation of convolved non-stationary sources.
    - A new method for resolving the frequency domain permutation problem.
    - A new method for improving the audio quality of the separated output signals.
    - Successful application of the proposed blind source separation algorithm for blind separation of audio signals in a real reverberant environment.

1.6 Outline of Thesis

This thesis has been divided into three main chapters, plus two chapters for the Introduction and the Conclusions.

1. The first chapter introduces the blind source separation and blind identification problems. It outlines their applications, some of the main concepts for solving these two problems, some of the short-comings of the past approaches and how these two problems are related to the joint diagonalization problem. The chapter also discusses the difficult task of blind separation of audio signals in a real reverberant room.
2. The focus of the second chapter is on the joint diagonalization problem. A survey of past methods has been presented. Some identifiability results in connection with the blind source separation problem have also been discussed. Four new algorithms have been introduced including computer simulations to show their performance.
3. In Chapter 3 we consider blind identification of MIMO systems. The first part of this chapter establishes new MIMO channel identifiability results based only on the second-order statistics and the quasi-stationarity property of the input signals. The rest of the chapter discusses a new two-step frequency-domain algorithm for blind identification of MIMO systems. At the end, simulation results are provided which verify some of the theoretical arguments presented in this chapter.
4. Chapter 4 deals with blind source separation of convolved sources for audio application. The first part of this chapter discusses a new method for non-orthogonal joint diagonalization, and the second part of the chapter discusses its application to frequency domain blind source separation of convolved sources. A new diadic permutation algorithm has also been discussed which can be used to remove the arbitrary permutations across the frequency spectrum. The last part of this chapter is dedicated to the experimental results gathered from applying the new algorithm to recordings done in reverberant environments.
5. Chapter 5 gives Conclusions.

6. Appendices are included at the end of the thesis.

Chapter 2

The Joint Diagonalization Problem

In this chapter we discuss the joint diagonalization problem for real and complex matrices and its application to the blind source separation problem. We present new algorithms for orthogonal joint diagonalization using optimization techniques over the Stiefel manifold. Simulation results are provided to demonstrate the performance of the new algorithms and also to compare the proposed methods with existing joint diagonalization techniques.

2.1 Introduction

Joint diagonalization of matrices has direct application in the blind source separation problem. The problem first was introduced by Flury (Flury, 1984) as a method to find the common eigenvectors of a set of covariance matrices $\mathcal{R}_q = \{\mathbf{R}_m \in \mathbb{C}^{N\times N} | \mathbf{R}_m = \mathbf{Q}\mathbf{\Lambda}_m\mathbf{Q}^\dagger, \;\; 1 \leq m \leq M\}$ for some orthogonal matrix $\mathbf{Q} \in \mathbb{C}^{N\times N}$ and diagonal matrices $\mathbf{\Lambda}_1, \ldots, \mathbf{\Lambda}_M \in \mathbb{R}^{N\times N}$[1]. Later on in (Cardoso and Souloumiac, 1993) the authors reinstated the joint diagonalization problem as maximizing the following criterion with respect to the orthogonal matrix $\mathbf{Q}$

$$\mathcal{C}_d(\mathbf{Q}, \mathcal{R}_q) \equiv \sum_{m=1}^{M} ||\operatorname{diag}\{\mathbf{Q}^\dagger \mathbf{R}_m \mathbf{Q}\}||_2^2. \tag{2.1}$$

[1] Note in practice the set $\mathcal{R}$ may not be available, nevertheless it can be estimated from the available data samples. In this case the exact joint diagonalization of the sample estimates $\hat{\mathbf{R}}_1, \ldots, \hat{\mathbf{R}}_M$ may not feasible and the objective is to find a matrix $\hat{\mathbf{Q}}$ that approximately jointly diagonalizes $\hat{\mathbf{R}}_1, \ldots, \hat{\mathbf{R}}_M$.

Using the fact that the Frobenious norm of a matrix is invariant to orthonormal transformation of that matrix, it can be easily shown that maximizing the criterion in (2.1) with respect to $\mathbf{Q}$ is equivalent to minimizing the sum of off-diagonal values of all matrices $\mathbf{Q}^\dagger\mathbf{R}_m\mathbf{Q}$; i.e, minimizing

$$\mathcal{C}_o(\mathbf{Q},\mathcal{R}_q) \equiv \sum_{m=1}^{M} \mathrm{Off}(\mathbf{Q}^\dagger\mathbf{R}_m\mathbf{Q}) \tag{2.2}$$

is equivalent to maximizing the criterion in (2.1) where $\mathrm{Off}(\mathbf{Q}^\dagger\mathbf{R}_m\mathbf{Q})$ represents the sum of squared off-diagonal values of $\mathbf{Q}^\dagger\mathbf{R}_m\mathbf{Q}$. Notice that in general $\mathcal{C}_o(\mathbf{Q},\mathcal{R}_q) \geq 0$ and becomes zero when $\mathbf{Q}^\dagger\mathbf{R}_m\mathbf{Q}$ is diagonal for all m. In (Wax, 1997) it is shown that the criterion in (2.2) is equivalent to the following least-squares criterion

$$\mathcal{C}_{LS}(\mathbf{Q},\mathcal{R}_q) = \sum_{m=1}^{M} ||\mathbf{R}_m - \mathbf{Q}\mathbf{\Lambda}_m\mathbf{Q}^\dagger||_F^2 \tag{2.3}$$

for estimating $\mathbf{Q}$.

The common point among the criteria introduced above is that all assume that the joint diagonalizer matrix $\mathbf{Q}$ is orthogonal. Orthogonal joint diagonalization methods have also been discussed in (Flury, 1984) (Cardoso and Souloumiac, 1993) (Rahbar and Reilly, 2000) (Rahbar and Reilly, 2001a) (Joho and Rahbar, 2002). In (Cardoso and Souloumiac, 1993) the authors propose an extended Jacobi algorithm to achieve joint diagonalization while (Rahbar and Reilly, 2000) (Rahbar and Reilly, 2001a) propose adaptive algorithms, using gradient based optimization methods over the Stiefel manifold, to estimate the orthogonal diagonalizer matrix.

In recent years there has been some interest in non-orthogonal joint diagonalization methods; i.e., when the joint diagonalizer is not necessarily an orthogonal matrix. In a manner similar to the orthogonal case, the non-orthogonal joint diagonalization problem can be expressed as finding a matrix $\mathbf{W} \in \mathbb{C}^{N\times J}$ such that it jointly diagonalizes the set of matrices $\mathcal{R}_a = \{\mathbf{R}_m \in \mathbb{C}^{J\times J} | \mathbf{R}_m = \mathbf{A}\mathbf{\Lambda}_m\mathbf{A}^\dagger;\ 1 \leq m \leq M\}$ where $\mathbf{A} \in \mathbb{C}^{J\times N}$ in general is a non-square, non-orthogonal matrix. The criteria (2.2) and (2.3), discussed for the orthogonal joint diagonalization, can be directly extended to the non-orthogonal case by substituting the orthogonal matrix $\mathbf{Q}$ with the general-form matrix $\mathbf{A}$. Note that for

(2.2) some additional constraints are required to prevent the trivial solution $\mathbf{A} = \mathbf{0}$. In (Pham, 2000), the author proposes a method for non-orthogonal joint diagonalization of sample estimates of real covariance matrices. It is known that sample covariance matrix estimates $\hat{\mathbf{R}}_1, \ldots, \hat{\mathbf{R}}_M$ which are estimated from M independent populations of multivariate observations distributed according to zero-mean Gaussian probability density functions with true covariance matrices $\mathbf{R}_1, \ldots, \mathbf{R}_M$, are distributed according to the Wishart distribution with log-likelihood function is given as:

$$C - \frac{1}{2}\sum_{m=1}^{M} n_m[\log\det \mathbf{R}_m + Tr(\mathbf{R}_m^{-1}\hat{\mathbf{R}}_m)] \tag{2.4}$$

where C is a constant and n_m is the sample size used to estimate the covariance matrix $\hat{\mathbf{R}}_m$. Now considering that the covariance matrices $\mathbf{R}_m$ are related as $\mathbf{R}_m = \mathbf{A}\mathbf{\Lambda}_m\mathbf{A}^T \;\; m = 1\ldots, M$, for some real matrix $\mathbf{A} \in \Re^{N\times N}$ and diagonal matrices $\mathbf{\Lambda}_m$ with positive diagonal elements, the maximum likelihood method for estimating $\mathbf{A}$ from sample estimates $\hat{\mathbf{R}}_m$ corresponds to minimizing the following criterion(Pham, 2000)

$$\mathcal{C}_{Ml}(\mathbf{A}, \hat{\mathcal{R}}_q) = \sum_{m=1}^{M} n_m[\log\det \mathbf{\Lambda}_m + Tr(\mathbf{\Lambda}_m^{-1}\mathbf{W}^T\hat{\mathbf{R}}_m\mathbf{W}) - \log\det(\mathbf{W}\mathbf{W}^T)] \tag{2.5}$$

where $\mathbf{W}$ is pseudo-inverse of $\mathbf{A}$. $\mathbf{W}$ can be considered to be the matrix that jointly diagonalizes the set of covariance matrices $\hat{\mathbf{R}}_1, \ldots, \hat{\mathbf{R}}_M$.

The algorithm in (Pham, 2000) is based on minimizing an upper-bound of (2.5) using successive sweeps where each sweep includes pair-wise transformations of the rows of the matrix $\mathbf{W}$ such that at each transformation the criterion in (2.5) decreases. The algorithm performs iterative sweeps until convergence is achieved. Another iterative method is the ACDC algorithm in (Yeredor, 2000) which uses a least-squares criterion similar to (2.3) for non-orthogonal diagonalization of a set of complex symmetric matrices[2]. Also the method in (van der Veen, 2001) solves the joint diagonalization problem via weighted subspace fitting techniques by minimizing a criterion similar to (2.3) using a Gauss-Newton optimization algorithm.

[2] See also (Yeredor, 2002) for a more in depth discussion of ACDC method.

It should be noted that in general for the joint diagonalization problem there is no unique solution. This can be easily seen by looking at the criteria discussed above. For example for the non-orthogonal criterion of (2.2), if $\mathbf{Q}_{opt}$ is the optimum minimizer of this criterion then $\mathbf{D\Pi Q}_{opt}$, where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{D}$ is diagonal matrix, is also a minimizer of (2.2).

In this chapter we first show some applications of joint diagonalization techniques to the blind source separation problem. We also derive some previously known identifiability results for second-order blind source separation methods in a new context using identifiability results for the joint diagonalization problem. We then discuss developing new adaptive, orthogonal joint diagonalization algorithms based on optimization methods over the Stiefel manifold. At the end of this chapter we also discuss briefly the non-orthogonal joint diagonalization problem. Simulation results are provided to demonstrate the performance of the new algorithms and comparisons are made to some existing joint diagonalization methods.

2.2 Joint Diagonalization and BSS

In this section we give an example of how the joint diagonalization can be used for blind source separation of instantaneous mixtures using only second-order statistics of the observed signals.

We consider the following instantaneous mixing model for the BSS problem

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{2.6}$$

where $\mathbf{x}(t) \in \mathbb{C}^{J\times 1}$ is the observed signals, $\mathbf{A} \in \mathbb{C}^{J\times N}$ is the mixing system, $\mathbf{s}(t) \in \mathbb{C}^{N\times 1}$ are the source signals (assumed to have zero mean), and $\mathbf{n}(t) \in \mathbb{C}^{J\times 1}$ is the additive noise, which for now we consider to be white, Gaussian, with zero mean. The objective of the source separation problem is to find a $\mathbf{W} \in \mathbb{C}^{N\times J}$ such that

$$\mathbf{WA} = \mathbf{\Pi D} \tag{2.7}$$

where $\mathbf{\Pi}$ is a $N \times N$ permutation matrix and $\mathbf{D} \in \mathbb{C}^{N\times N}$ is a diagonal matrix. To estimate the $\mathbf{W}$ we can use the second-order statistics of the observed signal $\mathbf{x}(t)$. To this end we use the covariance matrices of the observed signal evaluated for a range of lags.

$$\begin{aligned}\mathbf{R}_x(l) &= \mathbf{A}\mathbf{R}_s(l)\mathbf{A}^\dagger + \sigma^2\mathbf{I} \quad l = 0 \\ &= \mathbf{A}\mathbf{R}_s(l)\mathbf{A}^\dagger \quad 0 < l \leq L-1\end{aligned} \tag{2.8}$$

where $\mathbf{R}_s(l)$ is the covariance matrix of the sources at time lag l and σ^2 is the power of the noise. Assuming that the sources are uncorrelated then $\mathbf{R}_s(l)$ is a diagonal matrix for all l. We now show that under some assumptions on the sources and the mixing system, the separation matrix $\mathbf{W}$ can be obtained by joint diagonalization of set of covariance matrices

$$\mathcal{R}_x = \{\mathbf{R}_x(0) - \sigma^2\mathbf{I}, \mathbf{R}_x(1), \ldots, \mathbf{R}_x(L-1)\}. \tag{2.9}$$

Theorem 1 *Consider the set of matrices*

$$\mathcal{R} = \{\mathbf{R}_m \in \mathbb{C}^{J\times J} | \mathbf{R}_m = \mathbf{A}\mathbf{\Lambda}_m\mathbf{A}^\dagger \;\; m = 0, \ldots, M-1\} \tag{2.10}$$

where $\mathbf{A} \in \mathbb{C}^{J\times N}$ *is some full column rank matrix and* $\mathbf{\Lambda}_m \in \mathbb{R}^{N\times N}$ *are diagonal matrices such that the set of vectors* $\boldsymbol{\lambda}_m = \mathrm{diag}\{\mathbf{\Lambda}_m\}$ *spans* $\mathbb{R}^N$. *Now if the full row rank matrix* $\mathbf{W} \in \mathbb{C}^{J\times N}$ *is the joint diagonalizer of the set* $\mathcal{R}$ *such that*

$$\mathbf{W}\mathbf{R}_m\mathbf{W}^\dagger = \tilde{\mathbf{\Lambda}}_m \quad \forall \; m = 0, \ldots, M-1 \tag{2.11}$$

where $\tilde{\mathbf{\Lambda}}_m$ *are diagonal matrices; then we have:*

$$\mathbf{W}\mathbf{A} = \mathbf{\Pi}\mathbf{D} \tag{2.12}$$

where $\mathbf{\Pi} \in \mathbb{R}^{N\times N}$ *is a permutation matrix and* $\mathbf{D} \in \mathbb{C}^{N\times N}$ *is a non-singular diagonal matrix.*

Proof:

From (2.11) by substituting $\mathbf{R}_m$ with $\mathbf{A}\mathbf{\Lambda}_m\mathbf{A}^\dagger$ we can write

$$\mathbf{W}\mathbf{A}\mathbf{\Lambda}_m\mathbf{A}^\dagger\mathbf{W}^\dagger = \tilde{\mathbf{\Lambda}}_m \;\; m = 0, \ldots, M-1. \tag{2.13}$$

Based on (2.13) for any sequence of scalars $a = (a_0, ..., a_{M-1})$ we can write

$$\sum_{m=0}^{M-1} a_m(\mathbf{W}\mathbf{A}\mathbf{\Lambda}_m\mathbf{A}^\dagger\mathbf{W}^\dagger) = \sum_{m=0}^{M-1} a_m\tilde{\mathbf{\Lambda}}_m. \tag{2.14}$$

Now defining

$$\mathbf{\Sigma}_a = \sum_{m=0}^{M-1} a_m\mathbf{\Lambda}_m, \ \tilde{\mathbf{\Sigma}}_a = \sum_{m=0}^{M-1} a_m\tilde{\mathbf{\Lambda}}_m, \quad a_m \in \mathbb{R}, \tag{2.15}$$

and $\mathbf{C} = \mathbf{W}\mathbf{A}$ and rearranging the terms in (2.14) we have

$$\mathbf{C}\mathbf{\Sigma}_a\mathbf{C} = \tilde{\mathbf{\Sigma}}_a. \tag{2.16}$$

Since by assumption $\mathbf{A}$ and $\mathbf{W}$ have full rank, then $\mathbf{C}$ is a full rank matrix. Also since the diagonal values of $\mathbf{\Lambda}_m$ span $\mathbb{R}^N$, $\mathbf{\Sigma}_a$ can be made equal to any real valued diagonal matrix by an appropriate choice of a. In this instance for any i, choose a such that all elements of $\mathbf{\Sigma}_a$ are zero except for the i_{th} diagonal element which is unity. Then $\mathbf{C}\mathbf{\Sigma}_a\mathbf{C} = \mathbf{c}_i\mathbf{c}_i^\dagger$ where $\mathbf{c}_i$ is the i_{th} column of $\mathbf{C}$. Moreover, since the RHS of (2.16) is diagonal, all the off-diagonal elements of $\mathbf{c}_i\mathbf{c}_i^\dagger$ are zero. Because $\mathbf{c}_i\mathbf{c}_i^\dagger$ has rank one at most, it can have at most one non-zero diagonal element. It follows immediately that every column of $\mathbf{C}$ has precisely one non-zero element, and moreover, because $\mathbf{C}$ is invertible, every row has precisely one non-zero element too; i.e.,

$$\mathbf{C} = \mathbf{\Pi}\mathbf{D} \tag{2.17}$$

where $\mathbf{D}$ is some non-singular diagonal matrix and $\mathbf{\Pi}$ is a permutation matrix. Equation (2.12) follows immediately from (2.17). □

If we substitute the set of $\mathcal{R}$ in Theorem 1 with set of $\mathcal{R}_x$ from (2.9) we can easily deduce the identifiability conditions for instantaneous blind source separation based on second-order statistics. The first condition is that the mixing matrix $\mathbf{A}$ should have full column rank. The next condition is that the set of vectors $\text{diag}\{\mathbf{R}_s(l)\}$ $l = 0, \ldots, L-1$ should span $\mathbb{R}^N$, which by itself means that the autocorrelation coefficients of the sources, $r_{s_i}(l)$ $l = 0, \ldots, L-1$ $i = 1, \ldots, N$, should be mutually lineally independent. This is a known identifiability result which also has been reported in (Meraim *et al.*, 2000). However here we used joint diagonalization theorem to prove it. Using Theorem 1 we can also see that a sufficient minimum number of covariance matrices required is N, where N is the number of the sources. In practice we only have access to $\hat{\mathbf{R}}_x(l)$, the sample estimates of $\mathbf{R}_x(l)$. Hence the joint diagonalization will be approximate. Because of this in practice we need more than N covariance matrices to get a good estimate of $\mathbf{W}$.

2.2.1 Orthogonal Joint Diagonalization and BSS

In the blind source separation problem, since there is an inherent ambiguity in recovering the scale of the sources, without loss of generality we can assume the sources have unit power. We base the development on the assumption that the covariance matrix of the observed signal at lag zero can be written as:

$$\begin{aligned}\mathbf{R}_x(0) &= \mathbf{A}\mathbf{R}_s(0)\mathbf{A}^\dagger + \sigma^2\mathbf{I} \\ &= \mathbf{A}\mathbf{A}^\dagger + \sigma^2\mathbf{I}\end{aligned} \tag{2.18}$$

where we have set $\mathbf{R}_s(0) = \mathbf{I}$. Replacing $\mathbf{A}$ with its singular value decomposition

$$\mathbf{A} = \mathbf{U}\begin{pmatrix} \mathbf{\Sigma}_{N\times N} \\ \mathbf{0}_{(J-N)\times N} \end{pmatrix}\mathbf{V}^\dagger, \quad \mathbf{U} \in \mathbb{C}^{J\times J}, \quad \mathbf{V} \in \mathbb{C}^{N\times N} \tag{2.19}$$

(2.18) can be written as:

$$\mathbf{R}_x(0) = \mathbf{U}\begin{pmatrix} \mathbf{\Sigma}^2 + \sigma^2\mathbf{I} & \mathbf{0}_{(J-N)\times(J-N)} \\ \mathbf{0}_{(J-N)\times(J-N)} & \sigma^2\mathbf{I} \end{pmatrix}\mathbf{U}^\dagger \tag{2.20}$$

As can be seen from (2.20) for $J > N$, σ^2 can be estimated from the $J - N$ smallest eigenvalues of $\mathbf{R}_x(0)$. Nevertheless, assuming σ^2 is known, we can obtain $\mathbf{\Sigma}$ from N dominant eigenvalues of $\mathbf{R}_x(0) - \sigma^2\mathbf{I}$ and from there we can define the whitening matrix $\mathbf{K}$ as

$$\mathbf{K} = \begin{pmatrix}\mathbf{\Sigma}^{-1} & \mathbf{0}_{N\times(J-N)}\end{pmatrix}\mathbf{U}^\dagger. \tag{2.21}$$

By applying $\mathbf{K}$ to the observed signals we have

$$\begin{aligned}\mathbf{z}(t) &\triangleq \mathbf{K}\mathbf{x}(t) = \mathbf{K}\mathbf{A}\mathbf{s}(t) + \mathbf{K}\mathbf{n}(t) \\ &= \mathbf{V}^\dagger\mathbf{s}(t) + \mathbf{K}\mathbf{n}(t).\end{aligned} \tag{2.22}$$

As can be seen using the whitening stage, the BSS problem can be simplified to finding an orthogonal matrix $\mathbf{V}$. Note the covariance matrix of the whitened data is now given as

$$\mathbf{R}_z(l) = \mathbf{V}^{\dagger}\mathbf{R}_s(l)\mathbf{V} + \sigma^2\mathbf{K}\mathbf{K}^{\dagger}, \qquad l = 0, \ldots, L-1. \tag{2.23}$$

We can estimate $\mathbf{V}$ by joint approximate diagonalization of the set $\mathcal{R}_z = \{\mathbf{R}_z(1), \ldots, \mathbf{R}_z(L-1)\}$. Note that assuming that a perfect estimate of $\mathbf{R}_z(l)$ is available then only in the noiseless case or at least when σ^2 is known can we exactly diagonalize $\mathcal{R}_z$.

The SOBI algorithm, explained in (Belouchrani *et al.*, 1997), is a second-order statistics blind source separation method which uses orthogonal diagonalization of set of whitened covariance matrices $\mathcal{R}_z$. The algorithm is based on an extension of Jacobi algorithm ((Golub and VanLoan, 1996)) to the joint approximate diagonalization of a set of covariance matrices. In the next section we propose an alternative approach for joint orthogonal diagonalization using optimization methods over the Stiefel manifold. In general the geometry of an orthogonality constraint can be represented by the Stiefel manifold. By exploiting this geometry, we can solve optimization problems with orthogonality constraints using unconstrained optimization methods (such as gradient descent) over the Stiefel manifold. Notice that an advantage of using such an optimization method is that we can develop adaptive joint diagonalization algorithms that can, for example, track the variation of a set of covariance matrices.

2.3 Joint Approximate Diagonalization Based on Geometric Optimization Methods

In this section we discuss how the joint diagonalization problem can be solved using optimization methods that exploit orthogonality constraints. We start by proposing cost functions for the joint diagonalization problem. We then explain some fundamentals of optimization methods over the Stiefel manifold and based on this, we develop gradient and Newton based algorithms for joint orthogonal diagonalization problems. The optimization methods used in this section are based on the works of (Edelman *et al.*, 1998) and (Manton, 2002). Since some of the optimization methods used in this section have been developed only for real matrices (for example the conjugate gradient method), throughout this section for consistency we assume real valued matrices. Nevertheless the extension to the complex case for the rest of algorithms (Algorithms I, III and IV) is straightforward and in most cases can be done by changing the "transpose" operator to "transpose and Hermitian". The details of derivations for the complex case can be found in Appendix C.

2.3.1 The Cost Function

Given the set of square symmetric matrices $\mathcal{R} = \{\mathbf{R}_1, \ldots, \mathbf{R}_M\}$, the orthogonal joint diagonalizer of this set, denoted as $\mathbf{Q}$, can be estimated by minimizing the sum of squared off -diagonal values of $\mathbf{Q}^T\mathbf{R}_m\mathbf{Q}$ for all $m = 1, \ldots, M$; i.e., we have the following optimization problem

$$\begin{array}{ll} \min\limits_{\mathbf{Q}} & \mathcal{C}_{off}(\mathcal{R}, \mathbf{Q}), \\ \text{subject to} & \mathbf{Q}^T\mathbf{Q} = \mathbf{I} \end{array} \tag{2.24}$$

where

$$\mathcal{C}_{off}(\mathcal{R}, \mathbf{Q}) = \sum_{m=1}^{M} ||\mathbf{Q}^T\mathbf{R}_m\mathbf{Q} - \text{ddiag}\{\mathbf{Q}^T\mathbf{R}_m\mathbf{Q}\}||_F^2, \tag{2.25}$$

and ddiag($\mathbf{X}$) is an operator which forms a diagonal matrix of diagonal values of the matrix $\mathbf{X}$. Since the Frobeneous norm of a matrix does not change by post- or pre-multiplication by a orthogonal matrix, rather than minimizing the sum of squared off-diagonal values we can maximize the sum of squared diagonal values; i.e.,

$$\begin{array}{ll} \max\limits_{\mathbf{Q}} & \mathcal{C}_{d}(\mathcal{R}, \mathbf{Q}), \\ \text{subject to} & \mathbf{Q}^T\mathbf{Q} = \mathbf{I} \end{array} \tag{2.26}$$

where $\mathcal{C}_d(\mathcal{R}, \mathbf{Q}) = \sum_{m=1}^{M} ||\,\text{diag}\{\mathbf{Q}^T\mathbf{R}_m\mathbf{Q}\}||_2^2$ and diag($\mathbf{X}$) is an operator which forms a column vector of diagonal values of the matrix $\mathbf{X}$.

Let $a_{ii}(m)$ represent the i_{th} diagonal value of $\mathbf{Q}^T\mathbf{R}_m\mathbf{Q}$. We then have

$$\begin{aligned} ||\operatorname{diag}\{\mathbf{Q}^T\mathbf{R}_m\mathbf{Q}\}||_2^2 &= \sum_{i=1}^{N} a_{ii}^2(m) \\ &= \frac{1}{N}\Big(\sum_{i=1}^{N} a_{ii}\Big)^2 + \frac{1}{N}\sum_{\substack{i<j \\ i,j=1}}^{N}(a_{ii}(m) - a_{jj}(m))^2. \end{aligned} \tag{2.27}$$

The first term of the RHS of equation (2.27) is equivalent to $\frac{1}{N}(Tr\{\mathbf{Q}^T\mathbf{R}_m\mathbf{Q}\})^2 = \frac{1}{N}(Tr\{\mathbf{R}_m\})^2$ and is invariant with respect to $\mathbf{Q}$. Hence for maximizing (2.27) we only need to maximize the second term of the RHS of (2.27). We therefore propose minimizing the following criterion subject to the orthogonality constraint $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$:

$$\tilde{C}_d(\mathcal{R}, \mathbf{Q}) = -\frac{1}{2N}\sum_{m=1}^{M}\sum_{\substack{i<j \\ i,j=1}}^{N}(a_{ii}(m) - a_{jj}(m))^2, \tag{2.28}$$

where the extra factor of $\frac{1}{2}$ is introduced for convenience in derivations appearing later on.

We now discuss the optimization methods that we can use to minimize the cost function given in (2.28).

2.3.2 Optimization Methods Over the Stiefel Manifold

The geometry of the orthogonality constraint $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ can be represented by the nonlinear space known as the Stiefel manifold. In (Edelman *et al.*, 1998) and (Manton, 2002) the authors provide a framework for solving optimization problems that involve such constraints. The key idea is that optimization problems with orthogonality constraints can be treated as unconstrained ones over the Stiefel manifold. In the linear Euclidean space, the update rule for smooth unconstrained optimization of an objective function $f(\mathbf{x})$ is given by

$$\mathbf{x}_k = \mathbf{x}_{k-1} + t\mathbf{h}_{k-1} \tag{2.29}$$

where $\mathbf{h}_{k-1}$ is the search direction at iteration $k-1$ and is calculated based on the knowledge of the gradient or Hessian of the objective function, and t is the step size parameter typically chosen using line search methods such as Armijo's rule (Bertsekas, 1999)(also see

Appendix B). Similar concepts can be carried over to optimization on a manifold by redefining operations such as line update (2.29) and the gradient or Hessian of a function, to the appropriate operators over the manifold. For example, analogous to the definition of a straight line in Euclidean space, the *geodesic* is defined as the curve with the shortest length between two points on a manifold. As shown in (Edelman *et al.*, 1998), on the Stiefel manifold the equation for the geodesic emanating from $\mathbf{Q}_{k-1}$ in the direction of $\mathbf{H}_{k-1}$, where $\mathbf{Q}_{k-1}$ and $\mathbf{H}_{k-1}$ are $J \times N$ matrices such that $\mathbf{Q}_{k-1}^T\mathbf{Q}_{k-1} = \mathbf{I}_N$ and $\mathbf{A} = \mathbf{Q}_{k-1}^T\mathbf{H}_{k-1}$ is skew-symmetric, is given by

$$\mathbf{Q}_k = \mathbf{Q}_{k-1}\mathbf{E}(t) + \mathbf{V}\mathbf{F}(t) \tag{2.30}$$

where $\mathbf{E}(t)$ and $\mathbf{F}(t)$ are $N \times N$ matrices calculated from

$$\begin{pmatrix} \mathbf{E}(t) \\ \mathbf{F}(t) \end{pmatrix} = \exp\left[t\begin{pmatrix} \mathbf{A} & -\mathbf{R}^T \\ \mathbf{R} & \mathbf{0} \end{pmatrix}\begin{pmatrix} \mathbf{I}_N \\ \mathbf{0} \end{pmatrix}\right] \tag{2.31}$$

and $\mathbf{V}$ and $\mathbf{R}$ are the QR decomposition of

$$\mathbf{V}\mathbf{R} = (\mathbf{I} - \mathbf{Q}_{k-1}\mathbf{Q}_{k-1}^T)\mathbf{H}_{k-1}. \tag{2.32}$$

When $\mathbf{Q}_{k-1}$ is a square orthogonal matrix, the update rule over a geodesic, given by equation (2.30),(2.31) and (2.32), simplifies to

$$\mathbf{Q}_k = \mathbf{Q}_{k-1}\exp\left[t\mathbf{Q}_{k-1}^T\mathbf{H}_{k-1}\right]. \tag{2.33}$$

In (Manton, 2002) an alternative approach is suggested by first using the standard linear update rule given in (2.29) and then projecting it onto the Stiefel manifold; i.e.,

$$\mathbf{Q}_k = \pi(\mathbf{Q}_{k-1} + t\mathbf{H}_{k-1}) \tag{2.34}$$

where π is the projection operator which is defined for a matrix $\mathbf{X} \in \mathbb{R}^{J\times N}$ as

$$\pi(\mathbf{X}) = \mathbf{U}_x\mathbf{I}_{J,N}\mathbf{V}_x^T \tag{2.35}$$

where $\mathbf{U}_x$ and $\mathbf{V}_x$ are obtained from singular value decomposition of $\mathbf{X}$; i.e., $\mathbf{X} = \mathbf{U}_x\mathbf{\Sigma}_x\mathbf{V}_x^T$.

For minimizing (2.28) we can use either equation (2.33) or (2.34) for updating the value of $\mathbf{Q}$. The remaining task is to calculate the search direction $\mathbf{H}_k$ on the Stiefel manifold.

Similar to the Euclidean space, to calculate the search direction on the Stiefel manifold, we can use the gradient or Hessian information of the function. For example, the steepest descent direction $\mathbf{H}_k = -\mathbf{G}_k$ for minimizing the function $f(\mathbf{X})$, $\mathbf{X}^T\mathbf{X} = \mathbf{I}$ on Stiefel manifold is calculated from: (Manton, 2002)(Edelman *et al.*, 1998)

$$\mathbf{G}_k = -\mathbf{X}_k\mathbf{D}_{X_k}^T\mathbf{X}_k + \mathbf{D}_{X_k} \tag{2.36}$$

where $\mathbf{D}_{X_k}$ is the gradient of $f(\mathbf{X})$ evaluated at $\mathbf{X}_k$.

We can also define the Newton direction over the Stiefel manifold. Compared to steepest descent, calculating the Newton direction is more complicated and computationally more expensive. Nevertheless the Newton method can have quadratic convergence, while for gradient descent, the rate of convergence is linear. In (Manton, 2002) a method for finding the Newton step over the Stiefel manifold has been proposed, that uses quadratic approximation of the function $f(\pi(\mathbf{X} + \mathbf{H}))$, where $\mathbf{H}$ is some vector in the tangent space of the manifold.

In optimization over a linear space, a method which is faster than steepest descent, but still only needs gradient information of the objective function, is the conjugate gradient method (Bertsekas, 1999). The search direction ($\mathbf{h}_k$) in the conjugate gradient method at each step is calculated using a linear combination of the gradient ($\mathbf{g}_k$) of the objective function at the current step and the search direction at the previous step; i.e,

$$\mathbf{h}_k = -\mathbf{g}_k + \beta_k\mathbf{h}_{k-1} \tag{2.37}$$

where β_k is given by:

$$\beta_k = \frac{\mathbf{g}_k^T\mathbf{g}_k}{\mathbf{g}_{k-1}^T\mathbf{g}_{k-1}}. \tag{2.38}$$

As is shown in (Bertsekas, 1999) for quadratic problems, conjugate gradient methods can converge in a finite number of steps. For non-quadratic problems convergence may not happen after a finite number of steps but nevertheless convergence can still be faster than a steepest descent method. In a manner similar to the Euclidean space, a conjugate gradient step on the Stiefel manifold is done first by parallel transporting the previous search direction to the point corresponding to the new step, and then choosing the new search direction to be a combination of the parallel transported version of the old search direction and the new gradient. In the Euclidean space, parallel transporting a vector is done by moving the base of the arrow. On an embedded manifold, if we use the same concept to move a tangent vector the result won't necessarily be a tangent vector. For parallel transport of tangent vectors on a manifold we can parallel transport the vector in infinitesimal steps in a similar way as in a Euclidean space, and then in each step we remove the normal component of the transferred vector such that the remaining portion is still tangent to the manifold. According to (Edelman *et al.*, 1998), for parallel translation along geodesics there is no simple, general formula. Nevertheless if the vector is tangent to a geodesic, then it is easy to find the parallel transport result by noticing that a geodesic always parallel transports its own tangent vector. Let $\bar{\mathbf{H}}_{k-1}$ be the parallel transform of the tangent vector $\mathbf{H}_{k-1}$ from point $\mathbf{Q}_{k-1}$ to point $\mathbf{Q}_k$. Then

$$\bar{\mathbf{H}}_{k-1} = \mathbf{H}_{k-1}\mathbf{E}(t) - \mathbf{Q}_k\mathbf{R}^T\mathbf{F}(t) \tag{2.39}$$

where $\mathbf{E}(t)$ and $\mathbf{F}(t)$ are calculated from (2.31) and $\mathbf{R}$ is given by (2.32). When $\mathbf{Q}_k$ is a square matrix then the above equation simplifies to

$$\bar{\mathbf{H}}_{k-1} = \mathbf{H}_{k-1}e^{[t\mathbf{Q}_{k-1}^T\mathbf{H}_{k-1}]}. \tag{2.40}$$

Having (2.40), the conjugate gradient step on the Stiefel manifold is given as

$$\mathbf{H}_k = -\mathbf{G}_k + \beta_k\bar{\mathbf{H}}_k \tag{2.41}$$

where β_k is given by

$$\beta_k = \frac{\langle \mathbf{G}_k, \mathbf{G}_k \rangle}{\langle \mathbf{G}_{k-1}, \mathbf{G}_{k-1} \rangle} \tag{2.42}$$

and $\langle \mathbf{\Delta}_1, \mathbf{\Delta}_2 \rangle$ represents the inner product between two tangent vectors on the Stiefel manifold and is defined as: (Edelman *et al.*, 1998)

$$\langle \mathbf{\Delta}_1, \mathbf{\Delta}_2 \rangle = Tr(\mathbf{\Delta}_1(\mathbf{I} - \frac{1}{2}\mathbf{Q}_k\mathbf{Q}_k^T)\mathbf{\Delta}_2). \tag{2.43}$$

Notice that when $\mathbf{Q}_k$ is a square matrix, the above equation simplifies to

$$\langle \mathbf{\Delta}_1, \mathbf{\Delta}_2 \rangle = \frac{1}{2}Tr(\mathbf{\Delta}_1^T\mathbf{\Delta}_2). \tag{2.44}$$

The conjugate gradient step with β_k given as (2.42) is known as the Fletcher-Reeves conjugate gradient. Another way of calculating β, known as Polak-Ribiere conjugate gradient, is

$$\beta_k = \frac{\langle \mathbf{G}_k - \mathbf{G}_{k-1}, \mathbf{G}_k \rangle}{\langle \mathbf{G}_{k-1}, \mathbf{G}_{k-1} \rangle}. \tag{2.45}$$

2.3.3 Algorithms

We can now derive orthogonal joint diagonalization algorithms based on the three optimization methods discussed above. To use these optimization methods we first need to calculate the gradient and Hessian of the cost function given in (2.28), rewritten below for ease of reference:

$$\tilde{\mathcal{C}}_d(\mathcal{R}, \mathbf{Q}) = -\frac{1}{2N} \sum_{m=1}^{M} \sum_{\substack{i<j \\ i,j=1}}^{N} (a_{ii}(m) - a_{jj}(m))^2. \tag{2.46}$$

Let $\mathbf{D}_Q$ represent the matrix of partial derivatives of $\tilde{\mathcal{C}}_d(\mathcal{R}, \mathbf{Q})$ with respect to elements of $\mathbf{Q}$. Then the rs_{th} element of $\mathbf{D}_Q$ is given as

$$[\mathbf{D}_Q]_{rs} = \frac{\partial \tilde{\mathcal{C}}_d(\mathcal{R}, \mathbf{Q})}{\partial q_{rs}} = -\frac{1}{N} \sum_{m=1}^{M} \frac{\partial a_{ss}(m)}{\partial q_{rs}} \sum_{\substack{j \neq s \\ j=1}}^{N} (a_{ss}(m) - a_{jj}(m)) \tag{2.47}$$

where here we use the fact that

$$\frac{\partial a_{ii}(m)}{\partial q_{rs}} = 0 \quad \text{for} \quad s \neq i, \tag{2.48}$$

where $a_{ii}(m) = \mathbf{q}_i^T \mathbf{R}_m \mathbf{q}_i$ and $\mathbf{q}_i$ is the i_{th} column of $\mathbf{Q}$. Note that

$$\begin{aligned} \sum_{\substack{j \neq s \\ j=1}}^{N} (a_{ss}(m) - a_{jj}(m)) &= N a_{ss}(m) - Tr(\mathbf{Q}^T \mathbf{R}_m \mathbf{Q}) \\ &= N(a_{ss}(m) - c(m)) \end{aligned} \tag{2.49}$$

where $c(m) = \frac{1}{N} Tr(\mathbf{Q}^T \mathbf{R}_m \mathbf{Q}) = \frac{1}{N} Tr(\mathbf{R}_m)$ and is independent of $\mathbf{Q}$. Also we can write[3]

$$\frac{\partial a_{ss}(m)}{\partial q_{rs}} = 2[\mathbf{R}_m \mathbf{Q}]_{rs}. \tag{2.50}$$

[3] Here we use the assumption that $\mathbf{R}_m = \mathbf{R}_m^T$

From (2.47), (2.49) and (2.50) we have:

$$\mathbf{D}_Q = -2\sum_{m=1}^{M} \mathbf{R}_m\mathbf{Q}\boldsymbol{\Sigma}(m) \tag{2.51}$$

where $\boldsymbol{\Sigma}(m) = \text{ddiag}\{\mathbf{Q}^T\mathbf{R}_m\mathbf{Q}\} - c(m)\mathbf{I}$. Having $\mathbf{D}_Q$, the gradient over the Stiefel manifold is calculated from (2.36) as

$$\begin{aligned}\mathbf{G} &= -2\sum_{m=1}^{M} \mathbf{R}_m\mathbf{Q}\boldsymbol{\Sigma}(m) + 2\mathbf{Q}\Big(\sum_{m=1}^{M} \mathbf{R}_m\mathbf{Q}\boldsymbol{\Sigma}(m)\Big)^T\mathbf{Q} \\ &= -2\sum_{m=1}^{M} \mathbf{Q}\mathbf{Q}^T\mathbf{R}_m\mathbf{Q}\boldsymbol{\Sigma}(m) + 2\sum_{m=1}^{M} \mathbf{Q}\boldsymbol{\Sigma}(m)\mathbf{Q}^T\mathbf{R}_m\mathbf{Q} \\ &= 2\mathbf{Q}\Big(\sum_{m=1}^{M}[\boldsymbol{\Sigma}(m)\mathbf{R}_y(m) - \mathbf{R}_y(m)\boldsymbol{\Sigma}(m)]\Big),\end{aligned} \tag{2.52}$$

where $\mathbf{R}_y(m) = \mathbf{Q}^T\mathbf{R}_m\mathbf{Q}$.

Using (2.52) we now can design steepest descent and conjugate gradient direction algorithms for optimizing the cost function in (2.46) as is summarized below. Note that for these two methods we use Armijo's rule for choosing the step size t.

Algorithm I: Joint Orthogonal Diagonalization Using The Steepest Descent Method Over The Stiefel Manifold

1. Initialize $\mathbf{Q}$ to some random matrix such that $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ and set $c(m) = \frac{1}{N}Tr(\mathbf{R}_m), \quad m = 1, \ldots, M$.

2. Choose the initial value for step size $t = 1$.

3. for $k = 1$ to Max_Numitr

   - Set

   $$\boldsymbol{\Sigma}(m) = \text{ddiag}\{\mathbf{Q}^T\mathbf{R}_m\mathbf{Q}\} - c(m)\mathbf{I}$$

   - Calculate

   $$\mathbf{G} = 2\mathbf{Q}\Big(\sum_{m=1}^{M}[\boldsymbol{\Sigma}(m)\mathbf{R}_y(m) - \mathbf{R}_y(m)\boldsymbol{\Sigma}(m)]\Big),$$

   where $\mathbf{R}_y(m) = \mathbf{Q}^T\mathbf{R}_m\mathbf{Q}$.

- Calculate the value of cost function at $\mathbf{Q}$ from[4]

$$\tilde{C}_d(\mathcal{R}, \mathbf{Q}) = \sum_{m=1}^{M} \left[\frac{1}{2N}[Tr(\mathbf{R}_m)]^2 - \frac{1}{2}\mathbf{d}_m^T(\mathbf{Q})\mathbf{d}_m(\mathbf{Q})\right]$$

  where $\mathbf{d}_m(\mathbf{Q}) = \text{diag}\{\mathbf{Q}^T\mathbf{R}_m\mathbf{Q}\}$.

- If $k > 1$ then
  - If $\dfrac{|\tilde{C}_d^k(\mathcal{R}, \mathbf{Q}) - \tilde{C}_d^{k-1}(\mathcal{R}, \mathbf{Q})|}{\tilde{C}_d^k(\mathcal{R}, \mathbf{Q})} < \epsilon$, where $0 < \epsilon \ll 1$, then STOP

- Use Armijo's rule to find the step size t

  (a) Propose a new point $\mathbf{Q}_p$ by setting $\mathbf{Q}_p = \mathbf{Q}\exp\left[t\mathbf{Q}^T\mathbf{H}\right]$ where $\mathbf{H} = -\mathbf{G}$.

  (b) if $\tilde{C}_d(\mathcal{R}, \mathbf{Q}_p) - \tilde{C}_d(\mathcal{R}, \mathbf{Q}) \leq t\alpha\frac{1}{2}Tr(\mathbf{G}^T\mathbf{H})$,where $0 < \alpha < 1$, then

  - $t = 2t$
  - $\mathbf{Q}_p = \mathbf{Q}\exp\left[t\mathbf{Q}^T\mathbf{H}\right]$ and repeat from step b.

  (c) if $\tilde{C}_d(\mathcal{R}, \mathbf{Q}_p) - \tilde{C}_d(\mathcal{R}, \mathbf{Q}) > t\alpha\frac{1}{4}Tr(\mathbf{G}^T\mathbf{H})$ then

  - $t = \frac{1}{2}t$
  - $\mathbf{Q}_p = \mathbf{Q}\exp\left[t\mathbf{Q}^T\mathbf{H}\right]$ and repeat from step c.

- $\mathbf{Q} = \mathbf{Q}_p$ and continue the loop from step 3.

Algorithm II: Joint Orthogonal Diagonalization Using The Conjugate Gradient Method Over The Stiefel Manifold

1. Initialize $\mathbf{Q}_0$ to some random matrix such that $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ and set $c(m) = \frac{1}{N}Tr(\mathbf{R}_m) \quad m = 1, \ldots, M$.

2. Choose the initial value for step size $t = 1$.

[4] To calculate $\tilde{C}_d(\mathcal{R}, \mathbf{Q})$ we use (2.27) by noticing that:

$$-\frac{1}{N}\sum_{\substack{i<j\\ i,j=1}}^{N} (a_{ii}(m) - a_{jj}(m))^2 = \frac{1}{N}\Big(\sum_{i=1}^{N} a_{ii}\Big)^2 - \sum_{i=1}^{N} a_{ii}^2(m).$$

Note also that $\sum_{i=1}^{N} a_{ii}^2(m) = \mathbf{d}_m^T(\mathbf{Q})\mathbf{d}_m(\mathbf{Q})$ and $\left(\sum_{i=1}^{N} a_{ii}\right)^2 = [Tr(\mathbf{R}_m)]^2$.

```
3. for $k = 1$ to Max_Numitr
    • Set
        $\mathbf{\Sigma}(m) = \text{ddiag}\{\mathbf{Q}^T \mathbf{R}_m \mathbf{Q}\} - c(m)\mathbf{I}$
    • Calculate the gradient over the Stiefel manifold
        $\mathbf{G} = 2\mathbf{Q}\Big( \sum_{m=1}^{M} [\mathbf{\Sigma}(m)\mathbf{R}_y(m) - \mathbf{R}_y(m)\mathbf{\Sigma}(m)] \Big),$
      where $\mathbf{R}_y(m) = \mathbf{Q}^T \mathbf{R}_m \mathbf{Q}$.
    • calculate the value of cost function at $\mathbf{Q}$ from
        $\tilde{C}_d(\mathcal{R}, \mathbf{Q}) = \sum_{m=1}^{M} \Big[ \frac{1}{2N} [Tr(\mathbf{R}_m)]^2 - \frac{1}{2} \mathbf{d}_m^T(\mathbf{Q}) \mathbf{d}_m(\mathbf{Q}) \Big]$
      where $\mathbf{d}_m(\mathbf{Q}) = \text{diag}\{\mathbf{Q}^T \mathbf{R}_m \mathbf{Q}\}$
    • If $k > 1$ then
        – if $\frac{|\tilde{C}_d^k(\mathcal{R}, \mathbf{Q}) - \tilde{C}_d^{k-1}(\mathcal{R}, \mathbf{Q})|}{\tilde{C}_d^k(\mathcal{R}, \mathbf{Q})} < \epsilon$, where $0 < \epsilon \ll 1$, then STOP
    • Initialize with a gradient step after $N(N-1)/2$ iterations
        – if $k = 1$ or $k$ Mod $N(N-1)/2 = 0$ then
            $\mathbf{H} = -\mathbf{G}$
        – else
            $\mathbf{H} = -\mathbf{G} + \beta \mathbf{H} e^{[t\mathbf{Q}^T \mathbf{H}]}$
          where $\beta = \frac{Tr(\mathbf{G}^T \mathbf{G})}{Tr(\mathbf{G}_{k-1}^T \mathbf{G}_{k-1})}$.
    • Use Armijo's rule to find the step size $t$
        (a) Propose a new point $\mathbf{Q}_p$ by setting $\mathbf{Q}_p = \mathbf{Q} \exp\big[t\mathbf{Q}^T \mathbf{H}\big]$.
        (b) if $\tilde{C}_d(\mathcal{R}, \mathbf{Q}_p) - \tilde{C}_d(\mathcal{R}, \mathbf{Q}) \le t\alpha \frac{1}{2} Tr(\mathbf{G}^T \mathbf{H})$, where $0 < \alpha < 1$ then
            – $t = 2t$.
            – $\mathbf{Q}_p = \mathbf{Q} \exp\big[t\mathbf{Q}^T \mathbf{H}\big]$ and repeat from step b.
```

(c) if $\tilde{C}_d(\mathcal{R}, \mathbf{Q}_p) - \tilde{C}_d(\mathcal{R}, \mathbf{Q}) > t\frac{1}{4}Tr(\mathbf{G}^T\mathbf{H})$ then

- $t = \frac{1}{2}t$
- $\mathbf{Q}_p = \mathbf{Q}\exp\left[t\mathbf{Q}^T\mathbf{H}\right]$ and repeat from step c

- Set $\mathbf{Q} = \mathbf{Q}_p$ and continue the loop from step 3

For the Newton method over the Stiefel manifold we first need to calculate the Hessian matrix of the objective function given in (2.46). To do this we can write the objective function of the optimization problem in (2.46) as:

$$\tilde{C}_d(\mathcal{R}, \mathbf{Q}) = -\frac{1}{2}\sum_{m=1}^{M}\sum_{\substack{i<j\\ i,j=1}}^{N}\left(\mathbf{q}^T\mathbf{A}_{ij}(m)\mathbf{q}\right)^2 \tag{2.53}$$

where $\mathbf{q} = \text{vec}\{\mathbf{Q}\}$ and

$$\mathbf{A}_{ij}(m) = \mathbf{E}_{ij} \otimes \mathbf{R}_m \tag{2.54}$$

where $\mathbf{E}_{ij}$ is an $N \times N$ diagonal matrix such that its i_{th} diagonal element is 1 and its j_{th} diagonal element is -1 and all other elements are equal to zero. Note that $\mathbf{A}_{ij}(m)$ is a symmetric matrix[5]; i.e., $\mathbf{A}_{ij}(m) = \mathbf{A}_{ij}(m)^T$. The Hessian of $\tilde{C}_d(\mathcal{R}, \mathbf{Q})$ is found to be (see Appendix C for derivation):

$$\mathbf{D}_{QQ} = -\sum_{m=1}^{M}\sum_{\substack{i<j\\ i,j=1}}^{N}\left[4\mathbf{A}_{ij}(m)\mathbf{q}\mathbf{q}^T\mathbf{A}_{ij}(m) + 2\mathbf{q}^T\mathbf{A}_{ij}(m)\mathbf{q}\mathbf{A}_{ij}(m)\right]. \tag{2.55}$$

In (Manton, 2002), new optimization methods over the Stiefel manifold have been proposed which are based on optimizing a local parameterization of the cost function $f(\mathbf{X})$ at each iteration. Based on this approach the Newton step over the Stiefel manifold is calculated by finding the turning points of the second-order Taylor series approximation of the local cost function $g(\mathbf{H}) = f(\pi(\mathbf{X}+\mathbf{H}))$, where $\pi(\mathbf{X})$ is defined in (2.35) and $\mathbf{H}$ is in the tangent space

[5] $\mathbf{A}_{ij}^T(m) = \mathbf{E}_{ij}^T \otimes \mathbf{R}_m^T = \mathbf{E}_{ij} \otimes \mathbf{R}_m = \mathbf{A}_{ij}(m)$ where here we have used the property of Kronecker products where for any two matrices $\mathbf{A}$ and $\mathbf{B}$ we have $(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T$.

of the Stiefel manifold at point $\mathbf{X}$[6]. As demonstrated in (Manton, 2002), the second-order Taylor series approximation of $g(\mathbf{H})$ is given as

$$g(\mathbf{H}) = f(\mathbf{X}) + Tr(\mathbf{H}^T \mathbf{D}_X) + \frac{1}{2}\operatorname{vec}\{\mathbf{H}\}^T \left[\mathbf{D}_{XX} - \frac{1}{2}((\mathbf{X}^T\mathbf{D}_X + \mathbf{D}_X^T\mathbf{X})^T \otimes \mathbf{I})\right] \operatorname{vec}\{\mathbf{H}\} + O(||H||^3) \tag{2.56}$$

where $\mathbf{D}_X$ and $\mathbf{D}_{XX}$ are respectively the gradient and Hessian of the $f(\mathbf{X})$. The Newton step is calculated by finding the turning point of the function $g(\mathbf{H})$, defined in (2.56), subject to the constraint that $\mathbf{H}$ is a tangent vector at point $\mathbf{X}$. A detailed description for finding the turning point of $g(\mathbf{H})$ is beyond the scope of this chapter, but further details can be found in (Manton, 2002). Also in (Manton, 2002) a Matlab function, *tpoint*, has been provided for calculating the turning point of (2.56). For ease of reference this Matlab function has been included in appendix A.

Note that the computational complexity of the Newton algorithm is much higher compared to the steepest descent or conjugate gradient methods; also, there is no guarantee that the Newton method will converge to a local minimum, since it can very well converge to a saddle point. To prevent this, as shown in our simulation results, we can use a few iterations of the steepest decent algorithm to minimize the cost function such that the value of the cost function is close enough to a local minimum before we apply the Newton algorithm.

Algorithm III: Joint Orthogonal Diagonalization Using The Newton Method Over The Stiefel Manifold

```
1. Initialize $\mathbf{Q}_0$ to some random matrix such that $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ and set $c(m) = \frac{1}{N}Tr(\mathbf{R}_m)$   $m = 1, \ldots, M$.
2. for $k = 1$ to Max_Numitr
    • Set
          $\boldsymbol{\Sigma}(m) = \operatorname{ddiag}\{\mathbf{Q}^T\mathbf{R}_m\mathbf{Q}\} - c(m)\mathbf{I}$
```

[6] The tangent space at point $\mathbf{X}$ on the Stiefel manifold is defined as (Manton, 2002) the set of matrices $T_X(N,N) = \left\{\mathbf{H} \in \mathbb{R}^{N\times N} | \mathbf{H} = \mathbf{X}\mathbf{A},\ \ \mathbf{A} \in \mathbb{R}^{N\times N},\ \ \mathbf{A} + \mathbf{A}^T = 0\right\}$.

- Calculate the first and second-order derivatives of $\tilde{\mathcal{C}}(\mathcal{R}, \mathbf{Q})$:

$$\mathbf{D}_Q = -2 \sum_{m=1}^{M} \mathbf{R}_m \mathbf{Q} \boldsymbol{\Sigma}(m)$$

$$\mathbf{D}_{QQ} = - \sum_{m=1}^{M} \sum_{\substack{i<j \\ i,j=1}}^{N} \left\{ 4\mathbf{A}_{ij}(m)\mathbf{q}\mathbf{q}^T \mathbf{A}_{ij}(m) + 2\mathbf{q}^T \mathbf{A}_{ij}(m)\mathbf{q}\mathbf{A}_{ij}(m) \right\}$$

  where $\mathbf{A}_{ij}(m) = \mathbf{E}_{ij} \otimes \mathbf{R}_m$ and $\mathbf{E}_{ij}$ is an $N \times N$ diagonal matrix such that its i_{th} diagonal element is 1 and its j_{th} diagonal element is -1 and all other elements are equal to zero.

- Calculate the value of cost function at $\mathbf{Q}$ from

$$\tilde{\mathcal{C}}_d(\mathcal{R}, \mathbf{Q}) = \sum_{m=1}^{M} \left[\frac{1}{2N} [Tr(\mathbf{R}_m)]^2 - \frac{1}{2} \mathbf{d}_m^T(\mathbf{Q}) \mathbf{d}_m(\mathbf{Q}) \right]$$

  where $\mathbf{d}_m(\mathbf{Q}) = \text{diag}\{\mathbf{Q}^T \mathbf{R}_m \mathbf{Q}\}$.

- If $k > 1$ then
  - if $\dfrac{|\tilde{\mathcal{C}}_d^k(\mathcal{R}, \mathbf{Q}) - \tilde{\mathcal{C}}_d^{k-1}(\mathcal{R}, \mathbf{Q})|}{\tilde{\mathcal{C}}_d^k(\mathcal{R}, \mathbf{Q})} < \epsilon$, where $0 < \epsilon \ll 1$, then STOP.

- Calculate the Newton step from

$$\mathbf{H} = \text{tpoint}(\mathbf{Q}, \mathbf{D}_Q, \mathbf{D}_{QQ} - \frac{1}{2} \left[(\mathbf{Q}^T \mathbf{D}_Q + \mathbf{D}_Q^T \mathbf{Q})^T \otimes \mathbf{I} \right])$$

  where *tpoint* is defined in Appendix A.

- Propose a new point $\mathbf{Q}_p$ by setting $\mathbf{Q}_p = \pi(\mathbf{Q} + \mathbf{H})$.

- if $\tilde{\mathcal{C}}_d(\mathcal{R}, \mathbf{Q}) \leq \tilde{\mathcal{C}}_d(\mathcal{R}, \mathbf{Q}_p)$ then abort.

- $\mathbf{Q}_p = \mathbf{Q}$ and continue the loop from step 2.

2.4 A Maximum Likelihood Approach to the Joint Diagonalization Problem

In the previous section we derived joint diagonalization algorithms based on least-squares criteria using optimization methods over the Stiefel manifold. In this section we derive a maximum likelihood algorithm to estimate the parameters of $\mathbf{Q}$ from the sample estimates of a set of positive definite covariance matrices $\hat{\mathcal{R}} = \{\hat{\mathbf{R}}_1, \ldots, \hat{\mathbf{R}}_M\}$, with true estimates given as the set $\mathcal{R} = \{\mathbf{R}_m | \mathbf{R}_m = \mathbf{Q}^T \mathbf{\Lambda}(m) \mathbf{Q}, \quad m = 1, \ldots, M\}$, where $\mathbf{\Lambda}(m)$ are diagonal matrices with positive diagonal values. Assuming that $\hat{\mathbf{R}}_1, \ldots, \hat{\mathbf{R}}_M$ are respectively sample covariance matrices of M independent populations of zero-mean Gaussian distributed multivariate observations of size L, then they are distributed according to the Wishart distribution which is given as (Anderson, 1971):

$$\hat{\mathbf{R}}_m \sim \prod_{m=1}^{M} \frac{\alpha_m}{\beta_m} e^{-\frac{1}{2} Tr(\mathbf{R}_m^{-1} \hat{\mathbf{R}}_m)} \tag{2.57}$$

where α_m is a nonnegative constant and $\beta_i = \det(\mathbf{R}_i)^{L/2}$, where L is the number of snapshots used to estimate the covariance matrices $\hat{\mathbf{R}}_m$. The log-likelihood function of (2.57) can be written as

$$\mathcal{C}_l(\hat{\mathcal{R}}, \mathbf{Q}) = C - \frac{1}{2} \sum_{m=1}^{M} \left[L \log(\det(\mathbf{R}_m)) + Tr(\mathbf{R}_m^{-1} \hat{\mathbf{R}}_m) \right] \tag{2.58}$$

where C is some constant. Substituting $\mathbf{R}_m = \mathbf{Q}^T \mathbf{\Lambda}(m) \mathbf{Q}$ and assuming an orthogonality constraint on $\mathbf{Q}$, the maximum likelihood method for estimating $\mathbf{Q}$ and $\mathbf{\Lambda}(m)$ can be expressed as maximization of the following log-likelihood function:

$$\mathcal{C}_l(\hat{\mathcal{R}}, \mathbf{Q}) = C - \frac{1}{2} \sum_{m=1}^{M} \left[L \log(\det(\mathbf{\Lambda}(m))) + Tr(\mathbf{Q} \mathbf{\Lambda}^{-1}(m) \mathbf{Q}^T \hat{\mathbf{R}}_m) \right] \tag{2.59}$$

where here we used the fact that for any square orthogonal matrix $\mathbf{Q}$, $\det(\mathbf{Q}) = 1$. For a given $\mathbf{Q}$, it can be verified that the criterion in (2.59) is maximized when $\mathbf{\Lambda}(m) = \text{ddiag}\{\mathbf{Q}^T \hat{\mathbf{R}}_m \mathbf{Q}\}$. Substituting $\mathbf{\Lambda}(m)$ with $\text{ddiag}\{\mathbf{Q}^T \hat{\mathbf{R}}_m \mathbf{Q}\}$, (2.59) can be written as

$$\mathcal{C}_l(\hat{\mathcal{R}}, \mathbf{Q}) = C - \frac{1}{2} \sum_{m=1}^{M} \left[L \log \Big(\det \big(\text{ddiag}\{\mathbf{Q}^T \hat{\mathbf{R}}_m \mathbf{Q}\} \big) \Big) + N \right]. \tag{2.60}$$

Based on (2.60) we can write the following optimization problem for estimating the $\mathbf{Q}$:

$$\begin{aligned} \hat{\mathbf{Q}} &= \arg\min_{\mathbf{Q}} \quad \tilde{\mathcal{C}}_l(\hat{\mathcal{R}}, \mathbf{Q}), \\ \text{subject to} & \quad \mathbf{Q}^T\mathbf{Q} = \mathbf{I}. \end{aligned} \tag{2.61}$$

where $\tilde{\mathcal{C}}_l(\hat{\mathcal{R}}, \mathbf{Q}) = \frac{1}{2}\sum_{m=1}^{M} \log\Big(\det\big(\operatorname{ddiag}\{\mathbf{Q}^T\hat{\mathbf{R}}_m\mathbf{Q}\big)\Big)$.

We use similar algorithms discussed in the previous section to optimize the above criterion. To do so we first calculate the gradient and Hessian of the above cost function. We have

$$\log\Big(\det(\operatorname{ddiag}\{\mathbf{Q}^T\hat{\mathbf{R}}_m\mathbf{Q}\})\Big) = \sum_{i=1}^{N} \log(\mathbf{q}_i^T\hat{\mathbf{R}}_m\mathbf{q}_i) \tag{2.62}$$

where $\mathbf{q}_i$ is the i_{th} column of $\mathbf{Q}$. Substituting (2.62) in (2.60), and taking the derivative with respect to $\mathbf{q}_i$ we get

$$\frac{\partial\tilde{\mathcal{C}}_l(\hat{\mathcal{R}}, \mathbf{Q})}{\partial\mathbf{q}_i} = \sum_{m=1}^{M} \frac{1}{\mathbf{q}_i^T\hat{\mathbf{R}}_m\mathbf{q}_i}\hat{\mathbf{R}}_m\mathbf{q}_i, \quad i = 1, \ldots, N. \tag{2.63}$$

and from this it follows that the gradient of the cost function in (2.61) can be calculated from:

$$\mathbf{D}_Q = \sum_{m=1}^{M} \hat{\mathbf{R}}_m\mathbf{Q}\mathbf{\Lambda}_y^{-1}(m) \tag{2.64}$$

where $\mathbf{\Lambda}_y(m) = \operatorname{ddiag}\{\mathbf{Q}^T\hat{\mathbf{R}}_m\mathbf{Q}\}$.

To calculate the Hessian we can use the same approach as in the previous section. First we write the objective function of the optimization problem in (2.61) as

$$\tilde{\mathcal{C}}_l(\hat{\mathcal{R}}, \mathbf{Q}) = \frac{1}{2}\sum_{m=1}^{M}\sum_{i=1}^{N} \log(\boldsymbol{\nu}^T\mathbf{\Phi}_i(m)\boldsymbol{\nu}) \tag{2.65}$$

where $\boldsymbol{\nu} = \operatorname{vec}\{\mathbf{Q}\}$ and $\mathbf{\Phi}_i(m) = \mathbf{E}_i \otimes \hat{\mathbf{R}}_m$ where $\mathbf{E}_i$ is a $N \times N$ matrix with the i_{th} diagonal element equal to one and all other elements equal to zero. The Hessian is then calculated to be (see Appendix C):

$$\mathbf{D}_{QQ} = \sum_{m=1}^{M}\sum_{i=1}^{N} \left[\frac{\mathbf{\Phi}_i(m)}{\boldsymbol{\nu}^T\mathbf{\Phi}_i(m)\boldsymbol{\nu}} - \frac{2\mathbf{\Phi}_i(m)\boldsymbol{\nu}\boldsymbol{\nu}^T\mathbf{\Phi}_i(m)}{(\boldsymbol{\nu}^T\mathbf{\Phi}_i(m)\boldsymbol{\nu})^2}\right] \tag{2.66}$$

Having the gradient and Hessian of the likelihood function, we can easily derive steepest decent or Newton algorithms over the Stiefel manifold, similar to the ones in the previous section. For example the steepest decent direction is easily calculated to be

$$\begin{aligned}\mathbf{H} &= \mathbf{Q}\mathbf{D}_Q^T\mathbf{Q} - \mathbf{D}_Q \\ &= \sum_{m=1}^{M} \mathbf{Q}(\mathbf{R}_m\mathbf{Q}\mathbf{\Lambda}_y^{-1}(m))^T\mathbf{Q} - \mathbf{R}_m\mathbf{Q}\mathbf{\Lambda}_y^{-1}(m) \\ &= \sum_{m=1}^{M} [\mathbf{Q}\mathbf{\Lambda}_y^{-1}(m)\mathbf{Q}^T\mathbf{R}_m\mathbf{Q} - \mathbf{Q}\mathbf{Q}^T\mathbf{R}_m\mathbf{Q}\mathbf{\Lambda}_y^{-1}(m)] \\ &= \sum_{m=1}^{M} \mathbf{Q}[\mathbf{\Lambda}_y^{-1}(m)\mathbf{R}_y(m) - \mathbf{R}_y(m)\mathbf{\Lambda}_y^{-1}(m)]\end{aligned} \tag{2.67}$$

where $\mathbf{R}_y(m) = \mathbf{Q}^T\mathbf{R}_m\mathbf{Q}$.

For the Newton step we can use the same procedure as in Algorithm III of the previous section. All that is required is to substitute $\mathbf{D}_Q$ and $\mathbf{D}_{QQ}$ with values given by (2.64) and (2.66) respectively.

Algorithm IV: Joint Orthogonal Diagonalization using the ML Criterion and Newton Method Over the Stiefel Manifold

1. Initialize $\mathbf{Q}_0$ to some random matrix such that $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$.

2. for $k = 1$ to Max_Numitr

   - Set

     $$\mathbf{\Lambda}(m) = \text{ddiag}\{\mathbf{Q}^T\hat{\mathbf{R}}_m\mathbf{Q}\}$$

   - Calculate the first and second-order derivatives of $\tilde{\mathcal{C}}_l(\hat{\mathcal{R}}, \mathbf{Q})$:

     $$\mathbf{D}_Q = \sum_{m=1}^{M} \hat{\mathbf{R}}_m\mathbf{Q}\mathbf{\Lambda}^{-1}(m)$$

     $$\mathbf{D}_{QQ} = \sum_{m=1}^{M}\sum_{i=1}^{N}\left[\frac{\mathbf{\Phi}_i(m)}{\boldsymbol{\nu}^T\mathbf{\Phi}_i(m)\boldsymbol{\nu}} - \frac{2\mathbf{\Phi}_i(m)\boldsymbol{\nu}\boldsymbol{\nu}^T\mathbf{\Phi}_i(m)}{(\boldsymbol{\nu}^T\mathbf{\Phi}_i(m)\boldsymbol{\nu})^2}\right]$$

where $\nu = \text{vec}\{\mathbf{Q}\}$, $\mathbf{\Phi}_i(m) = \mathbf{E}_i \otimes \hat{\mathbf{R}}_m$ and $\mathbf{E}_i$ is an $N \times N$ diagonal matrix such that only its i_{th} diagonal element is equal to 1 and all other elements are equal to zero.

- Calculate the value of cost function at $\mathbf{Q}$ from

$$\tilde{\mathcal{C}}_l(\hat{\mathcal{R}}, \mathbf{Q}) = \frac{1}{2}\sum_{m=1}^{M} \log\Big(\det\big(\text{ddiag}\{\mathbf{Q}^T\hat{\mathbf{R}}_m\mathbf{Q}\}\big)\Big)$$

- If $k > 1$ then
  - if $\dfrac{|\tilde{\mathcal{C}}_l^k(\hat{\mathcal{R}}, \mathbf{Q}) - \tilde{\mathcal{C}}_l^{k-1}(\hat{\mathcal{R}}, \mathbf{Q})|}{\tilde{\mathcal{C}}_l^k(\hat{\mathcal{R}}, \mathbf{Q})} < \epsilon$, where $0 < \epsilon \ll 1$ then STOP
- Calculate the Newton step from

$$\mathbf{H} = \text{tpoint}(\mathbf{Q}, \mathbf{D}_Q, \mathbf{D}_{QQ} - \frac{1}{2}\Big[(\mathbf{Q}^T\mathbf{D}_Q + \mathbf{D}_Q^T\mathbf{Q})^T \otimes \mathbf{I}\Big])$$

where *tpoint* is defined in Appendix A.

- Propose a new point $\mathbf{Q}_p$ by setting $\mathbf{Q}_p = \pi(\mathbf{Q} + \mathbf{H})$.
- if $\tilde{\mathcal{C}}_l(\hat{\mathcal{R}}, \mathbf{Q}) \leq \tilde{\mathcal{C}}_l(\hat{\mathcal{R}}, \mathbf{Q}_p)$ then abort.
- otherwise $\mathbf{Q}_p = \mathbf{Q}$.
- end

Although one can argue the proposed algorithms are computationally more expensive than the extended Jacobi method, nevertheless for applications where the objective is to track the joint diagonalizer rather to estimate it from scratch, the gradient-based algorithms can be more efficient than fixed point methods such as extended Jacobi algorithm used in JADE. Notice in general as shown in (Hori, 2000) a gradient step of the proposed algorithm over the Stiefel manifold is computationally less expensive than a sweep of extended Jacobi method.

2.5 Simulation Results

In this section we use numerical simulations to show the performance and convergence properties of the joint diagonalization algorithms discussed above, comparing the performance of the new algorithms to the JADE method (Cardoso and Souloumiac, 1993).

2.5.1 Example 1

In this example we compare the convergence of the Newton based algorithm with the one of steepest descent and conjugate gradient based methods. To this end we apply the Algorithms I,II and III to the joint diagonalization of a set of matrices $\mathcal{R} = \{\mathbf{R}_i|, \mathbf{R}_i = \mathbf{Q}\mathbf{\Lambda}(i)\mathbf{Q}^T, \quad i = 1, \ldots, M\}$ where in this case $\mathbf{Q} \in \mathbb{R}^{6\times 6}$ is a randomly selected orthogonal matrix and $\mathbf{\Lambda}(i)$ are randomly chosen diagonal matrices with diagonal values in the range $[0, 1]$. The objective is to calculate the common eigenvectors of $\mathbf{R}_1, \ldots \mathbf{R}_M$, which in fact are the columns of $\mathbf{Q}$, using the joint diagonalization algorithms I,II and III. We set the initial value for $\mathbf{Q}$ for the steepest descent method to be an identity matrix and we initialize the conjugate gradient and Newton algorithms using ten steps of the steepest descent algorithm such that the initial value of these algorithms is close enough to ensure convergence with high probability.

We use the following measure of performance

$$\text{Error} = ||\hat{\mathbf{Q}} - \mathbf{Q}||_F^2 \tag{2.68}$$

where $\hat{\mathbf{Q}}$ is the estimated value of $\mathbf{Q}$ using the respective joint diagonalization algorithm. Since we can only estimate $\mathbf{Q}$ up to a permutation ambiguity, before calculating the error using (2.68) we manually correct the permutations of the columns of $\hat{\mathbf{Q}}$. As can be seen from Figure 2.1, the steepest descent algorithm has the slowest convergence, the conjugate gradient algorithm has better convergence compared to the steepest descent method, but certainly the Newton algorithm is the one which demonstrates the fastest convergence.

2.5.2 Example 2

In this example we show the application of the joint diagonalization algorithm to an estimation problem. Assume that we have data vectors $\mathbf{x}_m, \quad m = 1, \ldots, M$ generated using the model

$$\mathbf{x}_m = \mathbf{Q}\mathbf{s}_m \quad m = 1, \ldots, M \tag{2.69}$$

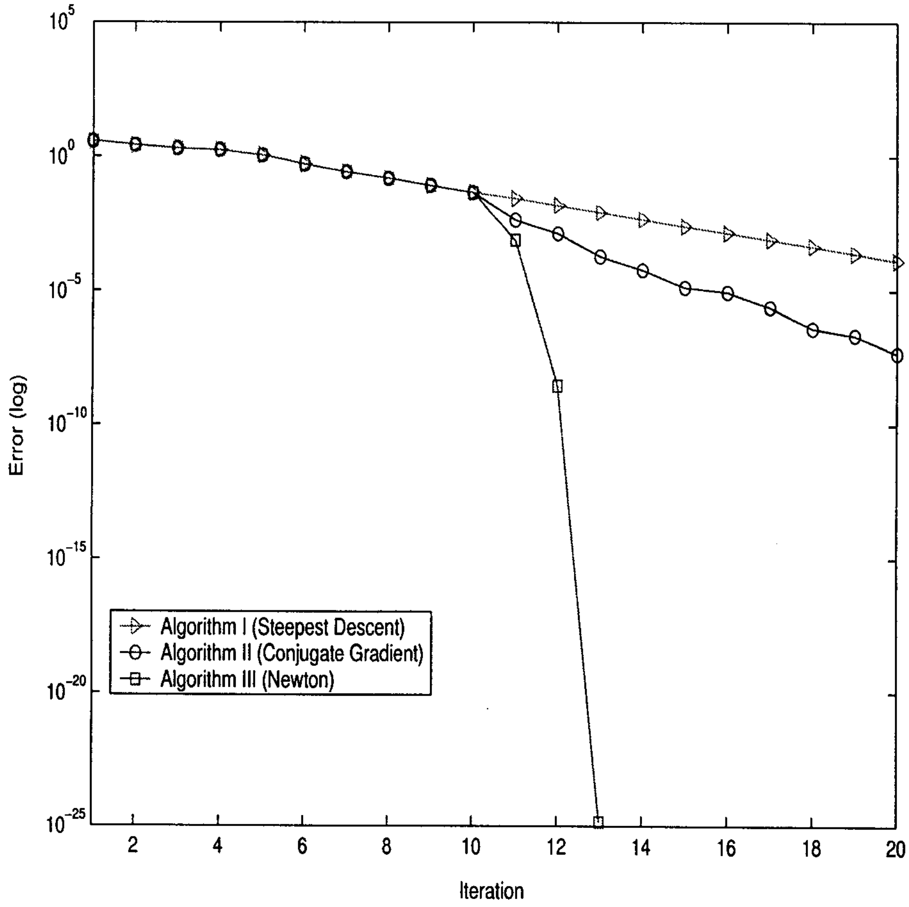

Figure 2.1: Estimation error for **Q** versus number of iterations for Algorithms I,II and III.

where in this case $\mathbf{s}_m \in \mathbb{R}^{6\times 6}$ are *iid*, zero-mean, Gaussian distributed with covariance matrix $\mathbf{\Sigma}_m$, where $\mathbf{\Sigma}_m$ are diagonal matrices, with diagonal values in this case uniformly distributed between zero and one. Given that $\mathbf{\Sigma}_m$ are unknown, the objective of this example is to estimate $\mathbf{Q}$ using only the data samples $\mathbf{x}_m$. The covariance matrix of $\mathbf{x}_m$ is given as

$$\mathbf{R}_m = \mathrm{E}[\mathbf{x}_m\mathbf{x}_m^T] = \mathbf{Q}\mathbf{\Sigma}_m\mathbf{Q}^T \quad m = 1, \ldots, M. \tag{2.70}$$

If $\mathbf{R}_m$ were given, as shown in Example 1, a perfect estimate of $\mathbf{Q}$ is achievable. Since we only have access to snapshots of $\mathbf{x}_i$, we can only calculate $\hat{\mathbf{R}}_m$, the sample estimate of $\mathbf{R}_m$, using

$$\hat{\mathbf{R}}_m = \frac{1}{N_x}\sum_{n=1}^{N_x} \mathbf{x}_m(n)\mathbf{x}_m^T(n) \quad m = 1, \ldots, M, \tag{2.71}$$

where N_x is the total number of snapshots. Note in this case we can only approximately diagonalize the set of matrices $\hat{\mathbf{R}}_1, \ldots, \hat{\mathbf{R}}_M$. In this example we use $N_x = 50$ snapshots and we use Algorithms III,IV and also the extended Jacobi method associated with the JADE algorithm (Cardoso and Souloumiac, 1993) to estimate the value of $\mathbf{Q}$. For Algorithms III and the extended Jacobi (JADE) method, we choose the initial $\mathbf{Q}$ to be the identity matrix. For Algorithm IV we use the estimated $\mathbf{Q}$, obtained using Algorithm III, as its initial value[7]. We choose the performance measure to be the mean squared error (mse) between the estimated and the true value of $\mathbf{Q}$ calculated from

$$\mathrm{mse} = \frac{1}{M_c}\sum_{i=1}^{M_c} ||\hat{\mathbf{Q}}_i - \mathbf{Q}||_F^2 \tag{2.72}$$

where $\mathbf{Q}_i$ is the estimated $\mathbf{Q}$ at i_{th} Monte Carlo run and M_c is the total number of the Monte Carlo runs. Table 2.1 shows the resulting mse for algorithms III and IV and JADE method versus M, the number of the covariance matrices used in the joint diagonalization process. As can be seen from the table, Algorithm III and the extended Jacobi of JADE method have exactly the same errors. This is not surprising because both algorithms are minimizing

[7] In our simulations we noticed that if we initialize Algorithm IV to an identity matrix it may not converge to its optimum value. This can be explained by noticing that the maximum likelihood function used in Algorithm IV is highly nonlinear and unless Algorithm IV is properly initialized it can very well converge to a local minimum. For Algorithms I,II and III this does not seem to be the case and they always converge to their optimal point, when initialized to the identity matrix.

M	20	40	60	100
Extended Jacobi Method (JADE)	0.0221	0.0103	0.0067	0.0038
Algorithm III	0.0221	0.0103	0.0067	0.0038
Algorithm IV	0.0031	0.0005	0.0003	0.0001

Table 2.1: Example 2, mse versus M for algorithms III, IV and the extended Jacobi of JADE method using $M_c = 50$ Monte Carlo runs.

the same criterion, which is the sum of squared off-diagonal values of $\mathbf{R}_1, \ldots, \mathbf{R}_M$. The difference between the two algorithms is that extended Jacobi algorithm is a fixed-point method based on recursive minimization of the least-squares criterion discussed above with respect to the Jacobi angles, while Algorithm III uses unconstrained minimization over a manifold to optimize the same criterion. Table 2.1 shows that these two algorithms converge to exactly the same point. On the other hand the mse for the Algorithm IV, which is based on a maximum likelihood criterion, is much lower than the other two algorithms, which use least-squares criteria. Note also that for all the algorithms, the mse decreases as M increases.

2.5.3 Example 3

In this example we demonstrate the application of the joint diagonalization to a blind source separation problem. We consider the following instantaneous mixing model

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \tag{2.73}$$

where $\mathbf{A} \in \mathbb{R}^{N \times N}$ is the mixing system, and $\mathbf{s}(t)$ is a vector of samples of N speech processes, which we assume to have zero mean and are statistically independent from each other. The objective is, given the observed signals $\mathbf{x}(t)$, to separate the speech signals up to a scaling and permutation ambiguity. As explained previously we can use orthogonal joint diagonalization method by first pre-whitening the observed signals using the matrix

$$\mathbf{W} = \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{V}^T \tag{2.74}$$

where $\mathbf{V}$ and $\mathbf{\Lambda}$ correspond respectively to the matrix of eigenvectors and eigenvalues of the matrix $\mathbf{R}_x$, the observed signals' covariance matrix. In practice $\mathbf{R}_x$ can be estimated from the observed signals using

$$\hat{\mathbf{R}}_x = \frac{1}{N_x} \sum_{n=0}^{N_x-1} \mathbf{x}(n)\mathbf{x}^T(n) \tag{2.75}$$

where N_x is the length of the data samples available. We have

$$\begin{aligned} \mathbf{z}(t) &= \mathbf{W}\mathbf{x}(t) \\ &= \mathbf{Q}\mathbf{s}(t) \end{aligned} \tag{2.76}$$

where $\mathbf{Q}$ is some orthogonal matrix and $\mathbf{z}(t)$ are the whitened signals. To estimate the orthogonal matrix $\mathbf{Q}$ we can exploit the non-stationarity of the speech signals[8]by choosing $\mathcal{R}_z = \{\mathbf{R}_z(0), \ldots, \mathbf{R}_z(M-1)\}$, to be the set of covariance matrices of the whitened signal $\mathbf{z}(t)$, evaluated at time epochs $0, \ldots, M-1$, using

$$\hat{\mathbf{R}}_z(m) = \frac{1}{L_z} \sum_{t=0}^{Lz-1} \mathbf{z}_m(t)\mathbf{z}_m^T(t) \quad m = 0, \ldots, M-1 \tag{2.77}$$

where $\mathbf{z}_m(0 : L_z - 1) = \mathbf{z}((m-1)L_z : mL_z - 1)$ and L_z is the epoch length. The orthogonal matrix $\mathbf{Q}$ is then estimated by joint diagonalization of the set of matrices $\hat{\mathbf{R}}_z(0), \ldots, \hat{\mathbf{R}}_z(M-1)$. The separating matrix $\mathbf{B}$ is then calculated from

$$\mathbf{B} = \mathbf{Q}^T\mathbf{W}. \tag{2.78}$$

In this example we choose the sources to be two male and one female speech signals. We also choose the elements of $\mathbf{A} \in \mathbb{R}^{3\times 3}$ randomly from a zero-mean, unit variance Gaussian distribution. We use Algorithm III and Algorithm IV in combination with the pre-whitening stage described above to separate the sources. We also compare our results with the well-known BSS algorithm JADE (Cardoso and Souloumiac, 1993). Note that JADE is a higher-order statistics method and is based on the joint diagonalization of the fourth order *cumulant* matrices of the whitened observed signals. To measure the performance we use the following

[8] We can also use the fact that speech signals are colored signals and so the set of covariance matrices $\tilde{\mathcal{R}}_z = \{\tilde{\mathbf{R}}_z(1), \ldots, \tilde{\mathbf{R}}_z(L-1)\}$, where $\tilde{\mathbf{R}}_z(l) = E[\mathbf{z}(t)\mathbf{z}(t-l)]$ can be used to estimate $\mathbf{Q}$. Notice that in this case $\tilde{\mathbf{R}}_z(l)$ are not necessary positive definite and so Algorithm IV is not applicable.

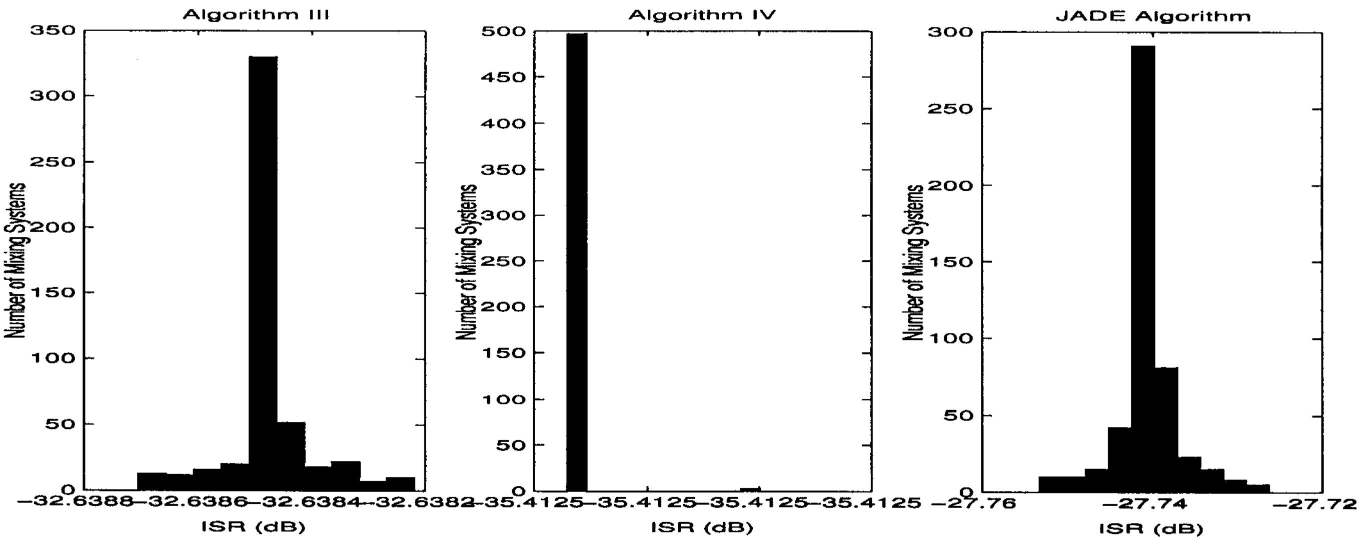

Figure 2.2: Distribution of ISR for random mixing systems for Algorithm III, Algorithm IV and JADE method

separation measure

$$\mathrm{ISR} = 10\log_{10}\left(\frac{1}{N}\sum_{i=1}^{N}\left[\frac{\sum_{j=1}^{N}|c_{ij}|^2}{\max_k(|c_{ik}|^2)} - 1\right]\right) \quad k = 1,\ldots,N \tag{2.79}$$

where c_{ij} is the ij_{th} element of global system $\mathbf{C} = \mathbf{BA}$. Note that equation (2.79) measures the average interference to signal ratio for the all outputs of the separating system $\mathbf{B}$. For this experiment we generated 500 randomly chosen mixing systems, and for each one the resulting mixed signals were used as the inputs for all three algorithms mentioned above. We then measured the separation performance using (2.79). Figure 2.2 shows the distribution of the measured ISR for each of the three algorithms. Also Table 2.2 shows the average performance for each of these three algorithms. As can be seen from the results, Algorithm III and IV outperform the JADE method. Note that, as mentioned before, Algorithm III

	Algorithm III	Algorithm IV	JADE Method
Average ISR (dB)	-32.64	-35.41	-27.74

Table 2.2: Example 3, Average ISR using 500 random mixing systems for algorithms III, IV and JADE method

and the extended Jacobi method, associated with JADE algorithm, both use the same least-squares criterion for the joint diagonalization and as is shown in the previous example, they have the same performance. Nevertheless for blind source separation, the JADE method is based on joint diagonalization of fourth-order cumulant matrices of the observed signals, while in this example we separate the speech signals by jointly diagonalizing the covariance matrices of the observed signals evaluated at different time epochs. For this example the results show that compared to using higher-order statistics, exploiting the second-order non-stationarity of the observed signals gives better separation performance.

Another interesting result of this example is the invariance performance of the proposed BSS algorithms with respect to the mixing system. By "invariance", we mean the performance of the BSS algorithm does not depend on the mixing system. This property is clearly shown in Figure 2.2, where it can be seen the separation performance for all three algorithms does not change (within some tolerance) with any of the 500 mixing systems used in this example. As has been indicated in (Cardoso, 1994), this invariance property is an attribute of the orthogonal joint diagonalization methods.

Notice that in this example, Algorithm IV is the least sensitive to the mixing system compared with the other two algorithms. For this example the standard deviation of the measured ISR for 500 mixing systems for Algorithm IV is 1.3×10^{-7} dB while for the JADE method the same quantity is found to be 3.5×10^{-3} dB.

It is worthy to note that the maximum likelihood procedure of Algorithm IV is only optimum with Gaussian source signals. In this example the sources are speech signals which are known to be non-Gaussian. Nevertheless it seems Algorithm IV still behaves the best among the other algorithms tested. Although here the performance gap between the Algorithm IV and the two other algorithms seems to be smaller compared to previous example where we used Gaussian signals.

2.6 Joint Non-Orthogonal Diagonalization

In the previous sections we discussed methods for joint diagonalization of a set of matrices $\mathbf{R}_1, \ldots, \mathbf{R}_M$ using an orthogonal square matrix $\mathbf{Q}$. We also discussed the application of orthogonal joint diagonalization to the blind source separation problem. Notice that to be able to use an orthogonal joint diagonalizer in a blind source separation context, an extra pre-whitening stage is required. One of the advantages of using an orthogonal joint diagonalizer for a BSS problem is its potential invariance property in the absence of noise. This property was clearly demonstrated through simulation in Example III of the previous section. The down side of using orthogonal BSS algorithms[9], as discussed in (Cardoso, 1994), is that in noisy environments their performance is bounded by the prewhitening stage. In other words, the errors in the whitening step, which may occur due to the noise, cannot be compensated later on by the second step, which is an orthogonal estimator, no matter how well the second step can perform the estimation task. It is shown in (Cardoso, 1994) that for the case where the noise is *iid* the performance of an orthogonal BSS algorithm depends on the matrix $\sigma^2(\mathbf{A}^\dagger\mathbf{A})^{-1}$ where $\mathbf{A}$ is the mixing system and σ^2 is the variance of the noise. Because of this, one may think of improving the whitening stage or using a non-orthogonal stage; i.e., estimating the matrix $\mathbf{A}$, the mixing system, directly and in one step. This explains the motivation behind finding a more general form for the joint diagonalizer matrix $\mathbf{Q}$ rather than assuming that it is square and orthogonal. Following a similar path for finding an orthogonal joint diagonalizer, we can propose least-squares or maximum likelihood criteria. A least-squares criterion for estimating the parameters of a $J \times N$ complex matrix $\mathbf{A}$ from a set of measured covariance matrices $\hat{\mathbf{R}}_1, \ldots, \hat{\mathbf{R}}_M$ is given as

$$\hat{\mathbf{A}} = \arg \min_{\mathbf{\Lambda}(m),\, \mathbf{A}} \sum_{m=1}^{M} ||\hat{\mathbf{R}}_m - \mathbf{A}\mathbf{\Lambda}(m)\mathbf{A}^\dagger||_F^2 \tag{2.80}$$

where $\mathbf{\Lambda}(m) \in \mathbb{R}^{N\times N}$ are diagonal matrices. A few methods have been discussed for optimizing the above criterion including methods in (van der Veen, 2001) and (Yeredor,

[9] By orthogonal BSS algorithms we mean those algorithms that perform a whitening step as a preprocessing stage to an orthogonal matrix estimation.

2000). A simple although effective method has also been developed in Chapter 4 of this thesis. A detailed description of this new algorithm is postponed until later, when we discuss the blind source separation problem for convolutive mixing.

Another possible choice for a non-orthogonal criterion is a maximum likelihood procedure similar to the one discussed in the previous section. A suboptimal method has been discussed in (Pham, 2000) which minimizes an upper bound of the mentioned likelihood criterion.

For the special case of joint diagonalization of two matrices using a non-orthogonal matrix $\mathbf{A}$, a closed form solution may exist as shown below. Assume $\mathbf{R}_1 = \mathbf{R}_1^\dagger \in \mathbb{C}^{J\times J}$ and $\mathbf{R}_2 = \mathbf{R}_2^\dagger \in \mathbb{C}^{J\times J}$ are two Hermitian matrices with $\mathbf{R}_1$ being non-singular. Then it can be shown (Horn and Johnson, 1985) that $\mathbf{R}_1$ and $\mathbf{R}_2$ are jointly diagonalizable if the generalized characteristic polynomial $\det(\lambda\mathbf{R}_1 - \mathbf{R}_2)$ has J distinct zeros. The joint diagonalizer of the set of matrices $(\mathbf{R}_1, \mathbf{R}_2)$ is given by the matrix of the eigenvectors of $\mathbf{R}_1^{-1}\mathbf{R}_2$.

Note when $\mathbf{R}_1 = \mathbf{A}\mathbf{D}_1\mathbf{A}^\dagger$ and $\mathbf{R}_2 = \mathbf{A}\mathbf{D}_2\mathbf{A}^\dagger$ where $\mathbf{D}_1$ and $\mathbf{D}_2$ are diagonal matrices and $\mathbf{A}$ is a square complex matrix then

$$\begin{aligned}\det(\lambda\mathbf{R}_1 - \mathbf{R}_2) &= \det(\mathbf{A}(\lambda\mathbf{D}_1 - \mathbf{D}_2)\mathbf{A}^\dagger)\\ &= (\det(\mathbf{A}))^2 \det(\lambda\mathbf{D}_1 - \mathbf{D}_2)\\ &= (\det(\mathbf{A}))^2 \prod_{i=1}^{J}(\lambda d_{ii}^1 - d_{ii}^2)\end{aligned} \tag{2.81}$$

where d_{ii}^1 and d_{ii}^2 are respectively the i_{th} diagonal elements of $\mathbf{D}_1$ and $\mathbf{D}_2$. If the above polynomial has J distinct zeros then we should have $\mathbf{D}_1 \neq \lambda\mathbf{D}_2$ for some non-zero λ. Notice this is the same condition we obtained in Theorem 1. When applied to this case, it requires the two vectors $\text{vec}\{\mathbf{D}_1\}$ and $\text{vec}\{\mathbf{D}_2\}$ to be linearly independent. Also note that this condition is only necessary for the above polynomial to have J distinct zeros but it is not sufficient (it is sufficient only when $J = 2$). This closed form joint diagonalization algorithm is useful for initialization purposes as is explained in later chapters.

2.7 Summary

In this chapter we discussed joint diagonalization methods and their application to the blind source separation problem. We proposed four new methods for joint orthogonal diagonalization of a set of symmetric matrices based on optimization methods over the Stiefel manifold. The first three algorithms are based on a least-squares criterion while the fourth algorithm uses a maximum likelihood method. We showed that the maximum likelihood method shows superior performance compared to least-squares based algorithms at least for white Gaussian noise and speech signals. We also compared our results with the extended Jacobi method which is used for the joint diagonalization in the JADE method. The simulation showed that the first three algorithms have exactly the same performance as the extended Jacobi method, while the fourth algorithm outperforms it. We also compared our results with the JADE method for a blind source separation scenario where we exploited the non-stationarity of speech signals, in the case of algorithms III and IV, and their non-Gaussianity, in the case of the JADE method, to separate them.

The results showed that, for this example, exploiting the non-stationarity of speech signals can result in a better separation performance. We also showed the invariance property of the orthogonal joint diagonalizer through these simulations.

Chapter 3

Blind Identification of MIMO Systems

In this Chapter we discuss a frequency domain method for blind identification of MIMO convolutive channels driven by white quasi-stationary sources. We demonstrate that by using the second-order statistics of the channel outputs, under mild conditions on the non-stationarity of sources, and under the condition that channel is column-wise coprime, the impulse response of the MIMO channel can be identified up to an inherent scaling and permutation ambiguity. We further present an efficient, two step frequency domain algorithm for identifying the channel. We show that the new algorithm, under the stated assumptions, does not experience the problem of frequency-dependent, arbitrary permutations and scaling factors across the frequency spectrum as is the case with previous frequency domain algorithms. Numerical simulations are presented to demonstrate the performance of the new algorithm.

3.1 Introduction

Multichannel blind identification has been of great interest to both the communications and signal processing communities and there have been numerous publications in both societies on this subject. Some of the literature on blind identification was already discussed in the introductory chapter of this thesis. See also (Tong and Perreau, 1998) for a review of recent blind channel estimation and identification techniques.

In this chapter we consider the problem of blind identification of MIMO channels with finite memory. Previous work in this area can be divided into two groups. The first group uses higher-oder statistical (HOS) methods that exploit the higher–order moments (or higher–order spectra) of the output signals to identify the channel; e.g., (Tugnait, 1999)(Chen and Petropulu, 2001). The second group are the second–order statistical (SOS) methods that rely only on the second–order moments of the output signals to identify the channel (Sahlin and Broman, 2000)(Hua *et al.*, 2001)(Gorokhov and Loubaton, 1997).

The proposed method is a frequency domain approach that exploits second-order non-stationarity of the input signals. Previously, for both HOS and SOS identification methods, the inputs have mostly been assumed stationary. However, some methods have been proposed that exploit non-stationarity of the input signals. So far, most blind identification methods that exploit non–stationarity (mostly in the form of *cyclostationarity*) address only the SISO case. In (Pham and Cardoso, 2001) (Abed_Meraim *et al.*, 2001), methods have been proposed for blind source separation of instantaneously mixed, non-stationary (cyclostationary in the second reference) signals. References (Parra and Spence, 2000) and (Rahbar and Reilly, 2001b) (Ma *et al.*, 2000) consider blind source separation of colored non-stationary signals when the mixing system is convolutive.

In this chapter we exploit the non-stationarity of the observed signals for blind identification of MIMO systems. We assume that the statistics of the input signals are slowly varying with time; i.e., we assume that they are *quasi-stationary* (Papoulis, 1984). Furthermore, we rely only on the second–order statistics of the input signal. This gives us the advantages of SOS methods, and also permits identification in situations where stationary-based SOS and HOS methods fail; e.g., when the input signals are temporally white and Gaussian distributed. Although the main focus of this chapter is on white, non-stationary signals, we show that under some additional conditions the same algorithm can also be applied to colored non-stationary signals.

In this chapter we demonstrate sufficient identifiability conditions for blind identifiability of a MIMO system in the frequency domain under a second-order non-stationarity assumption of the inputs. We also prove that a limited number of frequency samples are enough to identify the channel and to this end we derive an upper bound on the smallest number of frequency samples sufficient for blind identification of the MIMO system. This bound is lower than what has been perviously used in frequency domain blind identification or blind source separation methods ((Chen and Petropulu, 2001) (Parra and Spence, 2000)) and results in significant computational savings for the proposed algorithm.

As mentioned in the introductory chapter of this thesis, the main difficulties with frequency–domain blind identification of MIMO channels are the arbitrary column permutations and scaling ambiguities of the estimated frequency response of the channel at each frequency bin. In this chapter we exploit the quasi–stationary nature of input signals, such that the proposed algorithm results in a *common* permutation for the estimated channel frequency response across all frequency bins. Further we demonstrate that if the channel is column-wise coprime, then the problem of arbitrary scaling factors across the frequency bins can be resolved, thus avoiding the limitations of frequency domain methods.

This chapter is organized in the following manner: The problem formulation including the set of required assumptions is presented in Section 3.2. Section 3.3 establishes channel identifiability results based on only the second–order statistics and the quasi-stationarity property of the input signals. In Section 3.4 we present a two-stage frequency domain algorithm for blind identification of MIMO channels. Simulation results are described in Section 3.5. The first simulation scenario is a synthetic data case where two inputs are quasi–stationary zero–mean Gaussian noise signals. The second simulation uses colored sources, which are created by passing the white signals in the first simulation through an AR filter. The third simulation uses two speech signals as inputs. In each of these cases, the underlying channel was successfully identified. We also compare our results with those obtained using the HOS blind identification method in (Chen and Petropulu, 2001). Conclusions and final remarks are presented in Section 3.6.

3.2 Problem statement

We consider the following N-source J-sensor MIMO linear model for the received signal for the convolutive mixing problem:

$$\mathbf{x}(t) = \sum_{l=0}^{L} \mathbf{H}(l)\mathbf{s}(t-l) + \mathbf{n}(t) \quad t \in \mathbb{Z} \tag{3.1}$$

where $\mathbf{x}(t) = (x_1(t), \cdots, x_J(t))^T \in \mathbb{R}^{J\times 1}$ is the vector of observed signals,
$\mathbf{s}(t) = (s_1(t), \cdots, s_N(t))^T \in \mathbb{R}^{N\times 1}$ is the vector of sources, $\mathbf{H}(t) \in \mathbb{R}^{J\times N}$ is the channel matrix where the maximum order of its elements is L and $\mathbf{n}(t) = (n_1(t), \cdots, n_J(t))^T \in \mathbb{R}^{J\times 1}$ is the additive noise vector. The objective is to estimate the $\mathbf{H}(t)$ up to a scaling and permutation factor from the observed signals $\mathbf{x}(t)$. In other words, we are interested in finding $\hat{\mathbf{H}}(t)$ such that for all $0 \leq t \leq L$ we have

$$\hat{\mathbf{H}}(t) = \mathbf{H}(t)\mathbf{\Pi}\mathbf{D} \tag{3.2}$$

where $\mathbf{D} \in \mathbb{R}^{N\times N}$ and $\mathbf{\Pi} \in \mathbb{R}^{N\times N}$ are respectively constant diagonal and permutation matrices. In the frequency domain this is equivalent to finding an $\hat{\mathbf{H}}(\omega) \in \mathbb{C}^{J\times N}$ such that:

$$\hat{\mathbf{H}}(\omega) = \mathbf{H}(\omega)\mathbf{\Pi}\mathbf{D} \quad \forall\, \omega \in [0, \pi) \tag{3.3}$$

where $\mathbf{H}(\omega)$ is the DTFT of the $\mathbf{H}(t)$. Notice that in (3.3), since we assume that the elements of the channel are real numbers, we only need to estimate $\mathbf{H}(\omega)$ over half of the frequency range; i.e., $\omega \in [0, \pi)$.

3.2.1 Main Assumptions

A0: $J \geq N \geq 2$; i.e, we have at least as many sensors as sources and number of the sources is at least two.

A1: The sources $\mathbf{s}(t)$ are zero mean, second-order quasi-stationary white signals. The cross–spectral density matrices of the sources $\mathbf{P}_s(\omega, m)$ are diagonal for all ω and m where ω denotes frequency and m is the time epoch index.

A2: Let $\lambda_i(m)$ denote the variance of the i_{th} source at epoch m. We assume the matrix $\mathbf{\Gamma}$ given by:

$$\mathbf{\Gamma} = \begin{pmatrix} \lambda_1(0) & . \; . \; . & \lambda_1(M-1) \\ . & & . \\ . & & . \\ \lambda_N(0) & . \; . \; . & \lambda_N(M-1) \end{pmatrix} \in \mathbb{R}^{N \times M} \quad M > N \tag{3.4}$$

has full row rank where M is the total number of epochs, available from the observed data.

A3: The channel is modelled by a causal FIR system of the form $\mathbf{H}(t) = [\mathbf{h}_1(t), ..., \mathbf{h}_N(t)]$ and does not change over the entire observation interval. Also $\mathbf{H}(\omega)$, the DTFT of $\mathbf{H}(t)$, has full column rank for all $\omega \in [0, 2\pi)$.

A4: The noise $\mathbf{n}(t)$ is zero mean, *iid* across sensors, with power σ^2. The noise is assumed independent of the sources.

A5: $\mathbf{H}(z)$, the z-transform of $\mathbf{H}(t)$, is column–wise coprime, i.e. the elements in each column of $\mathbf{H}(z)$ do not share common zeros.

Assumption $A1$ is the core assumption here. As is shown later, this non-stationarity assumption enables us to identify a MIMO channel using only the second-order statistics of the observed signal. Although in our assumptions we consider white signals, the identifiability results and the algorithm can be extended to the colored signal case under the condition that the spectra of the sources stay constant over the observation interval and only their variances change between epochs. As is shown later, this condition will guarantee a uniform permutation across all frequency bins. The reason behind imposing assumptions $A2, \ldots, A4$ will become clear when we explain the identifiability proof. Notice that assumptions $A1, \ldots, A4$ are sufficient to identify the frequency response of a MIMO channel up to a constant permutation but a frequency dependent scaling ambiguity. This means that if we use the estimated channel to recover the sources, the outputs correspond to a separated but filtered version of the original sources. Assumption A5 enables us to remove the frequency dependent scaling ambiguity so the channel can be identified up to a constant scaling and permutation ambiguity which is the best that can be achieved in MIMO blind identification problems.

3.3 Blind Identifiability

In this section we present frequency domain blind identifiability results based on the above assumptions using only the second-order statistics of the observed signals. Let $\mathbf{P}_x(\omega, m)$ represent the cross-spectral density matrix of the observed signal at frequency ω and time epoch m. Using $A1, A3$ and $A4$ we have:

$$\mathbf{P}_x(\omega, m) = \mathbf{H}(\omega)\mathbf{P}_s(\omega, m)\mathbf{H}^\dagger(\omega) + \sigma^2\mathbf{I} \tag{3.5}$$

where $\mathbf{P}_s(\omega, m)$ by assumption is diagonal for all ω and m. Notice that for white sources we have $\mathbf{P}_s(\omega, m) = \mathbf{\Lambda}(m)$ where $\mathbf{\Lambda}(m) \in \mathbb{R}^{N\times N}$ is a diagonal matrix for each m and its i_{th} diagonal value, $\lambda_i(m)$, represents the variance of the i_{th} source at epoch m. Based on assumption $A2$ we can immediately see that the vectors $\text{diag}\{\mathbf{\Lambda}(m)\}, \ \ m = 0, \ldots, M-1$, span $\mathbb{R}^N$. For identifiability purposes, we assume that σ^2 is known although for $J > N$, σ^2 can be estimated from the smallest eigenvalue of the matrix $\mathbf{P}_x(\omega, m)$; so for now we consider the following noise free case

$$\mathbf{P}_x(\omega_k, m) = \mathbf{H}(\omega_k)\mathbf{\Lambda}(m)\mathbf{H}^\dagger(\omega_k) \tag{3.6}$$

where $\omega_k = (2\pi k)/K$ is the discretized version of ω and K is the number of frequency samples.

Theorem 2 *Consider the cross spectral density matrices*

$$\mathbf{P}_x(\omega_k, m) = \mathbf{H}(\omega_k)\mathbf{\Lambda}(m)\mathbf{H}^\dagger(\omega_k) \tag{3.7}$$

for $k = 0, \ldots, K-1$ and $m = 0, \ldots, M-1$. Under the assumptions that the $\mathbf{H}(\omega_k) \in \mathbb{C}^{J\times N}$ have full column rank and the vectors $\text{diag}\{\mathbf{\Lambda}(m)\} \in \mathbb{R}^N, \ \ m = 0, \ldots, M-1$, span $\mathbb{R}^N$, if there exist matrices $\mathbf{B}(\omega_k) \in \mathbb{C}^{J\times N}$ and $\tilde{\mathbf{\Lambda}}(m) \in \mathbb{R}^{N\times N}$, with $\tilde{\mathbf{\Lambda}}(m)$ diagonal, such that

$$\mathbf{P}_x(\omega_k, m) = \mathbf{B}(\omega_k)\tilde{\mathbf{\Lambda}}(m)\mathbf{B}^\dagger(\omega_k) \tag{3.8}$$

then $\mathbf{B}(\omega_k)$ *must be related to* $\mathbf{H}(\omega_k)$ *in the following way:*

$$\mathbf{B}(\omega_k) = \mathbf{H}(\omega_k)\mathbf{\Pi}\mathbf{D}e^{-j\mathbf{S}_k} \tag{3.9}$$

where $\mathbf{\Pi} \in \mathbb{R}^{N\times N}$ *is a permutation matrix,* $\mathbf{S}_k \in \mathbb{R}^{N\times N}$ *and* $\mathbf{D} \in \mathbb{R}^{N\times N}$ *are diagonal matrices with* $\mathbf{D}$ *being non-singular.*

Proof:

The proof is similar to that of Theorem 1. It must be shown that

$$\mathbf{B}(\omega_k)\tilde{\mathbf{\Lambda}}(m)\mathbf{B}^\dagger(\omega_k) = \mathbf{H}(\omega_k)\mathbf{\Lambda}(m)\mathbf{H}^\dagger(\omega_k) \tag{3.10}$$

implies (3.9).

For any sequence of scalars $a = (a_0, ..., a_{M-1})$, define the diagonal matrices

$$\mathbf{\Sigma}_a = \sum_{m=0}^{M-1} a_m\mathbf{\Lambda}(m), \;\; \tilde{\mathbf{\Sigma}}_a = \sum_{m=0}^{M-1} a_m\tilde{\mathbf{\Lambda}}(m), \quad a_m \in \mathbb{R}. \tag{3.11}$$

Therefore, an arbitrary linear combination of (3.10) over different epochs can be written as

$$\mathbf{B}(\omega_k)\tilde{\mathbf{\Sigma}}_a\mathbf{B}^\dagger(\omega_k) = \mathbf{H}(\omega_k)\mathbf{\Sigma}_a\mathbf{H}^\dagger(\omega_k). \tag{3.12}$$

Since the vectors diag$\{\mathbf{\Lambda}(m)\}$ $m = 0, \ldots, M-1$ span $\mathbb{R}^N$, $\mathbf{\Sigma}_a$ can be made equal to any real valued diagonal matrix by an appropriate choice of a. In this instance, choose a such that $\mathbf{\Sigma}_a$ is the identity matrix. Then since by assumption $\mathbf{H}(\omega_k)$ has full column rank for all $k = 0, ..., K-1$, the RHS of (3.12) has rank N implying $\mathbf{B}(\omega_k)$ has full column rank for all k. In particular, $\mathbf{B}^+(\omega_k)\mathbf{B}(\omega_k)$ is the identity matrix. Thus $\mathbf{B}(\omega_k)$ can be cancelled from the LHS of (3.12), giving

$$\tilde{\mathbf{\Sigma}}_a = \mathbf{C}_k\mathbf{\Sigma}_a\mathbf{C}_k^\dagger, \tag{3.13}$$

where $\mathbf{C}_k = \mathbf{B}^+(\omega_k)\mathbf{H}(\omega_k)$. Observe that if $\mathbf{\Sigma}_a$ is the identity matrix then (3.12) implies $\tilde{\mathbf{\Sigma}}_a$ has full rank and thus (3.13) implies $\mathbf{C}_k$ is invertible for all k.

It is first shown that

$$\mathbf{C}_k^{-1} = \mathbf{\Pi}\mathbf{D}e^{-j\mathbf{S}_k} \tag{3.14}$$

where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{D}$ and $\mathbf{S}_k$ are diagonal matrices. For any i, choose a such that all elements of $\mathbf{\Sigma}_a$ are zero except for the i_{th} diagonal element which is unity. Then $\mathbf{C}_k\mathbf{\Sigma}_a\mathbf{C}_k^\dagger = \mathbf{c}_i(k)\mathbf{c}_i(k)^\dagger$ where $\mathbf{c}_i(k)$ is the i_{th} column of $\mathbf{C}_k$. Moreover, since the LHS of (3.13) is diagonal, all the off-diagonal elements of $\mathbf{c}_i(k)\mathbf{c}_i^\dagger(k)$ are zero. Because $\mathbf{c}_i(k)\mathbf{c}_i^\dagger(k)$ has rank one at most, it can have at most one non-zero diagonal element. It follows immediately that every column of $\mathbf{C}_k$ has precisely one non-zero element, and moreover, because $\mathbf{C}_k$ is invertible, every row has precisely one non-zero element too. Clearly the same is true for $\mathbf{C}_k^{-1}$; i.e.,

$$\mathbf{C}_k^{-1} = \mathbf{\Pi}\mathbf{D}_k \tag{3.15}$$

where $\mathbf{D}_k$ are non-singular diagonal matrices for all k. Because the LHS of (3.13) is independent of k, $\mathbf{C}_k\mathbf{C}_k^\dagger$ and thus $\mathbf{D}_k\mathbf{D}_k^\dagger$ is independent from k too; this means only the phase and not the magnitude of elements of $\mathbf{D}_k$ change with k; i.e.,

$$\mathbf{D}_k = \mathbf{D}e^{-j\mathbf{S}_k}. \tag{3.16}$$

Substituting $\mathbf{C}_k = \mathbf{B}^+(\omega_k)\mathbf{H}(\omega_k)$ into (3.14) and rearranging gives

$$\mathbf{B}(\omega_k) = \mathbf{B}(\omega_k)\mathbf{B}^+(\omega_k)\mathbf{H}(\omega_k)\mathbf{\Pi}\mathbf{D}e^{-j\mathbf{S}_k}. \tag{3.17}$$

Notice that $\mathbf{B}(\omega_k)\mathbf{B}^+(\omega_k)$ is a projector matrix onto the range space of $\mathbf{B}(\omega_k)$. Choosing $\mathbf{\Sigma}_a$ to be the identity matrix in (3.12) reveals that the range space of $\mathbf{B}(\omega_k)$ must contain the range space of $\mathbf{H}(\omega_k)$. Therefore,

$$\mathbf{B}(\omega_k)\mathbf{B}^+(\omega_k)\mathbf{H}(\omega_k) = \mathbf{H}(\omega_k) \tag{3.18}$$

and (3.9) follows immediately from (3.18) and (3.17). □

In (3.9) $\mathbf{D}e^{-j\mathbf{S}_k}$ represents a frequency dependent diagonal matrix where the magnitudes of the diagonal values are constant and only their phase varies with k. Theorem 2 has one important implication which is that under the assumptions $A0, \ldots, A4$, equation (3.6) can be used to estimate the channel up to a constant permutation, but a frequency dependent phase ambiguity, across all frequency bins. Thus, the commonly–experienced difficulty with frequency domain approaches to blind identification problems of ensuring a constant permutation over all frequency bins can be alleviated with the proposed approach.

We can easily extend the above theorem to colored sources under a more restrictive assumption. More specifically we assume that only the scale of the power spectral density of each source changes with time. In other words we have:

$$\mathbf{P}_s(\omega_k, m) = \mathbf{\Lambda}_1(\omega_k)\mathbf{\Lambda}(m) \tag{3.19}$$

where $\mathbf{\Lambda}_1(\omega_k)$ and $\mathbf{\Lambda}(m)$ are diagonal matrices for ω_k and m. Based on this, the power spectral density of the observed signals can be written as:

$$\mathbf{P}_x(\omega_k, m) = \mathbf{H}(\omega_k)\mathbf{\Lambda}_1(\omega_k)\mathbf{\Lambda}(m)\mathbf{H}^\dagger(\omega_k). \tag{3.20}$$

Define $\mathbf{H}_1(\omega_k) = \mathbf{H}(\omega_k)\mathbf{\Lambda}_1^{\frac{1}{2}}(\omega_k)$ then $\mathbf{P}_x(\omega_k, m) = \mathbf{H}_1(\omega_k)\mathbf{\Lambda}(m)\mathbf{H}_1^\dagger(\omega_k)$ and based on Theorem 2 we have, for any $\mathbf{B}(\omega_k)$ satisfying (3.8)

$$\begin{aligned} \mathbf{B}(\omega_k) &= \mathbf{H}_1(\omega_k)\mathbf{\Pi}\mathbf{D}e^{-j\mathbf{S}_k} \\ &= \mathbf{H}(\omega_k)\mathbf{\Lambda}_1^{\frac{1}{2}}(\omega_k)\mathbf{\Pi}\mathbf{D}e^{-j\mathbf{S}_k}. \end{aligned} \tag{3.21}$$

Notice here that (3.21) in its general form can be written as

$$\mathbf{B}(\omega_k) = \mathbf{H}(\omega_k)\mathbf{\Pi}\mathbf{D}(\omega_k) \tag{3.22}$$

where $\mathbf{D}(\omega_k) = \mathbf{\Lambda}_1^{\frac{1}{2}}(\omega_k)\mathbf{D}e^{-j\mathbf{S}_k}$ is diagonal for all ω_k. In other words, when the sources are colored $\mathbf{H}(\omega)$ can be identified up to a constant permutation and a frequency dependent scaling factor of it's columns. Notice that Equation (3.9) can be considered as a special case of (3.22).

We now show that under the additional assumption $A5$, $\mathbf{D}(\omega_k)$ in equation (3.22) is constant for all frequency bins.

Theorem 3 *Let* $\mathbf{H}(\omega) = \sum_{t=0}^{L} \mathbf{H}(t)e^{-j\omega t} \in \mathbb{C}^{J\times N}$ *be the transfer function of a MIMO FIR channel of order* L*. Similarly let* $\mathbf{B}(\omega) \in \mathbb{C}^{J\times N}$ *be the transfer function of a MIMO FIR channel of unknown order. Assume that* $\mathbf{B}(\omega)$ *and* $\mathbf{H}(\omega)$*, evaluated at* K *uniformly spaced samples, satisfy*

$$\mathbf{B}(\omega_k) = \mathbf{H}(\omega_k)\mathbf{\Pi}\mathbf{D}(\omega_k), \quad \omega_k = \frac{2\pi k}{K}, \quad k = 0, \ldots, K-1 \tag{3.23}$$

for some permutation matrix $\mathbf{\Pi} \in \mathbb{R}^{N\times N}$ *and non-singular diagonal matrices* $\mathbf{D}(\omega_k) \in \mathbb{C}^{N\times N}$*. If* $K \geq 2L+1$ *and* $\mathbf{B}(z)$ *and* $\mathbf{H}(z)$*, the corresponding z-transforms of* $\mathbf{B}(\omega)$ *and* $\mathbf{H}(\omega)$*, are column-wise coprime then* (3.23) *implies*

$$\mathbf{B}(t) = \mathbf{H}(t)\mathbf{\Pi}\mathbf{D}, \quad t = 0,\dots,L \tag{3.24}$$

for some non-singular diagonal matrix $\mathbf{D} \in \mathbb{R}^{N\times N}$.

Proof:

Let $\mathbf{b}(\omega_k)$ be an arbitrary column of $\mathbf{B}(\omega_k)$ and let $\mathbf{h}(\omega_k)$ be the corresponding column of $\mathbf{H}(\omega)\mathbf{\Pi}$. It assumed that elements of $\mathbf{b}(z)$, the corresponding z-transform of $\mathbf{b}(\omega)$, are coprime, as are the elements of $\mathbf{h}(z)$, the corresponding z-transform of $\mathbf{h}(\omega)$. It will be proved that

$$\mathbf{b}(\omega_k) = d_k\mathbf{h}(\omega_k), \quad \omega_k = \frac{2\pi k}{K}, \; d_k \in \mathbb{C}, \; d_k \neq 0, \; k = 0,\dots,K-1 \tag{3.25}$$

implies $\mathbf{b}(t) = \alpha\mathbf{h}(t), \; t = 0,\dots,L$, for some non-zero $\alpha \in \mathbb{C}$ provided $K \geq 2L+1$, where L is the order of $\mathbf{h}(z)$. The theorem then follows immediately.

Let $\mathbf{b}(t)$ and $\mathbf{h}(t)$ be the impulse responses of $\mathbf{b}(\omega)$ and $\mathbf{h}(\omega)$ respectively, so that

$$\mathbf{b}(\omega) = \sum_{t=0}^{\tilde{L}} \mathbf{b}(t)e^{-j\omega t}, \qquad \mathbf{h}(\omega) = \sum_{t=0}^{L} \mathbf{h}(t)e^{-j\omega t} \tag{3.26}$$

where the order $\tilde{L}$ is unknown but finite. The proof below repeatedly uses the fact that, for any $Q \geq \tilde{L}-1$, a SIMO FIR channel of order $\tilde{L}$ having impulse response $\mathbf{b}(0),\dots,\mathbf{b}(\tilde{L}) \in \mathbb{R}^J$ is coprime if and only if the block Sylvester matrix

$$\mathcal{S}_Q(\mathbf{b}_i) = \begin{pmatrix} \mathbf{b}(0) & \mathbf{b}(1) & \dots & \mathbf{b}(\tilde{L}) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{b}(0) & \dots & \mathbf{b}(\tilde{L}-1) & \mathbf{b}(\tilde{L}) & \dots & \mathbf{0} \\ \vdots & \ddots & \ddots & \dots & \ddots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{b}(0) & \mathbf{b}(1) & \dots & \mathbf{b}(\tilde{L}) \end{pmatrix} \in \mathbb{R}^{J(Q+1)\times(\tilde{L}+Q+1)} \tag{3.27}$$

has full column rank (Serpedin and Giannakis, 1999).

It is first proved that $\tilde{L} \leq L$. Assume to the contrary that $\tilde{L} > L$. By substituting (3.26) into (3.25), it follows that, for the choice $Q = K - \tilde{L} - 1$,

$$\mathcal{S}_Q(\mathbf{b})\mathbf{F} = [\mathcal{S}_Q(\mathbf{h}) \;\; \mathbf{0}_{J(Q+1)\times(\tilde{L}-L)}]\mathbf{F}\mathbf{\Delta} \tag{3.28}$$

where $\mathcal{S}_Q(\mathbf{h})$ is a $J(Q+1) \times (L+Q+1)$ matrix having the same form as $\mathcal{S}_Q(\mathbf{b})$, $\mathbf{F}$ is the non-singular DFT matrix

$$\mathbf{F} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & e^{-j2\pi/K} & \dots & e^{-j2\pi(K-1)/K} \\ 1 & e^{-j4\pi/K} & \dots & e^{-j4\pi(K-1)/K} \\ \vdots & \vdots & \dots & \vdots \\ 1 & e^{-j2\pi(K-1)/K} & \dots & e^{-j2\pi(K-1)^2/K} \end{pmatrix} \tag{3.29}$$

and $\mathbf{\Delta} = \operatorname{diag}\{d_0, \dots, d_{K-1}\}$. The LHS of (3.28) has full column rank because $\mathbf{b}(z)$ is coprime and $Q = K - \tilde{L} - 1 \geq L - 1$ by assumption that $K \geq 2\tilde{L} + 1$ and $\tilde{L} > L$. However, the RHS of (3.28) clearly does not have full column rank, a contradiction.

This time, choose $Q = K - L - 1$. Analogous to (3.28), but this time because it is known $\tilde{L} \geq L$,

$$[\mathcal{S}_Q(\mathbf{b}) \;\; \mathbf{0}_{J(Q+1)\times(L-\tilde{L})}]\mathbf{F} = \mathcal{S}_Q(\mathbf{h})\mathbf{F}\mathbf{\Delta}. \tag{3.30}$$

Define $\mathbf{C} = \mathbf{F}\mathbf{\Delta}\mathbf{F}^{-1}$; since $\mathbf{F}$ is a DFT matrix, $\mathbf{C}$ is circulant:

$$\mathbf{C} = \begin{pmatrix} c(0) & c(1) & \dots & c(K-1) \\ c(K-1) & c(0) & \dots & c(K-2) \\ \vdots & \ddots & \dots & \vdots \\ c(1) & c(2) & \dots & c(0) \end{pmatrix}. \tag{3.31}$$

Then

$$[\mathcal{S}_Q(\mathbf{b}) \;\; \mathbf{0}_{J(Q+1)\times(L-\tilde{L})}] = \mathcal{S}_Q(\mathbf{h})\mathbf{C}. \tag{3.32}$$

Even if $\tilde{L} = L$, the first JQ elements of the last column of the LHS of (3.32) are zero. Therefore,

$$\mathcal{S}_{Q-1}(\mathbf{h})\mathbf{c} = \mathbf{0}, \tag{3.33}$$

where $\mathbf{c} \in \mathbb{R}^{K-1}$ is the vector formed from the first $K-1$ elements of the last column of $\mathbf{C}$. Because $Q - 1 = K - L - 2 \geq L - 1$ by assumption that $K \geq 2L + 1$, $\mathcal{S}_{Q-1}(\mathbf{h})$ has full column rank, and in particular, $\mathbf{c} = \mathbf{0}$. Since $\mathbf{C}$ is circulant, $\mathbf{c} = \mathbf{0}$ implies $\mathbf{C} = \alpha\mathbf{I}$ for some $\alpha \in \mathbb{R}$. It follows from (3.32) that $\mathbf{b}(t) = \alpha\mathbf{h}(t)$ for $t = 0, ..., L$. Notice that $\alpha \neq 0$

for otherwise the coprimeness of the elements of $\mathbf{b}(z)$ would be contradicted. The theorem then follows □

One implication of Theorem 3 is that under the further assumption that the columns of $\mathbf{H}(z)$ are coprime we can remove the frequency dependent scaling ambiguity $\mathbf{D}(\omega_k)$ given in equation (3.22). Note that $d_{ii}(\omega_k)$, the i_{th} diagonal value of $\mathbf{D}(\omega_k)$, is common to the elements of the i_{th} column of $\mathbf{H}(\omega_k)$. If $d_{ii}(\omega_k)$ varies with frequency then it will appear as common zeros (or poles) between the elements of the i_{th} column of $\mathbf{H}(z)$, thus violating the assumptions.

Another important result that can be deduced from Theorem 3 is that the number of frequency bins required for identification need only be at least $2L+1$. This number is significantly less than what was used in previous frequency domain approaches (Chen and Petropulu, 2001)(Parra and Spence, 2000); hence, significant computational savings can be realized with the proposed identification procedure.

3.4 The Algorithm

In this section we propose a two step algorithm. The first step estimates the channel up to a frequency dependent scaling ambiguity and constant permutation factor. In other words, the first step finds a $\mathbf{B}(\omega_k)$ such that:

$$\mathbf{B}(\omega_k) = \mathbf{H}(\omega_k)\mathbf{\Pi}\mathbf{D}(\omega_k) \tag{3.34}$$

where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{D}(\omega_k)$ represents the frequency dependent scaling ambiguity. The second step removes the frequency dependent scaling ambiguity $\mathbf{D}(\omega_k)$ by exploiting the column-wise coprimeness of the channel $\mathbf{H}(z)$.

3.4.1 Step I

For the first part of the algorithm we propose to estimate $\mathbf{B}(\omega_k)$ via the following weighted least-squares criterion:

$$\min_{\mathbf{B}(\omega_k),\mathbf{\Lambda}(m)} \sum_{k=0}^{K-1}\sum_{m=0}^{M-1} W_k||\hat{\mathbf{P}}_x(\omega_k,m) - \mathbf{B}(\omega_k)\mathbf{\Lambda}(m)\mathbf{B}^\dagger(\omega_k)||_F^2 \tag{3.35}$$

where $\hat{\mathbf{P}}_x(\omega_k,m)$ is a sample estimate of the observed signal cross spectral density matrix at frequency bin ω_k and time epoch m, $\mathbf{\Lambda}(m)$ is a diagonal matrix representing the unknown cross-spectral density matrix of the sources at epoch m, and $W_k, k = 0,\dots,K-1$ are positive scalars.

The rational for introducing the weight factor W_k into the optimization criterion is to emphasize the contribution of those frequency bins that are known to give a more reliable estimate of the channel. In case where such prior information is not available we set $W_k = 1$ for all $k = 0,\dots,K-1$.

To estimate the observed signals' cross-spectral density matrices, $\mathbf{P}_x(\omega_k,m)$, $m = 0,\dots,M-1$, $k = 0,\dots,K-1$, we first divide the observed sequence into M epochs, where stationarity can be assumed within the epoch but not over more than one epoch. We then apply the following formula to estimate the cross-spectral density matrix at the m_{th} epoch:

$$\hat{\mathbf{P}}_x(\omega_k,m) = \frac{1}{N_s}\sum_{i=0}^{N_s-1} \mathbf{x}_i(\omega_k,m)\mathbf{x}_i^\dagger(\omega_k,m) \tag{3.36}$$

where

$$\mathbf{x}_i(\omega_k,m) = \sum_{t=-\infty}^{\infty} \mathbf{x}(t)w(t-iT_s-mT_b)e^{-j\omega_k t} \quad k = 0,\dots,K-1 \tag{3.37}$$

where N_s is the number of overlapping windows inside each epoch, T_b is the size of each epoch, T_s is the time shift between two overlapping windows and $w(t)$ is the windowing sequence. Note that $\mathbf{x}_i(\omega_k,m))$ in (3.37) is computed using the FFT.

To minimize (3.35) we propose an alternating least-squares method (ALS). The basic idea behind ALS is that in the optimization process we divide the parameter space into multiple sets. At each iteration of the algorithm we minimize the criterion with respect to one set conditioned on the previously estimated sets of the parameters. The newly estimated set is then used to update the remaining sets. This process continues until convergence is achieved. Notice that the convergence of ALS is guaranteed because at each iteration we either improve or maintain the value of the cost function (Sidiropoulos *et al.*, 2000). Alternating least-squares methods have been used for blind source separation of finite alphabet signals in (Talwar *et al.*, 1996) and (Li and Sidiropoulos, 2000) and *parallel factor analysis* (PARAFAC) in (Sidiropoulos *et al.*, 2000). The advantage of using ALS (rather than gradient based optimization methods) is that it is simple to implement and there are no parameters to adjust. One disadvantage, shared by most non-linear optimization techniques, is that unless it is properly initialized, it can fall into a local minimum. Later on in this section we introduce a procedure for initializing the algorithm to diminish this possibility.

The quantity $\mathbf{B}(\omega_k)\mathbf{\Lambda}(m)\mathbf{B}^{\dagger}(\omega_k)$ in (3.35) can be written as $\sum_{i=1}^{N}\lambda_i(m)\mathbf{b}_i(\omega_k)\mathbf{b}_i(\omega_k)^{\dagger}$, where $\mathbf{b}_i(\omega_k)$ is the i_{th} column of $\mathbf{B}(\omega_k)$. Then equation (3.35) can be written as:

$$\min_{\mathbf{g}_i(\omega_k)\in\Omega,\mathbf{d}(m)} \sum_{k=0}^{K-1}\sum_{m=0}^{M-1} W_k||\hat{\mathbf{p}}_x(\omega_k,m)-\mathbf{G}(\omega_k)\mathbf{d}(m)||_2^2 \tag{3.38}$$

where $\hat{\mathbf{p}}_x(\omega_k,m)=\text{vec}\{\hat{\mathbf{P}}_x(\omega_k,m)\}$ is a $J^2\times 1$ column vector, $\mathbf{g}_i(\omega_k)$ is the i_{th} column of $\mathbf{G}(\omega_k)=[\text{vec}\{\mathbf{b}_1(\omega_k)\mathbf{b}_1^{\dagger}(\omega_k)\},\ldots,\text{vec}\{\mathbf{b}_N(\omega_k)\mathbf{b}_N^{\dagger}(\omega_k)\}]$ which is a $J^2\times N$ tall matrix, and $\mathbf{d}(m)=\text{diag}(\mathbf{\Lambda}(m))$ is an $N\times 1$ column vector. Since there is an inherent scaling ambiguity in calculating $\mathbf{b}_i(\omega_k)$ form (3.35), without loss of generality we can assume $||\mathbf{b}_i(\omega_k)||_2^2=1$. Also the constraint set $\Omega\subset\mathbb{C}^{J^2\times 1}$ is defined as:

$$\Omega=\{\text{vec}\{\mathbf{\Phi}\}|\mathbf{\Phi}=\boldsymbol{\nu}\boldsymbol{\nu}^{\dagger},\ \boldsymbol{\nu}\in\mathbb{C}^{J\times 1},\ ||\boldsymbol{\nu}||_2^2=1\}. \tag{3.39}$$

Following the ALS procedure, we can first minimize (3.38) with respect to $\mathbf{g}_i(\omega_k)$ conditioned on $\hat{\mathbf{d}}(m)$, the previously estimated values of $\mathbf{d}(m)$. To do this we form the matrices $\mathbf{T}(\omega_k)=[\hat{\mathbf{p}}(\omega_k,0),\ldots,\hat{\mathbf{p}}(\omega_k,M-1)]$ and $\mathbf{F}=[\hat{\mathbf{d}}(0),\ldots,\hat{\mathbf{d}}(M-1)]$ and we write equation (3.35) as:

$$\min_{\mathbf{g}_i(\omega_k)\in\Omega} \sum_{k=0}^{K-1} W_k||\mathbf{T}(\omega_k)-\mathbf{G}(\omega_k)\mathbf{F}||_F^2. \tag{3.40}$$

Notice that (3.40) is a constrained least-squares problem. One simple, although approximate, way to minimize (3.40) is to first find the unconstrained least-squares minimizer of

(3.40) by setting

$$\tilde{\mathbf{G}}(\omega_k) = \mathbf{T}(\omega_k)\mathbf{F}^{+}. \tag{3.41}$$

We then project each column of $\tilde{\mathbf{G}}(\omega_k)$ onto Ω; i.e.,

$$\hat{\mathbf{g}}_i(\omega_k) = \text{proj}_{\,\Omega}[\tilde{\mathbf{g}}_i(\omega_k)] \tag{3.42}$$

where $\tilde{\mathbf{g}}_i(\omega_k)$ is the i_{th} column of $\tilde{\mathbf{G}}(\omega_k)$.

We now discuss a convenient method of performing the projection operation. The projection operation can be effected by the following minimization:

$$\min_{\mathbf{g}_i(\omega_k)\in\Omega} ||\tilde{\mathbf{g}}_i(\omega_k) - \mathbf{g}_i(\omega_k)||_2^2. \tag{3.43}$$

Since $\mathbf{g}_i(\omega_k) = \text{vec}\{\mathbf{b}_i(\omega_k)\mathbf{b}_i^{\dagger}(\omega_k)\}$, by defining $\mathbf{Y}_i(\omega_k) = \text{mat}\{\tilde{\mathbf{g}}_i(\omega_k)\}$ we can write the above equation as:

$$\begin{aligned} &\min_{||\mathbf{b}_i(\omega_k)||_2=1} ||\mathbf{Y}_i(\omega_k) - \mathbf{b}_i(\omega_k)\mathbf{b}_i^{\dagger}(\omega_k)||_F^2 \equiv \\ &\min_{||\mathbf{b}_i(\omega_k)||_2=1} (\mathbf{b}_i^{\dagger}(\omega_k)\mathbf{b}_i(\omega_k))^2 + Tr\big(\mathbf{Y}_i^{\dagger}(\omega_k)\mathbf{Y}_i(\omega_k)\big) - 2\mathbf{b}_i^{\dagger}(\omega_k)\mathbf{Y}_i(\omega_k)\mathbf{b}_i(\omega_k) \equiv \\ &\min_{||\mathbf{b}_i(\omega_k)||_2=1} C - 2\mathbf{b}_i^{\dagger}(\omega_k)\mathbf{Y}_i(\omega_k)\mathbf{b}_i(\omega_k) \end{aligned} \tag{3.44}$$

where $C = 1 + Tr\big(\mathbf{Y}_i^{\dagger}(\omega_k)\mathbf{Y}_i(\omega_k)\big)$ is a constant term. The above minimization can be done easily by choosing $\hat{\mathbf{b}}_i(\omega_k)$, the estimated i_{th} column of $\mathbf{B}(\omega_k)$, to be the dominant eigenvector of $\mathbf{Y}_i(\omega_k)$. To find the dominant eigenvector of a matrix we can use the power iteration method described in (Golub and VanLoan, 1996). Since an initial estimate of $\mathbf{b}_i(\omega_k)$ is available (as is explained later), $\mathbf{Y}_i(\omega_k)$ is nearly a rank one matrix. Hence, the ratio of the largest eigenvalue of $\mathbf{Y}_i(\omega_k)$ to the second-largest (this ratio determines the convergence of the power method), is large. Hence, we need to apply only few iterations of the power method to minimize (3.44) [1].

To minimize (3.35) with respect to $\mathbf{d}(m)$ conditioned on the previous estimate of $\mathbf{B}(\omega_k)$ we concatenate the vectors $\hat{\mathbf{p}}(\omega_k, m)$ and the matrices

[1] In our simulations we use only one power iteration per ALS iteration. Increasing the number of iterations beyond one did not noticeably improve the convergence nor the performance of the algorithm.

$\hat{\mathbf{G}}(\omega_k) = [\text{vec}\{\hat{\mathbf{b}}_1(\omega_k)\hat{\mathbf{b}}_1^\dagger(\omega_k)\}, \ldots, \text{vec}\{\hat{\mathbf{b}}_N(\omega_k)\hat{\mathbf{b}}_N^\dagger(\omega_k)\}]$ for all values of $k = 0, ..., K-1$. For each m we have:

$$\min_{\mathbf{d}(m)} \left\| \begin{bmatrix} \sqrt{W_0}\hat{\mathbf{p}}(\omega_0, m) \\ \cdot \\ \cdot \\ \sqrt{W_{K-1}}\hat{\mathbf{p}}(\omega_{K-1}, m) \end{bmatrix} - \begin{bmatrix} \sqrt{W_0}\hat{\mathbf{G}}(\omega_0) \\ \cdot \\ \cdot \\ \sqrt{W_{K-1}}\hat{\mathbf{G}}(\omega_{K-1}) \end{bmatrix} \mathbf{d}(m) \right\|_2^2 \tag{3.45}$$

Minimizing (3.45) with respect to $\mathbf{d}(m)$ we get:

$$\hat{\mathbf{d}}(m) = \begin{bmatrix} \sqrt{W_0}\hat{\mathbf{G}}(\omega_0) \\ \cdot \\ \cdot \\ \sqrt{W_{K-1}}\hat{\mathbf{G}}(\omega_{K-1}) \end{bmatrix}^+ \begin{bmatrix} \sqrt{W_0}\hat{\mathbf{p}}(\omega_0, m) \\ \cdot \\ \cdot \\ \sqrt{W_{K-1}}\hat{\mathbf{p}}(\omega_{K-1}, m) \end{bmatrix} \qquad m = 0, ..., M-1. \tag{3.46}$$

Using equation (3.41), (3.42) and (3.46) we can repeatedly update the values of $\mathbf{d}(m)$ and $\mathbf{G}(\omega_k)$ until convergence is achieved.

As mentioned previously, to avoid being trapped in local minima, we need to properly initialize the algorithm. One simple way of doing this is to use the following closed form algorithm for joint diagonalization of two matrices based on the previous discussion in Chapter 2:

3.4.2 Initialization

To initialize the algorithm we can select two matrices $\hat{\mathbf{P}}(\omega_k, m_1)$, $\hat{\mathbf{P}}(\omega_k, m_2)$, $m_1 \neq m_2$. We then choose the initial estimate for $\hat{\mathbf{B}}(\omega_k)$ to be the matrix consisting of the N dominant generalized eigenvectors of the matrix couple $(\hat{\mathbf{P}}(\omega_k, m_1), \hat{\mathbf{P}}(\omega_k, m_2))$. Although no optimum selection for m_1 and m_2 can be given at this stage, in our simulations we choose $\hat{\mathbf{P}}(\omega_k, m_1)$ and $\hat{\mathbf{P}}(\omega_k, m_2)$ such that their non-zero generalized eigenvalues are not all repeated.

Summary of Step I of the Algorithm for Blind Identification

```
1. Estimate the observed signals' cross spectral density matrices, $\hat{\mathbf{P}}_x(\omega_k, m)$, based on (3.36) and set $\mathbf{T}(\omega_k) = [\hat{\mathbf{p}}_x(\omega_k, 0), ..., \hat{\mathbf{p}}_x(\omega_k, M-1)]$ where $\hat{\mathbf{p}}(\omega_k) = \text{vec}\{\hat{\mathbf{P}}_x(\omega_k, m)\}$

2. Set $\hat{\mathbf{B}}^0(\omega_k)$, the initial value for the $\mathbf{B}(\omega_k)$, based on the method described in Section 3.4.2

3. for $\nu = 0$ to Max_itr

    • Calculate $\hat{\mathbf{d}}^{\nu}(m) \quad m = 0, ..., M-1$ from (3.46)
    • Set $\mathbf{F}^{\nu} = [\hat{\mathbf{d}}^{\nu}(0), ..., \hat{\mathbf{d}}^{\nu}(M-1)]$
    • for $k = 0$ to $K-1$
        – $\hat{\mathbf{G}}^{\nu}(\omega_k) = [\text{vec}\{\hat{\mathbf{b}}_1^{\nu}(\omega_k)\hat{\mathbf{b}}_1^{\dagger\nu}(\omega_k)\}, ..., \text{vec}\{\hat{\mathbf{b}}_N^{\nu}(\omega_k)\hat{\mathbf{b}}_N^{\dagger\nu}(\omega_k)\}]$
        – $\tilde{\mathbf{G}}^{\nu}(\omega_k) = \mathbf{T}(\omega_k)(\mathbf{F}^{\nu})^{+}$
        – for $i = 1$ to $N$
            * $\mathbf{Y} = \text{mat}\{\tilde{\mathbf{g}}_i^{\nu}(\omega_k)\}$
            * $\mathbf{q} = \mathbf{Y}\mathbf{b}_i^{\nu}(\omega_k)$
            * $\hat{\mathbf{b}}_i^{\nu+1}(\omega_k) = \frac{\mathbf{q}}{||\mathbf{q}||_2}$
        – end
    • end

4. Calculate the cost value $C^{\nu} = \sum_{k=0}^{K-1} W_k ||\mathbf{T}(\omega_k) - \hat{\mathbf{G}}^{\nu}(\omega_k)\mathbf{F}^{\nu}||_F^2$

5. if $\frac{|C^{\nu} - C^{\nu-1}|}{C^{\nu}} < \epsilon$ where $0 < \epsilon \ll 1$ then stop

6. end
```

3.4.3 Step II

Step II removes the frequency dependent scaling ambiguity by exploiting A5 via Theorem 3. Let the frequency domain quantity $\mathbf{b}_i(\omega_k)$ denote the i_{th} column of $\mathbf{B}(\omega_k)$ obtained in

Step I of the algorithm. Without loss of generality we can assume the permutation matrix $\mathbf{\Pi}$ in (3.34) is an identity matrix and we can write:

$$\mathbf{b}_i(\omega_k) = \mathbf{h}_i(\omega_k) d_{ii}(\omega_k) \tag{3.47}$$

where $\mathbf{h}_i(\omega_k)$ is the i_{th} column of $\mathbf{H}(\omega_k)$ and $d_{ii}(\omega_k)$ is the i_{th} diagonal element of $\mathbf{D}(\omega_k)$. Assumption A5 states the elements of $\mathbf{h}_i(z)$ are coprime; i.e., they do not share common zeros. In the time domain this corresponds to the matrix

$$\mathcal{S}_Q(\mathbf{h}_i) = \begin{pmatrix} \mathbf{h}_i(0) & . & . & . & \mathbf{h}_i(L_i) & \mathbf{0} & . & . & \mathbf{0} \\ & & . & & & & . & & \\ & & & . & & & & . & \\ \mathbf{0} & . & . & \mathbf{0} & \mathbf{h}_i(0) & . & . & . & \mathbf{h}_i(L_i) \end{pmatrix} \in \mathbb{R}^{J(Q+1)\times(L_i+Q+1)} \tag{3.48}$$

having full column rank for $Q \geq L_i - 1$ where L_i is the maximum order of the elements of $\mathbf{h}_i(t)$ (Serpedin and Giannakis, 1999),(Tong *et al.*, 1995). To remove the frequency dependent scaling ambiguities, $d_{ii}(\omega_k)$, we use the following steps. In the time domain (3.47) can be expressed as the circular convolution of $d_{ii}(t)$, the K-point IDFT of $d_{ii}(\omega_k)$, with $\mathbf{h}_i(t)$, the K-point IDFT of $\mathbf{h}_i(\omega_k)$. Assuming that $K > L_i$, (3.47) can therefore be written as:

$$(\mathbf{b}_i(0), ..., \mathbf{b}_i(K-1)) = (\mathbf{h}_i(0), ..., \mathbf{h}_i(L_i), \mathbf{0}_{J\times(K-L_i-1)})\mathbf{D}_C^i \tag{3.49}$$

where

$$\mathbf{D}_C^i = \begin{pmatrix} d_i(0) & d_i(1) & . & . & d_i(K-1) \\ d_i(K-1) & d_i(0) & . & . & d_i(K-2) \\ . & . & . & . & . \\ d_i(1) & . & . & d_i(K-1) & d_i(0) \end{pmatrix} \in \mathbb{R}^{K\times K} \tag{3.50}$$

is a circulant matrix. To remove the scaling ambiguities $d_{ii}(\omega_k)$ we need to find a circulant matrix $\mathbf{\Phi}_C^i$ given as:

$$\mathbf{\Phi}_C^i = \begin{pmatrix} \phi_i(0) & \phi_i(1) & . & . & \phi_i(K-1) \\ \phi_i(K-1) & \phi_i(0) & . & . & \phi_i(K-2) \\ . & . & . & . & . \\ \phi_i(1) & . & . & \phi_i(K-1) & \phi_i(0) \end{pmatrix} \in \mathbb{R}^{K\times K} \tag{3.51}$$

such that

$$\mathbf{D}_C^i \mathbf{\Phi}_C^i = \alpha \mathbf{I} \tag{3.52}$$

where α is a constant scalar. Having found such a $\mathbf{\Phi}_C^i$ we can then calculate $\hat{\mathbf{h}}_i(t)$, the estimated i_{th} column of $\mathbf{H}(t)$, by setting

$$\hat{\mathbf{h}}_i(t) = \mathbf{b}_i(t) \circledast \phi_i(t), \qquad t = 0, ..., K-1 \tag{3.53}$$

where $\circledast$ represents the circular convolution operation. Notice that in general $\mathbf{D}_C^i$ is unknown so we cannot find $\mathbf{\Phi}_C^i$ from (3.52). To calculate $\mathbf{\Phi}_C^i$ we exploit the full column rank property of the matrix $\mathcal{S}_Q(\mathbf{h}_i)$ for $Q \geq L_i - 1$. For $K = L_i + Q + 1$, from equations (3.48) and (3.49) we can write:

$$\mathbf{B}_C^i = [\mathcal{S}_{Q-1}(\mathbf{h}_i) \;\; \mathbf{0}_{JQ\times 1}]\mathbf{D}_C^i \tag{3.54}$$

where

$$\mathbf{B}_C^i = \begin{pmatrix} \mathbf{b}_i(0) & \mathbf{b}_i(1) & . & . & \mathbf{b}_i(K-1) \\ \mathbf{b}_i(K-1) & \mathbf{b}_i(0) & . & . & \mathbf{b}_i(K-2) \\ . & . & . & . & . \\ \mathbf{b}_i(K-Q+1) & . & . & . & \mathbf{b}_i(K-Q) \end{pmatrix} \in \mathbb{R}^{JQ\times K}. \tag{3.55}$$

Multiplying both sides of equation (3.54) by $\mathbf{\Phi}_C^i$ yields:

$$\mathbf{B}_C^i \mathbf{\Phi}_C^i = [\mathcal{S}_{Q-1}(\mathbf{h}_i) \;\; \mathbf{0}_{JQ\times 1}]\mathbf{D}_C^i \mathbf{\Phi}_C^i. \tag{3.56}$$

Define $\hat{\mathbf{H}}_C^i = \mathbf{B}_C^i \mathbf{\Phi}_C^i$. We now show for $Q \geq L_i$, if we find a non-zero $\mathbf{\Phi}_C^i$ that makes the last column of $\hat{\mathbf{H}}_C^i$ equal to zero, then it also satisfies equation (3.52); i.e., $\hat{\mathbf{H}}_C^i$ becomes a scaled version of $\mathbf{B}_C^i$. Assign $\mathbf{\Lambda}_C^i = \mathbf{D}_C^{iT} \mathbf{\Phi}_C^i$, which is also a circulant matrix written as

$$\mathbf{\Lambda}_C^i = \begin{pmatrix} \lambda_i(0) & \lambda_i(1) & . & . & \lambda_i(K-1) \\ \lambda_i(K-1) & \lambda_i(0) & . & . & \lambda_i(K-2) \\ . & . & . & . & . \\ \lambda_i(1) & . & . & \lambda_i(K-1) & \lambda_i(0) \end{pmatrix} \in \mathbb{R}^{K\times K}. \tag{3.57}$$

Then the last column of $\hat{\mathbf{H}}_C^i$ is equal to:

$$\begin{pmatrix} \hat{\mathbf{h}}_i(K-1) \\ \cdot \\ \cdot \\ \hat{\mathbf{h}}_i(K-Q) \end{pmatrix} = \mathcal{S}_{Q-1}(\mathbf{h}_i) \begin{pmatrix} \lambda_i(K-1) \\ \cdot \\ \cdot \\ \lambda_i(1) \end{pmatrix}. \tag{3.58}$$

For $Q \geq L_i$, $\mathcal{S}_{Q-1}(\mathbf{h}_i)$ has full column rank. Therefore, if the vector on the LHS of (3.58) is to have all zero elements, then $\lambda(K-1), ..., \lambda(1)$ must also be all zeros; i.e., the matrix $\mathbf{\Lambda}_C^i$ is diagonal with all diagonal elements equal to $\lambda(0)$, and the proof is complete. Following what was said above we need to choose $\mathbf{\Phi}_C^i$ such that the elements of last column of $\hat{\mathbf{H}}_C^i$ become all zeros. In other words, we find the vector $\phi_i = (\phi_i(K-1), .., \phi_i(0))^T$, the last column of the circulant matrix $\mathbf{\Phi}_C^i$, such that

$$\mathbf{B}_C^i \phi_i = \mathbf{0}. \tag{3.59}$$

To find ϕ_i we minimize the quantity $||\mathbf{B}_C^i \phi_i||_2^2$, or equivalently:

$$\min_{||\phi_i||_2=1} \phi_i^T \mathbf{B}_C^{i^T} \mathbf{B}_C^i \phi_i. \tag{3.60}$$

The solution is given by choosing ϕ_i to be the eigenvector of $\mathbf{B}_C^{i^T} \mathbf{B}_C^i$ corresponding to it's minimum eigenvalue. The last step is to compute $\hat{\mathbf{h}}_i(t)$ from (3.53).

Summary of Step II of the Algorithm for Blind Identification

1. for i=1 to N
   - Choose $K \geq 2L_i + 1$
   - Calculate $Q = K - L_i - 1$
   - Calculate $\mathbf{b}_i(t) = IDFT\{\mathbf{b}_i(\omega_k)\}$ $k = 0, \ldots, K-1$ $t = 0, \ldots, K-1$
   - Form the matrix $\mathbf{B}_C^i$ from (3.55)
   - Set ϕ_i to be the minimum eigenvector of $\mathbf{B}_C^{i^T} \mathbf{B}_C^i$

```
• Calculate $\phi_i(\omega_k) = DFT\{\phi_i(0), ..., \phi_i(K-1)\}$
• Calculate $\hat{\mathbf{h}}_i(t) = IDFT\{\mathbf{b}_i(\omega_k)\phi_i(\omega_k)\}$  $k = 0, \ldots, K-1$  $t = 0, \ldots, K-1$
• end
```

3.5 Simulation Results

3.5.1 Example I, White Sources

For the first simulation, the sources are two independent white Gaussian signals, multiplied by slowly varying sine and cosine signals to create the desired quasi–stationary effect. The purpose of this example is to show that the algorithm is capable of identifying the channel even when the sources are white and Gaussian. Note that none of the previous SOS and HOS methods can identify the channel in this case because they require the sources to be colored in the case of SOS methods and non-Gaussian in the case of HOS methods. We choose the channel to be a 3×2 system whose impulse response $\mathbf{H}(t)$ is given in Table 3.1. The epoch size is kept constant at 500 and the data length was varied between 10000 and 50000 samples, corresponding to M, the number of epochs, ranging between 20 to 100 epochs. White Gaussian noise was added to the output of the system at a level corresponding to the desired value of averaged SNR over all epochs[2]. At each epoch, 128–point FFTs, applied

t	0	1	2	3	4	5	6	7
$H_{11}(t)$	-0.528	-0.153	0.631	0.942	-0.221	-0.701	0.274	-0.681
$H_{12}(t)$	0.696	1.952	0.234	-0.938	0.856	1.347	0.341	0.213
$H_{21}(t)$	0.963	-0.927	-0.085	0.322	-0.963	0.049	-0.614	0.000
$H_{22}(t)$	0.675	0.056	-0.143	0.180	1.054	0.230	1.704	0.704
$H_{31}(t)$	0.719	0.538	-1.070	-1.351	0.105	-1.493	0.224	0.144
$H_{32}(t)$	0.774	0.047	-0.147	-0.381	0.287	-0.047	0.649	0.147

Table 3.1: Impulse Response of the MIMO system for Examples I & II & III

[2] The power of the noise was kept constant at all epochs.

to time segments overlapping by 50%, weighted by Hanning windows were used to estimate the cross-spectral density matrices. At each epoch, only $K = 16$ cross-spectral density matrices[3], evaluated at uniformly spaced frequency samples, were calculated as input to the algorithm. We also choose the W_k in (3.35) to be:

$$W_k = \frac{1}{\sum_{m=0}^{M-1} ||\mathbf{P}_x(\omega_k, m)||_F^2}. \tag{3.61}$$

Notice that by this choice of W_k we put more emphasis on those frequency bins where average norm of the cross spectral density matrices is small. For white sources this corresponds to those frequency bins where channel parameters have small values and as a result are harder to estimate. Compared to the case when all W_k are set to ones, our simulations show that this choice of W_k improves the overall estimation error. To measure the estimation error, since a scaling ambiguity exists in the final results, we use the following measure for mean–squared error (mse) based on a method suggested in (Morgan *et al.*, 1998) for evaluating the estimated impulse responses[4]

$$MSE = 1 - \frac{1}{JNM_c} \sum_{k=1}^{M_c} \sum_{j=1}^{J} \sum_{i=1}^{N} \left(\frac{\mathbf{h}_{ij}^T \hat{\mathbf{h}}_{ij}^k}{||\mathbf{h}_{ij}|| ||\hat{\mathbf{h}}_{ij}^k||} \right)^2 \tag{3.62}$$

where $\mathbf{h}_{ij} = (h_{ij}(0), ..., h_{ij}(L_{ij}))^T$ is the true ij_{th} impulse response of the channel and $\hat{\mathbf{h}}_{ij}^k$ is the estimated response at Monte Carlo run k. The quantity M_c is the total number of Monte Carlo runs.

Table 3.2 shows the mse, calculated from (3.62), for different SNR's and varying M, using $M_c = 50$ Monte-Carlo runs. As can be seen from Table 3.2, by increasing the number of epochs, which corresponds to increasing the data length, the mse decreases.

To get a visual impression of the results in Table 3.2, Figures 3.1 and 3.2 illustrate the corresponding time domain impulse responses and frequency domain responses of the estimated and true channel for SNR=30dB, $M = 60$ epochs, and a data length of 10000 samples. In all of the simulations the algorithm converges between 7 and 17 iterations.

[3] This value satisfies the bound $K \geq 2L + 1$, where in this case $L = 7$.

[4] Also refer to (Manton, 2001) where the author proves that the mathematically correct way of measuring errors when scale ambiguity is present is to use a distance function on the complex projective space. In fact, (3.62) is one of a number of well known distance functions on the complex projective space.

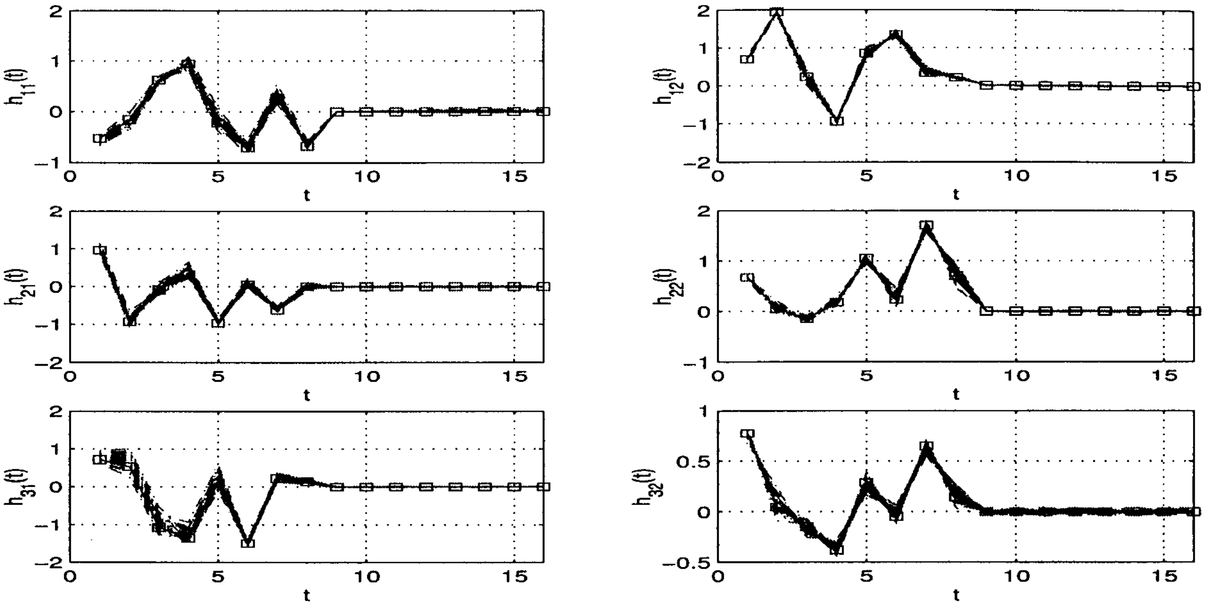

Figure 3.1: 50 superimposed independently estimated time-domain responses $\hat{\mathbf{H}}(t)$ shown as dot-dashed lines, along with the true $\mathbf{H}(t)$, shown by the solid lines with squares at the true data points. The horizontal axis is the time index. $M = 60$, SNR=30dB, and $K = 16$.

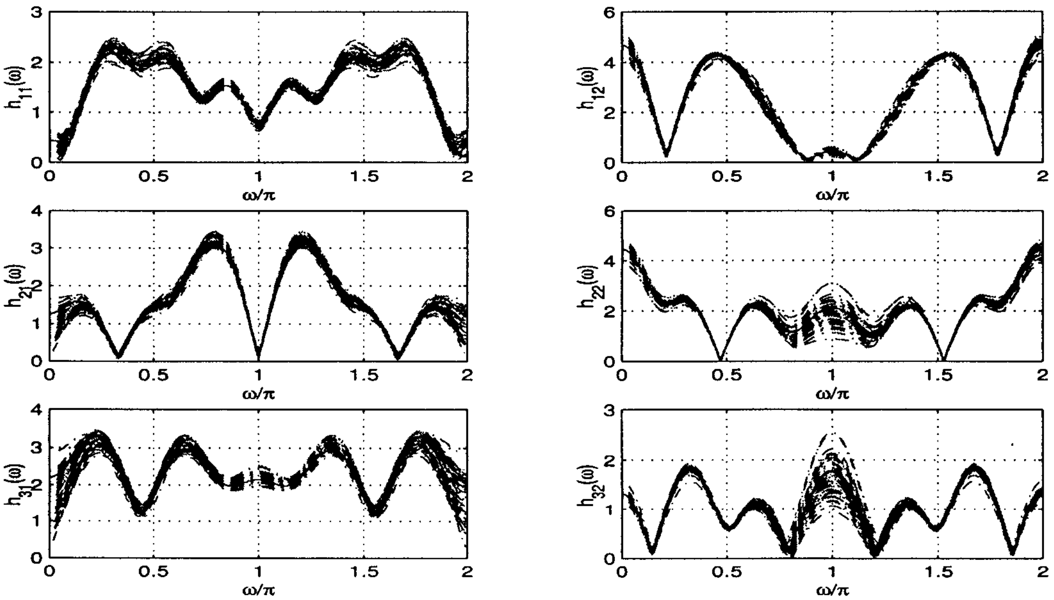

Figure 3.2: Same as Figure 3.1, except the impulse responses are shown in the frequency domain.

M	20	40	60	100
SNR=30 dB	0.0423	0.0212	0.0148	0.0068
SNR=15 dB	0.0531	0.0282	0.0180	0.0114
SNR=10 dB	0.0714	0.0324	0.0235	0.0153

Table 3.2: Example I, showing mse for different SNRs and varying M using $M_c = 50$ Monte Carlo runs.

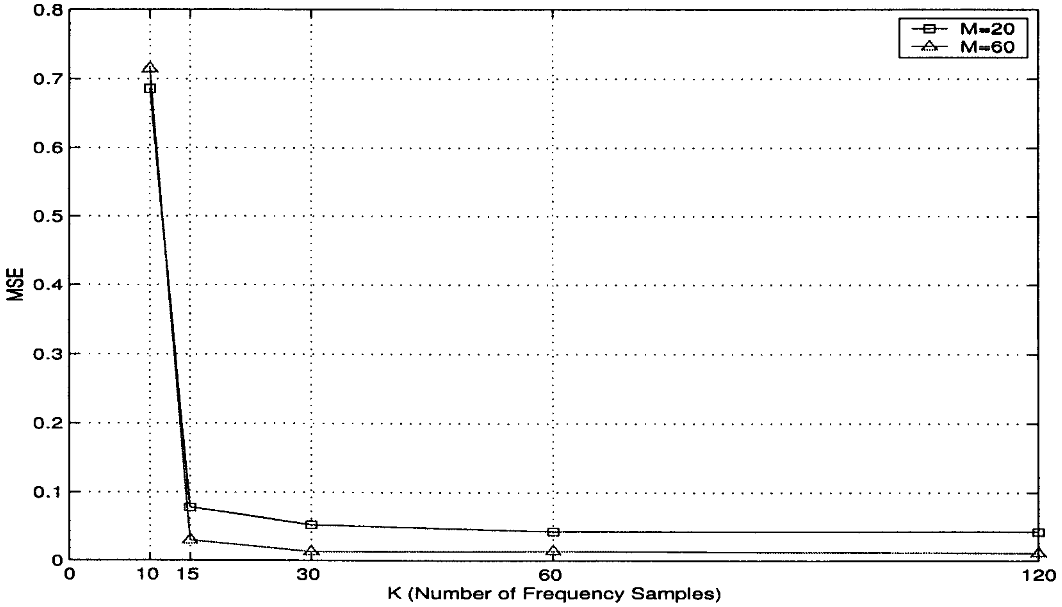

Figure 3.3: mse versus K, number of frequency samples, for M=20 and M=60 and for SNR=10dB using $M_c = 50$ Monte Carlo.

Further, in Figure 3.3, we show the results for varying K. As can be seen from the figure, for $K = 15$ we have a sudden drop in the estimation error. These results match our theoretical bound, derived in Theorem 3, which states that $K \geq 2L + 1$ frequency samples are required for identifiability of $\mathbf{H}(t)$. Since $L = 7$, in theory the number of frequency samples should be $K \geq 14 + 1 = 15$.

3.5.2 Example II, Colored Sources

In this example we show the performance of the algorithm when the sources are colored signals. For this simulation, to create the colored sources we pass the source signals in the previous example through a first order AR filter with the pole located at $z = 0.8$. Notice that contrary to some second-order statistics methods which require the sources to have distinct spectral shapes (Hua and Tugnait, 2000), here the shape of spectrum of the sources can be identical. The sources are mixed through the same channel as example I and we apply the algorithm using the same parameters used in the previous example. The results are shown in Table 3.3. It can be seen the results are close to what were obtained for the white source case.

M	20	40	60	100
SNR=30 dB	0.0566	0.0231	0.0154	0.0124
SNR=15 dB	0.0695	0.0298	0.0236	0.0160
SNR=10 dB	0.0983	0.0774	0.0511	0.0358

Table 3.3: Example II, showing mse for different SNR's and varying M using $M_c = 50$ Monte Carlo runs.

3.5.3 Example III, Speech Signals

The purpose of this next example is to show the performance of the algorithm when the sources are speech signals. Note that for speech signals not only the power but the whole spectrum of the signal changes with time. This violates (3.19) which we require for colored signals; therefore, we can expect degradation in performance. We use the same channel given in example 1; however, for the sources, we use two female speech sequences sampled at 8.0 KHz with a total duration of 2.0 seconds. White Gaussian noise is added to the output of the system commensurate with the specified value of SNR. In a manner similar to Example 1, we use 128–point FFTs, with the CPSD matrices being computed at only 16 FFT points for each epoch. Table 3.4 shows the computed mse for $M = 50$ epochs for varying values of SNR. As can be seen from the results, the performance has been degraded somewhat compared to the previous example, especially at low values of SNR. However, for high signal to noise ratios it can be observed that the channel can be identified to within a reasonable error. One explanation for this is that in general the speech signals do not satisfy the spectral condition given in (3.19); i.e., the shape of spectrum of speech signals may change over the observation time. Due to this modelling error, we can expect more errors in estimating the channel, using the criterion given in (3.35), when the inputs to the channel are speech signals. We can make the criterion (3.35) more general, so it is also applicable to speech signals, through representing the cross power spectral density matrices of the sources by $\boldsymbol{\Lambda}(\omega_k, m)$ rather then $\boldsymbol{\Lambda}(m)$. The down-side of doing this is that the new criterion will then be prone to the permutation problem; i.e., at each frequency bin we may get a different permutation. Nevertheless since the spectral envelope of speech signals are correlated across frequency spectrum, one can exploit this property to remove the arbitrarily frequency dependent permutations. The details of this new approach for the case when the input to the mixing system are speech (or in general audio) signals are given in the next chapter. Note that in current example no post-processing for removing permutation errors has been done.

SNR (dB)	30	15	10
	0.1291	0.4478	0.5195

Table 3.4: Example II, showing the estimation mse for varying SNR, for $M_c = 50$ Monte Carlo runs in the case of speech sources, with $M = 50$.

3.5.4 Example IV

In this example we compare the performance of our method with other existing MIMO channel identification approaches. Since most of the current methods for MIMO blind identification assume stationary sources, a direct comparison with our method, which explicitly exploits the non-stationarity of the sources, is not possible. The closest method to the proposed approach is the one recently proposed in (Chen and Petropulu, 2001). Their method is also a frequency domain approach with the major difference that they use higher-oder statistics of the observed signal to identify the channel. Note that the method in (Chen and Petropulu, 2001) can only be applied to non-Gaussian signals with non-symmetrical pdfs while the proposed method can be applied to signals with arbitrary pdfs as long as they satisfy the non-stationarity assumption. Also the proposed method can be directly extended to colored signals under assumption (3.19) while the method in (Chen and Petropulu, 2001) is restricted to white signals. To compare our results we use the same channel, data length, signal to noise ratio and number of FFT points used in example (1) of (Chen and Petropulu, 2001). The only exceptions are the sources; in (Chen and Petropulu, 2001) the sources are non-Gaussian stationary signals while for the proposed method we use non-stationary white Gaussian sources. Also to measure the estimation error we use the same performance measure given by equation (53) in (Chen and Petropulu, 2001). The comparative results are shown in Figure (3.4), where we have used the mean square error data in Table 1 of (Chen and Petropulu, 2001) to compare with our results. As can be seen from the figure, the performance of the two methods are very close. For shorter data lengths the method in (Chen and Petropulu, 2001) has a slightly better performance over the proposed method while for a higher number of data samples (more epochs), specially at low SNR, the proposed method has the advantage over the method in (Chen and Petropulu, 2001).

3.6 Summary

In this chapter we have derived sufficient conditions for identifiability of a MIMO system, driven by white quasi-stationary sources, in the frequency domain using only second-order statistics of the observed signals. We also showed that the same results can be directly extended to quasi stationary colored sources when only the power of the signal is slowly varying with time. We also proposed a two stage algorithm. The first stage estimates the channel parameters up to a constant permutation and frequency dependent scaling factor, based on a alternating least-squares method, while the second stage removes the frequency dependent scaling ambiguity using a closed form algorithm. The results of applying the new

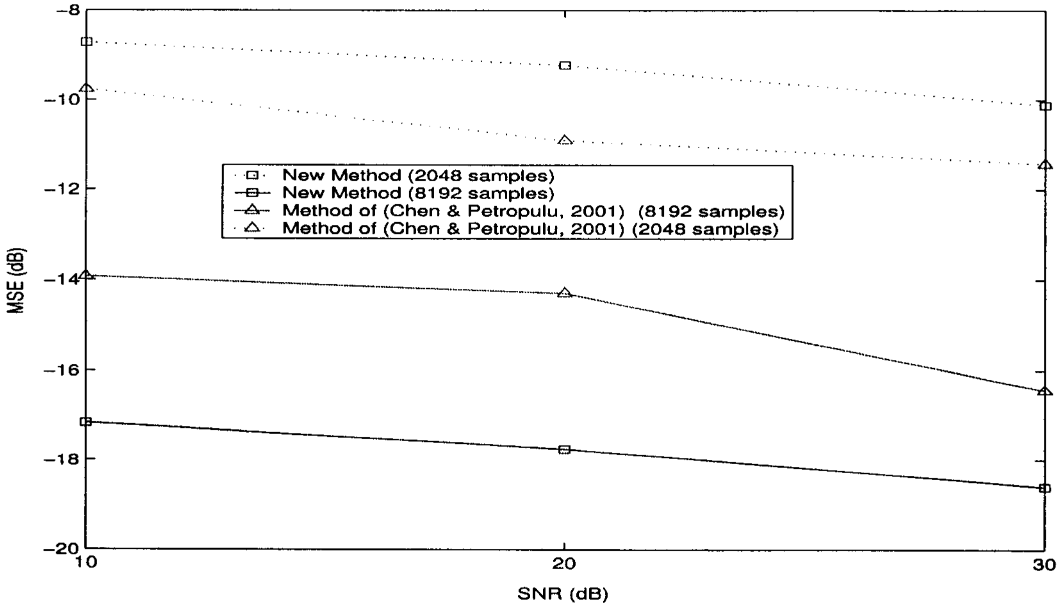

Figure 3.4: Comparison of the new proposed algorithm with the method in (Chen and Petropulu, 2001).

algorithm to white and colored sources under the stated assumptions verifies the identifiability conditions, as well as the performance of the algorithm. Application of the proposed algorithm to speech signals was also demonstrated.

Chapter 4

Real Room Blind Source Separation Problem

In this chapter we propose a new frequency domain algorithm for blind source separation of audio signals, mixed in a reverberant environment. The proposed algorithm is an extension of the MIMO identification procedure described in previous chapter. The first part of the algorithm uses joint diagonalization of the cross spectral density matrices of the output of the mixing system to identify the mixing system at each frequency bin up to a scaling and permutation ambiguity. The inverse of the estimated mixing system is then used to separate the sources. The second part of the algorithm uses a novel procedure to resolve the frequency dependent permutation problem by exploiting the inherent non-stationarity of the audio sources. Also the frequency dependent scaling ambiguity problem is partially resolved by means of a novel initialization procedure for the first step of the algorithm. We demonstrate the performance of the proposed algorithm using real room experiments, by blind separation of audio signals mixed in reverberant environments. In all of our experiments the algorithm demonstrates good separation performance and enhanced output audio quality. We also compare the proposed algorithm to the one in (Parra and Spence, 2000). The results show fast convergence and superior separation performance of the new algorithm.

4.1 Introduction

The MIMO blind identification method, discussed in the previous chapter, has direct application in blind source separation of convolved mixed sources, in that having the estimated mixing system we can use standard multichannel equalization techniques to recover the sources. As mentioned before an advantage of using the MIMO blind identification method, discussed in the previous section, is that we not only can separate the sources but we also can equalize them; i.e., we can can recover the sources up to a constant scaling and permutation ambiguity.

In this chapter we are interested in one particular application of blind source separation, which is blind separation of mixed audio signals in a reverberant environment. In theory this case can be modeled as a convolutive mixing blind source separation problem. Nevertheless, as has already been discussed in the introductory chapter of this thesis, when one actually uses a convolutive BSS algorithm in a real reverberant environment there are few challenges that need to be considered. For example the MIMO blind identification algorithm presented in the previous chapter cannot be used directly for blind separation of audio signals in a reverberant environment. One important reason, which is also discussed in the previous chapter, is that audio signals, although inherently non-stationary, do not necessarily satisfy the condition that the shape of the their spectrum stays the same with time, and only the scale of the spectrum changes. Another reason that we may not be able to use the algorithm in the previous chapter is that, we assume the mixing system has an FIR structure and the elements in each column of the mixing system do not share common zeros. We also assume that the order of the mixing system is known beforehand. For a real reverberant acoustic environment, none of these assumptions are practical. Note that acoustic impulse responses (AIR) are not FIR and even if we approximate them with FIR filters, their orders are not known beforehand.

Due to these problems, in this chapter we propose a new convolutive BSS algorithm that can be applied in a real reverberant acoustic environment. The proposed algorithm is a modified version of the algorithm presented in the previous chapter and resolves some of the problems mentioned above. Note that using the proposed algorithm, we can only separate the sources up to some filter ambiguity. Nevertheless the distortions caused by this filtering ambiguity, as shown by our listening tests, can be minimized using a novel initialization approach as discussed in this chapter.

Comparison to existing methods: So far, the results shown for most of the convolutive blind source separation methods, especially the HOS methods, are limited to computer simulations, using synthetically generated mixing channels with small orders, far from resembling a real mixing environment. Among these methods there are some that only consider two-input two-output (TITO) mixing systems (Lindgren and Broman, 1998)(Yellin and Weinstein, 1994)(Weinstein *et al.*, 1993) and some that assume mixing systems with the same number of inputs and outputs (Lee *et al.*, 1997). There are few methods that consider blind source separation of audio signals in real room situations (Parra and Spence, 2000)(Schobben and Sommen, 1998). Nevertheless the real performance of these algorithms seems to be poor when they are operating in a highly reverberant environment.

In this chapter we show the results of our new BSS algorithm for real room experiments with long reverberation times. The proposed method is not limited to the mixing systems with fixed dimensions nor to ones with same number of outputs and inputs. The only requirement is that the number of outputs for the mixing system should be greater than or equal to the number of inputs.

The method in (Parra and Spence, 2000) is rare among methods presented so far in that it considers blind separation in a real room environment, with some promising results. The method exploits the non-stationarity of the observed signals in the frequency domain and proposes a constrained least-squares criterion at each frequency bin which then is minimized using a steepest descent approach. One of the assumptions made in (Parra and Spence, 2000) is that the length of the un-mixing filters is finite and is much smaller than the number of frequency bins. Although not explicitly proven, it's been shown in (Parra and Spence, 2000) that the constraint on the filter size helps to resolve the frequency domain permutation problem. The main limitation of using this approach to solve the permutation problem has been discussed in (Ikram and Morgan, 2000)and later on in (Ikram and Morgan, 2001). Also in (Araki *et al.*, 2001), through experimental results, the authors show for the long room reverberation case, as the length of the un-mixing filters increases, the separation performance decreases and they conclude that the length constraint is not efficient in a long reverberant environment.

The algorithm proposed in this chapter does not experience the problems mentioned above, mainly because of the different approach that has been used to solve the permutation problem. In this method, rather than constraining the length of the un-mixing filters, we exploit the non-stationarity of the input signals to resolve permutation errors. In the previous chapter we proved theoretically that for white non-stationary signals, mixed through a convolutive system, a frequency domain algorithm can be derived which has a uniform permutation across all frequency bins. Although the results do not hold for speech signals, nevertheless, as is shown in this chapter, by using the spectral correlation between the adjacent frequency bins the permutation problem can be resolved in new way through solving a discrete optimization problem for each pair of frequency bins (see also (Rahbar and Reilly, 2001a)). To prevent any catastrophic situation (as will be explained later in this chapter), that can happen as a result of a missed or wrongly adjusted permutation, we propose a diadic sorting scheme to achieve a uniform permutation across all frequency bins.

Similar to the algorithm for MIMO blind identification, here for the first step of the algorithm we use an alternating least-squares (ALS) method to optimize a criterion used for joint diagonalization of the cross-spectral density matrices (CPSD)estimated at different time epochs. The advantage of the new algorithm compared to the one presented in (Parra and Spence, 2000) is its fast convergence and significantly better performance as is demonstrated in our real room experimental results.

The organization of this chapter is as follows: The problem formulation including the set of required assumptions is presented in Section 4.2. In Section 4.3 we present a frequency domain algorithm for convolutive blind source separation. The permutation problem, including the proposed solution, is discussed in Section 4.4. Simulation results for synthetic convolutive mixing scenarios are described in section 4.5 and real room experimental results are described in Section 4.6. The first set of experiments are based on real room recordings done in a small office environment with moderate reverberation time. The second set of experiments are done in a conference room with a highly reverberant characteristic. In all these experiments the original sources (speech signals) are successfully recovered with good audio quality. We also compare our results with those obtained using the method in (Parra and Spence, 2000). Conclusions and final remarks are presented in Section 4.7.

4.2 Problem statement

We consider the following N-source J-sensor MIMO linear model for the received signal for the convolutive mixing problem[1]:

$$\mathbf{x}(t) = [\mathbf{H}(z)]\mathbf{s}(t) + \mathbf{n}(t) \quad t \in \mathbb{Z} \tag{4.1}$$

where $\mathbf{x}(t) = (x_1(t), \cdots, x_J(t))^T$ is the vector of observed signals, $\mathbf{s}(t) = (s_1(t), \cdots, s_N(t))^T$ is the vector of sources, $\mathbf{H}(z)$ is the $J \times N$ transfer function of mixing system and $\mathbf{n}(t) = (n_1(t), \cdots, n_J(t))^T$ is the additive noise vector. Notice that here, contrary to previous chapter, we assume $h_{ij}(z)$, the ij_{th} element of $\mathbf{H}(z)$, to be a rational function of z. For the special case where the $h_{ij}(z)$ are causal FIR filters, we have $\mathbf{H}(z) = \sum_{t=0}^{L} \mathbf{H}(t)z^{-t}$ where L is the highest polynomial degree of $h_{ij}(z)$ for all $i, j = 0, \ldots, N$. The objective of the blind source separation algorithm is to estimate the un-mixing filters $\mathbf{W}(z)$ from the observed signals $\mathbf{x}(t)$ such that

$$\mathbf{W}(z)\mathbf{H}(z) = \mathbf{\Pi}\mathbf{D}(z) \tag{4.2}$$

where $\mathbf{\Pi} \in \mathbb{R}^{N\times N}$ is a permutation matrix and $\mathbf{D}(z)$ is a diagonal matrix with diagonal elements which are rational functions of z. In the frequency domain this is equivalent to finding an $\mathbf{W}(\omega) \in \mathbb{C}^{J\times N}$ such that:

$$\mathbf{W}(\omega)\mathbf{H}(\omega) = \mathbf{\Pi}\mathbf{D}(\omega) \quad \forall\, \omega \in [0, \pi) \tag{4.3}$$

where $\mathbf{H}(\omega)$ is the corresponding DTFT for $\mathbf{H}(z)$. Notice that in (4.3), since we assume that the elements of the channel are real numbers, we only need to estimate $\mathbf{W}(\omega)$ over

[1] Here we use the notation $[\mathbf{H}(z)]\mathbf{s}(t)$ to denote the convolution between a system with z-transform $\mathbf{H}(z)$ and source vectors $\mathbf{s}(t)$.

half of the frequency range; i.e., $\omega \in [0, \pi)$. Equation (4.3) corresponds to the case when the outputs of the un-mixing filter, although separated, are a filtered version of the original sources.

4.2.1 Main Assumptions

A0: $J \geq N \geq 2$; i.e, we have at least as many sensors as sources and the number of sources are at least two.

A1: The sources $\mathbf{s}(t)$ are zero mean, second-order non-stationary signals. The cross–spectral density matrices of the sources $\mathbf{P}_s(\omega, m)$ are diagonal for all ω and m where ω denotes frequency and m is the epoch index.

A2: The mixing system is modelled by a causal system of the form $\mathbf{H}(z) = [\mathbf{h}_1(z), ..., \mathbf{h}_N(z)]$ and does not change over the entire observation interval.

A3; $\mathbf{H}(\omega_k)$, the DFT of $\mathbf{H}(z)$, has full column rank for all ω_k, $k = 0, \ldots, K-1$, $\omega_k = \frac{(2\pi)k}{K}$.

A5: The noise $\mathbf{n}(t)$ is zero mean, *iid* across sensors, with power σ^2. The noise is assumed independent of the sources.

Note that the assumptions used in this chapter are similar to the ones in Chapter 3, and so the identifiability conditions are similar to the ones given by Theorem 2 in that chapter. Note that here we use a less restrictive set of assumptions; e.g., we do not make any assumption on the structure of the channel and because of this we can only estimate the channel at each frequency bin up to some arbitrary scaling ambiguity. Also since we do not make any assumption on the temporal structure of the sources (they can be colored or white), the proposed identification algorithm is somewhat different from the one presented in the previous chapter and hence requires a second step to eliminate the frequency dependent permutation errors. In the next section we discuss a method whereby jointly diagonalizing the set of matrices $\mathbf{P}_x(\omega_k, m)$ $m = 0, \ldots, M-1$ we can estimate $\mathbf{H}(\omega_k)$ up to a permutation and scaling ambiguity. The separating matrix $\mathbf{W}(\omega_k)$ is then calculated by finding the pseudo inverse of $\hat{\mathbf{H}}(\omega_k)$, the estimated value of $\mathbf{H}(\omega_k)$. In the following sections we also discuss efficient methods for eliminating the frequency dependent permutation and alleviating the effect of frequency dependent scaling ambiguities.

4.3 The Algorithm

Based on the assumptions in the previous section, the cross-spectral density matrix of the observed signal at frequency ω_k and time epoch m can be written as

$$\mathbf{P}_x(\omega, m) = \mathbf{H}(\omega)\mathbf{P}_s(\omega, m)\mathbf{H}^{\dagger}(\omega) + \sigma^2\mathbf{I}, \tag{4.4}$$

where $\mathbf{P}_s(\omega, m)$ is a diagonal matrix which represents the cross-spectral density matrices of the sources at epoch m. To estimate the separating matrix $\mathbf{W}(\omega_k)$ we propose the following least-squares based joint diagonalization criterion for the case when a sample estimate of $\mathbf{P}_x(\omega_k, m)$ is available.

$$\min_{\mathbf{B}(\omega_k), \mathbf{\Lambda}(m)} \sum_{k=0}^{K-1}\sum_{m=0}^{M-1} ||\hat{\mathbf{P}}_x(\omega_k, m) - \mathbf{B}(\omega_k)\mathbf{\Lambda}(\omega_k, m)\mathbf{B}^{\dagger}(\omega_k)||_F^2, \tag{4.5}$$

where $\mathbf{B}(\omega)$ is an estimate of the mixing system $\mathbf{H}(\omega_k)$, $\hat{\mathbf{P}}_x(\omega_k, m)$ is a sample estimate of the observed signals cross spectral density matrix at frequency bin ω_k and time epoch m, $\mathbf{\Lambda}(\omega_k, m)$ is a diagonal matrix, representing the unknown cross-spectral density matrix of the sources at epoch m. Note that there are some differences between the above criterion and the one presented in the previous chapter. The main difference is that in the above criterion, the cross power spectral density matrices of sources are modeled by $\mathbf{\Lambda}(\omega_k, m)$ which is a function of both ω_k and m while for the criterion used in previous chapter, the cross power spectral density matrices of sources are modeled by $\mathbf{\Lambda}(m)$ which is only a function of m. As mentioned before, the use of $\mathbf{\Lambda}(m)$ is not a good model to represent audio signals.

In (Parra and Spence, 2000) a similar criterion has been used with the main difference being that their proposed criterion uses a backward model which directly estimates the separating matrix $\mathbf{W}(\omega_k)$. Using the criterion in (4.5) allows us to implement the ALS algorithm which is described later in this section. In (Parra and Spence, 2000), an additional FIR constraint on the de-mixing matrix is required to prevent arbitrary frequency dependent permutations. As shown in (Ikram and Morgan, 2000) and (Araki *et al.*, 2001) such a constraint is not effective for a long reverberant environment and the performance of the algorithm may degrade as the length of the separating filter increases. In the proposed method we do not require an FIR constraint on the mixing model nor on the un-mixing system, mainly because we use a different approach for resolving the permutation problem.

To resolve the permutation problem we follow the same approach as in (Rahbar and Reilly, 2001a) by exploiting the inherent non-stationarity of the input signals, which basically is done through the second stage of the algorithm by solving a discrete optimization problem.

For the first stage of the algorithm, similar to the approach used in the previous chapter, we optimize the criterion given by (4.5) using an alternating least-squares method(ALS). Note that the separation algorithm in (Parra and Spence, 2000) is based on using a gradient method to minimize the suggested cost function. The advantage of using ALS (rather than gradient based optimization methods) is that it usually has fast convergence (as is demonstrated in simulations) and there are no parameters to adjust.

Most of the steps for minimizing the criterion (4.5) are similar to the ones in the previous chapter. Nevertheless to be complete and for ease of reference the whole procedure including those steps that are similar to the ones in the previous chapter are written below.

Using the properties of Kronecker products (Brewer, 1979), the quantity $\mathbf{B}(\omega_k)\mathbf{\Lambda}(\omega_k, m)\mathbf{B}^{\dagger}(\omega_k)$ in (4.5) can be written as

$$\mathrm{vec}\{\mathbf{B}(\omega_k)\mathbf{\Lambda}(\omega_k, m)\mathbf{B}^{\dagger}(\omega_k)\} = \mathbf{B}(\omega_k) \odot \mathbf{B}(\omega_k)\,\mathrm{diag}\{\mathbf{\Lambda}(\omega_k, m)\} \tag{4.6}$$

where $\odot$ is the Khatri-Rao product and is defined as:

$$\mathbf{B}(\omega_k) \odot \mathbf{B}(\omega_k) = [\mathbf{b}_1(\omega_k) \otimes \mathbf{b}_1(\omega_k), \dots, \mathbf{b}_N(\omega_k) \otimes \mathbf{b}_N(\omega_k)] \tag{4.7}$$

where $\mathbf{b}_i(\omega_k)$ is the i_{th} column of $\mathbf{B}(\omega_k)$ and $\otimes$ represents the Kronecker product. Setting $\mathbf{G}(\omega_k) = \mathbf{B}(\omega_k) \odot \mathbf{B}(\omega_k)$, $\mathbf{d}(\omega_k, m) = \mathrm{diag}\{\mathbf{\Lambda}(\omega_k, m))\}$ and $\hat{\mathbf{p}}_x(\omega_k, m) = \mathrm{vec}\{\hat{\mathbf{P}}_x(\omega_k, m)\}$ we can rewrite (4.5) as:

$$\min_{\mathbf{g}_i(\omega_k)\in\Omega,\mathbf{d}(\omega_k,m)} \sum_{k=0}^{K-1}\sum_{m=0}^{M-1} ||\hat{\mathbf{p}}_x(\omega_k,m) - \mathbf{G}(\omega_k)\mathbf{d}(\omega_k,m)||_2^2 \tag{4.8}$$

where $\mathbf{g}_i(\omega_k)$ is the i_{th} column of $\mathbf{G}(\omega_k)$. Since there is an inherent scaling ambiguity between $\mathbf{b}_i(\omega_k)$ and $d_i(\omega_k, m)$, the i_{th} diagonal value of $\mathbf{d}(\omega_k, m)$, in (4.5), without loss of generality we can assume $||\mathbf{b}_i(\omega_k)||_2^2 = 1$. Based on this we define the constraint set $\Omega \subset \mathbb{C}^{J^2\times 1}$ as:

$$\Omega = \{\text{vec}\{\mathbf{\Phi}\}|\mathbf{\Phi} = \boldsymbol{\nu}\boldsymbol{\nu}^\dagger,\ \boldsymbol{\nu} \in \mathbb{C}^{J\times 1},\ ||\boldsymbol{\nu}||_2^2 = 1\}. \tag{4.9}$$

In defining the constraint set Ω we used the fact that for the column vector $\boldsymbol{\nu}$ we have $\boldsymbol{\nu}\otimes\boldsymbol{\nu} = \text{vec}\{\boldsymbol{\nu}\boldsymbol{\nu}^\dagger\}$. We first minimize (4.8) with respect to $\mathbf{g}_i(\omega_k)$ conditioned on $\hat{\mathbf{d}}(\omega_k, m)$, the previously estimated values of $\mathbf{d}(\omega_k, m)$. To do this we form the matrices $\mathbf{T}(\omega_k) = [\hat{\mathbf{p}}(\omega_k, 0), ..., \hat{\mathbf{p}}(\omega_k, M-1)]$ and $\mathbf{F}(\omega_k) = [\hat{\mathbf{d}}(\omega_k, 0), ..., \hat{\mathbf{d}}(\omega_k, M-1)]$ and we write equation (4.8) as:

$$\min_{\mathbf{g}_i(\omega_k)\in\Omega} \sum_{k=0}^{K-1} ||\mathbf{T}(\omega_k) - \mathbf{G}(\omega_k)\mathbf{F}(\omega_k)||_F^2. \tag{4.10}$$

To minimize (4.10) we first find the unconstrained least-squares minimizer of (4.10) by setting

$$\tilde{\mathbf{G}}(\omega_k) = \mathbf{T}(\omega_k)\mathbf{F}^+(\omega_k). \tag{4.11}$$

We then project each column of $\tilde{\mathbf{G}}(\omega_k)$ onto Ω; i.e.,

$$\hat{\mathbf{g}}_i(\omega_k) = \text{proj}_{\,\Omega}[\tilde{\mathbf{g}}_i(\omega_k)] \tag{4.12}$$

where $\tilde{\mathbf{g}}_i(\omega_k)$ is the i_{th} column of $\tilde{\mathbf{G}}(\omega_k)$.

Similar to what is discussed in the previous chapter a convenient method of performing the projection operation is to solve the following minimization:

$$\hat{\mathbf{g}}_i(\omega_k) = \arg\min_{\mathbf{g}_i(\omega_k)\in\Omega} ||\tilde{\mathbf{g}}_i(\omega_k) - \mathbf{g}_i(\omega_k)||_2^2. \tag{4.13}$$

Since $\mathbf{g}_i(\omega_k) = \text{vec}\{\mathbf{b}_i(\omega_k)\mathbf{b}_i^\dagger(\omega_k)\}$, by defining $\mathbf{Y}_i(\omega_k) = \text{mat}\{\tilde{\mathbf{g}}_i(\omega_k)\}$ we can write the above equation as:

$$\begin{aligned} &\min_{||\mathbf{b}_i(\omega_k)||_2=1} ||\mathbf{Y}_i(\omega_k) - \mathbf{b}_i(\omega_k)\mathbf{b}_i^\dagger(\omega_k)||_F^2 \equiv \\ &\min_{||\mathbf{b}_i(\omega_k)||_2=1} C - 2\mathbf{b}_i^\dagger(\omega_k)\mathbf{Y}_i(\omega_k)\mathbf{b}_i(\omega_k) \end{aligned} \tag{4.14}$$

where $C = 1 + Tr\big(\mathbf{Y}_i^\dagger(\omega_k)\mathbf{Y}_i(\omega_k)\big)$ is a constant term. The above minimization can be done by choosing $\hat{\mathbf{b}}_i(\omega_k)$, the estimated i_{th} column of $\mathbf{B}(\omega_k)$, to be the dominant eigenvector of $\mathbf{Y}_i(\omega_k)$. In the manner used in the previous chapter, to find the dominant eigenvector we use the power iteration method, described in (Golub and VanLoan, 1996), with one power iteration per ALS iteration.

To minimize (4.5) with respect to $\mathbf{d}(\omega_k, m)$, conditioned on the previous estimate of $\mathbf{G}(\omega_k)$, we solve the following least-squares problem.

$$\hat{\mathbf{d}}(\omega_k, m) = \arg \min_{\mathbf{d}(\omega_k, m)} ||\hat{\mathbf{p}}_x(\omega_k, m) - \hat{\mathbf{G}}(\omega_k)\mathbf{d}(\omega_k, m)||_2^2 \tag{4.15}$$

Minimizing (4.15) with respect to $\mathbf{d}(\omega_k, m)$ we get:

$$\hat{\mathbf{d}}(\omega_k, m) = \hat{\mathbf{G}}^+(\omega_k)\hat{\mathbf{p}}(\omega_k, m) \quad m = 0, ..., M-1, \;\; k = 0, \ldots, K-1. \tag{4.16}$$

Using equations (4.11), (4.12) and (4.16), we can repeatedly update the values of $\mathbf{d}(m)$ and $\mathbf{G}(\omega_k)$ until convergence is achieved.

As mentioned previously, to avoid being trapped in local minima, we need to properly initialize the algorithm. The initialization procedure can be similar to that in the previous chapter where at each frequency bin we can choose the initial estimate for $\mathbf{H}(\omega_k)$ using a closed-form joint diagonalization procedure as described before. Note that since we need to apply this initialization for each frequency bin, for real room experiments where we need a large number of frequency bins, this initialization procedure is computationally inefficient. In the following we describe another initialization method which not only is computationally efficient but also improves dramatically the perceptual quality of the separated audio signals.

4.3.1 Initialization

To initialize the algorithm we have different options. The first option is that a rough estimate of the mixing system at each frequency bin (up to some scaling and permutation ambiguity) can be obtained using the closed-form, exact joint diagonalization procedure described in previous chapter. Since we need to use this initialization procedure at each frequency bin, one draw-back of this initialization method will be its computational complexity.

An alternative, novel, *ad hoc* initialization method which not only requires less computation, but also dramatically improves the quality of the separated audio signals, is described as follows. The main idea of this initialization procedure, is that first we choose the initial value of $\mathbf{B}(\omega_0)$, the first frequency bin, using the exact closed form joint diagonalization method mentioned above. We then apply the ALS algorithm to find the final estimate of $\mathrm{B}(\omega_0)$. This final estimate is then used as an initial value for the next adjacent frequency bin, which is $\mathbf{B}(\omega_1)$. The outcome of this frequency bin is also used as an initial value for the next frequency bin and this procedure continues until all the frequency bins have been covered. Note that in this way we need to apply the exact closed form joint diagonalization algorithm only for one frequency bin.

As has been demonstrated in our simulation results, this initialization procedure significantly improves the quality of the separated audio signals. An intuitive explanation for this is as follows. We realize that the estimate of $\boldsymbol{B}(\omega_k)$ is not unique because each column $\boldsymbol{b}_i(\omega_k)$ is still subject to a multiplicative phase ambiguity, even though the condition $\| \boldsymbol{b}_i \|_2 = 1$ in the solution of (4.5) has been enforced[2]. Fast variation of this phase ambiguity in frequency can cause the resulting time–domain estimate $\hat{\boldsymbol{H}}(t)$ of the channel to be excessively long. By initializing the algorithm in the manner proposed, this phase ambiguity varies smoothly with frequency, therefore creating an $\hat{\boldsymbol{H}}(t)$ which can be of moderate length. As is shown in our simulations, the resulting overall system (channel + inverse) is then much more localized in time. This property is known to minimize degradation in audio quality due to reverberative effects.

Summary of Stage I of the Algorithm for Blind Source Separation

[2] Note that placing a constraint on both the norm and the phase of the columns $\boldsymbol{b}_i$ would lead to a less computationally efficient algorithm.

```
1. Estimate the normalized observed signals' cross spectral density matrices, $\hat{\mathbf{P}}_x(\omega_k, m)$ and set $\mathbf{T}(\omega_k) = [\hat{\mathbf{p}}_x(\omega_k, 0), ..., \hat{\mathbf{p}}_x(\omega_k, M-1)]$ where $\hat{\mathbf{p}}_x(\omega_k) = \text{vec}\{\hat{\mathbf{P}}_x(\omega_k, m)\}$

2. Set $\hat{\mathbf{B}}^0(\omega_k)$, the initial value for the $\mathbf{B}(\omega_k)$, based on the method described in section 4.3.1

3. for $k = 0$ to $K-1$
    • for $\nu = 0$ to Max_itr
        – Set $\hat{\mathbf{G}}^\nu(\omega_k) = [\hat{\mathbf{b}}_1^\nu(\omega_k) \otimes \hat{\mathbf{b}}_1^\nu(\omega_k), \ldots, \hat{\mathbf{b}}_N^\nu(\omega_k) \otimes \hat{\mathbf{b}}_N^\nu(\omega_k)]$
        – Calculate $\hat{\mathbf{d}}^\nu(\omega_k, m) = (\hat{\mathbf{G}}^\nu(\omega_k))^+ \hat{\mathbf{p}}(\omega_k, m) \quad m = 0, ..., M-1$
        – Set $\mathbf{F}^\nu(\omega_k) = [\hat{\mathbf{d}}^\nu(\omega_k, 0), ..., \hat{\mathbf{d}}^\nu(\omega_k, M-1)]$
        – Calculate $\tilde{\mathbf{G}}^\nu(\omega_k) = \mathbf{T}(\omega_k)(\mathbf{F}^\nu(\omega_k))^+$
        – for $i = 1$ to $N$
            * $\mathbf{Y} = \text{mat}\{\tilde{\mathbf{g}}_i^\nu(\omega_k)\}$
            * $\mathbf{q} = \mathbf{Y}\mathbf{b}_i^\nu(\omega_k)$
            * $\hat{\mathbf{b}}_i^{\nu+1}(\omega_k) = \frac{\mathbf{q}}{||\mathbf{q}||_2}$
        – end
        – Calculate the cost value $C_k^\nu = ||\mathbf{T}(\omega_k) - \hat{\mathbf{G}}^\nu(\omega_k)\mathbf{F}^\nu(\omega_k)||_F^2$
        – if $\frac{|C_k^\nu - C_k^{\nu-1}|}{C_k^\nu} < \epsilon$, where $0 < \epsilon \ll 1$, then stop, go to the next frequency bin
        – end
    • end
```

4.4 Resolving Permutations

One potential problem with the cost function in (4.5) is that it is insensitive to permutations of the columns of $\mathbf{B}(\omega_k)$. More specifically if $\mathbf{B}_{opt}(\omega_k)$ is an optimum solution to (4.5) then

$\mathbf{B}_{opt}(\omega_k)\mathbf{\Pi}_k$, where $\mathbf{\Pi}_k$ is an arbitrary permutation matrix for each ω_k, will also be a optimum solution. Since in general $\mathbf{\Pi}_k$ can vary for different frequency bins, this will result in overall poor separation performance.

In this section we suggest a novel solution for solving the permutation problem which exploits the cross-frequency correlation between diagonal values of $\mathbf{\Lambda}(\omega_k, m)$ and $\mathbf{\Lambda}(\omega_{k+1}, m)$ given in (4.5). Notice that $\mathbf{\Lambda}(\omega_k, m)$ can be considered as an estimate of the sources' cross-power spectral density at epoch m. When the sources are speech signals the temporal trajectories of the power spectral density of speech, known as spectrum modulation of speech, are correlated across the frequency spectrum. Using this correlation we can adjust the wrong permutations as shown in this example for the two source case.

Assume that $\mathbf{\Lambda}(\omega_k, m) \quad m = 0, \cdots, M-1$ represents the estimated cross-spectral density of the two sources at frequency bin ω_k. We want to adjust the permutation at frequency ω_j such that it has the same permutation as in frequency bin ω_k. To do so we first calculate the cross frequency correlation between the diagonal elements of $\mathbf{\Lambda}(\omega_j, m)$ and $\mathbf{\Lambda}(\omega_k, m)$ using the following measure:

$$\rho_{qp}(\omega_k, \omega_j) = \frac{\sum_{m=0}^{M-1} \lambda_q(\omega_k, m)\lambda_p(\omega_j, m)}{\sqrt{\sum_{m=0}^{M-1} \lambda_q^2(\omega_k, m)}\sqrt{\sum_{m=0}^{M-1} \lambda_p^2(\omega_j, m)}} \tag{4.17}$$

where $\rho_{qp}(\omega_k, \omega_j)$ represents the cross frequency correlation between$\lambda_q(\omega_k, m)$, the q_{th} diagonal element of $\mathbf{\Lambda}(\omega_k, m)$, and $\lambda_p(\omega_j, m)$, the p_{th} diagonal element of $\mathbf{\Lambda}(\omega_j, m)$. If the frequency bins ω_k and ω_j have the same permutation then we expect that

$$\frac{\rho_{11}(\omega_k, \omega_j) + \rho_{22}(\omega_k, \omega_j)}{\rho_{12}(\omega_k, \omega_j) + \rho_{21}(\omega_k, \omega_j)} > 1; \tag{4.18}$$

otherwise, we need to change the permutation at one of frequency bins ω_k or ω_j such that the above condition is satisfied. We can apply the above ratio test to all frequency bins to detect and adjust the wrong permutations. In general when number of sources are greater than two, the ratio test given in (4.18) can be written as following discrete optimization problem

$$\max_{\mathbf{\Pi}_k \in \mathcal{P}} \quad \text{trace}(\mathbf{\Pi}_k \mathbf{E}(\omega_k)\mathbf{E}^T(\omega_j)) \tag{4.19}$$

where $\mathcal{P}$ is the set of $N \times N$ permutation matrices including the identity matrix, and $\mathbf{E}(\omega_k)$ is an $N \times M$ matrix given as

$$\mathbf{E}(\omega_k) = \left(\sum_{m=0}^{M-1} \mathbf{\Lambda}^2(\omega_k, m)\right)^{-\frac{1}{2}} \begin{pmatrix} \lambda_1(\omega_k, 0) & \dots & \lambda_1(\omega_k, M-1) \\ \vdots & \ddots & \vdots \\ \lambda_N(\omega_k, 0) & \dots & \lambda_N(\omega_k, M-1) \end{pmatrix}. \tag{4.20}$$

The discrete optimization criterion in (4.19) can be solved by enumerating over all possible selections for $\mathbf{\Pi}_k$. This means for a set of $N \times N$ permutation matrices we need to calculate the criterion in (4.19) $N!$ times to find the optimum solution for $\mathbf{\Pi}_k$. For large values of N this may not be computationally efficient. A more computationally efficient but less optimal approach to estimate the permutation matrix between the two frequency bins is given by the following algorithm.

Adjusting Permutations

1. initialize the $N \times N$ matrix $\mathbf{\Pi}_k$ to an all zeros matrix.

2. For $i = j$ and $i = k$ set up the matrices

$$\mathbf{E}(\omega_i) = \left(\sum_{m=0}^{M-1} \mathbf{\Lambda}^2(\omega_i, m)\right)^{-\frac{1}{2}} \begin{pmatrix} \lambda_1(\omega_i, 0) & \dots & \lambda_1(\omega_i, M-1) \\ \vdots & \ddots & \vdots \\ \lambda_N(\omega_i, 0) & \dots & \lambda_N(\omega_i, M-1) \end{pmatrix} \tag{4.21}$$

3. Form the multiplication $\mathbf{T}_{kj} = \mathbf{E}(\omega_k)\mathbf{E}^T(\omega_j)$

4. Find the row number r_{max} and column number c_{max} corresponding to the element of $\mathbf{T}$ with largest absolute value. Zero all elements of $\mathbf{T}$ corresponding to this row and column numbers and set $\mathbf{\Pi}_k(c_{max}, r_{max}) = 1$.

5. Recursively repeat the previous step for the remaining elements of matrix $\mathbf{T}_{kj}$ until only one non-zero element remains. Set

$$\mathbf{\Pi}_k(c_f, r_f) = 1 \tag{4.22}$$

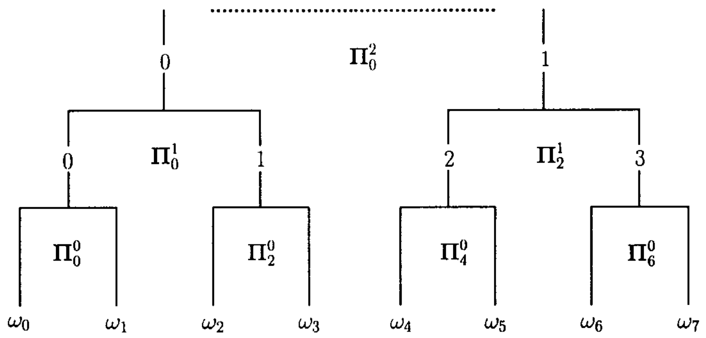

Figure 4.1: An example of the diadic permutation sorting algorithm for the case when the total number of frequency bins is eight.

where r_f and c_f are the corresponding row and column numbers of the remaining non-zero element.

Notice that the above algorithm calculates the permutation matrix $\mathbf{\Pi}_k$ between two frequency bins ω_k and ω_j. To obtain a uniform permutation across the whole frequency spectrum, we need to apply the above algorithm repeatedly to all pairs of frequency bins. One way of doing this is to adjust the permutation between adjacent frequency bins in a sequential order where, for example, starting from frequency bin ω_0 we adjust the permutation of each bin relative to it's previous bin. This approach, although simple, has a major drawback as explained as follows. Consider the situation where an error is made in estimating the correct permutation matrix for frequency bin ω_k. In this case, all frequency bins placed after ω_k will receive a different permutation than the ones placed before ω_k. In the worst case scenario we will have half of the frequency bins with one permutation and the other half with different permutation, which will result in no or very poor separation. To prevent such a catastrophic situation we propose following hierarchical sorting scheme to sort the permutations across all frequency bins. For clarity we explain the algorithm for the case when we have only eight frequency bins (Figure 4.1). Extension to general case of arbitrary frequency bins can be easily deduced.

Diadic Sorting Algorithm

---

1. Divide the frequency bins into groups of two bins[3] each with group index p.

2. Let k and $k+1$ be the indices to the frequency bins inside the group p and $\mathbf{\Pi}_k^0$ be the permutation matrix estimated from the criterion given in (4.19). Also let $\mathbf{\Sigma}_k^0(m) = \mathbf{\Lambda}(\omega_k, m)$ and $\mathbf{\Sigma}_{k+1}^0(m) = \mathbf{\Lambda}(\omega_{k+1}, m)$. Then for all $p = 0, \ldots, 3$, update the order of diagonal values of $\mathbf{\Sigma}_k^0(m)$ using

$$\mathbf{\Sigma}_k^0(m) = \mathbf{\Pi}_k^0 \mathbf{\Sigma}_k^0(m) \mathbf{\Pi}_k^{0^T} \quad k = 0, 2, 4, 6$$

3. Update the order of columns of $\mathbf{B}(\omega_k)$ using

$$\mathbf{B}(\omega_k) = \mathbf{\Pi}_k^0 \mathbf{B}(\omega_k) \quad k = 0, 2, 4, 6$$

4. For each group calculate

$$\mathbf{\Sigma}_p^1(m) = \mathbf{\Sigma}_k^0(m) + \mathbf{\Sigma}_{k+1}^0(m) \quad k = 0, 2, 4, 6 \quad p = 0, 1, 2, 3$$

5. Divide the set of $\mathbf{\Sigma}_0^1(m), \ldots, \mathbf{\Sigma}_3^1(m)$ into groups of two elements and for each of the new groups estimate the permutation matrices $\mathbf{\Pi}_p^1$ $p = 0, 2$ using the diagonal values of $\mathbf{\Sigma}_p^1(m)$ based on the criterion given in (4.19). Also for all $p = 0, 2$ update the order of diagonal values of $\mathbf{\Sigma}_p^1(m)$ using

$$\mathbf{\Sigma}_p^1(m) = \mathbf{\Pi}_p^1 \mathbf{\Sigma}_p^1(m) \mathbf{\Pi}_p^{1^T} \quad p = 0, 2$$

6. Update the order of columns of $\mathbf{B}(\omega_k)$ using

$$\mathbf{B}(\omega_{2p}) = \mathbf{\Pi}_p^1 \mathbf{B}(\omega_{2p})$$
$$\mathbf{B}(\omega_{2p+1}) = \mathbf{\Pi}_p^1 \mathbf{B}(\omega_{2p+1}) \quad p = 0, 2$$

[3] Here we assume that K, the total number of the frequency bins, is a multiple integer of 2.

7. For the new groups calculate

$$\mathbf{\Sigma}_q^2(m) = \mathbf{\Sigma}_p^1(m) + \mathbf{\Sigma}_{p+1}^1(m) \qquad p = 0, 2 \quad q = 0, 1 \tag{4.23}$$

8. Finally calculate $\mathbf{\Pi}_0^2$, by substituting $\mathbf{\Sigma}_0^2(m)$ and $\mathbf{\Sigma}_1^2(m)$ in (4.19). Update the columns of $\mathbf{B}(\omega_k) \quad k = 0, \ldots, 3$ using

$$\mathbf{B}(\omega_k) = \mathbf{\Pi}_0^2 \mathbf{B}(\omega_k) \qquad k = 0, \ldots, 3$$

4.5 Simulation Results

4.5.1 Example I, FIR Convolutive Mixing

The objective of this first simulation is to characterize the performance of the algorithm under a controlled mixing environment. For this purpose we use an 8-tap FIR convolutive mixing system where the impulse responses of its elements are selected randomly from a uniform distribution (Table 4.1). For the sources, we use two independent white Gaussian signals, multiplied by slowly varying sine and cosine signals to create the required non–stationary effect. For this example we kept the epoch size at 500 and the total data length was varied between 10000 and 50000 samples, corresponding to M, the number of epochs, ranging between 20 to 100 epochs. White Gaussian noise was added to the output of the

n	0	1	2	3	4	5	6	7
$H_{11}(n)$	-0.143	0.232	-0.129	-0.075	0.116	0.484	-0.131	0.147
$H_{12}(n)$	-0.227	-0.183	-0.107	0.473	0.465	-0.087	0.416	-0.199
$H_{21}(n)$	0.180	0.060	0.366	0.049	-0.183	0.483	-0.218	0.346
$H_{22}(n)$	0.088	0.287	-0.130	0.187	-0.112	0.442	0.046	-0.227
$H_{31}(n)$	-0.460	-0.4334	0.293	0.456	0.407	0.310	-0.193	-0.176
$H_{32}(n)$	-0.245	-0.340	0.129	0.487	-0.108	0.132	0.058	0.022

Table 4.1: Impulse Response of the MIMO system for Examples I & II

system at a level corresponding to the desired value of averaged SNR over all epochs[4]. At each epoch, 128–point FFTs, applied to time segments overlapping by 50%, weighted by Hanning windows were used to estimate the cross-spectral density matrices. For the joint diagonalization algorithm, for all frequency bins, we chose the initial estimate of the channel to be an identity matrix[5].

Let $\mathbf{C}(\omega_k)$ represents the global system frequency response

$$\mathbf{C}(\omega_k) = \mathbf{W}(\omega_k)\mathbf{H}(\omega_k) \tag{4.24}$$

where ij_{th} element is $c_{ij}(\omega_k)$. To measure the separation performance we can use following formula:

$$\mathrm{SIR}(i) = \frac{\sum_{q=0}^{M_c-1} \max_j(S_{ij}^q)}{\sum_{q=0}^{M_c-1} \left\{ \sum_{j=1}^{N} S_{ij}^q - \max_j(S_{ij}^q) \right\}} \tag{4.25}$$

where $S_{ij}^q = \sum_{k=0}^{K-1} |c_{ij}^q(\omega_k)|^2$, q is an index to the q_{th} Monte Carlo run and the quantity M_c is the total number of Monte Carlo runs. Notice that (4.25) implicitly measures the signal to interference ratio (SIR) for each output of the separating system. Here, at each output, the signal is the separated source that has the maximum power and the interference is considered as the contribution of the other sources. Figure 4.2 shows the variation of each outputs' SIR with M, number of epochs for a fixed signal to noise ratio (SNR=20dB). As can be seen from the figure, by increasing the number of epochs, which corresponds to increasing the data length, the output SIR improves (increases). Also Figure 4.3 shows how the separation performance changes with observed signals' signal to noise ratio for a fixed number of epochs $M = 50$. To demonstrate the effectiveness of the permutation algorithm, Table 4.2 shows the separation performance before and after the permutations have been resolved. Also refer to Figures 4.4 and 4.5 for a graphical visualization of the effects of arbitrary permutations at different frequency bins and the improvement caused by using the proposed permutation algorithm. As can be seen from the table and also the figures, the frequency dependent permutation ambiguity can severely degrade the overall separation

[4] The power of the noise was kept constant at all epochs.

[5] Of course we can use the initialization method described in previous section. Nevertheless the motivation of using an identity matrix for initialization in this example is to show that the algorithm may still converge, with reasonable error, even if we do not know of a good initialization point.

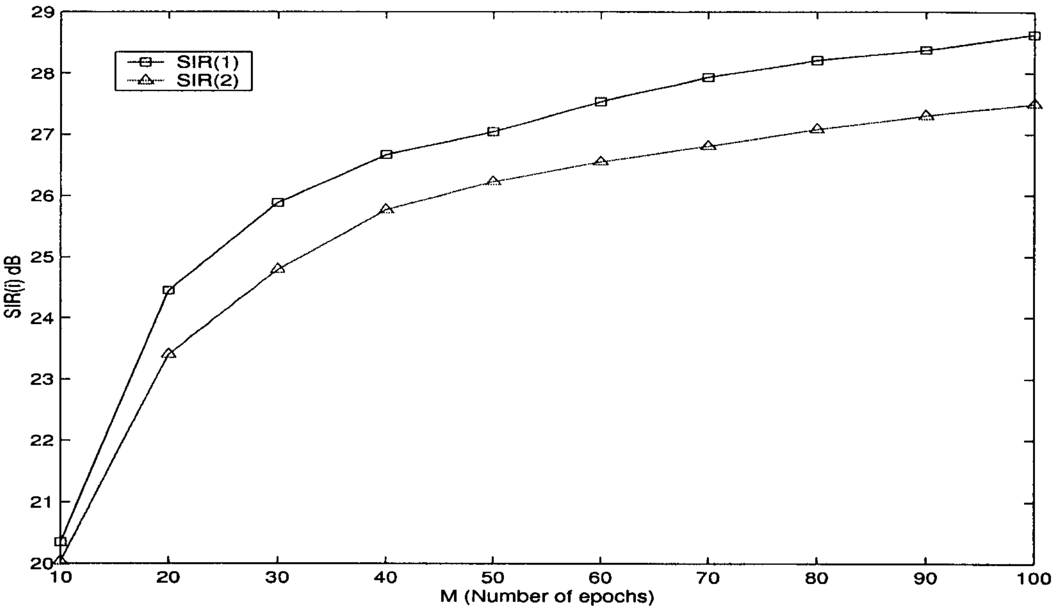

Figure 4.2: Example I, SIR versus M, number of epochs, for SNR=20dB and using $M_c = 50$ Monte Carlo runs.

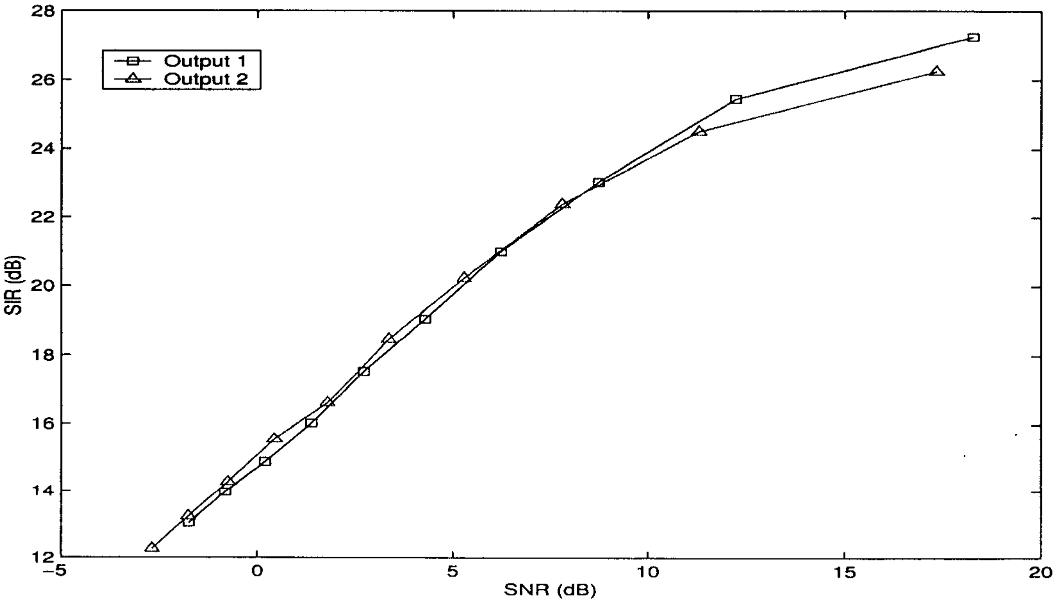

Figure 4.3: Example I, SIR versus SNR, for M=50 and using $M_c = 50$ Monte Carlo runs.

performance. Nevertheless the proposed algorithm is able to significantly improve the results by resolving these permutation ambiguities.

Output SIR (dB)	SIR(1)	SIR(2)
Before applying the permutation algorithm	3.5 dB	4.2 dB
After applying the permutation algorithm	27 dB	26 dB

Table 4.2: Example I, Output SIR before and after applying the permutation algorithm. SNR=20dB, M=50 and using $M_c = 50$ Monte Carlo Runs.

4.5.2 Example II, IIR Convolutive Mixing

The purpose of this example is to show how the algorithm can perform when there are more than two sources and when the mixing system is a matrix of stable IIR filters. Most of the existing algorithms in the literature consider only the case when the mixing system is a matrix of FIR filters. This simulation shows that the proposed algorithm, given enough data samples at each epoch to estimate the cross spectral density matrices, can perform well for the IIR case. For this example we use a 3×3 IIR mixing system where the impulse response of its ij_{th} element is given as:

$$h_{ij}(z) = \frac{b_{ij}}{1 - \alpha_{ij} z^{-1}} \tag{4.26}$$

where $0 < \alpha_{ij} < 1$ and b_{ij} are given as:

$$\alpha_{ij} \in \begin{pmatrix} 0.555 & 0.761 & 0.198 \\ 0.921 & 0.432 & 0.633 \\ 0.144 & 0.188 & 0.231 \end{pmatrix}, \; b_{ij} \in \begin{pmatrix} -1.50 & -0.19 & 1.21 \\ -0.52 & 2.53 & -0.27 \\ -0.95 & -0.03 & 0.38 \end{pmatrix}. \tag{4.27}$$

We use three sources here where the first two are the same as example I and third one is a white Gaussian signal with a slowly decaying exponential envelope. Source 3 is also statistically independent from the previous two sources. The sources were mixed using the Matlab "filter" command and similar to example I, white Gaussian noise was added to the results of mixture at a variable power depending on the desired average signal to noise ratio.

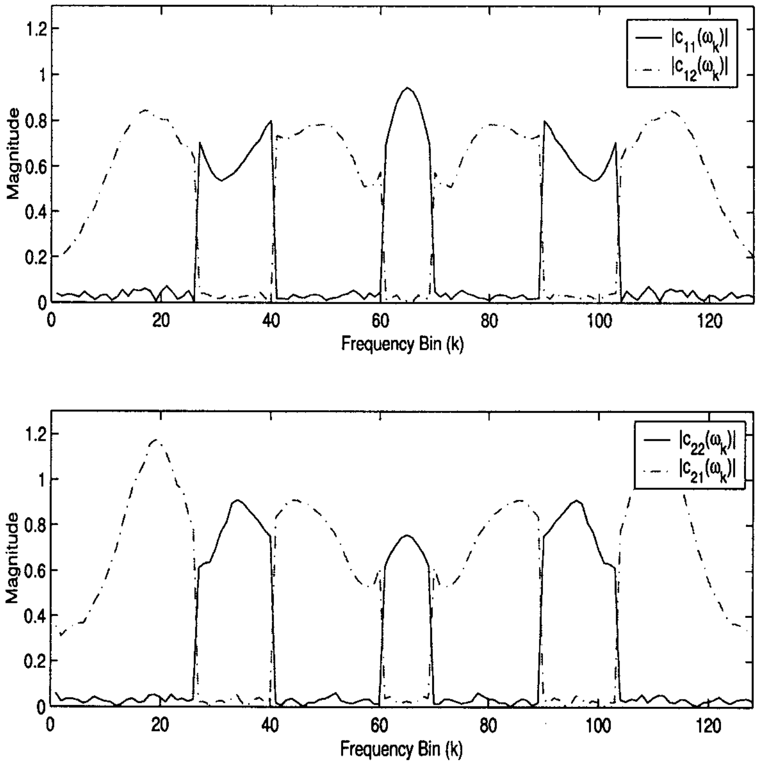

Figure 4.4: Example 1: Effects of random permutations in the frequency domain before applying the permutation algorithm,(M=50, SNR=20dB), $\mathbf{c}_{ij}(\omega_k)$ is the ij_{th} element of global system $\mathbf{C}(\omega_k) = \mathbf{W}(\omega_k)\mathbf{H}(\omega_k)$.

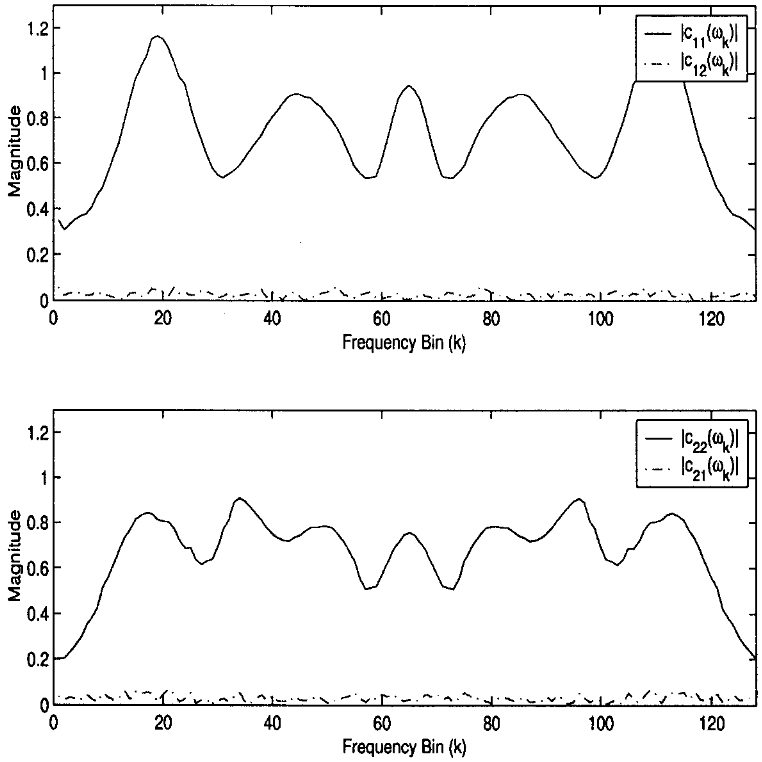

Figure 4.5: Example 1: Results after applying the permutation algorithm,(M=50, SNR=20dB), $\mathbf{c}_{ij}(\omega_k)$ is the ij_{th} element of global system $\mathbf{C}(\omega_k) = \mathbf{W}(\omega_k)\mathbf{H}(\omega_k)$.

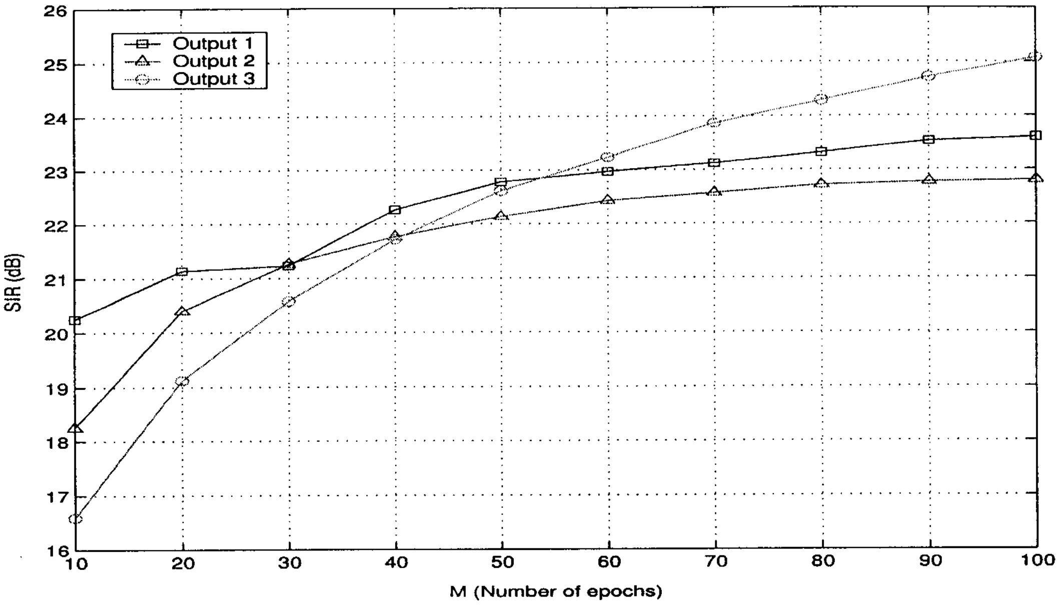

Figure 4.6: Example II, SIR versus M, for SNR=20dB and using M_c = 50 Monte Carlo runs.

For this example the epoch size was increased to 4000 data samples and at each epoch we used 256–point FFTs, applied to time segments overlapping by 80%, weighted by Hanning windows to estimate the cross-spectral density matrices. For each output the signal to interference ratio was measured using the criterion in (4.25) and the results are shown in Figure 4.6 for varying number of epochs and 20dB signal to noise ratio. Also Figure 4.7 shows the variation of output SIR with respect to the observed signals' signal to noise ratio for a fixed number of epochs (M=50).

4.6 Real Room Experiments

In this section we present the results of applying our algorithm to blind source septation of speech signals in a real reverberant environment. All recordings were done using 8.0 Khz, 16 bits sampling format.

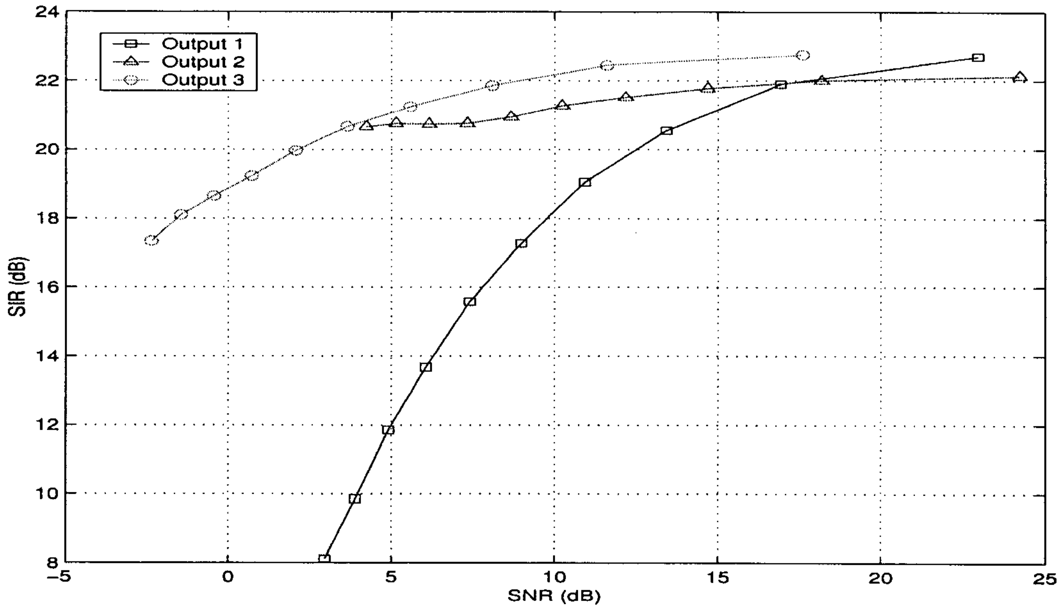

Figure 4.7: Example II, SIR versus SNR, for M=50 and using $M_c = 50$ Monte Carlo runs.

4.6.1 Real Room Experiment I

For the first set of experiments the recording were performed in an office room. We used two speakers as the sources and four *omnidirectional* microphones for recording the signals (Figure 4.8). Sources were created by catenating multiple speech segments from the TIMIT speech database. The speech signals then were played simultaneously through two speakers with approximately the same sound volume. The duration of the recording was around three minutes. To measure the separation performance, using the same setup, we also recorded white noise signals that were played through each speaker one at a time (One source was active at each time). Let $\hat{\sigma}_x^2(x_i, s_j) = \sum_{t=0}^{T-1} x_i^2(t)$ represent the power of the recorded signal at the i_{th} microphone when only speaker j is active and all other speakers (sources) are inactive. By playing white noise through each speaker at a time we can measure the signal to interference ratio for the recorded signal at the output of the i_{th} microphone using

$$SIR_x(i) = \frac{\max_j \hat{\sigma}_x^2(x_i, s_j)}{\sum_{j=1}^{N} \hat{\sigma}_x^2(x_i, s_j) - \max_j \hat{\sigma}_x^2(x_i, s_j)}. \tag{4.28}$$

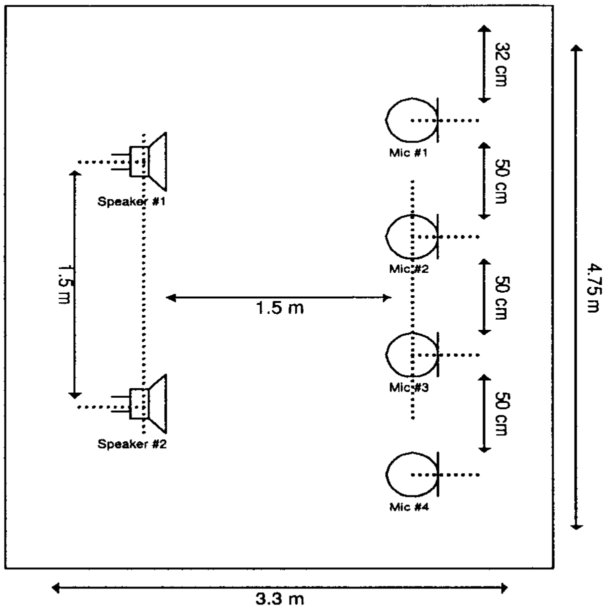

Figure 4.8: Real Room Experiments, Recording setup in an office room.

Using the above formula we can also measure the SIR for each output of the separating network by substituting $\hat{\sigma}_x^2(x_i, s_j)$ with $\hat{\sigma}_y^2(y_i, s_j)$, the power of the signal at the i_{th} output of the separating matrix when only source j is active. To perform the separation, in a manner similar to examples I and II, we first divided the recorded signal into multiple time segments (epochs), where we chose the size of each epoch to be around 10000 data samples long. For each epoch we calculated the cross spectral density matrices in a similar way as previous examples with these parameters: Number of FFT-points=4096, and overlap-percentage=80%. Figure 4.9 shows the output SIR versus M, the number of epochs. As can be seen for $M = 100$ an average SIR of more than 20dB is reached for each output. As a reference, Table 4.3 shows the SIRs for the recorded signals. By comparing the SIRs before and after applying the separating algorithm, it can easily be seen that the output SIRs have been improved by 19 to 20dB.

As mentioned in previous sections in this algorithm we can recover the sources up to a frequency dependent scaling ambiguity. The effect of the scaling ambiguity can deteriorate the quality of the separated audio signals.

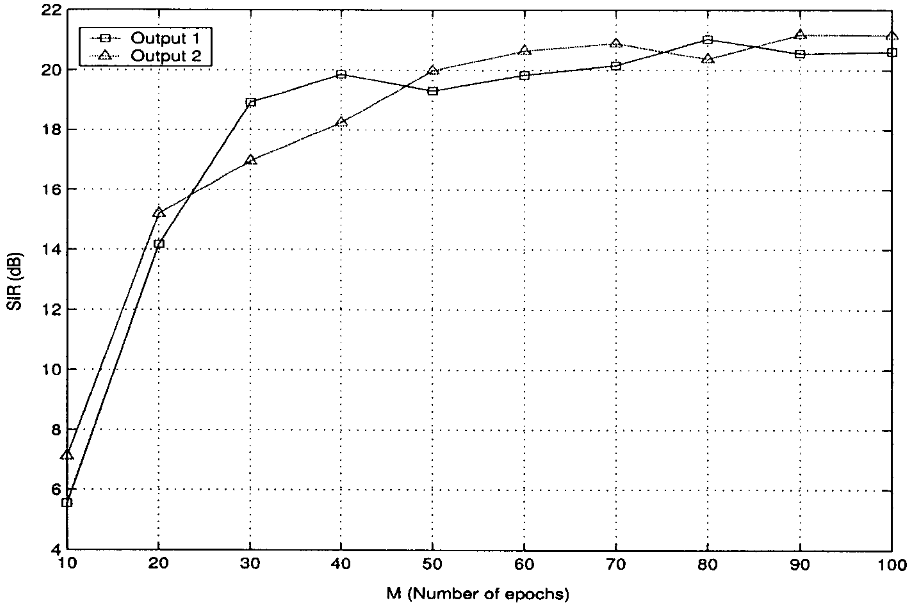

Figure 4.9: Results of separation for recordings in an Office room: SIR versus M, number of epochs, $K = 4096$.

	MIC 1	MIC 2	MIC 3	MIC 4
SIR	2.6945 dB	1.2282 dB	0.3266 dB	1.4031 dB

Table 4.3: Input SIRs for the recorded signals in an office environment

The listening tests show that the sequential initialization suggested in our algorithm dramatically improves the quality of the separated speeches in these experiments[6] .

4.6.2 Real Room Experiment II

In the next set of experiments, we performed real recordings in a highly reverberant conference room. The recording setup is similar to the previous experiment with the difference that the room dimension and the distance between the microphones and speakers are increased for this experiment (Figure 4.10). To compare the reverberation characteristic of the room used in this experiment to the one in the previous experiment, we measured the

[6] To hear the recordings and also the separation results please refer to the following website: "www.ece.mcmaster.ca/˜reilly/kamran"

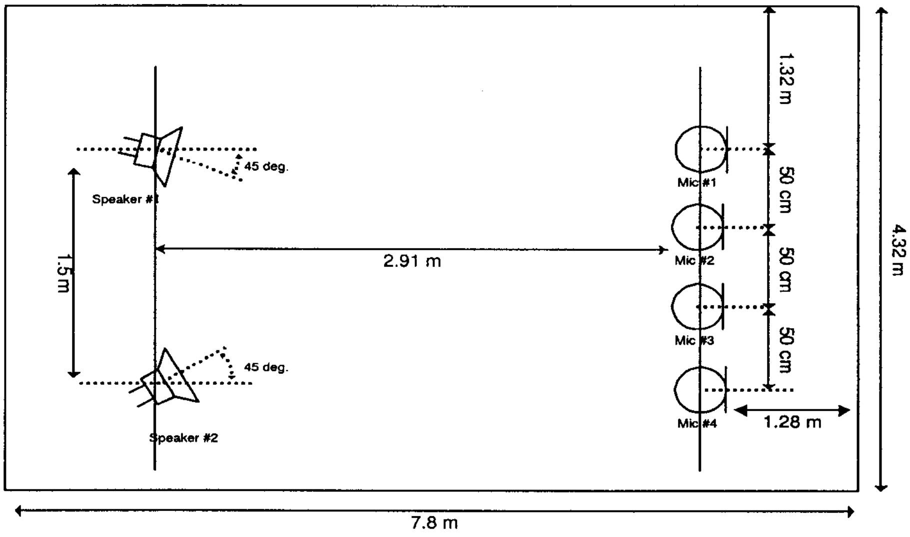

Figure 4.10: Real Room Experiments, Recording setup in a conference room.

reverberation times[7]of both rooms using our recording system and the software package "WINMLS2000". The results are shown in Figure 4.11. As can be seen from the figure, on average the reverberation time of the conference room, used in this experiment, is much higher than the reverberation time of the office room, used in the previous experiment. Due to the long reverberation time of the room, we expect more frequency bins are needed to estimate the cross spectral density matrices of the recorded signals. In this experiment the size of the epochs were kept the same as in the previous experiment[8] . Figure 4.12 shows how the performance of the algorithm improves by increasing the number of frequency bins used to estimate the cross spectral density matrices. As can be seen, a total number of 16384 frequency bins is needed to achieve a separation performance around 16dB. Also similar to the previous examples, the algorithm shows consistency with regards to improving the separation performance versus increasing the number of the epochs (Figure 4.13).

[7] The reverberation time in a room at a given frequency is the time required for the mean-square sound pressure in that room to decay from a steady state value by 60dB after the sound suddenly ceases(see also (Schroeder, 1965)).

[8] For $K = 16384$, since the epoch size was 10000, we used zero padding to compensate for the missing samples

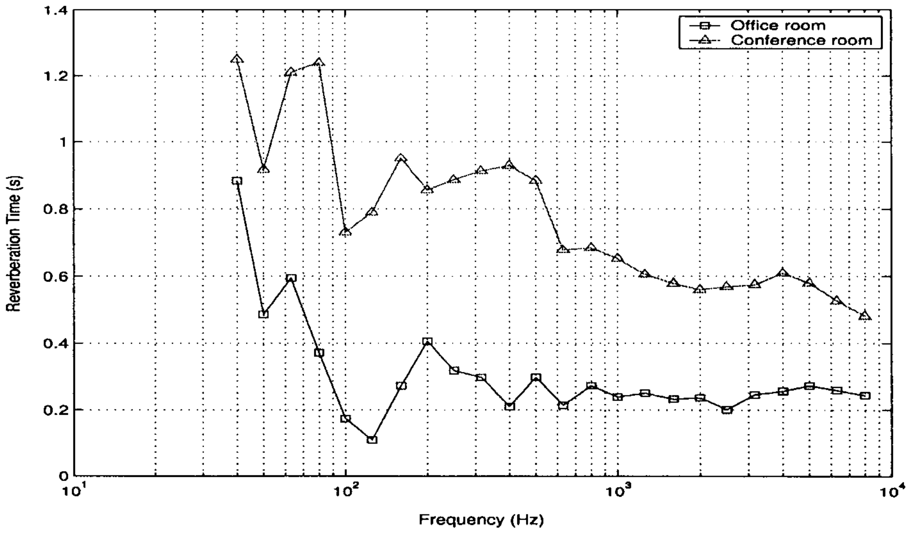

Figure 4.11: Comparison between the reverberation times of the rooms used in experiments I and II.

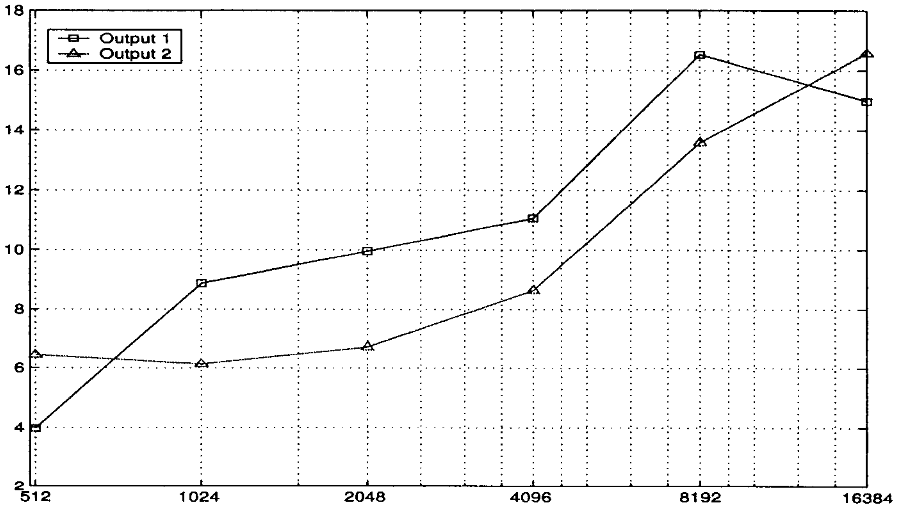

Figure 4.12: Separation performance versus number of frequency bins (K) for recordings in a conference room, $M = 140$.

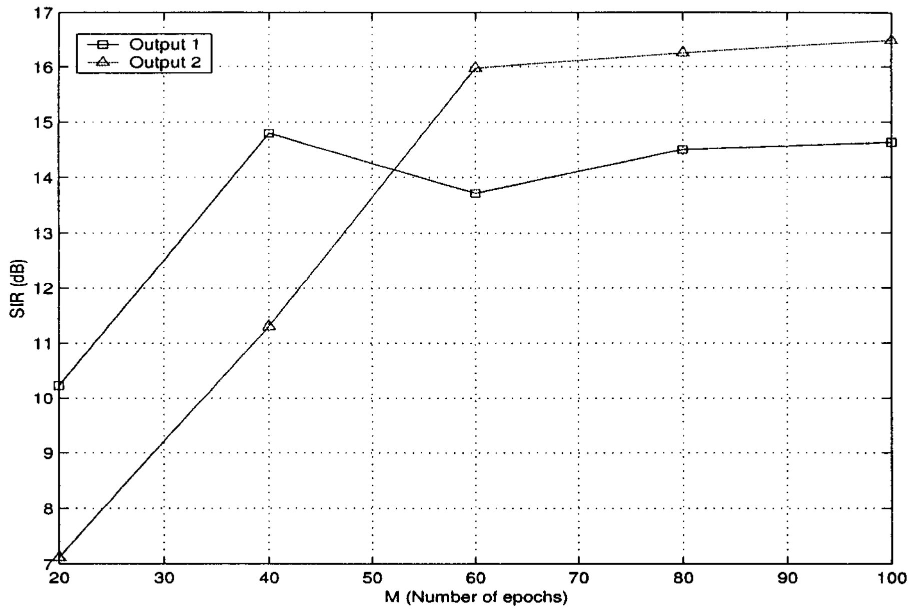

Figure 4.13: Results of separation for recordings in an Office room: SIR versus M, number of epochs, $K = 16384$

4.6.3 Comparisons with existing methods

In this section we compare the performance of our method with that of (Parra and Spence, 2000). Similar to the proposed algorithm, the method in (Parra and Spence, 2000) uses the estimated CPSD matrices over different time segments to calculate the un-mixing filters. Because of this we use the set of the CPSD matrices, evaluated previously in real room experiment I, as the input to both the proposed method and the algorithm in (Parra and Spence, 2000). Notice since the algorithm in (Parra and Spence, 2000) uses a finite length constraint on the size of un-mixing filters, the length of the un-mixing filters needs to be set beforehand in the program. Figure 4.14 shows the separation performance results for the algorithm in (Parra and Spence, 2000) versus the size of the un-mixing filters. As can be seen from the figure, the maximum SIR, which is around 4.5 dB, happens when the length of the un-mixing filters is around 512. Increasing the filter lengths after that degrades the separation performance. The comparative results, for the proposed method and the method in (Parra and Spence, 2000), are shown in Figure 4.15. For this experiment, based on the

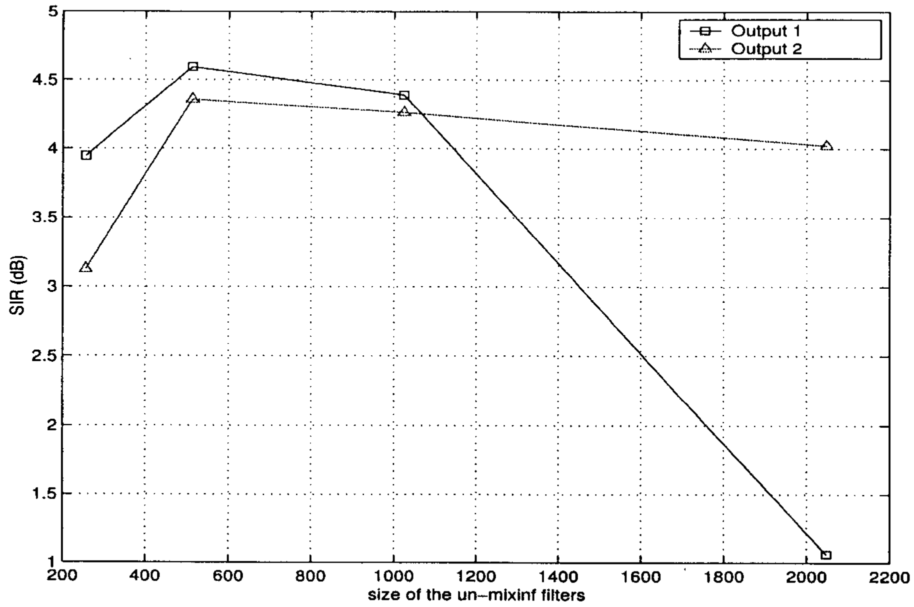

Figure 4.14: Results of separation for method in (Parra and Spence, 2000) for the set of recordings in an office room: SIR versus Q, size of un-mixing filters, $M = 20$.

previous simulation, we chose the length of un-mixing filter to be 512 for the algorithm in (Parra and Spence, 2000). As can be seen from the figure, the proposed algorithm outperforms the method in (Parra and Spence, 2000) by more than 15dB.

4.7 Summary

In this chapter we discussed a new recursive algorithm for blind source separation of convolved non-stationary sources. We proposed a two stage frequency domain algorithm to estimate the un-mixing filters. The main contribution of this work is that, unlike most of the existing algorithms, we do not make any assumptions on the structure of the mixing system; i.e. the elements of the mixing system can be FIR or IIR filters. Another strong point of the algorithm is its use of a recursive least-squares algorithm which has the advantage of fast convergence and no parameter tuning is required. We discussed the frequency domain permutation problem and we proposed efficient methods for solving this problem. We proposed an initialization procedure that helps alleviate the frequency dependent scaling

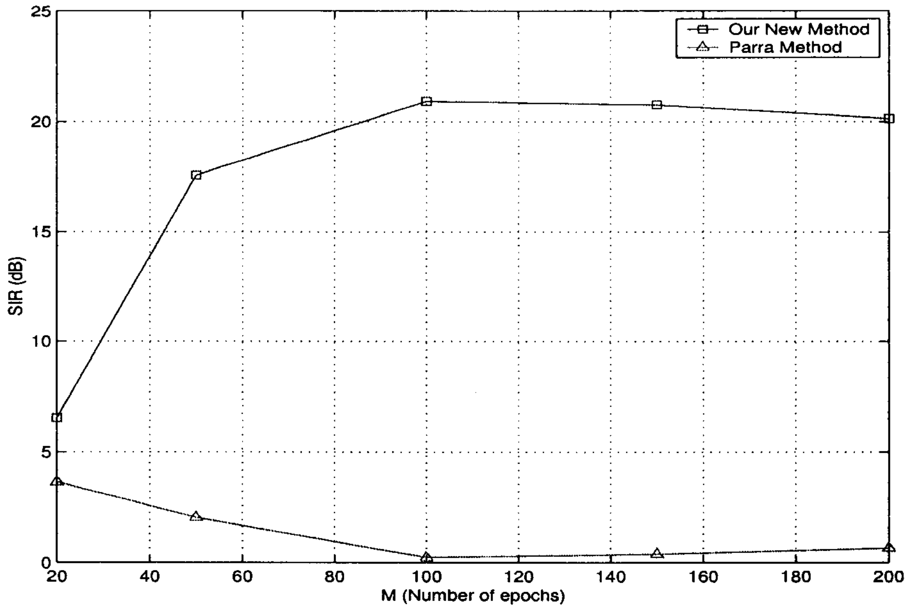

Figure 4.15: Comparison of the new proposed algorithm with the method in (Parra and Spence, 2000), Averaged output SIR versus M, number of epochs.

ambiguity problem and results in significantly improved perceptual quality of the output audio signal. We demonstrated the performance of the new algorithm using computer generated sources and mixing systems for different mixing scenarios. We also demonstrated the performance of the algorithm using real world data obtained by recording speech signals in different acoustic environments. We showed that algorithm can perform well in real environments with more than 20dB improvement in signal to interference ratio for a moderately reverberant office area. We showed for the real world source separation case the proposed algorithm outperforms the method in (Parra and Spence, 2000).

Chapter 5

Conclusions

This thesis discussed new algorithms for blind source separation of convolved mixtures and blind identification of MIMO systems. We showed how these problems can be solved by exploiting the non-stationarity of sources and using only the second order statistics of the observed signals. We also proposed efficient frequency domain approaches for solving these two problems.

This thesis presented a novel extension of joint diagonalization methods to the blind source separation of convolved mixtures and blind identification of MIMO systems. Because of the close relationship between the joint diagonalization problem, the blind source separation and MIMO blind identification problems, part of the material presented in this thesis was dedicated to developing new algorithms for the joint diagonalization problem including adaptive joint diagonalization methods using unconstrained optimization over the Stiefel manifold.

One of the key contributions of this thesis is that the proposed blind source separation algorithms can be applied in real world situations. Very few BSS methods so far have been proposed that can actually work in the real world scenarios and their performance is poor compare to the proposed algorithm.

Overall in the course of this thesis we addressed the following problems:

1. A set of adaptive algorithms for joint diagonalization problem was developed. We

employed the recently developed unconstrained optimization methods over the Stiefel manifold to minimize the proposed least-squares and maximum likelihood joint diagonalization criteria.

As shown by simulation results, the least-squares based algorithms have the same performance as that of the extended Jacobi joint diagonalization method of JADE. Nevertheless the new proposed methods have the additional advantage of being adaptive. Also the presented maximum likelihood method shows superior performance compared to the extended Jacobi joint diagonalization method of JADE.

2. The next problem that was discussed in this thesis is blind identification of FIR MIMO systems. We derived sufficient conditions for identifying a FIR MIMO system, driven by non-stationary sources, in the frequency domain, using only the second-order statistics of the observed signals. We also discussed the minimum number of frequency bins sufficient to identify the MIMO system in this case. A two stage frequency domain algorithm was presented. The first stage of algorithm, based on a alternating least-squares method, identifies the channel at each frequency bin up to a constant permutation, but frequency dependent scaling ambiguity. The second stage of algorithm removes the frequency dependent scaling ambiguity using a closed form method.

3. The last problem discussed in this thesis is the blind source separation of convolved audio(speech) signals. The presented algorithm is an extension of our frequency domain algorithm for blind identification of FIR MIMO systems. In this algorithm we do not make any assumptions on the structure of the mixing system; i.e., the impulse response of the mixing system can be FIR or IIR. Compared to the MIMO blind identification algorithm, we also use a less restrictive set of assumptions on the non-stationarity of sources. Due to this, the algorithm requires an extra step for removing the frequency dependent arbitrary permutations. Nevertheless, as has been verified by numerous real world experiments, the proposed algorithm can successfully separate recorded, mixed audio signals in a real reverberant environment.

5.0.1 Contribution to the Scientific Literature

Most of the work presented in Chapters 2, 3 and 4 of this thesis has been published in various conference papers. Part of the material in Chapter 2 has not yet been previously published but is under preparation to be submitted for publication. The material in Chapter 3 of this thesis has been submitted as a full journal paper. The material of Chapter 4 has also been organized into a full journal paper.

- Journal Papers
  - *Blind identification of MIMO FIR systems driven by quasi-stationary sources using second order statistics: A frequency domain approach.* (Rahbar *et al.*, 2002b) Submitted to IEEE Transactions on Signal Processing, Feb 2002.
  - *A new frequency domain method for blind source separation of convolutive audio mixtures* Under Preparation.
- Conference Papers
  - *A frequency domain approach to blind identification of MIMO FIR channels driven by non-stationary sources* (Rahbar *et al.*, 2002c) SAM2002.
  - *Joint diagonalization of correlation matrices by using Newton methods with application to blind signal separation* (Joho and Rahbar, 2002) SAM2002.
  - *A frequency domain approach to blind identification of MIMO FIR systems driven by quasi-stationary signals* (Rahbar *et al.*, 2002a) ICASSP2002.
  - *A new blind source separation algorithm for MIMO convolutive mixtures* (Rahbar and Reilly, 2001a) ICA2001.
  - *Blind source separation of convolved sources by joint approximate diagonalization of cross spectral density matrices* (Rahbar and Reilly, 2001b) ICASSP2001.
  - *Geometric Optimization Methods for Blind Source Separation of Convolutive Mixtures* (Rahbar and Reilly, 2000) ICA2000.

5.0.2 Suggestions for Future Research

- *Blind MIMO Identification*: The proposed MIMO blind identification algorithm is based on the batch processing of the observed data. For some practical applications it is desirable that the algorithm be adaptive; i.e., the algorithm should update its current estimate of the unknown MIMO system with each new sample of the observed signals. Note that in this case the adaptive algorithm can be used both for identification and tracking the unknown MIMO system. Further research is required to make the necessary changes into the algorithm such that it can be used in an adaptive application. Other suggestions are

    1. Investigate the case where the order of the system is unknown or is underestimated.
    2. Further research on improving the performance of the algorithm by using a maximum likelihood criterion rather than the least-squares criterion used in the current algorithm.
    3. Further research on how to take advantage of the a-priori known sensor array geometry to improve the performance of the blind identification algorithm.

- *Blind Source Separation*: Similar to the above, it is desirable to make the proposed convolutive blind source separation algorithm adaptive such that it can cope with a changing acoustic environment or moving speakers. Other related topics that can be suggested for future research are:

    1. Investigating the case when the number of sources is unknown.
    2. Further research for the case where the number of sources are greater than the number of the sensors.
    3. Investigating the effect of the diffused noise source on the performance of the BSS algorithm.

4. Further research on how to incorporate the geometry of the array of sensors as a constraint into the optimization criteria for the BSS algorithm.

Appendix A

Matlab Program for *tpoint* function

The following code has been extracted from (Manton, 2002) and it has been modified for complex square orthogonal matrices. The program calculates the turning points of a quadratic function

$$g(\mathbf{Z}) = \text{Real}\{Tr(\mathbf{Z}^{\dagger}\mathbf{D}) + \frac{1}{2}\,\text{vec}\{\mathbf{Z}\}^{\dagger}\mathbf{H}\,\text{vec}\{\mathbf{Z}\} + \frac{1}{2}\text{Real}\{\text{vec}\{\mathbf{Z}\}^{T}\mathbf{C}\,\text{vec}\{\mathbf{Z}\}\}$$

where $\mathbf{D} \in \mathbb{C}^{N\times N}$, $\mathbf{H}$,$\mathbf{C} \in \mathbb{C}^{N^2\times N^2}$ are arbitrary matrices such that $\mathbf{H} = \mathbf{H}^{\dagger}$ and $\mathbf{C} = \mathbf{C}^{T}$ and $\mathbf{Z}$ is restricted to be of the form $\mathbf{Z} = \mathbf{XA}$ with $\mathbf{A}$ being skew-hermitian and $\mathbf{X} \in \mathbb{C}^{N\times N}$ require to be an orthogonal matrix such that $\mathbf{XX}^{\dagger} = \mathbf{I}$

- *function [Z]=tpoint(X,D,H,C); [n,n]=size(X);*
- *d=n*n;*

  % Form basis for tangent space
- *E = zeros(n,n,d); i=1;*
- *for r=1:n*

  %Diagonal elements of A

```
    – M=zeros(n,n); M(r,r)=1j; E(:,:,i)=X*M; i=i+1;

• end

• for r=1:n-1

  % Off diagonal elements of A

    – for c=r+1:n

        * M=zeros(n,n); M(r,c)=1; M(c,r)=-1; E(:,:,i)=X*M; i=i+1;

        * M(r,c)=1j; M(c,r)=1j; E(:,:,i)=X*M; i=i+1;

    – end

• end

  % Form linear equation and solve for alpha

• A=zeros(d,d); b=zeros(d,1); vD=reshape(D,d,1);

• for r=1:d

    – vEr=reshape(E(:,:,r),d,1); vErHC=vEr'*H+vEr.'*C;

    – for c=1:d

        * A(r,c)=real(vErHC*reshape(E(:,:,c),d,1));

    – end

    – b(r)=real(vEr'*vD);

• end

• alpha=-(A \ b);

  %Recover Z

• Z=zeros(n,n);

• for i=1:d
```

  - *Z=Z+alpha(i)*E(:,:,i);*

- *end*

Appendix B

Armijo's Rule

In this section we discuss Armijo's rule for selecting the step size in gradient based unconstraint optimization method. The material in this section are taken form (Bertsekas, 1999). Consider the unconstrained minimization of a function $f(\mathbf{x})$. At each iteration we have

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \alpha^k \mathbf{h}_{k-1} \tag{B.1}$$

where $\mathbf{h}_k$ is the search direction and α^k is the step size at iteration k. There are a number of rules for choosing the step size α^k in a gradient based method. One way is to select α^k by line minimization; i.e.,

$$\alpha^k = \arg\min_{\alpha \geq 0} f(\mathbf{x}_{k-1} + \alpha \mathbf{h}_{k-1}). \tag{B.2}$$

To avoid the considerable computational burden of line minimization methods, in practice a successive reduction of step size is employed. One simple method for successive reduction of step size is to select an initial value for α^k; e.g. $\alpha^k = s$. Now if $f(\mathbf{x}_{k-1} + s\mathbf{h}_{k-1}) \geq f(\mathbf{x}_{k-1})$ then the value of s is reduced by some certain factor until $f(\mathbf{x}_{k-1} + s\mathbf{h}_{k-1}) < f(\mathbf{x}_{k-1})$. This method although simple, in theory may not guarantee convergence to minimum. Refer to (Bertsekas, 1999) for more details on this issue. Armijo's rule eliminates the theoretical convergence problem of successive reduction rule described above by modifying it so that for fixed scalars s, $0 < \beta < 1$ and $0 < \sigma < 1$, the step size α^k is chosen as $\alpha^k = \beta^m s$ where

m is the first nonnegative integer for which

$$f(\mathbf{x}_{k-1}) - f(\mathbf{x}_{k-1} + \beta^m s\mathbf{h}_{k-1}) \geq -\sigma\beta^m s \nabla f(\mathbf{x}_{k-1})^T \mathbf{h}_{k-1}. \tag{B.3}$$

where $\nabla f(\mathbf{x}_{k-1})$ is the gradient of $f(\mathbf{x})$ at point $\mathbf{x}_{k-1}$.

The initial value of s is usually chosen to be one, the reduction factor β is chosen usually from 1/2 to 1/10 and σ is chosen close to zero; i.e., $\sigma \in [10^{-5}, 10^{-1}]$.

Appendix C

Supplement to Chapter 2

Newton and gradient descent optimization methods over complex Stiefel manifold have been explained in (Manton, 2002). In this section we derive the complex version of Hessian and gradient of the cost functions used in Algorithms I, III, and IV, presented in Chapter 2.

C.0.3 Algorithm I

When $\mathbf{Q}$ is a complex matrix, $\mathbf{D}_Q$, the matrix of partial derivatives of $\tilde{\mathcal{C}}_d(\mathcal{R}, \mathbf{Q})$ with respect to elements of $\mathbf{Q}$, is evaluated from

$$[\mathbf{D}_Q]_{rs} = \frac{\partial \tilde{\mathcal{C}}_d(\mathcal{R}, \mathbf{Q})}{\partial \Re q_{rs}} + j \frac{\partial \tilde{\mathcal{C}}_d(\mathcal{R}, \mathbf{Q})}{\partial \Im q_{rs}} \tag{C.4}$$

and following the same procedure given in Section 2.3.3 of Chapter 2 it is easily found to be equal to:

$$\mathbf{D}_Q = -2 \sum_{m=1}^{M} \mathbf{R}_m \mathbf{Q} \mathbf{\Sigma}(m) \tag{C.5}$$

which is the same as its real counterpart. Note that for complex $\mathbf{Q}$, the steepest descent search direction over the complex Stiefel manifold is calculated from (Manton, 2002):

$$\begin{aligned} \mathbf{G} &= -\mathbf{Q}\mathbf{D}_Q^{\dagger}\mathbf{Q} + \mathbf{D}_Q \\ &= 2\mathbf{Q}\Big(\sum_{m=1}^{M} [\mathbf{\Sigma}(m)\mathbf{R}_y(m) - \mathbf{R}_y(m)\mathbf{\Sigma}(m)] \Big) \end{aligned} \tag{C.6}$$

C.0.4 Algorithm III

For algorithm III we also need to calculate the Hessian of the objective function $\tilde{\mathcal{C}}_d(\mathcal{R},\mathbf{Q})$. The second-order Taylor series approximation of the function $\tilde{\mathcal{C}}_d(\mathcal{R},\mathbf{Q})$ for the case $\mathbf{Q}$ is a complex orthogonal matrix is given as:

$$\tilde{\mathcal{C}}_d(\mathcal{R},\mathbf{Q}+t\mathbf{H}) = \tilde{\mathcal{C}}_d(\mathcal{R},\mathbf{Q}) + t\Re\{Tr(\mathbf{H}^\dagger\mathbf{D}_Q)\} + \frac{t^2}{2}\Big(\mathbf{h}^\dagger\mathbf{D}_{QQ}\mathbf{h} + \Re\{\mathbf{h}^T\mathbf{C}_{QQ}\mathbf{h}\}\Big) + O(t^3) \tag{C.7}$$

where $\mathbf{h} = \text{vec}\{\mathbf{H}\}$ and the matrices $\mathbf{D}_{QQ} \in \mathbb{C}^{N^2 \times N^2}$, $\mathbf{C}_{QQ} \in \mathbb{C}^{N^2 \times N^2}$ are the Hessian of $\tilde{\mathcal{C}}_d$ evaluated at $\mathbf{Q}$ and to ensure uniqueness they must satisfy $\mathbf{D}_{QQ} = \mathbf{D}_{QQ}^\dagger$ and $\mathbf{C}_{QQ} = \mathbf{C}_{QQ}^T$.

We have

$$\begin{aligned}
\tilde{\mathcal{C}}_d(\mathcal{R},\mathbf{Q}+t\mathbf{h}) &= -\frac{1}{2}\sum_{m=1}^{M}\sum_{\substack{i<j\\ i,j=1}}^{N}\Big((\mathbf{q}+t\mathbf{h})^\dagger\mathbf{A}_{ij}(m)(\mathbf{q}+t\mathbf{h})\Big)^2 \\
&= -\frac{1}{2}\sum_{m=1}^{M}\sum_{\substack{i<j\\ i,j=1}}^{N}\Big(\mathbf{q}^\dagger\mathbf{A}_{ij}(m)\mathbf{q} + t^2\mathbf{h}^\dagger\mathbf{A}_{ij}(m)\mathbf{h} + 2\Re\{t\mathbf{h}^\dagger\mathbf{A}_{ij}(m)\mathbf{q}\}\Big)^2 \\
&= -\frac{1}{2}\sum_{m=1}^{M}\sum_{\substack{i<j\\ i,j=1}}^{N}\Big\{(\mathbf{q}^T\mathbf{A}_{ij}(m)\mathbf{q})^2 + 4t(\mathbf{q}^T\mathbf{A}_{ij}(m)\mathbf{q})\Re\{\mathbf{h}^T\mathbf{A}_{ij}(m)\mathbf{q}\} + \\
&\quad t^2[4(\mathbf{h}^T\mathbf{A}_{ij}(m)\mathbf{q})^2 + 2(\mathbf{q}^T\mathbf{A}_{ij}(m)\mathbf{q})(\mathbf{h}^T\mathbf{A}_{ij}(m)\mathbf{h})]\Big\} + O(t^3) \\
&= \tilde{\mathcal{C}}_d(\mathcal{R},\mathbf{Q}) + 4t\Re\Big\{\mathbf{h}^\dagger\Big(-\frac{1}{2}\sum_{m=1}^{M}\sum_{\substack{i<j\\ i,j=1}}^{N}\mathbf{q}^\dagger\mathbf{A}_{ij}(m)\mathbf{q}\mathbf{A}_{ij}(m)\mathbf{q}\Big)\Big\} + \\
&\quad t^2\mathbf{h}^\dagger\Big(-\frac{1}{2}\sum_{m=1}^{M}\sum_{\substack{i<j\\ i,j=1}}^{N}\Big\{4\mathbf{A}_{ij}(m)\mathbf{q}\mathbf{q}^\dagger\mathbf{A}_{ij}(m) + 2\mathbf{q}^\dagger\mathbf{A}_{ij}(m)\mathbf{q}\mathbf{A}_{ij}(m)\Big\}\Big)\mathbf{h} + O(t^3)
\end{aligned} \tag{C.8}$$

and by comparing the last equation with (C.7) it follows that

$$\mathbf{D}_{QQ} = -\sum_{m=1}^{M}\sum_{\substack{i<j\\ i,j=1}}^{N}\Big[4\mathbf{A}_{ij}(m)\mathbf{q}\mathbf{q}^\dagger\mathbf{A}_{ij}(m) + 2\mathbf{q}^\dagger\mathbf{A}_{ij}(m)\mathbf{q}\mathbf{A}_{ij}(m)\Big], \tag{C.9}$$

and $\mathbf{C}_{QQ} = \mathbf{0}$. Similar to real case the Newton search direction over the complex Stiefel manifold can be obtained by inserting $\mathbf{D}_{QQ}$ and $\mathbf{D}_Q$ into the *tpoint* function.

C.0.5 Algorithm IV

The matrix of partial derivatives of $\tilde{C}_l(\hat{\mathcal{R}}, \mathbf{Q})$ for the case where $\mathbf{Q}$ is complex is similar to its real counterpart and is given by

$$\mathbf{D}_Q = \sum_{m=1}^{M} \hat{\mathbf{R}}_m \mathbf{Q} \mathbf{\Lambda}_y^{-1}(m) \tag{C.10}$$

where $\mathbf{\Lambda}_y(m) = \mathrm{ddiag}\{\mathbf{Q}^\dagger \hat{\mathbf{R}}_m \mathbf{Q}\}$. To calculate the Hessian we use the second-order Taylor series expansion of the cost function $\tilde{C}_l(\hat{\mathcal{R}}, \mathbf{Q})$, which is similar to (C.7). We have

$$\begin{aligned}\tilde{C}_l(\hat{\mathcal{R}}, \mathbf{Q} + t\mathbf{h}) &= \frac{1}{2} \sum_{m=1}^{M} \sum_{i=1}^{N} \log[(\boldsymbol{\nu} + t\mathbf{h})^\dagger \mathbf{\Phi}_i(m)(\boldsymbol{\nu} + t\mathbf{h})] \\ &= \frac{1}{2} \sum_{m=1}^{M} \sum_{i=1}^{N} \Big(\log(\boldsymbol{\nu}^\dagger \mathbf{\Phi}_i(m)\boldsymbol{\nu}) + \log(1 + \frac{2t\Re\{\mathbf{h}^\dagger \mathbf{\Phi}_i(m)\boldsymbol{\nu}\} + t^2 \mathbf{h}^\dagger \mathbf{\Phi}_i(m)\mathbf{h}}{\boldsymbol{\nu}^\dagger \mathbf{\Phi}_i(m)\boldsymbol{\nu}}) \Big).\end{aligned} \tag{C.11}$$

Using the power series expansion $\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots$ we can continue

$$\begin{aligned}\tilde{C}_l(\hat{\mathcal{R}}, \mathbf{Q} + t\mathbf{h}) &= \frac{1}{2} \sum_{m=1}^{M} \sum_{i=1}^{N} \Big(\log(\boldsymbol{\nu}^\dagger \mathbf{\Phi}_i(m)\boldsymbol{\nu}) + \frac{t^2 \mathbf{h}^\dagger \mathbf{\Phi}_i(m)\mathbf{h} + 2t\Re\{\mathbf{h}^\dagger \mathbf{\Phi}_i(m)\boldsymbol{\nu}\}}{\boldsymbol{\nu}^\dagger \mathbf{\Phi}_i(m)\boldsymbol{\nu}} - \\ &\quad \frac{(t^2 \mathbf{h}^\dagger \mathbf{\Phi}_i(m)\mathbf{h} + 2t\Re\{\mathbf{h}^\dagger \mathbf{\Phi}_i(m)\boldsymbol{\nu}\})^2}{2(\boldsymbol{\nu}^\dagger \mathbf{\Phi}_i(m)\boldsymbol{\nu})^2} + \ldots \Big) \\ &= \tilde{C}_l(\hat{\mathcal{R}}, \mathbf{Q}) + t\Re\{\mathbf{h}^\dagger \Big(\sum_{m=1}^{M} \sum_{i=1}^{N} \frac{\mathbf{\Phi}_i(m)\boldsymbol{\nu}}{\boldsymbol{\nu}^\dagger \mathbf{\Phi}_i(m)\boldsymbol{\nu}} \Big)\} + \\ &\quad t^2 \mathbf{h}^\dagger \Big(\frac{1}{2} \sum_{m=1}^{M} \sum_{i=1}^{N} \Big\{ \frac{\mathbf{\Phi}_i(m)}{\boldsymbol{\nu}^\dagger \mathbf{\Phi}_i(m)\boldsymbol{\nu}} - \frac{2\mathbf{\Phi}_i(m)\boldsymbol{\nu}\boldsymbol{\nu}^\dagger \mathbf{\Phi}_i(m)}{(\boldsymbol{\nu}^\dagger \mathbf{\Phi}_i(m)\boldsymbol{\nu})^2} \Big\} \Big) \mathbf{h} + O(t^3).\end{aligned} \tag{C.12}$$

Comparing the last equation of the above with the second-order Taylor series of $\tilde{C}_l(\hat{\mathcal{R}}, \mathbf{Q})$ we can easily see that

$$\mathbf{D}_{QQ} = \sum_{m=1}^{M} \sum_{i=1}^{N} \Big[\frac{\mathbf{\Phi}_i(m)}{\boldsymbol{\nu}^\dagger \mathbf{\Phi}_i(m)\boldsymbol{\nu}} - \frac{2\mathbf{\Phi}_i(m)\boldsymbol{\nu}\boldsymbol{\nu}^\dagger \mathbf{\Phi}_i(m)}{(\boldsymbol{\nu}^\dagger \mathbf{\Phi}_i(m)\boldsymbol{\nu})^2} \Big], \tag{C.13}$$

and $\mathbf{C}_{QQ} = \mathbf{0}$.

Bibliography

Abed-Meraim, K., Qiu, W., and Hua, Y. (1997a). Blind System Identification. *Proceedings of the IEEE*, **85**, 1310–1322.

Abed-Meraim, K., Cardoso, J., Gorokhov, A., Loubaton, P., and Moulines, E. (1997b). On subspace methods for blind identification of single-input/multiple-output FIR filters. *IEEE Transactions on Signal Processing*, **45**, 42–55.

Abed_Meraim, K., Xiang, Y., Manton, J. H., and Hua, Y. (2001). Blind Source Separation Using Second-Order Cyclostationary Statistics. *IEEE Transactions on Signal Processing*, **49**, 694–701.

Anderson, T. (1971). *An introduction to multivariate statistical analysis.* John Wiley & Sons, New York, second edition.

Araki, S., Makino, S., Nishikawa, T., and Saruwatari, H. (2001). Fundamental limitation of frequency domain blind source separation for convolutive mixtures of speech. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 2737–2740.

Bell, A. and Sejnowski, T. J. (1995). An information maximization approach to blind separation and blind deconvolution. *Neural Computation*, (7), 1129–1159.

Belouchrani, A., Abed_Meraim, K., and Cardoso, J. (1997). A blind source separation technique using second order statistics. *IEEE Transactions on Signal Processing*, **45**, 434–444.

Bertsekas, D. (1999). *Nonlinear Programming.* Athena Scientific, Belmont Mass., second edition.

Brewer, J. (1979). Kronecker products and matrix calculus in system theory. *IEEE Transactions on Circuits and Systems*, **25**(9), 772–780.

Cardoso, J. (1994). On the performance of orthogonal source separation technique. In *EUSIPCO European Signal Processing Conference*, volume VII, pages 776–779.

Cardoso, J. (1998). Blind Signal Separation: Statistical Principles. *Proceedings of the IEEE*, **86**, 2009–2025.

Cardoso, J. and Laheld, B. (1996). Equivariant adaptive source separation. *IEEE Transactions on Signal Processing*, **44**, 3017–3030.

Cardoso, J. and Souloumiac, A. (1993). Blind beamforming for non Gaussian signals. In *IEE-F*, volume 140, pages 362–370.

Chan, D., Rayner, P., and Godsill, S. (1996). Multi-channel signal separation. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 649–652.

Chang, C., Ding, Z., Yau, S., and Chan, F. (2000). A matrix-pencil approach to blind separation of colored non-stationary signals. *IEEE Transactions on Signal Processing*, **48**, 900–907.

Chen, B. and Petropulu, A. (2001). Frequency Domain Blind MIMO System Identification Based on Second and Higher Order Statistics. *IEEE Transactions on Signal Processing*, **49**, 1677–1688.

Comon, P. (1994). Independent component analysis, A new concept? *SIGNAL PROCESSING*, **36**, 287–314.

Ding, Z. and Li, Y. (2001). *Blind Equalization and Identification.* Marcel Dekker, New York.

Edelman, A., Arias, T., and Smith, S. T. (1998). The geometry of algorithms with orthogonality constraints. *SIAM Journal on Matrix Analysis and Applications*, **20**, 303–353.

Feng, M. and Kammeyer, K. (1999). Application of source separation algorithms for mobile communication environment. In *International Workshop on Independent Component Analysis and Signal Separation*, pages 431–436, Aussois, France.

Flury, B. (1984). Common principal components in k groups. *Journal of the American Statistical Association*, **79**, 892–897.

Giannakis, G. and Mendel, J. (1989). Identification of nonminimum phase systems using higher order statistics. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **37**, 7360–7377.

Giannakis, G., Hua, Y., Stoica, P., and Tong, L. (2001). *Signal Processing Advances in Wireless & Mobile Communication*, volume 1, chapter 3. Prentice Hall.

Golub, G. and VanLoan, C. (1996). *Matrix Computations.* John Hopkins, Baltimore and London, third edition.

Gorokhov, A. and Loubaton, P. (1997). Subspace based techniques for second order blind separation of convolutive mixtures with temporally correlated sources. *IEEE Transactions on Circuits and Systems—I: Fundamental Theory and Applications*, **44**, 813–820.

Hori, G. (2000). New approach to joint diagonalization. In *International Workshop on Independent Component Analysis and Signal Separation.*

Horn, R. and Johnson, C. (1985). *Matrix analysis.* Cambridge university press.

Hua, Y. and Tugnait, J. (2000). Blind Identifiability of FIR-MIMO systems with colored input using second order statistics . *IEEE Signal Processing Letters*, **7**, 348–350.

Hua, Y. and Wax, M. (1996). Strict identifiability of multiple FIR channels driven by an unkown arbitrary sequence. *IEEE Transactions on Signal Processing*, **44**, 756–759.

Hua, Y., An, S., and Xiang, Y. (2001). Blind Identification and Equalization of FIR MIMO Channels By BIDS . In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing.*

Ikram, M. Z. and Morgan, D. R. (2000). Exploring Permutation Inconsistency in Blind Separation Of Signals In a Reverberant Environment. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1041–1044.

Ikram, M. Z. and Morgan, D. R. (2001). A multiresolution approach to blind separation of speech signals in a reverberant environment. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 2757–2760.

Joho, M. and Rahbar, K. (2002). Joint diagonalization of correlation matrices by using newton methods with application to blind signal separation. In *IEEE Workshop on Sensor Array and Multichannel Signal Processing.*

Jung, T., Makeig, S., Lee, T., McKeown, M. J., Brown, G., Bell, A., and Sejnowski, T. (2000). Idependent component analysis of biomedical signals. In *International Workshop on Independent Component Analysis and Signal Separation*, pages 633–644.

Kailath, T. (1980). *Linear Systems.* Prentice Hall, New Jersey.

Lee, T.-W., Bell, A. J., and Lambert, R. H. (1997). Blind separation of delayed and convolved sources. In *Advances in Neural Information Processing System*, pages 758–764. MIT Press.

Li, T. and Sidiropoulos, N. (2000). Blind Digital Signal Separation Using Successive Interference Cancellation Iterative Least Squares. *IEEE Transactions on Signal Processing*, **48**, 3146–3152.

Lindgren, U. A. and Broman, H. (1998). Source separation using a criterion based on second-order statistics. *IEEE Transactions on Signal Processing*, **46**(7), 1837–1850.

Loubaton, P. and Moulines, E. (1999). Application of blind second order statistics MIMO identification methods to the blind CDMA forward link channel estimation . In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 2543–2546.

Loubaton, P. and Moulines, E. (2000). On blind multiuser forward link channel estimation by the subspace method: identifiability results. *IEEE Transactions on Signal Processing*, **48**, 2366–2376.

Ma, C. T., Ding, Z., and Yau, S. F. (2000). A Two-stage Algorithm for MIMO Blind Deconvolution of Nonstationary Colored Signals. *IEEE Transactions on Signal Processing*, **48**, 1187–1192.

Makeig, S., Enghoff, S., Jung, T., and Sejnowsky, T. J. (2000). Moving-window ica decomposition of eeg data reveals event-related changes in oscillatory brain activity. In *International Workshop on Independent Component Analysis and Signal Separation*, pages 627–632, Helsinki, Finland.

Manton, J. (2002). Optimization algorithms exploiting unitary constraints. *IEEE Transactions on Signal Processing*, **50**(3), 636–650.

Manton, J. H. (2001). A Packet Based Channel Identification Algorithm For Wireless Multimedia Communications. In *IEEE International Conference on Multimedia and Expo*, Tokyo, Japan.

Mendel, J. M. (1991). Tutorials on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE*, **79**(3), 278–305.

Meraim, K., Xiang, Y., and Hua, Y. (2000). Generalized second order identifiability condition and relevant testing technique. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 2989–2992.

Morgan, D. R., Benesty, J., and Sondhi, M. M. (1998). On the Evaluation of Estimated Impulse Responses. *IEEE Signal Processing Letters*, **5**, 174–176.

Moulines, E., Duhamel, P., Cardoso, J.-F., and Mayrargue, S. (1995). Subspace methods for the blind identification of multichannel fir filters. *IEEE Transactions on Speech and Audio Processing*, **43**, 516–525.

Neely, S. and Allen, J. (1979). Invertibility of a room impulse response. *Journal Acoustic Society America*, **66**, 165–169.

Papoulis, A. (1984). *Probability, Random Variables and Stochastic Processes.* McGraw-Hill, Singapore.

Parra, L. and Spence, C. (2000). Convolutive blind separation of non-stationary sources. *IEEE Transactions on Speech and Audio Processing*, **8**, 320–327.

Pham, D. (1996). Blind separation of instantaneous mixture of sources via an independent component analysis. *IEEE Transactions on Signal Processing*, **44**(11), 2768–2779.

Pham, D. (2000). Joint approximate diagonalization of positive definite hermitian matrices. In *Technical Report*, Grenoble University.

Pham, D. T. and Cardoso, J. (2001). Blind source separation of instantaneous mixtures of nonstationary sources. *IEEE Transactions on Signal Processing*, **49**, 1837–1848.

Proakis, J. (2001). *Digital Communications.* McGrawHill, fourth edition.

Rahbar, K. and Reilly, J. (2000). Geometric optimization methods for blind source separation of signals. In *International Workshop on Independent Component Analysis and Signal Separation*, pages 375–380.

Rahbar, K. and Reilly, J. (2001a). Blind source separation algorithm for mimo convolutive mixtures. In *International Workshop on Independent Component Analysis and Signal Separation*, pages 242–247.

Rahbar, K. and Reilly, J. (2001b). Blind source separation of convolved sources by joint approximate diagonalization of cross-spectral density matrices. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 2745–2748.

Rahbar, K., Reilly, J., and Manton, J. (2002a). A Frequency Domain Approach to Blind Identification of MIMO FIR Systems Driven By Quasi-Stationary Signals. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1717–1720.

Rahbar, K., Reilly, J., and Manton, J. H. (2002b). Blind identification of mimo fir systems driven by quasi-stationary sources using second order statistics: A frequency domain approach. *Submitted to IEEE Transaction on Signal Processing.*

Rahbar, K., Reilly, J., and Manton, J. (2002c). A frequency domian approach to blind identification of mimo fir channels driven by non-stationary sources. In *IEEE Workshop on Sensor Array and Multichannel Signal Processing.*

Reilly, J. and Mendoza, L. (1999). Blind source separation for convolutive mixing environments using spatial-temporal processing. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1437–1440.

Sahlin, H. and Broman, H. (2000). MIMO signal separation for FIR channels: a criterion and performance analysis. *IEEE Transactions on Signal Processing*, **48**, 642–649.

Schobben, D. and Sommen, P. (1998). A new blind signal separation algorithm based on second order statistics. In *IASTED International Conference on Signal and Image Processing*, pages 564–569.

Schroeder, M. (1965). New method for measuring reverberation time. *Journal Acoustic Society America*, **37**, 409–412.

Serpedin, E. and Giannakis, G. (1998). Blind channel identification and equalization with modulation-induced cyclostationarity. *IEEE Transactions on Signal Processing*, **46**, 1930–1944.

Serpedin, E. and Giannakis, G. (1999). A simple proof of a known blind channel identifiability result. *IEEE Transactions on Signal Processing*, **47**, 591–593.

Serviere, C. (1998). Feasibility of source separation in frequency domain. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 2085–2088.

Shamsunder, S. and Giannakis, G. (1997). Multichannel blind signal separation and reconstruction. *IEEE Transactions on Speech and Audio Processing*, **5**(6), 515–528.

Shen, J. and Ding, Z. (2001). Zero-forcing blind equalization based on channel subspace estimates for multiuser systems. *IEEE Transactions on Communications*, pages 262–271.

Sidiropoulos, N., Giannakis, G., and Bro, R. (2000). Blind PARAFAC Receivers for DS-CDMA Systems. *IEEE Transactions on Signal Processing*, **48**, 810–823.

Sidiropoulos, N. D., Giannakis, G. B., and Bro, R. (1998). Deterministic waveform - preserving blind separation of ds-cdma signals using an antenna array. In *IEEE SP Workshop on Statistical Signal and Array Processing*, pages 304–307, Portland, Oregon.

Souloumiac, A. (1995). Blind source detection and separation using second order non-stationarity. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1912–1915.

Talwar, S., Viberg, M., and Paulraj, A. (1996). Blind separation of synchronous co-channel digital signals using an antenna array-Part I: Algorithms. *IEEE Transactions on Signal Processing*, **44**, 1184–1197.

Tong, L. and Perreau, S. (1998). Multichannel blind identification: From subspace to maximum likelihood methods. *Proceedings of the IEEE*, **86**, 1951–1968.

Tong, L., Liu, R., Soon, V., and Huang, Y. (1991). Indeterminacy and identifiability of blind identification. **38**, 499–509.

Tong, L., Xu, G., and Kailath, T. (1994). Blind identification and equalization based on second-order statistics: A time domain approach. *IEEE Transactions on Information Theory*, **40**, 340–349.

Tong, L., Xu, G., Hassibi, B., and Kailath, T. (1995). Blind Channel Identification based on second order statistics: A frequency domain approach. *IEEE Transactions on Information Theory*, **41**, 329–334.

Torkkola, K. (1999). Blind source separation for audio signals- are we there yet? In *International Workshop on Independent Component Analysis and Signal Separation*, pages 239–243.

Tsatsanis, K. and Zhang, R. (2001). A second-order method for blind separation of non-stationary sources. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 2797–2800.

Tugnait, J. K. (1997). Identification and deconvolution of multichannel linear non-gaussian processes using higher order statistics and inverse filter criteria. *IEEE Transactions on Signal Processing*, **45**(3), 658–672.

Tugnait, J. K. (1999). Adaptive blind separation of convolutive mixtures of independent linear signals. *SIGNAL PROCESSING*, **73**, 139–152.

Vaidyanathan, P. (1993). *Multirate Systems and Filter Banks*. Prentice Hall, New Jersey.

van der Veen, A.-J. (2001). Joint diagonalization via subspace fitting techniques. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 2773–2776.

Wax, M. (1997). A least-squares approach to joint diagonalization. *IEEE Signal Processing Letters*, **4**, 52–53.

Weinstein, E., Feder, M., and Oppenheim, A. V. (1993). Multi-channel signal separation by decorrelation. *IEEE Transactions on Signal Processing*, 1(4), 404–413.

Westner, A. G. (1998). *Object-based audio capture: separating acoustic sounds.* Master's thesis, Massachusetts Institute of Technology Media Laboratory.

Wilbur, M. (2000). *The decomposition of large blind equalization problems using GDFT filter banks.* Master's thesis, McMaster University, Hamilton, Ontario, Canada.

Yellin, D. and Weinstein, E. (1994). Criteria for multichannel signal separation. *IEEE Transactions on Signal Processing*, **42**, 2158–2168.

Yeredor, A. (2000). Approximate joint diagonalization using non-orthogonal matrices. In *International Workshop on Independent Component Analysis and Signal Separation*, pages 33–38.

Yeredor, A. (2002). Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation. *IEEE Transactions on Signal Processing*, **50**, 1545–1553.

Zhu, J., Ding, Z., and Cao, X. (1998). Column anchored zeroforcing blind equalization for multiuser wireless fir channels. *IEEE Journal on Selected Areas in Communications*, pages 411–423.

Ziehe, A., Muller, K., Nolte, G., Mackert, B., and Cuiro, G. (1998). Artifact reduction in magnetoneurography based on time-delayed second order correlations. Technical report, GMD-Forshungszentrum Informationstechnik GmbH.